



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**  
**ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ**  
**Π.Μ.Σ "ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ & ΥΠΗΡΕΣΙΕΣ"**  
**ΚΑΤΕΥΘΥΝΣΗ "ΜΕΓΑΛΑ ΔΕΔΟΜΕΝΑ ΚΑΙ ΑΝΑΛΥΤΙΚΗ"**

---

**«ΤΕΧΝΙΚΕΣ ΕΞΟΡΥΞΗΣ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΣΕ ΟΙΚΟΝΟΜΙΚΑ ΔΕΔΟΜΕΝΑ»**

---

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**ΠΑΓΟΥΡΤΖΗ ΟΥΡΑΝΙΑ**

**ΠΕΙΡΑΙΑΣ, ΦΛΕΒΑΡΗΣ 2021**

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή του τμήματος Ψηφιακά Συστήματα του Πανεπιστημίου Πειραιώς στην υπ' αριθμό ..... συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών.

Τα μέλη της Επιτροπής ήταν:

**Μιχαήλ Φιλιππάκης**  
Αναπληρωτής Καθηγητής  
(Επιβλέπων)

**Δημοσθένης Κυριαζής**  
Αναπληρωτής Καθηγητής

**Χαλκίδη Μαρία**  
Αναπληρώτρια Καθηγήτρια

Η έγκριση της Διπλωματικής Εργασίας από το τμήμα Ψηφιακών Συστημάτων του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνωμών του συγγραφέα.



**UNIVERSITY OF PIRAEUS**  
**DEPARTMENT OF DIGITAL SYSTEMS**  
**POST GRADUATE PROGRAM "INFORMATION SYSTEMS AND**  
**SERVICES"**  
**AREA "BIG DATA & ANALYTICS"**

---

**«DATA MINING AND NEURAL NETWORK TECHNIQUES IN FINANCIAL  
DATA»**

---

**MSc Dissertation**

**PAGOURTZI OURANIA**

**PIRAEUS, FEBRUARY 2021**

This thesis was approved unanimously by the three – member committee appointed by the Department of Digital Systems, University of Piraeus, in accordance with the rules of the MSc program.

Committee members were:

**Michael Filippakis**  
Associate Professor  
(Supervisor)

**Dimosthenis Kyriazis**  
Associate Professor

**Halkidi Maria**  
Associate Professor

Approval of this thesis from the Department of Digital Systems, University of Piraeus, does not imply any endorsement of the opinions of the author.

Στην οικογένεια μου,



## Ευχαριστίες

*Θα ξεκινήσω ευχαριστώντας την οικογένεια μου και τους φίλους μου, που με στήριξαν καθ' όλη την διάρκεια της προσπάθειας μου.*

*Φυσικά τον κ.Μιχάλη Φιλιππάκη, για την καθοδήγηση και τις συμβουλές του προκειμένου να διεκπεραιώσω ορθά την παρούσα εργασία, όπως και για την κατανόηση που έδειξε στην πορεία της προσπάθειας μου.*

*Τέλος, θέλω να ευχαριστήσω όλους τους καθηγητές του Μεταπτυχιακού Προγράμματος για την διάθεση τους να μεταδώσουν τις γνώσεις τους και τις εμπειρίες τους προς εμάς.*

*Ράνια Παγουρτζή*





## Περίληψη

Σκοπός της παρούσας εργασίας αποτελεί η μελέτη και κατανόηση των νευρωνικών δικτύων και της μηχανικής μάθησης γενικότερα, έπειτα από την εφαρμογή τους σε ένα σετ χρονοσειρών χρηματιστηρίου, το οποίο αποτελείται από γνωστές εταιρίες πληροφορικής.

Το πρόβλημα της δημιουργίας προβλέψεων για τιμές χρηματιστηρίου, γίνεται ολοένα και πιο περιζήτητο, αφού αυξάνεται συνεχώς το μέγεθος των δεδομένων που μπορούν να χρησιμοποιηθούν για την εκπαίδευση και μελέτη αλγορίθμων. Επιπρόσθετα, αποτελεί πρόκληση η εύστοχη πρόβλεψη χρηματιστηριακών τιμών τόσο για τις ίδιες τις εταιρείες όσο και για ιδιώτες που ασχολούνται ενεργά με αυτό τον κλάδο.

Τα νευρωνικά δίκτυα με τα οποία θα ασχοληθούμε, αποτελούν μέθοδο Βαθιάς Μάθησης και είναι μέρος της Τεχνητής Νοημοσύνης. Είναι υπολογιστικά συστήματα εμπνευσμένα από τα βιολογικά νευρωνικά δίκτυα του ανθρώπινου εγκεφάλου. Έτσι ο αρχικός σκοπός τους ήταν να δώσουν λύσεις σε προβλήματα, με τον ίδιο τρόπο που δίνει λύσεις ο ανθρώπινος εγκέφαλος. Στη συγκεκριμένη εργασία, θα ασχοληθούμε εκτενέστερα με δύο είδη επαναλαμβανόμενων νευρωνικών δικτύων και ένα είδος Perceptron.

Ξεκινάμε την εργασία, παρουσιάζοντας κάποια εισαγωγικά στοιχεία, όσον αφορά την μηχανική μάθηση και τα είδη της. Στο Κεφάλαιο 2, αναφερόμαστε στα χαρακτηριστικά των χρονοσειρών. Στη συνέχεια στο Κεφάλαιο 3 αναλύονται οι διάφοροι αλγόριθμοι πρόβλεψης, αναλόγως το διαθέσιμο σύνολο δεδομένων κάθε φοράς και η εργασία προχωρά στο Κεφάλαιο 4 όπου αναλύονται βασικά Νευρωνικά Δίκτυα που θα χρησιμοποιηθούν στην πορεία για το πειραματικό μέρος. Τέλος, στο Κεφαλαίο 5 με την βοήθεια των χρονοσειρών Amazon, Google, Apple μοντελοποιούνται τα νευρωνικά δίκτυα του κεφαλαίου 4 και αξιολογούνται με κατάλληλα μέτρα αξιολόγησης ως προς την ικανότητα πρόβλεψης μελλοντικών τιμών.



# Περιεχόμενα

Ευχαριστίες .....	vii
Περίληψη .....	ix
<b>ΚΕΦΑΛΑΙΟ 1</b> .....	1
Μηχανική Μάθηση.....	1
1.1 Εισαγωγή στη Μηχανική Μάθηση.....	1
1.2 Είδη Μηχανικής Μάθησης .....	2
<b>ΚΕΦΑΛΑΙΟ 2</b> .....	3
Χρονοσειρές .....	3
2.1 Ορισμός .....	3
2.2 Κατηγορίες Χρονοσειρών .....	3
2.3 Ποιοτικά Χαρακτηριστικά.....	4
2.4 Στασιμότητα – Μη Στασιμότητα .....	7
2.5 Ανάλυση Χρονοσειράς.....	8
2.6 Στατιστικοί Δείκτες.....	9
<b>ΚΕΦΑΛΑΙΟ 3</b> .....	11
Μέθοδοι Πρόβλεψης .....	11
3.1 Λευκός Θόρυβος (White Noise) .....	11
3.2 Τυχαίος Περίπατος (Random Walk) .....	13
3.3 Μέθοδοι Πρόβλεψης με Κινητούς Μέσους ( Moving Average) .....	14
3.4 Μέθοδοι εξομάλυνσης (Smoothing Methods).....	15
3.5 Μοντέλα για Στάσιμες Χρονοσειρές.....	17
3.6 Μοντέλα για μη στάσιμες χρονοσειρές.....	21
3.7 Επιλογή Μεθόδου Πρόβλεψης.....	25
3.8 Μέτρα Αξιολόγησης Προβλέψεων .....	26
<b>ΚΕΦΑΛΑΙΟ 4</b> .....	28
Τεχνητά Νευρωνικά Δίκτυα.....	28
4.1 Ορισμός .....	28
4.2 Αρχιτεκτονική Νευρωνικών Δικτύων .....	28
4.3 Βασικά Νευρωνικά Δίκτυα .....	31
4.3.1 Πολυεπίπεδος Perceptron (Multilayer Perceptron - MLP).....	31
4.3.2 Επαναλαμβανόμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks-RNN) ....	32
4.3.3 Μακράς Βραχυχρόνιας Μνήμης Νευρωνικά Δίκτυα (Long Short-Term Memory -LSMT).....	34
4.4.4 Φραγμένα Επαναλαμβανόμενα Νευρωνικά Δίκτυα (Gated Recurrent Units - GRUs).....	35

<b>ΚΕΦΑΛΑΙΟ 5</b> .....	36
<b>Πειραματικός Έλεγχος</b> .....	36
<b>5. 1 Πρόβλημα – Σκοπός</b> .....	36
<b>5.2 Εργαλεία</b> .....	37
<b>5.3 Διερευνητική Ανάλυση (Exploratory Analysis)</b> .....	37
<b>5.3.1 Σύνολο Δεδομένων (dataset)</b> .....	37
<b>5.3.2 Επεξεργασία Δεδομένων</b> .....	38
<b>5.3.3 Οπτικοποίηση Δεδομένων (Data Visualization)</b> .....	38
<b>5.3.4 Ποιοτικά Χαρακτηριστικά</b> .....	40
<b>5.3.5 Augmented Dickey – Fuller Test</b> .....	40
<b>5.3.6 Συσχετίσεις μεταξύ των μεταβλητών</b> .....	41
<b>5.4 Μοντελοποίηση Νευρωνικών Δικτύων</b> .....	44
<b>5.4.1 Keras Configuration</b> .....	44
<b>5.5 Μετατροπή δεδομένων για εισαγωγή στα Νευρωνικά Δίκτυα</b> .....	45
<b>5.6 Αρχιτεκτονική Long-Short Term Memory Network (LSTMN)</b> .....	46
<b>5.6.1 LSTMN για την Amazon</b> .....	47
<b>5.6.2 LSTMN για την Google</b> .....	48
<b>5.6.3 LSTMN για την Apple</b> .....	49
<b>5.7 Αρχιτεκτονικά Gated Recurrent Neural Networks (GRUs)</b> .....	51
<b>5.7.1 GRU για την Amazon</b> .....	51
<b>5.7.2 GRU για την Google</b> .....	52
<b>5.7.3 GRU για την Apple</b> .....	53
<b>5.8 Αρχιτεκτονική Multilayer Perceptron (MLP)</b> .....	54
<b>5.8.1 MLP για την Amazon</b> .....	55
<b>5.8.2 MLP για Google</b> .....	56
<b>5.8.3 MLP Apple</b> .....	57
<b>5.9 Σύνοψη Αποτελεσμάτων</b> .....	58
<b>Βιβλιογραφία</b> .....	lx

# ΚΕΦΑΛΑΙΟ 1

## Μηχανική Μάθηση

### 1.1 Εισαγωγή στη Μηχανική Μάθηση

Η μάθηση σε ένα γνωστικό σύστημα, μπορεί να συνδεθεί με δύο βασικές ιδιότητες, την ικανότητά στην πρόσκτηση γνώσης κατά την αλληλεπίδρασή του με το περιβάλλον και την ικανότητά να βελτιώνει με την επανάληψη τον τρόπο εκτέλεσης μία ενέργειας.

Για τον όρο Μηχανική Μάθηση έχουν διατυπωθεί διάφοροι ορισμοί ανά τα χρόνια, όπως:

«Η μελέτη υπολογιστικών μεθόδων για την απόκτηση νέας γνώσης, νέων δεξιοτήτων και νέων τρόπων οργάνωσης της υπάρχουσας γνώσης.» - Carbonell (1987)

«Ένα πρόγραμμα υπολογιστή θεωρείται ότι μαθαίνει από την εμπειρία E σε σχέση με μια κατηγορία εργασιών T και μια μετρική απόδοσης P, αν η απόδοση του σε εργασίες της T, όπως μετριοούνται από την P, βελτιώνονται με την εμπειρία E.» - Mitchell (1997)

«Κάτι μαθαίνει όταν αλλάζει τη συμπεριφορά του κατά τέτοιο τρόπο ώστε να αποδίδει καλύτερα στο μέλλον.» - Witten & Frank (2000)

Η Μηχανική Μάθηση ανήκει στον κλάδο της Τεχνητής Νοημοσύνης και δίνει την δυνατότητα στον άνθρωπο να επιλύσει πολύπλοκα προβλήματα με την βοήθεια αλγορίθμων και υπολογιστών, που σε άλλη περίπτωση θα ήταν δύσκολο. Ουσιαστικά η Μηχανική Μάθηση ασχολείται με τη μελέτη αλγορίθμων που βελτιώνουν τη συμπεριφορά τους σε κάποια εργασία που τους έχει ανατεθεί χρησιμοποιώντας την εμπειρία τους, σ' ένα υπολογιστικό σύστημα. Τέλος, η μηχανική μάθηση μπορεί να θεωρηθεί ως «η αναζήτηση σε ένα χώρο πιθανών υποθέσεων, εκείνης της υπόθεσης που ταιριάζει κατά τον καλύτερο τρόπο στα υπό εξέταση δεδομένα και στην πιθανώς προϋπάρχουσα γνώση» . [1]

## 1.2 Είδη Μηχανικής Μάθησης

Η Μηχανική Μάθηση διακρίνεται σε τρεις κατηγορίες ως προς την εκπαίδευση των αλγορίθμων, που μοιάζουν με τον τρόπο που μαθαίνει ο άνθρωπος. Οι κατηγορίες είναι η επιβλεπόμενη μάθηση, η μη επιβλεπόμενη και η ενισχυτική μάθηση.

### ❖ **Επιβλεπόμενη Μάθηση (Supervised Learning)**

Είναι η διαδικασία όπου ο αλγόριθμος Μηχανικής Μάθησης κατασκευάζει μια συνάρτηση που αναπαριστά συγκεκριμένες εισόδους (σύνολο εκπαίδευσης) σε γνωστές επιθυμητές εξόδους, με στόχο τη γενίκευση της συνάρτησης αυτής και για εισόδους με άγνωστη έξοδο. Χρησιμοποιείται σε προβλήματα Ταξινόμησης (Classification), Πρόγνωσης (Prediction), Διερμηνείας (Interpretation). Ένα παράδειγμα αποτελεί η κατηγοριοποίηση των email σε ανεπιθύμητη ή μη αλληλογραφία.

### ❖ **Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)**

Είναι η διαδικασία όπου ο αλγόριθμος Μηχανικής Μάθησης κατασκευάζει ένα μοντέλο για κάποιες παρατηρήσεις, ως δεδομένα εισόδου, χωρίς να γνωρίζει τα επιθυμητά αποτελέσματα εξόδου. Χρησιμοποιείται σε προβλήματα Ανάλυσης Συσχετισμών (Association Analysis) και Ομαδοποίησης (Clustering).

### ❖ **Ενισχυτική Μάθηση (Reinforcement Learning)**

Είναι η διαδικασία όπου ο αλγόριθμος Μηχανικής Μάθησης μέσα από την άμεση αλληλεπίδραση με το περιβάλλον, εκπαιδεύεται ως προς μια στρατηγική ενεργειών. Αποτελεί ίσως την πιο κοινή μορφή μάθησης στον πραγματικό κόσμο, αφού αφορά στη μάθηση μέσω της δοκιμής (trial) και της αποτυχίας(error). Χρησιμοποιείται κυρίως σε προβλήματα σχεδιασμού (Planning), όπως για παράδειγμα ο έλεγχος κίνησης ενός ρομπότ, η αυτόματη οδήγηση ενός οχήματος και η εκπαίδευση ενός ρομποτ να παίζει ένα παιχνίδι εναντίον ενός αντιπάλου. [2]

[3]

## ΚΕΦΑΛΑΙΟ 2

### Χρονοσειρές

#### 2.1 Ορισμός

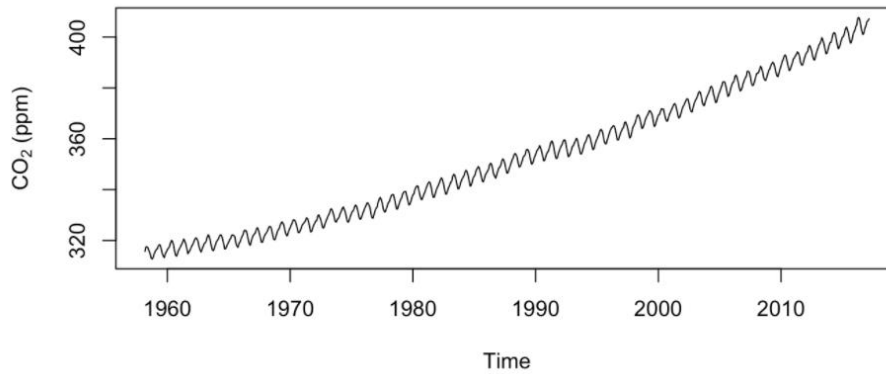
Ως χρονοσειρά ή χρονολογική σειρά ορίζεται μία ακολουθία από παρατηρήσεις που συλλέχθηκαν διαδοχικά σε ισαπέχουσες χρονικές στιγμές (ημέρες, μήνες, έτη κλπ) . Η χρονοσειρά εκφράζει την εξέλιξη μίας μεταβλητής  $X_T$  , όπου  $t = 1,2,3,\dots,T$  χρονικά διαστήματα, μέσα στο χρόνο.

#### 2.2 Κατηγορίες Χρονοσειρών

Οι χρονοσειρές διακρίνονται στις συνεχείς και στις διακριτές. Συνεχείς είναι οι χρονοσειρές που η τιμή της μεταβλητής παρατηρείται και καταγράφεται συνεχώς ακόμα και αν παίρνει διακριτές τιμές, ενώ ως διακριτές ορίζονται οι χρονοσειρές που η τιμή της μεταβλητής καταγράφεται σε ορισμένα χρονικά διαστήματα ακόμα και αν η τιμή της μεταβλητής είναι συνεχής. Παράδειγμα συνεχούς χρονοσειράς είναι η συνεχής παρακολούθηση των σεισμών, ενώ παράδειγμα διακριτής χρονοσειράς είναι η καταγραφή των τιμών των μετοχών του χρηματιστηρίου, όπου οι διαδοχικοί όροι ισαπέχουν.

Επιπρόσθετα, οι χρονοσειρές χαρακτηρίζονται βάσει των μεταβλητών που καταγράφονται. Όταν μία χρονοσειρά αποτελείται από καταγραφές μόνο μίας μεταβλητής ονομάζεται univariate, ενώ όταν αποτελείται από καταγραφές περισσότερων ονομάζεται multivariate.

Για την καλύτερη κατανόηση μίας χρονοσειράς συνηθίζεται η γραφική απεικόνιση όπως φαίνεται στο παρακάτω παράδειγμα, για την μηνιαία καταγραφή του CO<sub>2</sub> στην πόλη Manua Loa της Χαβάης.



Μηνιαία απεικόνιση από το 1958 έως το 2010 του CO<sub>2</sub> στην πόλη Mauna Loa της Χαβάη [4]

## 2.3 Ποιοτικά Χαρακτηριστικά

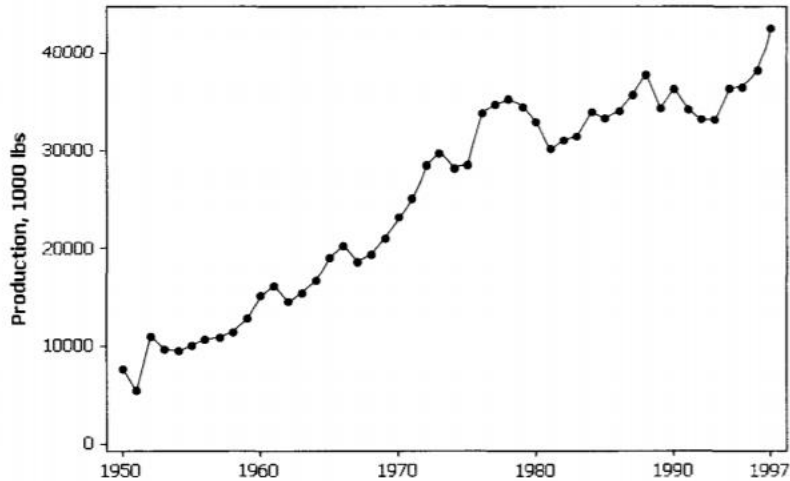
Ως ποιοτικά χαρακτηριστικά μίας χρονοσειράς ορίζονται οι διαφορετικοί τύποι των διακυμάνσεων που τη συνθέτουν. Τα βασικά ποιοτικά χαρακτηριστικά είναι η συχνότητα, η τάση, η εποχικότητα, η κυκλικότητα και η τυχαιότητα. Αυτά τα χαρακτηριστικά αποτελούν σημαντικές συνιστώσες για τις προβλέψεις που θα ακολουθήσουν. [5] [6] [7]

Η **συχνότητα** (frequency) δείχνει κάθε πότε καταγράφονται οι τιμές που σχηματίζουν την χρονοσειρά.

### Τάση

Ως τάση (trend) ορίζεται η μακροπρόθεσμη μεταβολή της μέσης τιμής της χρονοσειράς. Το χαρακτηριστικό αυτό σχετίζεται με τα δομικά στοιχεία της χρονοσειράς και μπορεί να είναι ανοδική, καθοδική ή μηδενική. Προκειμένου να αξιολογηθεί η τάση με ασφαλή συμπεράσματα, πρέπει να υπάρχει επαρκής αριθμός παρατηρήσεων. Ακολουθεί παράδειγμα χρονοσειράς με ανοδική τάση.

