

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
Σχολή Χρηματοοικονομικής και Στατιστικής



Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΠΡΟΒΛΕΨΕΙΣ ΑΠΟΔΟΣΕΩΝ
ΜΕΤΟΧΩΝ ΤΟΥ ΕΛΛΗΝΙΚΟΥ
ΧΡΗΜΑΤΙΣΤΗΡΙΟΥ ΜΕ ΤΗ ΧΡΗΣΗ
ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ**

Ραφαήλ Νικόλαος Τζανετής

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού
Διπλώματος Ειδίκευσης στην *Εφαρμοσμένη Στατιστική*

Πειραιάς
Ιανουάριος 2024

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
Σχολή Χρηματοοικονομικής και Στατιστικής



Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

ΠΡΟΒΛΕΨΕΙΣ ΑΠΟΔΟΣΕΩΝ
ΜΕΤΟΧΩΝ ΤΟΥ ΕΛΛΗΝΙΚΟΥ
ΧΡΗΜΑΤΙΣΤΗΡΙΟΥ ΜΕ ΤΗ ΧΡΗΣΗ
ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

Ραφαήλ Νικόλαος Τζανετής

Διπλωματική Εργασία

*που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού
Διπλώματος Ειδίκευσης στην Εφαρμοσμένη Στατιστική*

Πειραιάς
Ιανουάριος 2024

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Καθηγητής Χρήστος Ν. Αγιακλόγλου (Επιβλέπων)
- Αναπληρωτής Καθηγητής Κωνσταντίνος Πολίτης
- Καθηγητής Βασίλειος Σεβρόγλου

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

UNIVERSITY OF PIRAEUS
School of Finance and Statistics



Department of Statistics and Insurance Science

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

**FORCASTING STOCK RETURNS OF
THE GREEK STOCK EXCHANGE
USING MACHINE LEARNING
TECHNIQUES**

By
Rafail Nikolaos Tzanetis

MSc Dissertation

submitted to the Department of Statistics and Insurance
Science of the University of Piraeus in partial fulfilment of
the requirements for the degree of Master of Science in
Applied Statistics

Piraeus, Greece
January 2024

Στην οικογένεια μου

Στη Χρύσα

Στον Κόνορ

Ευχαριστίες

Με την παρούσα διπλωματική εργασία ολοκληρώνονται οι σπουδές μου στο Μεταπτυχιακό Πρόγραμμα Σπουδών «Εφαρμοσμένη Στατιστική», του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης, του Πανεπιστημίου Πειραιώς.

Στο σημείο αυτό θα ήθελα να εκφράσω τις θερμότερες ευχαριστίες μου προς τον κ. Αγιακλόγλου Ν. Χρήστο, Καθηγητή του Τμήματος Οικονομικής Επιστήμης του Πανεπιστημίου Πειραιώς, όπου με επιστημονική και συμβουλευτική καθοδήγηση καθ' όλη την διάρκεια, με βοήθησε στην κατανόηση και ανάπτυξη του θέματος για την εκπόνηση της διπλωματικής μου εργασίας

Παράλληλα θα ήθελα να ευχαριστήσω τα μέλη της εξεταστικής επιτροπής, τον κ. Πολίτη Κωνσταντίνο, Αναπληρωτή Καθηγητή του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης, καθώς και τον κ. Σεβρόγλου Βασίλειο, Καθηγητή του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης, για τις απόλυτα εύστοχες υποδείξεις τους, όπως επίσης θα ήθελα να τους ευχαριστήσω για την πολύτιμη γνώση που μου πρόσφεραν σε μεταπτυχιακό, αλλά και σε προπτυχιακό επίπεδο σπουδών αντίστοιχα.

Περίληψη

Σε ένα κόσμο που κυριαρχεί η αβεβαιότητα και τα πάντα είναι ρευστά, η ανάγκη του ανθρώπου για την δημιουργία ορθών προβλέψεων αποτελεί προτεραιότητα στην μείωση του ρίσκου. Η δημιουργία νέων τεχνολογιών σε συνδυασμό με την αύξηση της υπολογιστικής ισχύος, έδωσε νέα τροπή στην δημιουργία πιο εύστοχων προβλέψεων. Σε αυτή την διπλωματική εργασία, θα παρουσιαστεί αρχικά το θεωρητικό υπόβαθρο γύρω από την ανάλυση και την πρόβλεψη χρονοσειρών με κλασσικές μεθόδους. Πιο συγκεκριμένα, θα αναλυθούν τα συνθετικά στοιχεία των χρονοσειρών, τα είδη των υποδειγμάτων που χρησιμοποιούνται, διάφορα στατιστικά μέτρα, η ύπαρξη στασιμότητας, καθώς και διάφορες μέθοδοι πρόβλεψης για κάθε τύπο προβλήματος. Στο δεύτερο κεφάλαιο, θα γίνει αναφορά στη μηχανική μάθηση, στα είδη της, καθώς και στις διαφορές της σε σχέση με την βαθιά μάθηση. Παράλληλα, θα γίνει αναφορά στον βιολογικό νευρώνα, θα παρουσιαστεί εκτεταμένα το θεωρητικό πλαίσιο και ο τρόπος λειτουργίας τεσσάρων αρχιτεκτονικών τεχνητών νευρωνικών δικτύων και πιο συγκεκριμένα, των LSTM, GRU, RNN και MLP. Επιπροσθέτως, θα παρουσιαστούν οι πιο γνωστές συναρτήσεις βελτιστοποίησης, καθώς και συναρτήσεις απώλειας. Στο τρίτο κεφάλαιο, θα δοθούν κάποια γενικά γνωρίσματα γύρω από τις αποδόσεις μετοχών, όπως επίσης και τρόποι υπολογισμού τους. Στη συνέχεια, θα μελετηθούν οι λογαριθμικές αποδόσεις τριών εταιρειών της Ελληνικής χρηματιστηριακής αγοράς και συγκεκριμένα του ομίλου της ΓΕΚ ΤΕΡΝΑ Α.Ε., της ΑΕΡΟΠΟΡΙΑΣ ΑΙΓΑΙΟΥ Α.Ε. και του ΟΠΑΠ Α.Ε., με την χρήση τεσσάρων αρχιτεκτονικών νευρωνικών δικτύων και συγκεκριμένα των LSTM, GRU, RNN και MLP, καθώς και έξι διαφορετικών αλγορίθμων βελτιστοποίησης. Ακολούθως, θα συγκριθούν βάσει των πληροφοριακών κριτηρίων σφάλματος, όπου θα βρεθούν τα πιο αποδοτικά μοντέλα, ανά δίκτυο. Κλείνοντας, θα διεξαχθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, τα αποτελέσματα των οποίων, θα απεικονισθούν σε διαγράμματα.

Abstract

In a world dominated by uncertainty, where every aspect is in constant flux, the human imperative to formulate accurate forecasts takes precedence in risk reduction. The advent of new technologies, coupled with the surge in computing power, has revolutionized the creation of more precise predictions. This thesis delves into the theoretical groundwork surrounding the analysis and forecasting of time series using classical methods. It explores synthetic elements, model types, various statistical measures, the identification of stagnation, and diverse forecasting methods tailored to specific problems. The second chapter addresses machine learning, its types and distinctions from deep learning. Simultaneously, it discusses the biological neuron, theoretical frameworks and the operational modes of four architectural artificial neural networks, LSTM, GRU, RNN and MLP, in depth. In addition, this section will present the most well-known optimization functions and loss functions. In the third chapter, the thesis will provide general features about stock returns and ways to calculate them. Subsequently, the logarithmic returns of three companies in the Greek stock market, specifically, the group of GEK TERNA S.A., AEGEAN AIRLINES S.A., and OPAP S.A. will be studied. This analysis will involve the use of four architectural neural networks (LSTM, GRU, RNN and MLP) along with six different optimization algorithms. The efficiency of these models will be compared based on specific informative error criteria, leading to the identification of the most effective models for each network. In conclusion, the conversion of predicted logarithmic returns to predicted closing values will be executed and the results will be presented through charts.

Πίνακας περιεχομένων

Περίληψη	xxi
Abstract	xxiii
Κατάλογος Πινάκων	xxx
Κατάλογος Διαγραμμάτων	xxxiii
Κατάλογος Συντομογραφιών	xxxvii
ΚΕΦΑΛΑΙΟ 1: Θεωρία Χρονοσειρών	1
1.1 Εισαγωγή	1
1.2 Ορισμός των χρονοσειρών	2
1.3 Είδη χρονοσειρών ανάλογα την ανεξάρτητη μεταβλητή	3
1.4 Διαχωρισμός δεδομένων ανάλογα την φύση τους	3
1.4.1 Δομημένα και μη δομημένα δεδομένα	3
1.4.2 Συνεχή και διακριτά δεδομένα	3
1.5 Διάσπαση χρονοσειρών	4
1.5.1 Τάση.....	5
1.5.2 Εποχικότητα	6
1.5.3 Κυκλικότητα	7
1.5.4 Μη κανονικότητα	7
1.6 Ανάλυση χρονοσειρών	9
1.6.1 Προσθετικό υπόδειγμα.....	9
1.6.2 Πολλαπλασιαστικό υπόδειγμα.....	9
1.7 Στατιστικά μέτρα	10
1.7.1 Αναμενόμενη τιμή	10
1.7.2 Διακύμανση.....	11

1.7.3 Αυτοσυνδιακύμανση	11
1.7.4 Αυτοσυσχέτιση	11
1.8 Στασιμότητα.....	12
1.8.1 Στατιστικός έλεγχος Dickey - Fuller.....	14
1.8.2 Στατιστικός έλεγχος KPSS.....	15
1.8.3 Λευκός θόρυβος.....	16
1.8.4 Τυχαίος περίπατος.....	17
1.9 Μέθοδοι προβλέψεων.....	19
1.9.1 Μέθοδος Naive.....	19
1.9.2 Απλός κινητός μέσος	20
1.9.3 Διπλός κινητός μέσος	21
1.9.4 Σταθμισμένος κινητός μέσος	22
1.9.5 Απλή εκθετική εξομάλυνση	23
1.9.6 Διπλή εκθετική εξομάλυνση.....	24
1.9.7 Εκθετική εξομάλυνση με προσαρμογή στην τάση	25
1.9.8 Εκθετική εξομάλυνση με προσαρμογή στην τάση και την εποχικότητα	27
1.9.9 Γραμμική παλινδρόμηση	28
1.9.10 Μοντέλα AR, MA, ARMA.....	31
1.9.11 Μετασχηματισμός μη στάσιμης χρονοσειράς σε στάσιμη	37
1.9.12 Ολοκληρωμένο υπόδειγμα ARIMA.....	38
1.9.13 Μέθοδος Box & Jenkins	40
1.9.14 Κριτήρια επιλογής υποδείγματος.....	42
1.10 Υποδείγματα SARIMA	44
1.11 Σφάλμα πρόβλεψης.....	45
1.11.1 Μέσο τετραγωνικό σφάλμα.....	46

1.11.2 Τετραγωνική ρίζα μέσου τετραγωνικού σφάλματος.....	46
1.11.3 Μέση Απόκλιση.....	47
1.11.4 Μέσο απόλυτο ποσοστιαίο σφάλμα	47
1.12 Επιλογή κατάλληλου μοντέλου πρόβλεψης	47
1.13 Ανακεφαλαίωση	48
ΚΕΦΑΛΑΙΟ 2: Μηχανική Μάθηση – Νευρωνικά Δίκτυα	49
2.1 Εισαγωγή	49
2.2 Μηχανική Μάθηση	50
2.2.1 Κατηγορίες μηχανικής μάθησης.....	50
2.2.2 Διαφορές μεταξύ βαθιάς μάθησης και μηχανικής μάθησης	52
2.3 Νευρωνικά Δίκτυα	53
2.3.1 Εισαγωγή στα νευρωνικά δίκτυα	53
2.3.2 Πολυεπίπεδο Perceptron.....	57
2.3.3 Παλινδρομικά Νευρωνικά Δίκτυα	60
2.3.4 Διαφορές μεταξύ ενός MLP και ενός RNN νευρωνικού δικτύου	62
2.3.5 Δίκτυα Μακράς Βραχύχρονης Μνήμης.....	62
2.3.6 Φραγμένα ανατροφοδοτούμενα νευρωνικά δίκτυα.....	66
2.4 Συναρτήσεις Νευρωνικών Δικτύων.....	68
2.4.1 Συνάρτηση απώλειας.....	68
2.4.2 Συναρτήσεις ενεργοποίησης.....	69
2.4.3 Εκπαίδευση τεχνητών νευρωνικών δικτύων.....	71
2.5 Ανακεφαλαίωση	73
ΚΕΦΑΛΑΙΟ 3: Έρευνα και αποτελέσματα	75
3.1 Εισαγωγή	75
3.2 Η απόδοση σαν όρος στο χρηματιστήριο	75

3.3 Γενικά γνωρίσματα της μελέτης της διπλωματικής εργασίας	79
3.3.1 Εργαλεία που χρησιμοποιήθηκαν	79
3.3.2 Στόχος της μελέτης	79
3.3.3 Δομή της μελέτης	79
3.4 Διερευνητική Ανάλυση	80
3.4.1 Επιλογή συνόλων δεδομένων	80
3.4.2 Προεπεξεργασία δεδομένων	81
3.4.3 Οπτικοποίηση των τριών συνόλων δεδομένων	82
3.4.4 Επαυξημένος έλεγχος των Dickey - Fuller	84
3.4.5 Μετατροπή των τιμών κλεισίματος των μετοχών σε λογαριθμικές αποδόσεις	85
3.5 Μοντελοποίηση Νευρωνικών Δικτύων	86
3.5.1 Keras Configuration	86
3.5.2 Αλγόριθμοι που χρησιμοποιήθηκαν.....	88
3.5.3 Αλγόριθμοι βελτιστοποίησης.....	90
3.5.4 Πληροφορικά κριτήρια σφάλματος	91
3.6 Μελέτη των τριών εταιρειών	92
3.6.1 Αρχική φάση της μελέτης	92
3.6.2 Μελέτη και παρουσίαση αποτελεσμάτων για την ΓΕΚ ΤΕΡΝΑ Α.Ε.....	92
3.6.3 Μελέτη και παρουσίαση αποτελεσμάτων για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.....	107
3.6.4 Μελέτη και παρουσίαση αποτελεσμάτων για τον ΟΠΑΠ Α.Ε.	115
3.7 Συμπεράσματα της μελέτης και μελλοντικοί στόχοι	123
ΒΙΒΛΙΟΓΡΑΦΙΑ	126

Κατάλογος Πινάκων

	Σελίδα
Πίνακας 1.1: Ιδιότητες ACF και PACF στα μοντέλα AR(p) και MA(q)	36
Πίνακας 3.1: Αποτελέσματα του επαυξημένου ελέγχου Dickey Fuller για τις τρεις εταιρείες	85
Πίνακας 3.2: Αποτελέσματα του επαυξημένου ελέγχου Dickey Fuller στις αποδόσεις των τριών εταιρειών	86
Πίνακας 3.3: Αποτελέσματα των κριτηρίων σφάλματος του LSTM δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.	93
Πίνακας 3.4: Αποτελέσματα των κριτηρίων σφάλματος του GRU δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.	95
Πίνακας 3.5: Αποτελέσματα των κριτηρίων σφάλματος του RNN δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.	97
Πίνακας 3.6: Αποτελέσματα των κριτηρίων σφάλματος του MLP δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.	99
Πίνακας 3.7: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο LSTM δίκτυο για την ΓΕΚ ΤΕΡΝΑ Α.Ε.	101
Πίνακας 3.8: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο GRU δίκτυο για την ΓΕΚ ΤΕΡΝΑ Α.Ε.	103
Πίνακας 3.9: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο RNN δίκτυο για την ΓΕΚ ΤΕΡΝΑ Α.Ε.	104
Πίνακας 3.10: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο MLP δίκτυο για την ΓΕΚ ΤΕΡΝΑ Α.Ε.	106
Πίνακας 3.11: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο LSTM δίκτυο για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	108
Πίνακας 3.12: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο GRU δίκτυο για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	110
Πίνακας 3.13: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο RNN δίκτυο για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	112
Πίνακας 3.14: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο MLP δίκτυο για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	114

Πίνακας 3.15: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο LSTM δίκτυο για τον ΟΠΑΠ Α.Ε.	116
Πίνακας 3.16: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο GRU δίκτυο για τον ΟΠΑΠ Α.Ε.	118
Πίνακας 3.17: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο RNN δίκτυο για τον ΟΠΑΠ Α.Ε.	120
Πίνακας 3.18: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο MLP δίκτυο για τον ΟΠΑΠ Α.Ε.	122
Πίνακας 3.19: Σύνοψη των καλύτερων αποτελεσμάτων ανά δίκτυο για κάθε εταιρεία	123
Πίνακας 3.20: Σύνοψη των χειρότερων αποτελεσμάτων ανά δίκτυο για κάθε εταιρεία	124

Κατάλογος Διαγραμμάτων

	Σελίδα
Διάγραμμα 1.1: Μεταφερόμενοι επιβάτες ανά μήνα των διεθνών αερογραμμών κατά την περίοδο 1949 έως 1960	4
Διάγραμμα 1.2: Μεταβολή στην παραγωγή τυριού ανά τα χρόνια	6
Διάγραμμα 1.3: Λιανικές πωλήσεις μεταχειρισμένων αυτοκινήτων στις ΗΠΑ από τον Γενάρη του 1992 έως τον Δεκέμβριο του 2019	7
Διάγραμμα 1.4: Ζήτηση ταξί στην Νέα Υόρκη για το διάστημα Αύγουστος 2014 - Φεβρουάριος 2015	8
Διάγραμμα 1.5: Διάγραμμα λευκού θορύβου τυχαίου δείγματος από Γκαουζιανή κατανομή	16
Διάγραμμα 1.6: Διάγραμμα αυτοσυσχετίσεων τυχαίου δείγματος από Γκαουζιανή κατανομή	17
Διάγραμμα 1.7: Διάγραμμα τυχαίου περιπάτου τυχαίου δείγματος από κανονική κατανομή	18
Διάγραμμα 1.8: Διάγραμμα αυτοσυσχετίσεων τυχαίου περιπάτου τυχαίου δείγματος από κανονική κατανομή	18
Διάγραμμα 1.9: Διάγραμμα διασποράς	30
Διάγραμμα 1.10: Διάγραμμα από τυχαία παραγμένες χρονοσειρές με $\varphi=0.9$ και $\varphi=0.1$	32
Διάγραμμα 1.11: Κορελογράμματα των τυχαίων χρονοσειρών με $\varphi=0.9$ και $\varphi=0.1$	33
Διάγραμμα 1.12: Διάγραμμα αυτοσυνδιακυμάνσεων για $\varphi=0.6$ και $\varphi=-0.7$	33
Διάγραμμα 1.13: Γραφική απεικόνιση των αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων του μοντέλου MA(1) με $\theta_1=0.7$	35
Διάγραμμα 2.1: Γραφική απεικόνιση των διαφορών μεταξύ μηχανικής μάθησης και βαθιάς μάθησης	53
Διάγραμμα 2.2: Φυσικός νευρώνας	54
Διάγραμμα 2.3: Βασικό μοντέλο τεχνητού νευρώνα	55
Διάγραμμα 2.4: Απεικόνιση βηματικής συνάρτησης	56
Διάγραμμα 2.5: Σιγμοειδής συνάρτηση	57

Διάγραμμα 2.6: Τρόπος μετάδοσης της πληροφορίας σε ένα MLP	57
Διάγραμμα 2.7: Απεικόνιση ενός MLP που περιέχει μια στιβάδα εισόδου μια εξόδου και μια κρυφή στιβάδα	58
Διάγραμμα 2.8: Τοπολογία παλινδρομικού τεχνητού νευρωνικού δικτύου	61
Διάγραμμα 2.9: Απεικόνιση της δομής της εσωτερικής μονάδας ενός δικτύου LSTM	63
Διάγραμμα 2.10: Εσωτερική δομή ενός δικτύου GRU	68
Διάγραμμα 3.1: Τιμή μετοχής της ΓΕΚ ΤΕΡΝΑ Α.Ε. από 12/04/2017 ως 03/11/2023	82
Διάγραμμα 3.2: Τιμή μετοχής της ΑΕΡΟΠΟΡΙΑΣ ΑΙΓΑΙΟΥ Α.Ε από 12/04/2017 ως 03/11/2023	82
Διάγραμμα 3.3: Τιμή μετοχής της ΟΠΑΠ Α.Ε. από 12/04/2017 έως 3/11/2023	83
Διάγραμμα 3.4: Διαγράμματα σύγκρισης μεταξύ πραγματικών και προβλεπόμενων τιμών κλεισίματος του δικτύου LSTM, για τις έξι συναρτήσεις βελτιστοποίησης	94
Διάγραμμα 3.5: Διαγράμματα σύγκρισης μεταξύ πραγματικών και προβλεπόμενων τιμών κλεισίματος του δικτύου GRU, για τις έξι συναρτήσεις βελτιστοποίησης	96
Διάγραμμα 3.6: Διαγράμματα σύγκρισης μεταξύ πραγματικών και προβλεπόμενων τιμών κλεισίματος του δικτύου RNN, για τις έξι συναρτήσεις βελτιστοποίησης	98
Διάγραμμα 3.7: Διαγράμματα σύγκρισης μεταξύ πραγματικών και προβλεπόμενων τιμών κλεισίματος του δικτύου MLP, για τις έξι συναρτήσεις βελτιστοποίησης	100
Διάγραμμα 3.8: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το LSTM δίκτυο της ΓΕΚ ΤΕΡΝΑ Α.Ε.	102
Διάγραμμα 3.9: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το GRU δίκτυο της ΓΕΚ ΤΕΡΝΑ Α.Ε.	104
Διάγραμμα 3.10: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το RNN δίκτυο της ΓΕΚ ΤΕΡΝΑ Α.Ε.	105

Διάγραμμα 3.11: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το MLP δίκτυο της ΓΕΚ ΤΕΡΝΑ Α.Ε.	107
Διάγραμμα 3.12: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το LSTM δίκτυο της ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	109
Διάγραμμα 3.13: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το GRU δίκτυο της ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	111
Διάγραμμα 3.14: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το MLP δίκτυο της ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	113
Διάγραμμα 3.15: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το MLP δίκτυο της ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	115
Διάγραμμα 3.16: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το LSTM δίκτυο του ΟΠΑΠ Α.Ε.	117
Διάγραμμα 3.17: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το GRU δίκτυο του ΟΠΑΠ Α.Ε.	119
Διάγραμμα 3.18: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το RNN δίκτυο του ΟΠΑΠ Α.Ε.	121
Διάγραμμα 3.19: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το MLP δίκτυο του ΟΠΑΠ Α.Ε.	123

Κατάλογος Συντομογραφιών

Adam	Adaptive Moment Estimation
AIC	Akaike's Information Criterion
AR	Autoregressive Process
ARMA	Autoregressive Moving Average Process
BIC	Bayesian Information Criterion
CNN	Convolutional Neural Network
i.i.d	Independently Identically Distributed
LSTM	Long Short-Term Memory
MA	Moving Average Process
MLP	Multilayer Perceptron
Nadam	Nesterov-Accelerated Adaptive Moment Estimation
NLP	Natural Language Processing
NN	Neural Network
RNN	Recurrent Neural Network
SGD	Stochastic Gradient Descent
TN	Τεχνητή Νοημοσύνη
TNΔ	Τεχνητό Νευρωνικό Δίκτυο

ΚΕΦΑΛΑΙΟ 1: Θεωρία Χρονοσειρών

1.1 Εισαγωγή

Τα τελευταία χρόνια οι εξελίξεις σε παγκόσμιο επίπεδο είναι ραγδαίες δημιουργώντας ένα ασταθές οικονομικό περιβάλλον. Η πανδημία του COVID-19 ήρθε για πρώτη φορά τον Φεβρουάριο του 2020, όπου προκάλεσε μεγάλες αναταραχές στην οικονομία σε παγκόσμιο επίπεδο, αφού ήταν η δεύτερη μεγαλύτερη ύφεση που έχει παρατηρηθεί στην ιστορία. Σχεδόν όλοι οι κλάδοι υπέστησαν μεγάλο πλήγμα, αφού η επιχειρηματική δραστηριότητα ήταν αισθητά μειωμένη σε σχέση με την προ COVID εποχή. Παρ' όλα αυτά πολλές εταιρείες εκμεταλλεύτηκαν αυτή την κατάσταση και κατάφεραν με την δημιουργία ορθής στρατηγικής να αναπτυχθούν και να κερδοφορήσουν (Λιαγκόρβας & λοιποί, 2022). Παράλληλα ένα ακόμα εξίσου σημαντικό γεγονός αποτελεί η ενεργειακή κρίση, η οποία έπληξε και συνεχίζει να πλήττει την παγκόσμια οικονομία σε συνδυασμό με τον πόλεμο στην Ουκρανία, κάνοντας τις τιμές του πετρελαίου και του φυσικού αερίου συνεχώς να αυξάνονται, αφού σύμφωνα με τα δεδομένα του Διεθνούς Οργανισμού Ενέργειας, η Ρωσία είναι ο μεγαλύτερος εξαγωγέας φυσικού αερίου στον κόσμο και ο δεύτερος μεγαλύτερος αργού πετρελαίου μετά την Σαουδική Αραβία (Zhou et al., 2022). Επιπροσθέτως, ραγδαία είναι η εξέλιξη της τεχνολογίας, αφού ολόκληρη η παγκόσμια επιστημονική κοινότητα ασχολείται με την εύρεση τεχνολογιών που θα οδηγήσουν στην δημιουργία πιο ασφαλών προβλέψεων, οι οποίες θα μειώσουν το ρίσκο για αστοχίες και την δημιουργία απώλειας.

Η μελέτη των χρονοσειρών αποτελεί χρήσιμο εργαλείο πολλών επιστημονικών κλάδων, όπως της ιατρικής, των οικονομικών, του αθλητισμού και της πληροφορικής, αφού η φύση των δεδομένων αυτών είναι τέτοια, που αντικατοπτρίζουν την πρόοδο του αντικειμένου, το οποίο μελετάται σε πραγματικό χρόνο. Για παράδειγμα στον αθλητισμό, πάρα πολλές μεγάλες ομάδες ποδοσφαίρου χρησιμοποιούν μοντέλα χρονοσειρών με σκοπό να προβλέψουν την καριέρα των αθλητών στους οποίους θέλουν να επενδύσουν.

Τα τελευταία χρόνια έχουν εισέλθει λόγω της ραγδαίας τεχνολογικής εξέλιξης τα αυτοκινούμενα αυτοκίνητα, τα οποία χρησιμοποιούν δεδομένα χρονοσειρών για την βελτίωση της απόδοσης του κινητήρα και της μπαταρίας, αφού το σύστημα, το οποίο

διαθέτουν τα οχήματα αυτά προσαρμόζεται με την πάροδο του χρόνου στην γήρανση του οχήματος. Στον κλάδο των οικονομικών, λόγω της μεγάλης αβεβαιότητας, την οποία διανύει η παγκόσμια κοινότητα, η χρήση μεθόδων ανάλυσης και πρόβλεψης δεδομένων από χρονοσειρές είναι αναγκαία για την αποφυγή σφαλμάτων και την δημιουργία λανθασμένης στρατηγικής που θα οδηγήσει σε ζημιά. Η μελέτη μετοχών, η μελέτη πωλήσεων και αγοράς προϊόντων και πρώτων υλών αφορούν μερικές περιπτώσεις εφαρμογής των χρονοσειρών στον οικονομικό κλάδο.

Σε αυτό το κεφάλαιο θα γίνει προσπάθεια μελέτης και ανάλυσης των χρονοσειρών σε θεωρητικό επίπεδο. Η ανάλυση αυτή αποτελεί μέρος της ποσοτικής ανάλυσης, με την διαφορά, ότι κάθε χρονοσειρά χρήζει ιδιαίτερης μεταχείρισης, λόγω του ότι ο χρόνος αποτελεί κρίσιμη παράμετρο, αφού δείχνει πως τα δεδομένα προσαρμόζονται ανά χρονική στιγμή. Αρχικά, θα δοθεί ο ορισμός, η εξέλιξη και τα είδη των χρονοσειρών, ενώ ακολούθως ο ερευνητικός φακός θα στραφεί σε μεθόδους διάσπασης και ανάλυσης χρονοσειρών, στασιμότητας και πρόβλεψης των τιμών αλλά και των σφαλμάτων.

1.2 Ορισμός των χρονοσειρών

Χρονοσειρά ορίζεται ως μια ακολουθία παρατηρήσεων x_t , όπου η κάθε μια παρατήρηση καταγράφεται σε μια συγκεκριμένη χρονική στιγμή t και το χαρακτηριστικό τους γνώρισμα είναι, ότι οι παρατηρήσεις αυτές ισαπέχουν χρονικά μεταξύ τους. Παραδείγματος χάριν οι μέρες τις εβδομάδας. Ο σκοπός της μελέτης των χρονοσειρών αποσκοπεί στον έλεγχο της εξέλιξης των παρατηρήσεων συναρτήσει του χρόνου.

Η πρώτη εμφάνιση της μελέτης των χρονοσειρών εμφανίζεται στα έργα των G. U. Yule και J. Walker στις δεκαετίες του 1920 και 1930, αφού ήταν οι πρώτοι που εφάρμοσαν αυτοπαλίνδρομα μοντέλα (Makridakis, 1970). Παράλληλα εκείνη την περίοδο εισήχθη για πρώτη φορά η έννοια του κινητού μέσου όρου. Ο Herman Wold εισήγαγε τα μοντέλα ARMA (AutoRegressive Moving Average) για στάσιμες σειρές, αλλά δεν μπόρεσε να εξάγει μια συνάρτηση πιθανοφάνειας που να επιτρέπει την εκτίμηση των παραμέτρων με μέγιστη πιθανοφάνεια. Αυτό επιτευχθεί το 1970, όπου βγήκε στην κυκλοφορία το βιβλίο με τίτλο «Time Series Analysis» των G. E. P. Box και G. M. Jenkins (1970), το οποίο περιείχε την πλήρη διαδικασία μοντελοποίησης για μεμονωμένες σειρές.

1.3 Είδη χρονοσειρών ανάλογα την ανεξάρτητη μεταβλητή

Ανάλογα του πλήθους των μεταβλητών που περιέχονται στις χρονοσειρές, αυτές διακρίνονται σε δύο κατηγορίες. Στην πρώτη κατηγορία περιέχονται οι μονομεταβλητές και στη δεύτερη οι πολυμεταβλητές χρονοσειρές. Αρχικά οι χρονοσειρές ορίζονται ως μονομεταβλητές (univariate), όταν η μεταβλητή, η οποία εξετάζεται είναι μονός αριθμός, όπως η τιμή κλεισίματος μιας μετοχής ανά ημέρα. Αντίστοιχα πολυμεταβλητές (multivariate) ορίζονται οι χρονοσειρές, όπου ο αριθμός των μεταβλητών που μπορούν να μελετηθούν έχει μεγαλύτερο πλήθος του ενός. Για παράδειγμα η ταυτόχρονη μελέτη της τιμής θερμοκρασίας, η ταχύτητα του ανέμου συναρτήσει του ποσοστού βροχόπτωσης (Bergman, 1979).

Η μεταβλητή εισόδου είναι ενδογενής, εφόσον αυτή επηρεάζεται από άλλες μεταβλητές και αντίστοιχα επηρεάζει με την σειρά της και την μεταβλητή απόκρισης. Όταν σε μια χρονοσειρά, η κάθε παρατήρηση εξαρτάται από την προηγούμενη της, δηλαδή η παρατήρηση την χρονική στιγμή t εξαρτάται από την παρατήρηση την χρονική στιγμή $t-1$, όπως επίσης και η παρατήρηση την χρονική στιγμή $t-1$ θα εξαρτάται από την $t-2$ χρονική στιγμή, τότε η μεταβλητή εισόδου είναι ενδογενής. Αντίθετα οι εξωγενείς μεταβλητές δεν επηρεάζονται από άλλες μεταβλητές. Συνήθως οι μεταβλητές αυτό δεν λαμβάνονται υπόψη, διότι στη μελέτη χρονοσειρών περιέχονται κατά κύριο λόγο ενδογενείς μεταβλητές.

1.4 Διαχωρισμός δεδομένων ανάλογα την φύση τους

1.4.1 Δομημένα και μη δομημένα δεδομένα

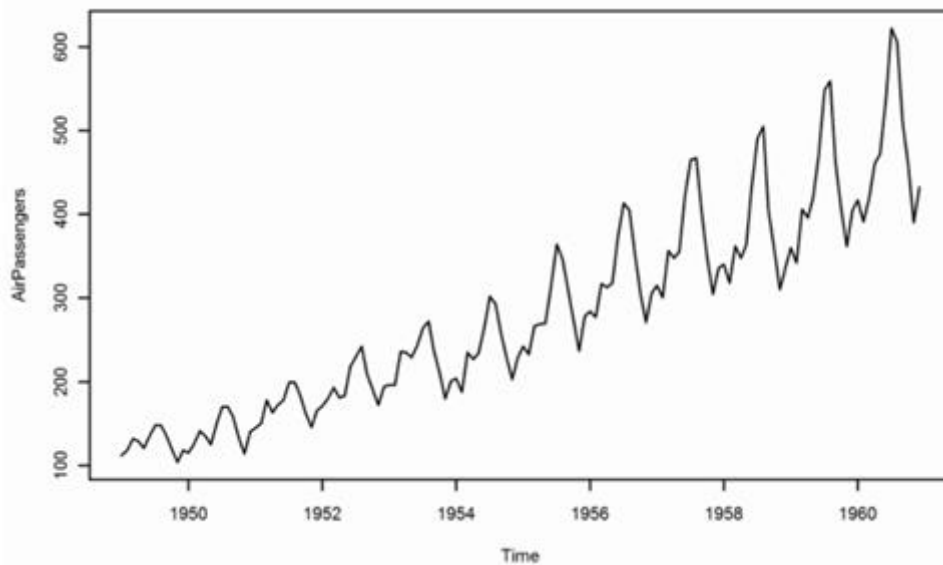
Κατά την μελέτη των χρονοσειρών, είναι πολύ σημαντικό, να ελεγχθεί κυρίως με την βοήθεια ενός γραφήματος, αν διακρίνεται στα δεδομένα κάποιο συγκεκριμένο μοτίβο (structured data), όπως για παράδειγμα η τάση, η κυκλικότητα και η εποχικότητα. Στην περίπτωση που δεν υπάρχει κάποιο προφανές μοτίβο, τότε τα δεδομένα δεν ακολουθούν καμία δομή που να εξαρτάται από τον χρόνο (unstructured data).

1.4.2 Συνεχή και διακριτά δεδομένα

Μια χρονοσειρά ορίζεται ως διακριτή, όταν μεταξύ των παρατηρήσεων, η χρονική διάρκεια μεταξύ των παρατηρήσεων υπερβαίνει το ένα δευτερόλεπτο, όπως τιμή μιας μετοχής

κλεισίματος, η οποία υπολογίζεται ανά ημέρα. Από την άλλη πλευρά μια συνεχής χρονοσειρά περιέχει παρατηρήσεις ανά δευτερόλεπτο, όπως για παράδειγμα η πίεση του αίματος.

Για την καλύτερη κατανόηση του μοτίβου μιας χρονοσειράς πολύ σημαντικό ρόλο παίζει η απεικόνιση της σε διάγραμμα. Στο Διάγραμμα 1.1 παρουσιάζονται οι μεταφερόμενοι επιβάτες ανά μήνα των διεθνών αερογραμμών κατά την περίοδο 1949 έως 1960. Οι αριθμοί στον άξονα των επιβατών είναι σε χιλιάδες.



Διάγραμμα 1.1: Μεταφερόμενοι επιβάτες ανά μήνα των διεθνών αερογραμμών κατά την περίοδο 1949 έως 1960 (Researchgate, 2012)

1.5 Διάσπαση χρονοσειρών

Η διάσπαση χρονοσειρών (time series decomposition) είναι μια μέθοδος προσδιορισμού των τιμών της χρονοσειράς, κάνοντας χρήση των συνθετικών στοιχείων της. Τα στοιχεία αυτά είναι τα ακόλουθα:

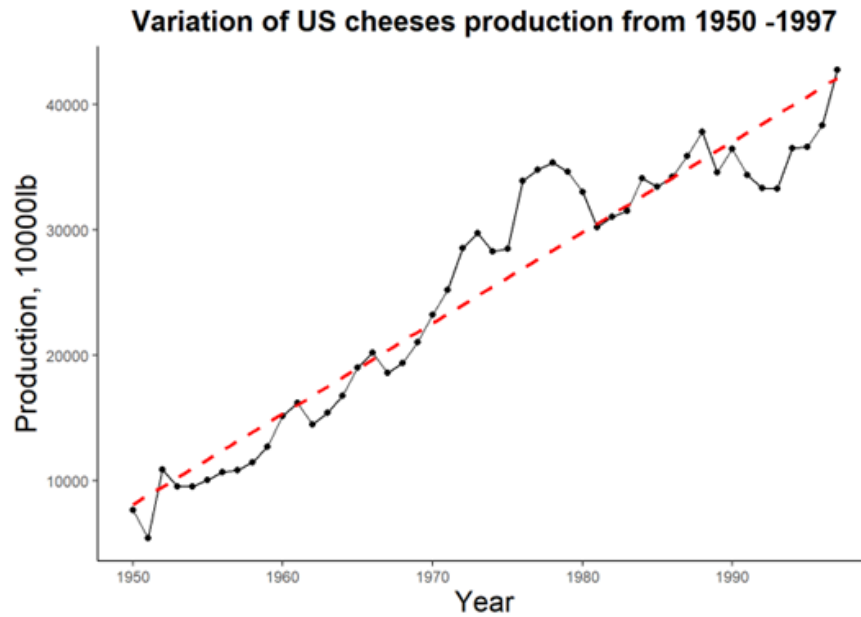
- Τάση (Trend)
- Εποχικότητα (Seasonality)
- Κυκλικότητα (Cyclic)
- Μη-Κανονικότητα (Outliers)

Ο σκοπός αυτής της μεθόδου είναι η δημιουργία προβλέψεων κάνοντας διάσπαση των συνθετικών στοιχείων ξεχωριστά, μελετώντας τον τρόπο που επηρεάζουν τις τιμές της χρονοσειράς. Στη διάσπαση χρονοσειρών καθοριστικό βήμα είναι να γίνει η αναγνώριση των συνθετικών στοιχείων της. Όσο καλύτερα γίνει η εξακρίβωση, τόσο καλύτερη κατανόηση θα υπάρχει στον τρόπο δημιουργίας των παρατηρήσεων (Αγιακλόγλου & Οικονόμου, 2004). Παρακάτω παρουσιάζονται αναλυτικά τα συνθετικά στοιχεία μιας χρονοσειράς:

1.5.1 Τάση

Η τάση είναι το συνθετικό στοιχείο, το οποίο δείχνει την μετακίνηση των τιμών της χρονοσειράς, είτε με αύξουσα κλίση, είτε με φθίνουσα για μεγάλο χρονικό διάστημα. Ουσιαστικά η τάση εκφράζει την μακροχρόνια εξέλιξη των τιμών μιας χρονοσειράς. Για να είναι ασφαλής η εκτίμηση και η εξαγωγή συμπερασμάτων, ο αριθμός των παρατηρήσεων πρέπει να είναι επαρκής. Το φαινόμενο της τάσης παρατηρείται σε γεγονότα που αφορούν τεχνολογικές αλλαγές, όπως επίσης και σε γεγονότα που επηρεάζονται από οικονομικούς παράγοντες.

Στο Διάγραμμα 1.2 παρουσιάζεται η μεταβολή της παραγωγής του μπλε τυριού (blue cheese) και της Gorgonzola στις ΗΠΑ για την περίοδο 1950 έως 1997. Η κόκκινη γραμμή του γραφήματος είναι η ευθεία παλινδρόμησης. Αποτελεί χαρακτηριστικό παράδειγμα χρονοσειράς που παρουσιάζει αυξητική τάση.



Διάγραμμα 1.2: Μεταβολή στην παραγωγή τυριού ανά τα χρόνια (Mominul, 2020)

1.5.2 Εποχικότητα

Η εποχικότητα ορίζεται ως το φαινόμενο που εμφανίζεται σε χρονοσειρές για χρονικές περιόδους μικρότερες του έτους, π.χ. μήνες, τρίμηνα, όπου παρατηρούνται εποχικές διακυμάνσεις με κύριο χαρακτηριστικό τους να εμφανίζονται με παρόμοια ή και ίδια μορφή από έτος σε έτος. Αυτές οι εποχικές διακυμάνσεις είναι σχετικά εύκολο να παρατηρηθούν, να μελετηθούν και να συμβάλλουν και αυτές με την σειρά τους, όπως και η τάση στη δημιουργία προβλέψεων. Παραδείγματος χάριν η μηνιαία κατανάλωση πέδιλων θάλασσας είναι αυξημένη τους καλοκαιρινούς μήνες σε αντίθεση με τους χειμερινούς. Από την άλλη τα πέδιλα του σκι η μηνιαία κατανάλωση τους είναι αυξημένη τους χειμερινούς μήνες σε σχέση με τους καλοκαιρινούς.

Το Διάγραμμα 1.3 παρουσιάζει τις λιανικές πωλήσεις μεταχειρισμένων αυτοκινήτων στις ΗΠΑ από το Γενάρη του 1992, έως το Δεκέμβριο του 2019. Πιο συγκεκριμένα, παρατηρούνται έντονα στοιχεία εποχικότητας στην παραπάνω χρονοσειρά.



Διάγραμμα 1.3: Λιανικές πωλήσεις μεταχειρισμένων αυτοκινήτων στις ΗΠΑ από τον Γενάρη του 1992 έως τον Δεκέμβριο του 2019 (Census, 2024).

1.5.3 Κυκλικότητα

Οι κυκλικές διακυμάνσεις εμφανίζονται σε κυματοειδή μορφή και οφείλονται σε «σκαμπανεβάσματα» των τιμών σε μη σταθερά διαστήματα, καθώς επίσης και σε διαφορετική συχνότητα. Αυτή είναι η κύρια διαφορά της με την εποχικότητα, όπου η συχνότητα είναι αμετάβλητη. Κυκλικότητα υπάρχει, όταν το φαινόμενο διαρκεί για διάστημα πολύ μεγαλύτερο του ενός έτους (π.χ. δύο έτη ή και περισσότερο). Σημαντικό δε στοιχείο αποτελεί το γεγονός ότι η κυκλικότητα δεν χρησιμοποιείται τόσο για πρόβλεψη, όσο για να γίνει μια αποτίμηση της μέχρι τώρα πορείας της χρονοσειράς.

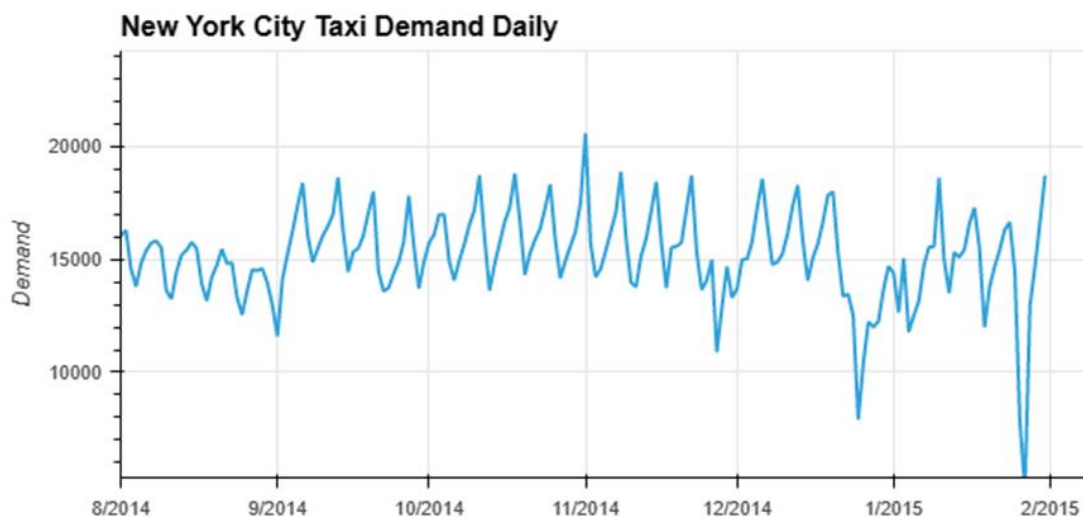
Η κυκλικότητα βρίσκεται κυρίως σε οικονομικές χρονοσειρές, όπως για παράδειγμα μετοχές σε εστιατόρια, αεροπορικές εταιρείες, κατασκευαστικές εταιρείες και αυτό συμβαίνει, διότι όταν υπάρχει άνθιση στην οικονομία, υπάρχει μεγαλύτερη ζήτηση γι' αυτές τις υπηρεσίες, με αποτέλεσμα οι μετοχές να αυξάνονται, σε αντίθεση με περιόδους ύφεσης της οικονομίας, όπου η ζήτηση μειώνεται, με αποτέλεσμα οι μετοχές να εμφανίζουν πτώση.

1.5.4 Μη κανονικότητα

Το τέταρτο και τελευταίο συνθετικό στοιχείο των χρονοσειρών είναι η μη κανονικότητα, η οποία εκφράζει ένα τυχαίο γεγονός που πραγματοποιείται με μη συστηματικό τρόπο και προφανώς δεν μπορεί να προβλεφθεί. Τα τυχαία γεγονότα αυτά ονομάζονται και outliers.

Η χρονική τους διάρκεια δεν είναι μεγάλη, όμως χρήζει ιδιαίτερης προσοχής στην διαχείρισή τους, καθώς κάθε περίπτωση είναι διαφορετική. Επιδρούν σημαντικά στη μελέτη της χρονοσειράς, γι' αυτό το λόγο χρειάζεται μια πολύπλευρη γνώση γύρω από το θέμα που μελετάται. Τέτοια φαινόμενα παρατηρούνται σε πολέμους, δυσμενείς καιρικές συνθήκες, όπως και μεταβολές στις προτιμήσεις των καταναλωτών. Μια πανδημία μπορεί να προκαλέσει τεράστια ζημία σε μια αεροπορική εταιρεία και γενικά σε όλη την οικονομία, όπως αυτή του Covid 19.

