



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ**  
**ΣΧΟΛΗ ΑΝΘΡΩΠΙΣΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ**  
**ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ**  
**ΚΑΙ ΤΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ**

**ΘΕΜΑ: «Υπολογιστικά υποστηριζόμενη εκπαιδευτική αξιολόγηση σε περιβάλλοντα ηλεκτρονικής μάθησης»**

**ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ**  
του Θεολόγου Τσιγάρου

**ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ**

Γεώργιος Φεσάκης	Καθηγητής	Πανεπιστήμιο Αιγαίου	Επιβλέπων Τριμελούς Συμβουλευτικής Επιτροπής
Αγγελική Δημητρακοπούλου	Καθηγήτρια	Πανεπιστήμιο Αιγαίου	Μέλος Τριμελούς Συμβουλευτικής Επιτροπής
Νικόλαος Ανδρεαδάκης	Καθηγητής	Πανεπιστήμιο Αιγαίου	Μέλος Τριμελούς Συμβουλευτικής Επιτροπής
Χρυσή Βιτσιλάκη	Καθηγήτρια	Πανεπιστήμιο Αιγαίου	Μέλος Επταμελούς Εξεταστικής Επιτροπής
Φραγκίσκος Καλαβάσης	Καθηγητής	Πανεπιστήμιο Αιγαίου	Μέλος Επταμελούς Εξεταστικής Επιτροπής
Αναστάσιος Κοντάκος	Καθηγητής	Πανεπιστήμιο Αιγαίου	Μέλος Επταμελούς Εξεταστικής Επιτροπής
Ανδρέας Μούτσιος Ρέντζος	Επίκουρος Καθηγητής	Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών	Μέλος Επταμελούς Εξεταστικής Επιτροπής

**ΡΟΔΟΣ, 2023**



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ**  
**ΣΧΟΛΗ ΑΝΘΡΩΠΙΣΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ**  
**ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ**  
**ΚΑΙ ΤΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ**

**ΘΕΜΑ: «Υπολογιστικά υποστηριζόμενη εκπαιδευτική αξιολόγηση σε περιβάλλοντα ηλεκτρονικής μάθησης»**

**ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ**  
του Θεολόγου Τσιγάρου

**ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ**

Γεώργιος Φεσάκης	Καθηγητής	Πανεπιστήμιο Αιγαίου	Επιβλέπων Τριμελούς Συμβουλευτικής Επιτροπής
Αγγελική Δημητρακοπούλου	Καθηγήτρια	Πανεπιστήμιο Αιγαίου	Μέλος Τριμελούς Συμβουλευτικής Επιτροπής
Νικόλαος Ανδρεαδάκης	Καθηγητής	Πανεπιστήμιο Αιγαίου	Μέλος Τριμελούς Συμβουλευτικής Επιτροπής
Χρυσή Βιτσιλάκη	Καθηγήτρια	Πανεπιστήμιο Αιγαίου	Μέλος Επταμελούς Εξεταστικής Επιτροπής
Φραγκίσκος Καλαβάσης	Καθηγητής	Πανεπιστήμιο Αιγαίου	Μέλος Επταμελούς Εξεταστικής Επιτροπής
Αναστάσιος Κοντάκος	Καθηγητής	Πανεπιστήμιο Αιγαίου	Μέλος Επταμελούς Εξεταστικής Επιτροπής
Ανδρέας Μούτσιος Ρέντζος	Επίκουρος Καθηγητής	Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών	Μέλος Επταμελούς Εξεταστικής Επιτροπής

**ΡΟΔΟΣ, 2023**

Κωδικός: 5\_GR1bNKFPcNn4gptXfq6Q

Επιβεβαιώνεται το γνήσιο. Υπουργείο  
Ψηφιακής Διακυβέρνησης / Verified by the Ministry  
of Digital Governance, Hellenic Republic  
20230405092153+03'00'



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

Υπεύθυνη Δήλωση  
(άρθρο 8 Ν.1599/1986)



Η ακρίβεια των στοιχείων που υποβάλλονται με αυτή τη δήλωση μπορεί να ελεγχθεί με βάση το αρχείο άλλων υπηρεσιών (άρθρο 8 παρ. 4 Ν. 1599/1986).

Προς <sup>(1)</sup> :	ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ, ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ ΚΑΙ ΤΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ						
Όνομα:	ΘΕΟΛΟΓΟΣ	Επώνυμο:	ΤΣΙΓΑΡΟΣ				
Όνομα και Επώνυμο Πατέρα:	ΙΩΑΝΝΗΣ ΤΣΙΓΑΡΟΣ						
Όνομα και Επώνυμο Μητέρας:	ΜΑΡΙΑ ΤΣΙΓΑΡΟΥ						
Ημερομηνία γέννησης:	07/06/1970						
Τόπος Γέννησης:	ΡΟΔΟΣ						
Αριθμός Δελτίου Ταυτότητας:	ΑΡ035107	Τηλ:	+306945996980				
Τόπος Κατοικίας:	ΡΟΔΟΣ	Οδός:	ΠΑΥΛΟΥ ΜΕΛΛΑ	Αριθ:	42	ΤΚ: 85100	
ΑΦΜ:	046840184	Δ/ση Ηλεκτρ. Ταχυδρομείου (E-mail):					

Με ατομική μου ευθύνη και γνωρίζοντας τις κυρώσεις<sup>(2)</sup>, που προβλέπονται από τις διατάξεις της παρ. 6 του άρθρου 22 του Ν. 1599/1986, δηλώνω ότι:

Είμαι ο αποκλειστικός συγγραφέας της υποβληθείσας Διδακτορικής Διατριβής με τίτλο «Υπολογιστικά υποστηριζόμενη εκπαιδευτική αξιολόγηση σε περιβάλλοντα ηλεκτρονικής μάθησης». Η συγκεκριμένη Διδακτορική Διατριβή είναι πρωτότυπη και εκπονήθηκε αποκλειστικά για την απόκτηση του Διδακτορικού διπλώματος του Τ.Ε.Π.Α.Ε.Σ.. Κάθε βοήθεια, την οποία είχα για την προετοιμασία της, αναγνωρίζεται πλήρως και αναφέρεται επακριβώς στην εργασία. Επίσης, επακριβώς αναφέρω στην εργασία τις πηγές, τις οποίες χρησιμοποίησα, και μνημονεύω επώνυμα τα δεδομένα ή τις ιδέες που αποτελούν προϊόν πνευματικής ιδιοκτησίας άλλων, ακόμη κι εάν η συμπερίληψή τους στην παρούσα εργασία υπήρξε έμμεση ή παραφρασμένη. Γενικότερα, βεβαιώνω ότι κατά την εκπόνηση της Διδακτορικής Διατριβής έχω τηρήσει απαρέγκλιτα όσα ο νόμος ορίζει περί διανοητικής ιδιοκτησίας και έχω συμμορφωθεί πλήρως με τα προβλεπόμενα στο νόμο περί προστασίας προσωπικών δεδομένων και τις αρχές Ακαδημαϊκής Δεοντολογίας.

05/04/2023

Ο - Η Δηλ.

ΘΕΟΛΟΓΟΣ ΤΣΙΓΑΡΟΣ

(1) Αναγράφεται από τον ενδιαφερόμενο παλίτη η αρχή ή η υπηρεσία του δημόσιου τομέα όπου απευθύνεται η αίτηση.

(2) Γνωρίζω ότι: Όποιος εν γνώσει του δηλώνει ψευδή γεγονότα ή αρνείται ή αποκρύπτει τα αληθινά με έγγραφη υπεύθυνη δήλωση του άρθρου 8 τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών. Εάν ο υπαίτιος αυτών των πράξεων σκόπευε να προσπορίσει στον εαυτόν του ή σε άλλον περιουσιακό όφελος βλάπτοντας τρίτον ή σκόπευε να βλάψει άλλον, τιμωρείται με κάθειρξη μέχρι 10 ετών.

## Ευχαριστίες

Στον Καθηγητή Γιώργο Φεσάκη, επιβλέποντα της διατριβής, είμαι ευγνώμων. Οι θερμές ευχαριστίες μου αποτελούν ελάχιστη ανταπόδοση έναντι της γεναιόδωρης, πολύτιμης και συνεχούς υποστήριξής του. Χωρίς την παρότρυνσή του, ενδεχομένως να μην επιχειρούσα την εκπόνηση διδακτορικής διατριβής. Επιπλέον, μου πρότεινε το εξαιρετικά ενδιαφέρον θέμα της έρευνας, με καθοδήγησε σε κάθε στάδιο της προσπάθειάς μου και κυρίως με ενθάρρυνε και μου ενέπνεε αισιοδοξία.

Οφείλω επίσης να ευχαριστήσω θερμά την Καθηγήτρια Αγγελική Δημητρακοπούλου, μέλος της τριμελούς επιτροπής για τη βοήθειά της καθόλη τη διάρκεια της προσπάθειάς μου, για τις πολύτιμες και εύστοχες συμβουλές της, για το αμέριστο ενδιαφέρον της ως προς την ευόδωση της διατριβής.

Ευχαριστώ θερμά τον Καθηγητή Νίκο Ανδρεαδάκη, μέλος της τριμελούς επιτροπής για τη καθοδήγησή του στα ζητήματα της εκπαιδευτικής αξιολόγησης και της στατιστικής ανάλυσης, για την άμεση ανταπόκρισή του κάθε φορά που ζητούσα τη βοήθειά του και γενικά για την πολύπλευρη υποστήριξη που πρόθυμα μου προσέφερε κατά την ερευνητική προσπάθειά μου.

Ευχαριστώ επίσης, τα μέλη της επταμελούς συμβουλευτικής επιτροπής, την Καθηγήτρια Χρυσή Βιτσιλάκη, Πρύτανη του Πανεπιστημίου Αιγαίου, τον Καθηγητή Φραγκίσκο Καλαβάση, τον Καθηγητή Αναστάσιο Κοντάκο και τον Επίκουρο Καθηγητή Ανδρέα Μούτσιο Ρέντζο, για τη διαχρονική και πολύπλευρη υποστήριξή τους και διότι με τίμησαν συμμετέχοντας στη διαδικασία υποστήριξης και αξιολόγησης της διατριβής.

Η εκπόνηση της διατριβής θα ήταν αδύνατη χωρίς την υποστήριξη της οικογένειάς μου προς την οποία εκφράζω τις ευχαριστίες και την αγάπη μου.

## ΠΕΡΙΛΗΨΗ

Οι ψηφιακές τεχνολογίες μετασχηματίζουν το χώρο της εκπαίδευσης. Υποστηρίζουν την ανάπτυξη και τη λειτουργία νέων ψηφιακών περιβαλλόντων μάθησης, εντός των οποίων οι διαδικασίες της εκπαιδευτικής αξιολόγησης παρέχονται μέσω υπολογιστικών συστημάτων. Η Υπολογιστικά Υποστηριζόμενη Αξιολόγηση σχετίζεται με ένα ευρύ φάσμα συστημάτων και διαδικασιών. Περιλαμβάνει συστήματα και εφαρμογές που αυτοματοποιούν παραδοσιακές μεθόδους αξιολόγησης, αλλά και καινοτόμους μεθόδους, μέσω των οποίων η αξιολόγηση αποκτά νέες μορφές και ενσωματώνεται στα νέα ψηφιακά περιβάλλοντα μάθησης.

Η διατριβή παρουσιάζει αναλυτικά τις βασικές μορφές που μπορεί να λάβει η Υπολογιστικά Υποστηριζόμενη Αξιολόγηση, αλλά επικεντρώνεται στις καινοτόμες μορφές της. Σκοπός της διατριβής είναι να συνεισφέρει στο πεδίο της μελέτης των προβλεπτικών μοντέλων Μηχανικής Μάθησης σε ψηφιακά περιβάλλοντα μάθησης και των δυνατοτήτων αξιοποίησής τους στην εκπαιδευτική αξιολόγηση. Ειδικότερα, η διατριβή διερευνά τις δυνατότητες των μοντέλων να αξιοποιήσουν δεδομένα χρήσης των Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ) και να προβλέψουν τις επιδόσεις των φοιτητών, όχι μόνο με ακρίβεια, αλλά και εγκαίρως, ώστε να μπορούν να χρησιμοποιηθούν ως εργαλεία διαμορφωτικής αξιολόγησης, τα οποία αφενός εντοπίζουν ζητήματα που προοιωνίζονται αποτυχημένη φοίτηση και αφετέρου καθοδηγούν στοχευμένες διδακτικές παρεμβάσεις.

Συλλέχθηκαν δεδομένα πρόσβασης στο ΣΔΜ Moodle για 4 μαθήματα προπτυχιακού και μεταπτυχιακού επιπέδου. Τα μοντέλα πρόβλεψης αξιολογήθηκαν, ως προς την ακρίβεια και ως προς τη χρονική στιγμή της πρόβλεψης. Διερευνήθηκε επίσης, ο διαφορετικός βαθμός επίδρασης συγκεκριμένων χαρακτηριστικών στην απόδοση των μοντέλων και οι συσχετίσεις των χαρακτηριστικών αυτών με διαστάσεις που αφορούν την ενεργή συμμετοχή των φοιτητών στην εκπαιδευτική διαδικασία.

Τα κύρια αποτελέσματα της έρευνας έδειξαν ότι, με την αξιοποίηση μοντέλων Μηχανικής Μάθησης στα μαθήματα που μελετήθηκαν, προβλέφθηκε, με αρκετά καλή ακρίβεια και αξιοπιστία, η τελική επίδοση των σπουδαστών. Επιπλέον, εφικτή ήταν η έγκαιρη πρόβλεψη, αρκετά πριν την ολοκλήρωση των μαθημάτων. Προέκυψε, ότι κάποια από τα χαρακτηριστικά

στα σύνολα δεδομένων των μοντέλων, επηρεάζουν περισσότερο από τα υπόλοιπα την ικανότητα πρόβλεψης. Τα επιδραστικά χαρακτηριστικά είναι κοινά μεταξύ συναφών μαθημάτων και ο υψηλός βαθμός επίδρασης τους μπορεί να αποδοθεί στη σχέση που έχουν με δείκτες ενεργής συμμετοχής των φοιτητών στα μαθήματα. Προέκυψε επίσης, ότι οι διαφορές στη δομή και στην οργάνωση των μαθημάτων, ως αποτέλεσμα διαφορετικών σχεδιαστικών επιλογών, επηρεάζουν τη βαρύτητα των χαρακτηριστικών στα σύνολα δεδομένων και τελικά επηρεάζουν και την απόδοση των μοντέλων.

### *Λέξεις ευρετηριασμού*

Υπολογιστικά Υποστηριζόμενη Αξιολόγηση, Ψηφιακή Αξιολόγηση, Ψηφιακή Μάθηση, Μαθησιακή Αναλυτική, Εξόρυξη Εκπαιδευτικών Δεδομένων, Προβλεπτική Αναλυτική, Επιδραστικά Χαρακτηριστικά, Ενεργή Συμμετοχή.

## ABSTRACT

Digital technologies are transforming education. They are supporting the development and operation of new learning environments, where the processes of educational assessment are provided through computer systems. Computer Assisted Assessment is associated with a wide range of systems and processes. It includes systems and applications that automate traditional assessment methods, but it also includes innovative methods, where assessment acquires new forms and is integrated into the new digital learning environments.

The thesis presents in detail the basic forms that Computer Assisted Assessment can take, focusing on its most innovative forms. The aim of the thesis is to contribute to the study of predictive Machine Learning models in digital learning environments and the abilities of their utilization in educational assessment.

In particular, the thesis explores the potential of the models to utilize data from Learning Management Systems (LMS) and to predict student performance, not only with accuracy, but also in time, so that they can be used as formative assessment tools, which can identify factors affecting academic failure and guide targeted educational interventions.

Data from the LMS Moodle was collected for 4 undergraduate and postgraduate courses. The predictive models were evaluated for accuracy and for timely prediction. Dataset features were identified, which affect the performance of the models to a different degree and which are related to student engagement.

The main results of the research showed that, Machine Learning models, predict student's performance with quite good accuracy and reliability. In addition, early prediction was possible, well before the completion of the courses.

It emerged that some of the dataset features affect the prediction performance more than others. These features are common among relevant courses and their high degree of influence can be attributed to their relationship with indicators of student engagement. It also emerged

that differences in the structure and design of courses, affect the importance of dataset features and ultimately affect the performance of the models.

*Keywords*

Computer Assisted Assessment, e-assessment, e-learning, Learning Analytics, Educational Data Mining, Prediction Analytics, influential features, Engagement.



## Πίνακας Περιεχομένων

ΠΕΡΙΛΗΨΗ.....	5
ABSTRACT.....	7
Κατάλογος Πινάκων.....	14
Κατάλογος Σχημάτων.....	17
Κατάλογος Γραφημάτων.....	18
Κατάλογος Εικόνων.....	19
Ακρωνύμια.....	20
Γλωσσάρι ξενόγλωσσων όρων.....	21
<b>ΕΙΣΑΓΩΓΗ.....</b>	<b>22</b>
<b>1. Η ΥΠΟΛΟΓΙΣΤΙΚΑ ΥΠΟΣΤΗΡΙΖΟΜΕΝΗ ΕΚΠΑΙΔΕΥΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ / ΨΗΦΙΑΚΗ ΑΞΙΟΛΟΓΗΣΗ.....</b>	<b>26</b>
1.1 Η αξιολόγηση στα νέα περιβάλλοντα μάθησης.....	26
1.2. ΤΠΕ και Αξιολόγηση.....	28
1.3. Ορολογία.....	29
<b>2. ΕΞΕΛΙΞΗ ΚΑΙ ΤΑΞΙΝΟΜΗΣΗ ΤΩΝ ΣΥΣΤΗΜΑΤΩΝ ΤΗΣ ΨΗΦΙΑΚΗΣ ΑΞΙΟΛΟΓΗΣΗΣ.....</b>	<b>34</b>
2.1. Εξελικτικές φάσεις.....	35
2.2. Στρατηγικές Μετάβασης.....	42
<b>3. ΔΙΑΓΝΩΣΤΙΚΗ, ΔΙΑΜΟΡΦΩΤΙΚΗ, ΤΕΛΙΚΗ ΨΗΦΙΑΚΗ ΑΞΙΟΛΟΓΗΣΗ.....</b>	<b>47</b>
3.1. Διαμορφωτική έναντι τελικής αξιολόγησης.....	50
3.2. Αξία διαμορφωτικής ψηφιακής αξιολόγησης.....	51
3.3. Αλληλεπίδραση στη διαμορφωτική ψηφιακή αξιολόγηση.....	54
3.3.1. Εγκυρότητα, Αξιοπιστία, Ανεντιμότητα.....	55
3.4. Ο ρόλος της ανατροφοδότησης στη ψηφιακή αξιολόγηση.....	58
3.4.1. Ανατροφοδότηση και αρχές διαμορφωτικής ψηφιακής αξιολόγησης.....	61
<b>4. ΜΟΡΦΕΣ ΨΗΦΙΑΚΗΣ ΑΞΙΟΛΟΓΗΣΗΣ.....</b>	<b>69</b>
4.1. Ψηφιακές αξιολογήσεις κλειστού τύπου.....	69
4.1.1. Πλεονεκτήματα, Προκλήσεις.....	73
4.1.2. Αρχές σχεδιασμού καλών ερωτήσεων.....	74
4.1.3. Τυχαίες απαντήσεις.....	75
4.1.4. Μέθοδοι βαθμολόγησης.....	76

4.1.5. Βαθμολόγηση βάση βεβαιότητας.....	78
4.1.6. Προσαρμοζόμενα διαγωνίσματα .....	81
4.1.7. Τράπεζες ερωτήσεων .....	84
4.2. Ψηφιακές αξιολογήσεις ανοικτού τύπου .....	88
4.3. Ψηφιακοί φάκελοι επιτευγμάτων (e-portfolios).....	93
4.4. Συνεργατική ανάπτυξη περιεχομένου (wiki).....	98
4.5. Χώροι δημόσιου διαλόγου (Forum) .....	100
4.6. Σοβαρά Παιχνίδια και Προσομοιώσεις.....	104
<b>5. ΜΑΘΗΣΙΑΚΗ ΑΝΑΛΥΤΙΚΗ.....</b>	<b>110</b>
5.1.Ορισμός και επισκόπηση πεδίου.....	110
5.2. Η μαθησιακή αναλυτική και η αξιολόγηση .....	119
5.3. Προβλεπτική Αναλυτική και Μηχανική Μάθηση .....	125
5.3.1. Ταξινόμηση και Μοντέλα Πρόβλεψης.....	128
5.3.2. Η έρευνα στο πεδίο της Προβλεπτικής Αναλυτικής .....	132
5.4 Μαθησιακή Αναλυτική και Ενεργή Συμμετοχή .....	142
<b>6. ΟΦΕΛΗ ΤΗΣ ΨΗΦΙΑΚΗΣ ΑΞΙΟΛΟΓΗΣΗΣ .....</b>	<b>146</b>
6.1. Οφέλη στις βασικές μορφές ψηφιακής αξιολόγησης .....	146
6.2. Οφέλη στις καινοτόμες μορφές ψηφιακής αξιολόγησης.....	150
<b>7. ΠΡΟΚΛΗΣΕΙΣ .....</b>	<b>155</b>
7.1. Ηθικές και Πολιτικές προκλήσεις.....	155
7.2. Παιδαγωγικές προκλήσεις .....	156
7.3. Οργανωτικές προκλήσεις.....	158
7.4. Τεχνικές προκλήσεις .....	160
7.5. Η αξιολόγηση ανώτερων νοητικών ικανοτήτων.....	160
7.6. Η ακαδημαϊκή ακεραιότητα και η ασφάλεια υποδομών και υπηρεσιών.....	167
7.6.1. Προστασία ακαδημαϊκής ακεραιότητας.....	168
7.6.2. Ασφάλεια υποδομών και υπηρεσιών .....	175
<b>8. ΠΡΟΒΛΗΜΑΤΙΚΗ .....</b>	<b>178</b>
8.1. Το ερευνητικό πεδίο .....	178
8.2. Σκοπός και στόχοι .....	181
8.3. Ερευνητικά ερωτήματα.....	182

<b>9. ΜΕΘΟΔΟΣ</b> .....	184
9.1. Ο υπό μελέτη πληθυσμός .....	186
9.2. Περιγραφή μαθημάτων .....	186
9.3. Ερευνητική δεοντολογία.....	191
9.4. Εργαλεία.....	194
9.4.1. Το Σύστημα Διαχείρισης Μάθησης Moodle του Πανεπιστημίου Αιγαίου .....	194
9.4.2. Τα αρχεία καταγραφής δεδομένων χρήσης (log files) στο ΣΔΜ Moodle .....	197
9.4.3. Το λογισμικό αλγόριθμων Μηχανικής Μάθησης WEKA.....	199
9.4.4. Το Σύστημα φοιτητολογίου του Πανεπιστημίου Αιγαίου .....	202
9.4.5. Το λογισμικό υπολογιστικών φύλλων MS Excel .....	203
9.4.6. Το λογισμικό στατιστικής ανάλυσης IBM SPSS.....	204
9.5. Δημιουργία συνόλων δεδομένων.....	205
9.5.1. Αναπαράσταση συνόλου δεδομένων .....	205
9.5.2. Άντληση δεδομένων, αποθήκευση σε αρχείο. ....	209
9.5.3. Συσχέτιση εγγραφών αρχείων καταγραφής με εγγραφές φοιτητολογίου .....	214
9.5.4. Αντικατάσταση ονομάτων χρηστών με ψευδώνυμο.....	215
9.5.5. Τα χαρακτηριστικά των συνόλων δεδομένων .....	215
9.5.6. Αναλυτική παρουσίαση χαρακτηριστικών των συνόλων δεδομένων.....	218
9.5.7. Επιλογή χαρακτηριστικών .....	246
9.5.8. Υπολογισμός τιμών για τα χαρακτηριστικά των συνόλων δεδομένων .....	248
9.5.9. Η κλάση των οντοτήτων .....	249
9.5.10. Κανονικοποίηση Τιμών .....	251
9.6. Δημιουργία μοντέλων ταξινόμησης .....	253
9.6.1. Αλγόριθμοι ταξινόμησης.....	253
9.6.2. Κριτήριο απόδοσης και τιμή βάσης .....	256
9.6.3. Επιπλέον κριτήρια για την απόδοση των μοντέλων.....	257
9.6.4. Το ζήτημα της «υπερπροσαρμογής» .....	258
9.6.5. Ικανότητα των μοντέλων για γενίκευση. ....	260
9.6.6. Χρήση του WEKA Experimenter έναντι του WEKA Explorer.....	260
9.7. Εντοπισμός επιδραστικών χαρακτηριστικών .....	260
<b>10. ΑΝΑΛΥΣΗ ΕΡΕΥΝΗΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ</b> .....	263

10.1. Τα μοντέλα πρόβλεψης για το Μάθημα AA005 2020-2021 .....	264
10.1.1. Ακρίβεια και χρόνος πρόβλεψης των βασικών μοντέλων .....	264
10.1.2. Επιπλέον κριτήρια για την απόδοση των βασικών μοντέλων .....	275
10.1.3. Ακρίβεια και χρόνος πρόβλεψης του LSTM μοντέλου .....	280
10.1.4. Επιπλέον κριτήρια για την απόδοση του LSTM μοντέλου .....	283
10.2. Τα μοντέλα πρόβλεψης για το Μάθημα AA005 2021-2022 .....	285
10.2.1. Ακρίβεια και χρόνος πρόβλεψης των βασικών μοντέλων .....	285
10.2.2. Επιπλέον κριτήρια για την απόδοση των βασικών μοντέλων .....	293
10.2.3. Ακρίβεια και χρόνος πρόβλεψης του LSTM μοντέλου .....	298
10.2.4. Επιπλέον κριτήρια για την απόδοση του LSTM μοντέλου .....	301
10.3. Τα μοντέλα πρόβλεψης για τα Μαθήματα A3 2019-2020 & 2020-2021 .....	302
10.3.1. Ακρίβεια και χρόνος πρόβλεψης των βασικών μοντέλων .....	302
10.3.2. Επιπλέον κριτήρια για την απόδοση των βασικών μοντέλων .....	313
10.3.3. Ακρίβεια και χρόνος πρόβλεψης του LSTM μοντέλου .....	317
10.3.4. Επιπλέον κριτήρια για την απόδοση του LSTM μοντέλου .....	320
10.4. Έλεγχος γενίκευσης .....	321
10.5. Επιδραστικά χαρακτηριστικά .....	328
<b>11. ΑΠΑΝΤΗΣΕΙΣ ΣΤΑ ΕΡΕΥΝΗΤΙΚΑ ΕΡΩΤΗΜΑΤΑ .....</b>	<b>335</b>
11.1. Απάντηση στο Ερευνητικό Ερώτημα EEP1 .....	336
11.2. Απάντηση στο Ερευνητικό Ερώτημα EEP2 .....	340
11.3. Απάντηση στο Ερευνητικό Ερώτημα EEP3 .....	344
11.4. Απάντηση στο Ερευνητικό Ερώτημα EEP4 .....	349
11.5. Απάντηση στο Ερευνητικό Ερώτημα EEP5 .....	354
11.6. Απάντηση στο Ερευνητικό Ερώτημα EEP6 .....	357
11.7. Απάντηση στο Ερευνητικό Ερώτημα EEP7 .....	360
11.8. Απάντηση στο Ερευνητικό Ερώτημα EEP8 .....	366
<b>12. ΣΥΜΠΕΡΑΣΜΑΤΑ – ΣΥΖΗΤΗΣΗ .....</b>	<b>370</b>
12.1. Σύνοψη αποτελεσμάτων .....	370
12.2. Τα μοντέλα ως εργαλεία διαμορφωτικής αξιολόγησης .....	372
12.3. Προκλήσεις .....	373
12.3.1. Ηθικά ζητήματα .....	373

12.3.2. Ερμηνευσιμότητα .....	374
12.3.3. Γενικευσιμότητα .....	375
12.3.4. Ανάγκη θεωρητικής θεμελίωσης .....	376
12.4. Συμβολή της διατριβής .....	377
12.5. Περιορισμοί της έρευνας.....	381
12.6. Μελλοντικές επεκτάσεις .....	382
12.7.Επίλογος .....	384
<b>ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ .....</b>	<b>385</b>
ΠΑΡΑΡΤΗΜΑ Α .....	414
ΠΑΡΑΡΤΗΜΑ Β .....	419
ΠΑΡΑΡΤΗΜΑ Γ.....	421

## Κατάλογος Πινάκων

Πίνακας 1. Τρεις Γενιές Συστημάτων Αξιολόγησης (Bennett, 1998, σ. 16).....	39
Πίνακας 2. Η κατά Farrell και Rushby (2016) ταξινόμηση μορφών ψηφιακής αξιολόγησης σύμφωνα με τον μοντέλο SAMR .....	45
Πίνακας 3. Αρχές για καλές πρακτικές διαμορφωτικής αξιολόγησης ( Nicol & Macfarlane-Dick, 2006) .....	61
Πίνακας 4. Παραλλαγή ερώτησης από Τράπεζα Ερωτήσεων .....	85
Πίνακας 5. Παράδειγμα ερώτησης με ελεύθερες μεταβλητές .....	85
Πίνακας 6. Δείκτες ενεργής συμμετοχής.....	145
Πίνακας 7. Νοητικά επίπεδα και καταλληλότερες μορφές δοκιμασίας (Torres, Lopes, Babo & Azevedo, 2011) .....	163
Πίνακας 8. Μέθοδοι διασφάλισης ακαδημαϊκής ακεραιότητας (Amigud et al. 2017) .....	173
Πίνακας 9. Τα υπό μελέτη μαθήματα .....	189
Πίνακας 10. Είδη καταγραφόμενων δεδομένων για τα συμβάντα χρηστών στο ΣΔΜ Moodle	197
Πίνακας 11. Τμήμα αρχείου καταγραφής συμβάντων από το ΣΔΜ Moodle όπως εμφανίζεται σε περιβάλλον φυλλομετρητή. ....	198
Πίνακας 12. Παράδειγμα συνόλου δεδομένων με μορφή διδιάστατου πίνακα (Κ γραμμές, N+1 Στήλες) .....	206
Πίνακας 13. Παράδειγμα συνόλου δεδομένων με μορφή διδιάστατου πίνακα ( Κ γραμμές, (WxN)+1 Στήλες) με ενσωμάτωση της διάστασης του χρόνου για μελέτη με Περιοδικά Νευρωνικά Δίκτυα .....	209
Πίνακας 14. Πλήθος γραμμών ανά αρχείο καταγραφής .....	210
Πίνακας 15. Χρονική περίοδος και πλήθος εβδομάδων για τα σύνολα δεδομένων .....	211
Πίνακας 16. Σύνολα δεδομένων. Έναρξη και λήξη εβδομάδων .....	212
Πίνακας 17. Ημερομηνίες διαλέξεων και τελικής εξέτασης .....	212
Πίνακας 18. Πλήθος συνόλων δεδομένων .....	213
Πίνακας 19. Συνοπτική παρουσίαση των χαρακτηριστικών που συμπεριλήφθηκαν στα σύνολα δεδομένων .....	219
Πίνακας 20. (Ομάδα Πινάκων) Αναλυτική παρουσίαση των χαρακτηριστικών και του υπολογισμού των τιμών τους.....	222
Πίνακας 21. Αλγόριθμοι ταξινόμησης .....	254
Πίνακας 22. Ακρίβεια πρόβλεψης μοντέλων για το μάθημα AA005 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό).....	265
Πίνακας 23. Βασικοί αλγόριθμοι ταξινόμησης για την κατασκευή μοντέλων .....	267
Πίνακας 24. Τιμές μεταβλητής AUC για τα μοντέλα του μαθήματος AA005 2020-2021 .....	275
Πίνακας 25. Τιμές μεταβλητής Cohen’s Kappa για τα μοντέλα του μαθήματος AA005 2020-2021 .....	278
Πίνακας 26. Ακρίβεια πρόβλεψης LSTM μοντέλου με προκαθορισμένες παραμέτρους για το μάθημα AA005 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό).....	280

Πίνακας 27. Ακρίβεια πρόβλεψης LSTM μοντέλου με βελτιστοποίηση ως προς τις παραμέτρους για το μάθημα AA005 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό) .....	281
Πίνακας 28. Τιμές μεταβλητής AUC του LSTM μοντέλου για το μάθημα AA005 2020-2021....	283
Πίνακας 29. Τιμές μεταβλητής Cohen’s Kappa του LSTM μοντέλου για το μάθημα AA005 2020-2021 .....	284
Πίνακας 30. Ακρίβεια πρόβλεψης μοντέλων για το μάθημα AA005 2021-2022 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό).....	286
Πίνακας 31. Τιμές μεταβλητής AUC για τα μοντέλα του μαθήματος AA005 2021-2022 .....	294
Πίνακας 32. Τιμές μεταβλητής Cohen’s Kappa για τα μοντέλα του μαθήματος AA005 2021-2022 .....	296
Πίνακας 33. Ακρίβεια πρόβλεψης LSTM μοντέλου με προκαθορισμένες παραμέτρους για το μάθημα AA005 2021-2022 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό).....	298
Πίνακας 34. Ακρίβεια πρόβλεψης LSTM μοντέλου με «βελτιστοποίηση» ως προς τις παραμέτρους για το μάθημα AA005 2021-2022 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό).....	299
Πίνακας 35. Τιμές μεταβλητής AUC του LSTM μοντέλου για το μάθημα AA005 2021-2022....	301
Πίνακας 36. Τιμές μεταβλητής Cohen’s Kappa του LSTM μοντέλου για το μάθημα AA005 2021-2022 .....	301
Πίνακας 37. Ακρίβεια πρόβλεψης μοντέλων για τα μαθήματα A3 2019-2020 & 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό) .....	305
Πίνακας 38. Τιμές μεταβλητής AUC για τα μοντέλα των μαθημάτων A3 2019-2020 & 2020-2021 .....	313
Πίνακας 39. Τιμές μεταβλητής Cohen’s Kappa για τα μοντέλα των μαθημάτων A3 2019-2020 & 2021-2022 .....	315
Πίνακας 40. Ακρίβεια πρόβλεψης LSTM μοντέλου με προκαθορισμένες παραμέτρους για τα μαθήματα A3 2019-2020 & 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό) .	317
Πίνακας 41. Ακρίβεια πρόβλεψης LSTM μοντέλου με «βελτιστοποίηση» ως προς τις παραμέτρους για τα μαθήματα A3 2019-2020 & 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό).....	318
Πίνακας 42. Τιμές μεταβλητής AUC του LSTM μοντέλου για τα μαθήματα A3 2019-2020 & 2020-2021 .....	320
Πίνακας 43. Τιμές μεταβλητής Cohen’s Kappa του LSTM μοντέλου για τα μαθήματα A3 2019-2020 & 2020-2021 .....	320
Πίνακας 44. Ακρίβεια πρόβλεψης μοντέλου SimpleLogistic AA005 2021-2022 (17η εβδομάδα), που εφαρμόστηκε σε συνόλων δεδομένων του μαθήματος AA005 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό).....	323
Πίνακας 45. Τιμές μεταβλητών AUC & Cohen’s Kappa του μοντέλου SimpleLogistic AA005 2021-2022 (17η εβδομάδα) επί συνόλων δεδομένων του μαθήματος AA005 2020-2021 .....	325
Πίνακας 46. Τμήμα αναφοράς συστήματος WEKA για πρόβλεψη επίδοσης ανά οντότητα.....	326
Πίνακας 47. Επιδραστικά χαρακτηριστικά μοντέλων σύμφωνα με τον συντελεστή συσχέτισης (φθίνουσα διάταξη) μεταξύ χαρακτηριστικών και κλάσης ταξινόμησης .....	329

Πίνακας 48. Ακρίβεια πρόβλεψης του μοντέλου functions.SimpleLogistic για σύνολα δεδομένων που διαφέρουν ως προς τα επιδραστικά χαρακτηριστικά.....	330
Πίνακας 49. Συσχέτιση των επιδραστικών χαρακτηριστικών με τις διαστάσεις αλληλεπίδρασης .....	331
Πίνακας 50. Υψηλότερες τιμές ακρίβειας πρόβλεψης, τιμές AUC και τιμές Cohen’s Kappa ανά μάθημα. ....	338
Πίνακας 51. Εβδομάδα, τιμή ακρίβειας, AUC, Cohen’s Kappa πρωιμότερης πρόβλεψης ανά μάθημα .....	341
Πίνακας 52. Κατάταξη αλγόριθμων ανά μάθημα κατά ακρίβεια πρόβλεψης. Τελική και πρωιμότερη πρόβλεψη. ....	350
Πίνακας 53. Επιδραστικότερα χαρακτηριστικά για κάθε μάθημα .....	358
Πίνακας 54. Δείκτες ενεργής συμμετοχής.....	362
Πίνακας 55. Σχέσεις επιδραστικών χαρακτηριστικών και δεικτών ενεργής συμμετοχής .....	364
Πίνακας 56. Συναρτήσεις στο φύλλο Excel για τον υπολογισμό της τιμής των χαρακτηριστικών του συνόλου δεδομένων του μαθήματος AA005 Χρήση Βασικών Εφαρμογών ΗΥ. ....	414
Πίνακας 57. Συναρτήσεις στο φύλλο Excel για τον υπολογισμό της τιμής των χαρακτηριστικών του συνόλου δεδομένων του μαθήματος A3 Ηλεκτρονική Μάθηση. ....	416



## Κατάλογος Σχημάτων

Σχήμα 1. Εξελικτική πορεία συστημάτων αξιολόγησης (Redecker, 2013, σ. 11).....	41
Σχήμα 2. Διαστάσεις καινοτομίας στην ψηφιακή αξιολόγηση (Binkley et al., 2012) .....	43
Σχήμα 3. Αποτίμηση επιπέδου ικανότητας σε προσαρμοζόμενο διαγώνισμα (Linacre, 2000, σ. 6) .....	83
Σχήμα 4. Οι εξελικτική πορεία των συστημάτων αξιολόγησης (Siarova, Sternadel & Mavsidlauskaite, 2017, σ. 72) .....	120
Σχήμα 5. Κλασικές προγραμματιστικές τεχνικές έναντι Μηχανικής Μάθησης (Chollet, 2019 σ.3) .....	126
Σχήμα 6. Φάσεις εκμάθησης και δοκιμής στην τεχνική της ταξινόμησης της Μηχανικής Μάθησης.....	128
Σχήμα 7. Η δομή Περιοδικού Νευρωνικού Δικτύου (Recurrent Neural Network) .....	131
Σχήμα 8. Δείκτες ενεργής συμμετοχής (Wong & Chong, 2018, σ. 8) .....	143
Σχήμα 9. Η κατά τον Bloom (1956) ταξινόμηση των νοητικών λειτουργιών και η αναθεώρησή της (Anderson & Krathwohl, 2001).....	162
Σχήμα 10. Το κατά Kereluik et al. (2013) πλαίσιο μάθησης για τον 21 <sup>ο</sup> αιώνα .....	166
Σχήμα 11. Το συνεχές των εκπαιδευτικών περιβαλλόντων (Φεσάκης & Κωνσταντοπούλου, 2022, σσ. 34-35).....	191
Σχήμα 12. Αναπαράσταση τριδιάστατου συνόλου δεδομένων με αποτύπωση της χρονικής διάστασης .....	208
Σχήμα 13. Παράδειγμα καμπύλης ROC (Witten, Frank & Hall, 2011, p. 173).....	258
Σχήμα 14. Διασταυρούμενη επικύρωση k πτυχών (Ren, Li & Han, 2019) .....	259

## Κατάλογος Γραφημάτων

Γράφημα 1. Ακρίβεια Πρόβλεψης Μοντέλων για το μάθημα AA005 2020-2021 .....	272
Γράφημα 2. Καμπύλες ROC για τα μοντέλα του μαθήματος AA005 2020-2021 .....	277
Γράφημα 3. Ακρίβεια Πρόβλεψης LSTM Μοντέλου μετά από βελτιστοποίηση παραμέτρων .	282
Γράφημα 4. Καμπύλη ROC για το μοντέλο LSTM, Εβδομάδα 6 .....	283
Γράφημα 5. Ακρίβεια Πρόβλεψης Μοντέλων για το μάθημα AA005 2021-2022 .....	291
Γράφημα 6. Καμπύλες ROC για τα μοντέλα του μαθήματος AA005 2021-2022 .....	295
Γράφημα 7. Ακρίβεια Πρόβλεψης LSTM Μοντέλου .....	300
Γράφημα 8. Τιμές βάσεων για την ακρίβεια πρόβλεψης στα μαθήματα AA005 και A3.....	303
Γράφημα 9. Ακρίβεια πρόβλεψης μοντέλων για το μάθημα AA005 2021-2022.....	310
Γράφημα 10. Καμπύλες ROC για τα μοντέλα των μαθημάτων A3 2019-2020 & 2020-2021 ....	315
Γράφημα 11. Ακρίβεια Πρόβλεψης LSTM Μοντέλου .....	319
Γράφημα 12. Ακρίβεια πρόβλεψης μοντέλου SimpleLogistic AA005 2021-2022 (17η εβδομάδα) επί συνόλων δεδομένων του μαθήματος AA005 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό) .....	324
Γράφημα 13. Απόδοση τεχνικών «βαθιάς μάθησης» σε σχέση με την ποσότητα των δεδομένων (Alom, et al., 2019, σ. 7).....	346

## Κατάλογος Εικόνων

Εικόνα 1. Η μηχανή του Sidney L. Pressey (Petrina, 2004).....	34
Εικόνα 2. Διαδικτυακή σελίδα όπου οι φοιτητές μπορούσαν να δηλώσουν άρνηση για τη μελέτη δεδομένων πρόσβασής τους.....	193
Εικόνα 3. Παράδειγμα διαδικτυακής σελίδας μαθήματος στο ΣΔΜ Moodle (έκδοση 3.9) του Πανεπιστημίου Αιγαίου.....	195
Εικόνα 4. Το γραφικό περιβάλλον χρήσης του WEKA (έκδοση 3.9.6).....	201
Εικόνα 5. Το σύστημα φοιτητολογίου του Πανεπιστημίου Αιγαίου.....	202
Εικόνα 6. Το περιβάλλον χρήσης του Microsoft Excel (έκδοση 2209).....	203
Εικόνα 7. Το περιβάλλον χρήσης του IBM SPSS (έκδοση 28).....	204

## Ακρωνύμια

<b>AUC</b>	Area Under Curve
<b>LMS</b>	Learning Management System
<b>LSTM Networks</b>	Long Short-Term Memory Networks
<b>MOOC</b>	Massive Open Online Course
<b>Moodle</b>	Modular Object-Oriented Dynamic Learning Environment
<b>RNN</b>	Recursive Neural Networks
<b>ROC</b>	Receiver Operating Characteristic
<b>WEKA</b>	Waikato Environment for Knowledge Analysis
<b>ΑΕΙ</b>	Ανώτατα Εκπαιδευτικά Ιδρύματα
<b>ΣΔΜ</b>	Συστήματα Διαχείρισης Μάθησης
<b>ΤΕΠΑΕΣ</b>	Τμήμα Επιστημών της Προσχολικής Αγωγής και του Εκπαιδευτικού Σχεδιασμού
<b>ΤΠΕ</b>	Τεχνολογίες Πληροφορικής και Επικοινωνίας

## Γλωσσάρι ξενόγλωσσων όρων

<b>Academic Analytics</b>	Ακαδημαϊκή Αναλυτική
<b>Attribute</b>	Χαρακτηριστικό
<b>Class</b>	Κλάση / Κατηγορία
<b>Classifier</b>	Ταξινομητής
<b>Computer Aided Assessment</b>	Υπολογιστικά Υποβοηθούμενη Αξιολόγηση
<b>Computer Assisted Assessment</b>	Υπολογιστικά Υποστηριζόμενη Αξιολόγηση
<b>Computer Based Assessment</b>	Υπολογιστικά Βασισμένη Αξιολόγηση
<b>Computer Enhanced Assessment</b>	Υπολογιστικά Βελτιούμενη Αξιολόγηση
<b>Dataset</b>	Σύνολο δεδομένων
<b>Deep Learning</b>	Βαθιά Μάθηση
<b>E-assessment</b>	Ψηφιακή αξιολόγηση
<b>Educational data mining</b>	Εξόρυξη εκπαιδευτικών δεδομένων
<b>E-learning</b>	Ψηφιακή μάθηση
<b>E-portfolio</b>	Ψηφιακός φάκελος επιτευγμάτων
<b>Feature</b>	Χαρακτηριστικό
<b>Forum</b>	Χώρος δημόσιου διαλόγου
<b>Learning Analytics</b>	Μαθησιακή Αναλυτική
<b>Machine Learning</b>	Μηχανική Μάθηση
<b>Precision accuracy</b>	Ακρίβεια πρόβλεψης
<b>Prediction</b>	Πρόβλεψη
<b>Prediction model</b>	Μοντέλο πρόβλεψης
<b>Predictive Analytics</b>	Προβλεπτική αναλυτική
<b>Technology Enabled Assessment</b>	Τεχνολογικά Ενεργοποιούμενη Αξιολόγηση
<b>Technology Enhanced Assessment</b>	Τεχνολογικά Βελτιούμενη Αξιολόγηση
<b>Wiki</b>	Εργαλείο συνεργατικής ανάπτυξης περιεχομένου

## ΕΙΣΑΓΩΓΗ

Ένα από τα φραστικά σχήματα με τα οποία συχνά ξεκινούν οι διδακτορικές διατριβές είναι το «*Η ανά χείρας διατριβή ...*». Προς στιγμή, μπήκαμε στον πειρασμό να το χρησιμοποιήσουμε, ώστε, με μια φράση που συνηθίζεται, να υπερβούμε την αμηχανία μπροστά στο ερώτημα: «Πώς αρχίζουμε»;

Τελικά, δεν θα χρησιμοποιήσουμε το «*Η ανά χείρας διατριβή...*». Διότι, η φράση έχει νόημα για τις διατριβές που έχουν έντυπη μορφή, όπου το «*ανά χείρας*» κυριολεκτεί. Τη διατριβή μας, ελάχιστοι θα την κρατήσουν στα χέρια τους. Σχεδόν όλοι θα τη διαβάσουν στην οθόνη κάποιας ψηφιακής συσκευής. Πλέον, οι διατριβές, οι ερευνητικές μελέτες, τα επιστημονικά άρθρα, τα εκπαιδευτικά συγγράμματα, το εκπαιδευτικό υλικό παράγονται, διακινούνται και μελετώνται ψηφιακά. Αποτελούν δομικά στοιχεία ενός ψηφιακού εκπαιδευτικού περιβάλλοντος, το οποίο συγκροτείται από ψηφιακούς πόρους, ψηφιακές υπηρεσίες, ψηφιακές αλληλεπιδράσεις. Ο κόσμος της εκπαίδευσης, όπως και όλος ο κόσμος, μετασχηματίζεται. Παραμέτρους αυτού του ψηφιακού εκπαιδευτικού περιβάλλοντος και διαστάσεις αυτού του μετασχηματισμού, διαπραγματεύεται η διατριβή, επιχειρώντας να συνεισφέρει στην έρευνα για τις καινοτόμες μορφές εκπαιδευτικής αξιολόγησης στα νέα ψηφιακά εκπαιδευτικά περιβάλλοντα.

Τα εκπαιδευτικά ιδρύματα, ιδίως αυτά της τριτοβάθμιας εκπαίδευσης, αξιοποιούν συστήματα ψηφιακής μάθησης και προσφέρουν μαθήματα, είτε αποκλειστικώς διαδικτυακά, είτε με μεικτούς τρόπους διδασκαλίας. Καθώς, οι νέοι τρόποι μάθησης διευρύνονται και καθιερώνονται, προκύπτουν ανάγκες, αλλά και δυνατότητες για τον ανασχεδιασμό των αξιολογικών διαδικασιών. Με τη βοήθεια νέων τεχνολογικών μεθόδων και εργαλείων, αναπτύσσονται μέθοδοι ψηφιακής εκπαιδευτικής αξιολόγησης για να υποστηρίξουν τις μορφές της ψηφιακής μάθησης.

Το πεδίο της ψηφιακής εκπαιδευτικής αξιολόγησης είναι εξαιρετικά ευρύ. Περιλαμβάνει συστήματα και εφαρμογές που, αυτοματοποιώντας παραδοσιακές μεθόδους, έχουν βελτιώσει εντυπωσιακά την αποδοτικότητα των αξιολογικών διαδικασιών. Περιλαμβάνει επίσης καινοτόμα συστήματα, τα οποία προτείνουν νέες μεθόδους, μέσω των οποίων η αξιολόγηση

αποκτά νέες πρωτόγνωρες μορφές και ενσωματώνεται στα νέα ψηφιακά εκπαιδευτικά περιβάλλοντα.

Στα πλαίσια της διατριβής εξετάσαμε και σχολιάσαμε τις βασικές μορφές που μπορεί να λάβει η ψηφιακή αξιολόγηση, αλλά επικεντρώσαμε την ερευνητική προσπάθειά μας στις καινοτόμες μορφές της. Ειδικότερα, σχεδιάσαμε και υλοποιήσαμε έρευνα στα πλαίσια του πεδίου της Μαθησιακής Αναλυτικής, του πεδίου που μελετά τον τεράστιο όγκο δεδομένων που παράγουν και διαχειρίζονται τα εκπαιδευτικά ιδρύματα. Με τη βοήθεια μεθόδων, εργαλείων και τεχνικών της Μηχανικής Μάθησης κατασκευάσαμε μοντέλα πρόβλεψης για την επίδοση των σπουδαστών και διερευνήσαμε δυνατότητες για τη χρήση τους ως εργαλείων, που θα υποστηρίξουν διαμορφωτικές αξιολογικές διαδικασίες και κατάλληλες διδακτικές ενισχυτικές παρεμβάσεις.

Η διατριβή είναι οργανωμένη σε θεωρητικό και ερευνητικό μέρος και απαρτίζεται από τα ακόλουθα δώδεκα κεφάλαια:

### **Θεωρητικό μέρος:**

Το πρώτο κεφάλαιο εξετάζει τις αλλαγές που επιφέρουν στην εκπαίδευση τα νέα ψηφιακά περιβάλλοντα μάθησης και διερευνά τους σχετικούς μετασχηματισμούς στις διαδικασίες της εκπαιδευτικής αξιολόγησης. Οριοθετεί την έννοια της ψηφιακής αξιολόγησης καταγράφοντας τις διαφορετικές οπτικές των ερευνητών και σχολιάζει τους ορισμούς που κατά καιρούς έχουν προταθεί για να ορίσουν το πεδίο.

Το δεύτερο κεφάλαιο αναφέρεται στην εξέλιξη των συστημάτων της ψηφιακής αξιολόγησης και καταγράφει τη μετάβαση από τη χρήση συστημάτων που απλώς επιδιώκουν την αυτοματοποίηση αξιολογήσεων παραδοσιακού τύπου, προς τη χρήση συστημάτων που επαναπροσδιορίζουν την αξιολόγηση και την ενσωματώνουν στη μαθησιακή διαδικασία.

Το τρίτο κεφάλαιο εστιάζει στον ρόλο της ψηφιακής αξιολόγησης στην αποδοτικότερη λειτουργία των διαμορφωτικών αξιολογικών διαδικασιών, στη βελτίωση της ανατροφοδότησης, στην ενδυνάμωση αλληλεπιδραστικών και συνεργατικών αξιολογικών διαδικασιών.

Το τέταρτο κεφάλαιο αναφέρεται στις μορφές με τις οποίες υλοποιείται και εφαρμόζεται η ψηφιακή αξιολόγηση. Παρουσιάζονται μορφές ψηφιακής αξιολόγησης που ψηφιοποιούν παραδοσιακές αξιολογικές διαδικασίες, αλλά και καινοτόμες μορφές που υλοποιούνται μόνο με τη βοήθεια σύγχρονων τεχνολογιών.

Το πέμπτο κεφάλαιο αναφέρεται στην Μαθησιακή Αναλυτική, μεθόδους της οποίας αξιοποιήσαμε στα πλαίσια της διατριβής. Οριοθετεί το πεδίο, καταγράφει πώς τα σχετικά εργαλεία αξιοποιούνται για αξιολογικούς σκοπούς και παραθέτει συνοπτικά σχετικές ερευνητικές προσπάθειες.

Το έκτο κεφάλαιο καταγράφει τα οφέλη που προκύπτουν από τη χρήση συστημάτων ψηφιακής αξιολόγησης, τα οποία βελτιώνουν την αποτελεσματικότητα και την αποδοτικότητα και αναβαθμίζουν ποιοτικά τις αξιολογικές διαδικασίες.

Το έβδομο κεφάλαιο αναφέρεται στις ηθικές, πολιτικές, παιδαγωγικές, διοικητικές και τεχνικές προκλήσεις που προκύπτουν από τη χρήση των συστημάτων ψηφιακής αξιολόγησης.

### **Ερευνητικό μέρος:**

Στο όγδοο κεφάλαιο παρουσιάζεται η προβληματική επί της οποίας αναπτύχθηκε η έρευνα. Περιγράφονται το πεδίο, οι σκοποί, οι στόχοι και τα ερευνητικά ερωτήματα της διατριβής.

Το ένατο κεφάλαιο περιγράφει τη μεθοδολογία της έρευνας. Γίνεται αναφορά στον πληθυσμό που εστίασε η έρευνα, στον τύπο και στο τρόπο συλλογής των δεδομένων, στα εργαλεία που χρησιμοποιήθηκαν. Επιπλέον, περιγράφονται τα μεθοδολογικά διλήμματα που ανέκυψαν και οι σχετικές αποφάσεις που λήφθηκαν.

Στο δέκατο κεφάλαιο παρουσιάζονται, αναλύονται και σχολιάζονται τα αποτελέσματα της έρευνας.

Στο ενδέκατο κεφάλαιο, αναπτύσσονται οι απαντήσεις στα ερευνητικά ερωτήματα της διατριβής, όπως αυτές προέκυψαν από την ανάλυση των ερευνητικών αποτελεσμάτων.

Το δωδέκατο κεφάλαιο ολοκληρώνει τη διατριβή. Ειδικότερα, συνοψίζει τα αποτελέσματα της έρευνας, αναφέρει τις προκλήσεις που αναδείχθηκαν κατά τη διάρκειά της, καταγράφει τη συμβολή της διατριβής και προτείνει μελλοντικές προεκτάσεις της.



---

## **ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ**

---

# 1. Η ΥΠΟΛΟΓΙΣΤΙΚΑ ΥΠΟΣΤΗΡΙΖΟΜΕΝΗ ΕΚΠΑΙΔΕΥΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ / ΨΗΦΙΑΚΗ ΑΞΙΟΛΟΓΗΣΗ

## 1.1 Η αξιολόγηση στα νέα περιβάλλοντα μάθησης

Η αξιολόγηση κατέχει κεντρική θέση στην εκπαίδευση. Παρέχει απτές πληροφορίες για τις γνώσεις και τις δεξιότητες των σπουδαστών και κρίνει την επίτευξη των στόχων που καθορίζονται στα προγράμματα σπουδών. Η αξιολόγηση δεν είναι μόνον μέσο για τη μέτρηση εκπαιδευτικών επιδόσεων, αλλά συχνά ορίζει τις προτεραιότητες της εκπαίδευσης (Kapsalis, et al., 2019) και αποκαλύπτει τον τρόπο με τον οποίο οι εκπαιδευτικοί οργανισμοί αντιλαμβάνονται τη διαδικασία της μάθησης (Oldfield, Broadfoot, Sutherland & Timmis, 2012). Παράλληλα, αποτελεί βασικό στοιχείο της μαθησιακής εμπειρίας των σπουδαστών (Rust, O'Donovan & Price, 2005) και επηρεάζει σε μεγάλο βαθμό τον τρόπο με τον οποίο οι σπουδαστές προσεγγίζουν τη μάθηση και τη μελέτη. Όπως αναφέρουν οι Russell, Elton, Swinglehurst & Greenhalgh (2006), η αξιολόγηση βοηθά τους σπουδαστές να επεξεργασθούν με ενεργό τρόπο τι έμαθαν, να ανακαλύψουν τι γνωρίζουν, να εστιάσουν τη μελέτη τους και να ασχοληθούν με το γνωστικό αντικείμενο, ενώ παράλληλα, η επιβράβευση μέσω της βαθμολογίας και η λήψη ανατροφοδότησης αποτελούν κίνητρα για μάθηση (Nicol, 2008). Συνεπώς, οι μέθοδοι και τα συστήματα αξιολόγησης έχουν μεγάλη σημασία, όχι μόνο για την μέτρηση των επιδόσεων των σπουδαστών και για την αποτίμηση των γνωστικών ικανοτήτων τους, αλλά για το σύνολο της εκπαιδευτικής διαδικασίας (Lafuente, Remesal & Álvarez Valdivia, 2014).

Οι μέθοδοι και τα συστήματα της αξιολόγησης συχνά αποτελούν αντικείμενο κριτικής (τόσο από τους ερευνητές όσο και από τους σπουδαστές), η οποία συνοδεύεται από αιτήματα για σχετικές μεταρρυθμίσεις (Oldfield et al., 2012). Οι Oldfield et al. (2012) αναφέρουν ότι οι χρησιμοποιούμενες μέθοδοι αξιολόγησης, ιδιαίτερες αυτές που σχετίζονται με υψηλής σπουδαιότητας τελικές εξετάσεις κάθε εκπαιδευτικής βαθμίδας, συχνά χαρακτηρίζονται ως απαρχαιωμένες, αναποτελεσματικές και καταστροφικές. Οι Nicol & Macfarlane-Dick (2006) σημειώνουν ότι, παρά τις αλλαγές που παρατηρούνται στην ανώτατη εκπαίδευση προς την κατεύθυνση ενός μαθητοκεντρικού μοντέλου μάθησης (όπου οι σπουδαστές θα συμμετέχουν ενεργά και θα νιώσουν υπεύθυνοι για τη μάθησή τους), οι απαιτούμενες μεταρρυθμίσεις στην

αξιολόγηση υλοποιούνται με βραδείς ρυθμούς. Οι Ferrell & Sheppard (2012) αναφέρουν ότι οι σπουδαστές των ανώτατων εκπαιδευτικών ιδρυμάτων στο Ηνωμένο Βασίλειο, δηλώνουν ότι περισσότερο από οποιαδήποτε άλλη εκπαιδευτική εμπειρία, δεν είναι ικανοποιημένοι με την αξιολόγηση. Συχνά, οι σπουδαστές περιγράφουν την αξιολόγηση ως διαδικασία που προκαλεί «φόβο και άγχος» (Vaughan, Martha & Garrison, 2013, σ. 81).

Την ανάγκη για μεταρρυθμίσεις επιτείνουν οι νέες προκλήσεις που αντιμετωπίζουν οι εκπαιδευτικοί οργανισμοί κάθε εκπαιδευτικής βαθμίδας και ιδιαιτέρως της ανώτατης εκπαίδευσης, όπου είναι ευδιάκριτες δύο ισχυρές και συμπληρωματικές τάσεις: η τάση της ενσωμάτωσης της διαδικτυακής (online) διδασκαλίας σε καθημερινές εκπαιδευτικές πρακτικές και η τάση της συνεχώς αυξανόμενης εμπλοκής των ανώτατων εκπαιδευτικών ιδρυμάτων σε δράσεις εξ αποστάσεως εκπαίδευσης (Larreamendy-Joerns & Leinhardt, 2006). Συνεπώς, καθώς οι μορφές της διαδικτυακής (online) και της μεικτής (blended) μάθησης καθιερώνονται στις εκπαιδευτικές πρακτικές της ανώτατης εκπαίδευσης, προκύπτει η ανάγκη για αναθεώρηση βασικών ζητημάτων της διδασκαλίας, της μάθησης και της αξιολόγησης στα νέα, ψηφιακά, μη παραδοσιακά εκπαιδευτικά περιβάλλοντα (Gikandi, Morrow & Davis, 2011).

Για την αντιμετώπιση των νέων προκλήσεων που αντιμετωπίζει η αξιολόγηση (π.χ. εξ αποστάσεως εκπαίδευση, μεγάλος αριθμός σπουδαστών, αντικειμενική και υψηλής ποιότητας ανατροφοδότηση) απαιτούνται αλλαγές στις παραδοσιακές αξιολογικές πρακτικές (Whitelock & Watt, 2008), οι οποίες είναι αδύνατον πλέον να παραμένουν ουσιαστικώς ίδιες, όπως ήταν δύο αιώνες πριν, όταν η εκπαιδευτική αξιολόγηση άρχισε να εφαρμόζεται σε μεγάλη κλίμακα (Broadfoot, Oldfield, Sutherland, & Timmis, 2014). Το νέο τοπίο που διαμορφώνεται στην εκπαίδευση και οι ανατροπές που επιφέρει η ψηφιακή μάθηση (e-learning) σε καθιερωμένες δομές και πρακτικές, οδηγούν στην αξιοποίηση τεχνολογικών καινοτομιών στην εκπαιδευτική αξιολόγηση (Russell, Elton, Swinglehurst, & Greenhalgh, 2006). Τα νέα, ψηφιακά εκπαιδευτικά περιβάλλοντα αυξάνουν την πολυπλοκότητα του εγχειρήματος για τον ανασχεδιασμό των αξιολογικών διαδικασιών. Ο ανασχεδιασμός πρέπει ταυτοχρόνως να επιδιώξει την αποβολή των αρνητικών χαρακτηριστικών της παραδοσιακής αξιολόγησης, την διατήρηση των θετικών χαρακτηριστικών της και την αξιοποίηση των δυνατοτήτων που προσφέρει η τεχνολογία για νέες καινοτόμες μορφές αξιολόγησης, που δεν μπορούν να υλοποιηθούν στα πλαίσια των

παραδοσιακών περιβαλλόντων μάθησης (Conrad & Oropo, 2018). Ταυτοχρόνως, οι απαιτούμενες αλλαγές είναι αναγκαίο να συμβαδίζουν με τις σύγχρονες παιδαγωγικές, κοινωνικές και τεχνολογικές εξελίξεις οι οποίες επηρεάζουν την εκπαίδευση και τη μάθηση (Sweeney, et al., 2017).

## 1.2. ΤΠΕ και Αξιολόγηση

Οι Τεχνολογίες της Πληροφορικής και της Επικοινωνίας (ΤΠΕ) μετασχηματίζουν όλες τις βαθμίδες της εκπαίδευσης (Lafuente, Remesal & Álvarez Valdivia, 2014). Η επίδραση τους είναι καταλυτική σε όλα τα στάδια της εκπαιδευτικής διαδικασίας, συνεπώς και στην αξιολόγηση, η οποία αναπόφευκτα επηρεάζεται από τις σύγχρονες εξελίξεις και προσαρμόζεται στο νέο συνεχώς μεταβαλλόμενο εκπαιδευτικό τοπίο. Η αξιολόγηση προκειμένου να επιτελέσει τον σύνθετο ρόλο της, οφείλει να αξιοποιήσει τις τεχνολογικές καινοτομίες που ευρέως χρησιμοποιούνται στην εκπαίδευση και να διερευνήσει τις δυνατότητες καινοτόμων τεχνολογικών εργαλείων. Η ευθυγράμμιση των καινοτόμων μαθησιακών εμπειριών με τις μεθόδους αξιολόγησης, σημαίνει ότι όσο διευρύνεται η χρήση των ΤΠΕ στην εκπαιδευτική διαδικασία, τόσο θα διευρύνεται και η χρήση των ΤΠΕ στην εκπαιδευτική αξιολόγηση (James, 2016).

Οι πρώτες προσπάθειες αξιοποίησης των ΤΠΕ στην εκπαιδευτική αξιολόγηση αφορούσαν στην αυτοματοποίηση της βαθμολόγησης και αποσκοπούσαν στην ελάφρυνση του φόρτου εργασίας των διδασκόντων (Chiou, Hwang & Tseng, 2009. Noorberhbahani & Kardan, 2011. Stödberg, 2012). Ωστόσο, η αξιοποίηση των ΤΠΕ στην αξιολόγηση μπορεί πλέον να υποστηρίξει, όχι μόνο τη μέτρηση των επιδόσεων των σπουδαστών, αλλά επίσης να βελτιώσει όλα τα στάδια της εκπαιδευτικής διαδικασίας (Daly, Pachler, Yishay & Harvey, 2010. Russell, Elton, Swinglehurst & Greenhalgh, 2006) και να προσφέρει διοικητικά οφέλη (Tomas, Borg & McNeil, 2015). Στην αξιολόγηση, όπου τα κίνητρα συμμετοχής των σπουδαστών είναι κυρίως εξωγενή, οι ΤΠΕ μπορούν μεταξύ άλλων να πολλαπλασιάσουν και να εμπλουτίσουν τις συμμετοχικές και αλληλεπιδραστικές εμπειρίες των σπουδαστών, ώστε να ενισχυθεί το ενδιαφέρον και η ενεργή συμμετοχή τους (Bull, et al., 2002). Επίσης είναι δυνατόν να ενισχύσουν την ομότιμη αξιολόγηση, την αυτό-αξιολόγηση, της εξ αποστάσεως και ευέλικτη μάθηση, την ποιοτική και έγκαιρη ανατροφοδότηση (Sweeney, et al., 2017). Προφανώς, η συνεισφορά των ΤΠΕ στον

μετασχηματισμό της αξιολόγησης δεν πρέπει να επικεντρώνεται αποκλειστικά στη ψηφιακή υλοποίηση παραδοσιακών αξιολόγησης πρακτικών, αλλά πρέπει να αναπροσαρμόζεται ανάλογα με τις δυνατότητες και της προκλήσεις της τεχνολογίας, κυρίως δε, να λειτουργεί στα πλαίσια του ασύγχρονου και εξ' αποστάσεως τρόπου συνεργασίας των συμμετεχόντων (Perera-Diltz & Moe, 2014). Είναι επίσης απαραίτητο, κατά την αξιοποίηση των ΤΠΕ σε αξιολογικές διαδικασίες, να λαμβάνονται υπόψη ζητήματα σχετικά με την εγκυρότητα και την αξιοπιστία της διαδικασίας (Gikandi, Morrow & Davis, 2011). Επιπλέον, οι σχετικές προσπάθειες πρέπει να βασίζονται σε στέρεες παιδαγωγικές θεωρίες και να μην παραγνωρίζουν την πολυπλοκότητα του εγχειρήματος, την πληθώρα των εμπλεκόμενων παραγόντων και τις σύνθετες αλληλεπιδράσεις που διέπουν την εκπαιδευτική διαδικασία (Oldfield, Broadfoot, Sutherland & Timmis, 2012).

### 1.3. Ορολογία

Βασική απότοκος της αξιοποίησης των ΤΠΕ στο πεδίο της εκπαιδευτικής αξιολόγησης είναι η τεχνολογική καινοτομία που φέρει το όνομα **ψηφιακή / ηλεκτρονική αξιολόγηση** (ελληνική απόδοση του όρου **e-assessment**). Σύμφωνα με την «Κοινή Επιτροπή Πληροφοριακών Συστημάτων» των Ακαδημαϊκών και Ερευνητικών Ιδρυμάτων του Ηνωμένου Βασιλείου (JISC), με τον όρο «ψηφιακή αξιολόγηση» ορίζουμε τις *«από άκρο σε άκρο (end to end) ηλεκτρονικές διαδικασίες αξιολόγησης, όπου οι ΤΠΕ χρησιμοποιούνται για να προσφερθεί η δραστηριότητα της αξιολόγησης και για να καταγραφούν οι αποκρίσεις. Αυτό περιλαμβάνει την από άκρο σε άκρο διαδικασία αξιολόγησης, όπως την αντιλαμβάνονται οι σπουδαστές, οι διδάσκοντες, οι εκπαιδευτικοί φορείς, οι εξεταστικοί φορείς, οι ελεγκτικοί φορείς και το ευρύ κοινό»* (JISC, e-Assessment Glossary (Extended) v1.1, 2006, σ. 43). Σύμφωνα με την ίδια Επιτροπή (JISC, Effective Practice with e-Assessment, 2007), οι δραστηριότητες που καλύπτει ο όρος «ψηφιακή αξιολόγηση» περιλαμβάνουν τη σχεδίαση και τη διανομή των αξιολογήσεων, τη βαθμολόγηση – που διεκπεραιώνουν υπολογιστές ή άνθρωποι που υποστηρίζονται από σαρωτές και δικτυακά εργαλεία - και όλες τις διαδικασίες που σχετίζονται με την αναφορά, την αποθήκευση και τη διανομή των δεδομένων τα οποία αφορούν δημόσιες ή εσωτερικές αξιολογήσεις.

Η Επιτροπή, προκειμένου να διαχωρίσει τους διαφορετικούς τύπους των σχετικών δραστηριοτήτων, αναφέρει ότι γίνεται χρήση και των όρων α) Υπολογιστικά Βασισμένη

Αξιολόγηση (Computed Based Assessment), η οποία σχετίζεται με αξιολογήσεις που διανέμονται και βαθμολογούνται από υπολογιστές και β) Υπολογιστικά Υποστηριζόμενη Αξιολόγηση (Computer Assisted Assessment), η οποία εν μέρει σχετίζεται με υπολογιστές – π.χ. με διαδικτυακούς χώρους συζητήσεων για ομότιμη αξιολόγηση (peer-assessment), με συστήματα απόκρισης κοινού (audience response systems), με ολοκλήρωση και υποβολή εργασιών ψηφιακά, με αποθήκευση εργασιών σε ψηφιακούς φάκελους επιτευγμάτων (e-portfolios). Η Επιτροπή όμως σημειώνει, ότι οι αναφερόμενοι δύο όροι συχνά αντιμετωπίζονται ως εναλλάξιμοι. Το 2010, σε νεότερη σχετική έκθεσή της (JISC, Effective Assessment in a Digital Age, 2010), η Επιτροπή, παρατηρεί ότι ο όρος «ψηφιακή αξιολόγηση» συχνά χρησιμοποιείται - όπως για παράδειγμα από τους Cook & Jenkins (2009) - για να δηλώσει αποκλειστικά την αξιολόγηση μπροστά σε οθόνη (on-screen assessment), οπότε προκρίνει τη χρήση του όρου Τεχνολογικά Ενεργοποιούμενη Αξιολόγηση (Technology Enabled Assessment) και του όρου Τεχνολογικά Βελτιούμενη Αξιολόγηση (Technology Enhanced Assessment), ο οποίος δηλώνει «τη χρήση της τεχνολογίας για να επεκτείνει ή να προσθέσει αξία στις διαδικασίες της αξιολόγησης και της ανατροφοδότησης». Εντούτοις, σύμφωνα με την ίδια έκθεση της Επιτροπής (JISC 2010), ο όρος «ψηφιακή αξιολόγηση», με την ευρεία του έννοια, είναι δυνατόν να αναφέρεται σε όλες τις τεχνολογικά ενεργοποιούμενες διαδικασίες αξιολόγησης. Τον ορισμό της JISC (2010) για την Τεχνολογικά Βελτιούμενη Αξιολόγηση (Technology Enhanced Assessment), υποστηρίζουν οι Sweeney, West, Groessler, Haynie, Higgs, Macaulay et al. (2017), τονίζοντας ότι πρόκειται για ορισμό που βασίζεται στην ιδέα ότι οι σχετικές με την αξιολόγηση εκπαιδευτικές και παιδαγωγικές θεωρίες αποτελούν θεμελιώδη στοιχεία.

Τον ορισμό της Επιτροπής JISC (2006) υιοθετεί η Redecker (2013) και ο Stödberg, (2012, p. 591) ο οποίος αναφέρει ότι η ψηφιακή αξιολόγηση *«περιλαμβάνει όλη τη διαδικασία της αξιολόγησης, από τον σχεδιασμό των εργασιών, ως την καταχώρηση των αποτελεσμάτων με τη βοήθεια των ΤΠΕ»*, παρατήρηση με την οποία συμφωνούν και οι Alrunarus, Gills & Wald (2018), οι οποίοι τονίζουν ότι ο ορισμός της επιτροπής JISC (2006), δηλώνει τη χρήση των ΤΠΕ σε όλη τη διαδικασία αξιολόγησης, από την προσφορά των ερωτήσεων, ως την αποθήκευση των απαντήσεων. Η Jordan (2013, σ. 88) τονίζει ότι η ευρύτητα του ορισμού της JISC (2006) επιτρέπει να συμπεριλάβουμε στον όρο *«οποιαδήποτε χρήση υπολογιστή για οποιαδήποτε*

*δραστηριότητα σχετίζεται με την αξιολόγηση (τελική, διαμορφωτική, διαγνωστική)».* Συμπληρώνει επιπλέον, ότι ο όρος περιλαμβάνει την διαδικτυακή (online) υποβολή εργασιών (τις οποίες θα βαθμολογήσει άνθρωπος), την αξιολόγηση ψηφιακών φακέλων επιτευγμάτων (e-portfolios) ή ιστολογίων (blog), την ανατροφοδότηση μέσω αρχείων ήχου αποθηκευμένων σε υπολογιστή και, περισσότερο από οτιδήποτε άλλο, τα διαδικτυακά (online) διαγωνίσματα. Οι Ridgway, McCusker & Pead (2004, σ. 41) ορίζουν τη ψηφιακή αξιολόγηση ως «*τις διαδικασίες που συμμετέχουν στην αξιοποίηση των ΤΠΕ για την καταγραφή, μετάδοση, παρουσίαση και επεξεργασία του υλικού της αξιολόγησης*». Ο Crisp (2011, σ. 5) αναφέρει ότι «*ψηφιακή αξιολόγηση περιλαμβάνει τη χρήση ψηφιακών συσκευών στην υποστήριξη της κατασκευής, διανομής, αποθήκευσης ή παρουσίασης των εργασιών, των απαντήσεων, των βαθμών ή των ανατροφοδοτήσεων της αξιολόγησης*». Συμπληρώνει ότι, η ψηφιακή αξιολόγηση μπορεί να υλοποιηθεί με συσκευές όπως ηλεκτρονικοί υπολογιστές, τηλέφωνα, ταμπλέτες, ή κονσόλες ψηφιακών παιχνιδιών και μπορεί να χρησιμοποιήσει κείμενο ή πολυμεσικό υλικό, να περιλάβει προσομοιώσεις και παιχνίδια, να υλοποιηθεί ατομικά ή ομαδικά, σύγχρονα ή ασύγχρονα.

Οι Oldfield et al. (2012), χρησιμοποιούν τον όρο «ψηφιακή αξιολόγηση» για να ορίσουν οποιαδήποτε χρήση ψηφιακής τεχνολογίας για τους σκοπούς της εκπαιδευτικής αξιολόγησης και τονίζουν ότι ο όρος δεν προσδιορίζει μόνο τα ψηφιακά διαγωνίσματα, αλλά και άλλες τεχνολογίες, που επίσης μπορούν να αξιοποιηθούν στην εκπαιδευτική αξιολόγηση, όπως ψηφιακά περιβάλλοντα εμπύθισης, προσομοιώσεις, σοβαρά παιχνίδια, εργαλεία Web 2.0, κινητές συσκευές, ψηφιακούς φακέλους επιτευγμάτων (e-portfolios), διαδικτυακά συστήματα ψηφοφορίας ή συσκευές απόκρισης σπουδαστών (learner response units). Στον ορισμό της JISC (2006) βασίζονται και οι Tomas, Borg & McNeil (2015), οι οποίοι ορίζουν τη ψηφιακή αξιολόγηση ως τη μεσολάβηση («to mediate») των ΤΠΕ σε οποιοδήποτε τμήμα της αξιολόγησης. Αναφέρουν ότι πρόκειται για έναν ευρύ και γενικό όρο («umbrella term»), που περιγράφει ένα σύνθετο σύστημα με εργαλεία διαφορετικών δυνατοτήτων και καλύπτει αφενός την Υπολογιστικά Βασιζόμενη Αξιολόγηση (Computer Based Assessment), παράδειγμα της οποίας αποτελούν τα ψηφιακά διαγωνίσματα και αφετέρου την Υπολογιστικά Υποστηριζόμενη Αξιολόγηση (Computer Assisted Assessment), η οποία σχετίζεται με ποικιλία εργαλείων και περιλαμβάνει διαφορετικές μεταξύ τους εφαρμογές, από την διαδικτυακή υποβολή εργασιών, ως τη χρήση

wiki και ψηφιακών φακέλων επιτευγμάτων (e-portfolios). Οι Boitshwarelo, Reedy & Billny (2017) αναφέρουν ότι ο όρος «ψηφιακή αξιολόγηση» χρησιμοποιείται για να περιγράψει γενικά οποιαδήποτε πρακτική, όπου η τεχνολογία χρησιμοποιείται για να βελτιώσει ή να υποστηρίξει τις διαδικασίες της αξιολόγησης και της ανατροφοδότησης. Οι Azevedo et al. (2019) ορίζουν τη ψηφιακή αξιολόγηση, ως τη χρήση των ΤΠΕ στη διαδικασία αξιολόγησης των σπουδαστών, συμπληρώνοντας ότι περιλαμβάνει το σύνολο της διαδικασίας αξιολόγησης και καλύπτει ένα ευρύ πεδίο δραστηριοτήτων και συγκεκριμένα το σχεδιασμό και τη διανομή των δραστηριοτήτων που θα αξιολογηθούν, τη βαθμολόγηση, τις αναφορές, τη μεταφορά και αποθήκευση των αποτελεσμάτων. Σημειώνουν επίσης ότι οι όροι «Ψηφιακή Αξιολόγηση», «Υπολογιστικά Υποστηριζόμενη Αξιολόγηση» και «Υπολογιστικά Βασισμένη Αξιολόγηση» **«είναι ισοδύναμοι και αντιπροσωπεύουν την ίδια έννοια»** (σ. 5). Οι Sim, Holifield & Brown (2004) προτιμούν τον όρο Υπολογιστικά Υποστηριζόμενη Αξιολόγηση, ο οποίος θεωρούν ότι περιλαμβάνει την χρήση των υπολογιστών στη διανομή, βαθμολόγηση και ανάλυση εργασιών ή διαγωνισμάτων. Οι Timmis, Broadfoot, Sutherland & Oldfield (2016) επιλέγουν τον όρο Τεχνολογικά Βελτιούμενη Αξιολόγηση, τον οποίο ορίζουν ως την οποιαδήποτε χρήση ψηφιακών τεχνολογιών με σκοπό τη βελτίωση της τυπικής ή άτυπης, διαμορφωτικής ή τελικής εκπαιδευτικής αξιολόγησης.

Προκύπτει από όσα αναφέρθηκαν, ότι ο όρος που περιγράφει την αξιοποίηση των ΤΠΕ στην εκπαιδευτική αξιολόγηση διαχρονικά έχει υποστεί μεταβολές και συναντάται στη βιβλιογραφία με αρκετές μορφές. Συγκεκριμένα, εκτός από τον όρο «ψηφιακή αξιολόγηση» έχει προταθεί η χρήση όρων όπως Υπολογιστικά Βασισμένη Αξιολόγηση (Computer Based Assessment), Υπολογιστικά Υποστηριζόμενη Αξιολόγηση (Computer Assisted Assessment), Υπολογιστικά Υποβοηθούμενη Αξιολόγηση (Computer Aided Assessment), Υπολογιστικά Βελτιούμενη Αξιολόγηση (Computer Enhanced Assessment), Τεχνολογικά Ενεργοποιούμενη Αξιολόγηση (Technology Enabled Assessment) και Τεχνολογικά Βελτιούμενη Αξιολόγηση (Technology Enhanced Assessment) κ.ά. Η Επιτροπή JISC (2006, 2007, 2010), οι Oldfield et al. (2012), η Redecker (2013), η Jordan (2013), οι Timmis et al. (2016), οι Boitshwarelo et al. (2017), οι Perrotta & Whitelock (2017), οι Azevedo et al. (2019) θεωρούν τους όρους **συνώνυμους ή έστω ισοδύναμους**. Αξιοσημείωτη είναι η παρατήρηση των Timmis et al. (2016), σύμφωνα με τους

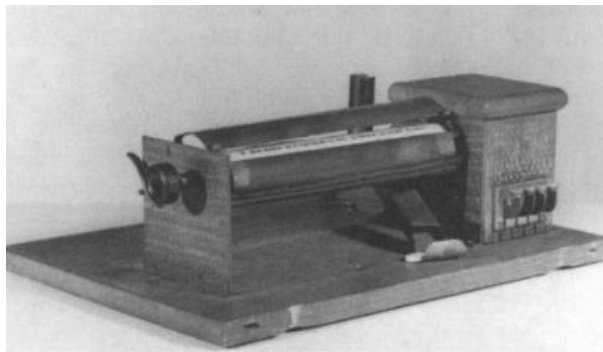


οποίους η εξέλιξη της ορολογίας υποδηλώνει ότι με την πάροδο του χρόνου διευρύνθηκαν οι αντιλήψεις για τις δυναμικές χρήσεις της τεχνολογίας στην αξιολόγηση· εντούτοις οι κατά καιρούς προτιμώμενοι όροι δεν προτείνουν κάποια θεμελιώδη αλλαγή στο πεδίο της αξιολόγησης μέσω της χρήσης της τεχνολογίας.

Στα πλαίσια της παρούσας εργασίας, **συντασσόμαστε με την άποψη περί συνωνυμίας και ισοδυναμίας των όρων**. Συνεπώς, επιλέγουμε εφεξής να χρησιμοποιούμε τον όρο «ψηφιακή / ηλεκτρονική αξιολόγηση» (e-assessment) ή απλώς «**ψηφιακή αξιολόγηση**», κυρίως διότι είναι αντίστοιχος με τον ευρέως καθιερωμένο όρο «ψηφιακή / ηλεκτρονική μάθηση» (e-learning) ή απλώς «ψηφιακή μάθηση» (Φεσάκης & Κωνσταντοπούλου, 2022, σ. 32), αλλά και διότι είναι συνοπτικός και επαρκώς εύστοχος. Επιπλέον, αντιλαμβανόμαστε και χρησιμοποιούμε τον όρο «ψηφιακή αξιολόγηση» με την ευρεία του έννοια (όπως π.χ. τον αντιλαμβάνονται οι Oldfield et al. (2012)), δηλαδή **με την έννοια της αξιοποίησης των ΤΠΕ σε οποιαδήποτε διαδικασία περιλαμβάνεται στην εκπαιδευτική αξιολόγηση**.

## 2. ΕΞΕΛΙΞΗ ΚΑΙ ΤΑΞΙΝΟΜΗΣΗ ΤΩΝ ΣΥΣΤΗΜΑΤΩΝ ΤΗΣ ΨΗΦΙΑΚΗΣ ΑΞΙΟΛΟΓΗΣΗΣ

Η αναζήτηση των απαρχών της αξιοποίησης τεχνολογικών μέσων για την υποστήριξη της εκπαιδευτικής αξιολόγησης οδηγεί στα μέσα της δεκαετίας του 1920, όταν ο Αμερικανός καθηγητής ψυχολογίας Sidney L. Pressey εφηύρε και παρουσίασε μια (καινοτόμο για την τότε εποχή, πρωτόγονη για τα σημερινά δεδομένα) μηχανική διάταξη, στην οποία έδωσε το αρκετά εντυπωσιακό, αλλά μάλλον άστοχο, όνομα «αυτόματη μηχανή ελέγχου ευφυΐας» (automatic intelligence testing machine). Η μηχανή, που είχε μέγεθος και μηχανικά μέρη γραφομηχανής, εμφάνιζε διαδοχικά τριάντα ερωτήσεις πολλαπλής επιλογής και βαθμολογούσε τις αντίστοιχες απαντήσεις. Οι ερωτήσεις ήταν τυπωμένες σε χαρτί και οι απαντήσεις εισάγονταν με τη βοήθεια τεσσάρων πλήκτρων.



*Εικόνα 1. Η μηχανή του Sidney L. Pressey (Petrina, 2004)*

Η μηχανή δεν γνώρισε εμπορική επιτυχία, αλλά ο Pressey, που διακατεχόταν από πάθος για την εισαγωγή αυτοματισμών στην εκπαίδευση, προέβλεψε (με αρκετή δόση υπερβολής), ότι σε είκοσι χρόνια, ειδικά μηχανικά μέσα θα επιφέρουν αλλαγές στην εκπαίδευση ανάλογες με τη βιομηχανική επανάσταση (Petrina, 2004). Τις επόμενες δεκαετίες, κατασκευάστηκαν μηχανές ικανές να ελέγχουν την ορθότητα των απαντήσεων σε ερωτήσεις πολλαπλής επιλογής, σαρώνοντας ειδικές έντυπες φόρμες, όπου οι εξεταζόμενοι είχαν σημειώσει τις απαντήσεις τους (Watters, 2015). Η επιτυχία τέτοιων μηχανών υπήρξε μεγάλη (είναι αξιοσημείωτο ότι εξακολουθούν να είναι σε χρήση μέχρι και σήμερα), κυρίως διότι διευκόλυναν τη γρήγορη, με χαμηλό κόστος και με μεγάλη ακρίβεια βαθμολόγηση σε εξετάσεις μεγάλης κλίμακας. Με την

πάροδο των ετών, οι μηχανές εξελίχθηκαν και οι σύγχρονες εκδοχές τους είναι ψηφιακές. Ωστόσο, η μόνη πρόοδος που προσφέρει η ψηφιακή μετασκευή τους είναι η ψηφιοποίηση μιας παλιάς μηχανικής διαδικασίας και η απλή μετατροπή διαγωνισμάτων από έντυπη σε ψηφιακή μορφή. Η επιτυχία και η πολύχρονη κυριαρχία τέτοιων μηχανών, όπου οι ερωτήσεις πολλαπλής επιλογής ήταν η μόνη προσφερόμενη μορφή ερωτήσεων, είχε ως επακόλουθο, όπως αναφέρουν οι Brown, Race & Bull (1999), να είναι ταυτισμένη για πολύ καιρό η έννοια της ψηφιακής αξιολόγησης με τα ψηφιακά διαγωνίσματα ερωτήσεων πολλαπλής επιλογής. Τις δεκαετίες του 1980 και του 1990 κατασκευάστηκαν περισσότερο σύνθετα, ψηφιακά συστήματα αξιολόγησης, ικανά να διαχειριστούν συνθετότερες μορφές ερωτήσεων (συνεπώς ικανότερα να αξιολογούν πιο σύνθετες νοητικές ικανότητες), να ενσωματώνουν πολυμεσικό υλικό και να προσαρμόζουν το επίπεδο δυσκολίας των ερωτήσεων ανάλογα με τις απαντήσεις στις ερωτήσεις που προηγήθηκαν (Jordan, 2013).

## 2.1. Εξελικτικές φάσεις

Η εξέλιξη των συστημάτων ψηφιακής αξιολόγησης περιγράφεται από τους Bunderson, Inouye & Olsen (1989) με ένα σχήμα τεσσάρων εξελικτικών γενεών (ή σταδίων), όπου κάθε γενιά διαθέτει ανώτερη ποιότητα και ισχύ από τις προηγούμενες. Η Πρώτη Γενιά περιλαμβάνει συστήματα που απλώς διαχειρίζονται παραδοσιακά διαγωνίσματα μέσω υπολογιστών. Η Δεύτερη Γενιά περιλαμβάνει συστήματα που προσαρμόζουν δυναμικά το επίπεδο δυσκολίας της αξιολόγησης, βάση της προηγούμενης απόδοσης των αξιολογούμενων. Η Τρίτη Γενιά περιλαμβάνει συστήματα που χρησιμοποιούν μηχανισμούς μέτρησης, ενσωματωμένους στα προγράμματα σπουδών με τη βοήθεια των οποίων εκτιμούν συνεχώς και με διακριτικότητα τις επιδόσεις των σπουδαστών. Η Τέταρτη Γενιά καταγράφει και αναλύει τα μαθησιακά χαρακτηριστικά των σπουδαστών και συμβουλεύει τους σπουδαστές και τους διδάσκοντες με τη βοήθεια τεχνικών που χρησιμοποιούνται στα λεγόμενα «έμπειρα» συστήματα, όπως βάσεις γνώσεων (knowledge bases) και μηχανές παραγωγής λογικών συμπερασμάτων (inference engines). Ο Martin (2008) σχολιάζοντας την παραπάνω ταξινόμηση αναφέρει ότι η υλοποίηση και χρήση συστημάτων που ανήκουν στην Τρίτη και Τέταρτη Γενιά εξαρτώνται, αφενός από τις αναγκαίες τεχνολογικές δυνατότητες και αφετέρου από την πρόοδο στους τομείς της γνωστικής

επιστήμης και της ψυχομετρικής, οι οποίες θα προσφέρουν τα εργαλεία για την ανάλυση γνωστικών διαδικασιών.

Από την εξελεγκτική πορεία που προτείνουν οι Bunderson et al. (1989) απουσιάζουν (προφανώς λόγω της χρονικής περιόδου που εκπονήθηκε η σχετική έρευνα) αναφορές στο διαδίκτυο, το οποίο επέφερε αλλαγές σε όλους τους τομείς της εκπαιδευτικής διαδικασίας και συνεπώς στην αξιολόγηση. Μια από τις βασικές αλλαγές αφορά τη δυνατότητα συμμετοχής σε διαδικασίες αξιολόγησης, από οπουδήποτε και οποτεδήποτε, με μόνη προϋπόθεση τη διαθεσιμότητα διαδικτυακής σύνδεσης. Η φυσική παρουσία των σπουδαστών σε εξεταστικά κέντρα δεν είναι πλέον απαραίτητη, δεδομένου ότι η συμμετοχή τους υλοποιείται διαδικτυακά. Η βαθμολόγηση και η ανατροφοδότηση, μπορούν επίσης να υλοποιηθούν διαδικτυακά και άμεσα. Τα αποτελέσματα των αξιολογήσεων αποθηκεύονται σε ψηφιακά μέσα, γεγονός που διευκολύνει την περαιτέρω επεξεργασία και ανάλυσή τους, ενώ είναι διαδικτυακά εφικτή και η βαθμολόγηση, στις περιπτώσεις όπου απαιτείται ανθρώπινη συμμετοχή. Παράλληλη εξέλιξη αποτελεί η ανάπτυξη και ευρεία χρήση των Συστημάτων Διαχείρισης Μάθησης (Learning Management Systems), τα οποία προσφέρουν το μέσο με το οποίο είναι δυνατόν να υλοποιηθούν αρκετές διαδικασίες ψηφιακής αξιολόγησης. Σημαντικές αλλαγές επέφερε επίσης η ανάπτυξη εργαλείων συγγραφής εκπαιδευτικού περιεχομένου (authoring tools), καθώς διευκόλυνε την κατασκευή ερωτήσεων και διαδικτυακών (online) διαγωνισμάτων από διδάσκοντες (Timmis, Broadfoot, Sutherland, & Oldfield, 2016). Σε συνέχεια των αλλαγών που επέφερε το διαδίκτυο, οι σύγχρονες δικτυακές πλατφόρμες και εφαρμογές και ειδικότερα η ανάπτυξη Web 2.0 τεχνολογιών επιτρέπουν διαδικτυακές συμμετοχικές και αλληλεπιδραστικές εμπειρίες, οι οποίες πολλαπλασιάζουν τις μορφές που μπορεί να λάβει, τις ικανότητες που μπορεί να αξιολογήσει και τους σκοπούς που μπορεί να εξυπηρετήσει η ψηφιακή αξιολόγηση (Dohn, 2009. Elliott, 2008. Jordan, 2013). Οι Whitelock & Brasher (2006) αναφέρουν ότι οι μελλοντικές μεσοπρόθεσμες τάσεις της ψηφιακής αξιολόγησης θα αφορούν «κατ' απαίτηση» (on-demand) δοκιμασίες, αυτονομία των σπουδαστών μέσω της διαμορφωτικής αξιολόγησης και περαιτέρω ανάπτυξη της διαγνωστικής αξιολόγησης. Υπογραμμίζουν ωστόσο την ανάγκη για βελτιωμένη ανατροφοδότηση που θα στηρίζεται σε παιδαγωγικές αρχές.

Ο Bennett (1998, 2010, 2015), ακολούθησε αντίστοιχη προσέγγιση με τους Bunderson et al. (1989) και περιέγραψε την εξελικτική πορεία με τη βοήθεια τριών γενιών συστημάτων κατάλληλων για αξιολογήσεις μεγάλης κλίμακας.

Η **Πρώτη Γενιά** (First Generation) συστημάτων έχει κυρίως μορφή διαγωνισμάτων, που αξιολογούν απλές γνωστικές ικανότητες χρησιμοποιώντας σχεδιασμό και μορφές αξιολογικών δοκιμασιών που μοιάζουν με τις παραδοσιακές έντυπες δοκιμασίες. Τα θεμελιώδη χαρακτηριστικά των ψηφιακών αξιολογικών δοκιμασιών παραμένουν σχεδόν τα ίδια με αυτά των παραδοσιακών αξιολογήσεων, οπότε μετριάζεται ο βαθμός πολυπλοκότητας της μετάβασης. Οι αλλαγές αφορούν κυρίως τις τεχνολογικές υποδομές που υποστηρίζουν τη διαδικασία και διανέμουν το υλικό της αξιολόγησης. Αξιοσημείωτη εξαίρεση αποτελεί η δυνατότητα για προσαρμογή της δυσκολίας των δοκιμασιών ανάλογα με τις προηγούμενες απαντήσεις του αξιολογούμενου. Το δυναμικά προσαρμοζόμενο επίπεδο δυσκολίας αυξάνει την ακρίβεια της αξιολόγησης και κατανέμει ορθότερα επί της αξιολογικής κλίμακας τις επιδόσεις των αξιολογούμενων. Οι αξιολογικές διαδικασίες είναι σχεδιασμένες ώστε πρωτίστως να εξυπηρετούν τις ανάγκες των εκπαιδευτικών οργανισμών. Διεξάγονται ως διακριτά και μεμονωμένα εκπαιδευτικά συμβάντα. Παρά τις ομοιότητες με τον παραδοσιακό τρόπο αξιολόγησης, το κόστος μετάβασης σε σύστημα αξιολόγησης της πρώτης γενιάς δεν είναι αμελητέο. Απαιτούνται επενδύσεις σε υπολογιστικό εξοπλισμό, σε δικτυακές υποδομές, σε άδειες χρήσης λογισμικού. Απαιτείται εκπαίδευση τεχνικών στελεχών, διδασκόντων και σπουδαστών. Απαιτείται προσπάθεια για τον σχεδιασμό, τη δημιουργία, την αποθήκευση και τη διανομή των αντικειμένων της αξιολόγησης (π.χ. ερωτήσεων).

Η **Δεύτερη γενιά** συστημάτων είναι αναβαθμισμένη ποιοτικά και αυξάνει την αποδοτικότητα των διαδικασιών. Υποστηρίζει πολύ περισσότερες μορφές δοκιμασιών σε σχέση με τις παραδοσιακές αξιολογήσεις. Οι διαφορές αφορούν στη μορφή των δοκιμασιών (π.χ. ενσωματώνουν πολυμεσικό υλικό, επιτρέπουν ερωτήσεις ανοικτού τύπου, υποβολή αναφορών ή εκθέσεων) και τις γνωστικές ικανότητες που είναι δυνατόν να αξιολογηθούν. Η αύξηση της αποδοτικότητας επιτυγχάνεται μέσω δυνατοτήτων αυτόματης βαθμολόγησης. Αξιοποιούνται διαδικτυακές υποδομές για τη διανομή των διαγωνισμάτων, για την καταγραφή των απαντήσεων και για την επικοινωνία με τους αξιολογούμενους. Συχνά, η χρήση καινοτομιών δεν

συνοδεύεται από επαρκή τεκμηρίωση, ως προς την καταλληλότητα και τη σκοπιμότητα χρήσης τους.

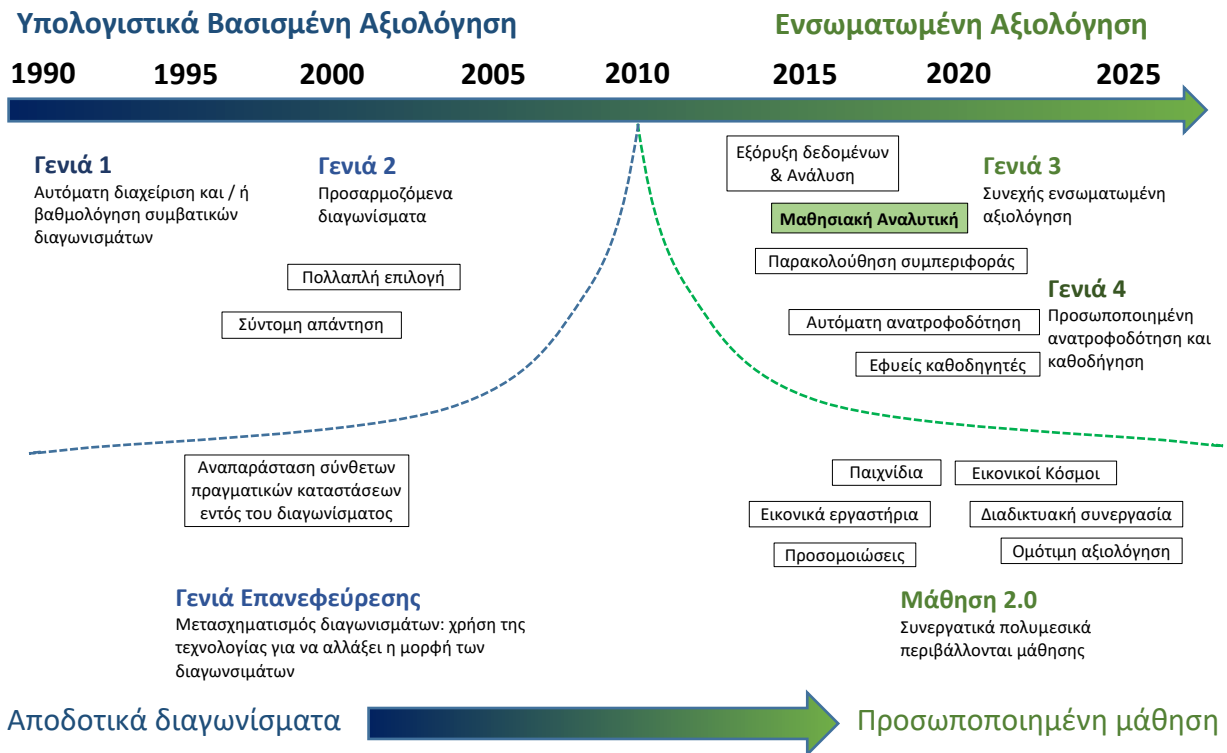
Η **Τρίτη γενιά** συστημάτων, την οποία ο Bennett (1998) αποκαλεί γενιά R (R -> Reinvention = Επανεφεύρεση), παρουσιάζει ουσιαστικές αλλαγές σε πολλά επίπεδα ταυτοχρόνως. Οι τεχνολογικές καινοτομίες αξιοποιούνται μόνο αν προσφέρουν ουσιαστικό όφελος, ή αν αναβαθμίζουν τη διαδικασία και όχι διότι απλώς είναι διαθέσιμες ή διότι είναι εντυπωσιακές. Οι διαδικασίες είναι σχεδιασμένες ώστε εκτός από τους εκπαιδευτικούς οργανισμούς να εξυπηρετούν και τις ανάγκες μάθησης των αξιολογούμενων (π.χ. μέσω διαδικασιών και εργαλείων που υποστηρίζουν διαγνωστικές και διαμορφωτικές μορφές αξιολόγησης που παρέχουν ανατροφοδότηση σε διδάσκοντες και σπουδαστές). Ο σχεδιασμός ακολουθεί τεκμηριωμένες παιδαγωγικές και γνωστικές θεωρητικές αρχές. Τα συστήματα αξιοποιούν τεχνολογίες προσομοίωσης και εικονικής πραγματικότητας, που επιτρέπουν περισσότερη φυσική διάδραση συστήματος και αξιολογούμενου. Οι δοκιμασίες διεξάγονται από απόσταση και είναι ικανές να αξιολογήσουν ανώτερες γνωστικές ικανότητες. Ενσωματώνονται αρμονικά στη εκπαιδευτική διαδικασία και επιτρέπουν τη συνεχή μέτρηση της απόδοσης των σπουδαστών. Ωστόσο, όπως παρατηρεί ο Bennett (2015), η εξέλιξη προς τα συστήματα τρίτης γενιάς είναι αντιμέτωπη με θεωρητικές και σχεδιαστικές προκλήσεις. Πρώτον, είναι αναγκαία η δημιουργία θεωρητικά τεκμηριωμένων μοντέλων για τα κριτήρια απόδοσης των συστημάτων. Δεύτερον, είναι αναγκαίο να αποτιμηθεί ο βαθμός της θετικής επίδρασης των συστημάτων στη μαθησιακή διαδικασία. Τρίτον, πρέπει να σχεδιασθούν και να εφαρμοσθούν κατάλληλοι μηχανισμοί για τη συχνή δειγματοληψία δεδομένων, που αφορούν την απόδοση των αξιολογούμενων λαμβάνοντας υπόψη ότι πρέπει να ικανοποιείται η αξιοπιστία των μετρήσεων, το παιδαγωγικό όφελος της διαδικασίας και η ιδιωτικότητα των αξιολογούμενων. Τέταρτον, η αυτόματη βαθμολόγηση πρέπει να λαμβάνει υπόψη τις ιδιαιτερότητες του γνωστικού πεδίου και να είναι όσο το δυνατόν ακριβής. Και τέλος, τα νέα συστήματα αξιολόγησης πρέπει να είναι έγκυρα, δίκαια έναντι όλων των αξιολογούμενων (ανεξαρτήτως προσωπικών ιδιαιτεροτήτων) και επωφελή για όλους.

Πίνακας 1. Τρεις Γενιές Συστημάτων Αξιολόγησης (Bennett, 1998, σ. 16)

ΤΡΕΙΣ ΓΕΝΙΕΣ ΣΥΣΤΗΜΑΤΩΝ ΓΙΑ ΑΞΙΟΛΟΓΗΣΗ ΜΕΓΑΛΗΣ ΚΛΙΜΑΚΑΣ	
Γενιά Συστημάτων	Βασικά Χαρακτηριστικά
<p>Πρώτη Γενιά Υπολογιστικά Βασισμένων Διαγωνισμάτων (Υλοποίηση Υποδομών)</p>	<ol style="list-style-type: none"> <li>1. Πρωτίστως εξυπηρετούν ανάγκες εκπαιδευτικών οργανισμών.</li> <li>2. Αξιολογούν απλές ικανότητες και χρησιμοποιούν μορφές διαγωνισμάτων που μοιάζουν με τις παραδοσιακές, με εξαίρεση τη δυνατότητα για δυναμικά προσαρμοζόμενο επίπεδο δυσκολίας.</li> <li>3. Διεξάγονται σε συγκεκριμένα εξεταστικά κέντρα ως δοκιμασίες μιας επιτρεπόμενης προσπάθειας.</li> <li>4. Δεν αξιοποιούν στο έπακρο τεχνολογικές δυνατότητες.</li> </ol>
<p>Δεύτερη Γενιά Διαδικτυακών Διαγωνισμάτων (Ποιοτική αναβάθμιση)</p>	<ol style="list-style-type: none"> <li>1. Πρωτίστως εξυπηρετούν ανάγκες εκπαιδευτικών οργανισμών.</li> <li>2. Χρησιμοποιούν νέες μορφές διαγωνισμάτων, επιτρέπουν αυτόματη βαθμολόγηση, αξιοποιούν δικτυακές υποδομές, αξιολογούν ευρύτερο σύνολο ικανοτήτων.</li> <li>3. Διεξάγονται σε συγκεκριμένα εξεταστικά κέντρα ως δοκιμασίες μιας επιτρεπόμενης προσπάθειας.</li> <li>4. Επιτρέπουν την επικοινωνία με μεταξύ του αξιολογούμενου και του φορέα αξιολόγησης.</li> </ol>
<p>Τρίτη Γενιά (ή Γενιά R: Reinvention = Επανεφεύρεση)</p>	<ol style="list-style-type: none"> <li>1. Εξυπηρετούν ανάγκες εκπαιδευτικών οργανισμών και αξιολογούμενων.</li> <li>2. Ενσωματώνονται στην εκπαιδευτική διαδικασία και αξιολογούν συνεχώς τους επιδόσεις των σπουδαστών. Είναι σχεδιασμένα βάση γνωστικών αρχών.</li> <li>3. Χρησιμοποιούν τεχνολογίες προσομοίωσης και βελτιώνουν τη διάδραση μεταξύ αξιολογούμενου και συστήματος.</li> <li>4. Είναι δυνατή, η από απόσταση διαχείρισή τους.</li> <li>5. Αξιολογούν νέες ικανότητες.</li> </ol>

Την οριοθέτηση της εξελεγκτικής πορείας των συστημάτων αξιολόγησης σε διαδοχικές γενιές που υπέδειξαν οι Bunderson et al. (1989), ο Martin (2008) και ο Bennett (1998, 2010, 2015), υιοθετεί και η Redecker (2013). Αναπαριστά γραφικά (Σχήμα 1) την εξέλιξη με χρονικό βέλος, το οποίο οδηγεί από τη χρονική περίοδο της «Υπολογιστικά Βασισμένης Αξιολόγησης» στη χρονική περίοδο της «**Ενσωματωμένης Αξιολόγησης**», από τη χρήση της τεχνολογίας για απλό μετασχηματισμό παραδοσιακών διαδικασιών στη χρήση συνεργατικών πολυμεσικών περιβαλλόντων μάθησης και από την επιδίωξη για αποδοτικές πρακτικές αξιολόγησης στην επιδίωξη της προσωποποιημένης μάθησης. Ορίζει ως συμβατικό χρονικό ορόσημο το έτος 2010, το οποίο διαχωρίζει τις πρώτες γενιές συστημάτων αξιολόγησης από τις επόμενες. Οι καινοτομίες που εισήγαγαν οι πρώτες γενιές συστημάτων αφορούν την αυτόματη διαχείριση και βαθμολόγηση συμβατικών διαγωνισμάτων και τον δυναμικά προσαρμόσιμο βαθμό δυσκολίας. Σε αυτά τα συστήματα, οι κυρίαρχοι τύποι των δοκιμασιών είναι οι ερωτήσεις πολλαπλών επιλογών ή σύντομων απαντήσεων. Οι καινοτομίες που εισάγουν ή θα εισάγουν οι νέες γενιές συστημάτων σχετίζονται με συνεχείς και ενσωματωμένες στην εκπαιδευτική διαδικασία αξιολογικές δοκιμασίες, με προσωποποιημένη ανατροφοδότηση και καθοδήγηση. Οι τύποι των δοκιμασιών των νέων συστημάτων περιλαμβάνουν προσομοιώσεις, εικονικά εργαστήρια, παιχνίδια, εικονικούς κόσμους, διαδικτυακές συνεργατικές δραστηριότητες και ομότιμη αξιολόγηση. Τα τεχνολογικά εργαλεία που αξιοποιούν τα συστήματα της νέας γενιάς περιλαμβάνουν εξόρυξη δεδομένων (data mining), Μαθησιακή Αναλυτική (Learning Analytics), παρακολούθηση μαθησιακής συμπεριφοράς, αυτοματοποιημένη ανατροφοδότηση, ευφυείς καθοδηγητές. Ειδικότερα, σε σχέση με τη Μαθησιακή Αναλυτική η Redecker (2013, σ. 7) αναφέρει ότι πρόκειται για μια από τις «πολλά υποσχόμενες αναδυόμενες τεχνολογικές τάσεις για την συνολική αξιολόγηση σύνθετων ικανοτήτων» και ότι μπορεί να υποστηρίξει την ερμηνεία δεδομένων που συλλέγονται σε ψηφιακά περιβάλλοντα μάθησης, ώστε να αξιολογηθεί η πρόοδος των σπουδαστών και να υλοποιηθούν στοχευμένες διδακτικές παρεμβάσεις σε σπουδαστές που έχουν σχετική ανάγκη. Η Μαθησιακή Αναλυτική, συμπληρώνει η Redecker (2013, σ. 7), παρέχει τη δυνατότητα για την ενσωμάτωση της αξιολόγησης σε περιβάλλοντα εμπύθισης, σε ψηφιακά παιχνίδια με πολλούς συμμετέχοντες και σε προσομοιώσεις.





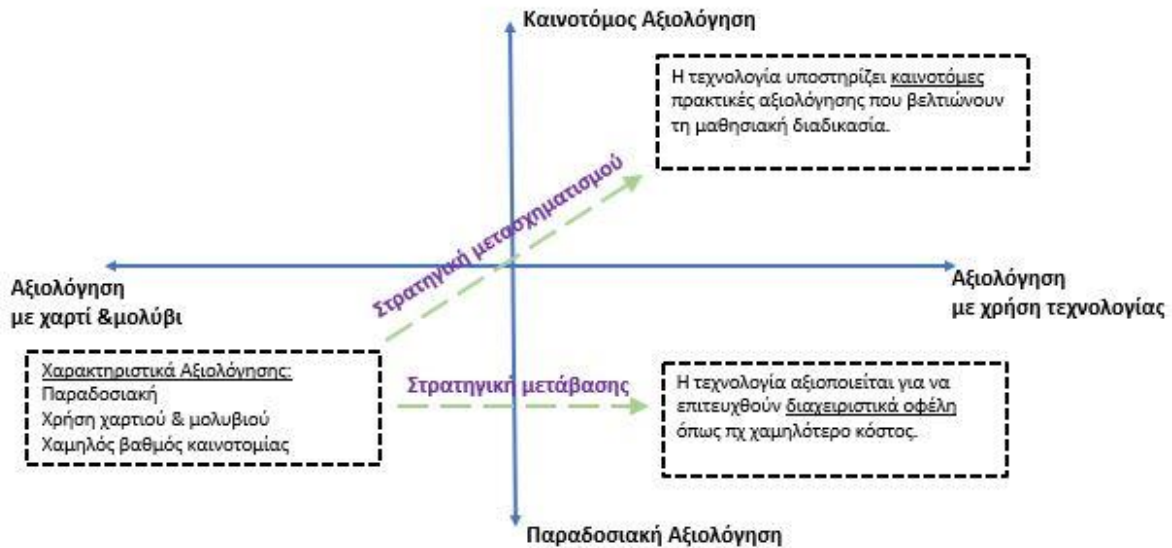
Σχήμα 1. Εξελικτική πορεία συστημάτων αξιολόγησης (Redecker, 2013, σ. 11)

Η Redecker (2013) αναφέρει ότι τα συστήματα των πρώτων γενιών έχουν ήδη εδραιωθεί και ότι το ζητούμενο πλέον είναι η μετάβαση από την εποχή της υπολογιστικά βασισμένης αξιολόγησης - όπου οι υπολογιστές απλώς βελτιώνουν την απόδοση των αξιολογικών διαδικασιών - στην εποχή της τεχνολογικά βελτιούμενης μάθησης - όπου τα συστήματα **ενσωματώνουν** την «ολιστική» και προσωποποιημένη αξιολόγηση εντός της μάθησης. Ο σχεδιασμός των συστημάτων των νέων γενιών διέπεται από το όραμα ότι οι διακριτές διαδικασίες αξιολόγησης θα είναι περιττές. Διότι οι σπουδαστές θα αξιολογούνται **εμμέσως** και **συνεχώς**, μέσω συστημάτων και διαδικασιών **ενσωματωμένων** στην εκπαιδευτική διαδικασία. Δηλαδή, η αξιολόγηση δεν θα αποτελεί διακριτό τμήμα της εκπαιδευτικής διαδικασίας, αλλά θα **διαχέεται** μέσα σε αυτήν. Τα συστήματα που θα υποστηρίζουν τη μάθηση, (π.χ. τα Συστήματα Διαχείρισης Μάθησης) ταυτοχρόνως θα αξιολογούν, θα ανατροφοδοτούν και γενικά θα υποστηρίζουν τους σπουδαστές για την επίτευξη των μαθησιακών στόχων. Παράλληλα, τα συστήματα των νέων γενιών θα πρέπει να είναι προσαρμοσμένα στους μαθησιακούς στόχους των σύγχρονων

εκπαιδευτικών συστημάτων. Συγκεκριμένα θα στοχεύουν στην καλλιέργεια και στην αξιολόγηση δεξιοτήτων που συμβατικά ονομάζονται «κεντρικές δεξιότητες για τον 21<sup>ο</sup> αιώνα» (21<sup>st</sup> century key competencies) και περιλαμβάνουν την κριτική σκέψη, τη δημιουργικότητα, την πρωτοβουλία, την επίλυση προβλημάτων, την αξιολόγηση κινδύνων, τη λήψη αποφάσεων και την επικοινωνιακή διαχείριση των συναισθημάτων. Η αξιολόγηση τέτοιων ικανοτήτων δεν είναι εφικτή, όσο τα συστήματα αξιολόγησης χρησιμοποιούν τις ΤΠΕ απλώς για να υλοποιήσουν με μεγαλύτερη αποδοτικότητα αξιολογικές δοκιμασίες παραδοσιακού τύπου. Η αξιολόγηση τέτοιων ικανοτήτων είναι εφικτή, αν οι ΤΠΕ χρησιμοποιηθούν για να μεταβάλουν τον τρόπο και τις διαδικασίες της αξιολόγησης.

## 2.2. Στρατηγικές Μετάβασης

Οι Binkley et al. (2012) υποστηρίζουν ότι η εξελικτική πορεία των συστημάτων της ψηφιακής αξιολόγησης ακολουθεί δύο τάσεις. Η πρώτη τάση που συμβατικά αποκαλείται «στρατηγική μετάβασης» χρησιμοποιεί τις ΤΠΕ με μόνο στόχο την αποτελεσματικότερη και αποδοτικότερη υλοποίηση αξιολογήσεων παραδοσιακού τύπου. Η δεύτερη τάση που συμβατικά αποκαλείται «στρατηγική μετασχηματισμού» χρησιμοποιεί τις ΤΠΕ για να αναπροσαρμόσει τον τρόπο με τον οποίο αξιολογούνται οι γνωστικές ικανότητες και να αξιοποιήσει καινοτόμες μεθόδους και πρακτικές που επιφέρουν μαθησιακά οφέλη. Οι δύο τάσεις αποτυπώνονται γραφικά από τους Binkley et al. (2012) σε διάγραμμα δύο αξόνων, όπου ο κάθετος άξονας αντιστοιχεί στον βαθμό καινοτομίας και ο οριζόντιος άξονας στον βαθμό χρήσης της τεχνολογίας. Το κάτω αριστερά τεταρτημόριο περιλαμβάνει τις αξιολογήσεις με περιορισμένη χρήση τεχνολογίας και χαμηλό βαθμό καινοτομίας. Η «στρατηγική μετάβασης» οδηγεί στο κάτω δεξιό τεταρτημόριο, εκεί όπου η τεχνολογία δεν παράγει καινοτόμες πρακτικές και αξιοποιείται μόνο για την επίτευξη διαχειριστικών οφελών. Η «στρατηγική μετασχηματισμού» οδηγεί στο πάνω δεξιό τεταρτημόριο, εκεί όπου η τεχνολογία αξιοποιείται για να υποστηρίξει καινοτομίες που βελτιώνουν τη μαθησιακή διαδικασία.



**Σχήμα 2. Διαστάσεις καινοτομίας στην ψηφιακή αξιολόγηση (Binkley et al., 2012)**

Οι «στρατηγικές μετάβασης και μετασχηματισμού» των Binkley et al. (2012) απηχούν έννοιες που περιλαμβάνει το εξελικτικό μοντέλο SAMR (**S**ubstitution, **A**ugmentation, **M**odification, **R**edefinition) του Puentedura (2010), το οποίο διαχωρίζει την έννοια της «τεχνολογικής βελτίωσης» από την έννοια του «τεχνολογικού μετασχηματισμού». Το μοντέλο SAMR είναι ευρέως αποδεκτό (JISC, 2015) ως πλαίσιο ταξινόμησης, όχι μόνο για τη ψηφιακή αξιολόγηση, αλλά γενικά για την εκπαιδευτική τεχνολογία. Περιγράφει τα παρακάτω τέσσερα επίπεδα αξιοποίησης της τεχνολογίας:

- **Υποκατάσταση (Substitution)**: οι ψηφιακές τεχνολογίες υποκαθιστούν παλαιότερα εργαλεία χωρίς να επιφέρουν διαδικαστικές αλλαγές.
- **Επαύξηση (Augmentation)**: οι ψηφιακές τεχνολογίες αντικαθιστούν παλαιότερα εργαλεία και επιφέρουν διαδικαστικές βελτιώσεις.
- **Τροποποίηση (Modification)**: οι ψηφιακές τεχνολογίες χρησιμοποιούνται για τον ανασχεδιασμό των διαδικασιών.
- **Επαναπροσδιορισμός (Redefinition)**: δημιουργούνται νέες διαδικασίες που πριν δεν ήταν υλοποιήσιμες.

Στα πρώτα δύο επίπεδα (Υποκατάσταση, Επαύξηση), η ψηφιακή τεχνολογία αξιοποιείται με απλό τρόπο, βελτιώνοντας τις διαδικασίες, χωρίς να μεταβάλλει βασικές παιδαγωγικές πρακτικές, ενώ στα επόμενα δύο επίπεδα (Τροποποίηση, Επαναπροσδιορισμός), η ψηφιακή τεχνολογία μετασχηματίζει τις παιδαγωγικές πρακτικές (Φεσάκης & Κωνσταντοπούλου, 2022, σ. 34).

Οι Morales-Martinez, Lopez-Ramirez, Castro-Campos, Villarreal-Trevino & Gonzales-Trujillo (2017) υποστηρίζουν α) ότι τα υφιστάμενα συστήματα της ψηφιακής αξιολόγησης βασίζονται κυρίως σε ψηφιακές υλοποιήσεις παλαιών μεθόδων, που αποσκοπούν απλώς στην αποτίμηση των μαθησιακών αποτελεσμάτων, β) ότι τα περισσότερα συστήματα διαθέτουν χαρακτηριστικά των πρώτων δύο επιπέδων, αλλά σπανίως χαρακτηριστικά που αντιστοιχούν σε τεχνολογικούς μετασχηματισμούς και γ) ότι οι ισχυρισμοί σύμφωνα με τους οποίους υφίστανται συστήματα ικανά να κρίνουν ότι οι εκπαιδευόμενοι έχουν κατακτήσει ανώτερες δεξιότητες και γνώσεις δεν τεκμηριώνονται με επαρκή δεδομένα. Οι Sweeney et al. (2017), στα πλαίσια της ευρείας βιβλιογραφικής επισκόπησης που πραγματοποίησαν στο πεδίο της ψηφιακής αξιολόγησης, επιχείρησαν να ταξινομήσουν τις σχετικές ερευνητικές προσπάθειες εντός των τεσσάρων επιπέδων του μοντέλου SAMR. Εντόπισαν 1713 άρθρα, από τα οποία 139 αφορούσαν σε αξιολόγηση με τη χρήση τεχνολογίας. Δέκα έξι άρθρα (12%) αναφέρουν υποκατάσταση παραδοσιακών μεθόδων από την τεχνολογία, χωρίς να καταγράφεται κάποιο όφελος. Πενήντα επτά άρθρα (41%) αναφέρουν ότι η χρήση της τεχνολογίας επέφερε διαδικαστικές βελτιώσεις (επαύξηση). Σαράντα ένα άρθρα (30%) αναφέρουν ότι η χρήση της τεχνολογίας οδήγησε σε σημαντική τροποποίηση των διαδικασιών, ενώ 23 άρθρα (17%) περιγράφουν κάποιου είδους επαναπροσδιορισμό της αξιολόγησης και μετασχηματισμό της μαθησιακής διαδικασίας. Συνολικά, η πλειονότητα των άρθρων (ποσοστό μεγαλύτερο του 70%) αντιστοιχεί στα ενδιάμεσα επίπεδα του μοντέλου SAMR, δηλαδή στο δεύτερο επίπεδο (Επαύξηση) και στο τρίτο επίπεδο (Τροποποίηση), όπου η τεχνολογία παρέχει δυνατότητες για τη βελτίωση της μαθησιακής διαδικασίας. Οι Sweeney et al. (2017) αναφέρουν ότι για να αξιοποιηθούν όσο το δυνατόν πιο αποδοτικά αυτές οι δυνατότητες είναι απαραίτητη η επίγνωση των ιδιαίτερων χαρακτηριστικών κάθε τεχνολογίας και των συγκεκριμένων τρόπων με τους οποίους μπορεί να υποστηρίξει την αξιολόγηση που στοχεύει στη μάθηση. Συγκεκριμένα, οι διδάσκοντες είναι

ανάγκη να επιλέξουν τις κατάλληλες τεχνολογίες και διαδικασίες τις οποίες θα αξιοποιήσουν καθ' όλη τη διάρκεια της εκπαιδευτικής διαδικασίας, ώστε να δημιουργήσουν αυθεντικές δραστηριότητες στις οποίες θα συμμετέχουν ενεργά οι σπουδαστές. Ακολούθως, να αξιολογήσουν την απόδοση των σπουδαστών, να ερμηνεύσουν τα αποτελέσματα και να αξιοποιήσουν τις σχετικές πληροφορίες για να βελτιώσουν τη μαθησιακή διαδικασία. Τα μισά περίπου άρθρα (47%) αντιστοιχούν στα δύο ανώτερα επίπεδα του μοντέλου SAMR, δηλαδή στο τρίτο επίπεδο (Μετατροπή) και στο τέταρτο επίπεδο (Επαναπροσδιορισμός), δηλαδή στα επίπεδα τα οποία θεωρούνται ότι επιφέρουν μετασχηματισμούς των διαδικασιών. Όσα άρθρα, περιγράφουν μετασχηματισμούς διαδικασιών, αναφέρονται σε συστήματα τα οποία αξιοποιούν διαμορφωτικές μορφές ομότιμης αξιολόγησης και αυτοαξιολόγησης μέσω πρακτικών συνεργατικής μάθησης και διαλόγων που περιλαμβάνουν πολλές μορφές ανατροφοδότησης. Αξιοποιούν μέσα κοινωνικής δικτύωσης και Συστήματα Διαχείρισης Μάθησης (ΣΔΜ), τα οποία ενισχύουν τις αλληλεπιδράσεις μεταξύ των συμμετεχόντων, δίνουν τη δυνατότητα στους σπουδαστές να συμμετάσχουν ενεργά στη μαθησιακή διαδικασία και να ενεργούν με αυτονομία, αυτοπεποίθηση και αυτογνωσία ως προς τα επιτεύγματά τους.

Σχετική με το μοντέλο SAMR είναι η κατάταξη των συστημάτων ψηφιακής αξιολόγησης που προτείνουν οι Farrell και Rushby (2016, σ. 118).

**Πίνακας 2. Η κατά Farrell και Rushby (2016) ταξινόμηση μορφών ψηφιακής αξιολόγησης σύμφωνα με τον μοντέλο SAMR**

	Υποκατάσταση	Επαύξηση	Τροποποίηση	Επαναπροσδιορισμός
Ψηφιακή διαχείριση αξιολόγησης	✓			
Υπολογιστικά υποστηριζόμενη βαθμολόγηση κειμένου		✓		
Ερωτήσεις πολλαπλής επιλογής			✓	
Στοχαστικά ιστολόγια			✓	

Ψηφιακοί φάκελοι επιτευγμάτων			✓	
Σύστημα απόκρισης κοινού				✓
Μοντελοποίηση σπουδαστών				✓
<b>Μαθησιακή Αναλυτική</b>				✓

Συγκεκριμένα, οι Farrell & Rushby (2016) αναφέρουν ότι η ψηφιακή διαχείριση της αξιολόγησης κατατάσσεται στο επίπεδο της «Επαύξησης» και δεν θα πρέπει να εκλαμβάνεται μόνο ως απλή αυτοματοποίηση. Διότι οι πόροι που εξοικονομούνται είναι δυνατόν να επενδυθούν στη βελτίωση της ανατροφοδότησης και γενικά στην υποστήριξη των σπουδαστών. Στο επίπεδο της «Επαύξησης» κατατάσσεται και η χρήση της τεχνολογίας για τη βαθμολόγηση ανοικτών απαντήσεων, διότι αυξάνει τη διαφάνεια και την αντικειμενικότητα. Η χρήση προηγμένων μορφών ερωτήσεων πολλαπλής επιλογής, ιστολογίων και ψηφιακών φακέλων επιτευγμάτων κατατάσσεται στο επίπεδο της «Τροποποίησης». Στο επίπεδο του «Επαναπροσδιορισμού» κατατάσσονται τα συστήματα απόκρισης κοινού, διότι επιτρέπουν την επανασχεδίαση εκπαιδευτικών μεθόδων. Στο επίπεδο του «Επαναπροσδιορισμού» κατατάσσονται επίσης η «μοντελοποίηση σπουδαστών», δηλαδή η δομημένη αναπαράσταση των γνώσεων, των παρανοήσεων και των δυσχερειών του σπουδαστή (Bull S., 2004) και η Μαθησιακή Αναλυτική για την οποία αναφέρει, ότι προσφέρει δυνατότητες για διαδικασίες αξιολόγησης, που πρωτύτερα ήταν «αδιανόητες» (inconceivable).

### 3. ΔΙΑΓΝΩΣΤΙΚΗ, ΔΙΑΜΟΡΦΩΤΙΚΗ, ΤΕΛΙΚΗ ΨΗΦΙΑΚΗ ΑΞΙΟΛΟΓΗΣΗ

Ανάλογα με τον σκοπό που επιδιώκει η ψηφιακή αξιολόγηση, κατηγοριοποιείται (κατ' αντιστοιχία με την εκπαιδευτική αξιολόγηση) σε διαγνωστική, διαμορφωτική και τελική (ή «αθροιστική»).

Η **διαγνωστική** (ή «αρχική») αξιολόγηση υλοποιείται πριν την έναρξη της εκπαιδευτικής διαδικασίας και επιδιώκει να συλλέξει πληροφορίες για τον προσδιορισμό του αρχικού επίπεδου γνώσεων, τις προγενέστερες ιδέες και αντιλήψεις, τις δυνατότητες και δεξιότητες των σπουδαστών σε ένα συγκεκριμένο γνωστικό αντικείμενο (Πετροπούλου, Κασιμάτη & Ρετάλης, 2005). Είναι η λιγότερο συχνά χρησιμοποιούμενη μορφή αξιολόγησης στην Ανώτατη Εκπαίδευση (Crisp, *Teacher's Handbook on e-Assessment*, 2011), ενώ υφίσταται η άποψη ότι αποτελεί ιδιαίτερο τύπο της διαμορφωτικής αξιολόγησης (Knight, 2001. McAlpine, 2002). Ο Crisp (2011) υποστηρίζει ότι η διαγνωστική αξιολόγηση πρέπει να εγκαταλείψει το αρνητικό μοντέλο στο οποίο βασίζεται, δηλαδή το μοντέλο που αναζητά αδυναμίες του σπουδαστή. Αν ενσωματωθεί ως βασική δραστηριότητα στο αρχικό στάδιο όλων των μαθημάτων, είναι δυνατόν να παρωθήσει τους σπουδαστές να αναλάβουν την ευθύνη της μάθησης και της αξιολόγησής τους, να τους υποδείξει ποιες είναι οι βασικές αρχές και κρίσιμες έννοιες του γνωστικού αντικειμένου που θα σπουδάσουν, να τους βοηθήσει να διαπιστώσουν το βαθμό ετοιμότητάς τους για την παρακολούθηση του μαθήματος. Επιπλέον μπορεί να συνδράμει τους διδάσκοντες για να προσαρμόσουν τις εισαγωγικές διδακτικές δραστηριότητες, ώστε η πλειονότητα των σπουδαστών να συμμετέχει σε αυτές επαρκώς (Crisp, 2011).

Η **διαμορφωτική** αξιολόγηση υλοποιείται σε όλη τη διάρκεια της εκπαιδευτικής διαδικασίας και επιδιώκει να παράσχει συνεχή ανατροφοδότηση στους σπουδαστές και στους διδάσκοντες. Δεν αποφάινεται για την επιτυχή ή μη επιτυχή επίτευξη εκπαιδευτικών στόχων, αλλά υποστηρίζει τη μάθηση παρέχοντας ανατροφοδότηση στους σπουδαστές και στους διδάσκοντες. Η ανατροφοδότηση δίνει τη δυνατότητα στους σπουδαστές να ανιχνεύσουν τις μαθησιακές αδυναμίες τους και γενικότερα να αναστοχαστούν για τη μαθησιακή πορεία τους. Βοηθά τους διδάσκοντες να προσαρμόσουν τις διδακτικές παρεμβάσεις τους με σκοπό την πρόοδο της μαθησιακής διαδικασίας (Πετροπούλου, Κασιμάτη & Ρετάλης, 2005). Σύμφωνα με τους Daly, Pachler, Mor & Mellar (2010) η μαθησιακή πρόοδος δεν μπορεί να επιτευχθεί, αν οι σπουδαστές

δεν κατανοήσουν τους σκοπούς και τις διαδικασίες της μάθησης. Οι σπουδαστές πρέπει, για παράδειγμα, να γνωρίζουν ποιο είναι το γνωστικό επίπεδό τους σε σχέση με τα επιδιωκόμενα μαθησιακά αποτελέσματα και πώς θα τα επιτύχουν. Η διαμορφωτική αξιολόγηση παρέχει πληροφορίες που βοηθούν τους διδάσκοντες να προβούν σε απαραίτητες βελτιωτικές προσαρμογές στην εκπαιδευτική διαδικασία. Βοηθά επίσης τους σπουδαστές να κατανοήσουν τι επιδιώκουν να μάθουν, τι γνωρίζουν ήδη, τι δεν γνωρίζουν, πώς θα βελτιώσουν την απόδοσή τους, πώς θα κατακτήσουν τις γνώσεις που δεν κατέχουν. Συνεπώς, η διαμορφωτική αξιολόγηση βασίζεται σε συνεχείς, πλούσιες με πληροφορίες αλληλεπιδράσεις και σε πολυτροπική επικοινωνία του διδάσκοντα με τους σπουδαστές, του σπουδαστή με τους άλλους σπουδαστές και του σπουδαστή με τον εαυτό του. Με αυτόν τον τρόπο η διαμορφωτική αξιολόγηση αποτελεί αναπόσπαστο τμήμα της μαθησιακής διαδικασίας.

Η **τελική** αξιολόγηση διενεργείται κατά την ολοκλήρωση της εκπαιδευτικής διαδικασίας και επιδιώκει να αποτιμήσει συνολικά τα επιτεύγματα του σπουδαστή, καταλήγοντας σε αξιολογική κρίση ή σε βαθμολογία (Stödberg, 2012). Η τελική αξιολόγηση εφαρμόζεται κυρίως σε εξετάσεις υψηλής κρισιμότητας, οι οποίες αποτελούν την καθοριστική δοκιμασία για να διαπιστωθεί η επίτευξη των μαθησιακών στόχων και οι οποίες κρίνουν την απονομή τίτλων σπουδών, πιστοποιήσεων, την εγγραφή σε πρόγραμμα σπουδών ανώτερης εκπαιδευτικής βαθμίδας. Η κρισιμότητα και οι τυπικές συνέπειες των τελικών αξιολογήσεων (απονομή τίτλων σπουδών και πιστοποιήσεων), αφενός είναι η αιτία για τη μεγάλη έμφαση που τα ανώτατα εκπαιδευτικά ιδρύματα προσδίδουν σε αυτές (Gikandi, Morrow, & Davis, 2011) και αφετέρου αποτελούν ανασταλτικό παράγοντα για πειραματισμούς και καινοτομίες. Οι τελικές αξιολογήσεις είναι, ως επί το πλείστον, ρυθμισμένες από εθνικά ή διεθνή κανονιστικά πλαίσια και επομένως είναι λιγότερο ευέλικτες και ανοικτές προς την καινοτομία σε σύγκριση με τις διαμορφωτικές αξιολογήσεις (Perrotta & Whitelock, 2017). Οπότε, ευεξήγητη είναι η διαπίστωση των Boyle & Hutchison (2009), σύμφωνα με την οποία, οι περισσότερες έρευνες για την εισαγωγή τεχνολογικών καινοτομιών στην αξιολόγηση αφορούν διαμορφωτικούς και λιγότερο συχνά «τελικούς» σκοπούς. Σε αντίστοιχο συμπέρασμα καταλήγουν η James (2016) και ο Stödberg (2012) ο οποίος, στα πλαίσια της ευρείας και λεπτομερούς βιβλιογραφικής έρευνας που υλοποίησε επί του πεδίου της ψηφιακής αξιολόγησης στην ανώτατη εκπαίδευση, διαπίστωσε



ότι οι ερευνητικές εργασίες που αφορούν τη διαμορφωτική αξιολόγηση είναι περισσότερες από αυτές που αφορούν μόνο την τελική αξιολόγηση. Το συμπέρασμα συμφωνεί με την αναφορά των Perera-Diltz & Moe (2014) σύμφωνα με την οποία, η ψηφιακή τελική αξιολόγηση έχει υποστεί περιορισμένη θεωρητική μελέτη (under-theorized). Στη βιβλιογραφία (π.χ. Azevedo et al. 2019. Oliveira & Beites, 2019. Ridgway, McCusker & Pead, 2004. William, 2018) είναι ευρέως διαδεδομένη η άποψη, ότι η τελική αξιολόγηση υπηρετεί την αντίληψη «της μάθησης για την αξιολόγηση», ενώ η διαμορφωτική αξιολόγηση υπηρετεί την αντίληψη «της αξιολόγησης για τη μάθηση». Οι Azevedo, Oliveira & Beites (2019) θεωρούν ότι η ψηφιακή αξιολόγηση αποτελεί καταλύτη για την μετάβαση από την πρώτη αντίληψη, στη δεύτερη.

Η τελική αξιολόγηση ως διαδικασία ελέγχου της επίτευξης των μαθησιακών στόχων σε υψηλής σπουδαιότητας εξετάσεις, είναι κυρίαρχη στην ανώτατη εκπαίδευση. Ωστόσο, υφίστανται επιφυλάξεις ως προς τη χρήση της ως κύριου και μοναδικού τρόπου αξιολόγησης στα ψηφιακά περιβάλλοντα μάθησης (Stewart, Waight, Norwood & Ezell, 2004). Οι Perera-Diltz & Moe (2014), εγείρουν παιδαγωγικά και γνωστικά ζητήματα. Συγκεκριμένα, θεωρούν ότι ο υψηλός βαθμός δυσκολίας που ενέχει η ενσωμάτωση στην αξιολόγηση αρχών σχετικών με την εποικοδομιστική, τη μαθητοκεντρική και την αυθεντική εκπαίδευση, αυξάνεται περαιτέρω στα ψηφιακά περιβάλλοντα μάθησης. Αναφέρουν ότι σύμφωνα με πολλούς υποστηρικτές της ψηφιακής μάθησης, οι αλληλοεπιδραστικές και επικοινωνιακές δυνατότητες των εργαλείων της ψηφιακής μάθησης καθιστούν δυσδιάκριτα τα όρια μεταξύ των διαδικασιών της μάθησης και της αξιολόγησης. Επιπλέον, υποστηρίζουν ότι απαιτείται δέσμευση σε εποικοδομιστικές παιδαγωγικές αρχές και ευχέρεια χρήσης των τεχνολογικών εργαλείων προκειμένου να επαναπροσδιοριστούν οι τρόποι που υλοποιείται, οι πρακτικές που χρησιμοποιεί και οι σκοποί που εξυπηρετεί η ψηφιακή τελική αξιολόγηση. Αναφέρουν επίσης ότι προκειμένου να αξιολογούνται τα ουσιαστικά επιτεύγματα των σπουδαστών και όχι απλώς η ικανότητα τους για στείρα απομνημόνευση, απαιτείται η βέλτιστη χρήση εργαλείων που θα εστιάζουν στην επίλυση προβλημάτων, στην κριτική ανάλυση πηγών και στην αναβάθμιση του ρόλου των σπουδαστών ως συνδημιουργών της εκπαιδευτικής εμπειρίας. Για να σχεδιαστούν βασικές διαδικασίες ψηφιακής τελικής αξιολόγησης απαιτούνται κατάλληλες δεξιότητες από τον διδάσκοντα, προκειμένου να χειρίζεται με ευχέρεια τα τεχνολογικά εργαλεία, που θα του επιτρέψουν να

εφαρμόσει αποτελεσματικές αξιολογικές στρατηγικές και να επιτύχει την ενεργή συμμετοχή των σπουδαστών και τη διευκόλυνση της συνδημιουργίας έργων και εργασιών. Η James (2016) αναδεικνύει τα, συχνώς αναφερόμενα στη βιβλιογραφία, ζητήματα που σχετίζονται με την ασφάλεια και την ταυτοποίηση των σπουδαστών στις ψηφιακές τελικές αξιολογήσεις, τα οποία θεωρεί ότι περιορίζουν τη διάδοση των ψηφιακών τελικών αξιολογήσεων. Υποτίθεται, ότι χωρίς άμεση επιτήρηση οι σπουδαστές είναι επιρρεπείς σε πρακτικές ανεντιμότητας. Ωστόσο, η κριτική, ως προς τη διασφάλιση της ακεραιότητας είναι περισσότερο εξαντλητική και αυστηρή για τις ψηφιακές μεθόδους αξιολόγησης έναντι των παραδοσιακών μεθόδων, που επίσης πλήττονται από συμβάντα ανεντιμότητας (McCabe, 2005. Schmelkin, Gilbert, Spencer, Pincus & Silva, 2008. Whitley, 1998).

### 3.1. Διαμορφωτική έναντι τελικής αξιολόγησης

Όπως αναφέρουν οι Parchler et al. (2010) και ο Stodberg (2012), η διάκριση μεταξύ των τύπων της αξιολόγησης δεν έγκειται στον τρόπο, στη μέθοδο ή στο μέσο της αξιολόγησης, αλλά στον επιδιωκόμενο σκοπό. Έτσι, το ίδιο μέσο αξιολόγησης μπορεί να χρησιμοποιηθεί, είτε για διαμορφωτικούς, είτε για τελικούς σκοπούς. Για παράδειγμα, ένα διαδικτυακό διαγώνισμα με ερωτήσεις πολλαπλής επιλογής είναι δυνατόν να χρησιμοποιηθεί από τους σπουδαστές για να ανιχνεύσουν τυχόν αδυναμίες τους στην κατανόηση του γνωστικού αντικείμενου κατά τη διάρκεια ενός προγράμματος σπουδών. Οπότε, εξυπηρετεί σκοπό διαμορφωτικής αξιολόγησης. Είναι όμως δυνατόν, να χρησιμοποιηθεί για να κρίνει την απονομή ή όχι ενός τίτλου σπουδών κατά την ολοκλήρωση του προγράμματος σπουδών. Οπότε, εξυπηρετεί σκοπούς τελικής αξιολόγησης. Ανάλογη είναι και άποψη των Daly et al. (2010) που υποστηρίζουν ότι οι διαφορές μεταξύ διαμορφωτικής και τελικής αξιολόγησης δεν είναι αρκετά ευδιάκριτες στα πλαίσια της ψηφιακής μάθησης και κυρίως στις περιπτώσεις όπου χρησιμοποιούνται συνεργατικά και προσωποποιημένα μαθησιακά περιβάλλοντα. Οι διαφορές δεν σχετίζονται με εγγενή χαρακτηριστικά των διαδικασιών, αλλά με το χρόνο που διεξάγεται η αξιολόγηση και με την ύπαρξη και την αξιοποίηση της βαθμολογίας και της ανατροφοδότησης. Συνεπώς, η ψηφιακή αξιολόγηση είναι δυνατόν να οδηγήσει σε σύγκλιση τις διαμορφωτικές και τις τελικές αξιολογικές διαδικασίες. Υφίστανται παραδείγματα διαμορφωτικών αξιολογήσεων που

μοιάζουν με διαδοχικές τελικές αξιολογήσεις, όπου η ανατροφοδότηση συνήθως έχει τη μορφή βαθμολογίας ή αξιολογικής κρίσης.

Οι Ridgway, McCusker & Pead (2004) παρουσιάζουν έναν εκτενή κατάλογο χαρακτηριστικών που διαφοροποιούν τη διαμορφωτική από την τελική ψηφιακή αξιολόγηση. Τα χαρακτηριστικά αφορούν στις συνέπειες, στην εξωτερική «ανταλλακτική αξία», στο ενδιαφερόμενο κοινό, στην αυτενέργεια, στις μεθόδους εγκυρότητας, στην ποιότητα, στους απαιτούμενους πόρους, στη σαφήνεια της αξιολογούμενης γνώσης, στο κύρος, στις αξιολογούμενες ικανότητες, στη θεωρητική τεκμηρίωση, στα χρησιμοποιούμενα εργαλεία και στο βαθμό απόκρυψης γνωστικών αδυναμιών («mendacity quotient») (Knight, 2001), δηλαδή στο μικρότερο δισταγμό των σπουδαστών να αποκαλύψουν τις μαθησιακές αδυναμίες τους κατά τη διαμορφωτική αξιολόγηση, καθώς δεν θα υποστούν βαθμολογική συνέπεια, αντιθέτως θα οφηληθούν λόγω της ανατροφοδότησης.

Ως προς τη βαθμολογία στη διαμορφωτική αξιολόγηση, ο Smith (2009) θεωρεί ότι διατηρεί ενεργό το ενδιαφέρον του σπουδαστή για την επίδοσή του και ενισχύει τα κίνητρά του για να βελτιώνεται, ενώ ο Wosley (2008) υποστηρίζει ότι η βαθμολογία αποσπά την προσοχή του σπουδαστή από τη μάθηση. Κοινό χαρακτηριστικό της διαμορφωτικής και της τελικής ψηφιακής αξιολόγησης είναι η επιδίωξη της προσαρμοστικότητας, δηλαδή η υποστήριξη πρακτικών που βοηθούν τον διδάσκοντα και τους σπουδαστές να αναπροσαρμόζουν διαρκώς τις ενέργειές τους για την επίτευξη των μαθησιακών στόχων. Οι προσαρμογές που επιχειρεί ο διδάσκοντας πρέπει να αφορούν όχι μόνο τεχνολογικές, αλλά και παιδαγωγικές παραμέτρους, ενώ οι προσαρμογές που η αυτοματοποιημένη ανατροφοδότηση υποδεικνύει στους σπουδαστές πρέπει να έχουν πραγματικό διαμορφωτικό όφελος (Daly et al., 2010).

### 3.2. Αξία διαμορφωτικής ψηφιακής αξιολόγησης

Στη βιβλιογραφία καταγράφονται διαφορετικές προσεγγίσεις σχετικά με τη φύση και την αξία της διαμορφωτικής ψηφιακής αξιολόγησης. Κάποιες φορές παρουσιάζεται κυρίως, ως ένας σχετικά εύκολος τρόπος για να εισαχθεί η ψηφιακή αξιολόγηση στις πρακτικές ενός εκπαιδευτικού οργανισμού, διότι υποτίθεται ότι σχετίζεται με αξιολογικές διαδικασίες μικρής σπουδαιότητας, οι οποίες δεν ενέχουν μεγάλους κινδύνους. Συχνά επίσης εκλαμβάνεται και

εφαρμόζεται ως επαναλαμβανόμενη τελική αξιολόγηση (Pachler, Daly, Mor, & Mellar, 2010). Ωστόσο, υπάρχουν απόψεις που αμφισβητούν αυτήν την προσέγγιση και επιχειρούν να αναδείξουν τον σημαντικό ρόλο της διαμορφωτικής ψηφιακής αξιολόγησης στην παροχή ανατροφοδότησης και στην ανάληψη από τους σπουδαστές της ευθύνης για τη μαθησιακή τους πρόοδο και γενικά να τονίσουν τις δυνατότητες που η διαμορφωτική ψηφιακή αξιολόγηση προσφέρει για την υλοποίηση σημαντικών αλλαγών, ως προς το τρόπο που παρέχεται η εκπαίδευση στα τριτοβάθμια εκπαιδευτικά ιδρύματα (Nicol & Macfarlane-Dick, 2006., Sadler 1989). Σύμφωνα με τον Whitelock (2007) και τους Nicol & Macfarlane-Dick (2006), όπως παρατίθενται από τους Pachler et al. (2010), η διαμορφωτική ψηφιακή αξιολόγηση βοηθά τους σπουδαστές να αυτορρυθμίσουν και να αποκτήσουν έλεγχο στις διαδικασίες της μάθησής τους και υποστηρίζει τη συνεχή συλλογή, ανάλυση και αποτίμηση πληροφοριών.

Οι Blanco & Ginovart (2012) υποστηρίζουν ότι η διαμορφωτική ψηφιακή αξιολόγηση κατέχει κεντρική θέση στις αρχές του Ευρωπαϊκού Χώρου Ανώτατης Εκπαίδευσης, σύμφωνα με τις οποίες, οι διδάσκοντες, κατά τον σχεδιασμό και την υλοποίηση διαδικασιών διαμορφωτικής αξιολόγησης, πρέπει να λαμβάνουν υπόψη τις ενέργειες και τις στάσεις των σπουδαστών πριν, κατά τη διάρκεια και μετά την μαθησιακή διαδικασία, να αξιολογούν τόσο τα αποτελέσματα όσο και τις διαδικασίες της μάθησης, να παρέχουν ανατροφοδότηση για να βελτιώνουν την μαθησιακή διαδικασία, να αξιοποιούν πρακτικές αυτό-αξιολόγησης και αυτορρύθμισης, να κοινοποιούν και να διευκρινίζουν στους σπουδαστές τα κριτήρια της αξιολόγησης.

Οι Daly et al. (2010) εντόπισαν επαναλαμβανόμενα μοτίβα (patterns) σε θεμελιώδη χαρακτηριστικά των διαμορφωτικών ψηφιακών αξιολογικών διαδικασιών, τα οποία παραπέμπουν σε τεχνολογικές ιδιότητες και τεχνολογικούς πόρους που διαμορφώνουν το μαθησιακό περιβάλλον και τα οποία παρέχουν δυνατότητες για τη βελτίωση των σπουδαστών μέσω διαδικασιών αλληλεπίδρασης και ανατροφοδότησης. Οι εν λόγω ιδιότητες και πόροι είναι οι παρακάτω:

- Ταχύτητα: Η ταχύτητα της ανατροφοδότησης είναι κρίσιμη για την αποτελεσματικότητα της διαδικασίας. Η ταχύτητα επιτρέπει την αύξηση στη συχνότητα της διαδικασίας. Όσο γρηγορότερα λάβει ανατροφοδότηση ο σπουδαστής, τόσο γρηγορότερα θα προχωρήσει

στην επόμενη δοκιμασία. Όσο περισσότερες είναι οι δοκιμασίες στις οποίες θα λάβει μέρος ο σπουδαστής, τόσο περισσότερο αυξάνεται η απόδοσή του.

- Διαχείριση μεγάλου όγκου δεδομένων: Η δυνατότητα πρόσβασης και διαχείρισης μεγάλου όγκου δεδομένων επιτρέπουν την ευρεία αναζήτηση και επιλογή των καταλληλότερων δεδομένων για προσωποποιημένη ανατροφοδότηση προς τον σπουδαστή (π.χ. διαμόρφωση συνόλου πηγών από ψηφιακές βιβλιοθήκες με βάση τις ιδιαίτερες ανάγκες του σπουδαστή).
- Επεξεργασία: Σημαντικός είναι ρόλος των αυτοματισμών α) στην άμεση ανάλυση των απαντήσεων και στην παροχή κατάλληλης ανατροφοδότησης, β) στην κλιμάκωση και επέκταση των συστημάτων και γ) στην προσαρμογή των συστημάτων στις ιδιαίτερες ανάγκες των σπουδαστών.
- Επικοινωνία: Οι επικοινωνιακές δυνατότητες που παρέχει η τεχνολογία επιταχύνουν τη διακίνηση ιδεών μεταξύ όσων συμμετέχουν στην εκπαιδευτική διαδικασία. Οι πληροφορίες και τα ψηφιακά αντικείμενα που διακινούνται είναι δυνατόν να αποθηκευθούν και είναι δικτυακώς προσβάσιμα οποιαδήποτε χρονική στιγμή.
- Κατασκευή και αναπαράσταση: Η τεχνολογία διευκολύνει τους σπουδαστές να κατασκευάσουν ψηφιακά αντικείμενα με τα οποία αναπαριστούν τις ιδέες τους. Οι αναπαραστάσεις μπορούν να υλοποιηθούν με εναλλακτικούς τρόπους και να χρησιμοποιούν σύμβολα, τα οποία επιτρέπουν να σχηματοποιηθούν και να κοινοποιηθούν οι έννοιες.
- Ευμεταβλητότητα: Τα διαμοιραζόμενα ψηφιακά αντικείμενα είναι ευμετάβλητα. Εύκολα και γρήγορα μπορούν να αλλάξουν, όπως π.χ. συμβαίνει στα wiki, δηλαδή στα εργαλεία συνεργατικής ανάπτυξης περιεχομένου, που χρησιμοποιούνται σε συνεργατικά περιβάλλοντα μάθησης.

Οι παραπάνω ιδιότητες και πόροι αποτελούν τμήματα ενός σύνθετου και εξελισσόμενου περιβάλλοντος μάθησης, όπου οι σπουδαστές και οι διδάσκοντες αλληλεπιδρούν με πολλούς τρόπους και όπου η τεχνολογία προσφέρει δυνατότητες για ενεργή συμμετοχή. Συνεπώς, σύμφωνα με τους Daly et al. (2010) η ψηφιακή διαμορφωτική αξιολόγηση συνιστά ένα σύνολο

διαδικασιών που επιδιώκουν διαμορφωτικούς μαθησιακούς σκοπούς και σχετίζονται με τεχνολογικούς πόρους, αλλά και με ανθρώπινες αλληλεπιδράσεις. Οι τεχνολογικοί πόροι αδυνατούν από μόνοι τους να οδηγήσουν σε διαμορφωτικές διαδικασίες. Δρουν διαμορφωτικά μόνο στις περιπτώσεις που η αξιοποίησή τους από τους διδάσκοντες και τους σπουδαστές ακολουθεί κατευθύνση που βασίζεται σε στέρεες παιδαγωγικές αρχές και όχι απλώς σε καινοτόμες τεχνολογικές δυνατότητες.

### 3.3. Αλληλεπίδραση στη διαμορφωτική ψηφιακή αξιολόγηση

Στην βιβλιογραφική επισκόπηση σχετικά με την διαμορφωτική ψηφιακή αξιολόγηση που υλοποίησαν οι Gikandi et al. (2011) αναδεικνύουν τη σημασία της ψηφιακής διαμορφωτικής αξιολόγησης στην ανάπτυξη και στην ενδυνάμωση αλληλεπιδραστικών και συνεργατικών διαδικασιών. Συγκεκριμένα, αναφέρουν ότι η ψηφιακή αξιολόγηση (διαμορφωτική ή τελική), λόγω της ασύγχρονης αλληλεπίδρασης μεταξύ των συμμετεχόντων (σπουδαστών και διδασκόντων), παρουσιάζει διαφορετικά χαρακτηριστικά σε σύγκριση με την αξιολόγηση σε παραδοσιακά μαθησιακά περιβάλλοντα (Vonderwell, Liang, & Alderman, 2007), γεγονός που καθιστά απαραίτητη την εφαρμογή παιδαγωγικών αρχών συμβατών με τα ψηφιακά περιβάλλοντα μάθησης.

Οι ουσιώδεις αλληλεπιδράσεις εντός μιας ενεργής εκπαιδευτικής κοινότητας αποτελούν προαπαιτούμενο για την αλληλεπιδραστική συνεργασία, η οποία είναι κρίσιμη διαδικασία για την επίτευξη κριτικής σκέψης και την κατάκτηση υψηλών μαθησιακών στόχων στα ψηφιακά περιβάλλοντα μάθησης, ιδιαίτερα στις ανώτατες εκπαιδευτικές βαθμίδες (Akyol, Garrison & Yasar, 2009. Kehrwald, 2010). Ωστόσο, η δημιουργία αποτελεσματικών μαθησιακών κοινοτήτων, που θα αλληλεπιδρούν ουσιωδώς εντός διαδικτυακών και μεικτών περιβαλλόντων μάθησης δεν επιτυγχάνεται εύκολα. Διότι απαιτείται καλά δομημένη εκπαιδευτική στρατηγική, η οποία δεν είναι πάντα προφανής για τους διδάσκοντες (Akyol et al. 2009). Η αποτελεσματική ενσωμάτωση της ψηφιακής διαμορφωτικής αξιολόγησης παρέχει αυτή τη δυνατότητα. Συγκεκριμένα, παρέχει τις κατάλληλες δομές για την υποστήριξη ουσιωδών αλληλεπιδράσεων μεταξύ των σπουδαστών και των διδασκόντων και για την ανάπτυξη αποτελεσματικών μαθησιακών κοινοτήτων που θα επιδιώξουν υψηλά μαθησιακά αποτελέσματα (Sorensen & Takle, 2005). Η συνεχής υποστήριξη των σπουδαστών μέσω της κατάλληλης ανατροφοδότησης

είναι εξαιρετικά σημαντική σε περιβάλλοντα ψηφιακής μάθησης και είναι δυνατόν να παρασχεθεί μέσω της αλληλεπιδραστικής συνεργασίας μεταξύ σπουδαστών και διδασκόντων (Ludwig-Hardman & Dunlap, 2003). Διότι δίνει τη δυνατότητα στους σπουδαστές να έχουν ενεργό ρόλο στην εκπαιδευτική διαδικασία, να αυτορρυθμίζουν τη συμμετοχή τους και να αναλαμβάνουν την ευθύνη για την μαθησιακή πορεία τους, στοιχείο απαραίτητο στην ψηφιακή μάθηση.

### 3.3.1. Εγκυρότητα, Αξιοπιστία, Ανεπιτητικότητα

Οι Gikandi et al. (2011) θεωρούν ότι στα ψηφιακά περιβάλλοντα μάθησης, εξαιτίας της ασύγχρονης φύσης των αλληλεπιδράσεων μεταξύ των σπουδαστών και των διδασκόντων, είναι απαραίτητοι ο επαναπροσδιορισμός και ο εμπλουτισμός (υπό το πρίσμα των ιδιαιτεροτήτων της ψηφιακής μάθησης) τριών βασικών ζητημάτων σχετικών με τη διαμορφωτική ψηφιακή αξιολόγηση: *της εγκυρότητας, της αξιοπιστίας και της ανεπιτητικότητας (dishonesty)*.

Η *εγκυρότητα*, σύμφωνα με όσα υποστηρίζουν, αποτελεί τον βαθμό κατά τον οποίο η διαμορφωτική αξιολόγηση προάγει τη μάθηση. Συνεπώς, η εγκυρότητα κρίνεται εκ του αποτελέσματος και όχι εκ του επιδιωκόμενου σκοπού. Δηλαδή, η πρόθεση και μόνο για σχεδιασμό διαμορφωτικής ψηφιακής αξιολόγησης δεν προσδίδει στη διαδικασία την ιδιότητα της εγκυρότητας. Η ειδοποιός διαφορά που κρίνει την επίτευξη της εγκυρότητας, βρίσκεται στην αξιοποίηση των πορισμάτων της αξιολόγησης, δηλαδή στη χρήση τους για διαμορφωτικούς σκοπούς (Pachler, Daly, Mor & Mellar, 2010). Η εγκυρότητα επιτυγχάνεται αν οι σπουδαστές και οι διδάσκοντες αναγνωρίζουν ότι το μαθησιακό αποτέλεσμα είναι αξιόλογο. Η εγκυρότητα πρέπει να περιλαμβάνει τέσσερα χαρακτηριστικά: την *αυθεντικότητα*, την *αποτελεσματική ανατροφοδότηση*, την *πολυδιάστατη προοπτική* και την *υποστήριξη των σπουδαστών*.

