

Διπλωματική Εργασία

του φοιτητή του Τμήματος Ηλεκτρολόγων Μηχανικών και
Τεχνολογίας Υπολογιστών της Πολυτεχνικής Σχολής του
Πανεπιστημίου Πατρών

Παύλου Πανταζή

ΑΜ: 1053563

Θέμα

**«Πρόβλεψη χρηματιστηριακών τιμών μετοχών με
χρήση στατιστικών μοντέλων και μεθόδων μηχανικής
μάθησης για την εκτίμηση της επίδοσης
ανταγωνιστικών εταιρειών»**

Επιβλέπων

Δημήτρης Λυμπερόπουλος,

Καθηγητής

Αριθμός Διπλωματικής Εργασίας:
(Αριθμός μητρώου /2021)

Πάτρα, Ιούλιος 2021
...../...../.....

ΠΙΣΤΟΠΟΙΗΣΗ

Πιστοποιείται ότι η Διπλωματική Εργασία με θέμα

**«Πρόβλεψη χρηματιστηριακών τιμών μετοχών με
χρήση στατιστικών μοντέλων και μεθόδων μηχανικής
μάθησης για την εκτίμηση της επίδοσης
ανταγωνιστικών εταιρειών»**

Του φοιτητή του Τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας
Υπολογιστών

Παύλου Πανταζή του Νικολάου

Αριθμός Μητρώου: 1053563

Παρουσιάστηκε δημόσια και εξετάστηκε στο Τμήμα Ηλεκτρολόγων
Μηχανικών και τεχνολογίας Υπολογιστών στις

...../...../.....

Ο Επιβλέπων

Δημήτρης Λυμπερόπουλος,

Καθηγητής

Αριθμός Διπλωματικής Εργασίας:

Θέμα: «Πρόβλεψη χρηματιστηριακών τιμών μετοχών με χρήση στατιστικών μοντέλων και μεθόδων μηχανικής μάθησης για την εκτίμηση της επίδοσης ανταγωνιστικών εταιρειών»

Φοιτητής: Παύλος Πανταζής

Περίληψη

Η αυξανόμενη διαθεσιμότητα υπολογιστικών πόρων σε συνδυασμό με τις εξελίξεις των τελευταίων ετών τόσο σε επίπεδο λογισμικού όσο και σε επίπεδο υλικού, μας παρέχουν τη δυνατότητα υλοποίησης σύνθετων διεργασιών αλλά και χρήσης υπολογιστικών μοντέλων, τα οποία στο παρελθόν είχαν πολύ υψηλό υπολογιστικό κόστος. Έτσι, η εφαρμογή των εργαλείων της επιστήμης των υπολογιστών έχει αυξηθεί ραγδαία αλλάζοντας πολλά ερευνητικά και επαγγελματικά πεδία. Ένας από τους πλέον δημοφιλέστερους τομείς της επιστήμης των υπολογιστών είναι η μηχανική μάθηση και πιο συγκεκριμένα τα νευρωνικά δίκτυα. Τα νευρωνικά δίκτυα χρησιμοποιούνται σε ένα ευρύ φάσμα εφαρμογών από την ανάλυση εικόνων σε πραγματικό χρόνο για την ταυτοποίηση προσώπων μέχρι την πρόβλεψη χρονοσειρών.

Σκοπός της διπλωματικής εργασίας είναι η πρόβλεψη των ημερήσιων τιμών των μετοχών που συμπεριλαμβάνονται στο δείκτη Standard and Poor's 500. Οι προβλέψεις παράγονται

από μεθόδους μηχανικής μάθησης και κυρίως από νευρωνικά δίκτυα, οι επιδόσεις των οποίων επιδιώκεται να βελτιωθούν.

Στη συνέχεια με βάση τις παραγόμενες προβλέψεις εξετάζεται εάν μπορεί να υλοποιηθεί μία κερδοφόρα στρατηγική επενδύσεων κάνοντας χρήση μεθόδων βελτιστοποίησης χαρτοφυλακίου.

Τελικά αποδεικνύεται πως ο συνδυασμός προβλέψεων υψηλής ακρίβειας με μεθόδους βελτιστοποίησης χαρτοφυλακίου μπορεί να αποτελέσει μία κερδοφόρα στρατηγική ωστόσο δεν είναι βέβαιο πως η στρατηγική αυτή θα αποδίδει περισσότερο από τον δείκτη S&P 500.

Λέξεις κλειδιά: Μηχανική Μάθηση; Νευρωνικά Δίκτυα; Πρόβλεψη Μετοχών; Βελτιστοποίηση Χαρτοφυλακίου;

Abstract

The increasing availability of computational resources along with recent technological advancements both in software and hardware, enables us to complete complex tasks and use computational models, which would have been impossible in the past due to their high computational cost. Therefore, the applications of a variety of computer science tools have increased rapidly, changing many professions and research fields. One of the most popular aspects of computer science is machine learning and artificial neural networks. Neural Networks are applied in a wide spectrum of tasks ranging from real time image analysis for facial recognition to timeseries forecasting.

Following this trend, this dissertation aims to forecast the daily prices of the stocks that are part of the S&P 500 Index, with the use of machine learning algorithms and specifically neural networks. Then an automated trading strategy will be created based on the forecasts with the use of portfolio optimization methods.

Keywords: Machine Learning; Neural Networks; Forecasting; Portfolio Optimization;

Ευχαριστίες

Αρχικά θα ήθελα να ευχαριστήσω τον καθηγητή Βασίλειο Ασημακόπουλο για την ευκαιρία που μου έδωσε να εκπονήσω τη Διπλωματική Εργασία αυτή και τον επιβλέποντα καθηγητή Δημήτρη Λυμπερόπουλο.

Επίσης, σημαντικός αρωγός στην υλοποίηση της παρούσας Διπλωματικής Εργασίας υπήρξε ο Δρ. Ευάγγελος Σπηλιώτης, ερευνητικός συνεργάτης της Μονάδας Προβλέψεων και Στρατηγικής.

Περιεχόμενα

Ευρετήριο Σχημάτων	12
Ευρετήριο Πινάκων	14
Κεφάλαιο 1: Εισαγωγή	16
1.1 Μετοχές	16
1.2 Δείκτες	17
1.3 Λειτουργία του Χρηματιστηρίου	18
1.4 Στρατηγικές Επενδύσεων	20
1.5 Αλγοριθμικές Συναλλαγές	22
Κεφάλαιο 2: Νευρωνικά Δίκτυα	26
2.1 Εισαγωγή	26
2.2 Μοντέλο Τεχνητού Νευρώνα	28
2.3 Δομή Τεχνητών Νευρωνικών Δικτύων	34
2.4 Εκπαίδευση Τεχνητών Νευρωνικών Δικτύων	36
Αλγόριθμος Οπισθοδρομικής Διάδοσης του Σφάλματος (Back Propagation Algorithm)	37
2.5 Multi-Layer Perceptron (MLP)	44
2.6 Recurrent Neural Networks (RNNs)	45
2.6.1 Εισαγωγή	45
2.6.2 Παράδειγμα Χρήσης	46
2.6.3 Εκπαίδευση Αναδρομικών Δικτύων	48
2.7 Παράμετροι που επηρεάζουν την απόδοση των Νευρωνικών Δικτύων	52
Κεφάλαιο 3: Δέντρα Απόφασης	56
3.1 Εισαγωγή	56
3.2 Κατασκευή	57
3.3 Παράδειγμα Χρήσης	62
3.4 Πλεονεκτήματα και Μειονεκτήματα Δέντρων Απόφασης	66
Κεφάλαιο 4: Προβλέψεις	68
4.1 Εισαγωγή στις προβλέψεις	68
4.2 Χρονοσειρές	70
4.3 Ακρίβεια Προβλέψεων	73
Κεφάλαιο 5: Πειραματική Διαδικασία	77
5.1 Εισαγωγή	77
5.2 Σύνολο Δεδομένων	78
5.3 Μοντέλα Τοπικών Προβλέψεων	79
5.4 Γενικό Μοντέλο Προβλέψεων	82

5.5 Δέντρα Απόφασης.....	86
5.6 Recurrent Neural Networks	88
5.7 Σύνοψη Αποτελεσμάτων	89
Κεφάλαιο 6: Βελτιστοποίηση Χαρτοφυλακίου	91
6.1 Εισαγωγή	91
6.2 Μοντέλο του Markowitz.....	93
6.2.1 Μαθηματικός Ορισμός	93
6.2.2 Υλοποίηση	94
6.3 Το Μοντέλο του Young	100
6.3.1 Μαθηματική Περιγραφή	100
6.3.2 Υλοποίηση του μοντέλου	102
6.3.3 Μορφοποίηση του προβλήματος	102
6.3.4 Αποτέλεσμα.....	103
Κεφάλαιο 7: Αποτελέσματα Επενδύσεων	105
7.1 Εισαγωγή	105
7.2 Στρατηγική με χρήση Markowitz.....	106
7.3 Στρατηγική με χρήση Young	109
7.4 Σύνοψη Αποτελεσμάτων	113
7.5 Μελλοντικές Προεκτάσεις	114
Επίλογος	116
Παράρτημα	118
Βιβλιογραφία	135

Ευρετήριο Σχημάτων

Σχήμα 1: Η διαφορά τιμής αγοράς-πώλησης	20
Σχήμα 2: Ποσοστό των αγοροπωλησιών που εκτελούνται αλγοριθμικά στις χρηματιστηριακές αγορές	23
Σχήμα 3: Μέγεθος της Αγοράς των Αλγοριθμικών Συναλλαγών	24
Σχήμα 4: Μοντέλο Τεχνητού Νευρώνα.....	28
Σχήμα 5: Καμπύλη Εκθετικής Σιγμοειδούς για τρεις διαφορετικές τιμές του α . Με μπλε χρώμα απεικονίζεται η καμπύλη με το μεγαλύτερο α , με κόκκινο η καμπύλη με το μικρότερο α . Για την πράσινη καμπύλη το α είναι μεταξύ αυτών της μπλε και της κόκκινης.....	30
Σχήμα 6: Καμπύλη Υπερβολικής Εφαπτομένης	30
Σχήμα 7: Καμπύλη Συνάρτησης Προσήμου	31
Σχήμα 8: Καμπύλη Τμηματικά Γραμμικής Συνάρτησης	32
Σχήμα 9: Καμπύλη Συνάρτησης Δύο Κλάδων	32
Σχήμα 10: Καμπύλη Βηματικής Συνάρτησης	33
Σχήμα 11: Καμπύλη Συνάρτησης Relu	33
Σχήμα 12: Νευρωνικό Δίκτυο Ενός Επιπέδου.....	34
Σχήμα 13: Νευρωνικό Δίκτυο Πολλών Επιπέδου	35
Σχήμα 14: : Επανατροφοδοτούμενο Νευρωνικό Δίκτυο	36
Σχήμα 15: Οπισθοδρομική Διάδοση Σφάλματος	39
Σχήμα 16: Μοντέλο Τεχνητού Νευρώνα.....	39
Σχήμα 17: Μοντέλο Multi-Layer Perceptron.....	44
Σχήμα 18: Μοντέλο Recurrent Neural Network	45
Σχήμα 19: Μοντέλο Recurrent Neural Network	46
Σχήμα 20: Αρχιτεκτονική Recurrent Neural Network.....	47
Σχήμα 21: Σχήμα Εκπαίδευσης Αναδρομικών Δικτύων	49
Σχήμα 22: Παράδειγμα Δέντρου Απόφασης	56
Σχήμα 23: Μορφολογικά Χαρακτηριστικά Δέντρου Απόφασης	58
Σχήμα 24: Διάγραμμα Συνάρτησης $\log(P(x))$	59
Σχήμα 25: Διαδικασία Επέκτασης Δέντρου Απόφασης	64
Σχήμα 26: Συνέχεια Επέκτασης Δέντρου Απόφασης	65
Σχήμα 27: Τελικό Δέντρο Απόφασης	66
Σχήμα 28: Adjacent Closing Price της Google	71
Σχήμα 29: Παράδειγμα Κυκλικότητας μίας Χρονοσειράς	72
Σχήμα 30: Παράδειγμα Χαρακτηριστικών μίας Χρονοσειράς.....	73
Σχήμα 31: Διάγραμμα Αλγορίθμου Εύρεσης Βέλτιστου ΤΝΔ	79
Σχήμα 32: Δύο Διαφορετικοί Ορίζοντες Πρόβλεψης	81
Σχήμα 33: Διαδικασία Εύρεσης Βέλτιστης Αρχιτεκτονικής.....	82
Σχήμα 34: Διαδικασία Εύρεσης Βέλτιστων Παραμέτρων	84
Σχήμα 35: Διαδικασία Εύρεσης Βέλτιστων Παραμέτρων για τα Δέντρα Απόφασης	87
Σχήμα 36: MAPE Score για τα διάφορα μοντέλα πρόβλεψης.....	89
Σχήμα 37: Διάγραμμα Κέρδους Ρίσκου για τυχαία Χαρτοφυλάκια.....	97
Σχήμα 38: Συνολική Διαδικασία Υλοποίησης	99
Σχήμα 39: Efficient Frontier	100
Σχήμα 40: Διάγραμμα Στρατηγικής με χρήση Markowitz.....	106

Σχήμα 41: Διάγραμμα Στρατηγικής με χρήση Markowitz και εκ νέου εκπαίδευση.....	107
Σχήμα 42: Ημερήσια Κερδοφορία του S&P500 σε σχέση με την στρατηγική Markowitz.	108
Σχήμα 43: Διάγραμμα Στρατηγικής με χρήση Young	109
Σχήμα 44: Διάγραμμα Στρατηγικής με χρήση Young και εκ νέου εκπαίδευση	110
Σχήμα 45: Boxplot διαφορετικών Χαρτοφυλακίων που προκύπτουν από την αλλαγή των αρχικών συνθηκών του Young	110
Σχήμα 46: Ημερήσια Returns διαφορετικών Χαρτοφυλακίων που προκύπτουν από Young με διαφορετικές αρχικές συνθήκες	111
Σχήμα 47: Κερδοφορία των δύο διαφορετικών προσεγγίσεων και του S&P 500 για το 2019	113

Ευρετήριο Πινάκων

Πίνακας 1: Τύποι RNN.....	52
Πίνακας 2: Σύνολο Δεδομένων που χρησιμοποιείται στο παράδειγμα των Δέντρων Απόφασης	63
Πίνακας 3: Αποτελέσματα υπολογισμών πρώτου βήματος.....	64
Πίνακας 4: Αποτελέσματα υπολογισμών για το δεύτερο διαχωρισμό	65
Πίνακας 5: Βέλτιστες Παράμετροι και τιμές Ορίζοντα πρόβλεψης για 3 μετοχές	81
Πίνακας 6: Βέλτιστη Αρχιτεκτονική Γενικού Νευρωνικού Δικτύου για κάθε Ορίζοντα Πρόβλεψης.....	83
Πίνακας 7: Βέλτιστη Αρχιτεκτονική Γενικού Νευρωνικού Δικτύου για κάθε Ορίζοντα Πρόβλεψης με την δεύτερη προσέγγιση	84
Πίνακας 8: Χαρακτηριστικά Χρονοσειρών με το μικρότερο MAPE	85
Πίνακας 9: Αποτελέσματα Δέντρων Απόφασης.....	88
Πίνακας 10: Αποτελέσματα Πορβλέψεων με Αναδρομικά Νευρωνικά Δίκτυα	88
Πίνακας 11: Μέσο ετήσιο Κέρδος Μετοχών για τα έτη 2013-2019.....	95
Πίνακας 12: Covariance Μετοχών	95
Πίνακας 13:Μορφολογία δομής δεδομένων τυχαίων χαρτοφυλακίων. Ο πλήρης πίνακας αποτελεσμάτων διαθέτει 1000 γραμμές (μία για κάθε χαρτοφυλάκιο) και 469 στήλες με τα βάρη των μετοχών	96
Πίνακας 14: Αποτελέσματα Βελτιστοποίησης Χαρτοφυλακίου με το Μοντέλο του Young	104
Πίνακας 15: Ανάλυση Χαρτοφυλακίου 10ης μέρας 1ης στρατηγικής	107
Πίνακας 16: Ανάλυση Χαρτοφυλακίου 246ης μέρας 2ης προσέγγισης.....	112
Πίνακας 17: Tickers των μετοχών που συμπεριλαμβάνονται στο τελικό σύνολο δεδομένων.....	118
Πίνακας 18: Αποτελέσματα διερεύνησης βέλτιστου νευρωνικού δικτύου για κάθε μετοχή	128
Πίνακας 19: Αποτελέσματα διερεύνησης βέλτιστου ολικού δικτύου	129
Πίνακας 20: Αποτελέσματα πειραμάτων με πρόσθετα χαρακτηριστικά	134

Κεφάλαιο 1: Εισαγωγή

1.1 Μετοχές

Οι μετοχές είναι μερίδια ιδιοκτησίας μίας επιχείρησης και δεν έχουν συγκεκριμένη ημερομηνία λήξεως. Παρέχουν στον ιδιοκτήτη ένα ποσοστό των περιουσιακών στοιχείων και των κερδών της εταιρείας ανάλογα με την ποσότητα των μετοχών που αυτός διαθέτει σε σχέση με τον συνολικό αριθμό των μετοχών της επιχείρησης. Περιοδικά οι εταιρείες διανέμουν ένα μέρος ή το σύνολο των κερδών στους μετόχους υπό μορφή μερίσματος.

Το βασικό πλεονέκτημα των μετοχών είναι ότι οι κάτοχοι τους συμμετέχουν πλήρως στην αύξηση της κερδοφορίας ή του ενεργητικού της επιχείρησης, ενώ το βασικό τους μειονέκτημα είναι ότι σε περίπτωση εκκαθάρισης η επιχείρηση πρέπει να πληρώσει τους κατόχους των δανείων και των ομολόγων πριν από τους μετόχους.

Οι αγορές και οι πωλήσεις των μετοχών πραγματοποιούνται κυρίως δημόσια στις χρηματιστηριακές αγορές. Μερικές από τις δημοφιλέστερες χρηματιστηριακές αγορές είναι το Χρηματιστήριο της Νέας Υόρκης, το Χρηματιστήριο του Λονδίνου κ.α.

Οι αγοροπωλησίες των μετοχών στις χρηματιστηριακές αγορές πραγματοποιούνται κυρίως κατά τη διάρκεια λειτουργίας τους, τυπικά για 7 ώρες κάθε μέρα. Η τιμή μίας μετοχής διαμορφώνεται μέσω του νόμου προσφοράς και ζήτησης και μπορεί να αλλάζει συνεχώς κατά τη διάρκεια της ημέρας. Τα τελευταία χρόνια με την εξέλιξη της τεχνολογίας η επικοινωνία των επενδυτών με τις χρηματιστηριακές αγορές γίνεται μέσω του διαδικτύου και έτσι οι τόσο οι αγορές όσο και οι πωλήσεις πραγματοποιούνται με πολύ μεγαλύτερη ταχύτητα γεγονός που επηρεάζει τις τιμές των μετοχών αλλά και άλλων προϊόντων.

Μερικά από τα βασικά μεγέθη που χαρακτηρίζουν την τιμή μίας μετοχής κατά τη διάρκεια μίας ημέρας είναι:

- Τιμή ανοίγματος μετοχής (Opening Price): Η τιμή της μετοχής κατά το άνοιγμα της αγοράς
- Μέγιστη τιμή: Η μέγιστη τιμή μίας μετοχής κατά τη διάρκεια της ημέρας
- Ελάχιστη τιμή: Η ελάχιστη τιμή μίας μετοχής κατά τη διάρκεια της ημέρας
- Τιμή μετοχής στο κλείσιμο: Η τιμή της μετοχής τη στιγμή κλεισίματος της αγοράς
- Όγκος (Volume): Ο αριθμός των μετοχών που αγοράστηκαν ή πωλήθηκαν

Επιπλέον ένα ακόμα σημαντικό μέγεθος είναι το Adjacent Closing Price. Σε αντίθεση με την τιμή κλεισίματος της μετοχής, το Adjacent Closing Price λαμβάνει υπόψιν του εταιρικές δράσεις όπως τα μερίσματα ή τον διαχωρισμό των μετοχών έτσι ώστε να αντικατοπτρίζει καλύτερα την αξία της μετοχής. Στα πλαίσια της διπλωματικής εργασίας θα χρησιμοποιηθεί το μέγεθος αυτό και όχι η απλή τιμή κλεισίματος της μετοχής.

1.2 Δείκτες

Χρηματιστηριακός δείκτης είναι μία μέτρηση ενός τομέα (ή της συνολικής αξίας) του χρηματιστηρίου. Οι δείκτες χρησιμοποιούνται από τους επενδυτές για την περιγραφή της γενικής κατάστασης της αγοράς και τη σύγκριση με άλλες επενδύσεις ή άλλες χρονικές περιόδους.

Οι δείκτες δεν αποτελούν πραγματικό περιουσιακό στοιχείο και έτσι δεν μπορούν να γίνουν επενδύσεις κατευθείαν σε αυτούς. Ωστόσο υπάρχουν επενδυτικά στοιχεία που συνδέουν την απόδοσή τους με αυτήν ενός συγκεκριμένου δείκτη. Πολλές φορές οι δείκτες είναι σταθμισμένοι μέσοι όροι των τιμών των μετοχών που αντιπροσωπεύουν.

Ο Standard & Poor's 500 (S&P 500) είναι ένας δείκτης για τις 500 εισηγημένες εταιρείες στις αγορές του NYSE και του NASDAQ με τη μεγαλύτερη κεφαλαιοποίηση. Αποτυπώνει το 80% της διαθέσιμης κεφαλαιοποίησης της αγοράς και συχνά χρησιμοποιείται από τους επενδυτές για την περιγραφή της γενικότερης εικόνας της οικονομίας των ΗΠΑ.

Στα πλαίσια της συγκεκριμένης εργασίας θα μελετηθεί το σύνολο των μετοχών που συμπεριλαμβάνονται στο δείκτη S&P 500 από το 2013 έως το 2019.

1.3 Λειτουργία του Χρηματιστηρίου

Στην παρούσα ενότητα θα γίνει παρουσίαση των βασικών στοιχείων της λειτουργίας του χρηματιστηρίου και κάποιων βασικών εννοιών.

Στις χρηματιστηριακές αγορές οι επενδυτές έχουν στη διάθεσή τους χιλιάδες προϊόντα στα οποία μπορούν να επενδύσουν. Οι βασικές κατηγορίες των προϊόντων αυτών είναι οι ακόλουθες:

- Μετοχές

Κάθε ημέρα λειτουργίας του χρηματιστηρίου ιδιοκτήτες μετοχών επιθυμούν να πουλήσουν το μερίδιό τους ενώ άλλοι θέλουν να αγοράσουν. Για κάθε εισηγημένη εταιρεία στο χρηματιστήριο υπάρχει συνεχής προσφορά μετοχών είτε από την ίδια την εταιρεία είτε από επενδυτές που έχουν αποκτήσει μερίδιο. Οι τιμή μίας μετοχής καθορίζεται από την προσφορά και τη ζήτηση αυτής στη χρηματιστηριακή αγορά.

- Ομόλογα

Το ομόλογο αποτελεί ένα χρεόγραφο που εκδίδεται είτε από δημόσιο είτε από ιδιωτικό οργανισμό και χρησιμοποιείται για το δανεισμό κεφαλαίων στους επενδυτές. Υπάρχουν πολλοί τύποι ομολόγων με διάφορες διάρκειες ζωής. Γενικά, τα ομόλογα θεωρούνται ασφαλέστερες επενδύσεις συγκριτικά με τις μετοχές γιατί προσφέρουν σταθερά εισοδήματα με πολύ μικρότερο κίνδυνο απώλειας του αρχικού κεφαλαίου.

- Παράγωγα Προϊόντα

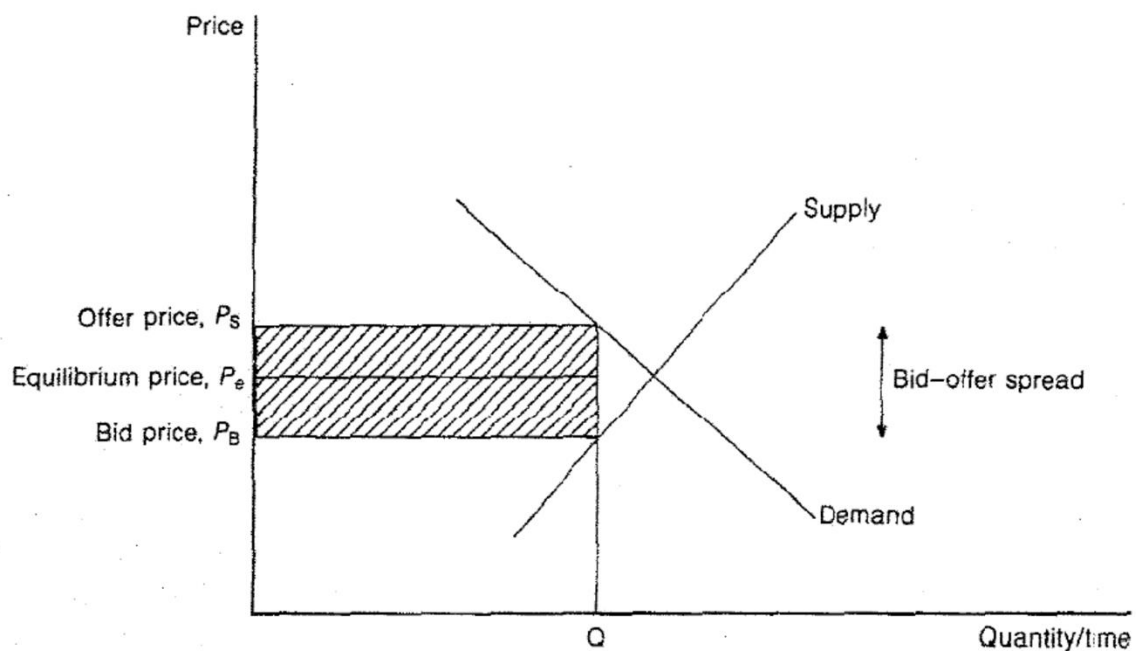
Ως παράγωγα προϊόντα στις χρηματιστηριακές αγορές ορίζονται τα συμβόλαια, η αξία των οποίων εξαρτάται από την αξία κάποιου άλλου βασικότερου προϊόντος (το βασικό προϊόν ονομάζεται υποκείμενο προϊόν). Η αξία ενός παραγωγού λοιπόν, υπολογίζεται με άμεσο τρόπο από την τιμή του υποκείμενου τίτλου. Τα υποκείμενα προϊόντα μπορεί να είναι οτιδήποτε από μετοχές και ομόλογα μέχρι αγροτικά προϊόντα (καφές) και μέταλλα (τιτάνιο). Οι δημοφιλέστερες κατηγορίες παραγωγών προϊόντων είναι:

- α. Τα προθεσμιακά συμβόλαια (Forward Contracts)
- β. Συμβόλαια Μελλοντικής Εκπλήρωσης (Future Contracts)
- γ. Δικαιώματα Προαίρεσης (Option Contracts)
- δ. Swaps

Ο χρηματιστής (stock broker) είναι εκείνος που πραγματοποιεί τις συναλλαγές κάθε επενδυτή στο χρηματιστήριο. Οι αγοροπωλησίες μετοχών, ομολόγων και των λοιπών προϊόντων εκτελούνται από τους χρηματιστές εκ μέρους των επενδυτών. Η αμοιβή των χρηματιστηριακών εταιρειών για τις υπηρεσίες που προσφέρουν ορίζεται σε κάποιο ποσοστό των συναλλαγών που εκτελούν κατόπιν συμφωνίας με τους επενδυτές. Παλαιότερα, οι υπηρεσίες αυτές ήταν αρκετά ακριβές με αποτέλεσμα πολλοί επενδυτές μικρών κεφαλαίων να μην μπορούν να εκτελέσουν χρηματιστηριακές συναλλαγές. Ωστόσο, σήμερα με την εξέλιξη της τεχνολογίας υπάρχουν πολλές χρηματιστηριακές εταιρείες που λειτουργούν καθαρά διαδικτυακά και απαιτούν πολύ μικρότερες αμοιβές. Για παράδειγμα, κάποιες από αυτές τις εταιρείες είναι η Degiro και η Interactive Brokers. Ενδεικτικά αναφέρεται πως για έναν Έλληνα επενδυτή που επιθυμεί να επενδύσει σε μετοχές των Ηνωμένων Πολιτειών η χρέωση σε μία από τις παραπάνω εταιρείες είναι €0.50 και 0.004\$ ανά μετοχή.

Ιδιάζων ρόλο στις σύγχρονες χρηματιστηριακές αγορές διαδραματίζουν οι εταιρείες που ονομάζονται ειδικοί διαπραγματευτές (market makers). Οι εταιρείες αυτές ενισχύουν την αποτελεσματικότητα και τη ρευστότητα των μετοχών για τις οποίες είναι υπεύθυνες και έτσι όλη την αγορά. Πρακτικά οι εταιρείες αυτές εισάγουν εντολές πώλησης και αγοράς συγκεκριμένων μετοχών ή άλλων χρεογράφων και έτσι μόλις γίνει εντολή πώλησης σε κάποιο επενδυτή, η εταιρεία πουλάει τις μετοχές ενώ αν δεν υπάρχουν παραγγελίες οι

μετοχές παραμένουν ως απόθεμα. Το κέρδος των εταιρειών αυτών προέρχεται από τη διαφορά στην τιμή αγοράς και πώλησης. Η διαφορά αυτή αποτελεί ένα αποδεκτό κόστος για την αγορά και τους επενδυτές ώστε να διαφυλαχθεί η ρευστότητα και η ομαλή λειτουργία της αγοράς. Στο παρακάτω σχήμα φαίνεται η διαφορά τιμής αγοράς-πώλησης.



Σχήμα 1: Η διαφορά τιμής αγοράς-πώλησης

1.4 Στρατηγικές Επενδύσεων

Οι επενδυτές έχουν στη διάθεσή τους πολλούς τρόπους με τους οποίους μπορούν να επενδύσουν στις χρηματιστηριακές αγορές. Με την εξέλιξη της τεχνολογίας και την ψηφιοποίηση της αγοράς έχουν αναπτυχθεί νέες στρατηγικές επενδύσεων ενώ οι παλαιότερες χρησιμοποιούνται όλο και λιγότερο. Στη συνέχεια της ενότητας αυτής θα γίνει μία περιγραφή των βασικών στρατηγικών επενδύσεων που χρησιμοποιούνται σήμερα και κυρίως του algorithmic trading.

Η τεχνική μακροπρόθεσμων επενδύσεων ή αλλιώς buy & hold είναι μία τεχνική με βάση την οποία οι επενδυτές αγοράζουν ένα σύνολο χρηματιστηριακών προϊόντων (μετοχές, ομόλογα κλπ) και τα διατηρούν για μεγάλο χρονικό διάστημα το οποίο συχνά ξεπερνάει το ένα έτος. Οι επενδυτές που ακολουθούν αυτήν την στρατηγική δεν ενδιαφέρονται για

βραχυπρόθεσμες αλλαγές στην τιμή των προϊόντων που έχουν στο χαρτοφυλάκιο τους, σκοπεύουν να διατηρήσουν τα προϊόντα αυτά μέχρι οι επενδύσεις τους να έχουν το απαιτούμενο κέρδος και έτσι συχνά καλούνται παθητικοί επενδυτές. Για να επενδύσουν σε ένα χρηματιστηριακό προϊόν εξετάζουν τα θεμελιώδη χαρακτηριστικά του, οδηγούνται σε ένα συμπέρασμα για την πραγματική του αξία και αν η τιμή της αγοράς για το συγκεκριμένο προϊόν είναι μικρότερη από αυτή που οι ίδιοι θεωρούν ότι είναι η πραγματική τότε επενδύουν. Ένα από τα πλεονεκτήματα της τεχνικής αυτής είναι η ελαχιστοποίηση των εξόδων σε προμήθειες προς τις χρηματιστηριακές εταιρείες καθώς και οι μειωμένες εισφορές στην εφορία. Ιστορικά είναι αποδεδειγμένο πως τέτοιες στρατηγικές επενδύσεων σε βάθος δεκαετιών αποφέρουν τα επιθυμητά αποτελέσματα. Ο Warren Buffett και ο Peter Lynch είναι από τους πιο επιτυχημένους επενδυτές τέτοιου τύπου.

Εκτός από τους παθητικούς επενδυτές υπάρχουν και επενδυτές που ακολουθούν την αντίθετη στρατηγική οι λεγόμενοι ενεργοί επενδυτές (active traders). Η τεχνική επένδυσης αυτή χρησιμοποιείται κυρίως από έμπειρους και επαγγελματίες επενδυτές και συνήθως βασίζεται στην αγοροπωλησία μετοχών, συμβολαίων μελλοντικής εκπλήρωσης και την αγορά συναλλάγματος. Οι active traders ανάλογα με το ρίσκο που επιθυμούν να πάρουν και τους οικονομικούς τους στόχους βασίζονται σε διάφορα κριτήρια, για παράδειγμα στην αγοροπωλησία μετοχών τα πιο σημαντικά κριτήρια είναι η ρευστότητα (liquidity), η μεταβλητότητα (volatility) που παρουσιάζει κάθε μετοχή και ο όγκος των συναλλαγών (volume). Μία από τις μεγαλύτερες κατηγορίες ενεργών επενδυτών είναι οι day traders που αγοράζουν και πωλούν χρηματιστηριακά προϊόντα μέσα στην ίδια μέρα. Η υλοποίηση μιας τέτοιας στρατηγικής απαιτεί εξαιρετική γνώση της αγοράς και μεγάλο διαθέσιμο κεφάλαιο διότι οι ποσοστιαίες αλλαγές στις τιμές των προϊόντων εντός μίας ημέρας συνήθως δεν είναι μεγάλες.

Σήμερα πολλοί μικροεπενδυτές αλλά και μεγάλοι επενδυτικοί οργανισμοί χρησιμοποιούν και ποσοτικές μεθόδους (quantitative methods) για την ανάλυση της χρηματιστηριακής αγοράς και για να πάρουν αποφάσεις σχετικά με τις διάφορες επενδύσεις τους. Οι επενδυτικές εταιρείες που πραγματοποιούν μεγάλο όγκο συναλλαγών κάθε μέρα βασίζονται σε μεγάλο βαθμό σε ποσοτικές αναλύσεις. Η εφαρμογή των ποσοτικών τεχνικών είναι άρρηκτα συνδεδεμένη με τις αλγοριθμικές συναλλαγές που αναλύονται στη συνέχεια.

1.5 Αλγοριθμικές Συναλλαγές

Ως algorithmic trading ορίζεται ένα εξελιγμένο επενδυτικό σύστημα του οποίου η λειτουργία βασίζεται σε σύνθετα μαθηματικά μοντέλα και στόχους του είναι η κερδοφόρα αγοροπωλησία χρηματιστηριακών προϊόντων. Η ύπαρξη αυτών των συστημάτων οφείλεται στην εξέλιξη των υπολογιστών και του διαδικτύου και στο γεγονός ότι η εξέλιξη αυτή δίνει στους επενδυτές τη δυνατότητα να υλοποιούν κερδοφόρες συναλλαγές σε ταχύτητα και συχνότητα που είναι αδύνατο να πραγματοποιηθούν χειροκίνητα.

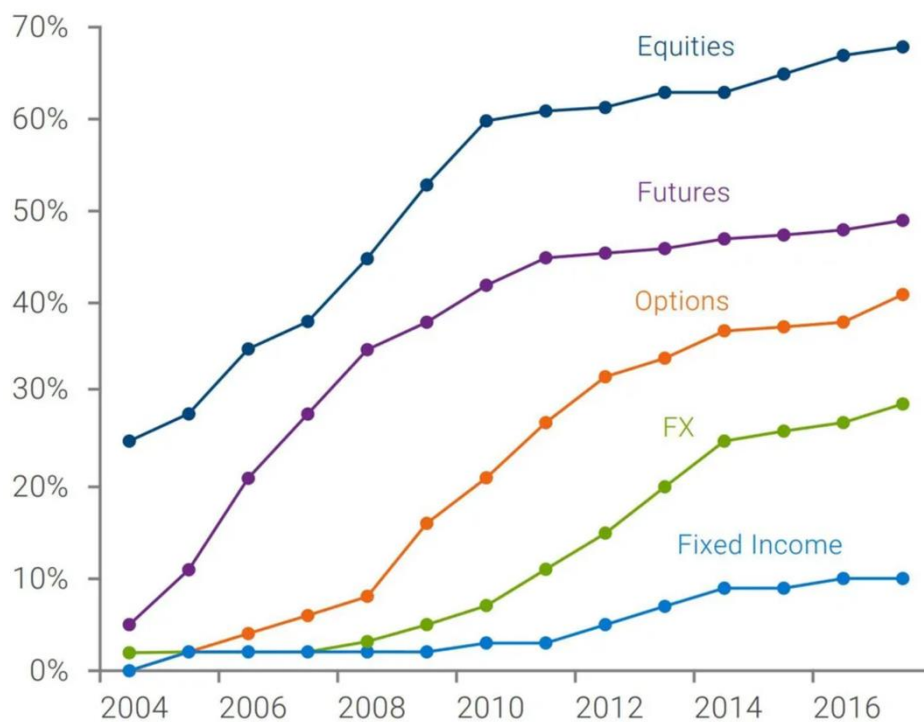
Το algorithmic trading χρησιμοποιείται από πολλούς οργανισμούς οι οποίοι μπορούν να έχουν διαφορετικούς ρόλους και στόχους στην αγορά. Οι μακροπρόθεσμοι και μεσοπρόθεσμοι επενδυτές χρησιμοποιούν αλγορίθμους έτσι ώστε να αγοράζουν αποθέματα σε μεγάλες ποσότητες όταν δεν επιθυμούν να επηρεάσουν τις τιμές των μετοχών με επενδύσεις μεγάλου όγκου. Εάν προσπαθήσουν να αγοράσουν μεγάλο όγκο μετοχών τότε εξαιτίας του νόμου προσφοράς και ζήτησης η τιμή των μετοχών αυτών ενδεχομένως να επηρεαστεί και έτσι αξιοποιούν διάφορους αλγορίθμους ώστε να αποκτήσουν τις επιθυμητές μετοχές χωρίς να επηρεάσουν την αγορά. Αντίθετα οι βραχυπρόθεσμοι επενδυτές και οι εταιρείες market makers επωφελούνται από την αυτόματη εκτέλεση συναλλαγών και την παροχή ρευστότητας στην αγορά. Επίσης, συστημικοί επενδυτές όπως τα hedge funds αξιοποιούν τις αλγοριθμικές συναλλαγές προκειμένου να κάνουν αποτελεσματικότερο προγραμματισμό των θέσεών τους και αυτές να εκτελούνται αυτοματοποιημένα. Μία από τις πρώτες προσπάθειες για την δημιουργία αυτοματοποιημένων στρατηγικών επενδύσεων οι οποίες αξιοποιούν ανεκμετάλλευτα πρότυπα στις αγορές έγινε από τον μαθηματικό Jim Simons ο οποίος ίδρυσε την εταιρεία επενδύσεων «Renaissance Technologies» της οποίας το κύριο fund ονομάζεται «Medallion», το 1988. Μέχρι σήμερα το Medallion έχει αποφέρει κέρδη πάνω από \$100 δισεκατομμύρια και έχει μέση ετήσια απόδοση πάνω από 66%. Ο Jim Simons θεωρείται ένας από τους πιο επιτυχημένους επενδυτές στην ιστορία των χρηματιστηριακών αγορών και αποκαλείται ως ο άνθρωπος που έλυσε την αγορά (ο τίτλος της βιογραφίας του από τον δημοσιογράφο Gregory Zuckerman της Wall Street Journal είναι “The man who solved the market”).

Μερικές από τις πιο διαδεδομένες στρατηγικές Αλγοριθμικών συναλλαγών είναι:

- Εύκαιρες Arbitrage
- Ακολουθία Τάσης (Trend Following)
- Στρατηγικές Βασισμένες σε μαθηματικά μοντέλα
- Στρατηγικές Αντιστροφής Μέσου Όρου
- Στρατηγικές βασισμένες στην «ορμή» μίας τιμής κατά την αύξηση ή την μείωσή της

Με την πάροδο του χρόνου και την εξέλιξη της τεχνολογίας το ποσοστό των αγοροπωλησιών στην αγορά που υλοποιούνται αλγοριθμικά αυξάνεται. Στο παρακάτω σχήμα φαίνεται το ποσοστό αυτό για διάφορα χρηματιστηριακά προϊόντα από το 2004 μέχρι και το 2016.