Στο Διάγραμμα 1.4 παρουσιάζεται η καθημερινή ζήτηση ταξί στη Νέα Υόρκη για το χρονικό διάστημα Αύγουστος 2014 - Φεβρουάριος 2015. Παρατηρείται ότι υπάρχει μια μεγάλη πτώση στη ζήτηση τον Φεβρουάριο του 2015, λόγω της μεγάλης άνθησης της εταιρείας μεταφορών «Uber», η οποία προσλάμβανε οδηγούς αυτοκινήτων για την μεταφορά επιβατών.



Διάγραμμα 1.4: Ζήτηση ταξί στην Νέα Υόρκη για το διάστημα Αύγουστος 2014 - Φεβρουάριος 2015 (Josh, 2023)

Για τα συνθετικά στοιχεία της χρονοσειράς, όπου έγινε εκτενής αναφορά, θα γίνει χρήση των παρακάτω συμβολισμών:

Y_t = πραγματική τιμή της χρονοσειράς

T_t = Τάση

$S_t =$ Εποχικότητα

$C_t =$ Κυκλικότητα

$I_t =$ Μη Κανονικότητα

Για $t = 1, 2, 3, \dots, n$.

1.6 Ανάλυση χρονοσειρών

Η Ανάλυση χρονοσειρών είναι μια συγκεκριμένη διαδικασία, η οποία αποσκοπεί στην κατανόηση της συμπεριφοράς της χρονοσειράς, καθώς και του μηχανισμού που διαμορφώνονται οι τιμές της, με απώτερο σκοπό την δημιουργία προβλέψεων. Γι' τον σκοπό αυτό δημιουργήθηκαν δύο υποδείγματα. Τα υποδείγματα αυτά είναι: το προσθετικό υπόδειγμα (additive model) και το πολλαπλασιαστικό υπόδειγμα (multiplicative model).

1.6.1 Προσθετικό υπόδειγμα

Η βασική υπόθεση του προσθετικού υποδείγματος στηρίζεται στο γεγονός, ότι οι πραγματικές τιμές της χρονοσειράς υπολογίζονται ως το άθροισμα των τεσσάρων συνθετικών στοιχείων της. Παράλληλα στηρίζεται επίσης, στην ύπαρξη ανεξαρτησίας μεταξύ των συνθετικών στοιχείων αυτών. Ένα σύνηθες παράδειγμα είναι ότι: η τάση δεν έχει επίδραση στην εποχικότητα για τον υπολογισμό των τιμών της χρονοσειράς. Είναι λογικό πως όλα τα συνθετικά στοιχεία έχουν την ίδια μονάδα μέτρησης σε σχέση με την τιμή της πραγματικής τιμής (Y_t). Πρακτικά αυτό το υπόδειγμα δεν χρησιμοποιείται συχνά, διότι προϋποθέτει σχετικά σταθερή εποχική διακύμανση, κάτι το οποίο δεν ισχύει σε μελέτη οικονομικών χρονοσειρών, αλλά βρίσκει εφαρμογή σε μελέτη φυσικών φαινομένων. Ο τύπος υπολογισμού του προσθετικού υποδείγματος είναι:

$$Y_t = T_t + S_t + C_t + I_t \quad (1.1)$$

1.6.2 Πολλαπλασιαστικό υπόδειγμα

Στο πολλαπλασιαστικό υπόδειγμα σε αντίθεση με το προσθετικό υπόδειγμα, τα συνθετικά στοιχεία της χρονοσειράς σχετίζονται (είναι εξαρτημένα), όπως επίσης και η πραγματική τιμή

της χρονοσειράς (Y_t) έχει ίδια μονάδα μέτρησης μόνο με την τάση (T_t), αφού τα υπόλοιπα συνθετικά στοιχεία λειτουργούν ως δείκτες.

Το πολλαπλασιαστικό υπόδειγμα εκφράζεται ως το γινόμενο των τεσσάρων συνθετικών στοιχείων του. Σε αυτό το υπόδειγμα η εποχικότητα μεταβάλλεται σύμφωνα με την τάση. Είναι φυσικό, πώς προτιμάται σε αντίθεση με το προσθετικό, αφού έχει μεγαλύτερη πληθώρα θεωρητικών και πρακτικών εφαρμογών. Ο τύπος υπολογισμού του πολλαπλασιαστικού υποδείματος είναι:

$$Y_t = T_t \cdot S_t \cdot C_t \cdot I_t \quad (1.2)$$

1.7 Στατιστικά μέτρα

Τα στατιστικά μέτρα είναι ιδιαίτερα χρήσιμα στην μελέτη των χρονοσειρών και ο ρόλος τους κρίσιμος, αφού αποτελούν θεμελιώδη εργαλεία που βοηθούν στην ανάλυση των δεδομένων, στην αναγνώριση των μοτίβων της χρονοσειράς όπως επίσης και στη λήψη αποφάσεων. Κάποια σημαντικά στατιστικά μέτρα που είναι ιδιαίτερα χρήσιμα στην μελέτη των χρονοσειρών είναι τα ακόλουθα:

- Αναμενόμενη τιμή
- Διακύμανση
- Αυτοσυνδιακύμανση
- Αυτοσυσχέτιση

1.7.1 Αναμενόμενη τιμή

Η αναμενόμενη τιμή (mean) ανήκει στα μέτρα κεντρικής τάσης και ορίζεται ως το άθροισμα των παρατηρήσεων και συγκεκριμένα των τιμών της χρονοσειράς, διαιρούμενο με το πλήθος των τιμών της. Παράλληλα συνδέεται άμεσα με την τάση, αφού είναι εξαρτώμενη από κάθε χρονική στιγμή t για κάθε παρατήρηση.

Αυτό σημαίνει, ότι αν μια χρονοσειρά παρουσιάζει ανοδική ή καθοδική τάση, αυτό θα αποτυπώνεται και στη μέση τιμή της. Παρακάτω παρουσιάζεται ο τρόπος υπολογισμού της μέσης τιμής:

$$\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t \quad (1.3)$$

$$\bar{Y} = E(Y_t) = \int_{-\infty}^{+\infty} y_t f_{Y_t}(y_t) dy_t \quad (1.4)$$

1.7.2 Διακύμανση

Η διακύμανση (variance) αποτελεί μέτρο διασποράς και εκφράζει την κατανομή των τιμών της χρονοσειράς γύρω από τον μέσο όρο της. υπολογίζεται από το άθροισμα του τετραγώνου της διαφοράς των τιμών της χρονοσειράς με τον μέσο όρο της προς το πλήθος μείον ένα.

$$\sigma^2 = \frac{1}{n-1} \sum_{t=1}^n (Y_t - \bar{Y})^2 \quad (1.5)$$

1.7.3 Αυτοσυνδιακύμανση

Οι ποσότητες γ_k ονομάζονται αυτοσυνδιακυμάνσεις (autocovariances) για $k = \dots, -1, 0, 1, \dots$

$\gamma_0 = \text{Var}(Y_t) = \sigma^2$, όπου σ^2 η διακύμανση της χρονοσειράς

$\gamma_{-k} = \gamma_k$ για την ύπαρξη στασιμότητας (θα γίνει ανάλυση παρακάτω) (Αγιακλόγλου, 2022)

Η ποσότητες c_k καλούνται τυχαίες αυτοσυνδιακυμάνσεις (autocovariances)

$$c_k = \frac{1}{n} \sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y}) \quad (1.6)$$

1.7.4 Αυτοσυσχέτιση

Στην ανάλυση παλινδρόμησης προκύπτουν διάφορα προβλήματα παραβίασης σε βασικές υποθέσεις. Μια εκ των βασικών είναι η υπόθεση ανεξαρτησίας των τιμών του τυχαίου σφάλματος του υποδείγματος, για την οποία ισχύει η σχέση:

$$E(\varepsilon_i \varepsilon_j) = 0, \text{ για κάθε } i, j \quad (1.7)$$

Όταν παραβιάζεται αυτή η συνθήκη εμφανίζεται το πρόβλημα της αυτοσυσχέτισης, δηλαδή εκφράζει την ύπαρξη σχέσης μεταξύ των τιμών του τυχαίου σφάλματος. Για το λόγο αυτό δημιουργήθηκε ο συντελεστής αυτοσυσχέτισης, όπου αποτελεί έναν από τους πιο σημαντικούς

δείκτες στην θεωρία χρονοσειρών και ο ρόλος του είναι να ελέγξει την ύπαρξη τυχαιότητας στην χρονοσειρά. Λαμβάνει τιμές στο διάστημα $[-1,1]$, το οποίο σημαίνει ότι όσο οι τιμές που λαμβάνει πλησιάζουν στις ακραίες τιμές, τόσο συσχετισμένες είναι οι παρατηρήσεις X_t και X_{t-k} της χρονοσειράς. Από την άλλη αν βρίσκεται στην τιμή 0, τότε οι τιμές είναι ασυσχέτιστες. Παρακάτω παρουσιάζεται ο συντελεστής αυτοσυσχέτισης ρ_k , για κάποια χρονική υστέρηση k , που ορίζεται ως το πηλίκο της αυτοσυνδιακύμανσης γ_k προς την διακύμανση σ^2 .

$$\rho_k = \text{Corr}(X_t, X_{t-k}) = \frac{\text{Cov}(X_t, X_{t-k})}{\sqrt{\text{VAR}(X_t)}\sqrt{\text{Var}(X_{t-k})}} = \frac{\gamma_k}{\sigma^2} = \frac{\gamma_k}{\gamma_0} \quad (1.8)$$

Επιπροσθέτως πολύ σημαντικό στοιχείο αποτελεί, πως η αυτοδιασπορά και η αυτοδιακύμανση έχει νόημα μόνο σε στάσιμες χρονοσειρές, διότι όταν δεν υπάρχει στασιμότητα δεν μπορούν να οριστούν αυτές ως συναρτήσεις υστέρησης. Αν γίνει η χρήση μιας μη στάσιμης χρονοσειράς και υπολογιστεί η συνάρτηση αυτοσυσχέτισης ως προς την υστέρηση παρατηρείται ότι έχει πολύ υψηλές τιμές και φθίνει στη συνέχεια πάρα πολύ αργά, λόγω ισχυρών συσχετίσεων.

Ένα πολύ σημαντικό στοιχείο είναι μελέτη της συνάρτησης μερικής αυτοσυσχέτισης (Partial Autocorrelation function). Η συνάρτηση μερικής αυτοσυσχέτισης ουσιαστικά είναι μια συνάρτηση αυτοσυσχέτισης υπό όρους. Δηλαδή εκτιμάει κατά πόσο η τιμή της χρονοσειράς για μια συγκεκριμένη χρονική περίοδο σχετίζεται με την τιμή k χρονικών υστερήσεων, δίχως να λαμβάνονται υπ' όψη οι χρονικές υστερήσεις που έχουν μεσολαβήσει. Για $k = 1$ ο συντελεστής μερικής αυτοσυσχέτισης ταυτίζεται με τον συντελεστή αυτοσυσχέτισης. Παρακάτω παρουσιάζεται ο τύπος υπολογισμού της μερικής αυτοσυσχέτισης για k χρονικές υστερήσεις.

$$\text{PACF}(T_t, k) = \frac{\text{COV}([T_t | T_{(t-1)}, T_{(t-2)}, \dots, T_{(t-k+1)}], [T_{(t-k)} | T_{(t-1)}, T_{(t-2)}, \dots, T_{(t-k+1)}])}{\sigma([T_t | T_{(t-1)}, T_{(t-2)}, \dots, T_{(t-k+1)}])\sigma([T_{(t-k)} | T_{(t-1)}, T_{(t-2)}, \dots, T_{(t-k+1)}])} \quad (1.9)$$

1.8 Στασιμότητα

Στάσιμη θεωρείται μια χρονοσειρά, αν για κάθε χρονική στιγμή η μέση τιμή και η διακύμανση της παραμένει σταθερή, όπως επίσης και οι αυτοσυνδιακυμάνσεις για δύο χρονικές στιγμές εξαρτώνται μόνο από τις χρονικές υστερήσεις και όχι από το χρονικό σημείο που υπολογίζονται. Αν μια από αυτές τις συνθήκες παραβιάζεται, τότε η χρονοσειρά θεωρείται μη

στάσιμη. Είναι απόλυτα φυσικό, αν μια χρονοσειρά δεν είναι στάσιμη εμφανίζει φαινόμενα τάσης και περιοδικότητας, τότε είναι δύσκολη η μελέτη της, άρα και η δημιουργία ορθών προβλέψεων. Βέβαια υπάρχουν τεχνικές που μπορούν να εφαρμοστούν, με σκοπό οι μη στάσιμες χρονοσειρές να μετατραπούν σε στάσιμες, οι οποίες θα αναλυθούν σε επόμενη ενότητα αυτού του κεφαλαίου.

Παρακάτω παρουσιάζονται οι σχέσεις δύο χρονικών στιγμών t και $t-k$ μιας στάσιμης χρονοσειράς βάσει κάποιων στατιστικών δεικτών, όπως η μέση τιμή, η διακύμανση και οι αυτοσυνδιακυμάνσεις:

Μέση τιμή για τις χρονικές στιγμές t και $t-k$:

$$E(y_t) = E(y_{t-k}) = \mu \quad (1.10)$$

Διακύμανση για τις χρονικές στιγμές t και $t-k$:

$$E[(y_t - \mu)^2] = E[(y_{t-k} - \mu)^2] = \sigma_y^2 \quad \{ \text{var}(\cdot) = \text{var}(y_{t-k}) = \sigma_y^2 \} \quad (1.11)$$

Αυτοσυνδιακυμάνσεις για τις χρονικές στιγμές t και $t-k$:

$$E[(y_t - \mu)(y_{t-k} - \mu)] = E[(y_{t-j} - \mu)(y_{t-j-k} - \mu)] = \gamma_k \quad (1.12)$$

Ένα πολύ χρήσιμο εργαλείο, για τον αρχικό έλεγχο της στασιμότητας αποτελεί το κορελόγραμμα. Ως κορελόγραμμα ορίζεται το διάγραμμα των αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων. Το πιο σημαντικό χαρακτηριστικό μη στάσιμων χρονοσειρών είναι ότι οι μερικές αυτοσυσχετίσεις και οι αυτοσυσχετίσεις στο κορελόγραμμα δεν «σβήνουν» ή όταν «σβήνουν», τότε αυτό πραγματοποιείται πολύ αργά.

Η έλλειψη στασιμότητας δημιουργεί δύο σημαντικά προβλήματα. Αρχικά, αν στη χρονοσειρά επέλθει μια απροσδόκητη μεταβολή (shock), τότε η μεταβολή αυτή δεν φθίνει στο πέρασμα του χρόνου και έτσι η σειρά δεν συγκλίνει στη μακροχρόνια μέση τιμή της. Σε αντίθεση με τις στάσιμες χρονοσειρές όπου με τον καιρό συγκλίνουν στη μέση τιμή τους. Επιπροσθέτως μια μη στάσιμη χρονοσειρά είναι πολύ πιθανόν να οδηγήσει σε αναξιόπιστα αποτελέσματα. Παράλληλα δεν ισχύουν οι υποθέσεις της γραμμικής παλινδρόμησης,

αφού η t και η F δεν ακολουθούν τις αντίστοιχες κατανομές, άρα οι στατιστικοί έλεγχοι δεν είναι έγκυροι.

1.8.1 Στατιστικός έλεγχος Dickey - Fuller

Ο έλεγχος των Dickey και Fuller (1969) αποτελεί τον πιο γνωστό έλεγχο, που δημιουργήθηκε με σκοπό να ελέγξει, την ύπαρξη στασιμότητας σε μια χρονοσειρά. Έστω το παρακάτω αυτοπαλίνδρομο μοντέλο πρώτου βαθμού:

$$y_t = \alpha + \beta t + \rho y_{t-1} + e_t \quad (1.13)$$

Αν $|\rho| < 1$ και $\beta = 0$ η χρονοσειρά y θα είναι στάσιμη. Αν $\beta \neq 0$ τότε η χρονοσειρά δεν είναι στάσιμη, καθώς δεν είναι σταθερή η μέση τιμή της. Αν $|\rho| = 1$, η χρονοσειρά έχει μοναδιαία ρίζα (δεν είναι στάσιμη), ενώ αν $|\rho| > 1$, τότε η διακύμανση της y θα αυξάνεται, καθώς ο χρόνος προχωρά και θα προσεγγίζει το άπειρο.

Με σκοπό τον έλεγχο σε μοντέλα με περισσότερες χρονικές υστερήσεις, δημιουργήθηκε ο επαυξημένος έλεγχος των Dickey-Fuller(ADF):

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{j=1}^p \lambda_j \Delta y_{t-j} + e_t, \quad t = \frac{\hat{Y}}{se(\hat{Y})} \quad (1.14)$$

Η αρχική υπόθεση (null hypothesis, H_0) του ελέγχου ορίζει, ότι η χρονοσειρά είναι μη στάσιμη, δηλαδή έχει μοναδιαία ρίζα. Η εναλλακτική υπόθεση (alternative hypothesis, H_1) ορίζει, ότι η χρονοσειρά είναι στάσιμη, δηλαδή δεν έχει μοναδιαία ρίζα. Μια χρονοσειρά για να είναι στάσιμη πρέπει να απορριφθεί η αρχική υπόθεση, δηλαδή ότι η χρονοσειρά δεν έχει μοναδιαία ρίζα (Αγγελάκη, 2020)

p-value $\leq 0,05$: Απορρίπτεται H_0 , στάσιμη χρονοσειρά

p-value $> 0,05$: Δεκτή H_0 , μη-στάσιμη χρονοσειρά

Η εγκυρότητα του ελέγχου επιτυγχάνεται με την μη ύπαρξη αυτοσυσχέτισης στα κατάλοιπα. Παράλληλα η ορθή επιλογή του αριθμού των χρονικών υστερήσεων πρέπει να γίνεται σύμφωνα με το κριτήριο του AIC (Agiakloglou & Newbold, 1992).

Παράλληλα ο ιδανικός αριθμός χρονικών υστερήσεων προέρχεται από τα δεδομένα και δεν είναι καθορισμένος από άλλους παράγοντες.

1.8.2 Στατιστικός έλεγχος KPSS

Οι Kwiatkowski, Phillips, Schmidt and Shin το 1992 έφτιαξαν έναν έλεγχο μοναδιαίας ρίζας, διαφορετικής προσέγγισης σε σχέση με αυτόν των Dickey και Fuller, αφού για αρχική υπόθεση (H_0) λαμβάνεται η υπόθεση, ότι η χρονοσειρά είναι στάσιμη $I(0)$ (Αγιακλόγλου & Γκάρος, 2022).

Το υπόδειγμα που βασίζεται το KPSS test είναι:

$$y_t = d_t + r_t + e_t \quad (1.15)$$

$$r_t = r_{t-1} + u_t \quad (1.16)$$

όπου:

- d_t : Προσδιοριστική τάση και ο σταθερός όρος (ντετερμινιστικά στοιχεία της χρονοσειράς)
- r_t : Τυχαίος περίπατος με διακύμανση σ_u^2
- e_t : είναι $iidN(0, \sigma_e^2)$
- u_t είναι $iidN(0, \sigma_u^2)$

Το KPSS τεστ στηρίζεται στον πολλαπλασιαστική Lagrange (Lagrange Multiplier, LM), αφού ελέγχεται η υπόθεση αν ο όρος του τυχαίου περιπάτου (r_t) έχει μηδενική διακύμανση εναντίον της εναλλακτικής υπόθεσης ότι δεν έχει.

$$H_0: \sigma_v^2 = 0, \text{ στάσιμη χρονοσειρά}$$

$$H_1: \sigma_v^2 > 0, \text{ Ύπαρξη μοναδιαίας ρίζας (μη στάσιμη χρονοσειρά)}$$

Ο έλεγχος LM :

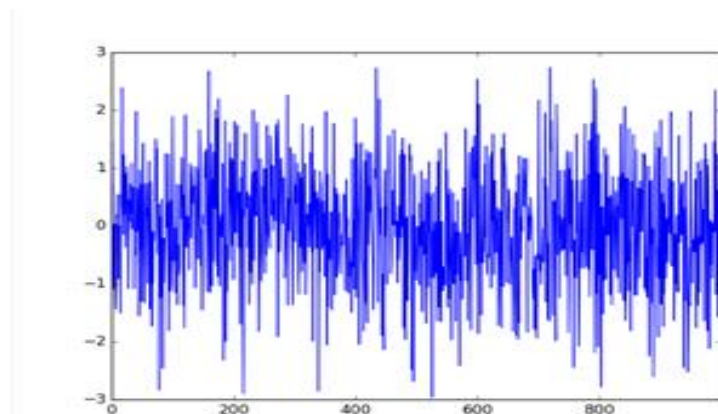
$$LM = LM = = \frac{s_t^2}{\hat{\sigma}_e^2} \quad (1.17)$$

όπου $s_t^2 = \sum_{t=1}^T \varepsilon_t$, $t = 1, 2, \dots, T$ και $\hat{\sigma}_\varepsilon^2$ η εκτιμώμενη διακύμανση των σφαλμάτων ε_t .

1.8.3 Λευκός θόρυβος

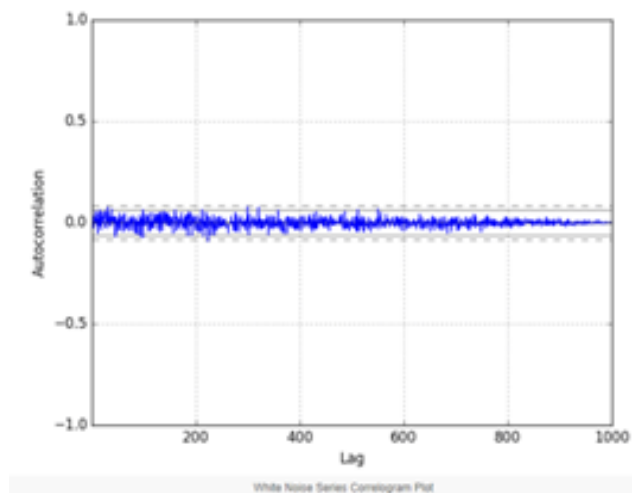
Έστω μια χρονοσειρά που αποτελείται από διαδοχικές ανεξάρτητες τυχαίες μεταβλητές $x_t, x_{t+1}, \dots, x_{t+\tau}$ με $\tau > 1$ που ακολουθούν την ίδια κατανομή (independent and identically distributed, iid), οι οποίες δεν περιέχουν αυτοσυσχετίσεις ονομάζεται λευκός θόρυβος (white noise) και συμβολίζεται με $WN(0, \sigma_\varepsilon^2)$. Αυτό σημαίνει πως η μέση τιμή ισούται με μηδέν και η διασπορά σ_ε^2 . Στην περίπτωση που τα στοιχεία της χρονοσειράς λευκού θορύβου ακολουθούν την κανονική (Γκαουσιανή) κατανομή, τότε η χρονοσειρά καλείται Γκαουσιανός λευκός θόρυβος (Gaussian White Noise) (Κουγιουμτζής, 2019-2020).

Στο Διάγραμμα 1.5 παρουσιάζεται παράδειγμα λευκού θορύβου τυχαίου Γκαουσιανού δείγματος κανονικής κατανομής μεγέθους 1000 και τυπικής απόκλισης 1, με την βοήθεια της γλώσσας Python:



Διάγραμμα 1.5: Διάγραμμα λευκού θορύβου τυχαίου δείγματος από Γκαουσιανή κατανομή (Brownlee, 2020)

Είναι φανερό, ότι από το Διάγραμμα 1.6, δηλαδή το διάγραμμα των αυτοσυσχετίσεων δεν υπάρχει καμία στατιστικά σημαντική μεταβλητή, άρα το παράδειγμα αυτό είναι χαρακτηριστικό λευκού θορύβου.



Διάγραμμα 1.6: Διάγραμμα αυτοσυσχετίσεων τυχαίου δείγματος από Γκαουζιανή κατανομή (Brownlee, 2020)

1.8.4 Τυχαίος περίπατος

Τυχαίος περίπατος (random walk) ορίζεται ως μια μη-στάσιμη χρονοσειρά, όπου κάθε στοιχείο x_t προκύπτει, όταν στο προηγούμενο στοιχείο της x_{t-1} προστεθεί μια τυχαία μεταβλητή ε_t . Μια χρονοσειρά είναι τυχαίος περίπατος αν:

$$x_t = x_{t-1} + \varepsilon_t = x_0 + \sum_{j=1}^t \varepsilon_j \quad (1.18)$$

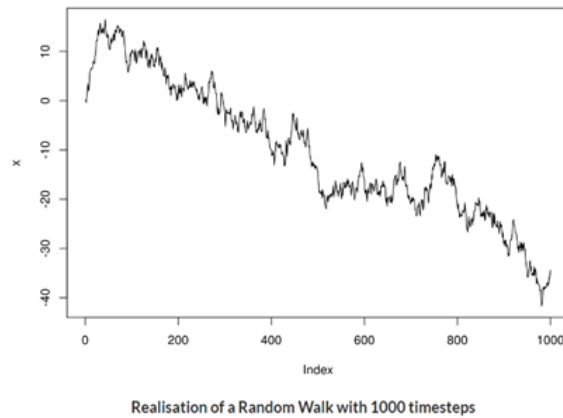
όπου $\{ \varepsilon_t \}$ είναι χρονοσειρά λευκού θορύβου, $\varepsilon_t \sim \text{WN}(0, \sigma_\varepsilon^2)$

Έτσι γίνεται αντιληπτό, ότι η χρονοσειρά παράγεται από τυχαία κίνηση πάνω σε ευθεία γραμμή (στο σύνολο των πραγματικών αριθμών), που σε κάθε χρονική στιγμή κάνει ένα τυχαίο βήμα μπρος ή πίσω (ε_t) από το σημείο που βρίσκεται (x_{t-1}) στο επόμενο (x_t). Στην περίπτωση που η τιμή x_0 είναι 0, τότε η χρονοσειρά λευκού θορύβου είναι ανεξάρτητη από την x_0 . Για αυτό τον λόγο η μέση τιμή του τυχαίου περιπάτου είναι 0 και η διακύμανση αυξάνεται με την πάροδο του χρόνου, αφού είναι ανάλογη του χρόνου t .

$$E[x_t] = 0 \quad (1.19)$$

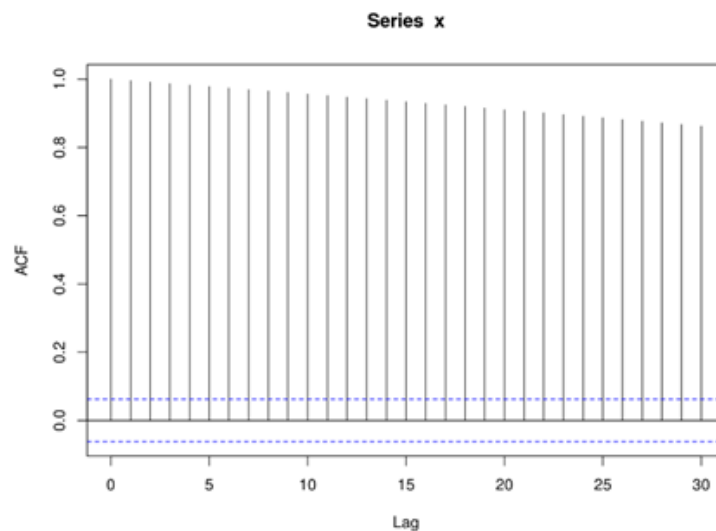
$$\text{Var}(x_t) = E(x_t^2) = t\sigma_\varepsilon^2 \quad (1.20)$$

Στο Διάγραμμα 1.7 παρουσιάζεται παράδειγμα τυχαίου περιπάτου τυχαίου δείγματος κανονικής κατανομής μεγέθους 1000, με την βοήθεια της γλώσσας Python:



Διάγραμμα 1.7: Διάγραμμα τυχαίου περιπάτου τυχαίου δείγματος από κανονική κατανομή (Quantsart – Web)

Αυτό που παρατηρείται έντονα στο Διάγραμμα 1.8 είναι πως όλα τα lags, στο διάγραμμα των αυτοσυσχετίσεων (ACF) είναι στατιστικά σημαντικά, πράγμα που δηλώνει πως η χρονοσειρά έχει την εικόνα τυχαίου περιπάτου.



Διάγραμμα 1.8: Διάγραμμα αυτοσυσχετίσεων τυχαίου περιπάτου τυχαίου δείγματος από κανονική κατανομή (Quantsart – Web)

1.9 Μέθοδοι προβλέψεων

Η ανάγκη των ανθρώπων να προβλέψουν μελλοντικές καταστάσεις με σκοπό την καλύτερη δυνατή οργάνωση στρατηγικής για την αποφυγή αποτυχιών ή την αύξηση κερδών έκανε απαραίτητη την δημιουργία μεθόδων και τεχνικών προβλέψεων. Είναι αναγκαίο για κάθε επιχείρηση να κάνει χρήση τεκμηριωμένων προβλέψεων με σκοπό την αποφυγή αποτυχιών. Οι προβλέψεις βρίσκουν εφαρμογή σε πολλούς κλάδους, όπως στις πωλήσεις, στις δαπάνες, τα κέρδη, αλλά και τις ζημιές. Για αυτό το λόγο η ανάγκη για την δημιουργία πιο εύστοχων προβλέψεων και σε συνδυασμό με την βελτίωση της τεχνολογίας αναπτύχθηκαν νέοι αλγόριθμοι που χρησιμοποιούν την τεχνητή νοημοσύνη (AI) με σκοπό την μείωση στο ελάχιστο του σφάλματος και την ακρίβεια του αποτελέσματος της πρόβλεψης.

Παραδείγματα που βρίσκουν εφαρμογή οι προβλέψεις είναι:

- Πρόβλεψη πωλήσεων
- Ανάλυση ρίσκου
- Μελέτη των κλιματικών αλλαγών
- Μελέτη του ποσοστού ανεργίας
- Αριθμός επιβατών αεροπλάνων ανά χώρα κ.α

1.9.1 Μέθοδος Naive

Η «απλοϊκή» πρόβλεψη (Naive forecast) αποτελεί την πιο απλή μέθοδο πρόβλεψης, όπου η πρόβλεψη της επόμενης περιόδου ισούται με την πραγματική τιμή της τρέχουσας περιόδου και πολλές φορές χρησιμοποιείται ως «κατώφλι», δηλαδή όταν γίνεται χρήση άλλων πιο σύνθετων μεθόδων υπάρχει σύγκριση με αυτήν με σκοπό να γίνει έλεγχος αν δίνονται καλύτερα αποτελέσματα. Αν η πιο σύνθετη μέθοδος δεν εξάγει καλύτερα αποτελέσματα, τότε σημαίνει πως η προβλεπτική της ικανότητα δεν πρέπει να ληφθεί υπόψιν από τον μελετητή. Τα πλεονεκτήματα της μεθόδου Naïve είναι η απλότητα και το χαμηλό κόστος. Αν ο μέσος όρος, η τάση, η εποχικότητα είναι σταθερά, καθώς και τα τυχαία γεγονότα μικρά έχουν μικρό εύρος, τότε τα αποτελέσματα της μεθόδου αυτής είναι αρκετά αξιόπιστα.

Στην περίπτωση που τα τυχαία γεγονότα είναι αρκετά μεγάλα, τότε θα υπάρχουν υψηλές διακυμάνσεις για την πρόβλεψη της επόμενης περιόδου, άρα η μέθοδος αυτή δεν θα αποτελεί αξιόπιστη πηγή προβλέψεων για τον ορθό στρατηγικό σχεδιασμό της επιχείρησης.

Ο μαθηματικός τύπος της μεθόδου Naive είναι ο παρακάτω:

$$\hat{Y}_{t+1} = Y_t \quad (1.21)$$

1.9.2 Απλός κινητός μέσος

Ο απλός κινητός μέσος (Simple Moving Average, SMA) αποτελεί μέθοδος πρόβλεψης, η οποία είναι πολύ απλή στην εφαρμογή της. Χρησιμοποιεί ως πρόβλεψη την τιμή του αριθμητικού μέσου όρου m πρόσφατων παρατηρήσεων της χρονοσειράς. Βασίζεται στο γεγονός πως οι πιο πρόσφατες παρατηρήσεις εκφράζουν πιο ξεκάθαρα την εικόνα της χρονοσειράς σε σχέση με τις πιο μακρινές παρατηρήσεις. Ένα μειονέκτημα της μεθόδου αυτής είναι πως όλες οι παρατηρήσεις έχουν τα ίδια βάρη, ακόμα και αν περιέχουν πιο σημαντική πληροφορία. Η ονομασία «κινητός μέσος όρος» προέρχεται από το γεγονός πως, με την εισαγωγή μιας νέας παρατήρησης, η τιμή του αναπροσαρμόζεται.

Έστω μια χρονοσειρά Y_t για $t=1,2,3,4,\dots,n$. Οι προβλέψεις του κινητού μέσου ορίζονται ως:

$$\hat{Y}_{t+1} = M_{t+1} = \frac{1}{m} \sum_{j=1}^m Y_{t-j+1} = \frac{1}{m} (Y_t + Y_{t-1} + \dots + Y_{t-m+1}) \quad (1.22)$$

όπου \hat{Y}_{t+1} , η πρόβλεψη της Y_t για την χρονική περίοδο $(t+1)$.

Με σκοπό την υλοποίηση της παραπάνω μεθόδου είναι απαραίτητη η γνώση του m . Σε πραγματικά δεδομένα είναι πολύ πιθανόν η τιμή του να μην είναι καθορισμένη, αλλά να χρειαστεί ο υπολογισμός της.

Με σκοπό τον ορθό υπολογισμό του κινητού μέσου όρου, ώστε να μην υπάρχουν αυθαιρεσίες στην πρόβλεψη, γίνεται χρήση του κριτηρίου του μέσου τετραγωνικού Σφάλματος (MSE) ή κάποιου άλλου αντίστοιχου κριτηρίου. Επιλέγεται η τιμή του m , που ελαχιστοποιεί την τιμή του μέσου τετραγωνικού σφάλματος. Εφόσον το m είναι γνωστό, ο τύπος αναπροσαρμόζεται κάθε φορά που εισέρχεται μια νέα παρατήρηση και δίνεται από τον τύπο:

$$\hat{Y}_{t+1} = \hat{Y}_t + Y_{t-m} + \frac{Y_t}{m} - \frac{Y_{t-m}}{m}. \quad (1.23)$$

Για μεγάλο αριθμό m ($m=30$), η τιμή της πρόβλεψης εξαρτάται από τις τιμές: Y_t , Y_{t-m} και όχι από όλες τις τιμές των m πιο πρόσφατων παρατηρήσεων. Όσο μικραίνει το m , τόσο πιο ευαίσθητη θα είναι και η μέθοδος στις μεταβολές των τιμών. Αντίστοιχα όσο αυξάνεται το m , τόσο καλύτερα θα εξομαλύνονται. Ο απλός κινητός μέσος χρησιμοποιείται συχνά για τον προσδιορισμό της κατεύθυνσης της τάσης. Αν κινείται ανοδικά, τότε και η τάση ανοδική. Αντίστοιχα αν κινείται καθοδικά, τότε και η τάση είναι καθοδική. Παράλληλα οι μικρές τιμές του m χρησιμοποιούνται για τον υπολογισμό βραχυπρόθεσμων τάσεων, ενώ αντίστοιχα για μεγάλες τιμές του m γίνεται υπολογισμός της τάσης σε μακροχρόνια περίοδο. Οι πιο συνηθισμένες τιμές του m είναι 10, 20, 50, 100 και 200.

1.9.3 Διπλός κινητός μέσος

Ο διπλός κινητός μέσος (Double Moving Average, DMA) είναι μέθοδος πρόβλεψης, ο οποίος χρησιμοποιείται σε χρονοσειρές, όπου υπάρχει γραμμική τάση, αλλά δεν παρατηρείται εποχικότητα. Ο υπολογισμός του διπλού κινητού μέσου γίνεται μέσω του απλού κινητού μέσου, όπου στη συνέχεια γίνεται ο υπολογισμός του κινητού μέσου των πρώτων κινητών μέσων λαμβάνοντας υπόψη και την γραμμική τάση. Τρόπος υπολογισμού του διπλού κινητού μέσου:

α) Απλός κινητός μέσος m περιόδων, M_t :

$$M_{t+1} = \frac{1}{m} \sum_{j=1}^m Y_{t-j+1} = \frac{1}{m}(Y_t + Y_{t-1} + \dots + Y_{t-m+1}) \quad (1.24)$$

β) Διπλός κινητός μέσος m περιόδων, M'_t :

$$M'_{t+1} = \frac{1}{m} \sum_{j=1}^m M_{t-j+1} = \frac{1}{m}(M_t + M_{t-1} + \dots + M_{t-m+1}) \quad (1.25)$$

γ) Υπολογισμός διαφοράς a_t

$$a_t = 2M_t - M'_t \quad (1.26)$$

δ) Υπολογισμός παράγοντα της τάσης b_t :

$$b_t = \frac{2}{m-1} (M_t - M'_t) \quad (1.27)$$

Υπολογισμός πρόβλεψης \hat{Y}_{t+h} για h μελλοντικής περιόδου

$$\hat{Y}_{t+1} = a_t + hb_t, h > 0 \quad (1.28)$$

Ο διπλός κινητός μέσος δίνει την δυνατότητα πρόβλεψης μεγαλύτερης της μιας περιόδου, εφόσον το h έχει τιμή μεγαλύτερη της μονάδας ($h > 1$). Αν το h ισούται με 1 ($h=1$), τότε η πρόβλεψη αφορά μια περίοδο. Ο τρόπος υπολογισμού του βέλτιστου m είναι ίδιος με αυτόν του απλού κινητού μέσου, αφού και στις δύο περιπτώσεις η ιδανική επιλογή του m εμφανίζεται στην ελαχιστοποίηση της τιμής ενός πληροφοριακού κριτηρίου, όπως για παράδειγμα του μέσου τετραγωνικού σφάλματος (MSE).

1.9.4 Σταθμισμένος κινητός μέσος

Ο σταθμισμένος κινητός μέσος (Weighted Moving Average, WMA) δίνει μεγαλύτερη βαρύτητα στις πιο πρόσφατες παρατηρήσεις και λιγότερη στις πιο απομακρυσμένες παρατηρήσεις. Αυτό πραγματοποιείται τοποθετώντας «βάρη» ή αλλιώς συντελεστές στις παρατηρήσεις, με τα μεγαλύτερα «βάρη» να βρίσκονται στις πρόσφατες παρατηρήσεις. Το άθροισμα όλων των σταθμίσεων πρέπει να ισούνται με την μονάδα (Ορφανού, 2015) Ο ορισμός του σταθμισμένου κινητού μέσου είναι ο ακόλουθος:

$$WMA = \sum_{k=1}^n w_k Y_{t-k} = w_1 \cdot Y_{t-1} + w_2 \cdot Y_{t-2} + \dots + w_n \cdot Y_{t-n} \quad (1.29)$$

όπου:

n = ο αριθμός των παρατηρήσεων

Y_{t-1} = η πραγματική τιμή της περιόδου $t-1$

w_k = το βάρος για την περίοδο $t-k$

1.9.5 Απλή εκθετική εξομάλυνση

Ο απλός κινητός μέσος (SMA) δημιουργεί προβλέψεις αντιμετωπίζοντας τις παρατηρήσεις ισάξια, δηλαδή το «βάρος» τους είναι κοινό και δεν εξαρτάται από την χρονική σειρά. Αυτό το γεγονός συχνά δημιουργεί προβλήματα, με συνέπεια οι προβλέψεις να είναι ανακριβείς, ιδιαίτερα όταν το m μεγαλώνει, π.χ. $m=50$. Αυτό το πρόβλημα ήρθε να το λύσει η μέθοδος της απλής εκθετικής εξομάλυνσης (Simple Exponential Smoothing, SMA), όπου η πρόβλεψη στηρίζεται σε ένα είδος σταθμικού μέσου όρου, όπου αυτός δίνει πολύ μεγαλύτερη έμφαση σε πρόσφατες παρατηρήσεις σε εν αντιθέσει με τις απομακρυσμένες.

Η μέθοδος αυτή κατάλληλη για δεδομένα που τα συνθετικά στοιχεία της χρονοσειράς δεν είναι σαφή. Παράλληλα αποτελεί συνήθης επιλογή, αφού έχει εύκολο υπολογισμό και ο αριθμός των παρατηρήσεων που απαιτούνται για την χρήση της δεν είναι αρκετά μεγάλος. Οι προβλέψεις της χρονοσειράς δημιουργούνται ως εξής:

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \dots \quad (1.30)$$

όπου α η σταθερά εξομάλυνσης (smoothing constant), η οποία λαμβάνει τιμές από 0 έως 1.

Το άθροισμα $\alpha + \alpha(1-\alpha) + \alpha(1-\alpha)^2 + \dots$ αποτελεί άθροισμα όρων γεωμετρικής προόδου και λόγω του ότι $(1-\alpha) < 1$, το άθροισμα αυτό ισούται με 1, δηλαδή: $\frac{\alpha}{\alpha - (1-\alpha)} = 1$. Όσο αυξάνεται η τιμή του α , τόσο μεγαλύτερη βαρύτητα δίνεται στις πρόσφατες παρατηρήσεις και κατά συνέπεια πολύ μικρότερη στις πιο απομακρυσμένες παρατηρήσεις, αφού το "βάρος" κάθε παρατήρησης μειώνεται εκθετικά κατά $(1-\alpha)$.

Αν τα δεδομένα παρουσιάζουν μεγάλες διακυμάνσεις, τότε το α θα είναι σχετικά μεγάλο, ενώ αν παρουσιάζουν σταθερότητα το α θα λαμβάνει σχετικά χαμηλές τιμές.

Η σχέση (1.31) μπορεί να γραφτεί και στην εξής μορφή:

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha) \hat{Y}_t \quad (1.32)$$

Η πρόβλεψη της τιμής της περιόδου $t+1$ μπορεί να υπολογιστεί ως ο σταθμικός μέσος της πραγματικής τιμής και της τιμής της πρόβλεψης την χρονική στιγμή t , με αντίστοιχα βάρη 1 και $1-\alpha$. Επίσης ισχύει ότι: $\hat{Y}_2 = Y_1$

Έστω ότι το α ισούται με 0.3 , δηλαδή $\alpha=0.3$. Αυτό σημαίνει ότι κατά 30% η τιμή υπολογίζεται από την πραγματική τιμή και 70% από την προβλεπόμενη τιμή της τρέχουσας περιόδου.

Ειδικές περιπτώσεις:

- Αν $\alpha=1$ τότε η τιμή της πρόβλεψης ($t+1$) ταυτίζεται με την πραγματική τιμή της περιόδου t .
- Αν $\alpha=0$ η πρόβλεψη της περιόδου ($t+1$) ταυτίζεται με την πρόβλεψη της περιόδου t .

Είναι στην κρίση του μελετητή, η τιμή του α που θα επιλέξει. Αυτό εξαρτάται από το μέγεθος της εξομάλυνσης που θέλει να κάνει χρήση. Εφόσον κρίνει ότι χρειάζεται μεγάλη εξομάλυνση στις τιμές, τότε η τιμή του α θα είναι μικρή, σε αντίθεση με το αν κρίνει πως χρειάζεται μεγαλύτερη προσέγγιση στις πραγματικές μεταβολές των τιμών, τότε θα κάνει χρήση μεγάλης τιμής στο α , πράγμα που δημιουργεί πολλές φορές προβλήματα στις ορθές προβλέψεις.

1.9.6 Διπλή εκθετική εξομάλυνση

Η διπλή εκθετική εξομάλυνση (Double Exponential Smoothing, Brown Method) είναι μια μέθοδος προβλέψεων που χρησιμοποιείται σε χρονοσειρές που υπάρχει τάση. Η φιλοσοφία της εφαρμογής της είναι παραπλήσια με αυτή του διπλού κινητού μέσου, αφού εφαρμόζεται δύο φορές εξομάλυνση στις παρατηρήσεις, όμως λαμβάνεται υπόψιν και η τάση.

Σημαντική διαφορά αποτελεί, πως στην διπλή εξομάλυνση χρειάζονται πολύ λιγότερες παρατηρήσεις σε σχέση με τον διπλό κινητό μέσο. Παράλληλα η μέθοδος Brown παρέχει προβλέψεις για πάνω από μια χρονική περίοδο. Αντίστοιχα χρησιμοποιούνται τα πληροφοριακά κριτήρια σφάλματος, όπως το μέσο τετραγωνικό σφάλμα με σκοπό την ελαχιστοποίηση της τιμής του για την εύρεση της βέλτιστης τιμής του α .

