



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ  
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ**

**ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ**

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ  
ΠΛΗΡΟΦΟΡΙΑΚΑ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΑΚΑ ΣΥΣΤΗΜΑΤΑ**

**Ανάλυση στιγμών αισθητικής σημασίας σε ακουσματική  
μουσική με χρήση μεθόδων μηχανικής μάθησης και  
επεξεργασία σήματος μουσικής**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

της

**Μαρίας Καλλιόπης Ντόμπρη**

**Επιβλέπων : Θεόδωρος Κωστούλας**

**Μέλη εξεταστικής επιτροπής: Ακριβή Βλάχου, Ευστάθιος Σταματάτος**

Σάμος, Φεβρουάριος 2022

Η σελίδα αυτή είναι σκόπιμα λευκή.

## Πρόλογος

Η αναζήτηση του τρόπου που αλληλοεπιδρά το ανθρώπινο όν με τον κόσμο, αλλά και επικοινωνεί με αυτόν αποτελεί μια από τις σπουδαιότερες επιστημονικές προκλήσεις στην σημερινή εποχή. Οι άνθρωποι άλλωστε, επηρεάζονται συναισθηματικά είτε θετικά, είτε αρνητικά πολλές φορές μέσα στη μέρα τους και η εναλλαγή αυτή των συναισθημάτων τους αποτελεί ενδιαφέρουσα προσέγγιση, η οποία τα τελευταία χρόνια αναλύεται μέσω των μεθόδων Μηχανικής Μάθησης. Πολλών δε ερευνητών η προσοχή φαίνεται να επικεντρώνεται στον τρόπο που αλληλοεπιδρούν με την μουσική και τα συναισθήματα που προκαλούν. Ωστόσο, για την πρόκληση συναισθημάτων υπάρχουν πολλά στοιχεία που συνθέτουν ένα μουσικό κομμάτι και τα εντείνουν και μερικά από αυτά είναι η μελωδία, ο ρυθμός και η αρμονία. Αυτά τα χαρακτηριστικά μπορούν να δημιουργήσουν συναισθήματα όπως είναι η νοσταλγία, η χαρά, λύπη, φόβος, θυμός, άγχος και άλλα. Βέβαια, η πρόκληση αυτών των συναισθημάτων είναι απόρροια του μουσικού κομματιού, του χαρακτήρα και της ψυχοσύνθεσης του ίδιου του ανθρώπου που την ακούει και τέλος του περιβάλλοντος όπου ανήκει την παρούσα φάση (Juslin P., 2008).

Ένα είδος μουσικής με ερευνητικό ενδιαφέρον είναι και η ακουσματική μουσική, μια μορφή ηλεκτρο-ακουστικής μουσικής. Ο όρος αυτός εμφανίζεται για πρώτη φορά στην Αρχαία Ελλάδα, όπου ο φιλόσοφος Πυθαγόρας έκανε διαλέξεις στους μαθητές του πίσω από μια κουρτίνα, διότι πίστευε πως έτσι θα εξασφάλιζε την προσοχή των μαθητών του, αφού θα έδιναν σημασία στους ήχους της φωνής του και τα λόγια του και όχι στην εικόνα τους ή την κίνηση του, τα οποία αναγκαστικά θα φαντάζονταν. Ωστόσο ήταν το 1974, τότε που ο όρος αυτός προτάθηκε από τον François Bayle ως ο καταλληλότερος. Ειδικότερα, η ακουσματική μουσική χρησιμοποιεί ήχους χωρίς οπτική συνοδεία, δηλαδή οι ακροατές δεν είναι σε θέση να δουν το αίτιο παραγωγής ή τις κινήσεις του ερμηνευτή που παράγουν τον ήχο, με αποτέλεσμα να εξασφαλίζεται η προσοχή των ακροατών στις ιδιότητες του ήχου. Αυτοί οι αφηρημένοι ήχοι συνήθως δεν ήταν αναγνωρίσιμοι και πολλές φορές παρατηρούνταν πως οι ακροατές έχαναν το ενδιαφέρον τους. Το πρόβλημα αυτό ήρθε να λύσει η διάχυση (diffusion) του ήχου μέσα από ένα σύνολο ηχείων που περιέβαλαν τον ακροατή (Αλεξοπούλου I. 2020). Ο συνθέτης, διαμοιράζοντας μέσω της κονσόλας με διαφορετικό τρόπο την ένταση της μουσικής στο κάθε ηχείο, δημιουργεί στον ακροατή την εντύπωση ότι ο ήχος κινείται στον χώρο (Αλεξοπούλου I. 2020). Από τη δεκαετία του '90 και μετά, δημοφιλία απέκτησε η οκτακάναλη σύνθεση. Στις μέρες μας, η διάχυση του ήχου μπορεί να επιτευχθεί με αυτοματισμό στον υπολογιστή. Έτσι, δεν είναι απαραίτητη η χειροκίνητη διάχυση από την κονσόλα την ώρα της μιας συναυλίας (Αλεξοπούλου I. 2020). Ο πρώτος χώρος με μόνιμα εγκατεστημένα ηχεία για τη διάχυση της μουσικής δημιουργήθηκε το 1973 στη Μπουρζ και λεγόταν Gmebaphone. Έναν χρόνο αργότερα, ιδρύθηκε το Acousmonium από το GRM στο Παρίσι, ενώ στο παρελθόν είχαν υπάρξει κάποια εφήμερα συστήματα διάχυσης ήχου. Ένα από αυτά ήταν το σφαιρικό αμφιθέατρο ηχείων στο Osaka World's Fair το 1970 (Αλεξοπούλου I. 2020).

Δεδομένου όλων των παραπάνω είναι φανερό το ενδιαφέρον που προκύπτει για τα συναισθήματα που προκαλεί ένα τέτοιο είδος μουσικής. Ωστόσο, υπάρχει περιορισμένη βιβλιογραφία που μελετά την συναισθηματική απόκριση του ακροατή σε ακουσματικές μουσικές

συνθέσεις. Η παρούσα διπλωματική εργασία θα προσπαθήσει να ρίξει λίγο φως σε αυτήν την αχαρτογράφητη περιοχή ανίχνευσης των συναισθημάτων που βιώνουν οι ακροατές όταν ακούνε αυτό το συγκεκριμένο είδος μουσικής. Κεντρικό της θέμα αποτελεί η έρευνα γύρω από την απόδοση που είχαν συγκεκριμένοι αλγόριθμοι για την ταξινόμηση των συναισθημάτων αυτών, ύστερα από την ακρόαση ενός ακουσματικού κομματιού με την βοήθεια εξαγωγής των χαρακτηριστικών του.

Φυσικά η εργασία αυτή δεν θα μπορούσε να υλοποιηθεί χωρίς την προσφορά δεδομένων από τον επιβλέποντα καθηγητή μου κ. Θεόδωρο Κωστούλα και την ομάδα του. Τα δεδομένα αυτά μάλιστα, πραγματοποιήθηκαν στο Πανεπιστήμιο Bournemouth, όπου η ομάδα αυτή των ερευνητών καλούσε φοιτητές, σπουδαστές και προσωπικό να συμμετάσχει. Η διαδικασία αυτή αποτελούταν από δύο μέρη:

1. ελεγχόμενες συνεδρίες ακρόασης ακουσματικής μουσικής με χρήση της πλατφόρμας crowdsourcing<sup>1</sup> για τη συλλογή δεδομένων επισημειωμένων συναισθημάτων, με το τελευταίο να μην περιλαμβάνεται στην παρούσα εργασία και
2. την πολυτροπική (multimodal) συλλογή δεδομένων κατά την ακρόαση ακουσματικής μουσικής.

Το δεύτερο μέρος περιλάμβανε την επισημείωση/σχολιασμό των συλλεγόμενων δεδομένων, χρησιμοποιώντας αυτοσχολιασμό με τη ρητή συμφωνία των χρηστών με κατάλληλο εργαλείο. Η διπλωματική αυτή εργασία, προσεγγίζει τα συναισθήματα των εννέα συμμετεχόντων κατά την ακρόαση ενός μουσικού κομματιού, διάρκειας 7 λεπτών και 19 δευτερολέπτων και την επισημείωση των χαρακτηριστικών του.

---

<sup>1</sup> Crowdsourcing ή πληθοπορισμός είναι η πράξη της εξωτερικής ανάθεσης καθηκόντων, που παραδοσιακά εκτελούνταν από υπάλληλο ή εργολάβο, σε μια μεγάλη ομάδα εθελοντών ή μία κοινότητα, μέσω ανοικτής πρόσκλησης (BIKIPAIΔΕΙΑ).

## ***Ευχαριστίες***

*Στο σημείο αυτό οφείλω να ευχαριστήσω και πάλι τον καθηγητή Θεόδωρο Κωστούλα και την ομάδα του για την σπουδαία συνεργασία και τις πολύτιμες συμβουλές τους. Παράλληλα, θα ήθελα να ευχαριστήσω την οικογένεια μου για την αδιάκοπη υποστήριξή τους μέσα σε όλα αυτά τα χρόνια σπουδών.*

© 2022

της

ΜΑΡΙΑ ΚΑΛΛΙΟΠΗ ΝΤΟΜΠΡΗ

Τμήμα Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων

ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

Η σελίδα αυτή είναι σκόπιμα λευκή.

## Περιεχόμενα

<b>1</b>	<b>Εισαγωγή</b>	<b>1</b>
1.1	Αντικείμενο της διπλωματικής εργασίας	1
1.1.1	Στόχος	2
1.2	Κίνητρο	2
1.3	Δομή της διπλωματικής εργασίας	3
<b>2</b>	<b>Υπόβαθρο</b>	<b>4</b>
2.1	Συναισθηματική ανάλυση	4
2.1.1	Ανάλυση συναισθηματικής κατάστασης	4
2.2	Μηχανική μάθηση	5
2.3	Αλγόριθμοι μηχανικής μάθησης	6
2.3.1	SVM	6
2.3.2	KNN	7
2.3.3	J48 Δέντρα απόφασης	9
<b>3</b>	<b>Δεδομένα</b>	<b>11</b>
3.1	Πηγές δεδομένων	11
3.2	Εξαγωγή παραθύρων ήχου	13
3.3	Εξαγωγή χαρακτηριστικών	13
3.3.1	Εξαγωγή χαρακτηριστικών μουσικού κομματιού	14
3.3.2	Ανάλυση χαρακτηριστικών ψηφιακών σημάτων	14
3.4	Επεξεργασία μετρικών δεδομένων των χρηστών	17
3.4.1	Δημιουργία ετικετών για την πρώτη υπόθεση	18
3.4.2	Δημιουργία ετικετών για την δεύτερη υπόθεση	18
3.5	Τρίτη υπόθεση για δημιουργία νέων κλάσεων	19
3.6	Σύγκριση επισημειώσεων του ακουσματικού κομματιού με τα αντίστοιχα σήματα των μετρικών Arousal-Valence	20
<b>4</b>	<b>Πειράματα εξόρυξης δεδομένων</b>	<b>27</b>
4.1	Προετοιμασία πειραμάτων	27
4.2	Κλάσεις ανά αρχείο	28
4.2.1	Πρώτη υπόθεση	28
4.2.2	Δεύτερη υπόθεση	30
4.3	Μετρικές κατηγοριοποίησης	32
4.3.1	Επιδόσεις ανά κλάση	32
4.3.2	Καμπύλες ROC	33

4.4	Δομή πειραμάτων.....	34
4.5	Αποτελέσματα πρώτης προσέγγισης.....	35
4.5.1	<i>SVM</i> .....	35
4.5.2	<i>KNN</i> .....	38
4.5.3	<i>J48</i> .....	41
4.5.4	Συγκριτικά αποτελέσματα .....	45
4.6	Αποτελέσματα δεύτερης προσέγγισης.....	46
4.6.1	<i>SVM</i> .....	46
4.6.2	<i>KNN</i> .....	49
4.6.3	<i>J48</i> .....	52
4.6.4	Συγκριτικά αποτελέσματα .....	55
5	Συζήτηση.....	57
	<b>Βιβλιογραφία.....</b>	<b>59</b>



## Λίστα Εικόνων

Εικόνα 1: Ο κυκλικός δίσκος των συναισθημάτων του Russell, όπως απεικονίστηκε στην πρωτότυπη δημοσίευσή του, αποτυπώνοντας 8 συναισθήματα (αριστερή εικόνα (α)) και 28 συναισθήματα (δεξιό εικόνα (β)) (Russell, 1980).....	5
Εικόνα 2: Κατάταξη με βάση τον KNN. ....	7
Εικόνα 3: Δέντρα απόφασης (Faiza Bawah, N. Ussiph, 2018). ....	9
Εικόνα 4: Ηλεκτρόκαρδιογράφημα .....	12
Εικόνα 5: Μέτρηση της ηλεκτροδερμικής δραστηριότητας.....	12
Εικόνα 6: Αντιστοίχιση συναισθημάτων με βάση την τρίτη υπόθεση .....	20
Εικόνα 7: Πίνακας σύγχυσης (Κύρκου Ε. 2015).....	32
Εικόνα 8: Η καμπύλη ROC (Κύρκου Ε. 2015). ....	34
Εικόνα 9: Το δέντρο κατά τον έλεγχο στις κλάσεις της Arousal .....	42
Εικόνα 10: Το δέντρο που προκύπτει κατά τον έλεγχο στις κλάσεις της Valence.....	43
Εικόνα 11: Το δέντρο που προκύπτει κατά τον έλεγχο των συναισθημάτων .....	45
Εικόνα 12: Το δέντρο που σχηματίζεται κατά τον έλεγχο με τη χρήση J48. ....	53
Εικόνα 13: Το δέντρο κατά τον έλεγχο με τη χρήση J48. ....	54
Εικόνα 14: Το δέντρο του J48 κατά την ταξινόμηση σε κλάσεις συναισθημάτων .....	55
Εικόνα 15: Οι εντολές που δόθηκαν στην openSMILE για την εξαγωγή δεδομένων.....	63
Εικόνα 16: Το μήνυμα που εμφανίστηκε για την εξαγωγή 1580 χαρακτηριστικών. ....	63
Εικόνα 17: Η δημιουργία δέκατης στήλης με τις μέσες τιμές της διέγερσης κατά την πρώτη υπόθεση .....	64
Εικόνα 18: Η εντολή για έλεγχο θετικών και αρνητικών ετικετών και η επιστροφή Boolean τιμών .....	65
Εικόνα 19: Η δημιουργία ετικετών της μετρικής Arousal κατά την δεύτερη υπόθεση.....	65
Εικόνα 20: Η προσθήκη των κλάσεων Arousal στο αρχείο .....	66
Εικόνα 21: Η προσθήκη ετικετών συναισθήματος κατά την τρίτη υπόθεση .....	66



## Λίστα Γραφημάτων

Γράφημα 1: Το σήμα του Arousal της πρώτης υπόθεσης από όλους τους χρήστες.....	21
Γράφημα 2: Τα χαρακτηριστικά του ακουσματικού κομματιού που σημειώνονται κατά τη διάρκεια του ..	21
Γράφημα 3: Το σήμα του Valence για όλους τους χρήστες, σύμφωνα με την πρώτη υπόθεση.....	22
Γράφημα 4: Τα σήματα των Arousal για τους πρώτους τρεις χρήστες .....	22
Γράφημα 5: Τα σήματα των Arousal για τους τελευταίους τρεις χρήστες.....	23
Γράφημα 6: Τα σήματα των Arousal για τους τρεις επόμενους χρήστες.....	23
Γράφημα 7: Τα σήματα των Valence για τους πρώτους χρήστες.....	24
Γράφημα 8: Τα σήματα των Valence για τους επόμενους τρεις χρήστες.....	24
Γράφημα 9: Τα σήματα Valence για τους τελευταίους τρεις χρήστες .....	25
Γράφημα 10: Οι κλάσεις για την μετρική Arousal ανά πλήθος δεδομένων .....	29
Γράφημα 11: Οι κλάσεις για την μετρική Valence ανά πλήθος δεδομένων.....	29
Γράφημα 12: Οι κλάσεις για τα συναισθήματα ανά πλήθος δεδομένων.....	30
Γράφημα 13: Οι κλάσεις για την μετρική Arousal ανά πλήθος δεδομένων .....	30
Γράφημα 14: Οι κλάσεις για την μετρική Valence ανά πλήθος δεδομένων.....	31
Γράφημα 15: Οι κλάσεις για τα συναισθήματα ανά πλήθος δεδομένων.....	31
Γράφημα 16: Όλες οι τιμές των ROC Area που προέκυψαν στα τεστ ελέγχου κατά την πρώτη υπόθεση...45	
Γράφημα 17: Όλες οι τιμές των ROC Area που προέκυψαν στα τεστ ελέγχου κατά την δεύτερη υπόθεση	56



## Λίστα Πινάκων

Πίνακας 1: Τα χαρακτηριστικά του αρχείου .....	16
Πίνακας 2: Συνδυασμός Arousal-Valence για αντιστοίχιση συναισθημάτων.....	19
Πίνακας 3: Αντιστοίχιση κλάσεων με τα χαρακτηριστικά του ακουσματικού κομματιού .....	26
Πίνακας 4: Τα μέτρα απόδοσης συναρτήσεων καλύτερων τιμών των παραμέτρων του SVM.....	35
Πίνακας 5: Τα μέτρα απόδοσης ΤΟΥ Arousal κατά τον έλεγχο με τον SVM.....	36
Πίνακας 6: Τα μέτρα απόδοσης συναρτήσεων καλύτερων τιμών των παραμέτρων του SVM.....	36
Πίνακας 7: Τα μέτρα απόδοσης κατά τον έλεγχο με τον SVM.....	37
Πίνακας 8: Τα μέτρα απόδοσης συναρτήσεων καλύτερων τιμών των παραμέτρων του SVM.....	37
Πίνακας 9: Τα μέτρα απόδοσης κατά τον έλεγχο για τα συναισθήματα με τον SVM.....	38
Πίνακας 10: Τα μέτρα απόδοσης εκπαίδευσης συναρτήσεων των καλύτερων τιμών των παραμέτρων του KNN .....	39
Πίνακας 11: Τα μέτρα απόδοσης του ελέγχου με τον KNN .....	39
Πίνακας 12: Τα μέτρα απόδοσης εκπαίδευσης συναρτήσεων των καλύτερων τιμών των παραμέτρων του KNN .....	39
Πίνακας 13: Τα μέτρα απόδοσης κατά τον έλεγχο του Valence με τον KNN .....	40
Πίνακας 14: Τα μέτρα απόδοσης εκπαίδευσης συναρτήσεων των καλύτερων τιμών των παραμέτρων του KNN .....	40
Πίνακας 15.....	41
Πίνακας 16: Τα μέτρα απόδοσης της εκπαίδευσης συναρτήσεων των καλύτερων τιμών των παραμέτρων του J48 .....	41
Πίνακας 17: Τα μέτρα απόδοσης ελέγχου του J48.....	42
Πίνακας 18: Τα μέτρα απόδοσης κατά την εκπαίδευσης συναρτήσεων των καλύτερων τιμών των παραμέτρων του J48.....	43
Πίνακας 19: Τα μέτρα απόδοσης ελέγχου του J48 για την μετρική Valence.....	43
Πίνακας 20: Τα μέτρα απόδοσης για την εκπαίδευση συναρτήσεων των καλύτερων τιμών των παραμέτρων του J48.....	44
Πίνακας 21: Τα μέτρα απόδοσης ελέγχου του J48 για την μετρική των συναισθημάτων .....	44
Πίνακας 22: Τα μέτρα απόδοσης εκπαίδευσης συναρτήσεων των καλύτερων τιμών των παραμέτρων του SVM .....	46
Πίνακας 23: Τα μέτρα απόδοσης ελέγχου για τον SVM.....	46
Πίνακας 24: Τα μέτρα απόδοσης εκπαίδευσης του κατηγοριοποιητή SVM συναρτήσεων των καλύτερων τιμών των παραμέτρων του .....	47
Πίνακας 25: Τα μέτρα απόδοσης ελέγχου της μετρικής Valence του SVM .....	47
Πίνακας 26: Τα μέτρα απόδοσης εκπαίδευσης του κατηγοριοποιητή SVM συναρτήσεων των καλύτερων τιμών των παραμέτρων του .....	48
Πίνακας 27: Τα μέτρα απόδοσης ελέγχου του SVM για την μετρική των συναισθημάτων .....	48
Πίνακας 28: Τα μέτρα απόδοσης συναρτήσεων των καλύτερων τιμών των παραμέτρων του KNN.....	49
Πίνακας 29: Τα μέτρα απόδοσης ελέγχου του KNN.....	49
Πίνακας 30: Τα μέτρα απόδοσης συναρτήσεων των καλύτερων τιμών των παραμέτρων του KNN.....	50

Πίνακας 31: Τα μέτρα απόδοσης ελέγχου του KNN για την μετρική Valence.....	50
Πίνακας 32: Τα μέτρα απόδοσης κατά την εκπαίδευση συναρτήσεων των καλύτερων τιμών των παραμέτρων του KNN.....	51
Πίνακας 33: Τα μέτρα απόδοσης ελέγχου του KNN για την μετρική των συναισθημάτων .....	51
Πίνακας 34: Τα μέτρα απόδοσης της εκπαίδευσης συναρτήσεων των καλύτερων τιμών των παραμέτρων του J48 .....	52
Πίνακας 35: Τα μέτρα απόδοσης ελέγχου του J48 για την μετρική Arousal .....	52
Πίνακας 36: Τα μέτρα απόδοσης της εκπαίδευσης συναρτήσεων των καλύτερων τιμών των παραμέτρων του J48 .....	53
Πίνακας 37: Τα μέτρα απόδοσης ελέγχου του J48 για την μετρική J48 .....	53
Πίνακας 38: Πίνακας εκπαίδευσης για τα συναισθήματα με τον κατηγοριοποιητή J48.....	54
Πίνακας 39: Πίνακας ελέγχου για τα συναισθήματα του J48 .....	54



## Ακρωνύμια

ARFF	Attribute-Relation File Format
J48	Decision Tree
KNN/IBK	K Nearest Neighbor
LLD	Low-Level descriptors
MATLAB	Matrix Laboratory
NLP	Natural Language Processing
OPENSMMILE	Munich Open-Source Media Interpretation by Large feature-space Extraction
SMO	Sequential Minimal Optimization
SVM	Support Vector Machines
WEKA	Waikato Environment for Knowledge Analysis



## Περίληψη

Οι τελευταίες δεκαετίες χαρακτηρίζονται από την ιλιγγιώδη ταχύτητα ανάπτυξης διαφόρων επιστημών χάρη στην παράλληλη πρόοδο της τεχνολογίας. Ένας από τους πιο δημοφιλείς τομείς είναι αυτός της Συναισθηματικής Ανάλυσης σε συνδυασμό με τη Μηχανική Μάθηση. Με την παροχή διαφόρων εργαλείων πλέον ο άνθρωπος είναι ικανός να πειραματιστεί και επεξεργαστεί μεγάλο όγκο δεδομένων εφαρμόζοντας μεθόδους από πολλά ερευνητικά πεδία ταυτόχρονα.

Η παρούσα μεταπτυχιακή διπλωματική εργασία πραγματεύεται την αναγνώριση και κατηγοριοποίηση του συναισθήματος που προκαλεί η ακουσματική μουσική. Η αναγνώριση βασίστηκε πρώτα σε εκείνα τα γνωρίσματα που προκύπτουν από την ανάλυση ψηφιακού σήματος ακουσματικής μουσικής. Ειδικότερα, η εξαγωγή και επεξεργασία των χαρακτηριστικών αυτών έγινε με την εισαγωγή του ακουσματικού κομματιού στο εργαλείο openSMILE με τη χρήση της βάσης audiovisual-audio.

Τα δεδομένα που προκύπτουν είναι η βάση, ώστε να ακολουθήσει η κατηγοριοποίηση δεδομένων αυτών με βάση τα συναισθήματα που προέκυψαν από κάποιους ακροατές, των οποίων η επεξεργασία γίνεται στο περιβάλλον της Matlab μέσω των μετρικών διέγερσης-arousal και valence.

Σε ένα επόμενο πλαίσιο γίνεται συγκριτική ανάλυση μιας σειράς αλγορίθμων Μηχανικής Μάθησης με το εργαλείο εξόρυξης δεδομένων, την WEKA. Οι αλγόριθμοι που θα αναλυθούν και δοκιμαστούν είναι τα Δέντρα Απόφασης (J48), Μηχανές Διανυσμάτων Υποστήριξης (SVM) και οι K-πλησιέστεροι γείτονες (KNN). Συγκεκριμένα, στοχεύεται κατά πόσο αυτοί οι αλγόριθμοι διαφοροποιούνται στα αποτελέσματά τους, ενώ παραθέτονται συγκεντρωτικά διαγράμματα των αποδόσεων τους.

**Λέξεις Κλειδιά:** *Μηχανική Μάθηση, συναισθήματα, διέγερση, ψηφιακά σήματα, ακουσματική μουσική, WEKA, openSMILE.*

## Abstract

In recent years, Machine Learning has taken the technological world by storm, as researchers have plethora tools which make them able to experiment with a big amount of data. One of the most popular research fields is that one of emotion recognition through audio signals. Music emotion recognition is a subfield of Music Information Retrieval that deals with music classification and develops music similarity measures, using signal processing and machine learning techniques.

The purpose of this thesis is to present the recognition and classification of the emotion expressed through acousmatic music. At first, the feature extraction and processing performed using the openSMILE software, by utilizing a prerecorded acousmatic song that capture specific characteristics from the configuration file audiovisual-audio. Secondly, this thesis followed an experimental procedure to extract information regarding to another database of arousal and valence of each listener. The environment of Matlab was used to extract emotional labels, describing musical audio.

The database, which had extracted, was used by WEKA, the software for data analysis and classification. Moreover, the classification of samples was done using this software and with the utilization of three different metrics for classification, arousal, valence and emotions. The classification algorithms that will be analyzed and used are the following: Decision Tree (J48), K Nearest Neighbors (KNN) and Support Vector Machines (SVM). Furthermore, we analyzed the effectiveness of each technique per data set and we presented detailed diagrams of their results.

**Keywords:** *Machine Learning, arousal, valence, emotions, acousmatic music, signal processing, WEKA, openSMILE.*

# 1

## *Εισαγωγή*

### *1.1 Αντικείμενο της διπλωματικής εργασίας*

Η αναγνώριση των συναισθημάτων των ανθρώπων, εκτεθειμένων στην μουσική είναι ένας δημοφιλής τομέας έρευνας αλλά και απαιτητικός στον κλάδο της τεχνολογίας. Οι έρευνες που έχουν γίνει γύρω από τα διάφορα είδη μουσικής με ή χωρίς στίχους προτείνουν ποικίλες και διαφορετικές τεχνικές, στις οποίες χρησιμοποιούνται διαφορετικοί αλγόριθμοι μηχανικής μάθησης σε συνδυασμό με τα χαρακτηριστικά (features).

Από τη σκοπιά της επιστήμης της μουσικής, αυτή η εργασία θα δώσει μια εικόνα για τα συναισθήματα του ακροατή στην ακουσματική μουσική. Από την άποψη της συναισθηματικής υπολογιστικής, είναι σημαντικό να προσδιοριστούν οι μέθοδοι που απαιτούνται για την αποτύπωση συναισθημάτων ως απάντηση στην ακουσματική μουσική. Το τελευταίο θα επέτρεπε την ερμηνεία των απαντήσεων σε αυτό το είδος καλλιτεχνικών ερεθισμάτων. Ειδικότερα, στην παρούσα εργασία εξετάζεται η περίπτωση κατά την οποία εννέα χρήστες εκτίθενται σε ένα κομμάτι ακουσματικής μουσικής και οι αντιδράσεις που είχαν κατά την ακρόασή του. Αρχικά γίνεται η ψηφιακή ανάλυση του ακουσματικού κομματιού για την εξαγωγή των χαρακτηριστικών του. Μελετώνται επίσης, διάφοροι αλγόριθμοι μηχανικής μάθησης καθώς γίνεται και η σύγκριση των αποτελεσμάτων τους. Δεδομένου πως γύρω από την αναγνώριση συναισθημάτων υπό την επήρεια μουσικής υπάρχει αρκετή έρευνα, αλλά για το συγκεκριμένο είδος ελάχιστη, η διπλωματική αυτή θα προσφέρει μια επιπλέον προσέγγιση και πληροφορία. Όλα τα πειράματα έγιναν σε πλατφόρμες που ήδη πολλοί επιστήμονες έχουν προτείνει σε δική τους έρευνα, όπως και οι επιλογές αλγορίθμων.

Αυτό το μοντέλο είναι και αυτό που η παρούσα διπλωματική θα ακολουθήσει προκειμένου να προσεγγίσει στα πειράματα τα συναισθήματα που νιώθει ένας ακροατής ακουσματικής μουσικής

μέσω των δοσμένων μετρικών Arousal και Valence των συμμετεχόντων, μετρικές που αποτυπώνουν ανάλογα με την τιμή τους γνωστά συναισθήματα. Μάλιστα, προκειμένου να προσεγγιστούν τα πειράματα έχοντας ως βάση τις θετικές ή αρνητικές ετικέτες των Valence-Arousal, έγινε αντιστοίχιση μέσω μίας υπόθεσης-παραδοχής, με τα αντίστοιχα συναισθήματα, όπως δίνονται από τον πίνακα του Russell (Εικόνα 1).

Η διπλωματική εργασία αυτή βασίστηκε σε πληροφορίες που δόθηκαν από μια ερευνητική ομάδα στο Πανεπιστήμιο του Bournemouth, καθώς συμμετείχαν εννέα σπουδαστές ή και εργαζόμενοι. Αξίζει, ωστόσο, να σημειωθεί πως τα αποτελέσματα που θα αποκαλυφθούν παρακάτω είναι μερικώς ικανοποιητικά, καθώς το δείγμα δεν ήταν τόσο μεγάλο σε πλήθος, πράγμα που θα αύξανε το διάστημα εμπιστοσύνης των πειραμάτων.

### **1.1.1 Στόχος**

Ο στόχος της παρούσας εργασίας είναι πρώτα η ψηφιακή επεξεργασία σημάτων ακουσματικής μουσικής μέσω ενός δημοφιλούς εργαλείου και έπειτα η συγκριτική αξιολόγηση των αποδόσεων αλγορίθμων Μηχανικής Μάθησης κατόπιν ρύθμισης των παραμέτρων τους και όλα αυτά υπό το πλαίσιο των συναισθημάτων που προέκυψαν από τους ακροατές.

Παρά τους στόχους που άμεσα στοχεύει η παρούσα διπλωματική, δημιουργεί και την ανάγκη για περαιτέρω συνέχεια του αντικειμένου της, ώστε σε ένα μελλοντικό πλαίσιο να γίνει έλεγχος των χαρακτηριστικών που εξήχθησαν, ώστε να γίνει επιλογή εκείνων που επηρεάζουν τα συναισθήματα και τις μετρικές τους.