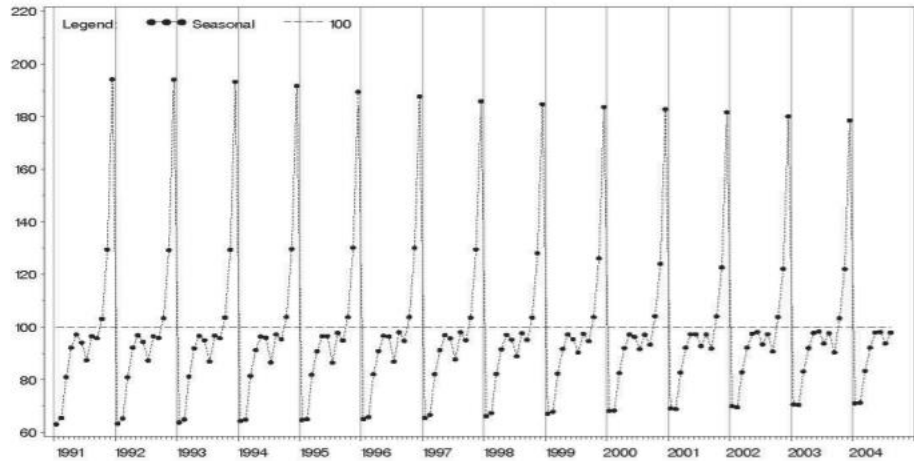




Ετήσια παραγωγή των blue cheese & gorgonzola cheese στις Ηνωμένες Πολιτείες από το 1950 – 1997 [8]

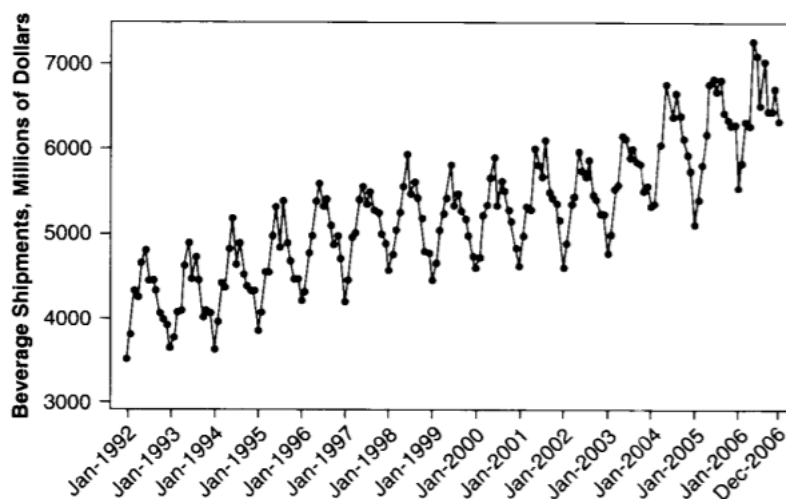
### Εποχικότητα

Η εποχικότητα (seasonality) εμφανίζεται όταν στην χρονοσειρά παρατηρείται μία επαναληψιμότητα, σε σταθερά διαστήματα μικρότερα του ενός έτους (πχ. εβδομάδα, μήνες, τετράμηνα). Η περιοδική διακύμανση είναι προβλέψιμη καθώς οι τιμές επαναλαμβάνονται με τον ίδιο τρόπο στην πάροδο του χρόνου. Η εποχικότητα επηρεάζεται συνήθως από το κλίμα και τις συμβατικές εποχές, όπως θρησκευτικές, κοινωνικές και πολιτικές εκδηλώσεις, οι οποίες επαναλαμβάνονται από χρόνο σε χρόνο. Για παράδειγμα, στη Βόρεια Αμερική η αγορά και κατανάλωση πετρελαίου αυξάνεται το χειμώνα και η κατανάλωση ηλεκτρικής ενέργειας αυξάνεται τους καλοκαιρινούς μήνες λόγω του κλιματισμού. Ακολουθεί διάγραμμα απεικόνισης των πωλήσεων Καναδέζικου Καταστήματος, όπου παρατηρείται κάθε χρόνο ο ίδιο ρυθμός πωλήσεων, με ιδιαίτερη αύξηση τους τελευταίους μήνες κάθε έτους. [6] [7]



Πωλήσεις καταστήματος στο Καναδά από το 1991 ως το 2004

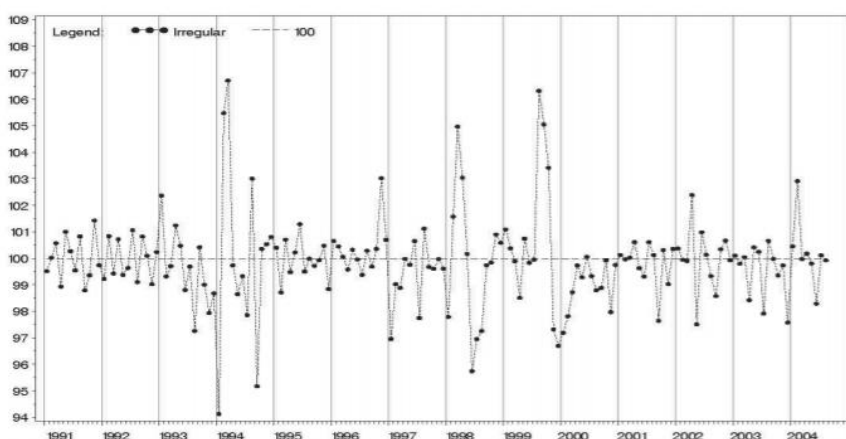
Η **κυκλικότητα** (cyclic ) παρατηρείται σε μία χρονοσειρά όταν οι τιμές έχουν διακυμάνσεις αλλά όχι σε σταθερά διαστήματα. Συνήθως οι περίοδοι επανάληψης ξεπερνούν το ένα έτος (πχ. διετία, πενταετία). Αυτό μπορεί να συμβαίνει σε χρονοσειρές όπως είναι οι τιμές των μετοχών του χρηματιστηρίου, εξαιτίας της μεταβλητότητας της οικονομίας κάθε κράτους. Στις γραφικές παραστάσεις των χρονοσειρών αυτών παρατηρείται μία κυματοειδής γραμμή που κινείται ανάμεσα στην υψηλότερη και χαμηλότερη τιμή. Ακολουθεί παράδειγμα χρονοσειράς με τις μηνιαίες αποστολές ποτών στις Ηνωμένες Πολιτείες με κυκλικό μοτίβο που επαναλαμβάνεται κάθε τρία έτη.



Απεικόνιση της αποστολής ποτών από παραγωγή στις Ηνωμένες Πολιτείες από τον Ιανουάριο του 1992 έως το Ιανουάριο του 2007

## Τυχειότητα

Η τυχειότητα (irregular – random) εμφανίζεται στις χρονοσειρές όταν υπάρχουν μη κανονικές διακυμάνσεις, επηρεασμένες από τυχαίους εξωτερικούς παράγοντες. Διακρίνονται ως απότομες αλλαγές στην συμπεριφορά της χρονοσειράς. Τέτοιες τιμές ονομάζονται **ακραίες τιμές** (outliers), για παράδειγμα μία ακραία τιμή θα μπορούσε να είναι μία ξαφνική καταιγίδα κατά την διάρκεια του καλοκαιριού όπου παρατηρούνται πολλές ημέρες ηλιοφάνειας. Οι ακραίες τιμές δεν μπορούν να προβλεφθούν και έχουν μικρή χρονική διάρκεια, ωστόσο μπορούν να επηρεάσουν σημαντικά τα αποτελέσματα των μοντέλων πρόβλεψης. Για το λόγο αυτό χρήζουν μελέτης και απομόνωσης. Στο διάγραμμα που ακολουθεί απεικονίζονται οι πωλήσεις από κατάσταση στον Καναδά. Παρατηρούνται ακραίες τιμές το 1994, 1998, 1999 και 2000, πιθανόν από το κλείσιμο και άνοιγμα άλλων παρόμοιων καταστημάτων στην ευρύτερη περιοχή, που επηρέασαν τις πωλήσεις του.



Πωλήσεις σε κατάσταση στον Καναδά από το 1991 έως το 2004

## 2.4 Στασιμότητα – Μη Στασιμότητα

Η **στασιμότητα** (stationary) δείχνει για μία χρονοσειρά αν αλλάζουν ή παραμένουν σταθεροί οι στατιστικοί δείκτες της στο χρόνο. Πιο συγκεκριμένα, αν αλλάζει συστηματικά η μέση τιμή και η διασπορά η χρονοσειρά είναι μη-στάσιμη, σε αντίθετη περίπτωση όταν παραμένουν σταθερά θεωρείται στάσιμη. Μία **μη-στάσιμη** χρονοσειρά μπορεί να εμφανίζει τάσεις και περιοδικότητα. Είναι πολύ δύσκολο να μελετηθούν και να εφαρμοστούν προβλέψεις σε μη στάσιμες χρονοσειρές. Για το λόγο αυτό υπάρχουν διάφορες τεχνικές, που θα αναλυθούν παρακάτω, προκειμένου να μετατραπεί μία χρονοσειρά από μη-στάσιμη σε στάσιμη.

## Augmented Dickey - Fuller Test

Το Augmented Dickey - Fuller Test ελέγχει αν μία χρονοσειρά είναι στάσιμη. Δηλαδή, αν οι τιμές κυμαίνονται γύρω από μία μέση τιμή. Η αρχική υπόθεση (null hypothesis,  $H_0$ ) του τεστ είναι ότι η χρονοσειρά είναι μη-στάσιμη και έχει κάποια εξάρτηση από τον χρόνο, ενώ η εναλλακτική υπόθεση (alternative hypothesis,  $H_1$ ) λέει ότι η χρονοσειρά είναι στάσιμη. Ο συντελεστής του τεστ παίρνει αρνητικές και θετικές τιμές. Όσο πιο αρνητική τιμή, τόσο πιο σίγουρα απορρίπτεται η αρχική υπόθεση. Επιπλέον, στα αποτελέσματα εμφανίζεται η τιμή p-value.

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

## 2.5 Ανάλυση Χρονοσειράς

Κατά την ανάλυση μιας χρονοσειράς διερευνάται η διαχρονική συμπεριφορά των τιμών της μεταβλητής της, με στόχο την πρόβλεψη, την μοντελοποίηση και τον χαρακτηρισμό. Κατά την πρόβλεψη εκτιμώνται οι μελλοντικές τιμές της μεταβλητής της χρονοσειράς για κάποιο σύντομο χρονικό διάστημα, θεωρώντας ότι οι παλαιότερες τιμές βοηθούν στην πρόβλεψη. Με την διαδικασία της μοντελοποίησης κατανοούνται η συμπεριφορά και οι ιδιότητες του συστήματος για μεγάλο χρονικό διάστημα, ενώ κατά τον χαρακτηρισμό στόχος είναι ο καθορισμός θεμελιωδών ιδιοτήτων του συστήματος.

Προκειμένου να γίνει η ανάλυση της χρονοσειράς, ακολουθεί διάσπαση (decomposition) της χρονοσειράς στις τέσσερις συνιστώσες (τάση, εποχικότητα, κυκλικότητα, τυχαιότητα), ώστε να εντοπιστούν ποιες επηρεάζουν περισσότερο την μεταβολή της τιμής της μεταβλητής και τελικά σχηματίζουν την πραγματική τιμή της χρονοσειράς. Δεν είναι απαραίτητο να συνυπάρχουν σε κάθε χρονοσειρά και οι τέσσερις συνιστώσες.

Έστω:

$Y_t$  η πραγματική τιμή της χρονοσειράς

$T_t$  η τάση της χρονοσειράς

$S_t$  η εποχικότητα της χρονοσειράς

$C_t$  η κυκλικότητα της χρονοσειράς

$E_t$  η τυχαιότητα της χρονοσειράς

όπου  $t = 1, 2, 3, \dots, n$  το χρονικό διάστημα.

Οι παραπάνω συνιστώσες συνδέονται με δύο μοντέλα μεταξύ τους.

### **Αθροιστικό μοντέλο (addictive model)**

Το πρώτο είναι το αθροιστικό μοντέλο, που συνεπάγεται ότι οι συνιστώσες είναι ανεξάρτητες μεταξύ τους και η τιμή της μίας συνιστώσας είναι ασυσχέτιστη με την τιμή της άλλης, καθώς η μονάδα μέτρησης των τεσσάρων συνιστωσών είναι ίδια με την μονάδα μέτρησης της πραγματικής τιμής ( $Y_t$ ). Αυτή η περίπτωση ισχύει όταν για παράδειγμα η εποχικότητα παραμένει σταθερή με την αλλαγή της τάσης, κυρίως για βραχυπρόθεσμες προβλέψεις (3 έως 5 μήνες). Δηλαδή στο προσθετικό μοντέλο το εύρος των μεταβολών στη χρονοσειρά παραμένει σταθερό ως προς το χρόνο.

$$Y_t = T_t + S_t + C_t + E_t$$

### **Πολλαπλασιαστικό μοντέλο (multiplicative model)**

Το δεύτερο μοντέλο που συνδέει τις παραπάνω συνιστώσες είναι το πολλαπλασιαστικό μοντέλο στο οποίο οι συνιστώσες συσχετίζονται. Στο εκάστοτε μοντέλο μόνο η μεταβλητή τάση ( $T_t$ ) είναι στην ίδια μονάδα μέτρησης με την πραγματική τιμή ( $Y_t$ ), οι υπόλοιπες συνιστώσες λειτουργούν ως δείκτες, δηλαδή καθαροί αριθμοί. Αυτό το μοντέλο ταιριάζει όταν η εποχικότητα μίας χρονοσειράς αλλάζει βάσει της τάσης. Δηλαδή, το εύρος των μεταβολών στη χρονοσειρά αυξάνει ως προς το χρόνο.

$$Y_t = T_t \cdot S_t \cdot C_t \cdot E_t$$

## **2.6 Στατιστικοί Δείκτες**

Παρακάτω παρατίθεται ορισμένοι στατιστικοί δείκτες που θα χρησιμοποιηθούν στα παρούσα εργασία.

### **Μέση τιμή (mean)**

Η μέση τιμή προκύπτει από το πηλίκο του αθροίσματος των τιμών της χρονοσειράς προς το πλήθος τους.

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

### Διακύμανση ή Διασπορά (variance)

Η διακύμανση αποτελεί ένα μέτρο διασποράς. Ως διασπορά ορίζεται η κατανομή των τιμών της χρονοσειράς γύρω από την μέση τιμή και υπολογίζεται από τον μέσο όρο των τετραγώνων των αποκλίσεων  $Y_t$  των από την μέση τιμή  $\bar{Y}$ .

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

### Αυτοσυσχέτιση (autocorrelation)

Ο συντελεστής αυτοσυσχέτισης είναι ένας στατιστικός δείκτης που περιγραφεί την τυχαιότητα της χρονοσειράς. Παίρνει τιμές από  $[-1,1]$ , με τα άκρα να σημαίνουν ότι οι μεταβλητές σχετίζονται ισχυρά μεταξύ τους, ενώ το 0 να δηλώνει ότι είναι ασυσχέτιστες. Με μία καθυστέρηση  $k$  ορίζεται ως εξής:

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y}_1) \cdot (Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

Ο συντελεστής αυτοσυσχέτισης έχει νόημα μόνο σε στάσιμες χρονοσειρές. Αν προσπαθήσουμε να υπολογίσουμε την αυτοσυσχέτιση σε μη στάσιμες χρονοσειρές με τάση ως προς την υστέρηση, παρατηρούμε πολύ υψηλές τιμές που ξεκινούν να φθίνουν μετά την υστέρηση. Αυτό οφείλεται στην παρουσία τάσης, που δημιουργεί ισχυρές συσχετίσεις μεταξύ κοντινών χρονικών διαστημάτων.

Από το γράφημα της αυτοσυσχέτισης συναρτήσει της καθυστέρησης  $k$ , (autocorrelation function – ACF) μπορεί κανείς να μελετήσει την ύπαρξη τάσης και εποχικότητας, καθώς και την συσχέτιση μεταξύ των μεταβλητών.

Πιο συγκεκριμένα, υψηλές τιμές στον συντελεστή αυτοσυσχέτισης μπορεί να σημαίνει την ύπαρξη εποχικότητας. Αν τα δεδομένα παρουσιάζουν εποχικότητα κάθε 12 μήνες, θα παρατηρηθούν υψηλές τιμές στην καθυστέρηση 12. Από την άλλη, αν οι τιμές των αυτοσυσχετίσεων φθίνουν στο μηδέν με γρήγορο ρυθμό, σημαίνει ότι η σειρά είναι στάσιμη και αντίστοιχα αν φθίνουν με αργό ρυθμό καθώς αυξάνει ο αριθμός των καθυστερήσεων, τότε πιθανόν η σειρά να είναι μη στάσιμη.

## ΚΕΦΑΛΑΙΟ 3

### Μέθοδοι Πρόβλεψης

Στην σύγχρονη εποχή που διανύουμε είναι φλέγον ζήτημα η πρόβλεψη μελλοντικών συνθηκών, κυρίως στον τομέα της βιομηχανίας. Την τελευταία δεκαετία, η επιστήμη μελετά τους διάφορους αλγορίθμους πρόβλεψης, με σκοπό την βελτιστοποίηση τους αλλά και την δημιουργία νέων. Στόχος είναι αφενός η ορθή αξιολόγηση των αποτελεσμάτων, και αφετέρου η επιλογή της μεθόδου πρόβλεψης που θα ελαχιστοποιεί το σφάλμα και θα προσφέρει ακριβείς προβλέψεις. [9] [10] [11]

Σημαντικές προβλέψεις αποτελούν:

- οι τιμές των μετοχών του χρηματιστηρίου,
- η ώρα με την μέγιστη κατανάλωση ηλεκτρικής ενέργειας,
- ο αριθμός των χρηστών ενός κοινωνικού δικτύου με στόχο την έγκαιρη αναβάθμιση των συστημάτων
- η τιμές της θερμοκρασίας, όχι τόσο για καθημερινή χρήση όσο και για την μακροπρόθεσμη πρόβλεψη κλιματικών αλλαγών
- ο όγκος των ηλεκτρικών εγγράφων σε ηλεκτρονικές βιβλιοθήκες
- ο αριθμός των γεννήσεων για το επόμενο έτος ανά πόλη και νοσοκομείο
- η τιμή των καυσίμων σε μία πόλη
- το ποσοστό της ανεργίας για κάθε επόμενο τρίμηνο κ.ά.

Η πρόβλεψη των μελλοντικών τιμών μπορεί να γίνει με τις ακόλουθες μεθόδους.

### 3.1 Λευκός Θόρυβος (White Noise)

Μια χρονοσειρά ονομάζεται λευκός θόρυβος αν δεν έχει κανένα ευκρινές σχήμα ή πρότυπο. Συμβολίζεται ως εξής:

$$Y(t)=\varepsilon_t$$

Όπου  $\varepsilon_t$  μία ακολουθία τυχαίων μεταβλητών με  $t = 0, \pm 1, \pm 2, \dots$ , για την οποία ισχύει ότι η μέση τιμή ίση με μηδέν και η διακύμανση είναι σταθερή όλες τις χρονικές στιγμές. Επιπλέον, όλες οι παρατηρήσεις της σειράς είναι ασυσχέτιστες μεταξύ τους, δηλαδή η αυτοσυνδιακύμανση της s-οστής είναι μηδενική.

$$E(\varepsilon_t) = 0$$

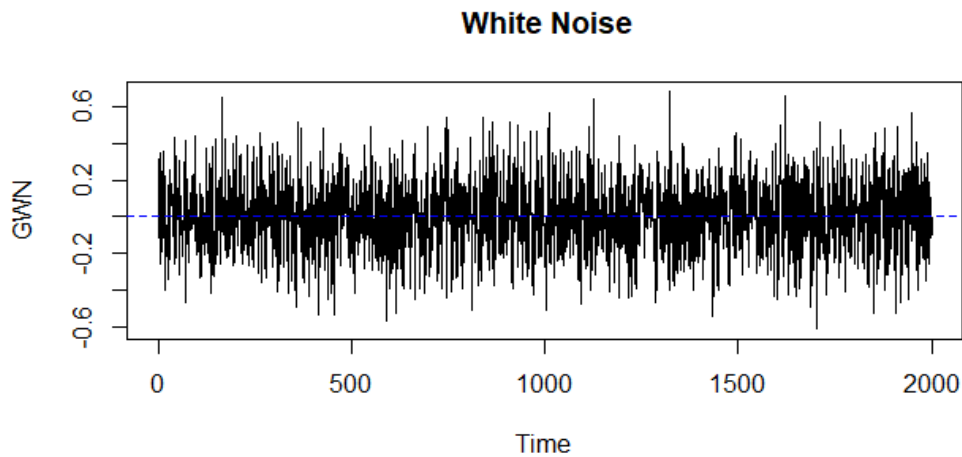
$$V(\varepsilon_t) = \sigma^2$$

$$\text{Cov}(\varepsilon_t, \varepsilon_{t-s}) = 0, s \neq 0$$

Άρα, μια τέτοια σειρά είναι πάντα στάσιμη και έχει μηδενικούς συντελεστές αυτοσυσχέτισης και μερικής αυτοσυσχέτισης. Χαρακτηριστικότερο παράδειγμα αποτελούν οι τυχεροί αριθμοί ΤΖΟΚΕΡ που καταγράφονται κάθε εβδομάδα. Σε αυτές τις περιπτώσεις οι προηγούμενες τιμές της χρονοσειράς δεν χρησιμεύουν και δεν σχετίζονται καθόλου με την πρόβλεψη των μελλοντικών τυχερών αριθμών που θα κληρωθούν.

Ακολουθεί παράδειγμα Λευκού Θορύβου:

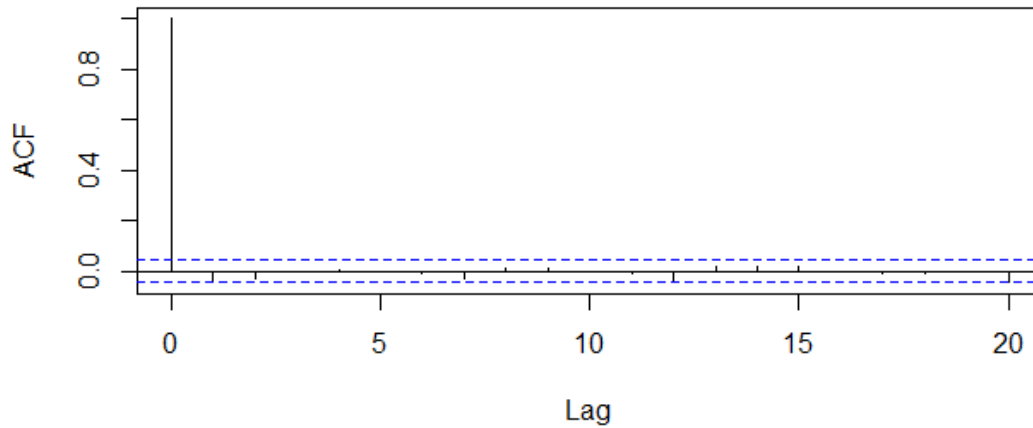
Έγινε παραγωγή τυχαίου Γκαουσιανού δείγματος κανονικής κατανομής στη R, με μέγεθος 2000 και τυπική απόκλιση 0.2.