- Η *αυθεντικότητα* εξαρτάται καταρχήν από τη συνάφεια των δραστηριοτήτων με το γνωστικό αντικείμενο. Οι σπουδαστές πρέπει να συμμετέχουν σε διαδικασίες επίλυσης προβλημάτων που αντιστοιχούν σε ρεαλιστικές καταστάσεις σχετικές με το γνωστικό αντικείμενο και σε διαδικασίες όπου απαιτείται λήψη αποφάσεων. Με αυτόν τον τρόπο αφενός δεν προσεγγίζουν επιφανειακά το γνωστικό αντικείμενο και αφετέρου εξασκούν ανώτερες γνωστικές ικανότητες.

- Η *αποτελεσματική ανατροφοδότηση* οφείλει να είναι γρήγορη, έγκαιρη, συνεχής, χρήσιμη και κατανοητή (Wolsey, 2008), ειδάλλως οι σπουδαστές τείνουν να την αγνοούν. Η αποτελεσματική ανατροφοδότηση είναι απαραίτητη προκειμένου να ξεπεραστούν τα, εγγενή στη ψηφιακή μάθηση, εμπόδια που προκαλεί η απουσία της φυσικής εγγύτητας και της προσωπικής επαφής, εξαιτίας των οποίων οι σπουδαστές διστάζουν ή αποφεύγουν εντελώς να ζητήσουν βοήθεια και διευκρινήσεις. Ιδιαίτερη σημαντική είναι η παροχή έμμεσης ανατροφοδότησης (π.χ. μέσω βιβλιογραφικών αναφορών και παράλληλων πηγών), που παρωθεί και εξασκεί τους σπουδαστές στη συνεχή αναζήτηση και αυτορρύθμιση της μάθησής τους. Την αποτελεσματική ανατροφοδότηση ενισχύει η διαφάνεια των επιδιωκόμενων μαθησιακών στόχων (Wolsey, 2008), των κριτηρίων και της μεθοδολογίας της αξιολόγησης.
- Η *πολυδιάστατη προοπτική* σχετίζεται με τη χρήση πολλαπλών εναλλακτικών διαδικασιών αξιολόγησης, η οποία επιτρέπει βαθμούς αυτονομίας και ευελιξίας που επιτρέπουν στους σπουδαστές να επιλέξουν το πρόσφορο μέσο και τον κατάλληλο τρόπο για να παρουσιάσουν τις γνώσεις και τις ικανότητές τους. Ωστόσο, το συγκεκριμένο χαρακτηριστικό οφείλει να λειτουργεί σε ισορροπία με το χαρακτηριστικό της αυθεντικότητας, καθώς η αυτονομία και η ευελιξία στην επιλογή των μέσων και των διαδικασιών της αξιολόγησης, ενδεχομένως να μην είναι πάντα επιτεύξιμες σε περιπτώσεις, όπου η οφειλόμενη αυθεντικότητα της αξιολόγησης ως προς το γνωστικό αντικείμενο περιορίζει τις διαθέσιμες επιλογές στα μέσα και στις διαδικασίες της αξιολόγησης.
- Η *υποστήριξη των σπουδαστών* πρέπει να αποτελεί συνεχή μέριμνα των διδασκόντων με στόχο, όχι μόνο την κατάκτηση των γνώσεων που προβλέπει το πρόγραμμα σπουδών, αλλά και την ανάπτυξη κριτικής σκέψης και άλλων ανώτερων γνωστικών ικανοτήτων απαραίτητων για τις ανώτερες εκπαιδευτικές βαθμίδες. Η υποστήριξη των σπουδαστών είναι δυνατόν να προέρχεται όχι μόνο από τους διδάσκοντες, αλλά και από άλλους σπουδαστές μέσω αλληλεπιδράσεων, τις οποίες μπορεί να ενθαρρύνει ο διδάσκοντας μέσω συνεργασιών και ομότιμης ανατροφοδότησης (Gikandi et al., 2011)



Η *αξιοπιστία* στα πλαίσια της διαμορφωτικής ψηφιακής αξιολόγησης αποτελεί, σύμφωνα με τους Gikandi et al. (2011), τον βαθμό κατά τον οποίο, αυτό που εξετάζεται είναι επαρκές για να μετρήσει το επίπεδο της γνώσης που έχει κατακτηθεί. Η αξιοπιστία πρέπει να περιλαμβάνει τρία χαρακτηριστικά : *δυνατότητα για τεκμηρίωση της μάθησης, πολλαπλές πηγές τεκμηρίων της μάθησης και σαφήνεια και ενημέρωση για τους μαθησιακούς σκοπούς και για τα αναλυτικά κριτήρια αξιολόγησης.*

- Η *δυνατότητα για τεκμηρίωση* είναι απαραίτητη προκειμένου να παρουσιαστεί η μαθησιακή πρόοδος και να διαγνωσθούν τυχόν αδυναμίες, ώστε να ληφθούν κατάλληλες ενέργειες, που θα οδηγήσουν στο επιθυμητό εκπαιδευτικό αποτέλεσμα. Με τον τρόπο αυτό, εκτός από τον διδάσκοντα, και οι σπουδαστές παρακολουθούν συνεχώς την μαθησιακή πρόοδό τους, γεγονός που τους παρωθεί στην αυτορρύθμιση της μάθησής τους.
- Οι *πολλαπλές πηγές τεκμηρίων* επιτρέπουν στους σπουδαστές να παρουσιάσουν τις ικανότητές τους με εναλλακτικούς τρόπους και να παράσχουν πολλαπλά πειστήρια για το επίπεδο των γνώσεων τους.
- Η *σαφήνεια και η ενημέρωση για τους μαθησιακούς στόχους και τα κριτήρια αξιολόγησης* καθοδηγούν τους σπουδαστές και τους επιτρέπουν, αφενός να αποσαφηνίσουν ποιο είναι το επιδιωκόμενο αποτέλεσμα και αφετέρου να έχουν ενεργό και υπεύθυνο ρόλο στη μαθησιακή διαδικασία (Vonderwell, Liang, & Alderman, 2007). Επιπλέον, η βαθμολόγηση δεν πρέπει να είναι γενική, δηλαδή να στηρίζεται σε γενικές και αόριστες αντιλήψεις για την απόδοση των σπουδαστών. Αλλά, πρέπει να είναι αναλυτική, τεκμηριωμένη, να ακολουθεί συγκεκριμένα κριτήρια και να συνοδεύεται από επιπλέον αξιολογικές πληροφορίες.

Η *ανεπιμότητα* στα πλαίσια της ψηφιακής διαμορφωτικής αξιολόγησης αφορά, σύμφωνα με τους Gikandi et al. (2011), την εξακρίβωση της πραγματικής ταυτότητας των αξιολογούμενων προσώπων και της πατρότητας των αξιολογούμενων έργων. Συμβάντα ανεπιμότητας περιορίζονται κατά πολύ, αν ο διαμορφωτικός χαρακτήρας της αξιολόγησης καταστεί σαφής

από τους διδάσκοντες και κατανοητός από τους σπουδαστές και αν η προσδοκώμενη απόδοση καταστεί γνωστή και κατανοητή μέσω της κοινοποίησης των κριτηρίων αξιολόγησης (Oosterhof, Conrad & Ely, 2008, σ. 138). Περαιτέρω περιορισμός των συμβάντων ανεντιμότητας επιτυγχάνεται, αν καλλιεργηθεί θετική στάση έναντι της μάθησης και αν αποθαρρυνθεί η επιφανειακή μελέτη και η επιπόλαιη ενασχόληση με το γνωστικό αντικείμενο. Επιπλέον, συμβάντα ανεντιμότητας είναι λιγότερο συχνά στις περιπτώσεις που επιτυγχάνεται το χαρακτηριστικό της αυθεντικότητας της αξιολόγησης (Oosterhof, Conrad & Ely, 2008). Αφενός, διότι αυξάνεται το ενδιαφέρον των σπουδαστών και η ενεργή και υπεύθυνη ενασχόληση τους με το γνωστικό αντικείμενο. Αφετέρου, διότι στις αξιολογήσεις αυθεντικών δοκιμασιών είναι ευκολότερο να παραχθούν αρκετές παραλλαγές διαγωνισμάτων, (εντελώς ή μερικώς διαφορετικών ανά αξιολογούμενο), γεγονός που δυσχεραίνει τις απόπειρες ανέντιμων πρακτικών.

Οι Gikandi et al. (2011) συνοψίζουν, αναφέροντας ότι η αποτελεσματική αξιοποίηση της ψηφιακής διαμορφωτικής αξιολόγησης μπορεί να προσφέρει μια καινοτόμο παιδαγωγική στρατηγική, που θα διευκολύνει σταθερές, ουσιώδεις αλληλεπιδράσεις και συνεργασίες μεταξύ σπουδαστών και διδασκόντων. Στόχος είναι η ενεργή ενασχόληση στα περιβάλλοντα ψηφιακής μάθησης και η μαθησιακή πρόοδος των σπουδαστών.

### 3.4. Ο ρόλος της ανατροφοδότησης στη ψηφιακή αξιολόγηση

Προκειμένου να αναδείξουν το ρόλο της ανατροφοδότησης στις λειτουργίες της ψηφιακής αξιολόγησης, οι Perrotta και Whitelock (2017) αντλούν πορίσματα από την πρωτοποριακή θεωρητική συνεισφορά του Sadler (1989) ο οποίος, όπως αναφέρουν, προετοίμασε θεωρητικά το πεδίο της ψηφιακής αξιολόγησης και ανέδειξε την έννοια της ανατροφοδότησης, ως ειδοποιό διαφορά μεταξύ διαμορφωτικής και τελικής αξιολόγησης. Στο θεωρητικό πλαίσιο του Sadler, η ανατροφοδότηση στην τελική αξιολόγηση έχει τη μορφή της αποτίμησης, ενώ στη διαμορφωτική αξιολόγηση προσφέρει ευκαιρίες όχι μόνο για την αποτίμηση, αλλά και για την περαιτέρω βελτίωση και ανάπτυξη φυσικών, γνωστικών και κοινωνικών δεξιοτήτων. Το ίδιο θεωρητικό πλαίσιο, απηχεί παιδαγωγικές αρχές οι οποίες πρεσβεύουν: Οτι οι σπουδαστές

αποτελούν το κέντρο της εκπαιδευτικής διαδικασίας και ότι πρέπει να έχουν τη δυνατότητα εποπτείας της ποιότητας των εργασιών τους. Ότι η επίγνωση των λειτουργιών που ενισχύουν τη μάθηση, όπως τα κίνητρα και η ενεργή συμμετοχή, είναι απαραίτητη για την αυτόνομη και αυτορρυθμιζόμενη μάθηση. Ότι τα άτομα και το κοινωνικό περιβάλλον συνεχώς αλληλοεπηρεάζονται και αλληλοδιαμορφώνονται.

Οι παραπάνω αρχές είναι απαραίτητες στον σχεδιασμό και στην ανάπτυξη τεχνολογικών εργαλείων για την υποστήριξη και τη βελτίωση της ψηφιακής αξιολόγησης. Οι σχεδιαστές των εργαλείων της ψηφιακής αξιολόγησης πρέπει να λαμβάνουν υπόψη ότι η μάθηση δεν αποτελεί μεταφορά γνώσης. Ότι η μάθηση προάγεται μέσω μαθητοκεντρικών μοντέλων και αυτορρυθμιζόμενων εκπαιδευτικών πρακτικών. Ότι οι ατομικές ικανότητες των σπουδαστών δεν είναι σταθερές και αμετάβλητες. Ότι είναι μεγάλη η σημασία των παραγόντων που ρυθμίζουν την αλληλεπίδραση, την επικοινωνία και την συνεργασία στα πλαίσια αποτελεσματικών διαδικασιών ανατροφοδότησης.

Το ζήτημα της ανατροφοδότησης στα ψηφιακά περιβάλλοντα μάθησης διερευνούν και οι Gikandi et al. (2011) ο οποίοι στην βιβλιογραφική επισκόπησή τους για την ψηφιακή διαμορφωτική αξιολόγηση παραθέτουν και σχολιάζουν σχετικά ερευνητικά πορίσματα. Συγκεκριμένα, αναφέρουν ότι σύμφωνα με τον Wosley (2008), στα ψηφιακά περιβάλλοντα μάθησης, η ανατροφοδότηση βοηθά τους σπουδαστές να εντοπίσουν τις αδυναμίες τους, να βελτιώσουν την απόδοσή τους, να υιοθετήσουν ενεργή και αυτορρυθμιζόμενη στάση για την μάθηση. Ο Wosley (2008) επίσης, περιέγραψε τρόπους μέσω των οποίων η ανατροφοδότηση ενισχύεται στα ψηφιακά περιβάλλοντα μάθησης με τη βοήθεια εφαρμογών που ενσωματώνουν στις εργασίες των σπουδαστών λεπτομερή και σαφή σχόλια ανατροφοδότησης και γενικότερα ενισχύουν την αλληλεπίδραση μεταξύ διδάσκοντα και σπουδαστών. Σύμφωνα με τον Koh (2008), ο οποίος μελέτησε ψηφιακά περιβάλλοντα μάθησης στην ανώτατη εκπαίδευση, η ανατροφοδότηση ενισχύει τα κίνητρα για μάθηση, την αυτοεκτίμηση των σπουδαστών, την αυτορρυθμιζόμενη μάθηση. Οι Sorensen & Takle (2005) και οι Vonderwell et al. (2007) μελέτησαν την αποτελεσματικότητα της ανατροφοδότησης που παρέχεται στα ψηφιακά περιβάλλοντα μάθησης σε σύγκριση με τα παραδοσιακά περιβάλλοντα. Συμπέραναν ότι οι αλληλεπιδράσεις και οι συνεργατικές δραστηριότητες που λαμβάνουν χώρα στα ψηφιακά

περιβάλλοντα παράγουν περισσότερο άμεση, επαρκή και συνεχόμενη ανατροφοδότηση. Οι Vonderwell et al. (2007) οδηγήθηκαν σε ίδια συμπεράσματα με τους Nicol & Milligan (2006) ως προς τη σημασία των διαδικτυακών χώρων δημόσιων συζητήσεων. Ο ασύγχρονος τρόπος λειτουργίας των διαδικτυακών δημόσιων συζητήσεων παρέχει στους σπουδαστές άνεση χρόνου για να μελετήσουν όσα περιλαμβάνει η δημόσια συζήτηση, για να συντάξουν τα δικά τους σχόλια και για να στοχαστούν πριν τα αναρτήσουν, διαδικασίες που λειτουργούν ως προσωπική ανατροφοδότηση. Ανάλογο είναι το συμπέρασμα των Van der Pol, Van den Berg, Admiraal & Simons (2008) οι οποίοι μελέτησαν διαδικασίες ομότιμης ανατροφοδότησης σε ψηφιακά περιβάλλοντα μάθησης στην ανώτατη εκπαίδευση. Συμπέραναν ότι ο βαθμός και η ποιότητα των αλληλεπιδράσεων επηρεάζει την αποτελεσματικότητα της ανατροφοδότησης και ότι τα ψηφιακά περιβάλλοντα μάθησης είναι δυνατόν να παράσχουν μεγαλύτερη αλληλεπίδραση, μέσω της οποίας αυξάνονται οι δυνατότητες για παροχή ομότιμης ανατροφοδότησης. Οι Chung, Shel & Kaizer (2006) υποστηρίζουν ότι μέσω της ανατροφοδότησης ενισχύεται η αυτορρύθμιση της μάθησης και ο αναστοχασμός, που αποτελούν σημαντικούς παράγοντες για την ανώτατη εκπαίδευση και για τους δια βίου εκπαιδευόμενους που επιδιώκουν τη συνεχή αυτοβελτίωση και την κατάκτηση νέων γνώσεων και επαγγελματικών δεξιοτήτων.

Οι Gikandi et al. (2011) συμπληρώνουν ότι οι δυναμικές αλληλεπιδράσεις, που αναπτύσσονται στα πλαίσια ασύγχρονων διαδικτυακών διαλόγων μεταξύ των σπουδαστών και μεταξύ του διδάσκοντα και των σπουδαστών, ενισχύουν την ανατροφοδότηση. Διότι παρέχουν κίνητρα και ικανοποίηση, οπότε παρωθούν τους σπουδαστές σε περισσότερο ενεργή συμμετοχή στην εκπαιδευτική διαδικασία. Υποστηρίζουν επίσης ότι τα διαδικτυακά και μεικτά περιβάλλοντα μάθησης επιτρέπουν στον διδάσκοντα να συλλέξει τεκμήρια της μαθησιακής πορείας των σπουδαστών, να εντοπίσει τις περιοχές όπου συχνότερα προκύπτουν αδυναμίες και να παρέμβει παρέχοντας την κατάλληλη γενική ανατροφοδότηση. Η πρακτική της γενικής ανατροφοδότησης είναι ιδιαίτερα αποδοτική, διότι εξοικονομεί πολύτιμο χρόνο, ωστόσο δεν μπορεί να αποτελέσει καθολική πρακτική δεδομένου ότι, σχεδόν πάντοτε είναι αναγκαία η εξατομικευμένη καθοδήγηση. Γενικά, τα ψηφιακά περιβάλλοντα μάθησης παρέχουν πολλές δυνατότητες για παροχή ανατροφοδότησης και συνεχούς υποστήριξης προς τους σπουδαστές.

Οι δυνατότητες αυτές διευκολύνουν τη μαθησιακή διαδικασία και ενισχύουν την ενεργή συμμετοχή των σπουδαστών.

Οι Carless, Salter, Yang & Lam (2001) μελετούν πρακτικές που ενισχύουν την αποτελεσματικότητα της ανατροφοδότησης σε διαδικασίες ψηφιακής αξιολόγησης στην ανώτατη εκπαίδευση. Μεταξύ άλλων, προτείνουν τη χρήση ψηφιακών εργαλείων, όπως των ασύγχρονων διαδικτυακών διαλόγων ή των ιστολογίων. Αναφέρουν ότι οι διαδικτυακοί διάλογοι και τα ιστολόγια αποτελούν αποτελεσματικότερο μέσο ανατροφοδότησης από ό,τι η παραδοσιακή (προφορική ή γραπτή) ανατροφοδότηση, διότι ενισχύουν την ενεργή συμμετοχή των σπουδαστών και τη γόνιμη ανταλλαγή απόψεων ενώ ο διδάσκοντας συμμετέχει για να παράσχει κατευθύνσεις και διευκρινήσεις. Την άποψη, ότι η ομότιμη ανατροφοδότηση, ως τμήμα διαδικασιών ψηφιακής αξιολόγησης, είναι ιδιαίτερα χρήσιμη και ισοδύναμη (ως προς αξιοπιστία και την αξία) με την ανατροφοδότηση που προέρχεται από τον διδάσκοντα, υιοθετούν και οι Fontanillas, Carbonell & Catusus (2016).

#### 3.4.1. Ανατροφοδότηση και αρχές διαμορφωτικής ψηφιακής αξιολόγησης

Σύμφωνα με τους Perrotta & Whitelock (2017) ο κυρίαρχος ρόλος της ανατροφοδότησης αναδεικνύεται στο μοντέλο της αυτορρυθμιζόμενης μάθησης των Nicol & Macfarlane-Dick (2006) οι οποίοι προτείνουν επτά αρχές για καλές πρακτικές διαμορφωτικής αξιολόγησης και ανατροφοδότησης στην ανώτατη εκπαίδευση. Το μοντέλο συμπληρώνουν οι Nicol & Milligan (2006) που προτείνουν τρόπους με τους οποίους οι επτά αρχές των καλών πρακτικών ανατροφοδότησης εφαρμόζονται σε περιβάλλοντα ψηφιακής μάθησης.

**Πίνακας 3. Αρχές για καλές πρακτικές διαμορφωτικής αξιολόγησης ( Nicol & Macfarlane-Dick, 2006)**

A/A	Αρχές για καλές πρακτικές διαμορφωτικής αξιολόγησης
1	Διασαφήνιση καλής απόδοσης.
2	Διευκόλυνση αυτοαξιολόγησης
3	Παροχή ανατροφοδότησης

4	Ενθάρρυνση διαλόγου
5	Ανάπτυξη κινήτρων και αυτοεκτίμησης
6	Άμβλυνση διαφοράς μεταξύ υφιστάμενης και επιθυμητής απόδοσης
7	Παροχή πληροφοριών προς τους διδάσκοντες

Η *πρώτη αρχή* αφορά τη διασαφήνιση για το τι συνιστά καλή απόδοση. Συχνά, σε όλα τα περιβάλλοντα μάθησης (παραδοσιακά ή ψηφιακά), υφίστανται παρανοήσεις για τους σκοπούς, τα αναμενόμενα επιτεύγματα και τα κριτήρια αξιολόγησης της μαθησιακής διαδικασίας, γεγονός που όχι μόνο δημιουργεί προβλήματα στους σπουδαστές να ανταποκριθούν στα ζητούμενα των εργασιών και των αξιολογικών διαδικασιών, αλλά επιπλέον περιορίζει την αποτελεσματικότητα της ανατροφοδότησης που λαμβάνουν. Αν οι σπουδαστές δεν κατανοούν ή παρανοούν τις απαιτήσεις που θέτουν οι διδάσκοντες, τότε η ανατροφοδότηση δεν βοηθά, ούτε να αποτιμήσουν το σημείο στο οποίο βρίσκονται, ούτε να ακολουθήσουν την πορεία που θα τους οδηγήσει στο σημείο που πρέπει να φτάσουν. Στην άρση των παρανοήσεων βοηθούν οι αναρτήσεις στα ψηφιακά περιβάλλοντα μάθησης και οι οδηγίες με σαφείς και αναλυτικές περιγραφές των κριτηρίων αξιολόγησης και των μέτρων για τα διάφορα επίπεδα επίδοσης. Όταν όμως, εξαιτίας της πολυπλοκότητας των αξιολογούμενων έργων, τα κριτήρια είτε είναι ασαφή, είτε είναι δύσκολο να εκφραστούν γραπτώς, τότε είναι αναγκαίο να αξιοποιηθούν και συμπληρωματικές πρακτικές. Ιδιαίτερα αποτελεσματική είναι η επίδειξη υποδειγματικών εργασιών, έναντι των οποίων οι σπουδαστές συγκρίνουν και αξιολογούν τις εργασίες τους. Στα ψηφιακά περιβάλλοντα μάθησης η επίδειξη υποδειγματικών εργασιών είναι εύκολα υλοποιήσιμη και η αποτελεσματικότητά της μπορεί να ενισχυθεί περαιτέρω με την αξιοποίηση διαδικτυακών δημοσίων διαλόγων, όπου οι σπουδαστές δημοσίως, μελετούν τα αναρτημένα υποδείγματα, καταγράφουν κριτήρια αξιολόγησης, αξιολογούν τα υποδείγματα και τεκμηριώνουν τις αξιολογήσεις τους, ενώ οι διδάσκοντες παρεμβαίνουν για να παράσχουν σχετικές διευκρινήσεις (Nicol & Milligan, 2006).

Η *δεύτερη αρχή* αφορά τη διευκόλυνση της αυτοαξιολόγησης. Καθώς οι σπουδαστές συμμετέχουν στη μαθησιακή διαδικασία, αναπτύσσουν την ικανότητα να αντιλαμβάνονται μόνοι τους την απόσταση μεταξύ της τρέχουσας απόδοσής τους και των επιδιωκόμενων στόχων. Οι διδάσκοντες πρέπει να αξιοποιήσουν τη συγκεκριμένη ικανότητα και να δημιουργήσουν όσο το δυνατόν περισσότερες δραστηριότητες, οι οποίες θα συμβάλουν με δομημένο τρόπο στον συνεχή αυτοέλεγχο των σπουδαστών. Ιδιαίτερως χρήσιμες προς αυτή την κατεύθυνση είναι οι αυτοαξιολογικές δραστηριότητες με τις οποίες οι σπουδαστές εξοπλίζονται με ικανότητες αυτοαξιολόγησης, αυτοβελτίωσης και αυτορρύθμισης της μαθησιακής πορείας τους και παύουν να είναι μονίμως εξαρτημένοι από τη καθοδήγηση του διδάσκοντα. Προϋπόθεση για την αυτορρύθμιση και την αυτοαξιολόγηση αποτελεί η ικανότητα των σπουδαστών για ορθή αναγνώριση των μαθησιακών στόχων και των κριτηρίων αξιολόγησης. Συνεπώς, η πρώτη αρχή, που σχετίζεται με τη διασαφήνιση του τι συνιστά καλή απόδοση, αποτελεί προαπαιτούμενο για τη δεύτερη αρχή.

Στα ψηφιακά περιβάλλοντα μάθησης η πιο κοινή πρακτική αυτοαξιολόγησης είναι η αξιοποίηση ψηφιακών διαγωνισμάτων με ερωτήσεις πολλαπλών επιλογών, τα οποία συνίσταται να συνοδεύονται από κατάλληλη ανατροφοδότηση. Ωστόσο, τα εν λόγω διαγωνίσματα παρουσιάζουν περιορισμούς. Συγκεκριμένα, οι σπουδαστές δεν έχουν συνήθως ρόλο στη θέσπιση των στόχων των διαγωνισμάτων, οπότε πιθανόν να προκύψουν προβλήματα διασαφήνισης των ερωτήσεων ή των σκοπών τους, κάτι που έρχεται σε αντίθεση με την πρώτη αρχή. Επιπλέον, η ανατροφοδότηση που παρέχεται είναι αυτοματοποιημένη και καθορισμένη πριν τη διεξαγωγή του διαγωνίσματος, οπότε δεν είναι ιδιαίτερα αποτελεσματική. Και τέλος, η δυνατότητα των ερωτήσεων πολλαπλής επιλογής να αξιολογήσουν ανώτερες γνωστικές λειτουργίες είναι αμφισβητήσιμη, ιδιαίτερα στις περιπτώσεις που οι ερωτήσεις δεν είναι σχεδιασμένες καταλλήλως. Ωστόσο, ο ρόλος των ψηφιακών διαγωνισμάτων δεν είναι ασήμαντος. Οι ίδιοι οι σπουδαστές τα θεωρούν χρήσιμα για να ελέγχουν τις γνώσεις τους, ιδιαίτερα αν τους δίνεται η δυνατότητα να τα επαναλαμβάνουν. Εξάλλου, υφίστανται τρόποι για ενίσχυση της μαθητοκεντρικής διάστασης και της ενεργής συμμετοχής των σπουδαστών στη χρήση ψηφιακών διαγωνισμάτων. Συγκεκριμένα, οι σπουδαστές μπορούν να συνθέσουν ερωτήσεις και να τις συμπεριλάβουν σε τράπεζες ερωτήσεων. Έτσι, οι σπουδαστές αφενός

βελτιώνουν τις γνώσεις τους για το γνωστικό αντικείμενο και ενισχύουν το μαθησιακό κίνητρό τους. Παράλληλα οι διδάσκοντες έχουν στη διάθεσή τους μια διευρυμένη τράπεζα ερωτήσεων, την οποία μπορούν να αξιοποιήσουν για χρήση στην ίδια ή σε άλλη ομάδα σπουδαστών.

Άλλο μέσο, που μπορεί να αξιοποιηθεί στα πλαίσια της δεύτερης αρχής για την ενίσχυση της αυτοαξιολόγησης, είναι οι προσομοιώσεις, οι οποίες θεωρούνται περισσότερο αποτελεσματικές από τα ψηφιακά διαγωνίσματα ως προς το στόχο της αυτορρύθμισης των σπουδαστών και συγκεκριμένα ως προς το ζήτημα της ανατροφοδότησης. Διότι, στις προσομοιώσεις η ανατροφοδότηση είναι σαφής, άμεση, δυναμική και εστιασμένη προς τους στόχους της αξιολογικής διαδικασίας. Άλλη καλή πρακτική για την ενίσχυση της αυτοαξιολόγησης, αποτελούν οι δραστηριότητες ομότιμης αξιολόγησης, όπου κάθε σπουδαστής αξιολογεί τη δική του εργασία ή την εργασία άλλου σπουδαστή και παράσχει σχετική ανατροφοδότηση. Οι συγκεκριμένες πρακτικές ομότιμης αξιολόγησης αναπτύσσουν την ικανότητα των σπουδαστών να αξιολογούν με ορθότητα συγκρίνοντας την αξιολογούμενη εργασία έναντι των κριτηρίων αξιολόγησης. Τα ψηφιακά περιβάλλοντα μάθησης προσφέρουν πολλές δυνατότητες που διευκολύνουν την ομότιμη αξιολόγηση. Αξιοσημείωτο είναι το σύστημα ομότιμης αξιολόγησης που ανέπτυξαν οι Bhalerao & Ward (2001), όπου κάθε εργασία κατατίθεται ψηφιακά, καθίσταται ανώνυμη, διανέμεται με τυχαίο τρόπο σε άλλους σπουδαστές, που καθένας τους χωριστά και ανώνυμα την αξιολογεί και παρέχει ανατροφοδότηση. Στο τέλος της διαδικασίας, η εργασία με τις αξιολογήσεις και όλα τα σχόλια ανατροφοδότησης επιστρέφεται στον δημιουργό της. Έτσι, κάθε σπουδαστής έχει την ευκαιρία, αφενός να αξιολογήσει πολλές εργασίες άλλων σπουδαστών και να παράσχει σχετική ανατροφοδότηση και αφετέρου να λάβει αξιολογήσεις και σχόλια ανατροφοδότησης από πολλούς σπουδαστές.

Η ενίσχυση της αυτοαξιολόγησης επιτυγχάνεται επίσης με την αξιοποίηση των ψηφιακών φακέλων επιτευγμάτων (e-portfolios). Διότι, κατά την διαμόρφωση και ανάπτυξη των ψηφιακών φακέλων επιτευγμάτων, οι σπουδαστές αναστοχάζονται για τα επιτεύγματά τους και επιλέγουν έργα που πληρούν τα κριτήρια της αξιολόγησης. Επιπλέον, ενδυναμώνεται το αίσθημα κυριότητας των σπουδαστών επί των έργων τους, τα οποία συχνά ανέπτυξαν με τη βοήθεια διαφορετικών τεχνικών και μέσων (π.χ. κείμενο, εικόνα, βίντεο). Με τη βοήθεια του ψηφιακού



φάκελου επιτευγμάτων, οι σπουδαστές είναι δυνατόν να συνδυάζουν και να συσχετίζουν διαφορετικά έργα, αποδεικνύοντας ότι είναι ικανοί να προσεγγίζουν το ίδιο γνωστικό αντικείμενο με πολλαπλές θεωρήσεις. Προφανώς, η διατήρηση και η διανομή των ψηφιακών φακέλων επιτευγμάτων είναι ευκολότερη από τη διατήρηση και διανομή των έντυπων φακέλων επιτευγμάτων, συνεπώς, οι σπουδαστές μπορούν αποτελεσματικά να διαχειρίζονται και να παρουσιάζουν τα έργα τους και τα μαθησιακά επιτεύγματά τους (Nicol & Milligan, 2006).

Η *τρίτη αρχή* αφορά την παροχή προς τους σπουδαστές υψηλής ποιότητας πληροφοριών σχετικά με την μάθησή τους. Ο διδάσκοντας έχει κεντρικό ρόλο στη παροχή πληροφοριών ανατροφοδότησης, που βοηθούν τους σπουδαστές να κατανοήσουν ποιοι είναι οι μαθησιακοί στόχοι, να ελέγξουν τις αδυναμίες τους και να βελτιώσουν την απόδοσή τους. Προκειμένου να έχει θετικά αποτελέσματα, η ανατροφοδότηση πρέπει να είναι έγκαιρη, να εστιάζει σε ανώτερες γνωστικές ικανότητες, να μην είναι επικριτική, να είναι περιγραφική και όχι αποτιμητική. Τα ψηφιακά περιβάλλοντα μάθησης παρέχουν στους διδάσκοντες πολλά εργαλεία για την παροχή ανατροφοδότησης. Για παράδειγμα, τα Συστήματα Διαχείρισης Μάθησης (ΣΔΜ) αυτοματοποιούν και επιταχύνουν την παροχή ανατροφοδότησης προς μεγάλες ομάδες σπουδαστών. Με τα συγκεκριμένα εργαλεία οι διδάσκοντες μπορούν να παράσχουν γενικά σχόλια ανατροφοδότησης προς μεγάλες ομάδες σπουδαστών, αλλά και εξειδικευμένα σχόλια προς μεμονωμένους σπουδαστές. Δεδομένου ότι υπάρχει δυνατότητα διαδικτυακής παρακολούθησης στην ανάπτυξη κάθε εργασίας, ο διδάσκοντας μπορεί να παράσχει ανατροφοδότηση, όχι μόνο για τελικό αποτέλεσμα, αλλά και για τα ενδιάμεσα στάδια ανάπτυξης. Ως προς την δυνατότητα αυτοματοποιημένης παροχής τυποποιημένων και εκ των προτέρων κατασκευασμένων σχολίων ανατροφοδότησης, είναι απαραίτητη η τεκμηρίωση για την αποτελεσματικότητα της συγκεκριμένης μεθόδου (Nicol & Milligan, 2006).

Η *τέταρτη αρχή* αφορά την ενθάρρυνση του διαλόγου για τη μάθηση μεταξύ του διδάσκοντα και των σπουδαστών και των σπουδαστών μεταξύ τους. Συχνά, οι σπουδαστές δεν κατανοούν την ανατροφοδότηση που λαμβάνουν από τους διδάσκοντες, οπότε δεν μπορούν να την αξιοποιήσουν για να βελτιώσουν τη μαθησιακή τους πορεία. Συνεπώς, η ανατροφοδότηση προκειμένου να είναι χρήσιμη δεν πρέπει να εξαντλείται σε μονόδρομη (από τον διδάσκοντα προς τον σπουδαστή) επικοινωνία. Πρέπει να παρέχεται η δυνατότητα στον σπουδαστή να

αναπτύσσει διάλογο, που θα τον βοηθήσει να διαλευκάνει σημεία της ανατροφοδότησης που δεν κατανοεί. Ωστόσο, λόγω του μεγάλου πλήθους των σπουδαστών στα μαθήματα των ιδρυμάτων της ανώτατης εκπαίδευσης, η ανάπτυξη διαλογικής επικοινωνίας με τους σπουδαστές προκαλεί μεγάλο φόρτο στους διδάσκοντες, πρόβλημα στο οποίο μπορούν να αναζητηθούν λύσεις με την αξιοποίηση της τεχνολογίας. Για παράδειγμα, μετά την επίλυση ψηφιακών διαγωνισμάτων, ο διδάσκοντας αναρτά σε κοινή θέα τη συχνότητα των απαντήσεων σε κάθε ερώτηση, χωρίς να αποκαλύπτει την ορθή απάντηση. Ζητά από τους σπουδαστές να δικαιολογήσουν τις επιλογές τους και προτρέπει ανταλλαγή επιχειρημάτων μεταξύ σπουδαστών που έδωσαν διαφορετικές απαντήσεις στην ίδια ερώτηση. Παρεμβαίνει για να δώσει σχετικές διευκρινήσεις ή για να συντονίσει τη διαδικασία. Η μέθοδος μπορεί να υλοποιηθεί, είτε εξ ολοκλήρου διαδικτυακά, είτε με μεικτό τρόπο, όπου για παράδειγμα ο διάλογος λαμβάνει χώρα στα πλαίσια ενός παραδοσιακού μαθήματος. Η μέθοδος περιλαμβάνει τρία διαφορετικά είδη ανατροφοδότησης: α) ποσοτική ανατροφοδότηση μέσω πληροφόρησης για τη συχνότητα των απαντήσεων, β) ανατροφοδότηση μέσω ομότιμου διαλόγου μεταξύ των σπουδαστών και γ) ανατροφοδότηση μέσω των διευκρινήσεων και οδηγιών που παρέχει ο διδάσκοντας.

Συνηθέστερη μέθοδος διεξαγωγής διαλόγων είναι αυτή που αξιοποιεί διαδικτυακούς χώρους δημόσιων συζητήσεων, όπου οι διάλογοι διεξάγονται με ασύγχρονο τρόπο. Ο ασύγχρονος τρόπος διαλογικής επικοινωνίας διαθέτει δύο μοναδικά χαρακτηριστικά. Πρώτον, διατηρεί τα αναρτημένα σχόλια επ' αόριστον, αφενός δίνοντας τη δυνατότητα στον διδάσκοντα να τα χρησιμοποιήσει μελλοντικά και αφετέρου δημιουργώντας μόνιμη πηγή ανατροφοδότησης, στην οποία οι σπουδαστές μπορούν να ανατρέξουν οποτεδήποτε. Δεύτερον, παρέχει άνεση χρόνου στους σπουδαστές να μελετήσουν τα αναρτημένα σχόλια, να συντάξουν τα δικά τους και να στοχαστούν επί αυτών, πριν τα αναρτήσουν. Οι διαδικτυακοί χώροι δημόσιων συζητήσεων είναι εξαιρετικά χρήσιμοι, διότι οι σπουδαστές μπορούν να παρακολουθήσουν και να διδαχθούν από όλη την εξέλιξη της δημόσιας συζήτησης και να ωφεληθούν όχι μόνο από τα σχόλια που αφορούν δικές τους μαθησιακές ανάγκες, αλλά και από αυτά που αφορούν άλλους σπουδαστές (Nicol & Milligan, 2006).

Η *πέμπτη αρχή* αφορά την ανάπτυξη κινήτρων και αυτοεκτίμησης, σημαντικών υποστηρικτικών παραγόντων για τη μάθηση. Η ανατροφοδότηση μπορεί να διαδραματίσει θετικό ή αρνητικό

ρόλο στην ανάπτυξη κινήτρων και να επηρεάσει τον βαθμό αυτοεκτίμησης των σπουδαστών. Οι δυνατότητες που παρέχουν τα ψηφιακά περιβάλλοντα μάθησης είναι δυνατόν να ενδυναμώσουν και τα κίνητρα και την αυτοεκτίμηση των σπουδαστών. Για παράδειγμα, τα ψηφιακά διαγωνίσματα διεξάγονται με εμπιστευτικό τρόπο, οπότε ο σπουδαστής δεν συγκρίνει την απόδοσή του με την απόδοση άλλων σπουδαστών. Επιδιώκοντας την αυτοβελτίωση και όχι τον ανταγωνισμό με τους άλλους σπουδαστές, το μαθησιακό κίνητρο ενισχύεται. Η δυνατότητα πολλών επαναλήψεων του διαγωνίσματος, ενισχύει το κίνητρο των σπουδαστών και εκτονώνει το άγχος της αξιολόγησης. Καθώς επιτυγχάνουν υψηλότερη απόδοση μέσω των επαναλαμβανόμενων προσπαθειών, οι σπουδαστές αποκτούν περισσότερα κίνητρα και αυτοεκτίμηση. Τα κίνητρα και η αυτοεκτίμηση ενισχύονται, διότι οι σπουδαστές βλέπουν τη βελτίωση της επίδοσής τους και λαμβάνουν άμεση και σαφή ανατροφοδότηση. Επιπλέον, οι δυνατότητες που παρέχουν τα ψηφιακά περιβάλλοντα μάθησης για την ευρεία και λεπτομερή συλλογή και καταγραφή των μαθησιακών επιτευγμάτων διευκολύνουν τους σπουδαστές να εποπτεύσουν τη πρόοδο της μαθησιακής πορείας, να αποκτήσουν αυτοεκτίμηση και να πειστούν ότι τα επιτεύγματα εξαρτώνται, όχι από προϋπάρχουσες νοητικές ικανότητες, αλλά από την προσπάθεια που καταβάλει κάποιος και την προσήλωση προς τον στόχο (Nicol & Milligan, 2006).

Η *έκτη αρχή* αφορά την άμβλυση της διαφοράς μεταξύ υφιστάμενης και επιθυμητής απόδοσης. Η ανατροφοδότηση για να φέρει το επιδιωκόμενο αποτέλεσμα δεν αρκεί να είναι απλώς ποιοτική. Πρέπει να είναι αποτελεσματική, δηλαδή να επιτυγχάνει βελτίωση στην απόδοση του σπουδαστή. Προκειμένου να διαπιστωθεί η βελτίωση, ο σπουδαστής πρέπει να έχει τη δυνατότητα να υποβάλλει ξανά την εργασία του. Συγκεκριμένα, μετά την πρώτη υποβολή, ο σπουδαστής δέχεται ανατροφοδότηση, βάση της οποίας προβαίνει σε βελτιώσεις και επανέρχεται υποβάλλοντας ξανά την εργασία. Με τον τρόπο αυτό, κλείνει ένας κύκλος ανατροφοδότησης, ο οποίος είναι δυνατόν να επαναληφθεί. Στα ανώτατα εκπαιδευτικά ιδρύματα, η επανάληψη υποβολής των εργασιών δεν συνηθίζεται, γεγονός που όχι απλώς μειώνει την αποτελεσματικότητα της ανατροφοδότησης, αλλά καθιστά δύσκολο να αποτιμηθεί η επιτυχία της διαδικασίας. Επίσης, η αποτελεσματικότητα της ανατροφοδότησης ενισχύεται, όταν παρέχεται κατά τη διάρκεια της ενασχόλησης του σπουδαστή με την εργασία. Σε

περιβάλλοντα ψηφιακής μάθησης, υπάρχουν εργαλεία τα οποία δίνουν τη δυνατότητα για επαναλαμβανόμενες υποβολές των εργασιών. Διευκολύνουν επίσης την άμεση ανατροφοδότηση κατά την ενασχόληση του σπουδαστή με την εργασία (χαρακτηριστικό παράδειγμα αποτελούν οι προσομοιώσεις), δίνοντας τους την ευκαιρία να προβεί σε άμεσες επανορθωτικές κινήσεις (Nicol & Milligan, 2006).

Η *έβδομη αρχή* αφορά την παροχή πληροφοριών που βοηθούν τον διδάσκοντα να προσαρμόσει τη διδασκαλία. Η ανατροφοδότηση δεν βελτιώνει μόνο την απόδοση των σπουδαστών, αλλά παρέχει στους διδάσκοντες πολύτιμες πληροφορίες σχετικές με την πρόοδο των σπουδαστών. Οι πληροφορίες είναι απαραίτητες για να γίνουν κατάλληλες προσαρμογές στις μεθόδους, στο ρυθμό και στους στόχους της εκπαιδευτικής διαδικασίας. Συχνές διαμορφωτικές ψηφιακές αξιολογήσεις, που εύκολα μπορούν υλοποιηθούν με χρήση ψηφιακών μέσων παρέχουν συνεχείς και συσσωρευόμενες πληροφορίες για την μαθησιακή πορεία των σπουδαστών. Για παράδειγμα, στα ψηφιακά διαγωνίσματα, η μεγάλη συχνότητα λανθασμένων απαντήσεων σε συγκεκριμένες ερωτήσεις, πρέπει να οδηγήσει σε διδακτικές παρεμβάσεις για να θεραπευθεί η υστέρηση που διαγνώστηκε μέσω του διαγωνίσματος. Είναι δυνατόν να διαγνωστούν σπουδαστές με συνολική απόδοση που υπολείπεται κατά πολύ έναντι του μέσου όρου της τάξης, οπότε θα πρέπει να τύχουν σχετικής υποστήριξης. Πολύτιμη πηγή πληροφοριών αποτελούν και οι διαδικτυακοί χώροι δημόσιων συζητήσεων, όπου και εκεί ο διδάσκοντας είναι δυνατόν να εντοπίσει μαθησιακές δυσκολίες, ώστε να αναπροσαρμόσει την εκπαιδευτική πρακτική (Nicol & Milligan, 2006).

Συμπερασματικά, οι Nicol & Milligan (2006), αναφέρουν ότι τα τεχνολογικά εργαλεία που χρησιμοποιούνται στις διάφορες μορφές ψηφιακής αξιολόγησης είναι περισσότερο αποτελεσματικά αν συνδυαστούν με πρακτικές που ενισχύουν την ανατροφοδότηση και βοηθούν τους σπουδαστές να μάθουν στα πλαίσια αυτόνομων και αυτορρυθμιζόμενων πρακτικών μάθησης.

## 4. ΜΟΡΦΕΣ ΨΗΦΙΑΚΗΣ ΑΞΙΟΛΟΓΗΣΗΣ

Βασικές μορφές ψηφιακής αξιολόγησης είναι οι αξιολογήσεις κλειστού τύπου, οι αξιολογήσεις ανοικτού τύπου, οι ψηφιακοί φάκελοι επιτευγμάτων, η συνεργατική ανάπτυξη περιεχομένου, η διαδικτυακοί χώροι δημόσιου διαλόγου, τα σοβαρά παιχνίδια, οι προσομοιώσεις και η Μαθησιακή Αναλυτική (Oldfield, et al., 2012. Jordan, 2013. Redecker, 2013. Kapsalis, et al., 2019). Ακολουθεί η περιγραφή καθεμιάς από τις μορφές, εκτός από τη Μαθησιακή Αναλυτική, η οποία περιγράφεται σε ξεχωριστό κεφάλαιο.

### 4.1. Ψηφιακές αξιολογήσεις κλειστού τύπου

Ο όρος «αξιολογήσεις κλειστού τύπου» περιγράφει αξιολογήσεις με ερωτήσεις πολλαπλής επιλογής, ή αντιστοίχισης ή κατάταξης. Ο αξιολογούμενος καλείται να απαντήσει σε ερώτηση, επιλέγοντας την ορθή απάντηση από ένα πεπερασμένο / κλειστό σύνολο διαθέσιμων απαντήσεων (JISC, 2006). Ή καλείται να δημιουργήσει ζεύγη λέξεων ή φράσεων τις οποίες παρέχει το σύστημα. Ή καλείται να κατατάξει σε σωστή σειρά λέξεις ή φράσεις. Σε όλες τις περιπτώσεις, ο αξιολογούμενος δεν επινοεί την απάντηση. Αλλά, αναζητά τη σωστή επιλογή μεταξύ διαθέσιμων πεπερασμένων επιλογών. Το πεπερασμένο, κλειστό σύνολο των επιλογών δίνει το όνομα στον όρο «αξιολογήσεις κλειστού τύπου», ο οποίος χρησιμοποιείται σε αντιδιαστολή με τον όρο «αξιολογήσεις ανοικτού τύπου», όπου το σύνολο των επιλογών είναι μη πεπερασμένο και ανοικτό και όπου ο αξιολογούμενος ανακαλεί από τη μνήμη του ή συνθέτει την απάντηση. Εκτός από τον όρο «αξιολογήσεις κλειστού τύπου», χρησιμοποιείται και όρος «αντικειμενικές αξιολογήσεις», διότι η σωστή απάντηση και ο αντίστοιχος αποδιδόμενος βαθμός έχουν καθοριστεί εκ των προτέρων, κατά το σχεδιασμό της ερώτησης. Οπότε, η βαθμολόγηση είναι αυτόματη, ακαριαία και αντικειμενική, δηλαδή δεν επηρεάζεται από τυχόν υποκειμενικές απόψεις του αξιολογητή. Η χρήση του όρου «αντικειμενική αξιολόγηση» υπογραμμίζει επίσης το γεγονός ότι, με τη χρήση τέτοιου τύπου αξιολογήσεων, επιδιώκεται η αύξηση του βαθμού αντικειμενικότητας της αξιολόγησης.

Οι αξιολογήσεις κλειστού τύπου θεωρούνται ότι είναι ισότιμες με τις αξιολογήσεις ανοικτού τύπου, όπως προκύπτει από τις έρευνες που αναφέρουν οι Azevedo et al. (2019), όπου οι δύο τύποι αξιολόγησης συγκρίνονται βάση του αποτελέσματος και της επίδρασης που έχουν στις

μαθησιακές στρατηγικές. Ειδικότερα, οι Bible, Simkin & Kuechler (2008) διαπίστωσαν ότι η απόδοση των σπουδαστών σε αξιολογήσεις κλειστού τύπου είναι παρόμοια με την απόδοση σε αξιολογήσεις ανοικτού τύπου, γεγονός που σύμφωνα με τους ερευνητές επιτρέπει την με μέτρο αξιοποίησή τους. Η Ferrao (2010) συνέκρινε αποτελέσματα αξιολογήσεων και των δύο τύπων και συμπέρανε ότι μεταξύ τους υφίσταται ισχυρή συσχέτιση, οπότε οι αξιολογήσεις κλειστού τύπου μπορούν να χρησιμοποιηθούν εναλλακτικά των αξιολογήσεων ανοικτού τύπου. Οι Lee, Liu & Linn (2011) συμπέραναν ότι οι αξιολογήσεις ανοικτού τύπου είναι καλύτερες για να αξιολογήσουν την ενοποίηση («intergration») των γνώσεων. Οι Mora, Sancho-Bru, Iserte και Sanchez (2012) δεν διαπίστωσαν σοβαρές διαφορές στα αποτελέσματα που προέκυψαν από τη χρήση αξιολογήσεων και των δύο τύπων. Οι Heron & Lerpiniere (2013) χρησιμοποίησαν αξιολογήσεις και των δύο τύπων προκειμένου να υποστηρίξουν διδακτικές μεθοδολογίες που επιδιώκουν ουσιαστική μάθηση και δεν διαπίστωσαν να υπάρχουν διαφορές μεταξύ των αξιολογήσεων ανοικτού τύπου, έναντι των αξιολογήσεων κλειστού τύπου.

Οι αξιολογήσεις κλειστού τύπου προϋπήρχαν της ψηφιακής αξιολόγησης. Χρησιμοποιήθηκαν μαζικά για πρώτη φορά στις αρχές του εικοστού αιώνα από τον στρατό των Η.Π.Α. για την αξιολόγηση νεοσύλλεκτων στρατιωτών κατά τον Α΄ Παγκόσμιο Πόλεμο (Jordan, 2013). Έκτοτε, η χρήση τους συνεχώς διευρυνόταν και αξιοποιούνταν για εκπαιδευτικούς σκοπούς. Αφενός, διότι τεχνολογικές καινοτομίες, όπως οι οπτικοί αναγνώστες, καθιστούσαν εφικτή την αυτόματη αναγνώριση και βαθμολόγηση των απαντήσεων. Αφετέρου, διότι ήταν αναγκαίο να αντιμετωπιστούν προβλήματα στην αντικειμενικότητα των αξιολογήσεων ανοικτού τύπου, όπου συχνά ανέκυπταν αποκλίσεις μεταξύ των βαθμών που έδιναν διαφορετικοί αξιολογητές για την ίδια απάντηση. Συγκεκριμένα, ενώ η αυτόματη βαθμολόγηση στις αξιολογήσεις κλειστού τύπου χαρακτηρίζεται από συνέπεια, η βαθμολόγηση στις αξιολογήσεις ανοικτού τύπου, όταν διενεργείται από φυσικά πρόσωπα, επηρεάζεται από πολλούς υποκειμενικούς παράγοντες και είναι εγγενώς ασυνεπής. Συχνά, ανακύπτουν αποκλίσεις όχι μόνο στη βαθμολόγηση της ίδιας απάντησης από διαφορετικούς αξιολογητές, αλλά και αποκλίσεις όταν ο ίδιος αξιολογητής βαθμολογεί την ίδια απάντηση σε διαφορετικές χρονικές στιγμές (Jordan, 2013).

Οι αξιολογήσεις κλειστού τύπου αποτελούνται από ερωτήσεις που έχουν δομή δύο μερών: α) το θέμα, το οποίο προηγείται και το οποίο συνήθως έχει τη μορφή ερώτησης, ή ελλιπούς

πρότασης και β) το σύνολο των Επιλογών, το οποίο έπεται και το οποίο αποτελείται από τις απαντήσεις, εκ των οποίων μια τουλάχιστον είναι ορθή και οι υπόλοιπες είναι λανθασμένες (Torres et al., 2011). Σε μια καλά σχεδιασμένη ερώτηση, η ορθή επιλογή πρέπει να είναι αδιαμφισβήτητη. Ταυτοχρόνως, οι λανθασμένες επιλογές πρέπει να κρίνονται πράγματι ως λανθασμένες από όσους γνωρίζουν καλά το γνωστικό αντικείμενο της ερώτησης, ενώ ταυτοχρόνως να είναι όλες αληθοφανείς για όσους δεν γνωρίζουν καλά το γνωστικό αντικείμενο (Azevedo et al., 2019). Ο τύπος που κυριαρχεί ως προς τη μορφή των Επιλογών είναι είτε λέξεις, είτε αριθμητικές παραστάσεις, είτε φράσεις που απαντούν στην ερώτηση ή συμπληρώνουν το κενό της ελλιπούς πρότασης. Ο κυρίαρχος αυτός τύπος είναι γνωστός ως ερώτηση πολλαπλής επιλογής.

**Πίνακας 4 Παράδειγμα ερώτησης πολλαπλής επιλογής (Chamberlain, 2019)**

<b>ΘΕΜΑ</b>	
Βρείτε τη λύση της εξίσωσης: $\frac{7x+6}{2} = \frac{-3x-6}{3} - \frac{-8x-8}{5}$	
<b>ΕΠΙΛΟΓΕΣ</b>	
A. $x = -\frac{40}{29}$	Λανθασμένη επιλογή
B. $x = -\frac{34}{29}$	Ορθή επιλογή
Γ. $x = -\frac{66}{29}$	Λανθασμένη επιλογή
Δ. $x = -\frac{17}{10}$	Λανθασμένη επιλογή

Λιγότερο συχνά, χρησιμοποιούνται τύποι που έχουν τη μορφή λεκτικών ζευγών ή ταξινομημένης λίστας (ή σχετικών παραλλαγών), όπου ο αξιολογούμενος συνθέτει πάντοτε από τις διαθέσιμες επιλογές. Δηλαδή, δεν έχει τη δυνατότητα να προτείνει άλλες απαντήσεις, πέραν από αυτές που προτείνει το σύστημα.

Οι βασικοί τύποι των ερωτήσεων πολλαπλής επιλογής καταγράφονται στον παρακάτω Πίνακα 5 (Azevedo et al., 2019):

*Πίνακας 5. Μορφές Ερωτήσεων Πολλαπλής Επιλογής (Azevedo et al. 2019)*

Τύπος	Περιγραφή
Συμβατική πολλαπλή επιλογή	Το θέμα ακολουθούν πολλές επιλογές, εκ των οποίων μία είναι η ορθή.
Εναλλακτικές επιλογές	Το θέμα ακολουθούν δύο επιλογές, εκ των οποίων μία είναι η ορθή.
Σωστό / Λάθος	Οι διαθέσιμες επιλογές είναι δύο. Η μία επιβεβαιώνει και η άλλη διαψεύδει την ορθότητα θέματος.
Πολλαπλές επιλογές με Σωστό / Λάθος	Το θέμα ακολουθούν πολλές επιλογές για την καθεμιά εκ των οποίων δηλώνεται η ορθότητα ή μη ορθότητα της.
Σύζευξη	Προσφέρονται δύο λίστες από επιλογές και ζητείται να κατασκευαστούν ζεύγη, συνδυάζοντας συναφείς επιλογές από τις δύο λίστες. Το πλήθος των επιλογών της μιας λίστας δεν είναι απαραίτητως το ίδιο, με το πλήθος των επιλογών της άλλης λίστας.
Σύνθετη πολλαπλή επιλογή	Το θέμα ακολουθούν επιλογές που είναι οργανωμένες σε σύνολα προτάσεων, εκ των οποίων το ένα είναι το ορθό.



Σύνολο από σχετιζόμενα ζητήματα	Το θέμα αποτελεί εισαγωγή σε ζήτημα στο οποίο αναφέρονται οι συμβατικής μορφής ερωτήσεις πολλαπλής επιλογής που ακολουθούν.
---------------------------------	-----------------------------------------------------------------------------------------------------------------------------

#### 4.1.1. Πλεονεκτήματα, Προκλήσεις

Οι αξιολογήσεις κλειστού τύπου και ιδιαίτερα οι ερωτήσεις πολλαπλής επιλογής αποτελούν την περισσότερο διαδεδομένη μορφή ψηφιακής αξιολόγησης (Stödberg, 2012). Σύμφωνα με έρευνα που διενήργησε ο Hunt (2012), μεταξύ 50'000'000 ερωτήσεων που εντόπισε σε 2'500 Συστήματα Διαχείρισης Μάθησης (ΣΔΜ), το 90% ήταν κλειστού τύπου. Η κυριαρχία των αξιολογήσεων κλειστού τύπου οφείλεται εν πολλοίς στα πλεονεκτήματά τους, τα οποία καταγράφουν οι Torres et al. (2011) και συνοψίζουν οι Azevedo et al. (2019) ως εξής: α) μπορούν να χρησιμοποιηθούν σε πολλά γνωστικά αντικείμενα, β) είναι κατάλληλες για διαφορετικά είδη εκπαίδευσης (τυπικής και μη τυπικής) γ) προσαρμόζονται σε πολλά επίπεδα νοητικών επιπέδων, δ) διευκολύνουν την αξιολόγηση πολυπληθών τμημάτων, ε) διευκολύνουν την αυτόματη βαθμολόγηση και τη στατιστική ανάλυση των αποτελεσμάτων, στ') αποτελούν την καταλληλότερη μορφή αξιολόγησης για τη διαχρονική σύγκριση αποτελεσμάτων.

Ωστόσο, υφίστανται και περιορισμοί όπως (Azevedo et al., 2019): α) ο σχεδιασμός ερωτήσεων που αξιολογούν ανώτερες νοητικές ικανότητες είναι δύσκολος, β) ο σχεδιαστής των ερωτήσεων πρέπει να είναι καλός χειριστής του γραπτού λόγου, ώστε οι ερωτήσεις να είναι σαφείς, γ) οι αξιολογούμενοι πρέπει να διαθέτουν δεξιότητες κατανόησης γραπτού λόγου για να μην παρερμηνεύουν τις ερωτήσεις, δ) δεν είναι εφικτό να αξιολογήσουν δεξιότητες όλων των τύπων, όπως για παράδειγμα τις επικοινωνιακές δεξιότητες, ε) συχνά, είναι δύσκολο θα σχεδιασθούν λανθασμένες, πλην όμως αληθοφανείς επιλογές, στ') συχνά, οι αξιολογούμενοι επιλέγουν τυχαία απάντηση. Η Jordan (2013) επεκτείνει το σύνολο των περιορισμών, προσθέτοντας και τα εξής: α) η ορθή απάντηση είναι δυνατόν να αποκαλυφθεί με αντίστροφα νοητικά βήματα, οπότε η ερώτηση δεν καταφέρνει να αξιολογήσει τη νοητική δεξιότητα, που επιδιώκει να αξιολογήσει (για παράδειγμα μια ερώτηση που ζητά το ολοκλήρωμα μιας συνάρτησης, μπορεί να απαντηθεί με τον υπολογισμό της παραγώγου των διαθέσιμων επιλογών), β) στις ερωτήσεις που απαιτούν

αριθμητικό – αλγεβρικό υπολογισμό, αν η απάντηση στην οποία καταλήγει ο αξιολογούμενος δεν βρίσκεται μεταξύ των επιλογών, τότε είναι φανερό ότι υπήρξε λάθος στον υπολογισμό και ο αξιολογούμενος μπορεί να επαναλάβει τους υπολογισμούς, έως ότου καταλήξει σε απάντηση που περιλαμβάνεται στις επιλογές, γ) δεν είναι πάντοτε εφικτή η αναπαράσταση των λύσεων ενός αυθεντικού προβλήματος με ένα κλειστό σύνολο επιλογών. Η Jordan (2013) θεωρεί ως αιτία των αμφιβολιών για την αποτελεσματικότητα των ερωτήσεων κλειστού τύπου, όχι τους αναφερόμενους περιορισμούς, αλλά κυρίως την ανομοιογένεια των υπό μελέτη ερωτήσεων. Αναφέρει ότι οι διαφορετικές ερωτήσεις απαιτούν διαφορετική σχεδίαση και επιπλέον επισημαίνει τη σημασία της ποιότητας και της τήρησης σχεδιαστικών αρχών κατά την κατασκευή των ερωτήσεων. Ο Burton (2005) συμπληρώνει ότι κάποια διαγωνίσματα, συμπεριλαμβανομένων των ερωτήσεων και των μεθόδων βαθμολόγησης, κρίνονται ως αναξιόπιστα, πιθανόν εξαιτίας σχεδιαστικών ελαττωμάτων και άστοχων αποφάσεων κατά τον σχεδιασμό τους.

#### 4.1.2. Αρχές σχεδιασμού καλών ερωτήσεων

Οι προαναφερόμενοι περιορισμοί και ιδιαίτερα τα ζητήματα που αφορούν α) το σχεδιασμό «καλών» ερωτήσεων και β) το περιορισμό του φαινομένου επιλογής τυχαίων απαντήσεων, αποτελούν βασικό αντικείμενο στη σχετική βιβλιογραφία. Ως προς το ζήτημα του σχεδιασμού, σχετικές κατευθυντήριες οδηγίες περιλαμβάνονται στις εργασίες των Clegg & Cashin (1986), των Burton, Sudweeks, Merrill & Wood (1991), των Camilo & Silva (2008), των Haladyna, Downing & Rodriguez (2002). Οι Azevedo et al. (2019) θεωρούν πληρέστερη την εργασία των Haladyna, Downing & Rodriguez (2002), οι οποίοι διενέργησαν μια εξαντλητική και συστηματική ανάλυση σχετικών ερευνητικών εργασιών, που απέφερε ένα σύνολο 31 οδηγιών για τη σχεδίαση «καλών» ερωτήσεων. Οι οδηγίες κατανέμονται σε πέντε κατηγορίες, οι οποίες αφορούν το γνωστικό αντικείμενο, τη μορφή, τη σύνταξη, την κατασκευή του θέματος και την κατασκευή του συνόλου των επιλογών. Μεταξύ άλλων, προτείνεται: Η παρουσίαση της βασικής ιδέας στο θέμα και όχι στις επιλογές της ερώτησης. Η παράφραση των προτάσεων που αντλούνται από τα βιβλία, ώστε να αποθαρρύνεται η μηχανική απομνημόνευση. Η χρήση κατανοητών προτάσεων, χωρίς αμφίσημες, δυσνόητες λέξεις, φράσεις, έννοιες. Η χρήση σύντομων, λιτών προτάσεων, χωρίς εντυπωσιοθηρικά λεκτικά σχήματα. Η χρήση καταφατικών προτάσεων και ο περιορισμός

αρνητικών προτάσεων (αποφυγή του συνδέσμου «δεν») – ιδίως, η αποφυγή διπλής άρνησης, η οποία είναι καλό να αντικαθίσταται από κατάφαση. Η χρήση πολλών (κατά προτίμηση 3 ή περισσότερων) επιλογών. Η χρήση επιλογών που είναι ισότιμα αληθοφανείς. Η εναλλαγή στη θέση όπου τοποθετείται η ορθή επιλογή (π.χ. η ορθή επιλογή να μην τοποθετείται πάντα ή σχεδόν πάντα στην πρώτη θέση). Η αποφυγή, ή τουλάχιστον η προσεκτική, χρήση των επιλογών «*όλα τα παραπάνω*» και «*κανένα από τα παραπάνω*» κ.α.

#### 4.1.3. Τυχαίες απαντήσεις

Ως προς το ζήτημα της επιλογής τυχαίας απάντησης, οι Azevedo et al. (2019) αναφέρουν ότι κάθε φορά που ο αξιολογούμενος καλείται να απαντήσει, α) είτε γνωρίζει την ορθή απάντηση και την επιλέγει, β) είτε γνωρίζει ότι κάποιες από τις επιλογές είναι λανθασμένες και τις αποκλείει, γ) είτε δεν γνωρίζει οτιδήποτε και επιλέγει με τυχαίο τρόπο. Στην τελευταία περίπτωση, η πιθανότητα να επιλέξει ορθή απάντηση σε μια ερώτηση είναι αντιστρόφως ανάλογη του πλήθους των επιλογών (π.χ. 25% αν οι επιλογές είναι τέσσερις, 20% αν οι επιλογές είναι πέντε κ.ο.κ.). Οι εν λόγω πιθανότητες είναι υψηλές και δεν μπορεί να θεωρηθούν ανεκτές, όμως μειώνονται προοδευτικά και δραστικά, όσο ο αξιολογούμενος συνεχίζει να επιλέγει τυχαίες απαντήσεις και σε επόμενες ερωτήσεις. Για παράδειγμα, οι πιθανότητες μειώνονται στο 1,56% για ορθές απαντήσεις σε ένα διαγώνισμα με τρεις ερωτήσεις (που περιλαμβάνουν τέσσερις επιλογές) και στο 0,39% για ορθές απαντήσεις σε ένα διαγώνισμα με τέσσερις ερωτήσεις (που περιλαμβάνουν τέσσερις επιλογές). Βέβαια, οι αναφερόμενες πιθανότητες αφορούν περιπτώσεις όπου ο αξιολογούμενος επιδιώκει την άριστη επίδοση (100%). Συνεπώς, είναι περισσότερο ρεαλιστικό να υπολογιστούν οι πιθανότητες για την επίτευξη όχι άριστης επίδοσης, αλλά απλής αποδεκτής επίδοσης (τουλάχιστον 50%). Σε διαγώνισμα δέκα ερωτήσεων, η πιθανότητα επίτευξης αποδεκτής επίδοσης είναι περίπου 8% και σε διαγώνισμα είκοσι ερωτήσεων είναι περίπου 1% (Azevedo et al. 2019, σ. 14). Συνεπώς, όσο αυξάνεται το πλήθος των ερωτήσεων και το πλήθος των επιλογών, τόσο μειώνεται (και μάλιστα εκθετικά) η πιθανότητα αποδεκτής επίδοσης σε διαγωνίσματα πολλαπλών επιλογών. Οπότε, ο Downing (2003) υποστηρίζει ότι η επίδραση της τυχαίας επιλογής είναι τελικά ασήμαντη στα διαγωνίσματα με πολλές ερωτήσεις και δεν πρέπει να αποτελεί αντικείμενο προβληματισμού, δεδομένου ότι είναι εξαιρετικά δύσκολο για έναν αξιολογούμενο να επιτύχει αποδεκτή επίδοση,

αν οι επιλογές του βασίζονται μόνο στην τύχη. Αλλά, ο Burton (2005) αναφέρει ότι εύστοχες τυχαίες επιλογές μπορούν εν δυνάμει να δημιουργήσουν κρίσιμες διαφορές στη βαθμολογία αξιολογούμενων, των οποίων ο βαθμός κινείται κοντά στο όριο της αποδεκτής επίδοσης.

#### 4.1.4. Μέθοδοι βαθμολόγησης

Προκειμένου να αποθαρρυνθούν όσοι επιλέγουν τυχαίες απαντήσεις, έχουν προταθεί διάφορες μέθοδοι βαθμολόγησης στις ερωτήσεις πολλαπλής επιλογής. Η απλούστερη μέθοδος είναι η αρνητική βαθμολογία, δηλαδή η αφαίρεση βαθμού (ακέραιου ή κλασματικού) στις περιπτώσεις που ο αξιολογούμενος επιλέγει λανθασμένη απάντηση. Οι Azevedo et al. (2019) θεωρούν ότι η συγκεκριμένη στρατηγική μειώνει την πιθανότητα επίτευξης αποδεκτής επίδοσης για όσους επιλέγουν απαντήσεις με τυχαίο τρόπο, ωστόσο ο Brown (2001) υποστηρίζει ότι ευνοεί τους αξιολογούμενους που είναι διατεθειμένοι να ρισκάρουν.

Ο Bush (2015) αναφέρει ότι οι αξιολογούμενοι που ρέπουν περισσότερο στην τυχαία επιλογή είναι όσοι έχουν μειωμένες προσδοκίες για επίτευξη υψηλού βαθμού. Παράλληλα, όσοι νιώθουν σχετική αβεβαιότητα για την ορθή απάντηση, αλλά διαθέτουν μερική γνώση επί του αξιολογούμενου ζητήματος, θεωρούν ότι η αρνητική βαθμολογία είναι άδικη. Διότι τους αποτρέπει να επιλέξουν την απάντηση που θεωρούν πιθανότερη. Ωστόσο, ο Bush (2015) υποστηρίζει ότι η αρνητική βαθμολογία είναι απαραίτητη προκειμένου να ενισχυθεί η αξιοπιστία της βαθμολογικής μεθόδου και προτείνει τα παρακάτω εναλλακτικά βαθμολογικά σχήματα ανάλογα με τους παρακάτω τύπους ερωτήσεων και απαντήσεων:

1. Παραδοσιακός τύπος, με επιλογή μόνο των ορθών απαντήσεων.
2. Επιλογή συνόλου πιθανών ορθών απαντήσεων.
3. Απόρριψη λανθασμένων απαντήσεων.
4. Ταξινόμηση απαντήσεων με φθίνουσα πιθανότητα ορθότητας.
5. Επαναλαμβανόμενες προσπάθειες, έως τον εντοπισμό της ορθής απάντησης.
6. Επαναλαμβανόμενες προσπάθειες απόρριψης λανθασμένων απαντήσεων.
7. Ταξινόμηση απαντήσεων με φθίνουσα πιθανότητα ορθότητας, με δυνατότητα για απόδοση ίδιας πιθανότητας σε διαφορετικές απαντήσεις.
8. Επαναλαμβανόμενες προσπάθειες για επιλογή συνόλου πιθανών ορθών απαντήσεων.

Τα βαθμολογικά σχήματα του Bush (2015) διαθέτουν πρωτοτυπία, ποιότητα, και προσαρμοστικότητα. Ωστόσο, το τίμημα είναι η δυσκολία στην κατασκευή των ερωτήσεων και συχνά η δυσκολία στην κατανόησή τους από τους αξιολογούμενους (Azevedo et al. 2019). Τρία από τα σχήματα που προτείνει ο Bush (2015) -συγκεκριμένα τα υπ' αριθμόν 5, 6, 8 - επιτρέπουν στον αξιολογούμενο να επαναλάβει την προσπάθειά του στην περίπτωση που η αρχική απάντησή του είναι λανθασμένη. Σε αυτές τις περιπτώσεις, ο βαθμός που απονέμεται είναι αντιστρόφως ανάλογος των προσπαθειών που απαιτήθηκαν για την ορθή απάντηση. Τα τρία σχήματα που επιτρέπουν επαναλαμβανόμενες προσπάθειες είναι κατάλληλα για διαμορφωτική αξιολόγηση, δεδομένου ότι οι αξιολογούμενοι τελικά ανακαλύπτουν ποια είναι η σωστή απάντηση. Αντιθέτως, τα σχήματα που δεν επιτρέπουν επαναλήψεις, δεν προσφέρουν κάποια πληροφορία στους αξιολογούμενους για το ποια είναι η ορθή απάντηση, παρά μόνο με την ολοκλήρωση της δοκιμασίας.

Γενικά, τα βαθμολογικά σχήματα που προτείνει ο Bush (2015) επιδιώκουν να καταστήσουν το διαγώνισμα «ουδέτερο» έναντι των τυχαίων επιλογών. Δηλαδή, οι εύστοχες τυχαίες επιλογές εξουδετερώνονται βαθμολογικά από τις άστοχες τυχαίες επιλογές, συνεπώς, η στρατηγική της τυχαίας επιλογής, κατά μέσο όρο, δεν επιφέρει όφελος σε όσους την υιοθετούν, αλλά ούτε προκαλεί ζημιά σε όσους την αποφεύγουν. Για παράδειγμα, σε ερώτηση παραδοσιακού τύπου με τέσσερις επιλογές όπου, η πιθανότητα άστοχης τυχαίας επιλογής είναι τρεις φορές μεγαλύτερη από την πιθανότητα εύστοχης τυχαίας επιλογής, ο Bush (2015) προτείνει την απόδοση τριών μονάδων για την ορθή απάντηση και την αφαίρεση μιας μονάδας για την λανθασμένη απάντηση. Ο Bush (2015) επίσης διαχωρίζει την «ενημερωμένη τυχαία επιλογή» (informed guesswork) από την «αμιγώς τυχαία επιλογή» (pure guesswork). Κατά την «ενημερωμένη τυχαία επιλογή», ο αξιολογούμενος, παρότι δεν γνωρίζει την ορθή απάντηση, θεωρεί ότι κάποιες επιλογές φαίνονται να είναι σωστότερες από άλλες. Οπότε, επιλέγοντας τυχαία μόνο μεταξύ αυτών, αυξάνει τις πιθανότητες εύστοχης απάντησης. Κατά την «αμιγώς τυχαία επιλογή», ο αξιολογούμενος δεν εντοπίζει κάποιες επιλογές που να φαίνονται λιγότερο ή περισσότερο ορθές, οπότε επιλέγει τυχαία από το αρχικό σύνολο των επιλογών.

Οι Ventouras, Triantis, Tsiakas & Stergiopoulos (2010) προτείνουν μέθοδο βαθμολόγησης κατά την οποία, για κάθε ερώτηση εντός του διαγωνίσματος υπάρχει δεύτερη ερώτηση που αξιολογεί

ακριβώς το ίδιο ζήτημα. Οι ερωτήσεις που συνιστούν το ζεύγος, φαινομενικά είναι διαφορετικές, δηλαδή έχουν διαφορετική μορφή και σύνταξη και δεν εμφανίζονται διαδοχικά. Οπότε, οι αξιολογούμενοι που δεν κατέχουν επαρκώς το γνωστικό αντικείμενο είναι δύσκολο να τις αναγνωρίσουν ως ίδιες, παρότι είναι ουσιαστικά ίδιες. Οι αξιολογούμενοι, που απαντούν ορθά και στις δύο ερωτήσεις του ζεύγους, επιβραβεύονται βαθμολογικά. Αντιθέτως, οι αξιολογούμενοι, που απαντούν ορθά στη μία ερώτηση και λανθασμένα στην άλλη, υφίστανται βαθμολογική ποινή. Με αυτή τη μέθοδο οι Ventouras et al. (2010) θεωρούν ότι απαλείφεται η στρέβλωση που προκαλεί η τυχαία επιλογή.

#### 4.1.5. Βαθμολόγηση βάση βεβαιότητας

Διαφορετική προσέγγιση ως προς τη μέθοδο βαθμολόγησης ακολουθεί ο Gardner-Medwin (2006, 2019) ο οποίος προτείνει τη χρήση θετικών και αρνητικών συντελεστών βαρύτητας, ανάλογων με το επίπεδο «βεβαιότητας» (certainty) στην απάντηση του αξιολογούμενου. Στη μέθοδο αρχικά αποδόθηκε το όνομα «*βαθμολόγηση βάση πεποίθησης*» (confidence based marking), αλλά τελικά ο Gardner-Medwin (2019) κατέληξε στο όνομα «*βαθμολόγηση βάση βεβαιότητας*» (certainty based marking). Διότι όπως αναφέρει, στόχος δεν είναι η ενίσχυση ή η επιβράβευση της αυτοπεποίθησης του αξιολογούμενου, αλλά ο εντοπισμός της αβεβαιότητας στην κατοχή της γνώσης.

Σύμφωνα με τον Gardner-Medwin (2019), κάθε απάντηση χρήζει περαιτέρω διερεύνησης, για αυτό και η μέθοδος απαιτεί από τον αξιολογούμενο να αναζητήσει τεκμηρίωση για την απάντησή του και να στοχαστεί για τυχόν αμφιβολίες του. Οι συγκεκριμένες νοητικές λειτουργίες δεν απαιτούνται στις παραδοσιακές αξιολογήσεις κλειστού τύπου, στις οποίες οι αξιολογούμενοι ανταποκρίνονται βασιζόμενοι σε επιφανειακές γνώσεις και δεν έχουν κίνητρα για να στοχαστούν ως προς την αξιοπιστία των γνώσεων τους.

Η μέθοδος διασφαλίζει ότι ο αξιολογούμενος, προκειμένου να επιτύχει τη βέλτιστη βαθμολογία, πρέπει να διαχωρίσει τις απαντήσεις για τις οποίες είναι βέβαιος, δηλαδή τις απαντήσεις που βασίζονται σε στέρεη, καλά εμπεδωμένη γνώση, από τις απαντήσεις για τις οποίες έχει αμφιβολία. Οι απαντήσεις που δίνονται με βεβαιότητα, βαθμολογούνται με υψηλά θετικό βαθμό αν είναι ορθές, αλλά με χαμηλά αρνητικό βαθμό αν είναι λανθασμένες. Αντιστοίχως, οι απαντήσεις που δίνονται με δισταγμό, βαθμολογούνται με χαμηλά θετικό βαθμό αν είναι ορθές

και με χαμηλά αρνητικό βαθμό ή με μηδενικό βαθμό (δηλαδή χωρίς ποινή), αν είναι λανθασμένες. Ως αποτέλεσμα, ανταμείβονται οι αξιολογούμενοι που «... αναγνωρίζουν τη βασιμότητα της αιτιολόγησης ή της αμφιβολίας τους και όχι αυτοί που σταθερά είναι βέβαιοι ή σταθερά αβέβαιοι» (Gardner-Medwin, 2016, σ. 141).

Στον Πίνακα 6 παρουσιάζεται το βαθμολογικό σχήμα της μεθόδου. Βασίζεται σε κλίμακα τριών επιπέδων βεβαιότητας: χαμηλό, μέτριο, υψηλό. Οι βαθμοί για την ορθή απάντηση είναι: 1, για χαμηλό επίπεδο βεβαιότητας, 2, για μέτριο επίπεδο βεβαιότητας, 3, για υψηλό επίπεδο βεβαιότητας. Αντίστοιχη του επιπέδου βεβαιότητας είναι και η αρνητική βαθμολογία (ποινή) για την λανθασμένη απάντηση: καμία ποινή για χαμηλό επίπεδο βεβαιότητας, -2, για μέτριο επίπεδο βεβαιότητας, -6, για υψηλό επίπεδο βεβαιότητας. Ο αξιολογούμενος δεν υπόκειται σε ποινή αν δεν απαντήσει στην ερώτηση.

*Πίνακας 6. Βαθμολογικό σχήμα στη βαθμολόγησης βάση βεβαιότητας (Gardner-Medwin, 2006, σ. 142)*

<b>Βαθμολογικό Σχήμα στη Βαθμολόγηση Βάση Βεβαιότητας</b>				
	<b>Επίπεδο Βεβαιότητας</b>			<b>Αναπάντητη ερώτηση</b>
	<b>C=1 (χαμηλό)</b>	<b>C=2 (μέτριο)</b>	<b>C=3 (υψηλό)</b>	
Βαθμός για ορθή απάντηση	1	2	3	0
Ποινή για λανθασμένη απάντηση	0	-2	-6	0

Είναι δυνατόν να υπάρξουν συνθετότερες παραλλαγές στο παραπάνω βαθμολογικό σχήμα, οι οποίες περιλαμβάνουν ευρύτερη κλίμακα επιπέδων βεβαιότητας και διαφορετικούς βαθμούς και ποινές. Ωστόσο, σύμφωνα με τον Gardner-Medwin (2006), το προτεινόμενο από αυτόν σχήμα, παρότι σχεδιάστηκε και εφαρμόστηκε για ερωτήσεις δύο επιλογών (σωστό / λάθος), υπερτερεί κάθε παραλλαγής, λόγω της απλότητας και της διαφάνειάς του. Αναφέρει επίσης, ότι προκειμένου να επιτευχθούν μαθησιακά οφέλη, η μέθοδος πρέπει να συνοδεύεται από

ανατροφοδότηση σχετικά με τις ορθές απαντήσεις και με ανάλυση για το ποσοστό επιτυχίας σε κάθε επίπεδο βεβαιότητας.

Ο αξιολογούμενος που βαθμολογείται με τη συγκεκριμένο μέθοδο, συνυπολογίζει ότι για να επιτύχει την μέγιστη βαθμολογία, πρέπει να είναι σε θέση να τεκμηριώσει την απάντηση του, ώστε να κατοχυρώσει την ορθότητά της. Ειδάλλως, κινδυνεύει να υποστεί βαθμολογική απώλεια. Συνεπώς, δεν μπορεί να βασίζεται σε γνώση που έχει απλώς αποστηθίσει, αλλά μόνο στη γνώση που έχει κατανοήσει και εμπεδώσει. Γνωρίζοντας, ότι έχει τη δυνατότητα να λάβει θετικό βαθμό ακόμα και για αμφίβολες επιλογές, στοχάζεται επί της αμφιβολίας του και προσαρμόζει κατάλληλα το επίπεδο βεβαιότητας που θα δηλώσει. Αναγνωρίζει, ότι μια τυχαίως εύστοχη απάντηση δεν συνεπάγεται την κατοχή γνώσης και αποδέχεται χαμηλότερο βαθμό για τις απαντήσεις που σχετίζονται με χαμηλό και μέτριο επίπεδο βεβαιότητας. Αιφνιδιάζεται όταν διαπιστώνει ότι η απάντηση που έδωσε με μεγάλη βεβαιότητα είναι λανθασμένη, αναζητά την αιτία του λάθους και επανέρχεται στη μελέτη του σχετικού ζητήματος.

Η μέθοδος βασίζεται στην ιδέα ότι η γνώση έχει διαβαθμίσεις, δηλαδή δεν υφίσταται πάντα το δίπολο «γνωρίζω - δεν γνωρίζω». Συχνά, ο σπουδαστής είναι δύσκολο να ανακαλέσει τη γνώση, αλλά αναγνωρίζει την ορθότητά κάποιου ισχυρισμού ή κάποιου γεγονότος, όταν κάποιος τρίτος του το παρουσιάσει. Συχνά επίσης, γνωρίζει κάτι, αλλά δεν το κατανοεί πλήρως. Γενικά, οι επιφανειακές γνώσεις, η απουσία νοητικών συνδέσεων μεταξύ συναφών γνώσεων, η έλλειψη επίγνωσης για το βαθμό και την ποιότητα κατανόησης ενός γνωστικού ζητήματος, αποτελούν σημαντικά εμπόδια στην στέρεη οικοδόμηση της γνώσης. Η μέθοδος βοηθά τον σπουδαστή να αναπτύξει τις συνδέσεις και τις στρατηγικές, στις οποίες βασίζονται η ανάκληση και η κατανόηση. Καλλιεργεί τις δεξιότητες για τη στάθμιση της βεβαιότητας και της αμφιβολίας και για τον συνεχή έλεγχο της ορθότητας μιας ιδέας. Εκπαιδεύει τους σπουδαστές να σκέφτονται με κριτικό τρόπο, να εντοπίζουν τις γνωστικές αδυναμίες τους και να συνάπτουν συνδέσεις μεταξύ διαφορετικών, αλλά συναφών γνωστικών στοιχείων. Ενισχύει τη μάθηση και επιπλέον επιτυγχάνει δίκαιη αξιολόγηση. Συγκεκριμένα, ανταμείβει την κατοχή της γνώσης, αλλά επίσης ανταμείβει, έστω μερικώς, την ατελή κατοχή της γνώσης, υπό την προϋπόθεση ότι υφίσταται επίγνωση της ατέλειας. Δεδομένου ότι, η πλάνη είναι χειρότερη από την άγνοια, δεν «τιμωρεί» την άγνοια, αλλά «τιμωρεί» την πλάνη, ιδίως αν η πλάνη συνοδεύεται από βεβαιότητα. Γενικά,



ανταμείβει την ακριβή αποτίμηση της αξιοπιστίας (Gardner-Medwin, 2006, 2019) και ενισχύει τη ενεργή συμμετοχή του αξιολογούμενου στη διαδικασία (Boitshwarelo et al., 2017).

Ο Rosewell (2011) αξιοποιεί τη βασική ιδέα της βαθμολόγησης βάση βεβαιότητας, προκειμένου να προσδώσει χαρακτηριστικά αξιολογήσεων ανοικτού τύπου στις αξιολογήσεις κλειστού τύπου. Η μέθοδος που προτείνει είναι όμοια με τη βαθμολόγηση βάση βεβαιότητας ως προς τη δομή. Η ερώτηση αποτελείται από «Θέμα» και «Επιλογές», ενώ ως απάντηση καταχωρείται μια από τις επιλογές, μαζί με δήλωση για το επίπεδο βεβαιότητας. Αλλάζει όμως η αλληλουχία στα βήματα της διαδικασίας. Συγκεκριμένα, η διαδικασία αποτελείται από δυο στάδια. Στο πρώτο στάδιο εμφανίζεται μόνο το «Θέμα» (όχι όμως οι επιλογές) και ζητείται η δήλωση του επιπέδου βεβαιότητας. Στο δεύτερο στάδιο, ζητείται να καταχωρηθεί μια από τις επιλογές. Με αυτόν τον τρόπο, κατά το πρώτο στάδιο, ο αξιολογούμενος στοχάζεται επί του θέματος, αναζητά (αλλά δεν καταχωρεί) την ορθή απάντηση και δηλώνει το επίπεδο βεβαιότητάς του. Το γεγονός ότι δεν έχουν αποκαλυφθεί οι επιλογές, τον αναγκάζει να ανακαλέσει εξολοκλήρου από τη μνήμη του ή να συνθέσει την απάντηση, όπως θα έκανε σε αξιολόγηση ανοικτού τύπου. Αυτό που καταχωρεί κατά το πρώτο στάδιο είναι το επίπεδο βεβαιότητάς του. Στο δεύτερο στάδιο, αμέσως μετά την καταχώρηση του επιπέδου βεβαιότητας, εμφανίζονται οι επιλογές και ο αξιολογούμενος επιλέγει μια από αυτές ως απάντησή του. Για την υλοποίηση της μεθόδου δεν απαιτούνται αλλαγές στη δομή των ερωτήσεων. Τεχνικά, οι ερωτήσεις παραμένουν ίδιες. Είναι ερωτήσεις πολλαπλής επιλογής. Η κατασκευή τους και η επαναχρησιμοποίησή τους είναι εύκολη, η βαθμολόγησή τους είναι ακαριαία και αντικειμενική, τα αποτελέσματα διαχειρίζονται εύκολα, οι απαιτούμενες ενέργειες από τους αξιολογούμενους είναι λίγες και απλές. Όμως, με τον διαχωρισμό της διαδικασίας σε δύο στάδια, την απόκρυψη των επιλογών κατά το πρώτο στάδιο, την αποκάλυψη τους κατά το δεύτερο στάδιο (δηλαδή μετά τον στοχασμό επί της ερώτησης), η αξιολόγηση μετατρέπεται, σύμφωνα με τον Rosewell (2011), από κλειστού τύπου σε ανοικτού τύπου. Διότι, οι αξιολογούμενοι κατασκευάζουν νοητικά την απάντησή τους, χωρίς να βοηθηθούν ή να αποπροσανατολιστούν από τις πληροφορίες που εμπεριέχουν οι επιλογές.

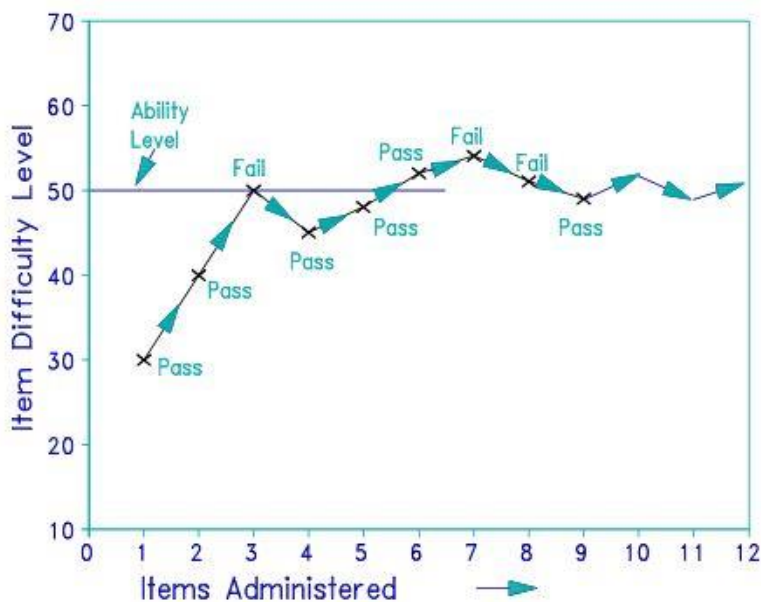
#### 4.1.6. Προσαρμοζόμενα διαγωνίσματα

Σύμφωνα με την Jordan (2013), είναι σημαντικό να δοθεί προσοχή όχι μόνο στα ζητήματα της τυπολογίας των ερωτήσεων και της βαθμολογικής μεθόδου, αλλά και στο ζήτημα του τρόπου

με τον οποίο επιλέγονται οι ερωτήσεις που συγκροτούν το διαγώνισμα. Ενδιαφέρουσα είναι η μέθοδος που ακολουθείται στα λεγόμενα *προσαρμοζόμενα διαγωνίσματα* («adaptive tests»). Δηλαδή, στα διαγωνίσματα όπου το σύστημα αξιολόγησης αποτιμά την ικανότητα του αξιολογούμενου βασιζόμενο στις προηγούμενες απαντήσεις του και προσαρμόζει αναλόγως (αυξομειώνει) το επίπεδο δυσκολίας των επόμενων ερωτήσεων, τις οποίες επιλέγει με κριτήριο το καταλληλότερο για τον αξιολογούμενο επίπεδο δυσκολίας. Προφανώς, για να είναι εφικτή η συγκεκριμένη λειτουργία είναι απαραίτητο να έχει αποδοθεί επίπεδο δυσκολίας σε κάθε ερώτηση. Το επίπεδο δυσκολίας αποδίδεται συνήθως κατά την κατασκευή της ερώτησης, αλλά είναι δυνατόν να επικαιροποιείται με βάση τα δεδομένα που προκύπτουν από τη στατιστική ανάλυση των απαντήσεων που σχετίζονται με την ερώτηση. Με την αξιοποίηση προσαρμοζόμενων διαγωνισμάτων κάθε αξιολογούμενος καλείται να επιλύσει διαγώνισμα, το οποίο είναι κατάλληλο για το επίπεδο των ικανοτήτων του. Βασική αρχή των προσαρμοζόμενων διαγωνισμάτων είναι ότι η αποτελεσματική και ακριβής αξιολόγηση δεν είναι δυνατή, αν χρησιμοποιηθούν εύκολες ερωτήσεις για την αξιολόγηση σπουδαστών με υψηλό επίπεδο ικανότητας και αντιστρόφως, δύσκολες ερωτήσεις για αξιολόγηση σπουδαστών με χαμηλό επίπεδο ικανότητας. Αν το επίπεδο δυσκολίας των ερωτήσεων δεν είναι ανάλογο των ικανοτήτων του σπουδαστή, τότε οι ληφθείσες πληροφορίες είναι χαμηλής αξίας και δεν μπορούν να οδηγήσουν σε αξιολογικά συμπεράσματα. Αντιθέτως, αν το επίπεδο δυσκολίας είναι ανάλογο των ικανοτήτων του σπουδαστή, τότε αυξάνει η αποδοτικότητα και ακρίβεια του διαγωνίσματος, ενώ παράλληλα είναι δυνατόν να μειωθεί και η διάρκειά του. Επιπλέον, η αξιοποίηση προσαρμοζόμενων διαγωνισμάτων αυξάνει το ενδιαφέρον, την αλληλεπίδραση, το κίνητρο και την ενεργή συμμετοχή των σπουδαστών στην αξιολογική διαδικασία. Διότι, ούτε οι ικανοί σπουδαστές πλήττουν, απαντώντας σε εύκολες ερωτήσεις, ούτε οι αδύναμοι σπουδαστές απογοητεύονται και παραιτούνται, αντιμετωπίζοντας δύσκολες ερωτήσεις (Lilley, Barker & Britton, 2004).

Η ποιότητα και η ακρίβεια ενός προσαρμοζόμενου διαγωνίσματος εξαρτάται σημαντικά από τον αλγόριθμο επιλογής των ερωτήσεων. Στην απλούστερη μορφή τους οι αλγόριθμοι επιλογής ερωτήσεων ακολουθούν λογική απλής διακλάδωσης («fixed branching») και λειτουργούν συνοπτικά ως εξής: Κατά την έναρξη του διαγωνίσματος το σύστημα επιλέγει μια τυχαία

ερώτηση μέτριου επιπέδου δυσκολίας. Αν ο αξιολογούμενος απαντήσει ορθά, η εκτίμηση για την ικανότητά του αναβαθμίζεται, οπότε το σύστημα επιλέγει ερώτηση με το αμέσως υψηλότερο επίπεδο δυσκολίας και καλεί τον αξιολογούμενο να την απαντήσει. Αντιστρόφως, αν ο αξιολογούμενος δώσει λανθασμένη απάντηση, η εκτίμηση για την ικανότητά του υποβαθμίζεται και καλείται να απαντήσει στην αμέσως ευκολότερη ερώτηση. Το διαγώνισμα συνεχίζει με τον ίδιο τρόπο. Δηλαδή, κάθε φορά που δίδεται ορθή απάντηση επιλέγεται ως επόμενη ερώτηση αυτή που έχει τον αμέσως μεγαλύτερο βαθμό δυσκολίας και που δεν έχει απαντηθεί. Και κάθε φορά που δίδεται λανθασμένη απάντηση, επιλέγεται ως επόμενη ερώτηση, αυτή με τον αμέσως μικρότερο βαθμό δυσκολίας και που δεν έχει απαντηθεί (Pyper & Willey 2004). Όσο το διαγώνισμα εξελίσσεται, οι συνεχείς αναθεωρήσεις (αυξομειώσεις) στην εκτίμηση της ικανότητας του αξιολογούμενου οδηγούν σε σύγκλιση προς ένα σημείο της βαθμολογικής κλίμακας (Σχήμα 3) το οποίο θεωρείται ότι αποτιμά με σχετική ακρίβεια το επίπεδο ικανότητας του αξιολογούμενου (Linacre, 2000).



Σχήμα 3. Αποτίμηση επιπέδου ικανότητας σε προσαρμοζόμενο διαγώνισμα (Linacre, 2000, σ. 6)

#### 4.1.7. Τράπεζες ερωτήσεων

Τα προσαρμοζόμενα διαγωνίσματα, συχνά όμως και όλα τα υπόλοιπα διαγωνίσματα που χρησιμοποιούν ερωτήσεις κλειστού είτε ανοικτού τύπου, αξιοποιούν «τράπεζες ερωτήσεων» οι οποίες σύμφωνα με τους Azevedo et al. (2019), ενισχύουν την εγκυρότητα και την αξιοπιστία της αξιολογικής διαδικασίας και βοηθούν στην εξοικονόμηση χρόνου και χρήματος. Οι τράπεζες ερωτήσεων είναι ειδικά ψηφιακά αποθετήρια ή βάσεις δεδομένων, όπου οι ερωτήσεις όχι μόνο αποθηκεύονται, αλλά και κατατάσσονται σε κατηγορίες, ανάλογα με το γνωστικό αντικείμενο, το γνωστικό επίπεδο, τον τύπο, τον βαθμό δυσκολίας, τις αξιολογούμενες δεξιότητες ή γνώσεις. Σημαντικός παράγοντας στη χρήση των τραπεζών ερωτήσεων είναι η ποιότητα των ερωτήσεων, για την εξασφάλιση της οποίας είναι απαραίτητη η τήρηση προδιαγραφών κατά τον σχεδιασμό τους και η συνεχής στατιστική ανάλυση των απαντήσεων (π.χ. υψηλά ποσοστά αποτυχίας, συχνές παρανοήσεις), μέσω της οποίας μπορούν να εντοπιστούν ελαττωματικές ερωτήσεις, να βελτιωθούν ή να αντικατασταθούν (Azevedo et al. 2019).

Οι τράπεζες ερωτήσεων βοηθούν στη συστηματική διαχείριση, ταξινόμηση και επαναχρησιμοποίηση των ερωτήσεων, αλλά και στην υλοποίηση μορφών αξιολόγησης, όπως τα προσαρμοσμένα διαγωνίσματα. Βοηθούν επιπλέον στην αντιμετώπιση συμβάντων παραβίασης της ακαδημαϊκής ακεραιότητας και συγκεκριμένα της συνεννόησης μεταξύ των αξιολογούμενων. Συγκεκριμένα, διευκολύνουν την κατασκευή διαγωνισμάτων που είναι εντελώς ή μερικώς διαφορετικά για κάθε αξιολογούμενο. Το σύστημα αξιολόγησης, αντλώντας ερωτήσεις από την τράπεζα ερωτήσεων, κατασκευάζει πολλές διαφορετικές εκδοχές του διαγωνίσματος, όπου η εκδοχή για κάθε αξιολογούμενο είναι εντελώς ή μερικώς διαφορετική από των υπολοίπων. Προφανώς, το σύστημα διασφαλίζει ότι για κάθε εκδοχή, οι ερωτήσεις έχουν τον ίδιο βαθμό δυσκολίας, αντιστοιχούν στο ίδιο γνωστικό επίπεδο, αξιολογούν τις ίδιες δεξιότητες και γνώσεις. Αναιρεί δηλαδή αδικίες που τυχόν προκύπτουν αν οι ερωτήσεις επιλεχθούν με τυχαίο, αλλά άνισο τρόπο (Azevedo et al. 2019). Αν οι εκδοχές δεν είναι εντελώς διαφορετικές για κάθε αξιολογούμενο, αν δηλαδή αντληθούν από την τράπεζα ερωτήσεων λίγες ή πολλές κοινές ερωτήσεις, τότε το σύστημα μπορεί να τοποθετήσει τις κοινές ερωτήσεις σε διαφορετική θέση ανά διαγώνισμα. Επιπλέον, μπορεί να αλλάξει τη σειρά με την οποία εμφανίζονται οι επιλογές των κοινών ερωτήσεων (για τον λόγο αυτό, δεν ενδείκνυται, η χρήση

επιλογών όπως «όλα τα παραπάνω», «κανένα από τα παραπάνω» ή επιλογών που περιλαμβάνουν αυτοαναφορές, π.χ. «το Β και το Γ»).

Επιπλέον, είναι εφικτό, η τράπεζα ερωτήσεων να εμπλουτισθεί με πολλές παραλλαγές της ίδιας ερώτησης. Οι παραλλαγές κατασκευάζονται αυτομάτως από το σύστημα με τη χρήση ελεύθερων μεταβλητών στη θέση των σταθερών αριθμητικών τιμών στο θέμα και στις επιλογές της ερώτησης. Για παράδειγμα, έστω ερώτηση που εμφανίζεται στον αξιολογούμενο με τη μορφή του παρακάτω Πίνακα 4.

*Πίνακας 4. Παραλλαγή ερώτησης από Τράπεζα Ερωτήσεων*

<b>ΘΕΜΑ</b>	
Σε πόσο χρόνο, ένα ποδήλατο που κινείται με ταχύτητα <b>10</b> χλμ./ώρα, διανύει απόσταση <b>40</b> χλμ.;	
<b>ΕΠΙΛΟΓΕΣ</b>	
A. 2 ώρες	Λανθασμένη επιλογή
B. 4 ώρες	Ορθή επιλογή
Γ. 0,25 ώρες	Λανθασμένη επιλογή
Δ. 8 ώρες	Λανθασμένη επιλογή

Η ερώτηση εισάγεται στο σύστημα, ως *ερώτηση ελεύθερων μεταβλητών* («calculated question») (de Sande, 2015., Heyde & Siebrits, 2019), με τη μορφή του παρακάτω Πίνακα 5.

*Πίνακας 5. Παράδειγμα ερώτησης με ελεύθερες μεταβλητές*

<b>ΘΕΜΑ</b>	
Σε πόσο χρόνο, ένα ποδήλατο που κινείται με ταχύτητα <b>X</b> χλμ./ώρα, διανύει απόσταση <b>Y</b> χλμ.;	
<b>ΕΠΙΛΟΓΕΣ</b>	

A.	$\frac{Y}{2X}$ ώρες	Λανθασμένη επιλογή
B.	$\frac{Y}{X}$ ώρες	Ορθή επιλογή
Γ.	$\frac{X}{Y}$ ώρες	Λανθασμένη επιλογή
Δ.	$\frac{2Y}{X}$ ώρες	Λανθασμένη επιλογή

Δηλαδή, στη θέση των σταθερών αριθμητικών τιμών (10 και 40) τοποθετούνται ελεύθερες μεταβλητές (X και Y αντιστοίχως), που αντλούν αριθμητικές τιμές από καθορισμένο εύρος επιτρεπόμενων τιμών, π.χ.  $X \in \{2, 4, 6, \dots, 100\}$  και  $Y \in \{20, 40, 60, \dots, 1000\}$ . Με τη μέθοδο αυτή, (η οποία προφανώς είναι υλοποιήσιμη σε γνωστικά αντικείμενα που απαιτούν μαθηματικούς υπολογισμούς), είναι δυνατόν να κατασκευαστούν αυτομάτως τόσες διαφορετικές παραλλαγές της ερώτησης, όσες επιτρέπει το εύρος του συνόλου από το οποίο αντλούν τιμές οι ελεύθερες μεταβλητές.

Συχνά, οι εκδοτικοί οίκοι, παρέχουν τράπεζες ερωτήσεων μαζί με τα εκπαιδευτικά βιβλία, δηλώνοντας ότι προορίζονται για «προσωπική μελέτη» (homework) και ότι αποσκοπούν στην καλύτερη κατανόηση του γνωστικού αντικειμένου και στην αυτοαξιολόγηση των αναγνωστών του βιβλίου. Αλλά, επειδή οι ερωτήσεις συχνά προσφέρονται σε ψηφιακή μορφή και με διαμόρφωση (format) κατάλληλη για ενσωμάτωση στις τράπεζες ερωτήσεων των συστημάτων αξιολόγησης ή γενικότερα των Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ), αξιοποιούνται και από τους διδάσκοντες ή γενικότερα από τα εκπαιδευτικά ιδρύματα (Jordan, 2013).

Οι Boitshwarelo et al. (2017) εκφράζουν σοβαρούς προβληματισμούς για την αξιοποίηση από τους διδάσκοντες των τραπεζών ερωτήσεων που περιλαμβάνονται στα εκπαιδευτικά βιβλία και παραθέτουν σχετικές απόψεις ερευνητών. Συγκεκριμένα αναφέρουν, ότι η χρήση ερωτήσεων που κατασκευάζονται από κερδοσκοπικούς εκδοτικούς οργανισμούς εγείρει ανησυχίες σχετικά με την ποιότητα και την καταλληλότητα των σχετικών αξιολογήσεων, ιδιαίτερα ως προς την απότιμηση δεξιοτήτων του 21<sup>ου</sup> αιώνα (Vista & Care, 2017), δεδομένου ότι το γνωστικό επίπεδο στο οποίο εστιάζουν οι ερωτήσεις των εκδοτικών οίκων, συχνά αφορά απομνημόνευση και

ανάκληση. Επιπλέον, αν οι διδάσκοντες χρησιμοποιήσουν απaráλλακτες τις ερωτήσεις του εκδοτικού οίκου, τότε όσοι σπουδαστές έχουν αποκτήσει πρόσβαση στις ίδιες ερωτήσεις πριν το διαγώνισμα αποκτούν αθέμιτο πλεονέκτημα έναντι των υπολοίπων. Το ζήτημα είναι οξύτερο στις περιπτώσεις που απαιτείται χρηματικό κόστος για την πρόσβαση στις ερωτήσεις του εκδοτικού οίκου, οπότε προκύπτουν οικονομικές και κοινωνικές ανισότητες μεταξύ των σπουδαστών. Οι Boitshwarelo et al. (2017) παραπέμποντας στους Masters et al. (2001) και στους Ibbet & Wheldon (2016), σχολιάζουν επίσης την ποιότητα των ερωτήσεων που κατασκευάζουν οι εκδοτικοί οίκοι. Συγκεκριμένα, αναφέρουν ότι αφενός δεν πληρούνται πάντοτε τα απαραίτητα επίπεδα αξιοπιστίας και εγκυρότητας για τελικές αξιολογήσεις. Και ότι αφετέρου έχουν εντοπιστεί ελαττώματα, με σημαντικότερο τη συχνή (σε αναλογία έως και δύο τρίτων) ύπαρξη «σημάτων ορθότητας» (clueing signals), δηλαδή λέξεων ή φράσεων που εμμέσως καθοδηγούν τους αξιολογούμενους και αυξάνουν τις πιθανότητες τους να επιλέξουν τη σωστή απάντηση. Για τους λόγους αυτούς, οι ερωτήσεις εκδοτικών οίκων που συνοδεύουν εκπαιδευτικά βιβλία πρέπει να αξιοποιούνται με προσοχή, να έλεγχονται επιμελώς για λάθη και να χρησιμοποιούνται κυρίως για διαμορφωτικές και όχι για τελικές αξιολογήσεις.

## 4.2. Ψηφιακές αξιολογήσεις ανοικτού τύπου

Στις «αξιολογήσεις ανοικτού τύπου», ο αξιολογούμενος απαντά σε ερώτηση, συνθέτοντας την απάντηση, την οποία ακολούθως καταχωρεί στο σύστημα συνήθως με τη μορφή κειμένου. Το σύνολο των δυνατών απαντήσεων, που μπορεί να συνθέσει και να καταχωρήσει στο σύστημα ο αξιολογούμενος, δεν είναι «κλειστό», δηλαδή πεπερασμένο, όπως συμβαίνει στις «αξιολογήσεις κλειστού τύπου». Αντιθέτως, είναι «ανοικτό», δηλαδή μη πεπερασμένο, γεγονός στο οποίο οφείλεται ο όρος «αξιολογήσεις ανοικτού τύπου». Ωστόσο, το σύνολο των ορθών απαντήσεων είναι περιορισμένο και συνήθως η ορθή απάντηση είναι μόνο μία. Οι αξιολογήσεις ανοικτού τύπου έχουν δύο βασικές μορφές.

Στην πρώτη μορφή, ο αξιολογούμενος καλείται να συμπληρώσει λέξη ή λέξεις που απουσιάζουν από μια πρόταση (π.χ. «Η [Αθήνα] είναι η πρωτεύουσα της Ελλάδος»). Η συγκεκριμένη μορφή εύκολα μετασχηματίζεται σε ερώτηση, όπου η απάντηση αντιστοιχεί στη λέξη που απουσιάζει (π.χ. «Ποια είναι η πρωτεύουσα της Ελλάδος;»).

Στη δεύτερη μορφή, ο αξιολογούμενος καλείται, ανάλογα με τις οδηγίες, είτε να συνθέσει μια σύντομη απάντηση («short answer»), η οποία αποτελείται από λέξη ή φράση και η οποία, σύμφωνα με την Jordan (2013), δεν πρέπει να υπερβαίνει τις είκοσι λέξεις. Είτε να συνθέσει «έκθεση» (essay), η οποία αποτελείται από μια ή πολλές παραγράφους.

Σύμφωνα με τους Oosterhof, Conrad & Ely (2008, σ. 84), η μορφή της «έκθεσης» διευκολύνει την παρακολούθηση των διανοητικών διεργασιών του αξιολογούμενου. Όμως, τόσο η «έκθεση», όσο και οι άλλες μορφές αξιολογήσεων ανοικτού τύπου δυσχεραίνουν την αυτόματη και αντικειμενική βαθμολόγηση, που αποτελούν σημαντικά χαρακτηριστικά της ψηφιακής αξιολόγησης. Αυτός είναι ένας από τους βασικούς λόγους για τους οποίους οι αξιολογήσεις ανοικτού τύπου χρησιμοποιούνται λιγότερο συχνά στις ψηφιακές αξιολογήσεις από ό,τι στις παραδοσιακές αξιολογήσεις, όπου αποτελούν την πιο κοινή μορφή αξιολόγησης.

Η μορφή της σύντομης απάντησης διαθέτει, σύμφωνα με τους Oosterhof, Conrad & Ely (2008, σ. 85), τρία πλεονεκτήματα έναντι των άλλων μορφών. Πρώτον, κατασκευάζεται εύκολα. Αφενός, διότι προορίζεται κυρίως για να αξιολογήσει τη γνωστική ικανότητα για ανάκληση πληροφοριών (όπως π.χ. ονόματα, γεγονότα, ορισμοί) και όχι για να αξιολογήσει γνωστικές



ικανότητες, που αφορούν εντοπισμό διαφορών, κατανόηση εννοιών, εφαρμογή κανόνων που διέπουν σχέσεις (Oosterhof, Conrad & Ely, 2008, σ.15, 16, 85). Αφετέρου, διότι δεν απαιτεί πολύπλοκα σχέδια βαθμολόγησης, όπως οι εκθέσεις, ούτε εναλλακτικές (ορθές και λανθασμένες) επιλογές, όπως οι ερωτήσεις πολλαπλών επιλογών. Δεύτερον, οι σπουδαστές συνθέτουν εκ του μηδενός την απάντηση. Οπότε, δεν έχουν τη δυνατότητα να επιλέξουν απάντηση βασιζόμενοι στην τύχη, όπως συχνά συμβαίνει στα διαγωνίσματα πολλαπλών επιλογών. Δεδομένου ότι η πρακτική της τυχαίας επιλογής υποβαθμίζει την αξιοπιστία της αξιολόγησης (Bush, 2015), οι αξιολογήσεις ανοικτού τύπου θεωρούνται πιο αξιόπιστες έναντι των αξιολογήσεων κλειστού τύπου. Τρίτον, το μικρό μέγεθος των απαντήσεων (σε σύγκριση με τη μορφή της έκθεσης) επιτρέπει τη συμπερίληψη πολλών ζητούμενων σε ένα διαγώνισμα, συνεπώς υπάρχει δυνατότητα για αξιολόγηση που καλύπτει πολλά πεδία του γνωστικού αντικείμενου. Η συγκεκριμένη δυνατότητα επιτρέπει τη γενίκευση της αξιολόγησης και δυνητικά την αύξηση της αξιοπιστίας της.

Δύο είναι τα βασικά μειονεκτήματα που οι Oosterhof, Conrad & Ely (2008, σ. 85) εντοπίζουν στη μορφή της σύντομης απάντησης. Πρώτον, οι απαντήσεις απαιτούν συνήθως ανάκληση απλών γνώσεων όπως π.χ. ονόματα ανθρώπων, τόπων, διαδικασιών. Ο σπουδαστής καλείται να απαντήσει με μια απλή λέξη ή φράση, οπότε αξιολογούνται απλές γνωστικές ικανότητες. Δεύτερον, η ορθή βαθμολόγηση είναι δυσκολότερη σε σχέση με τις ερωτήσεις κλειστού τύπου. Συγκεκριμένα, σε μια ερώτηση είναι δυνατόν να δοθούν πολλές εναλλακτικές, εξίσου ορθές, απαντήσεις. Επιπλέον, η ίδια απάντηση (λέξη ή φράση) είναι δυνατόν να εισαχθεί στο σύστημα με διαφορετική μορφή ως προς την ορθογραφία, τον τονισμό, τη μορφή των γραμμάτων (πεζά ή κεφαλαία), τα σημεία στίξης, τους τυχαίους ή αδιάφορους χαρακτήρες (π.χ. συνεχόμενα κενά). Το γεγονός, καθιστά δυσχερή τον αυτόματο έλεγχο της ορθής απάντησης και ακολούθως την αυτόματη βαθμολόγηση από το σύστημα. Η αντιμετώπιση του προβλήματος απαιτεί ενέργειες όπως: α) κατάλληλο σχεδιασμό της ερώτησης, ώστε να ελαχιστοποιούνται οι εν δυνάμει διαφορετικές, αλλά τελικά ορθές, απαντήσεις που μπορεί να δώσει ένας σπουδαστής, β) καταχώρηση στο σύστημα όλων των εν δυνάμει ορθών απαντήσεων που μπορεί να δώσει ο σπουδαστής και γ) ρύθμιση του συστήματος ώστε να αναγνωρίζει ισοδύναμες απαντήσεις (π.χ. να μετατρέπει τα πεζά γράμματα σε κεφαλαία, ή να αγνοεί σημεία τονισμού) και να ανέχεται

ασήμαντες αποκλίσεις (π.χ. μικρά ορθογραφικά λάθη). Οι συγκεκριμένες ενέργειες αυξάνουν σημαντικά την πολυπλοκότητα της κατασκευής καλών ερωτήσεων, απαιτούν κατάλληλες δεξιότητες και αυξάνουν τον απαιτούμενο χρόνο κατασκευής των ερωτήσεων.

Η ανάπτυξη συστημάτων που αυτομάτως βαθμολογούν σύντομες απαντήσεις έχει αυξήσει τη συχνότητα χρήσης του συγκεκριμένου τύπου αξιολόγησης. Η Jordan (2013) περιγράφει σχετικά συστήματα που αξιοποιούν πολύπλοκες τεχνικές, οι οποίες βασίζονται στην υπολογιστική γλωσσολογία και συγκρίνουν την απάντηση του σπουδαστή με συντακτικούς και γραμματικούς τύπους των πρότυπων απαντήσεων ή απλούστερες τεχνικές που βασίζονται σε συστήματα «ταύτισης μοτίβων» (pattern matching) και συγκρίνουν την απάντηση του σπουδαστή με «λέξεις κλειδιά» και τις συνώνυμές τους. Η Jordan (2013) αναφέρει ότι τέτοια συστήματα βαθμολογούν με ακρίβεια ανάλογη (ή και υψηλότερη) της ακρίβειας βαθμολόγησης που υλοποιούν φυσικά πρόσωπα.

Η μορφή της «έκθεσης» είναι αρκετά κοινή στις παραδοσιακές αξιολογήσεις, αλλά αξιοποιείται και στην ψηφιακή αξιολόγηση, δεδομένου ότι είναι κατάλληλη για την αξιολόγηση μεγάλου εύρους γνωστικών ικανοτήτων, τουλάχιστον όσων είναι αξιολογήσιμες μέσω γραπτών εξετάσεων (Oosterhof, Conrad & Ely, 2008, σ. 89). Τα βασικά πλεονεκτήματα της έκθεσης είναι τρία. Πρώτον, η έκθεση αποτιμά με μεγαλύτερη αμεσότητα, έναντι άλλων μορφών, την επίτευξη των μαθησιακών στόχων. Διότι, οι περισσότεροι μαθησιακοί στόχοι μπορούν εύκολα να μεταγραφούν ως ερωτήσεις κατάλληλες για τη μορφή της έκθεσης. Δεύτερον, η έκθεση είναι η καταλληλότερη μορφή για να αξιολογηθεί η ικανότητα στην επικοινωνία των ιδεών. Ωστόσο, δεν πρέπει να χρησιμοποιείται για να αξιολογηθεί ταυτόχρονα η ποιότητα του γραπτού λόγου και η κατανόηση κειμένου. Διότι, η χρήση της ίδιας δοκιμασίας για την αξιολόγηση διαφορετικών ικανοτήτων μειώνει την αξιοπιστία της αξιολόγησης. Τρίτον, η έκθεση επιτρέπει στους σπουδαστές να παραθέσουν τους συλλογισμούς που χρησιμοποίησαν για να καταλήξουν στην απάντησή τους, οπότε αξιολογούνται επιπλέον γνωστικές ικανότητες. Επιπλέον, οι σπουδαστές δεν μπορούν να δώσουν σωστή απάντηση, αν στηριχθούν σε εσφαλμένους συλλογισμούς, όπως είναι δυνατόν να συμβεί σε αξιολογήσεις κλειστού τύπου (Oosterhof, Conrad & Ely, 2008, σ. 89).

Οι περιορισμοί που έχει η μορφή της έκθεσης (Oosterhof, Conrad & Ely, 2008, σ. 90) είναι οι εξής: Πρώτον, ο χρόνος που απαιτείται για την ανάπτυξη μιας απάντησης είναι μεγάλος σε σχέση με άλλους τύπους αξιολόγησης. Εντός του περιορισμένου χρόνου μιας αξιολογικής διαδικασίας, είναι λίγες οι ερωτήσεις που μπορούν συμπεριληφθούν, συνεπώς ένα διαγώνισμα δεν μπορεί να καλύψει μεγάλο εύρος του γνωστικού αντικειμένου. Επιπλέον, οι αξιολογούμενοι καταναλώνουν περισσότερο χρόνο για να καταγράψουν την απάντηση, παρά για να στοχαστούν επί της ερώτησης ή για να επιλύσουν το πρόβλημα. Δεύτερον, η βαθμολόγηση των απαντήσεων είναι αρκετά υποκειμενική, ιδιαίτερα στις περιπτώσεις που δεν έχουν συνταχθεί και διανεμηθεί σαφείς οδηγίες και λεπτομερές πλάνο βαθμολόγησης. Το πρόβλημα είναι ιδιαίτερα έντονο στις εξετάσεις μεγάλης κλίμακας, όπου πολλοί βαθμολογητές με διαφορετικό γνωστικό υπόβαθρο, διαφορετική εκπαίδευση και διαφορετική εμπειρία αξιολογούν μεγάλο αριθμό εκθέσεων (Rudner & Gagne, 2000). Η υποκειμενικότητα, ζημιώνει τη συνέπεια της βαθμολόγησης και γενικότερα την αξιοπιστία της έκθεσης ως μορφής αξιολόγησης. Τρίτον, η βαθμολόγηση των εκθέσεων από φυσικό πρόσωπο είναι χρονοβόρος διαδικασία. Ο απαιτούμενος χρόνος αυξάνεται περαιτέρω, στις περιπτώσεις που, για λόγους αξιοπιστίας, στη βαθμολόγηση συμμετέχουν δύο ή και τρεις αξιολογητές. Η αυτόματη βαθμολόγηση μέσω συστήματος, προφανώς, μειώνει κατά πολύ τον απαιτούμενο χρόνο και απαλείφει τον συγκεκριμένο περιορισμό. Αλλά, η συγκεκριμένη δυνατότητα, συνήθως προσφέρεται σε αξιολογικές διαδικασίες μεγάλης κλίμακας και όχι σε αξιολογήσεις που σχεδιάζονται και υλοποιούνται από έναν διδάσκοντα στα πλαίσια ενός μαθήματος.

Η αυτόματη βαθμολόγηση εκθέσεων μειώνει τον χρόνο και το κόστος της αξιολόγησης, αυξάνει την αξιοπιστία της και επιπλέον διευκολύνει την αναβαθμολόγηση, αν επέλθει αλλαγή στα κριτήρια αξιολόγησης (Rudner & Gagne, 2000). Η Jordan (2013) αναφέρει ότι η αυτόματη βαθμολόγηση εκθέσεων διαφέρει από την αυτόματη βαθμολόγηση σύντομων απαντήσεων. Διότι, η βαθμολόγηση εκθέσεων αφενός επικεντρώνεται συχνά στο στυλ του γραπτού λόγου και αφετέρου δεν είναι εφικτό να προβλεφθούν όλες οι εν δυνάμει ορθές απαντήσεις, όπως συμβαίνει στη μορφή των σύντομων απαντήσεων. Οι Valenti, Neri & Cucchiarelli (2003) περιγράφουν και συγκρίνουν αρκετά συστήματα αυτόματης βαθμολόγησης εκθέσεων, σημειώνοντας ότι κάποια αξιολογούν το στυλ του γραπτού λόγου, ενώ κάποια άλλα αξιολογούν

το περιεχόμενο. Αναφέρουν ότι πολύ σημαντικό ζήτημα για την πρόοδο στο ερευνητικό πεδίο της αυτόματης βαθμολόγησης εκθέσεων είναι η συγκρότηση ενός μεγάλου σώματος εκθέσεων, βαθμολογημένων από φυσικά πρόσωπα, το οποίο θα χρησιμοποιηθεί για τη δοκιμή, την αξιολόγηση και τη βελτίωση των σχετικών αλγόριθμων.

### 4.3. Ψηφιακοί φάκελοι επιτευγμάτων (e-portfolios)

Οι ψηφιακοί φάκελοι επιτευγμάτων (e-portfolios) είναι συστήματα αποθήκευσης και διαχείρισης ψηφιακών αρχείων, που περιλαμβάνουν, όχι μόνο ψηφιοποιημένα πιστοποιητικά και βεβαιώσεις σπουδών, αλλά και επιλεγμένα από τον σπουδαστή πολυμεσικά αντικείμενα, που σχετίζονται με τις γνώσεις και τις δεξιότητές του (Mason, Pegler & Weller, 2004). Σύμφωνα με τον Wiedmer (1998), ο ψηφιακός φάκελος επιτευγμάτων είναι μια ψηφιακή συλλογή από τις εργασίες του σπουδαστή, μέσω της οποίας παρουσιάζει τις προσπάθειες, την πρόοδο και τα επιτεύγματά του. Σύμφωνα με τις Oldfield et al. (2012), ο όρος ψηφιακός φάκελος επιτευγμάτων συνήθως αναφέρεται σε έναν προσωπικό διαδικτυακό χώρο, που λειτουργεί ως πλαίσιο αξιολόγησης και προσφέρει δυνατότητες όπως αποθήκευση πληροφοριών, οργάνωση της μάθησης και συνεργασία. Οι σπουδαστές αποθέτουν στους φάκελους, τεκμήρια μάθησης, τα οποία συσχετίζουν με συγκεκριμένες γνώσεις και δεξιότητες.

Όπως υποστηρίζουν οι Cotterill, Bradley & Hammond (2006) οι ψηφιακοί φάκελοι επιτευγμάτων τεκμηριώνουν μαθησιακά επιτεύγματα και παρέχουν στους σπουδαστές προσωποποιημένη θέαση όλων των αποτελεσμάτων, που έλαβαν κατά τη διάρκεια αξιολογικών διαδικασιών. Επιπλέον, βοηθούν τον διαδικτυακό διαμοιρασμό των περιεχομένων του φάκελου σε αξιολογητές ή άλλους σπουδαστές, οι οποίοι μπορούν να παράσχουν διαμορφωτική ανατροφοδότηση. Ο διαδικτυακός διαμοιρασμός και η πρόσβαση αποτελεί βασικό πλεονέκτημα για την καταγραφή και τεκμηρίωση των επιτευγμάτων σε διαδικασίες διαμορφωτικής αξιολόγησης. Η Jordan (2013) αναφέρει ότι οι ψηφιακοί φάκελοι επιτευγμάτων χρησιμοποιούνται ευρέως, προκειμένου οι σπουδαστές να καταγράψουν τη μαθησιακή πορεία τους, να στοχαστούν για αυτήν και να αποθηκεύσουν τεκμήρια των επιτευγμάτων τους. Με αυτόν τον τρόπο, είναι δυνατόν να αξιολογηθούν δεξιότητες που δύσκολα αξιολογούνται αλλιώς και να ενθαρρυνθεί η στοχαστική προσέγγιση της μάθησης. Οι σπουδαστές είναι υπεύθυνοι για το τι θα συμπεριληφθεί στον φάκελο, οπότε επικεντρώνονται σε θετικούς στόχους. Επιπλέον, οι ψηφιακοί φάκελοι επιτευγμάτων, όχι μόνο αποτελούν μέσο αξιολόγησης των σπουδαστών, αλλά τους βοηθούν στην αναζήτηση και στην εύρεση θέσεων εργασίας. Οι Lorenzo & Ittelson (2005) χαρακτηρίζουν τους ψηφιακούς φάκελους επιτευγμάτων, ως μια από τις σημαντικότερες καινοτομίες της εκπαιδευτικής τεχνολογίας. Ο Jafari (2004) αναφέρει ότι

παρέχουν νέους τρόπους για την παρουσίαση και την αξιολόγηση των σπουδαστών και για τη συσχέτιση των μαθησιακών στόχων και των επιτευγμάτων. Τη δυνατότητα συσχέτισης μεταξύ των μαθησιακών στόχων και των επιτευγμάτων των σπουδαστών τονίζει και η Barret (2000), η οποία επιπλέον αναφέρει ότι με τη βοήθεια των ψηφιακών φάκελων επιτευγμάτων, οι σπουδαστές συμμετέχουν ενεργά στη μαθησιακή διαδικασία, διότι μπορούν να ορίσουν και να αναθεωρήσουν στόχους, να στοχαστούν για τη μαθησιακή πορεία τους και να αναλάβουν την ευθύνη της μάθησής τους.

Οι Oldfield et al. (2012) αναφέρουν ότι οι ψηφιακοί φάκελοι επιτευγμάτων χρησιμοποιούνται συχνά για σκοπούς αξιολόγησης και συγκέντρωσης επαγγελματικών πιστοποιήσεων και ότι συνιστούν σημαντική πρόοδο σε σχέση με τους παραδοσιακούς έντυπους φάκελους, τόσο για πρακτικούς λόγους (ευκολία και ταχύτητα διαχείρισης), όσο και για παιδαγωγικούς λόγους (ενίσχυση στοχασμού και ταχύτερη ανατροφοδότηση). Συνεπώς, παρέχουν δυνατότητες εναλλακτικών τρόπων αξιολόγησης για όσους σπουδαστές αντιμετωπίζουν δυσκολίες με τους παραδοσιακούς τρόπους αξιολόγησης. Οι Mason, Pegler και Weller (2004) υποστηρίζουν ότι οι ψηφιακοί φάκελοι επιτευγμάτων μπορούν να αξιοποιηθούν ως εργαλεία προσωπικής βελτίωσης, αυτοπαρουσίασης, αλλά και αξιολόγησης των σπουδαστών. Οι σπουδαστές συλλέγουν τεκμήρια της μαθησιακής διαδρομής τους, τα διαμοιράζονται με άλλους σπουδαστές, με διδάσκοντες, με εργοδότες, τα αξιοποιούν για να αποδείξουν την κατοχή γνώσεων και δεξιοτήτων και τα συσχετίζουν με εκπαιδευτικούς, επαγγελματικούς, προσωπικούς και κοινωνικούς στόχους.

Οι Mason, Pegler και Weller (2004) καταγράφουν μια σειρά πλεονεκτημάτων που προκύπτουν από την αξιοποίηση των ψηφιακών φάκελων επιτευγμάτων. Υποστηρίζουν ότι η επιλογή, η συγκέντρωση, η οργάνωση και η πρόσβαση των τεκμηρίων που εντάσσονται στους ψηφιακούς φάκελους επιτευγμάτων είναι εύκολη και γρήγορη, διότι τα τεκμήρια αφενός είναι σε ψηφιακή μορφή και αφετέρου είναι διαχειρίσιμα μέσω ψηφιακών συστημάτων. Ο διαδικτυακός διαμοιρασμός των ψηφιακών φάκελων επιτευγμάτων διευκολύνει τον συνεργατικό σχολιασμό και την αναθεώρησή τους. Οι ψηφιακοί φάκελοι επιτευγμάτων προσφέρουν πολλούς τρόπους παρουσίασης. Με τη χρήση αρχείων πολυμέσων και αναφορών σε εξωτερικές πηγές, αυξάνεται η ποικιλία και η ελκυστικότητα των τεκμηρίων. Οι Oldfield et al. (2012) αναφερόμενες στα

πλεονεκτήματα των ψηφιακών φάκελων επιτευγμάτων που καταγράφονται στη βιβλιογραφία, τονίζουν την ανάπτυξη της δημιουργικότητας των σπουδαστών και την ενίσχυση της συνεργατικής και προοδευτικής μάθησης. Υποστηρίζουν ότι ως μέθοδος αξιολόγησης, οι ψηφιακοί φάκελοι επιτευγμάτων παρέχουν δυνατότητες για την ενοποίηση και την εποπτεία του συνόλου των εργασιών του σπουδαστή για όλη τη διάρκεια του μαθήματος. Σύμφωνα με αναφορά του ePortConsurtium (2003), την οποία καταγράφουν οι Sweat-Guy & Buzzeto-More (2007), οι ψηφιακοί φάκελοι επιτευγμάτων βοηθούν τους σπουδαστές να αναπτύξουν οργανωτικές δεξιότητες, να εντοπίζουν ατομικές ικανότητες και αδυναμίες, να παρακολουθήσουν τη διαχρονική ανάπτυξη των δεξιοτήτων τους, να λάβουν επαγγελματικές αποφάσεις, να αποδείξουν ότι πληρούν τις ακαδημαϊκές προϋποθέσεις ενός προγράμματος σπουδών, να προωθήσουν επαγγελματικά τους εαυτούς τους. Η αναφορά υποστηρίζει επίσης ότι οι ψηφιακοί φάκελοι επιτευγμάτων είναι καταλληλότεροι από άλλες μορφές αξιολόγησης, διότι περιλαμβάνουν ποσότητα και ποικιλία στοιχείων, σχετικών με τη συνολική μαθησιακή εμπειρία των σπουδαστών, διότι παρέχουν δυνατότητες για συσχετίσεις με πρότυπα και με βάσεις δεδομένων αξιολόγησης και διότι επιτρέπουν πρόσβαση από εξωτερικούς αξιολογητές. Γενικά, οι Sweat-Guy & Buzzeto-More (2007) υποστηρίζουν ότι οι ψηφιακοί φάκελοι επιτευγμάτων αποτελούν μοναδικά εργαλεία για την τεκμηρίωση της μαθησιακής προόδου των σπουδαστών και για την ενθάρρυνσή της ενεργής συμμετοχής στη μαθησιακή διαδικασία. Στα πλεονεκτήματα των ψηφιακών φάκελων επιτευγμάτων οι Green, Wyllie & Jackson (2014) συγκαταλέγουν την καλλιέργεια της υπευθυνότητας και της αυτονομίας των σπουδαστών, διότι τους ενθαρρύνουν να αναλαμβάνουν την ευθύνη για τις μαθησιακές ανάγκες τους και για την κατεύθυνση, την πρόοδο και την ποιότητα της μαθησιακής πορείας τους. Ως πλεονέκτημα αναφέρουν επίσης τη δυνατότητα της φορητότητας και της προσαρμοστικότητας των τεκμηρίων, διότι αποθηκεύονται σε ψηφιακά αποθετήρια, τα οποία επιτρέπουν τη διανομή και την αναδιοργάνωση των τεκμηρίων, ανάλογα με το κοινό στο οποίο απευθύνονται.

Οι Oldfield et al. (2012) αναφέρουν παραδείγματα στα οποία οι ψηφιακοί φάκελοι επιτευγμάτων αξιοποιούνται σε διαδικασίες ομότιμης αξιολόγησης και αυτοαξιολόγησης, όπου καταγράφηκε, αφενός υψηλός βαθμός ικανοποίησης τόσο των σπουδαστών όσο και των διδασκόντων και αφετέρου βελτίωση των μαθησιακών αποτελεσμάτων. Επίσης καταγράφηκε

υψηλός βαθμός αναστοχασμού των σπουδαστών επί των επιτευγμάτων τους, ο οποίος αυξάνεται περαιτέρω, αν στη διαδικασία συμμετέχουν και οι άλλοι σπουδαστές. Σύμφωνα με την Redecker (2013), οι ψηφιακοί φάκελοι επιτευγμάτων μπορούν να ενισχύσουν τη συνεργασία των σπουδαστών εντός ομάδων εργασιών, οπότε είναι κατάλληλοι για τον συνδυασμό συνεργατικών μορφών μάθησης και ατομικών μεθόδων αξιολόγησης. Μπορούν επίσης να συνδυάσουν τη διαμορφωτική αξιολόγηση με την αυτοαξιολόγηση και την ομότιμη αξιολόγηση.

Οι Green, Wyllie & Jackson (2014) υποστηρίζουν ότι η χρήση των ψηφιακών φάκελων επιτευγμάτων ως μέσων αξιολόγησης, ενισχύει την αυθεντικότητα της αξιολόγησης. Διότι επιτρέπει την αξιολόγηση της ικανότητας του σπουδαστή να ορίσει ατομικούς στόχους, να αναπτύξει κριτική σκέψη, να επιλύσει προβλήματα, να δομήσει νέα γνώση. Αν οι στόχοι ορισθούν εγκαίρως, τότε οι σπουδαστές, με την ολοκλήρωση των σπουδών τους, έχουν καταρτήσει φάκελο έτοιμο προς υποβολή για αναζήτηση θέσεως εργασίας ή για περαιτέρω σπουδές. Επιπλέον, με τους ψηφιακούς φάκελους επιτευγμάτων είναι δυνατόν να αξιολογηθούν, όχι μόνο γνωστικές, αλλά και συναισθηματικές δεξιότητες.

Σύμφωνα με τους Cotterill, Bradley & Hammond (2006), οι ψηφιακοί φάκελοι επιτευγμάτων είναι δυνατόν να αμβλύνουν την αντίθεση μεταξύ διαμορφωτικής και τελικής αξιολόγησης, δεδομένου ότι ο σπουδαστής δομεί σταδιακά το φάκελο και λαμβάνει διαμορφωτική ανατροφοδότηση για τη βελτίωση ή την προσθήκη υλικού, αλλά επιλέγει το τελικό υποσύνολο των τεκμηρίων που θα υποβάλλει για την τελική αξιολόγησή του. Επίσης, οι ψηφιακοί φάκελοι επιτευγμάτων, αν έχουν την κατάλληλη δομή, μπορούν να περιλάβουν τεκμήρια για διαφορετικά γνωστικά αντικείμενα και για διαφορετικές δεξιότητες, που δύσκολα μπορούν να αξιολογηθούν ταυτοχρόνως από τις παραδοσιακές μορφές αξιολόγησης. Η κατασκευή και η διαχείριση ενός ψηφιακού φάκελου επιτευγμάτων αποτελεί από μόνη της αντικείμενο προς αξιολόγηση, διότι τεκμηριώνει τη δεξιότητα της ανεξάρτητης μάθησης.

Οι Green, Wyllie & Jackson (2014) δεν παραλείπουν να αναφερθούν στους προβληματισμούς που καταγράφονται στη βιβλιογραφία σχετικά με την εγκυρότητα και την αξιοπιστία της αξιολόγησης μέσω ψηφιακών φάκελων επιτευγμάτων. Τα σημεία στα οποία επικεντρώνεται η σχετική κριτική αφορούν τα εξής ζητήματα: Η κατάρτιση του φακέλου απαιτεί ευχέρεια στον



γραπτό λόγο, την οποία δεν διαθέτουν όλοι οι σπουδαστές με αποτέλεσμα να ευνοούνται όσοι υπερτερούν στον γραπτό λόγο. Κάποιοι σπουδαστές ενδεχομένως φοβούνται και αποφεύγουν να καταγράψουν τις εμπειρίες τους σε γραπτό κείμενο και να τις κοινοποιήσουν σε τρίτους, ενώ άλλοι, γνωρίζοντας ότι ο φάκελος αποτελεί αντικείμενο αξιολόγησης, δεν ενεργούν με αυθεντικότητα και καταγράφουν μόνο όσα θεωρούν ότι θα εκτιμηθούν από τον αξιολογητή. Έτερο πρόβλημα, που αναφέρεται από τους Green, Wyllie & Jackson (2014), είναι η χρονική διάρκεια που απαιτείται για την κατάρτιση και την αξιολόγηση του φακέλου και η σύγχυση που συχνά προκαλείται στους σπουδαστές ως προς το τι θα περιληφθεί και τι θα παραληφθεί κατά την κατάρτιση του φακέλου. Προβληματισμός εκφράζεται από τις Oldfield et al. (2012) για την τάση αξιοποίησης των ψηφιακών φακέλων επιτευγμάτων, περισσότερο για αξιολογικούς και λιγότερο για παιδαγωγικούς σκοπούς. Οι Weleschuk, Dyjur & Kelly (2019) καταγράφουν προκλήσεις που προκύπτουν κατά τη χρήση των ψηφιακών φακέλων επιτευγμάτων, οι οποίες σχετίζονται με το άγχος που νιώθουν οι σπουδαστές όταν καλούνται να δημιουργήσουν έναν μεγάλο φάκελο με πολλά τεκμήρια, με τη δυσκολία χρήσης των συστημάτων που διαχειρίζονται τους φακέλους και με τη δυσκολία κατανόησης από τους σπουδαστές της σκοπιμότητας δημιουργίας φακέλου. Σημαντική πρόκληση σύμφωνα με τη Redecker (2013) αποτελεί η αξιολόγηση των ψηφιακών φακέλων επιτευγμάτων, η οποία είναι δύσκολο να αυτοματοποιηθεί, οπότε προκαλεί μεγάλο φόρτο εργασίας στον διδάσκοντα.

#### 4.4. Συνεργατική ανάπτυξη περιεχομένου (wiki)

Τα wiki είναι συστήματα συνεργατικής ανάπτυξης ψηφιακού περιεχομένου, τα οποία επιτρέπουν σε πολλούς χρήστες να αναπτύξουν από κοινού ψηφιακό περιεχόμενο αναρτημένο σε ιστοσελίδα (Oldfield et al., 2012). Το περιεχόμενο, που διαχειρίζονται τα wiki, διατίθεται σε οποιονδήποτε διαθέτει απλή πρόσβαση (δικαίωμα ανάγνωσης) στη σχετική ιστοσελίδα και μπορεί να τροποποιηθεί από χρήστες / συντάκτες που διαθέτουν λογαριασμό πρόσβασης με δικαίωμα τροποποίησης. Βασικό χαρακτηριστικό των wiki είναι η τήρηση αρχείου, όπου καταγράφεται το ιστορικό των τροποποιήσεων, ώστε να είναι διαθέσιμες οι πληροφορίες για το ποιος και πότε τροποποίησε την ιστοσελίδα. Με τον τρόπο αυτό, τα wiki επιτρέπουν, όχι μόνο την συνεργατική ανάπτυξη και διαδικτυακή διανομή περιεχομένου, αλλά επιπλέον την ιχνηλάτηση κάθε παρέμβασης (προσθήκης ή τροποποίησης), δηλαδή τη συσχέτισή της με συγκεκριμένο συντάκτη. Επίσης, η τήρηση του ιστορικού των τροποποιήσεων επιτρέπει την πρόσβαση, όχι μόνο στην τρέχουσα έκδοση, αλλά και σε παλαιότερες εκδόσεις της ιστοσελίδας. Η έκταση που μπορεί να λάβει ένα wiki κυμαίνεται από λίγες σελίδες, ως γιγάντιους ιστότοπους, με χαρακτηριστικότερο παράδειγμα τη διαδικτυακή εγκυκλοπαίδεια Wikipedia, που περιλαμβάνει εκατομμύρια λήμματα. Οι εν λόγω ιδιότητες οδηγούν τους Carle & Bogle (2013) να χαρακτηρίσουν τα wiki ως ανοικτά, συνεργατικά και διαφανή εργαλεία και να αναφέρουν ότι η χρήση τους απαιτεί ελάχιστες τεχνικές δεξιότητες.

Η προσβασιμότητα σε τρέχουσες και παλαιότερες εκδόσεις των ιστοσελίδων και η συσχέτιση κάθε προσθήκης ή τροποποίησης με συγκεκριμένους συντάκτες καθιστούν τα wiki εξαιρετικά εργαλεία για την αξιολόγηση ομάδων σπουδαστών, οι οποίοι υλοποιούν συνεργατικά κάποιο κοινό έργο. Το κοινό έργο αφενός αξιολογείται συνολικά, αλλά αφετέρου είναι δυνατή η αξιολόγηση της ατομικής συνεισφοράς κάθε σπουδαστή (Carle & Bogle, 2013). Η δυνατότητα για την αξιολόγηση της ατομικής συνεισφοράς, βοηθά να επιλυθεί το πρόβλημα της «άδικης» αξιολόγησης των συμμετεχόντων σε ομαδικά έργα, όπου συνήθως είναι δύσκολος ο εντοπισμός της συνεισφοράς κάθε μέλους της ομάδας, με αποτέλεσμα οι σπουδαστές που συνεισέφεραν ελάχιστα να ανταμείβονται με τον ίδιο βαθμό, που λαμβάνουν και οι σπουδαστές που είχαν μεγάλη συνεισφορά στο ομαδικό έργο (James, McInnis & Devlin, 2002).

Ως προς τη δυνατότητα αξιολόγησης της ατομικής συνεισφοράς, οι De Wever, Van Keer, Schellens & Valcke (2011) σημειώνουν ότι, αν το πλήθος των σπουδαστών είναι μεγάλο, αν τα wiki είναι πολλά, αν σε κάθε wiki συνεργάζονται πολλοί σπουδαστές, τότε είναι πρακτικά δύσκολο (αν και τεχνικά εφικτό) να αξιολογηθεί ξεχωριστά κάθε συνεισφορά. Διότι πρόκειται για διαδικασία, όπου ο διδάσκοντας πρέπει να αφιερώσει πολύ χρόνο και κόπο. Για τις περιπτώσεις όπου το πλήθος των σπουδαστών είναι μεγάλο, αντιπροτείνουν, ως πρακτικώς εφικτή λύση, την ομότιμη αξιολόγηση μεταξύ των σπουδαστών που αναπτύσσουν το ίδιο wiki. Διότι, οι ίδιοι οι σπουδαστές γνωρίζουν καλύτερα τη δομή και το περιεχόμενο του δικού τους wiki και τις συνέργειες που οδήγησαν στην ανάπτυξή του, συνεπώς μπορούν να αξιολογήσουν την συνεισφορά καθενός. Επομένως, τα wiki, αν αξιοποιηθούν ως εργαλεία διαδικτυακής ομότιμης αξιολόγησης, ενισχύουν την υπευθυνότητα και τα κίνητρα των σπουδαστών και δομούν κοινότητες μάθησης, όπως υποστηρίζουν οι Weleschuk, Dyjur & Kelly (2019).

Οι Robles, Pons, Borrell, Freixas, Fernandez-Cordoba, Sebastian et al. (2008) τονίζουν τη χρησιμότητα των wiki ως εργαλείων διαμορφωτικής αξιολόγησης. Ειδικότερα, αναφέρουν ότι διευκολύνουν την επικοινωνία μεταξύ των σπουδαστών και του διδάσκοντα, ο οποίος έχει τη δυνατότητα να παρεμβαίνει και να καθοδηγεί τους σπουδαστές, ενόσω το ομαδικό έργο αναπτύσσεται. Οι προτεινόμενες βελτιώσεις και οι οδηγίες, που παρέχονται από τον διδάσκοντα ή ακόμα και από τους άλλους σπουδαστές, είναι άμεσες και επομένως αποδοτικές και είναι δυνατόν να παρέχονται καθ' όλη τη χρονική περίοδο ανάπτυξης του ομαδικού έργου. Την άποψη ότι τα wiki είναι κατάλληλα για διαμορφωτική αξιολόγηση υποστηρίζουν και οι Oldfield et al. (2012), που αναφέρουν ότι τα wiki διευκολύνουν τον διάλογο μεταξύ ομάδων σπουδαστών, ότι αναπτύσσουν τον αναστοχασμό και την αυτονομία των σπουδαστών και ότι είναι ιδιαιτέρως κατάλληλα για να υποστηρίξουν συνεργατικές μαθησιακές δραστηριότητες. Ωστόσο, συμμερίζονται τις επιφυλάξεις των Barton & Heiman (2012) σύμφωνα με τις οποίες η αξιολόγηση με την χρήση των wiki είναι ιδιαιτέρως απαιτητική διαδικασία, εξαιτίας της δυσκολίας στην ανάλυση του μεγάλου όγκου πληροφοριών, που παράγονται κατά τις φάσεις της συνεχούς ανάπτυξης του περιεχομένου.

#### 4.5. Χώροι δημόσιου διαλόγου (Forum)

Οι χώροι δημόσιου διαλόγου είναι χώροι ασύγχρονων διαδικτυακών συζητήσεων, όπου οι χρήστες αναρτούν μικρής έκτασης κείμενα, ή σύντομα σχόλια επί συγκεκριμένων ζητημάτων, καταθέτουν τις απόψεις και τις εμπειρίες τους, επιχειρηματολογούν και απαντούν ο ένας στον άλλον. Συνήθως φιλοξενούνται εντός Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ) και μπορούν να αξιοποιηθούν για συζητήσεις στις οποίες συμμετέχουν όλοι οι σπουδαστές, ή για συζητήσεις μεταξύ μικρότερων ομάδων σπουδαστών.

Οι χώροι δημόσιου διαλόγου είναι από τα παλαιότερα εργαλεία που χρησιμοποιήθηκαν στην ψηφιακή μάθηση και είναι ιδιαίτερα αποδεκτά και οικεία στους σπουδαστές, διότι προσομοιώνουν τον φυσικό διάλογο εντός μιας παραδοσιακής τάξης (Shaul, 2008). Ο Baleni (2015) τονίζει τη χρησιμότητα τους στην υποστήριξη της διαμορφωτικής ψηφιακής αξιολόγησης, διότι διευκολύνουν τον διαμοιρασμό πληροφοριών και τη συνεργατική δόμηση της γνώσης. Οι Vonderwell et al. (2007) αναφέρουν ότι οι χώροι δημόσιου διαλόγου μπορούν να χρησιμοποιηθούν για την ενσωμάτωση διαδικασιών αξιολόγησης και ότι προσφέρουν στους σπουδαστές τη δυνατότητα να απαντούν σε ερωτήσεις, να συμμετέχουν ισότιμα και να συνεισφέρουν στη συνεργατική δόμηση της γνώσης, μέσω ουσιαστικών συζητήσεων. Επιπλέον, υποστηρίζουν ότι οι χώροι δημόσιου διαλόγου διευκολύνουν μια πολυδιάστατη διαδικασία αξιολόγησης, οι διαστάσεις της οποίας είναι η δομή της συζήτησης, οι αυτορρυθμιζόμενες γνωστικές λειτουργίες και δραστηριότητες, η αυτονομία των σπουδαστών, η συμμετοχή στη μαθησιακή κοινότητα και οι δεξιότητες στον γραπτό λόγο. Άξιες αναφοράς είναι οι επισημάνσεις των Weleschuk, Dyjur & Kelly (2019) που υποστηρίζουν ότι κάποιοι σπουδαστές ανησυχούν για τα σχόλια που θα λάβουν και ενοχλούνται με την κριτική που τυχόν θα δεχτούν οι απόψεις τους, ενώ κάποι άλλοι νιώθουν μειονεκτικά, αν οι αναρτήσεις δεχθούν λιγότερα σχόλια σε σύγκριση με αναρτήσεις άλλων. Τέτοιες στάσεις είναι δυνατόν να αλλοιώσουν τον συνεργατικό χαρακτήρα των χώρων δημόσιων διαλόγων και να τους προσδώσουν ανταγωνιστική χροιά.

Ο Shaul (2008) προσφέρει μια αναλυτική παρουσίαση και αποτίμηση των τρόπων, μέσω των οποίων είναι δυνατή η αξιοποίηση των χώρων δημόσιου διαλόγου ως εργαλείων αξιολόγησης. Καταρχήν, αναφέρει ότι υφίστανται διάφοροι τύποι χώρων δημόσιου διαλόγου και ότι όλοι οι τύποι δεν είναι κατάλληλοι ως εργαλεία αξιολόγησης. Οι «κοινωνικοί» χώροι χρησιμοποιούνται

από τους σπουδαστές για ανεπίσημες, ελεύθερες και «χαλαρές» συζητήσεις, οι οποίες συνήθως δεν σχετίζονται με το μάθημα ή, ακόμα και αν σχετίζονται, φιλοξενούν θετικά ή αρνητικά σχόλια για τον διδάσκοντα, για το μάθημα, για το εκπαιδευτικό υλικό, για τις εκπαιδευτικές υποδομές, για τον εκπαιδευτικό οργανισμό. Οι χώροι δημόσιου διαλόγου που ανήκουν σε αυτόν τον τύπο, προφανώς δεν πρέπει να αποτελούν πεδίο αξιολόγησης. Κατάλληλοι για αξιολόγηση είναι οι χώροι δημόσιου διαλόγου που σχετίζονται σαφώς με το γνωστικό αντικείμενο. Στους χώρους αυτούς, η συζήτηση είναι καλά δομημένη, έχει συγκεκριμένο προσανατολισμό και διέπεται από κανόνες τους οποίους ορίζει ο διδάσκοντας. Οι σπουδαστές αξιολογούνται ανάλογα με τη συνεισφορά τους στον διάλογο και την ευστοχία και την ποιότητα των στοχασμών και των επιχειρημάτων τους.

Ο Shaul (2008) αναλύει δύο μεθόδους για τη χρήση των χώρων δημόσιου διαλόγου ως μεθόδου αξιολόγησης: την ποσοτική, αντικειμενική μέθοδο και την ποιοτική, υποκειμενική μέθοδο. Στα πλαίσια της ποσοτικής, αντικειμενικής μεθόδου, η πιο απλή και συχνή προσέγγιση είναι η καταμέτρηση του πλήθους των αναρτήσεων κάθε σπουδαστή και ακολούθως η σύγκριση του πλήθους με προκαθορισμένα επίπεδα που αντιστοιχούν σε αντίστοιχους βαθμούς. Τα πλεονεκτήματα είναι προφανή. Ο διδάσκοντας εύκολα καταγράφει το πλήθος των αναρτήσεων κάθε σπουδαστή, δεδομένου ότι τα ΣΔΜ συνήθως εμφανίζουν αυτομάτως τη σχετική πληροφορία. Επιπλέον, τα κριτήρια αξιολόγησης είναι απλά και ευκόλως κατανοητά από τους σπουδαστές, οι οποίοι αντιλαμβάνονται πλήρως τον ποσοτικό στόχο, που πρέπει να επιτύχουν. Όμως, υφίσταται το μειονέκτημα, ότι όλες οι αναρτήσεις δεν είναι ποιοτικώς ισότιμες. Δεδομένου ότι το ΣΔΜ απλώς απαριθμεί τις αναρτήσεις, οι μονολεκτικές αναρτήσεις (π.χ. «συμφωνώ»), θεωρούνται ισότιμες με τις αναρτήσεις που περιλαμβάνουν εύστοχα σχόλια και πειστικά επιχειρήματα. Συνεπώς, η αξιολόγηση που βασίζεται αποκλειστικά στο πλήθος των αναρτήσεων είναι άδικη, όπως παραδέχονται τόσο οι σπουδαστές, όσο και οι διδάσκοντες. Επιπλέον, η απλή απαρίθμηση αναρτήσεων παρωθεί τους σπουδαστές να καταθέσουν απόψεις, αλλά δεν εξασφαλίζει ότι θα μελετήσουν, ή έστω θα διαβάσουν τις αναρτήσεις των υπολοίπων. Ως αποτέλεσμα, ο χώρος δημόσιου διαλόγου πιθανόν να εκφυλιστεί σε παράλληλα ασύνδετα νήματα μονολόγων, ή έστω χαλαρά συνδεδεμένων διαλόγων, ενώ ο πρωταρχικός σκοπός είναι να αποτελέσει διαδραστικό, επικοινωνιακό εργαλείο.

Οι διδάσκοντες, προκειμένου, να αντιμετωπίσουν την έλλειψη διαδραστικότητας, συχνά απαιτούν από τους σπουδαστές να κάνουν αναρτήσεις, εκ των οποίων κάποιες να ξεκινούν νέο νήμα συζήτησης και κάποιες να σχολιάζουν αναρτήσεις άλλων σπουδαστών. Ωστόσο, ακόμα και έτσι, δεν μπορεί να διασφαλιστεί ότι οι αναρτήσεις θα είναι ουσιαστικές και ότι η αλληλεπίδραση θα έχει βάθος και ποιότητα. Επίσης, δεν μπορεί να διασφαλιστεί ότι τα σχόλια επί των αναρτήσεων των άλλων σπουδαστών, θα είναι απολύτως σχετικά. Αντιθέτως, είναι πιθανό, τα νήματα των συζητήσεων να περιλαμβάνουν αρκετές άσχετες, ή ελάχιστα σχετικές απαντήσεις, με αποτέλεσμα να προκύπτει πρόχειρη, διεκπεραιωτική, δηλαδή απατηλή αλληλεπίδραση, η οποία απέχει πολύ από το επιδιωκόμενο εύρος και βάθος που θα προσδώσει ποιότητα στη συζήτηση. Το επόμενο αντίμετρο, που είναι δυνατόν να εφαρμόσουν οι διδάσκοντες για να διασφαλίσουν τη διακίνηση ουσιαστικών και ποιοτικών σχολίων, είναι να απαιτήσουν τη συμπερίληψη βιβλιογραφικών αναφορών για κάθε ανάρτηση. Η υλοποίηση της συγκεκριμένης απαίτησης προσδίδει στα σχόλια κάποιο βαθμό ακαδημαϊκής εγκυρότητας και εν μέρει εξασφαλίζει ότι οι σπουδαστές ερευνούν και στοχάζονται, προκειμένου να προβούν στο σχολιασμό των ζητημάτων. Ωστόσο, το ΣΔΜ αδυνατεί να διακρίνει τις βιβλιογραφικές αναφορές από οποιοδήποτε άλλο κείμενο, οπότε η ευθύνη και ο φόρτος εργασίας για την καταμέτρηση των έγκυρων αναρτήσεων και σχολίων μεταφέρεται στον διδάσκοντα.

Τις αδυναμίες της ποσοτικής, αντικειμενικής μεθόδου, επιχειρεί να υπερβεί η ποιοτική, υποκειμενική μέθοδος. Η αξιολόγηση βασίζεται σε κριτήρια με τα οποία αποτιμώνται ποιοτικά χαρακτηριστικά των αναρτήσεων, όπως η σχετικότητα με το θέμα συζήτησης, η δομή της ανάρτησης, η αρτιότητα του επιχειρήματος, η τεκμηρίωση του συλλογισμού, η εκφραστική ποιότητα, η συνεισφορά στη μαθησιακή κοινότητα. Με αυτή τη μέθοδο, αξιολογούνται ανώτερες γνωστικές ικανότητες που σχετίζονται με την ανάλυση, τη σύνθεση, την αποτίμηση. Οι σπουδαστές προκειμένου να ικανοποιήσουν τα κριτήρια της μεθόδου, αυξάνουν το επίπεδο στην ποιότητα της συμμετοχής τους, με αποτέλεσμα την αύξηση της αλληλεπίδρασης, της αποδοτικότητας και της ποιότητας της διαδικασίας. Δεδομένου ότι η διαδικασία είναι δύσκολο να αυτοματοποιηθεί, το τίμημα για τα εμφανή πλεονεκτήματα της μεθόδου είναι η αύξηση του χρόνου και της προσπάθειας που πρέπει να καταβάλλει ο διδάσκοντας για να αξιολογήσει τις

αναρτήσεις, των οποίων το πλήθος είναι της τάξεως των εκατοντάδων για τις τυπικές τάξεις που αποτελούνται από δεκάδες σπουδαστές (Shaul, 2008).

#### 4.6. Σοβαρά Παιχνίδια και Προσομοιώσεις

Η ενιαία προσέγγιση και μελέτη των σοβαρών παιχνιδιών (serious games) και των προσομοιώσεων (simulations), ως εργαλείων ψηφιακής αξιολόγησης, είναι συνήθης στη βιβλιογραφία (Bellotti, Kapralos, Lee, Moreno-Ger & Berta, 2013. Crisp, 2013. Vlachopoulos & Makri, 2017).