## **1.2 Κίνητρο**

Τις τελευταίες δεκαετίες, παρατηρείται μια αυξανόμενη έρευνα στους τομείς της νευροεπιστήμης, της ψυχολογίας και της συναισθηματικής επιστήμης, καθώς στο παρελθόν ο άνθρωπος έτεινε να αγνοεί τα συναισθήματα του και να υποτιμά την σημασία τους. Στα πλαίσια αυτά είναι φανερή πρόκληση που υπάρχει για το ερευνητικό πεδίο αυτό, καθώς όσο μεγάλη αξία έχει, άλλη τόση πολυπλοκότητα φέρει. Τι συμβαίνει όταν ο στόχος γίνεται πιο συγκεκριμένος καθώς ξεπερνά κατά πολύ την κλασική ανθρώπινη αλληλεπίδραση με το περιβάλλον, αλλά ζητηθεί και ο καθορισμός εκείνων των παραγόντων που προκαλούν τις εκάστοτε ανθρώπινες αντιδράσεις, με σκοπό την βελτίωση του τομέα που μελετάται;

Η αναγνώριση συναισθημάτων αποτελεί ένα άκρως ενδιαφέρον ερευνητικό πρόβλημα μιας και σε συνδυασμό με την αυξανόμενη ζήτηση στη βιομηχανία της μουσικής παραγωγής, μια τέτοια πληροφορία για τα συναισθήματα των ακροατών μπορεί να συνεισφέρει δραστικά στην παραγωγή μουσικών κομματιών σύμφωνα με το σκοπό που έχουν οι παραγωγοί ανά το είδος της μουσικής. Σαφώς λοιπόν, μια τέτοια πληροφορία είναι ένα ισχυρό κίνητρο, παράλληλα όμως και ενώ πολλά είδη μουσικής έχουν μελετηθεί υπό αυτό το πρίσμα, η ακουσματική μουσική αποτελεί το είδος με τη λιγότερη, αν όχι καθόλου πληροφορία. Δεδομένου δε, πως σε πληθώρα μουσικών κομματιών έχουμε εικόνα για το μέσο που παράγει τον ήχο ή ακόμα κι αν το μουσικό κομμάτι συνδυάζεται με στίχους, η ακουσματική μουσική μας προκαλεί ακόμα περισσότερο ενδιαφέρον για τον λόγο πως όλα αυτά λείπουν κατά την παραγωγή της.

### 1.3 Δομή της διπλωματικής εργασίας

Η παρούσα διπλωματική έχει διαρθρωθεί με τρόπο ώστε ο αναγνώστης να ακολουθήσει ομαλά το θέμα και το αντικείμενο που πραγματεύεται. Για το σκοπό αυτό στο αμέσως επόμενο κεφάλαιο δίνονται πληροφορίες για την συναισθηματική ανάλυση και την έρευνα γύρω από αυτό τον τομέα. Ειδικότερα παρατίθενται πληροφορίες γύρω από το μοντέλο συναισθημάτων πάνω σε ένα κυκλικό δίσκο-τριγωνομετρικό κύκλο, που νιώθει το άτομο σύμφωνα με την θεωρία του Russell (Εικόνα 1). Έπειτα, ο αναγνώστης εισάγεται στην έννοια της Μηχανικής Ανάλυσης και μερικών αλγορίθμων που εφαρμόζονται συχνά σε πειράματα. Πιο συγκεκριμένα, ακολουθούν οι ορισμοί των αλγορίθμων SVM, KNN και J48 μαζί με τις παραμέτρους τους και το εύρος της εκάστοτε και τον ρόλο τους.

Στο επόμενο κεφάλαιο, ακολουθεί η εξήγηση της προσέγγισης των πειραμάτων στην εργασία. Πρώτα, δίνονται οι συνθήκες και η διαδικασία που ακολουθήθηκε για τα πειράματα στο Πανεπιστήμιο του Bournemouth. Επίσης, δίνονται πληροφορίες για το πώς προέκυψαν τα δεδομένα για την εργασία και πως επεξεργάστηκαν. Πρώτα, δίνονται πληροφορίες για το πρώτο εργαλείο που χρησιμοποιήθηκε, αυτό της openSMILE και έπειτα για την τεχνική. Αναλυτικότερα, εξηγείται πως έγινε η εξαγωγή παραθύρων ήχου, πως χρησιμοποιήθηκε το εργαλείο και με ποια φίλτρα και τέλος δίνονται λεπτομερώς όσα χαρακτηριστικά προέκυψαν στο εξαγόμενο αρχείο. Στην επόμενη ενότητα περιγράφεται η επεξεργασία των δεδομένων των συναισθημάτων των εννέα ανθρώπων στο περιβάλλον της Matlab, με σκοπό την δημιουργία ετικετών στο προηγούμενο αρχείο. Βέβαια, για τη δημιουργία ετικετών υπήρξαν τρεις υποθέσεις και οι οποίες δίνονται στα επόμενα υποκεφάλαια.

Στο προτελευταίο κεφάλαιο παρουσιάζονται τα πειράματα που έγιναν με το εργαλείο εξόρυξης δεδομένων, αυτό της WEKA και κάθε βήμα που ακολουθήθηκε, καθώς και τα αποτελέσματα που έφεραν οι δυο υποθέσεις για κάθε ένα αλγόριθμο τόσο στο πρώτο βήμα των πειραμάτων κατά την εκμάθηση, όσο και στο τελευταίο του ελέγχου.

Στο τελευταίο κεφάλαιο γίνεται μια συζήτηση από όσα καταγράφηκαν σε όλες τις ενότητες καθώς και τα αποτελέσματα που έφερε η προσέγγιση της εργασίας, δίνοντας έναυσμα για περαιτέρω έρευνα.

# 2

## Υπόβαθρο

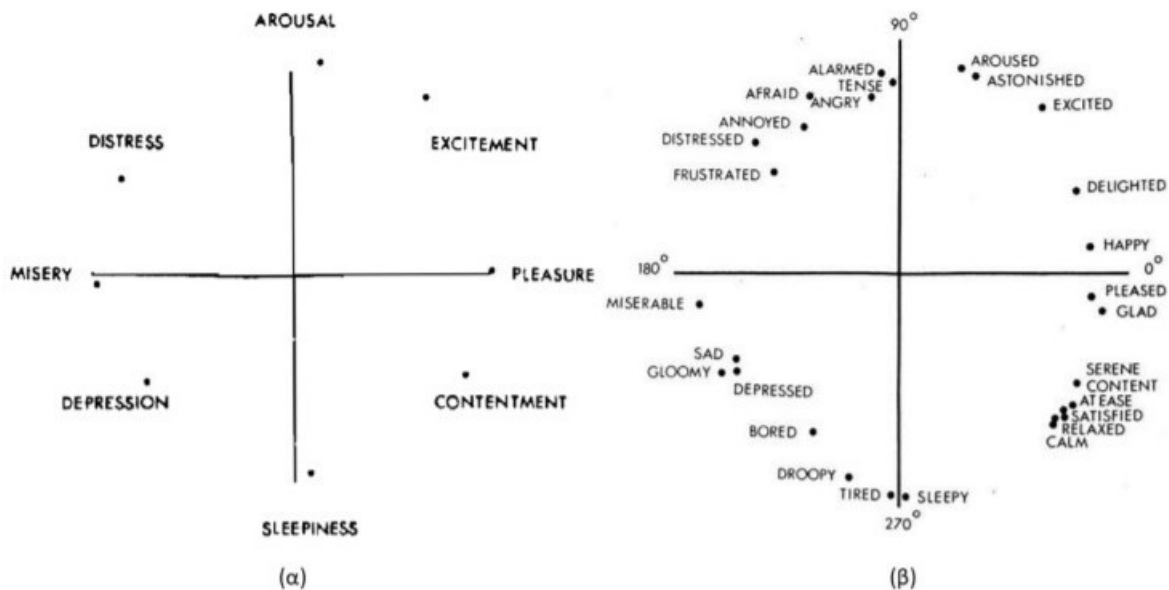
### 2.1 Συναισθηματική ανάλυση

Η συναισθηματική ανάλυση είναι ένα πεδίο που αποσκοπεί στην εύρεση και συλλογή συναισθημάτων που προέρχονται από τη προσωπική άποψη του εκάστοτε ατόμου. Μάλιστα, υπάρχει ένα μοντέλο της, αυτό της Συναισθηματικής Υπολογιστικής-Affective Computing, το οποίο προτάθηκε από την Picard και ασχολήθηκε με τη σημασία των συναισθημάτων του ανθρώπου συναρτήσει του υπολογιστή, προωθώντας περαιτέρω έρευνα σε τομείς της πληροφορικής και ψυχολογίας (Alex G. C. 2020, Picard R. W. 2003). Σκοπός την Συναισθηματικής Υπολογιστικής είναι να εφοδιαστούν οι υπολογιστές με την ικανότητα να αναγνωρίζουν και να εκφράζουν συναισθήματα, γεφυρώνοντας το κενό μεταξύ της συναισθηματικής φύσης του ανθρώπου και του υπολογιστή, αναπτύσσοντας υπολογιστικά συστήματα που αναγνωρίζουν και προσαρμόζονται στη συναισθηματική κατάσταση του χρήστη (Calvo R. 2010).

#### 2.1.1 Ανάλυση συναισθηματικής κατάστασης

Το μοντέλο που θα υιοθετήσει η εργασία αυτή αναφέρεται στην 2D απεικόνιση συναισθημάτων (στο επίπεδο) για την προσέγγιση των συναισθημάτων. Η πρόταση αυτή έγινε από τον James A. Russell, ο οποίος απέδωσε σε ένα τριγωνομετρικό κύκλο κάθε συναίσθημα, τον ονομαζόμενο Circumflex of Affect, και με το κάθε συναίσθημα να απεικονίζεται από μια συντεταγμένη. Στην πραγματικότητα, οι άξονες  $x$  και  $y$  απεικονίζονται από τις μετρικές Valence και Arousal αντίστοιχα, με την πρώτη να αναφέρεται στο πόσο θετικά ή αρνητικά είναι τα συναισθήματα, δηλαδή από ευχάριστα ως δυσάρεστα, και η δεύτερη να αντανακλά τον βαθμό της διέγερσης που το κάθε συναίσθημα μας δημιουργεί (Μαυροφόρος Ν. 2019). Δίνεται η

γραφική παράσταση Valence-Arousal (Διάθεσης – Διέγερσης), στην οποία απεικονίζονται διαφορετικά συναισθήματα ανάλογα με τις τιμές των Valence και Arousal. Αναλυτικότερα, όπως φαίνεται και στην Εικόνα 1, όταν προκύπτουν θετικές τιμές valence-arousal, το άτομο έχει και τα ανάλογα θετικά συναισθήματα, όπως χαρά, ενθουσιασμό. Για αρνητικές τιμές arousal-valence, τα συναισθήματα είναι αυτά που αναγράφονται στο τρίτο τεταρτημόριο και για θετικές τιμές arousal σε συνδυασμό με αρνητικές τιμές valence, παρατηρούνται συναισθήματα, όπως απελπισία, θυμός και στην αντίθετη περίπτωση συναισθήματα όπως ηρεμία, χαλάρωση.



Εικόνα 1: Ο κυκλικός δίσκος των συναισθημάτων του Russell, όπως απεικονίστηκε στην πρωτότυπη δημοσίευσή του, αποτυπώνοντας 8 συναισθήματα (αριστερή εικόνα (α)) και 28 συναισθήματα (δεξιά εικόνα (β)) (Russell, 1980).

## 2.2 Μηχανική μάθηση

Η μηχανική μάθηση είναι ένα από τα ισχυρότερα εργαλεία τεχνητής νοημοσύνης που επιτρέπει να επεξεργαζόμαστε δεδομένα, προκειμένου να αποδίδουν νόημα στον περίπλοκο κόσμο. Η χρήση της γίνεται ευρέως και σε τέτοια κλίμακα παγκοσμίως πάνω σε πληθώρα προβλήματα που έχουν σκοπό να επιλύσουν ζητήματα που προκύπτουν. Στα πλαίσια της μηχανικής μάθησης, υπάρχει και η εξόρυξη δεδομένων, μία από τις πιο σημαντικές εφαρμογές της. Ωστόσο κατά τη χρήση της, συχνά οι χρήστες διαπράττουν λάθη κατά τη διάρκεια αναλύσεων ή κατά την προσπάθεια δημιουργίας σχέσεων μεταξύ πολλαπλών χαρακτηριστικών ή μεταβλητών και έτσι είναι δύσκολη η επίλυση πολλών προβλημάτων. Σε αυτό το σημείο έρχεται η μηχανική μάθηση, η οποία μπορεί να χρησιμοποιηθεί για την επιτυχή επίλυση αυτών των προβλημάτων, βελτιώνοντας την αποτελεσματικότητα των δεδομένων. Το ίδιο σύνολο χαρακτηριστικών αντιπροσωπεύει κάθε εμφάνιση οποιουδήποτε συνόλου δεδομένων που χρησιμοποιείται από τους αλγόριθμους μηχανικής μάθησης. Τα χαρακτηριστικά μπορούν να είναι συνεχή, διακριτά ή και δυαδικής μορφής (Pradeep Kandhasamy J., Balamurali S., 2015). Στα πλαίσια της μηχανικής μάθησης έχουν δημιουργεί αλγόριθμοι ικανοί να λύσουν μερικά

προβλήματα, ανάλογα με τη φύση του προβλήματος ή των χαρακτηριστικών που φέρουν τα δεδομένα.

## 2.3 Αλγόριθμοι μηχανικής μάθησης

Στην εργασία χρησιμοποιήθηκαν τρεις αλγόριθμοι μηχανικής μάθησης και πιο συγκεκριμένα μηχανές διανυσμάτων υποστήριξης (Support Vector Machines, SVM), K πλησιέστεροι γείτονες (KNN) και τα δέντρα απόφασης (J48). Σε όλους τους αλγορίθμους έγινε προσέγγιση με βάση την καλύτερη τιμή για τις παραμέτρους του καθένα. Αξίζει να σημειωθεί πως η επιλογή αυτών των αλγορίθμων οφείλεται στο γεγονός πως στις περισσότερες προτεινόνταν οι παραπάνω αλγόριθμοι.

### 2.3.1 SVM

Οι μηχανές διανυσμάτων υποστήριξης ή μηχανές διανυσματικής υποστήριξης (Support Vector Machines – SVM) θεωρούνται ως οι πιο επιτυχημένοι αλγόριθμοι κατηγοριοποίησης (Russell & Norvig, 2016). Αν υποθέσουμε ότι υπάρχει γραμμική διαχωριστικότητα των δεδομένων προς κατηγοριοποίηση, τότε αυτό που επιτυγχάνει ο αλγόριθμος είναι η επιλογή του βέλτιστου υπερεπιπέδου, όσο αφορά την απόσταση των δύο κλάσεων. Αρχικά επιλέγεται ένας μικρός αριθμός δεδομένων εκπαίδευσης (στιγμιότυπα) από κάθε κλάση, τα οποία ορίζουν το μέγιστο περιθώριο (margin) μεταξύ των δύο κλάσεων. Με τα δεδομένα αυτά κατασκευάζεται μία γραμμική συνάρτηση διάκρισης (discriminant function), η οποία θα διαχωρίζει τα δεδομένα όσο καλύτερα γίνεται (Κωτσάκη Π. 2020). Ωστόσο δεδομένα που δεν είναι γραμμικά διαχωρίσιμα στον αρχικό χώρο εισόδου είναι εύκολα διαχωρίσιμα σε υψηλότερο χώρο διαστάσεων (Κωτσάκη Π. 2020). Ο γραμμικός διαχωριστής είναι στην πραγματικότητα μη γραμμικός στον αρχικό χώρο.

Αξίζει να σημειωθεί πως στα πειράματα της τρέχουσας εργασίας έγινε χρήση του SMO (Sequential Minimal Optimization), ο οποίος είναι η βελτίωση του SVM (Μίγκου Α, 2017). Έτσι, ένας χρήστης έχει την δυνατότητα να διαλέξει τον τύπο πυρήνα (kernel) ανάμεσα σε τέσσερις επιλογές, όπως:

- Linear:  $K(x, x_i) = x^T x_i$
- Polynomial: Ο πολυωνυμικός πυρήνας δύναμης  $d$  είναι της μορφής:  $K(x, x_i) = (x, x_i)^d$
- RBF: Ο γκαουσιανός πυρήνας (Gaussian), γνωστός ως radial basis function με τύπο:  $K(x, x_i) = \exp(-\|x_i - x_j\|^2) / 2\sigma^2$
- Sigmoid:  $K(x, x_i) = \tanh(k(x, x_i) + \theta)$

Πρακτικά, εντός μιας εφαρμογής εξόρυξης δεδομένων, ο αλγόριθμος αυτός έχει πολλές παραμέτρους. Οι παράμετροι, σύμφωνα με τους Alex de Sa et al. 2020 που ο χρήστης μπορεί να προσαρμόσει είναι:

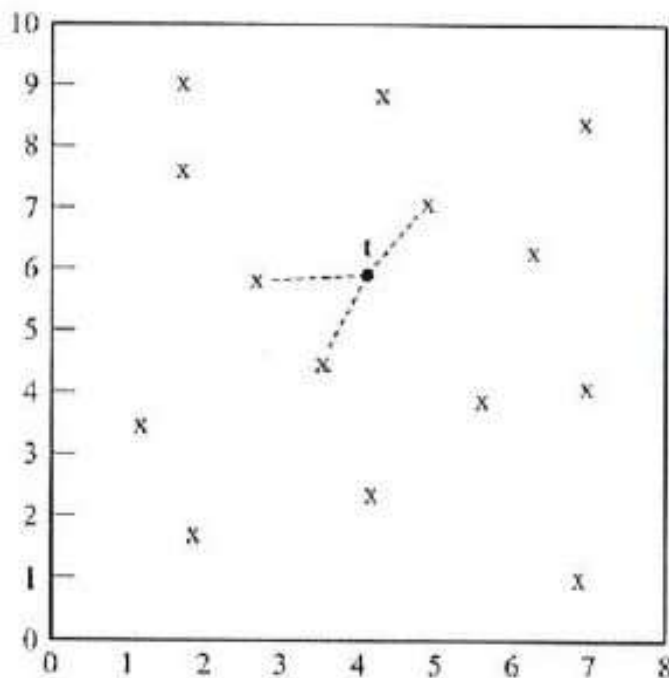
- Κόστος [-C]: Καθορίζει την παράμετρο πολυπλοκότητας, η οποία είναι η παράμετρος τιμωρίας του όρου σφάλματος και έχει εύρος  $\{ C \in \mathbb{R} \mid 0,5 \leq C \leq 1,5 \}$ . Αυτή είναι μια παράμετρος που ελέγχει την αντιστάθμιση μεταξύ σφαλμάτων εκπαίδευσης και πολυπλοκότητας του μοντέλου. Είναι σημαντικό να αναφερθεί ότι μια χαμηλή τιμή του  $C$  θα αυξήσει τον αριθμό των εκπαιδευτικών λαθών, ενώ ένα υψηλή τιμή θα οδηγήσει σε συμπεριφορά παρόμοια με αυτή ενός hard-margin SVM. Προεπιλεγμένη τιμή: 1,0.



- Kernel [-K]: Ο πυρήνας όπως προαναφέρθηκε έχει τις εξής επιλογές όπως και τις δικές του αντίστοιχες παραμέτρους:
  - PolyKernel: είναι ο τυπικός πολυωνυμικός πυρήνας με παραμέτρους:
    1. Εκθέτης [-E]: Καθορίζει την τιμή του εκθέτη και ορίζεται από το διάστημα:  $\{E \in \mathbb{R} \mid 0,2 \leq E \leq 5,0\}$ .
    2. Χρήση χαμηλότερης τάξης [-L]: που αποφασίζει εάν ο αλγόριθμος θα χρησιμοποιήσει όρους χαμηλότερης τάξης ή όχι. Μπορεί να πάρει τιμές Boolean (true ή false).
  - NormalizedPolyKernel: είναι ο κανονικοποιημένος πολυωνυμικός πυρήνας με παραμέτρους:
    1. Εκθέτη [-E]: Καθορίζει την τιμή του εκθέτη και ορίζεται από το διάστημα:  $\{E \in \mathbb{R} \mid 0,2 \leq E \leq 5,0\}$ .
    2. Χρήση χαμηλότερης τάξης [-L]: που αποφασίζει εάν ο αλγόριθμος θα χρησιμοποιήσει όρους χαμηλότερης τάξης ή όχι. Μπορεί να πάρει τιμές Boolean (true ή false).
  - RBF: Ο πυρήνας RBF. Παράμετροι:
    1. Γάμμα [-G]: Η τιμή γάμμα και ορίζεται από το διάστημα:  $\{G \in \mathbb{R} \mid 0,0001 \leq G \leq 1,0\}$ .

Προεπιλεγμένη τιμή: PolyKernel με  $E=1,0$  και  $L=True$ .

### 2.3.2 KNN



Εικόνα 2: Κατάταξη με βάση τον KNN.

Ο αλγόριθμος K πλησιέστερου γείτονα αποτελεί μία από τις γνωστότερες μεθόδους κατηγοριοποίησης, βασισμένη στην απόσταση και τις διαφορές μετρικές της (Μίγκου Α. 2017). Το όνομα του προέρχεται από την λειτουργία του, μιας και η τιμή της συνάρτησης-στόχου για ένα νέο στιγμιότυπο βασίζεται στις αντίστοιχες τιμές των K πλησιέστερων στιγμιότυπων εκπαίδευσης, τα οποία είναι οι λεγόμενοι γείτονες του (Μίγκου Α. 2017).

Ο αλγόριθμος αυτός λειτουργεί πολύ απλά, καθώς απαιτεί μόνο δυο παραμέτρους για να λειτουργήσει. Η πρώτη είναι ο ορισμός της απόστασης των γειτόνων, δηλαδή η μετρική που θα επιλεγθεί για την εύρεση της εγγύτητας

ενός σημείου από τους γείτονες του. Η πιο δημοφιλής μετρική είναι η ευκλείδεια απόσταση (Euclidian distance), ενώ εναλλακτικές μετρικές είναι η Chebychev, Manhattan, Minkowski και

άλλες (Μίγκου Α. 2017). Μία αναλυτική σύγκριση μεταξύ των διαφορετικών μέτρων απόστασης που μπορούν να χρησιμοποιηθούν, έχει πραγματοποιηθεί στην εργασία των Chomboon et al., (2015). Αφού επιλεγθεί η μετρική, υπολογίζεται η απόσταση του σημείου (στόχου) από τους  $K$  γείτονές του (Εικόνα 2). Οι αποστάσεις αυτές τοποθετούνται σε σειρά από την μικρότερη στη μεγαλύτερη και επιλέγεται ως πρόβλεψη η κλάση με την πλειονότητα των τιμών εντός αυτής της περιοχής, σε περίπτωση ταξινόμησης (Μίγκου Α. 2017). Η δεύτερη παράμετρος είναι η τιμή του  $K$ . Το εύρος των τιμών του  $K$  προφανώς είναι μεγάλο, ωστόσο διακρίνονται οι περιπτώσεις ενός πολύ μικρού  $K$ , υπό την προϋπόθεση δε ότι σύνολο δεδομένων είναι μεγάλο, τότε τα αποτελέσματά δεν θα είναι έγκυρα, και ενός  $K$  πολύ μεγάλου, το οποίο επηρεάζει και την απόσταση, καθώς το νέο δείγμα θα εξαρτάται σε μεγάλο βαθμό από ολόκληρο το σύστημα (Cover & Hart, 1967). Στα περιβάλλοντα μηχανικής μάθησης, ο χρήστης καλείται να χειριστεί τις εξής παραμέτρους σύμφωνα με τους Alex de Sa et al. 2020:

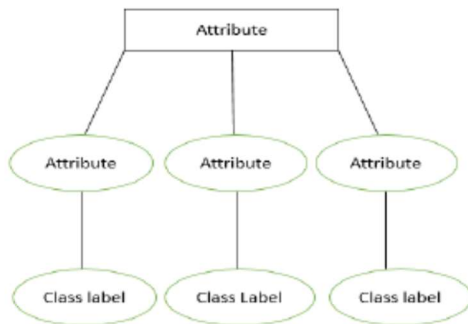
- Αριθμός γειτόνων [-K]: Ο αριθμός των γειτόνων που χρησιμοποιούνται και το εύρος του  $K$  γείτονα προκύπτει από το διάστημα  $\{K \in \mathbb{Z} \mid 1 \leq K \leq 64\}$ .  
Προεπιλεγμένη τιμή  $K=1$ .
- Leave-one-out [-X]: Αποφασίζει αν η leave-one-out τεχνική θα χρησιμοποιηθεί ή όχι στα δεδομένα εκπαίδευσης για επιλογή καλύτερης τιμής  $K$  μεταξύ του 1 και της τιμής που καθορίστηκε ως παράμετρος KNN. Εάν οριστεί ως false, χρησιμοποιείται η επιλεγμένη τιμή  $K$ . Μπορεί να πάρει μόνο τιμές Boolean (true ή false).  
Προεπιλεγμένη τιμή: false.
- Αλγόριθμος πλησιέστερου γείτονα – nearestNeighbourSearchAlgorithm: Προεπιλεγμένος αλγόριθμος ο γραμμικός, ο οποίος μπορεί να τρέξει με διαφορετικές συναρτήσεις/ μέτρα απόστασης, όπως η Ευκλείδεια, η Manhattan (Μίγκου Α. 2007).

- Ευκλείδεια απόσταση: 
$$dis(t_i, t_j) = \sqrt{\sum_{h=1}^k (t_{ih} - t_{jh})^2}$$

- Manhattan απόσταση: 
$$dis(t_i, t_j) = \sum_{h=1}^k |t_{ih} - t_{jh}|$$

Δεν υπάρχουν εξαρτήσεις μεταξύ των παραμέτρων του KNN.

### 2.3.3 J48 Δέντρα απόφασης



Εικόνα 3: Δέντρα απόφασης (Faiza Bawah, N. Ussiph, 2018).

Τα δέντρα αποφάσεων είναι ένα από τους πιο ευρέως χρησιμοποιούμενους αλγόριθμους για επιβλεπόμενα δεδομένα, τα οποία ταξινομούν τα κατηγορικά δεδομένα βασισμένα στα χαρακτηριστικά τους (Faiza Bawah, N. Ussiph, 2018). Τα δέντρα απόφασης, όπως μαρτυρά και το όνομα τους είναι ιεραρχικά δέντρα στη δομή τους αλλά με ρίζα- αρχικό κόμβο, φύλλα κόμβων και κλαδιά (Εικόνα 3). Μάλιστα, κάθε κόμβος αναπαριστά και μια ετικέτα κλάσης (Faiza Bawah, N. Ussiph, 2018). Η τεχνική που ακολουθείται είναι αρκετά δημοφιλής και ευκατανόητη από τους ερευνητές που την χρησιμοποιούν σε μικρά και μεγάλα δεδομένα, τα οποία κατά την ταξινόμηση τους ακολουθούν κανόνες τύπου if-then (Faiza Bawah, N. Ussiph, 2018). Στην πραγματικότητα, κάθε δέντρο απόφασης αντιπροσωπεύει μια συνάρτηση που λαμβάνει ως είσοδο ένα διάλυσμα των τιμών χαρακτηριστικών και επιστρέφει μια "απόφαση" - μια ενιαία τιμή εξόδου (Faiza Bawah, N. Ussiph, 2018). Οι τιμές εισόδου και εξόδου μπορούν να είναι διακριτές ή συνεχείς. Ουσιαστικά, μπορεί να σκεφτεί κανείς ένα δέντρο ταξινόμησης ως ένα μοντέλο πρόβλεψης στο οποίο κάθε κλαδί αποτελεί μία ερώτηση ταξινόμησης, ενώ τα φύλλα αντιπροσωπεύουν τον διαχωρισμό των δεδομένων εισόδου με την ταξινόμησή τους (Jain & Srivastava, 2013).

Παράμετροι που έχει ο αλγόριθμος αυτός σύμφωνα με τους Alex de Sa et al. 2020:

- Συντελεστής εμπιστοσύνης – ConfidenceFactor [-C]: Χρησιμοποιείται για του δέντρου απόφασης το σφάλμα, βασισμένο στην μέθοδο κλαδέματος (οι μικρότερες τιμές απαιτούν περισσότερο κλάδεμα) και ορίζεται από το διάστημα:  $\{C \in \mathbb{R} | 0.0 < C < 1.0\}$ .

Προεπιλεγμένη τιμή: 0,25.

- Ελάχιστος αριθμός αντικειμένων - minNumObj [-M]: Ο ελάχιστος αριθμός των στιγμιοτύπων ανά φύλλο. Μπορεί να πάρει τιμές στο διάστημα:  $\{M \in \mathbb{Z} | 1 \leq M \leq 64\}$ . Προεπιλεγμένη τιμή: 2.

- Σύμπτυξη δέντρου - collapseTree [-O]: Χρησιμοποιείται προκειμένου να αποφασιστεί αν υπάρχουν εσωτερικοί κόμβοι που θα αναδιπλωθούν, ώστε να αποφευχθεί η υπερβολική τοποθέτηση/overfitting. Αυτή η παράμετρος χρησιμοποιείται με τη μέθοδο κλαδέματος βάσει σφαλμάτων του J48 για βελτίωση του τελικού δέντρου απόφασης. Συμπύσσει ένα υποδέντρο σε έναν κόμβο, μόνο εάν το σφάλμα εκπαίδευσης του υποδέντρου δεν αυξάνεται εν συγκρίσει με το ολόκληρο δέντρο. Για παράδειγμα, εάν υπάρχει ένα υποδέντρο με δύο κόμβους φύλλων που έχουν την ίδια ταξινόμηση των δεδομένων στην

εκπαίδευση, αυτό το υποδέντρο θα αντικατασταθεί από ένα μόνο φύλλο. Μπορεί να πάρει τιμές Boolean (true ή false). Προεπιλεγμένη τιμή: true.

- Unpruned [-U]: Αποφασίζει εάν θα πραγματοποιηθεί το κλάδεμα ή όχι. Μπορεί να πάρει Boolean τιμές (true ή false).

Προεπιλεγμένη τιμή: false.

- Binary splits [-B]: Αποφασίζει εάν ο J48 θα χρησιμοποιήσει binary splits (δυναμικό διαχωρισμό), στα ονομαστικά χαρακτηριστικά κατά την κατασκευή των δέντρων. Μπορεί να πάρει Boolean τιμές (true ή false).

Προεπιλεγμένη τιμή: false.

- Ανύψωση υπόδεντρου [-S]: Χρησιμοποιείται για του J48 το σφάλμα βασισμένο στο κλάδεμα και αποφασίζει πότε ο αλγόριθμος θα κάνει ανύψωση υπόδεντρου κατά το κλάδεμα. Μπορεί να πάρει Boolean τιμές (true ή false).