Παρακάτω παρατίθενται οι υπολογισμοί που επιτελούνται για την εφαρμογή της μεθόδου Brown:

α) Εξομάλυνση των αρχικών παρατηρήσεων με την μέθοδο της Απλής Εκθετικής Εξομάλυνσης

$$A_t = \alpha Y_t + (1-\alpha)A_{t-1} \quad (1.33)$$

όπου A_t οι εξομαλυμένες τιμές της χρονοσειράς και $t = 2, 3, \dots, n$. Αν $t=1$ τότε $A_t = Y_t$

β) Εξομάλυνση των ήδη εξομαλυσμένων τιμών:

$$A'_t = \alpha A_t + (1-\alpha)A'_{t-1} \quad (1.34)$$

όπου α η σταθερά εξομάλυνσης και $t \geq 2$. Αν $t=1$ τότε $A'_1 = A_1$

γ) Γίνεται ο υπολογισμός της διαφοράς a_t :

$$a_t = 2A_t - A'_t \quad (1.35)$$

δ) Υπολογισμός του b_t που αποτελεί τον παράγοντα προσαρμογής της τάσης:

$$b_t = \frac{\alpha}{1-\alpha} (A_t - A'_t) \quad (1.36)$$

ε) Υπολογισμός της πρόβλεψης \hat{Y}_{t+h} για την χρονική στιγμή $t+h$ (για την μελλοντική περίοδο h)

$$\hat{Y}_{t+h} = a_t + hb_t \quad (1.37)$$

1.9.7 Εκθετική εξομάλυνση με προσαρμογή στην τάση

Η Εκθετική εξομάλυνση με προσαρμογή στην τάση (Exponential Smoothing adjusted for Trend, Holt method) ή μέθοδος Holt όπως συνηθίζεται είναι μέθοδος εξομάλυνσης, που βρίσκει εφαρμογή σε χρονοσειρές που περιέχουν τάση, όπως αντίστοιχα και η διπλή εκθετική εξομάλυνση. Η σημαντική τους διαφορά αποτελεί, πως στη μέθοδο Holt περιέχονται δύο σταθερές, το α και το β σε σχέση με την διπλή εκθετική εξομάλυνση. Παράλληλα άλλη μια διαφορά είναι, πως η μέθοδος Holt εξομαλύνει κατευθείαν τις τιμές της τάσης σε σχέση με την

μέθοδο Brown που χρησιμοποιεί δύο εξομαλύνσεις σε συνδυασμό με τον προσδιορισμό του παράγοντα b . Για την εφαρμογή της μεθόδου Holt δημιουργούνται δύο εξισώσεις. Η πρώτη υπολογίζει την τιμή των χρονοσειρών και η δεύτερη υπολογίζει την τάση (κλίση) των παρατηρήσεων της χρονοσειράς. Αναλυτικότερα:

α) Εξομάλυνση των τιμών της χρονοσειράς:

$$A_t = \alpha Y_t + (1-\alpha)(A_{t-1} + T_{t-1}) \quad (1.38)$$

όπου α η σταθερά εξομάλυνσης και λαμβάνει τιμές στο $0 < \alpha < 1$, A_t οι εξομαλυνθείσες τιμές της χρονοσειράς για $t= 2,3, \dots, n$, αφού για $t=1$ $A_1=Y_1$

β) Εξομάλυνση της τάσης (κλίσης)

$$T_t = \beta(A_t - A_{t-1}) + (1-\beta) T_{t-1} \quad (1.39)$$

όπου β η σταθερά για την εξομάλυνση της τάσης, η οποία και αυτή λαμβάνει τιμές μεταξύ $0 \leq \beta \leq 1$, T_t οι εξομαλυνθείσες τιμές της τάσης για $t \geq 2$, αφού για $t=1$ $T_1=0$.

γ) Ως τελικός σκοπός αποτελεί η πρόβλεψη της \hat{Y}_{t+h} τιμής για την h περίοδο πρόβλεψης:

$$\hat{Y}_{t+h} = A_t + hT_t \quad (1.40)$$

Με σκοπό την βέλτιστη πρόβλεψη η μέθοδος Holt, όπως και οι υπόλοιπες μέθοδοι χρησιμοποιούν την ελαχιστοποίηση της τιμής κάποιου πληροφοριακού κριτηρίου σφάλματος, όπως για παράδειγμα του μέσου τετραγωνικού σφάλματος (MSE).

Η μέθοδος Holt χρησιμοποιείται πολύ συχνότερα σε σχέση με την μέθοδο Brown, αφού δεν είναι τόσο επιρρεπής σε τυχαίες διακυμάνσεις.

1.9.8 Εκθετική εξομάλυνση με προσαρμογή στην τάση και την εποχικότητα

Η εκθετική εξομάλυνση με προσαρμογή στην τάση και την εποχικότητα ή αλλιώς μέθοδος Winters (Exponential smoothing adjusted for trend and seasonality, Winters Method) χρησιμοποιείται σε χρονοσειρές που περιέχουν στοιχεία τάσης και εποχικότητας.

Αποτελεί μια επέκταση της μεθόδου Holt, αφού περιέχει τρεις παραμέτρους, τις α , β και γ . Τέτοιες περιπτώσεις συναντιούνται συνήθως σε οικονομικά δεδομένα, όπως στις μετοχές του χρηματιστηρίου. Η μορφή των δεδομένων επαναλαμβάνεται σε χρονικές περιόδους μικρότερες του έτους, όπως τρίμηνα, τετράμηνα, αφού δημιουργούνται εποχικές διακυμάνσεις. Για να αντιμετωπιστεί το πρόβλημα της εποχικότητας δημιουργήθηκαν μέθοδοι εξομάλυνσης για την καλύτερη δυνατή πρόβλεψη και την μείωση στο ελάχιστο του σφάλματος.

Παρακάτω παρουσιάζεται ο τρόπος εφαρμογής της μεθόδου Winters:

α) Εξομάλυνση τιμών της χρονοσειράς:

$$A_t = \alpha \frac{Y_t}{S_{t-L}} + (1-\alpha)(A_{t-1} + T_{t-1}) \quad (1.41)$$

όπου α η σταθερά εξομάλυνσης των τιμών της χρονοσειράς με $0 \leq \alpha \leq 1$, $t = 2, 3, \dots, n$, A_t οι εξομαλυνθείσες τιμές της χρονοσειράς, S_t ο εποχικός συντελεστής της περιόδου t και L η περιοδικότητα της εποχικότητας. Για παράδειγμα $L=4$ που αφορά τριμηνιαία δεδομένα και $L=2$ που αφορά εξαμηνιαία δεδομένα.

β) Εξομάλυνση της Τάσης (ομοίως με τη μέθοδο Holt):

$$T_t = \beta(A_t - A_{t-1}) + (1-\beta) T_{t-1} \quad (1.42)$$

όπου β είναι η σταθερά εξομάλυνσης της τάσης, ισχύει $0 \leq \beta \leq 1$, T_t είναι οι εξομαλυνθείσες τιμές της τάσης.

γ) Εξομάλυνση της εποχικότητας:

$$S_t = \gamma \frac{Y_t}{A_t} + (1-\gamma)S_{t-L} \quad (1.43)$$

όπου γ η σταθερά για την εξομάλυνση της εποχικότητας με $0 \leq \gamma \leq 1$

δ) Ως τελικός σκοπός είναι η πρόβλεψη \hat{Y}_{t+h}

Για τις h μελλοντικές περιόδους του πρώτου έτους η πρόβλεψη προσδιορίζεται από την παρακάτω σχέση:

$$\hat{Y}_{t+h} = (A_t + hT_t)S_{t+h-L} \quad (1.44), \text{ για } h = 1, 2, \dots, L$$

Για τις h μελλοντικές περιόδους του δεύτερου έτους η πρόβλεψη προσδιορίζεται από την παρακάτω σχέση:

$$\hat{Y}_{t+h} = (A_t + hT_t)S_{t+h-2L} \quad (1.45)$$

για $h = L+1, L+2, \dots, 2L$

Η μέθοδος Winters μπορεί να προβλέψει περισσότερες από μια περιόδους. Όπως στις προηγούμενες μεθόδους έτσι και εδώ για τον υπολογισμό των βέλτιστων παραμέτρων α, β και γ γίνεται μέσω της ελαχιστοποίησης κάποιου πληροφοριακού κριτηρίου σφάλματος, όπως του μέσου τετραγωνικού σφάλματος (MSE).

Οι τιμές A_t δεν προσδιορίζονται για $t = 1, 2, \dots, L-1$, ενώ για $t = L$ το A_L ορίζεται ως:

$$A_L = \frac{Y_1 + Y_2 + Y_3 + \dots + Y_L}{L} \quad (1.46)$$

1.9.9 Γραμμική παλινδρόμηση

Η γραμμική παλινδρόμηση αποτελεί το πιο απλό είδος παλινδρόμησης που συναντάται. Εκφράζει την σχέση μέσω μιας γραμμικής συνάρτησης μιας ή περισσότερων μεταβλητών (x_1, x_2, \dots, x_n) με την μεταβλητή στόχο (y) . Οι μεταβλητές x_1, x_2, \dots, x_n ονομάζονται ανεξάρτητες ή ερμηνευτικές μεταβλητές, ενώ η μεταβλητή y ονομάζεται εξαρτημένη μεταβλητή ή μεταβλητή απόκρισης. Αν η ανεξάρτητη μεταβλητή που μελετάται είναι μια τότε γίνεται αναφορά σε απλή γραμμική παλινδρόμηση, ενώ αν οι ανεξάρτητες μεταβλητές είναι περισσότερες από μια, γίνεται αναφορά για πολλαπλή γραμμική παλινδρόμηση.

Υπάρχουν δύο μοντέλα για να περιγράψουν την ανάλυση στην οποία υπόκειται το κάθε πρόβλημα:

- Προσδιοριστικό ή Μαθηματικό μοντέλο
- Στοχαστικό ή Στατιστικό μοντέλο

Το προσδιοριστικό μοντέλο στο οποίο η μεταβλητή Y ορίζεται πλήρως όταν οι τιμές των ερμηνευτικών μεταβλητών είναι γνωστές και δίνεται από τον ακόλουθο τύπο:

$$y_i = \beta_0 + \beta_1 x_i \quad (1.47)$$

- y_i η τιμή της μεταβλητής απόκρισης για το i ζεύγος των παρατηρήσεων
- x_i η τιμή της επεξηγηματικής μεταβλητής
- β_0, β_1 άγνωστες παράμετροι

Το στατιστικό μοντέλο περιγράφει την σχέση των ερμηνευτικών μεταβλητών με την μεταβλητή απόκρισης προσθέτοντας τον όρο του σφάλματος (ε). Με τον όρο σφάλμα προσδιορίζεται το γεγονός, ότι ακόμα και αν είναι γνωστές οι τιμές β_0 και β_1 δεν μπορεί να υπολογιστεί επακριβώς η μεταβλητή απόκρισης, αλλά προσδιοριστικά.

Η μαθηματική απεικόνιση του στατιστικού μοντέλου είναι η παρακάτω:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (1.48)$$

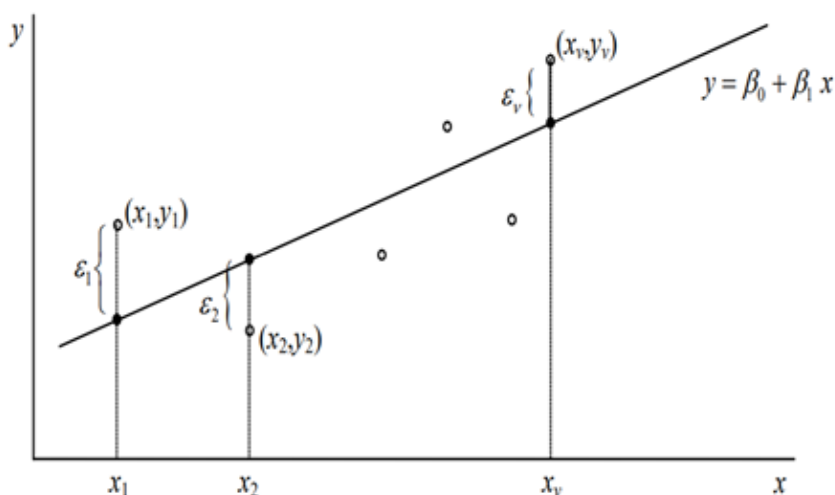
όπου:

- Y_i η τιμή της μεταβλητής απόκρισης για το i ζεύγος των παρατηρήσεων
- X_i η τιμή της επεξηγηματικής μεταβλητής
- β_0, β_1 άγνωστες παράμετροι
- $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ τα τυχαία σφάλματα, τα οποία κατανέμονται σύμφωνα με την κανονική κατανομή $N(0, \sigma^2)$, άρα είναι και ανεξάρτητα.

Πρακτικά σε πραγματικές συνθήκες το προσδιοριστικό μοντέλο είναι ανεπαρκές, διότι δεν λαμβάνει υπόψιν του τον όρο του σφάλματος, αφού σε πραγματικά δεδομένα δεν γίνεται να

είναι ακριβής σε πλήρες ποσοστό η σχέση μεταξύ ερμηνευτικών μεταβλητών και μεταβλητής απόκρισης.

Με σκοπό την εύρεση του κατάλληλου μοντέλου για να απεικονισθεί η σχέση μεταξύ μεταβλητής απόκρισης και επεξηγηματικής μεταβλητής γίνεται κατασκευή ενός διαγράμματος διασποράς (scatterplot) . Στο Διάγραμμα διασποράς τοποθετούνται οι τιμές της μεταβλητής x στον οριζόντιο άξονα και της μεταβλητής y στον κατακόρυφο άξονα, όπως εμφανίζεται στο Διάγραμμα 1.9 (προερχόμενο από σημειώσεις του μαθήματος «Ανάλυση Παλινδρόμησης και ανάλυση Διακύμανσης» από το ΠΜΣ Εφαρμοσμένης Στατιστικής με καθηγητή τον κ. Κούτρα Μάρκο). Ειδικότερα στο διάγραμμα αυτό απεικονίζονται οι τιμές της σχέσης μεταξύ της μεταβλητής x και της y με άσπρη κουκίδα. Παράλληλα η γραμμή που διαπερνά αυτά τα σημεία ονομάζεται ευθεία παλινδρόμησης. Ως σφάλματα ορίζονται οι αποστάσεις των τιμών των μεταβλητών από την ευθεία παλινδρόμησης.



Διάγραμμα 1.9: Διάγραμμα διασποράς (Μάρκος Κούτρας, Σημειώσεις ΠΜΣ «Εφαρμοσμένη Στατιστική»)

Η μέθοδος που χρησιμοποιείται για την εύρεση της καλύτερης δυνατής ευθείας ονομάζεται Μέθοδος Ελαχίστων Τετραγώνων (Method of Ordinary Least Squares, OLS), όπου σύμφωνα με αυτή η καλύτερη δυνατή ευθεία είναι αυτή που ελαχιστοποιεί το άθροισμα των τετραγώνων των σφαλμάτων ή καταλοίπων ϵ_i , δηλαδή το:

$$\sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i))^2 \quad (1.49)$$

Στην ανάλυση χρονοσειρών χρησιμοποιούνται κατά κύριο λόγο μοντέλα παλινδρόμησης (regression models), τα ευρέως γνωστά $AR(p)$, $MA(q)$, $ARMA(p,q)$, $ARIMA(p,d,q)$, τα οποία θα αναλυθούν αναλυτικά παρακάτω.

1.9.10 Μοντέλα AR, MA, ARMA

1.9.10.1 Αυτοπαλίνδρομα υποδείγματα

Τα αυτοπαλίνδρομα υποδείγματα (Autoregressive Models, $AR(p)$), μπορούν να θεωρηθούν ως μοντέλα παλινδρόμησης, αφού οι ανεξάρτητες μεταβλητές είναι οι τιμές της μεταβλητής απόκρισης X_t . Αυτό σημαίνει ότι η τρέχουσα τιμή βασίζεται στις προηγούμενες. Για παράδειγμα σε ένα αυτοπαλίνδρομο υπόδειγμα πρώτης τάξης ($AR(1)$), η τρέχουσα τιμή βασίζεται στην προηγούμενη τιμή, Αντίστοιχα σε ένα αυτοπαλίνδρομο υπόδειγμα δεύτερης τάξης ($AR(2)$), η τρέχουσα τιμή βασίζεται στις δύο προηγούμενες τιμές. Παρακάτω παρουσιάζεται το αυτοπαλίνδρομο υπόδειγμα $AR(p)$ με $\varepsilon_t \sim WN(0, \sigma^2)$:

$$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (1.50)$$

Ως p ορίζεται το μήκος της χρονικής υστέρησης, ή αλλιώς η τάξη του αυτοπαλίνδρομου υποδείγματος και ϕ_1, \dots, ϕ_p οι συντελεστές του μοντέλου, με $|\phi| < 1$. Όταν το ϕ λάβει την τιμή 1 ($\phi=1$), τότε η χρονοσειρά δεν είναι στάσιμη, αλλά έχει εικόνα τυχαίου περιπάτου, αφού η εξίσωση της βρίσκεται πάνω στον μοναδιαίο κύκλο (μοναδιαία ρίζα). Όταν το ϕ λάβει την τιμή 0 ($\phi=0$), τότε η χρονοσειρά έχει εικόνα λευκού θορύβου. Ως μ η μέση τιμή της χρονοσειράς και προφανώς, αν ισούται με 0 ($\mu=0$), τότε η χρονοσειρά είναι στάσιμη. Με την βοήθεια της ελαχιστοποίησης του αθροίσματος των τετραγώνων των σφαλμάτων $S(\mu, \phi_1, \phi_2, \dots, \phi_p)$ γίνεται η εκτίμηση των παραμέτρων $\mu, \phi_1, \dots, \phi_p$, και η μαθηματική έκφραση είναι η ακόλουθη:

$$\text{Min} S(\mu, \phi_1, \phi_2, \dots, \phi_p) = \min \sum_{t=p+1}^n (x_t - \mu - \phi_1(x_{t-1} - \mu) + \dots + \phi_p(x_{t-p} - \mu))^2 \quad (1.51)$$

Έστω μια στάσιμη χρονοσειρά X_t :

Χρησιμοποιώντας τον τελεστή υστέρησης B η χρονοσειρά λαμβάνει την παρακάτω μορφή:

$$X_t = \frac{1}{(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p)} \varepsilon_t \rightarrow$$

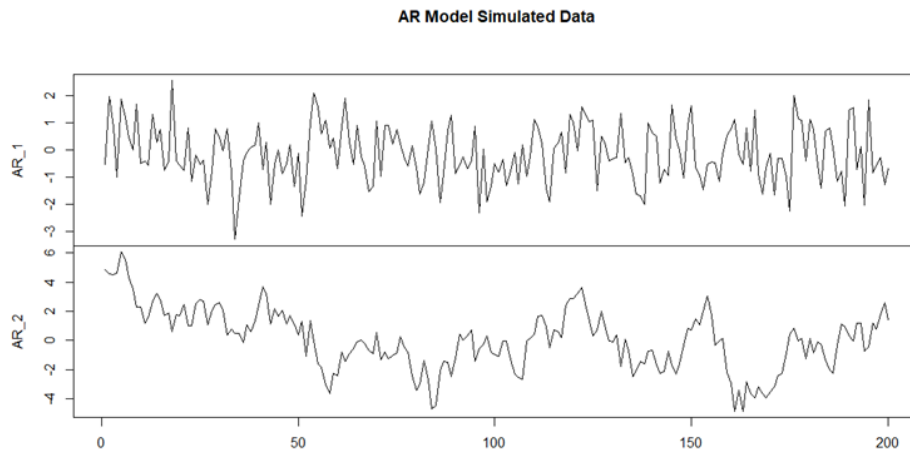
$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p) X_t = \varepsilon_t \rightarrow \quad (1.52)$$

$$\Phi(B) X_t = \varepsilon_t$$

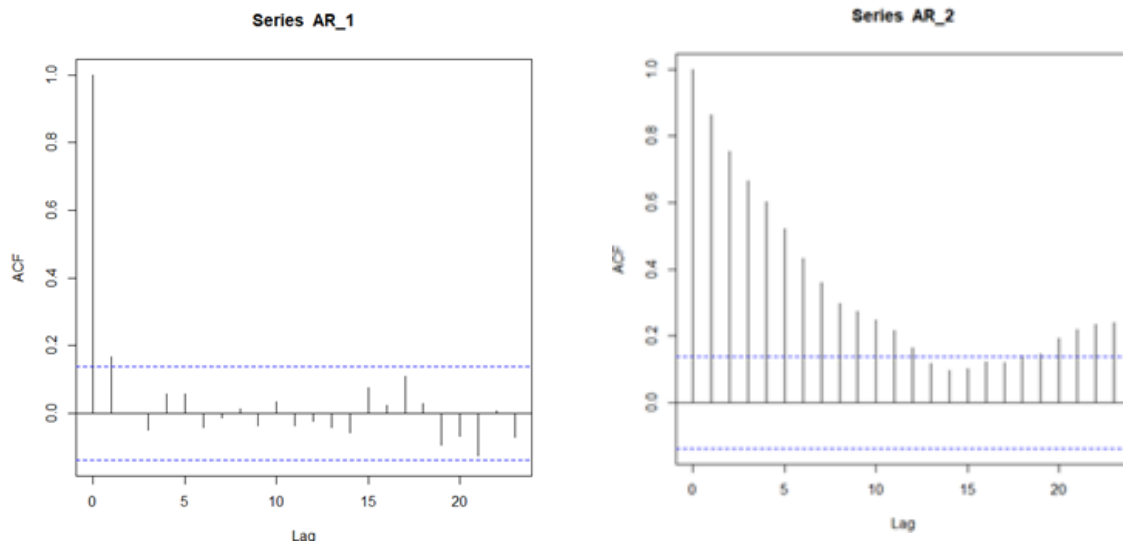
όπου $\Phi(B)$ το χαρακτηριστικό πολυώνυμο της $AR(p)$

Οι ρίζες του πολυωνύμου $\frac{1}{B^p} - \varphi_1 \frac{1}{B^{p-1}} - \dots - \varphi_p = 0$ θα πρέπει να είναι μικρότερες της μονάδας, ώστε να μη βρίσκονται εντός του μοναδιαίου κύκλου.

Στα Διαγράμματα 1.10 και στα αντίστοιχα Διαγράμματα 1.11 αυτοσυνδιακυμάνσεων παρουσιάζονται δύο παραδείγματα από 2 ίδιες τυχαία παραγμένες χρονοσειρές μεγέθους $n=200$ με την βοήθεια της γλώσσας R, ορίζοντας το φ για τιμές 0.9 στην πρώτη περίπτωση και 0.1 στην δεύτερη, με σκοπό της σύγκρισή τους. Παρατηρείται ότι στο πρώτο κορελλόγραμμα, όπου η τιμή του φ είναι 0.1, η εμφάνιση του μοιάζει με εικόνα λευκού θορύβου, αφού η τιμή του φ είναι κοντά στο 0. Από την άλλη πλευρά το δεύτερο κορελλόγραμμα έχει εικόνα τυχαίου περιπάτου, αφού η τιμή του φ πλησιάζει στο 0.

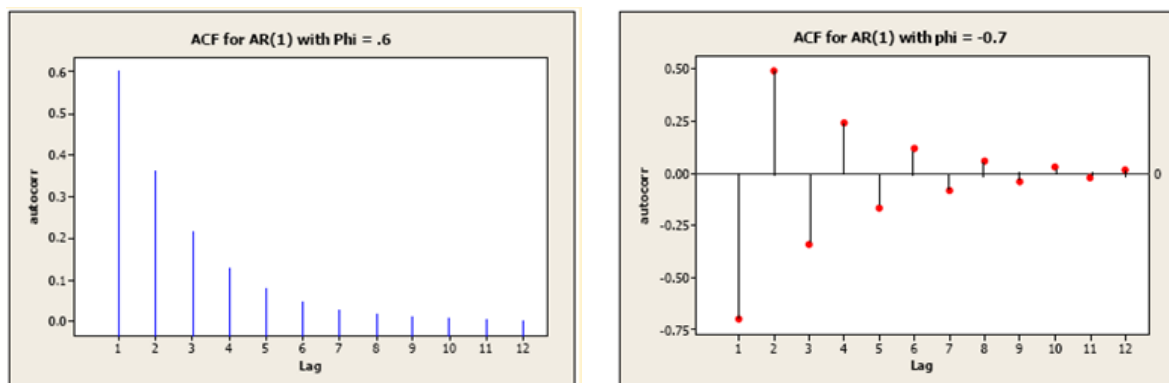


Διάγραμμα 1.10: Διάγραμμα από τυχαία παραγμένες χρονοσειρές με $\varphi=0.9$ και $\varphi=0.1$



Διάγραμμα 1.11: Κορρολογράμματα των τυχαίων χρονοσειρών με $\phi=0.9$ και $\phi=0.1$

Στο ακόλουθο παράδειγμα παρουσιάζονται τα κορρολογράμματα των αυτοσυνδιακυμάνσεων για $\phi=0.6$ και $\phi=-0.7$ στο Διάγραμμα 1.12. Αυτό που παρατηρείται είναι ότι στην πρώτη περίπτωση για θετικό ϕ , οι τιμές των αυτοσυνδιακυμάνσεων φθίνουν προς το 0, ενώ στην δεύτερη περίπτωση για αρνητικό ϕ , οι τιμές των αυτοσυνδιακυμάνσεων εναλλάσσονται γύρω από το 0.



Διάγραμμα 1.12: Διάγραμμα αυτοσυνδιακυμάνσεων για $\phi=0.6$ και $\phi=-0.7$ (PennState, n.d.)

1.9.10.2 Υποδείγματα κινητού Μέσου

Ο κινητός μέσος όρος (MA(q)) ορίζεται ως το σταθμισμένο άθροισμα μεταξύ της τρέχουσας παρατήρησης που μελετάται και των σφαλμάτων σε προηγούμενες χρονικές περιόδους.

Η μαθηματική απεικόνιση του μοντέλου είναι η ακόλουθη:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1.53)$$

όπου θ_i οι συντελεστές μερικής αυτοσυσχέτισης του μοντέλου MA για q χρονικές υστερήσεις, οι οποίοι λαμβάνουν τιμές στο διάστημα $[-1,1]$. Το q απεικονίζει το μήκος της χρονικής υστέρησης, όπως επίσης και τα σφάλματα αποτελούν χρονοσειρά λευκού θορύβου με μέση τιμή 0 και σταθερή διακύμανση σ^2 . Ισχύει ότι: $\varepsilon_t \sim WN(0, \sigma^2)$.

Στην περίπτωση που η μέση τιμή της χρονοσειράς δεν είναι 0, τότε η μαθηματική απεικόνιση του μοντέλου είναι η εξής:

$$X_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1.54)$$

όπου μ η μέση τιμή της χρονοσειράς, η οποία λαμβάνει τιμές σε όλο το σύνολο των πραγματικών αριθμών. Επίσης σε ένα μοντέλο MA(q) η μέθοδος ελαχίστων τετραγώνων εκτιμά τις παραμέτρους $\mu, \theta_1, \theta_2, \dots, \theta_q$ και στη συνέχεια πραγματοποιεί την ελαχιστοποίηση του αθροίσματος των τετραγώνων των σφαλμάτων της προσαρμογής με τον ακόλουθο τρόπο:

$$\text{MinS}(\mu, \theta_1, \theta_2, \dots, \theta_q) = \min \sum_{t=q+1}^n (x_t - \mu + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q})^2 \quad (1.55)$$

Στα μοντέλα αυτά, η εξαρτημένη μεταβλητή δίνεται ως γραμμικός συνδυασμός των ανεξάρτητων μεταβλητών.

Έστω μια στάσιμη χρονοσειρά:

Με την βοήθεια του συντελεστή υστέρησης η χρονοσειρά λαμβάνει την κάτωθι μορφή:

$$X_t = \mu + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t \rightarrow \quad (1.56)$$

$$X_t = \Theta(B) \varepsilon_t$$

όπου $\Theta(B)$ το χαρακτηριστικό πολυώνυμο της χρονοσειράς MA(q).

Οι ρίζες του πολυωνύμου: $\frac{1}{B^q} - \theta_1 \frac{1}{B^{q-1}} - \dots - \theta_q = 0$ θα πρέπει να είναι μικρότερες της μονάδας, ώστε να μη βρίσκονται εντός του μοναδιαίου κύκλου.

Υπόδειγμα MA(1):

Το υπόδειγμα MA(1) είναι το απλούστερο μοντέλο κινητού μέσου MA(q), αφού περιέχει μόνο μια χρονική υστέρηση (q=1) και η μαθηματική του έκφραση είναι:

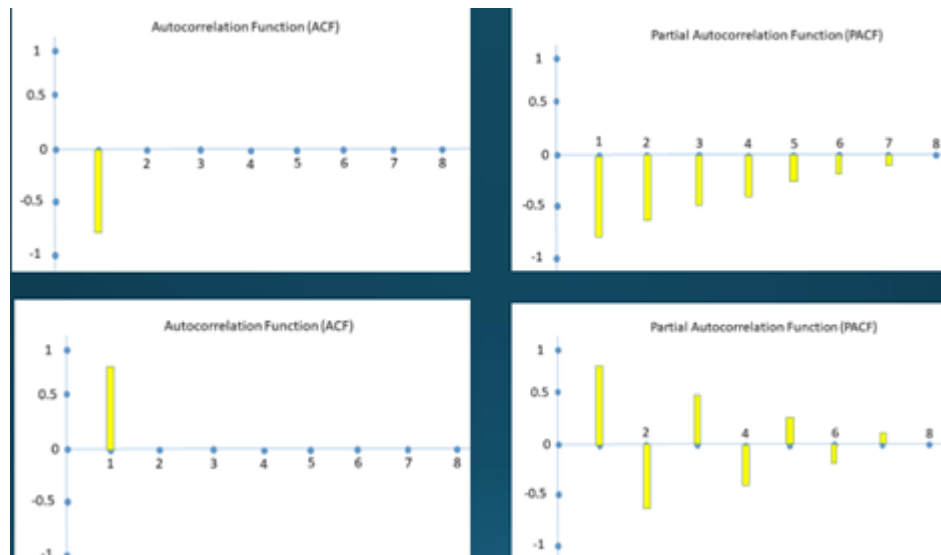
$$X_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (1.57)$$

Αντίστοιχα το υπόδειγμα κινητού μέσου με δύο χρονικές υστερήσεις MA(2) δίνεται από τον παρακάτω τύπο:

$$X_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \quad (1.58)$$

Το πιο αποτελεσματικό βοήθημα για την εύρεση του καλύτερου υποδείγματος είναι το κορελόγραμμα, αφού η απεικόνιση της συνάρτησης αυτοσυσχέτισης (Auto Correlation Function, ACF) γίνεται 0 μετά από q χρονικές υστερήσεις, όπως επίσης και η συνάρτηση μερικής αυτοσυσχέτισης φθίνει εκθετικά προς το μηδέν.

Στο Διάγραμμα 1.13 παρουσιάζεται η γραφική απεικόνιση ενός μοντέλου MA(1) με $\theta_1 = 0,7$



Διάγραμμα 1.13: Γραφική απεικόνιση των αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων του μοντέλου MA(1) με $\theta_1 = 0.7$ (Καρακασίδης και Φράγκου, n.d)

Αυτό που παρατηρείται στο Διάγραμμα 1.13 είναι, πως όσο τα lags απομακρύνονται από την q χρονική υστέρηση, τόσο ο δείκτης ACF τείνει στο μηδέν.

Πίνακας 1.1: Ιδιότητες ACF και PACF στα μοντέλα AR(p) και MA(q)

		Μοντέλα	
		AR(p)	MA(q)
Δείκτες	ACF	Φθίνει εκθετικά μέχρι το 0 ή πρότυπο φθίνουσας ημιτονοειδούς συνάρτησης	Γίνεται απότομα 0 μετά από q χρονικές υστερήσεις
	PACF	Γίνεται απότομα 0 μετά από p χρονικές υστερήσεις	Φθίνει εκθετικά μέχρι το 0 ή πρότυπο φθίνουσας ημιτονοειδούς συνάρτησης

Στον Πίνακα 1.1 συνοψίζονται οι ιδιότητες των συναρτήσεων ACF και PACF για τα μοντέλα . Ειδικότερα από τον πίνακα αυτόν γίνεται κατανοητή η διάκριση, αυτών των δύο υποδειγμάτων μέσω των συναρτήσεων ACF, PACF.

1.9.10.3 Αυτοπαλίνδρομα μοντέλα κινητού μέσου όρου

Τα αυτοπαλίνδρομα μοντέλα κινητού μέσου όρου (Autoregressive Moving Average, ARMA) συνδυάζουν τις ιδιότητες των υποδειγμάτων MA(q) και AR(p), αφού πολλές στάσιμες χρονοσειρές δεν μπορούν να μοντελοποιηθούν σε ένα από τα δύο υποδείγματα, δηλαδή υπάρχουν χρονοσειρές που συνδυάζουν χαρακτηριστικά και των δύο μοντέλων. Έτσι γεννάται ένα νέο υπόδειγμα, το ARMA(p,q), όπου το AR(p) δηλώνει την στασιμότητα και το MA(q) την αντιστρεψιμότητα. Σημαντικό επίσης να τονιστεί, είναι ότι για να ισχύει η στασιμότητα, πρέπει οι ρίζες του πολυωνύμου να βρίσκονται έξω από τον μοναδιαίο κύκλο, όπως και στο μοντέλο AR(p). Αντίστοιχα ισχύει και για το δεύτερο σκέλος του μοντέλου, όπως σε ένα MA(q) υπόδειγμα για να ισχύει η αντιστρεψιμότητα, οι ρίζες του πολυωνύμου πρέπει να βρίσκονται έξω από τον μοναδιαίο κύκλο. Η μαθηματική του ARMA(p,q) έκφραση είναι η ακόλουθη:

$$X_t = \mu + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1.59)$$

όπου $\varphi_i, \theta_i \in [-1, 1]$.

Με την βοήθεια του συντελεστή υστέρησης το υπόδειγμα λαμβάνει την ακόλουθη μορφή:

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p) X_t = \mu + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t \quad (1.60)$$

$$\varphi(B) X_t = \mu + \theta(B) \varepsilon_t$$

Ειδικές Περιπτώσεις ARMA υποδειγμάτων:

- ARMA(p,0) → AR(p)
- ARMA(0,q) → MA(q)

Τα τρία αυτά μοντέλα AR(p), MA(q), ARMA(p,q) λειτουργούν αποτελεσματικά σε στάσιμες χρονοσειρές. Σε πραγματικές συνθήκες είναι πολύ δύσκολο έως απίθανο μια χρονοσειρά να είναι στάσιμη.

Για παράδειγμα η τιμή μιας μετοχής στο χρηματιστήριο επηρεάζεται καθημερινά από ποικίλους παράγοντες και έτσι συνήθως έχει αυξητική ή πτωτική τάση, όπως επίσης οι καθημερινές πωλήσεις ενός καταστήματος για κάποιο χρονικό διάστημα. Στα παραπάνω υποδείγματα, όπως αναφέρθηκε η μέση τιμή και διακύμανση δεν μεταβάλλονται με τον χρόνο και παραμένουν σταθερές, καθώς επίσης και οι αυτοδιακυμάνσεις εξαρτώνται μόνο από την πρώτη χρονική υστέρηση. Είναι πολύ δύσκολο να πραγματοποιηθεί μια πρόβλεψη, αν η χρονοσειρά δεν είναι στάσιμη, αφού είναι πολύ επιρρεπής σε τυχαία γεγονότα που δημιουργούν μεγάλες μεταβολές. Για τον λόγο αυτό δημιουργήθηκαν τεχνικές μετατροπής μιας χρονοσειράς από μη στάσιμη σε στάσιμη με σκοπό οι προβλέψεις να γίνουν περισσότερο ακριβείς.

1.9.11 Μετασχηματισμός μη στάσιμης χρονοσειράς σε στάσιμη

Όπως αναφέρθηκε παραπάνω είναι πολύ δύσκολο σε μια μη στάσιμη χρονοσειρά να γίνουν ασφαλείς προβλέψεις, για τον λόγο ότι η χρονοσειρά αυτή περιέχει τάση, εποχικές διακυμάνσεις κ.λπ.. Έτσι δημιουργήθηκαν μέθοδοι απαλοιφής της τάσης και των εποχικών διακυμάνσεις, όμως όχι της ερμηνείας τους. Σε αυτήν εδώ την ενότητα θα γίνει ανάλυση

χρονοσειρών που περιέχουν μόνο ανοδική ή καθοδική τάση και δεν περιέχουν εποχικές διακυμάνσεις. Η μέθοδος που χρησιμοποιείται για την απαλοιφή της τάσης είναι η μέθοδος της διαφόρισης.

Όταν μια χρονοσειρά μετατρέπεται σε στάσιμη λαμβάνοντας τις πρώτες διαφορές ονομάζεται ολοκληρωμένη πρώτης τάξης και συμβολίζεται με $I(1)$ και υπολογίζονται ως εξής:

$$X'_t = X_t - X_{t-1} \quad (1.61)$$

Με την χρήση του τελεστή ολίσθησης οι πρώτες διαφορές ορίζονται ακολούθως:

$$X'_t = X_t - X_{t-1} = (1-B) X_t \quad (1.62)$$

Πολλές φορές όμως δεν επιτυγχάνεται η στασιμότητα με τις πρώτες διαφορές, οπότε χρειάζεται να ληφθούν και οι δεύτερες διαφορές, οπότε η χρονοσειρά ονομάζεται ολοκληρώσιμη δεύτερης τάξεως, συμβολίζεται με $I(2)$ και υπολογίζονται ως εξής:

$$X''_t = X_t - 2X_{t-1} + X_{t-2} \quad (1.63)$$

Με την χρήση του τελεστή ολίσθησης οι δεύτερες διαφορές ορίζονται ακολούθως:

$$X''_t = X_t - 2X_{t-1} + X_{t-2} = (1+2B+B^2) X_t = (1-B)^2 X_t \quad (1.64)$$

Κάθε φορά που χρησιμοποιείται η μέθοδος της διαφόρισης, χάνεται και μια παρατήρηση ανά φορά χρήσης. Αυτή η μέθοδος μπορεί να γενικευτεί υπολογίζοντας d διαφορές. Η χρονοσειρά ονομάζεται ολοκληρώσιμη d τάξεως και συμβολίζεται με $I(d)$. Συνήθως λαμβάνονται μόνο οι πρώτες δύο διαφορές για την μετατροπή της χρονοσειράς.

1.9.12 Ολοκληρωμένο υπόδειγμα ARIMA

Το μοντέλο $ARIMA(p,d,q)$ είναι το ακρωνύμιο των λέξεων Autoregressive Integrated Moving Average που σημαίνει ότι συνδυάζει τα αυτοπαλίνδρομα υποδείγματα $AR(p)$ και τα υποδείγματα κινητού Μέσου $MA(q)$. Η διαφορά σε σχέση με τα μοντέλα $ARMA(p,q)$ είναι ότι χρησιμοποιεί την μέθοδο της διαφόρισης για να μετατρέψει τις μη στάσιμες χρονοσειρές σε στάσιμες, κάτι το οποίο δεν μπορεί να πραγματοποιήσει το $ARMA(p,q)$ υπόδειγμα, αφού λειτουργεί μόνο για στάσιμες χρονοσειρές. Το μοντέλο $ARIMA$ ορίζεται ως:

$$x_t = x_{t-1} + \varphi_2 x_{t-2} + \dots + \varphi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1.65)$$

Ένα μοντέλο ARIMA(p,d,q) περιέχει τρεις παραμέτρους:

- Η παράμετρος p του αυτοπαλίνδρομου υποδείγματος ARIMA δηλώνει την τάξη του αυτοπαλίνδρομου υποδείγματος
- Η παράμετρος d υποδηλώνει την τάξη της διαφορίσης του υποδείγματος ARIMA (τον αριθμό των διαφορών που απαιτούνται έως ότου φτάσει η χρονοσειρά στη στασιμότητα)
- Η παράμετρος q του υποδείγματος του κινητού μέσου ARIMA δηλώνει την τάξη του κινητού μέσου

Είναι προφανές ότι με την κατάλληλη μετατροπή των I(d) διαφορών, ένα μοντέλο ARIMA(p,d,q) μετατρέπεται σε ένα μοντέλο ARMA(p,q).

Με την χρήση της τεχνικής της διαφορίσης η χρονοσειρά μετατρέπεται σε στάσιμη. Έτσι το υπόδειγμα ARIMA διαθέτει σταθερή μέση τιμή (μ), διακύμανση (σ^2) και συνάρτηση αυτοσυσχέτισης (ACF) καθ' όλη την διάρκεια της χρονικής περιόδου που μελετάται για τον λόγο αυτό δεν εξαρτάται από τις αντίστοιχες χρονικές στιγμές. Τα υποδείγματα ARIMA μπορούν να πραγματοποιήσουν βραχυπρόθεσμες προβλέψεις της χρονικής στιγμής t, αφού χρειάζεται να είναι γνωστές οι: $x_{t-1}, x_{t-2}, \dots, x_{t-n}$. Στη συνέχεια για να γίνει πρόβλεψη της επόμενης χρονικής στιγμής x_{t+1} χρειάζεται να είναι γνωστές οι: $x_{t-1}, x_{t-2}, \dots, x_{t-n-1}$. Στο σημείο αυτό είναι σημαντικό να τονιστεί, πως η τιμή της x_{t+1} αποτελεί τιμή πρόβλεψης, άρα περιέχει το στοιχείο της αβεβαιότητας. Όσο αυξάνεται το χρονικό διάστημα προβλέψεων και χρησιμοποιούνται προβλέψεις για τιμές μελέτης είναι βέβαιο πως τα αποτελέσματα των μακροχρόνιων προβλέψεων θα είναι σε κάποιο βαθμό ανακριβή.

Ειδικές περιπτώσεις ενός μοντέλου ARIMA:

- ARIMA (1,0,0) = AR (1)
- ARIMA (0,0,1) = MA (1)
- ARIMA(0,0,0) = Λευκός Θόρυβος
- ARIMA(0,1,0) = Τυχαίος Περίπατος

1.9.13 Μέθοδος Box & Jenkins

Η μέθοδος Box & Jenkins εισήχθη για πρώτη φορά το 1970. Δημιουργοί της οι δύο στατιστικοί George E.P. Box και Gwilym M. Jenkins. Η μέθοδος αυτή χρησιμοποιεί υποδείγματα ARIMA με σκοπό να βρεθεί το βέλτιστο υπόδειγμα, το οποίο θα κάνει την καλύτερη δυνατή προσαρμογή στις παρελθοντικές τιμές της χρονοσειράς που γίνεται η μελέτη με τελικό στόχο την πιο εύστοχη πρόβλεψη. Εφαρμόζεται σε χρονοσειρές, όπου δεν υπάρχει στασιμότητα, όπως η τιμή μιας μετοχής στο χρηματιστήριο και χωρίζεται σε τέσσερα στάδια:

- Ταυτοποίηση (identification)

Στο πρώτο στάδιο της μεθόδου γίνεται ο προσδιορισμός και η επιλογή των παραμέτρων p, d, q του υποδείματος. Αρχικά γίνεται εύρεση και εφαρμογή των d διαφορών με σκοπό να επιτευχθεί η στασιμότητα (αν δεν υπάρχει ήδη). Αυτό ελέγχεται δια μέσου της συνάρτησης αυτοσυσχετίσης. Αν οι αυτοσυσχετίσεις συγκλίνουν με υψηλό ρυθμό στο 0, τότε υπάρχει μεγάλη πιθανότητα να είναι στάσιμη η χρονοσειρά. Αντίθετα αν φθίνουν με χαμηλό ρυθμό υπάρχει μεγάλη πιθανότητα η χρονοσειρά να μην είναι στάσιμη και να χρειαστεί μετατροπή με την μέθοδο της διαφορίσης. Στη συνέχεια αφού υπάρχει η βεβαιότητα πως η χρονοσειρά είναι στάσιμη, γίνεται ο προσδιορισμός των παραμέτρων p, q βάσει των απλών και των μερικών αυτοσυσχετίσεων.

- Εκτίμηση (Estimation)

Σε συνέχεια του σταδίου της ταυτοποίησης εισέρχεται το στάδιο της εκτίμησης της μεθόδου Box & Jenkins. Σε αυτό το στάδιο σκοπός είναι η εκτίμηση των συντελεστών του υποδείματος ARIMA(p, d, q). Ποιο συγκεκριμένα γίνεται η εκτίμηση των p παραμέτρων $\phi_0, \phi_1, \dots, \phi_p$ για το αυτοπαλίνδρομο υπόδειγμα και αντίστοιχα για το υπόδειγμα κινητού μέσου q παραμέτρων $\theta_0, \theta_1, \dots, \theta_q$. Πολύ σημαντικό σε αυτό το σημείο να αναφερθεί, πως αν η χρονοσειρά που μελετάται περιέχει μόνο την αυτοπαλίνδρομη διαδικασία, τότε για τον υπολογισμό της χρησιμοποιείται η μέθοδος των ελαχίστων τετραγώνων. Αν η χρονοσειρά περιέχει και το υπόδειγμα του κινητού μέσου, τότε η εκτίμηση των παραμέτρων θα γίνει με την χρήση μη γραμμικών μεθόδων, όπως η εκτίμηση μέγιστης πιθανοφάνειας, όπου στόχος είναι να βρεθούν οι τιμές των παραμέτρων που μεγιστοποιούν την συνάρτηση πιθανοφάνειας.

- Διαγνωστικός Έλεγχος

Σε αυτό το στάδιο της μεθόδου εξετάζεται αν το υπόδειγμα ARIMA(p,d,q) που βρέθηκε στα προηγούμενα στάδια είναι ικανοποιητικό και ανταποκρίνεται όσο το δυνατόν περισσότερο στο μοτίβο και την φύση των δεδομένων (fitting). Με σκοπό να γίνει ο πλήρης έλεγχος για την καταλληλότητα και την αποδοτικότητα του μοντέλου χρησιμοποιούνται τέσσερις πυλώνες οι οποίοι είναι:

- Οι ιδιότητες των καταλοίπων (residuals)
- Η σταθερότητα και σημαντικότητα των συντελεστών του υποδείγματος
- Η ικανότητα του μοντέλου να προβλέπει με ακρίβεια

Εφόσον το υπόδειγμα που μελετάται θεωρηθεί αξιόπιστο και κατάλληλο πρέπει ελεγχθεί, ότι τα κατάλοιπα συμπεριφέρονται ως μια διαδικασία λευκού θορύβου.

Αυτό διαπιστώνεται με τον έλεγχο των αυτοσυσχετίσεων των καταλοίπων, όπου εκεί πρέπει να αποφανθεί ότι δεν υπάρχουν αυτοσυσχετίσεις μεταξύ τους. Ο παραπάνω έλεγχος καταλοίπων ακολουθείται η στατιστική συνάρτηση Q των Box & Pierce, όπου με την χρήση της γίνεται ο έλεγχος σημαντικότητας n πλήθους συντελεστών αυτοσυσχέτισης.

Η στατιστική Συνάρτηση δίνεται από την έκφραση:

$$Q_{BP} = T \sum_{s=1}^n \hat{\rho}_s^2 \quad (1.66)$$

όπου:

$\hat{\rho}_s$: οι δειγματικές αυτοσυσχετίσεις των καταλοίπων

T: το πλήθος των συντελεστών αυτοσυσχέτισης

n: Το πλήθος των συντελεστών αυτοσυσχέτισης

Ο αριθμός των αυτοσυσχετίσεων ισούται με την τετραγωνική ρίζα του αριθμού των παρατηρήσεων, δηλαδή ισχύει $n=\sqrt{T}$. Η στατιστική συνάρτηση QBP ακολουθεί την X^2 κατανομή με n-p-q βαθμούς ελευθερίας.

