



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ

ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ ΚΑΙ ΥΠΗΡΕΣΙΕΣ**

**ΜΕΛΕΤΗ ΟΙΚΟΝΟΜΙΚΩΝ ΣΤΟΙΧΕΙΩΝ ΓΙΑ ΤΗΝ ΠΡΟΒΛΕΨΗ
ΣΥΜΠΕΡΙΦΟΡΑΣ ΤΟΥΣ ΣΤΗΝ ΣΥΓΧΡΟΝΗ ΟΙΚΟΝΟΜΙΚΗ
ΠΡΑΓΜΑΤΙΚΟΤΗΤΑ ΜΕ ΧΡΗΣΗ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ**

Από

ΚΥΡΙΑΚΟ ΚΥΡΙΑΚΟΥ

Υποβάλλεται

για την εκπλήρωση των προϋποθέσεων λήψης

Μεταπτυχιακού Διπλώματος

στην ειδικευση «Προηγμένα Πληροφοριακά Συστήματα»

του ΠΜΣ “Πληροφοριακά Συστήματα & Υπηρεσίες”

στο

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΜΑΙΟΣ 2024

Επιβλέπων/Επιβλέπουσα: ΜΙΧΑΗΛ ΦΙΛΙΠΠΑΚΗΣ

Copyright © Κυριακού Κυριάκος, 2024

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα.

Οι απόψεις και θέσεις που περιέχονται σε αυτήν την εργασία εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Πανεπιστημίου Πειραιώς.

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου, καθηγητή κο.Φιλιππάκη Μιχάλη, για την παρότρυνση, την ενθάρρυνση, την προθυμία του για συνεχείς συζητήσεις, την καθοδήγηση και τη βοήθεια του καθ' όλη τη διάρκεια εκπόνησης της παρούσας μεταπτυχιακής εργασίας. Θα ήθελα επιπλέον να εκφράσω την εκ των προτέρων εκτίμησή μου στα υπόλοιπα μέλη της Επιτροπής για τις παρατηρήσεις τους για τη μεταπτυχιακή μου εργασία.

Την εργασία θα ήθελα να αφιερώσω στους γονείς, στον αδελφό μου και στον επιστήθιο φίλο μου Τ.Χ, χωρίς την έμπρακτη υποστήριξη των οποίων δεν θα είχε ολοκληρωθεί αυτό το δύσκολο έργο.

Αθήνα, Μάιος 2024

Κυριακού Κυριάκος

Περιεχόμενα

Περίληψη.....	6
Summary.....	7
Κεφάλαιο 1. Εισαγωγή	8
Κεφάλαιο 2. Υπόβαθρο Έρευνας	10
2.1 Ιστορική αναδρομή.....	10
2.2 Βιβλιογραφική Ανασκόπηση	12
2.3 Περιγραφή Δεδομένων	13
Κεφάλαιο 3. Πιθανά προβλήματα -Δυσκολίες ερευνάς	21
3.1 Υπερπροσαρμογή (overfitting).....	21
3.1.1 Τρόποι αποτροπής υπερπροσαρμογής (overfitting):.....	22
3.2 Adam Optimizer:	23
3.2.1 Το μέρος του Adam στο overfitting.....	23
3.3 Η σημασία του πειράγματος των υπερπαραμετρών (hyperparameter -tuning).....	24
3.3.1 Προ επεξεργασία Δεδομένων.....	26
Κεφάλαιο 4. Τεχνικές που χρησιμοποιήθηκαν.....	30
4.1 Μοντέλο LSTM.....	30
4.2 Τμήματα Μοντέλου	31
4.3 XGBoost	33
4.4 MLP: Εισαγωγή στα Multilayered Perceptrons (MLP) στην Πρόβλεψη Χρονοσειρών	35
4.4.1. Σημασία στην πρόβλεψη χρονοσειρών:	36
4.4.2 Αρχιτεκτονική Μοντέλου: Αποκάλυψη του Ρόλου των Κρυφών Επιπέδων και Ενεργοποίηση ReLU	37
4.4.3 Υπερπαραμέτροι σε MLP μοντέλο για την πρόβλεψη χρονοσειρών:.....	39
4.5 Arima	40
4.5.1 Επιλογή μοντέλου	41
4.5.2 Αποκρυπτογράφηση AIC και PDQ στην ανάλυση χρονοσειρών.....	41
4.5.3. Η σημασία του AIC:	42
4.5.5. Ο Μαθηματικός Χορός του PDQ:	43
4.5.7. Τασεις	44
4.5.8 Προσαρμογή και αξιολόγηση μοντέλου (Model Fitting and Evaluation).....	46
4.6. Μετασχηματιστές (Transformers)	47

Κεφαλαίο 5 .Αποτελέσματα και σύγκριση	5.1 Αποτελέσματα MLP	50
	5.2 Αποτελέσματα ARIMA	52
	5.3 Αποτελέσματα XGBOOST	53
	5.4. Αποτελέσματα LSTM.....	56
Κεφάλαιο 6 . Πόρισμα.....		58
Βιβλιογραφία.....		60

ΛΙΣΤΑ ΕΙΚΟΝΩΝ

<i>Εικόνα 1Η ιστορία των νευρωνικών δικτύων[6] (ann, n.d.).....</i>	12
Εικόνα 2.Απεικόνιση των τιμών μέτοχων	15
Εικόνα 3.Γραφική αναπαράσταση τιμών από το 2016 εως το 2024.	15
Εικόνα 4.Γραφημα Τιμών κλεισίματος (2016-2024)	16
Εικόνα 5Autocorrelation γράφημα που χρησιμοποιήθηκε στην ερμηνεία των τιμών κλεισίματος στον αλγόριθμο ARIMA.	16
Εικόνα 6.Τιμες μετοχων (2016-2024)	17
Εικόνα 7.Τιμές μέτοχων εβδομαδιαίως	18
Εικόνα 8.Τιμές μέτοχων Μηνιαίως	19
Εικόνα 9.Θετική αυτοσυσχέτιση: Η διαγώνια γραμμή υποδηλώνει θετική σχέση μεταξύ των παρατηρήσεων και των υστερήσεων τους. Με άλλα λόγια, όταν αυξάνεται η υστέρηση (κινούμενη προς τα δεξιά κατά μήκος του άξονα x), αυξάνεται και η συσχέτιση μεταξύ των παρατηρήσεων και των υστερήσεων τους.	27
Εικόνα 10 . Απεικόνιση διαχωρισμού δεδομένων για το κομμάτι εκπαίδευσης και test.	28
Εικόνα 11.Αρχιτεκτονική μοντέλου [5]	48
Εικόνα 12Τιμές απωλειών στην διαδικασία εκπαίδευσης για 150 epochs.	50
Εικόνα 13.Μετρικές.....	51
Εικόνα 14Γράφημα αποτελεσμάτων προβλεψεων	52
Εικόνα 15Παράγει ένα γράφημα Plotly με δύο ίχνη: το ένα αντιπροσωπεύει τις πραγματικές τιμές κλεισίματος και το άλλο αντιπροσωπεύει τις προβλεπόμενες τιμές, επιτρέποντας την οπτική σύγκριση.	53
Εικόνα 16.Μετρικές.....	54
Εικόνα 17Απόκλιση στους τελευταίους μήνες του prediction	55
Εικόνα 18Απόκλιση στους τελευταίους μήνες του prediction	55
Εικόνα 19.Τιμές απωλειών στην διαδικασία εκπαίδευσης.....	56
Εικόνα 20.Μετρικές.....	57
Εικόνα 21.Πρόβλεψη πάνω στο target προβλέψεων.....	57

Περίληψη

Σε μια εποχή που χαρακτηρίζεται από την εκθετική ανάπτυξη δεδομένων και την αδιάκοπη πορεία της τεχνολογικής καινοτομίας, τα τεχνητά νευρωνικά δίκτυα (ΤΝΔ) έχουν αναδειχθεί ως μια μετασχηματιστική δύναμη στη σφαίρα της μηχανικής μάθησης και της τεχνητής νοημοσύνης (ΤΝ). Αυτά τα ευέλικτα υπολογιστικά μοντέλα, εμπνευσμένα από την πολύπλοκη δομή του ανθρώπινου εγκεφάλου, έχουν ταξιδέψει σε ένα συναρπαστικό ταξίδι από την ταπεινή αρχή τους για να γίνουν ο ακρογωνιαίος λίθος της σύγχρονης ανάλυσης δεδομένων, της αναγνώρισης προτύπων και της προγνωστικής μοντελοποίησης.

Σε αυτή τη δυναμική εποχή, όπου τα δεδομένα ρέουν άφθονα, η επίλυση μη γραμμικών προβλημάτων έχει καταστεί απόλυτη αναγκαιότητα, ιδιαίτερα στον χρηματοπιστωτικό τομέα. Το οικονομικό τοπίο παρουσιάζει μυριάδες προκλήσεις, πολλές από τις οποίες είναι εγγενώς μη γραμμικές και πολύπλοκες. Σε αυτό το πλαίσιο τα τεχνητά νευρωνικά δίκτυα λάμπουν ως ισχυρά εργαλεία για την ανάλυση περίπλοκων οικονομικών πληροφοριών.

Αυτή η ερευνητική εργασία διερευνά την εφαρμογή των τεχνητών νευρωνικών δικτύων στη σύγχρονη οικονομική πραγματικότητα χρησιμοποιώντας τεχνικές μηχανικής μάθησης και πρόβλεψη χρονοσειρών. Επιδιώκει να αποκαλύψει τις δυνατότητες των ΤΝΔ ως μέσο αντιμετώπισης των περίπλοκων προκλήσεων που αντιμετωπίζει σήμερα ο χρηματοπιστωτικός τομέας. Μέσω ενός συνδυασμού ιστορικού πλαισίου, θεωρητικών θεμελίων, πρακτικής εφαρμογής και κριτικής ανάλυσης, αυτή η διατριβή στοχεύει να συμβάλει στον εξελισσόμενο λόγο για την ενσωμάτωση της τεχνολογίας αιχμής στον οικονομικό τομέα. Τελικά, προσπαθεί να διαθέσει πληροφορίες που μπορούν να παρέχουν πιο ακριβείς οικονομικές προβλέψεις και τη λήψη αποφάσεων βάσει δεδομένων σε μια εποχή που ορίζεται από την αφθονία των δεδομένων και την τεχνολογική πρόοδο.

Summary

In an era characterized by exponential data growth and the relentless march of technological innovation, artificial neural networks (ANNs) have emerged as a transformative force in the realm of machine learning and artificial intelligence (AI). These flexible computational models, inspired by the complex structure of the human brain, have embarked on an exciting journey from their humble beginnings to become the cornerstone of modern data analysis, pattern recognition and predictive modeling. In this dynamic era where data flows abundantly, solving nonlinear problems has become an absolute necessity, especially in the financial sector. The financial landscape presents myriad challenges, many of which are inherently nonlinear and complex. In this context, artificial neural networks shine as powerful tools for analyzing complex financial information. This research paper explores the application of artificial neural networks to modern economic reality using machine learning techniques and time series prediction. It seeks to uncover the potential of AI as a means of addressing the complex challenges facing the financial sector today. Through a combination of historical context, theoretical foundations, practical application and critical analysis, this thesis aims to contribute to the evolving discourse on the integration of cutting-edge technology in the financial sector. Ultimately, it seeks to make available information that can provide more accurate economic forecasting and data-driven decision making in an era defined by data abundance and technological advances.

Κεφάλαιο 1. Εισαγωγή

Αυτή η διατριβή υποκινείται από τη συνειδητοποίηση ότι τα τεχνητά νευρωνικά δίκτυα (ANN) διαθέτουν απaráμιλλη ευελιξία και προσαρμοστικότητα, καθιστώντας τα απαραίτητα εργαλεία για την αντιμετώπιση των πολύπλοκων προβλημάτων που αντιμετωπίζει ο χρηματοπιστωτικός κλάδος. Μέσα σε ένα ολοένα και πιο περίπλοκο χρηματοοικονομικό περιβάλλον που χαρακτηρίζεται από μη γραμμικά ζητήματα, υπάρχει μια αυξανόμενη ανάγκη για προηγμένα μέσα που μπορούν να προσδιορίσουν διακριτικά μοτίβα. Αυτή η εργασία στοχεύει να αποκαλύψει τις δυνατότητες διαφορετικών σχεδίων νευρωνικών δικτύων και να δώσει μια εικόνα για το πόσο καλά μπορούν να χειριστούν την περίπλοκη φύση των οικονομικών δεδομένων χρησιμοποιώντας έναν εμπειρικό φακό.

Η έρευνα αυτή έχει διπλό σκοπό. Καταρχάς, στοχεύει να δώσει μια διεξοδική θεωρητική βάση και ιστορικό υπόβαθρο, ακολουθώντας την ανάπτυξη των ANN από την αρχή τους έως τη σημερινή τους θέση ως ακρογωνιαίο λίθο της σύγχρονης ανάλυσης δεδομένων. Δεύτερον, ελπίζει να προωθήσει την τρέχουσα συζήτηση σχετικά με την ενσωμάτωση τεχνολογίας αιχμής στον χρηματοπιστωτικό κλάδο παρέχοντας χρήσιμες πληροφορίες σχετικά με τη χρήση των νευρωνικών δικτύων.

Η έρευνα χρησιμοποιεί μια εμπειρική τεχνική για την επίτευξη αυτών των στόχων. Διαφορετικές τοπολογίες νευρωνικών δικτύων όπως Μακροπρόθεσμη μνήμη (Long-Short-Term Memory/LSTM) και Πολυστρωματικό Perceptron (Multi-Layer Perceptron/MLP) καθώς επίσης η στατιστική μέθοδος Αυτοπαλινδρομικός ολοκληρωμένος κινούμενος μέσος όρος (AutoRegressive Integrated Moving Average/ARIMA) και ο αλγόριθμος XGBoost είναι προσεκτικά σχεδιασμένες και γεμάτες με ποικιλία συνόλων δεδομένων που αντικατοπτρίζουν την περίπλοκη φύση των οικονομικών δεδομένων. Αυτή η αναλυτική προσέγγιση διευκολύνει την αξιολόγηση του πόσο καλά αποδίδει κάθε μοντέλο στον χειρισμό της πολυπλοκότητας των οικονομικών δεδομένων με διαφοροποιημένο τρόπο.

Μέσω μιας διεξοδικής συγκριτικής εξέτασης πολλών νευρωνικών δικτύων, αυτή η μελέτη στοχεύει στον προσδιορισμό του βέλτιστου μοντέλου για την αποκρυπτογράφηση της πολυπλοκότητας των οικονομικών δεδομένων. Η παροχή πρακτικών πληροφοριών στους ενδιαφερόμενους του χρηματοπιστωτικού τομέα είναι υψίστης σημασίας, καθώς διευκολύνει την ακριβέστερη πρόβλεψη και τη λήψη αποφάσεων βάσει δεδομένων. Αυτή η έρευνα στοχεύει να συμβάλει στην ανάπτυξη της συνέργειας μεταξύ τεχνολογίας και χρηματοοικονομικά σε μια εποχή που χαρακτηρίζεται από πληθώρα δεδομένων και τεχνολογικής καινοτομίας.

Η διατριβή αναπτύσσεται μεθοδικά, παρασύροντας τους αναγνώστες μέσα από θεωρητικές βάσεις, εφαρμογές πραγματικού κόσμου, ιστορικό υπόβαθρο και κριτική ανάλυση. Κάθε ενότητα έχει σχεδιαστεί προσεκτικά για να παρέχει στους αναγνώστες μια ολοκληρωμένη κατανόηση του τρόπου με τον οποίο η τεχνολογία αιχμής έχει ενσωματωθεί στον χρηματοπιστωτικό κλάδο. Ο τρόπος οργάνωσης της διατριβής διασφαλίζει ότι επιτυγχάνει τον συγκεκριμένο στόχο της, που είναι να αναδείξει τις δυνατότητες των τεχνητών νευρωνικών δικτύων (ANN) και πόσο σημαντικά είναι για τη βελτίωση της χρηματοοικονομικής πρόβλεψης και λήψης αποφάσεων σε ένα ταχέως μεταβαλλόμενο τεχνολογικό περιβάλλον.

Κεφάλαιο 2. Υπόβαθρο Έρευνας

2.1 Ιστορική αναδρομή

Η πρόοδος των νευρωνικών δικτύων κατά τη διάρκεια των δεκαετιών χαρακτηρίζεται τόσο από σημαντικές ανακαλύψεις όσο και από περιόδους σκεπτικισμού, καταδεικνύοντας τη συνεχή αναζήτηση της τεχνητής νοημοσύνης. Οι Warren McCulloch και Walter Pitts συνέλαβαν τα νευρωνικά δίκτυα χρησιμοποιώντας ηλεκτρικά κυκλώματα τη δεκαετία του 1940, θέτοντας τα θεμέλια για περαιτέρω έρευνα. Το 1949, ο Donald Hebb διεξήγαγε έρευνα που έριξε φως στην ενίσχυση των νευρωνικών κυκλωμάτων με τη συχνή χρήση, προωθώντας έτσι την κατανόησή μας για τον τρόπο λειτουργίας του εγκεφάλου. Όμως οι πρώτες προσπάθειες αντιμετώπισαν δυσκολίες, μεταξύ των οποίων ήταν οι μεγάλοι χρόνοι υπολογισμού και οι περιορισμένες ικανότητες μάθησης. Αυτό έγινε ιδιαίτερα αισθητό στις προσπάθειες μίμησης απλών κυκλωμάτων αποκλεισμού-ή-αντιγραφής Boole τη δεκαετία του 1960.