Σύμφωνα με τον Φεσάκη (2019), οι προσομοιώσεις και τα παιχνίδια συχνά συγχέονται, ωστόσο δεν είναι όλα τα παιχνίδια προσομοιώσεις, ούτε όλες οι προσομοιώσεις παιχνίδια. Η προσομοίωση είναι πλήρης αν υπάρχει αντιστοιχία με την πραγματικότητα, πιστότητα, ακρίβεια και εγκυρότητα. Αλλά, στα παιχνίδια πρέπει επιπλέον να υπάρχουν *«παίχτες, ψυχαγωγία, διασκέδαση, ανταγωνισμός, πρόκληση, περιέργεια, φαντασία, συγκρούσεις, κανόνες, προκαθορισμένες καταστάσεις, στόχοι και συνδεδεμένο σύστημα επιβράβευσης»* (Φεσάκης, 2019, σ.136). Οι Vlachopoulos & Makri (2017) αναφέρουν ότι οι δύο όροι ούτε είναι πλήρως συνώνυμοι, ούτε εντελώς διαφορετικοί και θεωρούνται ως σημεία κατά μήκος της γραμμής του ίδιου πολυδιάστατου φάσματος (Aldrich, 2009), δεδομένου ότι συγκαταλέγονται μεταξύ των διαδραστικών εικονικών περιβαλλόντων που αξιοποιεί η ψηφιακή μάθηση. Αναφέρουν επίσης ότι η κύρια διαφορά, μεταξύ των παιχνιδιών και των προσομοιώσεων, είναι ότι τα παιχνίδια συνιστούν τεχνητά, παιδαγωγικά εργαλεία που περιλαμβάνουν κανόνες, προδιαγεγραμμένους στόχους, αντιτιθέμενες στρατηγικές, ενώ οι προσομοιώσεις είναι δυναμικά τεχνητά περιβάλλοντα που αναπαριστούν την πραγματικότητα, επιδιώκοντας πιστότητα, ακρίβεια και εγκυρότητα.

«Σοβαρά» είναι τα παιχνίδια τα οποία έχουν ουσιαστική σχέση με την απόκτηση γνώσεων και των οποίων ο κύριος σκοπός δεν είναι η ψυχαγωγία (Sawyer, 2002. Zyda, 2005). Προορίζονται για να εφαρμοσθούν στην εκπαίδευση και έχουν σαφείς μαθησιακούς στόχους (Φεσάκης, 2019, σ.156). Ωστόσο, τα σοβαρά παιχνίδια είναι σχεδιασμένα ώστε να είναι εξίσου ελκυστικά με τα «εμπορικά» και ψυχαγωγικά παιχνίδια (Bellotti et al., 2013). Τα σοβαρά παιχνίδια, τα εκπαιδευτικά παιχνίδια και οι εικονικοί κόσμοι που σχεδιάζονται για εκπαιδευτικούς σκοπούς προσφέρουν δυνατότητες για ενεργή συμμετοχή και παρακίνηση πέρα από ψυχαγωγικές δραστηριότητες (Anderson et al. 2009). Γενικά, η μάθηση που βασίζεται σε παιχνίδια παρέχει



δυνατότητες για τη διανομή, την υποστήριξη και την ενίσχυση της μάθησης και της αξιολόγησης (Connolly, 2007).

Τα συστήματα προσομοιώσεων δημιουργούν εικονικά περιβάλλοντα βασισμένα σε σενάρια, τα οποία μιμούνται τον πραγματικό κόσμο και τα οποία μπορούν να χρησιμοποιηθούν ως εκπαιδευτικά εργαλεία, προκειμένου οι σπουδαστές να εφαρμόσουν τις γνώσεις τους, να ασκήσουν τις δεξιότητές τους, να υποδυθούν χαρακτήρες, να πειραματισθούν με διαφορετικές στρατηγικές στα πλαίσια ελεγχόμενων και ασφαλών συνθηκών, να παρακολουθήσουν τα αποτελέσματα των αποφάσεων και των ενεργειών τους. Συχνά, οι προσομοιώσεις ενσωματώνονται σε εκπαιδευτικά ψηφιακά παιχνίδια και θεωρούνται ιδανικά εργαλεία για τη μετάδοση γνώσης σε εργασιακούς χώρους και για την εκπαίδευση στη λήψη αποφάσεων σε πολύπλοκες και δυναμικές καταστάσεις. Επιπλέον, θεωρούνται καινοτόμα και υψηλής αξίας εργαλεία αξιολόγησης για τις περιπτώσεις όπου οι σπουδαστές πρέπει να εξασκήσουν τις δεξιότητές τους και να χρησιμοποιήσουν τις γνώσεις τους σε καταστάσεις που μοιάζουν πολύ με τις αυθεντικές (Karsalis et al., 2019). Επιτρέπουν τον έλεγχο υποθέσεων και τη διερεύνηση σεναρίων του τύπου «τι θα γίνει αν» (what if) (Φεσάκης, 2019, σ. 137).

Τόσο τα σοβαρά παιχνίδια, όσο και οι προσομοιώσεις διευκολύνουν τους σπουδαστές να κατανοήσουν θεωρητικές έννοιες και να τις συσχετίσουν με ρεαλιστικές καταστάσεις, υπό την προϋπόθεση ότι τα εν λόγω καινοτόμα εργαλεία συμπληρώνουν παραδοσιακές εκπαιδευτικές πρακτικές. Ιδιαίτερως, οι προσομοιώσεις ενισχύουν την ενεργή και συνεργατική συμμετοχή των σπουδαστών και συνεισφέρουν στη βελτίωση της κριτικής σκέψης και των ανώτερων γνωστικών διαδικασιών. Επιπλέον, παρέχουν στους σπουδαστές τη δυνατότητα να παρακολουθήσουν τα αποτελέσματα των ενεργειών τους, να λάβουν αποφάσεις, να επιλύσουν προβλήματα και γενικά να αποκτήσουν γνώσεις μέσω ενεργής, βιωματικής διαδικασίας (Vlachopoulos & Makri, 2017).

Σύμφωνα με τους Cook et al. (2013), οι προσομοιώσεις ενισχύουν την αποτελεσματικότητα του εκπαιδευτικού σχεδιασμού, μέσω των ακόλουθων χαρακτηριστικών: παραλλαγές στην πολυπλοκότητα της ανατιθέμενης αποστολής, επαναλαμβανόμενη και διαρκής εξάσκηση, ενεργή συμμετοχή, χρήση πολλαπλών στρατηγικών μάθησης, προσαρμογή στις μαθησιακές ανάγκες του σπουδαστή, υψηλής ποιότητας μάθηση για πεδία απόδοσης με καλά καθορισμένα

πρότυπα, παροχή ανατροφοδότησης κατά τη διάρκεια ή μετά την ολοκλήρωση της προσομοίωσης, δυνατότητα για χρονικά μεγαλύτερη πρακτική άσκηση. Οι Arias-Aranda, Domiguez & Martinez (2010) υποστηρίζουν ότι οι προσομοιώσεις βοηθούν τον διδάσκοντα να παρατηρήσει τη συμπεριφορά των σπουδαστών και να αποτιμήσει τις δυνατότητες και τις στάσεις τους κατά τη διάρκεια ρεαλιστικών δοκιμασιών που απαιτούν λήψη αποφάσεων.

Οι Karsalis et al. (2019) αναφέρονται σε μορφές διαμορφωτικής αξιολόγησης, που αξιοποιούν συστήματα προσομοίωσης μέσω τεχνικών που βασίζονται κυρίως σε συνεντεύξεις και συζητήσεις, οι οποίες διεξάγονται μετά την ολοκλήρωση της προσομοίωσης. Οι συζητήσεις διεξάγονται μεταξύ δύο ή περισσότερων προσώπων, συνήθως μεταξύ του διδάσκοντα και του σπουδαστή (ή των σπουδαστών), όπου ο διδάσκοντας παρέχει ανατροφοδότηση, η οποία βοηθά τον σπουδαστή να βελτιώσει την απόδοσή του. Στη συζήτηση αναλύονται οι στόχοι που έπρεπε να επιτευχθούν, οι αποφάσεις που λήφθηκαν και οι ενέργειες που υλοποιήθηκαν. Η αξιολόγηση μπορεί να έχει και τη μορφή της αυτοαξιολόγησης, όπου ο σπουδαστής καλείται να διαγνώσει τις δυσκολίες που αντιμετώπισε και τα προβλήματα που προέκυψαν και να προτείνει λύσεις για να βελτιώσει την απόδοσή του. Προφανώς, η επιτυχία της διαδικασίας εξαρτάται από παράγοντες, όπως ο κατάλληλος σχεδιασμός, η χρονική αμεσότητα και η ποιότητα της ανατροφοδότησης και ο βαθμός της ενεργής συμμετοχής του σπουδαστή. Ιδιαίτερη αναφορά γίνεται από τους Karsalis et al. (2019) σε χρήση των προσομοιώσεων ως εργαλείων αξιολόγησης στις επιστήμες της υγείας, όπου είναι δυνατή η αποτίμηση των γνώσεων, των δεξιοτήτων και γενικά της απόδοσης των σπουδαστών χωρίς να τίθεται σε κίνδυνο η υγεία των ασθενών. Οι Nicol & Milligan (2006) χαρακτηρίζουν τις προσομοιώσεις ως αποτελεσματικό εργαλείο αυτοαξιολόγησης και αυτορρύθμισης, διότι οι σπουδαστές λαμβάνουν άμεση και δυναμική ανατροφοδότηση για τα αποτελέσματα των ενεργειών τους.

Την άποψη ότι *«η αξιολόγηση μπορεί να ενσωματωθεί στην μαθησιακή διαδικασία εντός εικονικών περιβάλλοντων και παιχνιδιών»* υποστηρίζει η Redecker (2013, σ. 34), η οποία αναλυτικότερα αναφέρει ότι τα δεδομένα που συλλέγονται στις προσομοιώσεις είναι χρήσιμα, όχι μόνο για να μετρήσουν την ορθότητα κάποιας απάντησης, αλλά και για να αποτιμήσουν την ικανότητα του σπουδαστή να εφαρμόζει δεξιότητες, να επιλύει προβλήματα, να χρησιμοποιεί αποδοτικά πληροφορίες και εργαλεία, να λαμβάνει ορθές αποφάσεις. Οι προσομοιώσεις

αποτελούν σημαντικό εργαλείο για την αξιολόγηση των επιστημονικών γνώσεων με τρόπο ρεαλιστικό και προσανατολισμένο στην επίλυση προβλημάτων. Οι τεχνικές εξελίξεις στις προσομοιώσεις και στην συλλογή και ανάλυση δεδομένων, τα οποία σχετίζονται με τη συμπεριφορά και τις ενέργειες του σπουδαστή σε εικονικά περιβάλλοντα, καθιστούν δυνατό τον σχηματισμό μιας μερικής έστω άποψης για τις γνώσεις, τον τρόπο σκέψης και τις ικανότητες του και μπορούν να αξιοποιηθούν ως βάση διαμορφωτικής και τελικής αξιολόγησης. Ειδικότερα, η ικανότητα του σπουδαστή να διενεργεί επιστημονικά πειράματα, να καταγράφει πειραματικά δεδομένα και ακολούθως να τα αναλύει και να εξάγει σχετικά πορίσματα, μπορεί να αξιολογηθεί αποδοτικότερα μέσω προσομοιώσεων και εικονικών εργαστηρίων, όπου είναι εύκολη η αύξηση στην ποικιλία και στο εύρος των πειραμάτων. Γενικά, η Redecker (2013) υποστηρίζει ότι οι προσομοιώσεις, οι εικονικοί κόσμοι και τα παιχνίδια παρέχουν δυνατότητες για την αξιολόγηση (κυρίως διαμορφωτική) πολύπλοκων δεξιοτήτων, όπως η επιστημονική έρευνα, η ανάλυση, η ερμηνεία, ο στοχασμός. Ωστόσο, όπως σημειώνει η ίδια και όπως άλλωστε προκύπτει από τον μικρό αριθμό παραδειγμάτων που αναφέρει, οι συγκεκριμένες προοπτικές δεν έχουν πλήρως διερευνηθεί και δοκιμασθεί εντός των εκπαιδευτικών ιδρυμάτων.

Ευρεία αναφορά για την αξιοποίηση των σοβαρών παιχνιδιών ως εργαλείων αξιολόγησης γίνεται από τους Bellotti et al. (2013) οι οποίοι υποστηρίζουν ότι τα παιχνίδια πρέπει να περιλαμβάνουν και λειτουργίες αξιολόγησης, έστω και αν πρόκειται για ιδιαιτέρως απαιτητική διαδικασία. Περιγράφουν τρεις μορφές αξιολόγησης: την αξιολόγηση ολοκλήρωσης, την εν εξελίξει αξιολόγηση και την αξιολόγηση του διδάσκοντα. Η αξιολόγηση ολοκλήρωσης σχετίζεται με την επιτυχημένη, ή όχι, ολοκλήρωση της δοκιμασίας από τον σπουδαστή. Η εν εξελίξει αξιολόγηση εξετάζει το πώς, το πότε, το γιατί, ο σπουδαστής έλαβε συγκριμένες αποφάσεις για να φέρει σε πέρας τη δοκιμασία. Η αξιολόγηση του διδάσκοντα σχετίζεται με κρίσεις του διδάσκοντα σχετικά τη συμπεριφορά και τη στάση του σπουδαστή με στόχο να αξιολογηθούν παράγοντες, που δεν καταγράφονται στη λογική και στη λειτουργία του παιχνιδιού. Οι Bellotti et al. (2013), υποστηρίζουν ότι καταλληλότερη μορφή αξιολόγησης στα σοβαρά παιχνίδια είναι η διαμορφωτική αξιολόγηση, η οποία πρέπει να ενσωματώνεται στις λειτουργίες των παιχνιδιών μέσω κατάλληλης ανατροφοδότησης. Σημαντικός παράγοντας στην αποδοτικότητα της αξιολόγησης είναι οι πληροφορίες που αφορούν σημαντικές ενέργειες του σπουδαστή και

συμβάντα του παιχνιδιού. Μέσω αυτών των πληροφοριών διενεργείται η ανάλυση και η κατανόηση των αποφάσεων του σπουδαστή και ακολούθως, μετά την ολοκλήρωση του παιχνιδιού, παράχεται σχετική ανατροφοδότηση. Επιπλέον, αναφέρουν ότι μια από τις προκλήσεις στο σχεδιασμό των σοβαρών παιχνιδιών είναι η ενσωμάτωση σε αυτά εναλλακτικών, έμμεσων και λιγότερο προφανών μορφών αξιολόγησης, τις οποίες ο σπουδαστής δεν θα αντιλαμβάνεται ως αξιολόγηση. Για παράδειγμα, ο σπουδαστής θα μπορεί να προχωρήσει σε επόμενο, ανώτερο στάδιο του παιχνιδιού, μόνο όταν οι ενέργειές του αποδεικνύουν ότι κατέχει συγκεκριμένες γνώσεις. Εν κατακλείδι, σημειώνουν ότι η αξιολόγηση μέσω σοβαρών παιχνιδιών και προσομοιώσεων αποτελεί σύνθετο ζήτημα. Ωστόσο, η αύξηση της ισχύος και η μείωση του κόστους των υπολογιστικών συστημάτων παρέχει δυνατότητες αξιολόγησης μέσω της καταγραφής των ενεργειών και της απόδοσης των σπουδαστών. Η αξιολόγηση είναι περισσότερη αποδοτική και χρήσιμη, αν ενσωματώνεται στη λογική της εκπαιδευτικής δραστηριότητας, δεν αποσπά την προσοχή του σπουδαστή και καθιστά εφικτή τη παροχή άμεσης ανατροφοδότησης και την προσαρμογή της διαδικασίας. Γενικά, ο σχεδιασμός της αξιολόγησης πρέπει να περιλαμβάνει σαφείς στόχους και τεχνικές, που θα συλλέξουν πληροφορίες για τον έλεγχο εκπλήρωσης των στόχων (Belloti et al. 2013).

Προερχόμενα από το περιβάλλον και την κουλτούρα των παιχνιδιών (σοβαρών και ψυχαγωγικών), αλλά διαρκώς επεκτεινόμενα και σε άλλους τύπους ψηφιακής και γενικά καινοτόμου αξιολόγησης είναι τα λεγόμενα «ψηφιακά παράσημα» (digital badges) (Karsalis et al., 2019). Πρόκειται για ψηφιακά αντικείμενα που σχετίζονται και πιστοποιούν επιτεύγματα, διακρίσεις, συνεργασίες. Αναρτώνται στο διαδίκτυο και περιλαμβάνουν πληροφορίες (συνήθως με την μορφή διαδικτυακών συνδέσμων), που αναφέρουν το περιεχόμενό τους, τον εκδότη τους, τις διαδικασίες, τις δοκιμασίες και τα αποτελέσματα της δραστηριότητας με την οποία σχετίζονται και την οποία αξιολογούν. Η πρακτική της δημιουργίας, απονομής και επίδειξης των ψηφιακών παρασήμων πηγάζει από πρακτικές ψηφιακών παιχνιδιών, από ψηφιακά συστήματα φήμης και αξιοπιστίας (όπως π.χ. αυτά που χρησιμοποιούν οι πλατφόρμες ψηφιακού εμπορίου για να χαρακτηρίσουν και να αξιολογήσουν τα ψηφιακά καταστήματα), από την κοινότητα των Μέσων ψυχαγωγίας και φυσικά από την παλαιά κοινωνική παράδοση να απονέμονται παράσημα, βραβεία, μετάλλια ως αναγνωριστικά αξιολογών πράξεων.

Ψηφιακά παράσημα είναι δυνατόν να απονεμηθούν σε περιπτώσεις, όπως η επιτυχής ολοκλήρωση ενός ψηφιακού διαγωνίσματος, η συμμετοχή σε ομαδική δραστηριότητα, η επίτευξη ενός στόχου στα πλαίσια μιας τυπικής ή άτυπης αξιολόγησης. Κατόπιν, τα παράσημα συνήθως αναρτώνται και επιδεικνύονται σε προσωπικές ιστοσελίδες ή σε προσωπικά προφίλ στα Μέσα κοινωνικής δικτύωσης προκειμένου να ενισχύσουν τη φήμη του κατόχου τους. Επίσης, τα ψηφιακά παράσημα συχνά συμπεριλαμβάνονται σε ψηφιακούς φάκελους επιτευγμάτων. Στα πλαίσια της εκπαιδευτικής διαδικασίας, τα ψηφιακά παράσημα είναι δυνατόν να κινητοποιήσουν τους σπουδαστές, ώστε να υιοθετήσουν παραγωγικές μαθησιακές συμπεριφορές, να αποτυπώσουν την πρόοδο της μαθησιακής πορείας τους, να πιστοποιήσουν τη συμμετοχή τους και τα επιτεύγματά τους. Προφανώς, η ισχύς και η χρησιμότητα ενός συστήματος εκπαιδευτικών ψηφιακών παρασήμων σχετίζεται με τη γενικότερη στρατηγική αξιολόγησης του εκπαιδευτικού οργανισμού. Εξαρτάται επίσης, από παράγοντες όπως η αξιοπιστία του εκδότη και από το πλήθος των χρηστών και των φορέων που τα αναγνωρίζουν, τα αποδέχονται και τα αξιοποιούν (Gibson, Ostashewski, Flintoff, Grant & Knight, 2015).

## 5. ΜΑΘΗΣΙΑΚΗ ΑΝΑΛΥΤΙΚΗ

### 5.1. Ορισμός και επισκόπηση πεδίου

Όπως συμβαίνει συχνά στα νέα ερευνητικά πεδία, έχουν κατά καιρούς προταθεί διαφορετικοί ορισμοί για τη Μαθησιακή Αναλυτική. Κάποιοι από τους προτεινόμενους ορισμούς εστιάζουν αποκλειστικά στη χρήση των δεδομένων των σπουδαστών με σκοπό τη πρόβλεψη μαθησιακών αποτελεσμάτων και την κατάλληλη προσαρμογή της εκπαιδευτικής διαδικασίας, ενώ άλλοι ορισμοί προσεγγίζουν τη Μαθησιακή Αναλυτική ως μέσο που βοηθά τους διδάσκοντες να εξετάζουν, να κατανοούν και να υποστηρίζουν τις μαθησιακές συμπεριφορές των σπουδαστών και να προσαρμόζουν τα μαθησιακά περιβάλλοντα (Viberg, Hatakka, Balter & Mavroudi, 2018). Ο πλέον αποδεκτός ορισμός είναι αυτός που προτάθηκε στο 1ο διεθνές συνέδριο για τη Μαθησιακή Αναλυτική (International Conference on Learning Analytics & Knowledge (LAK), 2011), σύμφωνα με τον οποίον *«Μαθησιακή Αναλυτική είναι το πεδίο της καταγραφής, συλλογής, ανάλυσης και αναφοράς επί των δεδομένων των μαθητών και του περικειμένου τους, με στόχο την κατανόηση και τη βελτίωση της μάθησης και των περιβαλλόντων μέσα στα οποία λαμβάνει χώρα»* (Siemens & Gasevic, 2012).

Σύμφωνα με τη Ferguson (2012, σ.3), ο συγκεκριμένος ορισμός μπορεί δυνητικά να καλύψει την πλειονότητα των κατευθύνσεων της εκπαιδευτικής έρευνας, οπότε, προκειμένου να προσδιοριστεί με σαφήνεια η Μαθησιακή Αναλυτική, ο ορισμός είναι απαραίτητο να συνοδευτεί από δύο προϋποθέσεις. Πρώτον, ότι η Μαθησιακή Αναλυτική αξιοποιεί προϋπάρχοντα δεδομένα που μπορούν να τύχουν αυτοματοποιημένης επεξεργασίας και δεύτερον, ότι οι τεχνικές της Μαθησιακής Αναλυτικής μπορούν να διαχειριστούν τεράστια σύνολα δεδομένων (big data), τα οποία είναι πρακτικώς αδύνατο για οποιονδήποτε να τα διαχειριστεί χειρακτικώς (manually).

Επιπλέον, όπως εύστοχα επισημαίνει η Δημητρακοπούλου (2017, σ. 5), ο ορισμός δεν αποσαφηνίζει επαρκώς την έννοια του «περιβάλλοντος», το οποίο δεν πρέπει να περιορίζεται σε στενά τεχνολογικά πλαίσια, αλλά πρέπει να *«διευρύνεται στην ομάδα, στην τάξη, στην εκπαιδευτική μονάδα, στην εκπαιδευτική περιφέρεια, στο εθνικό εκπαιδευτικό σύστημα κ.λ.π.»*. Επίσης, είναι αναγκαίες, α) η διεύρυνση της έννοιας των συμμετεχόντων, διότι δεν είναι μόνο

σπουδαστές και β) η ρητή αναφορά σε εκπαιδευτική διαδικασία. Οπότε, προτείνει την ακόλουθη διεγερμένη διατύπωση, την οποία υιοθετούμε: *«Μαθησιακή Αναλυτική είναι το πεδίο της καταγραφής, συλλογής, ανάλυσης και αναφοράς επί των δεδομένων των συμμετεχόντων σε μια εκπαιδευτική διαδικασία και του περικειμένου αυτών, με στόχο την κατανόηση και τη βελτίωση της μάθησης και των περιβαλλόντων μέσα στα οποία λαμβάνει χώρα».*

Η Μαθησιακή Αναλυτική είναι διεπιστημονικό πεδίο με βασικούς επιστημονικούς άξονες, αφενός τις επιστήμες που σχετίζονται με την «Αναλυτική» (Στατιστική, Επιχειρηματική Νοημοσύνη, Επιχειρησιακή Έρευνα, Εξόρυξη Δεδομένων, Τεχνητή Νοημοσύνη, Οπτικοποίηση Πληροφοριών, Βιβλιομετρική) και αφετέρου με τις «Επιστήμες της Αγωγής» (Γνωσιακή Επιστήμη, Εκπαιδευτική Ψυχολογία, Παιδαγωγική, Διδακτική των Επιστημών, Κοινωνιολογία και Οικονομία της Εκπαίδευσης, Εκπαιδευτική Μηχανική, Τεχνολογίες της Πληροφορίας και της Επικοινωνίας στην Εκπαίδευση) (Δημητρακοπούλου, 2017, σ. 6). Σύμφωνα με τη Ferguson (2012), το επίκεντρο της Μαθησιακής Αναλυτικής μετατοπίζεται σταδιακά από τα τεχνολογικά προς τα παιδαγωγικά ζητήματα, όπου όμως τα μοντέλα που χρησιμοποιούνται είναι απαραίτητο να μην αγνοούν την πολυπλοκότητα της σχέσης μεταξύ σπουδαστή και μάθησης (Oldfield, Broadfoot, Sutherland & Timmis, 2012, σ. 31).

Παρόλο που η Μαθησιακή Αναλυτική βασίζεται σε καλά εδραιωμένα, κλασικά επιστημονικά πεδία διαφέρει από τις παραδοσιακές εκπαιδευτικές αναλύσεις για μια σειρά από λόγους. Πρώτον, ο μεγάλος όγκος των συνόλων δεδομένων που μελετά, ενισχύει την πεποίθηση ότι τα πορίσματα είναι γενικεύσιμα. Δεύτερον, επειδή τα δεδομένα συλλέγονται ως επί το πλείστον από ψηφιακά συστήματα, επιτυγχάνεται μεγάλος βαθμός λεπτομέρειας στην καταγραφή των υπό μελέτη παραγόντων, ο οποίος δεν είναι εύκολο να επιτευχθεί με παραδοσιακές τεχνικές παρατήρησης (π.χ. συνεντεύξεις). Τρίτον, τα δεδομένα είναι κατά κύριο λόγο χρονικώς διατεταγμένα, εξαιτίας του γεγονότος ότι στις μεθόδους και στα εργαλεία συλλογής δεδομένων κυριαρχεί η χρονική διάσταση. Βασικό χαρακτηριστικό της Μαθησιακής Αναλυτικής είναι ότι θεωρείται εφαρμοσμένη επιστήμη, διότι, αξιοποιώντας διεπιστημονικές προσεγγίσεις, επιδιώκει την κατανόηση και τη βελτιστοποίηση της διαδικασίας της μάθησης (Joksimovic, Kovanovic & Dawson, 2019, σ. 40). Παράλληλα, θεωρείται και παρεμβατική επιστήμη, διότι

μπορεί να επιφέρει κοινωνική αλλαγή, διαμέσου ποικίλων ερευνητικών στόχων και ευρημάτων από τις μαθησιακές ως τις πολιτικές στρατηγικές (Δημητρακοπούλου, σ. 27).

Η Ferguson (2012, σ. 4) καταγράφει τέσσερις παράγοντες που καθοδηγούν τις εξελίξεις στην Μαθησιακή Αναλυτική, τους οποίους συσχετίζει με ισάριθμες αντίστοιχες προκλήσεις. Πρώτος παράγοντας είναι τα τεράσια σύνολα δεδομένων, των οποίων ο όγκος υπερβαίνει τις δυνατότητες διαχείρισης από κλασικές εφαρμογές λογισμικού και κλασικές βάσεις δεδομένων. Ο πρώτος παράγοντας σχετίζεται με την πρόκληση: *Πώς μπορεί να προκύψει όφελος από τα δεδομένα*. Δεύτερος παράγοντας είναι η, συνεχώς διευρυνόμενη, χρήση της διαδικτυακής εκπαίδευσης (online learning), η οποία παρέχει πλεονεκτήματα, αλλά δεν είναι αμιγής προβλημάτων, όπως για παράδειγμα οι δυσκολίες προσαρμογής στα τεχνολογικά εκπαιδευτικά περιβάλλοντα, οι τεχνικές δυσλειτουργίες, η απουσία προσωπικής επαφής μεταξύ σπουδαστών και διδασκόντων και η μειωμένη ποιότητα επικοινωνίας συγκριτικά με την «πρόσωπο με πρόσωπο» διδασκαλία. Ο δεύτερος παράγοντας σχετίζεται με την πρόκληση: *Πώς θα βελτιστοποιηθούν οι δυνατότητες της διαδικτυακής μάθησης*. Τρίτος παράγοντας είναι οι ευρύτερες πολιτικές συνθήκες που απαιτούν από τα εκπαιδευτικά ιδρύματα να καταμετρούν, να παρουσιάζουν και να βελτιώνουν την απόδοσή τους. Ο τρίτος παράγοντας σχετίζεται με την πολιτική και παράλληλα οικονομική πρόκληση: *Πώς θα βελτιστοποιηθεί η μάθηση και τα εκπαιδευτικά αποτελέσματα σε εθνικό και διεθνές επίπεδο*. Τέταρτος παράγοντας είναι οι επιδιώξεις των τριών εμπλεκόμενων ομάδων, δηλαδή των κυβερνήσεων, των εκπαιδευτικών ιδρυμάτων και των διδασκόντων - σπουδαστών. Οι επιδιώξεις των ομάδων συχνά επικαλύπτονται, είτε μερικώς, είτε απολύτως, ωστόσο οδηγούν σε ερευνητικές προσεγγίσεις, που διαφέρουν ως προς την κλίμακα και ως προς τον βαθμό ανάλυσης. Ο τέταρτος παράγοντας σχετίζεται με την πρόκληση: *Ποιος οφελείται από τη Μαθησιακή Αναλυτική*.

Πρόσφορο έδαφος για τη Μαθησιακή Αναλυτική δημιουργεί η συνεχής, ραγδαία αύξηση των δεδομένων τα οποία παράγονται και αποθηκεύονται στα συστήματα, που υποστηρίζουν τις λειτουργίες των εκπαιδευτικών οργανισμών. Τα συστήματα αυτά, έως πρόσφατα, κατέγραφαν και διαχειρίζονταν μόνο δημογραφικά, οικονομικά και ακαδημαϊκά δεδομένα των σπουδαστών. Όμως, οι συνήθεις πλέον διαδικτυακές και μεικτές μορφές διδασκαλίας, που υποστηρίζονται από τις ΤΠΕ και ειδικότερα από τα ψηφιακά εκπαιδευτικά περιβάλλοντα, έχουν ως αποτέλεσμα



την παραγωγή και την αποθήκευση δεδομένων, τα οποία καταγράφουν τις αλληλεπιδράσεις των σπουδαστών με το εκπαιδευτικό υλικό, με τις εκπαιδευτικές δραστηριότητες, με τους άλλους σπουδαστές, με τους διδάσκοντες. Πρόκειται για δεδομένα που δεν σχετίζονται με δημογραφικά, οικονομικά και ακαδημαϊκά χαρακτηριστικά των σπουδαστών, αλλά για δεδομένα που καταγράφουν τις μαθησιακές συμπεριφορές και τις δραστηριότητές τους και γενικότερα καταγράφουν διαστάσεις της εκπαιδευτικής διαδικασίας. Όπως αναφέρουν οι Sclater, Peasgood & Mullan (2016, σ. 17), η κύρια πηγή των δεδομένων που αξιοποιεί η Μαθησιακή Αναλυτική είναι τα Συστήματα Διαχείρισης Μάθησης (ΣΔΜ), τα οποία οι σπουδαστές χρησιμοποιούν για να έχουν πρόσβαση σε εκπαιδευτικό υλικό, για να συμμετάσχουν σε εκπαιδευτικές δραστηριότητες, για να αναρτούν εργασίες, για να επιλύουν ασκήσεις, για να συμμετέχουν σε δημόσιους διαλόγους, για να διαβάζουν ανακοινώσεις. Ο όγκος των δεδομένων που παράγονται εξαρτάται από τον βαθμό χρήσης των ΣΔΜ κατά την εκπαιδευτική διαδικασία. Αν τα ΣΔΜ αξιοποιούνται λίγο, τότε ο όγκος των δεδομένων είναι μικρός. Όμως, αν τα ΣΔΜ αποτελούν βασικό μέσο για τη διεξαγωγή των μαθημάτων, όπως για παράδειγμα σε εξ' αποστάσεως ακαδημαϊκά προγράμματα και σε μαζικά ανοικτά διαδικτυακά μαθήματα (ΜΟΟC), τότε ο όγκος των δεδομένων είναι τεράστιος. Ωστόσο, ακόμα και στις περιπτώσεις που η χρήση των ΣΔΜ δεν είναι συχνή, η Μαθησιακή Αναλυτική μπορεί να αξιοποιήσει δεδομένα και από άλλες πηγές. Για παράδειγμα, μπορεί να αξιοποιήσει δεδομένα από συστήματα φοιτητολογίων που περιλαμβάνουν δημογραφικά, οικονομικά και ακαδημαϊκά δεδομένα των σπουδαστών ή από συστήματα βιβλιοθηκών που καταγράφουν το πλήθος και το είδος των βιβλίων που χρησιμοποίησαν οι σπουδαστές. Γενικά, τα δεδομένα προέρχονται από ενέργειες και δραστηριότητες στα πλαίσια τεχνολογικών και διαδικτυακών εφαρμογών και υπηρεσιών, από ψηφιοποίηση εντύπων και αντικειμένων ή και από βιο-φυσιολογικές ενέργειες (π.χ. ομιλία, κίνηση) (Δημητρακοπούλου, 2017, σ. 18).

Αναλόγα με το υποκείμενο στο οποίο απευθύνονται και τον σκοπό που υπηρετούν, οι εφαρμογές της Μαθησιακής Αναλυτικής μπορούν να κατηγοριοποιηθούν σε εφαρμογές μάθησης (όταν απευθύνονται στον σπουδαστή), σε εφαρμογές διδασκαλίας (όταν απευθύνονται στον διδάσκοντα), σε εφαρμογές εκπαιδευτικής διοίκησης (όταν απευθύνονται στη διοίκηση των εκπαιδευτικών οργανισμών) και σε εφαρμογές έρευνας (όταν απευθύνονται

σε ερευνητές που μελετούν ζητήματα μαθησιακών και εκπαιδευτικών διαδικασιών). Η εν λόγω κατηγοροποίηση μπορεί να έχει και δεύτερο επίπεδο ανάλυσης, όπου κριτήριο αποτελεί ο σκοπός της εφαρμογής. Συγκεκριμένα, οι εφαρμογές που απευθύνονται στους μαθητές αφορούν συστήματα επίγνωσης, αναστοχασμού και αυτορρύθμισης, συστήματα υποδείξεων, προσαρμοστικά συστήματα και συστήματα αυτοαξιολόγησης. Οι εφαρμογές που απευθύνονται στους διδάσκοντες αφορούν συστήματα διαχείρισης διδακτικών στρατηγικών, συστήματα αξιολόγησης μαθητών, συστήματα επίγνωσης και αναστοχασμού διδακτικών παρεμβάσεων. Οι εφαρμογές που απευθύνονται στις ομάδες διοίκησης αφορούν συστήματα διοίκησης, συστήματα διαχείρισης πόρων, συστήματα προγραμματισμού και αξιολόγησης εκπαιδευτικών μονάδων, συστήματα διαχείρισης δημογραφικών δεδομένων εκπαιδευτικών μονάδων (Δημητρακοπούλου, 2017, σ. 10, 15).

Όσες από τις εφαρμογές αφορούν τις διοικητικές λειτουργίες των εκπαιδευτικών μονάδων και αξιοποιούν δεδομένα τοπικής, περιφερειακής, εθνικής και διεθνικής κλίμακας κατατάσσονται σε ένα ιδιαίτερο κλάδο της Μαθησιακής Αναλυτικής που φέρει τον τίτλο «**Ακαδημαϊκή Αναλυτική**» (Academic Analytics). Η Ακαδημαϊκή Αναλυτική βρίσκει εφαρμογή στο επίπεδο των εκπαιδευτικών μονάδων στηρίζοντας αποφάσεις της διοίκησης, μέχρι και στο επίπεδο των εθνικών εκπαιδευτικών συστημάτων, στηρίζοντας αποφάσεις που αφορούν εθνικές εκπαιδευτικές πολιτικές (Oldfield, Broadfoot, Sutherland, & Timmis, 2012, σ. 29). Η Μαθησιακή Αναλυτική σχετίζεται επίσης με την «**Εξόρυξη Εκπαιδευτικών Δεδομένων**» (Educational Data Mining), η οποία εστιάζει κυρίως στις τεχνολογίες για την εξαγωγή συμπερασμάτων από μεγάλους όγκους εκπαιδευτικών δεδομένων που παράγονται από σπουδαστές και διδάσκοντες (Linan & Perez, 2015, σ. 100) και αξιοποιεί υπολογιστικές τεχνικές για τον εντοπισμό μοτίβων σε μεγάλες εκπαιδευτικές βάσεις δεδομένων (Oldfield, Broadfoot, Sutherland, & Timmis, 2012, σ. 29). Συχνή εφαρμογή της Μαθησιακής Αναλυτικής και της Εξόρυξης Εκπαιδευτικών Δεδομένων (Linan & Perez, 2015, σ. 103. Lopez-Zambrano, Lara & Romero, 2020, σ.2) είναι τα **Μοντέλα Πρόβλεψης**<sup>1</sup>, τα οποία αναζητούν μοτίβα και συσχετίσεις στα εκπαιδευτικά δεδομένα και

---

<sup>1</sup> Ο όρος «πρόβλεψη» που χρησιμοποιούμε στα πλαίσια της διατριβής είναι η μετάφραση του αγγλικού όρου «prediction», ο οποίος χρησιμοποιείται χωρίς εξαίρεση στη διεθνή βιβλιογραφία για να περιγράψει τη λειτουργία των σχετικών μοντέλων.

επιχειρούν πιθανοκρατικές προβλέψεις για μελλοντικά αποτελέσματα που σχετίζονται με ακαδημαϊκές επιδόσεις ή συμπεριφορές των σπουδαστών. Οι προβλέψεις μπορούν να χρησιμοποιηθούν για την κατάλληλη προσαρμογή των συστημάτων μάθησης και για την υλοποίηση παρεμβάσεων που θα βελτιώσουν την εκπαιδευτική διαδικασία (Oldfield, Broadfoot, Sutherland, & Timmis, 2012, σ. 29). Σύμφωνα με την Ferguson (2012, σ. 9) η Μαθησιακή Αναλυτική εστιάζει σε εκπαιδευτικές προκλήσεις, η Εξόρυξη Εκπαιδευτικών Δεδομένων εστιάζει σε τεχνικές προκλήσεις και η Ακαδημαϊκή Αναλυτική εστιάζει σε πολιτικές και οικονομικές προκλήσεις. Ωστόσο οι επικαλύψεις των πεδίων είναι συχνές.

Σύμφωνα με τους Joksimovic, Kovanovic & Dawson (2019, σ. 43) οι ερευνητικές προσπάθειες στα πλαίσια της Μαθησιακής Αναλυτικής εστιάζουν σε τρεις κατευθύνσεις: στην «προβλεπτική αναλυτική», στην «αναλυτική της κοινωνικής μάθησης» και στην «αναλυτική των συζητήσεων».

Η «προβλεπτική αναλυτική» κυριαρχεί από τα πρώτα χρόνια εμφάνισης του πεδίου. Επιδιώκει την κατασκευή μοντέλων που επιχειρούν να προβλέψουν μαθησιακά αποτελέσματα, όπως η ακαδημαϊκή επίδοση των σπουδαστών και να εντοπίσουν εγκαίρως σπουδαστές που κινδυνεύουν να αποτύχουν. Βασίζεται σε τεχνικές της Εξόρυξης Εκπαιδευτικών Δεδομένων για να καταγράψει και να αναλύσει δεδομένα, που σχετίζονται με δημογραφικά στοιχεία των σπουδαστών, με την ακαδημαϊκή πορεία τους με τις ενέργειες και τις αλληλεπιδράσεις τους στα ΣΔΜ. Αξιοποιεί κυρίως αλγόριθμους Μηχανικής Μάθησης, που παράγουν μοντέλα πρόβλεψης, τα οποία επιχειρούν προβλέψεις με πιθανοκρατικές μεθόδους.

Η «αναλυτική της κοινωνικής μάθησης» εστιάζει στη μάθηση που συντελείται εντός διαδικτυακών κοινοτήτων. Έχει τις απαρχές της στο ερευνητικό πεδίο της «ανάλυσης των κοινωνικών δικτύων» (social network analysis), το οποίο μελετά τα κοινωνικά δίκτυα, τους δρώντες εντός αυτών και τις αλληλεπιδράσεις τους (Oldfield, Broadfoot, Sutherland, & Timmis, 2012, σ. 31) και το οποίο εξελισσόμενο εντός των πλασίων της Μαθησιακής Αναλυτικής επιχειρεί την αυτόματη αποτύπωση μεγάλης κλίμακας δικτύων που δημιουργούν οι αλληλεπιδράσεις οι οποίες συντελούνται σε ΣΔΜ και σε δημοφιλείς διαδικτυακές πλατφόρμες κοινωνικής δικτύωσης. Όπως αναφέρουν οι Joksimovic, Kovanovic & Dawson (2019, σ. 44), η «αναλυτική της κοινωνικής μάθησης» αναλύει τις ψηφιακές κοινωνικές αλληλεπιδράσεις των σπουδαστών με τους άλλους σπουδαστές και με τους διδάσκοντες και ιδιαιτέρως τις

αλληλεπιδράσεις που συντελούνται στους δημόσιους διαδικτυακούς διαλόγους. Επιχειρεί να κατανοήσει τις σχέσεις μεταξύ ψηφιακών κοινωνικών συμπεριφορών και μαθησιακών αποτελεσμάτων ή να μελετήσει κανονικότητες στις αλληλεπιδράσεις που προκύπτουν κατά τη συμμετοχή σε μαθησιακές δραστηριότητες. Επιχειρεί επίσης να αναπτύξει στατιστικά μοντέλα σχετικά με τους παράγοντες και τους κοινωνικούς μηχανισμούς, οι οποίοι επηρεάζουν τον σχηματισμό και τη δυναμική των κοινωνικών δικτύων μεταξύ των σπουδαστών, στα πλαίσια τυπικών και άτυπων περιβαλλόντων μάθησης

Η «αναλυτική των συζητήσεων» μελετά δεδομένα κειμένων, τα οποία παράγονται στα πλαίσια εκπαιδευτικών διαδικασιών. Παρότι είναι δυνατόν να αξιοποιηθεί στη μελέτη κάθε είδους κειμένων, όπως για παράδειγμα εκθέσεις ιδεών, επικεντρώνεται κυρίως σε κείμενα από διαδικτυακές επικοινωνίες σπουδαστών, δηλαδή διαδικτυακούς δημόσιους διαλόγους (Forum), συνομιλίες (chat), αναρτήσεις σε κοινωνικά δίκτυα και σε ιστολόγια (blog). Επιδιώκει να καταγράψει τις κοινωνικές διαστάσεις της μάθησης, αναλύοντας τα κείμενα με τη χρήση τεχνικών επεξεργασίας της φυσικής γλώσσας και αλγορίθμων Μηχανικής Μάθησης (Joksimovic, Kovanovic & Dawson, 2019, σ. 44).

Καθώς η Μαθησιακή Αναλυτική εξελίσσεται και εδραιώνεται ως διεπιστημονικό ερευνητικό πεδίο, τίθεται αντιμέτωπη με μια σειρά από προκλήσεις μεταξύ των οποίων η Ferguson (2012, σ. 10) διαβλέπει τις εξής:

- Τη σύναψη ισχυρής σύνδεσης με τις επιστήμες της μάθησης. Συγκεκριμένα, οι μελέτες που αφορούν τη Μαθησιακή Αναλυτική σε σχέση με τις γνωστικές και τις μεταγνωστικές λειτουργίες και την παιδαγωγική δεν είναι πολλές και πρέπει να ενισχυθούν. Ο στόχος για τη βελτιστοποίηση της μάθησης δεν μπορεί να προσεγγιστεί, αν δεν είναι κατανοητές οι διαδικασίες με τις οποίες συντελείται η μάθηση. Το όφελος που θα προκύψει από τη σύνδεση είναι ότι η Μαθησιακή Αναλυτική θα βοηθήσει στον καλό εκπαιδευτικό σχεδιασμό, στις αποτελεσματικές διδακτικές παρεμβάσεις και στην αυτεπίγνωση των σπουδαστών.
- Την ανάπτυξη μεθόδων μελέτης για ευρύ φάσμα συνόλων δεδομένων, ώστε να βελτιστοποιηθούν τα εκπαιδευτικά περιβάλλοντα. Συγκεκριμένα, οι μέθοδοι θα κληθούν

να διαχειριστούν δεδομένα που θα αφορούν όχι μόνο σπουδαστές που χρησιμοποιούν ΣΔΜ, αλλά και σπουδαστές που συμμετέχουν σε ανοικτά, άτυπα ή μεικτά εκπαιδευτικά περιβάλλοντα. Συνεπώς θα προκύψει ανάγκη διαχείρισης πολύπλοκων συνόλων δεδομένων, που θα περιλαμβάνουν βιομετρικά δεδομένα, δεδομένα κίνησης και δεδομένα «συναισθηματικής κατάστασης» (mood data).

- Την πρόταξη των προτεραιοτήτων και των αναγκών των σπουδαστών. Συγκεκριμένα, η ανάπτυξη της Μαθησιακής Αναλυτικής πρέπει να θέσει στο επίκεντρο τις ανάγκες των σπουδαστών, όχι των εκπαιδευτικών ιδρυμάτων. Τα κριτήρια της ακαδημαϊκής επιτυχίας πρέπει να διευρυνθούν και να περιλαμβάνουν όχι μόνο βαθμολογικές κλίμακες, αλλά και τα κίνητρα των σπουδαστών, την ικανοποίηση, την απόλαυση, την επίτευξη των επαγγελματικών στόχων τους. Απαιτούνται μέθοδοι, που μέσω διάφανων και ευεξήγητων μηχανισμών, θα προσφέρουν εύληπτες και προσωποποιημένες αναλύσεις, οι οποίες πρέπει να προτείνουν στους σπουδαστές τρόπους βελτίωσης της μάθησής τους.
- Τη θέσπιση και εφαρμογή σαφών οδηγιών για την ηθική χρήση των δεδομένων και των μεθόδων. Συγκεκριμένα, απαιτείται να καθορισθούν κανόνες σχετικά με την ιδιοκτησία και την εποπτεία των δεδομένων, ώστε να είναι σαφές ποια δικαιώματα έχουν οι σπουδαστές επί των δεδομένων που τους αφορούν. Στις περισσότερες περιπτώσεις δεν υπάρχουν θεσπισμένες διαδικασίες για το αν και πώς μπορούν οι ερευνητές να έχουν χρονικά περιορισμένη ή εσαεί πρόσβαση σε δεδομένα και για το αν και πώς οι σπουδαστές μπορούν να ζητήσουν την απάλειψη των δεδομένων τους από τα συστήματα. Παρότι τα σχετικά ζητήματα έχουν επισημανθεί προ πολλού και από πολλούς, υφίσταται υστέρηση στη θέσπιση και στην εφαρμογή κατάλληλου θεσμικού πλαισίου.

Το ζήτημα ηθικής χρήσης των δεδομένων στα πλαίσια της Μαθησιακής Αναλυτικής αποτελεί έκφανση του ευρύτερου ζητήματος της προστασίας των προσωπικών δεδομένων που καταγράφουν τα τεχνολογικά συστήματα κάθε είδους. Τα υφιστάμενα θεσμικά πλαίσια είναι δύσκολο να ρυθμίσουν ένα διαρκώς μεταβαλλόμενο πεδίο. Οσο τα ψηφιακά δεδομένα αυξάνονται και αφορούν συνεχώς διευρυνόμενες διαστάσεις του ιδιωτικού και δημόσιου βίου

των χρήστων, τόσο εντονότερα οι πανίσχυροι κρατικοί και επιχειρηματικοί φορείς θα επιδιώκουν να αποκτήσουν ισχύ ή οικονομικό κέρδος από αυτά, οπότε τόσο πιο ευάλωτη θα καθίσταται η προστασία των προσωπικών δεδομένων. Στο χώρο των εκπαιδευτικών ιδρυμάτων και ως προς τις εφαρμογές της Μαθησιακής Αναλυτικής, παρατηρείται είτε έλλειψη, είτε ανεπάρκεια, είτε μη τήρηση των κανόνων δεοντολογίας με ποικίλες συνέπειες ως προς την εμπέλεια και το βαθμό επικινδυνότητας τους. Στο επίπεδο της σχολικής ή πανεπιστημιακής τάξης, η ηθική διαχείριση των δεδομένων είναι εφικτή αν τηρούνται οι όροι της συναίνεσης, της διαφάνειας, της πρόσβασης, της υπευθυνότητας, της ιδιωτικότητας, της διαχείρισης του συστήματος, ώστε κατά τη λειτουργία των συστημάτων της Μαθησιακής Αναλυτικής να διασφαλίζεται η προστασία των υποκειμένων και η αβλαβής χρήση των δεδομένων που τους αφορούν (Δημητρακοπούλου, σ. 23).

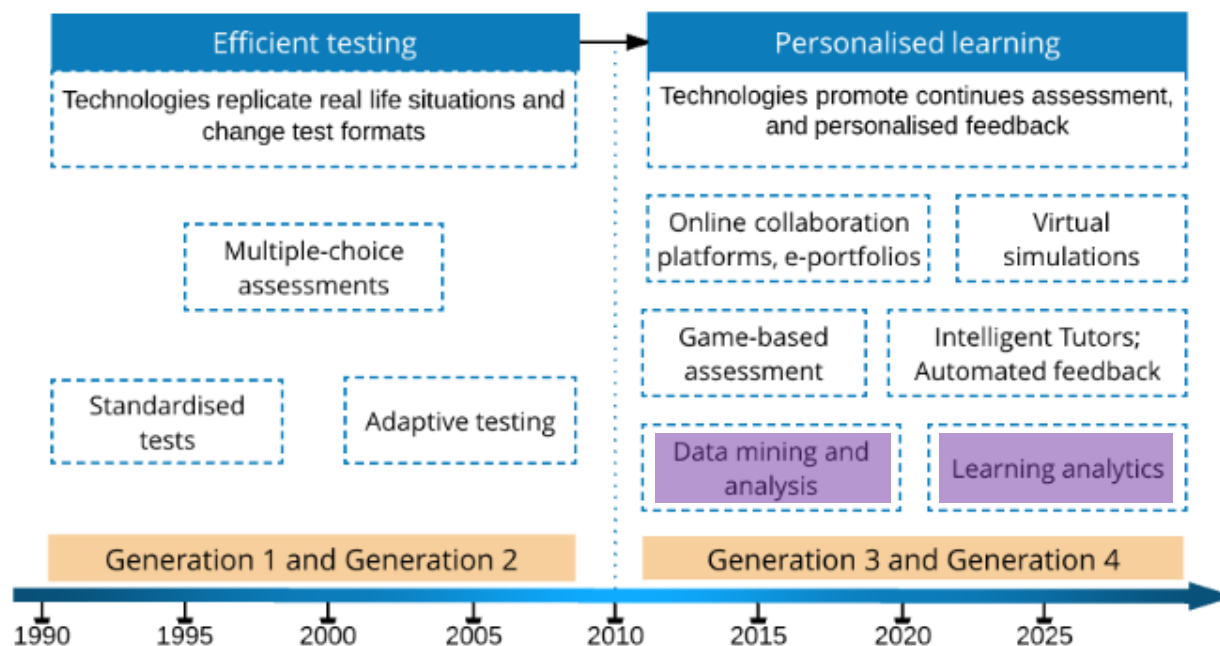
## 5.2. Η μαθησιακή αναλυτική και η αξιολόγηση

Η Redecker (2013, σ. 11), περιγράφοντας την εξέλιξη των συστημάτων ψηφιακής αξιολόγησης, αναφέρει ότι οι σύγχρονες προκλήσεις οδηγούν στην ανάπτυξη μεθόδων αξιολόγησης, που δεν θα αποτελούν διακριτό στάδιο της εκπαιδευτικής διαδικασίας, αλλά θα διαχέονται μέσα σε αυτήν ώστε οι σπουδαστές να αξιολογούνται εμμέσως και διαρκώς. Τεχνολογίες όπως η εξόρυξη δεδομένων, η Μαθησιακή Αναλυτική, η παρακολούθηση μαθησιακής συμπεριφοράς, η αυτοματοποιημένη ανατροφοδότηση, οι ευφυείς καθοδηγητές θα υποστηρίζουν τις νέες γενιές συστημάτων αξιολόγησης, τα οποία θα υλοποιούν συνεχείς και ενσωματωμένες στην εκπαιδευτική διαδικασία αξιολογικές δοκιμασίες, με προσωποποιημένη ανατροφοδότηση και καθοδήγηση (βλ. Κεφ. 2.1, Σχήμα 1).

Σύμφωνα με τη Redecker (2013, σ. 32), συστήματα που συνδυάζουν τεχνολογίες εξόρυξης εκπαιδευτικών δεδομένων χρησιμοποιούνται για να αποτιμήσουν τις συνήθειες δραστηριότητας σπουδαστών σε εκπαιδευτικά περιβάλλοντα. Τα πορίσματα που προκύπτουν είναι δυνατόν να βοηθήσουν στον εντοπισμό «αδύναμων» σπουδαστών, στη σύνταξη διαγνωστικών αναφορών σχετικά με την επίδοση των σπουδαστών, στην καταγραφή και αξιολόγηση των αλληλεπιδράσεων των σπουδαστών, αλλά και στην αναθεώρηση και βελτίωση προγραμμάτων σπουδών. Ωστόσο, η αξιοποίηση της θεωρητικής δυνατότητας για καταγραφή κάθε δραστηριότητας, που συντελείται εντός τεχνολογικών περιβαλλόντων μάθησης, προϋποθέτει την ιεράρχηση και την ανάδειξη των ενεργειών που σχετίζονται με γνωστικές διεργασίες και τον προσδιορισμό των δεδομένων, που οδηγούν στη συγκρότηση πολύτιμων και αξιοποιήσιμων πληροφοριών. Η Redecker (2013, σ. 40) αναφέρει επίσης, ότι η Μαθησιακή Αναλυτική είναι μια από τις πολλά υποσχόμενες νέες τεχνολογικές τάσεις για την αξιολόγηση σύνθετων ικανοτήτων και επιπλέον βοηθά την ενσωμάτωση της αξιολόγησης σε περιβάλλοντα εμπύθισης, σε ομαδικά ψηφιακά εκπαιδευτικά παιχνίδια και σε προσομοιώσεις.

Την ίδια εξελικτική πορεία, η οποία εκκινεί από τα κλασικά υπολογιστικά συστήματα, που υποστηρίζουν κυρίως ερωτήσεις πολλαπλών επιλογών ή σύντομων απαντήσεων και κατευθύνεται προς συνεχείς και ενσωματωμένες στην εκπαιδευτική διαδικασία ψηφιακές αξιολογικές δοκιμασίες, όπου κεντρικό ρόλο κατέχουν η Εξόρυξη Εκπαιδευτικών Δεδομένων και

η Μαθησιακή Αναλυτική, περιγράφουν και οι Siarova, Sternadel & Mavsidlauskaitė (2017, σ. 72) (Σχήμα 4).



Σχήμα 4. Οι εξελικτική πορεία των συστημάτων αξιολόγησης (Siarova, Sternadel & Mavsidlauskaitė, 2017, σ. 72)

Οι Oldfield, Broadfoot, Sutherland & Timmis (2012, σ. 30) υποστηρίζουν ότι με τη Μαθησιακή Αναλυτική είναι δυνατόν να διερευνηθούν μέθοδοι αξιολόγησης, που υπερβαίνουν τον διαχωρισμό μεταξύ διαμορφωτικής και τελικής αξιολόγησης, εξαιτίας των πολλών τρόπων για την καταγραφή και την ανάλυση των δεδομένων και την ενημέρωση των σπουδαστών. Ο στόχος επιτυγχάνεται με εργαλεία, που παρέχουν στους σπουδαστές αναπαραστάσεις των δεδομένων, ικανές να υποστηρίξουν την αυτοαξιολόγηση και την αυτορρύθμιση της μάθησης. Υποστηρίζουν, επίσης, ότι η Μαθησιακή Αναλυτική μπορεί να αξιοποιηθεί σε διαδικασίες ανατροφοδότησης, όπου η ανάλυση των δραστηριοτήτων των σπουδαστών βοηθά, αφενός τους διδάσκοντες να προσαρμόσουν τις διδακτικές παρεμβάσεις τους και να παράσχουν προσωποποιημένη υποστήριξη και αφετέρου τους σπουδαστές να αυτοαξιολογήσουν τις δραστηριότητές τους και να προβούν σε κατάλληλες αλλαγές.



Ως προς την αξιοποίηση της Μαθησιακής Αναλυτικής σε διαδικασίες αξιολόγησης, οι Φεσάκης & Κωνσταντοπούλου (2022, σ. 141) αναφέρουν ότι η σύνδεση της Μαθησιακής Αναλυτικής και της εκπαιδευτικής αξιολόγησης «έχει ανοίξει νέες προοπτικές στην έρευνα και στις εφαρμογές της αξιολόγησης και του μαθησιακού σχεδιασμού δυναμικά προσαρμοσμένων εκπαιδευτικών σεναρίων». Συγκεκριμένα, αναφέρουν ότι η Μαθησιακή Αναλυτική μπορεί να αξιοποιηθεί στην υπολογιστικά υποστηριζόμενη συνεχή διαμορφωτική αξιολόγηση, όπου τα δεδομένα από επιλεγμένα χρονικά σημεία της υλοποίησης των εκπαιδευτικών σεναρίων τροφοδοτούν αυτοματοποιημένες αναλυτικές μεθόδους, που παρέχουν ανατροφοδότηση και γενικότερα πληροφορίες ενημέρωσης στους σπουδαστές και στους διδάσκοντες για την πορεία τους σε σχέση με τους στόχους των εκπαιδευτικών σεναρίων. Επιπλέον, τονίζουν τη σημασία της ισορροπίας μεταξύ της αυτοματοποιημένης και της «χειροκίνητης» συλλογής και επεξεργασίας των δεδομένων, ζήτημα το οποίο αναδεικνύουν και οι Bienkowski, Feng & Means (2021, σ. 40), οι οποίοι αναφέρουν ότι είναι εξαιρετικά σημαντικός ο ρόλος της ανθρώπινης κρίσης ως προς τα δεδομένα που θα συλλεχθούν, ως προς τη διερμηνεία των δεδομένων και ως προς τις αποφάσεις που θα ληφθούν (Oldfield, Broadfoot, Sutherland, & Timmis, 2012, σ. 32)

Η Jordan (2013, σ. 99) θεωρεί ότι η εκπαιδευτική κοινότητα πρέπει να προχωρήσει πέρα από τις κλασικές αξιολογικές διαδικασίες μέσω διαγωνισμάτων (testing paradigm) και να αξιοποιήσει εφαρμογές της Μαθησιακής Αναλυτικής που θα αξιολογούν τους σπουδαστές βάση των αλληλεπίδρασεών τους. Προτείνει επίσης την αξιοποίηση της «Αξιολογικής Αναλυτικής» (assessment analytics), δηλαδή την ανάλυση δεδομένων που παράγονται από αξιολογικές διαδικασίες και σχολιάζει ότι η χρήση της Μαθησιακής Αναλυτικής για αξιολόγηση και παράλληλα, η χρήση της Αξιολογικής Αναλυτικής για τη μελέτη της μάθησης, καθιστούν δυσδιάκριτα τα όρια μεταξύ της αξιολόγησης και της μάθησης.

Οι Oldfield et al. (2012, σ. 30) αναφέρουν ότι η Μαθησιακή Αναλυτική μπορεί να λειτουργήσει ως «βρόχος ανατροφοδότησης» (feedback loop), που παρέχει ανάλυση δραστηριοτήτων και αποδόσεων, έτσι ώστε, αφενός οι διδάσκοντες να αναμορφώσουν τις διδακτικές παρεμβάσεις τους ή να παράσχουν προσωποποιημένη υποστήριξη προς τους σπουδαστές και αφετέρου οι σπουδαστές να αναστοχαστούν σχετικά με την απόδοσή τους και να ενεργήσουν για τη βελτίωσή της. Επιπλέον, αναφέρουν ότι η Μαθησιακή Αναλυτική υποστηρίζει την ανάπτυξη

μεθόδων αξιολόγησης, που αμβλύνουν τις διαφορές μεταξύ διαμορφωτικής και τελικής αξιολόγησης, δεδομένου ότι υφίστανται πολλοί τρόποι με τους οποίους τα δεδομένα μπορούν να καταγραφούν και να ανατροφοδοτηθούν προς τους σπουδαστές. Επιπλέον, υφίστανται σχετικά εργαλεία που παρέχουν στους σπουδαστές αναπαραστάσεις των δεδομένων και βοηθούν στην αυτοαξιολόγηση των σπουδαστών και στην αυτορρύθμιση της μάθησής τους Oldfield et al. (2012, σ. 30).

Οι Gavsevic, Greiff & Shaffer (2022, σ. 1) αναφέρουν ότι για κάποιους ερευνητές (Knight, Buckingham Shum & Littleton, 2013. Milligan, 2020), η Μαθησιακή Αναλυτική είναι εγγενώς μια μορφή αξιολόγησης. Ειδικότερα, οι Knight et al. (2013, σ. 75) υποστηρίζουν ότι η Μαθησιακή Αναλυτική είναι μια νέα τεχνολογία αξιολόγησης και ότι η χρήση της, ως συνεχούς αξιολογικής διαδικασίας με παράλληλη διαμορφωτική ανατροφοδότηση, διευκολύνει τη μάθηση. Η Milligan (2020, σ. 182) θεωρεί ότι η ανάλυση των δεδομένων που συλλέγονται από τα συστήματα των εκπαιδευτικών ιδρυμάτων μπορεί θεωρητικά να χρησιμοποιηθεί για αξιολογικούς σκοπούς. Σύμφωνα με τους Barthakur, et al. (2021, σ. 1) τα δεδομένα χρήσης που συλλέγονται στα διαδικτυακά περιβάλλοντα μάθησης (συγκεκριμένα στα Μαζικά Ανοικτά Διαδικτυακά Μαθήματα) και αναλύονται με τεχνικές Μαθησιακής Αναλυτικής παρέχουν λεπτομερή αξιολόγηση για την μαθησιακή πρόοδο των σπουδαστών. Οι Dowell & Roquet (2021, σ. 1) θεωρούν ότι τεχνικές της Μαθησιακής Αναλυτικής που μελετούν ομότιμες αλληλεπιδράσεις, οι οποίες συντελούνται σε τεχνολογικά περιβάλλοντα, μπορούν να αξιολογήσουν τις κοινωνικό-γνωσιακές διεργασίες των σπουδαστών. Οι Albreiki, Zaki & Alashwal (2021, σ. 2) υποστηρίζουν ότι τα δεδομένα που καταγράφουν οι πλατφόρμες ψηφιακής μάθησης μπορούν να χρησιμοποιηθούν για την αξιολόγηση σπουδαστών και μαθημάτων. Οι Liu et al. (2023) υποστηρίζουν ότι τα μοντέλα πρόβλεψης μπορούν να βοηθήσουν τους διδάσκοντες να κατανοήσουν τις συμπεριφορές των σπουδαστών και να προβούν σε κατάλληλες αναπροσαρμογές στο σχεδιασμό των μαθημάτων. Μπορούν επίσης να αξιοποιηθούν για την παροχή στοχευμένης εκπαιδευτικής στήριξης προς τους σπουδαστές που αντιμετωπίζουν δυσκολίες.

Εργαλεία και τεχνικές για την αξιολόγηση σπουδαστών μέσω Μαθησιακής Αναλυτικής παραθέτουν αρκετοί ερευνητές. Ένα από τα πρώτα και τα πιο γνωστά παραδείγματα είναι το

σύστημα «Course Signals» (Arnold & Pistilli, 2012), που χρησιμοποιήθηκε στο Πανεπιστήμιο Purdue των Η.Π.Α. και το οποίο θεωρείται ως μια από τις εμβληματικές εφαρμογές της Μαθησιακής Αναλυτικής (Ferguson, 2012, σ. 8). Απαρτιζόταν από ένα προβλεπτικό μοντέλο που εντόπιζε «αδύναμους» σπουδαστές και μια «διεπαφή» (interface), η οποία εμφάνιζε την πιθανότητα αποτυχίας χρησιμοποιώντας το «ανάλογο» ενός φωτεινού σηματοδότη (πράσινο: μικρή πιθανότητα αποτυχίας, πορτοκαλί: μέση πιθανότητα αποτυχίας, κόκκινο: μεγάλη πιθανότητα αποτυχίας). Οι σπουδαστές ενημερώνονταν μέσω μηνυμάτων αλληλογραφίας που έστελνε ο διδάσκοντας. Το σύστημα αξιοποιούσε δεδομένα, όχι μόνο από το ΣΔΜ που χρησιμοποιούσαν οι σπουδαστές στα πλαίσια των μαθημάτων, αλλά και δεδομένα τα οποία αφορούσαν δημογραφικά στοιχεία και πληροφορίες από την προηγούμενη ακαδημαϊκή πορεία του σπουδαστή. Ο στόχος του συστήματος ήταν η παροχή έγκαιρης και προσωποποιημένης ανατροφοδότησης προς τους σπουδαστές. Το σύστημα λειτούργησε με αποδοτικό τρόπο στις περιπτώσεις κατά τις οποίες η ανατροφοδότηση που λάμβαναν οι σπουδαστές συνοδευόταν και από αξιοποιήσιμες συμβουλές που παρείχαν οι διδάσκοντες με μη αυτόματοποιημένο τρόπο.

Οι Petropoulou, Kasimatis, Dimopoulos & Retalis (2014) ανέπτυξαν εργαλείο διαμορφωτικής αξιολόγησης το οποίο κατέγραφε τις αλληλεπιδράσεις και τις ενέργειες των σπουδαστών στο ΣΔΜ και βοηθούσε τους διδάσκοντες στην αξιολόγηση γνώσεων και δεξιοτήτων των σπουδαστών, βάση κριτηρίων που θέτει ο διδάσκοντας και τα οποία σχετίζονται με δείκτες όπως, η συμμετοχή σε συνεργατικές δραστηριότητες, η βαθμολογία αξιολογούμενων δραστηριοτήτων και η μελέτη εκπαιδευτικού υλικού. Η Redecker (2013) περιγράφει την εγκατάσταση εφαρμογής στους υπολογιστές των σπουδαστών, η οποία καταγράφει δεδομένα χρήσης του διαδικτύου, μελετά τις στρατηγικές διαδικτυακής έρευνας που υιοθετούν οι σπουδαστές και ακολούθως παρέχει κατάλληλη ανατροφοδότηση για τη βελτίωση των στρατηγικών. Οι Amigud et al. (2017) αξιοποιούν μεθόδους Μαθησιακής Αναλυτικής για να αναγνωρίσουν μοτίβα γραπτού λόγου που παράγουν οι σπουδαστές, ώστε να πιστοποιήσουν την ταυτότητα τους σε διαδικασίες ψηφιακής αξιολόγησης και ακολούθως να εντοπίσουν παραβιάσεις της ακαδημαϊκής ακεραιότητας. Η Jordan (2013) αναφέρει παραδείγματα εφαρμογών που αναλύουν δεδομένα συμπεριφοράς σπουδαστών που συμμετέχουν σε

ψηφιακές αξιολογήσεις κλειστού τύπου, προκειμένου να αναδειχθούν οι παράμετροι που ενισχύουν την ενεργή συμμετοχή των σπουδαστών.

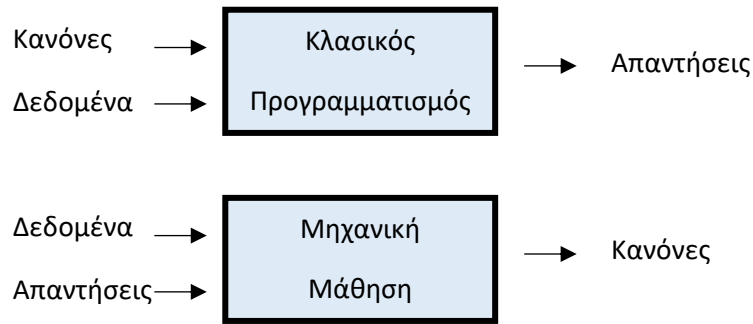
Οι Kapsalis et al. (2019) εκφράζουν επιφυλάξεις αφενός για τον βαθμό ωριμότητας των συγκεκριμένων εργαλείων και αφετέρου για την αποτελεσματικότητά τους. Συγκεκριμένα αναφέρουν ότι παρότι υφίστανται υλοποιήσεις που είναι ενθαρρυντικές, ο τομέας της Μαθησιακής Αναλυτικής είναι κατακερματισμένος ως προς τις χρήσεις, που μπορούν να βελτιώσουν και να εκσυγχρονίσουν την εκπαίδευση. Επιπλέον υπάρχει ευρύ χάσμα μεταξύ αφενός των δυνητικών χρήσεων και αφετέρου των προτάσεων που έχουν φτάσει στο στάδιο της υλοποίησης και της δοκιμής. Σημειώνουν επίσης την ανάγκη να προκύψουν απτές αποδείξεις - πέρα από μικρής κλίμακας μελέτες - που θα επικυρώσουν ότι τα σχετικά εργαλεία εκπληρώνουν τους επιδιωκόμενους στόχους και ότι επιτυγχάνουν θετικά αποτελέσματα ως προς τον στόχο της βελτίωσης της αποτελεσματικότητας και της αποδοτικότητας της εκπαιδευτικής διαδικασίας.

### 5.3. Προβλεπτική Αναλυτική και Μηχανική Μάθηση

Όπως έχει ήδη αναφερθεί η «προβλεπτική αναλυτική», δηλαδή η ανάπτυξη μοντέλων που αναλύουν μοτίβα και συσχετίσεις στα εκπαιδευτικά δεδομένα και επιχειρούν πιθανοκρατικές προβλέψεις για μελλοντικά αποτελέσματα που σχετίζονται με ακαδημαϊκές επιδόσεις ή με συμπεριφορές των σπουδαστών, κυριαρχεί από τα πρώτα χρόνια εμφάνισης της Μαθησιακής Αναλυτικής (Joksimovic, Kovanovic & Dawson, 2019, σ. 43).

Η κατασκευή των μοντέλων βασίζεται σε τεχνικές Μηχανικής Μάθησης, δηλαδή στον πιο δημοφιλή και πιο επιτυχημένο κλάδο της Τεχνητής Νοημοσύνης (Chollet, 2017, σ. 2), όπου τα υπολογιστικά συστήματα «μαθαίνουν» από τα δεδομένα, αναγνωρίζουν «μοτίβα» (patterns) και προβλέπουν αποτελέσματα (Albreiki, Zaki & Alashwal, 2021, σ. 2). Η Μηχανική Μάθηση είναι ένα σύνολο τεχνικών που επιτρέπουν στις υπολογιστικές μηχανές να κατασκευάζουν κανόνες (να «μαθαίνουν»), χωρίς να προγραμματίζονται από άνθρωπο (Rastrollo-Guerrero, Gomez-Pulido & Duran-Dominguez, 2020, σ. 5). Αξιοποιείται σε πλήθος ερευνητικών πεδίων, όπως για παράδειγμα στην αναγνώριση εικόνων και στην επεξεργασία γλώσσας, αλλά και σε εφαρμογές όπως οι μηχανές αναζήτησης, το φιλτράρισμα ανεπιθύμητης ηλ. αλληλογραφίας, η προστασία από κακόβουλες ενέργειες κατά τις συναλλαγές πιστωτικών καρτών, η στάθμιση κινδύνων σε ασφαλιστικά και τραπεζικά προϊόντα. Εκατομμύρια αποφάσεις που επηρεάζουν τις ζωές των ανθρώπων, λαμβάνονται καθημερινά από συστήματα που βασίζονται σε τεχνικές Μηχανικής Μάθησης (Carabantes, 2020, σ. 315).

Η Μηχανική Μάθηση βασίζεται σε ένα νέο «προγραμματιστικό υπόδειγμα» (programming paradigm), το οποίο ο Chollet (2017, σ. 2) περιγράφει απλουστευτικά ως εξής: Στις κλασικές προγραμματιστικές τεχνικές, ο «προγραμματιστής» εισάγει στο σύστημα α) κανόνες και β) δεδομένα. Τα δεδομένα υφίστανται επεξεργασία, βάση των κανόνων, και με τον τρόπο αυτό προκύπτουν απαντήσεις. Στη Μηχανική Μάθηση, ο «προγραμματιστής» εισάγει στο σύστημα α) δεδομένα και β) τις αναμενόμενες από τα δεδομένα απαντήσεις. Με τον τρόπο αυτό, προκύπτουν κανόνες (Σχήμα 5). Ακολούθως, οι κανόνες μπορούν να εφαρμοσθούν σε νέα δεδομένα και να οδηγήσουν σε απαντήσεις.



**Σχήμα 5. Κλασικές προγραμματιστικές τεχνικές έναντι Μηχανικής Μάθησης (Chollet,2019 σ.3)**

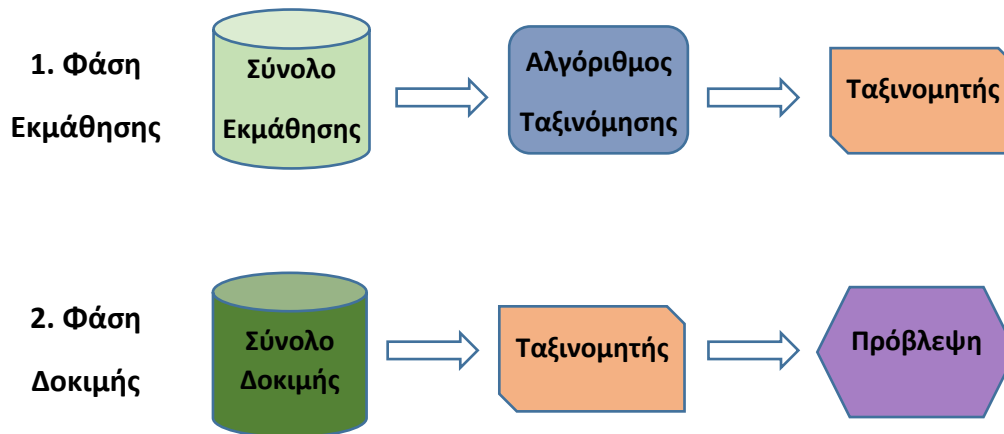
Συνεπώς, ένα σύστημα Μηχανικής Μάθησης «μαθαίνει», «εκπαιδεύεται» και δεν προγραμματίζεται. Κατά το στάδιο της «εκμάθησης» ή «εκπαίδευσης», τροφοδοτείται με πολλά «παραδείγματα», δηλαδή ζεύγη δεδομένων και απαντήσεων, εντός των οποίων εντοπίζει στατιστικές συσχετίσεις, από όπου συνάγει κανόνες. Στο επόμενο στάδιο, το στάδιο της «δοκιμής», όταν δηλαδή τροφοδοτηθεί με νέα δεδομένα, επιχειρεί να δημιουργήσει ζεύγη μεταξύ νέων, άγνωστων ως τότε, δεδομένων και απαντήσεων που, βάση των κανόνων, φαίνεται ότι ταιριάζουν στα νέα δεδομένα.

Η Μηχανική Μάθηση σχετίζεται στενά με την επιστήμη της Στατιστικής, αλλά διαφέρει από αυτήν σε αρκετά σημεία. Συγκεκριμένα, αναλύει τεράστιες ποσότητες δεδομένων (π.χ. εκατομμύρια εικόνες με χιλιάδες εικονοστοιχεία η καθέμια), εργασία που θα ήταν πρακτικώς αδύνατο να υλοποιηθεί με τη Στατιστική. Εμφανίζει λίγη μαθηματική θεωρία και είναι κυρίως προσανατολισμένη στη «μηχανική» (engineering-oriented). Συνιστά επιστημονικό πεδίο, όπου «οι ιδέες αποδεικνύονται περισσότερο εμπειρικά, παρά θεωρητικά» (Chollet,2019 σ.3). Επίσης, ενώ η Στατιστική εξάγει συμπεράσματα από δείγματα, η Μηχανική Μάθηση εντοπίζει γενικεύσιμα μοτίβα (Bzdok, Altman & Krzywinski, 2018, σ. 1). Γενικά, τα στατιστικά μοντέλα εστιάζουν σε συμπεράσματα που προκύπτουν μετά από δημιουργία και εφαρμογή πιθανοτικών μοντέλων, τα οποία ελέγχουν αν οι σχέσεις μεταξύ των μεταβλητών περιγράφουν ένα πραγματικό φαινόμενο ή αν προκύπτουν τυχαίως. Αντιθέτως, η Μηχανική Μάθηση επικεντρώνεται σε προβλέψεις που προκύπτουν από μοτίβα και απαιτεί ελάχιστες προϋποθέσεις ως προς την καταλληλότητα των δεδομένων. Είναι αποτελεσματική ακόμα και αν τα δεδομένα συλλεχθούν χωρίς να υπάρχει αυστηρός πειραματικός σχεδιασμός, ακόμα και αν

υφίστανται σύνθετες μη γραμμικές συσχετίσεις. Βέβαια, μεταξύ Μηχανικής Μάθησης και Στατιστικής υπάρχουν αρκετά σημεία σύγκλισης και συνέργειας (Witten, Frank & Hall, 2011, σ. 29). Για παράδειγμα, κατά τη συγκρότηση των συνόλων δεδομένων, εφαρμόζονται στατιστικές μέθοδοι για την επιλογή των χαρακτηριστικών και την απόρριψη ακραίων τιμών. Επίσης, στατιστικοί έλεγχοι εφαρμόζονται για να αποτιμήσουν τους αλγόριθμους και τα μοντέλα Μηχανικής Μάθησης και για να διορθώσουν στρεβλώσεις που ενίοτε προκύπτουν, λόγω της υπερ-προσαρμογής των μοντέλων κατά το στάδιο της «εκμάθησης» ή «εκπαίδευσης».

### 5.3.1. Ταξινόμηση και Μοντέλα Πρόβλεψης

Βασική τεχνική της Μηχανικής Μάθησης είναι η «επιβλεπόμενη μάθηση» (supervised learning) που αφορά προβλήματα «ταξινόμησης» (classification). Η «ταξινόμηση», δηλαδή η κατάταξη μιας οντότητας σε μία από πολλές κλάσεις / τάξεις / κατηγορίες (class), είναι διαδικασία που συντελείται σε δύο φάσεις / στάδια (Σχήμα 6).



*Σχήμα 6. Φάσεις εκμάθησης και δοκιμής στην τεχνική της ταξινόμησης της Μηχανικής Μάθησης*

Στην πρώτη φάση, η οποία ονομάζεται φάση «εκμάθησης» ή «εκπαίδευσης» (learning phase / training step), ένας αλγόριθμος ταξινόμησης τροφοδοτείται με δεδομένα από το σύνολο εκμάθησης, το οποίο αποτελείται από  $n$  διατεταγμένες  $n$ -άδες (tuples) (π.χ. δυάδες, τριάδες, τετράδες, κτλ), που η κάθε μία περιγράφει μια οντότητα και  $\beta$ ) τις αντίστοιχες κλάσεις / τάξεις / κατηγορίες (class), στις οποίες ανήκουν οι οντότητες. Καθένα από τα  $n$  μέλη κάθε διατεταγμένης  $n$ -άδας εκφράζει την τιμή ενός χαρακτηριστικού της οντότητας. Η κλάση κάθε  $n$ -άδας εκφράζει την κατηγορία, στην οποία ανήκει η οντότητα και είναι κατηγορική μεταβλητή. Οι  $n$ -άδες του συνόλου εκμάθησης επιλέγονται τυχαία από το ευρύτερο σύνολο δεδομένων. Στα παραπάνω, οφείλεται η ονομασία «επιβλεπόμενη μάθηση», δεδομένου ότι ο αλγόριθμος ταξινόμησης λειτουργεί «υπό επίβλεψη», διότι του γνωστοποιείται η κλάση στην οποία ανήκει κάθε  $n$ -άδα. Ο αλγόριθμος ταξινόμησης κατασκευάζει έναν «ταξινομητή» ή ένα προβλεπτικό μοντέλο, δηλαδή κανόνες για την ταξινόμηση οντοτήτων. Ο «ταξινομητής» μπορεί να νοηθεί ως συνάρτηση  $y=f(X)$ , η οποία υπολογίζει (ή «προβλέπει»), την κλάση  $y$  από την  $n$ -άδα, η οποία συμβολίζεται με το  $X$ . Συνήθως, η συνάρτηση συνίσταται από κανόνες ταξινόμησης, δέντρα αποφάσεων,



μαθηματικές αναπαραστάσεις κ.α. (Han, Kamber & Pei, 2011, σ. 272,273). Ας σημειωθεί, ότι αν η μεταβλητή  $y$  δεν είναι κατηγορική, αλλά συνεχής μεταβλητή, τότε δεν πρόκειται για πρόβλημα «ταξινόμησης», αλλά για πρόβλημα «παλινδρόμησης» (regression) (Han, Kamber & Pei, 2011, σ. 27)

Στην δεύτερη φάση, η οποία ονομάζεται φάση «δοκιμής», ο «ταξινομητής», δηλαδή το προβλεπτικό μοντέλο, τροφοδοτείται με  $n$ -άδες από το σύνολο «δοκιμής» και τις ταξινομεί με βάση τους κανόνες που προέκυψαν κατά τη φάση «εκμάθησης». Οι  $n$ -άδες του συνόλου «δοκιμής» πρέπει να είναι διαφορετικές από τις  $n$ -άδες του συνόλου «εκμάθησης», δηλαδή δεν πρέπει να έχουν χρησιμοποιηθεί για την κατασκευή του «ταξινομητή». Αν, κατά τη φάση «δοκιμής», ο «ταξινομητής» τροφοδοτηθεί με  $n$ -άδες από το σύνολο «εκμάθησης» και όχι από το σύνολο «δοκιμής», τότε η πρόβλεψη θα είναι αφύσικα ακριβής. Διότι, ο «ταξινομητής» τείνει να υπερ-προσαρμόζεται στα δεδομένα του συνόλου «εκμάθησης», δηλαδή να αποδίδει βαρύτητα σε ιδιαιτερότητες του συνόλου «εκμάθησης», οι οποίες δεν είναι αντιπροσωπευτικές του ευρύτερου συνόλου δεδομένων. Μετά την ταξινόμηση των  $n$ -άδων του συνόλου «δοκιμής», συγκρίνεται η κλάση που πρόβλεψε ο «ταξινομητής» με την «πραγματική» κλάση κάθε  $n$ -άδας. Το ποσοστό των ορθώς ταξινομημένων  $n$ -άδων ορίζεται ως *ακρίβεια πρόβλεψης* (prediction accuracy) του «ταξινομητή». Αν η ακρίβεια πρόβλεψης θεωρηθεί αποδεκτή, τότε ο «ταξινομητής», μπορεί να χρησιμοποιηθεί για την ταξινόμηση νέων  $n$ -άδων άγνωστης κλάσης (class) από άλλα σύνολα δεδομένων (Han, Kamber & Pei, 2011, σ. 272,273).

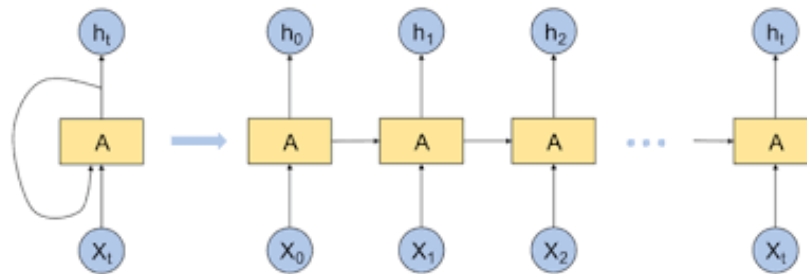
Άλλη τεχνική της Μηχανικής Μάθησης είναι η «μη επιβλεπόμενη μάθηση» (unsupervised learning), όπου οι κατηγορίες των  $n$ -άδων δεν είναι γνωστές, οπότε οι αλγόριθμοι Μηχανικής Μάθησης «ανακαλύπτουν» ομαδοποιήσεις, δηλαδή κατατάσσουν τις οντότητες σε ομάδες (clusters), έτσι ώστε οι οντότητες μιας ομάδας να μοιάζουν αρκετά μεταξύ τους και ταυτοχρόνως να διαφέρουν αρκετά από τις οντότητες των άλλων ομάδων. Οι ομοιότητες και οι διαφορές αποτιμώνται με βάση τα χαρακτηριστικά στις  $n$ -άδες των οντοτήτων (Han, Kamber & Pei, 2011, σ. 361).

### 5.3.1.1. Τεχνικές Βαθιάς Μάθησης (Deep Learning)

Η «Βαθιά Μάθηση» (deep learning) είναι μια μορφή της Μηχανικής Μάθησης, όπου τα μοντέλα πρόβλεψης, μέσω επάλληλων στρωμάτων επεξεργασίας (processing layers), αυξάνουν σταδιακά τη νοηματική πληρότητα στις αναπαραστάσεις των δεδομένων. Οπότε, για την αντιστοίχιση των ν-άδων του συνόλου δεδομένων με τις αντίστοιχες κλάσεις, απαιτούνται πολλαπλές διαδοχικές μεταβολές στη μορφή των δεδομένων, οι οποίες υλοποιούνται από τα επάλληλα στρώματα. Ο επιθετικός προσδιορισμός «βαθιά» δεν υπονοεί βαθύτερη κατανόηση της πληροφορίας, αλλά αναφέρεται στο πλήθος των επάλληλων στρωμάτων επεξεργασίας. Όσο περισσότερα είναι τα στρώματα, τόσο μεγαλύτερο το βάθος του μοντέλου. Ας σημειωθεί, ότι παρά το γεγονός ότι τα δίκτυα που συγκροτούν τα στρώματα αποκαλούνται «νευρωνικά δίκτυα» (όρος που προέρχεται από την επιστήμη της νευροφυσιολογίας), τα μοντέλα της «βαθιάς μάθησης» δεν αποτελούν προσομοιώσεις του ανθρώπινου εγκεφάλου (Chollet, 2017, σ. 6, 7). Η «βαθιά μάθηση» θεωρείται ως η πιο πρωτοποριακή προσέγγιση για προβλήματα ταξινόμησης, παλινδρόμησης και κυρίως πρόγνωσης χρονοσειρών (Coelho & Silveira, 2017, σ. 1).

Από τους βασικότερους αλγόριθμους «βαθιάς μάθησης» για τις χρονοσειρές είναι ο LSTM (Long Short Term Memory) (Chollet, 2017, σ. 12). Δεδομένου ότι τα δεδομένα που παράγονται κατά τη χρήση των Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ), σχετίζονται με χρονοσειρές, ο LSTM αλγόριθμος ελκύει το ενδιαφέρον ερευνητών (Chen & Cui, 2020. Fei & Yeung, 2015. Kim, Vitezai & Ganarathi, 2018. Liu, et al. 2022. Okubo, Yamashita, Shimada & Ogata, 2017). Τα δίκτυα που παράγει ο αλγόριθμος LSTM ανήκουν σε μια ευρύτερη κατηγορία δικτύων που ονομάζονται «Περιοδικά Νευρωνικά Δίκτυα» (RNN – Recurrent Neural Networks). Οι βρόχοι που συγκροτούν τα RNN δίκτυα προωθούν την πληροφορία μεταξύ διαδοχικών κόμβων, ο καθένας εκ των οποίων αντιστοιχεί σε μια χρονική στιγμή (Σχήμα 7) . Έτσι, τα RNN έχουν καλή απόδοση, όταν αναλύουν δεδομένα χρονοσειρών. Όμως, όταν οι χρονοσειρές καλύπτουν μακρά χρονικά διαστήματα, η απόδοση των RNN δικτύων κάμπτεται, οπότε σε τέτοιες περιπτώσεις αξιοποιούνται τα LSTM δίκτυα, δεδομένου ότι έχουν την ικανότητα να αναλύουν μακρές χρονοσειρές. Τα LSTM δίκτυα επιδεικνύουν εντυπωσιακά αποτελέσματα στη ρομποτική, στην ιατρική, στην ανάλυση ομιλίας, αλλά δεν έχουν μελετηθεί και αξιοποιηθεί αρκετά στη μελέτη

εκπαιδευτικών δεδομένων, παρότι φαίνεται να ταιριάζουν στα δεδομένα των ΣΔΜ, όπου οι συμπεριφορές και αλληλεπιδράσεις των σπουδαστών συμβαίνουν κατά τη διάρκεια μακρών χρονικών περιόδων και όπου η συμμετοχή των σπουδαστών πιθανόν περιγράφεται καλύτερα όχι από τις συνολικές σωρευμένες τιμές, αλλά από τη χρονική εξέλιξη των τιμών (Chen & Cui, 2020, σ. 4, 6).



**Σχήμα 7. Η δομή Περιοδικού Νευρωνικού Δικτύου (Recurrent Neural Network)**

### 5.3.2. Η έρευνα στο πεδίο της Προβλεπτικής Αναλυτικής

Έχουμε ήδη αναφέρει (βλ. Κεφ. 5.1) ότι η «Προβλεπτική Αναλυτική», κυριαρχεί από τα πρώτα χρόνια εμφάνισης του πεδίου της Μαθησιακής Αναλυτικής (Joksimovic, Kovanovic & Dawson, 2019, σ. 43). Συνεπώς, αρκετές είναι οι μελέτες που περιγράφουν σχετικές ερευνητικές προσεγγίσεις, οι οποίες έχουν μεν κοινή ερευνητική στόχευση, αλλά επιμέρους μεθοδολογικές διαφοροποιήσεις. Ο βασικός άξονας των ερευνών αφορά την ανάπτυξη μοντέλων τα οποία αναλύουν μοτίβα και συσχετίσεις στα εκπαιδευτικά δεδομένα και επιχειρούν πιθανοκρατικές προβλέψεις για ακαδημαϊκές επιδόσεις ή για συμπεριφορές των σπουδαστών. Οι διαφορές αφορούν κυρίως τον τύπο των δεδομένων που αναλύονται, τα χαρακτηριστικά που συγκροτούν τα σύνολα δεδομένων, τις μεθόδους ταξινόμησης των οντοτήτων, τα εργαλεία που χρησιμοποιούνται για την κατασκευή των μοντέλων, τα κριτήρια συγκριτικής αξιολόγησης των μοντέλων.

Οι έρευνες που σταχυολογήθηκαν και παρουσιάζονται παρακάτω, εντοπίστηκαν μέσω της μηχανής αναζήτησης Google Scholar, η οποία τηρεί βάση δεδομένων με ερευνητικά άρθρα επιστημονικών περιοδικών και πρακτικών συνεδρίων, με επιστημονικές αναφορές και με βιβλία. Οι λέξεις ευρετηριασμού που χρησιμοποιήθηκαν ήταν: Learning Analytics, Predictive Analytics, Educational Data Mining, Machine Learning, Supervised Learning, Student Performance Prediction, Dropout Prediction, Student Classification, Student Performance, Student Online Engagement, Deep Learning, LSTM. Για τον εντοπισμό των ερευνών, αξιοποιήθηκαν επίσης οι συστηματικές βιβλιογραφικές επισκοπήσεις των Coelho & Silveira (2017), Cui, Chen, Shiri & Fan (2019), Zulkifli, Mohamed & Azmee (2019), Doleck, Lemay, Basnet & Bazalais (2020), Namoun & Alshantqi (2020), Rastrollo-Guerrero, Gomez-Pulido & Duran-Dominguez (2020), Albreiki, Zaki & Alashwal (2021), Baashar, Alkawsji, Ali, Alhussian & Bahboub (2021). Επιπλέον, ακολουθήσαμε τη μέθοδο της «βιβλιογραφικής ιχνηλασίας» (reference tracking / backward snowballing method) (Wohlin, 2014), δηλαδή αξιοποιήσαμε τους καταλόγους των βιβλιογραφικών αναφορών στα άρθρα που εντοπίζαμε, διευρύνοντας συνεχώς το σώμα των υπό μελέτη ερευνών. Από τα κείμενα που προέκυψαν και τα οποία διατίθενταν με ελεύθερη διαδικτυακή πρόσβαση, εξαιρέσαμε καταρχάς όσα αφορούσαν σύντομες παρουσιάσεις και περιλήψεις σε συνέδρια και δεν αναφέρονταν στις μεθόδους που

εφάρμοσαν. Εξαιρέσαμε επίσης έρευνες που δεν αφορούσαν εκπαιδευτικά δεδομένα, έρευνες που δεν αξιοποιούσαν δεδομένα χρήσης ΣΔΜ και έρευνες που είχαν θεωρητικό χαρακτήρα και δεν περιλάμβαναν πειραματικά δεδομένα και αποτελέσματα. Ακολουθούν τα άρθρα των μελετών κατά χρονολογική σειρά.

Οι Macfadyen & Dawson (2010) μελέτησαν δεδομένα χρήσης ΣΔΜ από πέντε πλήρως διαδικτυακά προπτυχιακά μαθήματα. Οι μεταβλητές που χρησιμοποιήσαν εξήγησαν το 30% στην απόκλιση των τελικών βαθμών και το μοντέλο που κατασκεύασαν προέβλεψε με επιτυχία το 70.3% των αποτυχόντων σπουδαστών.

Ο Kotsiantis (2012) περιγράφει την κατασκευή μοντέλου για την πρόβλεψη της επιτυχίας ή της αποτυχίας σπουδαστών βασιζόμενος σε δημογραφικά και ακαδημαϊκά δεδομένα και επιπλέον σε δεδομένα χρήσης ΣΔΜ. Τα δεδομένα αφορούσαν σπουδαστές από τεχνολογικό μάθημα του Ελληνικού Ανοικτού Πανεπιστημίου. Η τιμή της ακρίβειας πρόβλεψης που επιτεύχθηκε ήταν 80%, στο μέσο του χρονικού διαστήματος του μαθήματος.

Οι Romero, Zafra, Romero & Ventura (2013) επιχείρησαν να προβλέψουν την επίδοση των σπουδαστών αναλύοντας αφενός τους βαθμούς διαγωνισμάτων και εργασιών και αφετέρου τη διάρκεια χρήσης εκπαιδευτικών πόρων στο ΣΔΜ. Χρησιμοποίησαν δεδομένα από επτά μαθήματα και συνέκριναν την απόδοση εναλλακτικών μεθόδων ταξινόμησης. Η ακρίβεια πρόβλεψης που πέτυχαν ήταν 65%.

Οι Hu, Lo & Shih (2014) χρησιμοποίησαν αποκλειστικά δεδομένα χρήσης ΣΔΜ από δύο προπτυχιακά τεχνολογικά μαθήματα. Χρησιμοποιώντας βασικούς αλγόριθμους Μηχανικής Μάθησης κατασκεύασαν μοντέλα των οποίων η ακρίβεια πρόβλεψης, ως προς τις κατηγορίες «επιτυχία» και «αποτυχία», ήταν 95.3%.

Οι Jayaprakash, Moody, Lauria, Regan & Baron (2014) μελέτησαν δημογραφικά και ακαδημαϊκά δεδομένα και δεδομένα χρήσης ΣΔΜ για όλα τα μαθήματα που προσφέρθηκαν σε τρία ακαδημαϊκά εξάμηνα. Χρησιμοποίησαν ως μέτρο απόδοσης των μοντέλων την τιμή AUC και πέτυχαν τιμές σημαντικά υψηλότερες από την τιμή βάσης, ως προς την πρόβλεψη των σπουδαστών που κινδυνεύουν να αποτύχουν.

Οι Azcona & Casey (2015) μελέτησαν δεδομένα από 111 φοιτητές που συμμετείχαν σε προπτυχιακό τεχνολογικό μάθημα. Κατασκεύασαν ταξινομητή για τις κατηγορίες «επιτυχία» και «αποτυχία» και πέτυχαν ακρίβεια πρόβλεψης 91%. Με δεδομένα που αφορούσαν χρονική στιγμή στο μέσο του μαθήματος (8<sup>η</sup> εβδομάδα), η ακρίβεια πρόβλεψης ήταν μεταξύ 65% και 70%. Όπως προέκυψε από την έρευνα, η ανάδραση που έλαβαν οι σπουδαστές βάση των προβλέψεων του μοντέλου, τους βοήθησε να έχουν καλύτερες επιδόσεις στην τελική αξιολόγηση του μαθήματος.

Οι Dewan, Lin & Wen (2015) προέβλεψαν το ποσοστό επιτυχίας των σπουδαστών σε διαδικτυακά μαθήματα αξιοποιώντας δεδομένα από ΣΔΜ και συνδυάζοντας τρεις διαφορετικούς αλγόριθμους Μηχανικής Μάθησης. Το μοντέλο που κατασκεύασαν συνδύαζε διαφορετικές μεθόδους Μηχανικής Μάθησης και προέβλεψε τους σπουδαστές που δεν ολοκλήρωσαν τη φοίτηση στα μαθήματα με ακρίβεια μεταξύ 85% και 90%.

Οι Fei & Yeung (2015) θεωρώντας ότι οι δραστηριότητες των σπουδαστών στο ΣΔΜ συμβαίνουν κατά τη διάρκεια συγκεκριμένης χρονικής περιόδου και χρονοσημούνται, επέλεξαν να αξιοποιήσουν περιοδικά νευρωνικά δίκτυα και συγκεκριμένα LSTM αλγόριθμο για να κατασκευάσουν μοντέλο πρόβλεψης. Η έρευνά τους έδειξε ότι τα εν λόγω μοντέλα υπερτερούν σε απόδοση έναντι μοντέλων που παράγονται με βασικούς αλγόριθμους Μηχανικής Μάθησης.

Οι Kuzilek, et al. (2015) χρησιμοποίησαν δεδομένα από το Open University Learning Analytics Dataset και κατασκεύασαν δύο διαφορετικά σύνολα δεδομένων, όπου το ένα περιλάμβανε δημογραφικά και ακαδημαϊκά δεδομένα και το δεύτερο δεδομένα χρήσης ΣΔΜ. Δημιούργησαν μοντέλα πρόβλεψης για καθένα από τα σύνολα και συνδύασαν τις προβλέψεις, ώστε τελικά να προβλέψουν ποιοι από τους σπουδαστές κινδυνεύουν να αποτύχουν.

Οι Tempelaar, Rienties & Giesbers (2015) συνέκριναν την προβλεπτική ικανότητα δεδομένων που αφορούσαν μαθησιακές διαθέσεις, δημογραφικά στοιχεία, βαθμούς από αρχικές και διαμορφωτικές αξιολογήσεις και δεδομένα από ΣΔΜ. Διαπίστωσαν ότι τα δεδομένα που αφορούσαν βαθμούς από διαμορφωτικές αξιολογήσεις προέβλεπαν καλύτερα την επίδοση των σπουδαστών.

Ο Zacharis (2015) διαπίστωσε ότι τα δεδομένα που αφορούσαν τη χρήση του ΣΔΜ εξηγούσαν περισσότερο από το 52% της απόκλισης στον τελικό βαθμό των σπουδαστών. Το μοντέλο πρόβλεψης, που κατασκεύασε για τις κατηγορίες «επιτυχία» / «αποτυχία», πέτυχε ακρίβεια πρόβλεψης 81.3% .

Οι Conijn, Snijders, Kleingeld & Matzat (2016) ανέλυσαν δεδομένα από ΣΔΜ που αφορούσαν δέκα επτά μαθήματα. Επιδίωξαν να προβλέψουν την επίδοση των φοιτητών ως προς τις κατηγορίες «επιτυχία» και «αποτυχία» αλλά και ως προς τον τελικό βαθμό. Για τα σύνολα δεδομένων που αφορούσαν όλα τα μαθήματα, προέκυψε ότι στην αρχή της περιόδου και ειδικότερα την 3<sup>η</sup> εβδομάδα, η ακρίβεια πρόβλεψης για την «επιτυχία»/ «αποτυχία» ήταν 67% και αυξανόταν σταθερά ιδίως μετά την 5<sup>η</sup> εβδομάδα, όταν ήταν διαθέσιμοι και οι βαθμοί από ενδιαμέση αξιολόγηση του μαθήματος. Η ακρίβεια πρόβλεψης παρουσίαζε σημαντικές διαφορές όταν τα μοντέλα εφαρμόζονταν σε σύνολα δεδομένων που αφορούσαν κάθε μάθημα ξεχωριστά.

Οι Gasevic, Dawson, Rogers & Gasevic (2016) κατασκεύασαν μοντέλα λογιστικής παλινδρόμησης για εννέα προπτυχιακά μαθήματα. Χρησιμοποίησαν δεδομένα από πληροφοριακά συστήματα του πανεπιστημίου και δεδομένα χρήσης ΣΔΜ και βρήκαν ότι η απόκλιση στους τελικούς βαθμούς μπορούσε να εξηγηθεί κατά 5%, όταν στα μοντέλα τους περιλαμβάνονταν χαρακτηριστικά των σπουδαστών και έως 16% όταν περιλαμβάνονταν και δεδομένα χρήσης του ΣΔΜ. Το μοντέλο που ανέλυσε δεδομένα όλων των μαθημάτων πέτυχε ικανοποιητική επίδοση, ενώ τα μοντέλα που ανέλυσαν δεδομένα για κάθε μάθημα ξεχωριστά πέτυχε εξαιρετική επίδοση.

Οι Liang, Li & Zheng (2016) μελέτησαν δεδομένα χρήσης ΣΔΜ για τριάντα εννέα διαδικτυακά μαθήματα και επιχειρήσαν να προβλέψουν ποιοι από τους σπουδαστές δεν θα ολοκλήρωναν τη φοίτησή τους. Κατασκεύασαν σύνολα δεδομένων τα οποία περιλάμβαναν εκατοντάδες χαρακτηριστικά. Τα μοντέλα πρόβλεψης που κατασκεύασαν πέτυχαν τιμή ακρίβειας πρόβλεψης 89%.

Οι Marbouti, Diefes-Dux & Madhavan (2016) χρησιμοποίησαν δεδομένα για δύο μαθήματα που διεξήχθησαν δύο συνεχόμενα έτη. Χρησιμοποίησαν τεχνικές επιλογής μεταβλητών, ώστε να

αναλύσουν μόνο τα δεδομένα που φαίνονταν να έχουν επίδραση στην απόδοση των μοντέλων. Κατασκεύασαν επτά μοντέλα πρόβλεψης εκ των οποίων, το αποδοτικότερο πέτυχε ακρίβεια πρόβλεψης για τις κατηγορίες «επιτυχία» / «αποτυχία» με ποσοστό 84.6%.

Οι Nespereira, Elhariri, El-Bendary, Vilas & Redondo (2016) μελέτησαν δεδομένα που αποκλειστικά προέρχονταν από τη χρήση του ΣΔΜ για δύο μαθήματα. Κατασκεύασαν μοντέλα πρόβλεψης, τα οποία πέτυχαν ακρίβεια 78.1% ως προς την πρόβλεψη για επιτυχή, ή μη επιτυχή τελική επίδοση του σπουδαστών.

Η Cohen (2017) μελέτησε δεδομένα από σπουδαστές που συμμετείχαν σε τρία προπτυχιακά μαθήματα. Διερεύνησε αν τα δεδομένα χρήσης του ΣΔΜ μπορούσαν να αξιοποιηθούν για να προβλεφθεί ποιοι από τους σπουδαστές δεν θα ολοκλήρωναν τη φοίτησή τους. Τα μοντέλα που κατασκευάσε εφαρμόστηκαν αφενός σε ξεχωριστά σύνολα δεδομένων, ένα για κάθε μάθημα, όπου επιτεύχθηκε ακρίβεια πρόβλεψης από 55% ως 80% και αφετέρου σε σύνολο δεδομένων για όλα τα μαθήματα όπου επιτεύχθηκε ακρίβεια πρόβλεψης 66%.

Οι Costa, Fonseca, Santana, de Araujo & Rego (2017) μελέτησαν δεδομένα από δύο προπτυχιακά τεχνολογικά μαθήματα, εκ των οποίων το ένα διδάχτηκε δια ζώσης και το άλλο εξ αποστάσεως. Χρησιμοποιήθηκαν δημογραφικά και ακαδημαϊκά δεδομένα και δεδομένα χρήσης του ΣΔΜ. Ως μέτρο απόδοσης των μοντέλων χρησιμοποιήθηκε το κριτήριο F, η τιμή του οποίου, ανάλογα με το πόσο πρώιμα επιχειρήθηκε η πρόβλεψη, κυμάνθηκε μεταξύ 55% και 82% για το εξ αποστάσεως μάθημα και 50% ως 79% για το δια ζώσης μάθημα.

Οι Leppanen, Leinonen, Ihantola & Hellas, (2017) μελέτησαν σύνολα δεδομένων που σχετιζόνταν μόνο με χρήση του ΣΔΜ για προπτυχιακό τεχνολογικό μάθημα. Διαπίστωσαν ότι η πρόβλεψη των σπουδαστών που κινδυνεύουν να αποτύχουν ήταν δυσκολότερη από την πρόβλεψη των σπουδαστών που επιτυγχάνουν. Με βάση το κριτήριο F το ποσοστό πρόβλεψης στο τέλος του μαθήματος ήταν αντιστοίχως 76% και 95%.

Οι Okubo, Yamashita, Shimada & Ogata (2017) κατασκεύασαν μοντέλα πρόβλεψης βασισμένα σε LSTM αλγόριθμο για να αναλύσουν δεδομένα χρήσης ΣΔΜ για προπτυχιακό τεχνολογικό μάθημα. Στο σύνολο δεδομένων, εκτός από δεδομένα σχετικά με τη συχνότητα χρήσης περιλαμβάνονταν και βαθμοί από διαγωνίσματα και εργασίες. Το μοντέλο που κατασκεύασαν



προέβλεψε τον τελικό βαθμό των φοιτητών με ποσοστό μεγαλύτερο από 90% στο μέσο της χρονικής διάρκειας του μαθήματος.

Οι Raga & Raga (2017) χρησιμοποίησαν δεδομένα από ΣΔΜ για τρία προπτυχιακά μαθήματα και κατασκεύασαν μοντέλο πρόβλεψης το οποίο ταξινομούσε σε τρεις κατηγορίες. Για το σύνολο των μαθημάτων, η ακρίβεια πρόβλεψης ήταν 72.8%, ενώ η μεγαλύτερη ακρίβεια που επιτεύχθηκε ανά μάθημα ήταν 87%.

Οι Howard, Meehan & Parnell (2018) συνέκριναν 8 διαφορετικές μεθόδους πρόβλεψης για την επίδοση των σπουδαστών. Το σύνολο δεδομένων που χρησιμοποίησαν αφορούσε διαδικτυακό μάθημα και περιλάμβανε δεδομένα από το ΣΔΜ, αλλά και δεδομένα που αφορούσαν το ακαδημαϊκό προφίλ των σπουδαστών και τους βαθμούς τους σε ενδιάμεσες αξιολογικές δραστηριότητες. Το μοντέλο που κατασκεύασαν ήταν ικανό να προβλέψει το βαθμό των σπουδαστών στο μέσο της χρονικής διάρκειας του μαθήματος με τιμή 6.5 ως προς το κριτήριο MAE (Mean Absolute Error).

Οι Kim, Vizitei & Ganapathi (2018) μελέτησαν δεδομένα χρήσης από σπουδαστές δύο διαδικτυακών μαθημάτων στο MOOC (Massive Open Online Course ) Udacity. Για την κατασκευή των μοντέλων χρησιμοποίησαν LSTM αλγόριθμο και προέβλεψαν την τελική επίδοση των σπουδαστών χρησιμοποιώντας ως κριτήριο την τιμή AUC η οποία υπερέβει σημαντικά την τιμή βάσης.

Οι Ljubobratovic & Matetic (2019) κατασκεύασαν μοντέλα πρόβλεψης που αναλύουν δεδομένα από τα ΣΔΜ πετυχαίνοντας ακρίβεια πρόβλεψης 96.3% και επιπλέον διαπίστωσαν ότι τη μεγαλύτερη επίδραση στην τελική επίδοση έχουν οι βαθμοί των ενδιάμεσων αξιολογήσεων.

Οι Buschetto Macarini et al. (2019) μελέτησαν σύνολα δεδομένων από τέσσερα προπτυχιακά τεχνολογικά μαθήματα που διδάχθηκαν με διαδικτυακές και δια ζώσης συνεδρίες. Τα δεδομένα προέρχονταν από το ΣΔΜ, από βαθμούς ενδιάμεσων αξιολογήσεων και από ερωτηματολόγια μέσω των οποίων καταγράφηκαν πληροφορίες για το κοινωνικό και δημογραφικό προφίλ των σπουδαστών και για τα κίνητρά τους. Τα μοντέλα που κατασκεύασαν πέτυχαν ικανοποιητική ακρίβεια πρόβλεψης για τις κατηγορίες «επιτυχία / αποτυχία», ήδη από την 8<sup>η</sup> εβδομάδα του μαθήματος.

Οι Hussain M., Zhu, Zhang, Abidi & Ali (2019) αξιοποίησαν δεδομένα από μάθημα, όπου οι σπουδαστές χρησιμοποίησαν ΣΔΜ για να μελετήσουν εκπαιδευτικό υλικό και να συμμετάσχουν σε δραστηριότητες αξιολόγησης. Επέλεξαν να μελετήσουν πέντε χαρακτηριστικά που περιγράφουν τη χρήση του ΣΔΜ από τους σπουδαστές. Κατασκεύασαν μοντέλα για την πρόβλεψη του τελικού βαθμού των σπουδαστών, τα οποία πέτυχαν μέγιστη ακρίβεια πρόβλεψης 82%.

Οι Sheshadri, Gitinabard, Lynch, Barnes & Heckman (2019) επιχείρησαν να προβλέψουν την απόδοση των σπουδαστών που συμμετείχαν σε τρία τεχνολογικά μαθήματα τα οποία διδάχθηκαν με μεικτό τρόπο (δια ζώσης και διαδικτυακά). Αξιοποίησαν δεδομένα από ΣΔΜ όπως τη διάρκεια πρόσβασης και τη συμμετοχή σε δημόσιους διαδικτυακούς διαλόγους. Τα μοντέλα που κατασκεύασαν πρόβλεψαν την «επιτυχία» ή «αποτυχία» των σπουδαστών με μέγιστο ποσοστό 80.77% ως προς το κριτήριο F.

Οι Umer, Mathrani, Susnjak & Lim (2019) αξιοποίησαν δεδομένα από ΣΔΜ και από αξιολογήσεις στις οποίες συμμετείχαν σπουδαστές, οι οποίοι παρακολούθησαν τέσσερα προπτυχιακά μαθήματα. Τα μοντέλα που κατασκεύασαν πέτυχαν ακρίβεια πρόβλεψης ως προς τις κατηγορίες «επιτυχία» / «αποτυχία» με ποσοστό που υπερβαίνει το 80% στο μέσο χρονικής διάρκειας του μαθήματος και το 90% στο τέλος του μαθήματος. Η ακρίβεια πρόβλεψης παρουσίαζε σημαντική αύξηση, όταν στα σύνολα δεδομένων προσθέτονταν οι βαθμοί από τις ενδιάμεσες αξιολογικές δραστηριότητες των σπουδαστών.

Οι Chui, Fung, Lytras & Lam (2020) κατασκεύασαν μοντέλο για την πρόβλεψη της επίδοσης των σπουδαστών ως προς τις κατηγορίες επιτυχία και αποτυχία με το οποίο πέτυχαν ακρίβεια πρόβλεψης 93.8%. Αξιοποίησαν δημογραφικά δεδομένα και δεδομένα χρήσης του ΣΔΜ από το Open University Learning Analytics Dataset.

Οι Lopez-Zambrano, Lara & Romero (2020) διερεύνησαν τη μεταφερσιμότητα μοντέλων πρόβλεψης μεταξύ διαφορετικών μαθημάτων. Κατασκεύασαν μοντέλα χρησιμοποιώντας αλγόριθμο δέντρου αποφάσεων (συγκεκριμένα τον αλγόριθμο J48) και μελέτησαν δεδομένα από 24 προπτυχιακά μαθήματα, τα οποία ομαδοποίησαν ανάλογα με τη συνάφεια του γνωστικού αντικειμένου και ανάλογα με το βαθμό χρήσης του ΣΔΜ. Για κάθε μάθημα

κατασκεύασαν δύο διαφορετικά σύνολα δεδομένων, όπου στο ένα σύνολο, τα χαρακτηριστικά εκφράζονται με αριθμητικές τιμές και στο άλλο, με κατηγορικές τιμές. Με βάση το κριτήριο της μεταβλητής AUC, αξιολόγησαν την προβλεπτική ικανότητα του μοντέλου κάθε μαθήματος, όταν αυτό το μοντέλο εφαρμοζόταν στα σύνολα δεδομένων των άλλων μαθημάτων. Διαπίστωσαν, ότι, γενικά, οι μέσες τιμές AUC δεν ήταν υψηλές και υπερβαίνουν ελάχιστα το όριο της «τυχαιότητας» ( $AUC=0.5$ ) (Lopez-Zambrano, Lara & Romero, 2020, σ. 17). Η μεταφερσιμότητα βελτιωνόταν μεταξύ μαθημάτων, που έχουν το ίδιο γνωστικό αντικείμενο και ίδιους διδάσκοντες, πιθανόν λόγω ομοιοτήτων στη μέθοδο υλοποίησης και αξιολόγησης του μαθήματος. Στις περιπτώσεις αυτές, η βέλτιστη μέση τιμή απώλειας στην τιμή AUC ήταν 0.22. Διαπίστωσαν επίσης ότι τα καλύτερα, ως προς τη μεταφερσιμότητα, μοντέλα δεν είναι εύκολο να αξιοποιηθούν από διδάσκοντες άλλων μαθημάτων, διότι περιλαμβάνουν χαρακτηριστικά που αντιστοιχούν σε δραστηριότητες οι οποίες δεν υφίστανται στα άλλα μαθήματα.

Οι Quinn & Gray (2020) χρησιμοποίησαν, λόγω του μικρού πλήθους των εγγεγραμμένων φοιτητών, δεδομένα από ΣΔΜ για είκοσι εννέα μαθήματα, τα οποία διεξήχθησαν σε περίοδο οκτώ χρόνων και τα οποία διέφεραν και ως προς το αντικείμενο και ως προς τη χρονική διάρκεια υλοποίησης, γεγονός που όπως αναφέρουν μπορεί να επηρεάσει την ακρίβεια της πρόβλεψης. Ωστόσο, θεωρούν πως αρκεί ότι προσφέρονται από τον ίδιο διδάσκοντα και ανήκουν στο ίδιο επιστημονικό πεδίο. Τα μοντέλα ταξινόμησης που δημιούργησαν πέτυχαν ακρίβεια 92.2% για την κατάταξη στις κατηγορίες «επιτυχία» και «αποτυχία» και 60.5% για την πρόβλεψη ως προς τον τελικό βαθμό των σπουδαστών.

Οι Tomasevic, Gvozdenovic & Vranes (2020) χρησιμοποίησαν ανοικτά σύνολα δεδομένων από το Open University Learning Analytics Dataset, τα οποία περιλαμβάνουν όχι μόνο δεδομένα χρήσης του ΣΔΜ, αλλά δημογραφικά και ακαδημαϊκά δεδομένα. Επέλεξαν να μελετήσουν δύο μαθήματα και επιχείρησαν προβλέψεις στην αρχή και στο μέσο της χρονικής διάρκειας του μαθήματος. Με βάση το κριτήριο F πέτυχαν ακρίβεια πρόβλεψης (για κατηγορίες «επιτυχία» / «αποτυχία»), που κυμάνθηκε από 78% στο μέσο της χρονικής διάρκειας του μαθήματος, κατά τη χρονική στιγμή της πρώτης ενδιάμεσης αξιολόγησης και 96.6% στο τέλος του μαθήματος.

Οι Riestra-Gonzalez, del Puerto Paule-Ruiz & Ortin (2021) μελέτησαν δεδομένα μόνο από ΣΔΜ που αφορούσαν 5112 μαθήματα. Πρόκειται για δια ζώσης, διαδικτυακά και μεικτά μαθήματα

που προσφέρθηκαν στα πλαίσια προπτυχιακών και μεταπτυχιακών κύκλων σπουδών και σχετίζονται με πολλά γνωστικά αντικείμενα. Όπως αναφέρουν, τα μαθήματα διαφέρουν στο γνωστικό αντικείμενο, στον τρόπο διδασκαλίας, στη διάρκεια, στη μεθοδολογία. Επιχείρησαν πρόβλεψη στην αρχή και στο μέσο του μαθήματος. Τα μοντέλα που κατασκεύασαν πέτυχαν ακρίβεια πρόβλεψης που κυμάνθηκε από 80.1% στην αρχή των μαθημάτων, κατά τη στιγμή της πρώτης ενδιάμεσης αξιολόγησης, έως 90.1% στο μέσο χρονικής διάρκειας των μαθημάτων.

Οι Liu, et al. (2022) χρησιμοποίησαν δεδομένα από το Open University Learning Analytics Dataset και τα μορφοποίησαν ως χρονοσειρές. Συνέκριναν την απόδοση βασικών μοντέλων ταξινόμησης με μοντέλα που βασίστηκαν σε αλγόριθμο LSTM. Το LSTM μοντέλο πέτυχε ακρίβεια πρόβλεψης 90.25% ως προς την κατηγορία «επιτυχία» / «αποτυχία» και ξεπέρασε σε απόδοση τα βασικά μοντέλα.

Οι παραπάνω έρευνες αναδεικνύουν το συνεχώς αυξανόμενο ενδιαφέρον της ερευνητικής κοινότητας για το πεδίο της Μαθησιακής Αναλυτικής και ειδικότερα για τα μοντέλα που επιχειρούν να προβλέψουν τις επιδόσεις των σπουδαστών.

- Προκύπτει, ότι τα δεδομένα στα οποία επικεντρώθηκαν οι έρευνες κατατάσσονται σε τρεις κατηγορίες: σε δημογραφικά δεδομένα (τα οποία περιλαμβάνουν πληροφορίες για το οικονομικό και κοινωνικό προφίλ των σπουδαστών), σε δεδομένα για τις ακαδημαϊκές επιδόσεις και γενικά για την ακαδημαϊκή πορεία των σπουδαστών και σε δεδομένα χρήσης του ΣΔΜ. Λίγες ήταν οι έρευνες που χρησιμοποίησαν αποκλειστικά δεδομένα χρήσης του ΣΔΜ.
- Λίγες ήταν οι έρευνες που διερεύνησαν ζητήματα ρύθμισης και επιλογής χαρακτηριστικών στα σύνολα δεδομένων για να αυξήσουν την ακρίβεια πρόβλεψης των μοντέλων ή για να αναδείξουν χαρακτηριστικά που φαίνεται να σχετίζονται περισσότερο από ό,τι άλλα με την επίδοση των σπουδαστών.
- Προκύπτει επίσης, ότι το πλήθος των οντοτήτων στα σύνολα δεδομένων είναι σημαντικός παράγοντας για την απόδοση των μοντέλων. Αρκετές έρευνες βασίστηκαν σε μικρά δείγματα, γεγονός που εμποδίζει την κατασκευή μοντέλων με υψηλή απόδοση και γενικά μειώνει την αποδεικτική ισχύ των πορισμάτων. Επιπλέον, οι περισσότερες

έρευνες εστίασαν σε ένα μόνο σύνολο δεδομένων. Δεν κατασκεύασαν περισσότερα σύνολα, ώστε να είναι δυνατές συγκρίσεις, ως προς την απόδοση των μοντέλων, όταν αυτά κληθούν να διαχειριστούν διαφορετικά δεδομένα.

- Στις περισσότερες έρευνες χρησιμοποιήθηκαν βασικοί αλγόριθμοι Μηχανικής Μάθησης. Συγκεκριμένα, οι αλγόριθμοι που κυρίως χρησιμοποιήθηκαν για την κατασκευή των μοντέλων ήταν οι: Decision Tree, Logistic Regression, K-Nearest Neighbor, Support Vector Machines, Naïve Bayes, Random Forest, Boosted Trees, Adaptive Boosting. Πέραν της αποδοτικότητάς τους, η δημοφιλία κάποιων αλγόριθμων ίσως οφείλεται και στην ικανότητά τους να κατασκευάζουν εύληπτα μοντέλα. Παρότι στα δεδομένα χρήσης του ΣΔΜ ενυπάρχει σημαντική χρονική διάσταση, λίγες ήταν οι έρευνες που αξιοποίησαν τεχνικές οι οποίες διαθέτουν δυνατότητα ανάλυσης δεδομένων που εκφράζονται με χρονοσειρές (π.χ. τεχνικές «βαθιάς μάθησης»).
- Οι περισσότερες μελέτες αντιμετώπισαν το ζήτημα της πρόβλεψης ως ζήτημα ταξινόμησης, επιχειρώντας να προβλέψουν τις κατηγορίες «επιτυχία» / «αποτυχία» ή «ολοκλήρωση» / «μη ολοκλήρωση» των σπουδών.

#### 5.4 Μαθησιακή Αναλυτική και Ενεργή Συμμετοχή

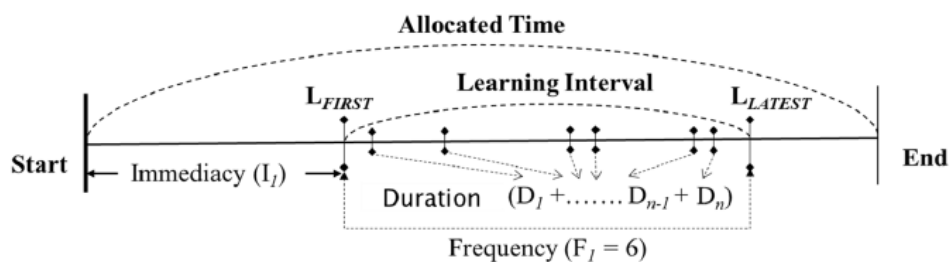
Η ενεργή συμμετοχή (engagement) αποτελεί βασικό πεδίο μελέτης της παιδαγωγικής επιστήμης. Τα οφέλη που προκύπτουν, όταν οι σπουδαστές συμμετέχουν ενεργά στην εκπαιδευτική διαδικασία, είναι πολλαπλά και προοιωνίζονται την επιτυχημένη φοίτηση (Sinatra, Heddy & Lombardi, 2015, σ. 1). Ωστόσο, δεν υπάρχει ομοφωνία, ούτε ως προς τον ακριβή ορισμό της ενεργής συμμετοχής, ούτε ως προς την αποτελεσματική μέτρησή της. Στη βιβλιογραφία, καταγράφονται πολλές εννοιολογικές αναπαραστάσεις (μοντέλα) της ενεργής συμμετοχής. Ένα από τα πιο λεπτομερή και περιεκτικά μοντέλα (Henrie, Bodily, Larsen, & Graham, 2018, σ. 345), είναι αυτό που προτείνουν οι Fredricks, Blumenfeld & Paris, (2004), σύμφωνα με τους οποίους, υφίστανται τρεις διαστάσεις ενεργής συμμετοχής: η συμπεριφοριστική (behavioral) διάσταση, η γνωστική (cognitive) διάσταση και η συναισθηματική (emotional) διάσταση.

Η μέτρηση της ενεργής συμμετοχής είναι σύνθετο ζήτημα (Sinatra, Heddy & Lombardi, 2015, σ. 1), αλλά είναι συνθετότερο στα ψηφιακά περιβάλλοντα μάθησης. Στα πλαίσια της Μαθησιακής Αναλυτικής, η ερευνητική κοινότητα, ενώ καταλήγει σε πορίσματα, που τεκμηριώνουν συσχετίσεις μεταξύ δεδομένων χρήσης και ακαδημαϊκής επίδοσης, εκφράζει σκεπτικισμό ως προς την τεκμηρίωση στέρεων συσχετίσεων μεταξύ δεδομένων χρήσης και «ενεργής συμμετοχής» (Motz, Quick, Schroeder, Zook & Gunkel, 2019, σ. 300, 307), κυρίως διότι η ακαδημαϊκή επίδοση είναι μετρήσιμη, ενώ η ενεργή συμμετοχή εμπεριέχει ψυχολογικές, γνωστικές και συναισθηματικές παραμέτρους, οι οποίες είναι δύσκολο να ανιχνευθούν, πόσο δε μάλλον να μετρηθούν. Εξάλλου, η έννοια της ενεργής συμμετοχής είναι υπό συνεχή διερεύνηση και τα επιμέρους συστατικά στοιχεία της συνεχώς, αναζητούνται, καταγράφονται, μελετώνται, αναλύονται, επανα-ορίζονται.

Οι Vytasek, Patzak & Winne (2020, σ. 27,28) αναφέρουν ότι στα ψηφιακά περιβάλλοντα μάθησης, η μέτρηση της ενεργής συμμετοχής μπορεί να βασιστεί στα δεδομένα χρήσης των συστημάτων, από τα οποία εξάγονται δείκτες όπως: α) ο χρόνος και συγκεκριμένα η χρονική διάρκεια για την εκτέλεση μιας ανατιθέμενης εργασίας (task), ή για την πρόσβαση και μελέτη ενός μαθησιακού πόρου, β) το πλήθος των μαθησιακών πόρων που οι σπουδαστές επισκέφθηκαν και των «τεχνουργημάτων» (artifacts, π.χ. εργασίες, κείμενα) που δημιούργησαν,

γ) οι αλληλεπιδράσεις μεταξύ των σπουδαστών, όπως π.χ. οι αναρτήσεις σε δημόσιους διαλόγους, στις οποίες, αν επιχειρηθεί ποιοτική και όχι μόνο ποσοτική ανάλυση, είναι δυνατόν να μετρηθεί και η συναισθηματική διάσταση της ενεργής συμμετοχής, δ) τα μαθησιακά αποτελέσματα και το ποσοστό ολοκλήρωσης ανατιθέμενων εργασιών, που χρησιμοποιούνται ως δείκτες ενεργής συμμετοχής κυρίως στα μαζικά ανοικτά διαδικτυακά μαθήματα (MOOC).

Οι Wong & Chong (2018) προσεγγίζουν την ενεργή συμμετοχή στα ψηφιακά εκπαιδευτικά περιβάλλοντα, ως εκδήλωση ενός φάσματος από αλληλεπιδράσεις μεταξύ α) σπουδαστή και εκπαιδευτικού υλικού, β) σπουδαστή και άλλων σπουδαστών, γ) σπουδαστή και διδασκόντων. Αναπαριστούν την ενεργή συμμετοχή, ως μια τριδιάστατη εννοιολογική κατασκευή, που απαρτίζεται από τρεις δείκτες: την Αμεσότητα (Immediacy), τη Συχνότητα (Frequency) και τη Διάρκεια (Duration) (Σχήμα 8). Θεωρούν ότι η μελέτη αυτών των τριών δεικτών προσφέρει μια επαρκή μέθοδο για τη μέτρηση της ενεργής συμμετοχής σε ψηφιακά εκπαιδευτικά περιβάλλοντα.



Σχήμα 8. Δείκτες ενεργής συμμετοχής (Wong & Chong, 2018, σ. 8)

Η Αμεσότητα είναι η χρονική απόσταση μεταξύ της στιγμής που το εκπαιδευτικό υλικό, ή η εκπαιδευτική δραστηριότητα είναι διαθέσιμα και της στιγμής που ο σπουδαστής πραγματοποιεί την πρώτη πρόσβαση στο υλικό, ή στη δραστηριότητα. Με αυτή την έννοια, είναι αναμενόμενο ότι ο αυξημένος βαθμός ενεργής συμμετοχής σχετίζεται με βραχύτερη Αμεσότητα, διότι ο ενεργά συμμετέχων φοιτητής αναμένεται να σπεύσει για να αποκτήσει πρόσβαση στο υλικό, ή στη δραστηριότητα. Η Συχνότητα είναι το πλήθος των συνόδων σύνδεσης, εντός συγκεκριμένης χρονικής περιόδου (π.χ από την έναρξη έως το τέλος του μαθήματος, ή από τη στιγμή της πρώτης έως τη στιγμή της τελευταίας συνόδου σύνδεσης). Με αυτήν την έννοια, είναι αναμενόμενο ότι

ο αυξημένος βαθμός ενεργής συμμετοχής σχετίζεται με μεγαλύτερη Συχνότητα, διότι ο ενεργά συμμετέχων σπουδαστής πραγματοποιεί περισσότερες συνόδους σύνδεσης, προκειμένου να αξιοποιήσει, όσο δυνατόν περισσότερες από τις μαθησιακές δυνατότητες, που του προσφέρονται. Η Διάρκεια είναι ο συνολικός χρόνος που διαρκούν όλοι οι σύνοδοι σύνδεσης του σπουδαστή. Με αυτήν την έννοια, είναι αναμενόμενο ότι ο αυξημένος βαθμός ενεργής συμμετοχής σχετίζεται με μεγαλύτερη Διάρκεια, διότι ο ενεργά συμμετέχων σπουδαστής αφιερώνει περισσότερο χρόνο στη χρήση του συστήματος για να μελετήσει το εκπαιδευτικό υλικό και για να υλοποιήσει τις εκπαιδευτικές δραστηριότητες. Οι Wong & Chong (2018, σ. 6) επιπλέον αναφέρουν, ότι η Κανονικότητα των μεσοδιαστημάτων μεταξύ των συνόδων σύνδεσης αναπαριστά ενεργή συμμετοχή των σπουδαστών.

Οι Motz, Quick, Schroeder, Zook & Gunkel (2019, σ. 301) εντοπίζουν συμπεριφοριστικές διάστασεις της ενεργής συμμετοχής στα ψηφιακά εκπαιδευτικά περιβάλλοντα, όταν ο σπουδαστής παρίσταται (*attending*), παίρνει μέρος (*participating*) και διεκπεραιώνει ανατιθέμενη εργασία (*completing coursework*). Σύμφωνα με τα ευρήματά τους, σημαντικοί δείκτες της ενεργής συμμετοχής είναι το πλήθος των υποβληθέντων εργασιών, ο μεγάλος αριθμός συνόδων σύνδεσης και η διάρκεια αδράνειας στη χρήση του ΣΔΜ. Οι Chen & Cui (2020, σ. 5), μελετώντας τις μακράς διάρκειας συμπεριφορές στη χρήση του ΣΔΜ από τους σπουδαστές, συμπεραίνουν ότι η ασυνέπεια (*discrepancy*) χρήσης αποτελεί δείκτη ενεργής συμμετοχής. Οι Raga Jr & Raga (2017, σ. 6) θεωρούν ως δείκτες ενεργής συμμετοχής, τις ενέργειες που σχετίζονται με συμμετοχή σε δημόσιους διαδικτυακούς διαλόγους. Οι Hussain, Zhu, Zhang & Abidi (2018, σ. 7) αναφέρουν, ότι ο συνολικός αριθμός των «κλικ», παλαιότερα, «*θεωρείτο*» (*was assumed*) δείκτης ενεργής συμμετοχής, αλλά πλέον θεωρείται ανεπαρκής για να την αποτιμήσει. Εναλλακτικά, προτείνουν ως δείκτη, το πλήθος των δραστηριοτήτων, στις οποίες απέκτησε πρόσβαση ο σπουδαστής (σ.8). Ο Zacharis (2015) αναφέρει, ως δείκτες ενεργής συμμετοχής, τη συχνότητα πρόσβασης και το συνολικό χρόνο χρήσης του ΣΔΜ. Οι Buschetto Macarini et al. (2019) υποστηρίζουν, ότι οι αλληλεπιδράσεις των σπουδαστών συνδέονται στενά με την ενεργή συμμετοχή. Οι Soffer & Cohen (2019) θεωρούν ότι ενεργή συμμετοχή σχετίζεται με την πρόσβαση σε εκπαιδευτικό υλικό και με την αλληλεπίδραση με τους άλλους σπουδαστές.

Ο Πίνακας 6 συνοψίζει του δείκτες της ενεργής συμμετοχής που προτείνουν οι ερευνητές.



**Πίνακας 6. Δείκτες ενεργής συμμετοχής**

A.A.	Δείκτης ενεργής συμμετοχής	Vytasek et al. (2020)	Wong & Chong (2018)	Zacharis (2015)	Buschetto Macarini et al. (2019)	Soffer & Cohen (2019)	Motz et al. (2019)	Chen & Cui (2020)	Raga Jr & Raga (2017)	Hussain et al. (2018)
1	Διάρκεια σύνδεσης	√	√	√						
2	Πλήθος πόρων ή δραστηριοτήτων ή εργασιών	√					√			√
3	Αλληλεπιδράσεις	√			√	√			√	
4	Μαθησιακά αποτελέσματα	√								
5	Ποσοστό ολοκλήρωσης εργασιών	√								
6	Αμεσότητα σύνδεσης		√							
7	Συχνότητα σύνδεσης		√	√						
8	Κανονικότητα		√							
9	Σύνοδοι σύνδεσης						√			
10	Διάρκεια Αδράνειας						√			
11	Ασυνέπεια							√		

## 6. ΟΦΕΛΗ ΤΗΣ ΨΗΦΙΑΚΗΣ ΑΞΙΟΛΟΓΗΣΗΣ

Οι πρώτες προσπάθειες για την αξιοποίηση της τεχνολογίας στην εκπαιδευτική αξιολόγηση επιδίωκαν πρωτίστως την αύξηση της απόδοσης των αξιολογικών διαδικασιών μεγάλης κλίμακας. Επιδίωκαν δηλαδή να επιταχύνουν τη διαδικασία, να ελαφρύνουν τον φόρτο εργασίας των διδασκόντων και των διοικητικών στελεχών, να μειώσουν το κόστος υλοποίησης (Pellegrino & Quellmalz, 2010). Όμως, πλέον, η ψηφιακή αξιολόγηση αποφέρει οφέλη που δεν σχετίζονται μόνο με ποσοτικές παραμέτρους. Παρέχει δυνατότητες που σχετίζονται και με ποιοτικές παραμέτρους και αναβαθμίζουν τις πρακτικές αξιολόγησης, ώστε να είναι περισσότερο αξιόπιστες, έγκυρες, αυθεντικές, ευέλικτες, παιδαγωγικά ορθές (Oldfield et al., 2012).

### 6.1. Οφέλη στις βασικές μορφές ψηφιακής αξιολόγησης

Η βιβλιογραφία βρίθει αναφορών για τα οφέλη που προκύπτουν από τη χρήση της ψηφιακής αξιολόγησης. Οι περισσότερες αναφορές περιγράφουν οφέλη που σχετίζονται με τις λεγόμενες «κλειστού τύπου» ή «αντικειμενικές» ψηφιακές αξιολογήσεις. Δηλαδή με αξιολογήσεις που περιλαμβάνουν ερωτήσεις «κλειστού τύπου», όπου οι ορθές απαντήσεις είναι προκαθορισμένες. Παράδειγμα τέτοιου είδους αξιολόγησης είναι οι ερωτήσεις πολλαπλών επιλογών, που όπως έχει αναφερθεί (βλ. Κεφ. 2) κυριαρχούσαν στις πρώτες γενιές συστημάτων ψηφιακής αξιολόγησης (Bennett, 2015. Redecker, 2013). Ακολουθεί κατάλογος με τα οφέλη που προκύπτουν από τις «κλειστού τύπου» ψηφιακές αξιολογήσεις. Στον κατάλογο προτάσσονται τα οφέλη που αφορούν κυρίως τους εκπαιδευτικούς οργανισμούς και τους διδάσκοντες και ακολουθούν τα οφέλη που αφορούν τους σπουδαστές.

- Μείωση του φόρτου εργασίας που απαιτείται για τη διόρθωση, τη βαθμολόγηση των διαγωνισμάτων, την έκδοση και την ανακοίνωση των αποτελεσμάτων. Το όφελος αυξάνει, όσο αυξάνει το πλήθος των αξιολογούμενων σπουδαστών και το πλήθος των αντικειμένων (π.χ. των απαντήσεων) που πρέπει να ελεγχθούν και να βαθμολογηθούν. Το συγκεκριμένο όφελος αποτέλεσε την βασικότερη αρχική επιδίωξη κατά την αξιοποίηση συστημάτων ψηφιακής αξιολόγησης. Στις περιπτώσεις διαγωνισμάτων με ερωτήσεις κλειστού τύπου, η αυτοματοποίηση της διαδικασίας πρακτικά μηδενίζει τον

απαιτούμενο χρόνο διότι η διόρθωση και η αξιολόγηση γίνονται ακαριαία (Arriah και van Tonder, 2018. Azevedo et al., 2019. Bull & Danson, 2001. Conrad & Openo, 2018. Crisp, 2010. JISC, 2010. Jordan, 2013. Mora et al., 2012. Redecker, 2013).

- Εύκολη και σχεδόν ανέξοδη κλιμάκωση της αξιολόγησης για μεγαλύτερο πλήθος σπουδαστών με την εξαίρεση των περιπτώσεων ταυτόχρονης, επιτηρούμενης αξιολογικής δοκιμασίας (Azevedo et al., 2019).
- Η διαχείριση των ερωτήσεων με τη βοήθεια Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ) διευκολύνει την αποθήκευση, την επεξεργασία, την αναπαραγωγή, την ανασύνθεση και την επαναχρησιμοποίηση τους. Ο διδάσκοντας μπορεί εύκολα να αναδιαμορφώνει τις ερωτήσεις και να τις αναπροσαρμόζει συνεχώς για επόμενες αξιολογήσεις και για διαφορετικές ομάδες αξιολογούμενων (Azevedo et al., 2019. Cook & Jenkins, 2010. Redecker, 2013).
- Η δημιουργία και η χρήση μεγάλων τραπεζών ερωτήσεων επιτρέπει την αξιολόγηση μεγαλύτερου μέρους της διδακτέας ύλης και εντός μικρότερου χρόνου εξέτασης σε σχέση με τα παραδοσιακά διαγωνίσματα, που συνήθως περιλαμβάνουν λίγες ερωτήσεις ελεύθερης ανάπτυξης κειμένου (Azevedo et al., 2019).
- Επιτυγχάνεται υψηλός βαθμός αξιοπιστίας, ιδιαίτερα στις περιπτώσεις που γίνεται χρήση των λεγόμενων «κλειστού τύπου» αξιολογήσεων, όπου δεν μπορεί να υπάρξει αμφιβολία για την ορθότητα της απάντησης και επίσης δεν τίθεται ζήτημα μερικώς ορθής απάντησης. Ο βαθμός αξιοπιστίας αυξάνει περαιτέρω, όταν κατά την αξιολόγηση μειώνεται η συμμετοχή του αξιολογητή, ο οποίος είναι δυνατόν να υποπέσει σε λάθη λόγω απροσεξίας ή κόπωσης ή να βαθμολογήσει με υποκειμενικά κριτήρια. Αντιθέτως, τα συστήματα δεν υποπίπτουν σε λάθη, δεν κουράζονται, δεν βαθμολογούν με υποκειμενισμό (Arriah και van Tonder, 2018. Azevedo et al., 2019. Conrad & Openo, 2018. Cook & Jenkins, 2010. JISC, 2010. Jordan, 2013).
- Υποστηρίζονται πλουσιότερες (σε σχέση με τις παραδοσιακές αξιολογήσεις) μορφές παρουσίασης των αντικειμένων της αξιολόγησης. Για παράδειγμα, αξιοποιούνται κινούμενες εικόνες (animation), ήχοι, βίντεο. Αντιστοίχως, πολλές είναι και οι επιλογές

ως προς τη μέθοδο με την οποία ο σπουδαστής υποβάλει την απάντησή του. Για παράδειγμα είναι δυνατόν να αλληλοεπιδράσει με το σύστημα μέσω του πληκτρολογίου ή του ποντικιού, μέσω οθονών αφής ή άλλων ψηφιακών συσκευών εισόδου, να εισάγει κείμενο, να μετακινήσει εικόνες (Azevedo et al., 2019. Bull & Danson, 2001. Cook & Jenkins, 2010. Jordan, 2013. Mora et al., 2012. Redecker, 2013).

- Η δυσκολία της δοκιμασίας είναι δυνατόν να προσαρμόζεται δυναμικά, ανάλογα με γνωστικό επίπεδο του σπουδαστή. Το σύστημα χρησιμοποιεί τις προηγούμενες απαντήσεις του σπουδαστή ώστε να υποβάλλει ερωτήσεις (ή γενικά δοκιμασίες), οι οποίες είναι ανάλογες με το γνωστικό επίπεδό του, όπως αυτό συνεχώς αποτιμάται από την τρέχουσα επίδοσή του. Με το τρόπο αυτό, η διαδικασία δεν είναι πληκτική για τους σπουδαστές με υψηλό επίπεδο, ούτε εξαιρετικά δύσκολη για τους σπουδαστές με μέτριο επίπεδο. Για την υλοποίηση της συνεχώς προσαρμοζόμενης αξιολόγησης, είναι απαραίτητη η ύπαρξη τράπεζας ερωτήσεων (ή γενικά αντικειμένων αξιολόγησης) διαβαθμισμένης δυσκολίας, από όπου αντλούνται κάθε φορά οι ερωτήσεις. Για την επιλογή των κατάλληλων κάθε φορά ερωτήσεων, χρησιμοποιούνται αλγόριθμοι βασισμένοι σε στατιστικά μοντέλα και σε δέντρα αποφάσεων (Jordan, 2013).
- Υφίσταται δυνατότητα συνεχών επαναλήψεων της αξιολογικής διαδικασίας, είτε απaráλλακτα, είτε με μικρές διαφορές στο περιεχόμενο των ερωτήσεων. Η συμμετοχή του σπουδαστή σε επαναλαμβανόμενες δοκιμασίες τον βοηθά να διαγνώσει ποιο είναι το επίπεδο των γνώσεων του, να συγκρίνει την τρέχουσα με τις προηγούμενες επιδόσεις του και να καταγράφει την πρόοδο που έχει επιτελέσει. Επιπλέον, οι επαναλήψεις είναι δυνατόν να υλοποιούνται αυτόματα και άκοπα για τους σπουδαστές και τους διδάσκοντες και με μειωμένο διαχειριστικό κόστος για τον εκπαιδευτικό οργανισμό. Οι επαναλήψεις δεν απαιτούν δέσμευση ανθρώπινων πόρων. Διότι το αξιολογικό υλικό έχει ήδη παραχθεί, ενώ οι τεχνολογικές υποδομές είναι διαθέσιμες και ρυθμίζονται εύκολα για να επαναλάβουν την αξιολογική διαδικασία (Arriah και van Tonder, 2018. Azevedo et al., 2019. Bull & Danson, 2001).

- Η αξιολόγηση διεξάγεται από οποιοδήποτε γεωγραφικό σημείο και οποιαδήποτε χρονική στιγμή. Ο σπουδαστής δεν είναι απαραίτητο να παρευρίσκεται σε συγκεκριμένο χώρο. Μπορεί να συνδεθεί από οποιοδήποτε, αρκεί να διαθέτει κατάλληλο τεχνολογικό εξοπλισμό και πρόσβαση στο διαδίκτυο, μέσω υπολογιστή ή μέσω άλλης ψηφιακής συσκευής. Η συμμετοχή στην αξιολόγηση δεν γίνεται απαραίτητως σε συγκεκριμένη χρονική στιγμή. Μπορεί να υλοποιηθεί ασύγχρονα, δηλαδή σε άλλη χρονική στιγμή από αυτήν που θα διανεμηθούν τα αντικείμενα της αξιολόγησης, δυνατότητα που αξιοποιείται συχνότατα σε διαδικασίες διαμορφωτικής αξιολόγησης. Επιπλέον, ο διδάσκοντας, ή γενικότερα ο αξιολογητής συντάσσει και διανέμει τα αντικείμενα αξιολόγησης και αξιολογεί τις απαντήσεις ή τις υποβληθείσες εργασίες μέσω του διαδικτύου, χωρίς να είναι απαραίτητη η παρουσία του σε συγκεκριμένο χώρο. Περαιτέρω ευελιξία είναι δυνατή, μέσω της χρήσης φορητών ψηφιακών συσκευών (τηλεφώνων ή «ταμπλετών»), που προσφέρουν στους σπουδαστές απεριόριστες δυνατότητες ως προς το χρόνο και τον τόπο της αλληλεπίδρασης με το σύστημα της αξιολόγησης. Η ευελιξία, ως προς τον χρόνο και τον τόπο της αξιολόγησης, διευκολύνει την προσαρμογή της διαδικασίας στις ιδιαίτερες ανάγκες των σπουδαστών, ιδιαίτερα αυτών που έχουν παράλληλες εργασιακές, οικογενειακές, κοινωνικές υποχρεώσεις και συμμετέχουν σε προγράμματα σπουδών «μερικής» (part time) φοίτησης (Arriah και van Tonder, 2018. Azevedo et al., 2019. Conrad & Openo, 2018. Cook & Jenkins, 2010. JISC, 2010. Weleschuk, Dyjur & Kelly, 2019).
- Παρέχεται άμεση ανατροφοδότηση, που υποστηρίζει τους σπουδαστές και τους βοηθά να βελτιώσουν τη μαθησιακή τους πορεία. Η άμεση ανατροφοδότηση φαίνεται να αποτελεί το χαρακτηριστικό που οι σπουδαστές εκτιμούν ιδιαίτερα. Επιπλέον η άμεση ανατροφοδότηση παρέχει πληροφορίες στον διδάσκοντα σχετικά με την πρόοδο των σπουδαστών και την αποτελεσματικότητα των διδακτικών μεθόδων του και τον βοηθά να προβεί σε αναγκαίες αναπροσαρμογές. Η άμεση ανατροφοδότηση προς τους σπουδαστές και τους διδάσκοντες είναι εξαιρετικά χρήσιμη σε διαδικασίες διαμορφωτικής και συνεχούς αξιολόγησης (Arriah και van Tonder, 2018. Azevedo et al., 2019. Crisp, 2010. Conrad & Openo, 2018. JISC, 2010).

- Η ανατροφοδότηση που λαμβάνουν οι σπουδαστές είναι απρόσωπη, μη δημόσια και συνήθως δεν περιλαμβάνει αξιολογική κρίση, γεγονός που δεν προκαλεί αίσθημα μειονεξίας στους σπουδαστές και δεν τους αποθαρρύνει (υπό τον φόβο του λάθους και των δημόσιων δυσμενών σχολίων) να συμμετάσχουν σε συνεχείς επαναλήψεις της αξιολογικής διαδικασίας (Azevedo et al., 2019. Jordan, 2013. Weleschuk, Dyjur & Kelly, 2019).
- Οι σπουδαστές μπορούν να επισκοπούν με ευκολία προηγούμενες ερωτήσεις και να αναθεωρούν τις απαντήσεις τους πριν την οριστικοποίησή τους, χωρίς το φόβο να αλλοιώσουν το έντυπο των απαντήσεων, διότι δεν υφίσταται έντυπο, όπως αυτό των παραδοσιακών αξιολογήσεων (Azevedo et al., 2019. Cook & Jenkins, 2010).

## 6.2. Οφέλη στις καινοτόμες μορφές ψηφιακής αξιολόγησης

Επιπλέον των παραπάνω οφελών που αφορούν κυρίως τις λεγόμενες «κλειστού τύπου» ψηφιακές αξιολογήσεις, στη βιβλιογραφία καταγράφονται οφέλη που προκύπτουν από την αξιοποίηση νέων καινοτόμων μορφών ψηφιακής αξιολόγησης.

- Η συλλογή και ανάλυση μεγάλου όγκου δεδομένων σχετικών με τους σπουδαστές (στα πλαίσια αξιοποίησης μεθόδων Μαθησιακής Αναλυτικής), ως εργαλείο αξιολόγησης και αντιστρόφως, η συλλογή και ανάλυση δεδομένων από την αξιολόγηση («Αξιολογική Αναλυτική» / assessment analytics), ως εργαλείο για τη βελτίωση της μάθησης, οδηγεί σε ώσμωση των δύο διαδικασιών (Jordan, 2013). Με τον τρόπο αυτό είναι υλοποιήσιμη η μορφή αξιολόγησης, που περιγράφεται ως «αόρατη» (stealth) αξιολόγηση (Shute & Kim, 2014), δηλαδή αξιολόγηση που δεν αποτελεί διακριτή δραστηριότητα, αλλά διαχέεται στην μαθησιακή διαδικασία, με όφελος αφενός την ευπροσάρμοστη και τη συνεχή αξιολόγηση των σπουδαστών (Timmis et al. 2016) και αφετέρου τη μείωση του άγχους και της αναστάτωσης, που συνήθως προκαλεί η αξιολόγηση στους σπουδαστές (Oldfield et al., 2012).
- Η ανάλυση ποσοτικών και ποιοτικών δεδομένων που αφορούν τους σπουδαστές, στα πλαίσια αξιοποίησης μεθόδων Μαθησιακής Αναλυτικής οδηγεί σε πολύτιμα συμπεράσματα σχετικά με την απόδοση και τη μαθησιακή πρόοδο τους. Διευκολύνει

συγκριτικές μελέτες σχετικά με την απόδοση και την πρόοδο ομάδων σπουδαστών με διαφορετικά προσωπικά, κοινωνικά, μαθησιακά χαρακτηριστικά. Τα δεδομένα σχετίζονται όχι μόνο με την απόδοση, την πρόοδο, τη μαθησιακή πορεία των σπουδαστών, αλλά και με τις αλληλοεπιδράσεις τους. Παρέχουν δυνατότητες για να μελετηθούν οι στάσεις των σπουδαστών έναντι της μάθησης και τα κίνητρα που τους παρωθούν. Με τον τρόπο αυτό είναι δυνατόν να αναπτυχθούν εξειδικευμένες στρατηγικές μάθησης προσαρμοσμένες σε συγκεκριμένους σπουδαστές ή σε ομάδες σπουδαστών με κοινά χαρακτηριστικά, με στόχο την αποτελεσματικότερη μάθηση και την ενθάρρυνση της ενεργής συμμετοχής στη μαθησιακή διαδικασία. Συμπεράσματα είναι δυνατόν να προκύψουν και για την ποιότητα του υλικού της αξιολόγησης και για την αποτελεσματικότητα, την εγκυρότητα και την αξιοπιστία της αξιολογικής διαδικασίας (Jordan, 2013).

- Η συλλογή και ανάλυση μεγάλου όγκου δεδομένων από τις απαντήσεις των σπουδαστών σε τοπικό, περιφερειακό, εθνικό και διεθνές επίπεδο («Ακαδημαϊκή Αναλυτική») δίνει τη δυνατότητα στους διδάσκοντες και στους σχεδιαστές της εκπαιδευτικής πολιτικής να εξάγουν συμπεράσματα, σχετικά με την πρόοδο μεγάλων ομάδων σπουδαστών και σχετικά με την αποτελεσματικότητα των προγραμμάτων σπουδών και των εκπαιδευτικών πολιτικών. Επιπλέον σε εθνικό επίπεδο, οι σχεδιαστές της εκπαιδευτικής πολιτικής χρησιμοποιούν τα δεδομένα για να αποτιμήσουν τη συνολική απόδοση του εκπαιδευτικού συστήματος και να συγκρίνουν τους εθνικούς με τους διεθνείς δείκτες.
- Συστήματα και διαδικασίες αξιολόγησης που αξιοποιούν τα ψηφιακά κοινωνικά δίκτυα, τα ιστολόγια, τους χώρους ανάπτυξης συνεργατικού περιεχομένου (wikis), τους ψηφιακούς φάκελους επιτευγμάτων (e-portfolios) παρέχουν δυνατότητες για την υλοποίηση αυτοαξιολόγησης και ομότιμης αξιολόγησης. Με την αυτοαξιολόγηση, οι σπουδαστές αναλαμβάνουν την ευθύνη της μάθησής τους και έχουν μεγαλύτερο κίνητρο για να συμμετάσχουν στην εκπαιδευτική διαδικασία. Με την ομότιμη αξιολόγηση, οι σπουδαστές δεν αποτελούν παθητικά υποκείμενα αξιολόγησης. Η θέση τους αναβαθμίζεται, αποκτούν ενεργό ρόλο και διαμοιράζονται την ευθύνη της αξιολογικής διαδικασίας. Έτσι, ενισχύονται η κριτική σκέψη και η ικανότητα λήψης αποφάσεων

(Timmis et al., 2016) και αναπτύσσονται στάσεις αυτόνομης, αυτορρυθμιζόμενης και διαρκούς μάθησης, οι οποίες βελτιώνουν την ποιότητα και την αποτελεσματικότητα της μαθησιακής διαδικασίας (Karsalis, et al., 2019). Γενικά, η χρήση διαδικτυακών αλληλεπιδραστικών συστημάτων ενισχύει την ενεργή συμμετοχή των σπουδαστών στην μαθησιακή διαδικασία, αξιοποιώντας συμμετοχικές, διαδραστικές και ομότιμες μορφές αξιολόγησης.

- Οι προσομοιώσεις, οι εικονικοί κόσμοι, τα «σοβαρά» παιχνίδια διευκολύνουν την αξιολόγηση ανώτερων γνωστικών ικανοτήτων και σύνθετων μαθησιακών λειτουργιών, όπως ο έλεγχος της ορθότητας υποθέσεων, η επιτέλεση ρόλων (role playing), η επίλυση σύνθετων προβλημάτων, η λήψη αποφάσεων στα πλαίσια πολύπλοκων και δυναμικών καταστάσεων. Επιτρέπουν την αξιολόγηση των σπουδαστών υπό ρεαλιστικές συνθήκες, δηλαδή υπό συνθήκες κατά το δυνατόν όμοιες με αυτές που ισχύουν σε πραγματικά περιβάλλοντα και σε πραγματικούς χώρους εργασίας. Οι σπουδαστές αντιλαμβάνονται σε πραγματικό χρόνο τις συνέπειες των αποφάσεων που έχουν λάβει και των ενεργειών στις οποίες έχουν προβεί. Είναι δυνατόν να πειραματιστούν με διαφορετικές αποφάσεις και ενέργειες και να συγκρίνουν τα αποτελέσματα διαφορετικών στρατηγικών, δηλαδή να λάβουν άμεση, διαδραστική, δυναμικά προσαρμοζόμενη ανατροφοδότηση. Μέσω προσομοιώσεων, εικονικών κόσμων και εκπαιδευτικών παιχνιδιών, είναι δυνατόν να αξιολογηθούν οι ικανότητες της συνεργασίας, της ενσυναίσθησης, της διαπραγμάτευσης της στρατηγικής σκέψης, της λήψης πρωτοβουλιών, του πειραματισμού (Redecker, 2013). Παρέχεται η δυνατότητα αξιολόγησης δοκιμασιών, οι οποίες είναι δύσκολο, ή επικίνδυνο, ή αδύνατο να υλοποιηθούν σε πραγματικές συνθήκες (Sweeney, et al., 2017). Με τον τρόπο αυτόν, ενισχύεται η ενεργή συμμετοχή και η ικανοποίηση των σπουδαστών και αυξάνεται ο βαθμός αυθεντικότητας της αξιολόγησης (Conrad & Oropo, 2018. JISC, 2010. Karsalis, et al., 2019). Αξιολογήσεις που διενεργούνται στα πλαίσια εικονικών κόσμων αυξάνουν το ενδιαφέρον των σπουδαστών να συμμετάσχουν. Για αρκετούς, συμπεριλαμβανομένων των σπουδαστών με ειδικές ανάγκες, η ψηφιακή αξιολόγηση στα πλαίσια εικονικών κόσμων μπορεί να προσφέρει πλουσιότερες, πιο ενδιαφέρουσες και εγκυρότερες εμπειρίες, σε σύγκριση με τις παραδοσιακές



αξιολογήσεις (JISC, 2007), ωστόσο ζητούμενο αποτελεί η αξιοποίηση τέτοιων συστημάτων στις τελικές, κρίσιμης σημασίας αξιολογήσεις.