Market Share of Algorithmic Trading by Asset Class

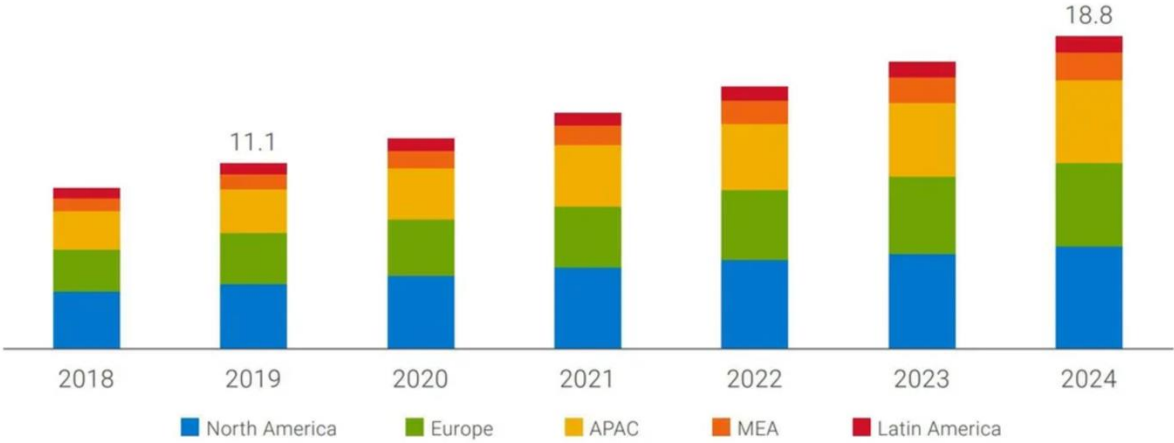


As of 2017
Source: Goldman Sachs, Aite Group

Σχήμα 2: Ποσοστό των αγοροπωλησιών που εκτελούνται αλγοριθμικά στις χρηματιστηριακές αγορές

Ο όγκος των αγοροπωλησιών που γίνονται αλγοριθμικά στις αγορές αναμένεται να αυξηθεί ακόμα περισσότερο στο μέλλον. Συγκεκριμένα, οι αναλυτές εκτιμούν ότι η αγορά των αλγοριθμικών συναλλαγών θα αυξηθεί από \$11.1 δισεκατομμύρια το 2019 σε \$18.8 δισεκατομμύρια το 2024.

Algorithmic Trading Market, by Region (USD Billion)



Source: MarketsandMarkets Analysis

Σχήμα 3: Μέγεθος της Αγοράς των Αλγοριθμικών Συναλλαγών

Κεφάλαιο 2: Νευρωνικά Δίκτυα

2.1 Εισαγωγή

Το 1911 ο Ισπανός νευροεπιστήμονας Santiago Ramón y Cajal ανέπτυξε την θεωρία ότι οι λειτουργίες του ανθρώπινου εγκεφάλου εκτελούνται από υπολογιστικές μονάδες που ονομάζονται νευρώνες. Πειράματα που έγιναν στη φυσιολογία των νευρικών κυττάρων έδειξαν ότι η λειτουργία τους μπορεί, σε απλοποιημένη μορφή, να προσομοιωθεί με μη γραμμικούς τελεστές που δέχονται σήματα από έναν πεπερασμένο αριθμό εισόδων και διαθέτουν μονάχα μία έξοδο. Η εγκεφαλική ουσία των έμβιων όντων αποτελείται από εκατομμύρια διασυνδεδεμένους νευρώνες οι οποίοι αποτελούν ένα πολύπλοκο και μη γραμμικό νευρωνικό δίκτυο. Οι νευρώνες του δικτύου, μέσω ηλεκτρικών παλμών, μεταδίδουν πληροφορίες μεταξύ των διαφορετικών περιοχών του εγκεφάλου και μεταξύ του εγκεφάλου και του νευρικού συστήματος.

Έχουν παρατηρηθεί ποσοτικές αλλά και ποιοτικές διαφοροποιήσεις στο είδος των συνδέσεων, το μηχανισμό επεξεργασίας σημάτων και το πλήθος εισόδων των νευρικών κυττάρων. Για αυτό το λόγο τα νευρικά κύτταρα έχουν ομαδοποιηθεί σε ένα μικρό αριθμό κατηγοριών. Έτσι, μπορούμε να δούμε τον ανθρώπινο εγκέφαλο σαν ένα σύνθετο παράλληλο και μη γραμμικό υπολογιστή, ο οποίος εξαιτίας της μορφολογίας του, μπορεί να ανταποκριθεί σε πολλές περιπτώσεις με πολλαπλάσια ταχύτητα από έναν ψηφιακό υπολογιστή. Η παρατήρηση αυτή ώθησε τους ερευνητές στην διαδικασία κατασκευής ενός υπολογιστικού δικτύου εμπνευσμένο από αυτό του εγκεφάλου.

Το επιστημονικό πεδίο των «Τεχνητών Νευρωνικών Δικτύων» βασίστηκε στην παρατήρηση ότι ο ανθρώπινος εγκέφαλος μπορεί να εκτελέσει ταχύτατα χιλιάδες υπολογισμούς, γεγονός που επιτρέπει στους ανθρώπους να υλοποιούν, σε καθημερινή βάση, πολύ σύνθετες διεργασίες. Το 1943 οι Walter Pitts και Warren McCulloch κατασκεύασαν ένα μοντέλο υπολογιστή βασισμένο στα νευρωνικά δίκτυα του εγκεφάλου.

Στη συνέχεια του 20^{ου} αιώνα έγιναν εξαιρετικά σημαντικά βήματα στην εξέλιξη των Τεχνητών Νευρωνικών Δικτύων όπως η εκπαίδευση κατά Hebbian (Hebbian Learning) και ο αλγόριθμος οπισθοδρομικής διάδοσης του σφάλματος. Από το 1990 και μετά τα νευρωνικά δίκτυα αναπτύσσονται ραγδαία και χρησιμοποιούνται σε πλήθος εφαρμογών.

Στο κεφάλαιο αυτό θα αναλυθούν τα τεχνητά νευρωνικά δίκτυα ως προς τη δομή τα χαρακτηριστικά και τη λειτουργία τους. Θα δοθεί περισσότερη έμφαση στην περιγραφή και ανάλυση των δικτύων που χρησιμοποιήθηκαν στα πλαίσια της εργασίας.

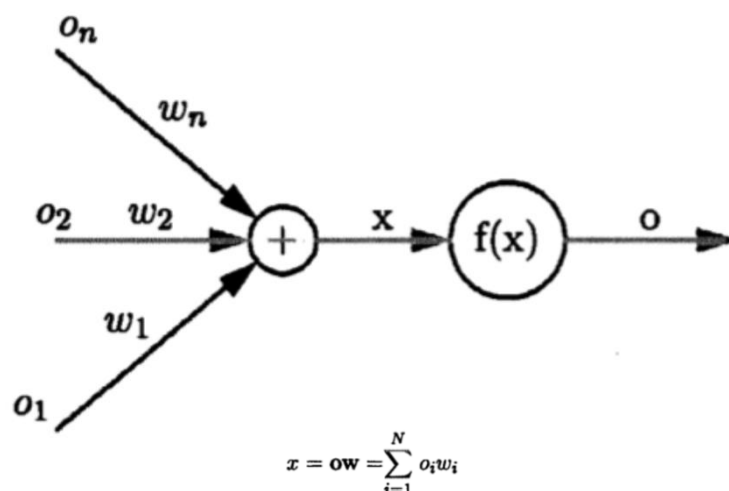
2.2 Μοντέλο Τεχνητού Νευρώνα

Κάθε νευρωνικό δίκτυο αποτελείται από διασυνδεδεμένες υπολογιστικές μονάδες που ονομάζονται νευρώνες. Οι τεχνητοί νευρώνες είναι το κύριο συστατικό των τεχνητών νευρωνικών δικτύων και σκοπός τους είναι να μετασχηματίσουν την είσοδο που δέχονται, με συγκεκριμένο τρόπο, στην έξοδο.

Όπως απεικονίζεται στο σχήμα 1, ένας τεχνητός νευρώνας δέχεται ένα σύνολο τιμών o_1, o_2, \dots, o_n , οι οποίες μπορεί να είναι συνεχείς μεταβλητές. Το σύνολο των τιμών o_1, o_2, \dots, o_n είναι το διάνυσμα εισόδου. Κάθε τιμή εισόδου αντιστοιχίζεται με κάποιο βάρος w_i (weight), η τιμή του οποίου μπορεί να είναι είτε θετική είτε αρνητική. Το σώμα του τεχνητού νευρώνα αποτελείται από δύο μέρη:

- τον αθροιστή (sum) ο οποίος προσθέτει τα γινόμενα των τιμών εισόδου με τα αντίστοιχα βάρη και παράγει την ποσότητα: $x = o * w = \sum_{i=1}^N o_i * w_i$
- τη συνάρτηση ενεργοποίησης ή κατωφλίου (activation, threshold or transformation function). Η συνάρτηση λειτουργεί ως ένα μη γραμμικό φίλτρο το οποίο διαμορφώνει την τιμή της εξόδου του νευρώνα σε συνάρτηση με την ποσότητα x . Έτσι η έξοδος ή οι έξοδοι του νευρώνα θα είναι $f(x)$.

Στην παρακάτω εικόνα παρουσιάζεται το μοντέλο ενός τεχνητού νευρώνα.



Σχήμα 4: Μοντέλο Τεχνητού Νευρώνα

Η μη γραμμική συνάρτηση ενεργοποίησης ακολουθεί τα χαρακτηριστικά των σιγμοειδών συναρτήσεων:

1. Είναι αύξουσα συνάρτηση

$$\forall x_1, x_2 \in \mathfrak{R}, \quad x_1 > x_2 \Rightarrow f(x_1) \geq f(x_2)$$

Σχέση 3.1

2. Έχει πεπερασμένα απειροστικά όρια*

$$\begin{aligned} \lim_{x \rightarrow \infty} f(x) &= a, & a \in \mathbb{R} - \{-\infty, +\infty\} \\ \lim_{x \rightarrow -\infty} f(x) &= b, & b \in \mathbb{R} - \{-\infty, +\infty\} \end{aligned}$$

Σχέση 3.2

3. Έχει πεδίο ορισμού το σύνολο των πραγματικών αριθμών και έχει φραγμένο πεδίο τιμών. Η ιδιότητα αυτή προκύπτει από τα δύο προηγούμενα χαρακτηριστικά

$$f: \mathbb{R} \rightarrow [a, b]$$

Σχέση 3.3

Μερικές από τις πιο διαδεδομένες συναρτήσεις ενεργοποίησης είναι οι ακόλουθες:

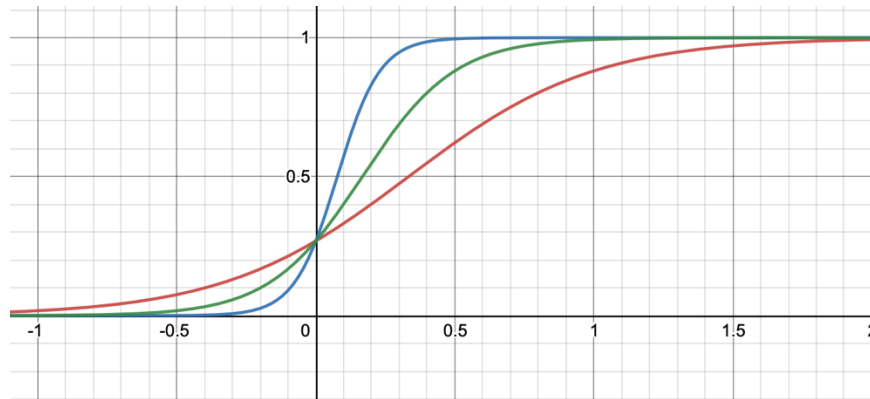
- Εκθετική Σιγμοειδής:

$$f(x) = \frac{1}{e^{-a*x + b} + c}, \quad a, b \in \mathbb{R}^+$$

Σχέση 3.4

*Σημείωση: Η ιδιότητα ισχύει για τις κλασσικές συναρτήσεις ενεργοποίησης. Για μερικές από τις συναρτήσεις που χρησιμοποιούνται στην πράξη όπως η relu δεν ισχύει

όπου a είναι ένας συντελεστής ρύθμισης της ταχύτητας μετάβασης μεταξύ δύο ασύμπτωτων τιμών. Στην παρακάτω εικόνα φαίνεται το διάγραμμα της εκθετικής σιγμοειδούς για διαφορετικές τιμές του a . Η καμπύλη που συγκλίνει πιο γρήγορα στην τιμή 1 (μπλε καμπύλη) έχει το μεγαλύτερο a , η κόκκινη καμπύλη έχει το μικρότερο a και η τιμή του a της πράσινης βρίσκεται μεταξύ των δύο.

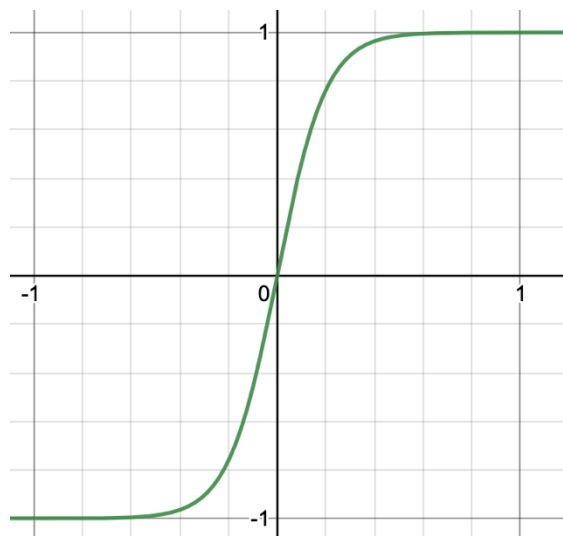


Σχήμα 5: Καμπύλη Εκθετικής Σιγμοειδούς για τρεις διαφορετικές τιμές του a . Με μπλε χρώμα απεικονίζεται η καμπύλη με το μεγαλύτερο a , με κόκκινο η καμπύλη με το μικρότερο a . Για την πράσινη καμπύλη το a είναι μεταξύ αυτών της μπλε και της κόκκινης.

- Υπερβολική Εφαπτομένη:

$$f(x) = \frac{e^{a*x} - e^{-a*x}}{e^{a*x} + e^{-a*x}}, \quad a \in \mathcal{R}^+$$

Σχέση 3.5



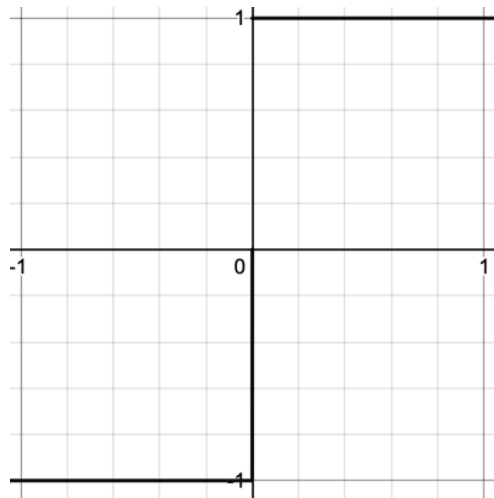
Σχήμα 6: Καμπύλη Υπερβολικής Εφαπτομένης

Η διαδικασία υπολογισμού εκθετικών όρων είναι σε γενικές γραμμές μία χρονοβόρα διαδικασία. Έτσι, η εκθετική σιγμοειδής αλλά κυρίως η υπερβολική εφαπτομένη, σε πολλές περιπτώσεις αποφεύγονται.

- Η συνάρτηση προσήμου:

$$f(x) = \text{sign}(x)$$

Σχέση 3.6

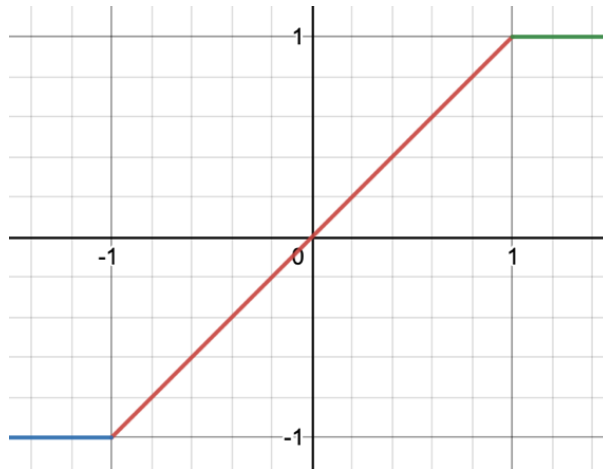


Σχήμα 7: Καμπύλη Συνάρτησης Προσήμου

- Τμηματικά γραμμική συνάρτηση:

$$f(x) = \begin{cases} \frac{1}{a} * x, & -a \leq x \leq a \\ -1, & x < -a \\ 1, & x > a \end{cases} \quad a \in \mathbb{R}^+$$

Σχέση 3.7

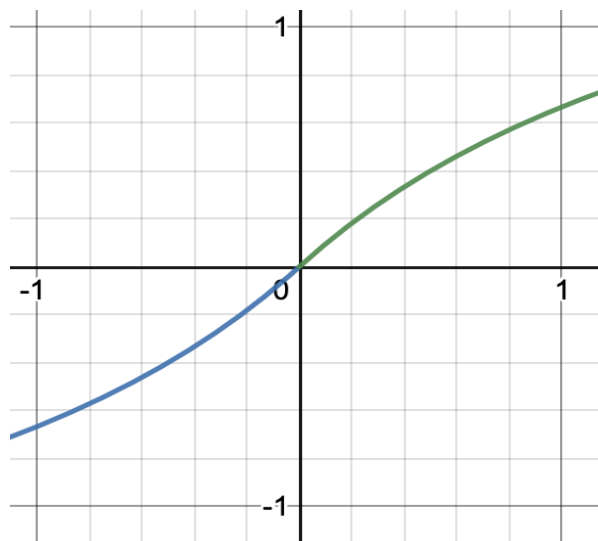


Σχήμα 8: Καμπύλη Τμηματικά Γραμμικής Συνάρτησης

- Συνάρτηση δύο κλάδων:

$$f(x) = \begin{cases} \frac{a * x}{x + b}, & x \geq 0 \\ \frac{-a * x}{x - b}, & x < 0 \end{cases} \quad a, b \in \mathbb{R}^+$$

Σχέσης 3.7

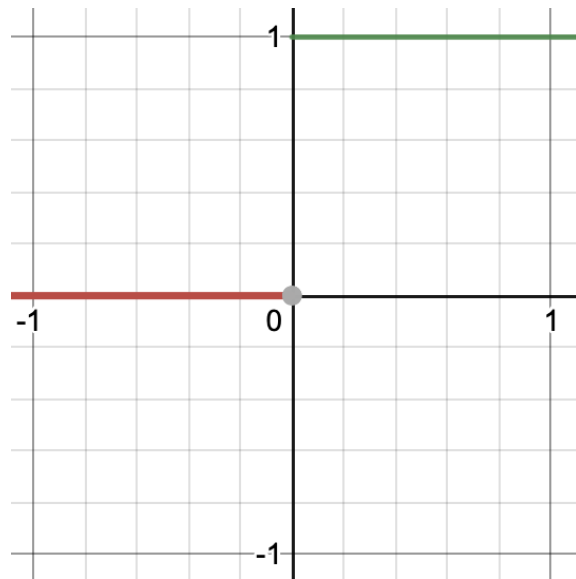


Σχήμα 9: Καμπύλη Συνάρτησης Δύο Κλάδων

- Βηματική συνάρτηση:

$$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

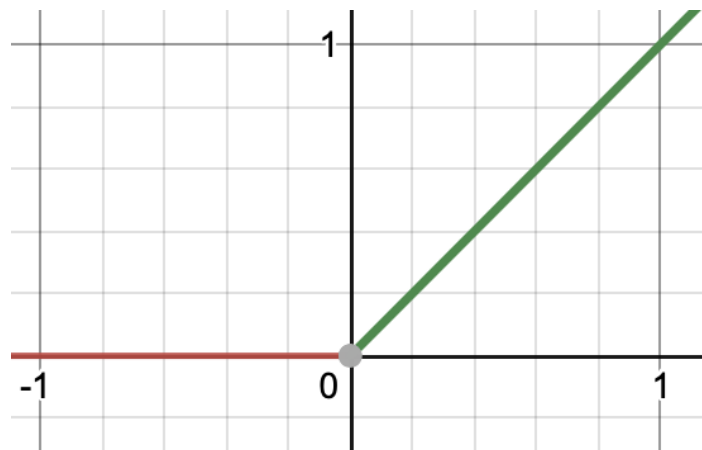
Σχέση 3.8



Σχήμα 10: Καμπύλη Βηματικής Συνάρτησης

- Συνάρτηση Rectified Linear Units ή αλλιώς Relu: $f(x) = \max(0, x)$

Σχέση 3.9

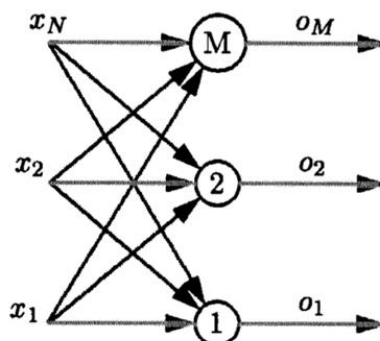


Σχήμα 11: Καμπύλη Συνάρτησης Relu

2.3 Δομή Τεχνητών Νευρωνικών Δικτύων

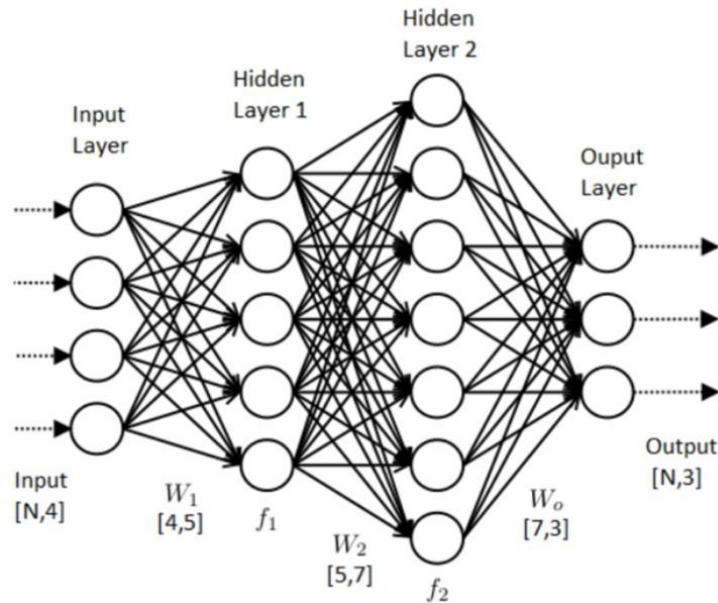
Τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) αποτελούνται από ένα πλήθος νευρώνων οι οποίοι είναι οργανωμένοι σε δομές, όπως και ο ανθρώπινος εγκέφαλος. Οι νευρώνες των δικτύων οργανώνονται σε στρώματα ή επίπεδα (layers). Το πρώτο επίπεδο, ή αλλιώς επίπεδο εισόδου (input layer) χρησιμοποιείται για την εισαγωγή των δεδομένων στο δίκτυο. Τα στοιχεία από τα οποία αποτελείται δεν μπορούν να θεωρηθούν νευρώνες, εφόσον δεν έχουν βάρη εισόδου ούτε συναρτήσεις ενεργοποίησης και δεν πραγματοποιούν υπολογισμούς. Στο επόμενο επίπεδο, το οποίο ονομάζεται κρυφό επίπεδο, μπορούν να υπάρχουν κανένα, ένα ή και περισσότερα στρώματα νευρώνων (hidden layers). Επιπλέον, κάθε στρώμα στο κρυφό επίπεδο μπορεί να αποτελείται από οποιονδήποτε αριθμό νευρώνων. Τελευταίο επίπεδο του νευρωνικού δικτύου είναι το επίπεδο εξόδου. Δεν υπάρχει περιορισμός για τον τρόπο με τον οποίο πρέπει να οργανωθεί ένα νευρωνικό δίκτυο, αλλά υπάρχουν συγκεκριμένες κατηγορίες δομών οι οποίες έχουν μελετηθεί εκτενέστερα και χρησιμοποιούνται συχνότερα.

- I. Νευρωνικά Δίκτυα ενός επιπέδου. Είναι η απλούστερη περίπτωση οργάνωσης ενός ΤΝΔ. Κάθε νευρώνας δέχεται ως είσοδο τις εισόδους του δικτύου και η έξοδός του αποτελεί μία από τις εξόδους του δικτύου. Στο διάγραμμα 9 απεικονίζεται ένα ΤΝΔ ενός επιπέδου. Όπως φαίνεται κάθε ένας από τους νευρώνες δέχεται ως είσοδο τις εισόδους του δικτύου και παράγει μία έξοδο η οποία είναι και έξοδος όλου του δικτύου.



Σχήμα 12: Νευρωνικό Δίκτυο Ενός Επιπέδου

II. Πολυεπίπεδα νευρωνικά δίκτυα. Στην περίπτωση αυτή το ΤΝΔ αποτελείται από τουλάχιστον 3 επίπεδα. Το επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα και το επίπεδο εξόδου. Αξίζει να σημειωθεί πως με την προσθήκη ενός ή παραπάνω κρυφών επιπέδων οι βαθμοί ελευθερίας του ΤΝΔ μπορούν να αυξηθούν απεριόριστα. Επιπλέον, τα κρυφά επίπεδα δεν είναι απαραίτητο να περιέχουν τον ίδιο αριθμό νευρώνων. Στην εικόνα 10 φαίνεται ένα πολυεπίπεδο ΤΝΔ.

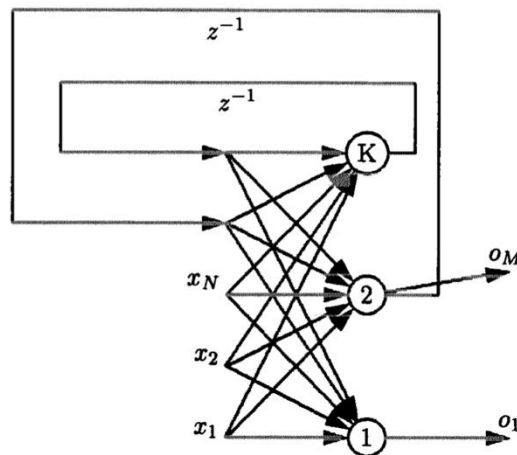


Σχήμα 13: Νευρωνικό Δίκτυο Πολλών Επιπέδου

Εκτός από τον αριθμό των νευρώνων ανά στρώμα, μπορεί να αλλάζει και ο τρόπος σύνδεσης των νευρώνων. Οι νευρώνες των δικτύων που αναλύθηκαν παραπάνω είναι πλήρως συνδεδεμένοι (fully connected) καθώς κάθε νευρώνας συνδέεται με όλους τους νευρώνες του επόμενου επιπέδου. Εάν κάποιοι από τους νευρώνες τους δικτύου δεν συνδέονται με όλους τους νευρώνες του επόμενου τότε είναι μερικώς συνδεδεμένοι (partially connected). Επιπλέον, στις εικόνες 9 και 10 δεν υπάρχουν συνδέσεις μεταξύ νευρώνων ενός επιπέδου και νευρώνων του προηγούμενου επιπέδου και έτσι η ροή της πληροφορίας έχει πρόσθια κατεύθυνση. Τα ΤΝΔ αυτά χαρακτηρίζονται ως feedforward δίκτυα.

Στην περίπτωση που υπάρχει έστω και μία διαδρομή μέσω της οποίας, ξεκινώντας από έναν νευρώνα και μέσω των συνάψεων, κατά τη φορά ενεργοποίησης του δικτύου, μπορούμε να επανέλθουμε στο νευρώνα εκκίνησης, τότε το νευρωνικό δίκτυο θα λέγεται επανατροφοδοτούμενο (recurrent). Ο τύπος του δικτύου με ανατροφοδότηση διαφέρει

από τον τύπο της απλούς τροφοδότησης στο ότι περιλαμβάνει βρόχους ανάδρασης όπου νευρώνες τροφοδοτούν το σήμα εξόδου τους στις εισόδους άλλων νευρώνων και έτσι το δίκτυο αποκτά ιδιαίτερα χαρακτηριστικά. Οι χρονικές καθυστερήσεις δίνουν τη δυνατότητα στο ΤΝΔ να αναγνωρίσει χρονικά μεταβαλλόμενα πρότυπα. Στην εικόνα 11 φαίνεται η τυπική μορφή ενός επανατροφοδοτούμενο δικτύου.



Σχήμα 14: : Επανατροφοδοτούμενο Νευρωνικό Δίκτυο

2.4 Εκπαίδευση Τεχνητών Νευρωνικών Δικτύων

Ένας ορισμός για την εκπαίδευση των νευρωνικών δικτύων δόθηκε από τους Mendel και McLaren: Η εκπαίδευση είναι η διαδικασία κατά την οποία αλλάζουν οι ελεύθερες μεταβλητές του νευρωνικού δικτύου μέσω μιας διαρκούς κατάστασης διέγερσης από το περιβάλλον. Το είδος της εκπαίδευσης εξαρτάται από τον τρόπο με τον οποίο γίνεται η αλλαγή των μεταβλητών.

Ανάλογα με τη μέθοδο εκτέλεσης τα ΤΝΔ ταξινομούνται σε δύο κατηγορίες:

- **Εκπαίδευση με Επίβλεψη (Supervised Learning):** Για ένα σύνολο δεδομένων είναι γνωστή η είσοδος του νευρωνικού δικτύου αλλά και η αναμενόμενη έξοδος. Το δίκτυο δέχεται τα δεδομένα αυτά στην είσοδο, τα επεξεργάζεται και στην συνέχεια παράγει την έξοδο. Η διαδικασία αυτή εκτελείται μέχρι το σφάλμα, δηλαδή η διαφορά μεταξύ της πραγματικής τιμής και της εξόδου του δικτύου, να είναι

χαμηλότερο από το μέγιστο επιτρεπτό. Η μέθοδος ελαχιστοποίησης του σφάλματος καθορίζεται από τον αλγόριθμο εκπαίδευσης ο οποίος μεταβάλλει τις ελεύθερες μεταβλητές του συστήματος. Στόχος των συγκεκριμένων δικτύων είναι ο μηδενισμός του σφάλματος μέσω καθορισμού των κατάλληλων βαρών.

- **Εκπαίδευση χωρίς επίβλεψη (Unsupervised Learning):** Για ένα σύνολο δεδομένων μεγάλου όγκου υπάρχουν τιμές διανυσμάτων εισόδου αλλά όχι τα διανύσματα εξόδου. Το δίκτυο δέχεται τα δεδομένα και πρέπει μέσω αλλαγών στα βάρη του να εντοπίσει τα κριτήρια ταξινόμησης των παραδειγμάτων. Ουσιαστικά ο αλγόριθμος εκπαίδευσης σε αυτόν τον τύπο δικτύου πραγματοποιεί τυχαίες αλλαγές στις τιμές των συνάψεων με σκοπό να προσεγγίσει τη βέλτιστη λύση.

Το βασικότερο συστατικό της εκπαίδευσης ενός τεχνητού νευρωνικού δικτύου είναι ο αλγόριθμος μάθησης (learning algorithms). Ο αλγόριθμος μάθησης επηρεάζει τον τρόπο με τον οποίο γίνεται η αλλαγή στα βάρη των συνάψεων του δικτύου. Υπάρχουν πολλοί και διαφορετικοί αλγόριθμοι, κάθε ένας με πλεονεκτήματα και μειονεκτήματα.

Στα πλαίσια της εργασίας θα αναλυθεί με λεπτομέρεια ο αλγόριθμος οπισθοδρομικής διάδοσης του σφάλματος (Back Propagation).

Αλγόριθμος Οπισθοδρομικής Διάδοσης του Σφάλματος (Back Propagation Algorithm)

Ο αλγόριθμος οπισθοδρομικής διάδοσης του σφάλματος είναι στην ουσία μία μέθοδος διόρθωσης του σφάλματος εξόδου.

Έστω ότι διαθέτουμε ένα σύστημα $f(x,g)$ όπου x είναι η είσοδος, g η είναι οι ελεύθερες μεταβλητές του συστήματος τις οποίες μπορούμε να αλλάξουμε και K το σύνολο των δεδομένων που διαθέτουμε. Το σύνολο K αποτελείται από παραδείγματα εισόδου-εξόδου. Αρχικά, το δίκτυο έχει τυχαίες τιμές στις ελεύθερες μεταβλητές του.

Τοποθετείται κάθε ένα από τα παραδείγματα στην είσοδο του συστήματος και υπολογίζεται η έξοδος του συστήματος και έτσι για ένα παράδειγμα εισόδου x_i όπου $i=1,2,\dots,M$ έχουμε την έξοδο του συστήματος $s(x_i, w)$ και την πραγματική του τιμή t_i , $i=1,2,\dots,M$. Το μέσο τετραγωνικό σφάλμα του συστήματος είναι:

$$\text{Σφάλμα} = \frac{1}{M} \sum_{i=1}^M (t_i - s(x_i, w))^T (t_i - s(x_i, w))$$

Σχέση 3.9

Σκοπός λοιπόν της μεθόδου διόρθωσης του σφάλματος είναι ο υπολογισμός των ελεύθερων παραμέτρων w έτσι ώστε το τετραγωνικό σφάλμα εξόδου να ελαχιστοποιείται. Αυτό το διάνυσμα g δίνεται από την επίλυση του συστήματος:

$$\frac{\partial \text{Σφάλμα}}{\partial w} = 0$$

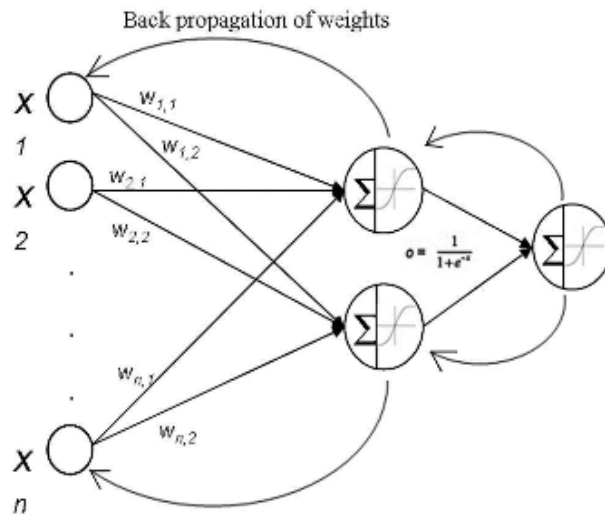
Σχέση 3.10

Αλγόριθμος

Ο αλγόριθμος οπισθοδρομικής διάδοσης του σφάλματος έχει ως βασική αρχή λειτουργίας την ελαχιστοποίηση του τετραγωνικού σφάλματος μεταξύ της επιθυμητής εξόδου και της εξόδου του τεχνητού νευρωνικού δικτύου. Αποτελείται από τα εξής βήματα:

1. **Πρόσθια διάδοση:** Ένα παράδειγμα από το σύνολο των δεδομένων εισάγεται στο δίκτυο, το οποίο έπειτα από υπολογισμούς παράγει την έξοδο.
2. **Αντίστροφη Διάδοση:** Υπολογίζεται η διαφορά μεταξύ των δύο εξόδων (του δικτύου και της επιθυμητής). Στη συνέχεια η τιμή του σφάλματος «ταξιδεύει» προς τα πίσω έτσι ώστε να αλλάξουν οι τιμές των βαρών και να προσεγγιστεί καλύτερα το επιθυμητό αποτέλεσμα.

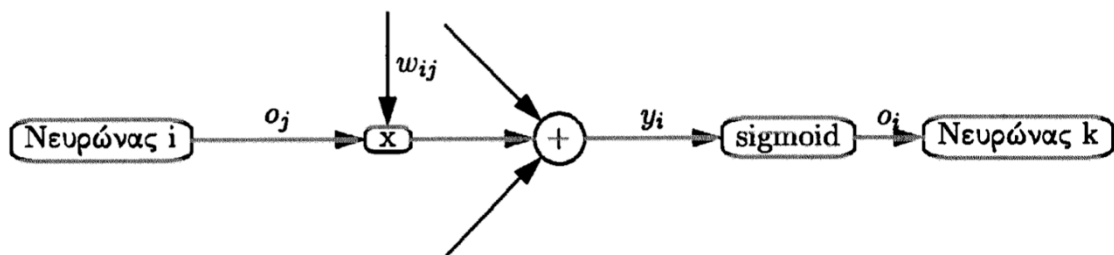
Στην εικόνα 12 φαίνεται η διάδοση της πληροφορίας στο δίκτυο.



Σχήμα 15: Οπισθοδρομική Διάδοση Σφάλματος

Ωστόσο, για να επεξηγηθεί με μεγαλύτερη λεπτομέρεια ο αλγόριθμος είναι απαραίτητο να εξεταστούν οι ενέργειες σε κάθε επίπεδο αλλά και το μαθηματικό υπόβαθρο. Θεωρούμε ότι το νευρωνικό δίκτυο στο οποίο εξετάζουμε τον αλγόριθμο είναι ένα πλήρως διασυνδεδεμένο δίκτυο πολλών επιπέδων.

Όπως απεικονίζεται στο σχήμα 13, έστω ένας νευρώνας ο οποίος συνδέεται με τον νευρώνα i μέσω της σύναψης o_j και του βάρους w_{ij} . Η εσωτερική κατάσταση του νευρώνα είναι y_i ενώ έξοδος του νευρώνα, ο οποίος τροφοδοτείται στο νευρώνα k .



Σχήμα 16: Μοντέλο Τεχνητού Νευρώνα

Το σφάλμα για μια τυχαία είσοδο a είναι το τετράγωνο των διαφορών των δύο εξόδων, της εκτιμώμενης o και της πραγματικής b :

$$\text{Σφάλμα} = \frac{1}{2}(b - o)^T(b - o) = \frac{1}{2} \sum_{i=1}^M (b_i - o_i)^2$$

Σχέση 3.11

όπου M είναι ο αριθμός νευρώνων του επιπέδου εξόδου. Η παραπάνω σχέση είναι το σφάλμα για ένα παράδειγμα από το σύνολο δεδομένων. Εάν έχουμε στη διάθεσή μας Q παραδείγματα τότε το συνολικό σφάλμα είναι:

$$E[\text{Σφάλμα}] = \frac{1}{2} E\left[\sum_{i=1}^M (b_i - o_i)^2\right] = \frac{1}{2} \sum_{i=1}^M E[(b_i - o_i)^2] = \frac{1}{2Q} \sum_{i=1}^M \sum_{j=1}^Q (b_{ij} - o_{ij})^2$$

Σχέση 3.12

Για τον υπολογισμό των συντελεστών βαρύτητας των συνάψεων καταφεύγουμε σε μία μέθοδο διαδοχικής προσέγγισης κατά τη διεύθυνση της αρνητικής τιμής της πρώτης παραγώγου του σφάλματος. Η αριθμητική τιμή των συντελεστών βαρύτητας υπολογίζεται με την ακόλουθη σχέση:

$$\Delta w_{ij} = -n \frac{\partial \text{Σφάλμα}}{\partial w_{ij}}$$

Σχέση 3.13

όπου n είναι ένας μικρός θετικός πραγματικός αριθμός που ονομάζεται ρυθμός εκπαίδευσης του αλγορίθμου.

Η σχέση X μπορεί να εκφραστεί σαν συνάρτηση της εξόδου και της εσωτερικής κατάστασης του νευρώνα y_i :

$$\Delta w_{ij} = -n \frac{\partial \text{Σφάλμα}}{\partial y_i} * \frac{\partial y_i}{\partial w_{ij}}$$

Σχέση 3.14

Αν ορίσουμε σαν δ_i την αρνητική τιμή της πρώτης παραγώγου του σφάλματος ως προς την εσωτερική κατάσταση του νευρώνα γι τότε:

$$\delta_i = - \frac{\partial \Sigma \text{φάσμα}}{\partial y_i} = - \frac{\partial \Sigma \text{φάσμα}}{\partial o_i} * \frac{\partial o_i}{\partial y_i}$$

Σχέση 3.15

Οπότε η σχέση 3.14 γίνεται:

$$\Delta w_{ij} = n * \delta_i * \frac{\partial y_i}{\partial w_{ij}} = n * \delta_i * \frac{\partial \sum_k w_{ij} * o_k}{\partial w_{ij}} = n * \delta_i * o_j$$

Σχέση 3.16

Για να γίνει ευκολότερη η αναλυτική αναπαράσταση της συνάρτησης επαναπροσδιορισμού των συντελεστών βαρύτητας των συνάψεων θα χωρίσουμε τους νευρώνες σε δύο κατηγορίες.

