



ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΚΑΙ
ΕΠΙΚΟΙΝΩΝΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΜΕ
ΕΞΕΙΔΙΚΕΥΣΗ ΣΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ ΓΙΑ ΑΠΟΦΟΙΤΟΥΣ
ΑΛΛΩΝ ΣΧΟΛΩΝ

ΕΡΓΑΣΙΑ ΜΕ ΘΕΜΑ
**ΜΕΛΕΤΗ ΕΠΙΡΡΟΗΣ ΚΟΙΝΩΝΙΚΩΝ
ΔΕΔΟΜΕΝΩΝ ΣΤΗΝ ΤΕΧΝΙΚΗ ΑΝΑΛΥΣΗ
ΧΡΗΜΑΤΙΣΤΗΡΙΑΚΩΝ ΔΕΙΚΤΩΝ ΜΕ
ΜΕΘΟΔΟΥΣ ΕΠΕΞΕΡΓΑΣΙΑΣ ΧΡΟΝΟΣΕΙΡΩΝ**

ΧΥΔΕΡΙΩΤΟΥ ΣΤΕΛΛΑ
ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ
ΜΑΡΑΓΚΟΥΔΑΚΗΣ Ε.

ΑΠΡΙΛΙΟΣ 2018
ΣΑΜΟΣ

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω θερμά τον κ. Μαραγκουδάκη που με εμπιστεύτηκε για την εργασία και μου παρείχε τις πολύτιμες γνώσεις του και την βοήθειά του σε οποιοδήποτε πρόβλημα αντιμετώπισα. Η καθοδήγησή του υπήρξε καταπληκτική και η εν γένει συνεργασία μας καθ' όλη την διάρκεια εκπόνησης της υπήρξε άψογη.

Κλείνοντας, θα ήθελα να ευχαριστήσω τους γονείς μου και την αδερφή μου για την υποστήριξη και την εμπύχωση όλα αυτά τα χρόνια. Είναι βέβαιο ότι η συμπαράσταση τους υπήρξε πολύ σημαντική σ' αυτό το κομμάτι των σπουδών μου.

Πίνακας περιεχομένων

ΕΥΧΑΡΙΣΤΙΕΣ	2
ΠΕΡΙΛΗΨΗ	5
SUMMARY	6
ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ.....	7
Κατάλογος διαγραμμάτων.....	8
Κατάλογος Πινάκων.....	12
Καταλογος Εικόνων	16
Εισαγωγή.....	17
1. Παλαιότερες ενασχολήσεις.....	20
2.Θεωρητικό υπόβαθρο.....	22
2.1 Βασικά Χαρακτηριστικά Χρονοσειράς.....	22
2.2 Μέθοδοι απαλοιφής της τάσης.....	25
3. Μοντελοποίηση με χρήση ARIMA.....	31
3.1 Πρόβλεψη της τιμής κλεισίματος της εταιρείας Facebook	31
3.2 Πρόβλεψη της τιμής κλεισίματος της εταιρείας Google	39
3.3 Πρόβλεψη της τιμής κλεισίματος της εταιρείας MSFT	44
3.4 Πρόβλεψη της τιμής κλεισίματος της εταιρείας BA	49
3.5 Πρόβλεψη της τιμής κλεισίματος της εταιρείας LMT	54
3.6 Πρόβλεψη της τιμής κλεισίματος της εταιρείας MMM.....	59
3.7 Πρόβλεψη της τιμής κλεισίματος της εταιρείας BLK.....	63
3.8 Πρόβλεψη της τιμής κλεισίματος της εταιρείας JPM	67
3.9 Πρόβλεψη της τιμής κλεισίματος της εταιρείας GS	71
4.Νευρωνικά δίκτυα	79
4.1 Ορισμός των νευρωνικών δικτύων.....	79
5. Μοντελοποίηση με χρήση νευρωνικών δικτύων.....	86
5.1 Πρόβλεψη της τιμής κλεισίματος της εταιρείας BA	86
5.2 Πρόβλεψη της τιμής κλεισίματος της εταιρείας BLK.....	91
5.3 Πρόβλεψη της τιμής κλεισίματος της εταιρείας FB.....	97
5.4 Πρόβλεψη της τιμής κλεισίματος της εταιρείας GOOGLE	102

5.5 Πρόβλεψη της τιμής κλεισίματος της εταιρείας GS	107
5.6 Πρόβλεψη της τιμής κλεισίματος της εταιρείας JMP	112
5.7 Πρόβλεψη της τιμής κλεισίματος της εταιρείας MMM.....	117
5.8 Πρόβλεψη της τιμής κλεισίματος της εταιρείας LMT	122
5.9 Πρόβλεψη της τιμής κλεισίματος της εταιρείας MSFT	128
ΣΥΝΟΨΗ.....	133
6. ΣΥΜΠΕΡΑΣΜΑΤΑ.....	135
ΠΑΡΑΡΤΗΜΑ (Κώδικας για μία εταιρεία)	140
• ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΜΕ ΜΕΘΟΔΟ ARIMA ΣΕ ΓΛΩΣΣΑ R.....	140
• ΚΩΔΙΚΑΣ ΓΙΑ ΜΙΑ ΕΤΑΙΡΕΙΑ ΜΕ ΜΕΘΟΔΟ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΣΕ ΠΡΟΓΡΑΜΜΑ RAPIDMINER.....	144

ΠΕΡΙΛΗΨΗ

Η παρούσα μελέτη ασχολήθηκε με την επιρροή των κοινωνικών δεδομένων στην τεχνική ανάλυση χρηματιστηριακών δεικτών με μεθόδους επεξεργασίας χρονοσειρών και ολοκληρώθηκε μέσα από επτά ενότητες όπου πιο συγκεκριμένα, η πρώτη ενότητα αναφέρθηκε στο εισαγωγικό σημείωμα. Συγκεκριμένα, η στοχαστικότητα δε μπορεί να αγνοηθεί όταν έχουμε πραγματικές μετρήσεις και μελετάμε πραγματικές διαδικασίες καθώς υπάρχει τυχαιότητα και στην εξέλιξη της διαδικασίας, φαινόμενο που ονομάζεται δυναμικός θόρυβος. Επίσης υπάρχει τυχαιότητα και στην μέτρηση του μεγέθους ενδιαφέροντος και ονομάζεται θόρυβος μέτρησης.

Η δεύτερη ενότητα αναφέρθηκε σε παλαιότερες ενασχολήσεις με παρόμοιες μελέτες. Στη συνέχεια, η τρίτη ενότητα ανέλυσε το θεωρητικό υπόβαθρο εμβαθύνοντας στα βασικά χαρακτηριστικά χρονοσειράς και στις μεθόδους απαλοιφής της τάσης αλλά και στη μέθοδο των πρώτων διαφορών και διαφορών p τάξης .

Η τέταρτη ενότητα έκανε λόγο για τη μοντελοποίηση με τη μέθοδο ARIMA, ενώ η πέμπτη ενότητα εστίασε στον ορισμό των νευρωνικών δικτύων. Πιο αναλυτικά τα τεχνητά νευρωνικά δίκτυα είναι συσκευές επεξεργασίας (αλγόριθμοι ή πραγματικό υλικό) που διαμορφώνονται σύμφωνα με την νευρωνική δομή του εγκεφαλικού φλοιού των θηλαστικών, αλλά σε πολύ μικρότερες κλίμακες. Το παγκόσμιο ελάχιστο είναι αυτή η θεωρητική λύση με το μικρότερο δυνατό λάθος. Τα νευρωνικά δίκτυα είναι καθολικά προσεγγιστικά και λειτουργούν καλύτερα αν το σύστημα στο οποίο χρησιμοποιούνται για την μοντελοποίηση έχει υψηλή ανοχή στο σφάλμα.

Η έκτη ενότητα αφορά τη μοντελοποίηση της χρονοσειράς με νευρωνικά δίκτυα και η τελευταία κατέληξε με τα συμπεράσματα της σύγκρισης των δύο μεθοδών.

Λέξεις Κλειδιά: χρονοσειρά, προβλέψεις, τεχνική ανάλυση, τάση, στοχαστικότητα, ARIMA, MA, AR, νευρωνικά δίκτυα, θόρυβος, σφάλμα.

SUMMARY

This study dealt with the influence of social data on the technical analysis of stock indices with time series processing methods and was completed through seven modules, where the first section was mentioned in the introductory note. In particular, the stochasticity can not be ignored when we have actual measurements and we are studying real processes as there is randomness in the evolution of the process called phenomenon called dynamic noise. There is also randomness in measuring the magnitude of interest and is called measurement noise.

The second section referred to previous studies with similar studies. Subsequently, the third module analyzed the theoretical background by deepening the basic time series and trend-elimination methods, as well as the method of the first differences and p differences.

The fourth module spoke about ARIMA modeling, while the fifth module focused on the definition of neural networks. More specifically, artificial neural networks are processing devices (algorithms or real material) formulated according to the neural structure of the mammalian brain cortex, but on much smaller scales. The world minimum is this theoretical solution with the slightest possible error. Neural networks are universally approximate and work best if the system they are used for modeling has high fault tolerance.

The seventh module concerns the modeling of the time series with neural networks and the latter concluded with the conclusions of the comparison of the two methods.

Key words: time series, predictions, technical analysis, trend, stochasticity, READY, MA, AR, neural networks, noise, error.

ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

ANNs	Artificial neural network (νευρωνικά δίκτυα)
BA	The Boeing Company
LMT	Lockheed Martin Corporation
MMM	3M Company
JMP	JMP Group LLC
GS	The Goldman Sachs Group
BLK	BlackRock
FB	Facebook,
MSFT	Microsoft Corporation
ARIMA	Autoregressive Integrated Moving Average Model
MA	Moving Average
AR	Autoregression

Κατάλογος διαγραμμάτων

A/A	ΤΙΤΛΟΣ	Σελίδα
1.	Χρονοσειρά με τάση.....	22
2.	Χρονοσειρά με εποχικότητα.....	23
3.	Χρονοσειρά με τάση και εποχικότητα.....	24
4.	Χρονοσειρά χωρίς τάση και εποχικότητα.....	24
5.	Χρονοσειρά του δείκτη GNP των Ηνωμένων Πολιτειών απαλλαγμένη από την τάση με χρήση πρώτων διαφορών.....	26
6.	Χρονοσειρά με τάση πολωνύμου β βαθμού, χρονοσειρά μετά από χρήση πρώτων διαφορών, χρονοσειρά μετά από χρήση διαφορών δεύτερης τάξης.....	27
7.	Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας Facebook.....	31
8.	Χρονοσειρά με τάση της εταιρείας Facebook κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων.....	32
9.	Χρονοσειρά χωρίς τάση της εταιρείας Facebook	33
10.	Γράφημα αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία Facebook	34
11.	Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία Facebook.....	35
12.	Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας Google	39
13.	Χρονοσειρά με τάση της εταιρείας Google κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων.....	40
14.	Χρονοσειρά χωρίς τάση της εταιρείας Google.....	40
15.	Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία Google.....	41
16.	Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας MSFT.....	45
17.	Χρονοσειρά με τάση της εταιρείας MSFT κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων.....	45
18.	Χρονοσειρά χωρίς τάση της εταιρείας MSFT.....	46
19.	Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία MSFT	47
20.	Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας BA	49
21.	Χρονοσειρά με τάση της εταιρείας BA κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων.....	50

22. Χρονοσειρά χωρίς τάση της εταιρείας BA.....	50
23. Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία BA....	51
24. Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας LMT.....	55
25. Χρονοσειρά με τάση της εταιρείας LMT κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων.....	55
26. Χρονοσειρά χωρίς τάση της εταιρείας LMT.....	56
27. Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία LMT.....	57
28. Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας MMM.....	59
29. Χρονοσειρά με τάση της εταιρείας MMM κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων.....	59
30. Χρονοσειρά χωρίς τάση της εταιρείας MMM.....	60
31. Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία MMM	61
32. Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας BLK.....	63
33. Χρονοσειρά με τάση της εταιρείας BLK κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων.....	63
34. Χρονοσειρά χωρίς τάση της εταιρείας BLK.....	64
35. Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία BLK...	65
36. Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας JPM.....	67
37. Χρονοσειρά με τάση της εταιρείας JPM κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων...	68
38. Χρονοσειρά χωρίς τάση της εταιρείας JPM.....	68
39. Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία JPM...	69
40. Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας GS.....	72
41. Χρονοσειρά με τάση της εταιρείας GS κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων.....	72
42. Χρονοσειρά χωρίς τάση της εταιρείας GS.....	73
43. Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία GS....	74
44. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία BA.....	86
45. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία BA.....	88
46. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14	89

για την εταιρεία BA.....	
47. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία BLK.....	91
48. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία BLK.....	93
49. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία BLK.....	95
50. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία Facebook.....	97
51. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία Facebook.....	99
52. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία Facebook.....	101
53. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία Google.....	103
54. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία Google.....	104
55. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία Google.....	106
56. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία GS.....	108
57. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία GS.....	109
58. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία GS.....	111
59. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία JMP.....	113
60. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία JMP.....	114
61. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14.. για την εταιρεία JMP.....	116
62. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία MMM.....	117
63. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία MMM.....	119
64. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14	120

για την εταιρεία MMM.....	
65. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία LMT.....	122
66. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία LMT.....	124
67. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία LMT.....	126
68. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία MSFT.....	128
69. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία MSFT.....	130
70. Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία MSFT.....	131

Κατάλογος Πινάκων

ΠΙΝΑΚΑΣ	ΤΙΤΛΟΣ	Σελίδα
1.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία FB για την πρόβλεψη της επόμενης τιμής με υστέρηση 3	36
2.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία FB για την πρόβλεψη της επόμενης τιμής με υστέρηση 7	36
3.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία FB για την πρόβλεψη της επόμενης τιμής με υστέρηση 14	36
4.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία Google για την πρόβλεψη της επόμενης τιμής με υστέρηση 3	42
5.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία Google για την πρόβλεψη της επόμενης τιμής με υστέρηση 7	42
6.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία Google για την πρόβλεψη της επόμενης τιμής με υστέρηση 14	42
7.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία MSFT για την πρόβλεψη της επόμενης τιμής με υστέρηση 3	48
8.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία MSFT για την πρόβλεψη της επόμενης τιμής με υστέρηση 7	48
9.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία MSFT για την πρόβλεψη της επόμενης τιμής με υστέρηση 14	48
10.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BA για την πρόβλεψη της επόμενης τιμής με υστέρηση 3	52
11.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BA για την πρόβλεψη της επόμενης τιμής με υστέρηση 7	52
12.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BA για την πρόβλεψη της επόμενης τιμής με υστέρηση 14	52
13.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία LMT για την πρόβλεψη της επόμενης τιμής με υστέρηση 3	57
14.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία LMT για την πρόβλεψη της επόμενης τιμής με υστέρηση 7	58
15.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία LMT	58

	για την πρόβλεψη της επόμενης τιμής με υστέρηση 1	
16.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία MMM για την πρόβλεψη της επόμενης τιμής με υστέρηση 3	62
17.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία MMM για την πρόβλεψη της επόμενης τιμής με υστέρηση 7	62
18.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία MMM για την πρόβλεψη της επόμενης τιμής με υστέρηση 14	62
19.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BLK για την πρόβλεψη της επόμενης τιμής με υστέρηση 3	66
20.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BLK για την πρόβλεψη της επόμενης τιμής με υστέρηση 7	66
21.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BLK για την πρόβλεψη της επόμενης τιμής με υστέρηση 14	66
22.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία JMP για την πρόβλεψη της επόμενης τιμής με υστέρηση 3	70
23.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία JMP για την πρόβλεψη της επόμενης τιμής με υστέρηση 7	70
24.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία JMP για την πρόβλεψη της επόμενης τιμής με υστέρηση 14	71
25.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία GS για την πρόβλεψη της επόμενης τιμής με υστέρηση 3	75
26.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία GS για την πρόβλεψη της επόμενης τιμής με υστέρηση 7	75
27.	Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία GS για την πρόβλεψη της επόμενης τιμής με υστέρηση 14	75
28.	Αποτελέσματα RMSE με υστέρηση 3,7 και 14 για κάθε εταιρεία με χρήση μεθόδου ARIMA	77
29.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία BA με χρήση νευρωνικών δικτύων με υστέρηση 3	87
30.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία BA με χρήση νευρωνικών δικτύων με υστέρηση 7	89
31.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία BA με χρήση νευρωνικών δικτύων με υστέρηση 14	90

32.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία BLK με χρήση νευρωνικών δικτύων με υστέρηση 3	92
33.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία BLK με χρήση νευρωνικών δικτύων με υστέρηση 7	94
34.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία BLK με χρήση νευρωνικών δικτύων με υστέρηση 14	96
35.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία Facebook με χρήση νευρωνικών δικτύων με υστέρηση 3	98
36.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία Facebook με χρήση νευρωνικών δικτύων με υστέρηση 7	100
37.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία Facebook με χρήση νευρωνικών δικτύων με υστέρηση 14	102
38.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία Google με χρήση νευρωνικών δικτύων με υστέρηση 3	104
39.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία Google με χρήση νευρωνικών δικτύων με υστέρηση 7	105
40.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία Google με χρήση νευρωνικών δικτύων με υστέρηση 14	107
41.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία GS με χρήση νευρωνικών δικτύων με υστέρηση 3	109
42.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία GS με χρήση νευρωνικών δικτύων με υστέρηση 7	110
43.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία GS με χρήση νευρωνικών δικτύων με υστέρηση 14	112
44.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία JMP με χρήση νευρωνικών δικτύων με υστέρηση 3	114
45.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία JMP με χρήση νευρωνικών δικτύων με υστέρηση 7	115
46.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία JMP με χρήση νευρωνικών δικτύων με υστέρηση 14	117
47.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία MMM με χρήση νευρωνικών δικτύων με υστέρηση 3	118
48.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία MMM με χρήση	120

	νευρωνικών δικτύων με υστέρηση 7	
49.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία MMM με χρήση νευρωνικών δικτύων με υστέρηση 14	121
50.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία LMT με χρήση νευρωνικών δικτύων με υστέρηση 3	123
51.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία LMT με χρήση νευρωνικών δικτύων με υστέρηση 7	125
52.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία LMT με χρήση νευρωνικών δικτύων με υστέρηση 14	127
53.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία MSFT με χρήση νευρωνικών δικτύων με υστέρηση 3	129
54.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία MSFT με χρήση νευρωνικών δικτύων με υστέρηση 7	131
55.	Πραγματική και προβλεπόμενη τιμή για την εταιρεία MSFT με χρήση νευρωνικών δικτύων με υστέρηση 14	132
56.	Αποτελέσματα RMSE με υστέρηση 3,7 και 14 για κάθε εταιρεία με χρήση Νευρωνικών δικτύων	133
57.	Σύγκριση αποτελεσμάτων RMSE με υστέρηση 3,7 και 14 για κάθε εταιρεία με χρήση μεθόδου ARIMA και Νευρωνικών δικτύων	134

Κατάλογος Εικόνων

ΕΙΚΟΝΑ	ΤΙΤΛΟΣ	ΣΕΛΙΔΑ
1.	Σχέδιο Νευρωνικού δικτύου	80
2.	'Μάθηση' νευρωνικού δικτύου με Κανόνα Δέλτα	81
3.	Η επιφάνεια σφάλματος της οπισθοδιάδοσης	82

Εισαγωγή

Η λήψη απόφασης για μία συναλλαγή είναι ένα θέμα συνεχούς καινοτομίας. Η ταχεία εξέλιξη της χρηματιστηριακής πραγματικότητας με την είσοδο νέων δυναμικών επιχειρήσεων στο χρηματιστήριο καθώς και η διεθνοποίηση των οικονομιών οδήγησε στην εφαρμογή σύγχρονων επενδυτικών εργαλείων αξιολόγησης των τιμών των μετοχών. Η Θεμελιώδης και η **τεχνική ανάλυση** είναι δύο τέτοια εργαλεία. Αποτελούν δύο συμπληρωματικές αλλά ταυτόχρονα και ανταγωνιστικές, μεθόδους λήψης επενδυτικών αποφάσεων, οι οποίες προσπαθούν να προβλέψουν με διαφορετικό τρόπο τη μελλοντική πορεία της χρηματιστηριακής αγοράς.

Η τεχνική ανάλυση με την οποία ασχολήθηκε η παρούσα μελέτη εμφανίστηκε στα τέλη του 19^{ου} αιώνα από το θεμελιωτή της μεθόδου Charles Dow. Πρόκειται ουσιαστικά για ένα εργαλείο που μελετά και ερευνά τους τρόπους εκείνους με τους οποίους μπορούμε να προβλέψουμε τις μελλοντικές χρηματιστηριακές τάσεις.

Πληθώρα φαινομένων χρήζουν κάποιας μελέτης, με σκοπό την αποκάλυψη των μηχανισμών κίνησης των παραμέτρων που τα ορίζουν. Διάφοροι κλάδοι των μαθηματικών συμβάλλουν σημαντικά στην εύρεση μιας σχέσης, η οποία περιγράφει τη διαδικασία εξέλιξης με όσο το δυνατόν μικρότερο σφάλμα, όπως επίσης και στην πρόβλεψη των τιμών του φαινομένου για μελλοντικούς χρόνους. Υπάρχουν μοντέλα τα οποία στηρίζονται στην παρατήρηση των σχετικών μεγεθών, που καλούνται **εμπειρικά**. Υπάρχουν επίσης μοντέλα που εκφράζουν τις σχέσεις μεταξύ των υπό μελέτη μεγεθών, τα οποία ονομάζονται **μοντέλα βασικών αρχών**. Αν εστιάσουμε στη μελέτη φαινομένων, που παράγουν δεδομένα με χρονολογική σειρά, τότε αναφερόμαστε **στον κλάδο της ανάλυσης χρονοσειρών**. Με τον όρο χρονοσειρά αναφερόμαστε σε ένα σύνολο δεδομένων, τα οποία συλλέγονται διαχρονικά και παρουσιάζουν την εξέλιξη που έχουν οι τιμές μιας μεταβλητής κατά τη διάρκεια ίσων διαδοχικών χρονικών περιόδων [15]. Πιο συγκεκριμένα, η χρονοσειρά αποτελείται από ένα σύνολο παρατηρήσεων, οι τιμές της οποίας λαμβάνονται σε ίσες χρονικές στιγμές ή περιόδους, π.χ. μέρα, μήνα, τρίμηνο εξάμηνο, έτος κ.ά.

Βασικός σκοπός του κλάδου αυτού, είναι η ανάδειξη της **στοχαστικής διαδικασίας ή του δυναμικού συστήματος** μέσω των οποίων παράγεται το υπό

μελέτη φαινόμενο, με τελικό σκοπό την πρόβλεψη. Στόχος των προβλέψεων είναι να είναι όσο το δυνατόν πιο εύστοχες και να ελαχιστοποιηθεί η απόκλισή τους από τις πραγματικές μελλοντικές τιμές που προσπαθούν να περιγράψουν. Έτσι, οι μέθοδοι πρόβλεψης έχουν καθοριστικό ρόλο στη σωστή και έγκαιρη λήψη αποφάσεων. Γίνεται φανερό πως αν τα δεδομένα που χρησιμοποιούμε στη διαδικασία της πρόβλεψης, είναι υψηλής ποιότητας, τότε μπορούμε να έχουμε ικανοποιητικά ακριβείς προβλέψεις. Στην πραγματικότητα όμως υπάρχουν παράγοντες που εισάγουν σφάλμα στις προβλέψεις. Ένας τέτοιος παράγοντας, είναι η αβεβαιότητα. Η αβεβαιότητα αναφέρεται σε απροσδόκητα γεγονότα που επηρέασαν την πραγματική μελλοντική πορεία της χρονοσειράς που εξετάζουμε. Το γεγονός αυτό έχει ως αποτέλεσμα τα διάφορα μοντέλα να εμφανίζουν από μικρή απόκλιση, έως και ολική αστοχία στις εκτιμήσεις τους. **Στοχαστικότητα**, ονομάζεται η τυχαιότητα με την οποία μεταβάλλονται οι τιμές του παρατηρούμενου μεγέθους και η ύπαρξή της, μας υποχρεώνει να θεωρούμε το παρατηρούμενο μέγεθος ως **τυχαία μεταβλητή** [16]. Η στοχαστικότητα δε μπορεί να αγνοηθεί όταν έχουμε πραγματικές μετρήσεις και μελετάμε πραγματικές διαδικασίες καθώς υπάρχει τυχαιότητα και στην εξέλιξη της διαδικασίας φαινόμενο που ονομάζεται **δυναμικός θόρυβος**. Επίσης υπάρχει τυχαιότητα και στην μέτρηση του μεγέθους ενδιαφέροντος και ονομάζεται **θόρυβος μέτρησης**.

Στόχος της παρούσας εργασίας, είναι η δημιουργία μοντέλων πρόβλεψης των τιμών κλεισίματος της επόμενης ημέρας, για 9 δείκτες μετοχών, χρησιμοποιώντας ως βάση, τις 3, 7 ή 14 προηγούμενες ημέρες (3 μοντέλα για κάθε περίπτωση). Πιο συγκεκριμένα, τα μοντέλα σε πρώτη φάση προσαρμόστηκαν στους δείκτες κλεισίματος των μετοχών των εταιρειών : 1) BA, 2) BLK, 3) Facebook, 4) Google, 5) GS, 6) JPM, 7) LMT, 8) MMM και 9) MSFT. Για κάθε μετοχή, γίνεται επίσης και μια κριτική ως προς την επιτυχία της προσαρμογής τους στα δεδομένα, χρησιμοποιώντας μέσα τετραγωνικά σφάλματα (mse ή rmse).

Τα δεδομένα για κάθε μία εταιρεία, εξήχθησαν από τη σελίδα <https://finance.yahoo.com/quote/fb/history?ltr=1> και περάστηκαν στο προγραμματιστικό περιβάλλον της R, στο οποίο έγινε και η παρακάτω ανάλυση.

Σε κάθε μεταβλητή υπάρχουν 6 στήλες: Date, Open, High, Low, Close, Adj.Close και Volume. Οι μεταβλητές που περιέχουν τις τιμές κλεισίματος του δείκτη της εκάστοτε

εταιρείας (που είναι οι τιμές τις οποίες στοχεύουμε να προβλέψουμε) για τα τελευταία 5 έτη (από 14/01/2013 έως 12/01/2018), μετασχηματίστηκαν ως κλάσεις χρονοσειρών και αποθηκεύτηκαν με σχετική ονοματολογία στην R.

Στη δεύτερη φάση τα δεδομένα μας περάστηκαν στο προγραμματιστικό περιβάλλον του Rapidminer στο οποίο έγινε όμοια ανάλυση και τα αποτελέσματα αποθηκεύτηκαν με σχετική ονοματολογία του Rapidminer.

Η χρησιμότητα-σημαντικότητα της επιτυχίας τέτοιου είδους μοντέλων, είναι σαφής. Η πρόβλεψη, αν είναι επιτυχής, μπορεί να χρησιμοποιηθεί τόσο για πρόληψη, όσο και για δραστηριοποίηση από μέρους των επενδυτών. Η επανάπαυση σε μία επιρρεπή πρόβλεψη, μπορεί να στοιχίσει χιλιάδες ή και εκατομμύρια ευρώ, ενώ η εύρεση ευνοϊκών χρονικών στιγμών ή περιόδων, μπορεί να αποφέρει μεγάλα κέρδη.

1. Παλαιότερες ενασχολήσεις

Η πρόβλεψη μετοχών έχει απασχολήσει πάρα πολύ τη μαθηματική και την οικονομική επιστήμη. Παρακάτω παραθέτουμε προηγούμενες εργασίες πάνω στο συγκεκριμένο αντικείμενο:

- Shakira Green-Georgia Southern University 2011: “Time Series Analysis of Stock Prices Using the BoxJenkins Approach”. Σε αυτό το σύγγραμμά, έχει γίνει εφαρμογή των ίδιων μοντέλων που χρησιμοποιήσαμε κι εμείς στην ανάλυσή μας, για τον ίδιο λόγο (προσαρμογή σε δεδομένα μετοχών) [17].
- Uma Devi, D.Sundar and Dr Alli 2013: “An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50”. Στο παρόν σύγγραμμα, γίνεται πολύ καλή θεμελίωση όλων των εισαγωγικών και διαδικαστικών θεμάτων που αφορούν τα ίδια μοντέλα. Το έργο καταλήγει στην εφαρμογή των μοντέλων για κατανόηση συσχετίσεων μεταξύ των τιμών κάποιων μετοχών και του κεφαλαίου των αγορών [18].
- Mahantesh Angadi και Amogh Kulkarni 2015: “Time Series Data Analysis for Stock Market Prediction using Data Mining Techniques with R. Στο παρόν σύγγραμμα γίνεται και πάλι εφαρμογή μοντέλων χρονοσειρών και μάλιστα με χρήση της γλώσσας R. Επιπλέον, εισάγονται στοιχεία από θεωρία εξόρυξης δεδομένων [19].
- Abhinav Pathak 2014 “Predictive time series analysis of stock prices using neural network classifier”. Εδώ παρατίθεται μια εργασία η οποία εφαρμόζει μοντέλα χρονοσειρών με χρήση νευρωνικών δικτύων.
- Μια ομάδα ερευνητών, Sadi Evren Seker, Cihan Mert, Khaled Al-Naami, Nuri Ozalp, Ugur Ayan. 2014 “Time series analysis on stock market for text mining correlation of economy news”. Αυτή η εργασία, εκτελεί μοντελοποίηση με χρονοσειρές, προκειμένου να μελετηθούν συσχετίσεις που υπάρχουν μεταξύ των οικονομικών νέων (μέσω εφημερίδων κλπ) και των τιμών κλεισίματος μετοχών. [9]
- Wen Cao, Ying Han, and Sarah S. Lam Department of Systems Science and Industrial Engineering State University of New York at Binghamton

Binghamton, NY 13902 2013 “Short-Term Stock Price Trend Prediction of an Emerging Market Using Neural Networks”. Η έρευνα αυτή προτείνει μια προσέγγιση νευρωνικού δικτύου για τη βραχυπρόθεσμη τάση των τιμών των μετοχών πρόβλεψη της εταιρείας εξόρυξης άνθρακα.

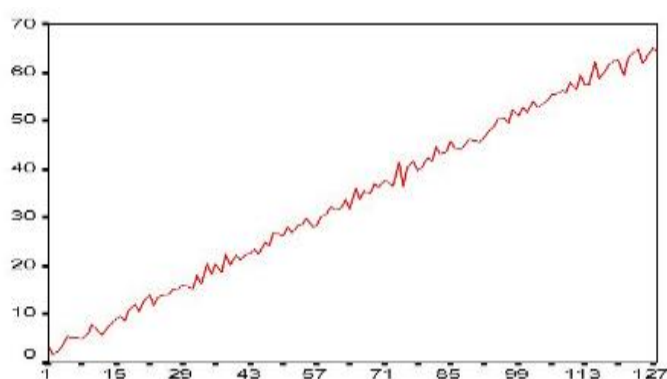
- Fahim Faisal *May 2012* “Forecasting Bangladesh's Inflation Using Time Series ARIMA Models”. Αυτή η έρευνα μελετά την πρόβλεψη του δείκτη τιμών ενός καταναλωτή. Συντάσσεται μια δομή για το μοντέλο πρόβλεψης ARIMA. Η ισχύς του μοντέλου δοκιμάστηκε χρησιμοποιώντας τυποποιημένες στατιστικές τεχνικές και το καλύτερο μοντέλο προτείνεται με βάση διάφορα κριτήρια διάγνωσης και επιλογής και αξιολόγησης [20].
- Hemanth Kumar P.S. Basavaraj Patil “Estimation & forecasting of volatility using ARIMA, ARFIMA and Neural Network based techniques”. Στόχος του εγγράφου είναι η πρόβλεψη της μεταβλητότητας με μεγάλη ακρίβεια. Τεχνικές πρόβλεψης χρονοσειρών ARIMA, ARFIMA και τεχνικές που βασίζονται σε Νευρωνικό Δίκτυο χρησιμοποιούνται για την πρόβλεψη της μεταβλητότητας. Τα αποτελέσματα συγκρίνονται για να βρεθεί μια ακριβής μέθοδος εκτίμησης και πρόβλεψης. Η καλύτερη τεχνική πρόβλεψης επιλέγεται με σύγκριση των αποτελεσμάτων σφάλματος όλων των τεχνικών πρόγνωσης με παραμέτρους μέτρησης σφάλματος όπως ME, RMSE, MAE, MPE, MAPE, MASE και ACF1.

Υπάρχουν προφανώς πάρα πολλές αναφορές που μπορούν να γίνουν, παρ' όλα αυτά οι παραπάνω είναι τόσο αξιόλογες, όσο και εξαιρετικά πρόσφατες, γι' αυτό και επιλέχθηκαν.

2.Θεωρητικό υπόβαθρο

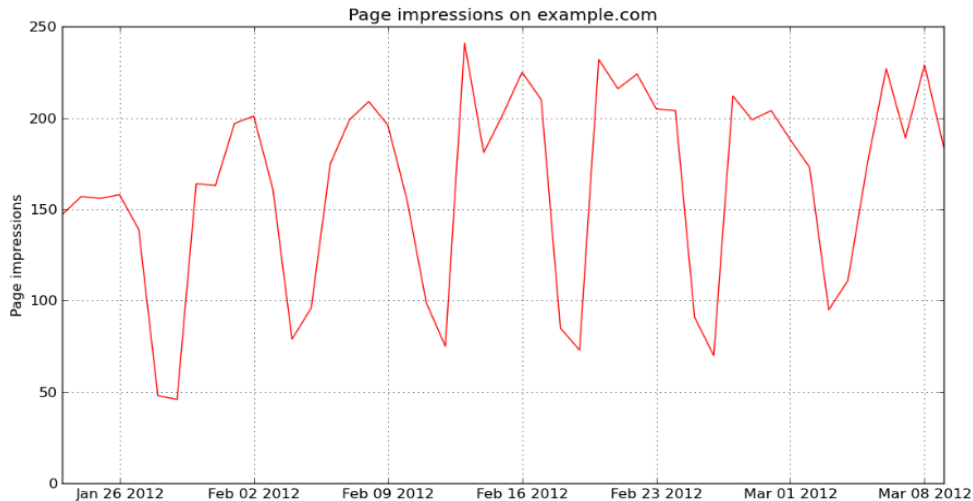
2.1 Βασικά Χαρακτηριστικά Χρονοσειράς

Όπως αναφέραμε στόχος της ανάλυσης χρονοσειρών είναι να μελετηθεί η **στοχαστική διαδικασία ή το δυναμικό σύστημα** που παράγει την ακολουθία δεδομένων. Προσπαθούμε δηλαδή να βρούμε τον «μηχανισμό» που παράγει την χρονοσειρά. Υπάρχουν όμως κάποια βασικά προβλήματα σε αυτήν την ανάλυση και το κυριότερο είναι η **έλλειψη στασιμότητας**. Αυτό το πρόβλημα σχετίζεται με την ύπαρξη εξωτερικών επιδράσεων ως προς το σύστημα που μελετάμε που ενδεχομένως προσθέτουν χαρακτηριστικά ξένα ως προς το σύστημα. Τα ξένα αυτά χαρακτηριστικά είναι η **τάση** και η **εποχικότητα**. Με τον όρο **τάση** αναφερόμαστε στις μεταβολές της χρονοσειράς, είναι δηλαδή η σταδιακή αύξηση ή μείωση της τιμής της χρονοσειράς με το χρόνο. Η τάση μπορεί να είναι **καθοριστική** ή **στοχαστική**. **Καθοριστική** ονομάζεται η τάση η οποία μπορεί να περιγραφεί από κάποια γνωστή ή εκτιμώμενη συνάρτηση του χρόνου. **Στοχαστική** ονομάζεται η τάση η οποία δε μπορεί να περιγραφεί από μία γνωστή συνάρτηση του χρόνου και γενικά δε μπορεί να περιγραφεί με κάποιο καθορισμένο τρόπο. Με τον όρο **εποχικότητα** αναφερόμαστε στην ιδιότητα της χρονοσειράς να επαναλαμβάνεται ανά συγκεκριμένο χρονικό διάστημα που είναι η **περίοδος** της χρονοσειράς. Παρακάτω δίνονται παραδείγματα χρονοσειρών όπου φαίνονται οι έννοιες που αναφέρθηκαν. Στο Γράφημα 1 έχουμε ένα παράδειγμα χρονοσειράς με ανοδική τάση ενώ στο Γράφημα 2 ένα παράδειγμα χρονοσειράς με εποχικότητα.



Γράφημα 1: Χρονοσειρά με τάση

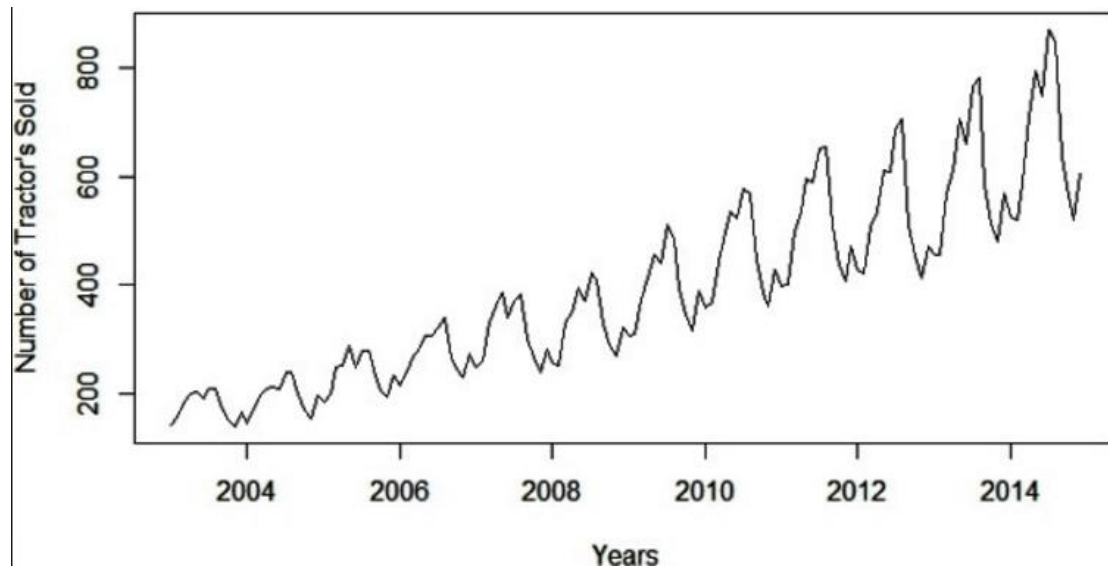
Είναι φανερό ότι η τιμή της χρονοσειράς διαρκώς αυξάνει και αυτό οφείλεται στην ύπαρξη τάσης, πιθανόν γραμμική και όχι τόσο στον «μηχανισμό» που παράγει την χρονοσειρά.



Γράφημα 2 : Χρονοσειρά με εποχικότητα

Στο Γράφημα 2 βλέπουμε τις εμφανίσεις μιας ανώνυμης σελίδας από τις 23 Ιανουαρίου του 2012 έως και τις 10 Μαρτίου του ίδιου χρόνου. Μπορούμε να παρατηρήσουμε στο συγκεκριμένο διάγραμμα, ότι υπάρχει κάποια επαναληπτικότητα. Πιο συγκεκριμένα μπορούμε να παρατηρήσουμε πως έχω μέγιστα και ελάχιστα τα οποία κατά κάποιο τρόπο ισαπέχουν. Όσον αφορά τα ελάχιστα βλέπουμε ότι 28 και 29 Ιανουαρίου είχα κάτω από 50 εμφανίσεις, που είναι μία κάτω κορυφή. Επίσης, 4 και 5 Φεβρουαρίου είχα επίσης κάτω κορυφή αν και οι τιμές ήταν αυξημένες σε σχέση με την προηγούμενη κάτω κορυφή. Όμοια, ελάχιστες εμφανίσεις έχουμε 11 και 12 Φεβρουαρίου, 17-18 Φεβρουαρίου, 24-25 Φεβρουαρίου και 2-3 Μαρτίου. Θα μπορούσαμε λοιπόν να πούμε, ότι η πορεία των εμφανίσεων της συγκεκριμένης σελίδας, έχει κάποια εβδομαδιαία επαναληπτικότητα καθώς ανά 7 μέρες ακολουθεί παρόμοια συμπεριφορά. Ομοίως αν παρατηρήσουμε τα μέγιστα θα δούμε ότι αυτά συναντώνται στις 26 Ιανουαρίου, 2 Φεβρουαρίου, 8-9 Φεβρουαρίου, 15 Φεβρουαρίου, 22 Φεβρουαρίου και 6 Μαρτίου και καταλήγουμε ότι η περίοδος της χρονοσειράς είναι ίση με 7 (1 εβδομάδα). Συνεπώς έχουμε εποχικότητα.

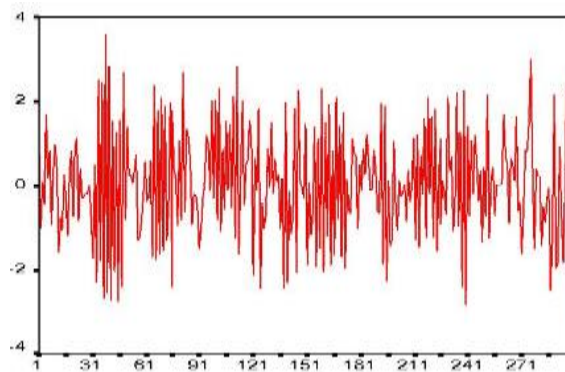
Παρακάτω στο Γράφημα 3 έχουμε ένα παράδειγμα χρονοσειράς που έχει τάση και εποχικότητα μαζί.