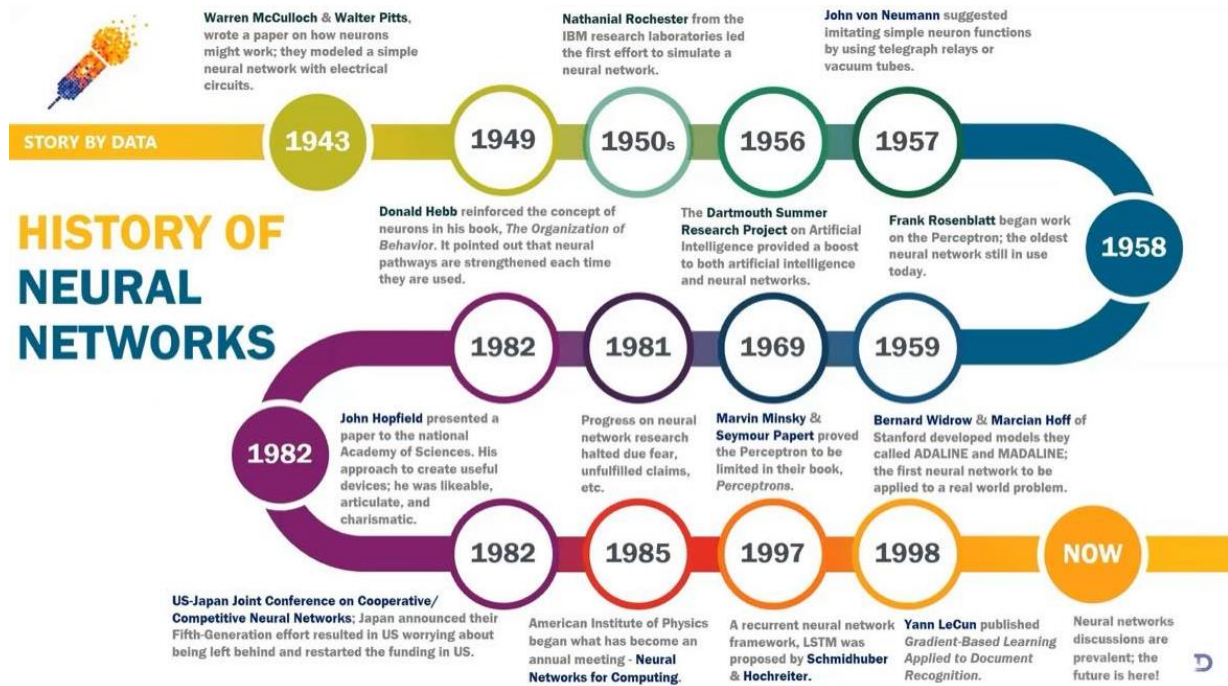
Το βιβλίο "Perceptrons" των Marvin Minsky και Seymour Papert του 1969 αποτέλεσε σημείο καμπής στον τομέα, καθώς αξιολόγησε κριτικά την τεχνική του ενός perceptron και τις δυσκολίες της κατά την κλιμάκωση σε δίκτυα πολλαπλών επιπέδων. Αυτή η κριτική σηματοδότησε την έναρξη μιας σχεδόν δεκαετούς εποχής σκεπτικισμού, γνωστής ως "χειμώνας της τεχνητής νοημοσύνης", κατά τη διάρκεια της οποίας η χρηματοδότηση και η έρευνα για τα νευρωνικά δίκτυα μειώθηκε. Ωστόσο, το 1982 παρατηρήθηκε η αναζωπύρωση του ενδιαφέροντος και της χρηματοδότησης για το θέμα χάρη στην ανάπτυξη των δικτύων Hopfield από τον Jon Hopfield και την αφοσίωση της Ιαπωνίας στο πρόγραμμα νευρωνικών δικτύων πέμπτης γενιάς.

Η εκ νέου ανακάλυψη του οπίσθιου πολλαπλασιασμού backpropagation, μιας μεθόδου που εφευρέθηκε τη δεκαετία του 1960 αλλά δεν ήταν γνωστή μέχρι τη δεκαετία του 1980, ήταν μια από τις σημαντικότερες στιγμές καμπής. Η Gradient Descent και ο οπίσθιος πολλαπλασιασμός (Backpropagation) μαζί άλλαξαν την εκπαίδευση των νευρωνικών δικτύων επιτρέποντας αποτελεσματικές αλλαγές στα βάρη ανάλογα με την κλίση της συνάρτησης κόστους. Ο σχηματισμός ετήσιων συνεδρίων και παγκόσμιων συνεργασιών ως αποτέλεσμα αυτού του ανανεωμένου ενδιαφέροντος και της χρηματοδότησης επιτάχυνε την πρόοδο της έρευνας των νευρωνικών δικτύων. Οι ομιλίες για τα νευρωνικά δίκτυα είναι δημοφιλείς στις μέρες μας, αναδεικνύοντας τη συνεχή ανάπτυξη της τεχνολογίας και την αυξανόμενη σημασία της στα έργα μηχανικής μάθησης και τεχνητής νοημοσύνης.

Αυτό που ακολουθήσε ήταν το concept RNN το 1986 στο οποίο βασίστηκε το μοντέλο LSTM .Η αρχιτεκτονική των recurrent neural network (RNN) παρουσιάστηκε για πρώτη φορά από τους Sepp Hochreiter και Jürgen Schmidhuber στα τέλη της δεκαετίας του 1990, και τότε δημιουργήθηκαν τα δίκτυα Long Short-Term Memory (LSTM). Λόγω του προβλήματος της εξαφανιζόμενης κλίσης(vanishing gradient), οι εξαρτήσεις μεγάλης εμβέλειας ήταν δύσκολο να συλληφθούν από τα τυπικά RNN- ωστόσο, τα δίκτυα LSTM κατασκευάστηκαν για να αντιμετωπίσουν αυτό το πρόβλημα. Στην πρωτοποριακή δημοσίευσή τους το 1997, οι Hochreiter και Schmidhuber πρότειναν το LSTM ως έναν τρόπο αποτελεσματικής καταγραφής και αποθήκευσης δεδομένων για μεγάλες χρονικές περιόδους.

Η ικανότητα των δικτύων LSTM να διαχειρίζονται ακολουθίες διαφορετικού μήκους και να αποτυπώνουν την αλληλεξάρτηση σε μεγάλους χρονικούς ορίζοντες είναι ένα από τα διακριτικά χαρακτηριστικά τους. Πολλές διαφορετικές εφαρμογές, όπως η αναγνώριση ομιλίας, η πρόβλεψη χρονοσειρών, η επεξεργασία φυσικής γλώσσας και η μοντελοποίηση διαδοχικών δεδομένων, έχουν υιοθετήσει ευρέως αυτό το χαρακτηριστικό. Τα δίκτυα LSTM έχουν βελτιωθεί σημαντικά όσον αφορά την απόδοση και την προσαρμοστικότητα με διάφορες βελτιώσεις και βελτιστοποιήσεις από την εισαγωγή τους. Προκειμένου να επιλύσουν συγκεκριμένα ζητήματα και να ενισχύσουν την αποτελεσματικότητα του μοντέλου, οι ερευνητές και οι επαγγελματίες έχουν διερευνήσει μια ποικιλία επεκτάσεων και αλλαγών στον αρχικό σχεδιασμό LSTM, όπως τα gated recurrent units (GRUs), οι συνδέσεις κλειδαρότρυπα (peephole) και οι μέθοδοι προσοχής.

Η βαθιά μάθηση σημείωσε σημαντική πρόοδο με την εμφάνιση των δικτύων LSTM, τα οποία επέτρεψαν σε ακαδημαϊκούς και επαγγελματίες να χειριστούν με μεγαλύτερη ακρίβεια και αποτελεσματικότητα δύσκολες εργασίες διαδοχικών δεδομένων. Στη σύγχρονη εποχή, τα δίκτυα LSTM συνεχίζουν να αποτελούν τη βάση των αρχιτεκτονικών των επαναλαμβανόμενων νευρωνικών δικτύων και να ωθούν τις εξελίξεις σε τομείς που απαιτούν ισχυρές ικανότητες διαδοχικής μοντελοποίησης και πρόβλεψης.



Εικόνα 1Η ιστορία των νευρωνικών δικτύων[6] (ann, n.d.)

2.2 Βιβλιογραφική Ανασκόπηση

Για τους επενδυτές, τους διαχειριστές κεφαλαίων και άλλους χρηματοπιστωτικούς οργανισμούς, είναι ζωτικής σημασίας η ακριβής εκτίμηση των τιμών των περιουσιακών στοιχείων, όπως των μετοχών, των ομολόγων, των δικαιωμάτων προαίρεσης και των συναλλαγματικών ισοτιμιών. Ως αποτέλεσμα, συνεχίζονται οι συνεχείς προσπάθειες για τη βελτίωση της ακρίβειας των χρηματοοικονομικών μοντέλων πρόβλεψης. Ακόμα κι έτσι, η έρευνα σημειώνει ότι η επίτευξη ακριβών χρηματοοικονομικών προβλέψεων παραμένει πρόκληση, ακόμα και με τις συνεχείς προσπάθειες βελτίωσης. Μια προσέγγιση που πραγματοποιήθηκε στην ερευνά του Khashei, M., & Hajirahimi ([1] Khashei), αποσκοπούσε στην σύζευξη και σύγκριση δυο μοντελων ως ένα hybrid. Στην παρούσα μελέτη αξιολογείται η προγνωστική αποτελεσματικότητα δύο υβριδικών γραμμικών/μη γραμμικών μοντέλων, ARIMA-MLP και MLP-ARIMA, τα οποία

αναμένεται να αποδώσουν αποτελέσματα ανώτερα από τουλάχιστον ένα από τα μοντέλα που τα απαρτίζουν και είναι απίθανο να έχουν χειρότερες επιδόσεις από όλα τα επιμέρους στοιχεία. Η σύγκριση μεταξύ των υβριδικών μοντέλων της σειράς αποκαλύπτει ότι το MLP-ARIMA γενικά υπερτερεί του ARIMA-MLP, παρουσιάζοντας βελτιώσεις 16,50%, 2,57%, 2,57% και 3,71% όσον αφορά τα MAE, MSE, MAPE και RMSE τόσο στα σύνολα δεδομένων εκπαίδευσης όσο και στα σύνολα δεδομένων δοκιμής. Παρά την επικρατούσα χρήση των μοντέλων ARIMA-MLP στη βιβλιογραφία, τα αποτελέσματα αυτά συνηγορούν υπέρ της εξέτασης του MLP-ARIMA ως βιώσιμης εναλλακτικής λύσης, ιδίως στην πρόβλεψη χρηματοοικονομικών χρονοσειρών. Μια ακόμα ερευνητική εργασία του ([2] Zhang, n.d.) συνηγορεί υπέρ μιας συνδυασμένης προσέγγισης για την πρόβλεψη χρονολογικών σειρών με την ενσωμάτωση του γραμμικού μοντέλου ARIMA με το μη γραμμικό μοντέλο ANN. Αυτό το υβριδικό μοντέλο αξιοποιεί τα διακριτά πλεονεκτήματα του ARIMA και του ANN στη γραμμική και μη γραμμική μοντελοποίηση, αντίστοιχα, για να συλλάβει ποικίλες μορφές σχέσεων σε δεδομένα χρονοσειρών. Τα εμπειρικά ευρήματα καταδεικνύουν την ανώτερη απόδοση πρόβλεψης του υβριδικού μοντέλου σε σύγκριση με τα μεμονωμένα μοντέλα που χρησιμοποιούνται μεμονωμένα. Όσον αφορά τα LSTM ενώ το βαθύ νευρωνικό δίκτυο παρουσιάζει αποτελεσματικότητα στην πρόβλεψη χρονοσειρών και αντιμετωπίζει το ζήτημα της εξαφάνισης της κλίσης που είναι εγγενές στα παραδοσιακά RNN, η τρέχουσα ακρίβεια πρόβλεψής του παραμένει ένα σημείο προς βελτίωση. Οι χρονοσειρές του χρηματιστηρίου υπόκεινται σε πολυάριθμους θορύβους, με πλήθος παραγόντων να επηρεάζουν τις τάσεις τους ([3] Yan, n.d.). Τέλος, η ερευνά Yun, K. K., Yoon, S. W., & Won, D. ([4] Yun, n.d.) αποδεικνύει την αποτελεσματικότητα του αλγορίθμου GA-XGBoost στην πρόβλεψη της κατεύθυνσης των τιμών των μετοχών. Η μελέτη επιτυγχάνει ένα λιτό βέλτιστο σύνολο χαρακτηριστικών με μόνο 33 χαρακτηριστικά, υπογραμμίζοντας τη σημασία της επιλογής χαρακτηριστικών για τον μετριασμό του προβλήματος της διαστατικότητας. Ενώ η μελέτη αναγνωρίζει τις δυνατότητες για περαιτέρω διερεύνηση των τεχνικών δεικτών και την εφαρμογή τεχνικών βελτιστοποίησης του μοντέλου, υπογραμμίζει τον καθοριστικό ρόλο της μηχανικής των χαρακτηριστικών για την επίτευξη επιτυχημένων αποτελεσμάτων πρόβλεψης.

2.3 Περιγραφή Δεδομένων

Στην παρούσα διατριβή, χρησιμοποιήθηκαν οικονομικά δεδομένα που προέρχονται από τον ιστότοπο Yahoo Finance για την εξαγωγή σχετικών χαρακτηριστικών για την προγνωστική μοντελοποίηση με νευρωνικά δίκτυα. Η επιλογή των χρηματοοικονομικών δεδομένων, ιδίως των τιμών των μετοχών, έγινε για να διευκολυνθεί μια ολοκληρωμένη σύγκριση των διαφόρων προσεγγίσεων μοντελοποίησης στο πλαίσιο του εργαστηρίου. Η πρωταρχική ημερομηνία-στόχος που επιλέχθηκε για την παρούσα ανάλυση ήταν η "2016-06-14", που χρησιμεύει ως σημείο αναφοράς για την αξιολόγηση και τη σύγκριση των

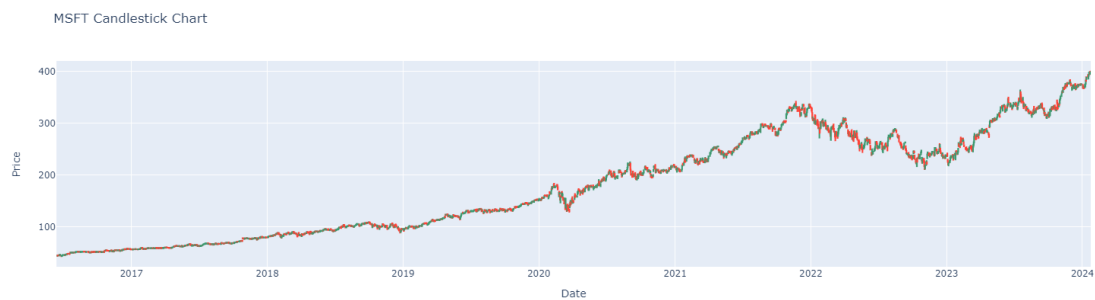
μοντέλων. Επιπλέον, χρησιμοποιήθηκαν πολλαπλές ημερομηνίες ζήτησης για την αντιπροσώπευση διαφορετικών στόχων και σεναρίων πρόβλεψης, επιτρέποντας την ενδελεχή αξιολόγηση των επιδόσεων του μοντέλου σε διαφορετικές χρονικές περιόδους και συνθήκες της αγοράς. Με την ενσωμάτωση ενός εύρους ημερομηνιών ζήτησης, η μελέτη είχε ως στόχο να αποτυπώσει τη δυναμική φύση των χρηματοπιστωτικών αγορών και να αξιολογήσει την ανθεκτικότητα των μοντέλων πρόβλεψης σε διαφορετικά χρονικά πλαίσια. Η προσέγγιση αυτή επέτρεψε μια ολοκληρωμένη αξιολόγηση της ακρίβειας του μοντέλου, της ικανότητας γενίκευσης και της καταλληλότητας για εφαρμογές στον πραγματικό κόσμο. Συνολικά, η αξιοποίηση σχολαστικά επιλεγμένων συνόλων χρηματοοικονομικών δεδομένων λειτούργησε ως θεμελιώδες στοιχείο στη διερεύνηση των τεχνικών προγνωστικών μοντέλων και της αποτελεσματικότητάς τους σε εργασίες χρηματοοικονομικής πρόβλεψης.

Οι ιστορικές τιμές των μετοχών, συμπεριλαμβανομένων των open, high, low, close , επιλέχθηκαν ως χαρακτηριστικά για την εργασία μοντελοποίησης πρόβλεψης. Αυτά τα χαρακτηριστικά επιλέχθηκαν λόγω της σημασίας τους για την καταγραφή των υποκείμενων προτύπων και τάσεων στα δεδομένα της χρηματιστηριακής αγοράς. Η τιμή ανοίγματος (open) αντικατοπτρίζει την αρχική τιμή διαπραγμάτευσης μιας μετοχής στην αρχή μιας συνεδρίασης διαπραγμάτευσης, παρέχοντας εικόνα του κλίματος των επενδυτών και των προσδοκιών της αγοράς. Η υψηλότερη (high) και η χαμηλότερη τιμή (low) αντιπροσωπεύουν το ανώτατο και το κατώτατο επίπεδο στο οποίο έφτασε η μετοχή κατά τη διάρκεια της συνεδρίασης, υποδεικνύοντας το εύρος της κίνησης των τιμών και την πιθανή μεταβλητότητα. Τέλος, η τιμή κλεισίματος (close), η οποία είναι η τελική τιμή διαπραγμάτευσης που καταγράφεται στο τέλος της συνεδρίασης, χρησιμεύει ως κρίσιμο μέτρο για την αξιολόγηση της συνολικής απόδοσης και του κλίματος προς τη μετοχή. Με την ενσωμάτωση αυτών των χαρακτηριστικών στα μοντέλα πρόβλεψης, στόχος είναι η αξιοποίηση των ιστορικών κινήσεων και τάσεων των τιμών για την ακριβή πρόβλεψη των μελλοντικών τιμών των μετοχών. Επιπλέον, αξιοποιήθηκαν χαρακτηριστικά υστέρησης που προέρχονται από την τιμή κλεισίματος για την καταγραφή των χρονικών εξαρτήσεων και των ιστορικών προτύπων, ενισχύοντας περαιτέρω την προβλεπτική ικανότητα των μοντέλων.