Προεπιλεγμένη τιμή: true.

- Χρήση Laplace [-A]: Αποφασίζει εάν οι σύνολο των στιγμιότυπων στα φύλλα γίνονται πιο ομαλοποιημένα με βάση τη διόρθωση Laplace. Μπορεί πάρει τιμές Boolean (true ή false).

Προεπιλεγμένη τιμή: false.

*Εξαρτήσεις και περιορισμοί υπάρχουν όταν η παράμετρος unpruned έχει οριστεί ως "true", τότε οι παράμετροι «παράγοντας εμπιστοσύνης» δεν χρησιμοποιείται (παραλείπεται).*

# 3

## Δεδομένα

### 3.1 Πηγές δεδομένων

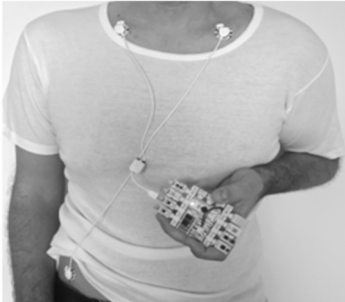
Τα δεδομένα που χρησιμοποιήθηκαν για την παρούσα εργασία προέρχονται από την επεξεργασία της ερευνητικής ομάδας του κ. Θεόδωρου Κωστούλα. Αξίζει να σημειωθεί πως το κομμάτι της ακουσματικής μουσικής είναι επίσης δημιούργημα τους. Αρχικά, με έδρα το πανεπιστήμιο Bournemouth συμμετείχαν εννέα άτομα, τα οποία υποβλήθηκαν στην παρακάτω διαδικασία, αποτελούμενη από δυο μέρη. Πρώτα, υποβλήθηκαν σε ελεγχόμενες συνεδρίες ακρόασης ακουσματικής μουσικής για την συλλογή των συναισθημάτων τους και έπειτα σε πολυτροπική συλλογή δεδομένων. Μάλιστα, αυτό το μέρος περιελάμβανε την επισημείωση και τον σχολιασμό των συλλεγόμενων δεδομένων από τους ίδιους τους χρήστες για τον εαυτό τους. Πιο συγκεκριμένα, ζητήθηκε από τους συμμετέχοντες αυτούς να τοποθετήσουν ένα καλώδιο που συνδέεται με τους αισθητήρες, με τους οποίους πρόκειται να καταγράψουν τα βιομετρικά σήματα. Έπειτα, τους ζητήθηκε να χαλαρώσουν ώστε να ξεκινήσει η αναπαραγωγή της ακουσματικής μουσικής. Ο τρόπος που τοποθετήθηκαν τα καλώδια, φαίνεται στις Εικόνες 4 και 5.

Οι αισθητήρες μέτρησαν την επιτάχυνση, την ηλεκτροδερμική δραστηριότητα και συγχρόνως έγινε ηλεκτροκαρδιογράφημα. Ως μέρος των πειραμάτων χρησιμοποιήθηκαν ένα ή και περισσότερα φυσιολογικά και συμπεριφορικά συστήματα καταγραφής σημάτων. Τα φυσιολογικά σήματα μπορεί να είναι η αντίσταση δέρματος (GSR- Galvanic Skin Response), θερμοκρασία δέρματος, καρδιακός ρυθμός και ρυθμός αναπνοής. Τα σήματα συμπεριφοράς είναι αυτά που επέτρεψαν να αναλυθούν οι δραστηριότητες, όπως η επιτάχυνση, το βίντεο.

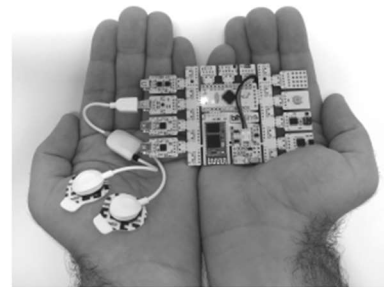
Όσον αφορά τους ανιχνευτές που χρησιμοποιήθηκαν, ήταν:

- Η Bitalino Physiological συσκευή, η οποία είναι ένα ενσωματωμένο σύστημα συλλογής φυσιολογικών δεδομένων,

- Το βίντεο. Σε ορισμένα πειράματα, γίνεται λήψη βίντεο (κανονική ή/και θερμική) που μπορεί να περιλαμβάνει τα πρόσωπα και την ηχογράφιση του πειράματος για λόγους ελέγχου και σχολιασμού- προκειμένου να επαληθευτεί ότι το πείραμα διεξήχθη όπως είχε προγραμματιστεί- και επίσης να μελετήσει πτυχές συμπεριφοράς και
- Τα ενσωματωμένα συστήματα για τη συλλογή δεδομένων στάσης και άλλων δεδομένων συμπεριφοράς.



Εικόνα 4: Ηλεκτροκαρδιογράφημα



Εικόνα 5: Μέτρηση της ηλεκτροδερμικής δραστηριότητας

Επιπλέον, συμπλήρωσαν τεστ προσωπικότητας, τεστ παρούσας διάθεσης, τεστ αλεξυθυμίας καθώς και το συναισθηματικό τεστ PANAS<sup>2</sup>. Αυτά τα ερωτηματολόγια εξυπηρέτησαν στην απόκτηση εικόνας για το ψυχολογικό προφίλ του εκάστοτε ανθρώπου σε σχέση με τις συναισθηματικές του απαντήσεις. Αξίζει να σημειωθεί πως τα ερεθίσματα που δημιουργούνταν κατά την καταγραφή, δεν ήταν δυσάρεστα και μάλιστα υπήρχε η δυνατότητα αυξομείωσης της έντασης του ήχου, αφού ελεγχόταν από τους ίδιους τους ακροατές σύμφωνα με την προτίμηση τους.

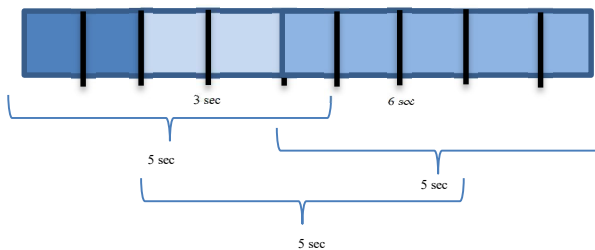
Από τα παραπάνω δεδομένα που προέκυψαν, μετά από την επεξεργασία τους δόθηκαν και χρησιμοποιήθηκαν μόνο οι πίνακες με τους μέσους όρους των τιμών arousal και valence με τις επισημειώσεις κάθε χρήστη κατά τη διάρκεια του μουσικού κομματιού, δημιουργώντας τελικώς 145 στιγμιότυπα. Ουσιαστικά, αυτοί οι πίνακες προεπεξεργάστηκαν από την ερευνητική ομάδα με τρόπο ώστε να γίνει κανονικοποίηση των δεδομένων με εύρος τιμής από 0 ως 1, αντικατοπτρίζοντας το κάθε στιγμιότυπο την κάθε χρονική στιγμή κατά την ακρόαση του κομματιού με την επιλεγείσα επικάλυψη-βήμα. Τα στιγμιότυπα προέκυψαν από την διάρκεια του κομματιού, επιλέγοντας συγκεκριμένο μήκος παράθυρου και βήμα. Περισσότερες λεπτομέρειες θα δοθούν παρακάτω.

---

<sup>2</sup> Το Positive and Negative Affect Schedule (PANAS) είναι ένα ερωτηματολόγιο αυτοαναφοράς που αποτελείται από δύο κλίμακες 10 σημείων για τη μέτρηση τόσο του θετικού όσο και του αρνητικού συναισθήματος. Κάθε στοιχείο βαθμολογείται από μια κλίμακα 5 βαθμών, από 1 (καθόλου) έως το 5 (πολύ). Το μέτρο έχει χρησιμοποιηθεί κυρίως ως ερευνητικό εργαλείο σε ομαδικές μελέτες. Μελέτες έχουν βρει ότι το PANAS είναι ένα αξιόπιστο και έγκυρο εργαλείο για την αξιολόγηση του θετικού και του αρνητικού συναισθήματος (ΒΙΚΙΠΑΙΔΕΙΑ).

### 3.2 Εξαγωγή παραθύρων ήχου

Η πρώτη παράμετρος, που προέκυψε κατά τη διαδικασία εξαγωγής χαρακτηριστικών στην εργασία αυτή, είναι το μήκος (η χρονική διάρκεια) του παραθύρου και η δεύτερη είναι το βήμα (step). Αυτό ορίζει την απόσταση του πρώτου msec ενός παραθύρου από το πρώτο msec του αμέσως επόμενου. Τροποποιώντας κατάλληλα το βήμα και το μέγεθος του παραθύρου, επιτυγχάνεται η επικάλυψη μεταξύ των διαδοχικών παραθύρων και το διάστημα μεταξύ τους. Στα πειράματα για την παρούσα εργασία ορίστηκε μέγεθος παραθύρου ίσο με 5 sec και βήμα ίσο με 3 sec, προκειμένου να συμβαδίζουν με την προϋπάρχουσα επεξεργασία δεδομένων που έγινε κατά τη δημιουργία των πινάκων arousal και valence.



### 3.3 Εξαγωγή χαρακτηριστικών

Η εξαγωγή των χαρακτηριστικών του μουσικού κομματιού έγινε με τη βοήθεια της εφαρμογής OpenSMILE 3.0.0<sup>3</sup>, ένα ευέλικτο και γρήγορο εργαλείο ανοιχτού πηγαίου κώδικα (open source) για την εξαγωγή χαρακτηριστικών ήχου, γραμμένο σε γλώσσα C++ (F. Eyben 2014), το οποίο χτίστηκε πάνω στα Windows 10 για τους σκοπούς του πειράματος της παρούσας διπλωματικής.

Η Munich Open-Source Media Interpretation by Large feature-space Extraction (openSMILE) είναι μια εργαλειοθήκη εξαγωγής χαρακτηριστικών με σκοπό την επεξεργασία σήματος του εισαγόμενου δεδομένου και την εφαρμογή μηχανικής εκμάθησης (F. Eyben 2014, 2015). Η κεντρική ιδέα του εργαλείου αυτού είναι η εξαγωγή χαρακτηριστικών ηχητικού σήματος (F. Eyben 2014). Ωστόσο, η εφαρμογή της openSMILE μπορεί επίσης να χρησιμοποιηθεί και για την ανάλυση σημάτων, όπως τα φυσιολογικά σήματα, οπτικά σήματα και άλλους φυσικούς αισθητήρες, δοθέντων με τις κατάλληλες συνιστώσες εισόδου (F. Eyben 2014). Η openSMILE έχει σχεδιαστεί για επεξεργασία σε πραγματικό χρόνο, αλλά μπορεί επίσης να χρησιμοποιηθεί off-line για την επεξεργασία μεγάλων συνόλων δεδομένων και μάλιστα κατά μικρότερα δεδομένα εισόδου. Αυτό είναι ένα μεγάλο πλεονέκτημα έναντι σε παρόμοιες άλλες εφαρμογές, που έχουν σχεδιαστεί για εξαγωγή off-line, απαιτώντας ολόκληρο το δεδομένο εισαγωγής (F. Eyben 2014). Το openSMILE μπορεί να εξάγει χαρακτηριστικά σταδιακά καθώς νέα δεδομένα εισάγονται. Με τη χρήση δε, της βιβλιοθήκης PortAudio, η openSMILE παρέχει την είσοδο ζωντανού ήχου (Live), δυνατότητα που προσφέρει άμεσα την εξαγωγή χαρακτηριστικών ήχου σε πραγματικό χρόνο (F. Eyben 2014, 2015). Άλλα χαρακτηριστικά που φέρει αυτό το εργαλείο είναι η

<sup>3</sup> <https://github.com/audeering/opensmile>

ανάγνωση και γραφή διαφόρων μορφών δεδομένων-αρχείων, ευρέως χρησιμοποιούμενα από άλλες εργαλειοθήκες εξόρυξης δεδομένων και μηχανικής μάθησης (F. Eyben 2014, 2015). Πιο συγκεκριμένα, μερικά από αυτά που μπορεί να εξάγει είναι αρχεία μορφής PCM WAVE για αρχεία ήχου, CSV (τιμή διαχωρισμένη με κόμμα, μορφή υπολογιστικού φύλλου) και ARFF (Weka Data Mining) για δεδομένα βασισμένα σε κείμενο (F. Eyben 2014).

### 3.3.1 Εξαγωγή χαρακτηριστικών μουσικού κομματιού

Το συγκεκριμένο εργαλείο χρησιμοποιεί ως είσοδο ένα αρχείο configuration και ένα αρχείο ήχου για να εξαγάγει από τον ήχο εκείνα τα χαρακτηριστικά που υποδεικνύει το configuration αρχείο. Στην παρούσα εργασία, εισήχθη το αρχείο ακουσματικής μουσικής με όνομα Ballistichory και χρησιμοποιήθηκε ένα configuration file για την εξαγωγή των χαρακτηριστικών από τον φάκελο AudioVisual, το audio (Εικόνα 15 – Παράρτημα I). Τέλος, ζητήθηκε να γίνει εξαγωγή των χαρακτηριστικών αυτών σε αρχείο ARFF, προκειμένου σε ένα επόμενο στάδιο να εισαχθεί στην εφαρμογή WEKA. Το πλήθος των χαρακτηριστικών που εξήγαγε ήταν 1580 (Εικόνα 16 - Παράρτημα I, Πίνακας 1).

Σημειώνεται πως προτού τρέξει η διαδικασία αυτή στην openSMILE, έγινε τροποποίηση στο default αρχείο του audio configuration προκειμένου να συμπίπτει η προσέγγιση της εργασίας με τα προϋπάρχοντα δεδομένα των χρηστών, όπως αναφέρθηκε παραπάνω. Ειδικότερα, το default αρχείο είχε τις τιμές  $frameSizeFrames = 500$  και  $frameStepFrames = 200$ , το οποίο αυξήθηκε σε  $frameStepFrames = 300$ .

### 3.3.2 Ανάλυση χαρακτηριστικών ψηφιακών σημάτων

Η μέθοδος εξαγωγής ακουστικών χαρακτηριστικών είναι μια συχνή προσέγγιση σε τέτοια πειράματα, κατά τα οποία χρησιμοποιούνται χαρακτηριστικά ως είσοδο ταξινόμησης βασισμένα στην ακουστική συσχέτιση. Διακρίνονται σε δυο κατηγορίες, η πρώτη είναι οι περιγραφείς χαμηλού επιπέδου-LLD και η δεύτερη είναι οι λειτουργικοί-functionals.

Προτού αναλυθούν τα αποτελέσματα που έδωσε το εργαλείο της OpenSMILE, ας οριστούν οι LLD- περιγραφείς χαμηλού επιπέδου. Ένας LLD είναι μια παράμετρος που υπολογίζεται από το σύντομο χρονικό πλαίσιο ορισμένου μήκους από ένα ακουστικό σήμα τη στιγμή. Το μήκος του πλαισίου πρέπει να επιλεγεί με τρόπο ώστε: α) να εξασφαλιστεί η σταθερότητα του σήματος εντός του πλαισίου σε σχέση με το ενδιαφερόμενο LLD, και β) να βεβαιώνει ότι το πλαίσιο περιέχει αρκετά δεδομένα για τον υπολογισμό του LLD. Τα τυπικά μήκη πλαισίου είναι 25–32 ms για τους περισσότερους LLD (F. Eyben 2015).

Μάλιστα, τα χαρακτηριστικά που προέκυψαν από την πρώτη φάση πειράματος φαίνονται στον Πίνακα 1, όπου αποδεικνύεται και το πλήθος τους, όπου:  $(68 * 21) + (8 * 19) = 1.580$ . Αναλυτικότερα, το αποτέλεσμα αυτό προκύπτει καθώς οι πρώτοι 68 συνδυάζονται και με τους 21 functionals, ενώ οι τελευταίοι 8 (από τη θέση 13-20 του Πίνακα 1) εμφανίζονται μόνο με τους 19 functionals, καθώς το  $pctrange$  0-1 και το  $upleveltime75$  λείπουν. Τα χαρακτηριστικά που εξάγονται, σύμφωνα με τον (F. Eyben 2014) είναι:

- $pcm$  loudness είναι η κανονικοποιημένη ένταση υψωμένη εις την δύναμη 0.3.
- $logMelFreqBand$  είναι η λογαριθμική δύναμη των Mel-frequency bands 0-7 (κατανεμημένη σε μια περιοχή από 0 έως 8 kHz)



- `lspFreq` είναι οι συχνότητες φασματικών ζευγών 8 γραμμών που υπολογίζονται από 8 συντελεστές LPC.
- `F0finEnv` είναι το περίβλημα του ομαλοποιημένου περιγράμματος της θεμελιώδους συχνότητας.
- `VoicingFinalUnclipped` είναι η φωνητική πιθανότητα της τελικής υπονήφιας θεμελιώδους συχνότητας. `Unclipped` σημαίνει ότι δεν μηδενίζεται, όταν πέσει κάτω από το όριο φωνής.

Οι παρακάτω τέσσερις LLD που σχετίζονται με τον τόνο και γι' αυτό συναντώνται ως 0 για περιοχές χωρίς φωνή, αντιθέτως όταν υπάρχει ήχος θα εμφανιστούν σύμφωνα με τον F. Eyben 2014 οι:

- `F0final` είναι το ομαλοποιημένο περίγραμμα της θεμελιώδους συχνότητας
- `jitterLocal` είναι το τοπικό (από καρέ σε καρέ) Jitter (αποκλίσεις μήκους περιόδου τόνου)
- `jitterDDP` είναι το διαφορετικό από καρέ σε καρέ Jitter
- `shimmerLocal` είναι το τοπικό (από καρέ σε καρέ) Shimmer (αποκλίσεις πλάτους μεταξύ των τονικών περιόδων).

Το `suffix_sma` που προκύπτει δίπλα από τα ονόματα των LLD, υποδηλώνει ότι ομαλοποιήθηκαν, μετακινώντας το μέσο φίλτρο με μήκος παραθύρου 3. Το `suffix_de`, αμέσως μετά το `sma`, υποδηλώνει ότι το τρέχον χαρακτηριστικό είναι 1ης τάξης συντελεστής δέλτα (διαφορικό) των ομαλοποιημένων LLD (F. Eyben 2014). Παρακάτω δίνονται οι επεξηγήσεις των 21 functionals, όπως εμφανίστηκαν στο αρχείο, σύμφωνα με τον οδηγό της `openSmile` του F. Eyben 2014:

- `maxPos`: Η απόλυτη θέση της μέγιστης τιμής (σε καρέ)
- `minPos`: Η απόλυτη θέση της ελάχιστης τιμής (σε καρέ)
- `Amean`: Ο αριθμητικός μέσος όρος του περιγράμματος
- `linregc1`: Η κλίση (m) μιας γραμμικής προσέγγισης του περιγράμματος
- `linregc2`: Το σύνολο (t) μιας γραμμικής προσέγγισης του περιγράμματος
- `linregerrA`: Το γραμμικό σφάλμα που υπολογίζεται ως η διαφορά της γραμμικής προσέγγισης και της πραγματικό περίγραμμα
- `linregerrQ`: Το τετραγωνικό σφάλμα που υπολογίζεται ως η διαφορά της γραμμικής προσέγγισης και το πραγματικό περίγραμμα
- `Stddev`: Η τυπική απόκλιση των τιμών στο περίγραμμα
- `Skewness`: Η λοξότητα (3<sup>η</sup> σε σειρά στιγμής).
- `Kurtosis`: Η κύρτωση (4<sup>η</sup> σε σειρά στιγμής).
- `quartile1` είναι το πρώτο τεταρτημόριο (25% εκατοστημόριο)
- `quartile2` είναι το πρώτο τεταρτημόριο (50% εκατοστημόριο)
- `quartile3` είναι το πρώτο τεταρτημόριο (75% εκατοστημόριο)
- `iqr1-2`: Το εύρος μεταξύ τεταρτημορίων: τεταρτημόριο2-τεταρτημόριο1
- `iqr2-3`: Το εύρος μεταξύ τεταρτημορίων: τεταρτημόριο3-τεταρτημόριο2

- iqr1-3: Το εύρος μεταξύ τεταρτημορίων: τεταρτημόριο3-τεταρτημόριο1
- percentile1.0 είναι η ελάχιστη τιμή του περιγράμματος που αντιπροσωπεύεται από το 1% του εκατοστημόριου.
- percentile99.0 είναι η μέγιστη τιμή του περιγράμματος, που αντιπροσωπεύεται από το 99% του εκατοστημόριου.
- pctlrange0-1 είναι το ακραίο εύρος ισχυρού σήματος «max-min» αντιπροσωπεύεται από το εύρος του 1% και το 99% εκατοστημόριο.
- upleveltime75 είναι το ποσοστό του χρόνου όπου το σήμα είναι παραπάνω (75% \* εύρος + λεπτά).
- upleveltime90 είναι το ποσοστό του χρόνου όπου το σήμα είναι παραπάνω (90% \* εύρος + λεπτά).

Σημειώνεται, πως από αυτά αφαιρέθηκαν πέντε χαρακτηριστικά, όπως οι κλάσεις που δίνονταν από το configuration file που επιλέχθηκε και το name και frameIndex, προκειμένου και να προστεθεί αργότερα η επιθυμητή κλάση ώστε να μπορέσει το τελευταίο εργαλείο εξόρυξης δεδομένων που χρησιμοποιήθηκε να το ανοίξει, κάνοντας αντιστοιχίσεις κάθε τιμής στο αντίστοιχο χαρακτηριστικό.

**Πίνακας 1: Τα χαρακτηριστικά του αρχείου**

<u>LLD</u>			<u>FUNCTIONALS</u>	
	<b>ΟΝΟΜΑΣΙΑ ΣΤΗΝ OPENSMILE</b>	<b>ΠΛΗΘΟΣ</b>	<b>ΟΝΟΜΑΣΙΑ ΣΤΗΝ OPENSMILE</b>	<b>ΠΛΗΘΟΣ</b>
1	pcm_loudness_sma	1	maxPos, minPos	2
2	pcm_fftMag_mfcc_sma	15	Amean	1
3	logMelFreqBand_sma	8	Linregc1, Linregc2	4
4	lspFreq_sma	8	LinregerrA, LinregerrQ	
5	F0finEnv_sma	1	Stddev	3
6	voicingFinalUnclipped_sma	1	Kurtosis	
7	pcm_loudness_sma_de	1	Skewness	

8	pcm_fftMag_mfcc_sma_de	15	Quartile1-3	6
9	logMelFreqBand_sma_de	8	IqrN, N=1-2,2-3,1-3	
10	lspFreq_sma_de	8	Percentile1.0, Percentile99.0	3
11	F0finEnv_sma_de	1	Pctrange0-1	
12	voicingFinalUnclipped_sma_de	1	Upleveltime75-90	2
13	<i>F0final_sma</i>	1		
14	<i>jitterLocal_sma</i>	1		
15	<i>jitterDDP_sma</i>	1		
16	<i>shimmerLocal_sma</i>	1		
17	<i>F0final_sma_de</i>	1		
18	<i>jitterLocal_sma_de</i>	1		
19	<i>jitterDDP_sma_de</i>	1		
20	<i>shimmerLocal_sma_de</i>	1		
21	Classes	3		
22	frameIndex	1		

### 3.4 Επεξεργασία μετρικών δεδομένων των χρηστών

Αρχικά, κατά την εισαγωγή των δεδομένων των χρηστών (Arousal και Valence) στην Matlab (Matrix Laboratory) που δημιουργήθηκε από τον Cleve Moler σε γλώσσα προγραμματισμού Fortran την δεκαετία του 1970. Το 1984 η εταιρεία Mathworks ιδρύεται αναλαμβάνοντας το έργο που δημιούργησε ο Cleve Moler και η χρήση του Matlab από αποκλειστικά εκπαιδευτική γίνεται τώρα και εμπορική η χρήση του. Σήμερα το πρόγραμμα Matlab είναι γραμμένο σε γλώσσα προγραμματισμού C και Java (Wikipedia, Matlab εγχειρίδιο). Στο περιβάλλον της Matlab έγιναν δυο υποθέσεις:

- Η πρώτη υπόθεση θεωρήθηκε πως όλοι οι χρήστες συμμετείχαν με τον ίδιο τρόπο στην έρευνα, έχοντας κατανοήσει τις οδηγίες και τον τρόπο λειτουργίας των μέσων, ενώ βρίσκονταν σε παρόμοιες καταστάσεις ψυχολογικές και περιβαλλοντικές.
- Η δεύτερη υπόθεση είναι αφορά την αντιμετώπιση τα δεδομένα κάθε χρήστη ως ξεχωριστή οντότητα, ώστε να αποκλείσουμε γενικεύσεις και σφάλματα στα δεδομένα μας και περιπτώσεις που κάποιος χρήστης ενδεχομένως να μην κατανόησε σωστά τις οδηγίες για την συμμετοχή του στο πείραμα.

Σε πρώτη φάση, εισήχθησαν όλα τα δεδομένα τόσο για το Arousal, όσο και για το Valence των εννέα χρηστών σε έναν ενιαίο πίνακα διάστασης 9x145, με σκοπό να προκύψει μια νέα στήλη με τα τελικά δεδομένα, τα οποία και θα κρίνουν την ετικέτα και στις δυο περιπτώσεις. Επίσης επιλέχθηκε να γίνει διαχωρισμός δεδομένων με το 60% να δίνεται για την εκπαίδευση-training και το 40% για τον έλεγχο-test. Έτσι λοιπόν, προέκυψαν 87 instances για την εκμάθηση-training και 58 instances για το test.

### 3.4.1 Δημιουργία ετικετών για την πρώτη υπόθεση

Η πρώτη υπόθεση υπολογίστηκε με τη βοήθεια εντολών στην Matlab. Πρώτος στόχος ήταν ο υπολογισμός του μέσου όρου για κάθε γραμμή του πίνακα. Αυτοί οι μέσοι όροι τοποθετήθηκαν σε μια νέα στήλη και για τα πρώτα 87 instances (του training) και βρέθηκε ο αντίστοιχος δικός τους μέσος όρος, ο οποίος αντιστοιχούσε στη τιμή 0,6955. Αυτή η τιμή καθόρισε τις ετικέτες, ώστε όσες είναι μικρότερες από αυτήν να πάρουν την ετικέτα *negative\_Arousal* και αντίστοιχα για ίσες και μεγαλύτερες τιμές *positive\_Arousal*. Το ίδιο έγινε και για τα υπόλοιπα 58 instances (test) με τον δικό τους μέσο όρο να αντιστοιχεί στην τιμή 0,6304. Η ίδια διαδικασία ακολουθήθηκε και στον πίνακα του Valence, όπου ο τελικός μέσος όρος για το δείγμα εκπαίδευσης ήταν η τιμή 0,6619, ενώ για του ελέγχου η τιμή 0,6871 και έτσι προέκυψαν οι ετικέτες *positive\_Valence* για όσα φέρουν μεγαλύτερη ίση τιμή από τον εκάστοτε μέσο και *negative\_Valence* για μικρότερη τιμή από αυτόν. Αφού έγινε ετικετοποίηση, προστέθηκαν οι κλάσεις στο αρχείο που δημιουργήθηκε από την *openSMILE*. Έτσι, προέκυψαν δυο νέα αρχεία, όπου το καθένα από αυτά έχει ως κλάση το *positive/negative Arousal* ή *Valence* αντιστοίχως. Έτσι, τα αρχεία αυτά είναι έτοιμα για εισαγωγή στο τελευταίο εργαλείο εξόρυξης δεδομένων και εκμάθησης, την WEKA. Ο τρόπος που έγινε ετικετοποίηση καθώς και οι κλάσεις με μερικά χαρακτηριστικά στο αρχείο που αφορά τη μετρική Arousal παρουσιάζονται στο Παράρτημα II από τις Εικόνες 17-18.

### 3.4.2 Δημιουργία ετικετών για την δεύτερη υπόθεση

Στην δεύτερη υπόθεση οι ετικέτες δημιουργήθηκαν με διαφορετικό τρόπο. Πιο συγκεκριμένα, για τα πρώτα 89 instances ελήφθη από κάθε χρήστη, δηλαδή κάθε στήλη του πίνακα, η μέση τιμή και αυτό έγινε εννέα φορές, όσοι και οι χρήστες. Στη συνέχεια, ο κάθε μέσος χρησιμοποιήθηκε ως τρόπος διαχωρισμού για τις ετικέτες για κάθε στήλη. Για παράδειγμα, για τον πίνακα των Arousal ο πρώτος χρήστης είχε μέση τιμή στα πρώτα 87 instances, το 0.8252, και έτσι όσα είναι μικρότερα πήραν ετικέτα *negative\_Arousal* και για μεγαλύτερη ή ίση *positive\_Arousal*. Η ίδια τακτική ακολουθήθηκε και για τους άλλους οκτώ χρήστες για τα πρώτα 87 instances (το απαιτούμενο πλήθος για την εκμάθηση). Έπειτα, με τον ίδιο τρόπο προέκυψε για κάθε χρήστη ο δεύτερος

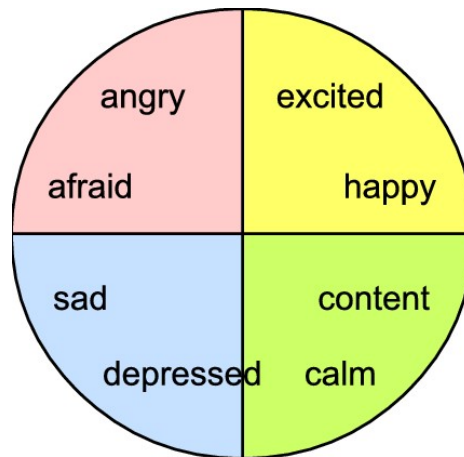
μέσος όρος του για τα υπόλοιπα 58 instances, το απαιτούμενο δηλαδή πλήθος για το τεστ ελέγχου. Ο μέσος όρος του κάθε ένα χρήστη έγινε και το κριτήριο για την ετικέτα. Τέλος, για να προκύψει η δέκατη στήλη με τις τελικές και πλέον ζητούμενες ετικέτες, εφαρμόστηκε η μέθοδος της πλειοψηφίας του πρόσημου. Αν δηλαδή οι χρήστες σε μια γραμμή είχαν περισσότερες από πέντε (5) θετικές ετικέτες και οι υπόλοιποι από τέσσερις ή λιγότερες αρνητικές, τότε η τελική ετικέτα προέκυπε θετική των Arousal-Valence. Αυτές οι ετικέτες προστέθηκαν στο αρχείο της openSMILE. Ένα δείγμα αντιπροσωπευτικό για αυτή την υπόθεση μετά από τον έλεγχο μέσω των και ετικετοποίησης φαίνεται στην Εικόνα 19 στο Παράρτημα II της εργασίας.

