



ΤΜΗΜΑ ΛΟΓΙΣΤΙΚΗΣ ΚΑΙ  
ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ  
ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ

ΠΡΟΓΡΑΜΜΑΤΑ  
ΜΕΤΑΠΤΥΧΙΑΚΩΝ  
ΣΠΟΥΔΩΝ

## Πρόγραμμα Μεταπτυχιακών Σπουδών στη Λογιστική Φορολογία και Χρηματοοικονομική Διοίκηση

Όνοματεπώνυμο: ΒΑΡΕΛΑΣ ΧΑΡΑΛΑΜΠΟΣ

---

Επιβλέπων: ΑΧΙΛΛΕΑΣ ΖΑΠΡΑΝΗΣ

Θέμα Διπλωματικής Εργασίας:

**ΠΡΟΒΛΕΨΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΩΝ ΜΕΓΕΘΩΝ ΜΕ ΧΡΗΣΗ  
ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ**

Αφιερώνεται στους Γονείς μου,  
Αθανάσιο & Μαρία

## Ευχαριστίες

Ευχαριστώ τον Δρ. Αχιλλέα Ζαπράνη, Καθηγητή του Τμήματος Λογιστικής και Χρηματοοικονομικής για την ανάθεση της εργασίας αυτής και την συμπαράστασή του κατά την διάρκεια της εκπόνησης της.

Ευχαριστώ επίσης, την Κα. Ευτυχία Ιωακειμίδου την γραμματέα του Τμήματος Λογιστικής Φορολογίας και Χρηματοοικονομικής Διοίκησης για την επίλυση των θεμάτων κατά την διάρκεια των σπουδών.

# Περιεχόμενα

Ευχαριστίες.....	3
Πρόλογος.....	7
1 Εισαγωγικό Κεφάλαιο.....	8
2 Τεχνητή Νοημοσύνη σε Χρηματοοικονομικές Εφαρμογές.....	11
2.1 Εισαγωγή.....	11
2.2 Ορισμός της ΤΝ.....	11
2.3 Ιστορική Αναδρομή.....	13
2.4 Πλαίσιο Ταξινόμησης συστημάτων Τεχνητής Νοημοσύνης.....	16
2.5 Αξιοπιστία της Τεχνητής Νοημοσύνης.....	17
2.6 Παράγοντες καθιέρωσης Τεχνητής Νοημοσύνης.....	19
2.7 Επενδύσεις στην Τεχνητή Νοημοσύνη.....	20
2.8 Η Τεχνητή Νοημοσύνη στον Τομέα των Χρηματοοικονομικών.....	21
2.9 Πρόβλεψη Μεγεθών Χρηματοοικονομικών Χρονοσειρών.....	23
2.10 Θεωρίες για την πρόβλεψη χρηματαγορών.....	24
2.11 Μέθοδοι Πρόβλεψης.....	25
2.12 Αυτοματοποιημένη Πρόβλεψη .....	25
2.12.1 Αλγοριθμικό Εμπόριο / Algorithmic Trading.....	26
2.12.2 Υψηλής Συχνότητας Εμπορία/ High Frequency Trading.....	26
2.13 Επεξηγήσιμη Τεχνητή Νοημοσύνη (ETN).....	27
3 Ανασκόπηση Μεθόδων Πρόβλεψης .....	29
3.1 Θεμελιώδης Ανάλυση / Fundamental Analysis .....	29
3.2 Τεχνική Ανάλυση / Technical Analysis.....	30
3.3 Στατιστικές τεχνικές Πρόβλεψης.....	33
3.4 Αυτοπαλινδρόμηση / Auto Regressive (AR) .....	33
3.5 Μοντέλο Κινούμενων Μέσων Όρων / Moving Average Model (MA).....	34
3.6 Auto-Regressive Moving Average (ARMA) .....	34
3.7 Auto-Regressive Integrated Moving Average (ARIMA).....	34
3.8 Ασαφής Λογική/Fuzzy Logic.....	35
3.9 Μηχανική Μάθηση (MM) / Machine Learning(ML).....	35
3.10 Χαρακτηριστικά Δεδομένων.....	37
3.10.1 Προεργασία των Δεδομένων / Data Preprocessing.....	38
3.10.2 Μέγεθος Δείγματος Δεδομένων.....	39
3.10.3 Επιλογή Μεταβλητών.....	39
3.10.4 Προκατάληψης-Διακύμανσης (Bias -Variance).....	40
3.11 Εκπαίδευση του μοντέλου.....	41
3.11.1 Διαμερισμός των δεδομένων.....	41
3.12 Επιλογή Μοντέλου .....	43
3.12.1 Μετρικές αξιολόγησης απόδοσης μοντέλου / Model performance evaluation metrics.....	43
3.13 Μέθοδοι Μηχανικής Μάθησης.....	46
3.13.1 Γενετικοί Αλγόριθμοι (ΓΑ) / Genetic Algorithms (GA).....	46
3.13.2 Μέθοδοι Πυρήνων/ Kernel methods .....	47
3.13.2.1 Support vector machine (SVM).....	47
3.13.3 Άλλες μέθοδοι μηχανικής Μάθησης.....	48
3.13.3.1 K-nearest neighbor (KNN).....	48
3.13.3.2 Random forest.....	48
3.14 Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) / Artificial Neural Networks (ANN).....	49
3.14.1 Αλγόριθμος Οπισθοδιάδοσης / Backpropagation Algorithm.....	50
3.14.2 Αλγόριθμοι βελτιστοποίησης .....	51
3.14.2.1 Stochastic Gradient Descent.....	51
3.14.3 Είδη Νευρωνικών Δικτύων.....	52

3.14.4	Νευρωνικό Δίκτυο Προσω τροφοδότησης/Feed-Forward Neural Network.....	52
3.14.5	Multi-Layer Perceptron (MLP).....	53
3.14.6	Radial-basis function network (RBF).....	53
3.14.7	Bayesian Regularized Artificial Neural Network (BRANNs).....	53
3.14.8	Βαθιά Μάθηση (BM) / Deep Learning (DL).....	53
3.14.9	Νευρωνικό Δίκτυο Ανατροφοδότησης / Recurrent Neural Networks (RNN).....	55
3.14.10	Bidirectional Recurrent Neural Network (BRNN).....	56
3.14.11	Πολυεπίπεδο Perceptron Βαθιάς Μάθησης/ Deep Multilayer Perceptron (DMLP)..	56
3.14.12	Νευρωνικό Δίκτυο Περιελιγμού / Convolutional Neural Network (CNN).....	56
3.14.13	Long-Short Term Memory (LSTM) /Μνήμη Μακροπρόθεσμη-Βραχυπρόθεσμη.....	57
3.14.14	Gated Recurrent Units (GRUs).....	59
3.14.15	Περιορισμένες μηχανές Boltzmann / Restricted Boltzmann Machines (RBM).....	59
3.14.16	Deep Belief Network (DBN).....	60
3.14.17	Αυτοκωδικοποιητής / Autoencoder (AE) .....	61
3.14.18	Graph neural network (GNN).....	62
3.14.19	Βαθια Εξαναγκασμένη Μάθηση / Deep Reinforcement learning (RL) .....	62
3.15	Μηχανισμός Προσοχής /Attention Mechanism.....	63
3.16	Εκτίμηση Τεχνητών Νευρωνικών Δικτύων.....	63
3.16.1	Υπερπαράμετροι & Παράμετροι και Νευρωνικών Δικτύων.....	64
3.16.1.1	Υπερπαράμετροι Δομής Δικτύου .....	65
3.16.1.1.1	Αριθμός Κρυφών Επιπέδων και Μονάδων.....	65
3.16.1.1.2	Απόρριψη / Dropout.....	65
3.16.1.1.3	Συνάρτηση Ενεργοποίησης/Activation function.....	66
3.16.1.1.4	Αρχικοποίηση Βαρών Δικτυου / Network Weight Initialization.....	66
3.16.1.2	Υπερπαράμετροι Μάθησης Δικτύου.....	67
3.16.1.2.1	Ρυθμός Μάθησης/ Learning rate.....	67
3.16.1.2.2	Αριθμός Εποχών/ Number of epochs.....	67
3.16.1.2.3	Δείγμα εκπαίδευσης/ Batch.....	68
3.16.1.2.4	Momentum.....	68
3.17	Κριτική Τεχνητών Νευρωνικών Δικτύων .....	68
	Πλεονεκτήματα Τεχνητών Νευρωνικών Δικτύων .....	68
	Μειονεκτήματα Τεχνητών Νευρωνικών Δικτύων .....	69
3.18	Υβριδικά-Συνδυαστικά μοντέλα .....	69
3.18.1	Neuro Fuzzy Συστήματα.....	70
3.18.2	Modular Neural Networks .....	70
3.19	Text Mining.....	70
3.20	Ανάλυση Συναισθήματος / Sentiment Analysis.....	71
4	Βιβλιογραφική έρευνα.....	73
4.1	Μεθοδολογία .....	74
4.2	Περιπτώσεις Ανασκοπήσεων / Case Reviews.....	75
4.3	Χαρακτηριστικά της μελέτης.....	79
4.3.1	Περιοχές Θεμάτων Πρόβλεψης.....	79
4.3.2	Αντικείμενο μελέτης πρόβλεψης.....	80
4.3.3	Χαρακτηριστικά αγορών Χρηματοοικονομικών Μεγεθών.....	81
4.3.4	Πρόβλεψη Τιμής Μετοχής(Stock).....	82
4.3.5	Πρόβλεψη Τιμής Δείκτη / Index Forecating.....	83
4.3.6	Πρόβλεψη Τιμών Αγαθών (Commodities).....	84
4.3.7	Πρόβλεψη Τιμών Αγορας FOREX / Forex price Forecasting.....	85
4.3.8	Προβλεψης Τιμής Ομολόγων / Bond price forecasting.....	85
4.3.9	Πρόβλεψη Τιμής Κρυπτονομισμάτων/ Cryptocurrency price forecasting.....	86
4.3.10	Πρόβλεψη Τάσης / Trend forecasting .....	87
4.3.11	Πρόβλεψη Μεταβλητότητας/Volatility.....	88

4.3.12 Πρόβλεψη Κρίσεων /Forecasting Stock Market Crisis.....	90
4.3.13 Αλγοριθμικό Εμπόριο Αξιών / Algorithmic Trading.....	92
4.4 Δεδομένα .....	93
4.4.1 Τύπος Δεδομένων.....	93
4.4.2 Μέγεθος δείγματος.....	95
4.4.3 Επεξεργασία Δεδομένων.....	95
4.4.4 Εξαγωγή και Επιλογή Χαρακτηριστικών.....	96
4.4.5 Αριθμός Μεταβλητών Εισόδου.....	96
4.4.6 Διαμερισμός δεδομένων .....	97
4.5 Περιβάλλον Ανάπτυξης/Development Environment.....	97
4.6 Υλικό Εκτέλεσης Τεχνικών Βαθιάς Μάθησης.....	97
4.7 Αξιολόγηση Μοντέλων.....	97
4.7.1 Μετρικές ταξινόμησης.....	97
4.7.1.1 Μετρικές Παλινδρόμησης/ Regression metrics.....	98
4.7.1.2 Ανάλυση Κερδοφορίας.....	98
4.7.1.3 Ανάλυση Σημαντικότητας/ Significance Analysis.....	98
4.8 Αναπαραγωγή των Μοντέλων.....	99
4.9 Υβριδικές Μέθοδοι.....	99
5 Συμπεράσματα.....	101
6 Επίλογος.....	104
Παραρτημα.....	110

## Πρόλογος

Εν έτει 2022, η ένταση των εφαρμογών Τεχνητής Νοημοσύνης επιβεβαιώνεται σε παγκόσμιο επίπεδο από τον αριθμό των επιστημονικών δημοσιεύσεων, την έρευνα και ανάπτυξη αλγορίθμων και σχετικού λογισμικού από τις εταιρείες τεχνολογίας και τα εκπαιδευτικά ιδρύματα, και της πληθώρας των εφαρμογών που αναπτύσσονται.

Η Τεχνητή Νοημοσύνη αποτελώντας τώρα τομέα επιχειρηματικής δραστηριότητας και εργαλείο καινοτομίας αιχμής, εμπλέκει την έρευνα και την τεχνολογία με χρηματοδοτήσεις από οργανισμούς και ιδιώτες και πολιτικές προώθησης κρατών. Προσδίδει ανταγωνιστικό πλεονέκτημα και αξία και καθιερώνει την χρήση της με γνώμονα την ηθική.

Σημείο αναφοράς των τελευταίων δεκαετιών αποτελεί παράλληλα με τις μεθόδους Μηχανικής Μάθησης (Machine Learning) η ανάπτυξη των εφαρμογών Τεχνητών Νευρωνικών Δικτύων (Artificial Neural Networks) και η τελευταία εξέλιξη τους στον τομέα της Βαθιάς Μάθησης (Deep Learning).

Το παγκόσμιο χρηματοοικονομικό σύστημα δέχεται ισχυρή επίδραση από τις τεχνολογίες εφαρμογών Τεχνητής Νοημοσύνης σε κάθε τομέα δραστηριότητας του. Ο τομέας της πρόβλεψης τιμών χρηματοοικονομικών μεγεθών αποτελεί σημείο σημαντικών προσπαθειών από ερευνητές, επιστήμονες, μηχανικούς και τεχνικούς που υλοποιούν τις εφαρμογές. Ειδικότερα ο τομέας της πρόβλεψης των μεγεθών χρηματοοικονομικών χρονοσειρών των αγορών.

Στις δημοσιεύσεις σε επιστημονικά περιοδικά και τεχνικές οδηγίες εταιρειών στο διαδίκτυο δημιουργείται ένας σημαντικός όγκος γνώσης που προκύπτει από την δοκιμή και εφαρμογή των αλγορίθμων που συνεχώς αναπτύσσονται, πάνω σε μεγάλο όγκου διαθέσιμα δεδομένα χρηματαγορών.

Η ανάλυση των ερευνών (surveys), ανασκοπήσεων (reviews) και δημοσιεύσεων της χρήσης των Τεχνητών Νευρωνικών Δικτύων στον τομέα της πρόβλεψης χρηματοοικονομικών μεγεθών χρονοσειρών προς εξαγωγή κριτικών και χρήσιμων συμπερασμάτων αποτελεί το αντικείμενο αυτής της εργασίας.

2022-12-15, Ν. Ορεστιάδα, Έβρος

Βαρελάς Χαράλαμπος

# 1 Εισαγωγικό Κεφάλαιο

Η Τεχνητή Νοημοσύνη από προσπάθεια επιστημόνων την δεκαετία του 1950, να αναπτύξουν μηχανές, μέσω της ψηφιακής τεχνολογίας και της επιστήμης των υπολογιστών που μιμούνται την ανθρώπινη νοημοσύνη, ενεργούν και δρουν αυτόνομα, αποτελεί πλέον την 5η βιομηχανική επανάσταση όπου οι μηχανές συνεργάζονται με τους ανθρώπους για την επίλυση των θεμάτων [(Noble et al 2022)]. Αποτελεί μοχλό οικονομικής ανάπτυξης όπως φαίνεται από τις στατιστικές του ΟΟΣΑ και διεπιστημονικό πεδίο όπως διαπιστώνεται από το πλήθος των δημοσιευμένων μελετών. λειτουργώντας προς όφελος και μετασχηματίζοντας την ανθρώπινη κοινωνία.

Η Τεχνητή Νοημοσύνη αποτελεί διασπαστική τεχνολογία (disruptive technology) και κινητήριος δύναμη (driving force) στην επιστημονική έρευνα και στις επιχειρήσεις, με ισχυρές άμεσες και έμμεσες επιδράσεις σε όλους τους τομείς της ανθρώπινης ζωής.

Η λειτουργία της βασίζεται σε αλγορίθμους Μηχανικής Μάθησης που μαθαίνουν ένα θέμα από το παράδειγμα που δίνουν διαθέσιμα δεδομένα σε συνδυασμό με αναλυτικές και στατιστικές τεχνικές. Ειδικότερη ανάπτυξη παρουσιάζει ο τομέας των Τεχνητών Νευρωνικών Δικτύων και της πρόσφατης εξέλιξης τους στην Βαθιά Μάθηση.

Το 2022, η χρήση του όρου Fintech (από το Financial Technology) χρησιμοποιείται για να περιγράψει καινοτομίες, ψηφιακών τεχνολογιών, στον τομέα των χρηματοοικονομικών. Αφορά την δημιουργία νεοτερικών εφαρμογών υπηρεσιών, μοντέλων, διαδικασιών που υποστηρίζονται από τις αναβαθμίσεις των ψηφιακών τεχνολογιών. Χαρακτηριστικό τους είναι η σημαντική διαφοροποίηση από τις παραδοσιακές χρηματοοικονομικές υπηρεσίες και η ισχυρή επίδραση μετασχηματισμού στις αγορές, με αντικρουόμενα συμφέροντα με την παραδοσιακή πρακτική των χρηματοοικονομικών.

Οι ψηφιακές τεχνολογίες που πυροδοτούν αυτή την ανάπτυξη αναφέρονται ως ABCD, αντιστοιχίζοντας τα πρώτα γράμματα του αλφάβητου με τις σημαντικότερες τεχνολογίες της Fintech : Artificial Intelligence (AI)/Τεχνητή Νοημοσύνη (TN), η Blockchain, το Cloud Computing, και τα Big Data (δεδομένα μεγάλου όγκου).

Χρησιμοποιείται σε μεγάλο εύρος στον χρηματοοικονομικό τομέα, στον ασφαλιστικό τομέα, στον τραπεζικό τομέα, στις επενδύσεις, στην εμπορία χρηματοοικονομικών εργαλείων (trading), στην διαχείριση κινδύνου και περιλαμβάνει εφαρμογές που έχουν την ικανότητα αυτοματοποιημένης και ευφυούς εξαγωγής συμπερασμάτων από τα δεδομένα.

Οι επιχειρήσεις εξετάζουν στρατηγικές με τις οποίες θα αξιοποιήσουν τα δεδομένα που διαθέτουν και θα ενσωματώσουν τις τεχνολογίες Τεχνητής Νοημοσύνης, στην επιχειρησιακή τους δομή. Εντύπωση προκαλεί το γεγονός πως μέσα στην υγειονομική κρίση του COVID-19 πολλές αγορές



και ειδικά οι αναπτυσσόμενες εξέλιξαν τις εφαρμογές Fintech ως αντίμετρα στις πολιτικές που ακολουθήθηκαν.

Η νεότερη επανάσταση που συντελείται στην Τεχνητή Νοημοσύνη είναι η εξέλιξη των μεθόδων των Τεχνητών Νευρωνικών Δικτύων σε Δίκτυα Βαθιάς Μάθησης (Deep Learning). Τα μοντέλα αυτά των νευρωνικών δικτύων με την επαυξημένη δυνατότητα μάθησης συνέβαλλαν σε μια πλημμυρίδα νέων εφαρμογών, νέων μεθοδολογιών και στρατηγικών δημιουργίας των μοντέλων.

Ο τομέας της Τεχνητής Νοημοσύνης αποτελεί ένα πεδίο που χαρακτηρίζεται για την ανοιχτή πρόσβαση στις εφαρμογές του, με κοινοποιήσεις των εφαρμογών σε συνέδρια και ιστοσελίδες αποθετηρίων κώδικα. Η πρόσβαση στις επιτυχημένες στρατηγικές δημιουργίας των μοντέλων μπορούν να βοηθήσουν τους ερευνητές να εξελίξουν την εργασία τους. Η δημιουργία ενός μοντέλου με συστηματικό τρόπο αποτελεί ένα σημαντικό τομέα της ανάπτυξης της ΤΝ. Παράλληλα υπάρχει μεγάλο ενδιαφέρον στους ερευνητές για να μοιραστούν τα συμπεράσματά τους, τις έρευνες τους και τον κώδικά τους.

Ο τομέας της Πρόβλεψης χρηματοοικονομικών μεγεθών έχει ευρύ πεδίο τομέων εφαρμογών, μια από τις οποίες είναι και η πρόβλεψη μεγεθών χρηματοοικονομικών χρονοσειρών. Η προσέγγισή τους γίνεται με χρήση παραδοσιακών εργαλείων όπως στατιστική, τεχνική ανάλυση, θεμελιώδης ανάλυση, παραδοσιακές μέθοδοι Μηχανικής Μάθησης, Τεχνητά Νευρωνικά Δίκτυα και Νευρωνικών Δικτύων Βαθιάς Μάθησης.

Η παρακάτω εργασία πραγματοποιεί έρευνα σε εφαρμογές και ανασκοπήσεις από την ακαδημία, τον τομέα των επιχειρήσεων και τις σύγχρονες τάσεις από ερευνητικά κέντρα, πανεπιστήμια, επιχειρήσεις και την βιομηχανία. Επικεντρώνει σε εφαρμογές τεχνητής νοημοσύνης σε χρηματοοικονομικά δεδομένα χρονοσειρών ως προς την ικανότητα πρόβλεψης των μεθόδων και αλγορίθμων. Η ανασκόπηση των ερευνών που έχουν πραγματοποιηθεί μπορεί να οδηγήσει στην ταχύτερη και αποδοτικότερη ανάπτυξη των εφαρμογών ΤΝ.

Η υπόθεση που διατυπώνεται είναι αν η μελέτη προηγούμενων μελετών σχετικά με την εφαρμογή της ΤΝ στην περίπτωση της πρόβλεψης οροσειρών χρηματοοικονομικών δεδομένων μπορεί να οδηγήσει σε συμπεράσματα για την δημιουργία καλών μοντέλων και να αναδείξει τα στοιχεία της επιτυχίας τους..

Η διπλωματική αυτή εργασία προσπαθεί να απαντήσει στα παρακάτω ερωτήματα:

- ▶ *Ποιες είναι οι εφαρμογές μοντέλων, μέσω αλγορίθμων και μεθόδων, για την πρόβλεψη χρηματοοικονομικών μεγεθών χρονοσειρών;*
- ▶ *Ποιες είναι οι παράμετροι που επηρεάζουν την δημιουργία μοντέλων στον χρηματοοικονομικό τομέα και ειδικότερα στην πρόβλεψη χρηματοοικονομικών μεγεθών χρονοσειρών;*

Σκοπός της παρακάτω εργασίας είναι η ανάλυση ερευνών, των βασικών μεθόδων και παραμέτρων θεμάτων Τεχνητής Νοημοσύνης με έμφαση στα Τεχνητά Νευρωνικά Δίκτυα και τα Νευρωνικά

Δίκτυα Βαθιάς Μάθησης, που χρησιμοποιούνται στην πρόβλεψη μεγεθών Χρηματοοικονομικών χρονοσειρών και η εξαγωγή χρήσιμων συμπερασμάτων.

**Στόχοι της εργασίας είναι:**

- ▶ Να περιγράψει την διαδικασία πρόβλεψης χρηματοοικονομικών μεγεθών χρονοσειρών και τις βασικές μεθόδους που χρησιμοποιούνται.
- ▶ Να αναλύσει την εφαρμογή των μεθόδων Τεχνητής Νοημοσύνης και ειδικά των Νευρωνικών Δικτύων και των Δικτύων Βαθιάς Μάθησης.
- ▶ Να αναλύσει τα αποτελέσματα των ερευνών και των δημοσκοπήσεων αυτών.
- ▶ Να εξάγει και να παρουσιάσει τα αποτελέσματα της έρευνας.

**Στην πτυχιακή αυτή εργασία παρουσιάζονται τα εξής :**

Το 1ο κεφάλαιο αποτελεί μια εισαγωγή της δομής αυτής της μελέτης και του τρόπου εργασίας, παρουσιάζει τις έρευνες που εξετάστηκαν και την μεθοδολογία έρευνας.

Το 2ο κεφάλαιο περιγράφει την Τεχνητή Νοημοσύνη και την εφαρμογή της στις χρηματοοικονομικές χρονοσειρές.

Το 3ο κεφάλαιο αναλύει τους κυριότερους αλγόριθμους Τεχνητής Νοημοσύνης και ιδιαίτερα τα Τεχνητά Νευρωνικά Δίκτυα και τα Δίκτυα Βαθιάς Μάθησης-Deep Neural Networks.

ΤΟ 4ο κεφάλαιο παρουσιάζει την βιβλιογραφική έρευνα.

ΤΟ 5ο κεφάλαιο παρουσιάζει τα συμπεράσματα των ερευνών.

## 2 Τεχνητή Νοημοσύνη σε Χρηματοοικονομικές Εφαρμογές

### 2.1 Εισαγωγή

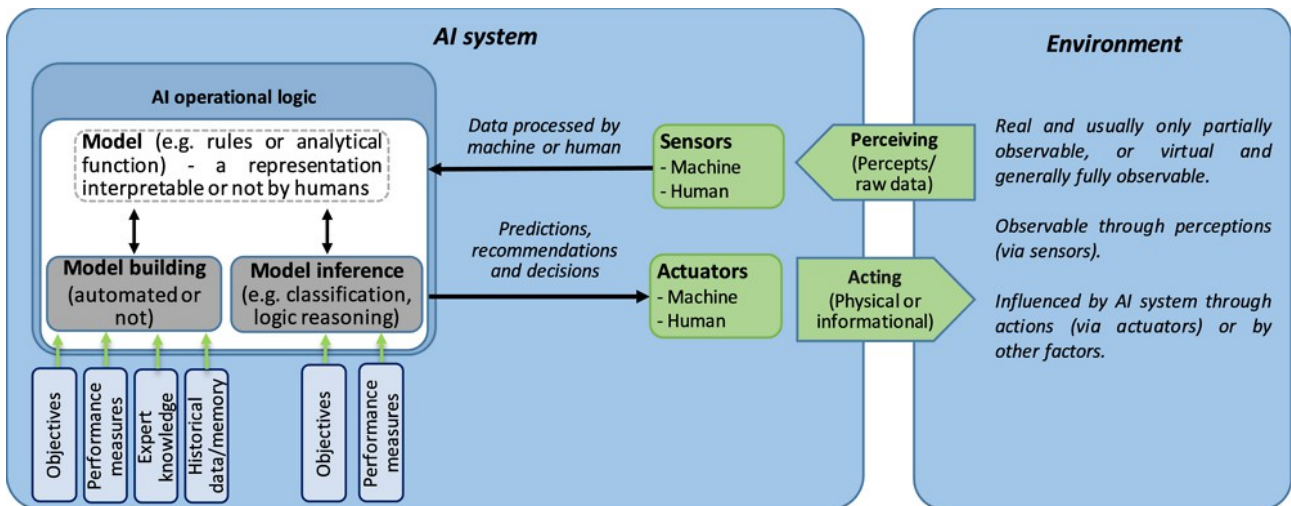
Η Τεχνητή Νοημοσύνη αποτελεί μια διαχρονική επιστημονική προσπάθεια η οποία συνίσταται στην δημιουργία ικανότητας αυτόνομης μάθησης και νόησης στις μηχανές και στους ηλεκτρονικούς υπολογιστές. Εφαρμόζεται σε διαθέσιμα δεδομένα δημιουργώντας μοντέλα που έχουν την δυνατότητα να εξάγουν γνωστικά συμπεράσματα, η εφαρμογή των οποίων διαμορφώνει το εξωτερικό περιβάλλον. Η υιοθέτηση αυτών των μοντέλων προϋποθέτει ότι είναι ασφαλή προς χρήση, ότι μπορούν να αιτιολογηθούν ως προς τον τρόπο που δημιούργησαν το αποτέλεσμα, ότι θα λειτουργήσουν με τρόπο επωφελή για τον άνθρωπο. Η ευρεία ανάπτυξη της ΤΝ θέτει την ανάγκη για την δημιουργία ενός πλαισίου για αξιόπιστη, δίκαιη, ηθική και εύρωστη εφαρμογή της Τεχνητής Νοημοσύνης. Εταιρείες και οργανισμοί καθιερώνουν πολιτικές και διαδικασίες για την ομαλή χρήση των μοντέλων Τεχνητής Νοημοσύνης.

Η εφαρμογή της στο πεδίο των χρηματοοικονομικών πραγματοποιείται εδώ και δεκαετίες με διάφορες προσεγγίσεις. Τα χρηματοοικονομικά ενσωματώνουν κάθε τεχνολογική καινοτομία όπως και την Τεχνητή Νοημοσύνη σε όλους του τομείς τους. Η πρόβλεψη μεγεθών χρηματοοικονομικών χρονοσειρών είναι μια από τις εφαρμογές που συγκεντρώνει το ενδιαφέρον της Ακαδημίας και των εταιρειών.

### 2.2 Ορισμός της ΤΝ

Η Τεχνητή Νοημοσύνη αποτελεί την μίμηση της ανθρώπινης νοημοσύνης στην λειτουργία ενός συστήματος υπολογιστών, δίνοντας την δυνατότητα επίλυσης θεμάτων χωρίς την ανθρώπινη παρέμβαση, εκτελώντας εργασίες όπως η οπτική αντίληψη, η αναγνώριση φωνής, η λήψη αποφάσεων, ανακαλύπτοντας νοήματα και εξάγοντας συμπεράσματα.

Ο Οργανισμός Οικονομικής Συνεργασίας και Ανάπτυξης (ΟΟΣΑ) ορίζει την Τεχνητή Νοημοσύνη ως “ένα μηχανο-κεντρικό (machine-based) σύστημα το οποίο παράγει ένα αποτέλεσμα που είναι ικανό να επηρεάσει το περιβάλλον (με προβλέψεις, συστάσεις, αποφάσεις) για ένα δοθέν σύνολο από στόχους. Χρησιμοποιεί δεδομένα και εισόδους ανθρωπογενή και δημιουργημένα από μηχανές για να (1) αντιληφθεί πραγματικά η εικονικά περιβάλλοντα (2) να περιλάβει αυτές τις αντιλήψεις σε μοντέλα μέσω ανάλυσης με έναν αυτοματοποιημένο τρόπο (μηχανική μάθηση) ή με ανθρώπινη παρέμβαση και (3) να χρησιμοποιήσει την αναφορά των μοντέλων για να δημιουργήσει επιλογές για αποτελέσματα. Τα συστήματα ΤΝ έχουν σχεδιαστεί για να λειτουργούν με πολλαπλά επίπεδα αυτονομίας”. [(OECD 2022)] Ως σύστημα θα εννοείται κάθε μεμονωμένο στοιχείο, λογισμικό ή υλικό, που βασίζεται στην ΤΝ.



Εικόνα 1: Λεπτομερές θεματική αναπαράσταση ενός Συστήματος ΤΝ. Πηγή :[(OECD 2019)]

Η Ευρωπαϊκή Επιτροπή δίνει τον παρακάτω ορισμό για την Τεχνητή Νοημοσύνη:Ευρωπαϊκής Επιτροπής για την ΤΝ:

«Τα συστήματα τεχνητής νοημοσύνης (ΤΝ) είναι συστήματα λογισμικού (ή ενδεχομένως και υλικού) που σχεδιάζονται από ανθρώπους και, βάσει ενός δεδομένου σύνθετου στόχου, ενεργούν στην υλική ή ψηφιακή διάσταση με το να αντιλαμβάνονται το περιβάλλον τους μέσω της απόκτησης δεδομένων, να ερμηνεύουν τα δομημένα ή αδόμητα δεδομένα που έχουν συλλεχθεί, να προβαίνουν σε συλλογισμούς με βάση τις γνώσεις ή να επεξεργάζονται τις πληροφορίες που εξάγονται από αυτά τα δεδομένα και να αποφασίζουν ποια είναι η βέλτιστη ενέργεια (ή οι βέλτιστες ενέργειες) που θα πρέπει να εκτελέσουν για να επιτύχουν τον δεδομένο στόχο. Τα συστήματα ΤΝ μπορεί είτε να χρησιμοποιούν συμβολικούς κανόνες είτε να μαθαίνουν ένα αριθμητικό μοντέλο, και μπορεί επίσης να προσαρμόζουν τη συμπεριφορά τους με το να αναλύουν πώς επηρεάζεται το περιβάλλον από τις προηγούμενες ενέργειές τους.

Ως επιστημονικό πεδίο, η ΤΝ περιλαμβάνει διάφορες προσεγγίσεις και τεχνικές, όπως η μηχανική μάθηση (συγκεκριμένα παραδείγματα της οποίας είναι η βαθιά μάθηση και η ενισχυτική μάθηση), η μηχανική συλλογιστική (που περιλαμβάνει τον σχεδιασμό, τον προγραμματισμό, την αναπαράσταση και τη συλλογιστική γνώσης, την αναζήτηση και τη βελτιστοποίηση) και η ρομποτική (που περιλαμβάνει έλεγχο, αντίληψη, αισθητήρες και ενεργοποιητές, καθώς και την ενσωμάτωση όλων των άλλων τεχνικών σε κυβερνο-υλικά συστήματα).» [(ΕΕ 2019)]

Μια εναλλακτική σημασία δίνεται από τον Kaplan (2016) το νόημα της ΤΝ, είναι η ικανότητα μιας μηχανής να κάνει κατάλληλες γενικεύσεις, σε έναν εύλογο χρόνο βασιζόμενος σε ελάχιστα δεδομένα. “Όσο μεγαλύτερο είναι το εύρος των εφαρμογών, το γρηγορότερο σχηματίζονται ι διαπιστώσεις με την ελάχιστη πληροφορία, και πιο νοήμων είναι η συμπεριφορά.”.

Η ΤΝ χωρίζεται σε ασθενής, ισχυρή και γενική αναλόγως του “βαθμού νοημοσύνης” του συστήματος σε σύγκριση με την ανθρώπινη νοημοσύνη.

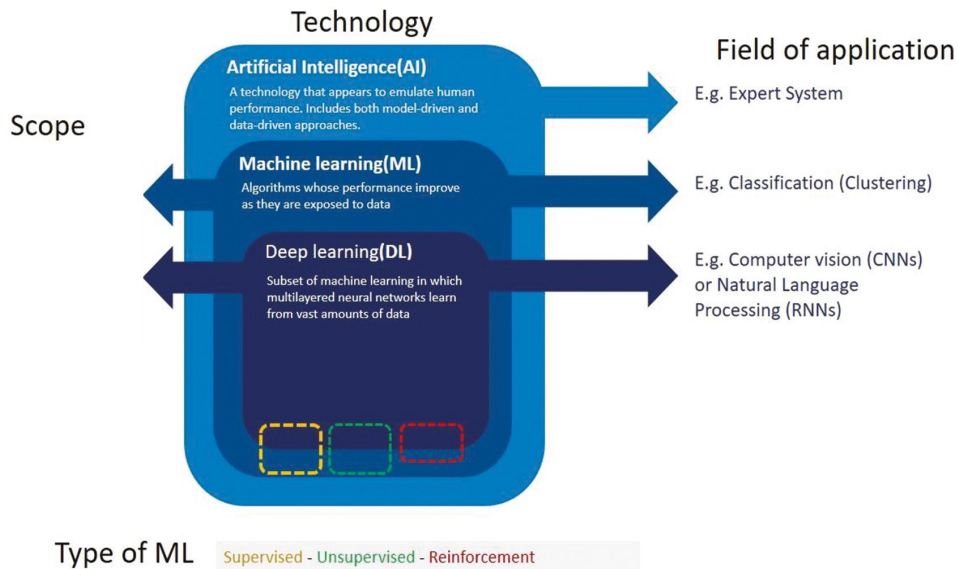
## 2.3 Ιστορική Αναδρομή

Παρόλο που η ύπαρξη Τεχνητής Νοημοσύνης, αναφέρεται στην Ελληνική Μυθολογία, στο δημιούργημα Ηφαίστου, του Τάλου μια μηχανή φρουρό, που περιέτρεχε το νησί της Κρήτης και τα μηχανικά αυτόματα αγάλματα της Ελλάδας και Αιγύπτου, η ανάπτυξη την ψηφιακής τεχνητής νοημοσύνη συντελέστηκε τον 20ο αιώνα μ.Χ.

Η αρχή της ΤΝ προσδιορίζεται το 1943 με την εργασία των McCulloch-Pitts που περιγράψανε μαθηματικά μοντέλα (perceptrons) των νευρώνων του εγκεφάλου. Το 1949 ο Καναδός Φυσιολόγος Donald Hebb πρότεινε μια θεωρία μάθησης με νευρωνικά δίκτυα. Το 1950 ο Βρετανός μαθηματικός, Allan Turing εξέφρασε την απορία αν “Μπορούν να σκεφτούν οι μηχανές;” και δημιούργησε το Test Turing μέσω του οποίου ορίζεται αν μια μηχανή μπορεί να σκεφτεί. Το 1951 ο γνωστός επιστήμονας Marvin Minsky και ο φυσικός Dean Edmonds δημιούργησαν το πρώτο σύστημα νευρωνικών δικτύων βασισμένο στον κανόνα μάθησης του Hebb και των μοντέλων των McCulloch-Pitts με την ονομασία SNARC. Το 1959 ο John McCarthy καθιερώνει τον όρο Τεχνητή Νοημοσύνη, ο Frank Rosenblatt εφηύρε το perceptron, ενώ δημιουργείται με την συμβολή των Newell Allen, J.A. Shaw και Herbert Simon ο General Problem Solver, ένα πρόγραμμα προσομοίωσης ανθρωπίνων μεθόδων επίλυσης προβλημάτων. Το 1960 παρουσιάστηκε από τους Widrow και Hoff ο κανόνας μάθησης Δέλτα. Το 1961 αναπτύσσεται το πρώτο βιομηχανικό ρομπότ το Unimate που χρησιμοποιήθηκε σε γραμμή παραγωγής αυτοκινήτων. Ο ενθουσιασμός της δεκαετίας 1960 ακολούθησε η στασιμότητα της ΤΝ την δεκαετία του 1970 με τις τεχνικές δυσκολίες που συνάντησε το πεδίο. Το 1980 άρχισε η αναβίωση της ΤΝ, με την δημιουργία των δικτύων ανατροφοδότησης από τον John Hopfield και οι δημοσιεύσεις των White και Hornik, Stinchcombe και White για τα δίκτυα radial basis.

Το 2010 εισάγεται ο όρος Deep Learning για να περιγράψει τα Δίκτυα Βαθιάς Μάθησης (DNN) που αποτελούν νευρωνικά δίκτυα περισσότερων επιπέδων ικανά να μαθαίνουν εις βάθος από μεγάλο όγκο δεδομένα και να εξάγουν πολύπλοκες σχέσεις από τα δεδομένα. Καθιερώθηκε ως νέα μέθοδος ανάλυσης ως συνισταμένη τριών παραγόντων, της συγκέντρωσης μεγάλων όγκων δεδομένων (Big Data), την καθιέρωση των GPUs ως υλικό μέρος υπολογιστών με την δυνατότητα παράλληλης επεξεργασίας και την εξέλιξη της οικογένειας αλγορίθμων CNN. [(Jiang 2022)]

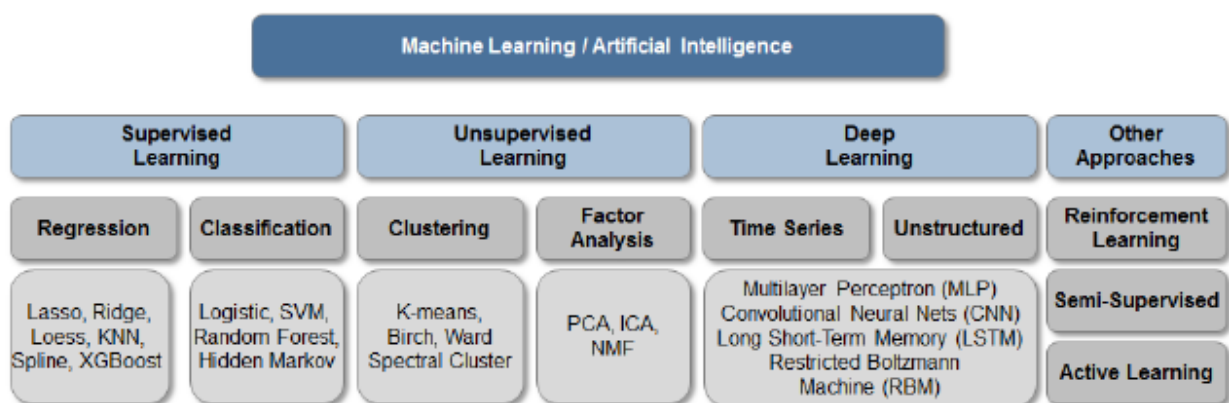
Περαιτέρω εξέλιξη των μεθόδων ΤΝ αναμένεται να γίνει με την επέκταση της συγκέντρωσης των δεδομένων μέσω αισθητήρων Internet of Things (Io) και η εκτέλεση διαδικασιών τεχνητής νοημοσύνης μέσω Κβαντικών Υπολογιστών.



**Εικόνα 2: Τεχνητή Νοημοσύνη, Μηχανική Μάθηση, Βαθιά Μάθηση, Πηγή [(OECD 2022)]**

Η Τεχνητή Νοημοσύνη αποτελεί ένα ευρύ πεδίο μεθόδων όπως η μηχανική μάθηση, επεξεργασία φυσικής γλώσσας (ΦΕΦ ή NLP), σύνθεση κειμένου και λόγου, αναπαράσταση γνώσης ρομποτική, έμπειρα συστήματα, σχεδιασμός υπό την έννοια της εξαγωγής συμπερασμάτων για δοθέντες στόχους. Η Μηχανική Μάθηση (MM) οδηγεί τον αλγόριθμο να μάθει από την εμπειρία σε διαθέσιμα δεδομένα, να δημιουργήσει ένα μοντέλο και να μετρήσει την εφαρμογή του. Κάποιες παράμετροί του μοντέλου ορίζονται από τον άνθρωπο, οδηγώντας την εξέλιξή του. Στο μοντέλο εξετάζεται η ικανότητά του να προβλέψει νέα δεδομένα. Κεντρικό σημείο στην ΤΝ αποτελεί η επεξεργασία μεγάλου όγκου δεδομένων, πολλών τύπων.

Παράλληλα, η ΤΝ αποτελεί ξεχωριστό κλάδο της επιστήμης υπολογιστών ο οποίος ασχολείται με την δημιουργία λογισμικού υπολογιστών που ενσωματώνει την Μηχανική Μάθηση, και μπορεί να μαθαίνει αυτόνομα [Britannica (2022)],



**Εικόνα 3: Κατηγορίες ΤΝ, Πηγή: [(Kolanovic & Krishnamachari 2017)]**

Ένας διαχωρισμός της TN αποτελεί η Συμβολική TN/Symbolic AI, η οποία προσπαθεί να αντιγράψει την νοημοσύνη αναλύοντας την γνώση ανεξάρτητα από την βιολογική δομή του εγκεφάλου με όρους συμβόλων. Η δεύτερη προσέγγιση αποτελεί την Connectionist AI /Συνδεδετική TN, περιλαμβάνει την προσομοίωση των τεχνητών νευρώνων του ανθρωπίνου εγκεφάλου [(Britannica 2022d)].

Αρχικά οι αλγόριθμοι Μηχανικής Μάθησης αφορούσαν αλγόριθμους και στατιστικές τεχνικές που μαθαίνουν από τα δεδομένα, όπως τα Τεχνητά Νευρωνικά Δίκτυα, τα οποία εξελίχθηκαν στα Τεχνητά Νευρωνικά Δίκτυα Βαθιάς Μάθησης (Deep Learning) τα οποία έχουν την ικανότητα να εξάγουν πολύπλοκα πρότυπα και δομές από μεγαλύτερου όγκου δεδομένα. Τα Τεχνητά Νευρωνικά Δίκτυα μπορούν να αντιγράψουν και να μιμηθούν την λειτουργία του ανθρώπινου εγκεφάλου καθώς λειτουργούν με τον ίδιο τρόπο που λειτουργούν τα νευρικά κύτταρα στον ανθρώπινο εγκέφαλο. Ωστόσο δεν είναι πλήρως αυτόματα καθώς ο άνθρωπος πρέπει να ορίσει τις παραμέτρους του ΤΝΔ για να βρει το βέλτιστο μοντέλο και το αντιπροσωπευτικό δείγμα των δεδομένων.. Τα Νευρωνικά Δίκτυα Βαθιάς απεναντίας διαχειρίζονται μεγαλύτερα σύνολα εκπαίδευσης, χαρακτηρίζονται από λιγότερη ή καθόλου ανθρώπινη παρέμβαση, μεγαλύτερη ακρίβεια αλλά και μεγαλύτερους χρόνους εκπαίδευσης, χρήση περισσότερων υπολογιστικών πόρων.

Οι δυνατότητες που δημιουργούνται από την χρήση της Τεχνητής Νοημοσύνης είναι :

- Καλύτερη, πολύπλευρη και πολύπλοκη ανάλυση των δεδομένων, καλύτερη κατανόηση της φύσης των δεδομένων.
- Αυτοματοποίηση εργασιών που χρειαζόταν την ανθρώπινη παρέμβαση
- Ικανότητα μοντέλων να μαθαίνουν μόνα τους και να επάγονται σε αλλαγές του περιβάλλοντος των δεδομένων
- Η ανακάλυψη νοήματος και χρήσιμων προτύπων σε μεγάλες ποσότητες δεδομένων κάθε τύπου όπως κείμενο, εικόνα ήχος, βίντεο.
- Καλύτερες αποφάσεις, αξιοποιώντας την αξία των δεδομένων

Στην μελέτη των [(Annor-Antwi & Al-Dherasi 2022)] συμπεραίνεται ότι η πρόβλεψη έχει σημαντική επίδραση στις επιχειρήσεις. Ακριβής πρόβλεψη συνεπάγεται καλύτερος σχεδιασμός των επιχειρήσεων, καλύτερη χρηματοοικονομική ανάλυση για το μέλλον, ικανότητα για μείωση και αποφυγή ρίσκων, και ευημερία των επιχειρήσεων. Η πρόβλεψη έχει σημαντική επίδραση στις συναλλαγές μετοχών, καθώς οι αγοραστές ή πωλητές που χρησιμοποιούν την πρόβλεψη έχουν την ικανότητα μεγιστοποίησης των κερδών τους.