Currency in USD

Date	Open	High	Low	Close*	Adj Close**
Feb 23, 2024	415.67	415.86	408.97	410.34	410.34
Feb 22, 2024	410.19	412.83	408.57	411.65	411.65
Feb 21, 2024	400.17	402.29	397.22	402.18	402.18
Feb 20, 2024	403.24	404.49	398.01	402.79	402.79
Feb 16, 2024	407.96	408.29	403.44	404.06	404.06
Feb 15, 2024	408.14	409.13	404.29	406.56	406.56
Feb 14, 2024	408.07	409.84	404.57	409.49	409.49
Feb 14, 2024	0.75 Dividend				

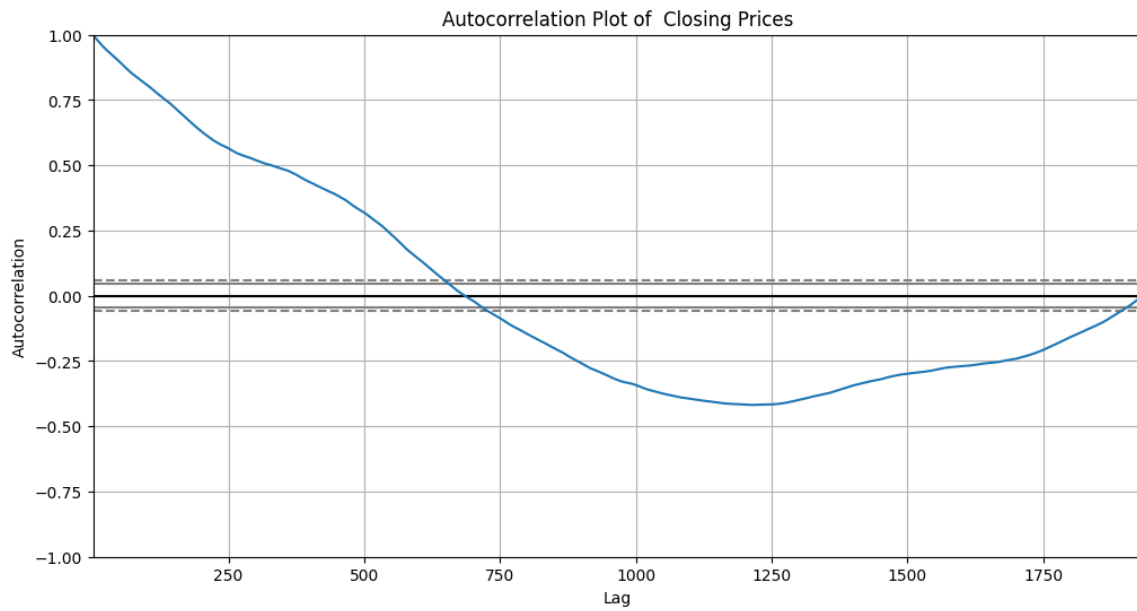
Εικόνα 2. Απεικόνιση των τιμών μετοχών



Εικόνα 3. Γραφική αναπαράσταση τιμών από το 2016 έως το 2024.

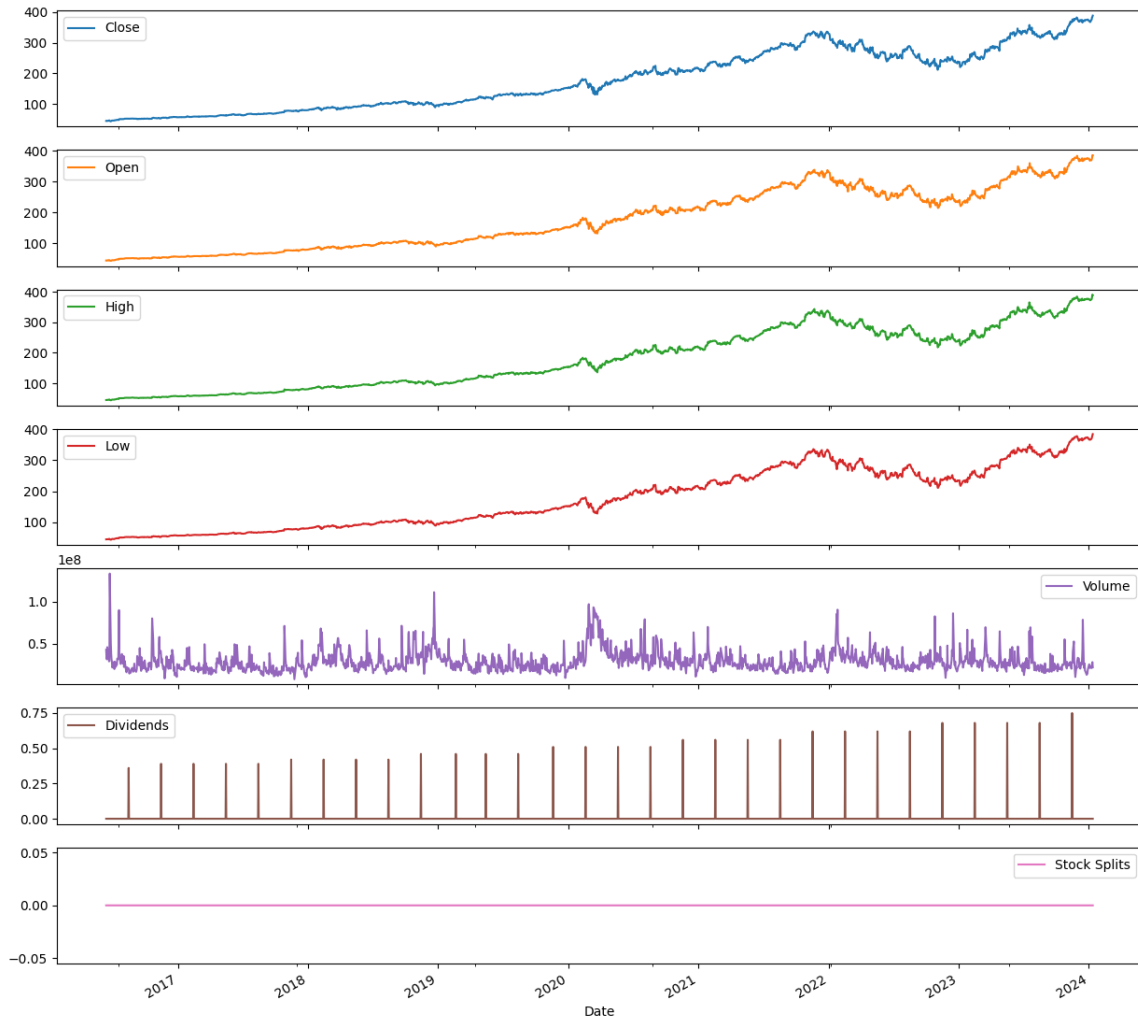


Εικόνα 4.Γραφημα Τιμών κλεισίματος (2016-2024)



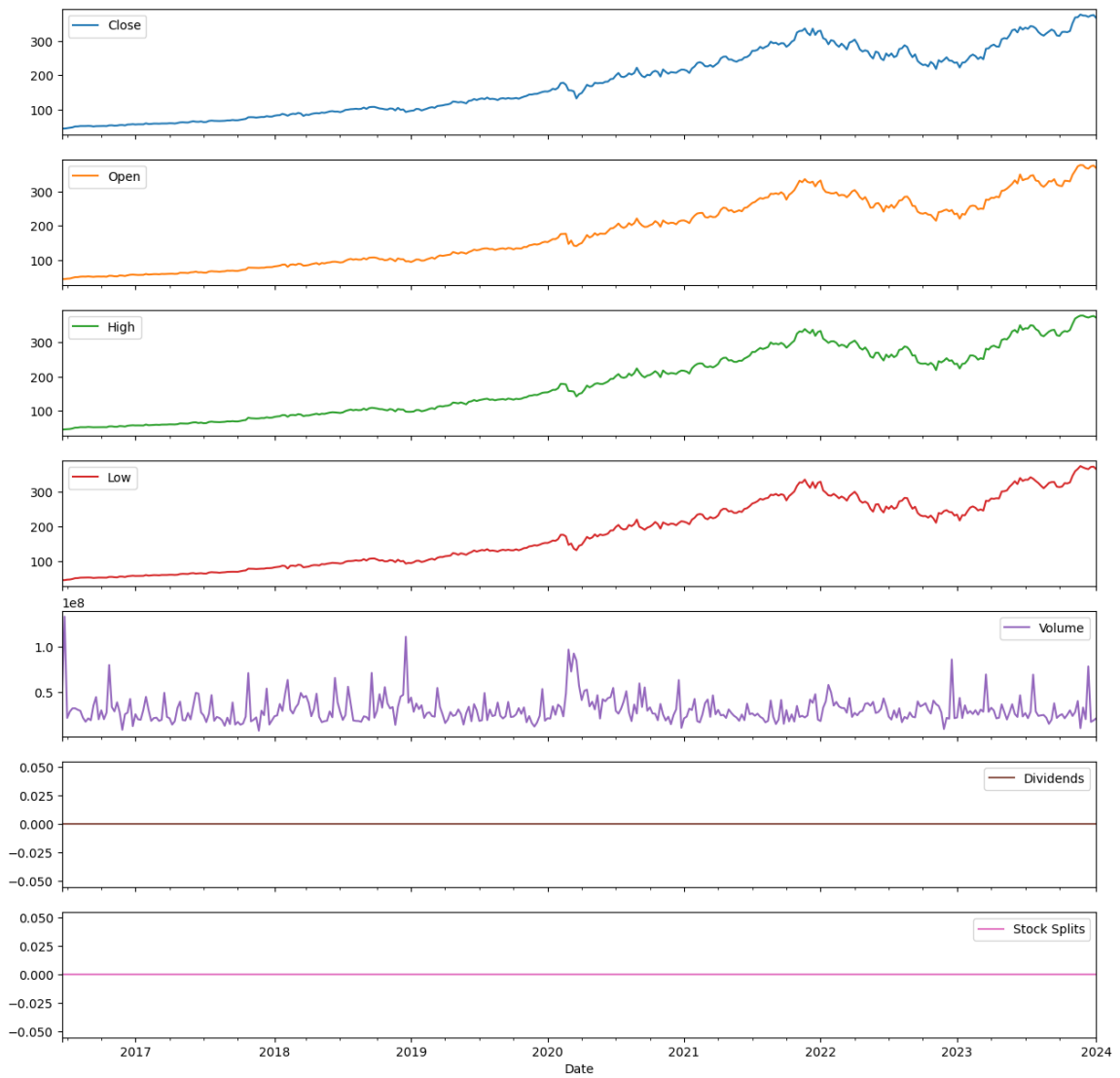
Εικόνα 5 Autocorrelation γράφημα που χρησιμοποιήθηκε στην ερμηνεία των τιμών κλεισίματος στον αλγόριθμο ARIMA.

MICROSOFT stock attributes from 2016 to 2023

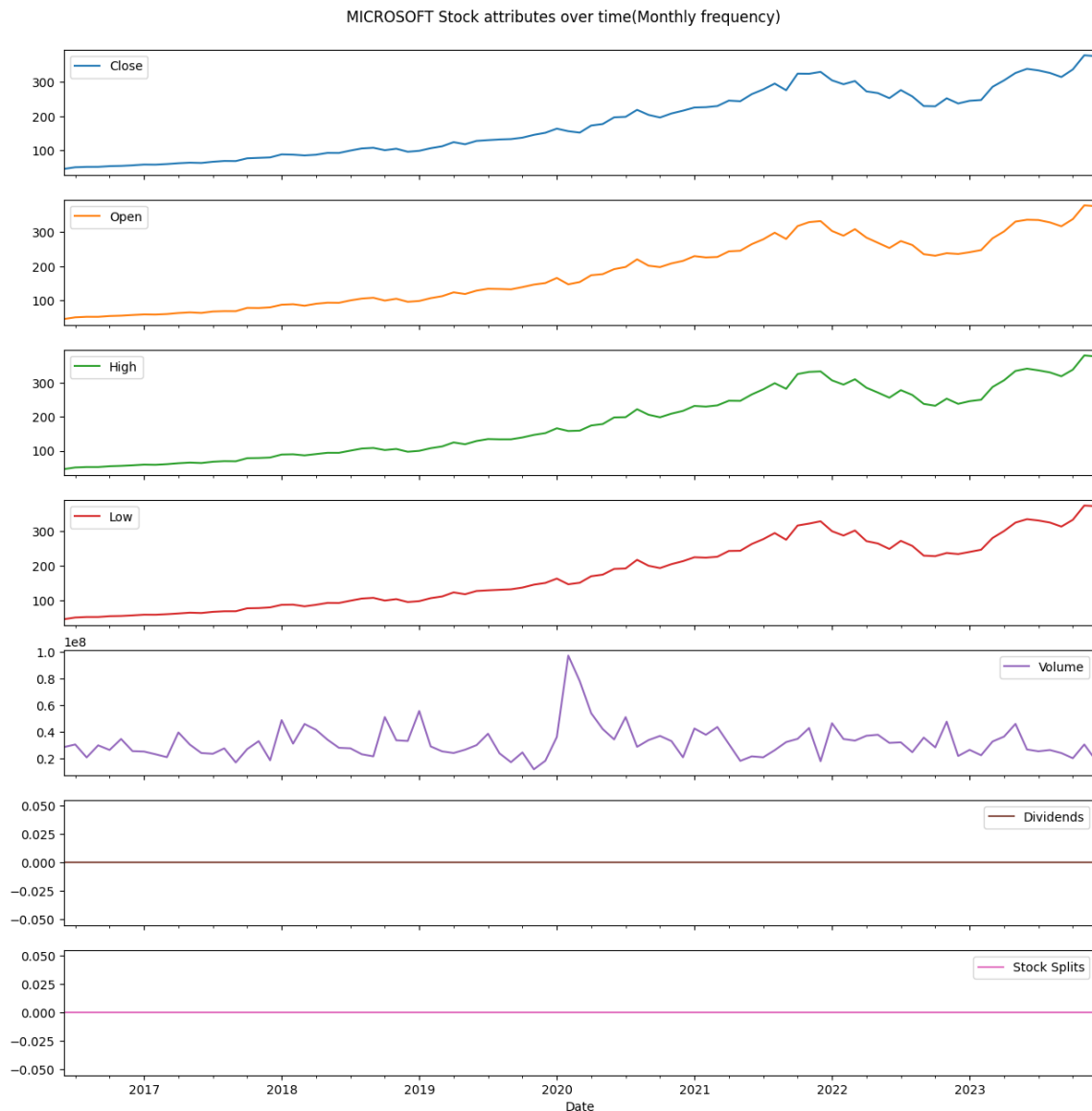


Εικόνα 6. Τιμές μετοχών (2016-2024)

MICROSOFT Stock attributes over time(Weekly frequency)



Εικόνα 7. Τιμές μετοχών εβδομαδιαίως



Εικόνα 8. Τιμές μέτοχων Μηνιαίως

Γραφικές απεικονίσεις των τιμών κλεισίματος ,ανοίγματος καθώς επίσης και των ψηλότερων και χαμηλότερων τιμών. Τα συγκεκριμένα γραφήματα είναι απαραίτητα για την ερμηνεία των δεδομένων προκειμένου να ακολουθήσει η σωστή επιλογή των χαρακτηριστικών (features) που θα χρησιμοποιηθούν στο κομμάτι της προ επεξεργασίας δεδομένων .Αναλυτικότερα ,εκ πρώτης όψεως γίνεται αντιληπτό πως οι αιχμές (spikes) που εμφανίζονται δεν υποδηλώνουν κάποια στασιμότητα αλλά αντίθετα τυχαιότητα (randomness) καθώς ενδεχομένως επηρεάζονται από μη σχετικούς οικονομικούς παράγοντες (covid ,πόλεμος κτλ.) Στο πλαίσιο χρηματοοικονομικών δεδομένων, όπως οι

τιμές κλεισίματος, η αυτοσυσχέτιση (autocorrelation) βοηθά στον εντοπισμό μοτίβων ή εξαρτήσεων μεταξύ παρελθουσών και τρεχουσών τιμών(εικονα5) . Στην συγκεκριμένη έρευνα παρατηρήθηκε ότι οι τιμές αυτοσυσχέτισης μειώνονται σταδιακά όσο αυξάνεται η υστέρηση(lag). Αυτό δείχνει ότι οι παρελθούσες τιμές προβλέπουν κάπως τις μελλοντικές τιμές, αλλά αυτή η προβλεψιμότητα μειώνεται όσο αυξάνεται η χρονική υστέρηση. Τα διαγράμματα αυτοσυσχέτισης είναι χρήσιμα για την κατανόηση των χρονικών εξαρτήσεων σε δεδομένα χρονοσειρών και μπορούν να ενημερώσουν για την επιλογή κατάλληλων μοντέλων, όπως το ARIMA, τα οποία μοντελοποιούν ρητά αυτές τις εξαρτήσεις.

Κεφάλαιο 3.

Πιθανά προβλήματα -Δυσκολίες ερευνάς

3.1 Υπερπροσαρμογή (overfitting)

Ένα τυπικό πρόβλημα στη μηχανική μάθηση είναι η υπερβολική προσαρμογή, όταν ένα μοντέλο μαθαίνει πολύ καλά τα δεδομένα εκπαίδευσης, συλλέγοντας χαρακτηριστικά και θόρυβο μοναδικά για αυτό το σύνολο δεδομένων, αλλά δεν μπορεί να τα γενικεύσει καλά σε νέα, άγνωστα δεδομένα. Εν ολίγοις, ένα μοντέλο υπερπροσαρμογής (overfitting) έχει χαμηλή απόδοση σε άγνωστα δεδομένα ενώ τα πάει εξαιρετικά καλά στο σετ προπόνησης. Όταν συμβαίνει αυτό το φαινόμενο, η απόδοση του συνόλου δεδομένων δοκιμής ή επικύρωσης είναι συνήθως πολύ χειρότερη παρά την υψηλή ακρίβεια εκπαίδευσης.

Γιατί συμβαίνει overfitting:

- *Σύνθετη Αρχιτεκτονική Μοντέλων:*
Το overfitting είναι πιο πιθανό σε μοντέλα με υπερβολική πολυπλοκότητα, όπως πολλά επίπεδα ή παραμέτρους. Αντί να κατανοεί τα υποκείμενα μοτίβα, μπορεί να θυμάται μόνο το σετ προπόνησης.
- *Ανεπαρκή Δεδομένα:*
Ένα πολύπλοκο μοντέλο μπορεί να μάθει εύκολα όλες τις λεπτομέρειες, συμπεριλαμβανομένου του θορύβου, από μια μικρή ποσότητα δεδομένων εκπαίδευσης. Η υπερβολική τοποθέτηση μπορεί να οφείλεται σε ανεπαρκή δεδομένα, ιδιαίτερα εάν το μοντέλο είναι πιο ικανό από ό,τι απαιτείται για την εργασία.
- *Έλλειψη τακτοποίησης:*
Οι μέθοδοι τακτοποίησης όπως η μείωση του βάρους και η εγκατάλειψη αποσκοπούν στο να σταματήσουν την υπερβολική προσαρμογή (overfitting) θέτοντας περιορισμούς στις παραμέτρους του μοντέλου. Η πιθανότητα overfitting αυξάνεται όταν η τακτοποίηση είναι ακατάλληλη.