### 3.5 Τρίτη υπόθεση για δημιουργία νέων κλάσεων

Κατά την ενασχόληση των μετρικών Arousal και Valence, δημιουργήθηκε η ανάγκη για βαθύτερη κατανόηση των συναισθημάτων των χρηστών. Η ανάγκη αυτή προέκυψε, καθώς ένας απλός αναγνώστης πιθανόν να μην είναι σε θέση να καταλάβει την έννοια της διέγερσης, είτε είναι θετική, είτε αρνητική για παράδειγμα. Δεδομένου πως όλα τα δεδομένα των χρηστών είναι κανονικοποιημένα, δηλαδή έχουν εύρος από 0 έως 1 και δεν ανήκουν στο σύνολο  $[-1,1]$ , το οποίο θα βοηθούσε στην αντιστοίχιση των συναισθημάτων σύμφωνα με την πρόταση του Russell. Έτσι, υιοθετήθηκε η υπόθεση: *Όσοι θα έχουν ετικέτα positive\_Arousal και positive\_Valence για την δεδομένη στιγμή, θα νιώθουν θετικά συναισθήματα αντίστοιχα με αυτά του πρώτου τεταρτημόριου. Όσοι έχουν αρνητικές ετικέτες και στις δυο μετρικές, θα ανήκουν στο τρίτο τεταρτημόριο συναισθημάτων. Ενώ για θετική διέγερση και αρνητική απόκριση (valence), θα νιώθουν αρνητικά του τρίτου τεταρτημόριου και αντίστροφα τα συναισθήματα του τέταρτου τεταρτημόριου.* Ουσιαστικά, μια υπόθεση που προσπαθεί να προσεγγίσει την απόδοση των συναισθημάτων κατά Russell. Με άλλα λόγια, δίνεται ο τρόπος που θα δημιουργηθούν οι νέες κλάσεις συναρτήσεων των δυο παραπάνω υποθέσεων για καθεμία μεμονωμένα (Πίνακας 2). Σημειώνεται πως τα χρώματα του πίνακα συμβαδίζουν με αυτά της Εικόνας 6 με τα αντίστοιχα συναισθήματα. Τελικά, δημιουργήθηκαν δυο νέα αρχεία με τέσσερις κλάσεις των παρακάτω συναισθημάτων και είχαν την μορφή της Εικόνας 21 του Παραρτήματος III, τα οποία και αυτά με τη σειρά τους ήταν χωρισμένα σε αναλογία 60/40, προκειμένου να συμφωνούν με τις δυο παραπάνω υποθέσεις αλλά και να ακολουθήσουν την ίδια τεχνική κατηγοριοποίησης στο περιβάλλον της WEKA.

Πίνακας 2: Συνδυασμός Arousal-Valence για αντιστοίχιση συναισθημάτων

Positive Arousal	Positive Valence	Emotions1
Positive Arousal	Negative Valence	Emotions2
Negative Arousal	Positive Valence	Emotions4
Negative Arousal	Negative Valence	Emotions3



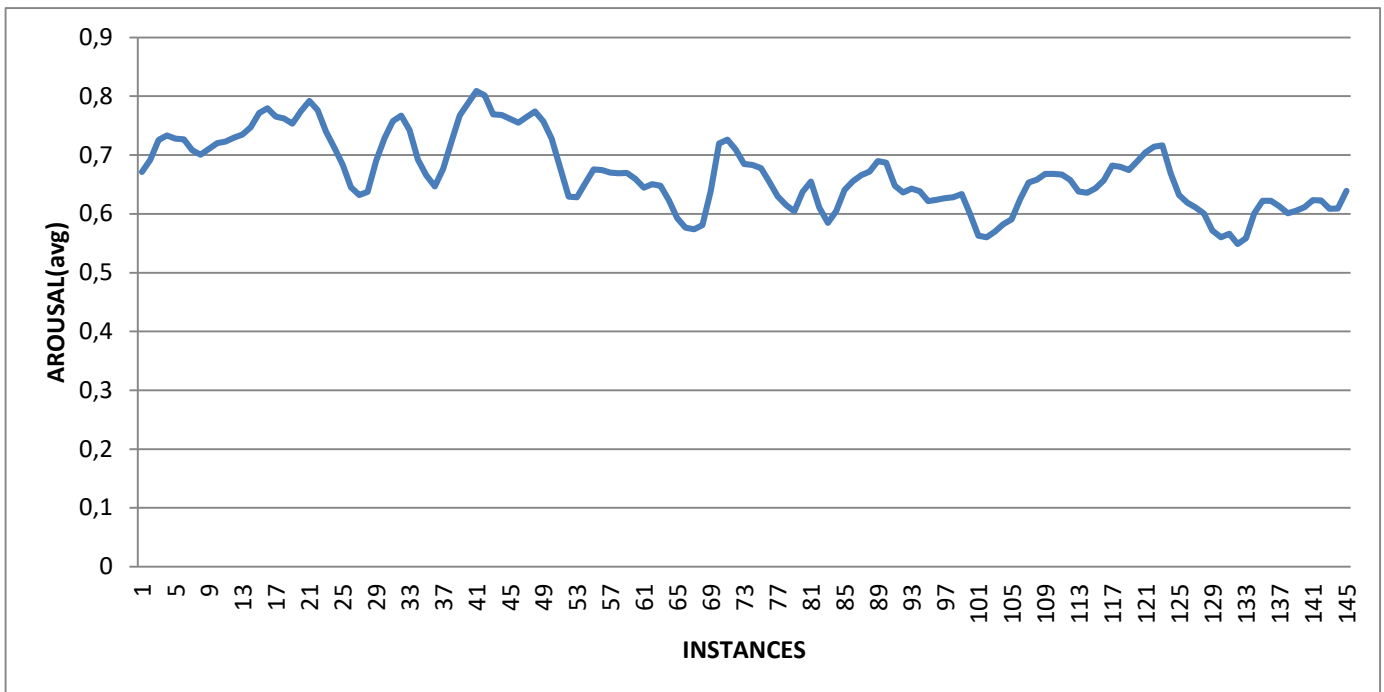
Εικόνα 6: Αντιστοίχιση συναισθημάτων με βάση την τρίτη υπόθεση

### ***3.6 Σύγκριση επισημειώσεων του ακουσματικού κομματιού με τα αντίστοιχα σήματα των μετρικών Arousal-Valence***

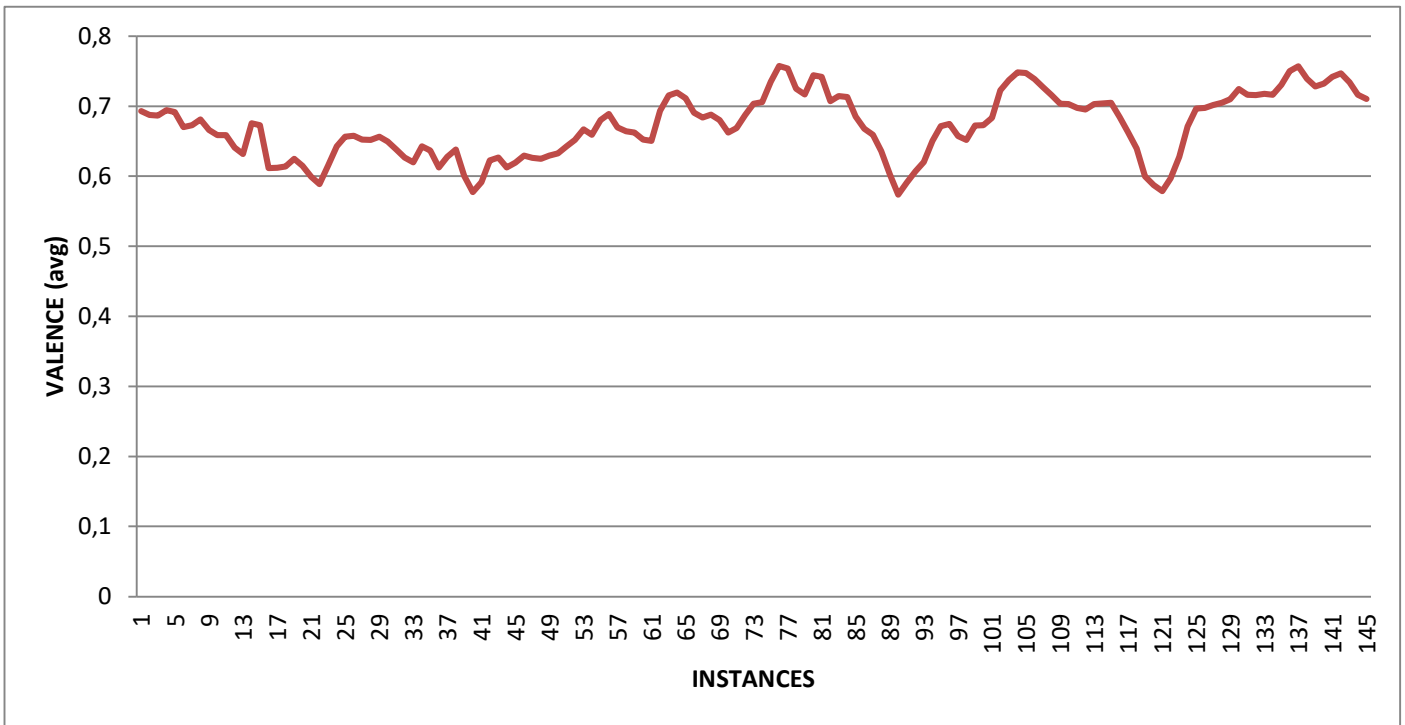
Το Γράφημα 1 παρουσιάζει όλα τα χαρακτηριστικά που εμφανίζει το ακουσματικό κομμάτι Ballichistory που δόθηκε για ακρόαση στους εννέα χρήστε εντός των 7:14:16 λεπτών που έχει διάρκεια. Στον Πίνακα 3 δίνεται η αντιστοίχιση των χαρακτηριστικών ως προς την ακουστική του κομματιού με την αντίστοιχη αρίθμηση σε κλάσεις, που δημιουργήθηκαν στα πλαίσια της παρούσας εργασίας, για λόγους ευκολίας κατά την αναπαράστασή τους σε γράφημα. Δίνονται επίσης, παρακάτω τα σήματα των μετρικών για τις αντίστοιχες μετρικές των Arousal και Valence. Αναλυτικότερα, τα Γραφήματα 2 και 3 αφορούν τις μετρικές αυτές σύμφωνα με την πρώτη υπόθεση, ενώ τα Γραφήματα 4 έως 9 αφορούν τις μετρικές αυτές σύμφωνα με την δεύτερη υπόθεση και επίσης χωρίστηκαν με τρόπο, ώστε κάθε γράφημα να περιέχει τρεις χρήστες για λόγους ευκολίας ανάγνωσης και παρατήρησης. Αξίζει να τονιστεί πως το Γράφημα 1 έχει οριζόντιο άξονα την διάρκεια του τραγουδιού, ενώ όλα τα υπόλοιπα Γραφήματα το πλήθος των instances (145), τα οποία αν πολλαπλασιαστούν με το βήμα της επικάλυψης που είναι 3 sec και διαιρεθούν με το ένα λεπτό (60 sec) δίνουν κατά προσέγγιση την διάρκεια του τραγουδιού. Έτσι έχει νόημα ο παραλληλισμός των γραφημάτων και η σύγκριση τους.



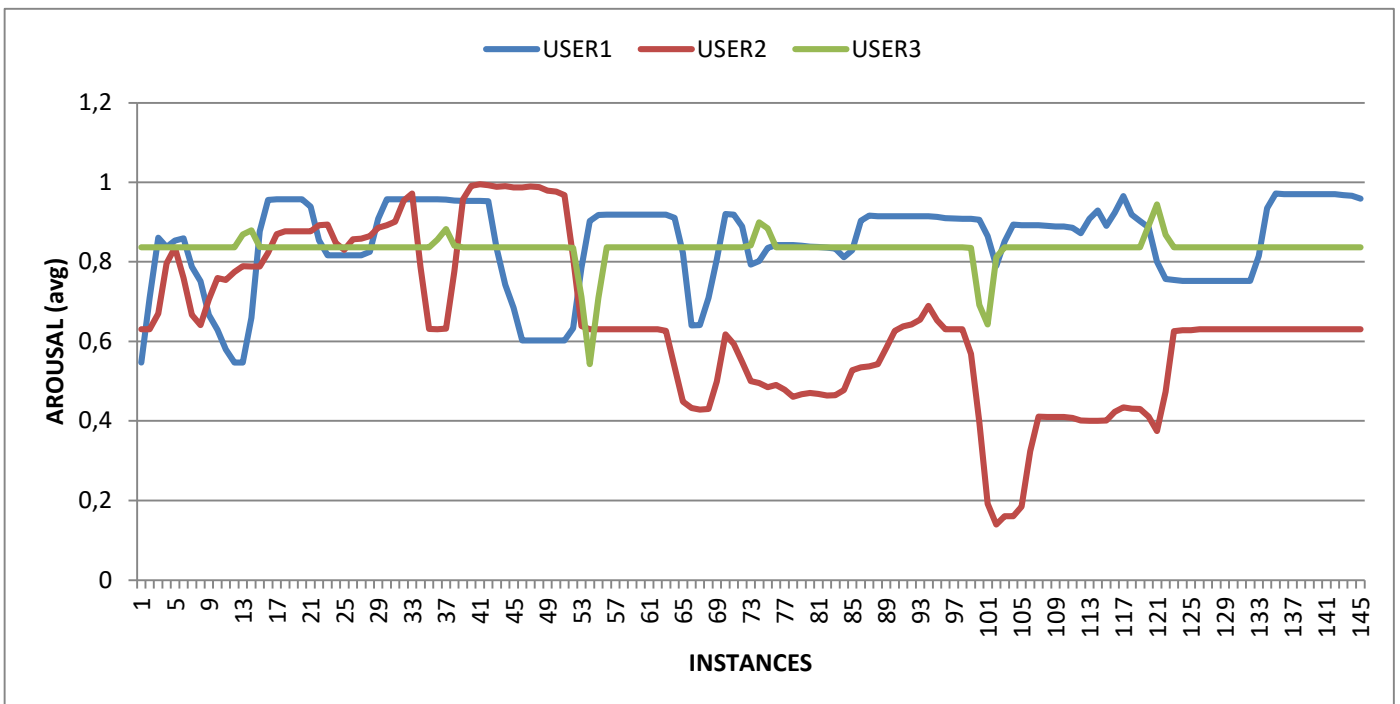
Γράφημα 2: Τα χαρακτηριστικά του ακουσματικού κομματιού που σημειώνονται κατά τη διάρκεια του



Γράφημα 1: Το σήμα του Arousal της πρώτης υπόθεσης από όλους τους χρήστες

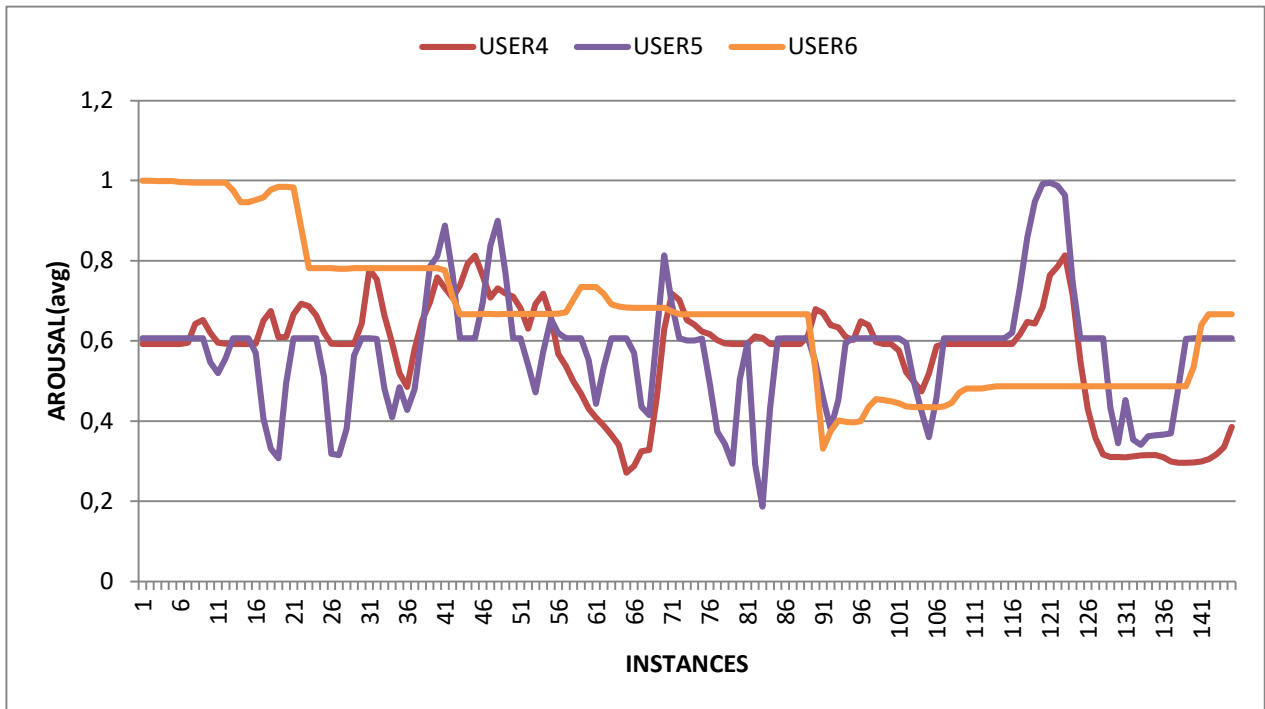


Γράφημα 3: Το σήμα του Valence για όλους τους χρήστες, σύμφωνα με την πρώτη υπόθεση.

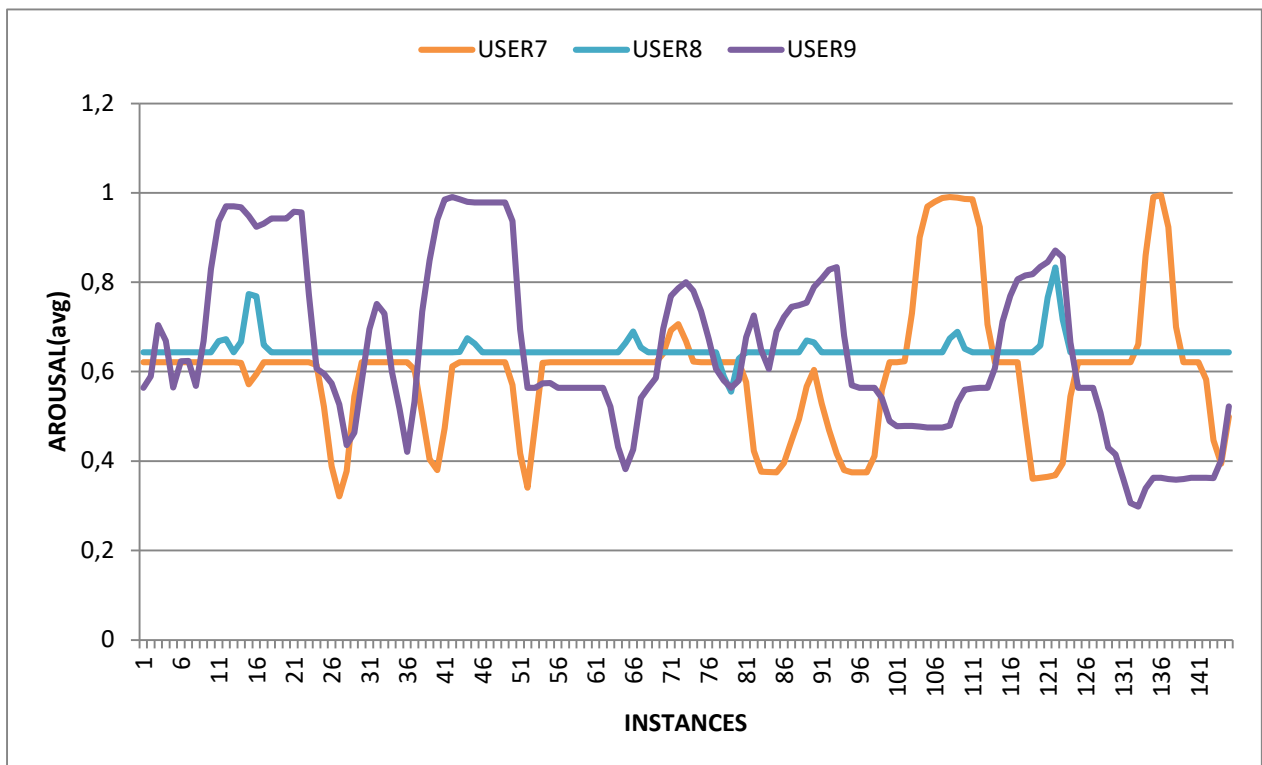


Γράφημα 4: Τα σήματα των Arousal για τους πρώτους τρεις χρήστες

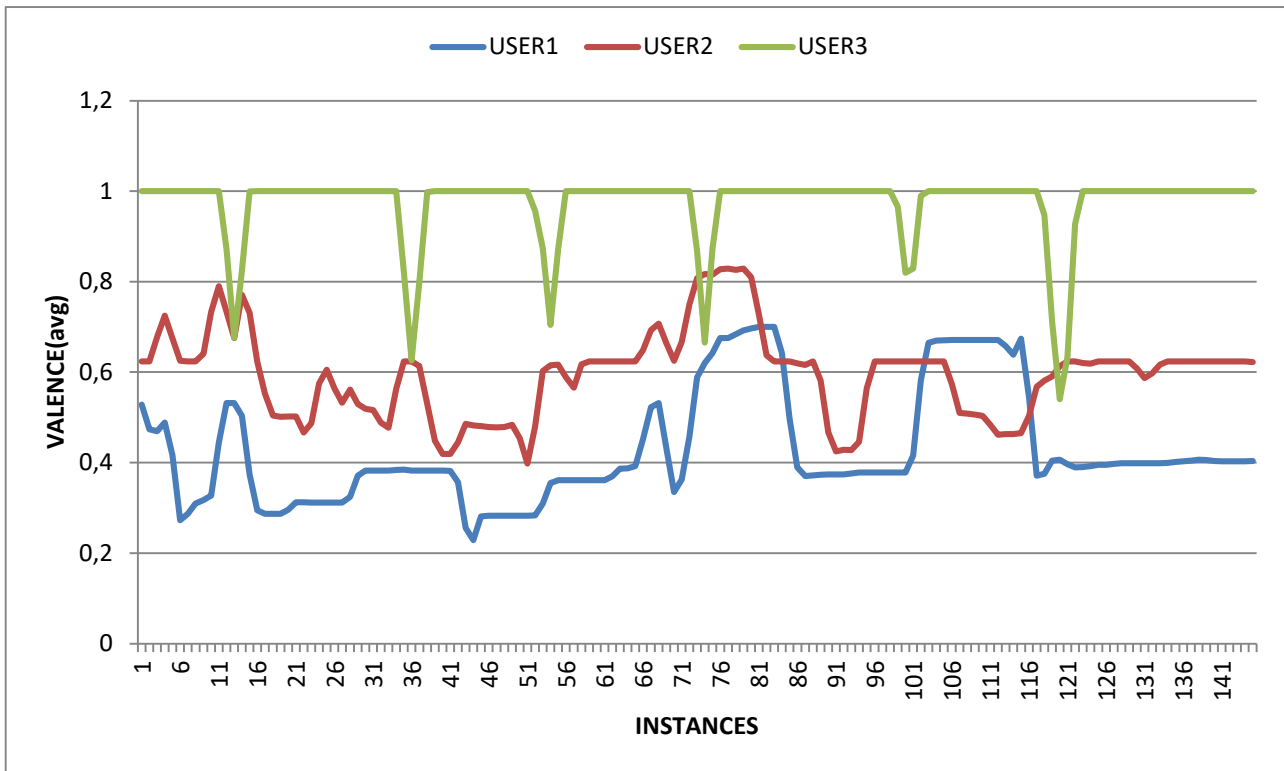




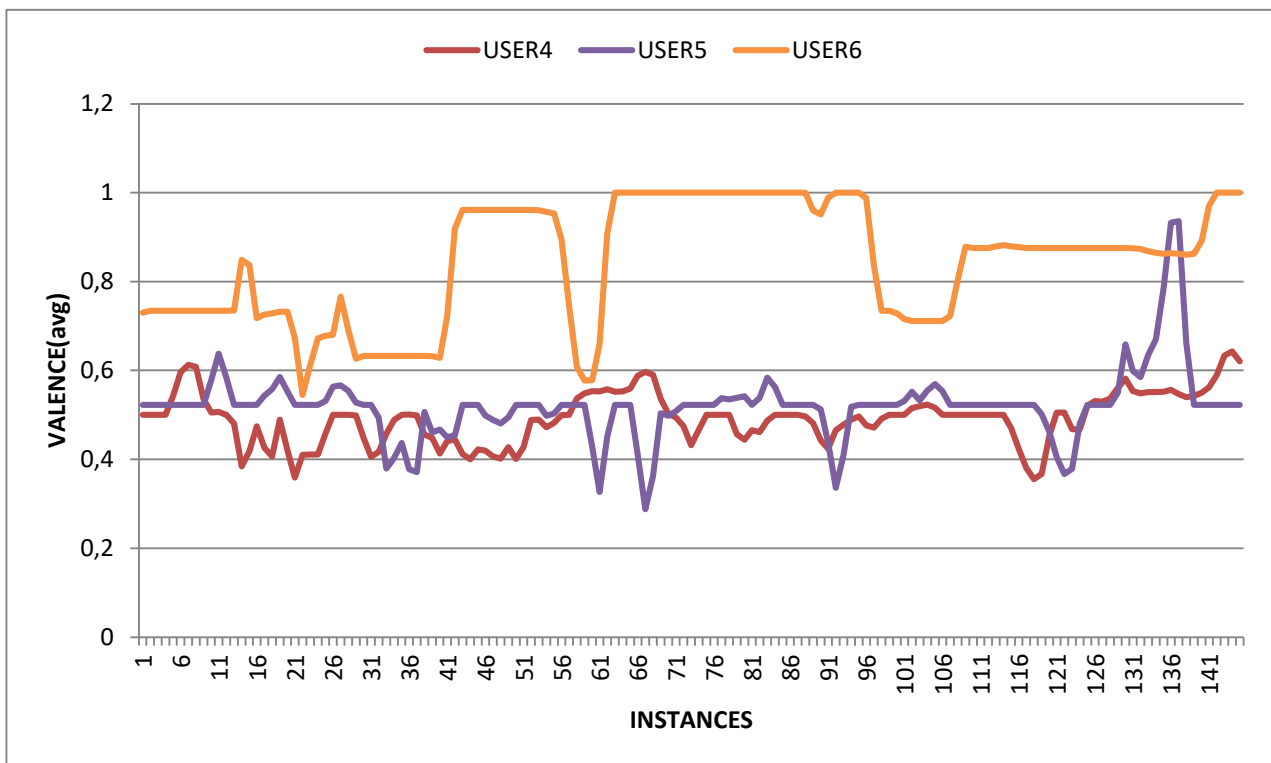
Γράφημα 6: Τα σήματα των Arousal για τους τρεις επόμενους χρήστες



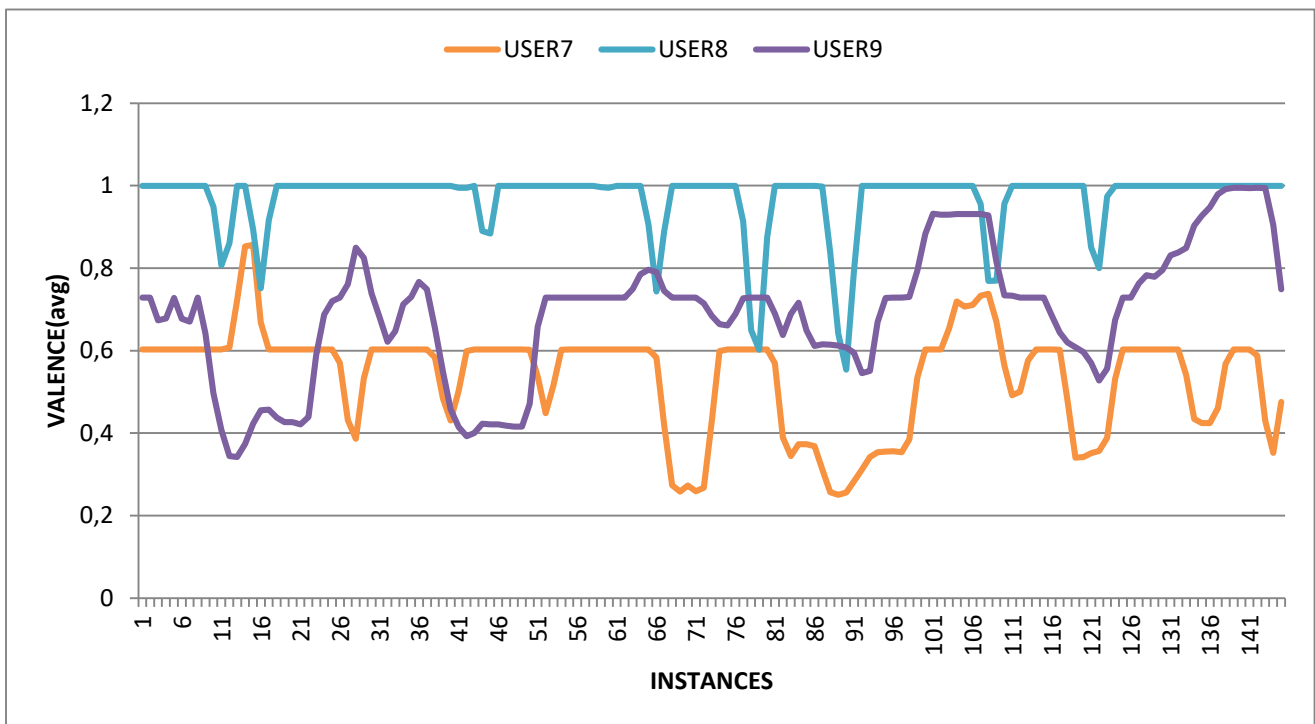
Γράφημα 5: Τα σήματα των Arousal για τους τελευταίους τρεις χρήστες



Γράφημα 7: Τα σήματα των Valence για τους τρεις πρώτους χρήστες



Γράφημα 8: Τα σήματα των Valence για τους επόμενους τρεις χρήστες



Γράφημα 9: Τα σήματα Valence για τους τελευταίους τρεις χρήστες

Παρατηρείται σχετική ανομοιογένεια του εύρους των μετρικών τόσο για κάθε χρήστη, όσο και της πρώτης υπόθεσης συναρτήσεως της δεύτερης. Μάλιστα, για την πρώτη υπόθεση οι μετρικές Arousal και Valence έχουν εύρος τιμών από 0.5 ως 0.8, ενώ για κάθε χρήστη υπάρχει περισσότερη αυξομειώση των τιμών. Η μετρική Arousal υποδηλώνει ανάλογα με τις τιμές της την αντανάκλαση του βαθμού διέγερσης, ενώ η μετρική Valence το πόσο θετικά ή αρνητικά αισθάνεται ο χρήστης την εκάστοτε στιγμή. Με βάση όλα αυτά μπορεί κανείς να προσέξει πως περίπου πριν την έναρξη του 5<sup>ου</sup> λεπτού και την λήξη του (00:05:00-00:06:00) για τον οριζόντιο άξονα του Γραφήματος 1 και τους οριζόντιους άξονες των άλλων γραφημάτων μεταξύ 90-105 instances [ $\sim(100 \cdot 3)/60=5$ ], τόσο οι μετρικές στην πρώτη υπόθεση, όσο και στην δεύτερη για όλους χρήστες, έχουν μια αισθητή πτώση και μετά την ολοκλήρωσης του 5<sup>ου</sup> λεπτού οι περισσότερες ανεβαίνουν εξίσου αισθητά. Για εκείνα τα δευτερόλεπτα εφαρμόζεται το χαρακτηριστικό rhythmic passage, από το οποίο μπορεί κανείς να υποθέσει πως σαν χαρακτηριστικό προκαλεί αρνητικά συναισθήματα, αντίστοιχα με αυτά του 3<sup>ου</sup> τεταρτημόριου, στους ακροατές. Αντίστοιχα, μια αισθητή πτώση των μετρικών και στις δυο υποθέσεις στα ίδια δευτερόλεπτα εμφανίζεται μεταξύ των 67-73 instances ( $\sim 00:03:50$ ), όπου και εφαρμόζεται το χαρακτηριστικό drone-βουητό. Σε αυτά τα σημεία λοιπόν, μπορεί κανείς να υποθέσει πως τα χαρακτηριστικά αυτά προκάλεσαν χαμηλότερη διέγερση στους χρήστες και αρνητικά συναισθήματα, δεδομένου ότι υπήρξαν στις στιγμές αυτές χαμηλές τιμές Arousal και Valence.