Ο έλεγχος υπόθεσης της στατιστικής συνάρτησης QBP είναι ο ακόλουθος:

$$H_0 : Q_{BP} > X_a^2 \quad (1.67)$$

$$H_1 : Q_{BP} \leq X_a^2 \quad (1.68)$$

Η αρχική υπόθεση απορρίπτεται, όταν η τιμή της στατιστικής συνάρτησης QBP είναι μεγαλύτερη από την κρίσιμη τιμή της κατανομής X_a^2 για επίπεδο σημαντικότητας α . Αντίστοιχα απορρίπτεται η εναλλακτική υπόθεση, όταν η τιμή της στατιστικής συνάρτησης QBP είναι μικρότερη της κρίσιμης τιμής κατανομής X_a^2 για επίπεδο σημαντικότητας α .

Το κριτήριο που χρησιμοποιείται ευρέως στην μέθοδο Box – Jenkins είναι η στατιστική Q των Ljung & Box η οποία απεικονίζεται παρακάτω:

$$Q_{LB} = n(n+2) \sum_{j=1}^k \frac{\rho_j^2}{n-j} \quad (1.69)$$

Η στατιστική συνάρτηση αυτή ακολουθεί προσεγγιστικά την κατανομή X_a^2 με $n-p-q$ βαθμούς ελευθερίας. Όσο αναφορά τον έλεγχο υπόθεσης είναι ο ίδιος με της στατιστικής συνάρτησης των Box & Pierce, δηλαδή η Αρχική υπόθεση H_0 απορρίπτεται όταν απορρίπτεται όταν η τιμή της στατιστικής συνάρτησης Q_{LB} είναι μεγαλύτερη από την κρίσιμη τιμή της κατανομής X_a^2 για επίπεδο σημαντικότητας α . Αντίστοιχα απορρίπτεται η εναλλακτική υπόθεση, όταν η τιμή της στατιστικής συνάρτησης Q_{LB} είναι μικρότερη της κρίσιμης τιμής κατανομής X_a^2 για επίπεδο σημαντικότητας α .

1.9.14 Κριτήρια επιλογής υποδείγματος

Σε αυτό το στάδιο του ελέγχου θα γίνει σχολιασμός της καταλληλότητας του εκτιμώμενου υποδείγματος μέσω της διαδικασίας υπερπροσαρμογής (overfitting). Ο τρόπος ελέγχου της διαδικασίας αυτής είναι η σύγκριση του εκτιμώμενου υποδείγματος με ένα υπόδειγμα υψηλότερης τάξης. Για παράδειγμα το εκτιμηθέν υπόδειγμα είναι το: ARMA(1,2).

Το συγκεκριμένο υπόδειγμα πρέπει να συγκριθεί με τα μικτά υποδείγματα ARMA(2,2) και ARMA(1,3). Αν εξαχθεί το συμπέρασμα πως οι πρόσθετοι όροι δεν είναι στατιστικά σημαντικοί, τότε το εκτιμηθέν μοντέλο είναι το καταλληλότερο. Στην αντίπερα όχθη αν οι

πρόσθετοι όροι βγουν στατιστικά σημαντικοί, τότε αποδεικνύεται πως το εκτιμηθέν υπόδειγμα δεν είναι το καταλληλότερο, αλλά υπάρχουν καλύτερα υποδείγματα που προσαρμόζονται καλύτερα στα δεδομένα. Ένα σημαντικό στοιχείο που πρέπει να γίνει αντιληπτό είναι, πώς για κάθε παραπάνω παράμετρο που προστίθεται στο υπόδειγμα αντίστοιχα γίνεται μείωση των βαθμών ελευθερίας.

Για την εύρεση του κατάλληλου υποδείματος υπάρχουν κριτήρια, τα οποία είναι χρήσιμα και βοηθούν στην επιλογή του πιο κατάλληλου υποδείματος:

α) Το πληροφοριακό κριτήριο του Akaike (Akaike Information Criterion – AIC)

$$AIC = -2 \cdot \ln \hat{\sigma}^2 + \frac{2k}{T} \quad (1.70)$$

όπου:

- Το σ^2 δηλώνει την εκτίμηση της διακύμανσης του σφάλματος του υποδείματος
- Το k δηλώνει τον αριθμό των παραμέτρων του μοντέλου.
- Το T δηλώνει το μέγεθος του δείματος.

Ο δεύτερος όρος αφορά την «ποινή» του πληροφοριακού κριτηρίου (Τσιμπάνος, 2005).

Η εξίσωση του AIC δηλώνει πως όσες περισσότερες παραμέτρους υπάρχουν στο υπόδειγμα, τόση πληροφορία εξάγεται. Παράλληλα το υπόδειγμα που διαθέτει την μικρότερη τιμή στο πληροφοριακό κριτήριο του AIC είναι και αυτό που είναι το καταλληλότερο.

β) Το πληροφοριακό κριτήριο του Bayes (Bayesian Information Criterion -BIC)

$$BIC = \ln \hat{\sigma}^2 + \frac{k \ln T}{T} \quad (1.71)$$

όπου:

- Το σ^2 δηλώνει την εκτίμηση της διακύμανσης του σφάλματος του υποδείματος
- Το k δηλώνει τον αριθμό των παραμέτρων του μοντέλου.
- Το T δηλώνει το μέγεθος του δείματος.
- Ο δεύτερος όρος αφορά την «ποινή» του πληροφοριακού κριτηρίου, όπως και στο AIC.

Στη σύγκριση των υποδειγμάτων, καλύτερο αναδεικνύεται εκείνο που έχει τη μικρότερη τιμή στο BIC.

Η βασική διαφορά μεταξύ AIC και BIC είναι ότι η εισαγωγή επιπρόσθετων παραμέτρων στο BIC αποθαρρύνεται σε σχέση με το AIC, αφού το κριτήριο BIC επιβάλλει μεγαλύτερη ποινή στις πρόσθετες εκτιμώμενες παράμετροι. Η επιλογή του υποδείγματος BIC θα έχει λιγότερες ή στην ακραία περίπτωση ίσες παραμέτρους σε σχέση με τις αντίστοιχες παραμέτρους του AIC.

Συνεπώς το κριτήριο του AIC είναι πιο ορθό να επιλέγεται σε υποδείγματα με μεγάλο πλήθος παραμέτρους, ενώ το κριτήριο του BIC σε μικρά.

Τα μοντέλα ARIMA είναι αποτελεσματικά στην πρόβλεψη τους σε χρονοσειρές που περιέχουν αποκλειστικά τάση. Όμως υπάρχουν χρονοσειρές που περιέχουν εποχικότητα και έτσι τα υποδείγματα ARIMA δεν επαρκούν ώστε η πρόβλεψη να είναι εύστοχη. Γι αυτό το λόγω δημιουργήθηκαν τα υποδείγματα SARIMA.

1.10 Υποδείγματα SARIMA

Τα υποδείγματα SARIMA(p,d,q)(P,D,Q) αποτελούν μια επέκταση των υποδειγμάτων ARIMA(p,d,q), αφού στην δημιουργία των προβλέψεων τους λαμβάνουν υπόψιν και τις εποχικές διακυμάνσεις. Έτσι ένα μοντέλο SARIMA αποτελείται από τις μη εποχικές παραμέτρους p,d,q και της εποχικές (P,D,Q)s.

όπου :

- p και εποχικό P: υποδεικνύουν τον αριθμό των αυτοπαλίνδρομων όρων
- q και εποχικό Q: υποδεικνύουν τον αριθμό των όρων κινητού μέσου όρου
- s: υποδεικνύει τον αριθμό των περιόδων για μια εποχή
- d: υποδεικνύει την τάξη της διαφορίσης του υποδείγματος

Στα μοντέλα ARIMA η εφαρμογή της μεθόδου διαφορίσης για την εύρεση της τάξης του υποδείγματος ήταν το πρώτο βήμα εφαρμογής τους. Αντίστοιχα και στα εποχικά μοντέλα SARIMA το πρώτο βήμα εφαρμογής τους είναι ο καθορισμός της εποχικής διαφοράς. Να

σημειωθεί ότι είναι σημαντικό να μην γίνεται χρήση πάνω από μια εποχικές διαφορές, ούτε πάνω από δύο διαφορίσεις (είτε εποχικές διαφορές, είτε διαφορές στην τάση). Σε ένα εποχικό υπόδειγμα οι όροι SAR και SMA προσδιορίζονται από τον έλεγχο αυτοσυσχετίσεων. Ο όρος SAR προστίθεται στο μοντέλο, αν για την εποχική περίοδο που μελετάται η αυτοσυσχέτιση είναι θετική, ενώ αν είναι αρνητική προθέεται ένας όρος SMA. Συνήθως ένας όρος SAR και ένας όρος SMA είναι αρκετοί να περιγράψουν επαρκώς το μοντέλο και έτσι σπάνια συναντιούνται περισσότεροι όροι στο υπόδειγμα.

Το υπόδειγμα SARIMA που έχει χρησιμοποιείται συνήθως είναι το ARIMA(0,1,1)_x(0,1,1), όπου είναι ένας συνδυασμός ενός MA(1) x SMA(1), το οποίο συνδυάζει μια εποχική και μια μη εποχική διαφορά (Παναγοπούλου,2020; Πανάνα, 2018-2019)

Με την βοήθεια του συντελεστή υστέρησης το μοντέλο SARIMA:

$$\varphi(B)(1-B)^d x_t = \theta(B)\varepsilon_t \quad (1.72)$$

$$\varphi(B^s)(1-B^s)^D x_t = \theta(B^s)\varepsilon_t \quad (1.73)$$

όπου:

- $\varphi(B), \theta(B)$ τα χαρακτηριστικά πολυώνυμα για το υπόδειγμα με την τάση
- $\varphi(B^s), \theta(B^s)$ τα χαρακτηριστικά πολυώνυμα για το υπόδειγμα με την εποχικότητα.

Το d, D αφορούν την τάξη διαφόρισης του μοντέλου της τάσης και της εποχικότητας αντίστοιχα.

Το συνδυαστικό μοντέλο SARIMA(p,d,q)(P,D,Q)_s ορίζεται ως:

$$\Phi(B)[\varphi(B^s)(1-B^s)^D](1-B)^d x_t = \theta(B)\Theta(B^s)\varepsilon_t \quad (1.74)$$

1.11 Σφάλμα πρόβλεψης

Ως σφάλμα (e_t) πρόβλεψης ορίζεται το αποτέλεσμα της διαφοράς μεταξύ παρατηρούμενης τιμής (Y_t) και προβλεπόμενης τιμής (\hat{Y}_t) της χρονοσειράς και δίνεται από τον τύπο:

$$e_t = Y_t - \hat{Y}_t \quad (1.75)$$

για n περιόδους ο τύπος λαμβάνει την μορφή:

$$e_{t+n} = Y_{t+n} - \hat{Y}_{t+n} \quad (1.76)$$

Για την ορθή αξιολόγηση της πρόβλεψης, ένα σημαντικό μέτρο αποτελεί το $(1-\alpha)100\%$ διάστημα εμπιστοσύνης της διακύμανσης του σφάλματος, το οποίο δείχνει την ακρίβεια της πρόβλεψης και δίνεται από τον ακόλουθο τύπο:

$$\hat{Y}_{t+n} - z_{\alpha/2} \sigma_k \leq Y_{t+k} \leq \hat{Y}_{t+n} + z_{\alpha/2} \sigma_n \quad (1.77)$$

όπου σ_k^2 η διακύμανση του σφάλματος πρόβλεψης. Η διακύμανση αυτή είναι συνάρτηση διακύμανσης του διαταρακτικού όρου ε_t η οποία είναι άγνωστη. Έτσι γίνεται εκτίμηση της μέσω του αθροίσματος των τετραγώνων των καταλοίπων:

$$\hat{\sigma}^2 = \frac{\sum \varepsilon_t^2}{T-p-q} \quad (1.78)$$

Συνήθως η τιμή που λαμβάνει το $(1-\alpha)100\%$ διάστημα εμπιστοσύνης είναι 95%.

1.11.1 Μέσο τετραγωνικό σφάλμα

Το μέσο τετραγωνικό σφάλμα είναι μια κοινώς αποδεκτό κριτήριο για πολλές μεθόδους πρόβλεψης, όπως αυτής της εκθετικής εξομάλυνσης και ορίζεται ως το άθροισμα των τετραγώνων των σφαλμάτων διαιρεμένο με το πλήθος του αριθμού των προβλέψεων και η μαθηματική του έκφραση είναι η ακόλουθη:

$$MSE = \frac{1}{n} \sum_{t=1}^n \varepsilon_t^2 \quad (1.79)$$

Το Μέσο τετραγωνικό σφάλμα, λόγω του τετραγώνου που υπάρχει στον υπολογισμό επηρεάζεται περισσότερο από μεγάλα σφάλματα. Παράλληλα το MSE χρησιμοποιείται στην εύρεση των κατάλληλων παραμέτρων στις μεθόδους εξομάλυνσης, όπου επιλέγονται αυτοί που το ελαχιστοποιούν.

1.11.2 Τετραγωνική ρίζα μέσου τετραγωνικού σφάλματος

Ορίζεται ως η τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος (RMSE).

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{n} \sum_{t=1}^n \varepsilon_t^2} \quad (1.80)$$

1.11.3 Μέση Απόκλιση

Η μέθοδος αυτή βασίζεται στο γεγονός, ότι η σοβαρότητα του σφάλματος σχετίζεται γραμμικά με το μέγεθός του και ορίζεται ως το αποτέλεσμα της διαίρεσης του αθροίσματος των απόλυτων τιμών των σφαλμάτων με τον αριθμό των προβλέψεων και δίνεται από τον κάτωθι τύπο:

$$\text{MAD} = \frac{1}{n} \sum_{t=1}^n |\varepsilon_t| \quad (1.81)$$

Η μέθοδος MAD αποτελεί λιγότερο αξιόπιστη επιλογή από το μέσο τετραγωνικό σφάλμα, λόγω του ότι το μέσο τετραγωνικό σφάλμα είναι στατιστικά πιο θεμελιώδης δείκτης.

1.11.4 Μέσο απόλυτο ποσοστιαίο σφάλμα

Ορίζεται ως:

$$\text{MPE} = \frac{1}{n} \sum_{t=1}^n \frac{\varepsilon_t}{Y_t} \quad (1.82)$$

όπου αν:

- $\text{MPE} > 0$ τότε η μέθοδος αυτή υποεκτιμά την πρόβλεψη, αφού δίνει μεροληπτικά μεγάλες τιμές.
- $\text{MPE} < 0$ τότε η μέθοδος αυτή υπεκτιμά την πρόβλεψη, αφού δίνει μεροληπτικά μικρές τιμές.
- MPE κοντά στο 0, τότε η μέθοδος είναι αμερόληπτη, ως προς τις τιμές που δίνει.

1.12 Επιλογή κατάλληλου μοντέλου πρόβλεψης

Κάθε χρονοσειρά είναι μοναδική και περιέχει τα δικά της χαρακτηριστικά. Έτσι χρειάζεται να γίνει ορθή αναγνώριση αυτών των χαρακτηριστικών, ώστε να χρησιμοποιηθεί το κατάλληλο μοντέλο πρόβλεψης: Σημαντικό δεν είναι να γίνει χρήση παραπάνω από ένα μοντέλων

πρόβλεψης, για να μπορέσει να διαπιστωθεί ποιο δίνει πιο ορθά αποτελέσματα (Παγουρτζή, 2021).

Με μια σύνοψη των χαρακτηριστικών των χρονοσειρών οι μέθοδοι πρόβλεψης μπορούν να ταξινομηθούν ακολούθως:

- Στάσιμες χρονοσειρές: Μοντέλα ARMA και Απλός κινητός μέσος
- Χρονοσειρές με τάση: Μέθοδος HOLT και μοντέλα ARIMA
- Χρονοσειρές με τάση και εποχικότητα: Κλασική διάσπαση (decomposition) ή Μέθοδος Winters ή μοντέλα SARIMA
- Χρονοσειρές με κυκλικότητα: Κλασική διάσπαση (decomposition) ή Μοντέλα ARIMA

1.13 Ανακεφαλαίωση

Η μελέτη και η πρόβλεψη χρονοσειρών ήταν ανέκαθεν ένας πολύ σημαντικός τομέας έρευνας σε πολλούς επιστημονικούς κλάδους, όπως η ιατρική, η φυσική, τα οικονομικά, η μετεωρολογία κ.λπ. Στο κεφάλαιο αυτό έγινε μια προσπάθεια να δοθεί ο ορισμός των χρονοσειρών, όπως επίσης και οι έννοιες σχετικά με την διάσπαση και ανάλυση των χρονοσειρών.

Να γίνει κατανοητή η διάκριση μεταξύ της στάσιμης και μη στάσιμης χρονοσειράς, καθώς και να δοθούν διάφοροι μέθοδοι πρόβλεψης ανάλογα την φύση που έχει- κάθε χρονοσειρά, αφού η κάθε μία χρήζει ιδιαίτερης μεταχείρισης, λόγω της μοναδικότητας της και της επιρροής της από τον χρόνο.

Τα τελευταία χρόνια η τεράστια αύξηση των διαθέσιμων δεδομένων, καθώς και η συνεχόμενη άνοδος της τεχνολογίας έδωσε κίνητρο στην παγκόσμια επιστημονική κοινότητα να αναπτύξει νέους αλγορίθμους και μεθόδους πρόβλεψης που θα μπορέσει να καλύψει τις ανάγκες της παγκόσμιας αγοράς. Έτσι τα τελευταία χρόνια η Μηχανική μάθηση (ML) και τα Νευρωνικά Δίκτυα (NN) έχουν γίνει βασικός πυλώνας της νέας γενιάς μοντέλων για την Ανάλυση και πρόβλεψη Χρονοσειρών, όπου επιτυγχάνονται εξαιρετικά αποτελέσματα. Στο επόμενο κεφάλαιο θα γίνει εκτενής αναφορά σε αλγορίθμους πρόβλεψης χρονοσειρών με την χρήση μηχανικής μάθησης.

ΚΕΦΑΛΑΙΟ 2: Μηχανική Μάθηση – Νευρωνικά Δίκτυα

2.1 Εισαγωγή

Η μηχανική μάθηση (machine learning) αποτελεί κλάδος της τεχνητής νοημοσύνης (AI), όπου βασική του χρήση είναι η δημιουργία και η ανάπτυξη μοντέλων και αλγορίθμων που έχουν στόχο την δημιουργία μιας πρόβλεψης ή μιας απόφασης. Αυτό γίνεται με την χρήση δεδομένων, αφού οι αλγόριθμοι και τα μοντέλα που χρησιμοποιούνται, εκπαιδεύονται μέσα από αυτά, με τελικό στόχο να εξάγουν νέα ασφαλή δεδομένα για την δημιουργία της πρόβλεψης ή της απόφασης. Το εύρος των χρήσεων της είναι μεγάλο, αφού χρησιμοποιείται σε πολλούς τομείς, όπως η αναγνώριση προτύπων, η υπολογιστική όραση, επεξεργασία φωνής και εικόνας και η πρόβλεψη διαφόρων τιμών, όπως η πρόβλεψη αποδόσεων μετοχών. Υπάρχουν πολλοί ορισμοί για την μηχανική μάθηση, οι οποίοι παρόλο το κοινό τους χαρακτηριστικό είναι πολλές φορές διαφορετικοί, αφού οι άνθρωποι που τους δημιούργησαν προέρχονταν από διαφορετικούς εκπαιδευτικούς τομείς. Ο πρώτος ορισμός μηχανικής μάθησης, δόθηκε από τον Αμερικανό Athrur Samuel (1959), που θεωρείται ότι ήταν ένας από τους πρωτοπόρους στο πεδίο της τεχνητής νοημοσύνης, έδωσε τον ακόλουθο ορισμό «πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί». Επίσης ο Carbonelli (1897), έδωσε τον ακόλουθο ορισμό μηχανικής μάθησης: «...η μελέτη υπολογιστικών μεθόδων για την απόκτηση νέας γνώσης, νέων δεξιοτήτων και νέων τρόπων οργάνωσης της υπάρχουσας γνώσης». Μετά από δέκα χρόνια ο Mitchell (1997) υποστήριξε πως «ένα πρόγραμμα H/Y θεωρείται ότι μαθαίνει μέσω εμπειρίας E που αποκτά κάνοντας κάποιες δραστηριότητες T και σε συνδυασμό με κάποια μετρική απόδοσης P , αν οι επιδόσεις του στις δραστηριότητες T , όπως καταμετρώνται από την P , βελτιώνονται με την εμπειρία E ».

Η φιλοσοφία της μηχανικής μάθησης και των νευρωνικών δικτύων διαφέρει κατά πολύ σε σχέση με τον «κοινό προγραμματισμό», αφού δεν είναι υπεύθυνος ο χρήστης (προγραμματιστής) να προγραμματίσει τον υπολογιστή, ώστε να εξάγει μια πληροφορία. Αντίθετα κάνει χρήση δεδομένων και αλγορίθμων, με σκοπό ο υπολογιστής να διεξάγει τη διαδικασία της εκπαίδευσης μόνος του και να «μάθει» μέσα από τα δεδομένα χωρίς την παρέμβαση του χρήστη και στην συνέχεια να εξάγει την επιθυμητή πληροφορία.

Τα τελευταία χρόνια η βαθιά μάθηση, η οποία αποτελεί υποσύνολο της μηχανικής μάθησης έχει στρέψει την προσοχή και αποτελεί πολύ σημαντικό εργαλείο σε πολλούς επιστημονικούς κλάδους, λόγω της μεγάλης ακρίβειας των αλγορίθμων της σε μεγάλα σύνολα δεδομένων. Παράλληλα οι αλγόριθμοι της είναι πολύ ευέλικτοι σε πολλούς και διαφορετικούς τύπους προβλημάτων, αφού βρίσκει εφαρμογές στην βιολογία, την ιατρική, τα οικονομικά, την ρομποτική κ.λπ.. Ακόμα ένα βασικό πλεονέκτημα της βαθιάς μάθησης είναι η συνεχόμενη εξέλιξη των αλγορίθμων της και εύρεση νέων, οι οποίοι καταφέρνουν να είναι εξαιρετικά αποτελεσματικοί, μειώνοντας το σφάλμα στο ελάχιστο.

2.2 Μηχανική Μάθηση

2.2.1 Κατηγορίες μηχανικής μάθησης

Η μηχανική μάθηση χωρίζεται σε τέσσερις βασικές κατηγορίες:

Η πρώτη κατηγορία μηχανικής μάθησης είναι η επιβλεπόμενη μάθηση (Supervised Learning). Η επιβλεπόμενη μάθηση (Supervised learning) ή αλλιώς μάθηση με παραδείγματα χρησιμοποιεί ένα σύνολο δεδομένων της μορφής (x_i, y_i) , $i=1,2,\dots,n$, όπου κάθε x_i είναι ένα d -διάστατο διάνυσμα εισόδου και y_i η αντίστοιχη έξοδος. Ανάλογα με το είδος της μεταβλητής εξόδου, η επιβλεπόμενη μάθηση χωρίζεται σε δύο κατηγορίες:

- Αν οι κλάσεις της μεταβλητής εξόδου y είναι διακριτές, τότε αποτελεί πρόβλημα κατηγοριοποίησης
- Αν οι κλάσεις της μεταβλητής εξόδου y είναι συνεχείς, τότε αποτελεί πρόβλημα παλινδρόμησης

Δύο γνωστά παραδείγματα επιβλεπόμενης μάθησης αποτελούν:

α) η ανάλυση συναισθήματος πελατών (Customer Sentiment Analysis)

β) η ανίχνευση Spams στα Emails, δηλαδή η εύρεση της ανεπιθύμητης αλληλογραφίας και ο διαχωρισμός της από την επιθυμητή.

Η δεύτερη κατηγορία μηχανικής μάθησης είναι η μη επιβλεπόμενη μάθηση (Unsupervised Learning). Η μη επιβλεπόμενη Μάθηση δεν χρησιμοποιεί δεδομένα με ετικέτα (labeled data),

αφού προσπαθεί να εντοπίσει ένα μοτίβο στα δεδομένα με σκοπό την ομαδοποίηση τους, ανιχνεύοντας όμοια ή διαφορετικά χαρακτηριστικά. Κάποια παραδείγματα μη επιβλεπόμενης μάθησης αποτελούν:

- Συσταδοποίηση (clustering)
- Μείωση Διαστάσεων (Dimensionality Reduction)
- Αναγνώριση Προτύπων (Pattern Recognition)
- Ανάλυση Συσχετίσεων

Η τρίτη κατηγορία μηχανικής μάθησης είναι η ενισχυτική μάθηση (reinforcement learning). Σε αυτή την κατηγορία μάθησης, ο αλγόριθμος που εκπαιδεύεται έχει αλληλεπίδραση με το περιβάλλον, αφού για κάθε κίνηση που πραγματοποιεί, λαμβάνει μια ανταμοιβή ή μια ποινή. Ανάλογα το είδος της, καταλαβαίνει αν η κίνηση που πραγματοποίησε ήταν ορθή ή λανθασμένη. Στόχος είναι να μπορέσει ο αλγόριθμος να κατανοήσει την ακολουθία που οδηγεί σε σωστά αποτελέσματα, ώστε να τα εντοπίζει με μεγάλη ακρίβεια. Χρησιμοποιείται κυρίως σε προβλήματα σχεδιασμού (planning) Η ενισχυτική μάθηση δίνει έμφαση στην κατανόηση του εκάστοτε προβλήματος και όχι στην εκμάθηση μεθόδων επίλυσης του. Παράλληλα έχει μεγάλο πεδίο εφαρμογών σε σύγχρονα πεδία, όπως τα αυτόνομα αυτοκίνητα (self - driving cars), την επεξεργασία φυσικής γλώσσας (Natural Language processing, NLP), τον έλεγχο της κυκλοφορίας (Traffic Control), όπως επίσης και στο χειρισμό ρομποτικής (robotics manipulation).

Η τέταρτη κατηγορία μηχανικής μάθησης είναι η βαθιά μάθηση (Deep Learning). Η Βαθιά μάθηση αποτελεί μια εξειδικευμένη μορφή μηχανικής μάθησης, η οποία προσπαθεί μέσω των αλγορίθμων της να καταφέρει να κάνει τον υπολογιστή να πράττει, όπως τον ανθρώπινο εγκέφαλο. Στη βαθιά μάθηση πραγματοποιείται χρήση νευρωνικών δικτύων με πολλαπλά επίπεδα. Η χρήσης της βαθιάς μάθησης έχει αυξηθεί ραγδαία τα τελευταία χρόνια, αφού πλέον τα διαθέσιμα δεδομένα έχουν αυξηθεί και αυτά άρδην. Κάποια παραδείγματα που βρίσκει εφαρμογή είναι: στο χρηματιστήριο για την πρόβλεψη μετοχών, στην ρομποτική, η αναγνώριση εικόνας και ομιλίας, όπως επίσης και η επεξεργασία φυσικής γλώσσας.

2.2.2 Διαφορές μεταξύ βαθιάς μάθησης και μηχανικής μάθησης

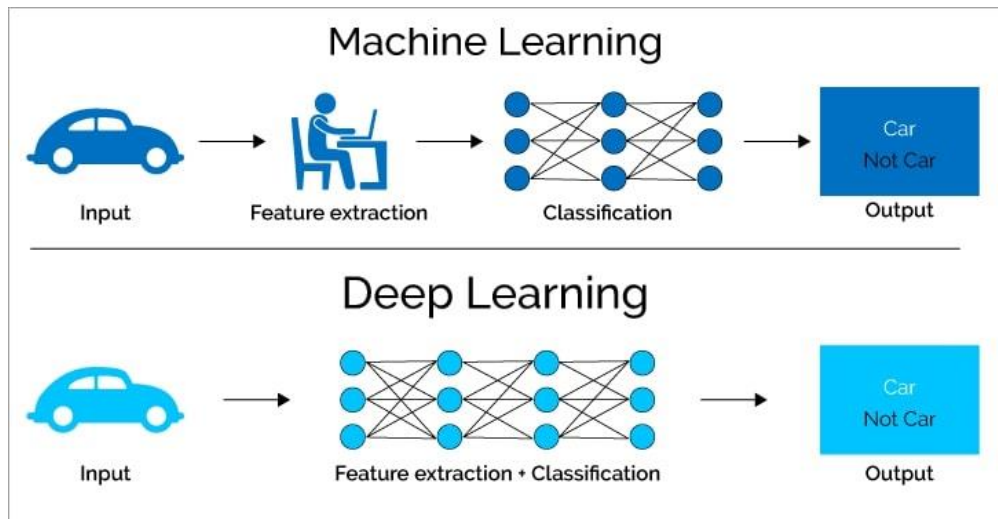
Μηχανική Μάθηση:

- Χρήση μικρότερου όγκου δεδομένων
- Αυτόματη ταξινόμηση
- Εξαγωγή χαρακτηριστικών από τον άνθρωπο

Βαθιά Μάθηση:

- Οι νευρώνες εκπαιδεύονται από τα δεδομένα
- Το μοντέλο «μαθαίνει» από τα σφάλματά του
- Πλήρης αυτόματη διαδικασία
- Απαιτείται πολύ μεγάλος όγκος δεδομένων

Στο Διάγραμμα 2.1 αναπαρίσταται η διαφορά μεταξύ της μηχανικής μάθησης και της βαθιάς μάθησης σε ένα πρόβλημα ταξινόμησης, για το αν τα δεδομένα που εισήλθαν αφορούν «αυτοκίνητο» ή όχι. Στην πρώτη περίπτωση, δηλαδή της μηχανικής μάθησης, τα δεδομένα εισέρχονται στον υπολογιστή, γίνεται εξαγωγή των χαρακτηριστικών από τον άνθρωπο και στην συνέχεια πραγματοποιείται η ταξινόμηση, ώστε να ληφθεί η απόφαση αν τα δεδομένα που εισήλθαν ανήκουν στην παραπάνω κατηγορία. Στην δεύτερη περίπτωση της βαθιάς μάθησης υπάρχει αυτοματοποίηση ως προς την εξαγωγή των χαρακτηριστικών, όπου δεν χρειάζεται ανθρώπινη παρέμβαση. Στη συνέχεια πραγματοποιείται η ταξινόμηση για την λήψη απόφασης, με απώτερο σκοπό την εξαγωγή συμπερασμάτων, για το αν τα δεδομένα αφορούν «αυτοκίνητο» (Χασάπη, 2022).



Διάγραμμα 2.1 Γραφική απεικόνιση των διαφορών μεταξύ μηχανικής μάθησης και βαθιάς μάθησης (Χασάπη, n.d.)

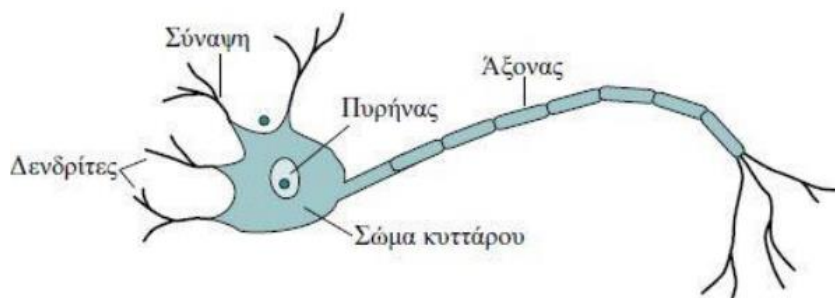
2.3 Νευρωνικά Δίκτυα

2.3.1 Εισαγωγή στα νευρωνικά δίκτυα

Το μέρος αυτού του κεφαλαίου στρέφεται στην παρουσίαση των σχετικών πληροφοριών για τα νευρωνικά δίκτυα. Τα τεχνητά νευρωνικά δίκτυα που θα παρουσιαστούν στην παρούσα διπλωματική είναι τα Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης (FeedForward Neural Networks), τα Παλινδρομικά Νευρωνικά δίκτυα (Recurrent Neural Networks), τα Δίκτυα Μακράς Βραχύχρονης Μνήμης (LSTM) καθώς και τα Φραγμένα Ανατροφοδοτούμενα Δίκτυα (GRU). Αρχικά θα γίνει μια μικρή αναφορά στα βιολογικά νευρωνικά δίκτυα και ακολούθως θα παρουσιαστούν οι τοπολογίες των παραπάνω δικτύων. Επιπροσθέτως θα γίνει ανάλυση της σε θεωρητικό επίπεδο της συνάρτησης απώλειας (loss function), όπως επίσης και οι τρόποι εκπαίδευσης των δικτύων που βοηθούν στην αποτελεσματικότητα των μοντέλων.

Νευρών ονομάζεται το νευρικό κύτταρο με όλες τις αποφυάδες του. Οι νευρώνες είναι κύτταρα χαρακτηριζόμενα από τη διεγερσιμότητα και εξειδικευμένα να παραλαμβάνουν ερεθίσματα και να άγουν την νευρική ώση. Ως το προς το μέγεθος και το διάγραμμα διαφέρουν σημαντικά, αλλά όλα έχουν κυτταρικό σώμα, από την επιφάνεια του οποίου εκφύονται οι νευρίτες. Οι νευρίτες, οι οποίοι παραλαμβάνουν την πληροφορία και την άγουν προς το

κυτταρικό σώμα, ονομάζονται δεντρίτες. Στο Διάγραμμα 2.2 παρουσιάζεται ο φυσικός νευρώνας (Snell, 2008).



Διάγραμμα 2.2: Φυσικός νευρώνας (Κοντόπουλος, 2016)

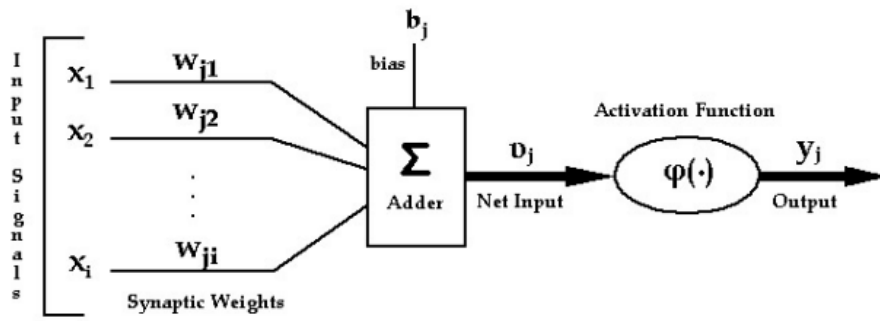
Τα τεχνητά νευρωνικά δίκτυα (ΤΝΔ) είναι ένα κλάδος της μηχανικής μάθησης, που ο σκοπός δημιουργίας και χρήσης τους είναι να διδάσκουν στους υπολογιστές να μπορούν να κάνουν επεξεργασία των δεδομένων εμπνευσμένα από την ανθρώπινη βιολογία και τον τρόπο με τον οποίο οι νευρώνες του ανθρώπινου εγκεφάλου λειτουργούν από κοινού για να κατανοήσουν τις εισροές από τις ανθρώπινες αισθήσεις. Υπάρχουν όμως διαφορές μεταξύ βιολογικών και τεχνητών νευρώνων ως προς την αρχιτεκτονική και τις ιδιότητες τους.

Ένα ΤΝΔ για να μπορέσει να παρέχει πληροφορία μέσα από αυτόματες διαδικασίες, όπου αυτός είναι και ο σκοπός του πραγματοποιεί αυτή τη διαδικασία μέσα από δύο φάσεις:

- (a) Φάση εκπαίδευσης (training)
- (b) Φάση ελέγχου (testing)

Ένα ΤΝΔ αποτελείται από μια πληθώρα στοιχείων, τους νευρώνες. Ως νευρώνας ορίζεται η μαθηματική απεικόνιση, η οποία λαμβάνει δεδομένα εισόδου (inputs). Κάθε είσοδος έχει διαφορετική τιμή βάρους (weight), ανάλογα με το πόσο ισχυρά συνδεδεμένοι είναι οι νευρώνες που συνδέονται με το βάρος αυτό. Στη συνέχεια γίνεται η επεξεργασία των δεδομένων και γίνεται η εξαγωγή της πληροφορίας (outputs) (Sopelsa et al., 2021).

Στο Διάγραμμα 2.3 παρουσιάζεται το βασικό μοντέλο ενός νευρώνα, όπου μπορεί να γίνει η διάκριση τριών φάσεων λειτουργίας του.



Διάγραμμα 2.3: Βασικό μοντέλο τεχνητού νευρώνα (Τσουχνικά, 2007)

Στην πρώτη φάση κάθε είσοδος πολλαπλασιάζεται με το συνοπτικό βάρος που της αντιστοιχεί.

Στη δεύτερη φάση το σταθμισμένο άθροισμα των δεδομένων της εισόδου με τα αντίστοιχα βάρη αθροίζεται με τον όρο της μεροληψίας ή αλλιώς της πόλωσης (bias), όπου δημιουργείται το τοπικό πεδίο (net input).

Στην τρίτη φάση εφαρμόζεται η συνάρτηση ενεργοποίησης ή συνάρτηση μεταφοράς (activation function ή squashing function) στο τοπικό πεδίο και το αποτέλεσμα τους δίνουν την έξοδο του νευρώνα .

Με μαθηματικούς όρους μπορεί να περιγραφεί ακολούθως:

$$\hat{u}_j = \sum_{i=1}^N w_{j,i} x_i$$

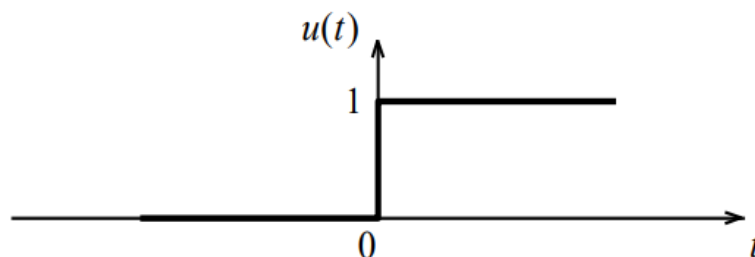
$$\hat{y}_j = \varphi(\hat{u}_j + b_j)$$

όπου x_1, x_2, \dots, x_i είναι τα σήματα εισόδου, $w_{j1}, w_{j2}, \dots, w_{ji}$ είναι τα συναπτικά βάρη του νευρώνα j , \hat{u}_j η έξοδος του αθροιστή που αφορά τα σήματα εισόδου ή αλλιώς το τοπικό πεδίο, b_j η μεροληψία ή πόλωση (bias), $\varphi(\cdot)$ η συνάρτηση ενεργοποίησης και \hat{y}_j η έξοδος του νευρώνα.

$$\varphi(u_j) = \begin{cases} 0, & u \leq \theta \\ 1, & u > \theta \end{cases} \quad (2.1)$$

Η συνάρτηση ενεργοποίησης που χρησιμοποιεί το συγκεκριμένο μοντέλο είναι η συνάρτηση βήματος (step function), η οποία λαμβάνει δύο τιμές: Μηδέν (0) αν η τιμή που λαμβάνει το x

είναι μικρότερη ή ίση με το θ , όπου θ η οριακή τιμή της κλάσης και ένα (1) αν είναι μεγαλύτερη του θ , δηλαδή το θ λειτουργεί σαν κατώφλι. Έτσι με αυτόν τον τρόπο κατατάσσει τα δεδομένα στην αντίστοιχη κλάση που ανήκουν και η γραφική απεικόνιση παρουσιάζεται στο Διάγραμμα 2.4:



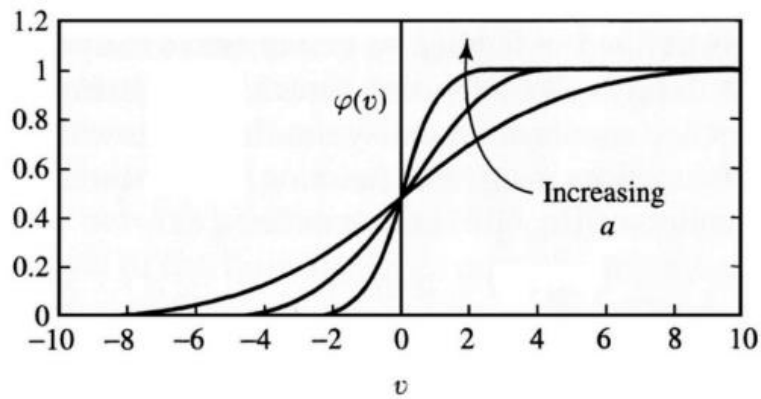
Διάγραμμα 2.4: Απεικόνιση βηματικής συνάρτησης (Παρασκευάς, n.d.)

Το μοντέλο αυτό ονομάζεται και μοντέλο McCulloch-Pitts.

Η μελέτη των ΤΝΔ αυξανόταν όλο ένα και περισσότερο με την πάροδο των χρόνων και γεννήθηκε το ερώτημα, αν η συνάρτηση ενεργοποίησης έπαιζε σημαντικό ρόλο στην βελτίωση αυτών. Έτσι αποφάνθηκε ότι κάνοντας χρήση μιας παραγωγίσιμης συνάρτησης, θα έδινε καλύτερα αποτελέσματα σε σχέση με μια βηματική. Η πιο γνωστή και ευρέως χρησιμοποιούμενη συνάρτηση ενεργοποίησης είναι η σιγμοειδής συνάρτηση. Το πεδίο ορισμού της είναι το σύνολο των πραγματικών αριθμών, αλλά συνήθως χρησιμοποιείται ένα μικρό υποσύνολο, το $[-1,1]$ ή το $[0,1]$ και δίνεται από τον παρακάτω τύπο:

$$\varphi(u_j) = \frac{1}{1+e^{-au_j}} \quad (2.2)$$

όπου a η παράμετρος της κλίσης και χαρακτηριστικά απεικονίζεται στο Διάγραμμα 2.5. Όταν το a τείνει προς το άπειρο, η λογιστική τείνει προς την βηματική συνάρτηση.

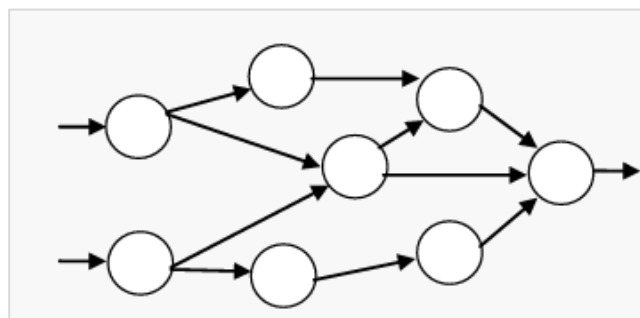


Διάγραμμα 2.5 Σιγμοειδής συνάρτηση (Γερολυμάτος, 2022)

Από το σύνολο των σιγμοειδών συναρτήσεων η πιο διαδεδομένη είναι η λογιστική συνάρτηση (logistic function).

2.3.2 Πολυεπίπεδο Perceptron

Το Πολυεπίπεδο Μοντέλο Perceptron (Multilayer Perceptron - MLP) είναι ένας από τους πιο διαδεδομένους τύπους νευρωνικών δικτύων. Ανήκει στα δίκτυα πρόσθιας τροφοδότησης (Feed-Forward Neural Networks - FNN), αφού η πληροφορία μεταδίδεται από την είσοδο προς την έξοδο, όπως αντίστοιχα φαίνεται και στο Διάγραμμα 2.6 (Λύκας, 2014).



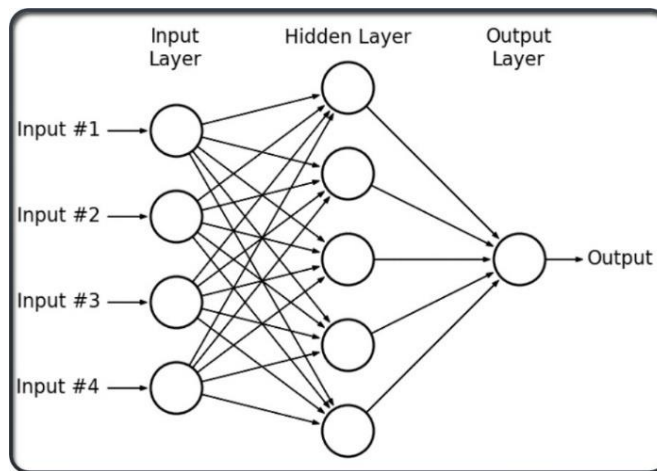
Διάγραμμα 2.6 Τρόπος μετάδοσης της πληροφορίας σε ένα MLP (Λύκας, 2014)

Παράλληλα σε αυτό το είδος ΤΝΔ δεν υπάρχει ανατροφοδότηση προς τους νευρώνες. Τα MLP νευρωνικά δίκτυα αποτελούν την εξέλιξη των μόνοστρωματικών νευρωνικών δικτύων (single-layer perceptrons). Ένα αρνητικό τους στοιχείο είναι ότι η διαδικασία εκμάθησής τους πραγματοποιείται με αργούς ρυθμούς, όπως επίσης χρειάζονται μεγάλο όγκο δεδομένων.

Τα MLPs αποτελούνται από:

- Ένα σύνολο σημείων συνένωσης (κόμβοι) εισόδου που αποτελούν την στιβάδα εισόδου (inputs)
- Μια ή περισσότερες κρυφές στιβάδες υπολογιστικών κόμβων (hidden layers)
- Μια στιβάδα εξόδου υπολογιστικών κόμβων (outputs)

Στο Διάγραμμα 2.7 παρουσιάζεται ένα ΤΝΔ MLP, που αποτελείται από μια στιβάδα εισόδου η οποία αποτελείται από τέσσερα χαρακτηριστικά, μια κρυφή στιβάδα και μια στιβάδα εξόδου.