## 2.4 Πλαίσιο Ταξινόμησης συστημάτων Τεχνητής Νοημοσύνης

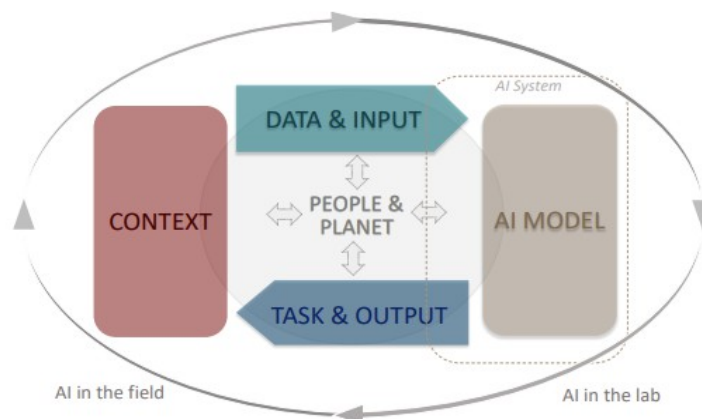
Ο ΟΟΣΑ για να προωθήσει την κοινή κατανόηση της ΤΝ, έχει αναπτύξει ένα πλαίσιο χαρακτηρισμού και ταξινόμησης των συστημάτων ΤΝ, την καταγραφή τους σε ανάλογα μητρώα και ευρετήρια, έτσι ώστε να υποστηρίξει ανάλογες εφαρμογές ανά τομέα ΤΝ και την αξιολόγηση και την διαχείριση των κινδύνων. [(OECD 2022)]

Το πλαίσιο αυτό ταξινομεί τα συστήματα ΤΝ βάση των εξής κατηγοριών:

1. **Της ωφέλειας του ανθρώπου και του πλανήτη**, από την εφαρμογή της ΤΝ.
2. **Το οικονομικό περιβάλλον και τομέας δραστηριότητας**, στο οποίο υλοποιείται ένα σύστημα ΤΝ, η επιχειρηματική λειτουργία και το μοντέλο, η κρισιμότητα της εφαρμογής, η κλίμακα και επίπτωση εφαρμογής, και η τεχνολογική του ωριμότητα.
3. **Τα δεδομένα και οι μεταβλητές εισόδου**, βάση των οποίων δημιουργούνται τα μοντέλα. Χαρακτηριστικά τους είναι η προέλευσή τους, ο τρόπος συλλογής τους (μηχανικός ή με παρέμβαση του ανθρώπου), η δομή (structure) και η μορφή τους (format) και οι ιδιότητές τους. Περαιτέρω χωρίζονται σε δεδομένα που συγκεντρώνονται στο εργαστήριο και δεδομένα που προέρχονται από το πεδίο δραστηριότητας.
4. **Το μοντέλο Τεχνητής Νοημοσύνης**, η διασαφήνιση όλων των συστατικών μερών του μοντέλου ΤΝ όπως για παράδειγμα διαδικασίες, αλληλεπιδράσεις, τεχνικά χαρακτηριστικά, τρόπος δημιουργίας και τρόπος χρήσης του μοντέλου.
5. **Το θέμα που πραγματοποιεί το σύστημα ΤΝ**, όπως πρόβλεψη αναγνώριση, βελτιστοποίηση και αποτελέσματα ή έξοδοι, όπως για παράδειγμα όραση υπολογιστών.

Επιπρόσθετα αναφέρεται ο κύκλος ζωής του συστήματος Τεχνητής Νοημοσύνης, σαν συμπληρωματική δομή για την κατανόηση των τεχνικών χαρακτηριστικών του συστήματος.

[(OECD 2022)]



Εικόνα4: Εφαρμογή του συστήματος ΤΝ στο πεδίο και στο εργαστήριο, Πηγή:[(OECD 2022)]



Επιπλέον πληροφορίες για το πλαίσιο ταξινόμησης φαίνονται στο παράρτημα.

## 2.5 Αξιοπιστία της Τεχνητής Νοημοσύνης

Τα μοντέλα Τεχνητής Νοημοσύνης χρησιμοποιούνται σε ένα μεγάλο αριθμό από κρίσιμους τομείς όπως χρηματοοικονομική αξιολόγηση ρίσκου και ιατρική διάγνωση και προϋποθέτει ένα πλαίσιο αξιοπιστίας καθώς μπορεί να προβούν σε μεροληπτικές αποφάσεις.

Η μεροληψία (bias) των μοντέλων αποτελεί κεντρικό θέμα στην ΤΝ που αποτελεί την απόκλιση των μοντέλων σε εξαγωγή πραγματιστικών συμπερασμάτων. Ανιχνεύεται σε όλο τον κύκλο της υλοποίησης των μοντέλων και περιλαμβάνει τις εξής κατηγορίες μεροληψίας: λόγω μετρήσεων, λόγω εξαίρεσης σημαντικών μεταβλητών στο προς εκπαίδευση μοντέλο, αναπαράστασης, συγκέντρωσης δεδομένων, εφαρμογής αλγορίθμου, χρήσης συστήματος ΤΝ, δημοτικότητας, αλλαγής των δεδομένων παρατήρησης, εκτίμησης μοντέλου. [(MEHRABI et al 2021)]

Παράλληλα, η ελλιπής πληροφόρηση σε κάποιες τεχνικές μάθησης, ως προς το πώς το μοντέλο δημιουργήθηκε παρά την καλή εφαρμογή του δεν παρέχει αιτιολόγηση του αποτελέσματος, ώστε να εξηγή συγκεκριμένες αποφάσεις και να είναι σίγουρη η εφαρμογή του, μια έννοια που στην ΤΝ αναφέρεται ως “Μαύρο Κουτί ή Black Box-AI”. Ένα σύνολο από παράγοντες του εξωτερικού περιβάλλοντος του μοντέλου δημιουργεί κινδύνους ως προς την συμβατότητα με κανονισμούς, ως προς την πρόθεση χρήσης, ως προς την συνολική του ακρίβεια/accuracy, ως προς την δικαιοσύνη του μοντέλου, ως προς την προέλευση του συνόλου δεδομένων, ως προς την διαφύλαξη της ιδιωτικότητας και των προσωπικών δεδομένων.

Η Κυβερνητικότητα /Governance στην ΤΝ αφορά την ανάγκη για τον καθορισμό πώς ένα μοντέλο ή μια υπηρεσία ΤΝ έχει δημιουργηθεί, την συμβατότητα του με την περίσταση, την αξιοπιστία του και την αποδοτικότητα του. Η κυβερνητικότητα με κατάλληλες μεθοδολογίες μπορεί να αποτρέψει την χρήση μη κατάλληλων συνόλων δεδομένων, την μεροληψία των μοντέλων, ή μοντέλα που να έχουν απροσδόκητες μεταβολές στην εφαρμογή τους.

.Ο ΟΟΣΑ έχει θεσπίσει ένα σύνολο από αρχές για την προώθηση της Αξιοπιστής Τεχνητής Νοημοσύνης οι οποίες είναι:

1. Ωφέλεια στους ανθρώπους και στον πλανήτη, πρόοδος, βιώσιμη ανάπτυξη ευημερία.
2. Ανθρωποκεντρικές αξίες, σεβασμός στις δημοκρατικές αξίες, στα ανθρώπινα δικαιώματα στην προάσπιση της ιδιωτικότητας και της δικαιοσύνης
3. Διαφάνεια και επεξηγηματικότητα, των μοντέλων που δημιουργούνται
4. Ευρωστία, ασφάλεια,

5. Υπευθυνότητα και λογοδοσία ως προς την λειτουργία τους από τους υπεύθυνους των εφαρμογών περιλαμβάνοντας και τις παραπάνω αρχές στην υλοποίησή τους.

[(OECD 2022b)]

Η ανάγκη για την δημιουργία αξιόπιστης TN οδηγεί στην καθιέρωση πρακτικών σε τεχνικό, διαδικαστικό και εκπαιδευτικό επίπεδο .[(OECD 2021)]

Σε ένα διαφορετικό πλαίσιο η Ευρωπαϊκή Ένωση θέτει κατευθύνσεις για την επίτευξη Αξιόπιστης Τεχνητής Νοημοσύνης, κατά τον κύκλο ζωής του συστήματος:

1. Θα πρέπει να είναι νόμιμη συμβαδίζοντας με τους εφαρμοζόμενους νόμους και κανονισμούς.
2. Θα πρέπει να είναι ηθική, εξασφαλίζοντας προσκόλληση σε ηθικές αξίες και αρχές
3. Θα πρέπει να είναι εύρωστη, από τεχνική και κοινωνική άποψη, ώστε να μην προκαλεί φθορά.

[(EE 2019)]

Άλλα θέματα που σχετίζονται με την αξιοπιστία της TN είναι η **Ασφάλεια Δημιουργίας Μοντέλων**, αποτρέποντας κακόβουλες ενέργειες που επηρεάζουν τον τρόπο δημιουργίας του μοντέλου.

Η ιδιωτικότητα των δεδομένων και η ύπαρξη ευαίσθητων προσωπικών δεδομένων σ' αυτά δημιουργεί προβλήματα ως προς τον τρόπο με τον οποίο οι ιδιοκτήτες δεδομένων τα μοιράζονται Ένας τρόπος για να διαφυλαχτούν τα εμπόδια της ιδιωτικότητας είναι το **Federated Learning**, που αποτελεί έναν τρόπο εκπαίδευσης των μοντέλων όπου οι χρήστες του μοντέλου μπορούν να παρέχουν τα δεδομένα τους με ασφάλεια ώστε να συνεκπαιδευόσουν ένα μοντέλο χωρίς να έχει πρόσβαση κάποιος σε αυτά (όπως για παράδειγμα να παραχωρήσουν ιατρικές γνωματεύσεις ή προσωπικά δεδομένα).[(IBM 2022c)]

Εταιρείες όπως η IBM παρέχει οδηγίες σχετικά με το πως θα μπορέσει να επιτευχθεί η αξιοπιστία της TN δημιουργώντας ένα Δελτίο Ενημέρωσης (Factsheet) δίνοντας πληροφόρηση για τα βασικά στοιχεία του μοντέλου που έχει σχηματιστεί, αυξάνοντας την διαφάνεια και την εμπιστοσύνη. [(IBM 2022b)]

Μια άλλη μέθοδος για την χρήση της TN με αύξηση της αξιοπιστίας αποτελεί την χρήση βασικών Μοντέλων / Foundation Models, η οποία αφορά προ εκπαιδευμένα μοντέλα τα οποία οι εταιρείες μπορούν να τα χρησιμοποιούν από ένα αποθετήριο μοντέλων αντί να δημιουργούνε ένα μοντέλο από την αρχή και να το προσαρμόζουν σε εξειδικευμένα θέματα. Με την χρήση αυτοεποπτευόμενης μάθησης και μεταφορά γνώσης το μοντέλο μπορεί να μάθει πληροφορία από την μία περίπτωση στην άλλη. [(IBM 2022a)]

## 2.6 Παράγοντες καθιέρωσης Τεχνητής Νοημοσύνης

Η ΤΝ μεταμορφώνει το επιχειρηματικό περιβάλλον εισάγοντας ένα νέο επιχειρηματικό μοντέλο ως προς την εταιρεία και ως προς τους πελάτες. Η έρευνα σε πάνω από 500 μελέτες περιπτώσεων απέδειξε ότι “η ΤΝ καλύπτει ολόκληρη την αλυσίδα αξίας μιας επιχείρησης, βελτιώνει τις επιχειρηματικές διαδικασίες κάνοντάς τις πιο αποδοτικές και λιγότερο κοστοβόρες και βελτιώνει την αντιδραστικότητα του οργανισμού (reactivity)” [(Milana & Ashta,2021)] “Ένα σύνολο από δυνατότητες της ΤΝ ενσωματώνονται σε λειτουργίες στις επιχειρήσεις. Όπως φαίνεται και στην εικόνα.

### AI CAPABILITIES EMBEDDED in STANDARD BUSINESS PROCESSES, 2021

Source: McKinsey & Company, 2021 | Chart: 2022 AI Index Report

Industry	Computer Vision	Deep Learning	Facial Recognition	Knowledge Graphs	NL Generation	NL Speech Understanding	NL Text Understanding	Physical Robotics	Recommender Systems	Reinforcement Learning	Robotic Process Automation	Simulations	Transfer Learning	Virtual Agents
All Industries	23%	19%	11%	17%	12%	14%	24%	12%	17%	16%	26%	17%	12%	23%
Automotive and Assembly	15%	14%	9%	16%	3%	11%	12%	24%	12%	5%	33%	27%	6%	12%
Business, Legal, and Professional Services	29%	24%	15%	20%	23%	18%	19%	13%	22%	27%	31%	18%	21%	19%
Consumer Goods/Retail	23%	12%	14%	17%	11%	13%	14%	4%	8%	8%	16%	9%	1%	15%
Financial Services	17%	16%	11%	16%	12%	18%	32%	4%	13%	16%	33%	12%	12%	28%
Healthcare Systems/Pharma and Medical Products	30%	25%	12%	19%	10%	8%	26%	28%	22%	13%	28%	22%	19%	31%
High Tech/Telecom	28%	22%	6%	17%	17%	18%	34%	5%	19%	15%	23%	14%	11%	25%

% of Respondents (AI Capability)

Εικόνα 5: Δεξιότητες ΤΝ σε επιχειρήσεις Πηγή [(OECD.AI 2022)]

Η ΤΝ εισάγει τον αυτοματισμό στην εργασία και στις υπηρεσίες δίνοντας την ικανότητα στις εταιρείες να βελτιώσουν τις επιχειρηματικές τους λειτουργίες, το οποίο συνεπάγεται αφ' ενός την εξοικονόμηση πόρων και αφ' ετέρου την αποδέσμευση ανθρωπίνου δυναμικού. Η επίδραση αυτή στο εργατικό δυναμικό έρχεται σε παράλληλη ανάδραση ως ανάγκη της επιχείρησης είτε για να εκπαιδέψει το προσωπικό σε μεθόδους ΤΝ είτε να προσλάβει νέους εξειδικευμένους εργαζόμενους. Η εξοικονόμηση κόστους εργαζομένων σε καθιερωμένες εργασίες που τώρα αυτοματοποιήθηκαν αντιδιαστέλλεται με το κόστος σε επενδύσεις ΤΝ και απόκτησης κατάλληλης υλικοτεχνικής υποδομής. Παράλληλα οι επιχειρήσεις μπορούν να αποδεσμεύσουν τους εργαζόμενους από εργασίες, μονότονες ή δύσκολες δεσμεύοντας τους σε εργασίες που δίνουν αξία στην επιχείρηση ή που μπορούν να αναπτύξουν την επιχείρηση. [(OECD 2021)]

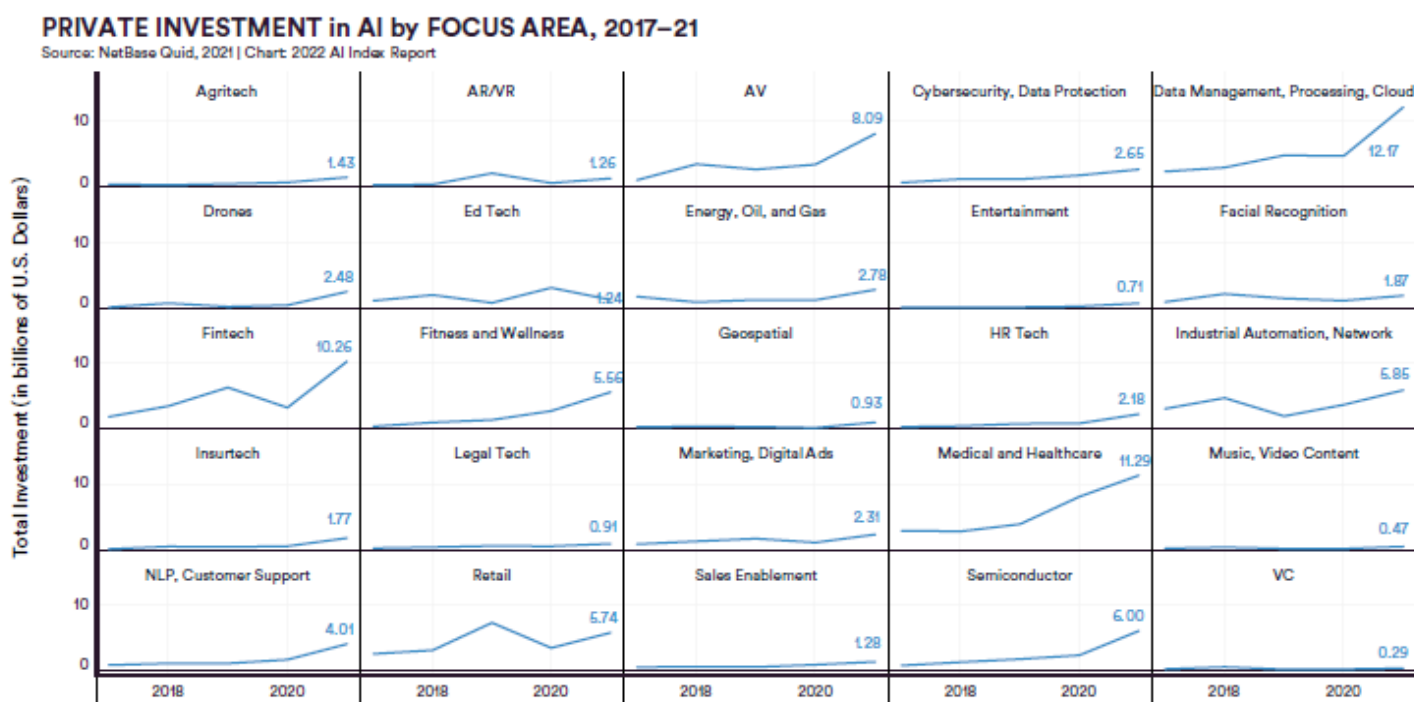
Ως προς τις επιχειρήσεις, το νέο επιχειρηματικό μοντέλο υιοθετεί την πρακτική να αξιοποιούνται τα αποθηκευμένα δεδομένα για να εξάγεται γνώση από αυτά, να αναζητούνται άλλα δημόσια διαθέσιμα δεδομένα ή να γίνονται ενέργειες ώστε να αποκτηθούν νέα δεδομένα που θα βοηθήσουν,

από την επεξεργασία τους, να ληφθούν καλύτερες αποφάσεις. Με την χρήση της TN είναι δυνατόν να εξαχθούν συμπεράσματα με διαφορετικό τρόπο από τις συνηθισμένες μεθόδους ανάλυσης όπως για παράδειγμα για τον υπολογισμό της αξίας μια επιχείρησης, ή της χρηματοοικονομικής της κατάστασης. Αυτό οδηγεί σε καλύτερη κατανόηση των πελατών και των αναγκών τους και μεγαλύτερο μερίδιο αγοράς, την αύξηση των κερδών, την προσθήκη αξίας στην επιχείρηση την διαφοροποίηση από τον ανταγωνισμό με την δημιουργία ανταγωνιστικού πλεονεκτήματος.

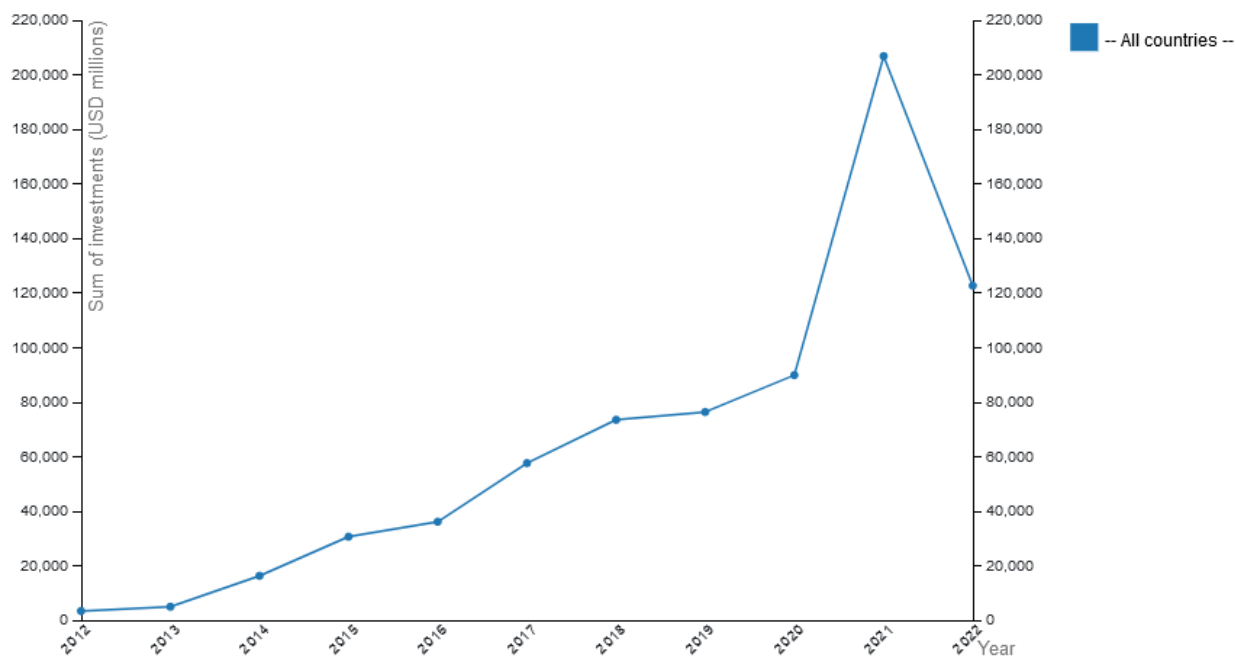
Ως προς τον πελάτη οδηγεί στην αυτοματοποίηση των εργασιών, την προσωποποίηση στις ανάγκες του, την απλοποίηση των διαδικασιών, την μείωση του κόστους, ταχύτερες υπηρεσίες και την υποβοήθηση των αποφάσεων του. Πλέον μια εργασία μετατρέπεται σε μια εφαρμογή mobile, που αντιστοιχεί σε μια εταιρεία που λειτουργεί την εφαρμογή και που τροποποιείται ανάλογα με τις πληροφορίες που έχει για τον χρήστη.

## 2.7 Επενδύσεις στην Τεχνητή Νοημοσύνη

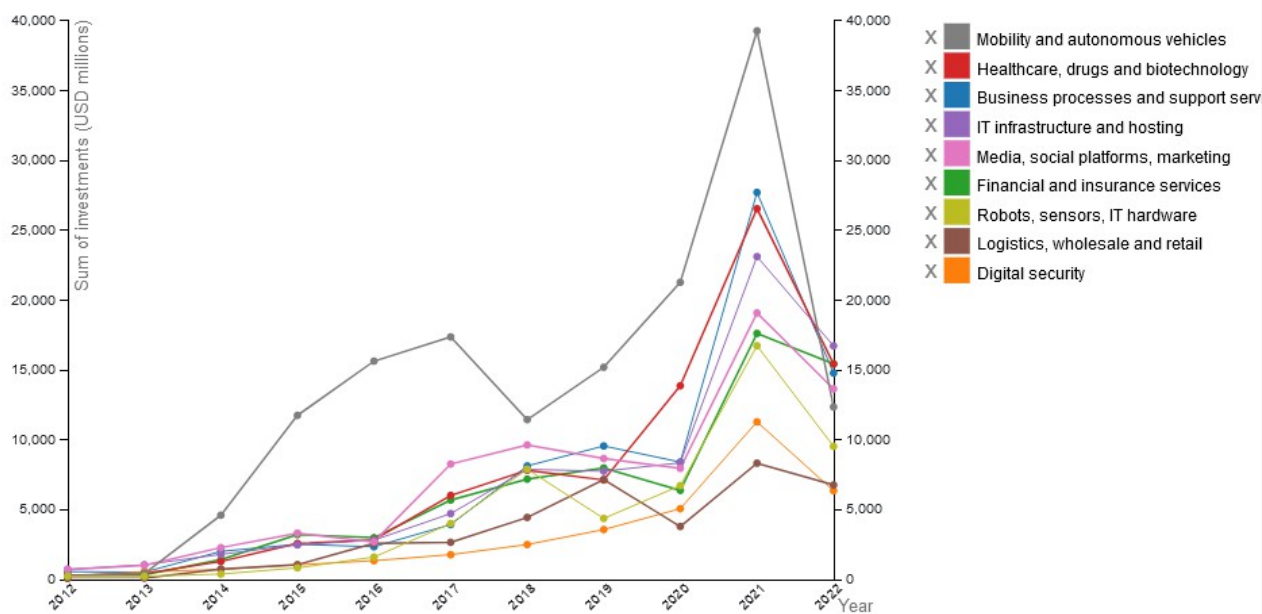
Από τις επενδύσεις που γίνονται στον τομέα της TN στις επιχειρήσεις τόσο επιχειρηματικών αγγέλων (venture capital) όσο και επενδυτών μετοχών (equity investors) διαφαίνεται η ανάπτυξη του τομέα της Τεχνητή Νοημοσύνης.



Εικόνα 6: Επενδύσεις στην TN Πηγή [(OECD.AI 2022)]



Εικόνα 7: Παγκόσμιες επενδύσεις Venture Capital [(OECD.AI 2022)]



Εικόνα 8: Επενδύσεις ΤΝ ανά τομέα δραστηριότητας [(OECD.AI 2022)]

## 2.8 Η Τεχνητή Νοημοσύνη στον Τομέα των Χρηματοοικονομικών

Στον χρηματοοικονομικό τομέα η ΤΝ δημιουργεί μεγάλη ώθηση για καινοτομία και επενδύσεις. Το παγκόσμιο χρηματοπιστωτικό σύστημα μετασχηματίζεται ως προς την λειτουργία του από την επίδραση των νέων τεχνολογιών. Οι τομείς στους οποίους γίνεται χρήση ΤΝ είναι οι εξής:

- Πρόβλεψη τιμών μεγεθών Χρηματοοικονομικών Χρονοσειρών αξιών
- Πρόβλεψη Χρηματοοικονομικών μεγεθών μιας εταιρείας
- Διαδικασία Αποφάσεων/Decision Making, σχετικά με επενδύσεις
- Πρόβλεψη Χρεοκοπίας/Bankruptcy, πρόβλεψη πτώχευσης βάσει δεδομένων λογιστικών καταστάσεων
- Credit Scoring/Χρηματοπιστωτική αξιοπιστία, εκτίμηση της πιστοληπτικής ικανότητας αιτούμενων δανείων, πρόβλεψη της αθέτησης της εκπλήρωσης της υποχρέωσης εξόφλησης δανείου
- Λογιστική και Ελεγκτική/Accounting and Auditing, σχετίζεται με την ανάπτυξη έμπειρων συστημάτων/η μελέτης ετήσιων χρηματοοικονομικών καταστάσεων για την εύρεση μη κανονικοτήτων στις επιχειρήσεις, συστήματα αποφάσεων λογιστικής
- Φοροδιαφυγή/Tax fraud, από την μελέτη των χαρακτηριστικών των εταιρειών που διαφεύγουν μπορεί να γίνει πρόβλεψη.
- Χρηματοοικονομική Διοίκηση / Financial Management, με θέματα όπως
- Τραπεζική, η TN μπορεί να είναι ευνοϊκή για την μείωση των απωλειών δανεισμού, αυξημένη ασφάλεια σε πληρωμές συναλλαγών, στόχευση πελατών/customer targeting, διαχείριση διαθεσίμων και συνδεδεμένες υπηρεσίες.
- Αυτοματοποιημένα Συστήματα Εμπορίας Αξιών (Automated Trading Systems), τα οποία περιλαμβάνουν στη αγορά πώληση μετοχών, ανίχνευση σημάτων από τις αγορές, robot-advisory systems που αποφασίζουν τί θα αγοραστεί χωρίς ανθρώπινη παρέμβαση [Sezer B. et al, 2019]
- High Frequency Trading (HFT), εμπορία αξιών με πολύ μεγάλη συχνότητα συναλλαγών
- Chat bots, αυτόματη συνομιλία προς παροχή βοήθειας
- Κυβερνοασφάλεια / Cybersecurity
- Εκτίμηση χρηματοοικονομικού κινδύνου, σε προϊόντα σταθερού εισοδήματος που διαπραγματεύονται στο χρηματιστήριο ή γενικότερα σε επενδύσεις
- Η μοντελοποίηση της χρηματοοικονομικής δομής μιας επιχείρησης
- Εκτίμηση τιμής για τους στόχους εξαγορών και συγχωνεύσεων
- Διαπραγμάτευση αξιογράφων
- Διαχείριση χαρτοφυλακίου/ Portfolio management
- Προσδιορισμός τιμής εισαγωγής στην χρηματιστηριακή αγορά
- Επιλογή μετοχών, πρόβλεψη απόδοσης μελλοντικών μετοχών
- Πρόβλεψη τιμής μετοχής για αρχική εισαγωγή στο χρηματιστήριο

- Θεμελιώδης Ανάλυση
- Διαχείριση χρηματοοικονομικών κινδύνων
- Έλεγχος χρηματοοικονομικών καταστάσεων
- Πιστωτικός κίνδυνος
- Αξιολόγηση κινδύνου απάτης (Fraud Detection)
- Πρόβλεψη μελλοντικών χρηματοροών
- Έλεγχος των πληροφοριών και τεχνολογιών επικοινωνίας (ICT) στον χρηματοοικονομικό τομέα για τον έλεγχο παράνομων δοσοληψιών  
[Ζαπράνης Α.]

## 2.9 Πρόβλεψη Μεγεθών Χρηματοοικονομικών Χρονοσειρών

Η πρόβλεψη τιμών χρηματοοικονομικών μεγεθών χρονοσειρών αποτελεί αντικείμενο ανάλυσης και έρευνας της Ακαδημίας, αλλά και των επιχειρήσεων καθώς συνδέεται με την πρόβλεψη μεγεθών που οι αποφάσεις βάσει αυτών θα έχουν επιπτώσεις πιθανού κέρδους ή απώλειας. Η συνολική κεφαλαιοποίηση των εταιρειών ή η συνολική αξία των μετοχών των δημόσιων εγγεγραμμένων εταιρειών στις Ηνωμένες Πολιτείες για το 2020 είναι 93,69trillion US\$.

Οι χρηματοοικονομικές αγορές εμφανίζουν, δυναμική, μη γραμμική, χαοτική συμπεριφορά, μη στατικότητα, ακολουθιακούς συσχετισμούς ενώ επιρρεάζονται απο πολλές άλλες παραμέτρους. Αποτελούν έναν χώρο συνεχών συναλλαγών και μεγάλου όγκου διακινήσιμων κεφαλαίων.

Οι παράγοντες που επιδρούν στην διαμόρφωση των χρηματοοικονομικών αξιών μπορούν να είναι μακροοικονομικοί όπως η γενική οικονομική κατάσταση, η τιμή του ΑΕΠ, οι τιμές των επιτοκίων, ο πληθωρισμός ή εσωτερικά χρηματοοικονομικά στοιχεία μιας εταιρείας όπως προκύπτουν από τις χρηματοοικονομικές καταστάσεις μιας εταιρείας, όπως οι χρηματοροές. Επιπλέον παράγοντες είναι η πολιτική κατάσταση, η ψυχολογία των επενδυτών, οι οικονομικές ειδήσεις ή οι δημοσιεύσεις σε μέσα κοινωνικής δικτύωσης ή ακόμα και οι φήμες. Η επίδραση αυτών των παραγόντων στην τιμή και την τάση των χρηματοοικονομικών αξιών προσπαθεί να μοντελοποιηθεί στην χρηματοοικονομική πρόβλεψη.

Η δυναμική των χρηματοοικονομικών αγορών αυξήθηκε με την πρόσβαση του απλού χρήστη στις χρηματαγορές, όπως αυτών σε mobile και pc πλατφόρμες αγορών συναλλαγών αλλά και την έκδοση κρυπτονομισμάτων με την τεχνολογία Blockchain.

Παράλληλα με τον όρο του επενδυτή αναπτύχθηκε για τις ηλεκτρονικές κυρίως αγορές περισσότερο ο όρος του trader αυτού που εκτελεί τις ενέργειες αγοράς και λήψης μετοχών. Μια

κατηγορία επενδυτών δημιουργεί επενδυτικές κινήσεις βάσει της πεποίθησης ότι η κίνηση της αγοράς μπορεί να προβλεφθεί και ότι μπορούν να οφεληθούν από τις επιλογές πρόβλεψης βραχυπρόθεσμα, ενώ μια κατηγορία επενδυτών επενδύει στις μετοχές μιας εταιρείας προσδοκώντας την μακροπρόθεσμη οφέλεια της από την παραγωγικότητα και κερδοφορία μιας επιχείρησης. Παράλληλα οι επενδυτές είναι και θεσμικοί όπως κυβερνήσεις, τράπεζες, ιδρύματα και hedge funds.

## 2.10 Θεωρίες για την πρόβλεψη χρηματαγορών

Διαχρονικά διατυπώθηκαν διάφορες θεωρίες σχετικά με την πρόβλεψη των χρηματαγορών:

Οι Malkiel and Fama(1970) διατύπωσαν την Υπόθεση Αποτελεσματικής Αγοράς / Efficient Market Hypothesis (EMH) σύμφωνα με την οποία στις χρηματοοικονομικές αγορές, οι τιμές των μετοχών απορροφούν όλη την γνωστή γνώση σε δεδομένη στιγμή. Η τιμή μιας μετοχής μια δεδομένη χρονική στιγμή αντανακλά όλη την διαθέσιμη πληροφορία για αυτήν και συνεπώς δεν μπορεί να προβλεφθεί.

Υπάρχουν τρεις μορφές EMH βασιζόμενες στον βαθμό της απόδοσης των μετοχών της αγοράς:

**1) Weak Form**, η οποία θεωρεί ότι η αγορά αντανακλά όλη την προηγούμενη πληροφορία της αγοράς. Η υπόθεση υποθέτει ότι οι παλιές συχνότητες αποδόσεων δεν έχουν καμία επίδραση στις μελλοντικές τιμές.

**2) Semi strong Form**, θεωρεί ότι η αγορά αντανακλά αποδοτικά όλη την δημόσια πληροφορία. Η τιμή της μετοχής προσαρμόζεται γρήγορα για να απορροφήσει την νέα πληροφορία. Η μορφή αυτή περιλαμβάνει την ασθενή μορφή.

**3) Strong Form**, θεωρεί ότι η αγορά αντανακλά όλη την διαθέσιμη πληροφορία τόσο ιδιωτική όσο και δημόσια. Η θεωρία αυτή ότι κανένας επενδυτής δεν θα μπορούσε να επιτύχει θετικές αποδόσεις ακόμα και αν είχε διαθέσιμη δημόσια πληροφορία. Η μορφή αυτή καλύπτει τις δύο προηγούμενες μορφές

Η **θεωρία του Τυχαίου Περιπάτου/Random Walk**, αναφέρει ότι οι μελλοντικές τιμές των μετοχών δεν ακολουθούν κάποια τάση ή πρότυπο αλλά είναι αυθόρμητες εκδηλώσεις από τις προηγούμενες τιμές και συνεπώς δύσκολο να προβλεφθούν.

Το μοντέλο **Martingale Paul Samuelson**, δηλώνει ότι δεδομένης όλης της διαθέσιμης πληροφορίας οι τωρινές τιμές αποτελούν την καλύτερη πρόβλεψη για μελλοντικές τιμές και ο μοναδικός συντελεστής πρόβλεψης είναι η τωρινή τιμή των μετοχών.

Η **Fractal market Hypothesis(FMH)**, σε αντίθεση με την EMH η οποία θεωρεί ότι οι αγορές βρίσκονται σε κατάσταση δυναμικής ισορροπίας-equilibrium, είναι σταθερές και συνυπολογίζουν την προσφορά και ζήτηση των επενδυτών. [(Chopra & Sharma 2022)]



## 2.11 Μέθοδοι Πρόβλεψης

Το θέμα της πρόβλεψης Χρηματοοικονομικών Μεγεθών περιλαμβάνει τομείς όπως η πρόβλεψη τιμής, κατεύθυνση τιμής, πρόβλεψη μεταβλητότητας αγοράς. Με τα αποτελέσματα της πρόβλεψης διαμορφώνονται οι στρατηγικές εμπορίας χρηματοοικονομικών εργαλείων και περαιτέρω οι αυτοματοποιημένες στρατηγικές εμπορίας. Δοθείσας της πρόβλεψης τιμής ή κατεύθυνσης τιμής μέσω της δημιουργίας μοντέλων μηχανικής μάθησης, μπορεί να δημιουργηθεί μια στρατηγική long-short η οποία να εκτελεί την εμπορία μετοχών (trading).

Με την πάροδο του χρόνου διαμορφώθηκαν εργαλεία ανάλυσης και πρόβλεψης χρονοσειρών όπως στατιστικές τεχνικές, οικονομετρικές τεχνικές, η τεχνική ανάλυση και η θεμελίωσης ανάλυση, η οποίες έδωσαν την θέση τους σε μεθόδους βελτιστοποίησης, Γενετικούς Αλγόριθμους, μεθόδους Μηχανικής Μάθησης όπως οι Support Vector Machines, και τα Νευρωνικά Δίκτυα. Η πρόσφατη εξέλιξη των νευρωνικών δικτύων βαθιάς μάθησης (Deep Learning) αποτελεί μια σημαντική αναβάθμιση των μεθόδων πρόβλεψης που δίνει την δυνατότητα διαχείρισης μεγάλου όγκου δεδομένων, με αυτοματοποιημένο τρόπο, σε σύντομο χρονικό διάστημα. Νέες προσεγγίσεις αποτελούν η πρόβλεψη συναισθήματος (sentiment analysis) ενώ ως προς την κατανόηση των αποτελεσμάτων της TN.

Η πρόσβαση σε δεδομένα των παγκόσμιων αγορών δίνει την δυνατότητα συστηματικής μελέτης στις χρονοσειρές με μεθόδους Τεχνητής Νοημοσύνης. Το θέμα της ανάπτυξης μοντέλων χρηματοοικονομικής πρόβλεψης χρησιμοποιεί παρελθοντικές ιστορικές τιμές μεγεθών για να προβλέψει μελλοντικές τιμές όπως μετοχές, ισοτιμίες και ξένο συνάλλαγμα (FOREX), αγαθά, πολύτιμα μέταλλα δεδομένων και άλλους παράγοντες, επίσης μπορεί να χρησιμοποιήσει τεχνικούς δείκτες, δείκτες θεμελιώδους ανάλυσης και την παράμετρο του συναισθήματος (sentiment) των επενδυτών. Χρησιμοποιεί κυρίως εποπτευόμενη μάθηση και είναι επαγωγικό στα δεδομένα βοηθώντας να βελτιωθεί η πρόβλεψη με τον χρόνο.

Ωστόσο η χρήση της TN ακόμα και αν τα μοντέλα που δημιουργούνται είναι αποδοτικά δεν συνεπάγεται πάντα κερδοφορία από την εφαρμογή τους. Το μοντέλο που δημιουργείται μπορεί να κάνει αποτελεσματική πρόβλεψη βάσει των ήδη υπαρχόντων δεδομένων, στο χρονικό πλαίσιο αναφοράς ενώ δεν υπολογίζει νέους αστάθμητους εξωτερικούς παράγοντες όπως για παράδειγμα μι φυσική καταστροφή ή μια πανδημία. Εντούτοις τα ιστορικά δεδομένα που θα προκύψουν από τέτοια απροσδόκητα γεγονότα θα μπορέσουν αν βελτιώσουν την ακρίβεια μελλοντικών μοντέλων.

## 2.12 Αυτοματοποιημένη Πρόβλεψη

Το μέλλον της χρηματοοικονομική πρόβλεψης μεταμορφώνεται από την τεχνολογία έτσι ώστε να δημιουργηθεί ένα περιβάλλον στο οποίο οι άνθρωποι θα εργάζονται σε ένα περιβάλλον που θα

τροφοδοτείται από δεδομένα, με αλγορίθμους που θα αξιοποιούν τα δεδομένα για να ανακαλύψουν γνώση. Η συνεργασία αυτή ανθρώπων και αλγορίθμων, με αποφάσεις που βασίζονται στα διαθέσιμα δεδομένα αποτελεί και το νόημα της 5ης βιομηχανικής επανάστασης, που θα δώσει την δυνατότητα στους και να δράσουν με περισσότερη σιγουριά, με στρατηγικό τρόπο αποφάσεων σε πραγματικό χρόνο για την υποστήριξη των καθημερινών αποφάσεων. [(Deloitte 2022)]

Κεντρικό σημείο στην θεώρηση αυτή αποτελεί η συστηματική, αυτοματοποιημένη συγκέντρωση δεδομένων από πολλές πηγές και με διάφορους τύπους δεδομένων ώστε να υποστηρίξουν την πολύπλευρη ανάλυση των δεδομένων όπως μέσω μηχανικής μάθησης, επεξεργασίας φυσικής γλώσσας, και εξελιγμένων τεχνικών ανάλυσης. [(Deloitte 2022)] Τα δεδομένα τίθενται προς επεξεργασία για τον σχηματισμό μοντέλων και στρατηγικών εμπορίας χωρίς την ανθρώπινη παρέμβαση ή με την ελάχιστη ανθρώπινη παρέμβαση.

### **2.12.1 Αλγοριθμικό Εμπόριο / Algorithmic Trading**

Αποτελεί την χρήση ηλεκτρονικών συστημάτων για την εκτέλεση μιας στρατηγικής εμπορίας στις χρηματοοικονομικές ηλεκτρονικές αγορές με μακροπρόθεσμους στόχους χρησιμοποιώντας αυτόματες και προγραμματισμένες οδηγίες εμπορίας, αλγορίθμους εκτέλεσης/ execution algorithms με κριτήρια την τιμή, τον χρονισμό και τον όγκο των συναλλαγών. Μέσω αυτού είναι δυνατόν ένας επενδυτής να κάνει αποφάσεις εμπορίας και να διενεργεί όλο το εμπόριο των μετοχών. Η προσέγγιση του αυτοματοποιημένου εμπορίου μετοχών μπορεί να γίνει με χρήση TN, MM ή BM.

Πλεονεκτήματά του είναι η αυτόματη επαγωγή στην κατάσταση της αγοράς, η γρηγορότερη εκτέλεση εντολών, η κερδοφορία από τις μικρές κινήσεις στην τιμή εμπορευόμενων μεγεθών, η αύξηση της ταχύτητας, τα μικρότερα κόστη συναλλαγών. Μειονεκτήματά του είναι η απώλεια ρευστότητας στις αγορές [(Chen 2022)][(Velvetech 2022)]

Οι κυριότερες αλγοριθμικές στρατηγικές εμπορίας είναι Ακολουθία Τάσης/Trend-following , Μόχλευση παρέκκλισης τιμών/Arbitrage leveraging, Επανεξισσορόπηση κεφαλαίων δεικτών/ Index fund rebalancing, Εμπορία βάση Μαθηματικών μοντέλων/Mathematical model-based , Χρονικά Σταθμισμένο μέση τιμής/ Volume-weighted average pricing, Χρονικά-βαρυμένη σταθμισμένη μέση τιμή/Time-weighted average pricing, Μέση Αναστροφή/ Mean reversion, Ποσοστό Όγκου Πωλήσεων/ Percentage of volume, Ελλείψεις Υλοποίησης / Implementation shortfall. [(Velvetech 2022)]

### **2.12.2 Υψηλής Συχνότητας Εμπορία/ High Frequency Trading**

Αποτελεί την εκτέλεση μιας στρατηγικής εμπορίας σε πολύ μεγάλη συχνότητα συναλλαγών στην μονάδα του χρόνου. Για την πραγμάτωσή του χρειάζεται πλήρη μή καθυστερούμενα δεδομένα,

Είναι χρήσιμο για τεχνικές arbitrage και οι χρήστες μπορούν να οφεληθούν από μικρά κέρδη σε κάθε εμπορία. Σε έρευνα για την πρόβλεψη διαφόρων μοντέλων μηχανική μάθησης για υψηλής συχνότητας εμπορία τιμών χαρτονομισμάτων Bitcoin υποδείχτηκε η καλύτερη πρόβλεψη χρηματοοικονομικών μεγεθών. [(Rundo 2022)]

### **2.13 Επεξηγήσιμη Τεχνητή Νοημοσύνη (ETN)**

Η ανάπτυξη της TN και ειδικότερα με την εξέλιξη της BM δίνει την δυνατότητα στην προσέγγιση μεγάλου όγκου δεδομένων υψηλών διαστάσεων από τα οποία εξάγονται πολύπλοκα γνωρίσματα με εξαιρετική εφαρμογή. Ωστόσο ο τρόπος δημιουργίας των μοντέλων δεν είναι ευκρινής με αποτέλεσμα την ελλιπή αιτιολόγηση των αποτελεσμάτων στους χρήστες και παράλληλα την δημιουργία αβεβαιότητας.

Παράλληλα το θέμα προστασίας δεδομένων ιδιαίτερα σε όσα αναφέρονται σε προσωπικά δεδομένα θέτει το ερώτημα με ποιο τρόπο η TN μπορεί να καταστεί ηθικά εφαρμόσιμη. Για την περαιτέρω υιοθέτηση των μοντέλων BM, και την αύξηση της εμπιστοσύνης τους η Επεξηγησιμότητα είναι ιδιαίτερα σημαντική, ιδιαίτερα σε κρίσιμες εφαρμογές.

Ως Επεξηγήσιμη Τεχνητή Νοημοσύνη (ETN) / Explainable AI (XAI) είναι εκείνη στην οποία οι άνθρωποι μπορούν να κατανοήσουν τις αποφάσεις ή τις προβλέψεις που δημιουργήθηκαν από αυτή. Η θεώρηση αυτή έρχεται σε αντίθεση με την θεώρηση της TN ως ένα “μαύρο κουτί/blackbox” όπου ακόμα και οι σχεδιαστές της δεν μπορούν να εξηγήσουν πως κατέληξε σ αυτή την συγκεκριμένη επιλογή. Κατά μια εκδοχή η επεξηγησιμότητα ακολουθεί τις αρχές της διαφάνειας, διερμηνείας, επεξήγησης. [(Wikipedia e, 2022)] Οι λόγοι χρήσης της είναι για την επικύρωση του συστήματος TN, την διάγνωση ατελειών στο μοντέλο και μεροληψία στα δεδομένα, για την επικύρωση των προβλέψεων, για την βελτίωση περαιτέρω του συστήματος και για την μάθηση εκ νέου του χρήστη από το σύστημα. [(Samek et al, 2017)]

Ως προς την προστασία των Πολιτών, η Ευρωπαϊκή Ένωση, θεωρεί ως βασικά για την ορθή υλοποίηση της TN στην κοινωνία, αρχικά την διαφάνεια των μοντέλων, η οποία αφορά τον τρόπο υλοποίησης, τις ειδικές οδηγίες (specifications), την διερμηνεία ως προς την γνώση των μηχανισμών υλοποίησης. Ως δεύτερο σημείο την αξιοπιστία των μοντέλων, που αφορά την καλή εφαρμογή, και την γνώση των τρωτών σημείων των μοντέλων, τις παραλείψεις. Τρίτον, την προστασία δεδομένων στα μοντέλα, την εξασφάλιση των σωστών δεδομένων, την απουσία λανθασμένων τιμών ή θορύβου, την κατακεκομμένη (distributed) μάθηση και την μάθηση βάση κοινών παραμέτρων (federated learning) και την μάθηση σε κρυπτογραφημένα δεδομένα. [(Ronan et al, 2022)]

Μια άλλη προσέγγιση προσδιορίζει τις τέσσερις αρχές της επεξηγήσιμης TN ως: 1)επεξήγηση, με στοιχεία αιτιολόγησης, 2)νοηματοδότηση ως προς την κατανόηση του χρήστη, 3)ακρίβεια πρόβλεψης με αιτιολόγηση δημιουργίας του αποτελέσματος εξόδου, και 4) γνώση των ορίων γνώσης.[(Ronan et al, 2022)]

Η ETN αναμένεται να προκαλέσει την δημιουργία μιας νέας γενιάς συστημάτων TN,με ύπαρξη πιο επεξηγήσιμων μοντέλων, με ικανότητα αιτιολόγησης, διατηρώντας το επίπεδο μάθησης και την ακρίβεια πρόβλεψης, ώστε οι χρήστες να έχουν την δυνατότητα να κατανοούν, να εμπιστεύονται και να διαχειρίζονται αιτιολογικά τις εφαρμογές που χρησιμοποιούν.