Γράφημα 3 : Χρονοσειρά με τάση και εποχικότητα

Παρατηρούμε ότι υπάρχει τάση η οποία είναι ανοδική και οδηγεί στην αύξηση των τιμών της χρονοσειράς ωστόσο υπάρχει και εποχικότητα δηλαδή συναντάμε περιόδους όπου η χρονοσειρά επαναλαμβάνεται. Συγκεκριμένα η περίοδος αυτή είναι 1 χρόνος αν παρατηρήσουμε τα ελάχιστα και τα μέγιστα συνεπώς έχουμε εποχικότητα.

Παρακάτω στο Γράφημα 4 έχουμε ένα παράδειγμα χρονοσειράς που δεν έχει ούτε τάση ούτε εποχικότητα.



Γράφημα 4 : Χρονοσειρά χωρίς τάση και εποχικότητα

Παρατηρούμε ότι δεν υπάρχει κάποια τάση ανοδική ή πτωτική αλλά ούτε και κάποια επαναληπτικότητα , συνεπώς συμπεραίνουμε ότι δεν έχουμε τάση και εποχικότητα.

Άλλα χαρακτηριστικά χρονοσειράς

Η **κυκλικότητα** και η **μη κανονικότητα** είναι δύο επιπλέον χαρακτηριστικά μιας χρονοσειράς. Η **κυκλικότητα** είναι το χαρακτηριστικό που εμφανίζει η χρονοσειρά να επαναλαμβάνεται χωρίς ωστόσο να είναι σταθερή η περίοδος και η μορφή της. Στην ουσία πρόκειται για έναν στοχαστικό κύκλο και διαφέρει από την **εποχικότητα** όπου είναι πλήρως προσδιορισμένη η περίοδος. Η **κυκλικότητα** αποδίδεται κυρίως στους οικονομικούς κύκλους, οι οποίοι οφείλονται στις μεταβαλλόμενες οικονομίες και τεχνολογίες. Η κυκλικότητα χρησιμοποιείται για την κατανόηση της δημιουργίας των παρατηρούμενων τιμών έως σήμερα και δεν συμπεριλαμβάνεται άμεσα στη διαδικασία σχηματισμού πρόβλεψης, σε αντίθεση με την τάση και την εποχικότητα. Στην πραγματικότητα η κυκλικότητα θεωρείται ότι βρίσκεται μέσα στον «μηχανισμό» που παράγει την χρονοσειρά απαλλαγμένο από τάση και εποχικότητα.

Η **μη κανονικότητα** είναι εκείνο το συστατικό στοιχείο της χρονοσειράς όπου επηρεάζει τις τιμές της χρονοσειράς κατά ένα τυχαίο και μη συστηματικό τρόπο, που δεν μπορεί να προσδιοριστεί. Έτσι λοιπόν γίνεται αντιληπτό πως η μη κανονικότητα δεν μπορεί να χρησιμοποιηθεί στην διαδικασία παραγωγής πρόβλεψης καθώς αναπαριστά όλους εκείνους τους τυχαίους και απρόσμενους παράγοντες που επηρεάζουν τις τιμές της χρονοσειράς, όπως πχ πόλεμοι, απρόσμενες αλλαγές στην νομοθεσία, φυσικές καταστροφές, τρομοκρατικές επιθέσεις, κ.α.

2.2 Μέθοδοι απαλοιφής της τάσης

Όπως αναφέρθηκε η ύπαρξη τάσης στη χρονοσειρά υποδηλώνει μη στασιμότητα και είναι κάτι που δεν θα θέλαμε, καθώς μας ενδιαφέρει κυρίως ο «μηχανισμός» που παράγει την χρονοσειρά απαλλαγμένος από την επίδραση κάποιας τάσης. Παρ' όλα αυτά, υπάρχουν μελέτες που ζητούν αποκλειστικά πληροφορία όσον αφορά την τάση και τότε περιοριζόμαστε απλά στην εκτίμηση της και θεωρούμε ότι η υπόλοιπη πληροφορία της χρονοσειράς δεν έχει καμία σημασία. Στα παρακάτω θα συμβολίζουμε με Y_t την αρχική χρονοσειρά και με X_t την χρονοσειρά απαλλαγμένη από την τάση. Υπάρχουν αρκετές μέθοδοι απαλοιφής, παρακάτω θα αναφέρουμε τη μέθοδο που χρησιμοποιήθηκε στην εργασία.

2.2.1 Η μέθοδος των πρώτων διαφορών και διαφορών p τάξης

Έστω ότι η χρονοσειρά Y_t είναι μη στάσιμη αποκλειστικά λόγω τάσης. Τότε ισχύει ότι $Y_t = \mu_t + X_t$ όπου μ_t είναι η τάση. Τότε μία μέθοδος απαλοιφής της τάσης είναι με χρήση πρώτων διαφορών. Θεωρούμε δηλαδή σαν νέα χρονοσειρά απαλλαγμένη από την τάση, την $X_t = Y_t - Y_{t-1}$.

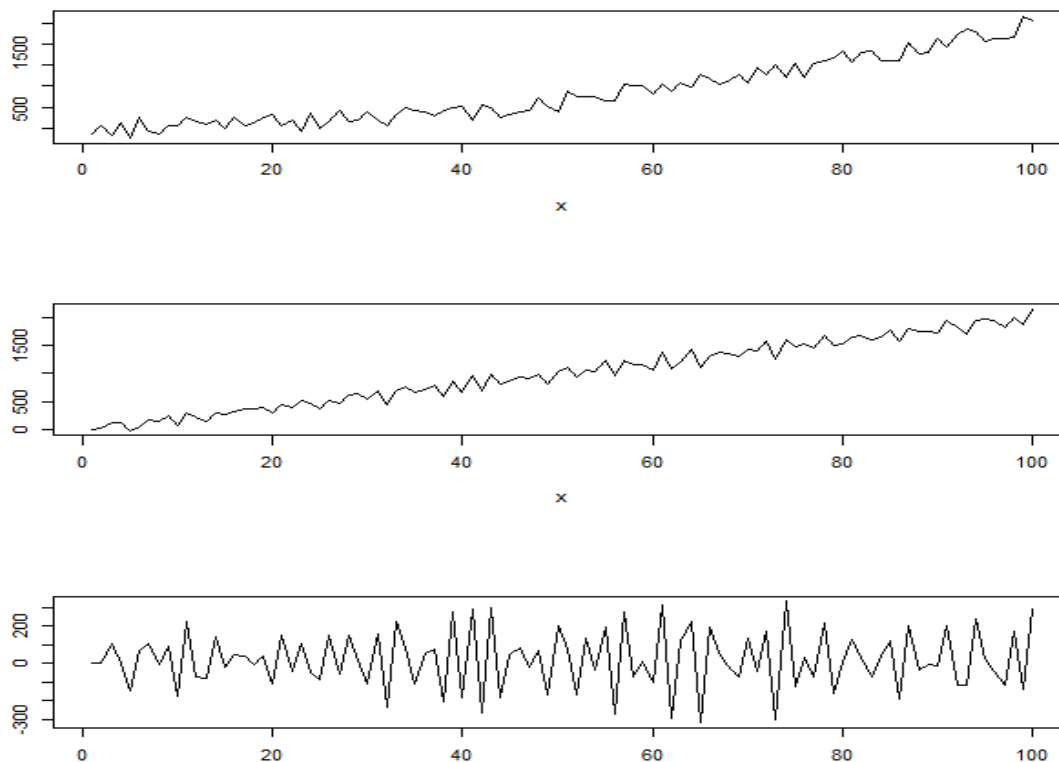
Στο Γράφημα 5 βλέπουμε την Χρονοσειρά του δείκτη GNP των Ηνωμένων Πολιτειών απαλλαγμένη από την τάση χρησιμοποιώντας αυτή τη φορά όχι την προσέγγιση πολυωνύμου αλλά την μέθοδο των πρώτων διαφορών.



Γράφημα 5 :Χρονοσειρά του δείκτη GNP των Ηνωμένων Πολιτειών απαλλαγμένη από την τάση με χρήση πρώτων διαφορών

Βλέπουμε πως δεν υπάρχει κάποια ανοδική ή πτωτική τάση. Γενικά αν η τάση είναι γραμμική τότε με χρήση πρώτων διαφορών θα απαλειφτεί. Αυτό αποδεικνύεται καθώς αν η τάση είναι γραμμική τότε θα ισχύει $\mu_t = a_0 + a_1t$ συνεπώς με τη χρήση πρώτων διαφορών θα έχουμε $Y_t - Y_{t-1} = \mu_t + X_t - \mu_{t-1} - X_{t-1} = \mu_t - \mu_{t-1} + X_t - X_{t-1}$. Όμως $\mu_t - \mu_{t-1} = a_0 + a_1t - a_0 - a_1(t-1) = a_1$, δηλαδή η τάση στη νέα χρονοσειρά είναι ανεξάρτητη του χρόνου, άρα έχει απαλειφτεί. Αν η τάση όμως δεν ήταν γραμμική τότε με πρώτες διαφορές δε θα ήταν ανεξάρτητη του χρόνου και ενδέχεται να μην είχε απαλειφτεί. Τότε θα έπρεπε να πάρουμε ξανά πρώτες διαφορές στη νέα χρονοσειρά, χρησιμοποιώντας ουσιαστικά την μέθοδο των διαφορών δεύτερης τάξης. Γενικά αν η τάση εκφράζεται τοπικά με πολυώνυμο βαθμού p τότε απαλείφεται με χρήση διαφορών p τάξης. Ωστόσο στις περισσότερες περιπτώσεις η τάση είναι γραμμική και απαλείφεται με διαφορές πρώτης τάξης. Στο Γράφημα 6

δίνεται μία χρονοσειρά η οποία φαίνεται να έχει τάση που να εκφράζεται από πολυώνυμο δευτέρου βαθμού, συνεπώς δεν απαλείφεται με διαφορές πρώτης τάξης αλλά με διαφορές δεύτερης τάξης.



Γράφημα 6: Χρονοσειρά με τάση πολυώνυμου β βαθμού, χρονοσειρά μετά από χρήση πρώτων διαφορών, χρονοσειρά μετά από χρήση διαφορών δεύτερης τάξης

2.2.2 Μοντέλο AR

Από τα πιο γνωστά μοντέλα είναι το μοντέλο AR(p) δηλαδή το αυτοπαλινδρομούμενο μοντέλο τάξης p. Τα αυτοπαλινδρομούμενα μοντέλα, προσπαθούν να εκτιμήσουν την επόμενη, ή τις επόμενες τιμές της χρονοσειράς, χρησιμοποιώντας μια γραμμική έκφραση των p προηγούμενων τιμών της.

Το μοντέλο ορίζεται ως εξής:

$$X_t = c + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t$$

Όπου c μια σταθερά, φ_i οι συντελεστές του μοντέλου (εκτιμώνται από την προσαρμογή στα πραγματικά δεδομένα) και ε_t ο θόρυβος των παρατηρήσεων.

2.2.3 Μοντέλο MA

Το μοντέλο MA(q) δηλαδή το μοντέλου κινητού μέσου τάξης q, επίσης χρησιμοποιεί ένα παράθυρο των q προηγούμενων μεταβλητών, ώστε να επιτύχει εκτίμηση της επόμενης τιμής. Λέγεται κινητού μέσου, καθώς η εκτίμηση προκύπτει από τον υπολογισμό των μέσων τιμών των παρατηρήσεων στο q παράθυρο.

Το μοντέλο ορίζεται ως εξής:

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

Όπου μ η εκτίμηση του μέσου όρου των παρατηρήσεων στις προηγούμενες q τιμές της χρονοσειράς, θ_i οι συντελεστές που θα εκτιμηθούν από την προσαρμογή στα δεδομένα και ε_t ο θόρυβος των παρατηρήσεων.

2.2.4 Μοντέλο ARMA

Ο συνδυασμός των δύο παραπάνω μοντέλων, δίνει το αυτοπαλινδρομούμενο μοντέλου κινητού μέσου ARMA(p,q). Το μοντέλο ορίζεται ως το άθροισμα των δύο διαδικασιών, δηλαδή:

$$X_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$$

Στο οποίο αν προστεθεί το στοιχείο των πρώτων διαφορών που έχουμε αναφέρει, μετασχηματίζεται στο μοντέλο ARIMA(p,d,q) δηλαδή σε ένα μοντέλο AR(p) συνδυασμένο με το I(d), δηλαδή πρώτες διαφορές τάξης d και ένα μοντέλο MA(q). Έτσι, συμβαίνει μια πιο σφαιρική εποπτεία επί του μοντέλου και οι επόμενες τιμές μπορούν να εκτιμηθούν από τη σχέση:

$$X_t = \left(1 - \sum_{i=1}^p \varphi_i L^i + \right) (1-L)^d X_t = \delta + \left(1 + \sum_{i=1}^q \theta_i L^i + \right) \varepsilon_t$$

Το παραπάνω μοντέλο, είναι και αυτό που χρησιμοποιήθηκε στην εργασία μας.

Εκτίμηση των p,d και q

Στην προσπάθεια εύρεσης του μοντέλου που θα προσαρμοστεί καλύτερα στα δεδομένα μας, πρέπει να εκτιμηθούν αρχικά οι παράμετροι p, d και q .

Αρχικά δίνονται δύο ορισμοί:

Ορισμός 1

Συσχέτιση των μεταβλητών X_t και X_{t-1} , είναι η ποσότητα:

$$\text{cor}(x_t, x_{t-1}) = \rho_{x_t, x_{t-1}} = r_{x_t, x_{t-1}} = \frac{s_{x_t, x_{t-1}}}{s_{x_t} s_{x_{t-1}}},$$

$$\text{όπου } s_{xy} = \text{COV}(x, y) = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{n-1},$$

$$s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \text{ και}$$

$$s_y = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}$$

Αν λοιπόν για όποια υστέρηση $d=1,2,\dots$ υπολογιστούν οι συσχετίσεις των X_t, X_{t-d} , θεωρούμε ως τάξη του αυτοπαλινδρομούμενου μοντέλου, την υστέρηση για την οποία πρώτη φορά υπάρχει αρνητική συσχέτιση, μειωμένη κατά ένα. Αν δηλαδή η υστέρηση 2 δίνει πρώτη φορά αρνητική συσχέτιση μεταξύ των μεταβλητών X_t, X_{t-2} , σημαίνει πως η βέλτιστη τάξη του AR τμήματος, θα είναι το $p=1$.

Ορισμός 2

Η μερική αυτοσυσχέτιση υστέρησης d , είναι η απευθείας αυτοσυσχέτιση των τιμών X_t και X_{t-d} , χωρίς τον συνυπολογισμό των ενδιάμεσων υστερήσεων. Ο τύπος που υπολογίζει μερικές αυτοσυσχετίσεις, δίνεται παρακάτω:

$$\phi_{\tau,\tau} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{\tau-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{\tau-3} & \rho_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{\tau-1} & \rho_{\tau-2} & \rho_{\tau-3} & \dots & \rho_1 & \rho_\tau \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{\tau-2} & \rho_{\tau-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{\tau-3} & \rho_{\tau-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{\tau-1} & \rho_{\tau-2} & \rho_{\tau-3} & \dots & \rho_1 & 1 \end{vmatrix}},$$

όπου ρ οι αντίστοιχες αυτοσυσχετίσεις.

Μετά τον υπολογισμό των μερικών αυτοσυσχετίσεων κάθε υστέρησης, η βέλτιστη τάξη του MA μοντέλου, επιλέγεται ακριβώς με την ίδια λογική που επιλέχθηκε το p .

Τέλος, η τάξη d είναι το πλήθος μετασχηματισμών πρώτων διαφορών που απαιτήθηκε, ώστε να γίνει απαλοιφή της τάσης της χρονοσειράς, ή προς καταστροφή περιοδικότητας. Αν λοιπόν χρειάστηκαν 2 μετασχηματισμοί πρώτων διαφορών στην αρχική χρονοσειρά, τότε επιλέγεται $d=2$.

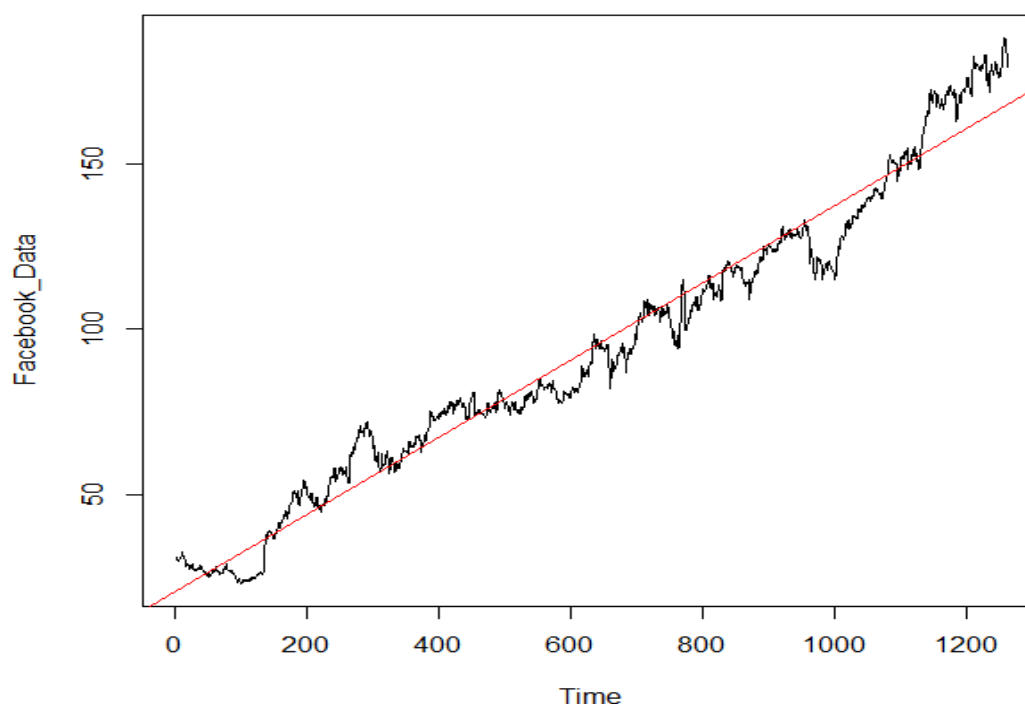
Με βάση αυτά λοιπόν, μπορούμε να έχουμε μία εποπτεία όσον αφορά το τι θα γίνει στο ερευνητικό κομμάτι, στο οποίο χρησιμοποιήθηκαν τα μοντέλα ARIMA(p,d,q), οι παράμετροι των οποίων εκτιμήθηκαν όπως περιγράψαμε.

3. Μοντελοποίηση με χρήση ARIMA

3.1 Πρόβλεψη της τιμής κλεισίματος της εταιρείας Facebook

Προετοιμασία για την εφαρμογή του μοντέλου

Αρχικά μπορούμε να οπτικοποιήσουμε τα δεδομένα μας, αναπαριστώντας τη χρονοσειρά των τιμών στους άξονες. Έτσι προκύπτει το εξής γράφημα (Γράφημα 7):



Γράφημα 7: Χρονοσειρά με τάση της τιμής κλεισίματος της εταιρείας Facebook

Η μαύρη γραμμή δηλώνει τις τιμές της χρονοσειράς, ενώ η κόκκινη δηλώνει τη γραμμική εκτίμηση της τάσης που εμφανίζεται.

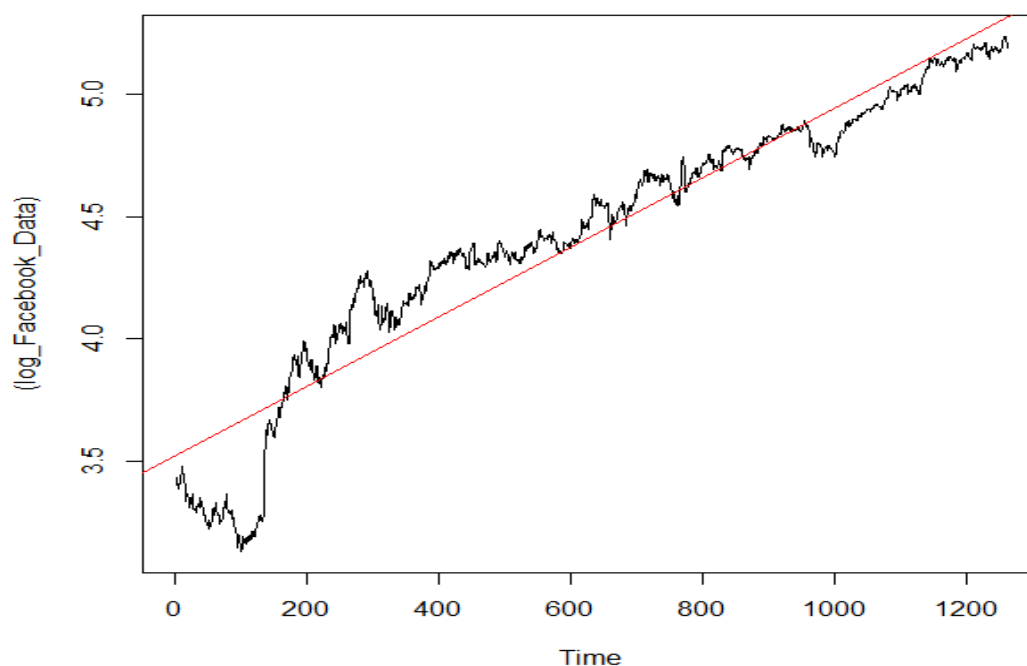
Παρατηρούμε πως η συγκεκριμένη χρονοσειρά δεν έχει κάποια εμφανή περιοδικότητα ή εποχικότητα, παρ' όλα αυτά εμφανίζει μια σημαντική αύξουσα τάση.

Είναι γνωστό πως για την εφαρμογή ενός μοντέλου πρόβλεψης, είναι απαραίτητη η διάσπαση της χρονοσειράς στις κύριες συνιστώσες της, δηλαδή στη συνιστώσα της τάσης, της εποχικότητας, καθώς επίσης και η ρύθμιση των τυχαίων διακυμάνσεων.

Απαλοιφή τυχαίων διακυμάνσεων

Η απαλοιφή των τυχαίων διακυμάνσεων γίνεται με την εφαρμογή της λογαριθμικής συνάρτησης.

Έτσι, η αρχική χρονοσειρά μετασχηματίζεται στην παρακάτω χρονοσειρά (Γράφημα 8), στην οποία έχει ρυθμιστεί το ζήτημα αυτών των διακυμάνσεων:

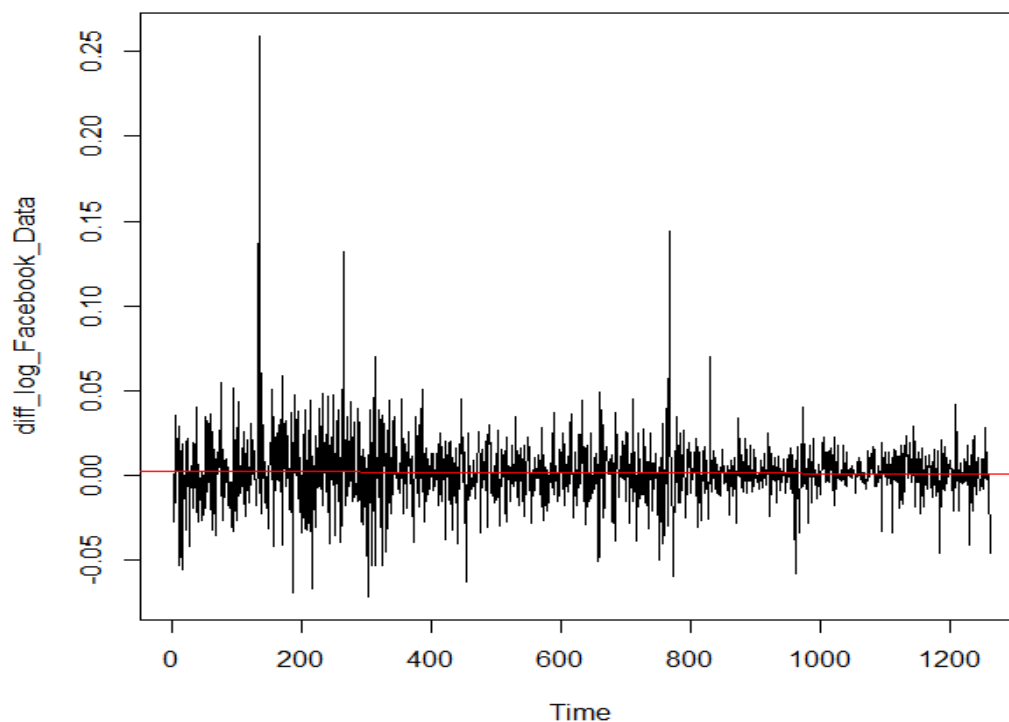


Γράφημα 8: Χρονοσειρά με τάση της εταιρείας Facebook κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων

Απαλοιφή τάσης

Έπειτα, ακολουθεί η διαδικασία απαλοιφής τάσης. Αυτό θα συμβεί παίρνοντας το μετασχηματισμό των πρώτων διαφορών στην αρχική ακολουθία τιμών.

Μέσω του μετασχηματισμού αυτού, προκύπτει μια νέα χρονοσειρά, που όπως φαίνεται παρακάτω ((Γράφημα 9), έχει μηδενική τάση:



Γράφημα 9: Χρονοσειρά χωρίς τάση της εταιρείας Facebook

Το μοντέλο

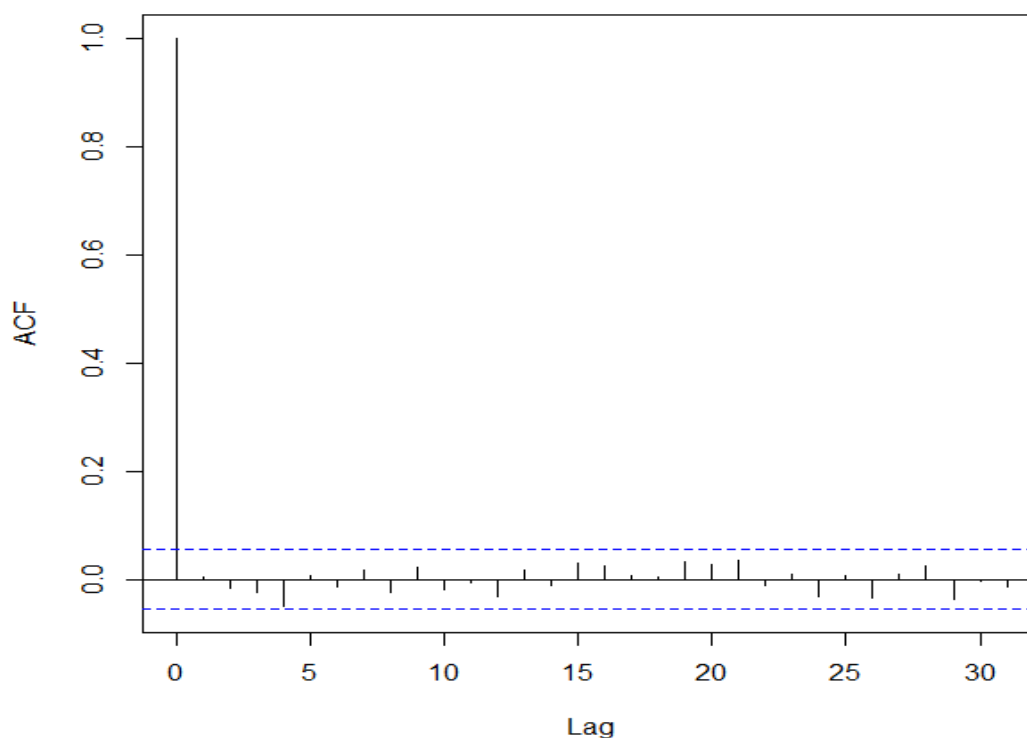
Το μοντέλο που θα χρησιμοποιηθεί είναι το ARIMA(p,d,q).

Η συνήθης διαδικασία, περιλαμβάνει την εκτίμηση των παραμέτρων του μοντέλου και έπειτα την εφαρμογή του.

Στη συγκεκριμένη άσκηση ζητείται πρόβλεψη της επόμενης τιμής μόνο με συγκεκριμένη υστέρηση (3, 7 και 14 ημέρες), επομένως η εκτίμηση της παραμέτρου p είναι άσκοπη.

Παρ όλα αυτά, χάριν πληρότητας, παρακάτω δίνεται το γράφημα (Γράφημα 10), αυτοσυσχετίσεων της χρονοσειράς:

Series diff_log_Facebook_Data

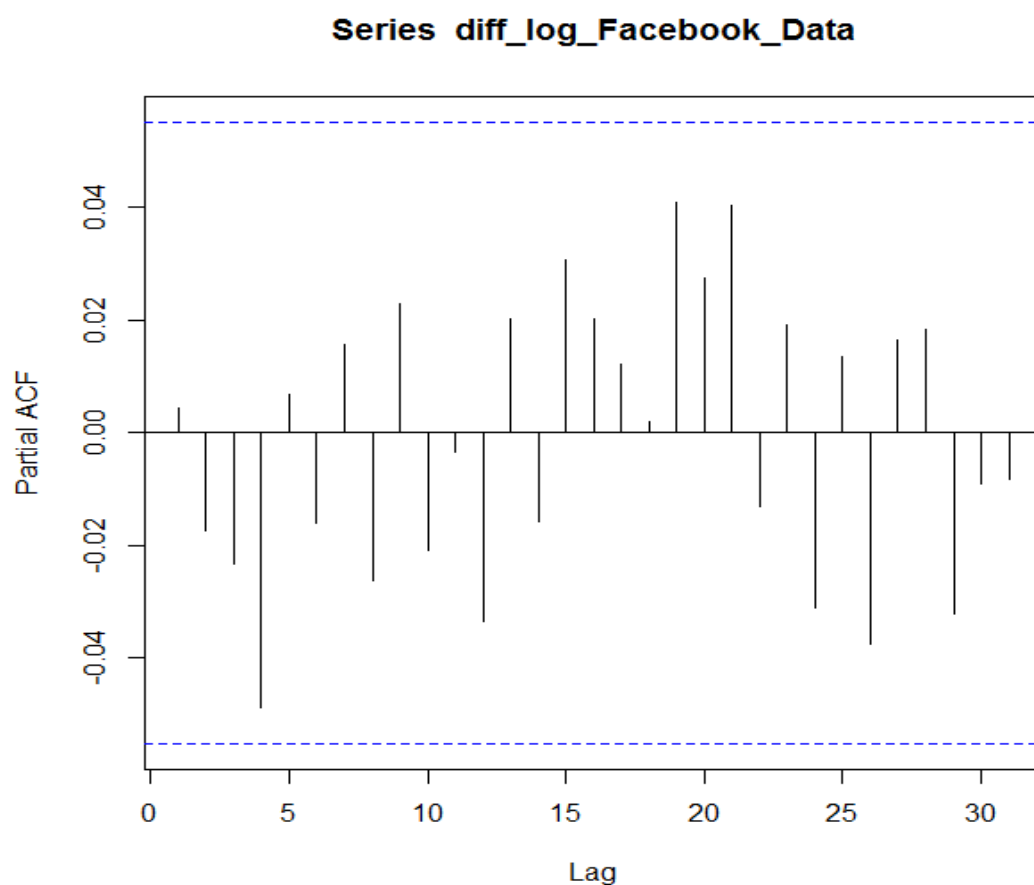


Γράφημα 10: Γράφημα αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία Facebook

Από το γράφημα αυτό, ως τάξη για το αυτοπαλινδρομούμενο τμήμα του μοντέλου, υποδεικνύεται το $p=1$, καθώς όπως φαίνεται, η τρίτη γραμμή (δηλαδή το $p=2$) εμφανίζει για πρώτη φορά αλλαγή προσήμου αυτοσυσχέτισης, όπου όπως είδαμε στην εισαγωγή είναι δείκτης για την τάξη του μοντέλου.

Έτσι, θα επιλέγαμε $p=1$

Επίσης, για το MA τμήμα του μοντέλου, παίρνουμε το παρακάτω γράφημα των μερικών αυτοσυσχετίσεων (Γράφημα 11):



Γράφημα 11: Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία Facebook

Εδώ φαίνεται πως υπάρχει αντιστροφή του προσήμου στην υστέρηση 2, γεγονός που υποδεικνύει ως τάξη του τμήματος κινούμενου μέσου, το $q=1$.

Εφαρμογή του μοντέλου

Με βάση λοιπόν τα παραπάνω, το MA τμήμα έχει τάξη 1 και η τάση απαλείφθηκε με την εφαρμογή πρώτων διαφορών ($d=1$).

Έτσι, τα μοντέλα που προκύπτουν για αυτές τις παραμέτρους, σε συνδυασμό με υστερήσεις 3, 7 και 14 ως προς το αυτοπαλινδρομούμενο τμήμα, που απαιτείται από την εκφώνηση, παίρνουμε τα τρία παρακάτω μοντέλα:

ΠΙΝΑΚΑΣ 1: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στη εταιρεία FB για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 3.

Model 1				
Κώδικας που χρησιμοποιήθηκε	fit3=arima(log_Facebook_Data,c(3,1,1))			
Συντελεστές				
	AR 1	AR 2	AR 3	MA 1
	0.5399	-0.0171	-0.0203	-0.5311

ΠΙΝΑΚΑΣ 2: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στη εταιρεία FB για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 7.

Model 2								
Κώδικας που χρησιμοποιήθηκε	fit7=arima(log_Facebook_Data,c(7,1,1))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	MA 1
	-0.9554	-0.0041	-0.0287	-0.0599	-0.0285	0.0020	0.0128	0.9729

ΠΙΝΑΚΑΣ 3: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στη εταιρεία FB για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 14.

Model 3								
Κώδικας που χρησιμοποιήθηκε	fit14=arima(log_Facebook_Data,c(14,1,1))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	AR 8
	-0.9557	-0.0039	-0.0286	-0.0599	-0.0280	0.0028	0.0120	-0.0003
	AR 9	AR 10	AR 11	AR 12	AR 13	AR 14	MA 1	
	0.0085	0.0128	-0.0112	-0.0252	-0.0007	0.0083	0.9731	

Έλεγχος προσαρμογής και σύγκριση των μοντέλων

Για να ελέγξουμε την εγκυρότητα των μοντέλων και το κατά πόσο αυτά προσεγγίζουν καλά τα δεδομένα μας, κάνουμε το εξής:

Δημιουργήθηκε ένας αλγόριθμος στην R ο οποίος με τις συγκεκριμένες παραμέτρους, χρησιμοποιεί ως δείγμα εκπαίδευσης τις πρώτες 1210 παρατηρήσεις (από τις 1260 συνολικά) και με βάση αυτές κάνει πρόβλεψη για τις υπόλοιπες 50 (από την παρατήρηση 1211 έως και την 1260). Έπειτα, υπολογίζει τη ρίζα του μέσου τετραγωνικού σφάλματος (RMSE), το οποίο αποτελεί και ένα δείκτη της προσαρμογής αυτής.

1ο μοντέλο

> COMPARE3

	Obs 1211	Obs 1212	Obs 1213	Obs 1214	Obs 1215	Obs 1216	Obs 1217
Original Data	182.6600	178.9200	178.92	180.1700	180.2500	179.5600	179.3000
Model Prediction	179.8417	179.6871	179.61	179.5792	179.5686	179.5655	179.5648
	Obs 1218	Obs 1219	Obs 1220	Obs 1221	Obs 1222	Obs 1223	Obs 1224
Original Data	178.4600	178.7700	178.0700	177.9500	179.5900	179.0000	178.7400
Model Prediction	179.5648	179.5648	179.5649	179.5649	179.5649	179.5649	179.5649
	Obs 1225	Obs 1226	Obs 1227	Obs 1228	Obs 1229	Obs 1230	Obs 1231
Original Data	181.8600	180.8700	182.7800	183.0300	182.4200	175.1300	177.1800
Model Prediction	179.5649	179.5649	179.5649	179.5649	179.5649	179.5649	179.5649
	Obs 1232	Obs 1233	Obs 1234	Obs 1235	Obs 1236	Obs 1237	Obs 1238
Original Data	175.1000	171.4700	172.8300	176.0600	180.1400	179.0000	179.0400
Model Prediction	179.5649	179.5649	179.5649	179.5649	179.5649	179.5649	179.5649
	Obs 1239	Obs 1240	Obs 1241	Obs 1242	Obs 1243	Obs 1244	Obs 1245
Original Data	176.9600	178.3000	178.3900	180.1800	180.8200	179.5100	177.8900
Model Prediction	179.5649	179.5649	179.5649	179.5649	179.5649	179.5649	179.5649
	Obs 1246	Obs 1247	Obs 1248	Obs 1249	Obs 1250	Obs 1251	Obs 1252
Original Data	177.4500	177.2000	175.9900	177.6200	177.9200	176.4600	181.4200
Model Prediction	179.5649	179.5649	179.5649	179.5649	179.5649	179.5649	179.5649
	Obs 1253	Obs 1254	Obs 1255	Obs 1256	Obs 1257	Obs 1258	Obs 1259
Original Data	184.6700	184.3300	186.8500	188.2800	187.8700	187.8400	187.7700
Model Prediction	179.5649	179.5649	179.5649	179.5649	179.5649	179.5649	179.5649
	Obs 1260						
Original Data	179.3700						
Model Prediction	179.5649						

2ο μοντέλο

> COMPARE7

	Obs 1211	Obs 1212	Obs 1213	Obs 1214	Obs 1215	Obs 1216	Obs 1217
Original Data	182.6600	178.920	178.9200	180.1700	180.2500	179.5600	179.3000
Model Prediction	180.0096	179.543	179.6881	179.5425	179.7893	179.6019	179.7878
	Obs 1218	Obs 1219	Obs 1220	Obs 1221	Obs 1222	Obs 1223	Obs 1224
Original Data	178.4600	178.7700	178.0700	177.9500	179.5900	179.0000	178.7400
Model Prediction	179.6055	179.7688	179.6144	179.7597	179.6249	179.7511	179.6333
	Obs 1225	Obs 1226	Obs 1227	Obs 1228	Obs 1229	Obs 1230	Obs 1231
Original Data	181.8600	180.8700	182.7800	183.0300	182.4200	175.1300	177.1800
Model Prediction	179.7431	179.6405	179.7362	179.6469	179.7303	179.6525	179.7251
	Obs 1232	Obs 1233	Obs 1234	Obs 1235	Obs 1236	Obs 1237	Obs 1238
Original Data	175.1000	171.4700	172.8300	176.0600	180.1400	179.0000	179.0400
Model Prediction	179.6574	179.7206	179.6616	179.7166	179.6653	179.7132	179.6685
	Obs 1239	Obs 1240	Obs 1241	Obs 1242	Obs 1243	Obs 1244	Obs 1245
Original Data	176.9600	178.3000	178.3900	180.1800	180.8200	179.5100	177.8900
Model Prediction	179.7102	179.6712	179.7076	179.6737	179.7053	179.6758	179.7034
	Obs 1246	Obs 1247	Obs 1248	Obs 1249	Obs 1250	Obs 1251	Obs 1252
Original Data	177.4500	177.2000	175.9900	177.6200	177.9200	176.4600	181.4200
Model Prediction	179.6776	179.7016	179.6792	179.7001	179.6806	179.6988	179.6818
	Obs 1253	Obs 1254	Obs 1255	Obs 1256	Obs 1257	Obs 1258	Obs 1259
Original Data	184.6700	184.3300	186.8500	188.2800	187.8700	187.8400	187.7700
Model Prediction	179.6977	179.6829	179.6967	179.6838	179.6959	179.6846	179.6951
	Obs 1260						
Original Data	179.3700						
Model Prediction	179.6853						

3ο μοντέλο

> COMPARE14

Obs 1211	Obs 1212	Obs 1213	Obs 1214	Obs 1215	Obs 1216	Obs 1217	
Original Data	182.6600	178.9200	178.9200	180.1700	180.2500	179.5600	179.3000
Model Prediction	179.9679	179.4714	179.6326	179.4747	179.7776	179.6562	179.8906
Obs 1218	Obs 1219	Obs 1220	Obs 1221	Obs 1222	Obs 1223	Obs 1224	
Original Data	178.4600	178.7700	178.0700	177.950	179.5900	179.0000	178.7400
Model Prediction	179.7881	179.8298	179.5616	179.743	179.6263	179.7716	179.6589
Obs 1225	Obs 1226	Obs 1227	Obs 1228	Obs 1229	Obs 1230	Obs 1231	
Original Data	181.8600	180.8700	182.7800	183.0300	182.4200	175.1300	177.1800
Model Prediction	179.7647	179.6587	179.7516	179.6618	179.7421	179.6669	179.7428
Obs 1232	Obs 1233	Obs 1234	Obs 1235	Obs 1236	Obs 1237	Obs 1238	
Original Data	175.1000	171.4700	172.830	176.060	180.1400	179.000	179.0400
Model Prediction	179.6758	179.7366	179.678	179.732	179.6815	179.729	179.6849
Obs 1239	Obs 1240	Obs 1241	Obs 1242	Obs 1243	Obs 1244	Obs 1245	
Original Data	176.9600	178.3000	178.3900	180.18	180.8200	179.5100	177.8900
Model Prediction	179.7261	179.6878	179.7236	179.69	179.7212	179.6921	179.7194
Obs 1246	Obs 1247	Obs 1248	Obs 1249	Obs 1250	Obs 1251	Obs 1252	
Original Data	177.450	177.2000	175.9900	177.6200	177.9200	176.4600	181.4200
Model Prediction	179.694	179.7177	179.6955	179.7162	179.6969	179.7149	179.6981
Obs 1253	Obs 1254	Obs 1255	Obs 1256	Obs 1257	Obs 1258	Obs 1259	
Original Data	184.6700	184.3300	186.8500	188.2800	187.8700	187.8400	187.7700
Model Prediction	179.7138	179.6992	179.7128	179.7001	179.7119	179.7009	179.7112
Obs 1260							
Original Data	179.3700						
Model Prediction	179.7016						

Για περαιτέρω επεξήγηση της λειτουργίας του συγκεκριμένου αλγορίθμου, αναφέρουμε το εξής:

Στον πρώτο πίνακα, βλέπουμε ότι η παρατήρηση 1211 έχει πραγματική τιμή 182.6600, ενώ το μοντέλο μας, προέβλεψε ότι θα είχε 179.9679. Αυτή η πρόβλεψη έγινε μέσω ενός μοντέλου ARIMA με τάξη $p=3$, $d=1$ και $q=1$ όπου τα d και q αναδείχθηκαν από την ανάλυση, ενώ το $p=3$ προκύπτει από την εκφώνηση καθώς θέλουμε χρήση των 3^{ων} προηγούμενων ημερών. Στο μοντέλο αυτό, χρησιμοποιήθηκαν δεδομένα από την αρχή έως και την 1210 παρατήρηση και έπειτα έγινε πρόβλεψη για την 1211. Στην αμέσως επόμενη στήλη του ίδιου πίνακα, χρησιμοποιήθηκε ακριβώς το ίδιο μοντέλο από το οποίο έχει καταγραφεί η πρόβλεψη 2 βημάτων μετά και ούτω καθ' εξής.