3.1.1 Τρόποι αποτροπής υπερπροσαρμογής (overfitting):

- *Επαύξηση Δεδομένων (Data Augmentation):*
Χρησιμοποιώντας στρατηγικές όπως η αύξηση δεδομένων, το αποτελεσματικό μέγεθος του συνόλου δεδομένων εκπαίδευσης μπορεί να αυξηθεί. Αυτό συνεπάγεται τη δημιουργία παραλλαγών από τα διαθέσιμα δεδομένα, μειώνοντας επομένως την πιθανότητα το μοντέλο να δεσμεύσει συγκεκριμένες περιπτώσεις στη μνήμη.
- *Διασταυρωμένη επικύρωση (Cross-Validation):*
Για την αξιολόγηση της απόδοσης του μοντέλου σε πολλά υποσύνολα δεδομένων, αρωγός καθίσταται διασταυρούμενη επικύρωση. Αυτό βοηθά στην αξιολόγηση της απόδοσης γενίκευσης του μοντέλου σε διάφορα τμήματα δεδομένων.
- *Τεχνικές τακτοποίησης(Regularization Technique):*
Η χρήση Regularization Techniques διευκολύνει στο να μην εξαρτάται υπερβολικά το μοντέλο από συγκεκριμένα χαρακτηριστικά, όπως η εγκατάλειψη, η οποία αφαιρεί τυχαία τις συνδέσεις κατά τη διάρκεια της εκπαίδευσης.
- *Απλούστερη Αρχιτεκτονική Μοντέλων:*
Μια αρχιτεκτονική μοντέλου που είναι πιο απλή και έχει λιγότερα επίπεδα ή παραμέτρους πολλές φορές οδηγεί στο να υπάρχουν λιγότερες πιθανότητες το μοντέλο να ταιριάζει με το σετ εκπαίδευσης με υπερβολική ακρίβεια.
- *Πρόωρη διακοπή:*
Καθ' όλη τη διάρκεια της εκπαίδευσης, η παρακολούθηση απόδοσης του μοντέλου σε ένα σετ επικύρωσης(validation) έχει σαν αποτέλεσμα την αποφυγή της υπερβολικής προσαρμογής.

3.2 Adam Optimizer:

Οι προσαρμοστικοί αλγόριθμοι βελτιστοποίησης, όπως ο Adam βελτιστοποιητής, χρησιμοποιούνται συχνά για την εκπαίδευση σε βαθιά νευρωνικά δίκτυα. Ενημερώνει αποτελεσματικά τις παραμέτρους του μοντέλου κατά τη διάρκεια της προπόνησης συνδυάζοντας έννοιες από την ορμή και το RMSprop. Ο Adam διατηρεί διακριτούς ρυθμούς εκμάθησης για κάθε παράμετρο, τους τροποποιεί ως απόκριση σε ιστορικές κλίσεις και χρησιμοποιεί ορμή για να επιταχύνει τη σύγκλιση.

3.2.1 Το μέρος του Adam στο overfitting

- *Προσαρμοστικά ποσοστά μάθησης(Adaptive Learning Rates):*
Ο Adam τροποποιεί τους ρυθμούς εκμάθησης για κάθε παράμετρο ξεχωριστά, γεγονός που το καθιστά πιο ικανό στην πλοήγηση στο έδαφος βελτιστοποίησης. Ρυθμίζοντας το μέγεθος των ενημερώσεων παραμέτρων, οι προσαρμοστικοί ρυθμοί εκμάθησης μπορούν να βοηθήσουν στην πρόληψη της υπερβολικής προσαρμογής.
- *Ορμή (Momentum):*
Ο Adam αποτρέπει τις ταλαντώσεις κατά τη βελτιστοποίηση χρησιμοποιώντας ορμή(Momentum) για την εξομάλυνση των ενημερώσεων. Αυτό μπορεί να μειώσει την πιθανότητα υπερβολικής προσαρμογής και να οδηγήσει σε πιο σταθερή σύγκλιση.
- *Εφέ τακτοποίησης:*
Οι προσαρμοστικοί ρυθμοί μάθησης και η ορμή του Adam συνεργάζονται για τη διαχείριση της δυναμικής ενημέρωσης κατά τη διάρκεια της εκπαίδευσης, αποτρέποντας έτσι την υπερβολική προσαρμογή με τρόπο παρόμοιο με την σιωπηρή τακτοποίηση(implicit fit).

Συμπερασματικά, τεχνικές όπως η τακτοποίηση, η αύξηση δεδομένων, διασταυρωμένη επικύρωση και η έγκαιρη παύση βοηθούν στη μείωση υπερπροσαρμογής. Αυτό γίνεται εφικτό από το Adam βελτιστοποιητής, το οποίο ενσωματώνει ορμή και προσφέρει ρυθμιζόμενους ρυθμούς εκμάθησης, με αποτέλεσμα πιο σταθερές και αποτελεσματικές διαδικασίες βελτιστοποίησης που μειώνουν την πιθανότητα υπερπροσαρμογής.

3.3 Η σημασία του πειράγματος των υπερπαραμέτρων (hyperparameter - tuning)

Για να εξισορροπήσουμε την πολυπλοκότητα του μοντέλου με τα εγγενή χαρακτηριστικά του συνόλου δεδομένων, είναι κρίσιμο να τροποποιηθεί ο αριθμός των στρωμάτων σε ένα νευρωνικό δίκτυο, συμπεριλαμβανομένων των στρώσεων LSTM, ανάλογα με το μέγεθος των δεδομένων. Το μέγεθος και η πολυπλοκότητα των δεδομένων εισόδου αναφέρονται ως "κλίμακα των δεδομένων" και αποτελούν σημαντικούς παράγοντες για τον καθορισμό της καλύτερης αρχιτεκτονικής νευρωνικών δικτύων. Το παρακάτω δικαιολογεί τη σημασία της αλλαγής του αριθμού των στρωμάτων.

Αντιστοίχιση χωρητικότητας (Capacity Matching):

Η ικανότητα ενός νευρικού δικτύου να αναγνωρίζει και να εντοπίζει τα πρότυπα στα δεδομένα επηρεάζεται από τον αριθμό των στρωμάτων στο δίκτυο. Για ένα περίπλοκο σύνολο δεδομένων, ένα μοντέλο με πολύ λίγα στρώματα μπορεί να δυσκολεύεται να αντιπροσωπεύει πολύπλοκες σχέσεις και μπορεί να υποβάλει υπόψη τα δεδομένα. Αντίθετα, εάν το μοντέλο περιέχει πάρα πολλά στρώματα για ένα απλοποιημένο σύνολο δεδομένων, μπορεί να υπερβεί τα δεδομένα με εκμάθηση εξωτερικών πληροφοριών ή θορύβου. Η πραγματοποίηση αλλαγών στον αριθμό των στρωμάτων εγγυάται ότι η ικανότητα του μοντέλου αντιστοιχεί στην πολυπλοκότητα των δεδομένων.

Αποφυγή υπερπροσαρμογής (Avoiding Overfitting):

Ένα μοντέλο λέγεται ότι έχει υπερπροσαρμογή αν γίνει υπερβολικά εξειδικευμένο στην εκμάθηση του σετ κατάρτισης και των χαμηλών επιδόσεων όταν πρόκειται να γενικευτεί σε νέα, μη αναμφισβήτητα δεδομένα. Η υπερπροσαρμογή μπορεί να αποφευχθεί μειώνοντας τον αριθμό των στρωμάτων, ιδιαίτερα όταν κάποιος δουλεύει σε ένα μικρότερο σύνολο δεδομένων. Ένα πιο σύνθετο μοντέλο (περισσότερα στρώματα) μπορεί να απομνημονεύσει το θόρυβο σε ένα μικρό σύνολο δεδομένων, ενώ ένα απλούστερο μοντέλο (λιγότερα στρώματα) μπορεί να έχει καλύτερη γενίκευση (generalisation).

Υπολογιστική Αποδοτικότητα (Computational Efficiency):

Η υπολογιστική πολυπλοκότητα του μοντέλου αυξάνεται κανονικά με τον αριθμό των στρωμάτων. Η τροποποίηση των στρωμάτων εγγυάται ότι το μοντέλο είναι υπολογιστικά αποδοτικό σε καταστάσεις όπου ο χρόνος κατάρτισης είναι ένα κρίσιμο στοιχείο ή οι υπολογιστικοί πόροι είναι λίγοι. Προκειμένου να καταγράψουν πιο πολύπλοκα πρότυπα, μπορεί να είναι απαραίτητη μια βαθύτερη αρχιτεκτονική για μεγαλύτερα σύνολα δεδομένων.

Ιεραρχία χαρακτηριστικών (Feature Hierarchy):

Τα μοντέλα που διαθέτουν βαθύτερες αρχιτεκτονικές και το σωστό ποσό των στρωμάτων επιτρέπουν στο μοντέλο να μάθει παραστάσεις χαρακτηριστικών που είναι ιεραρχικές. Τα βαθύτερα στρώματα σε δεδομένα χρονοσειρών μπορούν να βοηθήσουν στην ταυτοποίηση των προτύπων σε διάφορες χρονικές κλίμακες. Τα βαθύτερα στρώματα μπορούν να επικεντρωθούν σε μακροπρόθεσμες τάσεις, ενώ τα χαμηλότερα στρώματα μπορούν να συλλάβουν βραχυπρόθεσμες εξαρτήσεις. Η εκμάθηση ιεραρχικών χαρακτηριστικών που σχετίζονται με την κλίμακα δεδομένων διευκολύνεται με την προσαρμογή της αρχιτεκτονικής στρώματος.

Διαβάθμιση ροής και βαθμίδες εξαφάνισης/έκρηξης (Gradient Flow and Vanishing/Exploding Gradients):

Στα βαθιά δίκτυα, μπορεί να συμβεί το πρόβλημα της εξαφάνισης ή της έκρηξης, ιδιαίτερα στις εφαρμογές LSTM. Η εξαφάνιση των κλίσεων μπορεί να προκαλέσει προβλήματα για βαθιά δίκτυα με πολύ λίγα στρώματα, τα οποία θα καθιστούσαν την εκπαίδευση πιο

δύσκολη. Ωστόσο, ένας υπερβολικός αριθμός στρωμάτων θα μπορούσε να οδηγήσει σε κλίσεις που εκρήγνυνται. Αυτά τα θέματα που σχετίζονται με τη κλίση μπορούν να επιλυθούν με την τροποποίηση της αρχιτεκτονικής των στρωμάτων.

Σαν συμπέρασμα, ένα κρίσιμο πρώτο βήμα προς την ανάπτυξη ενός καλά ισορροπημένου μοντέλου που μπορεί να καταγράψει με επιτυχία τα υποκείμενα πρότυπα και τις εξαρτήσεις στο σύνολο δεδομένων, ενώ αποτρέποντας τα προβλήματα όπως η υπερφόρτωση και η υπολογιστική αναποτελεσματικότητα είναι ο καθορισμός των στρωμάτων που θα έπρεπε να έχει ένα μοντέλο LSTM με βάση το μέγεθος των δεδομένων. Ευθυγραμμίζει την ικανότητα του μοντέλου με την πολυπλοκότητα των δεδομένων, συμβάλλοντας στην καλύτερη γενίκευση και προβλεπτική απόδοση.

3.3.1 Προ επεξεργασία Δεδομένων

Τα μοντέλα LSTM, ARIMA, XGBoost και MLP, υποβάλλονται σε ξεχωριστά βήματα προ επεξεργασίας προσαρμοσμένα στις συγκεκριμένες απαιτήσεις τους, ωστόσο μοιράζονται ομοιότητες σε ορισμένες πτυχές. Αναλυτικότερα :

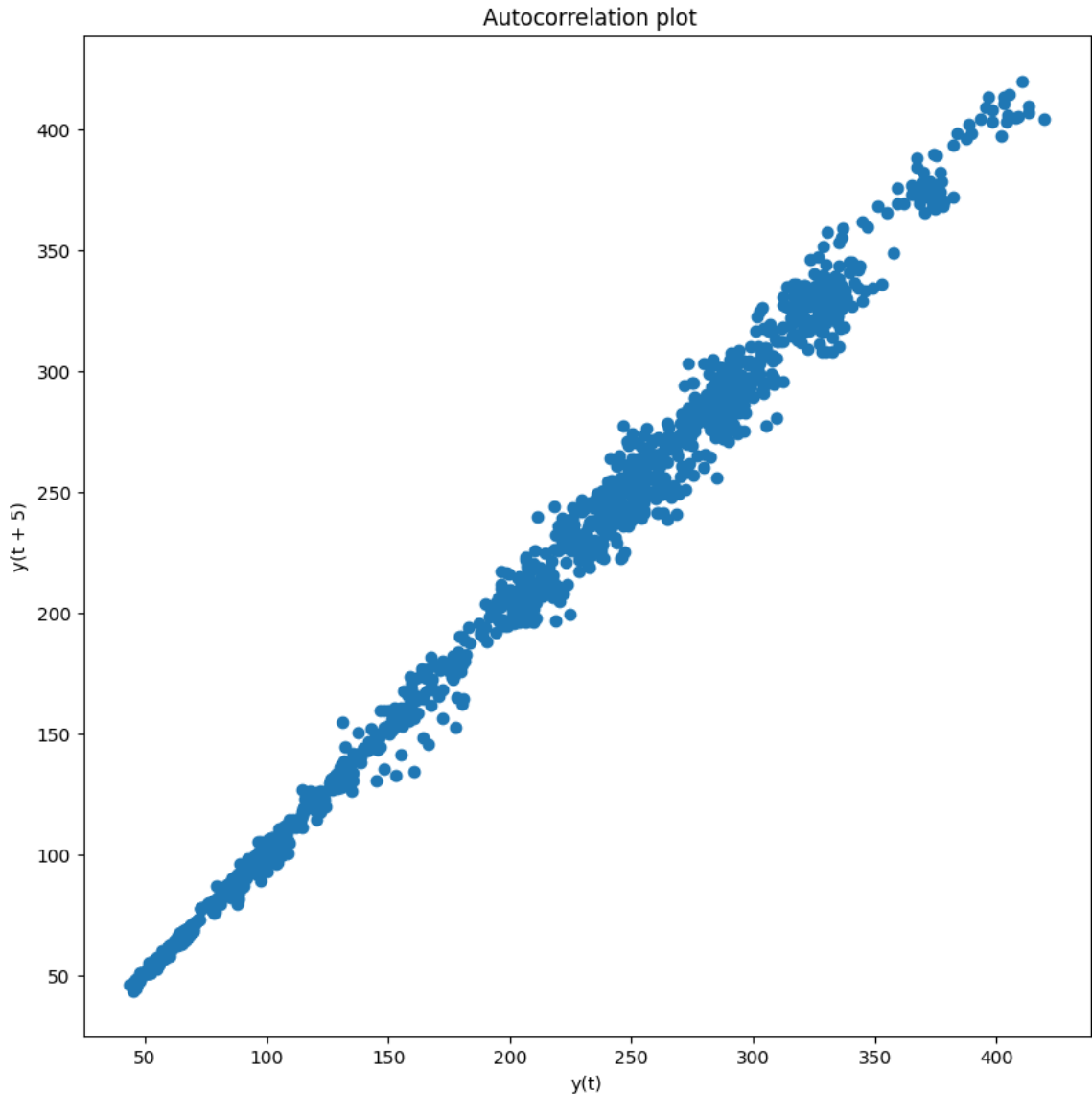
Για την LSTM, η προ επεξεργασία περιλαμβάνει:

- Κανονικοποίηση των δεδομένων με τη χρήση Min-Max Scaling σε ένα καθορισμένο εύρος (0 έως 15).
- Δημιουργία χρονικών ακολουθιών με μήκος ακολουθίας 11, αποτυπώνοντας βραχυπρόθεσμα χρονικά μοτίβα.
- Διαχωρισμό των δεδομένων σε σύνολα εκπαίδευσης και επικύρωσης, με το 80% να διατίθεται για εκπαίδευση και το υπόλοιπο για επικύρωση.
- Μετατροπή των δεδομένων σε τένσορες PyTorch και δημιουργία φορτωτών δεδομένων για αποδοτική εκπαίδευση κατά δέσμη.

Η προ επεξεργασία ARIMA περιλαμβάνει:

- αναζήτηση πλέγματος για την εύρεση των βέλτιστων παραμέτρων (p, d, q) χρησιμοποιώντας το κριτήριο πληροφοριών Akaike (AIC).
- Κλιμάκωση Min-Max των τιμών "Close".

- Διαχωρισμό των δεδομένων σε σύνολα εκπαίδευσης και δοκιμής για αξιολόγηση.

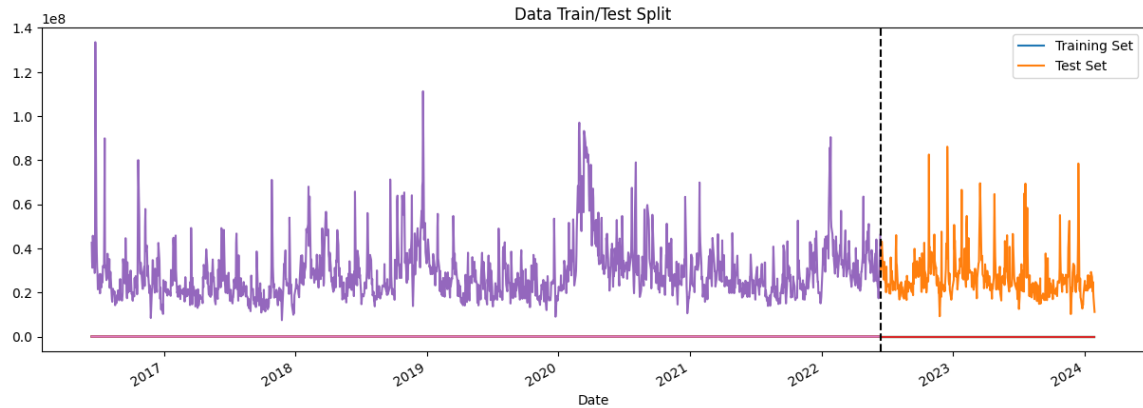


Εικόνα 9. Θετική αυτοσυσχέτιση: Η διαγώνια γραμμή υποδηλώνει θετική σχέση μεταξύ των παρατηρήσεων και των υστερήσεων τους. Με άλλα λόγια, όταν αυξάνεται η υστέρηση (κινούμενη προς τα δεξιά κατά μήκος του άξονα x), αυξάνεται και η συσχέτιση μεταξύ των παρατηρήσεων και των υστερήσεων τους.

Για το XGBoost:

- Διαχωρισμός των δεδομένων σε σύνολα εκπαίδευσης και δοκιμής με βάση μια καθορισμένη ημερομηνία ("2022-06-14").
- Κανονικοποίηση των αριθμητικών χαρακτηριστικών ('Open', 'High', 'Low', 'Close') σε ένα εύρος από 0 έως 15 χρησιμοποιώντας Min-Max Scaling.
- Δημιουργία χαρακτηριστικών υστέρησης ('Close(t-i)') για την καταγραφή των χρονικών εξαρτήσεων.

- Ρύθμιση υπερπαραμέτρων μέσω Grid Search Cross-Validation για τον εντοπισμό των βέλτιστων υπερπαραμέτρων του μοντέλου.



Εικόνα 10 . Απεικόνιση διαχωρισμού δεδομένων για το κομμάτι εκπαίδευσης και test.