Διάγραμμα 2.7 Απεικόνιση ενός MLP που περιέχει μια στιβάδα εισόδου μια εξόδου και μια κρυφή στιβάδα (Nisha, 2022)

Τα MLPs χρησιμοποιούν μη γραμμικές συναρτήσεις ενεργοποίησης, που κυρίως να ανήκουν στις σιγμοειδείς συναρτήσεις (Χονδροδήμα, 2017). Οι δύο πιο διαδεδομένες είναι:

Λογιστική συνάρτηση:

$$\hat{y}_j = \frac{1}{1 + \exp(-u_j)}$$

Εφαπτόμενη Συνάρτηση:

$$\hat{y}_j = \frac{1 - \exp(-u_j)}{1 + \exp(-u_j)}$$

όπου u_j είναι ο σταθμικός μέσος όρος των συνάψεων του j νευρώνα και \hat{y}_j η έξοδος του νευρώνα.

Για ένα νευρωνικό 2 κρυφών στιβάδων, ένα επίπεδο εξόδων και ένα διάνυσμα $X \in \mathbb{R}^n$, η έξοδος $Y \in \mathbb{R}^m$ καθορίζεται όπως παρακάτω:

$$h(1) = f(1)(W(1)T x + b(1)) \quad (2.3)$$

$$h(2) = f(2)(W(2)T h(1) + b(2)) \quad (2.4)$$

$$y = f(3)(W(3)T h(2) + b(3)) \quad (2.5)$$

όπου το $h(i)$ είναι η έξοδος του επιπέδου i , η $F(i)$ είναι η συνάρτηση ενεργοποίησης του επιπέδου i , το $b(i)$ αποτελεί το bias του επιπέδου i , και το $W(i)$ είναι ο πίνακας βαρών του επιπέδου i . Το νευρωνικό δίκτυο είναι μια σύνθεση συναρτήσεων.

Η διαδικασία εκπαίδευσης τους γίνεται με επιβλεπόμενο τρόπο, αφού χρησιμοποιούν τον αλγόριθμο σφάλματος οπισθοδιάδοσης (error back-propagation). Αυτός ο αλγόριθμος βασίζεται στον κανόνα διόρθωσης σφάλματος (error-correction learning rule).

Ο αλγόριθμος back-propagation απαρτίζεται από δύο περάσματα διαμέσου του δικτύου: Ένα πέρασμα προς τα εμπρός και ένα πέρασμα προς τα πίσω. Στο πρώτο πέρασμα ένα διάνυσμα εισόδου εφαρμόζεται στα σημεία συνένωσης του δικτύου και η επίδραση του μεταφέρεται από επίπεδο στιβάδα σε στιβάδα. Οι τιμές του διανύσματος εισόδου μετατρέπονται σε τιμές εξόδου μετά από κατάλληλους υπολογισμούς. Σε αυτό το πέρασμα τα βάρη του δικτύου είναι σταθερά. Κατά το πίσω πέρασμα τα βάρη μεταβάλλονται και προσαρμόζονται με βάση τον κανόνα διόρθωσης σφάλματος. Παράλληλα η πραγματική απόκριση της εξόδου αφαιρείται από την επιθυμητή τιμή της εξόδου, δημιουργώντας ένα σήμα σφάλματος, όπου και μεταδίδεται προς τα πίσω, διαμέσου του δικτύου, αντίθετα στην κατεύθυνση των συναπτικών συνδέσεων. Επιπροσθέτως γίνεται αναπροσαρμογή των βαρών του δικτύου με τελικό στόχο η παραγόμενη έξοδος του δικτύου να πλησιάζει όσο το δυνατόν περισσότερο την επιθυμητή τιμή της εξόδου (Χονδροδήμα, 2017).

Στάδια εκπαίδευσης των Νευρωνικών Δικτύων MLP:

- Επιλογή αριθμού κρυφών στιβάδων
- Επιλογή νευρώνων που θα στελεχώσουν τις στιβάδες
- Εύρεση βέλτιστης λύσης για την αποφυγή των τοπικών ελαχίστων και συνεπώς την μείωση του σφάλματος
- Εύρεση της καλύτερης δυνατής λύσης στο μικρότερο χρονικό διάστημα που απαιτείται
- Έλεγχος του δικτύου για την αποφυγή υπερ προσαρμογής (over fitting)

Σε αυτό εδώ το σημείο είναι πολύ σημαντικό να αναφερθεί, ότι για την αποφυγή υπερ προσαρμογής (over fitting) πρέπει να γίνει κατάλληλη επιλογή των νευρώνων. Παράλληλα, πρέπει να γίνει η κατάλληλη επιλογή των epochs (εποχών). Τα epochs είναι μια παράμετρος του νευρωνικού δικτύου που δείχνει πόσες φορές φορές έχει επεξεργαστεί ο αλγόριθμος ολόκληρο το σύνολο δεδομένων εκπαίδευσης.

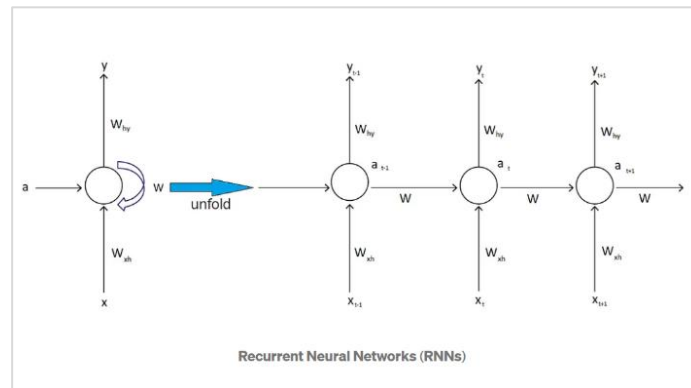
Σε κάθε εποχή μεταβάλλονται τα βάρη του νευρωνικού δικτύου με σκοπό την εύρεση του καλύτερου μοντέλου. Επιπροσθέτως πολύ σημαντικό ρόλο λαμβάνει μια ακόμα παράμετρος του νευρωνικού δικτύου, το batch size. Ως batch size ορίζεται ο αριθμός των παραδειγμάτων εκπαίδευσης που υπάρχουν σε κάθε παρτίδα (batch). Όσο μεγαλύτερη τιμή έχει το batch size, τόσο γρηγορότερη γίνεται η εκπαίδευση, βέβαια χρειάζεται μεγαλύτερο πλήθος δεδομένων εκπαίδευσης (training data).

2.3.3 Παλινδρομικά Νευρωνικά Δίκτυα

Το παλινδρομικό νευρωνικό δίκτυο (Recurrent neural Network, RNN) είναι μια κατηγορία τεχνητού νευρωνικού δικτύου, τα οποία βρίσκουν εφαρμογή στην επεξεργασία ακολουθιακών δεδομένων. Έχουν χρησιμοποιηθεί με επιτυχία σε προβλήματα μετάφρασης κειμένου αλλά και αναγνώρισης φωνής. Η τοπολογία του νευρωνικού δικτύου είναι αρκετά κοινή με του MLP, αλλά επιτρέπει συνδέσεις μεταξύ των κρυφών στιβάδων με πάνω από μια χρονική καθυστέρηση. Με αυτόν τον τρόπο το δίκτυο μπορεί να «μάθει» από προηγούμενα δεδομένα. Αντίθετα τα MLPs δημιουργούν την πρόβλεψη από την είσοδό τους.

Στο Διάγραμμα 2.8 παρουσιάζεται ένα τυπικό νευρωνικό δίκτυο RNN (Simple RNN), πως ξεδιπλώνεται σε μια πλήρη ακολουθία. Αποτελείται από μια στιβάδα εισόδου, μια κρυφή στιβάδα και στιβάδα εξόδου. Η κρυφή στιβάδα ενημερώνεται σε κάθε βήμα από τα αντίστοιχα

δεδομένα εισόδου εκείνης της χρονικής στιγμής και από την κρυφή στιβάδα της προηγούμενης χρονικής στιγμής (Πελέκης & Χονδροδήμα, 2022) Η μορφή του είναι η ακόλουθη:



Διάγραμμα 2.8 Τοπολογία παλινδρομικού τεχνητού νευρωνικού δικτύου (Manish, 2019)

όπου:

- W_{xh} : ο πίνακας των βαρών για τη σύνδεση της στιβάδας εισόδου με την κρυφή στιβάδα
- W : ο πίνακας των βαρών ανάμεσα στη κρυφή στιβάδα και στον εαυτό της σε διαφορετικές χρονικές στιγμές
- W_{hy} : ο πίνακας των βαρών για τη σύνδεση της κρυφής στιβάδας με τη στιβάδα εξόδου
- a : συνάρτηση ενεργοποίησης της στιβάδας

Οι παράμετροι που χρησιμοποιεί ένα RNN για κάθε χρονικό βήμα είναι κοινές. Στη παραπάνω εικόνα, οι παράμετροι W_{hx} , W_{hy} και W είναι ίδιες για κάθε χρονικό βήμα, αλλά με διαφορετικές εισόδους, αφού ένα RNN είναι μια αλυσίδα αντιγράφων του ίδιου δικτύου. Τα RNN δίκτυα λαμβάνουν υπόψιν τους τον χρόνο, όπως επίσης συνδυάζουν την τρέχουσα είσοδο με την κατάσταση του προηγούμενου βήματος σε μια συνάρτηση η οποία παράγει την καινούρια τρέχουσα κατάσταση.

Η λειτουργία ενός τυπικού νευρωνικού δικτύου RNN είναι η ακόλουθη:

Την χρονική στιγμή t τα δεδομένα εισόδου από την τρέχουσα μονάδα δεδομένων (data point) $x^{(t)}$ εισέρχονται στις στιβάδες με επαναλαμβανόμενες άκρες, όπως επίσης εισέρχονται και τιμές από την κρυφή στιβάδα $a^{(t-1)}$, όπου αυτές οι τιμές απεικονίζουν την κατάσταση του δικτύου μια χρονική στιγμή πριν. Ο υπολογισμός του $y^{(t)}$ καθορίζεται με βάση τα

αποτελέσματα που έλαβε το $a^{(t)}$ της κρυφής στιβάδας την χρονική στιγμή t . Είναι φυσικό ότι τα δεδομένα εισόδου $x^{(t-1)}$ που θα ληφθούν την χρονική στιγμή $t-1$ μπορούν να επηρεάσουν τα δεδομένα εξόδου $y^{(t)}$ τη χρονική στιγμή t και έπειτα με την πάροδο των επαναλαμβανόμενων συνδέσεων.

Υπάρχουν δύο συναρτήσεις που καθορίζουν τους υπολογισμούς σε κάθε χρονικό βήμα του περάσματος προς τα εμπρός ενός τυπικού RNN:

$$a^{(t)} = \sigma(W^{hx} \cdot x^{(t)} + Wa^{(t-1)} + b_a) \quad (2.6)$$

$$y^{(t)} = \text{softmax}(Wy^h a^{(t)} + b_y) \quad (2.7)$$

όπου b_a , b_y είναι οι παράμετροι της μεροληψίας ή των σφαλμάτων (bias parameters) που παράγονται και βοηθούν το νευρωνικό δίκτυο να προσαρμόσει πιο ορθά τα βάρη του μέσω του back-propagation για να παρέχει πιο στοχευμένη πρόβλεψη.

2.3.4 Διαφορές μεταξύ ενός MLP και ενός RNN νευρωνικού δικτύου

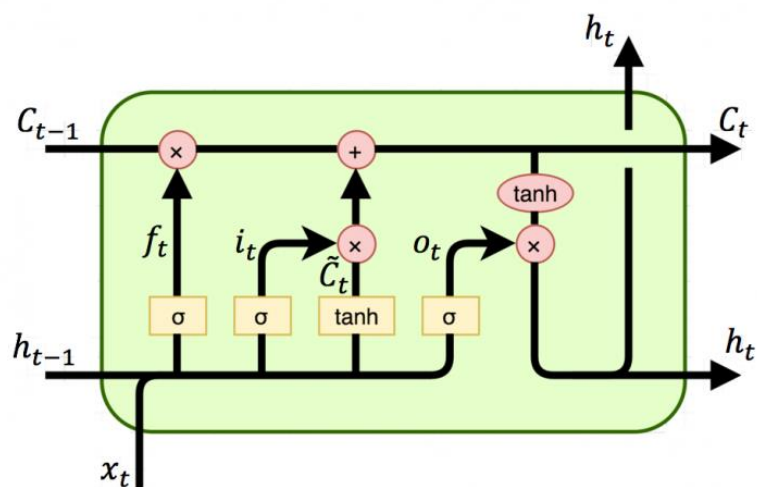
Ένα παράδειγμα για να γίνει καλύτερα κατανοητή η διαφορά μεταξύ ενός τυπικού νευρωνικού δικτύου RNN, ενός νευρωνικού Δικτύου MLP είναι το ακόλουθο: Έχοντας δεδομένα από εικόνες ζώων και θέλοντας να γίνει ορθή κατανομή τους σε είδη, όπως για παράδειγμα σκύλος, γάτα, κ.λπ., το νευρωνικό δίκτυο το οποίο ενδείκνυται για τέτοιου είδους προβλήματα είναι το MLP. Από την άλλη πλευρά, αν φωτογραφίες ζώων αλλάζουν ανά λεπτό και πρέπει να ταξινομηθούν ανά δραστηριότητες, π.χ. αν ο σκύλος ή γάτα κάθεται ή τρέχει, τότε χρειάζεται ένα παλινδρομικό νευρωνικό Δίκτυο. Παράλληλα ένα ακόμα παράδειγμα για την ορθότερη κατανόηση είναι το ακόλουθο: Ένας άνθρωπος διαβάζει ένα βιβλίο, διαβάζοντας τις προηγούμενες λέξεις, κατανοώντας και το νόημα όσων διαβάζει, τότε μπορεί με ακρίβεια να προβλέψει την επόμενη λέξη. Έτσι λειτουργεί και ένα παλινδρομικό νευρωνικό δίκτυο, λαμβάνοντας τις προηγούμενες λέξεις μπορεί να κάνει πρόβλεψη για την επόμενη.

2.3.5 Δίκτυα Μακράς Βραχύχρονης Μνήμης

Τα νευρωνικά δίκτυα μακράς βραχύχρονης μνήμης (LSTM) ανήκουν στην οικογένεια των παλινδρομικών νευρωνικών δικτύων (RNN). Τα RNN αν και είναι αποτελεσματικά σε διάφορες λύσεις προβλημάτων, υστερούν στο γεγονός ότι είναι επιρρεπή στην μεταφορά της

πληροφορίας για μεγάλο χρονικό διάστημα, αφού συγκρατούν την πληροφορία για δέκα περίπου χρονικές στερήσεις. Αυτό το πρόβλημα ήρθαν να λύσουν τα νευρωνικά δίκτυα LSTM. Η πρώτη τους αναφορά έγινε το 1997 από τους Sepp Hochreiter και Jürgen Schmidhuber. 33 (Kühnert et al., 2021).

Τα νευρωνικά δίκτυα μακράς βραχύχρονης Μνήμης (LSTM) είναι σχεδόν ίδια με τα RNN, αφού και αυτά διαθέτουν παλινδρομικές συνδέσεις, με την διαφορά τους να βρίσκεται ότι η μοντελοποίηση τους είναι καλύτερη σε ακολουθιακά δεδομένα που περιέχουν μακροπρόθεσμες συσχετίσεις. Τα LSTM περιέχουν μικρότερα συστήματα (νευρώνες) που ονομάζονται πύλες (gates) και η χρήση τους είναι ρυθμίζουν το είδος της πληροφορίας που θα εισέρχεται στην εσωτερική κατάσταση του δικτύου. Με άλλα λόγια θα προσθέτουν ή θα αφαιρούν την πληροφορία που θα εισέρχεται στο δίκτυο. Με αυτό τον τρόπο αυτό που επιτυγχάνεται είναι η ορθότερη διαλογή πληροφορίας κάνοντας το δίκτυο να θυμάται μακροχρόνιες συσχετίσεις. Στο Διάγραμμα 2.9 απεικονίζεται η εσωτερική δομή ενός δικτύου LSTM.



Διάγραμμα 2.9: Απεικόνιση της δομής της εσωτερικής μονάδας ενός δικτύου LSTM (wikidocs, n.d.)

όπου:

X : element – wise πολλαπλασιασμός

+ : element – wise πρόσθεση

σ : Επίπεδο NN λογιστικής συνάρτησης

tanh : Επίπεδο NN συνάρτησης than

tahn: element wise υπολογισμός

Ένα από τα πιο βασικά στοιχεία ενός LSTM δικτύου είναι η εσωτερική κατάσταση της μονάδας (cell states, c). Ο ρόλος της είναι να συγκρατεί την πληροφορία από παλαιότερους κύκλους σε πιο πρόσφατους καταφέροντας με αυτόν τον τρόπο να επιλύσει το πρόβλημα της βραχυχρόνιας μνήμης. Με το πέρασμα των κύκλων μέσω των πυλών γίνεται η προσθαφαίρεση της πληροφορίας στην εσωτερική κατάσταση της μονάδας και στη συνέχεια γίνεται η προώθηση τους στη ροή επεξεργασίας. Η κάθε πύλη είναι ένα επίπεδο νευρωνικού δικτύου, το οποίο έχει συνάρτηση ενεργοποίησης μια λογιστική συνάρτηση. Οι τιμές που λαμβάνουν οι εξόδοι των νευρώνων κυμαίνονται μεταξύ 0 και 1. Σημαντικό στοιχείο αποτελεί πως ο element – wise πολλαπλασιασμός της εξόδου μιας πύλης με κάποιο διάστημα έστω u (ίδιας διάστασης) ελέγχει την διέλευση της πληροφορίας του διανύσματος u . Όταν η τιμή της εξόδου λάβει το 0 καταστέλλουν την πληροφορία αντίστοιχης θέσης στο διάστημα u , ενώ τιμές ίσες με 1 επιτρέπουν την διέλευση της πληροφορίας. Παράλληλα εκτός της εσωτερικής κατάστασης της μονάδας τα LSTM δίκτυα περιέχουν και την κρυφή κατάσταση (hidden state), όπου γίνεται η μεταφορά των στοιχείων στους διαδοχικούς κύκλους επεξεργασίας, με τελικό σκοπό την εξαγωγή της πληροφορίας (πρόβλεψη) (Βλαχάβας και λοιποί, 2002)

- Πύλη λήθης

Η πύλη λήθης (forget gate) πραγματοποιεί τον έλεγχο, για το αν τα δεδομένα από μια προηγούμενη κατάσταση που βρισκόταν το LSTM δίκτυο είναι χρήσιμα για να προωθηθούν στα επόμενα στάδια του δικτύου ή πρέπει να «ξεχαστούν».

Πρακτικά η πύλη λήθης ελέγχει τον βαθμό επιρροής της εσωτερικής κατάστασης της μονάδας την στιγμή c_{t-1} ως προς την την κατάσταση c_t , στο διάστημα $[h_{t-1}, x_t]$. Η έξοδος του θα είναι η ακόλουθη:

$$f_t = \sigma_g(U_f * h_{t-1} + W_f * x_t + b_f) \quad (2.8)$$

όπου :

σ_g : λογιστική συνάρτηση

b_f : διάνυσμα πολώσεων του δικτύου

x_t, h_{t-1} : οι εισοδοι το δικτύου. Οι τιμές τους διαμορφώνονται από πίνακες βαρών

- Πύλη εισόδου:

Η πύλη εισόδου (input gate) πραγματοποιεί τον έλεγχο της νέας πληροφορίας που εισέρχεται στο δίκτυο μέσω ενός διανύσματος i_t , το οποίο αξιολογεί την ποιότητα της νέα πληροφορίας και αποφασίζει αν πρέπει να ενταχθεί στην εσωτερική μονάδα c_t . Η νέα πληροφορία προκύπτει από την ακόλουθη σχέση:

$$\hat{C}_t = \tanh(U_c * h_{t-1} + W_c * x_t + b_c) \quad (2.9)$$

Η εξίσωση της πύλης εισόδου είναι η ακόλουθη:

$$i_t = c_g (U_i * h_{t-1} + W_x * x_t + b_c) \quad (2.10)$$

Η νέα κατάσταση c_t θα είναι η ακόλουθη:

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \quad (2.11)$$

- Πύλη εξόδου

Ο ρόλος της πύλης εξόδου (output gate) είναι να καθορίζει, ποια στοιχεία της εσωτερικής μονάδας c_t θα μεταφερθούν στην έξοδο του δικτύου. Το διάνυσμα c_t ως element - wise άθροισμα δύο άλλων διανυσμάτων περνά δια μέσω ενός στρώματος που έχει ως συνάρτηση ενεργοποίησης την tanh και έτσι παράγεται το τελικό διάνυσμα εξόδου. Με αυτόν τον τρόπο οι τιμές του διανύσματος βρίσκονται εντός του διαστήματος [-1,1] (Βλαχάβας et al., 2002).

Η μαθηματική απεικόνιση της πύλης εξόδου έχει την ακόλουθη μορφή:

$$o_t = c_g(U_o * h_{t-1} + W_x * x_t + b_o) \quad (2.12)$$

Αντίστοιχα η τελική πρόβλεψη (έξοδος) του δικτύου είναι:

$$h_t = o_t \cdot \tan(c_t) \quad (2.13)$$

Οι παραπάνω εξισώσεις εξηγούν τον τρόπο λειτουργίας ενός LSTM δικτύου, το οποίο είναι σε θέση να διατηρεί μακροπρόθεσμες συσχετίσεις σε ακολουθίες δεδομένων και να ελέγχει την ροή της πληροφορίας με βάση την σημασία της.

Η κύρια διαφορά σε σχέση με τα παλινδρομικά νευρωνικά δίκτυα (RNN) βρίσκεται, στο ότι τα LSTM κάθε παλινδρομικό τμήμα έχουν προστεθεί τρεις πύλες, η πύλη εισόδου, η πύλη εξόδου και η πύλη λήθης. Αυτές οι τρεις πύλες βοηθούν στη μείωση του σφάλματος για την διάρκεια που το νευρωνικό δίκτυο εκπαιδεύεται και καταφέρνουν να κρατήσουν την πληροφορία για 1000 χρονικές στερήσεις, κάτι που δεν μπορεί να πραγματοποιήσει ένα RNN δίκτυο.

2.3.6 Φραγμένα ανατροφοδοτούμενα νευρωνικά δίκτυα

Τα φραγμένα ανατροφοδοτούμενα δίκτυα (GRU) αποτελούν μια ακόμα σημαντική αρχιτεκτονική παλινδρομικών νευρωνικών δικτύων, η οποία αποτελεί μια επέκταση των RNN δικτύων. Η πρώτη τους εμφάνιση έλαβε χώρα για πρώτη φορά το 2014 από τον Cho Kyunghyun. Η δομή τους μοιάζει πολύ με αυτά των LSTM δικτύων, αφού σχεδιάστηκαν, ώστε να έχουν την δυνατότητα να πραγματοποιούν επαναφορά της μνήμης τους, όπως και να μπορούν και να την αναβαθμίσουν, με την ιδιαιτερότητα, πως χρειάζονται λιγότερες παραμέτρους σε σχέση με τα LSTM δίκτυα, το οποίο τα κάνει αποδοτικότερα σε υπολογιστική ισχύ.

Τα GRU δίκτυα διαθέτουν δύο πύλες, η πρώτη ονομάζεται πύλη ενημέρωσης (update gate) που συμβολίζεται με z_t και η δεύτερη ονομάζεται πύλη επαναφοράς (reset gate), η οποία συμβολίζεται με r_t . Τα GRU δίκτυα, όπως και το LSTM δίκτυο σχεδιάστηκαν για να μπορούν να διαγράφουν από την μνήμη τους πληροφορία, η οποία δεν είναι χρήσιμη.

Τα δίκτυα GRU εκθέτουν πλήρως το περιεχόμενο της μνήμης τους και με την μέθοδο ολοκλήρωση διαρροής (Leaky Integration), ισορροπούν μεταξύ του προηγούμενου περιεχομένου μνήμης και του νέου περιεχομένου μνήμης.

- Πύλη ενημέρωσης:

Η Πύλη ενημέρωσης είναι υπεύθυνη για τον έλεγχο των δεδομένων και την απόφαση, για το ποια από αυτά θα περάσει από τα προηγούμενά στάδια στα επόμενα. Ο υπολογισμός της επιτελείται με τον κάτωθι τύπο:

$$z_t = \sigma_g(W_z x_t + U_z h_{t-1} + b_z) \quad (2.14)$$

- Πύλη επαναφοράς:

Η πύλη επαναφοράς είναι υπεύθυνη για το ποια δεδομένα δεν είναι χρήσιμα πια για τον αλγόριθμο και πρέπει να διαγραφτούν. Ο υπολογισμός της επιτελείται με τον κάτωθι τύπο:

$$r_t = \sigma_g(W_r x_t + U_r h_{t-1} + b_r) \quad (2.15)$$

- Υποψήφια έξοδος:

$$\tilde{h}_t = \sigma_h(W_h x_t + U_h(r_t \circ h_{t-1}) + b_h) \quad (2.16)$$

- Έξοδος:

$$h_t = z_t \circ h_{t-1} + (1 - z_t) \circ \tilde{h}_t \quad (2.17)$$

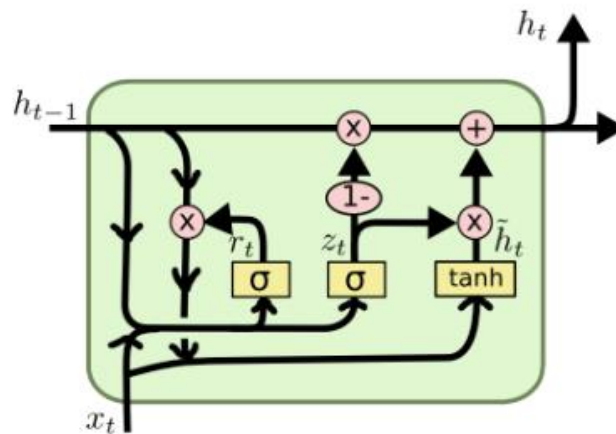
όπου:

x_t : το διάνυσμα της εισόδου

h_t : το διάνυσμα της εξόδου

z_t : το διάνυσμα ενημέρωσης

r_t : το διάνυσμα επαναφοράς



Διάγραμμα 2.10: Εσωτερική δομή ενός δικτύου GRU (Primo ai, 2023)

Σημαντικό στοιχείο αποτελεί, πως οι πύλες του αλγορίθμου χρησιμοποιούν την σιγμοειδή συνάρτηση, αντίθετα με την έξοδο που χρησιμοποιεί την υπερβολική εφαπτομένη.

- Σύγκριση RNN με GRU:

Τα RNN είναι αποτελεσματικά σε πολλές εφαρμογές, όμως εμφανίζουν δυσκολίες στη διατήρηση πληροφορίας για μεγάλα χρονικά διαστήματα, λόγω της εξαφάνισης της κλίσης (vanishing gradient problem). Το πρόβλημα αυτό γίνεται ιδιαίτερα αισθητό, όταν οι νευρώνες αυξάνονται κατά πολύ, τότε οι κλίσεις πλησιάζουν στο 0, και έτσι τα νευρωνικά δίκτυα είναι δύσκολο να εκπαιδευτούν. Αυτό το πρόβλημα κατάφεραν να ξεπεράσουν τα GRU, χρησιμοποιώντας τις πύλες τους, επιτρέποντάς τους να διατηρούν την πληροφορία για πολλά βήματα.

2.4 Συναρτήσεις Νευρωνικών Δικτύων

2.4.1 Συνάρτηση απώλειας

Στη μηχανική μάθηση είναι πολύ χρήσιμος ο ορισμός ενός μέτρου απόδοσης P , που ορίζει τον βαθμό της ακρίβειας του αλγορίθμου. Ο στόχος των περισσότερων αλγορίθμων είναι η βελτιστοποίηση του μοντέλου συναρτήσει του βαθμού ακρίβειας του, δηλαδή το P . Για τον σκοπό αυτό πραγματοποιείται εισαγωγή μιας συνάρτησης απώλειας, η οποία δείχνει τον βαθμό αποτελεσματικότητας του μοντέλου υπολογίζοντας την απόκλιση της προβλεπόμενης τιμής της

εξόδου από την πραγματική. Στόχος του δικτύου είναι η ελαχιστοποίηση της συνάρτησης απώλειας, ώστε να πετύχει την μεγαλύτερη δυνατή ακρίβεια (El Hajji et al., 2019)

2.4.2 Συναρτήσεις ενεργοποίησης

- ReLU

Η πιο δημοφιλής συνάρτηση ενεργοποίησης στην έρευνα για τα τεχνητά νευρωνικά δίκτυα, είναι η συνάρτηση ReLU (Rectified linear activation function). Η Συνάρτηση της ReLU περιγράφεται από την παρακάτω συνάρτηση:

$$f(x) = \max(0, x) \quad (2.18)$$

Η συνάρτηση αυτή παρουσιάζει αρκετά πλεονεκτήματα. Αρχικά είναι πολύ αποδοτική υπολογιστικά. Πραγματοποιείται η σύγκλιση του δικτύου πολύ γρήγορα. Παράλληλα είναι μη γραμμική συνάρτηση. Έχει παράγωγο και επιτρέπει back-propagation.

Ένα μειονέκτημα που παρουσιάζει η παραπάνω συνάρτηση, είναι ότι αν κάποια χαρακτηριστικά είναι αρνητικά, τότε τα αντίστοιχα βάρη τους δεν θα επικαιροποιηθούν καθόλου, οπότε δεν μπορεί να εκτελέσει back propagation (Maniatiopoulos et al., 2021)

- LReLU

Μια άλλη αρκετά δημοφιλής συνάρτηση ενεργοποίησης είναι η συνάρτηση γραμμικής ReLU με διαρροή (leaky ReLU) (Maniatiopoulos et al., 2021). Η συνάρτηση παρουσιάζεται παρακάτω:

$$f(x) = \begin{cases} x & \text{αν } x \geq 0 \\ ax & \text{αν } x \leq 0 \end{cases} \quad (2.19)$$

Η LReLU εξαλείφει το πρόβλημα της μη ενημέρωσης των βαρών για αρνητικές τιμές διότι προσφέρει μικρές τιμές για αρνητικά χαρακτηριστικά, αφού το a κυμαίνεται στο διάστημα $[0.01, 0.03]$ και μεγαλύτερες για θετικά. Το a εκφράζει το ποσό της διαρροής, όταν η είσοδος είναι αρνητική.

Ένα θετικό χαρακτηριστικό της LReLU είναι πως διαθέτει μια μικρή θετική κλίση στην αρνητική περιοχή, οπότε επιτρέπει το backpropagation, ακόμη και για αρνητικές τιμές εισόδου.

Το αρνητικό χαρακτηριστικό της LReLU είναι, πως τα αποτελέσματα τα οποία εξάγει δεν είναι συνεπή, αφού δεν παράγει ασφαλείς προβλέψεις για αρνητικές τιμές εισόδου.

- Κλιμακωτή scaled tanh συνάρτηση

Παράλληλα μια δημοφιλής συνάρτηση ενεργοποίησης, επίσης η είναι η κλιμακωτή scaled tanh συνάρτηση (Scaled Hyperbolic Tangent), που περιγράφεται από τον ακόλουθο τύπο:

$$f(x) = \alpha \cdot \tanh(b \cdot x) \quad (2.20)$$

Μεγάλο πλεονέκτημα της κλιμακωτής tanh συνάρτησης αποτελεί το γεγονός, ότι έχει εύκολα υπολογίσιμη παράγωγο. Η συνάρτηση αυτή λαμβάνει τιμές στο διάστημα $[-1,1]$. Αυτό έχει ως αποτέλεσμα την καταπολέμηση της μη - γραμμικότητας, αφού η συνάρτηση αυτή «συμπιέζει» τις τιμές εισόδου εντός του διαστήματος $[-1,1]$. Από την άλλη πλευρά είναι ευαίσθητη σε περιπτώσεις, όπου οι τιμές εισόδου είναι αρνητικές, αφού το εύρος της είναι συγκεκριμένο. Σε μεγάλους εύρους αρνητικές τιμές δεν είναι ιδιαίτερα αξιόπιστη, αφού οι αποκλίσεις είναι μεγάλες στην εκπαίδευση του νευρωνικού δικτύου.

- Σιγμοειδής συνάρτηση

Τέλος, μια κλασική συνάρτηση ενεργοποίησης είναι η σιγμοειδής συνάρτηση. Η μαθηματική της απεικόνιση είναι η παρακάτω:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.21)$$

Ένα μεγάλο πλεονέκτημα της σιγμοειδούς συνάρτησης αποτελεί η εύκολα υπολογίσιμη παράγωγος. Παρουσιάζονται όμως και κάποια μειονεκτήματα, παρόμοια με των tanh συναρτήσεων, αλλά και το γεγονός ότι η έξοδος είναι ορισμένη στο διάστημα $[0,1]$, κάτι που μπορεί να εισάγει μεροληψία στο τεχνητό νευρωνικό δίκτυο.

2.4.3 Εκπαίδευση τεχνητών νευρωνικών δικτύων

- Gradient Descent:

Ένας από τους πιο ευρέως χρησιμοποιούμενους αλγόριθμους βελτιστοποίησης είναι ο αλγόριθμος Gradient Descent. Βρίσκει εφαρμογή σε προβλήματα παλινδρόμησης, αλλά και ταξινόμησης. Ο αλγόριθμός αυτός έχει σκοπό την εύρεση όλων των βαρών και των πολώσεων με στόχο η συνάρτηση απώλειας να τείνει στο 0. Υπάρχουν πολλοί αλγόριθμοι Gradient Descent, διαφορετικής φιλοσοφίας με τους πιο διαδεδομένους να είναι:

α) Κατάβαση δυναμικού κατά παρτίδες (Batch Gradient Descent),

β) Στοχαστική κατάβαση δυναμικού (Stochastic Gradient Descent)

γ) Κατάβαση δυναμικού σε μικρές παρτίδες (Mini-batch Gradient Descent)

- Κατάβαση δυναμικού κατά παρτίδες

Η κατάβαση δυναμικού κατά παρτίδες χρησιμοποιεί το σύνολο των δεδομένων εκπαίδευσης προκειμένου να πραγματοποιήσει τον υπολογισμό της συνάρτησης απώλειας και της κλίσης με απώτερο σκοπό να προσαρμόσει κατάλληλα τα βάρη. Στο τέλος του υπολογισμού ανανεώνει τα βάρη. Ο αλγόριθμος αυτός είναι ιδιαίτερα αργός και κοστοβόρος, ειδικά σε μεγάλα σύνολα δεδομένων, αφού ο αλγόριθμος αυτός υπολογίζει και διατηρεί στη μνήμη του τις κλίσεις έως ότου ολοκληρωθεί μια εποχή εκπαίδευσης.

Στη συνέχεια αυτοί οι παράμετροι ανανεώνονται. Αυτό το γεγονός αποτελεί πρόβλημα, διότι η υπολογιστική ισχύς που χρειάζεται είναι πολύ μεγάλη (Κεράστας, 2021).

- Στοχαστική κατάβαση δυναμικού

Η προσέγγιση, την οποία περιέχει η Στοχαστική κατάβαση δυναμικού (Stochastic Gradient Descent, SGD) είναι αντίθετη σε σχέση με αυτή της κατάβασης δυναμικού κατά παρτίδες. Ο αλγόριθμος αυτός ανανεώνει τις παραμέτρους ανά επανάληψη πραγματοποιώντας τυχαία επιλογή σε αντίθεση με την κατάβαση δυναμικού κατά παρτίδες που επαναπροσδιορίζει κλίσεις για παρόμοια παραδείγματα, κάτι που κάνει την στοχαστική κατάβαση δυναμικού να

είναι πιο αποδοτική στην ταχύτητα εκπαίδευσης, ιδιαίτερα για μεγάλα σύνολα δεδομένων. Αρνητικό στοιχείο αποτελεί, πως η συνάρτηση απώλειας μπορεί να μην έχει τόσο ομαλή πορεία προς το ελάχιστο (Κεράστας, 2021).

- Κατάβαση δυναμικού κατά μικρές παρτίδες:

Η κατάβαση δυναμικού κατά μικρές παρτίδες αποτελεί μια μίξη των δύο προηγούμενων μεθόδων που αναφέρθηκαν, αφού αντί οι παράμετροι να ανανεώνονται στο τέλος της εποχής εκπαίδευσης, ή ανά επανάληψη, το πραγματοποιούν με την ολοκλήρωση μιας παρτίδας. Αυτό έχει το όφελος, ότι η ταχύτητα του αλγορίθμου παραμένει σε πολύ καλά επίπεδα, όπως επίσης και η αποτελεσματικότητά του είναι πολύ ικανοποιητική. Παράλληλα οι διαδικασίες της ενημέρωσης των παραμέτρων γίνεται με πιο αργό ρυθμό γεγονός, που μπορεί να οδηγήσει σε πιο σταθερή σύγκλιση.

- Ρυθμός μάθησης

Ο ρόλος του ρυθμού μάθησης (learning rate, συμβολίζεται με d) στην εκπαίδευση του μοντέλου είναι ιδιαίτερα σημαντικός, αφού επηρεάζει την ταχύτητα του. Όταν ο ρυθμός εκμάθησης έχει μεγάλο αριθμό μάθησης μπορεί να οδηγήσει σε γρηγορότερη σύγκλιση, αλλά και σε ταλάντωση γύρω από τις βέλτιστες τιμές βαρών. Αντίθετα όταν ο ρυθμός εκμάθησης είναι μικρός, τότε η σύγκλιση είναι πιο αργή, ενώ μπορεί να οδηγήσει και σε τοπικά ακρότατα.

- RMSProp:

Ο RMSProp είναι ένας μη δημοσιευμένος αλγόριθμος βελτιστοποίησης που εισήγαγε ο Geoffrey Hinton (Tieleman & Hinton, 2012). Η ανάπτυξη του ήρθε να αντιμετωπίσει ο πρόβλημα της επιλογής του ρυθμού μάθησης. Είναι ένας αλγόριθμος με ρυθμό προσαρμοστικής μάθησης και είναι εμπνευσμένος από έναν άλλο αλγόριθμο που ονομάζεται Adagrad. Η βασική ιδέα του RMSProp είναι η προσαρμογή του ποσοστού μάθησης σε σχέση με τις προηγούμενες κλίσεις. Για την υλοποίηση της προσαρμογής ο RMSprop διατηρεί έναν κινητό μέσο όρο του τετραγώνου της κλίσης, με τον οποίο διαιρεί την τρέχουσα κλίση.

- Adam

Ο αλγόριθμος Adam αποτελεί ίσως τον πιο διαδεδομένο αλγόριθμο βελτιστοποίησης, όπως επίσης πηγάζει από την σύνθεση δύο αλγορίθμων βελτιστοποίησης, αυτή του SGD, όπως και του RMSprop προσθέτοντας όμως, έναν όρο Ορμής (Momentum). Κάνει χρήση των τετραγωνικών κλίσεων, όπως και ο RMSprop με σκοπό την κλιμάκωση του ρυθμού μάθησης. Από την άλλη πλευρά διαφοροποιείται σε σχέση με τον αλγόριθμο SGD, κάνοντας χρήση του κινούμενου μέσου όρου της κλίσης (Γιαννακόπουλος, 2021; Χιούτας, 2022).

- Ορμή:

Η εκμάθηση με μια μέθοδο κλίσης μπορεί μερικές φορές να είναι πολύ αργή. Για τον λόγο αυτό ένας τρόπος επιτάχυνσης της διαδικασίας της εκπαίδευσης είναι η εισαγωγή μιας μεθόδου ορμής. Η μέθοδος αυτή ενδείκνυται σε περιπτώσεις, όπου υπάρχουν υψηλές καμπυλώσεις ή θορυβώδεις κλίσεις. Ο όρος «ορμή» αποτελείται από παρελθοντικές βαθμίδες που έχουν συσσωρευτεί με ένα εκθετικά κινούμενο μέσο όρο (Sun et al., 2022)

2.5 Ανακεφαλαίωση

Τα τεχνητά νευρωνικά δίκτυα αποτελούν την επιτομή της σύγχρονης τεχνολογίας αλγορίθμων, αφού αποτελούν ένα ιδιαίτερα χρήσιμο εργαλείο και δίνουν μια νέα διάσταση στον τρόπο επεξεργασίας και λύσης προβλημάτων, όπως στη μελέτη της χρηματιστηριακής αγοράς, στην ανάλυση εικόνων, την ανάλυση υπολογιστικής όρασης κλπ. Η χρήση τους ολοένα και αυξάνεται, καθώς μέρα με την μέρα δημιουργούνται νέοι αλγόριθμοι, όπου στόχο την ορθότερη εξαγωγή πληροφορίας.

Σε αυτό το κεφάλαιο έγινε αναφορά στα βασικά χαρακτηριστικά της μηχανικής μάθησης, όπως επίσης έγινε και διαχωρισμός στα είδη της. Έγινε διάκριση μεταξύ μηχανικής μάθησης και βαθιάς μάθησης, αν και η πρώτη αποτελεί υποσύνολο της δεύτερης. Πραγματοποιήθηκε μικρή αναφορά στον βιολογικό νευρώνα, αφού οι αρχιτεκτονικές των τεχνητών νευρωνικών δικτύων έχουν εμπνευστεί από τον τρόπο λειτουργίας τους καθώς και την επίδραση που έχει στον ανθρώπινο εγκέφαλο. Στη συνέχεια έγινε εκτενής ανάλυση του τρόπου λειτουργίας των νευρωνικών δικτύων, παρουσιάστηκαν οι μαθηματικές απεικονίσεις τους καθώς και οι διαφορές που παρουσιάζουν μεταξύ τους. Τέλος εξετάστηκαν τεχνολογίες εκπαίδευσης, όπως

οι αλγόριθμοι βελτιστοποίησης και οι συναρτήσεις βελτιστοποίησης, που χρησιμοποιούνται για την εκπαίδευση των δικτύων, αλλά και τεχνικές, που βοηθούν στην αποτελεσματική εκπαίδευση και προστασία από την υπέρ-εκπαίδευση.

ΚΕΦΑΛΑΙΟ 3: Έρευνα και αποτελέσματα

3.1 Εισαγωγή

Η ανάλυση και η πρόβλεψη των αποδόσεων της χρηματιστηριακής αγοράς αποτελούν κρίσιμα ζητήματα για επενδυτές, εταιρείες και οικονομικούς θεσμούς. Οι αποδόσεις αυτές αντικατοπτρίζουν την απόδοση των επενδύσεων και επηρεάζουν αποφάσεις σχετικά με την διαχείριση του χαρτοφυλακίου, την εκτίμηση του κινδύνου και την εκτέλεση στρατηγικών συναλλαγών. Η τεχνητή νοημοσύνη (TN) και η μηχανική μάθηση έχουν ανοίξει νέες προοπτικές για τη μοντελοποίηση αυτών των αποδόσεων. Η ενότητα αυτή εξετάζει τη χρήση της TN για την ανάλυση και πρόβλεψη των αποδόσεων της χρηματιστηριακής αγοράς και τη δυνατότητά της να προσφέρει νέες προοπτικές για την κατανόηση της δυναμικής της αγοράς. Στόχος μας είναι να αναδείξουμε τον ρόλο της τεχνητής νοημοσύνης στη βελτίωση της κατανόησης και πρόβλεψης των αποδόσεων της χρηματιστηριακής αγοράς και να προσφέρουμε νέα εργαλεία για την λήψη αποφάσεων σε αυτόν τον σημαντικό τομέα.

3.2 Η απόδοση σαν όρος στο χρηματιστήριο

Η απόδοση (return) στο χρηματιστήριο αναφέρεται στο κέρδος ή την απώλεια που προκύπτει από την επένδυση σε χρηματοοικονομικά εργαλεία όπως μετοχές, ομόλογα, και άλλα επενδυτικά προϊόντα. Είναι μια κρίσιμη μετρική για επενδυτές και εταιρείες, καθώς καθορίζει την απόδοση του κεφαλαίου τους και επηρεάζει τις αποφάσεις επενδυτικής στρατηγικής.

Οι αποδόσεις μπορούν να υπολογιστούν με διάφορους τρόπους, ανάλογα με το είδος του επενδυτικού προϊόντος και τη διάρκεια της επένδυσης. Δύο κύριοι τρόποι υπολογισμού των αποδόσεων είναι η απόδοση της κεφαλαίου (capital gain) και η απόδοση από το έκτακτο εισόδημα (income return).

- Απόδοση της κεφαλαίου (Capital Gain): Αυτή η μορφή απόδοσης αφορά την αύξηση (κέρδος) ή μείωση (απώλεια) της αξίας του επενδυτικού προϊόντος κατά την διάρκεια της κατοχής του. Υπολογίζεται ως η διαφορά μεταξύ της αρχικής τιμής αγοράς και της τρέχουσας αξίας του προϊόντος

- Απόδοση εισοδήματος (Income Return): Αυτός ο τύπος απόδοσης αφορά τα τακτικά εισοδήματα που προκύπτουν από την επένδυση, όπως τα μερίσματα από μετοχές ή τόκους ή από ομόλογα. Υπολογίζεται ως το συνολικό ποσό των εισοδημάτων που προκύπτουν από την επένδυση και εκφράζεται ως ποσοστό της ετήσιας απόδοσης προς την τρέχουσα αξία του προϊόντων

Οι αποδόσεις διαδραματίζουν έναν σημαντικό ρόλο στην αξιολόγηση της αποδοτικότητας μιας επένδυσης και τη λήψη αποφάσεων σχετικά με τον τρόπο διαχείρισης του επενδυτικού χαρτοφυλακίου. Είναι σημαντικό να εκτιμηθεί το ρίσκο, που συνδέεται με τις αποδόσεις και να εξεταστεί προσεκτικά η συνολική εικόνα της επένδυσης πριν ληφθούν αποφάσεις επενδυτικής στρατηγικής.