Στον τομέα της πρόβλεψης χρηματοοικονομικών μεγεθών με χρήση νευρωνικών δικτύων και επεξηγήσιμη TN ήδη υπάρχουν μελέτες που να ποσοτικοποιούν και να οπτικοποιούν τα αποτελέσματα προς κατανόηση του χρήστη. Ωστόσο μια μέθοδος που θεωρείται ότι αφορά την επεξηγησιμότητα του μοντέλου είναι η εφαρμογή σε παλιότερα πραγματικά δεδομένα (backtesting) [(Olorunnimbe et al,2022)].

### 3 Ανασκόπηση Μεθόδων Πρόβλεψης

Η ανάλυση χρηματοοικονομικών χρονοσειρών βρίσκεται στην τομή πολλών επιστημονικών τομέων όπως είναι η επιστήμη των υπολογιστών, η επιστήμη δεδομένων, η μηχανική μάθηση, η στατιστική και τα χρηματοοικονομικά. Σε αντίθεση με την υπόθεση της αποτελεσματικής αγοράς η οποία αναφέρει ότι η αγορά δεν μπορεί να προβλεφθεί, γίνονται προσπάθειες με διάφορες μεθόδους ώστε να ανιχνευθούν οι τάσεις της αγοράς. Από τις κλασσικές μεθόδους ανάλυσης όπως οι στατιστικές μέθοδοι, η θεμελιώδης ανάλυση και η τεχνική ανάλυση οι οποίες χρειάζονται την παρέμβαση ανθρώπων στους υπολογισμούς, γίνεται μετάβαση στην Τεχνητή Νοημοσύνη όπου οι αλγόριθμοι αυτοματοποιούν την διαδικασία της ανάλυσης και πρόβλεψης αλλά και της εμπορίας/trading. Η χρήση της Τεχνητής Νοημοσύνης και αναλυτικών μεγεθών για πρόβλεψη χρηματοοικονομικών μεγεθών δίνει μοντέλα με μεγάλη ακρίβεια πρόβλεψης και αναγνώριση πολύπλοκων, ακολουθιακών συσχετίσεων στα δεδομένα παρατήρησης. Οι νέες μέθοδοι ανάλυσης περιλαμβάνουν μοντέλα Μηχανικής Μάθησης όπως Τεχνητών Νευρωνικών Δικτύων ή η νεότερη μορφή τους των Νευρωνικών Δικτύων Βαθιάς Μάθησης, όπου οι παραδοσιακές μέθοδοι αντικαθίστανται ή χρησιμοποιούνται παράλληλα με τις νέες μεθόδους.

#### 3.1 Θεμελιώδης Ανάλυση / Fundamental Analysis

Η Θεμελιώδης ανάλυση μετράει την εσωτερική αξία/intrinsic value ενός χρηματοοικονομικού εργαλείου που είναι η αξία μιας επένδυσης με βάση την χρηματοοικονομική κατάσταση της εκδότριας εταιρείας και τις τρέχουσες συνθήκες της αγοράς και της οικονομίας [(Segal 2022)].

Η Θεμελιώδης Ανάλυση αποτελεί μέθοδο με την οποία εκτιμάται η τιμή μιας μετοχής ή ενός χρηματοοικονομικού προϊόντος ώστε να υπολογιστεί αν είναι υπερτιμημένη η υποτιμημένη σε σχέση με την τιμή της στην αγορά. Γίνεται με χρήση δημόσιων διαθέσιμων χρηματοοικονομικών δεδομένων όπως οι χρηματοοικονομικές εκθέσεις της εταιρείας, (Ισολογισμός, Κατάσταση Αποτελεσμάτων Χρήσης, Ανάλυση Ταμειακών Ροών, Εκθέσεις ελεγκτών), η κερδοφορία, η ρευστότητα, τα περιουσιακά στοιχεία, η μελλοντική ανάπτυξη, η απόδοση ανά μερίδιο, το περιθώριο κέρδους, χρηματοοικονομικούς δείκτες ή άλλες ποιοτικές πληροφορίες όπως η διοίκηση της εταιρείας, το επιχειρηματικό μοντέλο, το “brand name”, οι ευρεσιτεχνίες, οι αποδόσεις των μετοχών. Επίσης εξετάζει την γενική κατάσταση της οικονομίας, την ζήτηση στην αγορά, τη δυναμική του συγκεκριμένου κλάδου και συγκρίνει με άλλες εταιρείες με τα ίδια χαρακτηριστικά.

Η θεμελιώδης ανάλυση κάνει την υπόθεση ότι η τιμή της παρούσας αξίας μιας μετοχής συχνά δεν αντανακλά την τιμή μιας εταιρείας όταν συγκρίνεται με τα δημόσια δεδομένα. Η άλλη υπόθεση που γίνεται είναι ότι η τιμή που καθορίζεται από τα θεμελιώδη στοιχεία μιας εταιρείας είναι πιθανό να

βρίσκεται πιο κοντά στην πραγματική τιμή της εταιρείας. Η τρίτη υπόθεση είναι ότι μακροπρόθεσμα η αγορά θα αναδείξει την πραγματική τιμή. [(Segal 2022)]

Τα πιο σημαντικά μεγέθη θεμελιώδους ανάλυσης είναι οι αναλογίες/ratios που δίνουν αποτέλεσμα σύγκρισης μεταξύ δύο ή τριών σχετιζόμενων μετοχών. Οι αναλογίες και οι μετρικές που δημιουργούνται αποτελούν επίσης εργαλεία σύγκρισης της απόδοσης της εταιρείας με άλλες εταιρείες του ίδιου τομέα δραστηριότητας. [(Segal T., 2022)]

Χαρακτηριστικά παραδείγματα αναλογιών είναι η συχνότητα ανάπτυξης, η ανάπτυξη στις πωλήσεις, η ανάπτυξη στα κέρδη ανά μετοχή, η αναλογία τιμής προς κέρδη.

### **Αναλογία Τιμής προς Κέρδη / Price to Earnings Ration (P/E)**

Αποτελεί την αναλογία της τωρινής Τιμής (Price) αγοράς της μετοχής προς τα Ετήσια Κέρδη Ανα Μετοχή (Earnings per Share). Η αναλογία (P/E) δίνει την δυνατότητα σύγκρισης μεταξύ πολλαπλών μετοχών. Αποτελεί αντανάκλαση της επίκαιρης δημοτικότητας μιας μετοχής.

### **Αναλογία Τιμής προς Κέρδη ανάπτυξης/ Price to Earnings Growth Ratio (PEG)**

Αποτελεί εξέλιξη της P/E και προκύπτει από την διαίρεση της με την συχνότητα ανάπτυξης των κερδών  $PEG = (PRICE/EPS) / (EPS Growth)$

Όσο μικρότερη είναι η αναλογία τόσο περισσότερο η μετοχή μπορεί να είναι υποτιμημένη δεδομένης της απόδοσης των κερδών της. Μια τιμή ίση με 1 δείχνει μια σωστά τιμημένη μετοχή, κάτω από 1 θεωρείται πιθανώς υποτιμημένη και πάνω από 1 θεωρείται υπερτιμημένη.

Τα πρακτικά πλεονεκτήματα της Θεμελιώδους Ανάλυσης είναι ότι βασίζεται σε σημερινά δεδομένα των επιχειρήσεων, ορθολογικοποιεί την εκτίμηση για μια εταιρεία, είναι σημαντική για την αξιολόγηση της πορείας μιας επιχείρησης και τον προσδιορισμό της τιμής μιας αξίας. Επιπλέον βασίζεται στην χρηματοοικονομική απόδοση, στους χρηματοοικονομικούς πόρους, την Διοίκηση, τον ανταγωνισμό, την αξιολόγηση των επιχειρήσεων.

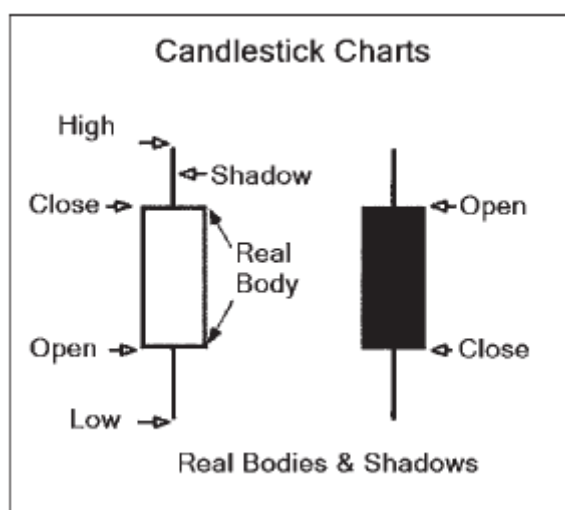
Τα μειονεκτήματα της είναι ότι τα χρηματοοικονομικά δεδομένα των επιχειρήσεων στα οποία στηρίζεται η Θεμελιώδης Ανάλυση, μπορεί να είναι ελλιπή ή ψευδή, η ερμηνεία τους μπορεί να είναι λανθασμένη, τα δεδομένα που εξετάζονται ανήκουν στο παρελθόν και δεν αντιπροσωπεύουν το μέλλον, δεν εξετάζεται την εικόνα που έχει η αγορά για μια εταιρεία. Επιπλέον περιορίζει το εύρος εξέτασης των επενδυτών με το να εξετάσουν τα στοιχεία περισσότερων εταιρειών.

## **3.2 Τεχνική Ανάλυση / Technical Analysis**

Η Τεχνική Ανάλυση (TA) αποτελεί μέθοδο ανάλυσης και πρόβλεψης τιμών χρηματοοικονομικών προϊόντων που βασίζεται στην μελέτη των ιστορικών τιμών των αγορών όπως είναι η τιμή/price, οι τάσεις τιμών/trend, και ο όγκος συναλλαγών/volume για τον υπολογισμό βραχυπρόθεσμων

μελλοντικών τιμών. Η Τεχνική Ανάλυση δεν εξετάζει τα χρηματοοικονομικά δεδομένα που εξετάζει η θεμελιώδης ανάλυση ή τους παράγοντες της οικονομίας. Μια άλλη πτυχή της Τεχνικής Ανάλυσης αφορά την αναγνώριση και τον χαρακτηρισμό οπτικών προτύπων διαγραμμάτων τιμών χρονοσειρών, κρίσιμων σημείων, επίπεδων τιμών όπου αναγνωρίζονται περιοχές πιθανής υποστήριξης ή αντίστασης των τιμών, χρονικά πλαίσια στα οποία οι τιμές ακολουθούν μια συγκεκριμένη τάση. Βάσει των τάσεων των τιμών και της τιμής δράσης δημιουργεί δείκτες, και σχεδιαστικά γραμμές στα διαθέσιμα χρηματοοικονομικά δεδομένα. Οι δείκτες αυτοί έχουν ονόματα τα οποία ομοιάζουν στα σχήματά τους και αποτελούν πρότυπα για τον προσδιορισμό της εξέλιξης της αγοράς. Τέτοιοι δείκτες είναι:

- Δείκτες τάσης/ Price trends, επιχειρούν να περιγράψουν μαθηματικά την τάση για μια χρονική περίοδο.
- Διαγράμματα Προτύπων / Chart patterns, που αποτελούν οπτικά πρότυπα στην εξέλιξη των τιμών με συγκεκριμένη μορφή που προσδιορίζει μια συμπεριφορά στην αγορά. Αυτά είναι οι αναλογίες Fibonacci, πρότυπα Κεφαλής και Ώμων ( Head and Shoulders), Κανάλια (Channels), Πολλαπλά Tops/Bottoms και Triangles /Wedges, Japanese Candlestick Patterns

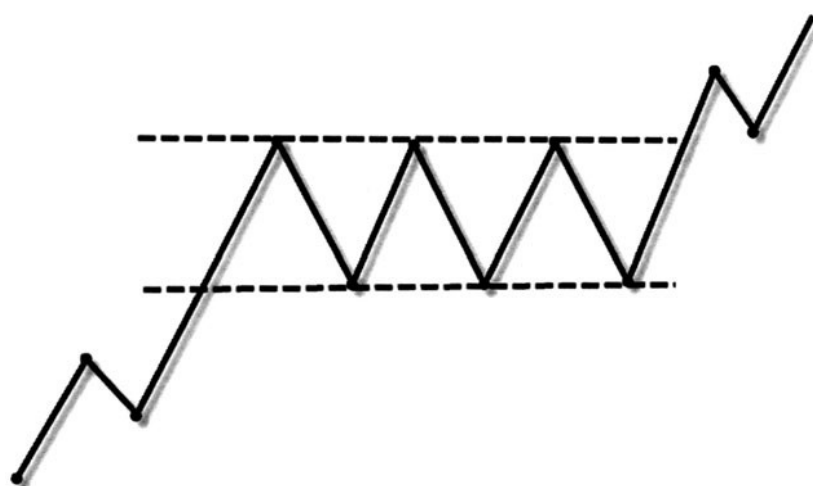


Εικόνα9: Candlestick Charts Πηγή: [(Brooks 2006)]

- Δείκτες Ποσοτήτων (Volume), όπως On-Balance Volume (OBV), Volume RSI (Relative Strength Index)
- Δείκτες Momentum δείκτες, που χρησιμοποιούνται για τον προσδιορισμό της συχνότητας επιτάχυνσης της τιμής μιας μετοχής και δείχνουν την δύναμη ή την αδυναμία της τιμής μιας μετοχής:
  - MACD (Moving Average Convergence/Divergence), αποτελεί έναν δείκτη momentum ο οποίος δείχνει την σχέση μεταξύ δύο εκθετικών Exponential Moving Average(EMA) μιας τιμής ενός χρηματοοικονομικού εργαλείου.

-Relative Strength Index (RSI), αποτελεί δείκτη ο οποίος μετράει τις αλλαγές στην τιμή ενός περιουσιακού στοιχείου σε 14 περιόδους, διαιρώντας το μέσο κέρδος με την μέση απώλεια που είχε σε εκείνο το διάστημα, που δίνει την δυνατότητα να εντοπιστούν υπερβολικές αγορές (overbought) ή υπερβολικές πωλήσεις (oversold). Η κλίμακα αξιολόγησης είναι 0-100 με χαρακτηριστικές τιμές 30-70.

- Ταλαντωτές/ Oscillators, δημιουργούν χαμηλές και υψηλές ζώνες μεταξύ δύο ακραίων τιμών και έπειτα δημιουργεί έναν δείκτη τάσης που ταλαντώνεται μεταξύ αυτών των ορίων. Χρησιμοποιείται για βραχυπρόθεσμες συνθήκες υπεραγοράς και υπερπώλησης.
- Κινούμενοι μέσοι όροι / Moving averages, που υποδεικνύουν την μέση τιμή ενός εμπορεύσιμου εργαλείου για μια συγκεκριμένη χρονική περίοδο και χρησιμοποιούνται για την ανίχνευση τάσεων.
- Η υποστήριξη / support ή επίπεδο υποστήριξης και η αντίσταση/resistance ή breakout ( η υψηλότερη προσδοκώμενη τιμή) η οποία αποτελεί κίνηση τιμής έξω από το καθορισμένο εύρος για να υποδείξουν πώς οι traders βλέπουν τις επενδύσεις και να υποδείξει τί πρόκειται να συμβεί. Παραδείγματα είναι το συμμετρικό τρίγωνο ή το wedge. [(Hayes 2022)]



Εικόνα10: Υποστήριξη & Αντίσταση Πηγή: [(Brooks 2006)]

Πλεονέκτημα της Τεχνικής Ανάλυσης είναι ότι οι Τεχνικοί δείκτες είναι άμεσοι, εύκολα κατανοήσιμοι με δυνατότητα έγκαιρου εντοπισμού των τάσεων των τιμών. Ωστόσο οι δείκτες περικλείουν υποκειμενικότητα στην ερμηνεία τους.

Σε κάποιες από τις μελέτες που εξετάστηκαν χρησιμοποιούνται οι δείκτες της τεχνικής ανάλυσης RSI (14 ημέρες, 70-30) και SMA(50 ημέρες) για τον σχεδιασμό στρατηγικών. [(Jiang 2022)]

Η ΤΑ στηρίζεται στην πεποίθηση ότι οι πρόσφατες τάσεις μπορούν να υποδείξουν τις μελλοντικές τιμές. Το πλαίσιο της Τεχνικής Ανάλυσης στηρίζεται στις παρακάτω υποθέσεις:

- Η τιμή του εξεταζόμενου χρηματοοικονομικού εργαλείου κινείται σε τάσεις.



- Οι αγορές είναι αποτελεσματικές με τιμές να αναπαριστούν τους παράγοντες που επηρεάζουν την τιμή ενός εργαλείου(security)
  - Ακόμα και οι τυχαίες τιμές κίνησης της αγοράς εμφανίζονται να κινούνται με αναγνωρίσιμα πρότυπα και τάσεις που τείνουν να επαναλαμβάνονται με τον χρόνο
  - Η τιμή της αγοράς αντανακλά όλη την διαθέσιμη πληροφορία για την αγορά (Efficient Market Hypothesis)
- [(Hayes 2022)]

Μειονεκτήματα της Τεχνικής Ανάλυσης είναι ότι οι δείκτες έχουν υποκειμενικότητα στην ερμηνεία τους. Επιπλέον η αναγνώριση των προτύπων στις μεταβολές των δεδομένων τίθεται στην ικανότητα του παρατηρητή. Η αυτοματοποίηση των μεθόδων αναγνώρισης προτύπων στις χρονοσειρές έχει αποτελέσει αντικείμενο έρευνας και εξετάζεται παρακάτω.

### **3.3 Στατιστικές τεχνικές Πρόβλεψης**

Οι στατιστικές μέθοδοι αποτελούν μέθοδο ανάλυσης χρονοσειρών και πρόβλεψης των τιμών. Οι πιο γνωστές από αυτές είναι η απλή παλινδρόμηση η οποία αποτελεί μια γραμμική μέθοδο που καθορίζει την σχέση μεταξύ μιας εξαρτημένης μεταβλητής  $Y$  ως προς μία ανεξάρτητη μεταβλητή  $X$ , ή την σχέση με περισσότερες ανεξάρτητες μεταβλητές  $X_i$  που ονομάζεται πολλαπλή παλινδρόμηση. Η μέθοδος λειτουργεί δημιουργώντας μια γραμμή χρησιμοποιώντας ως κριτήριο την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος των πραγματικών μεταβλητών από τις μεταβλητές που ορίζει η γραμμή της παλινδρόμησης. Οι υποθέσεις που γίνονται είναι ότι οι σχέσεις μεταξύ των μεταβλητών είναι γραμμικές, όλες οι μεταβλητές είναι ανεξάρτητες και ανήκουν σε κανονική κατανομή. Άλλες παραλλαγές από στατιστικά μοντέλα που χρησιμοποιούνται είναι η Λογιστική Παλινδρόμηση / Logistic regression, που αποτελεί ένα θέμα παλινδρόμησης με την διαφορά ότι οι μεταβλητές είναι δυαδικές κατηγορίες του τύπου “1-0”, “True-False”, “Yes-No” και το πρόβλημα που ορίζεται είναι αυτό της δυαδικής ταξινόμησης όπως είναι το σύστημα ανίχνευσης spam.

### **3.4 Αυτοπαλινδρόμηση / Auto Regressive (AR)**

Το μοντέλο της Αυτοπαλινδρόμησης / Autoregression (AR), δείχνει ότι η μεταβαλλόμενη μεταβλητή ή μεταβλητή εξόδου, μεταβάλλεται βάσει των δικών της προηγούμενων-καθυστερούμενων τιμών, η οποία και εκφράζεται βάση του γραμμικού συνδυασμού των τιμών της και ενός στοχαστικού όρου.

Η τιμή για κάθε χρονοσειρά μπορεί να υπολογιστεί ως γραμμικός συνδυασμός των  $p$  παλιότερων παρατηρήσεων  $AR(p)$  μαζί με ένα σφάλμα  $\eta$  και μια σταθερά  $c$ . Το  $p$  αναφέρεται και ως τάξη του μοντέλου.

### 3.5 Μοντέλο Κινούμενων Μέσων Όρων / Moving Average Model (MA)

Αποτελεί μια μέθοδο μοντελοποίησης χρονοσειρών μιας μεταβλητής ο οποίος στην απλούστερη μορφή αποτελεί τον αριθμητικό μέσο όρο σε μια συγκεκριμένη χρονική περίοδο (παράθυρο δεδομένων) (Απλός Μέσος Όρος). Χρησιμοποιείται για την ανίχνευση των τάσεων στα δεδομένα και για να οριστούν τα επίπεδα υποστήριξης και αντίστασης. Το μέγεθος των τιμών  $q$  του παραθύρου του κινούμενου μέσου αποτελεί παράμετρο του μοντέλου  $MA(q)$ . Ο **Εκθετικός Κινούμενος Όρος (EMA)** υπολογίζει περισσότερο βάρος στις πρόσφατες τιμές ώστε να ενσωματώνει την νεότερη πληροφορία.

### 3.6 Auto-Regressive Moving Average (ARMA)

Αποτελεί μια μέθοδο μοντελοποίησης χρονοσειρών που ενσωματώνει δύο πολυωνυμικούς όρους έναν για την **Αυτοπαλινδρόμηση (AR)** και έναν για τον **Κινούμενο Μέσο Όρο/Moving Average (MA)** και αναφέρεται σε στατικές χρονοσειρές δεδομένων δηλαδή αν οι στατιστικές τους ιδιότητες δεν εξαρτώνται με τον χρόνο.

### 3.7 Auto-Regressive Integrated Moving Average (ARIMA)

Το μοντέλο ARIMA αποτελεί ένα μοντέλο πρόβλεψης και ανάλυσης χρονοσειρών που βασίζεται στην παλινδρόμηση/regression και ενσωματώνει τα στοιχεία των Κινούμενων Μέσων Όρων (MA) και της Αυτοπαλινδρόμησης (AR) ενώ παράλληλα ο όρος Ολοκληρωμένο/Integrated(I), αναπαριστά την διαφορά μεταξύ των τιμών δεδομένων έτσι ώστε η χρονοσειρά να γίνει στάσιμη. Η παράμετρος  $d$  του  $I(d)$  αποτελεί τον αριθμό που η χρονοσειρά έχει διαφοροποιηθεί ή τον βαθμό των διαφορών μέχρι η χρονοσειρά να θεωρείται στάσιμη. Με αυτό τον τρόπο χρησιμοποιείται για μη στάσιμα δεδομένα.

Άλλες στατιστικές μέθοδοι που χρησιμοποιούνται στην στατιστική ανάλυση είναι η Generalized Autoregressive Conditional Heteroskedastic (GARCH), η Linear Discriminant Analysis (LDA), η Quadratic Discriminant Analysis (QDA).

Οι μέθοδοι αυτές οι οποίες χρησιμοποιούνται και στην Οικονομετρία, χρησιμοποιούν κάποια στατιστική δειγματοληψία για την δημιουργία των μοντέλων τους, η οποία όμως δεν μπορεί να καλύψει την δυναμική φύση των χρηματοοικονομικών χρονοσειρών. Παρόλο που κάποια απο τα μοντέλα μπορούν να λάβουν υπόψη τα ακολουθιακά στοιχεία των χρονοσειρών, η επεξεργασία τους χρειάζεται πολύπλοκη ανάλυση μέσω οικονομετρικών τεστ[(Tang et al, 2022)]

### 3.8 Ασαφής Λογική/Fuzzy Logic

Στηρίζεται στην λογική ότι οι αποφάσεις των ανθρώπων γίνονται υπό αβεβαιότητα, βασίζονται σε μη ακριβή και μη αριθμητική πληροφορία. Οι πιθανότητες των διαφορετικών ενδεχομένων δεν είναι γνωστές. Βασίζονται στην δημιουργία κανόνων από τους οποίους εξάγουν τα συμπεράσματα, με κανόνες if-then ανακριβή βάρη για την αναπαράσταση μεταξύ των μεταβλητών. Τα σύνολα Fuzzy ορίζονται μέσω των membership functions ( $\mu$ ) που χαρτογραφούν το σύνολο των στοιχείων στο διάστημα  $[0,1]$ .

- Παρέχει γρήγορα μεταφραζόμενα αποτελέσματα
- Χρησιμοποιείται σε ατελή και ανακριβή πληροφορία
- Ικανοποιητική ακρίβεια πρόβλεψης
- Μπορεί να αναγνωρίσει και να κατηγοριοποιήσει δυναμικά τεχνικούς κανόνες με εφαρμογή στις αποφάσεις κινδύνου αποφασέων σχετικών με διαχείριση και κατανομή κεφαλαίου

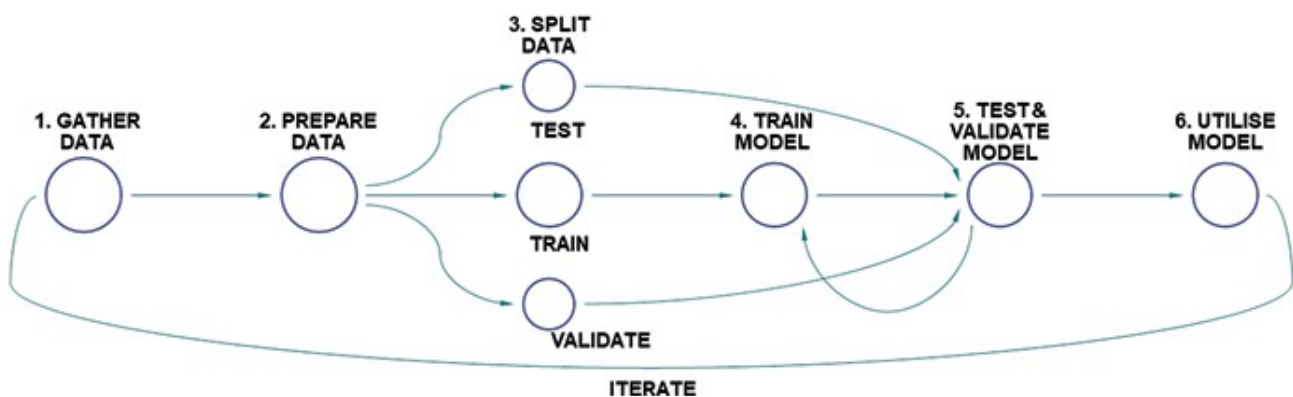
### 3.9 Μηχανική Μάθηση (MM) / Machine Learning(ML)

Η Μηχανική Μάθηση (MM) αποτελεί πεδίο της Τεχνητής Νοημοσύνης και ο όρος αναφέρεται στην αυτόματη προσπέλαση διαθέσιμων δεδομένων μέσω αλγορίθμων προσεγγίζοντας τα μέσω της δημιουργίας ενός μοντέλου με μια διαδικασία η οποία καλείται μάθηση. Η MM σε αντιδιαστολή με τις στατιστικές μεθόδους δεν λειτουργεί με την δημιουργία υποθέσεων προς επιβεβαίωση αλλά ανιχνεύει πρότυπα στα δεδομένα. Δεν κάνει στατιστική δειγματοληψία όπως οι στατιστικές μέθοδοι, ούτε έλεγχο υπόθεσης με χρήση αβεβαιότητας ούτε υποθέτει κάποια κατανομή στα δεδομένα, αλλά επεξεργάζεται τα δεδομένα υπο μορφή δικτύων ή γράφων, και χρησιμοποιεί πιθανότητες και μετρικές για την αξιολόγηση των καλύτερων μοντέλων. Τα διαθέσιμα δεδομένα χωρίζονται σε δεδομένα εκπαίδευσης, δεδομένα ελέγχου και δεδομένα επικύρωσης. Η μάθηση αυτή μπορεί να συμβαίνει σε πραγματικό χρόνο (on-line) σε μια διαδικασία όπου ελέγχονται όλα τα δεδομένα που δημιουργούνται είτε να συμβαίνει εκτός πραγματικού χρόνου (off-line).

Η MM χρησιμοποιείται για την δημιουργία μοντέλων πρόβλεψης χρηματοοικονομικών χρονοσειρών. Η διαδικασία της μάθησης από τα διαθέσιμα δεδομένα περιλαμβάνει τα παρακάτω βήματα:

- 1) **Συγκέντρωση Δεδομένων (Gather Data)**, που προορίζονται για το προς εξέταση θέμα πρόβλεψης. Η συγκέντρωση των δεδομένων αποτελεί σημαντική παράμετρο για την δημιουργία των μοντέλων, ενώ προϋποθέτει την οργανωτική δομή μιας επιχείρησης ώστε να συγκεντρώνει αυτόματα τα δεδομένα που χρειάζεται να χρησιμοποιήσει. Στην περίπτωση των χρονοσειρών η επιλογή του συνόλου δεδομένων θα πρέπει να είναι τέτοια ώστε να καλύπτει ώστε το χρονικό πλαίσιο μακρά χρονική διάρκεια.

- 2) **Προεργασία των Δεδομένων/ Data Preprocessing**, αφορά την διαμόρφωση των δεδομένων ώστε να επεξεργαστούν με τον πιο πρακτικό τρόπο όπως η επιλογή του κατάλληλου συνόλου δεδομένων η συμπλήρωση των ελλিপών τιμών, η κανονικοποίηση.
- 3) **Διαμερισμός των δεδομένων**, αφορά τον χωρισμό των δεδομένων σε σύνολο δεδομένων εκπαίδευσης (training set), σύνολο δεδομένων δοκιμής (test set) και σύνολο επικύρωσης δεδομένων ( validation set)
- 4) **Εκπαίδευση του μοντέλου (Train the model)**, με την εφαρμογή του αλγορίθμου επιλογής στα διαθέσιμα δεδομένα ώστε να γίνουν οι υπολογισμοί και να εξαχθούν τα πρότυπα και να δημιουργηθεί το μαθηματικό μοντέλο που περιγράφει τα δεδομένα. Σκοπός της δημιουργίας είναι να αναπτυχθεί το μοντέλο ώστε να μορφοποιεί την τιμή της μεταβλητής στόχου ή εξόδου.
- 5) **Δοκιμή και επικύρωση του μοντέλου (Test and Validate a model)**, εφαρμογή του μοντέλου που έχει δημιουργηθεί στα δύο σύνολα που έχουν εξαιρεθεί, και είναι η φάση στην οποία εξετάζεται η ακρίβεια της πρόβλεψης και εφαρμογή του μοντέλου σε νέα δεδομένα.
- 6) **Χρήση του μοντέλου (Utilize a model)**, αφορά την εισαγωγή το μοντέλου στις αναλυτικές διαδικασίες
- 7) **Επανάληψη**, η διαδικασία επαναλαμβάνεται με την συγκέντρωση νέων δεδομένων την επαναδημιουργία νέων μοντέλων



Εικόνα11: Διαδικασία Μάθησης [(Akerkar R., 2019)]

Η Μηχανική Μάθηση και η Τεχνητή Νοημοσύνη μπορούν να εξάγουν συμπεράσματα βασή μοντέλων από τα δεδομένα. Οι μέθοδοι μάθησης χωρίζονται στις εξής κατηγορίες:

- **Εποπτευόμενη μάθηση / Supervised Learning**

Αποτελεί την χρήση ενός αλγορίθμου σε ένα σύνολο δεδομένων εκπαίδευσης όπου για κάθε δεδομένο υπάρχει κάποιος κανόνας και έχει επισυναφθεί μια κατηγορία ή κλάση. Η μάθηση αξιοποιεί τα γνωρίσματα των δεδομένων και επισυνάπτει τα νέα δεδομένα τα οποία εξετάζει στις

κατηγορίες αυτές. Ως classification/ταξινόμηση αναφέρεται η διαδικασία επισύναψης των αλγορίθμων σε κατηγορίες. Παραδείγματα τέτοιων μεθόδων είναι τα Νευρωνικά Δίκτυα, η Random Forest, τα Δένδρα Αποφάσεων (Decision Trees). Ως regression/παλινδρόμηση αναφέρεται η διαδικασία όπου ο αλγόριθμος χρησιμοποιείται για την κατανόηση της σχέσης μεταξύ των ανεξάρτητων και των εξαρτημένων μεταβλητών αριθμητικών τιμών. Παραδείγματα τέτοιων μεθόδων είναι η γραμμική παλινδρόμηση, η λογιστική παλινδρόμηση, η πολυωνυμική παλινδρόμηση.

- **Μη εποπτευόμενη Μάθηση / Unsupervised learning**

Αφορά την ανίχνευση δομών ή προτύπων στα δεδομένα, στα οποία δεν υπάρχει κάποιος κανόνας ή κάποια κατηγορία. Τέτοιες μέθοδοι μάθησης είναι η Συσταδοποίηση/Clustering.

- **Εξαναγκασμένη Μάθηση / Reinforcement learning**

Αποτελεί είδος μη εποπτευόμενης μάθησης, η οποία πραγματοποιείται βάσει παραδειγμάτων, όπου οι αποφάσεις λαμβάνονται με την αλληλεπίδραση με το περιβάλλον με επικύρωση συνεπειών και η μάθηση γίνεται βάση της δοκιμής και σφάλματος (trial and error). Οι αποφάσεις λαμβάνονται ώστε να μεγιστοποιήσουν την συνολική ανταμοιβή από το περιβάλλον (όπως στην περίπτωση των ευφυών πρακτόρων).

### **3.10 Χαρακτηριστικά Δεδομένων**

Η Μηχανική Μάθηση βασίζεται στα διαθέσιμα δεδομένα τα οποία χρειάζονται να είναι αξιόπιστα ώστε να δημιουργούν βάσιμο αποτέλεσμα, καθώς θα αποτελέσουν τις μεταβλητές πάνω στις οποίες θα εφαρμοστούν οι τεχνικές των αλγορίθμων, η εκπαίδευση και η δοκιμή τους. Τα δεδομένα που αναφέρονται στις μελέτες αντλούνται από το ιστορικό των αγορών που υπάρχει διαθέσιμο σε ιστοσελίδες, τράπεζες και χρηματοοικονομικά ιδρύματα. Η φύση των δεδομένων, τα χαρακτηριστικά τους, το είδος τους, η δομή τους, οι ιδιότητες τους, ο τρόπος συλλογής τους είναι σημαντικά για τον σχηματισμό των μοντέλων. Η ποσότητα των δεδομένων όπως και ο αριθμός των μεταβλητών χρήσης επιδρά σημαντικά στον σχηματισμό των μοντέλων.

Η ονομασία Big Data αναφέρεται σε δεδομένα τα οποία είναι πολυάριθμα και πολύ δύσκολο να επεξεργαστούν με στατιστικά πακέτα επεξεργασίας.

Η χρήση μεγάλου αριθμού δεδομένων και μεγάλων διαστάσεων δημιουργεί το πρόβλημα των υψηλών διαστάσεων. Αυτό αναφέρεται ως ότι ο αλγόριθμος που εφαρμόζεται στο σύνολο δεδομένων φαίνεται να έχει καλή εφαρμογή στο σύνολο δεδομένων εκπαίδευσης αλλά δεν γενικεύει στο σύνολο δεδομένων ελέγχου. Το πρόβλημα αντιμετωπίζεται με τη επιλογή μεταβλητών από τα διαθέσιμα δεδομένα.

Ως προς την προέλευσή τους τα δεδομένα μπορεί να είναι ανθρωπογενή, να προέρχονται από τις αγορές, να συλλέγονται από αισθητήρες ή να είναι τεχνητά δηλαδή να δημιουργούνται από αλγορίθμους. Μπορεί να προέρχονται είτε από πειράματα σε εργαστήριο ή να προέρχονται από το πεδίο δραστηριότητας όπως είναι ένας χώρος παραγωγής. Τα δεδομένα μπορεί να είναι δημόσια ή ιδιωτικά ή να υπόκεινται σε προσωπικά δεδομένα.

Ως προς τους τύπους των δεδομένων μπορούν να είναι αδόμητα (κείμενο, εικόνα, ήχος) ημιαδόμητα, δομημένα, πολύπλοκα δεδομένα. Χωρίζονται περαιτέρω σε ποιοτικά και ποσοτικά.

Περαιτέρω επιχειρησιακά χρειάζεται να υπάρχει ένα τρόπος αξιόπιστης συγκέντρωσης διαφόρων τύπων δεδομένων από πολλαπλές πηγές (ενοποίηση των δεδομένων) και ένα σύστημα εύκολης διαχείρισής τους η οποία σύμφωνα με την κατεύθυνση που παίρνει η ΤΝ προορίζεται να είναι αυτόματη μέσω αισθητήρων όπως το Internet of Things, Chatbots.

Η δυναμική φύση των δεδομένων και ο ακολουθιακός χαρακτήρας τους απαιτεί διαφορετικό τρόπο προεργασίας. Στην περίπτωση των χρηματοοικονομικών μπορεί να είναι δεδομένα στην απλή μορφή όπως προκύπτουν από κάποιο αποθετήριο όπως το Yahoo Finance. Για χρηματοοικονομικά δεδομένα στους αλγορίθμους χρησιμοποιούνται είτε τα δεδομένα στην απλή τους μορφή είτε οι αποδόσεις των δεδομένων.

### 3.10.1 Προεργασία των Δεδομένων / Data Preprocessing

Η προεργασία των Δεδομένων που θα χρησιμοποιηθούν θα πρέπει να είναι αντιπροσωπευτικά των δεδομένων του πλαισίου αναφοράς. Περαιτέρω η επεξεργασία των δεδομένων περιλαμβάνει:

- **Απομάκρυνση των outliers** (εκτός ορίων) δεδομένων τα οποία συνήθως αποτελούν δεδομένα πολλαπλάσια κατά  $n$  φορές από την τυπική απόκλιση  $\sigma$  (ή από το μέσο όρο  $\mu$  (μεγαλύτερες ή μικρότερες) των υπολοίπων στοιχείων δηλαδή  $\mu - n\sigma$ , ή μεγαλύτερες από  $\mu + n\sigma$ ).
- **Διαχείριση των ελλιπών τιμών**, οι οποίες προέρχονται από σφάλματα από την εισαγωγή των δεδομένων σε διάφορους τρόπους αποθήκευσης, και για τα οποία γίνεται αντικατάστασή τους από μια τυπική τιμή όπως ο μέσος όρος ή η συχνότερη τιμή ή κάποια γειτονική τιμή.
- **Διακριτικοποίηση δεδομένων** η οποία δημιουργεί ομάδες από συσχετιζόμενες τιμές κάτω από μια μοναδική ετικέτα (single value ή bin) βάσει κάποιας κοινής κατηγοριοποίησης των γνωρισμάτων. Χρησιμοποιείται για την αφαίρεση του θορύβου από τα δεδομένα.
- **Αφαίρεση θορύβου**

- **Κανονικοποίηση (Normalization)**,σημαίνει την μεταφορά γνωρίσματος από την αρχική κλίμακα η οποία χρησιμοποιείται, σε μια κλίμακα μικρότερης εμβέλειας. Η κλίμακα αυτή είναι συνήθως είτε το διάστημα  $[-1,1]$  είτε το διάστημα  $[0,1]$ .Ο μετασχηματισμός Min/Max, μετατρέπει την τιμή μιας μεταβλητής  $x$  η οποία ανήκει στην περιοχή  $A[\min, \max]$  σε μια νέα μεταβλητή  $x'$  σε μια νέα περιοχή τιμών  $B[\min', \max']$ .
- Τοποθέτηση σε κάποια **Κλίμακα γνωρισμάτων / Feature Scaling**
- **Στατικοποίηση / Stationarity**, αποτελεί έναν μετασχηματισμό κατά τον οποίο οι στατιστικές ιδιότητες μιας χρονοσειράς δεν μεταβάλλονται με την πάροδο του χρόνου, και συνεπώς δεν αλλάζει το σχήμα της κατανομής όπως και ο μέσος όρος, η μεταβλητότητα και συνεπάγεται ότι είναι πιο εύκολα επεξεργάσιμη. Υπάρχουν πολλών ειδών στατικοποιήσεις όπως ισχυρή, ασθενής, κυκλοστατικότητα, στατικότητα τάσης
- **Διόρθωση του χρονικού πλαισίου**, γίνεται επιλογή του κατάλληλου πλαισίου αναφοράς των τιμών προς μελέτη.

### 3.10.2 Μέγεθος Δείγματος Δεδομένων

Το μέγεθος του δείγματος που επιλέγεται θα πρέπει να είναι αντιπροσωπευτικό της πρόβλεψης που επιχειρείται. Διάφοροι αλγόριθμοι μπορούν να κάνουν διαχείριση σε διαφορετικά μεγέθη δεδομένων. Χαρακτηριστικό παράδειγμα σχετικό με το μέγεθος του δείγματος είναι ότι ένα Recurrent Neural Network (RNN) μπορεί να επεξεργαστεί είτε ολόκληρη την χρονική σειρά είτε ένα μικρό μέρος αυτής με επιπτώσεις για την μάθηση του μοντέλου.[(Ge et al, 2022)]

### 3.10.3 Επιλογή Μεταβλητών

Έπειτα από την συγκέντρωση των δεδομένων ακολουθεί η επιλογή μεταβλητών που θα χρησιμοποιηθεί στην μάθηση ώστε να απομακρυνθούν τα δεδομένα που δεν σχετίζονται με το θέμα της εκπαίδευσης. Η διαδικασία αυτή οδηγεί στην απλοποίηση των μοντέλων, μικρότερους χρόνους εκπαίδευσης, μικρότερες διαστάσεις δεδομένων, βελτίωση της συμβατότητας των δεδομένων, κωδικοποίηση συμμετριών στα δεδομένα ( Wikipedia, 2022).Οι μέθοδοι με τις οποίες επιτυγχάνεται είναι:

- **Wrapper methods**, όπου ένα μοντέλο πρόβλεψης υπολογίζει υποσύνολα μεταβλητών πρόβλεψης,
- **Filter methods**, όπου χρησιμοποιείται μια μετρική για τον υπολογισμό των υποσυνόλων.
- **Embedded methods**, όπου πραγματοποιείται η διαδικασία της επιλογής μεταβλητών κατά την δημιουργία του μοντέλου.

- **Feature Selection**, επιλογή ενός υποσυνόλου απο τα αρχικά γνωρίσματα, με χρήση τεχνικών όπως Principle Component Analysis(PCA), Singular value decomposition(SVD), και Linear Discriminant Analysis (LDA).
- **Απόρριψη γνωρισμάτων/Feature Dropping**, Γνωρίσματα με μεγάλη πυκνότητα ελλিপών τιμών απορρίπτονται, όπως επίσης και όσα θεωρούνται ότι δεν έχει νόημα η χρήση τους για τον σχηματισμό του μοντέλου.

### 3.10.4 Προκατάληψης-Διακύμανσης (Bias -Variance)

Το ζήτημα με την δημιουργία των μοντέλων είναι να έχουν μικρό χρόνο δημιουργίας και καλή δυνατότητα **Γενίκευσης/ Generalization**, δηλαδή το μοντέλο που εξάγεται να μπορεί να προσεγγίσει μεγάλο αριθμό νέων δεδομένων. Το θέμα της επιλογής μεταξύ μεροληψίας/bias και διακύμανσης/variance επηρεάζει την εφαρμογή των μοντέλων.

- Η **Μεροληψία / Bias** είναι η συστηματική απόκλιση ή συστηματικό σφάλμα των προβλέψεων του μοντέλου απο τις σωστές τιμές, δείχνει κατά πόσο ο μέσος όρος των προβλέψεων διαφέρει από τις πραγματικές τιμές
- Η **Διακύμανση/ Variance** είναι η **αναμενόμενη τιμή** της τετραγωνικής απόκλισης της **τυχαίας μεταβλητής** από τη μέση τιμή και δείχνει την ευαισθησία του μοντέλου σε μικρές αλλαγές στο σύνολο εκπαίδευσης, δείχνει την ικανότητα γενίκευσης.

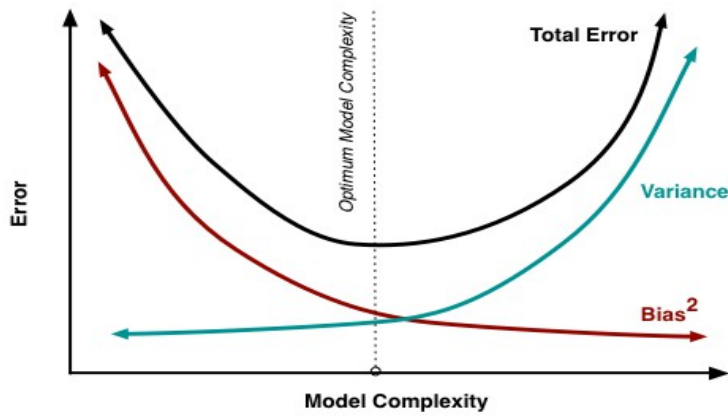
Το σφάλμα πρόβλεψης μιας μεταβλητης  $Y$  που προσπαθούμε να υπολογίσουμε είναι  $Y=f(x)+e$  όπου  $e$  όρος σφάλματος τότε το προσδοκόμενο σφάλμα είναι:

$$Err(x) = Bias^2 + Variance + Irreducible Error$$

Όπου το **Irreducible Error/ Αμείωτο σφάλμα** αφορά το σφάλμα που δεν μπορεί περαιτέρω να μειωθεί όσο καλό και να είναι το μοντέλο.