MLP περιλαμβάνει:

- Κανονικοποίηση των δεδομένων με χρήση της κλίμακας Min-Max σε ένα εύρος από 0 έως 15.
- Δημιουργία χρονικών ακολουθιών με μήκος ακολουθίας 11.
- Διαχωρισμό των δεδομένων σε σύνολα εκπαίδευσης, επικύρωσης και δοκιμής, με 80%, 10% και 10% των δεδομένων να κατανέμονται σε καθένα από αυτά, αντίστοιχα.
- Μετατροπή των δεδομένων σε τένσορες PyTorch και δημιουργία φορτωτών δεδομένων για εκπαίδευση κατά δέσμη.

Διαφορές:

- Το ARIMA επικεντρώνεται στον προσδιορισμό των βέλτιστων παραμέτρων μέσω αναζήτησης πλέγματος με βάση την AIC, ενώ άλλα μοντέλα δεν περιλαμβάνουν εγγενώς τέτοια βελτιστοποίηση παραμέτρων.
- Το XGBoost ενσωματώνει χαρακτηριστικά υστέρησης για την καταγραφή των χρονικών εξαρτήσεων, κάτι που δεν αποτελεί βήμα στην προεπεξεργασία άλλων μοντέλων.
- Τα LSTM και MLP χρησιμοποιούν τένσορες PyTorch και φορτωτές δεδομένων για την εκπαίδευση, ειδικά για το πλαίσιο βαθιάς μάθησης, ενώ τα ARIMA και XGBoost χρησιμοποιούν το scikit-learn για την εκπαίδευση του μοντέλου.
- Η προεπεξεργασία του ARIMA δεν περιλαμβάνει ρητό διαχωρισμό των δεδομένων σε σύνολα εκπαίδευσης και επικύρωσης, σε αντίθεση με τα άλλα μοντέλα.

Ομοιότητες:

- Όλα τα μοντέλα εκτελούν κανονικοποίηση δεδομένων για να διασφαλίσουν ότι τα χαρακτηριστικά εισόδου βρίσκονται σε συνεπή κλίμακα.
- Ο διαχωρισμός των δεδομένων είναι ένα κοινό βήμα μεταξύ όλων των μοντέλων για τον διαχωρισμό των δεδομένων σε υποσύνολα εκπαίδευσης και αξιολόγησης.
- Κάθε μοντέλο χειρίζεται τις χρονικές πτυχές των δεδομένων, είτε μέσω χαρακτηριστικών υστέρησης (XGBoost) είτε μέσω της δημιουργίας ακολουθιών (LSTM και MLP).

Συνοπτικά, ενώ η προ επεξεργασία κάθε μοντέλου διαφέρει σε ορισμένες πτυχές, μοιράζονται κοινά βήματα, όπως η κανονικοποίηση και ο διαχωρισμός των δεδομένων. Ωστόσο, χρησιμοποιούν επίσης μοναδικές τεχνικές προσαρμοσμένες στις συγκεκριμένες προσεγγίσεις μοντελοποίησης, αντανακλώντας την ποικιλομορφία των μεθοδολογιών πρόβλεψης χρονοσειρών.

Κεφάλαιο 4. Τεχνικές που χρησιμοποιήθηκαν

4.1 Μοντέλο LSTM

Το μοντέλο που έχει τεθεί σε εφαρμογή είναι ένα νευρωνικό δίκτυο βραχυπρόθεσμης μνήμης (LSTM) που δημιουργήθηκε ειδικά για πρόβλεψη χρονοσειρών, με έμφαση στην πρόβλεψη οικονομικών δεδομένων όπως οι τιμές των μετοχών. Το μοντέλο αποτελείται από βασικά στοιχεία για να συλλάβει τις χρονικές εξαρτήσεις που υπάρχουν στα δεδομένα χρονοσειρών και κατασκευάζεται χρησιμοποιώντας το πλαίσιο Deep Learning Pytorch.

Η αρχιτεκτονική LSTM είναι κατάλληλη για την εκμάθηση μακροπρόθεσμων εξαρτήσεων σε διαδοχικά δεδομένα, επειδή μπορεί να λύσει το πρόβλημα της κλίσης. Το πλήρως συνδεδεμένο στρώμα (FC) και το στρώμα LSTM αποτελούν τα δύο κύρια τμήματα του μοντέλου. Το στρώμα LSTM, το οποίο δημιουργείται χρησιμοποιώντας την μονάδα nn.LSTM, αποτελεί τον πυρήνα του μοντέλου. Οι παράμετροι για αυτό το στρώμα έχουν οριστεί σε επεξεργασία κατά δέσμες με `batch_first=True`, διάσταση hidden state dimension (ορισμένη σε 64) και input dimension (`num_feature`). Με την επεξεργασία ακολουθιών εισόδου και τη διατήρηση των εσωτερικών καταστάσεων, το στρώμα LSTM επιτρέπει στα δεδομένα να καταγράφουν εξαρτήσεις μεγάλης εμβέλειας.

Η έξοδος LSTM υποβάλλεται επιπλέον σε ένα στρώμα ομαλοποίησης παρτίδας πριν από την τροφοδοσία στο στρώμα FC. Με την ομαλοποίηση της εισόδου σε κάθε μίνι-παρτίδα και τη μείωση των εσωτερικών μεταφορών μεταβλητών, η ομαλοποίηση παρτίδας συμβάλλει στη σταθεροποίηση και την επιτάχυνση της διαδικασίας κατάρτισης.

Το FC πλήρως συνδεδεμένο (fully connected) στρώμα (nn.Linear) χρησιμοποιείται για να αντιστοιχίσει την έξοδο του LSTM στον επιθυμητό χώρο πρόβλεψης. Αυτό το στρώμα λειτουργεί ως τελικό βήμα μετασχηματισμού, παράγοντας την έξοδο του μοντέλου με βάση τις επεξεργασμένες πληροφορίες της ακολουθίας.

Η μέθοδος `forward pass` ορίζει τον υπολογισμό του μοντέλου `forward pass`. Λαμβάνει δεδομένα εισόδου (x) και τα περνάει μέσα από το στρώμα LSTM. Εξάγεται η έξοδος του στρώματος LSTM, μαζί με τις κρυφές καταστάσεις και τις καταστάσεις των κελιών. Στη συνέχεια, η κρυφή κατάσταση περνάει μέσω του γραμμικού στρώματος για να παραχθεί η τελική πρόβλεψη (x).

Κατά τη διάρκεια της φάσης εκπαίδευσης, το μοντέλο υφίσταται πολλαπλές εποχές, με κάθε εποχή να επαναλαμβάνεται πάνω από το σύνολο δεδομένων κατάρτισης. Η λειτουργία `train` διευκολύνει τη διαδικασία εκπαίδευσης του μοντέλου. Επαναλαμβάνει τις παρτίδες δεδομένων που παρέχονται από το `dataloader`, υπολογίζει τις προβλέψεις χρησιμοποιώντας το μοντέλο, υπολογίζει την απώλεια μεταξύ των προβλέψεων και των

πραγματικών ετικετών (y) και εκτελεί οπίσθιο πολλαπλασιασμό (backpropagation) για την ενημέρωση των παραμέτρων του μοντέλου με βάση τις υπολογισμένες κλίσεις. Η απώλεια εκπαίδευσης που συσσωρεύεται σε ολόκληρη την εποχή επιστρέφεται ως δείκτης της απόδοσης του μοντέλου κατά τη διάρκεια της εκπαίδευσης. Το μοντέλο αρχικοποιείται με τον καθορισμένο αριθμό χαρακτηριστικών εισόδου (`num_feature`). Κατά τη διάρκεια της εκπαίδευσης, χρησιμοποιείται ο ADAM optimizer για τη βελτιστοποίηση των παραμέτρων του μοντέλου.

Για να εκτιμηθεί η απόδοση του μοντέλου, υφίσταται επικύρωση σε ξεχωριστό σύνολο δεδομένων δοκιμών. Οι προβλέψεις συγκρίνονται με τις πραγματικές τιμές χρησιμοποιώντας διάφορες μετρήσεις αξιολόγησης, συμπεριλαμβανομένου του μέσου απόλυτου σφάλματος (MAE), της εξήγησης της διακύμανσης, του μέσου τετραγωνικού σφάλματος ρίζας (RMSE), του μέσου τετραγωνικού σφάλματος (MSE) και του R-τετράγωνο (R²).

Συνοψίζοντας, το μοντέλο νευρωνικού δικτύου που κατασκευάστηκε με την αρχιτεκτονική LSTM είναι προσαρμοσμένο για την επεξεργασία διαδοχικών δεδομένων, καθιστώντας το κατάλληλο για εργασίες όπως η πρόβλεψη χρονοσειρών, η επεξεργασία φυσικής γλώσσας και η αναγνώριση ομιλίας. Αξιοποιώντας τις δυνατότητες των στρωμάτων LSTM για τη σύλληψη των χρονικών εξαρτήσεων, σε συνδυασμό με πλήρως συνδεδεμένα στρώματα για την πρόβλεψη, το μοντέλο επιδεικνύει την ικανότητα εκμάθησης σύνθετων μοτίβων και ακριβών προβλέψεων από διαδοχικές εισόδους δεδομένων.

4.2 Τμήματα Μοντέλου

Σημαντικοί υπερπαραμέτροι στην αρχιτεκτονική μοντέλου LSTM επιλέγονται προσεκτικά, γεγονός που ενισχύει την ικανότητα του μοντέλου να συλλάβει χρονικές τάσεις στα δεδομένα των οικονομικών χρονοσειρών.

1. Κρυφές Μονάδες (Hidden Units):

Ο αριθμός των κρυφών μονάδων σε κάθε στρώμα LSTM καθορίζεται από την παράμετρο `HIDDEN_SIZE`. Το μοντέλο είναι σε θέση να μάθει πιο περίπλοκες αναπαραστάσεις των δεδομένων εισόδου όταν υπάρχουν περισσότερες κρυφές μονάδες σε αυτό. Αλλά επιμηκύνει επίσης την περίοδο εκπαίδευσης και αυξάνει την πολυπλοκότητα του μοντέλου. Λειτουργεί με ένα μέγεθος εισόδου που καθορίζεται από την παράμετρο `num_feature` και ένα κρυφό μέγεθος που ορίζεται σε 64, καθορίζοντας τη διαστατικότητα της κρυφής κατάστασης. Θέτοντας `batch_first=True`, το στρώμα προσαρμόζεται ώστε να δέχεται ταυστές εισόδου με τη διάσταση της παρτίδας ως κύρια διάσταση, επιτρέποντας την αποτελεσματική επεξεργασία διαδοχικών δεδομένων σε παρτίδες.

2. Πλήρως συνδεδεμένο (γραμμικό) στρώμα (Fully Connected (Linear) Layer):

Το μέγεθος εισόδου του γραμμικού στρώματος ορίζεται σε 64, που αντιστοιχεί στο μέγεθος εξόδου του στρώματος LSTM, ενώ το μέγεθος εξόδου καθορίζεται από την παράμετρο `num_feature` που παρέχεται κατά την αρχικοποίηση του μοντέλου, ευθυγραμμιζόμενο με τη διαστατικότητα της εξόδου-στόχου.

3. Αριθμός Εποχών (Number of Epochs):

Το μοντέλο υποβάλλεται σε εκπαίδευση για συνολικά 50 εποχές, όπως ορίζεται από την παράμετρο `n_epochs`. Κάθε εποχή περιλαμβάνει ένα πλήρες πέρασμα στο σύνολο δεδομένων εκπαίδευσης, διευκολύνοντας την επαναληπτική βελτίωση των παραμέτρων του μοντέλου μέσω της οπισθοδιάδοσης. Η επιλογή του αριθμού των εποχών αντικατοπτρίζει μια ισορροπία μεταξύ του χρόνου εκπαίδευσης και της σύγκλισης σε μια βέλτιστη λύση. Ενώ η αύξηση του αριθμού των εποχών μπορεί να προσφέρει περισσότερες ευκαιρίες για προσαρμογή των παραμέτρων, απαιτεί προσεκτική εξέταση των υπολογιστικών πόρων και του κινδύνου υπερπροσαρμογής.

4. Παρτίδες εκπαίδευσης (Training Batches):

Η διαδικασία εκπαίδευσης λειτουργεί με παρτίδες δεδομένων που προέρχονται από τον τροφοδότη δεδομένων (`train_dataloader`). Κάθε παρτίδα αποτελείται από ακολουθίες εισόδου (x) σε συνδυασμό με τις αντίστοιχες ετικέτες-στόχους (y). Με την επεξεργασία δεδομένων σε παρτίδες, το μοντέλο μπορεί να αξιοποιήσει τον παράλληλο υπολογισμό και να βελτιστοποιήσει τη χρήση μνήμης, διευκολύνοντας την αποτελεσματική εκπαίδευση σε μεγάλα σύνολα δεδομένων. Η επαναληπτική φύση της επεξεργασίας παρτίδων επιτρέπει στο μοντέλο να ενημερώνει σταδιακά τις παραμέτρους του, βελτιώνοντας σταδιακά την απόδοση σε διαδοχικές εποχές.

Συνοψίζοντας, η ρύθμιση των στρωμάτων LSTM, των κρυφών μονάδων, των εποχών και του μεγέθους παρτίδας έχει μελετηθεί προσεκτικά για να επιτύχει μια ισορροπία μεταξύ της τακτοποίησης, της αποτελεσματικότητας της κατάρτισης και της πολυπλοκότητας του μοντέλου. Μαζί, αυτά τα `hyperparameters` βελτιώνουν την ικανότητα του μοντέλου να εντοπίζουν και να προβάλλουν τα πρότυπα από τα δεδομένα των οικονομικών χρονοσειρών, τα οποία βελτιώνουν την πρόβλεψη.

4.3 XGBoost

Το XGBoost, που σημαίνει eXtreme Gradient Boosting, είναι μια δύναμη μηχανικής μάθησης που αποτελεί ερευνητική προσπάθεια του υποψήφιου διδάκτορα του Πανεπιστημίου της Ουάσινγκτον Tianqi Chen με στόχο τη βελτίωση της αποτελεσματικότητας και της αποδοτικότητας των αλγορίθμων Gradient Boosting. Ξεκίνησε την ευρεία χρήση του ως πακέτο λογισμικού ανοιχτού κώδικα (open source) το 2014 αποκτώντας υψηλή δημοτικότητα στη κοινότητα της επιστήμης των Μεγάλων Δεδομένων. Αποτέλεσε την go-to επιλογή ως προς την λύση πολλών προβλημάτων αλλά ακόμα και για την αποδοτικότητα του σε διαγωνισμούς machine learning συμπεριλαμβανομένων και αυτών που διεξάγονται στο kaggle. Οι εξαιρετικές επιδόσεις, η επεκτασιμότητα και η προσαρμοστικότητά του τον έφεραν στο προσκήνιο των αλγορίθμων μηχανικής μάθησης, όπου ξεπερνούσε συστηματικά τις συμβατικές τεχνικές για μια ποικιλία εργασιών, συμπεριλαμβανομένων των συστημάτων κατάταξης και συστάσεων, της παλινδρόμησης και της ταξινόμησης.

Οι αποδοτικές δυνατότητες παράλληλου και κατανεμημένου υπολογισμού του επιτρέπουν την ταχεία εκπαίδευση σε μεγάλα σύνολα δεδομένων, καθιστώντας το κατάλληλο για πραγματικές εφαρμογές σε κλάδους όπως οικονομικά, υγειονομική περίθαλψη και το ηλεκτρονικό εμπόριο. Οντάς ακόμα ένα καινούργιο μοντέλο μηχανικής μάθησης το XGBoost βελτιώνεται διαρκώς από μια ζωντανή κοινότητα προγραμματιστών που συμβάλλει στην τελειοποίηση του έτσι ώστε να συνεχίσει να αποτελεί αρωγό στους επαγγελματίες που ζητάνε γρήγορες και αποτελεσματικές λύσεις.

Τι είναι το XGBoost στον πυρήνα του

Το XGBoost είναι μέρος της οικογένειας boosting ενός συνόλου αλγορίθμων που συνδυάζουν αδύναμα μοντέλα (συνήθως decision trees) για να δημιουργήσουν ένα ισχυρό μοντέλο πρόβλεψης. Η βασική ιδέα που διέπει είναι η ικανότητά του να εκπαιδεύει πολλούς αδύναμους “learners” με διαδοχικό τρόπο, με έμφαση στη διόρθωση των λαθών που παράγονται από τα προηγούμενα μοντέλα. Αυτό επιτυγχάνεται μέσω μιας διαδικασίας που ονομάζεται gradient boosting, η οποία περιλαμβάνει την επαναληπτική προσαρμογή νέων μοντέλων στα κατάλοιπα (τις διαφορές μεταξύ των παρατηρούμενων και των προβλεπόμενων τιμών) των προηγούμενων μοντέλων. Αυτή η προσέγγιση έχει ως αποτέλεσμα την σταδιακή βελτίωση της συνολικής ακρίβειας εστιάζοντας στα πιο δύσκολα μοντέλα.

Μαθηματικά, το XGBoost ελαχιστοποιεί μια συνάρτηση απώλειας (loss function) που μετρά τη διαφορά μεταξύ των προβλεπόμενων και των πραγματικών τιμών της μεταβλητής-στόχου. Συνδυάζει τις προβλέψεις πολλαπλών ασθενών learners (δέντρα απόφασης) αθροίζοντας τις εξόδους τους, σταθμισμένες με συντελεστές που καθορίζονται κατά τη διαδικασία εκπαίδευσης. Για να αποτρέψει την υπερβολική προσαρμογή(overfitting), το XGBoost ενσωματώνει όρους κανονικοποίησης στην αντικειμενική του συνάρτηση, όπως οι ποινές L1 (Lasso) και L2 (Ridge), οι οποίες τιμωρούν την πολυπλοκότητα του μοντέλου.

Κατά τη χρήση του XGBoost, υπάρχουν διάφορες περιπτώσεις που πρέπει να έχουν κατά νου οι χρήστες για να μεγιστοποιήσουν την αποτελεσματικότητά του και να αποφύγουν τις συνήθεις παγίδες:

- **Υπερπροσαρμογή (Overfitting):** Το XGBoost είναι επιρρεπές στο overfitting, ειδικά όταν η πολυπλοκότητα του μοντέλου δεν ελέγχεται σωστά. Οι χρήστες θα πρέπει να ρυθμίζουν υπερπαραμέτρους όπως το tree depth, learning rate, regularization και τις παραμέτρους κανονικοποίησης για να αποτρέψουν την υπερπροσαρμογή στα δεδομένα εκπαίδευσης. Οι τεχνικές Cross-validation μπορούν να βοηθήσουν στην αξιολόγηση της απόδοσης του μοντέλου και στον εντοπισμό του overfitting.
- **Συντονισμός υπερπαραμέτρων (Hyperparameter Tuning):** Προσφέρει ένα ευρύ φάσμα υπερπαραμέτρων που μπορούν να επηρεάσουν σημαντικά την απόδοση του μοντέλου. Οι χρήστες θα πρέπει να πειραματιστούν με διαφορετικές τιμές παραμέτρων και να χρησιμοποιήσουν τεχνικές όπως grid search για την εύρεση του βέλτιστου συνδυασμού και της ισορροπίας μεταξύ πολυπλοκότητάς του μοντέλου και του generalization.
- **Προεπεξεργασία δεδομένων (Data Preprocessing):** Η κατάλληλη προεπεξεργασία των δεδομένων είναι ζωτικής σημασίας για τη βέλτιστη απόδοση του XGBoost. Οι χρήστες πρέπει να χειρίζονται κατάλληλα τις ελλείπουσες τιμές, τις ακραίες τιμές και τις κατηγορικές μεταβλητές.