Η μοντελοποίηση των αποδόσεων στον χρηματοοικονομικό τομέα έχει μεγάλη σημασία για πολλούς λόγους, κάποιους από τους οποίους είναι οι ακόλουθοι:

- Κατανόηση της αγοράς: Βοηθά στην κατανόηση του τρόπου λειτουργίας της χρηματιστηριακής αγοράς και των διακυμάνσεων της. Με την ανάλυση των αποδόσεων, μπορούμε να ανιχνεύσουμε τάσεις, μοτίβα και πιθανούς παράγοντες που επηρεάζουν τις αγορές
- Εκτίμηση του κινδύνου: Βοηθά στον προσδιορισμό και την αξιολόγηση του ρίσκου. Οι μοντέλα απόδοσης μπορούν να χρησιμοποιηθούν για την εκτίμηση της πιθανότητας απώλειας κεφαλαίου και την ανάλυση του αναμενόμενου κινδύνου
- Επιλογή επενδύσεων: Βοηθά τους επενδυτές να επιλέξουν τα κατάλληλα επενδυτικά προϊόντα. Με τη χρήση μοντέλων, μπορούν να αναλύσουν την απόδοση προηγούμενων επενδύσεων και να λάβουν αποφάσεις βάσει πληροφοριών
- Ανάπτυξη επενδυτικών στρατηγικών: Βοηθά τους επενδυτές και τους επαγγελματίες της χρηματοοικονομικής αγοράς να αναπτύξουν επενδυτικές στρατηγικές που θα βασίζονται σε ανάλυση δεδομένων και προβλέψεις αποδόσεων
- Διαχείριση χαρτοφυλακίου: Βοηθά τους επενδυτές να διαχειριστούν αποδοτικά τα χαρτοφυλάκια τους. Με τη χρήση μοντέλων, μπορούν να επιλέγουν ποια επενδύσεις πρέπει να διατηρήσουν, να πουλήσουν ή να αγοράσουν

- Πρόβλεψη μελλοντικών αποδόσεων: Επιτρέπει την πρόβλεψη των μελλοντικών αποδόσεων της αγοράς. Αυτό είναι ιδιαίτερα χρήσιμο για επενδυτές και επιχειρήσεις που προσπαθούν να κατανοήσουν τις μελλοντικές τάσεις της αγοράς

Στην ανάλυση των τιμών των μετοχών, υπάρχουν δύο κύριοι τρόποι για τον υπολογισμό των αποδόσεων: η απλή απόδοση (Simple Return) και η λογαριθμική απόδοση (Logarithmic Return). Κάθε μία από αυτές τις μεθόδους έχει τα πλεονεκτήματά της και χρησιμοποιείται ανάλογα με τον σκοπό της ανάλυσης. Ας εξηγήσουμε κάθε μία από αυτές τις μεθόδους:

- Απλή Απόδοση (Simple Return):

Η απλή απόδοση υπολογίζεται ως η απόλυτη διαφορά μεταξύ της αρχικής τιμής (αγοράς) και της τελικής τιμής (πώλησης) της μετοχής, διαιρεμένη με την αρχική τιμή.

$$\text{Simple Return} = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (3.1)$$

Παρόλο που είναι εύκολη στον υπολογισμό και στην κατανόηση της, χρησιμοποιείται συχνά για αναφορά της συνολικής απόδοσης επενδύσεων, όμως παρουσιάζει το πρόβλημα ότι δεν αντικατοπτρίζει πλήρως την αντιστροφή των τιμών, καθώς δεν λαμβάνει υπόψη τις συνεχείς και μικρές διακυμάνσεις στην αξία της μετοχής.

- Λογαριθμική Απόδοση (Logarithmic Return)

Η λογαριθμική απόδοση υπολογίζεται με βάση τον φυσικό λογάριθμο του πυλίκου της τρέχουσας τιμής της ποσότητας που μελετάται προς την προηγούμενη τιμή της ποσότητας που μελετάται. Αυτή η μέθοδος εκφράζει τις αποδόσεις σε λογαριθμική κλίμακα και η μαθηματική της έκφραση είναι η ακόλουθη:

$$\text{Log Return} = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (3.2)$$

όπου:

P_t : η τρέχουσα αξία της ποσότητας που μελετάται

P_{t-1} : η προηγούμενη αξία της ποσότητας που μελετάται

Η λογαριθμική απόδοση αντικατοπτρίζει καλύτερα τις αλλαγές στις τιμές της μετοχής και είναι χρήσιμη για την ανάλυση χρονοσειρών και την εκτίμηση του συνολικού κερδοσκοπικού δυναμικού. Παρόλα αυτά, μπορεί να είναι πιο δύσκολο να κατανοηθεί από τον μέσο επενδυτή λόγω της λογαριθμικής κλίμακας.

Και οι δύο μέθοδοι έχουν τη χρησιμότητά τους, αλλά για την παρούσα διπλωματική επιλέχθηκε η ανάλυση της λογαριθμικής απόδοσης.

Τα μοντέλα τεχνητής νοημοσύνης έχουν σημαντικό ρόλο στην ανάλυση και πρόβλεψη των αποδόσεων στη χρηματιστηριακή αγορά. Αυτά τα μοντέλα χρησιμοποιούνται για να εξάγουν προβλέψεις σχετικά με τις μελλοντικές τιμές των χρηματοοικονομικών εργαλείων, να διαχειριστούν το ρίσκο και να λάβουν αποφάσεις επενδυτικής στρατηγικής. Μερικά από τα συνήθη μοντέλα τεχνητής νοημοσύνης που χρησιμοποιούνται στη χρηματιστηριακή αγορά είναι τα παρακάτω:

- Νευρωνικά δίκτυα (Neural networks): Τα νευρωνικά δίκτυα είναι ένα από τα πιο δημοφιλή μοντέλα τεχνητής νοημοσύνης για την ανάλυση αποδόσεων. Χρησιμοποιούνται για τον προσδιορισμό μοτίβων και τάσεων στις τιμές των μετοχών και τη δημιουργία προβλέψεων
- Support vector machines - SVM: Τα SVM είναι μοντέλα που χρησιμοποιούνται για την κατηγοριοποίηση και πρόβλεψη των τιμών των μετοχών. Είναι ιδιαίτερα χρήσιμα για τον προσδιορισμό σημείων αντίστασης και υποστήριξης στις τιμές
- Συναρτησιακή παρεμβολή (Functional Approximation): Αυτή η τεχνική χρησιμοποιείται για την εκτίμηση των μελλοντικών τιμών με βάση την ανάλυση της ιστορικής απόδοσης και την παρεμβολή συναρτήσεων. Χρησιμοποιείται κυρίως παρεμβολή, extrapolation
- Σύνθετα μοντέλα (Ensemble Models): Αυτά τα μοντέλα συνδυάζουν πολλές διαφορετικές τεχνικές μοντελοποίησης για τη βελτιστοποίηση της ακρίβειας των προβλέψεων
- Μοντέλα βαθιάς μάθησης (Deep Learning models): Τα μοντέλα βαθιάς μάθησης, όπως τα παλινδρομικά νευρωνικά δίκτυα (Recurrent Neural Networks - RNN) και τα δίκτυα μακράς μνήμης (Long Short-Term Memory - LSTM), είναι ιδιαίτερα κατάλληλα για την ανάλυση χρονοσειρών, όπως οι αποδόσεις μετοχών

Στο πλαίσιο της μελέτης της διπλωματικής εργασίας αναπτύχθηκαν μοντέλα μηχανικής μάθησης με έμφαση στα νευρωνικά δίκτυα LSTM, MLP, RNN, όπως και GRU.

3.3 Γενικά γνωρίσματα της μελέτης της διπλωματικής εργασίας

Σε αυτή την ενότητα θα γίνει αναφορά στα εργαλεία που χρησιμοποιήθηκαν για την υλοποίηση της μελέτης της διπλωματικής εργασίας, καθώς τον στόχο και την δομή της.

3.3.1 Εργαλεία που χρησιμοποιήθηκαν

Η υλοποίηση αυτής της διπλωματικής εργασίας έγινε με την χρήση της γλώσσας προγραμματισμού R version 4.3.2 καθώς και με την χρήση της γλώσσας προγραμματισμού Python version 3.9 στην σχεδιασμένη πλατφόρμα για αναλύσεις δεδομένων Anaconda. Η χρήση αυτής της πλατφόρμας είναι επιβεβλημένη, καθώς για τις ανάγκες της ανάλυσης αυτής θα χρησιμοποιηθούν βιβλιοθήκες μηχανικής μάθησης, όπως το Keras και το Ternsoflow.

3.3.2 Στόχος της μελέτης

Ο στόχος της μελέτης είναι να εξερευνήσουμε τον τρόπο λειτουργίας των πιο βασικών αρχιτεκτονικών τεχνητών νευρωνικών δικτύων, καθώς και τα αποτελέσματα τα οποία εξάγουν, σε δεδομένα αποδόσεων που απορρέουν από πραγματικές τιμές μετοχών της Ελληνικής χρηματιστηριακής αγοράς.

3.3.3 Δομή της μελέτης

Η δομή της μελέτης θα χωριστεί στα ακόλουθα στάδια: Αρχικά θα γίνει εισαγωγή στα σύνολα δεδομένων (datasets), τα οποία θα μελετηθούν. Στη συνέχεια θα γίνει καθαρισμός δεδομένων, όπου και αν χρειάζεται. Θα ληφθούν οι λογαριθμικές αποδόσεις των τιμών κλεισίματος των μετοχών.

Στο δεύτερο στάδιο θα γίνει η επιλογή των αρχιτεκτονικών των τεχνητών νευρωνικών δικτύων που θα χρησιμοποιηθούν, θα υλοποιηθεί η κατασκευή τους, καθώς και θα γίνει η εξαγωγή και η αξιολόγηση των αποτελεσμάτων και της απόδοσης τους. Κλείνοντας στο τελευταίο στάδιο θα πραγματοποιηθεί παρουσίαση των αποτελεσμάτων των νευρωνικών δικτύων, όπως επίσης θα γίνει μετατροπή των λογαριθμικών αποδόσεων της πρόβλεψης σε τιμές κλεισίματος πρόβλεψης και θα παρουσιαστούν τα αντίστοιχα γραφήματα, με σκοπό να γίνουν πιο ξεκάθαρα τα αποτελέσματα.

3.4 Διερευνητική Ανάλυση

3.4.1 Επιλογή συνόλων δεδομένων

Για τις ανάγκες της μελέτης, χρησιμοποιήθηκαν τρία σύνολα δεδομένων από τρεις εταιρείες μεγαθήρια της Ελληνικής χρηματιστηριακής αγοράς. Από τον κλάδο της κατασκευαστής αγοράς επιλέχθηκε ο όμιλος της ΓΕΚ ΤΕΡΝΑ Α.Ε., από τον κλάδο των αεροπορικών εταιρειών επιλέχθηκε η ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε. (διακ: AEGEAN AIRLINES), καθώς και από τον κλάδο των εταιρειών τυχερών παιγνίων επιλέχθηκε ο ΟΠΑΠ Α.Ε.

Τα δεδομένα της μελέτης λήφθηκαν από τον διαδικτυακό ιστότοπο της οικονομικής εφημερίδας «Ναυτεμπορική» που ιδρύθηκε από τους αδελφούς Αθανασιάδη σηματοδοτώντας την εμφάνιση του οικονομικού τύπου στην Ελλάδα, το 1924. Αποτελεί την πρωτοπόρο εφημερίδα οικονομικού ενδιαφέροντος σε κυκλοφορία και αριθμό αναγνωστών (Wikipedia, n.d.). Οι παρατηρήσεις των τριών συνόλων δεδομένων ανήκουν στο διάστημα 12/04/2017 έως 03/11/2023. Κάθε σύνολο δεδομένων περιέχει επτά στήλες χαρακτηριστικών, οι οποίες περιέχουν από 1600 παρατηρήσεις εκάστη.

Οι επτά στήλες αντιπροσωπεύουν τα παρακάτω χαρακτηριστικά:

- Ημερομηνία
- Υψηλότερη τιμή μετοχής ανά ημέρα
- Χαμηλότερη τιμή μετοχής ανά ημέρα
- Τιμή ανοίγματος μετοχής ανά ημέρα
- Τιμή κλεισίματος μετοχής ανά ημέρα
- Όγκος διακίνησης μετοχής
- Τζίρος μετοχής σε €

3.4.2 Προεπεξεργασία δεδομένων

3.4.2.1 Εισαγωγή των δεδομένων

Η μορφή του αρχείου των δεδομένων που θα χρησιμοποιηθεί για επεξεργασία είναι σε μορφή csv. Για την επεξεργασία του θα χρησιμοποιήσουμε την γλώσσα προγραμματισμού Python και το περιβάλλον της Anaconda το Jupyter. Για να είναι πιο ομαλή η επεξεργασία του θα κάνουμε χρήση μιας βιβλιοθήκης της γλώσσας Python, την Pandas. Η βιβλιοθήκη Pandas είναι πάρα πολύ χρήσιμη και μερικά χαρακτηριστικά της είναι τα παρακάτω:

- Ταχύτητα και αποτελεσματικότητα στην επεξεργασία και ανάλυση δεδομένων
- Δυνατότητα επεξεργασίας μεγάλων σετ δεδομένων
- Περιέχει εργαλεία για εισαγωγή δεδομένων από διάφορους τύπους αρχείων

Παράλληλα κάνοντας χρήση της βιβλιοθήκης Pandas θα κάνουμε μετασχηματισμό των δεδομένων σε μορφή dataframe και αυτή η μορφή θα παραμείνει καθ' όλη την διάρκεια της ανάλυσης.

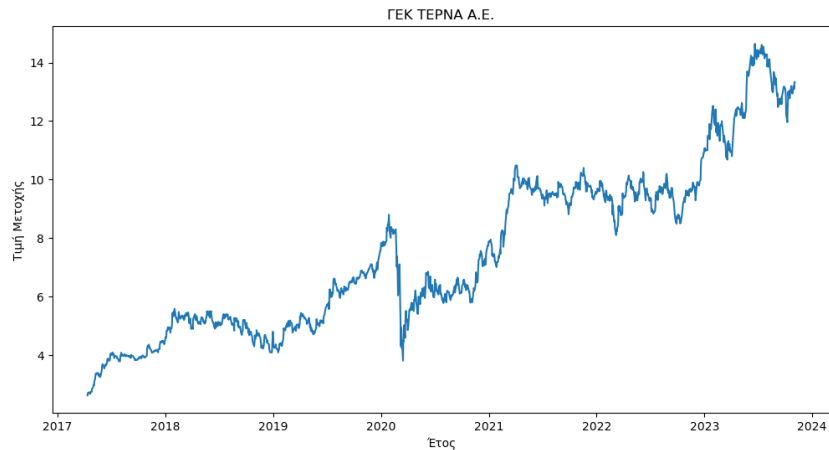
3.4.2.2. Έλεγχος ελλειπών τιμών

Ένας από τους πιο σημαντικούς ελέγχους στην προεπεξεργασία των δεδομένων είναι ο έλεγχος ελλειπών τιμών (missing values), διότι η ύπαρξη χαμένων τιμών προμηγνύει απώλεια πληροφορίας, ιδιαίτερα όταν το πλήθος τους είναι μεγάλο. Επίσης υπάρχει έντονος κίνδυνος για παραμόρφωση των αποτελεσμάτων οδηγώντας στην εξαγωγή λανθασμένων συμπερασμάτων. Υπάρχουν τεχνικές που αντικαθιστούν τις χαμένες τιμές, όπως η αντικατάσταση τους με τον μέσο όρο των παρατηρήσεων είτε με αντικατάσταση με την διάμεσο, είτε κάποιας δημοφιλής τεχνικής χαμένων τιμών, όπως ο KNN Imputer, όπου χρησιμοποιεί γειτονικές τιμές με κοινά χαρακτηριστικά, άλλωστε είναι εμπνευσμένος από τον αλγόριθμο K-Nearest Neighbors (KNN).

Ο έλεγχος των ελλειπών τιμών θα γίνει και πάλι με την χρήση της βιβλιοθήκης Pandas, όπου αποδεικνύεται πως δεν υπάρχουν χαμένες τιμές και στα τρία σύνολα δεδομένων. Αυτό το γεγονός είναι ιδιαίτερα σημαντικό, διότι η ανάλυση θα αρχίσει με τον καλύτερο δυνατό τρόπο.

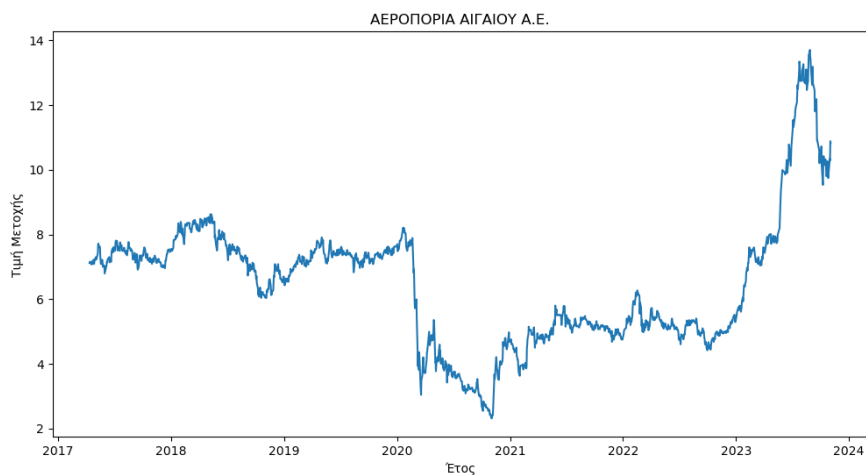
3.4.3 Οπτικοποίηση των τριών συνόλων δεδομένων

Στο Διάγραμμα 3.1 παρουσιάζεται η τιμή μετοχής της ΓΕΚ ΤΕΡΝΑ Α.Ε. για το χρονικό διάστημα από 12 Απριλίου 2017 έως 3 Νοεμβρίου 2023.



Διάγραμμα 3.1: Τιμή μετοχής της ΓΕΚ ΤΕΡΝΑ Α.Ε. από 12/04/2017 ως 03/11/2023

Στο Διάγραμμα 3.2 παρουσιάζεται η τιμή μετοχής της ΑΕΡΟΠΟΡΙΑΣ ΑΙΓΑΙΟΥ Α.Ε. για το χρονικό διάστημα από 12 Απριλίου 2017 έως 3 Νοεμβρίου 2023.



Διάγραμμα 3.2: Τιμή μετοχής της ΑΕΡΟΠΟΡΙΑΣ ΑΙΓΑΙΟΥ Α.Ε. από 12/04/2017 ως 03/11/2023

Στο Διάγραμμα 3.3 παρουσιάζεται η τιμή μετοχής του ΟΠΑΠ Α.Ε. για το χρονικό διάστημα από 12 Απριλίου 2017 έως 3 Νοεμβρίου 2023.



Διάγραμμα 3.3: Τιμή μετοχής της ΟΠΑΠ Α.Ε. από 12/04/2017 έως 3/11/2023

Τα δεδομένα, τα οποία έχουν ληφθεί για την ανάλυση της διπλωματικής εργασίας αφορούν αποδόσεις μετοχών κλεισίματος του Ελληνικού χρηματιστηρίου, για τρεις εταιρείες «μεγαθήρια» της Ελληνικής αγοράς. Η συχνότητα των τιμών αυτών είναι ανά ημέρα. Και στις τρεις εταιρείες παρατηρείται έντονα το φαινόμενο της τάσης.

Η πρώτη μετοχή, στην οποία θα αξιολογηθούν τα ποιοτικά της χαρακτηριστικά της είναι η μετοχή του Ομίλου της ΓΕΚ ΤΕΡΝΑ Α.Ε. Σύμφωνα με το Διάγραμμα 3.1 παρατηρείται πως το πρώτο τρίμηνο του 2020, που η παγκόσμια κοινότητα είχε πληγεί από την πανδημία του Covid-19, η τιμή της μετοχής της είχε πτωτική τάση, φτάνοντας στις 16 Μαρτίου 2020, η τιμή της να είναι 4,498. Στην συνέχεια η εταιρεία ανέκαμψε και η μετοχή της κατάφερε να έχει μια πολύ μεγάλη ανοδική πορεία για τα επόμενα χρόνια, φτάνοντας στις 17 Ιουλίου 2023 η τιμή μετοχής της να είναι € 14,425, δηλαδή 3,2 φορές μεγαλύτερη σε σχέση με την τιμή που είχε στις 16 Μαρτίου 2020.

Η δεύτερη μετοχή, στην οποία θα αξιολογηθούν τα ποιοτικά της χαρακτηριστικά είναι η μετοχή της Αεροπορίας Αιγαίου Α.Ε. Φαίνεται από το Διάγραμμα 3.2 ότι έχει αυξητική τάση, ενώ από αρχές του 2023, η τιμή της μετοχής της δείχνει να έχει λάβει μια «ξέφρενη ανοδική πορεία», όμως γίνεται αντιληπτό πως στις αρχές του 2022, από όταν ξέσπασε η πανδημία του Covid-19, η οποία δημιούργησε ένα πολύ μεγάλο πλήγμα στις αεροπορικές εταιρείες και όχι μόνο. Παρατηρείται μια μεγάλη ύφεση στην τιμή της μετοχής της, η οποία κυμαινόταν με σταθερό ρυθμό γύρω στα € 7,88 στις αρχές του έτους, ενώ στα τέλη

Νοεμβρίου η τιμή της είχε κατακυλήσει στα € 2,37. Στην συνέχεια όμως η μετοχή της είχε ανοδική πορεία μέχρι τον Σεπτέμβρη του 2023, όπου δείχνει να υπάρχει μια μικρή μείωση στην τιμή της. Σύμφωνα με την σελίδα Business daily η τιμή της μετοχής της Αεροπορίας Αιγαίου δεν οφείλεται μόνο στις κακές συνθήκες που βιώνει αυτήν την περίοδο η αγορά, αλλά και την πολεμική αναταραχή που έχει προκαλέσει ο πόλεμος στο Ισραήλ, ιδιαίτερα στον κλάδο των αερομεταφορών. Η Αεροπορία Αιγαίου Α.Ε., όπως και άλλες εταιρείες ανακοίνωσαν την διακοπή των δρομολογίων προς το Ισραήλ, για λόγους ασφαλείας.

Η τρίτη εταιρεία που θα αξιολογηθούν τα ποιοτικά της χαρακτηριστικά είναι η ΟΠΑΠ Α.Ε. Η τιμή της μετοχής της ΟΠΑΠ Α.Ε. έχει λάβει ραγδαία αύξηση δηλαδή η τάση είναι αυξητική τα τελευταία δύο χρόνια στρέφοντας το ενδιαφέρον αρκετών ξένων επενδυτών, όπως αντικατοπτρίζεται στο Διάγραμμα 3.3. Παράλληλα από το ίδιο διάγραμμα γίνεται κατανοητό, πως στο διάστημα του Μαρτίου 2020 έως τον Νοέμβριο του 2020 η πανδημία του Covid 19 είχε πολύ μεγάλο αντίκτυπο στην τιμή της μετοχής της εταιρείας, αφού στις 15 Μαρτίου του 2020 η τιμή της είχε φτάσει € 5,72. Η κακή πορεία της μετοχής συνεχίστηκε περίπου μέχρι τις αρχές Νοεμβρίου, ενώ στην συνέχεια ξεκίνησε την ανοδική της πορεία. Παρατηρείται το 2023, πως ήταν μια πολύ καλή χρονιά για τον ΟΠΑΠ Α.Ε., αφού κατάφερε να αυξήσει την τιμή της μετοχής του από € 13,1 ανά μετοχή την 1^η Ιανουαρίου 2023 στα 15,45 € ανά μετοχή την 3^η Νοεμβρίου 2023. Παρά το δύσκολο οικονομικό περιβάλλον που βιώνει η χώρα, λόγω της ακρίβειας, φαίνεται πως ο κόσμος έχει στραφεί προς τα τυχερά παίγνια. (σελ. Business Day)

3.4.4 Επαυξημένος έλεγχος των Dickey - Fuller

Με την βοήθεια των Διαγραμμάτων 3.1, 3.2 και 3.3 βγήκε ένα αρχικό συμπέρασμα, πως οι χρονοσειρές που θέλουμε να εξάγουμε συμπεράσματα δεν είναι στάσιμες, αφού υπάρχει το φαινόμενο της τάσης αλλά και της μη κανονικότητας. Με σκοπό να επιβεβαιωθεί ο αρχικός ισχυρισμός θα χρησιμοποιήσουμε, ίσως το πιο γνωστό στατιστικό έλεγχο για την ύπαρξη στασιμότητας, τον επαυξημένο έλεγχο των Dickey - Fuller (ADF test).

Πίνακας 3.1: Αποτελέσματα του επαυξημένου ελέγχου Dickey Fuller για τις τρεις εταιρείες.

ΓΕΚ ΤΕΡΝΑ Α.Ε.	ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	ΟΠΑΠ Α.Ε.
ADF Statistic: -1.31952	ADF Statistic: -2.34676	ADF Statistic: -2.07921
p-value: 0.6202	p-value: 0.1573	p-value: 0.2530
Critical Values:	Critical Values:	Critical Values:
1%: -3.43446	1%: -3.43446	1%: -3.43446
5%: -2.86335	5%: -2.86335	5%: -2.86335
10%: -2.56773	10%: -2.56773	10%: -2.56773

Για την εφαρμογή του ελέγχου στα δεδομένα έγινε χρήση της βιβλιοθήκης της γλώσσας προγραμματισμού R, την βιβλιοθήκη tseries και την συνάρτηση adfuller. Τα αποτελέσματα του ελέγχου παρουσιάζονται στον Πίνακα 3.1. Τα αποτελέσματα του ελέγχου έδειξαν, πως οι χρονοσειρές του ομίλου ΓΕΚ ΤΕΡΝΑ Α.Ε. και των εταιρειών ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε. και ΟΠΑΠ Α.Ε. δεν περιέχουν στασιμότητα, αφού το p-value του ελέγχου είναι μεγαλύτερο του επιπέδου σημαντικότητας $\alpha=5\%$.

3.4.5 Μετατροπή των τιμών κλεισίματος των μετοχών σε λογαριθμικές αποδόσεις

Η μετατροπή των τιμών κλεισίματος των μετοχών σε λογαριθμικές αποδόσεις πραγματοποιήθηκε με την βοήθεια της βιβλιοθήκης numpy, την πιο χρήσιμη βιβλιοθήκη αριθμητικών υπολογισμών στην γλώσσα προγραμματισμού Python, όπως και την Pandas.

Αρχικά χρησιμοποιήθηκε η εντολή pct_change, η οποία επιστρέφει ένα Dataframe με την ποσοστιαία διαφορά μεταξύ των τιμών της παρατήρησης την χρονική στιγμή t και της παρατήρησης την χρονική στιγμή t-1. Στην συνέχεια υπολογίστηκαν οι λογαριθμικές αποδόσεις με την βοήθεια της συνάρτησης np.log.

Σε αυτό το σημείο θα γίνει εκ νέου χρήση του επαυξημένου ελέγχου των Dickey - Fuller, αυτή την φορά στις λογαριθμικές αποδόσεις:

Πίνακας 3.2: Αποτελέσματα του επαυξημένου ελέγχου Dickey Fuller στις αποδόσεις των τριών εταιρειών

ΓΕΚ ΤΕΡΝΑ Α.Ε.	ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	ΟΠΑΠ Α.Ε.
ADF Statistic: -16.3868	ADF Statistic: -13.5052	ADF Statistic: -16.7767
p-value: 0.000	p-value: 0.000	p-value: 0.000
Critical Values:	Critical Values:	Critical Values:
1%: -3.43446	1%: -3.43446	1%: -3.43446
5%: -2.86335	5%: -2.86335	5%: -2.86335
10%: -2.56773	10%: -2.56773	10%: -2.56773

Τα αποτελέσματα του ελέγχου έδειξαν, πως οι αποδόσεις των χρονοσειρών των εταιρειών περιέχουν στασιμότητα, αφού το p-value του ελέγχου και στις τρεις εταιρείες είναι πάρα πολύ μικρό, συνεπώς απορρίπτεται η αρχική υπόθεση του ελέγχου πως η χρονοσειρά δεν είναι στάσιμη. Το γεγονός αυτό θα βοηθήσει τα νευρωνικά δίκτυα να πραγματοποιήσουν πιο ασφαλείς προβλέψεις.

3.5 Μοντελοποίηση Νευρωνικών Δικτύων

3.5.1 Keras Configuration

Το Keras API αποτελεί ένα από τα κορυφαία API στον κλάδο της μηχανικής μάθησης με ειδικευση στην βαθιά μάθηση. Ανήκει στην πλατφόρμα μηχανικής μάθησης Tensorflow. Μέσω αυτού του API, η διαδικασία για μοντελοποίησης νευρωνικών δικτύων έχει γίνει πολύ πιο εύχρηστη και πολύ προσιτή για τον χρήστη, αφού μέσα από αυτήν την πλατφόρμα χρησιμοποιήσει πολλά εργαλεία με σκοπό να κατασκευάζει και να επεξεργάζεται κάθε σκέλος των νευρωνικών δικτύων που επιθυμεί.

Υπάρχουν πολλές γλώσσες προγραμματισμού που χρησιμοποιούν το Keras API, όπως η Python, η R, η Javascript και Julia (Μελάς, 2020). Παράλληλα για την κατασκευή των νευρωνικών δικτύων το Keras χρησιμοποιεί διάφορες παραμέτρους οι οποίες είναι:

- Επιλογή μοντέλου (model)
- Επιλογή μεταγλωττιστή (compiler)
- Επιλογή συνάρτησης ενεργοποίησης (activation function)

- Επιλογή συνάρτησης βελτιστοποίησης (optimize function)
- Επιλογή συνάρτησης απώλειας (loss function)
- Επιλογή πληροφοριακού κριτηρίου σφάλματος (metrics)

Υπάρχουν δύο τρόποι δημιουργίας μοντέλων νευρωνικών δικτύων στο Keras:

- Ακολουθιακό μοντέλο (Sequential model)
- Λειτουργικό μοντέλο (Functional model)

Η ορθή επιλογή της συνάρτησης βελτιστοποίησης αποτελεί πολύ σημαντικό κομμάτι στην αποτελεσματικότητα του μοντέλου. Το Keras API έχει πολλές συναρτήσεις βελτιστοποίησης, που μπορούν να χρησιμοποιηθούν ανάλογα την φύση του προβλήματος. Οι πιο σημαντικοί είναι κάτωθι:

- SGD (Stochastic Gradient Descent):
- Adam (Adaptive Moment Estimation)
- RMSprop (Root Mean Square Propagation)
- Adagrad (Adaptive Gradient Algorithm)
- Adadelta
- Nadam

Επιπροσθέτως εξίσου σημαντική είναι και η επιλογή της κατάλληλης συνάρτησης ενεργοποίησης. Οι πιο σημαντικές είναι:

- ReLU (Rectified Linear Unit)
- Sigmoid (Σιγμοειδής)
- Tanh (Υπερβολική εφαπτομένη)
- LReLU (Leaky Rectified Linear Unit)
- ELU (Exponential Linear Unit)
- Softmax

Σπουδαίο ρόλο έχει και η παράμετρος fit, αφού δίνει την δυνατότητα επιλογής του αριθμού των εποχών (epochs), που θα εισέλθει μέσα στο δίκτυο ο αλγόριθμος, καθώς και το μέγεθος

της παρτίδας (batch size), δηλαδή τον διαχωρισμό του συνόλου δεδομένων σε παρτίδες, για την ομαλή διεξαγωγή της εκπαίδευσης.

3.5.2 Αλγόριθμοι που χρησιμοποιήθηκαν

1. LSTM

Τα LSTM (Long Short-Term Memory) δίκτυα είναι μια δημοφιλής επιλογή για τη μοντελοποίηση ακολουθιακών δεδομένων, λόγω της δυνατότητάς τους να καταγράφουν μακροπρόθεσμες εξαρτήσεις. Στο πλαίσιο των δεδομένων του χρηματιστηρίου, μπορούν να ανιχνεύσουν πολύπλοκα πρότυπα και σχέσεις με την πάροδο του χρόνου.

- **Πλεονεκτήματα:** Μπορούν να χειριστούν αποτελεσματικά μεγάλες ακολουθίες, ανιχνεύοντας μακροπρόθεσμες εξαρτήσεις στις τιμές των μετοχών.
- **Μειονεκτήματα:** Απαιτούν περισσότερους υπολογισμούς από απλούστερα μοντέλα, μπορούν να παρακολουθήσουν πολύ «σφιχτά» τα δεδομένα, και λόγω της πολυπλοκότητάς τους απαιτούν την ρύθμιση πολλών υπερ παραμέτρων. Παράλληλα όταν τα διαθέσιμα δεδομένα δεν είναι αρκετά, το μοντέλο κινδυνεύει από το φαινόμενο της υπερ προσαρμογής.

2. GRU

Τα GRU (Gated Recurrent Unit) δίκτυα είναι παρόμοια με τα LSTM δίκτυα, με την διαφορά, ότι διαθέτουν πιο απλούστερη αρχιτεκτονική. Συνήθως επιλέγονται, όταν οι υπολογιστικοί πόροι είναι περιορισμένοι ή όταν προτιμάται ένα απλούστερο μοντέλο.

- **Πλεονεκτήματα:** Απαιτεί λιγότερες παραμέτρους και υπολογισμούς σε σύγκριση με τα LSTM δίκτυα, γεγονός που μπορεί να οδηγήσει σε ταχύτερη εκπαίδευση. Επιπροσθέτως μπορούν να «κατανοήσουν» ακολουθιακές εξαρτήσεις αποτελεσματικά.
- **Μειονεκτήματα:** Υπάρχει πιθανότητα μη καταγραφής πολύ μακροπρόθεσμων εξαρτήσεων, όσο τα LSTM δίκτυα, όπως και ενδέχεται να μην είναι αρκετά αξιόπιστα σε πολύπλοκα μοτίβα δεδομένων.

3. RNN

Τα αναδρομικά νευρωνικά δίκτυα RNN είναι ο πιο απλός τύπος αναδρομικού νευρωνικού δικτύου και χρησιμοποιείται ως μοντέλο σύγκρισης σε δεδομένα χρονοσειρών, προκειμένου να συγκριθεί με ένα πιο πολύπλοκο μοντέλο, όπως το LSTM ή το GRU.

- Πλεονεκτήματα: Απλό και υπολογιστικά αποδοτικό, αποτελεί καλή βάση. Μπορεί να καταγράψει ακολουθιακές εξαρτήσεις στα δεδομένα.
- Μειονεκτήματα: Προσπαθεί να καταγράψει μόνο σύντομες ακολουθιακές εξαρτήσεις, και συχνά υποφέρει από το πρόβλημα της εξαφάνισης της κλίσης (vanishing gradient).

4. Πολυεπίπεδο Perceptron (MLP)

Το Πολυεπίπεδο Perceptron (MLP) ανήκει στα δίκτυα πρόσθιας τροφοδότησης (FNN). Το MLP υπόδειγμα μπορεί να χρησιμοποιηθεί για την πρόβλεψη χρονοσειρών ως μια διαφορετική προσέγγιση, αντί να χρησιμοποιηθεί κάποιος αλγόριθμος αναδρομικού μοντέλου. Παράλληλα το MLP δεν λαμβάνει υπόψη τις ακολουθιακές εξαρτήσεις εμπλέκοντας αντί αυτού μοντέλα προτύπων από τα εισερχόμενα δεδομένα.

- Πλεονεκτήματα: Έχει απλή μορφή και σε γενικές γραμμές είναι υπολογιστικά αποδοτικό. Μπορεί να καταγράψει γραμμικά και απλά μοτίβα στα δεδομένα και ενδέχεται να λειτουργεί με αξιοπιστία, εάν τα δεδομένα δεν έχουν ακολουθιακή δομή.
- Μειονεκτήματα: Δεν λαμβάνει υπόψη τις ακολουθιακές εξαρτήσεις στα δεδομένα, που είναι κρίσιμες στη κατανόηση του μοτίβου των δεδομένων. Ενδέχεται να μην καταγράψει αποτελεσματικά τα πολύπλοκα ή τα μοτίβα που δεν περιέχουν γραμμικότητα.

Η τελική επιλογή του καλύτερου μοντέλου εξαρτάται από ποικίλους παράγοντες. Δύο από τα πιο βασικά στοιχεία επιλογής του αποτελούν η πολυπλοκότητα του προβλήματος, όπως επίσης και οι διαθέσιμοι υπολογιστικοί πόροι. Σημαντικό δε στοιχείο αποτελεί η ικανότητα του να αντλεί την απαιτούμενη πληροφορία αποτελεσματικά. Σε δεδομένα που αφορούν μετοχές, αποδόσεις μετοχών κ.λπ.

Τα LSTM και GRU δίκτυα αποτελούν την πιο αξιόπιστη επιλογή αλγορίθμων. Το RNN δίκτυο αποτελεί απλούστερη εναλλακτική, ενώ το MLP δίκτυο μπορεί να χρησιμοποιηθεί για πιο απλές προβλέψεις. Στο υπό κεφάλαιο των αποτελεσμάτων θα πραγματοποιηθεί αξιολόγηση

των μοντέλων που χρησιμοποιήθηκαν, καθώς και θα παρουσιαστεί το μοντέλο με τα πιο αξιόπιστα αποτελέσματα.

3.5.3 Αλγόριθμοι βελτιστοποίησης

Στα πλαίσια των δοκιμών, εξετάστηκε η χρήση διαφόρων αλγορίθμων βελτιστοποίησης (optimizers) για την εκπαίδευση των νευρωνικών δικτύων, όπου στη συνέχεια επιλέχθηκαν οι πιο αποδοτικές για την πρόβλεψη των λογαριθμικών αποδόσεων, των τιμών κλεισίματος, των μετοχών στις τρεις εταιρείες, που μελετήθηκαν.

Κάθε αλγόριθμος βελτιστοποίησης ενημερώνει τα βάρη του δικτύου με διαφορετικό τρόπο κατά την εκπαίδευση, και αυτό μπορεί να επηρεάσει την απόδοση του μοντέλου. Χρησιμοποιήθηκαν οι εξής συναρτήσεις βελτιστοποίησης:

1. RMSprop (Root Mean Square Propagation): Ο RMSprop είναι ένας αλγόριθμος που διατηρεί ένα κύριο βάρος για κάθε παράμετρο του δικτύου και ενημερώνει τα βάρη, βάσει της τετραγωνικής μέσης τιμής των προηγούμενων κλίσεων. Αυτό βοηθά στην αποφυγή του πολύ μεγάλου ή πολύ μικρού βήματος.
2. Adam (Adaptive Moment Estimation): Ο αλγόριθμος βελτιστοποίησης Adam είναι μια αποδοτική και δημοφιλής επιλογή. Συνδυάζει την προσαρμογή του ρυθμού μάθησης του RMSprop προσθέτοντας έναν όρο Ορμής (Momentum) για αποδοτική ενημέρωση των βαρών. Παρουσιάζει καλή απόδοση σε πολλά προβλήματα.
3. SDG (Stochastic Gradient Descent): Ο αλγόριθμος βελτιστοποίησης SGD πραγματοποιεί τυχαία επιλογή από μικρά υποσύνολα δεδομένων (παρτίδες), αντί να γίνεται χρήση ολόκληρου συνόλου δεδομένων. Έτσι επιλέγεται μικρός αριθμός δειγμάτων. Στη συνέχεια γίνεται ανακατανομή των δεδομένων σε κάθε επανάληψη.
4. Adagrad (Adaptive Gradient Algorithm): Ο αλγόριθμος Adagrad προσαρμόζει τον ρυθμό μάθησης για κάθε παράμετρο βάσει της συχνότητας ενημέρωσης της παραμέτρου αυτής. Αυτό τον καθιστά κατάλληλο για προβλήματα με ανομοιογενή δεδομένα.
5. Adadelat (Adaptive Learning Rate Method): Ο αλγόριθμος Adadelat είναι μια εξέλιξη του Adagrad που ελέγχει τον ρυθμό μάθησης με μια κυλιόμενη κανονική κίνηση των γραμμικών κλίσεων. Αυτό τον καθιστά ακόμα πιο κατάλληλο για την εκπαίδευση δικτύων.

6. Nadam: Ο αλγόριθμος Nadam είναι μια συνδυαστική επιλογή των αλγορίθμων Nesterov Accelerated Gradient (NAG) και Adam. Αυτό τον καθιστά αποδοτικό σε προβλήματα που απαιτούν γρήγορη σύγκλιση.

Κάθε αλγόριθμος βελτιστοποίησης έχει τα πλεονεκτήματά της και τους περιορισμούς της. Η επιλογή της καλύτερης συνάρτησης εξαρτάται από την φύση του κάθε προβλήματος, καθώς και την αρχιτεκτονική του μοντέλου. Στην εργασία μας, εκτιμήσαμε την απόδοση των διάφορων αλγορίθμων βελτιστοποίησης στην πρόβλεψη αποδόσεων μετοχών από τρεις εταιρείες της Ελληνικής χρηματιστηριακής αγοράς, για να επιλέξουμε την κατάλληλη στρατηγική βελτιστοποίησης που θα κάνει τον αλγόριθμο πιο αποτελεσματικό.

3.5.4 Πληροφοριακά κριτήρια σφάλματος

Κατά τη διάρκεια της αξιολόγησης των μοντέλων έγινε χρήση τριών διαφορετικών πληροφοριακών κριτηρίων σφάλματος, τα οποία είναι:

- Μέσο Τετραγωνικό Σφάλμα (MSE)
- Ρίζα Μέσου Τετραγωνικού Σφάλματος (RMSE)

Το μέσο τετραγωνικό σφάλμα ορίζεται ως το άθροισμα των τετραγώνων των σφαλμάτων διαιρεμένο με το πλήθος του αριθμού των προβλέψεων και τρόπος υπολογισμού του είναι ο ακόλουθος:

$$MSE = \frac{1}{n} \sum_{t=1}^n \varepsilon_t^2 \quad (3.3)$$

Ως ρίζα μέσου τετραγωνικού σφάλματος (RMSE) ορίζεται η τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n \varepsilon_t^2} \quad (3.4)$$

3.6 Μελέτη των τριών εταιρειών

3.6.1 Αρχική φάση της μελέτης

Η εύρεση των αποτελεσματικότερων αλγορίθμων βελτιστοποίησης αποτελεί μια πολύ κρίσιμη παράμετρο για την ορθή εκπαίδευση των νευρωνικών δικτύων. Στην διπλωματική εργασία αυτή χρησιμοποιήθηκαν έξι διαφορετικοί βελτιστοποιητές, οι οποίοι θα αξιολογηθούν βάσει των πληροφοριακών κριτηρίων σφάλματος MSE και RMSE. Τα έξι αυτά μοντέλα που θα εκπαιδευτούν θα έχουν ακριβώς τις ίδιες παραμέτρους και το χαρακτηριστικό που θα αλλάζει κάθε φορά θα είναι μόνο ο αλγόριθμος βελτιστοποίησης. Το σύνολο δεδομένων θα χωριστεί σε σύνολο εκπαίδευσης (training set) που θα αποτελεί το 80% του συνόλου και σύνολο ελέγχου (test set) που θα αποτελεί το 20% του συνόλου των δεδομένων. Το σύνολο δεδομένων που θα μελετηθεί περιέχει τις λογαριθμικές αποδόσεις από 1600 τιμές. Ο ρυθμός εισαγωγής των δεδομένων (batch size) στο δίκτυο θα ισούται με 64. Η συνάρτηση ενεργοποίησης (activation function) θα είναι η ReLU. Ύστερα από πολλές δοκιμές τα καλύτερα αποτελέσματα παράγονται, όταν ο αριθμός των νευρώνων ισούται με 100. Είναι πολύ σημαντικό να τονιστεί, πως λόγω της φύσης των λογαριθμικών αποδόσεων των συγκεκριμένων χρονοσειρών, όπου οι τιμές τους δεν έχουν μεγάλες διακυμάνσεις δεν είναι απαραίτητο να χρησιμοποιηθεί κάποια συνάρτηση κανονικοποίησης, αφού ελέγχθηκε η μέγιστη και ελάχιστη τιμή τους.

Για το τελικό στάδιο της μελέτης θα ληφθούν τα δύο μοντέλα, των οποίων τα κριτήρια σφάλματος έχουν την μικρότερη τιμή και στην συνέχεια θα εκπαιδευτούν με διαφορετικές παραμέτρους, όπου θα αναδειχθεί το βέλτιστο μοντέλο για την κάθε εταιρεία για κάθε αλγόριθμο ξεχωριστά.

3.6.2 Μελέτη και παρουσίαση αποτελεσμάτων για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

3.6.2.1 Εύρεση βέλτιστων αλγορίθμων βελτιστοποίησης για το LSTM

Η πρώτη αρχιτεκτονική νευρωνικών δικτύων που θα εφαρμοσθεί είναι αυτή των δικτύων μακράς βραχύχρονης μνήμης (LSTM). Αρχικά θα γίνει μετατροπή των τιμών κλεισίματος των μετοχών σε λογαριθμικές αποδόσεις. Θα ορισθούν κάποιες σταθερές παράμετροι με σκοπό να αναδειχθούν οι καλύτεροι βελτιστοποιητές βάσει των πληροφοριακών κριτηρίων σφάλματος.