Η υψηλή μεροληψία οδηγεί σε μοντέλα που δεν εφαρμόζουν καλά στα δεδομένα (underfit) ενώ η υψηλή διακύμανση οδηγεί σε μοντέλο με χαμηλό σφάλμα που υπερεφαρμόζει (overfit). Εάν το μοντέλο είναι απλουστευμένο δεν θα μπορεί να βρει ικανοποιητικές λύσεις μεταξύ μεταβλητών εισόδου και μεταβλητών εξόδου. Αν το μοντέλο είναι πολύπλοκο τότε έχει καλύτερη εφαρμογή στα δεδομένα αλλά δεν θα προσεγγίζει ικανοποιητικά τα νεότερα δεδομένα. Επίσης πραγματοποιεί περισσότερους υπολογισμούς για την δημιουργία του. Το ιδανικό είναι ένα μοντέλο που δεν είναι πολύπλοκο, είναι γρήγορο στην εκπαίδευση και είναι αρκετό πολύπλοκο ώστε να ανιχνεύει όλες τις σχέσεις μεταξύ των εισόδων και εξόδων σε νέα δεδομένα. Η δημιουργία απλών μοντέλων συνεπάγεται μη ακολουθία των δεδομένων ενώ ένα πολύπλοκο μοντέλο συνεπάγεται στην υπερεφαρμογή στα δεδομένα και μεγάλη απόκλιση.





Εικόνα12: Σχέση μεταξύ σφάλματος Μεροληψίας και Διακύμανσης Πηγή [(Scott Fortmann-Roe, 2022)]

## 3.11 Εκπαίδευση του μοντέλου

### 3.11.1 Διαμερισμός των δεδομένων

Το θέμα αυτό σχετίζεται και με την κατάλληλη επιλογή της διαμέρισης των δεδομένων για εκπαίδευση και επικύρωση.

**Το σύνολο δεδομένων εκπαίδευσης (Training set)**, χρησιμοποιείται κατά την εκπαίδευση του μοντέλου για την προσαρμογή των παραμέτρων (βαρών). Περιλαμβάνει τις ανεξάρτητες και εξαρτημένες μεταβλητές σε κάθε επαναληπτικό πέρασμα στα δεδομένα (epoch). Δίνεται στο νευρωνικό δίκτυο επανειλημμένα, καθώς το μοντέλο μαθαίνει τα χαρακτηριστικά του από τα δεδομένα. Το σύνολο αυτό πρέπει να είναι αντιπροσωπευτικό όλων των καταστάσεων που συναντώνται στα δεδομένα. Οι μέθοδοι που επιλέγεται ένα σύνολο δεδομένων μέσω τυχαίας δειγματοληψίας η δειγματοληψίας σε στρώματα.

**Το σύνολο επικύρωσης δεδομένων (Validation set)**, είναι διαφορετικό από το σύνολο δεδομένων εκπαίδευσης και ελέγχου. Χρησιμοποιείται για την ρύθμιση των υπερπαραμέτρων του μοντέλου από την φάση εκπαίδευσης.

Η τεχνική **Cross Validation**, αποτελεί μέθοδο επικύρωσης του μοντέλου που χρησιμοποιεί ως σύνολο επικύρωσης διαφορετικά υποσύνολα από το σύνολο εκπαίδευσης για να εκπαιδεύσει το μοντέλο και να δοκιμάσει το μοντέλο σε διαφορετικές επαναλήψεις. Η cross validation k συνόλων χωρίζει τυχαία το σύνολο δεδομένων εκπαίδευσης σε k σύνολα από τα οποία τα k-1 χρησιμοποιούνται για εκπαίδευση και το 1 για validation. Η διαδικασία επαναλαμβάνεται k φορές με τα επόμενα σύνολα δεδομένων. Στην συνέχεια ως τελική μετρική για το μοντέλο υπολογίζεται ο Μέσος Όρος των μετρικών ακρίβειας που έχουν εξαχθεί. Κάποιες από τις μεθόδους cross validation είναι:

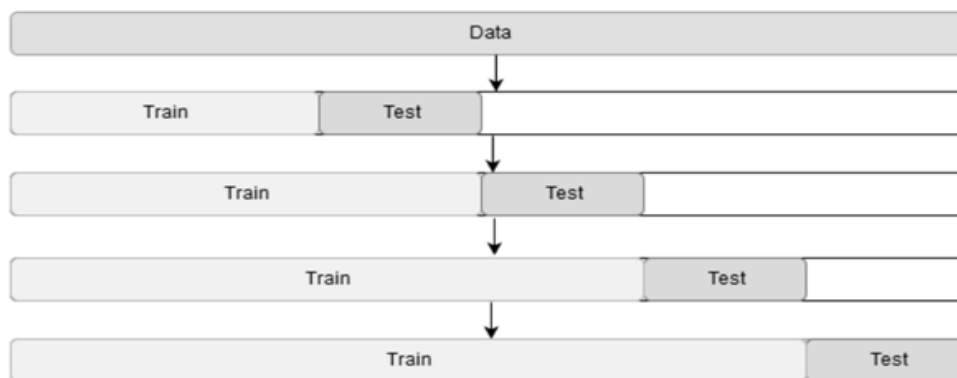
**Leave-p-out cross validation**, όπου εξαιρούνται  $p$  δεδομένα εκτός από το σύνολο δεδομένων που θα χρησιμοποιηθούν ως δεδομένα για validation.

**Leave-one-out cross validation**, εξαιρώντας ένα δεδομένο από το συνολικό για validation.

Διαφορετικές προσεγγίσεις εφαρμόζονται στο θέμα της εκπαίδευσης των νευρωνικών δικτύων. Ο ορισμός της αναλογίας δειγμάτων εκπαίδευσης προς το δείγμα επικύρωσης οδηγεί σε βελτίωση της απόδοσης και συνήθως τα ποσοστά που χρησιμοποιούνται είναι είτε 50%-50%, είτε 70%-30%. [Tkac & Verner] Μια τακτική εκπαίδευσης είναι να χωρίζεται το σύνολο εκπαίδευσης κατά το 80% για εκπαίδευση και 20 % για validation (out-of-sample-trading).

Για την περίπτωση δυναμικών δεδομένων όπως των χρηματοοικονομικών χρονοσειρών μια μέθοδος δοκιμής που χρησιμοποιείται είναι η “walk-forward” όπου τα νευρωνικά δίκτυα εκπαιδεύονται σε ένα παράθυρο δεδομένων που συνεχώς μεταβάλλεται χρησιμοποιώντας πιο πρόσφατα δεδομένα, αναβαθμίζοντας συνεχώς το σύνολο δεδομένων εκπαίδευσης και παρέχοντας πιο ρεαλιστική προσέγγιση για τις χρηματοοικονομικές τιμές. Μια επιπλέον μέθοδος αφορά τον μέσο όρο πολλαπλών προβλέψεων από διάφορα κινούμενα σύνολα δεδομένων. Σε μια διαφορετική εκδοχή εξετάζονται οι διαφορετικές συνέπειες από την χρήση διαφορετικής χρονική περιόδου καθώς μεγαλύτερα χρονικά παράθυρα αυξάνουν την πολυπλοκότητα του μοντέλου ενώ μικρότερα παράθυρα μπορούν να δημιουργούν καλύτερα μοντέλα εκπαίδευσης. [(DUNIS et al, 2016)]. Στην περίπτωση των χρονοσειρών μπορούν να χρησιμοποιηθούν πιο εξειδικευμένες μέθοδοι:

**Time Series Split Cross-Validation**, η μέθοδος μοιράζει το σύνολο δεδομένων σε ένα σύνολο για εκπαίδευση και ένα σύνολο για επικύρωση, όπου το σύνολο προς επικύρωση βρίσκεται πάντα χρονικά μπροστά από το σύνολο εκπαίδευσης. Τα σημεία που έχουν χρησιμοποιηθεί για επικύρωση της πρόβλεψης έπειτα χρησιμοποιούνται για εκπαίδευση και η διαδικασία επαναλαμβάνεται συνεχώς. [(Hyndman & Athanasopoulos, 2022)]



Εικόνα 13: Time Series Split Cross-Validation Πηγή [(Shrivastava, 2020)]

**Blocked Cross-Validation**, αποτελεί εξέλιξη της προηγούμενης μεθόδου μόνο που τώρα τίθενται περιθώρια μεταξύ των δεδομένων εκπαίδευσης και επικύρωσης ώστε να μην χρησιμοποιηθούν οι

τιμές δύο φορές, η δεύτερη εκδοχή είναι μεταξύ των δεδομένων εκπαίδευσης και επικύρωσης σε κάθε επανάληψη ώστε να αποτραπεί το μοντέλο να απομνημονεύει πρότυπα από μια αλληλεπίδραση στην επόμενη. [(Shrivastava, 2020)]

**Το σύνολο δεδομένων δοκιμής (Test Set)**, αποτελεί ξεχωριστό σύνολο δεδομένων από τα συνολικά δεδομένα. Χρησιμοποιεί μόνο τις ανεξάρτητες μεταβλητές και θα πρέπει να ακολουθεί την ίδια κατανομή δεδομένων με το σύνολο εκπαίδευσης. Χρησιμοποιείται για να διαπιστώσει την εφαρμογή το μοντέλο που δημιουργήθηκε κατά την διαδικασία επικύρωσης μέσω μετρικών συναρτήσεων όπως είναι η ακρίβεια/accuracy, precision/προσέγγιση, ευαισθησία /sensitivity, F-measure. Για την περίπτωση των μοντέλων για χρονοσειρές μπορεί να είναι προβληματική και η προτείνονται άλλες μέθοδοι όπως η rolling cross validation.

### **3.12 Επιλογή Μοντέλου**

Το μοντέλο που δημιουργείται θα πρέπει να έχει καλή ακρίβεια πρόβλεψης νέων δεδομένων να είναι όσο το δυνατόν απλούστερο, να έχει μεγάλη δυνατότητα γενίκευσης, δυνατότητα κλιμάκωσης σε μεγάλο σύνολο δεδομένων, να είναι εύρωστο δηλαδή να χειρίζεται θόρυβο ή δεδομένα με ελλειπείς τιμές και να έχει μεγάλη ταχύτητα δημιουργίας και μικρούς χρόνους δημιουργίας, δοκιμής και επικύρωσης όσο και επεξεργασίας σε πραγματικό χρόνο. Επίσης θα πρέπει να είναι αιτιολογήσιμος ο σχηματισμός του και επεξηγήσιμη η δομή και τα αποτελέσματά του.

#### **3.12.1 Μετρικές αξιολόγησης απόδοσης μοντέλου / Model performance evaluation metrics**

Οι Μετρικές συναρτήσεις χρησιμοποιούνται για την εκτίμηση ενός μοντέλου κατά την διάρκεια της δοκιμής και επικύρωσής του, και για την σύγκριση μεταξύ των μοντέλων, καθώς και για την εφαρμογή του μοντέλου σε διαφορετικά σύνολα δεδομένων. Η επιτυχία των μοντέλων μάθησης βασίζεται στην βελτιστοποίηση των μετρικών συναρτήσεων. Ανάλογα με το πρόβλημα της Μάθησης χρησιμοποιούνται και οι αντίστοιχες μετρικές συναρτήσεις. Κάποια παραδείγματα είναι:

- Classification Metrics (accuracy, precision, recall, F1-score, ROC, AUC)
- Regression Metrics (MSE, MAE)
- Ranking Metrics (MRR, DCG, NDCG)
- Statistical Metrics (Correlation)
- Computer Vision Metrics (PSNR, SSIM, IoU)
- NLP Metrics (Perplexity, BLEU score)
- Deep Learning Related Metrics (Inception score, Frechet Inception distance)

Οι Μετρικές είναι διαφορετικές από τις Συναρτήσεις Απώλειας/ Loss Functions που αποτελούν ένα μέτρο της εφαρμογής του μοντέλου κατά την διάρκεια της εκπαίδευσης.

Για την περίπτωση των θεμάτων Χρηματοοικονομικής πρόβλεψης καθώς τα περισσότερα προβλήματα πρόβλεψης κατηγοριοποιούνται ως πρόβλεψη τιμής χρηματοοικονομικού μεγέθους, που αντιμετωπίζεται ως θέμα παλινδρόμησης, υπολογίζοντας την διαφοροποίηση της τιμή πρόβλεψης από την πραγματική τιμή μέσω κάποιας μετρικής ή ως πρόβλεψης τάσης τιμής που αντιμετωπίζεται ως θέμα ταξινόμησης δύο ή τριών κατηγοριών ('Ανοδος/Up, Κάθοδος/Down, Σταθερή) με τις αντίστοιχες μετρικές συναρτήσεις.

Οι συχνότερα εφαρμοζόμενες μετρικές αριθμητικές συναρτήσεις για μοντέλα πρόβλεψης όπου  $Y_i$  είναι η πραγματική τιμή και  $\hat{Y}_i$  η τιμή πρόβλεψης είναι:

- **Μέσο Απόλυτο Σφάλμα/Mean Absolute Error,**

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|$$

- **Μέσο Τετραγωνικό Σφάλμα /Mean square error (MSE)** αποτελεί τον μέσο όρο του τετράγωνο της απόστασης της πραγματικής τιμής και την υπολογιζόμενης τιμής πρόβλεψης.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

- **Η τετραγωνική ρίζα του Μέσου Τετραγωνικού Σφάλματος/ Root mean squared error (RMSE)**

Για την περίπτωση θεμάτων ταξινόμησης χρησιμοποιείται ο Confusion Matrix ως εξής:

<b>Confusion Matrix</b>		Πραγματική Κλάση/Real Class	
		Y	Non Y
Προβλεπόμενη Κλάση/Predicted Class	Y	True Positive (TP)	False Positive (FP)
	Non Y	False Negative (FN)	True Negative (TN)

Όπου ο συνολικός αριθμός των προβλέψεων είναι :  $N = TP + FP + FN + TN$  ενώ οι θετικές περιπτώσεις  $K_T = TP + FN$  και ορίζονται οι παρακάτω μετρικές συναρτήσεις.

- Ακρίβεια/Accuracy = Αριθμός Ορθών Προβλέψεων/Συνολικός Αριθμός προβλέψεων= $(TP+TN)/N$
- Η Ευαισθησία ορίζεται ως Sensitivity= $TP/(TP+FN)$  (%)
- Η ιδιαιτερότητα ορίζεται ως Specificity= $TN= TN/(TN+FP)$  (%)
- Η ακρίβεια ισούται με Precision= $TP (TP+ FP)(\%)$
- Η ανάκληση ορίζεται ως Recall= $TP/ (TP+ FN)$  (%)
- Το F1 score= $2/(1/Precision)+ (1/Recall))$

Στις περιπτώσεις εμπορίας μεγεθών χρηματοοικονομικών χρονοσειρών μια υψηλή τιμή της ακρίβειας δεν συνεπάγεται υψηλή κερδοφορία καθώς θα πρέπει να συνυπολογιστούν και τα κόστη των συναλλαγών. Για αυτό υπάρχουν μετρικές συναρτήσεις που υπολογίζονται με βάση των κατηγοριών Ανόδου/Up και Καθόδου αν :

- UpAccuracy (%), το ποσοστό των πιθανών ακριβών Ανοδικών προβλέψεων κατά την επομενη περίοδο εμπορίας.
- DownAccuracy (%): το ποσοστό των πιθανών ακριβών Καθοδικών προβλέψεων κατά την επομενη περίοδο εμπορίας.
- LongAccuracy (%): έχει σχέση με την UpAccuracy καθώς μια εντολή buy/long trade προέρχεται απο μια Ανοδική πρόβλεψη, μπορεί να μετρήσει το θετικό ποσοστό κέρδους αυτών των εμπορικών κινήσεων χωρίς να προβλέπει τα κόστη εμπορίας.
- ShortAccuracy (%): σχετίζεται με την DownAccuracy καθώς μια εντολή sell/short trade προέρχεται απο μια Καθοδική πρόβλεψη, μπορεί να μετρήσει το θετικό ποσοστό κέρδους αυτών των εμπορικών κινήσεων χωρίς να προβλέπει τα κόστη εμπορίας.
- CumulativeReturn (%): Το ποσοστό των αποδόσεων που υπολογίζεται κατά το τέλος του κύκλου εμπορίας το οποίο υπολογίζεται ως εξής:
- CumulativeReturn για την  $i$  περίοδο:
- $CumulativeReturn(i) = (1 + CumulativeReturn(i-1)) * (1 + Return(i)) - 1$ .
- MaximumDrawdown: Δείχνει τη μέγιστη συσσωρευμένη απώλεια κατά την διάρκεια της εμπορίας
- AverageReturnPerTrade: Μέσος όρος των αποδόσεων κατά το τέλος του κύκλου εμπορίας. [(Gerlein et al 2016)]

Στις μελέτες γίνεται προσπάθεια να συνοψιστούν τα χαρακτηριστικά των επιτυχημένων μοντέλων για πρόβλεψη χρηματοοικονομικών χρονοσειρών. Κάποια από τα συμπεράσματα είναι τα εξής :

- Θα πρέπει να αντανακλά την μη γραμμική αλληλεπίδραση των παραγόντων που επιδρούν στις χρηματοοικονομικές χρονοσειρές και έπειτα να την ενσωματώνουν στο μοντέλο
- Θα πρέπει να ενσωματώνει τα μακροπρόθεσμα και βραχυπρόθεσμα ακολουθιακά χαρακτηριστικά εξάρτησης των χρηματοοικονομικών χρονοσειρών ώστε να μπορεί να αναγνωρίζει αυτόματα το μήκος της ακολουθίας εξάρτησης που βασίζονται στα δεδομένα εισόδου.
- Να έχει σημαντικά επίπεδα γενίκευσης και ακρίβεια πρόβλεψης σε δεδομένα εκτός δείγματος.

- Ισχυρή και εύρωστη ικανότητα μάθησης, παράκαμψη του θορύβου, αναγνώριση των μη γραμμικών, μη στατικών και άλλων πολύπλοκων χαρακτηριστικών των χρονοσειρών ώστε να μπορεί εξάγει τις σημαντικές πληροφορίες που ασκούν επιδράσεις στην αλλαγή των χρονοσειρών.

[(Tang et al, 2022)]

### 3.13 Μέθοδοι Μηχανικής Μάθησης

Οι παρακάτω μέθοδοι Μηχανικής Μάθησης για χρηματοοικονομική πρόβλεψη αναφέρονται στις δημοσιευμένες ανασκοπήσεις και εφαρμογές και χρησιμοποιούνται ως σημείο αναφοράς των αποδόσεων των μοντέλων Βαθιάς Μάθησης που αναφέρονται στην συνέχεια.

#### 3.13.1 Γενετικοί Αλγόριθμοι (ΓΑ) / Genetic Algorithms (GA)

Ο Γενετικός Αλγόριθμος αποτελεί μια ευρετική αναζήτηση που μπορεί να προσεγγίσει μια συνολική βέλτιστη λύση σε πολύπλοκους πολυδιάστατους χώρους. Έχει ως τρόπο λειτουργίας αρχές από την βιολογία, και την εξέλιξη. Λειτουργούν τροποποιώντας επανειλημμένα έναν πληθυσμό χρησιμοποιώντας γενετικούς τελεστές οι οποίοι είναι :

- Φυσικής επιλογή: η βελτίωση του πληθυσμού γίνεται με την επιβίωση του πιο προσαρμοσμένου (survival of the fittest). Οι δομές με την καλύτερη εφαρμογή διπλασιάζονται και δομές με την ελάχιστη εφαρμογή εξαλείφονται.
- Διασταύρωση/Crossover: όταν συνδυάζεται με την επιλογή, τα αποτελέσματα από τις καλές δομές συνδυάζονται για να δημιουργήσουν ακόμα καλύτερες δομές. Χρησιμοποιείται ο ρυθμός διασταύρωσης.
- Μετάλλαξη/Mutation: δημιουργούνται καινούρια γονότυπα, επιδιώκοντας την εύρεση της καλύτερης λύσης για ένα πρόβλημα όπως είναι ένα κάρνασ εμπορίας/ trading μετοχών Χρησιμοποιείται ο ρυθμός μετάλλαξης

Στα Χρηματοοικονομικά ο ΓΑ, χρησιμοποιείται για την πρόβλεψη τιμών αξιών, για την βελτιστοποίηση κανόνων trading με την επιλογή κατάλληλων τιμών παραμέτρων. Ο Γενετικός αλγόριθμος μπορεί να ανακαλύψει ότι μια μεταβλητή ή ένα δείκτης που χρησιμοποιείται ως μεταβλητή όπως η P/E, price-to-earnings ratio είναι καλύτερος δείκτης για μελλοντικές αποδόσεις από κάθε άλλη μεταβλητή στην οποία θα δώσει περισσότερη έμφαση. Στην μελέτη των [(Mahfoud & Mani, 1996)] διαπιστώνεται ότι ο Γενετικός Αλγόριθμος μπορεί αν κάνει πρόβλεψη και έχει ως πλεονέκτημά του ότι μπορεί να εξάγει αποτελέσματα με κατανοητούς κανόνες οι οποίοι μπορούν να ενσωματωθούν σε ένα το αποτέλεσμα του οποίου μπορεί να αξιοποιηθεί αυτόνομα ή να

χρησιμοποιηθεί περαιτέρω σε μια άλλα η διαδικασία μάθησης όπως ένα expert system ή ένα Νευρωνικό Δίκτυο.

### **3.13.2 Μέθοδοι Πυρήνων/ Kernel methods**

Οι μέθοδοι πυρήνων χαρτογραφούν τον χώρο των μεταβλητών εισόδου αναπαριστώντας τον σε έναν χώρο περισσότερων διαστάσεων με την χρήση συναρτήσεων πυρήνων. Οι συναρτήσεις πυρήνων αυτοί έχουν την ικανότητα να αναπαραστούν τα χαμηλότερης διάστασης δεδομένα σε υψηλότερης διάστασης δεδομένα. Τέτοιες συναρτήσεις πυρήνων είναι ο γραμμικός πυρήνας, ο πολυωνυμικός πυρήνας, ο πυρήνας Gauss, ο Σιγμοειδής, η Radial Basis Function. Με τον τρόπο αυτό μπορούν να δημιουργηθούν και να εκπαιδευτούν γραμμικά μοντέλα που έχουν ως αποτέλεσμα τον σχηματισμό μη πολύπλοκων μοντέλων, με καλή εφαρμογή που αντεπεξέρχονται στο δίλημα της προκατάληψης και απόκλισης (bias-variance).[(Chauhan S. N., 2019)]

#### **3.13.2.1 Support vector machine (SVM)**

Αποτελεί μέθοδο εποπτευόμενης μάθησης η οποία βασίζεται στην μέθοδο των πυρήνων που βασίζεται στο εμπειρικό σφάλμα και έναν όρο κανονικοποίησης (regularization) και στην αρχή της μείωσης του δομικού ρίσκου (structural risk minimization) η οποία χρησιμοποιείται για ταξινόμηση και παλινδρόμηση. Οι πυρήνες μπορεί να είναι γραμμικός, πολυωνυμικός, radial basis. Η μέθοδος δημιουργεί ένα N-διάστατο, υπερεπίπεδο αντιστοιχίζοντας τις τιμές των μεταβλητών με τέτοιο τρόπο ώστε η απόσταση μεταξύ των να είναι μέγιστη (κριτήριο maximum margin hyperplane) και να σχηματίζονται διακριτές κατηγορίες. Ο αριθμός των διαστάσεων του υπερεπιπέδου εξαρτάται από τον αριθμό των χαρακτηριστικών. Το υπερεπίπεδο αυτό αποτελεί και το όριο αποφάσεων χωρίζοντας τις κατηγορίες των σημείων. Ως Support Vector ονομάζονται τα σημεία που βρίσκονται στα όρια του υπερεπιπέδου και επηρεάζουν την θέση και τον προσανατολισμό του υπερεπιπέδου. Τα SVR χρησιμοποιείται για περιπτώσεις παλινδρόμησης αριθμητικών τιμών.

Πλεονεκτήματά τους είναι η ευκολία στην εκπαίδευση, η αποφυγή τοπικών ελαχίστων (σε αντίθεση με τα νευρωνικά δίκτυα μηχανική μάθησης) καθώς η λύση επιτυγχάνεται μέσω προγραμματισμού τετραγωνικών συναρτήσεων με γραμμικές παραμέτρους, προσεγγίζουν ολικά ακρότατα, η αποφυγή της υπερεφαρμογής (overfitting), έχουν καλή απόδοση πρόβλεψης μειώνοντας το άνω όριο του σφάλματος γενικοποίησης. Το μοντέλο έχει την ικανότητα να παρακάμπτει τα μη σχετικά δεδομένα και να βελτιώνει την ακρίβεια πρόβλεψης. Ωστόσο ο χρόνος που χρειάζεται για την προσέγγιση κάποιων συνόλων δεδομένων μπορεί να είναι μεγάλος. Επίσης δεν παρέχουν επαρκή αιτιολόγηση για τον τρόπο με τον οποίο οδηγούνται στα αποτελέσματα.

Μειονεκτήματά τους είναι ο μεγάλος χρόνος εκπαίδευσης, ειδικότερα σε μεγάλα σύνολα δεδομένων, η ρύθμιση των υπερπαραμέτρων του δικτύου σε περίπτωση διάσπαρτων μεταβλητών. Αποτελούν μαζί με τα νευρωνικά δίκτυα τις κυριότερες κατηγορίες μοντέλων πρόβλεψης για αριθμητικές και ονομαστικές τιμές. Χρησιμοποιήθηκαν σε εφαρμογές συστημάτων πιστοληπτικής ικανότητας.

### **3.13.3 Άλλες μέθοδοι μηχανικής Μάθησης**

#### **3.13.3.1 K-nearest neighbor (KNN)**

Η μέθοδος των K-εγγύτερων γειτόνων γνωστή και ως αλγόριθμος KNN αποτελεί έναν μη παραμετρικό αλγόριθμο που ταξινομεί τα δεδομένα βάσει την εγγύτητάς τους και της διασύνδεσης τους με τα υπόλοιπα διαθέσιμα δεδομένα. Ο αλγόριθμος θεωρεί ότι τα όμοια δεδομένα βρίσκονται κοντά μεταξύ τους. Με την χρήση μετρικής απόστασης επισυνάπτει μια κατηγορία ομάδας που βασίζεται στην συχνότερη κατηγορία ή μέσο όρο. Χρησιμοποιείται σε μεγάλα σύνολα δεδομένων μεγάλων διαστάσεων για εξαγωγή χαρακτηριστικών.

#### **3.13.3.2 Random forest**

Αποτελεί μέθοδο εποπτευόμενης εκμάθησης η οποία χρησιμοποιείται για ταξινόμηση και παλινδρόμηση, Η μέθοδος χρησιμοποιεί το μοντέλο δένδρων αποφάσεων “Decision Trees” ταξινόμησης και παλινδρόμησης, Δημιουργεί ένα “Δάσος” από δένδρα αποφάσεων “Decision Trees”.Κάθε δένδρο δημιουργεί μια πρόβλεψη από ένα σύνολο χαρακτηριστικών, όταν όλες οι προβλέψεις έχουν δημιουργηθεί η πιο κοινή κλάση αποτελεί την τελική πρόβλεψη.

Πλεονέκτημα της αποτελεί ότι εφαρμόζοντας τον συγκερασμό από πολλαπλά δένδρα αποφάσεων το μοντέλο δεν οδηγείται σε υπερεφαρμογή, και έχει καλή ακρίβεια πρόβλεψης.

Μειονεκτήματα της είναι ότι μικρές αλλαγές στο δείγμα εκπαίδευσης μπορούν αν οδηγήσουν σε μεγάλες αλλαγές στο μοντέλο, μπορεί να αγνοηθούν σημαντικά γνωρίσματα πληροφορίας, η κατανάλωση υπολογιστικών πόρων.

Επιπλέον υπάρχει μια πληθώρα από άλλες μεθόδους όπως είναι τα Δέντρα απόφασης / Decision trees), Κανόνες Συσχέτισης (Association Rules),ο Επαγωγικός λογικός προγραμματισμός (ILP), Συσταδοποίηση (Clustering), Δίκτυα Bayes, XGBoost.



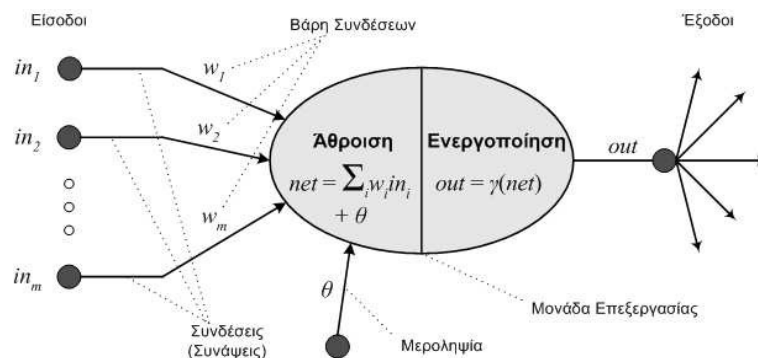
### 3.14 Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) / Artificial Neural Networks (ANN)

Τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) είναι μια κατηγορία αλγορίθμων μηχανικής μάθησης που μιμούνται την ανθρώπινη λειτουργία μάθησης. Δηλαδή μιμείται τον τρόπο που επεξεργάζονται τις πληροφορίες οι νευρώνες στον ανθρώπινο εγκέφαλο : Η γνώση αποκτάται από το δίκτυο από το περιβάλλον του μέσω μιας διαδικασίας μάθησης.

Τα ΤΝΔ λειτουργούν ως συστήματα Παράλληλης Κατανεμημένης Επεξεργασίας και απαρτίζεται από τα παρακάτω στοιχεία:

- Ένα σύνολο μονάδων επεξεργασίας
- Μια κατάσταση ενεργοποίησης
- Μια συνάρτηση Εξόδου για κάθε μονάδα
- Ένα κανόνα ενεργοποίησης που συνδυάζει τις εισόδους που επενεργούν σε μία μονάδα με την τρέχουσα κατάσταση ενεργοποίησης, ώστε να προκύψει η νέα κατάσταση ενεργοποίησης της μονάδας, σήμερα θεωρείται η συνάρτηση εξόδου
- Ένα μοντέλο συνδεσιμότητας μεταξύ των μονάδων
- Ένα κανόνα διάδοσης, του σήματος διαμέσου του δικτύου συνδέσεων μεταξύ των μονάδων
- Ένα κανόνα μάθησης, σύμφωνα με τον οποίο το μοντέλο συνδεσιμότητας μεταβάλλεται μέσω της εμπειρίας
- Ένα περιβάλλον μέσα στο οποίο το σύστημα πρέπει να λειτουργήσει.

[(Ζαπράνης 2005)]



Εικόνα 14: Η δομή ενός κόμβου νευρωνικού δικτύου [(Ζαπράνης 2005)]

Το Νευρωνικό Δίκτυο αποτελείται από διασυνδεδεμένες μονάδες, τους τεχνητούς νευρώνες ή κόμβους, σε παράλληλα επίπεδα (layers). Ο κάθε Τεχνητός Νευρώνας φέρει συνδέσεις με τους άλλους νευρώνες με τρόπο όμοιο με τις συνάψεις στον ανθρώπινο εγκέφαλο, μέσω ακμών που χαρακτηρίζονται από το βάρος (weight) και λειτουργούν ως εισοδοι και ως έξοδοι των πληροφοριών από και προς τους άλλους νευρώνες και χρησιμοποιούνται για την αποθήκευση

γνώσης[(Haykin S., 2009)]. Σε κάθε κόμβο υπάρχει και μια συνάρτηση ενεργοποίησης. Αν η πληροφορία που έρχεται από τις εισόδους ξεπερνάει κάποιο ποσοστό τότε πυροδοτεί την συνάρτηση ενεργοποίησης, η οποία μεταφέρει την τιμή της ως πληροφορία στο επόμενο επίπεδο κόμβων στο δίκτυο. Στην απλή μορφή του αποτελείται από ένα επίπεδο νευρώνων εισόδου, ένα κρυφό επίπεδο επεξεργασίας και ένα επίπεδο νευρώνων εξόδου.

### 3.14.1 Αλγόριθμος Οπισθοδιάδοσης / Backpropagation Algorithm

Δημοφιλής μέθοδος μάθησης στα νευρωνικά δίκτυα είναι η Backpropagation/ Οπισθοδιάδοση, η οποία αποτελεί έναν αλγόριθμο βελτιστοποίησης, υπολογίζοντας την gradient (διάνυσμα μερικών παραγώγων) μια συνάρτησης απώλειας / loss function σε σχέση με τις μεταβλητές του μοντέλου. Λειτουργεί με τον υπολογισμό της τιμής εξόδου/output και το σφάλμα/error κατά την εισαγωγή ενός συνόλου δεδομένων στο επίπεδο εισόδου του νευρωνικού δικτύου. Αν υπάρχει διαφορά μεταξύ της επιθυμητής τιμής και της πραγματικής, τότε τα σήματα σφαλμάτων μεταδίδονται προς τα πίσω στο νευρωνικό δίκτυο και τα βάρη αναδιαμορφώνονται από τις αρχικές τους συνθήκες για να μειώσουν αυτό το σφάλμα επαναλαμβάνοντας την διαδικασία. Ο συνδυασμός των βαρών που μηδενίζουν την συνάρτηση σφάλματος θεωρείται η λύση στο πρόβλημα της μάθησης.

Η μέθοδος χρησιμοποιεί τον αλγόριθμο βελτιστοποίησης της Κλίσης Καθόδου / Gradient Descent για την ελαχιστοποίηση του σφάλματος εξόδου της συνάρτησης απώλειας /loss function. Ο αλγόριθμος οπισθοδιάδοσης σφάλματος προσπαθεί να προσεγγίσει το ολικό ελάχιστο της συνάρτησης σφάλματος  $E(w)$  δηλαδή να ελαχιστοποιήσει το ολικό σφάλμα μεταξύ της μετρούμενης και την πραγματικής τιμής με παράμετρο τις τιμές των βαρών, καθώς το δίκτυο δημιουργεί τις εξόδους τη συνάρτηση απώλειας loss function  $C$ . Η μέθοδος της οπισθοδιάδοσης χρειάζεται μια διαφορίσιμη συνάρτηση ενεργοποίησης και ως τέτοιες χρησιμοποιούνται οι συναρτήσεις tanh-sigmoid, log-sigmoid, και περιστασιακά οι γραμμικές.

Για την εκπαίδευση το σύνολο εκπαίδευσης χωρίζεται στο σύνολο εκπαίδευσης (training set) βάση του οποίου υπολογίζονται τα σφάλματα κλίσης (error gradients) και η αναβάθμιση των βαρών και το σύνολο επικύρωσης (validation set) το οποίο καθορίζει το μέγιστο αριθμό επαναλήψεων ώστε να αποφευχθεί το overlearning. Καθώς ο αριθμός των επαναλήψεων αυξάνεται, το σφάλμα εκπαίδευσης πέφτει και το σφάλμα επικύρωσης πέφτει, φτάνει σε ένα ελάχιστο και τελικά αυξάνεται. Συνεχίζοντας την διαδικασία μάθησης μετά το σφάλμα επικύρωσης φτάνει σε ένα ελάχιστο οδηγεί σε υπερμάθηση / overlearning. Μετά το τέλος της διαδικασίας μάθησης ένα άλλο σύνολο δεδομένων χρησιμοποιείται για την επικύρωση και την επιβεβαίωση της ακρίβειας πρόβλεψης.

Η κανονικοποίηση των δεδομένων είναι απαραίτητη πριν την εκκίνηση της διαδικασίας εκπαίδευσης έτσι ώστε οι τιμές του μοντέλου να μην επηρεαστούν από τις αρνητικές τιμές ή την κλίμακα των μεταβολών. Η κλίμακα που επιλέγεται είναι συνήθως μεταξύ  $[0,1]$ .

Μειονέκτημα της είναι ότι μπορεί να συγκλίνει σε τοπικό ελάχιστο ενώ μπορεί κάτω από κάποιες συνθήκες η εκπαίδευση του δικτύου να σταματήσει (Παράλυση Δικτύου). Η ύπαρξη πολλών λύσεων που αντιστοιχούν στο ίδιο επίπεδο σφάλματος είναι προβληματική στα πλαίσια ενός γενικότερου πλαισίου επιλογής υποδειγμάτων.

### **3.14.2 Αλγόριθμοι βελτιστοποίησης**

Χρησιμοποιούνται για να βρεθεί ο βέλτιστος αριθμός παραμέτρων/μεταβλητών των Νευρωνικών Δικτύων, μέσω των οποίων αναβαθμίζονται τα βάρη των συνδέσεων μεταξύ των επιπέδων. Οι αλγόριθμοι που χρησιμοποιούνται για βελτιστοποίηση είναι

- Stochastic Gradient Descent (SGD), αποτελεί επαναληπτική μέθοδο εύρεσης των βέλτιστων παραμέτρων που ελαχιστοποιεί την συνάρτηση κόστους.
- SGD with Momentum, μνημονεύει την αναβάθμιση σε κάθε επανάληψη επιταχύνοντας την gradient descent.
- Adaptive Gradient Algorithm (AdaGrad), αποτελεί τροποποίηση της SGD που βελτιώνει το πρόβλημα της επικάλυψης εφαρμογής στον SGD αλγόριθμο
- Root Mean Square Propagation (RMSProp),επάγει την συχνότητα μάθησης για κάθε παράμετρο
- ADAM, αποτελεί αναβάθμιση του RMSProp που χρησιμοποιεί μέσους όρους των gradients και των δεύτερων στιγμών των gradients

#### **3.14.2.1 Stochastic Gradient Descent**

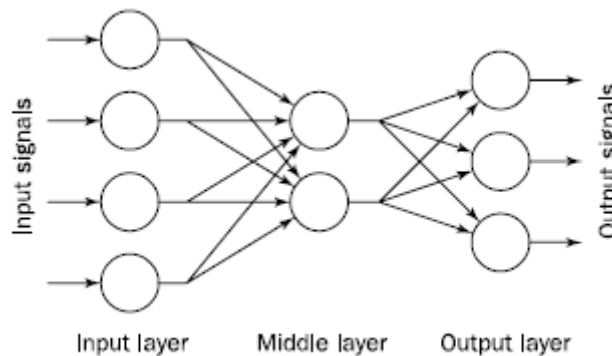
Η Gradient Descent μπορεί να χρησιμοποιηθεί για την μείωση της loss function ενός μοντέλου πρόβλεψης σε ένα σύνολο εκπαίδευσης όπως η ταξινόμηση ή παλινδρόμηση και η διαμόρφωση αυτή ονομάζεται Stochastic Gradient Descent και αποτελεί την πιο συχνό αλγόριθμος βελτιστοποίησης για την μηχανική μάθηση και την Βαθιά Μάθηση αντίστοιχα. [(Goddfellow et al 2022a)]

### 3.14.3 Είδη Νευρωνικών Δικτύων

Η πιο απλή μορφή νευρώνα είναι το Perceptron και βασίζεται στο νευρωνικό μοντέλο των McCulloch & Pitts που αποτελείται από έναν νευρώνα με προσαρμοζόμενα βάρη, εισόδους και μια έξοδο. Το perceptron λειτουργεί ως δυαδικός ταξινομητής και αποτελείται από έναν γραμμικό συνδυαστή που θέτει μια έξοδο μέσω ενός hard limiter με τιμές +1, -1. Η μάθηση πραγματοποιείται μέσω ρύθμισης των βαρών, από την αρχική τυχαία επισυναπτόμενη τιμή, για την μείωση της διαφοράς μεταξύ της πραγματικής και της επιδιωκόμενης τιμής. Το perceptron μπορεί να μάθει μόνο γραμμικά διαχωρίσιμες συναρτήσεις. Τα νευρωνικά δίκτυα περιέχουν επίπεδα από διασυνδεδεμένους κόμβους-perceptron.

### 3.14.4 Νευρωνικό Δίκτυο Προσω τροφοδότησης/Feed-Forward Neural Network

Το ΤΝΔ πρόσω τροφοδότησης (Feed-Forward Neural Networks -FNN) αποτελείται από ένα επίπεδο με κόμβους εισόδου, τουλάχιστον ένα επίπεδο με κρυφούς κόμβους και ένα επίπεδο με κόμβους εξόδου. Η κατασκευή τους μπορεί να γίνει με διάφορες μονάδες, όπως το απλό perceptron. Κάθε επίπεδο μπορεί να περιλαμβάνει μεγάλο αριθμό κόμβων.



Εικόνα 15 : Νευρωνικό Δίκτυο Προσω τροφοδότησης

Όταν το Δίκτυο πρόσω τροφοδότησης αποτελείται από 2 ή περισσότερα κρυφά επίπεδα θεωρείται δίκτυο Βαθιάς Μάθησης. Η πληροφορία λαμβάνεται από τους κόμβους εισόδου και κινείται μόνο προς μια κατεύθυνση προς τους κόμβους εξόδου. Τα βάρη των κόμβων αρχικοποιούνται τυχαία. Οι τιμές εισόδου πολλαπλασιάζονται από τα ανάλογα βάρη σε κάθε κόμβο του κρυφού επιπέδου και μια τιμή μεροληψίας προστίθεται στο άθροισμα του κάθε κόμβου. Η τιμή τελικά μετατρέπεται στην έξοδο του κόμβου χρησιμοποιώντας την συνάρτηση ενεργοποίησης

Παραδείγματα τέτοιων δικτύων είναι το Increasing-Decreasing-Linear Processing Units (IDLNN) όπου κάθε επίπεδο αποτελείται από ένα σύνολο από increasing-decreasing units. Stochastic Time Effective Function Neural Network, Radial Basis Function Network (RBFN).[(Jiang 2021)]

Τα δίκτυα τροφοδότησης με ένα τουλάχιστον κρυφό στρώμα και μη γραμμικές συναρτήσεις ενεργοποίησης διαθέτουν την ιδιότητα της καθολικής προσέγγισης, δηλαδή μπορούν να προσεγγίσουν οποιαδήποτε συνάρτηση και να χειριστούν γραμμικά διαχωρίσιμα και μη γραμμικά διαχωρίσιμα δεδομένα.

### **3.14.5 Multi-Layer Perceptron (MLP)**

Αποτελεί την πιο απλή μορφή Τεχνητού Νευρωνικού Δικτύου εμπρόσθιας τροφοδότησης που αποτελείται από ένα τουλάχιστον επίπεδο από νευρώνες εισόδου, ένα επίπεδο κρυφών νευρώνων και ένα επίπεδο νευρώνων εξόδου και μια συνάρτηση ενεργοποίησης. Η υλοποίηση του δικτύου γίνεται με χρήση της βασικής μονάδας νευρώνων perceptron, διασυνδεδεμένα σε πολλαπλά επίπεδα. Η μάθηση γίνεται όταν ένα σύνολο δεδομένων επισυναφθεί στο επίπεδο εισόδου, από το επίπεδο εισόδου προς την έξοδο, και αν η έξοδος δεν είναι η επιθυμητή, το σφάλμα (error) που υπολογίζεται τροφοδοτείται προς τα πίσω αναβαθμίζοντας τα βάρη του δικτύου και η διαδικασία επαναλαμβάνεται μέχρι το σφάλμα αυτό να μειωθεί στο μηδέν. Ο κάθε κόμβος σε κάθε επίπεδο είναι διασυνδεδεμένος με όλους τους κόμβους στο επόμενο επίπεδο.

### **3.14.6 Radial-basis function network (RBF)**

Αποτελούν νευρωνικό δίκτυο πρόσω τροφοδότησης που έχουν ως συνάρτηση ενεργοποίησης χρησιμοποιεί στο κρυφό επίπεδο την μη γραμμική συνάρτηση Radial Basis Function. Χαρακτηριστικά της είναι η καθολική προσέγγιση και η ταχύτητα. Χρησιμοποιείται για προσέγγιση συναρτήσεων, πρόβλεψη χρονοσειρών, ταξινόμηση και συστήματα ελέγχου.

### **3.14.7 Bayesian Regularized Artificial Neural Network (BRANNs)**

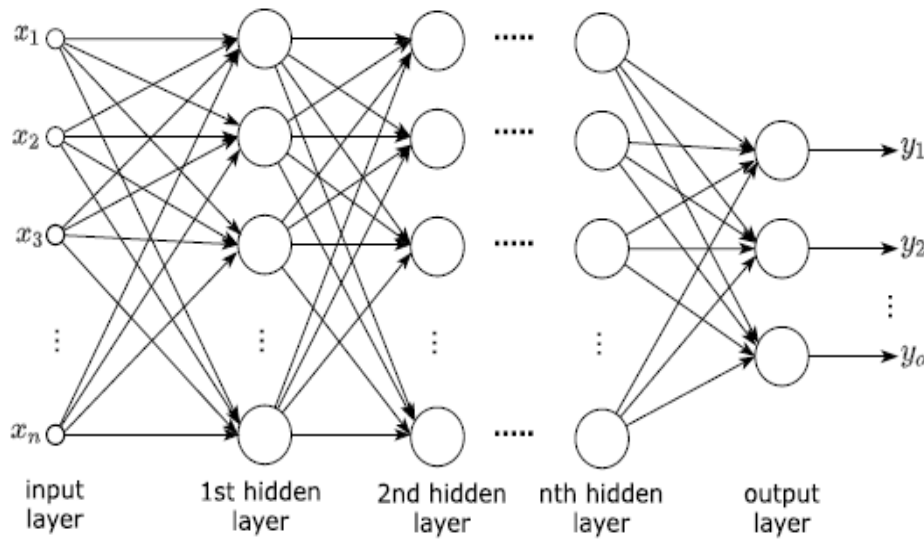
Αποτελεί ένα πιθανοκρατικό γραφικό μοντέλο, που αναπαριστά ένα σύνολο μεταβλητών με τις υπο συνθήκη εξαρτήσεις τους μέσω ενός ακυκλικού γραφήματος (Directed Acyclic Graph).

Είναι πιο εύρωστα από τα δίκτυα οπίσθιας διάδοσης (backpropagation) και αποφεύγουν την διαδικασία του cross-validation. Επιπλέον δεν οδηγούνται σε υπερπροσαρμογή/overfitting. Τα βάρη λαμβάνουν κατανομές πιθανότητας, οι οποίες κατανομές προσδιορίζονται με μια διαδικασία που ονομάζεται marginalization.