Πίνακας 3: Αντιστοίχιση κλάσεων με τα χαρακτηριστικά του ακουσματικού κομματιού

ΚΛΑΣΗ	ΧΑΡΑΚΤΗΡΙΣΤΙΚΟ
0	Structural – silence/ Ησυχία
1	Gestural/ Κίνηση
2	Attack/ Επίθεση
3	Climax/ Κορύφωση
4	End of phrase followed by silence/ Τέλος έκφρασης ακολουθούμενη από ησυχία
5	Debris/ Θραύσματα
6	Texture/ Υφή
7	Climax-attack/ Κορύφωση -επίθεση
8	Attack-gesture/ Επίθεση- κίνηση
9	Drone-texture/ Βουητό-υφή
10	Phrase/ έκφραση
11	Gesture-texture/ Κίνηση- Υφή
12	Tail/ Ουρά
13	Attack-texture/ Επίθεση- Υφή
14	Drone-gesture/ Βουητό-Κίνηση
15	Attack-debris-drone/ Επίθεση- Θραύσματα-Βουητό
16	Attack-debris/ Επίθεση- Θραύσματα
17	Drone/ Βουητό
18	Climax-texture/ Κορύφωση-Υφή
19	Stasis/ Στάση
20	Rhythmic passage/ Ρυθμός

# 4

## Πειράματα εξόρυξης δεδομένων

### 4.1 Προετοιμασία πειραμάτων

Το περιβάλλον της WEKA (Waikato Environment for Knowledge Analysis), αναπτύχθηκε στην Νέα Ζηλανδία το 1992 και είναι ένα από τα πιο κοινά δωρεάν λογισμικά για την εκτέλεση εργασιών μηχανικής μάθησης και εξόρυξης γνώσης και αυτό γιατί παρέχει την δυνατότητα στον χρήστη να εφαρμόσει διάφορους αλγορίθμους εκμάθησης στο δικό του σύνολο δεδομένων. Ένα εργαλείο ανοιχτού κώδικα βασισμένο στη γλώσσα JAVA (Κύρκου Ε. 2015, Remco 2013). Περιλαμβάνει επίσης, μια ποικιλία εργαλείων για την υλοποίηση των αλγορίθμων που περιλαμβάνονται στο πακέτο, είτε εφαρμόζοντάς τα απευθείας σε ένα σύνολο δεδομένων μέσα στην ίδια την εφαρμογή, είτε χρησιμοποιώντας τον κώδικα JAVA κάποιας τρίτης εφαρμογής (Eibe Frank 2016, Μίγκου 2007). Ένα άλλο πλεονέκτημα που έχει, είναι ότι μπορεί ο χρήστης να προεπεξεργαστεί το δείγμα του, να το ενημερώσει με ένα σχήμα εκμάθησης, να αναλύσει το αποτέλεσμα ταξινόμησης και της απόδοσής του, χωρίς να γράψει απαραίτητα κάποιο κώδικα, παρά μόνο με τη χρήση του περιβάλλοντός του και των διεπαφών του, όπως Explorer, Experimenter, Knowledge flow, Workbench και SimpleCli (Remco 2018, Μίγκου 2007). Ο χώρος εργασίας του περιλαμβάνει μεθόδους για κύρια προβλήματα εξόρυξης δεδομένων όπως αυτά της παλινδρόμησης, κατηγοριοποίησης, συσταδοποίησης, κανόνες συσχέτισης, οπτικοποίησης και επιλογής χαρακτηριστικών (Κύρκου Ε. 2015, Μίγκου 2007).

Η επιλογή για τα πειράματα της παρούσας εργασίας αφορά τα μοντέλα ταξινόμησης και αξιολόγησης. Ειδικότερα, επιλέχθηκε η διεπαφή Explorer και στη συνέχεια για το εκάστοτε πείραμα η επιλογή Classify. Κατά την εισαγωγή των δεδομένων στην WEKA 3.8.5<sup>4</sup>, επιλέχθηκε η

---

<sup>4</sup> [https://waikato.github.io/weka-wiki/downloading\\_weka/](https://waikato.github.io/weka-wiki/downloading_weka/)

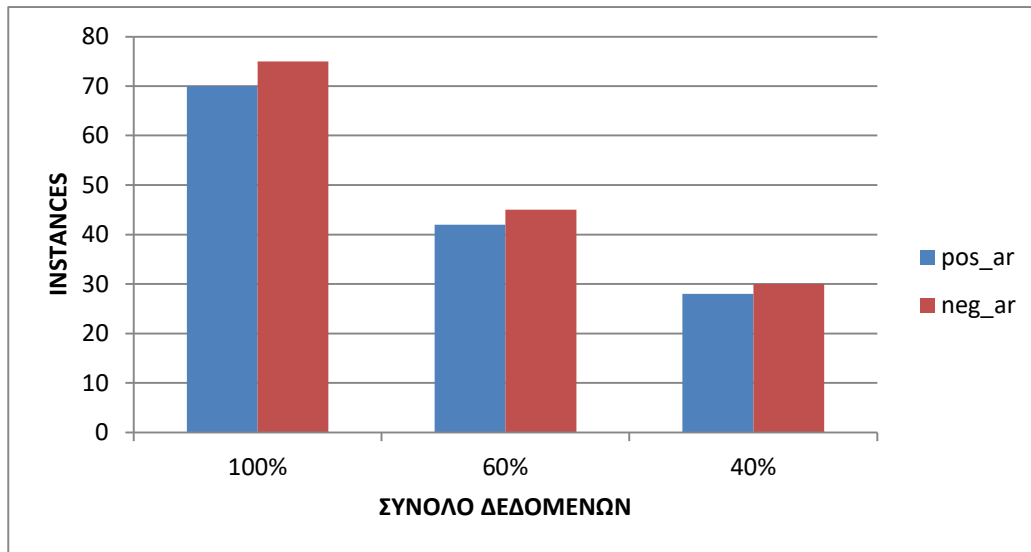
διαδικασία του διαχωρισμού των δεδομένων σε αναλογία 60/40, μια συνήθης επιλογή που εφαρμόζεται. Έτσι, ξεκινά η εκπαίδευση του μοντέλου με το σύνολο εκπαίδευσης (60% του συνόλου-87 instances) και έπειτα και αφού επιλεγθούν οι κατάλληλες παράμετροι, θα αξιολογηθεί με το σύνολο ελέγχου (40% των δεδομένων-58 instances). Για την υλοποίηση αυτών, ορίστηκε κατά την εκπαίδευση των δεδομένων το cross-validation να είναι ίσο με 5 folds και αυτό γιατί το δείγμα μας δεν είναι ιδιαίτερα μεγάλο. Ενώ για το τεστ, καλείται η supplied test set με τα κατάλληλα ορίσματα για τα διαφορετικά σύνολα δεδομένων που χρησιμοποιήθηκαν.

- **5folds cross-validation ή Διασταυρούμενη Επικύρωση 5 τμημάτων** είναι μια μέθοδος επικύρωσης 5 τμημάτων, όπου το σύνολο δεδομένων διαιρείται σε 5 υποσύνολα. Κάθε υποσύνολο περιέχει διαφορετικές παρατηρήσεις και η επιλογή γίνεται τυχαία. Ένα από τα υποσύνολα χρησιμοποιείται ως σύνολο επικύρωσης και τα υπόλοιπα τέσσερα συνενώνονται και δημιουργούν το σύνολο εκπαίδευσης (Κύρκου Ε. 2015). Το μοντέλο εκπαιδεύεται χρησιμοποιώντας το σύνολο εκπαίδευσης και δοκιμάζεται έναντι του συνόλου επικύρωσης. Η διαδικασία αυτή επαναλαμβάνεται 5 συνολικά φορές, κάθε φορά χρησιμοποιώντας ένα διαφορετικό σύνολο ως σύνολο επικύρωσης. Στο τέλος υπολογίζεται η μέση επίδοση του μοντέλου. Αυτή η διαδικασία θα χρησιμοποιηθεί για όλα τα σύνολα εκπαίδευσης που δημιουργήθηκαν σε κάθε υπόθεση, συνολικά έξι, και από όπου ανάλογα με τα αποτελέσματα που θα προκύπτουν κάθε φορά θα επιλεγθούν τα καλύτερα με βάση τις παραμέτρους για κάθε αλγόριθμο. Περισσότερες πληροφορίες θα δοθούν στην ενότητα 4.3-4.4.
- **Supplied test set** είναι μια επιλογή που παίρνει σαν είσοδο πέρα από το αρχείο της εκμάθησης και ένα άλλο αρχείο εξωτερικό, το οποίο θα χρησιμοποιηθεί για την επικύρωση, test-έλεγχο. Στα πειράματα της εργασίας αυτής και κατά την επιλογή αυτή εισήχθησαν για κάθε πείραμα τα αρχεία με το 40%, όπως κατασκευάστηκαν και παρουσιάστηκαν παραπάνω.

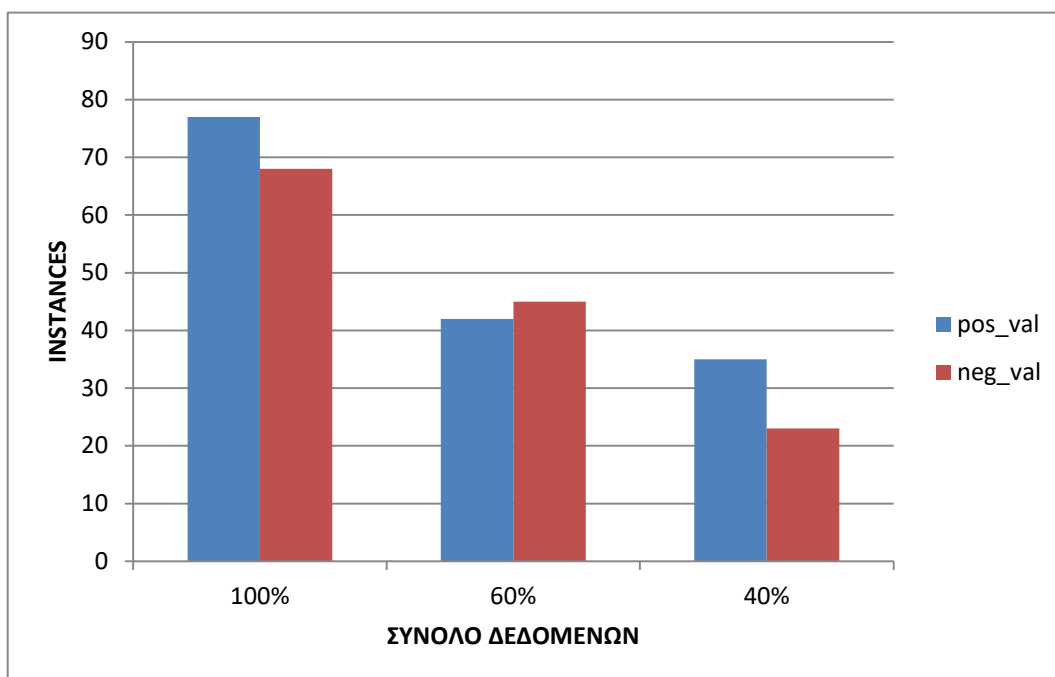
## 4.2 Κλάσεις ανά αρχείο

Παρακάτω απεικονίζονται τα αρχεία με τις κλάσεις τους σύμφωνα με την υπόθεση που ακολούθησαν, ώστε να προκύψουν οι ετικέτες ανά κλάση. Μάλιστα, για καλύτερη κατανόηση των δεδομένων αποδίδονται οι κλάσεις τόσο για όλο το πλήθος δεδομένων (100%), για το σύνολο εκπαίδευσης (60%) και τέλος για τον έλεγχο (40%).

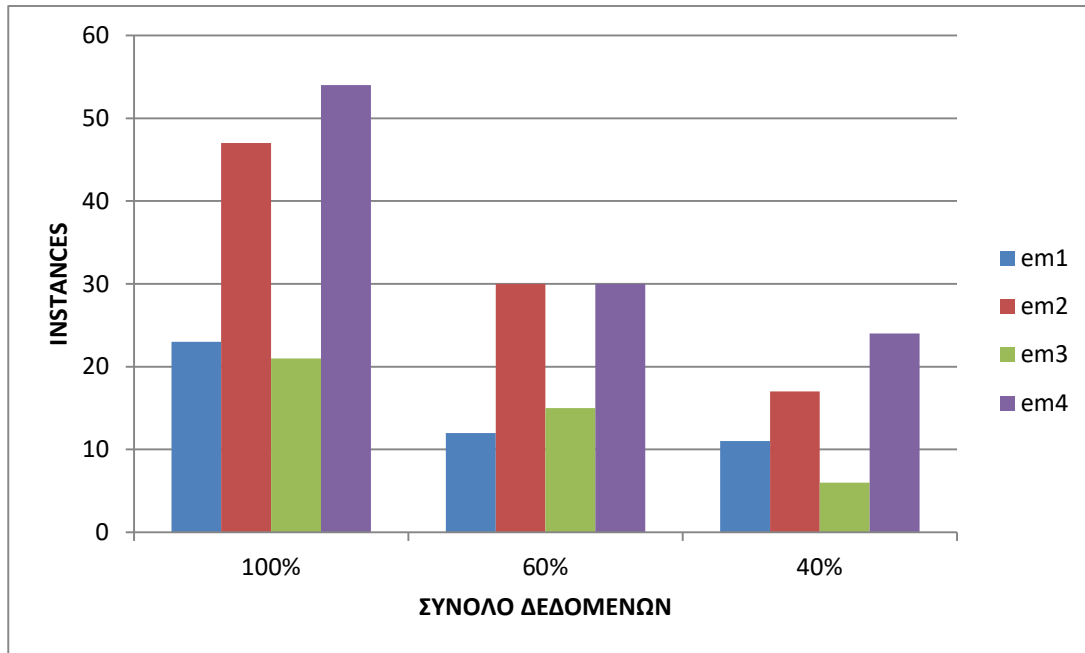
### 4.2.1 Πρώτη υπόθεση



Γράφημα 10: Οι κλάσεις για την μετρική Arousal ανά πλήθος δεδομένων

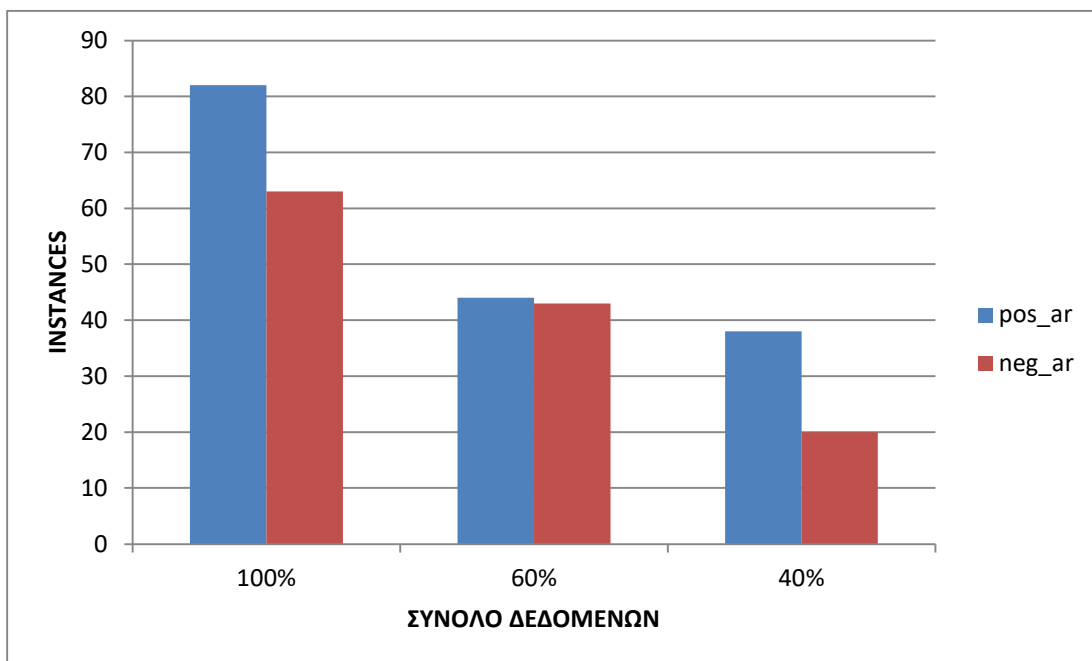


Γράφημα 11: Οι κλάσεις για την μετρική Valence ανά πλήθος δεδομένων



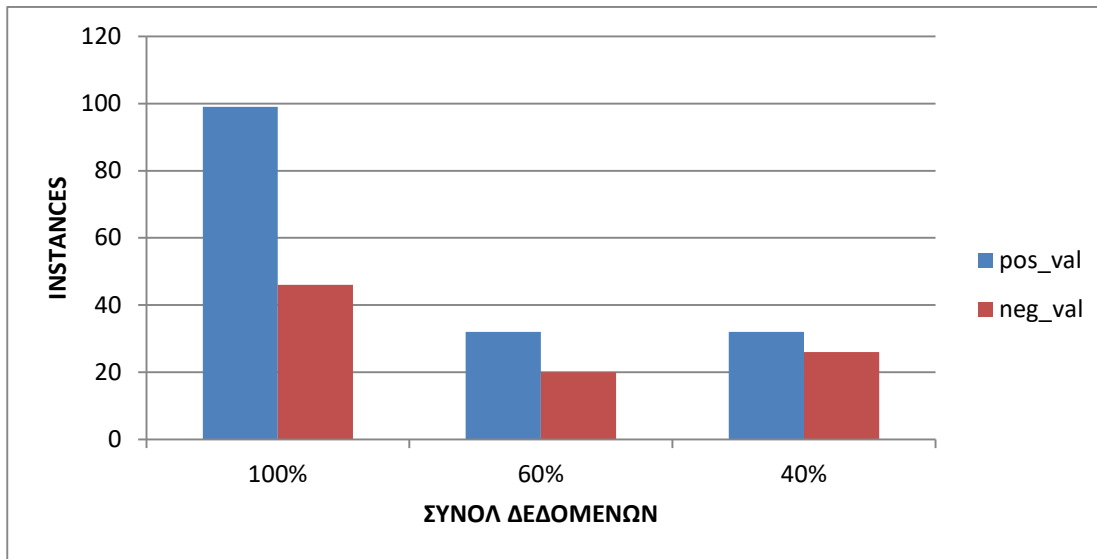
Γράφημα 12: Οι κλάσεις για τα συναισθήματα ανά πλήθος δεδομένων

#### 4.2.2 Δεύτερη υπόθεση

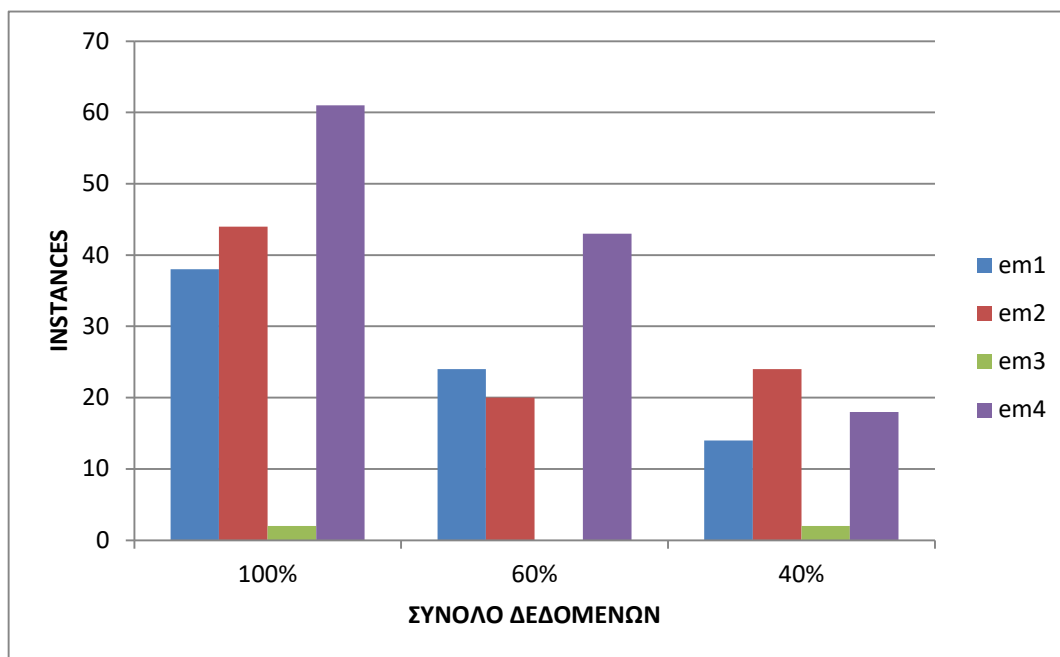


Γράφημα 13: Οι κλάσεις για την μετρική Arousal ανά πλήθος δεδομένων





Γράφημα 14: Οι κλάσεις για την μετρική Valence ανά πλήθος δεδομένων



Γράφημα 15: Οι κλάσεις για τα συναισθήματα ανά πλήθος δεδομένων

Παρατηρούμε πως το σύνολο δεδομένων με κλάσεις τα συναισθήματα, ακολουθούμενο την δεύτερη υπόθεση, έχει μικρό αριθμό ετικετών για την τρίτη κλάση, που αντιστοιχεί στα συναισθήματα του τρίτου τεταρτημορίου στον πίνακα κατά Russell. Πληροφορία που ίσως επηρεάσει τα αποτελέσματα κατά τον έλεγχο στο περιβάλλον της WEKA, δημιουργώντας την ανάγκη για περισσότερη έρευνα και περισσότερο πλήθος συμμετεχόντων για πιο αξιόπιστα δεδομένα.

### 4.3 Μετρικές κατηγοριοποίησης

Παρακάτω δίνονται όλες οι μετρικές που μπορούν να χρησιμοποιηθούν κατά την διαδικασία αξιολόγησης κατηγοριοποίησης δεδομένων σε περιβάλλοντα εκμάθησης. Μπορούν οι μετρικές να χωριστούν σε αυτές που δίνουν μέτρα ανά κλάση ή σε αυτές με βάση τον μέσο όρο επιδόσεων.

#### 4.3.1 Επιδόσεις ανά κλάση

Προτού δοθεί η περιγραφή της διαδικασίας των πειραμάτων και τα αποτελέσματά τους, θα παρουσιαστούν τα μέτρα απόδοσης που εμφανίζονται κατά την κατηγοριοποίηση των δεδομένων στο περιβάλλον Explorer της WEKA. Ειδικότερα, για να εκτιμηθούν οι ανά κλάση επιδόσεις ενός κατηγοριοποιητή, συναντώνται οι εξής έννοιες (Κύρκος Ε. 2015):

- Θετικές παρατηρήσεις (positive) ονομάζονται οι παρατηρήσεις, οι οποίες ανήκουν σε μια τιμή της κλάσης.
- Αρνητικές παρατηρήσεις (negative) ονομάζονται οι παρατηρήσεις, οι οποίες ανήκουν στην άλλη τιμή της κλάσης.
- Αληθινές Θετικές Προβλέψεις (true positive – TP) είναι το πλήθος των επιτυχημένων προβλέψεων για θετικές παρατηρήσεις.
- Αληθινές Αρνητικές Προβλέψεις (true negative – TN) είναι το πλήθος των επιτυχημένων προβλέψεων για αρνητικές παρατηρήσεις.
- Ψευδείς Θετικές Προβλέψεις (false positive – FP) είναι το πλήθος των αποτυχημένων προβλέψεων για αρνητικές παρατηρήσεις.
- Ψευδείς Αρνητικές Προβλέψεις (false negative – FN) είναι το πλήθος των αποτυχημένων προβλέψεων για θετικές παρατηρήσεις.

Ένας τρόπος παρουσίασης των επιδόσεων ανά κλάση ενός κατηγοριοποιητή είναι με τη χρήση του Πίνακα Σύγχυσης (confusion matrix). Ο Πίνακας Σύγχυσης είναι ένας διδιάστατος πίνακας, όπου οι στήλες αντιστοιχούν στις προβλέψεις και οι γραμμές στις πραγματικές τιμές κλάσης. Στα κελιά του πίνακα αναγράφονται οι αληθινές θετικές, οι αληθινές αρνητικές, οι ψευδείς θετικές και οι ψευδείς αρνητικές προβλέψεις (Εικόνα 7).

	Πρόβλεψη Αρνητικής Κλάσης	Πρόβλεψη Θετικής Κλάσης
Πραγματική Αρνητική Κλάση	tn	fp
Πραγματική Θετική Κλάση	fn	tp

tn = true negative  
tp = true positive  
fn = false negative  
fp = false positive

Εικόνα 7: Πίνακας σύγχυσης (Κύρκου Ε. 2015)

Επίσης, ορισμένα πρόσθετα μέτρα για τις επιδόσεις ενός κατηγοριοποιητή είναι τα ακόλουθα:

$$\begin{aligned} \text{Sensitivity} &= TP/POS & \text{Specificity} &= TN/NEG & \text{Precision} &= TP/(TP+FP) \\ \text{Accuracy} &= \text{sensitivity} * [POS/(POS+NEG)] + \text{specificity} * [NEG/(POS+NEG)] = (TP+TN)/(POS+NEG) \end{aligned}$$

Όπου POS είναι το πλήθος των θετικών παρατηρήσεων και NEG των αρνητικών. Σύμφωνα με τα παραπάνω, η ακρίβεια (accuracy) ορίζεται ως το ποσοστό των ορθών θετικών προβλέψεων επί το ποσοστό των θετικών παρατηρήσεων συν το ποσοστό των ορθών αρνητικών προβλέψεων επί το ποσοστό των αρνητικών παρατηρήσεων ή ισοδύναμα ως το πλήθος των ορθών προβλέψεων προς το πλήθος των παρατηρήσεων (Κύρκου Ε. 2015).

#### 4.3.2 Καμπύλες ROC

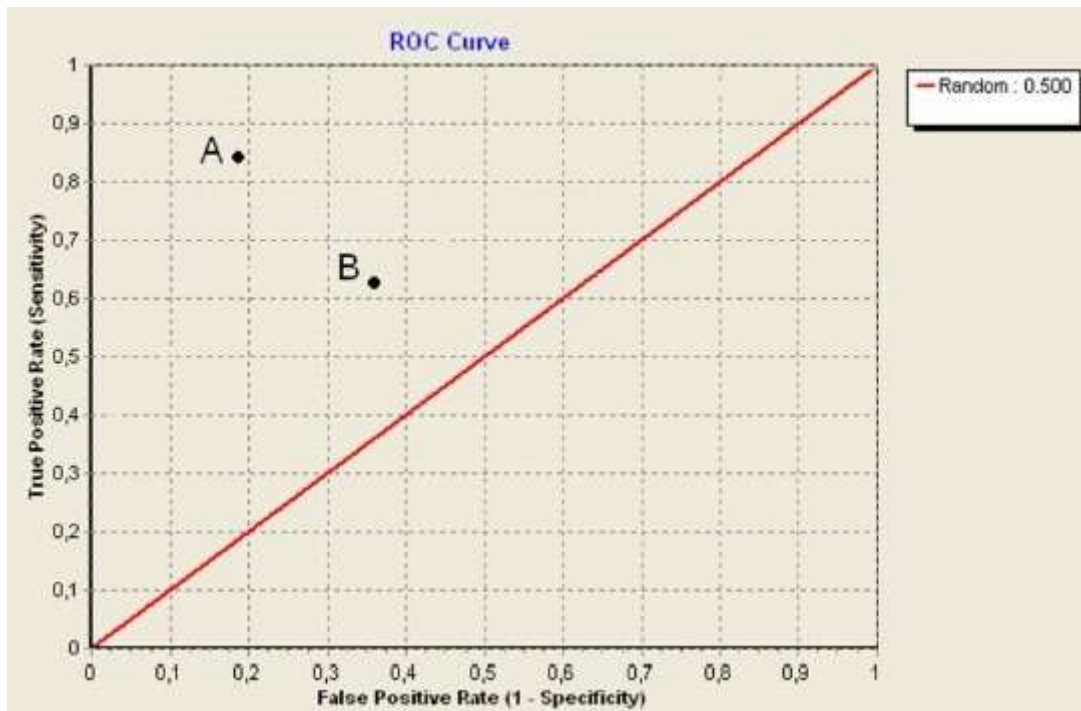
Ένα ισχυρό μέτρο για την εκτίμηση της ανά κλάσης ακρίβειας του κατηγοριοποιητή είναι οι λεγόμενες καμπύλες ROC (Receiver Operating Characteristics). Οι καμπύλες ROC σχεδιάζονται σε έναν δυδιάστατο επίπεδο χώρο. Ο οριζόντιος άξονας εκφράζει το μέγεθος 1-specificity, το οποίο ονομάζεται και False Positive Rate (Κύρκου Ε. 2015).

$$\text{False\_positive\_rate} = 1 - \text{specificity} = FP/NEG$$

Ο κατακόρυφος άξονας εκφράζει το μέγεθος sensitivity, το οποίο ονομάζεται και True Positive Rate

$$\text{True\_positive\_rate} = \text{Sensitivity} = TP/POS$$

Ουσιαστικά, ο οριζόντιος άξονας εκφράζει το ποσοστό των αρνητικών παρατηρήσεων, οι οποίες κατηγοριοποιήθηκαν λάθος, και ο κατακόρυφος άξονας εκφράζει το ποσοστό των θετικών παρατηρήσεων, οι οποίες κατηγοριοποιήθηκαν σωστά (Κύρκου Ε. 2015). Η Εικόνα 8 απεικονίζει τον δυδιάστατο χώρο καμπύλων ROC (Κύρκου Ε. 2015). Κάθε σημείο του χώρου αυτού εκφράζει ένα ισοζύγιο ανάμεσα στο ποσοστό ορθών θετικών προβλέψεων και εσφαλμένων θετικών προβλέψεων. Το σημείο (0,0) είναι ένας κατηγοριοποιητής, που δεν προβλέπει ποτέ θετική παρατήρηση (Κύρκου Ε. 2015). Το σημείο (1,1) είναι ένας κατηγοριοποιητής, που προβλέπει πάντα θετική παρατήρηση. Η διαγώνια γραμμή, από το σημείο (0,0) στο σημείο (1,1) είναι ένας κατηγοριοποιητής που προβλέπει τυχαία την κλάση (Κύρκου Ε. 2015). Οι κατηγοριοποιητές που βρίσκονται κάτω από τη διαγώνια γραμμή είναι χειρότεροι από την τυχαία πρόβλεψη. Οι κατηγοριοποιητές που βρίσκονται πάνω από τη διαγώνια γραμμή είναι καλύτεροι από την τυχαία πρόβλεψη (Κύρκου Ε. 2015). Το σημείο (0,1) είναι ο άριστος κατηγοριοποιητής, οι οποίος προβλέπει σωστά όλες τις θετικές και αρνητικές παρατηρήσεις. Γενικώς, όσο πιο μετατοπισμένο είναι προς τα επάνω και προς τα αριστερά ένα σημείο, τόσο καλύτερη θεωρείται η επίδοση. Στην Εικόνα 8 το σημείο Α είναι καλύτερο από το σημείο Β (Κύρκου Ε. 2015).