Στη συνέχεια θα πραγματοποιηθεί η μετατροπή τους εκ νέου σε τιμές κλεισίματος, όπου θα πραγματοποιηθεί η σύγκριση τους

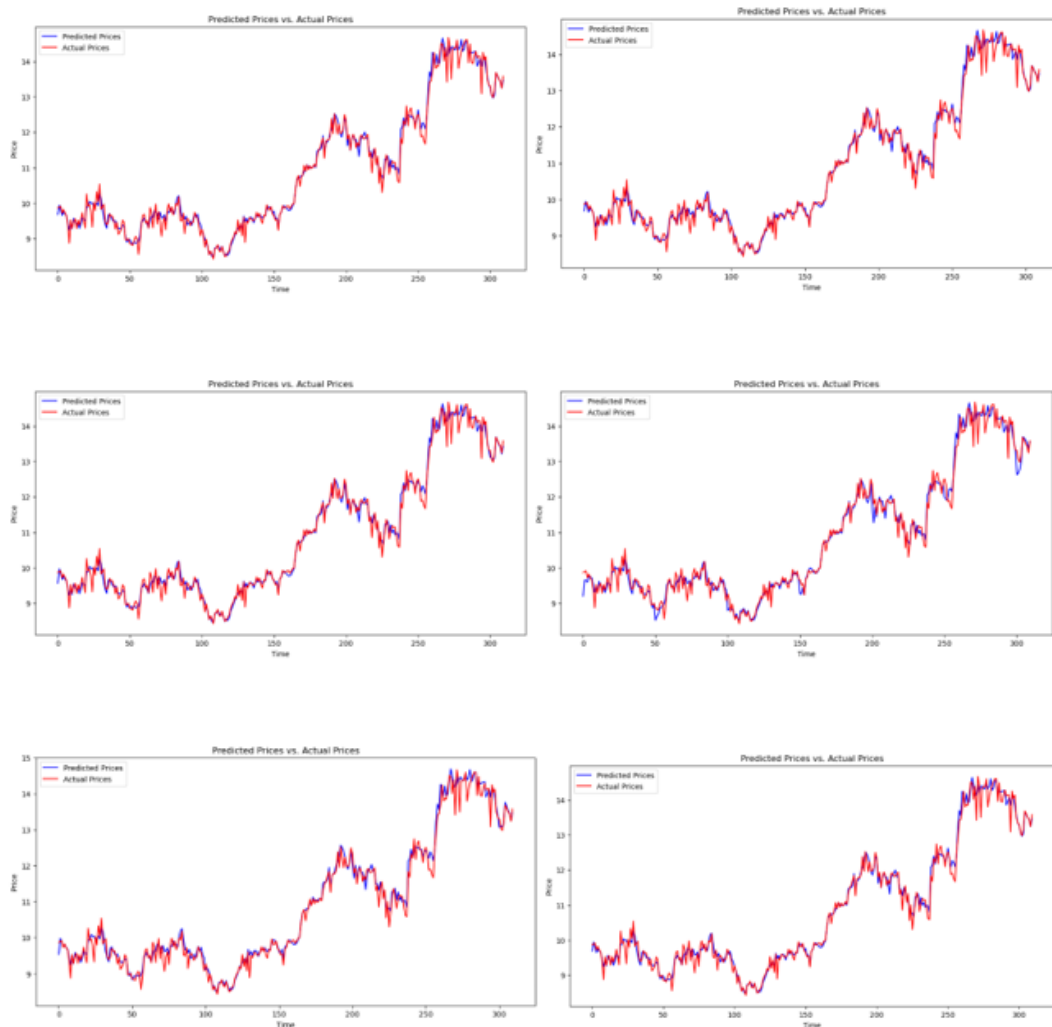
Στον Πίνακα 3.3 παρουσιάζονται τα αποτελέσματα των έξι αλγορίθμων βελτιστοποίησης, όπου έχουν σταθερή τιμή στις παραμέτρους batch size, που ισούται με 64 και την παράμετρο epochs που ισούται με 100.

Πίνακας 3.3: Αποτελέσματα των κριτηρίων σφάλματος του LSTM δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Neural Network	Optimizer	Metrics
LSTM	Adam	MSE: 0.000293 RMSE: 0.017127
LSTM	SGD	MSE: 0.000304 RMSE: 0.017461
LSTM	RMSprop	MSE: 0.000299 RMSE: 0.017302
LSTM	Adagrad	MSE: 0.000376 RMSE: 0.019399
LSTM	Adadelta	MSE: 0.000331 RMSE: 0.018200
LSTM	Nadam	MSE: 0.000289 RMSE: 0.017007

Αυτό που παρατηρείται στον Πίνακα 3.3 είναι ότι για το δίκτυο LSTM καλύτερη συμπεριφορά παρουσιάζουν οι συναρτήσεις βελτιστοποίησης Adam και Nadam, σύμφωνα με τα κριτήρια αξιολόγησης MSE και RMSE.

Στη συνέχεια έγινε μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε τιμές κλεισίματος, όπου παρουσιάστηκαν τα αντίστοιχα διαγράμματα για να γίνει πιο κατανοητή η σύγκριση μεταξύ των πραγματικών τιμών και των προβλεπόμενων. Η σειρά παρουσίασης των διαγραμμάτων είναι ίδια με την σειρά χρήσης των αλγορίθμων βελτιστοποίησης, από τα πάνω προς τα κάτω, δηλαδή η ακόλουθη σειρά είναι: Adam, SGD, RMSPROP, Adagrad, Adadelta, Nadam.



Διάγραμμα 3.4: Διαγράμματα σύγκρισης μεταξύ πραγματικών και προβλεπόμενων τιμών κλεισίματος του δικτύου LSTM, για τις έξι συναρτήσεις βελτιστοποίησης

3.6.2.2 Εύρεση βέλτιστων αλγορίθμων βελτιστοποίησης για το GRU

Αντίστοιχη διαδικασία, όπως στο LSTM πραγματοποιείται και στο νευρωνικό δίκτυο GRU κάνοντας χρήση των έξι αλγορίθμων βελτιστοποίησης, κάνοντας χρήση των ίδιων παραμέτρων σε σχέση με το LSTM δίκτυο.

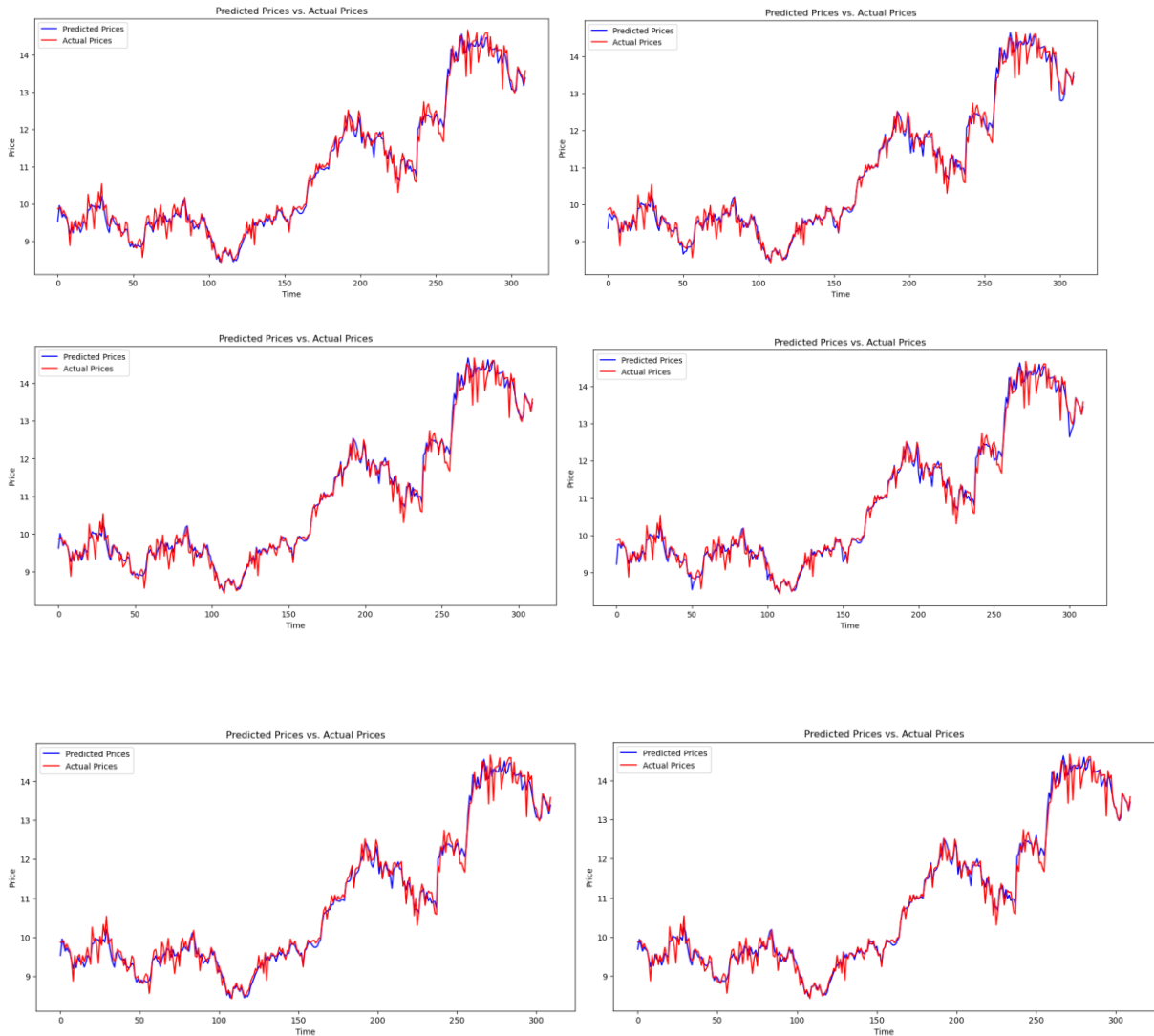
Στον Πίνακα 3.4 παρουσιάζονται τα αποτελέσματα των έξι αλγορίθμων βελτιστοποίησης, όπου έχουν σταθερή τιμή στις παραμέτρους batch size, που ισούται με 64 και την παράμετρο epochs που ισούται με 100.

Πίνακας 3.4: Αποτελέσματα των κριτηρίων σφάλματος του GRU δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Neural Network	Optimizer	Metrics
GRU	Adam	MSE: 0.000284 RMSE: 0.016861
GRU	SGD	MSE: 0.000323 RMSE: 0.017985
GRU	RMSprop	MSE: 0.000305 RMSE: 0.017483
GRU	Adagrad	MSE: 0.000345 RMSE: 0.018579
GRU	Adadelta	MSE: 0.000298 RMSE: 0.011726
GRU	Nadam	MSE: 0.000286 RMSE: 0.016939

Γίνεται αντιληπτό ότι και για το GRU δίκτυο καλύτερη συμπεριφορά παρουσιάζουν οι συναρτήσεις Adam και Nadam βάσει των κριτηρίων σφάλματος MSE και RMSE.

Στο Διάγραμμα 3.5 παρουσιάζονται τα διαγράμματα της μετατροπής των λογαριθμικών τιμών κλεισίματος σε τιμές κλεισίματος της Αεροπορίας Αιγαίου Α.Ε., όπου πραγματοποιείται σύγκριση μεταξύ πραγματικών και προβλέψιμων τιμών κλεισίματος, με την σειρά εμφάνισης να είναι ίδια με την σειρά χρήσης των αλγορίθμων βελτιστοποίησης, δηλαδή αντιστοιχούν στην ακόλουθη σειρά: Adam, SGD, RMSPROP, Adagrad, Adadelta, Nadam.



Διάγραμμα 3.5: Διαγράμματα σύγκρισης μεταξύ πραγματικών και προβλεπόμενων τιμών κλεισίματος του δικτύου GRU, για τις έξι συναρτήσεις βελτιστοποίησης

3.6.2.3 Εύρεση βέλτιστων αλγορίθμων βελτιστοποίησης για το RNN

Ακριβώς η ίδια διαδικασία, όπως στο LSTM και GRU δίκτυο πραγματοποιείται και στο νευρωνικό δίκτυο RNN κάνοντας χρήση των έξι αλγορίθμων βελτιστοποίησης, χρησιμοποιώντας τις ίδιες παραμέτρους.

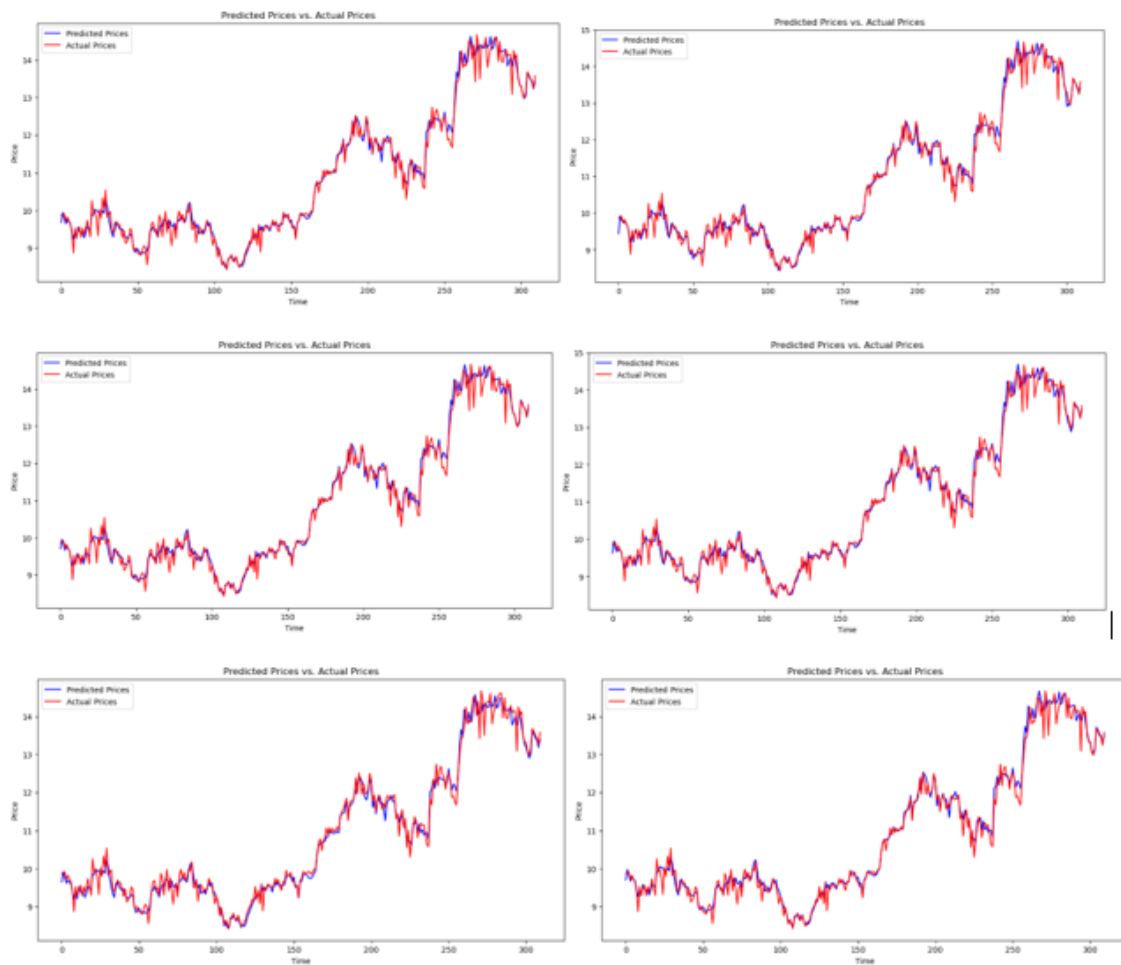
Στον Πίνακα 3.5 παρουσιάζονται τα αποτελέσματα των κριτηρίων σφάλματος MSE και RMSE για τις έξι συναρτήσεις βελτιστοποίησης στο παλινδρομικό νευρωνικό δίκτυο RNN.

Πίνακας 3.5: Αποτελέσματα των κριτηρίων σφάλματος του RNN δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Neural Network	Optimizer	Metrics
RNN	Adam	MSE: 0.000288 RMSE: 0.016986
RNN	SGD	MSE: 0.000314 RMSE: 0.017728
RNN	RMSprop	MSE: 0.000297 RMSE: 0.017257
RNN	Adagrad	MSE: 0.000295 RMSE: 0.017176
RNN	Adadelta	MSE: 0.000298 RMSE: 0.017263
RNN	Nadam	MSE: 0.000286 RMSE: 0.016939

Γίνεται αντιληπτό ότι και για το RNN δίκτυο καλύτερη συμπεριφορά παρουσιάζουν οι συναρτήσεις Adam και Nadam βάσει των κριτηρίων σφάλματος MSE και RMSE.

Ακολούθως στο Διάγραμμα 3.6 παρουσιάζονται τα διαγράμματα της μετατροπής των λογαριθμικών τιμών κλεισίματος σε τιμές κλεισίματος της ΓΕΚ ΤΕΡΝΑ Α.Ε., όπου πραγματοποιείται σύγκριση μεταξύ πραγματικών και προβλέψιμων τιμών κλεισίματος, με την σειρά εμφάνισης να είναι ίδια με την σειρά χρήσης των αλγορίθμων βελτιστοποίησης, δηλαδή αντιστοιχούν στην ακόλουθη σειρά: Adam, SGD, RMSPROP, Adagrad, Adadelta, Nadam.



Διάγραμμα 3.6: Διαγράμματα σύγκρισης μεταξύ πραγματικών και προβλεπόμενων τιμών κλεισίματος του δικτύου RNN, για τις έξι συναρτήσεις βελτιστοποίησης

3.6.2.4 Εύρεση βέλτιστων αλγορίθμων βελτιστοποίησης για το MLP

Η ίδια διαδικασία, όπως και στα προηγούμενα δίκτυα θα ακολουθηθεί και σε αυτόν τύπο νευρωνικού δικτύου.

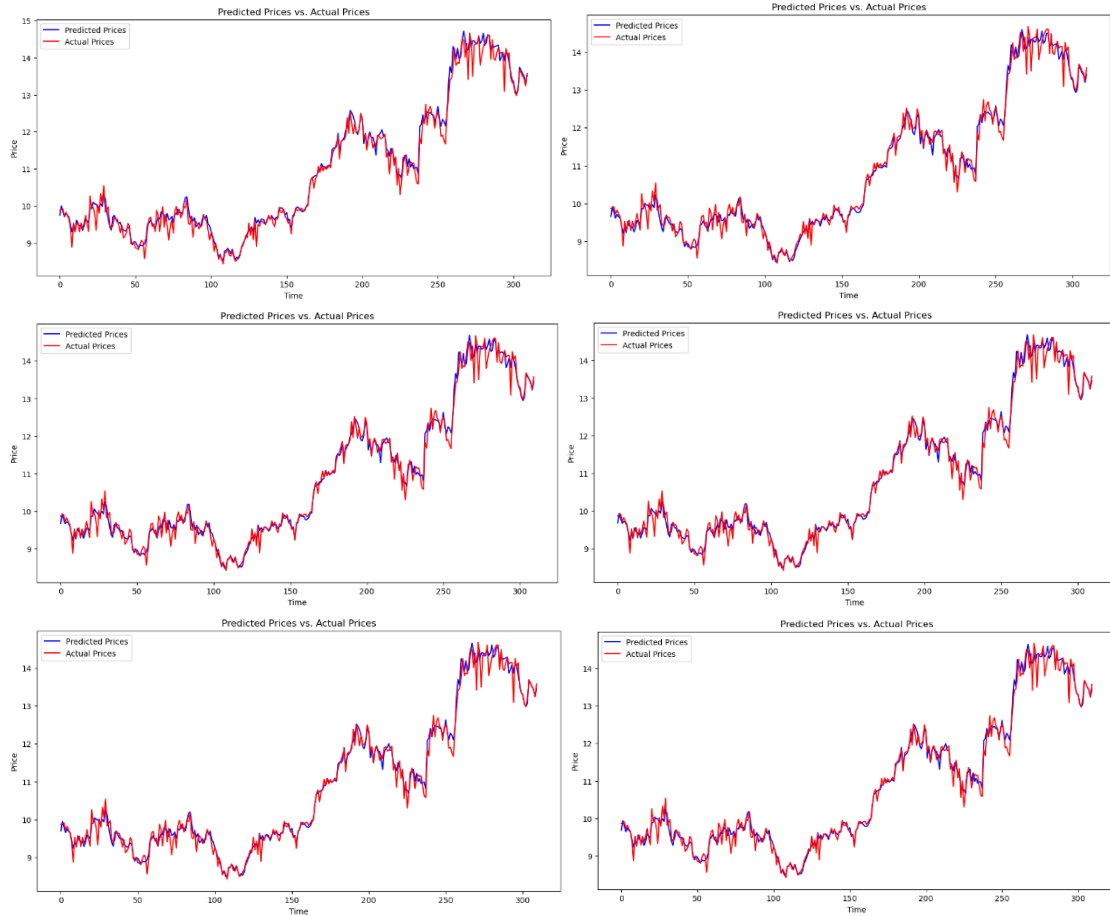
Στον Πίνακα 3.6 παρουσιάζονται τα αποτελέσματα των κριτηρίων σφάλματος MSE και RMSE για τις έξι συναρτήσεις βελτιστοποίησης στο Πολυεπίπεδο Perceptron (MLP).

Πίνακας 3.6: Αποτελέσματα των κριτηρίων σφάλματος του MLP δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Neural Network	Optimizer	Metrics
MLP	Adam	MSE: 0.000285 RMSE: 0.016888
MLP	SGD	MSE: 0.000289 RMSE: 0.017027
MLP	RMSprop	MSE: 0.000286 RMSE: 0.016911
MLP	Adagrad	MSE: 0.000294 RMSE: 0.017146
MLP	Adadelat	MSE: 0.000298 RMSE: 0.017092
MLP	Nadam	MSE: 0.000288 RMSE: 0.016970

Είναι εμφανές ότι στην περίπτωση των MLP δικτύων, οι δύο πιο αποτελεσματικές συναρτήσεις βελτιστοποίησης είναι οι Adam και RMSprop, βάσει των κριτηρίων σφάλματος MSE, RMSE. Άλλωστε η φύση της αρχιτεκτονικής των MLP's είναι τελείως διαφορετική σε σχέση με αυτές των προηγούμενων δικτύων, που ανήκουν στα παλινδρομικά δίκτυα.

Στη συνέχεια, στο Διάγραμμα 3.7 παρουσιάζονται τα διαγράμματα της μετατροπής των λογαριθμικών τιμών κλεισίματος σε τιμές κλεισίματος της ΓΕΚ ΤΕΡΝΑ Α.Ε., όπου πραγματοποιείται σύγκριση μεταξύ πραγματικών και προβλέψιμων τιμών κλεισίματος, με την σειρά εμφάνισης να είναι ίδια με την σειρά χρήσης των αλγορίθμων βελτιστοποίησης, δηλαδή αντιστοιχούν στην ακόλουθη σειρά: Adam, SGD, RMSPROP, Adagrad, Adadelat, Nadam.



Διάγραμμα 3.7: Διαγράμματα σύγκρισης μεταξύ πραγματικών και προβλεπόμενων τιμών κλεισίματος του δικτύου MLP, για τις έξι συναρτήσεις βελτιστοποίησης

Τα διαγράμματα επιβεβαιώνουν τα αποτελέσματα από τους παραπάνω πίνακες, αφού γίνεται αντιληπτό, ότι οι δύο καλύτερες συναρτήσεις βελτιστοποίησης είναι ο Adam και ο Nadam, αφού εξάγουν αποτελέσματα με το μικρότερο σφάλμα σύμφωνα με τα κριτήρια σφάλματος MSE και RMSE, όταν ο αλγόριθμος που εξετάζεται αφορά παλινδρομικό δίκτυο (RNN, LSTM, GRU), ενώ όταν ο αλγόριθμος που επιλέγεται αφορά MLP δίκτυο, οι συναρτήσεις που εξάγουν τα καλύτερα αποτελέσματα είναι οι: Adam και SDG.

3.6.2.5 Εύρεση βέλτιστων παραμέτρων για κάθε νευρωνικό δίκτυο

Για την εύρεση των κατάλληλων παραμέτρων των νευρωνικών δικτύων, έγινε επιλογή των δύο καλύτερων αλγορίθμων βελτιστοποίησης για κάθε νευρωνικό δίκτυο ξεχωριστά. Για τα παλινδρομικά δίκτυα RNN, GRU και LSTM, επιλέχθηκαν οι συναρτήσεις βελτιστοποίησης

Adam και Nadam, ενώ για το MLP δίκτυο, οι συναρτήσεις ενεργοποίησης Adam και SDG. Στη συνέχεια αυτές οι συναρτήσεις βελτιστοποίησης θα δοκιμαστούν σε πολλαπλές επιλογές στον αριθμό των εποχών (epochs) για κάθε περίπτωση νευρωνικού δικτύου. Ο αριθμός των epochs που θα δοκιμαστούν είναι: 100, 200 και 400.

3.6.2.6 LSTM για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Στον Πίνακα 3.7 παρουσιάζονται τα αποτελέσματα για το νευρωνικό δίκτυο LSTM, κάνοντας χρήση των δύο καλύτερων αλγορίθμων βελτιστοποίησης, τα αποτελέσματα, των οποίων βρίσκονται στον Πίνακα 3.4, όπως επίσης εναλλάσσοντας και των αριθμό των εποχών (epochs).

Πίνακας 3.7 Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο LSTM δίκτυο για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

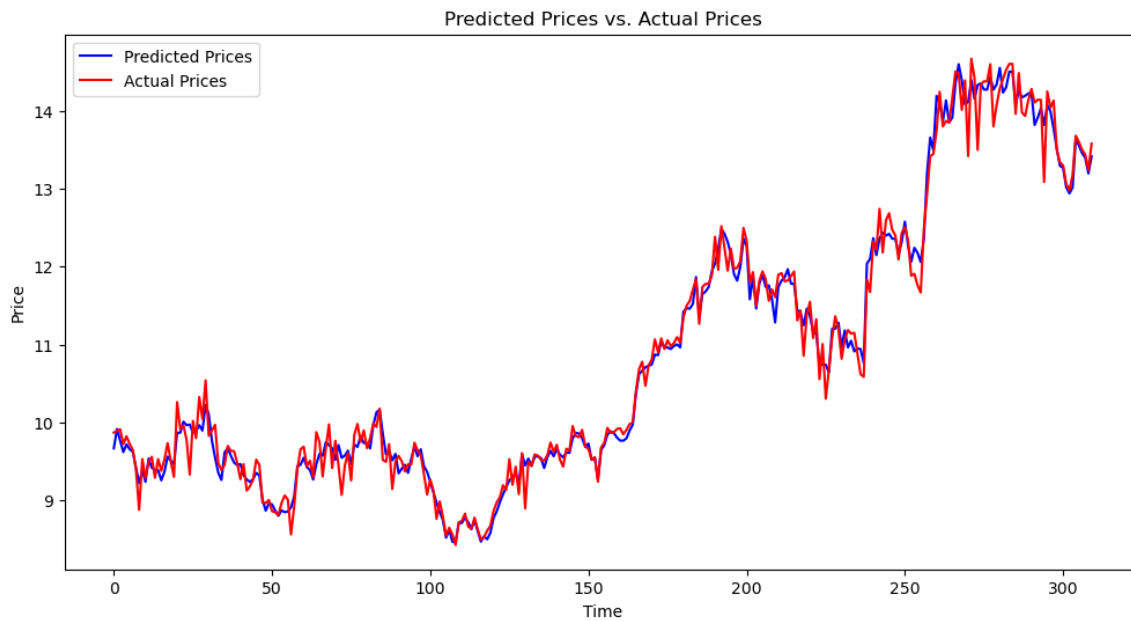
Test Number	Neural Network	Optimizer	Epochs	Metrics
1	LSTM	Adam	100	MSE: 0.000293370 RMSE: 0.021712085
2	LSTM	Adam	200	MSE: 0.000284917 RMSE: 0.016879504
3	LSTM	Adam	400	MSE: 0.000284262 RMSE: 0.016860079
4	LSTM	Nadam	100	MSE: 0.000284924 RMSE: 0.017007257
5	LSTM	Nadam	200	MSE: 0.000286331 RMSE: 0.016921339
6	LSTM	Nadam	400	MSE: 0.000284860 RMSE: 0.016877810

Από τον Πίνακα 3.7 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου LSTM είναι το μοντέλο 3. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000284262, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Adam, καθώς και τα epochs του 400. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Στη συνέχεια πραγματοποιείται μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές

κλεισίματος της ΓΕΚ ΤΕΝΡΑ Α.Ε.. Το σύνολο ελέγχου περιλαμβάνει 320 παρατηρήσεις, όσο και το 20% του συνόλου δεδομένων.

Στο Διάγραμμα 3.8 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 3, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.8: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το LSTM δίκτυο της ΓΕΚ ΤΕΡΝΑ Α.Ε.

3.6.2.7 GRU για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Ο Πίνακας 3.8 παρουσιάζει τα αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου GRU δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

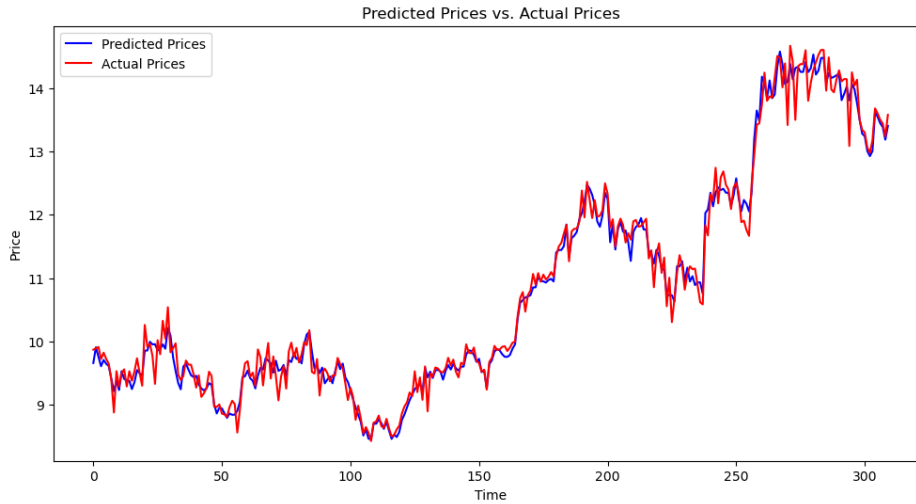
Πίνακας 3.8: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο GRU δίκτυο για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Test Number	Neural Network	Optimizer	Epochs	Metrics
1	GRU	Adam	100	MSE: 0.000284316 RMSE: 0.016861673
2	GRU	Adam	200	MSE: 0.000286923 RMSE: 0.016938821
3	GRU	Adam	400	MSE: 0.000284109 RMSE: 0.016865553
4	GRU	Nadam	100	MSE: 0.000286944 RMSE: 0.016939421
5	GRU	Nadam	200	MSE: 0.000286331 RMSE: 0.016921339
6	GRU	Nadam	400	MSE: 0.000284281 RMSE: 0.016860642

Από τον Πίνακα 3.8 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου GRU είναι το μοντέλο 3. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000284109, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Adam, καθώς και τα epochs του 400. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.9 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 3, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.9: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το GRU δίκτυο της ΓΕΚ ΤΕΡΝΑ Α.Ε.

3.6.2.8 RNN για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Ο Πίνακας 3.9 παρουσιάζει τα αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου RNN δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Πίνακας 3.9: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο RNN δίκτυο για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

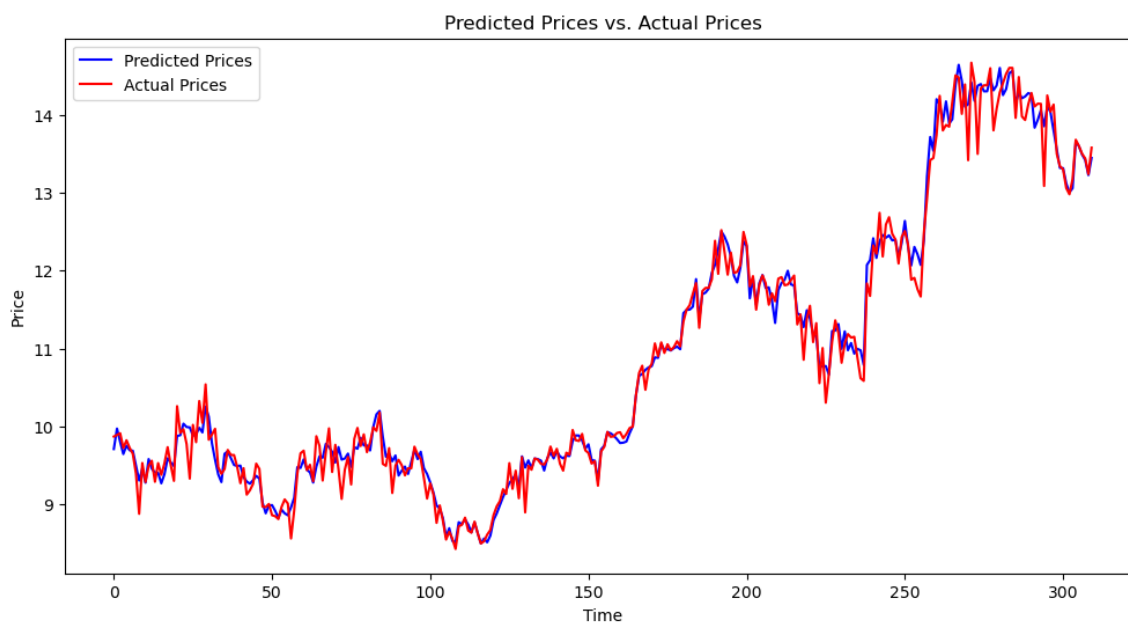
Test Number	Neural Network	Optimizer	Epochs	Metrics
1	RNN	Adam	100	MSE: 0.000288549 RMSE: 0.016867372
2	RNN	Adam	200	MSE: 0.000306865 RMSE: 0.017517556
3	RNN	Adam	400	MSE: 0.000284669 RMSE: 0.016872136
4	RNN	Nadam	100	MSE: 0.000284860 RMSE: 0.068878103
5	RNN	Nadam	200	MSE: 0.000292262 RMSE: 0.017095675
6	RNN	Nadam	400	MSE: 0.000291423 RMSE: 0.017071144

Από τον Πίνακα 3.9 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου RNN είναι το μοντέλο 3. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου

ισούται με 0.000284669, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Adam, καθώς και τα epochs του 400. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.10 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 3, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.10: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το RNN δίκτυο της ΓΕΚ ΤΕΡΝΑ Α.Ε.

3.6.2.9 MLP για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Ο Πίνακας 3.10 παρουσιάζει τα αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου MLP δικτύου για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

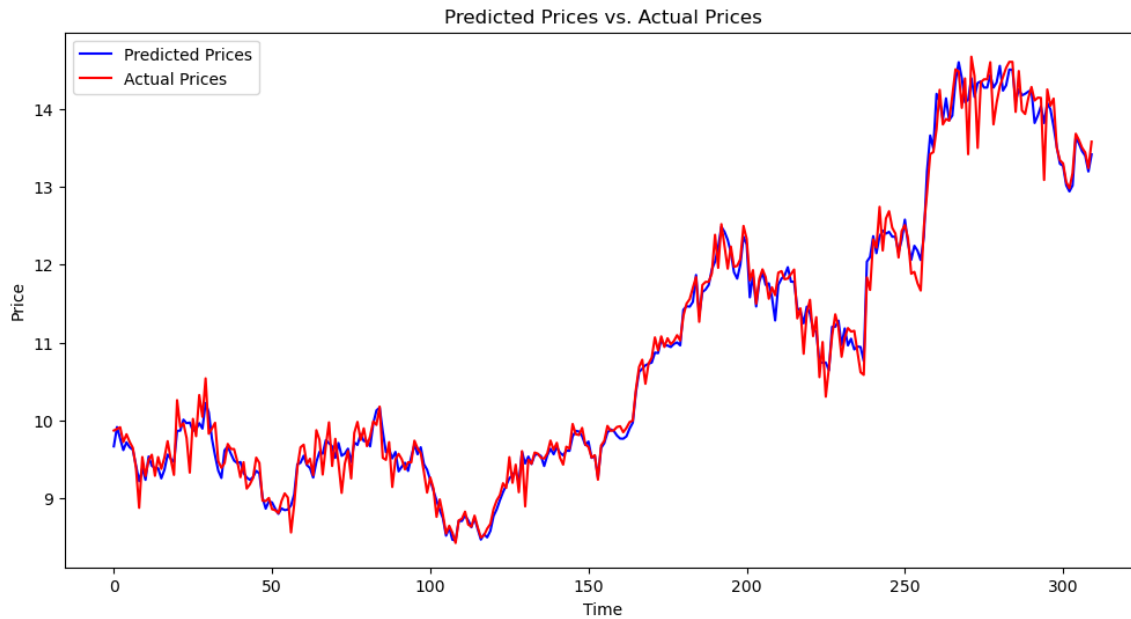
Πίνακας 3.10: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο MLP δίκτυο για την ΓΕΚ ΤΕΡΝΑ Α.Ε.

Test Number	Neural Network	Optimizer	Epochs	Metrics
1	MLP	Adam	100	MSE: 0.000285222 RMSE: 0.016888527
2	MLP	Adam	200	MSE: 0.000290934 RMSE: 0.016861310
3	MLP	Adam	400	MSE: 0.000284303 RMSE: 0.016861310
4	MLP	RMSprop	100	MSE: 0.000291453 RMSE: 0.017071995
5	MLP	RMSprop	200	MSE: 0.000290351 RMSE: 0.017039711
6	MLP	RMSprop	400	MSE: 0.000284373 RMSE: 0.016863375

Από τον Πίνακα 3.10 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου MLP είναι το μοντέλο 3. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000284303, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Adam, καθώς και τα epochs του 400. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.11 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 3, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.11: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το MLP δίκτυο της ΓΕΚ ΤΕΡΝΑ Α.Ε.

3.6.2.10 Συμπεράσματα

Με γνώμονα τα αποτελέσματα που αντλήθηκαν μέσω των δοκιμών, οι οποίες πραγματοποιήθηκαν, το συμπέρασμα το οποίο εξάγεται είναι, πως ο βέλτιστος αριθμός επαναλήψεων (epochs) για τα νευρωνικά δίκτυα που επιλέχθηκαν είναι 400, βάσει των πληροφοριακών κριτηρίων MSE, RMSE. Παράλληλα παρατηρήθηκε η πολύ καλή προβλεπτική ικανότητα του μοντέλου, λόγω των μικρών διαφορών που είχαν τα σφάλματα μεταξύ τους.

3.6.3 Μελέτη και παρουσίαση αποτελεσμάτων για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

Η δεύτερη εταιρεία που θα μελετηθεί είναι η μεγαλύτερη αεροπορική εταιρεία στην Ελλάδα, η οποία είναι η Αεροπορία Αιγαίου Α.Ε.. Ως οδηγό τα αποτελέσματα, καθώς και τα συμπεράσματα τα οποία εξήχθησαν από την μελέτη της ΓΕΚ ΤΕΡΝΑ Α.Ε., οι συναρτήσεις βελτιστοποίησης που θα χρησιμοποιηθούν για τα δίκτυα LSTM, GRU και RNN θα είναι ο Adam και ο Nadam, καθώς και για το δίκτυο MLP θα είναι ο Adam και ο RMSprop.

3.6.3.1 LSTM για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

Ο πρώτος αλγόριθμος που θα υλοποιηθεί και σε αυτή την περίπτωση είναι ο LSTM. Θα χρησιμοποιηθούν 2 συναρτήσεις βελτιστοποίησης, ο Adam και ο Nadam, καθώς και 3 διαφορετικά epochs, τα οποία είναι 100, 200 και 400.

Πίνακας 3.11: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο LSTM δίκτυο για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

Test Number	Neural Network	Optimizer	Epochs	Metrics
1	LSTM	Adam	100	MSE: 0.000221550 RMSE: 0.014884557
2	LSTM	Adam	200	MSE: 0.000284917 RMSE: 0.016879504
3	LSTM	Adam	400	MSE: 0.000220532 RMSE: 0.014850326
4	LSTM	Nadam	100	MSE: 0.000225017 RMSE: 0.015000576
5	LSTM	Nadam	200	MSE: 0.000222682 RMSE: 0.014225407
6	LSTM	Nadam	400	MSE: 0.000221305 RMSE: 0.014876352

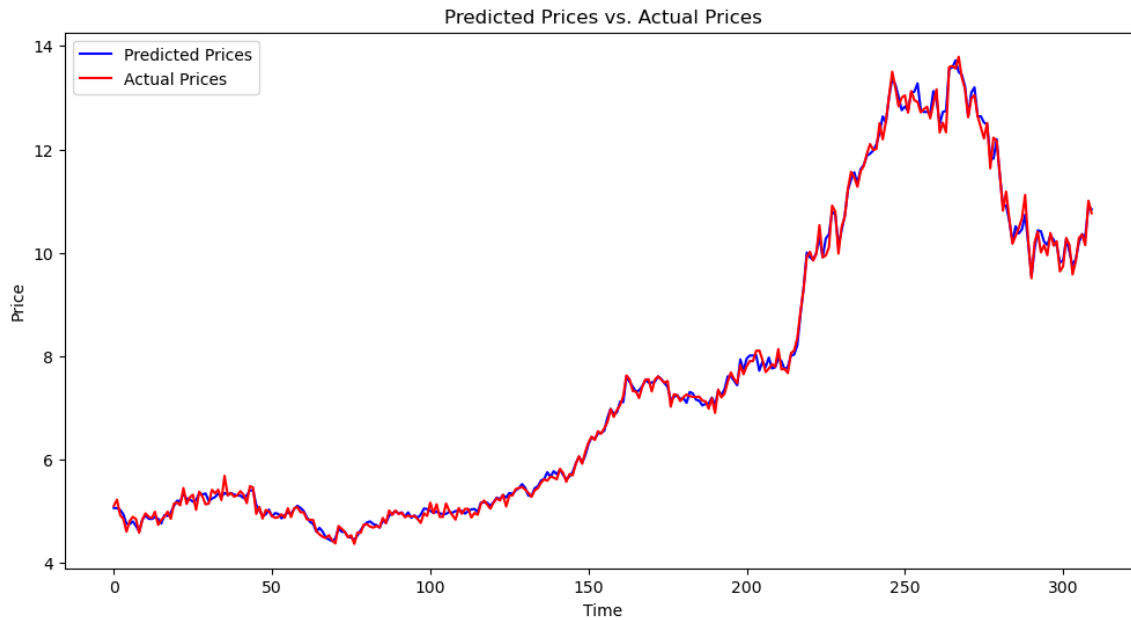
Από τον Πίνακα 3.11 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου LSTM είναι το μοντέλο 3. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000220532, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Adam, καθώς και τα epochs του 400.

Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.12 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 3, το οποίο σημείωσε το μικρότερο σφάλμα. Με

μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Πιο συγκεκριμένα, παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.12: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το LSTM δίκτυο της ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

3.6.3.2 GRU για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

Ο δεύτερος αλγόριθμος που θα παρουσιαστεί, όπως και στην προηγούμενη περίπτωση είναι το φραγμένο ανατροφοδοτούμενο νευρωνικό δίκτυο GRU, το οποίο θα εξάγει αποτελέσματα κάνοντας χρήση δύο αλγορίθμων βελτιστοποίησης, τον Adam και τον Nadam, καθώς και διαφορετικών epochs, τα οποία θα είναι 100,200,400. Τα αποτελέσματα των δοκιμών παρουσιάζονται στον Πίνακα 3.12.

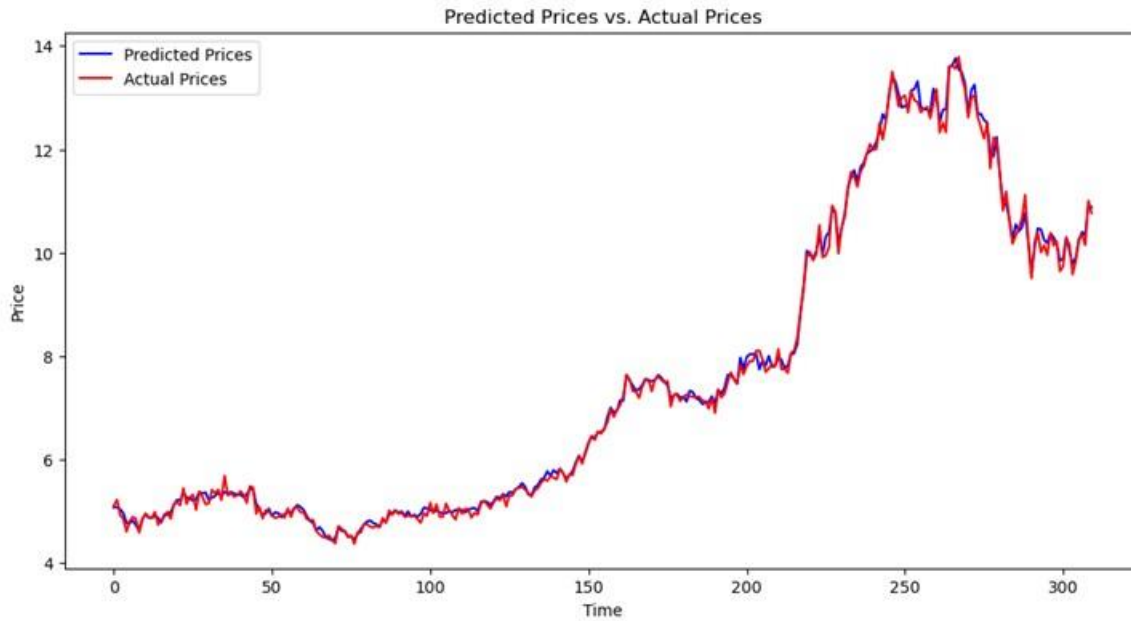
Πίνακας 3.12: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο GRU δίκτυο για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

Test Number	Neural Network	Optimizer	Epochs	Metrics
1	GRU	Adam	100	MSE: 0.000226945 RMSE: 0.015064703
2	GRU	Adam	200	MSE: 0.000220201 RMSE: 0.014839186
3	GRU	Adam	400	MSE: 0.000247189 RMSE: 0.015722254
4	GRU	Nadam	100	MSE: 0.000246363 RMSE: 0.015695979
5	GRU	Nadam	200	MSE: 0.000221983 RMSE: 0.014899113
6	GRU	Nadam	400	MSE: 0.000222864 RMSE: 0.014928660

Από τον Πίνακα 3.12 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου GRU είναι το μοντέλο 2. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000220201, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Adam, καθώς και τα epochs του 400. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.13 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 2, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.13: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το GRU δίκτυο της ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

3.6.3.3 RNN για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

Ο τρίτος αλγόριθμος που θα παρουσιαστεί είναι το παλινδρομικό νευρωνικό RNN, το οποίο θα εξάγει αποτελέσματα κάνοντας χρήση δύο αλγορίθμων βελτιστοποίησης, τον Adam και τον Nadam, καθώς και διαφορετικών epochs, τα οποία θα είναι 100,200,400. Τα αποτελέσματα των δοκιμών παρουσιάζονται στον Πίνακα 3.13.