### **3.14.8 Βαθιά Μάθηση (BM) / Deep Learning (DL)**

Η κατηγορία μεθόδων Βαθιάς Μάθησης (BM) / Deep Learning (DL) αποτελεί εξέλιξη των Τεχνητών Νευρωνικών Δικτύων των προηγούμενων δεκαετιών, η οποία κατέστη δυνατή από την αύξηση της υπολογιστικής ισχύος των υπολογιστών, αξιοποιώντας την αρχιτεκτονική των περισσότερων διασυνδεδεμένων επιπέδων (layers). Διαθέτουν περισσότερα πολλαπλά κρυφά

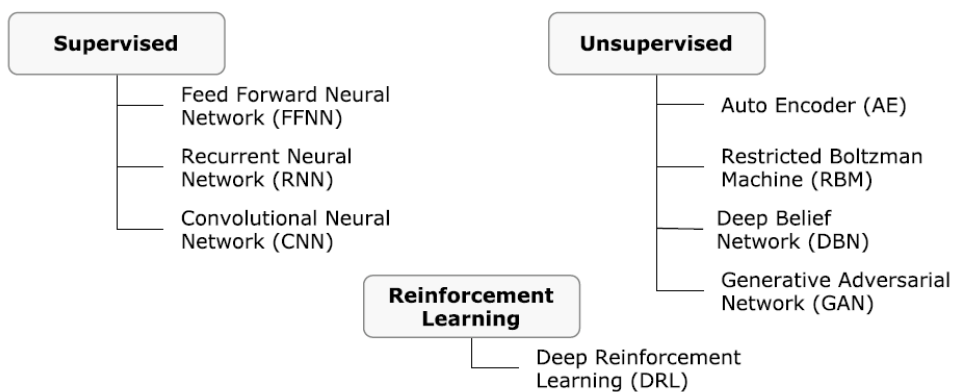
επίπεδα (layers) και ικανότητα παράλληλης επεξεργασίας τα οποία ονομάζονται Βαθιά Νευρωνικά Δίκτυα (BND) / Deep Neural Network (DNN). Έχουν την δυνατότητα να προσεγγίζουν πολλές χρονοσειρές, να ανιχνεύουν πολύπλοκες σχέσεις, πρότυπα, συσχετίσεις σε μεγάλου όγκου δεδομένα, πολλών διαστάσεων.



Εικόνα 16 : Νευρωνικά Δίκτυα Πρόσω τροφοδότησης-FFNN, Πηγή: [(Olorunnimbe et al,2022)]

Με τα BND αυτοματοποιήθηκε το θέμα της εξαγωγής χαρακτηριστικών καθώς δεν χρειάζονται να προσδιοριστούν χαρακτηριστικά γνωρίσματα.

Οι μέθοδοι BM αφορά μεθόδους εποπτευόμενης μάθησης, μη εποπτευόμενης μάθησης, εξαναγκασμένης. Ειδικότερα έχουν την ικανότητα να υπερκεράσει το πρόβλημα που εμφανίζεται με την υπερεφαρμογή/overfitting των μοντέλων. Επίσης έχουν την δυνατότητα να επεξεργάζονται αδόμητα δεδομένα / unstructured data όπως έγγραφα, εικόνες, κείμενο, φωνή με ανάλογες εφαρμογές σε επεξεργασία εικόνας, ήχου, video, χρονοσειρών. Ειδικότερα στα χρηματοοικονομικά εφαρμόζονται στην πρόβλεψη χρηματοοικονομικών χρονοσειρών, algorithmic trading, credit risk assessment, portfolio allocation, asset pricing, fraud detection, derivatives.

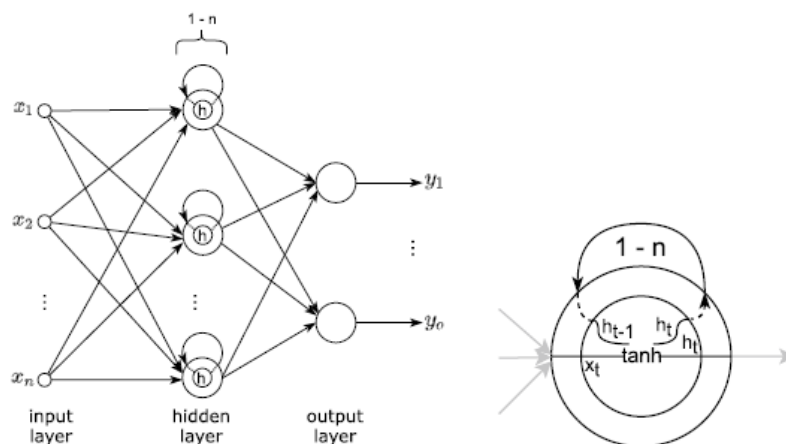


Εικόνα 17: Τύποι αλγορίθμων Μηχανικής Μάθησης για πρόβλεψη χρηματοοικονομικών χρονοσειρών

Πηγή: [(Olorunnimbe & Viktor 2022)]

### 3.14.9 Νευρωνικό Δίκτυο Ανατροφοδότησης / Recurrent Neural Networks (RNN)

Το Νευρωνικό Δίκτυο Ανατροφοδότησης (RNN), αποτελεί δομή ΒΜ και χρησιμοποιείται για αναγνώριση προτύπων σε ακολουθιακά δεδομένα όπως ήχος, χρονοσειρές, κείμενο. Στην αρχιτεκτονική τους περιέχουν στο κρυφό επίπεδο, βρόχους ανάδρασης, που έχουν την ικανότητα να μεταφέρουν την έξοδο από κάποιους κόμβους ως εισόδους στους ίδιους κόμβους, καθώς μεταφέρουν την πληροφορία από το ένα επίπεδο στο άλλο, λειτουργώντας ως δυναμική βραχυπρόθεσμη μνήμη. Έχουν την ικανότητα να κρατάνε την αναπαράσταση των προηγούμενων δεδομένων εισόδου με επιδράσεις στην μάθηση και βελτίωση του δικτύου.



Εικόνα 17: Νευρωνικά Δίκτυα Ανατροφοδότησης-RNN, Πηγή: [(Olorunnimbe et al,2022)]

Η διαφορά των RNN και των FNN είναι ότι η κάθε μονάδα λαμβάνει την παρούσα και προηγούμενη σειρά δεδομένων την ίδια στιγμή. Σε αντίθεση με τα δίκτυα προσω τροφοδότησης τα δίκτυα ανατροφοδότησης μοιράζονται την ίδια παράμετρο βάρους σε κάθε επίπεδο του δικτύου αλλά μεταβάλλονται μέσω των διαδικασιών backpropagation και gradient descent. Συναρτήσεις ενεργοποίησης που χρησιμοποιούν είναι η Σιγμοειδής, η εφαπτομενική, η ReLu. Υπάρχουν διάφορα είδη αρχιτεκτονικής RNN: Ένα-προς-ένα, Ένα προς πολλά, Πολλά προς ένα, Πολλά προς πολλά.

Μειονέκτημά του δικτύου είναι ότι δεν μπορεί να διατηρήσει την πληροφορία από προηγούμενες καταστάσεις για μεγάλες χρονικές περιόδους. [(Olorunnimbe et al,2022)] Τα RNN παρόλο που στόχος τους είναι η δημιουργία μεγάλης κλίμακας εξαρτήσεις, οι μελέτες δείχνουν όταν τα δεδομένα αποθηκεύονται για μεγάλο χρόνο η μάθηση αποτυγχάνει. Τα RNNs γίνονται πολύπλοκα όσο ο χρόνος της εκπαίδευσης αυξάνει.

Λόγω την μικρής μνήμη μπορεί να αναπτυχθεί το φαινόμενο vanishing gradient (εξαφανιζόμενη κλίση) όπου ενδέχεται να υπάρξει το φαινόμενο τα βάρη των νευρώνων να σταματήσουν να αναβαθμίζονται και ακόμα και το δίκτυο να σταματήσει την εκπαίδευσή του. Για να ξεπεραστεί

αυτό το πρόβλημα έχουν αναπτυχθεί οι μέθοδοι Long-Short term Memory (LSTM) και Gated Recurrent Unit (GRU).

### **3.14.10 Bidirectional Recurrent Neural Network (BRNN)**

Ειδικότερες αρχιτεκτονικές των RNN είναι η Bidirectional Recurrent Neural Networks (BRNN). Σε αντίθεση με τα RNN τα BRNN μπορούν να τροφοδοτούν μελλοντικούς κόμβους για να βελτιώσουν την ακρίβεια.

### **3.14.11 Πολυεπίπεδο Perceptron Βαθιάς Μάθησης/ Deep Multilayer Perceptron (DMLP)**

Το Νευρωνικό Δίκτυο MLP BM ή DMLP ανήκει στα νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης με πολλαπλά κρυφά επίπεδα που σχηματίζονται με την βασική δομή του Perceptron. Αποτελείται από το επίπεδο εισόδου, το επίπεδο εξόδου, και τα κρυφά επίπεδα. Χρησιμοποιεί την τεχνική back propagation. Ως συναρτήσεις ενεργοποίησης εφαρμόζονται οι συναρτήσεις ReLUs, sigmoid functions, και hyperbolic tangent.

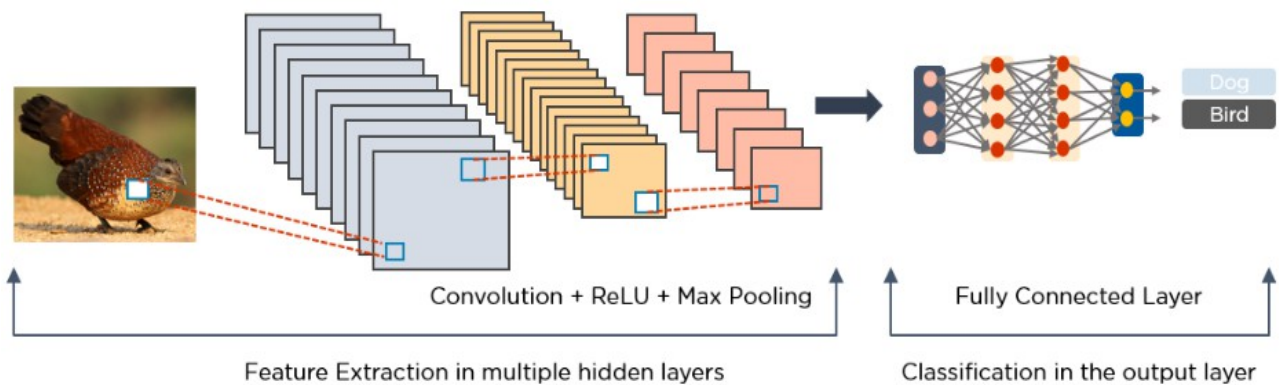
### **3.14.12 Νευρωνικό Δίκτυο Περιελιγμού / Convolutional Neural Network (CNN)**

Το Νευρωνικό Δίκτυο Περιελιγμού / Convolutional Neural Network (CNN), αποτελεί έναν από τους πρώτους αλγόριθμους ΤΝΔ που λόγω της δομής των πολλαπλών κρυφών επιπέδων που διαθέτει θεωρήθηκε Βαθιάς Μάθησης. Είναι νευρωνικό δίκτυο που χρησιμοποιεί την μαθηματική λειτουργία του περιελιγμού με χρήση πινάκων τιμών χαρακτηριστικών. Αποτελείται από πολλαπλά Multilayer Perceptrons. Η δομή του μοντέλου CNN αποτελείται από ένα επίπεδο εισόδου, ένα επίπεδο εξόδου, ένα επίπεδο convolution, ένα pooling layer, ένα επίπεδο fully connected layer, και ένα επίπεδο εξόδου. Το convolutional layer, και το pooling layer, χρησιμοποιούνται κατά επανάληψη. Το μέγεθος του φίλτρου και ο τρόπος που προσεγγίζει τα δεδομένα αποτελεί μέρος των υπερπαραμέτρων του δικτύου. Η μάθηση στηρίζεται στην μέθοδο της backpropagation με συνάρτηση ενεργοποίησης την ReLU. με χρήση φίλτρων (η κάθε ομάδα νευρώνων) και διεργασίες pooling. Είναι ειδικοί αλγόριθμοι για την διαχείριση δεδομένων εικόνας pixel με εφαρμογές σε αναγνώριση εικόνας και επεξεργασίας. Οι μεταβλητές εισόδου έχουν δομημένη μορφή-πίνακα όπως οι εικόνες και αποτελούν το πρώτο convolution layer. Ως έξοδο δίνει αποτέλεσμα ταξινόμησης/classification.

- Το επίπεδο Convolution αποτελείται από φίλτρα (μια ομάδα νευρώνων) για την εκτέλεση της λειτουργίας convolution/περιελιγμού με την οποία το διάλυμα εισόδου επισυνάπτεται



σε έναν χάρτη χαρακτηριστικών. Μπορεί να θεωρηθεί ως ένα σύνολο δισδιάστατων δεδομένων μορφής πίνακα, το οποίο περιέχει παραμέτρους μάθησης.



Εικόνα 19: Convolutional Neural Network (CNN), Πηγή [www.simplilearn.com](http://www.simplilearn.com)

- Το Rectified Linear Unit (ReLU) επίπεδο εκτελεί λειτουργίες στα δεδομένα και το αποτέλεσμα είναι ένας χάρτης χαρακτηριστικών.
- Το επίπεδο Pooling, μειώνει τις διαστάσεις των εισερχόμενων δεδομένων χάρτη χαρακτηριστικών εκτελώντας μια δειγματοληψία. Έπειτα μετατρέπει τους 2-διάστατους πίνακες από τον μειωμένο χάρτη χαρακτηριστικών σε ένα συνεχές γραμμικό διάνυσμα τιμών. Μ' αυτόν τον τρόπο μειώνεται ο αριθμός παραμέτρων μάθησης. Η μέθοδος που χρησιμοποιείται για την εξαγωγή της μέγιστης τιμής χαρακτηριστικών είναι η Max-Pooling. Το μέγιστο επίπεδο Pooling και το convolution επίπεδο αποτελούν το βάθος δικτύου.
- Το πλήρως συνδεδεμένο επίπεδο Fully Connected Layer(FC) MLPv ενώνει κάθε νευρώνα σε ένα επίπεδο σε κάθε νευρώνα στο άλλο επίπεδο. Δημιουργείται όταν το γραμμικό διάνυσμα από το pooling επίπεδο αποτελεί είσοδο στο νευρωνικό δίκτυο Multilayer Perceptron (MLP), το οποίο αναγνωρίζει και ταξινομεί τα αρχικά δεδομένα όπως εικόνες.

Το πλεονέκτημα των CNN σε σχέση με άλλους αλγόριθμους BM, είναι η μείωση του αριθμού των παραμέτρων προς εκπαίδευση, η δυνατότητα προσέγγισης των χωρικών χαρακτηριστικών των δεδομένων, η εξοικονόμηση βασικών πόρων για επεξεργασία,

Χρησιμοποιείται για επεξεργασία εικόνας και αναγνώριση αντικειμένων, εξαγωγή χαρακτηριστικών με εφαρμογές σε εικόνες ενώ χρησιμοποιούνται και για πρόβλεψη τιμών χρονοσειρών με καλύτερα αποτελέσματα από άλλες μεθόδους πρόβλεψης. [(Tang et al, 2022)]

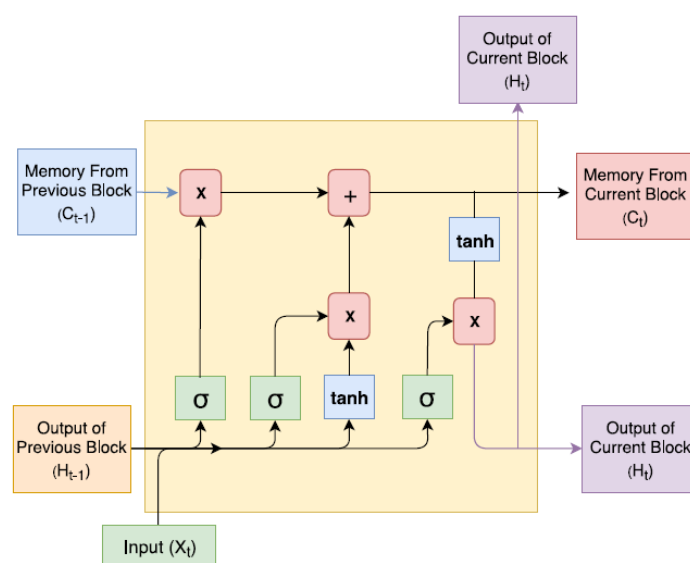
### 3.14.13 Long-Short Term Memory (LSTM) /Μνήμη Μακροπρόθεσμη-Βραχυπρόθεσμη

Τα Νευρωνικά Δίκτυα Μακράς-Βραχείας Μνήμης / Long Short Term Memory (LSTM) αποτελούν έναν εξειδικευμένο τύπο αρχιτεκτονικής RNN και χρησιμοποιούνται για ακολουθιακά δεδομένα. Τα LSTM καλύπτουν το πρόβλημα των μακροπρόθεσμων συσχετίσεων δεδομένων έχοντας

συνδέσεις ανάδρασης που μπορούν να θυμούνται τιμές των νευρώνων βραχυπρόθεσμα και μακροπρόθεσμα. Σε αντίθεση με τα RNN, τα οποία δεν είναι επαρκή για να θυμούνται μακροπρόθεσμες συσχετίσεις των δεδομένων.

Αποτελείται από ένα επίπεδο εισόδου, ένα ή περισσότερα επίπεδα κρυφά, και ένα επίπεδο εξόδου. Ο αριθμός των νευρώνων στο επίπεδο εισόδου είναι ίσος με τον αριθμό των επεξηγηματικών μεταβλητών (χώρος χαρακτηριστικών). Ο αριθμός των νευρώνων στο επίπεδο εξόδου αναπαριστά τον χώρο εξόδου (output space). Για να πραγματοποιεί την λειτουργία των μακροπρόθεσμων συσχετίσεων το LSTM αποτελείται από μονάδες μνήμης LSTM, που αποτελούν ουσιαστικά ένα σύνολο από δίκτυα RNN, οι οποίες η κάθε μία έχει μια θύρα εισόδου (input gate), μια θύρα εξόδου (output gate), και μία θύρα μνήμης (forget gate), που είναι ανάλογες με τις λειτουργίες εγγραφής (ποια πληροφορία θα αφαιρεθεί από την μνήμη), ανάγνωσης (ποια πληροφορία θα προστεθεί στην μνήμη) και ρύθμισης (ποια πληροφορία από την μνήμη θα χρησιμοποιηθεί ως έξοδος). Συνδυαζόμενες φτιάχνουν ένα LSTM επίπεδο. Η μακροπρόθεσμη μνήμη αφορά τα βάρη που μαθαίνονται και η βραχυπρόθεσμη αφορά την εσωτερική κατάσταση των κελιών. Η μακροπρόθεσμη μνήμη αναφέρεται στην εκμάθηση των βαρών ενώ η βραχυπρόθεσμη μνήμη αναφέρεται στην εσωτερική κατάσταση των κελιών. Η κάθε μονάδα θυμάται τις τιμές σε διάφορα χρονικά διαστήματα και ελέγχει την ροή της πληροφορίας που χρειάζεται για να προβλέψει την έξοδο επηρεάζοντας το τελικό αποτέλεσμα και βελτιώνοντας την απόδοση.

Πλεονεκτήματά των LSTM είναι η ικανότητα επαγωγικής μάθησης γνωρισμάτων στα πολύπλοκα χαρακτηριστικά των χρηματοοικονομικών χρονοσειρών σε σχέση με παλιότερες μεθόδους, έχουν καλύτερη δυνατότητα γενίκευσης και πρόβλεψης.



Εικόνα 20: Η βασική μονάδα LSTM( Πηγή:[(Ozbayoglu et al , 2022)]

Μειονεκτήματά τους αποτελούν η ανάγκη μετασχηματισμού των δεδομένων προς χρήση, η ρύθμιση των υπερπαραμέτρων του μοντέλου. [(Tang et al, 2022)]

Οι αναβαθμίσεις των βαρών για την εκπαίδευση του δικτύου γίνονται με διάφορες μεθόδους όπως η Gradient Descent και η Backpropagation Through Time (BPTT), χρησιμοποιείται χρησιμοποιείται για την αναβάθμιση των βαρών για τα LSTM.

Τα LSTM μπορούν να χρησιμοποιηθούν παράλληλα με την χρήση μεθόδων Αυτοκωδικοποιητών και Attention Mechanisms σε ανάλυση σειρών χρονικών δεδομένων με βελτίωση της απόδοσής τους. [(Sezer et al, 2019)]

### **3.14.14 Gated Recurrent Units (GRUs)**

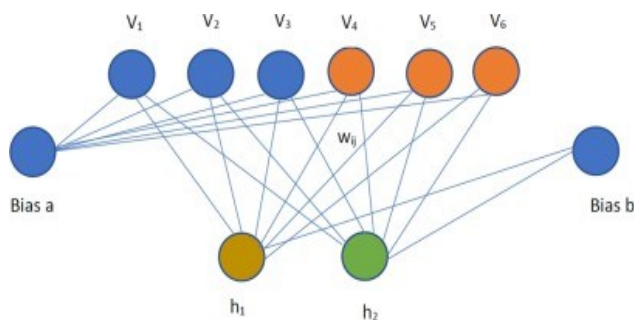
Τα Νευρωνικά Δίκτυα Gated Recurrent Unit (GRU) αποτελούν έναν εξειδικευμένο τύπο αρχιτεκτονικής RNN τα οποία μπορούν να διατηρήσουν μακροπρόθεσμες συσχετίσεις σε ακολουθιακά δεδομένα. Αυτή η RNN αρχιτεκτονική είναι παρόμοια με την LSTM, ενώ η διαφοροποίηση έγκειται στο γεγονός ότι αντί για νευρώνα κατάστασης για να ρυθμίσει την πληροφορία χρησιμοποιεί κρυφές καταστάσεις οι οποίες έχουν αντί για τρεις θύρες, δύο θύρες, μία θύρα reset και μια θύρα update οι οποίες ελέγχουν την ροή της πληροφορίας.

### **3.14.15 Περιορισμένες μηχανές Boltzmann / Restricted Boltzmann Machines (RBM)**

Ο τύπος TND Restricted Boltzmann Machine (RBM) αποτελεί στοχαστικό νευρωνικό δίκτυο, δύο επιπέδων, η αρχιτεκτονική του οποίου πραγματοποιεί μάθηση της κατανομής πιθανότητας των δεδομένων εισόδου, βασιζόμενη σε δείγματα αυτής της κατανομής και αποδίδοντας τις πιθανοκρατικές σχέσεις μεταξύ των μεταβλητών με γραφικό τρόπο (generative graphical model). Δοθείσες κάποιες τιμές δείγματος από τα δεδομένα εισόδου το δίκτυο είναι δυνατό να δημιουργήσει την κατανομή των τιμών των δεδομένων εισόδου.

Η αρχιτεκτονική του είναι εμπνευσμένη από την στατιστική φυσική και συγκεκριμένα από την συνάρτηση πυκνότητας πιθανότητας ή κατανομή Boltzmann. Αποτελείται από ένα διμερές, διπλά ανακατευθυνόμενο γραφικό μοντέλο (Boltzmann machine). Το πρώτο επίπεδο είναι το φανερό, που περιέχει  $n$  φανερές μονάδες που είναι και οι τιμές εισόδου και  $m$  κρυφές μονάδες που αποτελούν τις κρυφές μεταβλητές, για να φτιάξει τις εξαρτήσεις μεταξύ των παρατηρούμενων μεταβλητών. [(Fischer & Igel, 2012)] Οι κόμβοι του δικτύου που ανήκουν στο ίδιο επίπεδο δεν συνδέονται μεταξύ τους, ενώ όλοι οι κόμβοι του ενός επιπέδου συνδέονται με όλους τους κόμβους του επόμενου επιπέδου, αυτός είναι και ο περιορισμός που έχει δώσει στο δίκτυο αυτό το όνομα (restricted). Κάθε μονάδα αποτελεί μια υπολογιστική διαδικασία η οποία επεξεργάζεται την είσοδο και κάνει στοχαστικές αποφάσεις για το αν αυτός ο κόμβος θα μεταδώσει την είσοδο. Οι είσοδοι

πολλαπλασιάζονται από ειδικά βάρη (weights) ενώ προστίθεται και ένα ποσοστό μεροληψίας, ενώ στην συνέχεια οι τιμές εισέρχονται από μια συνάρτηση ενεργοποίησης. Τα αποτελέσματα των εξόδων επανέρχονται στο δίκτυο ως εισοδοί και ξαναοδηγούνται στις εξόδους. Οι τιμές των εισόδων πριν και μετά την επεξεργασία συγκρίνονται για την εύρεση της παραμέτρου βελτιστοποίησης η οποία ελαχιστοποιείται σε κάθε πέρασμα καθώς η μάθηση πραγματοποιείται πολλαπλές φορές στο δίκτυο.

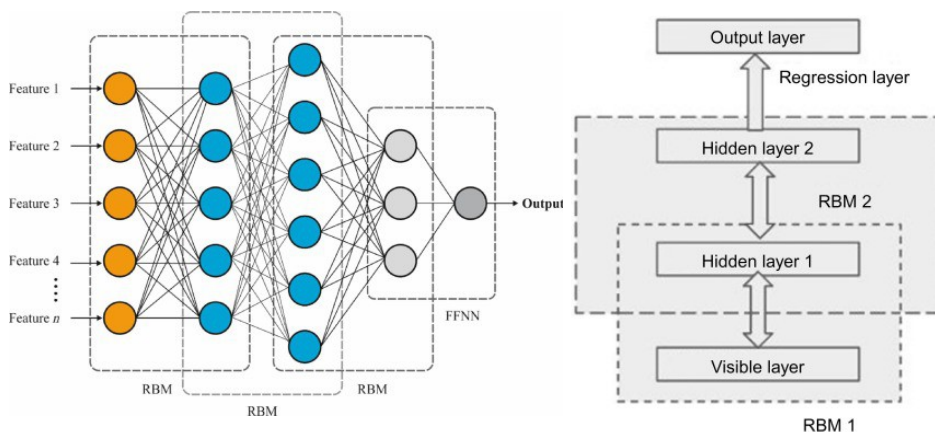


Εικόνα 21:RBM Πηγή ScienceDirect.com

Χρησιμοποιείται κυρίως για μείωση διαστάσεων/dimensionality reduction, ταξινόμηση/classification, παλινδρόμηση, μάθηση γνωρισμάτων/feature learning σε εφαρμογές εικόνας, κειμένου.

### 3.14.16 Deep Belief Network (DBN)

Το Deep Belief Network (DBN) αποτελεί ένα ΤΝΔ μή εποπτευόμενης βαθιάς μάθησης (Deep Neural Network), που χρησιμοποιεί πολλαπλά επίπεδα από το μοντέλο Restricted Boltzman Machines (RBMs) ή από το μοντέλο Autoencoder (AE). Μπορεί να πραγματοποιήσει μη εποπτευόμενη μάθηση της κατανομής πιθανότητας των δεδομένων εισόδου, βασισμένη σε δείγματα αυτής της κατανομής και αποδίδοντας τις συσχετίσεις μεταξύ των μεταβλητών με γραφικό τρόπο (generative graphical model).



Εικόνα22:DBN Πηγή sciencedirect.com

Η εκπαίδευση τους γίνεται μέσω δύο σταδίων αρχικά layer-by-layer που αναφέρεται στην μη εποπτευόμενη μάθηση μέσω της χρήσης του Greedy Αλγορίθμου και σε δεύτερο στάδιο στην εποπτευόμενη μάθηση μέσω fine-tuning όπου χρησιμοποιείται το σφάλμα της μεθόδου backpropagation. Κατά την διάρκεια της εκπαίδευσης στο πρώτο στάδιο ο χώρος εισόδου ανακατασκευάζεται σε μια πιθανοκρατική μορφή, κατά το δεύτερο στάδιο στο δίκτυο αρχίζουν να διακρίνονται τα διακριτικοποιημένα χαρακτηριστικά την εφαρμογή της ταξινόμησης.

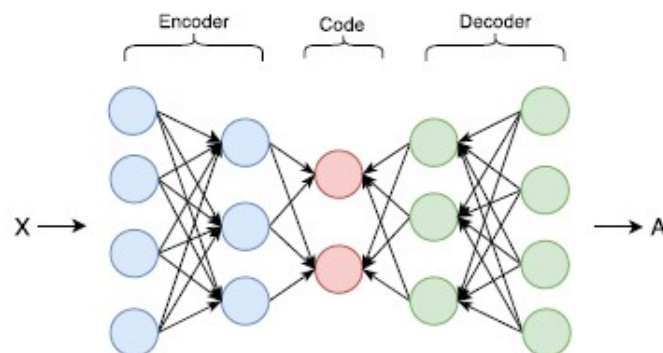
Χαρακτηριστικό του είναι ότι μπορεί να κάνει βαθιά διάγνωση στα δεδομένα, έχει την ικανότητα να παρακάμπτει τοπικά ελάχιστα και να συγκλίνει στο ολικό ελάχιστο, καλύτερη ακρίβεια πρόβλεψης. Ο μακρύτερος χρόνος εκπαίδευσης μπορεί να βελτιωθεί με την χρήση conjugate gradient method. [(Chen et al, 2019)]

Χρησιμοποιούνται για μείωση διαστάσεων χαρακτηριστικών, ταξινόμηση, παλινδρόμηση σε εφαρμογές αναγνώρισης εικόνων, ακολουθίες video, δεδομένα καταγραφής κίνησης

### 3.14.17 Αυτοκωδικοποιητής / Autoencoder (AE)

Αποτελούν ΤΝΔ μη εποπτευόμενης μάθησης που αποτελείται από τρία μέρη έναν encoder, τον κώδικα (code) και έναν decoder. Χρησιμοποιούν στην αρχιτεκτονική τους, ένα input layer, που πραγματοποιεί την κωδικοποίηση των δεδομένων εισόδων (features), και προσπαθούν να ανακατασκευάσουν τα ίδια δεδομένα με μικρότερους κόμβους (emphcode ή code), μια διαδικασία η οποία είναι γνωστή ως latent representation ή encoding. Η κωδικοποίηση αυτή αποτελεί και τα κρυφά δίκτυα (hidden layers). Ο Αποκωδικοποιητής δημιουργεί την έξοδο από τα κρυφά επίπεδα. Ο αριθμός των εισόδων και των εξόδων είναι ίδιος και έχουν συμμετρική δομή. Η μέθοδος μάθησης backpropagation χρησιμοποιείται για την αναβάθμιση των βαρών του δικτύου.

Το χαρακτηριστικό τους είναι η μείωση των διαστάσεων και η εκμάθηση γνωρισμάτων. Η εκπαίδευσή τους μπορεί να οδηγήσει σε απώλεια σημαντικών δεδομένων και να δημιουργήσει καθυστερήσεις στην χρόνο εκπαίδευσης.



Εικόνα23: Νευρωνικά Δίκτυα Auto Encoders (AE), Πηγή: [(Olorunnimbe et al,2022)]

### 3.14.18 Graph neural network (GNN)

Το Νευρωνικό Δίκτυο Γραφήματος (GNN) αποτελεί μια κατηγορία δικτύων τεχνητής νοημοσύνης βαθιάς μάθησης, που έχει σχεδιαστεί για να πραγματοποιεί αναφορές από δεδομένα που μπορούν να αναπαρασταθούν ως γραφήματα. Βασικό σημείο τους είναι η λειτουργία **pairwise message passing** κατά την οποία οι κόμβοι γραφημάτων αναβαθμίζουν επανειλημμένα τις αναπαραστάσεις τους ανταλλάσσοντας πληροφορίες με τους γείτονές τους. Μπορούν να πραγματοποιήσουν θέματα πρόβλεψης στο επίπεδο των κόμβων (node-level), στις αιχμές (edge-level) και στο γράφημα (graph level).

Η αρχιτεκτονική ενός γενικού GNN υλοποιεί τα παρακάτω βασικά επίπεδα:

- **Permutation equivariant:** δημιουργεί μια αναπαράσταση ενός γραφήματος σε μια αναβαθμισμένη αναπαράσταση του ίδιου γραφήματος
- **Local pooling:** ομαδοποίηση γραφημάτων σε τοπικό επίπεδο
- **Global pooling:** ομαδοποίηση γραφημάτων σε μικρό επίπεδο μια τροποποιημένη αναπαράσταση του ολικού γραφήματος.

Μέσω της χρήσης δεδομένων δομημένων σε γράφους έχει την δυνατότητα να ενσωματώσει τις διασυνδέσεις των χρηματοοικονομικών αγορών και να κάνει καλύτερες προβλέψεις. Επίσης χρησιμοποιούνται για την ανάλυση των κοινωνικών δικτύων.

Τα GNN αποτελούν ένα από τα δομικά στοιχεία του AlphaFold, ενός προγράμματος TN που αναπτύχθηκε από το πρόγραμμα της Google DeepMind για την επίλυση του προβλήματος περιέλιξης πρωτεϊνών.

Επέκταση των GNN αποτελεί η Graph Convolutional Network με την εφαρμογή της λειτουργίας Convolution σε δομές που αναπαρίστανται με γραφήματα.

### 3.14.19 Βαθεία Εξαναγκασμένη Μάθηση / Deep Reinforcement learning (RL)

Η μέθοδος DRL βασίζεται στην εξαναγκασμένη μάθηση Reinforcement Learning (RL) και έχει ως στόχο την επιλογή ενός συνόλου δράσεων που μεγιστοποιούν μια ανταμοιβή (reward). Το μοντέλο σε αντίθεση με την εποπτευόμενη μάθηση, μαθαίνει αλληλεπιδρώντας με το περιβάλλον μέσω ενός πράκτορα (agent) μια επαναλαμβανόμενη ακολουθία δράσεων, ποια σειρά βημάτων οδηγούν στο υψηλότερο κέρδος στο τέλος της διαδικασίας (reward). Συνεπώς δεν απαιτεί την ύπαρξη ενός

αρχικού συνόλου δεδομένων εκπαίδευσης. Η μέθοδος αποτελεί στοχαστική διαδικασία διακριτού χρόνου και σχηματοποιείται ως διαδικασία Markov Decision Process (MDP).

Στην περίπτωση του DRL, μπορεί να είναι είτε model-based όπου διατηρεί ένα μοντέλο του περιβάλλοντος που του δίνει την δυνατότητα να επιλέξει τις δράσεις που μεγιστοποιούν την ανταμοιβή είτε είναι model-free χωρίς να διατηρεί μοντέλο του περιβάλλοντος είτε να είναι συνδυασμός των δύο παραπάνω. [(Olorunnimbe et al,2022)]

### **3.15 Μηχανισμός Προσοχής /Attention Mechanism**

Στα Τεχνητά Νευρωνικά Δίκτυα ο Μηχανισμός Προσοχής / Attention Mechanism, αποτελεί μια τεχνική η οποία μιμείται την γνωστική προσοχή του εγκεφάλου, η οποία λειτουργεί εξαιρώντας κάποια μέρη των δεδομένων εισόδου εξετάζοντας άλλα πιο σημαντικά σε σχέση με το γενικό πλαίσιο.

Ως προς την πρόβλεψη ο μηχανισμός επικεντρώνεται σε περισσότερες από μιας αναπαραστάσεις χαρακτηριστικών, μέσω βαρών,για κάθε βήμα της ακολουθίας εισόδου και αποφασίζει ποια από αυτές τις αναπαραστάσεις είναι χρήσιμες (attention) για την πρόβλεψη του συγκεκριμένου στοιχείου της ακολουθίας.

### **3.16 Εκτίμηση Τεχνητών Νευρωνικών Δικτύων**

Η εκτίμηση των ΤΝΔ γίνεται βάση των χαρακτηριστικών που δίνει το αποτέλεσμα κατά την εφαρμογή του. Η ικανότητα να προσεγγίζει τα δεδομένα αντιπαρατίθεται στην πολυπλοκότητα του δικτύου, και η χρυσή τομή βρίσκεται στην εξισορρόπηση της μεροληψίας και της διακύμανσης βάσει της ελαχιστοποίησης της συνάρτησης του τετραγωνικού σφάλματος. Η εκτίμηση του κινδύνου πρόβλεψης εκφράζει την ικανότητα γενίκευσης του δικτύου. Το πρόβλημα αυτό παρουσιάζει δυσκολίες αναλόγως του προβλήματος που εξετάζεται και τα δεδομένα που χρησιμοποιούνται.

Η διαδικασία επιλογής νευρωνικού υποδείγματος γίνεται στα εξής στάδια

- Επιλογή του αρχικού υποδείγματος από μια ακολουθία μοντέλων αυξανόμενης περιπλοκότητας, επιλέγοντας τον αριθμό των κρυφών μονάδων ώστε να αντιστοιχεί στην

ελάχιστη τιμή του κινδύνου πρόβλεψης και έπειτα την αφαίρεση των πλεοναζουσών μονάδων εισόδου

- Περιλαμβάνει τον προσδιορισμό της καταλληλότητας της μεταβλητής εισόδου για το υπόδειγμα, την εκτίμηση της δειγματικής διακύμανσης της μέτρησης καταλληλότητας και τον έλεγχο της υπόθεσης ότι η μεταβλητή είναι ακατάλληλη.
- Έλεγχος των υποθέσεων σχετικά με την υποκείμενη σχέση που δημιουργήσε τις διαθέσιμες παρατηρήσεις. [(Ζαπράνης 2005)]

### **3.16.1 Υπερπαραμέτροι & Παράμετροι και Νευρωνικών Δικτύων**

Ένα μοντέλο Μηχανικής Μάθησης και συνεπώς ΤΝΔ, ορίζεται βάση των παραμέτρων του που υπολογίζονται κατά τον σχηματισμό του από την εκτέλεση των αλγορίθμων στα δεδομένα και είναι οι συντελεστές coefficients του μοντέλου.

Ωστόσο η διαδικασία της εκπαίδευσης ενός μοντέλου περιλαμβάνει την ρύθμιση των βέλτιστων **υπερπαραμέτρων**, ορολογία που αναφέρεται στις ρυθμίσεις που κάνει ο χρήστης, για την επιλογή της δομής του δικτύου και της διαδικασίας μάθησης. Η ρύθμιση αυτών των υπερπαραμέτρων, αναλόγως της πολυπλοκότητας ενός προβλήματος και των διαθέσιμων δεδομένων αποτελεί την αρχιτεκτονική του δικτύου. Το πρόβλημα της επιλογής κατάλληλης ρύθμισης των υπερπαραμέτρων νευρωνικού δικτύου αποτελεί την λύση του θέματος της εξισορρόπησης της μεροληψίας και της διακύμανσης με την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος της συνάρτησης loss function. Η βελτιστοποίηση αυτών των παραμέτρων οδηγεί στην δημιουργία μοντέλων με καλή εφαρμογή στα δεδομένα και μπορεί να γίνει και με αυτοματοποιημένο τρόπο. Αποτελεί σημαντικό κεφάλαιο της Βαθιάς Μάθησης.

Κάποιες από τις τεχνικές που χρησιμοποιούνται για την βελτιστοποίηση των υπερπαραμέτρων είναι:

1. Manual Search, αποτελεί την ρύθμιση των παραμέτρων από τον χρήστη μέσω δοκιμών
2. Grid Search, αποτελεί ρύθμιση παραμέτρων δοκιμάζοντας όλους τους συνδυασμούς των μεταβλητών, αυξάνοντας την πολυπλοκότητα των υπολογισμών.



3. Randomized Search, αποτελεί τυχαία επιλογή των υπερπαραμέτρων
4. Halving Grid Search, αποτελεί μια βελτιστοποιημένη μέθοδο του Grid Search χρησιμοποιώντας μια επιτυχημένη halving προσέγγιση.
5. Halving Randomized Search, αποτελεί εξέλιξη της προηγούμενης τεχνικής με τυχαία επιλογή συνόλου υπερπαραμέτρων.
6. HyperOpt-Sklearn, αποτελεί μέθοδο βελτιστοποίησης Bayes
7. Bayes Search, χρησιμοποιεί την τεχνική βελτιστοποίησης Bayes

### **3.16.1.1 Υπερπαραμέτροι Δομής Δικτύου**

#### **3.16.1.1.1 Αριθμός Κρυφών Επιπέδων και Μονάδων**

Τα κρυφά επίπεδα βρίσκονται μεταξύ των επιπέδων εισόδου και των επιπέδων εξόδου. Ένα μοντέλο με περισσότερα επίπεδα και περισσότερες κρυφές μονάδες έχει καλύτερη εμβέλεια αναπαράστασης και είναι ικανό να αναπαριστά πιο πολύπλοκες συναρτήσεις. Η εμβέλεια ενός μοντέλου εξαρτάται από τον αριθμό των κρυφών επιπέδων (βάθος) οδηγεί και τον αριθμό των κόμβων στα κρυφά επίπεδα. Ένα μοντέλο με περισσότερα επίπεδα και περισσότερες κρυφές μονάδες ανά επίπεδο είναι ικανό να αναπαραστήσει πιο πολύπλοκες συναρτήσεις. Ένα νευρωνικό δίκτυο με ένα επίπεδο και πολλούς κόμβους μπορεί να μάθει να προσεγγίζει κάθε συνάρτηση που αναπαριστά τα δεδομένα. Αυξάνοντας το βάθος του δικτύου αυξάνουμε την εμβέλεια του δικτύου. Ένα μοντέλο με πολλά επίπεδα μπορεί να είναι περισσότερο αποδοτικό από την εκπαίδευση ένα μοντέλο με ένα επίπεδο με πολλούς κόμβους. [(Goddfellow et al 2022a)]

#### **3.16.1.1.2 Απόρριψη / Dropout**

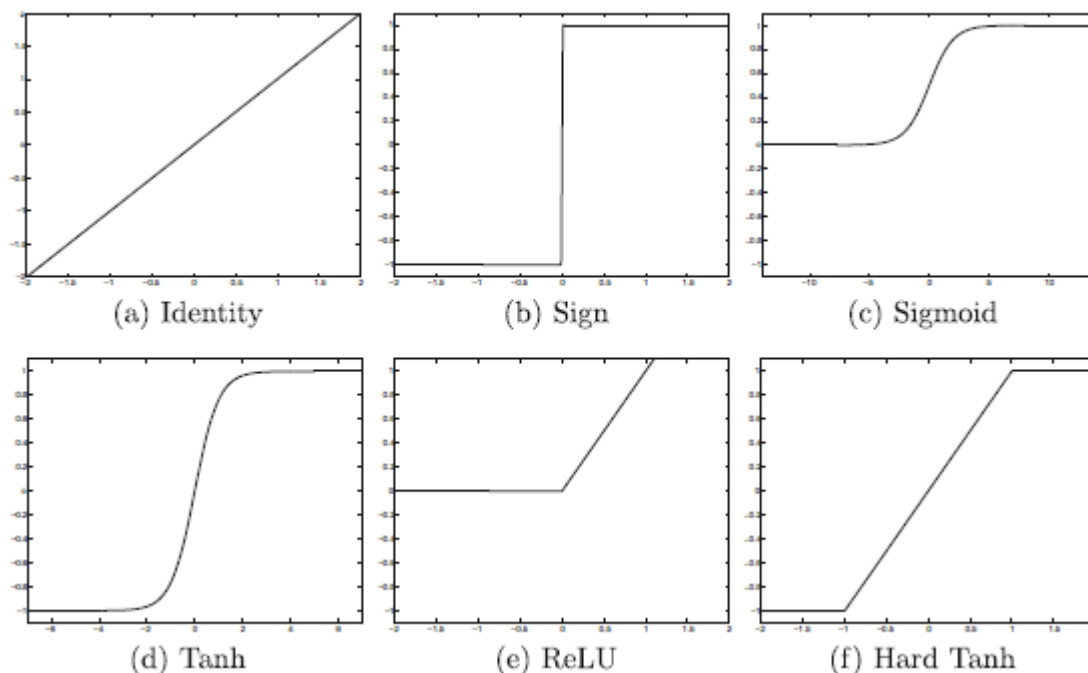
Το πρόβλημα της υπερεφαρμογής μπορεί να αντιμετωπιστεί με την χρήση ενός επιπέδου νευρώνων το οποίο σε κάποια σημεία του τυχαία αποκόπτονται από το συνολικό δίκτυο κατά την διαδικασία της εκπαίδευσης. Η διαδικασία αυτή λέγεται Απόρριψη/Dropout. Κατά την διαδικασία αυτή αποκόπτονται οι νευρώνες μαζί με τις συνάψεις τους. Η εκπαίδευση γίνεται με το εναπομείναν δίκτυο. Στην συνέχεια οι κόμβοι επανεισάγονται στο δίκτυο με τις αρχικές τιμές των κόμβων τους. Με αυτό τον τρόπο γίνεται αποφυγή του overfitting και αύξηση της ακρίβειας επικύρωσης.

### 3.16.1.1.3 Συνάρτηση Ενεργοποίησης/Activation function

Όπως έχει αναφερθεί ο κάθε νευρώνας έχει μια συνάρτηση ενεργοποίησης, οι πιο κοινές των οποίων είναι η σιγμοειδής/sigmoid για δυαδικές προβλέψεις, η υπερβολική/hyperbolic, η εφαπτόμενη/tangent, η Rectified Linear Unit (ReLU) (βοηθάει στην αποφυγή του vanishing gradients problem δηλαδή την αδυναμία μάθησης), η leaky-ReLU, η swish, η softmax για προβλέψεις πολλών κατηγοριών με διαφορετικά χαρακτηριστικά η κάθε μία.

Οι συναρτήσεις ενεργοποίησης χρησιμοποιούνται για να εισάγουν μη γραμμικότητα στα μοντέλα το οποίο καθιστά δυνατή την εκμάθηση μη γραμμικών συνόρων στα δεδομένα.

Μια πρακτική στα νευρωνικά δίκτυα εμπρόσθιας διάδοσης είναι να έχουν ένα ή περισσότερα κρυφά επίπεδα από sigmoid συναρτήσεις και ένα επίπεδο από γραμμικές συναρτήσεις.