Λευκός Θόρυβος με Γκαουσιανή Κατανομή

Στο δείκτη αυτοσυσχέτισης που ακολουθεί, γίνεται φανερό ότι πρόκειται για μία στάσιμη χρονοσειρά με μηδενική αυτοσυσχέτιση για  $\text{lag} \geq 0$ .





Συνάρτηση Συντελεστή Αυτοσυσχέτισης Λευκού Θορύβου

### 3.2 Τυχαίος Περίπατος (Random Walk)

Ο τυχαίος περίπατος (random walk) είναι μια μη-στάσιμη χρονοσειρά, όπου κάθε στοιχείο της  $y_t$  προκύπτει όταν στο προηγούμενο στοιχείο της  $y_{t-1}$  προστεθεί μια τυχαία μεταβλητή  $\varepsilon_t$ . Τέτοιες χρονοσειρές χαρακτηρίζονται από μακριές περιόδους, εμφανών τάσεων αυξητικών ή πτωτικών, με απότομες αλλαγές.

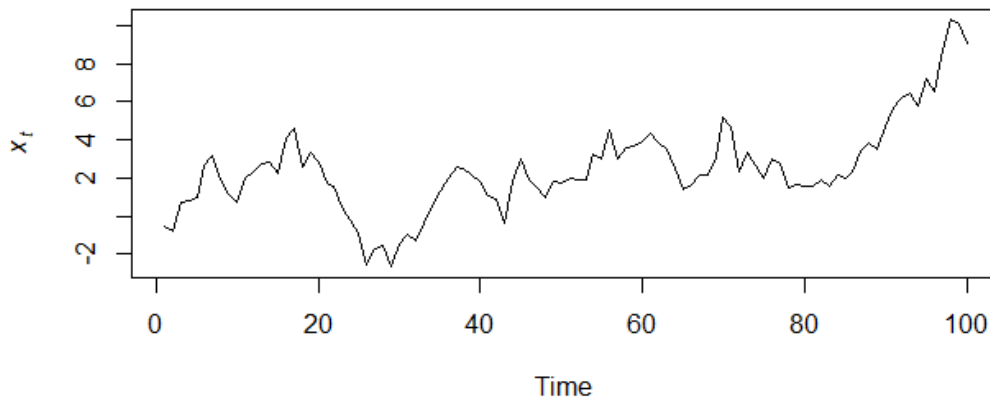
Δίνεται από την σχέση:

$$y_t = y_{t-1} + \varepsilon_t$$

Όπου  $\varepsilon_t$  η χρονοσειρά λευκού θορύβου. Αν υποθέσουμε ότι η τιμή  $y_0$  είναι μηδέν, τότε η χρονοσειρά του λευκού θορύβου είναι ανεξάρτητη από το  $y_0$ . Για το λόγο αυτό, η μέση τιμή του τυχαίου περιπάτου είναι μηδέν και η διακύμανση μεγαλώνει με το χρόνο.

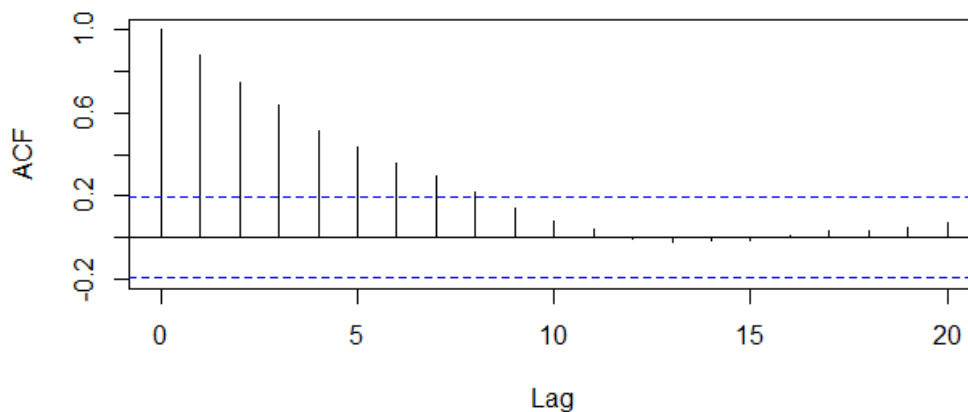
$$E(y_t) = 0$$

$$V(y_t^2) = t \sigma^2$$



Τυχαίος Περίπατος

Παρακάτω ο συντελεστής αυτοσυσχέτισης, υποδηλώνει την ύπαρξη υψηλής αυτοσυσχέτισης ως το lag = 8.



Συνάρτηση Συντελεστή Αυτοσυσχέτισης Τυχαίου Περίπατου

### 3.3 Μέθοδοι Πρόβλεψης με Κινητούς Μέσους ( Moving Average)

Οι τεχνικές αυτές ονομάζονται μέθοδοι εξομάλυνσης, διότι η δημιουργία των προβλέψεων προέρχεται από την εξομάλυνση των τιμών της χρονοσειράς και μέσω της εξομάλυνσης διακρίνεται καλύτερα η συμπεριφορά των τιμών.

Μέθοδοι εξομάλυνσης:

- **Μέθοδος του Απλού Κινητού Μέσου (Simple Moving Average)**

Είναι κατάλληλη για να απομακρύνει την τυχαία μεταβλητότητα στην πρόβλεψη, όταν η χρονοσειρά είναι στάσιμη δηλαδή οι τιμές κυμαίνονται γύρω από έναν σταθερό μέσο όρο και κατ' επέκταση απουσιάζει η τάση και η εποχικότητα. Είναι πολύ σημαντικό να

επιλεγεί το κατάλληλο διάστημα για τον κινητό μέσο. Όσο μεγαλύτερο είναι το επιλεγμένο διάστημα τόσο περισσότερο εξομαλύνεται το τυχαίο στοιχείο στην πρόβλεψη. Στα αρνητικά της μεθόδου είναι η εξάρτηση από  $n$  παρελθούσες παρατηρήσεις της σειράς, με ίδια βάρη σε όλες τις παρατηρήσεις, ακόμα και αν οι τελευταίες περιέχουν περισσότερη πληροφορία.

- **Μέθοδος του Διπλού Κινητού Μέσου (*Double Moving Average*)**

Η μέθοδος αυτή επιλέγεται για χρονοσειρές που εμφανίζουν ανοδική ή πτωτική γραμμική τάση. Υπολογίζεται ένας δεύτερος κινητός μέσος μετά τον απλό κινητό, λαμβάνοντας υπόψιν και την γραμμική τάση της χρονοσειράς. Η μέθοδος αυτή, μπορεί να χρησιμοποιηθεί για τη διενέργεια προβλέψεων για περισσότερες από μια μελλοντικές περιόδους ( $h > 1$ ), ενώ για  $h=1$  δίνει την πρόβλεψη για την επόμενη περίοδο.

### **3.4 Μέθοδοι εξομάλυνσης (Smoothing Methods)**

- **Μέθοδος της Απλής Εκθετικής Εξομάλυνσης (*Simple exponential smoothing*)**

Το πρόβλημα του απλού κινητού μέσου έρχεται να λύσει η απλή εκθετική εξομάλυνση. Σε ορισμένες χρονοσειρές, τα πιο πρόσφατα δεδομένα περιέχουν περισσότερες πληροφορίες από τα παλαιότερα, με συνέπεια να πρέπει να δίνεται μεγαλύτερο βάρος σε αυτά.

Ένα ρεαλιστικό μοντέλο για μια χρονοσειρά θα ήταν να θεωρήσουμε κάθε παρατήρηση ως συνιστώσα μιας σταθερής συνιστώσας  $b$  και ενός σφάλματος  $\epsilon$ .

$$X_t = b + \epsilon_t$$

Η σταθερά  $b$  είναι σχετικά σταθερή σε κάθε τμήμα της σειράς, αλλά μπορεί αργά να αλλάξει. Εάν είναι απαραίτητο, τότε ένας τρόπος απομόνωσης της πραγματικής τιμής του  $b$ , και συνεπώς του προβλέψιμου μέρους της σειράς, είναι να υπολογίσουμε ένα είδος κινητού μέσου όρου, όπου η τρέχουσα και η αμέσως προηγούμενη παρατήρηση αποδίδουν μεγαλύτερο βάρος από τις αντίστοιχες παλαιότερες παρατηρήσεις. Η απλή εκθετική εξομάλυνση επιτυγχάνει ακριβώς αυτή τη στάθμιση, τα βάρη μειώνονται

εκθετικά στη πάροδο του χρόνου, με σκοπό να εκχωρούνται εκθετικά μεγαλύτερα βάρη στα πιο πρόσφατα δεδομένα.

Δίνεται από τον τύπο:

$$S_t = \alpha \bullet X_t + (1 - \alpha) \bullet S_{t-1}$$

Όταν εφαρμοστεί αναδρομικά σε κάθε διαδοχική παρατήρηση της σειράς, κάθε νέα ομαλοποιημένη τιμή (πρόβλεψη) υπολογίζεται ως ο σταθμισμένος μέσος όρος της τρέχουσας παρατήρησης και της προηγούμενης ομαλοποιημένης παρατήρησης. η προηγούμενη ομαλοποιημένη παρατήρηση υπολογίστηκε με τη σειρά της από την προηγούμενη παρατηρούμενη τιμή και την εξομαλυνθείσα τιμή πριν από την προηγούμενη παρατήρηση και ούτω καθεξής. Έτσι, στην πραγματικότητα, κάθε εξομαλυνθείσα τιμή είναι ο σταθμισμένος μέσος όρος των προηγούμενων παρατηρήσεων, όπου τα βάρη μειώνονται εκθετικά ανάλογα με την τιμή της παραμέτρου  $\alpha$  (alpha).

Εάν το  $\alpha$  είναι ίσο με 1 τότε οι προηγούμενες παρατηρήσεις αγνοούνται εξ ολοκλήρου, εάν είναι ίσο με 0, τότε η τρέχουσα παρατήρηση αγνοείται εξ ολοκλήρου και η ομαλοποιημένη τιμή αποτελείται εξ ολοκλήρου από την προηγούμενη ομαλοποιημένη τιμή. Έτσι όλες οι ομαλοποιημένες τιμές θα να είναι ίσες με την αρχική ομαλοποιημένη τιμή  $S_0$ .

- **Μέθοδος της Διπλής Εκθετικής Εξομάλυνσης (*Double exponential smoothing*) (Brown)**

Η μέθοδος αυτή χρησιμοποιείται κυρίως σε χρονοσειρές που παρουσιάζουν τάση. Η λογική της μεθόδου μοιάζει με αυτή του Διπλού Κινητού Μέσου, δηλαδή εφαρμόζεται διπλή εξομάλυνση στις τιμές της χρονοσειράς λαμβάνοντας υπόψιν και την τάση της χρονοσειράς, με τη διαφορά ότι απαιτούνται πολύ λιγότερες παρατηρήσεις. Επιπλέον, συγκριτικά την Απλή Εκθετική Εξομάλυνση, η Διπλή μπορεί να προβλέψει για περισσότερες από μία μελλοντικές περιόδους (βραχυχρόνιες προβλέψεις).

- **Μέθοδος της Εκθετικής Εξομάλυνσης με προσαρμογή στην τάση (*Exponential smoothing adjusted for trend*) (Holt)**

Όταν διαθέτουμε χρονοσειρές που περιγράφονται από το αθροιστικό μοντέλο, με παρουσία τάσης και απουσία εποχικότητας, τότε συνιστάται η μέθοδος Holt για

βραχυπρόθεσμες προβλέψεις. Μία τέτοια χρονοσειρά μπορεί να είναι τα δεδομένα που αναφέρονται στη διάμετρο της μέσης των γυναικών για μία 10ετία. Η μέθοδος Holt έχει δύο παραμέτρους εξομάλυνσης, την παράμετρο  $\alpha$  για την εξομάλυνση των τιμών της χρονοσειράς και την παράμετρο  $\beta$  για την εξομάλυνση της τάσης, σε αντίθεση με τη μέθοδο της διπλής εκθετικής εξομάλυνσης που έχει μόνο μια την  $\alpha$ . Οι τιμές των παραμέτρων  $\alpha$  και  $\beta$  για μια συγκεκριμένη χρονοσειρά είναι αυτές που ελαχιστοποιούν την τιμή του κριτηρίου που έχει επιλεγεί για τα δεδομένα της χρονοσειράς (π.χ. το MSE). Οι τιμές όπως και στην απλή Εκθετική Εξομάλυνση κυμαίνονται ανάμεσα στο 0 και το 1. Όταν η τιμή βρίσκεται κοντά στο 0, σημαίνει ότι στις πιο πρόσφατες παρατηρήσεις τοποθετούνται τα μικρά βάρη, κατά την διαδικασία της πρόβλεψης. Αξιοσημείωτο είναι ότι, με τη μέθοδο Holt οι τιμές της τάσης εξομαλύνονται απευθείας, ενώ με τη μέθοδο Brown η τάση προσδιορίζεται από τον παράγοντα προσαρμογής  $\gamma$ , αφού προηγουμένως εξομαλυνθούν δύο φορές οι τιμές της χρονοσειράς. Αυτό σημαίνει ότι με τη μέθοδο Holt γίνεται καλύτερη εκτίμηση των τιμών της τάσης, απ' ό,τι με τη μέθοδο Brown που είναι πολύ ευαίσθητη στις τυχαίες διακυμάνσεις της χρονοσειράς. Για αυτό το λόγο η μέθοδος Holt χρησιμοποιείται πιο συχνά.

- **Μέθοδος της Εκθετικής Εξομάλυνσης με προσαρμογή στην τάση και στην εποχικότητα (Exponential smoothing adjusted for trend and seasonality)**

Η μέθοδος Winters είναι κατάλληλη όταν η χρονοσειρά παρουσιάζει τάση, εποχικότητα και χαρακτηρίζεται από το αθροιστικό μοντέλο. Μία τέτοια χρονοσειρά, μπορεί να είναι οι μηνιαίες πωλήσεις ενός καταστήματος με souvenir στην παραλία της Αυστραλίας. Η μέθοδος Holt έχει τρεις παραμέτρους εξομάλυνσης, την παράμετρο  $\alpha$  για την εξομάλυνση των τιμών της χρονοσειράς, την παράμετρο  $\beta$  για την εξομάλυνση της τάσης και την παράμετρο  $\gamma$  για την εξομάλυνση της εποχικότητας.

### 3.5 Μοντέλα για Στάσιμες Χρονοσειρές

- **Αυτοπαλίνδρομα Μοντέλα (Autoregressive (AR) Models)**

Ένα Αυτοπαλίνδρομο Υπόδειγμα τάξης  $p$  (AR( $p$ )) ορίζεται από την σχέση:

$$Y_t = a_0 + a_1 Y_{t-1} + \dots + a_p Y_{t-p} + \varepsilon_t$$

Η παράμετρος  $a_0$  σχετίζεται με τη μέση τιμή της χρονοσειράς και τα  $a_1, a_p$  είναι αυτοπαλινδρομούμενοι παράμετροι. Το  $p$  υποδηλώνει την τάξη του αυτοπαλινδρομού μοντέλου και αναφέρεται στο μήκος της χρονικής υστέρησης. Η μεταβλητή  $\varepsilon_t$  θεωρείται ότι είναι λευκός θόρυβος. Ο όρος αυτοπαλινδρομο δημιουργείται επειδή η εξαρτημένη μεταβλητή  $Y_t$  παλινδρομεί τις προηγούμενες τιμές της της της μεταβλητής  $Y_t$ .

### AR (1):

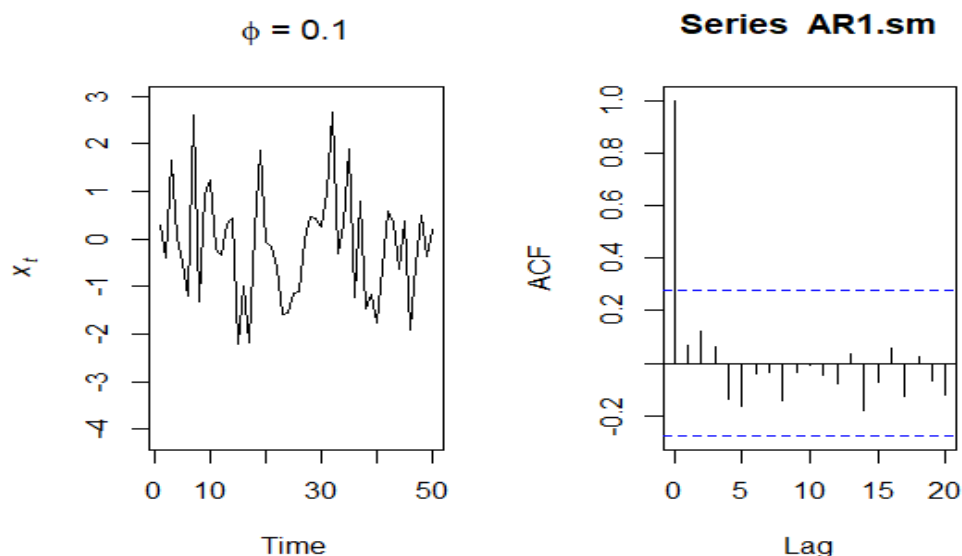
Εάν η τάξη είναι  $p=1$ , το μοντέλο δίνεται από την σχέση:

$$Y_t = a_0 + a_1 Y_{t-1} + \varepsilon_t$$

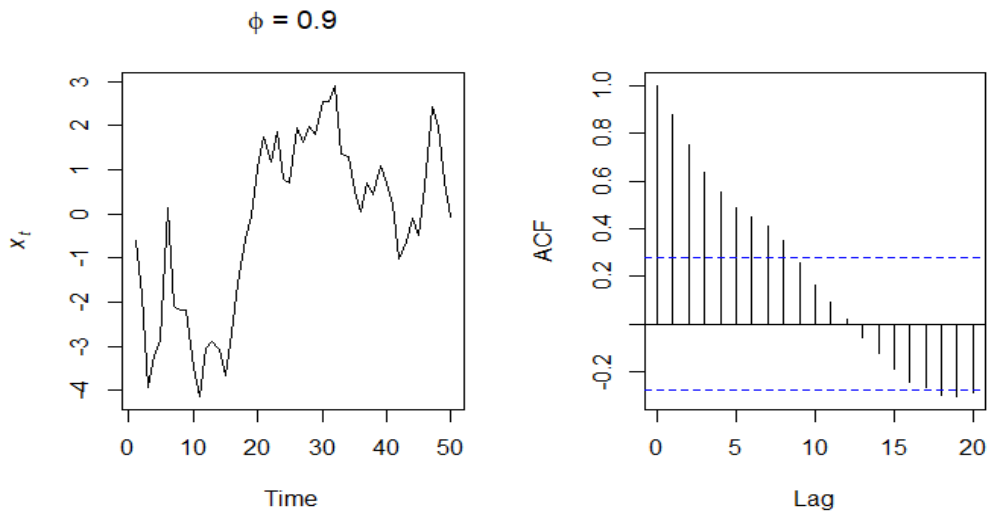
Ο συντελεστής  $a_1$  παίρνει τιμές ανάμεσα στο -1 και 1. Για  $a_1=0$  το μοντέλο είναι ισοδύναμο με ένα μοντέλο λευκού θορύβου, ενώ για  $a_1=1$  το μοντέλο είναι ισοδύναμο με ένα μοντέλο τυχαίου περιπάτου.

Γενικά, ένα AR(p) μοντέλο χρησιμοποιείται όταν οι συντελεστές αυτοσυσχέτισης φθίνουν εκθετικά στο μηδέν και ταυτόχρονα υπάρχουν  $p$  στατιστικοί σημαντικοί συντελεστές μερικής αυτοσυσχέτισης.

Παρακάτω αναπαρίσταται ένα μοντέλο AR (1), στο πρώτο σχήμα με μικρό δείκτη coefficient = 0.1 και στο δεύτερο με coefficient = 0.9 .



AR(1) με coefficient 0.1



AR(1) με coefficient 0.9

Βάσει των γραφημάτων, παρατηρούμε ότι η χρονοσειρά με το μικρότερο AR coefficient, παραμένει κοντά στην περιοχή του 0, γεγονός που υποδηλώνει ότι ακολουθεί την λογική Λευκού Θορύβου και είναι στάσιμη χρονοσειρά, ως της τη μέση τιμή και ως της την διακύμανση. Εν αντίθεσή, η χρονοσειρά με μεγαλύτερο coefficient που τείνει στο 1, κινείται περισσότερο στο χώρο και ακολουθεί την λογική Τυχαίου Περίπατου ως της τη μέση τιμή και διακύμανση ταυτόχρονα.

ο **Κινητοί Μέσοι (Moving Average (MA) Models)**

Ένα Υπόδειγμα Κινητών Μέσων τάξης q (MA(q)) ορίζεται από την σχέση:

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

Όπου οι παράμετροι  $\mu$  και τα  $\theta_1, \theta_q$  είναι οποιοιδήποτε πραγματικοί αριθμοί. Το q υποδηλώνει την τάξη της υψηλότερης υστέρησης, για την οποία γίνεται υπόθεση ότι είναι λευκός θόρυβος και το  $\varepsilon_t$  είναι η χρονοσειρά του λευκού θορύβου. Ο όρος κινητός μέσος οφείλεται στο γεγονός ότι η  $Y_t$  εμφανίζεται ως ένα σταθμισμένο άθροισμα των τιμών της  $\varepsilon_t$ .