- **Υπολογιστικοί Πόροι (Computational Resources):** Το XGBoost σε συγκεκριμένες περιπτώσεις ενδεχομένως να είναι πολύ απαιτητικό από την υπολογιστική δύναμή με αποτέλεσμα να καθιστά αναγκαίο για τους χρήστες να βεβαιωθούν ότι διαθέτουν αρκετή μνήμη και επεξεργαστική ικανότητα για την επαρκή εκπαίδευση και αξιολόγηση των μοντέλων XGBoost.

4.4 MLP: Εισαγωγή στα Multilayered Perceptrons (MLP) στην Πρόβλεψη Χρονοσειρών

Το μοντέλο perceptron του Frank Rosenblatt, το οποίο πρότεινε στα τέλη της δεκαετίας του 1950, είναι το θεμέλιο της πρωτοποριακής εργασίας στα τεχνητά νευρωνικά δίκτυα, όπου ξεκινά η ιστορία των MLP. Αρχικά προοριζόταν να είναι ένας απλός δυαδικός ταξινομητής, το perceptron μπορεί να αναγνωρίσει γραμμικά όρια απόφασης. Ωστόσο, περισσότερες προόδους προκλήθηκαν από τις ελλείψεις του στη διαχείριση περίπλοκων, μη γραμμικών αλληλεπιδράσεων.

Ένα σημαντικό σημείο καμπής για τα MLP ήρθε με την ανάπτυξη του αλγορίθμου οπίσθιο πολλαπλασιασμό (backpropagation) στη δεκαετία του 1980. Το δίκτυο ήταν σε θέση να αναγνωρίσει πολύπλοκα μοτίβα στα δεδομένα χάρη σε αυτή τη μέθοδο εκμάθησης και την ενσωμάτωση πολλών επιπέδων. Η πρόβλεψη χρονοσειρών είναι μια από τις πιο ανθεκτικές και προσαρμόσιμες εφαρμογές που κατέστη δυνατή χάρη σε αυτήν την καινοτομία.

Ένα επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα και ένα επίπεδο εξόδου συνθέτουν ένα MLP. Κόμβοι, ή νευρώνες, βρίσκονται σε κάθε στρώμα και συνδέονται με βάρη. Αφού ληφθούν τα δεδομένα από το επίπεδο εισόδου, μετασχηματίζονται μέσω σταθμισμένων συνδέσεων, με την προσθήκη μη γραμμικότητας μέσω της χρήσης συναρτήσεων ενεργοποίησης. Λόγω του πολυεπίπεδου σχεδιασμού τους, τα MLP είναι ιδιαίτερα κατάλληλα για εργασίες πρόβλεψης χρονοσειρών, καθώς μπορούν να αναγνωρίσουν πολύπλοκους συσχετισμούς και μοτίβα μέσα σε χρονικά δεδομένα.

4.4.1. Σημασία στην πρόβλεψη χρονοσειρών:

Η πρόβλεψη χρονοσειρών περιλαμβάνει την πρόβλεψη μελλοντικών αξιών με βάση ιστορικά δεδομένα, ένα έργο κρίσιμο σε πολλούς τομείς όπως τα χρηματοοικονομικά, η πρόβλεψη καιρού και η ανάλυση χρηματιστηρίων. Τα MLP υπερέχουν σε αυτόν τον τομέα λόγω της ικανότητάς τους να μοντελοποιούν περίπλοκες εξαρτήσεις και να προσαρμόζονται σε μεταβαλλόμενα μοτίβα με την πάροδο του χρόνου. Αξιοποιώντας την ιεραρχική δομή των MLP, αυτά τα μοντέλα μπορούν να αποκαλύψουν χρονικές εξαρτήσεις, καθιστώντας τα αποτελεσματικά εργαλεία για ακριβείς και προσαρμόσιμες προβλέψεις χρονοσειρών. Πιο συγκεκριμένα τα MLP :

- Με την ικανότητά τους να μοντελοποιούν μη γραμμικές σχέσεις, διαπρέπουν σε διακριτικά μοτίβα στα δεδομένα οικονομικών χρονοσειρών που μπορεί να διαφεύγουν από απλούστερα μοντέλα.
- Εξοπλισμένα με πολλαπλά κρυφά στρώματα, παρουσιάζουν υψηλό βαθμό προσαρμοστικότητας. Μπορούν να μάθουν και να προσαρμόσουν τις εσωτερικές τους αναπαραστάσεις καθώς γίνονται διαθέσιμες νέες πληροφορίες, καθιστώντας τις κατάλληλες για προβλέψεις σε δυναμικά οικονομικά περιβάλλοντα.
- Χάρη στη διαδοχική επεξεργασία τους σε κρυφά επίπεδα, διαπρέπουν στο να συλλαμβάνουν και να αξιοποιούν χρονικές εξαρτήσεις μέσα στα δεδομένα χρονοσειρών
- συμβάλλουν στην αξιόπιστη αξιολόγηση κινδύνου παρέχοντας διαφοροποιημένες προβλέψεις που εξηγούν τις περίπλοκες σχέσεις μεταξύ των οικονομικών δεικτών. Αυτό, με τη σειρά του, εξουσιοδοτεί τους λήπτες αποφάσεων με αξιόπιστες ιδέες, βοηθώντας στη διαμόρφωση ενημερωμένων στρατηγικών και πολιτικών.
- “Λάμπουν”, αξιοποιώντας την ικανότητά τους να μαθαίνουν και να προσαρμόζονται στα περίπλοκα μοτίβα που ενυπάρχουν στα δεδομένα χρηματοοικονομικών χρονοσειρών.

4.4.2 Αρχιτεκτονική Μοντέλου: Αποκάλυψη του Ρόλου των Κρυφών Επιπέδων και Ενεργοποίηση ReLU

Στον τομέα των Perceptrons Multi-Layer (MLPs), οι αρχιτεκτονικές επιλογές επηρεάζουν σημαντικά την ικανότητα του μοντέλου να συλλαμβάνει πολύπλοκα μοτίβα μέσα σε δεδομένα χρονοσειρών, καθιστώντας το ιδιαίτερα εφαρμόσιμο στις οικονομικές προβλέψεις. Η επιλεγμένη αρχιτεκτονική για αυτήν την εφαρμογή περιλαμβάνει τέσσερα στρώματα εισόδου, δυο κρυφά στρώματα με 8 νευρώνες το καθένα και ένα στρώμα εξόδου. Ενσωματωμένα στην αρχιτεκτονική MLP είναι τα στρώματα ομαδικής κανονικοποίησης (nn.BatchNorm1d), τα οποία συμβολίζονται ως self.bn1 και self.bn2 και παρεμβάλλονται μετά το πρώτο και το δεύτερο γραμμικό στρώμα, αντίστοιχα. Η εισαγωγή ενός κρυφού επιπέδου, σε συνδυασμό με τη λειτουργία ενεργοποίησης Rectified Linear Unit (ReLU), είναι μια στρατηγική απόφαση που ενισχύει την ικανότητα του μοντέλου να χειρίζεται μη γραμμικές σχέσεις που είναι εγγενείς στα οικονομικά δεδομένα.

Κανονικοποίηση παρτίδας (Batch Normalization):

Αυτά τα στρώματα ομαλοποιούν τις ενεργοποιήσεις κάθε στρώματος, σταθεροποιώντας τη διαδικασία εκπαίδευσης και επιταχύνοντας τη σύγκλιση. Με τον μετριασμό της εσωτερικής μετατόπισης των συνδιακυμάνσεων και τη μείωση της εξάρτησης από τα σχήματα αρχικοποίησης, η ομαλοποίηση των παρτίδων ενισχύει τη δυναμική και την ευρωστία της εκπαίδευσης του μοντέλου.

Κρυφό στρώμα (Hidden Layer) και νευρώνες:

Το Hidden Layer λειτουργεί ως κρίσιμος ενδιάμεσος μεταξύ των επιπέδων εισόδου και εξόδου, επιτρέποντας στο μοντέλο να μάθει περίπλοκες αναπαραστάσεις των υποκείμενων μοτίβων στα δεδομένα. Σε αυτή τη συγκεκριμένη αρχιτεκτονική MLP, χρησιμοποιούνται δύο κρυφά στρώματα, που αποτελούνται από 8 νευρώνες. Η επιλογή του αριθμού των νευρώνων βασίζεται συχνά σε πειραματισμούς και μπορεί να ποικίλλει ανάλογα με την πολυπλοκότητα της εργασίας πρόβλεψης. Ένας υψηλότερος αριθμός νευρώνων μπορεί δυνητικά να επιτρέψει στο μοντέλο να καταγράψει πιο περίπλοκες σχέσεις, αλλά ενέχει τον κίνδυνο υπερβολικής προσαρμογής στα δεδομένα εκπαίδευσης.

Λειτουργία ενεργοποίησης ReLU:

Η συνάρτηση ενεργοποίησης Rectified Linear Unit (ReLU) είναι ο ακρογωνιαίος λίθος στις σύγχρονες αρχιτεκτονικές νευρωνικών δικτύων και η συμπερίληψή της στο κρυφό στρώμα είναι ζωτικής σημασίας για την καταγραφή μη γραμμικοτήτων. Η συνάρτηση

ReLU λειτουργεί βάσει στοιχείων, αντικαθιστώντας τυχόν αρνητικές τιμές εισόδου με μηδέν και αφήνοντας τις θετικές τιμές αμετάβλητες. Μαθηματικά, μπορεί να οριστεί ως

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

Πλεονεκτήματα του ReLU:

- **Μη γραμμικότητα(Non-Linearity)** Το ReLU εισάγει τη μη γραμμικότητα στο μοντέλο, επιτρέποντάς του να μάθει και να αναπαραστήσει περίπλοκες σχέσεις που μπορεί να υπάρχουν στα δεδομένα των οικονομικών χρονοσειρών.
- **Αποτελεσματική Εκπαίδευση(Efficient Training)** : Η απλότητα του ReLU απλοποιεί τη διαδικασία εκπαίδευσης, επιτρέποντας ταχύτερη σύγκλιση κατά τη βελτιστοποίηση.
- **Αραιή ενεργοποίηση(Sparse Activation):** Μηδενίζοντας τις αρνητικές τιμές, το ReLU τείνει να ενεργοποιεί συγκεκριμένους νευρώνες επιλεκτικά, προωθώντας αραιές αναπαραστάσεις στο κρυφό στρώμα.

Αποτύπωση μη γραμμικών σχέσεων:

Τα οικονομικά δεδομένα συχνά παρουσιάζουν περίπλοκα, μη γραμμικά μοτίβα που τα γραμμικά μοντέλα δυσκολεύονται να καταγράψουν. Η λειτουργία ενεργοποίησης ReLU, σε συνδυασμό με το κρυφό επίπεδο, εξουσιοδοτεί το MLP να διακρίνει και να μοντελοποιεί αυτές τις πολύπλοκες σχέσεις. Αυτό είναι ιδιαίτερα σημαντικό στις οικονομικές προβλέψεις, όπου οι παράγοντες που επηρεάζουν τις τάσεις είναι πολύπλευροι και συχνά αλληλένδετοι.

Εξισορρόπηση πολυπλοκότητας και γενίκευσης:

Ενώ ένα μεμονωμένο κρυφό στρώμα με ενεργοποίηση ReLU επιτυγχάνει μια ισορροπία μεταξύ της πολυπλοκότητας και της γενίκευσης του μοντέλου, οι αρχιτέκτονες μοντέλων θα πρέπει να πειραματιστούν με διαφορετικές αρχιτεκτονικές για να βρουν τη βέλτιστη διαμόρφωση για τη συγκεκριμένη εργασία πρόβλεψης. Πρόσθετα κρυφά επίπεδα και εναλλακτικές λειτουργίες ενεργοποίησης μπορούν να διερευνηθούν για περαιτέρω βελτίωση της απόδοσης του μοντέλου.

4.4.3 Υπερπαράμετροι σε MLP μοντέλο για την πρόβλεψη χρονοσειρών:

Ποσοστό μάθησης (Learning Rate):

Ο ρυθμός εκμάθησης είναι μια κρίσιμη υπερπαράμετρος που διέπει το μέγεθος των βημάτων που λαμβάνονται κατά τη διαδικασία βελτιστοποίησης. Στο μοντέλο MLP, επιλέχθηκε ένα ποσοστό μάθησης 0,001. Αυτός ο μέτριος ρυθμός μάθησης επιτυγχάνει μια ισορροπία μεταξύ του κινδύνου προβλημάτων σύγκλισης με πολύ υψηλό ποσοστό μάθησης και της αργής σύγκλισης με πολύ χαμηλό ρυθμό μάθησης. Πειραματισμοί οδήγησαν στην επιλογή αυτής της τιμής, διασφαλίζοντας ότι το μοντέλο προσαρμόζει αποτελεσματικά τις παραμέτρους του χωρίς να υπερβαίνει τη βέλτιστη λύση.

Μέγεθος παρτίδας (Batch Size):

Το μέγεθος παρτίδας αντιπροσωπεύει τον αριθμό των σημείων δεδομένων που χρησιμοποιούνται σε κάθε επανάληψη του αλγορίθμου βελτιστοποίησης. Στον κώδικα, επιλέχθηκε ένα μέγεθος παρτίδας 32 για την εκπαίδευση του μοντέλου MLP. Αυτό το μέτριο μέγεθος παρτίδας προτιμάται συχνά, καθώς εξισορροπεί τα πλεονεκτήματα της υπολογιστικής απόδοσης με την ικανότητα του μοντέλου να καταγράφει περίπλοκα μοτίβα στα δεδομένα. Τα μεγαλύτερα μεγέθη παρτίδων μπορούν να επιταχύνουν την εκπαίδευση, αλλά μπορεί να οδηγήσουν σε προκλήσεις σύγκλισης, ενώ μικρότερες παρτίδες μπορεί να ενισχύσουν την προσαρμοστικότητα αλλά μπορεί να είναι υπολογιστικά ακριβές.

Αριθμός κρυφών στρωμάτων και νευρώνων (Number of Hidden Layers and Neurons):

Η επιλογή των layers και νευρώνων αντικατοπτρίζει μια ισορροπία μεταξύ της πολυπλοκότητας του μοντέλου και της αποφυγής υπερπροσαρμογής (overfitting). Δυο μόνο κρυφά στρώματα είναι συχνά αρκετά για την καταγραφή μη γραμμικών σχέσεων σε δεδομένα χρονοσειρών και ο αριθμός των νευρώνων είναι αρκετά μέτριος για να αποτρέψει την υπερβολική πολυπλοκότητα του μοντέλου.

Πράξη εξισορρόπησης (Balancing Act):

Ο επιλεγμένος ρυθμός εκμάθησης και το μέγεθος παρτίδας επιτυγχάνουν μια ισορροπία μεταξύ αποτελεσματικής σύγκλισης και υπολογιστικής απόδοσης.

Μέτρο στην πολυπλοκότητα (Moderation in Complexity):

Η χρήση των κρυφών στρώματων με 8 νευρώνες αντικατοπτρίζει ένα μέτριο επίπεδο πολυπλοκότητας του μοντέλου, αποτρέποντας την υπερπροσαρμογή ενώ επιτρέπει στο μοντέλο να συλλαμβάνει περίπλοκες σχέσεις στα δεδομένα χρονοσειρών.

Εμπειρική Επικύρωση (Empirical Validation):

Αυτές οι επιλογές υπερπαραμέτρων είναι συχνά το αποτέλεσμα εμπειρικής επικύρωσης μέσω πειραματισμού, όπου ελέγχονται διαφορετικές τιμές και επιλέγονται αυτές που αποδίδουν τη βέλτιστη απόδοση.

Οι επιλογές υπερπαραμέτρων, που επηρεάζονται από την προσεκτική εξέταση της συμπεριφοράς και της απόδοσης του μοντέλου κατά τη διάρκεια της εκπαίδευσης, συμβάλλουν στην ευρωστία και την αποτελεσματικότητα του μοντέλου πρόβλεψης χρονοσειρών MLP

4.5 Arima

Βασικός άξονας της ανάλυσης χρονοσειρών, το μοντέλο Αυτοπαλινδρομικός ολοκληρωμένος κινούμενος μέσος όρος (Autoregressive Integrated Moving Average/ARIMA) χρονολογείται από τα μέσα του 20ου αιώνα και αναπτύχθηκε από το πρωτοποριακό έργο των στατιστικολόγων George Box και Gwilym Jenkins. Στα τέλη της δεκαετίας του 1960, αυτή η στατιστική τεχνική εξελίχθηκε ως συνέχεια των προηγούμενων μοντέλων κινητού μέσου όρου (MA) και αυτοπαλινδρομικής (AR). Ο Box και ο Jenkins στόχευαν να αναπτύξουν ένα ενιαίο, συνεκτικό πλαίσιο που θα μπορούσε να αναγνωρίσει περίπλοκα χρονικά μοτίβα σε διάφορα σύνολα δεδομένων.

Όπως υποδηλώνει το όνομά του, το ARIMA συνδυάζει τρία θεμελιώδη στοιχεία: Κινητό μέσο όρο (MA), Ολοκλήρωση (I) και Αυτοπαλίνδρομο (AR). Ο αντίκτυπος των προηγούμενων παρατηρήσεων καταγράφεται από το στοιχείο AR, η μη σταθερότητα αντιμετωπίζεται από το στοιχείο I μέσω διαφοροποίησης και το αποτέλεσμα εξομάλυνσης

των προηγούμενων υπολειπόμενων σφαλμάτων αντιμετωπίζεται από το στοιχείο MA. Αυτός ο συνδυασμός παράγει ένα ευέλικτο και αποτελεσματικό μέσο πρόβλεψης.