Εικόνα 24: Συναρτήσεις Ενεργοποίησης[(Aggarwal 2022)]

### 3.16.1.1.4 Αρχικοποίηση Βαρών Δικτύου / Network Weight Initialization

Οι κόμβοι του δικτύου αποτελούνται από παραμέτρους οι οποίοι αναφέρονται ως βάρη/weight. Η χρήση του αλγορίθμου βελτιστοποίησης όπως η gradient descent μεταβάλλει αυτά τα βάρη ώστε να μειωθεί η συνάρτηση απώλειας ή loss function κατά την διάρκεια της εκπαίδευσης. Τα βάρη αυτά χρειάζονται μια αρχική τιμή από την οποία θα ξεκινήσει η διαδικασία βελτιστοποίησης. Η αρχική τιμή των βαρών μπορεί να καθορίσει αν ο αλγόριθμος θα προσεγγίζει τα δεδομένα ή θα συναντήσει

δυσκολίες στην εξέλιξη του στα δεδομένα. [(Goddfellow et al 2022a)] Η αρχικοποίηση των βαρών γίνεται με μια Κανονική Κατανομή ή Κατανομή Gauss, η οποία έχει σημαντική επίδραση στην διαδικασία βελτιστοποίησης και στην ικανότητα γενίκευσης του δικτύου. [(Goddfellow et al 2022a)]

### **3.16.1.2 Υπερπαράμετροι Μάθησης Δικτύου**

#### **3.16.1.2.1 Ρυθμός Μάθησης/ Learning rate**

Η μάθηση συμβαίνει κατά την διάρκεια της εκπαίδευσης με την χρήση στοχαστικού gradient descent αλγόριθμου με την χρήση της μεθόδου backpropagation. Με την μάθηση αναπροσαρμόζονται τα βάρη των κόμβων του δικτύου βάση κάποιου κανόνα αναπροσαρμογής για την ελαχιστοποίηση της συνάρτησης σφάλματος loss function. Ο ρυθμός με τον οποίον τα βάρη αναπροσαρμόζονται κατά την εκπαίδευση ονομάζεται ως ρυθμός μάθησης “ $\eta$ ”. Μεγάλος ρυθμός μάθησης, μπορεί να οδηγήσει στην αύξηση παρά την μείωση του σφάλματος εκπαίδευσης ενώ για μικρό ρυθμό μάθησης η εκπαίδευση είναι πιο αργή και μπορεί να κολλήσει με υψηλό σφάλμα μάθησης. Ο ρυθμός μάθησης δεν μπορεί να οριστεί a priori. Ο ρυθμός μάθησης βρίσκεται μεταξύ  $[0,1]$ . Για την επιλογή της ιδανικής συχνότητα μάθησης μπορεί να γίνει μια ανάλυση ευαισθησίας για την συχνότητα μάθησης η οποία ονομάζεται grid search για τις τιμές από 0.1 έως  $10^{-5}$ . [(Goddfellow et al 2022a)]

#### **3.16.1.2.2 Αριθμός Εποχών/ Number of epochs**

Η εκπαίδευση ενός Νευρωνικού Δικτύου ορίζεται είτε βάση epoch/εποχής είτε βάσει του ελαχίστου σφάλματος. Ως epoch/εποχή ορίζεται η εκπαίδευση του ΤΝΔ με όλα τα διαθέσιμα δεδομένα εκπαίδευσης για έναν κύκλο. Καθορίζει πότε η εκπαίδευση θα σταματήσει όταν ο αριθμός των επαναλήψεων / iterations ξεπεράσει τον αριθμό των εποχών/epochs. Η εκπαίδευση βάση ελαχίστου μέσου σφάλματος αυτό αναπαριστά τον μέγιστο αριθμό από επαναλήψεις μέχρι να μειωθεί το ελάχιστο μέσο σφάλμα.

### 3.16.1.2.3 Δείγμα εκπαίδευσης/ Batch

Ως Batch ορίζεται ο αριθμός των δειγμάτων εκπαίδευσης πριν την αναβάθμιση του εσωτερικού μοντέλου των παραμέτρων δηλαδή σε μια επανάληψη. Ένα σύνολο δεδομένων μπορεί να χωριστεί σε πολλά Batch. Τυπικά μεγέθη batch είναι 32, 64, 128.

### 3.16.1.2.4 Momentum

Αποτελεί παράμετρο στον Gradient Descent Optimization Algorithm που επιτρέπει την επιτάχυνση της διαδικασίας βελτιστοποίησης. Αποτρέπει το σύστημα να συγκλίνει σε ένα τοπικό ελάχιστο ή ένα διάσελο (saddle point). Αποτελεί επίσης υπερπαράμετρο που ελέγχει το ποσό της ιστορίας (momentum) για να κάνει ένα βήμα σε ένα νέο σημείο στον χώρο αναζήτησης. [(Brownlee 2022)]Μια μεγάλη τιμή Momentum βοηθάει να αυξηθεί η ταχύτητα σύγκλισης στο σύστημα με κίνδυνο αποσταθεροποίησης του δικτύου. Μικρή τιμή Momentum συνεπάγεται αποφυγή των τοπικών ελαχίστων αλλά και αργή εκπαίδευση. Η τιμή του βρίσκεται μεταξύ  $[0,1]$  και η συνήθη τιμή του είναι από 0,8 έως 0,99.

## 3.17 Κριτική Τεχνητών Νευρωνικών Δικτύων

### Πλεονεκτήματα Τεχνητών Νευρωνικών Δικτύων

- **Γενίκευση/Generalization**, αφορά την ικανότητα των νευρωνικών δικτύων να σχηματίζουν μοντέλα και έπειτα να τα εφαρμόζουν σε νεά δεδομένα
- **Ανοχή σε σφάλματα/ Fault tolerance**, τα παραδοσιακά συστήματα υπολογιστών είναι ευαίσθητα σε σφάλματα, ενώ τα TNN δεν επηρεάζονται από τα σφάλματα
- **Επαγωγή/Adaptation**, τα νευρωνικά δίκτυα μπορούν να μάθουν και επάγονται σε διαφορετικά περιβάλλοντα χωρίς να χρειάζονται να επανεκπαιδευθούν.
- **Παράλληλη κατανεμημένη επεξεργασία**, όλες οι μονάδες του νευρωνικού δικτύου εκτελούνται ταυτόχρονα και έχουν καλύτερη ταχύτητα επεξεργασίας.
- **Δεν χρειάζεται υπόθεση στα δεδομένα**, ως δεδομένα εισόδου μπορούν να είναι οποιαδήποτε δεδομένα.
- Ικανοποιητική ακρίβεια πρόβλεψης
- Επαγωγή σε αλλαγές στις τιμές των εισόδων

- **Ευελξία στην επεξεργασία δεδομένων**, μπορούν να λειτουργήσουν σε αδόμητα δεδομένα, με ελλιπείς τιμές, να ανεχθούν θόρυβο, χαοτική συμπεριφορά να ξεκινήσουν απο μια τυχαία αρχή
- **Ευρωστία/Robustness**, η οποία αφορά στην ικανότητα του TNN να ανταποκρίνεται σε ελλιπείς και εσφαλμένες εισόδους
- Ικανότητα να προσεγγίσουν **πολύπλοκες συσχετίσεις** των δεδομένων, μη γραμμικές επιδράσεις και να περιλάβουν διαφορετικούς τύπους πληροφοριών πέραν των αριθμητικών

### **Μειονεκτήματα Τεχνητών Νευρωνικών Δικτύων**

- **Αδυναμία πρόβλεψης αποτελεσμάτων με ακρίβεια**, η μέθοδος μπορεί να λειτουργήσει ανατιολόγητα και τα αποτελέσματα να είναι ασυσχέτιστα [(Yildiz B.,2010)] ειδικά στην περίπτωση των χρηματοοικονομικών χρονοσειρών μεγάλων διαστάσεων με θόρυβο [(Tang Y. et al, 2022)]
- **Σύγκλιση σε τοπικά ελάχιστα**
- Κάποιες φορές τα TNN **δεν μπορούν να εκπαιδευτούν**[(Yildiz B.,2010)], π.χ. λόγω του ότι ο αλγόριθμος οπισθοδιάδοσης έχει περιέλθει σε στάσιμη κατάσταση [(Ζαπράνης, 2005)]
- Η χρήση των TNΔ χρειάζεται **πρακτική εμπειρία στην ρύθμιση των υπερπαραμέτρων** ώστε να δημιουργηθούν τα κατάλληλα μοντέλα, ώστε να αποφεύγεται το overfitting και να επιλύεται το δίλημμα “bias-variance”.
- **Δυσκολία διερμηνείας, επεξήγησης και αιτιολόγησης των αποτελεσμάτων μοντέλων** που δημιουργούνται, λόγω του ότι είναι δύσκολο να ελεγχθούν όλες οι αλληλεπιδράσεις που έχουν κάνει οι νευρώνες πρόβλημα που καλείται “black box”. Δεν υπάρχει δυνατότητα άμεσων συμπερασμάτων. σε αντιδιαστολή με τις στατιστικές μεθόδους που αιτιολογούνται πλήρως
- **Η αιτιολόγηση των αποτελεσμάτων προϋποθέτει εξειδίκευση** στην τεχνολογία των TNΔ
- Τα TNΔ παρουσιάζουν **μεγαλύτερη υπολογιστική πομπυλοκότητα** από άλλες μεθόδους μηχανικής μάθησης.
- Μεγάλος χρόνος εκπαίδευσης

### **3.18 Υβριδικά-Συνδυαστικά μοντέλα**

Για να βελτιωθεί η ανάλυση σε ένα πρόβλημα που επιλύεται με μεθόδους TN, ένας υβριδικός μηχανισμός μπορεί να χρησιμοποιηθεί συνδυάζοντας την μέθοδο TN, είτε με μια ακόμα μέθοδο

μάθησης, είτε με κάποιες από τις παλιότερες μεθόδους ανάλυσης όπως είναι οι στατιστικές μέθοδοι.

Τέτοια παραδείγματα είναι π.χ. η χρήση ενός ΤΝΔ με έναν γενετικό αλγόριθμο ΓΑ, όπου ο γενετικός αλγόριθμος χρησιμοποιείται για τα χρονικά χαρακτηριστικά των χρονοσειρών, η χρήση Empirical Mode Decomposition (EMD) και ΤΝΔ για την πρόβλεψη τιμής δείκτη, χρήση CNN και LSTM.

Τα αποτελέσματα της χρήσης τους είναι καλύτερη εφαρμογή του μοντέλου, ικανότητα γενίκευσης, Σε μελέτη που έγινε σχετικά με την ικανότητα πρόβλεψης του επόμενου σημείου (one-step prediction) χρονοσειρών αποδείχτηκε ότι η χρήση στατιστικού γραμμικού μοντέλου σε συνδυασμό με την χρήση μη γραμμικού μοντέλου MM είναι αποτελεσματική για την πρόβλεψη της μελλοντικής τάσης της ακολουθίας δεδομένων. [(Tang et al, 2022)]

### **3.18.1 Neuro Fuzzy Συστήματα**

Neuro Fuzzy σύστημα είναι το σύστημα που χρησιμοποιεί ασαφή fuzzy λογική και έναν αλγόριθμο μάθησης που προκύπτει από την θεωρία νευρωνικών δικτύων για τον ορισμό των παραμέτρων του μοντέλου (fuzzy sets και fuzzy rules) από την επεξεργασία των δεδομένων. Ο συνδυασμός Fuzzy Logic και Νευρωνικών Δικτύων αποτελεί πεδίο της ΤΝ με την δημιουργία υβριδικού συστήματος.

### **3.18.2 Modular Neural Networks**

Τα Modular Neural Networks περιέχουν πολλαπλά νευρωνικά δίκτυα τα οποία εργάζονται παράλληλα το ένα από το άλλο και δεν αλληλεπιδρούν μεταξύ τους κατά την διαδικασία της ανάλυσης. Το καθένα είναι υπεύθυνο για ένα συγκεκριμένο μέρος των υπολογισμών.

## **3.19 Text Mining**

Εξόρυξη Κειμένου ή Text Mining, αποτελεί την ανάλυση κειμένου με μεθόδους μηχανικής μάθησης με σκοπό την ανίχνευση γνώσης και συμπερασμάτων. Η πηγή του κειμένου μπορεί να είναι ειδήσεις, άρθρα, βιβλία, έρευνες, όπως και δημοσιεύσεις στο διαδίκτυο όπως ιστοσελίδες και δημοσιεύσεις στα μέσα κοινωνικής δικτύωσης, blogs και άλλα. Κάποια κοινωνικά, πολιτικά και οικονομικά γεγονότα αποτυπώνονται στις ειδήσεις, στο διαδίκτυο, σε μέσα κοινωνικής δικτύωσης

Μέθοδοι που χρησιμοποιούνται είναι η επεξεργασία φυσικής γλώσσας, η ανάλυση συναισθήματος/sentiment analysis, η ανάκτηση δεδομένων.

Η τεχνική χρησιμοποιείται και για πρόβλεψη μεγεθών χρηματοοικονομικών χρονοσειρών. Οι εφαρμογές περιλαμβάνουν την εξαγωγή συμπερασμάτων από ειδήσεις, άρθρα για την πραγματοποίηση πρόβλεψης..Άλλες εφαρμογές είναι το θέμα της αναγνώρισης αντιδράσεων από την αγορά σε ειδήσεις και η ανάλυση των ειδήσεων για την πρόβλεψη αποδόσεων σε χρηματοοικονομικά μεγέθη.

Ως **Word Embedding** αναφέρεται η μέθοδος κατά την οποία στο διάνυσμα εισόδου σε έναν αλγόριθμο μάθησης μπορούν να ενσωματωθούν λέξεις από κείμενο αυξάνοντας την εφαρμογή των μοντέλων. Η διαδικασία αυτή πραγματοποιείται μέσω των μοντέλων νευρωνικών δικτύων word2vec και του GloVe(Global Vector for Word representation)

Κάποιες έρευνες επικεντρώνονται σε επεξεργασία φυσικής γλώσσας με χρήση μεθόδων Βαθιάς Μάθησης. Στην βιβλιογραφική έρευνα των [(Farimani et al. 2022)] εξετάζονται δημοσιεύσεις στον τομέα της επεξεργασίας φυσικής γλώσσας και των συμπεριφοριστικών χρηματοοικονομικών συγκριτικά με την ανάλυση δεδομένων αγοράς για υποστήριξη χρηματοοικονομικών αποφάσεων. Εξετάζεται η ανάκτηση πληροφορίας από διάφορες πηγές όπως νέα και κοινωνικά δίκτυα και η επίδρασή τους στις χρηματοοικονομικές αγορές, αναλύοντας την συμπεριφορά των επενδυτών και εξετάζοντας τα συστήματα που προτείνουν τις στρατηγικές εμπορίου.

### **3.20 Ανάλυση Συναισθήματος / Sentiment Analysis**

Η Sentiment Analysis/Ανάλυση Συναισθήματος, ή διαφορετικά Opinion Mining/ Εξόρυξη Γνώμης, χρησιμοποιεί την μέθοδο της Επεξεργασίας Φυσικής Γλώσσας/Natural Language Processing (NLP) η οποία περιλαμβάνει συντακτική και σημασιολογική ανάλυση για να κατανοήσει την δομή και το νόημα ενός κειμένου και της Μηχανικής Μάθησης για να το κατηγοριοποιήσει σε μια από τις κατηγορίες που έχουν οριστεί όπως για παράδειγμα “πολύ αρνητικό, αρνητικό, ουδέτερο, θετικό, πολύ θετικό” (sentiment factor). Το συναίσθημα ενός μηνύματος ή ενός κειμένου μπορεί να αναλυθεί με διαφορετικούς τρόπους όπως είναι: Ανάλυση Πόλωσης/ Polarized analysis,

Αναγνώριση Συναισθήματος/ Emotion recognition, Πλευρική Ανάλυση/Aspect-based, Ανάλυση Προθέσης/ Intent analysis. Πηγές κειμένων αποτελούν δημοσιεύσεις από social media,ειδήσεις, χρηματοοικονομικές ειδήσεις, blogs, rss feeds, μπορούν να χρησιμοποιηθούν περαιτέρω για πρόβλεψη.

Πρακτικά η έρευνα αποδεικνύει ότι οι κινήσεις των μετοχών σχετίζονται με το δημόσιο συναίσθημα απέναντι στις επιχειρήσεις. Η χρήση ανάλυσης συναισθήματος στα μοντέλα πρόβλεψης αυξάνει την ακρίβεια πρόβλεψης κατά 20%.Συνεπώς το συναίσθημα για μια εταιρεία στα μέσα μαζικής ενημέρωσης, στα μέσα κοινωνικής δικτύωσης, ή απόψεις επενδυτών, πελατών,προμηθευτών, συμβούλων, εργαζομένων μπορούν να αποτελέσουν ένα εργαλείο για το πώς θα κινηθούν οι τιμές των μετοχών των εταιρειών.[(Yilmaz, 2022)] Οι μέθοδοι πρόβλεψης χρηματοοικονομικών μεγεθών ενσωματώνουν στα μοντέλα τους δεδομένα κειμένου από διάφορες πηγές ώστε να εκπαιδεύσουν τους αλγορίθμους τους.

Στην μελέτη τους για χρήση sentiment ανάλυσης για πολυδιάστατες χρηματοοικονομικές χρονοσειρές οι [(Liapis et al, 2021)] διαπιστώνουν ότι μετά τη χρήση του συναισθήματος διαπιστώνεται μια βελτίωση της απόδοσης υπό όρους σε περιπτώσεις όπου οι μέθοδοι εφαρμόστηκαν για την πρόβλεψη μακροπρόθεσμων χρονικών πλαισίων.

Η μέθοδος χρησιμοποιείται με παραλλαγή σε άλλες εφαρμογές όπως είναι η εύρεση του Market Sentiment Indicator, για να ορίσει πώς μια ομάδα αισθάνεται για μια αγορά ή την οικονομία ή τι στάση έχουν οι επενδυτές απέναντι σε μια συγκεκριμένη χρηματιστηριακή αξία ή χρηματοοικονομική αγορά. Τετοια παραδείγματα δεικτών Sentiment είναι : CBOE Volatility Index (VIX), New York Exchange (NYSE) High/Low Indicator, NYSE 200-day Moving Average, Odd-Lot, Trading Statistics, The Commitment of Traders Report.[(Mitchell, 2022)]

Στην μελέτη του [(Jiang, 2021)] αναφέρεται η χρήση των CNN, Natural Language Toolkit (NLTK) για sentiment analysis με χρήση επιπλέον ανάλυσης του Harvard IV-4 ψυχολογικού λεξικού.



## 4 Βιβλιογραφική έρευνα

Η ένταση έρευνας και ανάπτυξης, εφαρμογών, επενδύσεων, στην Τεχνητή Νοημοσύνη δημιουργεί την ανάγκη να καθοριστεί ένα πλαίσιο χαρακτηριστικών, καλών πρακτικών και συμπερασμάτων. Το πλαίσιο αυτό διαφοροποιείται ανά τομέα εφαρμογής καθώς οι εφαρμογές θέτουν ζητήματα τεχνικής φύσεως των μοντέλων, της καταλληλότητας των μεθόδων, της αποτελεσματικότητας των πρακτικών. Ειδικότερα σε περιπτώσεις χρηματοοικονομικών αποφάσεων το ενδεχόμενο κόστος λανθασμένης χρήσης ενός μοντέλου TN είναι σημαντικό. Τα ζητήματα αυτά που συναντάνε όσοι αναπτύσσουν συστήματα TN αφορούν:

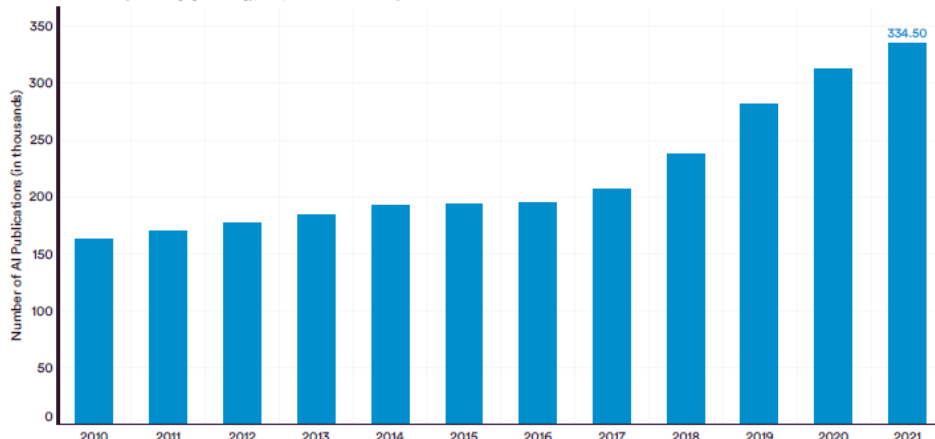
- 1) Την εύρεση κατάλληλων χρηματοοικονομικών δεδομένων για την δημιουργία των μοντέλων, την προετοιμασία τους και την επιλογή των κατάλληλων μεταβλητών εισόδου προς χρήση στα μοντέλα.
- 2) Την επιλογή των κατάλληλων μεθόδων-αλγορίθμων για την εφαρμογή πρόβλεψης στο θέμα ενδιαφέροντος, την ρύθμιση των υπερπαραμέτρων των μοντέλων ώστε να δημιουργούνται ιδανικά και αξιόπιστα μοντέλα με ικανότητα γενίκευσης και ικανότητα πρόβλεψης.
- 3) Την αξιολόγηση των μοντέλων, την ερμηνεία και την αιτιολόγηση των αποτελεσμάτων τους, την επανεκτίμηση των δεδομένων και των μοντέλων, των μεταβλητών εισόδου, των υπερπαραμέτρων ώστε να βρεθούν τα τα χαρακτηριστικά των επιτυχημένων μοντέλων
- 4) Τον έλεγχο της απόδοσης του μοντέλου σε παλιότερα δεδομένα (backtesting).

Εκτός από το θεωρητικό πλαίσιο των μεθόδων TN, οι ερευνητές προστρέχουν στις δημοσιευμένες μελέτες, για παρόμοιες εφαρμογές, για να συμπεράνουν και να αποφασίσουν την μέθοδο που ταιριάζει καλύτερα στις ανάγκες και στις απαιτήσεις του θέματος ενδιαφέροντος. Η ανάπτυξη των εφαρμογών TN, ανατροφοδοτείται από έναν μεγάλο αριθμός από έρευνες, των οποίων το περιεχόμενο δημοσιεύεται από τους ερευνητές. Η έρευνα αυτή των δημοσιεύσεων των εφαρμογών και των ανασκοπήσεων (reviews) των ερευνών TN είναι έντονη ειδικότερα τα τελευταία χρόνια. Επιπλέον τα μοντέλα BM, έχουν τις περισσότερες αναφορές την τελευταία πενταετία, αλλάζοντας το σκηνικό στην TN από τις παλιές μεθόδους MM.

Αναδεικνύεται το ζήτημα της αξιολόγησης των μοντέλων με συστηματικό τρόπο που να καλύπτει όλο το πλαίσιο της εφαρμογής μοντέλων TN από την επιλογή του δείγματος δεδομένων, την δημιουργία του μοντέλου ως την επικύρωση στα δεδομένα.

**NUMBER of AI PUBLICATIONS in the WORLD, 2010–21**

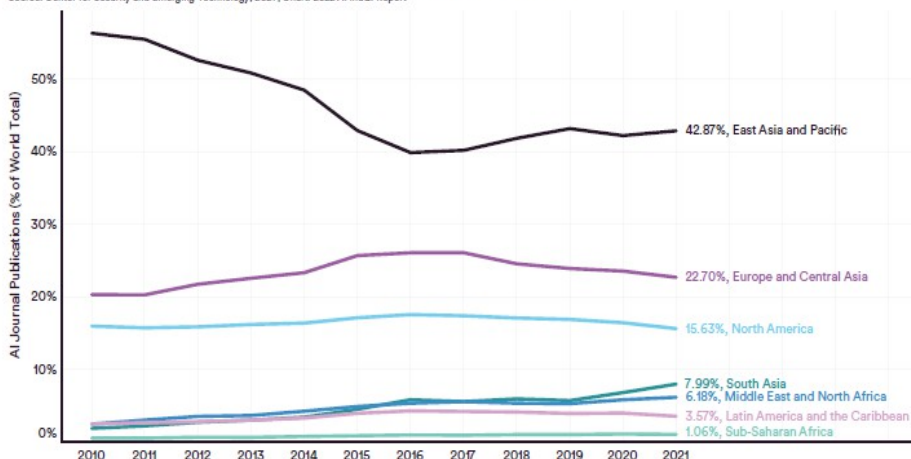
Source: Center for Security and Emerging Technology, 2021 | Chart: 2022 AI Index Report



**Εικόνα25: ο αριθμός των δημοσιεύσεων στην ΤΝ [(Zhang et al, 2022)]**

**AI JOURNAL PUBLICATIONS (% of WORLD TOTAL) by REGION, 2010–21**

Source: Center for Security and Emerging Technology, 2021 | Chart: 2022 AI Index Report



**Εικόνα26: ποσοστό του αριθμού δημοσιεύσεων στην ΤΝ ανά περιοχή [(Zhang et al, 2022)]**

## 4.1 Μεθοδολογία

Από την έρευνα των μελετών ΤΝ για την ανάλυση χρηματοοικονομικών, διαπιστώθηκε ότι αυτές χωρίζονται βάσει δύο κυριων χαρακτηριστικών: Των χαρακτηριστικών της μελέτης και των χαρακτηριστικών των μοντέλων.

### 1. Χαρακτηριστικά της μελέτης

1. Περιοχή πρόβλεψης
2. Μέγεθος πρόβλεψης
3. Η αγορά του μεγέθους
4. Δεδομένα
  1. Δεδομένα εισόδου (input variables)
  2. Μεταβλητή Πρόβλεψης

3. Μέγεθος δείγματος
  5. Φύση της μελέτης
  6. Περιβάλλον Ανάπτυξης/Development Environment
  7. Χρονικός ορίζοντας πρόβλεψης
- 2. Χαρακτηριστικά μοντέλο, περιλαμβάνουν στοιχεία που αποτελούν το μοντέλο και στοιχεία εκτίμησης του**
1. Στατιστικές Τεχνικές
  2. Προεπεξεργασία δεδομένων
  3. Μοντέλο και Τεχνική Τεχνητής Νοημοσύνης
    1. Μηχανική Μάθηση
    2. Τεχνητή Νοημοσύνης
    3. Deep Learning
  4. Συνδυαστικές Μέθοδοι
  5. Υπερπαράμετροι
  6. Παράμετροι
  7. Ανάλυση Σημαντικότητας

## **4.2 Περιπτώσεις Ανασκοπήσεων / Case Reviews**

Στην ανασκόπηση των [(Ozbayoglu et al , 2022)], σε δημοσιεύσεις διαπιστώνεται το ενδιαφέρον της ακαδημίας για χρήση μεθόδων BM και MM σε χρηματοοικονομικές εφαρμογές όπως το Financial Text Mining, Algorithmic trading, Risk Assessment, Sentiment Analysis, Portfolio Management, Fraud detection. Αναφέρεται ότι το πεδίο την έρευνας εξελίσσεται σε νέες περιοχές εφαρμογών όπως κρυπτονομίσματα, Blockchain, Behavioral Finance, και παράγωγα. Οι εφαρμογές BM έχουν καλύτερη απόδοση από τις κλασικές εφαρμογές MM ωστόσο σε κάποιες εφαρμογές διαπιστώθηκε η καλύτερη απόδοση των μοντέλων MM. Τα μοντέλα MM έχουν περιορισμούς που αφορούν την ακολουθία των εξαρτώμενων χαρακτηριστικών των χρηματοοικονομικών χρονοσειρών. Τα συχνότερα μοντέλα που χρησιμοποιούνται είναι η ομάδα των RNN και ειδικότερα τα LSTM για πρόβλεψη χρηματοοικονομικών χρονοσειρών, και ακολουθούν τα CNN και τα DMLP ειδικότερα για εφαρμογές ταξινόμησης ή τάσης τιμής.

Στον τομέα του Algorithmic Trading, για την πρόβλεψη τιμών μετοχών και δεικτών τα μοντέλα που χρησιμοποιούνται περισσότερο είναι τα LSTM, επιλογή μετοχών σε κάποια μελέτη αναφέρεται

με την χρήση CNN με χρήση δεδομένων, με δείκτες μικροδομών αγοράς τεχνικούς δείκτες. Άλλες μέθοδοι που αναφέρονται είναι RNN, GRU, CNN DMLP, DRL. Άλλες μέθοδοι που αναφέρονται είναι η χρήση μεθόδων RL για αλγοριθμική εμπορία, η χρήση text mining για εξαγωγή πληροφοριών από “tweets” και χρηματοοικονομικές ειδήσεις και χρήση LSTM, RNN, GRU για την δημιουργία σημάτων εμπορίας

Για θέματα κρυπτονομισμάτων επίσης αναφέρονται οι αλγόριθμοι LSTM, RNN, DMLP.

Αρκετές μελέτες αναφέρουν την χρήση των μοντέλων CNN για την προσέγγιση των μεγεθών χρονοσειρών με μετατροπή των δεδομένων χρονοσειρών σε διδιάστατες εικόνες χρησιμοποιώντας τεχνική ανάλυση. Παρόμοια προσέγγιση αναφέρεται η μετατροπή εικόνων Candlestick chart graphs σε διδιάστατες εικόνες με χρήση AE. Επίσης μια πρωτότυπη μέθοδος αναφέρεται η χρήση 100 εγγραφών από το limit order book για την δημιουργία 2διάστατων εικόνων και χρήση CNN. Ακόμα μια μέθοδος χρησιμοποίησε το limit order book με RNN για την πρόβλεψη μελλοντικού γεγονότος μετακίνησης τιμής.

. Η μελλοντική κατεύθυνση της έρευνας στην BM αφορά σε υβριδικά μοντέλα, βασιζόμενα σε Spatio-temporal αναπαραστάσεις δεδομένων και μοντέλα που βασίζονται στο Text Mining, NLP, semantics, sentiment analysis. Το θέμα της πρόβλεψης είναι η προβολή της τιμής και η πρόβλεψη της τάσης η οποία παρουσιάζει και μεγαλύτερο επιστημονικό ενδιαφέρον.

Στην έρευνα των [(Tang et al, 2022)] που περιλαμβάνει δημοσιεύσεις από το 2011 έως το 2021, διαπιστώνεται η αποτελεσματικότητα της χρήση μοντέλων μηχανικής μάθησης για πρόβλεψη χρηματοοικονομικών χρονοσειρών. Γίνεται σύγκριση των μοντέλων TNA, SVM, και RF τα οποία ως μη παραμετρικά, μη γραμμικά, μπορούν να διαχειριστούν τα μη γραμμικά, μη στατικά χαρακτηριστικά των χρηματοοικονομικών χρονοσειρών, όμως υστερούν στο ότι δεν μπορούν να διαχειριστούν τα ακολουθιακά χαρακτηριστικά των χρονοσειρών να ελιχθούν σε πολύ μεγάλα δεδομένα (Big Data), υψηλών διαστάσεων. Ειδικότερα δεν μπορούν έχουν καλή απόδοση μάθησης και πρόβλεψης και γενίκευσης σε πολύπλοκη συμπεριφορά, ακανόνιστη συμπεριφορά της αγοράς, απροσδόκητα γεγονότα στις αγορές. Για την βελτίωση της εφαρμογής των μεθόδων Μηχανικής Μάθησης η χρήση υβριδικών μοντέλων μπορεί να βελτιώσει τα μοντέλα καλύπτοντας τα θέματα της μή γραμμικότητας, αποφυγή υπερεφαρμογής, αποφυγή θορύβου στην εφαρμογή του μοντέλου. Τα υβριδικά μοντέλα αφορούν χρήση SVM και μοντέλων TNA με μεθόδους ARIMA, Γενετικών Αλγορίθμων, κρυφό μοντέλο Markov, Empirical Mode Decomposition, SVR και Self Organizing Map.

Επίσης αναφέρεται το θέμα των στόχων πρόβλεψης τάσης με χρονικούς ορίζοντες μακροπρόθεσμους- μηνιαίους και ενδιάμεσους-εβδομαδιαίους αντί των βραχυπρόθεσμων, καθώς είναι δύσκολο να βρεθούν οι μεταβλητές που επηρεάζουν την πρόβλεψη.

Σχετικά με το μέγεθος του δείγματος δεδομένων, υπάρχει μια διακύμανση μεταξύ της επάρκειας και της απόδοσης καθώς ένα μικρό δείγμα δεδομένων δεν είναι επαρκές για τον έλεγχο της αποτελεσματικότητάς του και της υπερεφαρμογής ενώ ένα μεγάλο περικλείει τον κίνδυνο να παρέχει αποτελέσματα εκτός των τωρινών τάσεων της αγοράς. Οι περισσότερες έρευνες υιοθετούν ένα μεγάλο σύνολο δεδομένων.

Το συμπέρασμα της έρευνας είναι ότι τα μοντέλα BM όπως είναι ειδικότερα το LSTM και ακολούθως το CNN αλλά και υβριδικών μοντέλων έχουν καλή απόκριση για πρόβλεψη χρονοσειρών, αποφεύγουν το πρόβλημα του overfitting. Το μοντέλο LSTM έχει την μεγαλύτερη εφαρμογή είτε μεμονωμένα είτε ως μέρος ενός υβριδικού μοντέλου. Η πρόβλεψη γίνεται για μακροπρόθεσμους χρόνους. Οι λόγοι που αναφέρονται για την επιτυχία του μοντέλου στην πρόβλεψη των χρηματοοικονομικών χρονοσειρών είναι η ισχυρή ικανότητα μάθησης του μοντέλου, η δυνατότητα να ενσωματώνει τα ακολουθιακά μακροπρόθεσμα και βραχυπρόθεσμα χαρακτηριστικά των δεδομένων, η καλύτερη ικανότητα γενίκευσης, βελτίωση της ικανότητας πρόβλεψης, δυνατότητα χρήσης σε πρακτικά προβλήματα, χρήση μεγάλου όγκου και αδόμητα χρηματοοικονομικά δεδομένα. να εξάγει χαρακτηριστικά από ακολουθιακά δεδομένα, επιπλέον μικρότερο υπολογιστικό κόστος στην διάρκεια της εκπαίδευσης.

Τα μοντέλα που ακολουθούν τα LSTM ως προς την συχνότητα των εφαρμογών είναι τα υβριδικά μοντέλα. Τα υβριδικά μοντέλα βελτιώνουν την απόδοση πρόβλεψης, μειώνουν την αβεβαιότητα και έχουν ικανότητα γενίκευσης, όμως έχουν μεγαλύτερο χρόνο σχηματισμού και χρειάζονται περισσότερους υπολογιστικούς πόρους. Η έρευνα εξετάζει την πρόβλεψη για τιμές μετοχής και δεικτών ενώ πρόκληση αποτελεί για τις εφαρμογές η χρησιμοποίηση κειμένου από ειδήσεις και άλλα κανάλια στα μοντέλα πρόβλεψης. Τα σύνολα δεδομένων που εξετάζονται προέρχονται με την σειρά με τις περισσότερες παρατηρήσεις ο S&P500, ο S&P Index, ο NASDAQ, ο Shanghai Composite Index και άλλοι. Τα δεδομένα προέρχονται περισσότερο από τις αναπτυγμένες χώρες και ακολουθούν οι αναπτυσσόμενες χώρες.

Η έρευνα συμπεραίνει ότι η μετρική του μέσου τετραγωνικού σφάλματος δεν είναι η πιο κατάλληλη για την περίπτωση της χρήσης του μοντέλου για το χρηματοοικονομικό εμπόριο. Αυτό που προτείνεται είναι η εξασφάλιση μιας σχέσης μεταξύ ακρίβειας και κερδοφορίας συμπεριλαμβανόμενων του κόστους που χρειάζεται να περιληφθεί στην αξιολόγηση της εφαρμογής, μια προσέγγιση που να λαμβάνει υπόψιν τα χαρακτηριστικά του risk-award "ή risk-return θα μπορούσε να αποτελέσει αντικείμενο έρευνας.

Η έρευνα εξετάζει επίσης τα πιο συχνά δημοσιευμένα άρθρα, καθώς αυτά χρησιμοποιούνται ως βάση αναφοράς για περαιτέρω έρευνα. Ειδικότερα στην έρευνα αναδεικνύεται η δημοσιευμένη μελέτη του Fischer Thomas (2018) ως το πιο συχνό αναφερόμενο άρθρο για την εφαρμογή των

LSTM. Επιπλέον παραθέτει και μια μέθοδο άλλη από το μοντέλο BM την εφαρμογή Γενετικού Αλγορίθμου (ΓΑ) με SVR μεθοδολογία την έρευνα των Chien Fung Huang (2012).και άλλους ερευνητές.

Μια καινοτόμος προσέγγιση της έρευνας είναι η απεικόνιση των μονοδιάστατων χρονοσειρών σε ένα δισδιάστατο επίπεδο και την εφαρμογή των αλγορίθμων CNN με τρεις καταστάσεις εξόδου buy, sell , hold.

Οι αλγόριθμοι MM έχουν προβλήματα με υψηλών διαστάσεων δεδομένων, (big data)και συναντάνε προβλήματα όπως καταστροφή διαστάσεων και μη κατάλληλη αναπαράσταση χαρακτηριστικών. Η λύση στις υψηλές διαστάσεις καλύπτεται από την ανάπτυξη των μοντέλων BM[(Tang et al, 2022)]

Στην ανασκόπηση των [(Sezer et al, 2020)] για την χρήση μοντέλων BM για πρόβλεψη χρηματοοικονομικών τιμών μεγεθών, διαπιστώνεται ότι τα μοντέλα που χρησιμοποιούνται περισσότερο είναι τα RNN και ειδικότερα τα LSTM,το οποίο έχει και απλή φάση ανάπτυξης μοντέλου μαζί με κάποια υβριδικά μοντέλα. Τα μοντέλα CNN και DMLP έχουν εκτεταμένη εφαρμογή σε υλοποιήσεις ταξινόμησης με το DMLP να προτιμάται σε πρόβλεψη ταξινόμησης, μετοχής ή δείκτη και τάσης τιμής σε σχέση με το MLP. Το CNN χρησιμοποιείται για χρονοσειρές με μετασχηματισμό των μονοδιάστατων τιμών σε δισδιάστατες. Για την περίπτωση εύρεσης τιμής χρησιμοποιούνται τα RNN μοντέλα ενώ για ταξινόμηση ενώ τα μοντέλα DMLP για ταξινόμηση τάσης. Η μελέτη συμπεραίνει ότι τα μοντέλα BM έχουν γενικότερα καλύτερη εφαρμογή από τα μοντέλα MM. Στην μελέτη εξετάζονται οι περιπτώσεις πρόβλεψης μετοχών, δεικτών χρηματιστηρίων (Indexes). Τα δεδομένα που χρησιμοποιήθηκαν για τα μοντέλα RNN είναι τα ανεπεξέργαστα. Άλλα δεδομένα που χρησιμοποιήθηκαν είναι ροή ειδήσεων , tweets, δηλώσεις για την εξαγωγή συναισθήματος. Μια στρατηγική στα μοντέλα είναι χρήση η εξαγωγή ανάλυσης συναισθήματος και η χρήση της με τεχνικούς δείκτες καθώς και θεμελιώδεις δείκτες.

Μια άλλη κατηγορία μάθησης που χρησιμοποιείται είναι η DRL για algorithmic trading, με χρήση μοντέλα agent-based μοντέλων με προεκτάσεις σε financial setniment analysis μέσω text mining για την εύρεση της ψυχολογίας των επενδυτών.

Επιπλέον το πεδίο του Text Mining χρησιμοποιείται με την ανάλυση χρηματοοικονομικών ειδήσεων, “tweets”,δηλώσεις και blogs με τον συνδυασμό με θεμελιώδη και τεχνική ανάλυση.

Μελλοντικές προοπτικές υπάρχουν για τα μοντέλα, 2-D CNN, Graph CNN, GAN.

Επίσης αναφέρεται ως συμπέρασμα ότι και οι μέθοδοι MM είχαν σε ορισμένες περιπτώσεις καλύτερη εφαρμογή από τις μελέτες BM.

Στην συστηματική έρευνα των [(Olorunnimbe et al,2022)] αναφέρεται η χρήση μεθόδων BM ως προς τα χρηματοοικονομικά κυρίως στην αγορά μετοχών,σε θέματα στρατηγική εμπορίας,

πρόβλεψη τιμής, διαχείριση χαρτοφυλακίου, και με ελάχιστες αναφορές σε μελέτες προσομοίωση αγοράς, επιλογή μετοχών, στρατηγική hedging και διαχείριση ρίσκου. Η έρευνα εξετάζει την εφαρμογή των μοντέλων σε παλιότερα δεδομένα για την επαλήθευση της ικανότητας των μοντέλων (backtesting) την οποία θέτει ως σημαντική για την αξιολόγηση του μοντέλου.

Η μελέτη καταλήγει ότι σημαντικό για την εμπιστοσύνη στις μεθόδους TN η ικανότητα δημιουργίας μοντέλων που μπορούν να αναπαράγονται, να είναι επεξηγήσιμα και αξιόπιστα. Επιπλέον σε σχέση με τις έρευνες προτείνεται η αύξηση του χρονικού ορίζοντα των προβλέψεων με χρήση συμπληρωματικών δεδομένων.

Η έρευνα αναφέρει την ευρεία χρήση των LSTM και DRL, LSTM+RL για πρόβλεψη ενώ άλλες μέθοδοι είναι ο CNN. Ως προς την στρατηγική εμπορίας ισχύουν τα ίδια συμπεράσματα LSTM, RL παράλληλα με μηχανισμό Attention, καθώς και CNN. Για την επιλογή μετοχών προτείνεται η CNN και η LSTM.

Στην ανασκόπηση του [(Jiang,2021)] γίνεται έλεγχος των δημοσιεύσεων της περιόδου 2017-2019, με θέμα χρηματοοικονομική πρόβλεψη και θέμα την πρόβλεψη τιμών κλεισίματος μετοχών διαπιστώνεται η χρήση υβριδικών μοντέλων BM. Τα περισσότερο χρησιμοποιούμενα είναι των LSTM,FFNN, CNN και RNN καθώς και των GAN, RL, και TL(Transfer Learning). Η Transfer Learning χρησιμοποιείται για την εκπαίδευση των DNN με μικρή ποσότητα χρόνου εκπαίδευσης και μικρό χρόνο εκπαίδευσης. Από τα υβριδικά μοντέλα τα πιο συχνά χρησιμοποιούμενα είναι τα CNN+LSTM και άλλοι συνδυασμοί είναι AE+LSTM,RNN+LSTM. Επιπλέον τίθενται θέματα όπως η μέτρηση της κερδοφορίας των μοντέλων, και η ανάλυση σημαντικότητας με μοντέλα βάσης, αναφοράς, η μεταφορά γνώσης στις αγορές, ενώ προτείνονται και νέες μέθοδοι από νευρωνικά δίκτυα όπως είναι τα Graph Neural Network(GNN), Capsule Network, Generative Adversarial Network (GAN).

Εκτός από τις μετρικές για περιπτώσεις ταξινόμησης και παλινδρόμησης που χρησιμοποιούνται αναφέρεται και η ανάλυση κερδοφορίας με χρήση των αποδόσεων και του ρίσκου καθώς και η ανάλυση σημαντικότητας / significance analysis για την σύγκριση των μοντέλων με τα μοντέλα βάσης τα οποία είναι η Logistic Regression,ARIMA, GARCH, SVM/SVR

### **4.3 Χαρακτηριστικά της μελέτης**

#### **4.3.1 Περιοχές Θεμάτων Πρόβλεψης**

Οι εφαρμογές TN σε θέματα χρηματοοικονομικής πρόβλεψης που αναφέρονται σε δημοσιευμένες μελέτες, με χρήση ευρύτερων μεθόδων Μηχανικής Μάθησης όπως Νευρωνικών Δικτύων και Βαθιάς Μάθησης προσδιορίζονται στις παρακάτω περιοχές:

- Στρατηγική Εμπορίου/Trade Strategy: αφορούν αλγοριθμικές μεθόδους για την πραγματοποίηση διαδικασιών αποφάσεων αγοράς και πώλησης στην αγορά μετοχών.
- Πρόβλεψη τιμής αξιών/Asset Price Prediction: Πρόβλεψη της μελλοντικής τιμής μιας χρηματοοικονομικής αξίας ή μετοχής στην χρηματοοικονομική αγορά.
- Διαχείριση Χαρτοφυλακίου / Portfolio Management: Επιλογή και διαχείριση ενός συνόλου από χρηματοοικονομικές αξίες για την πρόβλεψη της μακροχρόνιας κερδοφορίας.
- Προσομοίωση αγοράς / Market Simulation: Δημιουργία δεδομένων αγοράς κάτω από διαφορετικά σενάρια (what-if). Επιπλέον είναι δυνατόν να προβλεφτεί η εξέλιξη σε περίπτωση ενός απροσδόκητου σεναρίου όπως είναι μια επιδημία ή ένας πόλεμος.
- Επιλογή μετοχών / Stock Selection: Επιλογή μετοχών στην αγορά μετοχών ως μέρος ενός χαρτοφυλακίου που βασίζεται σε ανάλυση μελλοντικών αποδόσεων.
- Διαχείριση Ρίσκου Risk Management: Εκτίμηση των ρίσκων που εμπεριέχονται στην εμπορία για μεγιστοποίηση των αποδόσεων.
- Στρατηγική Hedging / Hedging Strategy: Μεταφορά του ρίσκου επένδυσης σε ένα χρηματοοικονομικό εργαλείο παίρνοντας μια αντίθετη επιχειρηματική κίνηση σε μια άλλη χρηματοοικονομική αξία. [(Olorunnimbe et al,2022)]  
Επιπλέον ανιχνεύονται τα θέματα:
- Πρόβλεψη χρεοκοπίας, η οποία μπορεί να οδηγήσει τους επενδυτές να απομακρυνθούν από την έκθεση σε εταιρείες που αναμένεται να κλείσουν.
- Πρόβλεψη Μεταβλητότητας/Volatility η οποία αφορά την πρόβλεψη της μεταβλητότητας των αγορών μέσω των οποίων μπορεί να προσδιοριστεί ενδεχόμενη χρηματοοικονομική κρίση. Ένας τομέας που είναι ιδιαίτερα σημαντικός είναι η πρόβλεψη bubble bursts (σκάσιμο της φούσκας) ή και κατάρρευση των αγορών.

Η παρούσα εργασία εστιάζει στην πρόβλεψη χρηματοοικονομικών μεγεθών αξιών με χρήση μεθόδων μηχανικής μάθησης.

### 4.3.2 Αντικείμενο μελέτης πρόβλεψης

Το θέμα της χρηματοοικονομικής πρόβλεψης έχει ως αντικείμενο ή στόχο:

- **Πρόβλεψη Τιμής (Price forecasting)**, η οποία διαχωρίζεται σε **Πρόβλεψη ενός μετέπειτα σημείου** ή **Πρόβλεψη πολλών μετέπειτα σημείων multi-step forecasting**. Παραδείγματα είναι η πρόβλεψη της τιμής ανοίγματος την επόμενη μέρα ή 1-λεπτού τιμή μεγεθών.
- **Πρόβλεψη Τάσης (Trend Forecasting)**, ή **κατεύθυνσης τιμής (price direction)**, ανάγεται σε ένα πρόβλημα ταξινόμησης με 2 ή 3 κατηγορίες ταξινόμησης (ανοδική, καθοδική και ουδέτερη) και ανάλογες μετρικές συναρτήσεις.



- **Πρόβλεψη Μεταβλητότητας (Volatility forecasting)**, η οποία αναφέρεται στην μεταβολή των μεγεθών της με την διάρκεια του χρόνου. και αφορά διαφορετικούς χρονικούς ορίζοντες όπως είναι η διάρκεια μιας μέρας ή μεγαλύτερες της μιας μέρας χρονικές περιόδους.