MA (1):

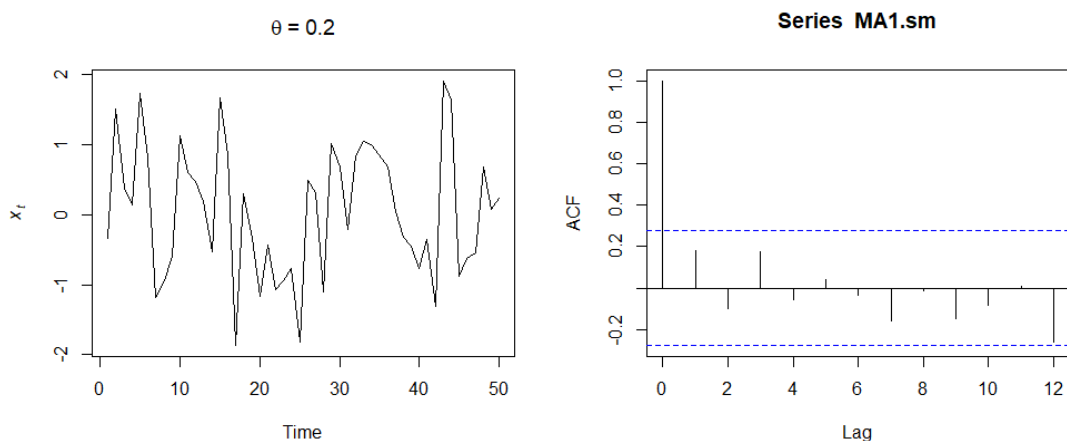
Εάν η τάξη είναι q=1, το μοντέλο δίνεται από την σχέση:

$$Y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1}$$

Η παρατήρηση  $Y_t$  εξαρτάται από την προηγούμενη παρατήρηση και την επόμενη. Για παράδειγμα η  $Y_3$  θα εξαρτάται από την  $Y_2$  και την  $Y_4$ . Ο συντελεστής  $\theta_1$  παίρνει τιμές

από -1 μέχρι 1. Μια MA(1) διαδικασία μπορεί να χαρακτηριστεί ως αντιστρέψιμη, αν μπορεί να διατυπωθεί ως μία αυτοπαλίνδρομη διαδικασία με άπειρους όρους. Σε αυτή την περίπτωση η MA(1) μετατρέπεται σε AR( $\infty$ ).

Γενικά, ένα μοντέλο MA(q) χρησιμοποιείται όταν οι συντελεστές μερικής αυτοσυσχέτισης φθίνουν εκθετικά στο μηδέν και συγχρόνως υπάρχουν q στατιστικοί σημαντικοί συντελεστές αυτοσυσχέτισης. Επιπρόσθετα, οι κινητοί μέσοι χρησιμοποιούνται για χρονοσειρές όπου τα γεγονότα επηρεάζουν άμεσα τη στιγμή που συμβαίνουν, αλλά μετά το τέλος των γεγονότων, τα αποτελέσματα συνεχίζουν να επηρεάζουν. Ένα τέτοιο παράδειγμα είναι, η οικονομία που επηρεάζεται κατά την διάρκεια της πανδημίας του κορωνοϊού, αλλά και κατά της επόμενους μήνες ακόμα και όταν θα έχει ξεπεραστεί.



MA(1) με coefficient 0.9

Παρατηρείται ότι για lag > q ο acf δείκτης τείνει στο 0.

ο **Αυτοπαλινδρόμενοι Κινητοί Μέσοι Όροι (Autoregressive Moving Average (ARMA) Models)**

Ορισμένες χρονοσειρές δεν μπορούν να μοντελοποιηθούν αποκλειστικά ως AR(p) ή MA(q) χρονοσειρές, εφόσον μπορεί να παρουσιάζουν χαρακτηριστικά και από της δύο κατηγορίες. Για το λόγο αυτό δημιουργείται το μοντέλο Αυτοπαλινδρόμενοι Κινητοί Μέσοι Όροι (ARMA) που έρχεται να περιγράψει τέτοιες χρονοσειρές.

Το μοντέλο ARMA(p,q) ορίζεται από την σχέση:

$$Y_t = a_0 + a_1 Y_{t-1} + \dots + a_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$



Οι παράμετροι  $\alpha_i$  και  $\theta_i$  περιορίζονται στο  $(-1,1)$ . Ουσιαστικά το μοντέλο είναι ο συνδυασμός  $p$  αυτοπαλίνδρομων όρων και  $q$  όρων κινητού μέσου. Η στασιμότητα της χρονοσειράς ορίζεται από το AR μέρος και η αντιστρεψιμότητα από το MA μέρος. Είναι προφανές ότι ένα καθαρά αυτοπαλίνδρομο μοντέλο ή ένα καθαρό μοντέλο κινητού μέσου μπορούν να θεωρηθούν ως ειδικές περιπτώσεις μιας ARMA χρονοσειράς ως εξής:

- $AR(p) = ARMA(p, 0)$
- $MA(q) = ARMA(0, q)$

### 3.6 Μοντέλα για μη στάσιμες χρονοσειρές

Τα τρία τελευταία μοντέλα που παρουσιάστηκαν αναφέρονται όλα σε στάσιμες χρονοσειρές, γεγονός που υποδηλώνει ότι ο μέσος, η διακύμανση και οι αυτοδιακυμάνσεις δεν εξαρτώνται από το χρόνο. Ο μέσος και η διακύμανση παραμένουν σταθεροί, ενώ οι αυτοδιακυμάνσεις εξαρτώνται μόνο από τη χρονική υστέρηση.

Είναι σημαντικό κατά την διαδικασία πρόβλεψης η χρονοσειρά να είναι στάσιμη προκειμένου να αποφευχθούν διάφορα λάθη. Ωστόσο, στην πραγματικότητα οι περισσότερες χρονοσειρές, όπως δείκτες τιμών, ανεργία, οικονομικές χρονοσειρές, στην αρχική τους μορφή είναι μη στάσιμες, για το λόγο αυτό χρειαζόμαστε έναν τρόπο για να τις μετατρέψουμε σε στάσιμες.

#### Μετατροπή Μη Στάσιμης Χρονοσειράς σε Στάσιμη

Μία μη στάσιμη χρονοσειρά μετατρέπεται σε στάσιμη παίρνοντας τις απαραίτητες διαφορές. Όταν μια χρονοσειρά μετατρέπεται σε στάσιμη, παίρνοντας τις πρώτες διαφορές η σειρά ονομάζεται ολοκληρωμένη πρώτης τάξης και συμβολίζεται με  $I(1)$ . Εάν μετατρέπεται σε στάσιμη παίρνοντας τις δεύτερες διαφορές, είναι ολοκληρωμένη δεύτερης τάξης και συμβολίζεται με  $I(2)$ . Μια στάσιμη σειρά, όπως ο λευκός θόρυβος, θεωρείται ολοκληρωμένη σειρά μηδενικής τάξεως δηλαδή  $I(0)$ . Γενικεύοντας, εάν  $d$  είναι ο αριθμός των διαφορών που μετατρέπει μια χρονοσειρά σε στάσιμη, η σειρά ονομάζεται **ολοκληρωμένη  $d$  τάξεως** (integrated order) και συμβολίζεται με  **$I(d)$** .

Ορίζεται από την σχέση:

$$\Delta y = y_t - y_{t-1}$$

### ○ **Ανάλυση ARIMA (Autoregressive Intergated Moving Average)**

Τα μοντέλα ARIMA περιγράφουν την διαχρονική συμπεριφορά ενός φυσικού μεγέθους. Χρησιμοποιούνται για να προβλέψουν τις μελλοντικές τιμές μίας στάσιμης σειράς, λαμβάνοντας υπόψιν τις παρελθοντικές τιμές. Για παράδειγμα, για μία τυχαία χρονοσειρά, τα μοντέλα ARIMA βασίζονται στην αλληλεξάρτηση μεταξύ των τιμών που λαμβάνει η χρονοσειρά για τις διάφορες χρονικές στιγμές. Τα ARMA μοντέλα που είδαμε παραπάνω μπορούν να χρησιμοποιηθούν μόνο για στάσιμες χρονοσειρές, αλλά επεκτείνονται και σε μη στάσιμες χρονοσειρές με τη χρήση της μεθόδου των διαφορών. Σε αυτή την περίπτωση τα μοντέλα μετατρέπονται σε ολοκληρωμένα αυτοπαλίνδρομα μοντέλα κινητού μέσου τάξεως  $(p, d, q)$  και συμβολίζονται με  $ARIMA(p, d, q)$ . Συνεπώς, το μοντέλο  $ARIMA(p, d, q)$ , είναι μια διαδικασία που διαφορίζει την σειρά  $d$  φορές και παράγει εντέλει ένα  $ARMA(p, q)$  μοντέλο. [11]

Συγκεκριμένα, οι τρεις παράμετροι αυτού του μοντέλου είναι: [12]

- **AR** :  $p$  = δηλώνει την τάξη του αυτοπαλίνδρομου μοντέλου
- **I** :  $d$  = δείχνει τον αριθμό των διαφορών (η τάξη της διαφορίσης) που απαιτούνται για να γίνει η σειρά στάσιμη
- **MA** :  $q$  = που δηλώνει την τάξη του κινητού μέσου

### ***Ειδικές Περιπτώσεις***

$ARIMA(0,0,0)$  = Λευκός Θόρυβος

$ARIMA(0,1,0)$  = Τυχαίος Περίπατος

$ARIMA(1,0,0)$  = AR (1)

$ARIMA(0,0,1)$  = MA (1)

### **Μέθοδος Box-Jenkins**

Τα μοντέλα ARIMA έχουν μελετηθεί εκτεταμένα από τους Box και Jenkins (1970), σε βαθμό που τα ονόματά τους είναι άρρηκτα συνδεδεμένα με τα μοντέλα ARIMA και τις

εφαρμογές τους στην ανάλυση και την πρόβλεψη χρονοσειρών. Η προσέγγιση Box-Jenkins αποτελεί μία μέθοδο εύρεσης του κατάλληλου στατιστικού μοντέλου ARIMA, που παριστάνει όσο πιο ρεαλιστικά γίνεται την διαδικασία από την οποία προήλθαν τα δεδομένα και θεωρείται ως καταλληλότερο μοντέλο, για την εκάστοτε σειρά, για την διαμόρφωση των προβλέψεων. Η μέθοδος αυτή εφαρμόζεται όταν η χρονοσειρά δεν είναι στάσιμη και περιλαμβάνει τα εξής τέσσερα στάδια: [13] [14]

#### ο *Ταυτοποίηση*

Σε αυτό το στάδιο καθορίζονται οι τιμές των παραμέτρων  $p$ ,  $d$ ,  $q$ . Αρχικά, ελέγχεται με την βοήθεια των συντελεστών αυτοσυσχέτισης αν η σειρά είναι στάσιμη. Αν είναι στάσιμη στο  $d$  αντιστοιχεί 0, καθώς δεν απαιτείται καμία διαφορίση προκειμένου να γίνει στάσιμη. Ωστόσο, αν είναι μη στάσιμη, πρέπει να πάρουμε τις κατάλληλες διαφορές προκειμένου να την μετατρέψουμε σε στάσιμη. Εν συνεχεία, καθορίζεται η τάξη  $p$  της αυτοπαλίνδρομης διαδικασίας και η τάξη  $q$  της διαδικασίας κινητού μέσου.

#### ο *Εκτίμηση*

Σ' αυτό το στάδιο γίνεται η εκτίμηση των  $p$  παραμέτρων  $\alpha_0, \alpha_1, \dots, \alpha_p$  της αυτοπαλίνδρομης διαδικασίας και των  $q$  παραμέτρων  $\theta_0, \theta_1, \dots, \theta_q$  της διαδικασίας κινητού μέσου. Αν η σειρά που εξετάζεται είναι μόνο αυτοπαλίνδρομη, οι παράμετροί της, μπορούν να εκτιμηθούν με τη μέθοδο των ελαχίστων τετραγώνων. Αν όμως, περιέχει και όρους κινητού μέσου τότε για την εκτίμηση των παραμέτρων του κινητού μέσου θα χρησιμοποιηθούν μη γραμμικές μέθοδοι εκτίμησης.

#### ο *Διαγνωστικός Έλεγχος*

Σ' αυτό το στάδιο ελέγχεται αν το συγκεκριμένο υπόδειγμα ARIMA που προσδιορίστηκε παραπάνω είναι ικανοποιητικό και ταιριάζει (fitting) στα δεδομένα μας, για τις μελλοντικές προβλέψεις. Ελέγχεται ως προς την σημαντικότητα και τη σταθερότητα των συντελεστών του υποδείγματος, τις ιδιότητες των καταλοίπων και την προβλεπτική ικανότητα του υποδείγματος. Ο έλεγχος της σημαντικότητας ενός αριθμού συντελεστών αυτοσυσχέτισης γίνεται παράλληλα με τον έλεγχο των καταλοίπων. Μία διαδικασία ARIMA ταιριάζει κατάλληλα στη σειρά, όταν τα κατάλοιπα συμπεριφέρονται σαν μία διαδικασία λευκού θορύβου, δηλαδή δεν αυτοσυσχετίζονται μεταξύ τους. Αυτός ο έλεγχος γίνεται με την στατιστική διαδικασία Q των Box-Pierce. Η μηδενική υπόθεση είναι :

$H_0 = \rho_1, \rho_2, \dots, \rho_m = 0$ , όπου  $\rho_i, i = 1, 2, \dots, m$  οι συντελεστές αυτοσυσχέτισης.

Η στατιστική Q των Box-Pierce ορίζεται ως:

$$Q_{BP} = T \sum_{s=2}^m \hat{\rho}^2$$

Όπου  $\hat{\rho}_s$  είναι δειγματικές αυτοσυσχετίσεις των καταλοίπων και T ο αριθμός των παρατηρήσεων. Ο αριθμός των αυτοσυσχετίσεων ισούται με την τετραγωνική ρίζα του αριθμού των παρατηρήσεων, δηλαδή ισχύει  $m = \sqrt{T}$ . Η στατιστική  $Q_{BP}$  ακολουθεί την κατανομή  $\chi^2$  με  $m - p - q$  βαθμούς ελευθερίας. Η μηδενική υπόθεση θα απορρίπτεται αν η τιμή  $Q_{BP}$  είναι μεγαλύτερη από την κρίσιμη τιμή της κατανομής  $\chi^2_\alpha$ , για δεδομένο επίπεδο σημαντικότητας  $\alpha$ . Οπότε θα ισχύουν τα παρακάτω:

Απόρριψη :  $H_0$ , αν  $Q_{BP} > \chi^2_\alpha$

Αποδοχή :  $H_0$ , αν  $Q_{BP} \leq \chi^2_\alpha$

Επιπλέον, σ' αυτό το στάδιο γίνεται ο έλεγχος της τάξης του υποδείγματος, μέσω της διαδικασίας της υπερπροσαρμογής (overfitting). Προκειμένου να ελεγχθεί η καταλληλότητα του εκτιμώμενου υποδείγματος, συγκρίνεται με ένα άλλο μοντέλο μεγαλύτερης τάξης. Δηλαδή, το εκτιμώμενο υπόδειγμα ARMA(p, q), συγκρίνεται με τα υποδείγματα ARIMA(p+1, q) και ARIMA(p, q+1) της αμέσως επόμενης τάξης. Το εκτιμώμενο υπόδειγμα είναι το καταλληλότερο για τα δεδομένα μας, αν οι συντελεστές στα μεγαλύτερα υποδείγματα δεν στατιστικά διαφορετικοί από το μηδέν. Αν αυτοί οι συντελεστές δεν είναι μηδέν, τότε υπάρχει κάποιο άλλο υπόδειγμα που να είναι πιο κατάλληλο για τα δεδομένα μας, απ' ότι το εκτιμημένο. [15]

### ○ **Πρόβλεψη**

Γνωρίζοντας πλέον το εκτιμώμενο υπόδειγμα και τις υπάρχουσες πληροφορίες για μια χρονοσειρά, μπορεί να παραχθεί η πρόβλεψη για τις επόμενες ζητούμενες περιόδους

### ○ **Ανάλυση SARIMA (Seasonal Autoregressive Intergated Moving Average)**

Τα μοντέλα SARIMA είναι μια άλλη κατηγορία των υποδειγμάτων ARIMA. Όταν μία σειρά έχει δεδομένα μικρής διάρκειας, όπως μηνιαία, τριμηνιαία τότε την χαρακτηρίζει η εποχικότητα. Το εποχικό μέρος του υποδείγματος ARIMA έχει παρόμοια δομή με αυτή ενός μη εποχικού υποδείγματος και γράφεται ως SARIMA (Sp, Sd, Sq) x ARIMA

$(p, d, q)$ . Στο εποχικό μέρος διεξάγονται πολλαπλασιασμοί της χρονικής υστέρησης  $S$  (δηλαδή τον αριθμό των περιόδων για μια εποχή) με τους συντελεστές  $p$  που είναι ο αριθμός των αυτοπαλίνδρομων εποχικών όρων (SAR) , τον αριθμό των εποχικών διαφορών  $d$  και το εύρος των στοιχείων των εποχικών όρων του κινητού μέσου  $q$  (SMA).

Σε τέτοιες σειρές, μπορεί να αφαιρεθεί η εποχικότητα και στη συνέχεια να χρησιμοποιηθεί η μέθοδος των Box-Jenkins ή, επικρατέστερη είναι η μέθοδος η οποία ενσωματώνει το εποχικό πρότυπο των στοιχείων στο κανονικό υπόδειγμα ARIMA και χρησιμοποιεί την μεθοδολογία Box-Jenkins. Η μέθοδος αυτή καταλήγει στην εκτίμηση υποδειγμάτων με περισσότερες παραμέτρους.

### 3.7 Επιλογή Μεθόδου Πρόβλεψης

Κατά την διαδικασία επιλογής της καταλληλότερης μεθόδου πρόβλεψης βασικός παράγοντας είναι η κατανόηση των κυρίων συνιστωσών μίας χρονολογικής σειράς. Ενδεικτικά ένας αναλυτής μπορεί να ακολουθεί το παρακάτω μοτίβο, όταν επιλέγει μέθοδο πρόβλεψης.

- i. Όταν μια χρονοσειρά είναι στάσιμη και η μέση τιμή της παραμένει σταθερή κατά το πέρας του χρόνου, τότε μπορεί να επιλεγθούν ο Απλός Κινητός Μέσος ή η μέθοδος ARMA.
- ii. Όταν η χρονοσειρά εμφανίζει τάση και η μέση τιμή της μεταβάλλεται κατά την διάρκεια του χρόνου , τότε μπορεί να επιλεγθεί Διπλός Κινητός Μέσος ή Εκθετική Εξομάλυνση με προσαρμογή στη τάση (Holt) ή τα μοντέλα ARIMA.
- iii. Όταν η χρονοσειρά εμφανίζει εποχικότητα, σε πρώτη φάση είτε με το αθροιστικό μοντέλο είτε με το πολλαπλασιαστικό γίνεται διάσπαση της χρονοσειράς και στη συνέχεια εκτιμάτε ο δείκτης εποχικότητας των δεδομένων. Αυτό μπορεί να γίνει με την απλή μέθοδο διάσπασης (decomposition) ή με την εκθετική εξομάλυνση (Winter) ή τα μοντέλα ARIMA.
- iv. Όταν η χρονοσειρά εμφανίζει κυκλικότητα, είναι πιο δύσκολη η διαδικασία πρόβλεψης καθώς τα μαθηματικά μοντέλα δυσχεραίνουν όταν υπάρχει ανομοιογένεια στην περίοδο. Ωστόσο, προτείνεται η κλασική διάσπαση (decomposition) ή τα μοντέλα ARIMA.

Επισημαίνεται, σε κάθε περίπτωση η παραπάνω μεθοδολογία δεν είναι πανάκεια και είναι σοφότερο να δοκιμάζονται παραπάνω από μία μέθοδος πρόβλεψης για μία χρονοσειρά και στο τέλος να αξιολογείται από τα διάφορα μέτρα αξιολόγησης.

### 3.8 Μέτρα Αξιολόγησης Προβλέψεων

Προκειμένου να γίνει η πρόβλεψη μελλοντικών τιμών της χρονοσειράς, χρησιμοποιούνται οι παρατηρήσεις της σειράς μέχρι την παρούσα χρονική στιγμή. Εάν αναφερόμαστε σε μία σειρά  $X_n$ , τότε η τελευταία παρατήρηση είναι  $x_n$  και η μελέτη γίνεται για την παρατήρηση,  $k$  χρονικές στιγμές μετά από την  $n$  στιγμή και συμβολίζεται με  $x_n(k)$ . Η παρατήρηση τη χρονική στιγμή  $n+k$  συμβολίζεται με  $x_{n+k}$ .

#### Σφάλμα πρόβλεψης (prediction error)

Το σφάλμα ορίζεται ως η διαφορά μεταξύ της πραγματικής τιμής και της πρόβλεψης για μια περίοδο και ο βασικός στόχος του είναι να τείνει στο μηδέν.

$$e_n = x_n - F_n$$

Για να αξιολογηθεί η απόδοση ενός μοντέλου πρόβλεψης σε μια χρονοσειρά για  $k$  χρονικές στιγμές μπροστά, υπολογίζεται κάποιο μέτρο που συγκεντρώνει τα σφάλματα πρόβλεψης για έναν ικανοποιητικό αριθμό χρονικών στιγμών. Γίνονται λοιπόν προβλέψεις για  $k$  χρονικά βήματα μπροστά, σε έναν αριθμό γνωστών παρατηρήσεων.

Δηλαδή, το σφάλμα για μία χρονική στιγμή  $n$  είναι  $e_n$ , με  $x_n$  να είναι η πραγματική τιμή της χρονοσειράς για την χρονική στιγμή  $n$  και  $F_n$  να 'ναι η προβλεπόμενη τιμή για αυτή την χρονική στιγμή.

Υπάρχουν διάφορα μέτρα, που μπορούν να υπολογίσουν το συνολικό σφάλμα πρόβλεψης για μία μέθοδο πρόβλεψης.