- Συστήματα που υποστηρίζουν αξιολόγηση, μέσω ψηφιακών φακέλων επιτευγμάτων (e-portfolio), ιστολογίων (blogs), χώρων ανάπτυξης συνεργατικού περιεχομένου (wiki), χώρων δημόσιου διαλόγου (forum) και άλλων Web 2.0 εφαρμογών, ενισχύουν τη συμμετοχή, τη συνεργασία και τον αναστοχασμό. Οι ψηφιακοί φάκελοι επιτευγμάτων παρέχουν στους σπουδαστές τη δυνατότητα να αποθηκεύσουν τα τεκμήρια της μαθησιακής πορείας τους και να στοχαστούν για αυτά. Με τον τρόπο αυτόν, είναι δυνατή η αξιολόγηση ικανοτήτων, που είναι δύσκολο να αξιολογηθούν διαφορετικά. Επιπλέον ενθαρρύνεται η στοχαστική προσέγγιση της μάθησης και οι σπουδαστές νιώθουν την ευθύνη για τα περιεχόμενα του φακέλου τους και εστιάζουν στα θετικά επιτεύγματά τους. Οι διδάσκοντες είναι εύκολο να αποκτήσουν πρόσβαση στα περιεχόμενα των φακέλων για να αξιολογήσουν, ή για να παράσχουν ανατροφοδότηση προς τους σπουδαστές. Επίσης, οι ψηφιακοί φάκελοι επιτευγμάτων ενδυναμώνουν το προφίλ του σπουδαστή κατά την αναζήτηση εργασίας. Τα οφέλη των ψηφιακών φακέλων επιτευγμάτων ενισχύονται με παράλληλη και συμπληρωματική χρήση ιστολογίων (blogs), μέσω των οποίων οι σπουδαστές καταγράφουν και δημοσιεύουν τις εμπειρίες και τις σκέψεις τους (Jordan, 2013). Οι χώροι ανάπτυξης συνεργατικού περιεχομένου (wikis) διευκολύνουν την αξιολόγηση έργων που έχουν αναπτυχθεί συνεργατικά. Τα έργα κατασκευάζονται ομαδικά από ομάδα σπουδαστών, αλλά η συνεισφορά κάθε μέλους, είτε πρόκειται για προσθήκη, είτε για διόρθωση, καταγράφεται. Με τον τρόπο αυτό είναι δυνατή, τόσο η συνολική αξιολόγηση της ομάδας, όσο και η ατομική αξιολόγηση κάθε μέλους της ομάδας. Με το ίδιο τρόπο, στους χώρους δημόσιου διαλόγου, όπου διεξάγονται ομαδικές συζητήσεις, είναι δυνατή η ατομική αξιολόγηση κάθε μέλους της ομάδας, το οποίο συμμετέχει στον διάλογο, αλληλοεπιδρά με τα υπόλοιπα μέλη και συνεισφέρει στο θέμα του διαλόγου (Jordan, 2013).
- Οι δυναμικές αλληλεπιδράσεις και οι συνεργατικές δραστηριότητες που λαμβάνουν χώρα στα πλαίσια διαμορφωτικών αξιολογικών διαδικασιών στα ψηφιακά

περιβάλλοντα μάθησης παράγουν περισσότερο άμεση, αποτελεσματική και συνεχόμενη διαμορφωτική ανατροφοδότηση, η οποία ενισχύει τα κίνητρα για μάθηση, την αυτοεκτίμηση των σπουδαστών, την αυτορρυθμιζόμενη μάθηση. Επιπλέον, παρέχουν ικανοποίηση στους σπουδαστές και τους παρωθούν σε περισσότερο ενεργή συμμετοχή στην εκπαιδευτική διαδικασία (Gikandi et al. 2011).

- Δίνονται δυνατότητες για πολλαπλούς και εναλλακτικούς τρόπους αξιολόγησης, με τη βοήθεια διαφορετικών τεχνολογικών εργαλείων και με ανάλυση διαφορετικών δεικτών απόδοσης. Με τον τρόπο αυτό διευκολύνεται η προσαρμογή της αξιολογικής διαδικασίας στις ιδιαίτερες ανάγκες κάθε σπουδαστή και κατά συνέπεια, προάγεται η ισότιμη μάθηση και περιορίζονται φαινόμενα αποκλεισμού, ή ανισονομίας για τους σπουδαστές που υστερούν σε μαθησιακές ικανότητες και οι οποίοι δεν είναι εύκολο να αξιολογηθούν με τις παραδοσιακές μεθόδους (Gikandi et al., 2011)
- Καινοτόμα εναλλακτικά συστήματα απονομής διακρίσεων, όπως τα «ψηφιακά παράσημα» (online badges) χρησιμοποιούνται ευρέως στις κοινότητες των διαδικτυακών παιχνιδιών και αξιοποιούνται από Μαζικά Δικτυακά Ανοικτά Μαθήματα (Massive Online Open Courses) ως ένα εναλλακτικό, κοινά αποδεκτό, σύστημα επικύρωσης και αναγνώρισης επιτευγμάτων και κατοχής γνώσεων ή δεξιοτήτων (Timmis et al. 2016).

Καθώς τα συστήματα εξελίσσονται και νέες τεχνολογίες συνεχώς αναπτύσσονται, ο κατάλογος των ωφελειών συνεχώς διευρύνεται. Ωστόσο, είναι αναγκαίο να επισημανθεί ότι τα οφέλη δεν θα επιτευχθούν αν η ψηφιακή αξιολόγηση υλοποιηθεί με τυχαίο και άναρχο τρόπο. Αν δηλαδή οι εκπαιδευτικοί οργανισμοί δεν σχεδιάσουν και δεν εφαρμόσουν συνεκτική στρατηγική βασισμένη σε στέρεες παιδαγωγικές αρχές, η οποία θα στοχεύει, όχι μόνο στην ανάπτυξη των απαραίτητων τεχνολογικών υποδομών, αλλά επιπλέον θα μεριμνά για την εκπαίδευση και τη συνεχή τεχνική, διοικητική, παιδαγωγική υποστήριξη των διδασκόντων και των άλλων στελεχών που θα κληθούν να αξιοποιήσουν τα συστήματα (Sim, Holifield & Brown, 2004).

## 7. ΠΡΟΚΛΗΣΕΙΣ

Η αξιοποίηση διαδικασιών ψηφιακής αξιολόγησης αντιμετωπίζει σειρά ηθικών, πολιτικών παιδαγωγικών, οργανωτικών, τεχνικών προκλήσεων, οι σπουδαιότερες από τις οποίες καταγράφονται στις επόμενες παραγράφους.

### 7.1. Ηθικές και Πολιτικές προκλήσεις

Η καταγραφή και η ανάλυση δεδομένων στα πλαίσια εφαρμογών της Μαθησιακής Αναλυτικής ενέχει ηθικούς κινδύνους, που απειλούν την ιδιωτικότητα των αξιολογούμενων. Εγείρονται ερωτήματα, για το ποιος δικαιούται να συλλέγει, να αναλύει, να αξιοποιεί τέτοιους είδους πληροφορίες, χωρίς να διαθέτει σχετική εξουσιότητα (Oosterhof, Conrad & Ely, σ. 137), χωρίς να τηρεί τις κανονιστικές διατάξεις και χωρίς συχνά να έχει εξασφαλίσει τη συναίνεση των αξιολογούμενων. Οι κίνδυνοι διευρύνονται, όταν στα πλαίσια συγκεκριμένων πρακτικών αξιολόγησης, τα δεδομένα που συλλέγονται δεν σχετίζονται μόνο με τις επιδόσεις των σπουδαστών σε ψηφιακά διαγωνίσματα, αλλά με τις ψηφιακές αλληλεπιδράσεις τους, τις προτιμήσεις τους και τις συμπεριφορές τους σε μέσα κοινωνικής δικτύωσης, σε εικονικούς κόσμους, σε παιχνίδια επιτέλεσης ρόλων. Σε τέτοιες περιπτώσεις, η κριτική δεν περιορίζεται στην επάρκεια και ορθότητα των κανόνων, που πρέπει να διέπουν την συλλογή και την ανάλυση τέτοιων δεδομένων, αλλά αμφισβητεί το δικαίωμα των εκπαιδευτικών οργανισμών να εφαρμόζουν πρακτικές καταγραφής τέτοιου είδους δεδομένων (Timmis et al., 2016).

Η καταγραφή, η ανάλυση και γενικά η αξιοποίηση των δεδομένων που παράγονται από τις διαδικασίες της ψηφιακής αξιολόγησης (assessment analytics), αξιοποιούνται σε πολλές χώρες, ως μηχανισμοί μέτρησης της απόδοσης και αξιολογικής κατάταξης των εκπαιδευτικών μονάδων. Υφίσταται αυξανόμενη ανησυχία για την αξιοπιστία αυτών των πρακτικών, οι οποίες είναι δυνατόν να οδηγήσουν σε στρεβλά, παραπλανητικά έως και απαξιωτικά συμπεράσματα, ως προς την ποιότητα του εκπαιδευτικού έργου που προσφέρουν οι εκπαιδευτικές μονάδες.

## 7.2. Παιδαγωγικές προκλήσεις

Η αξιολόγηση αποτελεί αναπόσπαστο τμήμα της μαθησιακής διαδικασίας. Οποιαδήποτε αλλαγή στον τρόπο με τον οποίο υλοποιείται η αξιολόγηση απαιτεί αναστοχασμό για όλα τα στάδια και τις παραμέτρους της μαθησιακής διαδικασίας (Sweeney et al., 2017). Αν το γενικότερο εκπαιδευτικό πλαίσιο και περιβάλλον εμμένει σε παραδοσιακές μορφές και διαδικασίες μάθησης, τότε η υιοθέτηση διαδικασιών ψηφιακής αξιολόγησης είναι αναποτελεσματική. Η ψηφιακή αξιολόγηση πρέπει να αντιμετωπίζεται, όχι ως απομονωμένος πειραματισμός, αλλά ως ένας από τους κρίκους στην αλυσίδα των αλλαγών που απαιτούνται για τη βελτίωση της εκπαιδευτικής διαδικασίας (Bull & Danson, 2004).

Παρά την διαδεδομένη άποψη ότι η αξιοποίηση των ΤΠΕ στην εκπαιδευτική αξιολόγηση αποτελεί βασικό παράγοντα αλλαγής, οι καινοτόμες πρακτικές αξιολόγησης είναι λιγοστές. Οι σχετικές προσπάθειες επικεντρώνονται κυρίως στην επίτευξη ποσοτικών στόχων, όπως η αύξηση της αποδοτικότητας και η μείωση του οικονομικού και του οργανωτικού κόστους. Επίσης, οι ΤΠΕ συνήθως αντιμετωπίζονται ως μέσο για την ψηφιοποίηση παραδοσιακών πρακτικών, αντί να αξιοποιούνται για την ριζική επανασχεδίαση των αξιολογικών διαδικασιών. Όμως έτσι, περιορίζεται η αναζήτηση ευφάνταστων και δημιουργικών προτάσεων, βασισμένων σε όλες τις δυνατότητες που μπορούν να προσφέρουν οι ΤΠΕ. Ακόμα και στις περιπτώσεις καινοτόμων περιβαλλόντων ψηφιακής μάθησης (όπως για παράδειγμα εκπαιδευτικών παιχνιδιών και εικονικών κόσμων), η αξιολόγηση δεν τυγχάνει ιδιαίτερης προσοχής και λαμβάνει μικρή προτεραιότητα. Περιορισμένη είναι η διεϊσδυση της ψηφιακής αξιολόγησης και στο πεδίο της συνεργατικής ψηφιακής μάθησης. Η έμφαση που δίδεται για την αξιολόγηση των αλληλοεπιδράσεων και της συνεργασίας, στα πλαίσια σχετικών ψηφιακών μαθησιακών διαδικασιών, είναι μικρή. Διότι, οι διδάσκοντες και οι εκπαιδευτικοί οργανισμοί, αφενός επικεντρώνονται στην ατομική εκπαιδευτική επίδοση των σπουδαστών και αφετέρου θεωρούν ότι η συνεργατική ή η ομότιμη αξιολόγηση δεν είναι πάντοτε δίκαιες και αντικειμενικές. Επιπλέον, όπου αναπτύσσονται καινοτόμες πρακτικές αξιολόγησης, αυτές είναι συνήθως περιχαρακωμένες εντός των στενών ορίων συγκεκριμένων εκπαιδευτικών ιδρυμάτων και ερευνητικών κέντρων. Δεν έχουν ευρεία διάδοση και συνήθως δεν χρηματοδοτούνται επαρκώς για την περαιτέρω ανάπτυξη και

εφαρμογή τους. Είναι επίσης διαχρονικά διαπιστωμένο, ότι η ψηφιακή αξιολόγηση δεν αξιοποιείται στον ίδιο βαθμό σε όλα τα γνωστικά αντικείμενα. Συγκεκριμένα, είναι διαδεδομένη πολύ περισσότερο στις θετικές επιστήμες, έναντι των άλλων γνωστικών αντικειμένων (Timmis et al., 2016).

Οι δράσεις και τα έργα στον τομέα της ψηφιακής αξιολόγησης, αλλά και γενικότερα στο ευρύτερο πεδίο της εκπαιδευτικής τεχνολογίας, συχνά διαπνέονται από τεχνολογικό ντετερμινισμό, δηλαδή από την υπεραπλουστευμένη άποψη ότι η τεχνολογία είναι πάντοτε φορέας θετικών αλλαγών, οι οποίες προκύπτουν αυτόματα και αναπόφευκτα από τη χρήση της τεχνολογίας και μόνο. Η άποψη αυτή, συνήθως, οδηγεί σε δράσεις που είναι αποκλειστικώς τεχνολογικά εστιασμένες και οι οποίες παραβλέπουν ότι κάθε καινοτομία πρέπει να εντάσσεται αρμονικά και να λαμβάνει υπόψη το οργανωτικό, κοινωνικό και πολιτισμικό πλαίσιο, το οποίο καλείται να υποστηρίξει. Συνεπώς, η ψηφιακή αξιολόγηση όπως και κάθε άλλη εκπαιδευτική τεχνολογία πρέπει να βασίζεται σε στέρεες μαθησιακές αρχές και θεωρίες, δηλαδή να είναι κυρίως παιδαγωγικά προσανατολισμένη και όχι απλώς να επιλέγει άκριτα και να υλοποιεί οτιδήποτε είναι τεχνικά υλοποιήσιμο. Πρέπει επίσης να σχεδιάζεται από ομάδες στις οποίες συμμετέχουν πρόσωπα προερχόμενα, τόσο από τεχνολογικά, όσο και από παιδαγωγικά γνωστικά αντικείμενα, ώστε οι λύσεις να αναδύονται μέσα από τη δημιουργική σύνθεση διαφορετικών ακαδημαϊκών αντιλήψεων και τη συνεργασία επιστημόνων διαφορετικών ειδικοτήτων (Timmis et al., 2016).

Η ψηφιακή αξιολόγηση πρέπει να διασφαλίζει, ότι όλοι οι σπουδαστές έχουν τις ίδιες ευκαιρίες να συμμετάσχουν στη διαδικασία (Bull & Danson, 2004). Παρά την ευρεία διάδοση και χρήση διαδικτυακών συνδέσεων και ψηφιακών συσκευών δεν είναι σπάνιες οι περιπτώσεις όπου σπουδαστές, λόγω οικονομικών, κοινωνικών, γεωγραφικών λόγων, δεν έχουν πρόσβαση σε αναγκαίες υποδομές και αναγκαίο εξοπλισμό για τη συμμετοχή τους στην αξιολογική διαδικασία. Η ισοτιμία στην πρόσβαση διαταράσσεται επίσης στις περιπτώσεις, που οι σπουδαστές υστερούν σε δεξιότητες στη χρήση των ΤΠΕ και δεν είναι σε θέση να χρησιμοποιήσουν αποδοτικά τα ψηφιακά συστήματα, ή στις περιπτώσεις που υφίστανται μαθησιακές δυσκολίες ή λόγοι υγείας (π.χ. κινητικά προβλήματα, προβλήματα όρασης) που θέτουν κάποιους σε μειονεκτική θέση έναντι των υπολοίπων. Η ισοτιμία στην

πρόσβαση δεν διαταράσσεται μόνο εξαιτίας ελλιπών δεξιοτήτων ή ανεπαρκούς εξοπλισμού. Είναι επίσης αποτέλεσμα δισταγμού και επιφύλαξης κάποιων από τους σπουδαστές να συμμετάσχουν σε ψηφιακά περιβάλλοντα και σε ψηφιακές ομαδικές δραστηριότητες, όπως για παράδειγμα στα εργαλεία ανάπτυξης συνεργατικού περιεχομένου (wikis) ή στους ψηφιακούς χώρους συζητήσεων. Έτσι, οι καινοτόμοι τρόποι ψηφιακής αξιολόγησης, που βασίζονται σε Web 2.0 τεχνολογίες και σε μέσα κοινωνικής δικτύωσης, είναι πιθανόν να οδηγήσουν σε αντίρροπες τάσεις, ανάλογα με τις ιδιαιτερότητες των σπουδαστών. Ειδικότερα, άλλοτε ενισχύουν τις δυνατότητες ενεργής συμμετοχής και αυθεντικών αξιολογικών εμπειριών των σπουδαστών που είναι εξοικειωμένοι και παραγωγικοί στα ψηφιακά περιβάλλοντα μάθησης και αξιολόγησης. Άλλοτε, δυσχεραίνουν τη συμμετοχή όσων δεν αισθάνονται άνετα και δεν αποδίδουν ικανοποιητικά σε περιβάλλοντα τα οποία διαφέρουν από τα παραδοσιακά. Επομένως, δεν είναι δεδομένο ότι οι καινοτόμες πρακτικές διευκολύνουν και ενισχύουν τη μάθηση και την αξιολόγηση, πάντοτε και για κάθε σπουδαστή, αλλά είναι πιθανόν να προκαλέσουν ανισότητα ως προς τη συμμετοχή και την απόδοση κάποιων (Timmis et al., 2016). Ανάλογα συμπεράσματα προκύπτουν σχετικά με τις γενικότερες μαθησιακές επιδόσεις των σπουδαστών. Ειδικότερα, οι σπουδαστές με καλές επιδόσεις προσαρμόζονται εύκολα σε νέους τρόπους αξιολόγησης και αποδίδουν εξίσου καλά με τις παραδοσιακές αξιολογήσεις, σε αντίθεση με τους σπουδαστές που έχουν μέτριες επιδόσεις, οι οποίοι δυσκολεύονται περαιτέρω με την ψηφιακή αξιολόγηση (Arriah & van Tonder, 2018). Η ισότιμη πρόσβαση είναι δυνατόν να διαταραχθεί και σε περιπτώσεις που δεν συντρέχουν λόγοι υποκειμενικών δυσκολιών, ή ιδιαιτεροτήτων του σπουδαστή. Για παράδειγμα, είναι δυνατόν να προκύψουν τυχαία γεγονότα (όπως π.χ. βλάβες στον εξοπλισμό του χρήστη, διακοπή ηλεκτροδότησης, διακοπή διαδικτυακής σύνδεσης), τα οποία μπορεί να εμποδίσουν για μικρό ή μεγάλο χρονικό διάστημα την πρόσβαση του σπουδαστή στην αξιολογική διαδικασία (Bull & Danson, 2004).

### 7.3. Οργανωτικές προκλήσεις

Οι διαδικασίες της αξιολόγησης και ιδιαίτερα της τελικής αξιολόγησης, αφενός οφείλουν να συμμορφώνονται με εθνικά ή διεθνή κανονιστικά πλαίσια (Perrotta & Whitelock, 2017). Αφετέρου, οι τυπικές συνέπειες της αξιολόγησης (απονομή πτυχίων και πιστοποιήσεων,

εισαγωγή σε ανώτερη εκπαιδευτική βαθμίδα) είναι εξαιρετικά κρίσιμες (Oldfield et al., 2012). Κατά συνέπεια, ο βαθμός ευελιξίας των τελικών αξιολογικών διαδικασιών είναι περιορισμένος, δεδομένου, ότι υφίσταται σοβαρός σκεπτικισμός έναντι των πειραματισμών και των καινοτομιών.

Η ψηφιακή αξιολόγηση προϋποθέτει στρατηγικό όραμα, λεπτομερή σχεδιασμό, υποστήριξη από την ηγεσία του εκπαιδευτικού οργανισμού, παιδαγωγική και τεχνική καθοδήγηση και επιμόρφωση των διδασκόντων, των τεχνικών και των διοικητικών στελεχών, που θα κληθούν να την υλοποιήσουν (Bull & Danson, 2004. Cook & Jenkins, 2010. JISC, 2007. Weleshuk et al., 2019). Οι διδάσκοντες και οι σπουδαστές πρέπει να ενημερωθούν, να εκπαιδευτούν και να προετοιμαστούν καταλλήλως, δεδομένου ότι οι αρκετοί δεν είναι εξοικειωμένοι με αντίστοιχες διαδικασίες και δεν διαθέτουν εμπειρίες από μη παραδοσιακές μορφές αξιολόγησης.

Η προετοιμασία των ψηφιακών αξιολογήσεων είναι διαδικασία απαιτητική και χρονοβόρα (JISC, Effective Practice with e-Assessment, 2007). Συνήθως, διαρκεί περισσότερο και απαιτεί μεγαλύτερο μόχθο, σε σύγκριση με την προετοιμασία παραδοσιακών αξιολογήσεων. Ωστόσο, τελικά το ισοζύγιο αποβαίνει θετικό, λόγω της δυνατότητας επανάληψης της ψηφιακής αξιολόγησης, της εύκολης επαναχρησιμοποίησης του αξιολογικού υλικού και της εύκολης κλιμάκωσης της διαδικασίας.

Το κόστος ανάπτυξης, συντήρησης και συνεχούς εκσυγχρονισμού των υπολογιστικών και δικτυακών υποδομών είναι υψηλό (Bull & Danson, 2004). Επιπλέον, λόγω της ταχείας, συνεχούς ανάπτυξης νέων και ανταγωνιστικών συστημάτων λογισμικού, οι σχετικές επενδύσεις σε υποδομές, σε εξοπλισμό και σε λογισμικό, ενέχουν κίνδυνο να αποδειχθούν μη αποδοτικές, αν το σύστημα αξιολόγησης αποδειχθεί ότι δεν ήταν κατάλληλο. Δεν πρέπει επίσης να υποτιμάται, αφενός το κόστος απόκτησης κατάλληλου εξοπλισμού από τους σπουδαστές και αφετέρου το κόστος για τη χρήση διαδικτυακών υπηρεσιών. Παρότι τα συγκεκριμένα κόστη (για αγορά υπολογιστή και για διαδικτυακή σύνδεση) καταγράφουν φθίνουσα πορεία, δεν πρέπει να θεωρείται δεδομένο, ότι όλοι οι σπουδαστές έχουν την οικονομική δυνατότητα για να ανταποκριθούν σε αυτά.

Ο σχεδιασμός και η δημιουργία ποιοτικών ψηφιακών αξιολογικών δοκιμασιών προϋποθέτουν γνώσεις, εμπειρία, προσπάθεια και χρόνο. Αν αυτοί οι παράγοντες δεν διατίθενται από όσους σχεδιάζουν και υλοποιούν ψηφιακές αξιολογήσεις, τότε τα παραγόμενα αποτελέσματα υστερούν σε ποιότητα (Arriah & van Tonder, 2018).

#### 7.4. Τεχνικές προκλήσεις

Η διόρθωση των λαθών, των παραλείψεων και των ασαφειών στα οποία τυχόν υποπίπτει ο σχεδιαστής της ψηφιακής αξιολογικής διαδικασίας, δεν είναι εύκολη. Διότι τα συστήματα είναι πολύπλοκα, απαιτούν εξειδικευμένες γνώσεις για την διαχείρισή τους, τα αντικείμενα της αξιολόγησης διανέμονται μέσω του διαδικτύου σε πραγματικό χρόνο (στις περιπτώσεις των σύγχρονων διαδικασιών), υφίσταται φυσική απόσταση μεταξύ σπουδαστή και διδάσκοντα, το πλήθος των αξιολογούμενων συχνά είναι μεγάλο και γεωγραφικά διεσπαρμένο. Οπότε, ενώ στις παραδοσιακές αξιολογικές διαδικασίες, τα λάθη, οι παραλήψεις, οι ασάφειες είναι δυνατόν να διορθωθούν εύκολα και γρήγορα με σχετικές προφορικές ή γραπτές ανακοινώσεις και διευκρινήσεις εντός του εξεταστικού κέντρου, οι αντίστοιχες διορθωτικές ενέργειες είναι δυσκολότερο να διεκπεραιωθούν μέσω των συστημάτων ψηφιακής αξιολόγησης, διότι απαιτούν πρόσβαση στα συστήματα, γνώσεις για τη χρήση τους και γρήγορες και εύστοχες τεχνικές παρεμβάσεις. (Oosterhof, Conrad & Ely, 2007, σ. 137)

Επιπροσθέτως των παραπάνω, η βιβλιογραφία αναδεικνύει, ως σημαντικές, δύο επιπλέον προκλήσεις: το ζήτημα της αξιολόγησης των ανώτερων νοητικών ικανοτήτων και το ζήτημα της ακαδημαϊκής ακεραιότητας. Δεδομένου ότι, όπως προκύπτει από τη μελέτη της βιβλιογραφίας, αυτά τα δύο ζητήματα προσελκύουν ιδιαίτερως την προσοχή της επιστημονικής κοινότητας, επιλέξαμε να τα παρουσιάσουμε χωριστά από τις υπόλοιπες προσκλήσεις, στις παραγράφους που ακολουθούν.

#### 7.5. Η αξιολόγηση ανώτερων νοητικών ικανοτήτων

Κεντρικό ζήτημα στη βιβλιογραφία είναι η δυνατότητα των διάφορων μορφών ψηφιακής αξιολόγησης να αξιολογήσουν ανώτερες νοητικές ικανότητες και προσωπικές δεξιότητες. Η σχετική συζήτηση κυριαρχείται από το ερώτημα, αν τη σχετική δυνατότητα διαθέτουν (ή δεν



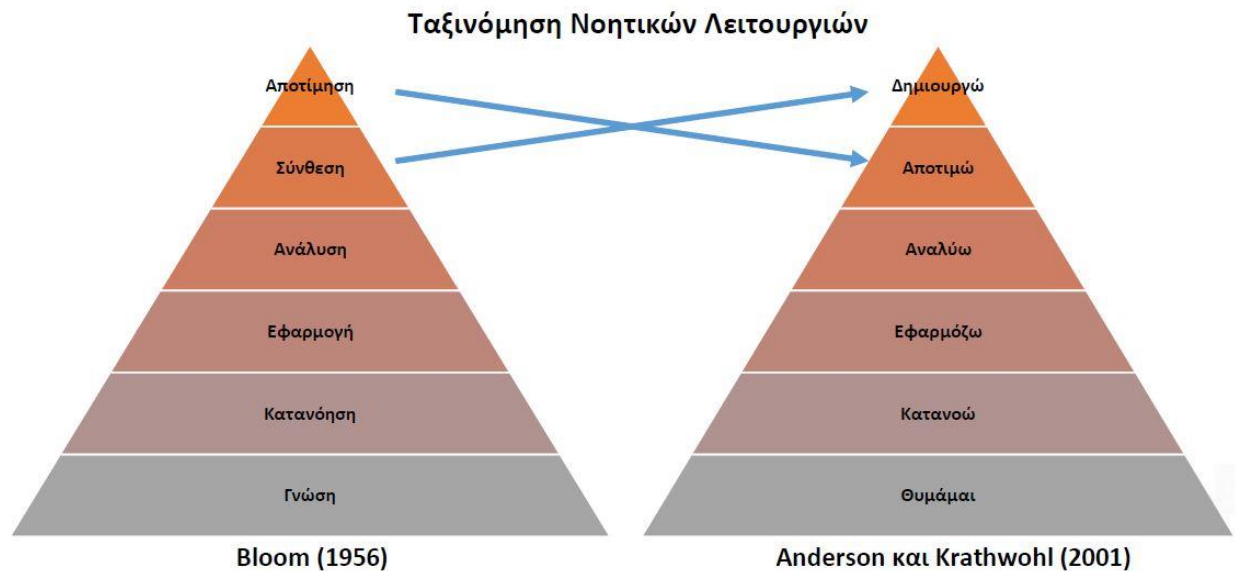
διαθέτουν) οι λεγόμενες «κλειστού τύπου» ψηφιακές αξιολογήσεις, οι οποίες κυριαρχούν στις διαδικασίες της ψηφιακής αξιολόγησης (Bennett, 2015. Redecker, 2013).

Πριν την παράθεση των απόψεων που συναντάμε στη βιβλιογραφία, είναι απαραίτητη η αναφορά στις ανώτερες και κατώτερες νοητικές ικανότητες, και ειδικότερα σε μοντέλα ταξινόμησής τους. Το επικρατέστερο μοντέλο ταξινόμησης προτάθηκε από τον Bloom (1956), ο οποίος κατέταξε τις νοητικές ικανότητες σε έξι ιεραρχικά επίπεδα (Σχήμα 9), ως εξής (Bloom, Krathwohl & Masia, 1984. Πετροπούλου, Κασσιμάτη & Ρετάλης, 2005):

- Επίπεδο 1 – Γνώση: Η απομνημόνευση πληροφοριών όπως ορολογίες, γεγονότα και στοιχεία, τάσεις και ακολουθίες, ταξινομήσεις και κατηγορίες, κριτήρια, μεθοδολογίες αφηρημένες έννοιες, γενικές αρχές και η μετέπειτα ανάκλησή τους από τη μνήμη.
- Επίπεδο 2 – Κατανόηση: Η περιγραφή, επεξήγηση, αναδιατύπωση, σύγκριση πληροφοριών που ανακαλούνται από τη μνήμη, χωρίς όμως απαραίτητα την εφαρμογή των πληροφοριών σε νέες καταστάσεις, ή την πλήρη αντίληψη των προεκτάσεων της γνώσης.
- Επίπεδο 3 - Εφαρμογή: Η χρήση γνωστών εννοιών, κανόνων, μεθόδων, αρχών, νόμων, θεωριών σε νέες, άγνωστες ως τότε καταστάσεις, προκειμένου να επιλυθούν προβλήματα και να διεκπεραιωθούν έργα.
- Επίπεδο 4 – Ανάλυση: Ο διαχωρισμός των γνώσεων στα συστατικά μέρη τους και η αναγνώριση των μεταξύ τους σχέσεων, αλληλεξαρτήσεων, συνεκτικών αρχών, κινήτρων, αιτίων, διαφορών, αντιθέσεων.
- Επίπεδο 5 – Σύνθεση: Η επίλυση αντιθέσεων και η κατασκευή νέων εννοιών, κατασκευών, σχημάτων, συνόλων, με χρήση συστατικών στοιχείων που προέρχονται από διαφορετικές πηγές.
- Επίπεδο 6 – Αποτίμηση: Η κριτική παρουσίαση και τεκμηρίωση απόψεων σχετικά με την αξία εννοιών, διαδικασιών, προϊόντων με βάση σύνολα θεσπισμένων κριτηρίων.

Οι Anderson & Krathwohl (2001) πρότειναν την αναθεώρηση της ταξινόμησης: μετέγραψαν τους τίτλους των Επιπέδων (συγκεκριμένα χρησιμοποίησαν ρήματα αντί ουσιαστικών), αντιμετάθεσαν τις λειτουργίες του Επιπέδου 5 και του Επιπέδου 6 και τοποθέτησαν στο Επίπεδο

6 τη λειτουργία της «Δημιουργίας», την οποία περιέγραψαν ως τη σύνδεση διαφορετικών στοιχείων προκειμένου να σχηματισθεί ένα συνεκτικό και λειτουργικό σύνολο και ως την αναδιοργάνωση διαφορετικών στοιχείων σε νέα δομή (Σχήμα 9).



**Σχήμα 9. Η κατά τον Bloom (1956) ταξινόμηση των νοητικών λειτουργιών και η αναθεώρησή της (Anderson & Krathwohl, 2001).**

Έχοντας ως σύστημα αναφοράς την ταξινόμηση του Bloom (1956), οι Douglas, Wilson & Ennis, (2002) και οι Simkin & Kuechler (2005), θεωρούν ότι οι λεγόμενες «κλειστού τύπου» ψηφιακές αξιολογήσεις, οι οποίες ως επί το πλείστον χρησιμοποιούν ερωτήσεις πολλαπλών επιλογών, είναι κατάλληλες για τα τρία κατώτερα νοητικά επίπεδα, δηλαδή αυτά της γνώσης, της κατανόησης και της εφαρμογής. Οι Buckles & Siegfried (2006) και η Brady (2005) θεωρούν ότι είναι μερικώς κατάλληλες και για το τέταρτο νοητικό επίπεδο, δηλαδή την ανάλυση. Οι Torres, Lopes, Babo & Azevedo (2011) υποστηρίζουν ότι υπάρχουν μορφές ψηφιακών δοκιμασιών «κλειστού τύπου», που είναι κατάλληλες για όλα τα νοητικά επίπεδα και παραθέτουν τον παρακάτω σχετικό Πίνακα 7. Σημειώνουν ωστόσο, ότι όσο υψηλότερο είναι το νοητικό επίπεδο, τόσο δυσκολότερη είναι η σχεδίαση δοκιμασίας που θα το αξιολογήσει.

*Πίνακας 7. Νοητικά επίπεδα και καταλληλότερες μορφές δοκιμασίας (Torres, Lopes, Babo & Azevedo, 2011)*

<b>Νοητικό Επίπεδο</b>	<b>Καταλληλότερη μορφή δοκιμασίας</b>
Γνώση, Κατανόηση (Επίπεδα 1, 2)	Πολλαπλή επιλογή, Σωστό / Λάθος, Αντιστοίχιση, Συμπλήρωση, Σύντομη απάντηση.
Εφαρμογή (Επίπεδο 3)	Πολλαπλή επιλογή, Σύντομη απάντηση, Επίλυση προβλήματος, Έκθεση, Υλοποίηση έργου.
Ανάλυση, Σύνθεση, Αποτίμηση (Επίπεδα 4, 5, 6)	Πολλαπλή επιλογή, Σύντομη απάντηση, Έκθεση.

Παρόμοια άποψη εκφράζουν και οι Azevedo et al. (2019), που υποστηρίζουν ότι οι ερωτήσεις πολλαπλής επιλογής είναι δυνατόν, αν προσαρμοστούν κατάλληλα, να αξιολογήσουν όλα τα νοητικά επίπεδα. Προσθέτουν όμως, ότι η κατασκευή ερωτήσεων για τα ανώτερα νοητικά επίπεδα είναι δύσκολη. Ίδια άποψη εκφράζει ο Rodriguez (2002), σύμφωνα με τον οποίον, οι «κλειστού τύπου» ψηφιακές δοκιμασίες μπορούν να αξιολογήσουν ικανότητες που αντιστοιχούν σε όλα τα νοητικά επίπεδα, αρκεί να σχεδιαστούν με αυτόν τον στόχο. Οι Morrison & Free (2001) αναφέρουν ότι οι ερωτήσεις που έχουν σχεδιαστεί, έτσι ώστε να απαιτούν σύνθετη σκέψη και να περιλαμβάνουν μεταξύ των επιλογών εκτός από την ορθή απάντηση και εναλλακτικές αληθοφανείς απαντήσεις, είναι ικανές να αξιολογήσουν σύνθετες νοητικές ικανότητες. Οι Farrell & Rushby (2016) υποστηρίζουν ότι οι ερωτήσεις πολλαπλής επιλογής μπορούν να αξιολογήσουν ικανότητες πολλών νοητικών επιπέδων: από την βασική ανάκληση γνώσεων, ως την ανάλυση και αποτίμηση πληροφοριών. Ωστόσο, αναγνωρίζουν ότι είναι δύσκολο να κατασκευαστούν ερωτήσεις πολλαπλής επιλογής, που θα αξιολογήσουν τη δημιουργικότητα των αξιολογούμενων, ή την ικανότητα τους να οργανώνουν τις σκέψεις τους. Σύμφωνα με την Επιτροπή JISC (JISC, Effective Practice with e-Assessment, 2007, σ. 28), η ελκυστικότητα της ψηφιακής αξιολόγησης έγκειται στην ικανότητά της να αποτυπώνει

χαρακτηριστικά της μάθησης, τα οποία έως πρόσφατα θεωρούνταν δύσκολο να αξιολογηθούν. Ειδικότερα, υπάρχουν μέθοδοι, όπως η «βαθμολόγηση βάση βεβαιότητας», μέσω των οποίων οι «κλειστού τύπου» αντικειμενικές ψηφιακές δοκιμασίες (που θεωρούνταν κατάλληλες για την αξιολόγηση μόνο των κατώτερων νοητικών επιπέδων) είναι δυνατόν να αξιολογήσουν ικανότητες και ανώτερων νοητικών επιπέδων.

Προτάσεις για αξιολόγηση ικανοτήτων ανώτερων νοητικών επιπέδων με τη χρήση «κλειστού τύπου» ψηφιακών δοκιμασιών, καταθέτουν η Donnelly (2014) και ο Hemming (2010). Συγκεκριμένα, η Donnelly (2014) προτείνει τη χρήση ερωτήσεων πολλαπλής επιλογής, οι οποίες σχετίζονται με μελέτες περίπτωσης (case based) και παρουσιάζει ερευνητικά πορίσματα σύμφωνα με τα οποία, η συγκεκριμένη μέθοδος διευκολύνει την αξιολόγηση ικανοτήτων που αντιστοιχούν σε ανώτερα νοητικά επίπεδα. Ο Hemming (2010) προτείνει τη χρήση ψηφιακών αξιολογήσεων με ερωτήσεις που στοχεύουν στα ανώτερα νοητικά επίπεδα και με απαντήσεις που αποτελούνται από ζεύγη ισχυρισμού – αιτιολόγησης (assertion – reason questions), τα οποία σύμφωνα με τον Williams (2006) είναι κατάλληλα για την «σε βάθος» μάθηση. Η Holmes (2015) επικαλείται τις έρευνες των Brady (2005), Draper (2009), Leung, Mok & Wong (2008) για να υποστηρίξει την άποψη ότι, οι καλά σχεδιασμένες ψηφιακές αξιολογήσεις, που χρησιμοποιούν ερωτήσεις πολλαπλής επιλογής, μπορούν να αξιολογήσουν ανώτερες νοητικές διεργασίες, όπως η κριτική σκέψη και η αναλυτική ικανότητα. Υποστηρίζει επίσης ότι, οι σπουδαστές που προετοιμάζονταν για αξιολόγηση με ερωτήσεις πολλαπλής επιλογής εστίαζαν στην κατανόηση του γνωστικού αντικείμενου, ενώ οι σπουδαστές που προετοιμάζονταν για αξιολόγηση με απαντήσεις ελεύθερου κειμένου εστίαζαν στην απομνημόνευση, προκειμένου να αναπαράγουν τις απαντήσεις.

Η Donnelly (2014) παραθέτει την άποψη των Master et al. (2001) σύμφωνα με τους οποίους, οι ερωτήσεις πολλαπλής επιλογής μπορούν επαρκώς να αξιολογήσουν ικανότητες που αντιστοιχούν σε όλα τα νοητικά επίπεδα, πλην αυτών της σύνθεσης και της αποτίμησης, για την αξιολόγηση των οποίων, είναι απαραίτητες απαντήσεις με τη μορφή κειμένου ελεύθερης ανάπτυξης. Παραθέτει όμως και την αντίθετη άποψη των Buckles και Siegfried (2006), οι οποίοι υποστηρίζουν ότι οι απαντήσεις με μορφή κειμένου ελεύθερης ανάπτυξης και οι ερωτήσεις πολλαπλής επιλογής αξιολογούν τα ίδια νοητικά επίπεδα.

Οι Conole & Warburton (2016) εκφράζουν επιφυλάξεις για την ικανότητα των αντικειμενικών ψηφιακών δοκιμασιών να αξιολογήσουν ανώτερα νοητικά επίπεδα, για τα οποία θεωρούν ότι είναι απαραίτητη σύνταξη κειμένων και εκθέσεων, όπου όμως υφίσταται δυσκολία ως προς την αυτόματη βαθμολόγησή τους. Με την ίδια άποψη συντάσσονται και οι McAllister & Guidice (2012), που υποστηρίζουν ότι οι ερωτήσεις πολλαπλής επιλογής συχνά δεν ενσωματώνουν, δεν ενθαρρύνουν και τελικά δεν αξιολογούν ανώτερες νοητικές διεργασίες και δεξιότητες. Επιπλέον, αναφέρουν ότι η πλειονότητα των ερωτήσεων, που περιλαμβάνονται στις τράπεζες ερωτήσεων, απαιτούν μόνο ανάκληση πληροφοριών και ότι ελάχιστες από τις ερωτήσεις σχετίζονται με τα τρία ανώτερα νοητικά επίπεδα της σύνθεσης, της εφαρμογής και της αποτίμησης. Αντιθέτως, οι εκθέσεις, ως μορφή αξιολόγησης, προσφέρουν στους σπουδαστές τη δυνατότητα να παραγάγουν και όχι να επιλέγουν τις απαντήσεις, οπότε μπορούν να επιδείξουν τις γνώσεις και τις δεξιότητές τους.

Οι Heyde & Siebrits (2019) συντάσσονται με την άποψη του Livingston (2009) που υποστηρίζει ότι οι ερωτήσεις πολλαπλής επιλογής δεν μπορούν να αξιολογήσουν όλες τις μορφές γνώσεων και δεξιοτήτων, δεδομένου ότι οι απαντήσεις προσφέρονται σε λίστα από επιλογές. Για την αξιολόγηση ανώτερων νοητικών επιπέδων προτείνουν, όπως και η Schultz (2011), τη χρήση ερωτήσεων, που απαιτούν από τον αξιολογούμενο να υπολογίσει, ή να κατασκευάσει την απάντηση, αντί να την επιλέξει από λίστα. Ομοίως, όπως αναφέρει η Jones (2014), οι ερωτήσεις πολλαπλής επιλογής παρότι μπορούν, αν σχεδιαστούν σωστά, να αξιολογήσουν ανώτερα νοητικά επίπεδα, εντούτοις τις περισσότερες φορές σχεδιάζονται για να αξιολογήσουν την ανάκληση γνώσεων, που σχετίζεται με τα κατώτερα νοητικά επίπεδα.

Οι Arriah & van Tonder (2018), υπογραμμίζουν ότι για την αξιολόγηση ανώτερων νοητικών επιπέδων και δεξιοτήτων, οι καταλληλότερες μορφές ψηφιακής αξιολόγησης είναι τα εργαλεία ανάπτυξης συνεργατικού περιεχομένου (wikis), τα ιστολόγια (blogs), τα συστήματα αυτοαξιολόγησης και ομότιμης αξιολόγησης, οι βασισμένες σε σενάρια ερωτήσεις, οι προσομοιώσεις, τα παιχνίδια επιτέλεσης ρόλων. Παρομοίως, η Jordan (2013) αναφέρει ότι οι ψηφιακοί φάκελοι επιτευγμάτων, οι προσομοιώσεις και τα διαδραστικά παιχνίδια υποστηρίζουν την αξιολόγηση ανώτερων νοητικών ικανοτήτων, που δεν μπορούν να αξιολογηθούν εύκολα με άλλο τρόπο και ο Crisp (2009) προτείνει την ενσωμάτωση

διαδραστικών τεχνικών στα ψηφιακά διαγωνίσματα, ώστε να είναι δυνατή η αξιολόγηση ανώτερων νοητικών ικανοτήτων.

Παράλληλα με το ερώτημα για το ποιες νοητικές λειτουργίες είναι δυνατόν να αξιολογηθούν με τις διάφορες μορφές της ψηφιακής αξιολόγησης, υφίσταται το αντίστοιχο ερώτημα για την αξιολόγηση των γνώσεων και των δεξιοτήτων, που είναι απαραίτητες τον 21<sup>ο</sup> αιώνα. Οι Kereluik, Mishra, Fahnoe & Terry (2013) κατηγοριοποίησαν τις απαραίτητες για τον 21<sup>ο</sup> αιώνα γνώσεις και δεξιότητες σε ένα μοντέλο που περιλαμβάνει τρεις τομείς γνώσεων: τη θεμελιώδη γνώση, τη μετα-γνώση, την ανθρωπιστική γνώση (Σχήμα 10).



**Σχήμα 10.** Το κατά Kereluik et al. (2013) πλαίσιο μάθησης για τον 21<sup>ο</sup> αιώνα

Οι Boitshwarelo et al. (2017) υποστηρίζουν ότι τα ψηφιακά διαγωνίσματα, δηλαδή οι ψηφιακές αξιολογικές δοκιμασίες που περιλαμβάνουν ερωτήσεις με προκαθορισμένες απαντήσεις, παρότι συνήθως είναι ανεπαρκώς σχεδιασμένες και χρησιμοποιούνται για την αξιολόγηση χαμηλού επιπέδου νοητικών ικανοτήτων, μπορούν να αξιοποιηθούν για την αξιολόγηση γνώσεων του 21ου αιώνα, κυρίως, αλλά όχι αποκλειστικώς, για τον τομέα της θεμελιώδους γνώσης. Αν μάλιστα στον σχεδιασμό τους επενδυθεί δημιουργική σκέψη, προσπάθεια και χρόνος, τότε είναι δυνατόν να αξιοποιηθούν και στους άλλους δύο τομείς, δηλαδή στον τομέα

της ανθρωπιστικής γνώσης και στον τομέα της μετα-γνώσης. Ειδικότερα, είναι δυνατόν να υποστηρίξουν «ισχυρές» μορφές αξιολόγησης, όπως επίλυση προβλημάτων, υλοποίηση αυθεντικών μαθησιακών έργων και μελέτες περίπτωσης (Scott, 2016). Η κατάλληλη επιλογή της μορφής των δοκιμασιών, του νοητικού επιπέδου των ερωτήσεων και της μαθησιακής προσέγγισης, που ενσωματώνεται στη δοκιμασία, βοηθούν στην ενεργοποίηση νοητικών λειτουργιών πέραν της απλής ανάκλησης γνώσης. Δηλαδή, μπορούν να αξιολογήσουν την ικανότητα των σπουδαστών, αφενός να κατακτούν γνώσεις και δεξιότητες από όλους τους τομείς και αφετέρου να τις αξιοποιούν σε έναν ρευστό και διαρκώς μεταβαλλόμενο κόσμο. Οι Binkley et al. (2012) υποστηρίζουν ότι η αξιολόγηση γνώσεων και δεξιοτήτων του 21<sup>ου</sup> αιώνα, είναι εφικτή, αν τα συστήματα αξιολόγησης υιοθετήσουν τη στρατηγική του μετασχηματισμού, όπου η αξιοποίηση της τεχνολογίας έχει ως στόχο, όχι απλώς την αύξηση της αποδοτικότητας, αλλά οδηγεί σε καινοτόμες μαθησιακές και αξιολογικές πρακτικές, που περιλαμβάνουν επίλυση σύνθετων προβλημάτων, επικοινωνία, ομαδική δουλειά και δημιουργικότητα.

#### 7.6. Η ακαδημαϊκή ακεραιότητα και η ασφάλεια υποδομών και υπηρεσιών

Ιδιαίτερη αναφορά είναι απαραίτητη στις προκλήσεις που σχετίζονται α) με την προστασία της ακαδημαϊκής ακεραιότητας και β) με την ασφάλεια των ψηφιακών υποδομών και υπηρεσιών. Καθώς η χρήση ψηφιακών αξιολογήσεων από εκπαιδευτικούς οργανισμούς διευρύνεται, ταυτοχρόνως αυξάνεται η ανάγκη για την προστασία της ακαδημαϊκής ακεραιότητας και της ασφάλειας των ψηφιακών υποδομών και υπηρεσιών, προκειμένου να μην τίθενται υπό αμφισβήτηση οι τίτλοι σπουδών και οι εκπαιδευτικές πιστωτικές μονάδες, που απονέμουν τα εκπαιδευτικά ιδρύματα, τόσο στα πλαίσια παραδοσιακών προγραμμάτων σπουδών (Harmon, Lambrios & Buffolino, 2010) όσο και στα πλαίσια Μαζικών Ανοικτών Διαδικτυακών Μαθημάτων (MOOC / Massive Open Online Courses), τα οποία ολοένα και συχνότερα απονέμουν τίτλους σπουδών και όχι απλώς βεβαιώσεις παρακολούθησης (James, 2016). Οι συγκεκριμένες προκλήσεις θεωρούνται ιδιαίτερες κρίσιμες (Brink & Lautenbach, 2011) και αποτελούν βασικό παράγοντα ανασχεσης για την ευρεία διάδοση των ψηφιακών αξιολογήσεων (Mellar, Peytcheva-Forsyth, Kocdar, Karadeniz & Yonkova, 2018).

Η ακαδημαϊκή ακεραιότητα απειλείται από ενέργειες με τις οποίες, κάποιος σπουδαστής παραβιάζει τους κανόνες της αξιολογικής διαδικασίας και αποκτά αθέμιτο πλεονέκτημα έναντι

των υπολοίπων. Η ασφάλεια των ψηφιακών υποδομών και υπηρεσιών («e-security») απειλείται από σκόπιμες, κακόβουλες ενέργειες ή τυχαία συμβάντα που παρεμποδίζουν την ομαλή λειτουργία των υποδομών και των υπηρεσιών της ψηφιακής αξιολόγησης, ή που θέτουν σε κίνδυνο την ακεραιότητα και την εμπιστευτικότητα των πληροφοριών που σχετίζονται με τις διαδικασίες αξιολόγησης.

Οι συγκεκριμένες προκλήσεις δεν αποτελούν καινοφανή ζητήματα, τα οποία απειλούν μόνο τη ψηφιακή αξιολόγηση (JISC, 2007). Αντίστοιχες προκλήσεις, ως προς την παραβίαση της ακαδημαϊκής ακεραιότητας, ή ως προς την ασφάλεια των υποδομών και των υπηρεσιών (π.χ. ασφάλεια κτηρίων, διασφάλιση απορρήτου δεδομένων), υφίστανται και στις παραδοσιακές αξιολογήσεις. Απειλούν την ακεραιότητα και την ασφάλεια και των παραδοσιακών αξιολογήσεων, αλλά η χρήση της τεχνολογίας εγείρει εντονότερες ανησυχίες για την αύξηση των σχετικών κινδύνων, αν και αναγνωρίζεται το γεγονός ότι οι τεχνολογίες παράλληλα προσφέρουν τρόπους και εργαλεία για τον περιορισμό τους (Mellar et al., 2018).

Η ακαδημαϊκή κοινότητα αναγνωρίζοντας την σπουδαιότητα του ζητήματος αναζητεί στρατηγικές για την αντιμετώπιση του. Ωστόσο, όπως παρατηρούν οι Oosterhof, Conrad & Ely (2008, σ. 138), δεν θα πρέπει να επιδιώκεται υψηλότερος βαθμός θωράκισης των ψηφιακών αξιολογήσεων σε σχέση με τις παραδοσιακές αξιολογήσεις. Διότι συχνά, το οικονομικό και διαχειριστικό κόστος, που καταβάλλεται για την επίτευξη υψηλότερου βαθμού θωράκισης, μπορεί να αποβεί δυσανάλογο και να περιορίσει ή και να ακυρώσει τα οφέλη των ψηφιακών αξιολογήσεων και γενικότερα των ψηφιακών περιβαλλόντων μάθησης, όπως για παράδειγμα να περιορίσει την ευελιξία στον χρόνο, στον τόπο και στον τρόπο σύνδεσης και συμμετοχής.

#### 7.6.1. Προστασία ακαδημαϊκής ακεραιότητας

##### 7.6.1.1 Συχνότητα απειλών

Ως προς την προστασία της ακαδημαϊκής ακεραιότητας, θα πρέπει να σημειωθεί ότι τα πορίσματα που προκύπτουν από την επισκόπηση της βιβλιογραφίας είναι διφορούμενα ως προς το εάν τα συμβάντα εξαπάτησης είναι συχνότερα στην ψηφιακή αξιολόγηση έναντι των παραδοσιακών αξιολογήσεων (James, 2016). Ειδικότερα, από την έρευνα των Watson & Sottile (2010) δεν προέκυψαν αποδείξεις ότι η εξαπάτηση είναι συχνότερη στην ψηφιακή αξιολόγηση. Σύμφωνα με τον Arnold (2016), η άποψη περί αύξησης στη συχνότητα των συμβάντων



εξαπάτησης στην ψηφιακή αξιολόγηση είναι διαδεδομένη, γεγονός που αποδίδεται στην αναποτελεσματική, εξαιτίας της απόστασης, επιτήρηση. Ωστόσο, ο Arnold (2016) σημειώνει ότι οι αυξημένες ευκαιρίες για εξαπάτηση, που υφίστανται στις μη επιτηρούμενες ψηφιακές αξιολογήσεις, δεν συνεπάγονται αυτομάτως ότι οι σπουδαστές επιλέγουν πράγματι να εκμεταλλευτούν αυτές τις ευκαιρίες. Οι Mellar et al. (2018) καταγράφουν αντικρουόμενα ερευνητικά πορίσματα, ως προς το εάν η εξαπάτηση είναι συχνότερη, ή όχι, στις επιτηρούμενες αξιολογήσεις και αποδίδουν τις αποκλίσεις σε παράγοντες, όπως ο σχεδιασμός της αξιολόγησης και όχι η ύπαρξη, ή η μη ύπαρξη, επιτήρησης.

Ο Lang (2013) υποστηρίζει ότι, παρότι τα περιβάλλοντα ψηφιακής μάθησης προσφέρουν περισσότερες ευκαιρίες για εξαπάτηση, η συχνότητα των συμβάντων είναι περίπου ίδια με τα παραδοσιακά περιβάλλοντα μάθησης. Εικάζει, ότι αυτό συμβαίνει, διότι τα περιβάλλοντα ψηφιακής μάθησης χρησιμοποιούνται περισσότερο από σπουδαστές μεγαλύτερης ηλικίας, οι οποίοι είναι λιγότερο επιρρεπείς σε πρακτικές εξαπάτησης. Ο Lang (2013) συμπληρώνει ότι στο μέλλον, καθώς η χρήση περιβαλλόντων ψηφιακής μάθησης θα γίνεται συχνότερη και για τους σπουδαστές μικρότερης ηλικίας, τα συμβάντα εξαπάτησης πιθανόν θα αυξηθούν, αλλά προς το παρόν δεν υπάρχουν σαφείς αποδείξεις ότι η εξαπάτηση στα περιβάλλοντα ψηφιακής μάθησης είναι συχνότερη. Παρόμοια είναι και η άποψη των Stuber-McEwen, Wiseley & Hoggatt (2009), που υποστηρίζουν ότι η εξαπάτηση στην ψηφιακή αξιολόγηση ίσως είναι λιγότερο συχνή σε σχέση από ό,τι πιστεύεται ότι είναι. Ωστόσο, θεωρούν πιθανή την αύξησή της, όσο θα διευρύνεται η αξιοποίησή της.

Ο Arnold (2016) αναφέρει ότι η συχνότητα περιπτώσεων εξαπάτησης είναι αντιστρόφως ανάλογη με την ωριμότητα των σπουδαστών. Ειδικότερα οι μεταπτυχιακοί φοιτητές, λόγω της συνειδητής επιλογής τους για παρακολούθηση του μεταπτυχιακού προγράμματος σπουδών και της βαθύτερης προσήλωσής τους στην απόκτηση γνώσεων, είναι λιγότεροι επιρρεπείς σε πρακτικές εξαπάτησης.

Οι Gikandi et al. (2011) αναφέρουν ότι τα φαινόμενα παραβίασης της ακαδημαϊκής ακεραιότητας εντείνονται στις υψηλής κρισιμότητας τελικές ψηφιακές αξιολογήσεις. Αυξάνονται επίσης στις περιπτώσεις, όπου οι σπουδαστές τηρούν διεκπεραιωτική στάση έναντι της μάθησης και επιδιώκουν απλώς την απόκτηση ενός τίτλου σπουδών. Αντιστρόφως, είναι

σπανιότερα στις περιπτώσεις, όπου οι σπουδαστές αναλαμβάνουν την ευθύνη της μάθησής τους και επιδιώκουν να αναπτύξουν γνώσεις και δεξιότητες. Γενικά, οι σχετικές ανησυχίες για την ακαδημαϊκή ακεραιότητα και τα ανάλογα αντίμετρα αφορούν κυρίως τις τελικές και όχι τις διαμορφωτικές ψηφιακές αξιολογήσεις. Άλλωστε, οι σπουδαστές που εξαπατούν στις διαμορφωτικές αξιολογήσεις δεν αποκτούν κάποιο μακροπρόθεσμο πλεονέκτημα, διότι δεν ωφελούνται από τις δυνατότητες που η διαμορφωτική αξιολόγηση προσφέρει για τη βελτίωσή τους. Αντιθέτως βλάπτονται, αν εξαπατούν κατά τη διαμορφωτική αξιολόγηση, με αποτέλεσμα οι επιδόσεις τους κατά την τελική αξιολόγηση να είναι χαμηλότερες σε σύγκριση με όσους δεν εξαπατούν (Arnold, 2016).

Ως προς τις ανησυχίες των διδασκόντων για τα φαινόμενα παραβίασης της ακαδημαϊκής ακεραιότητας, ο Arnold (2016) υποστηρίζει ότι ο βαθμός ανησυχίας είναι αντιστρόφως ανάλογος με τον βαθμό εμπειρίας στη ψηφιακή μάθηση. Επιπλέον αναφέρει, ότι, όσο πιο έμπειροι σε περιβάλλοντα ψηφιακής μάθησης είναι οι διδάσκοντες, τόσο λιγότερες είναι οι ανησυχίες που εκφράζουν για το ζήτημα, διότι εκτιμούν ότι οι παραβιάσεις δεν είναι περισσότερο πιθανές από ό,τι στις παραδοσιακές αξιολογήσεις.

#### *7.6.1.2 Αντίμετρα*

Οι απόπειρες εξαπάτησης («e-cheating»), λαμβάνουν τη μορφή α) της πλαστοπροσωπίας, β) της λήψης βοήθειας από τρίτο πρόσωπο ή από μη επιτρεπόμενη πηγή πληροφοριών και γ) της λογοκλοπής.

Ως προς την αντιμετώπιση της πλαστοπροσωπίας, αξιοποιούνται καταρχήν οι βασικές μέθοδοι ψηφιακής ταυτοποίησης, όπως για παράδειγμα η εισαγωγή ονόματος χρήστη και κωδικού πρόσβασης, ή η πρόσβαση μέσω πιστοποιημένης δικτυακής διεύθυνσης. Οι συγκεκριμένες μέθοδοι δεν μπορεί να θεωρηθούν επαρκείς, οπότε είναι αναγκαίες πιο εξελιγμένες μέθοδοι, οι οποίες ελέγχουν βιομετρικά χαρακτηριστικά του αξιολογούμενου (π.χ. σαρωτής ίριδας, αναγνώριση χαρακτηριστικών προσώπου, αναγνώριση φωνής, ρυθμός πληκτρολόγησης (Crisp, 2011). Παραδείγματα σχετικών συστημάτων σχεδίασαν οι Clarke, Dowland & Furnell (2013), οι Ahlawat, Pareek & Singh (2014), οι Li, Chang, Yuan & Hauptmann, (2015). Τα εν λόγω συστήματα επιβεβαιώνουν διαρκώς την ταυτότητα των αξιολογούμενων καθ' όλη τη διάρκεια της αξιολόγησης βασιζόμενα σε βιομετρικά χαρακτηριστικά, χωρίς να

απαιτούν την εγκατάσταση πολύπλοκων ψηφιακών συσκευών στον υπολογιστή του εξεταζόμενου. Ωστόσο, δεν είναι σπάνιες οι περιπτώσεις, όπου οι εκπαιδευτικοί οργανισμοί, κυρίως στις εξετάσεις υψηλής κρισιμότητας, απαιτούν από τους σπουδαστές να προσέλθουν αυτοπροσώπως σε εξεταστικά κέντρα. Εκεί, οι σπουδαστές ταυτοποιούνται και επιτηρούνται με φυσικό τρόπο, αλλά εξετάζονται με τη χρήση των ΤΠΕ. Προφανώς, η συγκεκριμένη πρακτική, όχι μόνο προκαλεί υψηλό οικονομικό και διαχειριστικό κόστος για τους εκπαιδευτικούς οργανισμούς, αλλά επιπλέον στερεί από τους σπουδαστές την ευελιξία ως προς τον τόπο και τον χρόνο της εξέτασης, η οποία αποτελεί από τα βασικότερα πλεονεκτήματα της ψηφιακής αξιολόγησης. Γενικά, οι δυσκολίες στο ζήτημα της επιτήρησης, η οποία αποτελεί απαραίτητη προϋπόθεση στις περισσότερες παραδοσιακές αξιολογήσεις, προβληματίζει έντονα τους εκπαιδευτικούς οργανισμούς και τους διδάσκοντες και τους αποτρέπει να υιοθετήσουν διαδικασίες ψηφιακής αξιολόγησης, ιδιαιτέρως στις περιπτώσεις τελικών και υψηλής σπουδαιότητας εξετάσεων.

Ως προς τη λήψη βοήθειας από τρίτο πρόσωπο, ή τη χρήση μη επιτρεπόμενης πηγής πληροφοριών, καταγράφονται διάφορες πρακτικές εξαπάτησης. Η συνηθέστερη, απλούστερη και ευκολότερη (ελλείψει αποτελεσματικής επιτήρησης) μορφή εξαπάτησης, είναι η αναζήτηση των ορθών απαντήσεων στο διαδίκτυο, πρακτική την οποία όπως αναφέρουν οι Boitshwarelo et al. (2017), αρκετοί σπουδαστές θεωρούν αποδεκτή. Σοβαρότερα κρούσματα εξαπάτησης είναι αυτά όπου οι σπουδαστές λαμβάνουν βοήθεια από τρίτο πρόσωπο ή δημιουργούν ομάδες για την από κοινού επίλυση του διαγωνίσματος. Παρόμοια πρακτική εξαπάτησης είναι η σύνδεση διαφορετικών προσώπων με τον λογαριασμό πρόσβασης του ίδιου σπουδαστή, προκειμένου να επιλύσουν από κοινού το διαγώνισμα (Boitshwarelo et al., 2017). Η αλληλοβοήθεια μεταξύ σπουδαστών αποτελεί συχνό φαινόμενο στα ψηφιακά διαγωνίσματα πολλαπλής επιλογής, στα οποία σύμφωνα με τους Watson & Sottile (2010) καταγράφονται τα περισσότερα συμβάντα εξαπάτησης. Γι' αυτές τις περιπτώσεις εφαρμόζονται αντίμετρα, όπως η τυχαία επιλογή ερωτήσεων από μεγάλες τράπεζες ερωτήσεων, η αλλαγή στη διάταξη των απαντήσεων για κάθε σπουδαστή, οι παραλλαγμένες ερωτήσεις, ως προς τις αριθμητικές τιμές τους (de Sande, 2015., Heyde & Siebrits, 2019). Η υλοποίηση τέτοιων αντιμέτρων έχει απλοποιηθεί μέσω μηχανισμών

που είναι ενσωματωμένοι στα Συστήματα Διαχείρισης Μάθησης (ΣΔΜ), τα οποία υποστηρίζουν την κατασκευή και τη διανομή ψηφιακών διαγωνισμάτων.

Ως προς την λογοκλοπή, η οποία αποτελεί πρόκληση, όχι μόνο για την ψηφιακή αξιολόγηση, αλλά για κάθε μορφή αξιολόγησης, οι Davies & Howard (2016) αναφέρουν ότι δεν υφίστανται εμπειρικές έρευνες, οι οποίες τεκμηριώνουν την ευρέως διαδεδομένη άποψη ότι το διαδίκτυο διευκολύνει και ενισχύει την λογοκλοπή. Απεναντίας, η ψηφιακή αξιολόγηση προσφέρει αποτελεσματικά αντίμετρα, όπως η αξιοποίηση εξειδικευμένων λογισμικών που εντοπίζουν ομοιότητες μεταξύ κειμένων στις εργασίες των σπουδαστών και κειμένων που βρίσκονται στο διαδίκτυο (π.χ. κειμένων που περιλαμβάνονται σε ιστοσελίδες, βιβλία, περιοδικά, ερευνητικές εργασίες, διατριβές κ.α.). Τα λογισμικά βασίζονται σε μηχανές αναζήτησης, που σαρώνουν το διαδίκτυο και σε αλγόριθμους που αναλύουν και συγκρίνουν πολύ μεγάλο πλήθος κειμένων. Εντοπίζουν ενδείξεις λογοκλοπής με ακρίβεια και ταχύτητα, που είναι αδύνατον να επιτευχθούν από φυσικό πρόσωπο. Επιπλέον, επιτρέπουν παραμετροποίηση ως προς την ευαισθησία της σύγκρισης (π.χ. επιτρέπουν τον καθορισμό ορίου στο πλήθος των όμοιων λέξεων, πάνω από το οποίο θα εμφανίζεται ενημέρωση για πιθανή λογοκλοπή). Ωστόσο, πρέπει να σημειωθεί ότι τα συγκεκριμένα λογισμικά, παρότι εξελίσσονται και βελτιώνονται συνεχώς, δεν στερούνται προβλημάτων. Για παράδειγμα, οι αναφορές που παράγουν δεν είναι πάντοτε εύκολα κατανοητές. Τμήματα κειμένων, παρότι συνοδεύονται από τις προβλεπόμενες βιβλιογραφικές αναφορές, υποδεικνύονται λανθασμένα ως προϊόντα λογοκλοπής. Τμήματα κειμένων που αποτελούν προϊόντα λογοκλοπής δεν εντοπίζονται. Διαφορετικά λογισμικά παράγουν διαφορετικές αναφορές για το ίδιο κείμενο (Mellar et al., 2018). Ο Lang (2013) εξάρει τη χρησιμότητα τέτοιων λογισμικών, αλλά σημειώνει ότι οι αναφορές που παράγουν, απλώς υποδεικνύουν ομοιότητες και δεν αποδεικνύουν συμβάντα λογοκλοπής. Η οριστική απόφαση, για το αν η υποδεικνυόμενη ομοιότητα συνιστά λογοκλοπή, ανήκει πάντοτε στην κρίση του διδάσκοντα. Υποστηρίζει επίσης, ότι τα λογισμικά μπορούν να έχουν και άλλη χρησιμότητα. Συγκεκριμένα, προτείνει τη μελέτη και τον σχολιασμό των αναφορών, που τα λογισμικά παράγουν, ως μέσο για να κατανοήσουν οι σπουδαστές τις δεοντολογικά ορθές πρακτικές για τη αξιοποίηση βιβλιογραφικών πηγών.

Σύμφωνα με τους Amigud, Arnedo-Moreno, Daradoumis & Guerrero-Roldan (2017) η διασφάλιση της ακαδημαϊκής ακεραιότητας συνιστά διττή πρόκληση. Αφενός, πρέπει να πιστοποιηθεί η ταυτότητα του σπουδαστή που συμμετέχει στη διαδικασία αξιολόγησης και αφετέρου να ελεγχθεί ότι τα έργα, που ο σπουδαστής υποβάλλει προς αξιολόγηση, δημιουργήθηκαν από αυτόν. Η αντιμετώπιση και των δύο προκλήσεων απαιτεί μεθόδους, που γενικά καταναλώνουν σημαντικούς πόρους και επιπλέον είναι παρεμβατικές. Ως αποτέλεσμα, η ακαδημαϊκή ακεραιότητα δεν διασφαλίζεται με ομοιόμορφο τρόπο σε όλες τις αξιολογικές δραστηριότητες και συχνά προκύπτουν κενά ασφάλειας. Οι Amigud et al. (2017) κατατάσσουν τις μεθόδους διασφάλισης της ακαδημαϊκής ακεραιότητας ως εξής: α) σε μεθόδους που πιστοποιούν την ταυτότητα των σπουδαστών, β) σε μεθόδους που επικυρώνουν την πνευματική πατρότητα μιας εργασίας και γ) σε μεθόδους που παρακολουθούν και ελέγχουν το περιβάλλον της μάθησης. Οι μέθοδοι, με τα πλεονεκτήματα και τα μειονεκτήματά τους, παρουσιάζονται στον παρακάτω Πίνακα 8.

**Πίνακας 8. Μέθοδοι διασφάλισης ακαδημαϊκής ακεραιότητας (Amigud et al. 2017)**

Τύπος	Μέθοδος	Πλεονεκτήματα	Μειονεκτήματα	Αναφορά
Πιστοποίηση Ταυτότητας	Βιομετρικός έλεγχος	Υψηλό επίπεδο ακρίβειας. Δυνατότητα αυτοματοποίησης	Απαιτεί ειδικό εξοπλισμό.	Arampa, Wills & Argles (2010)
	Ερωτήσεις πιστοποίησης	Εύκολη υλοποίηση. Δυνατότητα αυτοματοποίησης.	Ενοχλητικό, αν επαναλαμβάνεται συχνά.	Bailie & Jortberg (2009)
Επικύρωση πνευματικής πατρότητας	Εντοπισμός Λογοκλοπής	Δυνατότητα αυτοματοποίησης. Εύκολη υλοποίηση.	Δεν επικυρώνει την πνευματική πατρότητα, αλλά ελέγχει ισχυρισμούς περί πνευματικής πατρότητας.	Fiedler & Kaner (2010)
	Επικύρωση από διδάσκοντα	Συνεχής αξιολόγηση. Ενισχύει την εμπιστοσύνη και την ακεραιότητα.	Απαιτεί πολλούς πόρους. Κλιμακώνεται δύσκολα.	Barnes & Paris (2013)

Παρακολούθηση και έλεγχος	Επιτήρηση	Κατάλληλο για κάθε μορφή αξιολόγησης. Περιορίζει ανάγκες μετακίνησης αν υλοποιείται από απόσταση. Δυνατότητα αυτοματοποίησης.	Απαιτεί πολλούς πόρους. Κλιμακώνεται δύσκολα.	Li, Chang, Yuan & Hauptmann, (2015).
	Παρακολούθηση δραστηριότητας	Εύκολη υλοποίηση. Δυνατότητα αυτοματοποίησης.	Δεν επικυρώνει την πνευματική πατρότητα, αλλά ελέγχει ισχυρισμούς περί πνευματικής πατρότητας.	Gao (2012)

Οι Harmon, Lambrinos & Buffolino (2010), συνοψίζοντας πορίσματα της βιβλιογραφίας, αναφέρουν, ότι υπάρχουν δύο αντίθετες απόψεις ως προς ενδεδειγμένα αντίμετρα για την αντιμετώπιση της εξαπάτησης. Σύμφωνα με την πρώτη άποψη, κατά την οποία οι απειλές είναι το ίδιο συχνές, είτε στη ψηφιακή, είτε στην παραδοσιακή αξιολόγηση, τα αντίμετρα που μπορούν να εφαρμοσθούν είναι: κατάλληλος σχεδιασμός της αξιολόγησης, ώστε να είναι ελεύθερη η χρήση πηγών (ό,τι αποκαλείται «αξιολόγηση με ανοικτά βιβλία»), περιορισμός του χρόνου, που ο σπουδαστής έχει στη διάθεσή του για την ολοκλήρωση της διαδικασίας, περιορισμός των επαναλήψεων, που ο σπουδαστής δικαιούται να δει τις ερωτήσεις, περιορισμός του χρονικού διαστήματος, που η αξιολόγηση είναι διαθέσιμη, περιορισμός της βαθμολογικής βαρύτητας της ψηφιακής αξιολόγησης, τυχαία επιλογή υποσυνόλου ερωτήσεων από τράπεζα ερωτήσεων, ψηφιακή επιτήρηση. Επιπλέον των τεχνικών αντίμετρων, προτείνονται διδακτικές παρεμβάσεις, όπως: η ενίσχυση των κινήτρων των σπουδαστών, ώστε να επιδιώκουν τη μάθηση και όχι την βαθμοθηρία, η ανάδειξη του γεγονότος ότι η εξαπάτηση δεν προσφέρει μακροπρόθεσμο όφελος, η δέσμευση έναντι ηθικών κανόνων. Σύμφωνα με τη δεύτερη άποψη, σύμφωνα με την οποία οι απειλές είναι συχνότερες στην ψηφιακή αξιολόγηση, το βασικό αντίμετρο είναι η επιτήρηση, δηλαδή το μέτρο που χρησιμοποιείται και στις παραδοσιακές αξιολογήσεις.

Οι Oosterhof, Conrad & Ely (2008, σ. 138) υποστηρίζουν ότι η αποτελεσματικότερη στρατηγική για την αποτροπή παραβιάσεων της ακαδημαϊκής ακεραιότητας είναι ο περιορισμός στη χρήση συγκεκριμένων τύπων αξιολογήσεων, όπως οι ερωτήσεις πολλαπλής επιλογής και η σύνταξη εκθέσεων. Θεωρούν ότι καλύτερα προστατευμένες, έναντι παραβιάσεων της ακεραιότητας, είναι οι αξιολογήσεις απόδοσης («performance assessments»), όπου ο σπουδαστής διεκπεραιώνει έργο, επιτελεί ρόλο, επιλύει πρόβλημα, επιδεικνύει και εφαρμόζει δεξιότητες. Ωστόσο, οι Oosterhof, Conrad & Ely (2008, σ. 147) δεν παραλείπουν να αναφέρουν, ότι η χρήση των αξιολογήσεων απόδοσης ενέχει μειονεκτήματα, όπως μειωμένη αποδοτικότητα, υποκειμενική βαθμολόγηση, δυσκολία γενίκευσης. Ο Lang (2013) υποστηρίζει ότι τα τεχνικά αντίμετρα δεν μπορούν να εξασφαλίσουν πλήρως την ακαδημαϊκή ακεραιότητα. Κρούσματα εξαπάτησης υπήρχαν, υπάρχουν και θα υπάρχουν, τόσο στα παραδοσιακά, όσο και στα ψηφιακά περιβάλλοντα μάθησης. Η αποτελεσματικότερη στρατηγική για την αντιμετώπιση της εξαπάτησης πρέπει, σύμφωνα με τον Lang (2003), να επιδιώκει τον ανασχεδιασμό της μαθησιακής διαδικασίας, ώστε να προσφέρει δυνατότητες ουσιαστικής γνώσης και να ενθαρρύνει την ενεργή συμμετοχή των σπουδαστών. Ο Talbert (2020) αντιτίθεται στην αξιοποίηση τεχνικών αντίμετρων, όπως τα συστήματα δικτυακής επιτήρησης, τα οποία θεωρεί, αφενός αναποτελεσματικά και αφετέρου επικίνδυνα για την ιδιωτικότητα. Προτείνει την αντικατάσταση των παραδοσιακών βαθμολογικών πρακτικών, που θεωρεί ότι ευνοούν τη βαθμοθηρία, με εναλλακτικές μεθόδους αξιολόγησης, όπως η βαθμολόγηση βάση κριτηρίων («specifications grading») (Nilson, 2015).

#### 7.6.2. Ασφάλεια υποδομών και υπηρεσιών

Το ζήτημα της ασφάλειας των ψηφιακών υποδομών και υπηρεσιών ανάγεται στο γενικότερο ζήτημα της ασφάλειας των πληροφοριακών και δικτυακών συστημάτων, τα οποία παρέχουν τις υπολογιστικές και δικτυακές υπηρεσίες για τη ψηφιακή αξιολόγηση. Στη σχετική βιβλιογραφία (Kiennert et al., 2017. Miguel, Xhafa & Prieto, 2015. Marais, Argles & von Solms, 2006) καταγράφονται δύο κατηγορίες σχετικών απειλών. Η πρώτη κατηγορία απειλών σχετίζεται με την εύρυθμη διεξαγωγή της αξιολόγησης, η οποία είναι δυνατόν να διαταραχθεί, είτε από τυχαία συμβάντα, είτε από σκόπιμες και κακόβουλες ενέργειες. Τεχνικά προβλήματα, όπως για παράδειγμα, αστοχία του υλικού και του λογισμικού, διακοπές ηλεκτροδότησης, ή κακόβουλες

ενέργειες, όπως για παράδειγμα δικτυακές επιθέσεις άρνησης υπηρεσιών (denial of service), είναι δυνατόν να θέσουν εκτός λειτουργίας τους διακομιστές, ή τους δικτυακούς δρομολογητές και να προκαλέσουν παρακώλυση, ή διακοπή της αξιολόγησης. Οι δυσμενείς συνέπειες είναι σημαντικές, όταν οι αξιολογήσεις διεξάγονται εντός συγκεκριμένου και αμετάθετου χρονικού διαστήματος και δεν μπορούν να επαναληφθούν, ή να μετατεθούν χρονικά. Αξίζει να σημειωθεί ότι έκτακτα συμβάντα (όπως για παράδειγμα φυσικές καταστροφές, έντονα καιρικά φαινόμενα, διακοπές ηλεκτροδότησης, επεισόδια βίας) είναι δυνατόν να πλήξουν την εύρυθμη διεξαγωγή και των παραδοσιακών αξιολογήσεων, αλλά συνήθως τέτοια συμβάντα είναι λιγότερο συχνά, από τα τεχνικά προβλήματα, όπως για παράδειγμα η βλάβη ενός διακομιστή. Η δεύτερη κατηγορία απειλών σχετίζεται με μη εξουσιοδοτημένη πρόσβαση στα συστήματα με σκοπό την υποκλοπή ή την αλλοίωση δεδομένων της αξιολόγησης. Σχετικοί κίνδυνοι παραβίασης του απορρήτου ευαίσθητων πληροφοριών, υποκλοπής και αλλοίωσης δεδομένων υφίστανται και στις παραδοσιακές αξιολογήσεις, αλλά είναι δυσκολότερα υλοποιήσιμοι και ευκολότερα ανιχνεύσιμοι.

Η λήψη των αναγκαίων τεχνικών και διαδικαστικών αντίμετρων έναντι των απειλών είναι δυνατόν να μειώσει την πιθανότητα τόσο των έκτακτων συμβάντων, όσο και των σκόπιμων και κακόβουλων ενεργειών και να αμβλύνει τις συνέπειές τους. Για να διασφαλίσουν την εύρυθμη λειτουργία των ψηφιακών υποδομών και υπηρεσιών και την προστασία των δεδομένων της αξιολόγησης, οι εκπαιδευτικοί οργανισμοί οφείλουν να επενδύσουν στη θωράκιση των υποδομών, στην εκπαίδευση των διαχειριστών και των χρηστών. Γενικά, οφείλουν να σχεδιάσουν, να υλοποιήσουν και να επικαιροποιούν διαρκώς ένα γενικό και ολοκληρωμένο σχέδιο ασφάλειας των υπολογιστικών και δικτυακών συστημάτων και των υπηρεσιών που υποστηρίζουν την ψηφιακή αξιολόγηση.



---

## *ΕΡΕΥΝΗΤΙΚΟ ΜΕΡΟΣ*

---

## 8. ΠΡΟΒΛΗΜΑΤΙΚΗ

### 8.1. Το ερευνητικό πεδίο

Η χρήση των Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ) από τα εκπαιδευτικά ιδρύματα έχει γενικευτεί. Τα ΣΔΜ χρησιμοποιούνται σε όλα τα εκπαιδευτικά περιβάλλοντα (διαδικτυακά, μεικτά, συμβατικά) για τη διανομή εκπαιδευτικού υλικού, για τη συμμετοχή σε εκπαιδευτικές δραστηριότητες, για την επικοινωνία, για την αξιολόγηση. Έμμεσο αποτέλεσμα της λειτουργίας των ΣΔΜ, είναι η καταγραφή και η αποθήκευση δεδομένων που σχετίζονται με τη χρήση των εκπαιδευτικών πόρων, με τη συμμετοχή σε εκπαιδευτικές δραστηριότητες και γενικά με κάθε είδους αλληλεπίδραση των χρηστών με το σύστημα ή των χρηστών μεταξύ τους, μέσω του συστήματος.

Ποτέ άλλοτε τα εκπαιδευτικά ιδρύματα δεν είχαν στη διάθεσή τους τόσο πολλά και τέτοιου είδους δεδομένα χρήσης, αυτομάτως καταγεγραμμένα, ψηφιακώς μορφοποιημένα, αμέσως διαθέσιμα. Τα δεδομένα αποτυπώνουν μοτίβα (patterns) χρήσης, όπως για παράδειγμα μοτίβα πρόσβασης σε εκπαιδευτικούς πόρους, συμμετοχής σε δραστηριότητες, επικοινωνίας με τους σπουδαστές και τους διδάσκοντες. Εντός τέτοιων μοτίβων αναζητούνται συσχετίσεις μεταξύ του τρόπου χρήσης των ΣΔΜ και της επίδοσης των φοιτητών στα μαθήματα που παρακολουθούν (Cui, Chen, Shiri & Fan, 2019, σσ. 213,214).

Παράλληλα με τη διεύρυνση της χρήσης των ΣΔΜ, διευρύνεται η χρήση τεχνικών Μηχανικής Μάθησης. Η ύπαρξη εύχρηστων και προσιτών λογισμικών Μηχανικής Μάθησης, η αύξηση στην ισχύ των υπολογιστικών συστημάτων και ο μεγάλος όγκος δεδομένων χρήσης των ΣΔΜ καθιστούν εφικτή την κατασκευή μοντέλων Μηχανικής Μάθησης τα οποία, τροφοδοτούμενα με δεδομένα χρήσης των ΣΔΜ, επιχειρούν να προβλέψουν την επίδοση των σπουδαστών. Τα προβλεπτικά μοντέλα Μηχανικής Μάθησης βασίζονται (και συχνά εξακολουθούν να βασίζονται) σε δημογραφικά, κοινωνικό-οικονομικά δεδομένα και στην προηγούμενη εκπαιδευτική διαδρομή των φοιτητών (Baashar, Alkaws, Ali, Alhussian & Bahboub, 2021. Chen & Cui, 2020. Cui, Chen, Shiri & Fan, 2019). Τα δεδομένα αυτά αντλούνται από παραδοσιακές πηγές (π.χ. αρχεία γραμματειών, ερωτηματολόγια κ.α.), οι οποίες συνήθως είναι διασκορπισμένες, ασύμβατες, ετερογενείς, μη πλήρεις, μη ψηφιοποιημένες, γεγονός που

δυσχεραίνει την άντληση και την επεξεργασία των δεδομένων. Η γενίκευση της χρήσης των ΣΔΜ και η συλλογή μεγάλου όγκου δεδομένων χρήσης, διεύρυνε εντυπωσιακά το πλήθος και το είδος των δεδομένων, με τα οποία μπορούν να τροφοδοτηθούν τα μοντέλα. Συνεπώς, γίνεται ολοένα και περισσότερο ελκυστική και εφικτή η ανάπτυξη μοντέλων πρόβλεψης που χρησιμοποιούν δεδομένα τα οποία προέρχονται αποκλειστικά από τα ΣΔΜ και τα οποία είναι πολυάριθμα, ομοιογενή, ψηφιοποιημένα και γενικά συλλέγονται και επεξεργάζονται εύκολα.

Για την κατασκευή των προβλεπτικών μοντέλων χρησιμοποιούνται κυρίως βασικοί αλγόριθμοι Μηχανικής Μάθησης. Δηλαδή, αλγόριθμοι που βασίζονται σε δέντρα αποφάσεων, σε κανόνες, σε συναρτήσεις, σε μαθηματικά θεωρήματα (Bayes) κ.α. Τελευταία, η ερευνητική κοινότητα διερευνά την αποδοτικότητα μοντέλων, τα οποία χρησιμοποιούν, αντί των βασικών αλγόριθμων, τεχνικές «βαθιάς μάθησης» (deep learning) και ειδικότερα αλγόριθμους που βασίζονται σε περιοδικά νευρωνικά δίκτυα (Recursive Neural Networks) (βλ. Κεφ. 5.3.1). Τέτοιοι αλγόριθμοι είναι σχεδιασμένοι να αναλύουν, όχι μόνο τις συνολικές, τελικές τιμές των χαρακτηριστικών του συνόλου δεδομένων, αλλά και τις χρονικές ακολουθίες που αποτυπώνουν την εξέλιξη των τιμών. Δηλαδή, οι αλγόριθμοι αυτοί διαχειρίζονται τα σύνολα δεδομένων ως δομές, όχι στατικές, αλλά χρονικά εξελισσόμενες. Όπως αναφέρουν οι Doleck, Lemay, Basnet & Bazalais (2020), η μελέτη τέτοιων τεχνικών επί εκπαιδευτικών δεδομένων, αφενός αποτελεί πεδίο αυξανόμενου ερευνητικού ενδιαφέροντος και αφετέρου είναι σχετικά λίγες οι ερευνητικές προσπάθειες που διερευνούν την προβλεπτική ικανότητα τέτοιων τεχνικών (Botelho, Baker & Heffernan, 2017. Doleck, Poitras & Lajoie, 2019). Αναφέρουν επίσης, ότι οι έρευνες σχετικά με την αποδοτικότητα των τεχνικών «βαθιάς μάθησης» δεν συγκλίνουν προς κοινώς αποδεκτά συμπεράσματα. Την ίδια άποψη, περί περιορισμένων ερευνών και περί απόκλισης των συμπερασμάτων, εκφράζουν και οι Chen & Cui, (2020). Συγκεκριμένα, αναφέρουν, ότι οι εφαρμογές τεχνικών «βαθιάς μάθησης» στο πεδίο της προβλεπτικής αναλυτικής είναι σχετικά σπάνιες και ότι είναι εξαιρετικά λίγες οι μελέτες τις οποίες αναφέρει η βιβλιογραφία. Σημειώνουν επίσης, ότι η σχετική έρευνα είναι περιορισμένη συγκριτικά με την έρευνα που μελετά βασικούς αλγόριθμους Μηχανικής Μάθησης και ότι τα συμπεράσματα των ερευνών δεν αναδεικνύουν με σαφήνεια, πόσο καλύτερα αποδίδουν οι τεχνικές «βαθιάς μάθησης» συγκριτικά με τις βασικές τεχνικές Μηχανικής Μάθησης.

Κύριος στόχος των μοντέλων είναι η πρόβλεψη της επίδοσης των φοιτητών και ειδικότερα της πρόβλεψης, είτε του βαθμού, είτε γενικά της «επιτυχίας» ή της «αποτυχίας». Σημαντικό κριτήριο για τη χρησιμότητα των μοντέλων αποτελεί η ακρίβεια της πρόβλεψης, ωστόσο, εξίσου σημαντική είναι και η χρονική στιγμή που επιτυγχάνεται η πρόβλεψη. Για παράδειγμα, έστω ότι υφίσταται μοντέλο που επιτυγχάνει πρόβλεψη υψηλής ακρίβειας στο τέλος του εξαμήνου (ή γενικά όταν το μάθημα ολοκληρωθεί). Το μοντέλο αυτό θεωρείται μεν τεχνικώς επιτυχημένο, ωστόσο διαθέτει περιορισμένη εκπαιδευτική χρησιμότητα. Διότι, η αξία της πρόβλεψης εξαρτάται, όχι μόνο από το πόσο ακριβής είναι, αλλά και από το πόσο έγκαιρη είναι. Η χρησιμότητα των προβλεπτικών εκπαιδευτικών μοντέλων σχετίζεται κυρίως με τη δυνατότητα τους να επιτυγχάνουν πρώιμες προβλέψεις, έστω λιγότερο ακριβείς από τις τελικές. Διότι, με αυτόν τον τρόπο τα προβλεπτικά μοντέλα μπορούν να αξιοποιηθούν ως εργαλεία διαμορφωτικής αξιολόγησης. Συγκεκριμένα, τα μοντέλα, αναλύοντας και αξιολογώντας μοτίβα χρήσης του ΣΔΜ, εντοπίζουν «αδύναμους» φοιτητές, οι οποίοι κινδυνεύουν να αποτύχουν, οπότε χρειάζονται υποστήριξη μέσω διδακτικών παρεμβάσεων, όπως για παράδειγμα ενισχυτική διδασκαλία.

Ο άλλος τρόπος αξιοποίησης των μοντέλων, ως εργαλείων διαμορφωτικής αξιολόγησης, βασίζεται στον εντοπισμό των επιδραστικών χαρακτηριστικών στα σύνολα δεδομένων. Δηλαδή, εκείνων των χαρακτηριστικών, τα οποία περισσότερο από τα υπόλοιπα, συμβάλλουν στην επιτυχημένη λειτουργία των μοντέλων. Κάθε χαρακτηριστικό αποτυπώνει μοτίβα / πρακτικές στη χρήση του ΣΔΜ, όπως π.χ. όγκο χρήσης, συχνότητα χρήσης, περιόδους αδράνειας κ.α. Οπότε, μέσω του εντοπισμού των επιδραστικών χαρακτηριστικών, αποκαλύπτονται πρακτικές χρήσης που σχετίζονται και επηρεάζουν την ακαδημαϊκή επιτυχία. Ο εντοπισμός χαρακτηριστικών κρίσιμων για την ακαδημαϊκή επιτυχία (π.χ. η πρόσβαση σε συγκεκριμένους εκπαιδευτικούς πόρους, η συμμετοχή σε δραστηριότητες, οι περίοδοι αδράνειας), μπορεί να προσανατολίσει τον διδάσκοντα σε εστιασμένες αναπροσαρμογές στον σχεδιασμό του μαθήματος και να αποκαλύψει τη σπουδαιότητα μιας σχεδιαστικής αλλαγής (π.χ. την υποχρεωτική συμμετοχή σε δημόσιους διαλόγους που γίνονται σε τακτά χρονικά διαστήματα).

## 8.2. Σκοπός και στόχοι

**Σκοπός** της διατριβής είναι να συνεισφέρει στο πεδίο της μελέτης των προβλεπτικών μοντέλων Μηχανικής Μάθησης σε εκπαιδευτικά περιβάλλοντα και των δυνατοτήτων αξιοποίησής τους στην εκπαιδευτική αξιολόγηση. Ειδικότερα, η διατριβή διερευνά τις δυνατότητες των μοντέλων να αξιοποιήσουν δεδομένα χρήσης των Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ) και να προβλέψουν τις επιδόσεις των φοιτητών, όχι μόνο με ακρίβεια, αλλά και εγκαίρως, ώστε να μπορούν να χρησιμοποιηθούν ως εργαλεία διαμορφωτικής αξιολόγησης, τα οποία αφενός εντοπίζουν ζητήματα που προοιωνίζονται αποτυχημένη φοίτηση και αφετέρου καθοδηγούν στοχευμένες διδακτικές παρεμβάσεις.

**Επιμέρους στόχοι** της έρευνας είναι:

1. Η άντληση πρωτογενών δεδομένων από τα αρχεία καταγραφής των ΣΔΜ και η συγκρότηση συνόλων δεδομένων, που εμπεριέχουν μοτίβα χρήσης των ΣΔΜ από τους φοιτητές.
2. Η κατασκευή μοντέλων που προβλέπουν την επίδοση των φοιτητών με αξιοποίηση συνόλων δεδομένων, τα οποία συγκροτούνται από χαρακτηριστικά των οποίων οι τιμές σχετίζονται αποκλειστικά με τη χρήση των ΣΔΜ.
3. Η κατασκευή μοντέλων με τη χρήση τεχνικών «βαθιάς μάθησης» που αναλύουν, όχι μόνο τις συνολικές, τελικές τιμές των χαρακτηριστικών των μοντέλων, αλλά και τις χρονικές ακολουθίες της εξέλιξης των τιμών.
4. Η συγκριτική αξιολόγηση της απόδοσης των μοντέλων με κριτήριο την ακρίβεια της πρόβλεψης.
5. Η συγκριτική αξιολόγηση της απόδοσης των μοντέλων με κριτήριο τον χρόνο της πρόβλεψης.
6. Ο εντοπισμός χαρακτηριστικών τα οποία έχουν την μεγαλύτερη επίδραση στην απόδοση των μοντέλων και είναι τα καταλληλότερα για την κατασκευή προβλεπτικών μοντέλων στα εκπαιδευτικά περιβάλλοντα.

### 8.3. Ερευνητικά ερωτήματα

Τα ερωτήματα τα οποία τέθηκαν κατά την εκπόνηση της διατριβής και τα οποία η έρευνα επιχείρησε να απαντήσει είναι τα εξής:

#### *Ερευνητικό Ερώτημα EEP1:*

Είναι εφικτή η πρόβλεψη της επίδοσης των φοιτητών με τη βοήθεια μοντέλων Μηχανικής Μάθησης, τα οποία βασίζονται αποκλειστικά σε δεδομένα χρήσης του ΣΔΜ; Επιπλέον:

- 1.1 Τι τιμή ακρίβειας πρόβλεψης επιτυγχάνεται;
- 1.2 Είναι η πρόβλεψη αξιόπιστη;

#### *Ερευνητικό Ερώτημα EEP2:*

Είναι εφικτή η πρώιμη πρόβλεψη της επίδοσης των φοιτητών; Επιπλέον:

- 2.1 Πόσο πρώιμα επιτυγχάνεται η πρόβλεψη;
- 2.2 Τι τιμή ακρίβειας πρόβλεψης επιτυγχάνεται;
- 2.3 Είναι η πρόβλεψη αξιόπιστη;

#### *Ερευνητικό Ερώτημα EEP3:*

Είναι τα μοντέλα που προκύπτουν από τεχνικές «βαθιάς μάθησης» καλύτερα από τα μοντέλα που προκύπτουν από βασικούς αλγόριθμους Μηχανικής Μάθησης;

#### *Ερευνητικό Ερώτημα EEP4:*

Ποιοι αλγόριθμοι Μηχανικής Μάθησης παράγουν μοντέλα με υψηλή ακρίβεια και πρώιμη πρόβλεψη;

#### *Ερευνητικό Ερώτημα EEP5:*

Είναι ικανά τα παραγόμενα μοντέλα να επιτύχουν αξιολογη ακρίβεια πρόβλεψης, αν δοκιμαστούν σε «μη οικεία» σύνολα δεδομένων, δηλαδή σε σύνολα διαφορετικά από αυτά που χρησιμοποιήθηκαν για τη δημιουργία τους;

*Ερευνητικό Ερώτημα ΕΕΡ6:*

Ποια από τα χαρακτηριστικά που συγκροτούν τα σύνολα δεδομένων έχουν τη μεγαλύτερη επίδραση στα αποτελέσματα της πρόβλεψης;

*Ερευνητικό Ερώτημα ΕΕΡ7:*

Πώς ερμηνεύεται ο αυξημένος βαθμός επίδρασης κάποιων χαρακτηριστικών, που συγκροτούν τα σύνολα δεδομένων; Υπάρχει σχέση των επιδραστικών χαρακτηριστικών με την ενεργή συμμετοχή και τον ρόλο της στην επιτυχημένη φοίτηση;

*Ερευνητικό Ερώτημα ΕΕΡ8:*

Επηρεάζεται η απόδοση των μοντέλων πρόβλεψης από ζητήματα που αφορούν δομικά χαρακτηριστικά του μαθήματος και γενικότερα από ζητήματα που αφορούν τον εκπαιδευτικό σχεδιασμό;

## 9. ΜΕΘΟΔΟΣ

Η έρευνα που διεξαγάγαμε για την εκπόνηση της διατριβής εξελίχθηκε σε δύο στάδια:

A. Πραγματοποιήθηκε ανασκόπηση της διεθνούς βιβλιογραφίας με σκοπό, αφενός την περιγραφή του ερευνητικού πεδίου της διατριβής και αφετέρου την παράθεση και τον σχολιασμό των σημαντικότερων συμπερασμάτων από ανάλογες έρευνες επί του ερευνητικού πεδίου. Η ανασκόπηση της βιβλιογραφίας περιλαμβάνεται στο θεωρητικό μέρος της διατριβής.

B. Μελετήθηκαν δεδομένα που παράγει η χρήση των Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ), με βασικό σκοπό να διερευνηθούν πιθανές συσχετίσεις μεταξύ της χρήσης και της επίδοσης των φοιτητών. Η μελέτη των δεδομένων, δηλαδή ο εντοπισμός, η άντληση, η επεξεργασία των δεδομένων και ακολούθως η εφαρμογή τεχνικών Μηχανικής Μάθησης, καθώς και η παρουσίαση και ο σχολιασμός των αποτελεσμάτων που προέκυψαν, περιλαμβάνεται στο ερευνητικό μέρος της διατριβής.

Για την ανάπτυξη του θεωρητικού μέρους της διατριβής πραγματοποιήθηκε εκτενής βιβλιογραφική ανασκόπηση. Η ανασκόπηση ήταν *επιλεκτική* για το πεδίο της αξιοποίησης τεχνικών Μηχανικής Μάθησης στη Μαθησιακή Αναλυτική και *συστηματική* για το πεδίο της αξιοποίησης τεχνικών Μηχανικής Μάθησης στην κατασκευή μοντέλων πρόβλεψης της επίδοσης των φοιτητών.

Για την ανάπτυξη του ερευνητικού μέρους της διατριβής αξιοποιήθηκαν διάφορες μέθοδοι ανάλογα με τα στάδια της έρευνας.

Για την συλλογή και την επεξεργασία των πρωτογενών δεδομένων, για τον καθορισμό των χαρακτηριστικών στα σύνολα δεδομένων, για τον καθορισμό των κλάσεων της ταξινόμησης, για τη δημιουργία των συνόλων δεδομένων, για την επιλογή των αλγόριθμων ταξινόμησης και για την κατασκευή των μοντέλων πρόβλεψης, αξιοποιήθηκε η μέθοδος της *έρευνας με σχεδιασμό*. Η *έρευνα με σχεδιασμό* είναι, σύμφωνα με τους Barab & Squire (2004, σ. 2), μια σειρά από προσεγγίσεις με σκοπό την παραγωγή νέων θεωριών, αντικειμένων και πρακτικών, τα οποία εξηγούν και επηρεάζουν τη μάθηση και τη διδασκαλία σε πραγματικά περιβάλλοντα. Η *έρευνα με σχεδιασμό* επιχειρεί να ενισχύσει τις θεωρητικές βάσεις και την πρακτική αξία της έρευνας



στην εκπαιδευτική τεχνολογία (Herrington, McKenney, Reeves, & Oliver, 2007) και να αντιμετωπίσει πρακτικά προβλήματα διδακτικής φύσης, μέσα από μια επαναλαμβανόμενη διαδικασία σχεδιασμού, εφαρμογής και αξιολόγησης (Μαυρομάτη, 2019, σσ. 75-88).

Για τη συγκριτική αξιολόγηση των προβλεπτικών μοντέλων και την αναζήτηση επιδραστικών χαρακτηριστικών στα σύνολα δεδομένων, αξιοποιήθηκε η *πειραματική μέθοδος* και ειδικότερα, ο *πειραματισμός κατά αντιστροφή* ή *εκ των υστέρων έρευνα*, όπως ορίζεται από τους Cohen, Manion & Morrison (2007, σ. 361). Η *εκ των υστέρων έρευνα* διερευνά πιθανές σχέσεις αιτίου – αιτιατού, αναζητώντας πίσω στο χρόνο εύλογους αιτιώδεις παράγοντες και μελετά γεγονότα «τα οποία ήδη έχουν συμβεί και συνεπώς δεν μπορούν να κατασκευαστούν ή να χειραγωγηθούν από τον ερευνητή» (Cohen, Manion, & Morrison, 2007, σσ. 351-352). Σύμφωνα με τον Kerlinger (1970), στην *εκ των υστέρων έρευνα*, ο ερευνητής μελετά αναδρομικά τις εξαρτημένες και ανεξάρτητες μεταβλητές για να διαπιστώσει την πιθανή σχέση τους και εξετάζει αναδρομικά τις επιδράσεις ενός γεγονότος σε ένα μεταγενέστερο γεγονός (Cohen, Manion, & Morrison, 2007, σσ. 351-352).

Στα πλαίσια της *πειραματικής μεθόδου* και προκειμένου να μελετηθεί η αποδοτικότητα των μοντέλων στα διαφορετικά μαθήματα, αξιοποιήθηκαν *διερευνητικές μελέτες περίπτωσης*. Η *διερευνητική μελέτη περίπτωσης*, είναι, σύμφωνα με τον Yin (1984), ένα από τα τρία είδη των μελετών περίπτωσης και λειτουργεί πιλοτικά για άλλες μελέτες ή ερευνητικά ερωτήματα (Cohen, Manion & Morrison, 2007, σ. 312). Αξιοποιείται για τη βασική κατανόηση του ερευνητικού αντικειμένου και χρησιμεύει στην κατασκευή υποθέσεων και ερευνητικών ερωτημάτων και στον σχεδιασμό κατευθύνσεων, που θα μελετηθούν περαιτέρω από επόμενες έρευνες (Chopard & Przybylski, 2021).

### 9.1. Ο υπό μελέτη πληθυσμός

Μελετήθηκαν δεδομένα χρήσης **440** φοιτητών (362 προπτυχιακών και 78 μεταπτυχιακών) του Τμήματος Επιστημών της Προσχολικής Αγωγής και του Εκπαιδευτικού Σχεδιασμού (ΤΕΠΑΕΣ) του Πανεπιστημίου Αιγαίου, οι οποίοι για τις ανάγκες παρακολούθησης των μαθημάτων τους χρησιμοποίησαν το Σύστημα Διαχείρισης Μάθησης (ΣΔΜ) Moodle, στο οποίο έχουν δημιουργηθεί σελίδες για όλα τα προσφερόμενα από το Τμήμα μαθήματα.

Ειδικότερα μελετήθηκαν δεδομένα χρήσης για τις σελίδες των παρακάτω **τεσσάρων** μαθημάτων:

#### Προπτυχιακό Πρόγραμμα Σπουδών

- ΑΑ005, Χρήση Βασικών Εφαρμογών ΗΥ, Χειμερινό Εξάμηνο 2020-2021
- ΑΑ005, Χρήση Βασικών Εφαρμογών ΗΥ, Χειμερινό Εξάμηνο 2021-2022

#### Μεταπτυχιακό Πρόγραμμα Σπουδών, Νέες Μορφές Εκπαίδευσης και Μάθησης

- Α3, Ηλεκτρονική Μάθηση, Χειμερινό Εξάμηνο 2019-2020
- Α3, Ηλεκτρονική Μάθηση, Χειμερινό Εξάμηνο 2020-2021

### 9.2. Περιγραφή μαθημάτων

Το μάθημα «ΑΑ005, Χρήση Βασικών Εφαρμογών ΗΥ» είναι υποχρεωτικό μάθημα του Α' εξαμήνου στο προπτυχιακό πρόγραμμα σπουδών του ΤΕΠΑΕΣ. Προσφέρθηκε για πρώτη φορά το χειμερινό εξάμηνο 2020-2021. Πρόκειται για μάθημα που *«έχει ως στόχο την εκμάθηση δεξιοτήτων για τη χρήση α) βασικών δικτυακών εφαρμογών (π.χ. περιηγητές διαδικτύου, ηλεκτρονικό ταχυδρομείο) β) ακαδημαϊκών ψηφιακών υπηρεσιών (δήλωση μαθημάτων, σύστημα διαχείρισης μάθησης) και γ) βασικών εφαρμογών ΗΥ γενικής χρήσης (επεξεργασία κειμένου). Ο σκοπός του μαθήματος είναι να καταστούν ικανοί οι φοιτητές/τριες να χρησιμοποιούν βασικές ψηφιακές υπηρεσίες και εφαρμογές γενικής χρήσης τόσο για τις ανάγκες παρακολούθησης άλλων μαθημάτων όσο και για προσωπική τους χρήση»*. (Οδηγός Σπουδών ΤΕΠΑΕΣ, 2022). Το μάθημα διεξάγεται σε 13 δια ζώσης διαλέξεις, στις οποίες η συμμετοχή των φοιτητών είναι προαιρετική. Η αξιολόγηση γίνεται στο τέλος του εξαμήνου με γραπτές εξετάσεις, όπου οι φοιτητές καλούνται να υλοποιήσουν μικρή εργασία με τη βοήθεια

κατάλληλου λογισμικού (π.χ. να συντάξουν αρχείο κειμένου). Για το μάθημα έχει δημιουργηθεί σελίδα στο ΣΔΜ Moodle, στην οποία αναρτώνται εκπαιδευτικό υλικό (οδηγίες, κείμενα, ασκήσεις) και ανακοινώσεις. Το εκπαιδευτικό υλικό είναι οργανωμένο ανά διδακτική ενότητα και αναρτάται με την έναρξη της αντίστοιχης διδακτικής ενότητας. Το έτος 2020-2021, λόγω της πανδημίας του κορονοϊού Covid-19 και της αλλαγής στον τρόπο διεξαγωγής των μαθημάτων, οι διαλέξεις και οι γραπτές εξετάσεις πραγματοποιήθηκαν, όχι δια ζώσης, αλλά αποκλειστικά με τηλεδιάσκεψη, με το σύστημα τηλεδιασκέψεων Zoom. Το ΣΔΜ Moodle χρησιμοποιήθηκε για την ασύγχρονη μελέτη του εκπαιδευτικού υλικού. Οπότε, το έτος 2020-2021 το μάθημα προσφέρθηκε με διαδικτυακή διδασκαλία (online teaching), ενώ το έτος 2021-2022, όταν αποκαταστάθηκε η κανονική λειτουργία των εκπαιδευτικών ιδρυμάτων, το μάθημα προσφέρθηκε με Τεχνολογικά Ενισχυμένη Πρόσωπο με Πρόσωπο Διδασκαλία και Μάθηση (Technology Enhanced Face to Face Teaching and Learning) (Φεσάκης & Κωνσταντοπούλου, 2022, σ. 34).

Για το έτος 2020-2021, ο αριθμός των εγγεγραμμένων φοιτητών στο μάθημα ήταν **287**, ενώ για το έτος 2021-2022 ήταν **348**, σύμφωνα με τα στοιχεία του συστήματος φοιτητολογίου του Πανεπιστημίου Αιγαίου. Μελετήθηκαν δεδομένα πρόσβασης για όσους φοιτητές (**159** για το έτος 2020-2021 και **203** για το έτος 2021-2022) ίσχυαν οι ακόλουθες προϋποθέσεις: Α) Οι σπουδαστές είχαν χρησιμοποιήσει το ΣΔΜ Moodle, οπότε υπήρχαν δεδομένα πρόσβασης που τους αφορούσαν στο αρχείο καταγραφής. Β) Ήταν ενεργοί φοιτητές την ημέρα της εξέτασης, ανεξαρτήτως αν προσήλθαν, ή αν δεν προσήλθαν να εξεταστούν και ανεξαρτήτως αν διαγράφηκαν αργότερα. Γ) Δεν εξέφρασαν αντίρρηση για τη μελέτη των δεδομένων που τους αφορούσαν. Οι υπόλοιποι φοιτητές, είτε δεν χρησιμοποίησαν το ΣΔΜ (οπότε δεν εντοπίστηκαν δεδομένα πρόσβασης που τους αφορούσαν), είτε διαγράφηκαν από το μάθημα πριν την ολοκλήρωση του ακαδημαϊκού εξαμήνου (οπότε θεωρούμε ότι κάποια στιγμή κατά τη διάρκεια του εξαμήνου σταμάτησαν την ενασχόληση με το μάθημα, άρα τα δεδομένα πρόσβασής τους δεν μπορούν να χρησιμεύσουν στην έρευνά μας)<sup>2</sup>, είτε εισήχθησαν στο Τμήμα μέσω

---

<sup>2</sup> Ας σημειωθεί, ότι οι φοιτητές μπορεί να διαγραφούν (π.χ. λόγω μεταγραφής σε άλλο Πανεπιστημιακό Τμήμα) οποτεδήποτε. Είτε κατά τη διάρκεια εξαμήνου στο οποίο τα μαθήματα έχουν εγγραφεί, είτε αργότερα. Δεδομένου ότι το φοιτητολόγιο εμφανίζει πάντα την τρέχουσα κατάσταση των φοιτητών, είναι δυνατόν να

κατατακτηρίων εξετάσεων, οπότε λόγω της καθυστερημένης εγγραφής τους στο μάθημα δεν είχαν τη δυνατότητα να κάνουν επαρκή χρήση του ΣΔΜ, είτε εξέφρασαν αντίρρηση για τη μελέτη των δεδομένων που τους αφορούσαν.

Το μάθημα «Α3, Ηλεκτρονική Μάθηση» είναι υποχρεωτικό μάθημα του Α΄ Εξαμήνου στο Μεταπτυχιακό Προγραμμάτων Σπουδών Νέες Μορφές Εκπαίδευσης και Μάθησης του ΤΕΠΑΕΣ. Προσφέρθηκε για πρώτη φορά το χειμερινό εξάμηνο 2018-2019<sup>3</sup>. Το μάθημα «αναφέρεται στις νέες μορφές ηλεκτρονικής μάθησης υπό το πρίσμα παιδαγωγικά τεκμηριωμένων μεθοδολογιών. Δομείται σε δύο μέρη: α) σε ένα θεωρητικό και β) σε ένα ειδικό πλαίσιο αναφοράς. Στο πρώτο διευκρινίζονται βασικές έννοιες που σχετίζονται με την ηλεκτρονική μάθηση, παρουσιάζονται λειτουργίες του εκπαιδευτικού υλικού, βασικές παιδαγωγικές θέσεις για τη δημιουργία ηλεκτρονικών περιβαλλόντων μάθησης και μεθοδολογίες σχεδιασμού ηλεκτρονικού μαθήματος. Το δεύτερο μέρος εστιάζει σε ειδικά θέματα διδακτικής μεθοδολογίας ηλεκτρονικών μαθημάτων και παρουσιάζει διδακτικά μοντέλα με διακριτές θεωρητικές προσεγγίσεις. Κοινό στοιχείο όλων των νέων μορφών διδασκαλίας και μάθησης είναι ότι συνδέονται άμεσα με βασικές θέσεις της προοδευτικής Παιδαγωγικής και έχουν σημείο αναφοράς την ενεργή συμμετοχή των εκπαιδευομένων, τον προσανατολισμό στη δράση και στη βιωματική μάθηση» (Μαθήματα ΠΜΣ Νέες Μορφές Εκπαίδευσης και Μάθησης, 2022). Το μάθημα διεξάγεται σε 10 συνεδρίες (με υποχρεωτική συμμετοχή), εκ των οποίων η 1<sup>η</sup> και η 10<sup>η</sup> γίνονται δια ζώσης. Οι υπόλοιπες 8 συνεδρίες διεξάγονται με ασύγχρονη μελέτη εκπαιδευτικού υλικού του ΣΔΜ Moodle. Δηλαδή, στο μάθημα συνδυάζονται η διαδικτυακή και η πρόσωπο με πρόσωπο διδασκαλία (Μεικτή Μάθηση) (Φεσάκης & Κωνσταντοπούλου, 2022, σ. 35). Το 2020-2021, λόγω της πανδημίας του κορονοϊού Covid-19 και της αλλαγής στον τρόπο διεξαγωγής των μαθημάτων, όλες οι συνεδρίες πραγματοποιήθηκαν με τηλεδιάσκεψη με το σύστημα τηλεδιασκέψεων Zoom, οπότε το μάθημα προσφέρθηκε με διαδικτυακή διδασκαλία.

Για το μάθημα έχει δημιουργηθεί σελίδα στο ΣΔΜ Moodle του Πανεπιστημίου, όπου αναρτάται εκπαιδευτικό υλικό οργανωμένο σε διδακτικές ενότητες, αντίστοιχες με τις συνεδρίες του

---

εμφανίζει φοιτητές διαγραμμένους από το μάθημα και ταυτοχρόνως αξιολογημένους στο μάθημα. Για αυτούς τους φοιτητές, συνάγουμε το προφανές συμπέρασμα ότι διαγράφηκαν μετά την ημέρα της εξέτασης του μαθήματος.

<sup>3</sup> Μάθημα με το ίδιο όνομα προσφερόταν και πριν το 2018, αλλά είχε άλλη δομή και άλλον διδάσκοντα.

μαθήματος. Στο ΣΔΜ Moodle, υπό τον συντονισμό του διδάσκοντα, διεξάγονται οι δημόσιοι διάλογοι μεταξύ των φοιτητών. Για κάθε συνεδρία διεξάγεται ένας δημόσιος διάλογος. Κατά την διάρκεια της 5<sup>ης</sup> συνεδρίας πραγματοποιείται τηλεδιάσκεψη με τη συμμετοχή όλων των φοιτητών και του διδάσκοντα. Το υλικό κάθε συνεδρίας αναρτάται την πρώτη μέρα έναρξης της συνεδρίας. Οι φοιτητές αξιολογούνται βάσει της συμμετοχής τους στους δημόσιους διαλόγους και της τελικής εργασίας που οφείλουν να υλοποιήσουν. Για τα έτη που μελετήσαμε (2019-2020 και 2020-2021), η τελική αξιολογική εργασία ήταν ατομική.<sup>4</sup> Στον τελικό βαθμό των φοιτητών προσμετράται κατά 30% η συμμετοχή στους δημόσιους διαλόγους των 7 πρώτων συνεδριών και κατά 70% η τελική αξιολογική εργασία.

Για το έτος 2019-2020, ο αριθμός των εγγεγραμμένων φοιτητών στο μάθημα ήταν **39**, σύμφωνα με το σύστημα φοιτητολογίου του Πανεπιστημίου Αιγαίου. Για το έτος 2020-2021, ο αριθμός των εγγεγραμμένων φοιτητών στο μάθημα ήταν επίσης **39**. Μελετήσαμε δεδομένα πρόσβασης όλων των φοιτητών, δεδομένου ότι όλοι παρέμειναν ενεργοί για όλη τη διάρκεια των εξαμήνων και όλοι χρησιμοποίησαν το ΣΔΜ Moodle, διότι η χρήση του συστήματος ήταν υποχρεωτική.

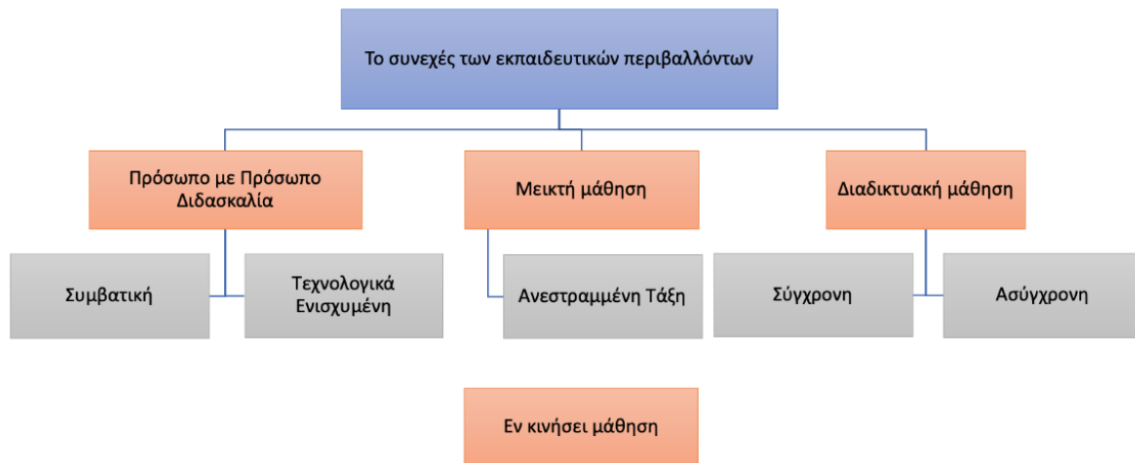
Συνοπτικά, τα μαθήματα και τα χαρακτηριστικά τους καταγράφονται στον Πίνακα 9.

**Πίνακας 9. Τα υπό μελέτη μαθήματα**

Τίτλος	Έτος	Βαθμίδα	Υποχρεωτική παρακολούθηση	Εγγεγραμμένοι φοιτητές	Φοιτητές υπό μελέτη	Αξιολόγηση	Διδασκαλία	Κατάταξη στο συνεχές στην εκπαιδευτικών περιβαλλόντων
AA005 Χρήση Βασικών Εφαρμογών ΗΥ	2020-2021	Προπτυχιακό	Όχι	287	159	Γραπτή Εξέταση	Σύγχρονες διαδικτυακές διαλέξεις (λόγω Covid 19), ασύγχρονη μελέτη εκπαιδευτικού υλικού	Διαδικτυακή (σύγχρονη & ασύγχρονη) Μάθηση
	2021-2022			348	203		Δια ζώσης διαλέξεις, ασύγχρονη μελέτη εκπαιδευτικού υλικού	Τεχνολογικά Ενισχυμένη πρόσωπο με πρόσωπο Διδασκαλία και Μάθηση
A3 Ηλεκτρονική Μάθηση	2019-2020	Μεταπτυχιακό	Ναι	39	39	Εργασία και συμμετοχή σε δημόσιο ηλεκτρονικό διάλογο	Δια ζώσης διαλέξεις και ασύγχρονες συνεδρίες	Μεικτή μάθηση
	2020-2021			39	39		Σύγχρονες διαδικτυακές διαλέξεις (λόγω Covid 19), ασύγχρονη μελέτη εκπαιδευτικού υλικού	Διαδικτυακή (σύγχρονη & ασύγχρονη) Μάθηση

<sup>4</sup> Αυτός ήταν ο λόγος που επιλέξαμε να μελετήσουμε το μάθημα για τα συγκεκριμένα έτη. Τα άλλα έτη, η εργασία ήταν ομαδική.

Επιδιώξαμε, τα μαθήματα που μελετήθηκαν να καλύπτουν (σύμφωνα με τον τρόπο διεξαγωγής τους) το συνεχές των τεχνολογικά υποστηριζόμενων εκπαιδευτικών περιβαλλόντων (Σχήμα 11) (Φεσάκης & Κωνσταντοπούλου, σσ. 34-35). Ειδικότερα, το μάθημα «ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ» κατά το έτος 2020-2021 κατατάσσεται στην Διαδικτυακή Διδασκαλία (Online teaching) με σύγχρονο και ασύγχρονο τρόπο, διότι λόγω των περιορισμών της πανδημίας του κορονοϊού Covid-19 υλοποιήθηκε εξ ολοκλήρου μέσω διαδικτύου (σύγχρονες διαδικτυακές διαλέξεις και ασύγχρονη μελέτη εκπαιδευτικού υλικού). Το ίδιο μάθημα, κατά το έτος 2021-2022, κατατάσσεται στην Τεχνολογικά Ενισχυμένη Πρόσωπο με Πρόσωπο Διδασκαλία και Μάθηση (Technology Enhanced Face to Face Teaching and Learning), διότι υλοποιήθηκε με διαζώσης διαλέξεις πρόσωπο με πρόσωπο και επιπλέον αξιοποιήθηκε το ΣΔΜ Moodle για την ανάρτηση εκπαιδευτικού υλικού και για την επικοινωνία των συμμετεχόντων. Το μάθημα «Α3 Ηλεκτρονική Μάθηση», κατά το έτος 2019-2020, κατατάσσεται στην Μεικτή διδασκαλία (blended teaching), διότι συνδυάστηκαν η διαδικτυακή και η πρόσωπο με πρόσωπο διδασκαλία. Το ίδιο μάθημα, κατά το επόμενο έτος 2020-2021, κατατάσσεται στην Διαδικτυακή Διδασκαλία (Online teaching) με σύγχρονο και ασύγχρονο τρόπο, διότι λόγω των περιορισμών της πανδημίας του κορονοϊού Covid-19, υλοποιήθηκε εξ ολοκλήρου μέσω διαδικτύου (σύγχρονες διαδικτυακές διαλέξεις και ασύγχρονη μελέτη εκπαιδευτικού υλικού). Συνεπώς, τα μαθήματα που μελετήσαμε καλύπτουν και τους 3 βασικούς τρόπους διδασκαλίας που αξιοποιούν την Τεχνικά Ενισχυμένη Μάθηση (Πρόσωπο με Πρόσωπο Διδασκαλία, Μεικτή Μάθηση, Διαδικτυακή Μάθηση) στο συνεχές των εκπαιδευτικών περιβαλλόντων (Σχήμα 11) (Φεσάκης & Κωνσταντοπούλου, 2022, σσ. 34-35).



**Σχήμα 11. Το συνεχές των εκπαιδευτικών περιβαλλόντων (Φεσάκης & Κωνσταντοπούλου, 2022, σσ. 34-35)**

### 9.3. Ερευνητική δεοντολογία

Η έρευνα διεξάχθηκε σύμφωνα με την Ελληνική και Ευρωπαϊκή νομοθεσία για την προστασία προσωπικών δεδομένων και διέπεται από τους κανόνες της ερευνητικής δεοντολογίας. Χορηγήθηκαν οι απαραίτητες άδειες<sup>5</sup> διεξαγωγής εκπαιδευτικής έρευνας. Λήφθηκαν όλα τα απαραίτητα μέτρα, ώστε η έρευνα και τα αποτελέσματά της να μην βλάψουν, ή να προκαλέσουν διακρίσεις ή άνιση μεταχείριση σε οποιονδήποτε. Τα δεδομένα που αντλήθηκαν από τα συστήματα αποσυσχετίστηκαν από προσωπικά στοιχεία. Οι σχετικές εγγραφές «ψευδωνυμοποιήθηκαν», δηλαδή αντιστοιχήθηκαν με ψευδώνυμες οντότητες (π.χ. Φ1, δηλαδή φοιτητής1, Φ2 δηλαδή, φοιτητής2). Όσα στοιχεία προσδιορίζουν προσωπικά τους συμμετέχοντες (π.χ. Ονοματεπώνυμο, Αριθμός Μητρώου, username) διαγράφηκαν οριστικά. Δεν υπάρχει δυνατότητα αντιστοίχισης των ψευδώνυμων οντοτήτων με το φυσικό πρόσωπο στο οποίο αντιστοιχούν και γενικότερα δεν υπάρχει δυνατότητα ταυτοποίησης των συμμετεχόντων.

<sup>5</sup> Από τη Συνέλευση του Τμήματος Επιστημών Προσχολικής Αγωγής και Εκπαιδευτικού Σχεδιασμού κατά την 2<sup>η</sup>/19.10.2022 (Θέμα 8.4) Συνέλευση, κατόπιν της υπ. αριθμό 2956/12.10.2022 εισήγησης της Επιτροπής Εκπαιδευτικής Έρευνας και κατά την 10<sup>η</sup>/26.6.2022 (Θέμα 7.3) Συνέλευση, κατόπιν της υπ. αριθμό 1758/11.6.2021 εισήγησης της Επιτροπής Εκπαιδευτικής Έρευνας και από τη Διευθύντρια του Προγράμματος Μεταπτυχιακών Σπουδών Νέες Μορφές Εκπαίδευσης και Μάθησης σύμφωνα με την από 9.6.2021 σχετική βεβαίωσή της.

Τα αρχεία καταγραφής, που αντλήθηκαν από το σύστημα και τα οποία περιλαμβάνουν τα πλήρη στοιχεία των φοιτητών, διαγράφηκαν οριστικά από τον υπολογιστή, όπου προσωρινά αποθηκεύτηκαν αμέσως μετά την «ψευδωνυμοποίηση» των εγγραφών. Για λόγους τεκμηρίωσης, διατηρήθηκαν μόνο αρχεία με ανώνυμα ή ψευδώνυμα δεδομένα, τα οποία τηρούνται σε ειδικά ψηφιακά μέσα, μη συνδεδεμένα στο διαδίκτυο.

Οι δημοσιεύεις που αφορούν την έρευνα περιλαμβάνουν μόνο ερευνητικά αποτελέσματα της επεξεργασίας δεδομένων και μετρήσεις, ή στατιστικά μέτρα που δεν επιτρέπουν ταυτοποίηση προσώπων. Δεν έγινε γνωστοποίηση των δεδομένων σε τρίτους.

Τα δεδομένα αντλήθηκαν από συμμετέχοντες που χορήγησαν σχετική συναίνεση, κατόπιν ενημέρωσης για τα είδη των δεδομένων, που θα αντληθούν καθώς και για τον τρόπο και τον σκοπό της επεξεργασίας τους. Οι φοιτητές του μαθήματος «ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ» έλαβαν, στο ακαδημαϊκό ηλεκτρονικό γραμματοκιβώτιο τους, μήνυμα με το οποίο ενημερώθηκαν για την επιθυμία του ερευνητή να μελετήσει τα εν λόγω δεδομένα, αποκλειστικά και μόνο για σκοπούς ακαδημαϊκής έρευνας. Με το ίδιο μήνυμα κλήθηκαν όσοι φοιτητές δεν επιθυμούσαν να μελετηθούν τα δεδομένα πρόσβασης και χρήσης που τους αφορούν, να επισκεφθούν διαδικτυακή σελίδα, όπου εύκολα και γρήγορα (με ένα «κλικ» του ποντικιού) μπορούσαν να δηλώσουν την άρνησή τους. Οι φοιτητές που δήλωσαν ότι δεν επιθυμούν να μελετηθούν δεδομένα που τους αφορούν εξαιρέθηκαν από την έρευνα (Εικόνα 2).



## ΑΑ005 ΧΡΗΣΗ ΒΑΣΙΚΩΝ ΕΦΑΡΜΟΓΩΝ Η/Υ [ΧΕΙΜΕΡΙΝΟ ΕΞΑΜΗΝΟ 2021-2022]

Dashboard / My courses / ΑΑ005 [ΧΕΙΜΕΡΙΝΟ ΕΞΑΜΗΝΟ 2021-2022] / General / Δήλωση αντίρρησης για μελέτη δεδομένων πρόσβασης στο μάθημα. / View 3 responses / Responses

Δήλωση αντίρρησης για μελέτη δεδομένων πρόσβασης στο μάθημα.

Responses

Choice options	Δεν επιθυμώ να μελετηθούν δεδομένα που αφορούν την πρόσβασή μου στο μάθημα ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ [ΧΕΙΜΕΡΙΝΟ ΕΞΑΜΗΝΟ 2021-2022] <input type="checkbox"/>
Number of responses	3
Users who chose this option	<ul style="list-style-type: none"><li><input type="checkbox"/> Κο</li><li><input type="checkbox"/> Β. Χακί</li><li><input type="checkbox"/> Μα. Χακί</li></ul>

**Εικόνα 2. Διαδικτυακή σελίδα όπου οι φοιτητές μπορούσαν να δηλώσουν άρνηση για τη μελέτη δεδομένων πρόσβασης τους**

Οι φοιτητές του μαθήματος «Α3 Ηλεκτρονική Μάθηση» συναίνεσαν επίσης στα πλαίσια παράλληλης ερευνητικής εργασίας του επιβλέποντος τη διατριβή.

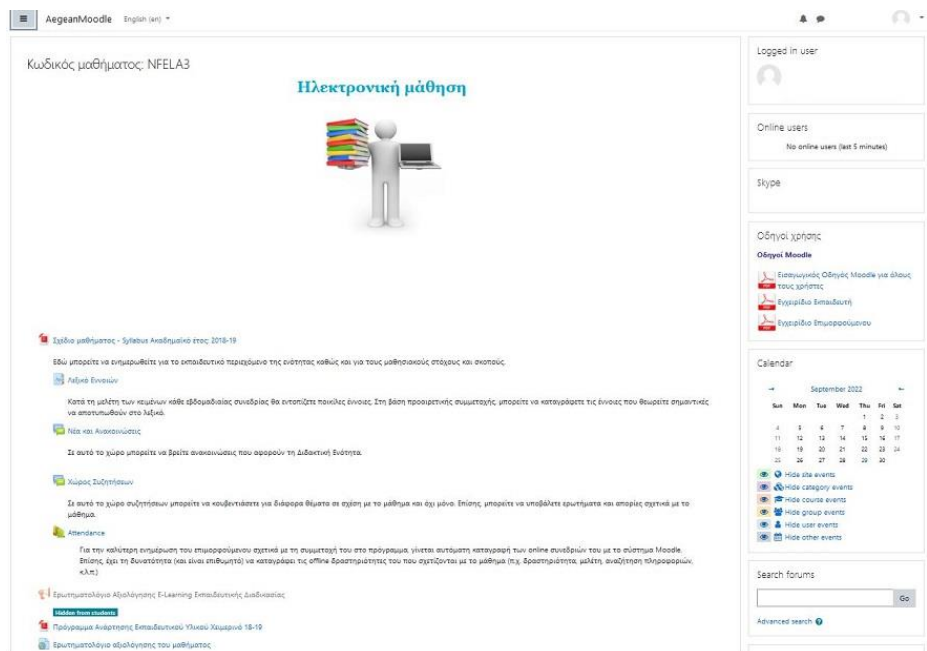
Συνοψίζοντας, η έρευνά μας εξασφάλισε τη συναίνεση των υποκειμένων, ψευδωνυμοποίησε τις εγγραφές, απέκλεισε την πρόσβαση στα αρχεία με τα πρωτογενή δεδομένα, έλαβε την απαραίτητη άδεια έρευνας από την αρμόδια επιτροπή του Τμήματος και γενικά τήρησε τους σχετικούς κανόνες της ερευνητικής δεοντολογίας.

## 9.4. Εργαλεία

### 9.4.1. Το Σύστημα Διαχείρισης Μάθησης Moodle του Πανεπιστημίου Αιγαίου

Το Σύστημα Διαχείρισης Μάθησης Moodle (Modular Object-Oriented Dynamic Learning Environment) είναι ένα ελεύθερο και ανοικτού κώδικα Σύστημα Διαχείρισης Μάθησης (ΣΔΜ), του οποίου η ανάπτυξη ξεκίνησε το 1999. Είναι μεταξύ των δημοφιλέστερων ΣΔΜ, με αριθμό χρηστών που υπερβαίνει τα 213 εκατομμύρια. Μεταξύ των πανεπιστημίων, οργανισμών και επιχειρήσεων που χρησιμοποιούν το Moodle συγκαταλέγονται το Open University, το London School of Economics, η Microsoft, η Shell (About Moodle, 2022).

Το Πανεπιστήμιο Αιγαίου χρησιμοποιεί το Moodle (έκδοση 3.9) ως το κύριο ΣΔΜ για την υποστήριξη των προπτυχιακών και μεταπτυχιακών μαθημάτων, τα οποία προσφέρονται από τα Τμήματα καθώς και των επιμορφωτικών προγραμμάτων τα οποία προσφέρονται από το Κέντρο Επιμόρφωσης και Δια Βίου Μάθησης του Ιδρύματος. Για κάθε προσφερόμενο μάθημα υφίσταται διαδικτυακή σελίδα (Εικόνα 3), στην οποία, αναλόγως με τη δομή, τα περιεχόμενα και τον σκοπό του μαθήματος, αναρτάται εκπαιδευτικό υλικό (π.χ. σημειώσεις, ασκήσεις) και ανακοινώσεις, διεξάγονται δημόσιοι διάλογοι, διενεργούνται διαγωνίσματα και δημιουργούνται διαδικτυακοί χώροι για την υποβολή από τους φοιτητές αρχείων με τις εργασίες τους. Η διαδικτυακή σελίδα δεν παραμένει ίδια για όλα τα έτη που προσφέρεται το μάθημα, αλλά κάθε έτος, για κάθε μάθημα, δημιουργείται νέα σελίδα. Προφανώς, ο τρόπος και η έκταση χρήσης των σελίδων των μαθημάτων στο Moodle δεν είναι ίδιος για όλα τα μαθήματα και για όλους τους φοιτητές. Εξαρτώνται από τον τρόπο οργάνωσης και διεξαγωγής του μαθήματος, όπως φυσικά και από τις επιλογές του κάθε φοιτητή για το αν, πώς και πόσο θα αξιοποιήσει τη σελίδα.



**Εικόνα 3. Παράδειγμα διαδικτυακής σελίδας μαθήματος στο ΣΔΜ Moodle (έκδοση 3.9) του Πανεπιστημίου Αιγαίου**

Στη σελίδα κάθε μαθήματος, πρόσβαση έχουν οι διδάσκοντες, οι εγγεγραμμένοι στο μάθημα φοιτητές και οι διαχειριστές του συστήματος. Η είσοδος στο σύστημα γίνεται με τον ακαδημαϊκό λογαριασμό πρόσβασης, ωστόσο η πρόσβαση στις σελίδες των μαθημάτων δεν είναι ελεύθερη για όλους τους φοιτητές του ιδρύματος. Ο διαχειριστής του συστήματος δίνει δικαιώματα πρόσβασης μόνο στους φοιτητές που είναι εγγεγραμμένοι στο μάθημα. Είναι επίσης διαδεδομένη η πρακτική της «αυτό-εγγραφής» των φοιτητών, μέσω «κλειδιού» αυτό-εγγραφής, που συνήθως ο διδάσκοντας του μαθήματος κοινοποιεί στους φοιτητές. Ο περιορισμός της πρόσβασης μόνο στους εγγεγραμμένους φοιτητές εφαρμόζεται, όχι για να περιοριστεί η πρόσβαση προς το εκπαιδευτικό υλικό μόνο μεταξύ των εγγεγραμμένων φοιτητών, αλλά κυρίως για να αποφευχθεί η σύγχυση που θα προέκυπτε, αν στις εκπαιδευτικές δραστηριότητες που διεξάγονται μέσω της σελίδας (π.χ. δημόσιοι διάλογοι, διαγωνίσματα, εργασίες) συμμετείχαν μη εγγεγραμμένοι στο μάθημα φοιτητές.

Ωστόσο, σε αρκετές περιπτώσεις το σύνολο των φοιτητών που τελικά έχουν πρόσβαση στη σελίδα είναι ευρύτερο του συνόλου των εγγεγραμμένων στο μάθημα φοιτητών. Και αυτό διότι,

ο κατάλογος των φοιτητών με δικαίωμα πρόσβασης στη σελίδα ενός μαθήματος στο Moodle, δεν είναι αυτομάτως συσχετισμένος - συγχρονισμένος με τον κατάλογο φοιτητών του συστήματος φοιτητολογίου των Τμημάτων. Οπότε, οι φοιτητές που δηλώνουν στο σύστημα του φοιτητολογίου ότι θα παρακολουθήσουν κάποιο μάθημα και εγγράφονται σε αυτό, δεν αποκτούν αυτομάτως πρόσβαση στη σελίδα του μαθήματος στο Moodle. Αντιστοίχως, οι φοιτητές που για οποιοδήποτε λόγο διαγράφονται από το μάθημα δεν χάνουν αυτομάτως την πρόσβαση στη σελίδα του μαθήματος, τουλάχιστον όσο ισχύει ο λογαριασμός πρόσβασής τους. Η μη αυτόματη συσχέτιση με το σύστημα φοιτητολογίου έχει ως αποτέλεσμα να περιλαμβάνονται στο αρχείο καταγραφής δεδομένων χρήσης (log file), όχι μόνο τα δεδομένα χρήσης φοιτητών που έχουν εγγραφεί στο μάθημα, αλλά και τα δεδομένα χρήσης φοιτητών που εγγράφτηκαν, αλλά τελικά διαγράφτηκαν από το μάθημα, όπως και δεδομένα χρήσης φοιτητών που ποτέ δεν εγγράφτηκαν τυπικά στο μάθημα, αλλά με κάποιο τρόπο γνώριζαν τον κωδικό αυτό-εγγραφής. Επίσης, το γεγονός ότι η πρόσβαση στη σελίδα του μαθήματος δεν τερματίζεται όταν ολοκληρώνεται το μάθημα, αλλά είναι συνεχώς προσβάσιμη, έχει ως αποτέλεσμα το αρχείο καταγραφής δεδομένων χρήσης να περιλαμβάνει και δεδομένα χρήσης που αφορούν χρονική περίοδο μεταγενέστερη από την ημερομηνία ολοκλήρωσης του μαθήματος.

#### 9.4.2. Τα αρχεία καταγραφής δεδομένων χρήσης (log files) στο ΣΔΜ Moodle

Κάθε συμβάν που προκύπτει μετά από ενέργεια χρήστη στο Moodle καταγράφεται ως διακριτή εγγραφή σε βάση δεδομένων, από την οποία είναι δυνατόν να εξαχθούν τα αρχεία καταγραφής δεδομένων χρήσης (log files). Τα είδη των δεδομένων που καταγράφονται από το σύστημα για να προσδιορίσουν κάθε συμβάν, περιλαμβάνονται στον Πίνακα 10:

**Πίνακας 10. Είδη καταγραφόμενων δεδομένων για τα συμβάντα χρηστών στο ΣΔΜ Moodle**

Είδος Δεδομένου	Περιγραφή
Ώρα	Η ημέρα και η ώρα του συμβάντος.
Πλήρες όνομα χρήστη	Το όνομα του χρήστη που επιτέλεσε το συμβάν και ειδικότερα το πλήρες όνομά του χρήστη (full name) και όχι το συστημικό όνομα χρήστη (username).
Χρήστης που επηρεάζεται	Το όνομα του χρήστη που επηρεάστηκε από το συμβάν. Συνήθως δεν έχει τιμή διότι, σπανίως, κάποιο συμβάν που επιτελείται από απλό χρήστη (φοιτητή) επηρεάζει άλλον χρήστη.
Πλαίσιο συμβάντος	Η δραστηριότητα ή ο πόρος που «στόχευσε» το συμβάν. Π.χ. «Μάθημα ΑΑ005» ή «Αρχείο File1.docx» κ.τ.λ.
Στοιχείο λογισμικού	Το υποσύστημα του Moodle που υλοποίησε το συμβάν. Π.χ. «πυρήνας συστήματος» ή «ανάθεση εργασίας» ή «αρχείο», κ.τ.λ.
Όνομα συμβάντος	Το όνομα του συμβάντος.
Περιγραφή	Σύντομη περιγραφή του συμβάντος.
Προέλευση	Το είδος της πλατφόρμας που χρησιμοποίησε ο χρήστης. Π.χ. «web» (browser) ή «ws» (smartphone app).
Διεύθυνση IP	Η διαδικτυακή διεύθυνση (IP address) της συσκευής που χρησιμοποίησε ο χρήστης για να επιτελέσει το συμβάν.

Οι διαχειριστές και οι διδάσκοντες του μαθήματος έχουν εξ αρχής δικαιώματα πρόσβασης, που τους επιτρέπουν να δουν τις εγγραφές των αρχείων με τον φυλλομετρητή που χρησιμοποιούν (Πίνακας 11). Μπορούν επίσης να αποθηκεύσουν τα αρχεία στη συσκευή τους, επιλέγοντας μεταξύ πολλών μορφών (format) αρχείου, όπως: csv, xlsx, html, json, ods, pdf.

**Πίνακας 11. Τμήμα αρχείου καταγραφής συμβάντων από το ΣΔΜ Moodle όπως εμφανίζεται σε περιβάλλον φυλλομετρητή.<sup>6</sup>**

Ώρα	Πλήρες όνομα χρήστη	Χρήστης που επηρεάζεται	Πλαίσιο συμβάντος	Στοιχείο λογισμικού	Όνομα συμβάντος	Περιγραφή	Προέλευση	Διεύθυνση IP
29/9/2021, 10:40 πμ	Όνομα1 Επίθετο1	-	AA00	Πυρήνας συστήματος	Εμφάνιση μαθήματος	The user with id '14871' viewed the course with id '3542'.	web	xxx.xxx.xxx.xxx
29/9/2021, 9:49 πμ	Όνομα2 Επίθετο2	-	AA00	Πυρήνας συστήματος	Εμφάνιση περίληψης μαθήματος	The user with id '0' viewed the course information for the course with id '3542'.	web	yyy.yyy.yyy.yyy

Επειδή τα αρχεία καταγραφής δεδομένων χρήσης περιλαμβάνουν χιλιάδες εγγραφές, το περιβάλλον διαχείρισης παρέχει δυνατότητα εφαρμογής φίλτρων, ώστε να επιλέγονται προς εμφάνιση μόνο οι εγγραφές που ενδιαφέρουν. Με την εφαρμογή φίλτρων, οι διαχειριστές και οι διδάσκοντες μπορούν κάθε φορά να επιλέγουν, προς εμφάνιση, συμβάντα συγκεκριμένων τύπων, ή/και συγκεκριμένων χρηστών, ή/και συγκεκριμένων χρονικών περιόδων. Η βάση δεδομένων για τα συμβάντα ενημερώνεται αμέσως, συνεπώς τα παραγόμενα αρχεία καταγραφής, αφενός εμφανίζουν πάντοτε τα παλαιότερα έως τα πιο πρόσφατα συμβάντα και αφετέρου συνεχώς διευρύνονται, όσο η πρόσβαση στο μάθημα παραμένει δυνατή. Στα αρχεία καταγραφής συνεχώς προστίθενται νέες εγγραφές. Δεν διαγράφονται ποτέ εγγραφές ακόμα και αν ανήκουν σε χρήστες που δεν έχουν πλέον δικαιώματα πρόσβασης στο μάθημα. Συνεπώς, τα αρχεία συχνά περιλαμβάνουν εγγραφές που σχετίζονται με χρήστες οι οποίοι στο παρελθόν είχαν πρόσβαση στη σελίδα του μαθήματος, αλλά δεν έχουν πλέον, διότι π.χ. διαγράφηκαν από το μάθημα.

<sup>6</sup> Σημείωση: τα ονόματα και οι δικτυακές διευθύνσεις των πραγματικών χρηστών έχουν αντικατασταθεί.

#### 9.4.3. Το λογισμικό αλγόριθμων Μηχανικής Μάθησης WEKA

Για να μελετήσουμε τα δεδομένα χρησιμοποιήσαμε το λογισμικό αλγόριθμων Μηχανικής Μάθησης WEKA (Waikato Environment for Knowledge Analysis) (έκδοση 3.9.6) (Weka 3: Machine Learning Software in Java, 2022), το οποίο αναπτύχθηκε από το Πανεπιστήμιο του Waikato στη Νέα Ζηλανδία. Το WEKA είναι ελεύθερο λογισμικό και διέπεται από την Γενική Δημόσια Άδεια (General Public Licence). Αποτελεί μια συλλογή από βασικούς, ευρείας χρήσης και από προηγμένους αλγόριθμους Μηχανικής Μάθησης. Επιπλέον, διαθέτει ένα πλούσιο σύνολο εργαλείων για την προετοιμασία, επεξεργασία και μελέτη των δεδομένων, όπως ταξινόμηση (classification), παλινδρόμηση (regression), ομαδοποίηση (clustering), οπτικοποίηση (visualization). Επιπλέον, προσφέρει τη δυνατότητα (μέσω του package manager) λήψης περισσότερων εργαλείων από μια πλούσια τράπεζα λογισμικού, τα οποία είναι ελευθέρως διαθέσιμα και τα οποία προορίζονται για την υλοποίηση ιδιαίτερων και περισσότερων εξειδικευμένων εργασιών. Οι εργασίες στο WEKA γίνονται μέσω εύχρηστου γραφικού περιβάλλοντος (Εικόνα 4), το οποίο επιτρέπει τη διεκπεραίωση έργων σχετικών με Μηχανική Μάθηση, χωρίς να είναι απαραίτητη η γνώση γλώσσας προγραμματισμού και η ανάπτυξη κώδικα. Εκτός από το γραφικό περιβάλλον, το WEKA μπορεί επιπλέον να χρησιμοποιηθεί με τη γραμμή εντολών (command line) και με το Java API (Application Programming Interface).

Τα βασικά υποσυστήματα του WEKA είναι α) το Explorer που αποτελεί το βασικό εργαλείο για την εισαγωγή των δεδομένων, την προεπισκόπησή τους, τη χρήση φίλτρων, την εφαρμογή αλγόριθμων Μηχανικής Μάθησης, την εξαγωγή και παρουσίαση αποτελεσμάτων και β) το Experimenter που επιτρέπει την ταυτόχρονη εφαρμογή πολλών αλγόριθμων επί πολλών συνόλων δεδομένων και τη συγκριτική παρουσίαση των αποτελεσμάτων (Witten, Frank, & Hall, 2011, σσ. 407, 505).

Το WEKA είναι αρκετά δημοφιλές, ιδιαίτέρως στην ακαδημαϊκή κοινότητα (Altalhi, Luna, Vallejo, & Ventura, 2017). Χρησιμοποιείται από ακαδημαϊκά ιδρύματα αλλά και από επιχειρήσεις (Hall, et al., 2009). Έχουν καταγραφεί περισσότερες από 12 εκατομμύρια λήψεις (downloads) του λογισμικού (2021 International Weka User Conference, 2022) και υπάρχουν περισσότερες από 18'000 επιστημονικές δημοσιεύσεις που αναφέρονται σε αυτό (Waikato Software, 2022). Είναι

από τα παλαιότερα ενεργά λογισμικά Μηχανικής Μάθησης, με έτος πρώτης διάθεσης το 1993 (Waikato Software, 2022). Υποστηρίζεται από έντυπη και ηλεκτρονική βιβλιογραφία (WEKA Literature, 2022) και από δραστήρια ψηφιακή κοινότητα χρηστών (WEKA Documentation, 2022). Η εκμάθηση χρήσης του WEKA υποστηρίζεται από βιβλίο (Witten, Frank, & Hall, 2011) και από διαδικτυακά βίντεο (Waikato, 2013). Η ιστοσελίδα του λογισμικού περιλαμβάνει ενεργή λίστα αλληλογραφίας, wiki, συχνές ερωτήσεις, αναφορές σφαλμάτων κ.α. (Weka Wiki, 2022).

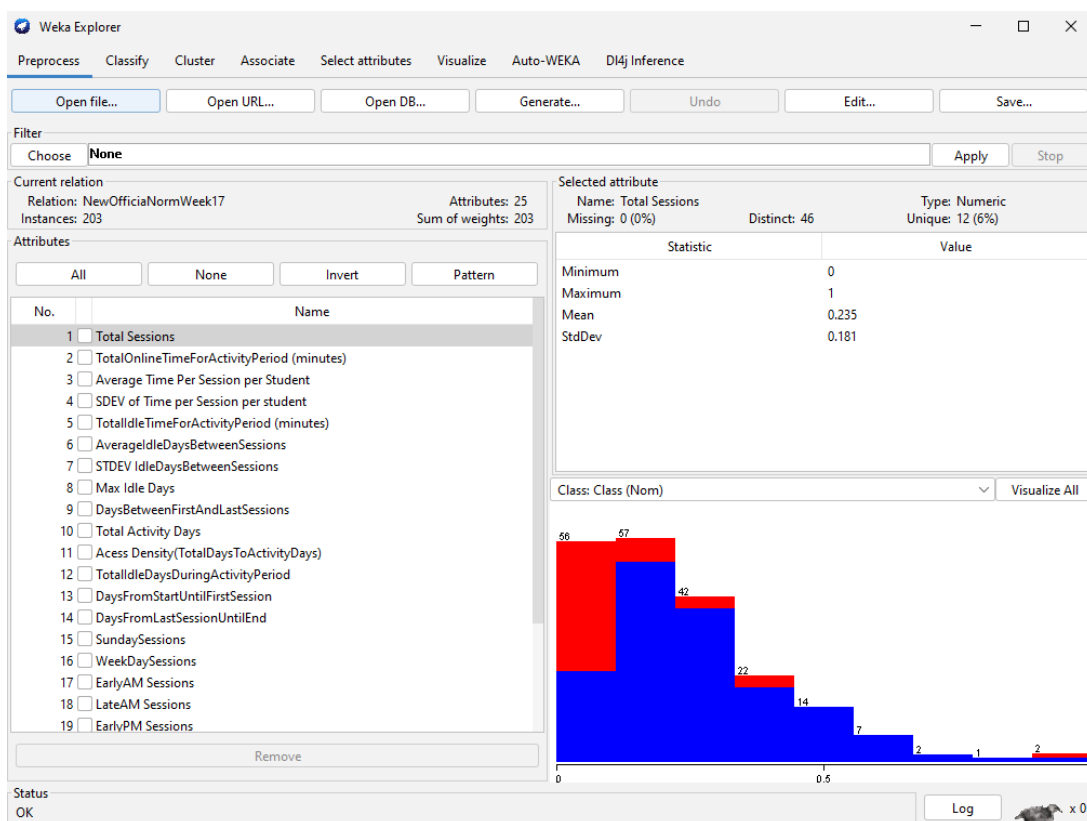
Το WEKA καταλαμβάνει υψηλή θέση σε αξιολογήσεις ελεύθερων λογισμικών Μηχανικής Μάθησης, όπου περιγράφεται ως ιδανική επιλογή για εκπαιδευτικούς και ερευνητικούς σκοπούς (Shravan, 2017), ως λογισμικό με πλούσια συλλογή αλγόριθμων και μοντέλων για τη μοντελοποίηση και πρόβλεψη διεργασιών και συσχετίσεων σε εκπαιδευτικά δεδομένα (Slater, Joksimonovic, Kovanovic, Baker, & Gasevic, 2017), ως εργαλείο που διαθέτει σε υψηλό βαθμό χαρακτηριστικά όπως απόδοση, λειτουργικότητα, ευχρηστία, συμπληρωματικές δυνατότητες, τεχνικές απαιτήσεις, επαρκές σύνολο αλγόριθμων (για ταξινόμηση, παλινδρόμηση, ομαδοποίηση, συσχετίσεις, εντοπισμό ανωμαλιών, ανάλυση χρονικών σειρών) (Altalhi, Luna, Vallejo, & Ventura, 2017).

Η απόφασή μας για την επιλογή του WEKA ως βασικού εργαλείου της έρευνας βασίστηκε στις παραπάνω αξιολογήσεις και στο ότι πρόκειται για ελεύθερο λογισμικό. Το λογισμικό αποδείχθηκε επαρκές και αποτελεσματικό για τις εργασίες που χρειάστηκε να διεκπεραιώσουμε. Αποδείχθηκε επίσης εύχρηστο, χάρη στο γραφικό περιβάλλον χρήσης.

Ωστόσο, οφείλουμε να σημειώσουμε ότι συναντήσαμε δυσκολίες στη δημιουργία και ανάλυση μοντέλων Περιοδικών Νευρωνικών Δικτύων (Recurrent Neural Networks) (Okubo, Yamashita, Shimada, & Ogata, 2017), (Chen & Cui, 2020), στα οποία κάθε χαρακτηριστικό (feature) του συνόλου δεδομένων αντιπροσώπευε λίστα χρονικών ακολουθιών και επομένως αντιστοιχούσε όχι σε μία, αλλά σε περισσότερες τιμές. Στα εν λόγω μοντέλα και ειδικότερα στον τρόπο αναπαράστασης του συνόλου δεδομένων, στη χρήση των σχετικών αλγόριθμων και στην παραμετροποίησή τους, οι σχετικές βιβλιογραφικές πηγές και τα υποστηρικτικά κείμενα ήταν ελάχιστα.



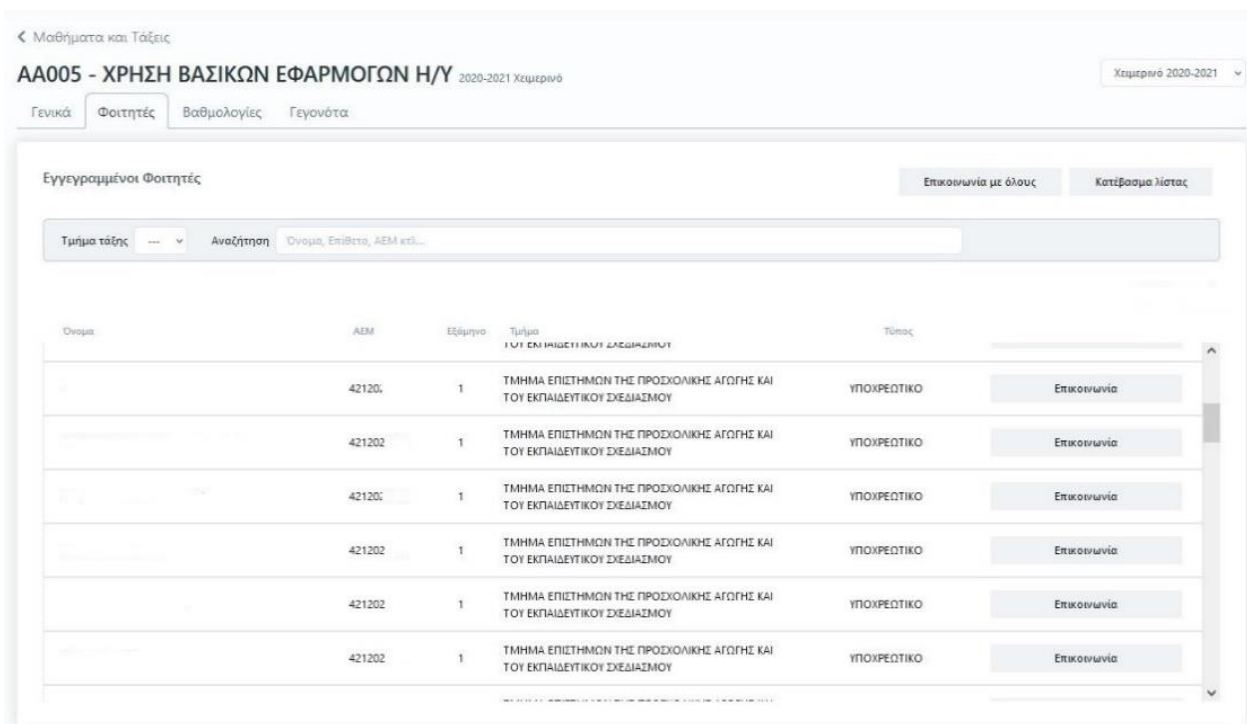
Η δημιουργία και η ανάλυση τέτοιων μοντέλων διεκπεραιώθηκε μετά από κοπιώδη διαδικτυακή αναζήτηση για παρόμοια παραδείγματα και μετά από προσπάθειες που βασίστηκαν στη μέθοδο «δοκιμής και λάθους» (trial and error). Ωστόσο, και σε αυτό το ζήτημα οφείλουμε να αναφερθούμε, αφενός στις αναρτήσεις της ενεργής λίστας αλληλογραφίας που περιλάμβανε βασικές έστω πληροφορίες και αφετέρου στην πρόθυμη απόκριση που έτυχαν τα ερωτήματα που υποβάλλαμε στους διαχειριστές της λίστας (δημιουργούς και έμπειρους χρήστες του WEKA) (Frank, 2021).



**Εικόνα 4.** Το γραφικό περιβάλλον χρήσης του WEKA (έκδοση 3.9.6)

#### 9.4.4. Το Σύστημα φοιτητολογίου του Πανεπιστημίου Αιγαίου

Το Πανεπιστήμιο Αιγαίου χρησιμοποιεί σύστημα φοιτητολογίου για τη διαχείριση των δεδομένων που αφορούν τους φοιτητές, τους διδάσκοντες και τα μαθήματα των Τμημάτων του Ιδρύματος. Από τα δεδομένα που διατηρούνται στη βάση δεδομένων του συστήματος, χρήσιμα για την έρευνά μας ήταν αυτά που αφορούσαν τους εγγεγραμμένους φοιτητές στα μαθήματα που μελετήθηκαν. Ο διδάσκων κάθε μαθήματος έχει τη δυνατότητα να δει σε περιβάλλον φυλλομετρητή (Εικόνα 5) και να αποθηκεύσει στον υπολογιστή του, κατάλογο με τα ονόματα, τους αριθμούς μητρώου και την επίδοση των εγγεγραμμένων στο μάθημα φοιτητών.