Ως πρώτη παρατήρηση μπορούμε να πούμε πως είναι εμφανής η αδυναμία πρόβλεψης των μοντέλων για πολλά βήματα μπροστά. Όπως φαίνεται, η προβλεπόμενη τιμή σταθεροποιείται σχετικά γρήγορα γύρω από μία σχεδόν σταθερή τιμή και σταματάει να προσεγγίζει ικανοποιητικά τις πραγματικές τιμές του φαινομένου.

Για μία σαφή εκτίμηση της προσαρμογής των τριών μοντέλων, παίρνουμε το δείκτη RMSE που αναφέραμε προηγουμένως:

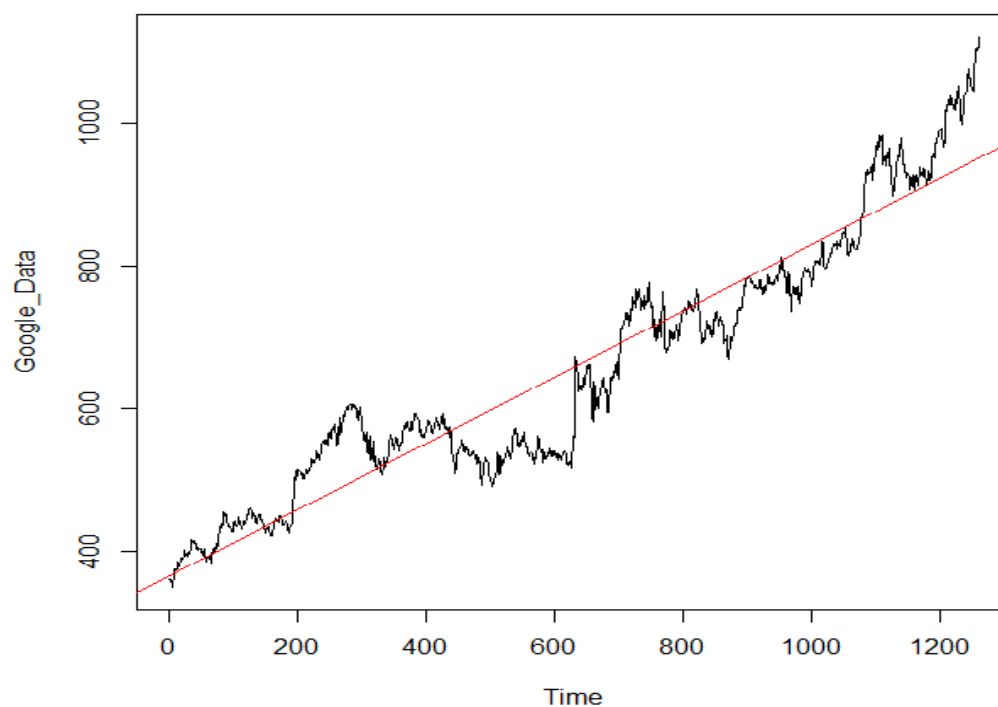
```
> RMSE3
[1] 2.5768
> RMSE7
[1] 2.570676
> RMSE14
[1] 2.571407
```

Όπως φαίνεται, το μοντέλο που ως είσοδο δέχεται τις 7 προηγούμενες ημέρες είναι καλύτερο από τα άλλα δύο. Ως προς την ποιότητα πρόβλεψης, ακολουθεί το μοντέλο με υστέρηση 14 και τέλος το μοντέλο με υστέρηση 3.

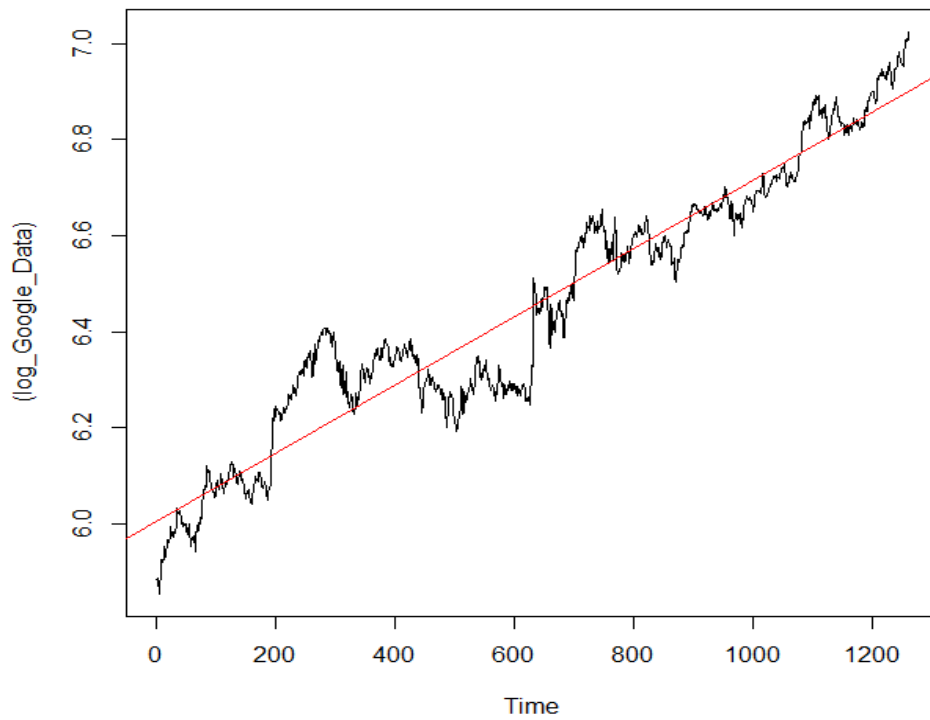
3.2 Πρόβλεψη της τιμής κλεισίματος της εταιρείας Google

Προετοιμασία για την εφαρμογή του μοντέλου

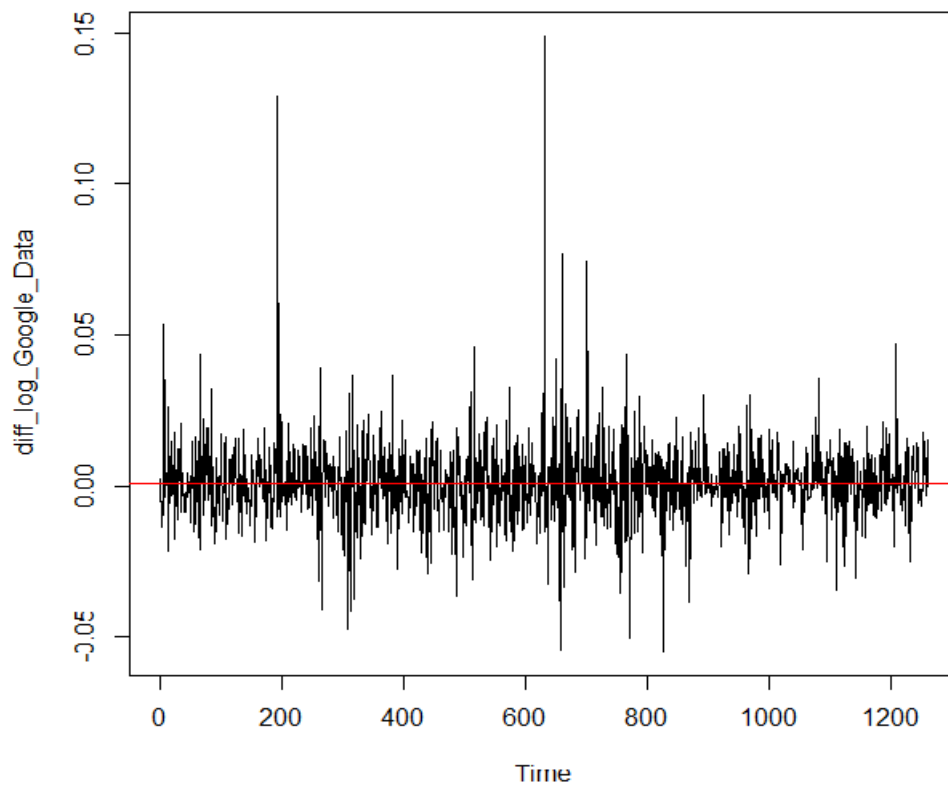
Παρακάτω βλέπουμε το γράφημα της χρονοσειράς (Γράφημα 12), και στη συνέχεια την ίδια χρονοσειρά απαλλαγμένη από τυχαίες διακυμάνσεις (Γράφημα 13), και κάτω επίσης την ίδια χρονοσειρά απαλλαγμένη και από τυχαίες διακυμάνσεις και από τάση (Γράφημα 14).



Γράφημα 12 :Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας Google



Γράφημα 13: Χρονοσειρά με τάση της εταιρείας Google κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων



Γράφημα 14: Χρονοσειρά χωρίς τάση της εταιρείας Google

Η μαύρη γραμμή δηλώνει τις τιμές της χρονοσειράς, ενώ η κόκκινη δηλώνει τη γραμμική εκτίμηση της τάσης που εμφανίζεται.

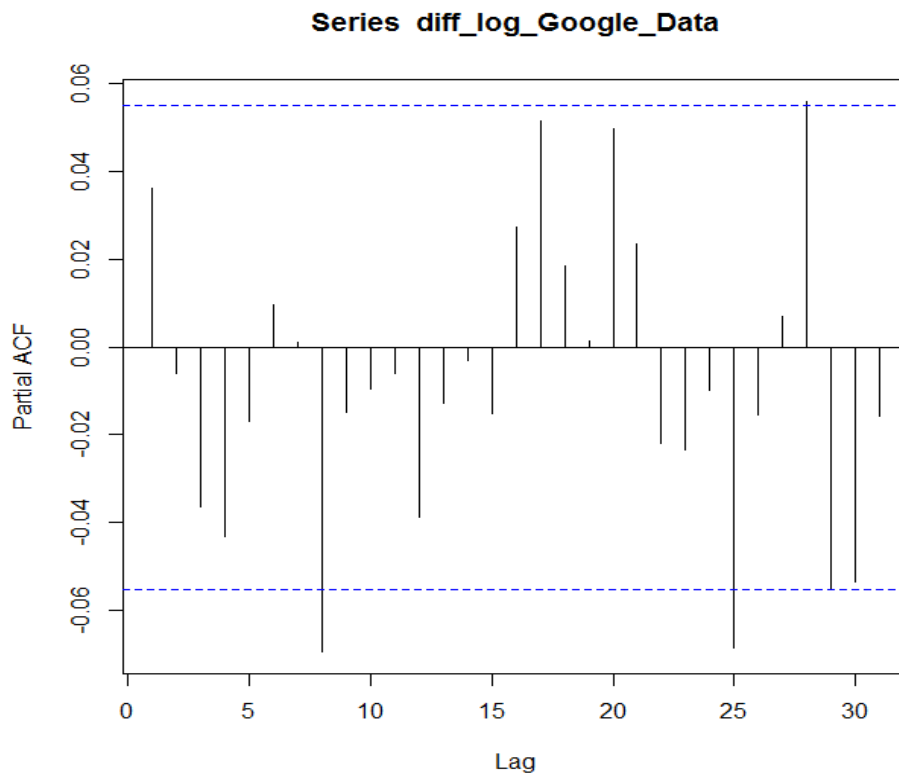
Παρατηρούμε πως η συγκεκριμένη χρονοσειρά δεν έχει κάποια εμφανή περιοδικότητα ή εποχικότητα, παρ'όλα αυτά εμφανίζει μια επίσης αύξουσα τάση, όπως και της προηγούμενης εταιρείας.

Το μοντέλο

Όπως και στην προηγούμενη περίπτωση, θα χρησιμοποιήσουμε ένα μοντέλο ARIMA(p,d,q).

Από εδώ και πέρα θα εκτιμούμε μόνο την τάξη q του MA τμήματος, καθώς η τάξη του αντίστοιχου αυτοπαλινδρομούμενου τμήματος είναι ορισμένη από την εκφώνηση (3,7 ή 14)

Το γράφημα λοιπόν των μερικών αυτοσυσχετίσεων (Γράφημα 15), δίνεται παρακάτω:



Γράφημα 15: Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία Google

Βλέπουμε πως υπάρχει αναστροφή για υστέρηση 2, επομένως η προτεινόμενη τάξη του MA τμήματος του μοντέλου θα είναι $q=1$.

Εφαρμογή του μοντέλου

Με βάση λοιπόν τα παραπάνω, το MA τμήμα έχει τάξη 1 και η τάση απαλείφθηκε με την εφαρμογή πρώτων διαφορών ($d=1$).

Έτσι, τα μοντέλα που προκύπτουν για αυτές τις παραμέτρους, σε συνδυασμό με υστερήσεις 3, 7 και 14 ως προς το αυτοπαλινδρομούμενο τμήμα, που απαιτείται από την εκφώνηση, παίρνουμε τα τρία παρακάτω μοντέλα:

ΠΙΝΑΚΑΣ 4: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στη εταιρεία Google για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 3

Model 1				
Κώδικας που χρησιμοποιήθηκε	fit3=arima(log_Google_Data,c(3,1,1))			
Συντελεστές				
	AR 1	AR 2	AR 3	MA 1
	0.7962	-0.0311	-0.0325	-0.7578

ΠΙΝΑΚΑΣ 5: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στη εταιρεία Google για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 7

Model 2								
Κώδικας που χρησιμοποιήθηκε	fit7=arima(log_Google_Data,c(7,1,1))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	MA 1
	-0.6218	0.0255	-0.0296	-0.0573	-0.0378	0.0055	0.0371	0.6614

ΠΙΝΑΚΑΣ 6: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στη εταιρεία Google για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 14

Model 3								
Κώδικας που χρησιμοποιήθηκε	fit14=arima(log_Google_Data,c(14,1,1))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	AR 8
	-0.2740	0.0124	-0.0310	-0.0523	-0.0283	0.0102	0.0122	-0.064
	AR 9	AR 10	AR 11	AR 12	AR 13	AR 14	MA 1	
	-0.0320	-0.0080	-0.0009	-0.0333	-0.0178	0.0046	0.3123	

Έλεγχος προσαρμογής και σύγκριση των μοντέλων

Με χρήση του αλγορίθμου που αναφέραμε και στην πρώτη περίπτωση, βλέπουμε παρακάτω τις συγκρίσεις μεταξύ των πραγματικών και προβλεπόμενων από τα μοντέλα τιμών:

1ο μοντέλο

```
> COMPARE3
Obs 1211 Obs 1212 Obs 1213 Obs 1214 Obs 1215 Obs 1216 Obs 1217
Original Data 1025.500 1025.580 1032.480 1025.900 1033.330 1039.850 1031.260
Model Prediction 1015.082 1013.952 1013.128 1012.569 1012.193 1011.943 1011.776
Obs 1218 Obs 1219 Obs 1220 Obs 1221 Obs 1222 Obs 1223 Obs 1224
Original Data 1028.070 1025.750 1026.000 1020.910 1032.500 1019.090 1018.380
Model Prediction 1011.666 1011.593 1011.545 1011.513 1011.491 1011.477 1011.468
Obs 1225 Obs 1226 Obs 1227 Obs 1228 Obs 1229 Obs 1230 Obs 1231
Original Data 1034.490 1035.960 1040.610 1054.210 1047.410 1021.660 1021.410
Model Prediction 1011.462 1011.458 1011.455 1011.453 1011.452 1011.451 1011.451
Obs 1232 Obs 1233 Obs 1234 Obs 1235 Obs 1236 Obs 1237 Obs 1238
Original Data 1010.17 998.68 1005.15 1018.38 1030.93 1037.05 1041.10
Model Prediction 1011.45 1011.45 1011.45 1011.45 1011.45 1011.45 1011.45
Obs 1239 Obs 1240 Obs 1241 Obs 1242 Obs 1243 Obs 1244 Obs 1245
Original Data 1040.48 1040.61 1049.15 1064.19 1077.14 1070.68 1064.95
Model Prediction 1011.45 1011.45 1011.45 1011.45 1011.45 1011.45 1011.45
Obs 1246 Obs 1247 Obs 1248 Obs 1249 Obs 1250 Obs 1251 Obs 1252
Original Data 1063.63 1060.12 1056.74 1049.37 1048.14 1046.40 1065.00
Model Prediction 1011.45 1011.45 1011.45 1011.45 1011.45 1011.45 1011.45
Obs 1253 Obs 1254 Obs 1255 Obs 1256 Obs 1257 Obs 1258 Obs 1259
Original Data 1082.48 1086.40 1102.23 1106.94 1106.26 1102.61 1105.52
Model Prediction 1011.45 1011.45 1011.45 1011.45 1011.45 1011.45 1011.45
Obs 1260
Original Data 1122.26
Model Prediction 1011.45
```

2ο μοντέλο

```
> COMPARE7
Obs 1211 Obs 1212 Obs 1213 Obs 1214 Obs 1215 Obs 1216 Obs 1217
Original Data 1025.50 1025.580 1032.480 1025.900 1033.330 1039.850 1031.260
Model Prediction 1014.35 1013.223 1012.454 1013.478 1014.726 1014.013 1014.533
Obs 1218 Obs 1219 Obs 1220 Obs 1221 Obs 1222 Obs 1223 Obs 1224
Original Data 1028.070 1025.750 1026.000 1020.910 1032.500 1019.090 1018.380
Model Prediction 1013.999 1014.232 1014.023 1014.229 1014.133 1014.192 1014.163
Obs 1225 Obs 1226 Obs 1227 Obs 1228 Obs 1229 Obs 1230 Obs 1231
Original Data 1034.490 1035.960 1040.610 1054.21 1047.410 1021.660 1021.410
Model Prediction 1014.164 1014.166 1014.159 1014.17 1014.161 1014.169 1014.162
Obs 1232 Obs 1233 Obs 1234 Obs 1235 Obs 1236 Obs 1237 Obs 1238
Original Data 1010.170 998.680 1005.150 1018.380 1030.930 1037.050 1041.100
Model Prediction 1014.167 1014.163 1014.166 1014.164 1014.165 1014.165 1014.165
Obs 1239 Obs 1240 Obs 1241 Obs 1242 Obs 1243 Obs 1244 Obs 1245
Original Data 1040.480 1040.610 1049.150 1064.190 1077.140 1070.680 1064.950
Model Prediction 1014.165 1014.165 1014.165 1014.165 1014.165 1014.165 1014.165
Obs 1246 Obs 1247 Obs 1248 Obs 1249 Obs 1250 Obs 1251 Obs 1252
Original Data 1063.630 1060.120 1056.740 1049.370 1048.140 1046.400 1065.000
Model Prediction 1014.165 1014.165 1014.165 1014.165 1014.165 1014.165 1014.165
Obs 1253 Obs 1254 Obs 1255 Obs 1256 Obs 1257 Obs 1258 Obs 1259
Original Data 1082.480 1086.400 1102.230 1106.940 1106.260 1102.610 1105.520
Model Prediction 1014.165 1014.165 1014.165 1014.165 1014.165 1014.165 1014.165
Obs 1260
Original Data 1122.260
Model Prediction 1014.165
```

3ο μοντέλο

```
> COMPARE14
Obs 1211 Obs 1212 Obs 1213 Obs 1214 Obs 1215 Obs 1216 Obs 1217
Original Data 1025.500 1025.580 1032.480 1025.90 1033.330 1039.850 1031.260
Model Prediction 1014.554 1014.051 1013.525 1014.61 1014.895 1012.473 1012.085
Obs 1218 Obs 1219 Obs 1220 Obs 1221 Obs 1222 Obs 1223 Obs 1224
Original Data 1028.070 1025.750 1026.00 1020.910 1032.500 1019.090 1018.380
Model Prediction 1011.589 1011.918 1010.48 1010.284 1010.402 1010.412 1010.661
Obs 1225 Obs 1226 Obs 1227 Obs 1228 Obs 1229 Obs 1230 Obs 1231
Original Data 1034.490 1035.960 1040.610 1054.210 1047.41 1021.660 1021.410
Model Prediction 1010.755 1010.753 1010.691 1010.855 1010.91 1010.931 1010.921
Obs 1232 Obs 1233 Obs 1234 Obs 1235 Obs 1236 Obs 1237 Obs 1238
Original Data 1010.17 998.680 1005.150 1018.380 1030.930 1037.050 1041.100
Model Prediction 1010.94 1010.945 1010.936 1010.941 1010.922 1010.911 1010.909
Obs 1239 Obs 1240 Obs 1241 Obs 1242 Obs 1243 Obs 1244 Obs 1245
Original Data 1040.480 1040.610 1049.150 1064.190 1077.140 1070.680 1064.950
Model Prediction 1010.912 1010.907 1010.904 1010.903 1010.903 1010.904 1010.905
Obs 1246 Obs 1247 Obs 1248 Obs 1249 Obs 1250 Obs 1251 Obs 1252
Original Data 1063.630 1060.120 1056.740 1049.370 1048.140 1046.400 1065.000
Model Prediction 1010.906 1010.905 1010.906 1010.907 1010.907 1010.907 1010.907
Obs 1253 Obs 1254 Obs 1255 Obs 1256 Obs 1257 Obs 1258 Obs 1259
Original Data 1082.480 1086.400 1102.230 1106.940 1106.260 1102.610 1105.520
Model Prediction 1010.907 1010.907 1010.907 1010.907 1010.907 1010.907 1010.907
Obs 1260
Original Data 1122.260
Model Prediction 1010.907
```

Βλέπουμε πως επίσης υπάρχει μία αρχική προσέγγιση των δεδομένων, η οποία αργότερα έχει απλά την αξία που έχει η πρόβλεψη με μέσο όρο. Προσπαθώντας να επιλέξουμε το καλύτερο μοντέλο, βρίσκουμε και πάλι τα RMSE

```
> RMSE3
[1] 32.89893
> RMSE7
[1] 31.48671
> RMSE14
[1] 33.15629
```

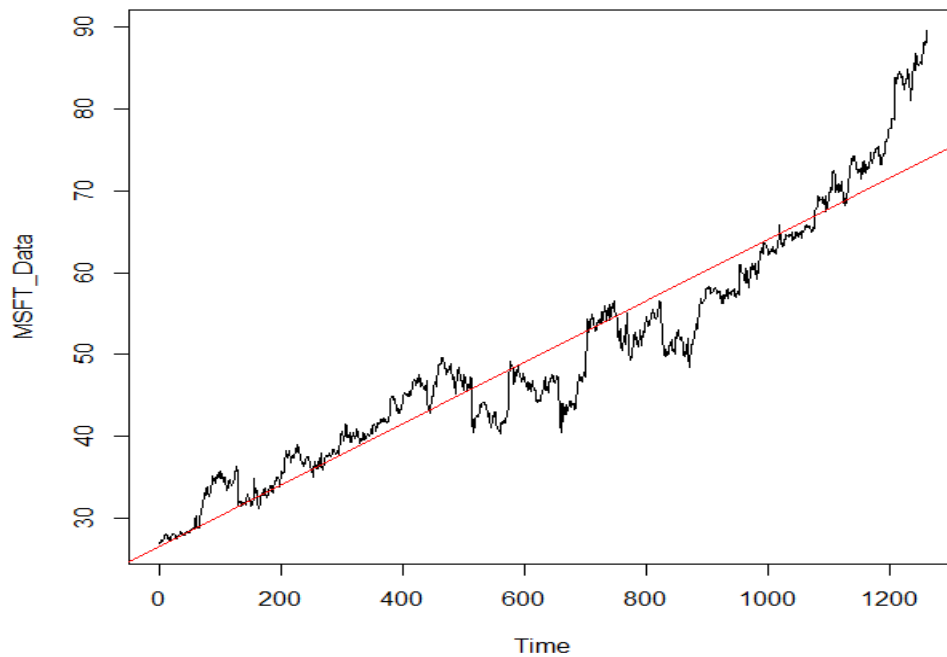
Εδώ, το μοντέλο με υστέρηση 7 είναι το καλύτερο, ακολουθεί το μοντέλο με υστέρηση 3 και τέλος το μοντέλο που κάνει πρόβλεψη με χρήση των 14ων προηγούμενων ημερών.

3.3 Πρόβλεψη της τιμής κλεισίματος της εταιρείας MSFT

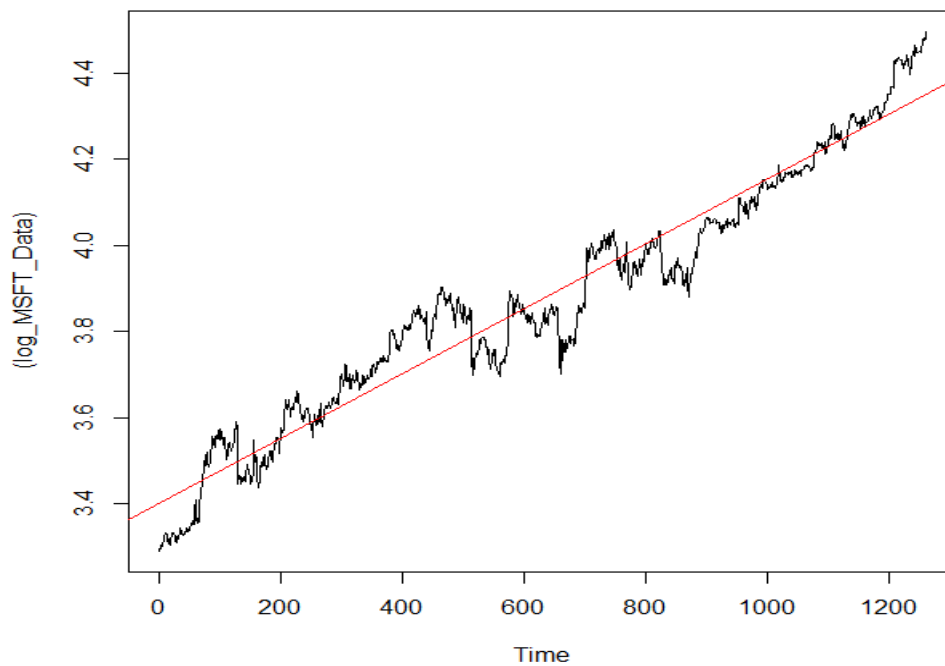
Προετοιμασία για την εφαρμογή του μοντέλου

Ομοίως για μια πρώτη άποψη, βλέπουμε το γράφημα της χρονοσειράς (Γράφημα 16), την ίδια χρονοσειρά απαλλαγμένη από τυχαίες διακυμάνσεις (Γράφημα 17) και κάτω

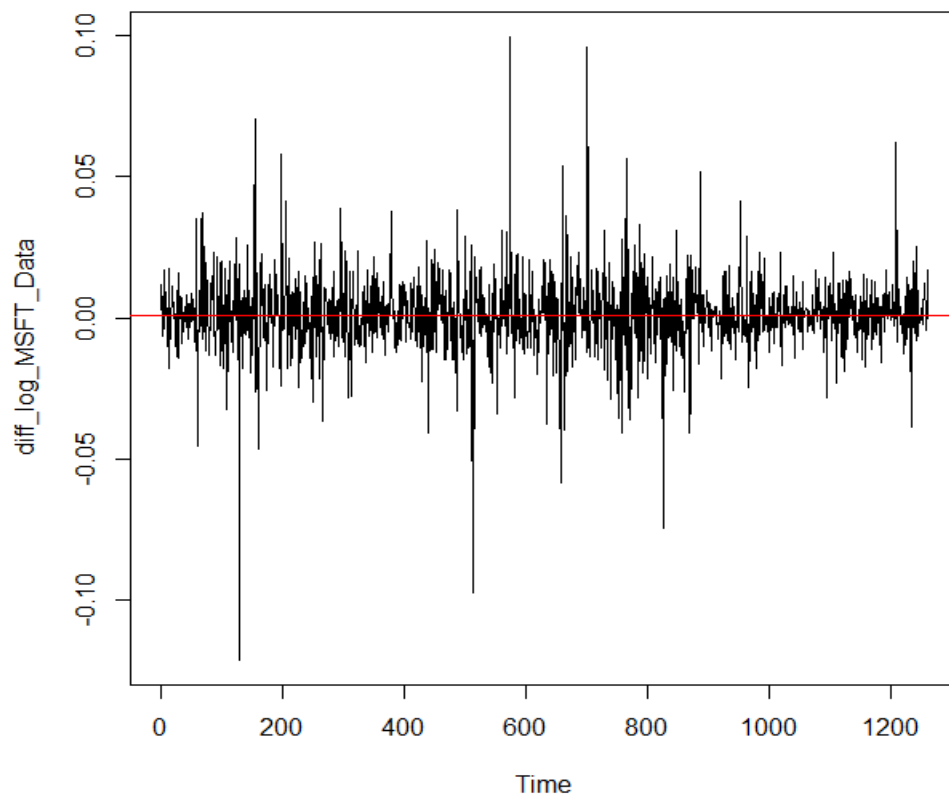
επίσης την ίδια χρονοσειρά απαλλαγμένη και από τυχαίες διακυμάνσεις και από τάση (Γράφημα 18).



Γράφημα 16 :Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας MSFT



Γράφημα 17: Χρονοσειρά με τάση της εταιρείας MSFT κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων



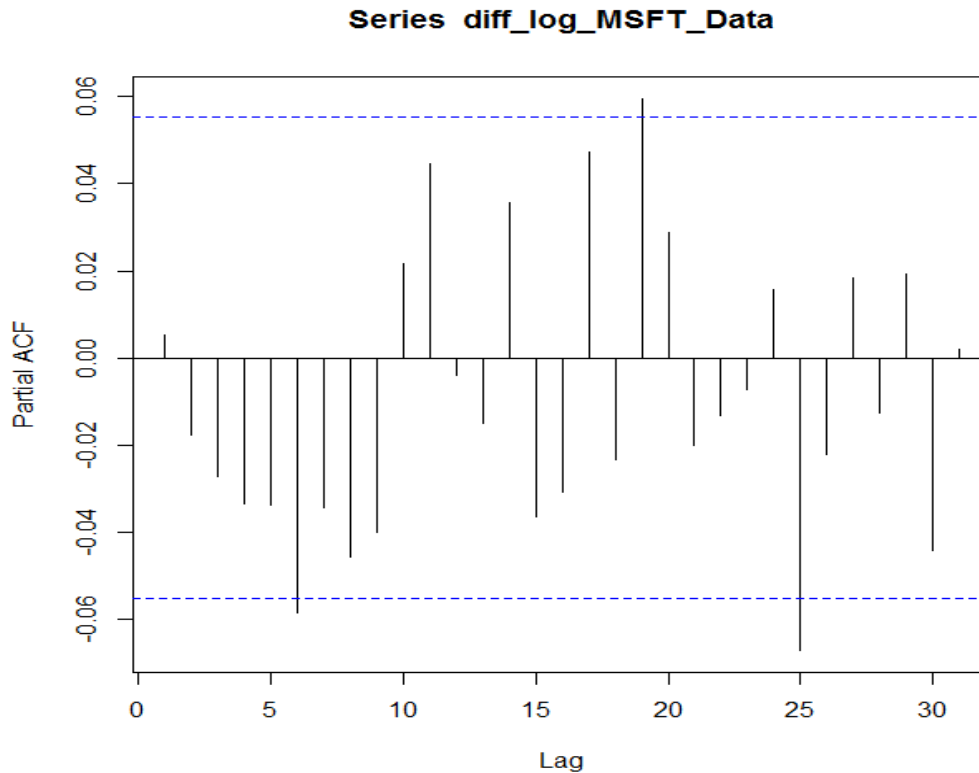
Γράφημα 18: Χρονοσειρά χωρίς τάση της εταιρείας MSFT

Η μαύρη γραμμή δηλώνει τις τιμές της χρονοσειράς, ενώ η κόκκινη δηλώνει τη γραμμική εκτίμηση της τάσης που εμφανίζεται.

Και σε αυτήν την περίπτωση δεν μπορεί να γίνει λόγος για σαφή περιοδικότητα.

Το μοντέλο

Για την εκτίμηση της παράμετρο του τμήματος κινούμενου μέσου του μοντέλου ARIMA(p,d,q), παίρνουμε το παρακάτω γράφημα μερικών αυτοσυσχετίσεων (Γράφημα 19).



Γράφημα 19: Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία MSFT

Βλέπουμε πως υπάρχει οι μερικές αυτοσυσχετίσεις φθίνουν όπως θα έπρεπε και γίνονται αρνητικές πρώτη φορά για υστέρηση 2. Αυτό υποδεικνύει πως η τάξη q θα είναι ίση με το 1.

Εφαρμογή του μοντέλου

Με βάση λοιπόν τα παραπάνω, το MA τμήμα έχει τάξη 1 και η τάση απαλείφθηκε με την εφαρμογή πρώτων διαφορών ($d=1$).

Έτσι, τα μοντέλα που προκύπτουν για αυτές τις παραμέτρους, σε συνδυασμό με υστερήσεις 3, 7 και 14 ως προς το αυτοπαλινδρομούμενο τμήμα, που απαιτείται από την εκφώνηση, παίρνουμε τα τρία παρακάτω μοντέλα:

ΠΙΝΑΚΑΣ 7: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στη εταιρεία MSFT για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 3

Model 1				
Κώδικας που χρησιμοποιήθηκε	fit3=arima(log_MSFT_Data,c(3,1,1))			
Συντελεστές				
	AR 1	AR 2	AR 3	MA 1
	0.7825	-0.0204	-0.0301	-0.7765

ΠΙΝΑΚΑΣ 8: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στη εταιρεία MSFT για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 7

Model 2								
Κώδικας που χρησιμοποιήθηκε	fit7=arima(log_MSFT_Data,c(7,1,1))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	MA 1
	0.4762	-0.0197	-0.0178	-0.0188	-0.0148	-0.0394	-0.0127	-0.4729

ΠΙΝΑΚΑΣ 9: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στη εταιρεία MSFT για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 14

Model 3								
Κώδικας που χρησιμοποιήθηκε	fit14=arima(log_MSFT_Data,c(14,1,1))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	AR 8
	-0.2151	-0.0159	-0.0305	-0.0358	-0.0318	-0.0561	-0.0366	-0.0409
	AR 9	AR 10	AR 11	AR 12	AR 13	AR 14	MA 1	
	-0.0392	0.0224	0.0582	0.0143	-0.0083	0.0426	0.2177	

Έλεγχος προσαρμογής και σύγκριση των μοντέλων

Τα μοντέλα στην περίπτωση αυτή, λόγω του μικρού εύρους τιμών της χρονοσειράς, έχουν ικανοποιητική προσέγγιση σε όλο το μήκος του διαστήματος ελέγχου. Για τη διάκριση μεταξύ αυτών και την επιλογή του καλύτερου βλέπουμε και πάλι τα μέσα τετραγωνικά σφάλματα:

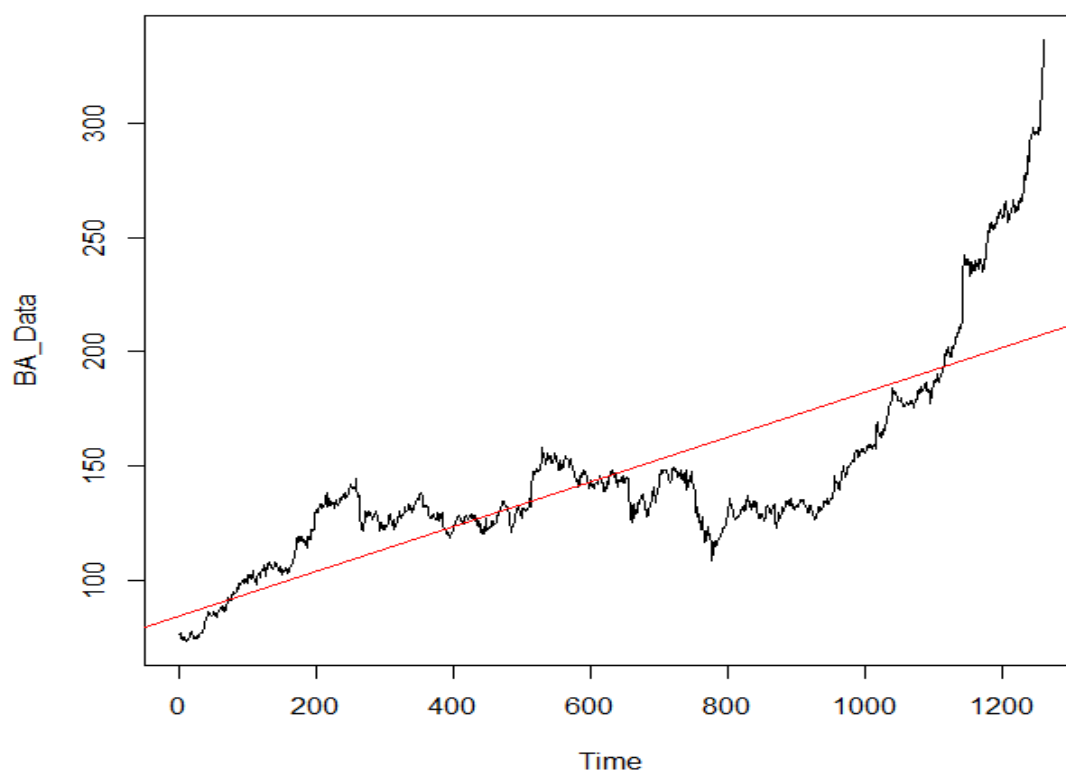
```
> RMSE3
[1] 2.183079
> RMSE7
[1] 2.193999
> RMSE14
[1] 1.967471
```


Το μοντέλο με υστέρηση 14 είναι το καλύτερο, ακολουθεί το μοντέλο με υστέρηση 3 και τέλος το μοντέλο με υστέρηση 7 ημερών.

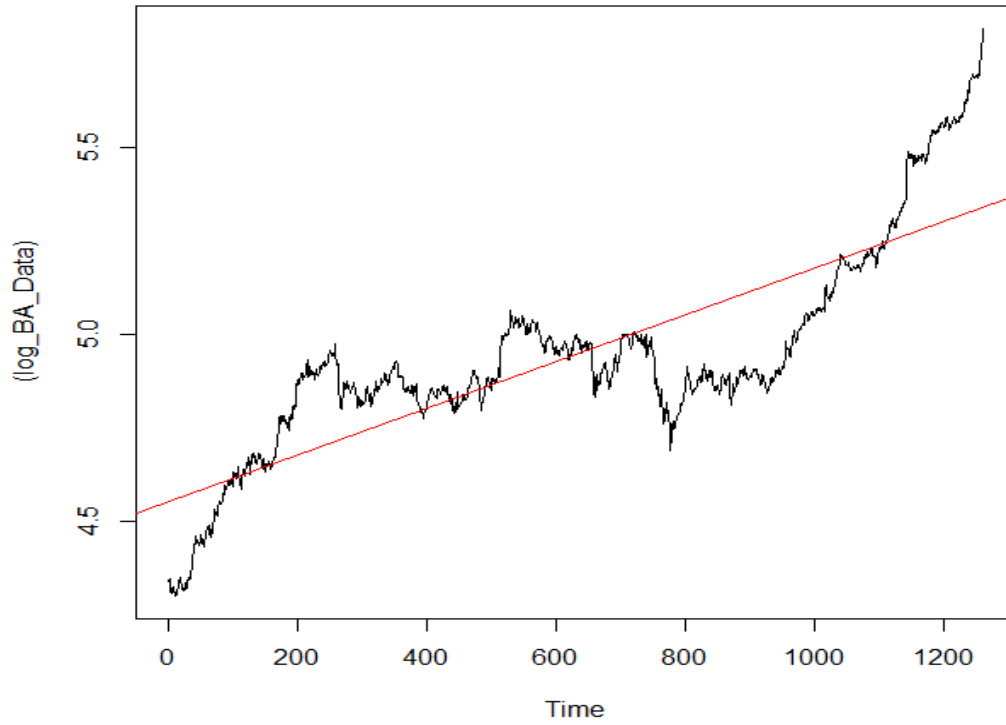
3.4 Πρόβλεψη της τιμής κλεισίματος της εταιρείας ΒΑ

Προετοιμασία για την εφαρμογή του μοντέλου

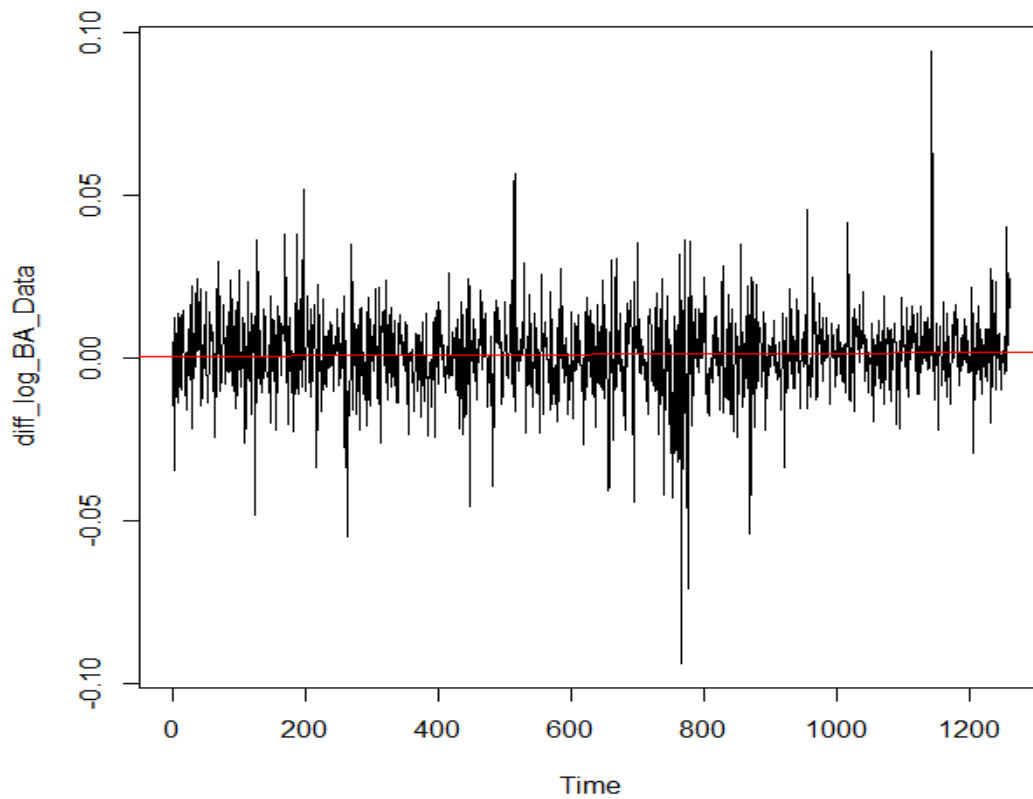
Όπως και πριν, βλέπουμε το γράφημα της χρονοσειράς (Γράφημα 20), βλέπουμε την ίδια χρονοσειρά απαλλαγμένη από τυχαίες διακυμάνσεις (Γράφημα 21) και κάτω επίσης την ίδια χρονοσειρά απαλλαγμένη και από τυχαίες διακυμάνσεις και από τάση (Γράφημα 22).