- Μέσο Σφάλμα (Mean Error (ME))

$$ME = \frac{1}{n} \sum_{t=1}^n e_n$$

- Μέση Απόλυτη Απόκλιση (Mean Absolute Deviation (MAD) )

Το μέτρο αυτό είναι χρήσιμο, κυρίως όταν θέλουμε να μετρήσουμε το σφάλμα της πρόβλεψης στις ίδιες μονάδες με τις τιμές της χρονοσειράς.

$$MAD = \frac{1}{n} \sum_{t=1}^n |e_n|$$

- Μέσο Τετραγωνικό Σφάλμα (*Mean Squared Error (MSE)*)

Αυτό το μέτρο επηρεάζεται περισσότερο από τα μεγάλα σφάλματα αφού τα τετραγωνίζει. Κατά συνέπεια μια μέθοδος πρόβλεψης με μέτρια σφάλματα μπορεί να αξιολογηθεί σαν καλύτερη έναντι μιας άλλης με μικρότερα σφάλματα αλλά και περιστασιακά με κάποια μεγάλα. Επιπλέον, το MSE βοηθά στην εύρεση των καταλληλότερων παραμέτρων στις διάφορες μεθόδους εξομάλυνσης. Επιλέγονται οι παράμετροι που ελαχιστοποιούν το MSE.

$$MSE = \frac{1}{n} \sum_{t=1}^n (e_n)^2$$

- Τετραγωνική Ρίζα Μέσου Τετραγωνικού Σφάλματος (*Root Mean Squared Error (RMSE)*)

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

- Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (*Mean Absolute Percentage Error (MAPE)*)

Το μέτρο αυτό είναι χρήσιμο όταν το μέγεθος της μεταβλητής πρόβλεψης έχει σημασία για την αξιολόγηση της μεθόδου πρόβλεψης.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|e_n|}{x_{n+k}}$$

- Μέσο Ποσοστιαίο Σφάλμα (*Mean Percentage Error (MPE)*)

Μερικές φορές πρέπει να εξεταστεί αν μια μέθοδος πρόβλεψης είναι μεροληπτική σε μικρές ή μεγάλες τιμές. Όταν η μέθοδος πρόβλεψης είναι αμερόληπτη, τότε η τιμή του MPE είναι κοντά στο 0. Όταν η τιμή του MPE είναι αρνητική, τότε η μέθοδος

πρόβλεψης δίνει μεροληπτικά μεγάλες τιμές και αντίστοιχα όταν είναι θετική δίνει μεροληπτικά μικρές τιμές.

$$MPE = \frac{1}{n} \sum_{t=1}^n \frac{e_t}{x_{n+k}}$$

Επισημαίνεται ότι αν αυτά τα μέτρα διασποράς είναι μικρά, η προβλεπόμενη τιμή βρίσκεται κοντά στην πραγματική.

## ΚΕΦΑΛΑΙΟ 4

### Τεχνητά Νευρωνικά Δίκτυα

#### 4.1 Ορισμός

Ορισμένοι αλγόριθμοι Μηχανικής Μάθησης μιμούνται τον τρόπο λειτουργίας των βιολογικών συστημάτων μάθησης που χρησιμοποιούνται από τον ανθρώπινο εγκέφαλο και αποτελούνται από ένα πολύπλοκο δίκτυο υπερσυνδεδεμένων νευρώνων. Οι αλγόριθμοι αυτοί ονομάζονται Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks). [16]

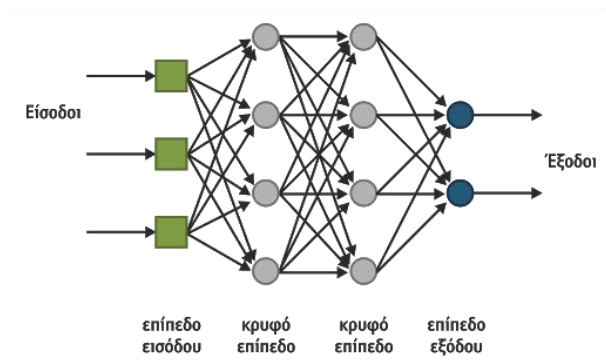
Τα τεχνητά νευρωνικά δίκτυα αποτελούν μια προσπάθεια προσέγγισης της λειτουργίας του ανθρώπινου εγκεφάλου από μια μηχανή και έχουν την ικανότητα να εκτελούν υπολογισμούς με μαζικό παράλληλο τρόπο [Hopfield,1985].

#### 4.2 Αρχιτεκτονική Νευρωνικών Δικτύων

Όλα τα τυπικά Τεχνητά Νευρωνικά Δίκτυα, δημιουργούνται συνδυάζοντας πολλούς νευρώνες μαζί με διασυνδεδεμένες συναπτικές συνδέσεις και αποτελούνται από τρία βασικά επίπεδα, το επίπεδο εισόδου, το υπολογιστικό ή κρυμμένο επίπεδο (hidden layer) και το επίπεδο εξόδου. Αντίστοιχα ονομάζονται και οι νευρώνες που δομούν το κάθε στρώμα, δηλαδή νευρώνες εισόδου, υπολογιστικοί ή κρυμμένοι νευρώνες και νευρώνες εξόδου. Ο όρος κρυμμένο, αναφέρεται στο ότι το στρώμα αυτό δεν είναι



άμεσα ορατό από τα επίπεδα εισόδου και εξόδου. Υπάρχουν νευρωνικά δίκτυα με περισσότερα από ένα κρυμμένο στρώμα και πολύ μεγάλο αριθμό νευρώνων. Η δημιουργία των Τεχνητών Νευρωνικών Δικτύων που παρουσιάζουν πολύ μεγάλη πολυπλοκότητα ως προς τη δομή τους συνθέτουν μία υποκατηγορία της Μηχανικής Μάθησης, την οποία ονομάζουμε **Βαθιά Μάθηση** (Deep Learning).



Δείγμα Νευρωνικού Δικτύου, με δύο κρυμμένα επίπεδα

Ένα τέτοιο δίκτυο, που παρουσιάζεται στο σχήμα, ονομάζεται πλήρως συνδεδεμένο, διότι κάθε κόμβος ενός επιπέδου συνδέεται με κάθε άλλο κόμβο του επόμενου επιπέδου. Οι συνδέσεις ανάμεσα στους κόμβους διαφέρουν ως προς την σημαντικότητά τους που προσδιορίζεται από τα συναπτικά βάρη.

Ανάμεσα στα τρία επίπεδα, πραγματοποιείται μία επεξεργασία των σημάτων, όπου τα σήματα εξόδου από ένα επίπεδο, χρησιμοποιούνται ως σήματα εισόδου για το επόμενο επίπεδο. Ειδικότερα, κάθε νευρώνας έχει πολλές εισόδους αλλά μόνο μία έξοδο. Οι νευρώνες εισόδου δεν πραγματοποιούν κανένα υπολογισμό, απλώς μεσολαβούν ανάμεσα στις περιβαλλοντικές εισόδους του δικτύου και τους υπολογιστικούς νευρώνες. Οι υπολογιστικοί νευρώνες πολλαπλασιάζουν κάθε είσοδό τους με το αντίστοιχο συναπτικό βάρος και υπολογίζουν το ολικό άθροισμα των γινομένων. Οι νευρώνες εξόδου εξάγουν τις τελικές μετρήσεις του δικτύου στο περιβάλλον.

Επιπρόσθετα, στην λειτουργία των νευρώνων συμβάλλουν η συνάρτηση εισαγωγής και ενεργοποίησης, δύο συναρτήσεις που συνήθως είναι ίδιες σε κάθε επίπεδο. Η συνάρτηση ενεργοποίησης, δέχεται ως όρισμα το ολικό άθροισμα των γινομένων και υλοποιείται εσωτερικά σε κάθε κόμβο. Η τιμή που παίρνει τελικά η συνάρτηση για το εν λόγω όρισμα αποτελεί και την έξοδο του νευρώνα για τις τρέχουσες εισόδους και βάρη.

Ειδικότερα, η Συνάρτηση Εισόδου έχει την παρακάτω μορφή:

$$u(x) = \sum_{i=1}^d w_{ij}x_j + \mu_i$$

Όπου,  $w_{ij}$  συμβολίζουμε τα βάρη που συνδέουν τον νευρώνα  $j$  με το νευρώνα  $i$ , με  $x_j$  το αποτέλεσμα εξόδου της μονάδας  $j$  και με  $\mu_i$  το κατώφλι για το νευρώνα  $i$ . Με τον όρο κατώφλι (threshold) ονομάζουμε την τιμή εισόδου που λειτουργεί σαν σημείο αναφοράς σε ένα κόμβο, απουσία άλλων τιμών εισόδου. Το κατώφλι έχει τιμή εισόδου την μονάδα. Εάν το συνολικό άθροισμα από τις υπόλοιπες εισόδους του νευρώνα είναι μεγαλύτερο από την τιμή αυτή, τότε ο νευρώνας ενεργοποιείται. Εάν είναι μικρότερο, τότε ο νευρώνας παραμένει ανενεργός. Στη συνέχεια κάθε μονάδα του Νευρωνικού δικτύου παίρνει τις τιμές από την μεταβλητή  $u(x)$  και εφαρμόζει μία συνάρτηση ενεργοποίησης πάνω σε αυτές. Το αποτέλεσμα της τιμής ενεργοποίησης, υπολογίζεται από τον τύπο  $g(u)$ . Όπου  $g$  η συνάρτηση ενεργοποίησης.

Ο ρόλος των συναρτήσεων ενεργοποίησης είναι πολύ σημαντικός γιατί ελέγχουν τα αποτελέσματα εξόδου να μην έχουν ακραίες τιμές, κάτι που θα οδηγούσε το δίκτυο να αδρανοποιηθεί διακόπτοντας την διαδικασία εκπαίδευσης τους. Συχνότερα χρησιμοποιούνται οι σιγμοειδής συναρτήσεις ενεργοποίησης (όπως sigmoid, tanh), γιατί είναι μη γραμμικές και συνεχώς διαφοροποιήσιμες, χαρακτηριστικά που τις αναγάγουν σε ιδανικές επιλογές για χρήση σε τεχνητά νευρωνικά δίκτυα και η συνάρτηση ReLu. Ένα από τα μεγαλύτερα πλεονεκτήματα που έχει η συνάρτηση ενεργοποίησης ReLu έναντι των υπολοίπων είναι ότι δεν ενεργοποιεί ταυτόχρονα όλους τους νευρώνες. Μετατρέπει όλες τις αρνητικές εισροές στο μηδέν και ο νευρώνας δεν ενεργοποιείται. Αυτό το καθιστά πολύ αποτελεσματικό υπολογιστικά καθώς λίγοι νευρώνες ενεργοποιούνται κάθε φορά. [17]

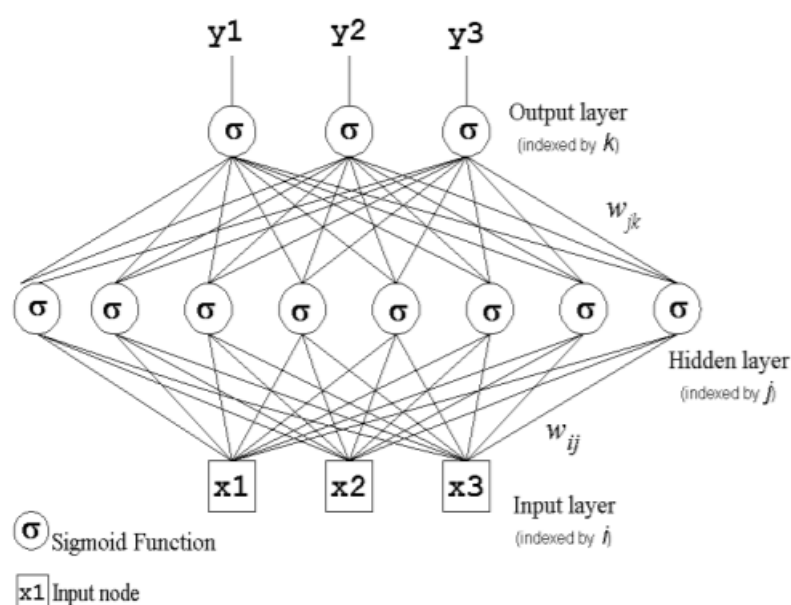
Είναι φανερό ότι το κάθε Τεχνητό Νευρωνικό Δίκτυο, μπορεί να ακολουθεί διαφορετική αρχιτεκτονική, ωστόσο κάθε αρχιτεκτονική έχει ως δομικά στοιχεία όσο αναφέρθηκαν παραπάνω. Η επιλογή του καταλληλότερου Τεχνητού Νευρωνικού Δικτύου, εξαρτάται από την φύση του προβλήματος και από τα αποτελέσματα των μετρικών.

## 4.3 Βασικά Νευρωνικά Δίκτυα

### 4.3.1 Πολυεπίπεδος Perceptron (Multilayer Perceptron - MLP)

Η απλούστερη περίπτωση ΤΝΔ που αποτελούν και την ευρύτερη χρησιμοποιούμενη κατηγορία ΤΝΔ είναι τα δίκτυα πρόσθιας τροφοδότησης (**feedforward networks**). Το χαρακτηριστικό των δικτύων αυτών είναι ότι δεν υπάρχει ανατροφοδότηση της εξόδου ενός νευρώνα προς τους νευρώνες από τους οποίους επηρεάζεται άμεσα ή έμμεσα. Τα δίκτυα αυτά ονομάζονται πρόσθιας τροφοδότησης γιατί, με δεδομένη κάποια είσοδο  $x$ , για τον υπολογισμό της εξόδου πραγματοποιούνται υπολογισμοί, όλοι προς την ίδια κατεύθυνση: από την είσοδο προς την έξοδο. [18]

Από τους πιο διαδεδομένους αλγορίθμους αυτού του είδους είναι ο πολυεπίπεδος Perceptron, όπου οι νευρώνες είναι οργανωμένοι σε επίπεδα (layers). Το χαρακτηριστικό των επιπέδων είναι ότι δεν υπάρχει διασύνδεση μεταξύ των νευρώνων του ίδιου επιπέδου, αλλά οι νευρώνες οποιουδήποτε επιπέδου τροφοδοτούν αποκλειστικά τους νευρώνες του επόμενου επιπέδου και τροφοδοτούνται από το προηγούμενο επίπεδο. Ο αλγόριθμος αποτελείται από το επίπεδο εισόδου, εξόδου και ένα ή περισσότερα ενδιάμεσα κρυμμένα επίπεδα (hidden layer). Επιπλέον, οι νευρώνες στο επίπεδο εισόδου δεν πραγματοποιούν υπολογισμούς, απλώς μεταδίδουν την τιμή του διανύσματος εισόδου στο επόμενο επίπεδο. Συνήθως οι συνδέσεις ξεκινούν από έναν νευρώνα και καταλήγουν σε νευρώνες στο επόμενο επίπεδο. Παρακάτω φαίνεται η αρχιτεκτονική ενός MLP.



#### Αρχιτεκτονική MLP [19]

Ένας MLP εκπαιδεύεται κυρίως με τον αλγόριθμο Back Propagation και χρησιμοποιεί μη γραμμικές συναρτήσεις ενεργοποίησης. Επικρατέστερες είναι οι σιγμοειδείς συναρτήσεις.

Σχεδιασμός και εκπαίδευση μοντέλου: [19]

1. Επιλογή αριθμού κρυμμένων επιπέδων
2. Επιλογή νευρώνων σε κάθε κρυμμένο επίπεδο
3. Επιλογή βέλτιστης αλγοριθμικής λύσης για ελαχιστοποίηση σφάλματος
4. Εφαρμογή της λύσης αυτής σε εύλογο χρονικό διάστημα
5. Έλεγχος του Νευρωνικού Δικτύου για αποφυγή υπερπροσαρμογής (overfitting), δηλαδή τα δεδομένα να μην επαρκούν για την εκπαίδευση του μοντέλου.

Ύστερα από το πέρας των παραπάνω βημάτων, στόχος είναι η μείωση του σφάλματος πρόβλεψης με την κατάλληλη επιλογή βαρών και ορίων. Σε αυτό βοηθούν οι αλγόριθμοι εκπαίδευσης που επιλέγονται για να εκπαιδεύσουν το δίκτυο σύμφωνα με την ιστορικότητα που έχουν συλλέξει. Συγκεκριμένα συγκρίνουν τα αποτελέσματα της κάθε στιγμής με συγκεκριμένες τιμές εξόδου. Όλες οι αποκλίσεις συλλέγονται σε μία συνάρτηση υπολογισμού σφάλματος του δικτύου. Αξίζει να σημειωθεί ότι ο MLP με μόνο ένα κρυφό επίπεδο μπορεί να παρουσιάσει πολύ καλά αποτελέσματα πρόβλεψης. Σπάνια είναι η επιλογή δύο κρυφών επιπέδων και όταν γίνεται είναι σε δεδομένα με ασυνέχεια. Τέλος, ο αριθμός των νευρώνων στα κρυμμένα επίπεδα, έχει καθοριστικό ρόλο στα αποτελέσματα του αλγορίθμου. Σε κάθε περίπτωση πρέπει να αποφεύγονται ο πολύ μικρός αριθμός νευρώνων, διότι υπάρχει κίνδυνος υποπροσαρμογής (underfitting) του μοντέλου, όπως και ο πολύ μεγάλος αριθμός νευρώνων γιατί μπορεί να οδηγηθεί σε υπερπροσαρμογή το δίκτυο (overfitting).

#### **4.3.2 Επαναλαμβανόμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks-RNN)**

Τα RNN είναι από τα πολύ αποδοτικά δίκτυα, όταν πρόκειται να πραγματοποιηθούν προβλέψεις σε ακολουθιακά δεδομένα, όπως χρονοσειρές, ήχος, καιρικά φαινόμενα,

οικονομικά δεδομένα, αναγνώριση φωνής, αναγνώριση χειρόγραφου κ.ά. όπως θα αναλυθεί παρακάτω.

Χαρακτηρίζονται ισχυρά δίκτυα στο χώρο των Νευρωνικών Δικτύων, καθώς αποτελούνται από κρυμμένους κόμβους όπου μπορούν να αποθηκεύουν πολλές παρελθοντικές πληροφορίες. Τα περισσότερα Νευρωνικά Δίκτυα θεωρούν ότι όλες οι τιμές εισόδου και εξόδου είναι ανεξάρτητες μεταξύ τους, γεγονός που δυσκολεύει τις περισσότερες φορές τις προβλέψεις. Αντιθέτως τα RNN λόγω της εσωτερικής μνήμης τους θυμούνται πολλές πληροφορίες που εισήχθησαν στο δίκτυο με αποτέλεσμα να δίνουν ακριβέστερες προβλέψεις.

### Λειτουργία

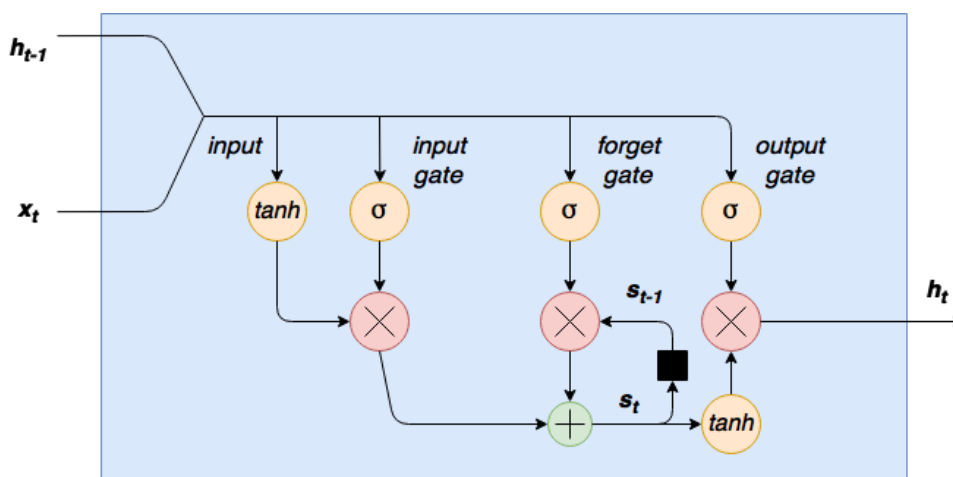
Στα κλασσικά Feed – Forward Νευρωνικά Δίκτυα η πληροφορία που εισέρχεται δεν περνάει δεύτερη φορά από έναν κόμβο και πηγαίνει προς μία μόνο κατεύθυνση, σε αντίθεση με τα RNN όπου η πληροφορία περιστρέφεται γύρω από έναν βρόγχο κάνοντας κύκλο στο δίκτυο, με αποτέλεσμα να λαμβάνει υπόψιν μαζί με την τρέχουσα τιμή και τις προηγούμενες. Συνεπώς, τα RNN έχουν δύο εισόδους, μία για την νέα πληροφορία και μία για την παρελθοντική, όπου συνδυάζονται προκειμένου να διαχειριστούν την νέα πληροφορία που θα εισαχθεί. Θα έλεγε κανείς ότι αυτός ο τρόπος λειτουργίας συνάδει αρκετά με του ανθρώπινου εγκεφάλου. Ουσιαστικά, με την βοήθεια της κρυμμένης εσωτερικής μνήμης, εντοπίζονται εξαρτήσεις μεταξύ πληροφοριών για διαδοχικές χρονικές στιγμές που ονομάζονται Μακροχρόνιες Εξαρτήσεις. Επιπλέον, σε κάθε κόμβο του δικτύου υπάρχουν τα αντίστοιχα βάρη που επηρεάζουν τα τελικά αποτελέσματα. Ακολουθεί η μαθηματική αναπαράσταση της παραπάνω διαδικασίας:

$$h_t = \Phi(W \cdot x_t + U \cdot h_{t-1})$$

Όπου  $h_t$  η κρυμμένη κατάσταση την χρονική στιγμή  $t$ ,  $W$  το βάρος μεταξύ των εισροών και του κρυμμένου επιπέδου,  $x_t$  η τιμή εισόδου την χρονική στιγμή  $t$ ,  $U$  ο συντελεστής μετάβασης μεταξύ του κρυμμένου επιπέδου και του εαυτού και  $h_{t-1}$  η κρυμμένη κατάσταση την χρονική στιγμή  $t-1$ . Το σφάλμα που παράγεται θα επιστρέψει μέσω του Backpropagation και το Νευρωνικό θα προσαρμόσει κατάλληλα τα βάρη του. [20]

### 4.3.3 Μακράς Βραχυχρόνιας Μνήμης Νευρωνικά Δίκτυα (Long Short-Term Memory -LSMT)

Τα Νευρωνικά Δίκτυα Long-Short Term Memory (LSTM) εμφανίστηκαν το 1997, ανήκουν στην οικογένεια RNN και μοντελοποιήθηκαν με σκοπό να διαχειριστούν τις Μακροχρόνιες Εξαρτήσεις, παρουσιάζοντας υψηλές αποδόσεις ακρίβειας. [20] Η αρχιτεκτονική τους αποτελείται από ένα σύνολο επαναλαμβανόμενων (recurrent) συνδεδεμένων υπό-δικτύων, γνωστά ως μπλοκ μνήμης. Κάθε τέτοιο μπλοκ περιέχει ένα ή περισσότερα αυτό-συνδεδεμένα κελιά μνήμης (memory cells) και τρεις πύλες, την πύλη εισόδου, την πύλη εξόδου και την forget πύλη. Μία πύλη εισόδου, είναι ένα στρώμα από σιγμοειδείς κόμβους, με τιμές εξόδου 0 ή 1, οι οποίοι μπορούν να διώξουν όποια στοιχεία του διανύσματος εισόδου θεωρούν αχρείαστα. Η πύλη forget, ελέγχει ποια πληροφορία από την προηγούμενη κατάσταση θα περάσει στην τωρινή ή θα ξεχαστεί. Οι δύο αυτές πύλες μαζί, επιτρέπουν στο δίκτυο να ελέγχει ποια πληροφορία θα αποθηκεύεται και ποια θα αντικατασταθεί σε κάθε βήμα. Τέλος η πύλη εξόδου, είναι αυτή που ελέγχει ποιες τιμές θα επιτραπούν ως εξοδοι από το μπλοκ.