Νευρώνες Εξόδου

Για τους νευρώνες εξόδου η τιμή δ_i γίνεται:

$$\delta_i = - \frac{\partial \Sigma \text{φάσμα}}{\partial o_i} * \frac{\partial o_i}{\partial y_i} = -(-(b_i - o_i)) \frac{\partial f(y_i)}{\partial y_i}$$

Σχέση 3.17

Παρατηρούμε ότι ο δεύτερος όρος του πολλαπλασιασμού της σχέσης 16 είναι η πρώτη παράγωγος της συνάρτησης ενεργοποίησης του νευρώνα. Έτσι, η συνάρτηση επαναπροσδιορισμού των συντελεστών βαρύτητας των συνάψεων γίνεται:

$$\Delta w_{ij} = n * \delta_i * o_j = n * (b_i - o_i) o_j \frac{\partial f(y_i)}{\partial y_i}$$

Σχέση 3.18

Κρυφοί Νευρώνες

Έστω ένα επίπεδο κρυφών νευρώνων το οποίο συνδέεται στο αμέσως επόμενο επίπεδο με K νευρώνες εξόδου, τότε η αρνητική τιμή της παραγώγου του σφάλματος ως προς την εσωτερική κατάσταση γι γίνεται:

$$\delta_i = - \frac{\partial \Sigma \text{φάλαμα}}{\partial o_i} \frac{\partial o_i}{\partial y_i} = - \frac{\partial o_i}{\partial y_i} \sum_{k=1}^K \frac{\partial \Sigma \text{φάλαμα}}{\partial y_k} \frac{\partial y_k}{\partial o_i} = \frac{\partial o_i}{\partial y_i} \sum_{k=1}^K \left(- \frac{\partial \Sigma \text{φάλαμα}}{\partial y_k} \right) \frac{\partial \sum_{m=1}^I w_{km} o_m}{\partial o_i}$$

Σχέση 3.19

όπου I είναι ο αριθμός των νευρώνων του κρυφού επιπέδου που μελετάμε. Απλοποιώντας την σχέση 19 προκύπτει η σχέση:

$$\delta_i = \frac{\partial o_i}{\partial y_i} \sum_{k=1}^K \delta_k w_{ki}$$

Σχέση 3.20

Έτσι, ο όρος δ κάθε νευρώνα στο επίπεδο I εξαρτάται από τους όρους δ των νευρώνων που βρίσκονται στο αμέσως επόμενο επίπεδο και τους συντελεστές βαρύτητας των συνάψεων που συνδέουν την έξοδο του νευρώνα. Το συμπέρασμα είναι λογικό εάν αναλογιστούμε πως στον αλγόριθμο οπισθοδρομικής διάδοσης του σφάλματος αφού υπολογιστεί το σφάλμα της εξόδου του δικτύου σε σχέση με την επιθυμητή έξοδο διαδίδεται αντίστροφα ξεκινώντας από τους νευρώνες εξόδου.

Η συνάρτηση επαναπροσδιορισμού των συντελεστών βαρύτητας των συνάψεων για τους κρυφούς νευρώνες γίνεται:

$$\Delta w_{ij} = n * \delta_i * o_j = n * o_j * \frac{\partial o_i}{\partial y_i} \sum_{k=1}^K \delta_k w_{ki}$$

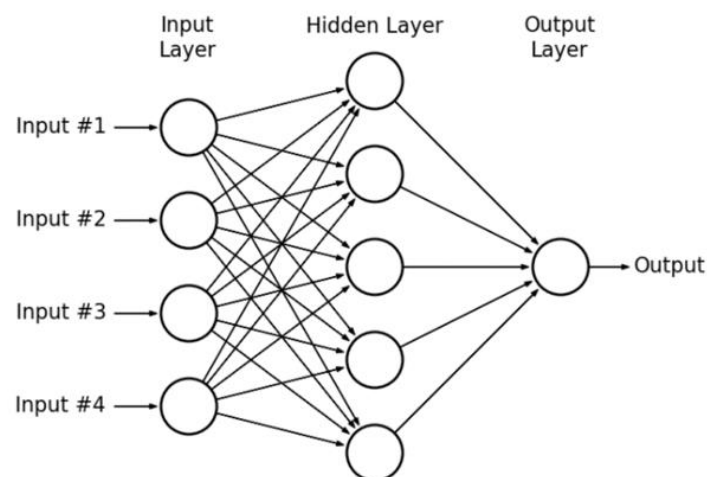
Σχέση 3.21

Όπως φαίνεται η διαδικασία υπολογισμού των συντελεστών βαρύτητας των συνάψεων είναι αναδρομική διαδικασία η οποία ξεκινά από τα υψηλότερα και συνεχίζει στα χαμηλότερα επίπεδα. Έτσι, ο αλγόριθμος οπισθοδρομικής διάδοσης του σφάλματος για ένα πλήρως συνδεδεμένο δίκτυο το οποίο διαθέτει τουλάχιστον ένα κρυφό επίπεδο ακολουθεί τα εξής βήματα:

- 1. Τοποθέτηση Αρχικών Τιμών:** Αφού οριστεί ο ρυθμός εκπαίδευσης η , τοποθετούνται τυχαίες αρχικές τιμές στους συντελεστές βαρύτητας των συνάψεων.
- 2. Ενεργοποίηση του Δικτύου:** Στην είσοδο του δικτύου τοποθετείται το διάνυσμα εισόδου ενός παραδείγματος από το σύνολο των δεδομένων και στη συνέχεια υπολογίζεται από το χαμηλότερο προς το υψηλότερο επίπεδο η έξοδος των νευρώνων κάθε επιπέδου.
- 3. Υπολογισμός των συντελεστών δ :** Αντιστρέφοντας την φορά των υπολογισμών υπολογίζονται οι συντελεστές δ από τα υψηλότερα στα χαμηλότερα επίπεδα.
- 4. Επαναπροσδιορισμός των συντελεστών βαρύτητας των συνάψεων:** Αφού έχουν υπολογιστεί οι συντελεστές δ και οι έξοδοι των κόμβων μπορούμε να επαναπροσδιορίσουμε τους συντελεστές βαρύτητας των συνάψεων όλων των νευρώνων του δικτύου.
- 5. Έλεγχος σύγκλισης:** Υπολογίζεται ο ρυθμός μεταβολής του σφάλματος στα παραδείγματα και η συνολική απόλυτη μεταβολή των συντελεστών βαρύτητας. Αν ο ρυθμός μεταβολής είναι μικρότερος από ένα κατώτατο όριο τότε ο αλγόριθμος τερματίζει. Επίσης, η συνολική μεταβολή των βαρών συγκρίνεται με ένα μικρό θετικό αριθμό και εάν είναι μικρότερη ο αλγόριθμος τερματίζει.
- 6. Επανάληψη του αλγορίθμου:** Τοποθετείται ένα νέο παράδειγμα στην είσοδο του δικτύου και επαναλαμβάνεται ο αλγόριθμος από το δεύτερο βήμα.

2.5 Multi-Layer Perceptron (MLP)

Ο multi-layer perceptron είναι ένας τύπος δικτύου πολλών επιπέδων, με πλήρως συνδεδεμένους νευρώνες. Ανήκει στην κατηγορία των δικτύων με επιβλεπόμενη μάθηση και μαθαίνει μία συνάρτηση $f(\cdot) : R^m \rightarrow R^o$, όπου m είναι ο αριθμός των διαστάσεων της εισόδου και o είναι οι διαστάσεις εξόδου, μέσα από την εκπαίδευση στο σύνολο των δεδομένων. Έτσι, για ένα σύνολο παραδειγμάτων $X = x_1, x_2, \dots, x_m$ και γνωστές εξόδους y μπορεί να προσεγγίσει και μη γραμμικές συναρτήσεις είτε σε προβλήματα πρόβλεψης τιμής (λειτουργία ως regressor) είτε σε προβλήματα κατηγοριοποίησης (λειτουργία ως classifier). Τα MLP δίκτυα μπορούν να έχουν ένα ή περισσότερα στρώματα κρυφών νευρώνων. Ένα δίκτυο MLP απεικονίζεται στο σχήμα 14.



Σχήμα 17: Μοντέλο Multi-Layer Perceptron

Τα MLPs χρησιμοποιούνται συχνά και έχουν τα εξής πλεονεκτήματα και μειονεκτήματα:

Πλεονεκτήματα: α) Έχουν την δυνατότητα να προσαρμόζονται σε μη γραμμικές συναρτήσεις
β) Μπορούν να προσαρμόζονται σε μοντέλα σε πραγματικό χρόνο

Μειονεκτήματα: α) Τυχαίες αρχικοποιήσεις των συναπτικών βαρών μπορεί να οδηγήσουν σε αποτελέσματα διαφορετικής ακρίβειας

β) Είναι απαραίτητος ο καθορισμός των υπερπαραμετρών όπως ο αριθμός των κρυφών νευρώνων, στρωμάτων και επαναλήψεων του αλγορίθμου μάθησης.

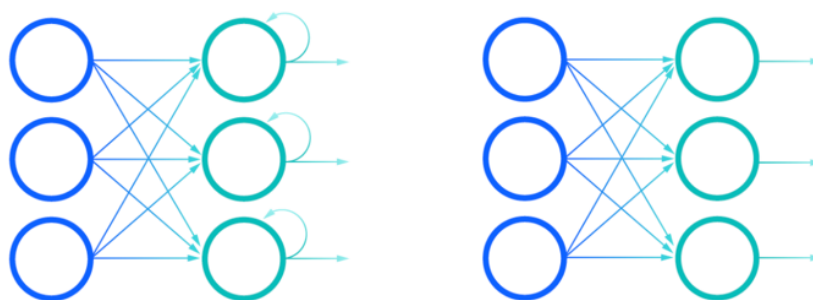
γ) Επηρεάζονται από την κλιμάκωση (scaling) των παραμέτρων.

2.6 Recurrent Neural Networks (RNNs)

2.6.1 Εισαγωγή

Τα Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks) είναι μία κατηγορία νευρωνικών δικτύων η οποία χρησιμοποιείται συχνά σε προβλήματα με ακολουθιακά δεδομένα και χρονοσειρές. Μερικά από τα προβλήματα στα οποία χρησιμοποιούνται είναι η μετάφραση από μία γλώσσα σε άλλη, η φυσική επεξεργασία γλώσσας (natural language processing) και η αναγνώριση λόγου (speech recognition) και έτσι διαδραματίζουν πολύ σημαντικό ρόλο σε εφαρμογές όπως το Siri της Apple και το Google Translate.

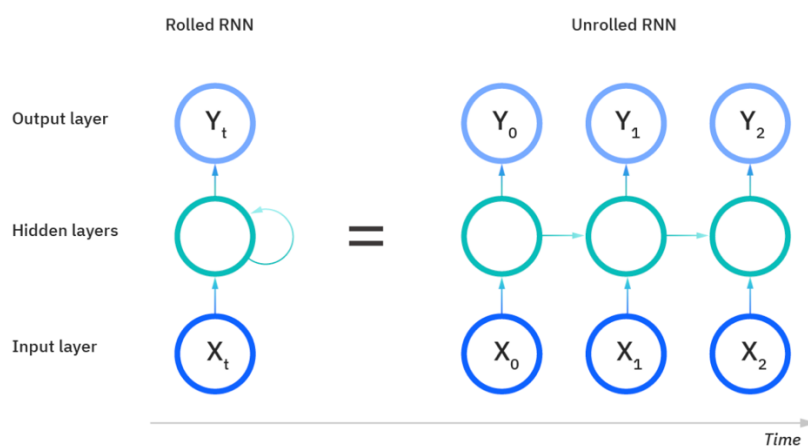
Ανήκουν στην κατηγορία των δικτύων μάθησης με επίβλεψη. Παρόλο που τα «παραδοσιακά» νευρωνικά δίκτυα που εξετάστηκαν παραπάνω θεωρούν πως οι εισόδοι είναι ανεξάρτητες από τις εξόδους, τα Αναδρομικά Νευρωνικά Δίκτυα ξεχωρίζουν λόγω της δυνατότητάς του να έχουν «μνήμη», αφού αξιοποιούν πληροφορίες από παλαιότερες εισόδους για να επηρεάσουν την τρέχουσα είσοδο και έξοδο. Στο σχήμα 15 απεικονίζεται ένα αναδρομικό δίκτυο (RNN) σε σχέση με ένα δίκτυο πρόσθιας διάδοσης (Feedforward Network).



Σχήμα 18: Μοντέλο Recurrent Neural Network

2.6.2 Παράδειγμα Χρήσης

Ας θεωρηθεί ως παράδειγμα η φράση «είναι καλός ο καιρός» η οποία θα μπορούσε να βρίσκεται εντός ενός συνόλου δεδομένων γραπτού ή προφορικού λόγου, έτσι ώστε να επεξηγηθεί αναλυτικά η λειτουργία των αναδρομικών δικτύων. Για να βγάλει νόημα το παραπάνω σύνολο λέξεων θα πρέπει η κάθε λέξη να βρίσκεται σε συγκεκριμένη θέση. Επομένως, το δίκτυο θα πρέπει να λάβει υπόψιν του τη θέση της κάθε λέξης και να τη χρησιμοποιήσει ώστε να προβλέψει την επόμενη λέξη.



Σχήμα 19: Μοντέλο Recurrent Neural Network

Όπως φαίνεται στην εικόνα 16 το δίκτυο που αναπαρίσταται στα αριστερά είναι μία απεικόνιση όλου του δικτύου και όλης της φράσης «είναι καλός ο καιρός». Το σχήμα στα δεξιά αναπαριστά το δίκτυο και την φράση με μεγαλύτερη λεπτομέρεια, αφού απεικονίζει όλα τα στρώματα ή αλλιώς χρονικά βήματα του δικτύου. Κάθε στρώμα χαρτογραφεί μία λέξη της φράσης, και έτσι 1^ο στρώμμα -> είναι, 2^ο στρώμμα -> καλός κ.ο.κ. Οι προηγούμενες εισοδοί όπως «είναι» και «καλός» βρίσκονται στο κρυφό επίπεδο του 3^{ου} στρώματος το οποίο προβλέπει τη λέξη «ο».

Ένα επιπλέον χαρακτηριστικό διαφοροποίησης των αναδρομικών δικτύων είναι το ότι έχουν κοινές παραμέτρους μεταξύ διαφορετικών στρωμάτων του δικτύου. Σε αντίθεση με τα δίκτυα πρόσθιας διάδοσης τα οποία έχουν διαφορετικά βάρη σε κάθε κόμβο, τα αναδρομικά δίκτυα «μοιράζονται» τις ίδιες παραμέτρους βάρους για κάθε στρώμα. Τα βάρη αυτά

διαμορφώνονται, όπως και στα απλά δίκτυα, από τη διαδικασία της διόρθωσης σφάλματος και τον αλγόριθμο οπισθοδρομικής διάδοσης έτσι ώστε να μπορεί να εκπαιδευτεί το δίκτυο.

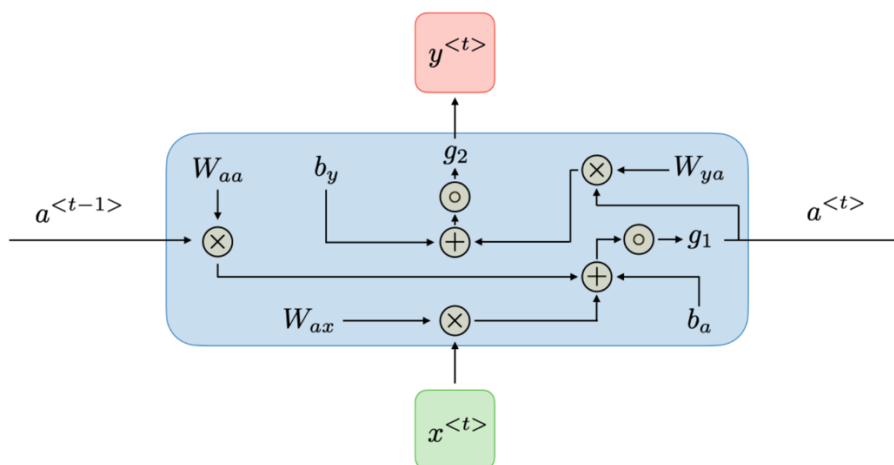
Στην παρακάτω εικόνα 17 φαίνεται το εσωτερικό ενός κρυφού στρώματος, ενός χρονικού βήματος για ένα αναδρομικό δίκτυο. Για κάθε χρονικό βήμα t η ενεργοποίηση του νευρώνα $a^{<t>}$ και $y^{<t>}$ εκφράζονται ως εξής:

$$\alpha^{<t>} = g_1(W_{aa}a^{<t-1>} + W_{ax}x^{<t>} + b_a)$$

$$y^{<t>} = g_2(W_{ya}\alpha^{<t>} + b_y)$$

Σχέση 3.22

όπου W_{ax} , W_{aa} , W_{ya} , b_a , b_y είναι σταθερές που «μοιράζονται» χρονικά και g_1 , g_2 συναρτήσεις ενεργοποίησης.



Σχήμα 20: Αρχιτεκτονική Recurrent Neural Network

Για τα αναδρομικά δίκτυα η συνάρτηση σφάλματος (Loss Function) L αντιπροσωπεύεται από τη σχέση:

$$\mathcal{L}(\hat{y}, y) = \sum_{t=1}^{T_y} \mathcal{L}(\hat{y}^{<t>}, y^{<t>})$$

Σχέση 3.23

όπου T_y είναι ο αριθμός των χρονικών βημάτων.

2.6.3 Εκπαίδευση Αναδρομικών Δικτύων

Τα αναδρομικά δίκτυα αξιοποιούν τον αλγόριθμο οπισθοδρομικής διάδοσης του σφάλματος μέσω του χρόνου, διαφοροποιώντας τα από τα δίκτυα πρόσθιας μετάδοσης. Η βασική ιδέα του αλγορίθμου είναι η ίδια εφόσον υπάρχει και πάλι εκπαίδευση του δικτύου μέσω της διάδοσης του σφάλματος από τους νευρώνες εξόδου στους νευρώνες εισόδου. Η διαφορά εντοπίζεται στο γεγονός ότι για τα αναδρομικά δίκτυα ο αλγόριθμος αθροίζει το σφάλμα για κάθε χρονικό βήμα (η ενέργεια αυτή είναι απαραίτητη εφόσον τα αναδρομικά δίκτυα «μοιράζονται» παραμέτρους μεταξύ των ίδιων στρωμάτων), ενώ στην απλή έκδοση τα σφάλματα δεν αθροίζονται και χρησιμοποιείται το μέσο τετραγωνικό.

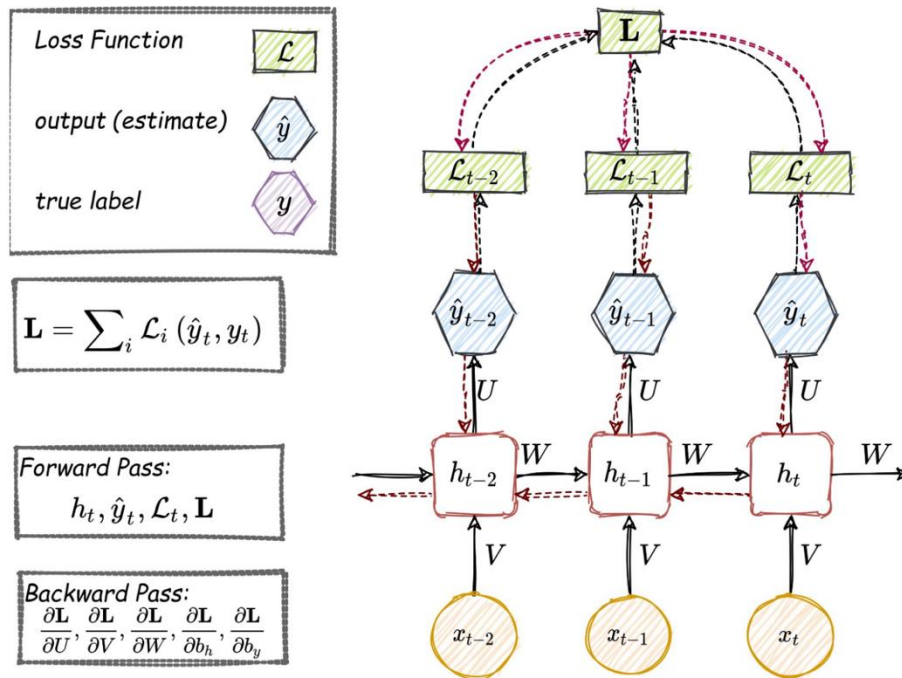
Για την αναλυτικότερη επεξήγηση του αλγορίθμου οπισθοδρομικής διάδοσης μέσω του χρόνου θα θεωρήσουμε ένα αναδρομικό νευρωνικό δίκτυο με τρία χρονικά βήματα. Το νευρωνικό δίκτυο αυτό δέχεται εισόδους x_{t-2} , x_{t-1} , x και παράγει εξόδους y_{t-1} , y_{t-2} και y_t . Η κρυφή κατάσταση κάθε νευρώνα σημειώνεται ως h_{t-2} , h_{t-1} , h_t και υπάρχουν 3 πίνακες για τιμές των συνάψεων των βαρών που «μοιράζονται» μεταξύ των στρωμάτων οι U , W και V . Επιπλέον, υπάρχουν δύο συναρτήσεις ενεργοποίησης οι f_x και f_y . Η συνάρτηση σφάλματος ορίζεται όπως και στη σχέση 3.11.

Για κάθε ένα χρονικό βήμα ακολουθείται πρώτα η εξής διαδικασία πρόσθιας διάδοσης:

- I. Το παράδειγμα τοποθετείται στην είσοδο του δικτύου
- II. Υπολογίζεται η εσωτερική κρυφή κατάσταση $h(t)$ και η έξοδος $y(t)$

- III. Μέσω της συνάρτησης σφάλματος υπολογίζεται η απόκλιση μεταξύ της εξόδου και της αναμενόμενης τιμής
- IV. Υπολογίζεται το συνολικό σφάλμα

Η διαδικασία πρόσθιας διάδοσης απεικονίζεται στο επόμενο σχήμα με τα μαύρα βέλη. Μετά το πέρας αυτής ακολουθεί η οπισθοδρομική διάδοση η οποία φαίνεται με κόκκινα βέλη.



Σχήμα 21: Σχήμα Εκπαίδευσης Αναδρομικών Δικτύων

Όπως και στη σχέση 3.13 για τον απλό αλγόριθμο οπισθοδρομικής διάδοσης έτσι και εδώ θα εξεταστεί η μερική παράγωγος της συνάρτησης του σφάλματος ως προς το βάρος W .

Η διαδικασία εκπαίδευσης του αναδρομικού δικτύου όμως δεν είναι τόσο απλή αφού γίνεται οπισθοδρομική διάδοση του σφάλματος όχι μόνο μέσα από τα στρώματα αλλά και από το χρόνο. Επομένως, για κάθε χρονικό βήμα θα πρέπει να αθροιστούν όλα τα προηγούμενα βήματα μέχρι το τωρινό όπως φαίνεται στην σχέση 3.24:

$$\frac{\partial L}{\partial W} = \sum_{i=0}^T \frac{\partial \mathcal{L}_i}{\partial W} \propto \sum_{i=0}^T \left(\prod_{i=0}^y \frac{\partial h_i}{\partial h_{i-1}} \right) \frac{\partial h_k}{\partial W}$$

Σχέση 3.24

όπου $i=0,1,2\dots T$ είναι τα χρονικά βήματα h_t είναι οι εσωτερικές καταστάσεις, L είναι η συνάρτηση σφάλματος και W είναι ο πίνακας με τα βάρη.

Μέσω αυτής της διαδικασίας εκπαίδευσης τα αναδρομικά δίκτυα ενδέχεται να αντιμετωπίσουν δύο προβλήματα. Το πρώτο είναι γνωστό ως το exploding gradient και το δεύτερο είναι γνωστό ως vanishing gradient. Τα προβλήματα αυτά προκύπτουν από το μέγεθος της κλίσης (gradient) της συνάρτησης σφάλματος ως προς την καμπύλη σφάλματος, δηλαδή εξαρτώνται από το ρυθμό μεταβολής της συνάρτησης σφάλματος ως προς το σφάλμα. Οι μαθηματικές εκφράσεις των δύο αυτών προβλημάτων φαίνονται στη σχέση 25.

Το vanishing gradient συναντάται όταν η κλίση είναι πολύ μικρή, συνεχίζει να μειώνεται αλλάζοντας τα βάρη, μέχρι αυτά να έχουν αμελητέες τιμές (κοντά στο 0). Όταν συμβαίνει αυτό το δίκτυο δεν μπορεί να «μάθει» καθώς τα βάρη δεν επηρεάζονται.

Το exploding gradient πρόβλημα συμβαίνει παρατηρείται όταν η κλίση είναι πολύ μεγάλη και έτσι το μοντέλο γίνεται ασταθές. Σε αυτήν την περίπτωση τα βάρη γίνονται πολύ μεγάλοι αριθμοί και από ένα σημείο και μετά φαίνονται ως NaN τιμές.

$$\begin{aligned} \text{Vanishing gradient} & \quad \left\| \frac{\partial h_i}{\partial h_{i-1}} \right\|_2 < 1 \\ \text{Exploding gradient} & \quad \left\| \frac{\partial h_i}{\partial h_{i-1}} \right\|_2 > 1 \end{aligned}$$

Σχέση 3.25

Τα αναδρομικά νευρωνικά δίκτυα έχουν τα εξής πλεονεκτήματα και μειονεκτήματα:

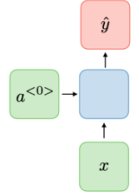
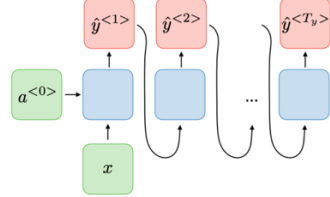
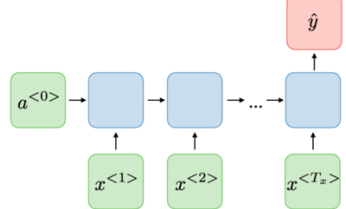
Πλεονεκτήματα:

- α) Δυνατότητα επεξεργασίας εισόδων ανεξάρτητα από το μήκος τους
- β) Το μέγεθος του δικτύου χρειάζεται να αυξηθεί όταν μεγαλώνει το μέγεθος της εισόδου
- γ) Οι υπολογισμοί συμπεριλαμβάνουν πληροφορίες από το παρελθόν
- δ) Τα βάρη «μοιράζονται» στο χρόνο,

Μειονεκτήματα:

- α) Οι υπολογισμοί των αναδρομικών δικτύων απαιτούν πολύ χρόνο
- β) Υπάρχει δυσκολία στην επεξεργασία πολύ παλαιών πληροφοριών
- γ) Δεν μπορούν να συνυπολογίσουν μελλοντικά γεγονότα στην τρέχουσα κατάσταση

Στον παρακάτω πίνακα διακρίνονται οι βασικότερες κατηγορίες αναδρομικών νευρωνικών δικτύων και ένα παράδειγμα χρήσης τους.

Τύπος RNN	Απεικόνιση	Εφαρμογή
Ένα προς ένα $T_x = T_y = 1$		Κλασικός τύπος RNN. Οι χρήσεις του συζητήθηκαν παραπάνω.
Ένα προς πολλά $T_x = 1, T_y > 1$		Παραγωγή Μουσικής
Πολλά προς ένα $T_x > 1, T_y = 1$		Κατηγοριοποίηση Λόγου/ Κειμένου

Πολλά προς Πολλά $T_x = T_y$		Αναγνώριση Ονόματος Οντοτήτων
Πολλά προς Πολλά $T_x \neq T_y$		Μετάφραση

Πίνακας 1: Τύποι RNN

Τα αναδρομικά δίκτυα χρησιμοποιούν τις ίδιες συναρτήσεις ενεργοποίησης με τα απλά δίκτυα, όπως συζητήθηκε παραπάνω.

2.7 Παράμετροι που επηρεάζουν την απόδοση των Νευρωνικών Δικτύων

Η ακρίβεια των προβλέψεων που γίνονται από τα νευρωνικά δίκτυα εξαρτώνται σε μεγάλο βαθμό από τη διαμόρφωση του δικτύου. Τα μορφολογικά χαρακτηριστικά ενός δικτύου τα οποία επηρεάζουν την ποιότητα των προβλέψεων είναι τα ακόλουθα:

1. Συναρτήσεις Ενεργοποίησης Νευρώνων: Οι συναρτήσεις Ενεργοποίησης των Νευρώνων ορίζουν την έξοδο του νευρώνα με βάση την εσωτερική του κατάσταση. Οι συναρτήσεις αυτές επιλέγονται κατά τον σχεδιασμό του δικτύου και δεν μπορούν να αλλάξουν όταν το δίκτυο τεθεί σε λειτουργία. Μερικές από τις διαθέσιμες συναρτήσεις ενεργοποίησης αναλύονται στην αρχή του κεφαλαίου.
2. Είδος και Αρχιτεκτονική του Δικτύου: Η επιλογή του κατάλληλου είδους δικτύου είναι απαραίτητη για την εξαγωγή καλύτερων αποτελεσμάτων. Αμέσως μετά την επιλογή του είδους του δικτύου ορίζεται και η αρχιτεκτονική του. Σε κάθε νευρωνικό δίκτυο οι νευρώνες εξόδου και εισόδου καθορίζονται από τον αριθμό των διαστάσεων των αντίστοιχων διανυσμάτων και έτσι είναι το πλήθος τους είναι συγκεκριμένο για κάθε

πρόβλημα. Ο αριθμός των κρυφών επιπέδων αλλά και το πλήθος των νευρώνων σε κάθε επίπεδο μπορούν να μεταβληθούν και επηρεάζουν σημαντικά την ποιότητα των προβλέψεων.

3. Συνάρτηση Μάθησης: Η συνάρτηση μάθησης είναι ένα από τα κυριότερα χαρακτηριστικά του δικτύου και καθορίζει τον τρόπο επίλυσης του προβλήματος της βελτιστοποίησης των βαρών των συνάψεων.
4. Ρυθμός Μάθησης: Ο ρυθμός μάθησης είναι, όπως αναφέρθηκε παραπάνω ο παράγοντας η στη σχέση 16 και χρησιμοποιείται για να μετριάσει ή για να αυξήσει τις αλλαγές που γίνονται στα βάρη μετά από κάθε επανάληψη του αλγορίθμου εκπαίδευσης. Συγκεκριμένα, καθορίζει το μέγεθος των μεταβολών στα βάρη των συνάψεων και κατ' επέκταση επηρεάζει το ρυθμό σύγκλισης. Ένα δίκτυο με πολύ μικρό ρυθμό μάθησης πραγματοποιεί πολύ μικρές αλλαγές στα βάρη και έτσι παρατηρείται σημαντική αύξηση του χρόνου σύγκλισης. Αντίθετα, πολύ μεγάλος ρυθμός μάθησης έχει ως αποτέλεσμα ταλαντώσεις στα βάρη και συνεπάγεται αργή σύγκλιση.
5. Αριθμός Εποχών: Ως εποχή ορίζεται μία πλήρης επανάληψη του αλγορίθμου εκπαίδευσης κατά την οποία το νευρωνικό δίκτυο έχει «δει» όλο το σύνολο δεδομένων. Συνήθως η εκπαίδευση των δικτύων απαιτεί αρκετές εποχές. Ο αριθμός των εποχών είναι μία από τις παραμέτρους που επηρεάζουν την απόδοση του δικτύου. Εάν ο αριθμός των εποχών είναι μικρός τότε η ακρίβεια του δικτύου θα είναι μικρή διότι το δίκτυο δεν προλαβαίνει να εκπαιδευτεί όσο θα έπρεπε. Αντίθετα, πολύ μεγάλος αριθμός εποχών έχει ως αποτέλεσμα την υπερ-εκπαίδευση του δικτύου (*overfitting*) στο σύνολο δεδομένων εκπαίδευσης. Ο βέλτιστος αριθμός εποχών είναι εκείνος για τον οποίο το δίκτυο παρουσιάζει τη μέγιστη δυνατή ακρίβεια πρόβλεψης στο σύνολο δεδομένων (*testing data set*).

Η διαδικασία της εκπαίδευσης ενός δικτύου θα αλλάξει τα βάρη των συνάψεων των νευρώνων, αλλά δεν θα επηρεάσει τα μορφολογικά χαρακτηριστικά του δικτύου, ή αλλιώς τις υπερ-παραμέτρους του δικτύου. Παρόλα αυτά τα μορφολογικά χαρακτηριστικά ενός

δικτύου που αναλύθηκαν παραπάνω επηρεάζουν την απόδοση του δικτύου. Όμως, δεν είναι εφικτό να γνωρίζει κανείς εκ των προτέρων την βέλτιστη δομή του δικτύου για κάθε πρόβλημα.

Η προφανής λύση για την εύρεση της βέλτιστης δομής ενός νευρωνικού δικτύου είναι η επανάληψη των προβλέψεων με πολλές και διαφορετικές προσεγγίσεις, ωστόσο η προσέγγιση αυτή απαιτεί πολύ χρόνο και πόρους. Για αυτό το λόγο υπάρχουν διάφορες προσεγγίσεις για την αναζήτηση των βέλτιστων υπερ-παραμέτρων.

Τυχαία Αναζήτηση (Random Search): Η τυχαία αναζήτηση, ψάχνει το πλαίσιο των δυνατών τιμών των υπερ-παραμέτρων με τυχαία σειρά. Συγκεκριμένα, αφού έχουν οριστεί κάποιες δυνατές τιμές για τα χαρακτηριστικά του δικτύου, εξετάζει τυχαίους συνδυασμούς διαφορετικών τιμών των παραμέτρων.

Εξελικτική Βελτιστοποίηση (Evolutionary Optimization): Η μέθοδος αυτή χρησιμοποιεί εξελικτικούς αλγορίθμους για να πραγματοποιήσει αναζητήσεις στο χώρο των υπερ-παραμέτρων.

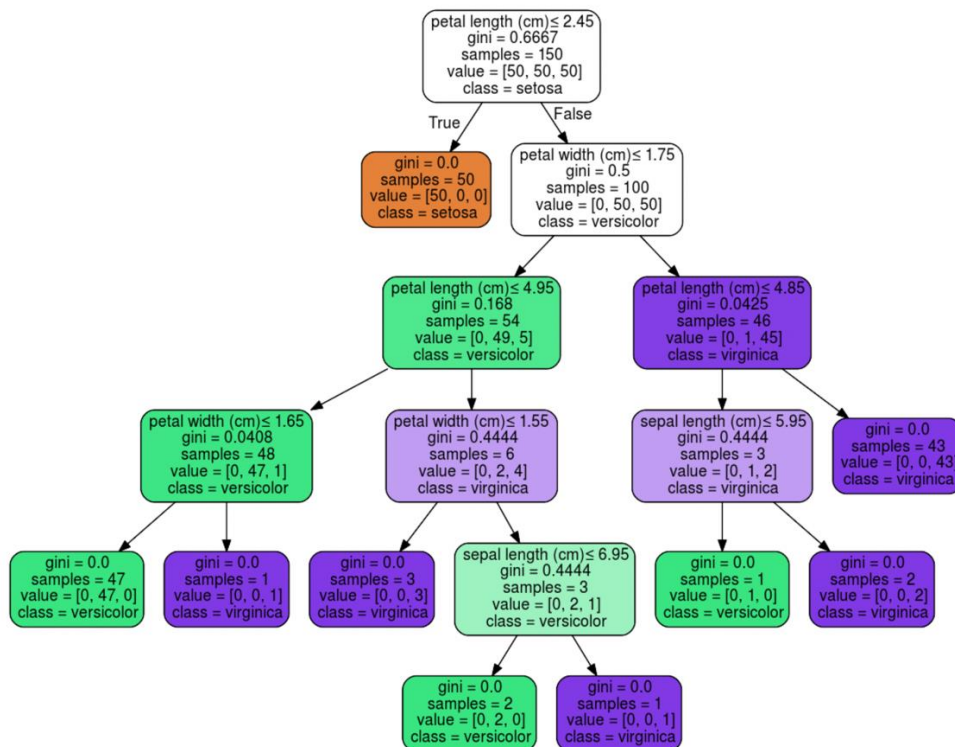
Αναζήτηση Πλαισίου (Grid Search): Το Grid Search αποτελεί μία εξαντλητική μέθοδο αναζήτησης μέσα στο σύνολο των δυνατών υπερ-παραμέτρων. Συγκεκριμένα ο σχεδιαστής του δικτύου φτιάχνει ένα σύνολο τιμών για κάθε μία από τις παραμέτρους, τοποθετώντας σε αυτό τιμές που θεωρεί ότι είναι πιθανό να αποτελούν τις βέλτιστες. Κάθε συνδυασμός των τιμών των παραμέτρων αυτών αποτελούν ένα πιθανό δίκτυο και έτσι το από το σύνολο των συνδυασμών των παραμέτρων δημιουργείται ένα σύνολο δικτύων. Στη συνέχεια κάθε μοντέλο εκπαιδεύεται και αξιολογείται.

Κεφάλαιο 3: Δέντρα Απόφασης

3.1 Εισαγωγή

Τα Δέντρα Αποφάσεων (Decision Trees) είναι μια μία μη παραμετρική μέθοδος επιβλεπόμενης μάθησης η οποία χρησιμοποιείται για προβλήματα κατηγοριοποίησης και πρόβλεψης. Στόχος του μοντέλου είναι να προβλέψει την τιμή μίας μεταβλητής του προβλήματος και το επιτυγχάνει «μαθαίνοντας» απλούς κανόνες για το σύνολο δεδομένων μέσω της διαδικασίας της μάθησης. Τα δέντρα απόφασης χρησιμοποιούνται τόσο για προβλήματα κατηγοριοποίησης όσο και για προβλήματα πρόβλεψης. Κάθε δέντρο αποτελείται από κόμβους και ακμές. Οι ακμές είναι βέλη μονής κατεύθυνσης, με διεύθυνση από την παλαιότερη στη νεότερη γενιά.

Στην παρακάτω εικόνα φαίνεται το δέντρο απόφασης που δημιουργείται για την κατηγοριοποίηση ενός φυτού από το σύνολο δεδομένων “Iris” με βάση τα χαρακτηριστικά (features) που περιέχει.



Σχήμα 22: Παράδειγμα Δέντρου Απόφασης

Η βασική ιδέα των δέντρων απόφασης είναι να εντοπίσουν τα χαρακτηριστικά που περιέχουν τις περισσότερες πληροφορίες σε σχέση με το προβλεπόμενο χαρακτηριστικό και να διαχωρίσουν το σύνολο δεδομένων. Ο διαχωρισμός θα πρέπει να γίνει με τέτοιο τρόπο ώστε το προβλεπόμενο χαρακτηριστικό να αντιπροσωπεύεται στους τελικούς κόμβους του δέντρου.

Το χαρακτηριστικό μέσω του οποίου μπορεί να διαχωριστεί το σύνολο δεδομένων καλύτερα ως προς το προβλεπόμενο χαρακτηριστικό, θεωρείται το βέλτιστο για το συγκριμένο βήμα διαχωρισμού. Η διαδικασία αναζήτησης για το βέλτιστο χαρακτηριστικό είναι συνεχής και πραγματοποιείται μέχρι και τους τελευταίους κόμβους του δέντρου.

Η κατασκευή ενός δέντρου απόφασης είναι μία διαδικασία που αποτελείται από πολλά “if statements”. Συγκεκριμένα σε κάθε παράδειγμα του συνόλου δεδομένων γίνονται ερωτήσεις με βάση τις τιμές που διαθέτει στα διάφορα χαρακτηριστικά του. Η ερώτηση που διαχωρίζει καλύτερα τα παραδείγματα επιλέγεται και το χαρακτηριστικό στο οποίο αυτή απευθύνεται είναι το χαρακτηριστικό του κόμβου.

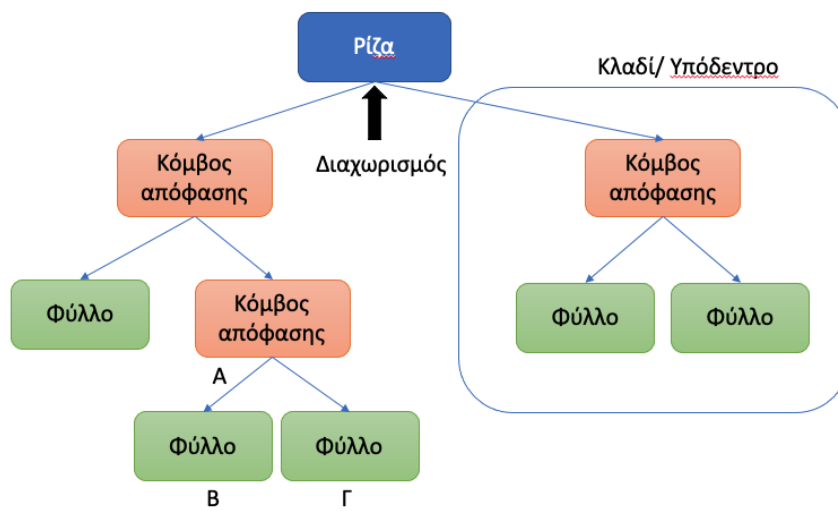
3.2 Κατασκευή

Προκειμένου να γίνει κατανοητή η διαδικασία κατασκευής των δέντρων απόφασης είναι απαραίτητο να δοθούν οι παρακάτω ορισμοί:

1. Ρίζα (Root Node): Ο πρώτος κόμβος αποκαλείται ρίζα του δέντρου.
2. Κλαδί ή Υπόδεντρο (Branch or Subtree): Ένα μέρος του δέντρου στο οποίο γίνεται μία απόφαση
3. Διαχωρισμός (Splitting): Ο διαχωρισμός ενός κόμβου σε δύο ή περισσότερους κόμβους επόμενης γενιάς
4. Φύλλο ή Τερματικός Κόμβος (Leaf or Terminal Node): Οι τελευταίοι κόμβοι του δέντρου απόφασης που δεν μπορούν να διαχωριστούν σε κόμβους επόμενης γενιάς.

5. Κόμβος Απόφασης (Decision Node): Ο τελευταίος κόμβος πριν από τα φύλλα του δέντρου. Ο κόμβος απόφασης μπορεί να διαχωρίζεται σε φύλλα ή και σε άλλους κόμβους.

Στο γράφημα 20 απεικονίζεται ένα τυπικό δέντρο απόφασης με τα στοιχεία που δόθηκαν στους παραπάνω ορισμούς. Ο κόμβος A λέγεται γονέας των κόμβων B και Γ.