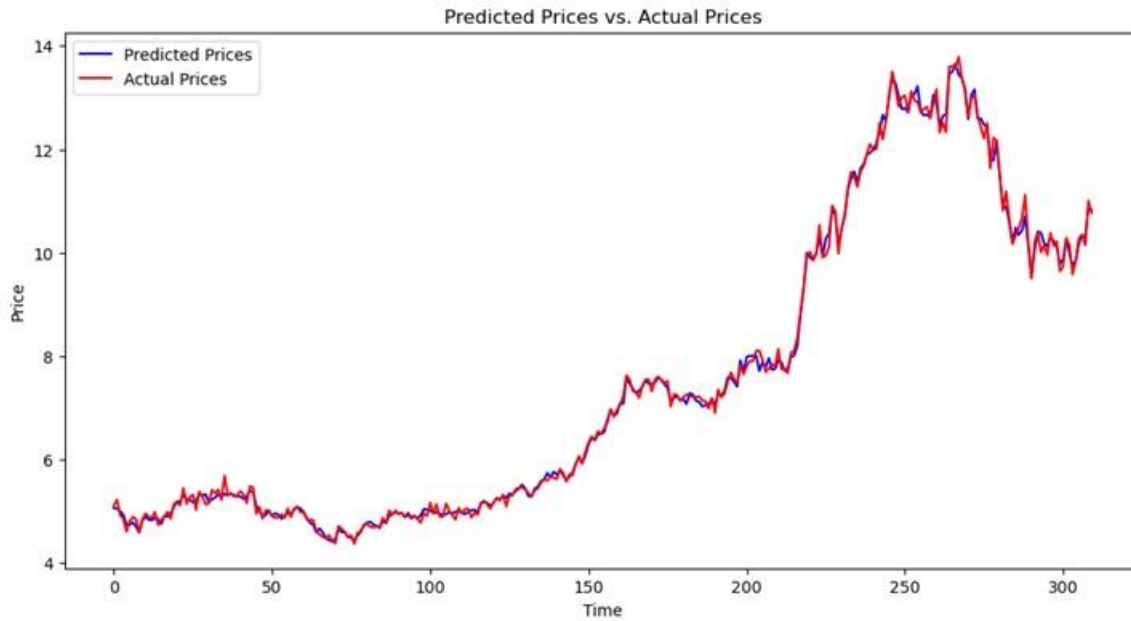
Πίνακας 3.13: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο RNN δίκτυο για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

Test Number	Neural Network	Optimizer	Epochs	Metrics
1	RNN	Adam	100	MSE: 0.000319421 RMSE: 0.017878249
2	RNN	Adam	200	MSE: 0.000237895 RMSE: 0.015423876
3	RNN	Adam	400	MSE: 0.000223015 RMSE: 0.014933716
4	RNN	Nadam	100	MSE: 0.000250457 RMSE: 0.015825762
5	RNN	Nadam	200	MSE: 0.000229707 RMSE: 0.015156102
6	RNN	Nadam	400	MSE: 0.000226035 RMSE: 0.015034477

Από τον Πίνακα 3.13 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου RNN είναι το μοντέλο 3. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000223015, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Adam, καθώς και τα epochs του 400. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.14 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 3, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.14: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το MLP δίκτυο της ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

3.6.3.4 MLP για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

Ο τέταρτος αλγόριθμος που θα παρουσιαστεί είναι το Πολυεπίπεδο Perceptron (MLP), το οποίο θα εξάγει αποτελέσματα κάνοντας χρήση δύο αλγορίθμων βελτιστοποίησης, τον Adam και τον RMSprop, καθώς και διαφορετικών epochs, τα οποία θα είναι 100,200,400. Τα αποτελέσματα των δοκιμών παρουσιάζονται στον Πίνακα 3.14.

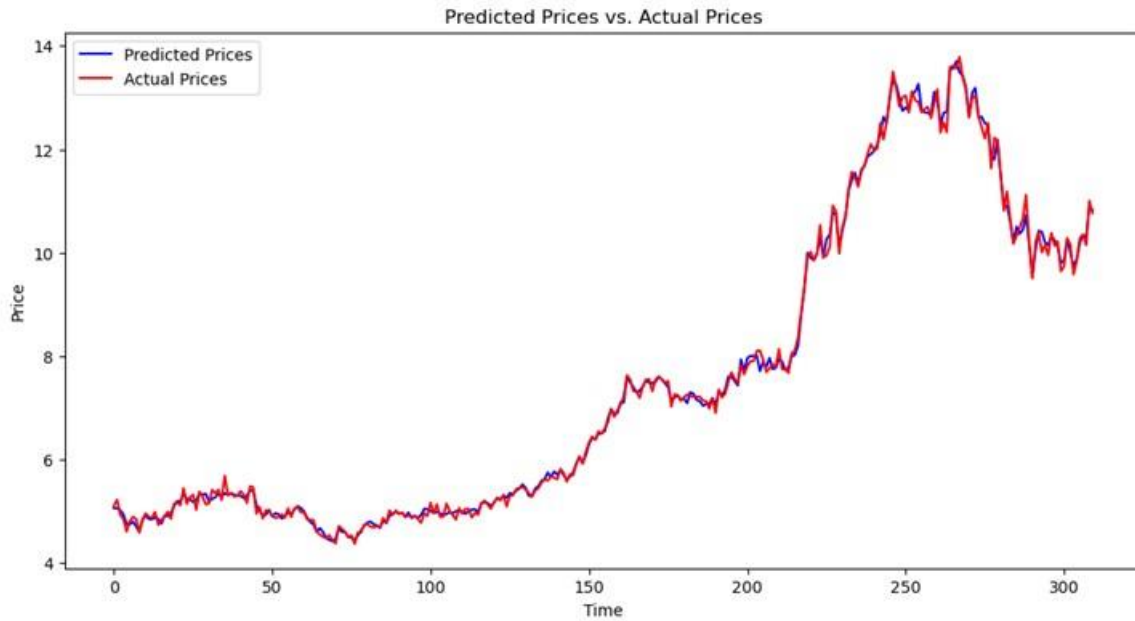
Πίνακας 3.14: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο MLP δίκτυο για την ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

Test Number	Neural Network	Optimizer	Epochs	Metrics
1	MLP	Adam	100	MSE: 0.000301923 RMSE: 0.017375931
2	MLP	Adam	200	MSE: 0.000227936 RMSE: 0.015097556
3	MLP	Adam	400	MSE: 0.000255269 RMSE: 0.015008976
4	MLP	RMSprop	100	MSE: 0.000223481 RMSE: 0.014949300
5	MLP	RMSprop	200	MSE: 0.000245171 RMSE: 0.015657941
6	MLP	RMSprop	400	MSE: 0.000220635 RMSE: 0.014853812

Από τον Πίνακα 3.14 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου MLP είναι το μοντέλο 6. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000220635, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο RMSprop, καθώς και τα epochs του 400. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.15 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 6, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.15: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το MLP δίκτυο της ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.

3.6.4 Μελέτη και παρουσίαση αποτελεσμάτων για τον ΟΠΑΠ Α.Ε.

Η τρίτη και τελευταία εταιρεία, η οποία θα μελετηθεί είναι η μεγαλύτερη εταιρεία τυχερών παιχνιδιών στην Ελλάδα, ο ΟΠΑΠ Α.Ε.

3.6.4.1 LSTM για τον ΟΠΑΠ Α.Ε.

Ο πρώτος αλγόριθμος που θα υλοποιηθεί και σε αυτή την περίπτωση είναι ο LSTM. Θα χρησιμοποιηθούν 2 συναρτήσεις βελτιστοποίησης, ο Adam και ο Nadam, καθώς και 3 διαφορετικά epochs, τα οποία είναι 100, 200 και 400.

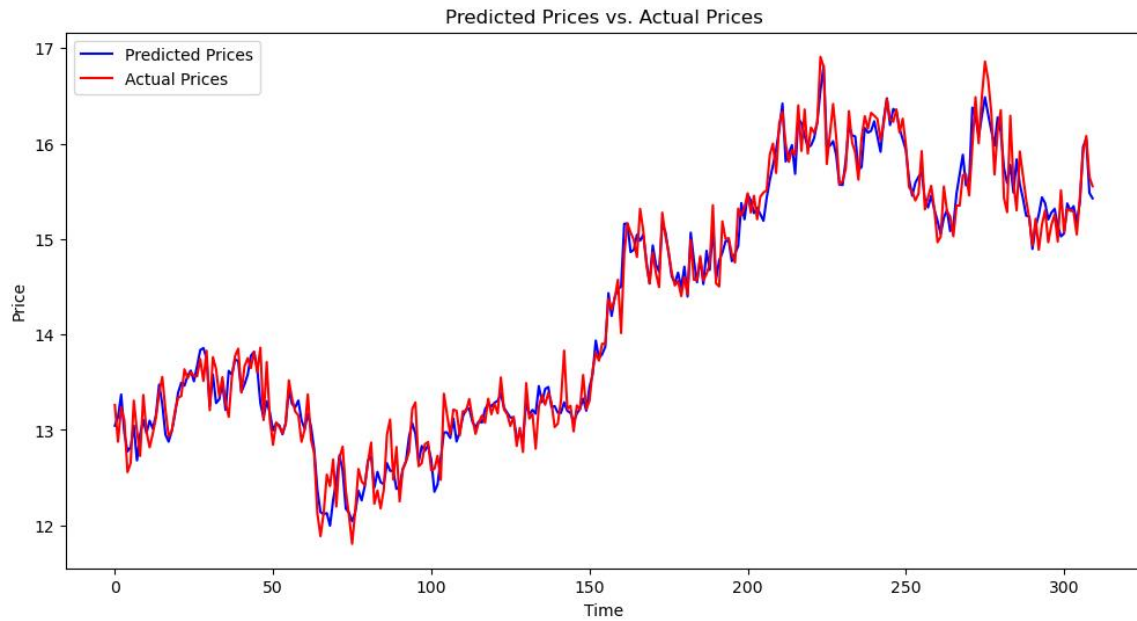
Πίνακας 3.15: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο LSTM δίκτυο για τον ΟΠΑΠ Α.Ε.

Test Number	Neural Network	Optimizer	Epochs	Metrics
1	LSTM	Adam	100	MSE: 0.000177254 RMSE: 0.013313692
2	LSTM	Adam	200	MSE: 0.000178675 RMSE: 0.013366945
3	LSTM	Adam	400	MSE: 0.000178469 RMSE: 0.013359231
4	LSTM	RMSprop	100	MSE: 0.000182216 RMSE: 0.013498751
5	LSTM	RMSprop	200	MSE: 0.000177697 RMSE: 0.013306635
6	LSTM	RMSprop	400	MSE: 0.000178257 RMSE: 0.013351292

Από τον Πίνακα 3.15 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου LSTM είναι το μοντέλο 1. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000177254, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Adam, καθώς και τα epochs του 100. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Στη συνέχεια ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.16 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 1, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.16: Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το LSTM δίκτυο του ΟΠΑΠ Α.Ε.

3.6.4.2 GRU για τον ΟΠΑΠ Α.Ε.

Ο δεύτερος αλγόριθμος που θα παρουσιαστεί, όπως και στην προηγούμενη περίπτωση είναι το φραγμένο ανατροφοδοτούμενο νευρωνικό δίκτυο GRU, το οποίο θα εξάγει αποτελέσματα κάνοντας χρήση δύο αλγορίθμων βελτιστοποίησης, τον Adam και τον Nadam, καθώς και διαφορετικών epochs, τα οποία θα είναι 100,200,400. Τα αποτελέσματα των δοκιμών παρουσιάζονται στον Πίνακα 3.16.

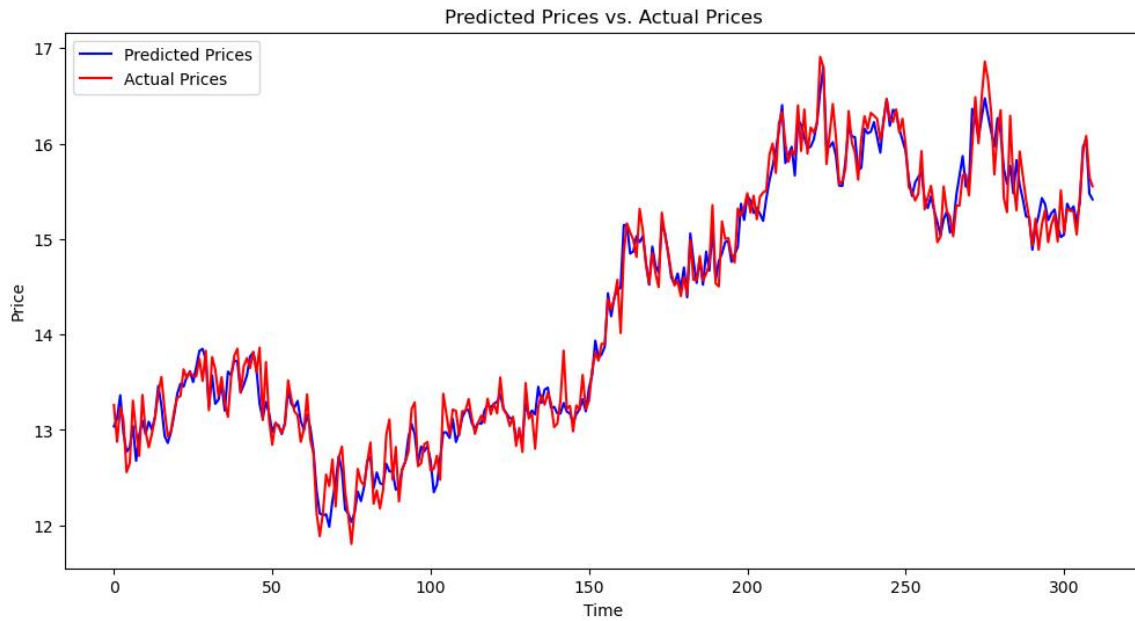
Πίνακας 3.16: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο GRU δίκτυο για τον ΟΠΑΠ Α.Ε.

Test Number	Neural Network	Optimizer	Epochs	Metrics
1	GRU	Adam	100	MSE: 0.000177554 RMSE: 0.013324952
2	GRU	Adam	200	MSE: 0.000177008 RMSE: 0.013304472
3	GRU	Adam	400	MSE: 0.000177673 RMSE: 0.013329424
4	GRU	Nadam	100	MSE: 0.000181793 RMSE: 0.013483064
5	GRU	Nadam	200	MSE: 0.000180196 RMSE: 0.013423740
6	GRU	Nadam	400	MSE: 0.000176944 RMSE: 0.013302052

Από τον Πίνακα 3.16 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου GRU είναι το μοντέλο 6. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000176944, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Nadam, καθώς και τα epochs του 400. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.17 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 6, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.17 Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το GRU δίκτυο του ΟΠΑΠ Α.Ε.

3.6.4.3 RNN για τον ΟΠΑΠ Α.Ε.

Ο τρίτος αλγόριθμος που θα παρουσιαστεί είναι το παλινδρομικό νευρωνικό RNN, το οποίο θα εξάγει αποτελέσματα κάνοντας χρήση δύο αλγορίθμων βελτιστοποίησης, τον Adam και τον Nadam, καθώς και διαφορετικών epochs, τα οποία θα είναι 100,200,400. Τα αποτελέσματα των δοκιμών παρουσιάζονται στον Πίνακα 3.17.

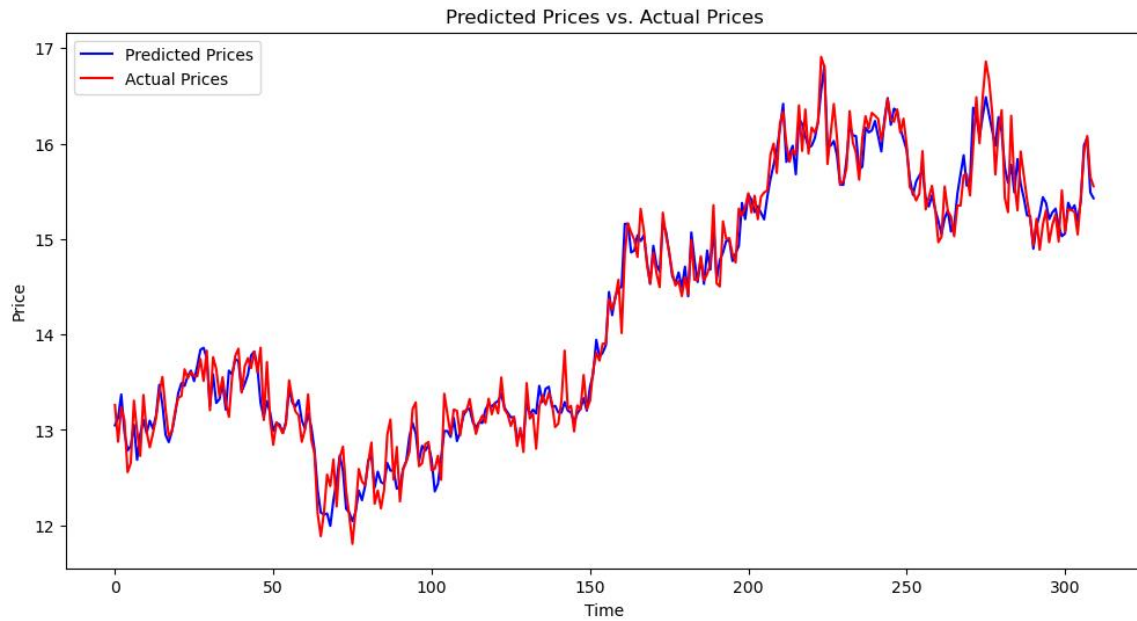
Πίνακας 3.17: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο RNN δίκτυο για τον ΟΠΑΠ Α.Ε.

Test Number	Neural Network	Optimizer	Epochs	Metrics
1	RNN	Adam	100	MSE: 0.000181026 RMSE: 0.013454603
2	RNN	Adam	200	MSE: 0.000177539 RMSE: 0.013332438
3	RNN	Adam	400	MSE: 0.000180464 RMSE: 0.013433709
4	RNN	Nadam	100	MSE: 0.000188620 RMSE: 0.013733916
5	RNN	Nadam	200	MSE: 0.000178075 RMSE: 0.013344478
6	RNN	Nadam	400	MSE: 0.000176890 RMSE: 0.013300001

Από τον Πίνακα 3.17 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου RNN είναι το μοντέλο 2. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000177539, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Adam, καθώς και τα epochs του 200. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.18 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 2, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.18 Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το RNN δίκτυο του ΟΠΑΠ Α.Ε.

3.6.4.4 MLP για τον ΟΠΑΠ Α.Ε.

Ο τέταρτος αλγόριθμος που θα παρουσιαστεί είναι το Πολυεπίπεδο Perceptron (MLP), το οποίο θα εξάγει αποτελέσματα κάνοντας χρήση δύο αλγορίθμων βελτιστοποίησης, τον Adam και τον RMSprop, καθώς και διαφορετικών epochs, τα οποία θα είναι 100,200,400. Τα αποτελέσματα των δοκιμών παρουσιάζονται στον Πίνακα 3.16.

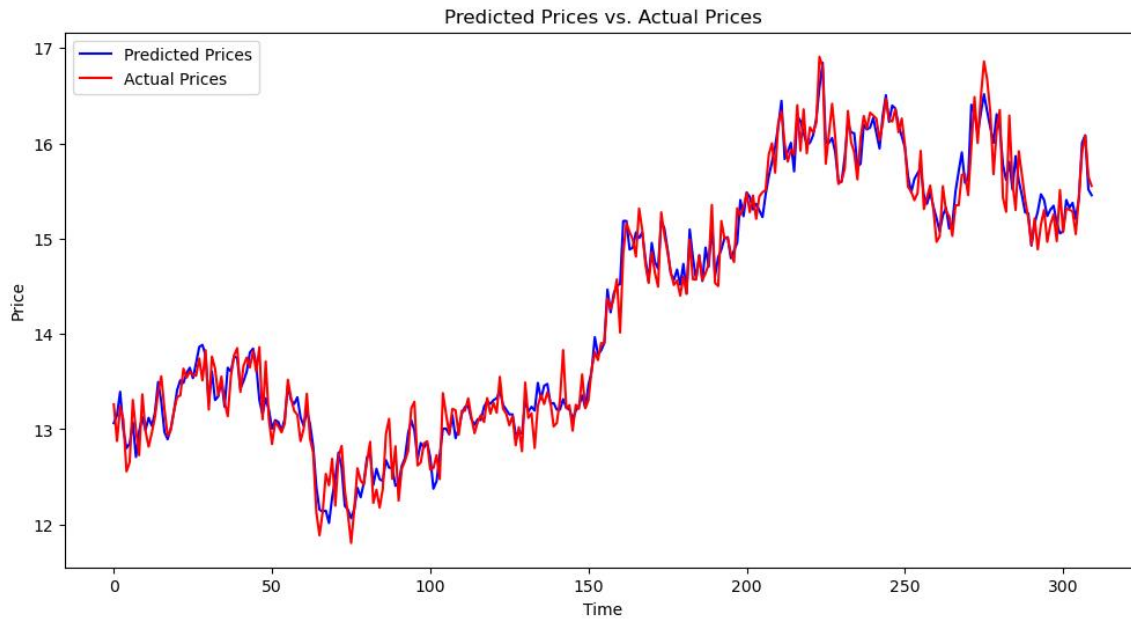
Πίνακας 3.18: Αποτελέσματα των κριτηρίων σφάλματος του βέλτιστου μοντέλου στο MLP δίκτυο για τον ΟΠΑΠ Α.Ε.

Test Number	Neural Network	Optimizer	Epochs	Metrics
1	MLP	Adam	100	MSE: 0.000177288 RMSE: 0.013315026
2	MLP	Adam	200	MSE: 0.000185706 RMSE: 0.013629036
3	MLP	Adam	400	MSE: 0.000183973 RMSE: 0.013563692
4	MLP	RMSprop	100	MSE: 0.000178684 RMSE: 0.013367280
5	MLP	RMSprop	200	MSE: 0.000180831 RMSE: 0.013447367
6	MLP	RMSprop	400	MSE: 0.000185336 RMSE: 0.013613835

Από τον Πίνακα 3.18 εξάγεται το συμπέρασμα, πως το αποδοτικότερο μοντέλο του δικτύου MLP είναι το μοντέλο 1. Το μέσο τετραγωνικό σφάλμα (MSE) του μοντέλου ισούται με 0.000177288, αλγόριθμος βελτιστοποίησης που περιέχει είναι ο Adam, καθώς και τα epochs του 100. Σημαντικό στοιχείο αποτελεί, πως οι διαφορές είναι πολύ μικρές, το οποίο δηλώνει πως το μοντέλο περιέχει πολύ καλή προβλεπτική ικανότητα για το συγκεκριμένο σύνολο δεδομένων.

Ακολουθεί η μετατροπή των προβλεπόμενων λογαριθμικών αποδόσεων σε προβλεπόμενες τιμές κλεισίματος, με σκοπό να συγκριθούν με τις πραγματικές τιμές.

Στο Διάγραμμα 3.19 παρουσιάζεται η γραφική απεικόνιση της πρόβλεψης των τιμών κλεισίματος των μετοχών του μοντέλου 1, το οποίο σημείωσε το μικρότερο σφάλμα. Με μπλε χρώμα απεικονίζονται οι προβλεπόμενες παρατηρήσεις, ενώ με κόκκινο χρώμα οι πραγματικές παρατηρήσεις. Παρατηρείται πως το μοντέλο ακολουθεί με πολύ μεγάλη ακρίβεια τις πραγματικές τιμές.



Διάγραμμα 3.19 Γραφική απεικόνιση του αποδοτικότερου μοντέλου για το MLP δίκτυο του ΟΠΑΠ Α.Ε.

3.7 Συμπεράσματα της μελέτης και μελλοντικοί στόχοι

- Συμπεράσματα

Στον Πίνακα 3.19 παρουσιάζεται μια σύνοψη των αποδοτικότερων μοντέλων βάσει του πληροφοριακού κριτηρίου του μέσου τετραγωνικού σφάλματος. Πιο αποδοτικό μοντέλο αναδεικνύεται αυτό που πετυχαίνει την μικρότερη τιμή του.

Πίνακας 3.19: Σύνοψη των καλύτερων αποτελεσμάτων ανά δίκτυο για κάθε εταιρεία

Neural Network	ΓΕΚ ΤΕΡΝΑ Α.Ε.	ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	ΟΠΑΠ Α.Ε.
LSTM	MSE: 0.000284262	MSE: 0.000220532	MSE: 0.000177254
GRU	MSE: 0.000284109	MSE: 0.000220201	MSE: 0.000176944
RNN	MSE: 0.000284669	MSE: 0.000223015	MSE: 0.000177539
MLP	MSE: 0.000284303	MSE: 0.000225269	MSE: 0.000177288

Σύμφωνα με τον Πίνακα 3.19 εξάγεται το συμπέρασμα πως και στις τρεις εταιρείες αποδοτικότερο νευρωνικό δίκτυο αναδεικνύεται το GRU, αφού σταθερά παράγγαγε

καλύτερα αποτελέσματα σε σχέση με τα υπόλοιπα, αν και οι διαφορές τους είναι πολύ μικρές.

Καλύτερος αλγόριθμος βελτιστοποίησης αναδεικνύεται ο Adam για όλα τα νευρωνικά δίκτυα, αφού παρήγαγε τα καλύτερα αποτελέσματα στις περισσότερες περιπτώσεις.

Στον Πίνακα 3.20 παρουσιάζεται μια σύνοψη των μοντέλων που παρουσίασαν τα χειρότερα αποτελέσματα, βάσει του πληροφοριακού κριτηρίου του μέσου τετραγωνικού σφάλματος.

Πίνακας 3.20: Σύνοψη των χειρότερων αποτελεσμάτων ανά δίκτυο για κάθε εταιρεία

Neural Network	ΓΕΚ ΤΕΡΝΑ Α.Ε.	ΑΕΡΟΠΟΡΙΑ ΑΙΓΑΙΟΥ Α.Ε.	ΟΠΑΠ Α.Ε.
LSTM	MSE: 0.000293370	MSE: 0.000284917	MSE: 0.000178674
GRU	MSE: 0.000286944	MSE: 0.000247189	MSE: 0.000181793
RNN	MSE: 0.000306865	MSE: 0.000319421	MSE: 0.000188620
MLP	MSE: 0.000291453	MSE: 0.000301923	MSE: 0.000185704

Από τον Πίνακα 3.20 γίνεται κατανοητό, πως το νευρωνικό δίκτυο που παρουσίασε σταθερά τα χειρότερα αποτελέσματα και στις τρεις εταιρείες είναι το RNN, αν και οι διαφορές τους είναι πολύ μικρές.

Παρατηρώντας τα αποτελέσματα των πινάκων, καθώς και τα αντίστοιχα διαγράμματα από την μετατροπή των λογαριθμικών αποδόσεων σε τιμές κλεισίματος εξάγεται το συμπέρασμα, πως οι προβλεπόμενες τιμές έχουν πολύ μικρή απόκλιση από τις πραγματικές, γεγονός που δείχνει την πολύ καλή προβλεπτική ικανότητα των μοντέλων. Παράλληλα παρατηρείται, πως, όσο μεγάλωνε ο αριθμός των epochs, τόσο αυξανόταν οι προβλεπτική ικανότητα των μοντέλων. Στις περισσότερες περιπτώσεις ο αριθμός των epochs που το μέσο τετραγωνικό σφάλμα λάμβανε την μικρότερη τιμή του ήταν 400.

- Μελλοντικοί στόχοι:

Πλησιάζοντας στην ολοκλήρωση της διπλωματικής εργασίας γίνεται αντιληπτό, πως η πρόβλεψη των αποδόσεων των μετοχών αποτελεί ένα πολυσύνθετο αντικείμενο μελέτης. Η αλματώδης αύξηση της τεχνολογίας, η δημιουργία νέας τεχνογνωσίας και κατασκευής

αλγορίθμων, οδηγεί μέρα με την μέρα στη δημιουργία πιο εύστοχων προβλέψεων. Τα τεχνητά νευρωνικά δίκτυα συνεχώς εξελίσσονται, γεγονός το οποίο θα δώσει πολλές δυνατότητες στο μέλλον. Ο κυριότερος σκοπός της διπλωματικής εργασίας αποτελεί η διερεύνηση της αξιοπιστίας των τεχνητών νευρωνικών δικτύων σε μεθόδους πρόβλεψης οικονομικών δεδομένων, που απορρέουν από την Ελληνική χρηματιστηριακή αγορά.

Σε ένα πιο μακροπρόθεσμο πλαίσιο, θα ήταν χρήσιμες κάποιες ενέργειες με σκοπό την εξαγωγή πιο ασφαλών συμπερασμάτων για τις τρεις εταιρείες. Αρχικά θα μπορούσε να πραγματοποιηθεί μελέτη όλων των μεταβλητών του αρχικού συνόλου δεδομένων, ώστε να πραγματοποιηθεί η κατασκευή αλγορίθμων τεχνητών νευρωνικών δικτύων για ανάλυση πολύ μεταβλητών χρονοσειρών. Παράλληλα θα μπορούσαν να δημιουργηθούν ολοκληρωμένα υποδείγματα ARIMA, των οποίων τα αποτελέσματα θα συγκριθούν με τα αποτελέσματα των παραπάνω αλγορίθμων. Κλείνοντας, ιδιαίτερη αξία θα είχε η εφαρμογή των παραπάνω μοντέλων σε ρεαλιστικές συνθήκες, δηλαδή την εφαρμογή τους, για την εύρεση της τιμής κλεισίματος την επόμενη ημέρα συνδυαστικά με τις συνθήκες τις οποίες επικρατούν (οικονομικές, πολιτικές, κοινωνικές κ.λπ.), αλλά και την αβεβαιότητα που επικρατεί.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική

1. Snell, S. R., Παπαδόπουλος, Γ., (μετ.). (2008). Κλινική Νευροανατομική, Ιατρικές Εκδόσεις Λίτσας.
2. Αγγελάκη, Κ. (2020). «Διερεύνηση σχέσης αιτιότητας μεταξύ του χρέους μιας χώρας και του γενικού χρηματιστηριακού δείκτη σε Ευρωπαϊκές χώρες», Διπλωματική εργασία ΠΜΣ «Εφαρμοσμένη Στατιστική», Πανεπιστήμιο Πειραιώς.
3. Αγιακλόγλου, Χ. Ν. (2022). Σημειώσεις μαθήματος «Πρόβλεψη-Χρονοσειρές» του ΠΜΣ «Εφαρμοσμένη Στατιστική», Πανεπιστήμιο Πειραιώς.
4. Αγιακλόγλου, Χ. Ν., Γκάρος, Π. Α. (2022). Σημειώσεις εργαστηρίου «Πρόβλεψη-Χρονοσειρές» του ΠΜΣ «Εφαρμοσμένη Στατιστική», Πανεπιστήμιο Πειραιώς.
5. Αγιακλόγλου, Χ. Ν., Οικονόμου, Γ. Σ. (Αθήνα 2004). Μέθοδοι προβλέψεων και ανάλυσης αποφάσεων, Β΄ έκδοση, εκδόσεις Γ. Μπένου.
6. Βλαχάβας, Ι., Κεφαλάς, Π., Βασιλειάδης, Ν., Κόκκορας, Φ., Σακελλαρίου, Η., (2002). Τεχνητή Νοημοσύνη, Δ΄ Έκδοση, Εκδόσεις Πανεπιστημίου Μακεδονίας.
7. Γαρδίκης, Π. (2022). «Αλγόριθμοι και τεχνικές βελτίωσης του Federated Learning», Διπλωματική Εργασία, Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής, Πανεπιστήμιο Πατρών.
8. Γερολυμάτος, Γ. (2022). Ανάπτυξη τεχνικών εκπαίδευσης νευρωνικών δικτύων ακτινικής συνάρτησης βάσης με τον αλγόριθμο Levenberg-Marquardt, Τμήμα Ηλεκτρολόγων & Ηλεκτρονικών Μηχανικών, Πανεπιστήμιο Δυτικής Αττικής.
9. Γεωργούλη, Κ. (2015). «Τεχνητή Νοημοσύνη – Μια εισαγωγική προσέγγιση», Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, Εθνικό Μετσόβιο Πολυτεχνείο.
10. Γιαννακόπουλος, Β. (2021). «Πυροδοτούμενα νευρωνικά δίκτυα και εφαρμογές τους», Διπλωματική Εργασία, ΠΜΣ Επιστήμη και Τεχνολογία της Πληροφορικής και των Υπολογιστών, Πανεπιστήμιο Δυτικής Αττικής.
11. Κεράστας, Ι. (2021). «Αναγνώριση φωνητικών εντολών με νευρωνικά δίκτυα» Διπλωματική Εργασία, τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Πανεπιστήμιο Θεσσαλίας.

12. Κοντόπουλος, Χ. (2015). «Τεχνικές βαθιάς μηχανικής μάθησης και Συνελικτικά Νευρωνικά Δίκτυα για την ταξινόμηση υπερφασματικών δεδομένων», Διπλωματική Εργασία, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, εργαστήριο Τηλεπισκόπησης, Εθνικό Μετσόβιο Πολυτεχνείο.
13. Κουγιουμτζής, Δ. (2019-2020). Κεφάλαιο 6ο «Χρονοσειρές», Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης.
14. Κούτρας, Μ. (2021-2022). Σημειώσεις μαθήματος «Ανάλυση παλινδρόμησης και ανάλυση Διακύμανσης», ΠΜΣ «Εφαρμοσμένη Στατιστική», Πανεπιστήμιο Πειραιώς.
15. Λιαγκόρβας, Π., Γουλάς, Χ., Αποστολόπουλος, Ν. (2022). Οι κοινωνικοοικονομικές επιπτώσεις της πανδημίας του Covid-19 στους εργαζόμενους του ιδιωτικού τομέα.
16. Λύκας, Α. (2014) Διαφάνειες μαθήματος «Νευρωνικά Δίκτυα». Πανεπιστήμιο Ιωαννίνων.
17. Μελάς, Χ. (2022). «Μελέτη νευρωνικών δικτύων και εφαρμογές στην πρόβλεψη χρονοσειρών στο χρηματιστήριο», Διπλωματική εργασία ΠΜΣ «Ψηφιακά Συστήματα και Υπηρεσίες», Πανεπιστήμιο Πειραιώς.
18. Μπαρδελής, Κ. (2019). «Εφαρμογή τεχνικών μηχανικής μάθησης για την πρόβλεψη της κατανομής των επισκεπτών σε μεγάλες εκδηλώσεις, κάνοντας χρήση δεδομένων τοποθεσίας», Διπλωματική Εργασία, τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Εθνικό Μετσόβιο Πολυτεχνείο.
19. ΟΠΑΠ, Α.Ε.. (2020). Ετήσιες Οικονομικές Καταστάσεις 2020.
20. Ορφανού, Φ. (Δεκ. 2015). «Μέθοδοι πρόβλεψης της ζήτησης - Εφαρμογή σε δεδομένα ζήτησης καταναλωτικών προϊόντων», Διπλωματική εργασία ΠΜΣ «Εφαρμοσμένη Στατιστική», Πανεπιστήμιο Πειραιώς.
21. Παγουρτζή, Ο. (2021), «Τεχνικές εξόρυξης δεδομένων και νευρωνικών δικτύων σε οικονομικά δεδομένα», Διπλωματική εργασία ΠΜΣ «Ψηφιακά Συστήματα και Υπηρεσίες», Πανεπιστήμιο Πειραιώς.
22. Παναγοπούλου, Ε. (2020). «Πρόβλεψη Χρονοσειρών με Έντονη Εποχικότητα», Διπλωματική Εργασία, τμήμα Μηχανολόγων και Αεροναυπηγών, Πανεπιστημίου Πατρών.
23. Πανάνα, Α. (2018-2019). «Χρονικές Σειρές», τμήμα Μαθηματικών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης.
24. Πελέκης, Ν. (2022). Σημειώσεις μαθήματος «Εξόρυξη Δεδομένων», ΠΜΣ «Εφαρμοσμένη Στατιστική», Πανεπιστήμιο Πειραιώς.

25. Πελέκης, Ν., Χονδροδήμα Ε. (2022). Σημειώσεις εργαστηρίου «Εξόρυξη Δεδομένων», ΠΜΣ «Εφαρμοσμένη Στατιστική», Πανεπιστήμιο Πειραιώς
26. Σημειώσεις μαθήματος «Τεχνικές Προβλέψεων», Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Εθνικό Μετσόβιο Πολυτεχνείο.
27. Τσιμπάνος, Α. Π. (2005). «Υποφαινόμενες Σχέσεις Μεταξύ Χρονοσειρών», Διπλωματική εργασία ΠΜΣ «Εφαρμοσμένη Στατιστική», Πανεπιστήμιο Πειραιώς.
28. Τσουχνικά, Μ. (2007). Νευρωνικά δίκτυα και εφαρμογές (Doctoral dissertation, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης).
29. Χασάπη, Λ. (2022). Παρουσίαση με θέμα «Βαθιά Μάθηση – Μηχανική Μάθηση», Πανεπιστήμιο Πατρών.
30. Χονδροδήμα, Ε. (2017). «Ανάπτυξη αλγορίθμων υπολογιστικής νοημοσύνης και εφαρμογές στην εξέλιξη του παγκόσμιου ιστού», Διδακτορική διατριβή, Εθνικό Μετσόβιο Πολυτεχνείο.
31. Χούτας, Β. (2017). «Ανάπτυξη Αυτόνομων Πρακτόρων με χρήση Προσαρμοστικών Μηχανισμών Εστίασης και Βαθέων Νευρωνικών Δικτύων», Διπλωματική Εργασία, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Εργαστήριο Επεξεργασίας Πληροφορίας και Υπολογισμών.

Ξένη

1. Agiakloglou, C., & Newbold, P. (1992). Empirical evidence on Dickey-Fuller-type tests. *Journal of Time Series Analysis*, 13(6), 471-483.
2. Bebis, G., & Georgiopoulos, M. (1994). Feed-forward neural networks. *Ieee Potentials*, 13(4), 27-31.
3. Bergman, K. H. (1979). Multivariate analysis of temperatures and winds using optimum interpolation. *Monthly Weather Review*, 107(11), 1423-1444.
4. de Camargo, A. A. R., & de Oliveira, M. A. (2023). Analysis of the Application of Different Forecasting Methods for Time Series in the Context of the Aeronautical Industry. *Engineering Proceedings*, 39(1), 74.
5. Dshalalow, J. H., & White, R. T. (2021). Random walk analysis in a reliability system under constant degradation and random shocks. *Axioms*, 10(3), 199.

6. Efron, B. (2004). The estimation of prediction error: covariance penalties and cross-validation. *Journal of the American Statistical Association*, 99(467), 619-632.
7. El Hajji, S., Moukafih, N., & Orhanou, G. (2019). Analysis of neural network training and cost functions impact on the accuracy of IDS and SIEM systems. In *Codes, Cryptology and Information Security: Third International Conference, C2SI 2019, Rabat, Morocco, April 22–24, 2019, Proceedings-In Honor of Said El Hajji 3* (pp. 433-451). Springer International Publishing.
8. Guo, Y., Wu, W., Zhang, B., & Sun, H. (2014). A fast solution for the lagrange multiplier-based electric power network parameter error identification model. *Energies*, 7(3), 1288-1299.
9. Jeong, J. J. (2019). A Minimum Mean-Square-Error (MMSE) Decoder for Quasi-Orthogonal Space–Time Block Code. *Electronics*, 8(7), 732.
10. Kazda, A., Badanik, B., & Serrano, F. (2022). Pandemic vs. Post-pandemic airport operations: Hard impact, slow recovery. *Aerospace*, 9(12), 810.
11. Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., ... & Gruszczyńska, K. (2023). What is machine learning, artificial neural networks and deep learning? —Examples of practical applications in medicine. *Diagnostics*, 13(15), 2582.
12. Kühnert, C., Gonuguntla, N. M., Krieg, H., Nowak, D., & Thomas, J. A. (2021). Application of LSTM networks for water demand prediction in optimal pump control. *Water*, 13(5), 644.
13. Makridakis, S. (1976). A survey of time series. *International Statistical Review/Revue Internationale de Statistique*, 29-70.
14. Maniatopoulos, A., & Mitianoudis, N. (2021). Learnable leaky relu (LeLeLU): An alternative accuracy-optimized activation function. *Information*, 12(12), 513.
15. Marhon, S. A., Cameron, C. J. F., & Kremer, S. C. (2013). Handbook on neural information processing. *Intelligent Systems Reference Library*, 49, 10-1007.
16. McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5, 115-133.
17. Okuyama, S., Yoshimura, M., Amakura, Y., Nakajima, M., & Furukawa, Y. (2022). Activation of Extracellular Signal-Regulated Kinase 2 and cAMP Response Element-Binding Protein in Cultured Neurons by the Macrocyclic Ellagitannin Oenothien B. *NeuroSci*, 3(3), 387-394.
18. Pham, H. (2019). A new criterion for model selection. *Mathematics*, 7(12), 1215.

19. Rhif, M., Ben Abbes, A., Farah, I. R., Martínez, B., & Sang, Y. (2019). Wavelet transform application for/in non-stationary time-series analysis: A review. *Applied Sciences*, 9(7), 1345.
20. Sopelsa Neto, N. F., Stefenon, S. F., Meyer, L. H., Bruns, R., Nied, A., Seman, L. O., ... & Yow, K. C. (2021). A study of multilayer perceptron networks applied to classification of ceramic insulators using ultrasound. *Applied Sciences*, 11(4), 1592.
21. Su, Y., Cui, C., & Qu, H. (2022). Self-attentive moving average for time series prediction. *Applied Sciences*, 12(7), 3602.
22. Sun, J., Liu, J., Hu, L., Wei, Z., & Xiao, L. (2022). A mutual teaching framework with momentum correction for unsupervised hyperspectral image change detection. *Remote Sensing*, 14(4), 1000.
23. Tieleman, T., & Hinton, G. (2012). Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning*, 4(2), 26-31.
24. Tomovski, Ž., Gerhold, S., Bansal, D., & Soni, A. (2022). Geometric Properties of Some Generalized Mathieu Power Series inside the Unit Disk. *Axioms*, 11(10), 568.
25. Zhou, X., Zhang, H., Zheng, S., Xing, W., Zhao, P., & Li, H. (2022). The crude oil international trade competition networks: Evolution trends and estimating potential competition links. *Energies*, 15(7), 2395.

Ηλεκτρονική

1. Brownlee J. (2020). White Noise Time Series with Python. Available at: <https://machinelearningmastery.com/white-noise-time-series-python/>
2. Census Bureau (2024). Retail Sales: *Used Car Dealers*. *Federal Reserve Bank of St. Louis*. Available at: <https://fred.stlouisfed.org/series/MRTSSM44112USN>
3. Josh (2023). Time Series Anomaly Detection. Available at: <https://www.kaggle.com/code/joshuaswords/time-series-anomaly-detection>
4. Manish N. (2019). Introduction to the Architecture of Recurrent Neural Networks (RNNs). Published in *Towards AI*. Available at: <https://pub.towardsai.net/introduction-to-the-architecture-of-recurrent-neural-networks-rnns-a277007984b7>

5. Mominul, Md., I., (2020). Introduction of Time Series. Available at: <https://rpubs.com/mominulislam2329/BASICACF>
6. Nisha A. A. (2022). Deep Learning Techniques you Should Know in 2022. Published in Heartbeat. Available at: <https://heartbeat.comet.ml/deep-learning-techniques-you-should-know-in-2022-94f33e62d922>
7. PennState: Statistics Online Courses (n.d.): Stat 510. Lesson 1.1.2 sample ACF and properties of AR(1) model. Available at: <https://online.stat.psu.edu/stat510/lesson/1/1.2>
8. Primo.ai. - Web (2023). Available at: [https://primo.ai/index.php?title=Gated Recurrent Unit %28GRU%29](https://primo.ai/index.php?title=Gated_Recurrent_Unit_%28GRU%29)
9. Quantsart – web. White Noise and Random Walks in Time Series Analysis. Available at: <https://www.quantstart.com/articles/White-Noise-and-Random-Walks-in-Time-Series-Analysis/>
10. Researchgate (2012): The air passengers time series data. Available at: https://www.researchgate.net/figure/The-air-passengers-time-series-data_fig4_261464969
11. Wikidocs (n.d.). Part C. LSTM, GRU, RNN, CNN – EN. Deep Learning Bible - 3. Natural Language Processing – Eng. Available at: <https://wikidocs.net/166316>
12. Wikipedia.org – web (n.d.). Available at: https://el.wikipedia.org/wiki/Η_Ναυτεμπορική
13. Η μοναχική πτωτική πορεία του ΧΑ, η Εθνική ξεχώρισε και η διόρθωση της Aegean | Business Daily
14. Καρακασίδης Θ. και Φράγκου Α. (n.d). Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών «Βιώσιμη Διαχείριση Περιβαλλοντικών Αλλαγών και Κυκλική Οικονομία». Ανάλυση χρονοσειρών και δεδομένων περιβαλλοντικών κινδύνων. Available at: https://eclass.uth.gr/modules/document/file.php/PRD_P_199/19_ΠΡΟΒΛΕΨΗ_ARIMA.pdf
15. Παρασκευάς Μ. (n.d.). Σήματα και Συστήματα. Διάλεξη 2: Στοιχειώδη Σήματα Συνεχούς Χρόνου. Τεχνολογικό Εκπαιδευτικό Ίδρυμα Δυτικής Ελλάδας. Available at: [http://opencourses.teiwest.gr/modules/document/file.php/CIED107/1.%20Διαλέξεις/Signals and Systems Lectures 2.pdf](http://opencourses.teiwest.gr/modules/document/file.php/CIED107/1.%20Διαλέξεις/Signals_and_Systems_Lectures_2.pdf)
16. Σπάει όλα τα ρεκόρ ο τζόγος στην Ελλάδα, πονταρίσματα 32 δισ. στο 11μηνο | Business Daily

17. Χασάπη Λ. (n.d.). Deep Learning. Πανεπιστήμιο Πατρών. Available at:
<https://eclass.upatras.gr/modules/document/file.php/PT187/11.DEEP%20LEARNING%2C%20MACHINE%20LEARNING.pdf>