Γράφημα 20 :Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας JMP



Γράφημα 21: Χρονοσειρά με τάση της εταιρείας BA κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων



Γράφημα 22: Χρονοσειρά χωρίς τάση της εταιρείας BA

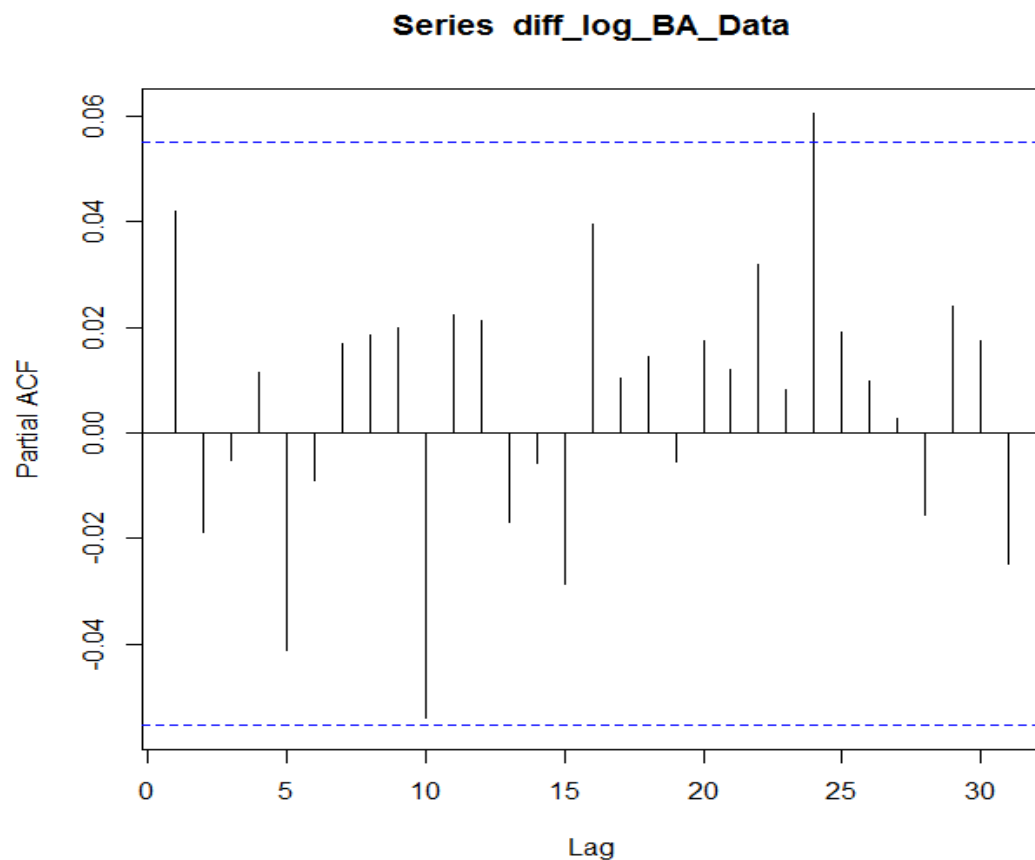
Η χρονοσειρά αυτή και πάλι δεν εμφανίζει σαφή περιοδικότητα, παρ' όλα αυτά βλέπουμε πως δομικά διαφέρει από τις προηγούμενες 3. Βλέπουμε πως ενώ αρχικά μέχρι την ημέρα 200 περίπου, είχε αύξουσα τάση, από την ημέρα 200 έως περίπου της ημέρα 800 είναι σχεδόν στάσιμη, ενώ από εκεί και πέρα συνεχίζει μια έντονη ανοδική πορεία. Αυτή η ιδιομορφία ίσως αποπροσανατολίσει το μοντέλο.

Το μοντέλο

Όπως και στην προηγούμενη περίπτωση, θα χρησιμοποιήσουμε ένα μοντέλο $ARIMA(p,d,q)$.

Από εδώ και πέρα θα εκτιμούμε μόνο την τάξη q του MA τμήματος, καθώς η τάξη του αντίστοιχου αυτοπαλινδρομούμενου τμήματος είναι ορισμένη από την εκφώνηση (3,7 ή 14)

Το γράφημα λοιπόν των μερικών αυτοσυσχετίσεων, δίνεται παρακάτω (Γράφημα 23):



Γράφημα 23: Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία BA

Βλέπουμε πως υπάρχει αναστροφή για υστέρηση 2, επομένως η προτεινόμενη τάξη του MA τμήματος του μοντέλου θα είναι $q=1$.

Εφαρμογή του μοντέλου

Με βάση λοιπόν τα παραπάνω, το MA τμήμα έχει τάξη 1 και η τάση απαλείφθηκε με την εφαρμογή πρώτων διαφορών ($d=1$).

Έτσι, τα μοντέλα που προκύπτουν για αυτές τις παραμέτρους, σε συνδυασμό με υστερήσεις 3, 7 και 14 ως προς το αυτοπαλινδρομούμενο τμήμα, που απαιτείται από την εκφώνηση, παίρνουμε τα τρία παρακάτω μοντέλα:

ΠΙΝΑΚΑΣ 10: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BA για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 3

Model 1				
Κώδικας που χρησιμοποιήθηκε	fit3=arima(log_BA_Data,c(3,1,1))			
Συντελεστές				
	AR 1	AR 2	AR 3	MA 1
	0.0246	-0.0103	0.0014	0.0253

ΠΙΝΑΚΑΣ 11: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BA για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 7

Model 2								
Κώδικας που χρησιμοποιήθηκε	fit7=arima(log_BA_Data,c(7,1,1))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	MA 1
	0.6311	-0.0398	0.0069	0.0201	-0.0455	0.0171	0.0301	-0.5814

ΠΙΝΑΚΑΣ 12: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BA για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 14

Model 3								
Κώδικας που χρησιμοποιήθηκε	fit14=arima(log_BA_Data,c(14,1,1))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	AR 8
	0.0255	-0.0077	-0.0002	0.0202	-0.0356	-0.0008	0.0239	0.0236
	AR 9	AR 10	AR 11	AR 12	AR 13	AR 14	MA 1	
	0.0308	-0.048	0.0275	0.0296	-0.0100	0.0019	0.026	

Έλεγχος προσαρμογής και σύγκριση των μοντέλων

Με χρήση του αλγορίθμου μας, βλέπουμε παρακάτω τις συγκρίσεις μεταξύ των πραγματικών και προβλεπόμενων τιμών από τα μοντέλα:

1ο μοντέλο

> COMPARE3

Original Data	Obs 1211	Obs 1212	Obs 1213	Obs 1214	Obs 1215	Obs 1216
Model Prediction	258.5000	262.6300	261.7500	264.0700	266.1300	265.5700
Original Data	Obs 1217	Obs 1218	Obs 1219	Obs 1220	Obs 1221	Obs 1222
Model Prediction	257.8424	257.8633	257.8717	257.8718	257.8715	257.8715
Original Data	Obs 1223	Obs 1224	Obs 1225	Obs 1226	Obs 1227	Obs 1228
Model Prediction	262.7000	260.8500	262.4200	261.7600	262.8600	263.7000
Original Data	Obs 1229	Obs 1230	Obs 1231	Obs 1232	Obs 1233	Obs 1234
Model Prediction	257.8715	257.8715	257.8715	257.8715	257.8715	257.8715
Original Data	Obs 1235	Obs 1236	Obs 1237	Obs 1238	Obs 1239	Obs 1240
Model Prediction	262.2600	264.6300	266.9900	265.1500	265.8800	265.5800
Original Data	Obs 1241	Obs 1242	Obs 1243	Obs 1244	Obs 1245	Obs 1246
Model Prediction	257.8715	257.8715	257.8715	257.8715	257.8715	257.8715
Original Data	Obs 1247	Obs 1248	Obs 1249	Obs 1250	Obs 1251	Obs 1252
Model Prediction	267.9900	269.3000	276.8000	271.3800	277.9700	275.5400
Original Data	Obs 1253	Obs 1254	Obs 1255	Obs 1256	Obs 1257	Obs 1258
Model Prediction	278.2700	281.9700	285.9000	283.1600	289.9400	291.8400
Original Data	Obs 1259	Obs 1260				
Model Prediction	293.8800	293.9400	296.1400	297.2500	297.9000	295.0300
	257.8715	257.8715	257.8715	257.8715	257.8715	257.8715
	295.1000	295.3600	295.6200	296.3500	294.9100	296.8400
	257.8715	257.8715	257.8715	257.8715	257.8715	257.8715
	297.8000	296.6700	308.8400	310.1500	318.4300	320.2600
	257.8715	257.8715	257.8715	257.8715	257.8715	257.8715
	328.1200	336.2100				
	257.8715	257.8715				

2ο μοντέλο

> COMPARE7

Original Data	Obs 1211	Obs 1212	Obs 1213	Obs 1214	Obs 1215	Obs 1216
Model Prediction	258.5000	262.6300	261.7500	264.0700	266.1300	265.5700
Original Data	Obs 1217	Obs 1218	Obs 1219	Obs 1220	Obs 1221	Obs 1222
Model Prediction	258.1256	258.2556	258.2495	258.0883	257.9966	258.017
Original Data	Obs 1223	Obs 1224	Obs 1225	Obs 1226	Obs 1227	Obs 1228
Model Prediction	262.7000	260.8500	262.4200	261.7600	262.8600	263.7000
Original Data	Obs 1229	Obs 1230	Obs 1231	Obs 1232	Obs 1233	Obs 1234
Model Prediction	257.9884	257.9731	257.9754	257.9793	257.9737	257.9685
Original Data	Obs 1235	Obs 1236	Obs 1237	Obs 1238	Obs 1239	Obs 1240
Model Prediction	262.2600	264.6300	266.9900	265.1500	265.8800	265.5800
Original Data	Obs 1241	Obs 1242	Obs 1243	Obs 1244	Obs 1245	Obs 1246
Model Prediction	257.9666	257.9645	257.9626	257.9618	257.9616	257.9613
Original Data	Obs 1247	Obs 1248	Obs 1249	Obs 1250	Obs 1251	Obs 1252
Model Prediction	267.990	269.3000	276.8000	271.3800	277.9700	275.5400
Original Data	Obs 1253	Obs 1254	Obs 1255	Obs 1256	Obs 1257	Obs 1258
Model Prediction	278.2700	281.9700	285.9000	283.1600	289.9400	291.8400
Original Data	Obs 1259	Obs 1260				
Model Prediction	257.9603	257.9603	257.9603	257.9603	257.9603	257.9603
Original Data	Obs 1247	Obs 1248	Obs 1249	Obs 1250	Obs 1251	Obs 1252
Model Prediction	293.8800	293.9400	296.1400	297.2500	297.9000	295.0300
Original Data	Obs 1253	Obs 1254	Obs 1255	Obs 1256	Obs 1257	Obs 1258
Model Prediction	257.9602	257.9602	257.9602	257.9602	257.9602	257.9602
Original Data	Obs 1259	Obs 1260				
Model Prediction	295.1000	295.3600	295.6200	296.3500	294.9100	296.8400
	257.9602	257.9602	257.9602	257.9602	257.9602	257.9602
	297.8000	296.6700	308.8400	310.1500	318.4300	320.2600
	257.9602	257.9602	257.9602	257.9602	257.9602	257.9602
	328.1200	336.2100				
	257.9602	257.9602				

3ο μοντέλο

```

> COMPARE14
Obs 1211 Obs 1212 Obs 1213 Obs 1214 Obs 1215 Obs 1216
Original Data 258.5000 262.6300 261.7500 264.0700 266.13 265.5700
Model Prediction 258.0246 258.3234 257.9895 258.1162 257.74 258.1991
Obs 1217 Obs 1218 Obs 1219 Obs 1220 Obs 1221 Obs 1222
Original Data 262.7000 260.8500 262.420 261.7600 262.8600 263.700
Model Prediction 257.9654 258.0163 257.808 257.9073 257.9397 257.873
Obs 1223 Obs 1224 Obs 1225 Obs 1226 Obs 1227 Obs 1228
Original Data 262.2600 264.630 266.9900 265.1500 265.8800 265.5800
Model Prediction 257.9073 257.901 257.9237 257.8791 257.8971 257.8965
Obs 1229 Obs 1230 Obs 1231 Obs 1232 Obs 1233 Obs 1234
Original Data 267.9900 269.3000 276.80 271.380 277.9700 275.5400
Model Prediction 257.9038 257.8938 257.89 257.899 257.8951 257.8943
Obs 1235 Obs 1236 Obs 1237 Obs 1238 Obs 1239 Obs 1240
Original Data 278.2700 281.9700 285.900 283.1600 289.9400 291.8400
Model Prediction 257.8933 257.8971 257.895 257.8943 257.8944 257.8952
Obs 1241 Obs 1242 Obs 1243 Obs 1244 Obs 1245 Obs 1246
Original Data 293.8800 293.9400 296.1400 297.250 297.9000 295.0300
Model Prediction 257.8953 257.8943 257.8948 257.895 257.8949 257.8945
Obs 1247 Obs 1248 Obs 1249 Obs 1250 Obs 1251 Obs 1252
Original Data 295.1000 295.3600 295.6200 296.3500 294.9100 296.8400
Model Prediction 257.8948 257.8949 257.8948 257.8947 257.8947 257.8948
Obs 1253 Obs 1254 Obs 1255 Obs 1256 Obs 1257 Obs 1258
Original Data 297.8000 296.6700 308.8400 310.1500 318.4300 320.2600
Model Prediction 257.8948 257.8947 257.8948 257.8948 257.8948 257.8948
Obs 1259 Obs 1260
Original Data 328.1200 336.2100
Model Prediction 257.8948 257.8948

```

Εδώ βλέπουμε, πως αν και η πρόβλεψη του αμέσως επόμενου βήματος είναι εξαιρετικά καλή, σε βάθος χρόνου το μοντέλο αποτυγχάνει πλήρως στην προσπάθειά του να προσεγγίσει την πραγματικότητα, καθώς υπάρχουν (προς το τέλος), αποκλίσεις σχεδόν 80 μονάδων. Για την επιλογή του καλύτερου, υπολογίζουμε τα RMSE:

```

> RMSE3
[1] 22.58198
> RMSE7
[1] 22.53026
> RMSE14
[1] 22.56647

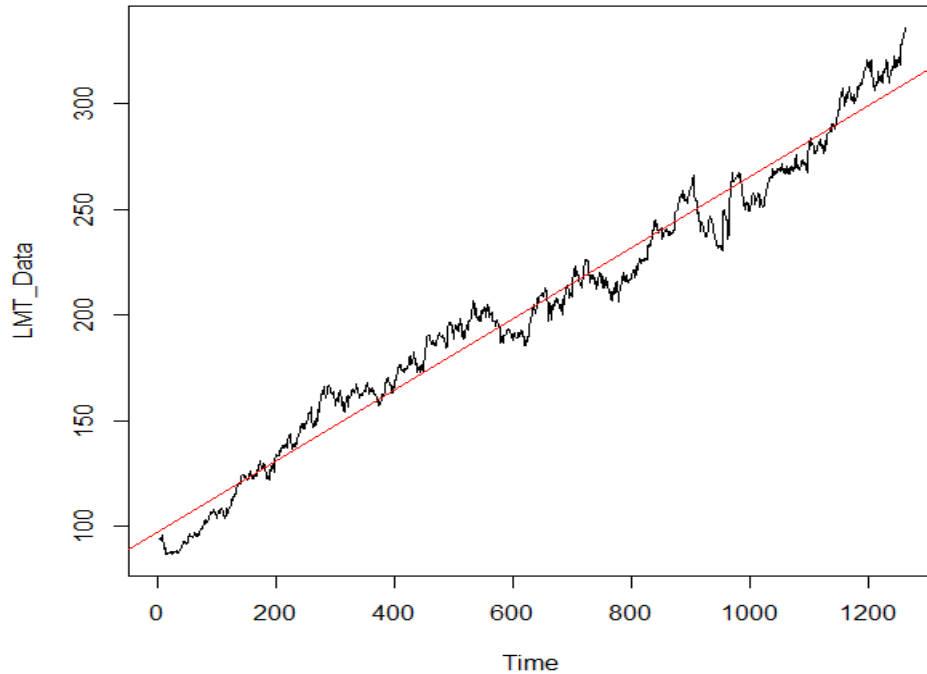
```

Εδώ, το καλύτερο είναι το μοντέλο με υστέρηση 7, ακολουθεί το μοντέλο με υστέρηση 14 και τέλος αυτό με υστέρηση 3.

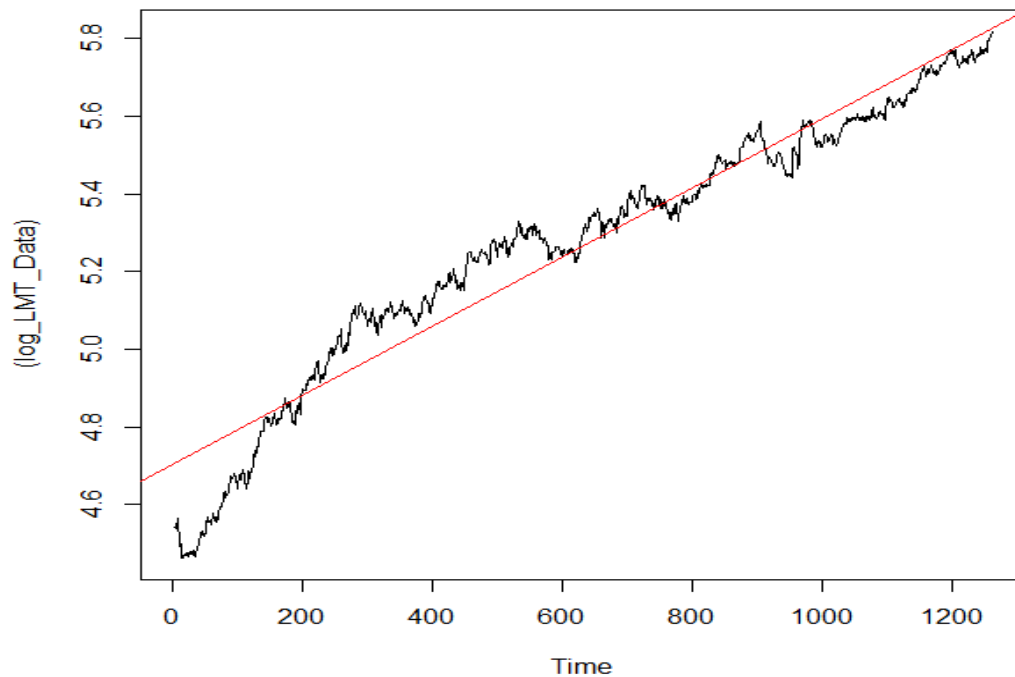
3.5 Πρόβλεψη της τιμής κλεισίματος της εταιρείας LMT

Προετοιμασία για την εφαρμογή του μοντέλου

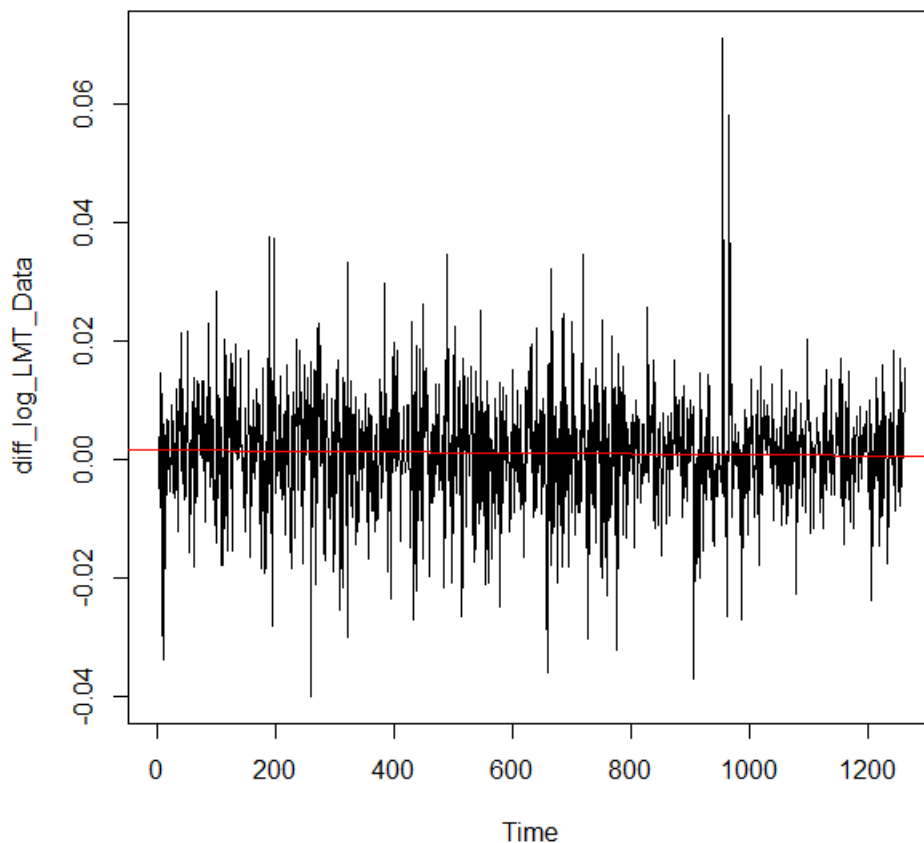
και πάλι βλέπουμε το γράφημα της χρονοσειράς (Γράφημα 24), βλέπουμε τη χρονοσειρά απαλλαγμένη από τυχαίες διακυμάνσεις (Γράφημα 25) και κάτω την ίδια χρονοσειρά απαλλαγμένη και από τάση (Γράφημα 26).



Γράφημα 24 :Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας LMT



Γράφημα 25: Χρονοσειρά με τάση της εταιρείας LMT κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων



Γράφημα 26: Χρονοσειρά χωρίς τάση της εταιρείας LMT

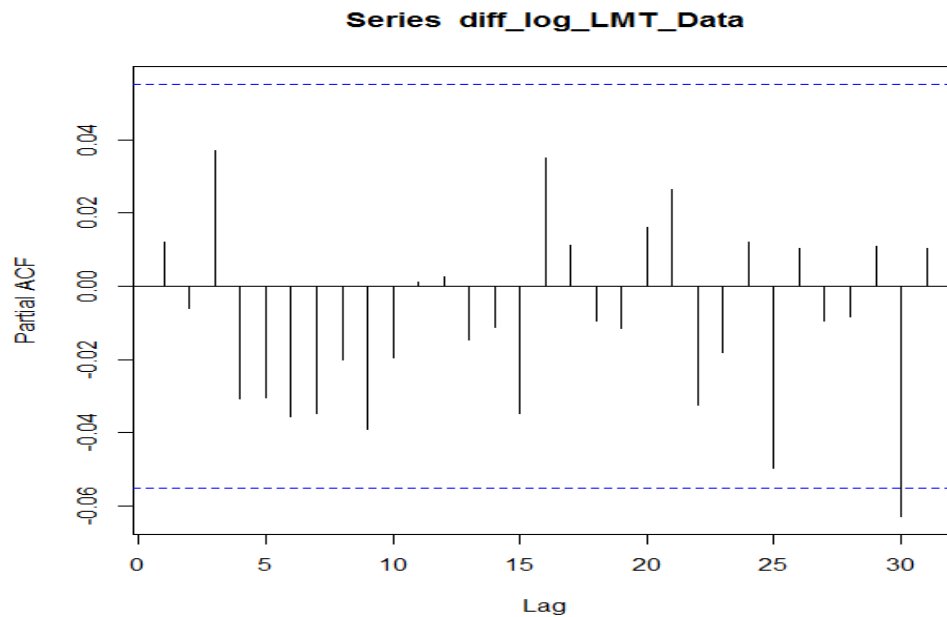
Η μαύρη γραμμή δηλώνει τις τιμές της χρονοσειράς, ενώ η κόκκινη δηλώνει τη γραμμική εκτίμηση της τάσης που εμφανίζεται.

Και πάλι δεν υπάρχει ένδειξη περί περιοδικότητας. Εδώ αξίζει να τονίσουμε το πόσο αποδοτική φαίνεται η μέθοδος των πρώτων διαφορών ως μέθοδος απαλοιφής τάσης. Βλέπουμε πως ακόμα και στο πέμπτο παράδειγμα με πραγματικά δεδομένα, η λήψη πρώτων διαφορών καταστρέφει ολοκληρωτικά την τάση.

Το μοντέλο

Το μοντέλο που θα χρησιμοποιήσουμε θα είναι και πάλι ένα μοντέλο ARIMA(p,d,q).

Το γράφημα λοιπόν των μερικών αυτοσυσχετίσεων, δίνεται παρακάτω (Γράφημα 27):



Γράφημα 27: Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία LMT

Βλέπουμε πως η αυτοσυσχέτιση στο δεύτερο βήμα γίνεται αρνητική, επομένως η τάξη του MA μοντέλου είναι $q=1$.

Εφαρμογή του μοντέλου

Με βάση λοιπόν τα παραπάνω, το MA τμήμα έχει τάξη 1 και η τάση απαλείφθηκε με την εφαρμογή πρώτων διαφορών ($d=1$).

Έτσι, τα μοντέλα που προκύπτουν για αυτές τις παραμέτρους, σε συνδυασμό με υστερήσεις 3, 7 και 14 ως προς το αυτοπαλινδρομούμενο τμήμα, που απαιτείται από την εκφώνηση, παίρνουμε τα τρία παρακάτω μοντέλα:

ΠΙΝΑΚΑΣ 13: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία LMT για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 3

Model 1				
Κώδικας που χρησιμοποιήθηκε	fit3=arima(log_LMT_Data,c(3,1,1))			
Συντελεστές				
	AR 1	AR 2	AR 3	MA 1
	-0.2353	0.0083	0.0486	0.2577

ΠΙΝΑΚΑΣ 14: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία LMT για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 7

Model 2								
Κώδικας που χρησιμοποιήθηκε	fit7=arima(log_LMT_Data,c(7,1,1))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	MA 1
	0.5564	-0.0090	0.0460	-0.0456	-0.0087	-0.0141	-0.0120	-0.5359

ΠΙΝΑΚΑΣ 15: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία LMT για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 14

Model 3								
Κώδικας που χρησιμοποιήθηκε	fit14=arima(log_LMT_Data,c(14,1,1))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	AR 8
	0.0088	0.0024	0.0472	-0.0197	-0.0196	-0.0233	-0.0234	-0.0087
	AR 9	AR 10	AR 11	AR 12	AR 13	AR 14	MA 1	
	-0.0280	-0.0071	0.0134	0.0156	-0.0015	0.0021	0.0116	

Έλεγχος προσαρμογής και σύγκριση των μοντέλων

Βλέπουμε πως επίσης υπάρχει προσέγγιση στα αρχικά βήματα, η οποία διατηρείται σχετικά σε όλο το διάστημα ελέγχου, λόγω του εύρους των τιμών του συγκεκριμένου δείκτη.

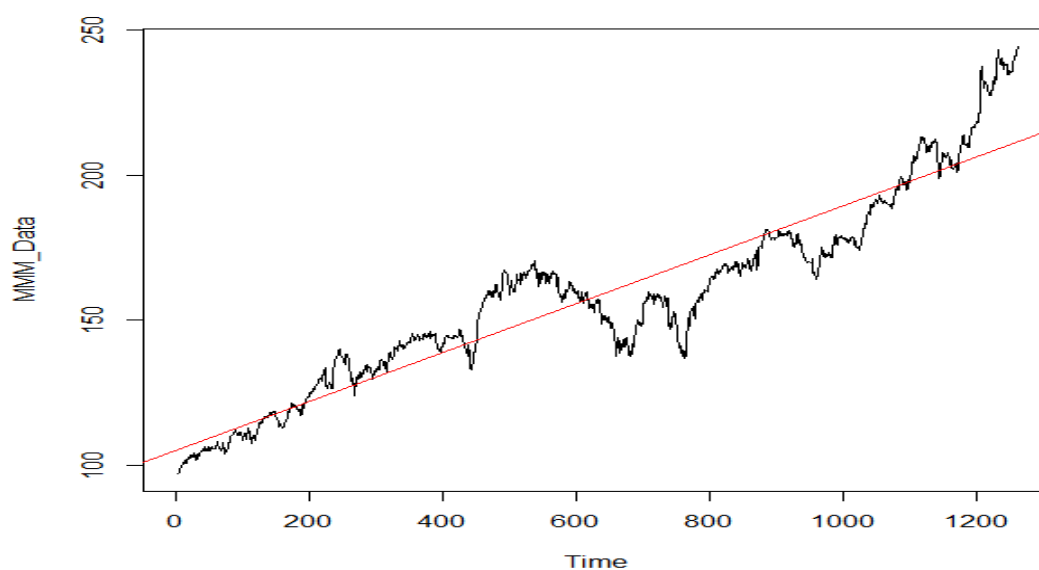
```
> RMSE3
[1] 8.083816
> RMSE7
[1] 7.546243
> RMSE14
[1] 7.621558
```

Εδώ, το μοντέλο με υστέρηση 7 είναι αυτό που επιτυγχάνει την καλύτερη προσέγγιση. Ακολουθεί αυτό από αυτό με υστέρηση 14, ενώ τελευταίο είναι το μοντέλο που κάνει πρόβλεψη με χρήση των 3 προηγούμενων ημερών.

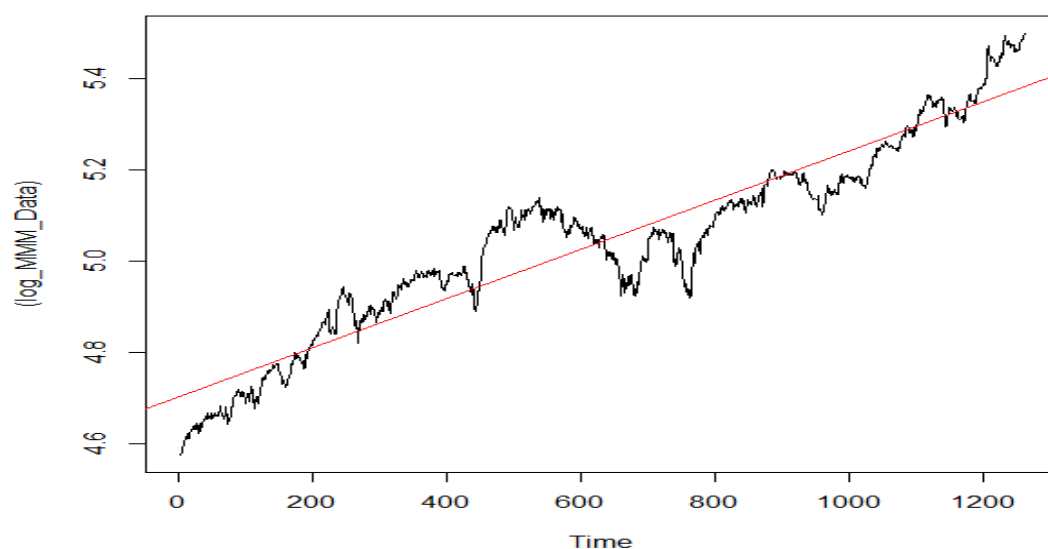
3.6 Πρόβλεψη της τιμής κλεισίματος της εταιρείας MMM

Προετοιμασία για την εφαρμογή του μοντέλου

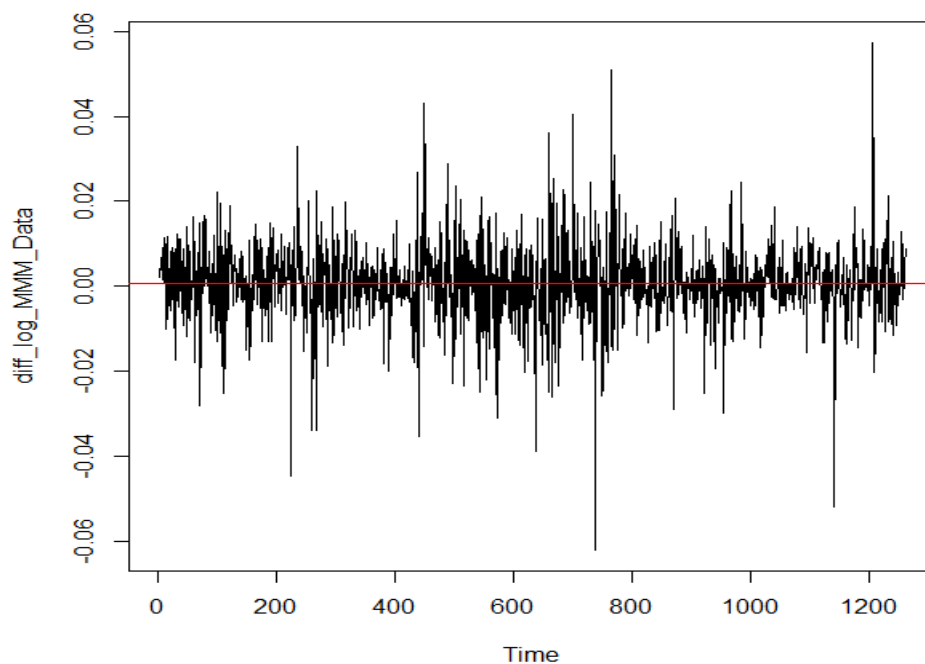
Όπως και πριν, βλέπουμε το γράφημα της χρονοσειράς (Γράφημα 28), βλέπουμε την ίδια χρονοσειρά απαλλαγμένη από τυχαίες διακυμάνσεις (Γράφημα 29) και κάτω επίσης την ίδια χρονοσειρά απαλλαγμένη και από τυχαίες διακυμάνσεις και από τάση (Γράφημα 30).



Γράφημα 28 :Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας MMM



Γράφημα 29: Χρονοσειρά με τάση της εταιρείας MMM κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων



Γράφημα 30: Χρονοσειρά χωρίς τάση της εταιρείας MMM

Η μαύρη γραμμή δηλώνει τις τιμές της χρονοσειράς, ενώ η κόκκινη δηλώνει τη γραμμική εκτίμηση της τάσης που εμφανίζεται.

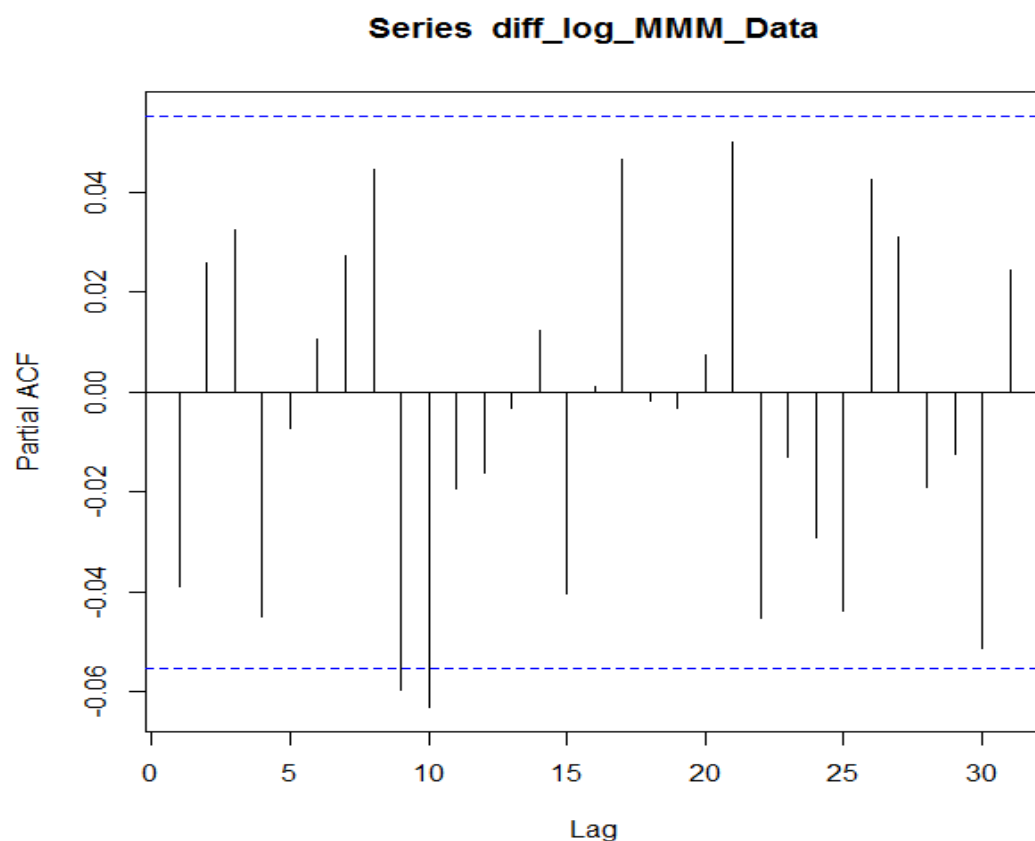
Κι εδώ δεν υπάρχει ένδειξη περιοδικότητας. Η συγκεκριμένη χρονοσειρά εμφανίζει επίσης αύξουσα τάση όπως οι περισσότερες μέχρι τώρα, αλλά έχει εμφανώς μεγαλύτερες αποκλίσεις κατά χρονικά διαστήματα απ ότι οι μέχρι τώρα χρονοσειρές.

Το μοντέλο

Όπως και στην προηγούμενη περίπτωση, θα χρησιμοποιήσουμε ένα μοντέλο $ARIMA(p,d,q)$.

Από εδώ και πέρα θα εκτιμούμε μόνο την τάξη q του MA τμήματος, καθώς η τάξη του αντίστοιχου αυτοπαλινδρομούμενου τμήματος είναι ορισμένη από την εκφώνηση (3,7 ή 14)

Το γράφημα λοιπόν των μερικών αυτοσυσχετίσεων, δίνεται παρακάτω (Γράφημα 31)



Γράφημα 31: Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία MMM

Βλέπουμε πως η πρώτη υστέρηση μας δίνει αρνητική αυτοσυσχέτιση, επομένως η διαδικασία κινούμενου μέσου του μοντέλου, θα έχει μηδενική τάξη.

Εφαρμογή του μοντέλου

Με βάση λοιπόν τα παραπάνω, το MA τμήμα έχει τάξη 0 και η τάση απαλείφθηκε με την εφαρμογή πρώτων διαφορών ($d=1$).