Σχήμα 23: Μορφολογικά Χαρακτηριστικά Δέντρου Απόφασης

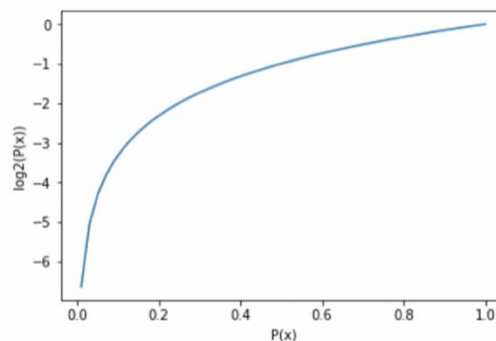
Κάθε κόμβος του δέντρου απόφασης αντιπροσωπεύει ένα χαρακτηριστικό του συνόλου δεδομένων. Για τον διαχωρισμό το μοντέλο διατυπώνει ένα “if” ερώτημα με βάση το οποίο διαχωρίζει τα στοιχεία του συνόλου. Για παράδειγμα όπως φαίνεται στη ρίζα του δέντρου του σχήματος 19, εάν το χαρακτηριστικό “petal length” έχει τιμή ≤ 2.45 τότε το παράδειγμα ανήκει στην κατηγορία “setosa”. Ο διαχωρισμός λειτουργεί με αντίστοιχο τρόπο και για μη συνεχείς μεταβλητές. Η διαδικασία επιλογής του χαρακτηριστικού που θα καθορίσει τον διαχωρισμό ενός κόμβου δεν είναι τυχαία και βασίζεται στις επόμενες μετρικές.

A. Εντροπία Πληροφορίας (Entropy):

Η εντροπία χρησιμοποιείται για την μέτρηση της τυχαιότητας και της καθαρότητας (purity) του συνόλου δεδομένων. Για παράδειγμα έστω ότι διαλέγουμε ένα κίτρινο αντικείμενο από

ένα κουτί το οποίο περιέχει μόνο κίτρινα αντικείμενα. Τότε, το κουτί έχει εντροπία ίση με το 0.

Αν υποθέσουμε ότι 50 από τα κίτρινα αντικείμενα αντικαθίστανται από 30 κόκκινα και 20 μπλε. Τότε, εάν επιλέξουμε τυχαία ένα αντικείμενο από το κουτί η πιθανότητα αυτό να είναι κίτρινο έχει μειωθεί από 1.0 (πιθανότητα πριν την πρόσθεση άλλων αντικειμένων) σε 0.5. Η καθαρότητα (purity) του συνόλου δεδομένων έχει μειωθεί και η εντροπία έχει αυξηθεί. Το μοντέλο της εντροπίας του Shannon χρησιμοποιεί τη λογαριθμική συνάρτηση $\log_2(P(x))$ για να μετρήσει την εντροπία καθώς όσο η πιθανότητα να επιλεγθεί ένα κίτρινο αντικείμενο $P(x)$ αυξάνεται τότε το αποτέλεσμα της εντροπίας πλησιάζει το 0 ($P(x)=1 \Rightarrow \log_2(P(x))=0$) όπως φαίνεται στο παρακάτω διάγραμμα.



Σχήμα 24: Διάγραμμα Συνάρτησης $\log(P(x))$

Όταν ένα χαρακτηριστικό περιέχει περισσότερα από ένα στοιχεία (για παράδειγμα αντικείμενα πολλών χρωμάτων και όχι μόνο κίτρινου), τότε είναι βοηθητικό η εντροπία να πολλαπλασιάζεται με την πιθανότητα να του κάθε ενδεχομένου (εντροπία να επιλεγεί κίτρινο αντικείμενο * πιθανότητα να επιλεγεί κίτρινο αντικείμενο) και τελικά να αθροίζονται. Έτσι, οδηγούμαστε στην πλήρη έκφραση της εντροπίας του Shannon η οποία φαίνεται στη σχέση X. Η εντροπία του Shannon αποτελεί τη βάση για τον υπολογισμό του κέρδους πληροφορίας.

$$Entropy(x) = - \sum (P(x = k) * \log_2(P(x = k)))$$

Σχέση 4.1

όπου $P(x=k)$ είναι η πιθανότητα το χαρακτηριστικό να λάβει την συγκεκριμένη τιμή k . Το αρνητικό πρόσημο οφείλεται στο γεγονός ότι οι λογαριθμικές συναρτήσεις έχουν αρνητικές τιμές, ενώ η εντροπία πρέπει να είναι θετική.

Η μέγιστη τιμή της εντροπίας ενός χαρακτηριστικού ορίζεται από τον αριθμό των κατηγοριών/τιμών που έχει το χαρακτηριστικό αυτό. Για παράδειγμα εάν στο κουτί υπάρχουν 2 χρώματα η μέγιστη τιμή της εντροπίας είναι 1 ενώ εάν υπάρχουν 16 χρώματα η μέγιστη τιμή της εντροπίας είναι 16.

B. Κέρδος Πληροφορίας (Information Gain):

**Κέρδος Πληροφορίας(ΧαρακτηριστικούA) = Εντροπία(Συνόλου Δεδομένων) –
Εντροπία(ΧαρακτηριστικούA)**

Αρχικά υπολογίζεται η εντροπία του συνόλου των δεδομένων. Στη συνέχεια για κάθε χαρακτηριστικό υλοποιείται ένας διαχωρισμός και υπολογίζεται η εντροπία του. Έτσι με την παραπάνω σχέση υπολογίζεται το κέρδος πληροφορίας για κάθε χαρακτηριστικό.

Το χαρακτηριστικό με το μεγαλύτερο κέρδος πληροφορίας επιλέγεται σαν ρίζα του δέντρου.

Ο αλγόριθμος ID2 χρησιμοποιεί το κέρδος πληροφορίας για την κατασκευή του δέντρου απόφασης.

Γ. Δείκτης Gini (Gini Index):

Ο δείκτης gini υπολογίζεται αφαιρώντας το άθροισμα των τετραγωνισμένων πιθανοτήτων κάθε κατηγορίας/τιμής από το 1. Η μαθηματική σχέση του δείκτη είναι η ακόλουθη:

$$Gini Index = 1 - \sum (P(x = k))^2$$

Σχέση 4.2

Το χαρακτηριστικό με τη μικρότερη τιμή Gini επιλέγεται ως η ρίζα του δέντρου.

Ο αλγόριθμος CART βασίζεται στο δείκτη Gini για την κατασκευή του δέντρου απόφασης.

Η δημιουργία του δέντρου, λοιπόν εξαρτάται από την επιλογή του αλγορίθμου κατασκευής του.

Το χαρακτηριστικό με το μεγαλύτερο κέρδος πληροφορίας ή την μικρότερη τιμή του δείκτη Gini επιλέγεται ως ρίζα του δέντρου και υλοποιείται ο πρώτος διαχωρισμός. Στη συνέχεια επαναλαμβάνεται η αντίστοιχη διαδικασία για όλους τους επόμενους διαχωρισμούς μέχρι τα φύλλα του δέντρου.

Αξίζει ωστόσο να διερευνηθεί πότε πρέπει να σταματήσει ο διαχωρισμός των κόμβων για ένα δέντρο απόφασης. Όταν ένα σύνολο δεδομένων έχει πολλά χαρακτηριστικά τότε είναι απαραίτητο να γίνουν πολλοί διαχωρισμοί κατά την κατασκευή και τη διαδικασία εκπαίδευσης. Αυτό όμως μπορεί να έχει ως αποτέλεσμα ένα σύνθετο δέντρο το οποίο είναι πλήρως προσαρμοσμένο στο σύνολο δεδομένων εκπαίδευσης (overfitting). Ένας τρόπος να αποφευχθεί το πρόβλημα αυτό είναι ο ορισμός του ελάχιστου αριθμού παραδειγμάτων εκπαίδευσης σε κάθε φύλλο του δέντρου. Έτσι, θα αγνοούνται τα φύλλα που περιέχουν λιγότερα παραδείγματα από το ελάχιστο όριο και το δέντρο θα μπορεί να πραγματοποιήσει πιο γενικές προβλέψεις κατά τον έλεγχο (testing). Ένας άλλος τρόπος για την επίλυση του προβλήματος είναι ο ορισμός του μέγιστου βάθους του δέντρου (maximum depth). Ως βάθος ορίζεται η απόσταση από τη ρίζα μέχρι το κάθε φύλλο και ο ορισμός του μέγιστου βάθους συνεπάγεται πως από ένα σημείο και κάτω το δέντρο δεν μπορεί να επεκταθεί περισσότερο.

Η αποδοτικότητα των δέντρων απόφασης μπορεί επιπλέον να βελτιωθεί μέσω της διαδικασίας κλαδέματος (pruning). Με το κλάδεμα αφαιρούνται τα κλαδιά του δέντρου τα οποία αφορούν χαρακτηριστικά μικρής σημασίας. Έτσι, μειώνεται η πολυπλοκότητα του δέντρου και πιθανότητα να προσαρμοστεί στο σύνολο δεδομένων εκπαίδευσης (overfitting) και αυξάνεται η δυνατότητα πρόβλεψης.

Το κλάδεμα μπορεί να ξεκινήσει από τη ρίζα ή από τα φύλλα. Η πιο απλή μέθοδος κλαδέματος ξεκινά από τα φύλλα και αφαιρεί τον κόμβο με την δημοφιλέστερη κατηγορία/τιμή. Στη συνέχεια ελέγχεται εάν η αλλαγή αυτή επηρεάζει την ακρίβεια πρόβλεψης και εάν αυτή δεν μειώνεται τότε η αλλαγή διατηρείται. Ο τρόπος κλαδέματος αυτός αποκαλείται κλάδεμα μειωμένου σφάλματος (reduced error pruning). Μία περισσότερο περίπλοκη μέθοδος κλαδέματος είναι το κλάδεμα του πιο αδύναμου κρίκου (weakest link pruning). Στον τρόπο αυτό χρησιμοποιείται μία μάθησης παράμετρος α η οποία

επηρεάζεται από το μέγεθος του υπόδεντρου και καθορίζει εάν ένα κλαδί μπορεί να αφαιρεθεί.

3.3 Παράδειγμα Χρήσης

Στον παρακάτω πίνακα φαίνεται ένα σύνολο δεδομένων. Συγκεκριμένα το σύνολο περιέχει 14 παραδείγματα και 5 χαρακτηριστικά. Θα εξεταστεί ο τρόπος κατασκευής ενός δέντρου απόφασης ώστε αυτό να προβλέπει την τελευταία κατηγορία με βάση τα υπόλοιπα χαρακτηριστικά. Για την κατασκευή του δέντρου θα χρησιμοποιηθεί το κέρδος πληροφορίας. Υποθέτοντας ότι τα χαρακτηριστικά των παραδειγμάτων μπορούν να έχουν δύο διαφορετικές τιμές, επομένως η εντροπία μπορεί να γραφτεί ως εξής:

$$\text{Εντροπία} = -\frac{a}{a+b} * \log_2\left(\frac{a}{a+b}\right) - \frac{b}{a+b} \log_2\left(\frac{b}{a+b}\right)$$

Σχέση 4.3

όπου a, b ο αριθμός των παραδειγμάτων που έχουν κάθε μία από τις πιθανές τιμές του χαρακτηριστικού που πρέπει να προβλεφθεί, a+b το σύνολο των παραδειγμάτων.

#	outlook	temperature	humidity	windy	play
1	sunny	hot	high	no	no
2	sunny	hot	high	yes	no
3	overcast	hot	high	yes	yes
4	rainy	mild	high	no	yes
5	rainy	cool	normal	no	yes

6	rainy	cool	normal	yes	no
7	overcast	cool	normal	yes	yes
8	sunny	cool	high	no	no
9	sunny	cool	normal	no	yes
10	rainy	mild	normal	no	yes
11	sunny	mild	normal	yes	yes
12	overcast	mild	high	yes	yes
13	overcast	hot	normal	no	yes
14	rainy	mild	high	yes	no

Πίνακας 2: Σύνολο Δεδομένων που χρησιμοποιείται στο παράδειγμα των Δέντρων Απόφασης

Πρώτο βήμα για την κατασκευή του δέντρου είναι ο υπολογισμός της εντροπίας της κατηγορία που πρέπει να προβλεφθεί.

$$\text{Εντροπία}(play) = -\frac{9}{14} * \log_2\left(\frac{9}{14}\right) - \frac{5}{14} \log_2\left(\frac{5}{14}\right) = 0.94$$

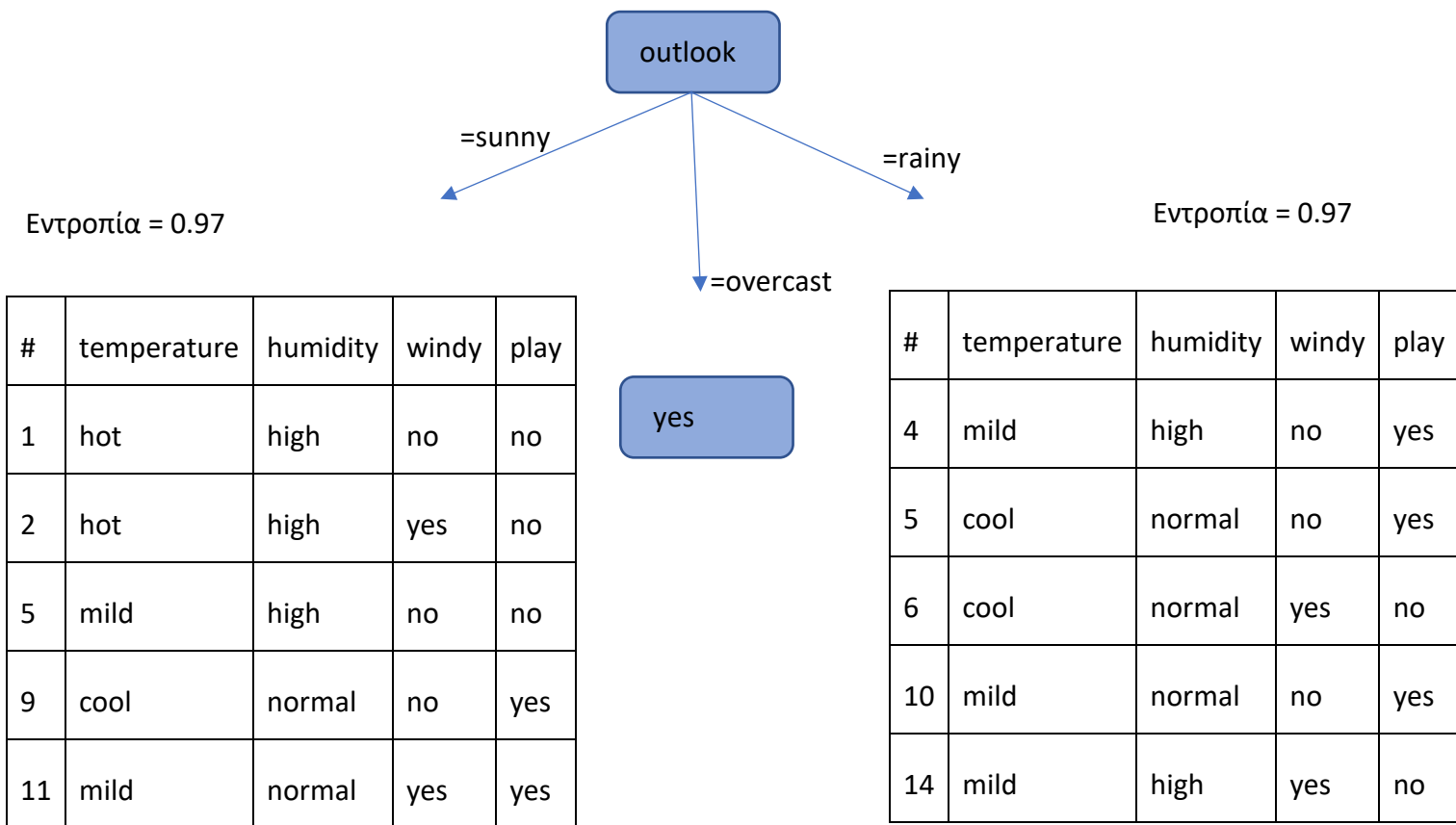
Για να επιλεγεί η ρίζα του δέντρου χρειάζεται να υπολογιστούν οι εντροπίες των άλλων χαρακτηριστικών και το κέρδος πληροφορίας για κάθε ένα από αυτά. Στον παρακάτω πίνακα φαίνονται τα αποτελέσματα των υπολογισμών.

Κατηγορία	Εντροπία	Κέρδος
outlook	0.69	0.25
temperature	0.90	0.04
humidity	0.79	0.15

windy	0.89	0.10
-------	------	------

Πίνακας 3: Αποτελέσματα υπολογισμών πρώτου βήματος

Ως ρίζα του δέντρου λοιπόν θα επιλεγθεί το χαρακτηριστικό outlook επειδή έχει το μεγαλύτερο κέρδος πληροφορίας. Στο παρακάτω σχήμα απεικονίζεται το δέντρο και το σύνολο δεδομένων αφού έχει οριστεί η ρίζα. Το επόμενο βήμα για την κατασκευή είναι η επιλογή του επόμενου χαρακτηριστικού - κόμβου για κάθε μία από τις τρεις πιθανές τιμές της ρίζας. Είναι φανερό ότι εάν outlook = overcast τότε η απάντηση είναι play και έτσι χρειάζεται να βρεθούν οι κόμβοι μόνο για τις άλλες δύο τιμές.



Σχήμα 25: Διαδικασία Επέκτασης Δέντρου Απόφασης

Η διαδικασία επέκτασης του δέντρου είναι παρόμοια με την επιλογή της ρίζας. Η μόνη διαφορά είναι ότι σαν εντροπία του συνόλου θα θεωρηθεί η εντροπία του κάθε υποσυνόλου για κάθε πιθανή τιμή του χαρακτηριστικού outlook. Έτσι για τον αριστερά κόμβο, θα

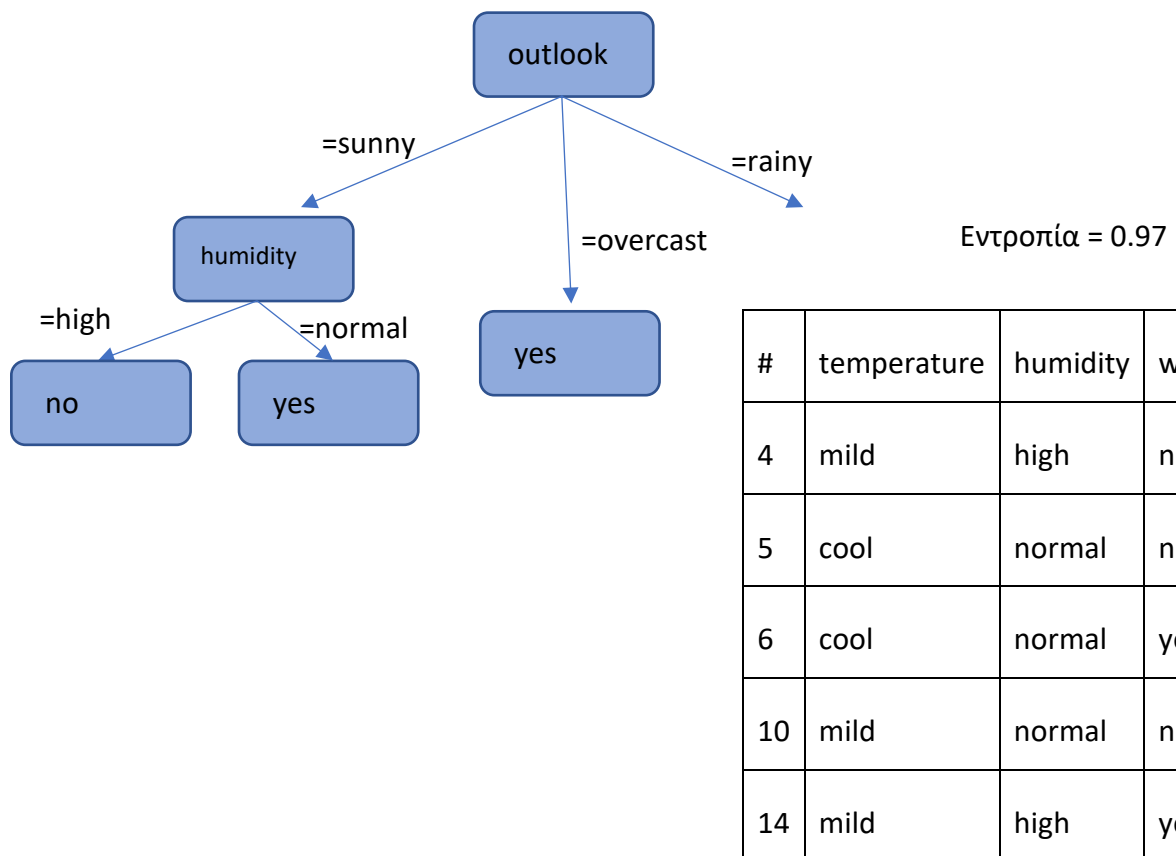
υπολογιστούν ξανά οι τιμές της εντροπίας και το κέρδος πληροφορίας για κάθε κατηγορία. Οι υπολογισμοί φαίνονται στον παρακάτω πίνακα.

Κατηγορία	Εντροπία	Κέρδος
temperature	0.4	0.57
humidity	0	0.97
windy	-	-

Πίνακας 4: Αποτελέσματα υπολογισμών για το δεύτερο διαχωρισμό

Από τη στιγμή που κατά τη διάρκεια των υπολογισμών προκύπτει πως η εντροπία του χαρακτηριστικού humidity είναι 0, τότε επιλέγεται αυτό σαν κόμβος καθώς το κέρδος για το χαρακτηριστικό αυτό έχει τη μέγιστη τιμή.

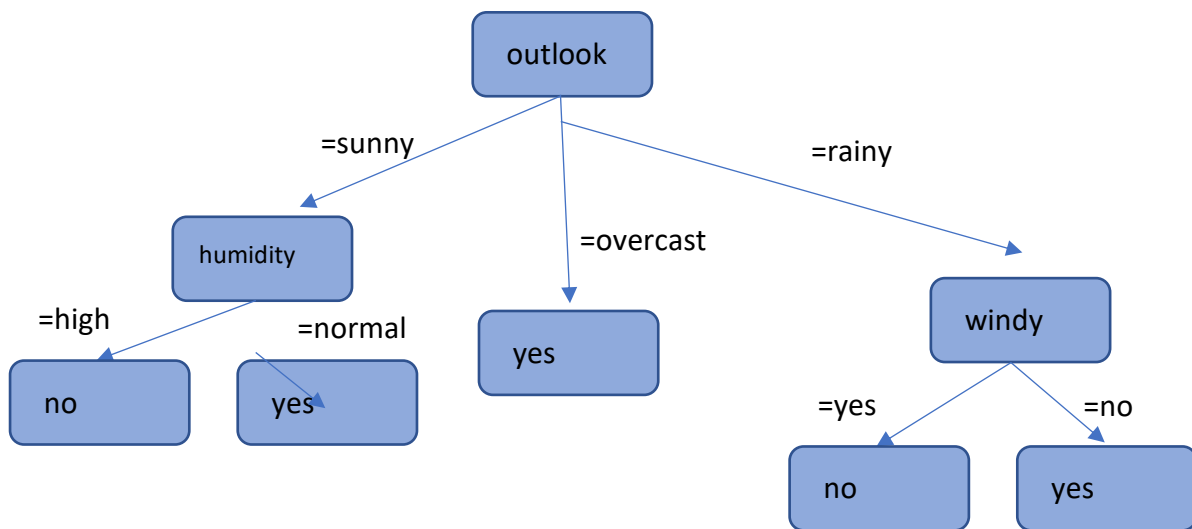
Έτσι, το δέντρο είναι το ακόλουθο:



Σχήμα 26: Συνέχεια Επέκτασης Δέντρου Απόφασης

Η ίδια διαδικασία επαναλαμβάνεται και για το άλλο υποσύνολο.

Η τελική μορφή του δέντρου είναι:



Σχήμα 27: Τελικό Δέντρο Απόφασης

3.4 Πλεονεκτήματα και Μειονεκτήματα Δέντρων Απόφασης

Πλεονεκτήματα των Δέντρων Απόφασης:

- Είναι κατανοητά και μπορούν να απεικονισθούν εύκολα
- Δεν απαιτούν προεπεξεργασία (preprocessing) του συνόλου δεδομένων. Άλλα μοντέλα απαιτούν κανονικοποίηση των δεδομένων (normalization) ή και συμπλήρωση τιμών, διαδικασίες που απαιτούν πόρους και χρόνο
- Μπορούν να διαχειριστούν αριθμητικά και κατηγορικά δεδομένα

Μειονεκτήματα των Δέντρων Απόφασης:

- Ενδέχεται να δημιουργηθούν πολύπλοκα δέντρα τα οποία προσαρμόζονται στο σύνολο εκπαίδευσης και δεν μπορούν να πραγματοποιήσουν γενικές προβλέψεις στο σύνολο ελέγχου (overfitting)
- Έχει αποδειχθεί πως αντιμετωπίζουν δυσκολίες στην μάθηση συγκεκριμένων προβλημάτων
- Εάν μία κατηγορία κυριαρχεί στο σύνολο των δεδομένων είναι πιθανό το αποτέλεσμα να μην είναι γενικό και να ευνοεί την κατηγορία αυτή.

Κεφάλαιο 4: Προβλέψεις

4.1 Εισαγωγή στις προβλέψεις

Η επιστήμη των προβλέψεων έχει ως βασικό αντικείμενο ανάλυσης το παρελθόν ενός μεγέθους και την επεξεργασία αυτού με τέτοιο τρόπο, ώστε να μπορεί να προσδιοριστεί ή να περιγραφεί το ρίσκο και η αβεβαιότητα για τις μελλοντικές εξελίξεις. Οι προβλέψεις αποτελούν αναπόσπαστο της καθημερινότητάς μας, τόσο το άτομο όσο και οι διάφοροι οργανισμοί και επιχειρήσεις χρησιμοποιούν άμεσες και έμμεσες προβλέψεις ώστε να προετοιμαστούν καλύτερα για το μέλλον. Η παραγωγή μίας πρόβλεψης για τις εβδομαδιαίες ανάγκες ενός νοικοκυριού, η πρόβλεψη για εξάντληση των αποθεμάτων καυσίμου ενός αυτοκινήτου ανάλογα με τις αποστάσεις σε χιλιόμετρα ή ακόμα και η τοποθέτηση ενός στοιχήματος είναι αποτελέσματα ανάλυσης εμπειρικών στοιχείων και δεδομένων που έχουν χρησιμοποιηθεί για να φτάσουμε σε ένα λογικό συμπέρασμα.

Υπάρχει μεγάλη ποικιλία στους διάφορους επιστημονικούς κλάδους που χρησιμοποιούν τις προβλέψεις. Οι χρηματοοικονομικοί αναλυτές και σύμβουλοι επιχειρήσεων παράγουν προβλέψεις ώστε να προχωρήσουν σε επενδύσεις υψηλού ρίσκου με μεγαλύτερη ασφάλεια και να αυτοματοποιήσουν τη διαδικασία λήψης αποφάσεων σε συνθήκες με μεγάλη αβεβαιότητα. Οι μετεωρολόγοι καλούνται καθημερινά στην παραγωγή προβλέψεων για τον καιρό, κάθε επιχείρηση καλείται να λάβει σημαντικές αποφάσεις για τα μελλοντικά της σχέδια με βάση τις προβλέψεις που έχει πραγματοποιήσει ενώ ακόμα και το κράτος παράγει ετήσιες προβλέψεις για το Ακαθάριστο Εγχώριο Προϊόν και την Ενεργειακή Κατανάλωση της χώρας

Στην παρούσα εργασία εξετάζεται κατά πόσο είναι δυνατόν να πραγματοποιηθούν προβλέψεις για τις ημερήσιες τιμές κλεισίματος μετοχών. Στη συνέχεια, με τη βοήθεια των προβλέψεων αυτών εξετάζεται αν γίνεται να δημιουργηθεί μία αυτοματοποιημένη διαδικασία λήψης επενδυτικών αποφάσεων.

Οι τεχνικές προβλέψεων κατηγοριοποιούνται σε ποσοτικές και ποιοτικές. Οι ποσοτικές μέθοδοι είναι αυτές που καταλήγουν σε ένα αποτέλεσμα μέσω υπολογισμών σε τιμές του παρελθόντος (χρονοσειρές) ενώ οι ποιοτικές μέθοδοι βασίζονται κυρίως στην εμπειρία και στις γνώσεις των ανθρώπων που παράγουν την πρόβλεψη.

Οι προβλέψεις υλοποιούνται από τα διάφορα μοντέλα πρόβλεψης. Ως μοντέλο πρόβλεψης ορίζεται η διαδικασία που ακολουθείται για να παραχθούν οι ζητούμενες προβλέψεις. Στα πλαίσια της συγκεκριμένης εργασίας χρησιμοποιούνται ποσοτικά μοντέλα πρόβλεψης. Συγκεκριμένα, οι προβλέψεις παράγονται από μοντέλα μηχανικής μάθησης και νευρωνικά δίκτυα.

Στην συνέχεια του κεφαλαίου θα γίνει μία εισαγωγή στις χρονοσειρές, τους διάφορους τύπους σφαλμάτων και την αξιολόγηση των προβλέψεων.

4.2 Χρονοσειρές

Με τον όρο χρονοσειρές περιγράφεται μια σειρά διαδοχικών παρατηρήσεων μιας τιμής ενός μεγέθους που λαμβάνονται σε ορισμένες χρονικές στιγμές ή περιόδους που ισαπέχουν μεταξύ τους.

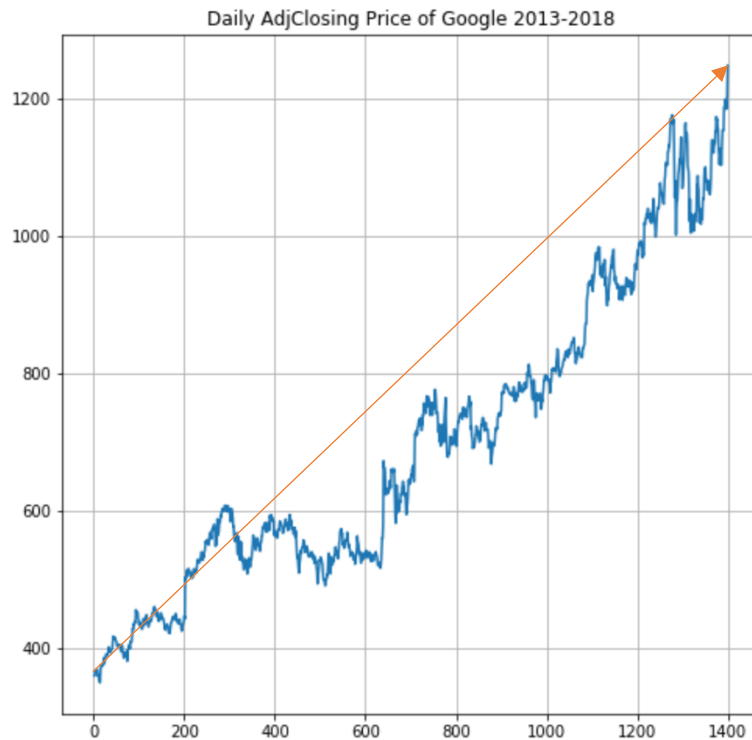
Συμβολίζοντας με X_i τις n χρονικές στιγμές (έτη, μήνες, μέρες κ.λπ.) και με Y_i τις τιμές των αντίστοιχων παρατηρήσεων, δημιουργούνται η ζεύγη της μορφής $M(X_i, Y_i)$ που μπορούν να παρασταθούν στο καρτεσιανό σύστημα. Ενώνοντας τα σημεία αυτά (M_1, M_2, \dots, M_n) δημιουργείται το χρονοδιάγραμμα.

Οι διαδοχικές αυτές παρατηρήσεις δεν είναι ανεξάρτητες μεταξύ τους, αλλά οι μελλοντικές τιμές τους μπορούν να προσδιορισθούν από τις προηγούμενες. Τα μοντέλα που περιγράφουν τη διαχρονική εξέλιξη ενός μεγέθους, για το οποίο υπάρχει πλήρης γνώση των παραγόντων που το επηρεάζουν ονομάζονται ντετερμινιστικά.

Στις πραγματικές χρονοσειρές, το μέλλον δεν καθορίζεται πλήρως από το παρελθόν, αλλά εξαρτάται σημαντικά και από τυχαίους παράγοντες, οι οποίοι αντιπροσωπεύουν στατιστικές μεταβλητές. Τα μοντέλα που περιέχουν τον τυχαίο παράγοντα καλούνται στοχαστικά. Τα βασικά ποιοτικά χαρακτηριστικά μιας χρονοσειράς, που μπορούν να αποσυντεθούν με τη χρήση διαφόρων μεθόδων ανάλυσης των χρονοσειρών είναι τα ακόλουθα:

- Τάση

Η τάση ορίζεται ως η μακροπρόθεσμη μεταβολή του μέσου επιπέδου τιμών της χρονοσειράς. Παρουσιάζει τη γενική εικόνα της χρονοσειράς δηλαδή εάν είναι ανοδική, πτωτική ή σταθερή. Με την ύπαρξη ενός μεγάλου αριθμού παρατηρήσεων μπορούμε να αποφανθούμε, για το αν μια χρονοσειρά παρουσιάζει τάση ή όχι.

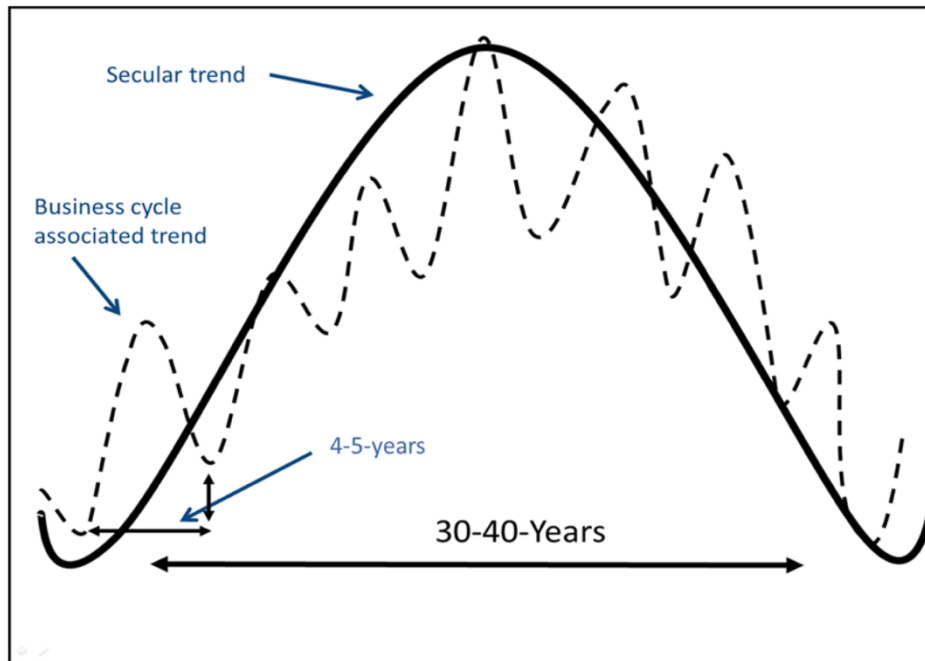


Σχήμα 28: Adjacent Closing Price της Google

Στην παραπάνω εικόνα φαίνεται η Adjacent Closing ημερήσια τιμή της μετοχής GOOG η οποία για το διάστημα 2013-2018 παρουσιάζει ανοδική τάση.

- Κυκλικότητα

Η κυκλικότητα ορίζεται ως μια «κυματοειδή» μεταβολή που οφείλεται σε ειδικές εξωγενείς συνθήκες και εμφανίζεται κατά περιόδους. Οι περίοδοι αυτές, συνήθως, δεν είναι σταθερές και το μήκος τους είναι μεγαλύτερο του έτους. Η κυκλικότητα παρατηρείται σε μακροοικονομικά μεγέθη όπως στο Ακαθάριστο Εθνικό Προϊόν και σε χρηματιστηριακούς δείκτες. Είναι αποτέλεσμα των γενικότερων οικονομικών συνθηκών που χαρακτηρίζονται από ανόδους της οικονομίας και υφέσεις.



Σχήμα 29: Παράδειγμα Κυκλικότητας μίας Χρονοσειράς

- Εποχικότητα

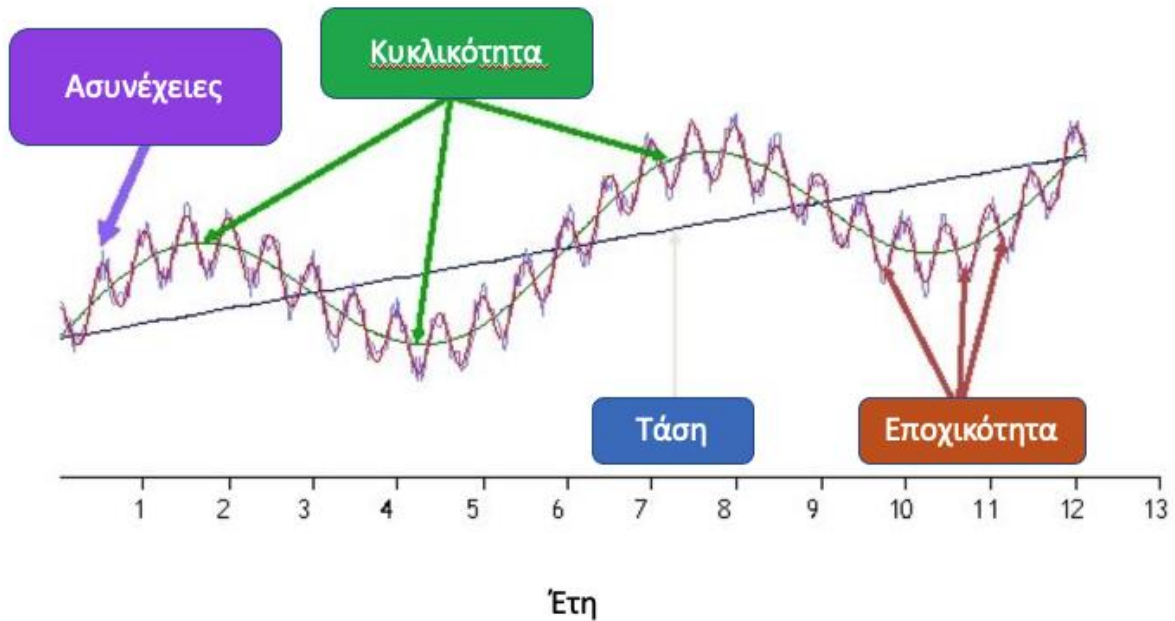
Η εποχικότητα αποτελεί μια περιοδική διακύμανση, η οποία έχει σταθερό μήκος και ταυτόχρονα μικρότερο του ενός έτους. Η διακύμανση αυτή είναι εύκολο να προβλεφθεί στις περισσότερες περιπτώσεις. Χαρακτηριστικά παραδείγματα εποχικών χρονοσειρών αποτελούν οι πωλήσεις καυσόξυλων, πετρελαίου θέρμανσης και φυσικού αερίου τον χειμώνα γιατί επαναλαμβάνονται με τον ίδιο κριβώς τρόπο.

- Ασυνέχεια

Ασυνέχεια ορίζεται το φαινόμενο όπου οι μεταβολές των τιμών μιας χρονοσειράς δεν ακολουθούν το ίδιο πρότυπο συμπεριφοράς με αυτήν.

- Τυχειότητα

Η τυχειότητα είναι η τελευταία συνιστώσα μετά την αφαίρεση των συνιστωσών της τάσης, της κυκλικότητας και της εποχικότητας. Σε μια χρονοσειρά όπου η τυχειότητα υπερκαλύπτει τις υπόλοιπες συνιστώσες, η διαδικασία της πρόβλεψής είναι δυσκολότερη και πιθανότατα με μεγαλύτερο σφάλμα.



Σχήμα 30: Παράδειγμα Χαρακτηριστικών μιας Χρονοσειράς

Στην παραπάνω εικόνα φαίνεται η ανάλυση μιας χρονοσειράς με βάση τις συνιστώσες της κυκλικότητας, της τάσης, της εποχικότητας και τις ασυνέχειες.

4.3 Ακρίβεια Προβλέψεων

Η ακρίβεια μιας πρόβλεψης εκτιμάται με βάση τη διαφορά της πραγματικής τιμής και της αντίστοιχης πρόβλεψης, δηλαδή μέσω του σφάλματος πρόβλεψης.

Το σφάλμα ορίζεται ως:

$$e_i = Y_i - F_i$$

Σχέση 5.1

όπου Y_i η πραγματική τιμή του μεγέθους και F_i η προβλεπόμενη τιμή.

Οι δημοφιλέστεροι στατιστικοί δείκτες σφάλματος είναι οι ακόλουθοι:

- Μέσο Σφάλμα (Mean Error)

$$ME = \frac{1}{n} \sum_{i=1}^n (Y_i - F_i) = \frac{1}{n} \sum_{i=1}^n e_i$$

Σχέση 5.2

όπου n το σύνολο των παραδειγμάτων για τα οποία πραγματοποιήθηκαν προβλέψεις. Το Μέσο Σφάλμα υπολογίζεται από τον μέσο όρο των σφαλμάτων των παραδειγμάτων και είναι προσημασμένο μέγεθος. Θετικές τιμές του δείκτη δηλώνουν απαισιοδοξία στις προβλέψεις, αφού οι προβλεπόμενες τιμές είναι μικρότερες από τις πραγματικές, ενώ αρνητικές απαισιοδοξία και τιμές κοντά στο μηδέν υποδηλώνουν ότι τα σφάλματα είναι τυχαία.