The screenshot shows a web interface for the student registration system. At the top, there is a navigation bar with the text "Μαθήματα και Τάξεις" and "ΑΑ005 - ΧΡΗΣΗ ΒΑΣΙΚΩΝ ΕΦΑΡΜΟΓΩΝ Η/Υ 2020-2021 Χειμερινό". Below this, there are tabs for "Γενικά", "Φοιτητές", "Βαθμολογίες", and "Γεγονότα". The "Φοιτητές" tab is active. The main content area is titled "Εγγεγραμμένοι Φοιτητές" and contains a search bar with a dropdown for "Τμήμα τάξης" and a search field for "Αναζήτηση". Below the search bar is a table with the following columns: "Όνομα", "ΑΕΜ", "Εξάμηνο", "Τμήμα", "Τύπος", and "Επικοινωνία". The table lists six rows of student data, all with the same course name: "ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ ΚΑΙ ΤΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ".

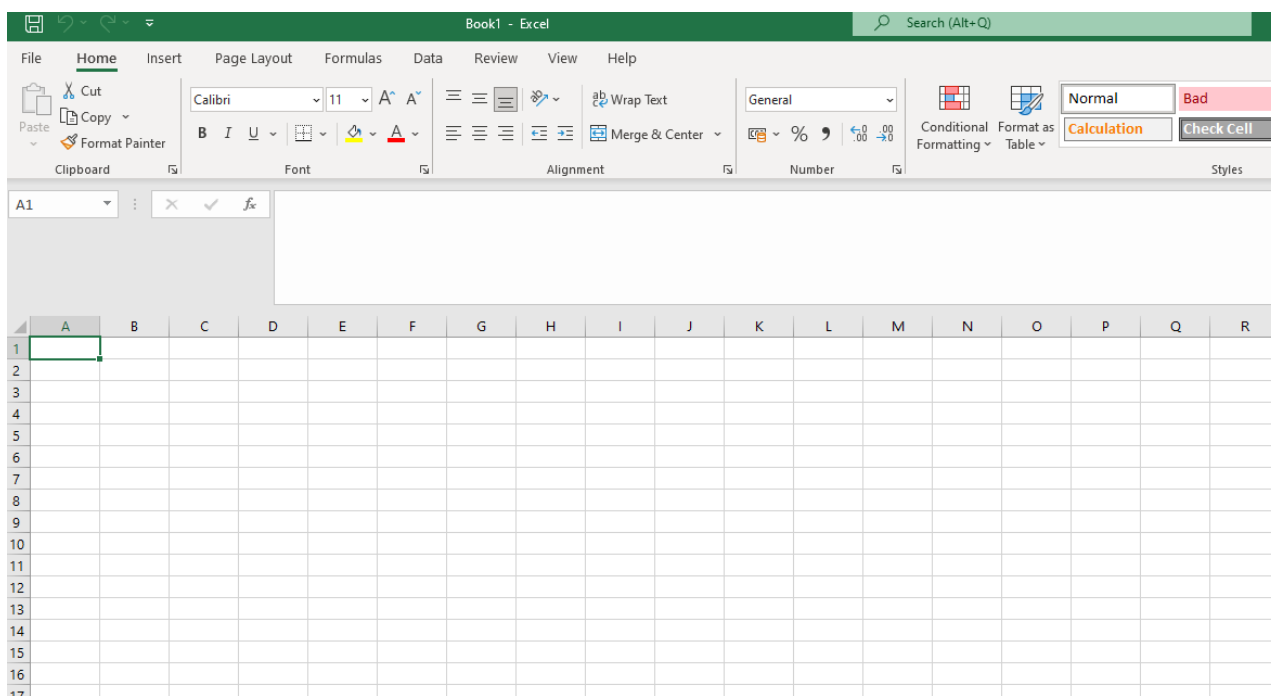
Όνομα	ΑΕΜ	Εξάμηνο	Τμήμα	Τύπος	Επικοινωνία
[Redacted]	42120	1	ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ ΚΑΙ ΤΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ	ΥΠΟΧΡΕΩΤΙΚΟ	Επικοινωνία
[Redacted]	421202	1	ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ ΚΑΙ ΤΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ	ΥΠΟΧΡΕΩΤΙΚΟ	Επικοινωνία
[Redacted]	42120	1	ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ ΚΑΙ ΤΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ	ΥΠΟΧΡΕΩΤΙΚΟ	Επικοινωνία
[Redacted]	421202	1	ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ ΚΑΙ ΤΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ	ΥΠΟΧΡΕΩΤΙΚΟ	Επικοινωνία
[Redacted]	421202	1	ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ ΚΑΙ ΤΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ	ΥΠΟΧΡΕΩΤΙΚΟ	Επικοινωνία
[Redacted]	421202	1	ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΤΗΣ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ ΚΑΙ ΤΟΥ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ	ΥΠΟΧΡΕΩΤΙΚΟ	Επικοινωνία

Εικόνα 5. Το σύστημα φοιτητολογίου του Πανεπιστημίου Αιγαίου<sup>7</sup>

<sup>7</sup> Έχουν αφαιρεθεί τα ονόματα και μέρος των αριθμών μητρώων των φοιτητών.

#### 9.4.5. Το λογισμικό υπολογιστικών φύλλων MS Excel

Για την επεξεργασία των δεδομένων που αντλήθηκαν από τις βάσεις δεδομένων του συστήματος του φοιτητολογίου και του συστήματος Moodle, ώστε να κατασκευαστούν σύνολα δεδομένων σε μορφή πινάκων και να εισαχθούν προς ανάλυση στο σύστημα WEKA, χρησιμοποιήθηκε το λογισμικό υπολογιστικών φύλλων Excel (έκδοση 2209) της εταιρείας Microsoft (Εικόνα 6). Επιπλέον, αξιοποιήθηκε η πλούσια βιβλιοθήκη συναρτήσεων του Excel για να υπολογιστούν οι τιμές όσων χαρακτηριστικών δεν προκύπτουν άμεσα από τα πρωτογενή δεδομένα των αρχείων καταγραφής του Moodle, αλλά απαιτούν σύνθετους υπολογισμούς.



**Εικόνα 6. Το περιβάλλον χρήσης του Microsoft Excel (έκδοση 2209)**

#### 9.4.6. Το λογισμικό στατιστικής ανάλυσης IBM SPSS

Ο υπολογισμός των συντελεστών συσχέτισης μεταξύ των χαρακτηριστικών των συνόλων δεδομένων και της «κλάσης», ο οποίος ήταν απαραίτητος για τον εντοπισμό των χαρακτηριστικών με την μεγαλύτερη επίδραση στην απόδοση των μοντέλων, έγινε κυρίως με το λογισμικό WEKA. Ωστόσο, η σχετική λειτουργία του WEKA δεν παρέχει οποιαδήποτε πληροφορία για τη στατιστική σημαντικότητα (significance) των υπολογιζόμενων συντελεστών συσχέτισης. Επιπλέον, πουθενά στο υλικό τεκμηρίωσης του WEKA δεν γίνεται αναφορά σε ποιο επίπεδο στατιστικής σημαντικότητας (significance level) υπολογίζονται οι συντελεστές. Οπότε, για να επιβεβαιώσουμε τις τιμές των συντελεστών συσχέτισης που υπολογίζει το WEKA και κυρίως για να ανακαλύψουμε το επίπεδο στατιστικής σημαντικότητάς τους, υπολογίσαμε τους ίδιους συντελεστές και με το λογισμικό στατιστικής ανάλυσης IBM SPSS. Διαπιστώθηκε, ότι το WEKA υπολογίζει τους συντελεστές συσχέτισης στο επίπεδο στατιστικής σημαντικότητας 0.01 (2-tailed).

Χρησιμοποιήθηκε η έκδοση 28 του λογισμικού IBM SPSS, το οποίο είναι διαθέσιμο για διαδικτυακή χρήση μέσω των πληροφοριακών συστημάτων του Πανεπιστημίου Αιγαίου (Εικόνα 7).

	TotalSessions	TotalOnlineTimeForPerSess	AverageTimePerSess	SDEVofTimePerSess	TotalIdleTimeForActiv	AverageIdleDaysBet	STDEVIdleDaysBet	MaxIdleDays	DaysBetweenFirst	TotalAccessDays	AccessDensity	TotalIdleDaysDuring	DaysFromStartUntilFirst	DaysFromLastSession	SundaySessions	WeekDaysSessions	EarlyAMSessions	LateAMSessions
1	.163842	194683	237143	324358	991218	043314	.114575	.146067	1.000000	182673	.153456	895652	0.000000	0.000000	.214286	.158537	.071429	254545
2	.231638	466771	406122	616577	985399	027662	.121900	.280899	991736	275510	.103270	808696	.011364	0.000000	.214286	.231707	0.000000	381818
3	.056497	066458	220779	222949	989531	136585	.319971	.337079	991736	081633	.321285	973913	.011364	0.000000	.071429	.054878	.071429	109091
4	.050847	015637	057143	088297	991377	151762	234965	292135	1.000000	091837	291566	973913	0.000000	0.000000	0.000000	.054878	0.000000	054545
5	.169492	120407	141935	289161	991742	040244	.132983	.224719	1.000000	224490	.126768	860870	0.000000	0.000000	.071429	.176829	0.000000	418182
6	.265537	197811	150595	229425	895686	020239	086061	.213483	900826	316327	.082078	678261	.079545	.041322	.071429	.280488	285714	363636
7	.118644	283815	471429	577079	987816	062718	.188413	.314607	991736	122449	.222428	939130	.011364	0.000000	0.000000	.128049	0.000000	381818
8	.118644	111806	185714	281333	987981	059814	.141031	.213483	991736	173469	.160643	895652	.011364	0.000000	.071429	.121951	0.000000	200000
9	.163842	144644	176190	386463	990397	042893	.140458	.235955	1.000000	193878	.145783	886957	0.000000	0.000000	.071429	.170732	.142857	345455
10	203390	312744	308880	441810	993520	032520	.149669	.303371	1.000000	255102	.112141	834783	0.000000	0.000000	.214286	.201220	.142857	309091
11	542373	204066	076878	180578	984641	007495	031781	087416	991736	622449	046838	513043	.011364	0.000000	285714	560976	0.000000	945455
12	016949	058640	535714	397653	502379	239837	695155	629213	504132	020408	489960	513043	681818	0.000000	0.000000	.018293	.071429	000000
13	.107345	318217	581429	670359	984624	068678	296483	516854	991736	132653	206540	930435	.011364	0.000000	.142857	.103659	0.000000	218182
14	.107345	066458	121429	187916	932399	064185	.195847	.314607	933884	132653	.194492	869565	.090909	0.000000	.071429	.109756	0.000000	254545
15	209040	272088	261654	343829	993030	032990	.117702	.213483	1.000000	214286	.132530	869565	0.000000	0.000000	0.000000	0.000000	0.000000	381818
16	293785	745895	514286	721399	988431	022749	.104947	.224719	1.000000	244899	.116627	843478	0.000000	0.000000	285714	292583	0.000000	454545
17	.152542	110242	143878	361731	915644	040199	.139143	.213483	925620	234694	.112450	773913	0.000000	.074380	0.000000	.164634	0.000000	345455
18	016949	000782	007143	009384	459802	215447	582319	539326	462810	030612	337349	460870	.011364	528926	0.000000	.018293	0.000000	072727
19	192090	441751	461224	493649	928404	033716	223069	528090	933884	193878	.136145	817391	.090909	0.000000	.357143	.176829	214286	218182
20	248588	294762	239365	439394	991697	026053	098564	.224719	1.000000	275510	.104131	817391	0.000000	0.000000	.071429	.262195	0.000000	436364
21	124294	125880	200000	296044	745884	042683	096334	.134831	.752066	142857	.146185	699565	.011364	239669	0.000000	.134146	0.000000	290909
22	203390	186083	183784	389325	994882	033198	.141389	.314607	1.000000	234694	.121486	852174	0.000000	0.000000	.142857	.207317	.071429	309091
23	225989	419859	374216	582490	990631	028354	.102629	.168539	1.000000	285714	.100540	808696	0.000000	0.000000	.357143	.213415	0.000000	381818

Εικόνα 7. Το περιβάλλον χρήσης του IBM SPSS (έκδοση 28)

## 9.5. Δημιουργία συνόλων δεδομένων

Για να μελετηθούν τα δεδομένα χρήσης των φοιτητών είναι απαραίτητη η δημιουργία αρχείων με τα σύνολα δεδομένων (datasets). Τα σύνολα έχουν μορφοποίηση διδιάστατου πίνακα, όπου κάθε γραμμή (σειρά) αντιστοιχεί σε έναν μόνο φοιτητή και κάθε στήλη αντιστοιχεί σε ένα χαρακτηριστικό / ιδιότητα (feature / attribute)<sup>8</sup>, το οποίο σχετίζεται με τη χρήση του συστήματος από τον φοιτητή. Τα αρχεία καταγραφής συστήματος δεν μπορούν ως έχουν, δηλαδή με την αρχική μορφή τους, να χρησιμοποιηθούν ως σύνολα δεδομένων, για πολλούς λόγους. Συγκεκριμένα, α) τα αρχεία καταγραφής περιλαμβάνουν πολλές εγγραφές για κάθε φοιτητή (όσες και οι ενέργειες που πραγματοποίησε στο σύστημα), ενώ στο σύνολο δεδομένων πρέπει υπάρχει μόνο μία γραμμή (σειρά) για κάθε φοιτητή, β) τα αρχεία καταγραφής δεν υπολογίζουν τις τιμές χαρακτηριστικών (όπως π.χ. συνολικός αριθμός ημερών πρόσβασης), οι οποίες είναι απαραίτητες για τη μελέτη της χρήσης του συστήματος από τους φοιτητές, γ) στα αρχεία καταγραφής το όνομα του φοιτητή είναι σε μορφή, που δεν είναι ταυτόσημη με τη μορφή του ονόματος στο φοιτητολόγιο δ) τα αρχεία καταγραφής περιλαμβάνουν δεδομένα, που ανάλογα με τους στόχους της έρευνας, άλλοτε έχουν ενδιαφέρουν και άλλοτε όχι (π.χ. δικτυακή διεύθυνση συσκευής χρήστη).

### 9.5.1. Αναπαράσταση συνόλου δεδομένων

Το σύνολο δεδομένων (έστω  $S$ ) που κατασκευάστηκε για κάθε μάθημα, περιγράφεται με τη βοήθεια αλγεβρικού φορμαλισμού, ως εξής (Hussain M. , Zhu, Zhang, & Abidi, 2018):

- Το  $S = \{X_i, Y_i\}, i \in [1, K]$  είναι το σύνολο δεδομένων. Αποτελείται από  $K$  στοιχεία (τα οποία στην μορφή του πίνακα αντιστοιχούν στις γραμμές), όπου  $K$  το πλήθος των οντοτήτων (φοιτητών).
- Το  $X_i$  είναι  $N$ -διάστατο διάνυσμα πραγματικών αριθμών, όπου το  $N$  είναι το πλήθος των χαρακτηριστικών / ιδιοτήτων (features) των οντοτήτων (φοιτητών). Σε μορφή διδιάστατου πίνακα (Πίνακας 12), το κάθε στοιχείο του διανύσματος (δηλαδή κάθε

---

<sup>8</sup> Στα πλαίσια της εργασίας χρησιμοποιούμε χωρίς διάκριση τους όρους «χαρακτηριστικό» (feature) και «ιδιότητα» (attribute). Το ίδιο εν γένει συμβαίνει στο πεδίο της Μηχανικής Μάθησης (Han, Kamber, & Pei, 2011, σ. 363), (Witten, Frank, & Hall, 2011, σ. 49), (Machine Learning Glossary, n.d.)

χαρακτηριστικό) αντιστοιχεί σε μία στήλη του πίνακα. (Σημείωση: η τελευταία στήλη του πίνακα είναι δεσμευμένη για την κατηγορία / κλάση της οντότητας).

- Το  $Y_i \in \{Yes, No\}$  είναι η κατηγορία / κλάση (class), όπου κατηγοριοποιείται κάθε οντότητα. Σε μορφή διδιάστατου πίνακα αντιστοιχεί στην τελευταία στήλη του. Για το μάθημα «ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ» έχει τον τίτλο «Επιτυχία» (Pass) και λαμβάνει τιμή «Yes» για τους φοιτητές που «πέρασαν» το μάθημα και «No» για τους υπόλοιπους. Για το μάθημα «Α3 Ηλεκτρονική Μάθηση», έχει τον τίτλο «Εύγε» (Well Done) και λαμβάνει τιμή «Yes» για τους φοιτητές που είχαν επίδοση μεγαλύτερη του 8,5 (στη δεκαβάθμια κλίμακα) και «No» για τους υπόλοιπους.

Το σύνολο δεδομένων  $S$  αναπαρίσταται ως διδιάστατος πίνακας διαστάσεων  $K \times (N+1)$ . Ειδικότερα, ο πίνακας αποτελείται από  $K$  γραμμές όσες και το πλήθος το οντοτήτων και  $N+1$  στήλες. Οι πρώτες  $N$  στήλες αντιστοιχούν στα χαρακτηριστικά / ιδιότητες της οντότητας και η τελευταία στήλη αντιστοιχεί στην κατηγορία / κλάση της οντότητας.

**Πίνακας 12. Παράδειγμα συνόλου δεδομένων με μορφή διδιάστατου πίνακα ( $K$  γραμμές,  $N+1$  Στήλες)**

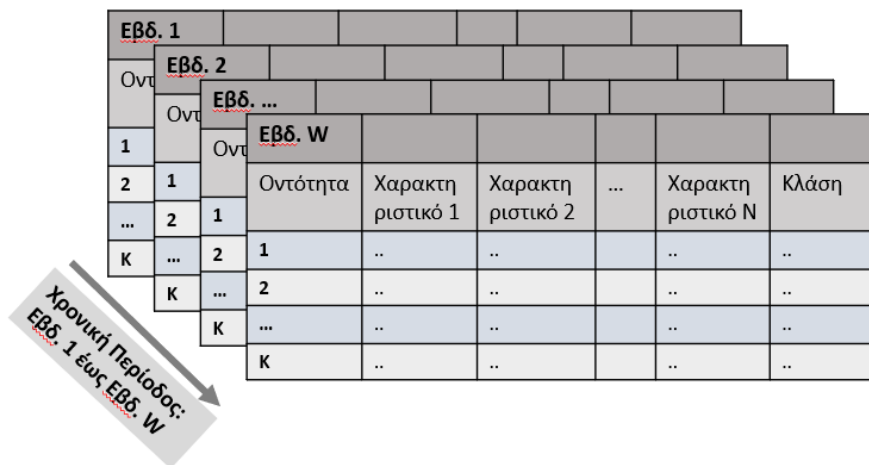
	Χαρακτηριστικό 1	Χαρακτηριστικό 2	...	Χαρακτηριστικό N	Κλάση
Οντότητα 1	τιμή $\in R$	τιμή $\in R$	...	τιμή $\in R$	τιμή $\in \{Yes, No\}$
Οντότητα 2	τιμή $\in R$	τιμή $\in R$	...	τιμή $\in R$	τιμή $\in \{Yes, No\}$
...	...	...	...	...	...
Οντότητα K	τιμή $\in R$	τιμή $\in R$		τιμή $\in R$	τιμή $\in \{Yes, No\}$

Ωστόσο για κάθε μάθημα, δεν περιοριστήκαμε στη δημιουργία μόνο ενός βασικού συνόλου δεδομένων, που καλύπτει όλες τις εβδομάδες από την αρχή έως το τέλος του μαθήματος. Προκειμένου να μελετήσουμε διαφοροποιήσεις που συνέβησαν σε χρονικές στιγμές πριν την ολοκλήρωση του μαθήματος, και προκειμένου να μελετήσουμε τη χρονική στιγμή της πρώιμης πρόβλεψης, δημιουργήσαμε περισσότερα σύνολα δεδομένων, που κάλυπταν μικρότερες χρονικές περιόδους. Δηλαδή, για κάθε μάθημα, εκτός από το βασικό, πλήρες σύνολο δεδομένων που καλύπτει όλες τις εβδομάδες του μαθήματος, δημιουργήθηκαν επιπλέον,  $M-1$  στο πλήθος, σύνολα δεδομένων (όπου  $M$  το πλήθος των εβδομάδων ως την τελική αξιολόγηση του

μαθήματος). Δηλαδή, δημιουργήθηκε σύνολο για την εβδομάδα 1, σύνολο για την εβδομάδα 2 κ.ο.κ. Με τον τρόπο αυτό δημιουργήθηκαν συνολικά  $M$  σύνολα δεδομένων για κάθε μάθημα.

Επεκτείναμε τη μελέτη μας και πέραν των βασικών αλγόριθμων Μηχανικής Μάθησης, που μελετούν διδιάστατα σύνολα δεδομένων. Μελετήσαμε τεχνικές «βαθιάς μάθησης» (deep learning) με αλγόριθμους που βασίζονται σε RNN δίκτυα (Recurrent Neural Networks / Περιοδικά Νευρωνικά Δίκτυα) και ειδικότερα σε LSTM δίκτυα (Μακρά Βραχείας Μνήμης Δίκτυα / Long Short-Term Memory Networks) (Okubo, Yamashita, Shimada, & Ogata, 2017), (Chen & Cui, 2020) (βλ. Κεφ. 5.3.1). Ειδικότερα, επιδιώξαμε να διαπιστώσουμε, αν τα συμπεράσματα της έρευνάς μας θα συμφωνούσαν με τα συμπεράσματα ερευνών, οι οποίες αναφέρουν ότι οι τεχνικές «βαθιάς μάθησης» επί εκπαιδευτικών δεδομένων είναι αποδοτικότερες από τις βασικές τεχνικές Μηχανικής Μάθησης, ή θα συμφωνούσαν με τα συμπεράσματα ερευνών που αναφέρουν το αντίθετο. Η μελέτη των Περιοδικών Νευρωνικών Δικτύων (Recurrent Neural Networks) οδήγησε στην ανάγκη για δημιουργία συνόλων δεδομένων με πιο σύνθετη μορφή, η οποία είναι ικανή να αποτυπώσει το γεγονός, ότι, οι τιμές των χαρακτηριστικών σωρεύονται και μεταβάλλονται με την πάροδο του χρόνου. Δηλαδή, προέκυψε ανάγκη για δημιουργία συνόλων δεδομένων με μορφή που πρέπει να περιλαμβάνει και τη διάσταση του χρόνου. Ειδικότερα, για κάθε οντότητα, έπρεπε να καταγραφούν, όχι μόνο οι τιμές όλων των χαρακτηριστικών όπως διαμορφώθηκαν στο τέλος της χρονικής περιόδου που μελετούμε, αλλά στο ίδιο σύνολο δεδομένων έπρεπε να καταγραφούν επίσης, οι τιμές όλων των χαρακτηριστικών στο τέλος κάθε εβδομάδας που περιλαμβάνεται στη χρονική περίοδο που καλύπτει το σύνολο (Σχήμα 12). Για παράδειγμα, το σύνθετο σύνολο δεδομένων που καλύπτει τη χρονική περίοδο μέχρι και την εβδομάδα  $W$  περιέχει, όχι μόνο τις τιμές των χαρακτηριστικών στο τέλος της εβδομάδας  $W$ , αλλά επιπλέον και τις τιμές στο τέλος των εβδομάδων που προηγήθηκαν. Με τον τρόπο αυτό, για κάθε μάθημα δημιουργήθηκαν  $W-1$  σύνθετα σύνολα δεδομένων και συγκεκριμένα για τα μαθήματα που μελετήσαμε δημιουργήθηκαν επιπλέον 67 σύνολα δεδομένων.

(Σημείωση: η δημιουργία σύνθετων συνόλων δεδομένων για την  $W^{\text{η}}$  εβδομάδα είναι άνευ νοήματος αν  $W=1$ , διότι τότε το σύνθετο σύνολο και το απλό σύνολο της  $1^{\text{ης}}$  εβδομάδας ταυτίζονται).



**Σχήμα 12. Αναπαράσταση τριδιάστατου συνόλου δεδομένων με αποτύπωση της χρονικής διάστασης**

Το σύνολο δεδομένων  $S$  που κατασκευάζουμε για να μελετηθεί από τα Περιοδικά Νευρωνικά Δίκτυα διαμορφώθηκε ώστε να είναι συμβατό με το εργαλείο WEKA που χρησιμοποιούμε (Frank, Converting univariate time series data for time series classification with WEKA, 2017). Η περιγραφή ενός τέτοιου συνόλου δεδομένων με τη βοήθεια αλγεβρικού φορμαλισμού έχει ως εξής:

- Το  $S = \{X_{ij}, Y_i\}$ ,  $i \in [1, K]$ ,  $j \in [1, W]$  είναι το σύνολο δεδομένων για τις εβδομάδες από την 1<sup>η</sup> έως την  $W^{\eta}$ . Αποτελείται από  $K$  στοιχεία (τα οποία στην μορφή του πίνακα αντιστοιχούν στις γραμμές), όπου  $K$  το πλήθος των οντοτήτων (φοιτητών). Περιλαμβάνει δεδομένα από τις  $W$  στο πλήθος εβδομάδες, που απαρτίζουν τη χρονική περίοδο που μελετούμε.
- Το  $X_{ij}$  είναι  $M$ -διάστατο διάνυσμα πραγματικών αριθμών, όπου το  $M$  είναι το γινόμενο του πλήθους ( $N$ ) των χαρακτηριστικών/ ιδιοτήτων (features) των οντοτήτων, επί του πλήθους ( $W$ ) των εβδομάδων που περιλαμβάνει η χρονική περίοδος που μελετούμε. Σε μορφή διδιάστατου πίνακα (Πίνακας 13), το κάθε στοιχείο του διανύσματος (δηλαδή κάθε χαρακτηριστικού, όπως έχει διαμορφωθεί κάποια εβδομάδα) αντιστοιχεί σε μία στήλη του πίνακα. (Σημείωση: η τελευταία στήλη του πίνακα είναι δεσμευμένη για την κατηγορία / κλάση της οντότητας).



- Το  $Y_i \in \{Yes, No\}$  είναι η κατηγορία / κλάση (class) όπου κατηγοριοποιείται κάθε οντότητα. Σε μορφή διδιάστατου πίνακα αντιστοιχεί στην τελευταία στήλη του. Για το μάθημα AA003 Χρήση Βασικών Εφαρμογών ΗΥ έχει τον τίτλο «Επιτυχία» (Pass) και λαμβάνει τιμή «Yes» για τους φοιτητές που «πέρασαν» το μάθημα και «No» για το υπόλοιπους. Για το μάθημα A3 Ηλεκτρονική Μάθηση, έχει τον τίτλο «Εύγε» (Well Done) και λαμβάνει τιμή «Yes» για τους φοιτητές που είχαν επίδοση μεγαλύτερη του 8,5 (στη δεκαβάθμια κλίμακα) και «No» για τους υπόλοιπους.

**Πίνακας 13. Παράδειγμα συνόλου δεδομένων με μορφή διδιάστατου πίνακα ( K γραμμές, (WxN)+1 Στήλες) με ενσωμάτωση της διάστασης του χρόνου για μελέτη με Περιοδικά Νευρωνικά Δίκτυα**

	Εβδομάδα 1			...	Εβδομάδα W			Κλάση
	Χαρακτηρ. 1	...	Χαρακτηρ. N		...	Χαρακτηρ. 1	...	
Οντότητα 1	τιμή $\in R$	...	τιμή $\in R$	...	τιμή $\in R$	...	τιμή $\in R$	τιμή $\in \{Yes, No\}$
Οντότητα 2	τιμή $\in R$	...	τιμή $\in R$	...	τιμή $\in R$	...	τιμή $\in R$	τιμή $\in \{Yes, No\}$
...	...	...	...	...	...	...	...	...
Οντότητα K	τιμή $\in R$	...	τιμή $\in R$	...	τιμή $\in R$	...	τιμή $\in R$	τιμή $\in \{Yes, No\}$

Για τη μελέτη των δεδομένων χρήσης των φοιτητών, αξιοποιήσαμε τα αρχεία καταγραφής και δημιουργήσαμε *σύνολα δεδομένων* για κάθε μάθημα, ακολουθώντας τη διαδικασία που περιγράφεται στις επόμενες παραγράφους.

#### 9.5.2. Άντληση δεδομένων, αποθήκευση σε αρχείο.

Τα αρχεία καταγραφής δεδομένων χρήσης προβλήθηκαν καταρχάς σε παράθυρο του φυλλομετρητή, μέσω συγκεκριμένης ακολουθίας εντολών στο περιβάλλον του Moodle (Διαχείριση Μαθήματος > Αναφορές > Καταγραφές). Δεν εφαρμόστηκε οποιοδήποτε φίλτρο, οπότε αντλήθηκαν όλα τα δεδομένα χρήσης, όλων των χρηστών, για όλες τις ημέρες, για όλες τις δραστηριότητες, για όλες τις ενέργειες, από όλες τις πηγές, για όλα τα συμβάντα. Οι εγγραφές που προέκυψαν και προβλήθηκαν στο παράθυρο του φυλλομετρητή, αποθηκεύτηκαν σε αποθηκευτικό μέσο (δίσκο) προσωπικού υπολογιστή με μορφή αρχείου .xlsx. Η συγκεκριμένη μορφή αρχείου επιλέχθηκε, διότι είναι εύκολα επεξεργάσιμη από την γνωστή εφαρμογή

Microsoft Excel, η οποία παρέχει δυνατότητες διαγραφών, ταξινομήσεων, αντιμεταθέσεων στηλών και γραμμών, προσωρινής απόκρυψης και επανεμφάνισης δεδομένων και χρήσης αλγεβρικών συναρτήσεων (functions), οι οποίες υπολογίζουν τιμές για χαρακτηριστικά που δεν υφίστανται πρωτογενώς στις εγγραφές της βάσης δεδομένων χρήσης.

Το πλήθος των γραμμών στα .xlsx αρχεία που προέκυψαν για κάθε μάθημα περιγράφονται στον Πίνακα 14.

**Πίνακας 14. Πλήθος γραμμών ανά αρχείο καταγραφής**

Μάθημα	Έτος	Πλήθος γραμμών	Ημερομηνία ανάκτησης αρχείου
AA005 Χρήση Βασικών Εφαρμογών ΗΥ	2020 - 2021	34332	22/2/2021
AA005 Χρήση Βασικών Εφαρμογών ΗΥ	2021 - 2022	28504	25/7/2022
A3 Ηλεκτρονική Μάθηση	2019 - 2020	33414	14/2/2022
A3 Ηλεκτρονική Μάθηση	2020 - 2021	37168	8/2/2022
Σύνολο Γραμμών		<b>133418</b>	

Οι εγγραφές των αρχείων υπέστησαν την παρακάτω επεξεργασία:

Διαγραφή χρηστών:

- Διαγράψαμε τις εγγραφές των μη φοιτητών (διδασκόντων και διαχειριστών μαθήματος).
- Διαγράψαμε τις εγγραφές των φοιτητών που δεν συναίνεσαν στη μελέτη των δεδομένων χρήσης τους.
- Διαγράψαμε τις εγγραφές των φοιτητών που δεν ήταν ενεργοί κατά την ημερομηνία εξέτασης του μαθήματος. (Σημείωση: διατηρήσαμε τις εγγραφές φοιτητών που διαγράφηκαν από το μάθημα μετά την ημερομηνία τελικής εξέτασης του μαθήματος).
- Διαγράψαμε τις εγγραφές των φοιτητών του μαθήματος «AA005 Χρήση Βασικών Εφαρμογών ΗΥ», οι οποίοι εισήχθησαν στο Τμήμα μέσω κατατακτηρίων εξετάσεων λίγες μέρες προ της τελικής εξέτασης, διότι θεωρήσαμε ότι, λόγω της καθυστερημένης εγγραφής τους στο μάθημα, δεν είχαν τη δυνατότητα να κάνουν επαρκή χρήση του Moodle.

Διαγραφή δεδομένων χρήσης μη σχετικών με την έρευνά μας:

- Διαγράψαμε, τις στήλες «Χρήστης που επηρεάζεται», «Περιγραφή», «Προέλευση», «Διεύθυνση IP», διότι δε σχετίζονται με την έρευνά μας.

Διαγραφή δεδομένων εκτός χρονικής περιόδου του μαθήματος:

- Διαγράψαμε δεδομένα που αφορούν χρονικές περιόδους πριν από την εβδομάδα έναρξης και μετά από την εβδομάδα λήξης του ακαδημαϊκού εξαμήνου. Ειδικότερα, για το μάθημα «ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ», μελετήσαμε δεδομένα ως την προηγούμενη μέρα από την τελική εξέταση του μαθήματος. Δεν συμπεριλάβαμε δεδομένα από συμβάντα κατά την ημέρα της τελικής εξέτασης, διότι υπήρχαν φοιτητές, που ενώ είχαν πρόωρα (πριν τη λήξη του εξαμήνου) διακόψει τη χρήση του Moodle, χρησιμοποίησαν το Moodle για να συμμετάσχουν στην τελική εξέταση. Τυχόν συμπερίληψη δεδομένων από την ημέρα της τελικής εξέτασης θα δυσχέραινε τη διάκριση μεταξύ όσων χρησιμοποιούσαν με συνέπεια το Moodle μέχρι την ολοκλήρωση του μαθήματος και όσων διέκοψαν πρόωρα τη χρήση και το χρησιμοποίησαν ξανά την ημέρα της εξέτασης.

Οι χρονικές περίοδοι στις οποίες ανήκουν τα δεδομένα που μελετήσαμε για κάθε μάθημα, καταγράφονται στον Πίνακα 15.

*Πίνακας 15. Χρονική περίοδος και πλήθος εβδομάδων για τα σύνολα δεδομένων*

Μάθημα	Έτος	Περίοδος μελέτης δεδομένων	Πλήθος εβδομάδων
ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ	2020 - 2021	12/10/2020 – 10/2/2021	18
ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ	2021 - 2022	11/10/2021 – 1/2/2022	17
Α3 Ηλεκτρονική Μάθηση	2019 - 2020	21/10/2019 – 23/2/2020	18
Α3 Ηλεκτρονική Μάθηση	2020 - 2021	19/10/2020 – 21/2/2021	18

Οι ημερομηνίες που ορίζουν την έναρξη και τη λήξη των εβδομάδων στις οποίες αντιστοιχούν τα σύνολα δεδομένων εμφανίζονται στον Πίνακα 16. Οι ημερομηνίες της πρώτης διάλεξης, της τελευταίας διάλεξης και της τελικής εξέτασης εμφανίζονται στον Πίνακα 17.

**Πίνακας 16. Σύνολα δεδομένων. Έναρξη και λήξη εβδομάδων**

Εβδομάδα	ΑΑ005 2020 - 2021		ΑΑ005 2021 - 2022		Α3 2019 - 2020		Α3 2020 - 2021	
	Έναρξη	Λήξη	Έναρξη	Λήξη	Έναρξη	Λήξη	Έναρξη	Λήξη
1	12/10/2020	18/10/2020	11/10/2021	17/10/2021	21/10/2019	27/10/2019	19/10/2020	25/10/2020
2	19/10/2020	25/10/2020	18/10/2021	24/10/2021	28/10/2019	03/11/2019	26/10/2020	01/11/2020
3	26/10/2020	01/11/2020	25/10/2021	31/10/2021	04/11/2019	10/11/2019	02/11/2020	08/11/2020
4	02/11/2020	08/11/2020	01/11/2021	07/11/2021	11/11/2019	17/11/2019	09/11/2020	15/11/2020
5	09/11/2020	15/11/2020	08/11/2021	14/11/2021	18/11/2019	24/11/2019	16/11/2020	22/11/2020
6	16/11/2020	22/11/2020	15/11/2021	21/11/2021	25/11/2019	01/12/2019	23/11/2020	29/11/2020
7	23/11/2020	29/11/2020	19/11/2021	25/11/2021	02/12/2019	08/12/2019	30/11/2020	06/12/2020
8	30/11/2020	06/12/2020	29/11/2021	05/12/2021	09/12/2019	15/12/2019	07/12/2020	13/12/2020
9	07/12/2020	13/12/2020	06/12/2021	12/12/2021	16/12/2019	22/12/2019	14/12/2020	20/12/2020
10	14/12/2020	20/12/2020	13/12/2021	19/12/2021	23/12/2019	29/12/2019	21/12/2020	27/12/2020
11	21/12/2020	27/12/2020	20/12/2021	26/12/2021	30/12/2019	05/01/2020	28/12/2020	03/01/2021
12	28/12/2020	03/01/2021	27/12/2021	02/01/2022	06/01/2020	12/01/2020	04/01/2021	10/01/2021
13	04/01/2021	10/01/2021	03/01/2022	09/01/2022	13/01/2020	19/01/2020	11/01/2021	17/01/2021
14	11/01/2021	17/01/2021	10/01/2022	16/01/2022	20/01/2020	26/01/2020	18/01/2021	24/01/2021
15	18/01/2021	24/01/2021	17/01/2022	23/01/2022	27/01/2020	02/02/2020	25/01/2021	31/01/2021
16	25/01/2021	31/01/2021	24/01/2022	30/01/2022	03/02/2020	09/02/2020	01/02/2021	07/02/2021
17	01/02/2021	07/02/2021	31/01/2022	01/02/2022	10/02/2020	16/02/2020	08/02/2021	14/02/2021
18	08/02/2021	10/02/2021			17/02/2020	23/02/2020	15/02/2021	21/02/2021

**Πίνακας 17. Ημερομηνίες διαλέξεων και τελικής εξέτασης**

	ΑΑ005 2020 - 2021	ΑΑ005 2021 - 2022	Α3 2019 - 2020	Α3 2020 - 2021
Πρώτη Διάλεξη	Τρίτη, 13/10/2020 <sup>9</sup>	Παρασκευή, 15/10/2021	Δευτέρα, 21/10/2019	Δευτέρα, 19/10/2021
Τελευταία Διάλεξη	Τρίτη, 12/01/2021 (14 <sup>η</sup> εβδομάδα)	Παρασκευή, 21/01/2022 (15 <sup>η</sup> εβδομάδα)	Δευτέρα, 13/01/2020 (13 <sup>η</sup> εβδομάδα)	Δευτέρα, 11/01/2021 (13 <sup>η</sup> εβδομάδα)
Τελική Εξέταση	Πέμπτη, 11/2/2021 (18 <sup>η</sup> εβδομάδα)	Τετάρτη, 02/02/2022 (17 <sup>η</sup> εβδομάδα)	Κυριακή, 23/02/2020 <sup>10</sup> (18 <sup>η</sup> εβδομάδα)	Κυριακή, 21/02/2021 <sup>11</sup> (18 <sup>η</sup> εβδομάδα)

Συνολικά δημιουργήθηκαν και μελετήθηκαν 103 σύνολα δεδομένων όπως αναφέρεται στον Πίνακα 18.

<sup>9</sup> Με απόφαση της Συγκλήτου, τα μαθήματα του πρώτους έτους ξεκίνησαν μία εβδομάδα μετά από τα μαθήματα των υπόλοιπων ετών.

<sup>10</sup> Παράδοση εργασίας.

<sup>11</sup> Παράδοση εργασίας.

**Πίνακας 18. Πλήθος συνόλων δεδομένων**

Μάθημα	Έτος	Σύνολα Δεδομένων		
		Βασικοί Αλγόριθμοι	Περιοδικά Νευρωνικά Δίκτυα	Σύνολο
ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ	2020 - 2021	18	17	35
	2021 - 2022	17	16	33
A3 Ηλεκτρονική Μάθηση	2019 - 2020 2020 - 2021	18	17	35
ΣΥΝΟΛΟ		53	50	<b>103</b>

Σημειώνουμε ότι στο μάθημα «Α3 Ηλεκτρονική Μάθηση» εγγράφηκαν 39 φοιτητές για το έτος 2019 -2020 και 39 για το έτος 2020 - 2021. Ο μικρός αριθμός των φοιτητών ανά έτος (39) και το μικρό σύνολο δεδομένων που προέκυψε δεν ενδείκνυνται για μελέτη με τεχνικές Μηχανικής Μάθησης. Οπότε, συγχωνεύσαμε τα σύνολα δεδομένων των δύο ετών και δημιουργήσαμε σύνολα δεδομένων που περιλαμβάνουν δεδομένα από τους 78 φοιτητές και των δύο ετών. Θεωρούμε ότι η συγκεκριμένη επιλογή είναι έγκυρη, δεδομένου ότι οι φοιτητές προέρχονται από το ίδιο μάθημα, το οποίο κατά τη διάρκεια των δύο ετών προσφέρθηκε με το ίδιο πρόγραμμα σπουδών, με το ίδιο εκπαιδευτικό υλικό, με τον ίδιο τρόπο αξιολόγησης, με τους ίδιους διδάσκοντες και με αντίστοιχο χρονοδιάγραμμα υλοποίησης. Η μόνη διαφορά μεταξύ των δύο ετών είναι ότι κατά το έτος 2020-2021, λόγω των περιοριστικών μέτρων κατά την πανδημία του κορονοϊού Covid-19, δύο δια ζώσης συνεδρίες υλοποιήθηκαν μέσω συστήματος τηλεδιάσκεψης.

Η συγχώνευση των εγγραφών δύο, ή και περισσότερων, μαθημάτων αποτελεί μεθοδολογική επιλογή που απαντάται στην βιβλιογραφία. Μάλιστα, σε αρκετές περιπτώσεις, τα μαθήματα δεν είναι καν τα ίδια ή έστω παρόμοια. Για παράδειγμα οι Quinn & Gray (2020), λόγω του μικρού πλήθους των εγγεγραμμένων φοιτητών, ενέταξαν στο ίδιο σύνολο δεδομένων 29 μαθήματα που διέφεραν και στο αντικείμενο και στη χρονική διάρκεια υλοποίησης, γεγονός που όπως αναφέρουν μπορεί να επηρεάσει την ακρίβεια της πρόβλεψης. Ωστόσο, θεωρούν πως αρκεί ότι προσφέρονται από τον ίδιο διδάσκοντα και ανήκουν στο ίδιο επιστημονικό πεδίο. Πολύ περισσότερο ανομοιογενή είναι τα σύνολα δεδομένων των Riestra-Gonzalez, del Puerto Paule-

Ruiz & Ortin (2021), οι οποίοι συγχώνευσαν δεδομένα από όλα τα μαθήματα (5112 στο πλήθος) που προσφέρθηκαν το έτος 2017-2018 στο Πανεπιστήμιο του Οβιέδο στην Ισπανία. Πρόκειται για δια ζώσης, διαδικτυακά και μεικτά μαθήματα, που προσφέρθηκαν στα πλαίσια προπτυχιακών και μεταπτυχιακών κύκλων σπουδών και σχετίζονται με πολλά γνωστικά αντικείμενα. Όπως, οι ίδιοι αναφέρουν τα μαθήματα διαφέρουν στο γνωστικό αντικείμενο, στον τρόπο διδασκαλίας, στη διάρκεια, στη μεθοδολογία. Ωστόσο, επέλεξαν να τα μελετήσουν συνολικά, με σκοπό να κατασκευάσουν γενικά μοντέλα, τα οποία θα είναι κατάλληλα για όλα τα μαθήματα και των οποίων η ακρίβεια δεν θα επηρεάζεται από τις ιδιαιτερότητες κάθε μαθήματος.

### 9.5.3. Συσχέτιση εγγραφών αρχείων καταγραφής με εγγραφές φοιτητολογίου

Δεδομένου ότι η έρευνά μας μελετά μεταξύ άλλων και την τελική επίδοση των φοιτητών είναι αναγκαία η άντληση σχετικών δεδομένων από το φοιτητολόγιο του Τμήματος, όπου είναι καταχωρημένες οι επιδόσεις των φοιτητών. Όμως, η μορφή με την οποία εμφανίζεται το όνομα κάθε φοιτητή στο αρχείο καταγραφής του Moodle δεν μπορεί να αντιστοιχηθεί άμεσα με τη μορφή του ονόματος στο φοιτητολόγιο. Ειδικότερα, στο αρχείο καταγραφής του Moodle, το όνομα του φοιτητή εμφανίζεται με λατινική γραφή και με πεζά - πλην του αρχικού - γράμματα, όπως ακριβώς έχει αποθηκευτεί στη βάση δεδομένων (Active Directory) του συστήματος που διαχειρίζεται την πρόσβαση σε όλα τα ψηφιακά συστήματα του Πανεπιστημίου Αιγαίου. Ας σημειωθεί επίσης, ότι συχνά, στο όνομα του φοιτητή, εξαιτίας αστοχιών, έχουν περιληφθεί χαρακτήρες, όπως διπλά κενά, παύλες (-), κάθετοι (/), οι οποίοι δεν υπάρχουν στο όνομα του φοιτητή, όπως αυτό εμφανίζεται στην ταυτότητά του. Επιπλέον, εκτός από το όνομα φοιτητή, το αρχείο καταγραφής του Moodle δεν περιλαμβάνει άλλα χαρακτηριστικά (όπως π.χ. Αριθμός Μητρώου, Όνομα Χρήστη, Διεύθυνση ηλεκτρονικού ταχυδρομείου), από τα οποία θα προέκυπτε ασφαλέστερη και αμεσότερη συσχέτιση με τον φοιτητή. Δεν υπάρχουν επίσης δεδομένα που αφορούν την επίδοση του φοιτητή.

Απεναντίας, στο φοιτητολόγιο (όπου υπάρχουν δεδομένα σχετικά με την επίδοση), το όνομα εμφανίζεται με ελληνική γραφή και με κεφαλαία γράμματα και επιπλέον εμφανίζεται ο Αριθμός

Μητρώου του φοιτητή. Συνεπώς, προκειμένου να συσχετίσουμε τις εγγραφές που αφορούν δεδομένα χρήσης των φοιτητών (όπως αντλήθηκαν από το Moodle), με τις εγγραφές που αφορούν τις επιδόσεις των φοιτητών (όπως αντλήθηκαν από το φοιτητολόγιο), υλοποιήσαμε τις εξής ενέργειες. Μεταγράψαμε<sup>12</sup> τα ονόματα των φοιτητών που αντλήσαμε από το φοιτητολόγιο από ελληνική γραφή σε λατινική γραφή, ακολουθώντας το πρότυπο EL0T 743 το οποίο έχει χρησιμοποιηθεί για την εισαγωγή των ονομάτων στη βάση δεδομένων (Active Directory) του Πανεπιστημίου Αιγαίου. Παράλληλα, διαγράψαμε περιττούς και λανθασμένους χαρακτήρες (διπλά κενά, παύλες, καθέτους) από τα ονόματα φοιτητών στο αρχείο καταγραφής του Moodle. Εντοπίσαμε τις εγγραφές όπου τα ονόματα (μετά την μεταγραφή τους) ταυτίζονταν και συσχετίσαμε τα δεδομένα χρήσης από το αρχείο καταγραφής του Moodle με την επίδοση από το φοιτητολόγιο του Πανεπιστημίου. Στις λίγες περιπτώσεις που οι παραπάνω ενέργειες δεν οδηγούσαν σε αυτόματο εντοπισμό των κοινών εγγραφών των φοιτητών, αναζητήσαμε τα ασυσχέτιστα ονόματα και προχωρήσαμε σε στοχευμένες ενέργειες διόρθωσης και συσχέτισης.

#### 9.5.4. Αντικατάσταση ονομάτων χρηστών με ψευδώνυμα

Για την προστασία των προσωπικών δεδομένων των φοιτητών (βλ. Κεφ. 9.3), αμέσως μετά τη συσχέτιση των δεδομένων χρήσης με τις επιδόσεις, αντικαταστήσαμε τα ονόματα των φοιτητών με ψευδώνυμα. Το όνομα κάθε φοιτητή αντικαταστάθηκε από το γράμμα Φ (εκ του Φοιτητής) ακολουθούμενου από μοναδικό τυχαίο αριθμό (ο οποίος λαμβάνει τιμές από το σύνολο [1..K], όπου K το πλήθος των φοιτητών που συγκροτούν το σύνολο δεδομένων). Οπότε, τα αναγνωριστικά των εγγραφών είχαν τιμή Φ1, Φ2, Φ3, κτλ.

#### 9.5.5. Τα χαρακτηριστικά των συνόλων δεδομένων

Πολλές από τις έρευνες στο πεδίο της Μαθησιακής Αναλυτικής, οι οποίες επιδιώκουν να προβλέψουν την ακαδημαϊκή επίδοση των φοιτητών, βασίζονται σε ανάλυση δεδομένων που αφορούν δημογραφικά στοιχεία των φοιτητών, προηγούμενες ακαδημαϊκές επιδόσεις τους,

---

<sup>12</sup> Χρησιμοποιήσαμε το σχετικό εργαλείο στη διαδικτυακή σελίδα της Ελληνικής Αστυνομίας <https://passport.gov.gr/passports/GrElotConverter/GrElotConverter.html>.

απαντήσεις σε ερωτηματολόγια σχετικά με προσωπικές ή εκπαιδευτικές πρακτικές τους κ.α. Ειδικότερα, οι Baashar, Alkawsji, Ali, Alhussian & Bahbouh (2021) στην βιβλιογραφική ανασκόπησή τους κατηγοριοποιούν τα χαρακτηριστικά (features), που χρησιμοποιούν οι σχετικές έρευνες σε επτά κατηγορίες: δημογραφικά χαρακτηριστικά (αξιοποιούνται σε 18 έρευνες), ακαδημαϊκά χαρακτηριστικά (αξιοποιούνται σε 20 έρευνες), βαθμολογικά χαρακτηριστικά (αξιοποιούνται σε 9 έρευνες), επικοινωνιακά χαρακτηριστικά (αξιοποιούνται σε 3 έρευνες), συμπεριφορικά χαρακτηριστικά (αξιοποιούνται σε 5 έρευνες), ψυχολογικά χαρακτηριστικά (αξιοποιούνται σε 2 έρευνες) και προσωπικά - οικογενειακά χαρακτηριστικά (αξιοποιούνται σε 11 έρευνες). Παραδείγματα τέτοιων χαρακτηριστικών είναι το φύλο, η ηλικία, ο μέσος όρος ακαδημαϊκής επίδοσης, ο χρόνος απόκρισης σε μηνύματα, η συμμετοχή σε δραστηριότητες, οι προτιμώμενες μαθησιακές στρατηγικές, το οικογενειακό κοινωνικό και οικονομικό επίπεδο. Επίσης όπως, αναφέρουν οι Chen & Cui (2020), η βιβλιογραφική επισκόπηση των Cui, Chen, Shiri & Fan (2019) ανέδειξε χαρακτηριστικά που σχετίζονται με ακαδημαϊκές επιδόσεις, (Luo, Korpinska, & Liu, 2015), με την συμπεριφορά κατά τη χρήση των Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ), με δεδομένα από ερωτηματολόγια σχετικά με κοινωνικές και οικονομικές παραμέτρους (Guarin, Guzman, & Gonzalez, 2015), με δημογραφικά στοιχεία (Evale, 2017) και με την προηγούμενη ακαδημαϊκή διαδρομή των φοιτητών (Ochoa, 2016). Επιπλέον, η βιβλιογραφική επισκόπηση των Rastrollo-Guerrero, Gomez-Pulido & Duran-Dominguez (2020) κατέγραψε, μελέτες, όπου η πρόβλεψη της επίδοσης βασίζεται σε δημογραφικά χαρακτηριστικά (Kotsiantis, Pierrakeas, & Pintelas, 2004), (Navamani & Kannammal, 2015), σε προηγούμενες επιδόσεις (Bydzovska, 2015a), (Bydzovska, 2015b) και σε συμπεριφορικά χαρακτηριστικά (Villagra-Arnedo, Gallego-Duran, Comran, Llorens Largo, & Molina-Carmona, 2016).

Η διατριβή μας αξιοποίησε δεδομένα που αποκλειστικά σχετίζονταν με τη χρήση του Moodle. Δηλαδή, η προσπάθειά μας για την πρόβλεψη της ακαδημαϊκής επίδοσης των φοιτητών βασιζόταν μόνο στις αλληλεπιδράσεις των φοιτητών με το σύστημα κατά τη χρονική περίοδο από την έναρξη του μαθήματος ως την ολοκλήρωσή του και ειδικότερα ως τη διεξαγωγή της τελικής αξιολόγησης. Οι πληροφορίες για τις αλληλεπιδράσεις προέκυψαν από τις εγγραφές του αρχείου καταγραφής και εκφράζονται με χαρακτηριστικά (όπως π.χ. το πλήθος συνόδων



σύνδεσης, οι συνολικές μέρες αδράνειας), τα οποία υπολογίζονται μετά από επεξεργασία και μελέτη του αρχείου καταγραφής.

Η επιλογή κατάλληλων χαρακτηριστικών συνιστά κρίσιμο ζήτημα, διότι η εύστοχη επιλογή βοηθά στην εξαγωγή γόνιμων συμπερασμάτων, ενώ αντιθέτως, η μη συμπερίληψη σημαντικών χαρακτηριστικών, πιθανόν δεν θα αναδείξει συσχετίσεις ή άλλες παραμέτρους που μελετάμε. Έχοντας ως γνώμονα τη σχετική βιβλιογραφία, εστίασαμε σε πέντε ομάδες χαρακτηριστικών.

- Η πρώτη ομάδα αποτελείται από βασικά χαρακτηριστικά, τα οποία αποτυπώνουν το συνολικό χρόνο διάδρασης με το σύστημα. Κάποια από τα χαρακτηριστικά φαίνεται να παρέχουν παρόμοιες πληροφορίες. Όμως, προκύπτουν με διαφορετικό τρόπο υπολογισμού, οπότε θεωρούμε ότι παρέχουν διαφορετικές «αποχρώσεις» του ίδιου βασικού μεγέθους ή της ίδιας βασικής ιδιότητας και ότι επιπλέον αυξάνουν την αξιοπιστία των μοντέλων, διότι το ίδιο μέγεθος εκφράζεται εναλλακτικούς τρόπους.
- Η δεύτερη ομάδα αποτελείται από συνθετότερα χαρακτηριστικά, που σχετίζονται με μοτίβα διάδρασης με το σύστημα και τα οποία αποτυπώνουν την ομαλότητα, την κανονικότητα και τον ρυθμό της διάδρασης και επίσης από χαρακτηριστικά που καταγράφουν αδράνεια, καθυστερημένη ενασχόληση ή πρόωρη διακοπή χρήσης.
- Η τρίτη ομάδα χαρακτηριστικών κατηγοριοποιεί τις διαδράσεις ως προς το χρόνο που επιτελούνται, δηλαδή καταγράφει την προτιμώμενη χρονική στιγμή της διάδρασης με το σύστημα π.χ. καταγράφει αν ο φοιτητής συνδέεται συνήθως πρωί ή βράδυ, αν συνδέεται τις «καθημερινές» ημέρες ή τις Κυριακές.
- Η τέταρτη ομάδα χαρακτηριστικών κατηγοριοποιεί τις αλληλεπιδράσεις ως προς το στόχο τους, καταγράφοντας ποιοι είναι οι εκπαιδευτικοί πόροι και οι δραστηριότητες (π.χ. αρχεία, δημόσιοι χώροι διαλόγου) στους οποίους κατευθύνεται το ενδιαφέρον του φοιτητή.
- Η πέμπτη ομάδα, αποτελείται από τα χαρακτηριστικά που καταγράφουν την επίδοση των φοιτητών στις ενδιάμεσες αξιολογικές δραστηριότητες (για την περίπτωση του μαθήματος «Α3 Ηλεκτρονική Μάθηση», όπου οι φοιτητές αξιολογούνται για τη συμμετοχή τους στους δημόσιου διαλόγους).

Συνοψίζοντας, αναφέρουμε ότι επιλέξαμε χαρακτηριστικά τα οποία προτείνει η βιβλιογραφία και ότι δεν ακολουθήσαμε την πρακτική όσων, σύμφωνα με τους Chen & Cui (2020), λαμβάνουν αποφάσεις ως προς την επιλογή των χαρακτηριστικών οι οποίες βασίζονται κυρίως σε εικασίες ή στη διαίσθηση σχετικά με την επιρροή κάθε χαρακτηριστικού στο τελικό αποτέλεσμα. Τα χαρακτηριστικά που απαρτίζουν τα σύνολα δεδομένων μας, αποτυπώνουν τις παρακάτω διαστάσεις των αλληλεπιδράσεων μεταξύ των φοιτητών και του συστήματος:

- τον όγκο / τη διάρκεια πρόσβασης (access volume / duration).
- τη συχνότητα πρόσβασης (access frequency).
- την ομαλότητα / κανονικότητα πρόσβασης (access regularity).
- την πυκνότητα πρόσβασης (access density).
- την ασυνέπεια πρόσβασης (access discrepancy).
- την αμεσότητα πρόσβασης (access immediacy).
- τον τύπο του μαθησιακού πόρου ή της δραστηριότητας όπου κατευθύνεται η ενέργεια πρόσβασης.
- την προτιμώμενη χρονική στιγμή (ημέρα και ώρα) της πρόσβασης (preferred access time).
- την επίδοση σε ενδιάμεσες αξιολογικές δραστηριότητες (μόνο για το μάθημα «Α3 Ηλεκτρονική Μάθηση», όπου οι φοιτητές αξιολογούνταν για τη συμμετοχή τους στους δημόσιους διαλόγους)

#### 9.5.6. Αναλυτική παρουσίαση χαρακτηριστικών των συνόλων δεδομένων

Ακολουθεί η παρουσίαση των χαρακτηριστικών και των συχνά πολύπλοκων ζητημάτων, που προκύπτουν κατά τον υπολογισμό των τιμών τους. Στον Πίνακα 19 παρουσιάζουμε συνοπτικά τα χαρακτηριστικά που περιλαμβάνονται στα σύνολα δεδομένων. Στον ίδιο πίνακα παρουσιάζουμε τις διαστάσεις των αλληλεπιδράσεων, που αποτυπώνονται από τα χαρακτηριστικά και τις ενδεικτικές βιβλιογραφικές αναφορές ερευνών, οι οποίες τα χρησιμοποιούν.

Ο Πίνακας 20 περιλαμβάνει την αναλυτική παρουσίαση των χαρακτηριστικών και του υπολογισμού των τιμών τους. Δείχνουμε, ότι συχνά οι τιμές των χαρακτηριστικών δεν προκύπτουν αμέσως και ακόπως, αλλά κατόπιν εκτεταμένης επεξεργασίας πρωτογενών δεδομένων από τα αρχεία καταγραφής και κατόπιν σύνθετων αριθμητικών υπολογισμών. Η παρουσίαση είναι ιδιαίτερος εκτενής. Όπως επισημαίνουν οι Chen & Cui (2020), δεν είναι πολλές οι έρευνες που αναφέρουν λεπτομερώς ποια ήταν τα στάδια επεξεργασίας των πρωτογενών δεδομένων και πώς υπολογίστηκαν οι τιμές των χαρακτηριστικών. Στον Πίνακα 20 που ακολουθεί και στους Πίνακες 56 και 57 του Παραρτήματος Α (με τις σχετικές συναρτήσεις Excel) παρουσιάζουμε αναλυτικά, τον τρόπο υπολογισμού των χαρακτηριστικών, όπως επίσης αναλυτικά παρουσιάσαμε (στις προηγούμενες παραγράφους) τις διαδικασίες επεξεργασίας των πρωτογενών δεδομένων.

**Πίνακας 19. Συνοπτική παρουσίαση των χαρακτηριστικών που συμπεριλήφθηκαν στα σύνολα δεδομένων**

A/A	Όνομα Χαρακτηριστικού	Όνομα στο σύνολο δεδομένων	Διάσταση αλληλεπίδρασης	Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού
1	Πλήθος συνόδων σύνδεσης	Total Sessions	όγκος / διάρκεια	(Conijn, Snijders, Kleingeld & Matzat, 2016), (Dawson, McWilliam & Tan, 2008), (Dawson, Mirriahi & Gasevic, 2015), (Gasevic, Dawson, Rogers & Gasevic, 2016), (Hu, Lo & Shih, 2014), (Jayaprakash, Moody, Lauria, Regan & Baron, 2014), (Kovanovic, 2015), (Lauria, Baron, Devireddy, Sundararaju & Jayaprakash, 2012), (Munoz-Organero, Munoz-Merino & Kloos, 2009), (You, 2016), (Yu & Jo, 2014), (Zacharis, 2015)
2	Συνολικός χρόνος σύνδεσης	TotalOnlineTimeForActivityPeriod	όγκος / διάρκεια	(Conijn, Snijders, Kleingeld & Matzat, 2016), (Hu, Lo & Shih, 2014), (Macfadyen & Dawson, 2010), (Yu & Jo, 2014), (Rienties, Toetenel & Bryan, 315-319), (Tempelaar, Rienties & Giesbers, 2015), (Zacharis, 2015)
3	Πλήθος λεπτών αδράνειας κατά τη χρονική περίοδο που ο φοιτητής ήταν ενεργός	TotalIdleTimeForActivityPeriod	όγκος / διάρκεια, ομαλότητα / κανονικότητα	(Conijn, Snijders, Kleingeld & Matzat, 2016)
4	Πλήθος ημερών όπου καταγράφηκε δραστηριότητα	Total Activity Days	όγκος / διάρκεια	(Raga Jr & Raga, 2017)

5	Πλήθος ημερών μεταξύ πρώτης και τελευταίας συνόδου σύνδεσης	DaysBetweenFirstAndLastSessions	αμεσότητα, ασυνέπεια	(Raga Jr & Raga, 2017)
6	Πυκνότητα πρόσβασης	Access Density	πυκνότητα	(Raga Jr & Raga, 2017)
7	Πλήθος ημερών αδράνειας κατά την διάρκεια της ενεργής περιόδου	TotalIdleDaysDuringActivityPeriod	ομαλότητα / κανονικότητα, ασυνέπεια	(Conijn, Snijders, Kleingeld & Matzat, 2016) κατά αναλογία με το χαρακτηριστικό με A/A 3
8	Μέγιστο πλήθος συνεχών ημερών αδράνειας μεταξύ συνόδων σύνδεσης	Max Idle Days	ομαλότητα / κανονικότητα, ασυνέπεια	(Mutz, Quick, Schroeder, Zook, & Gunkel, 2019), (Conijn, Snijders, Kleingeld & Matzat, 2016)
9	Μέσος όρος πλήθους ημερών αδράνειας μεταξύ των συνόδων σύνδεσης	AverageIdleDaysBetweenSessions	όγκος / διάρκεια, ομαλότητα / κανονικότητα	(Conijn, Snijders, Kleingeld & Matzat, 2016) κατά αναλογία με το χαρακτηριστικό με A/A 10
10	Τυπική απόκλιση των ημερών αδράνειας για τις μέρες αδράνειας του ίδιου φοιτητή	STDEV IdleDaysBetweenSessions	ομαλότητα / κανονικότητα, ασυνέπεια	(Conijn, Snijders, Kleingeld & Matzat, 2016)
11	Πλήθος ημερών μεταξύ έναρξης του μαθήματος και της πρώτης συνόδου σύνδεσης	DaysFromStartUntilFirstSession	κανονικότητα / ομαλότητα, ετοιμότητα, ασυνέπεια	(Conijn, Snijders, Kleingeld & Matzat, 2016)
12	Πλήθος ημερών μεταξύ της τελευταίας συνόδου και του τέλους χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων	DaysFromLastSessionUntilEnd	ομαλότητα / κανονικότητα, ασυνέπεια	(Conijn, Snijders, Kleingeld & Matzat, 2016) και ειδικότερα κατά αναλογία με το χαρακτηριστικό με A/A 11
13	Μέση χρονική διάρκεια συνόδων ανά φοιτητή	Average Time Per Session per Student	όγκος / διάρκεια, κανονικότητα / ομαλότητα	(Conijn, Snijders, Kleingeld & Matzat, 2016), (Dawson, McWilliam & Tan, 2008), (Hu, Lo & Shih, 2014)

14	Τυπική απόκλιση της διάρκειας των συνόδων του φοιτητή	SDEV of Time per Session per student	ομαλότητα / κανονικότητα	(Conijn, Snijders, Kleingeld & Matzat, 2016), (Yu & Jo, 2014), (Dollinger, Matyja & Huber, 2008), (Munoz-Organero, Munoz-Merino & Kloos, 2009)
15	Πλήθος συνόδων την Κυριακή	SundaySessions	προτιμώμενη χρονική στιγμή	(Howard, Meehan & Parnell, 2018), (Quinn & Gray, 2020)
16	Πλήθος συνόδων τις εργάσιμες ημέρες	WeekDaySessions	προτιμώμενη χρονική στιγμή	(Howard, Meehan & Parnell, 2018), (Quinn & Gray, 2020)
17	Πλήθος συνόδων τις πρώτες πρωινές ώρες	EarlyAMSessions	προτιμώμενη χρονική στιγμή	(Raga Jr & Raga, 2017), (Quinn & Gray, 2020)
18	Πλήθος συνόδων τις πρωινές ώρες	LateAMSessions	προτιμώμενη χρονική στιγμή	(Raga Jr & Raga, 2017), (Quinn & Gray, 2020)
19	Πλήθος συνόδων τις απογευματινές ώρες	EarlyPMSessions	προτιμώμενη χρονική στιγμή	(Raga Jr & Raga, 2017), (Quinn & Gray, 2020)
20	Πλήθος συνόδων τις βραδυνές ώρες	LatePMSessions	προτιμώμενη χρονική στιγμή	(Raga Jr & Raga, 2017), (Quinn & Gray, 2020)
21	Πλήθος συμβάντων πρόσβασης	TotalEvents	όγκος	(Conijn, Snijders, Kleingeld & Matzat, 2016), (Dollinger, Matyja & Huber, 2008), (Gasevic, Dawson, Rogers & Gasevic, 2016), (Jayaprakash, Moody, Lauria, Regan & Baron, 2014), (Kovanovic, 2015), (Lauria, Baron, Deviredy, Sundararaju & Jayaprakash, 2012), (Morris, Finnegan & Wu, 2005), (Quinn & Gray, 2020), (Rafaeli & Ravid, 1997), (Raga Jr & Raga, 2017)
22	Πλήθος συμβάντων πρόσβασης σε αρχεία	FileEvents	όγκος, τύπος μαθησιακού πόρου ή δραστηριότητας	(Conijn, Snijders, Kleingeld & Matzat, 2016), (Quinn & Gray, 2020), (Raga Jr & Raga, 2017)
23	Πλήθος συμβάντων πρόσβασης σε φάκελους	FolderEvents	όγκος, τύπος μαθησιακού πόρου ή δραστηριότητας	(Conijn, Snijders, Kleingeld & Matzat, 2016), (Quinn & Gray, 2020), (Raga Jr & Raga, 2017)
24	Πλήθος συμβάντων πρόσβασης σε σελίδες	PageEvents	όγκος, τύπος μαθησιακού πόρου ή δραστηριότητας	(Conijn, Snijders, Kleingeld & Matzat, 2016), (Quinn & Gray, 2020), (Raga Jr & Raga, 2017)
25	Πλήθος συμβάντων πρόσβασης στους χώρους δημόσιου διαλόγου	ForumEvents	όγκος, τύπος μαθησιακού πόρου ή δραστηριότητας	(Arbaugh, 2014), (Conijn, Snijders, Kleingeld & Matzat, 2016), (Kovanovic, 2015), (Lauria, Baron, Deviredy, Sundararaju & Jayaprakash, 2012), (Macfadyen & Dawson, 2010), (Morris, Finnegan & Wu, 2005), (Nandi, Hamilton, Harland & Warburton, 2011), (Quinn & Gray, 2020), (Raga Jr & Raga, 2017), (Zacharis, 2015)

26	Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου	PostedContent Events	όγκος, τύπος μαθησιακού πόρου ή δραστηριότητας	(Conijn, Snijders, Kleingeld & Matzat, 2016), (Dawson, Mirriahi & Gasevic, 2015), (Macfadyen & Dawson, 2010), (Munoz-Organero, Munoz-Merino & Kloos, 2009), (Nandi, Hamilton, Harland & Warburton, 2011), (Lauria, Baron, Devireddy, Sundararaju & Jayaprakash, 2012), (Raga Jr & Raga, 2017), (Zacharis, 2015)
27	Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο	ForumGrade	επίδοση	(Conijn, Snijders, Kleingeld & Matzat, 2016), (Jayaprakash, Moody, Lauria, Regan & Baron, 2014), (Rafaeli & Ravid, 1997), (Schell, Lukoff & Alvarado, 2014), (Tempelaar, Rienties & Giesbers, 2015)

**Πίνακας 20. (Ομάδα Πινάκων) Αναλυτική παρουσίαση των χαρακτηριστικών και του υπολογισμού των τιμών τους**

A/A	1
Όνομα Χαρακτηριστικού	Πλήθος συνόδων σύνδεσης
Όνομα στο σύνολο δεδομένων	Total Sessions
Περιγραφή	Αποτυπώνει το πλήθος των συνδέσεων που πραγματοποιεί ο φοιτητής από την αρχή ως το τέλος της χρονικής περιόδου, στην οποία αντιστοιχεί το σύνολο δεδομένων (π.χ. ως το τέλος του εξαμήνου ή ως το τέλος της 10 <sup>ης</sup> εβδομάδας). Κάθε σύνοδος αρχίζει με το συμβάν της σύνδεσης (login) του φοιτητή και τελειώνει με το συμβάν της αποσύνδεσής του (logout). Ενδιάμεσως, καταγράφονται συμβάντα πρόσβασης σε πόρους και δραστηριότητες.
Ζητήματα υπολογισμού	Στα αρχεία καταγραφής του Moodle, η σύνδεση και η αποσύνδεση φοιτητή δεν καταγράφονται ως διακριτά συμβάντα. Προκύπτουν εμμέσως από το γεγονός, ότι για κάποιον φοιτητή καταγράφονται συμβάντα πρόσβασης σε πόρο ή σε δραστηριότητα του συστήματος, τα οποία διαρκούν συγκεκριμένο χρονικό διάστημα, οπότε συνάγεται ότι ο φοιτητής συνδέθηκε στο σύστημα. Επίσης, στα αρχεία καταγραφής, οι εγγραφές των φοιτητών αλληλοπλέκονται χρονικά. Δηλαδή, για φοιτητές που ταυτοχρόνως χρησιμοποιούν το σύστημα, οι σχετικές εγγραφές εμφανίζονται κατά τη χρονική σειρά που πραγματικά συμβαίνουν, δηλαδή εναλλάξ ανά φοιτητή και όχι ομαδοποιημένες ανά φοιτητή. Οπότε, για τον εντοπισμό της αρχής και του τέλους μιας συνόδου πρέπει

	α) να ταξινομηθούν οι εγγραφές πρώτον ανά φοιτητή και δεύτερον ανά φθίνουσα χρονική στιγμή πρόσβασης και β) να εντοπισθούν οι αλλαγές φοιτητών στην αλληλουχία των εγγραφών στο αρχείο καταγραφής. Η σύνοδος διαρκεί όσο οι διαδοχικές εγγραφές του αρχείου καταγραφής αναφέρονται στον ίδιο φοιτητή και επιπλέον όσο οι διαδοχικές εγγραφές δεν απέχουν μεταξύ τους για χρονικό διάστημα μεγαλύτερο των 40 λεπτών. Αν μεταξύ δύο εγγραφών του ίδιου φοιτητή μεσολαβήσει διάστημα μεγαλύτερο των 40 λεπτών, θεωρούμε ότι ο φοιτητής αποσυνδέθηκε και συνδέθηκε ξανά αργότερα, δηλαδή θεωρούμε ότι εγκαινίασε νέα σύνοδο. Το διάστημα των 40 λεπτών αποτελεί ερευνητική σύμβαση που προτείνεται από τον Zacharis (2015) και τους Conijn, Snijders, Kleingeld & Matzat (2016).
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον όγκο / τη διάρκεια πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016), (Dawson, McWilliam, & Tan, 2008), (Dawson, Mirriahi, & Gasevic, 2015), (Gasevic, Dawson, Rogers, & Gasevic, 2016), (Hu, Lo, & Shih, 2014), (Jayaprakash, Moody, Lauria, Regan, & Baron, 2014), (Kovanovic, 2015), (Lauria, Baron, Devireddy, Sundararaju, & Jayaprakash, 2012), (Munoz-Organero, Munoz-Merino, & Kloos, 2009), (You, 2016), (Yu & Jo, 2014), (Zacharis, 2015)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	1η

A/A	2
Όνομα Χαρακτηριστικού	Συνολικός χρόνος σύνδεσης
Όνομα στο σύνολο δεδομένων	TotalOnlineTimeForActivityPeriod
Περιγραφή	Αποτυπώνει το άθροισμα της διάρκειας (καταμετρημένο σε λεπτά της ώρας) όλων των συνόδων σύνδεσης του φοιτητή. Ως μονάδα μέτρησης

	επιλέχθηκε το λεπτό της ώρας και όχι η ημέρα προκειμένου να επιτευχθεί μεγαλύτερη ακρίβεια στην ποσοτική αποτύπωση της αλληλεπίδρασης.
Ζητήματα υπολογισμού	Η χρονική περίοδος που διαρκεί μία σύνοδος σύνδεσης ισούται με τη διαφορά ώρας μεταξύ της τελευταίας εγγραφής της συνόδου και της πρώτης εγγραφής της συνόδου. Ο εντοπισμός της πρώτης και της τελευταίας εγγραφής της συνόδου γίνεται με υπολογισμό αντίστοιχο με αυτόν που χρησιμοποιείται στο χαρακτηριστικό με A/A 1 «Πλήθος συνόδων σύνδεσης».
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον όγκο / τη διάρκεια πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016), (Hu, Lo, & Shih, 2014), (Macfadyen & Dawson, 2010), (Yu & Jo, 2014), (Rienties, Toeteneel, & Bryan, 315-319), (Tempelaar, Rienties, & Giesbers, 2015), (Zacharis, 2015)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	2η

A/A	3
Όνομα Χαρακτηριστικού	Πλήθος λεπτών αδράνειας κατά τη χρονική περίοδο που ο φοιτητής ήταν ενεργός
Όνομα στο σύνολο δεδομένων	TotalIdleTimeForActivityPeriod
Περιγραφή	Αποτυπώνει το συνολικό χρόνο (καταμετρημένο σε λεπτά της ώρας) που ο φοιτητής παρέμεινε αδρανής. Δεν προσμετράται ο χρόνος μεταξύ της έναρξης του μαθήματος και της πρώτης συνόδου σύνδεσης του φοιτητή και ο χρόνος μεταξύ της τελευταίας συνόδου σύνδεσης του φοιτητή και του τέλους της χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων. Δηλαδή μελετάται το χρονικό διάστημα μεταξύ της πρώτης και της τελευταίας συνόδου σύνδεσης του φοιτητή. Όχι το συνολικό



	χρονικό διάστημα στο οποίο αντιστοιχεί το σύνολο δεδομένων. Ως μονάδα μέτρησης επιλέχθηκε το λεπτό της ώρας (και όχι π.χ. η ημέρα) κατά αντιστοιχία με τη μονάδα μέτρησης του χαρακτηριστικού με A/A 2 «Συνολικός χρόνος σύνδεσης».
Ζητήματα υπολογισμού	Ο υπολογισμός προϋποθέτει τον εντοπισμό των εγγραφών που αποτυπώνουν την έναρξη και το τέλος συνόδων σύνδεσης του φοιτητή. Κατόπιν, αθροίζονται τα χρονικά διαστήματα που μεσολαβούν μεταξύ του τέλους μιας συνόδου και της αρχής της επόμενης συνόδου του ίδιου φοιτητή.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον όγκο / τη διάρκεια και την ομαλότητα / κανονικότητα της πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	5η

A/A	4
Όνομα Χαρακτηριστικού	Πλήθος ημερών όπου καταγράφηκε δραστηριότητα
Όνομα στο σύνολο δεδομένων	Total Activity Days
Περιγραφή	Αποτυπώνει το πλήθος των ημερών που πραγματοποιήθηκαν σύνοδοι σύνδεσης (όχι το πλήθος των ημερών κατά τις οποίες ο φοιτητής παρέμενε ενεργός).

Ζητήματα υπολογισμού	Προϋποθέτει την καταγραφή αλλαγών στις ημερομηνίες σύνδεσης μεταξύ διαδοχικών εγγραφών του ίδιου χρήστη. Ισούται με το πλήθος των αλλαγών αυτών.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον όγκο / διάρκεια της πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Raga Jr & Raga, 2017)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	10η

A/A	5
Όνομα Χαρακτηριστικού	Πλήθος ημερών μεταξύ πρώτης και τελευταίας συνόδου σύνδεσης
Όνομα στο σύνολο δεδομένων	DaysBetweenFirstAndLastSessions
Περιγραφή	Αποτυπώνει το χρονικό διάστημα (με μονάδα μέτρησης την ημέρα) κατά το οποίο ο φοιτητής πραγματοποίησε συνόδους σύνδεσης, δηλαδή ήταν ενεργός. Ως μονάδα μέτρησης επιλέχθηκε η ημέρα, διότι θεωρήσαμε ότι θα ήταν υπερβολική η ακρίβεια που θα παρήγαγε η μέτρηση με μονάδα μέτρησης τα λεπτά της ώρας.
Ζητήματα υπολογισμού	Ο υπολογισμός απαιτεί τον εντοπισμό της ημερομηνίας της πρώτης και της τελευταίας συνόδου σύνδεσης του φοιτητή. Κατόπιν αφαιρείται η ημερομηνία της πρώτης από την ημερομηνία της τελευταίας συνόδου. Για τους φοιτητές που πραγματοποίησαν μόνο μια σύνοδο ή περισσότερες συνόδους εντός μιας και μόνης ημέρας, το χαρακτηριστικό λαμβάνει την τιμή 0.

Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την αμεσότητα της πρόσβασης και την ασυνέπεια της πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Raga Jr & Raga, 2017)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	9η

A/A	6
Όνομα Χαρακτηριστικού	Πυκνότητα πρόσβασης
Όνομα στο σύνολο δεδομένων	Access Density
Περιγραφή	Αποτυπώνει την πυκνότητα / συχνότητα της πρόσβασης κατά την περίοδο που ο φοιτητής είναι ενεργός. Είναι ο λόγος της χρονικής περιόδου (με μονάδα μέτρησης την ημέρα) που ο φοιτητής είναι ενεργός, προς το πλήθος των ημερών που καταγράφηκαν σύνοδοι σύνδεσης.
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει από τη διαίρεση της τιμής του χαρακτηριστικού με A/A 5 «Πλήθος ημερών μεταξύ πρώτης και τελευταίας σύνδεσης» προς την τιμή του χαρακτηριστικού με A/A 4 «Πλήθος ημερών όπου καταγράφηκε δραστηριότητα». Το κλάσμα θα ήταν περισσότερο κατανοητό αν ήταν αντεστραμμένο. Αλλά δεν μπορεί να αντιστραφεί, διότι το χαρακτηριστικό με A/A 5 «Πλήθος ημερών μεταξύ πρώτης και τελευταίας σύνδεσης» λαμβάνει και μηδενικές τιμές για όσους φοιτητές πραγματοποίησαν μόνο μια σύνοδο σύνδεσης, οπότε οι ημερομηνίες της πρώτης και της τελευταίας συνόδου σύνδεσης είναι ίδιες.

Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την πυκνότητα πρόσβασης
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Raga Jr & Raga, 2017)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	11η

A/A	7
Όνομα Χαρακτηριστικού	Πλήθος ημερών αδράνειας κατά τη διάρκεια της ενεργής περιόδου
Όνομα στο σύνολο δεδομένων	TotalIdleDaysDuringActivityPeriod
Περιγραφή	Αποτυπώνει το πλήθος των ημερών που ο φοιτητές ήταν αδρανής εντός της περιόδου που ορίζουν η πρώτη και η τελευταία σύννοδος πρόσβασης.
Ζητήματα υπολογισμού	<p>Ο υπολογισμός προϋποθέτει τον εντοπισμό των εγγραφών που αποτυπώνουν την έναρξη και το τέλος συνόδων σύνδεσης του φοιτητή. Αθροίζονται οι ημέρες που μεσολαβούν μεταξύ του τέλους μιας συνόδου και της αρχής της επόμενης συνόδου.</p> <p>Σημείωση: Ωστόσο, για τους φοιτητές που πραγματοποίησαν μόνο μια σύνοδο σύνδεσης, δεν μπορεί να οριστεί χρονικό διάστημα μεταξύ του τέλους της μιας μοναδικής σύνδεσης τους και της αρχής της επόμενης συνόδου, διότι δεν υπάρχει επόμενη σύνοδος. Στις περιπτώσεις αυτές επιλέξαμε να ορίσουμε ως χρονικό διάστημα μεταξύ των συνόδων του φοιτητή την τιμή 0 και όχι το πλήθος των ημερών της ενεργής περιόδου πλην 1. Η επιλογή αυτή, δημιουργεί παράδοξο, διότι κατά την κοινή λογική όσο λιγότερες είναι οι ημέρες αδράνειας μεταξύ των συνόδων, τόσο πιο ενεργός είναι ο φοιτητής. Επιλέγοντας να ορίσουμε την τιμή 0,</p>

	<p>αποδίδουμε αδιάλειπτη δραστηριότητα στον φοιτητή που συνδέθηκε στο σύστημα μόνο μια φορά!</p> <p>Αν όμως, αντί της τιμή 0, επιλέγαμε να ορίσουμε μια μεγάλη τιμή, π.χ. <math>179=180-1</math> (αν υποθέσουμε ότι η διάρκεια του μαθήματος είναι 180 μέρες), τότε προκύπτει μεγαλύτερο παράδοξο. Συγκεκριμένα, το «Πλήθος ημερών αδράνειας για την περίοδο που ο φοιτητής είναι ενεργός» θα είχε τιμή 179 ημέρες και ταυτοχρόνως η περίοδος που ο φοιτητής είναι ενεργός θα είχε τιμή 1 μέρα. Δηλαδή, οι ημέρες αδράνειας εντός της ενεργής περιόδου (που ορίζεται από την ημερομηνία της πρώτης ως την ημερομηνία της τελευταίας συνόδου) θα ήταν μεγαλύτερες από τις ημέρες της ενεργής περιόδου! Με άλλα λόγια το υποσύνολο (ημέρες αδράνειας) θα ήταν μεγαλύτερο από το σύνολο (ημέρες ενεργής περιόδου).</p> <p>Το πρώτο παράδοξο είναι δυνατόν να αναιρεθεί μερικώς από το γεγονός ότι η μειωμένη δραστηριότητα των φοιτητών με μια μόνο σύνοδο αποκαλύπτεται τελικώς από αρκετά άλλα χαρακτηριστικά όπως π.χ. αυτά που αποτυπώνουν το πλήθος των συνόδων, το συνολικό χρόνο πρόσβασης, την αδράνεια μεταξύ της έναρξης του μαθήματος και πρώτης συνόδου κ.α. Και εν πάση περιπτώσει, το παράδοξο προκύπτει μόνο σε ελάχιστες, οριακές περιπτώσεις, αυτές των σπουδαστών με μία μόνο σύνοδο.</p> <p>Το ζήτημα υφίσταται επίσης και για τα άλλα χαρακτηριστικά που σχετίζονται με ημέρες αδράνειας δηλαδή τα χαρακτηριστικά με A/A 8, 9, 10. Επιλέξαμε να το αντιμετωπίσουμε με τον ίδιο τρόπο.</p>
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την ομαλότητα / κανονικότητα της πρόσβασης και την ασυνέπεια πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016) κατά αναλογία με το χαρακτηριστικό με A/A 3
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	12η

A/A	8
Όνομα Χαρακτηριστικού	Μέγιστο πλήθος συνεχών ημερών αδράνειας μεταξύ συνόδων σύνδεσης
Όνομα στο σύνολο δεδομένων	Max Idle Days
Περιγραφή	Αποτυπώνει το μέγιστο από τα χρονικά διαστήματα (με μονάδα μέτρησης την ημέρα) που ο φοιτητής παρέμεινε αδρανής. Δεν προσμετράται ο χρόνος μεταξύ της έναρξης του μαθήματος ως την πρώτη σύνοδο σύνδεσης του φοιτητή και ο χρόνος μεταξύ της τελευταίας συνόδου σύνδεσης του φοιτητή και του τέλους της χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων. Ως μονάδα μέτρησης επιλέχθηκε η ημέρα, διότι οι περίοδοι αδράνειας συνήθως διαρκούν ημέρες. Αν διαρκούν χρονικό διάστημα μικρότερο της ημέρας, εκτιμούμε ότι δεν συνιστούν αδράνεια.
Ζητήματα υπολογισμού	Ο υπολογισμός προϋποθέτει τον εντοπισμό των εγγραφών που αποτυπώνουν την έναρξη και το τέλος συνόδων σύνδεσης του φοιτητή και τη σύγκριση των χρονικών διαστημάτων που μεσολαβούν μεταξύ του τέλους μιας συνόδου και της αρχής της επόμενης συνόδου.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την ομαλότητα / κανονικότητα της πρόσβασης και την ασυνέπεια πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Motz, Quick, Schroeder, Zook, & Gunkel, 2019) (Conijn, Snijders, Kleingeld, & Matzat, 2016),
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	8η

A/A	9
-----	---

Όνομα Χαρακτηριστικού	Μέσος όρος πλήθους ημερών αδράνειας μεταξύ των συνόδων σύνδεσης
Όνομα στο σύνολο δεδομένων	AverageldleDaysBetweenSessions
Περιγραφή	Αποτυπώνει το μέσο όρο των χρονικών διαστημάτων (καταμετρημένων σε ημέρες) κατά τα οποία ο φοιτητής ήταν αδρανής. Δεν προσμετράται ο χρόνος μεταξύ της έναρξης του μαθήματος ως την πρώτη σύνοδο σύνδεσης του φοιτητή και ο χρόνος μεταξύ της τελευταίας συνόδου σύνδεσης του φοιτητή και του τέλους της χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων. Ως μονάδα μέτρησης επιλέχθηκε η ημέρα, διότι οι περίοδοι αδράνειας συνήθως διαρκούν ημέρες. Αν διαρκούν χρονικό διάστημα μικρότερο της ημέρας, εκτιμούμε ότι δεν συνιστούν αδράνεια.
Ζητήματα υπολογισμού	Ο υπολογισμός προϋποθέτει τον εντοπισμό των εγγραφών που αποτυπώνουν την έναρξη και το τέλος συνόδων σύνδεσης του φοιτητή. Τα χρονικά διαστήματα που μεσολαβούν μεταξύ του τέλους μιας συνόδου και της αρχής της επόμενης συνόδου αθροίζονται και κατόπιν διαιρούνται με το πλήθος τους
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον όγκο / τη διάρκεια και την ομαλότητα / κανονικότητα της πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016) κατά αναλογία με το χαρακτηριστικό με A/A 10
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	6η

A/A	10
-----	----

Όνομα Χαρακτηριστικού	Τυπική απόκλιση των ημερών αδράνειας για τις μέρες αδράνειας του ίδιου φοιτητή
Όνομα στο σύνολο δεδομένων	STDEV IdleDaysBetweenSessions
Περιγραφή	Αποτυπώνει πόσο αποκλίνουν οι τιμές των ημερών αδράνειας του φοιτητή από το μέσο όρο των ημερών αδράνειας του ίδιου φοιτητή (όχι από το μέσο όρο των ημερών αδράνειας όλων των φοιτητών). Συνεπώς, αποτυπώνει κατά πόσο οι περίοδοι αδράνειας έχουν την ίδια περίπου διάρκεια ή αν ήταν άλλοτε μακρές και άλλοτε σύντομες.
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εφαρμογή της αντίστοιχης στατιστικής συνάρτησης η οποία τροφοδοτείται με τις τιμές ημερών αδράνειας του φοιτητή.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την ομαλότητα / κανονικότητα της πρόσβασης και την ασυνέπεια πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	7η

A/A	11
Όνομα Χαρακτηριστικού	Πλήθος ημερών μεταξύ έναρξης του μαθήματος και της πρώτης συνόδου σύνδεσης
Όνομα στο σύνολο δεδομένων	DaysFromStartUntilFirstSession



Περιγραφή	Αποτυπώνει πόσο έγκαιρα ή πόσο καθυστερημένα σε σχέση με την έναρξη του μαθήματος ξεκίνησαν οι σύνοδοι σύνδεσης του φοιτητή.
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει αφαιρώντας την ημερομηνία της έναρξης του μαθήματος από την ημερομηνία της πρώτης συνόδου σύνδεσης.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την κανονικότητα / ομαλότητα πρόσβασης, με την ετοιμότητα πρόσβασης και με την ασυνέπεια πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016)
Μάθημα όπου το χαρακτηριστικό μελετάται	
Στήλη στο Σύνολο Δεδομένων	13η

A/A	12
Όνομα Χαρακτηριστικού	Πλήθος ημερών μεταξύ της τελευταίας συνόδου και του τέλους χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων
Όνομα στο σύνολο δεδομένων	DaysFromLastSessionUntilEnd
Περιγραφή	Αποτυπώνει αν ο φοιτητής ήταν ενεργός ως το τέλος τη περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων ή αν τυχόν αδρανοποιήθηκε νωρίτερα.
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει αφαιρώντας την ημερομηνία της τελευταίας συνόδου σύνδεσης από την ημερομηνία ολοκλήρωσης της περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων.

Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την κανονικότητα / ομαλότητα πρόσβασης και με την ασυνέπεια πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016) κατά αναλογία με το χαρακτηριστικό με A/A 11
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	14η

	13
Όνομα Χαρακτηριστικού	Μέση χρονική διάρκεια συνόδων ανά φοιτητή
Όνομα στο σύνολο δεδομένων	Average Time Per Session per Student
Περιγραφή	Αποτυπώνει τη μέση διάρκεια των συνόδων σύνδεσης του φοιτητή.
Ζητήματα υπολογισμού	Είναι το κλάσμα του συνολικού χρόνου σύνδεσης (Χαρακτηριστικό με A/A 2) προς το πλήθος των συνόδων σύνδεσης του φοιτητή (Χαρακτηριστικό με A/A 1).
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον όγκο / τη διάρκεια πρόσβασης και με την κανονικότητα / ομαλότητα πρόσβασης
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016), (Dawson, McWilliam & Tan, 2008), (Hu, Lo & Shih, 2014)

Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	3η

A/A	14
Όνομα Χαρακτηριστικού	Τυπική απόκλιση της διάρκειας των συνόδων του φοιτητή.
Όνομα στο σύνολο δεδομένων	SDEV of Time per Session per student
Περιγραφή	Αποτυπώνει πόσο αποκλίνουν οι τιμές της διάρκειας των συνόδων του φοιτητή από τη μέση διάρκεια των συνόδων του ίδιου φοιτητή. Συνεπώς, αποτυπώνει κατά πόσο οι σύνοδοι σύνδεσης του φοιτητή έχουν την ίδια περίπου διάρκεια ή αν ήταν άλλοτε μακρές και άλλοτε σύντομες.
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εφαρμογή της αντίστοιχης στατιστικής συνάρτησης η οποία τροφοδοτείται με τις τιμές διάρκειας όλων των συνόδων
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την ομαλότητα / κανονικότητα της πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016), (Yu & Jo, 2014), (Dollinger, Matyja, & Huber, 2008), (Munoz-Organero, Munoz-Merino, & Kloos, 2009)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	4η

A/A	15
Όνομα Χαρακτηριστικού	Πλήθος συνόδων την Κυριακή
Όνομα στο σύνολο δεδομένων	SundaySessions
Περιγραφή	Αποτυπώνει το πλήθος των συνόδων σύνδεσης που πραγματοποιούνται τις Κυριακές.
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των συνόδων που πραγματοποιούνται τις Κυριακές. Προϋποθέτει την μετατροπή της ημερομηνίας των εγγραφών σε μορφή όπου εμφανίζεται η ημέρα της εβδομάδας.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την προτιμώμενη χρονική στιγμή πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Howard, Meehan, & Parnell, 2018), (Quinn & Gray, 2020)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	15η

A/A	16
Όνομα Χαρακτηριστικού	Πλήθος συνόδων τις εργάσιμες ημέρες.

Όνομα σύνολο δεδομένων	στο WeekDaySessions
Περιγραφή	Αποτυπώνει το πλήθος των συνόδων σύνδεσης που πραγματοποιούνται τις εργάσιμες ημέρες (συμπεριλαμβανομένου και του Σαββάτου).
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των συνόδων που πραγματοποιούνται τις εργάσιμες ημέρες συμπεριλαμβανομένου και του Σαββάτου. Προϋποθέτει την μετατροπή της ημερομηνίας των εγγραφών σε μορφή όπου εμφανίζεται η ημέρα της εβδομάδας.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την προτιμώμενη χρονική στιγμή πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Howard, Meehan, & Parnell, 2018), (Quinn & Gray, 2020)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη Σύνολο Δεδομένων	στο 15η

A/A	17
Όνομα Χαρακτηριστικού	Πλήθος συνόδων τις πρώτες πρωινές ώρες
Όνομα σύνολο δεδομένων	στο EarlyAMSessions
Περιγραφή	Αποτυπώνει το πλήθος των συνόδων σύνδεσης που ξεκίνησαν τις πρώτες πρωινές ώρες και συγκεκριμένα από 00:00 ως 07:00. Ο διαχωρισμός του 24ώρου σε 4 περιόδους (πρώτες πρωινές ώρες, πρωινές ώρες, απογευματινές ώρες, βραδινές ώρες) που εφαρμόσαμε στο παρόν αλλά

	και στα επόμενα χαρακτηριστικά (με A/A 18, 19, 20) επηρεάστηκε από το ανάλογο πρότυπο που χρησιμοποιούν οι Raga & Raga (2017).
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των συνόδων που ξεκίνησαν τις πρωινές ώρες και συγκεκριμένα από 00:00 ως 06:59.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την προτιμώμενη χρονική στιγμή πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Raga Jr & Raga, 2017)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	17η

A/A	18
Όνομα Χαρακτηριστικού	Πλήθος συνόδων τις πρωινές ώρες
Όνομα στο σύνολο δεδομένων	LateAMSessions
Περιγραφή	Αποτυπώνει το πλήθος των συνόδων σύνδεσης που ξεκίνησαν τις πρωινές ώρες και συγκεκριμένα από 07:00 ως 11:59.
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των συνόδων που ξεκίνησαν τις πρωινές ώρες και συγκεκριμένα από 07:00 ως 11:59.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την προτιμώμενη χρονική στιγμή πρόσβασης.
Βιβλιογραφικές αναφορές που	(Raga Jr & Raga, 2017)

προτείνουν τη μελέτη του χαρακτηριστικού	
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	18η

A/A	19
Όνομα Χαρακτηριστικού	Πλήθος συνόδων τις απογευματινές ώρες
Όνομα στο σύνολο δεδομένων	EarlyPMSessions
Περιγραφή	Αποτυπώνει το πλήθος των συνόδων σύνδεσης που ξεκίνησαν τις απογευματινές ώρες και συγκεκριμένα από 12:00 ως 20:59.
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των συνόδων που ξεκίνησαν τις απογευματινές ώρες και συγκεκριμένα από 12:00 ως 20:59.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την προτιμώμενη χρονική στιγμή πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Raga Jr & Raga, 2017)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	19η

A/A	20
Όνομα Χαρακτηριστικού	Πλήθος συνόδων τις βραδυνές ώρες
Όνομα στο σύνολο δεδομένων	LatePMSessions
Περιγραφή	Αποτυπώνει το πλήθος των συνόδων σύνδεσης που ξεκίνησαν τις βραδινές ώρες και συγκεκριμένα από 21:00 ως 23:59.
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των συνόδων που ξεκίνησαν τις βραδινές ώρες και συγκεκριμένα από 21:00 ως 23:59.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την προτιμώμενη χρονική στιγμή πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Raga Jr & Raga, 2017)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	20η

A/A	21
Όνομα Χαρακτηριστικού	Πλήθος συμβάντων πρόσβασης
Όνομα στο σύνολο δεδομένων	TotalEvents



Περιγραφή	Αποτυπώνει το πλήθος των συμβάντων που αντιστοιχούν στον φοιτητή
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των εγγραφών του αρχείου καταγραφής που αντιστοιχούν στον φοιτητή.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον όγκο πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016), (Dollinger, Matyja, & Huber, 2008), (Gasevic, Dawson, Rogers, & Gasevic, 2016), (Jayaprakash, Moody, Lauria, Regan, & Baron, 2014), (Kovanovic, 2015), (Lauria, Baron, Devireddy, Sundararaju, & Jayaprakash, 2012), (Morris, Finnegan, & Wu, 2005), (Quinn & Gray, 2020), (Rafaeli & Ravid, 1997), (Raga Jr & Raga, 2017)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	21η

A/A	22
Όνομα Χαρακτηριστικού	Πλήθος συμβάντων πρόσβασης σε αρχεία
Όνομα στο σύνολο δεδομένων	FileEvents
Περιγραφή	Αποτυπώνει το πλήθος των συμβάντων πρόσβασης σε πόρους τύπου «αρχείο»
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των εγγραφών που αντιστοιχούν στον φοιτητή και σχετίζονται με πρόσβαση σε αναρτημένα αρχεία.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον τύπο μαθησιακού πόρου ή δραστηριότητας όπου κατευθύνεται η πρόσβαση και με τον όγκο πρόσβασης.

Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016), (Quinn & Gray, 2020), (Raga Jr & Raga, 2017)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	22η

A/A	23
Όνομα Χαρακτηριστικού	Πλήθος συμβάντων πρόσβασης σε φάκελους
Όνομα στο σύνολο δεδομένων	FolderEvents
Περιγραφή	Αποτυπώνει το πλήθος των συμβάντων πρόσβασης σε πόρους τύπου «φάκελος»
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των εγγραφών που αντιστοιχούν στον φοιτητή και σχετίζονται με πρόσβαση σε φάκελους.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον τύπο μαθησιακού πόρου ή δραστηριότητας όπου κατευθύνεται η πρόσβαση και με τον όγκο πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016), (Quinn & Gray, 2020), (Quinn & Gray, 2020) (Raga Jr & Raga, 2017)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ

Στήλη Σύνολο Δεδομένων	στο 23η
------------------------------	---------

A/A	24
Όνομα Χαρακτηριστικού	Πλήθος συμβάντων πρόσβασης σε σελίδες
Όνομα στο σύνολο δεδομένων	PageEvents
Περιγραφή	Αποτυπώνει το πλήθος των συμβάντων πρόσβασης σε πόρους τύπου «σελίδα»
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των εγγραφών που αντιστοιχούν στον φοιτητή και σχετίζονται με πρόσβαση σε σελίδες.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον τύπο μαθησιακού πόρου ή δραστηριότητας όπου κατευθύνεται η πρόσβαση και με τον όγκο πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016), (Quinn & Gray, 2020), (Raga Jr & Raga, 2017)
Μάθημα όπου το χαρακτηριστικό μελετάται	AA005 Χρήση Βασικών Εφαρμογών ΗΥ
Στήλη Σύνολο Δεδομένων	στο 24η

A/A	25
Όνομα Χαρακτηριστικού	Πλήθος συμβάντων πρόσβασης στους χώρους δημόσιου διαλόγου

Όνομα σύνολο δεδομένων	στο ForumEvents
Περιγραφή	Αποτυπώνει το πλήθος των συμβάντων πρόσβασης στους χώρους δημόσιου διαλόγου είτε για ανάγνωση ανάρτησης, είτε για σχολιασμό ανάρτησης, είτε για δημιουργία ανάρτησης.
Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των εγγραφών που αντιστοιχούν στον φοιτητή και σχετίζονται με πρόσβαση στους χώρους δημόσιου διαλόγου.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον τύπο μαθησιακού πόρου ή δραστηριότητας όπου κατευθύνεται η πρόσβαση και με τον όγκο πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Arbaugh, 2014), (Conijn, Snijders, Kleingeld, & Matzat, 2016), (Kovanovic, 2015), (Lauria, Baron, Devireddy, Sundararaju, & Jayaprakash, 2012), (Macfadyen & Dawson, 2010), (Morris, Finnegan, & Wu, 2005), (Nandi, Hamilton, Harland, & Warburton, 2011), (Quinn & Gray, 2020), (Raga Jr & Raga, 2017), (Zacharis, 2015)
Μάθημα όπου το χαρακτηριστικό μελετάται	A3 Ηλεκτρονική Μάθηση
Στήλη Σύνολο Δεδομένων	στο 23η

A/A	26
Όνομα Χαρακτηριστικού	Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου
Όνομα σύνολο δεδομένων	στο PostedContentEvents
Περιγραφή	Αποτυπώνει το πλήθος των συμβάντων πρόσβασης στους χώρους δημόσιου διαλόγου για δημιουργία ανάρτησης.

Ζητήματα υπολογισμού	Ο υπολογισμός προκύπτει με εντοπισμό και άθροισμα των εγγραφών που αντιστοιχούν στον φοιτητή και σχετίζονται με δημιουργία ανάρτησης στους χώρους δημόσιου διαλόγου.
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με τον τύπο μαθησιακού πόρου ή δραστηριότητας όπου κατευθύνεται η πρόσβαση και με τον όγκο πρόσβασης.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016), (Dawson, Mirriahi, & Gasevic, 2015), (Macfadyen & Dawson, 2010), (Munoz-Organero, Munoz-Merino, & Kloos, 2009), (Nandi, Hamilton, Harland, & Warburton, 2011), (Lauria, Baron, Devireddy, Sundararaju, & Jayaprakash, 2012), (Raga Jr & Raga, 2017), (Zacharis, 2015)
Μάθημα όπου το χαρακτηριστικό μελετάται	A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	24η

A/A	27
Όνομα Χαρακτηριστικού	Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο
Όνομα στο σύνολο δεδομένων	ForumGrade
Περιγραφή	Αποτυπώνει το βαθμό που έλαβε ο φοιτητής για τη συμμετοχή του στους δημόσιου διαλόγους. Κάθε φοιτητής του μαθήματος A3 Ηλεκτρονική μάθηση για κάθε εβδομάδα από την 1 <sup>η</sup> ως την 7 <sup>η</sup> λαμβάνει βαθμό για τη συμμετοχή του στο δημόσιο διάλογο που διεξάγεται στα πλαίσια της εβδομαδιαίας εκπαιδευτικής συνεδρίας.
Ζητήματα υπολογισμού	Η τιμή του χαρακτηριστικού δεν προκύπτει από τις εγγραφές του αρχείου καταγραφής. Η τιμή αντιγράφεται (όχι αυτόματα) από τα αρχεία βαθμολογίου του μαθήματος. Υπολογίζουμε το μέσο όρο των βαθμών, υιοθετώντας την επιλογή των Conijn, Snijders, Kleingeld & Matzat (2016), που επέλεξαν να υπολογίσουν μέσους όρους για λόγους διαχειρισμότητας των μεταβλητών. Ειδικότερα, η τιμή του

	<p>χαρακτηριστικού στο σύνολο δεδομένων που αντιστοιχεί στην περίοδο από την εβδομάδα 1 ως την εβδομάδα X (όπου <math>X \leq 7</math>), είναι ο μέσος όρος των βαθμών της εβδομάδας 1, της εβδομάδας 2, ... και της εβδομάδας X. Στα σύνολα δεδομένων που αντιστοιχούν σε περίοδο από την εβδομάδα 1 ως την εβδομάδα Y (όπου <math>Y &gt; 7</math>), δεδομένου ότι δεν βαθμολογούνται οι δημόσιοι διάλογοι μετά την εβδομάδα 7, η τιμή του χαρακτηριστικού είναι ίδια με την τιμή του χαρακτηριστικού στο σύνολο δεδομένων της περιόδου από την εβδομάδα 1 ως την εβδομάδα 7.</p>
Διάσταση αλληλεπίδρασης	Το χαρακτηριστικό σχετίζεται με την επίδοση σε ενδιάμεσες αξιολογικές δραστηριότητες.
Βιβλιογραφικές αναφορές που προτείνουν τη μελέτη του χαρακτηριστικού	(Conijn, Snijders, Kleingeld, & Matzat, 2016), (Jayaprakash, Moody, Lauria, Regan, & Baron, 2014)], (Rafaeli & Ravid, 1997), (Schell, Lukoff, & Alvarado, 2014), (Tempelaar, Rienties, & Giesbers, 2015)
Μάθημα όπου το χαρακτηριστικό μελετάται	A3 Ηλεκτρονική Μάθηση
Στήλη στο Σύνολο Δεδομένων	25η

Ας σημειωθεί ότι από τα παραπάνω χαρακτηριστικά, δεν περιλήφθηκαν στα τριδιάστατα σύνολα δεδομένων τα χαρακτηριστικά με A/A: 5, 6, 9, 10, 11, 12, 13, 14. Ο λόγος της μη συμπερίληψης είναι ότι στα τριδιάστατα σύνολα αποτυπώνεται, όχι η συνολική τελική τιμή κάθε χαρακτηριστικού, αλλά (με μορφή αριθμητικής ακολουθίας) οι χρονικά διαδοχικές και αύξουσες τιμές του. Τα εν λόγω χαρακτηριστικά, δεν είναι συγκρίσιμα μεταξύ διαφορετικών χρονικών περιόδων, συνεπώς δεν έχει νόημα να εκφραστούν ως αύξουσες αριθμητικές ακολουθίες.

#### 9.5.7. Επιλογή χαρακτηριστικών

Συχνά, οι ερευνητές προβλεπτικών μοντέλων Μηχανικής Μάθησης, δεν διατηρούν, έως τα τελικά στάδια της μελέτης, όλα τα χαρακτηριστικά που αρχικώς επέλεξαν. Για λόγους απλότητας, κατασκευάζουν μικρότερα σύνολα δεδομένων, που αποτελούνται μόνο από τα χαρακτηριστικά που αξιολογούνται ως κύρια. Όπως σχετικά αναφέρουν οι Howard, Meehan &

Parnell (2018) υφίσταται δυο μεθοδολογικές επιλογές ως προς την καταγραφή των χαρακτηριστικών: η λεπτομερής (fine-grained) και η αδρή (coarse-grained). Η λεπτομερής καταγραφή των χαρακτηριστικών παράγει καλύτερα μοντέλα, αλλά με τίμημα την απλότητα, δεδομένου ότι τα απλά μοντέλα με περιορισμένο αριθμό χαρακτηριστικών είναι ευκολότερο να κατανοηθούν και να ερμηνευτούν. Έτσι, οι Raga Jr & Raga (2017) και οι Riestra-Gonzalez, del Puerto Paule-Ruiz & Ortin (2021) τελικώς μελετούν κάποια μόνο από τα αρχικώς επιλεγέντα χαρακτηριστικά. Αντιθέτως, οι Hussain, Zhu, Zhang & Abidi (2018) και ο Kovacic (2010) δεν αφαιρούν χαρακτηριστικά από τα μοντέλα τους, όπως ακριβώς πράττουν και οι Luan & Zhao (2006), οι οποίοι υποστηρίζουν ότι η πρακτική της παράλειψης χαρακτηριστικών πρέπει να αποφεύγεται, διότι ακόμα και αν κάποια χαρακτηριστικά δεν επηρεάζουν τη συνολική πρόβλεψη, μπορεί ωστόσο να είναι ουσιώδη για κάποια εγγραφή.

Επιλέξαμε τη λεπτομερή (fine-grained) καταγραφή των χαρακτηριστικών. Δηλαδή επιλέξαμε να διατηρήσουμε όλα τα χαρακτηριστικά στα σύνολα δεδομένων που κατασκευάσαμε, διότι επιδιώκουμε την καλύτερη δυνατή πρόβλεψη. Επιπλέον, διότι μελετήσαμε και συγκρίναμε σύνολα δεδομένων που κάλυπταν διαφορετικές χρονικές περιόδους. Δηλαδή, ακόμα και αν κάποιο χαρακτηριστικό αξιολογηθεί ως ασήμαντο για το σύνολο δεδομένων κάποιας χρονικής περιόδου, δεν είναι βέβαιο ότι θα ήταν επίσης ασήμαντο για το σύνολο δεδομένων άλλης χρονικής περιόδου. Για παράδειγμα, το χαρακτηριστικό με A/A 11 (*Πλήθος ημερών μεταξύ έναρξης του μαθήματος και της πρώτης συνόδου σύνδεσης*) αποκαλύπτει τους φοιτητές που δεν ξεκίνησαν εγκαίρως την παρακολούθηση του μαθήματος. Αναμένουμε ότι το χαρακτηριστικό θα είναι σημαντικό για τις χρονικές περιόδους που καλύπτουν τις πρώτες εβδομάδες του μαθήματος. Όμως, με την πάροδο του χρόνου, τα τυχόν γνωστικά κενά που προκάλεσε η αρχική αποχή από τις διαλέξεις μειώνονται, οπότε η σημασία του χαρακτηριστικού μειώνεται για τις χρονικές περιόδους που καλύπτουν σχεδόν όλο το μάθημα. Το αντίστροφο συμβαίνει με το χαρακτηριστικό με A/A 12 (*Πλήθος ημερών μεταξύ της τελευταίας συνόδου και του τέλους χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων*), το οποίο δηλώνει πρόωρη αποχώρηση από το μάθημα. Άλλο παράδειγμα, αποτελεί το χαρακτηριστικό με A/A 25 (*Πλήθος συμβάντων πρόσβασης στους χώρους δημόσιου διαλόγου*), στο μάθημα «Α3 Ηλεκτρονική Μάθηση». Η σημασία του χαρακτηριστικού είναι μεγάλη για τις πρώτες εβδομάδες του

μαθήματος, όπου οι φοιτητές είναι υποχρεωμένοι να συμμετάσχουν στους δημόσιους διαλόγους. Προς το τέλος του μαθήματος όταν η συμμετοχή δεν είναι υποχρεωτική, η σημασία του χαρακτηριστικού φθίνει.

#### 9.5.8. Υπολογισμός τιμών για τα χαρακτηριστικά των συνόλων δεδομένων

Οι τιμές των χαρακτηριστικών που συγκροτούν τα σύνολα δεδομένων δεν προκύπτουν άμεσα από τα πρωτογενή δεδομένα των αρχείων καταγραφής του Moodle. Τα αρχεία καταγραφής περιλαμβάνουν χιλιάδες εγγραφές, η καθεμιά εκ των οποίων αντιστοιχεί σε μία διακριτή ενέργεια του φοιτητή. Για παράδειγμα, περιλαμβάνουν εγγραφές για το ποια μέρα και ώρα καταγράφηκαν συμβάντα σχετικά με τον φοιτητή, ποια αρχεία είδε ο φοιτητής, σε ποιους χώρους δημόσιου διαλόγου ανάρτησε σχόλιο. Τα πρωτογενή δεδομένα κάθε εγγραφής αφορούν τον χρόνο του συμβάντος, το είδος του συμβάντος και τον φοιτητή που υλοποίησε το συμβάν (βλ. Πίνακα 11). Πληροφορίες, όπως για παράδειγμα το πλήθος των συνδέσεων του φοιτητή, η συνολική διάρκεια των συνδέσεων, ή η συχνότητά τους, δεν περιλαμβάνονται στα αρχεία καταγραφής του συστήματος, οπότε δεν μπορούν να αντληθούν άμεσα και αυτόματα. Είναι λοιπόν απαραίτητο, από το σύνολο των πρωτογενών δεδομένων, που βρίσκονται στα αρχεία καταγραφής, να εξαγάγουμε / εξορύξουμε δευτερογενή δεδομένα, τα οποία ενδιαφέρουν την έρευνά μας και ειδικότερα να εξαγάγουμε / εξορύξουμε πληροφορίες που αποκαλύπτουν μοτίβα (patterns) στη χρήση του συστήματος από τους φοιτητές. Οι πληροφορίες αυτές αποτυπώνονται στις τιμές των χαρακτηριστικών που αντιστοιχούν στις στήλες των συνόλων δεδομένων.

Οι υπολογισμοί με τους οποίους προκύπτουν οι τιμές των χαρακτηριστικών, που συγκροτούν τα σύνολα δεδομένων, πραγματοποιήθηκαν με τη βοήθεια του λογισμικού Excel και ειδικότερα με τη χρήση των συναρτήσεων που περιλαμβάνει η βιβλιοθήκη συναρτήσεων του λογισμικού. Το είδος των συναρτήσεων που χρησιμοποιήθηκαν, η συντακτική μορφή των συναρτήσεων και οι μεταβλητές που συμμετείχαν στους υπολογισμούς παρουσιάζονται αναλυτικά στους Πίνακες 56 και 57 του Παραρτήματος Α.



#### 9.5.9. Η κλάση των οντοτήτων

Η τελευταία στήλη του διδιάστατου και του τριδιάστατου πίνακα στα σύνολα δεδομένων αντιστοιχεί στην κατηγορία / κλάση (class), όπου κατατάσσεται κάθε οντότητα. Η κλάση είναι η *εξαρτημένη* μεταβλητή, που αντιστοιχεί στην επίδοση του φοιτητή και την οποία επιδιώκουμε να προβλέψουμε με τα μοντέλα που κατασκευάζουμε. Η επίσημη ακαδημαϊκή επίδοση κάθε φοιτητή εκφράζεται, όπως είναι γνωστό, με δεκαδικό αριθμό από το 0 έως το 10. Όμως, η έρευνα δεν επιδιώκει να προβλέψει τον βαθμό με τον οποίο αξιολογείται ο φοιτητής. Επιδιώκει να εντοπίσει τους φοιτητές με χαμηλές επιδόσεις, οπότε, η κλάση που κατατάσσουμε τους φοιτητές δεν εκφράζεται με ακέραιο αριθμό, αλλά με *κατηγορική* μεταβλητή. Ειδικότερα, στο μάθημα «ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ», η κατηγορική μεταβλητή που αποτελεί την κλάση τιτλοφορείται «Επιτυχία» (Pass) και λαμβάνει τιμές από το σύνολο {Yes, No}. Η τιμή «Yes» αντιστοιχεί σε βαθμούς μεγαλύτερους, ή ίσους από το 5 και η τιμή «No» στους υπόλοιπους βαθμούς. Στο μάθημα «Α3 Ηλεκτρονική Μάθηση», δεν βρέθηκαν φοιτητές με βαθμό μικρότερο του 5. Ειδικότερα, οι βαθμοί κυμαίνονται από 5 έως 10, με διάμεσο τον αριθμό 8.5. Συνεπώς, δεν είχε νόημα να κατηγοριοποιήσουμε, όπως ακριβώς στο μάθημα «ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ», δηλαδή να διαχωρίσουμε τους φοιτητές σε επιτυχόντες και μη, διότι όλοι οι φοιτητές είναι επιτυχόντες. Οπότε, ελλείψει αποτυχόντων, επιλέξαμε να δώσουμε στην κατηγορική μεταβλητή, που αποτελεί την κλάση, τον τίτλο «Εύγε» (Well Done) και να της εκχωρήσουμε τιμές από το σύνολο {Yes, No}. Προκειμένου να έχουμε ισόρροπη κατανομή των οντοτήτων στις δύο κατηγορίες, διαχωρίσαμε τις δύο κατηγορίες με κριτήριο τη διάμεσο. Δηλαδή, η τιμή «Yes» αντιστοιχεί σε βαθμούς μεγαλύτερους από τη διάμεσο (το 8.5) και η τιμή «No» στους υπόλοιπους βαθμούς.

Η διχοτόμηση της κλάσης, δηλαδή η κατανομή των οντοτήτων σε δύο κατηγορίες, αποτελεί μεθοδολογική επιλογή που συναντάται συχνά στην βιβλιογραφία. Συγκεκριμένα, οι Howard, Meehan & Parnell (2018) αναφέρουν ότι μια από τις συνήθεις εφαρμογές της Μαθησιακής Αναλυτικής είναι η πρόβλεψη της επιτυχίας, ή της αποτυχίας των φοιτητών με τη χρήση δυαδικής κατηγορικής μεταβλητής, η οποία διχοτομεί την επίδοση των φοιτητών. Ως αντιπροσωπευτικά παραδείγματα χρήσης δυαδικής κατηγορικής μεταβλητής, αναφέρουν τις μελέτες των Azcona & Casey (2015) των Marbouti, Diefes-Dux & Madhavan (2016) και των Calvo-

Flores, Galindo, Jimenez & Pineiro (2006). Οι Riestra-Gonzalez, del Puerto Paule-Ruiz & Ortin (2021), αναφέρουν ότι ο σκοπός τους είναι (όπως και της διατριβής μας), όχι να προβλέψουν επακριβώς την βαθμολογική επίδοση των φοιτητών, αλλά να διακρίνουν τους φοιτητές που θα έχουν καλή επίδοση, έναντι των φοιτητών που θα κινδυνέψουν να αποτύχουν. Για τις δύο αυτές κατηγορίες, όρισαν δύο διακριτές κατηγορικές τιμές και δημιούργησαν δυαδικά μοντέλα κατηγοριοποίησης. Οι Chen & Cui (2020) επίσης προτείνουν διχοτόμηση της κλάσης (Επιτυχία / Αποτυχία). Ο Konavic (2010) αναφέρει ότι χρησιμοποίησε διχοτομημένη (δυαδική) μεταβλητή για την κατηγορία της επίδοσης με τις τιμές «επιτυχία» και «αποτυχία». Όπως πράξαμε και εμείς, δεν διαχώρισε περαιτέρω τη δεύτερη κατηγορία («αποτυχία») σε αυτούς που παρακολούθησαν ως το τέλος το μάθημα και σε αυτούς που αποχώρησαν πρόωρα από το μάθημα, διότι εξαιτίας του σχετικά μικρού συνόλου δεδομένων, θα επηρεαζόταν η ακρίβεια της πρόβλεψης. Με το ίδιο τρόπο διχοτόμησαν την κλάση οι Quinn & Gray (2020, σ. 6) και οι Gasevic, Dawson, Rogers & Gasevic, (2016, σ. 73).

Ας σημειωθεί ότι η πρόβλεψη της αριθμητικής βαθμολογικής επίδοσης δεν αποτελεί ζητούμενο της έρευνάς μας. Ενδεχομένως, συνιστά ενδιαφέρουσα μαθηματική πρόκληση, αλλά όπως έχει ήδη αναφερθεί, η δική μας επιδίωξη είναι ο εντοπισμός φοιτητών που κινδυνεύουν να αποτύχουν στο μάθημα. Συνεπώς, η διχοτόμηση της κλάσης σε δύο κατηγορίες αποτελεί μεθολογική επιλογή, η οποία αφενός είναι έγκυρη επιλογή, δεδομένου ότι αποτελεί συνήθη πρακτική, όπως προκύπτει από τη βιβλιογραφία και αφετέρου είναι κατάλληλη επιλογή, δεδομένου ότι κατηγοριοποιεί σε κλάσεις που πραγματικά μας ενδιαφέρουν.

Δεδομένου ότι η κλάση των οντοτήτων εκφράζεται από κατηγορική μεταβλητή, της οποίας οι τιμές δεν έχουν διάταξη / ταξινομική σειρά (order), το πρόβλημα που μελετάμε κατάσσεται σε αυτά, που συμφωνα με την ορολογία της Μηχανικής Μάθησης, ονομάζονται «προβλήματα ταξινόμησης» (classification problems), τα οποία διαφέρουν από τα λεγόμενα «προβλήματα παλινδρόμησης» (regression problems), στα οποία η κλάση εκφράζεται από συνεχή, μη κατηγορική μεταβλητή (Han, Kamber, & Pei, 2011, σ. 272), (Riestra-Gonzalez, del Puerto Paule-Ruiz, & Ortin, 2021, σ. 6), (Witten, Transcript for Class1 - Lesson 3, 2013). Ως εκ τούτου, οι αλγόριθμοι Μηχανικής Μάθησης που χρησιμοποιούνται στα προβλήματα ταξινόμησης, παράγουν μοντέλα ταξινόμησης ή αλλιώς «ταξινομητές».

#### 9.5.10. Κανονικοποίηση Τιμών

Πριν την εφαρμογή των αλγόριθμων Μηχανικής Μάθησης επί των συνόλων δεδομένων, οι τιμές κανονικοποιήθηκαν. Σχετικά με την κανονικοποίηση των τιμών, οι Han, Kamber & Pei (2011, σ. 105) αναφέρουν, ότι οι αλλαγές στις μονάδες μέτρησης μπορεί να οδηγήσουν σε εντελώς διαφορετικά αποτελέσματα και ότι η έκφραση μια μεταβλητής με μικρότερες μονάδες μέτρησης προσδίδει μεγαλύτερο εύρος για την μεταβλητή επομένως μεγαλύτερη βαρύτητα. Για να αποφευχθεί η εξάρτηση από την μονάδα μέτρησης, οι τιμές πρέπει να κανονικοποιηθούν (normalized) ή να τυποποιηθούν (standardized)<sup>13</sup>. Δηλαδή, να μετασχηματισθούν, έτσι ώστε να ενταχθούν σε στενότερο εύρος τιμών όπως το [-1, 1] ή το [0.0, 1.0]. Επιπλέον, σύμφωνα με τον Brownlee (2019 a), η κανονικοποίηση αποτελεί καλή πρακτική στις περιπτώσεις, που οι τιμές δεν ακολουθούν την κατανομή Gauss.

Ως καλή πρακτική, η κανονικοποίηση υιοθετείται από τους Quinn & Gray (2020), που κανονικοποίησαν τις τιμές για να μην δοθεί υπερβολικό βάρος στις μεταβλητές με μεγάλο εύρος τιμών και από τους Riestra-Gonzalez, del Puerto Paule-Ruiz & Ortin (2021), που κανονικοποίησαν τις τιμές, διότι κάποιοι αλγόριθμοι εμφανίζουν καλύτερη απόδοση αν εφαρμοστούν σε κανονικοποιημένες τιμές. Όμως, όπως υποστηρίζουν οι Cui, Chen, Shiri & Fan (2019, σ. 215), λίγες είναι οι μελέτες που κανονικοποιούν τις μεταβλητές πριν την κατασκευή των μοντέλων.

Στα σύνολα δεδομένων που χρησιμοποιήσαμε, η κανονικοποίηση των τιμών κρίθηκε απαραίτητη, διότι υφίστανται διαφορετικές κλίμακες τιμών και διαφορετικές μονάδες μέτρησης. Για παράδειγμα, οι τιμές του χαρακτηριστικού με A/A 8 (*Μέγιστο πλήθος συνεχών ημερών αδράνειας μεταξύ συνόδων σύνδεσης*) είναι της τάξης των δεκάδων, οι τιμές του χαρακτηριστικού με A/A 2 (*Συνολικός χρόνος σύνδεσης*) είναι της τάξης των εκατοντάδων και οι τιμές του χαρακτηριστικού με A/A 3 (*Πλήθος λεπτών αδράνειας κατά τη χρονική περίοδο που ο φοιτητής ήταν ενεργός*) είναι της τάξης των χιλιάδων. Επιπλέον, σε αρκετά χαρακτηριστικά των συνόλων δεδομένων μας, οι τιμές δεν ακολουθούν την κατανομή Gauss. Χρησιμοποιώντας το

---

<sup>13</sup> Σύμφωνα με τους (Han, Kamber, & Pei, 2011, σ. 105) οι όροι κανονικοποίηση και τυποποίηση χρησιμοποιούνται χωρίς διάκριση στην προ-επεξεργασία δεδομένων (data preprocessing). Ωστόσο, στην στατιστική ο όρος τυποποίηση έχει άλλες συνδηλώσεις.

σχετικό φίλτρο που διαθέτει το WEKA (συγκεκριμένα: Preprocess> FilterChoose> Filters> Unsupervised> Attributes> Normalize) κανονικοποιήσαμε τις τιμές, δηλαδή τις μετασχηματίσαμε ώστε να ανήκουν στο εύρος [0.0 , 1.0]. Ας σημειωθεί, ότι αρκετοί αλγόριθμοι του WEKA, ούτως ή άλλως (by default), κανονικοποιούν τις τιμές των συνόλων δεδομένων πριν τα αναλύσουν.

## 9.6. Δημιουργία μοντέλων ταξινόμησης

Το στάδιο που έπεται της κατασκευής των συνόλων δεδομένων, είναι η ανάλυσή τους με το λογισμικό αλγόριθμων Μηχανικής Μάθησης WEKA και η δημιουργία προβλεπτικών μοντέλων. Το WEKA διαθέτει ευρύ σύνολο αλγόριθμων για τη δημιουργία προβλεπτικών μοντέλων.

### 9.6.1. Αλγόριθμοι ταξινόμησης

Δεδομένου ότι μελετάμε «προβλήματα ταξινόμησης» (βλ. Κεφ. 9.5.9) αξιοποιήσαμε τους βασικούς αλγόριθμους ταξινόμησης που διαθέτει το WEKA. Πρόκειται για τους αλγόριθμους, οι οποίοι συνήθως χρησιμοποιούνται για την κατασκευή προβλεπτικών μοντέλων και οι οποίοι αναφέρονται σε σχετικές έρευνες (Buschetto Macarini, Cechinel, Batista Machado, Faria Culmant Ramos, & Munoz, 2019), (Howard, Meehan, & Parnell, 2018), (Marbouti, Diefes-Dux, & Madhavan, 2016), (Quinn & Gray, 2020) και βιβλιογραφικές ανασκοπήσεις (Albreiki, Zaki, & Alashwal, 2021), (Baashar, Alkawsji, Ali, Alhussian, & Bahbouh, 2021), (Cui, Chen, Shiri, & Fan, 2019), (Rastrollo-Guerrero, Gomez-Pulido, & Duran-Dominguez, 2020) και οι οποίοι εμφάνισαν καλές επιδόσεις κατά την εφαρμογή τους στα σύνολα δεδομένων. Εκτός από τους βασικούς αλγόριθμους ταξινόμησης και προκειμένου να μελετήσουμε τα τριδιάστατα σύνολα δεδομένων (βλ. Κεφ. 9.5.1.) χρησιμοποιήσαμε και αλγόριθμους Περιοδικών Νευρωνικών Δικτύων (Chen & Cui, 2020).

Οι 21 αλγόριθμοι που χρησιμοποιήθηκαν (Πίνακας 21) ανήκουν στις παρακάτω κατηγορίες:

- Bayesian ταξινομητές, που βασίζονται σε πιθανότητες και χρησιμοποιούν το θεώρημα του Bayes (Han, Kamber, & Pei, 2011, σ. 291).
- Αλγόριθμοι συναρτήσεων, οι οποίοι μπορούν να εκφραστούν ως μαθηματικές εξισώσεις (Witten, Frank, & Hall, 2011, σ. 459).
- Οκνηροί (lazy) αλγόριθμοι, οι οποίοι δεν ξεκινούν την κατασκευή του μοντέλου κατά την τροφοδότησή τους με τις οντότητες από το σύνολο εκμάθησης (training dataset), αλλά απλώς τις αποθηκεύουν, έως ότου τροφοδοτηθούν με οντότητες από το σύνολο δοκιμών (test dataset), οπότε κατασκευάζουν το μοντέλο (Han, Kamber, & Pei, 2011, σ. 348).
- Αλγόριθμοι κανόνων, που βασίζονται σε κανόνες που αποτελούνται από δύο μέρη. Το πρώτο μέρος είναι η «προϋπόθεση» (precondition), δηλαδή σειρές ελέγχων, όπως π.χ.

οι έλεγχοι στα δέντρα αποφάσεων. Το δεύτερο μέρος είναι το «συμπέρασμα», το οποίο υποδεικνύει την κλάση στην οποία ταξινομείται η οντότητα (Witten, Frank, & Hall, 2011, σ. 67).

- Δέντρα αποφάσεων, που η λειτουργία τους αναπαρίσταται με δενδρική δομή, όπου κάθε εσωτερικός κόμβος δηλώνει έλεγχο επί ενός χαρακτηριστικού, κάθε κλαδί αναπαριστά το αποτέλεσμα του ελέγχου και κάθε φύλλο (τελικός κόμβος) αντιστοιχεί στην τιμή της κλάσης (Han, Kamber, & Pei, 2011, σ. 274).
- Περιοδικά Νευρωνικά Δίκτυα (Recursive Neural Networks), υποκατηγορία των νευρωνικών δικτύων, τα οποία τροφοδοτούνται με χρονοσημασμένα δεδομένα, οπότε είναι κατάλληλα για την αναπαράσταση ακολουθιών δεδομένων και γενικά δεδομένων που σχετίζονται χρονικά (Wang, Sy, Liu, & Piech, 2017).

**Πίνακας 21. Αλγόριθμοι ταξινόμησης**

Κατηγορία	Αλγόριθμος	Συνοπτική Περιγραφή <sup>14</sup>
Bayesian Αλγόριθμοι	BayesNet	Αλγόριθμος εκμάθησης Bayes Δικτύου που χρησιμοποιεί διάφορους αλγόριθμους αναζήτησης και μέτρα ποιότητας. Παρέχει δομές δεδομένων (δομή δικτύου, κατανομές πιθανοτήτων υπό όρους, κ.λπ.) και δυνατότητες κοινές στους αλγόριθμους εκμάθησης του Bayes Δικτύου όπως οι K2 & B.
	NaiveBayes	Αλγόριθμος Naive Bayes που χρησιμοποιεί εκτιμητή κλάσεων. Οι αριθμητικές τιμές του εκτιμητή επιλέγονται με ανάλυση των δεδομένων εκμάθησης.
	NaiveBayes Multinomial	Αλγόριθμος για την κατασκευή και χρήση πολυωνυμικού Naive Bayes ταξινομητή.
Αλγόριθμοι συναρτήσεων	Logistic	Αλγόριθμος για κατασκευή και χρήση πολυωνυμικών λογιστικών παλινδρομικών μοντέλων με εκτιμητή κορυφών.
	Multilayer Perceptron	Αλγόριθμος που χρησιμοποιεί οπισθόδρομη διάδοση για την εκμάθηση πολύ-επίπεδου δικτύου τύπου perceptron.
	SimpleLogistic	Αλγόριθμος για την κατασκευή και χρήση γραμμικών λογιστικών παλινδρομικών μοντέλων με χρήση της συνάρτησης logiboost.

<sup>14</sup> Η συνοπτική περιγραφή βασίστηκε στο αντίστοιχο πεδίο που συνοδεύει κάθε αλγόριθμο του WEKA (έκδοση 3.9.6). Αναλυτικότερες περιγραφές είναι διαθέσιμες στα (Frank, Hall, & Witten, The WEKA Workbench (4th ed.), 2016), (Han, Kamber, & Pei, 2011), (Witten, Frank, & Hall, 2011)

Οκνηροί (lazy) αλγόριθμοι	IBk	Αλγόριθμος που βασίζεται στις K-πλησιέστερες γειτονικές τιμές.
	KStar	Αλγόριθμος που σχετίζει την κλάση κάποιας οντότητας από το σύνολο δοκιμών με κλάσεις παρόμοιων οντοτήτων από το σύνολο εκμάθησης.
	LWL	Αλγόριθμος που αποδίδει βαρύτητα στις οντότητες και αναλόγως ταξινομεί σε κλάσεις.
Αλγόριθμοι κανόνων	ZeroR	Αλγόριθμος που προβλέπει τον μέσο όρο (σε αριθμητικές κλάσεις) ή την πολυπληθέστερη τιμή (σε κατηγορικές κλάσεις).
	DecisionTable	Αλγόριθμος που βασίζεται σε απλό πίνακα αποφάσεων.
	JRip	Αλγόριθμος που υλοποιεί εκμάθηση προτασιακών (propositional) κανόνων με επαναλαμβανόμενες και αυξανόμενες περικοπές.
	OneR	Αλγόριθμος που χρησιμοποιεί την ιδιότητα ελάχιστου σφάλματος για την πρόβλεψη και διακριτοποίηση (discretizing) ιδιοτήτων.
	PART	Αλγόριθμος που παράγει λίστες αποφάσεων βασιζόμενος στη στρατηγική «διαίρει και βασίλευε».
Δέντρα αποφάσεων	DecisionStump	Αλγόριθμος που χρησιμοποιεί δέντρο ενός επιπέδου, δηλαδή με μόνο έναν εσωτερικό κόμβο.
	HoeffdingTree	Αλγόριθμος που χρησιμοποιεί αυξανόμενα δέντρα για την ανάλυση μαζικών ροών δεδομένων.
	J48	Αλγόριθμος για τη δημιουργία ακέραιων ή περικομμένων δέντρων.
	RandomForest	Αλγόριθμος που κατασκευάζει δάσος από τυχαία δέντρα αποφάσεων.
	RandomTree	Αλγόριθμος που κατασκευάζει δέντρο με K τυχαίως επιλεγθείσες ιδιότητες σε κάθε κόμβο.
	REPTree	Αλγόριθμος για την κατασκευή δέντρων με τη χρήση αριθμητικών αποκλίσεων και με περικοπή για περιορισμό σφαλμάτων.
Περιοδικά Νευρωνικά Δίκτυα	RNNSequence Classifier	Αλγόριθμος για ταξινόμηση και παλινδρόμηση με περιοδικά νευρωνικά δίκτυα. Κατά την παραμετροποίησή του αλγόριθμου ενεργοποιήσαμε επίπεδο (layer) για εφαρμογή Long Short-Term Memory Network

### 9.6.2. Κριτήριο απόδοσης και τιμή βάσης

Ως κύριο κριτήριο απόδοσης των μοντέλων χρησιμοποιήθηκε η *ακρίβεια πρόβλεψης* (accuracy). Πρόκειται για το κριτήριο, που περισσότερο συχνά (Cui, Chen, Shiri, & Fan, 2019, σ. 218) χρησιμοποιούν οι σχετικές έρευνες. Ως ακρίβεια πρόβλεψης ορίζεται ο λόγος του πλήθους των ορθώς ταξινομημένων οντοτήτων, προς το συνολικό πλήθος των οντοτήτων (Witten, Frank & Hall, 2011, σ. 72).

$$\text{ακρίβεια πρόβλεψης} = \frac{\text{πλήθος ορθώς ταξινομημένων οντοτήτων}}{\text{συνολικό πλήθος οντοτήτων}}$$

Πρόκειται για αριθμητικό μέγεθος που κυμαίνεται στο διάστημα [0 , 1] και το οποίο στα πλαίσια της διατριβής εκφράζεται με τη μορφή ποσοστού επί τοις εκατό (%).

Προκειμένου να αξιολογηθούν τα μοντέλα, είναι απαραίτητο να οριστεί *τιμή βάσης* για την ακρίβεια πρόβλεψης. Γενικά, τα μοντέλα κρίνονται αποδεκτά, αν η ακρίβεια πρόβλεψής τους υπερβαίνει την τιμή βάσης. Επιπλέον, όσο μεγαλύτερη είναι η απόσταση από την τιμή βάσης, τόσο καλύτερα θεωρούνται τα μοντέλα. Τα μοντέλα που η ακρίβεια πρόβλεψής τους δεν υπερβαίνει την τιμή βάσης απορρίπτονται. Δηλαδή η τιμή βάσης αποτελεί τον «πήχη» που τα μοντέλα πρέπει να υπερβούν. Στα πλαίσια της διατριβής ορίσαμε αυστηρότερο κριτήριο για την αποδοχή των μοντέλων. Συγκεκριμένα, για να κριθεί ένα μοντέλο αποδεκτό δεν αρκεί να υπερβεί απλώς την τιμή βάσης, αλλά πρέπει να την υπερβεί με στατιστικώς σημαντική διαφορά.

Ως *τιμή βάσης* για την ακρίβεια πρόβλεψης (baseline accuracy), ορίζεται η ακρίβεια πρόβλεψης την οποία επιτυγχάνει το μοντέλο το οποίο προβλέπει «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Για παράδειγμα, αν σε ένα σύνολο 100 φοιτητών έχουμε 65 επιτυχόντες και 35 αποτυχόντες, τότε το εν λόγω μοντέλο κατατάσσει «τυφλά» όλους τους επιτυχόντες στην κλάση «Επιτυχόντες» και ομοίως όλους τους αποτυχόντες στην κλάση «Επιτυχόντες». Με αυτόν τον τρόπο, επιτυγχάνει ακρίβεια πρόβλεψης 65%, «τυφλά», χωρίς επί της ουσίας να αναλύσει τα χαρακτηριστικά των οντοτήτων. Η τιμή βάσης, αναφέρεται επίσης στη βιβλιογραφία και ως «*τιμή μη-πληροφορίας*» (no-information rate) (Quinn & Gray, 2020, σ. 8). Στο λογισμικό WEKA, η *τιμή βάσης* ή *τιμή μη-πληροφορίας* προκύπτει από τη χρήση του αλγόριθμου ZeroR (Witten, Transcript for Class2 -



Lesson 6, 2013), οπότε με την ακρίβεια πρόβλεψης του ZeroR μοντέλου συγκρίνονται τα μοντέλα που παράγονται με τη χρήση των άλλων αλγόριθμων.

### 9.6.3. Επιπλέον κριτήρια για την απόδοση των μοντέλων

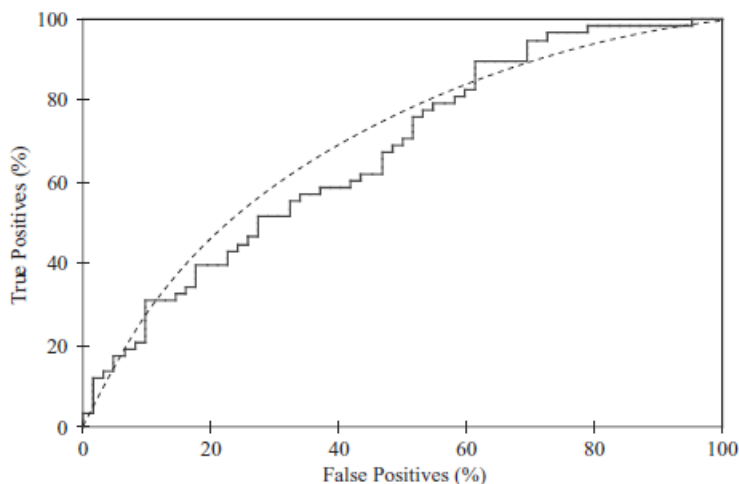
Για κάθε μοντέλο, εκτός από την ακρίβεια πρόβλεψης, την οποία χρησιμοποιήσαμε ως βασικό κριτήριο αξιολόγησης για την απόδοση των μοντέλων, υπολογίστηκαν επίσης οι τιμές της μεταβλητής AUC (Area Under Curve) (Witten, Frank & Hall, 2011, σ. 172) και της μεταβλητής Cohen's Kappa (Witten, Frank, & Hall, 2011, σ. 166). Πρόκειται για μεταβλητές, οι οποίες επίσης χρησιμοποιούνται για την αξιολόγηση της απόδοσης των μοντέλων.

Η μεταβλητή AUC αξιολογεί την διαγνωστική ικανότητα ενός μοντέλου ταξινόμησης με διχοτομημένη κλάση και μη ισορροπημένο (Hussain M. , Zhu, Zhang, & Abidi, 2018) σύνολο δεδομένων. Υπολογίζει την περιοχή που ορίζει η καμπύλη ROC (Receiver Operating Characteristic) (Σχήμα 13), η οποία σχηματίζεται σε σύστημα συντεταγμένων, όπου στον Y άξονα αποτυπώνεται ο δείκτης (rate) των αληθώς θετικών προβλέψεων<sup>15</sup> και στον X άξονα αποτυπώνεται ο δείκτης (rate) των ψευδώς θετικών προβλέψεων<sup>16</sup> (Cui, Chen, Shiri & Fan, 2019, σ. 218). Η μεταβλητή AUC λαμβάνει τιμές στο διάστημα [0 , 1]. Τιμές που ανήκουν στο διάστημα [0.5 , 0.69] κρίνονται αποδεκτές, στο διάστημα [0.70 , 0.79] καλές, στο διάστημα [0.8 , 0.89] εξαιρετικές και στο διάστημα [0.90 , 1] ιδανικές (Riestra-Gonzalez, del Puerto Paule-Ruiz, & Ortin, 2021). Η μεταβλητή AUC είναι χρήσιμη στα μη ισορροπημένα σύνολα, διότι εκεί ένα μοντέλο μπορεί να πετύχει υψηλή τιμή ακρίβειας, ταξινομώντας απλώς στην πολυπληθέστερη κλάση. Στη περίπτωση αυτή, εκτός από την ακρίβεια πρόβλεψης πρέπει να ληφθεί υπόψιν και η τιμή της μεταβλητής AUC (Chen & Cui, 2020, σ. 7. Jayaprakash, Moody, Lauria, Regan & Baron, 2014, σ. 22).

---

<sup>15</sup> Αληθώς Θετικές / (Αληθώς Θετικές + Ψευδώς Αρνητικές)

<sup>16</sup> Ψευδώς Θετικές / (Ψευδώς Θετικές + Αληθώς Αρνητικές)



**Σχήμα 13. Παράδειγμα καμπύλης ROC (Witten, Frank & Hall, 2011, p. 173)**

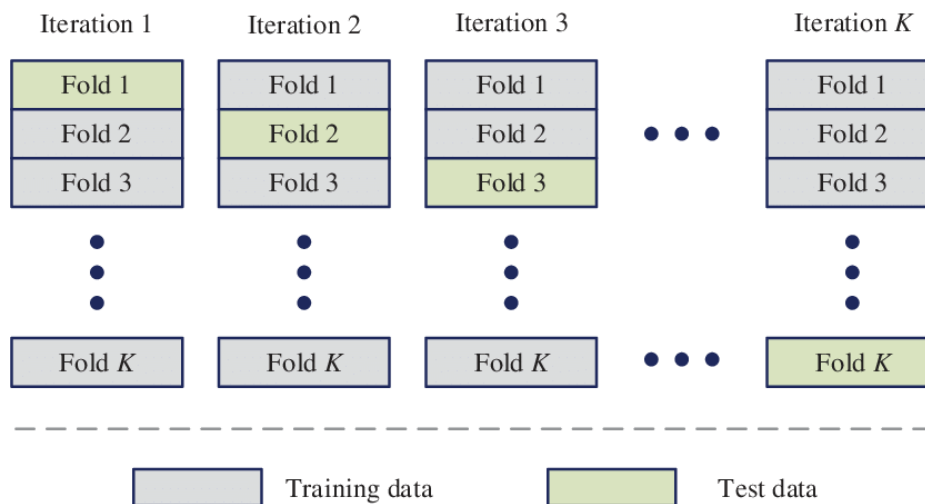
Η μεταβλητή Cohen's Kappa μετρά τη συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που προέβλεψε το μοντέλο, ενώ διορθώνει τυχόν συμφωνία που προέκυψε τυχαίως (Witten, Frank, & Hall, 2011, σ. 166). Η μεταβλητή λαμβάνει τιμές στο διάστημα  $[0, 1]$ . Τιμές που ανήκουν στο διάστημα  $[0, 0.2]$  δηλώνουν «αδύναμη» (slight) συμφωνία, στο διάστημα  $[0.21, 0.40]$  «δίκαιη» (fair) συμφωνία, στο διάστημα  $[0.41, 0.60]$  «μέση» (moderate) συμφωνία, στο διάστημα  $[0.61, 0.80]$  σημαντική συμφωνία και στο διάστημα  $[0.81, 1]$  σχεδόν τέλεια συμφωνία (Quinn & Gray, 2020, σ. 8).

#### 9.6.4. Το ζήτημα της «υπερπροσαρμογής»

Όπως αναφέρουν οι Howard, Meehan & Parnell (2018, σ. 68), ένα από τα βασικά ζητήματα στην κατασκευή προβλεπτικών μοντέλων είναι η «υπερπροσαρμογή» (overfitting), δηλαδή η υπερβολική προσαρμογή του μοντέλου στα δεδομένα του συνόλου εκμάθησης, που ως συνέπεια έχει την ανεπαρκή απόδοση του μοντέλου, όταν αυτό εφαρμοστεί σε νέα δεδομένα δοκιμών. Η υπερπροσαρμογή εμφανίζεται, συνήθως, στις περιπτώσεις όπου όλο το σύνολο δεδομένων χρησιμοποιείται για τη φάση εκμάθησης του μοντέλου.

Διαδεδομένη μέθοδος για την αντιμετώπιση του ζητήματος είναι η «διασταυρούμενη επικύρωση  $k$  πτυχών» ( $k$  fold cross validation). Κατά τη μέθοδο, το σύνολο δεδομένων χωρίζεται

σε  $k$  τυχαία ισομεγέθη υποσύνολα (ή πτυχές). Τα δεδομένα από τα  $k-1$  υποσύνολα χρησιμοποιούνται για την φάση της εκμάθησης του μοντέλου, ενώ η φάση της δοκιμής υλοποιείται επί των δεδομένων του  $k$  υποσυνόλου. Η διαδικασία επαναλαμβάνεται  $k$  φορές, με διαφορετικό κάθε φορά διαφορετικό συνδυασμό υποσυνόλων στη φάση της εκμάθησης και διαφορετικό υποσύνολο στη φάση της δοκιμής (Σχήμα 14). Κατά τη δημιουργία των μοντέλων της διατριβής μας, η διασταυρούμενη επικύρωση  $k$  πτυχών, υλοποιήθηκε με 10 επαναλήψεις ( $k=10$ ), τον θεωρούμενο βέλτιστο αριθμό επαναλήψεων (Witten, Frank & Hall, 2011, σ. 153) .



**Σχήμα 14. Διασταυρούμενη επικύρωση  $k$  πτυχών (Ren, Li & Han, 2019)**

Επιπλέον, για την επίτευξη της μέγιστης δυνατής αξιοπιστίας, αξιοποιώντας προεπιλεγμένη (by default) δυνατότητα του WEKA (Bouckaert, et al., 2020, p. 64), (Witten, Transcript for Class2 - Lesson 5, 2013), εφαρμόστηκε βελτιωμένη παραλλαγή της μεθόδου. Συγκεκριμένα, εφαρμόστηκε στρωματοποιημένη διασταυρούμενη επικύρωση  $k$  πτυχών (stratified  $k$  fold cross validation). Δηλαδή, η δειγματοληψία για τη δημιουργία των υποσυνόλων (πτυχών) έγινε, όχι με τυχαίο τρόπο, αλλά με στρωματοποίηση, δηλαδή τα ποσοστά των κλάσεων στα υποσύνολα ήταν ίδια με το ποσοστό των κλάσεων στο σύνολο των δεδομένων (Witten, Frank, & Hall, 2011, pp. 152-154).

#### 9.6.5. Ικανότητα των μοντέλων για γενίκευση.

Η κατασκευή των μοντέλων περιλαμβάνει τη φάση της εκμάθησης και τη φάση δοκιμής. Και οι δύο φάσεις υλοποιούνται σε υποσύνολα του ίδιου συνόλου δεδομένων (βλ. Κεφ. 9.6.4.). Όμως, η αξία ενός μοντέλου έγκειται στην ικανότητά του για γενίκευση, δηλαδή στην ικανότητά του να επιτυγχάνει πρόβλεψη, όχι μόνο όταν εφαρμόζεται στο «οικείο» σύνολο δεδομένων, αλλά και όταν εφαρμόζεται σε «άγνωστο» σύνολο δεδομένων (Gasevic, Dawson, Rogers & Gasevic, 2016. López-Zambrano, Lara & Romero, 2020).

Για να μελετήσουμε την ικανότητα γενίκευσης των μοντέλων που παράγουν οι αλγόριθμοι, εφαρμόσαμε το πιο «επιτυχημένο» από τα παραγόμενα μοντέλα σε σύνολα δεδομένων άλλου μαθήματος και αναλύσαμε την απόδοσή του, αφενός ως προς την ακρίβεια πρόβλεψης και αφετέρου ως προς τη χρονική στιγμή της πρόβλεψης.

#### 9.6.6. Χρήση του WEKA Experimenter έναντι του WEKA Explorer

Το λογισμικό WEKA διαθέτει δύο εργαλεία για την εφαρμογή αλγόριθμων επί των συνόλων δεδομένων, την κατασκευή μοντέλων και την εμφάνιση τιμών σχετικά με την ακρίβεια της πρόβλεψης. Το εργαλείο Explorer και το εργαλείο Experimenter. Χρησιμοποιήσαμε το εργαλείο Experimenter, διότι καταρχάς, δίνει δυνατότητα ταυτόχρονης και συγκριτικής εμφάνισης αποτελεσμάτων από πολλά σύνολα δεδομένων και από πολλούς αλγόριθμους. Κυρίως όμως διότι πληροφορεί αν οι διαφορές τιμών στην ακρίβεια πρόβλεψης είναι (ή δεν είναι) στατιστικά σημαντικές στο επίπεδο στατιστικής σημαντικότητας 0.05 (2-tailed). Συνεπώς, κατά την παρουσίαση των αποτελεσμάτων, όταν συγκρίνουμε την ακρίβεια ενός μοντέλου με την τιμή βάσης (baseline accuracy), είμαστε σε θέση να επιλέξουμε μόνο εκείνα τα μοντέλα, που η ακρίβειά τους υπερβαίνει την τιμή βάσης με στατιστικώς σημαντική διαφορά (βλ. Κεφ. 9.6.2).

#### 9.7. Εντοπισμός επιδραστικών χαρακτηριστικών

Για τον εντοπισμό των επιδραστικών χαρακτηριστικών των μοντέλων, δηλαδή εκείνων των χαρακτηριστικών, τα οποία περισσότερο από τα υπόλοιπα, επηρεάζουν την απόδοσή των μοντέλων χρησιμοποιήσαμε μέθοδο, η οποία προσφέρεται από το λογισμικό WEKA.

Συγκεκριμένα, χρησιμοποιήσαμε τη δυνατότητα εντοπισμού χαρακτηριστικών με τη μέθοδο «CorrelationAttributeEval» (Select Attributes > Attribute Evaluator = CorrelationAttributeEval). Η συγκεκριμένη μέθοδος αποτιμά την επιδραστικότητα / βαρύτητα κάθε χαρακτηριστικού υπολογίζοντας την τιμή του συντελεστή συσχέτισης (Pearson  $r$ ) μεταξύ του χαρακτηριστικού και της κλάσης.

Εφαρμόσαμε τη μέθοδο για τα σύνολα δεδομένων, τα οποία αντιστοιχούν στην περίοδο από την αρχή μέχρι το τέλος των μαθημάτων. Δεν υλοποιήσαμε ανάλογη ενέργεια σε σύνολα δεδομένων που κάλυπταν μικρότερες περιόδους, διότι στην αναζήτηση επιδραστικών χαρακτηριστικών δεν ενδιαφέρει η πρώιμη πρόβλεψη, ούτε ενδιαφέρει η επίδραση χαρακτηριστικών σε πρώιμες χρονικές στιγμές. Ενδιαφέρει αποκλειστικά, η επίδραση χαρακτηριστικών για όλη την περίοδο διεξαγωγής των μαθημάτων. Από τα χαρακτηριστικά που υπέδειξε η μέθοδος, επικεντρωθήκαμε σε αυτά με τους 5 μεγαλύτερους συντελεστές συσχέτισης.

Επειδή, ούτε το λογισμικό, ούτε το υλικό τεκμηρίωσης του WEKA, παρέχουν πληροφορίες για τη στατιστική σημαντικότητα (significance) των υπολογιζόμενων συντελεστών συσχέτισης, υλοποιήσαμε τις παραπάνω ενέργειες και με τη χρήση του λογισμικού στατιστικής ανάλυσης IBM SPSS. Προέκυψε ότι οι συντελεστές συσχέτισης που υπολογίζουν το WEKA και το IBM SPSS είναι ακριβώς ίδιοι και ότι υπολογίστηκαν στο επίπεδο στατιστικής σημαντικότητας 0.01 (2-tailed).

Δεδομένου ότι, αφενός ο συντελεστής συσχέτισης (Pearson  $r$ ) εφαρμόζεται σε αριθμητικές μεταβλητές (Ρούσσοι & Τσαούσης, 2002, σ. 165) και αφετέρου η κλάση στα σύνολα δεδομένων που χρησιμοποιούμε είναι κατηγορική μεταβλητή, εγείρεται το ζήτημα αν ενδείκνυται η χρήση της συγκεκριμένης μεθόδου. Σύμφωνα με τον Frank (2018a), εκ των συνδημιουργών του λογισμικού WEKA, η μέθοδος «CorrelationAttributeEval» μπορεί πράγματι να εφαρμοσθεί σε σύνολα δεδομένων με κατηγορικές κλάσεις. Ειδικότερα, η μέθοδος αποδίδει μία δυαδική αριθμητική τιμή σε κάθε κατηγορική τιμή της κλάσης και ακολούθως υπολογίζει τον συντελεστή συσχέτισης (Pearson  $r$ ) μεταξύ της κλάσης και του κάθε χαρακτηριστικού στο σύνολο δεδομένων (Frank, 2018 b). Επιπλέον, η μέθοδος μπορεί να εφαρμοσθεί ακόμα και στην περίπτωση (που δεν συναντάται στα σύνολα μας), κατά την οποία κάποιο από τα χαρακτηριστικά εκφράζεται με

κατηγορικές και όχι με αριθμητικές τιμές. Στην περίπτωση αυτή, οι  $k$  διαφορετικές τιμές του χαρακτηριστικού μετατρέπονται σε  $k$  διαφορετικούς αριθμητικούς δυαδικούς δείκτες (indicators), όπου ο καθένας λαμβάνει την τιμή 1 στην περίπτωση που η αντίστοιχη κατηγορική μεταβλητή εμφανίζεται στην τρέχουσα οντότητα, ενώ λαμβάνει την τιμή 0 στην αντίθετη περίπτωση. Η μέθοδος αντιμετωπίζει τους συγκεκριμένους δείκτες ως αριθμητικές μεταβλητές και με αυτόν τον τρόπο υπολογίζει τον συντελεστή συσχέτισης (Pearson  $r$ ) μεταξύ του χαρακτηριστικού και τη κλάσης. Ακολούθως, ο συνολικός συντελεστής συσχέτισης προκύπτει από τον σταθισμένο μέσο όρο των συντελεστών συσχέτισης κάθε δείκτη, όπου η στάθμιση είναι ανάλογη της συχνότητας κάθε κατηγορικής τιμής του χαρακτηριστικού (Hall, 2017). Αξίζει επίσης να επισημανθεί ότι η μέθοδος «CorrelationAttributeEval» χρησιμοποιείται ευρέως σε σύνολα δεδομένων με κατηγορικές κλάσεις, όπως προκύπτει από σχετικές βιβλιογραφικές αναφορές (Brownlee, 2019b. Gnanambal, Thangaraj, Meenatchi & Gayathri, 2018. Hussain, Dahan, Ba-Alwib & Ribata, 2018. Pehlivanova & Nedeva, 2021).

Αφού, με βάση τα παραπάνω, προκύπτει ότι η μέθοδος είναι έγκυρη και διαδεδομένη, επιχειρήσαμε επιπλέον να αποτιμήσουμε στην πράξη την ευστοχία της. Δηλαδή, να διαπιστώσουμε, αν τα χαρακτηριστικά που υποδεικνύει η μέθοδος ήταν πράγματι επιδραστικά στην απόδοση των μοντέλων. Οπότε, υλοποιώντας σχετικά πειράματα με τη χρήση του λογισμικού Weka Explorer, διερευνήσαμε αν η συμπερίληψη, ή η μη συμπερίληψη, των χαρακτηριστικών τα οποία υπέδειξε η μέθοδος, επηρεάζει την ακρίβεια πρόβλεψης των μοντέλων.