Απεικόνιση κελιού μνήμης [21]

Αναλυτικότερα, η είσοδος  $x_t$  συνδέεται με την προηγούμενη είσοδο  $h_{t-1}$  και η συνδυασμένη τιμή περνάει σ' ένα στρώμα συνάρτησης tanh, στο πρώτο βήμα του μπλοκ. Στο δεύτερο βήμα περνάει από την πύλη εισόδου. Η πύλη εισόδου αποτελείται από σιγμοειδή συνάρτηση ενεργοποίηση και τα αποτελέσματα των κόμβων που ενεργοποιούνται πολλαπλασιάζονται με την τιμή που παράχθηκε από το πρώτο

επίπεδο. Στο τρίτο βήμα σ' ένα κελί μνήμης βρίσκεται ο βρόγχος της πύλης forget. . Στο μπλοκ υπάρχει η εσωτερική μεταβλητή  $st$ , αυτή η μεταβλητή με μία χρονική υστέρηση  $st-1$ , προστίθεται στα δεδομένα και δημιουργείται ένα στρώμα επανάληψης, το οποίο ελέγχεται από την πύλη forget. Τέλος, υπάρχει το στρώμα εξόδου με την συνάρτηση  $\tanh$ , που ελέγχεται από την πύλη εξόδου. Η πύλη εξόδου είναι υπεύθυνη ποιες τιμές επιτρέπονται και ποιες απορρίπτονται ως εξόδοι από το μπλοκ.

Μαθηματικά εκφράζεται με την παρακάτω αλληλουχία: [22]

Όπου  $x_t$  οι τιμές εισόδου και  $y_t$  οι τιμές εξόδου, τις χρονικές στιγμές από  $t= 1$  έως  $T$ .  $U, V$  είναι τα βάρη της τιμής εισόδου και της προηγούμενης τιμής εξόδου αντίστοιχα, ενώ  $b$  η πόλωση της τιμής εισόδου.

$$g = \tanh(b^g + x_t U^g + h_{t-1} V^g)$$

$$i = \sigma(b^i + x_t U^i + h_{t-1} V^i)$$

$$g \circ i$$

$$f = \sigma(b^f + x_t U^f + h_{t-1} V^f)$$

$$s_t = s_{t-1} \circ f + g \circ i$$

$$o = \sigma(b^o + x_t U^o + h_{t-1} V^o)$$

$$h_t = \tanh(s_t) \circ o$$

#### 4.4.4 Φραγμένα Επαναλαμβανόμενα Νευρωνικά Δίκτυα (Gated Recurrent Units -GRUs)

Η αρχιτεκτονική δικτύου GRU παρουσιάστηκε από το Cho Kyunghuyn το 2014. Ανήκει στην οικογένειά RNN δικτύων και πιο συγκεκριμένα μοιάζει με τον LSTM αλγόριθμο σε απλούστερη μορφή. Το δίκτυο συνδυάζει τις πύλες εισόδου και forget σε μία πύλη, που ονομάζεται πύλη ενημέρωσης (update gate), παράλληλα η πύλη εξόδου ονομάζεται πύλη επαναφοράς (reset gate). Το μοντέλο είναι αρκετά πιο απλό από ένα LSTM ωστόσο έχει και αυτό σκοπό να αναβαθμίζει ή να επαναφέρει το περιεχόμενο της μνήμης. Το δίκτυο GRU εκθέτει πλήρως το περιεχόμενο της μνήμης του και με την μέθοδο ολοκλήρωση διαρροής (Leaky Integration), ισορροπεί μεταξύ του προηγούμενου περιεχομένου μνήμης και του νέου περιεχομένου μνήμης.

Η μαθηματική αλληλουχία που περιγράφει τα βάρη σ' ένα GRU δίκτυο είναι:

$$h_i^{(t)} = u_i^{(t-1)}h_i^{(t-1)} + (1 - u_i^{(t-1)})\sigma\left(b_i + \sum_j U_{i,j}x_j^{(t)} + \sum_j W_{i,j}r_j^{(t-1)}h_j^{(t-1)}\right)$$

Όπου  $u$  είναι η πύλη ανανέωσης και  $r$  η πύλη επαναφοράς και δίνονται από τις σχέσεις:

$$u_i^{(t)} = \sigma\left(b_i^u + \sum_j U_{i,j}^u x_j^{(t)} + \sum_j W_{i,j}^u h_j^{(t)}\right)$$
$$r_i^{(t)} = \sigma\left(b_i^r + \sum_j U_{i,j}^r x_j^{(t)} + \sum_j W_{i,j}^r h_j^{(t)}\right)$$

Οι πύλες ενημέρωσης μπορούν να περιορίσουν γραμμικά οποιαδήποτε κατάσταση και η πύλη επαναφοράς ρυθμίζει πόση από την προηγούμενη μνήμη θα χαθεί και πόση από την νέα θα συνυπολογιστεί στο δίκτυο.

## ΚΕΦΑΛΑΙΟ 5

### Πειραματικός Έλεγχος

#### 5.1 Πρόβλημα – Σκοπός

Στο κεφάλαιο αυτό θα περιγράψουμε και θα παρουσιάσουμε την ανάλυση που εφαρμόσαμε, με σκοπό την πρόβλεψη των μελλοντικών τιμών των μετοχών του χρηματιστηρίου. Στόχος είναι να πραγματοποιηθεί διερευνητική ανάλυση για τις χρονοσειρές και να εξετάσουμε τον τρόπο που λειτουργούν και τα αποτελέσματα που παράγουν βασικά Νευρωνικά Δίκτυα για τις χρονοσειρές των Amazon, Google και Apple.



## 5.2 Εργαλεία

Στα πλαίσια της εργασίας χρησιμοποιήθηκε η γλώσσα στατιστικού προγραμματισμού R, στην πλατφόρμα R Studio version 3.5.2 και η γλώσσα προγραμματισμού Python στην πλατφόρμα Jupyter.

## 5.3 Διερευνητική Ανάλυση (Exploratory Analysis)

### 5.3.1 Σύνολο Δεδομένων (dataset)

Το σύνολο δεδομένων που χρησιμοποιήθηκε ονομάζεται **Big Five Stock** και το προμηθευτήκαμε από την ιστοσελίδα Kaggle, η οποία παρέχει πληθώρα συνόλων δεδομένων ιδανικά για ανάλυση. [23]

Το σύνολο δεδομένων περιέχει δεδομένα σε μορφή χρονοσειρών για έξι μεγάλες εταιρείες που συμμετέχουν στον αμερικάνικο χρηματιστηριακό δείκτη NASDAQ. Οι εταιρείες είναι: Amazon, Google, Apple, Microsoft, Facebook, NASDAQ (IXIC). Αποτελείται από 41.660 εγγραφές και 7 στήλες χαρακτηριστικών. Οι εγγραφές αντιστοιχούν στο χρονικό διάστημα 1971-02-05 έως 2019-08-23. Οι επτά στήλες αντιστοιχούν στα παρακάτω χαρακτηριστικά:

*Date*: ημερομηνία

*Name*: τα ονόματα των εταιρειών

*Open*: η τιμή της μετοχής κατά το άνοιγμα του δείκτη για κάθε εταιρεία

*Close*: η τιμή της μετοχής κατά το κλείσιμο του δείκτη για κάθε εταιρεία

*High*: η υψηλότερη τιμή της μετοχής την συγκεκριμένη ημέρα για την κάθε εταιρεία

*Low*: η χαμηλότερη τιμή της μετοχής την συγκεκριμένη ημέρα για την κάθε εταιρεία

*Volume*: ο αριθμός των μετοχών που διακινήθηκαν ανά εταιρεία

### 5.3.2 Επεξεργασία Δεδομένων

Η επεξεργασία και ανάλυση του συνόλου δεδομένων υλοποιήθηκε στο περιβάλλον Jupyter με χρήση της γλώσσας Python. Το αρχείο είναι σε μορφή csv μεγέθους 1,97 MB και εισήχθη με την μέθοδο `read_csv`. Καθόλη την μελέτη και επεξεργασία τα δεδομένα ήταν σε μορφή Pandas Dataframe της βιβλιοθήκη Pandas. Έγινε έλεγχος για χαμένες τιμές (missing values) και δεν βρέθηκε καμία χαμένη τιμή. Αυτό βοηθάει στο ξεκίνημα της ανάλυσης καθώς αν υπήρχαν χαμένες τιμές θα έπρεπε να αντικατασταθούν ή να απαλειφθούν, προκειμένου να μην επηρεαστούν οι δοκιμές και να έχουμε αξιόπιστα αποτελέσματα.

### 5.3.3 Οπτικοποίηση Δεδομένων (Data Visualization)

Η οπτικοποίηση του συνόλου δεδομένου μας βοηθάει να κατανοήσουμε καλύτερα την εξελικτική πορεία των τιμών του χρηματιστηρίου και να εντοπίσουμε πιθανά μοτίβα. Για το υπόλοιπο του πειράματος επιλέχθηκαν να μελετηθούν εκτενέστερα από τις έξι εταιρείες οι τρεις: Amazon, Apple, Google.

Για την αναπαράσταση των τιμών χρησιμοποιήθηκαν, στην γλώσσα R η βιβλιοθήκη `quantmod` στο περιβάλλον Colaboratory [23] και στην Python η `Matplotlib`.

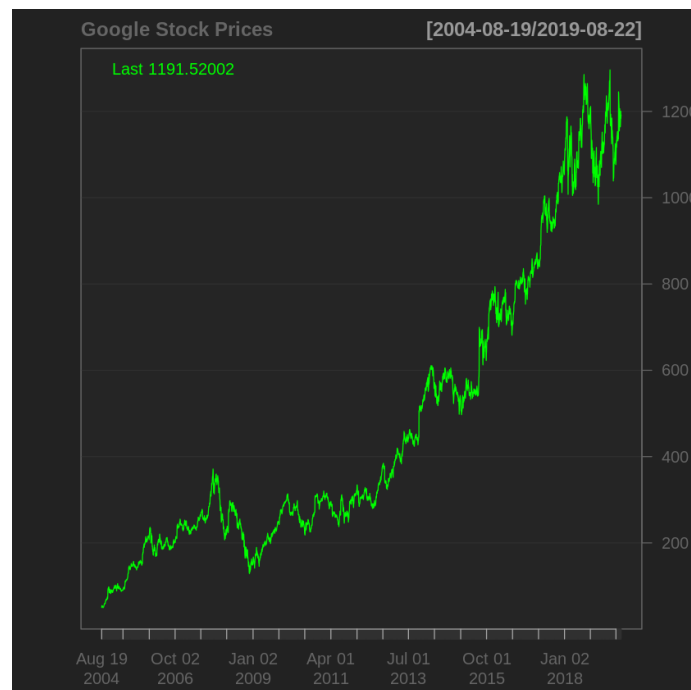
**Διαγραμματική Απεικόνιση των τιμών κλεισίματος για το σύνολο των ετών**  
[24]

## Amazon



Η διαγραμματική απεικόνιση της τιμής κλεισίματος της μετοχής για την Amazon συναρτήσκει του χρόνου

## Google



Η διαγραμματική απεικόνιση της τιμής κλεισίματος της μετοχής για την Google συναρτήσκει του χρόνου

## Apple



Η διαγραμματική απεικόνιση της τιμής κλεισίματος της μετοχής την Apple συναρτήσει του χρόνου

### 5.3.4 Ποιοτικά Χαρακτηριστικά

Έχοντας αποτυπώσει κάθε μία από τις τρεις χρονοσειρές, θα μελετήσουμε τα ποιοτικά χαρακτηριστικά τους. Η **συχνότητα** και για τις τρεις εταιρείες είναι η μία ημέρα, καθώς αναφέρεται στις καθημερινές τιμές του δείκτη NASDAQ. Η διαφορά ανάμεσα στις χρονοσειρές είναι ότι η Amazon ξεκινάει το 1997, η Google το 2006 και η Apple το 1980. Η **τάση** και για τις τρεις είναι κυρίως αυξητική με εντονότερη αύξηση τα τελευταία χρόνια. Παράλληλα παρατηρούνται μικρές διακυμάνσεις στο ενδιάμεσο. Καμία από τις τρεις δεν εμφανίζει **εποχικότητα**, ωστόσο παρατηρείται τα τελευταία χρόνια 2018 και 2019 να έχουν κοινή πτώση την περίοδο του Ιανουαρίου. Σε καμία χρονοσειρά δεν εντοπίζεται **κυκλικότητα** και **τυχειότητα**.

### 5.3.5 Augmented Dickey – Fuller Test

Εφόσον υπάρχει τάση στις χρονοσειρές και οι τιμές κλεισίματος δεν κυμαίνονται γύρω από έναν σταθερό μέσο όρο χαρακτηρίζονται ως **μη – στάσιμες** χρονοσειρές. Τα παραπάνω επιβεβαιώνεται με την βοήθεια της γλώσσας προγραμματισμού Python και την συνάρτηση `adf` της βιβλιοθήκης `statsmodels.tsa.stattools`, μέσω της οποίας

εφαρμόστηκε **Augmented Dickey-Fuller test**, με αρχική υπόθεση η χρονοσειρές να είναι μη-στάσιμες.

Προέκυψαν τα αποτελέσματα:

<b>Amazon</b>	<b>Google</b>	<b>Apple</b>
ADF Statistic: 1.991 689	ADF Statistic: 0.241 261	ADF Statistic: 2.088 275
p-value: 0.998662	p-value: 0.974483	p-value: 0.998777
Critical Values:	Critical Values:	Critical Values:
1%: -3.432	1%: -3.432	1%: -3.431
5%: -2.862	5%: -2.862	5%: -2.862
10%: -2.567	10%: -2.567	10%: -2.567

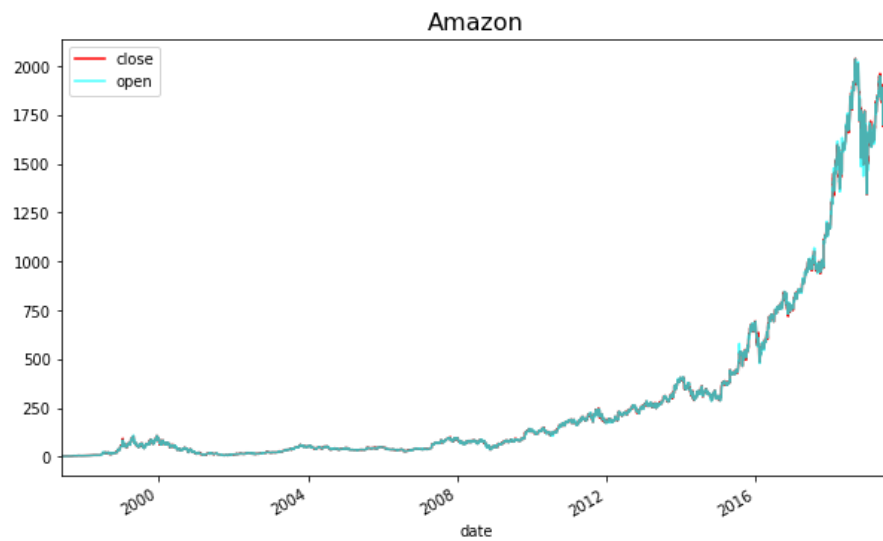
Συμπεραίνουμε, σύμφωνα με τον δείκτη ADF και p-value που είναι πολύ μεγαλύτερο από το 0,05 ότι είναι δεκτή η αρχική υπόθεση, δηλαδή οι χρονοσειρές είναι μη – στάσιμες.

### **5.3.6 Συσχετίσεις μεταξύ των μεταβλητών**

#### **Συγκριτική Απεικόνιση των τιμών ανοίγματος και τιμών κλεισίματος για κάθε εταιρία**

Παρακάτω ακολουθούν συγκριτικά διαγράμματα των τιμών κλεισίματος και ανοίγματος για κάθε μετοχή. Με τον τρόπο αυτό γίνεται καλύτερα κατανοητό αν υπάρχουν έντονες διακυμάνσεις ανάμεσα στην τιμή που ξεκινάει κάθε ημέρα η μετοχή και στην τιμή που κλείνει τελικά.

## Amazon



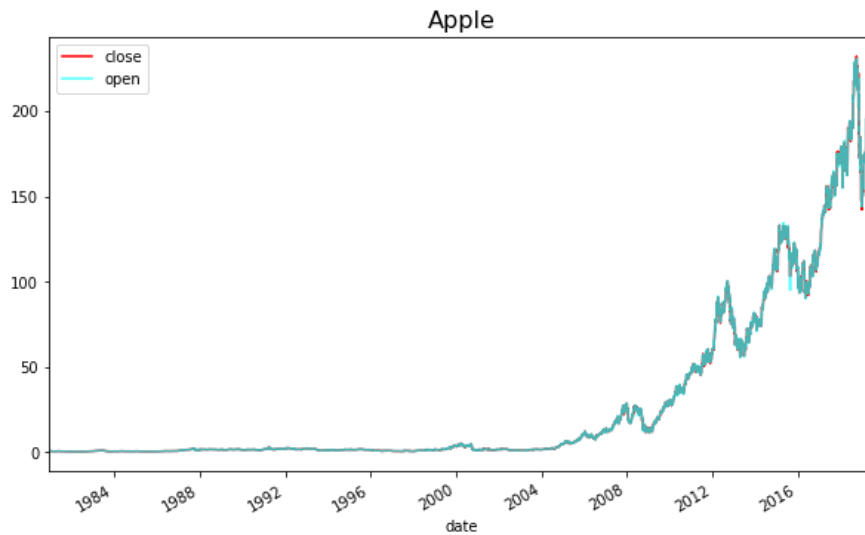
Γραφική απεικόνιση των τιμών ανοίγματος με γαλάζιο και των τιμών κλεισίματος με κόκκινο για την Amazon

## Google



Γραφική απεικόνιση των τιμών ανοίγματος με γαλάζιο και των τιμών κλεισίματος με κόκκινο για την Google

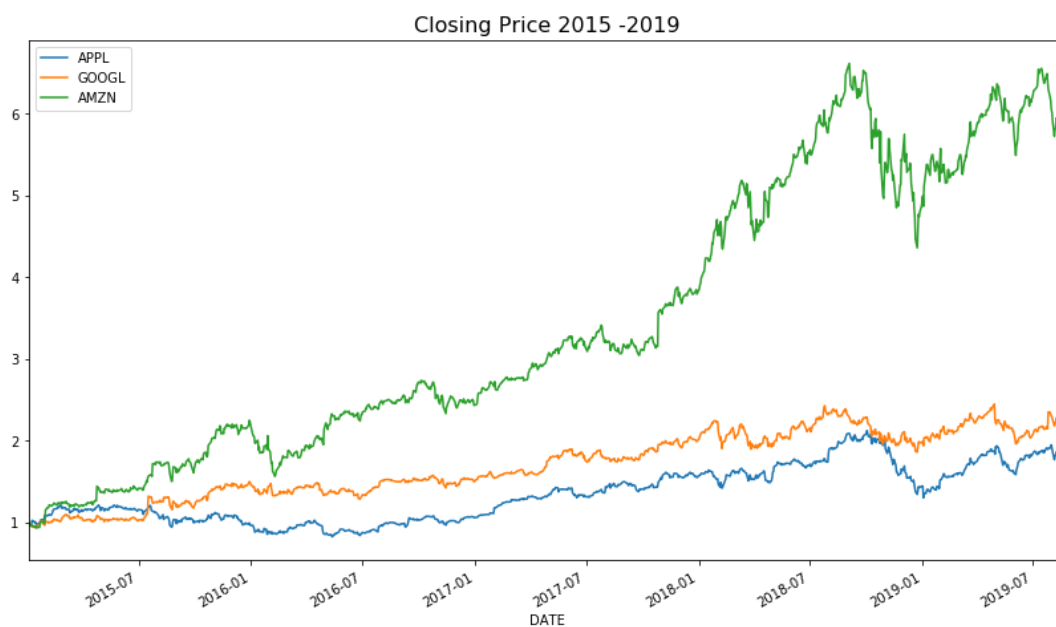
## Apple



Γραφική απεικόνιση των τιμών ανοίγματος με γαλάζιο και των τιμών κλεισίματος με κόκκινο για την Apple

Παρατηρούνται πολύ μικρές διακυμάνσεις ανάμεσα στις τιμές ανοίγματος και κλεισίματος και στις τρεις εταιρείες, ωστόσο εντονότερες είναι της Google.