- Μέσο Απόλυτο Σφάλμα (Mean Absolute Error – MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - F_i| = \frac{1}{n} \sum_{i=1}^n |e_i|$$

Σχέση 5.3

Το MAE είναι ένα μέτρο αστοχίας της προβλεπόμενης σε σχέση με την πραγματική τιμή, διατηρώντας τις μονάδες μέτρησης της αρχικής χρονοσειράς, χωρίς να λαμβάνεται υπόψη το πρόσημο και άρα η κατεύθυνση της πρόβλεψης. Μεγάλη τιμή του δείκτη υποδηλώνει μικρή ακρίβεια στις προβλέψεις.

- Μέσο τετραγωνικό σφάλμα (Mean Squared Error – MSE)

$$ME = \frac{1}{n} \sum_{i=1}^n (Y_i - F_i)^2 = \frac{1}{n} \sum_{i=1}^n e_i^2$$

Σχέση 5.4

Αντίστοιχα με το μέσο απόλυτο σφάλμα το MSE είναι ένας ένα μέτρο της ακρίβειας των προβλέψεων. Με την πράξη του τετραγωνισμού δίνει μεγαλύτερο βάρος σε μεγάλα σφάλματα και μικρότερο στα μικρά.

Για να εκφράζεται το σφάλμα σε μονάδες μέτρησης ίδιες με αυτές της αρχικής χρονοσειράς χρησιμοποιείται η ρίζα του MSE:

$$RMSE = \sqrt{MSE} \quad (RootMeanSquareError)$$

Σχέση 5.5

- Μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error – MAPE)

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - F_i}{Y_i} \right| * 100\% = \frac{1}{n} \sum_{i=1}^n \left| \frac{e_i}{Y_i} \right| * 100\%$$

Σχέση 5.6

Χρησιμοποιείται για τη σύγκριση της ακρίβειας χρονοσειρών με διαφορετικά επίπεδα μέσης τιμής. Μικρότερες τιμές του MAPE δηλώνουν μεγαλύτερη ακρίβεια πρόβλεψης.

Κεφάλαιο 5: Πειραματική Διαδικασία

5.1 Εισαγωγή

Στο κεφάλαιο αυτό θα γίνει αναλυτική περιγραφή των πειραμάτων πρόβλεψης που υλοποιήθηκαν.

Στόχος των πειραμάτων είναι η πρόβλεψη της ημερήσιας τιμής των μετοχών οι οποίες συμπεριλαμβάνονται στο δείκτη Standard and Poor's 500 (S&P 500). Για την υλοποίηση των προβλέψεων είναι απαραίτητη η ύπαρξη μεγάλου όγκου δεδομένων για την κάθε μετοχή έτσι ώστε να εκπαιδευτούν τα μοντέλα και στην συνέχεια να προβλέψουν τις τιμές κλεισίματος κάθε μέρας του επόμενου έτους.

Για την πρόβλεψη των ημερήσιων τιμών των μετοχών αξιοποιούνται μοντέλα μηχανικής μάθησης και διάφοροι τύποι νευρωνικών δικτύων. Για κάθε μοντέλο πρόβλεψης γίνεται βελτιστοποίηση των παραμέτρων του καθώς και του ορίζοντα πρόβλεψης, δηλαδή τον αριθμό των ημερών που αυτό «κοιτά» στο παρελθόν έτσι ώστε προβλέψει την τιμή κλεισίματος της επόμενης μέρας. Αφού ολοκληρωθεί η διαδικασία βελτιστοποίησης, κάθε μοντέλο υλοποιεί προβλέψεις με το βέλτιστο σύνολο παραμέτρων και αξιολογείται η ακρίβεια του.

Στη συνέχεια του κεφαλαίου θα γίνει περιγραφή του συνόλου δεδομένων αλλά και της πειραματικής διαδικασίας για κάθε μοντέλο πρόβλεψης.

5.2 Σύνολο Δεδομένων

Ο δείκτης Standard & Poor's 500 περιλαμβάνει τις 500 μεγαλύτερες, βάσει κεφαλαιοποίησης, εταιρείες εισηγμένες στις αγορές του NYSE και του NASDAQ. Αποτυπώνει περίπου το 80% της διαθέσιμης κεφαλαιοποίησης της αγοράς και έτσι φορές χρησιμοποιείται συχνά από τους αναλυτές ως ένα μέτρο για την γενικότερη πορεία της οικονομίας των Ηνωμένων Πολιτειών της Αμερικής.

Προκειμένου να υλοποιηθούν προβλέψεις για τις μετοχές που απαρτίζουν τον S&P 500 είναι απαραίτητο να υπάρχει μεγάλος όγκος δεδομένων. Έτσι το σύνολο δεδομένων αποτελείται από τις τιμές κλεισίματος των μετοχών αυτών για τα έτη 2013-2019.

Συγκεκριμένα, το διάστημα 2013-2017 χρησιμοποιήθηκε για την εκπαίδευση των μοντέλων (training dataset), το έτος 2018 για τον καθορισμό των βέλτιστων υπερπαραμέτρων και την αξιολόγηση της απόδοσης του μοντέλου και το 2019 ως το έτος στο οποίο θα αξιολογηθεί η κερδοφορία της τελικής στρατηγικής.

Επομένως οι προβλεπόμενες μετοχές θα πρέπει να βρίσκονται στον δείκτη S&P 500 και να έχουν τιμές από το 2013 μέχρι και το 2019. Κατά την συλλογή των δεδομένων παρατηρήθηκε το πρόβλημα ότι πολλές από τις εταιρείες που υπάρχουν στον δείκτη σήμερα δεν είχαν πραγματοποιήσει το Initial Public Offering το 2013. Για παράδειγμα η εταιρεία 'Qorvo' ανήκει στον δείκτη σήμερα αλλά η μετοχή της ήταν διαθέσιμη στο χρηματιστήριο από τις 2 Ιανουαρίου του 2015 και μετά.

Για να μην υπάρχουν κενά στο σύνολο των δεδομένων επιλέχτηκαν 469 μετοχές από τον δείκτη. Η διαφορά μετά την αφαίρεση είναι πολύ μικρή καθώς αφαιρείται λιγότερο από το 10% του συνόλου του δείκτη. Οι δείκτες (tickers) των μετοχών που συμπεριλαμβάνονται στο τελικό σύνολο δεδομένων υπάρχουν στον πίνακα 17 στο παράρτημα.

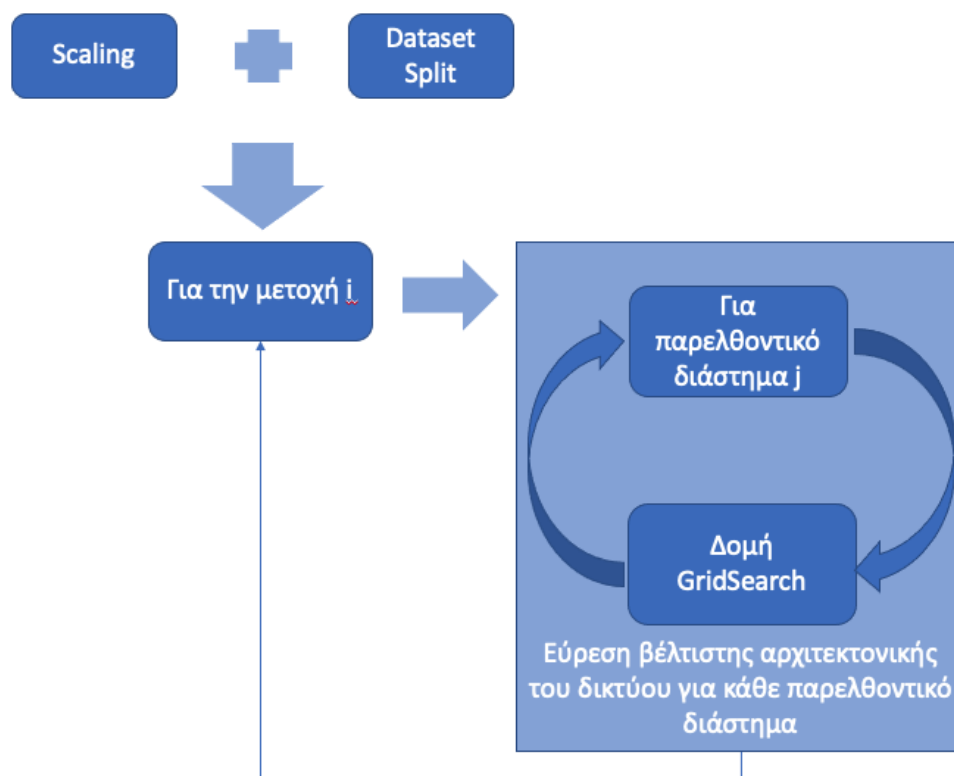
Έτσι, το τελικό σύνολο δεδομένων στο οποίο θα πραγματοποιηθούν οι προβλέψεις με τον τρόπο που αναφέρθηκε παραπάνω αποτελείται από 253 ημερήσιες τιμές κλεισίματος (253 είναι κατά μέσο όρο ο αριθμός των ημερών κάθε έτους για τις οποίες το χρηματιστήριο λειτουργεί) για κάθε μία από τις 469 μετοχές για το διάστημα 7 ετών 2013-2019.

5.3 Μοντέλα Τοπικών Προβλέψεων

Η πρώτη προσέγγιση που πραγματοποιήθηκε είναι η δημιουργία ενός μοντέλου για κάθε ξεχωριστή χρονοσειρά, δηλαδή κάθε μετοχή. Συγκεκριμένα, χρησιμοποιήθηκε πλήρως συνδεδεμένο νευρωνικό δίκτυο τύπου MLP Regressor και για τον καθορισμό των υπερπαραμέτρων του η δομή Gridsearch. Αξίζει να σημειωθεί πως πέρα από τις παραμέτρους του δικτύου, πρέπει να γίνει διερεύνηση και για τον βέλτιστο ορίζοντα πρόβλεψης. Έτσι, το πρόβλημα δημιουργίας ενός μοντέλου για κάθε χρονοσειρά, ανάγεται στην εύρεση της βέλτιστης αρχιτεκτονικής του νευρωνικού δικτύου και ορίζοντα πρόβλεψης για κάθε μετοχή.

Για την υλοποίηση της προσέγγισης το πρόβλημα διαχωρίστηκε σε 3 βήματα:

- Την κανονικοποίηση (scaling) των τιμών των χρονοσειρών έτσι ώστε οι τιμές να βρίσκονται στο σύνολο [0,1].
- Το διαχωρισμό των επιμέρους συνόλων training και testing.
- Τη δημιουργία της δομής επανάληψης η οποία θα υπολογίζει τις βέλτιστες υπερπαραμέτρους του νευρωνικού δικτύου για κάθε χρονοσειρά.



Σχήμα 31: Διάγραμμα Αλγορίθμου Εύρεσης Βέλτιστου TNM

Στην παραπάνω εικόνα φαίνεται το διάγραμμα λειτουργίας του αλγορίθμου εύρεσης του βέλτιστου νευρωνικού δικτύου για κάθε μετοχή.

Η κανονικοποίηση έγινε με Min-Max Scaling, όπως φαίνεται στην παρακάτω εξίσωση:

$$x_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Σχέση 5.7

όπου x η τιμή που κανονικοποιείται, $\min(x)$ η ελάχιστη τιμή της χρονοσειράς στο σύνολο δεδομένων και $\max(x)$ η μέγιστη. Η σχέση της κανονικοποίησης εφαρμόζεται σε κάθε τιμή που υπάρχει στο σύνολο των δεδομένων.

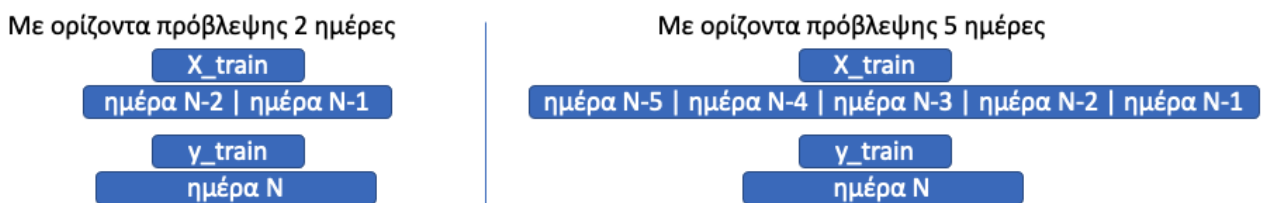
Συγκεκριμένα, η σχέση κανονικοποίησης εφαρμόζεται στο σύνολο δεδομένων εκπαίδευσης κάθε χρονοσειράς (δηλαδή στις τιμές της κάθε μετοχής για τα έτη 2013-2017). Το σύνολο δεδομένων επαλήθευσης (οι τιμές των μετοχών για το 2018) κανονικοποιείται με την ίδια σχέση απλά όπου $\min(x)$ και $\max(x)$ τοποθετούνται η ελάχιστη και μέγιστη τιμή της μετοχής από το σύνολο δεδομένων εκπαίδευσης. Έτσι, οι κανονικοποιημένες τιμές των μετοχών για το 2018 μπορούν να ξεπεράσουν το 1.

Ενδιαφέρον παρουσιάζει η δομή εύρεσης των βέλτιστων παραμέτρων του νευρωνικού δικτύου. Ο δείκτης σφάλματος που χρησιμοποιείται για την αξιολόγηση της ακρίβειας του δικτύου είναι το μέσο απόλυτο ποσοστιαίο σφάλμα. Για κάθε μετοχή λοιπόν, πρέπει να βρεθούν:

1. Ο βέλτιστος ορίζοντας πρόβλεψης, δηλαδή πόσο «παλιά» μπορεί να «δει» το δίκτυο προκειμένου να προβλέψει την τιμή της μετοχής την επόμενη μέρα. Για παράδειγμα, μία προσέγγιση είναι το νευρωνικό να διαθέτει 40 ημέρες και να καλείται να προβλέψει την ημέρα 41. Ενώ, σε άλλη περίπτωση το δίκτυο διαθέτει τιμές των δύο προηγούμενων ημερών και καλείται να προβλέψει την 3^η .

2. Οι βέλτιστες παράμετροι του νευρωνικού δικτύου, ο αριθμός των νευρώνων κρυφών επιπέδων, η συνάρτηση ενεργοποίησης, ο solver αλλά και ο μέγιστος αριθμός επαναλήψεων εκτέλεσης του αλγορίθμου.

Αξίζει να σημειωθεί πως κάθε φορά που αλλάζει ο ορίζοντας πρόβλεψης πρέπει να οργανωθούν εκ νέου τα σύνολα δεδομένων εκπαίδευσης (training) και επαλήθευσης (testing) σε νέα X_{train} , y_{train} και X_{test} , y_{test} αντίστοιχα όπως φαίνεται στο παρακάτω σχήμα.



Σχήμα 32: Δύο Διαφορετικοί Ορίζοντες Πρόβλεψης

Αφού βρεθούν οι βέλτιστες παράμετροι του νευρωνικού για κάθε ορίζοντα πρόβλεψης και κάθε μετοχή, τότε με βάση το MAPE συγκρίνονται οι αποδόσεις των διαφορετικών οριζόντων πρόβλεψης και βρίσκεται το βέλτιστο δίκτυο για κάθε χρονοσειρά. Τα αποτελέσματα για το σύνολο των δεδομένων φαίνονται στον αντίστοιχο πίνακα στο παράρτημα, ενώ ενδεικτικά για κάποιες από τις μετοχές φαίνονται στον παρακάτω πίνακα.

Μετοχή	Ορίζοντας Πρόβλεψης (Μέρες)	Κρυφό Επίπεδο	Solver	Μέγιστος Αριθμός Επαναλήψεων	Συνάρτηση Ενεργοποίησης	MAPE
LNT	6	(6,)	lbfgs	100	relu	0,824142
KO	2	(4,)	lbfgs	100	relu	0,718003
COP	10	(40,)	lbfgs	100	relu	1,398373

Πίνακας 5: Βέλτιστες Παράμετροι και τιμές Ορίζοντα πρόβλεψης για 3 μετοχές

Η μέση τιμή του μέσου απόλυτου ποσοστιαίου σφάλματος πρόβλεψης για όλες τις μετοχές του συνόλου δεδομένων είναι **1.24%**.

5.4 Γενικό Μοντέλο Προβλέψεων

Η δεύτερη προσέγγιση που πραγματοποιήθηκε είναι η δημιουργία ενός μοντέλου πρόβλεψης το οποίο εφαρμόζεται σε όλες τις μετοχές του συνόλου δεδομένων.

Συγκεκριμένα, το μοντέλο πρόβλεψης είναι ένα νευρωνικό δίκτυο τύπου MLP Regressor.

Ως σύνολο δεδομένων εκπαίδευσης χρησιμοποιούνται οι ημερήσιες τιμές των μετοχών στο διάστημα 2013-2017 και σύνολο δεδομένων επαλήθευσης οι τιμές για το 2018.

Αντίστοιχα με τα τοπικά μοντέλα πρόβλεψης, έτσι και το γενικό μοντέλο είναι απαραίτητο να γίνει διερεύνηση για να βρεθούν:

1. Ο βέλτιστος ορίζοντας πρόβλεψης
2. Οι βέλτιστες παράμετροι της αρχιτεκτονικής του νευρωνικού δικτύου, ο αριθμός των νευρώνων κρυφών επιπέδων, η συνάρτηση ενεργοποίησης, ο solver αλλά και ο μέγιστος αριθμός επαναλήψεων εκτέλεσης του αλγορίθμου.

Η κανονικοποίηση και η οργάνωση των X_{train} , y_{train} , X_{test} και y_{test} έγινε όπως και στα τοπικά μοντέλα πρόβλεψης. Για την αξιολόγηση της ακρίβειας των προβλέψεων χρησιμοποιήθηκε ο δείκτης MAPE. Στο παρακάτω διάγραμμα φαίνεται η διαδικασία εύρεσης των βέλτιστων παραμέτρων για το νευρωνικό δίκτυο.



Σχήμα 33: Διαδικασία Εύρεσης Βέλτιστης Αρχιτεκτονικής

Μέσω αυτής της διαδικασίας βρίσκεται η βέλτιστη αρχιτεκτονική του δικτύου για κάθε ορίζοντα πρόβλεψης. Στον παρακάτω πίνακα μπορείτε να βρείτε ενδεικτικά κάποια αποτελέσματα (ο πλήρης πίνακας αποτελεσμάτων υπάρχει στο παράρτημα).

Ορίζοντας Πρόβλεψης (Μέρες)	Κρυφό Επίπεδο	Συνάρτηση Ενεργοποίησης	Solver	Μέγιστος Αριθμός Επαναλήψεων	MAPE
2	(8,)	relu	lbfgs	100	14.65%
10	(6,)	relu	adam	100	14.72%
40	(20,)	logistic	adam	100	14.12%
80	(320,)	relu	adam	100	14.61%

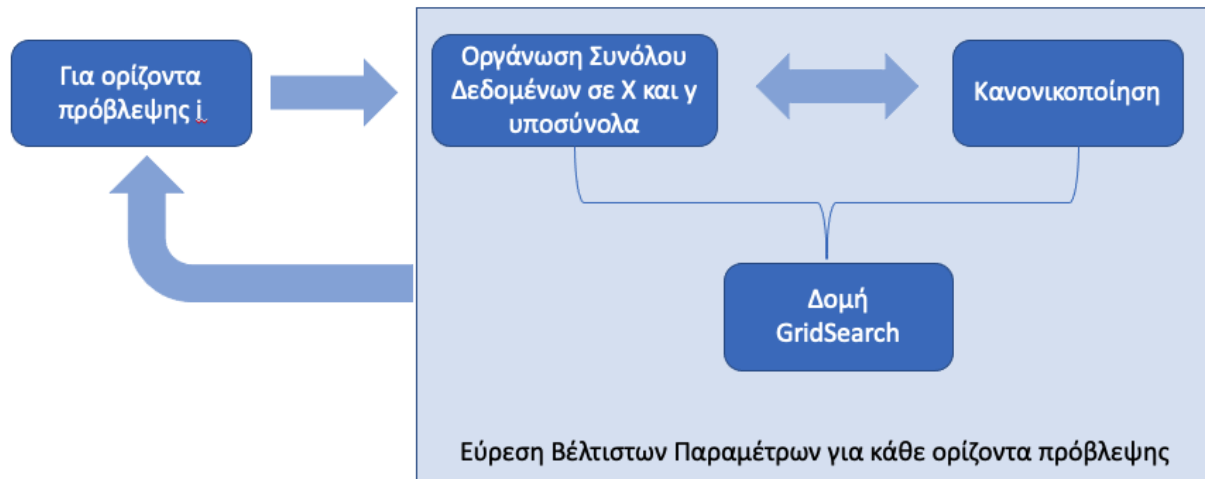
Πίνακας 6: Βέλτιστη Αρχιτεκτονική Γενικού Νευρωνικού Δικτύου για κάθε Ορίζοντα Πρόβλεψης

Για όλες τις διαφορετικές τιμές του ορίζοντα πρόβλεψης που εξετάστηκαν το απόλυτο ποσοστιαίο σφάλμα ήταν περίπου 14%. Η υψηλή τιμή του δείκτη δηλώνει πως οι προβλέψεις του μοντέλου έχουν μεγάλη απόσταση από τις πραγματικές τιμές.

Η χαμηλή ακρίβεια του μοντέλου οφείλεται στο γεγονός ότι η έξοδος του δικτύου αποδίδεται σε διαφορετική κλίμακα από αυτήν που βρίσκεται στην πραγματικότητα. Έτσι, για την βελτίωση της απόδοσης άλλαξε η διαδικασία της κανονικοποίησης των τιμών των χρονοσειρών.

Η κανονικοποίηση πλέον γίνεται σε επίπεδο ορίζοντα πρόβλεψης. Για παράδειγμα, εάν $train = \{1, 2, 3\}$ και $test = \{4\}$ τότε στο μοντέλο θα τροφοδοτηθούν οι τιμές $train_scaled = \{0, 0.5, 1\}$ και αυτό θα πρέπει να προβλέπει την τιμή 1.5 η οποία μετά τον αντίστροφο μετασχηματισμό θα γίνεται 4 και ίση με την πραγματική. Με αυτόν τον τρόπο στο παράδειγμα με το οποίο κάθε φορά τροφοδοτείται το δίκτυο, συμπεριλαμβάνεται και η κλίμακα της εξόδου σε σχέση με την είσοδο.

Στο παρακάτω διάγραμμα φαίνεται η βελτιωμένη διαδικασία εύρεσης των βέλτιστων παραμέτρων για το νευρωνικό δίκτυο.



Σχήμα 34: Διαδικασία Εύρεσης Βέλτιστων Παραμέτρων

Στον παρακάτω πίνακα μπορείτε να βρείτε ενδεικτικά κάποια αποτελέσματα.

Ορίζοντας Πρόβλεψης (Μέρες)	Κρυφό Επίπεδο	Συνάρτηση Ενεργοποίησης	Solver	Μέγιστος Αριθμός Επαναλήψεων	MAPE
2	(1,)	relu	adam	100	1.63%
8	(32,)	relu	lbfgs	200	1.21%
10	(40,)	relu	lbfgs	200	1.22%
20	(80,)	relu	adam	100	1.23%

Πίνακας 7: Βέλτιστη Αρχιτεκτονική Γενικού Νευρωνικού Δικτύου για κάθε Ορίζοντα Πρόβλεψης με την δεύτερη προσέγγιση

Όπως φαίνεται από τον πίνακα η νέα προσέγγιση για το καθολικό μοντέλο βελτίωσε την ικανότητα πρόβλεψης εφόσον το MAPE μειώθηκε από 14% σε 1.21%. Επιπλέον παρατηρείται πως η απόδοση του νευρωνικού δικτύου με αρχιτεκτονική $\{(32,), \text{relu}, \text{lbfgs}, 200\}$ και ορίζοντα πρόβλεψης τις 8 μέρες παρουσιάζει καλύτερη απόδοση από την προσέγγιση των τοπικών νευρωνικών.

Για την περαιτέρω βελτίωση της ακρίβειας των προβλέψεων, εξετάστηκε η προσθήκη χαρακτηριστικών των εξής τύπων:

- α. Χαρακτηριστικά και μεγέθη χρονοσειρών, όπως η κύρτωση ή ο μεγαλύτερος αριθμός ημερών που η μετοχή είχε τιμή πάνω από τη μέση

- β. Πληροφορίες για την βιομηχανία στην οποία δραστηριοποιείται η εταιρεία της κάθε μετοχής

Προκειμένου να πραγματοποιηθεί βελτίωση της απόδοσης με χαρακτηριστικά του πρώτου τύπου είναι απαραίτητη η διερεύνηση αυτών που μπορούν να βελτιώσουν περισσότερο την απόδοση του δικτύου. Για να γίνει αυτό υπολογίστηκαν οι τιμές 224 χαρακτηριστικών και μεγεθών για τις τιμές των χρονοσειρών στα έτη 2013-2017.

Στη συνέχεια για κάθε ένα από αυτά τα χαρακτηριστικά:

1. Το χαρακτηριστικό τοποθετήθηκε στα διανύσματα εισόδου του δικτύου
2. Ο αλγόριθμος MLP Regressor εφαρμόστηκε στα διανύσματα χρησιμοποιώντας και τα νέα χαρακτηριστικά
3. Υπολογίστηκε ο δείκτης MAPE

Με αυτόν τον τρόπο έγινε αναλυτική διερεύνηση για κάθε πιθανό χαρακτηριστικό των χρονοσειρών. Τα 3 χαρακτηριστικά με τα οποία το δίκτυο παρουσίασε τις καλύτερες αποδόσεις φαίνονται στον παρακάτω πίνακα:

Όνομα Χαρακτηριστικού	MAPE
Value__lempel_ziv_complexity__bins_100	1.26%
Value__first_location_of_maximum	1.26%
Value__linear_trend__attr_"stderr"	1.26%

Πίνακας 8: Χαρακτηριστικά Χρονοσειρών με το μικρότερο MAPE

Η προσθήκη των μεμονωμένων χαρακτηριστικών δεν βελτίωσε την ακρίβεια των προβλέψεων οπότε, διερευνητικά δοκιμάστηκαν οι γραμμικοί συνδυασμοί των καλύτερων. Ωστόσο, όπως και στην περίπτωση των μοναδικών χαρακτηριστικών έτσι και με τον συνδυασμό δύο ή τριών δεν παρατηρήθηκε βελτίωση του MAPE. Στο παράρτημα της εργασίας μπορείτε να βρείτε τον αναλυτικό πίνακα με τα αποτελέσματα όλων των πειραμάτων*.

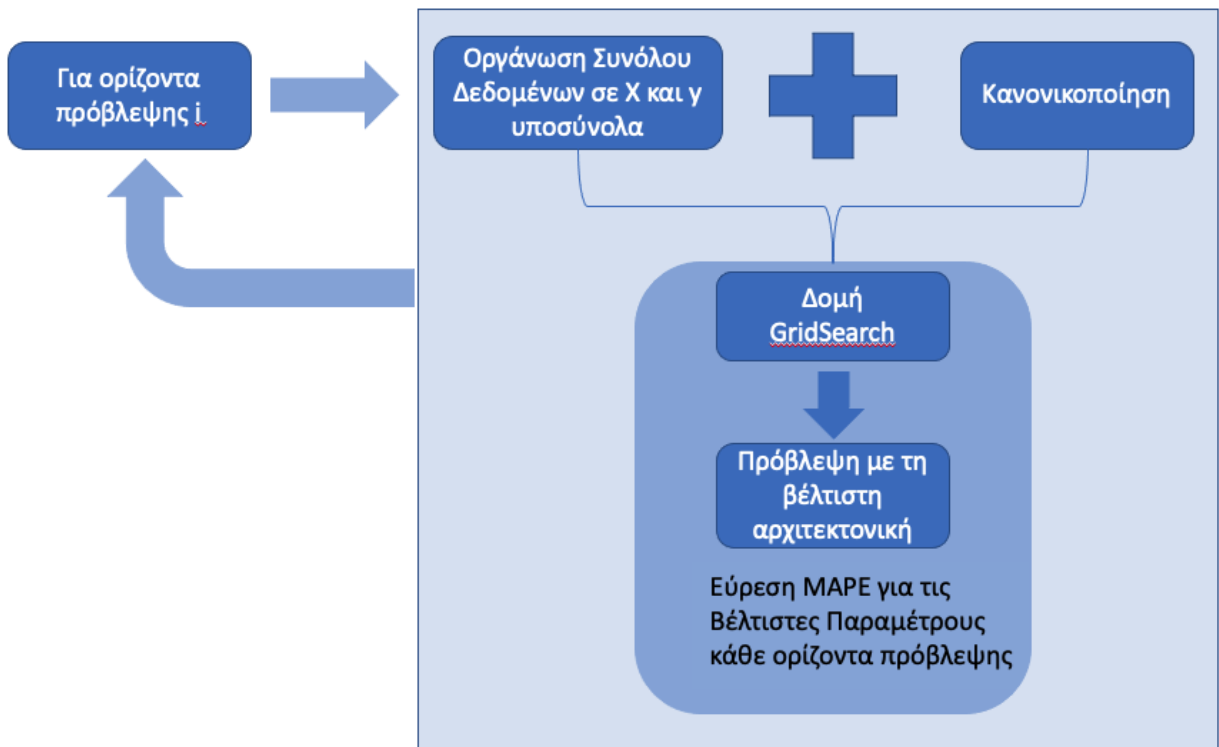
*Σημείωση: Για τον υπολογισμό των χαρακτηριστικών χρησιμοποιήθηκε η βιβλιοθήκη ts fresh. Τα χαρακτηριστικά που βρίσκονται στον πίνακα εξηγούνται αναλυτικά στο documentation της βιβλιοθήκης.

Για την βελτίωση της απόδοσης με χαρακτηριστικά του δεύτερου τύπου κρίνεται απαραίτητη η κατηγοριοποίηση των μετοχών με βάση την βιομηχανία στην οποία δραστηριοποιούνται. Το «Global Industry Classification Standard» αναπτύχθηκε το 1999 από την Standard & Poor's και παρέχει την κατηγοριοποίηση των μετοχών του δείκτη S&P 500 σε 11 τομείς και 159 υποκατηγορίες βιομηχανιών.

Η πληροφορία για την κατηγορία της κάθε μετοχής προστέθηκε στα δεδομένα εισόδου του νευρωνικού δικτύου με την τεχνική One Hot Encoding. Συγκεκριμένα, σε κάθε διάνυσμα εισόδου (κάθε υπολίστα του X_{train} και του X_{test}) έγινε προσθήκη ενός διανύσματος ίσο με τον αριθμό των κατηγοριών όπου όλα τα στοιχεία ήταν 0 και το στοιχείο που αντιστοιχούσε στην κατηγορία της εταιρείας ήταν 1. Η προσέγγιση με τις 158 κατηγορίες ήταν πολύ χειρότερη από το απλό νευρωνικό ενώ, η προσέγγιση με τις 11 κατηγορίες είχε απόδοση 1.2596%. Έτσι, παρατηρήθηκε πως ούτε με την προσθήκη αυτών των χαρακτηριστικών βελτιώνεται η ακρίβεια των προβλέψεων.

5.5 Δέντρα Απόφασης

Στην προσέγγιση αυτή εξετάστηκε η ακρίβεια των προβλέψεων που μπορεί να επιτευχθεί με τη χρήση δέντρων απόφασης. Ακριβώς όπως και στις προσεγγίσεις που αναλύθηκαν παραπάνω, πρέπει να γίνει πρώτα διερεύνηση σχετικά με την αρχιτεκτονική του υπερπαραμέτρων του δέντρου και τον ορίζοντα πρόβλεψης. Για την αξιολόγηση της ακρίβειας των προβλέψεων των δέντρων θα χρησιμοποιηθεί ο δείκτης MAPE.



Σχήμα 35: Διαδικασία Εύρεσης Βέλτιστων Παραμέτρων για τα Δέντρα Απόφασης

Στο παραπάνω σχήμα φαίνεται αναλυτικά η διαδικασία υλοποίησης, η οποία παρουσιάζει αρκετές ομοιότητες με την αντίστοιχη διαδικασία στα νευρωνικά δίκτυα. Αποτέλεσμα της διεργασίας είναι ο εύρεση των βέλτιστων παραμέτρων της αρχιτεκτονικής του δικτύου για κάθε διαφορετικό ορίζοντα πρόβλεψης. Στη συνέχεια με τις βέλτιστες παραμέτρους αυτές πραγματοποιούνται οι προβλέψεις και υπολογίζεται ο δείκτης MAPE. Στον πίνακα 9 μπορείτε να βρείτε τα αποτελέσματα.

Ορίζοντας Πρόβλεψης (Μέρες)	Criterion	Max_depth	Splitter	MAPE
2	mse	2	best	3.09%
4	mse	6	random	2.96%
8	mse	6	best	3.21%
15	mse	6	best	6.82%
30	mse	6	best	10.35%

50	mse	6	best	14.85%
70	mse	6	best	14.4%
100	mse	6	best	23.36%

Πίνακας 9: Αποτελέσματα Δέντρων Απόφασης

Το δέντρο απόφασης με την μεγαλύτερη ακρίβεια είναι αυτό με ορίζοντα πρόβλεψης 4 μέρες το οποίο έχει MAPE ίσο με 2.96%. Επομένως, παρατηρείται πως για όλα τα δέντρα που εξετάστηκαν, κανένα δεν παράγει προβλέψεις μεγαλύτερης ακρίβειας από το καλύτερο νευρωνικό δίκτυο. Έτσι, το βέλτιστο μέχρι στιγμής μοντέλο είναι το δίκτυο τύπου MLP Regressor με παραμέτρους: {(32,), relu, lbfgs, 200} και ορίζοντα πρόβλεψης τις 8 μέρες.

5.6 Recurrent Neural Networks

Οι προβλέψεις με τα αναδρομικά νευρωνικά δίκτυα έγιναν με παρόμοια διαδικασία σε σχέση με τα δέντρα απόφασης και τα νευρωνικά δίκτυα τύπου MLP Regressor. Δεν έγιναν αλλαγές στον τρόπο οργάνωσης του συνόλου δεδομένων σε X και Y υποσύνολα αλλά ούτε και στην κανονικοποίηση.

Έτσι, για κάθε ορίζοντα πρόβλεψης πραγματοποιήθηκε διερεύνηση σχετικά με την αρχιτεκτονική του δικτύου το οποίο μπορεί να κάνει ακριβέστερες προβλέψεις. Στη συνέχεια υλοποιήθηκαν οι προβλέψεις με την βέλτιστη αρχιτεκτονική για κάθε διαφορετικό ορίζοντα ώστε να υπολογιστεί το MAPE.

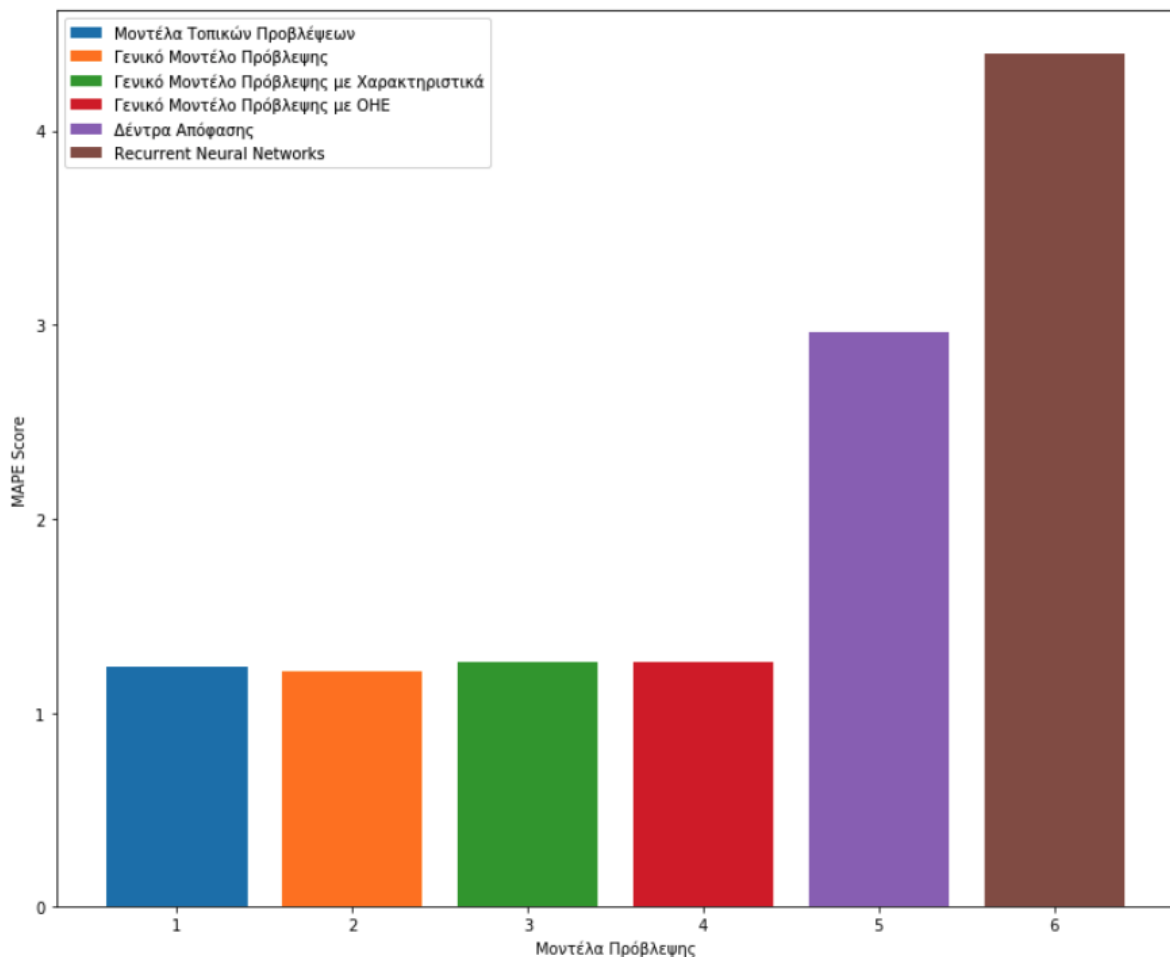
Μερικά από τα καλύτερα αποτελέσματα φαίνονται στον παρακάτω πίνακα:

Ορίζοντας Πρόβλεψης (Μέρες)	Αρχιτεκτονική Δικτύου	Συνάρτηση Ενεργοποίησης	Optimizer	MAPE
4	{1,5,1}	tanh	adam	4.40%
8	{1,10,5,1}	sigmoid	adam	4.66%
10	{1,15,1}	relu	adam	4.69%

Πίνακας 10: Αποτελέσματα Προβλέψεων με Αναδρομικά Νευρωνικά Δίκτυα

όπως φαίνεται από τον πίνακα των αποτελεσμάτων οι προβλέψεις με τα αναδρομικά έχουν ένα μέσο απόλυτο ποσοστιαίο σφάλμα γύρω στο 4.5%.

5.7 Σύνοψη Αποτελεσμάτων



Σχήμα 36: MAPE Score για τα διάφορα μοντέλα πρόβλεψης

Από τη διαδικασία των προβλέψεων λοιπόν προκύπτει πως οι ακριβέστερες προβλέψεις για το συγκεκριμένο σύνολο δεδομένων πραγματοποιούνται με το γενικό μοντέλο πρόβλεψης μετά την διόρθωση που έγινε στην κανονικοποίηση. Έτσι το μοντέλο που παράγει τις ακριβέστερες προβλέψεις και θα χρησιμοποιηθεί στη συνέχεια της εργασίας είναι το νευρωνικό δίκτυο τύπου MLP Regressor με παραμέτρους: $\{(32, \text{relu}, \text{lbfgs}, 200)\}$ και ορίζοντα πρόβλεψης τις 8 μέρες το οποίο έχει MAPE ίσο με 1.21%.

Κεφάλαιο 6: Βελτιστοποίηση Χαρτοφυλακίου

6.1 Εισαγωγή

Στη συνέχεια της διπλωματικής εργασίας θα κατασκευαστούν χαρτοφυλάκια μετοχών με βάση τις προβλέψεις που πραγματοποιήθηκαν.

Ως χαρτοφυλάκιο ορίζεται ο συνδυασμός διάφορων επενδυτικών προϊόντων. Τα εξεταζόμενα χαρτοφυλάκια στην εργασία αυτή αποτελούνται από σύνολα μετοχών οι οποίες ανήκουν στο δείκτη S&P 500 και οι τιμές τους προβλέπονται από τα μοντέλα που αναλύθηκαν στο προηγούμενο κεφάλαιο. Κύριος στόχος της κατασκευής ενός χαρτοφυλακίου είναι η διαφοροποίηση των επενδύσεων έτσι ώστε να μειωθεί ο κίνδυνος χωρίς απαραίτητα να επηρεαστεί το αναμενόμενο κέρδος. Συγκεκριμένα για να επιτευχθεί ο στόχος αυτός πρέπει να διερευνηθούν οι συνδυασμοί των μετοχών με τη μεγαλύτερη αναμενόμενη απόδοση και το μικρότερο δυνατό ρίσκο.