## 10. ΑΝΑΛΥΣΗ ΕΡΕΥΝΗΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

Όπως αναφέρεται στο Κεφάλαιο 9.6, το λογισμικό αλγόριθμων Μηχανικής Μάθησης WEKA αξιοποιήθηκε για τη δημιουργία προβλεπτικών μοντέλων. Οι αλγόριθμοι εφαρμόστηκαν όχι μόνο επί των βασικών συνόλων δεδομένων κάθε μαθήματος, δηλαδή όχι μόνο επί των συνόλων που κάλυπταν όλες τις εβδομάδες, από την αρχή έως το τέλος κάθε μαθήματος. Εφαρμόστηκαν και σε σύνολα δεδομένων που κάλυπταν μικρότερες χρονικές περιόδους και συγκεκριμένα τη χρονική περίοδο της 1<sup>ης</sup> εβδομάδας, τη χρονική περίοδο από την 1<sup>η</sup> ως τη 2<sup>η</sup> εβδομάδα, τη χρονική περίοδο από την 1<sup>η</sup> ως την 3<sup>η</sup> εβδομάδα κ.ο.κ. Με τον τρόπο αυτό, μελετήθηκαν οι διακυμάνσεις στην απόδοση των προβλεπτικών μοντέλων σε συνάρτηση με τη χρονική στιγμή κατά την οποία έγινε η πρόβλεψη. Συνολικά μελετήθηκαν **103** σύνολα δεδομένων (βλ. Κεφ. 9.5.2 Πίνακας 18) και χρησιμοποιήθηκαν **21** αλγόριθμοι Μηχανικής Μάθησης (βλ. Κεφ. 9.6.1 Πίνακας 21). Το εργαλείο Weka Experimenter προτιμήθηκε έναντι του Weka Explorer, αφενός διότι παρέχει τη δυνατότητα εφαρμογής πολλών αλγόριθμων σε πολλά σύνολα δεδομένων ταυτοχρόνως και αφετέρου διότι, όχι απλώς συγκρίνει τις αποδόσεις των αλγόριθμων, αλλά επιπλέον παρέχει την πληροφορία για τον αν οι διαφορές στις τιμές της ακρίβειας πρόβλεψης είναι στατιστικώς σημαντικές.

Ακολουθεί αναλυτική παρουσίαση των αποτελεσμάτων από την εφαρμογή των αλγόριθμων σε όλα τα σύνολα δεδομένων όλων μαθημάτων. Οι βασικοί αλγόριθμοι εφαρμόστηκαν με τις προκαθορισμένες (default) από το WEKA παραμέτρους. Προκειμένου να αντιμετωπιστεί το ζήτημα της «υπερπροσαρμογής», εφαρμόστηκε στρωματοποιημένη διασταυρούμενη επικύρωση k πτυχών (βλ. Κεφ. 9.6.4).

## 10.1. Τα μοντέλα πρόβλεψης για το Μάθημα AA005 2020-2021

### 10.1.1. Ακρίβεια και χρόνος πρόβλεψης των βασικών μοντέλων

Στον Πίνακα 22 εμφανίζονται οι τιμές για την ακρίβεια πρόβλεψης των μοντέλων που κατασκεύασαν 20 βασικοί αλγόριθμοι σε 18 σύνολα δεδομένων του μαθήματος AA005 Χρήση Βασικών Εφαρμογών ΗΥ, χειμερινό εξάμηνο 2020-2021



**Πίνακας 22. Ακρίβεια πρόβλεψης μοντέλων για το μάθημα ΑΑ005 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
Αλγορ.01	73.19	74.9	74.52	74.52	73.04	72.58	72.58	72.12	70.75	70.75	70.75	70.75	69.83	69.83	69.83	69.83	69.83	69.83
Αλγορ.02	71.07	71.78	72.89	74.05	76.25	76.3	75.71	75.75	76.29	80.06 v	78.72 v	80.44 v	81.90 v	78.95 v	80.21 v	80.90 v	80.08 v	84.29 v
Αλγορ.03	53.36 *	67.21 *	69.18	69.51	75.91	75.77	73.99	73.39	76.49	77.63	79.18 v	79.81 v	79.26 v	82.83 v	81.91 v	83.73 v	79.84 v	80.63 v
Αλγορ.04	73.32	74.84	74.52	74.11	74.22	73.57	79.33 v	78.69 v	79.94 v	80.62 v	81.07 v	81.65 v	81.40 v	83.35 v	82.67 v	81.85 v	81.03 v	82.21 v
Αλγορ.05	67.08	74.66	75.31	76.13	73.67	73.9	74.82	73.51	75.9	76	75.32	77.53	77.89 v	78.94 v	79.57 v	78.64 v	78.38 v	83.35 v
Αλγορ.06	68.97	70.11	69.02	68.83	68.54	67.83	66.74	67.94	68.82	73.33	71.52	75.11	76.83	77.69 v	77.89 v	79.20 v	80.63 v	81.05 v
Αλγορ.07	72.06	79.15	75.8	75.12	76.69	77.3	79.19 v	78.74 v	77.79 v	79.15 v	79.77 v	81.35 v	84.12 v	81.59 v	82.04 v	82.52 v	82.84 v	84.54 v
Αλγορ.08	69.27	70.45	64.65 *	61.16 *	66.51	67.64	66.33	68.15	66.05	69.79	69.74	72.83	74.44	76.71	78.22 v	77.90 v	76.68 v	82.50 v
Αλγορ.09	68.3	70.07	66.43	63.66 *	69.54	71.78	68.93	67.96	66.69	68.39	70.69	71.57	77.52 v	67.28	70.4	72.78	71.26	81.07 v
Αλγορ.10	69.41	78.81	76.15	72.09	74.25	75.89	78.60 v	74.42	75.63	74.66	76.67	83.25 v	83.86 v	76.06	76.19 v	77.25 v	78.31 v	83.86 v
Αλγορ.11	71.95	75.65	76.47	76.08	74.86	76.12	76.77	75.18	75.94	78.80 v	77.15 v	81.97 v	82.48 v	79.12 v	77.95 v	78.26 v	78.94 v	82.16 v
Αλγορ.12	74.71	75.85	73.84	72.54	74.51	73.35	74.84	74.64	76.59 v	75.56	79.35 v	82.44 v	83.30 v	79.01 v	78.17 v	78.26 v	78.50 v	82.09 v
Αλγορ.13	71.01	75.56	71.76	78.17	70.39	73.32	76.02	75.99	72.82	80.38 v	76.16	80.91 v	83.11 v	78.18 v	74.42	74.41	77.23 v	80.15 v
Αλγορ.14	66.92	67.48 *	68.72	67.38	69.14	66.01	68.61	70.93	68.34	75.15	72.49	74.57	75.18	75.05	76.74 v	77.32 v	75.25	78.31 v
Αλγορ.15	67.98	74.83	78.03	74.23	71.39	77.36	79.25 v	75.55	77.22 v	73.95	77.93 v	83.57 v	83.86 v	76.25 v	75.62 v	77.69 v	78.50 v	83.98 v
Αλγορ.16	73.19	74.28	73.39	74.52	72.1	71.01	71.97	70.28	75.6	78.08	79.11 v	79.24 v	79.07 v	82.08 v	81.78 v	83.03 v	80.09 v	80.32 v
Αλγορ.17	69.69	71.46	72.1	68.93	72.1	68	70.28	71.59	70.6	76.76	74.06	76.15	75.75	75.61	76.74 v	76.32	75.55	78.62 v
Αλγορ.18	75.17	75.23	72.5	72.43	74.55	74.07	75.59	75.31	76.68	80.39 v	79.44 v	84.13 v	84.69 v	82.85 v	83.54 v	84.23 v	83.29 v	85.80 v
Αλγορ.19	67.06	67.62 *	65.66 *	67.09 *	65.23	69.08	69.25	69.51	69.54	70.9	70.63	74.62	75.4	71.18	73.36	72.89	73.55	76.62
Αλγορ.20	70.52	73.62	75.41	74.87	74.33	73.82	77.4	74.9	75.56	76.66 v	78.96 v	80.59 v	81.55 v	77.75 v	78.69 v	77.81 v	77.93 v	80.82 v

Η πρώτη γραμμή του Πίνακα 22 δηλώνει την εβδομάδα του μαθήματος (π.χ. 1<sup>η</sup>, 2<sup>η</sup>, ..., 18<sup>η</sup>) στην οποία ολοκληρώνεται η χρονική περίοδος που καλύπτει το σύνολο δεδομένων και στην οποία γίνεται η πρόβλεψη. Τη 18<sup>η</sup> εβδομάδα (τελευταία στήλη), ολοκληρώνεται η χρονική περίοδος του μαθήματος με τη διεξαγωγή των τελικών εξετάσεων. Οπότε, στην τελευταία στήλη του πίνακα εμφανίζονται οι τιμές για την ακρίβεια πρόβλεψης στο σύνολο δεδομένων της 18<sup>ης</sup> εβδομάδας, δηλαδή στο σύνολο που περιλαμβάνει τα δεδομένα όλου του μαθήματος από την αρχή του (1<sup>η</sup> εβδομάδα) έως το τέλος του (18<sup>η</sup> εβδομάδα).

Στις επόμενες γραμμές του πίνακα εμφανίζονται οι τιμές για την ακρίβεια πρόβλεψης ανά αλγόριθμο. Η ακρίβεια πρόβλεψης ορίζεται ως ο λόγος του πλήθους των ορθώς ταξινομημένων οντοτήτων προς το συνολικό πλήθος των οντοτήτων (Witten, Frank & Hall, 2011, σ. 72) (βλ. Κεφ. 9.6.2). Στον Πίνακα 22, εμφανίζεται με τη μορφή ποσοστού επί τοις εκατό (%). Στη δεύτερη γραμμή του Πίνακα 22 εμφανίζονται οι τιμές για την *τιμή βάσης* της ακρίβειας πρόβλεψης (baseline accuracy). Δηλαδή, για την ακρίβεια πρόβλεψης, που επιτυγχάνει το μοντέλο, το οποίο προβλέπει «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες (εν προκειμένω, οι φοιτητές) ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Στο λογισμικό WEKA, η *τιμή βάσης* ή *τιμή μη-πληροφορίας* υπολογίζεται από το μοντέλο του αλγόριθμου rules.ZeroR (Witten, Transcript for Class2 - Lesson 6, 2013). Οπότε, με την ακρίβεια πρόβλεψης του ZeroR μοντέλου συγκρίνονται οι τιμές των μοντέλων που παράγονται με τη χρήση των άλλων αλγόριθμων ταξινόμησης (βλ. Κεφ. 9.6.1 Πίνακας 21). Η τιμή βάσης δεν είναι ίδια για όλες τις εβδομάδες. Τις τελευταίες 6 εβδομάδες η τιμή βάσης είναι 69.83%, ενώ στην αρχή του μαθήματος είναι μεγαλύτερη κατά 4 έως 5 εκατοστιαίες μονάδες. Η μεταβολή της τιμής βάσης οφείλεται στο ότι το πλήθος των φοιτητών δεν είναι το ίδιο σε όλα τα σύνολα δεδομένων. Ειδικότερα, στα σύνολα δεδομένων των πρώτων εβδομάδων, οι φοιτητές είναι λιγότεροι σε σύγκριση με τους φοιτητές στα σύνολα δεδομένων των τελευταίων εβδομάδων, διότι λόγω της μη υποχρεωτικής χρήσης του ΣΔΜ Moodle, όλοι οι φοιτητές δεν ξεκινούν τη χρήση της πλατφόρμας από την 1<sup>η</sup> εβδομάδα – αρκετοί ξεκινούν αργότερα. Επιπλέον, επειδή οι φοιτητές που επιτυγχάνουν στο μάθημα, συνήθως, ξεκινούν τη χρήση του ΣΔΜ νωρίτερα σε σχέση με αυτούς που αποτυγχάνουν, το ποσοστό των επιτυχόντων φοιτητών είναι μεγαλύτερο τις πρώτες εβδομάδες. Οπότε, η τιμή βάσης, λόγω της εξάρτησής της από το ποσοστό της πολυπληθέστερης κλάσης, είναι αυξημένο

τις πρώτες εβδομάδες και μειώνεται καθώς το μάθημα βαίνει χρονικά προς την ολοκλήρωσή του, όταν καθυστερημένα συνδέονται φοιτητές, οι οποίοι τελικά, συνήθως, αποτυγχάνουν. Η τιμή βάσης σταθεροποιείται τις τελευταίες εβδομάδες του μαθήματος, διότι τις εβδομάδες αυτές σταθεροποιείται το πλήθος φοιτητών που χρησιμοποιούν το ΣΔΜ Moodle.

Στις επόμενες γραμμές του Πίνακα 22 (γραμμή 3, ως γραμμή 21) εμφανίζονται οι τιμές της ακρίβειας πρόβλεψης που επιτυγχάνουν τα μοντέλα που δημιουργήθηκαν από τους βασικούς αλγόριθμους του WEKA (βλ. Κεφ. 9.6.1 Πίνακας 21). Στον Πίνακα 22, αντί των ονομάτων των αλγόριθμων εμφανίζονται αύξοντες αριθμοί. Τα ονόματα των αλγόριθμων εμφανίζονται στον Πίνακα 23, όπου επίσης εμφανίζονται και οι παράμετροι με τις οποίες εκτελέστηκαν οι αλγόριθμοι, οι οποίες είναι οι προκαθορισμένες (default) παράμετροι που προτείνει το WEKA.

**Πίνακας 23. Βασικοί αλγόριθμοι ταξινόμησης για την κατασκευή μοντέλων**

A/A	Αλγόριθμος	Προκαθορισμένες τιμές παραμέτρων εκτέλεσης αλγόριθμου
1	rules.ZeroR	" 48055541465867954
2	bayes.BayesNet	-D -Q bayes.net.search.local.K2 -- -P 1 -S BAYES -E bayes.net.estimate.SimpleEstimator -- -A 0.5' 746037443258775954
3	bayes.NaiveBayes	5995231201785697655
4	bayes.NaiveBayesMultinomial	' 5932177440181257085
5	functions.Logistic	-R 1.0E-8 -M -1 -num-decimal-places 4' 3932117032546553727
6	functions.MultilayerPerceptron	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a' - 5990607817048210779
7	functions.SimpleLogistic	-I 0 -M 500 -H 50 -W 0.0' 7397710626304705059
8	lazy.IBk	-K 1 -W 0 -A "\"weka.core.neighboursearch.LinearNNSearch -A \\\"weka.core.EuclideanDistance -R first-last\\\"\" - 3080186098777067172
9	lazy.KStar	-B 20 -M a' 332458330800479083
10	lazy.LWL	-U 0 -K -1 -A "\"weka.core.neighboursearch.LinearNNSearch -A \\\"weka.core.EuclideanDistance -R first-last\\\"\" -W trees.DecisionStump' 1979797405383665815
11	rules.DecisionTable	-X 1 -S "\"BestFirst -D 1 -N 5\" 2888557078165701326
12	rules.JRip	-F 3 -N 2.0 -O 2 -S 1' -6589312996832147161
13	rules.OneR	-B 6' -3459427003147861443
14	rules.PART	-C 0.25 -M 2' 8121455039782598361

15	trees.DecisionStump	' 1618384535950391
16	trees.HoeffdingTree	-L 2 -S 1 -E 1.0E-7 -H 0.05 -M 0.01 -G 200.0 -N 0.0' 7117521775722396251
17	trees.J48	-C 0.25 -M 2' -217733168393644444
18	trees.RandomForest	-P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1' 1116839470751428698
19	trees.RandomTree	-K 0 -M 1.0 -V 0.001 -S 1' -9051119597407396024
20	trees.REPTree	-M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0' -9216785998198681299

Στον Πίνακα 22, οι τιμές για την ακρίβεια πρόβλεψης έχουν την εξής χρωματική μορφοποίηση: Όσες τιμές υπερβαίνουν την τιμή βάσης (γραμμή 2) έχουν γαλάζιο (σκούρο ή ανοικτό) υπόβαθρο. Αναλυτικότερα, σκούρο γαλάζιο υπόβαθρο έχουν όσες τιμές υπερβαίνουν την τιμή βάσης με στατιστικώς σημαντική διαφορά, με ποσοστό στατιστικής σημαντικότητας 0.05. Σε αυτές τις τιμές, δίπλα στον αριθμό υπάρχει το σύμβολο «v». Όσες τιμές απλώς υπερβαίνουν την τιμή βάσης, χωρίς στατιστικώς σημαντική διαφορά, έχουν ανοικτό γαλάζιο υπόβαθρο. Λευκό υπόβαθρο έχουν οι τιμές που υπολείπονται της τιμής βάσης. Σε όσες από αυτές τιμές, η διαφορά από την τιμή βάσης είναι στατιστικώς σημαντική, δίπλα στον αριθμό υπάρχει το σύμβολο «\*».

#### *Γενική τάση των τιμών της ακρίβειας πρόβλεψης*

Μελετώντας την απόδοση των μοντέλων από την 1<sup>η</sup> προς την 18<sup>η</sup> εβδομάδα του μαθήματος, παρατηρούμε ότι είναι εμφανείς δύο αντίρροπες τάσεις.

Η πρώτη και σημαντικότερη είναι η αύξουσα εξέλιξη των τιμών της ακρίβειας πρόβλεψης, καθώς κινούμαστε χρονικά από την έναρξη προς το τέλος του μαθήματος. Πρόκειται για φαινόμενο αναμενόμενο και ευεξηγήτο. Τις πρώτες εβδομάδες του μαθήματος, τα δεδομένα που περιλαμβάνονται στα σύνολα δεδομένων είναι λιγότερα σε σχέση με τις τελευταίες εβδομάδες. Κατά συνέπεια είναι λιγότερες οι πληροφορίες, οι οποίες προσανατολίζουν τους αλγόριθμους να εντοπίσουν μοτίβα (patterns) χρήσης στο ΣΔΜ Moodle και να τα συσχετίσουν με τις κλάσεις ταξινόμησης. Δηλαδή, στην αρχή του μαθήματος, υφίσταται έλλειμμα πληροφοριών, το οποίο είναι τόσο μεγαλύτερο, όσο μελετούμε σύνολα δεδομένων κοντά στην έναρξη του μαθήματος και το οποίο δυσχεραίνει την «εκμάθηση» (training) των μοντέλων. Κατά συνέπεια, η προβλεπτική ικανότητα των μοντέλων οξύνεται, όταν κινούμαστε χρονικά προς το τέλος του

μαθήματος και αντιστρόφως αμβλύνεται, όταν κινούμαστε προς την αρχή του μαθήματος. Παρότι η μακροπρόθεσμη γενική τάση στην ακρίβεια πρόβλεψης είναι σαφώς αύξουσα, (αν κινούμαστε σύμφωνα με το βέλος του χρόνου), δεν απουσιάζουν μικρές σημειακές «ανωμαλίες», όπου η έως τότε συνεχώς αύξουσα τάση αναστρέφεται στιγμιαία και προσωρινά. Δηλαδή, υπάρχουν χρονικά σημεία, όπου αναπάντεχα, η ακρίβεια πρόβλεψης κατά την εβδομάδα  $W$  είναι μικρότερη από την εβδομάδα  $W-1$ . Δεδομένης της πολυπλοκότητας των αλγόριθμων και των μηχανισμών δημιουργίας των μοντέλων (εγγενές χαρακτηριστικό της Μηχανικής Μάθησης), δεν είναι εύκολο να αναζητηθούν τα αίτια των ανωμαλιών. Αν οι «ανωμαλίες» εμφανίζονται σε λίγα μόνο μοντέλα, τότε είναι δυνατόν να αποδοθούν σε συμπτώσεις και σε τυχαίους ή απροσδιόριστους παράγοντες. Αν εμφανίζονται την ίδια χρονική στιγμή σε πολλά μοντέλα, τότε προκύπτει η υποψία ότι το φαινόμενο δεν είναι τυχαίο και αναδύεται ο ερευνητικός πειρασμός να αναζητηθούν τυχόν αίτια, σχετικά με γεγονότα που αφορούν το μάθημα. Ωστόσο και σε αυτές τις περιπτώσεις, οι ερμηνείες (ή ορθότερα οι εικασίες) ούτε εύκολες είναι, ούτε στέρεες.

Η δεύτερη τάση είναι η φθίνουσα εξέλιξη της τιμής βάσης, αν κινηθούμε χρονικά προς την ολοκλήρωση του μαθήματος. Και αυτό το φαινόμενο, που αναλυτικά περιγράφηκε σε προηγούμενη παράγραφο, είναι αναμενόμενο και ευεξηγήτο. Δεδομένου ότι σχετίζεται με το πλήθος των οντοτήτων στα διαφορετικά σύνολα δεδομένων και το ποσοστό της πολυπληθέστερης κλάσης, το φαινόμενο είναι άσχετο με τη δυσχέρεια «εκμάθησης» των αλγόριθμων, λόγω ελλείμματος πληροφοριών και δεν επηρεάζει την τιμή στην ακρίβεια πρόβλεψης των μοντέλων. Ωστόσο, έχει άλλη σημαντική συνέπεια, η οποία αφορά την αξιολόγηση των μοντέλων. Συγκεκριμένα, επηρεάζει την στατιστική σημαντικότητα της απόδοσης των μοντέλων και κατά συνέπεια την αξιοπιστία τους. Διότι η τιμή βάσης, όπως έχουμε αναφέρει (βλ. Κεφ. 9.6.2) αποτελεί τον «πήχη», που πρέπει οπωσδήποτε να υπερβεί, με στατιστικώς σημαντική διαφορά, η ακρίβεια πρόβλεψης των μοντέλων.

Οι δύο αυτές ταυτόχρονες τάσεις, αφενός το έλλειμμα πληροφορίας κατά τις πρώτες εβδομάδες, που αμβλύνει την απόδοση των μοντέλων και αφετέρου η αύξηση της τιμής βάσης ( του «πήχη» ) καθιστούν αυστηρότερα τα κριτήρια αξιοπιστίας των μοντέλων.

### *Χρονική στιγμή πρόβλεψης*

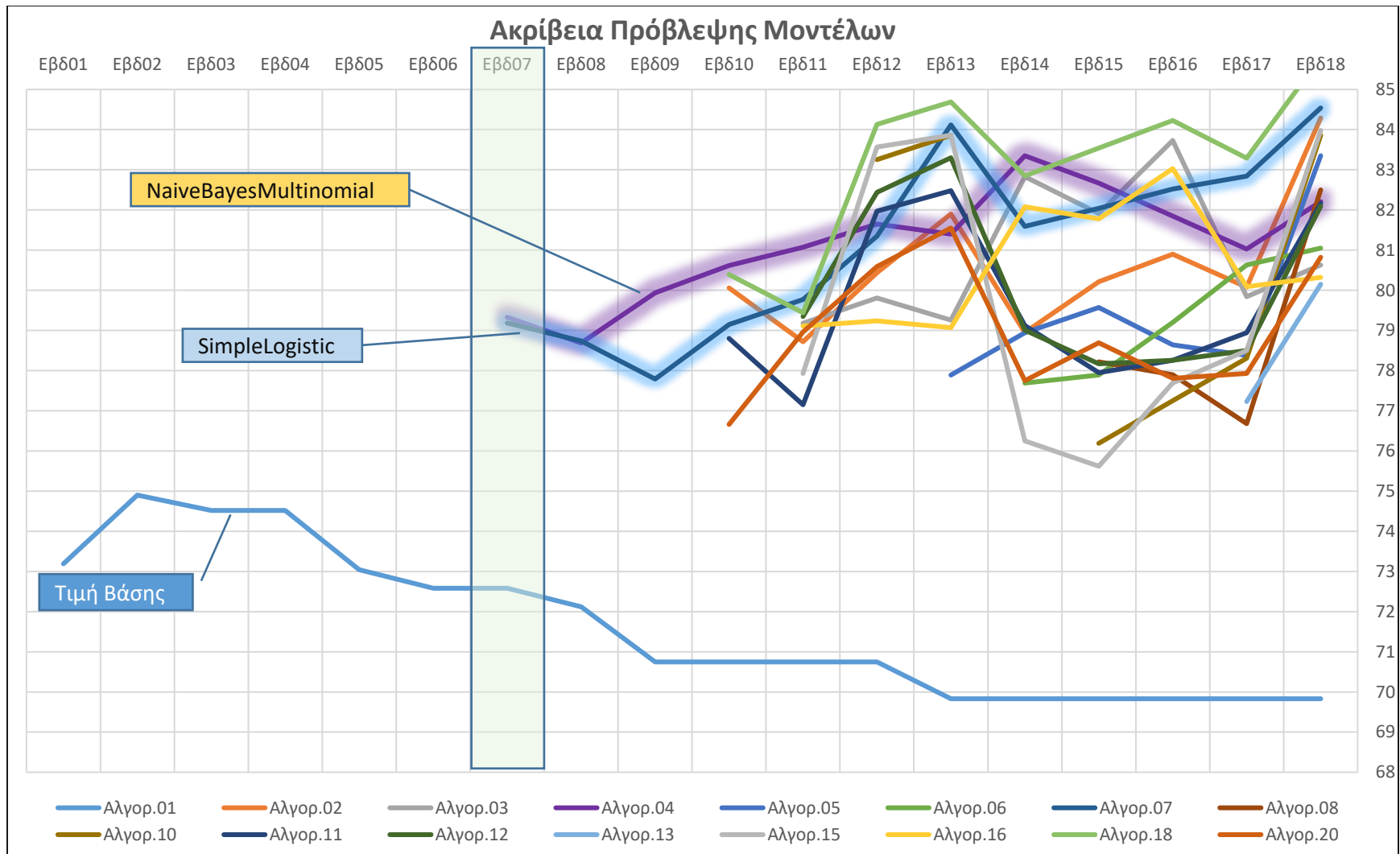
Παρατηρούμε στον Πίνακα 22, ότι την 18<sup>η</sup> εβδομάδα, όταν δηλαδή διαθέτουμε δεδομένα από την έναρξη του μαθήματος (1<sup>η</sup> εβδομάδα) έως την ημέρα της τελικής εξέτασης (18<sup>η</sup> εβδομάδα), όλα τα μοντέλα επιτυγχάνουν υψηλή ακρίβεια πρόβλεψης. Σε όλα τα μοντέλα, με την εξαίρεση ενός (A/A 19), η ακρίβεια πρόβλεψης είναι, όχι απλώς αρκετά μεγαλύτερη από την τιμή βάσης, αλλά η διαφορά της ακρίβειας πρόβλεψης από την τιμή βάσης, είναι στατιστικώς σημαντική. Από τα 19 μοντέλα, τα 16 επιτυγχάνουν τιμές μεγαλύτερες του 80%. Στα υπόλοιπα 3 μοντέλα, η τιμή δεν πέφτει κάτω από το 76,62%. Ωστόσο, αξίζει να σημειωθεί ότι ακόμα και αυτή η τιμή (76.62%), υπερβαίνει την τιμή βάση κατά 7 περίπου εκατοστιαίες μονάδες, που όμως δεν είναι αρκετές για να χαρακτηρίσουν τη διαφορά στατιστικώς σημαντική. Η υψηλότερη ακρίβεια πρόβλεψης είναι 85.80%, διαφέρει 16 εκατοστιαίες μονάδες από την τιμή βάσης και επιτυγχάνεται από το μοντέλο του αλγόριθμου `trees.RandomForest` (A/A 18). Ακολουθούν σε ελάχιστη απόσταση τα μοντέλα του αλγόριθμου `functions.SimpleLogistic` (A/A 7) με 84.54%, του αλγόριθμου `bayes.BayesNet` (A/A 2) με 84.29% και του αλγόριθμου `trees.DecisionStump` (A/A 15) με 83.98%.

Ωστόσο, για τους σκοπούς της έρευνάς μας, η ακρίβεια πρόβλεψης δεν είναι ούτε το μοναδικό, ούτε ίσως το σημαντικότερο κριτήριο για την αξιολόγηση των μοντέλων πρόβλεψης. Όπως αναφέρουμε στο κεφάλαιο 8.1, η έρευνά μας ενδιαφέρεται για μοντέλα που θα είναι, όχι μόνο μαθηματικώς ενδιαφέροντα και τεχνικώς επιτυχημένα, αλλά για μοντέλα που θα είναι και εκπαιδευτικώς χρήσιμα. Η χρησιμότητα των προβλεπτικών εκπαιδευτικών μοντέλων σχετίζεται κυρίως με την δυνατότητά τους να παράγουν πρώιμες προβλέψεις, έστω λιγότερο ακριβείς από τις τελικές. Διότι, με αυτόν τον τρόπο τα προβλεπτικά μοντέλα μπορούν να αξιοποιηθούν ως εργαλεία διαμορφωτικής αξιολόγησης.

Παρατηρούμε στον Πίνακα 22, ότι η ακρίβεια πρόβλεψης κάποιων μοντέλων (A/A 9, 14, 17) χάνει την στατιστική σημαντικότητά της μόλις την 17<sup>η</sup> εβδομάδα, δηλαδή μια εβδομάδα πριν την ολοκλήρωση του μαθήματος. Όμως, τα περισσότερα μοντέλα καταφέρνουν να διατηρήσουν στατιστικώς σημαντικές τιμές έως λίγο μετά το μέσο περίπου του μαθήματος (την 10<sup>η</sup> και την 11<sup>η</sup> εβδομάδα). Δύο μοντέλα κατορθώνουν (παρά το έλλειμμα πληροφορίας και παρά την υψηλή τιμή βάσης) να επιτύχουν στατιστικώς σημαντικές τιμές την 7<sup>η</sup> εβδομάδα, δηλαδή πριν

το μέσο του μαθήματος. Πρόκειται για τα μοντέλα των αλγόριθμων bayes.NaiveBayesMultinomial (A/A 4) με ακρίβεια πρόβλεψης 79.33% και functions.SimpleLogistic (A/A 7) με ακρίβεια πρόβλεψης 79.19%. Παρατηρούμε ότι και τα μοντέλα των αλγόριθμων lazy.LwL(A/A 10) και trees.DecisionStump (A/A 15) επιτυγχάνουν στατιστικώς σημαντικές τιμές για την ακρίβεια πρόβλεψης την 7<sup>η</sup> εβδομάδα. Ωστόσο η επιτυχία τους είναι «σημειακή», δηλαδή, υπάρχουν εβδομάδες μετά την 7<sup>η</sup> (π.χ. η 8<sup>η</sup>, η 10<sup>η</sup>) κατά τις οποίες τα εν λόγω μοντέλα δεν επιτυγχάνουν στατιστικώς σημαντικές αποδόσεις. Συνεπώς, δεν μπορούν να αξιοποιηθούν στην πρώιμη πρόβλεψη, δεδομένου ότι προσφέρουν ασφαλή πρόβλεψη την 7<sup>η</sup> εβδομάδα, αλλά μη ασφαλή την 8<sup>η</sup> ή τη 10<sup>η</sup> εβδομάδα. Κατ' αντιπαραβολή, η στατιστική σημαντικότητα των μοντέλων bayes.NaiveBayesMultinomial (A/A 4) και functions.SimpleLogistic (A/A 7) ισχύει αδιαλείπτως για όλες τις εβδομάδες από την 7<sup>η</sup> έως τη 18<sup>η</sup>. Αξίζει επίσης να σημειωθεί ότι κάποια μοντέλα υπερβαίνουν την τιμή βάσης στην αρχή του μαθήματος και συγκεκριμένα την 1<sup>η</sup> ή και την 2<sup>η</sup> εβδομάδα. Ωστόσο, η διαφορά από την τιμή βάσης δεν είναι στατιστικώς σημαντική, οπότε οι περιπτώσεις αυτές δεν τυγχάνουν περαιτέρω προσοχής.

Η εξέλιξη των τιμών της ακρίβειας πρόβλεψης των μοντέλων απεικονίζεται παραστατικότερα στο Γράφημα 1 που ακολουθεί.



Γράφημα 1. Ακρίβεια Πρόβλεψης Μοντέλων για το μάθημα ΑΑ005 2020-2021



Στο Γράφημα 1 απεικονίζονται μόνο οι τιμές, που η διαφορά τους από την τιμή βάσης είναι στατιστικώς σημαντική. Δηλαδή, απεικονίζονται οι τιμές οι οποίες έχουν σκούρο γαλάζιο υπόβαθρο στον Πίνακα 22. Στον οριζόντιο άξονα του γραφήματος απεικονίζεται η χρονική περίοδος από την έναρξη του μαθήματος, η οποία προσδιορίζεται με τον αντίστοιχο αριθμό της εβδομάδας (π.χ. Εβδ10: 10<sup>η</sup> εβδομάδα από την έναρξη του μαθήματος). Στον κάθετο άξονα του γραφήματος απεικονίζεται αριθμητική κλίμακα που αντιστοιχεί στο αριθμητικό εύρος των τιμών της ακρίβειας πρόβλεψης. Οι αριθμητικές τιμές αντιστοιχούν σε ποσοστά επί τοις εκατό (%). Κάθε μία έγχρωμη γραμμή, που διατρέχει το γράφημα, αντιστοιχεί στις τιμές μοντέλου το οποίο παράγεται από αλγόριθμο του Πίνακα 23. Δεν σχηματίζονται γραμμές για μία ή περισσότερες τιμές που είναι «απομονωμένες», δηλαδή για τιμή ή για τιμές για τις οποίες υφίσταται μεταγενέστερη μη στατιστικώς σημαντική τιμή. Κατά συνέπεια, δεν υπάρχει γραμμή για το μοντέλο του αλγόριθμου με A/A 19, διότι ποτέ δεν πέτυχε ακρίβεια πρόβλεψης στατιστικά υψηλότερη από την τιμή βάσης. Επίσης δεν υπάρχουν γραμμές για τα μοντέλα των αλγόριθμων με A/A 9, με A/A 14 και με A/A 17, διότι όλες οι στατιστικά σημαντικές τιμές τους είναι «απομονωμένες». Για το μοντέλο του αλγόριθμου με A/A 13 αποτυπώνονται μόνο όσες στατιστικά σημαντικές τιμές δεν είναι «απομονωμένες». Η γαλάζια γραμμή στο κάτω μέρος του γραφήματος αντιστοιχεί στις τιμές του μοντέλου του αλγόριθμου ZeroR (A/A 1), δηλαδή ορίζει τον «πήχη», που όλα τα μοντέλα πρέπει να υπερβούν με στατιστικώς σημαντική διαφορά. Με την εξαίρεση του μοντέλου του αλγόριθμου ZeroR (τιμή βάσης) δεν υπάρχουν τιμές (οπότε δεν σχηματίζονται γραμμές) από την 1<sup>η</sup> εβδομάδα έως και την 7<sup>η</sup> εβδομάδα, διότι κατά την συγκεκριμένη χρονική περίοδο κανένα μοντέλο δεν επιτυγχάνει ακρίβεια πρόβλεψης στατιστικώς σημαντική σε σύγκριση με την τιμή βάσης.

Η πρωιμότερη, στατιστικώς σημαντική διαφορά από την τιμή βάσης επιτυγχάνεται την 7<sup>η</sup> εβδομάδα από τα μοντέλα των αλγόριθμων `functions.SimpleLogistic` (A/A 7) και `Bayes.NaiveBayesMultinomial` (A/A 4). Για λόγους ευκρίνειας, οι γραμμές αυτών των δύο μοντέλων περιβάλλονται από ομοιόχρωμη αχλή. Όπως έχει ήδη αναφερθεί, την 7<sup>η</sup> εβδομάδα τα μοντέλα των αλγόριθμων `lazy.LwL` (A/A 10) και `trees.DecisionStump` (A/A 15) επίσης επιτυγχάνουν στατιστικώς σημαντικές τιμές για την ακρίβεια πρόβλεψης. Ωστόσο, η επιτυχία των δύο αυτών μοντέλων κατά την 7<sup>η</sup> εβδομάδα είναι «σημειακή», δεν επαναλαμβάνεται τις

αμέσως επόμενες εβδομάδες (π.χ. την 8<sup>η</sup>, ή την 10<sup>η</sup>), οπότε η γραμμή που αντιστοιχεί στα μοντέλα αυτών των αλγόριθμων, δεν ξεκινά από την 7<sup>η</sup> εβδομάδα. Η ανάδειξη στατιστικώς σημαντικών τιμών κατά την 7<sup>η</sup> εβδομάδα διευκολύνθηκε από τη μείωση της τιμής βάσης κατά τις προηγούμενες (5<sup>η</sup> και 6<sup>η</sup>) εβδομάδες, μείωση που είναι εμφανής στο γράφημα, μέσω της βύθισης της γραμμής του μοντέλου ZeroR.

Στο γράφημα είναι εμφανές, ότι τα περισσότερα μοντέλα επιτυγχάνουν για πρώτη φορά στατιστικώς σημαντικές τιμές τη 10<sup>η</sup> και την 11<sup>η</sup> εβδομάδα. Αυτό συμβαίνει κυρίως, διότι την προηγούμενη (10<sup>η</sup>) εβδομάδα η τιμή βάσης μειώνεται περαιτέρω, γεγονός που αναπαρίσταται από την ανάλογη βύθιση στην αντίστοιχη γραμμή του μοντέλου ZeroR. Κατά αναλογία, η επόμενη μείωση στην τιμή βάσης την 13<sup>η</sup> εβδομάδα (που απεικονίζεται με βύθιση της αντίστοιχης γραμμής) επιτρέπει σε περισσότερα μοντέλα να επιτύχουν ακρίβεια πρόβλεψης με στατιστικώς σημαντική διαφορά από την τιμή βάσης.

Στο γράφημα φαίνεται επίσης, ότι η ακρίβεια πρόβλεψης των μοντέλων αυξάνει ως την 13<sup>η</sup> εβδομάδα. Εμφανίζει πτώση τη 14<sup>η</sup> εβδομάδα, όταν δηλαδή ολοκληρώνονται οι διαλέξεις (βλ. Κεφ. 9.5.2, Πίνακας 17). Ακολουθεί σταθεροποίηση και αργότερα ανάκαμψη, που κορυφώνεται τις τελευταίες 2 εβδομάδες, όταν δηλαδή πλησιάζει η ημέρα της τελικής εξέτασης. Το φαινόμενο εμφανίζεται, με μικρές διαφοροποιήσεις, σε όλα τα μοντέλα, οπότε εικάζουμε ότι δεν είναι τυχαίο. Ως εξήγηση του φαινομένου προτείνουμε το εξής: Όσο χρονικό διάστημα πραγματοποιούνται διαλέξεις, οι φοιτητές χρησιμοποιούν τακτικά το ΣΔΜ Moodle και κατά συνέπεια, τα αρχεία καταγραφής εμπλουτίζονται συνεχώς με νέες πληροφορίες που βοηθούν στην «εκμάθηση» (training) των μοντέλων και στη βελτίωση της απόδοσής τους. Την 14<sup>η</sup> εβδομάδα, με την ολοκλήρωση των διαλέξεων και την επικείμενη έναρξη της εξεταστικής περιόδου, οι φοιτητές, είτε παύουν να χρησιμοποιούν το ΣΔΜ, ή το χρησιμοποιούν με τρόπο λιγότερο συστηματικό από ό,τι τις προηγούμενες περιόδους. Δεδομένου ότι η τελική εξέταση του μαθήματος είχε οριστεί στο τέλος της εξεταστικής περιόδου (βλ. Κεφ. 9.5.2, Πίνακας 17), οι φοιτητές επιλέγουν να μελετήσουν άλλα μαθήματα. Μεταθέτουν τη μελέτη του μαθήματος και τη σύνδεση στο σύστημα προς το τέλος της εξεταστικής περιόδου, όταν θα πλησιάζει η ημέρα της τελικής εξέτασης. Η εν λόγω τακτική των φοιτητών έχει ως αποτέλεσμα να εισάγεται στο σύστημα μη συστηματική, ακανόνιστη, άρρυθμη πληροφορία, η οποία προκαλεί αταξία στα

σύνολα δεδομένων και η οποία «θολώνει», αλλοιώνει, αποδιοργανώνει τα μοτίβα (patterns) οδηγώντας σε πτώση της ακρίβειας πρόβλεψης των μοντέλων. Η συστηματική, εύρυθμη, εντατική χρήση του ΣΔΜ Moodle τις δύο τελευταίες (17<sup>η</sup> και 18<sup>η</sup>) εβδομάδες, όταν οι φοιτητές προετοιμάζονταν για την τελική εξέταση του μαθήματος, έχει ως αποτέλεσμα την ανάκτηση των υψηλών τιμών στην ακρίβεια πρόβλεψης των μοντέλων.

#### 10.1.2. Επιπλέον κριτήρια για την απόδοση των βασικών μοντέλων

Προκειμένου να αποτιμήσουμε την απόδοση των μοντέλων δεν αρκεστήκαμε στη σύγκριση της ακρίβειας πρόβλεψης με την τιμή βάσης και στην ανάδειξη των τιμών με στατιστική σημαντικότητα. Υπολογίσαμε επιπλέον τις τιμές της μεταβλητής AUC (Area Under Curve) (Witten, Frank & Hall, 2011, σ. 172) και της μεταβλητής Cohen’s Kappa (Witten, Frank, & Hall, 2011, σ. 166).

Όπως έχουμε αναφέρει στο Κεφάλαιο 9.6.3, η μεταβλητή AUC αξιολογεί τη διαγνωστική ικανότητα ενός μοντέλου ταξινόμησης με διχοτομημένη κλάση και μη ισορροπημένο (Hussain M., Zhu, Zhang, & Abidi, 2018) σύνολο δεδομένων. Υπολογίζει την περιοχή που ορίζει η καμπύλη ROC (Receiver Operating Characteristic) (βλ. Κεφ. 9.6.3 Σχήμα 13), η οποία σχηματίζεται σε σύστημα συντεταγμένων, όπου στον Y άξονα αποτυπώνεται ο δείκτης (rate) των αληθώς θετικών προβλέψεων και στον X άξονα αποτυπώνεται ο δείκτης (rate) των ψευδώς θετικών προβλέψεων (Cui, Chen, Shiri & Fan, 2019, σ. 218). Η μεταβλητή AUC είναι χρήσιμη στα μη ισορροπημένα σύνολα, διότι εκεί ένα μοντέλο μπορεί να πετύχει υψηλή τιμή ακρίβειας, ταξινομώντας απλώς στην πολυπληθέστερη κλάση. Στη περίπτωση αυτή, για να επιβεβαιωθεί η προβλεπτική ισχύς του μοντέλου πρέπει να ληφθεί υπόψιν και η τιμή της μεταβλητής AUC (Chen & Cui, 2020, σ. 7).

Στον Πίνακα 24 καταγράφονται οι τιμές της μεταβλητής AUC για τα μοντέλα του μαθήματος AA005 Χρήση Βασικών Εφαρμογών ΗΥ, Χειμερινό Εξάμηνο 2020-2021.

**Πίνακας 24. Τιμές μεταβλητής AUC για τα μοντέλα του μαθήματος AA005 2020-2021**

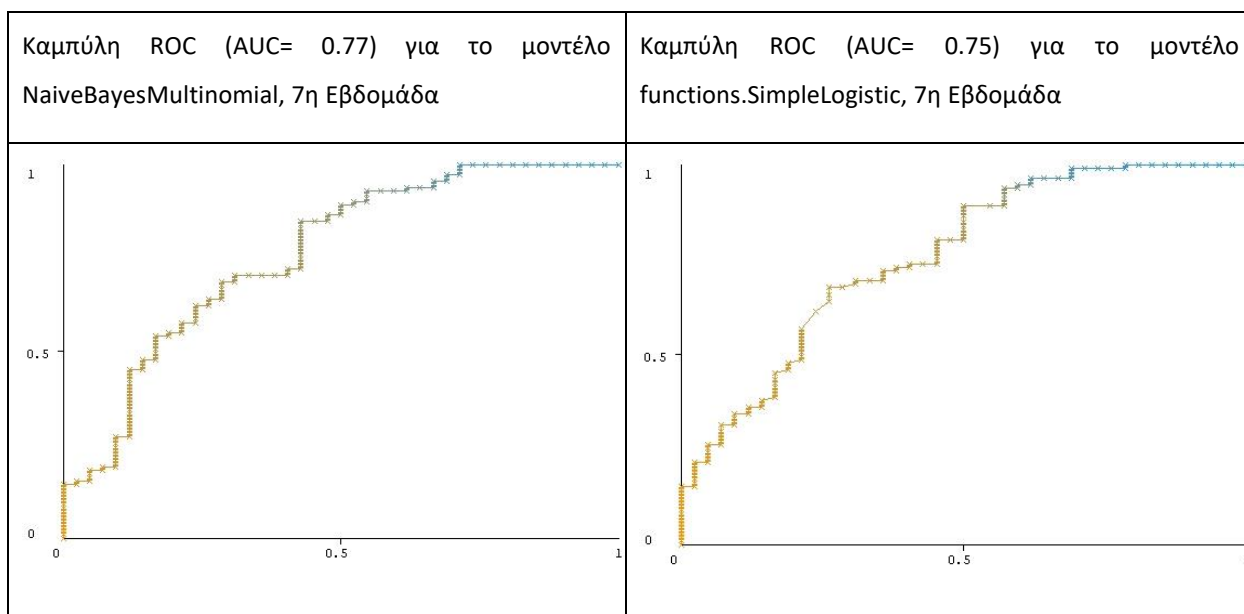
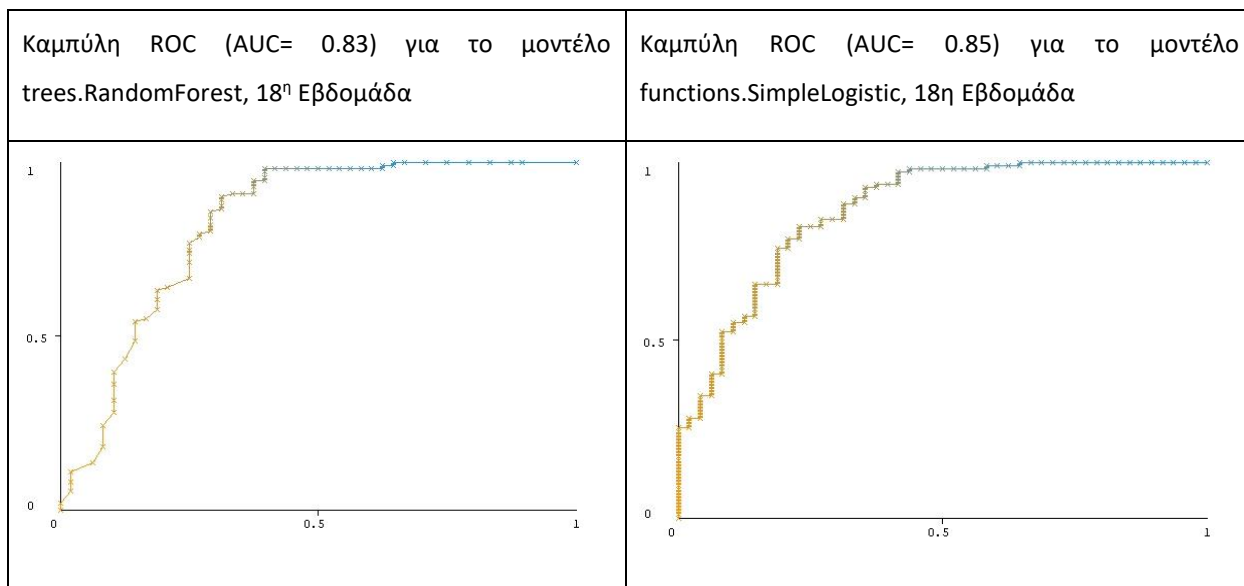
Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
Αλγορ.01	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Αλγορ.02	0.49	0.51	0.6	0.63	0.71	0.69	0.75	0.78	0.79	0.82	0.81	0.83	0.84	0.81	0.81	0.82	0.82	0.83
Αλγορ.03	0.7	0.65	0.68	0.7	0.74	0.74	0.77	0.8	0.81	0.81	0.81	0.82	0.84	0.84	0.85	0.84	0.85	0.85

Αλγορ.04	0.67	0.63	0.68	0.7	0.75	0.72	0.77	0.8	0.82	0.82	0.82	0.82	0.85	0.84	0.84	0.83	0.83	0.85
Αλγορ.05	0.67	0.68	0.63	0.66	0.7	0.66	0.74	0.75	0.77	0.76	0.75	0.78	0.79	0.79	0.78	0.78	0.81	0.85
Αλγορ.06	0.7	0.63	0.53	0.57	0.62	0.63	0.68	0.67	0.68	0.7	0.69	0.71	0.73	0.77	0.78	0.79	0.78	0.83
Αλγορ.07	0.6	0.68	0.65	0.67	0.75	0.74	0.75	0.77	0.79	0.8	0.77	0.81	0.83	0.83	0.83	0.83	0.84	0.85
Αλγορ.08	0.62	0.61	0.56	0.51	0.57	0.6	0.56	0.6	0.6	0.63	0.62	0.67	0.68	0.73	0.74	0.74	0.74	0.8
Αλγορ.09	0.62	0.67	0.56	0.56	0.63	0.67	0.64	0.68	0.68	0.67	0.68	0.73	0.79	0.7	0.7	0.74	0.74	0.8
Αλγορ.10	0.65	0.68	0.65	0.64	0.74	0.69	0.74	0.74	0.79	0.79	0.79	0.81	0.84	0.8	0.81	0.81	0.79	0.87
Αλγορ.11	0.49	0.56	0.57	0.56	0.64	0.63	0.63	0.64	0.72	0.73	0.74	0.77	0.79	0.76	0.75	0.72	0.71	0.77
Αλγορ.12	0.61	0.58	0.57	0.6	0.61	0.6	0.63	0.62	0.67	0.66	0.7	0.73	0.75	0.71	0.71	0.71	0.71	0.74
Αλγορ.13	0.54	0.57	0.55	0.63	0.56	0.58	0.61	0.63	0.62	0.71	0.66	0.71	0.74	0.69	0.65	0.65	0.68	0.73
Αλγορ.14	0.6	0.56	0.58	0.56	0.65	0.6	0.65	0.65	0.63	0.68	0.67	0.69	0.73	0.67	0.67	0.7	0.66	0.75
Αλγορ.15	0.52	0.56	0.58	0.57	0.55	0.62	0.64	0.61	0.67	0.65	0.67	0.74	0.75	0.67	0.67	0.68	0.69	0.76
Αλγορ.16	0.5	0.51	0.5	0.5	0.55	0.53	0.65	0.69	0.8	0.81	0.81	0.82	0.84	0.84	0.85	0.84	0.85	0.85
Αλγορ.17	0.57	0.56	0.58	0.52	0.61	0.62	0.61	0.62	0.65	0.67	0.65	0.68	0.72	0.67	0.67	0.68	0.69	0.73
Αλγορ.18	0.69	0.68	0.64	0.65	0.72	0.72	0.73	0.74	0.76	0.76	0.77	0.8	0.82	0.77	0.8	0.8	0.79	0.83
Αλγορ.19	0.59	0.57	0.56	0.57	0.57	0.61	0.62	0.62	0.64	0.64	0.64	0.68	0.71	0.67	0.69	0.69	0.69	0.72
Αλγορ.20	0.54	0.56	0.56	0.55	0.59	0.58	0.66	0.61	0.65	0.7	0.69	0.71	0.74	0.7	0.71	0.7	0.7	0.73

Στον Πίνακα 24, καταρχάς παρατηρούμε ότι η τιμή AUC για το μοντέλο του αλγόριθμου ZeroR (A/A 1) ισούται με 0.5, γεγονός απολύτως αναμενόμενο, διότι όπως έχουμε ήδη αναφέρει (βλ. Κεφ. 9.6.2) το συγκεκριμένο μοντέλο προβλέπει «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Συνεπώς η τιμή AUC δεν έχει νόημα στο συγκεκριμένο μοντέλο, το οποίο χρησιμοποιείται για τον υπολογισμό της τιμής βάσης. Όλα τα άλλα μοντέλα, σε όλες τις χρονικές περιόδους (με την εξαίρεση δύο τιμών την 1<sup>η</sup> εβδομάδα), σημειώνουν τιμές που είναι μεγαλύτερες από το 0.5, συνεπώς κρίνονται αποδεκτές (Riestra-Gonzalez, del Puerto Paule-Ruiz & Ortin, 2021, σ. 3). Την 18<sup>η</sup> εβδομάδα, κατά την ολοκλήρωση του μαθήματος, οι τιμές AUC, για τα περισσότερα μοντέλα, είναι μεγαλύτερες από 0.8, οπότε κρίνονται εξαιρετικές. Επιπλέον παρατηρούμε ότι, την 7<sup>η</sup> εβδομάδα για τα μοντέλα των αλγόριθμων functions.SimpleLogistic (A/A 7) και Bayes.NaiveBayesMultinomial (A/A 4), τα οποία επιτυγχάνουν πρώτη πρόβλεψη, οι αντίστοιχες τιμές AUC είναι 0.75 και 0.77 και κρίνονται καλές.

Στο Γράφημα 2 που ακολουθεί αναπαρίστανται οι καμπύλες ROC που υπολογίζουν ενδεικτικές τιμές AUC και συγκεκριμένα τις τιμές AUC της 18<sup>ης</sup> εβδομάδας για τα μοντέλα trees.RandomForest (A/A 18) και functions.SimpleLogistic (A/A 7) και τις τιμές AUC της 7<sup>ης</sup>

εβδομάδας για τα μοντέλα bayes.NaiveBayesMultinomial (A/A 4) και functions.SimpleLogistic (A/A 7).



*Γράφημα 2. Καμπύλες ROC για τα μοντέλα του μαθήματος AA005 2020-2021*

Όπως έχουμε αναφέρει στο Κεφάλαιο 9.6.3, η μεταβλητή Cohen's Kappa μετρά τη συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που προέβλεψε το μοντέλο, ενώ διορθώνει τυχόν συμφωνία που προέκυψε τυχαίως (Witten, Frank, & Hall, 2011, σ. 166).

**Πίνακας 25. Τιμές μεταβλητής Cohen's Kappa για τα μοντέλα του μαθήματος AA005 2020-2021**

Εβδομ	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
Αλγορ.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Αλγορ.02	-0.03	-0.01	0.21	0.17	0.35	0.3	0.32	0.34	0.4	0.51	0.46	0.52	0.56	0.49	0.52	0.54	0.51	0.6
Αλγορ.03	0.13	0.14	0.27	0.29	0.41	0.41	0.37	0.38	0.44	0.47	0.51	0.52	0.52	0.59	0.57	0.61	0.53	0.55
Αλγορ.04	0.03	0	0	-0.01	0.08	0.07	0.32	0.34	0.4	0.43	0.44	0.45	0.46	0.53	0.5	0.47	0.44	0.49
Αλγορ.05	0.12	0.24	0.28	0.28	0.29	0.28	0.32	0.29	0.36	0.37	0.35	0.41	0.45	0.46	0.48	0.45	0.46	0.59
Αλγορ.06	0.19	0.16	0.11	0.13	0.15	0.15	0.13	0.16	0.21	0.31	0.27	0.37	0.41	0.44	0.45	0.48	0.52	0.53
Αλγορ.07	0.07	0.29	0.21	0.18	0.29	0.32	0.39	0.37	0.37	0.4	0.42	0.48	0.56	0.51	0.52	0.54	0.55	0.58
Αλγορ.08	0.22	0.21	0.11	0.02	0.14	0.2	0.13	0.2	0.2	0.27	0.24	0.33	0.37	0.45	0.49	0.48	0.46	0.59
Αλγορ.09	0.15	0.18	0.13	0.08	0.22	0.28	0.22	0.22	0.22	0.24	0.29	0.33	0.45	0.25	0.31	0.37	0.31	0.54
Αλγορ.10	0.07	0.28	0.18	0.14	0.24	0.25	0.32	0.24	0.34	0.36	0.36	0.52	0.55	0.39	0.39	0.39	0.41	0.57
Αλγορ.11	0	0.15	0.18	0.18	0.24	0.27	0.28	0.25	0.33	0.41	0.36	0.49	0.52	0.45	0.42	0.41	0.43	0.53
Αλγορ.12	0.24	0.18	0.17	0.2	0.26	0.22	0.27	0.27	0.37	0.34	0.44	0.51	0.55	0.46	0.44	0.44	0.44	0.52
Αλγορ.13	0.08	0.17	0.12	0.3	0.13	0.18	0.25	0.29	0.26	0.45	0.35	0.46	0.53	0.41	0.32	0.32	0.39	0.48
Αλγορ.14	0.09	0.1	0.16	0.12	0.19	0.12	0.22	0.26	0.22	0.37	0.3	0.36	0.4	0.38	0.41	0.42	0.38	0.46
Αλγορ.15	-0.01	0.1	0.22	0.16	0.11	0.28	0.34	0.26	0.37	0.31	0.38	0.53	0.55	0.37	0.35	0.4	0.41	0.57
Αλγορ.16	0	0.01	0	0	0.09	0.03	0.21	0.21	0.41	0.49	0.51	0.51	0.52	0.58	0.56	0.6	0.54	0.55
Αλγορ.17	0.11	0.12	0.16	0.08	0.22	0.16	0.21	0.24	0.24	0.38	0.31	0.37	0.4	0.37	0.4	0.4	0.38	0.47
Αλγορ.18	0.24	0.17	0.13	0.13	0.26	0.25	0.3	0.29	0.36	0.48	0.44	0.56	0.59	0.55	0.57	0.59	0.56	0.62
Αλγορ.19	0.15	0.13	0.11	0.14	0.13	0.22	0.24	0.24	0.27	0.29	0.28	0.37	0.41	0.33	0.37	0.38	0.38	0.44
Αλγορ.20	0.04	0.08	0.13	0.11	0.17	0.18	0.29	0.21	0.31	0.36	0.41	0.45	0.5	0.41	0.43	0.41	0.4	0.5

Στον Πίνακα 25 καταρχάς παρατηρούμε ότι για το μοντέλο του αλγόριθμου ZeroR (A/A 1) η τιμή της μεταβλητής Cohen's Kappa ισούται με μηδέν, διότι, όπως έχουμε ήδη αναφέρει (βλ. Κεφ. 9.6.2), το συγκεκριμένο μοντέλο προβλέπει «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Παρατηρούμε επίσης ότι κατά την 18<sup>η</sup> εβδομάδα, η μεταβλητή Cohen's Kappa για όλα τα μοντέλα λαμβάνει τιμές από το διάστημα [0.44 , 0.62], γεγονός που σύμφωνα με τους Quinn & Gray (2020, σ. 8) δηλώνει «μέση» (moderate) συμφωνία. Ωστόσο, τις εβδομάδες πριν την ολοκλήρωση του μαθήματος, οι τιμές της μεταβλητής, όπως είναι αναμενόμενο, μειώνονται σε επίπεδα που δηλώνουν «δίκαιη» (fair) συμφωνία και τις πρώτες εβδομάδες μειώνονται σε επίπεδα που δηλώνουν «αδύναμη» (slight) συμφωνία. Συνεπώς, σύμφωνα με το κριτήριο της μεταβλητής Cohen's Kappa, κατά τη χρονική στιγμή ολοκλήρωσης του μαθήματος, υπάρχει «μέση» συμφωνία μεταξύ των

πραγματικών κλάσεων και των κλάσεων που προέβλεψαν τα μοντέλα, ενώ η ισχύς της συμφωνίας εξασθενεί βαθμιαία κατά τις προηγούμενες χρονικές περιόδους.

### 10.1.3. Ακρίβεια και χρόνος πρόβλεψης του LSTM μοντέλου

Όπως αναφέραμε στο Κεφάλαιο 8.1. διερευνήσαμε την ακρίβεια πρόβλεψης μοντέλων τα οποία παράγονται, όχι μόνο με τη βοήθεια βασικών αλγόριθμων (αυτών του Πίνακα 23), αλλά και αλγόριθμων που βασίζονται σε περιοδικά νευρωνικά δίκτυα (Recursive Neural Networks). Αυτοί οι αλγόριθμοι αναλύουν όχι μόνο τις συνολικές, τελικές τιμές των χαρακτηριστικών του συνόλου δεδομένων, αλλά και τις χρονικές ακολουθίες, οι οποίες αποτυπώνουν την εξέλιξη των τιμών. Η μελέτη τέτοιων τεχνικών επί εκπαιδευτικών δεδομένων αποτελεί πεδίο αυξανόμενου ερευνητικού ενδιαφέροντος (Doleck, Lemay, Basnet & Bazalais, 2020). Επιπλέον, οι ερευνητικές προσπάθειες που διερευνούν την προβλεπτική ικανότητα τέτοιων τεχνικών είναι σχετικά λίγες και δεν συγκλίνουν προς κοινώς αποδεκτά συμπεράσματα (Chen & Cui, 2020. Doleck, Lemay, Basnet & Bazalais, 2020). Επιλέξαμε την LSTM υλοποίηση των περιοδικών νευρωνικών δικτύων και καταρχάς χρησιμοποιήσαμε - όπως και στην περίπτωση των βασικών αλγόριθμων (βλ. Κεφ. 10.1) - τις προκαθορισμένες (default) παραμέτρους του LSTM αλγόριθμου.

Στον Πίνακα 26 που ακολουθεί καταγράφονται οι τιμές για την ακρίβεια πρόβλεψης του LSTM μοντέλου.

**Πίνακας 26. Ακρίβεια πρόβλεψης LSTM μοντέλου με προκαθορισμένες παραμέτρους για το μάθημα AA005 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
ZeroR	73.19	74.9	74.52	74.52	73.04	72.58	72.58	72.12	70.75	70.75	70.75	70.75	69.83	69.83	69.83	69.83	69.83	69.83
LSTM	73.19	74.9	74.52	74.52	73.04	72.58	72.58	72.12	70.75	70.75	70.75	70.75	69.83	69.83	69.83	69.83	69.83	69.83

Στον Πίνακα 26, παρατηρούμε ότι σε όλες τις χρονικές περιόδους η ακρίβεια πρόβλεψης του μοντέλου του LSTM αλγόριθμου δεν ξεπερνά την τιμή βάσης, δηλαδή το μοντέλο δεν δίνει γόνιμα αποτελέσματα. Προκειμένου να επιτύχουμε τιμές που θα υπερβαίνουν την τιμή βάσης, επιχειρήσαμε αλλαγές στις προκαθορισμένες (default) παραμέτρους του αλγόριθμου, ακολουθώντας εμπειρικούς κανόνες που υποδεικνύονται από τη σχετική βιβλιογραφία, σύμφωνα με την οποία, η εύρεση ενός αποδοτικού LSTM μοντέλου αποτελεί «άσκηση δοκιμής και λάθους» (an exercise in trial-and-error) (Frank, 2021). Μετά από πολλές δοκιμές με διαφορετικούς συνδυασμούς τιμών για τις παραμέτρους, κατασκευάστηκε νέο



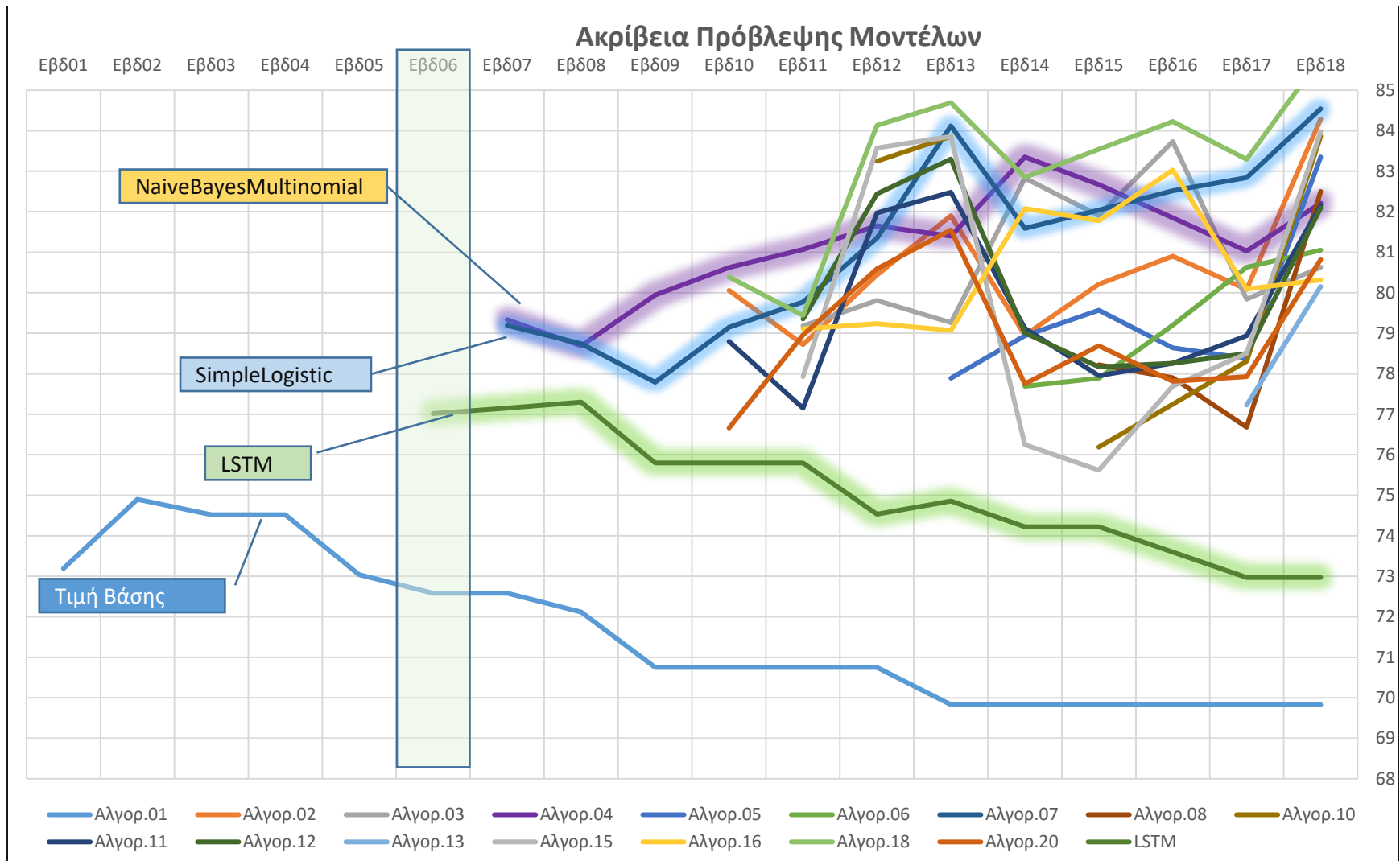
βελτιστοποιημένο μοντέλο (βλ. Παράρτημα Β), το οποίο επιτυγχάνει ακρίβεια πρόβλεψης που υπερβαίνει την τιμή βάσης με στατιστικώς σημαντική διαφορά. Οι τιμές καταγράφονται στον Πίνακα 27.

**Πίνακας 27. Ακρίβεια πρόβλεψης LSTM μοντέλου με βελτιστοποίηση ως προς τις παραμέτρους για το μάθημα ΑΑ005 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
ZeroR	73.19	74.9	74.52	74.52	73.04	72.58	72.58	72.12	70.75	70.75	70.75	70.75	69.83	69.83	69.83	69.83	69.83	69.83
LSTM	59.71 *	76.18	76.07	77.87 v	76.32	77.02 v	77.15 v	77.30 v	75.80 v	75.80 v	75.80 v	74.53 v	74.86 v	74.22 v	74.22 v	73.60 v	72.97 v	72.97 v

Στον Πίνακα 27 παρατηρούμε ότι το μοντέλο επιτυγχάνει συνεχόμενες στατιστικώς σημαντικές τιμές για την ακρίβεια πρόβλεψης από την 6<sup>η</sup> έως τη 18<sup>η</sup> εβδομάδα. Δηλαδή, το LSTM μοντέλο επιτυγχάνει πρώιμη πρόβλεψη μία εβδομάδα νωρίτερα από ό,τι τα μοντέλα των βασικών αλγόριθμων ταξινόμησης του Πίνακα 23. Παρατηρούμε επίσης, ότι οι τιμές ακρίβειας του LSTM μοντέλου είναι μικρότερες από τις τιμές των βασικών μοντέλων (βλ. Πίνακα 22). Συγκεκριμένα, κατά την 18<sup>η</sup> εβδομάδα, δηλαδή κατά την ολοκλήρωση του μαθήματος, η υψηλότερη τιμή είναι 77.30%, ενώ τα βασικά μοντέλα συχνά υπερβαίνουν το 80% και φτάνουν ως το 85%.

Στο Γράφημα 3, το οποίο παρουσιάζει τις γραφικές παραστάσεις των τιμών όλων των μοντέλων, φαίνεται ότι το LSTM μοντέλο επιτυγχάνει την πρωιμότερη στατιστικώς σημαντική πρόβλεψη κατά την 6<sup>η</sup> εβδομάδα, δηλαδή μία εβδομάδα νωρίτερα από ό,τι οι βασικοί αλγόριθμοι. Η ακρίβεια πρόβλεψής του είναι μικρότερη από όλων των άλλων μοντέλων. Αξιοσημείωτο είναι ότι καθώς κινούμαστε προς την ολοκλήρωση του μαθήματος, οι τιμές στην ακρίβεια πρόβλεψής του LSTM μοντέλου φθίνουν, ενώ αντιθέτων οι τιμές ακρίβειας πρόβλεψης των άλλων μοντέλων αυξάνουν.



Γράφημα 3. Ακρίβεια Πρόβλεψης LSTM Μοντέλου μετά από βελτιστοποίηση παραμέτρων

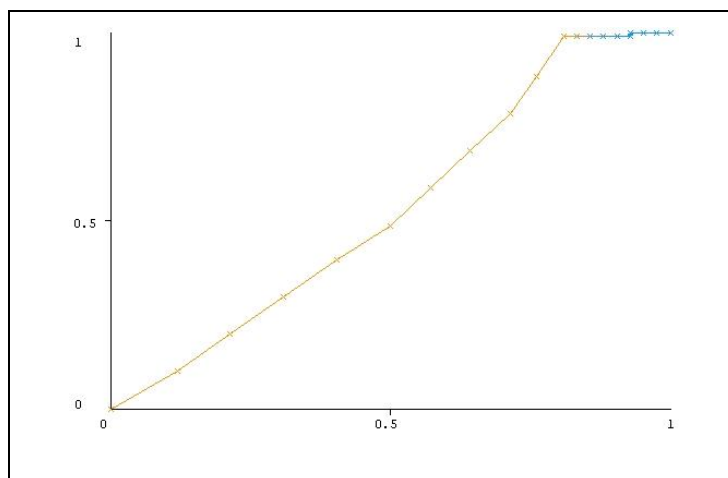
#### 10.1.4 Επιπλέον κριτήρια για την απόδοση του LSTM μοντέλου

Προκειμένου να αποτιμήσουμε την απόδοση του LSTM μοντέλου με περισσότερα κριτήρια υπολογίσαμε (όπως πράξαμε και για τα μοντέλα των βασικών αλγόριθμων) τις τιμές της μεταβλητής AUC (Area Under Curve) (Witten, Frank & Hall, 2011, σ. 172) και της μεταβλητής Cohen's Kappa (Witten, Frank, & Hall, 2011, σ. 166). Οι τιμές καταγράφονται στους Πίνακες 27 και 28. Επιπλέον στο Γραφημα 4 αναπαρίσταται η καμπύλη ROC της τιμής AUC σε μια ενδεικτική χρονική στιγμή (6<sup>η</sup> εβδομάδα).

Στους Πίνακες 28 και 29 παρατηρούμε ότι η μεταβλητή AUC λαμβάνει τιμές από το διάστημα [0.54 , 0.59], οπότε κρίνεται αποδεκτή (Riestra-Gonzalez, del Puerto Paule-Ruiz, & Ortin, 2021), ενώ η μεταβλητή Cohen's Kappa λαμβάνει τιμές από το διάστημα [0.13 , 0.23], γεγονός που δηλώνει «αδύναμη» (slight) συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που προέβλεψαν τα μοντέλα (Quinn & Gray, 2020, σ. 8). Παρατηρούμε επίσης, ότι οι τιμές και των δύο μεταβλητών, όπως συμβαίνει και με την τιμή της ακρίβειας πρόβλεψης, ακολουθούν φθίνουσα και όχι αύξουσα πορεία όταν κινούμαστε προς την ολοκλήρωση του μαθήματος.

**Πίνακας 28. Τιμές μεταβλητής AUC του LSTM μοντέλου για το μάθημα AA005 2020-2021**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
ZeroR	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
LSTM	0.54	0.55	0.55	0.56	0.57	0.59	0.58	0.59	0.58	0.58	0.58	0.56	0.58	0.57	0.57	0.56	0.55	0.55



**Γράφημα 4. Καμπύλη ROC για το μοντέλο LSTM, Εβδομάδα 6**

**Πίνακας 29. Τιμές μεταβλητής Cohen's Kappa του LSTM μοντέλου για το μάθημα AA005 2020-2021**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
ZeroR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LSTM	0.16	0.18	0.14	0.19	0.17	0.22	0.21	0.23	0.21	0.21	0.21	0.16	0.21	0.18	0.18	0.15	0.13	0.13

Συμπερασματικά, με την αξιοποίηση του LSTM μοντέλου, η πρόβλεψη μπορεί να επιτευχθεί νωρίτερα κατά μία εβδομάδα. Ωστόσο, οι τιμές στην ακρίβεια πρόβλεψης είναι χαμηλότερες από ό,τι των βασικών μοντέλων και φθίνουν, αντί να αυξάνουν, καθώς κινούμαστε προς την ολοκλήρωση του μαθήματος. Επιπλέον, σύμφωνα με την μεταβλητή Cohen's Kappa η προβλεπτική ικανότητα του μοντέλου κρίνεται ως «αδύναμη» (slight).

Ως προς την αξιοποίηση του LSTM μοντέλου, είναι απαραίτητο να επισημανθεί ότι η κατασκευή του δεν είναι εύκολη. Απαιτείται μεγάλη προσπάθεια και πολλές δοκιμές για την εύρεση των παραμέτρων που βελτιστοποιούν την απόδοση του LSTM αλγόριθμου. Επιπλέον, ο χρόνος εκτέλεσης του αλγόριθμου είναι πολύ μεγάλος, λόγω της μεγάλης επεξεργαστικής ισχύος που απαιτείται (Chollet, 2017, σ. 12). Ενδεικτικά, αναφέρουμε ότι για το σύνολο δεδομένων της 6<sup>ης</sup> εβδομάδας, στο περιβάλλον του WEKA Experimenter, η εκτέλεση του βελτιστοποιημένου LSTM αλγόριθμου διαρκεί 40 λεπτά, ενώ η εκτέλεση οποιουδήποτε από τους βασικούς αλγόριθμους διαρκεί μόλις 2 δευτερόλεπτα<sup>17</sup>.

<sup>17</sup> Η εκτέλεση του αλγόριθμου έγινε σε υπολογιστή με τα εξής χαρακτηριστικά.: Λειτουργικό Σύστημα MS Windows 10, 64-bit / Επεξεργαστής Intel(R) Core(TM) i7-4790 CPU@3.60GHz / RAM 16.0 GB.

## 10.2. Τα μοντέλα πρόβλεψης για το Μάθημα ΑΑ005 2021-2022

### 10.2.1. Ακρίβεια και χρόνος πρόβλεψης των βασικών μοντέλων

Στον Πίνακα 30 εμφανίζονται οι τιμές για την ακρίβεια πρόβλεψης των μοντέλων που κατασκεύασαν 20 βασικοί αλγόριθμοι (Πίνακας 23) σε 17 σύνολα δεδομένων του μαθήματος ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ, χειμερινό εξάμηνο 2021-2022.

**Πίνακας 30. Ακρίβεια πρόβλεψης μοντέλων για το μάθημα ΑΑ005 2021-2022 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>
Αλγορ.01	85.28	81.15	80.57	79.38	78.95	79.35	79.8	80	79.36	79.15	78.3	78.3	78.3	78.47	77.5	77.62	77.36
Αλγορ.02	85.28	81.15	80.04	79.38	68.99 *	70.99 *	74.5	80.83	81.06	80.32	82.06	82.13	82.28	83.86	85.65 v	85.61 v	85.45 v
Αλγορ.03	81.57	59.92 *	59.02 *	64.75 *	65.75 *	70.88 *	76.01	78.28	77.51	75.71	77.63	78.58	79.37	84.63 v	84.50 v	85.72 v	85.59 v
Αλγορ.04	82.25	80.37	80.1	79.38	78.95	79.35	79.13	80.94	78.96	79.15	79.9	79.74	80.32	85.83 v	87.85 v	89.71 v	89.53 v
Αλγορ.05	84.85	75.27 *	78.42	75.63	77.5	75.49	79.57	83	81.37	82.27	80.75	80.91	81.6	84.55 v	86.40 v	86.68 v	85.39 v
Αλγορ.06	81.94	71.81 *	75.39	71.06 *	72.06 *	72.20 *	73.39 *	79.72	79.31	80.34	80.62	81.85	80.86	84.77 v	84.90 v	87.46 v	86.66 v
Αλγορ.07	86.01	80.21	79.9	79.25	78.42	77.92	81.01	83.17	84.11 v	83.04 v	84.08 v	85.08 v	85.67 v	87.87 v	89.75 v	89.91 v	87.75 v
Αλγορ.08	74.01 *	70.99 *	72.62 *	71.37 *	69.01 *	67.42 *	71.37 *	77.33	81.79	78.39	77.3	77.83	75.78	83.79 v	85.90 v	86.48 v	82.94 v
Αλγορ.09	74.83 *	71.29 *	71.39 *	69.94 *	65.32 *	67.79 *	73.00 *	79.28	79.48	79.81	77.83	76.72	80.54	86.33 v	87.35 v	88.02 v	84.77 v
Αλγορ.10	86.5	75.37 *	79.9	78.62	78.95	78.9	79.01	84.5	83.24	80.11	82.23	82.55	84.65 v	79.83	84.25 v	89.72 v	86.26 v
Αλγορ.11	85.28	80.22	79.9	79.38	78.9	78.15	81.2	83.83	82.49	82.77	82.81	83.02	83.96 v	83.73 v	87.35 v	89.26 v	86.83 v
Αλγορ.12	85.15	79.74	78.42	79.12	77.32	76.3	80.36	84.33	81.86	81.06	81.22	81.43	82.06	83.72 v	87.00 v	88.02 v	88.24 v
Αλγορ.13	85.03	77.78	77	76.44	74.98	74.97	78.62	82.78	81.98	81.05	81.38	81.34	83.96 v	82.84	85.40 v	89.37 v	83.28 v
Αλγορ.14	81.6	72.29 *	73.37 *	72.37 *	74.83	74.42	74.16	78	77.66	78.09	79	78.49	79.11	83.4	83.65 v	84.75 v	86.00 v
Αλγορ.15	85.15	79.21	79.9	79.38	78.95	78.78	79.29	83.5	82.65	79.47	82.76	82.29	84.86 v	80.28	85.20 v	90.07 v	86.15 v
Αλγορ.16	84.14	81.15	80.57	79.38	78.95	79.35	79.69	79.06	77.98	78.03	77.37	77.67	79.04	83.86	84.45 v	85.72 v	85.83 v
Αλγορ.17	83.9	76.17 *	72.91 *	76.44	75.63	75.23	74.17	79.83	80.93	81.07	78.96	78.23	80.12	82.96	86.00 v	85.98 v	85.43 v
Αλγορ.18	82.04	75.22 *	78.89	76.06	73.57 *	76.21	78.18	84	84.01 v	81.58	82.01	83.23 v	84.81 v	86.03 v	87.70 v	88.66 v	89.24 v
Αλγορ.19	75.26 *	66.81 *	68.82 *	66.75 *	69.23 *	70.95 *	73.26	77.56	78.46	76.36	77.67	77.99	77.32	80.33	82	83.54 v	83.01
Αλγορ.20	85.03	80.76	79.7	78.88	77.84	78.08	78.33	81.33	80.49	80.21	79.79	80.9	81.54	82.8	87.10 v	88.22 v	87.08 v

Το πλήθος των εβδομάδων στον Πίνακα 30 είναι 17, δηλαδή λιγότερες κατά μία από τις εβδομάδες του Πίνακα 22, διότι το έτος 2021-2022 το χρονικό διάστημα μεταξύ της έναρξης του μαθήματος και της τελικής εξέτασης ήταν 17 εβδομάδες (βλ. Κεφ. 9.5.2, Πίνακας 16). Οπότε, στην τελευταία στήλη του Πίνακα 30 εμφανίζονται οι τιμές για την ακρίβεια πρόβλεψης στο σύνολο δεδομένων της 17<sup>ης</sup> εβδομάδας, δηλαδή στο σύνολο που περιλαμβάνει τα δεδομένα όλου του μαθήματος από την αρχή του (1<sup>η</sup> εβδομάδα), έως το τέλος του (17<sup>η</sup> εβδομάδα). Οι τιμές της ακρίβειας πρόβλεψης, η οποία ορίζεται ως ο λόγος του πλήθους των ορθώς ταξινομημένων οντοτήτων προς το συνολικό πλήθος των οντοτήτων (Witten, Frank & Hall, 2011, σ. 72) (βλ. Κεφ. 9.6.2), εμφανίζονται με μορφή ποσοστού επί τοις εκατό (%). Στη δεύτερη γραμμή του Πίνακα 30 εμφανίζονται οι τιμές για την *τιμή βάσης* της ακρίβειας πρόβλεψης (baseline accuracy), την οποία επιτυγχάνει το μοντέλο ZeroR προβλέποντας «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες (εν προκειμένω, οι φοιτητές) ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Οπότε με την ακρίβεια πρόβλεψης του ZeroR μοντέλου συγκρίνονται οι τιμές των μοντέλων που παράγονται με τη χρήση των άλλων αλγόριθμων ταξινόμησης (βλ. Κεφ. 9.6.1 Πίνακας 21). Όπως συμβαίνει και στο μάθημα ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ, χειμερινό εξάμηνο 2020-2021, η τιμή βάσης δεν είναι ίδια για όλες τις εβδομάδες. Αφενός, διότι όλοι οι φοιτητές δεν ξεκινούν τη χρήση της πλατφόρμας από την 1<sup>η</sup> εβδομάδα, οπότε άλλο είναι το πλήθος των φοιτητών τη 1<sup>η</sup> και άλλο τη 17<sup>η</sup> εβδομάδα. Αφετέρου, διότι οι επιτυχόντες φοιτητές (που συγκροτούν την πολυπληθέστερη κλάση), συνήθως ξεκινούν τη χρήση του ΣΔΜ νωρίτερα, οπότε το ποσοστό των επιτυχόντων είναι αυξημένο τις πρώτες εβδομάδες και μειώνεται καθώς το μάθημα βαίνει χρονικά προς την ολοκλήρωσή του, όταν καθυστερημένα συνδέονται φοιτητές, οι οποίοι τελικά, συνήθως, αποτυγχάνουν. Η υψηλότερη τιμή βάσης (85.28%) εμφανίζεται την 1<sup>η</sup> εβδομάδα και η χαμηλότερη (77.36%) την τελευταία (17<sup>η</sup>) εβδομάδα.

Στην 3<sup>η</sup> έως και στην 21<sup>η</sup> γραμμή του Πίνακα 30 εμφανίζονται οι τιμές της ακρίβειας πρόβλεψης που επιτυγχάνουν τα μοντέλα των βασικών αλγόριθμων του WEKA, οι οποίοι εκτελέστηκαν με τις προκαθορισμένες (default) παραμέτρους (βλ. Κεφ. 10.1.1 Πίνακας 23).

Ο χρωματικός κώδικας των τιμών του Πίνακα 30 είναι ίδιος με του Πίνακα 22. Δηλαδή, όσες τιμές υπερβαίνουν την τιμή βάσης (γραμμή 2) έχουν γαλάζιο (σκούρο ή ανοικτό) υπόβαθρο.

Αναλυτικότερα, σκούρο γαλάζιο υπόβαθρο έχουν όσες τιμές υπερβαίνουν την τιμή βάσης με στατιστικώς σημαντική διαφορά, με ποσοστό στατιστικής σημαντικότητας 0.05. Σε αυτές τις τιμές, δίπλα στον αριθμό υπάρχει το σύμβολο «ν». Όσες τιμές απλώς υπερβαίνουν την τιμή βάσης, χωρίς στατιστικώς σημαντική διαφορά, έχουν ανοικτό γαλάζιο υπόβαθρο. Λευκό υπόβαθρο έχουν οι τιμές που υπολείπονται της τιμής βάσης. Σε όσες από αυτές τιμές, η διαφορά από την τιμή βάσης είναι στατιστικώς σημαντική, δίπλα στον αριθμό υπάρχει το σύμβολο «\*».

#### *Γενική τάση των τιμών της ακρίβειας πρόβλεψης*

Στον Πίνακα 30 παρατηρούμε καταρχάς ότι οι τιμές της ακρίβειας πρόβλεψης που επιτυγχάνουν τα μοντέλα αυξάνονται καθώς κινούμαστε χρονικά από την 1<sup>η</sup> προς την τελευταία (17<sup>η</sup>) εβδομάδα του μαθήματος. Η αύξηση της προβλεπτικής ικανότητας των μοντέλων οφείλεται στο ότι τα σύνολα δεδομένων των τελευταίων εβδομάδων περιλαμβάνουν περισσότερες πληροφορίες, έναντι των συνόλων δεδομένων των πρώτων εβδομάδων. Οπότε, ο εντοπισμός μοτίβων (patterns) και η εκμάθηση (training) των μοντέλων είναι ευκολότερη τις τελευταίες, παρά τις πρώτες εβδομάδες. Ωστόσο, η αυξητική τάση δεν είναι συνεχώς αύξουσα, αλλά υπάρχουν σημειακές «ανωμαλίες», όπου η έως τότε αύξουσα τάση αναστρέφεται στιγμιαία και προσωρινά, Δηλαδή, υπάρχουν χρονικά σημεία, όπου αναπάντεχα, η ακρίβεια πρόβλεψης κατά την εβδομάδα W είναι μικρότερη από την εβδομάδα W-1.

Παρατηρούμε επίσης ότι η τιμή βάσης (η οποία καταγράφεται στη 2<sup>η</sup> γραμμή του Πίνακα) ακολουθεί φθίνουσα εξέλιξη. Η μεγαλύτερη τιμή παρατηρείται την 1<sup>η</sup> εβδομάδα του μαθήματος με ποσοστό 85.28% και η μικρότερη τιμή παρατηρείται την τελευταία (17<sup>η</sup>) εβδομάδα του μαθήματος με ποσοστό 77.36%. Όπως συμβαίνει και στο μάθημα AA005 Χρήση Βασικών Εφαρμογών ΗΥ, χειμερινό εξάμηνο 2020-2021, η μεταβολή της τιμής βάσης οφείλεται, αφενός στο ότι το πλήθος των φοιτητών δεν είναι το ίδιο σε όλα τα σύνολα δεδομένων και αφετέρου στο ότι το ποσοστό των επιτυχόντων φοιτητών (δηλαδή το ποσοστό της πολυπληθέστερης κλάσης το οποίο διαμορφώνει την τιμή βάσης) είναι μεγαλύτερο τις πρώτες εβδομάδες.

#### *Χρονική στιγμή πρόβλεψης*

Παρατηρούμε στον Πίνακα 30, ότι τη 17<sup>η</sup> εβδομάδα, όταν δηλαδή διαθέτουμε δεδομένα από την έναρξη του μαθήματος (1<sup>η</sup> εβδομάδα) έως την ημέρα της τελικής εξέτασης (17<sup>η</sup> εβδομάδα), όλα τα μοντέλα επιτυγχάνουν υψηλή ακρίβεια πρόβλεψης μεγαλύτερη του 80%. Επιπλέον σε

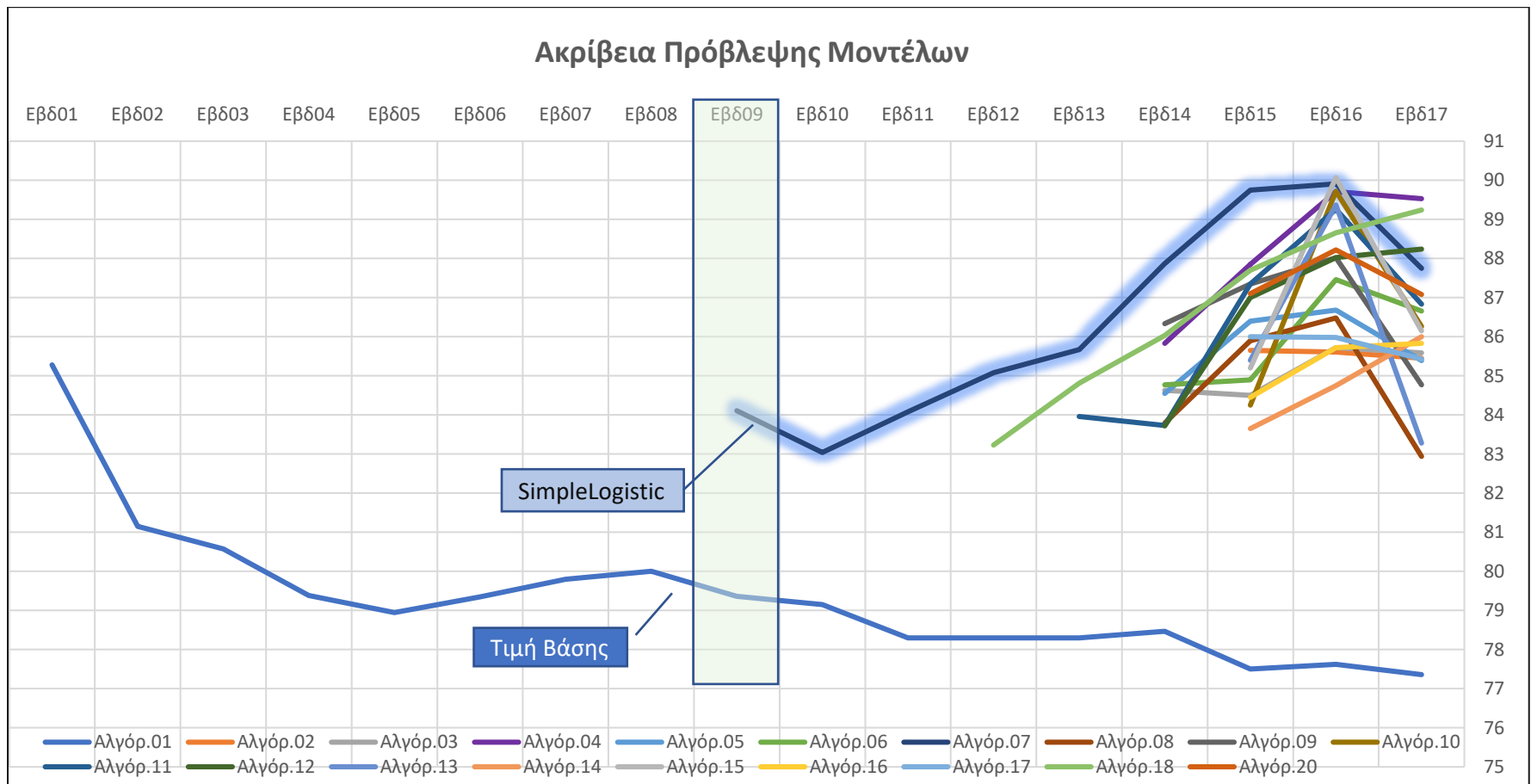


όλα τα μοντέλα, με την εξαίρεση ενός (A/A 19), η ακρίβεια πρόβλεψης είναι, όχι απλώς αρκετά μεγαλύτερη από την τιμή βάσης, αλλά η διαφορά της ακρίβειας πρόβλεψης από την τιμή βάσης, είναι στατιστικώς σημαντική. Ακόμα και η τιμή που επιτυγχάνει το μοντέλο με A/A 19, υπερβαίνει την τιμή βάσης κατά 5 περίπου εκατοστιαίες μονάδες, που όμως δεν είναι αρκετές για να χαρακτηρίσουν τη διαφορά στατιστικώς σημαντική. Η υψηλότερη τιμή ακρίβειας πρόβλεψης είναι 89.53%, διαφέρει 12 εκατοστιαίες μονάδες από την τιμή βάσης και επιτυγχάνεται από το μοντέλο του αλγόριθμου `bayes.NaiveBayesMultinomial` (A/A 4). Ακολουθούν σε ελάχιστη απόσταση τα μοντέλα του αλγόριθμου `trees.RandomForest` (A/A 18) με 89.24%, του αλγόριθμου `rules.JRip` (A/A 12) με 88.24% και του αλγόριθμου `functions.SimpleLogistic` (A/A 7) με 87.75%.

Ως προς τον χρόνο πρόβλεψης και συγκεκριμένα ως προς το πόσο νωρίς επιτυγχάνεται στατιστικώς σημαντική ακρίβεια πρόβλεψης, παρατηρούμε ότι η ακρίβεια πρόβλεψης όλων των μοντέλων, πλην ενός (A/A 19), διατηρεί την στατιστική σημαντικότητά της τη 16<sup>η</sup> και τη 15<sup>η</sup> εβδομάδα, δηλαδή 2 εβδομάδες πριν την ολοκλήρωση του μαθήματος. Τη 14<sup>η</sup> εβδομάδα, 10 από τα 19 μοντέλα επιτυγχάνουν τιμές ακρίβειας πρόβλεψης στατιστικά σημαντικές σε σχέση με την τιμή βάσης. Ένα μόνο μοντέλο, το `functions.SimpleLogistic`, κατορθώνει (παρά το έλλειμμα πληροφορίας και παρά την υψηλή τιμή βάσης) να επιτύχει στατιστικώς σημαντική τιμή (84.11%) την 9<sup>η</sup> εβδομάδα, δηλαδή στο μέσο περίπου του μαθήματος. Παρατηρούμε, ότι και το μοντέλο του αλγόριθμου `trees.RandomForest` (A/A 18) επιτυγχάνει στατιστικώς σημαντική τιμή την 9<sup>η</sup> εβδομάδα. Ωστόσο η επιτυχία του είναι «σημειακή», δηλαδή, υπάρχουν εβδομάδες μετά την 9<sup>η</sup> (π.χ. η 10<sup>η</sup>, η 11<sup>η</sup>), κατά τις οποίες το μοντέλο δεν επιτυγχάνει στατιστικώς σημαντικές τιμές. Συνεπώς, δεν μπορεί να αξιοποιηθεί στην πρώιμη πρόβλεψη, δεδομένου ότι προσφέρει ασφαλή πρόβλεψη την 9<sup>η</sup> εβδομάδα, αλλά μη ασφαλή τη 10<sup>η</sup> ή την 11<sup>η</sup> εβδομάδα. Κατ' αντιπαραβολή, η στατιστική σημαντικότητα του μοντέλου `functions.SimpleLogistic` ισχύει αδιαλείπτως για όλες τις εβδομάδες από την 9<sup>η</sup> έως τη 17<sup>η</sup>. Αξίζει επίσης να σημειωθεί, ότι κάποια μοντέλα υπερβαίνουν την τιμή βάσης νωρίτερα από την 9<sup>η</sup> εβδομάδα και ειδικότερα την 7<sup>η</sup> και την 8<sup>η</sup> εβδομάδα. Ωστόσο, η διαφορά από την τιμή βάσης δεν είναι στατιστικώς σημαντική, οπότε οι περιπτώσεις αυτές δεν τυγχάνουν περαιτέρω προσοχής.

Συγκρίνοντας τις τιμές ακρίβειας πρόβλεψης του μαθήματος 2020-2021 και του AA005 2021-2022, προκύπτει ότι τα μοντέλα του μαθήματος AA005 2021-2022 επιτυγχάνουν υψηλότερες τιμές. Ωστόσο η πρώιμη πρόβλεψη επιτυγχάνεται την 9<sup>η</sup> εβδομάδα, δηλαδή 2 εβδομάδες αργότερα από την πρώιμη πρόβλεψη στο μάθημα AA005 2020-2021. Το φαινόμενο οφείλεται στην υψηλότερη τιμή βάσης στο μάθημα AA005 2021-2022. Αν η τιμή βάσης, ήταν περίπου η ίδια, τότε και τα μοντέλα του μαθήματος AA005 2021-2022 θα επιτύγχαναν πρώιμη πρόβλεψη την 7<sup>η</sup> εβδομάδα, ίσως και νωρίτερα.

Η εξέλιξη των τιμών της ακρίβειας πρόβλεψης των μοντέλων απεικονίζεται παραστατικότερα στο Γράφημα 5 που ακολουθεί.



**Γράφημα 5. Ακρίβεια Πρόβλεψης Μοντέλων για το μάθημα ΑΑ005 2021-2022**

Στο Γράφημα 5 απεικονίζονται μόνο οι τιμές που η διαφορά τους από την τιμή βάσης είναι στατιστικώς σημαντική. Δηλαδή απεικονίζονται μόνο οι τιμές οι οποίες έχουν σκούρο γαλάζιο υπόβαθρο στον Πίνακα 30. Στον οριζόντιο άξονα του γραφήματος απεικονίζεται η χρονική περίοδος από την έναρξη του μαθήματος, η οποία προσδιορίζεται με τον αντίστοιχο αριθμό της εβδομάδας (π.χ. Εβδ10: 10<sup>η</sup> εβδομάδα από την έναρξη του μαθήματος). Στον κάθετο άξονα του γραφήματος απεικονίζεται αριθμητική κλίμακα, που αντιστοιχεί στο αριθμητικό εύρος των τιμών της ακρίβειας πρόβλεψης. Οι αριθμητικές τιμές αντιστοιχούν σε ποσοστά επί τοις εκατό (%). Κάθε μία έγχρωμη γραμμή που διατρέχει το γράφημα αντιστοιχεί στις τιμές μοντέλου το οποίο παράγεται από αλγόριθμο του Πίνακα 23. Δεν σχηματίζονται γραμμές για μία, ή περισσότερες τιμές, που είναι «απομονωμένες», δηλαδή για τιμή, ή τιμές, για τις οποίες υφίσταται μεταγενέστερη μη στατιστικώς σημαντική τιμή. Κατά συνέπεια, δεν υπάρχουν γραμμές για τα μοντέλα του αλγόριθμου με A/A 19, διότι η στατιστικά σημαντική τιμή είναι μόνο μία. Η γαλάζια γραμμή στο κάτω μέρος του γραφήματος αντιστοιχεί στις τιμές του μοντέλου του αλγόριθμου ZeroR (A/A 1), δηλαδή ορίζει την τιμή βάσης, τον «πήχη», που όλα τα μοντέλα πρέπει να υπερβούν με στατιστικώς σημαντική διαφορά. Με την εξαίρεση του μοντέλου του αλγόριθμου ZeroR (τιμή βάσης) δεν υπάρχουν τιμές (οπότε δεν σχηματίζονται γραμμές) από την 1<sup>η</sup> εβδομάδα έως και την 9<sup>η</sup> εβδομάδα, διότι κατά την συγκεκριμένη χρονική περίοδο, κανένα μοντέλο δεν επιτυγχάνει ακρίβεια πρόβλεψης στατιστικώς σημαντική σε σύγκριση με την τιμή βάσης.

Η πρωιμότερη, στατιστικώς σημαντική διαφορά από την τιμή βάσης επιτυγχάνεται την 9<sup>η</sup> εβδομάδα από το μοντέλο του αλγόριθμου functions.SimpleLogistic (A/A 7). Για λόγους ευκρίνειας, οι γραμμή του μοντέλου περιβάλλεται από ομοιόχρωμη αχλή. Η ανάδειξη στατιστικώς σημαντικής τιμής κατά την 9<sup>η</sup> εβδομάδα διευκολύνθηκε από τη μείωση της τιμής βάσης την 9<sup>η</sup> εβδομάδα, μείωση που είναι εμφανής στο γράφημα μέσω της βύθισης της γραμμής του μοντέλου ZeroR.

Στο Γράφημα 5 φαίνεται επίσης ότι η ακρίβεια πρόβλεψης των μοντέλων αυξάνει ως την 16<sup>η</sup> εβδομάδα δηλαδή μια εβδομάδα μετά την ολοκλήρωση των διαλέξεων (βλ Κεφ. 9.5.2, Πίνακας 17) και αμέσως μετά εμφανίζει πτώση. Το φαινόμενο εμφανίζεται, με μικρές διαφοροποιήσεις, σε όλα τα μοντέλα, οπότε εικάζουμε ότι δεν είναι τυχαίο. Ως εξήγηση του φαινομένου

προτείνουμε ό,τι προτείναμε για το ανάλογο φαινόμενο στο μάθημα του προηγούμενου έτους (AA005 2020-2021). Δηλαδή θεωρούμε, ότι όσο χρονικό διάστημα πραγματοποιούνται διαλέξεις, οι φοιτητές χρησιμοποιούν τακτικά το ΣΔΜ Moodle και κατά συνέπεια τα αρχεία καταγραφής εμπλουτίζονται συνεχώς με νέες πληροφορίες που βοηθούν στην «εκμάθηση» (training) των μοντέλων και στη βελτίωση της απόδοσής τους. Τη 16<sup>η</sup> εβδομάδα, όταν οι διαλέξεις έχουν μόλις ολοκληρωθεί, οι φοιτητές, είτε παύουν να χρησιμοποιούν το σύστημα, ή το χρησιμοποιούν με τρόπο λιγότερο συστηματικό από ό,τι τις προηγούμενες περιόδους, διότι επιλέγουν να μελετήσουν άλλα μαθήματα. Μεταθέτουν τη μελέτη του μαθήματός και τη σύνδεση στο σύστημα όταν θα πλησιάζει η ημέρα της τελικής εξέτασης. Η εν λόγω τακτική των φοιτητών έχει ως αποτέλεσμα να εισάγεται στο σύστημα μη συστηματική, ακανόνιστη, άρρυθμη πληροφορία η οποία προκαλεί αταξία στα σύνολα δεδομένων και η οποία «θολώνει», αλλοιώνει, αποδιοργανώνει τα μοτίβα (patterns), οδηγώντας σε πτώση την ακρίβεια πρόβλεψης των μοντέλων. Όμως, μια κρίσιμη διαφορά στο χρονικό σχεδιασμό του AA005 2020-2021 σε σχέση με το AA005 2021-2022 επηρεάζει την απόδοση των μοντέλων την τελευταία εβδομάδα. Ειδικότερα, το έτος 2020-2021, μεταξύ της τελευταίας διάλεξης και της τελικής εξέτασης μεσολαβεί χρονικό διάστημα ενός μηνός (βλ. Κεφ. 9.5.2, Πίνακας 17). Οπότε, εκτιμούμε ότι, λίγο πριν την τελική εξέταση, οι φοιτητές θεώρησαν αναγκαίο να συνδεθούν στο ΣΔΜ για να μελετήσουν ξανά το εκπαιδευτικό υλικό. Η συστηματική, εύρυθμη, εντατική χρήση του ΣΔΜ Moodle, ενόσω πλησιάζει η ημερομηνία για την τελική εξέταση του μαθήματος, έχει ως αποτέλεσμα την ανάκτηση των υψηλών τιμών στην ακρίβεια πρόβλεψης κατά την τελευταία εβδομάδα. Αντιστοίχως, κατά το έτος 2021-2022, μεταξύ της τελευταίας διάλεξης και της τελικής εξέτασης μεσολαβούν μόνο 12 ημέρες, κατά τις οποίες η χρήση του ΣΔΜ, όσο έντονη και να είναι, φαίνεται να μην επαρκεί για την ανάκτηση, κατά την 17<sup>η</sup> εβδομάδα, των μέγιστων τιμών ακρίβειας πρόβλεψης που επιτεύχθηκαν την 16<sup>η</sup>.

#### 10.2.2. Επιπλέον κριτήρια για την απόδοση των βασικών μοντέλων

Προκειμένου να αποτιμήσουμε την απόδοση των μοντέλων υπολογίσαμε εκτός από τις τιμές της ακρίβειας πρόβλεψης και τις τιμές της μεταβλητής AUC (Area Under Curve) (Witten, Frank & Hall, 2011, σ. 172) και της μεταβλητής Cohen's Kappa (Witten, Frank, & Hall, 2011, σ. 166).

Όπως έχουμε αναφέρει στο Κεφ. 9.6.3, η μεταβλητή AUC αξιολογεί τη διαγνωστική ικανότητα ενός μοντέλου ταξινόμησης με διχοτομημένη κλάση και μη ισορροπημένο σύνολο δεδομένων (Hussain M. , Zhu, Zhang, & Abidi, 2018) και υπολογίζει την περιοχή που ορίζει η καμπύλη ROC (Receiver Operating Characteristic) (Σχήμα 13).

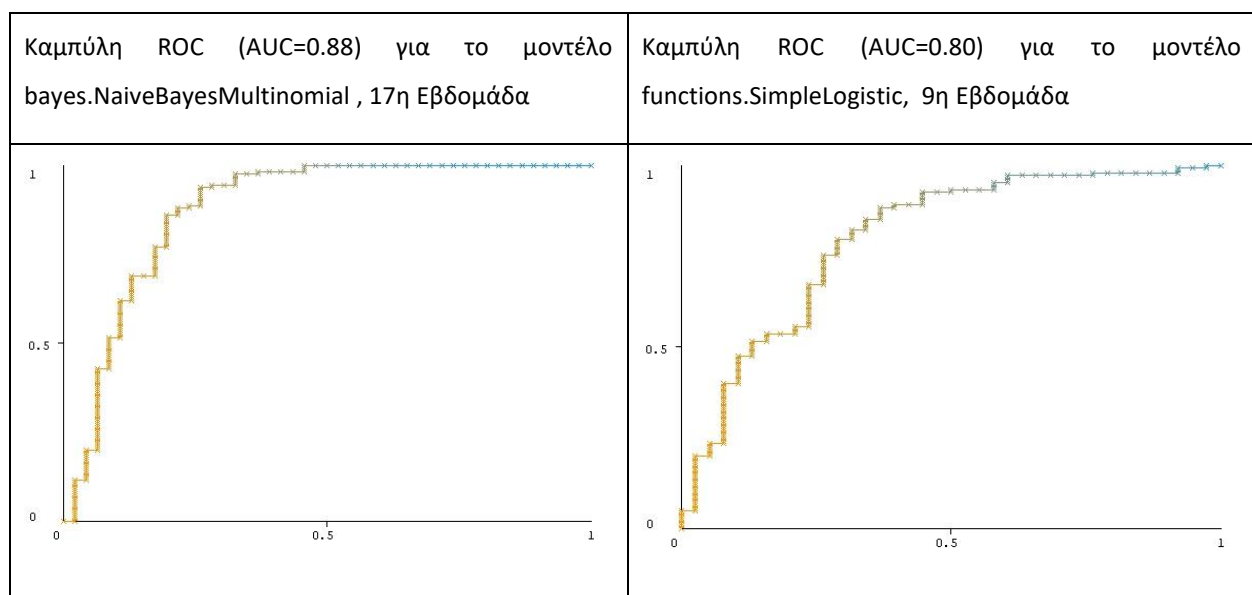
**Πίνακας 31. Τιμές μεταβλητής AUC για τα μοντέλα του μαθήματος AA005 2021-2022**

Εβδομάδα	01η	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>
Αλγορ.01	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Αλγορ.02	0.5	0.5	0.5	0.5	0.71	0.69	0.78	0.8	0.8	0.8	0.79	0.8	0.78	0.85	0.87	0.87	0.89
Αλγορ.03	0.63	0.45	0.56	0.61	0.64	0.66	0.75	0.77	0.73	0.74	0.78	0.78	0.78	0.82	0.85	0.87	0.89
Αλγορ.04	0.64	0.59	0.61	0.65	0.7	0.72	0.78	0.8	0.78	0.78	0.82	0.81	0.81	0.84	0.88	0.88	0.88
Αλγορ.05	0.48	0.5	0.58	0.6	0.72	0.72	0.79	0.82	0.82	0.79	0.81	0.8	0.8	0.85	0.84	0.87	0.88
Αλγορ.06	0.4	0.45	0.63	0.56	0.67	0.67	0.73	0.82	0.81	0.82	0.82	0.82	0.79	0.82	0.83	0.87	0.89
Αλγορ.07	0.54	0.5	0.57	0.5	0.51	0.56	0.78	0.81	0.8	0.81	0.83	0.83	0.83	0.87	0.88	0.89	0.89
Αλγορ.08	0.45	0.46	0.53	0.54	0.52	0.52	0.54	0.64	0.73	0.68	0.69	0.69	0.66	0.73	0.77	0.8	0.74
Αλγορ.09	0.31	0.52	0.58	0.55	0.61	0.64	0.71	0.75	0.73	0.72	0.73	0.75	0.71	0.79	0.84	0.87	0.85
Αλγορ.10	0.52	0.39	0.58	0.63	0.71	0.71	0.74	0.77	0.76	0.74	0.78	0.79	0.79	0.79	0.85	0.86	0.87
Αλγορ.11	0.5	0.5	0.5	0.5	0.5	0.51	0.76	0.79	0.77	0.79	0.79	0.78	0.81	0.81	0.82	0.83	0.84
Αλγορ.12	0.5	0.49	0.49	0.5	0.5	0.5	0.67	0.7	0.66	0.65	0.7	0.69	0.7	0.74	0.78	0.77	0.79
Αλγορ.13	0.5	0.48	0.51	0.49	0.51	0.51	0.6	0.64	0.63	0.67	0.66	0.66	0.72	0.74	0.74	0.77	0.73
Αλγορ.14	0.55	0.52	0.54	0.55	0.66	0.6	0.6	0.69	0.67	0.72	0.73	0.73	0.7	0.77	0.78	0.79	0.85
Αλγορ.15	0.48	0.51	0.49	0.64	0.64	0.61	0.64	0.68	0.68	0.64	0.72	0.71	0.74	0.76	0.75	0.78	0.74
Αλγορ.16	0.49	0.5	0.5	0.5	0.5	0.5	0.51	0.58	0.54	0.53	0.62	0.65	0.67	0.81	0.85	0.87	0.89
Αλγορ.17	0.51	0.49	0.5	0.6	0.65	0.63	0.56	0.68	0.7	0.72	0.73	0.7	0.71	0.72	0.77	0.8	0.81
Αλγορ.18	0.53	0.51	0.55	0.6	0.63	0.65	0.76	0.81	0.79	0.79	0.8	0.8	0.79	0.83	0.89	0.9	0.91
Αλγορ.19	0.47	0.46	0.5	0.5	0.55	0.55	0.59	0.64	0.66	0.64	0.66	0.68	0.68	0.7	0.74	0.76	0.77
Αλγορ.20	0.5	0.5	0.51	0.5	0.51	0.53	0.57	0.62	0.64	0.63	0.66	0.69	0.69	0.75	0.79	0.78	0.79

Στον Πίνακα 31, καταρχάς παρατηρούμε ότι η τιμή AUC για το μοντέλο του αλγόριθμου ZeroR (A/A 1) ισούται με 0.5, γεγονός απολύτως αναμενόμενο, διότι όπως έχουμε ήδη αναφέρει (βλ Κεφ. 9.6.2.) το συγκεκριμένο μοντέλο προβλέπει «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Συνεπώς, η τιμή AUC δεν έχει νόημα στο συγκεκριμένο μοντέλο, το οποίο χρησιμοποιείται για τον υπολογισμό της τιμής βάσης. Όλα τα άλλα μοντέλα, σε όλες τις χρονικές περιόδους (με την εξαίρεση λίγων τιμών τις

τέσσερις πρώτες εβδομάδες), σημειώνουν τιμές που είναι μεγαλύτερες από το 0.5, οπότε κρίνονται αποδεκτές (Riestra-Gonzalez, del Puerto Paule-Ruiz, & Ortin, 2021). Τη 17<sup>η</sup> εβδομάδα, με την ολοκλήρωση του μαθήματος, οι τιμές AUC για τα περισσότερα μοντέλα είναι μεγαλύτερες από 0.8, οπότε κρίνονται εξαιρετικές. Επιπλέον, παρατηρούμε ότι την 9<sup>η</sup> εβδομάδα για το μοντέλο του αλγόριθμου functions.SimpleLogistic (A/A 7), το οποίο επιτυγχάνει πρώτη πρόβλεψη, η τιμή AUC είναι 0.8 και κρίνεται εξαιρετική.

Στο Γράφημα 6 που ακολουθεί, αναπαρίστανται οι καμπύλες ROC, που υπολογίζουν ενδεικτικές τιμές AUC και συγκεκριμένα τις τιμές AUC της 17<sup>ης</sup> εβδομάδας για το μοντέλο bayes.NaiveBayesMultinomial (A/A 4) και οι τιμές AUC της 9<sup>ης</sup> εβδομάδας για το μοντέλο functions.SimpleLogistic (A/A 7).



**Γράφημα 6. Καμπύλες ROC για τα μοντέλα του μαθήματος AA005 2021-2022**

Όπως έχουμε αναφέρει στο Κεφ. 9.6.3, η μεταβλητή Cohen’s Kappa μετρά τη συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που πρόβλεψε το μοντέλο, ενώ διορθώνει τυχόν συμφωνία που προέκυψε τυχαίως (Witten, Frank, & Hall, 2011, σ. 166).

**Πίνακας 32. Τιμές μεταβλητής Cohen's Kappa για τα μοντέλα του μαθήματος AA005 2021-2022**

Εβδομ	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>
Αλγορ.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Αλγορ.02	0	0	-0.01	0	0.18	0.15	0.35	0.44	0.45	0.44	0.49	0.48	0.49	0.57	0.61	0.62	0.61
Αλγορ.03	0.15	-0.02	0.07	0.13	0.15	0.24	0.36	0.39	0.4	0.35	0.43	0.44	0.45	0.59	0.58	0.61	0.63
Αλγορ.04	0.11	-0.01	-0.01	0	0	0	0.12	0.18	0.09	0.05	0.17	0.19	0.23	0.49	0.55	0.62	0.63
Αλγορ.05	0.18	0.04	0.14	0.06	0.2	0.15	0.37	0.47	0.42	0.43	0.43	0.43	0.46	0.52	0.59	0.61	0.57
Αλγορ.06	0.05	-0.02	0.09	0.02	0.04	0.05	0.13	0.36	0.34	0.37	0.42	0.45	0.42	0.52	0.55	0.62	0.6
Αλγορ.07	0.08	-0.01	0.01	0	-0.01	-0.01	0.26	0.36	0.42	0.37	0.47	0.5	0.53	0.6	0.65	0.66	0.58
Αλγορ.08	-0.11	-0.09	0.07	0.07	0.03	0.04	0.08	0.27	0.44	0.36	0.35	0.36	0.31	0.48	0.56	0.6	0.49
Αλγορ.09	-0.1	-0.06	0	-0.02	-0.06	0.03	0.13	0.3	0.32	0.31	0.31	0.32	0.42	0.56	0.59	0.61	0.51
Αλγορ.10	0.14	-0.07	-0.01	-0.01	0	0	0.26	0.44	0.37	0.31	0.46	0.46	0.51	0.45	0.5	0.63	0.54
Αλγορ.11	0	-0.01	-0.01	0	0	-0.01	0.32	0.4	0.34	0.36	0.42	0.43	0.45	0.45	0.58	0.65	0.57
Αλγορ.12	0	-0.01	-0.02	0	-0.01	-0.01	0.34	0.44	0.36	0.34	0.41	0.41	0.43	0.48	0.59	0.59	0.62
Αλγορ.13	0	-0.04	0.02	-0.03	0.03	0.02	0.22	0.32	0.3	0.35	0.35	0.36	0.46	0.47	0.52	0.62	0.48
Αλγορ.14	0.07	0.05	0.05	0.03	0.14	0.14	0.16	0.28	0.29	0.3	0.36	0.35	0.36	0.49	0.51	0.55	0.58
Αλγορ.15	0	-0.02	-0.01	0	0	-0.01	0.21	0.36	0.34	0.24	0.46	0.44	0.51	0.47	0.52	0.64	0.54
Αλγορ.16	0.02	0	0	0	0	0	0.02	0.11	0.06	0.06	0.18	0.23	0.26	0.56	0.58	0.61	0.63
Αλγορ.17	-0.02	-0.01	-0.05	0.12	0.17	0.13	0.12	0.33	0.36	0.36	0.36	0.34	0.39	0.46	0.57	0.57	0.56
Αλγορ.18	-0.04	-0.08	0.03	-0.01	-0.05	0.04	0.21	0.41	0.39	0.31	0.39	0.43	0.49	0.53	0.6	0.63	0.65
Αλγορ.19	-0.03	-0.07	0	0	0.1	0.1	0.18	0.29	0.32	0.28	0.32	0.35	0.34	0.41	0.47	0.52	0.52
Αλγορ.20	0	0	0.01	-0.01	-0.01	0.01	0.11	0.27	0.24	0.25	0.28	0.34	0.38	0.46	0.58	0.6	0.58

Στον Πίνακα 32 καταρχάς παρατηρούμε ότι για το μοντέλο του αλγόριθμου ZeroR (A/A 1) η τιμή της μεταβλητής Cohen's Kappa ισούται με μηδέν, διότι όπως έχουμε ήδη αναφέρει (βλ. Κεφ. 9.6.2.) το συγκεκριμένο μοντέλο προβλέπει «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Παρατηρούμε επίσης, ότι κατά την 17<sup>η</sup> εβδομάδα, η μεταβλητή Cohen's Kappa για όλα τα μοντέλα λαμβάνει τιμές από το διάστημα [0.48 , 0.65], γεγονός που σύμφωνα με τους Quinn & Gray (2020, σ. 8) δηλώνει «μέση» (moderate) και «σημαντική» (substantial) συμφωνία. Ωστόσο, τις εβδομάδες πριν την ολοκλήρωση του μαθήματος, οι τιμές της μεταβλητής, όπως είναι αναμενόμενο, μειώνονται έως ότου κατά τις πρώτες εβδομάδες δηλώνουν «αδύναμη» (slight) συμφωνία. Ωστόσο, την 9<sup>η</sup> εβδομάδα, δηλαδή τη χρονική στιγμή της πρώιμης πρόβλεψης από το μοντέλο functions.SimpleLogistic, η μεταβλητή Cohen's Kappa έχει τιμή 0.42, η οποία δηλώνει «μέση»



(moderate) συμφωνία. Συνεπώς, σύμφωνα με το κριτήριο της μεταβλητής Cohen's Kappa, κατά τη χρονική στιγμή ολοκλήρωσης του μαθήματος, υπάρχει «μέση» και «σημαντική» συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που προέβλεψαν τα μοντέλα, ενώ η ισχύς της συμφωνίας εξασθενεί βαθμιαία κατά τις προηγούμενες χρονικές περιόδους.

### 10.2.3. Ακρίβεια και χρόνος πρόβλεψης του LSTM μοντέλου

Εκτός από τη μελέτη των μοντέλων τα οποία κατασκευάζονται από τους βασικούς αλγόριθμους (αυτών του Πίνακα 23), μελετήσαμε επίσης την απόδοση μοντέλων τα οποία παράγονται, όχι μόνο με τη βοήθεια βασικών αλγόριθμων (αυτών του Πίνακα 23), αλλά και αλγόριθμων που βασίζονται σε περιοδικά νευρωνικά δίκτυα (Recursive Neural Networks), που αναλύουν όχι μόνο τις συνολικές, τελικές τιμές των χαρακτηριστικών του συνόλου δεδομένων, αλλά και τις χρονικές ακολουθίες, οι οποίες αποτυπώνουν την εξέλιξη των τιμών. Η μελέτη τέτοιων τεχνικών επί εκπαιδευτικών δεδομένων αποτελεί πεδίο αυξανόμενου ερευνητικού ενδιαφέροντος (Doleck, Lemay, Basnet & Bazalais, 2020). Επιπλέον, οι ερευνητικές προσπάθειες που διερευνούν την προβλεπτική ικανότητα τέτοιων τεχνικών είναι σχετικά λίγες και δεν συγκλίνουν προς κοινώς αποδεκτά συμπεράσματα (Chen & Cui, 2020. Doleck, Lemay, Basnet & Bazalais, 2020). Επιλέξαμε την LSTM υλοποίηση των περιοδικών νευρωνικών δικτύων. Καταρχάς χρησιμοποιήσαμε τις προκαθορισμένες (default) παραμέτρους του αλγόριθμου με τις οποίες προέκυψαν οι τιμές που καταγράφονται στον Πίνακα 33

**Πίνακας 33. Ακρίβεια πρόβλεψης LSTM μοντέλου με προκαθορισμένες παραμέτρους για το μάθημα AA005 2021-2022 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>
ZeroR	85.28	81.15	80.57	79.38	78.95	79.35	79.8	80	79.36	79.15	78.3	78.3	78.3	78.47	77.5	77.62	77.36
LSTM	85.28	81.15	80.57	79.38	78.95	79.35	79.8	80	79.36	79.15	78.3	78.3	78.3	78.47	77.5	77.62	77.36

Από τις τιμές του Πίνακα 33 προκύπτει, ότι η ακρίβεια πρόβλεψης του LSTM μοντέλου, ποτέ δεν ξεπερνά την τιμή βάσης, δηλαδή το μοντέλο δεν δίνει γόνιμα αποτελέσματα. Προκειμένου να επιτύχουμε τιμές που θα υπερβαίνουν την τιμή βάσης, χρησιμοποιήσαμε τις «βελτιστοποιημένες» παραμέτρους του αλγόριθμου, ο οποίος κατασκεύασε το μοντέλο του μαθήματος AA005 2020-2021 (βλ. Παράρτημα Β). Όπως έχουμε αναφέρει (βλ. Κεφ. 10.1.3), οι βελτιστοποιημένες παράμετροι προέκυψαν με εφαρμογή εμπειρικών κανόνων, που υποδεικνύονται από τη σχετική βιβλιογραφία, σύμφωνα με την οποία, η εύρεση ενός

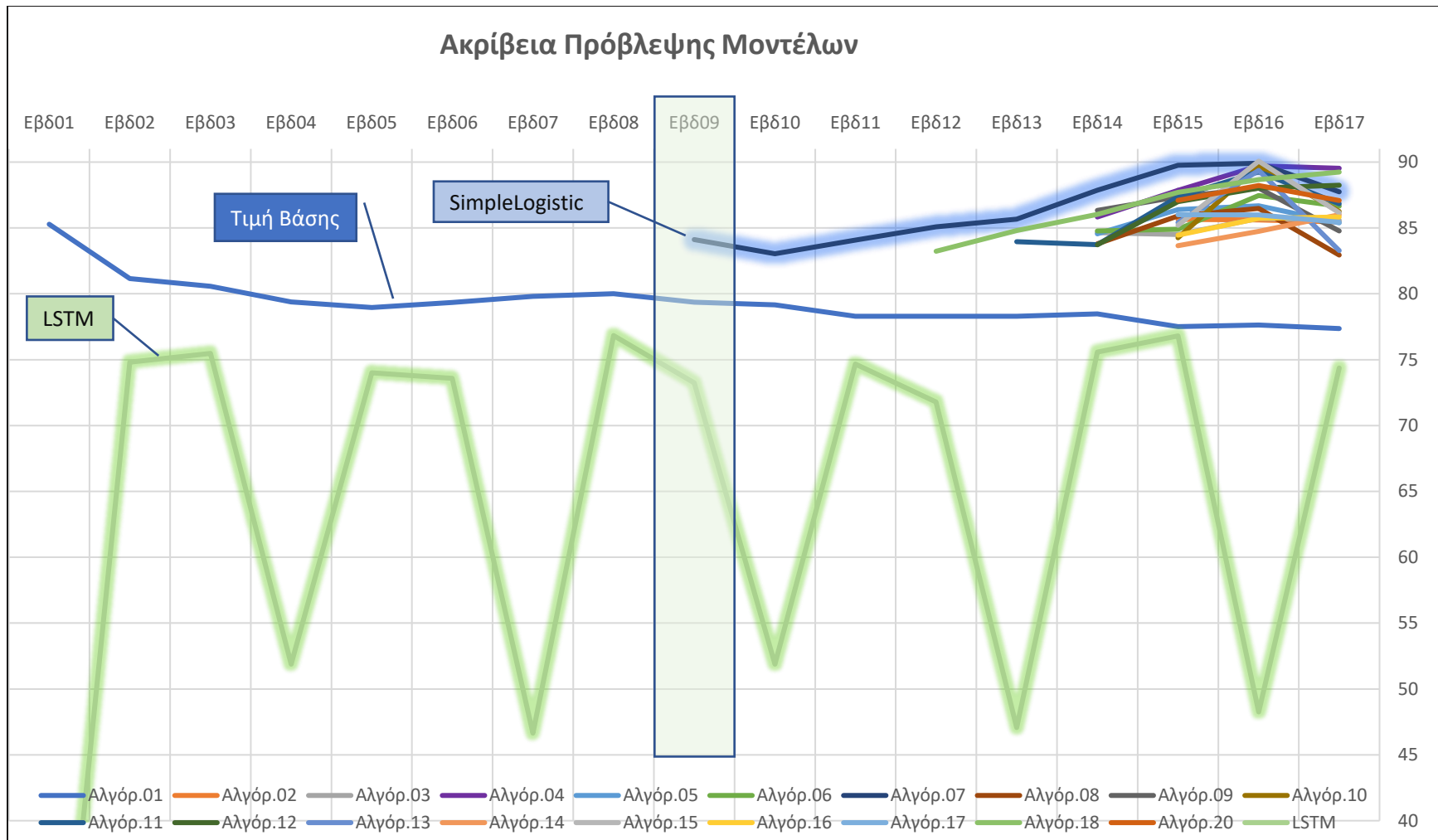
αποδοτικού LSTM μοντέλου αποτελεί «άσκηση δοκιμής και λάθους» (an exercise in trial-and-error) (Frank, 2021). Οι νέες τιμές καταγράφονται στον Πίνακα 34

**Πίνακας 34. Ακρίβεια πρόβλεψης LSTM μοντέλου με «βελτιστοποίηση» ως προς τις παραμέτρους για το μάθημα AA005 2021-2022 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>
ZeroR	85.28	81.15	80.57	79.38	78.95	79.35	79.8	80	79.36	79.15	78.3	78.3	78.3	78.47	77.5	77.62	77.36
LSTM	14.72	74.81	75.47	51.88	74.01	73.57	46.64	76.83	73.23	51.88	74.66	71.78	47.07	75.57	76.80	48.25	74.37

Στον Πίνακα 34 παρατηρούμε ότι σε όλες τις χρονικές στιγμές το LSTM μοντέλο, όχι μόνο δεν επιτυγχάνει στατιστικώς σημαντική τιμή, αλλά ούτε καν καταφέρνει να υπερβεί την τιμή βάσης. Το γεγονός δεν αποτελεί απόδειξη, ότι για τα σύνολα δεδομένων του μαθήματος AA005 2021-2022 δεν μπορούν να κατασκευαστούν LSTM μοντέλα που να επιτυγχάνουν καλές αποδόσεις. Ενδεχομένως, να υπάρχει παραμετροποίηση η οποία να αποδειχθεί επιτυχής. Ωστόσο, εμείς, παρότι προσπαθήσαμε, ακολουθώντας τους εμπειρικούς κανόνες που αποδείχθηκαν αποτελεσματικοί στο μάθημα AA005 2020-2021, δεν καταφέραμε να έχουμε γόνιμο αποτέλεσμα και στο μάθημα AA005 2021-2022.

Η αδυναμία του LSTM μοντέλου να επιτύχει τιμές έστω μεγαλύτερες από την τιμή βάσης αποτυπώνεται επίσης στο Γράφημα 7.



**Γράφημα 7. Ακρίβεια Πρόβλεψης LSTM Μοντέλου**

#### 10.2.4 Επιπλέον κριτήρια για την απόδοση του LSTM μοντέλου

Παρά τα άγωνα αποτελέσματα, ως προς την ακρίβεια πρόβλεψης του LSTM μοντέλου, υπολογίσαμε και τις τιμές των μεταβλητών AUC (Area Under Curve) και Cohen's Kappa. Στους Πίνακες 35 και 36 που ακολουθούν, φαίνεται ότι και για αυτές τις μεταβλητές, το LSTM μοντέλο δεν παράγει γόνιμα αποτελέσματα. Ειδικότερα η μεταβλητή AUC υπερβαίνει ελάχιστα την τιμή βάσης (υπάρχουν μάλιστα περιπτώσεις που υπολείπεται της βάσης), ενώ η μεταβλητή Cohen's Kappa λαμβάνει σχεδόν μηδενικές τιμές.

**Πίνακας 35. Τιμές μεταβλητής AUC του LSTM μοντέλου για το μάθημα AA005 2021-2022**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>
ZeroR	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
LSTM	0.5	0.52	0.51	0.53	0.51	0.5	0.49	0.5	0.51	0.50	0.49	0.51	.51	0.51	0.55	0.48	0.51

**Πίνακας 36. Τιμές μεταβλητής Cohen's Kappa του LSTM μοντέλου για το μάθημα AA005 2021-2022**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>
ZeroR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LSTM	0	0	0	0.04	0.01	-0.01	-0.02	0	0	0	0.02	-0.01	0.02	-0.01	0.09	-0.02	0.07

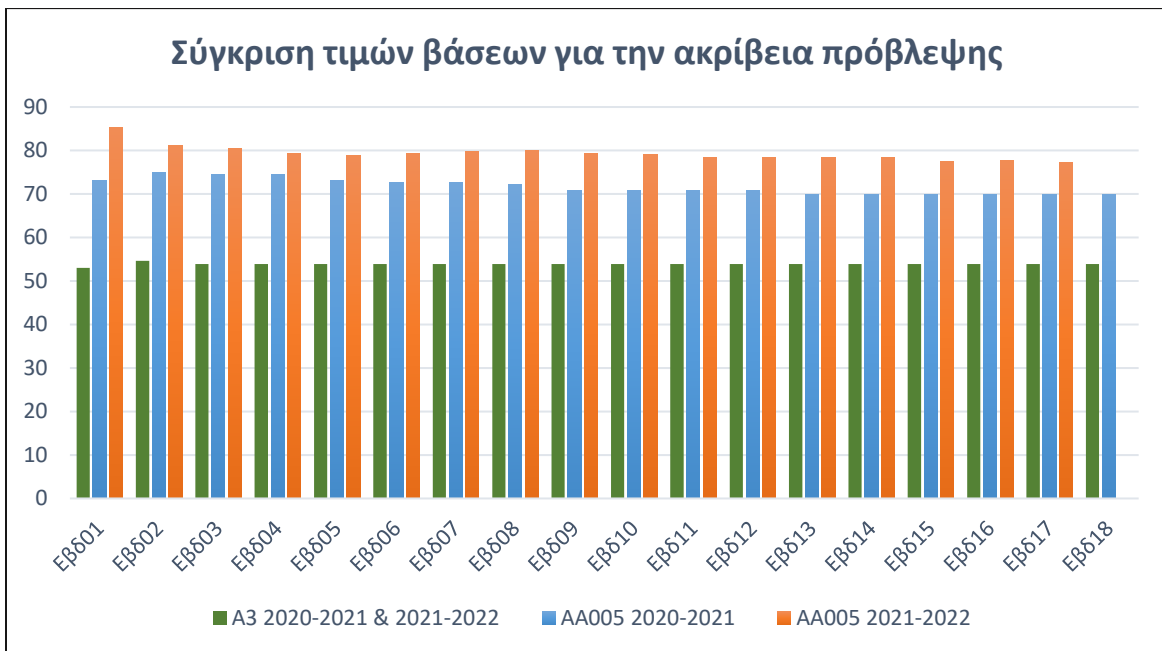
### 10.3. Τα μοντέλα πρόβλεψης για τα Μαθήματα Α3 2019-2020 & 2020-2021

#### 10.3.1. Ακρίβεια και χρόνος πρόβλεψης των βασικών μοντέλων

Για την πληρέστερη κατανόηση των αποτελεσμάτων που αφορούν τα μαθήματα Α3 Ηλεκτρονική Μάθηση, 2019-2020 & 2021-2022 είναι χρήσιμο να γίνει μνεία στις διαφορές των μαθημάτων σε σχέση με τα μαθήματα ΑΑ005, 2020-2021 και ΑΑ005, 2021-2022.

- Υφίσταται διαφορά ως προς το πλήθος των οντοτήτων στα σύνολα δεδομένων. Ο μικρός αριθμός των φοιτητών ανά έτος (39) και το μικρό σύνολο δεδομένων που προέκυψε δεν ενδείκνυται για μελέτη με τεχνικές Μηχανικής Μάθησης. Οπότε, συγχωνεύσαμε τα σύνολα δεδομένων των 2 ετών και δημιουργήσαμε σύνολα δεδομένων που περιλαμβάνουν δεδομένα από τους 78 φοιτητές και των 2 ετών. Η συγκεκριμένη επιλογή είναι έγκυρη, δεδομένου ότι οι φοιτητές προέρχονται από το ίδιο μάθημα, το οποίο κατά τη διάρκεια των δύο ετών προσφέρθηκε με το ίδιο πρόγραμμα σπουδών, με το ίδιο εκπαιδευτικό υλικό, με τον ίδιο τρόπο αξιολόγησης, με τους ίδιους διδάσκοντες και με αντίστοιχο χρονοδιάγραμμα υλοποίησης. Η συγχώνευση των εγγραφών δύο και περισσότερων μαθημάτων αποτελεί μεθοδολογική επιλογή που απαντάται στην βιβλιογραφία όπως αναλυτικά αναφέρεται στο κεφάλαιο 9.5.2.
- Υφίσταται διαφορά ως προς το εύρος των βαθμών (της τελικής επίδοσης των φοιτητών). Το εύρος των βαθμών των μαθημάτων Α3 Ηλεκτρονική Μάθηση είναι μικρότερο από το εύρος των βαθμών των μαθημάτων ΑΑ005 Χρήση Βασικών Εφαρμογών ΗΥ. Ειδικότερα, στα μαθήματα ΑΑ005, η κατηγορική μεταβλητή που αποτελεί την κλάση του συνόλου δεδομένων τιτλοφορείται «Επιτυχία» (Pass) και λαμβάνει τιμή «Yes» για τους βαθμούς που είναι μεγαλύτεροι ή ίσοι του 5 και τιμή «No» για τους υπόλοιπους βαθμούς. Στα μαθήματα Α3 Ηλεκτρονική Μάθηση, δεν υπάρχουν φοιτητές με βαθμό μικρότερο του 5. Ειδικότερα, οι βαθμοί κυμαίνονται από 5 ως 10, με διάμεσο τον αριθμό 8.5. Συνεπώς, δεν είχε νόημα να κατηγοριοποιήσουμε, όπως ακριβώς στο μάθημα ΑΑ005, δηλαδή να διαχωρίσουμε τους φοιτητές σε επιτυχόντες και μη, διότι όλοι οι φοιτητές είναι επιτυχόντες. Οπότε, ελλείψει αποτυχόντων, επιλέξαμε να δώσουμε στην κατηγορική μεταβλητή που αποτελεί την κλάση, τον τίτλο «Εύγε» (Well Done) και να της

εκχωρήσουμε τιμές από το σύνολο {Yes, No}. Προκειμένου να έχουμε ισόρροπη κατανομή των οντοτήτων στις δύο κατηγορίες, διαχωρίσαμε τις δύο κατηγορίες με κριτήριο τη διάμεσο. Δηλαδή, η τιμή «Yes» αντιστοιχεί σε βαθμούς μεγαλύτερους από τη διάμεσο (το 8.5) και η τιμή «No» στους υπόλοιπους βαθμούς (βλ. Κεφ. 9.5.9). Σημαντική συνέπεια του διαχωρισμού των δύο κατηγοριών με κριτήριο τη διάμεσο είναι η χαμηλή τιμή βάσης για την ακρίβεια πρόβλεψης στα μαθήματα Α3. Διότι, η πολυπληθέστερη κλάση (η οποία ως γνωστόν διαμορφώνει την τιμή βάσης) ελάχιστα υπερβαίνει (σε πλήθος φοιτητών) την άλλη κλάση. Οπότε, το ποσοστό όσων φοιτητών ανήκουν στην κατηγορία «Yes» είναι λίγο μεγαλύτερο από το 50%. Η χαμηλή τιμή βάσης για την ακρίβεια πρόβλεψης στα μαθήματα Α3 διευκολύνει τα μοντέλα να επιτύχουν στατιστικώς σημαντικές τιμές για την ακρίβεια πρόβλεψης. Στο Γράφημα 8 αναπαρίστανται οι διαφορές στις τιμές βάσης των 4 μαθημάτων.



**Γράφημα 8. Τιμές βάσεων για την ακρίβεια πρόβλεψης στα μαθήματα AA005 και Α3.**

- Υφίσταται διαφορά ως προς το φαινόμενο των διακυμάνσεων στο πλήθος των οντοτήτων στα σύνολα δεδομένων. Συγκεκριμένα, η παρακολούθηση στα μαθήματα Α3 (σε αντίθεση με την παρακολούθηση στα μαθήματα AA005) είναι υποχρεωτική για όλες

τις εβδομάδες και ως εκ τούτου είναι υποχρεωτική η σύνδεση των φοιτητών στο ΣΔΜ από την έναρξη των μαθημάτων. Ήδη από την 3<sup>η</sup> εβδομάδα τα σύνολα δεδομένων είναι πλήρη ως προς τους συμμετέχοντες φοιτητές, οπότε η τιμή βάσης για την ακρίβεια πρόβλεψης παραμένει σταθερή από τη 3<sup>η</sup> εβδομάδα ως την ολοκλήρωση των μαθημάτων, σε αντίθεση με τα μαθήματα AA005, όπου η τιμή βάσης αλλάζει συνεχώς έως λίγο πριν την ολοκλήρωση των μαθημάτων, διότι έως τότε συνδέονται φοιτητές για πρώτη φορά.

- Υφίσταται διαφορά ως προς τα χαρακτηριστικά του συνόλου δεδομένων, η οποία απορρέει από τις επιλογές στον εκπαιδευτικό σχεδιασμό του Α3. Συγκεκριμένα, στο ΣΔΜ, σε κάθε συνεδρία των μαθημάτων Α3 διεξάγονται δημόσιοι διάλογοι μεταξύ των φοιτητών. Η συμμετοχή στους δημόσιους διαλόγους αξιολογείται και προσμετράται στον τελικό βαθμό κατά ποσοστό 30%, ενώ κατά 70% προσμετράται η τελική αξιολογική εργασία. Κατά συνέπεια, η συμμετοχή στους δημόσιους διαλόγους συνιστά σημαντικό δείκτη για τη χρήση ΣΔΜ και γενικότερα για την ενασχόληση με το μάθημα. Τα δεδομένα του ΣΔΜ που αφορούν αναρτήσεις στους δημόσιους διαλόγους αποτυπώνουν μοτίβα (patterns) χρήσης που διευκολύνουν τους αλγόριθμους ταξινόμησης να δημιουργήσουν αποδοτικά μοντέλα. Οπότε, στα σύνολα δεδομένων του Α3 περιλαμβάνονται τα χαρακτηριστικά «Πλήθος συμβάντων πρόσβασης στους χώρους δημόσιου διαλόγου» (A/A 25) και «Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου» (A/A 26), τα οποία καταγράφουν δραστηριότητα σχετική με τους χώρους δημόσιου διαλόγου. Επιπλέον, η αξιολόγηση της συμμετοχής στους δημόσιους διαλόγους, δεδομένου ότι προσμετράται στον τελικό βαθμό, επηρεάζει σημαντικά το βαθμό των φοιτητών και κατά συνέπεια την κλάση κατάταξης. Οπότε, στα σύνολα δεδομένων του Α3 περιλαμβάνεται το χαρακτηριστικό «Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο» (A/A 27).

Στον Πίνακα 37 εμφανίζονται οι τιμές για την ακρίβεια πρόβλεψης των μοντέλων που κατασκεύασαν 20 βασικοί αλγόριθμοι σε 18 σύνολα δεδομένων των μαθημάτων Α3 Ηλεκτρονική Μάθηση, χειμερινό εξάμηνο 2019-2020 & 2020-2021



**Πίνακας 37. Ακρίβεια πρόβλεψης μοντέλων για τα μαθήματα Α3 2019-2020 & 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
Αλγορ.01	53	54.64	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93
Αλγορ.02	53	65.96	65.23 v	63.80 v	72.87 v	71.41 v	72.41 v	71.86 v	72.16 v	72.73 v	68.00 v	70.32 v	71.68 v	71.84 v	72.11 v	68.09 v	70.05 v	69.09 v
Αλγορ.03	46.33	63.55	62.98	62.29	67.45 v	67.93 v	70.25 v	72.11 v	69.27 v	70.55 v	68.21 v	72.91 v	69.61 v	68.82 v	70.68 v	69.52 v	70.20 v	69.64 v
Αλγορ.04	50.6	60.71	65.05 v	66.00 v	69.45 v	70.68 v	72.80 v	70.73 v	69.86 v	69.25 v	70.48 v	69.09 v	70.89 v	71.55 v	68.95 v	69.98 v	61.43	67.25 v
Αλγορ.05	53.47	64.68	59.18	65.14	69.66 v	69.73 v	71.30 v	71.16 v	71.11 v	68.57 v	70.23 v	70.43 v	69.30 v	69.18 v	69.64 v	69.86 v	68.50 v	69.48 v
Αλγορ.06	55.57	56.07	55.91	61.39	61.02	65.23	59.55	60.46	60.23	63.46	63.09	64.23	60.45	63.45	62.95	64.91	62.21	65.46
Αλγορ.07	47.73	63.39	63.89	65.95 v	71.71 v	71.50 v	77.70 v	79.14 v	75.91 v	76.68 v	77.27 v	77.21 v	77.68 v	77.32 v	75.37 v	75.14 v	75.02 v	75.27 v
Αλγορ.08	54.87	63.05	55.41	60.91	61.96	57.95	50.54	58.43	50.98	52.77	49.2	61.11	50.73	57.8	51.45	55.87	59.77	55.09
Αλγορ.09	55	57.02	49.48	57.59	60.07	64.5	59.73	67.68 v	64.93	66.14	59.04	64.55	65.59 v	64.48	65.05	57.55	61.34	58.77
Αλγορ.10	52.87	65.16	62.5	61.39	67.91 v	73.73 v	79.46 v	79.57 v	80.75 v	79.21 v	79.50 v	79.71 v	80.00 v	78.88 v	79.86 v	80.37 v	80.11 v	79.37 v
Αλγορ.11	52.3	68.52 v	60.32	61.54	67.23 v	73.73 v	79.73 v	80.09 v	80.73 v	80.11 v	80.34 v	80.61 v	80.73 v	79.73 v	80.37 v	79.00 v	76.18 v	79.36 v
Αλγορ.12	49.9	63.66	65.80 v	66.48 v	67.87 v	73.39 v	77.37 v	79.04 v	79.04 v	78.20 v	77.32 v	75.55 v	77.32 v	77.93 v	76.95 v	76.52 v	78.21 v	77.59 v
Αλγορ.13	44.03	60.37	50.21	59.27	56.8	66.89 v	81.50 v	79.50 v	80.88 v	81.00 v	81.25 v	74.98 v	78.16 v	77.50 v	79.21 v	79.82 v	73.68 v	73.68 v
Αλγορ.14	55.97	59.84	60.29	59.45	69.54 v	73.34 v	70.36 v	71.20 v	73.32 v	70.80 v	75.34 v	72.66 v	72.00 v	73.50 v	75.29 v	76.66 v	78.07 v	76.57 v
Αλγορ.15	48.5	64.18	59.7	59.41	66.29 v	71.75 v	79.50 v	77.23 v	80.88 v	76.89 v	79.25 v	80.00 v	79.88 v	77.63 v	79.75 v	79.75 v	80.12 v	79.25 v
Αλγορ.16	46.73	65.11	63	62.75	67.43 v	68.57 v	69.34 v	71.46 v	69.14 v	70.43 v	68.59 v	73.18 v	69.84 v	69.05 v	71.04 v	70.30 v	70.41 v	69.73 v
Αλγορ.17	59.07	56.79	59.79	58.3	74.52 v	70.57 v	73.73 v	73.16 v	71.14 v	73.34 v	75.32 v	74.04 v	72.89 v	76.30 v	76.14 v	76.29 v	77.86 v	74.64 v
Αλγορ.18	59.4	62.68	58.96	64.45	72.37 v	76.62 v	77.62 v	74.00 v	75.48 v	76.20 v	75.59 v	74.77 v	74.29 v	74.50 v	75.25 v	77.18 v	76.20 v	73.09 v
Αλγορ.19	54.3	57.86	55.64	58.12	62.29	64.75	68.61 v	66.75 v	64.64	72.04 v	67.52 v	66.73 v	67.05 v	70.91 v	67.27 v	67.66 v	68.50 v	70.11 v
Αλγορ.20	52	64.48	60.64	60.91	68.09	71.14	74.86	74.52	73.75	75.25	73.21	73.73	74.18	73.87	73.3	73.16	74.79	73.02

Στις γραμμές του Πίνακα 37 εμφανίζονται οι τιμές της ακρίβειας πρόβλεψης, η οποία ορίζεται ως ο λόγος του πλήθους των ορθώς ταξινομημένων οντοτήτων προς το συνολικό πλήθος των οντοτήτων (Witten, Frank & Hall, 2011, σ. 72) (βλ. Κεφ. 9.6.2). Στη δεύτερη γραμμή εμφανίζονται οι τιμές για την *τιμή βάσης* της ακρίβειας πρόβλεψης (baseline accuracy), την οποία επιτυγχάνει το μοντέλο ZeroR προβλέποντας «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες (εν προκειμένω, οι φοιτητές) ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Οπότε, με την ακρίβεια πρόβλεψης του ZeroR μοντέλου συγκρίνονται οι τιμές των μοντέλων που παράγονται με τη χρήση των άλλων αλγόριθμων ταξινόμησης (βλ. Κεφ. 9.6.1 Πίνακας 21). Η τιμή βάσης, σε αντίθεση με ότι συμβαίνει στα μαθήματα AA005, παραμένει σταθερή από την 3<sup>η</sup> εβδομάδα ως την ολοκλήρωση του μαθήματος, διότι η παρακολούθηση στα μαθήματα A3 είναι υποχρεωτική για όλες τις εβδομάδες. Την 1<sup>η</sup> εβδομάδα έχουν ήδη συνδεθεί 55 από τους 78 φοιτητές, την 2<sup>η</sup> εβδομάδα έχουν ήδη συνδεθεί 77 από τους 78 φοιτητές και την 3<sup>η</sup> εβδομάδα έχουν συνδεθεί όλοι οι φοιτητές. Οπότε, από την 3<sup>η</sup> εβδομάδα τα σύνολα δεδομένων είναι πλήρη ως προς τους συμμετέχοντες φοιτητές και η τιμή βάσης λαμβάνει την τιμή 53.93%, την οποία διατηρεί ως την ολοκλήρωση του μαθήματος.

Στην 3<sup>η</sup> έως και στην 21<sup>η</sup> γραμμή του Πίνακα 37 εμφανίζονται οι τιμές της ακρίβειας πρόβλεψης που επιτυγχάνουν τα μοντέλα των βασικών βασικών αλγόριθμων του WEKA οι οποίοι εκτελέστηκαν με τις προκαθορισμένες (default) παραμέτρους (βλ. Κεφ. 10.1.1 Πίνακας 23). Ο χρωματικός κώδικας των τιμών του Πίνακα 37 είναι ίδιος με των προηγούμενων αντίστοιχων πινάκων (Πίνακες 22 και 30). Δηλαδή, όσες τιμές υπερβαίνουν την τιμή βάσης (γραμμή 2) έχουν γαλάζιο (σκούρο ή ανοικτό) υπόβαθρο. Αναλυτικότερα, σκούρο γαλάζιο υπόβαθρο έχουν όσες τιμές υπερβαίνουν την τιμή βάσης με στατιστικώς σημαντική διαφορά, με ποσοστό στατιστικής σημαντικότητας 0.05. Σε αυτές τις τιμές, δίπλα στον αριθμό υπάρχει το σύμβολο «v». Οσες τιμές απλώς υπερβαίνουν την τιμή βάσης, χωρίς στατιστικώς σημαντική διαφορά, έχουν ανοικτό γαλάζιο υπόβαθρο. Λευκό υπόβαθρο έχουν οι τιμές που υπολείπονται της τιμής βάσης. Σε όσες από αυτές τιμές, η διαφορά από την τιμή βάσης είναι στατιστικώς σημαντική, δίπλα στον αριθμό υπάρχει το σύμβολο «\*».

### *Γενική τάση των τιμών της ακρίβειας πρόβλεψης*

Στον Πίνακα 37 παρατηρούμε ότι η τιμή της ακρίβειας πρόβλεψης, που επιτυγχάνουν τα μοντέλα, αυξάνεται κατά κανόνα, καθώς κινούμαστε χρονικά από την 1<sup>η</sup> προς την τελευταία (18<sup>η</sup>) εβδομάδα των μαθημάτων. Η αύξηση της προβλεπτικής ικανότητας των μοντέλων οφείλεται στο ότι τα σύνολα δεδομένων των τελευταίων εβδομάδων περιλαμβάνουν περισσότερες πληροφορίες έναντι των συνόλων δεδομένων των πρώτων εβδομάδων. Οπότε, ο εντοπισμός μοτίβων (patterns) και η εκμάθηση (training) των μοντέλων είναι ευκολότερη τις τελευταίες παρά τις πρώτες εβδομάδες. Ωστόσο, η αυξητική τάση δεν είναι συνεχώς αύξουσα, αλλά παρατηρούνται μικρές διακυμάνσεις, δηλαδή χρονικές περίοδοι κατά τις οποίες η τιμή κάμπτεται, επανακάμπτει, κάμπτεται κ.ο.κ. Όμως, ακόμα και τις χρονικές στιγμές που παρατηρείται κάμψη, οι τιμές ακρίβειας πρόβλεψης που επιτυγχάνουν τα μοντέλα, είναι πολύ υψηλότερες από την τιμή βάσης.

### *Χρονική στιγμή πρόβλεψης*

Παρατηρούμε στον Πίνακα 37, ότι την 18<sup>η</sup> εβδομάδα, όταν δηλαδή διαθέτουμε δεδομένα από την έναρξη των μαθημάτων (1<sup>η</sup> εβδομάδα) έως την προθεσμία κατάθεσης της τελικής εργασίας (18<sup>η</sup> εβδομάδα), όλα τα μοντέλα επιτυγχάνουν ακρίβεια πρόβλεψης μεγαλύτερη από την τιμή βάσης. Επιπλέον, σε 15 από τα 19 μοντέλα, η διαφορά της ακρίβειας πρόβλεψης από την τιμή βάσης, είναι στατιστικώς σημαντική, ενώ 11 από τα 19 επιτυγχάνουν αξιοσημείωτη ακρίβεια πρόβλεψης, μεγαλύτερη του 70%, η οποία σε κάποιες περιπτώσεις προσεγγίζει το 80%. Ειδικότερα, η υψηλότερη τιμή ακρίβειας πρόβλεψης είναι 79.37%, διαφέρει 25 εκατοστιαίες μονάδες από την τιμή βάσης και επιτυγχάνεται από το μοντέλο του αλγόριθμου lazy.LWL (A/A 10). Ακολουθούν σε ελάχιστη απόσταση τα μοντέλα του αλγόριθμου rules.DecisionTable (A/A 11) με 79.36%, του αλγόριθμου trees.DecisionStump (A/A 15) με 79.25% και του αλγόριθμου rules.JRip (A/A 12) με 77.59%.

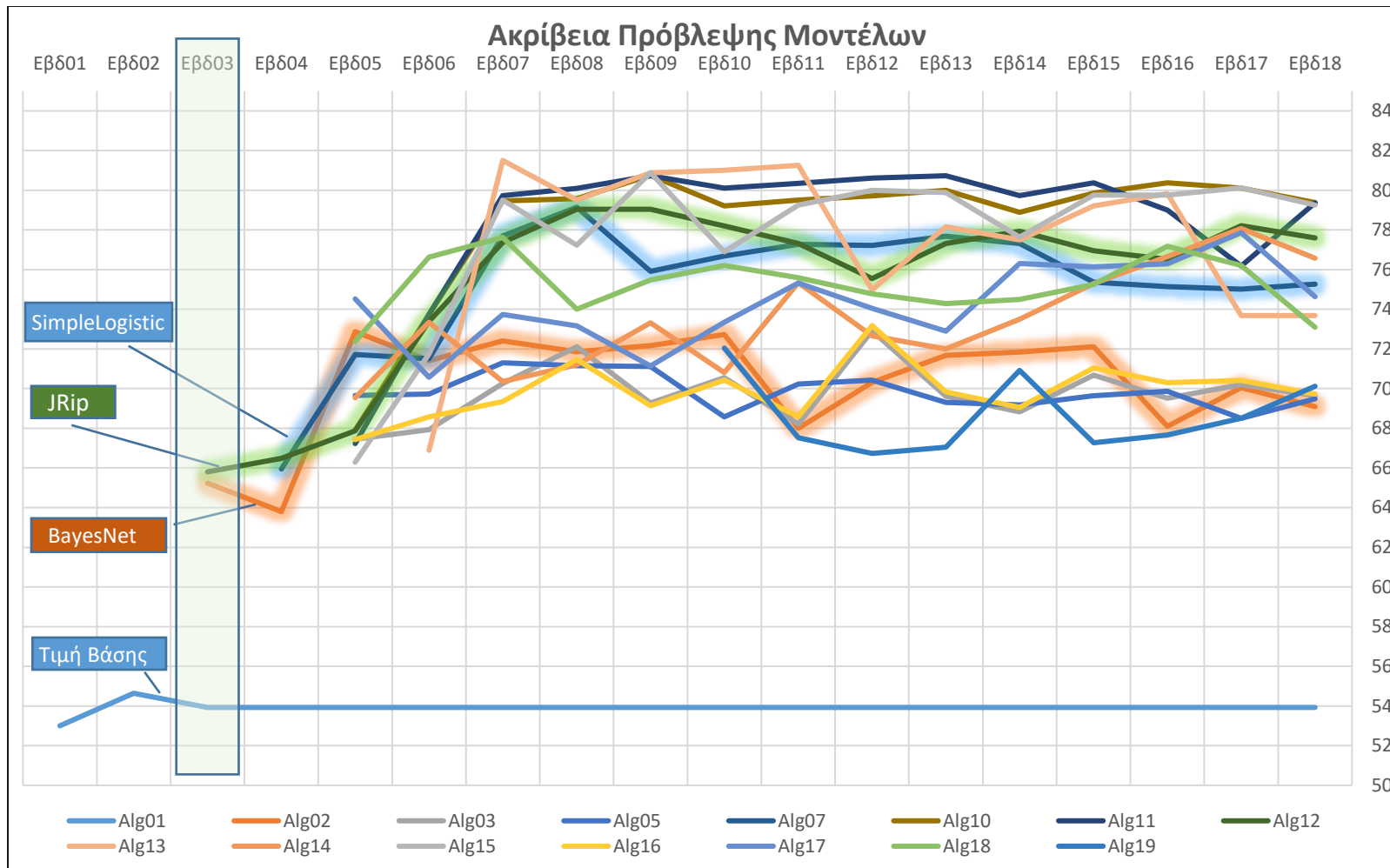
Ως προς τον χρόνο πρόβλεψης και συγκεκριμένα ως προς το πόσο νωρίς επιτυγχάνεται στατιστικώς σημαντική ακρίβεια πρόβλεψης, παρατηρούμε ότι η ακρίβεια πρόβλεψης των περισσότερων μοντέλων και συγκεκριμένα 13 μοντέλων από τα 19, εμφανίζει στατιστική σημαντικότητα πολύ πρώιμα και συγκεκριμένα την 5<sup>η</sup> εβδομάδα, δηλαδή 13 εβδομάδες πριν την ολοκλήρωση του μαθήματος. Κάποια μοντέλα επιτυγχάνουν στατιστικώς σημαντική

ακρίβεια πρόβλεψης ακόμα πιο πρώιμα. Πρόκειται για τα μοντέλα των αλγόριθμων bayes.BayesNet (A/A 2) και rules.JRip (A/A 12) με πρώιμη, στατιστικώς σημαντική πρόβλεψη την 3<sup>η</sup> εβδομάδα και του αλγόριθμου functions.SimpleLogistic (A/A 7) με πρώιμη, στατιστικώς σημαντική πρόβλεψη την 4<sup>η</sup> εβδομάδα. Μάλιστα οι πρώιμες προβλέψεις αυτών των μοντέλων δεν είναι απλώς στατιστικά σημαντικές, αλλά και αρκετά μεγαλύτερες από την τιμή βάσης. Παρατηρούμε ότι το μοντέλο του αλγόριθμου bayes.NaiveBayesMultinomial (A/A 4) επιτυγχάνει στατιστικώς σημαντική τιμή την 3<sup>η</sup> εβδομάδα και το μοντέλο του αλγόριθμου rules.DecisionTable (A/A 11) ακόμα νωρίτερα, την 2<sup>η</sup> εβδομάδα. Ωστόσο η επιτυχία τους δεν είναι αδιάλειπτη ως την τελευταία εβδομάδα, δηλαδή, υπάρχουν εβδομάδες (η 3<sup>η</sup> και η 4<sup>η</sup> για το μοντέλο rules.DecisionTable, η 17<sup>η</sup> για το μοντέλο bayes.NaiveBayesMultinomial), κατά τις οποίες τα μοντέλα δεν επιτυγχάνουν στατιστικώς σημαντικές τιμές. Συνεπώς, δεν μπορούν να αξιοποιηθούν για πρώιμη πρόβλεψη, δεδομένου ότι δεν προσφέρουν συνεχόμενες στατιστικώς σημαντικές προβλέψεις. Αξίζει επίσης να σημειωθεί, ότι όλα ανεξαιρέτως τα μοντέλα υπερβαίνουν την τιμή βάσης από τη 2<sup>η</sup> εβδομάδα. Ωστόσο, η διαφορά από την τιμή βάσης δεν είναι στατιστικώς σημαντική, συνεπώς αυτές οι τιμές δεν τυγχάνουν περαιτέρω προσοχής.