Παράγοντες που προκαλούν μεταβολές την τιμή μιας μετοχής, την τάση των χρηματοοικονομικών μεγεθών και την μεταβλητότητα της είναι πολιτικοί και οικονομικοί παράγοντες, παράγοντες της βιομηχανίας και τομέα δραστηριότητας, η απόδοση των εταιρειών, σημαντικά γεγονότα, οι ειδήσεις και οι πληροφορίες.

### 4.3.3 Χαρακτηριστικά αγορών Χρηματοοικονομικών Μεγεθών

Οι Χρηματοοικονομικές Αγορές/Financial markets, είναι οι αγορές που λαμβάνουν χώρα αγοραπωλησίες χρηματοοικονομικών εργαλείων με τις κυριότερες τις αγορές μετοχών (stocks), τις αγορές ομολόγων (bonds), την αγορά ξένου συναλλάγματος (forex), αγορές παραγώγων (derivatives). Οι Χρηματοοικονομικές Αγορές, χαρακτηρίζονται από ακραίες μεταβολές, μη γραμμική συμπεριφορά, χαοτική συμπεριφορά, μεγάλο όγκο συναλλαγών. Πολλοί παράγοντες διαμορφώνουν την λειτουργία και τις τιμές των αγορών όπως γεγονότα, οικονομικές πολιτικές, η ψυχολογία και η συμπεριφορά των επενδυτών και των θεσμικών επενδυτών, οι ειδήσεις, οι φήμες. Επιπλέον οι αγορές βρίσκονται σε συνεχή επίδραση από τις εξελίξεις της εφαρμογής και χρήσης νέων τεχνολογιών όπως είναι ο τρόπος πραγματοποίησης των συναλλαγών, η τεχνολογία blockchain και άλλες.

Οι εφαρμογές πρόβλεψης εστιάζουν σε διάφορες τιμές χρηματοοικονομικών αξιόγραφων των μετοχών:

- Τιμή Μετοχής /Stock price
- Τιμή Δείκτη/Index price
- Τιμή Κρυπτονομίσματος/Cryptocurrency, και μελλοντικά Τιμή Ψηφιακών Νομισμάτων Κεντρικών Τραπεζών/Central Bank Digital Currency
- Τιμή Οικονομικών αγαθών/ Commodities- όπως ορυκτά καύσιμα, μεταλλεύματα και Αγροτικά Αγαθά /Soft Commodities, όπως τιμές σιτηρών
- Ισοτιμίες ξένου συναλλάγματος/Forex-Foreign Exchange
- Τιμή Επενδυτικών Κεφαλαίων /Investment Fund
- Τιμή Ομολόγων /Bonds
- Τιμή Παραγώγων/Derivatives

#### 4.3.4 Πρόβλεψη Τιμής Μετοχής(Stock)

Η αγορά Μετοχών (Stock Markets), αποτελεί την αγορά των μεριδίων των επιχειρήσεων όπου οι μετοχές εκδίδονται, πωλούνται και αγοράζονται στα ανταλλακτήρια μετοχών (stock exchange). Μέσω αυτής παρέχεται μακροπρόθεσμη χρηματοδότηση για τις επιχειρήσεις. Μεγάλες αγορές μετοχών είναι στην Αμερική είναι New York Stock Exchange (NYSE), ο Nasdaq για εταιρείες τεχνολογίας, Στην Ευρώπη, Frankfurt Stock Exchange, SIX Swiss Exchange, London Stock Exchange. Ενώ στην Ασία, Shanghai Stock Exchange, Κίνα, Tokyo Stock Exchange, Hong Kong Stock Exchange. Η κάθε αγορά έχει και τον αντίστοιχο Δείκτη της Index ο οποίος εξετάζεται μετέπειτα. Οι αγορές μετοχών χαρακτηρίζονται από μεγάλη μεταβλητότητα (volatility) ακόμα και μέσα στην διάρκεια της μιας ημέρας.

Τα δεδομένα που επιλέγονται για πρόβλεψη μετοχών εξαρτώνται από τον χρονικό ορίζοντα και αφορούν από δεδομένα υψηλής συχνότητας High Frequency, έως ημερήσιες, εβδομαδιαίες μηνιαίες τιμές. Ως ακατέργαστα δεδομένα χαρακτηρίζονται οι τιμές Open, Close, High, Low, Volume οι οποίες αναφέρονται και ως OCHLV. Επίσης για τον σχηματισμό των μοντέλων χρησιμοποιούνται και καθυστερούμενες τιμές (lagged). Παράλληλα χρησιμοποιούνται διαφορετικού μεγέθη σύνολα. Άλλες μεταβλητές που επιλέγονται για τον σχηματισμό των μοντέλων είναι δείκτες τεχνικής ανάλυσης ή θεμελιώδης ανάλυσης. Επίσης χρησιμοποιείται και διαδικασία μάθησης με τροφοδοσία από social media, και κείμενα από όπου εξάγεται συναίσθημα sentiment analysis. Στην περίπτωση ενός μοντέλου LSTM οι μεταβλητές εισόδου περιλάμβαναν μεταξύ των άλλων και χρηματοοικονομικές καταστάσεις, μεταβλητές τεχνικής και θεμελιώδης ανάλυσης, μακροοικονομικά δεδομένα (ΑΕΠ, συχνότητα ανεργίας, inventories).[(Sezer et al, 2019)]

Οι εφαρμογές αφορούν την σύγκριση των μοντέλων ως προς τον ίδιο τύπο αλλά και ως προς άλλου τύπους, όπως για παράδειγμα DL με ML και στατιστικές μεθόδους όπως Autoregressive Integrated Moving Average (ARIMA).Επίσης αφορούν την χρήση κειμένου ως μεταβλητή στα μοντέλα που δημιουργούνται.

Από την επεξεργασία των μελετών διαπιστώνεται ότι οι εφαρμογές μοντέλων πρόβλεψης χρηματοοικονομικών μεγεθών μετοχών που χρησιμοποιούν DL μοντέλα χρησιμοποιούν νευρωνικά δίκτυα όπως LSTM, RNN, GRU, CNN, Stacked Recurrent Networks (SRNN), Deep Forward Neural Networks (DFNN),[(Sezer et al, 2019)]

Παράλληλα οι εφαρμογές περιλαμβάνουν και την διερεύνηση σε ML όπως SVM, MLP, AE, RBM, Function Neural Network (RBF) and Extreme Learning Machine (ELM), Gradient Boosted Trees (GBT) and Random Forest (RF).[(Sezer et al, 2020)]

Υβριδικά μοντέλα/hybrid models, αφορούν μοντέλα CNN+LSTM, smart indexing με AE, DBN+MLP. LSTM+MLP, GAN-FD (Direction Prediction Loss), Principal Component Analysis και

2-επίπεδο FFNN, Weighted Multichannel Time-series Regression (WMTR) και Multilinear Discriminant Analysis (MDA).[(Sezer et al, 2020)]

Επίσης έχουν αναπτυχθεί και Καινοτόμες μέθοδοι/Novel methods, Deep and Wide Neural Network (DWNN), που αποτελεί έναν συνδυασμό RNN και CNN, State Frequency Memory (SFM) recurrent network.[(Sezer et al, 2020)]

Μέθοδοι που αφορούν το text mining και sentiment analysis χρησιμοποιούν γεγονότα από τεο Reuters, Bloomberg, προσθήκη λέξεων από αναφορές αναλυτών, ιστοσελίδων, χρηματοοικονομικά νέα, γεγονότα εταιρειών, δημοσιεύσεις από το Twitter, πληροφορίες απο ιστοσελίδες. Τα θέματα των προβλέψεων αφορούν πρόβλεψη δείκτη, τιμή μετοχής και τάση τιμής, με μοντέλα LSTM, CNN, LSTM+CNN, Word2vec, Recurrent CNN.[(Sezer et al, 2020)]

### **4.3.5 Πρόβλεψη Τιμής Δείκτη / Index Forecating**

Οι εφαρμογές στον τομέα της χρηματοοικονομικής πρόβλεψης, εστιάζουν στην πρόβλεψη της τιμής των χρηματοοικονομικών Δεικτών/Indexes των αγορών, που αποτελούνται από πολλαπλές μετοχές απο διαφορετικούς τομείς. Παρουσιάζουν μικρότερη μεταβλητότητα από τις κοινές μετοχές και είναι ενδεικτικοί για την γενική κατάσταση της αγοράς και της οικονομίας.

Μια κατηγορία μελετών πρόβλεψης εστιάζει στην πρόβλεψη των δεικτών, για μελλοντικές χρονικές περιόδους, για να εξετάσουν την σταθερότητα της αγοράς. Οι δείκτες παρέχουν μια ευρύτατη αναπαράσταση του πώς οι αγορές λειτουργούν και φανερώνουν την άνοδο και την πτώση της αγοράς, χωρίς να χρειάζεται να εξετάζουν κάθε μία εμπορεύσιμη αξία της αγοράς.Οι επενδυτές δημιουργούν τα χαρτοφυλάκια τους βάση των τιμών του Γενικού Δείκτη της αγοράς και των επιμέρους Δεικτών. Το υπόδειγμα του Δείκτη της αγοράς συσχετίζει την απόδοση ενός χρεογράφου με την απόδοση του Γενικού Δείκτη της αγοράς. Ως Index Fund ορίζεται το ομόλογο που προσδοκεί να αντιγράψει την συμπεριφορά ενός δείκτη δημιουργώντας ένα χαρτοφυλάκιο που να αντανakλά τον δείκτη.

Οι μελέτες που αναφέρονται στις ανασκοπήσεις εξετάζουν μια πληθώρα δεικτών κάποιοι απο τους οποίους είναι S&P500, Tokyo Nikkei Index (NIKKEI), NYSE, , DJIA, NASDAQ, TOPIX, FTSE, DAX, TSX, ASX, NZX-50.

Τα δεδομένα που χρησιμοποιούνται για την πρόβλεψη τιμών δείκτη, είναι ανεπεξέργαστα δεδομένα (raw time series data) OCHLV, τιμές τεχνικών δεικτών, index data, social media feeds, ειδήσεις (Reuters), διαφορετικές πηγές δεδομένων όπως τεχνικοί δείκτες, τιμές index, social media feeds, ειδήσεις από το Reuters και το Bloomberg, στατιστικές τιμές δεδομένων (τυπική απόκλιση, λοξότητα, κύρτωση, Ωμεγα αναλογία, fund alpha), Return Sharpe-ratio (SR), μακροοικονομικοί δείκτες.[(Sezer et al, 2020)]

Τα μοντέλα που χρησιμοποιούνται είναι MLP, RNN, LSTM, και DNN (DFNN ή DMPLP), Fuzzy Deep Direct Reinforcement Learning (FDDR), RL και καινοτόμες (novel) μέθοδοι όπως Genetic DNN, stacked LSTM και Bidirectional LSTM (Bi-LSTM) wavelet decomposition. Μια μελέτη χρησιμοποιεί το μοντέλο LSTM σε συνδυασμό με τεχνικούς δείκτες, μηχανισμό attention, και προσθήκη (embedding) market vector. [(Sezer et al, 2020)]

Καινοτόμες μέθοδοι για την πρόβλεψη του δείκτη είναι RNN models, Recurrent Computationally, Efficient Functional Link Neural Network (RCEFLANN), Functional Link Neural network (FLANN), με αναβαθμίσεις βαρών μέσω αλγορίθμων EA όπως Particle Swarm Optimization (PSO), Modified Version of PSO (HMRPSO), RNN-Boost με Latent Dirichlet Allocation (LDA) features. [(Sezer et al, 2020)]

### **4.3.6 Πρόβλεψη Τιμών Αγαθών (Commodities)**

Αποτελείται από τα Πρωτογενή Αγαθά τα οποία είτε λαμβάνονται κατευθείαν από τους φυσικούς πόρους όπως το αργό πετρέλαιο, τα αγροτικά προϊόντα, και τα Δευτερογενή Αγαθά τα οποία παράγονται από τα πρωτεύοντα για να καλύψουν συγκεκριμένες ανάγκες της αγοράς. Οι κινήσεις των τιμών των αγαθών μπορούν να αντιπροσωπεύουν σημαντικές εξελίξεις για τις οικονομίες που βασίζονται στην εξαγωγή αυτών των αγαθών.

Η πρόβλεψη αγαθών αφορά Χρυσό, Ασήμι, Πετρέλαιο, Χαλκός. Αφορά την πρόβλεψη τιμής, πρόβλεψη κατεύθυνση τιμής με χρήση τιμών δεδομένων αγαθών, δεικτών, τεχνικών δεικτών όπως το Relative Strength Index, το Williams Percent Range (William%R), το Commodity Channel Index (CCI), το Percentage Price Oscillator (PPOSC), το momentum, το Exponential Moving Average (EMA), Η χρήση των μοντέλων αφορά DNN, RNN, Elman RNN, FDDR, and CNN. Επίσης υβριδικά μοντέλα χρησιμοποιήθηκαν FNN και Stacked Denoising Autoencoders (SDAE) και τα DL συγκρίθηκαν με Support Vector Regressor (SVR), Random Walk (RW), και Markov Regime Switching (MRS) για πρόβλεψη του WTI πετρελαίου. Μια εφαρμογή FDDR για πρόβλεψη τιμής μετοχής και δημιουργία σημάτων trading αφορά τον συνδυασμό των μεθόδων DNN και RL. με κέρδη. Profit, return, Sharpe ratio, καμπύλες κέρδους απωλειών ως κριτήρια εφαρμογής. [(Sezer et al, 2020)]

Ως κριτήρια εφαρμογής χρησιμοποιήθηκαν ως μετρικές η Ακρίβεια/Accuracy, Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) [(Sezer et al, 2020)]

Ένας τομέας που αποτελεί πρόκληση, αφορά τον υπολογισμό των τιμών των αγροτικών προϊόντων λαμβάνοντας δεδομένα από τις καιρικές συνθήκες, οικονομικούς παράγοντες. Στην μελέτη των [(Gu et al. 2022)] γίνεται χρήση μετεωρολογικών δεδομένων, μηνιαίων δεδομένων εμπορίας

αγαθών αγροτικών προϊόντων και της μεθόδου DIA-LSTM(Dual Input Attention Long-Short Term Memory) για την πρόβλεψη της τιμής των αγροτικών προϊόντων. Παράλληλα σε μια άλλη μελέτη των [(Paul et al, 2022)] σε μια σύγκριση κλασσικών στατιστικών μοντέλων όπως ARIMA συγκρίνεται με τεχνικές ML και ειδικότερα την GRNN Support Vector Regression (SVR), Random Forest (RF) and Gradient Boosting Machine (GBM) για τον υπολογισμό τιμών χονδρικής με δεδομένα από 16 αγορές της Ινδίας έδειξε την ικανότητα πρόβλεψης μέσω των μεθόδων αυτών.

#### **4.3.7 Πρόβλεψη Τιμών Αγορας FOREX / Forex price Forecasting**

Οι αγορές ξένου συναλλάγματος παρουσιάζουν μεγάλη μεταβλητότητα, θόρυβο, σε σχέση με τις αγορές μετοχών. Έχουν τον μεγαλύτερο όγκο συναλλαγών από όλες τις χρηματοοικονομικές αγορές και συνεχόμενη λειτουργία αγοράς 24 ώρες το 24ωρο, 7 μέρες την εβδομάδα. Οι συναλλαγές γίνονται μέσω on line πλατφόρμας με δυνατότητα μόχλευσης. Οι περισσότερες συναλλαγές γίνονται σε δολάριο USD και οι περισσότερες μελέτες πρόβλεψης αναφέρονται στο δολάριο στις αναλύσεις τους στα ισοτιμία βάσης.

Σύμφωνα με την αναφορά την Bank for International Settlements (BIS), υπάρχουν κατά μέσο όρο 6 τρισεκατομμύρια συναλλαγές σε καθημερινή βάση, με πολλούς συμμετέχοντες, οι περισσότεροι από τους οποίους τις κάνουν για τυχοδιωκτικούς σκοπούς.

Η πρόβλεψη στο FOREX παρουσιάζει πρόκληση ως προς την βραχυπρόθεσμη πρόβλεψη με χρονικούς ορίζοντες πολύ μικρούς όπως είναι 1-λεπτό, 5-λεπτά για την απόφαση συναλλαγών.

Οι μέθοδοι που αναφέρονται για την πρόβλεψη τιμών FOREX:

- Μοντέλα Continuous-valued Deep Belief Networks (CDBN), DBN με Continuous Restricted Boltzman machine (CRBM), DBN+RBM, AE, Fuzzy information granulation μαζί με CDBN
- Μοντελα DNN, RNN, RNN EVOLINO, Psi-Sigma Network (PSN), LSTM μοντέλα, χρήση time varying volatility leverage, orthogonal input data.
- CNN μοντέλα σε συνδυασμό με άλλα μοντέλα MLP, RNN, Wavelet +CNN, AE+CNN
- Υβριδικά Μοντέλα (TAR-VEC-RHE), LRNFIS ενώ μια από τις μετρικές συναρτήσεις που χρησιμοποιήθηκαν είναι η FHSO αντί για την LMS-based error minimization.

#### **4.3.8 Προβλεψης Τιμής Ομολόγων / Bond price forecasting**

Η αγορά ομολόγων αντικατοπτρίζει την κατάσταση της οικονομίας. Ιστορικά οι μακροχρόνιες αποδόσεις των ομολόγων είναι μεγαλύτερες από τις βραχυπρόθεσμες αποδόσεις, ενώ μετά τις

οικονομικές κρίσεις βραχυπρόθεσμες συχνότητες αποδόσεων ξεπερνούν τις μακροχρόνιες αποδόσεις. [(Sezer et al, 2019)]

Το θέμα της προβλεψης των ομολόγων δεν έχει εξεταστεί από τους ερευνητές επαρκώς. Προσεγγίσεις του θέματος περιλαμβάνουν χρήση αλγορίθμων RF, AE, PCA και DFNN.

### **4.3.9 Πρόβλεψη Τιμής Κρυπτονομισμάτων/ Cryptocurrency price forecasting**

Η αγορά κρυπτονομισμάτων είναι νέα, με μεγάλη ανάπτυξη κεφαλαιοποίησης και χαρακτηρίζεται από μεγάλο δυναμισμό και ισχυρές διακυμάνσεις. Δέχεται επίσης τις επιφυλάξεις των επενδυτών σχετικά με τα θέματα κρυπτογράφησης και υπάρχει διαμάχη σχετικά με την ύπαρξη τυχοδιωκτικών “bubbles/φούσκων” στις διακυμάνσεις της αγοράς κρυπτονομισμάτων καθώς η αγορά θεωρείται τυχοδιωκτική (speculative).

Η τεχνολογία Blockchain στην οποία στηρίζονται οι συναλλαγές, έφερε σημαντική αλλαγή στον τρόπο των συναλλαγών οι οποίες γίνονται αποκεντρωμένα μεταξύ των χρηστών “peer-to-peer”. Παράλληλα τμήματα αξιών κρυπτονομισμάτων συναλλάσσονται μέσα από τις πλατφόρμες ιδιωτικών συναλλαγών. Η αγορά χαρακτηρίζεται από έναν μεγάλο αριθμό κρυπτονομισμάτων που εκδίδονται συνεχώς. Πλέον χώρες, μεγάλες εταιρείες, οργανισμοί και τράπεζες και χώρες εκδίδουν δικά τους κρυπτονομίσματα.

Αρκετές εφαρμογές για πρόβλεψη τιμής έχουν ως ισοτιμία ενδιαφέροντος το Bitcoin ενώ άλλες τιμές είναι Litecoin, StockTwits. Οι μεταβλητές εισόδου είναι OCHLV, τεχνικοί δείκτες, sentiment analysis. Ως μετρικές χρησιμοποιήθηκαν sensitivity, specificity, precision, accuracy, RMSE.

Οι μέθοδοι πρόβλεψης για τα κρυπτονομίσματα είναι:

- DNN, LSTM, GRU, RNN
- ARMA, ARIMA, Autoregressive Conditional Heteroscedasticity (ARCH), GARCH.
- Τα μοντέλα CNN και LSTM χρησιμοποιήθηκαν για Opinion Market [(Sezer et al, 2019)]
- Takagi-Sugeno Fuzzy cognitive map και 3-επίπεδα DMLP [(Ozbayoglu et al , 2022)], χρήση μεταβλητών Hash Value, Bitcoin Address, Public/Private Key, Digital Signature
- Δημιουργία Bitcoin Transaction Graph, node embedding, ανίχνευση σημείων εκτός συνηθισμένων ορίων (outliers) μέσω AE [(Ozbayoglu et al , 2022)]
- Συνδιασμός δύο μοντέλων 1. Text Mining με δύο μοντέλα CNN και LSTM με χρήση των μεταβλητών OCHLV, τεχνικούς δείκτες και sentiment analysis [(Ozbayoglu et al , 2022)]

- Προσέγγιση μέσω Reinforcement Learning με χρήση 3 μοντέλων RNN, LSTM, CNN[(Ozbayoglu et al , 2022)]
- Διαχείριση χαρτοφυλακίου μέσω CNN και DRL με χρήση Bitcoin, Ethereum, Bitcoin Cash, Digital Cash[(Ozbayoglu et al , 2022)]
- Δημιουργία χαρτοφυλακίου με τις μεθόδους LSTM, RNN, DMLP χρησιμοποιώντας 8 κρυπτονομίσματα ( Bitcoin, Dash, Ripple, Monero, Litecoin, Dogecoin, Bitcoin Cash, Digital Cash)[(Ozbayoglu et al , 2022)]
- Πρόβλεψη κατεύθυνσης τιμής με Bayesian Optimised RNN, LSTM, και ARIMA, και χρήση των μετρικών Sensitivity, Specificity, Precision, Accuracy, Root Mean Square Error( RMSE). [(Ozbayoglu et al , 2022)]

Από την μελέτη των ερευνών διαφαίνεται ότι οι περισσότεροι ερευνητές χρησιμοποιούν DL μοντέλα όπως LSTM, CNN, DMLP, RNN, Bi-LSTM, DRSE, DeepClue, Emotional Analysis +LSTM, NN, word2vec, Naive Bayes +LSTM.

Σαν κριτήρια εφαρμογής χρησιμοποιούνται η Ακρίβεια/ Accuracy, Precision,  $R^2$ , F1-Score, Recall, AUROC, MSE, Error Rate, Root Mean Square Error( RMSE), Specificity, Sensitivity

Η ενσωμάτωση Text Mining αφορά την χρήση Text, Financial News, News, Twitter Sentiment, Features from news articles, Sentiments Posts, Twitter Moods.

#### **4.3.10 Πρόβλεψη Τάσης / Trend forecasting**

Η πρόβλεψη τάσης αποτελεί ένα πρόβλημα ταξινόμησης. Οι μεταβλητές δεδομένων εισόδου που χρησιμοποιούνται είναι ανεπεξέργαστες τιμές δεδομένων (raw time series), τιμές μετοχών (OCHLV), τεχνικοί δείκτες, δείκτες θεμελιωδών δεδομένων και χρήση text mining και διάφορα άλλα δεδομένα.

Οι διαφορετικές μέθοδοι-μοντέλα που χρησιμοποιούνται για πρόβλεψη τάσης περιλαμβάνουν τις παρακάτω ομάδες όπου αναφέρονται και κάποια σχετικά παραδείγματα τους:

- Μοντέλα ANN, DNN, FFNN, DFNN
- Μοντέλα LSTM, RNN, GRU, GRU-SVM, Πιθανοκρατικά ΝΔ/Probabilistic NN (PNN), Timedelay Neural Networks (TDNN), eXtreme Gradient Boosting (XGBoost) με χρήση ανεπεξέργαστων δεδομένων. [(Sezer et al, 2020)]
- Καινοτόμες Μέθοδοι / Novel Methods, Μέθοδος CNN με Gramian Angular Field (GAF), Moving Average Mapping (MAM) και Candlestick με μετασχηματισμένα δεδομένα εικόνας, CNN με αναπαράσταση γνωρισμάτων (feature imaging) για την πρόβλεψη buy/sell/hold

θέσεων, Empirical Mode Decomposition και Factorization Machine Based Neural Network (EMD2FNN) για πρόβλεψη της τάσης τιμών κλεισίματος[(Sezer et al, 2020)]

- Μέθοδοι με χρήση Text Mining, DNN, DMLP, and CNNGRU model LSTM, BiGRU (Bidirectional Gated Recurrent Unit), TGRU (Two-stream GRU), CNN, χρήση 2 LSTM, word2Vec για την δημιουργία ενός LSTM+CNN Stock2Vec models, novel methods περιλαμβάνουν κείμενο απο το Twitter, ειδήσεις και χρηματοοικονομικές ειδήσεις, sentiment, limit order book, χρηματοοικονομικών ειδήσεων γραφημάτων, social media tweets.[(Sezer et al, 2020)]
- Μια από τις εφαρμογές αφορά την χρήση δύο LSTM από τα οποία το ένα χρησιμοποιεί τιμές από το word2Vec αλγόριθμο για την τροφοδοσία ενός CNN ενώ το άλλο χρησιμοποιεί γνωρίσματα τεχνικής ανάλυσης [(Sezer et al, 2020)]
- Καινοτόμες μέθοδοι αφορούν
  - την χρήση συνδυασμού από RBM, DBN, και ενσωμάτωση λέξεων για την δημιουργία διανυσμάτων λέξεων για ενα RNN-RBM-DBN δικτύου.
  - η μέθοδος DeepClue που μεταφράζει οπτικά μοντέλα που περιλαμβάνουν κείμενο με χρήση χρηματοοικονομικών ειδήσεων γραφημάτων, social media tweets
  - χρήση Hybrid Attention Networks και Text Mining, βασισμένο σε χρηματοοικονομικές ειδήσεις.
  - χρήση τεχνικής και sentiment ανάλυσης από social media για ταξινόμηση
  - Deep Neural Generative Model (DGM), με νέα από άρθρα με χρήση Paragraph Vector για την δημιουργία διανύσματος εισόδου για την πρόβλεψη τιμών μετοχών.
  - LSTM, RNN, and GRU με χρήση δεδομένων που προέρχονται από limit order book flow και ιστορικές πληροφορίες, social media news, LDA features,expert recommendations
  - CNN μοντέλα με χρήση εγγραφών limit order book
  - HFT πρόβλεψη μικροδομής με χρήση CNN.

#### **4.3.11 Πρόβλεψη Μεταβλητότητας/Volatility**

Η πρόβλεψη της μεταβλητότητας της αγοράς ή συγκεκριμένων αξιογράφων είναι σημαντική για τους επενδυτές και τις ηλεκτρονικές συναλλαγές. Υψηλή μεταβλητότητα σημαίνει ταραχή στην αγορά, απότομες διακυμάνσεις των τιμών, αυξημένο ρίσκο ενώ χαμηλή μεταβλητότητα



συνεπάγεται ομαλές αλλαγές στην τιμή και μειωμένο ρίσκο. Παράλληλα η τιμή των χρεογράφων εξαρτάται από την μεταβλητότητα της αγοράς. Η μεταβλητότητα χρησιμοποιείται επίσης ως είσοδος σε χρηματοοικονομικά μοντέλα περιλαμβανόμενων των μοντέλων ρίσκου. Μελέτες αποδεικνύουν ότι μακροοικονομικές μεταβλητές μπορούν να επηρεάσουν την μεταβλητότητα.

Η έγκαιρη πρόβλεψη της οδηγεί τις χρηματοοικονομικές αποφάσεις και τις οικονομικές πολιτικές, Ενσωματωμένη σε ένα σύστημα έγκαιρης προειδοποίησης δημιουργεί ασφάλεια στους επενδυτές ως προς την χρονική στιγμή της απόφασης και τον χρονικό ορίζοντα των αποφάσεων τους. Πρακτική εφαρμογή της πρόβλεψης της μεταβλητότητας γίνεται σε συστήματα εταιρειών συναλλαγών μέσα από ηλεκτρονικές πλατφόρμες (trading) όπου ειδοποιούν τους επενδυτές για τις διακυμάνσεις της αγοράς και ειδικότερα σε περιπτώσεις υψηλής μεταβλητότητας όπου οι συναλλαγές παρακωλύονται ή διακόπτονται για κάποιο χρονικό διάστημα.

Η μεταβλητότητα μπορεί να υπολογιστεί με πολλούς τρόπους και διακρίνεται σε Ιστορική Μεταβλητότητα/ Historic Volatility (HV) η οποία αποτελεί την τυπική διακύμανση (variance) των λογαριθμικών τιμών των αποδόσεων, Realised Volatility (RV) που υπολογίζεται από το άθροισμα των τετραγώνων των λογαριθμικών αποδόσεων σε ένα χρονικό διάστημα, Implied Volatility (IV) η οποία υπολογίζεται ανάποδα από την τιμή μιας option μέσω κάποιου μοντέλου. Ως προς τον χρονικό ορίζοντα της μεταβλητότητας σε αγορές όπως είναι των κρυπτονομισμάτων είναι σημαντική η μεταβλητότητα κατά την διάρκεια της ημέρας ακόμα και κάποιον λεπτών της ώρας. Ο δείκτης Volatility Index (VIX), όπως αυτό του Chicago Board Options Exchange, αποτελεί δείκτη για την προσδοκώμενη μεταβλητότητα 30 ημερών των τιμών “bid”, “ask” εργαλείων options των οποίων η περίοδος λήξης βρίσκεται μεταξύ 23 και 27 ημέρες. [(Ge et al, 2022)]

Τα Αυτοπαλινδρομούμενο μοντέλο με Δεσμευμένη Ετεροσκεδαστικότητα (Autoregressive Conditionally Heteroscedastic – ARCH) είναι ένα μοντέλο που χρησιμοποιείται στην οικονομετρία και θεωρεί την διακύμανση του τρέχοντος σφάλματος ως συνάρτηση των όρων σφάλματος των προηγούμενων χρονικών περιόδων. Το μοντέλο προϋποθέτει το μοντέλο αυτοπαλινδρόμησης (AR autoregression) για την διακύμανση του σφάλματος. Επέκτασή του μοντέλου ARCH, και ένα από τα πιο συχνά χρησιμοποιημένα εργαλεία για πρόβλεψη της μεταβλητότητας στις χρηματοοικονομικές χρονοσειρές είναι το μοντέλο GARCH (Generalised Autoregressive Conditional Heteroscedasticity) το οποίο ενσωματώνει καθυστερούμενη υπο συνθήκη διακύμανση (variance) για την διάγνωση των μεταβολών σε χρονικές περιόδους. Επιπλέον επεκτάσεις των δύο παραπάνω αποτελεί το μοντέλο ARMA (Autoregressive Moving Average), όπου η διακύμανση εκφράζεται ως αυτοπαλινδρούμενος μέσος όρος. Μειονέκτημά τους είναι η αδυναμία τους να μετρήσουν επιδράσεις μόχλευσης λόγω της συμμετρικής υπόθεσης κατανομή που κάνουν.

Στην ανασκόπηση των [(Ge et al, 2022)] εξετάζεται η πρόβλεψη της μεταβλητότητας σε θέματα χρηματοοικονομικών μεγεθών, με χρήση νευρωνικών δικτύων, σε 35 μελέτες από το 2015 και έπειτα διαπιστώνεται ότι τα πιο συχνά χρησιμοποιούμενα νευρωνικά δίκτυα είναι τα MLP ακολουθούμενα από τα RNN, και μία εφαρμογή των CNN. Από τις υβριδικές μεθόδους οι πιο συχνά εφαρμόσιμες είναι τα μοντέλα AR (Autoregressive) με την συμβολή των GARCH. ήταν τα πιο συχνά χρησιμοποιούμενα. Οι μεταβλητές εισόδου ήταν πολλαπλές μεταβλητές από ένα αξιόγραφο ενώ το πιο συχνό αντικείμενο της μελέτης ήταν η πρόβλεψη το S&P500.

Στην δημοσίευση των [(Jia & Yang, 2021)] το θέμα της μεταβλητότητας τριών δεικτών των ΗΠΑ αντιμετωπίζεται με την χρήση μοντέλων LSTM με συνάρτηση σφάλματος την likelihood-based loss function με κλαύτερη απόδοση από το μοντέλο ARMA-GARCH.

Στην μελέτη των [(Song et al, 2022)] συμπεραίνεται ότι η χρήση μακροοικονομικών μεταβλητών με χρήση μεθόδων GARCH και μοντέλων BM όπως το GRU διαπιστώνεται ότι η ακρίβεια πρόβλεψης των GRU είναι η μεγαλύτερη ακολουθούμενη από μεθόδους MM ενώ τελευταίο σε απόδοση είναι το μοντέλο GARCH\_MIDAS μπορεί να βελτιώσει την ακρίβεια πρόβλεψης. Επιπλέον διαπιστώνεται μέσω ανάλυσης συσχέτισης και εξαγωγή σημαντικών χαρακτηριστικών αποδεικνύεται ότι ο μακροοικονομικός παράγοντας διαδραματίζει σημαντικό ρόλο στην πρόβλεψη της μεταβλητότητας των μετοχών.

Ωστόσο οι μελέτη των δημοσιεύσεων αποδεικνύει ότι οι διαφορετικές προσεγγίσεις των ερευνητών μέσω των μεταβλητών που χρησιμοποιούν, των δειγμάτων της αγοράς που λαμβάνουν ακόμα και η φύση της αγοράς παίζουν ρόλο στα αποτελέσματα των πειραμάτων. Ακόμα και αν διαφαίνεται ότι στις περισσότερες μεθόδους που εξετάστηκαν η χρήση των μοντέλων MM όπως TNΔ και BM μπορεί να ξεπερνάει τα παραμετρικά μοντέλα στην πρόβλεψη της μεταβλητότητας, επιπλέον ανάλυση χρειάζεται για την επιβεβαίωση των αποτελεσμάτων.

#### **4.3.12 Πρόβλεψη Κρίσεων /Forecasting Stock Market Crisis**

Μια από τις μελέτες που παρουσιάζει ενδιαφέρον ως προς την εφαρμογή των μεθόδων των μοντέλων DL και των επικυρώσεων των μοντέλων είναι σχετικά με τον προσδιορισμό των παραγόντων που οδηγούν σε χρηματοοικονομικές κρίσεις.

Η χρήση της MM στην μελέτη των [(Chatzis, 2018)] γίνεται προσπάθεια να προσδιοριστούν οι παράγοντες ή τα επεισόδια που οδηγούν στην κρίση της αγοράς μετοχών ή ακόμα και την κατάρρευση της. Η μελέτη παρουσιάζει την χρήση διαφορετικών αλγορίθμων όπως Classification Trees, Support Vector Machines, Random Forests, Neural Networks, Extreme Gradient Boosting και Neural Networks. Η μελέτη εξετάζει διάφορα χρονικά πλαίσια. Οι ανεξάρτητες μεταβλητές που επιλέγονται ως ιδανικά ως πρόδρομοι της χρηματοοικονομική πτώσεις είναι οι αποδόσεις και η

διακύμανση, η επιλογή των κατάλληλων μεταβλητών γίνεται μέσω αλγορίθμων μηχανικής μάθησης, ενώ δειγματοληψία Bootstrap εφαρμόζεται για την ρύθμιση της ανομοιομορφίας των δεδομένων. Τα δεδομένα αφορούν ημερήσιες τιμές μετοχών, ομολόγων, δεικτών (ομολόγων και αγαθών όπως πετρέλαιο, χρυσός) και ισοτιμιών από 39 χώρες που καλύπτουν ένα ευρύ φάσμα των οικονομιών.

Το ζήτημα της εύρεσης των παραγόντων που πυροδοτούν τις κρίσεις, δημιουργώντας ένα Σύστημα Έγκαιρης Προειδοποίησης (Early Warning System) είναι ιδιαίτερα πολύπλοκο λόγω των αλληλεπιδράσεων των οικονομιών των χωρών, της παγκοσμιοποίησης της κίνησης του κεφαλαίου αλλά και της ψυχολογίας της αγοράς.

Η έρευνα επικεντρώνει στα θέματα της μεταβλητότητας ως μηχανισμό αναγνώρισης χρηματοοικονομικών κρίσεων, στην χρήση των εξελιγμένων μεθόδων MM, την χρήση ημερησίων δεδομένων από τις αγορές, τον συνδυασμό πολλαπλών τεχνικών μηχανικής μάθησης, βρίσκοντας ως τρόπο μείωσης του model risk και την αύξηση της ακρίβειας πρόβλεψης, και εν τέλει πραγματοποιεί την μελέτη σε βάθος εικοσαετίας, σε διάφορες χώρες γεωγραφικά και διαφορετικές οικονομίες όπου είναι δυνατό να προσδιοριστούν οι περίοδοι με πίεση μεταβλητότητας.

Σημαντικά σημεία κατά την δημιουργία του μοντέλου είναι η επιλογή χαρακτηριστικών γνωρισμάτων από τα δεδομένα με την εφαρμογή διαφορετικών μεθόδων, η δημιουργία εξισορροπημένων συνόλων δεδομένων μέσω λογισμικού με κατάλληλες μεθόδους δειγματοληψίας και μια ομαλή bootstrap προσέγγιση, κανονικοποίηση των δεδομένων στο διάστημα [0,1]. Η μέθοδος που χρησιμοποιήθηκε για τον υπολογισμό της ακρίβειας του μοντέλου είναι η k- μερών cross validation.

Σχετικά με τις μεθόδους που χρησιμοποιήθηκαν

- Logistic Regression,
- Decision Trees(CART), η διαμόρφωση των κόμβων διαμόρφωση και μιας ιεραρχίας από συνθήκες ως προς τις ανεξάρτητες μεταβλητές
- Support Vector Machines
- Random Forests, η επιλογή των μεταβλητών πραγματοποιήθηκε βάση του Μέσου Τετραγωνικού Σφάλματος και του Node Purity
- Neural Networks, χρησιμοποιήθηκαν 1 έως 3 κρυφά επίπεδα, με βέλτιστο τα 2, με 2 έως 10 νευρώνες σε κάθε επίπεδο, η μέθοδος οπισθοδιάδοσης, με threshold 0.01 ή μέχρι το τέλος των επαναλήψεων, k-fold cross validation, σιγμοειδής συνάρτηση ενεργοποίησης

- Extreme Gradient Boosting, χρησιμοποιήθηκε που παρέχει καλύτερη πρόβλεψη και κλιμάκωση σε σχέση με την RF και y NN, παρέχοντας συναρτήσεις παλινδρόμησης, ταξινόμησης και διάταξης σε σειράς (ranking). Το εύρημα ήταν ότι οι διακυμάνσεις στην αγορά παραγώγων και αγορών ισοτιμιών αυξάνουν την πιθανότητα κρίσης.
- Deep Learning, έγινε χρήση του Multi-Layer Perceptron, με χρήση της βιβλιοθήκης MXNET με 5 κρυφά επίπεδα, 5-fold cross validation και μεγιστοποίηση της περιοχής κάτω από την καμπύλη, συνάρτηση ενεργοποίησης ReLU, χρήση Dropout μεθόδου για κανονικοποίηση

Ως μέτρα εγκυρότητας χρησιμοποιήθηκαν η Sensitivity και η Specificity, G-mean και άλλες

Από τα παραπάνω μοντέλα αυτά της BM είχε την καλύτερη εφαρμογή με αυξημένη ακρίβεια ταξινόμησης και συνέπεια εφαρμογής, που συνεπάγεται την ικανότητα γενίκευσης. Η μέθοδος διαφέρει σε σχέση με τις άλλες μεθόδους ως προς την συνδυαστική ανάλυση που παρουσιάζει έτσι ώστε να ελαχιστοποιηθούν τα περιθώρια σφαλμάτων.

Μια κατηγορία μελετών στοχεύει στην πρόβλεψη κατάρρευσης της αγοράς (stock market crash). Στην μελέτη των [(Chatzis, 2018)] γίνεται προσπάθεια να προσδιοριστούν η φύση και τα χαρακτηριστικά των παραγόντων που οδηγούν στην κρίση της αγοράς μετοχών ή της κατάρρευση της σε χρονικό ορίζοντα μιας και είκοσι ημερών.

### **4.3.13 Αλγοριθμικό Εμπόριο Αξιών / Algorithmic Trading**

Το αλγοριθμικό εμπόριο αναφέρεται σε συναλλαγές πώλησης-αγοράς χρεογράφων, με αποφάσεις που εκτελούνται από ένα αλγοριθμικό μοντέλο. Πραγματοποιούνται από τις online trading platforms. Οι μελέτες εστιάζουν σε μοντέλα πρόβλεψης τιμής ή τάσης, τιμές δεικτών και στην δυναμική της συναλλαγής βελτιστοποιώντας παραμέτρους trading όπως το bid-ask spread, ανάλυση limit order book, position sizing.

Στην μελέτη της μεθόδου ARIMA σε Υψηλής Συχνότητας Trading (High Frequency Trading) διαπιστώθηκε ότι παρουσιάζει μεγάλη απόκλιση για την πρόβλεψη χρηματοοικονομικών σειρών μεγάλης συχνότητας. Στην εφαρμογή συνδυασμού του μοντέλου ARIMA με BM για την πρόβλεψη χρονοσειρών υψηλής συχνότητας, διατήρησε την θεωρητική βάση του παραδοσιακού μοντέλου και την απόκριση του όρου σφάλματος του BM. Το συνδυαστικό μοντέλο απέδειξε με δοκιμές προσομοίωσης Monte Carlo, ότι έχει καλύτερη ακρίβεια πρόβλεψης, καλύτερη εφαρμογή σε σχέση με το ARIMA, SVM και LSTM και μειωμένη πολυπλοκότητα σε σχέση με το LSTM μοντέλο. Το μοντέλο αποτελεί εργαλείο για υψηλής συχνότητας στρατηγική για την μείωση του επιχειρηματικού ρίσκου του δείκτη μετοχών.

Το μοντέλο ARIMA-LSTM παρουσιάζει καλύτερη εφαρμογή, σταθερότητα και σε σχέση με το LSTM λιγότερη υπολογιστική πολυπλοκότητα και λιγότερο υπολογιστικό χρόνο και να ισοσταθμίσει τον χρόνο εκπαίδευσης του ARIMA. Το μοντέλο αποτελεί μια λύση για πρόβλεψη μετοχών, προσδιορίζοντας τον χρόνο επένδυσης, αποτρέποντας το ρίσκο και λαμβάνοντας σημαντικά οφέλη. [(Li, 2019)] Και πάλι εδώ ο συνδυασμός των μοντέλων οδηγεί σε καλύτερα αποτελέσματα μοντέλων.

Διαπιστώνεται μια τάση για χρήση Reinforcement Learning σε συνδυασμό με μοντέλα πρακτόρων σε αλγοριθμικό εμπόριο που συνδέονται με agent based models όπου διάφοροι agents αλληλεπιδρούν και μαθαίνουν από τις αλληλεπιδράσεις τους. Το πεδίο αυτό έχει περαιτέρω ευκαιρίες με αναβαθμίσεις σε financial sentiment analysis μέσω text mining για να καταλάβει την ψυχολογία του επενδυτή, με προεκτάσεις σε εφαρμογές behavioural finance. [(Sezer et al, 2019)]

## 4.4 Δεδομένα

Τα δεδομένα που χρησιμοποιούνται παρέχονται δωρεάν από το Διαδίκτυο, από τα ανταλλακτήρια των μετοχών, τις εταιρείες (Yahoo!), τις επενδυτικές εταιρείες. Οι μακροοικονομικοί δείκτες παρέχονται από International Monetary Fund (IMF), World Bank. Άλλες πηγές περιλαμβάνουν μέσα κοινωνικής δικτύωσης, το Twitter, Sina Weibo. Για relational data πηγή αποτελεί το Wikidata. Άλλες πηγές είναι το Bloomberg, Wind, Quantopian, Investing.com, Online brokers, το Kaggle αποτελεί μια ιστοσελίδα για δαιγωνισμούς πρόβλεψης μεγεθών. Βάσεις δεδομένων είναι η CSMAR, WRDS (Wharton Research Data Services). [(Jiang 2021)] Η διαθεσιμότητα τους δεν είναι πάντα δεδομένη καθώς υπάρχουν και περιπτώσεις στις οποίες συνεπάγεται σημαντικό κόστος απόκτησης.

### 4.4.1 Τύπος Δεδομένων

Διάφοροι τύποι δεδομένων μπορούν να χρησιμοποιηθούν για την εκμάθηση των μοντέλων:

**Δεδομένα Αγοράς/Market data:** περιλαμβάνουν τιμές δεδομένων από την αντίστοιχη αγορά που συμβαίνουν στις αγορές όπως Open-Τιμή Ανοίγματος μεγέθους, Close-Τιμή Κλεισίματος μεγέθους, High-Υψηλότερη τιμή κατά την διάρκεια της ημέρας, Low-Χαμηλότερη τιμή κατά την διάρκεια της ημέρας, Volume-όγκος συναλλαγών (OHLCV).

Στην περίπτωση των δεδομένων κατά την διάρκεια της ημέρας, αναφέρεται το κόστος που υπάρχει για την απόκτηση ποιοτικών δεδομένων και οι εφαρμογές χρησιμοποιούσαν δεδομένα ημέρας ενός χρόνου. [(Jiang, 2021)]

**Δεδομένα Κειμένου/Text data:** αφορά κείμενο από social media, ειδήσεις/news, αναζητήσεις στο διαδίκτυο/ web searches, το οποίο χρησιμοποιείται σε περιπτώσεις ανάλυσης συναισθήματος για την δημιουργία ενός δείκτη συναισθήματος/sentiment indicator.

**Μακροοικονομικά Δεδομένα/Macroeconomics data:** αντανακλούν τις οικονομικές συνθήκες μιας χώρας και είναι τιμές όπως Τιμή Δείκτη Καταναλωτή / Consumer Price Index (CPI), Ακαθάριστο Εθνικό Προϊόν/ Gross domestic product (GDP).

**Γνώση από δεδομένα από γραφήματα/ Knowledge graph data:** αποτελεί γνώση που προκύπτει από δεδομένα που αναπαρίστανται ως γραφήματα με χρήση μεθόδων όπως graph neural networks

**Δεδομένα Εικόνας/ Image data:** παράδειγμα αποτελεί η χρήση candlestick charts σαν δεδομένα εισόδου για πρόβλεψη μετοχών.

**Θεμελιώδη δεδομένα/ Fundamental data:** δεδομένα θεμελιώδους ανάλυσης που αφορούν την επιχείρηση, όπως λογιστικά δεδομένα.

**Αναλυτικά Δεδομένα/ Analytics data:** αναφέρεται σε δεδομένα που μπορούν να εξαχθούν από αναφορές/ reports π.χ. από τράπεζες που κάνουν μια σε βάθος ανάλυση των εταιρειών.