Στη βιβλιογραφία έχουν αναπτυχθεί πολλές υπολογιστικές διαδικασίες βελτιστοποίησης και σύνθεσης χαρτοφυλακίων. Στην παρούσα εργασία θα εξεταστούν δύο από τις δημοφιλέστερες οι οποίες είναι οι εξής:

- I. Η Μέθοδος μέσου διακύμανσης (Mean Variance)
- II. Η Μέθοδος Ελαχίστου-Μεγίστου (MinMax)

Οι μεθοδολογίες διαφοροποιούνται επειδή χρησιμοποιούν διαφορετικά μεγέθη ως μέτρα της καταλληλότητας των χαρτοφυλακίων.

Στη συνέχεια του κεφαλαίου θα γίνει ανάλυση του μοντέλου του Markowitz το οποίο ανήκει στην πρώτη κατηγορία και του μοντέλου του Young το οποίο ανήκει στη δεύτερη. Πέρα από τη μαθηματική τους ανάλυση θα πραγματοποιηθεί και η περιγραφή της υλοποίησής τους.

Για την περιγραφή της υλοποίησης των μοντέλων δεν θα χρησιμοποιηθούν οι προβλεπόμενες τιμές αλλά οι πραγματικές τιμές των μετοχών του δείκτη S&P 500 στο διάστημα 2013-2019.

6.2 Μοντέλο του Markowitz

Το πρώτο μοντέλο που υλοποιήθηκε προτάθηκε από τον Harry Markowitz στη δημοσίευσή του με τίτλο «Portfolio Selection» στο Journal of Finance το 1952. Στη συγκεκριμένη δημοσίευση εξηγεί γιατί η απλή μεγιστοποίηση του αναμενόμενου κέρδους δεν είναι ο βέλτιστος τρόπος επιλογής χαρτοφυλακίου.

Η θεωρία του προτείνει πως το ρίσκο και το αναμενόμενο κέρδος μίας επένδυσης δεν είναι χαρακτηριστικά τα οποία μπορούν να εξετάζονται ανεξάρτητα το ένα από το άλλο. Με βάση αυτήν την θεωρία, θα πρέπει για κάθε νέα επένδυση που προστίθεται στο χαρτοφυλάκιο να εξετάζεται πως επηρεάζει το συνολικό ρίσκο και αναμενόμενο κέρδος του.

Το ρίσκο που έχει κάθε υποψήφιο χαρτοφυλάκιο ή επένδυση, προσεγγίστηκε με το covariance και έτσι προέκυψε ότι για ένα συγκεκριμένο αναμενόμενο κέρδος υπάρχει μία ελάχιστη τιμή ρίσκου και αντίστροφα.

6.2.1 Μαθηματικός Ορισμός

Η μαθηματική έκφραση του μοντέλου φαίνεται παρακάτω:

Min-> $\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} * x_i * x_j$ με περιορισμούς:

Σχέση 6.1

$$\sum_{j=1}^n \bar{r}_j * x_j \geq \alpha * B$$

Σχέση 6.2

$$\sum_{j=1}^n x_j = 1$$

Σχέση 6.3

$$0 \leq x_j, j = 1, \dots, n$$

Σχέση 6.4

Όπου σ_{ij} είναι το covariance των μετοχών από τις οποίες θα απαρτίζεται το portfolio, \bar{r}_j είναι το μέσο return της μετοχής j , x_j είναι το ποσοστό του portfolio που έχει επενδυθεί στη μετοχή j , α είναι το ελάχιστο return και B είναι το συνολικό ποσό που υπάρχει για επένδυση.

Επομένως, το μοντέλο γραμμικού προγραμματισμού, δείχνει πως για ένα ελάχιστο ποσοστιαίο return α μπορούμε να αναζητήσουμε το portfolio με το ελάχιστο ρίσκο.

Για να γίνεται πιο εύκολα αντιληπτό το συμπέρασμα της θεωρίας μπορούμε να τοποθετήσουμε κάθε χαρτοφυλάκιο σε ένα καρτεσιανό επίπεδο όπου ο άξονας X είναι το ρίσκο και ο άξονας Y είναι το αναμενόμενο κέρδος. Έτσι θα δημιουργηθεί μία καμπύλη η οποία αποτελείται από το βέλτιστο χαρτοφυλάκιο για κάθε τιμή του X , δηλαδή για κάθε τιμή του ρίσκου θα μπορούμε εύκολα να βρούμε την μέγιστη τιμή του αναμενόμενου κέρδους για το ρίσκο αυτό.

Στη συνέχεια θα αναλυθεί ο τρόπος υλοποίησης του μοντέλου προγραμματιστικά και τα συμπεράσματα που προκύπτουν από αυτό.

6.2.2 Υλοποίηση

Αξίζει να σημειωθεί πως δεν χρησιμοποιήθηκαν τα έτη 2013-2019 αυτούσια. Χρησιμοποιήθηκαν τα δεδομένα από τα έτη αυτά και στη συνέχεια η μέση τιμή της επταετίας ώστε να δημιουργηθούν τα returns και το covariance κάθε μετοχής σε ένα τυπικό έτος.

Αρχικά συγκεντρώθηκε το data για όλα τα έτη και όλες τις μετοχές στην ίδια δομή δεδομένων και υπολογίστηκαν οι τιμές για τα daily returns, τα daily covariance των μετοχών στο χρονικό διάστημα αυτό και στη συνέχεια τα μέσα ετήσια return και covariance. Όλες οι εντολές που χρησιμοποιήθηκαν είναι από τις βιβλιοθήκες numpy και pandas.

Τα daily returns υπολογίζονται ως η ποσοστιαία διαφορά ανάμεσα στην τιμή μίας μετοχής για 2 διαδοχικές μέρες. Το πρώτο στοιχείο θα είναι NaN γιατί για την πρώτη ημέρα του dataset δεν μπορεί να υπολογιστεί η ποσοστιαία διαφορά καθώς δεν υπάρχει προηγούμενη

μέρα. Αυτό δεν επηρεάζει το μοντέλο καθώς η μία μέρα στα 7 χρόνια είναι πολύ μικρό ποσοστό και μπορούμε να την αγνοήσουμε.

Μετοχή	Ετήσιο Κέρδος
MMM	13.17%
ABBV	21.1%
YUM	14.81%
ZION	16.82%

Πίνακας 11: Μέσο ετήσιο Κέρδος Μετοχών για τα έτη 2013-2019

	MMM	ABMD	ACN
MMM	0.0333	0.0178	0.0175
ABMD	0.0178	0.0752	0.0206
ACN	0.0175	0.0206	0.0362

Πίνακας 12: Covariance Μετοχών

Οι δύο πίνακες που φαίνονται παραπάνω είναι οι πίνακες οι οποίοι θα χρησιμοποιηθούν στα μοντέλα. Πρακτικά αντιπροσωπεύουν τις τιμές ενός τυπικού χρόνου για την επταετία 2013-2019. Έτσι ώστε, τα αποτελέσματα να μην εξαρτώνται από μία συγκεκριμένη χρονιά.

Δομή Επανάληψης

Στη συνέχεια δημιουργήθηκε μία δομή επανάληψης η οποία κατασκευάζει πολλά τυχαία χαρτοφυλάκια.

Η δομή επανάληψης κάθε φορά που τρέχει:

1. Δημιουργεί τυχαία βάρη τα οποία έχουν άθροισμα 1 και αντιπροσωπεύουν το ποσοστό της επένδυσης για κάθε μετοχή από τα δεδομένα
2. Υπολογίζει το return και το risk του χαρτοφυλακίου
3. Τοποθετεί τα βάρη, την απόδοση και το ρίσκο σε δομές δεδομένων

Έτσι με κάθε επανάληψη δημιουργείται και αποθηκεύεται ένα χαρτοφυλάκιο.

Για τον υπολογισμό του return ανά μετοχή πολλαπλασιάζεται το βάρος της μετοχής, δηλαδή το ποσοστό του χαρτοφυλακίου που έχει επενδυθεί στη συγκεκριμένη μετοχή, με την απόδοση της μετοχής από τον πίνακα annual returns.

Ο υπολογισμός του ρίσκου είναι πιο σύνθετος. Παρακάτω φαίνονται αναλυτικά οι πράξεις πινάκων που πραγματοποιούνται.

$$\sigma = \sqrt{w_1^2 * \sigma_1^2 + 2 * w_1 * w_2 * \sigma_{1,2} + w_2^2 * \sigma_2^2} \quad \text{οπότε}$$

Σχέση 6.5

$$\sigma^2 = [w_1 \quad w_2] \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_2^2 \end{bmatrix} [w_1 \quad w_2]$$

Σχέση 6.6

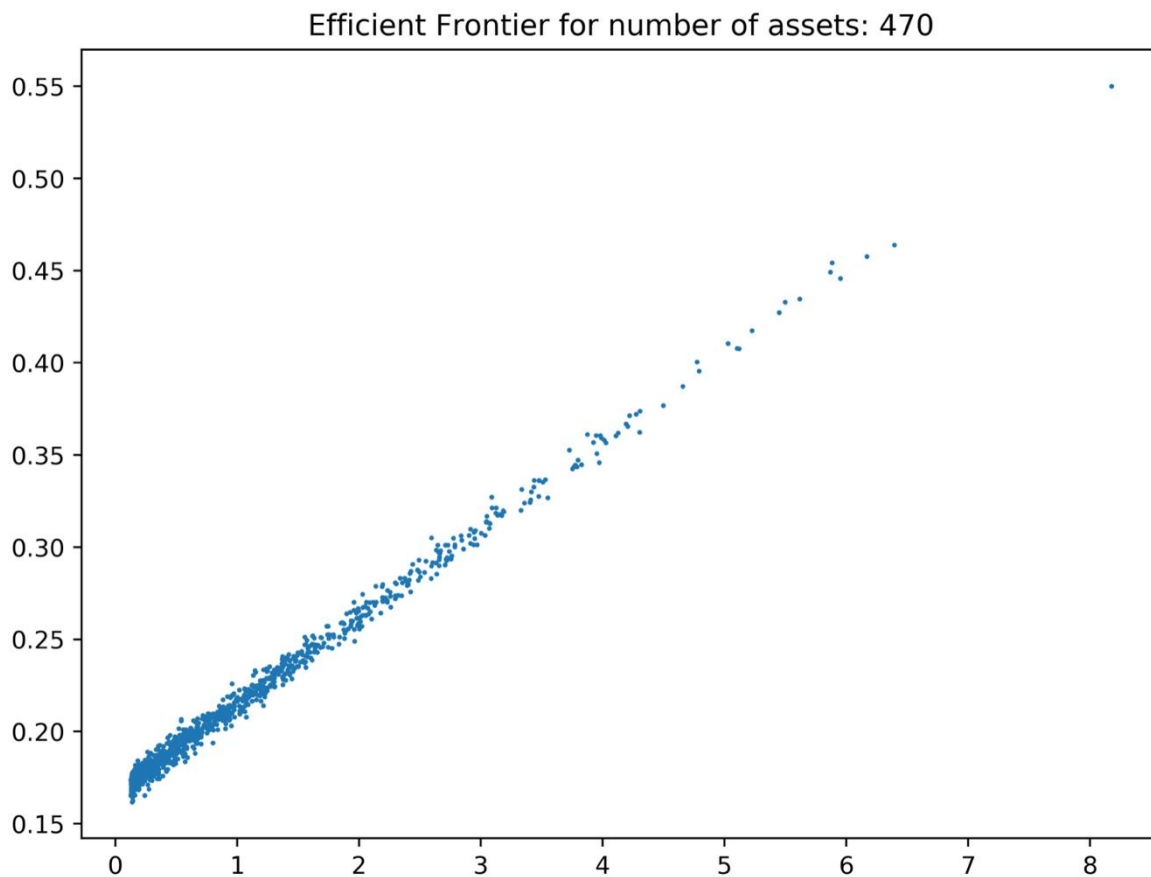
Έτσι υπολογίζεται το ρίσκο όπως φαίνεται και από τον κώδικα.

Η επανάληψη λοιπόν, δημιουργεί για τις 469 μετοχές τυχαία βάρη, υπολογίζει το return και το ρίσκο και τα τοποθετεί σε μια δομή δεδομένων. Η δομή δεδομένων επίσης περιέχει μία στήλη για κάθε μετοχή όπου φαίνεται το ποσοστό του χαρτοφυλακίου που έχει επενδυθεί στη μετοχή. Στον παρακάτω πίνακα φαίνονται ενδεικτικά τα πέντε πρώτα τυχαία χαρτοφυλάκια που δημιουργήθηκαν.

	Returns	Risk	MMM	ABT	ABBV	ABMD
1	20.1%	0.139732	1.83%	0.98%	3.96%	1.64%
2	19.2%	0.141216	1.11%	3.46%	1.08%	2.16%
3	19.5%	0.135443	1.0%	4.05%	3.19%	3.83%
4	18.2%	0.138487	2.4%	3.27%	1.69%	2.22%
5	20%	0.134208	2.3%	3.81%	0.96%	2.69%

Πίνακας 13:Μορφολογία δομής δεδομένων τυχαίων χαρτοφυλακίων. Ο πλήρης πίνακας αποτελεσμάτων διαθέτει 1000 γραμμές (μία για κάθε χαρτοφυλάκιο) και 469 στήλες με τα βάρη των μετοχών

Ενδιαφέρον έχει να σχεδιαστούν τα αποτελέσματα της δομής επανάληψης. Στον άξονα x τοποθετείται το ρίσκο και στον άξονα y τοποθετείται το κέρδος για το κάθε χαρτοφυλάκιο. Κάθε σημείο είναι ένα από τα 1000 τυχαία χαρτοφυλάκια που κατασκευάστηκαν.



Σχήμα 37: Διάγραμμα Κέρδους Ρίσκου για τυχαία Χαρτοφυλάκια

Από το διάγραμμα παρατηρείται πως με την αύξηση του ρίσκου υπάρχει και ανάλογη αύξηση του κέρδους.

Χρήση Solver

Όσες επαναλήψεις και να πραγματοποιηθούν με τυχαίες τιμές στα βάρη μπορεί να μην προκύψει η βέλτιστη λύση του προβλήματος. Έτσι, για να λυθεί το παραπάνω πρόβλημα θα πρέπει να γίνει χρήση κάποιου solver.

Εδώ χρησιμοποιήθηκε ο `optimize.minimize` από τη βιβλιοθήκη `scipy`. Επειδή όπως φαίνεται το ρίσκο, προγραμματιστικά περιέχει αρκετές εντολές ορίστηκε ως συνάρτηση. Επομένως, κρίνεται απαραίτητο ο solver που θα χρησιμοποιηθεί να μπορεί να δεχθεί ως αντικειμενική συνάρτηση, μία συνάρτηση ορισμένη στο πρόγραμμα. Μία μέθοδος η οποία πληροί το χαρακτηριστικό αυτό είναι η SLSQP.

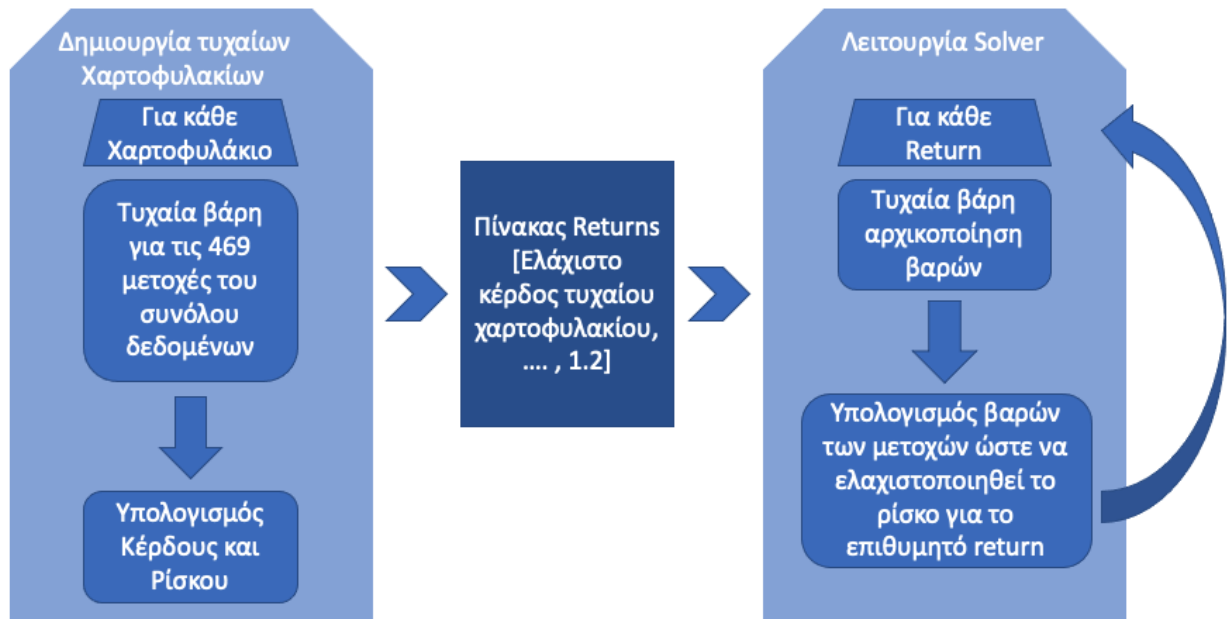
Ο solver αρχικοποιείται ως εξής:

- Τοποθετούνται τυχαία βάρη, τα οποία όλα έχουν άθροισμα 1. Αυτό είναι απαραίτητο διότι το βάρος μίας μετοχής αντιπροσωπεύει το ποσοστό του χαρτοφυλακίου που έχει επενδυθεί στην μετοχή αυτή.
- Ορίζονται τα όρια για την τιμή του κάθε βάρους να ανήκουν στο $[0,1]$

Η λειτουργία που πραγματοποιεί ο solver είναι να λύνει το πρόβλημα του γραμμικού προγραμματισμού. Δηλαδή, για ένα συγκεκριμένο return ελαχιστοποιεί το ρίσκο. Έτσι είναι απαραίτητο να οριστεί το return αυτό.

Ορίζεται η λίστα επιθυμητών κερδών, η οποία περιέχει τα returns για τα οποία ο solver θα λύσει το πρόβλημα. Φτιάχεται ως πίνακας που ξεκινάει από το ελάχιστο return που βρέθηκε με την επανάληψη τυχαίων χαρτοφυλακίων μέχρι το 1.2 και τοποθετεί στο διάστημα αυτό έναν αριθμό σημείων N.

Έτσι, ο solver θα τρέξει N φορές και κάθε φορά θα ελαχιστοποιεί το ρίσκο για το δεδομένο return. Το άνω άκρο (1.2) επιλέχτηκε με τέτοιο τρόπο ώστε να είναι μεγαλύτερο από το καλύτερο return που δίνει η δομή επανάληψης έτσι ώστε να παρατηρηθεί εάν ο solver δίνει καλύτερα αποτελέσματα. Στο παρακάτω σχήμα φαίνεται η διαδικασία δημιουργίας τυχαίων χαρτοφυλακίων και ο τρόπος λειτουργίας του solver για την ελαχιστοποίηση του ρίσκου για κάθε επιθυμητό κέρδος.

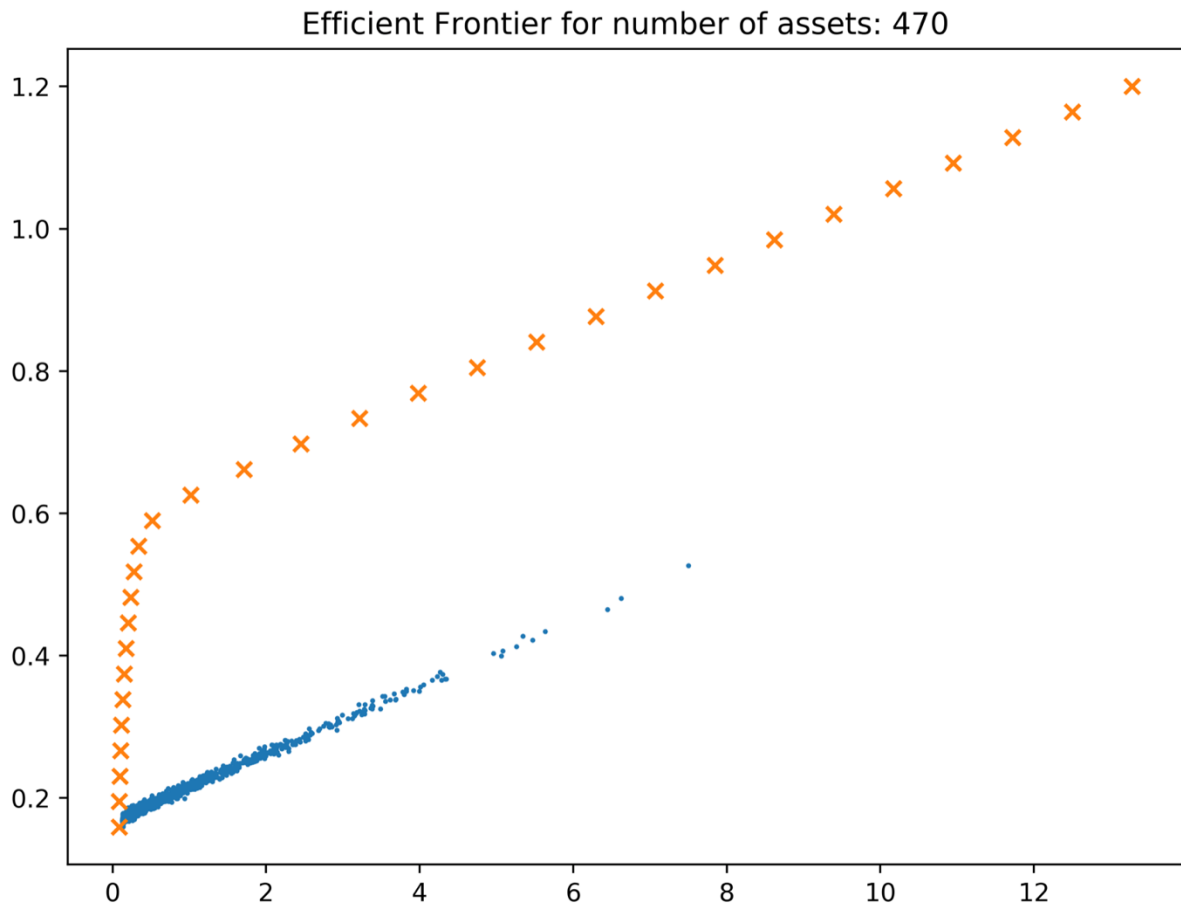


Σχήμα 38: Συνολική Διαδικασία Υλοποίησης

Τα αποτελέσματα του solver και της δομής επανάληψης φαίνονται στην παρακάτω εικόνα.

Τα πορτοκαλί X είναι τα χαρτοφυλάκια που φτιάχτηκαν με τον solver και τα μπλε σημεία είναι τα τυχαία χαρτοφυλάκια της δομής επανάληψης. Παρατηρείται πως για κάθε σημείο του άξονα x, για κάθε τιμή ρίσκου, το return των χαρτοφυλακίων που παράχθηκαν από τον solver είναι μεγαλύτερο.

Η καμπύλη που σχηματίζουν τα πορτοκαλί σημεία ονομάζεται Efficient Frontier και δείχνει για κάθε τιμή του return ποιο είναι το ελάχιστο ρίσκο που μπορεί το χαρτοφυλάκιο να έχει. Αποτελεί το πειραματικό συμπέρασμα της θεωρίας του Markowitz η οποία αποτέλεσε θεμέλιο του Portfolio Optimization.



Σχήμα 39: Efficient Frontier

6.3 Το Μοντέλο του Young

6.3.1 Μαθηματική Περιγραφή

Το πρόβλημα γραμμικού προγραμματισμού είναι το εξής:

- Max- $\rightarrow M_p$

- με περιορισμούς:

$$\sum_{j=1}^N w_j * y_{jt} - M_p \geq 0$$

Σχέση 6.7

- $\sum_{j=1}^N w_j \leq W$

Σχέση 6.8

$$\bullet \quad \sum_{j=1}^N w_j * \overline{y_{jt}} \geq G$$

Σχέση 6.9

Αξίζει να σημειωθεί πως τα βάρη των μετοχών w_j μπορεί να είναι φραγμένα σε περίπτωση που κάποιος επενδυτής δεν επιθυμεί να επενδύσει πολύ μεγάλο ποσοστό του χαρτοφυλακίου του σε μία συγκεκριμένη μετοχή.

Όπου y_{jt} είναι το return της μετοχής j σε χρονική περίοδο t , w_j το ποσοστό του χαρτοφυλακίου που έχει επενδυθεί στη μετοχή j , M_r το ελάχιστο return για το χαρτοφυλάκιο, W το συνολικό κεφάλαιο που διαθέτει το portfolio και G μία ελάχιστη τιμή για το μέσο return του χαρτοφυλακίου.

Από τον ορισμό του προβλήματος παρατηρείται πως στον πρώτο περιορισμό περιέχεται ο όρος y_{jt} , το κέρδος δηλαδή για μία συγκεκριμένη χρονική περίοδο. Άρα εφόσον το dataset αποτελείται από 7 διαφορετικά χρόνια, χρειάζεται για κάθε έτος να υπάρχει μία ξεχωριστή σχέση.

Επιπλέον από τον ορισμό φαίνεται ότι υπάρχει η δυνατότητα να αλλάξουν κάποια από τα μεγέθη του προβλήματος χωρίς να επηρεαστεί η μαθηματική έκφραση του προβλήματος. Τα μεγέθη αυτά είναι το G , δηλαδή το ελάχιστο απαιτούμενο κέρδος, το N , δηλαδή ο αριθμός των μετοχών από τις οποίες το μοντέλο μπορεί να επιλέξει για να κατασκευάσει το βέλτιστο χαρτοφυλάκιο και το ανώτατο φράγμα των w_j .

Για αυτό το λόγο μπορούμε να θεωρήσουμε πως τα μεγέθη αυτά είναι κατά κάποιο τρόπο οι αρχικές συνθήκες του μοντέλου, τις οποίες εάν αλλάξουμε θα καταλήξουμε σε διαφορετικά αποτελέσματα. Από τη σύγκριση των αποτελεσμάτων που θα γίνει στη συνέχεια θα δούμε πως η παρατήρηση αυτή είναι πολύ σημαντική.

6.3.2 Υλοποίηση του μοντέλου

Το πρώτο βήμα για την υλοποίηση του μοντέλου ήταν ο υπολογισμός του κέρδους για κάθε μετοχή για την επταετία 2013-2019. Αυτός πραγματοποιήθηκε ως εξής:

1. Για κάθε έτος του dataset και για κάθε μετοχή διατηρείται μόνο η πρώτη και η τελευταία ημέρα του έτους
2. Υπολογισμός της ποσοστιαίας διαφοράς μεταξύ των ημερών αυτών
3. Διατήρηση ποσοστιαίων διαφορών της τελευταίας ημέρας κάθε έτους. Έτσι υπολογίζεται το κέρδος κάθε μετοχής για κάθε χρόνο.

Με βάση όμως τον τελευταίο περιορισμό του προβλήματος χρειάζεται το μέσο κέρδος για την επταετία, το οποίο υπολογίζεται εύκολα από τα ετήσια returns της κάθε μετοχής.

6.3.3 Μορφοποίηση του προβλήματος

Ο πίνακας W περιέχει τα βάρη των μετοχών, δηλαδή το ποσοστό του χαρτοφυλακίου που θα επενδυθεί σε κάθε μετοχή. Ωστόσο το W περιέχει και την ποσότητα Mp .

Αντικειμενική συνάρτηση: Maximize $\rightarrow C * W$ έτσι ώστε $C * W = Mp$ οπότε ο πίνακας c θα πρέπει να έχει την μορφή $C = [0, 0, \dots, 0, 1]$.

Οι περιορισμοί θα πρέπει να δημιουργούνται από την εξίσωση: $A * W \leq b$

Για τους περιορισμούς υπάρχουν 3 κατηγορίες:

$$\sum_{j=1}^N w_j * y_{jt} - Mp \geq 0$$

- Για 7 έτη θα έχουμε 7 περιορισμούς αυτής της μορφής

$$\sum_{j=1}^N w_j \leq W$$

- Το συνολικό ποσό επένδυσης να μην ξεπερνά το διαθέσιμο κεφάλαιο
- Θα έχουμε 1 περιορισμό

$$\sum_{j=1}^N w_j * \bar{y}_{jt} \geq G$$

- Το κέρδος του χαρτοφυλακίου να είναι μεγαλύτερο από μία ελάχιστη τιμή
- Θα έχουμε 1 περιορισμό

Επίσης μπορεί να υπάρξει ανώτατο όριο στο ποσοστό του χαρτοφυλακίου που θα επενδυθεί ανά μετοχή και έτσι τα βάρη των μετοχών πρέπει να είναι φραγμένα.

Οπότε για τους 7 πρώτους περιορισμούς ο πίνακας A θα πρέπει να έχει τη μορφή $A = [a_1, a_2, a_3, \dots, a_n, -1]$ όπου: $a_{1\dots n}$ είναι το return της κάθε μετοχής για το συγκεκριμένο έτος και -1 είναι ο συντελεστής της ποσότητας M_p .

Για τον 8^ο περιορισμό θα πρέπει να έχει την μορφή: $A = [1, 1, 1, \dots, 1, 0]$ όπου κάθε μετοχή έχει συντελεστή 1 και η ποσότητα M_p έχει συντελεστή 0.

Για τον 9^ο περιορισμό θα πρέπει να έχει την μορφή: $A = [a_1, a_2, a_3, \dots, a_n, 0]$ όπου: $a_{1\dots n}$ είναι το μέσο return της κάθε μετοχής για την επταετία και 0 είναι ο συντελεστής της ποσότητας M_p .

6.3.4 Αποτέλεσμα

Στη συνέχεια αφού το πρόβλημα είχε μορφοποιηθεί και μπορούσαν να υπολογιστούν οι πίνακες για κάθε N, το πρόβλημα λύθηκε με τη χρήση του solver rymprog. Πραγματοποιήθηκαν εκτελέσεις του μοντέλου με διαφορετικές αρχικές συνθήκες και προέκυψε ο πίνακας αποτελεσμάτων.

Οι συνθήκες που άλλαζαν από εκτέλεση σε εκτέλεση ήταν ο αριθμός των μετοχών από τα οποία μπορεί να επιλέξει το μοντέλο (N), το ανώτατο ποσοστό για τα βάρη (U) και το ελάχιστο return (G). Τα αποτελέσματα που φαίνονται είναι η τιμή της αντικειμενικής συνάρτησης Z, το επιπλέον κέρδος του χαρτοφυλακίου (Additional Return) και το τελικό κέρδος του χαρτοφυλακίου (Final return).

(N,U,G)	50, 45%, 30%	100, 40%, 15%	200, 20%, 20%	300, 10%, 10%	470, 50%, 40%
Z	0.395390	0.336919	0.315974	0.427913	0.522928

Additional Average Return	56.01%	41.45%	37.91%	43.53%	61.48%
Final Average Return	86.1%	56.45%	57.91%	53.53%	101.48%

Πίνακας 14: Αποτελέσματα Βελτιστοποίησης Χαρτοφυλακίου με το Μοντέλο του Young

Κεφάλαιο 7: Αποτελέσματα Επενδύσεων

7.1 Εισαγωγή

Στο κεφάλαιο αυτό θα αναλυθεί ο συνδυασμός των προβλέψεων που πραγματοποιήθηκαν στο κεφάλαιο 5 με τα μοντέλα Markowitz και Young. Θα εξεταστεί πως θα μπορούσε να υλοποιηθεί μία πλήρως αυτοματοποιημένη στρατηγική επενδύσεων για το 2019.

Θεωρώντας λοιπόν ότι μόλις έχει τελειώσει το ημερολογιακό έτος 2018 και ξεκινά το 2019, με βάση τα όσα αναπτύχθηκαν στα προηγούμενα κεφάλαια πως θα πρέπει να κατασκευαστεί ένα τέτοιο μοντέλο;

Η στρατηγική αποτελείται από δύο μέρη, το πρώτο προβλέπει τις τιμές των μετοχών σε επίπεδο ημέρας και το δεύτερο αξιοποιεί τις προβλέψεις για την δημιουργία χαρτοφυλακίων. Συγκεκριμένα, για κάθε ημέρα του νέου έτους το πρώτο μέρος της στρατηγικής, χρησιμοποιώντας μόνο τις τιμές κλεισίματος των προηγούμενων ημερών, θα προβλέπει την τιμή κλεισίματος της επόμενης. Τότε, το δεύτερο μοντέλο για κάθε μέρα του 2019, με δεδομένα τις προβλέψεις του πρώτου, θα δημιουργεί ένα χαρτοφυλάκιο.

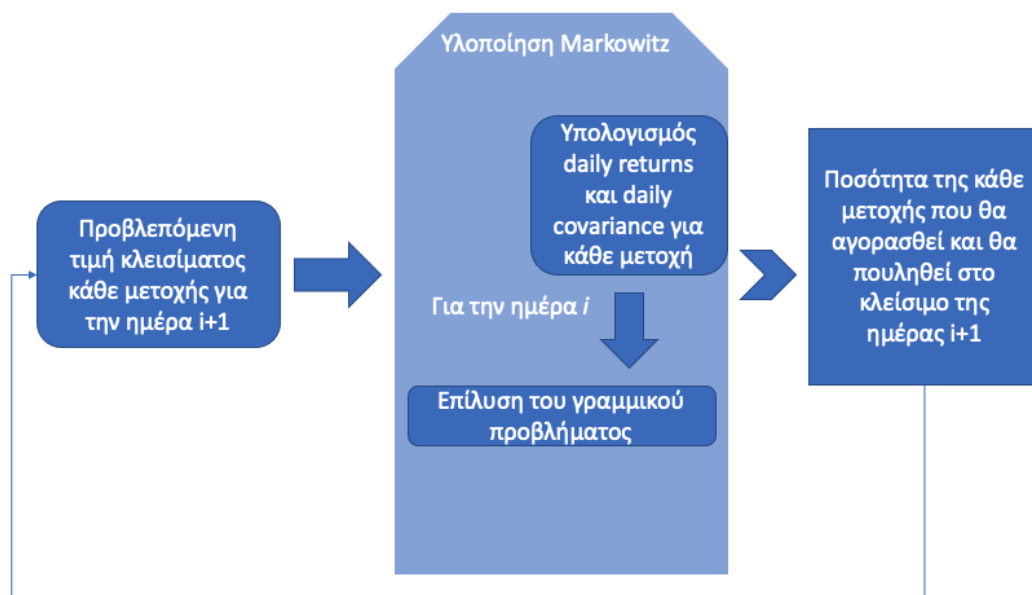
Για την εκτίμηση της κερδοφορίας της στρατηγικής έχουν γίνει οι εξής υποθέσεις:

- α. Το μέγεθος σύγκρισης της απόδοσης είναι η απόδοση του δείκτη S&P 500 για το 2019, η οποία ήταν 28.9%
- β. Η στρατηγική «αγοράζει» τις μετοχές που περιέχει το χαρτοφυλάκιο στην τιμή κλεισίματος της τελευταίας μέρας
- γ. Η στρατηγική «πουλάει» τις μετοχές του χαρτοφυλακίου στην τιμή κλεισίματος της επόμενης

Για τη διαδικασία των προβλέψεων θα χρησιμοποιηθεί το καλύτερο μοντέλο για το 2018, δηλαδή νευρωνικό δίκτυο τύπου MLP Regressor με ορίζοντα πρόβλεψης τις 8 ημέρες και αρχιτεκτονική της μορφής: `{(32,), relu, lbfgs, 200}`.

7.2 Στρατηγική με χρήση Markowitz

Δεδομένου ότι υπάρχουν τα δεδομένα για το 2019 ένας τρόπος για την υλοποίηση της στρατηγικής είναι να γίνουν πρώτα οι προβλέψεις με τη βοήθεια του νευρωνικού δικτύου και στη συνέχεια να αξιοποιηθούν από το μοντέλο του Markowitz. Η διαδικασία που πραγματοποιεί το μοντέλο του Markowitz φαίνεται στο παρακάτω διάγραμμα.

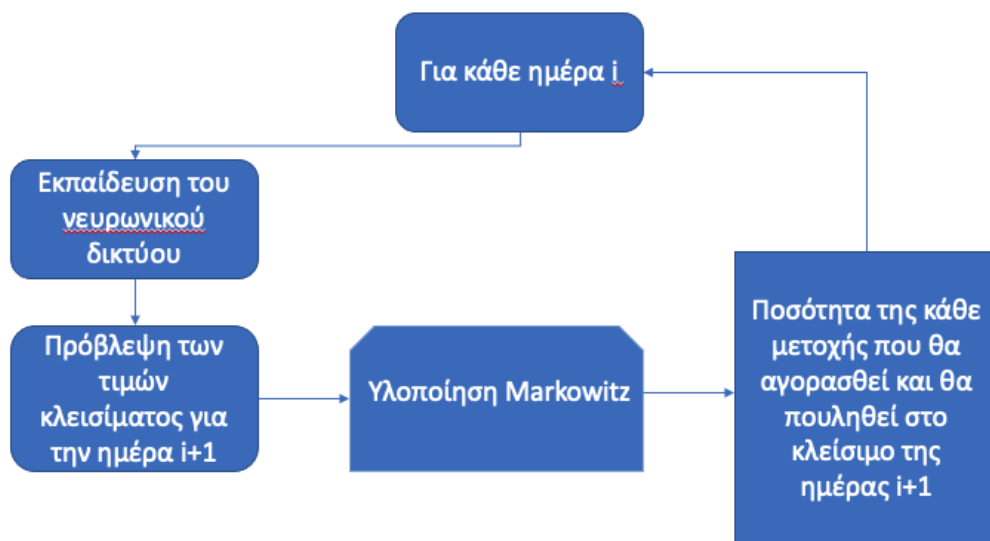


Σχήμα 40: Διάγραμμα Στρατηγικής με χρήση Markowitz

Η διαδικασία αυτή πραγματοποιείται για κάθε μέρα από τις 252 του έτους 2019 για τις οποίες η αγορά λειτούργησε. Παρόλο που δεν υπάρχει κάποιο λάθος στη συγκεκριμένη προσέγγιση εφόσον το μοντέλο δεν χρησιμοποιεί δεδομένα επόμενων ημερών από αυτή που εξετάζει κάθε φορά, υλοποιήθηκε και μία δεύτερη προσέγγιση με βάση την οποία:

- α. Οι προβλέψεις για τις τιμές της ημέρας $i+1$ δεν έχουν γίνει εκ των προτέρων αλλά πραγματοποιούνται σε πραγματικό χρόνο κατά την ημέρα i .
- β. Το νευρωνικό δίκτυο εκπαιδεύεται εκ νέου κάθε ημέρα i στα δεδομένα που έχει ήδη εκπαιδευτεί πριν και την ημέρα i .

Η νέα προσέγγιση φαίνεται στο παρακάτω σχήμα:



Σχήμα 41: Διάγραμμα Στρατηγικής με χρήση Markowitz και εκ νέου εκπαίδευση

Αποτέλεσμα της παραπάνω διαδικασίας είναι η δημιουργία ενός χαρτοφυλακίου για κάθε ημέρα του έτους.