Εικόνα 8: Η καμπύλη ROC (Κύρκου Ε. 2015).

Η επίδοση των κατηγοριοποιητών στον χώρο ROC συμβολίζεται με μία καμπύλη. Για να συγκρίνουμε κατηγοριοποιητές χρειαζόμαστε ένα μέτρο σύγκρισης. Τέτοιο μέτρο σύγκρισης είναι η Περιοχή Κάτω από την Καμπύλη ROC (Area Under ROC Curve (AUC)) (Κύρκου Ε. 2015). Η AUC εκφράζει το ποσοστό του χώρου που βρίσκεται κάτω από την καμπύλη, και παίρνει τιμές από 0 έως 1 (Κύρκου Ε. 2015). Η διαγώνια γραμμή τυχαίας πρόβλεψης έχει  $AUC = 0,5$ . Συνεπώς, κάθε κατηγοριοποιητής καλύτερος της τυχαίας πρόβλεψης έχει  $AUC > 0,5$ . Όσο μεγαλύτερη περιοχή AUC έχει ένας κατηγοριοποιητής τόσο καλύτερος είναι (Κύρκου Ε. 2015).

Σημειώνεται πως για τους σκοπούς των αξιολογήσεων των πειραμάτων θα χρησιμοποιηθούν η μετρική ROC area και η Ακρίβεια, ενώ ταυτόχρονα θα παρατίθεται και ο Πίνακας Σύγχυσης-Confusion Matrix. Επίσης, υπενθυμίζεται πως η τρίτη υπόθεση της προσέγγισης των συναισθημάτων πλέον προσαρτάται και στις άλλες δυο πρώτες υποθέσεις.

#### 4.4 Δομή πειραμάτων

Η γενική γραμμή που θα ακολουθηθεί στα πειράματα είναι:

- A. Εισαγωγή του εκάστοτε αρχείου που περιέχει το 60% των δεδομένων για εκμάθηση.
- B. Πραγματοποίηση εκμάθησης και ταυτόχρονη αναζήτηση καλύτερης τιμής των παραμέτρων, δηλαδή:
  - i) Εφαρμογή cross-validation ίσο με 5 folds.
  - ii) Εφαρμογή meta-αλγορίθμων. Πιο συγκεκριμένα θα χρησιμοποιηθεί ο meta-αλγόριθμος CVParameterSelection, που κατηγοριοποιεί τα δεδομένα υπό τις κατάλληλες τιμές παραμέτρων των ταξινομητών που επιλέγηκαν μέσω cross-validation, μέθοδο δηλαδή διασταυρούμενης επικύρωσης.

- iii) Σε κάθε εφαρμογή του αλγορίθμου αυτού θα εφαρμόζεται και ένας από τους τρεις αλγορίθμους που περιγράφηκαν στο Κεφάλαιο 2: SVM, KNN, J48.
  - iv) Για κάθε επιλογή αλγορίθμου θα ορίζεται στην επιλογή CVParameters το εύρος των παραμέτρων που έχει ο κάθε αλγόριθμος, ώστε να βρεθεί η καλύτερη τιμή τους που θα φέρει και τα καλύτερα αποτελέσματα.
  - v) Έναρξη ελέγχου.
  - vi) Παρατήρηση αποτελεσμάτων και σημείωση κατάλληλων τιμών των παραμέτρων
- C. Επιλογή Supplied test set και εισαγωγή του κατάλληλου αρχείου του συνόλου δεδομένων ελέγχου (40% test set) .
- vii) Εφαρμογή των τριών αλγορίθμων κάθε φορά με τις παραμέτρους που βρέθηκαν στην παραπάνω διαδικασία.
  - viii) Πραγματοποίηση ελέγχου.
- D. Σύγκριση αποτελεσμάτων και αποδόσεων αλγορίθμων.

Στα πλαίσια του πειράματος της εργασίας με χρήση του SVM θα χρησιμοποιηθούν οι παράμετροι Κόστος [C] και οι Kernels: Polykernel και RBF συναρτήσεως της συνάρτησης γάμμα (g). Το εύρος των τιμών για το κόστος και την γάμμα έχει οριστεί στο Κεφάλαιο 2 και αυτό θα χρησιμοποιηθεί. Σημειώνεται επίσης, πως μόνο η παράμετρος K θα χρησιμοποιηθεί στα πλαίσια των πειραμάτων για τον KNN, καθώς και η απόσταση Euclidean ή Manhattan. Τέλος, όσον αφορά το Δέντρο απόφασης J48 παρά το πλήθος των παραμέτρων που έχει, το παρόν πείραμα εξετάζει τις δυο πρώτες παραμέτρους, όπως παρουσιάστηκαν στο Κεφάλαιο 2, δηλαδή τον παράγοντα εμπιστοσύνης [C] και τον ελάχιστο αριθμό αντικειμένων [M], με τα γνωστά εύρη τιμών.

## 4.5 Αποτελέσματα πρώτης προσέγγισης

### 4.5.1 SVM

#### ➤ AROUSAL

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του Arousal με τα μέτρα απόδοσης.

**Πίνακας 4: Τα μέτρα απόδοσης συναρτήσεως των καλύτερων τιμών των παραμέτρων του SVM**

60%	Επιλογή κατάλληλου C	Γάμμα παράμετρος (g)	ACCURACY	ROC Area (Weighted avg.)
POLYKERNEL	0.5	-----	68.97%	0.687
RBFKERNEL	1.5	(default-0.01)	70.12%	0.698
RBFKERNEL	1.5	0.015	74.71%	0.745

**ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ**

**POLYKERNEL**  
 a b ← ταξινόμηση ως  
 34 11 | a = neg\_ar  
 16 26 | b = pos\_ar

**RBFKERNEL**  
 a b ← ταξινόμηση ως  
 35 10 | a = neg\_ar  
 16 26 | b = pos\_ar

**RBF/G**  
 a b ← ταξινόμηση ως  
 36 9 | a = neg\_ar  
 13 29 | b = pos\_ar

Άρα, καλύτερα αποτελέσματα έφερε με την παράμετρο του RBFKernel με  $g = 0.015$  και  $C = 1.5$ , τον οποίο και επιλέγεται για το τεστ ελέγχου με τις παραπάνω παραμέτρους.

Για το **test**:

**Πίνακας 5: Τα μέτρα απόδοσης TOY Arousal κατά τον έλεγχο με τον SVM**

30%	ACCURACY	ROC Area (Weighted avg.)
RBFKERNEL	63.79%	0.632

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b ← ταξινόμηση ως  
 24 6 | a = neg\_ar  
 15 13 | b = pos\_ar

Παρατηρείται πως η τιμή της ROC Area είναι σχετικά ικανοποιητική δεδομένου ότι είναι και μεγαλύτερη από τα 0.5 που αντιστοιχεί στην τυχαία κατηγοριοποίηση. Συγχρόνως, ο SVM τροποποιημένος σύμφωνα με τα παραπάνω, καταφέρνει να κατηγοριοποιήσει σωστά 24 instances και άλλα 13, που αντιστοιχούν στις κλάσεις αρνητικής και θετικής διέγερσης αντιστοίχως. Ωστόσο, 6 instances τα κατατάσσει ως θετικά ενώ αντιστοιχούν σε αρνητικά και άλλα 15 instances τα κατηγοριοποιεί αντιστρόφως.

➤ **VALENCE**

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του SVM με τα μέτρα απόδοσης.

**Πίνακας 6: Τα μέτρα απόδοσης συναρτήσει των καλύτερων τιμών των παραμέτρων του SVM**

60%	Επιλογή κατάλληλου C	Γάμμα παράμετρος (g)	ACCURACY	ROC Area (Weighted avg.)
POLYKERNEL	0.5	-----	81,61%	0.814
RBFKERNEL	1.25	0.01	85.05	0,849

**ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ**

<b>POLYKERNEL</b>
a b ← ταξινόμηση ως
39 6   a = neg_val
10 32   b = pos_val

<b>RBFKERNEL</b>
a b ← ταξινόμηση ως
40 5   a = neg_val
8 34   b = pos_val

Άρα, καλύτερα αποτελέσματα έφερε με την παράμετρο του RBFKernel με  $g = 0.01$  και  $C = 1.25$ .

Για το test:

**Πίνακας 7: Τα μέτρα απόδοσης κατά τον έλεγχο με τον SVM**

30%	ACCURACY	ROC Area (Weighted avg.)
RBFKERNEL	60.34%	0.597

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b ← ταξινόμηση ως
13 10   a = neg_val
13 22   b = pos_val

Τα αποτελέσματα για τον έλεγχο με τις παραπάνω επιλογές φαίνεται να είναι σχετικά ικανοποιητικά, αφού η τιμή της ROC Area είναι οριακά μεγαλύτερη του 0.5 και τείνει στο 0.6. Παρακάτω δε, φαίνεται και η ταξινόμηση στις κλάσεις, όπου 13 και 22 κατηγοριοποιεί σωστά στις αντίστοιχες κλάσεις, ενώ 13 και άλλα 10 instances τα κατατάσσει σε κλάσεις που δεν ανήκουν.

➤ **ΣΥΝΑΙΣΘΗΜΑΤΑ**

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του SVM με τα μέτρα απόδοσης.

**Πίνακας 8: Τα μέτρα απόδοσης συναρτήσει των καλύτερων τιμών των παραμέτρων του SVM**

60%	Επιλογή κατάλληλου C	Γάμμα παράμετρος (g)	ACCURACY	ROC Area (Weighted avg.)
POLYKERNEL	0.5	-----	71.26%	0.824
RBFKERNEL	1.5	(default-0.01)	70.11%	0.834
RBFKERNEL	1.5	0.012	68.97%	0.838

### ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ

POLYKERNEL	
a b c d	← ταξινόμηση ως
6 2 0 4	a = em1
1 26 0 3	b = em2
0 3 9 3	c = em3
3 3 3 21	d = em4

RBFKERNEL	
a b c d	← ταξινόμηση ως
1 4 0 7	a = em1
0 27 0 3	b = em2
0 3 9 3	c = em3
1 5 0 24	d = em4

RBFKERNEL/G	
a b c d	← ταξινόμηση ως
0 4 0 8	a = em1
0 27 0 3	b = em2
0 4 8 3	c = em3
1 4 0 25	d = em4

Άρα, καλύτερα αποτελέσματα έφερε με την παράμετρο του RBFKernel με  $g = 0.012$  και  $C = 1.5$ .

Για το **test**:

**Πίνακας 9: Τα μέτρα απόδοσης κατά τον έλεγχο για τα συναισθήματα με τον SVM**

30%	ACCURACY	ROC Area (Weighted avg.)
RBFKERNEL	46.55%	0.597

### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a b c d	← ταξινόμηση ως
0 4 3 4	a = em1
0 8 0 9	b = em2
0 2 2 2	c = em3
1 3 3 17	d = em4

Τα αποτελέσματα για τον έλεγχο με τις παραπάνω επιλογές φαίνεται να είναι σχετικά ικανοποιητικά, αφού η τιμή της ROC Area είναι οριακά μεγαλύτερη του 0.5 και τείνει στο 0.6. Παρακάτω δε, φαίνεται και η ταξινόμηση στις κλάσεις, όπου φαίνεται να έχει πρόβλημα να αναγνωρίσει τα συναισθήματα της πρώτης κατηγορίας- τεταρτημορίου και αυτό αποτυπώνεται και στην τιμή της ακρίβειας, της οποίας το ποσοστό είναι χαμηλό, 46,55%.

#### 4.5.2 KNN

##### ➤ AROUSAL

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του KNN με τα μέτρα απόδοσης.



**Πίνακας 10: Τα μέτρα απόδοσης εκπαίδευσης συναρτήσει των καλύτερων τιμών των παραμέτρων του KNN**

60%	K	ACCURACY	ROC Area (Weighted avg.)
ΕΥΚΛΕΙΔΕΙΑ	1	81.61%	0.810
MANHATTAN	1	86.21%	0.873

**ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ**

ΕΥΚΛΕΙΔΕΙΑ
a b ← ταξινόμηση ως
36 9   a = neg_ar
7 35   b = pos_ar

MANHATTAN
a b ← ταξινόμηση ως
40 5   a = neg_ar
7 35   b = pos_ar

Ο αλγόριθμος KNN έφερε καλύτερα αποτελέσματα με την Manhattan απόσταση και για πλήθος γειτόνων  $K = 1$ .

Άρα, για το **τεστ** ελέγχου:

**Πίνακας 11: Τα μέτρα απόδοσης του ελέγχου με τον KNN**

30%	ACCURACY	ROC Area (Weighted avg.)
MANHATTAN	55.17%	0.550

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b ← ταξινόμηση ως
18 12   a = neg_ar
14 14   b = pos_ar

Τα αποτελέσματα για τον έλεγχο με τις παραπάνω επιλογές φαίνεται να είναι σχετικά μέτρια, αφού η τιμή της ROC Area είναι οριακά μεγαλύτερη του 0.5. Παρακάτω δε, φαίνεται και η ταξινόμηση στις κλάσεις, όπου 18 και 14 instances κατηγοριοποιεί σωστά στις αντίστοιχες κλάσεις, ενώ 12 και άλλα 14 instances τα κατατάσσει σε κλάσεις που δεν ανήκουν.

➤ **VALENCE**

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του KNN με τα μέτρα απόδοσης.

**Πίνακας 12: Τα μέτρα απόδοσης εκπαίδευσης συναρτήσει των καλύτερων τιμών των παραμέτρων του KNN**

60%	K	ACCURACY	ROC Area (Weighted avg.)
ΕΥΚΛΕΙΔΕΙΑ	8	73.56%	0.779
MANHATTAN	1	85.06%	0.844

**ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ**

**ΕΥΚΛΕΙΔΕΙΑ**

a b ← ταξινόμηση ως

41 4 | a = neg\_val

19 23 | b = pos\_val

**MANHATTAN**

a b ← ταξινόμηση ως

42 3 | a = neg\_val

10 32 | b = pos\_val

Ο αλγόριθμος KNN έφερε καλύτερα αποτελέσματα με την Manhattan απόσταση και για πλήθος γειτόνων  $K = 1$ .

Άρα, για το **τεστ** ελέγχου:

**Πίνακας 13: Τα μέτρα απόδοσης κατά τον έλεγχο του Valence με τον KNN**

30%	ACCURACY	ROC Area (Weighted avg.)
MANHATTAN	36.21%	0.382

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b ← ταξινόμηση ως

11 12 | a = neg\_val

25 10 | b = pos\_val

Στο πείραμα ελέγχου φαίνεται η κακή απόδοση του KNN, τόσο η ακρίβεια του, όσο και η μέση τιμή της ROC Area είναι αρκετά χαμηλές. Στον πίνακα δε σύγχυσης αποδεικνύεται γιατί έχει τέτοια αποτελέσματα, αφού ένα μεγάλο αριθμό instances τα κατηγοριοποιεί σε λάθος κλάσεις.

➤ **ΣΥΝΑΙΣΘΗΜΑΤΑ**

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του KNN με τα μέτρα απόδοσης.

**Πίνακας 14: Τα μέτρα απόδοσης εκπαίδευσης συναρτήσει των καλύτερων τιμών των παραμέτρων του KNN**

60%	K	ACCURACY	ROC Area (Weighted avg.)
ΕΥΚΛΕΙΔΕΙΑ	8	60.92%	0.722
MANHATTAN	1	70.11%	0.797

**ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ**

**ΕΥΚΛΕΙΔΕΙΑ**

a b c d ← ταξινόμηση ως

4 3 1 4 | a = em1

3 22 3 2 | b = em2

1 2 9 3 | c = em3

1 7 4 18 | d = em4

**MANHATTAN**

a b c d ← ταξινόμηση ως

4 3 1 4 | a = em1

3 24 2 1 | b = em2

0 3 10 2 | c = em3

1 2 4 23 | d = em4

Ο αλγόριθμος KNN έφερε καλύτερα αποτελέσματα με την μετρική Manhattan απόσταση και για πλήθος γειτόνων  $K = 1$ .

Άρα, για το **τεστ** ελέγχου:

**Πίνακας 15**

30%	ACCURACY	ROC Area (Weighted avg.)
MANHATTAN	18.97%	0.476

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b c d	← ταξινόμηση ως
0 5 6 0	a = em1
5 4 2 6	b = em2
1 4 1 0	c = em3
4 3 11 6	d = em4

Στο πείραμα ελέγχου φαίνεται για άλλη μια φορά η κακή απόδοση του KNN, τόσο η ακρίβεια του, όσο και η μέση τιμή της ROC Area είναι αρκετά χαμηλές. Επίσης, στον πίνακα σύγκρισης αποτυπώνονται ο λόγος αυτών των αποτελεσμάτων, αφού ένα μεγάλο αριθμό instances τα κατηγοριοποιεί σε λάθος κλάσεις, μάλιστα ότι αφορά την πρώτη κλάση συναισθημάτων δεν φαίνεται να αναγνωρίζει κανένα στην σωστή κλάση, παρά μόνο τα τοποθετεί σε λάθος κλάσεις.

**4.5.3 J48**

➤ **AROUSAL**

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του J48 με τα μέτρα απόδοσης.

**Πίνακας 16: Τα μέτρα απόδοσης της εκπαίδευσης συναρτήσκει των καλύτερων τιμών των παραμέτρων του J48**

60%	C	M	ACCURACY	ROC Area (Weighted avg.)
J48	0.001	15	57.47%	0.552

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b	← ταξινόμηση ως
25 20	a = neg_ar
17 25	b = pos_ar

Για το **τεστ** ελέγχου:

**Πίνακας 17: Τα μέτρα απόδοσης ελέγχου του J48**

30%	ACCURACY	ROC Area (Weighted avg.)
J48	55.17%	<b>0.621</b>

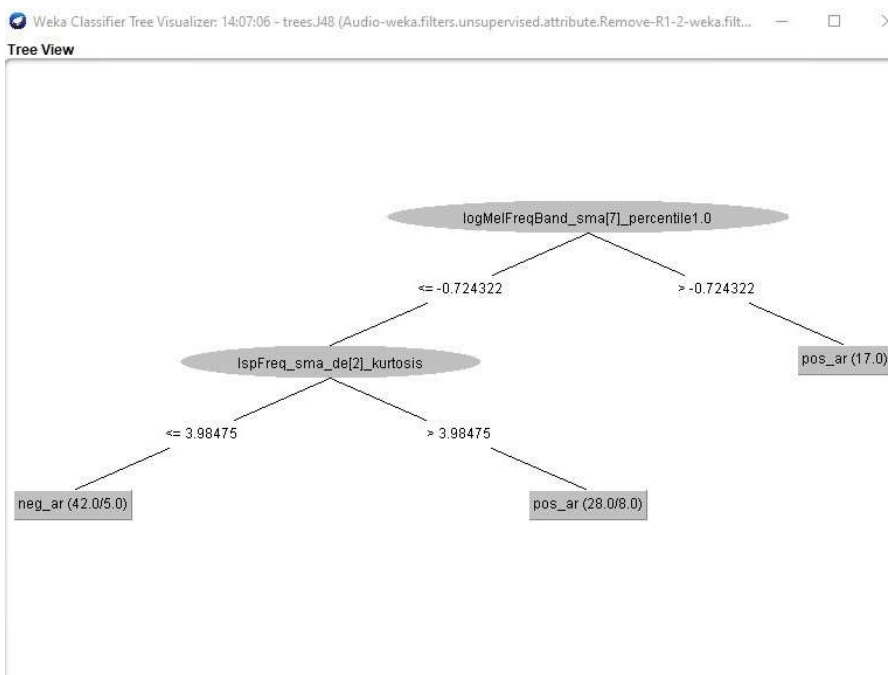
### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a b ← ταξινόμηση ως

14 16 | a = neg\_ar

10 18 | b = pos\_ar

Ο J48 δίνει μια καλή μέση τιμή ROC Area μεγαλύτερη από αυτής της τυχαίας κατηγοριοποίησης. Το δέντρο που προκύπτει έχει αριθμό φύλλων = 3 και μέγεθος δένδρου = 5. Παρακάτω, δίνεται και η απεικόνιση του (Εικόνα 9), όπως το σχηματίζει το εργαλείο της WEKA.



**Εικόνα 9: Το δέντρο κατά τον έλεγχο στις κλάσεις της Arousal**

#### ➤ VALENCE

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του J48 με τα μέτρα απόδοσης για την μετρική Valence.

**Πίνακας 18: Τα μέτρα απόδοσης κατά την εκπαίδευσης συναρτήσει των καλύτερων τιμών των παραμέτρων του J48**

60%	C	M	ACCURACY	ROC Area (Weighted avg.)
J48	0.001	36	79.31%	0.759

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b ← ταξινόμηση ως  
 37 8 | a = neg\_val  
 10 32 | b = pos\_val

Άρα, για το **τεστ** ελέγχου:

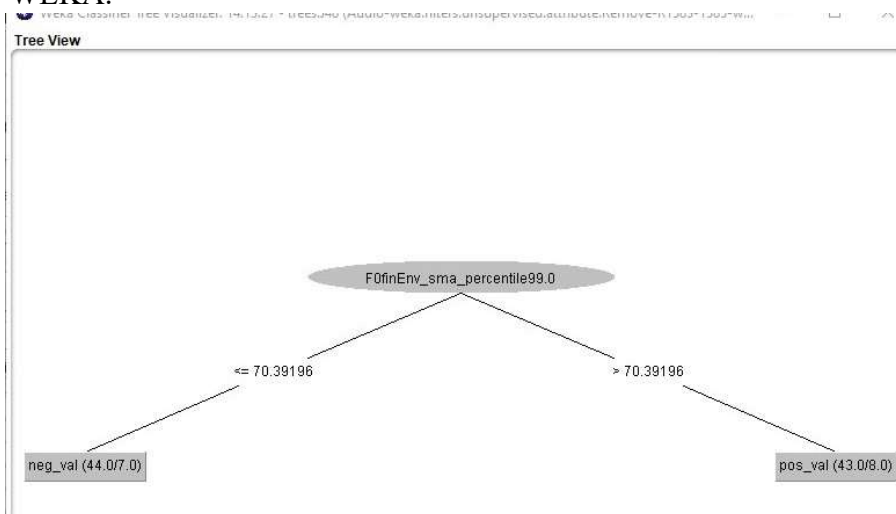
**Πίνακας 19: Τα μέτρα απόδοσης ελέγχου του J48 για την μετρική Valence**

30%	ACCURACY	ROC Area (Weighted avg.)
J48	46.55%	0.505

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b ← ταξινόμηση ως  
 16 7 | a = neg\_val  
 24 11 | b = pos\_val

Ο J48 δίνει μια οριακά μέτρια μέση τιμή ROC Area, μιας και είναι ελάχιστα μεγαλύτερη από αυτής της τυχαίας κατηγοριοποίησης. Το δέντρο που προκύπτει έχει αριθμό φύλλων = 2 και μέγεθος δένδρου = 3. Παρακάτω, δίνεται και η Εικόνα 10, όπως το σχηματίζει το εργαλείο της WEKA.



**Εικόνα 10: Το δέντρο που προκύπτει κατά τον έλεγχο στις κλάσεις της Valence**

### ➤ ΣΥΝΑΙΣΘΗΜΑΤΑ

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του J48 με τα μέτρα απόδοσης για την μετρική των συναισθημάτων.

**Πίνακας 20: Τα μέτρα απόδοσης για την εκπαίδευση συναρτήσει των καλύτερων τιμών των παραμέτρων του J48**

60%	C	M	ACCURACY	ROC Area (Weighted avg.)
J48	0.001	36	56.32%	0.711

### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a b c d ← ταξινόμηση ως 3 2 0 7   a = em1 1 23 0 6   b = em2 1 6 3 5   c = em3 2 5 3 20   d = em4
---------------------------------------------------------------------------------------------------------------

Άρα, για το **τεστ** ελέγχου:

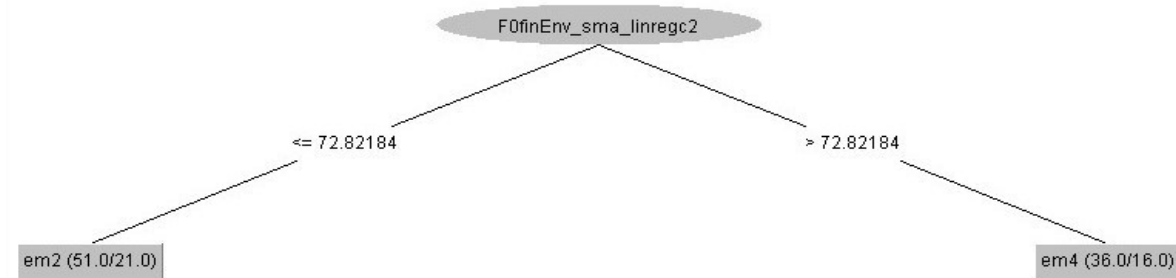
**Πίνακας 21: Τα μέτρα απόδοσης ελέγχου του J48 για την μετρική των συναισθημάτων**

30%	ACCURACY	ROC Area (Weighted avg.)
J48	29.31%	<b>0.489</b>

### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

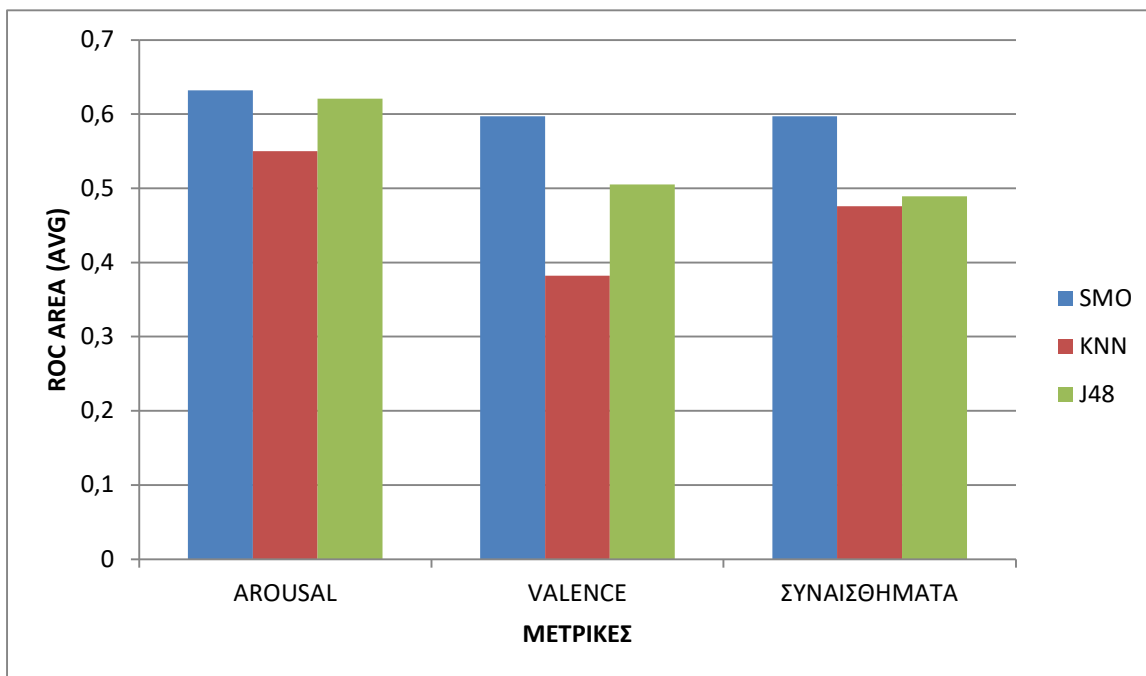
a b c d ← ταξινόμηση ως 0 10 0 1   a = em1 0 13 0 4   b = em2 0 6 0 0   c = em3 0 20 0 4   d = em4
----------------------------------------------------------------------------------------------------------------

Ο J48 δίνει μια όχι τόσο ικανοποιητική μέση τιμή ROC Area, μιας και είναι ελάχιστα μικρότερη από αυτής της τυχαίας κατηγοριοποίησης. Μάλιστα, φαίνεται ότι οι τέσσερις κλάσεις να μην βοηθούν την απόδοση του αλγορίθμου, καθώς κατά την κατηγοριοποίηση της πρώτης κλάσης αποτυγχάνει ολικώς, όπως επίσης και στην τρίτη κλάση. Το δέντρο που προκύπτει έχει αριθμό φύλλων = 2 και μέγεθος δένδρου = 3. Παρακάτω, δίνεται και η απεικόνιση του (Εικόνα 11), όπως το σχηματίζει το εργαλείο της WEKA.