## Συγκριτική απεικόνιση των τιμών κλεισίματος των τελευταίων 5 ετών για τις τρεις εταιρείες



Η τιμή κλεισίματος των τελευταίων πέντε ετών για την Amazon, Google και Apple

Η Amazon παρουσιάζει με διαφορά μεγαλύτερες τιμές κλεισίματος στο χρηματιστήριο συγκριτικά με τις Apple και Google και δείχνει εντονότερη μεταβολή τιμών. Η Google και η Apple κυμαίνονται γύρω από σταθερές τιμές.

Στην συνέχεια υπολογίστηκε με την βοήθεια της συνάρτησης `corr()`, η συσχέτιση της μεταβλητής `close` συγκριτικά με τα άλλα πεδία της χρονοσειράς.

Amazon		Google		Apple	
open	0.999832	open	0.999758	open	0.999899
close	1.000000	close	1.000000	close	1.000000
high	0.999913	high	0.999878	high	0.999949
low	0.999925	low	0.999902	low	0.999954
volume	-0.191652	volume	-0.47432	volume	0.596716

Η τιμή κλεισίματος (`close`) έχει υψηλή εξάρτηση με την τιμή ανοίγματος, την μέγιστη και την ελάχιστη ημερήσια τιμή της μετοχής.

## 5.4 Μοντελοποίηση Νευρωνικών Δικτύων

Στα πλαίσια την εργασίας θα μελετηθούν και θα αξιολογηθούν τα Νευρωνικά Δίκτυα Multilayer Perceptron (MLP), Long-Short Term Memory Networks (LSTM) και Gated Recurrent Neural Networks (GRUs) με την βοήθεια των τριών χρονοσειρών Amazon, Google, Apple.

Για την μοντελοποίηση των Νευρωνικών χρησιμοποιήθηκε το API Keras και για την αξιολόγηση η μέθοδος Cross Validation και ο στατιστικός δείκτης Mean Square Error, όπου αποτυπώνει το συνολικό σφάλμα του μοντέλου να προβλέψει σωστά τις παρατηρήσεις για το συγκεκριμένο σύνολο εκπαίδευσης (training set).

### 5.4.1 Keras Configuration

Το API Keras είναι ένα API που έχει σχεδιαστεί να παρέχει πολλές και εύχρηστες δυνατότητες στους προγραμματιστές για την κατασκευή και παραμετροποίηση Νευρωνικών Δικτύων. Μπορεί να χρησιμοποιηθεί με τις γλώσσες προγραμματισμού R



και Python και στηρίζεται στην βιβλιοθήκη μηχανικής μάθησης TensorFlow. Μέσω του Keras ο προγραμματιστής έχει πρόσβαση σε αρκετές βιβλιοθήκες που του επιτρέπουν με απλό τρόπο να επεξεργαστεί το κάθε επίπεδο του Νευρωνικού Δικτύου. Οι βασικοί πυλώνες που συνθέτουν ένα Νευρωνικό Δίκτυο και μπορούν να επεξεργαστούν μέσω Keras είναι το είδος του μοντέλου (model), ο μεταγλωττιστής (compiler), η συνάρτηση ενεργοποίησης (activation), η συνάρτηση εκπαίδευσης (optimizer) και οι επιλογές εκπαίδευσης του μοντέλου (fit).

Το Keras υποστηρίζει δύο μοντέλα Νευρωνικών Δικτύων, το Sequential Model και το Functional Model. [25] Το είδος του μεταγλωττιστή του Νευρωνικού Δικτύου μας επιτρέπει να επιλέξουμε τους εκάστοτε κατηγοριοποιητές. Στο Keras διαθέσιμοι είναι οι Stochastic Gradient Descent (Sgd), RMSprop, Adam, Adagrad, Adadelata, Adamax, Ftrl και Nadam. [26] Πολλές είναι οι επιλογές που αφορούν την συνάρτηση ενεργοποίησης. Ορισμένες είναι η exponential linear unit (elu), η relu, η softmax, η scaled exponential linear unit (selu), η tanh, η softplus, η softsign, η sigmoid, η hard\_sigmoid, η exponential και η linear. Τέλος, η επιλογή fit μας επιτρέπει να επιλέξουμε τον αριθμό των epochs και το batch size. Πιο συγκεκριμένα, τα epochs δείχνουν πόσες θα περάσει μέσα από το Νευρωνικό Δίκτυο το σύνολο δεδομένων και το batch size δείχνει πως θα χωριστεί το σύνολο δεδομένων προκειμένου να εισαχθεί σταδιακά στο Νευρωνικό Δίκτυο.

Στο παρόν πείραμα χρησιμοποιούνται το Sequential Model, το οποίο αποτελείται από μία γραμμική στοίβα πολλών επιπέδων. Οι Optimizers Adam, Sgd, RMSprop και η συνάρτηση ενεργοποίησης ReLu. Για την αξιολόγηση της απόδοσης των μοντέλων χρησιμοποιείται ο δείκτης Mean Square Error (MSE).

## **5.5 Μετατροπή δεδομένων για εισαγωγή στα Νευρωνικά Δίκτυα**

Αρχικά, δημιουργήθηκε ένα αρχείο μόνο με τις τιμές κλεισίματος της μετοχής, που θα χρησιμοποιηθεί για την εκπαίδευση του Νευρωνικού Δικτύου και την πρόβλεψη. Ως πρώτο βήμα κανονικοποιήθηκαν τα δεδομένα με την βοήθεια του MinMaxScaler από τη βιβλιοθήκη Sklearn, σε τιμές από 0 έως 1. Αυτό είναι απαραίτητο γιατί τα Νευρωνικά Δίκτυα είναι ευαίσθητα όταν δεν είναι όλα τα δεδομένα στη ίδια κλίμακα, κυρίως όταν χρησιμοποιείται η σιγμοειδής και η tanh ως συναρτήσεις ενεργοποίησης. [27]

Στη συνέχεια δημιουργήθηκε η συνάρτηση `Create_Dataset`, ώστε τα δεδομένα να είναι σε κατάλληλη μορφή για εισαγωγή στο δίκτυο, με την παρακάτω ακολουθία βημάτων:

#### *Δημιουργία Συνάρτησης `Create_Dataset`*

1. *Είσοδος: Όρισε το `Dataset` και το `Look_Back`*
2. *Δημιούργησε δύο κενούς πίνακες `X`, `Y`*
3. *Για κάθε  $i$  στο μήκος του `Dataset - LookBack-1`*  
$$a = Dataset[i : i + LookBack), 0]$$
*Κάνε `append` στον πίνακα `X` το  $a$*   
*Κάνε `append` στον πίνακα `Y` το `Dataset[(i + LookBack), 0]`*
4. *Έξοδος: Επέστρεψε τον πίνακα `X` και τον πίνακα `Y` ως `numpy arrays`.*

Τα δεδομένα εισόδου όπως φαίνεται είναι το `dataset`, σε μορφή `numpy` πίνακα και η μεταβλητή `Look_Back` που δηλώνει τον αριθμό των προηγούμενων χρονικών στιγμών που θα χρησιμοποιηθούν για την πρόβλεψη των επόμενων. Η αρχική τιμή της μεταβλητής ισούται με ένα. Κατά αυτόν τον τρόπο, θα δημιουργηθεί ένας πίνακας `X`, που θα περιλαμβάνει την τιμή κλεισίματος την χρονική στιγμή  $t$  και ένας πίνακας `Y` που θα αποτυπώνει την τιμή κλεισίματος της μετοχής την χρονική στιγμή  $t + 1$ .

Αξίζει να σημειωθεί, ότι εφόσον το σύνολο δεδομένων έχει μετατραπεί σε `numpy` πίνακα και η χρονοσειρά αποτελείται πλέον από μία μεταβλητή το πρόβλημα μετατρέπεται σε πρόβλημα παλινδρόμησης με συνεχόμενο αριθμό παρατηρήσεων που αντιστοιχούν σε συνεχόμενες ημέρες καταγραφής της τιμής κλεισίματος του χρηματιστηρίου.

## **5.6 Αρχιτεκτονική Long-Short Term Memory Network (LSTMN)**

Η πρώτη αρχιτεκτονική Τεχνητού Νεωρικού Δικτύου που θα χρησιμοποιηθεί για την αξιολόγηση της ως προς την πρόβλεψη μελλοντικών χρηματιστηριακών τιμών, χρονοσειράς με μία μεταβλητή είναι η LSTMN.

Ο σχεδιασμός ξεκινά δημιουργώντας ένα κλασσικό μοντέλο για τις τρεις χρονοσειρές και στη συνέχεια θα γίνουν πολλές δοκιμές με αλλαγές στις παραμέτρους. Ο

ασφαλέστερος τρόπος για τον προσδιορισμό της σωστής τιμής της παραμέτρου είναι μέσα από δοκιμές και αξιολόγησης των αποτελεσμάτων.

Το Τεχνητό Νευρωνικό Δίκτυο θα αποτελείται από τρία επίπεδα, δηλαδή το επίπεδο εισόδου, το κρυμμένο επίπεδο και το επίπεδο εξόδου που θα προβλέπει μία μεταβλητή. Ύστερα από δοκιμές, τα βέλτιστα αποτελέσματα έρχονται όταν έχει στο κρυφό επίπεδο 12 νευρώνες/blocks. Το σύνολο δεδομένων (dataset) χωρίστηκε σε δύο υποσύνολα train και test set, με αναλογία 80% - 20% αντίστοιχα και το batch size ανέρχεται στα 15. Αυτό συμβαίνει γιατί το μοντέλο χρειάζεται μεγαλύτερο όγκο δεδομένων στο στάδιο της εκπαίδευσης. Τέλος η συνάρτηση ενεργοποίησης όπως προαναφέρθηκε είναι η ReLu και η σιγμοειδής. [28]

### 5.6.1 LSTMN για την Amazon

Η χρονοσειρά Amazon ξεκινά από το 1997 και όπως παρατηρήθηκε παραπάνω τα πρώτα έτη η εταιρεία είχε σχεδόν μηδενικές τιμές κλεισίματος μετοχής. Ύστερα από κατάλληλες δοκιμές, αφαιρέθηκαν τα έτη πριν το 2000 όπου εμφάνιζαν πολύ διαφορετική συμπεριφορά από τα υπόλοιπα. Με τον τρόπο αυτό εκπαιδεύτηκε καλύτερα μοντέλο δίνοντας καλύτερα αποτελέσματα.

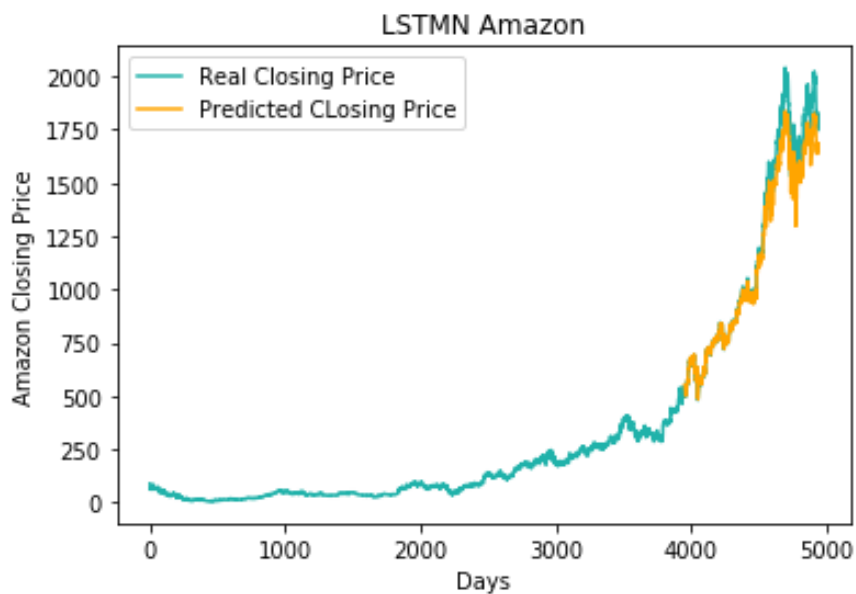
Έγιναν οι παρακάτω δοκιμές για την πρόβλεψη της τιμής κλεισίματος για την Amazon με τα αντίστοιχα αποτελέσματα:

Test	Optimizer	Epoch	Test MSE
1	adam	200	10736.26
2	adam	400	11567.88
3	adam	600	18268.31
4	sgd	200	773568.94
5	sgd	400	21185.64
6	sgd	600	16599.58
7	rmsprop	200	19850.19
8	rmsprop	400	39013.48

9	rmsprop	600	66942.32
---	---------	-----	----------

Από τα παραπάνω αποτελέσματα φαίνεται αποδοτικότερο να είναι το μοντέλο στα 200 epoch με optimizer adam και MSE 10736.26.

Στο παρακάτω διάγραμμα απεικονίζονται με πορτοκαλί οι προβλεπόμενες τιμές και με μπλε οι πραγματικές. Γίνεται κατανοητό ότι στις αρχικές παρατηρήσεις το μοντέλο παίζει σωστά, ωστόσο από τις 5000 και έπειτα οι διακυμάνσεις ανάμεσα στις προβλέψεις και στις πραγματικές τιμές είναι έντονες και ορατές, γεγονός που συνάδει και με την απότομη αύξηση του δείκτη MSE.



Γραφική Απεικόνιση των πραγματικών τιμών με γαλάζιο και των προβλεπόμενων με πορτοκαλί για την Amazon

### 5.6.2 LSTMN για την Google

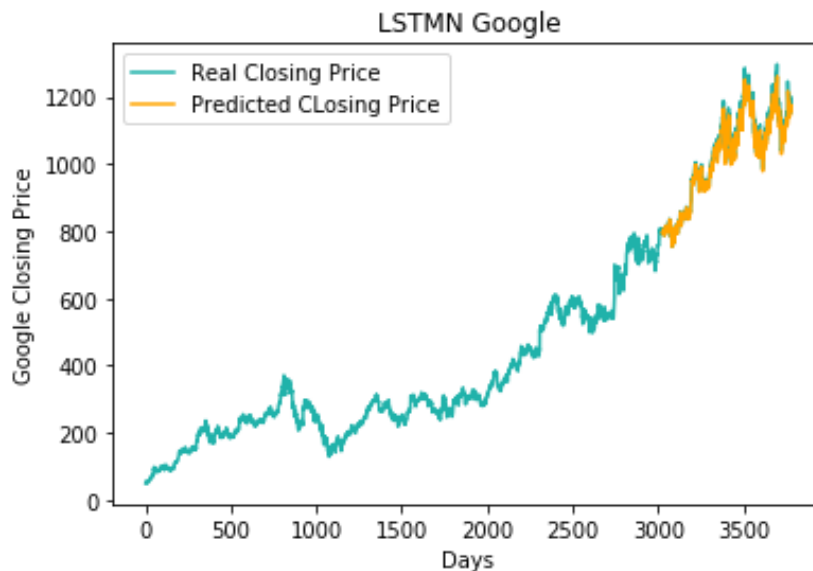
Έγιναν οι παρακάτω δοκιμές για την πρόβλεψη της τιμής κλεισίματος για την Google με τα αντίστοιχα αποτελέσματα:

Test	Optimizer	Epoch	Test MSE
1	adam	200	549.41
2	adam	400	634.91

3	adam	600	610.84
4	sgd	200	2864.52
5	sgd	400	1911.63
6	sgd	600	1381.22
7	rmsprop	200	960.61
8	rmsprop	400	837.97
9	rmsprop	600	377.83
10	rmsprop	800	1099.85

Από τα παραπάνω αποτελέσματα φαίνεται αποδοτικότερο να είναι το μοντέλο στα 600 epoch με optimizer rmsprop και MSE 377.83.

Γίνεται κατανοητό ότι το μοντέλο παρουσιάζει καλύτερα αποτελέσματα στην χρονοσειρά Amazon. Παρατηρούνται μικρές αποκλίσεις από τις 3500 παρατηρήσεις και έπειτα.



Γραφική Απεικόνιση των πραγματικών τιμών με γαλάζιο και των προβλεπόμενων με πορτοκαλί για την Google

### 5.6.3 LSTMN για την Apple

Η χρονοσειρά Apple ξεκινά από το 1980 και όπως παρατηρήθηκε παραπάνω τα πρώτα έτη η εταιρεία είχε σχεδόν μηδενικές τιμές κλεισίματος μετοχής. Ύστερα από

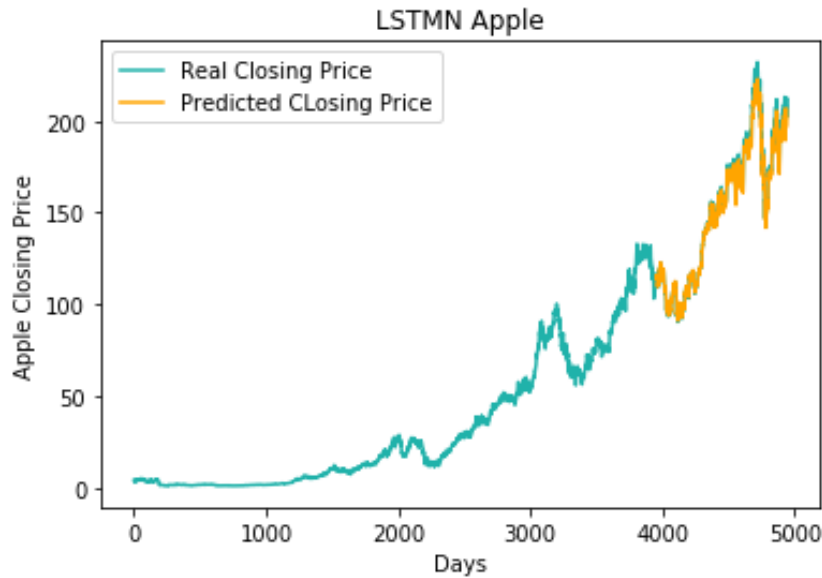
κατάλληλες δοκιμές, αφαιρέθηκαν τα έτη πριν το 2000 όπου εμφάνιζαν πολύ διαφορετική συμπεριφορά από τα υπόλοιπα, όπως και στην Amazon.

Έγιναν οι παρακάτω δοκιμές για την πρόβλεψη της τιμής κλεισίματος για την Apple με τα αντίστοιχα αποτελέσματα:

Test	Optimizer	Epoch	Test MSE
1	adam	200	18.97
2	adam	400	7.35
3	adam	600	21.26
4	sgd	200	65.67
5	sgd	400	30.99
6	sgd	600	79.98
7	sgd	800	26.12
8	rmsprop	200	21.46
9	rmsprop	400	34.12
10	rmsprop	600	14.69

Από τα παραπάνω αποτελέσματα αποδοτικότερο είναι το μοντέλο στα 400 epoch με optimizer adam και MSE 7.35.

Συγκριτικά στις τρεις χρονοσειρές το Τεχνητό Νευρωνικό Δίκτυο LSTMN παρουσίασε καλύτερη απόδοση στην χρονοσειρά Apple και χειρότερη στην Amazon έχοντας ως μέτρο σύγκρισης το MSE. Παράλληλα παρατηρώντας το παρακάτω γράφημα, είναι φανερό πως υπάρχουν έντονες αποκλίσεις των προβλεπόμενων τιμών που δεν ταιριάζουν με το πολύ μικρό MSE, το γεγονός αυτό πιθανόν να συνάδει με τις πολλές ημμεδινικές τιμές που είχε τα πρώτα χρόνια η Apple στο χρηματιστήριο.



Γραφική Απεικόνιση των πραγματικών τιμών με γαλάζιο και των προβλεπόμενων με πορτοκαλί για την Apple

## 5.7 Αρχιτεκτονικά Gated Recurrent Neural Networks (GRUs)

Η δεύτερη αρχιτεκτονική Τεχνητού Νεωρικού Δικτύου που θα χρησιμοποιηθεί για την αξιολόγηση της ως προς την πρόβλεψη μελλοντικών χρηματιστηριακών τιμών, χρονοσειράς με μία μεταβλητή είναι η GRU. Το Τεχνητό Νευρωνικό Δίκτυο θα αποτελείται από τρία επίπεδα, δηλαδή το επίπεδο εισόδου, το κρυμμένο επίπεδο και το επίπεδο εξόδου που θα προβλέπει μία μεταβλητή. Ύστερα από δοκιμές, τα βέλτιστα αποτελέσματα έρχονται όταν έχει στο κρυφό επίπεδο 12 νευρώνες/blocks. Το σύνολο δεδομένων (dataset) χωρίστηκε σε δύο υποσύνολα train και test set, με αναλογία 70% - 30% αντίστοιχα και το batch size ανέρχεται στα 15. Αυτό συμβαίνει γιατί το μοντέλο χρειάζεται μεγαλύτερο όγκο δεδομένων στο στάδιο της εκπαίδευσης. Τέλος η συνάρτηση ενεργοποίησης όπως προαναφέρθηκε είναι η ReLu και η σιγμοειδής.

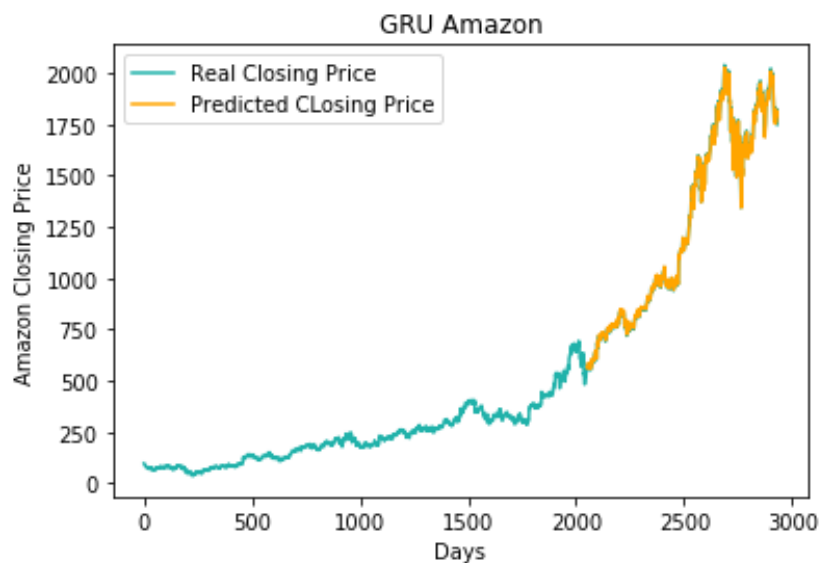
### 5.7.1 GRU για την Amazon

Τα αποτελέσματα από τις διάφορες δοκιμές της αρχιτεκτονικής GRU.