Έτσι, τα μοντέλα που προκύπτουν για αυτές τις παραμέτρους, σε συνδυασμό με υστερήσεις 3, 7 και 14 ως προς το αυτοπαλινδρομούμενο τμήμα, που απαιτείται από την εκφώνηση, παίρνουμε τα τρία παρακάτω μοντέλα:

ΠΙΝΑΚΑΣ 16: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία MMM για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 3

Model 1			
Κώδικας που χρησιμοποιήθηκε	fit3=arima(log_MMM_Data,c(3,1,0))		
Συντελεστές			
	AR 1	AR 2	AR 3
	-0.0331	0.0330	0.0381

ΠΙΝΑΚΑΣ 17: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία MMM για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 7

Model 2							
Κώδικας που χρησιμοποιήθηκε	fit7=arima(log_MMM_Data,c(7,1,0))						
Συντελεστές							
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7
	-0.0322	0.0351	0.0376	-0.0410	-0.0020	0.0174	0.0325

ΠΙΝΑΚΑΣ 18: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία MMM για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 14

Model 3								
Κώδικας που χρησιμοποιήθηκε	fit14=arima(log_MMM_Data,c(14,1,0))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	AR 8
	-0.0351	0.0378	0.0411	-0.0361	-0.0050	0.0144	0.0368	0.0494
	AR 9	AR 10	AR 11	AR 12	AR 13	AR 14		
	-0.0552	-0.0570	-0.0145	-0.0104	0.0039	0.0184		

Έλεγχος προσαρμογής και σύγκριση των μοντέλων

Τα μοντέλα και πάλι προσεγγίζουν σχετικά καλά την αρχική χρονοσειρά. Για την ανάδειξη του καλύτερου, υπολογίστηκαν τα μέσα τετραγωνικά σφάλματα:

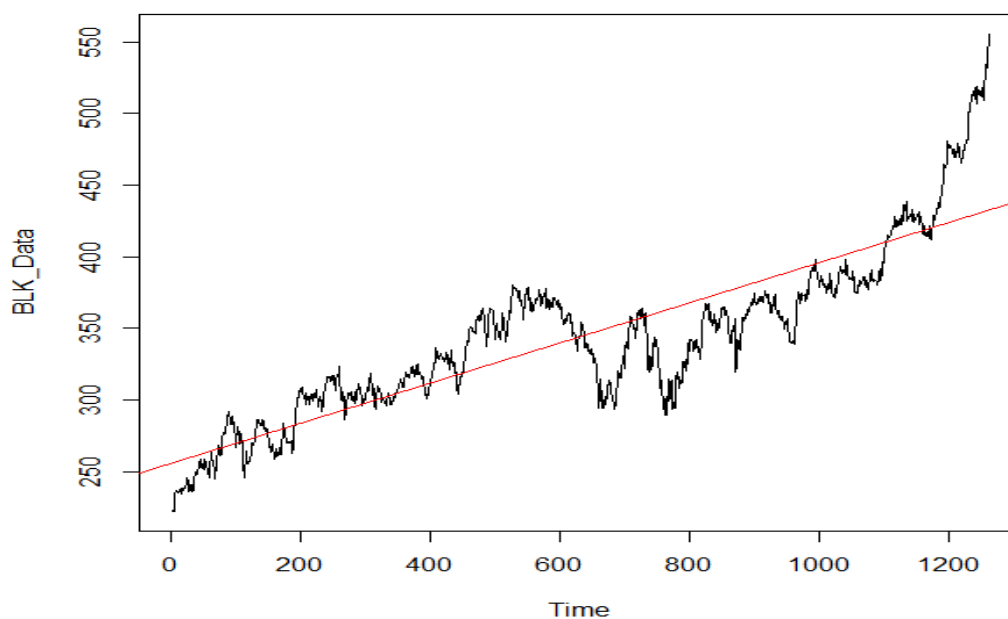
```
> RMSE3
[1] 5.022173
> RMSE7
[1] 4.55676
> RMSE14
[1] 5.050006
```

Το μοντέλο με υστέρηση 7 κάνει την καλύτερη προσαρμογή, ακολουθεί το μοντέλο με υστέρηση 3 και τέλος το μοντέλο υστέρησης 14.

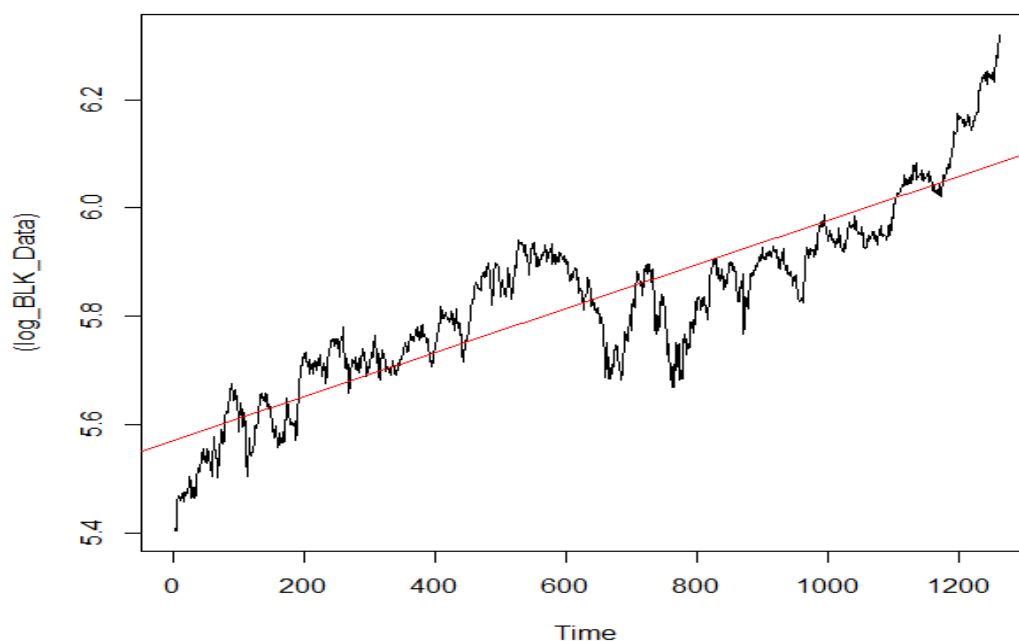
3.7 Πρόβλεψη της τιμής κλεισίματος της εταιρείας BLK

Προετοιμασία για την εφαρμογή του μοντέλου

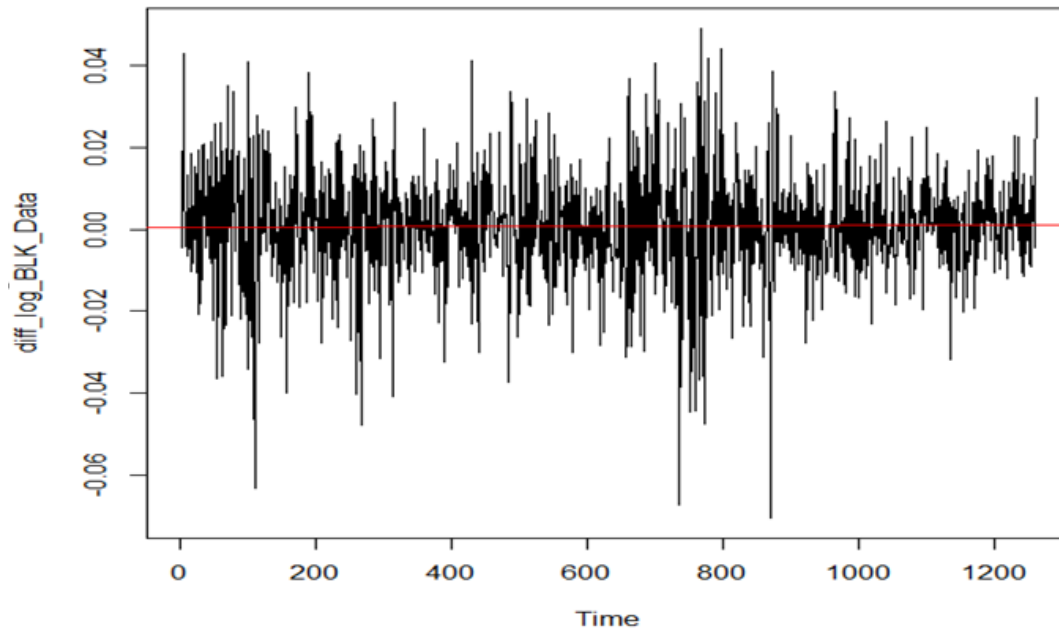
Με όμοια δομή, παρουσιάζουμε την αρχική χρονοσειρά αριστερά (Γράφημα 32), η χρονοσειρά απαλλαγμένη από τυχαίες διακυμάνσεις δεξιά (Γράφημα 33) και κάτω η ίδια χρονοσειρά απαλλαγμένη και από διακυμάνσεις και από τάση (Γράφημα 34):



Γράφημα 32 :Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας BLK



Γράφημα 33: Χρονοσειρά με τάση της εταιρείας BLK κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων



Γράφημα 34: Χρονοσειρά χωρίς τάση της εταιρείας BLK

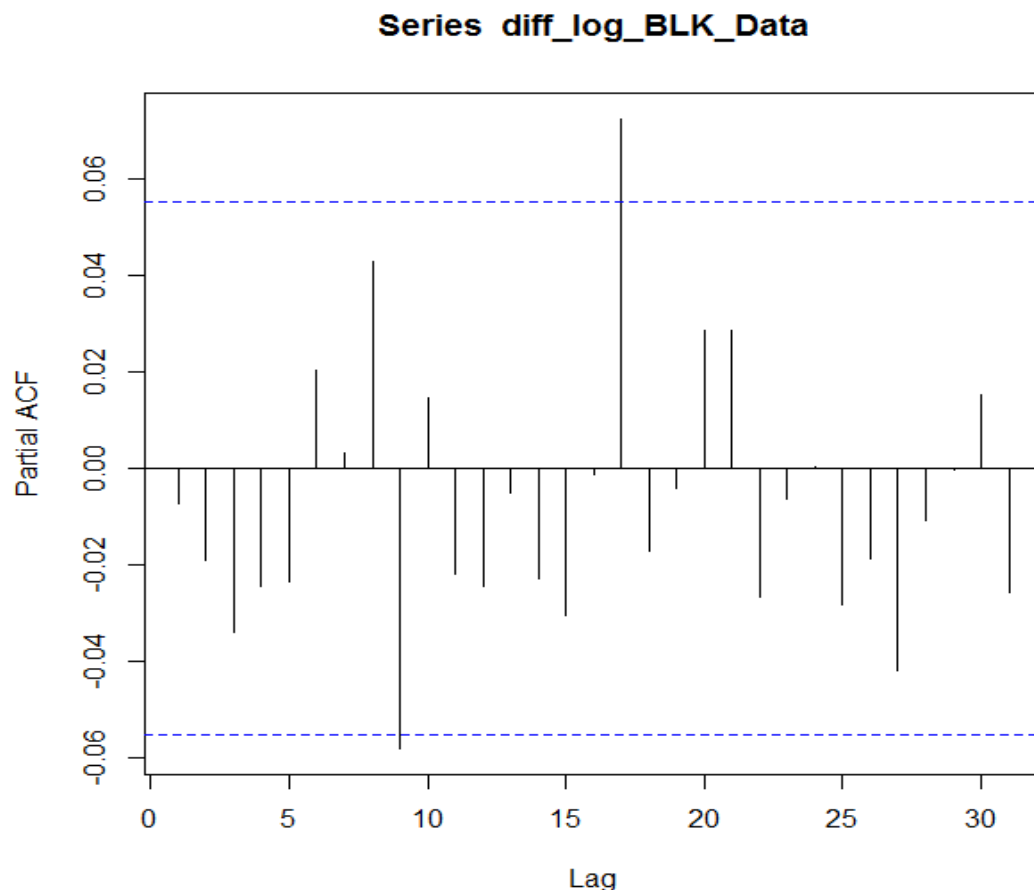
Η μαύρη γραμμή δηλώνει τις τιμές της χρονοσειράς, ενώ η κόκκινη δηλώνει τη γραμμική εκτίμηση της τάσης που εμφανίζεται.

Επίσης δεν λαμβάνουμε πληροφορία ως προς την περιδικότητα και θα μπορούσαμε να παρατηρήσουμε πως υπήρχε μια περίοδος, περίπου από την ημέρα 600 έως την ημέρα 900 από την αρχή του χρόνου, που η χρονοσειρά επηρεάστηκε από κάτι εξωγενές καθώς εμφανίζει δομική διαφορά από πριν και μετά.

Το μοντέλο

Όπως και στην προηγούμενη περίπτωση, θα χρησιμοποιήσουμε ένα μοντέλο ARIMA(p,d,q).

Για την τάξη του τμήματος κινούμενου μέσου, κατασκευάζουμε το γράφημα των μερικών αυτοσυσχετίσεων, το οποίο δίνεται παρακάτω (Γράφημα 35):



Γράφημα 35: Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία BLK

Η πρώτη μερική αυτοσυσχέτιση είναι αρνητική, επομένως επιλέγουμε τάξη του $q=0$.

Εφαρμογή του μοντέλου

Με βάση λοιπόν τα παραπάνω, το MA τμήμα έχει τάξη 0 και η τάση απαλείφθηκε με την εφαρμογή πρώτων διαφορών ($d=1$).

Έτσι, τα μοντέλα που προκύπτουν για αυτές τις παραμέτρους, σε συνδυασμό με υστερήσεις 3, 7 και 14 ως προς το αυτοπαλινδρομούμενο τμήμα, που απαιτείται από την εκφώνηση, παίρνουμε τα τρία παρακάτω μοντέλα:

ΠΙΝΑΚΑΣ 19: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BLK για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 3

Model 1			
Κώδικας που χρησιμοποιήθηκε	fit3=arima(log_BLK_Data,c(3,1,0))		
Συντελεστές			
	AR 1	AR 2	AR 3
	-0.0053	-0.0165	-0.0315

ΠΙΝΑΚΑΣ 20: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BLK για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 7

Model 2							
Κώδικας που χρησιμοποιήθηκε	fit7=arima(log_BLK_Data,c(7,1,0))						
Συντελεστές							
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7
	-0.0061	-0.0169	-0.0312	-0.0214	-0.0206	0.0235	0.0061

ΠΙΝΑΚΑΣ 21: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία BLK για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 14

Model 3								
Κώδικας που χρησιμοποιήθηκε	fit14=arima(log_BLK_Data,c(14,1,0))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	AR 8
	-0.0028	-0.0196	-0.0295	-0.0205	-0.0205	0.0241	0.0050	0.0449
	AR 9	AR 10	AR 11	AR 12	AR 13	AR 14		
	-0.0572	0.0165	-0.0201	-0.0219	-0.0020	-0.0199		

Έλεγχος προσαρμογής και σύγκριση των μοντέλων

Βλέπουμε πως επίσης υπάρχει σαφής προσέγγιση των πραγματικών τιμών μέσω των μοντέλων. Προσπαθώντας να επιλέξουμε το καλύτερο, βρίσκουμε και πάλι τα RMSE:

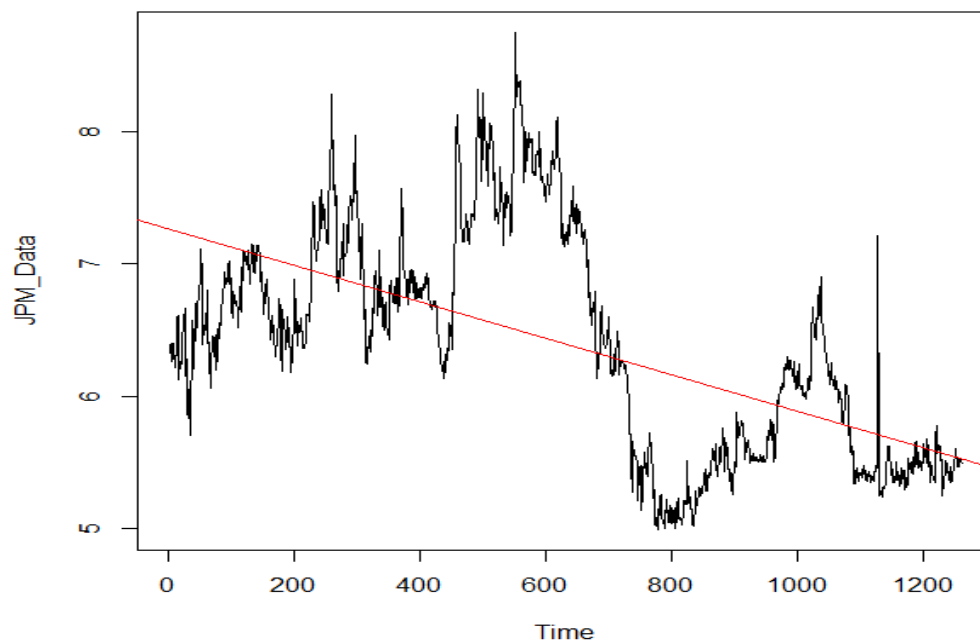
```
> RMSE3
[1] 26.47167
> RMSE7
[1] 26.53052
> RMSE14
[1] 26.3596
```

Το μοντέλο με υστέρηση 7 είναι το καλύτερο, ακολουθεί το μοντέλο με υστέρηση 14 και τέλος βρίσκεται το μοντέλο που κάνει πρόβλεψη με χρήση των τριών προηγούμενων ημερών.

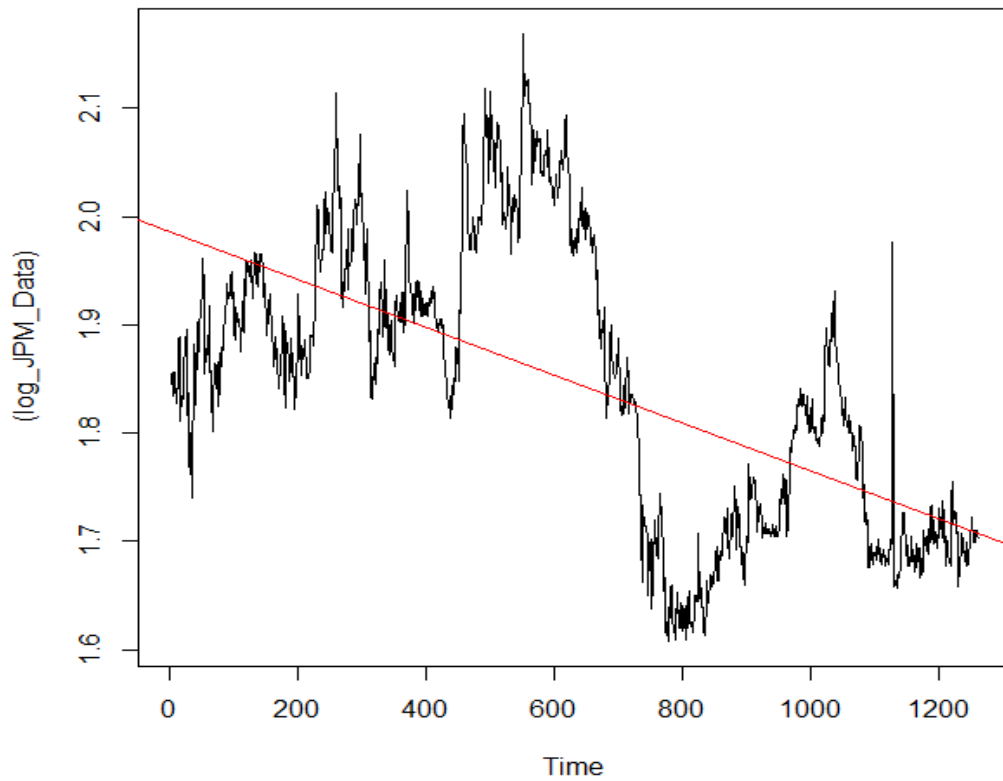
3.8 Πρόβλεψη της τιμής κλεισίματος της εταιρείας JPM

Προετοιμασία για την εφαρμογή του μοντέλου

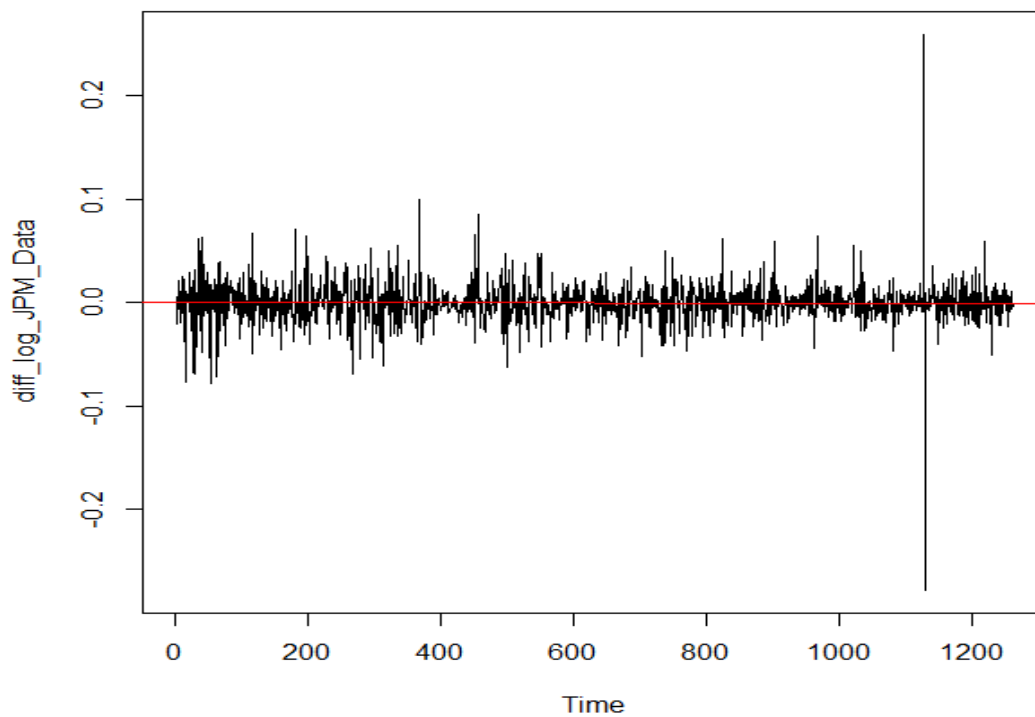
Και πάλι, βλέπουμε το γράφημα της αρχικής χρονοσειράς (Γράφημα 36), την ίδια χρονοσειρά απαλλαγμένη από τυχαίες διακυμάνσεις (Γράφημα 37) και κάτω επίσης την ίδια χρονοσειρά απαλλαγμένη και από τυχαίες διακυμάνσεις και από τάση (Γράφημα 38).



Γράφημα 36 :Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας JMP



Γράφημα 37: Χρονοσειρά με τάση της εταιρείας JMP κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων



Γράφημα 38: Χρονοσειρά χωρίς τάση της εταιρείας JMP

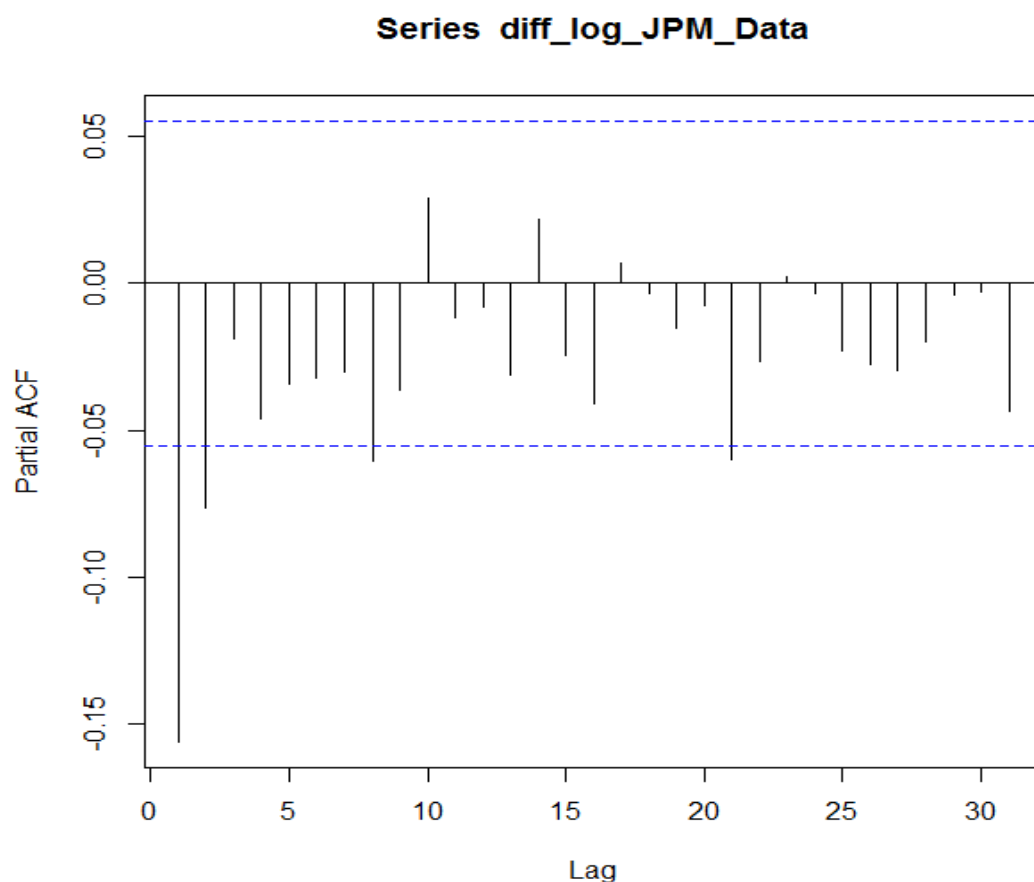
Η μαύρη γραμμή δηλώνει τις τιμές της χρονοσειράς, ενώ η κόκκινη δηλώνει τη γραμμική εκτίμηση της τάσης που εμφανίζεται.

Η χρονοσειρά και πάλι δεν εμφανίζει περιοδικότητα, παρ' όλα αυτά είναι σημαντικά διαφορετική από τις προηγούμενες περιπτώσεις που μελετήθηκαν. Από τις 07/08/2015 βλέπουμε μία πολύ απότομη πτώση, η οποία λήγει περίπου στις 24/03/16. Γενικά πάντως, υπάρχει πολύ αυξημένη «αστάθεια» σε σχέση με τις υπόλοιπες περιπτώσεις, η οποία ενδέχεται να επηρεάσει πολύ την προβλεπτική αξία του μοντέλου.

Το μοντέλο

Όπως και στην προηγούμενη περίπτωση, θα χρησιμοποιήσουμε ένα μοντέλο ARIMA(p,d,q).

Το γράφημα λοιπόν των μερικών αυτοσυσχετίσεων, προς εκτίμηση της παραμέτρου q, δίνεται παρακάτω (Γράφημα 39):



Γράφημα 39: Γράφημα μερικών αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία JPM

Από το γράφημα υποδεικνύεται πως η τάξη του μοντέλου κινούμενου μέσου που περιέχεται στο μοντέλο μας, είναι $q=0$, καθώς η πρώτη αυτοσυσχέτιση είναι αρνητική.

Εφαρμογή του μοντέλου

Με βάση λοιπόν τα παραπάνω, το MA τμήμα έχει τάξη 0 και η τάση απαλείφθηκε με την εφαρμογή πρώτων διαφορών ($d=1$).

Έτσι, τα μοντέλα που προκύπτουν για αυτές τις παραμέτρους, σε συνδυασμό με υστερήσεις 3, 7 και 14 ως προς το αυτοπαλινδρομούμενο τμήμα, που απαιτείται από την εκφώνηση, παίρνουμε τα τρία παρακάτω μοντέλα:

ΠΙΝΑΚΑΣ 22: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία JPM για την πρόβλεψη της επόμενης τιμής με υστέρηση 3

Model 1			
Κώδικας που χρησιμοποιήθηκε	fit3=arima(log_JPM_Data,c(3,1,0))		
Συντελεστές			
	AR 1	AR 2	AR 3
	-0.1692	-0.0793	-0.0189

ΠΙΝΑΚΑΣ 23: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία JPM για την πρόβλεψη της επόμενης τιμής μόνο με υστέρηση 7

Model 2							
Κώδικας που χρησιμοποιήθηκε	fit7=arima(log_JPM_Data,c(7,1,0))						
Συντελεστές							
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7
	-0.1797	-0.0867	-0.0321	-0.0555	-0.0423	-0.0373	-0.0302

ΠΙΝΑΚΑΣ 24: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία JMP για την πρόβλεψη της επόμενης τιμής με υστέρηση 14

Model 3								
Κώδικας που χρησιμοποιήθηκε	fit14=arima(log_JPM_Data,c(14,1,0))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	AR 8
	-0.1761	-0.0888	-0.0347	-0.0618	-0.0471	-0.0428	-0,0441	-0,0649
	AR 9	AR 10	AR 11	AR 12	AR 13	AR 14		
	-0.0331	0.0263	-0.0152	-0.0115	-0.0270	0.0219		

Έλεγχος προσαρμογής και σύγκριση των μοντέλων

Η παρούσα προσέγγιση, φαίνεται εξαιρετικά καλή, γεγονός το οποίο σίγουρα οφείλεται σε μεγάλο βαθμό στο πολύ περιορισμένο εύρος των τιμών του δείκτη. Τα μοντέλα έχουν πετύχει την καλύτερη ακρίβεια σε σχέση με τις προηγούμενες περιπτώσεις, σημειώνοντας πάρα πολύ καλές τιμές RMSE όπως βλέπουμε παρακάτω:

```
> RMSE3
[1] 0.08548606
> RMSE7
[1] 0.08455533
> RMSE14
[1] 0.08928585
```

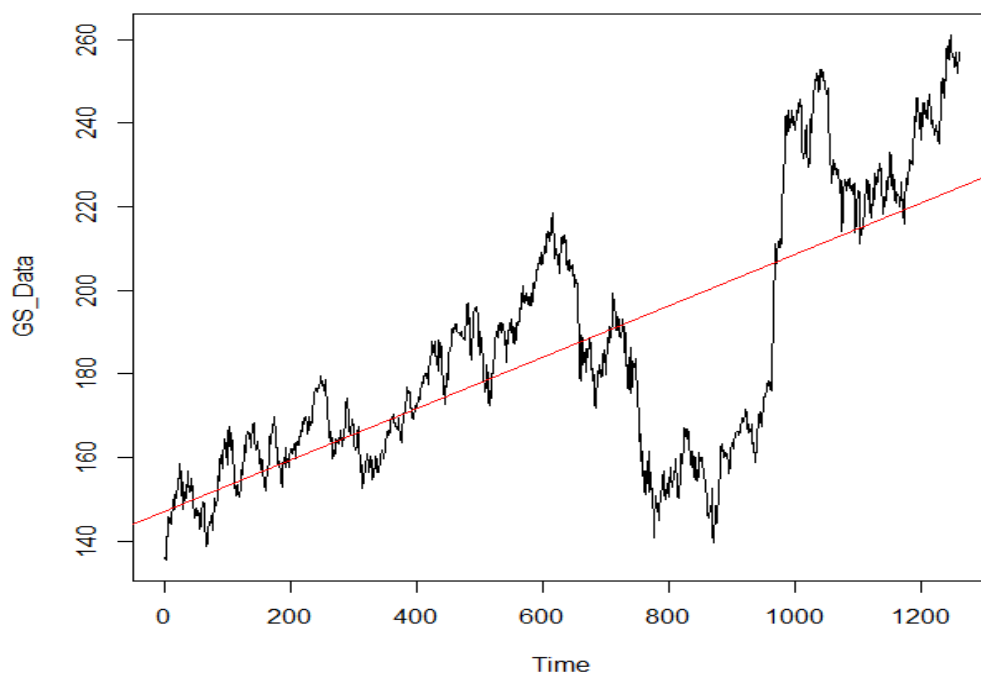
Το RMSE είναι της τάξης του 0.08, ενώ το μοντέλο με υστέρηση 7 είναι το καλύτερο. Ακολουθεί το μοντέλο με υστέρηση 3 και τέλος, ακολουθεί το μοντέλο που κάνει πρόβλεψη με χρήση των 14^{ov} προηγούμενων ημερών.

3.9 Πρόβλεψη της τιμής κλεισίματος της εταιρείας GS

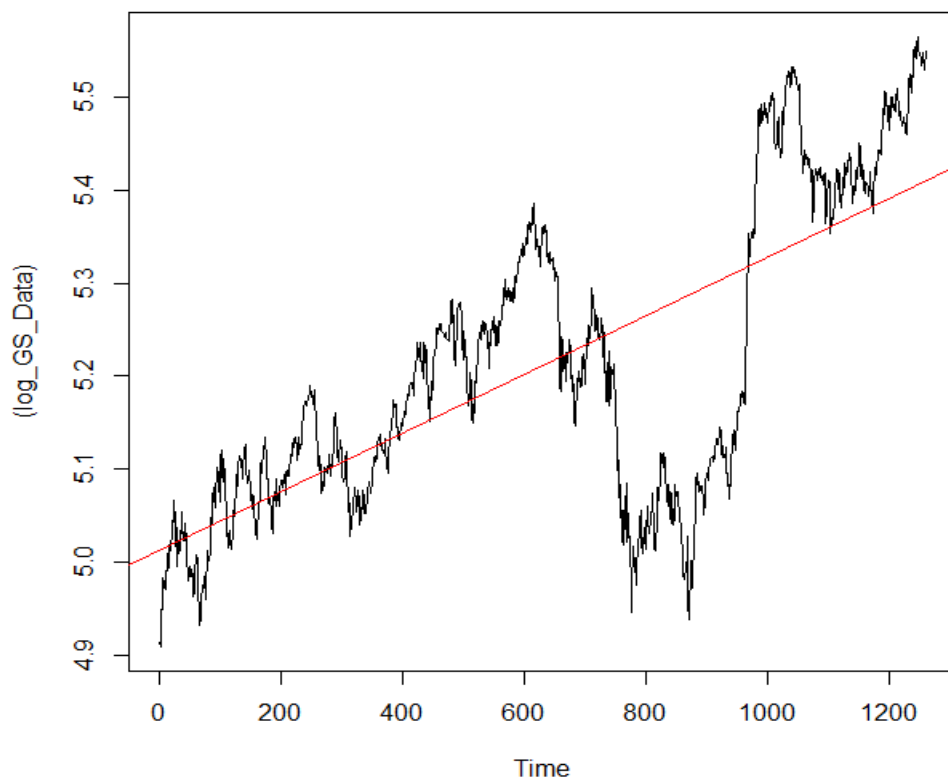
Προετοιμασία για την εφαρμογή του μοντέλου

Όπως και πριν, βλέπουμε το γράφημα της χρονοσειράς (Γράφημα 40), βλέπουμε την ίδια χρονοσειρά απαλλαγμένη από τυχαίες διακυμάνσεις (Γράφημα 41) και κάτω

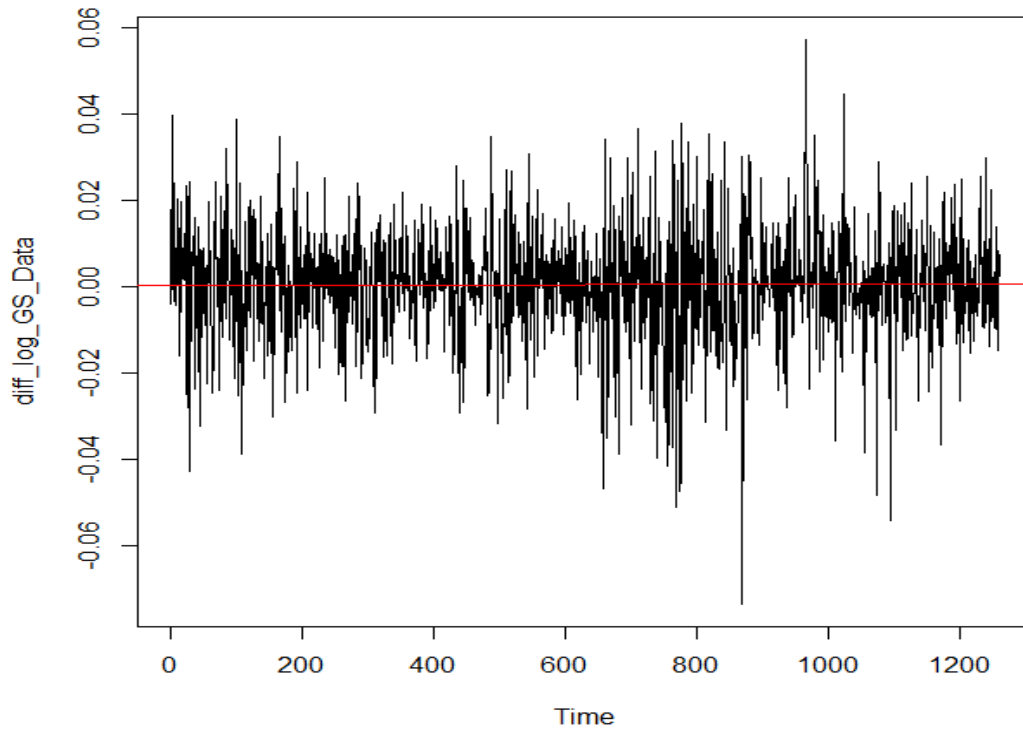
επίσης την ίδια χρονοσειρά απαλλαγμένη και από τυχαίες διακυμάνσεις και από τάση (Γράφημα 42).



Γράφημα 40 :Χρονοσειρά τιμής κλεισίματος με τάση της εταιρείας GS



Γράφημα 41: Χρονοσειρά με τάση της εταιρείας GS κατόπιν εφαρμογής της λογαριθμικής συνάρτησης για την απαλοιφή των τυχαίων διακυμάνσεων



Γράφημα 42: Χρονοσειρά χωρίς τάση της εταιρείας GS

Η μαύρη γραμμή δηλώνει τις τιμές της χρονοσειράς, ενώ η κόκκινη δηλώνει τη γραμμική εκτίμηση της τάσης που εμφανίζεται.

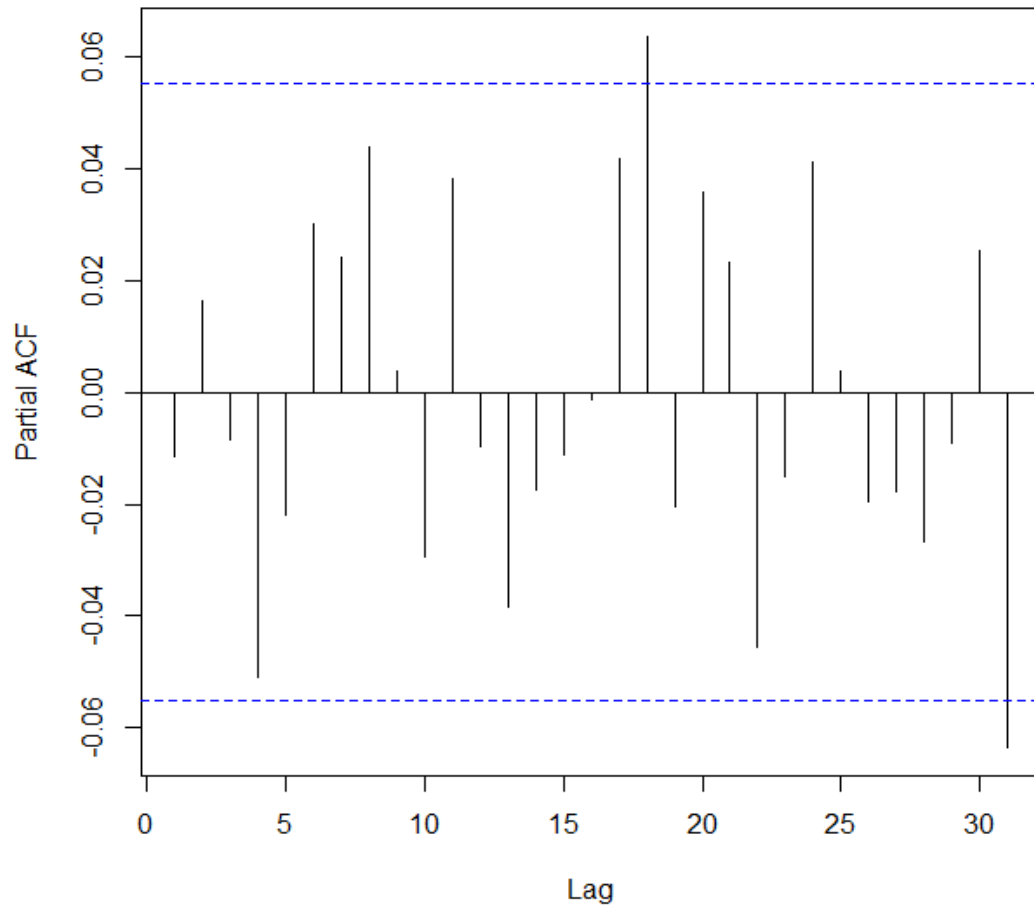
Παρατηρούμε πως η συγκεκριμένη χρονοσειρά δεν έχει κάποια εμφανή περιοδικότητα ή εποχικότητα, παρ' όλα αυτά έχει κάποια ιδιαιτερότητα λόγω της ύπαρξης διαστήματος στο οποίο υπήρξε μεγάλη πτώση και άνοδος, όπως και η προηγούμενη.

Το μοντέλο

Όπως και στην προηγούμενη περίπτωση, θα χρησιμοποιήσουμε ένα μοντέλο ARIMA(p,d,q).

Για την εύρεση της τάξης q, παίρνουμε το γράφημα των μερικών αυτοσυσχετίσεων, που δίνεται παρακάτω (Γράφημα 43):

Series diff_log_GS_Data



Γράφημα 43: Γράφημα αυτοσυσχετίσεων της χρονοσειράς για την εταιρεία GS

Η προτεινόμενη τάξη είναι $q=0$ καθώς η πρώτη αυτοσυσχέτιση είναι αρνητική.

Εφαρμογή του μοντέλου

Με βάση λοιπόν τα παραπάνω, το MA τμήμα έχει τάξη 1 και η τάση απαλείφθηκε με την εφαρμογή πρώτων διαφορών ($d=1$).