Εικόνα 11: Το δέντρο που προκύπτει κατά τον έλεγχο των συναισθημάτων

#### 4.5.4 Συγκριτικά αποτελέσματα



Γράφημα 16: Όλες οι τιμές των ROC Area που προέκυψαν στα τεστ ελέγχου κατά την πρώτη υπόθεση

Από το παραπάνω γράφημα συμπεραίνεται πως και οι τρεις αλγόριθμοι φέρουν μέτριες αποδόσεις και σχετικά παρόμοιες τιμές. Μάλιστα, δεδομένου ότι το σύνολο δεδομένων ήταν αριθμητικό και για τις δυο πρώτες μετρικές (arousal-valence) προέκυψαν δυο κλάσεις, παρατηρείται μια υπεροχή του SVM με τον J48 να ακολουθεί και τελευταίος ο KNN. Αν πάλι επιθυμείται η χρήση κάποιου αλγορίθμου που θα προσφέρει καλύτερα αποτελέσματα, φαίνεται πως αυτός είναι ο SVM. Για την μετρική συναισθήματα, όπου προέκυψαν τέσσερις κλάσεις, φαίνεται πως ο SVM είναι αυτός που και πάλι έχει τα καλύτερα αποτελέσματα, ωστόσο ακολουθούν ο J48 και KNN με μικρή διαφορά μεταξύ τους και μάλιστα ROC Area μικρότερη του 0.5, στοιχείο που δείχνει αυτοί οι αλγόριθμοι είναι χειρότεροι από την διαγώνια γραμμή – τυχαία κατηγοριοποίηση/πρόβλεψη.

## 4.6 Αποτελέσματα δεύτερης προσέγγισης

Δίνονται τα αποτελέσματα όπως προέκυψαν από την δεύτερη υπόθεση για τις μετρικές Arousal, Valence και συναισθήματα.

### 4.6.1 SVM

#### ➤ AROUSAL

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του SVM με τα μέτρα απόδοσης.

**Πίνακας 22: Τα μέτρα απόδοσης εκπαίδευσης συναρτήσει των καλύτερων τιμών των παραμέτρων του SVM**

60%	Επιλογή κατάλληλου C	Γάμμα παράμετρος (g)	ACCURACY	ROC Area (Weighted avg.)
POLYKERNEL	0.5	-----	65.52%	0.655
RBFKERNEL	1	(default-0.01)	70.11%	0.700

#### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

<p><b>POLYKERNEL</b></p> <p>a b ← ταξινόμηση ως</p> <p>27 16   a = neg_ar</p> <p>14 30   b = pos_ar</p>	<p><b>RBF/G</b></p> <p>a b ← ταξινόμηση ως</p> <p>25 18   a = neg_ar</p> <p>8 36   b = pos_ar</p>
---------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------

Άρα, καλύτερα αποτελέσματα έφερε με την παράμετρο του RBFKernel με  $g = 0.01$  και  $C = 1$ .  
Για το **test**:

**Πίνακας 23: Τα μέτρα απόδοσης ελέγχου για τον SVM**

30%	ACCURACY	ROC Area (Weighted avg.)
RBFKERNEL	70.69%	0.729

#### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

<p>a b ← ταξινόμηση ως</p> <p>16 4   a = neg_ar</p> <p>13 25   b = pos_ar</p>
-------------------------------------------------------------------------------

Παρατηρείται πως ο RBFkernel κατά τον έλεγχο έφερε αρκετά ικανοποιητικά αποτελέσματα με μέση τιμή ROC Area ίση με 0.729, από το δεν πίνακα σύγχυσης λαμβάνεται η πληροφορία πως κατηγοριοποιεί ένα αρκετά μεγάλο αριθμό instances στις σωστές κλάσεις, ενώ μόνο 17 Instances ταξινομούνται συνολικά σε λάθος κλάσεις.



➤ **VALENCE**

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του SVM με τα μέτρα απόδοσης.

**Πίνακας 24:** Τα μέτρα απόδοσης εκπαίδευσης του κατηγοριοποιητή SVM συναρτήσει των καλύτερων τιμών των παραμέτρων του

60%	Επιλογή κατάλληλου C	Γάμμα παράμετρος (g)	ACCURACY	ROC Area (Weighted avg.)
POLYKERNEL	0.5	-----	87.36%	0.795
RBFKERNEL	1.25	0.01	87.36%	0.743

**ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ**

<b>POLYKERNEL</b> a b ← ταξινόμηση ως 13 7   a = neg_val 4 63   b = pos_val
--------------------------------------------------------------------------------------

<b>RBFKERNEL</b> a b ← ταξινόμηση ως 10 10   a = neg_val 1 66   b = pos_val
--------------------------------------------------------------------------------------

Άρα, καλύτερα αποτελέσματα έφερε με την παράμετρο του POLYKernel με C = 0.5.

Για το test:

**Πίνακας 25:** Τα μέτρα απόδοσης ελέγχου της μετρικής Valence του SVM

30%	ACCURACY	ROC Area (Weighted avg.)
POLYKERNEL	56.897%	0.523

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b ← ταξινόμηση ως 2 24   a = neg_val 1 31   b = pos_val
-----------------------------------------------------------------

Μολονότι σε όλους στους ελέγχους ο SVM έφερνε καλύτερα αποτελέσματα με τον πυρήνα RBF, εδώ επιλέγεται ο POLYkernel, οποίος και στα πλαίσια του τεστ έχει μέση τιμή ROC Area ίση με 0.523, ελάχιστα πάνω από την τυχαία κατηγοριοποίηση. Φαίνεται μάλιστα, από τον πίνακα σύγκυσης να έχει πρόβλημα αναγνώρισης των instances εκείνων που ανήκουν κανονικά στην κλάση negative valence.

➤ ΣΥΝΑΙΣΘΗΜΑΤΑ

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του SVM με τα μέτρα απόδοσης.

**Πίνακας 26: Τα μέτρα απόδοσης εκπαίδευσης του κατηγοριοποιητή SVM συναρτήσει των καλύτερων τιμών των παραμέτρων του**

60%	Επιλογή κατάλληλου C	Γάμμα παράμετρος (g)	ACCURACY	ROC Area (Weighted avg.)
POLYKERNEL	0.5	-----	59.77%	0.682
RBFKERNEL	1.25	(default-0.01)	62.07%	0.648
RBFKERNEL	1.25	0.015	60.92%	0.655

**ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ**

POLYKERNEL
a b c d ← ταξινόμηση ως
7 3 0 14   a = em1
3 12 0 5   b = em2
0 0 0 0   c = em3
9 1 0 33   d = em4

RBFKERNEL
a b c d ← ταξινόμηση ως
0 1 0 23   a = em1
0 11 0 9   b = em2
0 0 0 0   c = em3
0 0 0 43   d = em4

RBFKERNEL/G
a b c d ← ταξινόμηση ως
2 2 0 20   a = em1
1 11 0 8   b = em2
0 0 0 0   c = em3
3 0 0 40   d = em4

Άρα, καλύτερα αποτελέσματα έφερε με την παράμετρο του POLYKernel με C = 0.5.  
Για το test:

**Πίνακας 27: Τα μέτρα απόδοσης ελέγχου του SVM για την μετρική των συναισθημάτων**

30%	ACCURACY	ROC Area (Weighted avg.)
POLYKERNEL	36.21%	0.588

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b c d ← ταξινόμηση ως
4 1 0 9   a = em1
9 2 0 13   b = em2
1 0 0 1   c = em3
2 1 0 15   d = em4

Τα αποτελέσματα για τον έλεγχο με τις παραπάνω επιλογές φαίνεται να είναι σχετικά ικανοποιητικά, αφού η τιμή της ROC Area είναι οριακά μεγαλύτερη του 0.5 και τείνει στο 0.6. Παρακάτω δε, φαίνεται και η ταξινόμηση στις κλάσεις, όπου φαίνεται να έχει πρόβλημα να αναγνωρίσει τα συναισθήματα της δεύτερης και τρίτης κατηγορίας- τεταρτημορίου και αυτό αποτυπώνεται και στην τιμή της ακρίβειας, της οποίας το ποσοστό είναι χαμηλό, 36,21%. Αξίζει

να σημειωθεί πως το πλήθος των στιγμιότυπων με ετικέτα τρίτης κλάσης-συναισθήματος ήταν εξαιρετικά μικρό, πράγμα που έδινε πιθανότητες για λάθη στην κατηγοριοποίηση.

#### 4.6.2 KNN

##### > AROUSAL

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του KNN με τα μέτρα απόδοσης.

**Πίνακας 28: Τα μέτρα απόδοσης συναρτήσε των καλύτερων τιμών των παραμέτρων του KNN**

60%	K	ACCURACY	ROC Area (Weighted avg.)
ΕΥΚΛΕΙΔΕΙΑ	1	73.56%	0.763
MANHATTAN	1	88.506%	0.872

#### ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ

ΕΥΚΛΕΙΔΕΙΑ
a b ← ταξινόμηση ως
29 14   a = neg_ar
9 35   b = pos_ar

MANHATTAN
a b ← ταξινόμηση ως
38 5   a = neg_ar
5 39   b = pos_ar

Ο αλγόριθμος KNN έφερε καλύτερα αποτελέσματα με την μετρική Manhattan απόσταση και για πλήθος γειτόνων  $K = 1$ .

Άρα, για το **τεστ** ελέγχου:

**Πίνακας 29: Τα μέτρα απόδοσης ελέγχου του KNN**

30%	ACCURACY	ROC Area (Weighted avg.)
MANHATTAN	62.07%	0.639

#### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a b ← ταξινόμηση ως
14 6   a = neg_ar
16 22   b = pos_ar

Ο KNN με απόσταση Manhattan κατά το τεστ, έδωσε μια καλή μέση τιμή της ROC Area, μεγαλύτερη από 0.5 και ακρίβειας, το οποίο αποτυπώνεται και στον πίνακα σύγκυσης.

➤ **VALENCE**

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του KNN με τα μέτρα απόδοσης.

**Πίνακας 30: Τα μέτρα απόδοσης συναρτήσει των καλύτερων τιμών των παραμέτρων του KNN**

60%	K	ACCURACY	ROC Area (Weighted avg.)
ΕΥΚΛΕΙΔΕΙΑ	15	83.908%	0.741
MANHATTAN	15	77.012%	0.741

**ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ**

<b>ΕΥΚΛΕΙΔΕΙΑ</b>	<b>MANHATTAN</b>
a b ← ταξινόμηση ως	a b ← ταξινόμηση ως
12 8   a = neg_val	12 8   a = neg_val
6 61   b = pos_val	12 55   b = pos_val

Ο αλγόριθμος KNN έφερε ίδια αποτελέσματα τόσο με την Ευκλείδεια απόσταση, όσο και με την Manhattan, ωστόσο θα δεχτούμε την επιλογή με την Ευκλείδεια μιας και ακόμη μετρική που κρίνει το αποτέλεσμα αυτή τα ακρίβειας είναι μεγαλύτερη από αυτής με απόσταση Manhattan. Έτσι, γίνεται έλεγχος με Ευκλείδεια απόσταση και για πλήθος γειτόνων K = 15.

Άρα, για το **τεστ** ελέγχου:

**Πίνακας 31: Τα μέτρα απόδοσης ελέγχου του KNN για την μετρική Valence**

30%	ACCURACY	ROC Area (Weighted avg.)
ΕΥΚΛΕΙΔΕΙΑ	55.17%	0.681

**ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ**

a b ← ταξινόμηση ως
0 26   a = neg_val
0 32   b = pos_val

Φαίνεται πως ο KNN δεν έχει κανένα πρόβλημα να ταξινομήσει σωστά όλα τα negative valence, όμως προκύπτει κάποια δυσκολία για την πρώτη κλάση, καθώς όλα τα στιγμιότυπα αναγνωρίζονται ως δεύτερης κλάσης λανθασμένα. Παρόλο που η ROC Area είναι ικανοποιητική, αξίζει να παρατηρηθεί το πλήθος το γειτόνων K=15, που βρήκε ως κατάλληλο ο meta CVParameter αλγόριθμος συναρτήσει του KNN. Ειδικότερα, είναι σχετικά μεγάλος αριθμός μπροστά στο πλήθος τόσο της εκπαίδευσης, όσο και στου ελέγχου. Συνεπώς, τίθεται το ερώτημα

αν το τεστ που έγινε έχει νόημα και είναι εξίσου αξιόπιστο ή θα έπρεπε να ερευνηθεί μια άλλη προσέγγιση για το εύρος των παραμέτρων συναρτήσει του εκάστοτε δείγματος των πειραμάτων σε μια έρευνα.

### ➤ ΣΥΝΑΙΣΘΗΜΑΤΑ

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του KNN με τα μέτρα απόδοσης.

**Πίνακας 32: Τα μέτρα απόδοσης κατά την εκπαίδευση συναρτήσει των καλύτερων τιμών των παραμέτρων του KNN**

60%	K	ACCURACY	ROC Area (Weighted avg.)
ΕΥΚΛΕΙΔΕΙΑ	1	59.77%	0.736
ΜΑΝΗΑΤΤΑΝ	1	64.37%	0.760

### ΠΙΝΑΚΕΣ ΣΥΓΧΥΣΗΣ

ΕΥΚΛΕΙΔΕΙΑ				
a	b	c	d	← ταξινόμηση ως
10	8	0	6	a = em1
5	13	0	2	b = em2
0	0	0	0	c = em3
9	5	0	29	d = em4

ΜΑΝΗΑΤΤΑΝ				
a	b	c	d	← ταξινόμηση ως
10	8	0	6	a = em1
6	12	0	2	b = em2
0	0	0	0	c = em3
6	3	0	34	d = em4

Ο αλγόριθμος KNN έφερε καλύτερα αποτελέσματα με την Manhattan απόσταση και για πλήθος γειτόνων  $K = 1$ .

Άρα, για το **τεστ** ελέγχου:

**Πίνακας 33: Τα μέτρα απόδοσης ελέγχου του KNN για την μετρική των συναισθημάτων**

30%	ACCURACY	ROC Area (Weighted avg.)
ΜΑΝΗΑΤΤΑΝ	39.66%	0.576

### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a	b	c	d	← ταξινόμηση ως
5	1	0	8	a = em1
11	5	0	8	b = em2
1	0	0	1	c = em3
4	1	0	13	d = em4

Τα αποτελέσματα για τον έλεγχο με τις παραπάνω επιλογές φαίνεται να είναι σχετικά ικανοποιητικά, αφού η τιμή της ROC Area είναι οριακά μεγαλύτερη του 0.5 και τείνει στο 0.6. Παρακάτω δε, φαίνεται και η ταξινόμηση στις κλάσεις, όπου φαίνεται να έχει πρόβλημα να αναγνωρίσει τα συναισθήματα της τρίτης κατηγορίας- τεταρτημορίου και αυτό αποτυπώνεται και στην τιμή της ακρίβειας, της οποίας το ποσοστό είναι χαμηλό, 39,66%. Αξίζει να σημειωθεί για άλλη μια φορά πως το πλήθος των στιγμιότυπων με ετικέτα τρίτης κλάσης-συναισθήματος ήταν εξαιρετικά μικρό, πράγμα που έδινε πιθανότητες για λάθη στην κατηγοριοποίηση.

#### 4.6.3 J48

##### ➤ AROUSAL

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του J48 με τα μέτρα απόδοσης.

**Πίνακας 34: Τα μέτρα απόδοσης της εκπαίδευσης συναρτήσκει των καλύτερων τιμών των παραμέτρων του J48**

60%	C	M	ACCURACY	ROC Area (Weighted avg.)
J48	0.001	8	63.22%	0.592

#### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a b ← ταξινόμηση ως  
28 15 | a = neg\_ar  
17 27 | b = pos\_ar

Για το **τεστ** ελέγχου:

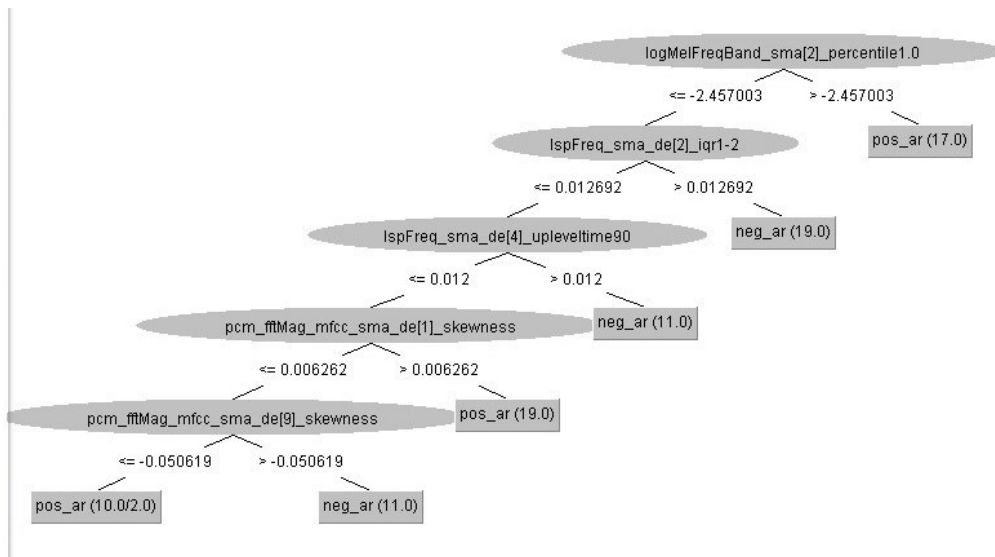
**Πίνακας 35: Τα μέτρα απόδοσης ελέγχου του J48 για την μετρική Arousal**

30%	ACCURACY	ROC Area (Weighted avg.)
J48	62.07%	0.680

#### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a b ← ταξινόμηση ως  
13 7 | a = neg\_ar  
15 23 | b = pos\_ar

Κατά τον έλεγχο των κλάσεων της διέγερσης ο J48, δίνει αρκετά καλά αποτελέσματα τόσο στη μέση τιμή της ROC, όσο και στην ακρίβεια και όμοια κατά την κατηγοριοποίηση στις κλάσεις, όπως αποδεικνύεται από τον πίνακα σύγχυσης. Επίσης, το δέντρο που προκύπτει έχει αριθμό φύλλων έξι και μέγεθος δέντρου έντεκα (Εικόνα 12).



Εικόνα 12: Το δέντρο που σχηματίζεται κατά τον έλεγχο με τη χρήση J48.

### ➤ VALENCE

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του J48 με τα μέτρα απόδοσης.

Πίνακας 36: Τα μέτρα απόδοσης της εκπαίδευσης συναρτήσκει των καλύτερων τιμών των παραμέτρων του J48

60%	C	M	ACCURACY	ROC Area (Weighted avg.)
J48	0.001	2	86.21%	0.794

### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a b ← ταξινόμηση ως
14 6   a = neg_val
6 61   b = pos_val

Άρα, για το **τεστ** ελέγχου:

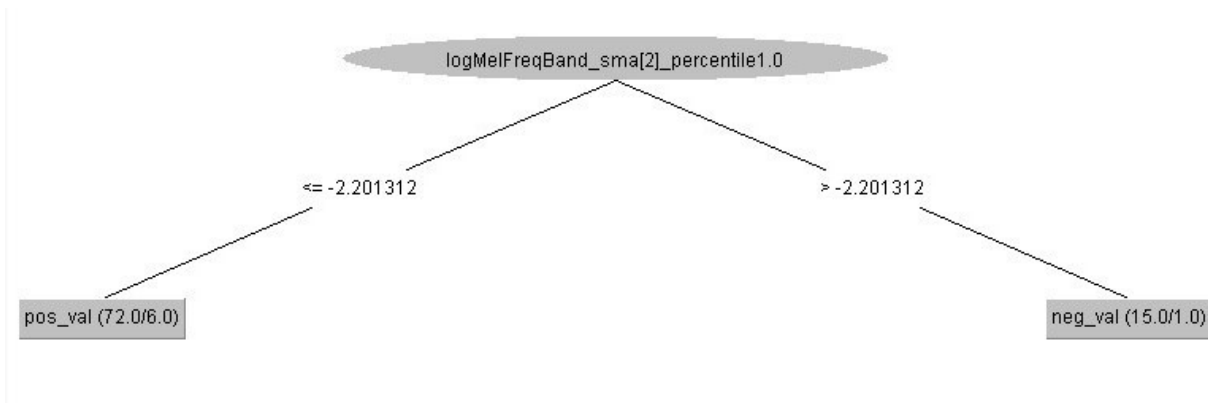
Πίνακας 37: Τα μέτρα απόδοσης ελέγχου του J48 για την μετρική J48

30%	ACCURACY	ROC Area (Weighted avg.)
J48	63.79%	0.603

### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a b ← ταξινόμηση ως
7 19   a = neg_val
2 30   b = pos_val

Κατά τον έλεγχο των κλάσεων της διέγερσης ο J48, δίνει αρκετά καλά αποτελέσματα τόσο στη μέση τιμή της ROC, όσο και στην ακρίβεια και όμοια κατά την κατηγοριοποίηση στις κλάσεις, όπως αποδεικνύεται από τον πίνακα σύγχυσης, από τον οποίο συμπεραίνεται και οποία δυσκολία ως προς την αναγνώριση της πρώτης κλάσης κατά την κατηγοριοποίηση. Επίσης, το δέντρο που προκύπτει έχει αριθμό φύλλων έξι και μέγεθος δέντρου έντεκα. Τέλος, δίνεται το δέντρο από το περιβάλλον της WEKA, όπου έχει μέγεθος τρία και αριθμό φύλλων δυο (Εικόνα 13).



Εικόνα 13: Το δέντρο κατά τον έλεγχο με τη χρήση J48.

### ➤ ΣΥΝΑΙΣΘΗΜΑΤΑ

Παρακάτω δίνεται ο πίνακας με τις επιλαχούσες τιμές των παραμέτρων του J48 με τα μέτρα απόδοσης.

Πίνακας 38: Πίνακας εκπαίδευσης για τα συναισθήματα με τον κατηγοριοποιητή J48

60%	C	M	ACCURACY	ROC Area (Weighted avg.)
J48	0.001	8	62.07%	0.689

### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a	b	c	d	← ταξινόμηση ως
8	3	0	13	a = em1
1	14	0	5	b = em2
0	0	0	0	c = em3
8	3	0	32	d = em4

Άρα, για το **τεστ** ελέγχου, προκύπτει ο παρακάτω πίνακας:

Πίνακας 39: Πίνακας ελέγχου για τα συναισθήματα του J48

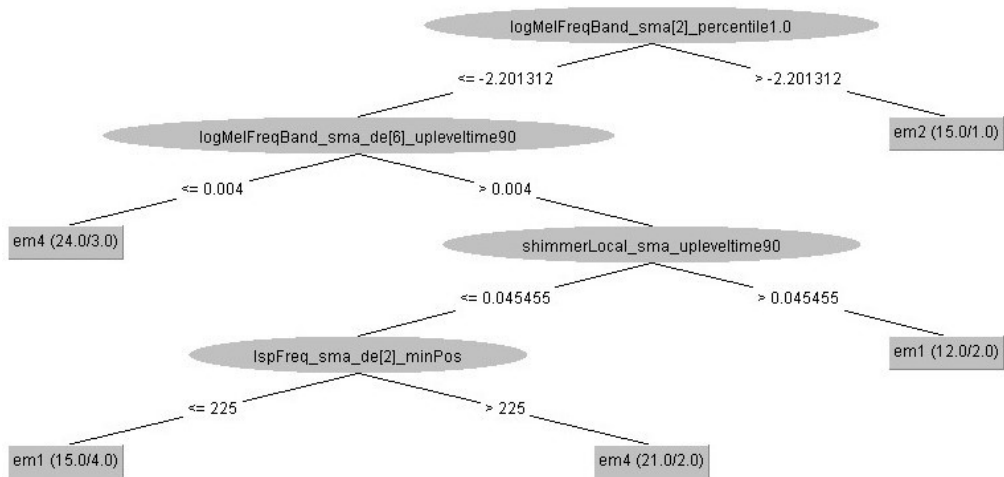
30%	ACCURACY	ROC Area (Weighted avg.)
J48	41.38%	0.583



### ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ

a	b	c	d	← ταξινόμηση ως
7	2	0	5	a = em1
9	7	0	8	b = em2
1	0	0	1	c = em3
8	0	0	10	d = em4

Στον τελευταίο έλεγχο, τα αποτελέσματα φαίνεται να είναι σχετικά ικανοποιητικά, αφού η τιμή της ROC Area είναι οριακά μεγαλύτερη του 0.5 και τείνει στο 0.6. Παρακάτω δε, φαίνεται και η ταξινόμηση στις κλάσεις, όπου φαίνεται να έχει πρόβλημα να αναγνωρίσει τα συναισθήματα της τρίτης κατηγορίας- τεταρτημορίου και αυτό αποτυπώνεται και στην τιμή της ακρίβειας, της οποίας το ποσοστό είναι χαμηλό, 39,66%. Αξίζει να σημειωθεί για άλλη μια φορά πως το πλήθος των στιγμιότυπων με ετικέτα τρίτης κλάσης-συναισθήματος ήταν εξαιρετικά μικρό, πράγμα που έδινε πιθανότητες για λάθη στην κατηγοριοποίηση. Τέλος, το δέντρο που προκύπτει, όπως φαίνεται και στην Εικόνα 14, έχει μέγεθος εννέα και πλήθος φύλλων πέντε.

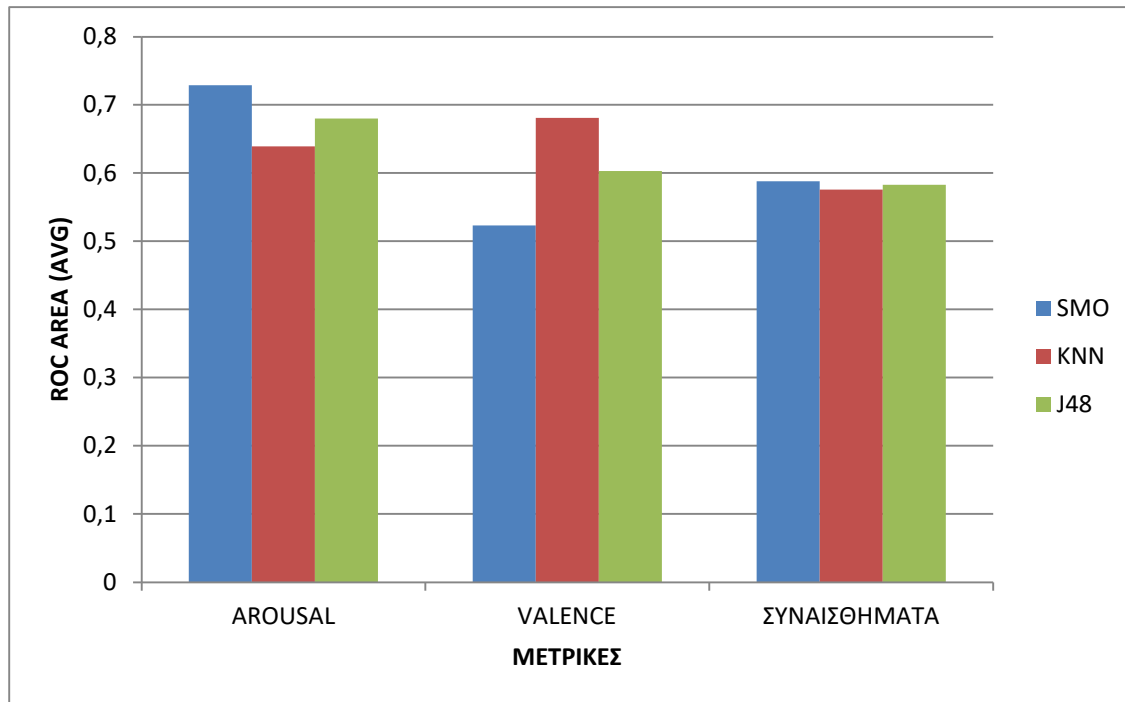


Εικόνα 14: Το δέντρο του J48 κατά την ταξινόμηση σε κλάσεις συναισθημάτων

#### 4.6.4 Συγκριτικά αποτελέσματα

Συμπερασματικά όπως παρατηρεί κάποιος στο Γράφημα 17, για ένα σύνολο δεδομένων με αριθμητικά δεδομένα και μόνο δυο κλάσεις δεν προκύπτει κάποιος αλγόριθμος, ο οποίος να χρησιμοποιήθηκε και να υπερτερεί με μεγάλη διαφορά σε απόδοση, μιας και όλοι έχουν μέση τιμή της ROC Area μεγαλύτερη από 0.5. Ωστόσο, φαίνεται πως κατά την κατηγοριοποίηση των κλάσεων της διέγερσης ο SVM είναι αυτός με τα καλύτερα δυνατά αποτελέσματα με ελαφρώς καλύτερη απόδοση από τους άλλους. Αν όμως κοιτάξουμε την κατηγοριοποίηση των κλάσεων της valence ο KNN όχι μόνο είναι ο καλύτερος, αλλά εν συγκρίσει με τον SVM που σε όλους τους ελέγχους υπερτερούσε, τον αφήνει τρίτο με τον J48 να είναι ελαφρώς χειρότερος από τον KNN.

Τέλος, για τα συναισθήματα φαίνεται πως όλοι κυμαίνονται στο ίδιο εύρος τιμής της ROC Area, αλλά με ένα μικρό βήμα υπεροχής του SVM.



Γράφημα 17: Όλες οι τιμές των ROC Area που προέκυψαν στα τεστ ελέγχου κατά την δεύτερη υπόθεση

# 5

## Συζήτηση

Εν κατακλείδι, δεν είναι εύκολο να συμπεράνει κανείς τον καλύτερο αλγόριθμο που πρέπει να χρησιμοποιηθεί για την εξόρυξη δεδομένων συναισθημάτων. Ο αλγόριθμος που θα χρησιμοποιηθεί εξαρτάται σε μεγάλο βαθμό από τις κλάσεις που δοθούν ως προς την κατηγοριοποίηση, μιας και φαίνεται πως για δυο κλάσεις ο σχετικά καλύτερος ήταν ο SVM, όσον αφορά την πρώτη υπόθεση. Μάλιστα, ο επόμενος με μικρή διαφορά ήταν ο J48. Τελευταίος ακολουθούσε ο KNN. Η δεύτερη υπόθεση διαφοροποιήθηκε εν μέρει, καθώς για πρώτη φορά ο KNN στον έλεγχο κατά την κατηγοριοποίηση των κλάσεων της Valence ήρθε πρώτος, με τον SVM να έχει μέση τιμή ROC Area μικρότερη από 0.5, της τυχαίας δηλαδή κατηγοριοποίησης. Όσον αφορά την ταξινόμηση με κλάσεις τα τέσσερα τεταρτημόρια συναισθημάτων κατά Russell, ο SVM ήταν ο καλύτερος αλλά με μικρή διαφορά, ωστόσο όλοι τους φάνηκαν να δυσκολεύονται να αναγνωρίσουν ορισμένες κλάσεις σωστά. Ειδικότερα, στην δεύτερη υπόθεση η τρίτη κλάση δεν αναγνωρίστηκε από κανέναν σωστά. Βέβαια, σημασία έχει και το πλήθος των ετικετών που αντιστοιχούν στην τρίτη κλάση, το οποίο και ήταν εξαιρετικά μικρό.