Ιδιαίτερη σημασία παρουσιάζει το μέγιστο ποσοστό που μπορεί να επενδυθεί σε μία συγκεκριμένη μετοχή. Το συγκεκριμένο ανώτατο όριο ορίζεται από τον επενδυτή ανάλογα με τη διαφοροποίηση που επιθυμεί να έχουν τα ημερήσια χαρτοφυλάκια. Στο συγκεκριμένο πείραμα το όριο αυτό τέθηκε στο 20%

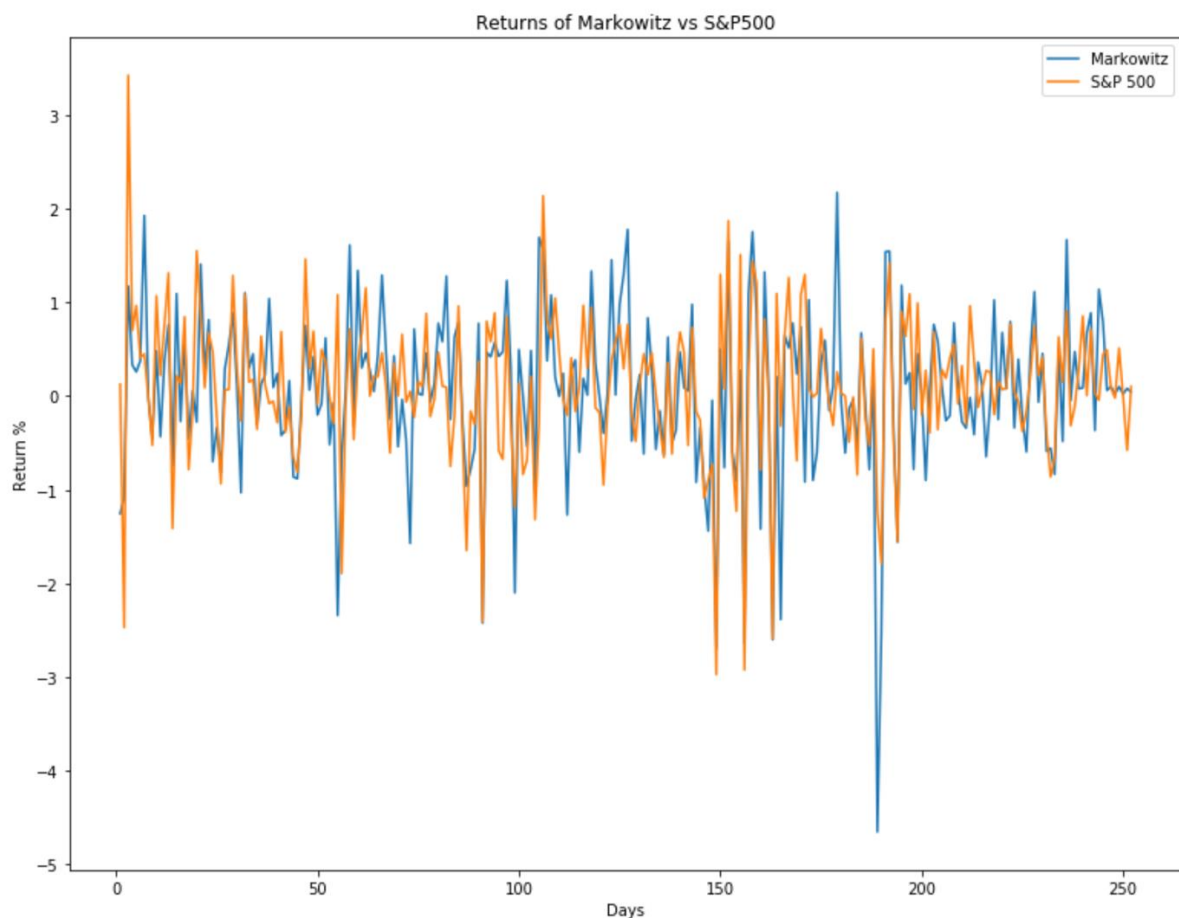
Ενδεικτικά το χαρτοφυλάκιο της 10^{ης} ημέρας φαίνεται στον παρακάτω πίνακα:

Μετοχή	Ποσοστό του Χαρτοφυλακίου που επενδύεται
C	3.911%
CBOE	8.4642%
MKTX	20%
MO	20%
PEP	7.6248%
STZ	20%
VZ	20%

Πίνακας 15: Ανάλυση Χαρτοφυλακίου 10ης μέρας 1ης στρατηγικής

Το κέρδος που σημείωσε η στρατηγική για την 10^η μέρα του 2019 είναι 0.4849%.

Η μέση απόδοση της στρατηγικής για το 2019 είναι 0.06224% ανά ημέρα. Συνολικά η στρατηγική, για όλο το έτος έχει κέρδος 15.3% και είναι κερδοφόρα. Παρόλο που έχει κέρδος, δεν κατάφερε να αποδώσει περισσότερο από τον S&P 500 και να «κερδίσει» την αγορά της οποίας η συνολική απόδοση για το έτος είναι 28.63% και η μέση ημερήσια απόδοσή της είναι 0.103%.



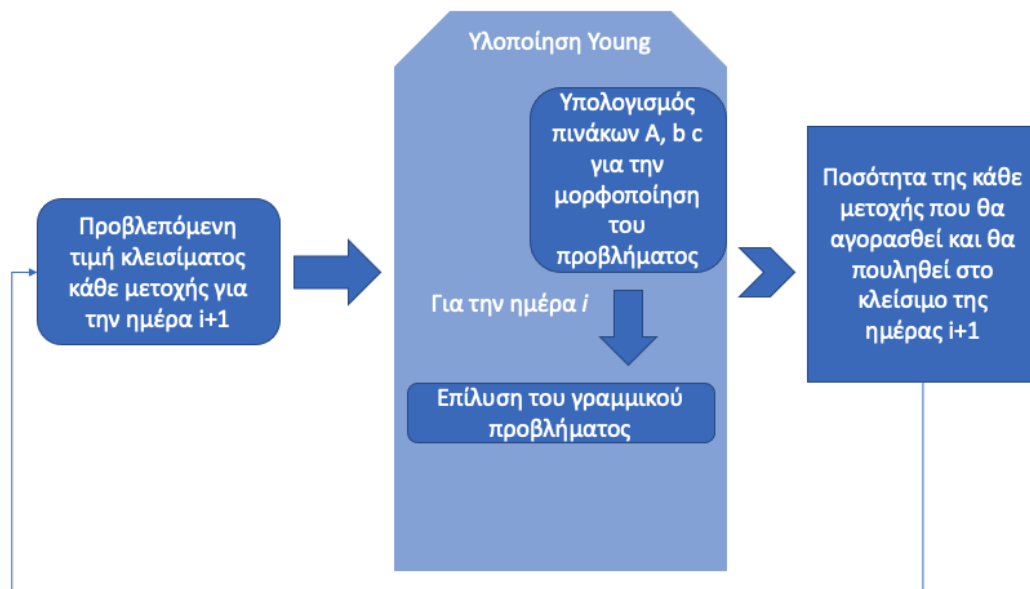
Σχήμα 42: Ημερήσια Κερδοφορία του S&P500 σε σχέση με την στρατηγική Markowitz

Στο παραπάνω σχήμα φαίνεται η σύγκριση της απόδοσης της στρατηγικής Markowitz με την απόδοση του S&P 500. Η καλύτερη απόδοση του S&P 500 είναι 3.44% και η χειρότερη -2.98% ενώ για τον Markowitz τα αντίστοιχα μεγέθη είναι 2.18% και -4.66%. Ο S&P 500 παρουσιάζει κερδοφορία για 150 ημέρες και το εύρος της κερδοφορίας για το 95% των ημερών είναι [-1.47%, 1.67%] ενώ η στρατηγική με Markowitz είναι κερδοφόρα για 149 μέρες και το αντίστοιχο εύρος είναι [-1.70%, 1.83%]. Έτσι, είναι προφανές πως παρόλο που οι δύο στρατηγικές είναι κερδοφόρες για τον ίδιο αριθμό ημερών στο έτος, η χαμηλότερη μέση

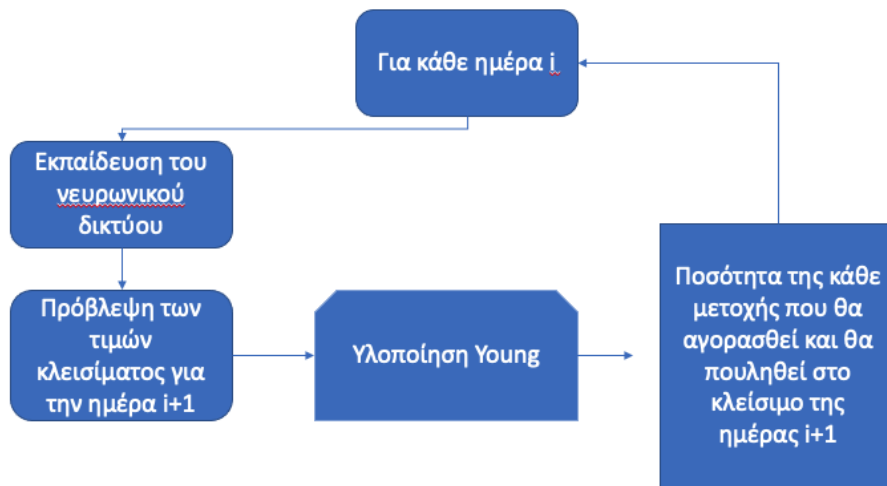
κερδοφορία της στρατηγικής με Markowitz οδηγεί σε διαφορά στην ετήσια απόδοση των δύο στρατηγικών.

7.3 Στρατηγική με χρήση Young

Αντίστοιχα με την προηγούμενη υλοποίηση, μπορούν να δημιουργηθούν δύο διαφορετικές προσεγγίσεις, οι οποίες δεν θα παρουσιάζουν σημαντικές διαφορές στα αποτελέσματα. Στις παρακάτω εικόνες φαίνονται η διαδικασία που υλοποιεί το μοντέλο του Young και η στρατηγική στην οποία οι προβλέψεις πραγματοποιούνται κάθε ημέρα για το έτος.



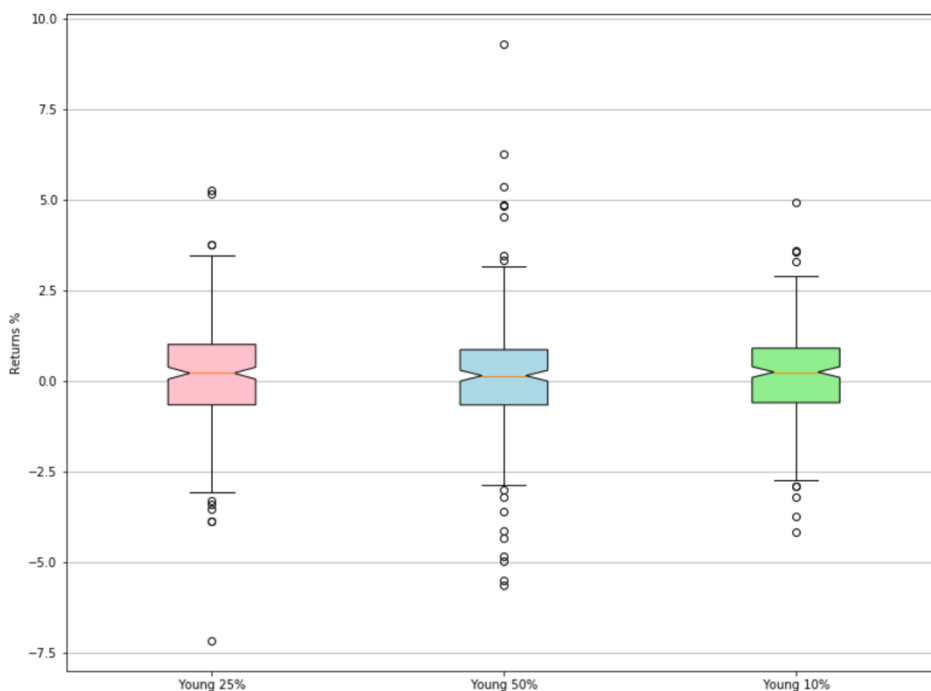
Σχήμα 43: Διάγραμμα Στρατηγικής με χρήση Young



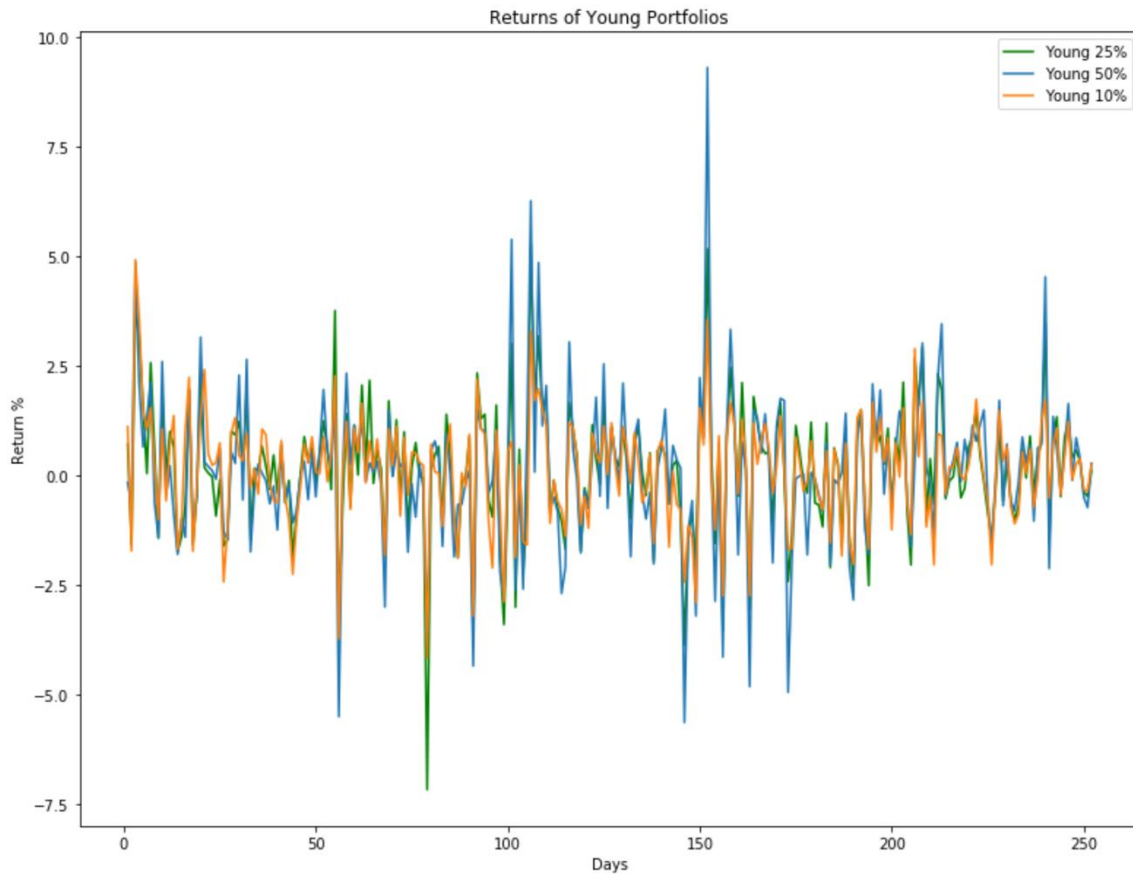
Σχήμα 44: Διάγραμμα Στρατηγικής με χρήση Young και εκ νέου εκπαίδευση

Αποτέλεσμα της παραπάνω διαδικασίας είναι η δημιουργία ενός χαρτοφυλακίου για κάθε ημέρα του έτους.

Σε αντίθεση με το μοντέλο του Markowitz, το μοντέλο του Young έχει περισσότερες αρχικές συνθήκες οι οποίες παρουσιάζουν ιδιαίτερη σημασία και οι τιμές που θα τεθούν επηρεάζουν το αποτέλεσμα. Για διαφορετικές τιμές των αρχικών συνθηκών η στρατηγική θα οδηγηθεί σε διαφορετικά χαρτοφυλάκια.



Σχήμα 45: Βοχρηlot διαφορετικών Χαρτοφυλακίων που προκύπτουν από την αλλαγή των αρχικών συνθηκών του Young



Σχήμα 46: Ημερήσια Returns διαφορετικών Χαρτοφυλακίων που προκύπτουν από Young με διαφορετικές αρχικές συνθήκες

Όπως φαίνεται από τα δύο παραπάνω σχήματα το χαρτοφυλάκιο που παράγεται από τη στρατηγική με χρήση του μοντέλου του Young επηρεάζεται αρκετά από τις αρχικές συνθήκες. Συγκεκριμένα, όσο μεγαλύτερο είναι το μέγιστο ποσοστό που μπορεί να επενδυθεί σε μία συγκεκριμένη μετοχή τότε τόσο μικρότερη είναι η διαφοροποίηση του χαρτοφυλακίου.

Τα χαρτοφυλάκια με μικρή διαφοροποίηση παρουσιάζουν μεγαλύτερη ευαισθησία σε «απότομες» μεταβολές στην τιμή των μετοχών στις οποίες έχουν επενδύσει. Έτσι, αναμένεται πως ένα χαρτοφυλάκιο που επενδύει σε μικρότερο αριθμό μετοχών θα παρουσιάζει περισσότερες πολύ θετικές και πολύ αρνητικές τιμές.

Αυτό επαληθεύεται και από τα παραπάνω σχήματα καθώς φαίνεται πως οι περισσότερες τιμές με τη μεγαλύτερη απόκλιση από τη μέση παρουσιάζονται στο χαρτοφυλάκιο που επενδύει μέχρι 50% σε μία μετοχή, ενώ το χαρτοφυλάκιο με μέγιστο ποσοστό επένδυσης σε

μία μετοχή το 25% παρουσιάζει λιγότερα outliers και το χαρτοφυλάκιο με 10% έχει τα λιγότερα outliers μεταξύ των τριών.

Η πιο κερδοφόρα προσέγγιση είναι αυτή στην οποία επενδύεται μέχρι και το 25% του χαρτοφυλακίου σε μία μεμονωμένη μετοχή και έτσι αυτή θα συγκριθεί με τον S&P 500.

Ενδεικτικά το χαρτοφυλάκιο της 246^{ης} ημέρας φαίνεται στον παρακάτω πίνακα:

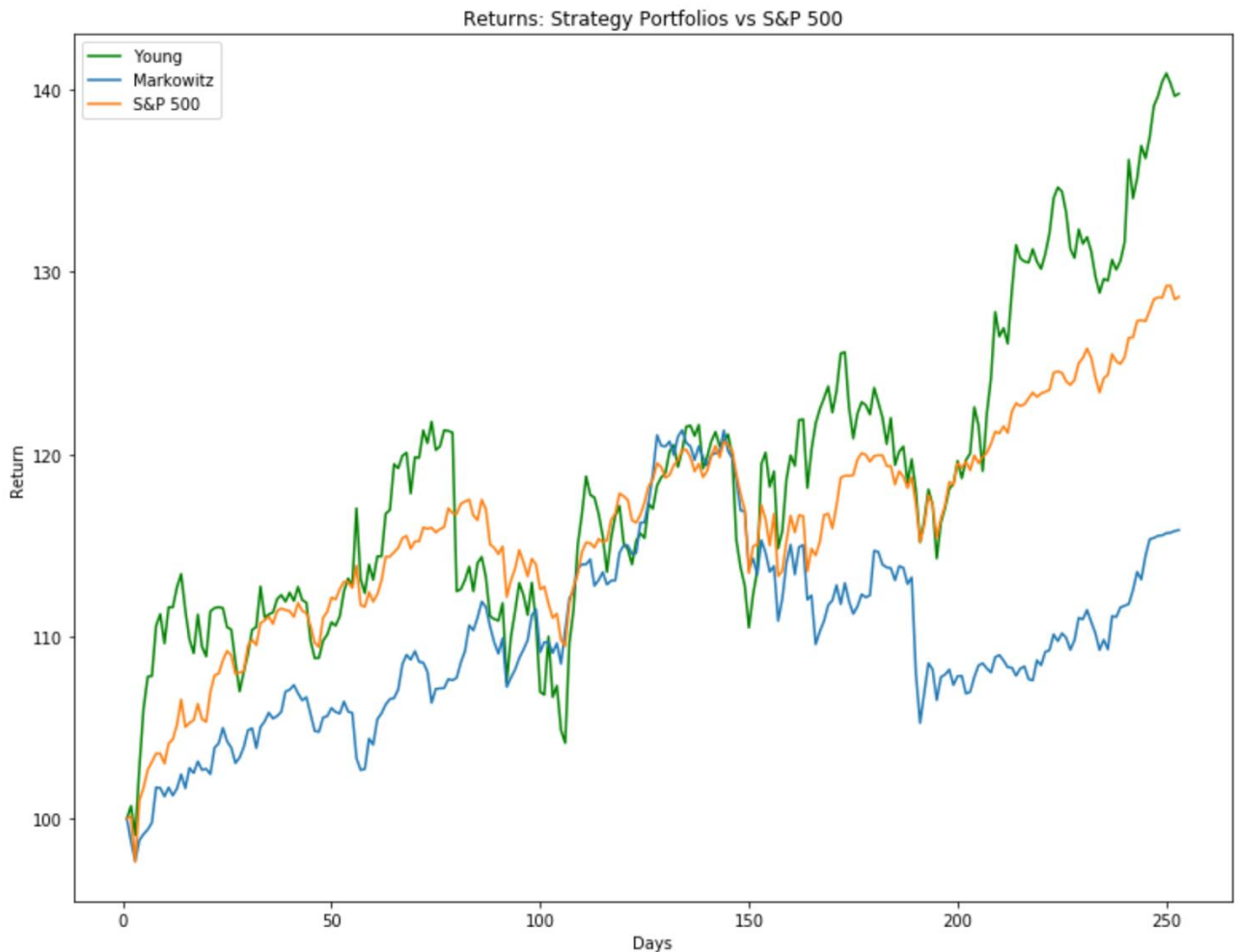
Μετοχή	Ποσοστό του Χαρτοφυλακίου που επενδύεται
AMD	25%
CMG	25%
TGT	25%
LRCX	25%

Πίνακας 16: Ανάλυση Χαρτοφυλακίου 246ης μέρας 2ης προσέγγισης

Το κέρδος της ημέρας αυτής είναι 1.25%.

Η μέση απόδοση της στρατηγικής για το 2019 είναι 0.144% ανά ημέρα. Συνολικά η στρατηγική, για όλο το έτος έχει κέρδος 39.78%. Επομένως, κατάφερε να αποδώσει περισσότερο από τον S&P και να κερδίσει την αγορά.

7.4 Σύνοψη Αποτελεσμάτων



Σχήμα 47: Κερδοφορία των δύο διαφορετικών προσεγγίσεων και του S&P 500 για το 2019

Όπως προκύπτει και από την αναλυτική περιγραφή των διαφορετικών προσεγγίσεων η πιο κερδοφόρα στρατηγική επενδύσεων είναι αυτή που βασίζεται στο μοντέλο βελτιστοποίησης χαρτοφυλακίου του Young.

Στο σχήμα 44 απεικονίζονται τα returns των δύο στρατηγικών σε σχέση με αυτά του S&P500. Ενδιαφέρον παρουσιάζει το γεγονός ότι παρόλο που η στρατηγική που βασίζεται στο Markowitz, έχει τελικά μικρή κερδοφορία υπάρχει μία χρονική στιγμή στη μέση του έτους 2019 στην οποία κερδίζει τόσο τον S&P 500 όσο και τη άλλη στρατηγική.

Επομένως, δεν είναι βέβαιο ότι εάν αλλάξει το χρονικό διάστημα που εξετάζεται ή το έτος στο οποίο γίνονται οι επενδύσεις, τα αποτελέσματα θα παραμείνουν τα ίδια και η στρατηγική που βασίζεται στο Young θα είναι η καλύτερη.

7.5 Μελλοντικές Προεκτάσεις

Μπορεί να υλοποιηθεί περαιτέρω έρευνα έτσι ώστε να δημιουργηθούν προεκτάσεις της συγκεκριμένης εργασίας.

Αρχικά μπορούν να εξεταστούν διαφορετικές προσεγγίσεις σε ότι αφορά τη δημιουργία των χαρτοφυλακίων και να αξιοποιηθούν και άλλα μοντέλα βελτιστοποίησης χαρτοφυλακίου. Ιδιαίτερο ενδιαφέρον θα έχει μία αντίστοιχη υλοποίηση με την μέθοδο των Κοππο και Yamazaki.

Επιπλέον, ενδιαφέρουσα θα είναι και η προέκταση της εργασίας για διαφορετικά έτη. Συγκεκριμένα, είναι δυνατό να πραγματοποιηθεί μία υλοποίηση όπου θα εξετάζονται περισσότερα έτη στο παρελθόν και θα γίνονται επενδύσεις σε περισσότερα από ένα έτη. Για παράδειγμα γίνεται να χρησιμοποιηθούν τα έτη 2012-2016 σαν training το έτος 2017 για testing και optimization και το έτος 2018 σαν έτος επένδυσης και αντίστοιχα να αξιοποιηθεί με παρόμοιο τρόπο και η επταετία 2014-2020 όπου έτος επενδύσεων θα είναι το τελευταίο. Έτσι, οι ίδιες αρχές λειτουργίας θα δημιουργήσουν επενδύσεις σε 3 διαφορετικά έτη 2018, 2019 και 2020. Με αυτό τον τρόπο μπορεί να εξετασθεί ποια στρατηγική και με ποιες αρχικές συνθήκες κερδίζει την αγορά με μεγαλύτερη συνέπεια.

Τέλος, στα πλαίσια επέκτασης της παρούσας διπλωματικής εργασίας, είναι δυνατόν να κατασκευαστεί ένα αυτοματοποιημένο σύστημα επενδύσεων το οποίο θα επικοινωνεί (μέσω κάποιου broker) με την αγορά και θα υλοποιεί επενδύσεις καθημερινά χωρίς να απαιτείται η επέμβαση του ίδιου του επενδυτή.

Επίλογος

Η παρούσα διπλωματική εργασία είχε ως στόχους:

- α. την πρόβλεψη των χρονοσειρών που απαρτίζονται από τις ημερήσιες τιμές των μετοχών που ανήκουν στο δείκτη Standard & Poor's 500.
- β. τη δημιουργία μίας αυτοματοποιημένης στρατηγικής επενδύσεων η οποία θα πρέπει να είναι κερδοφόρα και να έχει καλύτερη απόδοση από τον δείκτη S&P 500 για το ημερολογιακό έτος 2019.

Για την επίτευξη του πρώτου στόχου έγινε αρχικά μία διερεύνηση σχετικά με το αν είναι αποδοτικότερο κάθε μετοχή να προβλέπεται από διαφορετικό νευρωνικό δίκτυο ή εάν η χρήση του ίδιου δικτύου για όλο το σύνολο των μετοχών παράγει ακριβέστερες προβλέψεις. Στην πρώτη περίπτωση υπάρχει το πλεονέκτημα ότι κάθε δίκτυο είναι πλήρως προσαρμοσμένο στην κάθε χρονοσειρά, ενώ στην δεύτερη το δίκτυο έχει την ευκαιρία να αναγνωρίσει πρότυπα στην κίνηση των τιμών και να προβλέψει καλύτερα κάποιες μετοχές. Αποτέλεσμα της διερεύνησης αυτής ήταν η δημιουργία δύο πολύ αποδοτικών μοντέλων πρόβλεψης με MAPE: 1.24% για το πρώτο και 1.21% για το δεύτερο. Από την διαφορά αυτή συνεπάγεται πως η εφαρμογή ενός δικτύου σε όλο το σύνολο των μετοχών παράγει ελάχιστα καλύτερα αποτελέσματα.

Στη συνέχεια εξετάστηκε εάν κάποιο μοντέλο διαφορετικό του Multi Layer Perceptron μπορεί να προβλέψει ακριβέστερα τις χρονοσειρές των μετοχών. Συγκεκριμένα, εξετάστηκαν τα Δέντρα Απόφασης και τα Αναδρομικά Νευρωνικά Δίκτυα. Τελικά προκύπτει πως το MLP δίκτυο με απόδοση 1.21% είναι αποδοτικότερο για το συγκεκριμένο σύνολο δεδομένων.

Αξίζει να σημειωθεί πως για κάθε ένα από τα μοντέλα που εξετάστηκαν έγινε διερεύνηση για τις βέλτιστες παραμέτρους τους αλλά και το βέλτιστο ορίζοντα πρόβλεψης. Όλα τα αποτελέσματα που αναφέρονται παραπάνω, είναι προϊόν της βέλτιστης εκδοχής του κάθε μοντέλου.

Για την επίτευξη του δεύτερου στόχου της εργασίας χρησιμοποιήθηκαν δύο τεχνικές βελτιστοποίησης χαρτοφυλακίου. Σκοπός των τεχνικών αυτών είναι με βάση τις προβλέψεις

να δημιουργούν ένα χαρτοφυλάκιο για κάθε διαφορετική ημέρα του έτους 2019. Υπολογίζοντας την συνολική απόδοση για όλο το έτος φαίνεται η κερδοφορία της στρατηγικής.

Το πρώτο μοντέλο το οποίο βασίστηκε στη θεωρία βελτιστοποίησης χαρτοφυλακίου του Markowitz ήταν κερδοφόρο και σημείωσε κέρδος 15.3%. Το δεύτερο μοντέλο βασίστηκε στον τρόπο βελτιστοποίησης που προτάθηκε από τον Young και είχε κέρδος 59.90%. Παρόλο που και οι δύο προσεγγίσεις ήταν κερδοφόρες, μόνο η δεύτερη ήταν καλύτερη από τον S&P 500 ο οποίος για το 2019 σημείωσε κέρδος 28.9%.

Παράρτημα

A	AAL	AAP	AAPL	ABBV	ABC	ABMD	ABT	ACN	ADBE
ADI	ADM	ADP	ADS	ADSK	AEE	AEP	AES	AFL	AGN
AIG	AIV	AIZ	AJG	AKAM	ALB	ALGN	ALK	ALL	ALXN
AMAT	AMCR	AMD	AME	AMGN	AMP	AMT	AMZN	ANSS	ANTM
AON	AOS	APA	APD	APH	APTV	ARE	ATO	ATVI	AVB
AVGO	AVY	AWK	AXP	AZO	BA	BAC	BAX	BBY	BDX
BEN	BF-B	BIIB	BK	BKNG	BKR	BLK	BLL	BMY	BR
BRK-B	BSX	BWA	BXP	C	CAG	CAH	CAT	CB	COBE
CBRE	CCI	CCL	CDNS	CE	CERN	CF	CHD	CHRW	CHTR
CI	CINF	CL	CLX	CMA	CMCSA	CME	CMG	CMI	CMS
CNC	CNP	COF	COG	COO	COP	COST	CPB	CPRI	CPRT
CRM	CSCO	CSX	CTAS	CTL	CTSH	CTXS	CVS	CVX	CXO
D	DAL	DD	DE	DFS	DG	DGX	DHI	DHR	DIS
DISCA	DISCK	DISH	DLR	DLTR	DRE	DRI	DTE	DUK	DVA
DVN	DXC	EA	EBAY	ECL	ED	EFX	EIX	EL	EMN
EMR	EOG	EQIX	EQR	ES	ESS	ETFC	ETN	ETR	EVRG
EW	EXC	EXPD	EXPE	EXR	F	FANG	FAST	FB	FBHS
FCX	FDX	FE	FFIV	FIS	FISV	FITB	FLIR	FLS	FLT
FMC	FRC	FRT	FTI	FTNT	GD	GE	GILD	GIS	GL
GLW	GM	GOOG	GOOGL	GPC	GPN	GPS	GRMN	GS	GW
HAL	HAS	HBAN	HBI	HCA	HD	HES	HFC	HIG	HII
HOG	HOLX	HON	HP	HPQ	HRB	HRL	HSIC	HST	HSY
HUM	IBM	ICE	IDXX	IEX	IFF	IILMN	INCY	INTC	INTU
IP	IPG	IPGP	IRM	ISRG	IT	ITW	IVZ	J	JBHT
JCI	JKHY	JNJ	JNPR	JPM	JWN	K	KEY	KIM	KLAC
KMB	KMI	KMX	KO	KR	KSS	KSU	L	LB	LDOS
LEG	LEN	LH	LHX	LIN	LKQ	LLY	LMT	LNC	LNT
LOW	LRCX	LUV	LVS	LYB	LYV	MA	MAA	MAR	MAS
MCD	MCHP	MCK	MCO	MDLZ	MDT	MET	MGM	MHK	MKC
MKTX	MLM	MMC	MMM	MNST	MO	MOS	MPC	MRK	MRO
MS	MSCI	MSFT	MSI	MTB	MTD	MU	MXIM	MYL	NBL
NDAQ	NEE	NEM	NFLX	NI	NKE	NLSN	NOC	NOV	NOW
NRG	NSC	NTAP	NTRS	NUE	NVDA	NVR	NWL	O	ODFL
OKE	OMC	ORCL	ORLY	OXY	PAYX	PBCT	PCAR	PEAK	PEG
PEP	PFE	PFGE	PG	PGR	PH	PHM	PKG	PKI	PLD
PM	PNC	PNR	PNW	PPG	PPL	PRGO	PRU	PSA	PSX
PVH	PWR	PXD	QCOM	RCL	RE	REG	REGN	RF	RHI
RJF	RL	RMD	ROK	ROL	ROP	ROST	RSG	RTX	SBAC
SBUX	SCHW	SEE	SHW	SIVB	SJM	SLB	SLG	SNA	SNPS
SO	SPG	SPGI	SRE	STE	STT	STX	STZ	SWK	SWKS
SYK	SYY	T	TAP	TDG	TEL	TFC	TFX	TGT	TIF
TJX	TMO	TMUS	TPR	TROW	TRV	TSCO	TSN	TTWO	TXN
TXT	UAA	UAL	UDR	UHS	ULTA	UNH	UNM	UNP	UPS
URI	USB	V	VAR	VFC	VLO	VMC	VNO	VRSK	VRSN
VRTX	VTR	VZ	WAB	WAT	WBA	WDC	WEC	WELL	WFC
WHR	WLTW	WM	WMB	WMT	WRB	WU	WY	WYNN	XEL
XLNX	XOM	XRAY	XXR	XYL	YUM	ZBH	ZBRA	ZION	

Πίνακας 17: Tickers των μετοχών που συμπεριλαμβάνονται στο τελικό σύνολο δεδομένων

Ticker Loopback Window	Hidden Layers	Activasion	Solver	Max_iter	MAPE %
MMM 2	(1,)	relu	lbfgs	100	1,1875
ABT 3	(6,)	relu	lbfgs	100	1,0597
ABBV 6	(3,)	relu	lbfgs	100	1,5821
ABMD 2	(1,)	relu	lbfgs	100	2,2718
ACN 2	(1,)	relu	lbfgs	100	1,064
ATVI 3	(6,)	relu	lbfgs	100	1,6958
ADBE 2	(1,)	relu	lbfgs	100	1,6028
AMD 3	(3,)	relu	lbfgs	100	3,1381
AAP 3	(6,)	relu	lbfgs	100	1,333
AES 40	(40,)	relu	lbfgs	100	1,1788
AFL 3	(1,)	relu	lbfgs	100	0,9048
A 6	(3,)	relu	lbfgs	100	1,2544
APD 3	(6,)	relu	lbfgs	100	0,9538
AKAM 2	(8,)	relu	lbfgs	100	1,3689
ALK 3	(6,)	relu	lbfgs	100	1,5417
ALB 6	(3,)	relu	lbfgs	100	1,6129
ARE 6	(3,)	relu	lbfgs	100	0,9932
ALXN 40	(20,)	relu	lbfgs	200	1,7178
ALGN 2	(1,)	relu	lbfgs	100	2,1049
LNT 6	(3,)	relu	lbfgs	100	0,8241
ALL 2	(1,)	relu	lbfgs	100	0,9617
GOOGL 2	(1,)	relu	lbfgs	100	1,3595
GOOG 3	(3,)	relu	lbfgs	100	1,3087
MO 6	(3,)	relu	lbfgs	100	1,1876
AMZN 2	(1,)	relu	lbfgs	100	1,6643
AMCR 2	(4,)	relu	lbfgs	100	0,5711
AEE 6	(3,)	relu	lbfgs	100	0,8644
AAL 2	(8,)	relu	lbfgs	100	1,8053
AEP 6	(3,)	relu	lbfgs	100	0,8374
AXP 2	(8,)	relu	lbfgs	100	1,0693
AIG 2	(8,)	relu	lbfgs	100	1,0608
AMT 20	(80,)	relu	lbfgs	100	0,9505
AWK 6	(3,)	relu	lbfgs	100	0,8836
AMP 2	(1,)	relu	lbfgs	100	1,3811
ABC 3	(1,)	relu	lbfgs	100	1,4922
AME 3	(6,)	relu	lbfgs	100	1,0692
AMGN 3	(1,)	relu	lbfgs	100	1,1003
APH 6	(3,)	relu	lbfgs	100	0,9685
ADI 12	(48,)	relu	lbfgs	100	1,2856
ANSS 2	(1,)	relu	lbfgs	100	1,3478
ANTM 2	(1,)	relu	lbfgs	100	1,1224
AON 3	(6,)	relu	lbfgs	100	0,8533
AOS 2	(1,)	relu	lbfgs	100	1,236
APA 3	(12,)	relu	lbfgs	100	1,9438
AIV 2	(8,)	relu	lbfgs	100	0,9026
AAPL 6	(3,)	relu	lbfgs	100	1,3508

AMAT 2	(1,)	relu	lbfgs	100	1,9879
APTV 3	(6,)	relu	lbfgs	100	1,3979
ADM 2	(8,)	relu	lbfgs	100	0,8763
AJG 6	(3,)	relu	lbfgs	100	0,843
AIZ 3	(1,)	relu	lbfgs	100	1,1077
T 3	(12,)	relu	lbfgs	100	1,0216
ATO 6	(3,)	relu	lbfgs	100	0,8566
ADSK 3	(1,)	relu	lbfgs	100	1,7975
ADP 2	(1,)	relu	lbfgs	100	1,0406
AZO 10	(40,)	relu	lbfgs	100	1,1606
AVB 3	(12,)	relu	lbfgs	100	0,8729
AVY 2	(1,)	relu	lbfgs	100	1,1288
BKR 6	(3,)	relu	lbfgs	100	1,64
BLL 2	(8,)	relu	lbfgs	100	1,0455
BAC 3	(3,)	relu	lbfgs	100	1,1203
BK 3	(6,)	relu	lbfgs	100	1,0801
BAX 3	(3,)	relu	lbfgs	100	1,0409
BDX 2	(1,)	relu	lbfgs	100	1,0397
BRK-B 3	(1,)	relu	lbfgs	100	1,0105
BBY 2	(1,)	relu	lbfgs	100	1,6159
BIIB 2	(8,)	relu	lbfgs	100	1,3863
BLK 2	(1,)	relu	lbfgs	100	1,2656
BA 2	(1,)	relu	lbfgs	100	1,5826
BKNG 3	(6,)	relu	lbfgs	100	1,1884
BWA 3	(6,)	relu	lbfgs	100	1,5127
BXP 3	(1,)	logistic	lbfgs	100	0,9469
BSX 6	(3,)	relu	lbfgs	100	1,2913
BMY 2	(8,)	relu	lbfgs	100	1,1526
AVGO 6	(3,)	relu	lbfgs	100	1,7185
BR 2	(1,)	relu	lbfgs	100	1,0949
BF-B 2	(1,)	relu	lbfgs	100	1,181
CHRW 2	(8,)	relu	lbfgs	100	1,071
COG 3	(6,)	relu	lbfgs	100	1,4388
CDNS 6	(3,)	relu	lbfgs	100	1,2979
CPB 2	(8,)	relu	lbfgs	100	1,3077
COF 3	(6,)	relu	lbfgs	100	1,1066
CAH 2	(8,)	relu	lbfgs	100	1,4558
KMX 40	(20,)	relu	lbfgs	200	1,2332
CCL 3	(6,)	relu	lbfgs	100	1,156
CAT 3	(3,)	relu	lbfgs	100	1,554
CBOE 2	(1,)	relu	lbfgs	100	1,3342
CBRE 2	(4,)	relu	lbfgs	100	1,0939
CE 3	(6,)	relu	lbfgs	100	1,184
CNC 3	(6,)	relu	lbfgs	100	1,226
CNP 6	(3,)	relu	lbfgs	100	0,8973
CERN 6	(3,)	relu	lbfgs	100	1,164
CF 3	(6,)	relu	lbfgs	100	1,7502
SCHW 3	(1,)	relu	lbfgs	100	1,3646

CHTR 2	(1,)	relu	lbfgs	100	1,5117
CVX 2	(2,)	relu	lbfgs	100	1,1033
CMG 2	(8,)	relu	lbfgs	100	1,5431
CB 6	(3,)	relu	lbfgs	100	0,8873
CHD 3	(3,)	relu	lbfgs	100	1,0552
CI 2	(1,)	relu	lbfgs	100	1,232
CINF 3	(12,)	relu	lbfgs	100	0,9614
CTAS 2	(1,)	relu	lbfgs	100	1,0561
CSCO 2	(8,)	relu	lbfgs	100	1,2163
C 2	(8,)	relu	lbfgs	100	1,1255
CTXS 2	(4,)	relu	lbfgs	100	0,8244
CLX 2	(1,)	relu	lbfgs	100	1,165
CME 2	(1,)	relu	lbfgs	100	0,9906
CMS 6	(3,)	relu	lbfgs	100	0,8493
KO 2	(4,)	relu	lbfgs	100	0,718
CTSH 3	(3,)	relu	lbfgs	100	1,063
CL 2	(8,)	relu	lbfgs	100	0,928
CMCSA 3	(1,)	relu	lbfgs	100	1,3881
CMA 3	(3,)	relu	lbfgs	100	1,1607
CAG 3	(12,)	relu	lbfgs	100	1,0905
CXO 2	(4,)	relu	lbfgs	100	1,6941
COP 10	(40,)	relu	lbfgs	100	1,3984
ED 3	(3,)	relu	lbfgs	100	0,8614
STZ 6	(3,)	relu	lbfgs	100	1,0966
COO 2	(1,)	relu	lbfgs	100	1,2524
CPRT 2	(1,)	relu	lbfgs	100	1,3705
GLW 2	(1,)	relu	lbfgs	100	1,2958
COST 3	(6,)	relu	lbfgs	100	0,9638
CCI 3	(6,)	relu	lbfgs	100	0,9219
CSX 3	(6,)	relu	lbfgs	100	1,161
CMI 3	(6,)	relu	lbfgs	100	1,1733
CVS 3	(1,)	relu	lbfgs	100	1,4
DHI 3	(3,)	relu	lbfgs	100	1,468
DHR 2	(1,)	relu	lbfgs	100	1,012
DRI 2	(1,)	relu	lbfgs	100	1,1128
DVA 3	(1,)	relu	lbfgs	100	1,2773
DE 2	(1,)	relu	lbfgs	100	1,6158
DAL 30	(120,)	relu	lbfgs	100	1,3372
XRAY 6	(6,)	relu	lbfgs	100	1,2894
DVN 2	(8,)	relu	lbfgs	100	1,8277
FANG 2	(1,)	relu	lbfgs	100	1,9941
DLR 3	(12,)	relu	lbfgs	100	1,0337
DFS 2	(8,)	relu	lbfgs	100	1,1435
DISCA 2	(8,)	relu	lbfgs	100	1,7292
DISCK 2	(8,)	relu	lbfgs	100	1,6808
DISH 3	(6,)	relu	lbfgs	100	1,8184
DG 3	(1,)	relu	lbfgs	100	1,1601
DLTR 2	(4,)	relu	lbfgs	100	1,3215