Έτσι, τα μοντέλα που προκύπτουν για αυτές τις παραμέτρους, σε συνδυασμό με υστερήσεις 3, 7 και 14 ως προς το αυτοπαλινδρομούμενο τμήμα, που απαιτείται από την εκφώνηση, παίρνουμε τα τρία παρακάτω μοντέλα:

ΠΙΝΑΚΑΣ 25: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία GS για την πρόβλεψη της επόμενης τιμής με υστέρηση 3

Model 1			
Κώδικας που χρησιμοποιήθηκε	fit3=arima(log_GS_Data,c(3,1,0))		
Συντελεστές			
	AR 1	AR 2	AR 3
	-0.0098	0.0178	-0.0072

ΠΙΝΑΚΑΣ 26: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία GS για την πρόβλεψη της επόμενης τιμής με υστέρηση 7

Model 2							
Κώδικας που χρησιμοποιήθηκε	fit7=arima(log_GS_Data,c(7,1,0))						
Συντελεστές							
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7
	-0.0113	0.0203	-0.0058	-0.0503	-0.0207	0.0318	0.0254

ΠΙΝΑΚΑΣ 27: Αποτελέσματα της εφαρμογής του μοντέλου ARIMA στην εταιρεία GS για την πρόβλεψη της επόμενης τιμής με υστέρηση 14

Model 3								
Κώδικας που χρησιμοποιήθηκε	fit14=arima(log_GS_Data,c(14,1,0))							
Συντελεστές								
	AR 1	AR 2	AR 3	AR 4	AR 5	AR 6	AR 7	AR 8
	-0.0120	0.0208	-0.0064	-0.0481	-0.0200	0.0325	0.0288	0.0456
	AR 9	AR 10	AR 11	AR 12	AR 13	AR 14		
	0.0022	-0.0287	0.0401	-0.0090	-0.0374	-0.0159		

Έλεγχος προσαρμογής και σύγκριση των μοντέλων

Τα μοντέλα και πάλι προσεγγίζουν σχετικά ικανοποιητικά τις πραγματικές τιμές. Για την επιλογή του καλύτερου εκ των τριών, υπολογίζουμε τα RMSE:

```
> RMSE3  
[1] 6.989435  
> RMSE7  
[1] 6.975744  
> RMSE14  
[1] 7.08563
```

Το καλύτερο μοντέλο φαίνεται πως είναι αυτό με την υστέρηση των 7 ημερών, αμέσως επόμενο είναι αυτό με υστέρηση 3 ημερών και το λιγότερο καλό είναι το τελευταίο, δηλαδή αυτό με υστέρηση 14^{ov} ημερών.

Σύνοψη

Μέσω της παραπάνω ανάλυσης, έγινε εμφανής ο τρόπος δόμησης και λειτουργίας αυτών των μοντέλων. Αν και είναι σαφής η αδυναμία μακροπρόθεσμης πρόβλεψης, όπως φάνηκε σε μερικές περιπτώσεις η προσέγγιση ήταν πολύ ικανοποιητική. Στον παρακάτω πίνακα (βλ. Πίνακα 28), φαίνεται η σύνοψη από όλα τα μοντέλα που κατασκευάστηκαν.

ΠΙΝΑΚΑΣ 28: Αποτελέσματα RMSE για υστέρηση 3,7 και 14 για κάθε εταιρεία με χρήση μεθόδου ARIMA.

ARIMA Models				
Εταιρεία	MA	Difference	AR	RMSE
Facebook	d=1	q=1	p=3	2,576
			p=7	2,570
			p=14	2,571
Google	d=1	q=1	p=3	32,898
			p=7	31,486
			p=14	33,156
MSFT	d=1	q=1	p=3	2,183
			p=7	2,193
			p=14	1,967
BA	d=1	q=1	p=3	22,581
			p=7	22,530
			p=14	22,566
LMT	d=1	q=1	p=3	8,083
			p=7	7,546
			p=14	7,621
MMM	d=1	q=1	p=3	5,022
			p=7	4,556
			p=14	5,050
BLK	d=0	q=1	p=3	26,471
			p=7	26,530
			p=14	26,359

JPM	d=0	q=1	p=3	0,085
			p=7	0,084
			p=14	0,089
GS	d=1	q=1	p=3	6,989
			p=7	6,975
			p=14	7,085

Όπως βλέπουμε, όλα τα μοντέλα έχουν μικρή τάξη στο τμήμα του κινητού μέσου (0 ή 1), ενώ οι διαφορές πρώτης τάξης είναι αρκετές σε όλες τις περιπτώσεις για την επιβολή της στασιμότητας. Με κόκκινο έχουν σημειωθεί τα πιο επιτυχημένα μοντέλα που ανέδειξε η μοντελοποίηση για κάθε περίπτωση, από τα οποία παρατηρούμε πως κανένα δεν είναι τάξης 3. Αυτό σημαίνει πως οι 3 ημέρες δεν είναι επαρκείς για την αποθήκευση πληροφοριών που χρειάζονται για την βέλτιστη εκμάθηση του μοντέλου.

4.Νευρωνικά δίκτυα

4.1 Ορισμός των νευρωνικών δικτύων

Ο απλούστερος ορισμός ενός νευρωνικού δικτύου, αναφερόμενος καλύτερα ως «τεχνητό» νευρωνικό δίκτυο (ANN), προήλθε από τον εφευρέτη ενός από τους πρώτους νευροϋπολογιστές, τον Δρ. Robert Hecht-Nielsen, ο οποίος όρισε το νευρωνικό δίκτυο ως εξής: «... ένα υπολογιστικό σύστημα που αποτελείται από έναν αριθμό απλών, ιδιαίτερα διασυνδεδεμένων στοιχείων επεξεργασίας, τα οποία επεξεργάζονται τις πληροφορίες με την δική τους δυναμική απόκριση σε εξωτερικές πληροφορίες (στο «Neural Network Primer: Part I» του Maureen Caudill, AI Expert, Φεβρουάριος 1989).

Ένα μεγάλο τεχνητό νευρωνικό δίκτυο μπορεί να έχει εκατοντάδες ή χιλιάδες μονάδες επεξεργασίας, ενώ ένας εγκέφαλος θηλαστικών έχει δισεκατομμύρια νευρώνες με αντίστοιχη αύξηση του μεγέθους της συνολικής αλληλεπίδρασης και της αναδυόμενης συμπεριφοράς. Αν και οι ερευνητές των τεχνητών νευρωνικών δικτύων γενικά δεν ασχολούνται με το αν τα δίκτυά τους μοιάζουν με τα βιολογικά συστήματα, κάποιοι όντως ασχολούνται με αυτό. Για παράδειγμα, οι ερευνητές έχουν προσομοιώσει με ακρίβεια την λειτουργία του αμφιβληστροειδούς και έχουν μοντελοποιήσει τις λειτουργίες του οφθαλμού σχετικά καλά [10].

Αν και τα μαθηματικά που εμπλέκονται στην νευρωνική δικτύωση δεν είναι μια ασήμαντη υπόθεση, ο χρήστης μπορεί μάλλον εύκολα να αποκτήσει τουλάχιστον μία λειτουργική κατανόηση της δομής και της λειτουργίας τους [1]

Τα νευρωνικά δίκτυα είναι συνήθως οργανωμένα σε στρώματα. Τα στρώματα αποτελούνται από έναν αριθμό διασυνδεδεμένων «κόμβων» που περιέχουν μια «λειτουργία ενεργοποίησης». Τα μοτίβα παρουσιάζονται στο δίκτυο μέσω του «στρώματος εισόδου», το οποίο επικοινωνεί με ένα ή περισσότερα «κρυμμένα στρώματα» όπου η πραγματική επεξεργασία γίνεται μέσω ενός συστήματος σταθμισμένων συνδέσεων. Τα κρυμμένα στρώματα συνδέονται στην συνέχεια με ένα στρώμα εξόδου όπου η απάντηση εξάγεται όπως φαίνεται στο παρακάτω γράφημα.

- Σχέδιο νευρωνικού δικτύου

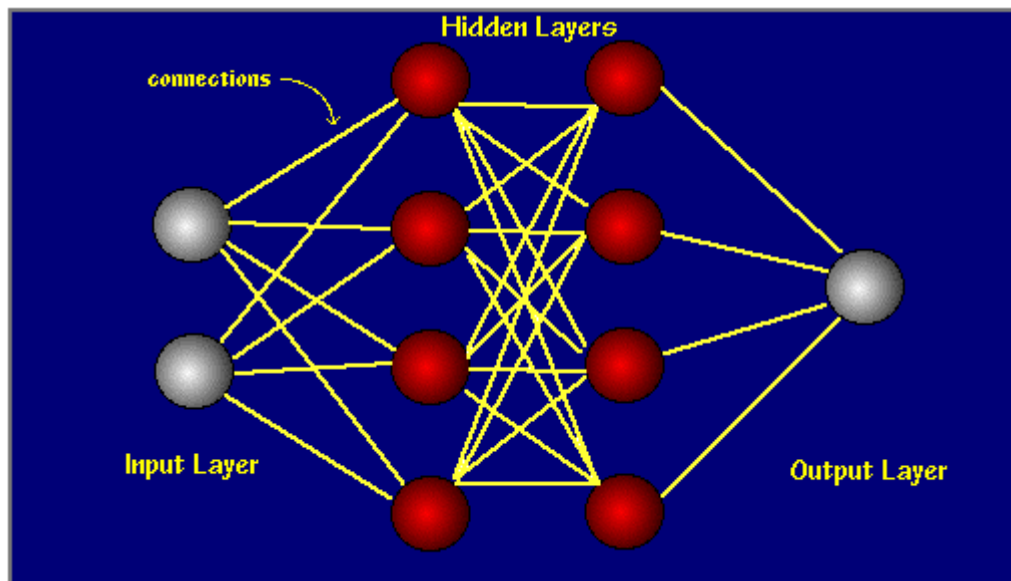


Figure 1: Σχέδιο τεχνητού νευρωνικού δικτύου

Τα περισσότερα τεχνητά νευρωνικά δίκτυα περιέχουν κάποια μορφή «κανόνα μάθησης» που τροποποιεί τα βάρη των συνδέσεων σύμφωνα με τα πρότυπα εισόδου που παρουσιάζονται [21]. Κατά μία έννοια, τα τεχνητά νευρωνικά δίκτυα μαθαίνουν με το παράδειγμα και τους βιολογικούς ομολόγους τους, όπως ένα παιδί μαθαίνει να αναγνωρίζει τα σκυλιά από παραδείγματα σκυλιών [6]

Αν και υπάρχουν πολλά διαφορετικά είδη μαθησιακών κανόνων που χρησιμοποιούνται από τα νευρωνικά δίκτυα, αυτή η επίδειξη αφορά μόνο έναν, τον κανόνα δέλτα. Ο κανόνας δέλτα συχνά χρησιμοποιείται από την πιο κοινή κλάση των τεχνητών νευρωνικών δικτύων που ονομάζεται «νευρωνικά δίκτυα οπισθοδιάδοσης». Η οπισθοδιάδοση είναι μια συντομογραφία για την οπισθοδιάδοση σφάλματος.

Με τον κανόνα δέλτα, όπως συμβαίνει και με άλλους τύπους οπισθοδιάδοσης, η «μάθηση» είναι μια εποπτευόμενη διαδικασία που συμβαίνει με κάθε κύκλο ή «εποχή» (δηλ. κάθε φορά που παρουσιάζεται στο δίκτυο ένα νέο μοτίβο στοιχείων εισόδου) μέσω μίας προς τα εμπρός ροής ενεργοποίησης των δεδομένων εξόδου και της οπισθοδιάδοσης σφάλματος των προσαρμογών βάρους. Πιο απλά, όταν ένα νευρωνικό δίκτυο παρουσιάζεται αρχικά με ένα μοτίβο κάνει μια τυχαία «εικασία» ως προς το τι θα μπορούσε να είναι. Στην συνέχεια, βλέπει πόσο απείχε η απάντησή

του από την πραγματική και κάνει την κατάλληλη προσαρμογή στα βάρη σύνδεσης. Πιο γραφικά, η διαδικασία φαίνεται κάπως έτσι [7]:

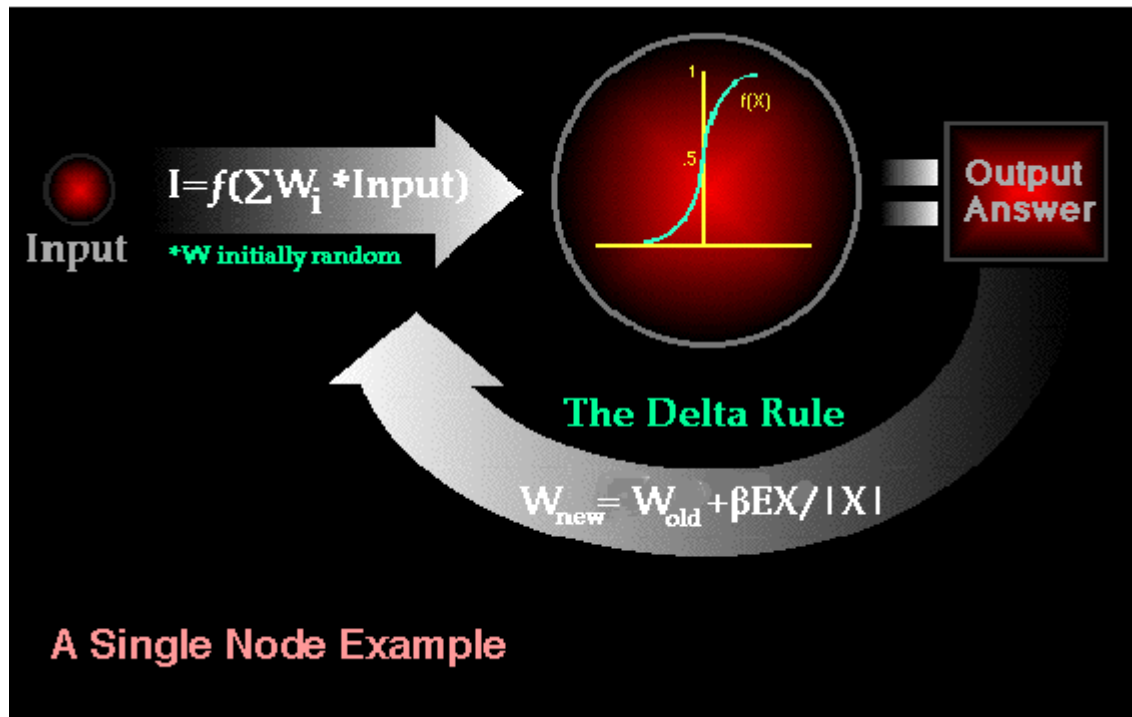


Figure 2: 'Μάθηση' νευρωνικού δικτύου με Κανόνα Δέλτα

Ας σημειωθεί επίσης ότι σε κάθε κόμβο του κρυμμένου στρώματος, υπάρχει μιας σιγμοειδής λειτουργία ενεργοποίησης που πολώνει την δραστηριότητα του δικτύου και βοηθά στην σταθεροποίηση.

Η οπισθοδιάδοση εκτελεί μια κλίση καθόδου εντός του διανυσματικού χώρου της λύσης προς ένα «παγκόσμιο ελάχιστο» κατά μήκος του πιο απότομου διανύσματος της επιφάνειας του σφάλματος. Η ίδια η επιφάνεια σφάλματος είναι ένα υπερ-παραβολοειδές, αλλά σπανίως είναι «ομαλή» όπως απεικονίζεται στο παρακάτω γράφημα. Πράγματι, στα περισσότερα προβλήματα, ο χώρος της λύσης είναι αρκετά ακανόνιστος με πολυάριθμες «κοιλότητες» και «κορυφές» που μπορεί να προκαλέσουν το δίκτυο να συμβιβαστεί με ένα «τοπικό ελάχιστο» που δεν είναι η καλύτερη συνολική λύση [4]

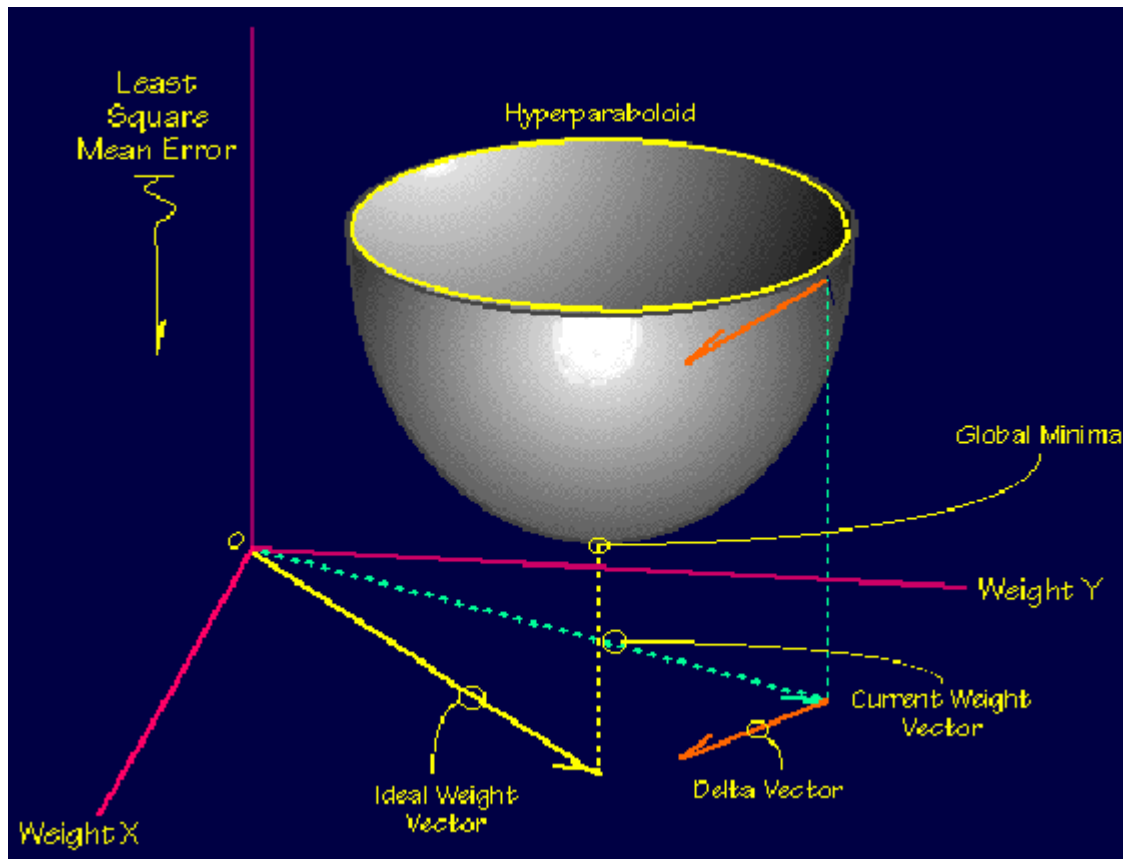


Figure 3 Η επιφάνεια σφάλματος της οπισθοδιάδοσης

Δεδομένου ότι η φύση του χώρου του σφάλματος δεν μπορεί να είναι γνωστή εκ των προτέρων, η ανάλυση νευρωνικού δικτύου απαιτεί συχνά έναν μεγάλο αριθμό μεμονωμένων διαδικασιών για να καθοριστεί η καλύτερη λύση. Οι περισσότεροι κανόνες μάθησης έχουν ενσωματωμένους μαθηματικούς όρους για να βοηθήσουν σε αυτήν την διαδικασία που ελέγχει την «ταχύτητα» (συντελεστής βήτα) και την «ορμή» της μάθησης. Η ταχύτητα της μάθησης είναι στην πραγματικότητα ο ρυθμός σύγκλισης μεταξύ της τρέχουσας λύσης και του παγκόσμιου ελάχιστου. Η ορμή βοηθά το δίκτυο να ξεπεράσει τα εμπόδια (τοπικά ελάχιστα) στην επιφάνεια σφάλματος και να συμβιβαστεί στο ή κοντά στο παγκόσμιο ελάχιστο [3]

Μόλις ένα νευρωνικό δίκτυο «εκπαιδευτεί» σε ικανοποιητικό επίπεδο, μπορεί να χρησιμοποιηθεί ως αναλυτικό εργαλείο σε άλλα δεδομένα. Για να γίνει αυτό, ο χρήστης δεν καθορίζει πλέον την εκπαίδευση και αντί αυτού επιτρέπει στο δίκτυο να λειτουργεί μόνο σε λειτουργία οπισθοδιάδοσης. Νέα δεδομένα εισόδου παρουσιάζονται στο μοτίβο τα οποία φιλτράρονται και επεξεργάζονται από τα μεσαία

στρώματα σαν να πραγματοποιούνταν εκπαίδευση, ωστόσο, σε αυτό το σημείο τα δεδομένα εξόδου διατηρούνται και δεν υπάρχει οπισθοδιάδοση. Τα δεδομένα εξόδου μιας οπισθοδιάδοσης είναι το προβλεπόμενο μοντέλο για τα δεδομένα τα οποία στην συνέχεια μπορούν να χρησιμοποιηθούν για περαιτέρω ανάλυση και ερμηνεία.

Είναι επίσης δυνατή η υπερ-εκπαίδευση ενός νευρωνικού δικτύου, το οποίο σημαίνει ότι το δίκτυο έχει εκπαιδευτεί ακριβώς για να ανταποκριθεί μόνο σε έναν τύπο δεδομένων εισόδου, κάτι που μοιάζει με την μηχανική αποστήθιση. Εάν συμβεί κάτι τέτοιο, τότε δεν μπορεί πλέον να υπάρξει εκμάθηση. Σε εφαρμογές πραγματικού κόσμου, αυτή η κατάσταση δεν είναι πολύ χρήσιμη, δεδομένου ότι θα χρειαζόταν ένα ξεχωριστό δίκτυο για κάθε νέο είδος δεδομένων εισόδου [5]

Για την καλύτερη κατανόηση των τεχνητών νευρωνικών υπολογιστών είναι σημαντικό να γνωρίζουμε πρώτα πώς ένας συμβατικός «σειριακός» υπολογιστής και το λογισμικό του επεξεργάζεται τις πληροφορίες. Ένας σειριακός υπολογιστής διαθέτει έναν κεντρικό επεξεργαστή ο οποίος μπορεί να απευθύνεται σε μια σειρά από θέσεις μνήμης όπου αποθηκεύονται δεδομένα και οδηγίες. Οι υπολογισμοί γίνονται από τον επεξεργαστή που διαβάζει μια εντολή καθώς και οποιαδήποτε δεδομένα απαιτούνται σύμφωνα με τις οδηγίες από τις διευθύνσεις μνήμης, εκτελείται η εντολή και τα αποτελέσματα αποθηκεύονται σε μια καθορισμένη θέση μνήμης όπως απαιτείται. Σε ένα σειριακό σύστημα (και ένα παράλληλο πρότυπο), τα υπολογιστικά βήματα είναι ντετερμινιστικά, διαδοχικά και λογικά και η κατάσταση μιας δεδομένης μεταβλητής μπορεί να παρακολουθηθεί από την μία λειτουργία στην άλλη.[8]

Συγκριτικά, τα τεχνητά νευρωνικά δίκτυα δεν είναι διαδοχικά ή αναγκαστικά ντετερμινιστικά. Δεν υπάρχουν περίπλοκοι κεντρικοί επεξεργαστές, μάλλον υπάρχουν πολλοί απλοί που γενικά δεν κάνουν τίποτα περισσότερο από να πάρουν το σταθμισμένο άθροισμα των δεδομένων εισαγωγής τους από άλλους επεξεργαστές. Τα τεχνητά νευρωνικά δίκτυα δεν εκτελούν προγραμματισμένες οδηγίες, αποκρίνονται παράλληλα (είτε προσομοιωμένα είτε πραγματικά) στο μοτίβο των δεδομένων εισαγωγής που παρουσιάζονται σε αυτό. Δεν υπάρχουν επίσης ξεχωριστές διευθύνσεις μνήμης για την αποθήκευση δεδομένων. Αντ' αυτού, οι πληροφορίες περιέχονται στη «κατάσταση» γενικής ενεργοποίησης του δικτύου. Ως εκ τούτου, η

«γνώση» αντιπροσωπεύεται από το ίδιο το δίκτυο, το οποίο είναι κυριολεκτικά κάτι περισσότερο από το άθροισμα των επιμέρους συστατικών του.

Τα νευρωνικά δίκτυα χρησιμεύουν αρκετά για να εντοπίζουν συσχετίσεις ή να ανακαλύπτουν κανονικότητες μέσα σε ένα σύνολο προτύπων, όπου ο όγκος, ο αριθμός μεταβλητών ή η ποικιλία των δεδομένων είναι πολύ μεγάλος, όπου οι σχέσεις μεταξύ των μεταβλητών είναι λίγο κατανοητές ή οι σχέσεις είναι δύσκολο να περιγραφούν επαρκώς με τις συμβατικές προσεγγίσεις.

Υπάρχουν πολλά πλεονεκτήματα και περιορισμοί στην ανάλυση νευρωνικών δικτύων και για να συζητηθεί σωστά αυτό το θέμα θα πρέπει να εξεταστεί κάθε τύπος δικτύου. Όσον αφορά στα δίκτυα στήριξης, ωστόσο, υπάρχουν ορισμένα ειδικά ζητήματα που πρέπει να γνωρίζουν οι δυνητικοί χρήστες.

Τα νευρωνικά δίκτυα οπισθοδιάδοσης (και πολλά άλλα δίκτυα) είναι κατά κάποιον τρόπο τα τελικά «μαύρα κουτιά». Εκτός από τον ορισμό της γενικής αρχιτεκτονικής ενός δικτύου και ίσως αρχικά την τροφοδοσία του με τυχαίους αριθμούς, ο χρήστης δεν έχει άλλο ρόλο παρά να το τροφοδοτήσει με δεδομένα εισόδου και να παρακολουθήσει την εκπαίδευση και να περιμένει τα δεδομένα εξόδου. Στην πραγματικότητα, έχει ειπωθεί ότι με την οπισθοδιάδοση ο χρήστης σχεδόν δεν ξέρει τι κάνει. Ορισμένα πακέτα λογισμικού που διατίθενται ελεύθερα (NevProp, bp, Mactivation) επιτρέπουν στον χρήστη να βλέπει την πρόοδο των δικτύων σε τακτά χρονικά διαστήματα, αλλά η ίδια η εκμάθηση προχωρά από μόνη της. Το τελικό προϊόν αυτής της δραστηριότητας είναι ένα εκπαιδευμένο δίκτυο που δεν παρέχει εξισώσεις ή συντελεστές που ορίζουν μια σχέση (όπως στην παλινδρόμηση) πέρα από τα δικά του εσωτερικά μαθηματικά. Το δίκτυο είναι η τελική εξίσωση της σχέσης [11]

Τα δίκτυα οπισθοδιάδοσης τείνουν να είναι πιο αργά στην εκπαίδευση από άλλους τύπους δικτύων και μερικές φορές απαιτούν αρκετά μεγάλο χρόνο. Εάν τρέχει σε ένα πραγματικά παράλληλο σύστημα ηλεκτρονικού υπολογιστή αυτό το ζήτημα δεν αποτελεί πραγματικό πρόβλημα, αλλά εάν το BPNN προσομοιώνεται σε ένα τυπικό σειριακό μηχάνημα (δηλαδή σε ένα SPARC, Mac ή PC), η εκπαίδευση μπορεί να διαρκέσει κάμποσο. Αυτό οφείλεται στο γεγονός ότι οι μηχανές CPU πρέπει να υπολογίζουν χωριστά την λειτουργία κάθε κόμβου και σύνδεσης, γεγονός που μπορεί να είναι προβληματικό σε πολύ μεγάλα δίκτυα με μεγάλο όγκο δεδομένων. Ωστόσο,

η ταχύτητα των περισσότερων σύγχρονων μηχανών είναι τέτοια που τυπικά δεν αποτελεί και μεγάλο πρόβλημα.

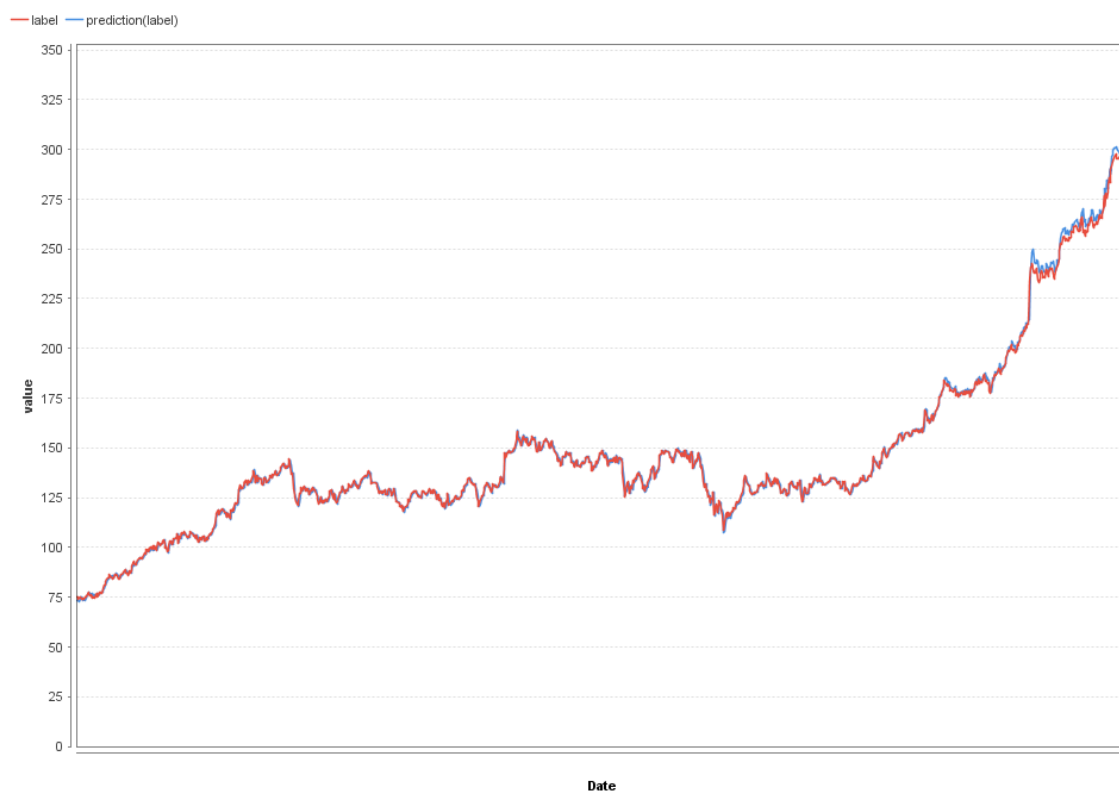
Ανάλογα με την φύση της εφαρμογής και την δύναμη των εσωτερικών προτύπων δεδομένων, αναμένεται ότι το δίκτυο θα εκπαιδεύεται αρκετά καλά. Αυτό ισχύει για προβλήματα όπου οι σχέσεις μπορεί να είναι αρκετά δυναμικές ή μη γραμμικές. Τα τεχνητά νευρωνικά δίκτυα παρέχουν μια αναλυτική εναλλακτική λύση έναντι των συμβατικών τεχνικών οι οποίες περιορίζονται συχνά από αυστηρές υποθέσεις κανονικότητας, γραμμικότητας, ανεξαρτησίας της μεταβλητής κλπ. Επειδή ένα τεχνητό νευρωνικό δίκτυο μπορεί να συλλάβει πολλά είδη σχέσεων επιτρέπει στο χρήστη να μοντελοποιεί φαινόμενα γρήγορα και σχετικά εύκολα κάτι που διαφορετικά μπορεί να ήταν πολύ δύσκολο ή απίθανο να εξηγηθεί διαφορετικά [2]

Τα τελευταία χρόνια τα τεχνητά Νευρωνικά δίκτυα καθιερώνονται όλο και περισσότερο στο χώρο της έρευνας. Η ικανότητά τους να προσεγγίζουν μη γραμμικές σχέσεις χωρίς να κάνουν εκ των προτέρων υποθέσεις αναφορικά με τη φύση της διαδικασίας που δημιουργεί το υπό εξέταση φαινόμενο είναι και το μεγάλο πλεονέκτημά τους .

5. Μοντελοποίηση με χρήση νευρωνικών δικτύων

5.1 Πρόβλεψη της τιμής κλεισίματος της εταιρείας BA

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο χρώμα και τη προβλεπόμενη τιμή με μπλε χρώμα (λαμβάνοντας υπόψη τις 3 προηγούμενες μέρες) (βλ. Γράφημα 44)



Γράφημα 44 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία BA

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό.



Το σφάλμα πρόβλεψης για τις 3 μέρες είναι :

root_mean_squared_error

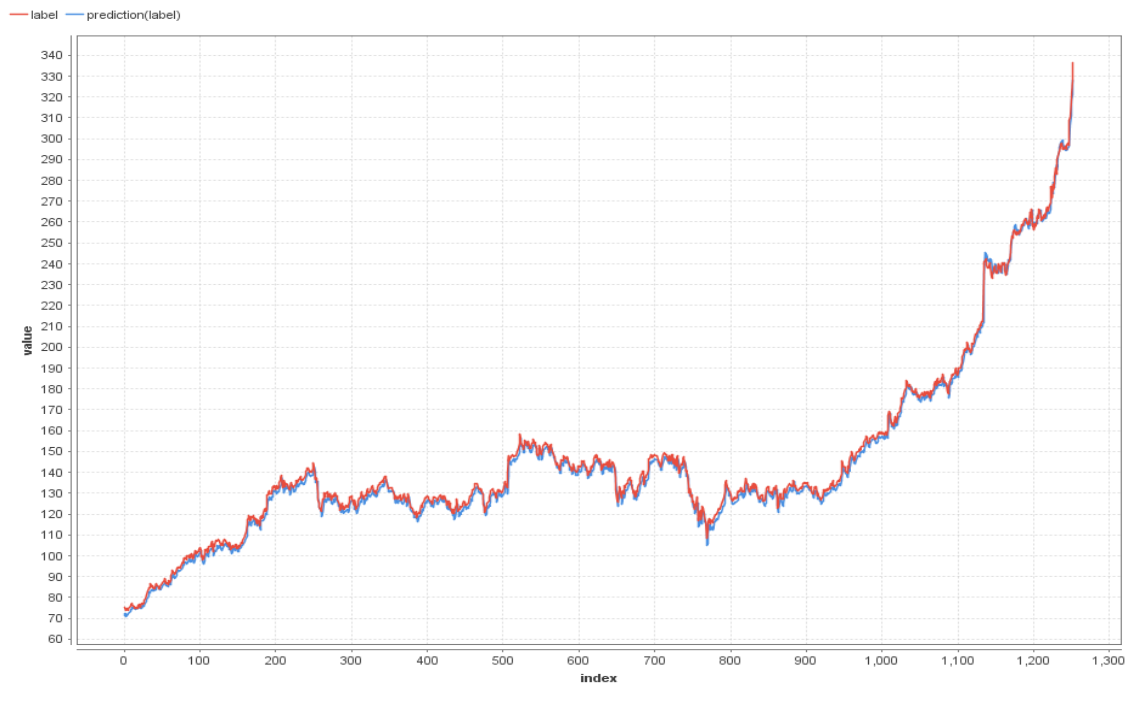
root_mean_squared_error: 3.302 +/- 3.705 (mikro: 4.967 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω είναι 146.413. Με βάση αυτό, το σφάλμα μας (3.302) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 29). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 3.302/146.413=0.02255$

ΠΙΝΑΚΑΣ 29: Πραγματική και προβλεπόμενη τιμή για την εταιρεία ΒΑ με χρήση νευρωνικών δικτύων με υστέρηση 3

Name	Type	Missing	Statistics		Filter (18 / 18 attributes):
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-16 (1)	Search for Attributes
Label label	Numeric	0	Min 73.650	Max 336.210	Average 146.413
Prediction prediction(label)	Numeric	0	Min 72.866	Max 332.454	Average 146.624

- Το μοντέλο μας λαμβάνοντας υπόψη της 7 προηγούμενες μέρες φαίνεται παρακάτω (βλ. Γράφημα 45)



Γράφημα 45 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία BA

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό.



Και εδώ φαίνεται μια πολύ προσέγγιση όπου το σφάλμα πρόβλεψης μας για 7 ημέρες είναι

root_mean_squared_error

```
root_mean_squared_error: 3.439 +/- 3.260 (mikro: 4.735 +/- 0.000)
```

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω είναι 146.642. Με βάση αυτό, το σφάλμα μας (3.439) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 30). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 3.439/146.642=0.02345$

ΠΙΝΑΚΑΣ 30: Πραγματική και προβλεπόμενη τιμή για την εταιρεία ΒΑ με χρήση νευρωνικών δικτύων με υστέρηση 7

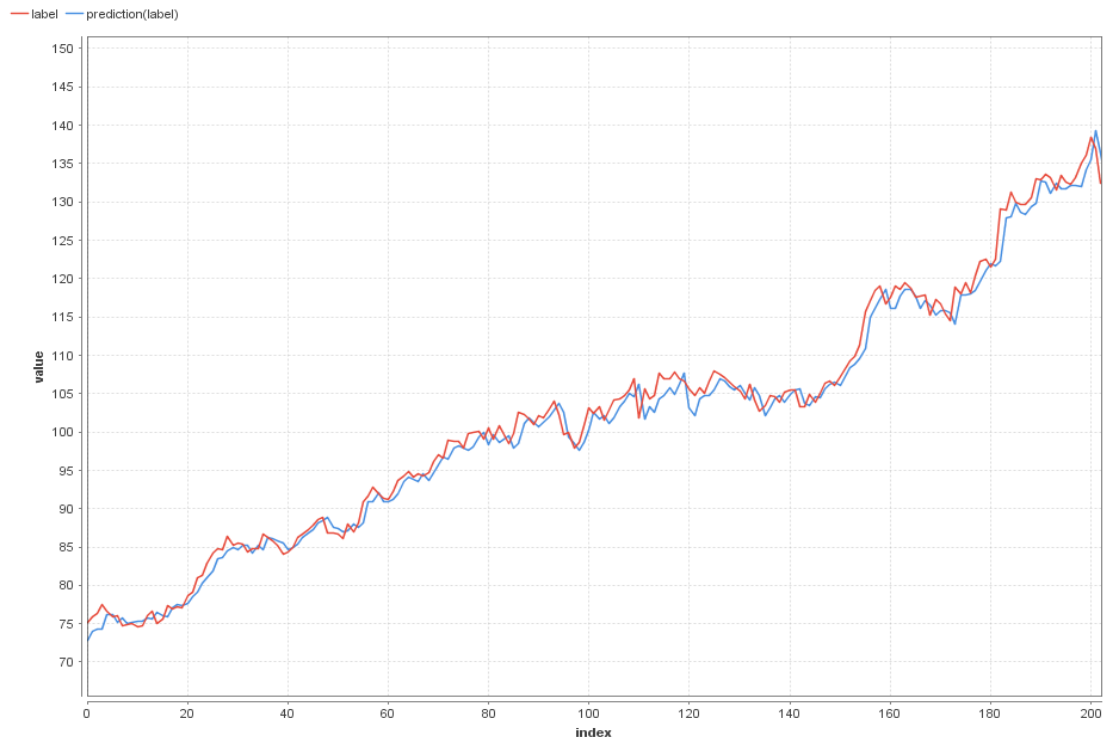
Name	Type	Missing	Statistics	Filter (38 / 38 attributes):
✓ Date	Polynomial	0	Least: 2018-01-12 (0) Most: 2013-01-23 (1)	Search for Attributes
✓ label	Numeric	0	Min: 73.650 Max: 336.210	Average: 146.642
✓ prediction(label)	Numeric	0	Min: 70.963 Max: 327.858	Average: 144.863

- Το μοντέλο μας λαμβάνοντας υπόψη της 14 προηγούμενες μέρες φαίνεται παρακάτω (βλ. Γράφημα 46)



Γράφημα 46 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία ΒΑ

Εστιάζοντας τη γραφική παράσταση θα δούμε ότι έχουμε μια πολύ καλή προσέγγιση



Το σφάλμα πρόβλεψης για τις 14 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 3.900 +/- 3.437 (mikro: 5.203 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω είναι 147.047. Με βάση αυτό, το σφάλμα μας (3.900) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 31). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 3.900/147.047=0.02652$

ΠΙΝΑΚΑΣ 31: Πραγματική και προβλεπόμενη τιμή για την εταιρεία BA με χρήση νευρωνικών δικτύων με υστέρηση 14

Name	Type	Missing	Statistics		
Id Date	Date time	1246	Earliest date Aug 17, 292278994 9:1...	Latest date Dec 2, 292269055 6:47 ...	Duration 0d 0h 0m 0s
Label label	Numeric	0	Min 74.650	Max 336.210	Average 147.047
Prediction prediction(label)	Numeric	0	Min 72.853	Max 330.964	Average 146.638

5.2 Πρόβλεψη της τιμής κλεισίματος της εταιρείας BLK

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 3 προηγούμενες μέρες) με μπλε χρώμα (βλ Γράφημα 47)



Γράφημα 47 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία BLK

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 3 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 4.957 +/- 1.180 (mikro: 5.096 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω είναι 344.322 ,ποσό αρκετά μεγάλο σε σχέση με πριν. Ωστόσο, το σφάλμα μας (4.957) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 32). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 4.957/344.322=0.01439$

ΠΙΝΑΚΑΣ 32: Πραγματική και προβλεπόμενη τιμή για την εταιρεία BLK με χρήση νευρωνικών δικτύων με υστέρηση 3