Συγκρίνοντας τα μοντέλα των μαθημάτων A3 με τα μοντέλα των μαθημάτων AA005 ως προς τη χρονική στιγμή πρώιμης πρόβλεψης, σημειώνουμε την εντυπωσιακή υπεροχή των μοντέλων των μαθημάτων A3, τα οποία επιτυγχάνουν στατιστικώς σημαντική ακρίβεια πρόβλεψης σχεδόν στην αρχή του μαθήματος (3<sup>η</sup> εβδομάδα), 4 εβδομάδες νωρίτερα από την αντίστοιχη χρονική στιγμή των μοντέλων του μαθήματος AA005 2020-2021 και 6 εβδομάδες νωρίτερα από μοντέλα του μαθήματος AA005 2021-2022. Είναι προφανές, ότι η κύρια αιτία του φαινομένου είναι η κατά πολύ χαμηλότερη τιμή βάσης στην τιμή πρόβλεψης (του «πήχη») στα μαθήματα A3, απόρροια του γεγονότος, ότι οι κατηγορικές μεταβλητές της κλάσης προέκυψαν με διχοτόμηση επί της διαμέσου των βαθμών που έλαβαν οι φοιτητές. Οπότε, η πολυπληθέστερη κατηγορία (που διαμορφώνει την τιμή βάσης) δεν υπερέβη κατά πολύ το 50% (βλ. Κεφ. 9.5.9). Η χαμηλή τιμή βάσης στην ακρίβεια πρόβλεψης διευκολύνει τα μοντέλα να επιτύχουν στατιστικώς σημαντικές τιμές για την ακρίβεια πρόβλεψης. Συγκεκριμένα, στα μαθήματα A3 η τιμή βάσης την 3<sup>η</sup> εβδομάδα είναι 53.93%, στο μάθημα AA005 2020-2021 είναι 74.52% και στο μάθημα AA005 2021-2022 είναι 80.57%. Συνεπώς, μοντέλο που, για παράδειγμα, επιτυγχάνει απόδοση

65% κατά την 3<sup>η</sup> εβδομάδα, θεωρείται ότι υπερβαίνει την τιμή βάσης με στατιστικώς σημαντική διαφορά στα μαθήματα A3, ενώ αντιθέτως θεωρείται ότι υπολείπεται της βάσης με στατιστικώς σημαντική διαφορά στα μαθήματα AA005. Οπότε, θεωρούμε ότι αν και στα μαθήματα AA005 οι κατηγορικές μεταβλητές της κλάσης προέκυπταν με διχοτόμηση επί της διαμέσου των βαθμών, τότε πιθανότατα θα υπήρχαν στατιστικώς σημαντικές προβλέψεις το ίδιο πρώιμα όπως και στα μαθήματα A3.

Η εξέλιξη των τιμών της ακρίβειας πρόβλεψης των μοντέλων απεικονίζεται παραστατικότερα στο Γράφημα 9 που ακολουθεί.



Γράφημα 9. Ακρίβεια πρόβλεψης μοντέλων για το μάθημα ΑΑ005 2021-2022

Στο Γράφημα 9 απεικονίζονται μόνο οι τιμές, που η διαφορά τους από την τιμή βάσης είναι στατιστικώς σημαντική. Δηλαδή απεικονίζονται οι τιμές οι οποίες έχουν σκούρο γαλάζιο υπόβαθρο στον Πίνακα 37. Στον οριζόντιο άξονα του γραφήματος απεικονίζεται η χρονική περίοδος από την έναρξη του μαθήματος, η οποία προσδιορίζεται με τον αντίστοιχο αριθμό της εβδομάδας (π.χ. Εβδ10: 10<sup>η</sup> εβδομάδα από την έναρξη του μαθήματος). Στον κάθετο άξονα του γραφήματος απεικονίζεται αριθμητική κλίμακα που αντιστοιχεί στο αριθμητικό εύρος των τιμών της ακρίβειας πρόβλεψης. Οι αριθμητικές τιμές αντιστοιχούν σε ποσοστά επί τοις εκατό (%). Κάθε μία έγχρωμη γραμμή που διατρέχει το γράφημα αντιστοιχεί στις τιμές μοντέλου το οποίο παράγεται από αλγόριθμο του Πίνακα 23. Δεν σχηματίζονται γραμμές για μία, ή περισσότερες τιμές, που είναι «απομονωμένες», δηλαδή για τιμή, ή για τιμές, για τις οποίες υφίσταται μεταγενέστερη μη στατιστικώς σημαντική τιμή. Κατά συνέπεια, δεν υπάρχουν γραμμές για το μοντέλο του αλγόριθμου με A/A 20, διότι ποτέ δεν πέτυχε ακρίβεια πρόβλεψης στατιστικά υψηλότερη από την τιμή βάσης. Επίσης, δεν υπάρχουν γραμμές για τα μοντέλα των αλγόριθμων με A/A 4, με A/A 6, με A/A 8 και με A/A 9, διότι όλες οι στατιστικά σημαντικές τιμές τους είναι «απομονωμένες». Για το μοντέλο του αλγόριθμου με A/A 19 αποτυπώνονται μόνο όσες στατιστικά σημαντικές τιμές δεν είναι «απομονωμένες». Η γαλάζια γραμμή στο κάτω μέρος του γραφήματος αντιστοιχεί στις τιμές του μοντέλου ZeroR (A/A 1), δηλαδή ορίζει τον «πήχη», που όλα τα μοντέλα πρέπει να υπερβούν με στατιστικώς σημαντική διαφορά. Με την εξαίρεση του μοντέλου του αλγόριθμου ZeroR (τιμή βάσης) δεν σχηματίζονται γραμμές από την 1η εβδομάδα έως και την 2η εβδομάδα, διότι κατά την συγκεκριμένη χρονική περίοδο κανένα μοντέλο δεν επιτυγχάνει ακρίβεια πρόβλεψης στατιστικώς σημαντική σε σύγκριση με την τιμή βάσης.

Η πρωιμότερη, στατιστικώς σημαντική διαφορά από την τιμή βάσης επιτυγχάνεται την 3<sup>η</sup> εβδομάδα από τα μοντέλα των αλγόριθμων bayes.BayesNet (A/A 2) και rules.JRip (A/A 12). Την επόμενη (4<sup>η</sup>) εβδομάδα και το μοντέλο του αλγόριθμου functions.SimpleLogistic (A/A 7) επιτυγχάνει στατιστικώς σημαντική ακρίβεια πρόβλεψης, η οποία διατηρείται ως την ολοκλήρωση του μαθήματος. Για λόγους ευκρίνειας, οι γραμμές αυτών των τριών μοντέλων περιβάλλονται από ομοιόχρωμη αχλή. Όπως έχει ήδη αναφερθεί, το μοντέλο του αλγόριθμου bayes.NaiveBayesMultinomial (A/A 4) επιτυγχάνει στατιστικώς σημαντική τιμή την 3<sup>η</sup> εβδομάδα και το μοντέλο του αλγόριθμου rules.DecisionTable ακόμα νωρίτερα (την 2<sup>η</sup> εβδομάδα).

Ωστόσο, η επιτυχία τους δεν είναι αδιάλειπτη ως την τελευταία εβδομάδα, οπότε δεν σχηματίζεται γραμμή για αυτά τα μοντέλα. Η ανάδειξη στατιστικώς σημαντικών τιμών από τα μοντέλα bayes.BayesNet (A/A 2) και rules.JRip (A/A 12) κατά την 3<sup>η</sup> εβδομάδα διευκολύνθηκε από την μείωση της τιμής βάσης, μείωση που είναι εμφανής στο γράφημα, μέσω της βύθισης της γραμμής του μοντέλου του αλγόριθμου ZeroR.

Στο Γράφημα 9 είναι εμφανές ότι η ακρίβεια πρόβλεψης των περισσότερων μοντέλων αυξάνει αλματωδώς έως την 7<sup>η</sup> εβδομάδα. Κατόπιν, ο ρυθμός αύξησης χάνει την ορμή του και η ακρίβεια πρόβλεψης κατά κανόνα σταθεροποιείται, παρουσιάζοντας μικρές μόνο διακυμάνσεις εντός μικρού εύρους τιμών. Το φαινόμενο εμφανίζεται, με μικρές διαφοροποιήσεις, σε όλα σχεδόν τα μοντέλα. Οπότε, εικάζουμε ότι δεν είναι τυχαίο και ως εξήγηση προτείνουμε το εξής: Σύμφωνα με τον σχεδιασμό των μαθημάτων, κάθε συνεδρία από την 1<sup>η</sup> έως την 7<sup>η</sup> περιλαμβάνει δραστηριότητα δημόσιου διαλόγου στο ΣΔΜ, όπου οι φοιτητές συμμετέχουν υποχρεωτικά αναρτώντας σχόλια. Επιπλέον, η συμμετοχή των φοιτητών σε κάθε δημόσιο διάλογο βαθμολογείται από τον διδάσκοντα του μαθήματος. Ο μέσος όρος των επτά βαθμών, που λαμβάνει κάθε φοιτητής προσμετράται κατά 30% στον τελικό βαθμό. Τα σύνολα δεδομένων του μαθήματος περιλαμβάνουν τα χαρακτηριστικά «Πλήθος συμβάντων πρόσβασης στους χώρους δημόσιου διαλόγου» (A/A 25) και «Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου» (A/A 26), τα οποία καταγράφουν δραστηριότητα σχετική με τους χώρους δημόσιου διαλόγου. Επίσης, περιλαμβάνουν το χαρακτηριστικό «Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο» (A/A 27), το οποίο αποτυπώνει τον αντίστοιχο βαθμό του φοιτητή. Με βάση τα παραπάνω, θεωρούμε ότι οι φοιτητές χρησιμοποιούν το ΣΔΜ περισσότερο εντατικά και συστηματικά τις πρώτες 7 εβδομάδες. Κατά συνέπεια, τα σύνολα δεδομένων των 7 πρώτων εβδομάδων διαθέτουν πληροφορία καλύτερης «ποιότητας» έναντι των επόμενων εβδομάδων. Επιπλέον, δεδομένου ότι την 7<sup>η</sup> εβδομάδα, ολοκληρώνεται η βαθμολόγηση των δραστηριοτήτων των δημόσιων διαλόγων, το αντίστοιχο σύνολο δεδομένων διαθέτει την πληροφορία που προδιαγράφει (έστω κατά 30%) τον τελικό βαθμό του φοιτητή. Συνεπώς, ήδη από την 7<sup>η</sup> εβδομάδα, τα μοντέλα έχουν στη διάθεσή τους σύνολα δεδομένων, που αφενός περιέχουν αποκαλυπτικά μοτίβα (patterns) χρήσης και αφετέρου βαρύνουσα ένδειξη (30%) για την τελική επίδοση των φοιτητών. Τα σύνολα δεδομένων των επόμενων εβδομάδων



προσφέρουν στα μοντέλα λίγες μόνο πρόσθετες χρήσιμες πληροφορίες. Οπότε, η ακρίβεια πρόβλεψης των μοντέλων φτάνει σε υψηλό επίπεδο την 7<sup>η</sup> εβδομάδα. Έκτοτε, όταν αφενός η χρήση του ΣΔΜ από τους φοιτητές ατονεί και αφετέρου παύει η βαθμολόγηση των φοιτητών, η ακρίβεια πρόβλεψης παρουσιάζει μόνο μικρές διακυμάνσεις. Σε κάποιες μάλιστα περιπτώσεις είναι εμφανής κάποια κάμψη, που πιθανόν οφείλεται στο ότι μετά το τέλος των διαλέξεων (13<sup>η</sup> εβδομάδα) οι φοιτητές συνδέονται στο ΣΔΜ σποραδικά.

### 10.3.2. Επιπλέον κριτήρια για την απόδοση των βασικών μοντέλων

Προκειμένου να αποτιμήσουμε την απόδοση των μοντέλων υπολογίσαμε εκτός από τις τιμές της ακρίβειας πρόβλεψης και τις τιμές της μεταβλητής AUC (Area Under Curve) (Witten, Frank & Hall, 2011, σ. 172) και της μεταβλητής Cohen's Kappa (Witten, Frank, & Hall, 2011, σ. 166).

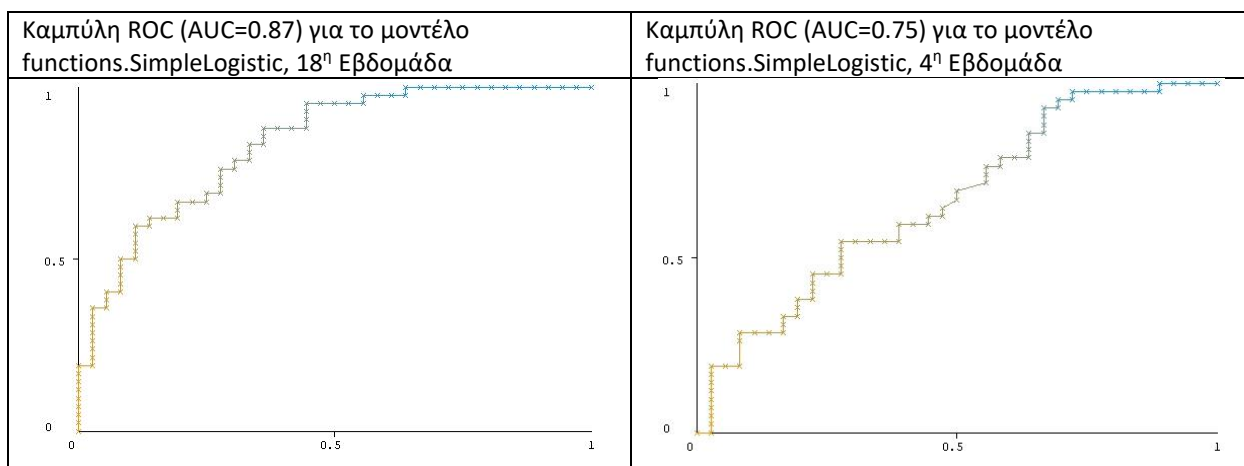
Όπως έχουμε αναφέρει στο Κεφ. 9.6.3, η μεταβλητή AUC αξιολογεί τη διαγνωστική ικανότητα ενός μοντέλου ταξινόμησης με διχοτομημένη κλάση και μη ισορροπημένο (Hussain M. , Zhu, Zhang, & Abidi, 2018) σύνολο δεδομένων και υπολογίζει την περιοχή που ορίζει η καμπύλη ROC (Receiver Operating Characteristic) (Σχήμα 13).

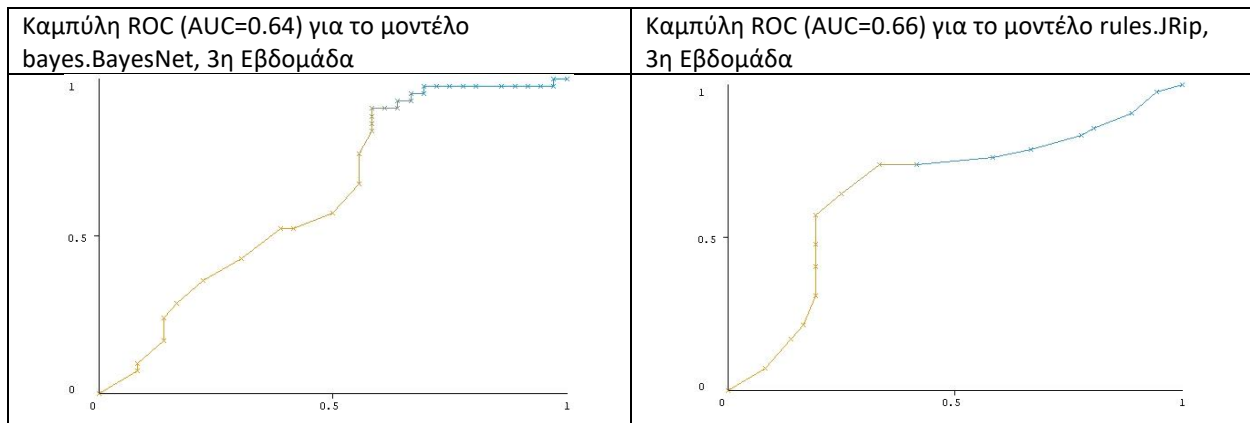
**Πίνακας 38. Τιμές μεταβλητής AUC για τα μοντέλα των μαθημάτων A3 2019-2020 & 2020-2021**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
Αλγορ.01	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Αλγορ.02	0.5	0.67	0.64	0.63	0.75	0.77	0.82	0.82	0.83	0.82	0.8	0.8	0.8	0.82	0.84	0.8	0.81	0.8
Αλγορ.03	0.49	0.68	0.64	0.63	0.68	0.72	0.71	0.73	0.71	0.73	0.73	0.74	0.75	0.73	0.73	0.74	0.73	0.74
Αλγορ.04	0.42	0.67	0.64	0.64	0.66	0.68	0.69	0.73	0.72	0.71	0.72	0.73	0.73	0.74	0.74	0.74	0.72	0.75
Αλγορ.05	0.59	0.67	0.61	0.68	0.74	0.76	0.78	0.74	0.73	0.76	0.77	0.78	0.75	0.74	0.75	0.75	0.77	0.75
Αλγορ.06	0.58	0.62	0.58	0.66	0.66	0.71	0.66	0.68	0.64	0.71	0.72	0.71	0.7	0.69	0.72	0.72	0.71	0.75
Αλγορ.07	0.51	0.68	0.74	0.75	0.79	0.84	0.86	0.86	0.87	0.87	0.87	0.87	0.88	0.88	0.88	0.87	0.87	0.87
Αλγορ.08	0.55	0.62	0.55	0.6	0.61	0.57	0.5	0.57	0.5	0.53	0.49	0.61	0.51	0.57	0.51	0.55	0.58	0.55
Αλγορ.09	0.56	0.58	0.51	0.65	0.62	0.68	0.67	0.73	0.72	0.69	0.66	0.72	0.7	0.71	0.68	0.63	0.64	0.62
Αλγορ.10	0.54	0.66	0.65	0.66	0.72	0.77	0.84	0.84	0.84	0.84	0.82	0.83	0.85	0.84	0.85	0.84	0.84	0.84
Αλγορ.11	0.49	0.67	0.59	0.63	0.71	0.76	0.81	0.81	0.82	0.83	0.81	0.82	0.83	0.84	0.83	0.82	0.81	0.83
Αλγορ.12	0.49	0.63	0.66	0.66	0.69	0.73	0.78	0.79	0.78	0.78	0.78	0.76	0.78	0.78	0.77	0.78	0.79	0.8
Αλγορ.13	0.44	0.6	0.5	0.59	0.56	0.66	0.81	0.79	0.8	0.81	0.81	0.75	0.78	0.77	0.79	0.79	0.73	0.73
Αλγορ.14	0.55	0.58	0.61	0.6	0.71	0.72	0.71	0.72	0.73	0.71	0.76	0.74	0.72	0.75	0.75	0.76	0.78	0.76
Αλγορ.15	0.47	0.62	0.57	0.58	0.65	0.71	0.79	0.77	0.8	0.76	0.79	0.79	0.79	0.77	0.79	0.79	0.8	0.79
Αλγορ.16	0.49	0.69	0.64	0.63	0.7	0.71	0.71	0.73	0.72	0.74	0.73	0.75	0.76	0.73	0.74	0.75	0.74	0.74
Αλγορ.17	0.59	0.55	0.58	0.59	0.74	0.69	0.74	0.75	0.74	0.74	0.76	0.75	0.73	0.78	0.78	0.76	0.79	0.74
Αλγορ.18	0.65	0.66	0.61	0.7	0.78	0.81	0.82	0.82	0.81	0.83	0.83	0.82	0.81	0.84	0.82	0.82	0.81	0.85
Αλγορ.19	0.55	0.58	0.55	0.58	0.62	0.65	0.69	0.67	0.65	0.72	0.67	0.67	0.67	0.71	0.67	0.68	0.68	0.7
Αλγορ.20	0.5	0.61	0.59	0.6	0.67	0.72	0.75	0.75	0.75	0.76	0.74	0.75	0.75	0.75	0.75	0.75	0.75	0.74

Στον Πίνακα 38, καταρχάς παρατηρούμε ότι η τιμή AUC για το μοντέλο του αλγόριθμου ZeroR (A/A 1) ισούται με 0.5, γεγονός απολύτως αναμενόμενο, διότι όπως έχουμε ήδη αναφέρει (βλ Κεφ. 9.6.2) το συγκεκριμένο μοντέλο προβλέπει «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Συνεπώς, η τιμή AUC δεν έχει νόημα στο συγκεκριμένο μοντέλο, το οποίο χρησιμοποιείται για τον υπολογισμό της τιμής βάσης. Όλα τα άλλα μοντέλα, σε όλες τις χρονικές περιόδους (με την εξαίρεση λίγων τιμών την πρώτη εβδομάδα), σημειώνουν τιμές που είναι μεγαλύτερες από το 0.5, συνεπώς κρίνονται αποδεκτές (Riestra-Gonzalez, del Puerto Paule-Ruiz, & Ortin, 2021). Τη 18<sup>η</sup> εβδομάδα, με την ολοκλήρωση των μαθημάτων, οι τιμές AUC όλων των μοντέλων είναι μεγαλύτερες από 0.7, οπότε κρίνονται καλές. Υπάρχει μάλιστα μοντέλο, αυτό του αλγόριθμου functions.SimpleLogistic (A/A 7) που επιτυγχάνει τιμή AUC 0.87, η οποία κρίνεται εξαιρετική. Επιπλέον, παρατηρούμε ότι, την 3<sup>η</sup> εβδομάδα, κατά την οποία επιτυγχάνεται η πρωιμότερη πρόβλεψη, οι τιμές AUC για τα μοντέλα bayes.BayesNet (A/A 2) και rules.JRip (A/A 12) είναι αντιστοίχως 0.64 και 0.66, οπότε κρίνονται ικανοποιητικές και την 4<sup>η</sup> εβδομάδα, η τιμή AUC του μοντέλου functions.SimpleLogistic (A/A 7) είναι 0.75, οπότε κρίνεται καλή.

Στο Γράφημα 10 που ακολουθεί αναπαρίστανται οι καμπύλες ROC που υπολογίζουν ενδεικτικές τιμές AUC και συγκεκριμένα τις τιμές AUC του μοντέλου functions.SimpleLogistic (A/A 7) για την 18<sup>η</sup> και την 4<sup>η</sup> εβδομάδα και τις τιμές AUC των μοντέλων bayes.BayesNet (A/A 2) και rules.JRip (A/A 12) για την 3<sup>η</sup> εβδομάδα.





**Γράφημα 10. Καμπύλες ROC για τα μοντέλα των μαθημάτων A3 2019-2020 & 2020-2021**

Όπως έχουμε αναφέρει στο Κεφ. 9.6.3, η μεταβλητή Cohen’s Kappa μετρά τη συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που προέβλεψε το μοντέλο, ενώ διορθώνει τυχόν συμφωνία που προέκυψε τυχαίως (Witten, Frank, & Hall, 2011, σ. 166).

**Πίνακας 39. Τιμές μεταβλητής Cohen’s Kappa για τα μοντέλα των μαθημάτων A3 2019-2020 & 2021-2022**

Εβδομ	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
Αλγορ.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Αλγορ.02	0	0.29	0.28	0.25	0.44	0.4	0.43	0.42	0.42	0.43	0.34	0.38	0.41	0.41	0.42	0.34	0.37	0.36
Αλγορ.03	-0.08	0.25	0.24	0.22	0.33	0.34	0.39	0.43	0.37	0.4	0.35	0.45	0.38	0.36	0.4	0.37	0.39	0.38
Αλγορ.04	-0.01	0.19	0.28	0.29	0.36	0.39	0.43	0.39	0.38	0.36	0.39	0.36	0.4	0.41	0.36	0.37	0.24	0.32
Αλγορ.05	0.08	0.28	0.18	0.3	0.39	0.4	0.43	0.42	0.42	0.37	0.4	0.41	0.38	0.38	0.39	0.39	0.36	0.39
Αλγορ.06	0.11	0.11	0.11	0.22	0.21	0.3	0.19	0.2	0.2	0.27	0.25	0.27	0.2	0.26	0.25	0.29	0.24	0.3
Αλγορ.07	-0.05	0.23	0.26	0.3	0.41	0.42	0.54	0.57	0.51	0.53	0.54	0.54	0.54	0.54	0.5	0.49	0.49	0.49
Αλγορ.08	0.1	0.25	0.1	0.2	0.23	0.14	0	0.15	0	0.05	-0.02	0.21	0.01	0.15	0.02	0.11	0.17	0.09
Αλγορ.09	0.09	0.14	-0.01	0.14	0.2	0.29	0.19	0.34	0.28	0.31	0.17	0.28	0.3	0.27	0.29	0.14	0.21	0.16
Αλγορ.10	0.05	0.26	0.22	0.22	0.35	0.46	0.58	0.58	0.61	0.58	0.58	0.59	0.59	0.57	0.59	0.6	0.59	0.58
Αλγορ.11	-0.01	0.34	0.19	0.22	0.34	0.46	0.59	0.6	0.61	0.59	0.6	0.61	0.61	0.58	0.6	0.57	0.52	0.58
Αλγορ.12	-0.02	0.25	0.31	0.32	0.35	0.46	0.54	0.58	0.58	0.56	0.54	0.51	0.54	0.55	0.53	0.52	0.56	0.55
Αλγορ.13	-0.13	0.2	0	0.17	0.12	0.32	0.62	0.58	0.61	0.61	0.62	0.49	0.56	0.54	0.58	0.59	0.46	0.47
Αλγορ.14	0.12	0.18	0.2	0.18	0.39	0.46	0.41	0.42	0.46	0.41	0.5	0.45	0.43	0.46	0.5	0.52	0.56	0.53
Αλγορ.15	-0.05	0.24	0.15	0.17	0.3	0.42	0.58	0.53	0.61	0.53	0.58	0.59	0.59	0.54	0.58	0.58	0.59	0.57
Αλγορ.16	-0.08	0.29	0.24	0.23	0.33	0.35	0.37	0.41	0.37	0.4	0.36	0.45	0.38	0.37	0.41	0.39	0.39	0.38
Αλγορ.17	0.19	0.13	0.19	0.16	0.49	0.41	0.47	0.46	0.42	0.46	0.5	0.47	0.45	0.52	0.51	0.52	0.55	0.49
Αλγορ.18	0.19	0.23	0.16	0.27	0.43	0.52	0.55	0.47	0.5	0.51	0.5	0.49	0.48	0.47	0.5	0.54	0.52	0.45
Αλγορ.19	0.09	0.15	0.11	0.16	0.24	0.29	0.37	0.33	0.29	0.44	0.34	0.33	0.34	0.42	0.34	0.35	0.37	0.4
Αλγορ.20	0.01	0.25	0.19	0.19	0.35	0.42	0.49	0.48	0.47	0.5	0.46	0.46	0.48	0.47	0.46	0.46	0.49	0.45

Στον Πίνακα 39 καταρχάς παρατηρούμε, ότι για το μοντέλο του αλγόριθμου ZeroR (A/A 1) η τιμή της μεταβλητής Cohen's Kappa ισούται με μηδέν, διότι όπως έχουμε ήδη αναφέρει (βλ Κεφ. 9.6.2) το συγκεκριμένο μοντέλο προβλέπει «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Παρατηρούμε επίσης, ότι κατά την 18<sup>η</sup> εβδομάδα, η μεταβλητή Cohen's Kappa λαμβάνει τιμές από το διάστημα [0.3 , 0.58], γεγονός που σύμφωνα με τους Quinn & Gray (2020, σ. 8) δηλώνει «μέση» (moderate) και «σημαντική» (substantial) συμφωνία. Εξάιρεση αποτελούν τα μοντέλα με A/A 8 και A/A 9, όπου οι τιμές είναι 0.09 και 0.16 αντιστοίχως και δηλώνουν «αδύναμη» (slight) συμφωνία. Την 7<sup>η</sup> εβδομάδα, όταν ολοκληρώνονται οι δημόσιοι διάλογοι και τα μοντέλα επιτυγχάνουν κατά κανόνα υψηλή ακριβεια πρόβλεψης, η τιμή της μεταβλητής Cohen's Kappa λαμβάνει τιμές, που κατά κανόνα δηλώνουν «μέση» (moderate) και «σημαντική» (substantial) συμφωνία. Την 4<sup>η</sup> εβδομάδα, δηλαδή τη χρονική στιγμή της πρώιμης πρόβλεψης από το μοντέλο functions.SimpleLogistic, η αντίστοιχη τιμή της μεταβλητής Cohen's Kappa είναι 0.3 η οποία δηλώνει «μέση» (moderate) συμφωνία και την 3<sup>η</sup> εβδομάδα, δηλαδή τη χρονική στιγμή της πρώιμης πρόβλεψης από τα μοντέλα Bayes.BayesNet και rules.Jrip, οι αντίστοιχες τιμές της μεταβλητής Cohen's Kappa είναι αντιστοίχως 0.28 και 0.31 οι οποίες δηλώνουν «δίκαιη» (fair) και «μέση» (moderate) συμφωνία. Συνεπώς, σύμφωνα με το κριτήριο της μεταβλητής Cohen's Kappa, κατά τη χρονική στιγμή ολοκλήρωσης του μαθήματος, υπάρχει «μέση» και «σημαντική» συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που προέβλεψαν τα μοντέλα, ενώ η ισχύς της συμφωνίας εξασθενεί τις χρονικές περιόδους που επιτυγχάνεται πρώιμη πρόβλεψη.

### 10.3.3. Ακρίβεια και χρόνος πρόβλεψης του LSTM μοντέλου

Εκτός από τη μελέτη των μοντέλων τα οποία κατασκευάζονται από τους βασικούς αλγόριθμους (αυτών του Πίνακα 23), μελετήσαμε επίσης την απόδοση μοντέλων τα οποία παράγονται, όχι μόνο με τη βοήθεια βασικών αλγόριθμων, αλλά και αλγόριθμων που βασίζονται σε περιοδικά νευρωνικά δίκτυα (Recursive Neural Networks). Αυτοί οι αλγόριθμοι αναλύουν, όχι μόνο τις συνολικές, τελικές τιμές των χαρακτηριστικών του συνόλου δεδομένων, αλλά και τις χρονικές ακολουθίες, οι οποίες αποτυπώνουν την εξέλιξη των τιμών. Η μελέτη τέτοιων τεχνικών επί εκπαιδευτικών δεδομένων αποτελεί πεδίο αυξανόμενου ερευνητικού ενδιαφέροντος (Doleck, Lemay, Basnet & Bazalais, 2020). Επιπλέον, οι ερευνητικές προσπάθειες, που διερευνούν την προβλεπτική ικανότητα τέτοιων τεχνικών, είναι σχετικά λίγες και δεν συγκλίνουν προς κοινώς αποδεκτά συμπεράσματα (Chen & Cui, 2020. Doleck, Lemay, Basnet & Bazalais, 2020). Επιλέξαμε την LSTM υλοποίηση των περιοδικών νευρωνικών δικτύων. Καταρχάς χρησιμοποιήσαμε τις προκαθορισμένες (default) παραμέτρους του αλγόριθμου, με τις οποίες προέκυψαν οι τιμές που καταγράφονται στον Πίνακα 40.

**Πίνακας 40. Ακρίβεια πρόβλεψης LSTM μοντέλου με προκαθορισμένες παραμέτρους για τα μαθήματα A3 2019-2020 & 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Εβδομ	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
ZeroR	53	54.64	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93
LSTM	53	54.64	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93

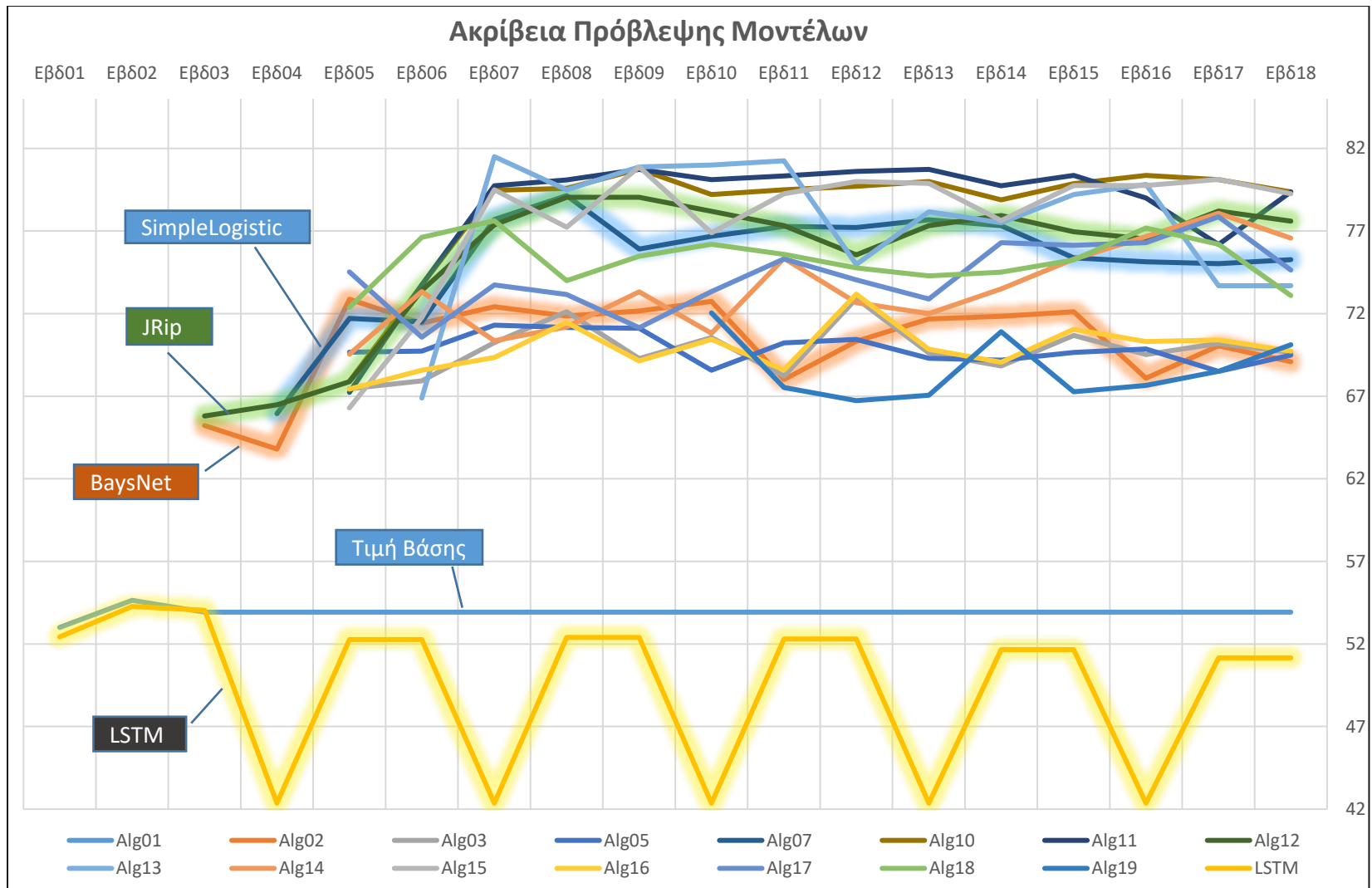
Από τις τιμές του Πίνακα 40 προκύπτει ότι η ακρίβεια πρόβλεψης του LSTM μοντέλου δεν ξεπερνά την τιμή βάσης, δηλαδή το μοντέλο δεν δίνει γόνιμα αποτελέσματα. Προκειμένου να επιτύχουμε τιμές που θα υπερβαίνουν την τιμή βάσης, χρησιμοποιήσαμε τις «βελτιστοποιημένες» παραμέτρους του αλγόριθμου, ο οποίος κατασκεύασε το μοντέλο του μαθήματος AA005 2020-2021 (βλ. Παράρτημα Β). Όπως έχουμε αναφέρει (βλ. Κεφ. 10.1.3), οι βελτιστοποιημένες παράμετροι προέκυψαν με εφαρμογή εμπειρικών κανόνων, που υποδεικνύονται από τη σχετική βιβλιογραφία, σύμφωνα με την οποία η εύρεση ενός αποδοτικού μοντέλου αποτελεί «άσκηση δοκιμής και λάθους» (an exercise in trial-and-error) (Frank, 2021). Οι νέες τιμές καταγράφονται στον Πίνακα 41.

**Πίνακας 41. Ακρίβεια πρόβλεψης LSTM μοντέλου με «βελτιστοποίηση» ως προς τις παραμέτρους για τα μαθήματα A3 2019-2020 & 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Εβδομ	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
ZeroR	53	54.64	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93	53.93
LSTM	52.43	54.27	54.04	42.36	52.27	52.27	42.36	52.39	52.39	42.36	52.3	52.3	42.36	51.66	51.66	42.36	51.16	51.16

Στον Πίνακα 41 παρατηρούμε, ότι σε όλες τις χρονικές στιγμές το LSTM μοντέλο, όχι μόνο δεν επιτυγχάνει στατιστικώς σημαντική τιμή, αλλά ούτε καν καταφέρνει να υπερβεί την τιμή βάσης. Το γεγονός δεν αποτελεί απόδειξη, ότι για τα σύνολα δεδομένων των μαθημάτων A3 2019-2020 & 2020-2021 δεν μπορούν να κατασκευαστούν LSTM μοντέλα, που να επιτυγχάνουν καλές αποδόσεις. Ενδεχομένως, να υπάρχει παραμετροποίηση η οποία να αποδειχθεί επιτυχής. Ωστόσο, εμείς, παρότι προσπαθήσαμε, ακολουθώντας τους εμπειρικούς κανόνες που αποδείχθηκαν αποτελεσματικοί στο μάθημα AA005 2020-2021, δεν καταφέραμε να έχουμε γόνιμο αποτέλεσμα και στα μαθήματα A3 2019-2020 & 2020-2021.

Η αδυναμία του LSTM μοντέλου να επιτύχει τιμές μεγαλύτερες από την τιμή βάσης αποτυπώνεται επίσης στο Γράφημα 11.



Γράφημα 11. Ακρίβεια Πρόβλεψης LSTM Μοντέλου

#### 10.3.4. Επιπλέον κριτήρια για την απόδοση του LSTM μοντέλου

Παρά τα άγωνα αποτελέσματα ως προς την ακρίβεια πρόβλεψης του LSTM μοντέλου, υπολογίσαμε και τις τιμές των μεταβλητών AUC (Area Under Curve) και Cohen's Kappa. Στους Πίνακες 42 και 43 που ακολουθούν, φαίνεται ότι και για αυτές τις μεταβλητές, το LSTM μοντέλο δεν παράγει γόνιμα αποτελέσματα. Ειδικότερα η μεταβλητή AUC δεν υπερβαίνει την τιμή βάσης, ενώ η μεταβλητή Cohen's Kappa λαμβάνει σχεδόν μηδενικές τιμές.

**Πίνακας 42. Τιμές μεταβλητής AUC του LSTM μοντέλου για τα μαθήματα A3 2019-2020 & 2020-2021**

Εβδομ	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
ZeroR	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
LSTM	0.51	0.48	0.48	0.42	0.47	0.47	0.42	0.47	0.47	0.42	0.48	0.48	0.42	0.46	0.46	0.42	0.46	0.46

**Πίνακας 43. Τιμές μεταβλητής Cohen's Kappa του LSTM μοντέλου για τα μαθήματα A3 2019-2020 & 2020-2021**

Εβδομ	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
ZeroR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LSTM	0.05	0.01	0.02	0.11	0.01	0.01	0.11	0.01	0.01	0.11	0.01	0.01	0.11	0.02	0.02	0.11	0.03	0.03



#### 10.4. Έλεγχος γενίκευσης

Η κατασκευή προβλεπτικών μοντέλων που περιγράφεται στις προηγούμενες παραγράφους του Κεφαλαίου 10, βασίστηκε σε δεδομένα εκμάθησης (training data) και δεδομένα ελέγχου (test data) που προέρχονται από το ίδιο σύνολο δεδομένων, δηλαδή από το ίδιο μάθημα. Κάθε μοντέλο συγκρίθηκε, ως προς την απόδοση και την αξιοπιστία του, με τα υπόλοιπα μοντέλα του ίδιου μαθήματος. Ωστόσο η απόδοση, η αξιοπιστία και κυρίως η χρηστική αξία ενός μοντέλου εξαρτάται και από την ικανότητά του για γενίκευση, δηλαδή την ικανότητα του να επιτυγχάνει πρόβλεψη, όχι μόνο όταν εφαρμόζεται στο «οικείο» σύνολο δεδομένων, αλλά και όταν δοκιμάζεται σε «μη οικείο» σύνολο δεδομένων (Gasevic, Dawson, Rogers & Gasevic, 2016. López-Zambrano, Lara & Romero, 2020), δηλαδή σε σύνολο δεδομένων επί του οποίου δεν έχει προηγηθεί φάση «εκμάθησης».

Για να ελέγξουμε την ικανότητα γενίκευσης των μοντέλων που προέκυψαν από τα μαθήματα που μελετήσαμε, εφαρμόσαμε το πιο «επιτυχημένο» από τα παραγόμενα μοντέλα επί συνόλου δεδομένων άλλου μαθήματος και αναλύσαμε την απόδοσή του, αφενός ως προς την τιμή ακρίβειας πρόβλεψης και αφετέρου ως προς τη χρονική στιγμή της πρόβλεψης. Δεδομένου, ότι τα σύνολα δεδομένων των μαθημάτων AA005 2020-2021 και AA005 2021-2022 αφενός και A3 2019-2020 & 2020-2021 αφετέρου, είναι ασύμβατα, παράγουν ασύμβατα μοντέλα. Η συμβατότητα θα μπορούσε να επιτευχθεί, αν από όλα τα σύνολα δεδομένων αφαιρούνταν τα μη κοινά χαρακτηριστικά. Τότε όμως, τα μοντέλα που θα προέκυπταν θα ήταν εντελώς διαφορετικά. Πιθανότατα θα είχαν και ασθενέστερη προβλεπτική ικανότητα, εξαιτίας της απουσίας σημαντικών χαρακτηριστικών. Οπότε, δεδομένης της ασυμβατότητας των συνόλων δεδομένων και των μοντέλων, δεν είναι δυνατόν να ελεγχθεί η ικανότητα γενίκευσης των μοντέλων των μαθημάτων A3 2019-2020 & 2021-2022 επί του μαθήματος AA005 2020-2021 ή επί του μαθήματος AA005 2021-2022. Για τον ίδιο λόγο δεν είναι δυνατόν να γίνει και το αντίστροφο, δηλαδή να ελεγχθεί η ικανότητα γενίκευσης των μοντέλων των μαθημάτων AA005 2020-2021 ή AA005 2021-2022 επί των μαθημάτων A3 2019-2020 & 2021-2022. Συνεπώς, εστίασαμε στα μοντέλα των μαθημάτων AA005 2020-2021 και AA005 2021-2022.

Αναζητήσαμε το αποδοτικότερο μοντέλο μεταξύ των μοντέλων του μαθήματος AA005 2021-2022, διότι πέτυχαν μεγαλύτερη τιμή στην ακρίβεια πρόβλεψης, έναντι των μοντέλων του

μαθήματος AA005 2019-2020 (βλ. Πίνακα 22 και Πίνακα 30). Ακολουθώντας, από τα μοντέλα αυτού του μαθήματος επιλέξαμε το μοντέλο του αλγόριθμου `functions.SimpleLogistic (A/A 7)`, διότι αφενός επιτυγχάνει την τέταρτη καλύτερη τιμή ακρίβειας πρόβλεψης (87.75%) την τελευταία (17<sup>η</sup>) εβδομάδα του μαθήματος και αφετέρου επιτυγχάνει την πρωιμότερη πρόβλεψη, την 9<sup>η</sup> εβδομάδα. Αυτό το μοντέλο (δηλαδή, το μοντέλο που παράγει ο αλγόριθμος `functions.SimpleLogistic`, όταν εφαρμοστεί στο σύνολο δεδομένων της 17<sup>ης</sup> εβδομάδας του μαθήματος AA005 2021-2022) το δοκιμάσαμε επί των συνόλων δεδομένων όλων των εβδομάδων του μαθήματος AA005 2020-2021. Οι τιμές για την ακρίβεια πρόβλεψης που προέκυψαν, εμφανίζονται στον Πίνακα 44 και αποτυπώνονται γραφικά στο Γράφημα 12.<sup>18</sup>

Συγκεκριμένα, στον Πίνακα 44 εμφανίζονται οι τιμές ακρίβειας πρόβλεψης που επιτυγχάνουν 3 διαφορετικά μοντέλα, όταν εφαρμοστούν επί ίδιων συνόλων δεδομένων και συγκεκριμένα επί των συνόλων δεδομένων του μαθήματος AA005 2020-2021. Στη 2<sup>η</sup> και στην 3<sup>η</sup> γραμμή του Πίνακα 44 εμφανίζονται οι τιμές της ακρίβειας πρόβλεψης που επιτυγχάνουν τα μοντέλα των αλγόριθμων `rules.ZeroR (A/A 1)` (τιμή βάσης) και `SimpleLogistic (A/A 7)`, τα οποία χάριν συντομίας ας τα ονομάσουμε αντιστοίχως «ZR2020-2021» και «SL2020-2021». Στην 4<sup>η</sup> γραμμή του Πίνακα 44 εμφανίζονται οι τιμές ακρίβειας πρόβλεψης του μοντέλου, του οποίου ελέγχουμε την ικανότητα γενίκευσής του. Δηλαδή, του μοντέλου που κατασκευάστηκε από την εφαρμογή του αλγόριθμου `functions.SimpleLogistic (A/A 7)` στο σύνολο δεδομένων της 17<sup>ης</sup> εβδομάδας του μαθήματος AA005 2021-2022, το οποίο χάριν συντομίας ας το ονομάσουμε «SL2021-2022». Δηλαδή, τα μοντέλα ZR2020-2021 (2<sup>η</sup> γραμμή του πίνακα) και SL2020-2021 (3<sup>η</sup> γραμμή του πίνακα) κατασκευάστηκαν και εφαρμόστηκαν στο ίδιο, «οικείο» σύνολο δεδομένων, ενώ το μοντέλο SL2021-2022 κατασκευάστηκε σε διαφορετικό σύνολο δεδομένων από αυτό που εφαρμόστηκε.

---

<sup>18</sup> Ας σημειωθεί ότι στο Γράφημα 12 απεικονίζονται όλες οι τιμές, ανεξαρτήτως αν η διαφορά τους από την τιμή βάσης είναι ή δεν είναι στατιστικώς σημαντική.

**Πίνακας 44. Ακρίβεια πρόβλεψης μοντέλου SimpleLogistic AA005 2021-2022 (17η εβδομάδα), που εφαρμόστηκε σε συνόλων δεδομένων του μαθήματος AA005 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Εβδομάδα	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
ZR2020-2021	73.19	74.9	74.52	74.52	73.04	72.58	72.58	72.12	70.75	70.75	70.75	70.75	69.83	69.83	69.83	69.83	69.83	69.83
SL2020-2021	72.06	79.15	75.8	75.12	76.69	77.3	79.19	78.74	77.79	79.15	79.77	81.35	84.12	81.59	82.04	82.52	82.84	84.54
SL2021-2022	39.55	45.57	63.75	72.48	78.28	67.32	75.81	77.9	77.7	78.98	80.25	82.16	83.64	82.38	83.01	81.13	79.24	81.13

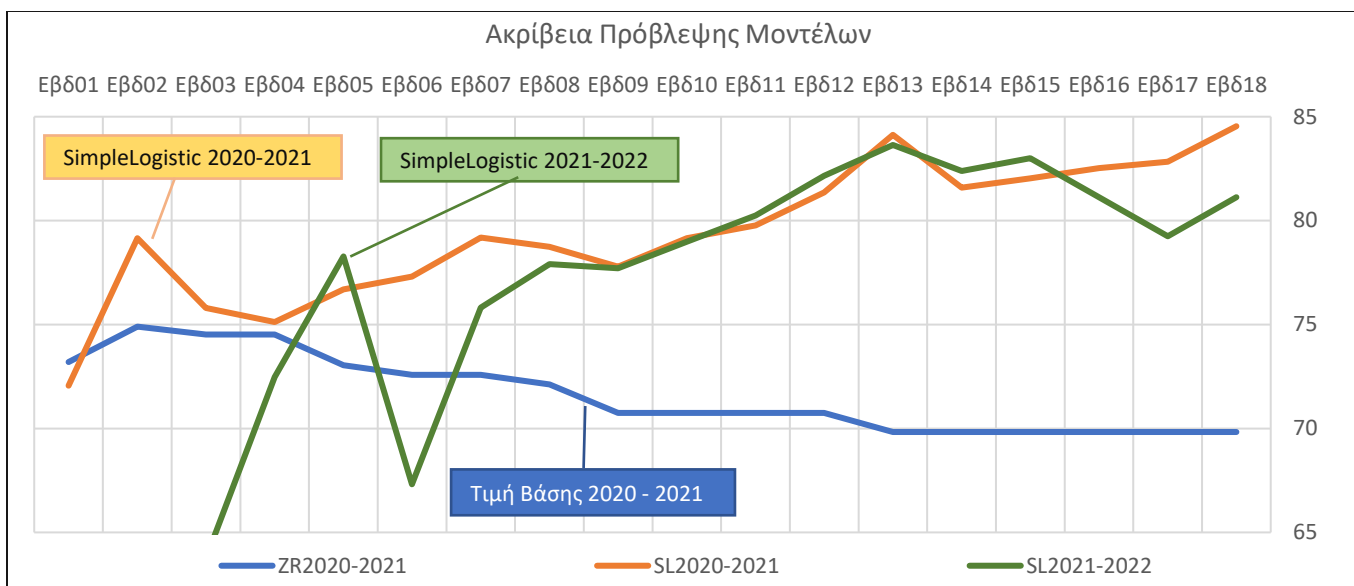
**ΥΠΟΜΝΗΜΑ**

ΣΥΝΤΟΜΟΓΡΑΦΙΑ ΠΕΡΙΓΡΑΦΗ

ZR 2021-2022 Μοντέλο αλγόριθμου ZeroR AA005 2020-2021 (Τιμή Βάσης)

SL 2021-2022 Μοντέλο αλγόριθμου SimpleLogistic AA005 2020-2021

SL 2021-2022 Μοντέλο αλγόριθμου SimpleLogistic AA005 2021-2022



**Γράφημα 12. Ακρίβεια πρόβλεψης μοντέλου SimpleLogistic AA005 2021-2022 (17η εβδομάδα) επί συνόλων δεδομένων του μαθήματος AA005 2020-2021 (οι αριθμητικές τιμές είναι ποσοστά επί τοις εκατό)**

Παρατηρούμε στον Πίνακα 44 και στο Γράφημα 12, ότι το μοντέλο SL2021-2022 επιτυγχάνει εξαιρετική τιμή ακρίβεια πρόβλεψης (81.13%) στο «μη οικείο σύνολο» δεδομένων της τελευταίας (18<sup>ης</sup>) εβδομάδας του μαθήματος AA005 2020-2021. Παρατηρούμε επίσης, ότι κατά τις προηγούμενες εβδομάδες αδιαλείπτως έως και την 7<sup>η</sup> εβδομάδα, οι τιμές ακρίβειας πρόβλεψης αφενός είναι πολύ κοντά στις τιμές του μοντέλου SL2020-2021 (δηλαδή του «οικείου» για το μάθημα μοντέλου) και αφετέρου υπερβαίνουν αρκετά την τιμή βάσης. Η απόκλιση των τιμών στην ακρίβεια πρόβλεψης από την 7<sup>η</sup> ως την 18<sup>η</sup> εβδομάδα δεν υπερβαίνει το 5% και είναι μικρότερη από την απόκλιση (9% ως 28%), που σύμφωνα με τους Riestra-Gonzalez, del Puerto Paule-Ruiz & Ortin (2021, σ. 4) καταγράφεται σε προηγούμενες ανάλογες μελέτες ο οποίες διερευνούν ικανότητα γενίκευσης των μοντέλων.

Επιπλέον, όπως φαίνεται στον Πίνακα 45, οι τιμές του μοντέλου SL2021-2022 για την μεταβλητή AUC κρίνονται καλές ως εξαιρετικές (Riestra-Gonzalez, del Puerto Paule-Ruiz & Ortin, 2021, σ. 3) και οι τιμές για την μεταβλητή Cohen's Kappa δηλώνουν «δίκαιη», έως «μέση» συμφωνία (Quinn & Gray, 2020, σ. 8) μεταξύ των πραγματικών κλάσεων και των κλάσεων που προβλέπει το μοντέλο.

**Πίνακας 45. Τιμές μεταβλητών AUC & Cohen's Kappa του μοντέλου SimpleLogistic AA005 2021-2022 (17η εβδομάδα) επί συνόλων δεδομένων του μαθήματος AA005 2020-2021**

Εβδομ	01 <sup>η</sup>	02 <sup>η</sup>	03 <sup>η</sup>	04 <sup>η</sup>	05 <sup>η</sup>	06 <sup>η</sup>	07 <sup>η</sup>	08 <sup>η</sup>	09 <sup>η</sup>	10 <sup>η</sup>	11 <sup>η</sup>	12 <sup>η</sup>	13 <sup>η</sup>	14 <sup>η</sup>	15 <sup>η</sup>	16 <sup>η</sup>	17 <sup>η</sup>	18 <sup>η</sup>
AUC	0.62	0.71	0.71	0.71	0.76	0.69	0.79	0.8	0.83	0.83	0.81	0.83	0.84	0.82	0.81	0.79	0.78	0.85
Kappa	0.08	0.14	0.22	0.32	0.42	0.28	0.33	0.4	0.36	0.44	0.46	0.51	0.57	0.51	0.54	0.5	0.43	0.46

Ως συμπέρασμα από τα παραπάνω, εκτιμούμε ότι για το μοντέλο SL2021-2022 προκύπτουν ενδείξεις ότι διαθέτει ικανότητα γενίκευσης, αφενός με υψηλές τιμές ακρίβειας πρόβλεψης και αφετέρου με πρώιμη πρόβλεψη στα μέσα περίπου του μαθήματος. Η εκτίμησή μας θα ήταν δυνατόν να ενισχυθεί περαιτέρω, αν η ικανότητα γενίκευσης του μοντέλου ελεγχθεί και διαπιστωθεί σε περισσότερα μαθήματα. Ωστόσο, ας σημειωθεί, ότι ο έλεγχος δεν μπορεί να γίνει επί οποιουδήποτε συνόλου δεδομένων, αλλά προϋποθέτει σύνολα δεδομένων αντίστοιχων με το σύνολο δεδομένων επί του οποίου κατασκευάστηκε το μοντέλο.

Πέρα από τη διαπίστωση της ικανότητας γενίκευσης του μοντέλου, είναι απαραίτητη η πρακτική αξιοποίηση του μοντέλου με τρόπο, ώστε πριν την ολοκλήρωση του (μη οικείου) μαθήματος να παραχθεί κατάλληλη αναφορά, στην οποία θα εμφανίζεται η προβλεπόμενη επίδοση κάθε φοιτητή. Η συγκεκριμένη δυνατότητα προσφέρεται από το λογισμικό WEKA και θα μπορούσε να αποτελέσει τον πυρήνα ενός πληρέστερου συστήματος έγκαιρης προειδοποίησης (early-warning system). Ένα τέτοιο σύστημα θα μπορούσε να αλληλεπιδρά με το Σύστημα Διαχείρισης Μάθησης (ΣΔΜ) και με το Σύστημα Φοιτητολογίου, να ενσωματώνει δυνατότητες κατασκευής συνόλων δεδομένων και να αξιοποιείται από τους διδάσκοντες των μαθημάτων. Η ανάπτυξη, ή έστω ο σχεδιασμός, ενός τέτοιου συστήματος είναι πέραν από τους σκοπούς της διατριβής. Περιοριζόμαστε στην παρουσίαση της σχετικής αναφοράς που παράγει το WEKA (Πίνακας 46), η οποία παρέχει τη βασική πληροφορία που χρειάζεται το σύστημα έγκαιρης προειδοποίησης.

**Πίνακας 46. Τμήμα αναφοράς συστήματος WEKA για πρόβλεψη επίδοσης ανά οντότητα**

inst#	actual	predicted	error	prediction
1	1:Yes	1:Yes		0.838
2	1:Yes	1:Yes		0.994
3	1:Yes	2:No	+	0.63
4	2:No	1:Yes	+	0.573
5	1:Yes	1:Yes		0.964
6	1:Yes	1:Yes		0.934
7	1:Yes	1:Yes		0.998
8	1:Yes	1:Yes		0.926
9	1:Yes	1:Yes		0.986
10	1:Yes	1:Yes		0.732
11	1:Yes	1:Yes		0.991
12	2:No	2:No		0.513
13	2:No	1:Yes	+	0.629
14	1:Yes	1:Yes		0.942
15	1:Yes	1:Yes		0.998
...				

Ο Πίνακας 46 αποτελεί τμήμα της αναφοράς του WEKA στην οποία εμφανίζονται οι προβλέψεις του μοντέλου SL2021-2022 για κάθε οντότητα του «μη οικείου» συνόλου δεδομένων του μαθήματος AA005 2020-2021, κατά την 9<sup>η</sup> εβδομάδα. Εμφανίζει πληροφορίες μόνο για τις 15 πρώτες οντότητες του συνόλου δεδομένων. Η πλήρης αναφορά βρίσκεται στο Παράρτημα Γ.

Κάθε γραμμή του Πίνακα 46 αντιστοιχεί σε μια οντότητα (φοιτητή) του συνόλου δεδομένων. Στην 1<sup>η</sup> στήλη εμφανίζεται ο αύξων αριθμός της οντότητας, στη 2<sup>η</sup> στήλη η πραγματική κλάση της οντότητας, στην 3<sup>η</sup> στήλη η προβλεπόμενη από το μοντέλο κλάση της οντότητας, στην 4<sup>η</sup> στήλη η ένδειξη για τυχόν αστοχία του μοντέλου (δηλαδή, η ένδειξη αν η πραγματική κλάση διαφέρει από την προβλεπόμενη) και στην 5<sup>η</sup> στήλη, η πιθανότητα σύμφωνα με την οποία, η προβλεπόμενη κλάση είναι ίδια με την πραγματική.

Για παράδειγμα, το μοντέλο προέβλεψε με πιθανότητα 0.838 ότι η κλάση της 1<sup>ης</sup> οντότητας είναι η κλάση Yes (επιτυχών φοιτητής). Πράγματι, η πραγματική κλάση της 1<sup>ης</sup> οντότητας είναι η κλάση Yes. Επίσης, το μοντέλο προέβλεψε με πιθανότητα 0.63 ότι η κλάση της 3<sup>ης</sup> οντότητας είναι η κλάση No (μη επιτυχών φοιτητής). Όμως, το μοντέλο αστόχησε στην 3<sup>η</sup> οντότητα, διότι η πραγματική κλάση της 3<sup>ης</sup> οντότητας είναι Yes (επιτυχών φοιτητής).

### 10.5. Επιδραστικά χαρακτηριστικά

Τα χαρακτηριστικά που περιλαμβάνονται στα σύνολα δεδομένων δεν έχουν όλα την ίδια βαρύτητα, ως προς την απόδοση των μοντέλων. Ο εντοπισμός των επιδραστικών χαρακτηριστικών αποτελεί κρίσιμο ζήτημα, διότι, αφενός επιτρέπει την κατασκευή μικρότερων συνόλων δεδομένων, που περιλαμβάνουν χαρακτηριστικά που αποτυπώνουν μόνο σημαντικές πληροφορίες. Συνεπώς, απαλλάσει τους αλγόριθμους από την ανάλυση δεδομένων τα οποία δεν έχουν βαρύνουσα σημασία, δεν συνεισφέρουν στην έρευνα μοτίβων χρήσης και απλώς εισάγουν «θόρυβο» στο σύστημα. Αφετέρου, διότι τα επιδραστικά χαρακτηριστικά προσανατολίζουν τον διδάσκοντα του μαθήματος να εστιάσει και να διερευνήσει αντίστοιχα κρίσιμα στοιχεία στον σχεδιασμό του μαθήματος, ώστε να προβεί σε καίριες εκπαιδευτικές προσαρμογές.

Για τον εντοπισμό των επιδραστικών χαρακτηριστικών των μοντέλων, χρησιμοποιήσαμε τη σχετική μέθοδο που προσφέρεται από το λογισμικό WEKA. Συγκεκριμένα, χρησιμοποιήσαμε τη δυνατότητα εντοπισμού χαρακτηριστικών με τη μέθοδο «CorrelationAttributeEval» (Select Attributes > Attribute Evaluator = CorrelationAttributeEval). Η συγκεκριμένη μέθοδος αποτιμά την επίδραση / βαρύτητα κάθε χαρακτηριστικού, υπολογίζοντας την τιμή του συντελεστή συσχέτισης (Pearson  $r$ ) μεταξύ του χαρακτηριστικού και της κλάσης (Zvornicanin, 2022). Η μέθοδος «CorrelationAttributeEval» μπορεί να χρησιμοποιηθεί (και πράγματι χρησιμοποιείται ευρέως) σε σύνολα δεδομένων με κατηγορικές κλάσεις (βλ. Κεφ. 9.7), όπως προκύπτει από σχετικές βιβλιογραφικές αναφορές (Brownlee, 2019b. Gnanambal, Thangaraj, Meenatchi & Gayathri, 2018. Hussain, Dahan, Ba-Alwib & Ribata, 2018. Pehlivanova & Nedeva, 2021).

Εφαρμόσαμε τη μέθοδο στα σύνολα δεδομένων, που αντιστοιχούν στην περίοδο από την αρχή μέχρι το τέλος των μαθημάτων. Δεν υλοποιήσαμε ανάλογη ενέργεια σε σύνολα δεδομένων που κάλυπταν μικρότερες περιόδους, διότι στην αναζήτηση επιδραστικών χαρακτηριστικών δεν ενδιαφέρει η πρώιμη πρόβλεψη, ούτε ενδιαφέρει η επίδραση χαρακτηριστικών σε πρώιμες χρονικές στιγμές. Ενδιαφέρει αποκλειστικά η επίδραση χαρακτηριστικών για όλη την περίοδο διεξαγωγής των μαθημάτων. Από τα χαρακτηριστικά που υπέδειξε η μέθοδος επικεντρωθήκαμε σε αυτά με τους 5 μεγαλύτερους συντελεστές συσχέτισης. Στον Πίνακα 47, παρουσιάζονται τα σχετικά αποτελέσματα. Οι τιμές των συντελεστών συσχέτισης που υπολογίστηκαν από το WEKA,



υπολογίστηκαν και από το λογισμικό στατιστικής ανάλυσης IBM SPSS, από όπου προέκυψε η πληροφορία ότι η τιμή στατιστικής σημαντικότητας είναι  $p < 0.01$ .

**Πίνακας 47. Επιδραστικά χαρακτηριστικά μοντέλων σύμφωνα με τον συντελεστή συσχέτισης (φθίνουσα διάταξη) μεταξύ χαρακτηριστικών και κλάσης ταξινόμησης**

	AA005 2020-2021		AA005 2021-2022		A3 2019-2020 & 2020-2021	
	Χαρακτηριστικό	r ( $p < 0.01$ )	Χαρακτηριστικό	r ( $p < 0.01$ )	Χαρακτηριστικό	r ( $p < 0.01$ )
1	Πλήθος λεπτών αδράνειας κατά τη χρονική περίοδο που ο φοιτητής ήταν ενεργός	0.555	Πλήθος ημερών μεταξύ της τελευταίας συνόδου και του τέλους χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων	0.6714	Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο	0.4311
2	Πλήθος ημερών μεταξύ πρώτης και τελευταίας συνόδου σύνδεσης	0.554	Πλήθος λεπτών αδράνειας κατά τη χρονική περίοδο που ο φοιτητής ήταν ενεργός	0.6134	Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου	0.4169
3	Πλήθος συνόδων τις πρωινές ώρες	0.497	Πλήθος ημερών μεταξύ πρώτης και τελευταίας συνόδου σύνδεσης	0.613	Πλήθος ημερών όπου καταγράφηκε δραστηριότητα	0.4118
4	Πλήθος ημερών μεταξύ της τελευταίας συνόδου και του τέλους χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων	0.452	Πλήθος ημερών αδράνειας κατά την διάρκεια της ενεργής περιόδου	0.5853	Τυπική απόκλιση των ημερών αδράνειας για τις ημέρες αδράνειας του ίδιου φοιτητή	0.4045
5	Πλήθος ημερών αδράνειας κατά την διάρκεια της ενεργής περιόδου	0.432	Πλήθος συνόδων τις εργάσιμες ημέρες	0.4019	Μέσος όρος πλήθους ημερών αδράνειας μεταξύ των συνόδων σύνδεσης	0.3729

Στον Πίνακα 47, εμφανίζονται τα χαρακτηριστικά με τους πέντε υψηλότερους συντελεστές συσχέτισης μεταξύ του χαρακτηριστικού και της κλάσης ταξινόμησης. Ο συντελεστής εμφανίζεται δίπλα στο όνομα κάθε χαρακτηριστικού. Τα χαρακτηριστικά που εμφανίζονται σε περισσότερα από ένα μαθήματα έχουν έγχρωμο υπόβαθρο.

Καταρχάς, παρατηρούμε ότι οι τιμές των συντελεστών είναι μεγαλύτερες (με την εξαίρεση ενός μόνο χαρακτηριστικού) από την τιμή 0.4 και για κάποια χαρακτηριστικά είναι μεγαλύτερες από την τιμή 0.6, γεγονός που δηλώνει ότι η συσχέτιση μεταξύ των χαρακτηριστικών και της κλάσης

είναι αξιοσημείωτη και όχι αμελητέα. Προκειμένου να επιβεβαιώσουμε στην πράξη την επίδραση των χαρακτηριστικών, που υποδεικνύονται από τη μέθοδο, υπολογίσαμε την ακρίβεια πρόβλεψης που επιτυγχάνεται όταν ο αλγόριθμος functions.SimpleLogistic (A/A 7) εφαρμοσθεί σε σύνολο δεδομένων χωρίς τα επιδραστικά χαρακτηριστικά. Μετά από σύγκριση με την ακρίβεια πρόβλεψης που επιτυγχάνεται όταν ο αλγόριθμος εφαρμοσθεί σε πλήρες σύνολο δεδομένων, προκύπτει ότι πράγματι, η μη συμπερίληψη των επιδραστικών χαρακτηριστικών επηρεάζει την απόδοση των μοντέλων (Πίνακας 48).

**Πίνακας 48. Ακρίβεια πρόβλεψης του μοντέλου functions.SimpleLogistic για σύνολα δεδομένων που διαφέρουν ως προς τα επιδραστικά χαρακτηριστικά**

Μάθημα	Πλήρες Σύνολο Δεδομένων	Σύνολο δεδομένων χωρίς επιδραστικά χαρακτηριστικά
AA005 2020-2021	84.54%	82.20%
AA005 2021-2022	87.75%	83.62%
A3 2019-2020 & 2021-2022	75.27%	60.04%

Παρατηρούμε επίσης στον Πίνακα 47, ότι τέσσερα από τα πέντε επιδραστικά χαρακτηριστικά των μαθημάτων AA005 2020-2021 και AA005 2021-2022 είναι κοινά. Επίσης, ότι τα μαθήματα A3 2019-2020 & 2021-2022 δεν έχουν κοινά επιδραστικά χαρακτηριστικά με τα μαθήματα AA005. Θεωρούμε ότι το γεγονός οφείλεται στις διαφορές στον σχεδιασμό των μαθημάτων και αναδεικνύει τη σημασία που έχει ο εκπαιδευτικός σχεδιασμός στην απόδοση των μοντέλων. Ειδικότερα, τα μαθήματα AA005 2020-2021 και AA005 2021-2022 έχουν ελάχιστες διαφορές μεταξύ τους ως προς τις θεματικές ενότητες, το εκπαιδευτικό υλικό, τον τρόπο επικοινωνίας του διδάσκοντα με τους φοιτητές, την αξιολόγηση. Οι ομοιότητες έχουν ως συνέπεια την ανάδειξη κοινών επιδραστικών χαρακτηριστικών. Τα μαθήματα A3 2019-2020 & 2021-2022 διαφέρουν από τα μαθήματα AA005, στην εκπαιδευτική βαθμίδα (μεταπτυχιακό μάθημα), στο γνωστικό αντικείμενο, στο εκπαιδευτικό υλικό, στην οργάνωση των ενοτήτων, στις εκπαιδευτικές δραστηριότητες, στην υποχρεωτικότητα παρακολούθησης, στον τρόπο αξιολόγησης. Οι διαφορές στον σχεδιασμό των μαθημάτων A3 2019-2020 & 2021-2022 σε σχέση με τον σχεδιασμό των μαθημάτων AA005, εκτιμούμε ότι αποτελεί την αιτία της ανάδειξης εντελώς διαφορετικών επιδραστικών χαρακτηριστικών.

Στο Πίνακα 49, παρουσιάζεται η συσχέτιση των επιδραστικών χαρακτηριστικών με τις διαστάσεις αλληλεπίδρασης όπως αυτές περιγράφηκαν και αναλύθηκαν στα κεφάλαια 9.5.5 και 9.5.6 και στους Πίνακες 19 και 20. Παρατηρούμε ότι τα επιδραστικά χαρακτηριστικά σχετίζονται συχνότερα με τις διαστάσεις του όγκου, της διάρκειας, της ομαλότητας, της κανονικότητας, της ασυνέπειας.

**Πίνακας 49. Συσχέτιση των επιδραστικών χαρακτηριστικών με τις διαστάσεις αλληλεπίδρασης**

Επιδραστικό Χαρακτηριστικό	Διάσταση Αλληλεπίδρασης								
	Όγκος	Διάρκεια	Ομαλότητα	Κανονικότητα	Ασυνέπεια	Αμεσότητα	Προτιμώμενη Χρονική Στιγμή	Επίδοση	Τύπος μαθησιακού πόρου
Πλήθος λεπτών αδράνειας κατά τη χρονική περίοδο που ο φοιτητής ήταν ενεργός	√	√	√	√					
Πλήθος ημερών μεταξύ πρώτης και τελευταίας συνόδου σύνδεσης					√	√			
Πλήθος ημερών μεταξύ της τελευταίας συνόδου και του τέλους χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων			√	√	√				
Πλήθος ημερών αδράνειας κατά την διάρκεια της ενεργής περιόδου			√	√	√				
Πλήθος συνόδων τις πρωινές ώρες							√		
Πλήθος συνόδων τις εργάσιμες ημέρες							√		
Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο								√	
Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου	√								√
Πλήθος ημερών όπου καταγράφηκε δραστηριότητα	√	√							
Τυπική απόκλιση των ημερών αδράνειας			√	√	√				

για τις ημέρες αδράνειας του ίδιου φοιτητή									
Μέσος όρος πλήθους ημερών αδράνειας μεταξύ των συνόδων σύνδεσης	√	√	√	√					

Ωστόσο, υπάρχουν και κάποια χαρακτηριστικά που δεν ακολουθούν τον κανόνα και σχετίζονται με διαστάσεις όπως η χρονική στιγμή, η επίδοση, ο τύπος του μαθησιακού πόρου. Ειδικότερα, στο μάθημα AA005 2020-2021, αιφνιδιάζει η ανάδειξη ως επιδραστικού του χαρακτηριστικού «Πλήθος συνόδων τις πρωινές ώρες». Γιατί άραγε, η ώρα σύνδεσης αποτελεί πληροφορία που επηρεάζει την ακρίβεια πρόβλεψης και γιατί αναδεικνύεται επιδραστική μόνο για το συγκεκριμένο μάθημα; Η άπαντηση ίσως σχετίζεται με την ιδιαιτερότητα των διαλέξεων του μαθήματος. Ειδικότερα, λόγω των περιοριστικών μέτρων της πανδημίας Covid-19, οι διαλέξεις του μαθήματος πραγματοποιήθηκαν αποκλειστικά μέσω τηλεδιάσκεψης. Οι διαλέξεις γίνονταν κάθε Τρίτη από τις 08.00 π.μ. έως τις 11.00 π.μ., δηλαδή τις ώρες οι οποίες συμπεριλαμβάνονται στο διάστημα 07.00 π.μ. έως 12.00, τις οποίες το σύνολο δεδομένων καταγράφει ως πρωινές ώρες. Συνεπώς, οι φοιτητές που συμμετείχαν στις τηλεδιασκέψεις είχαν υψηλή τιμή σε αυτό το χαρακτηριστικό και αντιθέτως οι φοιτητές που δεν συμμετείχαν στις τηλεδιασκέψεις είχαν χαμηλή τιμή σε αυτό το χαρακτηριστικό. Συνεπώς, το χαρακτηριστικό λειτουργεί ως οιονεί «παρουσιολόγιο» για τις διαλέξεις του μαθήματος και μόνο αυτού του μαθήματος. Κατά αναλογία, ως οιονεί «παρουσιολόγιο» στις διαλέξεις του μαθήματος AA005 2021-2022 λειτουργεί το χαρακτηριστικό «Πλήθος συνόδων τις εργάσιμες ημέρες». Διότι, οι διαλέξεις του μαθήματος AA005 2021-2022 έγιναν δια ζώσης, κάθε Παρασκευή από τις 08.00 π.μ. έως τις 11.00 π.μ. για την πρώτη ομάδα φοιτητών και από τις 11.00 π.μ. έως τις 02.00 μ.μ. για την δεύτερη ομάδα φοιτητών. Κατά τη διάρκεια των διαλέξεων, οι φοιτητές συνδέονταν στο σύστημα και επίλυαν ασκήσεις που είχαν αναρτηθεί στη σελίδα του μαθήματος στο Moodle. Το χαρακτηριστικό «Πλήθος συνόδων τις πρωινές ώρες», δεδομένου ότι καταγράφει συνόδους ως τις 12.00, δεν είναι ικανό να καταγράψει την παρουσία ή την απουσία των φοιτητών. Διότι, όσοι φοιτητές ήταν παρόντες στις δια ζώσης διαλέξεις και συνδέονταν στο σύστημα για να επιλύσουν ασκήσεις, καταγραφόνταν από το σύστημα ως συνδεδεμένοι, είτε τις πρωινές ώρες, είτε τις απογευματινές ώρες (12.00 έως 09.00 μ.μ). Ομως, το χαρακτηριστικό «Πλήθος συνόδων τις

εργάσιμες ημέρες», καταγράφει και πρωινές και απογευματινές συνόδους σύνδεσης, οπότε παρέχει πληροφορία για την παρουσία ή την απουσία φοιτητών από τις διαλέξεις. Τα παραπάνω, ίσως εξηγούν όχι μόνο την ανάδειξη των εν λόγω χαρακτηριστικών ως επιδραστικών, αλλά και το γεγονός ότι τα επιδραστικά χαρακτηριστικά των μαθημάτων διαφέρουν μόνο ως προς αυτό το σημείο. Θα μπορούσαμε μάλιστα να ισχυριστούμε, ότι αυτά δύο χαρακτηριστικά, στα οποία διαφέρουν τα δύο μαθήματα, ουσιαστικά σχετίζονται με την ίδια πληροφορία, δηλαδή την παρουσία ή την απουσία από τις διαλέξεις του μαθήματος. Θα μπορούσαμε επίσης να ισχυριστούμε ότι αν ορίζαμε χαρακτηριστικό, που στο μάθημα AA005 2020-2021 θα κατέγραφε συνδέσεις κατά το χρονικό διάστημα 08.00 π.μ. έως 11.00 π.μ. και στο μάθημα AA005 2021-2022 συνδέσεις κατά το χρονικό 08.00 π.μ. έως 02.00 μ.μ., τότε το νέο αυτό χαρακτηριστικό, αφενός θα αναδεικνυόταν επιδραστικό και αφετέρου θα ήταν το πέμπτο κοινό χαρακτηριστικό των δύο μαθημάτων.

Ως προς τα μαθήματα A3 2019-2020 & 2021-2022, τα χαρακτηριστικά που αναδεικνύονται ως επιδραστικά είναι ο «Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο» και το «Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου». Ούτε αυτά τα χαρακτηριστικά σχετίζονται με τις συχνότερα εμφανιζόμενες διαστάσεις του όγκου, της διάρκειας, της ομαλότητας, της κανονικότητας και της ασυνέπειας. Σχετίζονται το μεν πρώτο με την επίδοση, το δε δεύτερο με τον τύπο του μαθησιακού πόρου. Η ανάδειξή τους ως επιδραστικών, θεωρούμε ότι οφείλεται στον εκπαιδευτικό σχεδιασμό των μαθημάτων A3 2019-2020 & 2021-2022. Όπως έχει αναφερθεί (βλ. Κεφ. 9.2), σε κάθε συνεδρία των μαθημάτων A3 2019-2020 & 2021-2022, διεξάγεται διαδικτυακός δημόσιος διάλογος, όπου η συμμετοχή των φοιτητών, αφενός είναι υποχρεωτική και αφετέρου αξιολογείται και προσμετράται κατά 30% στον τελικό βαθμό των φοιτητών. Συνεπώς, η επιρροή του χαρακτηριστικού «Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο» είναι ευεξήγητη. Οφείλεται στην άμεση αλγεβρική σχέση της τιμής του χαρακτηριστικού με τον τελικό βαθμό, ο οποίος καθορίζει την κλάση στην οποία κατατάσσεται ο φοιτητής. Επίσης, ευεξήγητη είναι η επιρροή του χαρακτηριστικού «Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου». Διότι, δεδομένης της σπουδαιότητας των δημόσιων διαλόγων στον σχεδιασμό του μαθήματος, η συμμετοχή (ή η μη συμμετοχή) του φοιτητή στους δημόσιους διαλόγους αποτελεί ισχυρή

ένδειξη της γενικότερης ενασχόλησης και προσήλωσης του φοιτητή στο μάθημα. Θα μπορούσε κάποιος να υποστηρίξει, ότι το χαρακτηριστικό «Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο» αναδεικνύεται ως επιδραστικό, όχι διότι εκφράζει κάποιο μοτίβο χρήσης, αλλά διότι απλώς και μόνο σχετίζεται αλγεβρικά με τον τελικό βαθμό. Αντιτείνουμε, ότι ο βαθμός που ο φοιτητής λαμβάνει για τον δημόσιο διάλογο είναι αντιπροσωπευτικός για την ποιότητα των αναρτήσεων του. Οπότε, ακόμα και αν δεν επηρέαζε αλγεβρικά τον τελικό βαθμό, δηλαδή ακόμα και αν αποτελούσε μια απλή τιμή στα σύνολα δεδομένων, χωρίς άμεση αλγεβρική σχέση με τον τελικό βαθμό, θα ήταν οπωσδήποτε αποκαλυπτικός για την γενικότερη ενασχόληση και την προσήλωση του φοιτητή στο μάθημα.

Είναι αξιοσημείωτο ότι τα δύο χαρακτηριστικά που σχετίζονται με τον δημόσιο διάλογο αναδεικνύονται επιδραστικότερα έναντι των υπόλοιπων τριών χαρακτηριστικών, τα οποία σχετίζονται με τον όγκο, τη διάρκεια, την ομαλότητα, την κανονικότητα, την ασυνέπεια πρόσβασης. Είναι επίσης αξιοσημείωτο ότι στα επιδραστικά χαρακτηριστικά των μαθημάτων A3 2019-2020 & 2020-2021, σε αντίθεση με ό,τι συμβαίνει στα άλλα μαθήματα, δεν περιλαμβάνεται χαρακτηριστικό που να σχετίζεται με τυχόν καθυστερημένη έναρξη ενασχόλησης σύνδεσης με το σύστημα, ή με τυχόν πρόωρη διακοπή ενασχόλησης με το σύστημα, ή με το χρονικό διάστημα μεταξύ πρώτης και τελευταίας συνόδου σύνδεσης. Το εν λόγω γεγονός οφείλεται στην υποχρεωτικότητα της συμμετοχής στο μάθημα από την πρώτη ως την τελευταία συνεδρία, σε αντίθεση με τα μαθήματα AA005 2020-2021 και AA005 2021-2022, όπου η παρακολούθηση δεν είναι υποχρεωτική.

## 11. ΑΠΑΝΤΗΣΕΙΣ ΣΤΑ ΕΡΕΥΝΗΤΙΚΑ ΕΡΩΤΗΜΑΤΑ

Η ανάλυση των ερευνητικών δεδομένων επιτρέπει την εξαγωγή πορισμάτων ικανών να δώσουν απαντήσεις στα ερωτήματα στα οποία εστίασε η έρευνα. Δηλαδή, απαντήσεις που αφορούν τις δυνατότητες των μοντέλων Μηχανικής Μάθησης να αξιοποιήσουν δεδομένα χρήσης των Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ) και να προβλέψουν τις επιδόσεις των φοιτητών με ακρίβεια και εγκαίρως, έτσι ώστε να αξιοποιηθούν ως εργαλεία διαφορικής αξιολόγησης. Παράλληλα, η ανάλυση των ερευνητικών δεδομένων αναδεικνύει διαστάσεις σχετικές με τις σύνθετες συσχετίσεις που υφίστανται, αφενός μεταξύ των δεδομένων πρόσβασης στα ΣΔΜ και της ενεργής συμμετοχής των φοιτητών στην εκπαιδευτική διαδικασία και αφετέρου μεταξύ των δεδομένων πρόσβασης στα ΣΔΜ και του εκπαιδευτικού σχεδιασμού των μαθημάτων.

Πριν την ανάπτυξη των απαντήσεων, παραθέτουμε συνοπτικώς τα στάδια της έρευνας: Συλλέχθηκαν δεδομένα πρόσβασης στο ΣΔΜ Moodle από **440** φοιτητές που παρακολούθησαν **4** μαθήματα προπτυχιακού και μεταπτυχιακού επιπέδου. Τα δεδομένα πρόσβασης, μετά από κατάλληλη επεξεργασία, ενσωματώθηκαν σε σύνολα δεδομένων αποτελούμενα από χαρακτηριστικά, συνήθη σε ανάλογες έρευνες και από διχοτομημένη κατηγορική κλάση. Τα σύνολα δεδομένων αναλύθηκαν με αλγόριθμους Μηχανικής Μάθησης από τους οποίους προέκυψαν μοντέλα πρόβλεψης της επίδοσης των φοιτητών.

## 11.1. Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ1

### *Ερευνητικό Ερώτημα ΕΕΡ1:*

Είναι εφικτή η πρόβλεψη της επίδοσης των φοιτητών με τη βοήθεια μοντέλων Μηχανικής Μάθησης τα οποία βασίζονται αποκλειστικά σε δεδομένα χρήσης του ΣΔΜ; Επιπλέον:

1.1 Τι τιμή ακρίβειας πρόβλεψης επιτυγχάνεται;

1.2 Πόσο αξιόπιστη είναι η πρόβλεψη;

Τα δεδομένα πρόσβασης που μελετήσαμε αφορούσαν 4 διαφορετικά μαθήματα. Τα σύνολα δεδομένων περιλαμβάνουν δεδομένα που προέρχονται αποκλειστικώς από τα αρχεία καταγραφής (log files) του ΣΔΜ Moodle, συνεπώς αποτυπώνουν πληροφορίες σχετικές με τη χρήση του συστήματος από τους φοιτητές. Τα σύνολα δεν περιλαμβάνουν δημογραφικά και κοινωνικό-οικονομικά δεδομένα των φοιτητών, ούτε δεδομένα σχετικά με προηγούμενες ακαδημαϊκές επιδόσεις τους, με απαντήσεις τους σε ερωτηματολόγια κ.τ.λ. Η κλάση των συνόλων αποτυπώνει την επίδοση των φοιτητών και λαμβάνει δύο κατηγορικές τιμές, που αντιστοιχούν σε επιτυχημένη και σε μη επιτυχημένη τελική επίδοση. Για κάθε μάθημα, στο σύνολο δεδομένων που αντιστοιχεί σε χρονικό διάστημα από την αρχή ως το τέλος του μαθήματος, εφαρμόστηκαν 20 αλγόριθμοι Μηχανικής Μάθησης. Οι τιμές για την ακρίβεια πρόβλεψης που επιτεύχθηκαν από κάθε μοντέλο και για κάθε μάθημα καταγράφηκαν και συγκρίθηκαν.

Σε όλα τα μαθήματα, οι τιμές της ακρίβειας πρόβλεψης είναι αρκετά υψηλές, ώστε να θεωρήσουμε ότι η πρόβλεψη της επίδοσης των φοιτητών είναι εφικτή. Στο μάθημα AA005 2020-2021 η τιμή ακρίβειας πρόβλεψης ανήκει στο διάστημα [76.62% , 85.80%], στο μάθημα AA005 2021-2022 ανήκει στο διάστημα [83.81% , 89.53%] και στα μαθήματα A3 2019-2020 & 2021-2022 ανήκει στο διάστημα [55.09% , 79.37%]. Το εύρος στο οποίο κινείται η τιμή της ακρίβειας πρόβλεψης στα μαθήματα A3 2019-2020 & 2021-2022, παρότι είναι ικανοποιητικό και πάντοτε υψηλότερο από την τιμή βάσης, υπολείπεται εμφανώς από τις αντίστοιχες τιμές στα μαθήματα AA005 2020-2021 και AA005 2021-2022. Η διαφορά στην ακρίβεια πρόβλεψης, ίσως οφείλεται



στο κριτήριο διχοτόμησης των κλάσεων. Όπως έχει αναφερθεί, στα μαθήματα AA005 2020-2021 και AA005 2021-2022 η διχοτόμηση έγινε με κριτήριο την επιτυχία ή αποτυχία στην τελική επίδοση των φοιτητών, ενώ στα μαθήματα A3 2019-2020 & 2021-2022 η διχοτόμηση έγινε με κριτήριο τη διάμεσο της τελικής επίδοσης των φοιτητών. Τα μοντέλα πρόβλεψης ταξινόμησαν τις οντότητες μελετώντας διαφορές στα μοτίβα (patterns) χρήσης του ΣΔΜ, οι οποίες είναι εμφανέστερες στα μοτίβα χρήσης μεταξύ επιτυχόντων και αποτυχόντων φοιτητών, από ό,τι είναι στα μοτίβα χρήσης μεταξύ φοιτητών που έλαβαν βαθμό μεγαλύτερο από τη διάμεσο (π.χ. με βαθμό 9) και ίσο (π.χ. 8.5) ή μικρότερο από τη διάμεσο (Quinn & Gray, 2020, σ. 13). Επιπλέον, ο τρόπος βαθμολόγησης των μαθημάτων A3 2019-2020 & 2020-2021, που καθόριζεται από τη συμμετοχή στους δημόσιους διαλόγους και από την εκπόνηση γραπτής τελικής εργασίας, ενέχει μεγαλύτερη ρευστότητα από ό,τι ο τρόπος βαθμολόγησης των μαθημάτων AA005 2020-2021 και AA005 2021-2022, όπου οι φοιτητές αξιολογούνται με τελικές γραπτές εξετάσεις. Συνεπώς, η κατάταξη σε κλάσεις είναι εγγενώς δυσκολότερη στα μαθήματα A3 2019-2020 & 2020-2021, οπότε είναι αναμενόμενη η μικρότερη τιμή στην ακρίβεια πρόβλεψης. Η αξιολόγηση μέσω γραπτής εργασίας, δηλαδή δραστηριότητας ξένης με το ΣΔΜ, δυσχεραίνει περαιτέρω την απόδοση του μοντέλου, όπως αναφέρουν ότι συμβαίνει σε αντίστοιχη έρευνα οι Conijn, Snijders, Kleingeld & Matzat (2016, σ. 27). Εξάλλου, σύμφωνα με τους Gasevic, Dawson, Rogers & Gasevic (2016) στα σύνολα δεδομένων, όπου συγχωνεύονται δεδομένα από πολλά μαθήματα (ό,τι δηλαδή συμβαίνει για το A3 2019-2020 & 2020-2021), η ακρίβεια πρόβλεψης μειώνεται (Quinn & Gray, 2020, σ. 13). Ωστόσο, παρότι δεν φτάνει στα εξαιρετικά επίπεδα των μοντέλων των μαθημάτων AA005 2020-2021 και AA005 2021-2022, η τιμή ακρίβειας πρόβλεψης και για τα μαθήματα A3 2019-2020 & 2020-2021 είναι πολύ καλή. Για την πλειονότητα των μοντέλων, υπερβαίνει την τιμή βάσης με στατιστικώς σημαντική διαφορά και κινείται πάνω από 70%, ενώ σε κάποιες περιπτώσεις σχεδόν ισούται με 80%.

Η υψηλότερη τιμή στην ακρίβεια πρόβλεψης που επιτεύχθηκε για κάθε μάθημα παρουσιάζεται στη δεύτερη γραμμή του Πίνακα 50.

**Πίνακας 50. Υψηλότερες τιμές ακρίβειας πρόβλεψης, τιμές AUC και τιμές Cohen's Kappa ανά μάθημα.**

	AA005 2020-2021	AA005 2021-2022	A3 2019-2020 & 2021-2022
Υψηλότερη τιμή ακρίβειας πρόβλεψης	85.80%	89.53%	79.37%
Τιμή βάσης	69.83%	77.36%	53.93%
Στατιστική σημαντικότητα (p) υψηλότερης τιμής από την τιμή βάσης	<0.05	<0.05	<0.05
AUC	0.83	0.88	0.84
Cohen's Kappa	0.62	0.63	0.58

Για την αποτίμηση της αξιοπιστίας της πρόβλεψης καταρχάς συγκρίναμε την τιμή ακρίβειας πρόβλεψης με την τιμή βάσης, δηλαδή με την τιμή που επιτυγχάνει το μοντέλο του αλγόριθμου rules.ZeroR, το οποίο προβλέπει «τυφλά», ότι όλες ανεξαιρέτως οι οντότητες ανήκουν στην πολυπληθέστερη κλάση (Quinn & Gray, 2020, σ. 8). Η διαφορά μεταξύ των δύο τιμών είναι στατιστικώς σημαντική, με ποσοστό στατιστικής σημαντικότητας 0.05, ένδειξη ότι η πρόβλεψη είναι αξιόπιστη.

Επιπλέον, αποτιμήσαμε την απόδοση των μοντέλων μελετώντας και τα κριτήρια των τιμών AUC και Cohen's Kappa. Οι τιμές AUC και Cohen's Kappa που αντιστοιχούν στα μοντέλα που επιτυγχάνουν την υψηλότερη τιμή ακρίβειας πρόβλεψης παρουσιάζονται στην 5<sup>η</sup> και 6<sup>η</sup> γραμμή του Πίνακα 50. Σε όλα τα μαθήματα, οι τιμές της μεταβλητής AUC είναι μεγαλύτερες από 0.8 οπότε κρίνονται εξαιρετικές (Riestra-Gonzalez, del Puerto Paule-Ruiz & Ortin, 2021, σ. 3). Ως προς την μεταβλητή Cohen's Kappa, στα μαθήματα AA005 2020-2021 και AA005 2021-2022, οι τιμές είναι μεγαλύτερες από 0.6, οπότε δηλώνουν «σημαντική» (substantial) συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που προέβλεψε το μοντέλο. Στα μαθήματα A3 2019-2020 & 2021-2022 η τιμή είναι 0.58, οπότε δηλώνει «μέση» (moderate) συμφωνία (Quinn & Gray, 2020, σ. 8), ωστόσο η συγκεκριμένη τιμή είναι οριακά μικρότερη από την τιμή της σημαντικής συμφωνίας (0.6).

#### *Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ1*

Συνεπώς, από την ανάλυση των ερευνητικών αποτελεσμάτων που αφορούν τα μαθήματα που μελετήθηκαν, προκύπτει ότι είναι εφικτή η πρόβλεψη της επίδοσης των φοιτητών με τη βοήθεια μοντέλων Μηχανικής Μάθησης, τα οποία βασίζονται αποκλειστικά σε δεδομένα χρήσης του ΣΔΜ.

Η τιμή ακρίβειας πρόβλεψης, προσεγγίζει ή υπερβαίνει (ανάλογα με το μάθημα) το 80% και η αξιοπιστία της πρόβλεψης είναι ισχυρή.

## 11.2. Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ2

### *Ερευνητικό Ερώτημα ΕΕΡ2:*

Είναι εφικτή η πρώιμη πρόβλεψη της επίδοσης των φοιτητών; Επιπλέον:

- 2.1 Πόσο πρώιμα επιτυγχάνεται η πρόβλεψη;
- 2.2 Τι τιμή ακρίβειας πρόβλεψης επιτυγχάνεται;
- 2.3 Είναι η πρόβλεψη αξιόπιστη;

Για καθένα από τα τέσσερα μαθήματα, δεν περιορίσαμε την έρευνά μας μόνο στο πλήρες, βασικό σύνολο δεδομένων που καλύπτει όλες τις εβδομάδες από την αρχή έως το τέλος των μαθήματος. Προκειμένου μελετήσουμε διαφοροποιήσεις που συνέβησαν σε χρονικές στιγμές πριν την ολοκλήρωση των μαθημάτων, δημιουργήσαμε σύνολα δεδομένων που κάλυπταν και μικρότερες χρονικές περιόδους. Δηλαδή, για κάθε μάθημα, εκτός από το πλήρες σύνολο δεδομένων που καλύπτει όλες τις εβδομάδες του μαθήματος, επιπλέον δημιουργήθηκαν M-1 στο πλήθος σύνολα δεδομένων (όπου M το πλήθος των εβδομάδων ως την τελική αξιολόγηση του μαθήματος). Δηλαδή, δημιουργήθηκε σύνολο για την 1<sup>η</sup> εβδομάδα, σύνολο για την 2<sup>η</sup> εβδομάδα κ.ο.κ. Με τον τρόπο αυτό δημιουργήθηκαν συνολικά M σύνολα δεδομένων για κάθε μάθημα, στα οποία εφαρμόστηκαν 20 αλγόριθμοι Μηχανικής Μάθησης.

Οι τιμές για την ακρίβεια πρόβλεψης που επιτεύχθηκαν από κάθε μοντέλο, για κάθε μάθημα, για κάθε εβδομάδα του μαθήματος καταγράφηκαν και μελετήθηκαν, ώστε εκτός από την ακρίβεια πρόβλεψης να αναζητηθεί και η χρονική στιγμή κατά την οποία είναι εφικτή αξιόπιστη πρώιμη πρόβλεψη. Η χρονική στιγμή που συντελείται η πρόβλεψη είναι σημαντικό κριτήριο για τη χρησιμότητα των μοντέλων. Διότι, η αξία της πρόβλεψης εξαρτάται, όχι μόνο από το πόσο ακριβής είναι, αλλά και από το πόσο έγκαιρη είναι. Για παράδειγμα, έστω ένα μοντέλο που οδηγεί σε πρόβλεψη υψηλής ακρίβειας όταν το μάθημα ολοκληρωθεί. Το μοντέλο αυτό θεωρείται τεχνικώς επιτυχημένο, ωστόσο διαθέτει περιορισμένη εκπαιδευτική χρησιμότητα. Η χρησιμότητα των προβλεπτικών εκπαιδευτικών μοντέλων σχετίζεται κυρίως με την δυνατότητά τους να παράγουν πρώιμες προβλέψεις, έστω λιγότερο ακριβείς από τις τελικές.

Από τη μελέτη της διακύμανσης των τιμών στην ακρίβεια πρόβλεψης, προκύπτει ότι οι τιμές της ακρίβειας πρόβλεψης γενικά εμφανίζουν αυξητική τάση, καθώς κινούμαστε χρονικά από την έναρξη προς το τέλος του μαθήματος. Δηλαδή, οι υψηλές τιμές που επιτυγχάνονται στο τέλος του μαθήματος δεν επιτυγχάνονται κατά τις προηγούμενες, ενδιάμεσες εβδομάδες. Το φαινόμενο είναι αναμενόμενο και ευεξήγητο, δεδομένου ότι στις ενδιάμεσες εβδομάδες υφίσταται έλλειμμα πληροφοριών σε σχέση την τελευταία εβδομάδα. Το έλλειμμα πληροφοριών είναι τόσο μεγαλύτερο, όσο μελετούμε σύνολα δεδομένων κοντά στην έναρξη του μαθήματος και δυσχεραίνει την «εκμάθηση» (training) των μοντέλων. Κατά συνέπεια, η προβλεπτική ικανότητα των μοντέλων οξύνεται, όταν κινούμαστε χρονικά προς το τέλος του μαθήματος και αντιστρόφως αμβλύνεται, όταν κινούμαστε προς την αρχή του μαθήματος. Σε αρκετές περιπτώσεις, πολύ πριν την ολοκλήρωση των μαθημάτων, τα μοντέλα επιτυγχάνουν τιμές ακρίβειας πρόβλεψης που είναι μεγαλύτερες από την τιμή βάσης. Ωστόσο, η υπέρβαση της τιμής βάσης, θεωρούμε ότι, δεν αρκεί για να επιτευχθεί πρώιμη πρόβλεψη. Η διαφορά πρέπει επιπλέον να είναι στατιστικώς σημαντική για όλη τη χρονική περίοδο, από την ενδιάμεση εβδομάδα ως το τέλος του μαθήματος.

Στον Πίνακα 51 παρουσιάζεται, ανά μάθημα, η χρονική στιγμή της πρωιμότερης πρόβλεψης και η τιμή ακρίβειας που επιτεύχθηκε. Από τις στατιστικώς σημαντικές τιμές στην ακρίβεια πρόβλεψης, που επιτυγχάνονται α) στο μέσο των μαθημάτων AA005 2020-2021 και AA005 2021-2022 και β) σχεδόν στην αρχή των μαθημάτων A3 2019-2022 & 2020-2021, προκύπτει ότι η πρώιμη πρόβλεψη της επίδοσης των φοιτητών είναι εφικτή.

**Πίνακας 51. Εβδομάδα, τιμή ακρίβειας, AUC, Cohen's Kappa πρωιμότερης πρόβλεψης ανά μάθημα**

	AA005 2020-2021	AA005 2021-2022	A3 2019-2020 & 2021-2022	
Εβδομάδα	7 <sup>η</sup>	9 <sup>η</sup>	3 <sup>η</sup>	9 <sup>η</sup>
Τιμή ακρίβειας πρωιμότερης πρόβλεψης	79.33%	84.11%	65.80%	80.88%
Τιμή βάσης	72.58%	79.36%	53.93%	53.93%
Στατιστική σημαντικότητα (p) τιμής πρόβλεψης από την τιμή βάσης	<0.05	<0.05	<0.05	
AUC	0.77	0.80	0.66	0.8
Cohen's Kappa	0.32	0.42	0.31	0.61

Ως προς την αξιοπιστία της πρόβλεψης, διαπιστώσαμε ότι η τιμή ακρίβειας έχει στατιστικώς σημαντική διαφορά από την τιμή βάσης (με ποσοστό στατιστικής σημαντικότητας 0.05). Επιπλέον, υπολογίσαμε την τιμή AUC και την τιμή Cohen's Kappa, που αντιστοιχούν στα μοντέλα που επιτυγχάνουν την πρωιμότερη πρόβλεψη (Πίνακας 51). Η τιμή της μεταβλητής AUC είναι 0.77 για το μάθημα AA005 2020-2021 οπότε κρίνεται καλή, είναι 0.8 για το μάθημα AA005 2021-2022, οπότε κρίνεται εξαιρετική και είναι 0.66 για τα μαθήματα A3 2019-2020 & 2021-2022, οπότε κρίνεται αποδεκτή (Riestra-Gonzalez, del Puerto Paule-Ruiz & Ortin, 2021, σ. 3). Ως προς την μεταβλητή Cohen's Kappa, στα μαθήματα AA005 2020-2021 και A3 2019-2020 & 2021-2022, οι τιμές είναι αντιστοίχως 0.32 και 0.31 οπότε δηλώνουν «δίκαιη» (fair) συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που προέβλεψε το μοντέλο. Στο μάθημα A005 2021-2022 η τιμή είναι 0.42, οπότε δηλώνει «μέση» (moderate) συμφωνία (Quinn & Gray, 2020, σ. 8). Αξίζει να παρατηρήσουμε ότι στα μαθήματα A3 2019-2020 & 2020-2021, η εντυπωσιακά πρώιμη πρόβλεψη κατά την 3<sup>η</sup> εβδομάδα είναι κυρίως αποτέλεσμα της χαμηλής τιμή βάσης. Αν μεταφέρουμε το σημείο της προσοχής μας από την 3<sup>η</sup> εβδομάδα στο μέσο της χρονικής διάρκειας των μαθημάτων, δηλαδή στην 9<sup>η</sup> εβδομάδα (για λόγους σύγκρισης προστέθηκε αντίστοιχη στήλη στον Πίνακα 51, παρόλο που η πρωιμότερη πρόβλεψη δεν επιτυγχάνεται την 9<sup>η</sup> εβδομάδα), τότε παρατηρούμε ότι η τιμή ακρίβειας πρόβλεψης είναι 80.88%, η τιμή AUC είναι 0.8 (εξαιρετική) και η τιμή Cohen's Kappa είναι 0.61 (σημαντική).

#### *Απάντηση στο Ερευνητικό Ερώτημα EEP2*

Συνεπώς, από την ανάλυση των ερευνητικών αποτελεσμάτων που αφορούν τα μαθήματα που μελετήθηκαν, προκύπτει ότι είναι εφικτή η πρώιμη πρόβλεψη της επίδοσης των φοιτητών με τη βοήθεια μοντέλων Μηχανικής Μάθησης, τα οποία βασίζονται αποκλειστικά σε δεδομένα χρήσης του ΣΔΜ.

Η ακρίβεια και η αξιοπιστία της πρόβλεψης φθίνει όσο κινούμαστε χρονικά αντίθετα στο βέλος του χρόνου, δηλαδή από το τέλος προς την έναρξη του μαθήματος. Στο μέσο της χρονικής διάρκειας του μαθήματος, η τιμή ακρίβειας πρόβλεψης, προσεγγίζει, ή υπερβαίνει (ανάλογα με το μάθημα), το 80% και η αξιοπιστία της πρόβλεψης είναι ισχυρή. Η πρόβλεψη που

επιτυγχάνεται σε πρωιμότερη περίοδο (πριν το μέσο του μαθήματος) παρουσιάζει χαμηλότερη ακρίβεια και αξιοπιστία.

### 11.3. Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ3

#### *Ερευνητικό Ερώτημα ΕΕΡ3:*

Είναι τα μοντέλα που προκύπτουν από τεχνικές «βαθιάς μάθησης» καλύτερα από τα μοντέλα που προκύπτουν από τους βασικούς αλγόριθμους Μηχανικής Μάθησης;

Για καθένα από τα τέσσερα μαθήματα, εκτός από βασικούς αλγόριθμους Μηχανικής Μάθησης, μελετήσαμε και αλγόριθμους που βασίζονται σε τεχνικές «βαθιάς μάθησης» («deep learning»), δηλαδή αλγόριθμους RNN δικτύων (Recurrent Neural Networks / Περιοδικά Νευρωνικά Δίκτυα) και ειδικότερα LSTM δικτύων (Μακρά Βραχείας Μνήμης Δίκτυα / Long Short-Term Memory Networks) (Okubo, Yamashita, Shimada, & Ogata, 2017), (Chen & Cui, 2020). Αυτοί οι αλγόριθμοι αναλύουν, όχι μόνο τις συνολικές, τελικές τιμές των χαρακτηριστικών του συνόλου δεδομένων, αλλά και τις χρονικές ακολουθίες, οι οποίες αποτυπώνουν την εξέλιξη των τιμών. Οπότε, δημιουργήσαμε σύνολα δεδομένων με πιο σύνθετη μορφή, η οποία είναι ικανή να αποτυπώσει το γεγονός ότι οι τιμές των χαρακτηριστικών σωρεύονται και μεταβάλλονται με την πάροδο του χρόνου και εκφράζονται ως χρονοσειρές.

Εφαρμόσαμε τον LSTM αλγόριθμο στα σύνθετα σύνολα δεδομένων όλων των μαθημάτων για κάθε εβδομάδα του μαθήματος, ώστε να μελετήσουμε την μεταβολή των τιμών της ακρίβειας πρόβλεψης. Σε όλα τα σύνολα δεδομένων και για όλα τα μαθήματα, το μοντέλο που κατασκευάστηκε από τον LSTM αλγόριθμο, όταν αυτός εφαρμόσθηκε με την προκαθορισμένη (default) παραμετροποίηση, έδωσε τιμές ίδιες ακριβώς με την τιμή βάσης, δηλαδή απέτυχε να δώσει πρόβλεψη.

Προκειμένου να επιτύχουμε τιμές που θα υπερβαίνουν την τιμή βάσης, επιχειρήσαμε αλλαγές στις προκαθορισμένες (default) παραμέτρους του αλγόριθμου, οπότε κατασκευάστηκε νέο βελτιστοποιημένο μοντέλο (βλ. Παράρτημα Β), το οποίο για το μάθημα AA005 2020-2021, πέτυχε ακρίβεια πρόβλεψης, που υπερβαίνει την τιμή βάσης με στατιστικώς σημαντική διαφορά. Ειδικότερα, πέτυχε συνεχόμενες στατιστικώς σημαντικές τιμές ακρίβειας πρόβλεψης από την 6<sup>η</sup> έως τη 18<sup>η</sup> εβδομάδα, δηλαδή μια εβδομάδα νωρίτερα από την πρωιμότερη πρόβλεψη που πέτυχαν τα μοντέλα των βασικών αλγόριθμων. Όμως, η τιμή της ακρίβειας

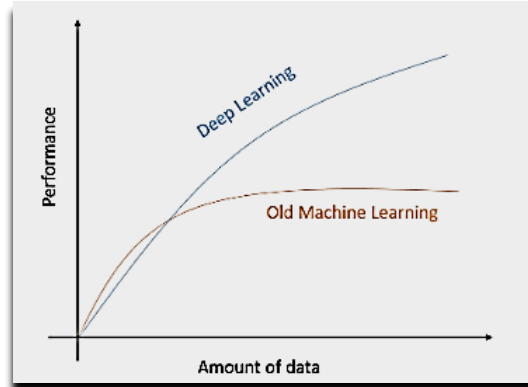


πρόβλεψης τις επόμενες εβδομάδες ήταν μικρότερη από αυτή των μοντέλων των βασικών αλγόριθμων. Για παράδειγμα, τη 18<sup>η</sup> εβδομάδα (κατά την ολοκλήρωση του μαθήματος), η ακρίβεια πρόβλεψης είναι 72.97 %, έναντι 85.80% που επιτεύχθηκε από μοντέλο βασικού αλγόριθμου. Ως προς τα άλλα κριτήρια απόδοσης, κατά τη 18<sup>η</sup> εβδομάδα η τιμή AUC είναι 0.55 οπότε κρίνεται αποδεκτή και η τιμή Cohen's Kappa είναι 0.13, γεγονός που δηλώνει «αδύναμη» (slight) συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που προέβλεψε το μοντέλο (Quinn & Gray, 2020, σ. 8). Αντιστοίχως, την 6<sup>η</sup> εβδομάδα, δηλαδή τη χρονική στιγμή της πρωιμότερης πρόβλεψης, η τιμή AUC είναι 0.59 (αποδεκτή τιμή) και η τιμή Cohen's Kappa είναι 0.22 («δίκαιη» συμφωνία μεταξύ των πραγματικών κλάσεων και των κλάσεων που προέβλεψε το μοντέλο). Στο μάθημα AA005 2021-2022 και στα μαθήματα A3 2019-2020 & 2021-2022, το ίδιο βελτιστοποιημένο μοντέλο δεν πέτυχε τιμή ακρίβειας μεγαλύτερη από την τιμή βάσης. Επιπλέον, η τιμή της μεταβλητής AUC ελάχιστα υπερβαίνει την τιμή βάσης και η τιμή μεταβλητής Cohen's Kappa λαμβάνει σχεδόν μηδενικές τιμές.

Συνεπώς, τα μοντέλα που βασίζονται σε τεχνικές «βαθιάς μάθησης», εφαρμοζόμενα στα σύνολα δεδομένων των μαθημάτων που μελετήσαμε, με την παραμετροποίηση που επιλέξαμε, δεν εμφανίζουν καλύτερη απόδοση από τα μοντέλα των βασικών αλγόριθμων. Το συγκεκριμένο πόρισμα, δεν αποκλίνει από αντίστοιχα πορίσματα που παρουσιάζονται σε βιβλιογραφικές πηγές (Doleck, Lemay, Basnet & Bazalais, 2020). Γενικά, οι έρευνες σχετικά με την αποδοτικότητα των τεχνικών «βαθιάς μάθησης» δεν συγκλίνουν προς κοινώς αποδεκτά συμπεράσματα και δεν αναδεικνύουν με σαφήνεια αν οι τεχνικές «βαθιάς μάθησης» αποδίδουν καλύτερα συγκριτικά με τις βασικές τεχνικές Μηχανικής Μάθησης (Chen & Cui, 2020).

Ωστόσο, πρέπει να επισημάνουμε ότι σύμφωνα με τους Alom, et al. (2019, σ. 7) οι τεχνικές «βαθιάς μάθησης» (επομένως και τα LSTM μοντέλα) υπερέχουν συνήθως σε απόδοση από τις βασικές τεχνικές Μηχανικής Μάθησης, όταν η ποσότητα των δεδομένων είναι πολύ μεγάλη (Γράφημα 13). Επίσης, σύμφωνα με τον Chollet (2017, σ. 13), οι τεχνικές «βαθιάς μάθησης» (επομένως και τα LSTM μοντέλα) υπερέχουν σε απόδοση από τις βασικές τεχνικές Μηχανικής Μάθησης, όταν έχουν μεγάλο «βάθος», δηλαδή περιλαμβάνουν πολλά επίπεδα αναπαράστασης. Συνεπώς, στην περίπτωση των μαθημάτων που μελετήσαμε, όπου αφενός το πλήθος των οντοτήτων είναι μικρό και αφετέρου τα LSTM δίκτυα που χρησιμοποιήσαμε ήταν

«αβαθή», διότι περιλάμβαναν δύο μόνο επίπεδα αναπαράστασης, είναι μάλλον αναμενόμενο να υπερέχουν οι βασικές τεχνικές Μηχανικής Μάθησης.



*Γράφημα 13. Απόδοση τεχνικών «βαθιάς μάθησης» σε σχέση με την ποσότητα των δεδομένων (Alom, et al., 2019, σ. 7)*

Εκτός από την αδυναμία των LSTM μοντέλων να επιτύχουν καλύτερη απόδοση από αυτή των βασικών μοντέλων, πρέπει να επισημανθούν τα εξής:

- Είναι δύσκολη η συγκρότηση των ανάλογων συνόλων δεδομένων, εξαιτίας της πολύπλοκης (τριδιάστατης) δομής τους. Για κάθε ένα χαρακτηριστικό πρέπει να σχηματισθούν αριθμητικές ακολουθίες με τόσους όρους, όσες είναι οι εβδομάδες που έχουν παρέλθει από την έναρξη του μαθήματος.
- Είναι δύσκολη η αναζήτηση των βέλτιστων τιμών για τις παραμέτρους του μοντέλου, δεδομένου ότι οι προκαθορισμένες (default) παράμετροι δεν παράγουν γόνιμα αποτελέσματα. Η αναζήτηση των βέλτιστων τιμών για τις παραμέτρους, συνήθως βασίζεται σε εμπειρικούς κανόνες, που υποδεικνύονται από τη σχετική βιβλιογραφία, σύμφωνα με την οποία η εύρεση ενός αποδοτικού μοντέλου αποτελεί «άσκηση δοκιμής και λάθους» (an exercise in trial-and-error) (Frank, 2021).
- Είναι πρακτικά αδύνατο να εξαντληθούν όλοι οι συνδυασμοί τιμών για τις παραμέτρους, οπότε δεν μπορεί να αποκλειστεί το ενδεχόμενο να υφίσταται μοντέλο παραμετροποιημένο με τιμές, που δεν ελέγξαμε (διότι δεν υποδεικνύονται από τους κανόνες που ακολουθήσαμε), το οποίο να αποδειχθεί αποδοτικότερο από αυτό που κατασκευάσαμε εμείς.

- Κατά τη διαδικασία των δοκιμών με τους διαφορετικούς συνδυασμούς στις τιμές των παραμέτρων, ήταν εμφανείς η αστάθεια και η ευαισθησία των μοντέλων, καθώς μια ελάχιστη αλλαγή τιμής σε μια από τις παραμέτρους, διαπιστώναμε ότι, είχε ως αποτέλεσμα μεγάλη αλλαγή στην απόδοση των μοντέλων.
- Η εγγενής πολυπλοκότητα των LSTM μοντέλων καθιστά εξαιρετικά δυσχερή την κατανόηση της λειτουργίας τους και γενικότερα της λογικής που διέπει τις προβλέψεις τους. Κατά συνέπεια, είναι αναπόφευκτη η δυσπιστία και ο δισταγμός στην αξιοποίησή τους, δεδομένου ότι βασική προϋπόθεση για την αξιοποίηση κάποιου μοντέλου είναι η εμπιστοσύνη στις προβλέψεις του, η οποία είναι ευθέως ανάλογη με την ερμηνευσιμότητά του.
- Η εκτέλεση των αλγόριθμων είναι χρονοβόρα. Συγκεκριμένα, η κατασκευή των μοντέλων με τον LSTM αλγόριθμο διαρκούσε δεκάδες λεπτά, έναντι ελάχιστων δευτερολέπτων που διαρκούσε η κατασκευή μοντέλων με τους βασικούς αλγόριθμους.
- Τα σύνολα δεδομένων των LSTM αλγόριθμων παρότι συνθετότερα ως προς την δομή τους, περιλαμβάνουν λιγότερα χαρακτηριστικά σε σύγκριση με τα σύνολα δεδομένων των βασικών αλγόριθμων. Ο λόγος είναι ότι στα συνθετότερα σύνολα αποτυπώνεται, όχι η συνολική τελική τιμή κάθε χαρακτηριστικού, αλλά (με μορφή αριθμητικής ακολουθίας) οι χρονικά διαδοχικές και αύξουσες τιμές του χαρακτηριστικού. Όμως, αρκετά χαρακτηριστικά (βλ. Κεφ. 9.5.6.) δεν είναι συγκρίσιμα μεταξύ διαφορετικών χρονικών περιόδων, οπότε δεν έχει νόημα να εκφρασθούν ως αύξουσες αριθμητικές ακολουθίες. Οπότε, το πλεονέκτημα που προκύπτει από το γεγονός, ότι οι LSTM αλγόριθμοι έχουν στη διάθεσή τους προς ανάλυση περισσότερη πληροφορία ως προς τη διάσταση τη χρονικής εξέλιξης των τιμών, αντισταθμίζεται από το μειονέκτημα της μειωμένης πληροφορίας ως προς μικρότερο πλήθος των χαρακτηριστικών.

#### *Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ3*

Συνεπώς, από την ανάλυση των ερευνητικών αποτελεσμάτων που αφορούν τα μαθήματα που μελετήθηκαν, προκύπτει ότι τα μοντέλα που κατασκευάζονται με τεχνικές «βαθιάς μάθησης» και τα οποία παραμετροποιήθηκαν με τιμές παραμέτρων,

τις οποίες κατόπιν δοκιμών επιλέξαμε ως βέλτιστες, δεν είναι καλύτερα από τα μοντέλα που προκύπτουν από τους βασικούς αλγόριθμους Μηχανικής Μάθησης.

#### 11.4. Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ4

##### *Ερευνητικό Ερώτημα ΕΕΡ4:*

Ποιοι αλγόριθμοι Μηχανικής Μάθησης παράγουν τα μοντέλα με υψηλή ακρίβεια και πρώιμη πρόβλεψη;

Από την ανάλυση των αποτελεσμάτων της έρευνας, προκύπτει ότι για όλα τα μαθήματα, η πλειονότητα των αλγόριθμων παράγει μοντέλα, που όταν εφαρμοσθούν στο πλήρες σύνολο δεδομένων, δηλαδή στο σύνολο δεδομένων που αντιστοιχεί στη χρονική περίοδο από την έναρξη ως την ολοκλήρωση των μαθημάτων, επιτυγχάνουν υψηλή και αξιόπιστη ακρίβεια πρόβλεψης (βλ. ΕΕΡ1). Επιπλέον προκύπτει, ότι υπάρχουν αλγόριθμοι που παράγουν μοντέλα τα οποία επιτυγχάνουν πρώιμη πρόβλεψη, δηλαδή πρόβλεψη με σημαντική τιμή ακρίβειας και αξιοπιστίας, όταν εφαρμοσθούν σε σύνολα δεδομένων που αντιστοιχούν σε χρονική περίοδο, που ως αρχή έχει την αρχή του μαθήματος και ως πέρασ έχει εβδομάδα αρκετά πριν την ολοκλήρωση του μαθήματος (βλ. ΕΕΡ2).

Ωστόσο, μεταξύ των μαθημάτων παρατηρούνται διαφορές ως προς το ποιος αλγόριθμος κατασκεύασε το αποδοτικότερο μοντέλο. Συμβαίνει, το μοντέλο κάποιου αλγόριθμου να επιτυγχάνει υψηλότερη απόδοση σε κάποιο μάθημα, αλλά στα άλλα μαθήματα η επίδοσή του να είναι απλώς καλή. Συμβαίνει, επίσης, κάποιος αλγόριθμός να κατασκευάζει μοντέλο που επιτυγχάνει κορυφαία απόδοση, όταν εφαρμοσθεί στο πλήρες σύνολο δεδομένων κάποιου μαθήματος, αλλά για το ίδιο μάθημα, το μοντέλο του ίδιου αλγόριθμου να μην επιτυγχάνει πρώιμη πρόβλεψη.

Στα πλαίσια της μελέτης των μοντέλων και της αξιολόγησης των σχετικών αλγόριθμων, θεωρούμε ότι εκτός από την ακρίβεια πρόβλεψης, την αξιοπιστία πρόβλεψης, την ικανότητα πρώιμης πρόβλεψης, εξίσου σημαντικό κριτήριο είναι η ικανότητα του αλγόριθμου να παράγει καλά μοντέλα, όχι μόνο σε ένα μάθημα αλλά πολλά μαθήματα.

Στον Πίνακα 52 οι αλγόριθμοι κατατάσσονται ανά μάθημα κατά φθίνουσα τιμή στην ακρίβεια πρόβλεψης. Η κατάταξη αφορά πρόβλεψη κατά την ολοκλήρωση του μαθήματος, αλλά και την πρωιμότερη πρόβλεψη που επιτεύχθηκε (βλ. Κεφ. 10.1.1, 10.2.1, 10.3.1)

Πίνακας 52. Κατάταξη αλγόριθμων ανά μάθημα κατά ακρίβεια πρόβλεψης. Τελική και πρωιμότερη πρόβλεψη.

AA005 2020-2021			AA005 2021-2022			A3 2019-2020 & 2021-2022		
Αλγόριθμος	Ακρίβεια Πρόβλεψης %	Εβδ.	Αλγόριθμος	Ακρίβεια Πρόβλεψης %	Εβδ.	Αλγόριθμος	Ακρίβεια Πρόβλεψης %	Εβδ.
RandomForest	85.8	18	NaiveBayesMultinom.	89.53	17	LWL	79.37	18
SimpleLogistic	84.54	18	RandomForest	89.24	17	DecisionTable	79.36	18
BayesNet	84.29	18	JRip	88.24	17	DecisionStump	79.25	18
DecisionStump	83.98	18	SimpleLogistic	87.75	17	JRip	77.59	18
LWL	83.86	18	REPTree	87.08	17	PART	76.57	18
Logistic	83.35	18	DecisionTable	86.83	17	SimpleLogistic	75.27	18
IBk	82.5	18	MultilayerPerceptron	86.66	17	J48	74.64	18
NaiveBayesMultinom.	82.21	18	LWL	86.26	17	OneR	73.68	18
DecisionTable	82.16	18	DecisionStump	86.15	17	RandomForest	73.09	18
JRip	82.09	18	PART	86	17	REPTree	73.02	18
KStar	81.07	18	HoeffdingTree	85.83	17	RandomTree	70.11	18
MultilayerPerceptron	81.05	18	NaiveBayes	85.59	17	HoeffdingTree	69.73	18
REPTree	80.82	18	BayesNet	85.45	17	NaiveBayes	69.64	18
NaiveBayes	80.63	18	J48	85.43	17	Logistic	69.48	18
HoeffdingTree	80.32	18	Logistic	85.39	17	BayesNet	69.09	18
OneR	80.15	18	KStar	84.77	17	NaiveBayesMultinom.	67.25	18
J48	78.62	18	OneR	83.28	17	MultilayerPerceptron	65.46	18
PART	78.31	18	RandomTree	83.01	17	KStar	58.77	18
RandomTree	76.62	18	IBk	82.94	17	IBk	55.09	18
Πρωιμότερη Πρόβλεψη			Πρωιμότερη Πρόβλεψη			Πρωιμότερη Πρόβλεψη		
Αλγόριθμος	Ακρίβεια Πρόβλεψης %	Εβδ.	Αλγόριθμος	Ακρίβεια Πρόβλεψης %	Εβδ.	Αλγόριθμος	Ακρίβεια Πρόβλεψης %	Εβδ.
NaiveBayesMultinom.	79.33	7	SimpleLogistic	84.11	9	BayesNet	65.23	3
SimpleLogistic	79.19	7				JRip	65.8	3
						SimpleLogistic	65.95	4

Στον Πίνακα 52, αναζητούμε αλγόριθμους, οι οποίοι έχουν υψηλή σειρά κατάταξης σε όλα τα μαθήματα και ταυτόχρονα επιτυγχάνουν πρώτη πρόβλεψη. Ο μόνος τέτοιος αλγόριθμος είναι ο αλγόριθμος **functions.SimpleLogistic**, το μοντέλο του οποίου, στο μάθημα AA005 2020-2021 επιτυγχάνει τη 2<sup>η</sup> καλύτερη τιμή ακρίβειας πρόβλεψης (84.54%), στο μάθημα AA005 2021-2022

επιτυγχάνει την 4<sup>η</sup> καλύτερη τιμή ακρίβειας πρόβλεψης (87.75%) και στα μαθήματα A3 2019-2020 & 2020-2021 επιτυγχάνει την 6<sup>η</sup> καλύτερη τιμή ακρίβειας πρόβλεψης.

Επιπλέον, το μοντέλο του ίδιου αλγόριθμου, είναι μεταξύ των δύο μοντέλων που επιτυγχάνουν την πρωιμότερη πρόβλεψη κατά την 7<sup>η</sup> εβδομάδα του μαθήματος AA005 2020-2021. Είναι το μόνο μοντέλο, που επιτυγχάνει την πρωιμότερη πρόβλεψη κατά την 9<sup>η</sup> εβδομάδα του μαθήματος AA005 2021-2022. Είναι το μοντέλο που επιτυγχάνει πρώιμη πρόβλεψη την 4<sup>η</sup> εβδομάδα, δηλαδή μια μόλις εβδομάδα μετά την πρωιμότερη πρόβλεψη στα μαθήματα A3 2019-2020 & 2020-2021.

Ο αλγόριθμος `functions.SimpleLogistic` βασίζεται στη μέθοδο της «λογιστικής παλινδρόμησης» (logistic regression) και συνιστά τη μία από τις δύο υλοποιήσεις της μεθόδου στο Weka. Ο άλλος αλγόριθμος, που υλοποιεί την ίδια μέθοδο είναι ο `functions.Logistic`. Η βασικότερη διαφορά των δύο υλοποιήσεων έγκειται στο ότι ο αλγόριθμος `functions.SimpleLogistic` έχει ενσωματωμένο μηχανισμό επιλογής χαρακτηριστικών (Frank, 2014) με συνέπεια να αναλύει με μεγαλύτερη αποδοτικότητα τα σύνολα δεδομένων, οπότε επιτυγχάνει υψηλότερες τιμές ακρίβειας πρόβλεψης. Άλλη διαφορά μεταξύ των δύο αλγόριθμων, είναι ότι ο `functions.Logistic` χρησιμοποιεί πολυωνυμική (multinomial) λογιστική παλινδρόμηση (`weka.classifiers.functions Class Logistic, n.d.`) (`weka.classifiers.functions Class SimpleLogistic, n.d.`), άρα είναι καταλληλότερος για σύνολα δεδομένων, όπου η κλάση λαμβάνει περισσότερες από δύο τιμές. Συνεπώς, δεν είναι αρκετά κατάλληλος για τα σύνολα δεδομένων που μελετήσαμε, όπου η κλάση είναι διχοτομημένη, εξ ου και η κατάταξή του `functions.Logistic` σε χαμηλές θέσεις στον Πίνακα 52.

Η λογιστική παλινδρόμηση είναι μια πολυπαραγοντική στατιστική μέθοδος, που υπολογίζει την τιμή μιας εξαρτημένης κατηγορικής μεταβλητής, με βάση τις τιμές ενός συνόλου από συνεχείς ή κατηγορικές ανεξάρτητες μεταβλητές. Στην απλή μορφή της επιλύει προβλήματα δυαδικής ταξινόμησης, ενώ στην πολυωνυμική (multinomial) μορφή της επιλύει προβλήματα πολλαπλής ταξινόμησης (Jayaprakash S. M., Moody, Lauría, Regan & Baron, 2014, σ. 17). Χρησιμοποιείται για την πρόβλεψη της πιθανότητας εμφάνισης ενός γεγονότος. Η πιθανότητα να λάβει η εξαρτημένη κατηγορική μεταβλητή την μία, ή την άλλη τιμή, εκφράζεται ως μια γραμμική συνάρτηση των ανεξάρτητων μεταβλητών. Η δυαδική λογιστική παλινδρόμηση μοντελοποιεί

μια εξίσωση  $f(z)$ ,  $z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$ , όπου  $X$  είναι οι ανεξάρτητες μεταβλητές και όπου  $\beta$  είναι οι συντελεστές παλινδρόμησης, οι οποίοι εκφράζουν το μέγεθος συνεισφοράς των αντίστοιχων μεταβλητών. Το  $f(z)$  λαμβάνει τιμές μεταξύ 0 και 1 και σχετίζεται με την πιθανότητα εμφάνισης μίας από τις δύο δυνητικές εκβάσεις του γεγονότος (Ahuja, 2020. Cui, Chen, Shiri & Fan, 2019, σ. 218. Jayaprakash S. M., Moody, Lauría, Regan & Baron, 2014, σ. 17. Ζήμερας, 2020). Η μοντελοποίηση με τη μορφή εξίσωσης, κάνει τους αλγόριθμους που βασίζονται στη λογιστική παλινδρόμηση εύκολα κατανοητούς, ιδιότητα πολύτιμη, την οποία δεν διαθέτουν πολλοί αλγόριθμοι Μηχανικής Μάθησης (Gunning & Aha, 2019). Η εν λόγω ιδιότητα αποτελεί ένα επιπλέον πλεοκέκτημα του αλγόριθμου `functions.SimpleLogistic`, πέραν αυτών που αφορούν την ακριβή και έγκαιρη πρόβλεψη.

Η ανάδειξη της λογιστικής παλινδρόμησης, μέσω του αλγόριθμου `functions.SimpleLogistic`, ως της επιτυχέστερης προβλεπτικής μεθόδου στα σύνολα δεδομένων που μελετήσαμε, επιβεβαιώνει ανάλογα πορίσματα που καταγράφονται στη βιβλιογραφία, σύμφωνα με τα οποία η λογιστική παλινδρόμηση είναι εξαιρετικά επιτυχής και προτιμάται στα εκπαιδευτικά σύνολα δεδομένων. Συγκεκριμένα, σύμφωνα με τους Jayaprakash S. M., Moody, Lauría, Regan & Baron (2014, σ. 17), η λογιστική παλινδρόμηση, πιθανόν, είναι η πιο δημοφιλής παραμετρική μέθοδος στις περιπτώσεις, όπου η εξαρτημένη μεταβλητή είναι κατηγορική, δηλαδή στις περιπτώσεις ταξινόμησης. Οι Gasevic, Dawson, Rogers & Gasevic (2016, σ. 69) αναφέρουν, ότι η πρόβλεψη της επίδοσης των φοιτητών «γενικά καθορίζεται» (*generally determined*) με την εφαρμογή της λογιστικής παλινδρόμησης. Οι Howard, Meehan & Parnell (2018, σ. 67) περιγράφουν τη λογιστική παλινδρόμηση, ως τη δημοφιλέστερη μέθοδο πρόβλεψης σε εκπαιδευτικά περιβάλλοντα και παραπέμπουν στους Marbouti, Diefes-Dux & Madhavan (2016, σ. 4), οι οποίοι επιλέγουν τη χρήση της λογιστικής παλινδρόμησης, διότι θεωρούν ότι είναι μια αξιόπιστη προβλεπτική μέθοδος, που συχνά χρησιμοποιείται στα εκπαιδευτικά περιβάλλοντα. Οι Cui, Chen, Shiri & Fan (2019, σ. 218) αναφέρουν, ότι για την πρόβλεψη φοιτητών που κινδυνεύουν να αποτύχουν (*at-risk students*), η λογιστική παλινδρόμηση υπερέχει σε απόδοση έναντι των άλλων τεχνικών.



#### Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ4

Συνεπώς, από την ανάλυση των ερευνητικών αποτελεσμάτων που αφορούν τα μαθήματα που μελετήθηκαν, προκύπτει ότι τα μοντέλα που κατασκευάζονται με τη μέθοδο της λογιστικής παλινδρόμησης, η οποία στο λογισμικό WEKA υλοποιείται με τον αλγόριθμο `functions.SimpleLogistic`, υπερέχουν των άλλων μοντέλων, διότι επιτυγχάνουν υψηλή τιμή ακρίβειας πρόβλεψης και πρώιμη πρόβλεψη σε όλα τα μαθήματα που μελετήθηκαν.

## 11.5. Απάντηση στο Ερευνητικό Ερώτημα EEP5

### *Ερευνητικό Ερώτημα EEP5:*

Είναι ικανά τα παραγόμενα μοντέλα να επιτύχουν αξιόλογη ακρίβεια πρόβλεψης, αν δοκιμαστούν σε «μη οικεία» σύνολα δεδομένων, δηλαδή σε σύνολα διαφορετικά από αυτά που χρησιμοποιήθηκαν για τη δημιουργία τους;

Σε όλα μαθήματα που μελετήσαμε, τα μοντέλα των βασικών αλγόριθμων που κατασκευάζονται και εφαρμόζονται στο «οικείο» σύνολο δεδομένων επιτυγχάνουν ακριβή και αξιόπιστη πρόβλεψη, όχι μόνο κατά την ολοκλήρωση του μαθήματος, αλλά και πριν, σε χρονικές περιόδους περίπου στο μέσο του μαθήματος. Ωστόσο, σημαντική ιδιότητα ενός προβλεπτικού μοντέλου είναι να ανταποκρίνεται με επιτυχία, όχι μόνο όταν δοκιμάζεται στο «οικείο» σύνολο δεδομένων, δηλαδή στο σύνολο από όπου προέκυψε, αλλά και όταν δοκιμάζεται σε σύνολα δεδομένων, τα οποία αφορούν άλλα μαθήματα. Διότι, η «εκμάθηση» και η «δοκιμή» στο «οικείο» σύνολο δεδομένων, ακόμα και αν οδηγήσει σε επιτυχημένη πρόβλεψη, δεν έχει μεγάλη πρακτική αξία. Η χρηστική αξία ενός μοντέλου προκύπτει από την ικανότητά του να επιτύχει ακριβή, αξιόπιστη, πρόβλεψη σε δεδομένα άλλων μαθημάτων.

Ωστόσο, η ικανότητα γενίκευσης δεν είναι εύκολα επιτεύξιμη, δεδομένου ότι κάθε αλγόριθμος αναζητά μοτίβα χρήσης (patterns) σε ένα σύνολο δεδομένων, το οποίο αποτελείται από συγκεκριμένα χαρακτηριστικά. Τα μοντέλα δεν μπορούν να λειτουργήσουν, όταν τα χαρακτηριστικά, στα οποία βασίζονται, απουσιάζουν από το «μη οικείο» σύνολο, όπου θα εφαρμοσθούν. Συνεπώς, τα σύνολα δεδομένων (το «οικείο» και το «ξένο») πρέπει να είναι όμοια, ως προς το πλήθος και ως προς το είδος των χαρακτηριστικών και να διαφέρουν μόνο ως προς τις τιμές που λαμβάνουν τα χαρακτηριστικά και ως προς το πλήθος των οντοτήτων. Αν, χάριν της γενικευσιμότητας, επιχειρηθεί κατασκευή συμβατών συνόλων δεδομένων, δηλαδή συνόλων που θα περιλαμβάνουν μόνο τα κοινά χαρακτηριστικά, τότε επέρχεται ποσοτική και ποιοτική υποβάθμιση στην πληροφορία που έχουν στη διάθεσή τους οι αλγόριθμοι για να κατασκευάσουν τα μοντέλα, με συνέπεια την εξασθένηση της προβλεπτικής ικανότητάς τους. Επιπλέον, ακόμα και αν το σύνολο δεδομένων, στο οποίο θα εφαρμοσθεί το μοντέλο, είναι

συμβατό, εξακολουθεί να είναι «μη οικείο» σύνολο, το οποίο αναφέρεται σε άλλο μάθημα με ενδεχομένως διαφορετικό εκπαιδευτικό υλικό, δομή, οργάνωση και στο οποίο φοιτούν άλλοι φοιτητές και διδάσκουν άλλοι διδάσκοντες. Δηλαδή, δεδομένου ότι τα μαθήματα είναι πάντα διαφορετικά, τα σύνολα δεδομένων ακόμα και αν είναι συμβατά διαφέρουν ως προς τις τιμές τους, οπότε οι αστοχίες του μοντέλου είναι αναπόφευκτες. Γενικά, η ικανότητα γενίκευσης και η ικανότητα πρόβλεψης συνιστούν δύο πόλους ενός ισοζυγίου (Gasevic, Dawson, Rogers, & Gasevic, 2016, σ. 83). Οσο μεγαλύτερη είναι η ικανότητα γενίκευσης, τόσο μικρότερη είναι η ικανότητα πρόβλεψης και αντίστροφα.

Από την ανάλυση των αποτελεσμάτων της έρευνας προκύπτει ότι, λόγω ασυμβατότητας στα σύνολα δεδομένων, δεν είναι δυνατόν να ελεγχθεί η ικανότητα γενίκευσης των μοντέλων των μαθημάτων A3 2019-2020 & 2021-2022 επί του μαθήματος AA005 2020-2021 ή επί του μαθήματος AA005 2021-2022. Για τον ίδιο λόγο, δεν είναι δυνατόν να γίνει και το αντίστροφο, δηλαδή να ελεγχθεί η ικανότητα γενίκευσης των μοντέλων του μαθήματος AA005 2020-2021 ή του μαθήματος AA005 2021-2022, επί των μαθημάτων A3 2019-2020 & 2021-2022. Οπότε, ελέγξαμε την ικανότητα γενίκευσης των μοντέλων, που προέκυψαν από τα μαθήματα AA005 2020-2021 και AA005 2021-2022. Ειδικότερα, επιλέξαμε το αποδοτικότερο μοντέλο που προέκυψε από τα δύο μαθήματα, δηλαδή το μοντέλο του αλγόριθμου `functions.SimpleLogistic` από το μάθημα AA005 2021-2022 (βλ. Κεφ. 10.4) και το εφαρμόσαμε στο σύνολο δεδομένων όλων των εβδομάδων του μαθήματος AA005 2020-2021 (Πίνακας 44). Προέκυψε, ότι το μοντέλο διαθέτει ικανότητα γενίκευσης. Η απόδοση του μοντέλου υπολείπεται ελαφρώς από την απόδοση του «οικείου» μοντέλου, αλλά αφενός οι τιμές ακρίβειας και αξιοπιστίας που επιτεύχθηκαν κινούνται σε υψηλά επίπεδα και αφετέρου επιτυγχάνεται πρώιμη πρόβλεψη στα μέσα περίπου του μαθήματος.

#### *Απάντηση στο Ερευνητικό Ερώτημα EEP5*

Συνεπώς, από την ανάλυση των ερευνητικών αποτελεσμάτων που αφορούν τα μαθήματα που μελετήθηκαν, προκύπτει ότι η ακριβής και αξιόπιστη πρόβλεψη από μοντέλα που δοκιμάζονται σε «μη οικεία» σύνολα δεδομένων είναι εφικτή, υπό προϋποθέσεις. Κατασκευάστηκε μοντέλο που πέτυχε πρόβλεψη σε «μη οικείο» σύνολο δεδομένων, γεγονός που αποτελεί ένδειξη ικανότητας γενίκευσης, η οποία όμως πρέπει

να ελεγχθεί περαιτέρω με σχετικές δοκιμές σε πολλά διαφορετικά σύνολα δεδομένων. Προκύπτει επίσης, ότι δεν είναι εφικτή η εφαρμογή μοντέλων σε σύνολα δεδομένων που δεν είναι συμβατά με τα σύνολα από τα οποία κατασκευάστηκαν. Γενικά, η ικανότητα γενίκευσης ενός μοντέλου αντισταθμίζει την προβλεπτική ικανότητά του.

## 11.6. Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ6

### Ερευνητικό Ερώτημα ΕΕΡ6:

Ποια από τα χαρακτηριστικά που συγκροτούν τα σύνολα δεδομένων έχουν τη μεγαλύτερη επίδραση στα αποτελέσματα της πρόβλεψης;

Από την ανάλυση των αποτελεσμάτων της έρευνας προκύπτει ότι τα χαρακτηριστικά που περιλαμβάνονται στα σύνολα δεδομένων δεν έχουν όλα την ίδια βαρύτητα ως προς την απόδοση των μοντέλων. Κάποια χαρακτηριστικά έχουν υψηλό βαθμό συσχέτισης με την κλάση των συνόλων, οπότε επηρεάζουν την απόδοση του μοντέλου, κάποια έχουν μέσο βαθμό συσχέτισης και κάποια έχουν πολύ μικρό βαθμό συσχέτισης, οπότε δεν έχουν καμιά επίδραση στην απόδοση του μοντέλου. Για τον εντοπισμό των επιδραστικών χαρακτηριστικών ακολουθήθηκε η, συχνά αναφερόμενη στην βιβλιογραφία (Brownlee, 2019b. Gnanambal, Thangaraj, Meenatchi & Gayathri, 2018. Hussain, Dahan, Ba-Alwib & Ribata, 2018. Pehlivanova & Nedeva, 2021), μέθοδος «CorrelationAttributeEval» η οποία βασίζεται στον υπολογισμό του συντελεστή συσχέτισης (Pearson  $r$ ) μεταξύ κάθε χαρακτηριστικού και της κλάσης του συνόλου δεδομένων.

Βασικό εύρημα των αποτελεσμάτων είναι τα αξιοσημείωτα επίπεδα τιμών στους συντελεστές συσχέτισης και στη στατιστική σημαντικότητά τους, τα οποία δηλώνουν αντίστοιχο σημαντικό βαθμό επίδρασης των χαρακτηριστικών στην απόδοση του μοντέλου (Πίνακας 47). Ο σημαντικός βαθμός επίδρασης επιβεβαιώθηκε και πρακτικά μετά από σύγκριση της απόδοσης των μοντέλων, όταν για την κατασκευή τους χρησιμοποιηθεί, αφενός πλήρες σύνολο δεδομένων και αφετέρου σύνολο δεδομένων που δεν περιλαμβάνει τα επιδραστικότερα χαρακτηριστικά (Πίνακας 48).

Βασικό εύρημα των αποτελεσμάτων είναι επίσης η ανάδειξη κοινών επιδραστικών χαρακτηριστικών στα σύνολα δεδομένων των συναφών μαθημάτων. Δηλαδή των μαθημάτων ΑΑ005 2020-2021 και ΑΑ005 2021-2022, τα οποία είναι ίδια ως προς την οργάνωση, το υλικό, τον τρόπο αξιολόγησης κ.λ.π. Ειδικότερα, αν, σε κάθε ένα από τα δύο παραπάνω μαθήματα, κατατάξουμε με φθίνουσα σειρά τα πέντε επιδραστικότερα χαρακτηριστικά τους (δηλαδή συνολικά εμφανίσουμε δέκα χαρακτηριστικά), τότε παρατηρούμε ότι οκτώ από τα δέκα

χαρακτηριστικά εμφανίζονται και στα δύο μαθήματα. Επιπλέον, τα δύο από τα δέκα, μη κοινά χαρακτηριστικά, είναι παρόμοια ως προς το φαινόμενο που καταγράφουν. Αντιθέτως, από τα πέντε επιδραστικότερα χαρακτηριστικά των μαθημάτων Α3 2019-2020 & 2021-2022, κανένα δεν ανήκει στις ομάδες των επιδραστικότερων χαρακτηριστικών των μαθημάτων ΑΑ005 2020-2021 και ΑΑ005 2021-2022 (Πίνακας 47).

#### Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ6

Συνεπώς, από την ανάλυση των ερευνητικών αποτελεσμάτων που αφορούν τα μαθήματα που μελετήθηκαν, προκύπτει ότι τα χαρακτηριστικά που απαρτίζουν κάθε σύνολο δεδομένων δεν είναι ισότιμα ως προς την επίδρασή τους στην απόδοση των μοντέλων πρόβλεψης. Κοινά επιδραστικά χαρακτηριστικά αναδεικνύονται μεταξύ συναφών μαθημάτων. Αντιθέτως, τα επιδραστικά χαρακτηριστικά διαφέρουν μεταξύ των μαθημάτων που δεν μοιάζουν.

Τα πέντε επιδραστικότερα χαρακτηριστικά, για κάθε μάθημα που μελετήθηκε, εμφανίζονται στον Πίνακα 53.

**Πίνακας 53. Επιδραστικότερα χαρακτηριστικά για κάθε μάθημα**

Χαρακτηριστικό	ΑΑ005 2020-2021	ΑΑ005 2021-2022	Α3 2019-2020 & 2020-2021
Πλήθος λεπτών αδράνειας κατά τη χρονική περίοδο που ο φοιτητής ήταν ενεργός	✓	✓	
Πλήθος ημερών μεταξύ πρώτης και τελευταίας συνόδου σύνδεσης	✓	✓	
Πλήθος ημερών μεταξύ της τελευταίας συνόδου και του τέλους χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων	✓	✓	
Πλήθος ημερών αδράνειας κατά την διάρκεια της ενεργής περιόδου	✓	✓	
Πλήθος συνόδων τις πρωινές ώρες	✓		
Πλήθος συνόδων τις εργάσιμες ημέρες		✓	
Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο			✓
Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου			✓

Πλήθος ημερών όπου καταγράφηκε δραστηριότητα			✓
Τυπική απόκλιση των ημερών αδράνειας για τις ημέρες αδράνειας του ίδιου φοιτητή			✓
Μέσος όρος πλήθους ημερών αδράνειας μεταξύ των συνόδων σύνδεσης			✓

## 11.7. Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ7

### Ερευνητικό Ερώτημα ΕΕΡ7:

Πώς ερμηνεύεται ο αυξημένος βαθμός επίδρασης κάποιων χαρακτηριστικών που συγκροτούν τα σύνολα δεδομένων; Υπάρχει σχέση των επιδραστικών χαρακτηριστικών με την ενεργή συμμετοχή και τον ρόλο της στην επιτυχημένη φοίτηση;

Η ανάδειξη χαρακτηριστικών που έχουν μεγαλύτερη επίδραση από τα υπόλοιπα στην απόδοση των μοντέλων, ωθεί σε αναζήτηση των λόγων για τους οποίους προκύπτει η συγκεκριμένη επίδραση. Οι στατιστικοί δείκτες που αποδίδουν διαφορετικό βαθμό επίδρασης των χαρακτηριστικών (Πίνακας 47), απλώς περιγράφουν στατιστικές συσχετίσεις, αλλά αδυνατούν να παράσχουν σαφείς εξηγήσεις, ως προς τους λόγους της επίδρασης. Ωστόσο, είναι χρήσιμοι, διότι υποδεικνύουν κατευθύνσεις προς τις οποίες θα μπορούσε να προσανατολισθεί η σχετική διερεύνηση.

Η καταδεικνυόμενη, από τους στατιστικούς δείκτες, επίδραση κάποιου χαρακτηριστικού στην απόδοση του μοντέλου (βλ. ΕΕΡ6), δηλώνει ότι συγκεκριμένα επίπεδα τιμών του χαρακτηριστικού σχετίζονται κατ' επανάληψη με συγκεκριμένη κλάση στο σύνολο δεδομένων. Δηλαδή, συγκεκριμένα επίπεδα των τιμών του χαρακτηριστικού σε κάποια οντότητα αυξάνουν την πιθανότητα, η οντότητα να ανήκει σε συγκεκριμένη κλάση. Δεδομένου, ότι αφενός τα χαρακτηριστικά αποτυπώνουν πραγματικές συμπεριφορές, ως προς τη χρήση του ΣΔΜ και αφετέρου ότι τα μοντέλα καταφέρνουν να προβλέψουν (έστω πιθανοκρατικώς) εκβάσεις της φοίτησης (επιτυχία / αποτυχία), είναι εύλογο να θεωρήσουμε, ότι οι εμφανείς συσχετίσεις μεταξύ χαρακτηριστικών και κλάσεων (π.χ. επιτυχία, μη επιτυχία) υποδηλώνουν συσχετίσεις μεταξύ παραγόντων συμπεριφοράς στη χρήση του ΣΔΜ και επιτυχημένης ή μη επιτυχημένης φοίτησης.

Οι παράγοντες συμπεριφοράς που επηρεάζουν την επιτυχημένη φοίτηση σε ένα περιβάλλον μάθησης, όχι απαραίτητως ψηφιακό, είναι πολλοί, σύνθετοι, εσωτερικοί (αφορούν την ιδιοσυστασία του φοιτητή) και εξωτερικοί (αφορούν το κοινωνικό περιβάλλον του φοιτητή και τις εκπαιδευτικές συνθήκες και παραμέτρους). Η μελέτη όλων αυτών των παραγόντων δεν περιλαμβάνεται στους σκοπούς της διατριβής. Η έρευνα που διενεργήθηκε στα πλαίσια της



διατριβής είχε στη διάθεσή της δεδομένα μόνο για τη χρήση του ΣΔΜ. Συνεπώς, μπορεί να μελετήσει μόνο εκείνους τους παράγοντες συμπεριφοράς που υφίστανται σε ψηφιακά περιβάλλοντα μάθησης και οι οποίοι είναι δυνατόν να αφήσουν αποτύπωμα στα διαθέσιμα δεδομένα χρήσης του ΣΔΜ.

Σε κάθε μάθημα που μελετήσαμε, η χρήση του ΣΔΜ αποτελούσε όχι μόνο αναγκαία, αλλά και κύρια συνθήκη «ενεργής συμμετοχής» στο μάθημα. Συγκεκριμένα, η χρήση του ΣΔΜ ήταν κύρια συνθήκη για την ενεργή συμμετοχή στο μάθημα ΑΑ005, διότι κατά το έτος 2020-2021 λόγω των περιοριστικών μέτρων της πανδημίας Covid-19, το μάθημα προσφέρθηκε με διαδικτυακή διδασκαλία, ενώ το έτος 2021-2022, όταν αποκαταστάθηκε η κανονική λειτουργία των εκπαιδευτικών ιδρυμάτων, το μάθημα προσφέρθηκε με Τεχνολογικά Ενισχυμένη Πρόσωπο με Πρόσωπο Διδασκαλία και Μάθηση. Το εκπαιδευτικό υλικό του μαθήματος ήταν διαθέσιμο μόνο μέσω του ΣΔΜ. Η χρήση του ΣΔΜ ήταν κύρια συνθήκη για την ενεργή συμμετοχή στο μάθημα Α3, διότι κατά το έτος 2019-2020, το μάθημα προσφέρθηκε με μεικτή μάθηση, αλλά η πλειονότητα των συνεδριών (8 από τις 10) πραγματοποιήθηκαν ασύγχρονα. Το έτος 2020-2021, λόγω των περιοριστικών μέτρων της πανδημίας Covid-19, το μάθημα προσφέρθηκε με διαδικτυακή διδασκαλία. Το εκπαιδευτικό υλικό των μαθημάτων ήταν διαθέσιμο μόνο μέσω του ΣΔΜ και επιπλέον υπήρχε υποχρεωτική συμμετοχή όλων των φοιτητών σε δημόσιους διάλογους που διεξάγονταν στο μάθημα Α3 2019-2020 & 2021-2022. Επομένως, σε όλα τα μαθήματα που μελετάμε, το ΣΔΜ αποτέλεσε το κυριότερο μέσο ενεργής συμμετοχής στο μάθημα. Συνεπώς, θεωρούμε ότι τα δεδομένα των αρχείων καταγραφής, από τα οποία κατασκευάστηκαν τα σύνολα δεδομένων που μελετάμε, αποτυπώνουν κατά κύριο λόγο την ενεργή συμμετοχή των φοιτητών. Οπότε, στα πλαίσια της έρευνάς μας, λόγω του είδους των δεδομένων που έχουμε στη διάθεσή μας, ο κύριος παράγοντας συμπεριφοράς που μπορούμε να μελετήσουμε είναι η «ενεργή συμμετοχή».

Ας σημειωθεί επίσης ότι τα δεδομένα που συλλέξαμε κατά την έρευνά μας δεν μπορούν να αποτυπώσουν όλες τις διαστάσεις της ενεργής συμμετοχής (βλ. Κεφ. 5.4). Ειδικότερα, μπορούν να αποτυπώσουν τη συμπεριφοριστική (behavioral) διάσταση, όχι όμως τη γνωστική (cognitive) και τη συναισθηματική (emotional) διάσταση (Fredricks, Blumenfeld & Paris, 2004). Διότι, τα δεδομένα αφορούν μόνο χρονοσημασμένες ενέργειες πρόσβασης σε πόρους και σε υπηρεσίες

του ΣΔΜ. Δεν συλλέξαμε, για παράδειγμα, δεδομένα από ερωτηματολόγια, τα οποία θα μπορούσαν να καταγράψουν απόψεις και στάσεις των φοιτητών. Οπότε, εκ των πραγμάτων, περιοριζόμαστε στην καταγραφή και ανάλυση ποσοτικών κυρίως μεγεθών, όπως ο χρόνος και το πλήθος των επισκέψεων σε ψηφιακούς πόρους και σε ψηφιακές υπηρεσίες, με συνέπεια τη μονομερή μελέτη μόνο της συμπεριφοριστικής διάστασης της ενεργής συμμετοχής στο ψηφιακό περιβάλλον μάθησης που ορίζεται από το ΣΔΜ (Vytasek, Patzak & Winne, 2020, σ. 33).

Επιπλέον, όπως έχουμε αναφέρει (βλ. Κεφ. 5.4) η μέτρηση της ενεργής συμμετοχής αποτελεί εξαιρετικά σύνθετο ζήτημα στα ψηφιακά περιβάλλοντα μάθησης. Η ερευνητική κοινότητα, ενώ καταλήγει σε πορίσματα που τεκμηριώνουν συσχετίσεις μεταξύ δεδομένων χρήσης και ακαδημαϊκής επίδοσης, συχνά εκφράζει σκεπτικισμό ως προς την τεκμηρίωση στέρεων συσχετίσεων μεταξύ δεδομένων χρήσης και «ενεργής συμμετοχής» (Motz, Quick, Schroeder, Zook & Gunkel, 2019, σ. 300, 307), κυρίως διότι η ακαδημαϊκή επίδοση είναι μετρήσιμη, ενώ η ενεργή συμμετοχή εμπεριέχει ψυχολογικές, γνωστικές, συναισθηματικές παραμέτρους, οι οποίες είναι δύσκολο να ανιχνευθούν, πόσο δε μάλλον να μετρηθούν. Όμως, ο εύλογος σκεπτικισμός και οι επιφυλάξεις, επ' ουδενί δεν πρέπει να οδηγήσουν σε «ερευνητική παράλυση», δηλαδή στην παραίτηση από κάθε προσπάθεια σχετικής έρευνας, εξαιτίας των δυσχερειών, που ενέχει η μελέτη ενός εννοιολογικά ρευστού αντικειμένου. Οπότε, επιχειρούμε, εντός του ερευνητικού πεδίου των μαθημάτων που μελετάμε, να αναζητήσουμε ευρήματα, που θα συμβάλουν στον ερευνητικό διάλογο. Στην αναζήτησή μας, αξιοποιούμε δεδομένα χρήσης του ΣΔΜ, γεγονός που παρέχει το πλεονέκτημα της, έστω ατελούς, ποσοτικής αναπαράστασης σύνθετων διεργασιών. Ωστόσο, λαμβάνουμε υπόψιν, ότι αναπόφευκτα η συγκεκριμένη προσέγγιση δεν είναι πλήρης, διότι περιορίζεται στη μελέτη μόνο της συμπεριφορικής διάστασης της ενεργής συμμετοχής.