Από τη δημιουργία της, η ARIMA έχει καθιερωθεί ως ένα αξιόπιστο εργαλείο σε διάφορους κλάδους, συμπεριλαμβανομένων της μετεωρολογίας, των οικονομικών και άλλων. Το γεγονός ότι συνεχίζει να είναι δημοφιλές δείχνει πόσο καλά μοντελοποιεί και προβλέπει γεγονότα που εξαρτώνται από το χρόνο. Η ανάλυση χρονοσειρών συνεχίζει να χρησιμοποιείται χάρη στη διαχρονική αξία και την κατάσταση πολύτιμων περιουσιακών στοιχείων της ARIMA, παρόλο που οι τεχνολογικές βελτιώσεις συνεχίζουν να μεταμορφώνουν τον κόσμο της ανάλυσης δεδομένων.

4.5.1 Επιλογή μοντέλου

Για ακριβή πρόβλεψη, πρέπει να επιλεγεί το σωστό μοντέλο. Ο αλγόριθμος Autoregressive Integrated Moving Average (ARIMA) είναι ένα ευρέως χρησιμοποιούμενο εργαλείο στην ανάλυση χρονοσειρών. Αλλά για να βρεθούν καλύτερες τιμές για τις παραμέτρους του (p , d και q) απαιτείται μια μεθοδική προσέγγιση. Αυτή είναι μια αναζήτηση πλέγματος στην οποία αξιολογούνται πολλές μεταθέσεις των p , d και q . Οι χαμηλότερες τιμές AIC υποδεικνύουν μοντέλα που ταιριάζουν καλύτερα και το Akaike Information Criterion (AIC) χρησιμοποιείται συχνά ως κριτήριο για την επιλογή μοντέλου.

4.5.2 Αποκρυπτογράφηση AIC και PDQ στην ανάλυση χρονοσειρών

Στη σφαίρα της ανάλυσης χρονοσειρών, η επιλογή του σωστού μοντέλου μοιάζει με το ξεκλείδωμα των μυστικών που κρύβονται στα χρονικά δεδομένα. Το μοντέλο Autoregressive Integrated Moving Average (ARIMA) ξεχωρίζει ως μια εύρωστη επιλογή, αλλά ο καθορισμός των βέλτιστων παραμέτρων (p , d , q) απαιτεί μια λεπτή κατανόηση του κριτηρίου πληροφοριών Akaike (AIC) και της περίπλοκης διαδικασίας εύρεσης του αόριστου PDQ

4.5.3. Η σημασία του AIC:

Κατά την αναζήτηση του μοντέλου ARIMA που ταιριάζει καλύτερα, το AIC λειτουργεί ως φάρος φωτός. Προκειμένου να αποφευχθεί η υπερβολική προσαρμογή των δεδομένων, αυτό το κριτήριο επιτυγχάνει μια ισορροπία μεταξύ της πολυπλοκότητας και της καλής προσαρμογής του μοντέλου. Οι χαμηλότερες τιμές AIC, γενικά, υποδεικνύουν ένα πιο λιτό μοντέλο που αποτυπώνει αποτελεσματικά τα βασικά μοτίβα που εμφανίζονται στις χρονοσειρές. Το AIC χρησιμεύει ως πυξίδα στην περιπέτεια επιλογής μοντέλου, δείχνοντάς προς την κατεύθυνση της ιδανικής διαμόρφωσης παραμέτρων.

4.5.4. Ερμηνεία PDQ: Οι τρεις παράμετροι του μοντέλου ARIMA p , d και q .

Το ποσό των παρατηρήσεων υστέρησης που ενσωματώνονται στο μοντέλο υποδεικνύεται από τη συνάρτηση p (Αυτοπαλινδρομική σειρά). Με πιο απλά λόγια, μετρά τον αντίκτυπο των προηγούμενων παρατηρήσεων στην τρέχουσα. Όταν το p είναι υψηλότερο, το μοντέλο λαμβάνει υπόψη περισσότερα δεδομένα του παρελθόντος.

Ο αριθμός των φορών που διαφοροποιούνται οι πρωτογενείς παρατηρήσεις υποδεικνύεται με το σύμβολο d (Ολοκληρωμένη σειρά). Μια μέθοδος για τη σταθεροποίηση της χρονοσειράς και τη σταθεροποίηση της διακύμανσης και του μέσου όρου ονομάζεται διαφοροποίηση. Μια χρονοσειρά που είναι πιο σταθερή υποδεικνύεται με μεγαλύτερο d .

Το μέγεθος του παραθύρου του κινούμενου μέσου όρου αντικατοπτρίζεται σε q (τάξη κινούμενου μέσου όρου). Δείχνει τον βαθμό στον οποίο τα υπολειπόμενα σφάλματα από προηγούμενες παρατηρήσεις και την τρέχουσα παρατήρηση εξαρτώνται το ένα από το άλλο. Ένα ισχυρότερο αποτέλεσμα εξομάλυνσης υπονοείται από ένα υψηλότερο q .

4.5.5.Ο Μαθηματικός Χορός του PDQ:

Η αναζήτησή μας για το βέλτιστο PDQ περιλαμβάνει μια εξαντλητική αναζήτηση πλέγματος, λαμβάνοντας υπόψη διάφορους συνδυασμούς p , d και q . Η εξίσωση στα παρασκήνια προέρχεται από το ίδιο το μοντέλο ARIMA(p , d , q):

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t$$

- ϕ (Phi): η παράμετρος αυτή αντιπροσωπεύει την αυτοπαλίνδρομη (AR) συνιστώσα του μοντέλου. Ενσωματώνει τη γραμμική σχέση μεταξύ της τρέχουσας παρατήρησης και των προηγούμενων τιμών της. Στην ουσία, το ϕ καθορίζει πόσο επηρεάζουν οι προηγούμενες παρατηρήσεις την τρέχουσα. Για παράδειγμα, σε ένα σενάριο πρόβλεψης των τιμών των μετοχών, μια υψηλότερη τιμή ϕ υποδηλώνει ότι οι πρόσφατες κινήσεις των τιμών ασκούν ισχυρότερη επιρροή στην πρόβλεψη.
- θ (Theta): Από την άλλη πλευρά, το θ δηλώνει τη συνιστώσα του κινητού μέσου όρου (KM). Αποτυπώνει την εξάρτηση μεταξύ της τρέχουσας παρατήρησης και των παρελθοντικών όρων σφάλματος. Μια μεγαλύτερη τιμή του θ υποδηλώνει μεγαλύτερη εξάρτηση από τα σφάλματα του παρελθόντος για την πρόβλεψη των μελλοντικών αποτελεσμάτων. Στο παράδειγμά για τις τιμές των μετοχών, μια υψηλότερη τιμή του θ υποδηλώνει ότι τα σφάλματα πρόβλεψης του παρελθόντος διαδραματίζουν σημαντικότερο ρόλο στη διαμόρφωση των μελλοντικών προβλέψεων.
- ϵ (Εψιλον): Τέλος, το ϵ αντιπροσωπεύει τον όρο σφάλματος λευκού θορύβου στο υπόδειγμα. Υποδηλώνει τις τυχαίες διακυμάνσεις ή το απρόβλεπτο που ενυπάρχει στα δεδομένα και δεν μπορεί να εξηγηθεί από τις συνιστώσες AR και MA. Ουσιαστικά, το ϵ αντικατοπτρίζει την εναπομένουσα μεταβλητότητα στη χρονοσειρά μετά τη συνεκτίμηση των επιδράσεων AR και MA. Στο δικό μας πλαίσιο, το ϵ ενσωματώνει την απρόβλεπτη δυναμική της αγοράς ή τους εξωτερικούς παράγοντες που επηρεάζουν τις τιμές των μετοχών πέρα από τις μοντελοποιημένες συνιστώσες.

Με την κατανόηση των διαφοροποιημένων ρόλων των ϕ , θ και ϵ , οι αναλυτές μπορούν να αποκρυπτογραφήσουν τα υποκείμενα μοτίβα εντός των δεδομένων των χρονοσειρών τους και να δημιουργήσουν ακριβέστερες προβλέψεις προσαρμοσμένες στις ιδιαιτερότητες του τομέα.

4.5.6. Πλοήγηση στο λαβύρινθο PDQ:

Καθώς πλοηγούμαστε στον λαβύρινθο PDQ, ο επιλεγμένος συνδυασμός θα πρέπει να αποδίδει το χαμηλότερο AIC, υποδεικνύοντας ένα μοντέλο που εξισορροπεί την πολυπλοκότητα και την εφαρμογή. Ένα χαμηλότερο AIC σημαίνει ένα μοντέλο που εξηγεί τη μεταβλητότητα των δεδομένων χωρίς περιττή πολυπλοκότητα - μια βασική αρχή στην τέχνη της ανάλυσης χρονοσειρών.

Συμπερασματικά, η αλληλεπίδραση μεταξύ του AIC και του διαφοροποιημένου χορού του PDQ στη διαδικασία επιλογής μοντέλου ARIMA αντιπροσωπεύει μια περίπλοκη αλλά κρίσιμη πτυχή της ανάλυσης χρονοσειρών. Η αγκαλιά των μαθηματικών περιπλοκών πίσω από αυτές τις παραμέτρους δίνει τη δυνατότητα στους αναλυτές να ξεδιαλύνουν τα χρονικά μυστήρια που κρύβονται στα σύνολα δεδομένων τους. Το ταξίδι συνεχίζεται με μια σχολαστική εξερεύνηση της προσαρμογής και της επικύρωσης του μοντέλου, καθώς πλησιάζουμε περισσότερο στην εξοικείωση με την τέχνη της πρόβλεψης χρονοσειρών.

4.5.7. Τάσεις

Στο πλαίσιο των δεδομένων χρονοσειρών, οι τάσεις αναφέρονται σε μακροπρόθεσμα μοτίβα ή κινήσεις που παρουσιάζουν μια σταθερή κατεύθυνση για μια εκτεταμένη περίοδο. Ακολουθούν ορισμένα παραδείγματα τάσεων που παρατηρούνται συνήθως σε διάφορους τύπους δεδομένων χρονοσειρών:

Ανοδική τάση (Upward Trend):

Μια ανοδική τάση υποδηλώνει μια σταθερή αύξηση των τιμών των χρονοσειρών με την πάροδο του χρόνου. Για παράδειγμα, οι τιμές κλεισίματος ενός αυξανόμενου τεχνολογικού αποθέματος, όπως το Amazon, μπορεί να παρουσιάζουν μια ανοδική τάση καθώς η εταιρεία επεκτείνει και δημιουργεί αυξανόμενα έσοδα.

Πτωτική τάση (Downward Trend):

Αντίθετα, μια πτωτική τάση συνεπάγεται διαρκή μείωση των τιμών των χρονοσειρών. Η ζήτηση για ένα προϊόν που μειώνεται ή τα έσοδα μιας προβληματικής εταιρείας μπορεί να παρουσιάσουν πτωτική τάση.

Εποχιακή τάση (Seasonal Trend):

Οι εποχικές τάσεις εμφανίζονται με κυκλικό τρόπο, συχνά επηρεαζόμενες από τακτικούς, επαναλαμβανόμενους παράγοντες όπως εποχές, διακοπές ή γεγονότα. Οι λιανικές πωλήσεις, για παράδειγμα, ενδέχεται να παρουσιάσουν άνοδο κατά την περίοδο των εορτών κάθε χρόνο, δημιουργώντας μια επαναλαμβανόμενη εποχιακή τάση.

Κυκλική τάση (Cyclical Trend):

Οι κυκλικές τάσεις περιλαμβάνουν περιοδικές διακυμάνσεις που δεν συνδέονται απαραίτητα με εποχές που βασίζονται στο ημερολόγιο. Οι οικονομικοί κύκλοι, όπως οι φάσεις της άνθησης και της κατάρρευσης, αντιπροσωπεύουν κυκλικές τάσεις στις χρηματοοικονομικές χρονοσειρές.

Επίπεδη ή Χωρίς Τάσεις (Flat or No Trend):

Σε ορισμένες περιπτώσεις, τα δεδομένα χρονοσειρών μπορεί να μην παρουσιάζουν σαφή ανοδική ή καθοδική τάση, υποδεικνύοντας μια επίπεδη τάση. Αυτό μπορεί να συμβεί όταν μια αγορά ή ένα προϊόν διατηρεί σταθερό επίπεδο δραστηριότητας με την πάροδο του χρόνου.

- **Τάση μεταβλητότητας (Volatility Trend):**

Οι τάσεις μεταβλητότητας αντιπροσωπεύουν διακυμάνσεις στο μέγεθος των διακυμάνσεων. Οι χρηματοπιστωτικές αγορές συχνά βιώνουν περιόδους υψηλής ή χαμηλής μεταβλητότητας, επηρεάζοντας το εύρος των διακυμάνσεων των τιμών.

- **Εκθετική Ανάπτυξη (Exponential Growth):**

Οι εκθετικές τάσεις ανάπτυξης συνεπάγονται ταχεία αύξηση των τιμών με την πάροδο του χρόνου, με το ρυθμό ανάπτυξης να είναι ανάλογος της τρέχουσας αξίας. Αυτό θα μπορούσε να ισχύει για την υιοθέτηση μιας νέας τεχνολογίας ή την εξάπλωση μιας επιδημίας ιού.

Η κατανόηση αυτών των τάσεων είναι ζωτικής σημασίας για την πρόβλεψη των χρονοσειρών, καθώς επιτρέπει την επιλογή των κατάλληλων μοντέλων (όπως η ARIMA) να συλλάβει και να προβλέψει τα υποκείμενα πρότυπα στα δεδομένα. Διαφορετικές τάσεις

μπορεί να απαιτούν συγκεκριμένες τεχνικές μοντελοποίησης για την επίτευξη ακριβών προβλέψεων.

4.5.8 Προσαρμογή και αξιολόγηση μοντέλου (Model Fitting and Evaluation)

Υπολειμματική Ανάλυση (Residual Analysis):

Τα residuals, οι διαφορές μεταξύ προβλεπόμενων και πραγματικών τιμών, αναλύονται. Τα residual διαγράμματα και τα ιστογράμματα παρέχουν πληροφορίες για την απόδοση του μοντέλου, βοηθώντας στον εντοπισμό προτύπων ή προκαταλήψεων.

Ελάχιστη-Μέγιστη Κλιμάκωση (Min-Max Scaling) Ο κώδικας χρησιμοποιεί κλίμακα Min-Max για να ομαλοποιήσει τα δεδομένα εκπαίδευσης και δοκιμών. Αυτό διασφαλίζει ότι το μοντέλο εκπαιδεύεται σε συνεπή κλίμακα, διευκολύνοντας τη δίκαιη αξιολόγηση.

Εκπαίδευση μοντέλου ARIMA (Model Training):

Το μοντέλο ARIMA εκπαιδεύεται στα κλιμακωμένα δεδομένα εκπαίδευσης. Το ιστορικό των παρατηρήσεων διατηρείται και οι προβλέψεις γίνονται επαναληπτικά.

Αντίστροφος μετασχηματισμός (Inverse Transform):

Οι προβλέψεις μετασχηματίζονται αντίστροφα στην αρχική τους κλίμακα για ουσιαστική αξιολόγηση σε σχέση με τις πραγματικές τιμές των μετοχών.

Μετρήσεις αξιολόγησης (Evaluation Metrics):

Ο κώδικας υπολογίζει διάφορες μετρήσεις αξιολόγησης, όπως το μέσο τετράγωνο σφάλμα (MSE), το ριζικό μέσο τετράγωνο σφάλμα (RMSE), το μέσο απόλυτο σφάλμα (MAE) και το μέσο απόλυτο ποσοστό σφάλματος (MAPE). Αυτές οι μετρήσεις ποσοτικοποιούν την ακρίβεια και την αξιοπιστία του μοντέλου.

Σωστές προβλέψεις για σημάδια (Correct Sign Predictions):

Το ποσοστό των σωστών προβλέψεων πρόσημου προσδιορίζεται, αξιολογώντας την ικανότητα του μοντέλου να συλλαμβάνει τις κατευθυντικές τάσεις στις τιμές των μετοχών.

4.6. Μετασχηματιστές (Transformers)

Η ιστορία των Transformers στον τομέα της τεχνητής νοημοσύνης είναι σχετικά πρόσφατη αλλά πρωτοποριακή. Περιστρέφεται γύρω από μια αρχιτεκτονική νευρωνικού δικτύου που ονομάζεται Transformer, η οποία έφερε επανάσταση στην επεξεργασία φυσικής γλώσσας (NLP) και σε άλλες εργασίες διαδοχικών δεδομένων. Το μοντέλο Transformer παρουσιάστηκε από τους Vaswani et al. στην εργασία τους του 2017 με τίτλο "Attention is All You Need" (Η προσοχή είναι το μόνο που χρειάζεστε). Σε αντίθεση με προηγούμενα διαδοχικά μοντέλα, όπως τα επαναλαμβανόμενα νευρωνικά δίκτυα (RNN) και τα δίκτυα μακράς βραχυπρόθεσμης μνήμης (LSTM), το μοντέλο Transformer βασίζεται εξ ολοκλήρου σε μηχανισμούς προσοχής χωρίς καμία επανάληψη. Αυτή η απομάκρυνση από την επανάληψη ήταν μια σημαντική αλλαγή στον τομέα της βαθιάς μάθησης. Η βασική ιδέα πίσω από τους Transformers είναι ο μηχανισμός self-attention, ο οποίος επιτρέπει στο μοντέλο να σταθμίζει τη σημασία των διαφορετικών tokens εισόδου κατά την επεξεργασία μιας ακολουθίας. Αυτός ο μηχανισμός attention επιτρέπει στα Transformers να συλλαμβάνουν εξαρτήσεις μεγάλης εμβέλειας στα δεδομένα πιο αποτελεσματικά από τα παραδοσιακά επαναλαμβανόμενα μοντέλα. Επιπλέον, τα Transformers αξιοποιούν κωδικοποιήσεις θέσης για να κατανοήσουν τη σειρά των tokens σε μια ακολουθία. Ένα από τα βασικά πλεονεκτήματα των Transformers είναι η ικανότητά τους να παραλληλίζουν τον υπολογισμό σε όλα τα tokens μιας ακολουθίας, οδηγώντας σε ταχύτερη εκπαίδευση και εξαγωγή συμπερασμάτων σε σύγκριση με τα επαναλαμβανόμενα μοντέλα. Αυτός ο παραλληλισμός διευκολύνεται από τον μηχανισμό self-attention, ο οποίος μπορεί να επεξεργάζεται παράλληλα όλα τα tokens.