Test	Optimizer	Epoch	Test MSE
1	adam	200	1429.50
2	adam	400	1102.03
3	adam	600	1351.76

4	sgd	200	75197.73
5	sgd	400	1146.03
6	sgd	600	2279.56
8	rmsprop	200	1877.86
9	rmsprop	400	1233.22
10	rmsprop	600	1753.88

Αποδοτικότερα αποτελέσματα πρόβλεψης δίνει ο optimizer adam στα epoch 400 με 1102.03 MSE. Είναι φανερό ότι το δίκτυο GRU αποδίδει καλύτερα συγκριτικά με το δίκτυο LSTM για την Amazon. Το παρακάτω διάγραμμα έχει μικρές αποκλίσεις μεταξύ πραγματικών και προβλεπόμενων τιμών.



Γραφική Απεικόνιση των πραγματικών τιμών με γαλάζιο και των προβλεπόμενων με πορτοκαλί για την Amazon

### 5.7.2 GRU για την Google

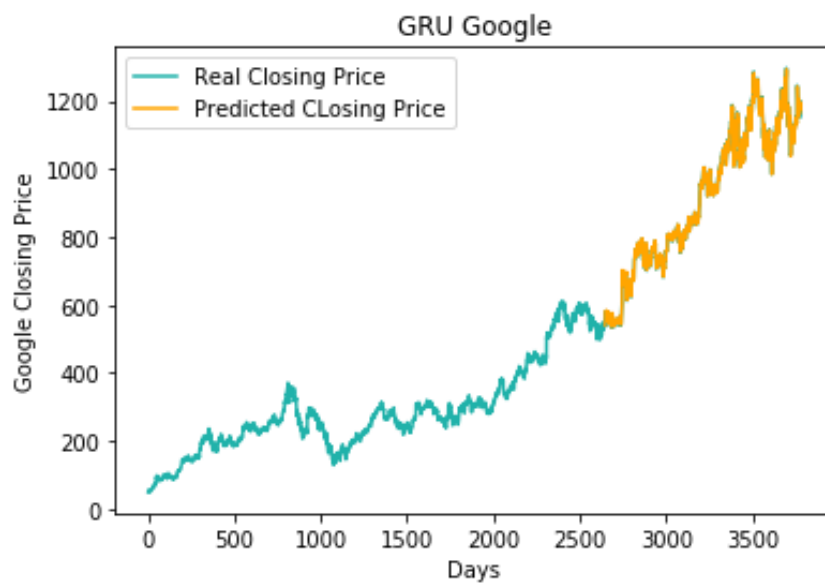
Ακολουθεί συγκριτικός πίνακας με τις δοκιμές που υλοποιήθηκαν για την χρονοσειρά της Google.

Test	Optimizer	Epoch	Test MSE
1	adam	200	253.57
2	adam	400	766.29
3	adam	600	268.08



4	sgd	200	688.33
5	sgd	400	4614.36
6	sgd	600	984.88
8	rmsprop	200	461.13
9	rmsprop	400	1181.89
10	rmsprop	600	205.22

Για την χρονοσειρά Google αποδοτικότερο ήταν το μοντέλο με optimizer rmsprop, στα 600 epochs , με 205.22 MSE. Το μοντέλο είχε πολύ καλά αποτελέσματα προβλέψεων.



Γραφική Απεικόνιση των πραγματικών τιμών με γαλάζιο και των προβλεπόμενων με πορτοκαλί για την Google

### 5.7.3 GRU για την Apple

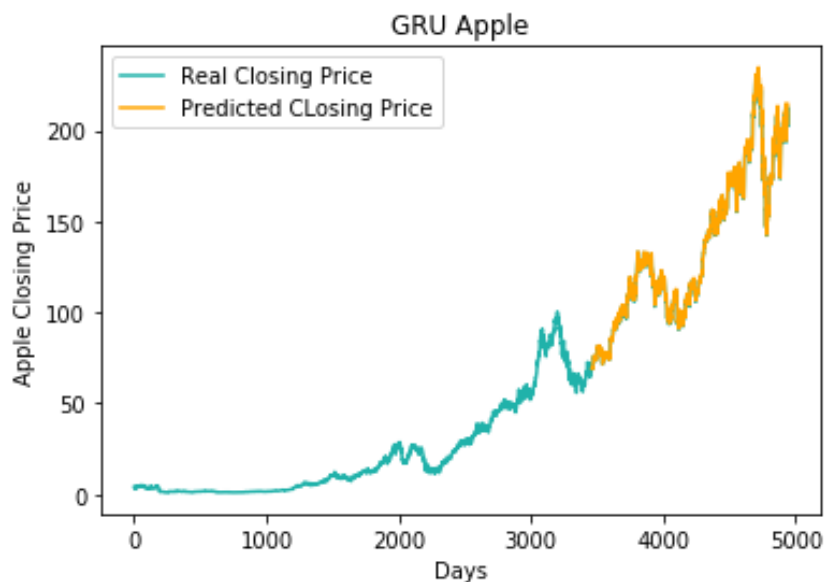
Για την χρονοσειρά Apple διατηρήθηκαν μόνο τα έτη 2000 και έπειτα.

Test	Optimizer	Epoch	Test MSE
1	adam	200	13.71
2	adam	400	11.36
3	adam	600	30.39
4	sgd	200	24.79
5	sgd	400	110.93

6	sgd	600	19.27
7	sgd	800	5.51
8	rmsprop	200	45.02
9	rmsprop	400	7.13
10	rmsprop	600	102.73

Από τα παραπάνω αποτελέσματα αποδοτικότερο είναι το μοντέλο στα 800 epoch με optimizer sgd και MSE 5.51.

Συγκριτικά στις τρεις χρονοσειρές το Τεχνητό Νευρωνικό Δίκτυο GRU παρουσίασε καλύτερη απόδοση στην χρονοσειρά Apple και χειρότερη στην Amazon έχοντας ως μέτρο σύγκρισης το MSE.



Γραφική Απεικόνιση των πραγματικών τιμών με γαλάζιο και των προβλεπόμενων με πορτοκαλί για την Apple

## 5.8 Αρχιτεκτονική Multilayer Perceptron (MLP)

Το Νευρωνικό δίκτυο MLP που θα χρησιμοποιηθεί, αποτελείται από τρία επίπεδα, το επίπεδο εισόδου, εξόδου και το κρυμμένο επίπεδο. Το σύνολο δεδομένων έχει χωριστεί σε 70% για εκπαίδευση και 30% για αξιολόγηση του δικτύου, με ρυθμό εισόδου στο

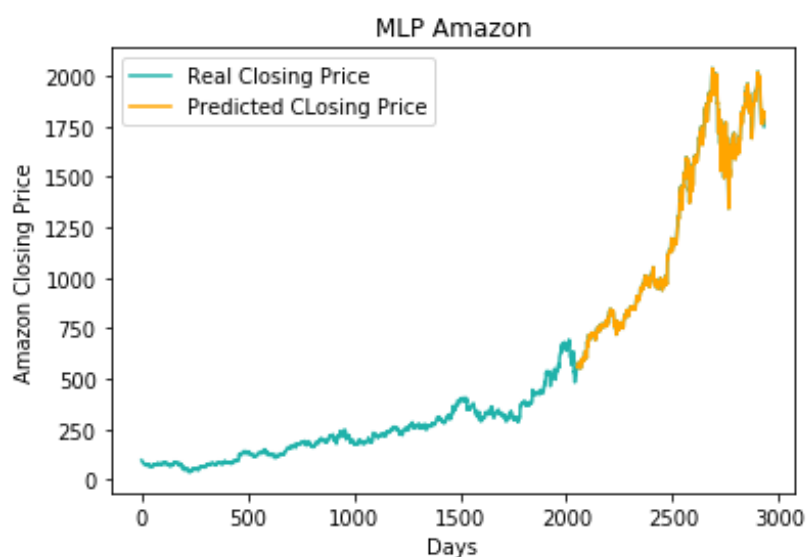
δίκτυο 15 batch size για την Amazon και την Google και 50 batch size για την Apple. Επιπλέον, το δίκτυο αποτελείται από 12 νευρώνες.

### 5.8.1 MLP για την Amazon

Για την χρονοσειρά Amazon, διατηρήθηκαν μόνο τα έτη από το 2000 και έπειτα. Τα αποτελέσματα του MLP φαίνονται παρακάτω:

Test	Optimizer	Epoch	Test MSE
1	adam	200	610.80
2	adam	400	656.95
3	adam	600	606.59
4	sgd	200	606.95
5	sgd	400	72476.72
6	sgd	600	606.04
8	rmsprop	200	44153.24
9	rmsprop	400	49242.37
10	rmsprop	600	1824.91

Το μοντέλο MLP είναι αποδοτικότερο για την Amazon με optimizer adam στα 600 epochs, με 606.59 MSE. Τελικώς η μοντελοποίηση αυτή είχε πιο καλές προβλέψεις και από δίκτυο LSTM και το δίκτυο GRU.



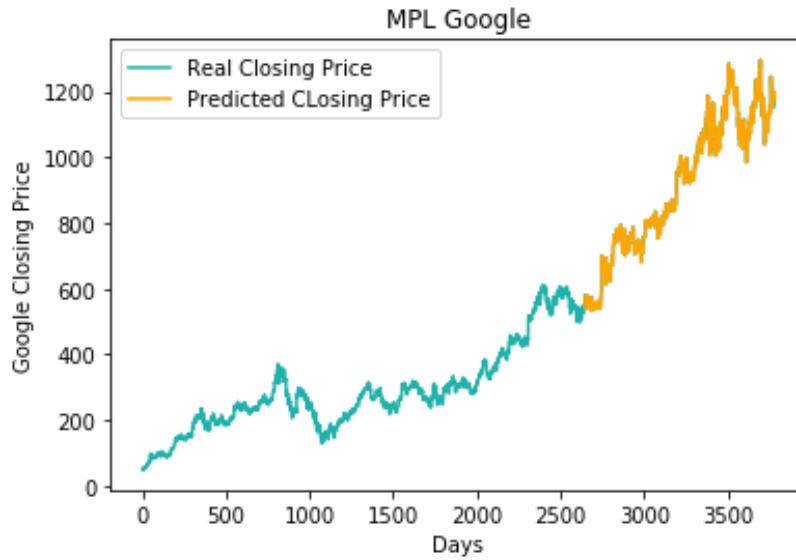
Γραφική Απεικόνιση των πραγματικών τιμών με μπλε και των προβλεπόμενων με πορτοκαλί για την Amazon με το MLP

### 5.8.2 MLP για Google

Τα αποτελέσματα για την χρονοσειρά φαίνονται παρακάτω:

Test	Optimizer	Epoch	Test MSE
1	adam	200	366.24
2	adam	400	232.15
3	adam	600	1793.72
4	sgd	200	445.37
5	sgd	400	204.74
6	sgd	600	2320.88
8	rmsprop	200	323.17
9	rmsprop	400	216.80
10	rmsprop	600	507.33

Σύμφωνα με τα αποτελέσματα καλύτερα αποτελέσματα παρουσιάζονται με τον optimizer sgd ,στα 400 epochs, με 204.74 MSE. Τα 400 epochs έφεραν και για τους τρεις optimizers καλύτερα αποτελέσματα. Από το παρακάτω διάγραμμα είναι φανερό πως το μοντέλο έχει ελάχιστες αποκλίσεις στις προβλεπόμενες τιμές συγκριτικά με τις πραγματικές.



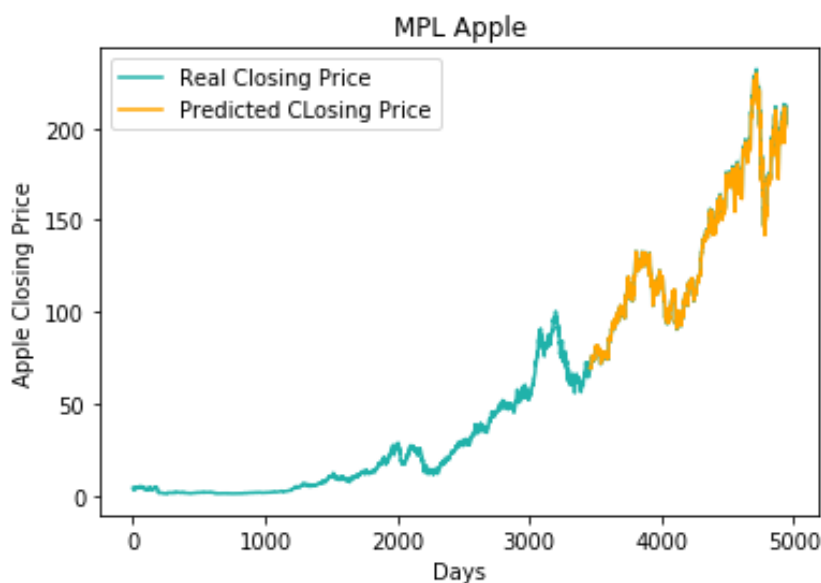
Γραφική Απεικόνιση των πραγματικών τιμών με μπλε και των προβλεπόμενων με πορτοκαλί για την Google με το MLP δίκτυο

### 5.8.3 MLP Apple

Για εφαρμογή του MLP στην χρονοσειρά Apple, έγινε εισαγωγή των δεδομένων με 50 batch size. Από την χρονοσειρά διατηρήθηκαν μόνο οι τιμές από το 2000 και έπειτα.

Test	Optimizer	Epoch	Test MSE
1	adam	200	94.69
2	adam	400	81.90
3	adam	600	5.62
4	adam	800	13.97
5	sgd	200	951.60
6	sgd	400	552.85
7	sgd	600	60.65
8	sgd	800	10.28
9	rmsprop	200	13.97
10	rmsprop	400	142.57
11	rmsprop	600	5.93
12	rmsprop	800	277.32

Το μοντέλο ήταν αποδοτικότερο με τον optimizer adam, στα 600 epochs, με 5.62 MSE. Από το παρακάτω διάγραμμα είναι φανερό πως υπάρχουν ελάχιστες αποκλίσεις και οι προβλεπόμενες τιμές είναι πολύ κοντά στις πραγματικές.



Γραφική Απεικόνιση των πραγματικών τιμών με μπλε και των προβλεπόμενων με πορτοκαλί για την Apple με το MLP δίκτυο

Ο Multilayer Perceptron παρουσίασε καλύτερα αποτελέσματα για την χρονοσειρά Apple. Ο δείκτης MSE ήταν με διαφορά πολύ μικρότερος από ότι στις χρονοσειρές Amazon και Google.

## 5.9 Σύνοψη Αποτελεσμάτων

Neural Network/ Timeseries	Amazon	Google	Apple
LSTM	10736.26	377.83	7.31
GRU	1102.03	205.22	5.51
MLP	606.59	204.74	5.62

Από το πίνακα σύνοψης φαίνεται πως αποδοτικότερα αποτελέσματα προβλέψεων έγιναν για την χρονοσειρά της Apple και για τα τρία δίκτυα που μοντελοποιήθηκαν. Παράλληλα ο Νευρωνικό Δίκτυο MLP είχε καλές αποδόσεις και για τις τρεις χρονοσειρές, χειρότερα αποτελέσματα παρήχθησαν για την Amazon.

Συμπερασματικά, αξιολογώντας τα αποτελέσματα της πειραματικής διαδικασίας γίνεται κατανοητό ότι η αρχιτεκτονική των τεχνητών νευρωνικών δικτύων που παράγει τα καλύτερα αποτελέσματα ως προς την παραγωγή προβλέψεων σε χρονολογικά δεδομένα μετοχών, είναι ο Multilayer Perceptron (MLP). Αυτό αποδεικνύεται από τις επιδόσεις του μοντέλου που παρήγαγε γενικότερα το μικρότερο μέσο τετραγωνικό σφάλμα στα σύνολα δεδομένων που χρησιμοποιήσαμε. Το αμέσως επόμενο μοντέλο από τον MLP ως προς τις επιδόσεις, είναι αυτό που βασίστηκε στην αρχιτεκτονική των Gated Recurrent Neural Networks, ενώ τη χειρότερη επίδοση την κατέγραψε το μοντέλο που βασίστηκε στην αρχιτεκτονική των Long Short Term Memory Network. Επίσης είναι πολύ σημαντικό να αναφέρουμε τις μεγάλες διακυμάνσεις στα σφάλματα ανάμεσα στα τρία σύνολα δεδομένων. Τα μικρότερα σφάλματα παρατηρήθηκαν στο σύνολο δεδομένων της εταιρείας Apple, ενώ τα μεγαλύτερα παρατηρήθηκαν στο σύνολο δεδομένων της εταιρείας Amazon. Το γεγονός αυτό είναι απολύτως λογικό, αφού το σύνολο δεδομένων της Apple είχε τις λιγότερες παρατηρήσεις και τη μικρότερη διακύμανση από τα τρία σύνολα δεδομένων, ενώ το σύνολο δεδομένων της Amazon είχε τις περισσότερες παρατηρήσεις και τη μεγαλύτερη διακύμανση από τα τρία σύνολα δεδομένων. Επομένως οδηγούμαστε στο συμπέρασμα ότι ένα μοντέλο τεχνητού νευρωνικού δικτύου που μπορεί να παράγει πολύ καλά αποτελέσματα για ένα συγκεκριμένο σύνολο χρονολογικών δεδομένων και ταυτόχρονα πολύ κακά αποτελέσματα για ένα άλλο σύνολο, που πιθανώς να χρειάζεται περισσότερη επεξεργασία πριν την εισαγωγή του στο Νευρωνικό Δίκτυο.

## Βιβλιογραφία

- [1] Μηχανική Μάθηση (Machine Learning) Τεχνητή Νοημοσύνη - Β' Έκδοση Ι. Βλαχάβας, Π. Κεφαλάς, Ν. Βασιλειάδης, Φ. Κόκκορας, Η. Σακελλαρίου, Κεφάλαιο 18.
- [2] Αικατερίνη Γεωργούλη (2015 Αθήνα). Μηχανική Μάθηση, Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, Κεφάλαιο 4.
- [3] Sodhi, Pinky and Awasthi, Naman and Sharma, Vishal, Introduction to Machine Learning and Its Basic Application in Python (January 6, 2019).
- [4] <https://nwfsc-timeseries.github.io/atsa-labs/sec-tslab-time-series-plots.html>.
- [5] <http://users.auth.gr/dkugiu/Teach/DataAnalysis/Chp6.pdf> .
- [6] [https://www.researchgate.net/publication/307663962\\_Time\\_Series\\_Modelling\\_and\\_Decomposition](https://www.researchgate.net/publication/307663962_Time_Series_Modelling_and_Decomposition).
- [7] <https://ec.europa.eu/eurostat/documents/64157/4374310/29-SEASONAL-ADJUSTMENT-METHODS-PRACTICES-2007.pdf/6628a64e-2160-4e6f-a34a-499d0f5cdcfe> .
- [8] <https://www.stat.ipb.ac.id/en/uploads/KS/S2%20-%20ADW/3%20Montgomery%20-%20Introduction%20to%20Time%20Series%20Analysis%20and%20Forecasting.pdf>.
- [9] <http://www.statsoft.com/Textbook/Time-Series-Analysis#simple>.
- [10] Douglas C. Montgomery, "Introduction to Time Series Analysis and Forecasting", 2008.
- [11] Sio-long Ao, "Applied Time Series Analysis and Innovative Computing", 2009.
- [12] Engle R.F., "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*. 50, pp.987-1008, 1982 .
- [13] Hamilton J.D., "Times Series Analysis", Princeton University Press, 1994. .
- [14] C. Chatfield, "The Analysis of time series: An Introduction", Third Edition, 1984 .
- [15] ΑΤΕΙ ΠΑΤΡΩΝ ΤΜΗΜΑ ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ & ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΣΗΜΕΙΩΣΕΙΣ ΜΑΘΗΜΑΤΟΣ ΤΕΧΝΙΚΩΝ ΠΡΟΒΛΕΨΕΩΝ & ΕΛΕΓΧΟΥ ,ΚΩΝΣΤΑΝΤΙΝΟΣ Η. ΚΟΥΝΕΤΑΣ ΠΑΤΡΑ ΔΕΚΕΜΒΡΙΟΣ 2012 .
- [16] Tom M. Mitchell (March 1, 1997). Machine Learning, McGraw-Hill Science/Engineering/Math, chapter 4. .
- [17] Artificial Neural Networks and Its Applications, Girish Jha, New Delhi, August 2014.



- [18] ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ «Υλοποίηση και εφαρμογή Τεχνητών Νευρωνικών Δικτύων για την πρόβλεψη χρονοσειρών συναλλαγματικών ισοτιμιών» Σαλάτας Ιωάννης, ΠΑΤΡΑ 2011.
- [19] Behaviour Analysis of Multilayer Perceptrons with Multiple Hidden Neurons and Hidden Layers Article in International Journal of Computer Theory and Engineering · January 2011.
- [20] Zachary C. Lipton, John Berkowitz, Charles Elkan, A Critical Review of Recurrent Neural Networks for Sequence Learning (17 October, 2015).
- [21] <http://adventuresinmachinelearning.com/keras-lstm-tutorial/>.
- [22] Long Short-Term Memory Recurrent Neural Network Architectures, Hasim Sak, Andrew Senior, Françoise Beaufays, Google USA, 2014.
- [23] <https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-packages>.
- [24] <https://rpubs.com/kapage/523169>.
- [25] <https://keras.io/api/models/>.
- [26] <https://keras.io/api/optimizers/>.
- [27] <https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/>.
- [28] <https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/>.
- [29] Μέθοδοι προβλέψεων και ανάλυσης αποφάσεων, Χρήστος Ν. Αγιακόγλου Γεώργιος Σ. Οικονόμου, Β' έκδοση, εκδόσεις Γ. Μπένου Αθήνα 2004.
- [30] [http://computationalfinance.lsi.upc.edu/?page\\_id=87](http://computationalfinance.lsi.upc.edu/?page_id=87)
- [31] <http://adventuresinmachinelearning.com/keras-lstm-tutorial/>