Οι δείκτες ενεργής συμμετοχής στα ψηφιακά περιβάλλοντα μάθησης τους οποίους προτείνει η βιβλιογραφία (βλ. Κεφ. 5.4, Πίνακας 6), περιλαμβάνονται στον παρακάτω Πίνακα 54.

*Πίνακας 54. Δείκτες ενεργής συμμετοχής*

A.A.	Δείκτης ενεργής συμμετοχής
1	Διάρκεια σύνδεσης

2	Πλήθος πόρων ή δραστηριοτήτων ή εργασιών
3	Αλληλεπιδράσεις
4	Μαθησιακά αποτελέσματα
5	Ποσοστό ολοκλήρωσης εργασιών
6	Αμεσότητα σύνδεσης
7	Συχνότητα σύνδεσης
8	Κανονικότητα
9	Σύνοδοι σύνδεσης
10	Διάρκεια Αδράνειας
11	Ασυνέπεια

Παρατηρούμε, ότι αρκετοί από τους δείκτες ενεργής συμμετοχής (Πίνακας 54) σχετίζονται με επιδραστικά χαρακτηριστικά των συνόλων δεδομένων (βλ. Πίνακα 47). Για παράδειγμα, στα μαθήματα ΑΑ005 2020-2021 και ΑΑ005 2021-2022, το επιδραστικό χαρακτηριστικό «Πλήθος λεπτών αδράνειας κατά τη χρονική περίοδο που ο φοιτητής ήταν ενεργός», σχετίζεται με τον δείκτη ενεργής συμμετοχής «Διάρκεια αδράνειας» (Α.Α. 10), διότι καταγράφει τη χρονική περίοδο που ο φοιτητής δεν χρησιμοποιούσε το σύστημα, δηλαδή αδρανούσε. Σχετίζεται επίσης με τον δείκτη ενεργής συμμετοχής «Ασυνέπεια» (Α.Α. 11), διότι ο φοιτητής που αδρανεύει ως προς τη χρήση του συστήματος είναι ασυνεπής στις υποχρεώσεις ενός μαθήματος, το οποίο κυρίως προσφέρεται μέσω του ΣΔΜ. Επίσης στα μαθήματα Α3 2019-2020 & 2020-2021, τα χαρακτηριστικά «Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο» και «Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου», σχετίζονται με τον δείκτη ενεργής συμμετοχής «Αλληλεπιδράσεις» (Α.Α 3), διότι καταγράφουν την αλληλεπίδραση των φοιτητών με άλλους φοιτητές μέσω των διαδικτυακών δημόσιων διαλόγων του μαθήματος, την οποία θεωρούμε ως την πιο απτή ένδειξη ενεργής συμμετοχής σε μάθημα. Στον Πίνακα 55 καταγράφονται οι σχέσεις των επιδραστικών χαρακτηριστικών όλων των μαθημάτων (βλ. Κεφ. 10.5, Πίνακα 47) και των δεικτών ενεργής συμμετοχής.

**Πίνακας 55. Σχέσεις επιδραστικών χαρακτηριστικών και δεικτών ενεργής συμμετοχής**

Επιδραστικό Χαρακτηριστικό	Α.Α. Δείκτη Ενεργής Συμμετοχής (από Πίνακα 54)										
	1	2	3	4	5	6	7	8	9	10	11
Πλήθος λεπτών αδράνειας κατά τη χρονική περίοδο που ο φοιτητής ήταν ενεργός										√	√
Πλήθος ημερών μεταξύ πρώτης και τελευταίας συνόδου σύνδεσης	√										√
Πλήθος ημερών μεταξύ της τελευταίας συνόδου και του τέλους χρονικής περιόδου στην οποία αντιστοιχεί το σύνολο δεδομένων	√							√			√
Πλήθος ημερών αδράνειας κατά την διάρκεια της ενεργής περιόδου								√		√	√
Πλήθος συνόδων τις πρωινές ώρες									√		
Πλήθος συνόδων τις εργάσιμες ημέρες									√		
Βαθμός αξιολόγησης για τη συμμετοχή σε δημόσιο διάλογο			√								
Πλήθος συμβάντων ανάρτησης στους χώρους δημόσιου διαλόγου		√	√								
Πλήθος ημερών όπου καταγράφηκε δραστηριότητα	√										
Τυπική απόκλιση των ημερών αδράνειας για τις ημέρες αδράνειας του ίδιου φοιτητή								√		√	√
Μέσος όρος πλήθους ημερών αδράνειας μεταξύ των συνόδων σύνδεσης								√		√	

Παρατηρούμε στον Πίνακα 55 ότι όλα ανεξαιρέτως τα επιδραστικά χαρακτηριστικά, όπως προέκυψαν από την ανάλυση των ερευνητικών δεδομένων (βλ. ΕΕΡ6, Πίνακα 53), σχετίζονται με έναν τουλάχιστον δείκτη ενεργής συμμετοχής, γεγονός που θεωρούμε ότι αποτελεί εξήγηση για τους λόγους που τα συγκεκριμένα χαρακτηριστικά αναδείχθηκαν επιδραστικότερα άλλων. Δηλαδή, εφόσον η ενεργή συμμετοχή προσιωνίζεται επιτυχή φοίτηση (Sinatra, Heddy & Lombardi, 2015, σ. 1), τότε είναι εύλογο να θεωρήσουμε ότι τα χαρακτηριστικά που υποδηλώνουν ενεργή συμμετοχή, αποτελούν την κύρια πηγή πληροφορίας βάση της οποίας τα μοντέλα προβλέπουν την έκβαση της επιτυχημένης ή μη φοίτησης.

Συνεπώς, από την ανάλυση των ερευνητικών αποτελεσμάτων που αφορούν τα μαθήματα που μελετήθηκαν, προκύπτει ότι ο αυξημένος βαθμός επίδρασης κάποιων χαρακτηριστικών μπορεί να αποδοθεί στη σχέση τους με δείκτες ενεργής συμμετοχής. Δεδομένης της συμβολής της ενεργής συμμετοχής στην επιτυχημένη φοίτηση, τα επιδραστικά χαρακτηριστικά, ως έκφραση ενεργής συμμετοχής, αποτελούν βασική πηγή πληροφορίας για τα μοντέλα που προβλέπουν την έκβαση της φοίτησης.

## 11.8. Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ8

### Ερευνητικό Ερώτημα ΕΕΡ8:

Επηρεάζεται η απόδοση των μοντέλων πρόβλεψης από ζητήματα που αφορούν δομικά χαρακτηριστικά του μαθήματος και γενικότερα από ζητήματα που αφορούν τον εκπαιδευτικό σχεδιασμό;

Από την έρευνα που πραγματοποιήσαμε προέκυψε καταρχάς ότι είναι δυνατή η πρόβλεψη της επίδοσης των φοιτητών με τη βοήθεια μοντέλων Μηχανικής Μάθησης, τα οποία βασίζονται αποκλειστικά σε δεδομένα χρήσης του ΣΔΜ (βλ. ΕΕΡ1). Επιπλέον, προέκυψε ότι μεταξύ των μοντέλων που κατασκευάστηκαν για κάθε μάθημα υφίστανται αρκετές διαφορές (βλ. Κεφ. 10.1, 10.2, 10.3). Συγκεκριμένα:

- Υπάρχουν διαφορές ως προς την ακρίβεια πρόβλεψης. Επί του πλήρους συνόλου δεδομένων, το οποίο αντιστοιχεί στη χρονική περίοδο από την έναρξη ως το τέλος των μαθημάτων, η μέγιστη ακρίβεια πρόβλεψης είναι 85.80% στο μάθημα AA005 2020-2021, είναι 89.53% στο μάθημα AA005 2021-2022 και είναι 89.53% στα μαθήματα A3 2019-2020 & 2021-2022.
- Υπάρχουν διαφορές ως προς τη χρονική στιγμή και ως προς την ακρίβεια της πρώιμης στατιστικώς σημαντικής πρόβλεψης. Η πρώιμη στατιστικώς σημαντική πρόβλεψη επιτυγχάνεται στο μάθημα AA005 2020-2021 την 7<sup>η</sup> εβδομάδα με τιμή ακρίβειας, 79.33%, στο μάθημα AA005 2021-2022 την 9<sup>η</sup> εβδομάδα με τιμή ακρίβειας 84.11% και στα μαθήματα A3 2019-2020 & 2021-2022 την 3<sup>η</sup> εβδομάδα με τιμή ακρίβειας 65.80%.
- Υπάρχουν διαφορές ως προς τους αλγόριθμους, οι οποίοι κατασκεύασαν τα μοντέλα με την καλύτερη απόδοση. Παρατηρείται ότι το μοντέλο κάποιου αλγόριθμου επιτυγχάνει τη μέγιστη απόδοση σε κάποιο μάθημα, αλλά στα άλλα μαθήματα επιτυγχάνει απλώς ικανοποιητική απόδοση. Επιπλέον, προκύπτει ότι η ικανότητα γενίκευσης του μοντέλου AA005 2021-2022 ελέγχθηκε και επιβεβαιώθηκε στο μάθημα AA005 2020-2021, αλλά δεν μπορεί να ελεγχθεί στα μαθήματα A3 2019-2020 & 2021-2022, λόγω ασυμβατότητας των συνόλων δεδομένων.

- Υπάρχουν διαφορές ως προς τα επιδραστικότερα χαρακτηριστικά των συνόλων δεδομένων. Ενώ, τα επιδραστικότερα χαρακτηριστικά των μαθημάτων AA005 2020-2021 και AA005 2021-2022 είναι σχεδόν τα ίδια, τα μαθήματα A3 2019-2020 & 2021-2022 δεν έχουν κάποιο κοινό επιδραστικό χαρακτηριστικό με τα άλλα μαθήματα.

Βέβαια, τα μαθήματα που μελετάμε δεν είναι ίδια. Συγκεκριμένα, τα μαθήματα A3 2019-2020 & 2021-2022 διαφέρουν από το μάθημα AA005 2020-2021 και το μάθημα AA005 2021-2022, στην εκπαιδευτική βαθμίδα (μεταπτυχιακά μαθήματα), στο γνωστικό αντικείμενο, στο εκπαιδευτικό υλικό, στην οργάνωση των ενοτήτων. Διαφέρουν στο ότι τα μαθήματα A3 2019-2020 & 2021-2022 σχεδιάστηκαν ώστε να προσφερθούν με Μεικτή Διδασκαλία (blended teaching), ενώ το AA005 2020-2021 και το AA005 2021-2022 σχεδιάστηκαν ώστε να προσφερθούν με Τεχνολογικά Ενισχυμένη Πρόσωπο με Πρόσωπο Διδασκαλία και Μάθηση (Technology Enhanced Face to Face Teaching and Learning) (βλ. Κεφ 9.2). Διαφέρουν επίσης στο ότι στα μαθήματα A3 2019-2020 & 2021-2022, η παρακολούθηση ήταν υποχρεωτική, οπότε οι επιμορφούμενοι είχαν ως υποχρέωση να χρησιμοποιούν το ΣΔΜ, ενώ στα AA005 2020-2021 και AA005 2021-2022 η παρακολούθηση ήταν προαιρετική. Διαφέρουν στο ότι στα μαθήματα A3 2019-2020 & 2021-2022, οι φοιτητές όφειλαν να συμμετάσχουν σε δημόσιους ηλεκτρονικούς διαλόγους, ενώ στα AA005 2020-2021 και AA005 2021-2022 δεν υπήρχε αλληλεπίδραση των φοιτητών στο ΣΔΜ. Διαφέρουν στο ότι στα μαθήματα A3 2019-2020 & 2021-2022, οι φοιτητές αξιολογήθηκαν με βάση τη συμμετοχή τους στους δημόσιους διαλόγους και την τελική εργασία που συνέγραψαν, ενώ στα AA005 2020-2021 και AA005 2021-2022, οι φοιτητές αξιολογήθηκαν με τελικές γραπτές εξετάσεις. Τα μαθήματα AA005 2020-2021 και AA005 2021-2022 είναι σχεδόν ίδια. Έχουν το ίδιο αντικείμενο, την ίδια δομή, το ίδιο εκπαιδευτικό υλικό, τον ίδιο τρόπο αξιολόγησης. Διέφεραν στο ότι, λόγω των περιοριστικών μέτρων της πανδημίας Covid-19, οι διαλέξεις του AA005 2020-2021 πραγματοποιήθηκαν αποκλειστικά μέσω τηλεδιάσκεψης.

Επομένως, στα μαθήματα που μελετήσαμε φαίνεται να συμβαίνει ό,τι καταγράφουν και άλλες αντίστοιχες έρευνες. Ειδικότερα, όπως αναφέρουν οι Motz, Quick, Schroeder, Zook & Gunkel (2019, σ. 301), η χρήση του ΣΔΜ από τους φοιτητές εξαρτάται καθοριστικά από τη δομή του μαθήματος, οπότε οι δείκτες που επιχειρούν να αναπαραστήσουν τη χρήση του ΣΔΜ δεν είναι εύκολα μεταφέρσιμοι (portable) μεταξύ των μαθημάτων. Οι ίδιοι προσθέτουν ότι η δυσκολία

που ενέχει η μεταφορά μοντέλων σε «μη οικεία» μαθήματα αποτελεί πτυχή του ευρύτερου ζητήματος της απόδοσης των γενικών έναντι των ειδικών προβλεπτικών μοντέλων. Το συγκεκριμένο ζήτημα έχει αναδειχθεί emphatically από τους Gasevic, Dawson, Rogers & Gasevic (2016) σύμφωνα με τους οποίους, οι περιορισμοί των γενικών μοντέλων να προβλέψουν την επίδοση των φοιτητών σε «μη οικεία» μαθήματα, εγείρει ερωτήματα σχετικά με την καταλληλότητα (relevance) και την εφαρμοσιμότητα (applicability) των μεθόδων, οι οποίες αξιοποιούν δεδομένα χρήσης των ΣΔΜ για να προβλέψουν οποιαδήποτε συμπεριφορά των σπουδαστών, χωρίς όμως να λαμβάνουν υπόψιν τις σχεδιαστικές (instructional) και παιδαγωγικές συνθήκες των μαθημάτων. Συνεπώς, θεωρούν ότι είναι απαραίτητη η δημιουργία μοντέλων τα οποία θα ενσωματώνουν τις συνθήκες του εκπαιδευτικού σχεδιασμού, διότι μόνο έτσι είναι δυνατόν να κατανοηθεί το «πότε», το «πώς» και το «γιατί» οι εκπαιδευόμενοι χρησιμοποιούν τους πόρους και τις υπηρεσίες του ΣΔΜ. Επιπλέον, όπως υποστηρίζουν οι Joksimovic, Gavsevic, Loughin, Kovanovic & Hatala (2015), οι τύποι αλληλεπίδρασης των σπουδαστών με τους άλλους σπουδαστές, με τους διδάσκοντες, με το εκπαιδευτικό υλικό, με τις δραστηριότητες που σχεδιάζονται και υλοποιούνται στα πλαίσια του μαθήματος έχουν μεγάλη επίδραση στα μαθησιακά αποτελέσματα. Οπότε, η κατασκευή ενός μοντέλου που επιχειρεί προβλέψεις σε πολλά μαθήματα, άλλοτε υποτιμά και άλλοτε υπερτιμά επιδραστικούς παράγοντες, που αναδείχθηκαν κατά τη φάση «εκμάθησης» του μοντέλου. Γενικά, κατά την κατασκευή ενός προβλεπτικού μοντέλου είναι απαραίτητο να λαμβάνεται υπόψη ο εκπαιδευτικός σχεδιασμός των μαθημάτων.

Οι Conijn, Snijders, Kleingeld & Matzat (2016, σ. 27) υπογραμμίζουν την ανάγκη θεωρητικής θεμελίωσης των μεθόδων πρόβλεψης. Ειδικότερα, αναφέρουν ότι η χαμηλή απόδοση κάποιων προβλεπτικών μοντέλων ίσως οφείλεται στο ότι δεν είναι πάντοτε σαφές τι ακριβώς μετρούν οι μετρήσεις. Επιπλέον, υποστηρίζουν ότι οι μέθοδοι ανάλυσης των δεδομένων δεν εντάσσονται σε ένα στέρεο θεωρητικό υπόβαθρο και ότι για να βελτιωθεί η απόδοση των μοντέλων είναι απαραίτητη η εμβάθυνση στο τι πραγματικά αναπαριστούν τα δεδομένα και στο πώς μπορούν να αντιστοιχηθούν σε σαφή μεγέθη μέτρησης εννοιών. Όπως αναφέρουν οι Joksimovic, Gavsevic, Loughin, Kovanovic & Hatala (2015), τα προβλεπτικά μοντέλα θα είναι περισσότερο χρήσιμα, αν, εκτός από προβλέψεις, προσφέρουν και εξηγήσεις.



Συμπερασματικά, μελετώντας τις διαφορές των μοντέλων που κατασκευάσαμε, προκύπτει ότι η σημασία του εκπαιδευτικού σχεδιασμού είναι πρόδηλη, όπως άλλωστε υποστηρίζουν τα πορίσματα της βιβλιογραφίας. Η απόδοση των μοντέλων επηρεάζεται από την ύπαρξη δραστηριοτήτων που υποχρεώνουν σε συστηματική χρήση του ΣΔΜ. Επηρεάζεται από την υποχρέωση παρακολούθησης του μαθήματος. Επηρεάζεται από αξιολογικές δραστηριότητες που εμπλουτίζουν το σύνολο δεδομένων με βαθμούς που συνυπολογίζονται στην τελική επίδοση των φοιτητών.

#### *Απάντηση στο Ερευνητικό Ερώτημα ΕΕΡ8*

Συνεπώς, από την ανάλυση των ερευνητικών αποτελεσμάτων που αφορούν τα μαθήματα που μελετήθηκαν, προκύπτει ότι ο εκπαιδευτικός σχεδιασμός και ειδικότερα παράμετροι όπως ο τρόπος υλοποίησης των διαλέξεων, η υποχρεωτικότητα της φοίτησης, η υλοποίηση αλληλεπιδραστικών δραστηριοτήτων, ο τρόπος αξιολόγησης επηρεάζουν την απόδοση των προβλεπτικών μοντέλων.

## 12. ΣΥΜΠΕΡΑΣΜΑΤΑ – ΣΥΖΗΤΗΣΗ

### 12.1. Σύνοψη αποτελεσμάτων

Συνοψίζοντας τα αποτελέσματα της έρευνάς μας, καταρχάς σημειώνουμε ότι ο βασικός σκοπός ήταν η μελέτη των προβλεπτικών μοντέλων Μηχανικής Μάθησης σε εκπαιδευτικά περιβάλλοντα και των δυνατοτήτων αξιοποίησής τους στην εκπαιδευτική αξιολόγηση. Ειδικότερα μελετήσαμε τις δυνατότητες των μοντέλων να αναλύσουν δεδομένα χρήσης των Συστημάτων Διαχείρισης Μάθησης (ΣΔΜ) και να προβλέψουν τις επιδόσεις των φοιτητών, όχι μόνο με ακρίβεια, αλλά και εγκαίρως. Αξιοποιήσαμε διερευνητικές μελέτες περίπτωσης, κατά τις οποίες συλλέξαμε δεδομένα χρήσης του ΣΔΜ Moodle τα οποία αφορούσαν τέσσερα μαθήματα σε προπτυχιακό και μεταπτυχιακό επίπεδο, οργανώσαμε τα ανάλογα σύνολα δεδομένων και με τη βοήθεια αλγορίθμων Μηχανικής Μάθησης κατασκευάσαμε μοντέλα που προβλέπουν την επίδοση των φοιτητών. Μελετήσαμε τα μοντέλα ως προς την ακρίβεια και την αξιοπιστία της πρόβλεψης, ως προς την ικανότητά τους για πρώιμη πρόβλεψη και για γενίκευση. Επίσης, εντοπίσαμε χαρακτηριστικά των συνόλων δεδομένων, τα οποία έχουν μεγάλη επίδραση στην απόδοση των μοντέλων. Από την ανάλυση των αποτελεσμάτων προέκυψαν τα εξής:

Τα μοντέλα επιτυγχάνουν υψηλές και αξιόπιστες τιμές πρόβλεψης με στατιστικώς σημαντική διαφορά από την τιμή βάσης. Οι υψηλότερες τιμές πρόβλεψης, που ανάλογα με το μάθημα προσεγγίζουν ή υπερβαίνουν το 80% , επιτυγχάνονται από τα μοντέλα που προκύπτουν από πλήρη σύνολα δεδομένων, δηλαδή σύνολα που περιλαμβάνουν δεδομένα χρήσης από την έναρξη ως την ολοκλήρωση του μαθήματος. Όλοι σχεδόν οι βασικοί αλγόριθμοι Μηχανικής Μάθησης παράγουν μοντέλα με υψηλές τιμές ακρίβειας και αξιοπιστίας, όταν εφαρμοσθούν σε πλήρη σύνολα δεδομένων. Η ακρίβεια και η αξιοπιστία της πρόβλεψης φθίνει όσο κινούμαστε χρονικά αντίθετα στο βέλος του χρόνου, δηλαδή από το τέλος προς την έναρξη του μαθήματος. Η πρωιμότερη χρονική στιγμή που μπορεί να επιτευχθεί αξιόλογη και στατιστικώς σημαντική ακρίβεια πρόβλεψης είναι περίπου στο μέσο του μαθήματος. Μοντέλα με ικανότητα πρώιμης πρόβλεψης παράγονται από μερικούς μόνο αλγόριθμους.

Τα μοντέλα που προκύπτουν από τεχνικές «βαθιάς μάθησης» και συγκεκριμένα από LSTM αλγόριθμους, τα οποία έχουν την ικανότητα να αναλύουν όχι μόνο τις συνολικές, τελικές τιμές

των χαρακτηριστικών του συνόλου δεδομένων, αλλά και τις χρονικές ακολουθίες, οι οποίες αποτυπώνουν την εξέλιξη των τιμών, επιτυγχάνουν χαμηλότερη απόδοση και ως προς την ακρίβεια πρόβλεψης και ως προς την αξιοπιστία, από ό,τι τα βασικά μοντέλα. Επιπλέον, η κατασκευή συνόλων δεδομένων για τα LSTM μοντέλα είναι επίπονη διαδικασία, η εκτέλεση των αλγορίθμων είναι χρονοβόρα και τα παραγόμενα μοντέλα δεν ερμηνεύονται εύκολα.

Το περισσότερο επιτυχημένο μοντέλο, δηλαδή το μοντέλο που σε όλα τα μαθήματα που μελετήθηκαν επιτυγχάνει υψηλές τιμές ακρίβειας πρόβλεψης και ταυτοχρόνως πρώιμη πρόβλεψη, είναι αυτό που κατασκευάζεται από αλγόριθμο που βασίζεται στη μέθοδο της λογιστικής παλινδρόμησης (logistic regression), η οποία θεωρείται ως η δημοφιλέστερη μέθοδος πρόβλεψης σε εκπαιδευτικά περιβάλλοντα. Η λογιστική παλινδρόμηση, είναι μια πολυπαραγοντική στατιστική μέθοδος, όπου η πιθανότητα να λάβει η εξαρτημένη κατηγορική μεταβλητή την μία ή την άλλη τιμή, εκφράζεται ως μια γραμμική συνάρτηση των ανεξάρτητων μεταβλητών.

Η ικανότητα γενίκευσης των μοντέλων, δηλαδή η επιτυχημένη πρόβλεψη σε σύνολα δεδομένων διαφορετικών από αυτά στα οποία βασίστηκε η κατασκευή του μοντέλου, συνιστά στόχο δύσκολα επιτεύξιμο. Μεταξύ των μοντέλων που κατασκευάστηκαν, υπάρχει μοντέλο το οποίο πέτυχε πρόβλεψη σε «μη οικείο» σύνολο δεδομένων, γεγονός που αποτελεί απλή ένδειξη για την ικανότητα γενίκευσης του συγκεκριμένου μοντέλου. Η ικανότητα γενίκευσης πρέπει να αξιολογηθεί περαιτέρω με σχετικές δοκιμές σε πολλά διαφορετικά σύνολα δεδομένων.

Παρατηρήθηκε, ότι μεταξύ των χαρακτηριστικών που συγκροτούν τα σύνολα δεδομένων υπάρχουν κάποια, τα οποία επηρεάζουν περισσότερο από τα υπόλοιπα την ικανότητα πρόβλεψης των μοντέλων. Τα χαρακτηριστικά είναι κοινά μεταξύ συναφών μαθημάτων. Ο υψηλός βαθμός επίδρασης αυτών των χαρακτηριστικών μπορεί να αποδοθεί στη σχέση που έχουν με δείκτες ενεργής συμμετοχής των φοιτητών στα μαθήματα.

Παρατηρήθηκε, ότι παράγοντας με καθοριστική σημασία στον βαθμό απόδοσης των μοντέλων και γενικά στις διαφορές μεταξύ των μοντέλων είναι ο εκπαιδευτικός σχεδιασμός. Οι διαφορές ή οι ομοιότητες που έχουν τα μαθήματα, ως αποτέλεσμα διαφορετικών σχεδιαστικών επιλογών για τη δομή και την οργάνωσή τους, επηρεάζουν τον βαθμό και τον τρόπο χρήσης των ΣΔΜ,

επηρεάζουν τη βαρύτητα των χαρακτηριστικών στα σύνολα δεδομένων και τελικά επηρεάζουν και την απόδοση των μοντέλων.

## 12.2. Τα μοντέλα ως εργαλεία διαμορφωτικής αξιολόγησης

Η ικανότητα των μοντέλων να επιτυγχάνουν πρόβλεψη α) με υψηλό βαθμό ακρίβειας και αξιοπιστίας β) σε πρώιμη χρονική στιγμή, πολύ πριν την ολοκλήρωση του μαθήματος και γ) σε «μη οικεία» μαθήματα, δηλαδή σε μαθήματα διαφορετικά από αυτά που δημιουργήθηκαν, καθιστά δυνατή την αξιοποίησή τους ως εργαλείων διαμορφωτικής αξιολόγησης, κυρίως μέσω του εντοπισμού των φοιτητών που εμφανίζουν αυξημένες πιθανότητες να αποτύχουν. Ο εντοπισμός προκύπτει με την ενεργοποίηση του μοντέλου και με την τροφοδοσία του με τα δεδομένα χρήσης του ΣΔΜ, χωρίς να υπάρχει ανάγκη για σχεδιασμό και υλοποίηση επιπλέον διαμορφωτικών αξιολογικών διαδικασιών. Το επόμενο βήμα είναι η προληπτική διδακτική ενισχυτική παρέμβαση. Ωστόσο, προκύπτουν μια σειρά από ζητήματα.

Πρώτον, τα μοντέλα λειτουργούν πιθανοκρατικά, οπότε αναπόφευκτες είναι οι αστοχίες. Για παράδειγμα, είναι πιθανόν, το μοντέλο να χαρακτηρίσει «αδύναμο», κάποιον «καλό» φοιτητή. Πρόκειται για την περίπτωση του αποκαλούμενου «λανθασμένου συναγερμού», ο οποίος θα προκαλέσει άσκοπη διδακτική παρέμβαση. Είναι επίσης πιθανόν, το μοντέλο να χαρακτηρίσει ως «καλό», κάποιον «αδύναμο» φοιτητή. Πρόκειται για πλάνη βαρύτερη από την προηγούμενη, διότι η διδακτική παρέμβαση, την οποία ο «αδύναμος» φοιτητής έχει ανάγκη, δεν θα υλοποιηθεί. Προφανώς, το πλήθος των αστοχιών είναι αντιστρόφως ανάλογο με την απόδοση του μοντέλου. Αλλά, ας σημειωθεί, ότι αστοχίες αναπόφευκτα θα προκύψουν και στην περίπτωση που για τον εντοπισμό «αδύναμων» φοιτητών χρησιμοποιηθούν συμβατικές διαδικασίες (π.χ. διαγώνισμα).

Δεύτερον, τα μοντέλα επιχειρούν να προβλέψουν ποιοι φοιτητές πιθανόν θα αποτύχουν, αλλά δεν προβλέπουν τους λόγους για τους οποίους θα αποτύχουν. Χρειάζεται να αναπτυχθούν άλλες τεχνικές προκειμένου να προκύψουν ενδείξεις για τους λόγους της πιθανής αποτυχίας. Μια από τις τεχνικές, την οποία αξιοποιήσαμε στην έρευνά μας, είναι ο εντοπισμός επιδραστικών χαρακτηριστικών, δηλαδή των χαρακτηριστικών στα σύνολα δεδομένων, τα οποία έχουν την μεγαλύτερη επίδραση στην απόδοση των μοντέλων. Από την έρευνά μας

προέκυψε ότι στα μαθήματα που μελετήσαμε, τα επιδραστικότερα χαρακτηριστικά σχετίζονται με την ενεργή συμμετοχή των φοιτητών. Συνεπώς, οι διδακτικές παρεμβάσεις που θα ενισχύουν χαρακτηριστικά που σχετίζονται με την ενεργή συμμετοχή, πιθανόν να συμβάλουν στην επιτυχημένη φοίτηση. Για παράδειγμα, στη έρευνά μας αναδείχθηκε η μεγάλη σημασία της συμμετοχής στους δημόσιους διαλόγους, οπότε μια διδακτική παρέμβαση που θα ενίσχυε τη συμμετοχή στους δημόσιους διαλόγους, πιθανόν να ενίσχυε και τις πιθανότητες επιτυχούς φοίτησης. Άλλα επιδραστικά χαρακτηριστικά τα οποία σχετίζονται με τη χρήση ψηφιακών εκπαιδευτικών πόρων και υπηρεσιών και τα οποία πιθανόν να αναδειχθούν από άλλα μοντέλα, θα μπορούσαν επίσης να υποδείξουν προς το «πού» θα πρέπει να στοχεύσουν αντίστοιχες διδακτικές παρεμβάσεις.

### 12.3. Προκλήσεις

#### 12.3.1. Ηθικά ζητήματα

Τα προβλεπτικά μοντέλα συνιστούν θεωρητικές κατασκευές ανάλυσης δεδομένων που αφορούν φυσικά πρόσωπα, τα οποία μπορεί να υποστούν βλάβη από την (ακούσια ή σκόπιμη) «κακή» χρήση των προσωπικών δεδομένων τους. Οπότε, η μελέτη του σχετικού ερευνητικού πεδίου δεν μπορεί να αδιαφορεί για τα σχετικά ζητήματα ηθικής και δεοντολογίας. Όπως αναφέρει η Δημητρακοπούλου (2017, σ. 22), η χρήση συστημάτων Μαθησιακής Αναλυτικής αποτελεί «αντικείμενο αναστοχασμού και κριτικής ως προς το ζήτημα της προστασίας των προσωπικών δεδομένων». Εγείρονται ερωτήματα ως προς το τι συνιστά ηθικώς αποδεκτή μελέτη δεδομένων στα πλαίσια της Μαθησιακής Αναλυτικής και ως προς το πώς και σε τι βαθμό θα πρέπει οι σπουδαστές να ενημερώνονται σχετικά με τη χρήση των δεδομένων τους και να παρέχουν τη συναίνεσή τους (Mathrani, Susnjak, Ramaswami & Barczak, 2021, σ. 6). Επιπλέον, στα προβλεπτικά μοντέλα είναι κρίσιμο το ζήτημα του ρόλου και της ασφάλειας του χαρακτηρισμού της επίδοσης των φοιτητών (Δημητρακοπούλου, 2017, σ. 22). Διότι, παρότι η πρόβλεψη της επίδοσης και η υλοποίηση υποστηρικτικών εκπαιδευτικών παρεμβάσεων έχει ως στόχο τη βελτίωση του εκπαιδευτικού αποτελέσματος είναι δυνατόν να θεωρηθεί ως μορφή παρακολούθησης. Η θεσμοθέτηση και η εφαρμογή κανόνων που θα ρυθμίζουν τα σχετικά ζητήματα είναι απαραίτητη. Οι Mathrani, Susnjak, Ramaswami & Barczak (2021, σ. 6), προτείνουν τη σύνταξη σχετικών πρωτοκόλλων που θα διασφαλίζουν τη συναίνεση των

φοιτητών. Σχετικοί κανονισμοί όπως ο Γενικός Κανονισμός για την Προστασία των Δεδομένων (GDPR) μπορούν να αποτελέσουν το ευρύτερο πλαίσιο, εντός του οποίου θα εξειδικευθούν τα σχετικά ακαδημαϊκά πρωτόκολλα. Ως προς το θέμα της συναίνεσης, θα πρέπει να ληφθούν κατάλληλες πρόνοιες και για τους απόφοιτους με τους οποίους το εκπαιδευτικό ίδρυμα δεν είναι εύκολο να επικοινωνήσει για να εξασφαλίσει τη συναίνεσή τους. Σε αντίστοιχες περιπτώσεις, αλλά και γενικότερα όταν απαιτείται η διασφάλιση της ιδιωτικότητας και της εμπιστευτικότητας, αποδοτική είναι η τεχνική της ψευδωνυμοποίησης (π.χ. με hash αλγόριθμους), όπου το φυσικό αναγνωριστικό κάθε φοιτητή αντικαθίσταται από τεχνικό αναγνωριστικό, μέσω του οποίου δεν μπορεί να υπάρξει συσχέτιση με τον φοιτητή. Ζήτημα επίσης αποτελεί ο διαμοιρασμός δεδομένων με άλλους οργανισμούς. Δεν αναφερόμαστε σε ιδιωτικούς οργανισμούς που αξιοποιούν τα δεδομένα για να προσκομίσουν οικονομικό όφελος, περίπτωση κατά την οποία ο διαμοιρασμός προσωπικών δεδομένων είναι παράνομος και ηθικά επιλήψιμος, αλλά αναφερόμαστε στη χρήση δεδομένων από κρατικούς, εθνικούς ή υπερεθνικούς οργανισμούς, αρμόδιους για την υλοποίηση εκπαιδευτικών πολιτικών. Και σε αυτές τις περιπτώσεις, η θέσπιση σαφούς λειτουργικού και καθολικού κανονιστικού πλαισίου είναι απαραίτητη.

### 12.3.2. Ερμηνευσιμότητα

Τα προβλεπτικά μοντέλα παρέχουν πληροφορίες που μπορεί να χρησιμεύσουν σε αποφάσεις, όπως στοχευμένες διδακτικές παρεμβάσεις προς τους φοιτητές για τους οποίους προκύπτουν αυξημένες πιθανότητες μη επιτυχημένης φοίτησης. Ωστόσο, η κατά τέτοιο τρόπο αξιοποίηση των μοντέλων, προϋποθέτει ότι χαίρουν της εμπιστοσύνης του διδάσκοντα, ή γενικότερα του εκπαιδευτικού οργανισμού. Η εμπιστοσύνη έναντι των μοντέλων είναι ανάλογη με τον βαθμό διαφάνειας της λειτουργίας τους. Συνεπώς, εξαιτίας της εγγενούς πολυπλοκότητάς τους, τα μοντέλα συχνά παρουσιάζονται και αντιμετωπίζονται ως «αδιαφανή κουτιά» (black boxes), έναντι των οποίων υπάρχει εύλογη δυσπιστία και δισταγμός για την αξιοποίησή τους. Η ερμηνευσιμότητα των μοντέλων αποτελεί απαραίτητη προϋπόθεση για την αξιοποίησή τους, διότι αν τα μοντέλα δεν είναι ερμηνεύσιμα, τότε δεν μπορούν να πείσουν για τη σταθερότητα, την αξιοπιστία και την αμεροληψία τους. Σύμφωνα με τους Mathrani, Susnjak, Ramaswami & Barczak (2021, σ. 5), η ερμηνευσιμότητα έχει δύο διαστάσεις. Την συνολική ερμηνευσιμότητα

του μοντέλου, που σχετίζεται με την ικανότητα του να περιγράφει τον τρόπο λειτουργίας του σε αφηρημένο επίπεδο και την ερμηνευσιμότητα της πρόβλεψης, που σχετίζεται με την ικανότητα του μοντέλου να εξηγεί με ποιο τρόπο και γιατί οδηγήθηκε σε συγκεκριμένη πρόβλεψη (π.χ. ότι κάποιος φοιτητής κινδυνεύει να αποτύχει). Υπάρχουν μοντέλα (όπως αυτά που βασίζονται σε κανόνες, ή σε δέντρα αποφάσεων) ικανά να παράσχουν ευκολότερα εξηγήσεις για τις προβλέψεις τους, δηλαδή είναι περισσότερο διαφανή από άλλα.

Δυστυχώς, η ερμηνευσιμότητα των μοντέλων είναι συνήθως αντιστρόφως ανάλογη της αποδοτικότητάς τους (Gunning & Aha, 2019). Ωστόσο, ακόμα και στις περιπτώσεις αδιαφανών μοντέλων, υφίσταται τεχνικές, οι οποίες μέσω ενδιάμεσων μοντέλων επιχειρούν να διαφωτίσουν τα ενδότερα της λειτουργίας τους (Mathrani, Susnjak, Ramaswami & Barczak (2021, σ. 5).

### 12.3.3. Γενικευσιμότητα

Η ικανότητα ενός μοντέλου για γενίκευση αποτελεί απαραίτητη προϋπόθεση για την αξιοποίηση του. Διότι, η επίτευξη αποδοτικότητας και αξιοπιστίας δεν έχουν πρακτική αξία (προφανώς έχουν θεωρητική αξία), αν το μοντέλο δεν μπορεί να παραγάγει προβλέψεις για δεδομένα διαφορετικά από αυτά που χρησιμοποιήθηκαν για την κατασκευή του. Τα μοντέλα προκύπτουν από τη φάση «εκμάθησης» επί ενός συνόλου δεδομένων (σύνολο εκμάθησης) και δοκιμάζονται σε άλλο σύνολο δεδομένων (σύνολο δοκιμής). Σύμφωνα με τους Mathrani, Susnjak, Ramaswami & Barczak (2021, σ. 4), η βασική αιτία για τη δυσχέρεια στη γενίκευση των μοντέλων είναι οι διαφορές που προκύπτουν λόγω των αποκλίσεων μεταξύ του συνόλου εκμάθησης και του συνόλου δοκιμής. Αν το σύνολο εκμάθησης σχετίζεται με μάθημα που διαφέρει σημαντικά από το μάθημα του συνόλου δοκιμής, τότε είναι αναμενόμενο η απόδοση του μοντέλου να αμβλυνθεί, διότι το μοντέλο προσπαθεί να εντοπίσει μοτίβα διαφορετικά από αυτά που γνωρίζει και στα οποία εκπαιδεύτηκε. Ακόμα και σε μαθήματα ίδιου γνωστικού αντικειμένου, οι διαφορές σε εκπαιδευτικές δραστηριότητες, ή στον τρόπο αξιολόγησης, ή στο προφίλ των φοιτητών επιφέρουν διαφορές μεταξύ των συνόλων εκμάθησης και των συνόλων δοκιμής, οπότε τα μοντέλα δυσκολεύονται να παράγουν ακριβείς προβλέψεις.

Το ίδιο πρόβλημα μπορεί να προκύψει αν το σύνολο εκμάθησης είναι σχετικά μικρό. Τότε, η εκπαίδευση του μοντέλου είναι ατελής, διότι τα μοτίβα που αναζητά είναι σχεδόν αδιαμόρφωτα

εντός των μικρών συνόλων. Ατελής είναι επίσης η εκπαίδευση του μοντέλου, όταν τα σύνολα δεδομένων δεν είναι ισορροπημένα, δηλαδή όταν (όπως συνήθως συμβαίνει στα εκπαιδευτικά σύνολα δεδομένων) η αντιπροσώπευση μια κλάσης (π.χ. «επιτυχία») είναι δυσανάλογα μεγάλη ως προς την άλλη κλάση (π.χ. «αποτυχία»). Σε αυτές τις περιπτώσεις, (δηλαδή πολύ συχνά στα εκπαιδευτικά σύνολα δεδομένων), η απόδοση του μοντέλου για την ολιγάριθμη κλάση υπολείπεται από την απόδοση για την πολυάριθμη κλάση.

Επιπλέον, σημαντικός είναι ο ρόλος της κατάλληλης επιλογής των χαρακτηριστικών στα σύνολα δεδομένων, ο οποίος ίσως είναι σημαντικότερος από την επιλογή των κατάλληλων αλγόριθμων και από το μέγεθος του συνόλου εκμάθησης. Η επιλογή των χαρακτηριστικών αποτελεί δύσκολο εγχείρημα, που απαιτεί γνώση της δομής των μαθημάτων και εξοικείωση με τις τεχνικές σχεδιασμού προβλεπτικών μοντέλων. Η επιλογή των χαρακτηριστικών είναι μια άσκηση ισορροπίας, μεταξύ αφενός αποδοτικών, αλλά υπερ-προσαρμοσμένων και αφετέρου γενικεύσιμων, αλλά λιγότερο αποδοτικών μοντέλων.

Γενικά, η επίτευξη της ικανότητας γενίκευσης απαιτεί συνεχή αναπροσαρμογή των μοντέλων. Όπως επισημαίνουν οι Mathrani, Susnjak, Ramaswami & Barczak (2021, σ. 8), ο κύκλος ανάπτυξης ενός μοντέλου δεν ολοκληρώνεται «δια μιας» (*is not a one-off task*), αλλά συνιστά μια επαναλαμβανόμενη εμπειρική διαδικασία δοκιμής και λάθους (a process of trial and error).

#### 12.3.4. Ανάγκη θεωρητικής θεμελίωσης

Σχετική με το ζήτημα της γενικευσιμότητας, είναι η ανάγκη θεωρητικής θεμελίωσης του πεδίου των εκπαιδευτικών μοντέλων πρόβλεψης. Όπως υποστηρίζουν οι Gasevic, Dawson, Rogers & Gasevic (2016, σ. 69), ευδιάκριτη είναι η απουσία θεωρητικής συζήτησης στη βιβλιογραφία για τα χαρακτηριστικά των συνόλων δεδομένων, που συνιστούν κρίσιμους δείκτες των αλληλεπιδράσεων και της ακαδημαϊκής επιτυχίας. Θεωρούν ότι σύμφωνα με τις θεωρίες της μάθησης, η πρόβλεψη των επιδόσεων δεν μπορεί να προκύπτει μόνο από την ανάλυση των εκπαιδευτικών δεδομένων που καταγράφονται στα ΣΔΜ. Αναφέρουν ότι, για παράδειγμα, οι μετά-συμπεριφοριστικές (post-behaviorist) θεωρίες τονίζουν τη σπουδαιότητα των ειδικών κάθε φορά μαθησιακών συνθηκών και των προθέσεων των σπουδαστών και των διδασκόντων, ενώ οι εποικοδομιστικές (constuctivist) θεωρίες τονίζουν τη σπουδαιότητα του εκπαιδευτικού σχεδιασμού και των εσωτερικών συνθηκών των σπουδαστών.



Οι Chen & Cui (2020, σ. 2) επιχειρούν να συνεισφέρουν στη θεωρητική συζήτηση του πεδίου, υποστηρίζοντας ότι σύμφωνα με τη θεωρία του εποικοδομισμού, τα ΣΔΜ είναι σημαντικά για την επιτυχημένη μάθηση. Αναφέρουν ότι η μάθηση είναι ενεργή διαδικασία, όπου οι σπουδαστές δεν είναι παθητικοί αποδέκτες γνώσεων, αλλά ενεργά συμμετέχουν στη μάθηση και δομούν έννοιες από μαθησιακές πηγές. Συνεπώς, τα ΣΔΜ, δεδομένου ότι ενσωματώνουν πολλούς εκπαιδευτικούς πόρους και τους παρέχουν με ευέλικτους τρόπους, μπορούν να δομήσουν εκπαιδευτικά περιβάλλοντα, όπου οι σπουδαστές δομούν γνώση αυτενεργώντας. Ειδικότερα, τα ΣΔΜ επιτρέποντας εύκολη και γρήγορη πρόσβαση σε εκπαιδευτικό υλικό και δραστηριότητες, διευκολύνουν τους σπουδαστές να αφομοιώσουν αποτελεσματικότερα ό,τι έμαθαν στο μάθημα. Επιπλέον, τα ΣΔΜ υποστηρίζουν την επικοινωνία των σπουδαστών με τους άλλους σπουδαστές και με τους διδάσκοντες, συνεπώς διευκολύνουν, αφενός τη δόμηση γνώσης μέσω της αλληλεπίδρασης με την κοινότητα και αφετέρου την παροχή ανατροφοδότησης και μαθησιακής υποστήριξης.

Γενικά, η θεωρητική θεμελίωση του πεδίου της Μαθησιακής Αναλυτικής δεν είναι επαρκώς εύρωστη, γεγονός που ίσως σχετίζεται με τη δυσκολία της γενικευσιμότητας των μοντέλων και την απόκλιση που παρατηρείται σε πορίσματα ερευνών που μελετούν μοντέλα πρόβλεψης.

#### 12.4. Συμβολή της διατριβής

1. Καταρχάς, η διατριβή διερευνά πτυχές ζητημάτων ενός αρκετά νέου επιστημονικού πεδίου. Έχουν περάσει 12 μόνο χρόνια από πρώτο διεθνές συνέδριο, το οποίο παρουσίασε τη Μαθησιακή Αναλυτική ως διακριτό επιστημονικό πεδίο (International Conference on Learning Analytics & Knowledge (LAK), 2011), ενώ το 2014 εκδόθηκε το πρώτο τεύχος του διεθνούς επιστημονικού περιοδικού Journal of Learning Analytics [ISSN 1929-7750]. Έκτοτε, οι ερευνητικές κατευθύνσεις που προσφέρει η Μαθησιακή Αναλυτική ελκύουν το ενδιαφέρον ολοένα και περισσότερων ερευνητών. Όπως αναφέρουν οι Cui, Chen, Shiri & Fan (2019, σ. 219), η Μαθησιακή Αναλυτική εξακολουθεί να είναι ένα αναπτυσσόμενο πεδίο στην εκπαίδευση, οι σχετικές εφαρμογές είναι μικρής κλίμακας και προκαταρκτικές και τα δεδομένα που συλλέγονται δεν έχουν ακόμα

αξιοποιηθεί πλήρως για να παράσχουν πολύτιμες γνώσεις που θα διευκολύνουν και θα υποστηρίξουν την εκπαίδευση. Επί μέρους πεδίο της Μαθησιακής Αναλυτικής είναι η ανάπτυξη προβλεπτικών μοντέλων Μηχανικής Μάθησης σε εκπαιδευτικά περιβάλλοντα, τα οποία μελετά η διατριβή.

2. Πολλά από τα προβλεπτικά μοντέλα (Baashar, Alkawsi, Ali, Alhussian, & Bahbouh, 2021) χρησιμοποιούν δημογραφικά και κοινωνικό-οικονομικά δεδομένα και δεδομένα από την προηγούμενη εκπαιδευτική διαδρομή των φοιτητών (Baashar, Alkawsi, Ali, Alhussian & Bahbouh, 2021. Chen & Cui, 2020. Cui, Chen, Shiri & Fan, 2019). Η διατριβή χρησιμοποιεί δεδομένα που προέρχονται αποκλειστικά από τα αρχεία καταγραφής των ΣΔΜ. Η άντληση και η επεξεργασία δεδομένων αποκλειστικά από τα ΣΔΜ πλεονεκτεί έναντι της άντλησης και επεξεργασίας δεδομένων από άλλες πηγές, διότι αυτές είναι συνήθως διασκορπισμένες, ασύμβατες, ετερογενείς, μη πλήρεις, μη ψηφιοποιημένες. Συνεπώς, η συγκρότηση συνόλων δεδομένων από πολλές πηγές καθιστά δυσχερή την ανάπτυξη και εφαρμογή προβλεπτικών μοντέλων. Αντιθέτως, τα μοντέλα που βασίζονται αποκλειστικά σε δεδομένα από τα ΣΔΜ (όπως αυτά που μελετήσαμε) είναι περισσότερο εύχρηστα και εφαρμόσιμα.
3. Η διατριβή αντλεί από τα πρωτογενή αρχεία καταγραφής των ΣΔΜ ένα ευρύ σύνολο από χαρακτηριστικά (27 στο πλήθος), αρκετά περισσότερα από τα χαρακτηριστικά άλλων ανάλογων μελετών. Επιπλέον, χρησιμοποιεί χαρακτηριστικά που δεν σχετίζονται μόνο με απλά αριθμητικά μεγέθη (όπως π.χ. το πλήθος συνόδων πρόσβασης, ή το πλήθος συμβάντων), τα οποία συνήθως αξιοποιούν οι περισσότερες σχετικές έρευνες και τα οποία είναι σχετικά εύκολο να προκύψουν από τα πρωτογενή αρχεία του ΣΔΜ. Η διατριβή χρησιμοποιεί επιπλέον χαρακτηριστικά που σχετίζονται με τη συχνότητα, την πυκνότητα, την κανονικότητα, την ασυνέπεια πρόσβασης, τα οποία απαιτούν μεν σύνθετους αριθμητικούς υπολογισμούς, αλλά είναι αρκετά αποκαλυπτικά, ως προς τις αλληλεπιδράσεις των φοιτητών με το σύστημα.

4. Τα περισσότερα προβλεπτικά μοντέλα παράγονται από βασικούς, ευρείας χρήσης αλγόριθμους Μηχανικής Μάθησης (Chen & Cui, 2020). Η διατριβή συγκρίνει τα μοντέλα που παράγονται από τους βασικούς αλγόριθμους, με μοντέλα που παράγονται από τεχνικές «βαθιάς μάθησης», τα οποία αναλύουν τις χρονικές ακολουθίες εξέλιξης στις τιμές των χαρακτηριστικών. Τα συμπεράσματα ανάλογων ερευνών δεν καταδεικνύουν με σαφήνεια, αν και κατά πόσο οι τεχνικές «βαθιάς μάθησης» είναι αποδοτικότερες (Chen & Cui, 2020), οπότε η διατριβή επιχειρεί να συνεισφέρει στο ζήτημα. Επιπλέον, τα σύνολα, με τα οποία η διατριβή τροφοδοτεί τα μοντέλα «βαθιάς μάθησης», περιλαμβάνουν όχι ένα ή δύο, ως συνήθως, αλλά πολύ περισσότερα χαρακτηριστικά.
5. Η διατριβή προτείνει τρόπους για την αξιοποίηση των προβλεπτικών μοντέλων ως εργαλείων διαμορφωτικής αξιολόγησης. Ειδικότερα:
  - α. Πολλά από τα προβλεπτικά μοντέλα χρησιμοποιούν δεδομένα που αφορούν όλο το χρονικό διάστημα που διαρκεί το μάθημα (Riestra-Gonzalez, del Puerto Paule-Ruiz & Ortin, 2021). Ακόμα και αν οι προβλέψεις είναι ακριβείς, τέτοια μοντέλα υστερούν σε πρακτική αξία και δεν μπορούν να αξιοποιηθούν επαρκώς σε πραγματικά εκπαιδευτικά περιβάλλοντα. Διότι, όταν οι προβλέψεις παράγονται μετά την ολοκλήρωση του μαθήματος, είναι πλέον αργά για διδακτικές παρεμβάσεις. Η διατριβή αναζητά μοντέλα ικανά να δώσουν πρώιμες προβλέψεις, χάρη στις οποίες ένα μοντέλο μπορεί να αξιοποιηθεί ως εργαλείο διαμορφωτικής αξιολόγησης, εντοπίζοντας εγκαίρως φοιτητές που διατρέχουν κίνδυνο να αποτύχουν και προσανατολίζοντας διδακτικές ενισχυτικές παρεμβάσεις.
  - β. Η διατριβή αναζητά εντός των συνόλων δεδομένων επιδραστικά χαρακτηριστικά. Δηλαδή χαρακτηριστικά που σχετίζονται περισσότερο από τα υπόλοιπα με την ακαδημαϊκή επιτυχία και την επηρεάζουν. Τέτοια χαρακτηριστικά, τα οποία αποτυπώνουν κρίσιμα μοτίβα και πρακτικές στην χρήση των ΣΔΜ, παρέχουν στον διδάσκοντα πληροφορίες για αντίστοιχα κρίσιμα στοιχεία στον εκπαιδευτικό

σχεδιασμό του μαθήματος και τον προσανατολίζουν να προβεί σε χρήσιμες προσαρμογές.

6. Η διατριβή συγκρίνει αποτελέσματα που αφορούν μαθήματα τα οποία καλύπτουν το συνεχές των τεχνολογικά υποστηριζόμενων εκπαιδευτικών περιβαλλόντων (Φεσάκης & Κωνσταντοπούλου, 2022, σσ. 34-35). Ειδικότερα τα μαθήματα που μελετώνται καλύπτουν και τους τρεις βασικούς τρόπους διδασκαλίας που αξιοποιούν την Τεχνικά Ενισχυμένη Μάθηση (Πρόσωπο με Πρόσωπο Διδασκαλία, Μεικτή Μάθηση, Διαδικτυακή Μάθηση).
7. Η διατριβή παρουσιάζει λεπτομερώς τους υπολογισμούς για τις τιμές των χαρακτηριστικών που συγκροτούν τα σύνολα δεδομένων, ώστε να είναι δυνατή η επανάληψη της ακολουθούμενης μεθόδου και η επαληθευσσιμότητα των αποτελεσμάτων. Όπως επισημαίνουν στην βιβλιογραφική ανασκόπησή τους, οι Cui, Chen, Shiri & Fan (2019, σ. 220), αλλά και οι Chen & Cui (2020), δεν είναι πολλές οι έρευνες που αναφέρουν λεπτομερώς, ποια ήταν τα στάδια επεξεργασίας των πρωτογενών δεδομένων και πώς υπολογίστηκαν οι τιμές των χαρακτηριστικών.
8. Η διατριβή μελετά τα χαρακτηριστικά των συνόλων δεδομένων, ως προς τη διαφορετική επίδραση τους στην ακρίβεια της πρόβλεψης. Αναδεικνύεται, με αυτόν τον τρόπο, ο κρίσιμος ρόλος της ενεργής συμμετοχής, εκφρασμένης σε διαστάσεις αλληλεπιδράσεων μεταξύ φοιτητών και συστήματος, οι οποίες σχετίζονται με τα επιδραστικότερα χαρακτηριστικά.
9. Η ανάδειξη χαρακτηριστικών με σημαντική επίδραση στην ακρίβεια της πρόβλεψης, τα οποία σχετίζονται με τη δομή του μαθήματος, το περιεχόμενό του, τον τρόπο αλληλεπίδρασης των συμμετεχόντων, τον τρόπο αξιολόγησης, παρέχει χρήσιμες πληροφορίες για τον εκπαιδευτικό σχεδιασμό του μαθήματος, οι οποίες αφού

αξιολογηθούν από τον σχεδιαστή του μαθήματος μπορεί να οδηγήσουν σε αναπροσαρμογές και βελτιώσεις.

### 12.5. Περιορισμοί της έρευνας

Οι περιορισμοί που της έρευνας που πραγματοποιήσαμε είναι οι παρακάτω:

- Τα δεδομένα που μελετήσαμε αφορούν πληθυσμό με μικρό σχετικά μέγεθος. Οι αλγόριθμοι Μηχανικής Μάθησης παράγουν πιο εύστοχα και πιο αξιόπιστα μοντέλα, όσο μεγαλύτερο είναι το πλήθος των οντοτήτων στα σύνολα δεδομένων. Διότι, κάθε αλγόριθμος μελετά και αναλύει επαναλαμβανόμενα μοτίβα (patterns). Όσο μεγαλύτερος είναι ο υπό μελέτη πληθυσμός, τόσο πιο ευδιάκριτα και σαφή γίνονται τα μοτίβα, οπότε τόσο περισσότερο εύστοχα και αξιόπιστα είναι τα μοντέλα. Ιδιαίτερα, τα LSTM μοντέλα, τα οποία βασίζονται σε τεχνικές «βαθιάς μάθησης» και τα οποία συγκρίναμε με τα βασικά μοντέλα, αποδίδουν όταν εφαρμόζεται σε πολύ μεγάλα σύνολα δεδομένων (Alom, et al., 2019, σ. 7).
- Οι κλάσεις των συνόλων δεδομένων δεν είναι ισορροπημένες. Ειδικότερα, όπως συνήθως συμβαίνει σε όλα στα εκπαιδευτικά συστήματα, ο αριθμός των επιτυχόντων φοιτητών είναι αρκετά μεγαλύτερος από τον αριθμό των αποτυχόντων. Ως αποτέλεσμα, τα μοντέλα, κατά το στάδιο εκμάθησης, δεν έχουν στη διάθεσή τους πολλά δεδομένα από την ολιγάριθμη κλάση και επομένως δεν μπορούν να προσαρμοσθούν το ίδιο καλά όσο στην πολυάριθμη κλάση. Αναπόφευκτα, η απόδοση των μοντέλων είναι καλύτερη στην πολυάριθμη, παρά στην ολιγάριθμη κλάση.
- Τα δεδομένα που μελετήσαμε προέρχονται από τέσσερα μαθήματα με τεχνολογικό γνωστικό αντικείμενο. Και στα τέσσερα μαθήματα, το ΣΔΜ αξιοποιήθηκε ως κύριο εκπαιδευτικό εργαλείο, οπότε η χρήση του από τους φοιτητές ήταν εντατική και ενδεικτική της ενασχόλησης τους με το μάθημα. Έτσι, κατέστη δυνατή η ανάπτυξη μοντέλων, τα οποία αξιοποιούν δεδομένα από το ΣΔΜ και επιχειρούν να προβλέψουν τις επιδόσεις των φοιτητών. Όμως, σε μαθήματα, όπου η χρήση του ΣΔΜ είναι περιορισμένη, διότι το ΣΔΜ δεν αποτελεί κεντρικό εκπαιδευτικό εργαλείο, ίσως να μην προκύψουν γόνιμα αποτελέσματα, λόγω ανεπάρκειας δεδομένων. Το συγκεκριμένο

ζήτημα αποτελεί πτυχή του γενικότερου ζητήματος της ικανότητας γενίκευσης των μοντέλων. Η ικανότητα γενίκευσης δεν εξαρτάται μόνο από την επάρκεια των δεδομένων. Διότι, ακόμα και σε μαθήματα, όπου το ΣΔΜ χρησιμοποιείται έντονα, οπότε υφίστανται πλούσια σύνολα δεδομένων, το είδος των αλληλεπιδράσεων με το σύστημα διαφέρει από μάθημα σε μάθημα και εξαρτάται από τις παραμέτρους του εκπαιδευτικού σχεδιασμού του μαθήματος. Κάθε μάθημα είναι μοναδικό, οπότε τα επιδραστικά χαρακτηριστικά του συνόλου δεδομένων του, δεν είναι πάντοτε συμβατά με τα σύνολα δεδομένων των άλλων μαθημάτων. Κατά συνέπεια, τα ασύμβατα σύνολα δεδομένων παράγουν ασύμβατα μοντέλα. Γενικά, υφίσταται λεπτή ισορροπία, μεταξύ της απόδοσης των μοντέλων και της ικανότητας τους για γενίκευση.

#### 12.6. Μελλοντικές επεκτάσεις

Με την έρευνα που διενεργήσαμε μελετήσαμε την απόδοση των μοντέλων πρόβλεψης σε τέσσερα μαθήματα υλοποιώντας διερευνητικές μελέτες περίπτωσης, που σκοπό είχαν τη βασική κατανόηση του ερευνητικού αντικειμένου. Τα πορίσματα που προέκυψαν μπορούν να αποτελέσουν σημεία αφετηρίας για μελλοντικές προεκτάσεις της έρευνας.

Συγκεκριμένα, προτείνεται η συλλογή και ανάλυση δεδομένων και από άλλα μαθήματα προκειμένου να επιβεβαιωθεί η αποδοτικότητα και η αξιοπιστία των μοντέλων που κατασκευάστηκαν. Η μελέτη και άλλων μαθημάτων, με διαφορετική δομή και γνωστικό αντικείμενο, θα συμβάλει στη βελτίωση της απόδοσης των μοντέλων και θα διερευνήσει περαιτέρω τη σχέση των μοντέλων με τον εκπαιδευτικό σχεδιασμό των μαθημάτων.

Προτείνεται η περαιτέρω αναζήτηση κατάλληλων τιμών για τις παραμέτρους του LSTM μοντέλου, ώστε να ενισχυθεί η προβλεπτική ικανότητά του. Όπως έχει αναφερθεί, δεν ήταν δυνατόν, στα πλαίσια της έρευνά μας, να εξαντλήσουμε όλους τους δυνατούς συνδυασμούς, οπότε δεν αποκλείουμε την ύπαρξη συνδυασμού τιμών για τις παραμέτρους, ο οποίος θα επιτύχει απόδοση καλύτερη από αυτή των βασικών μοντέλων. Επίσης, η απόδοση των LSTM μοντέλων αναμένεται να βελτιωθεί, αν σχηματισθούν σύνολα με πολύ περισσότερα δεδομένα από αυτά που είχαμε διαθέσιμα.

Προτείνεται η σχεδίαση και η ανάπτυξη ενός πιλοτικού συστήματος έγκαιρης προειδοποίησης (early-warning system), το οποίο θα βασίζεται στα μοντέλα που αναδείχθηκαν από την έρευνά μας ή σε βελτιωμένες εκδόσεις τους. Ένα τέτοιο σύστημα, θα μπορούσε να ενσωματωθεί ως πρόσθετη δυνατότητα (π.χ. ως plugin) στο Σύστημα Διαχείρισης Μάθησης (ΣΔΜ) και να αξιοποιηθεί από τους διδάσκοντες για τον εντοπισμό φοιτητών για τους οποίους είναι αυξημένες οι πιθανότητες μη επιτυχούς φοίτησης.

## 12.7.Επίλογος

Μελετήσαμε ένα συνεχώς αναπτυσσόμενο και ελκυστικό ερευνητικό πεδίο, το οποίο συνδυάζει στοιχεία της παιδαγωγικής επιστήμης, της επιστήμης των υπολογιστών και της μαθηματικής επιστήμης. Επιχειρήσαμε να φωτίσουμε κάποιες από τις πτυχές του.

Πολλές φορές, βρεθήκαμε ενώπιων ζητημάτων που δεν ήταν δυνατόν να διερευνήσουμε περαιτέρω, διότι ήταν πέραν των ερευνητικών και χρονικών πλαισίων της διατριβής.

Προφανώς, υφίστανται και άλλες πτυχές που θα ήταν ενδιαφέρον να διερευνηθούν.

Ωστόσο, η διαδρομή της έρευνας και της συγγραφής πρέπει να ολοκληρωθεί. Αν όχι με τελεία, τουλάχιστον με άνω τελεία\*



## ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

- 2021 *International Weka User Conference*. (2022, Οκτώβριος 05). Ανάκτηση από The University of Waikato: <https://events.waikato.ac.nz/events/2021-international-weka-user-conference-making-ai-accessible>
- About Moodle*. (2022, Σεπτέμβριος 27). Ανάκτηση Σεπτέμβριος 27, 2022, από moodle.org: [https://docs.moodle.org/400/en/About\\_Moodle](https://docs.moodle.org/400/en/About_Moodle)
- Ahlawat, V., Pareek, A., & Singh, S. (2014). Online Invigilation: A Holistic Approach. *International Journal of Computer Applications*, 90(17).
- Ahuja, H. (2020, 3 31). *Logistic Regression — Simple, Multinomial And Ordinal*. Ανάκτηση 1 16, 2023, από DataDrivenInvestor: <https://medium.datadriveninvestor.com/logistic-regression-simple-multinomial-and-ordinal-b2bc886bb974>
- Akyol, Z., Garrison, R. D., & Yasar, O. M. (2009). Online and blended communities of inquiry: Exploring the developmental and perceptual differences. *The International Review of Research in Open and Distributed Learning*, 10(6), σσ. 65-83.
- Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student' performance prediction using machine learning techniques. *Education Sciences*, 11(9), 552.
- Aldrich, C. (2009). Virtual worlds, simulations, and games for education: A unifying view. *Innovate: Journal of Online Education*, 5(5).
- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., . . . Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics*, 8(3), 292.
- Alruwais, N., Wills, G., & Wald, M. (2018). Advantages and challenges of using e-assessment. *International Journal of Information and Education Technology*, 8(1), σσ. 34-37.
- Altalhi, A. H., Luna, J. M., Vallejo, M., & Ventura, S. (2017). Evaluation and comparison of open source software suites for data mining and knowledge discovery. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(3), e1204.
- Amigud, A., Arnedo-Moreno, J., Daradoumis, T., & Guerrero-Roldan, A.-E. (2017). Using learning analytics for preserving academic integrity. *International Review of Research in Open and Distributed Learning: IRRODL*, 18(5), σσ. 192-210.
- Anderson, E. F., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, C., & de Freitas, S. (2009). Serious games in cultural. Στο M. Ashley, & F. Liarokapis (Επιμ.), *VAST 2009: 10th International Symposium on Virtual Reality, Archaeology* (σσ. 29-48). St. Julians, Malta: Eurographics Association.

- Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Apampa, K. M., Wills, G., & Argles, D. (2010). User security issues in summative e-assessment security. *International Journal of Digital Society (IJDS)*, 2, σσ. 1-13.
- Appiah, M., & van Tonder, F. (2018). E-Assessment in Higher Education: A Review. *International Journal of Business Management and Economic Research (IJBMER)*, 9, σσ. 1454-1460.
- Arbaugh, J. B. (2014). System, scholar or students? Which most influences online MBA course effectiveness? *Journal of Computer Assisted Learning*, 39(4), 349-362.
- Arias-Aranda, D., Domiguez, C., & Martinez, M. (2010, 09). An innovative approach to the learning process in management: The use of simulators in higher education. *Revista de Educación (Madrid)*, 353, σσ. 333-334.
- Arnold, I. J. (2016). Cheating at online formative tests: Does it pay off? *The Internet and Higher Education*, 29, σσ. 98-106.
- Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. *Proceedings of the 2nd international conference on learning analytics and knowledge*, (σσ. 267-270).
- Azcona, D., & Casey, K. (2015). Micro-analytics for student performance prediction. *International Journal of Computer Science and Software Engineering*, 4(8), 218-223.
- Azevedo, A., & Azevedo, J. (2019). *Handbook of Research on E-Assessment in Higher Education*. Hershey, PA.
- Azevedo, J. (2015). E-assessment in mathematics courses with multiple-choice questions tests. *7th International Conference on Computer Supported Education* (σσ. 260-266). SCITEPRESS (Science and Technology Publications, Lda.).
- Azevedo, J., Oliveira, E. P., & Beites, P. D. (2019). E-Assessment and Multiple-Choice Questions: A Literature Review. Στο *Handbook of Research on E-Assessment in Higher Education* (σσ. 1-27). IGI Global.
- Baashar, Y., Alkaws, G., Ali, N. A., Alhussian, H., & Bahbouh, H. T. (2021). Predicting student's performance using machine learning methods: A systematic literature review. *2021 International Conference on Computer & Information Sciences (ICCOINS)* (σσ. 357-362). IEEE.
- Babo, R., Babo, L. V., Suhonen, J. T., & Tukiainen, M. (2020). E-Assessment with Multiple-Choice Questions: A 5 Year Study of Students' Opinions and Experience. *Journal of Information Technology Education: Innovations in Practice*, 19, σσ. 1-29.

- Bailie, J. L., & Jortberg, M. A. (2009). Online learner authentication: Verifying the identity of online users. *Journal of Online Learning and Teaching*, 5(2), σσ. 197-207.
- Baleni, Z. G. (2015). Online formative assessment in higher education: Its pros and cons. *Electronic Journal of e-Learning*, 13(4), σσ. 228-236.
- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The journal of the learning sciences*, 13(1), 1-14.
- Barnes, C., & Paris, B. L. (2013). An analysis of academic integrity techniques used in online courses at a southern university. *Northwest Decision Sciences Institute Annual Meeting Proceedings*.
- Barrett, H. (2004). Differentiating electronic portfolios and online assessment management systems. Στο *Society for Information Technology & Teacher Education International Conference* (σσ. 46-50). Association for the Advancement of Computing in Education (AACE).
- Barthakur, A., Kovanovic, V., Joksimovic, S., Siemens, G., Richey, M., & Dawson, S. (2021). Assessing program-level learning strategies in MOOCs. *Computers in Human Behavior*, 117, 106674.
- Barton, M. D., & Heiman, J. R. (2012). Process, product, and potential: The archaeological assessment of collaborative, wiki-based student projects in the technical communication classroom. *Technical Communication Quarterly*, 21(1), σσ. 46-60.
- Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., & Berta, R. (2013). Assessment in and of serious games: an overview. *Advances in human-computer interaction*, 2013.
- Bennett, R. E. (1998). Reinventing Assessment. Speculations on the Future of Large-Scale Educational Testing. *A Policy Information Perspective*.
- Bennett, R. E. (2010). Technology for large-scale assessment. Στο P. L. Peterson, E. Baker, & B. McGaw (Επιμ.), *International encyclopedia of education* (3 εκδ., Τόμ. 8, σσ. 48-55). Oxford, England: Elsevier.
- Bennett, R. E. (2015). The changing nature of educational assessment. *Review of Research in Education*, 39(1), σσ. 370-407.
- Bhalerao, A., & Ward, A. (2001). Towards electronically assisted peer assessment: a case study. *ALT-j*, 9(1), σσ. 26-37.
- Bible, L., Simkin, M. G., & Kuechler, W. L. (2008). Using multiple-choice tests to evaluate students' understanding of accounting. *Accounting Education: an international journal*, 17(S1), σσ. S55-S68.

- Bienkowski, M., Feng, M., & Means, B. (2021). *Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics: An Issue Brief*. Center for Technology in Learning, US Department of Education Office of Educational Technology . ERIC.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., & Rumble, M. (2012). Defining twenty-first century skills. Στο *Assessment and teaching of 21st century skills* (σσ. 17-66). Springer, Dordrecht.
- Blanco, M., & Ginovart, M. (2012). On how moodle quizzes can contribute to the formative e-assessment of first-year engineering students in mathematics courses. *RUSC, Universities and Knowledge Society Journal*, 9(1), σσ. 354-370.
- Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. *Handbook 1: Cognitive domain*.
- Bloom, B. S., Krathwohl, D. R., & Masia, B. B. (1984). Bloom taxonomy of educational objectives. Στο *Allyn and Bacon*. Pearson Education.
- Boitshwarelo, B., Reedy, A. K., & Billny, T. (2017). Envisioning the use of online tests in assessing twenty-first century learning: a literature review. *Research and Practice in Technology Enhanced Learning*, 12(1).
- Botelho, A. F., Baker, R. S., & Heffernan, N. T. (2017). Improving sensor-free affect detection using deep learning. *International conference on artificial intelligence in education* (σσ. 40-51). Springer.
- Bouckaert, R. R., Frank, E., Hall, M., Krikby, R., Reutmann, P., Seewald, A., & Scuse, D. (2020, 12 21). WEKA Manual for Version 3-8-5. Hamilton, New Zealand: University of Waikato.
- Boyle, A., & Hutchison, D. (2009). Sophisticated tasks in e-assessment: what are they and what are their benefits? *Assessment & Evaluation in Higher Education*, 34(3), σσ. 305-319.
- Brady, A. M. (2005). Assessment of learning with multiple-choice questions. *Nurse Education in Practice*, 5(4), σσ. 238-242.
- Brink, R., & Lautenbach, G. (2011). Electronic assessment in higher education. *Educational Studies*, 37(5), σσ. 503-512.
- Broadfoot, P., Oldfield, A., Sutherland, R., & Timmis, S. (2014). Seeds of change: The potential of the digital revolution to promote enabling assessment. Στο C. Wyatt-Smith, V. Klenowski, & P. Colbert (Επιμ.), *Designing Assessment of Quality Learning* (σσ. 373-386). Dordrecht: Springer.
- Brown, G. (2001). Assessment series no 3-assessment: a guide for lectures. *York: Learning and Teaching Support Network (LTNS)*.

- Brown, S., Race, P., & Bull, J. (1999). *Computer-Assisted Assessment in Higher Education. Staff and Educational Development Series*. Stylus Publishers Inc., PO Box 605, Herndon, VA 20172-0605.
- Brownlee, J. (2019, Δεκέμβριος 11). *How to Normalize and Standardize Your Machine Learning Data in Weka*. Ανάκτηση Νοέμβριος 22, 2022, από Machine Learning Mastery: <https://machinelearningmastery.com/normalize-standardize-machine-learning-data-weka/>
- Brownlee, J. (2019, Δεκέμβριος 13). *How to Perform Feature Selection With Machine*. Ανάκτηση Ιανουάριος 2, 2023, από Machine Learning Mastery: <https://machinelearningmastery.com/perform-feature-selection-machine-learning-data-weka/>
- Buckles, S., & Siegfried, J. J. (2006). Using multiple-choice questions to evaluate in-depth learning of economics. *The Journal of Economic Education*, 37(1), σσ. 48-57.
- Bull, J., & Danson, M. (2004). *Assessment Series No14 Computer Assisted Assessment (CAA)*. York: Learning and Teaching Support Network (LTSN).
- Bull, J., Danson, M., Conole, G., Davis, H., White, S., & Sclater, N. (2002). Rethinking Assessment Through Learning Technologies. *Proceedings of Winds of Change in the Sea of Learning: Charting the Course of Digital Learning (ASCILITE 2002) UNITEC, Institute of Technology, Auckland, New Zealand*.
- Bull, S. (2004). Supporting learning with open learner models. *Planning*, 29(14), σ. 1.
- Bunderson, C. V., Inouye, D. K., & Olsen, J. B. (1989). *The four generations of computerized educational measurement*. American Council on Education.
- Burton, R. F. (2005). Multiple-choice and true/false tests: myths and misapprehensions. *Assessment & Evaluation in Higher Education*, 30(1), σσ. 65-72.
- Burton, S. J., Sudweeks, R. R., Merrill, P. F., & Wood, B. (1991). *How to prepare better multiple-choice test items: Guidelines for university faculty*. Provo, UT: Brigham Young University Testing Services.
- Buschetto Macarini, L. A., Cechinel, C., Batista Machado, M. F., Faria Culmant Ramos, V., & Munoz, R. (2019). Predicting students success in blended learning—evaluating different interactions inside learning management systems. *Applied Sciences*, 9(24), 5523.
- Bush, M. (2015). Reducing the need for guesswork in multiple-choice tests. *Assessment & Evaluation in Higher Education*, 40(2), σσ. 218-231.
- Bydzovska, H. (2015). Are collaborative filtering methods suitable for student performance prediction? *Portuguese conference on artificial intelligence*, 425-430.

- Bydzovska, H. (2015). Student performance prediction using collaborative filtering methods. *International Conference on Artificial Intelligence in Education*, 550-553.
- Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics versus machine learning. *Nature Methods*, 233-234.
- Calvo-Flores, M. D., Galindo, E. G., Jimenez, M. P., & Pineiro, O. P. (2006). Predicting students' marks from Moodle logs using neural network models. *Current Developments in Technology-Assisted Education*, 1(2), 586-590.
- Caple, H., & Bogle, M. (2013). Making group assessment transparent: what wikis can contribute to collaborative projects. *Assessment & Evaluation in Higher Education*, 38(2), σσ. 198-200.
- Carabantes, M. (2020). Black-box artificial intelligence: an epistemological and critical analysis. *AI & Society*, 35(2), 309-317.
- Carless, D., Salter, D., Yang, M., & Lam, J. (2001). Developing sustainable feedback practices. *Studies in higher education*, 36(4), σσ. 395-407.
- Chamberlain Jr, D. (2019). *Leveraging Cognitive Theory to Create Large-Scale Learning Tools*. University of Florida.
- Chen, F., & Cui, Y. (2020). Utilizing Student Time Series Behaviour in Learning Management Systems for Early Prediction of Course Performance. *Journal of Learning Analytics*, 7(2), 1-17.
- Chiou, C.-K., Hwang, G.-J., & Tseng, J. C. (2009). An auto-scoring mechanism for evaluating problem-solving ability in a web-based learning environment. *Computers & Education*, 53(2), σσ. 261-272.
- Chollet, F. (2017). *Deep Learning with Python*. Manning Publications.
- Chopard, K., & Przybylski, R. (2021). *Methods Brief: Case Studies*. Justice Research and Statistics Association, US Department of Justice. Ανάκτηση Δεκέμβριος 07, 2022, από <https://www.jrsa.org/pubs/factsheets/jrsa-research-methods-brief-case-studies.pdf>
- Chui, K. T., Fung, D. C., Lytras, M. D., & Lam, T. M. (2020). Predicting at-risk university students in a virtual learning environment via a machine learning algorithm. *Computers in Human Behavior*, 107, 105584.
- Chung, G. K., Shel, T., & Kaiser, W. J. (2006). An exploratory study of a novel online formative assessment and instructional tool to promote students' circuit problem solving. *The Journal of Technology, Learning and Assessment*, 5(6).

- Clarke, N. L., Dowland, P., & Furnell, S. M. (2013). e-Invigilator: A biometric-based supervision system for e-Assessments. Στο *International Conference on Information Society (i-Society 2013)* (σσ. 238-242). IEEE.
- Clegg, V. L., & Cashin, W. E. (1986). *Improving multiple-choice tests*. Center for Faculty Evaluation & Development, Kansas State University.
- Coelho, O. B., & Silveira, I. (2017). Deep learning applied to learning analytics and educational data mining: A systematic literature review. *Brazilian Symposium on Computers in Education (Simposio Brasileiro de Informatica na Educacao-SBIE)*, 28, σ. 143.
- Cohen, A. (2017). Analysis of student activity in web-supported courses as a tool for predicting dropout. *Educational Technology Research and Development*, 65, 1285-1304.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Μεθοδολογία εκπαιδευτικής έρευνας*. Αθήνα: Μεταίχμιο.
- Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2016). Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *IEEE Transactions on Learning Technologies*, 10(1), 17-29.
- Connolly, T. M., Stansfield, M., & Hainey, T. (2007). An application of games-based learning within software engineering. *British Journal of Educational Technology*, 38(3), σσ. 416-428.
- Conole, G., & Warburton, B. (2005). A review of computer-assisted assessment. *ALT-J*, 13(1), σσ. 17-31.
- Conrad, D., & Openo, J. (2018). *Assessment strategies for online learning: Engagement and authenticity*. Athabasca University Press.
- Cook, D. A., Hamstra, S. J., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., . . . Hatala, R. (2013). Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. *{Medical teacher*, 35(1), σσ. e867-e898.
- Cook, J., & Jenkins, V. (2010). *Getting started with e-assessment*. Bath: University of Bath.
- Costa, E. B., Fonseca, B., Santana, M. A., de Araujo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in human behavior*, 73, 247-256.
- Cotterill, S., Bradley, P., & Hammond, G. (2006). ePortfolios: Supporting assessment in complex educational environments. Στο C. Bryan, & K. Clegg (Επιμ.), *Innovative assessment in higher education* (σσ. 191-199). New York: Routledge.

- Crisp, G. (2009). Interactive e-Assessment: moving beyond multiple-choice questions. *Centre for Learning and Professional Development. Adelaide: University of Adelaide, 3*, σσ. 12-31.
- Crisp, G. (2011). *Teacher's Handbook on e-Assessment*. An ALTC Fellowship Activity, 18.
- Cui, Y., Chen, F., Shiri, A., & Fan, Y. (2019). Predictive analytic models of student success in higher education: A review of methodology. *Information and Learning Sciences, 208-227*.
- Daly, C., Pachler, N., Mor, Y., & Mellar, H. (2010). Exploring formative e-assessment: using case stories and design patterns. *Assessment & Evaluation in Higher Education, 35*(5), σσ. 619-636.
- Davies, L. J., & Howard, R. M. (2016). Plagiarism and the internet: fears, facts, and pedagogies. *Handbook of academic integrity, 591*.
- Dawson, S. P., McWilliam, E., & Tan, J. P.-L. (2008). Teaching smarter: How mining ICT data can inform and improve learning and teaching practice. *Proc. Annu. Conf. Australasian Soc. Comput.*, (σσ. 221-230).
- Dawson, S., Mirriahi, N., & Gasevic, D. (2015). Importance of theory in learning analytics in formal and workplace settings. *Journal of Learning Analytics, 2*(2), 1-4.
- de Sande, J. C. (2015). Calculated questions and e-cheating: A case study. Στο M. Carmo (Επιμ.), *Education Applications & Developments* (σσ. 91-99). Lisboa, Portugal: inScience Press.
- De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2011). Assessing collaboration in a wiki: The reliability of university students' peer assessment. *The Internet and Higher Education, 14*(4), σσ. 201-206.
- Dewan, M. A., Lin, F., & Wen, D. (2015). Predicting dropout-prone students in E-learning education system. *2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)* (σσ. 1735-1740). IEEE.
- Dohn, N. B. (2009). Web 2.0: Inherent tensions and evident challenges for education. *International journal of computer-supported collaborative learning, 4*(3), σσ. 343-363.
- Doleck, T., Lemay, D. J., Basnet, R. B., & Bazelais, P. (2020). Predictive analytics in education: a comparison of deep learning frameworks. *Education and Information Technologies, 25*(3), 1951-1963.
- Doleck, T., Poitras, E., & Lajoie, S. (2019). Assessing the utility of deep learning: Using learner-system interaction data from BioWorld. *EdMedia+ Innovate Learning, 734-738*.



- Dollinger, S. J., Matyja, A. M., & Huber, J. L. (2008). Which factors best account for academic success: Those which college students can control or those they cannot? *Journal of research in Personality*, 42(4), 872-885.
- Donnelly, C. (2014). The use of case based multiple choice questions for assessing large group teaching: implications on student's learning. *Irish Journal of Academic Practice*, 3(1), σ. 12.
- Douglas, M., Wilson, J., & Ennis, S. (2002). Multiple-choice question tests: a convenient, flexible and effective learning tool? A case study. *Innovations in Education and Teaching International*, 49(2), σσ. 111-121.
- Dowell, N. M., & Poquet, O. (2021). SCIP: Combining group communication and interpersonal positioning to identify emergent roles in scaled digital environments. *Computers in Human Behavior*, 106709.
- Downing, S. M. (2003). Guessing on selected-response examinations. *Medical Education*, 37(8), σσ. 670-671.
- Draper, S. W. (2009). *British Journal of Educational Technology*, 40(2), σσ. 285-293.
- Elliott, B. (2008). Assessment 2.0: Modernising assessment in the age of Web 2.0. *Scottish Qualifications Authority*, 28.
- ePortConsortium.Org. (2003). *Electronic Portfolio White Paper [Version 1.0]*.
- Evale, D. S. (2017). Learning management system with prediction model and course-content recommendation module. *Journal of Information Technology Education: Research*, 16, 437.
- Farrell, T., & Rushby, N. (2016). Assessment and learning technologies: An overview. *British Journal of Educational Technology*, 47(1), σσ. 106-120.
- Fei, M., & Yeung, D.-Y. (2015). Temporal models for predicting student dropout in massive open online courses. *2015 IEEE international conference on data mining workshop (ICDMW)* (σσ. 256-263). IEEE.
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5-6), σσ. 304-317.
- Ferrao, M. (2010). E-assessment within the Bologna paradigm: evidence from Portugal. *Assessment & Evaluation in Higher Education*, 35(7), σσ. 819-830.
- Ferrell, G., & Sheppard, M. (2012). *A view of the Assessment and Feedback Landscape: baseline analysis of policy and practice from the JISC Assessment & Feedback programme*.  
Ανάκτηση Απρίλιος 19, 2020, από  
<https://pdfs.semanticscholar.org/4a7d/a4408e7d59963e74aa90d8a4a306d70ed35e.pdf>

- Fiedler, R. L., & Kaner, C. (2010). Plagiarism-detection services: How well do they actually perform? *IEEE technology and society magazine*, 29(4), σσ. 37-43.
- Fontanillas, T. R., Carbonell, M. R., & Catusus, M. G. (2016). E-assessment process: giving a voice to online learners. *International Journal of Educational Technology in Higher Education*, 13(1), σ. 20.
- Frank, E. (2014). *Logistic VS Simple Logistic*. Ανάκτηση Ιανουάριος 16, 2023, από Narkive MailingList Archive: <https://wekalist.scms.waikato.ac.narkive.com/6ukR9FIE/logistic-vs-simple-logistic#post1>
- Frank, E. (2017, Δεκέμβριος 20). *Converting univariate time series data for time series classification with WEKA*. Ανάκτηση Δεκέμβριος 12, 2022, από Weka machine learning workbench list: <https://list.waikato.ac.nz/hyperkitty/list/wekalist@list.waikato.ac.nz/thread/KYHGRP2WPWHF5S4H3KOWM5ZDWQ55VKWJ/#GADFQL4MS7DE6WQ52TOYTZLLAJQTN5BA>
- Frank, E. (2018, Μάρτιος 24). *Choosing attribute selection with filter methods*. Ανάκτηση Ιανουάριος 2, 2023, από Weka machine learning workbench list: <https://list.waikato.ac.nz/hyperkitty/list/wekalist@list.waikato.ac.nz/thread/6BBFAAKJV4NMEW5JYJANYVQJTV4Y2UFN/#Q3O773EAA5U5IIRPQX6HRYTYUBFFV7P>
- Frank, E. (2018, Μάιος 13). *CorrelationAttributeEval and InfoGainAttributeEval*. Ανάκτηση Ιανουάριος 2, 2023, από Weka machine learning workbench list: <https://list.waikato.ac.nz/hyperkitty/list/wekalist@list.waikato.ac.nz/thread/TIYMM2JFU7YPTTRLCS6OQTDXCVCVZWHZ7/#EN7CSBNCTVRVFUE63RJHLM6VY2SDMWKL>
- Frank, E. (2021, Απρίλιος 16). *WekaDeepLearning4j classifier is greyed out*. Ανάκτηση Δεκέμβριος 23, 2022, από Weka machine learning workbench list: <https://list.waikato.ac.nz/hyperkitty/list/wekalist@list.waikato.ac.nz/message/RAC5AL2LYJIAYIYS6LCTUYA6QXKGGAN5/>
- Frank, E., Hall, M. A., & Witten, I. H. (2016). *The WEKA Workbench (4th ed.)*. Morgan Kaufmann.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, 74(1), 59-109.
- Gao, Q. (2012). Using IP addresses as assisting tools to identify collusions. *International Journal of Business, Humanities and Technology*, σσ. 70-75.
- Gardner-Medwin, A. R. (2006). Confidence-based marking: towards deeper learning and better exams. Στο C. Bryan, & K. Clegg (Επιμ.), *Innovative Assessment in Higher Education* (σσ. 141-149). London: Routledge.

- Gardner-Medwin, T. (2019). Certainty-based marking: Stimulating thinking and improving objective tests. Στο C. Bryan, & K. Clegg (Επιμ.), *Innovative Assessment in Higher Education: A Handbook for Academic Practitioners* (σσ. 141-150). London: Routledge.
- Gasevic, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education, 28*, 68-84.
- Gavsevic, D., Greiff, S., & Shaffer, D. W. (2022). Towards strengthening links between learning analytics and assessment: Challenges and potentials of a promising new bond. *Computers in Human Behavior, 107304*.
- Gibson, D., Ostaszewski, N., Flintoff, K., Grant, S., & Knight, E. (2015). Digital badges in education. *Education and Information Technologies, 20*(2), σσ. 403-410.
- Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & education, 57*(4), σσ. 2333-2351.
- Gnanambal, S., Thangaraj, M., Meenatchi, V. T., & Gayathri, V. (2018). Classification algorithms with attribute selection: an evaluation study using WEKA. *International Journal of Advanced Networking and Applications, 9*(6), 3640-3644.
- Green, J., Wyllie, A., & Jackson, D. (2014). Electronic portfolios in nursing education: A review of the literature. *Nurse education in practice, 14*(1), σσ. 4-8.
- Guarin, C. E., Guzman, E. L., & Gonzalez, F. A. (2015). A model to predict low academic performance at a specific enrollment using data mining. *IEEE Revista Iberoamericana de tecnologias del Aprendizaje, 10*(3), 119-125.
- Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI magazine, 40*(2), 44-58.
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied measurement in education, 15*(3), σσ. 309-333.
- Hall, M. (2017, Μάιος 2). *CorrelationAttributeEval*. Ανάκτηση Ιανουάριος 2, 2023, από Weka machine learning workbench list:  
<https://list.waikato.ac.nz/hyperkitty/list/wekalist@list.waikato.ac.nz/thread/XTNEUTYYL EYMS5X4XI5CO5TDSEPQFSGP/#XTNEUTYYLEYMS5X4XI5CO5TDSEPQFSGP>
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA Data Mining Software: An Update. *ACM SIGKDD explorations newsletter, 11*(1), 10-18.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques (3rd Ed)*. Morgan Kaufmann.

- Harmon, O. R., Lambrinos, J., & Buffolino, J. (2010). Assessment design and cheating risk in online instruction. *Online Journal of Distance Learning Administration, 13*(3).
- Hemming, A. (2010). Online tests and exams: lower standards or improved learning? *The Law Teacher, 44*(3), σσ. 283-308.
- Henrie, C. R., Bodily, R., Larsen, R., & Graham, C. R. (2018). Exploring the potential of LMS log data as a proxy measure of student engagement. *Journal of Computing in Higher Education, 30*(2), 344-362.
- Heron, G., & Lerpiniere, J. (2013). Re-engineering the multiple choice question exam for social work. *European Journal of Social Work, 16*(4), σσ. 521-535.
- Herrington, J., McKenney, S., Reeves, T., & Oliver, R. (2007). on, J., McKenney, S., Reeves, T., & Oliver, R. (2007, June). Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. Στο C. Montgomerie, & J. Seale (Επιμ.), *Proceedings of World Conference on Educational Multimedia*, (σσ. 4089-4097). Association for the Advancement of Computing in Education (AACE).
- Heyde, V. v., & Siebrits, A. (2019). Higher-order e-assessment for physics in the digital age using Sakai. *The Physics Teacher, 57*(1), σσ. 32-34.
- Holmes, N. (2015). Student perceptions of their learning and engagement in response to the use of a continuous e-assessment in an undergraduate module. *Assessment & Evaluation in Higher Education, 40*(1), σσ. 1-14.
- Howard, E., Meehan, M., & Parnell, A. (2018). Contrasting prediction methods for early warning systems at undergraduate level. (Elsevier, Επιμ.) *The Internet and Higher Education, 37*, 66-75.
- Hu, Y.-H., Lo, C.-L., & Shih, S.-P. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior, 36*, 469-478.
- Hussain, M., Zhu, W., Zhang, W., & Abidi, S. (2018). Student engagement predictions in an e-learning system and their impact on student course assessment scores. *Computational intelligence and neuroscience*.
- Hussain, M., Zhu, W., Zhang, W., Abidi, S. M., & Ali, S. (2019). Using machine learning to predict student difficulties from learning session data. *Artificial Intelligence Review, 52*, 381-407.
- Hussain, S., Dahan, N. A., Ba-Alwib, F. M., & Ribata, N. (2018). Educational data mining and analysis of students' academic performance using WEKA. *Indonesian Journal of Electrical Engineering and Computer Science, 9*(2), 447-459.

- Ibbett, N. L., & Wheldon, B. J. (2016). The Incidence of Clueing in Multiple Choice Testbank Questions in Accounting: Some Evidence from Australia. *e-Journal of Business Education and Scholarship of Teaching*, 10(1), σσ. 20-35.
- International Conference on Learning Analytics & Knowledge (LAK)*. (2011). Ανάκτηση Ιανουάριος 2, 2023, από Society for Learning Analytics Research (SoLAR): <https://www.solaresearch.org/events/lak/>
- Jafari, A. (2004). The "Sticky" ePortfolio System: Tackling Challenges and Identifying Attribute. *EDUCAUSE Review*, 39(4), σσ. 38-49.
- James, R. (2016). Tertiary student attitudes to invigilated, online summative examinations. *International Journal of Educational Technology in Higher Education*, 13(1), σ. 19.
- James, R., McInnis, C., & Devlin, M. (2002). *Assessing learning in Australian universities: Ideas, strategies and resources for quality in student assessment*. Australian, Universities Teaching Committee.
- Jayaprakash, S. M., Moody, E. W., Lauria, E. J., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, 1(1), 6-47.
- JISC. (2006). *e-Assessmnet Glossary (Extended) v1.1*. Ανάκτηση Απρίλιος 15, 2020, από [https://www.webarchive.org.uk/wayback/archive/20140614112453mp\\_/http://www.jisc.ac.uk/media/documents/themes/elearning/eassess\\_glossary\\_extendedv101.pdf](https://www.webarchive.org.uk/wayback/archive/20140614112453mp_/http://www.jisc.ac.uk/media/documents/themes/elearning/eassess_glossary_extendedv101.pdf)
- JISC. (2007). *Effective Practice with e-Assessment*. Ανάκτηση Απρίλιος 16, 2020, από [https://www.webarchive.org.uk/wayback/en/archive/20130414000900mp\\_/http://www.jisc.ac.uk/media/documents/themes/elearning/effpraceassess.pdf](https://www.webarchive.org.uk/wayback/en/archive/20130414000900mp_/http://www.jisc.ac.uk/media/documents/themes/elearning/effpraceassess.pdf)
- JISC. (2010). *Effective Assessment in a Digital Age*. Ανάκτηση Απρίλιος 17, 2020, από [https://www.webarchive.org.uk/wayback/archive/20140613220103/http://www.jisc.ac.uk/media/documents/programmes/elearning/digiassass\\_eada.pdf](https://www.webarchive.org.uk/wayback/archive/20140613220103/http://www.jisc.ac.uk/media/documents/programmes/elearning/digiassass_eada.pdf)
- JISC. (2015). *Learning and teaching considerations*. Ανάκτηση Μάιος 28, 2020, από <https://www.jisc.ac.uk/guides/mobile-learning/learning-and-teaching-considerations>
- Joksimovic, S., Gavsevic, D., Loughin, T. M., Kovanovic, V., & Hatala, M. (2015). Learning at distance: Effects of interaction traces on academic achievement. *Computers & Education*, 204-2017.
- Joksimovic, S., Kovanovic, V., & Dawson, S. (2019). The journey of learning analytics. *HERDSA Review of Higher Education*, 6, 27-63.
- Jones, S. M. (2014). Assessing the Science Knowledge of University Students: Perils, Pitfalls and Possibilities. *Journal of Learning Design*, 7(2), σσ. 16-27.

- Jordan, S. (2013). E-assessment: Past, present and future. *New Directions in the Teaching of Physical Sciences*, 9, σσ. 87-106.
- Kapsalis, G., Ferrari, A., Punie, Y., Conrads, J., Collado, A., Hotulainen, R., . . . Ilsley, P. (2019). *Evidence of innovative assessment: Literature review and case studies*. Joint Research Centre (Seville site).
- Kehrwald, B. (2010). Being online: Social presence as subjectivity in online learning. *London Review of Education*, 8(1), σσ. 39-50.
- Kereluik, K., Mishra, P., Fahnoe, C., & Terry, L. (2013). What knowledge is of most worth: Teacher knowledge for 21st century learning. *Journal of digital learning in teacher education*, 29(4), σσ. 127-140.
- Kerlinger, F. N. (1970). *Foundations of behavioral research*. New York: Holt, Rinehart & Winston.
- Kiennert, C., Rocher, P.-O., Ivanova, M., Rozeva, A., Durcheva, M., & Garcia-Alfaro, J. (2017). Security challenges in e-assessment and technical solutions. Στο *2017 21st International Conference Information Visualisation (IV)* (σσ. 366-371). IEEE.
- Kim, B.-H., Vizitei, E., & Ganapathi, V. (2018). GritNet: Student performance prediction with deep learning. *arXiv preprint arXiv:1804.07405*.
- Knight, P. (2001). *A briefing on key concepts: Formative and summative, criterion and norm-referenced assessment*. Ανάκτηση Απρίλιος 24, 2020, από [https://blogs.shu.ac.uk/teaching/files/2016/09/No7\\_briefing\\_on\\_key\\_concepts.docx](https://blogs.shu.ac.uk/teaching/files/2016/09/No7_briefing_on_key_concepts.docx)
- Knight, S., Buckingham Shum, S., & Littleton, K. (2013). Epistemology, pedagogy, assessment and learning analytics. *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, (σσ. 75-84).
- Koh, L. C. (2008). Refocusing formative feedback to enhance learning in pre-registration nurse education. *Nurse Education in Practice*, 8(4), σσ. 223-230.
- Kotsiantis, S. B. (2012). Use of machine learning techniques for educational proposes: a decision support system for forecasting students' grades. *Artificial Intelligence Review*, 37, 331-344.
- Kotsiantis, S., Pierrakeas, C., & Pintelas, P. (2004). Predicting students' performance in distance learning using machine learning techniques. *Applied Artificial Intelligence*, 18(5), 411-426.
- Kovacic, Z. (2010). Early prediction of student success: Mining students' enrolment data.
- Kovanovic, V. a. (2015). Penetrating the black box of time-on-task estimation. *Proceedings of the fifth international conference on learning analytics and knowledge*, (σσ. 184-193).

- Kuzilek, J., Hlosta, M., Herrmannova, D., Zdrahal, Z., Vaclavek, J., & Wolff, A. (2015). OU Analyse: analysing at-risk students at The Open University. *Learning Analytics Review*, 1-16.
- Lafuente, M., Remesal, A., & Álvarez Valdivia, I. M. (2014). Assisting learning in e-assessment: a closer look at educational supports. *Assessment & Evaluation in Higher Education*, 39(4), σσ. 443-460.
- Lang, J. M. (2013). *Cheating Lessons: Learning from academic dishonesty*. Cambridge, Massachusetts: Harvard University Press.
- Larreamendy-Joerns, J., & Leinhardt, G. (2006). Going the Distance With Online Education. *Review of Educational Research*, 76(4), σσ. 567-605.
- Lauria, E. J., Baron, J. D., Devireddy, M., Sundararaju, V., & Jayaprakash, S. M. (2012). Mining academic data to improve college student retention: An open source perspective. *Proceedings of the 2nd international conference on learning analytics and knowledge*, (σσ. 139-142).
- Lee, H.-S., Liu, O. L., & Linn, M. C. (2011). Validating measurement of knowledge integration in science using multiple-choice and explanation items. *Applied Measurement in Education*, 24(2), σσ. 115-136.
- Leitner, P., Mohammad, K., & Martin, E. (2017). Learning analytics in higher education - a literature review. Στο A. Pena-Ayala (Επιμ.). Springer.
- Leppanen, L., Leinonen, J., Ihantola, P., & Hellas, A. (2017). Predicting academic success based on learning material usage. *SIGITE 17: Proceedings of the 18th Annual Conference on Information Technology Education*, (σσ. 13-18).
- Leung, S. F., Mok, E., & Wong, D. (2008). The impact of assessment methods on the learning of nursing students. *Nurse education today*, 28(6), σσ. 711-719.
- Levine, J., & Pazdernik, V. (2018). Evaluation of a four-prong anti-plagiarism program and the incidence of plagiarism: a five-year retrospective study. *Assessment & Evaluation in Higher Education*, 43(7), σσ. 1094-1105.
- Li, X., Chang, K.-m., Yuan, Y., & Hauptmann, A. (2015). Massive open online proctor: Protecting the credibility of MOOCs certificates. *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, (σσ. 1129-1137).
- Liang, J., Li, C., & Zheng, L. (2016). Machine learning application in MOOCs: Dropout prediction. *2016 11th International conference on computer science & education (ICCSE)* (σσ. 52-57). IEEE.

- Lilley, M., Barker, T., & Britton, C. (2004). The development and evaluation of a software prototype for computer-adaptive testing. *Computers & Education, 43*(1-2), σσ. 109-123.
- Linacre, J. M. (2000). *Computer-Adaptive Testing: A methodology whose time has come*. Chicago: MESA Psychometric Laboratory, University of Chicago.
- Linan, L. C., & Perez, A. A. (2015). Educational Data Mining and Learning Analytics: differences, similarities, and time evolution. *RUSC. Universities and Knowledge Society Journal, 12*(3), 98-112.
- Liu, Y., Fan, S., Xu, S., Sajjanhar, A., Yeom, S., & Wei, Y. (2022). Predicting Student Performance Using Clickstream Data and Machine Learning. *Education Sciences, 13*(1), 17.
- Livingston, S. A. (2009). Constructed-Response Test Questions: Why We Use Them; How We Score Them. R&D Connections. Number 11. *Educational Testing Service*.
- Ljubobratovic, D., & Matetic, M. (2019). Using LMS Activity Logs to Predict Student Failure with Random Forest Algorithm. *The Future of Information Sciences, 113*.
- Lopez-Zambrano, J., Lara, J. A., & Romero, C. (2020). Towards portability of models for predicting students' final performance in university courses starting from moodle logs. *Applied Sciences, 10*(1), 354.
- Lorenzo, G., & Ittelson, J. (2005). *Demonstrating and Assessing Student Learning with E-Portfolios*. EDUCAUSE Learning Initiative.
- Lu, O. H., Huang, A. Y., Huang, J. C., Lin, A. J., Ogata, H., & Yang, S. J. (2018). Applying learning analytics for the early prediction of Students' academic performance in blended learning}. *Journal of Educational Technology & Society, 21*(2), 220-232.
- Luan, J., & Zhao, C. M. (2006). Practicing data mining for enrollment management and beyond. *New directions for institutional research, 131*, 117-122.
- Ludwig-Hardman, S., & Dunlap, J. C. (2003). Learner support services for online students: Scaffolding for success. *he International Review of Research in Open and Distributed Learning, 4*(1), σσ. 1-15.
- Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual review of public health, 23*(1), 151-169.
- Luo, L., Koprinska, I., & Liu, W. (2015). Discrimination-Aware Classifiers for Student Performance Prediction. *International Educational Data Mining Society*.
- Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an “early warning system” for educators: A proof of concept. *Computers & education, 54*(2), 588-599.



- Machine Learning Glossary*. (χ.χ.). Ανάκτηση Νοέμβριος 23, 2022, από <https://developers.google.com/machine-learning/glossary/#attribute>
- Marais, E., Argles, D., & von Solms, B. (2006). Security issues specific to e-assessments. *The International Journal for Infonomics*.
- Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Computers & Education* 103, 1-15.
- Martin, R. (2008). New possibilities and challenges for assessment through the use of technology. *Towards a research agenda on computer-based assessment*, σσ. 6-9.
- Mason, R., Pegler, C., & Weller, M. (2004). E-portfolios: an assessment tool for online courses. *British Journal of Educational Technology*, 35(6), σσ. 717-727.
- Masters, J. C., Hulsmeyer, B. S., Pike, M. E., Leichty, K., Miller, M. T., & Verst, A. L. (2001). Assessment of multiple-choice questions in selected test banks accompanying text books used in nursing education. *Journal of Nursing Education*, 40(1), σσ. 25-32.
- Mathrani, A., Susnjak, T., Ramaswami, G., & Barczak, A. (2021). Perspectives on the challenges of generalizability, transparency and ethics in predictive learning analytics. *Computers and Education Open*, 2, 100060.
- McAllister, D., & Guidice, R. M. (2012). This is only a test: a machine-graded improvement to the multiple-choice and true-false examination. *Teaching in Higher Education*, 17(2), σσ. 193-207.
- McAlpine, M. (2002). *Principles of assessment*. Luton: CAA Centre, University of Luton.
- McCabe, D. L. (2005). Cheating among college and university students: A North American perspective. *International Journal for Educational Integrity*, 1(1).
- Mellar, H., Peytcheva-Forsyth, R., Kocdar, S., Karadeniz, A., & Yovkova, B. (2018). Addressing cheating in e-assessment using student authentication and authorship checking systems: teachers' perspectives. *International Journal for Educational Integrity*, 14(2).
- Miguel, J. C., Xhafa, F., & Prieto, J. (2015). Security in online web learning assessment. *World Wide Web*, 18(6), σσ. 1655-1676.
- Milligan, S. (2020). Standards for developing assessments of learning using process data. *Re-imagining university assessment in a digital world*, 179-192.
- Mora, M. C., Sancho-Bru, J., Iserte, J., & Sanchez, F. (2012). An e-assessment approach for evaluation in engineering overcrowded groups. *Computers & Education*, 59(2), σσ. 732-740.

- Morales-Martinez, G. E., Lopez-Ramirez, E. O., Castro-Campos, C., Villarreal-Trevino, M. G., & Gonzales-Trujillo, C. J. (2017). Cognitive Analysis of Meaning and Acquired Mental Representations as an Alternative Measurement Method Technique to Innovate E-Assessment. *European Journal of Educational Research*, 6(4), σσ. 455-464.
- Morris, L. V., Finnegan, C., & Wu, S.-S. (2005). Tracking student behavior, persistence, and achievement in online courses. *The Internet and Higher Education*, 8(3), 221-231.
- Morrison, S., & Free, K. W. (2001). Writing multiple-choice test items that promote and measure critical thinking. *Journal of Nursing Education*, 40(1), σσ. 17-24.
- Motz, B., Quick, J., Schroeder, N., Zook, J., & Gunkel, M. (2019). The validity and utility of activity logs as a measure of student engagement. *Proceedings of the 9th international conference on learning analytics & knowledge*, (σσ. 300-309).
- Munoz-Organero, M., Munoz-Merino, P. J., & Kloos, C. D. (2009). Student behavior and interaction patterns with an LMS as motivation predictors in e-learning settings}. *IEEE Transactions on Education*, 53(3), 463-470.
- Namoun, A., & Alshantiti, A. (2020). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. *Applied Sciences*, 11(1), 237.
- Nandi, D., Hamilton, M., Harland, J., & Warburton, G. (2011). How active are students in online discussion forums? *Proceedings of the Thirteenth Australasian Computing Education Conference-Volume 114*, (σ. 125).
- Navamani, J. M., & Kannammal, A. (2015). Predicting performance of schools by applying data mining techniques on public examination results. *Research Journal of Applied Sciences, Engineering and Technology*, 9(4), 262-271.
- Nespereira, C. G., Elhariri, E., El-Bendary, N., Vilas, A. F., & Redondo, R. P. (2016). Machine learning based classification approach for predicting students performance in blended learning. *The 1st International Conference on Advanced Intelligent System and Informatics (AIS2015), November 28-30, 2015, Beni Suef, Egypt* (σσ. 47-56). Springer.
- Nicol, D. (2008). *Technology-supported Assessment: A Review of Research*. Δημοσιευτη αναφορά διαθέσιμη στο <http://www.reap.ac.uk/resources.html>. Ανάκτηση Απρίλιος 21, 2020, από [https://www.reap.ac.uk/Portals/101/Documents/REAP/Technology\\_supported\\_assessment.pdf](https://www.reap.ac.uk/Portals/101/Documents/REAP/Technology_supported_assessment.pdf)
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), σσ. 199-218.

- Nicol, D., & Milligan, C. (2006). Rethinking technology-supported assessment practices in relation to the seven principles of good feedback practice. Στο *Innovative assessment in higher education* (σσ. 84-98). Routledge.
- Nilson, L. (2015). *Specifications grading: Restoring rigor, motivating students, and saving faculty time*. Sterling, Virginia: Stylus Publishing, LLC.
- Noorberbahani, F., & Kardan, A. A. (2011). The automatic assessment of free text answers using a modified BLEU algorithm. *Computers & Education*, 56(2), σσ. 337-345.
- Ochoa, X. (2016). Adaptive multilevel clustering model for the prediction of academic risk. *2016 XI Latin American Conference on Learning Objects and Technology (LACLO)* (σσ. 1-8). IEEE.
- Okubo, F., Yamashita, T., Shimada, A., & Ogata, H. (2017). A neural network approach for students' performance prediction. *Proceedings of the seventh international learning analytics & knowledge conference*, (σσ. 598-599).
- Oldfield, A., Broadfoot, P., Sutherland, R., & Timmis, S. (2012). *Assessment in a digital age: A research review*. Bristol: University of Bristol.
- Oosterhof, A., Conrad, R.-M., & Ely, D. P. (2008). *Assessing learners online*. Pearson Merrill Prentice-Hall.
- Pachler, N., Daly, C., Mor, Y., & Mellar, H. (2010). Formative e-assessment: Practitioner cases. *Computers & Education*, 54(3), σσ. 715-721.
- Pehlivanova, T. I., & Nedeva, V. I. (2021). Attributes selection using machine learning for analysing students' dropping out of university: a case study. *IOP Conference Series: Materials Science and Engineering*, 1031, σ. 12055.
- Pellegrino, J. W., & Quellmalz, E. S. (2010). Perspectives on the integration of technology and assessment. *Journal of Research on Technology in Education*, 43(2), σσ. 119-134.
- Perera-Diltz, D. M., & Moe, J. L. (2014). Formative and summative assessment in online education. *Journal of research in innovative teaching*, 7(1), σσ. 130-142.
- Perrotta, C., & Whitelock, D. (2017). Assessment for learning. Στο *Technology Enhanced Learning* (σσ. 127-135). Cham: Springer.
- Petrina, S. (2004). Sidney Pressey and the automation of education, 1924-1934. *Technology and Culture*, 45(2), σσ. 305-330.
- Petropoulou, O., Kasimatis, K., Dimopoulos, I., & Retalis, S. (2014). LAe-R: A new learning analytics tool in Moodle for assessing students' performance. *Bulletin of the IEEE Technical Committee on Learning Technology*, 16(1), σ. 1.

- Puentedura, R. (2010). *Ruben R Puentedura's Weblog: Ongoing thoughts on education and technology*. Ανάκτηση Μάιος 25, 2020, από <http://www.hippasus.com/rrpweblog/archives/000049.html>
- Pyper, A., & Lilley, M. (2010). A comparison between the flexilevel and conventional approaches to objective testing. Στο D. Whitelock, W. Warburton, G. Wills, & L. Gilbert (Επιμ.), *2010 International Computer Assisted Assessment (CAA) Conference: Research into e-Assessment, Southampton, UK*.
- Quinn, R. J., & Gray, G. (2020). Prediction of student academic performance using Moodle data from a Further Education setting. *Irish Journal of Technology Enhanced Learning*, 5(1).
- Rafaeli, S., & Ravid, G. (1997). Online, web-based learning environment for an information systems course: Access logs, linearity and performance. *Proc. Inf. Syst. Educ. Conf*, 92-99.
- Raga Jr, R. C., & Raga, J. D. (2017). Monitoring Class Activity and Predicting Student Performance Using Moodle Action Log Data. *International Journal of Computing Sciences Research*, 1(3), 1-16.
- Rastrollo-Guerrero, J. L., Gomez-Pulido, J. A., & Duran-Dominguez, A. (2020). Analyzing and predicting students' performance by means of machine learning: A review. *Applied sciences*, 10(3), 1042.
- Redecker, C. (2013). The use of ICT for the assessment of key competences. *Joint Research Centre Institute for Prospective Technological Studies. Scientific and Technical Research series*.
- Ren, Q., Li, M., & Han, S. (2019). Ren, Q., Li, M., & Han, S. (2019). Tectonic discrimination of olivine in basalt using data mining techniques based on major elements: a comparative study from multiple perspectives. *Big Earth Data*, 3(1), 8-25.
- Ridgway, J., McCusker, S., & Pead, D. (2004). *Literature review of e-assessment*. Bristol: Futurelab.
- Rienties, B., Toetel, L., & Bryan, A. (315-319). Scaling up" learning design: impact of learning design activities on LMS behavior and performance. *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*, (σ. 2015).
- Riestra-Gonzalez, M., del Puerto Paule-Ruiz, M., & Ortin, F. (2021). Massive LMS log data analysis for the early prediction of course-agnostic student performance. *Computers & Education*, 163, 104-108.
- Robles, S., Pons, J., Borrell, J., Freixas, M., Fernandez-Cordoba, C., Sebastian, R., & Pelachs, A. (2008). New wiki tools for on-line formative assessment. *Sciences*, 2(2), σ. 1.

- Rodriguez, M. C. (2002). Choosing an item format. *Large-scale assessment programs for all students: Validity, technical adequacy, and implementation*, σσ. 213-231.
- Romero, C. a., Zafra, A., Romero, J. R., & Ventura, S. (2013). Web usage mining for predicting final marks of students that use Moodle courses. *Computer Applications in Engineering Education*, 21(1), 135-146.
- Rosewell, J. P. (2011). Opening up multiple-choice: assessing with confidence. Στο D. Whitelock, W. Warburton, G. Wills, & L. Gilbert (Επιμ.), *2011 International Computer Assisted Assessment (CAA) Conference: Research into e-Assessment, 5/6 July 2011, Southampton, UK*.
- Rudner, L. M., & Gagne, P. (2000). An overview of three approaches to scoring written essays by computer. *Practical Assessment, Research, and Evaluation*, 7(1), σ. 26.
- Russell, J., Elton, L., Swinglehurst, D., & Greenhalgh, T. (2006). Using the online environment in assessment for learning: a case-study of a web-based course in primary care. *Assessment & Evaluation in Higher Education*, 31(4), σσ. 465-478.
- Rust, C., O'Donovan, B., & Price, M. (2005). A social constructivist assessment process model: how the research literature shows us this could be best practice. *Assessment & Evaluation in Higher Education*, 30(3), σσ. 231-240.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, 18(2), σσ. 119-114.
- Sawyer, B., & Rejeski, D. (2002). *Serious games: Improving public policy through game-based learning and simulation*.
- Schell, J., Lukoff, B., & Alvarado, C. (2014). Using early warning signs to predict academic risk in interactive, blended teaching environments. *Internet Learn.*, 3(2).
- Schmelkin, L. P., Gilbert, K., Spencer, K. J., Pincus, H. S., & Silva, R. (2008). A multidimensional scaling of college students' perceptions of academic dishonesty. *The Journal of Higher Education*, 79(5), σσ. 587-607.
- Schultz, M. (2011). Sustainable Assessment for Large Science Classes: Non-Multiple Choice, Randomised Assignments through a Learning Management System. *Journal of Learning Design*, 4(3), σσ. 50-62.
- Sclater, N., Peasgood, A., & Mullan, J. (2016). *Learning analytics in higher education. A review of UK and international practise*. JISC.
- Scott, G. (2016). *Assuring the quality of achievement standards and their valid assessment in Australian Higher Education*. Australian Government, Department of Education and

- Training. Ανάκτηση Ιούλιος 26, 2020, από <http://flipcurric.edu.au/sites/flipcurric/media/109.pdf>
- Shaul, M. (2008). Assessing online discussion forum participation. Στο *Information communication technologies: Concepts, methodologies, tools, and applications* (σσ. 1459-1467). IGI Global.
- Sheshadri, A., Gitinabard, N., Lynch, C. F., Barnes, T., & Heckman, S. (2019). Predicting student performance based on online study habits: a study of blended courses. *arXiv preprint arXiv:1904.07331*.
- Shravan, I. (2017, Μάρτιος 25). Top 10 Open Source Data Mining Tools. *OpenSourceForU*. Ανάκτηση Οκτώβριος 25, 2022, από <https://www.opensourceforu.com/2017/03/top-10-open-source-data-mining-tools/>
- Shute, V. J., & Kim, Y. J. (2014). Formative and stealth assessment. Στο M. Spector, M. Merrill, J. Elen, & M. Bishop (Επιμ.), *Handbook of Research on Educational Communications and Technology* (σσ. 311-321). New York, NY: Springer.
- Siarova, H., Sternadel, D., & Mavsidlauskaitė, R. (2017). *Assessment practices for 21st century learning: Review of evidence, NESET II report*. Publications Office of the European Union.
- Siemens, G., & Gasevic, D. (2012). Guest Editorial- Learning and knowledge analytics. *Journal of Educational Technology & Society*, 15(3), 1-2.
- Sim, G., Holifield, P., & Brown, M. (2004). Implementation of computer assisted assessment: lessons from the literature. *ALT-J, Research in Learning Technology*, 12(3), σσ. 215-229.
- Simkin, M. G., & Kuechler, W. L. (2005). Multiple-choice tests and student understanding: What is the connection? *Decision Sciences Journal of Innovative Education*, 3(1), σσ. 73-98.
- Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. (T. & Francis}, Επιμ.) *Educational Psychologist*, 50(1), 1-13.
- Slater, S., Joksimovic, S., Kovanovic, V., Baker, R. S., & Gasevic, D. (2017). Tools for educational data mining: A review. *Journal of Educational and Behavioral Statistics*, 42, 85-106.
- Smith, G. (2007). How does student performance on formative assessments relate to learning assessed by exams? *Journal of College Science Teaching*, 36(7), σ. 28.
- Soffer, T., & Cohen, A. (2019). Students' engagement characteristics predict success and completion of online courses. *Journal of computer assisted learning*, 35(3), 378-389.
- Sorensen, E. K., & Takle, E. S. (2005). Investigating knowledge building dialogues in networked communities of practice. A collaborative learning endeavor across cultures. *Interactive educational multimedia: IEM*, 10, σσ. 50-60.

- Stewart, B. L., Waight, C. L., Norwood, M. M., & Ezell, S. D. (2004). Formative and summative evaluation of online courses. *Quarterly Review of Distance Education*, 5(2).
- Stödberg, U. (2012). A research review of e-assessment. *Assessment & Evaluation in Higher Education*, 37(5), σσ. 591-604.
- Stuber-McEwen, D., Wiseley, P., & Hoggatt, S. (2009). Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration*, 12(3).
- Sweat-Guy, R., & Buzzetto-More, N. A. (2007). A comparative analysis of common E-Portfolio features and available platforms. *Issues in Informing Science & Information Technology*, 4.
- Sweeney, T., West, D., Groessler, A., Haynie, A., Higgs, B. M., Macaulay, J., . . . Yeo, M. (2017). Where's the transformation? Unlocking the potential of technology-enhanced assessment. *Teaching and Learning Inquiry*, 5(1), σσ. 1-16.
- Talbert, R. (2020, Ιούλιος 20). *Mastery grading and academic honesty*. Ανάκτηση Αύγουστος 27, 2020, από <https://rtalbert.org/mastery-grading-and-academic-honesty/>
- Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior*, 47, 157-167.
- Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016). Rethinking assessment in a digital age: opportunities, challenges and risks. *British Educational Research Journal*, 42(3), σσ. 454-476.
- Tomas, C., Borg, M., & McNeil, J. (2015). E-assessment: Institutional development strategies and the assessment life cycle. *British Journal of Educational Technology*, 46(3).
- Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers & education*, 143, 103676.
- Torres, C., Lopes, A. P., Babo, L., & Azevedo, J. (2011). Improving Multiple-Choice Questions. *Online Submission*.
- Umer, R., Mathrani, A., Susnjak, T., & Lim, S. (2019). Mining activity log data to predict student's outcome in a course. *Proceedings of the 2019 international conference on big data and education*, (σσ. 52-58).
- Valenti, S., Neri, F., & Cucchiarelli, A. (2003). An overview of current research on automated essay grading. *Journal of Information Technology Education: Research*, 2(1), σσ. 319-330.

- Van der Pol, J., Van den Berg, B., Admiraal, W. F., & Simons, P. R.-J. (2008). The nature, reception, and use of online peer feedback in higher education. *Computers & Education*, 51(4), σσ. 1804-1817.
- Vaughan, N. D., Martha, C.-I., & Garrison, D. R. (2013). *Teaching in blended learning environments: Creating and sustaining communities of inquiry*. Athabasca University Press.
- Ventouras, E., Triantis, D., Tsiakas, P., & Stergiopoulos, C. (2010). Comparison of examination methods based on multiple-choice questions and constructed-response questions using personal computers. *Computers & Education*, 54(2), σσ. 455-461.
- Viberg, O., Hatakka, M., Balter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in human behavior*, 98-110.
- Villagra-Arnedo, C.-J., Gallego-Duran, F. J., Compan, P., Llorens Largo, F., & Molina-Carmona, R. (2016). Predicting academic performance from behavioural and learning data. *Int. J. Des. Nat. Ecodyn.*, 239-249.
- Vista, A., & Care, E. (2017). *It's time to mobilize around a new approach to educational assessment*. Stanford social innovation review.
- Vlachopoulos, D., & Makri, A. (2017). The effect of games and simulations on higher education: a systematic literature review. *International Journal of Educational Technology in Higher Education*, 14(1), σ. 22.
- Vonderwell, S., Liang, X., & Alderman, K. (2007). Asynchronous discussions and assessment in online learning. *Journal of Research on Technology in Education*, 39(3), σσ. 309-328.
- Vytasek, J. M., Patzak, A., & Winne, P. H. (2020). Analytics for student engagement. Στο M. V. al. (Επιμ.), *Machine learning paradigms* (σσ. 23-48). Springer.
- Waikato Software. (2022, Οκτώβριος 05). Ανάκτηση από The University of Waikato: <https://ai.waikato.ac.nz/software>
- Waikato, U. ο. (2013, Ιανουάριος 16). *Data Mining with Weka*. Ανάκτηση 11 26, 2022, από WekaMOOC: <https://www.youtube.com/@WekaMOOC/featured>
- Wang, L., Sy, A., Liu, L., & Piech, C. (2017). Deep knowledge tracing on programming exercises. *Proceedings of the fourth (2017) ACM conference on learning@ scale* (σσ. 201-204). ACM.
- Watson, G., & Sottile, J. (2010). Cheating in the Digital Age: Do students cheat more in on-line courses? *13*(1).



- Watters, A. (2015, Ιανουάριος 27). *Multiple Choice and Testing Machines: A History*. Ανάκτηση Απρίλιος 21, 2020, από <http://hackededucation.com/2015/01/27/multiple-choice-testing-machines>
- Weka 3: Machine Learning Software in Java*. (2022, Οκτώβριος 25). Ανάκτηση από University of Waikato, Departments of Computer Science and Software Engineering: <https://www.cs.waikato.ac.nz/~ml/weka/>
- WEKA Documentation*. (2022, Οκτώβριος 12). Ανάκτηση Οκτώβριος 12, 2022, από WEKA Wiki: <https://waikato.github.io/weka-wiki/documentation/>
- WEKA Literature*. (2022, Οκτώβριος 12). Ανάκτηση Οκτώβριος 12, 2022, από WEKA Wiki: <https://waikato.github.io/weka-wiki/literature/>
- Weka Wiki*. (2022, Οκτώβριος 25). Ανάκτηση από <https://waikato.github.io/weka-wiki/>
- weka.classifiers.functions Class Logistic*. (χ.χ.). Ανάκτηση Ιανουάριος 16, 2023, από SourceForge: <https://weka.sourceforge.io/doc.dev/weka/classifiers/functions/Logistic.html>
- weka.classifiers.functions Class SimpleLogistic*. (χ.χ.). Ανάκτηση Ιανουάριος 16, 2023, από SourceForge: <https://weka.sourceforge.io/doc.dev/weka/classifiers/functions/SimpleLogistic.html>
- Weleschuk, A., Dyjur, P., & Kelly, P. (2019). *Online Assessment in Higher Education*. Taylor Institute for Teaching and Learning, University of Calgary.
- Whitelock, D. (2007). Computer assisted formative assessment: Supporting students to become more reflective learners. Στο C. P. Constantinou, & Z. C. Zacharia (Επιμ.), *Proceedings of the 8th International Conference on Computer Based Learning in Science (CBLIS '07)* (σσ. 492-503). Crete, Greece: E-Media, University of Crete.
- Whitelock, D., & Brasher, A. (2006). *Roadmap for e-assessment*. Ανάκτηση Απρίλιος 25, 2020, από <https://www.webarchive.org.uk/wayback/archive/20140613220103/http://www.jisc.ac.uk/media/documents/programmes/elearningpedagogy/roadmapeassessment.doc>
- Whitelock, D., & Watt, S. (2008). Reframing e-assessment: adopting new media and adapting old frameworks. *Learning and Media Technology*, 33(3), σσ. 151-154.
- Whitley, B. E. (1998). Factors associated with cheating among college students: A review. *Research in higher education*, 39(3), σσ. 235-274.
- Wiedmer, T. L. (1998). Digital portfolios: capturing and demonstrating skills and levels of performance.(electronic validation of individual teacher performance). *Phi Delta Kappan*, 4.

- William, D. (2018). Assessment for learning: meeting the challenge of implementation. *Assessment in Education: Principles, Policy & Practice*, 25(6), σσ. 682-685.
- Williams, J. B. (2006). Assertion-reason multiple-choice testing as a tool for deep learning: a qualitative analysis. *Assessment & Evaluation in Higher Education*, 31(3), σσ. 287-301.
- Witten, I. H. (2013, Σεπτέμβριος 9). *Transcript for Class1 - Lesson 3*. Ανάκτηση Νοέμβριος 26, 2022, από Data Mining with Weka MOOC - Material: <https://www.cs.waikato.ac.nz/ml/weka/mooc/dataminingwithweka/transcripts/Transcript1-3.txt>
- Witten, I. H. (2013, Σεπτέμβριος 16). *Transcript for Class2 - Lesson 5*. Ανάκτηση Νοέμβριος 26, 2022, από Data Mining with Weka MOOC - Material: <https://www.cs.waikato.ac.nz/ml/weka/mooc/dataminingwithweka/transcripts/Transcript2-5.txt>
- Witten, I. H. (2013, Σεπτέμβριος 16). *Transcript for Class2 - Lesson 6*. Ανάκτηση Νοέμβριος 30, 2022, από Data Mining with Weka MOOC - Material: <https://www.cs.waikato.ac.nz/ml/weka/mooc/dataminingwithweka/transcripts/Transcript2-6.txt>
- Witten, I. H., Frank, E., & Hall, M. (2011). *Data Mining, Practical Machine Learning Tools and Techniques (3rd Ed.)* (3rd εκδ.). Elsevier.
- Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th international conference on evaluation and assessment in software engineering* (σσ. 1-10). New York: ACM.
- Wolsey, T. (2008). Efficacy of instructor feedback on written work in an online program. *International Journal on E-learning*, 7(2), σσ. 311-329.
- Wong, A., & Chong, S. (2018). Modelling adult learners' online engagement behaviour: proxy measures and its application. *Journal of Computers in Education*, 5(4), 463-479.
- Yin, R. K. (1984). *Case Study Research: Design and Methods* (5th εκδ.). Beverly Hills: Sage Publications.
- You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education*, 29, 23-30.
- Yu, T., & Jo, I.-H. (2014). Educational technology approach toward learning analytics: Relationship between student online behavior and learning performance in higher education. *Proceedings of the fourth international conference on learning analytics and knowledge*, (σσ. 269-270).

- Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education*, 27, 44-53.
- Zulkifli, F., Mohamed, Z., & Azmee, N. A. (2019). Systematic research on predictive models on students' academic performance in higher education. *International Journal of Recent Technology and Engineering*, 8(2), 357-363.
- Zvornicanin, E. (2022, Νοέμβριος 4). *What is Feature Importance in Machine Learning?* Ανάκτηση Ιανουάριος 18, 2023, από Baeldung: <https://www.baeldung.com/cs/ml-feature-importance>
- Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, 38(9), σσ. 25-32.
- Δημητρακοπούλου, Α. (2017). Ανάλυση ψηφιακών ιχνών και δεδομένων σε Εκπαιδευτικά Περιβάλλοντα: Προσεγγίσεις, Προοπτικές και Ζητήματα Ηθικής στο Πεδίο της Μαθησιακής Αναλυτικής. Στο Α. Κοντάκος, & Φ. Καλαβάσης, *Θέματα Εκπαιδευτικού Σχεδιασμού, 8ος Τόμος*. Διάδραση.
- Ζήμερας, Σ. (2020). *Ανάλυση Κατηγορικών Δεδομένων*. Ανάκτηση Ιανουάριος 16, 2023, από [eclass.aegean.gr](http://eclass.aegean.gr):  
[https://eclass.aegean.gr/modules/document/file.php/MATH110/ΣΥΜΠΛΗΡΩΜΑΤΙΚΟ\\_ΥΛΙΚΟ/ΛΟΓΙΣΤΙΚΗ\\_ΠΑΛΙΝΔΡΟΜΗΣΗ\\_-\\_ΠΑΡΑΔΕΙΓΜΑΤΑ.pdf](https://eclass.aegean.gr/modules/document/file.php/MATH110/ΣΥΜΠΛΗΡΩΜΑΤΙΚΟ_ΥΛΙΚΟ/ΛΟΓΙΣΤΙΚΗ_ΠΑΛΙΝΔΡΟΜΗΣΗ_-_ΠΑΡΑΔΕΙΓΜΑΤΑ.pdf)
- Μαθήματα ΠΜΣ Νέες Μορφές Εκπαίδευσης και Μάθησης*. (2022, Οκτώβριος 11). Ανάκτηση από Ιστοσελίδα ΠΜΣ Νέες Μορφές Εκπαίδευσης και Μάθησης:  
<https://nfel.aegean.gr/%CE%BC%CE%B1%CE%B8%CE%AE%CE%BC%CE%B1%CF%84%CE%B1/>
- Μαυρομάτη, Μ. (2019). Η έρευνα βασισμένη σε σχεδιασμό ως εργαλείο εκπαιδευτικής χειραφέτησης. *Επιστήμες Αγωγής*, 2, 75-88.
- Οδηγός Σπουδών ΤΕΠΑΕΣ*. (2022, 10 10). Ανάκτηση από Ιστοσελίδα ΤΕΠΑΕΣ:  
[http://www.pse.aegean.gr/?page\\_id=5421](http://www.pse.aegean.gr/?page_id=5421)
- Πετροπούλου, Ο., Κασιμάτη, Κ., & Ρετάλης, Σ. (2005). *Σύγχρονες μορφές εκπαιδευτικής αξιολόγησης με αξιοποίηση εκπαιδευτικών τεχνολογιών*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. Ανάκτηση Απρίλιος 23, 2020, από <http://hdl.handle.net/11419/232>
- Ρούσσο, Π. Λ., & Τσαούσης, Γ. (2002). *Στατιστική εφαρμοσμένη στις κοινωνικές επιστήμες*. Αθήνα: Ελληνικά Γράμματα.
- Φεσάκης, Γ. (2019). *Εισαγωγή στις εφαρμογές των ψηφιακών τεχνολογιών στην εκπαίδευση*. Αθήνα: Gutenberg.

Φεσάκης, Γ., & Κωνσταντοπούλου, Α. (2022). Σχεδιασμός τεχνολογικά ενισχυμένων εκπαιδευτικών σεναρίων για την προσχολική εκπαίδευση. Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις.

---

*ΠΑΡΑΡΤΗΜΑΤΑ*

---

## ΠΑΡΑΡΤΗΜΑ Α

**Πίνακας 56. Συναρτήσεις στο φύλλο Excel για τον υπολογισμό της τιμής των χαρακτηριστικών του συνόλου δεδομένων του μαθήματος AA005 Χρήση Βασικών Εφαρμογών ΗΥ.**

(Τα χαρακτηριστικά του συνόλου δεδομένων επισημαίνονται με κίτρινο χρώμα, ενώ οι βοηθητικές μεταβλητές με πράσινο χρώμα).

Στήλη	Χαρακτηριστικό	Συνάρτηση υπολογισμού τιμής χαρακτηριστικού
A	TimePeriodStarts (Date)	Έναρξη Περιόδου (σταθερή τιμή)
B	TimePeriodEndes (Date)	Λήξη Περιόδου (σταθερή τιμή)
C	TimePeriodStarts (WeekNumber)	1
D	TimePeriodEnds (WeekNumber)	Αριθμός τελευταίας εβδομάδας περιόδου
E	User full name	Κενό (για λόγους ανωνυμοποίησης)
F	ID	VLOOKUP(E2,Grades!\$A\$2:\$H\$188,2,FALSE) , ο ψευδωνυμοποιημένος αριθμός του φοιτητή.
G	Status	VLOOKUP(E2,Grades!\$A\$2:\$H\$188,6,FALSE), η κατάσταση εγγραφής του φοιτητή
H	Grade	VLOOKUP(E2,Grades!\$A\$2:\$H\$188,7,FALSE), ο βαθμός του φοιτητή
I	Class	VLOOKUP(E2,Grades!\$A\$2:\$H\$188,8,FALSE), η κατηγορία του φοιτητή
J	WeekDay	WEEKDAY(M2)
K	DAY	TEXT(C2,"dddd")
L	WeekNumber	Αριθμός της εβδομάδας που συμβαίνει η εγγραφή
M	Time	Ο χρόνος (μέρα / ώρα) της εγγραφής
N	AnotherDay	IF(DAYS(M3,M2),1,"")
O	AMorPM	IFS(HOUR(M2)>=21,"LatePM",HOUR(M2)>=12,"EarlyPM",HOUR(M2)>=7,"LateAM",HOUR(M2)>=0,"EarlyAM")
P	IdleMinutes	(M2-M3)*1440
Q	UserChange	IF(E2<>E3,"UserChange","NoUserChange")
R	NewSession	IF(Q2="UserChange","NewSession",IF(P2>=40,"NewSession"))
S	DistanceOfRowToNextNewSession	IF(R2="NewSession",0,S1+1)
T	RowsPerSession	IF(R2="NewSession",S1+1,"")

U	SessionDuration	IF(T2=1,0,IF(R2="NewSession",SUM(OFFSET(P2,-T2+1,0,T2-1,1)), ""))
V	Total Sessions	COUNTIFS(E:E,E2,R:R,"NewSession",L:L,"<="&D2,L:L,">="&C2)
W	TotalOnlineTimeForActivityPeriod (minutes)	SUMIFS(U:U,E:E,E2,L:L,"<="&D2,L:L,">="&C2)
X	Average Time Per Session per Student	W2/V2
Y	SDEV of Time per Session per student	IF(Q2="UserChange",IF(V2=1,0,STDEV.P(IF(E:E=E2,IF(L:L<=D2,IF(L:L>=C2,U:U))))), "")
Z	IdleDaysBetweenSessions	IF(AND(R2="NewSession",V2=1),0,IF(AND(E2=E3,R2="NewSession"),IF(DAYS(M2,M3)=0,0,DAYS(M2,M3)-1), ""))
AA	TotalIdleTimeForActivityPeriod (minutes)	IF(V2=1,0,SUMIFS(P:P,E:E,E2,R:R,"NewSession",Q:Q,"NoUserChange",L:L,"<="&D2,L:L,">="&C2))
AB	AverageIdleDaysBetweenSessions	AVERAGEIFS(Z:Z,E:E,E2,L:L,"<="&D2,L:L,">="&C2)
AC	STDEV IdleDaysBetweenSessions	IF(Q2="UserChange",STDEV.P(IF(E:E=E2,IF(L:L<=D2,IF(L:L>=C2,Z:Z))))), "")
AD	Max Idle Days	IF(Q2="UserChange",MAXIFS(Z:Z,E:E,E2,L:L,"<="&D2,L:L,">="&C2), "")
AE	FirstSession	MINIFS(M:M,E:E,E2,L:L,">="&C2)
AF	LastSession	MAXIFS(M:M,E:E,E2,L:L,"<="&D2)
AG	DaysBetweenFirstAndLast Sessions	DAYS(AF2,AE2)
AH	Total Activity Days	SUMIFS(N:N,E:E,E2,L:L,"<="&D2,L:L,">="&C2)
AI	Acess Density (TotalDaysToActivityDays)	AG2/AH2
AJ	TotalIdleDaysDuringActivityPeriod	SUMIFS(Z:Z,E:E,E2,L:L,"<="&D2,L:L,">="&C2)
AK	DaysFromStartUntilFirstSession	DAYS(AE2,A2)
AL	DaysFromLastSessionUntil End	DAYS(B2,AF2)
AM	SundaySession	IF(R2=FALSE,FALSE,IF(J2=1,"SundaySession","WeekdaySession"))
AN	Early/LateSession	IF(R2=FALSE,FALSE,CONCAT(O2,R2))
AO	SundaySessions	COUNTIFS(E:E,E2,AM:AM,"SundaySession",L:L,"<="&D2,L:L,">="&C2)
AP	WeekDaySessions	COUNTIFS(E:E,E2,AM:AM,"WeekdaySession",L:L,"<="&D2,L:L,">="&C2)
AQ	EarlyAM Sessions	COUNTIFS(E:E,E2,AN:AN,"EarlyAMNewSession",L:L,"<="&D2,L:L,">="&C2)
AR	LateAM Sessions	COUNTIFS(E:E,E2,AN:AN,"LateAMNewSession",L:L,"<="&D2,L:L,">="&C2)
AS	EarlyPM Sessions	COUNTIFS(E:E,E2,AN:AN,"EarlyPMNewSession",L:L,"<="&D2,L:L,">="&C2)
AT	LatePM Sessions	COUNTIFS(E:E,E2,AN:AN,"LatePMNewSession",L:L,"<="&D2,L:L,">="&C2)

AU	Total Events	COUNTIFS(E:E,E2,L:L,"<="&D2,L:L,">="&C2)
AV	File Events	COUNTIFS(E:E,E2,L:L,"<="&D2,L:L,">="&C2,AZ:AZ,"File")
AW	Folder Events	COUNTIFS(E:E,E2,L:L,"<="&D2,L:L,">="&C2,AZ:AZ,"Folder")
AX	PageEvents	COUNTIFS(E:E,E2,L:L,"<="&D2,L:L,">="&C2,AZ:AZ,"Page")

**Πίνακας 57. Συναρτήσεις στο φύλλο Excel για τον υπολογισμό της τιμής των χαρακτηριστικών του συνόλου δεδομένων του μαθήματος Α3 Ηλεκτρονική Μάθηση.**

(Τα χαρακτηριστικά του συνόλου δεδομένων επισημαίνονται με κίτρινο χρώμα, ενώ οι βοηθητικές μεταβλητές με πράσινο χρώμα).

Στήλη	Όνομα	Συνάρτηση υπολογισμού τιμής χαρακτηριστικού
A	TimePeriodStarts (Date)	Έναρξη Περιόδου (σταθερή τιμή)
B	TimePeriodEndes (Date)	Λήξη Περιόδου (σταθερή τιμή)
C	TimePeriodStarts (WeekNumber)	1
D	TimePeriodEnds (WeekNumber)	Αριθμός τελευταίας εβδομάδας περιόδου
E	User full name	Κενό (για λόγους ανωνυμοποίησης)
F	ID	VLOOKUP(E2,Grades!\$A\$2:\$H\$188,2,FALSE) , ο ψευδωνυμοποιημένος αριθμός του φοιτητή
G	Status	VLOOKUP(E2,Grades!\$A\$2:\$H\$188,6,FALSE), η κατάσταση εγγραφής του φοιτητή
H	Grade	VLOOKUP(E2,Grades!\$A\$2:\$H\$188,7,FALSE), ο βαθμός του φοιτητή
I	Class	VLOOKUP(E2,Grades!\$A\$2:\$H\$188,8,FALSE), η κατηγορία του φοιτητή
J	WeekDay	WEEKDAY(M2)
K	DAY	TEXT(C2,"ddd")
L	WeekNumber	Αριθμός της εβδομάδας που συμβαίνει η εγγραφή.
M	Time	Ο χρόνος (μέρα / ώρα) της εγγραφής
N	AnotherDay	IF(DAYS(M3,M2),1,"")
O	AMorPM	IFS(HOUR(M2)>=21,"LatePM",HOUR(M2)>=12,"EarlyPM",HOUR(M2)>=7,"LateAM",HOUR(M2)>=0,"EarlyAM")



P	IdleMinutes	$(M2-M3)*1440$
Q	UserChange	$IF(E2<>E3,"UserChange","NoUserChange")$
R	NewSession	$IF(Q2="UserChange","NewSession",IF(P2>=40,"NewSession"))$
S	DistanceOfRowToNextNewSession	$IF(R2="NewSession",0,S1+1)$
T	RowsPerSession	$IF(R2="NewSession",S1+1,"")$
U	SessionDuration	$IF(T2=1,0,IF(R2="NewSession",SUM(OFFSET(P2,-T2+1,0,T2-1,1)),""))$
V	Total Sessions	$COUNTIFS(E:E,E2,R:R,"NewSession",L:L,"<="&D2,L:L,">="&C2)$
W	TotalOnlineTimeForActivityPeriod (minutes)	$SUMIFS(U:U,E:E,E2,L:L,"<="&D2,L:L,">="&C2)$
X	Average Time Per Session per Student	$W2/V2$
Y	SDEV of Time per Session per student	$IF(Q2="UserChange",IF(V2=1,0,STDEV.P(IF(E:E=E2,IF(L:L<=D2,IF(L:L>=C2,U:U)))),""))$
Z	IdleDaysBetweenSessions	$IF(AND(R2="NewSession",V2=1),0,IF(AND(E2=E3,R2="NewSession"),IF(DAYS(M2,M3)=0,0,DAYS(M2,M3)-1),""))$
AA	TotalIdleTimeForActivityPeriod (minutes)	$IF(V2=1,0,SUMIFS(P:P,E:E,E2,R:R,"NewSession",Q:Q,"NoUserChange",L:L,"<="&D2,L:L,">="&C2))$
AB	AverageIdleDaysBetweenSessions	$AVERAGEIFS(Z:Z,E:E,E2,L:L,"<="&D2,L:L,">="&C2)$
AC	STDEV IdleDaysBetweenSessions	$IF(Q2="UserChange",STDEV.P(IF(E:E=E2,IF(L:L<=D2,IF(L:L>=C2,Z:Z)))),"")$
AD	Max Idle Days	$IF(Q2="UserChange",MAXIFS(Z:Z,E:E,E2,L:L,"<="&D2,L:L,">="&C2),"")$
AE	FirstSession	$MINIFS(M:M,E:E,E2,L:L,">="&C2)$
AF	LastSession	$MAXIFS(M:M,E:E,E2,L:L,"<="&D2)$
AG	DaysBetweenFirstAndLastSessions	$DAYS(AF2,AE2)$
AH	Total Activity Days	$SUMIFS(N:N,E:E,E2,L:L,"<="&D2,L:L,">="&C2)$
AI	Acess Density (TotalDaysToActivityDays)	$AG2/AH2$
AJ	TotalIdleDaysDuringActivityPeriod	$SUMIFS(Z:Z,E:E,E2,L:L,"<="&D2,L:L,">="&C2)$

AK	DaysFromStartUntilFirstSession	DAYS(AE2,A2)
AL	DaysFromLastSessionUntilEnd	DAYS(B2,AF2)
AM	SundaySession	IF(R2=FALSE,FALSE,IF(J2=1,"SundaySession","WeekdaySession"))
AN	Early/LateSession	IF(R2=FALSE,FALSE,CONCAT(O2,R2))
AO	SundaySessions	COUNTIFS(E:E,E2,AM:AM,"SundaySession",L:L,"<="&D2,L:L,">="&C2)
AP	WeekDaySessions	COUNTIFS(E:E,E2,AM:AM,"WeekdaySession",L:L,"<="&D2,L:L,">="&C2)
AQ	EarlyAM Sessions	COUNTIFS(E:E,E2,AN:AN,"EarlyAMNewSession",L:L,"<="&D2,L:L,">="&C2)
AR	LateAM Sessions	COUNTIFS(E:E,E2,AN:AN,"LateAMNewSession",L:L,"<="&D2,L:L,">="&C2)
AS	EarlyPM Sessions	COUNTIFS(E:E,E2,AN:AN,"EarlyPMNewSession",L:L,"<="&D2,L:L,">="&C2)
AT	LatePM Sessions	COUNTIFS(E:E,E2,AN:AN,"LatePMNewSession",L:L,"<="&D2,L:L,">="&C2)
AU	Total Events	COUNTIFS(E:E,E2,L:L,"<="&D2,L:L,">="&C2)
AV	File Events	COUNTIFS(E:E,E2,L:L,"<="&D2,L:L,">="&C2,BA:BA,"File")
AW	Forum Events	COUNTIFS(E:E,E2,L:L,"<="&D2,L:L,">="&C2,BA:BA,"Forum")
AX	PostedContentEvents	COUNTIFS(E:E,E2,L:L,"<="&D2,L:L,">="&C2,BB:BB,"Some content has been posted.")
AY	ForumGrade	VLOOKUP(E2,Grades!\$A\$2:\$H\$188,4,FALSE)

## ΠΑΡΑΡΤΗΜΑ Β

### Οι τιμές των παραμέτρων του LSTM αλγόριθμου.

```
functions.RnnSequenceClassifier '-S 1 -tBPTTBackward 3 -tBPTTForward 3 -cache-mode
MEMORY -early-stopping \"weka.dl4j.earlystopping.EarlyStopping -maxEpochsNoImprovement
0 -valPercentage 0.0\" -normalization \"No normalization/standardization\" -iterator
\"weka.dl4j.iterators.instance.sequence.RelationalInstanceIterator -relationalAttributeIndex 0 -
truncationLength 100 -bs 1\" -iteration-listener \"weka.dl4j.listener.EpochListener -eval true -n
5\" -layer \"weka.dl4j.layers.LSTM -gateActivation \\\"weka.dl4j.activations.ActivationSigmoid
\\\" -nOut 4 -activation \\\"weka.dl4j.activations.ActivationReLU \\\" -name \\\"LSTM
layer\\\" -layer \"weka.dl4j.layers.RnnOutputLayer -lossFn
\\\"weka.dl4j.lossfunctions.LossMCXENT \\\" -nOut 2 -activation
\\\"weka.dl4j.activations.ActivationSoftmax \\\" -name \\\"RnnOutput layer\\\" -logConfig
\"weka.core.LogConfiguration -append true -dl4jLogLevel WARN -logFile
C:\\\\Users\\\\\\\\tsigaros\\\\\\\\wekafiles\\\\\\\\wekaDeeplearning4j.log -nd4jLogLevel INFO -
wekaDL4jLogLevel INFO\" -config \"weka.dl4j.NeuralNetConfiguration -biasInit 0.0 -biasUpdater
\\\"weka.dl4j.updater.Sgd -lr 0.001 -lrSchedule \\\"weka.dl4j.schedules.ConstantSchedule -
scheduleType EPOCH\\\" -dist \\\"weka.dl4j.distribution.Disabled \\\" -dropout
\\\"weka.dl4j.dropout.Disabled \\\" -gradientNormalization None -gradNormThreshold 1.0 -l1
NaN -l2 NaN -minimize -algorithm STOCHASTIC_GRADIENT_DESCENT -updater
\\\"weka.dl4j.updater.Adam -beta1MeanDecay 0.9 -beta2VarDecay 0.999 -epsilon 1.0E-8 -lr
0.001 -lrSchedule \\\"weka.dl4j.schedules.ConstantSchedule -scheduleType
EPOCH\\\" -weightInit XAVIER -weightNoise \\\"weka.dl4j.weightnoise.Disabled \\\" -
numEpochs 20 -numGPUs 1 -averagingFrequency 10 -prefetchSize 24 -queueSize 0 -zooModel
\"weka.dl4j.zoo.CustomNet -channelsLast false -pretrained NONE\" 5643486590174837865
```

Η παραπάνω παραμετροποίηση διαφέρει από την προκαθορισμένη (default) ως προς τα εξής:

- LayerSpecification:
  - Upper Layer: LSTMLayer, Number of outputs: 4
  - Bottom Layer: RnnOutputLayer, Number of outputs: 2
- Number of epochs: 20
- Instance Iterator: RelationalInstanceIterator,

- SizeofMinibatch:1,
  - TruncationLength: 100
- Truncated backdrop through time backward: 3
- Truncated backdrop through time forward: 3

## ΠΑΡΑΡΤΗΜΑ Γ

**Αναφορά του συστήματος WEKA μετά από εφαρμογή του μοντέλου που δημιουργήθηκε από το σύνολο δεδομένων του μαθήματος ΑΑ005 2021-2022 κατά την 17<sup>η</sup> εβδομάδα, επί του συνόλου του μαθήματος ΑΑ0052020-2021 κατά την 9<sup>η</sup> εβδομάδα.**

User supplied test set

Relation: TestData2020-2021Week09

Instances: unknown (yet). Reading incrementally

Attributes: 25

=== Predictions on user test set ===

inst #	actual	predicted	error	prediction		inst #	actual	predicted	error	prediction
1	1:Pass	1:Pass		0.838		41	1:Pass	1:Pass		0.979
2	1:Pass	1:Pass		0.994		42	1:Pass	1:Pass		0.897
3	1:Pass	2:Fail	+	0.63		43	1:Pass	1:Pass		0.999
4	2:Fail	1:Pass	+	0.573		44	1:Pass	1:Pass		0.968
5	1:Pass	1:Pass		0.964		45	2:Fail	2:Fail		0.784
6	1:Pass	1:Pass		0.934		46	2:Fail	2:Fail		0.653
7	1:Pass	1:Pass		0.998		47	2:Fail	1:Pass	+	0.916
8	1:Pass	1:Pass		0.926		48	1:Pass	1:Pass		0.998
9	1:Pass	1:Pass		0.986		49	1:Pass	1:Pass		0.887
10	1:Pass	1:Pass		0.732		50	1:Pass	1:Pass		0.936
11	1:Pass	1:Pass		0.991		51	1:Pass	1:Pass		0.892
12	2:Fail	2:Fail		0.513		52	2:Fail	1:Pass	+	0.889
13	2:Fail	1:Pass	+	0.629		53	1:Pass	1:Pass		0.888

14	1:Pass	1:Pass		0.942		54	1:Pass	1:Pass		0.833
15	1:Pass	1:Pass		0.998		55	1:Pass	1:Pass		0.595
16	1:Pass	1:Pass		0.998		56	1:Pass	1:Pass		0.713
17	2:Fail	1:Pass	+	0.968		57	1:Pass	1:Pass		0.997
18	2:Fail	1:Pass	+	0.543		58	1:Pass	1:Pass		0.962
19	1:Pass	2:Fail	+	0.961		59	1:Pass	1:Pass		0.68
20	1:Pass	1:Pass		0.994		60	1:Pass	1:Pass		0.706
21	1:Pass	1:Pass		0.965		61	1:Pass	1:Pass		0.983
22	1:Pass	1:Pass		0.998		62	1:Pass	1:Pass		0.876
23	1:Pass	1:Pass		0.964		63	1:Pass	1:Pass		0.786
24	1:Pass	1:Pass		0.996		64	1:Pass	1:Pass		0.983
25	1:Pass	1:Pass		0.945		65	2:Fail	2:Fail		0.987
26	2:Fail	2:Fail		1		66	2:Fail	1:Pass	+	0.833
27	1:Pass	1:Pass		0.984		67	1:Pass	1:Pass		0.926
28	1:Pass	2:Fail	+	0.891		68	2:Fail	1:Pass	+	0.718
29	1:Pass	1:Pass		0.996		69	1:Pass	1:Pass		0.999
30	1:Pass	1:Pass		0.996		70	2:Fail	1:Pass	+	0.538
31	2:Fail	1:Pass	+	0.893		71	1:Pass	1:Pass		0.84
32	2:Fail	1:Pass	+	0.529		72	1:Pass	1:Pass		0.812
33	1:Pass	1:Pass		0.941		73	1:Pass	1:Pass		0.994
34	2:Fail	1:Pass	+	0.589		74	1:Pass	1:Pass		0.93
35	1:Pass	1:Pass		0.856		75	1:Pass	1:Pass		0.982
36	2:Fail	1:Pass	+	0.505		76	2:Fail	1:Pass	+	0.727

37	2:Fail	2:Fail		1		77	1:Pass	1:Pass		0.908
38	2:Fail	1:Pass	+	0.568		78	2:Fail	1:Pass	+	0.755
39	2:Fail	2:Fail		0.851		79	2:Fail	1:Pass	+	0.906
40	2:Fail	1:Pass	+	0.943		80	1:Pass	1:Pass		0.914
81	2:Fail	1:Pass	+	0.948		121	1:Pass	1:Pass		0.747
82	1:Pass	1:Pass		0.785		122	1:Pass	1:Pass		0.999
83	1:Pass	1:Pass		0.937		123	1:Pass	1:Pass		0.978
84	2:Fail	1:Pass	+	0.847		124	1:Pass	1:Pass		0.77
85	1:Pass	1:Pass		0.983		125	1:Pass	1:Pass		0.971
86	1:Pass	1:Pass		0.992		126	2:Fail	2:Fail		0.997
87	1:Pass	1:Pass		0.899		127	1:Pass	1:Pass		0.934
88	1:Pass	1:Pass		0.747		128	1:Pass	1:Pass		0.999
89	1:Pass	1:Pass		0.982		129	2:Fail	2:Fail		0.984
90	1:Pass	1:Pass		0.97		130	2:Fail	2:Fail		0.606
91	1:Pass	1:Pass		0.674		131	1:Pass	1:Pass		0.999
92	1:Pass	1:Pass		0.983		132	2:Fail	2:Fail		1
93	1:Pass	1:Pass		0.973		133	1:Pass	1:Pass		0.976
94	2:Fail	2:Fail		1		134	2:Fail	1:Pass	+	0.772
95	1:Pass	1:Pass		0.968		135	1:Pass	2:Fail	+	0.613
96	2:Fail	2:Fail		0.697		136	2:Fail	2:Fail		0.999
97	2:Fail	2:Fail		0.719		137	1:Pass	1:Pass		0.957
98	1:Pass	1:Pass		0.998		138	1:Pass	1:Pass		0.753

99	2:Fail	1:Pass	+	0.954		139	1:Pass	1:Pass		0.985
100	1:Pass	1:Pass		0.849		140	1:Pass	1:Pass		0.967
101	1:Pass	1:Pass		0.944		141	1:Pass	1:Pass		0.881
102	1:Pass	1:Pass		0.978		142	1:Pass	1:Pass		0.956
103	1:Pass	1:Pass		0.996		143	2:Fail	1:Pass	+	0.924
104	1:Pass	1:Pass		0.657		144	1:Pass	1:Pass		0.932
105	1:Pass	1:Pass		0.954		145	1:Pass	1:Pass		0.995
106	1:Pass	1:Pass		0.938		146	1:Pass	1:Pass		0.716
107	1:Pass	1:Pass		0.999		147	1:Pass	1:Pass		0.993
108	1:Pass	2:Fail	+	0.558		148	1:Pass	1:Pass		0.979
109	1:Pass	1:Pass		0.991		149	2:Fail	1:Pass	+	0.521
110	1:Pass	2:Fail	+	0.77		150	2:Fail	1:Pass	+	0.972
111	1:Pass	1:Pass		0.992		151	2:Fail	1:Pass	+	0.769
112	2:Fail	1:Pass	+	0.991		152	2:Fail	1:Pass	+	0.897
113	2:Fail	1:Pass	+	0.929		153	1:Pass	1:Pass		0.79
114	1:Pass	1:Pass		0.956		154	1:Pass	1:Pass		0.942
115	1:Pass	1:Pass		0.731		155	2:Fail	2:Fail		0.528
116	1:Pass	1:Pass		0.996		156	1:Pass	1:Pass		0.938
117	1:Pass	1:Pass		0.63		157	1:Pass	1:Pass		0.912
118	1:Pass	1:Pass		0.983						
119	1:Pass	1:Pass		0.992						
120	2:Fail	2:Fail		0.579						

=== Summary ===



Correctly Classified Instances	122	77.71%
Incorrectly Classified Instances	35	22.29%
Kappa statistic	0.3696	
Mean absolute error	0.2467	
Root mean squared error	0.3961	
Total Number of Instances	157	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.946	0.63	0.784	0.946	0.857	0.406	0.831	0.922	Pass
	0.37	0.054	0.739	0.37	0.493	0.406	0.831	0.702	Fail
Weighted Avg.	0.777	0.462	0.771	0.777	0.75	0.406	0.831	0.857	

=== Confusion Matrix ===

```

a    b    <--  Classified as
105  6  |  a      a=Yes
29  17 |  b      b=No

```