**Εναλλακτικά Δεδομένα/Alternative Data:** αφορούν δεδομένα όπως Google Trends.

Στις εφαρμογές των δεδομένων παρατηρούνται οι συνδυασμοί όλων των παραπάνω δεδομένων όπως Ιστορικές Τιμές & Τεχνικοί Δείκτες, Ιστορικές Τιμές & Κείμενο, Ιστορικές Τιμές με Τεχνικούς Δείκτες και Μακροοικονομικά στοιχεία, Ιστορικές Τιμές με Τεχνικούς Δείκτες και Κείμενο να είναι οι πιο συχνοί συνδυασμοί. [(Jiang 2021)]

Ανεξάρτητα από το πρόβλημα πρόβλεψης τα δεδομένα που χρησιμοποιούνται είναι ανεπεξέργαστες τιμές μετοχών (raw time series data) Επίσης σε πολλές μελέτες χρησιμοποιήθηκαν τιμές από δείκτες θεμελιώδους και τεχνική ανάλυσης όπως για πρόβλεψη τιμών μετοχών και δεικτών. Άλλα δεδομένα που χρησιμοποιήθηκαν είναι αυτά των κειμένων για την εξαγωγή των συναισθημάτων των επενδυτών/traders όπως χρηματοοικονομικά νέα, tweets, statements, blogs. [(Sezer et al, 2019)]

Στις περιπτώσεις των μελετών που εξετάστηκαν αφορούσαν χρηματοοικονομικές σειρές που προέρχονται από την Αμερική και την Ασία. Συνεπώς θα πρέπει να ληφθούν οι ιδιαιτερότητες αυτών των ερευνών στην εκτίμηση των αποτελεσμάτων. Ωστόσο πολλές από τις μελέτες που εξετάζονται λαμβάνουν δεδομένα χρονοσειρών από τις Ηνωμένες Πολιτείες Αμερικής.

Οι περισσότερες μελέτες που εξετάζονται αφορούν μια δοθείσα χρονοσειρά ενώ υπάρχουν και άλλες που εξετάζουν παράλληλα πολλές χρονοσειρές για κάποιο χρηματοοικονομικό μέγεθος.

Η μελέτη των [(Tang et al, 2022)] διαπιστώνει ότι ο πιο συχνός δείκτης πρόβλεψης είναι ο S&P500 όπως και οι περισσότερες μετοχές ανήκουν στον δείκτη αυτόν. Τα περισσότερα δεδομένα προέρχονται από αγορές αναπτυγμένων χωρών και ειδικότερα των ΗΠΑ αλλά και αναπτυσσόμενων χωρών όπως είναι ο Shanqai Composite Index.

Στην έρευνα τους εξετάζεται το θέμα από ποια περιοχή προέρχονται τα δεδομένα των μελετών. Οι Bruno et al(2019) διαπιστώνουν μεγάλο αριθμό των δεδομένων απο την αγορά της Βορείου Αμερικής.

Ανάλογα με τον χρονικό ορίζοντα πρόβλεψης διαφορετικές παράμετροι επιλέγονται που διαφοροποιούνται από το High Frequency Trading, intraday price movements, ημερήσιες τιμές, εβδομαδιαίες τιμές ή μηνιαίες τιμές κλεισίματος μετοχών. [(Sezer et al, 2019)]

Δεδομένα επιπλέον μεταβλητών, όπως οικονομικούς δείκτες ,δείκτες θεμελιώδους και τεχνικής ανάλυσης ή μετασχηματισμούς των δεδομένων αυτών. Χαρακτηριστικό παράδειγμα είναι η χρήση των χρηματοοικονομικών αποδόσεων των τιμών των μετοχών αντί της τιμής αγοράς.

Δεδομένα κειμένου, που προέρχονται από πηγές των Social Media, του Διαδικτύου και χρησιμοποιούνται σε τεχνικές Text Mining, Sentiment Analysis.

Ως προς την περιοδικότητα των δεδομένων υπάρχουν διαφορετικές επιμερίσεις αναλόγως του θέματος και των αλγορίθμων. Έτσι μπορεί να είναι τιμές κλεισίματος ή ανοίγματος ή μεγίστου ημέρας, εβδομαδιαίες, μηνιαίες, Στην περίπτωση του Algorithmic Trading ή και του High Frequency Trading, οι τιμές μπορεί να είναι αυτές που προσδιορίζονται ανα δευτερόλεπτο ή κλάσματα δευτερολέπτου.

#### **4.4.2 Μέγεθος δείγματος**

Το μέγεθος του δείγματος, δηλαδή η χρονική περίοδος των δεδομένων αποτελεί σημείο απόφασης στον σχηματισμό των μοντέλων. Επιλογή μικρής χρονικής περιόδου δεδομένων δεν είναι αντιπροσωπευτική και έχει ρίσκο υπερεφαρμογής ενώ η επιλογή μακράς χρονική περιόδου περιέχει τον κίνδυνο να συμπεριλάβει αναχρονικούς σχηματισμούς της αγοράς.

#### **4.4.3 Επεξεργασία Δεδομένων**

Τα δεδομένα στην μορφή που βρίσκονται μπορεί να παρουσιάζουν κάποια χαρακτηριστικά τα οποία δεν είναι επιθυμητά για όλους τους τύπους των μοντέλων και για αυτό ακολουθείται κάποια αρχική επεξεργασία. Ελλιπή δεδομένα, δεν παρατηρούνται με μεγάλη συχνότητα στις περιπτώσεις των χρονοσειρών. Ωστόσο η διαχείριση ελλιπών τιμών ή τιμών με τιμές πέραν των καθορισμένων γίνονται με συμπλήρωση από κάποια τιμή όπως ο μέσος όρος ή γειοτικές τιμές ή εξαιρούνται από τον υπολογισμό του μοντέλου. Δεδομένα, με μικρότερη συχνότητα δειγματοληψίας θα πρέπει να διαχειρίζονται διαφορετικά για την αποφυγή διαρροής δεδομένων, Τα δεδομένα της αγοράς έχουν υψηλό θόρυβο, ο οποίος μπορεί να οδηγήσει λανθασμένα στην πρόβλεψη. Την διαχείριση θορύβου η τεχνική που χρησιμοποιείται είναι η wavelet μετατροπή και η kNN-classifiers. [(Jiang,2021)] Η εφαρμογή κανονικοποίησης (Normalization), αφορά την μετατροπή των δεδομένων εισόδου έτσι

ώστε όλες οι τιμές να βρίσκονται στο διάστημα  $[0,1]$  ή  $[-1,1]$  Η εφαρμογή της τυποποίησης (Standardization), αναφέρεται στην αφαίρεση ενός μέτρου τοποθεσίας και διαίρεση με ένα μέγεθος κλίμακας. Κάποιοι ερευνητές μετατρέπουν τις χρονοσειρές σε πιο στατική μορφή, Για αυτές τις εφαρμογές τα Deep Forward Neural Networks και τα Convolution Neural Networks είναι τα συχνότερα εφαρμοζόμενα μοντέλα BM [(Sezer et al. 2019)] Η μετατροπή μια χρονοσειράς σε στατική γίνεται για τις εφαρμογές των μοντέλων CNN και DFFN. [(Sezer et al. 2019)]

Ένας καινοτόμος μετασχηματισμός που αναφέρεται σε χρηματοοικονομικές χρονοσειρές αφορά τον μετασχηματισμό της χρονοσειράς σε μια αναπαράσταση εικόνας και την εφαρμογή του μοντέλου CNN για εξαγωγή χαρακτηριστικών [(Sezer et al. 2019)]

#### **4.4.4 Εξαγωγή και Επιλογή Χαρακτηριστικών**

Η εξαγωγή χαρακτηριστικών αποτελεί τον μετασχηματισμό ενός αρχικού συνόλου μεταβλητών σε μια άλλη μορφή όπως το κείμενο ή η εικόνα μετατρέπονται σε αριθμητικά δεδομένα προς χρήση από τους αλγορίθμους. Εξαγόμενα χαρακτηριστικά από τα δεδομένα μπορούν να χρησιμοποιηθούν μαζί με τα πρωταρχικά δεδομένα της αγοράς

Στην περίπτωση των μελετών που εξετάζονται για την επιλογή των δεδομένων προς εκπαίδευση χρησιμοποιήθηκαν μέθοδοι εξαγωγή χαρακτηριστικών είναι μέσω Τεχνικής Ανάλυσης, με χαρακτηριστικά όπως κινούμενος μέσος όρος (Moving Average), moving average convergence/divergence (MACD) με χρήση της μεθόδου Chart Pattern Recognition. Οι Τεχνικοί Δείκτες μπορούν να χρησιμοποιηθούν για την κατασκευή δισδιάστατων εικόνων ως εισόδους στις μεθόδους Μηχανικής Μάθησης.

Για την Επιλογή Χαρακτηριστικών/Feature Selection, γίνεται χρήση των μεθόδων Chi-square μεθόδου, Maximum Relevance & Minimum Redundancy (MRMR), Rough Set Attribute Reduction (RSAR), Autocorrelation Function (ACF), Partial Correlation Function (PCF), ανάλυση διακύμανσης (ANOVA), Maximal Information Coefficient Feature Selection (MICFS), Principal Component Analysis (PCA), χρησιμοποιεί αποδόμηση του χώρου δεδομένων σε έναν χώρο μικρότερων διαστάσεων μέσω της μεθόδου Singular Value Decomposition και επιπλέον FRPCA: Fuzzy Robust Principal Component Analysis, KPCA: Kernel- Based Component Analysis, NLICA: Nonlinear Independent Component Analysis, SRA: Stepwise Regression Analysis ,GA: Genetic Algorithm, GRA: Grey Relation Analysis [(Berradi Z. et al, 2021)]

#### **4.4.5 Αριθμός Μεταβλητών Εισόδου**

Η βιβλιογραφία παρουσιάζει διαφορετικές εκδοχές ως προς την επιλογή των μεταβλητών για την πρόβλεψη των αποδόσεων των τιμών των μετοχών, από 2 έως 10 τιμές. Κάποιοι ερευνητές



χρησιμοποίησαν περισσότερες μεταβλητές για τα μοντέλα τους. Η βιβλιογραφία αναδεικνύει τον τρόπο που επιλέγονται οι μετοχές βασιζόμενη στην επίδραση που έχει κάθε μεταβλητή στο αποτέλεσμα που επιτυγχάνεται. Κάποιες μέθοδοι δεν παρέχουν εξήγηση για την επιλογή των μεταβλητών για την δημιουργία του μοντέλου αλλά βασίζονται στην διαθέσιμες μελέτες.

#### **4.4.6 Διαμερισμός δεδομένων**

Οι μέθοδοι διαμερισμού δεδομένων έχουν αναλυθεί στο προηγούμενο κεφάλαιο. Αυτό που διαφοροποιείται είναι τα ποσοστά που χρησιμοποιούνται για τα διαφορετικά σύνολα εκπαίδευσης, δοκιμή, επικύρωσης, την χρήση k-μέρη cross validation.

### **4.5 Περιβάλλον Ανάπτυξης/Development Environment**

Στις έρευνες που εξετάστηκαν και ανέφεραν τα εργαλεία ανάπτυξης το πιο συχνά χρησιμοποιούμενο είναι εργαλεία σχετικά με την γλώσσα προγραμματισμού Python. [(Jiang 2021)]

Αλλα εργαλεία είναι το Keras, TensorFlow, Pytorch, Matlab, Theano, MeCab, TALIB, Twitter API, H2O, Java, Sci-Kit Learn, R, Tensorflow in AWS.

### **4.6 Υλικό Εκτέλεσης Τεχνικών Βαθιάς Μάθησης**

Στην μελέτη του [(Jiang 2021)], διαπιστώνεται ότι τα μοντέλα BM χρειάζονται περισσότερους υπολογιστικούς πόρους για την υλοποίησή τους. Η υλοποίηση αυτή γίνεται με διάφορες Graphic Processing Units, ενώ η λύση της χρήσης Cloud computing από διάφορες εταιρείες δεν αναφέρεται στις ανασκοπήσεις.

### **4.7 Αξιολόγηση Μοντέλων**

Στις ανασκοπήσεις των μελετών τα μοντέλα που δημιουργούνται αξιολογούνται βάσει μετρικών συναρτήσεων, ανά τομέα εφαρμογής. Οι κατηγοριοποιήσεις που γίνονται είναι για θέματα Ταξινόμησης, για θέματα Παλινδρόμησης, ενώ εξετάζεται και το θέμα της Κερδοφορίας. Επιπρόσθετα στις δημοσιεύσεις αναφέρεται και το θέμα των χρήσεων παραδειγμάτων αναφοράς (Baseline).

#### **4.7.1 Μετρικές ταξινόμησης**

Οι μετρικές της απόδοσης ταξινόμησης του μοντέλου είναι:

Accuracy, precision, recall, sensitivity, specificity, F1 score, macro-average F-score, Matthews correlation coefficient, average AUC score, Theil's U coefficient, hit ratio, average relative variance, Confusion matrices και boxplots.

Για την ανασκόπηση του [(Jiang 2021)] οι πιο συχνές μετρικές που χρησιμοποιούνται είναι η Accuracy και το F1 Score ενώ οι αμέσως μετά χρησιμοποιούμενες είναι precision, recall, MCC.

#### **4.7.1.1 Μετρικές Παλινδρόμησης/ Regression metrics**

Οι μετρικές παλινδρόμησης χρησιμοποιούνται για την μέτρηση της εφαρμογής του μοντέλου σε αριθμητικές τιμές είναι:

(MAE), root mean absolute error (RMAE), mean squared error (MSE), normalized MSE (NMSE), root mean squared error (RMSE), relative RMSE, normalized RMSE (NRMSE), mean absolute percentage error (MAPE), root mean squared relative error (RMSRE), mutual information,  $R^2$  συντελεστής determination.

Για την ανασκόπηση του [(Jiang 2021)] οι πιο συχνές μετρικές που χρησιμοποιούνται είναι η RMSE και MAPE και στην συνέχεια η MAE and MSE

#### **4.7.1.2 Ανάλυση Κερδοφορίας**

Μια εναλλακτική μέθοδος αξιολόγησης ενός μοντέλου εκτιμά αν το μοντέλο πρόβλεψης μπορεί να είναι κερδοφόρος σε κάποια στρατηγική πρόβλεψης που θα ακολουθηθεί. Εκτιμάται σε όρους απόδοσης, που είναι οι διαφορές σε τιμές μετοχές ή χαρτοφυλακίου μετοχών και ρίσκου που εκτιμάται με διάφορες μεθόδους όπως είναι το μέτρο μέγιστης πιθανής απώλειας maximum drawdown (MDD), ή ετήσια μεταβλητότητα, η μέγιστη παρατηρούμενη απώλεια από κορυφή σε βύθιση  $MDD=(Trough\ Value-Peak\ Value)/Peak-Value$ .[(Jiang 2021)]

Μια μετρική που ενσωματώνει το return και το risk είναι η Sharpe Ratio, η οποία χρησιμοποιείται για να δείξει πως η απόδοση ενός εργαλείου πόσο καλά αποζημιώνει τον επενδυτή για το ρίσκο που πήρε.

#### **4.7.1.3 Ανάλυση Σημαντικότητας/ Significance Analysis**

Στην περίπτωση των εφαρμογών πρόβλεψης χρονολογικών σειρών αποδόσεων είναι σημαντική η χρήση στρατηγικών συναλλαγών αναφοράς ως σημείο σύγκρισης καθώς και στις πιθανοκρατικές συναλλαγές οι οποίες στηρίζονται στα διαστήματα πρόβλεψης.[(Ζαπράνης, 2005)] Τα στατιστικά μοντέλα και τα μοντέλα μηχανικής μάθησης στην μελέτη του [(Jiang 2021)] χρησιμοποιούνται ως σημείο μέτρησης αναφοράς (baseline) των μοντέλων BM. Η ανάλυση σημαντικότητας εξετάζει αν υπάρχει σημαντική διαφορά σε όρους πρόβλεψης συγκρίνοντας τα μοντέλα BM με τα μοντέλα βάσης αναφοράς. Παραδείγματα τέτοια είναι τα τεστ Kruskal–Wallis, Diebold-Mariano [(Jiang 2021)] που χρησιμοποιούνται για την στατιστική σημαντικότητα.

## 4.8 Αναπαραγωγή των Μοντέλων

Μια από τις παραμέτρους αξιολόγησης των μοντέλων είναι η δυνατότητα αναπαραγωγής από τους αναγνώστες της έρευνας που δίνει την δυνατότητα επιπλέον ανάλυσης και επικύρωσης των αποτελεσμάτων. Κάποιες από τις εφημερίδες δημοσιεύσεων έχουν θεσπίσει την υποβολή κώδικα και δεδομένων μαζί με την δημοσίευση ή την χρήση των αποθετηρίων κώδικα Github.

## 4.9 Υβριδικές Μέθοδοι

Ως Υβριδικές (Hybrid) ή Σύνθετες (Composite), μέθοδοι Μηχανικής Μάθησης αναφέρονται η χρήση δύο ή περισσότερων μεθόδων σε συνδυασμό σε ένα μοντέλο. Στην μελέτη διαπιστώθηκε η καλύτερη εφαρμογή των υβριδικών μεθόδων πρόβλεψης. Οι κατηγορίες των υβριδικών μοντέλων που συναντώνται είναι:

### **-Μοντέλο Βαθιάς Μάθησης μαζί με μοντέλο Βαθιάς Μάθησης**

Οι συνήθεις εφαρμογές που παρατηρούνται είναι συνδυασμός μεταξύ CNN και RNN καθώς και διαφορετικά RNN, ενώ το μοντέλο με την μεγαλύτερη χρήση είναι το LSTM.

### **-Μοντέλο Βαθιάς Μάθησης μαζί με τις κλασσικές μεθόδους ανάλυσης (στατιστική, θεμελιώδη, τεχνική)**

### **-Συνδυασμός κλασσικών μεθόδων ανάλυσης, με μοντέλο Μηχανικής Μάθησης και μοντέλο Βαθιάς Μάθησης.**

Τα υβριδικά νευρωνικά δίκτυα Fuzzy logic NNs (FNN) υπερτερούν σε σχέση από τα ΤΝΔ οπίσθια διάδοσης (backpropagation). Η διαφορά τους έγκειται στο γεγονός ότι η συσσώρευση γνώσης γίνεται μέσω ειδικών (experts) στην πρώτη περίπτωση και υπάρχει αιτιολόγηση στα αποτελέσματα και στην δεύτερη μέσω αλγορίθμου. Τα 3 πλεονεκτήματα που διαπιστώθηκαν σε εφαρμογή fuzzy NN σε investment risk management είναι ότι τα FNN έχει ισχυρή παράλληλη επεξεργασία, μαθαίνει σε σχέση με την αλλαγή του περιβάλλοντος και μπορεί να δημιουργήσει τους δικούς της κανόνες, καθώς είναι μη γραμμική διαδικασία δημιουργίας μοντέλου μπορεί να δημιουργήσει τα μοντέλα αρκετά γρήγορα. Ένας δημοφιλής συνδυασμός είναι η χρήση CNN και LSTM, RNN καθώς και ο συνδυασμό διαφορετικών μεθόδων RNN.

Η δημιουργία μοντέλων Νευρωνικών Δικτύων με χρήση δεικτών της Θεμελιώδους ανάλυσης αποτελεί άλλη μια συνδυαστική μεθοδολογία που εντοπίζεται σε έρευνες. Στην μελέτη των [(Yildiz B.,2010)], με χρήση δεδομένων από τις αγορές New York (NYSE), American (AMEX), Nasdaq(NSQ) που αφορούν τα έτη 1962 έως 2005, έγινε χρήση 18 financial ratios ως διάνυσμα εισόδου, που φαίνονται στον παρακάτω πίνακα για κάθε έτος των δεδομένων και την ετήσια απόδοση (buy-hold return). Για κάθε χρόνο υπολογίζεται η ετήσια buy-hold απόδοση 4 μήνες μετά το τέλος της χρονιάς. Ως διάνυσμα εξόδου τίθεται η απόδοση της εταιρείας αν δώδεκα μήνες. Το

σύνολο των δεδομένων χωρίστηκε σε τρία τμήματα εκπαίδευση/train(4000 τυχαία δείγματα από το 1962-1989), δοκιμή/test (70572 παρατηρήσεις από το 1962-1989), επικύρωση/validation (66.672 παρατηρήσεις από το 1990-2005). Το αποτέλεσμα για τις εταιρείες, ήταν να επισυναφθεί το αποτέλεσμα εξόδου για κάθε χρονικά σε μία κλίμακα με (1) το λιγότερο επιθυμητό και (10) το περισσότερο επιθυμητό. Το αποτέλεσμα ήταν ότι το σύστημα μπορούσε να προβλέψει μελλοντικές τιμές στις μετοχές των εταιρειών. Η μέτρηση των αποδόσεων έγινε με την χρήση της συνάρτησης παλινδρόμησης (Jensen's alpha)

**Πίνακας:** Δείκτες θεμελιώδους ανάλυσης

Current Ratio	Current Assets / Current Liabilities
Quick Ratio	(Cash and Short-Term Investments Plus Receivables – Total) / Current Liabilities
Short Term Debt to Equity	Current Liabilities / Stockholders' Equity
Debt to Equity	Total Liabilities / Stockholders' Equity
Debt to Assets	Total Liabilities / Total Assets
Interest Coverage	Income Before Interest and Tax Expense / Interest Expense
Liquid Asset Turnover	Net Sales / Liquid Assets
Current Asset Turnover	Turnover Net Sales / Current Assets
Turnover	Net Sales / Tangible Fixed Asset
Equity Turnover	Net Sales / Stockholders' Equity
Asset Turnover	Net Sales / Total Assets
Gross Profit Margin	Gross Profit / Net Sales
Operating Profit Margin	Operating Profit / Net Sales
Net Profit Margin	Net Profit / Net Sales
Return on Equity	Net Profit / Common Equity
P/E	Price/Earnings
B/M	Book/Market
P/S	Price/Sales

Με αυτόν τον τρόπο ήταν δυνατόν να δημιουργηθούν στρατηγικές από portfolio και να προβλεφθούν αποδόσεις χαρτοφυλακίων με ικανοποιητικές αποδόσεις.

## 5 Συμπεράσματα

- Από τις μελέτες των ερευνών αποδεικνύεται ότι στο πεδίο της χρηματοοικονομικής πρόβλεψης μεγεθών χρονοσειρών οι μέθοδοι που χρησιμοποιούνται περισσότερο είναι μοντέλα Βαθιάς Μάθησης και υβριδικές εφαρμογές τους. Η χρήση τους τροφοδοτήθηκε από την ύπαρξη των δεδομένων “Big data” και της αύξησης της υπολογιστικής ισχύος των υπολογιστών. Η σύγκριση με τις μέχρι τώρα μεθόδους MM και στατιστικής, αναδεικνύει τα πλεονεκτήματα της BM και των υβριδικών μοντέλων ως πιο αποτελεσματικής από την χρήση μεμονωμένων μεθόδων.
- Οι μέθοδοι Βαθιάς Μάθησης στην πρόβλεψη, κατορθώνουν να αντιμετωπίζουν προβλήματα πρόβλεψης με πολλές συνδεδεμένες χρονοσειρές και να εξάγουν ασθενή σήματα και πολύπλοκα πρότυπα από μεγάλα σύνολα δεδομένων. Η μέθοδος επιτρέπει τους ερευνητές να αποφύγουν τα προβλήματα trial and error που προκύπτουν κατά τις παραδοσιακές μεθόδους δημιουργίας μοντέλων όπου η επιλογή των μεταβλητών εισόδου στα δεδομένα, και τα επίπεδα των δικτύων ΝΔ αποτελεί πειραματική διαδικασία ή προκύπτει απο προηγούμενη μέθοδο εξαγωγής χαρακτηριστικών.[Benidis K. et al]
- Διαπιστώνεται το αυξημένο ενδιαφέρον της ακαδημίας για την Τεχνητή Νοημοσύνη και την Βαθιά Μάθηση, και διαφαίνεται η επιβεβαίωση του πλαισίου της 5ης βιομηχανικής επανάστασης σχετικά με την συνεργασία των ανθρώπων με την ΤΝ για την στήριξη των αποφάσεων σε θέματα χρηματοοικονομικής πρόβλεψης.
- Αναδεικνύεται επίσης το θέμα της Αξιόπιστης ΤΝ με την δημιουργία αλγορίθμων που θα προσφέρουν επεξηγησιμότητα το οποίο αναφέρεται συμβάλλοντας στην αξιοπιστία των μοντέλων ΤΝ.
- Από την έρευνα αναδεικνύεται το θέμα κάποιων δημοσιευμένων άρθρων να επιλέγονται ως άρθρα αναφοράς βάσης για τους ερευνητές έτσι ώστε να πραγματοποιήσουν και να εξελίξουν την δημιουργία των μοντέλων τους.
- Η περιοχή της πρόβλεψης τιμών μετοχών κυριαρχεί στο πεδίο έρευνας, με κυριότερους τομείς τιμής μετοχής, τάσης μετοχής και δείκτη.
- Σημαντικό ενδιαφέρον παρουσιάζουν οι έρευνες για πρόβλεψη μεταβλητότητας/ volatility αγοράς όπως και χρηματοοικονομικών κρίσεων με περιθώρια μελλοντικής έρευνας.
- Τα μοντέλα RNN είναι τα ευρέως χρησιμοποιούμενα για την προσέγγιση του θέματος της πρόβλεψης των ακολουθιακών δεδομένων των χρηματοοικονομικών χρονοσειρών είτε για θέματα πρόβλεψης τιμής είτε για θέματα πρόβλεψης τάσης.

- Από τις μελέτες που εξετάστηκαν η μέθοδος LSTM αναφέρεται ως η περισσότερο προτιμώμενη μέθοδος για την πρόβλεψη χρηματοοικονομικών χρονοσειρών είτε αυτόνομα είτε ως μέρος υβριδικού μοντέλου καθώς προσεγγίζει το θέμα των χρονικά σχετιζόμενων στοιχείων στις χρονοσειρές. [(Sezer et al, 2019)][Tkac & Verner].
- Επίσης τα μοντέλα CNN και DMLP χρησιμοποιούνται για το θέμα της εύρεσης τάσης το οποίο αποτελεί ένα θέμα ταξινόμησης. Ενώ παράλληλα για την εφαρμογή τους πραγματοποιείται και εξαγωγή χαρακτηριστικών. [(Sezer et al, 2019)] Τα μοντέλα CNN άρχισαν να χρησιμοποιούνται για χρονοσειρές με την εφαρμογή ενός μετασχηματισμού που μετατρέπει τα μονοδιάστατα δεδομένα σε διδιάστατα με την μέθοδο 2-CNN ενώ αρχικά αφορούσε επεξεργασία εικόνας. [(Sezer et al, 2019)]
- Τα μοντέλα CNN άρχισαν να χρησιμοποιούνται για χρονοσειρές με την εφαρμογή ενός μετασχηματισμού που μετατρέπει τα μονοδιάστατα δεδομένα σε διδιάστατα. Η μέθοδος 2-CNN ενώ αρχικά αφορούσε επεξεργασία εικόνας χρησιμοποιείται για την πρόβλεψη χρηματοοικονομικών μεγεθών καθώς μετατρέπει τις χρονοσειρές σε έναν τύπο δεδομένων εικόνας. [(Sezer et al, 2019)]
- Περαιτέρω εξέλιξη έχουν τα μοντέλα που αξιοποιούν και άλλες παραμέτρους όπως κείμενο από το διαδίκτυο, όπως οι μέθοδοι Sentiment Analysis και Text Mining. Από τις αρχικές τεχνικές που χρησιμοποιήθηκαν για την επίλυση του προβλήματος της πρόβλεψης χρηματοοικονομικών μεγεθών Θεμελιώδους ανάλυσης και Τεχνικής Ανάλυσης, οι νεότερες τεχνικές επικεντρώνουν στην sentiment ανάλυση και text mining.
- Μια νεότερη προσέγγιση στα δεδομένα που χρησιμοποιούνται στα μοντέλα είναι με την χρήση Graph Neural Networks.
- Η μεγαλύτερη ακρίβεια πρόβλεψης των μοντέλων των χρονοσειρών δεν συνεπάγεται και επικερδές μοντέλο για τους επενδυτές. [(Sezer et al, 2019)] Για επιπλέον ανάλυση προτείνεται η χρήση των τεχνικών risk-reward (ρίσκο-ανταπόδοση) και ενδεχόμενη χρήση του reinforcement learning.
- Το backtesting είναι σημαντικό ώστε να αποδειχθεί ότι τα μοντέλα που δημιουργούνται μπορούν να προσεγγίσουν παλιότερα δεδομένα.
- Επιπλέον χρειάζεται να γίνεται σύγκριση των μοντέλων που δημιουργούνται σε σύγκριση με κάποια μοντέλα βάσης (baseline) που αποτελούν άλλες μεθόδους που είχαν αναπτυχθεί παλιότερα για να διαπιστωθεί πιο αντικειμενικά η εφαρμογή τους.

- Η χρήση των μοντέλων DMMLP γίνεται απο πολλούς ερευνητές για περιπτώσεις ταξινόμησης ή πρόβλεψης τιμής τάσης λόγω της εξοικείωσης τους με την απλή χρήση των MLP για πρόβλεψη χρονοσειρών. Το μοντέλο χρησιμοποιείται για θέματα κυρίως ταξινόμησης όπως η πρόβλεψη της τάσης των μετοχών και παλινδρόμησης.
- Διαφορετική προσέγγιση στην χρήση του μοντέλου είναι ο συνδυασμός από μεταβλητές που προκύπτουν από την θεμελιώδη και την τεχνική ανάλυση, μαζί με δεδομένα χρηματοοικονομικού συναισθήματος μέσω Text Mining.

## 6 Επίλογος

Από την βιβλιογραφική έρευνα, την έρευνα στις σύγχρονες δημοσιευμένες μελέτες και ανασκοπήσεις σχετικά με την εφαρμογή των Τεχνητών Νευρωνικών Δικτύων διαπιστώνεται ότι η τεχνολογία αυτή και ειδικότερα η εξέλιξή της στα Νευρωνικά Δίκτυα Βαθιάς Μάθησης συνεχώς εξελίσσεται με νέες εφαρμογές και πρακτικές. Ειδικότερα στον τομέα της Χρηματοοικονομικής Πρόβλεψης η χρήση δικτύων RNN, LSTM και CNN αυτόνομα και σε υβριδικές εφαρμογές με άλλες μεθόδους δίνει μοντέλα με καλές αποδόσεις.

Η μελέτη των δημοσιευμένων μελετών εφαρμογών και ανασκοπήσεων μπορεί να δώσει χρήσιμα συμπεράσματα και να βοηθήσει τους ερευνητές στην περαιτέρω ανάπτυξη των εφαρμογών τους.



## Πηγές Αναφοράς

Stephanie M. Noble, Martin Mende, Dhruv Grewal, A. Parasuraman (2022), “The Fifth Industrial Revolution: How Harmonious Human–Machine Collaboration is Triggering a Retail and Service [R]evolution”, *Journal of Retailing*, 2, 199-208

OECD (2022), “OECD DIGITAL ECONOMY PAPERS, OECD FRAMEWORK FOR THE CLASSIFICATION OF AI SYSTEMS”, <https://www.oecd.org/publications/oecd-framework-for-the-classification-of-ai-systems-cb6d9eca-en.htm>

OECD (2019), “Artificial Intelligence in Society”, OECD, [https://www.oecd-ilibrary.org/science-and-technology/artificial-intelligence-in-society\\_eedfee77-en](https://www.oecd-ilibrary.org/science-and-technology/artificial-intelligence-in-society_eedfee77-en)

Ευρωπαϊκή Επιτροπή (2019), “ETHICS GUIDELINES FOR TRUSTWORTHY AI”, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

Weiwei Jiang (2021), "Applications of deep learning in stock market prediction: Recent progress", *Expert Systems with Applications*, 184, 115537, ELSEVIER,

Britannica (2022) Machine Learning

Marko Kolanovic, Rajesh T. Krishnamachari (2017), “Big Data and AI Strategies Machine Learning and Alternative Data Approach to Investing”, JPMorgan

Britannica (2022d), “Methods and goals in AI, Symbolic vs. connectionist approaches”

Albert Annor-Antwi and Ayman A. M. Al-Dherasi (2022), "Application of Artificial Intelligence in Forecasting: A Systematic Review", *American Journal of Computer Sciences and Applications*, 2019; 2:22. DOI: 10.28933/ajcsa-2019-11-0605

NINAREH MEHRABI, FRED MORSTATTER, NRIPSUTA SAXENA, KRISTINA LERMAN, and ARAM GALSTYAN (2021), "A Survey on Bias and Fairness in Machine Learning", *ACM Computing Surveys*, 54(6), 115, 2021

OECD (2022), Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449", OECD, 2022

OECD (2021), “Artificial Intelligence, Machine Learning and Big Data in Finance Opportunities, Challenges, and Implications for Policy Makers”, 2021

IBM (2022c), “IBM FactSheets Further Advances Trust in AI”, 2022, <https://research.ibm.com/blog/aifactsheets>

IBM (2022b), “Introduction to AI FactSheets”, 2022, <http://aifs360.mybluemix.net/introduction>

IBM (2022a), “What are foundation models?”, 2022, <https://research.ibm.com/blog/what-are-foundation-models>

Carlo Milana, Arvind Ashta (2021), "Artificial intelligence techniques in finance and financial markets: A survey of the literature", *Wiley On Line Library*.

OECD.AI (2022), "OECD.AI Policy Observatory Live Data", <https://oecd.ai/en/data>

Ζαπράνης Αχιλλέας(2005), "Χρηματοοικονομική και Νευρωνικά Συστήματα"

Ritika Chopra and Gagan Deep Sharma( 2021), "Application of Artificial Intelligence in Stock Market Forecasting: A Critique, Review, and Research Agenda",*J. Risk Financial Manag.*, **2021**, 14(11), 526; <https://doi.org/10.3390/jrfm14110526>

Deloitte (2021), "A path to automated financial forecasting", Deloitte Development LLC,

James Chen (2022), "Algorithmic Trading: Definition, How It Works, Pros & Cons", Investopedia,

Velvetech (2022), "The Role of High-Frequency and Algorithmic Trading", <https://www.velvetech.com/blog/high-frequency-algorithmic-trading/>

Francesco Rundo (2022), "Deep LSTM with Reinforcement Learning Layer for Financial Trend Prediction in FX High Frequency Trading Systems", 2019

Wikipedia (2022e), "Explainable artificial intelligence", 2022, [https://en.wikipedia.org/wiki/Explainable\\_artificial\\_intelligence#cite\\_note-1](https://en.wikipedia.org/wiki/Explainable_artificial_intelligence#cite_note-1)

Wojciech Samek, Thomas Wiegand, Klaus-Robert Muller (2017), "EXPLAINABLE ARTIFICIAL INTELLIGENCE: UNDERSTANDING, VISUALIZING AND INTERPRETING DEEP LEARNING MODELS"

Hamon Ronan, Junklewitz Henrik, Sanchez Ignacio (2022), "Robustness and Explainability of Artificial Intelligence", European Union

Kenny Olorunnimbe & Herna Viktor, (2022), "Deep learning in the stock market—a systematic survey of practice, backtesting, and applications",

Troy Segal (2022), "Fundamental Analysis: Principles, Types, and How to Use It", Investopedia, <https://www.investopedia.com/terms/f/fundamentalanalysis.asp>

John Brooks (2006), "Mastering Technical Analysis, Using the Tools of Technical Analysis for Profitable Trading", Mc Graw Hill

Adam Hayes (2022), "Technical Analysis: What It Is and How to Use It in Investing", <https://www.investopedia.com/terms/t/technicalanalysis.asp>

Yajiao Tang, Zhenyu Song , Yulin Zhu , Huaiyu Yuan, Maozhang Hou, Junkai Ji, Cheng Tang, Jianqiang Li (2022), "A survey on machine learning models for financial time series forecasting", *Neurocomputing*, 512, 363-380

Rajendra Akerkar (2018), "Artificial Intelligence for Business", SpringerLink

WENBO GE, POOIA LALBAKSHI, LEIGH ISAI, ARTEM LENSKIY, HANNA SUOMINEN (2022), "Neural Network-Based Financial Volatility Forecasting: A Systematic Review", *ACM Computing Surveys*, 55, 1, <https://doi.org/10.1145/3483596>

Scott Fortmann-Roe (2022), "Understanding the Bias-Variance Tradeoff", 2022, <https://scott.fortmann-roe.com/docs/BiasVariance.html>

Michal Tkáč, Robert Verner (2016), “Artificial neural networks in business: Two decades of research”, *Applied Soft Computing*, 38, 788-804

DUNIS L.C., MIDDLETON P.W., KARATHANASOPOULOS A., THEOFILATOS K.(2016), *Artificial Intelligence in Financial Markets*, Palgrave Macmillan.

Rob Hyndman, George Athanasopoulos (2022), “Forecasting Principles and Practices” 3<sup>rd</sup> edition, <https://otexts.com/fpp3/>

Soumya Shrivastava (2020), “Cross Validation in Time Series”, 2020,

Eduardo A. Gerlein, T.M. McGinnity, Ammar Belatreche, Sonya Coleman (2016), "Evaluating machine learning classification for financial trading: An empirical approach", *Expert Systems with Applications*, 54, 15

Sam Mahfoud & Ganesh Mani (1996) Financial forecasting using genetic algorithms, *Applied Artificial Intelligence*, 10:6, 543-566, DOI: 10.1080/088395196118425

Nagesh Singh Chauhan (2019), “A Friendly Introduction to Support Vector Machines”, 2019, <https://www.kdnuggets.com/2019/09/friendly-introduction-support-vector-machines.html>

Simon Haykin (2009), “Neural Networks and Learning Machines”, Third Edition, 2009

Ian Goodfellow, Yoshua Bengio, Aaron Courville (2022a), “Deep Learning”, *Adaptive Computation and Machine Learning series*

Weiwei Jiang (2021), “Applications of deep learning in stock market prediction: Recent progress”, *Expert System with Applications*, 184, 1

Kenny Olorunnimbe, Herna Viktor (2022), "Deep learning in the stock market—a systematic survey of practice, backtesting, and applications", *Artif Intell Rev*, 30;1-53. doi: 10.1007/s10462-022-10226-0.

Omer Berat Sezer, Mehmet Ugur Gudelek, Ahmet Murat Ozbayoglu (2020), "Financial time series forecasting with deep learning : A systematic literature review: 2005–2019", *Applied Soft Computing Journal*, 90, 106181

Asja Fischer, Christian Igel (2012), “An Introduction to Restricted Boltzmann Machines”, SpringerLink, CIARP 2012, Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications pp 14–36

Jun-Hua Chen, Yan-Hui Hao, Hao Wang, Tao Wang and Ding-Wen Zheng( 2019), "Futures price prediction modeling and decision-making based on DBN deeplearning", 53-65

Charu C. Aggarwal (2022), “Neural Networks and Deep Learning”, 2018

Jason Brownlee (2022), “Machine Learning Mastery”, 2022, <https://machinelearningmastery.com/dropout-for-regularizing-deep-neural-networks/>

Birol Yildiz, Ari Yezegel (2010), “FUNDAMENTAL ANALYSIS WITH ARTIFICIAL NEURAL NETWORK”, *The International Journal of Business and Finance Research*, The Institute for Business and Finance Research, vol. 4(1), pages 149-158.

Yajiao Tang, Zhenyu Song, Yulin Zhu, Huaiyu Yuan, Maozhang Hou, Junkai Ji, Cheng Tang, Jianqiang Li (2022), "A survey on machine learning models for financial time series forecasting", *Neurocomputing*, 512, 363-380

Saeede Anbaee Farimani, Majid Vafaei Jahan and Amin Milani Fard, (2022), "From Text Representation to Financial Market Prediction: A Literature Review", *Information* 2022, 13(10), 466; <https://doi.org/10.3390/info13100466>

Begüm Yılmaz (2022), "Stock Market Sentiment Analysis: How it works & 7 data sources", <https://research.aimultiple.com/sentiment-analysis-stock-market/>

Charalampos M. Liapis, Aikaterini Karanikola, Sotiris Kotsiantis, (2021): "A Multi-Method Survey on the Use of Sentiment Analysis in Multivariate Financial Time Series Forecasting", *Entropy*, 2021, 23(12), 1603; <https://doi.org/10.3390/e23121603>

Cory Mitchell (2022), "Market Sentiment Indicator: How It's Used in Analysis and Types", *Investopedia*, <https://www.investopedia.com/terms/s/sentimentindicator.asp>

Weiwei Jiang (2021), "Applications of deep learning in stock market prediction: Recent progress" 2021

(Zhang et al, 2022): Daniel Zhang, Nestor Maslej, Erik Brynjolfsson, John Etchemendy, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Michael Sellitto, Ellie Sakhaee, Yoav Shoham, Jack Clark, and Raymond Perrault, *The AI Index 2022 Annual Report*, 2022

(Sezer et al, 2020): Omer Berat Sezer, Mehmet Ugur Gudelek, Ahmet Murat Ozbayoglu, *Financial time series forecasting with deep learning: A systematic literature review: 2005–2019*, 2020

Yeong Hyeon Gu, Dong Jin, Helin Yin, Ri Zheng, Xianghua Piao and Seong Joon Yoo (2022), *Forecasting Agricultural Commodity Prices Using Dual Input Attention LSTM*, *Agriculture*, 2022, vol. 12, issue 2, 1-18

Ranjit Kumar Paul<sup>ID</sup>, Md. Yeasin<sup>ID</sup>, Pramod Kumar, Prabhakar Kumar, M. Balasubramanian, H. S. Roy<sup>1</sup>, A. K. Paul<sup>1</sup>, Ajit Gupta<sup>1</sup> (2022), *Machine learning techniques for forecasting agricultural prices: A case of brinjal in Odisha, India*, *PLoS ONE* 17(7): e0270553. <https://doi.org/10.1371/journal.pone.0270553>

(Jia & Yang, 2021): Fang Jia, Boli Yang, *Forecasting Volatility of Stock Index: Deep Learning Model with Likelihood-Based Loss Function*, 2021

Yuping Song, Xiaolong Tang, Hemin Wang, Zhiren Ma (2022), *Volatility forecasting for stock market incorporating macroeconomic variables based on GARCH-MIDAS and deep learning models*, *Journal of Forecasting*, 2023, vol. 42, issue 1, 51-59

Sotirios P. Chatzis, Vassilis Siakoulis, Anastasios Petropoulos, Evangelos Stavroulakis (2018), "Forecasting stock market crisis events using deep and statistical machine learning techniques", *Expert Systems with Applications*, ELSEVIER.

Zhenwei Li, Jing Han, Yuping Song (2019), "On the forecasting of high-frequency financial time series based on ARIMA model improved by deep learning," *Journal of Forecasting*, John Wiley & Sons, Ltd., vol.

Zahra Berradi, Mohamed Lazaar, Oussama Mahboub, Hicham Omara (2021), A Comprehensive Review of Artificial Intelligence Techniques in Financial Market, 2020 6th IEEE Congress on Information Science and Technology (CiSt) DOI: 10.1109/CiSt49399.2021.9357175

KONSTANTINOS BENIDIS ,SYAMA SUNDAR RANGAPURAM, VALENTIN FLUNKERT, YUYANG WANG,USA DANIELLE MADDIX,USA CANER TURKMEN, JAN GASTHAUS, MICHAEL BOHLKE-SCHNEIDER, DAVID SALINAS, LORENZO STELLA, FRANÇOIS-XAVIER AUBET, LAURENT CALLOT, TIM JANUSCHOWSKI(2018), “ Deep Learning for Time Series Forecasting: Tutorial andLiterature Survey”, ACM Comput. Surv., Vol. 55, No. 6, Article 121,

## – Παράρτημα

### AI FACTSHEET

<b>Model Name</b>	Text Sentiment Classifier				
<b>Overview</b>	This document is a FactSheet accompanying the <a href="#">Text Sentiment Classifier model</a> on IBM Developer <a href="#">Machine Learning eXchange</a> .				
<b>Purpose</b>	This model is able to detect whether an English-language sentence or set of sentences leans towards a positive or a negative sentiment.				
<b>Intended Domain</b>	The primary domain of this model is Natural Language Processing/Understanding.				
<b>Training Data</b>	<a href="#">IBM Claim Sentences Dataset</a> (Part of Project Debater).				
<b>Model Information</b>	<b>BERT</b> - Bidirectional Encoder Representations from Transformers ( <a href="#">paper</a> , <a href="#">implementation details</a> ). BERT makes use of Transformer, an attention mechanism that learns contextual relations between words (or sub-words) in a text.				
<b>Inputs and Outputs</b>	<b>Inputs:</b> An English language sentence or a batch of sentences as a JSON object. <b>Outputs:</b> A JSON with a "positive" and a "negative" score (which add up to 1.00) for each input sentence / batch.				
<b>Performance Metrics</b>	<table><thead><tr><th>Metric</th><th>Value</th></tr></thead><tbody><tr><td>Accuracy</td><td>94% on IBM Claim Sentences dataset</td></tr></tbody></table>	Metric	Value	Accuracy	94% on IBM Claim Sentences dataset
Metric	Value				
Accuracy	94% on IBM Claim Sentences dataset				
<b>Bias</b>	Possible bias due to the training dataset and possible bias associated with race and gender, but this has not been evaluated.				
<b>Robustness</b>	AI and ML models are susceptible to adversarial attacks where the output can be altered by a small number of changes in the input such as few characters or words, often overlooked by a human, or such that a human would still classify the input correctly.				
<b>Domain Shift</b>	The Domain Shift indicates how the model performs when scoring on a dataset other than the training dataset. It is a measure on how much the model can generalize across different domains. We report the accuracy metric on datasets from two other domains: - 81% accuracy on <a href="#">IMDb dataset</a> , containing sentiments from movie reviews - 83% accuracy on <a href="#">Sentiment140 dataset</a> , containing sentiments from Twitter tweets.				
<b>Test Data</b>	The data was split into train : val : test in 60 : 20 : 20 percent ratio.				
<b>Optimal Conditions</b>	<ul style="list-style-type: none"><li>• When the sentiment is objectively clear (for ex: swear words, harsh words, angry language for negative and mild language for positive).</li><li>• When the inputs are well-formed sentences similar to a typical "argument" in a debate or a well-articulated Wikipedia sentence.</li></ul>				
<b>Poor Conditions</b>	<ul style="list-style-type: none"><li>• Sarcasm</li><li>• Passive-aggressive statements</li><li>• Socially / racially offensive language</li><li>• Unknown terminology</li><li>• Sentences with both compliments and insults.</li></ul>				
<b>Explanation</b>	BERT does not provide explanations for its predictions.				
<b>Contact Information</b>	Any queries related to the operation of the MAX Text Sentiment Classifier model can be addressed on the <a href="#">model GitHub repo</a> .				