Μολονότι έγινε μια προσπάθεια προσέγγισης των συναισθημάτων κατά την ακρόαση ακουσματικής μουσικής με βάση τα χαρακτηριστικά που προκύπτουν από αυτό το κομμάτι, κρίνεται αναγκαία η περαιτέρω έρευνα. Αυτό οφείλεται στο πλήθος των συμμετεχόντων που ήταν μικρό και όπως φάνηκε και από την δεύτερη υπόθεση στην δημιουργία ετικετών προέκυψε μικρό πλήθος δεδομένων για κάποιες κλάσεις συναισθημάτων. Έτσι, η δυνατότητα εκμάθησης και άρα και ορθής πρόβλεψης ήταν λογικό να έχει τέτοια αποτελέσματα.

Κλείνοντας, αξίζει να σημειωθεί και η μελλοντική επέκταση που θα μπορούσε να λάβει η παρούσα εργασία. Πιο συγκεκριμένα, θα είχε ενδιαφέρον και η έρευνα γύρω από τα χαρακτηριστικά που προέκυψαν από το ακουσματικό κομμάτι, ώστε να βρεθούν εκείνα που προκαλούν ενδεχομένως θόρυβο. Ακόμα, θα μπορούσε να αξιολογηθούν εκείνα τα

χαρακτηριστικά που προκαλούν το εκάστοτε συναίσθημα είτε αναφέρεται κανείς στις μετρικές arousal και valence, είτε στον συνδυασμό αυτών, δηλαδή στα συναισθήματα, όπως δόθηκαν από τον Russell. Μια πρώτη επαφή έγινε κατά την μελέτη των διαγραμμάτων στο Κεφάλαιο 3, ώστε υπάρχουν κάποιες γενικεύσεις που δεν δίνουν ασφαλή και έγκυρα αποτελέσματα.

## ***Βιβλιογραφία***

- Alex G. C. de Sa, Cristiano G. Pimenta et al. (2020), A Robust Experimental Evaluation of Automated Multi-Label Classification Methods. *arXiv, Cornell University*.
- Alin G. Chitu. (), Comparison Between Different Feature Extraction Techniques for Audio-Visual Speech Recognition. *Journal on Multimodal User Interfaces*.
- Anna Aljanaki. (2016), Emotion in Music: representation and computational modeling.
- BAIXI XING. (2019), Exploiting EEG Signals and Audiovisual Feature Fusion for Video Emotion Recognition. *IEEE Access*.
- Brasma Jumaa Saleh. (2020), A review paper: Analysis of WEKA data mining techniques for heart disease prediction system. *University of Nebraska-Lincoln, Library Philosophy and Practice (e-journal)*.
- Bülent Üstün, Willem J. Melssen, and Lutgarde M. C. Buydens. (2006), Facilitating the Application of Support Vector Regression by Using a Universal Pearson VII Function based kernel. *Chemometrics and Intelligent Laboratory Systems* 81, 1 (2006), 29–40.
- Calvo R., A., & D'Mello S. (2010). Affect Detection: An Interdisciplinary Review of Models, methods.
- Dimitris Pappas. (2015), Anger Detection in Call Dialogues. *AUEB'S NLP Group*.
- Eibe Frank. (2016), The WEKA Workbench. *Fourth edition DATA MINING*.
- Faiza Umar Bawah, Ussiph. (2018), Appraisal of the Classification Technique in Data Mining of Student Performance using J48 Decision Tree, K-Nearest Neighbor and Multilayer Perceptron Algorithms. *International Journal of Computer Applications, Vol 179- No. 3*.
- Florian Eyben. (2015), The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing. *IEEE Transaction on Affective Computing*.
- Florian Eyben. Real-time Speech and Music Classification by large Audio Feature Space Extraction. *Springer*.
- Florian Eyben. (2014), openSMILE open-Source Media Interpretation by Large feature-space Extraction. Version 2.1. *audEERING UG*.
- Florian Eyben. (2010), openSMILE- The Munich Versatile and Fast Open-Source Audio Feature Extractor.
- Gwenaelle Cunha Sergio. (2020), Emotional Video to Audio Transformation Using Deep Recurrent Neural Networks and a Neuro-Fuzzy System. *Hindawi Mathematical Problems in Engineering*. <https://doi.org/10.1155/2020/8478527>
- J.A. Russell, J. Bachorowski, J. Fernandez-Dols. (2003), "Facial and Vocal Expressions of

Emotion".

James A. Russell. (1980), A circumplex model of affect.

J.A. Russell, J. Bachorowski, J. Fernandez-Dols. (2003), "Facial and Vocal Expressions of Emotion".

I. H. Witten, E. Frank, and M. A. Hall. (2011), Data Mining: Practical Machine Learning Tools and Techniques.

J. Pradeep Kandhasamy S. Balamurali. (2015), Performance Analysis of Classifier Models to Predict Diabetes Mellitus. *Procedia Computer Science* 47 45 – 51.

Jain & Srivastava, (2013). Broad host range plasmids. *FEMS Microbiology Letters*, Vol 348, Issue 2, Pages 87-96.

Juslin, P, N & Vastfj, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *The Behavioral and brain sciences*, 31(5):559–621.

Kittipong Chomboon, P. Chujai. (2015), An Empirical Study of Distance Metrics for k-Nearest Neighbor Algorithm.

Lytra Georgia. Hyperparameter Optimization on Supervised Learning Models. *Univeristy of Piraeus*.

Marc Escalona Mena. (2012), Emotion recognition from speech signals. *Univeristy of Ljubljana*.

Picard, R.W. (2003), Affective Computing. *M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 321*

Rab Nawaz. (2018), Acoustic feature extraction from music songs to predict emotions using neural networks. *2nd International Conference on BioSignal Analysis*.

Remco R. Bouckaert. (2013), WEKA Manual for Version 3-7-8. *THE UNIVERISTY OF WAIKATO*.

Russell SJ, Norvig P. (2016), Artificial intelligence: a modern approach. *Pearson Education Limited, London*.

Shakir Shabbir. (2015), Comparing Performance of J48, Multilayer Perceptron & Naïve Bayes Classifiers on Breast Cancer Data Set using WEKA

Simon Emmerson. (2001), Electro-acoustic music. *Groove Music Online*.

T. Cover, P. Hart. (1967), Nearest neighbor pattern classification. *Mathematics, Computer Science IEEE Trans. Inf. Theory*.

Αλέξανδρος Κυριαζής (), Ανάπτυξη αλγορίθμου κατηγοριοποίησης με αυξητικό τρόπο για ανάλυση συναισθήματος στο twitter. *ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ*.

Αννέτα Μίγκου. (2007), Δημιουργία μοντέλου γνώσης με την χρήση εργαλείων του data mining του WEKA. *Πανεπιστήμιο Μακεδονίας*.

Γεώργιος Κωνσταντίνου. (2012), Μελέτη κατηγοριοποίησης δεδομένων με Μηχανές Διανοσμάτων Υποστήριξης (Support Vector Machines) και υλοποίηση εφαρμογής. *ΤΕΙ ΣΕΡΡΩΝ*.

Γεώργιος Πάκος. (2015), Διαχείριση Γνώσης και Ασφάλεια Πληροφοριών. *ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ*.

Γρηγόριος Καλαντώνης. (2019), Διάγνωση διαβήτη με χρήση αλγορίθμων μηχανικής μάθησης. *Πανεπιστήμιο Πειραιώς*.

Δημήτριος Παππάς. (2015), Αναγνώριση συναισθημάτων σε προφορικούς διαλόγους, *Πανεπιστήμιο Αθηνών*.

- Ευστάθιος Γ. Κύρκος. (2015), Επιχειρηματική Ευφυΐα & Εξόρυξη Δεδομένων. *Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, Εθνικό Μετσόβιο Πολυτεχνείο.*
- Ιωάννα Αλεξοπούλου. (2020), Ηλεκτροακουστική μουσική: Από το 1948 έως τις μέρες μας. <https://artic.gr/hlektroakoustikh-mousikh/>
- Μαρίλνα Παπαδάκη. (2012), Μηχανές διανυσματικής υποστήριξης (SVMs) και εφαρμογές σε πραγματικά σεισμολογικά δεδομένα.
- Νικόλαος Μαυροφόρος. (2019), Συναισθηματική Ανάλυση σε Στίχους και Μουσική χρησιμοποιώντας Μηχανική Μάθηση. *Πανεπιστήμιο Πατρών.*
- Παναγιώτα Κωστάκη. (2020), Διαχείριση Δεδομένων στις πλατφόρμες KNIME & WEKA. *Πανεπιστήμιο Δυτικής Αττικής.*
- Πέτρος Ψαθόπουλος. (2020), Αυτόματη αναγνώριση και κατηγοριοποίηση του συναισθήματος της ανθρώπινης ομιλίας, μέσα από τεχνικές ανάλυσης, ψηφιακής επεξεργασίας και εξαγωγής χαρακτηριστικών φασματικού χαρακτήρα, MFCC (Mel Frequency Cepstral Coefficients) και αναγνώρισης προτύπων. *ΕΑΠ.*
- Wikipedia:<https://el.wikipedia.org/wiki/%CE%A0%CE%BB%CE%B7%CE%B8%CE%BF%CF%80%CE%BF%CF%81%CE%B9%CF%83%CE%BC%CF%8C%CF%82>
- Wikipedia, Matlab εγχειρίδιο. Διαθέσιμο στη διεύθυνση: <http://en.wikipedia.org/wiki/MATLAB>.
- Wikipedia, PANAS. Διαθέσιμο στην διεύθυνση: [https://en.wikipedia.org/wiki/Positive\\_and\\_Negative\\_Affect\\_Schedule](https://en.wikipedia.org/wiki/Positive_and_Negative_Affect_Schedule).





## Παράρτημα I-openSMILE

Παρακάτω παραθέτονται οι εντολές που δόθηκαν στο περιβάλλον της openSMILE για την εξαγωγή χαρακτηριστικών δεδομένων του ακουσματικού κομματιού (Εικόνα 15) και το μήνυμα για το πλήθος τους (Εικόνα 16).

```
C:\Users\Server\Desktop\opensmile-3.0.0\build\progsrc\smileextract\Release>SMILEExtract  
-C C:\Users\Server\Desktop\opensmile-3.0.0\config\audiovisual\audio.conf  
-A C:\Users\Server\Desktop\ACOUSMATIC\Tracks\Ballistichory_ST_48_24.WAV  
-P C:\Users\Server\Desktop\bal_opens.arff
```

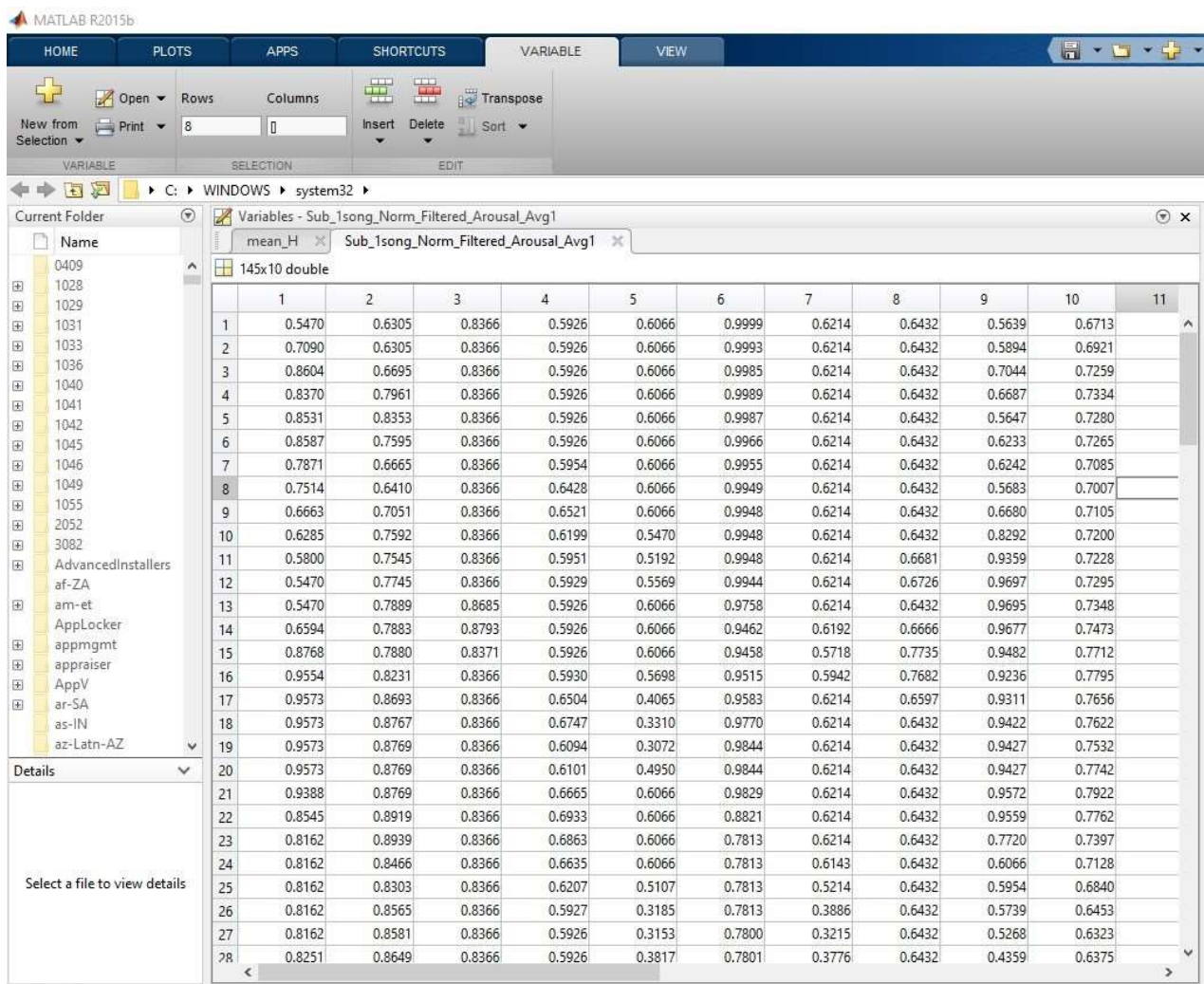
Εικόνα 15: Οι εντολές που δόθηκαν στην openSMILE για την εξαγωγή δεδομένων

```
(MSG) [2] instance 'arffsink': writing ARFF header (1580 features)...
```

Εικόνα 16: Το μήνυμα που εμφανίστηκε για την εξαγωγή 1580 χαρακτηριστικών.

## Παράρτημα II-Matlab

Παρακάτω εμφανίζεται το περιβάλλον της Matlab κατά την επεξεργασία των δεδομένων των χρηστών, όπως προσεγγίστηκαν στις δυο υποθέσεις. Στην Εικόνα 17 φαίνεται η δημιουργία του πίνακα (145x10) με όλα τα δεδομένα της μετρικής Arousal των εννέα χρηστών και η δημιουργία της δέκατης στήλης, η οποία προέκυψε από τον υπολογισμό της μέσης τιμής για κάθε γραμμή, κατά την πρώτη υπόθεση.

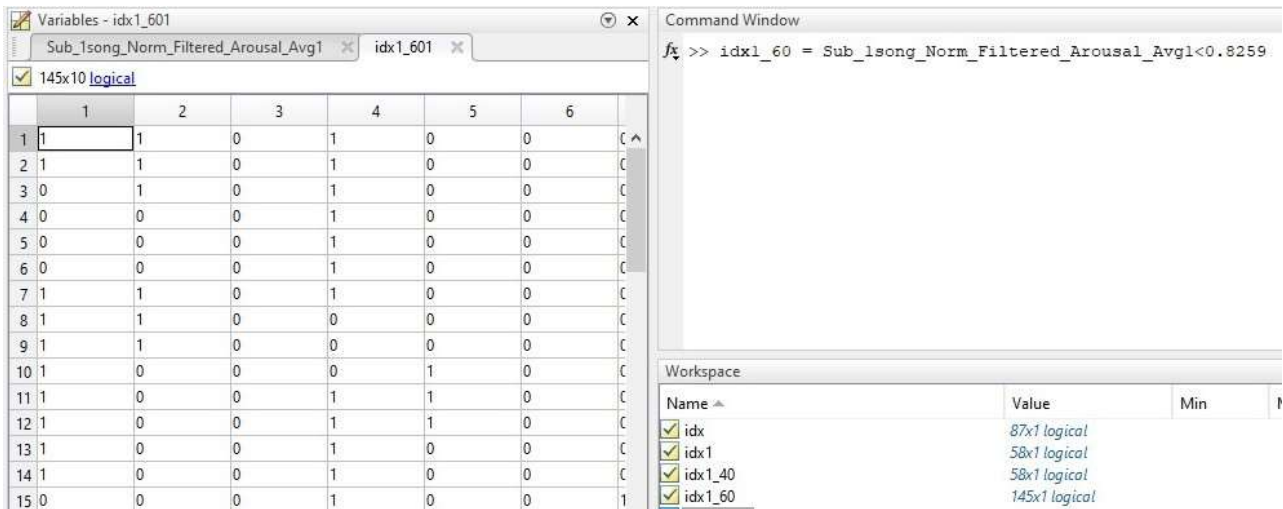


The screenshot shows the MATLAB R2015b environment. The 'Variables' window displays a variable named 'Sub\_1song\_Norm\_Filtered\_Arousal\_Avg1' of type '145x10 double'. The matrix data is as follows:

	1	2	3	4	5	6	7	8	9	10	11
1	0.5470	0.6305	0.8366	0.5926	0.6066	0.9999	0.6214	0.6432	0.5639	0.6713	
2	0.7090	0.6305	0.8366	0.5926	0.6066	0.9993	0.6214	0.6432	0.5894	0.6921	
3	0.8604	0.6695	0.8366	0.5926	0.6066	0.9985	0.6214	0.6432	0.7044	0.7259	
4	0.8370	0.7961	0.8366	0.5926	0.6066	0.9989	0.6214	0.6432	0.6687	0.7334	
5	0.8531	0.8353	0.8366	0.5926	0.6066	0.9987	0.6214	0.6432	0.5647	0.7280	
6	0.8587	0.7595	0.8366	0.5926	0.6066	0.9966	0.6214	0.6432	0.6233	0.7265	
7	0.7871	0.6665	0.8366	0.5954	0.6066	0.9955	0.6214	0.6432	0.6242	0.7085	
8	0.7514	0.6410	0.8366	0.6428	0.6066	0.9949	0.6214	0.6432	0.5683	0.7007	
9	0.6663	0.7051	0.8366	0.6521	0.6066	0.9948	0.6214	0.6432	0.6680	0.7105	
10	0.6285	0.7592	0.8366	0.6199	0.5470	0.9948	0.6214	0.6432	0.8292	0.7200	
11	0.5800	0.7545	0.8366	0.5951	0.5192	0.9948	0.6214	0.6681	0.9359	0.7228	
12	0.5470	0.7745	0.8366	0.5929	0.5569	0.9944	0.6214	0.6726	0.9697	0.7295	
13	0.5470	0.7889	0.8685	0.5926	0.6066	0.9758	0.6214	0.6432	0.9695	0.7348	
14	0.6594	0.7883	0.8793	0.5926	0.6066	0.9462	0.6192	0.6666	0.9677	0.7473	
15	0.8768	0.7880	0.8371	0.5926	0.6066	0.9458	0.5718	0.7735	0.9482	0.7712	
16	0.9554	0.8231	0.8366	0.5930	0.5698	0.9515	0.5942	0.7682	0.9236	0.7795	
17	0.9573	0.8693	0.8366	0.6504	0.4065	0.9583	0.6214	0.6597	0.9311	0.7656	
18	0.9573	0.8767	0.8366	0.6747	0.3310	0.9770	0.6214	0.6432	0.9422	0.7622	
19	0.9573	0.8769	0.8366	0.6094	0.3072	0.9844	0.6214	0.6432	0.9427	0.7532	
20	0.9573	0.8769	0.8366	0.6101	0.4950	0.9844	0.6214	0.6432	0.9427	0.7742	
21	0.9388	0.8769	0.8366	0.6665	0.6066	0.9829	0.6214	0.6432	0.9572	0.7922	
22	0.8545	0.8919	0.8366	0.6933	0.6066	0.8821	0.6214	0.6432	0.9559	0.7762	
23	0.8162	0.8939	0.8366	0.6863	0.6066	0.7813	0.6214	0.6432	0.7720	0.7397	
24	0.8162	0.8466	0.8366	0.6635	0.6066	0.7813	0.6143	0.6432	0.6066	0.7128	
25	0.8162	0.8303	0.8366	0.6207	0.5107	0.7813	0.5214	0.6432	0.5954	0.6840	
26	0.8162	0.8565	0.8366	0.5927	0.3185	0.7813	0.3886	0.6432	0.5739	0.6453	
27	0.8162	0.8581	0.8366	0.5926	0.3153	0.7800	0.3215	0.6432	0.5268	0.6323	
28	0.8251	0.8649	0.8366	0.5926	0.3817	0.7801	0.3776	0.6432	0.4359	0.6375	

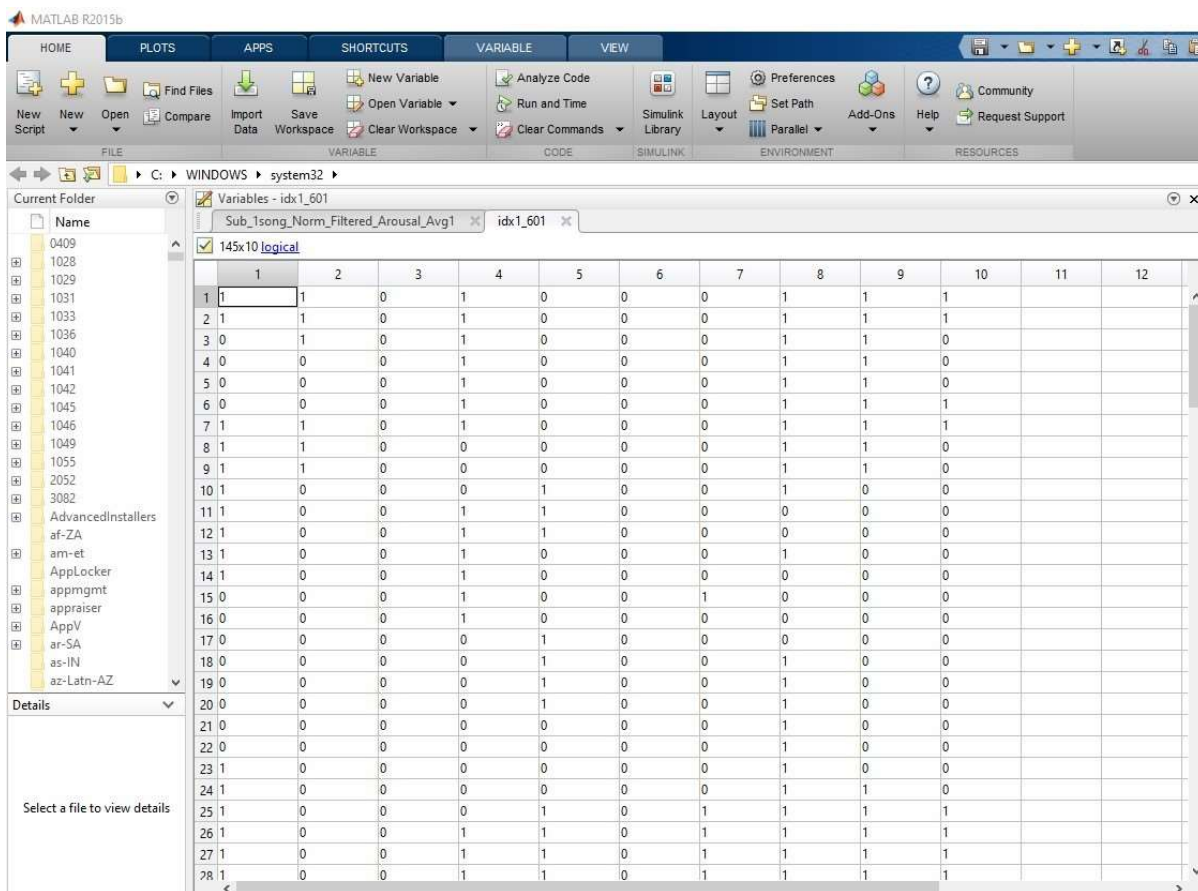
Εικόνα 17: Η δημιουργία δέκατης στήλης με τις μέσες τιμές της διέγερσης κατά την πρώτη υπόθεση

Στην Εικόνα 18, φαίνεται τόσο η εντολή που δόθηκε για την εύρεση εκείνων των τιμών της μετρικής Arousal που είναι μικρότερες από 0,8259, αλλά και το αποτέλεσμα που προέκυψε κατά τον αντίστοιχο έλεγχο για τις εκάστοτε μέσες τιμές. Σημειώνεται πως για τον έλεγχο θετικής ή αρνητικής διέγερσης έγινε το εξής: όσες τιμές είναι μικρότερες από την  $\chi$ -μέση τιμή, λαμβάνουν την Boolean τιμή 1 (true) και 0 όσες είναι μεγαλύτερες ή ίσες. Φυσικά η ίδια διαδικασία έγινε και για την δεύτερη μετρική Valence, ώστε τελικά να δοθούν όλες οι ετικέτες και να εισαχθούν στο αρχείο Arff της openSMILE.



Εικόνα 18: Η εντολή για έλεγχο θετικών και αρνητικών ετικετών και η επιστροφή Boolean τιμών

Στην Εικόνα 19 φαίνονται οι Boolean τιμές που προέκυψαν κατά τον έλεγχο θετικών και αρνητικών ετικετών της μετρικής Arousal στην δεύτερη υπόθεση με βάση τις μέσες τιμές στα αντίστοιχα σύνολο 60-40 που χρησιμοποιήθηκαν για κάθε χρήστη ξεχωριστά. Η δέκατη στήλη προέκυψε από την πλειοψηφία των πρόσημων-ετικετών, δηλαδή όπου 1 δίνει αρνητική διέγερση γιατί παραπάνω από πέντε χρήστες είχαν αρνητική ετικετα-πρόσημο διέγερσης και όπου 0 θετική διέγερση.



Εικόνα 19: Η δημιουργία ετικετών της μετρικής Arousal κατά την δεύτερη υπόθεση

## Παράρτημα II- WEKA

Στις Εικόνες 20 και 21 φαίνονται ορισμένα χαρακτηριστικά με τις τιμές τους, ενώ η 1581<sup>η</sup> στήλη περιλαμβάνει τις κλάσεις, θετική-αρνητική διέγερση και συναισθήματα 1<sup>ο</sup> - 4<sup>ο</sup> τεταρτημορίου αντίστοιχα, όπως προσεγγίστηκαν στην τρίτη υπόθεση.

Local_sma_de_iqr1-3	1578: shimmerLocal_sma_de_percentile99.0	1579: shimmerLocal_sma_de_upleveltime75	1580: shimmerLocal_sma_de_upleveltime90	1581: class
Numeric	Numeric	Numeric	Numeric	Nominal
0.092523	0.050467	0.5	0.333333	neg_ar
0.158976	0.119396	0.25	0.25	neg_ar
0.0	0.0	0.0	0.0	pos_ar
0.073661	0.117392	0.129032	0.032258	pos_ar
0.059037	0.102178	0.121622	0.013514	pos_ar
0.02669	0.02015	0.25	0.25	pos_ar
0.064261	0.173795	0.076923	0.019231	pos_ar
0.045023	0.143667	0.032258	0.008065	pos_ar
0.05363	0.185	0.043478	0.043478	pos_ar
0.094378	0.192333	0.083333	0.083333	pos_ar
0.0	0.0	0.0	0.0	pos_ar
0.072804	0.058537	0.3	0.1	pos_ar
0.072804	0.058537	0.3	0.1	pos_ar
0.0	0.0	0.0	0.0	pos_ar
0.082365	0.069775	0.25	0.25	pos_ar
0.082365	0.069775	0.25	0.25	pos_ar
0.03767	0.102446	0.125	0.083333	pos_ar
0.065705	0.103866	0.1875	0.125	pos_ar
0.0	0.0	0.0	0.0	pos_ar
0.0	0.0	0.0	0.0	pos_ar
0.0	0.0	0.0	0.0	pos_ar
0.163719	0.138693	0.25	0.25	pos_ar
0.0	0.0	0.0	0.0	pos_ar
0.0	0.0	0.0	0.0	pos_ar
0.010841	0.009184	0.25	0.25	neg_ar
0.066189	0.1673	0.084507	0.014085	neg_ar
0.065025	0.19864	0.040816	0.006803	neg_ar
0.075499	0.190492	0.035398	0.00885	neg_ar
0.067400	0.134444	0.076923	0.005644	neg_ar

Εικόνα 20: Η προσθήκη των κλάσεων Arousal στο αρχείο

Local_sma_de_iqr1-3	1578: shimmerLocal_sma_de_percentile99.0	1579: shimmerLocal_sma_de_upleveltime75	1580: shimmerLocal_sma_de_upleveltime90	1581: class3
Numeric	Numeric	Numeric	Numeric	Nominal
0.092523	0.050467	0.5	0.333333	em4
0.158976	0.119396	0.25	0.25	em4
0.0	0.0	0.0	0.0	em1
0.073661	0.117392	0.129032	0.032258	em1
0.059037	0.102178	0.121622	0.013514	em1
0.02669	0.02015	0.25	0.25	em1
0.064261	0.173795	0.076923	0.019231	em1
0.045023	0.143667	0.032258	0.008065	em1
0.05363	0.185	0.043478	0.043478	em1
0.094378	0.192333	0.083333	0.083333	em2
0.0	0.0	0.0	0.0	em2
0.072804	0.058537	0.3	0.1	em2
0.072804	0.058537	0.3	0.1	em2
0.0	0.0	0.0	0.0	em1
0.082365	0.069775	0.25	0.25	em1
0.082365	0.069775	0.25	0.25	em2
0.03767	0.102446	0.125	0.083333	em2
0.065705	0.103866	0.1875	0.125	em2
0.0	0.0	0.0	0.0	em2
0.0	0.0	0.0	0.0	em2
0.0	0.0	0.0	0.0	em2
0.163719	0.138693	0.25	0.25	em2
0.0	0.0	0.0	0.0	em2
0.0	0.0	0.0	0.0	em2
0.010841	0.009184	0.25	0.25	em3
0.066189	0.1673	0.084507	0.014085	em3
0.065025	0.19864	0.040816	0.006803	em3
0.075499	0.190492	0.035398	0.00885	em3
0.067400	0.134444	0.076923	0.005644	em3

Εικόνα 21: Η προσθήκη ετικετών συναισθήματος κατά την τρίτη υπόθεση