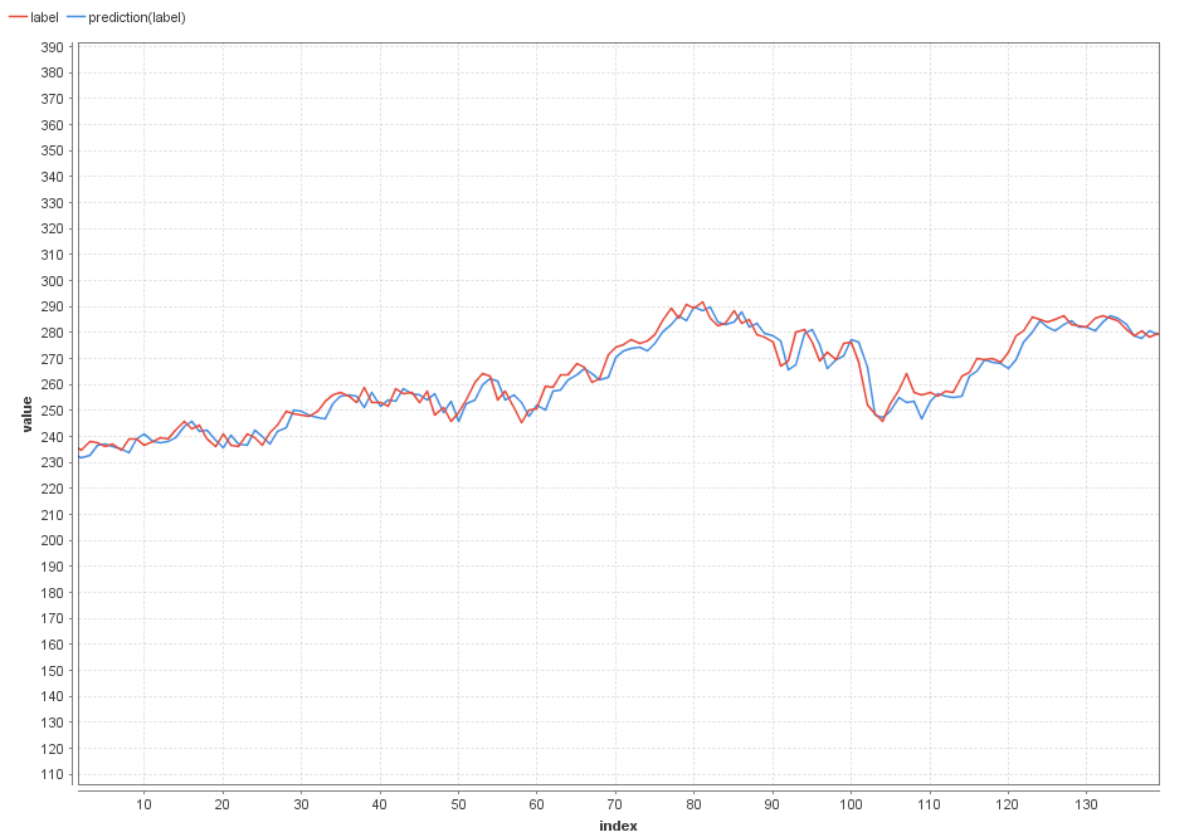
Name	Type	Missing	Statistics		Filter (18 / 18 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-16 (1)	Values 2013-01-16 (1), 2013-01-17 (1), ...[1258 more]
Label label	Numeric	0	Min 232	Max 555.530	Average 344.322
Prediction prediction(label)	Numeric	0	Min 226.149	Max 543.369	Average 345.554

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 7 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 48)



Γράφημα 48 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία BLK

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 7 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 6.204 +/- 2.794 (mikro: 6.800 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω είναι 344.672. Με βάση αυτό, το σφάλμα μας (6.204) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 33). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 6.204/344.672=0.01799$

ΠΙΝΑΚΑΣ 33: Πραγματική και προβλεπόμενη τιμή για την εταιρεία BLK με χρήση νευρωνικών δικτύων με υστέρηση 7

Name	Type	Missing	Statistics		Filter (38 / 38 attributes): <input type="text" value="Search for Attributes"/>
✓ Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-23 (1)	Values 2013-01-23 (1), 2013-01-24 (1), ... [1258 more]
✓ label	Numeric	0	Min 234.480	Max 555.530	Average 344.672
✓ prediction(label)	Numeric	0	Min 231.848	Max 537.789	Average 343.373

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 14 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 49)



Γράφημα 49 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία BLK

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 14 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 5.740 +/- 2.353 (mikro: 6.207 +/- 0.000)

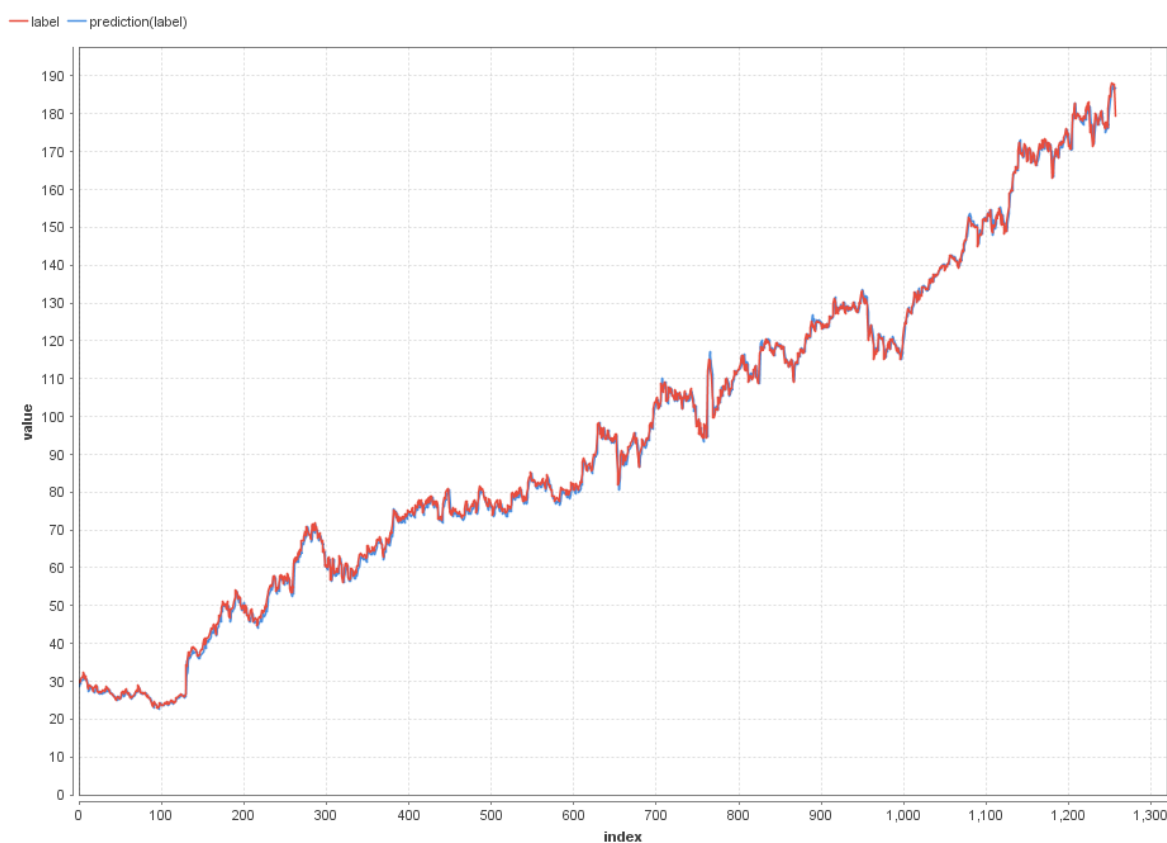
Ο μέσος όρος της πραγματικής τιμής είναι 345.279 , όπως φαίνεται και παρακάτω. Με βάση αυτό, το σφάλμα μας (5.740) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 34). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 5.740/345.279=0.01662$

ΠΙΝΑΚΑΣ 34: Πραγματική και προβλεπόμενη τιμή για την εταιρεία BLK με χρήση νευρωνικών δικτύων με υστέρηση 14

Name	Type	Missing	Statistics		Filter (73 / 73 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-02-01 (1)	Values 2013-02-01 (1), 2013-02-04 (1), ... [1258 more]
Label label	Numeric	0	Min 234.480	Max 555.530	Average 345.279
Prediction prediction(label)	Numeric	0	Min 235.055	Max 545.039	Average 345.425

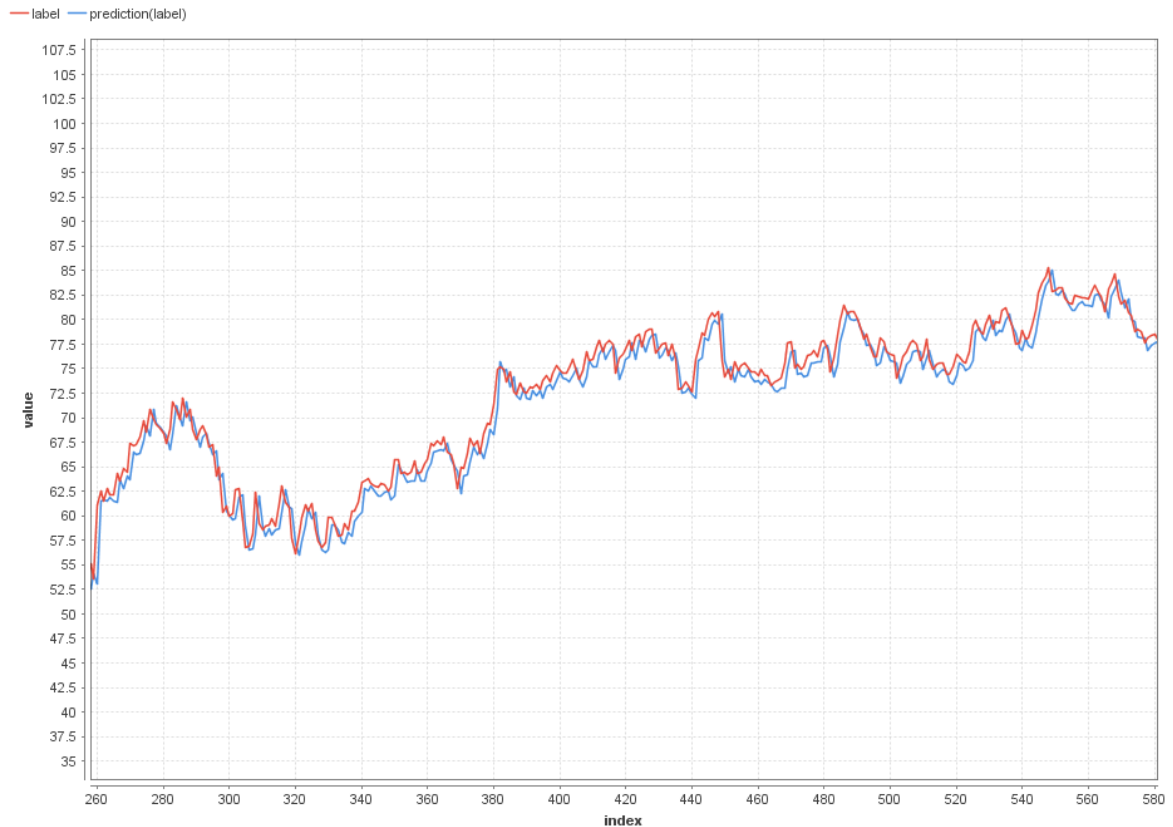
5.3 Πρόβλεψη της τιμής κλεισίματος της εταιρείας FB

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 3 προηγούμενες μέρες) με μπλε χρώμα (Γράφημα 50)



Γράφημα 50 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία Facebook

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 3 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 1.966 +/- 0.882 (mikro: 2.156 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής είναι 94.491, όπως φαίνεται και παρακάτω. Με βάση αυτό, το σφάλμα μας (1.966) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 35). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 1.966 / 94.491 = 0.02080$

ΠΙΝΑΚΑΣ 35: Πραγματική και προβλεπόμενη τιμή για την εταιρεία Facebook με χρήση νευρωνικών δικτύων με υστέρηση 3

Name	Type	Missing	Statistics		Filter (18 / 18 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-16 (1)	Values 2013-01-16 (1), 2013-01-17 (1), ... [1258 more]
Label label	Real	0	Min 22.900	Max 188.280	Average 94.491
Prediction prediction(label)	Real	0	Min 22.688	Max 187.353	Average 94.083

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 7 προηγούμενες μέρες) με μπλε χρώμα (Γράφημα 51)



Γράφημα 51 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία Facebook

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 7 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 2.091 +/- 0.820 (mikro: 2.245 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής είναι 94.695, όπως φαίνεται και παρακάτω. Με βάση αυτό, το σφάλμα μας (2.091) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 36). Το ποσοστό του σφάλματος είναι $\text{error/label} = 2.091 / 94.695 = 0.02281$

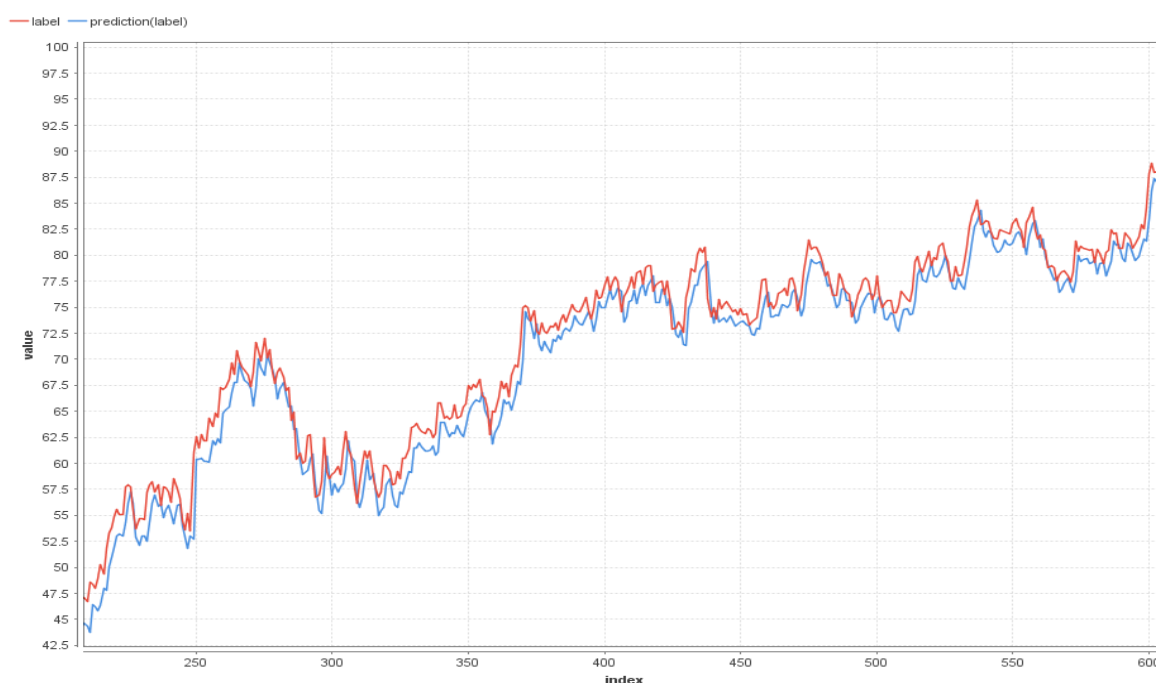
ΠΙΝΑΚΑΣ 36: Πραγματική και προβλεπόμενη τιμή για την εταιρεία Facebook με χρήση νευρωνικών δικτύων με υστέρηση 7

Name	Type	Missing	Statistics		Filter (38 / 38 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-23 (1)	Values 2013-01-23 (1), 2013-01-24 (1), ...[1258 more]
Label	Real	0	Min 22.900	Max 188.280	Average 94.695
Prediction prediction(label)	Real	0	Min 24.225	Max 185.960	Average 94.791

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 14 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 52)



Γράφημα 52 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία Facebook



Το αντίστοιχο σφάλμα πρόβλεψης φαίνεται παρακάτω

root_mean_squared_error

root_mean_squared_error: 2.558 +/- 1.216 (mikro: 2.833 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω είναι 95.053. Με βάση αυτό, το σφάλμα μας (2.558) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 37). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 2.588/95.053=0.02691$

ΠΙΝΑΚΑΣ 37: Πραγματική και προβλεπόμενη τιμή για την εταιρεία Facebook με χρήση νευρωνικών δικτύων με υστέρηση 14

Name	Type	Missing	Statistics	Filter (73 / 73 attributes)
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-02-01 (1)
Label label	Real	0	Min 22.900	Max 188.280
Prediction prediction(label)	Real	0	Min 21.816	Max 185.727

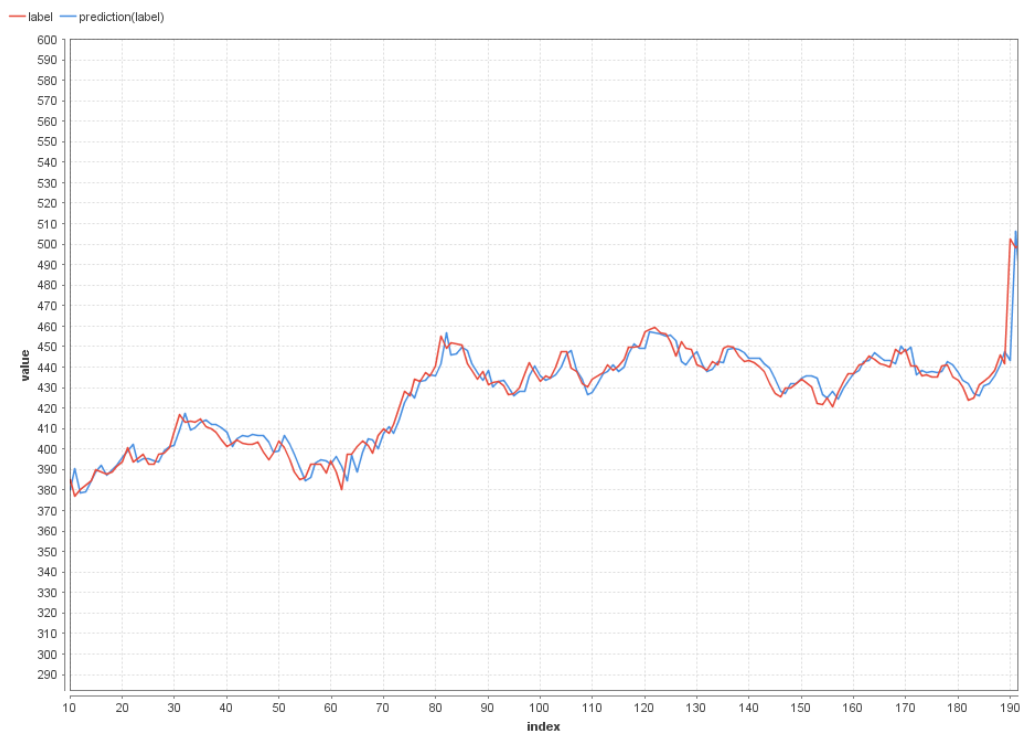
5.4 Πρόβλεψη της τιμής κλεισίματος της εταιρείας GOOGLE

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 3 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 53)



Γράφημα 53 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία Google

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 3 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 9.938 +/- 2.644 (mikro: 10.284 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής είναι 658.379, όπως φαίνεται και παρακάτω, ποσό αρκετά μεγάλο σε σχέση με πριν. Με βάση αυτό, το σφάλμα μας (9.938) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 38). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 9.938/658.379=0.01509$

ΠΙΝΑΚΑΣ 38: Πραγματική και προβλεπόμενη τιμή για την εταιρεία Google με χρήση νευρωνικών δικτύων με υστέρηση 3

Name	Type	Missing	Statistics	Filter (18 / 18 attributes):
Id				
▼ Date	Polynomial	0	Least: 2018-01-12 (0) Most: 2013-01-16 (1)	Values: 2013-01-16 (1), 2013-01-17 (1), ... [1258 more]
▼ label	Real	0	Min: 349.164 Max: 1122.260	Average: 658.379
▼ prediction(label)	Real	0	Min: 352.423 Max: 1107.123	Average: 658.569

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 7 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 54)



Γράφημα 54 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία Google

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 7 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 9.936 +/- 3.123 (mikro: 10.415 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω είναι 659.347. Με βάση αυτό, το σφάλμα μας (9.936) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 39). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 9.936/659.347=0.01396$

ΠΙΝΑΚΑΣ 39: Πραγματική και προβλεπόμενη τιμή για την εταιρεία Google με χρήση νευρωνικών δικτύων με υστέρηση 7

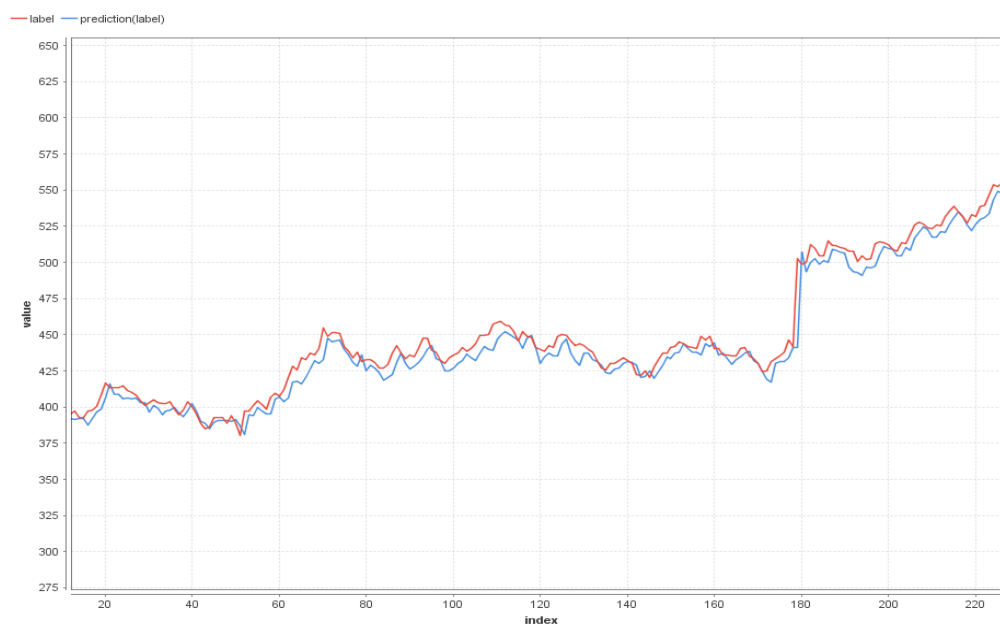
Name	Type	Missing	Statistics		Filter (38 / 38 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-23 (1)	Values 2013-01-23 (1), 2013-01-24 (1), ...[1258 more]
Label label	Real	0	Min 372.939	Max 1122.260	Average 659.347
Prediction prediction(label)	Real	0	Min 377.534	Max 1097.853	Average 662.860

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 14 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 55)



Γράφημα 55 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία Google

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 14 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 10.360 +/- 2.766 (mikro: 10.722 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω ,είναι 660.939. Με βάση αυτό, το σφάλμα μας (10.360) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 40). Το ποσοστό του σφάλματος είναι $\text{error/label} = 10.360/660.939=0.01567$

ΠΙΝΑΚΑΣ 40: Πραγματική και προβλεπόμενη τιμή για την εταιρεία Google με χρήση νευρωνικών δικτύων με υστέρηση 14

Name	Type	Missing	Statistics		Filter (73 / 73 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-02-01 (1)	Values 2013-02-01 (1), 2013-02-04 (1), ...[1258 more]
Label label	Real	0	Min 377.058	Max 1122.260	Average 660.939
Prediction prediction(label)	Real	0	Min 376.422	Max 1108.398	Average 659.005

5.5 Πρόβλεψη της τιμής κλεισίματος της εταιρείας GS

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 3 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 56)



Γράφημα 56 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία GS

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 3 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 2.616 +/- 0.455 (mikro: 2.655 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω ,είναι 186.032. Με βάση αυτό, το σφάλμα μας (2.616) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 41). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 2.616/186.032=0.01406$

ΠΙΝΑΚΑΣ 41: Πραγματική και προβλεπόμενη τιμή για την εταιρεία GS με χρήση νευρωνικών δικτύων με υστέρηση 3

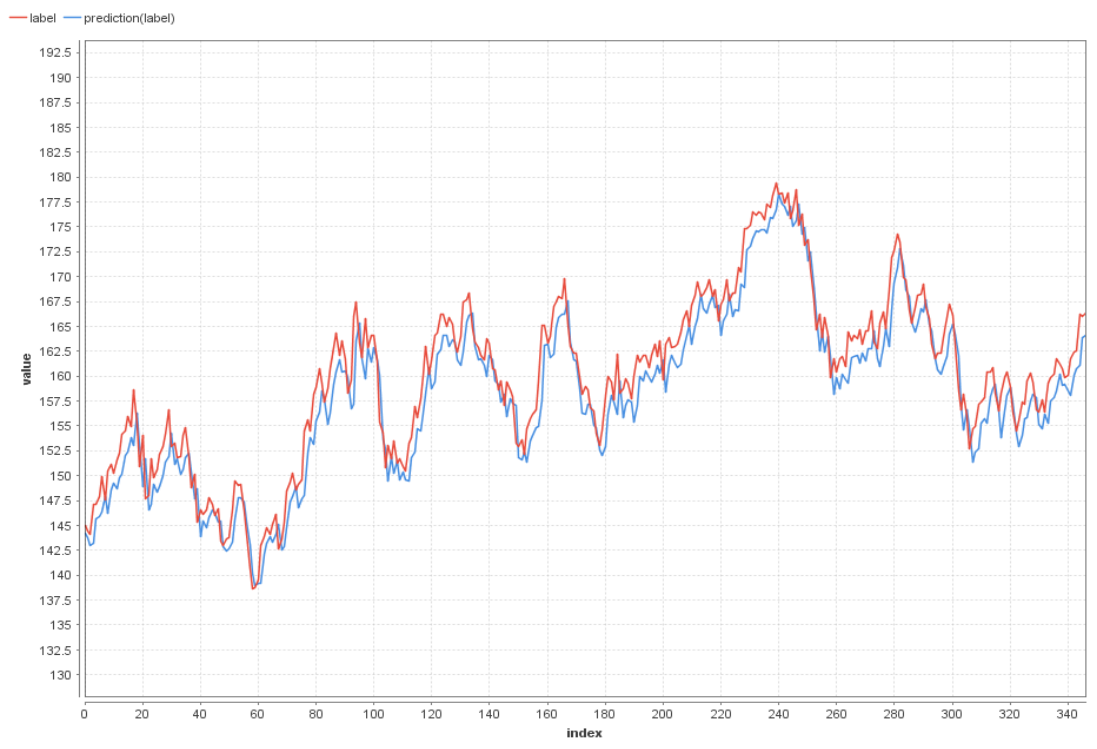
Name	Type	Missing	Statistics	Filter (18 / 18 attributes):
Id Date	Polynomial	0	Least 2018-01-12 (0) Most 2013-01-16 (1)	Search for Attributes
Label label	Numeric	0	Min 138.600 Max 261.010	Average 186.032
Prediction prediction(label)	Numeric	0	Min 140.548 Max 260.094	Average 186.565

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 7 προηγούμενες μέρες) με μπλε χρώμα



Γράφημα 57 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία GS

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 7 μέρες είναι :

root_mean_squared_error

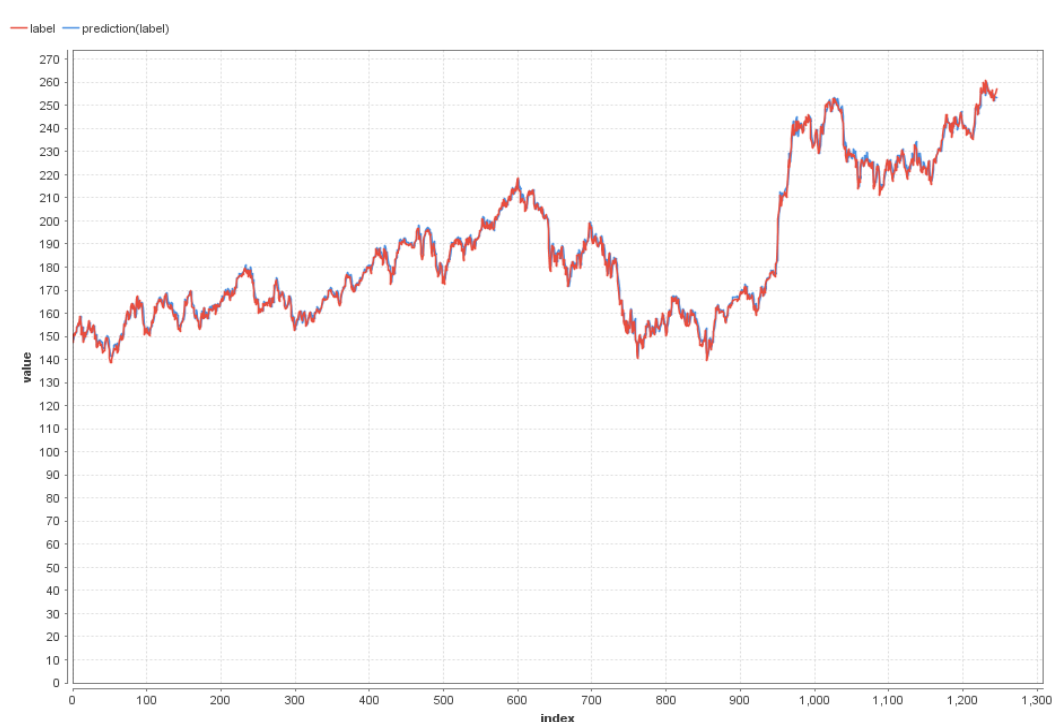
root_mean_squared_error: 3.040 +/- 0.740 (mikro: 3.130 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω ,είναι 186.165. Με βάση αυτό, το σφάλμα μας (3.040) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 42). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 3.040/186.165=0.01632$

ΠΙΝΑΚΑΣ 42: Πραγματική και προβλεπόμενη τιμή για την εταιρεία GS με χρήση νευρωνικών δικτύων με υστέρηση 7

Name	Type	Missing	Statistics		Filter (38 / 38 attributes): <input type="text" value="Search for Attributes"/>
<input checked="" type="checkbox"/> Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-23 (1)	Values 2013-01-23 (1), 2013-01-24 (1), ...[1258 more]
<input checked="" type="checkbox"/> label	Numeric	0	Min 138.600	Max 261.010	Average 186.165
<input checked="" type="checkbox"/> prediction(label)	Numeric	0	Min 138.937	Max 260.097	Average 184.883

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 14 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 58)



Γράφημα 58 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία GS

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 14 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 2.755 +/- 0.574 (mikro: 2.814 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω, είναι 186.388. Με βάση αυτό, το σφάλμα μας (2.755) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 43). Το ποσοστό του σφάλματος είναι $\text{error/label} = 2.755/186.388=0.01478$

ΠΙΝΑΚΑΣ 43: Πραγματική και προβλεπόμενη τιμή για την εταιρεία GS με χρήση νευρωνικών δικτύων με υστέρηση 14

Name	Type	Missing	Statistics		Filter (73 / 73 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-02-01 (1)	Values 2013-02-01 (1), 2013-02-04 (1), ... [1258 more]
Label label	Numeric	0	Min 138.600	Max 261.010	Average 186.388
Prediction prediction(label)	Numeric	0	Min 141.246	Max 259.754	Average 186.935

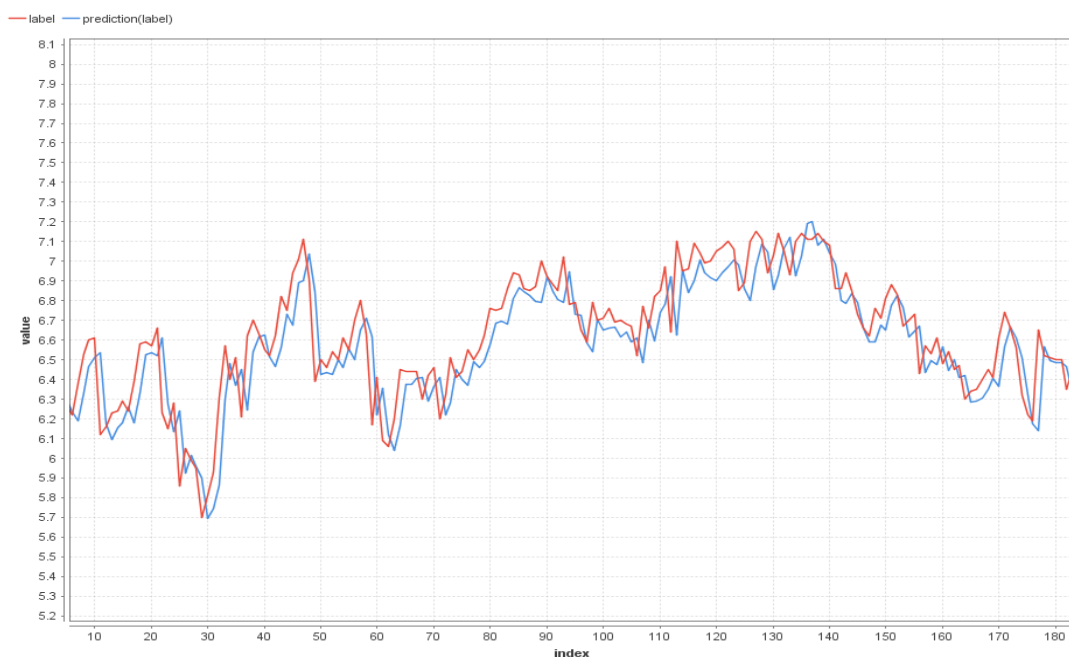
5.6 Πρόβλεψη της τιμής κλεισίματος της εταιρείας JMP

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 3 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 59)



Γράφημα 59 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία JMP

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 3 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 0.144 +/- 0.049 (mikro: 0.152 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω, είναι 6.394. Με βάση αυτό, το σφάλμα μας (0.144) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 44). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 0.144/6.394=0.02252$

ΠΙΝΑΚΑΣ 44: Πραγματική και προβλεπόμενη τιμή για την εταιρεία JMP με χρήση νευρωνικών δικτύων με υστέρηση 3

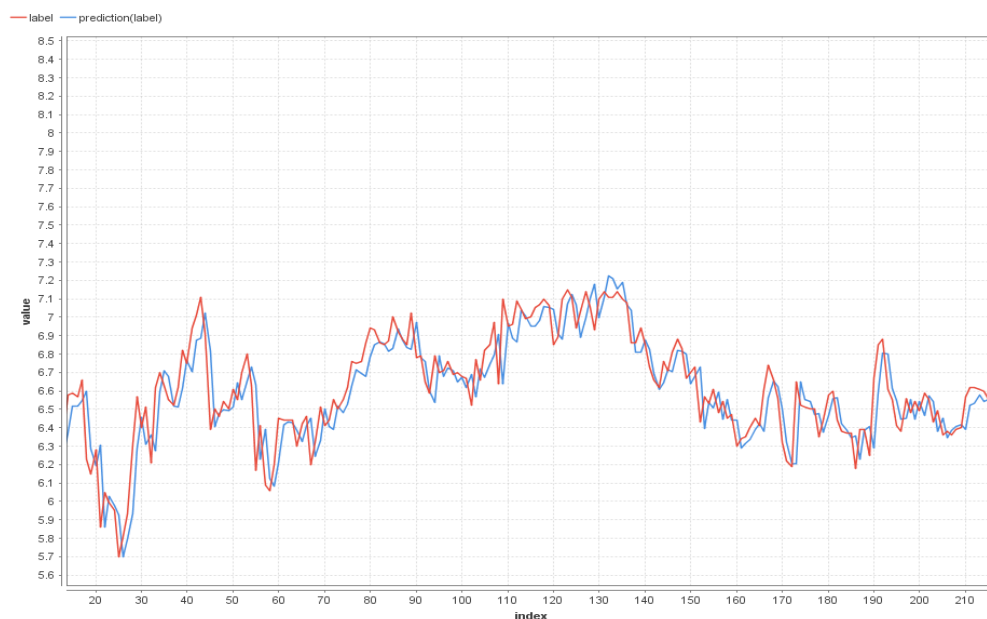
Name	Type	Missing	Statistics		Filter (18 / 18 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-16 (1)	Values 2013-01-16 (1), 2013-01-17 (1), ... [1258 more]
Label label	Numeric	0	Min 4.990	Max 8.750	Average 6.394
Prediction prediction(label)	Numeric	0	Min 4.990	Max 8.471	Average 6.339

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 7 προηγούμενες μέρες) με μπλε χρώμα (Γράφημα 60)



Γράφημα 60 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία JMP

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 7 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 0.151 +/- 0.042 (mikro: 0.157 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω, είναι 6.394. Με βάση αυτό, το σφάλμα μας (0.151) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 45). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 0.151/6.394=0.02361$

ΠΙΝΑΚΑΣ 45: Πραγματική και προβλεπόμενη τιμή για την εταιρεία JMP με χρήση νευρωνικών δικτύων με υστέρηση 7

Name	Type	Missing	Statistics		Filter (38 / 38 attributes): <input type="text" value="Search for Attributes"/>
<input checked="" type="checkbox"/> Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-23 (1)	Values 2013-01-23 (1), 2013-01-24 (1), ...[1258 more]
<input checked="" type="checkbox"/> label	Numeric	0	Min 4.990	Max 8.750	Average 6.394
<input checked="" type="checkbox"/> prediction(label)	Numeric	0	Min 4.961	Max 8.521	Average 6.364

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 14 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 61)



Γράφημα 61 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία JPM

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 14 μέρες είναι

root_mean_squared_error

root_mean_squared_error: 0.169 +/- 0.050 (mikro: 0.176 +/- 0.000)

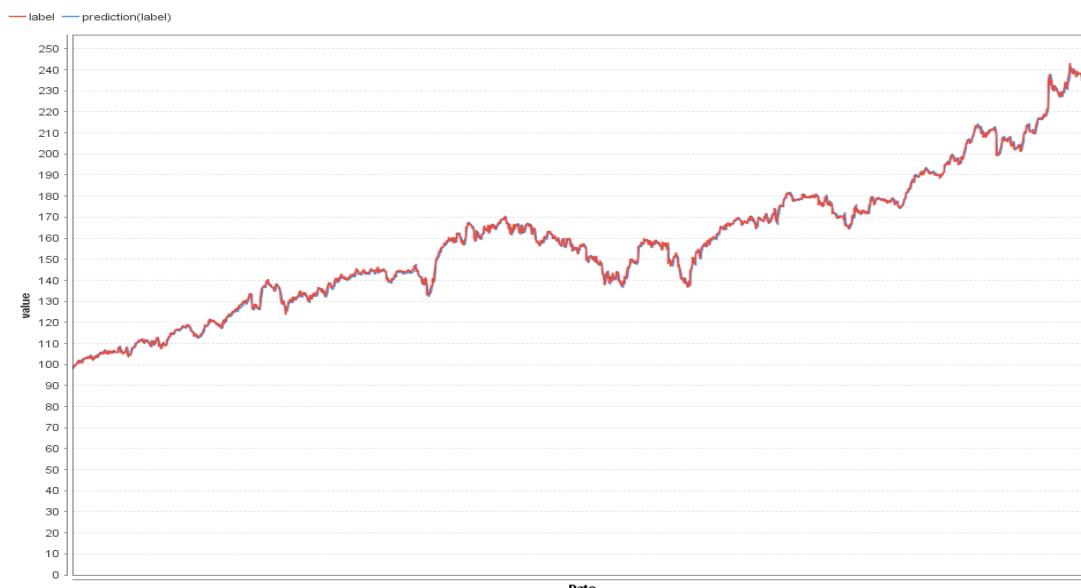
Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω, είναι 6.394. Με βάση αυτό, το σφάλμα μας (0.169) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 46). Το ποσοστό του σφάλματος είναι $\text{error/label} = 0.169/6.394=0.02643$

ΠΙΝΑΚΑΣ 46: Πραγματική και προβλεπόμενη τιμή για την εταιρεία JMP με χρήση νευρωνικών δικτύων με υστέρηση 14

Name	Type	Missing	Statistics		Filter (73 / 73 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-02-01 (1)	Values 2013-02-01 (1), 2013-02-04 (1), ... [1258 more]
Label label	Numeric	0	Min 4.990	Max 8.750	Average 6.394
Prediction prediction(label)	Numeric	0	Min 5.052	Max 8.611	Average 6.402

5.7 Πρόβλεψη της τιμής κλεισίματος της εταιρείας MMM

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 3 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 62)



Γράφημα 62 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία MMM

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 3 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 1.593 +/- 0.210 (mikro: 1.607 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω, είναι 158.425. Με βάση αυτό, το σφάλμα μας (1.593) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 47). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 1.593/158.425=0.0100$

ΠΙΝΑΚΑΣ 47: Πραγματική και προβλεπόμενη τιμή για την εταιρεία MMM με χρήση νευρωνικών δικτύων με υστέρηση 3

Name	Type	Missing	Statistics	Filter (18 / 18 attributes):
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-16 (1)
Label	Real	0	Min 98.080	Max 244.470
Prediction prediction(label)	Real	0	Min 98.265	Max 241.636
			Average 158.425	Average 158.222

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 7 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 63)



Γράφημα 63 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία MMM

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 7 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 1.660 +/- 0.218 (mikro: 1.673 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω, είναι 158.615. Με βάση αυτό, το σφάλμα μας (1.660) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 48). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 1.660/158.615=0.0104$

ΠΙΝΑΚΑΣ 48: Πραγματική και προβλεπόμενη τιμή για την εταιρεία MMM με χρήση νευρωνικών δικτύων με υστέρηση 7

Name	Type	Missing	Statistics	Filter (38 / 38 attributes):	
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-23 (1)	Values 2013-01-23 (1), 2013-01-24 (1), ...[1258 more]
Label label	Real	0	Min 99.670	Max 244.470	Average 158.615
Prediction prediction(label)	Real	0	Min 100.933	Max 242.093	Average 159.332

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 14 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 64)



Γράφημα 64 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία MMM

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 14 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 1.599 +/- 0.215 (mikro: 1.613 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής όπως φαίνεται και παρακάτω είναι 158.940. Με βάση αυτό, το σφάλμα μας (1.599) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 49). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 1.599/158.940=0.0100$

ΠΙΝΑΚΑΣ 49: Πραγματική και προβλεπόμενη τιμή για την εταιρεία MMM με χρήση νευρωνικών δικτύων με υστέρηση 14

Name	Type	Missing	Statistics		Filter (73 / 73 attributes): <input type="text" value="Search for Attributes"/>
<input checked="" type="checkbox"/> Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-02-01 (1)	Values 2013-02-01 (1), 2013-02-04 (1), ...[1258 more]
<input checked="" type="checkbox"/> label	Real	0	Min 100.770	Max 244.470	Average 158.940
<input checked="" type="checkbox"/> prediction(label)	Real	0	Min 100.228	Max 241.286	Average 158.155