D 3	(6,)	relu	lbfgs	100	0,8595
DTE 2	(1,)	relu	lbfgs	100	0,8435
DUK 2	(8,)	relu	lbfgs	100	0,8402
DRE 2	(1,)	relu	lbfgs	100	0,9664
DD 3	(1,)	relu	lbfgs	100	1,3502
DXC 2	(1,)	relu	lbfgs	100	1,3044
EMN 2	(4,)	relu	lbfgs	100	1,1208
ETN 3	(6,)	relu	lbfgs	100	1,0413
EBAY 2	(8,)	relu	lbfgs	100	1,2237
ECL 6	(3,)	relu	lbfgs	100	0,8498
EIX 3	(3,)	relu	lbfgs	100	1,1207
EW 2	(8,)	relu	lbfgs	100	1,4347
EA 2	(1,)	relu	lbfgs	100	1,5149
EMR 2	(4,)	relu	lbfgs	100	1,1375
ETR 2	(4,)	relu	lbfgs	100	0,8837
EOG 3	(1,)	relu	lbfgs	100	1,4461
EFX 2	(1,)	relu	lbfgs	100	1,163
EQIX 3	(3,)	relu	lbfgs	100	1,1467
EQR 3	(12,)	relu	lbfgs	100	0,8805
ESS 3	(6,)	relu	lbfgs	100	0,9345
EL 2	(1,)	relu	lbfgs	100	1,2223
EVRG 6	(3,)	relu	lbfgs	100	0,8661
ES 6	(3,)	relu	lbfgs	100	0,8681
RE 3	(1,)	relu	lbfgs	100	1,0006
EXC 40	(20,)	relu	lbfgs	100	0,7941
EXPE 2	(8,)	relu	lbfgs	100	1,2995
EXPD 3	(1,)	relu	lbfgs	100	1,1311
EXR 2	(8,)	relu	lbfgs	100	0,966
XOM 2	(2,)	relu	lbfgs	100	1,0329
FFIV 3	(3,)	relu	lbfgs	100	1,2289
FB 2	(1,)	relu	lbfgs	100	1,6349
FAST 2	(1,)	relu	lbfgs	100	1,2447
FRT 3	(6,)	relu	lbfgs	100	0,9441
FDX 6	(3,)	relu	lbfgs	100	1,2952
FIS 2	(1,)	relu	lbfgs	100	0,9205
FITB 2	(8,)	relu	lbfgs	100	1,1809
FE 15	(30,)	relu	lbfgs	100	0,9495
FRC 15	(30,)	relu	lbfgs	100	1,1105
FISV 6	(3,)	relu	lbfgs	100	0,9399
FLT 2	(4,)	relu	lbfgs	100	1,1113
FLIR 6	(3,)	relu	lbfgs	100	1,1393
FLS 3	(6,)	relu	lbfgs	100	1,5489
FMC 3	(3,)	relu	lbfgs	100	1,3475
F 2	(1,)	relu	lbfgs	100	1,2958
FTNT 2	(4,)	relu	lbfgs	100	1,6368
FBHS 2	(8,)	relu	lbfgs	100	1,2504
BEN 2	(4,)	relu	lbfgs	100	1,3132
FCX 3	(12,)	relu	lbfgs	100	2,0817

GPS 3	(12,)	relu	lbfgs	100	1,7287
GRMN 3	(6,)	relu	lbfgs	100	0,9495
IT 6	(3,)	relu	lbfgs	100	1,1572
GD 6	(3,)	relu	lbfgs	100	1,1336
GE 2	(8,)	relu	lbfgs	100	4,2936
GIS 10	(40,)	relu	lbfgs	100	1,2208
GM 3	(6,)	relu	lbfgs	100	1,3842
GPC 2	(4,)	relu	lbfgs	100	0,9074
GILD 3	(6,)	relu	lbfgs	100	1,2639
GL 6	(3,)	relu	lbfgs	100	0,9734
GPN 6	(3,)	relu	lbfgs	100	1,2595
GS 3	(6,)	relu	lbfgs	100	1,2051
GWW 30	(15,)	relu	lbfgs	100	1,3719
HAL 3	(6,)	relu	lbfgs	100	1,3329
HBI 2	(4,)	relu	lbfgs	100	1,6767
HIG 3	(1,)	relu	lbfgs	100	0,936
HAS 3	(1,)	relu	lbfgs	100	1,1837
HCA 3	(3,)	relu	lbfgs	100	1,1259
PEAK 40	(20,)	relu	lbfgs	200	1,02
HSIC 3	(1,)	relu	lbfgs	100	1,218
HSY 3	(6,)	relu	lbfgs	100	0,9468
HES 2	(4,)	relu	lbfgs	100	1,7839
HFC 2	(4,)	relu	lbfgs	100	1,8157
HOLX 2	(8,)	relu	lbfgs	100	1,0821
HD 2	(1,)	relu	lbfgs	100	1,0749
HON 6	(3,)	relu	lbfgs	100	0,942
HRL 6	(3,)	relu	lbfgs	100	0,9607
HST 20	(40,)	relu	lbfgs	100	1,0744
HPQ 6	(3,)	relu	lbfgs	100	1,2813
HUM 2	(1,)	relu	lbfgs	100	1,0971
HBAN 2	(8,)	relu	lbfgs	100	1,0693
HII 6	(3,)	relu	lbfgs	100	1,4038
IEX 6	(3,)	relu	lbfgs	100	1,0725
IDXX 2	(1,)	relu	lbfgs	100	1,436
ITW 2	(1,)	relu	lbfgs	100	1,1486
ILMN 3	(3,)	relu	lbfgs	100	1,6746
INCY 2	(8,)	relu	lbfgs	100	1,7793
INTC 2	(1,)	relu	lbfgs	100	1,5916
ICE 2	(1,)	relu	lbfgs	100	0,9397
IBM 3	(1,)	relu	lbfgs	100	1,0653
IP 3	(6,)	relu	lbfgs	100	1,2697
IPG 3	(1,)	relu	lbfgs	100	1,256
IFF 3	(6,)	relu	lbfgs	100	0,9876
INTU 2	(1,)	relu	lbfgs	100	1,3244
ISRG 3	(3,)	relu	lbfgs	100	1,5371
IVZ 3	(6,)	relu	lbfgs	100	1,4055
IPGP 2	(1,)	relu	lbfgs	100	2,1239
IRM 2	(8,)	relu	lbfgs	100	1,1367

JKHY 2	(1,)	relu	lbfgs	100	0,9455
J 2	(4,)	relu	lbfgs	100	1,253
JBHT 6	(3,)	relu	lbfgs	100	1,1385
SJM 2	(8,)	relu	lbfgs	100	1,1913
JNJ 2	(1,)	relu	lbfgs	100	0,9627
JCI 3	(6,)	relu	lbfgs	100	1,27
JPM 6	(3,)	relu	lbfgs	100	1,0883
JNPR 2	(8,)	relu	lbfgs	100	1,1052
KSU 2	(8,)	relu	lbfgs	100	1,2007
K 2	(8,)	relu	lbfgs	100	1,0544
KEY 20	(20,)	relu	lbfgs	100	1,2844
KMB 2	(8,)	relu	lbfgs	100	1,106
KIM 3	(6,)	relu	lbfgs	100	1,2872
KMI 2	(4,)	relu	lbfgs	100	1,1516
KLAC 2	(1,)	relu	lbfgs	100	1,7043
KR 30	(30,)	relu	lbfgs	100	1,4912
LB 6	(3,)	relu	lbfgs	100	1,9847
LHX 2	(1,)	relu	lbfgs	100	1,213
LH 2	(8,)	relu	lbfgs	100	0,9952
LRCX 2	(1,)	relu	lbfgs	100	2,0283
LVS 30	(120,)	relu	lbfgs	100	1,3688
LEG 3	(12,)	relu	lbfgs	100	1,1176
LDOS 6	(3,)	relu	lbfgs	100	1,1535
LEN 2	(4,)	relu	lbfgs	100	1,4928
LLY 2	(8,)	relu	lbfgs	100	1,0567
LNC 6	(3,)	relu	lbfgs	100	1,3211
LIN 2	(8,)	relu	lbfgs	100	1,2115
LYV 6	(3,)	relu	lbfgs	100	1,5573
LKQ 2	(4,)	relu	lbfgs	100	1,1971
LMT 2	(1,)	relu	lbfgs	100	1,1814
L 2	(4,)	relu	lbfgs	100	0,891
LOW 2	(1,)	relu	lbfgs	100	1,3126
LYB 2	(8,)	relu	lbfgs	100	1,2522
MTB 2	(8,)	relu	lbfgs	100	0,9788
MRO 2	(1,)	relu	lbfgs	100	2,0176
MPC 2	(4,)	relu	lbfgs	100	1,4767
MKTX 2	(1,)	relu	lbfgs	100	1,1382
MAR 2	(1,)	relu	lbfgs	100	1,2488
MMC 6	(3,)	relu	lbfgs	100	0,8371
MLM 2	(4,)	relu	lbfgs	100	1,2382
MAS 2	(1,)	relu	lbfgs	100	1,2142
MA 2	(1,)	relu	lbfgs	100	1,3244
MKC 2	(8,)	relu	lbfgs	100	0,9554
MXIM 2	(1,)	relu	lbfgs	100	1,4688
MCD 2	(1,)	relu	lbfgs	100	0,9711
MCK 3	(6,)	relu	lbfgs	100	1,2491
MDT 3	(3,)	relu	lbfgs	100	0,9284
MRK 3	(6,)	relu	lbfgs	100	0,944

MET 2	(4,)	relu	lbfgs	100	1,1813
MTD 6	(3,)	relu	lbfgs	100	1,2804
MGM 2	(4,)	relu	lbfgs	100	1,7658
MCHP 3	(6,)	relu	lbfgs	100	1,5543
MU 3	(3,)	relu	lbfgs	100	2,3414
MSFT 2	(1,)	relu	lbfgs	100	1,2773
MAA 2	(8,)	relu	lbfgs	100	0,9052
MHK 3	(1,)	relu	lbfgs	100	1,3477
TAP 6	(3,)	relu	lbfgs	100	1,2761
MDLZ 2	(4,)	relu	lbfgs	100	0,8723
MNST 3	(1,)	relu	lbfgs	100	1,2699
MCO 6	(3,)	relu	lbfgs	100	1,1271
MS 3	(6,)	relu	lbfgs	100	1,2677
MOS 2	(4,)	relu	lbfgs	100	1,5992
MSI 2	(8,)	relu	lbfgs	100	1,015
MSCI 6	(3,)	relu	lbfgs	100	1,247
MYL 3	(6,)	relu	lbfgs	100	1,4882
NDAQ 6	(3,)	relu	lbfgs	100	0,945
NOV 3	(12,)	relu	lbfgs	100	1,5023
NTAP 2	(1,)	relu	lbfgs	100	1,632
NFLX 2	(1,)	relu	lbfgs	100	2,2648
NWL 2	(8,)	relu	lbfgs	100	1,9389
NEM 3	(6,)	relu	lbfgs	100	1,205
NEE 6	(3,)	relu	lbfgs	100	0,8228
NLSN 6	(3,)	relu	lbfgs	100	1,4774
NKE 3	(1,)	relu	lbfgs	100	1,3052
NI 6	(3,)	relu	lbfgs	100	0,959
NSC 3	(6,)	relu	lbfgs	100	1,1988
NTRS 2	(8,)	relu	lbfgs	100	1,1261
NOC 2	(1,)	relu	lbfgs	100	1,2798
NRG 2	(8,)	relu	lbfgs	100	1,2836
NUE 2	(4,)	relu	lbfgs	100	1,3045
NVDA 2	(1,)	relu	lbfgs	100	2,07
NVR 2	(1,)	relu	lbfgs	100	1,6396
ORLY 15	(30,)	relu	lbfgs	100	1,2826
OXY 2	(4,)	relu	lbfgs	100	1,1587
ODFL 3	(1,)	relu	lbfgs	100	1,4679
OMC 2	(4,)	relu	lbfgs	100	1,0465
OKE 30	(15,)	relu	lbfgs	200	1,092
ORCL 3	(6,)	relu	lbfgs	100	1,1104
PCAR 3	(6,)	relu	lbfgs	100	1,2665
PKG 2	(1,)	relu	lbfgs	100	1,2351
PH 3	(3,)	relu	lbfgs	100	1,2962
PAYX 2	(1,)	relu	lbfgs	100	0,9877
PNR 3	(6,)	relu	lbfgs	100	1,0455
PBCT 2	(8,)	relu	lbfgs	100	0,9075
PEP 6	(3,)	relu	lbfgs	100	0,9239
PKI 6	(3,)	relu	lbfgs	100	1,2884

PRGO 3	(6,)	relu	lbfgs	100	1,4116
PFE 2	(4,)	relu	lbfgs	100	0,9219
PM 3	(3,)	relu	lbfgs	100	1,1812
PSX 2	(8,)	relu	lbfgs	100	1,2382
PNW 3	(3,)	relu	lbfgs	100	0,8509
PXD 2	(1,)	relu	lbfgs	100	1,5206
PNC 6	(3,)	relu	lbfgs	100	1,0785
PPG 2	(4,)	relu	lbfgs	100	1,0887
PPL 2	(8,)	relu	lbfgs	100	1,0402
PFG 2	(1,)	relu	lbfgs	100	1,2926
PG 3	(6,)	relu	lbfgs	100	0,8432
PGR 2	(1,)	relu	lbfgs	100	1,1308
PLD 6	(3,)	relu	lbfgs	100	1,006
PRU 2	(8,)	relu	lbfgs	100	1,2049
PEG 3	(1,)	relu	lbfgs	100	0,902
PSA 2	(1,)	relu	lbfgs	100	0,9295
PHM 3	(3,)	relu	lbfgs	100	1,4587
PVH 10	(40,)	relu	lbfgs	100	1,341
PWR 2	(4,)	relu	lbfgs	100	1,1581
QCOM 2	(4,)	relu	lbfgs	100	1,296
DGX 3	(3,)	relu	lbfgs	100	0,9424
RL 2	(8,)	relu	lbfgs	100	1,5138
RJF 2	(8,)	relu	lbfgs	100	1,2714
RTX 3	(6,)	relu	lbfgs	100	1,0185
O 3	(12,)	relu	lbfgs	100	0,9164
REG 3	(6,)	relu	lbfgs	100	0,9302
REGN 30	(120,)	relu	lbfgs	100	1,4892
RF 3	(3,)	relu	lbfgs	100	1,2729
RSG 2	(1,)	relu	lbfgs	100	0,7648
RMD 3	(6,)	relu	lbfgs	100	1,0867
RHI 2	(4,)	relu	lbfgs	100	1,2722
ROK 3	(6,)	relu	lbfgs	100	1,2441
ROL 2	(1,)	relu	lbfgs	100	1,0991
ROP 2	(1,)	relu	lbfgs	100	1,0992
ROST 3	(12,)	relu	lbfgs	100	1,185
RCL 3	(6,)	relu	lbfgs	100	1,3895
SPGI 6	(3,)	relu	lbfgs	100	1,0683
CRM 3	(6,)	relu	lbfgs	100	1,5279
SBAC 2	(1,)	relu	lbfgs	100	1,0331
SLB 3	(6,)	relu	lbfgs	100	1,3717
STX 6	(24,)	relu	lbfgs	100	1,6228
SEE 3	(1,)	relu	lbfgs	100	1,1112
SRE 3	(1,)	relu	lbfgs	100	0,9124
NOW 3	(6,)	relu	lbfgs	100	1,7977
SHW 6	(3,)	relu	lbfgs	100	1,1038
SPG 3	(6,)	relu	lbfgs	100	0,9465
SWKS 3	(12,)	relu	lbfgs	100	1,4919
SLG 40	(20,)	relu	lbfgs	100	0,9061

SNA 2	(1,)	relu	lbfgs	100	1,1787
SO 2	(8,)	relu	lbfgs	100	0,887
LUV 2	(1,)	relu	lbfgs	100	1,2957
SWK 6	(3,)	relu	lbfgs	100	1,4271
SBUX 3	(1,)	relu	lbfgs	100	1,0288
STT 2	(8,)	relu	lbfgs	100	1,2668
STE 6	(3,)	relu	lbfgs	100	0,956
SYK 2	(1,)	relu	lbfgs	100	1,0778
SIVB 3	(1,)	relu	lbfgs	100	1,5829
SNPS 2	(1,)	relu	lbfgs	100	1,1583
SYU 2	(1,)	relu	lbfgs	100	0,9263
TMUS 2	(1,)	relu	lbfgs	100	1,1302
TROW 2	(8,)	relu	lbfgs	100	1,1797
TTWO 2	(1,)	relu	lbfgs	100	1,9223
TPR 2	(8,)	relu	lbfgs	100	1,2207
TGT 30	(60,)	relu	lbfgs	200	1,2353
TEL 6	(3,)	relu	lbfgs	100	1,0872
FTI 10	(40,)	relu	lbfgs	100	1,6833
TFX 6	(3,)	relu	lbfgs	100	1,2698
TXN 2	(1,)	relu	lbfgs	100	1,4212
TXT 6	(3,)	relu	lbfgs	100	1,1677
TMO 3	(3,)	relu	lbfgs	100	1,0868
TIF 3	(6,)	relu	lbfgs	100	1,3135
TJX 3	(1,)	relu	lbfgs	100	1,0223
TSCO 2	(8,)	relu	lbfgs	100	1,3275
TDG 6	(3,)	relu	lbfgs	100	1,2378
TRV 3	(1,)	relu	lbfgs	100	0,9887
TFC 6	(3,)	relu	lbfgs	100	1,0389
TSN 6	(3,)	relu	lbfgs	100	1,0781
UDR 2	(8,)	relu	lbfgs	100	0,8854
ULTA 3	(3,)	relu	lbfgs	100	1,516
USB 2	(8,)	relu	lbfgs	100	0,8747
UAA 40	(20,)	relu	lbfgs	100	2,2138
UNP 3	(1,)	relu	lbfgs	100	1,1688
UAL 12	(6,)	relu	lbfgs	100	1,4188
UNH 6	(3,)	relu	lbfgs	100	1,1155
UPS 2	(4,)	relu	lbfgs	100	1,0744
URI 3	(3,)	relu	lbfgs	100	1,8552
UHS 2	(8,)	relu	lbfgs	100	1,0548
UNM 2	(1,)	relu	lbfgs	100	1,4782
VFC 3	(3,)	relu	lbfgs	100	1,1431
VLO 2	(1,)	relu	lbfgs	100	1,4089
VAR 2	(8,)	relu	lbfgs	100	1,1426
VTR 40	(20,)	relu	lbfgs	200	1,0232
VRSN 2	(1,)	relu	lbfgs	100	1,307
VRSK 2	(4,)	relu	lbfgs	100	0,8942
VZ 2	(8,)	relu	lbfgs	100	0,9632
VRTX 2	(8,)	relu	lbfgs	100	1,5021

V 2	(1,)	relu	lbfgs	100	1,1225
VNO 2	(4,)	relu	lbfgs	100	0,9375
VMC 2	(8,)	relu	lbfgs	100	1,3587
WRB 3	(3,)	relu	lbfgs	100	0,7794
WAB 3	(1,)	relu	lbfgs	100	1,3281
WMT 2	(8,)	relu	lbfgs	100	1,0272
WBA 3	(3,)	relu	lbfgs	100	1,2079
DIS 3	(3,)	relu	lbfgs	100	0,9862
WM 2	(1,)	relu	lbfgs	100	0,7927
WAT 6	(3,)	relu	lbfgs	100	1,0877
WEC 2	(1,)	relu	lbfgs	100	0,9101
WFC 3	(3,)	relu	lbfgs	100	1,053
WELL 6	(12,)	relu	lbfgs	100	1,0833
WDC 2	(8,)	relu	lbfgs	100	1,6601
WU 30	(30,)	relu	lbfgs	100	0,9112
WY 2	(4,)	relu	lbfgs	100	1,042
WHR 2	(4,)	relu	lbfgs	100	1,3779
WMB 2	(4,)	relu	lbfgs	100	1,1201
WLTW 3	(1,)	relu	lbfgs	100	0,9023
WYNN 3	(12,)	relu	lbfgs	100	1,9145
XEL 3	(3,)	relu	lbfgs	100	0,8617
XRX 2	(4,)	relu	lbfgs	100	1,3918
XLNX 2	(8,)	relu	lbfgs	100	1,5719
XYL 6	(3,)	relu	lbfgs	100	1,2219
YUM 2	(8,)	relu	lbfgs	100	0,8576
ZBRA 6	(3,)	relu	lbfgs	100	1,6432
ZBH 2	(4,)	relu	lbfgs	100	1,0501
ZION 2	(8,)	relu	lbfgs	100	1,189

Πίνακας 18: Αποτελέσματα διερεύνησης βέλτιστου νευρωνικού δικτύου για κάθε μετοχή

	Hidden Layers	Activation	Solver	Max_iter	MAPE %
Loopback Window: 2	(8,)	relu	lbfgs	100	14,648258
Loopback Window: 3	(6,)	relu	lbfgs	100	14,732882
Loopback Window: 4	(4,)	relu	adam	100	14,699353
Loopback Window: 5	(2,)	relu	adam	100	14,762276
Loopback Window: 10	(10,)'	relu	adam	100	14,717822
Loopback Window: 15	(30,)'	logistic	adam	100	14,609467
Loopback Window: 20	(40,)'	relu	adam	100	14,690912
Loopback Window: 40	(20,)'	logistic	adam	100	14,122172
Loopback Window: 60	(30,)'	logistic	adam	100	14,136541
Loopback Window: 80	(320,)'	relu	adam	100	14,610607
Loopback Window: 100	(50,)'	relu	adam	100	14,5433

Πίνακας 19: Αποτελέσματα διερεύνησης βέλτιστου ολικού δικτύου

Feature	MAPE %
Value__lempel_ziv_complexity__bins_100	1,258157484
Value__first_location_of_maximum	1,258187654
Value__linear_trend_attr_"stderr"	1,25823357
Value__sample_entropy	1,25829911
Value__agg_linear_trend_attr_"rvalue"__chunk_len_50__f_agg_"min"	1,25831027
Value__lempel_ziv_complexity__bins_5	1,258322683
Value__agg_linear_trend_attr_"rvalue"__chunk_len_50__f_agg_"mean"	1,258352394
Value__index_mass_quantile__q_0.6	1,258376989
Value__agg_linear_trend_attr_"stderr"__chunk_len_5__f_agg_"max"	1,258386212
Value__last_location_of_maximum	1,258398365
Value__ar_coefficient__coeff_2__k_10	1,258439479
Value__fourier_entropy__bins_5	1,258458815
Value__change_quantiles__f_agg_"mean"__isabs_False__qh_0.8__ql_0.0	1,258471016
Value__agg_linear_trend_attr_"slope"__chunk_len_10__f_agg_"min"	1,258473552
Value__index_mass_quantile__q_0.4	1,258476658
Value__ratio_beyond_r_sigma__r_0.5	1,258508018
Value__ar_coefficient__coeff_0__k_10	1,258516559
Value__agg_linear_trend_attr_"rvalue"__chunk_len_5__f_agg_"var"	1,258522797
Value__agg_autocorrelation__f_agg_"var"__maxlag_40	1,258551273
Value__change_quantiles__f_agg_"mean"__isabs_False__qh_0.4__ql_0.2	1,258557861
Value__index_mass_quantile__q_0.7	1,258558999

Value__ar_coefficient__coeff_10_k_10	1,258565275
Value__lempel_ziv_complexity__bins_2	1,258574629
Value__agg_linear_trend__attr_"rvalue"__chunk_len_5__f_agg_"max"	1,258577081
Value__agg_linear_trend__attr_"rvalue"__chunk_len_10__f_agg_"var"	1,258579401
Value__linear_trend__attr_"rvalue"	1,258579536
Value__partial_autocorrelation__lag_4	1,258590702
Value__approximate_entropy__m_2_r_0.9	1,258592576
Value__fourier_entropy__bins_100	1,25859578
Value__agg_autocorrelation__f_agg_"mean"__maxlag_40	1,258608049
Value__index_mass_quantile__q_0.8	1,258623307
Value__partial_autocorrelation__lag_6	1,258628705
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_1.0__ql_0.6	1,25863518
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_0.2__ql_0.0	1,258636133
Value__ar_coefficient__coeff_6_k_10	1,258637216
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_0.8__ql_0.4	1,25864239
Value__energy_ratio_by_chunks__num_segments_10__segment_focus_1	1,258645403
Value__change_quantiles__f_agg_"mean"__isabs_False__qh_0.8__ql_0.2	1,258659324
Value__approximate_entropy__m_2_r_0.5	1,258669991
Value__index_mass_quantile__q_0.3	1,258683201
Value__agg_linear_trend__attr_"rvalue"__chunk_len_10__f_agg_"max"	1,258683676
Value__agg_autocorrelation__f_agg_"median"__maxlag_40	1,25868509
Value__variation_coefficient	1,258695018
Value__agg_linear_trend__attr_"rvalue"__chunk_len_5__f_agg_"min"	1,258696406
Value__ratio_beyond_r_sigma_r_1	1,25871565
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_0.8__ql_0.0	1,258722246
Value__agg_linear_trend__attr_"stderr"__chunk_len_5__f_agg_"mean"	1,258722383
Value__energy_ratio_by_chunks__num_segments_10__segment_focus_8	1,258722656
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_1.0__ql_0.2	1,258753387
Value__agg_linear_trend__attr_"slope"__chunk_len_10__f_agg_"max"	1,258756859
Value__ar_coefficient__coeff_9_k_10	1,258766342
Value__last_location_of_minimum	1,258772442
Value__change_quantiles__f_agg_"mean"__isabs_False__qh_0.8__ql_0.6	1,258802391
Value__energy_ratio_by_chunks__num_segments_10__segment_focus_2	1,258805541
Value__first_location_of_minimum	1,258819443
Value__agg_linear_trend__attr_"slope"__chunk_len_50__f_agg_"min"	1,258821098
Value__energy_ratio_by_chunks__num_segments_10__segment_focus_5	1,258838884
Value__agg_linear_trend__attr_"stderr"__chunk_len_10__f_agg_"mean"	1,258844468
Value__fourier_entropy__bins_3	1,258861993
Value__lempel_ziv_complexity__bins_3	1,258863883
Value__ar_coefficient__coeff_7_k_10	1,258868826
Value__agg_linear_trend__attr_"stderr"__chunk_len_10__f_agg_"max"	1,258880282
Value__benford_correlation	1,258886802
Value__mean_abs_change	1,25889402
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_1.0__ql_0.0	1,25889402
Value__partial_autocorrelation__lag_7	1,258897842
Value__change_quantiles__f_agg_"mean"__isabs_False__qh_1.0__ql_0.8	1,258898589
Value__skewness	1,258908144
Value__linear_trend__attr_"slope"	1,258915683

Value__autocorrelation__lag_3	1,258920913
Value__ratio_beyond_r_sigma_r_1.5	1,258936222
Value__change_quantiles__f_agg_"mean"__isabs_False__qh_0.6__ql_0.0	1,258950097
Value__autocorrelation__lag_1	1,258958872
Value__ratio_beyond_r_sigma_r_2.5	1,258963883
Value__change_quantiles__f_agg_"mean"__isabs_False__qh_1.0__ql_0.0	1,258969971
Value__ar_coefficient__coeff_1_k_10	1,258972837
Value__change_quantiles__f_agg_"mean"__isabs_False__qh_0.4__ql_0.0	1,258979019
Value__ratio_beyond_r_sigma_r_2	1,258992162
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_0.4__ql_0.0	1,259003611
Value__fourier_entropy_bins_10	1,25900442
Value__mean_change	1,259017994
Value__agg_linear_trend__attr_"stderr"__chunk_len_5__f_agg_"var"	1,259031348
Value__agg_linear_trend__attr_"rvalue"__chunk_len_10__f_agg_"min"	1,25903733
Value__approximate_entropy__m_2_r_0.1	1,259040475
Value__ratio_value_number_to_time_series_length	1,259044656
Value__partial_autocorrelation__lag_9	1,259047604
Value__autocorrelation__lag_4	1,259050696
Value__energy_ratio_by_chunks__num_segments_10__segment_focus_4	1,259060169
Value__augmented_dickey_fuller__attr_"pvalue"__autolag_"AIC"	1,259079293
Value__approximate_entropy__m_2_r_0.7	1,259080147
Value__change_quantiles__f_agg_"mean"__isabs_False__qh_0.2__ql_0.0	1,259090748
Value__agg_linear_trend__attr_"stderr"__chunk_len_50__f_agg_"mean"	1,259092734
Value__autocorrelation__lag_7	1,25910995
Value__partial_autocorrelation__lag_5	1,259122685
Value__energy_ratio_by_chunks__num_segments_10__segment_focus_7	1,259124858
Value__fourier_entropy_bins_2	1,259131577
Value__change_quantiles__f_agg_"var"__isabs_True__qh_0.4__ql_0.0	1,259135109
Value__approximate_entropy__m_2_r_0.3	1,259153319
Value__partial_autocorrelation__lag_3	1,259166634
Value__mean_second_derivative_central	1,259176256
Value__autocorrelation__lag_2	1,259177891
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_1.0__ql_0.4	1,259180528
Value__change_quantiles__f_agg_"mean"__isabs_False__qh_0.6__ql_0.2	1,259189429
Value__energy_ratio_by_chunks__num_segments_10__segment_focus_6	1,259198181
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_0.4__ql_0.2	1,259233518
Value__augmented_dickey_fuller__attr_"teststat"__autolag_"AIC"	1,259242854
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_0.6__ql_0.2	1,25924876
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_0.8__ql_0.6	1,259250315
Value__kurtosis	1,259257599
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_0.8__ql_0.2	1,259270676
Value__agg_linear_trend__attr_"slope"__chunk_len_5__f_agg_"max"	1,259272531
Value__ar_coefficient__coeff_4_k_10	1,259273112
Value__percentage_of_reoccurring_values_to_all_values	1,259283008
Value__permutation_entropy_dimension_5_tau_1	1,259310336
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_1.0__ql_0.8	1,259319325
Value__agg_linear_trend__attr_"slope"__chunk_len_10__f_agg_"mean"	1,259327504
Value__index_mass_quantile__q_0.2	1,259341494

Value_ar_coefficient_coeff_3_k_10	1,259343584
Value_energy_ratio_by_chunks_num_segments_10_segment_focus_9	1,25934469
Value_permutation_entropy_dimension_3_tau_1	1,259346256
Value_change_quantiles_f_agg_"var"_isabs_True_qh_0.2_ql_0.0	1,259360274
Value_agg_linear_trend_attr_"stderr"_chunk_len_50_f_agg_"min"	1,259374562
Value_agg_linear_trend_attr_"slope"_chunk_len_5_f_agg_"min"	1,259377908
Value_autocorrelation_lag_9	1,259383796
Value_change_quantiles_f_agg_"mean"_isabs_False_qh_0.8_ql_0.4	1,259384077
Value_partial_autocorrelation_lag_2	1,259399132
Value_agg_linear_trend_attr_"rvalue"_chunk_len_5_f_agg_"mean"	1,259409831
Value_percentage_of_reoccurring_datapoints_to_all_datapoints	1,259423612
Value_energy_ratio_by_chunks_num_segments_10_segment_focus_0	1,259431209
Value_binned_entropy_max_bins_10	1,259441326
Value_agg_linear_trend_attr_"stderr"_chunk_len_50_f_agg_"var"	1,259481251
Value_change_quantiles_f_agg_"mean"_isabs_False_qh_0.6_ql_0.4	1,259512969
Value_index_mass_quantile_q_0.1	1,259523162
Value_change_quantiles_f_agg_"var"_isabs_True_qh_0.4_ql_0.2	1,259566395
Value_change_quantiles_f_agg_"mean"_isabs_True_qh_0.6_ql_0.0	1,259610693
Value_agg_linear_trend_attr_"slope"_chunk_len_10_f_agg_"var"	1,259622136
Value_ar_coefficient_coeff_5_k_10	1,259631254
Value_agg_linear_trend_attr_"slope"_chunk_len_5_f_agg_"mean"	1,259642615
Value_agg_linear_trend_attr_"rvalue"_chunk_len_10_f_agg_"mean"	1,259643203
Value_autocorrelation_lag_5	1,25964432
Value_agg_linear_trend_attr_"slope"_chunk_len_50_f_agg_"max"	1,259676248
Value_agg_linear_trend_attr_"stderr"_chunk_len_10_f_agg_"min"	1,25967988
Value_change_quantiles_f_agg_"mean"_isabs_False_qh_1.0_ql_0.2	1,25968598
Value_agg_linear_trend_attr_"rvalue"_chunk_len_50_f_agg_"var"	1,259691781
Value_autocorrelation_lag_8	1,259692692
Value_agg_linear_trend_attr_"slope"_chunk_len_50_f_agg_"mean"	1,259700868
Value_change_quantiles_f_agg_"mean"_isabs_False_qh_1.0_ql_0.4	1,259753788
Value_ar_coefficient_coeff_8_k_10	1,259759239
Value_agg_linear_trend_attr_"rvalue"_chunk_len_50_f_agg_"max"	1,259785448
Value_partial_autocorrelation_lag_1	1,259789514
Value_agg_linear_trend_attr_"slope"_chunk_len_5_f_agg_"var"	1,259796168
Value_change_quantiles_f_agg_"mean"_isabs_False_qh_1.0_ql_0.6	1,259812056
Value_permutation_entropy_dimension_6_tau_1	1,259813771
Value_lempel_ziv_complexity_bins_10	1,259835174
Value_fft_aggregated_aggttype_"kurtosis"	1,259886335
Value_energy_ratio_by_chunks_num_segments_10_segment_focus_3	1,259912983
Value_agg_linear_trend_attr_"stderr"_chunk_len_50_f_agg_"max"	1,259934573
Value_fft_aggregated_aggttype_"centroid"	1,259958183
Value_change_quantiles_f_agg_"var"_isabs_False_qh_0.4_ql_0.2	1,260017985
Value_permutation_entropy_dimension_7_tau_1	1,260034375
Value_quantile_q_0.9	1,260049379
Value_autocorrelation_lag_6	1,260067434
Value_quantile_q_0.6	1,260097568
Value_partial_autocorrelation_lag_8	1,26011114
Value_cid_ce_normalize_True	1,260145242

Value__change_quantiles__f_agg_"var"__isabs_True__qh_0.6__ql_0.4	1,260231681
Value__quantile__q_0.7	1,260237729
Value__quantile__q_0.4	1,260248094
Value__fft_aggregated__aggtype_"skew"	1,260260948
Value__change_quantiles__f_agg_"var"__isabs_True__qh_0.8__ql_0.2	1,260303089
Value__standard_deviation	1,260303098
Value__change_quantiles__f_agg_"mean"__isabs_True__qh_0.6__ql_0.4	1,260351018
Value__index_mass_quantile__q_0.9	1,260355921
Value__change_quantiles__f_agg_"var"__isabs_False__qh_0.4__ql_0.0	1,260447556
Value__change_quantiles__f_agg_"var"__isabs_True__qh_0.6__ql_0.0	1,260457137
Value__permutation_entropy_dimension_4_tau_1	1,260480684
Value__change_quantiles__f_agg_"var"__isabs_True__qh_0.8__ql_0.0	1,260487276
Value__agg_linear_trend_attr_"stderr"__chunk_len_5__f_agg_"min"	1,26052311
Value__agg_linear_trend_attr_"stderr"__chunk_len_10__f_agg_"var"	1,260633088
Value__change_quantiles__f_agg_"var"__isabs_True__qh_0.8__ql_0.6	1,260699459
Value__change_quantiles__f_agg_"var"__isabs_True__qh_0.6__ql_0.2	1,26082104
Value__quantile__q_0.1	1,260866207
Value__agg_linear_trend_attr_"slope"__chunk_len_50__f_agg_"var"	1,260896533
Value__change_quantiles__f_agg_"var"__isabs_False__qh_0.2__ql_0.0	1,260932292
Value__quantile__q_0.8	1,261296389
Value__change_quantiles__f_agg_"var"__isabs_False__qh_0.8__ql_0.6	1,26133517
Value__change_quantiles__f_agg_"var"__isabs_False__qh_0.6__ql_0.2	1,261474122
Value__change_quantiles__f_agg_"var"__isabs_True__qh_1.0__ql_0.0	1,261481768
Value__change_quantiles__f_agg_"var"__isabs_True__qh_1.0__ql_0.2	1,261510632
Value__change_quantiles__f_agg_"var"__isabs_False__qh_0.8__ql_0.2	1,261720696
Value__change_quantiles__f_agg_"var"__isabs_False__qh_0.6__ql_0.4	1,261759373
Value__change_quantiles__f_agg_"var"__isabs_True__qh_0.8__ql_0.4	1,261825076
Value__change_quantiles__f_agg_"var"__isabs_True__qh_1.0__ql_0.6	1,261995584
Value__change_quantiles__f_agg_"var"__isabs_False__qh_1.0__ql_0.4	1,262195074
Value__quantile__q_0.3	1,262356458
Value__agg_linear_trend_attr_"intercept"__chunk_len_5__f_agg_"var"	1,262402574
Value__change_quantiles__f_agg_"var"__isabs_True__qh_1.0__ql_0.4	1,262694024
Value__change_quantiles__f_agg_"var"__isabs_False__qh_0.8__ql_0.4	1,263049067
Value__change_quantiles__f_agg_"var"__isabs_False__qh_0.8__ql_0.0	1,263187363
Value__change_quantiles__f_agg_"var"__isabs_False__qh_1.0__ql_0.2	1,26322807
Value__change_quantiles__f_agg_"var"__isabs_False__qh_1.0__ql_0.0	1,263675555
Value__change_quantiles__f_agg_"var"__isabs_False__qh_0.6__ql_0.0	1,263797572
Value__quantile__q_0.2	1,264046736
Value__agg_linear_trend_attr_"intercept"__chunk_len_10__f_agg_"max"	1,264112937
Value__agg_linear_trend_attr_"intercept"__chunk_len_10__f_agg_"var"	1,264252624
Value__agg_linear_trend_attr_"intercept"__chunk_len_10__f_agg_"min"	1,264376236
Value__change_quantiles__f_agg_"var"__isabs_False__qh_1.0__ql_0.6	1,264859163
Value__change_quantiles__f_agg_"var"__isabs_True__qh_1.0__ql_0.8	1,265355441
Value__agg_linear_trend_attr_"intercept"__chunk_len_10__f_agg_"mean"	1,266293574
Value__agg_linear_trend_attr_"intercept"__chunk_len_50__f_agg_"max"	1,267747011
Value__agg_linear_trend_attr_"intercept"__chunk_len_50__f_agg_"min"	1,268907905
Value__agg_linear_trend_attr_"intercept"__chunk_len_5__f_agg_"mean"	1,270935547

Value__agg_linear_trend_attr_"intercept"__chunk_len_5_f_agg_"max"	1,271266182
Value__linear_trend_attr_"intercept"	1,272834922
Value__change_quantiles_f_agg_"var"__isabs_False__qh_1.0__ql_0.8	1,273135625
Value__agg_linear_trend_attr_"intercept"__chunk_len_50_f_agg_"mean"	1,273914863
Value__agg_linear_trend_attr_"intercept"__chunk_len_5_f_agg_"min"	1,278552678
Value__agg_linear_trend_attr_"intercept"__chunk_len_50_f_agg_"var"	1,282454741
Value__spkt_welch_density__coeff_8	1,308276173
Value__fft_aggregated__aggtype_"variance"	1,350833921
Value__spkt_welch_density__coeff_5	1,42412245
Value__spkt_welch_density__coeff_2	1,444046976
Value__time_reversal_asymmetry_statistic__lag_1	8,903942068

Πίνακας 20: Αποτελέσματα πειραμάτων με πρόσθετα χαρακτηριστικά

Βιβλιογραφία

Συγγράμματα:

- «Γραμμικός Προγραμματισμός» Γ. Σίσκος, Εκδόσεις Νέων Τεχνολογιών Ιδιωτική Κεφαλαιουχική Εταιρεία
- «Αναγνώριση Προτύπων» S. Theodoridis, K. Koutroumbas, Broken Hill Publishers
- «Εξόρυξη και Ανάλυση Δεδομένων: Βασικές Έννοιες και Αλγόριθμοι», Mohammed J. Zaki, Wagner Meira Jr., Εκδόσεις Κλειδάριθμος
- «Τεχνητή Νοημοσύνη, Μία σύγχρονη προσέγγιση», Stuart Russel, Peter Norvig, Εκδόσεις Κλειδάριθμος

Διδακτικό Υλικό:

- Σημειώσεις Ευάγγελου Δερματά για το μάθημα «Αναγνώριση Προτύπων»
- Stanford CS 230 – Deep Learning
- Stanford CS 231 – Mobile Computer Vision

Ιστοσελίδες:

- <https://www.investopedia.com>
- <https://www.ibm.com/learn>
- <https://www.spglobal.com>
- <https://tsfresh.readthedocs.io>
- <https://keras.io>
- <https://scikit-learn.org>

Δημοσιεύσεις:

- Optimal portfolios using Linear Programming models, Christos Papahristodoulou
- Markowitz H (1952). Portfolio selection, Journal of Finance
- Young (1998). A minimax-portfolio selection rule with linear programming solution, Management Science