5.8 Πρόβλεψη της τιμής κλεισίματος της εταιρείας LMT

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 3 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 65)



Γράφημα 65 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία LMT

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 3 μέρες είναι :

root_mean_squared_error

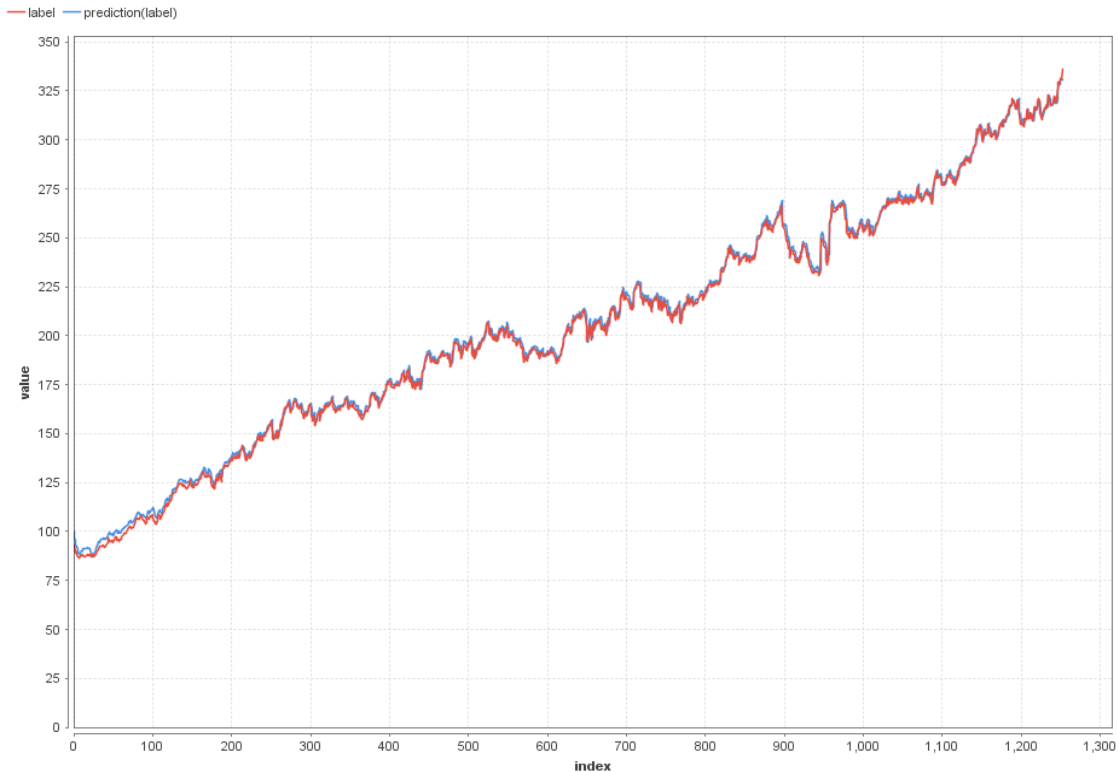
root_mean_squared_error: 3.045 +/- 1.737 (mikro: 3.507 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής, όπως φαίνεται και παρακάτω είναι 203.834. Με βάση αυτό, το σφάλμα μας (3.045) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 50). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 3.045/203.834 = 0.01493$

ΠΙΝΑΚΑΣ 50: Πραγματική και προβλεπόμενη τιμή για την εταιρεία LMT με χρήση νευρωνικών δικτύων με υστέρηση 3

Name	Type	Missing	Statistics		Filter (18 / 18 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-16 (1)	Values 2013-01-16 (1), 2013-01-17 (1), ...[1258 more]
Label label	Numeric	0	Min 86.700	Max 336.250	Average 203.834
Prediction prediction(label)	Numeric	0	Min 84.782	Max 331.402	Average 204.383

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 7 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 66)



Γράφημα 66 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία LMT

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 7 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 2.912 +/- 1.071 (mikro: 3.101 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής, όπως φαίνεται και παρακάτω είναι 204.181. Με βάση αυτό, το σφάλμα μας (2.912) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 51). Το ποσοστό του σφάλματος είναι $\text{error/label} = 2.912 / 204.181 = 0.01426$

ΠΙΝΑΚΑΣ 51: Πραγματική και προβλεπόμενη τιμή για την εταιρεία LMT με χρήση νευρωνικών δικτύων με υστέρηση 7

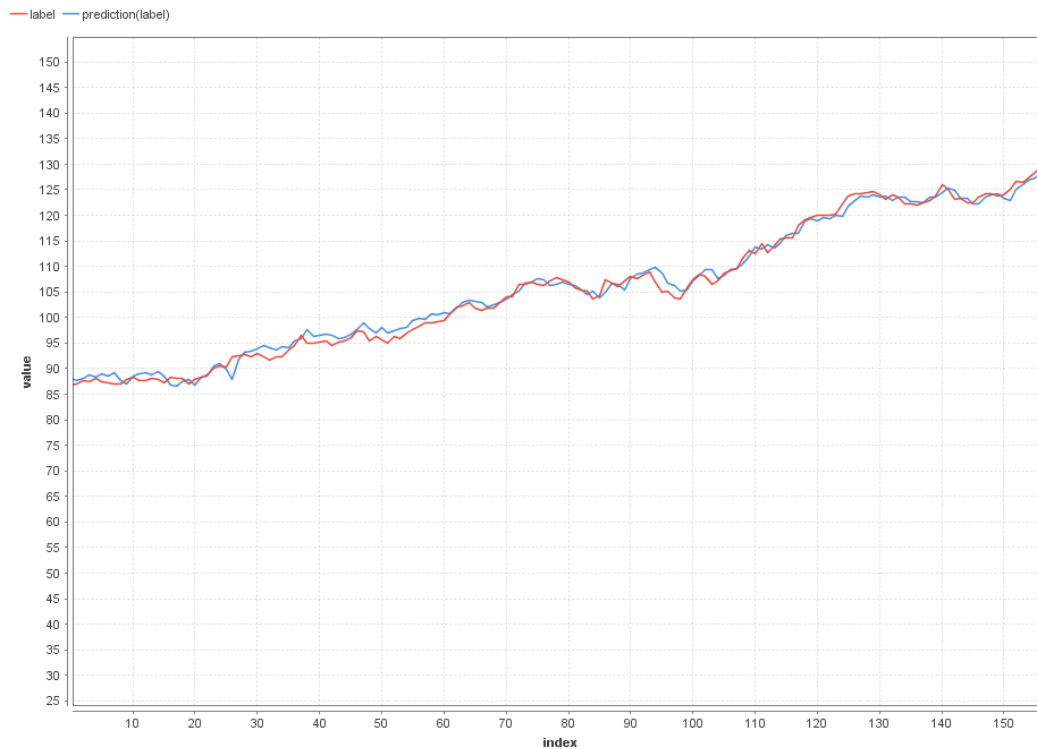
Name	Type	Missing	Statistics		Filter (38 / 38 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-23 (1)	Values 2013-01-23 (1), 2013-01-24 (1), ...[1258 more]
Label label	Numeric	0	Min 86.700	Max 336.250	Average 204.181
Prediction prediction(label)	Numeric	0	Min 87.153	Max 331.046	Average 205.483

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 14 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 67)



Γράφημα 67 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία LMT

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 14 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 3.067 +/- 2.129 (mikro: 3.738 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής, όπως φαίνεται και παρακάτω είναι 204.825. Με βάση αυτό, το σφάλμα μας (3.067) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 52). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 3.067/204.825 = 0.01494$

ΠΙΝΑΚΑΣ 52: Πραγματική και προβλεπόμενη τιμή για την εταιρεία LMT με χρήση νευρωνικών δικτύων με υστέρηση 14

Name	Type	Missing	Statistics		Filter (73 / 73 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-02-01 (1)	Values 2013-02-01 (1), 2013-02-04 (1), ... [1258 more]
Label label	Numeric	0	Min 86.700	Max 336.250	Average 204.825
Prediction prediction(label)	Numeric	0	Min 86.618	Max 328.785	Average 204.001

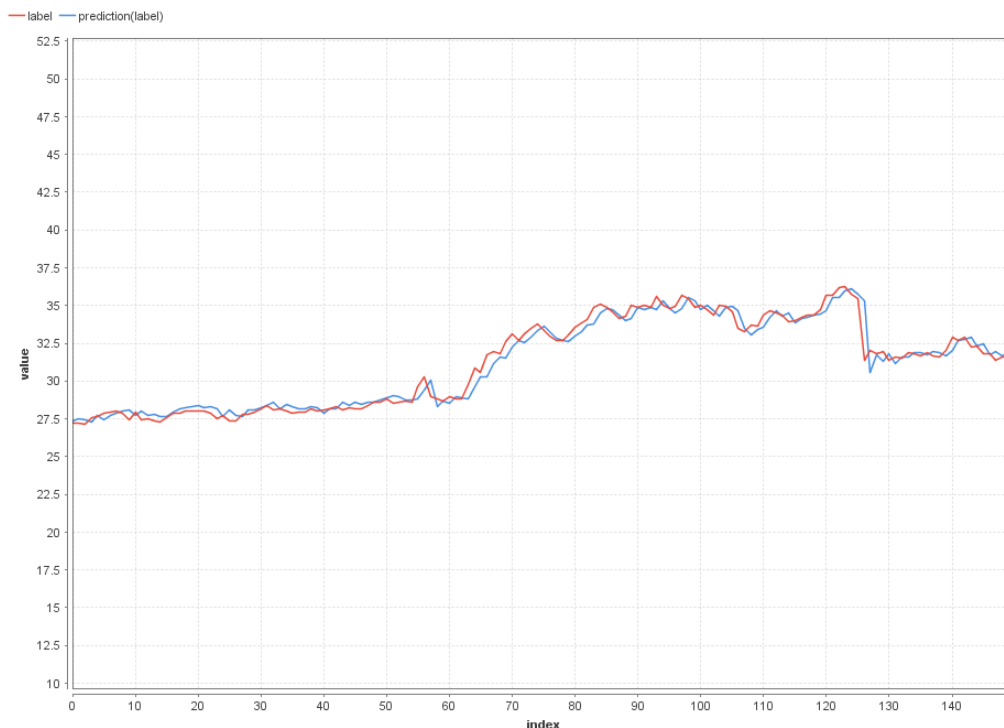
5.9 Πρόβλεψη της τιμής κλεισίματος της εταιρείας MSFT

Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 3 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 68)



Γράφημα 68 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 3 για την εταιρεία MSFT

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 3 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 0.834 +/- 0.342 (mikro: 0.901 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής, όπως φαίνεται και παρακάτω είναι 50.235. Με βάση αυτό, το σφάλμα μας (0.834) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 53) Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 0.834/50.235=0.01660$

ΠΙΝΑΚΑΣ 53: Πραγματική και προβλεπόμενη τιμή για την εταιρεία MSFT με χρήση νευρωνικών δικτύων με υστέρηση 3

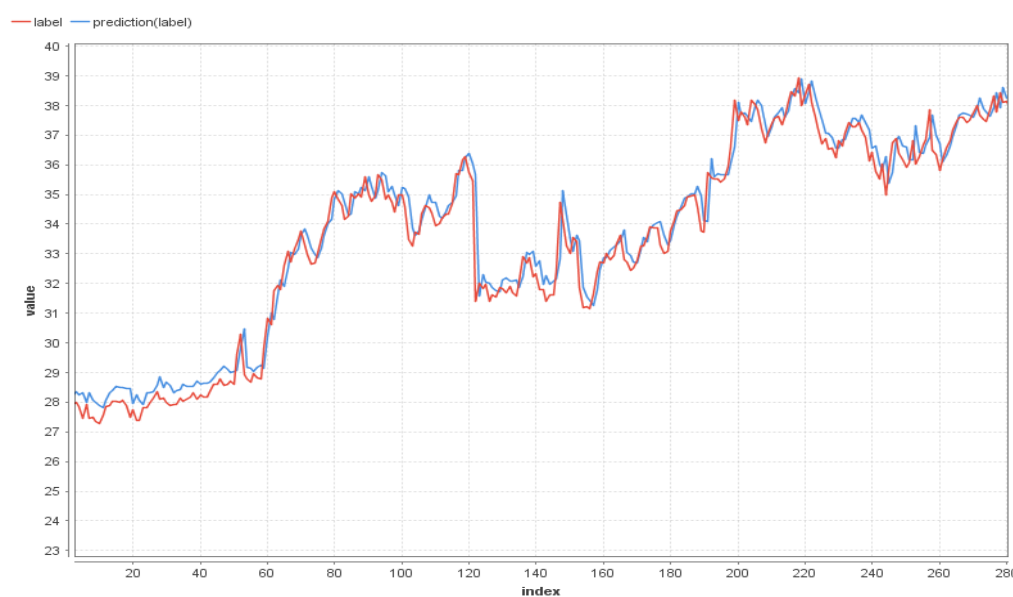
Name	Type	Missing	Statistics		Filter (18 / 18 attributes): <input type="text" value="Search for Attributes"/>
✓ Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-16 (1)	Values 2013-01-16 (1), 2013-01-17 (1), ...[1258 more]
✓ Label label	Numeric	0	Min 27.150	Max 89.600	Average 50.235
✓ Prediction prediction(label)	Numeric	0	Min 27.304	Max 88.458	Average 50.263

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 7 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 69)



Γράφημα 69 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 7 για την εταιρεία MSFT

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 7 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 0.798 +/- 0.261 (mikro: 0.839 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής, όπως φαίνεται και παρακάτω είναι 50.308. Με βάση αυτό, το σφάλμα μας (0.798) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή (βλ. Πίνακα 54). Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 0.798/50.308 = 0.01586$

ΠΙΝΑΚΑΣ 54: Πραγματική και προβλεπόμενη τιμή για την εταιρεία MSFT με χρήση νευρωνικών δικτύων με υστέρηση 7

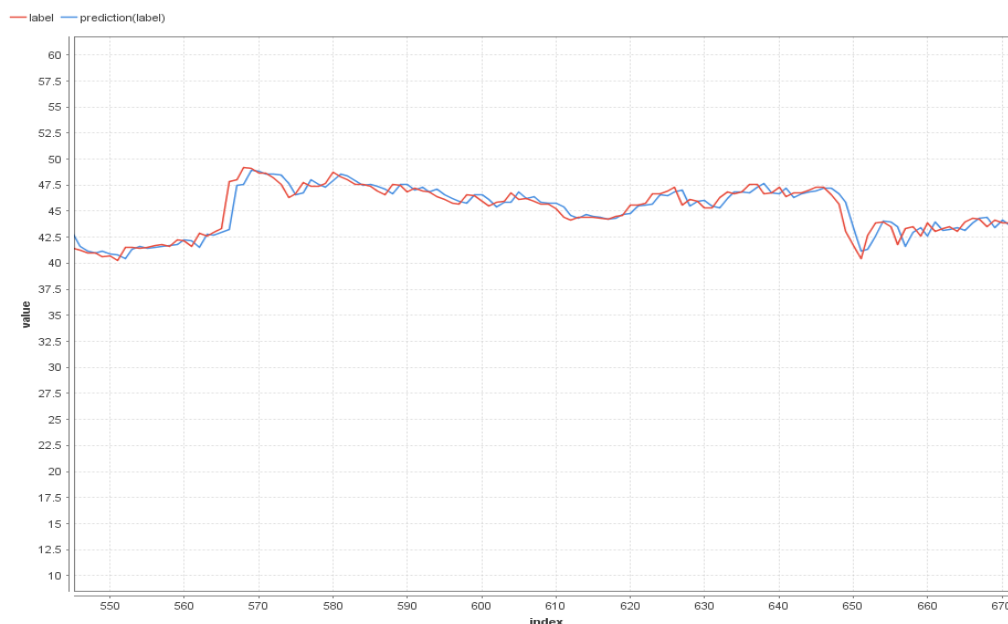
Name	Type	Missing	Statistics		Filter (38 / 38 attributes): <input type="text" value="Search for Attributes"/>
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-01-23 (1)	Values 2013-01-23 (1), 2013-01-24 (1), ...[1258 more]
Label label	Numeric	0	Min 27.280	Max 89.600	Average 50.308
Prediction prediction(label)	Numeric	0	Min 27.501	Max 88.087	Average 50.346

- Γραφική παράσταση που δείχνει την πραγματική τιμή με κόκκινο και τη προβλεπόμενη τιμή (λαμβάνοντας υπόψη τις 14 προηγούμενες μέρες) με μπλε χρώμα (βλ. Γράφημα 70)



Γράφημα 70 : Γραφική παράσταση πραγματικής και προβλεπόμενης τιμής με υστέρηση 14 για την εταιρεία MSFT

Αν εστιάσουμε λίγο καλύτερα θα δούμε η προβλεπόμενη τιμή πλησιάζει αρκετά την πραγματική οπότε το μοντέλο μας είναι αρκετά αποδοτικό



Το σφάλμα πρόβλεψης για τις 14 μέρες είναι :

root_mean_squared_error

root_mean_squared_error: 0.890 +/- 0.412 (mikro: 0.981 +/- 0.000)

Ο μέσος όρος της πραγματικής τιμής, όπως φαίνεται και παρακάτω είναι 50.434. Με βάση αυτό, το σφάλμα μας (0.890) δεν είναι μεγάλο σε σχέση με την πραγματική τιμή. Το ποσοστό του σφάλματος είναι $\text{error}/\text{label} = 0.890/50.434 = 0.01764$

ΠΙΝΑΚΑΣ 55: Πραγματική και προβλεπόμενη τιμή για την εταιρεία MSFT με χρήση νευρωνικών δικτύων με υστέρηση 14

Name	Type	Missing	Statistics		Filter (73 / 73 attributes):
Id Date	Polynomial	0	Least 2018-01-12 (0)	Most 2013-02-01 (1)	Search for Attributes
Label label	Numeric	0	Min 27.280	Max 89.600	Average 50.434
Prediction prediction(label)	Numeric	0	Min 27.680	Max 88.177	Average 50.504

ΣΥΝΟΨΗ

ΠΙΝΑΚΑΣ 56: Αποτελέσματα RMSE με υστέρηση 3,7 και 14 για κάθε εταιρεία με χρήση Νευρωνικών δικτύων.

Εταιρεία	NEURAL NETWORK	
	AR	RMSE
Facebook	p=3	1,966
	p=7	2,091
	p=14	2,558
Google	p=3	9,938
	p=7	9,936
	p=14	10,360
MSFT	p=3	0,834
	p=7	0,798
	p=14	0,890
BA	p=3	3,302
	p=7	3,439
	p=14	3,900
LMT	p=3	3,045
	p=7	2,912
	p=14	3,067
MMM	p=3	1,593
	p=7	1,660
	p=14	1,599
BLK	p=3	4,957
	p=7	6,204
	p=14	5,740
JPM	p=3	0,144
	p=7	0,151
	p=14	0,169
GS	p=3	2,616
	p=7	3,040

	p=14	2,755
--	------	-------

Στο κομμάτι των νευρωνικών δικτύων στον παραπάνω πίνακα, με κόκκινο έχουν σημειωθεί τα πιο επιτυχημένα μοντέλα που ανέδειξε η μοντελοποίηση για κάθε περίπτωση, από τα οποία παρατηρούμε πως κανένα δεν είναι τάξης 14 (βλ. Πίνακας 56). Αυτό σημαίνει πως οι 14 ημέρες δεν είναι επαρκείς για την αποθήκευση πληροφοριών που χρειάζονται για την βέλτιστη εκμάθηση του μοντέλου.

6. ΣΥΜΠΕΡΑΣΜΑΤΑ

ΠΙΝΑΚΑΣ 57: Σύγκριση αποτελεσμάτων RMSE με υστέρηση 3,7 και 14 για κάθε εταιρεία με χρήση μεθόδου ARIMA και Νευρωνικών δικτύων.

	ARIMA Models				NEURAL NETWORKS	
Εταιρεία	MA	Difference	AR	RMSE	AR	RMSE
Facebook	d=1	q=1	p=3	2,576	p=3	1,966
			p=7	2,570	p=7	2,091
			p=14	2,571	p=14	2,558
Google	d=1	q=1	p=3	32,898	p=3	9,938
			p=7	31,486	p=7	9,936
			p=14	33,156	p=14	10,360
MSFT	d=1	q=1	p=3	2,183	p=3	0,834
			p=7	2,193	p=7	0,798
			p=14	1,967	p=14	0,890
BA	d=1	q=1	p=3	22,581	p=3	3,302
			p=7	22,530	p=7	3,439
			p=14	22,566	p=14	3,900
LMT	d=1	q=1	p=3	8,083	p=3	3,045
			p=7	7,546	p=7	2,912
			p=14	7,621	p=14	3,067
MMM	d=1	q=1	p=3	5,022	p=3	1,593
			p=7	4,556	p=7	1,660
			p=14	5,050	p=14	1,599
BLK	d=0	q=1	p=3	26,471	p=3	4,957
			p=7	26,530	p=7	6,204
			p=14	26,359	p=14	5,740
JPM	d=0	q=1	p=3	0,085	p=3	0,144
			p=7	0,084	p=7	0,151
			p=14	0,089	p=14	0,169
GS	d=1	q=1	p=3	6,989	p=3	2,616
			p=7	6,975	p=7	3,040

		p=14	7,085	p=14	2,755
--	--	------	-------	------	-------

Αυτή η μελέτη συνέκρινε τα ANN και τα μοντέλα ARIMA για την πρόβλεψη των τιμών κλεισίματος εννέα εταιρειών. Τόσο το μοντέλο ARIMA όσο και το μοντέλο ANN μπορούν να επιτύχουν καλές προβλέψεις σε προβλήματα πραγματικής ζωής και έτσι να μπορούν να εμπλακούν αποτελεσματικά κερδοφόρα στην πρόβλεψη των τιμών των μετοχών. Ωστόσο το αναπτυγμένο μοντέλο των τιμών των μετοχών με την προσέγγιση που βασίζεται στο ANN έδειξε ανώτερη απόδοση έναντι των μοντέλων ARIMA. Πιο συγκεκριμένα σε όλες τις εταιρείες εκτός από την εταιρεία JMP η τιμή του σφάλματος του μοντέλου ANN γνωρίζοντας τις 3, 7 και 14 μέρες είναι εμφανώς αρκετά μικρότερη από αυτή του μοντέλου ARIMA. Μάλιστα στις περιπτώσεις των εταιρειών GOOGLE, BA και BLK η τιμή του σφάλματος με τη μέθοδο ARIMA ξεπερνά κατά πολύ αυτή των νευρωνικών δικτύων (βλ Πίνακας 57). Ο λόγος όπου το μοντέλο ANN είναι καλύτερα από το ARIMA μπορεί να οφείλεται στο γεγονός ότι τα δεδομένα περιέχουν μη γραμμική συμπεριφορά, η οποία δεν μπορεί να συλληφθεί πλήρως από το γραμμικό μοντέλο ARIMA.

Η βελτίωση της πρόβλεψης και κυρίως της ακρίβειας των προβλέψεων των χρονοσειρών, αποτελεί ένα σημαντικό αλλά συχνά δύσκολο έργο που καλούνται να αντιμετωπίσουν οι υπεύθυνοι λήψης αποφάσεων σε πολλούς τομείς. Το γραμμικό μοντέλο ARIMA και το μη γραμμικό μοντέλο ANN χρησιμοποιούνται από κοινού, με στόχο τη συλλογή διαφορετικών μορφών σχέσης στα δεδομένα χρονοσειρών. Τα τεχνητά νευρικά δίκτυα (ANN) έχουν αποδειχθεί ότι είναι μια αποτελεσματική προσέγγιση γενικής χρήσης για την αναγνώριση προτύπων, την ταξινόμηση, την ομαδοποίηση και ιδιαίτερα την πρόβλεψη με υψηλό βαθμό ακρίβειας.

Οι πραγματικές χρονοσειρές ωστόσο είναι σπάνια γραμμικές ή μη γραμμικές αφού συχνά περιέχουν τόσο γραμμικά όσο και μη γραμμικά πρότυπα. Σε αυτή την περίπτωση ούτε το ARIMA ούτε το ANNs μπορεί να είναι επαρκές για τη μοντελοποίηση και την πρόβλεψη χρονοσειρών, αφού το μοντέλο ARIMA δεν μπορεί να ασχοληθεί με τις μη γραμμικές σχέσεις, ενώ από την άλλη το μοντέλο νευρωνικού δικτύου από μόνο του δεν είναι σε θέση να χειριστεί τόσο γραμμικά όσο και μη γραμμικά πρότυπα εξίσου καλά. Για σύνθετα προβλήματα που έχουν τόσο γραμμικές όσο και μη γραμμικές δομές συσχέτισης, πλέον πραγματοποιούνται μέθοδοι συνδυασμού των μοντέλων. Το υβριδικό μοντέλο εκμεταλλεύεται τη μοναδική

δύναμη των ARIMA και ANN σε γραμμική και μη γραμμική μοντελοποίηση μπορεί να είναι ένας αποτελεσματικός τρόπος βελτίωσης των επιδόσεων πρόβλεψης.

Βιβλιογραφία

1. Alkon, D.L., Vogl, T.P., Blackwell, K.T. and Tam, D., 2016. Memory function in neural and artificial networks. *The Really Useful Book of Secondary Science Experiments: 101 Essential Activities to Support Teaching and Learning*.
2. Amato, F., López, A., Peña-Méndez, E. M., Vañhara, P., Hampl, A., & Havel, J. (2013). Artificial neural networks in medical diagnosis.
3. Basheer, I.A. & Hajmeer, M., (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of microbiological methods*, 43(1), pp.3-31.
4. Esfe, M. H., Yan, W. M., Afrand, M., Sarraf, M., Toghraie, D., & Dahari, M. (2016). Estimation of thermal conductivity of Al₂O₃/water (40%)–ethylene glycol (60%) by artificial neural network and correlation using experimental data. *International Communications in Heat and Mass Transfer*, 74, 125-128.
5. Graupe, D. (2013). *Principles of artificial neural networks* (Vol. 7). World Scientific.
6. Hirasawa, N., Takaya, K., Akiyama, Y., Amemiya, F., Sugiura, A., & Wada, O. (2016). Technical problems of the shunt-type asymmetric artificial network. *IEEE Transactions on Electromagnetic Compatibility*, 58(5), 1526-1534.
7. <http://users.auth.gr/dkugiu/Teach/TimeSeries/TimeSeries.pdf>
8. https://en.wikipedia.org/wiki/Autoregressive_model
9. https://en.wikipedia.org/wiki/Moving-average_model
10. Lee, S., Seong, W. J., Oh, K. Y., & Lee, M. J. (2016). The spatial prediction of landslide susceptibility applying artificial neural network and logistic regression models; A case study of Inje, Korea. *Open Geosciences*, 8(1), 117-132.
11. Valipour, M., Banihabib, M. E., & Behbahani, S. M. R. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network

- models in forecasting the monthly inflow of Dez dam reservoir. *Journal of hydrology*, 476, 433-441.
12. Κουγιουμτζης Δ. «Αναλυση Χρονοσειρων», Μάθημα Μεταπτυχιακού Προγράμματος Ειδίκευσης «Στατιστική Και Μοντελοποίηση»
 13. Νικολαου Ι.Ε. (2007) «Συγκριτική Αναλυση Και Εφαρμογη ΓραμμικΩν, Μη-ΓραμμικΩν Και Νευρο-Ασαφων Μεθοδων, Για Τη Βραχυπροθεσμη Προβλεψη Παραγωγιης Ενεργιαις Απο Αιολικα Παρκα », Διπλωματική Εργασία, Χανιά
 14. Nowrouz Kohzadi , Milton S. Boyd , Bahman Kermanshahi , Iebeling Kaastra
A comparison of artificial neural network and time series models for forecasting commodity prices
 15. Σοφία Δ. « Σύγχρονες μέθοδοι ανάλυσης χρονολογικών σειρών » Εκδόσεις Ο.Π.Α (Οικονομικό Πανεπιστήμιο Αθηνών), Αθήνα 2013
 16. Λουλάκης, Μ, « Στοχαστικές Διαδικασίες » Εκδόσεις Κάλλιπος, Αθήνα 2015
 17. Shakira Green-Georgia Southern University 2011: “*Time Series Analysis of Stock Prices Using the BoxJenkins Approach*”
 18. Uma Devi, D.Sundar and Dr Alli 2013: “*An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50*”
 19. Mahantesh Angadi και Amogh Kulkarni 2015: “*Time Series Data Analysis for Stock Market Prediction using Data Mining Techniques with R*”
 20. Fahim Faisal May 2012 “Forecasting Bangladesh's Inflation Using Time Series ARIMA Models”
 21. Κωνσταντίνος Διαμαντάρας «*Τεχνητά Νευρωνικά Δίκτυα*». Εκδόσεις Κλειδάριθμος, 2007
 22. Maureen Caudill, AI Expert «*Neural Network Primer: Part I*», Φεβρουάριος 1989

ΠΑΡΑΡΤΗΜΑ (Κώδικας για μία εταιρεία)

- **ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΜΕ ΜΕΘΟΔΟ ARIMA ΣΕ ΓΛΩΣΣΑ R**

```
###Πρόβλεψη της τιμής κλεισίματος για το BA###
```

```
###Εισαγωγή δεδομένων###
```

```
BA=read.csv(file.choose(),header=T)
```

```
###Μετατροπή των δεδομένων σε μορφή χρονοσειράς###
```

```
BA_test=ts(BA$Close,start=1,end=1210)
```

```
###Απεικόνιση της χρονοσειράς###
```

```
plot((BA_test))
```

```
abline( reg=lm(BA_test~time(BA_test)),col=2 )
```

```
###απαλοιφή τυχαίων διακυμάνσεων με χρήση της λογαριθμικής συνάρτησης###
```

```
log_BA_test=log(BA_test)
```

```
###Γράφημα της χρονοσειράς μετά την απαλοιφή των διακυμάνσεων###
```

```
plot((log_BA_test))
```

```
abline( reg=lm(log_BA_test~time(log_BA_test)),col=2 )
```

```
###Απαλοιφή τάσης με χρήση πρώτων διαφορών###
```

```
diff_log_BA_test=diff(BA_test)
```

```
###Γράφημα της χρονοσειράς μετά την απαλοιφή της τάσης###
```

```
plot(diff_log_BA_test)
```

```
abline( reg=lm(diff_log_BA_test~time(diff_log_BA_test)),col=2 )
```

```
###Γράφημα αυτοσυσχετίσεων προς εκτίμηση της τάξης του MA μοντέλου###
```

```
acf(diff_log_BA_test)
```

```
###Γράφημα μερικών αυτοσυσχετίσεων προς εκτίμηση της τάξης του AR μοντέλου
```

```
pacf(diff_log_BA_test)
```

```
###Δημιουργία των μοντέλων και πρόβλεψη###
```

```
fit3=arima(log_BA_test,c(3,1,1))
```

```
PRED3=predict(fit3,n.ahead=50)$pred
```

```
PRED3=exp(1)^PRED3
```

```
BA_test=ts(BA$Close,start=1,end=1210)
```

```
log_BA_test=log(BA_test)
```

```
diff_log_BA_test=diff(BA_test)
```

```
fit7=arima(log_BA_test,c(7,1,1))
```

```
PRED7=predict(fit7,n.ahead=50)$pred
```

```
PRED7=exp(1)^PRED7
```

```
BA_test=ts(BA$Close,start=1,end=1210)
```

```
log_BA_test=log(BA_test)
```

```

diff_log_BA_test=diff(BA_test)

fit14=arima(log_BA_test,c(14,1,1))

PRED14=predict(fit14,n.ahead=50)$pred

PRED14=exp(1)^PRED14

###Πίνακας με τις αρχικές και τις προβλεπόμενες τιμές###

COMPARE3=rbind(BA$Close[1211:1260],PRED3)

COMPARE3=data.frame(COMPARE3)

rownames(COMPARE3)=c("Original Data","Model Prediction")

for(i in 1:50)

colnames(COMPARE3)[i]=paste("Obs ",1210+i)

COMPARE7=rbind(BA$Close[1211:1260],PRED7)

COMPARE7=data.frame(COMPARE7)

rownames(COMPARE7)=c("Original Data","Model Prediction")

for(i in 1:50)

colnames(COMPARE7)[i]=paste("Obs ",1210+i)

COMPARE14=rbind(BA$Close[1211:1260],PRED14)

COMPARE14=data.frame(COMPARE14)

rownames(COMPARE14)=c("Original Data","Model Prediction")

for(i in 1:50)

colnames(COMPARE14)[i]=paste("Obs ",1210+i)

```

###Εκτύπωση των πινάκων σύγκρισης###

COMPARE3

COMPARE7

COMPARE14

###Εύρεση μέσου τετραγωνικού σφάλματος μεταξύ των προβλέψεων του μοντέλου και των αρχικών τιμών###

RMSE3= sqrt(sum(((BA\$Close[1211:1260]-PRED3)^2))/100)

RMSE7= sqrt(sum(((BA\$Close[1211:1260]-PRED7)^2))/100)

RMSE14= sqrt(sum(((BA\$Close[1211:1260]-PRED14)^2))/100)

RMSE3

RMSE7

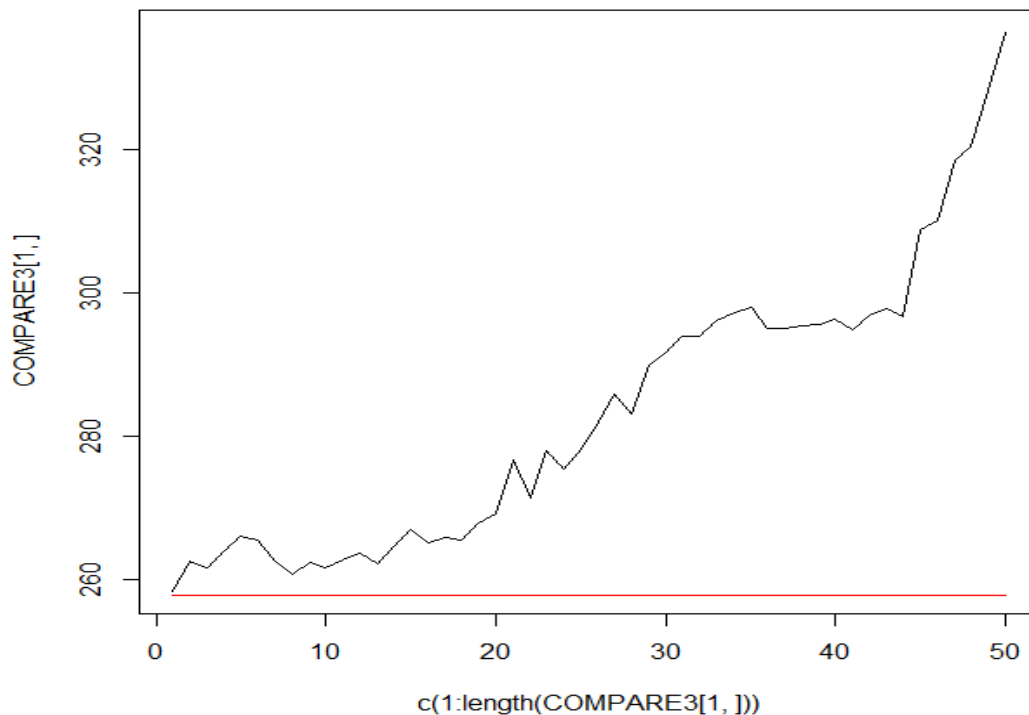
RMSE14

x=COMPARE3[1,]

plot(c(1:length(COMPARE3[1,])),COMPARE3[1,],pch=".",type="line")

points(c(1:length(COMPARE3[2,])),COMPARE3[2,],pch=".",type="line",col=2)

Γράφημα με ταυτόχρονη αποτύπωση της πραγματικής και της προβλεπόμενης τιμής για την εταιρεία BA με το μοντέλο 3^{ων} ημερών:



- **ΚΩΔΙΚΑΣ ΓΙΑ ΜΙΑ ΕΤΑΙΡΕΙΑ ΜΕ ΜΕΘΟΔΟ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΣΕ ΠΡΟΓΡΑΜΜΑ RAPIDMINER**

```
<?xml version="1.0" encoding="UTF-8"?><process version="8.1.000">
```

```
<context>
```

```
<input/>
```

```
<output/>
```

```
<macros/>
```

```
</context>
```

```
<operator activated="true" class="process" compatibility="8.1.000"
expanded="true" name="Process">
```

```
<process expanded="true">
```

```
<operator activated="true" class="read_excel" compatibility="8.1.000"
expanded="true" height="68" name="Read BA Data" width="90" x="112" y="238">
```



```
<parameter key="excel_file" value="C:\Users\Stella\Desktop\dedomena\BA-
_3_.xlsx"/>

<parameter key="imported_cell_range" value="A1:G1261"/>

<parameter key="first_row_as_names" value="false"/>

<list key="annotations">

  <parameter key="0" value="Name"/>

</list>

<list key="data_set_meta_data_information">

  <parameter key="0" value="Date.true.polynomial.id"/>

  <parameter key="1" value="Open.true.real.attribute"/>

  <parameter key="2" value="High.true.numeric.attribute"/>

  <parameter key="3" value="Low.true.numeric.attribute"/>

  <parameter key="4" value="Close.true.numeric.attribute"/>

  <parameter key="5" value="Adj Close.false.real.attribute"/>

  <parameter key="6" value="Volume.true.integer.attribute"/>

</list>

</operator>

<operator activated="true" class="series:windowing" compatibility="7.4.000"
expanded="true" height="82" name="Windowing" width="90" x="313" y="136">

  <parameter key="window_size" value="3"/>

  <parameter key="create_label" value="true"/>

  <parameter key="label_attribute" value="Close"/>

</operator>
```

```

<operator activated="true" class="concurrency:cross_validation"
compatibility="8.1.000" expanded="true" height="145" name="Validation"
width="90" x="581" y="136">

  <parameter key="sampling_type" value="linear sampling"/>

  <process expanded="true">

    <operator activated="true" class="neural_net" compatibility="8.1.000"
expanded="true" height="82" name="Neural Net" width="90" x="179" y="34">

      <list key="hidden_layers"/>

      <parameter key="training_cycles" value="750"/>

      <parameter key="learning_rate" value="0.1"/>

    </operator>

    <operator activated="false" class="h2o:gradient_boosted_trees"
compatibility="7.6.001" expanded="true" height="103" name="Gradient Boosted
Trees" width="90" x="112" y="340">

      <parameter key="number_of_trees" value="100"/>

      <list key="expert_parameters"/>

    </operator>

    <operator activated="false" class="h2o:deep_learning" compatibility="7.6.001"
expanded="true" height="82" name="Deep Learning" width="90" x="179" y="187">

      <parameter key="activation" value="Maxout"/>

      <enumeration key="hidden_layer_sizes">

        <parameter key="hidden_layer_sizes" value="100"/>

        <parameter key="hidden_layer_sizes" value="100"/>

      </enumeration>

      <enumeration key="hidden_dropout_ratios"/>

      <list key="expert_parameters"/>

```

```

    <list key="expert_parameters_"/>
</operator>

<operator activated="false" class="support_vector_machine"
compatibility="8.1.000" expanded="true" height="124" name="SVM" width="90"
x="313" y="238">

    <parameter key="kernel_type" value="radial"/>

    <parameter key="kernel_degree" value="3.0"/>

</operator>

<connect from_port="training set" to_op="Neural Net" to_port="training set"/>

<connect from_op="Neural Net" from_port="model" to_port="model"/>

<portSpacing port="source_training set" spacing="0"/>

<portSpacing port="sink_model" spacing="0"/>

<portSpacing port="sink_through 1" spacing="0"/>

</process>

<process expanded="true">

    <operator activated="true" class="apply_model" compatibility="8.1.000"
expanded="true" height="82" name="Apply Model" width="90" x="45" y="34">

        <list key="application_parameters"/>

</operator>

<operator activated="true" class="performance" compatibility="8.1.000"
expanded="true" height="82" name="Performance" width="90" x="246" y="34"/>

<connect from_port="model" to_op="Apply Model" to_port="model"/>

<connect from_port="test set" to_op="Apply Model" to_port="unlabelled
data"/>

<connect from_op="Apply Model" from_port="labelled data"
to_op="Performance" to_port="labelled data"/>

```

```

    <connect from_op="Performance" from_port="performance"
to_port="performance 1"/>
    <portSpacing port="source_model" spacing="0"/>
    <portSpacing port="source_test set" spacing="0"/>
    <portSpacing port="source_through 1" spacing="0"/>
    <portSpacing port="sink_test set results" spacing="0"/>
    <portSpacing port="sink_performance 1" spacing="0"/>
    <portSpacing port="sink_performance 2" spacing="0"/>
</process>

<description align="center" color="transparent" colored="false" width="126">A
cross validation including a linear regression.</description>

</operator>

<operator activated="true" class="apply_model" compatibility="8.1.000"
expanded="true" height="82" name="Apply Model (2)" width="90" x="782"
y="340">
    <list key="application_parameters"/>
</operator>

<connect from_op="Read BA Data" from_port="output" to_op="Windowing"
to_port="example set input"/>

<connect from_op="Windowing" from_port="example set output"
to_op="Validation" to_port="example set"/>

<connect from_op="Validation" from_port="model" to_op="Apply Model (2)"
to_port="model"/>

<connect from_op="Validation" from_port="example set" to_op="Apply Model
(2)" to_port="unlabelled data"/>

<connect from_op="Validation" from_port="performance 1" to_port="result 2"/>
<connect from_op="Apply Model (2)" from_port="labelled data" to_port="result
1"/>
    <portSpacing port="source_input 1" spacing="0"/>
    <portSpacing port="sink_result 1" spacing="0"/>
    <portSpacing port="sink_result 2" spacing="0"/>
    <portSpacing port="sink_result 3" spacing="0"/>
</process>

</operator>
</process>

```