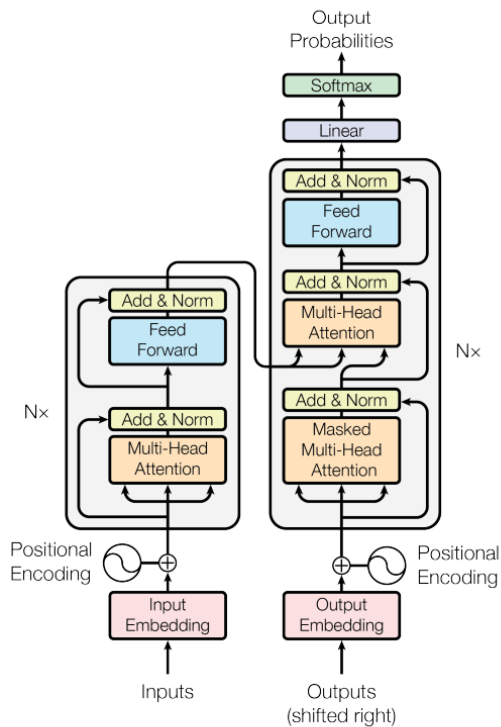


Figure 1: The Transformer - model architecture.

Εικόνα 11. Αρχιτεκτονική μοντέλου [5]

Αρχιτεκτονική: Το μοντέλο Transformer αποτελείται από στοιβαγμένα στρώματα κωδικοποιητή και αποκωδικοποιητή, καθένα από τα οποία περιέχει μηχανισμούς αυτοπροσοχής (self-attention) και δίκτυα προώθησης. Ο κωδικοποιητής και ο αποκωδικοποιητής έχουν και οι δύο $N = 6$ πανομοιότυπα στρώματα, με υπολειμματικές συνδέσεις και κανονικοποίηση στρώματος γύρω από κάθε υποστρώμα.

Μηχανισμοί προσοχής: Το μοντέλο χρησιμοποιεί self-attention πολλαπλών κεφαλών, όπου τα ερωτήματα, τα κλειδιά και οι τιμές προβάλλονται γραμμικά και υποβάλλονται σε παράλληλη επεξεργασία. Χρησιμοποιείται Scaled Dot-Product Attention και η προσοχή πολλαπλών κεφαλών επιτρέπει στο μοντέλο να παρακολουθεί ταυτόχρονα διαφορετικές πτυχές της εισόδου. Οι μηχανισμοί προσοχής χρησιμοποιούνται σε στρώματα κωδικοποιητή-αποκωδικοποιητή και σε στρώματα self-attention εντός του κωδικοποιητή και του αποκωδικοποιητή.

Δίκτυα τροφοδοσίας από άποψη θέσης (Position-wise Feed-Forward Networks): Εκτός από τα υποστρώματα προσοχής, κάθε στρώμα περιέχει ένα δίκτυο προώθησης που

εφαρμόζεται σε κάθε θέση ανεξάρτητα. Αυτό το δίκτυο αποτελείται από δύο γραμμικούς μετασχηματισμούς με συνάρτηση ενεργοποίησης ReLU.

Ενσωματώσεις και Softmax (Embeddings και Softmax): Οι μαθησιακές ενσωματώσεις μετατρέπουν τα tokens εισόδου και εξόδου σε διανύσματα και οι συναρτήσεις softmax παράγουν προβλεπόμενες πιθανότητες επόμενων tokens. Το μοντέλο μοιράζεται τον ίδιο πίνακα βαρών μεταξύ των στρωμάτων ενσωμάτωσης και των γραμμικών μετασχηματισμών προ softmax.

Κωδικοποίηση θέσης (Positional Encoding) : Δεδομένου ότι το μοντέλο δεν διαθέτει αναδρομή ή συνέλιξη, προστίθενται κωδικοποιήσεις θέσης για να υποδεικνύουν τη θέση των tokens στην ακολουθία. Χρησιμοποιούνται ημιτονοειδείς κωδικοποιήσεις θέσης, επιτρέποντας στο μοντέλο να μαθαίνει αποτελεσματικά τις σχετικές θέσεις.

Τα στρώματα self-attention προσφέρουν πλεονεκτήματα όσον αφορά την παραλληλοποίηση και τα μήκη διαδρομής, καθιστώντας τα κατάλληλα για την αποτελεσματική εκμάθηση εξαρτήσεων μεγάλης εμβέλειας. Συνοπτικά, το μοντέλο Transformer αξιοποιεί μηχανισμούς self-attention και κωδικοποιήσεις θέσης για την επεξεργασία διαδοχικών δεδομένων, προσφέροντας πλεονεκτήματα έναντι των παραδοσιακών αναδρομικών και συνελκτικών αρχιτεκτονικών όσον αφορά την υπολογιστική αποδοτικότητα και την ικανότητα σύλληψης εξαρτήσεων μεγάλης εμβέλειας. Συνολικά, ενώ οι Transformers υπερέχουν στην καταγραφή πολύπλοκων εξαρτήσεων σε διαδοχικά δεδομένα, μπορεί να μην είναι η καταλληλότερη επιλογή για εργασίες πρόβλεψης χρονοσειρών λόγω της μορφής εισόδου σταθερού μήκους, των περιορισμένων μηχανισμών χρονικής κωδικοποίησης, της υπολογιστικής επιβάρυνσης και της αποδοτικότητας των δεδομένων.

Κεφαλαίο 5 .Αποτελέσματα και σύγκριση

5.1 Αποτελέσματα MLP

Η διαδικασία εκπαίδευσης του μοντέλου Multilayer Perceptron (MLP) περιλάμβανε επανάληψη πάνω από 150 εποχές, κατά τη διάρκεια των οποίων το μοντέλο ενημέρωνε συνεχώς τις παραμέτρους του για να ελαχιστοποιήσει την απώλεια εκπαίδευσης. Κατά τη διάρκεια αυτών των εποχών, η απόδοση του μοντέλου βελτιωνόταν σταθερά, όπως υποδεικνύεται από τις μειούμενες τιμές των απωλειών. Ξεκινώντας με μια σχετικά υψηλή απώλεια 44,4062 στη 10η εποχή, το μοντέλο υπέστη σημαντική βελτίωση, με αποκορύφωμα μια σημαντικά μειωμένη απώλεια 2,1338 στην 150η εποχή. Αυτή η προοδευτική μείωση των απωλειών καταδεικνύει την ικανότητα του μοντέλου να προσαρμόζεται και να μαθαίνει από τα δεδομένα εκπαίδευσης σε διαδοχικές επαναλήψεις, υποδεικνύοντας αποτελεσματικές προσαρμογές παραμέτρων που καθοδηγούνται από την οπισθοδιάδοση(backpropagation).

```
Epoch [10/150], Loss: 44.4062
Epoch [20/150], Loss: 41.5807
Epoch [30/150], Loss: 38.7495
Epoch [40/150], Loss: 35.9651
Epoch [50/150], Loss: 33.0072
Epoch [60/150], Loss: 29.8465
Epoch [70/150], Loss: 26.4857
Epoch [80/150], Loss: 22.9597
Epoch [90/150], Loss: 19.3325
Epoch [100/150], Loss: 15.6971
Epoch [110/150], Loss: 12.1774
Epoch [120/150], Loss: 8.9222
Epoch [130/150], Loss: 6.0870
Epoch [140/150], Loss: 3.8023
Epoch [150/150], Loss: 2.1338
Training finished.
```

Εικόνα 12Τιμές απωλειών στην διαδικασία εκπαίδευσης για 150 epochs.

Μετά την ολοκλήρωση της εκπαίδευσης, η απόδοση του μοντέλου αξιολογήθηκε περαιτέρω χρησιμοποιώντας βασικές μετρικές αξιολόγησης. Το μέσο τετραγωνικό σφάλμα (MSE) ποσοτικοποίησε τη μέση τετραγωνική διαφορά μεταξύ των προβλέψεων του μοντέλου και των πραγματικών τιμών στην ακολουθία εκπαίδευσης, δίνοντας τιμή **2,1409**. Επιπλέον, το μέσο απόλυτο σφάλμα (MAE) παρείχε εικόνα του μέσου μεγέθους των σφαλμάτων μεταξύ των προβλέψεων και των πραγματικών τιμών, με το μοντέλο MLP

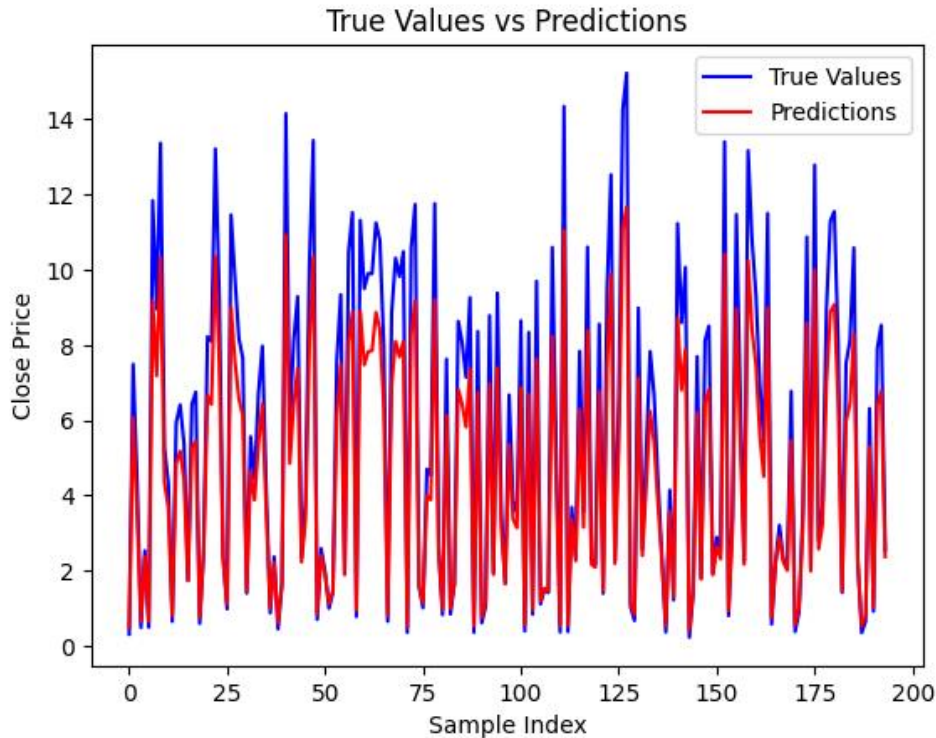
να επιτυγχάνει **MAE 1,0878**. Αυτές οι μετρικές προσφέρουν πολύτιμες πληροφορίες για την ακρίβεια και την ακρίβεια του μοντέλου στην πρόβλεψη των δεδομένων εκπαίδευσης.

Επιπλέον, η εξηγούμενη διακύμανση, η οποία μετρά το ποσοστό της διακύμανσης της μεταβλητής-στόχου (τιμές κλεισίματος στην προκειμένη περίπτωση) που καταγράφει το μοντέλο, υπολογίστηκε σε **0,9364**. Αυτή η υψηλή εξηγούμενη διακύμανση υπογραμμίζει την ικανότητα του μοντέλου MLP να συλλαμβάνει τα υποκείμενα πρότυπα και δυναμικές που υπάρχουν στην ακολουθία εκπαίδευσης. Ουσιαστικά, μια εξηγούμενη διακύμανση κοντά στο 1 υποδηλώνει ότι οι προβλέψεις του μοντέλου ταιριάζουν στενά με τη μεταβλητότητα των παρατηρούμενων δεδομένων.

```
MSE: 2.1409
MAE: 1.0878
Explained Variance: 0.9364
```

Εικόνα 13. Μετρικές

Συνοπτικά, τα αποτελέσματα αξιολόγησης αναδεικνύουν την αποτελεσματικότητα του μοντέλου MLP στην εκμάθηση και τη γενίκευση από την παρεχόμενη ακολουθία εκπαίδευσης. Ελαχιστοποιώντας τη συνάρτηση απωλειών και επιτυγχάνοντας ευνοϊκές επιδόσεις σε μετρικές αξιολόγησης όπως το **MSE**, το **MAE** και η Εξηγούμενη Διακύμανση, το μοντέλο MLP αποδεικνύει την ικανότητά του να συλλαμβάνει και να αναπαριστά αποτελεσματικά τα υποκείμενα πρότυπα στα δεδομένα εκπαίδευσης, αναδεικνύοντας έτσι τη χρησιμότητά του στην πρόβλεψη των τιμών των μετοχών.



Εικόνα 14 Γράφημα αποτελεσμάτων προβλεψεων

5.2 Αποτελέσματα ARIMA

Το μοντέλο ARIMA (Autoregressive Integrated Moving Average) εκπαιδεύτηκε στα παρεχόμενα δεδομένα εκπαίδευσης και η απόδοσή του αξιολογήθηκε στο σύνολο δεδομένων δοκιμής. Παρά την εκπαίδευσή του, το μοντέλο ARIMA παρουσίασε σημάδια υπερπροσαρμογής, οδηγώντας σε μη βέλτιστες επιδόσεις στα δεδομένα δοκιμής. Το μέσο τετραγωνικό σφάλμα (**MSE**) στην αρχική κλίμακα υπολογίστηκε σε **24,719**, υποδεικνύοντας σημαντική απόκλιση μεταξύ των προβλέψεων του μοντέλου και των πραγματικών τιμών. Το μέσο τετραγωνικό σφάλμα ρίζας (**RMSE**) ποσοτικοποίησε περαιτέρω την απόκλιση αυτή, δίνοντας τιμή **4,972**. Επιπλέον, το μέσο απόλυτο σφάλμα (**MAE**) στην αρχική κλίμακα υπολογίστηκε σε **3,848**, αντιπροσωπεύοντας το μέσο μέγεθος των σφαλμάτων μεταξύ των προβλέψεων του μοντέλου και των πραγματικών τιμών.

Επιπλέον, το μέσο απόλυτο ποσοστιαίο σφάλμα (**MAPE**) προσδιορίστηκε σε **1,331%**, υποδεικνύοντας τη μέση ποσοστιαία απόκλιση των προβλέψεων του μοντέλου από τις παρατηρούμενες τιμές. Αυτή η μετρική παρέχει πληροφορίες σχετικά με την ακρίβεια των προβλέψεων του μοντέλου σε σχέση με τα πραγματικά δεδομένα. Επιπλέον, το Ποσοστό σωστών προβλέψεων για το πρόσημο αξιολογήθηκε σε **52,33%**, υποδηλώνοντας ότι το μοντέλο προέβλεψε σωστά την κατεύθυνση (αύξηση ή μείωση) των κινήσεων των τιμών των μετοχών περίπου τις μισές φορές.

Ωστόσο, είναι ζωτικής σημασίας να σημειωθεί ότι παρά την ικανότητά του να καταγράφει μοτίβα στα δεδομένα εκπαίδευσης, το μοντέλο ARIMA δυσκολεύτηκε να γενικεύσει καλά σε άορατα δεδομένα, όπως αποδεικνύεται από τις σχετικά υψηλές μετρικές σφάλματος στο σύνολο δοκιμών. Αυτή η απόκλιση μεταξύ των επιδόσεων εκπαίδευσης και δοκιμής υποδεικνύει υπερπροσαρμογή, όπου το μοντέλο μπορεί να έχει μάθει θόρυβο ή συγκεκριμένα μοτίβα που υπάρχουν στα δεδομένα εκπαίδευσης και δεν γενικεύονται σε νέες παρατηρήσεις.



Εικόνα 15 Παράγει ένα γράφημα Plotly με δύο ίχνη: το ένα αντιπροσωπεύει τις πραγματικές τιμές κλεισίματος και το άλλο αντιπροσωπεύει τις προβλεπόμενες τιμές, επιτρέποντας την οπτική σύγκριση.

Ως αποτέλεσμα, ενώ το μοντέλο ARIMA μπορεί να αποδίδει επαρκώς στα δεδομένα εκπαίδευσης, η αποτελεσματικότητά του στην πραγματοποίηση ακριβών προβλέψεων σε άορατα δεδομένα, όπως το σύνολο δοκιμών, διακυβεύεται.

5.3 Αποτελέσματα XGBOOST

Το μοντέλο XGBoost επέδειξε ισχυρές επιδόσεις κατά τη διάρκεια της αξιολόγησης στα δεδομένα εκπαίδευσης, παρουσιάζοντας υποσχόμενες μετρήσεις ενδεικτικές της αποτελεσματικότητάς του στην πρόβλεψη των τιμών των μετοχών. Το μέσο απόλυτο

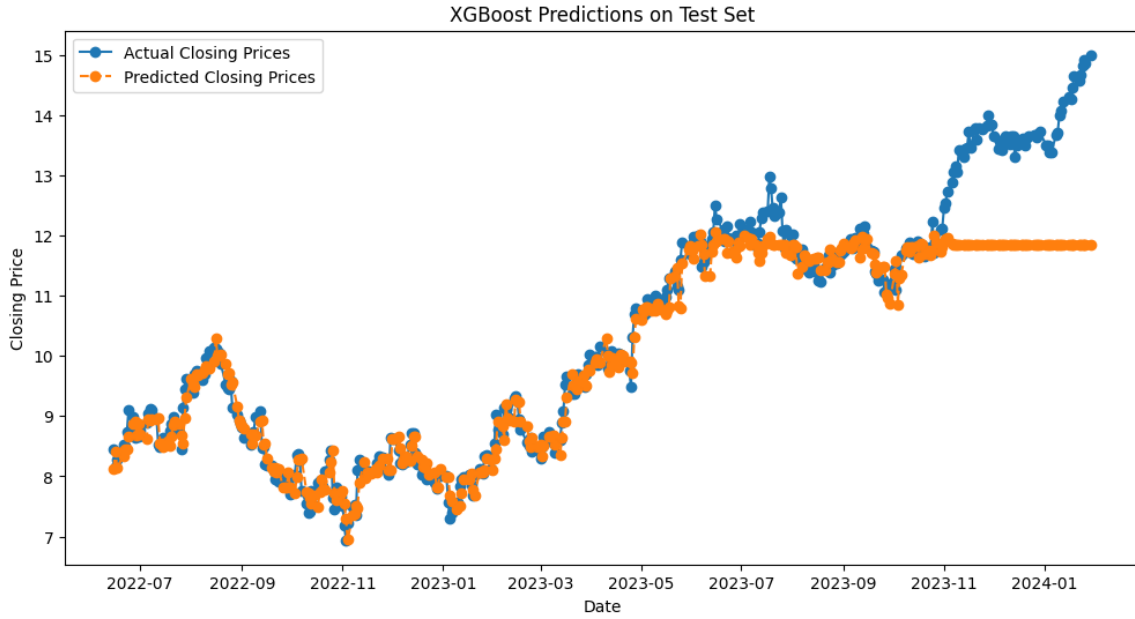
σφάλμα (MAE) υπολογίστηκε σε **0,44**, αντικατοπτρίζοντας το μέσο μέγεθος των σφαλμάτων μεταξύ των προβλέψεων του μοντέλου και των πραγματικών τιμών των μετοχών. Ομοίως, το μέσο τετραγωνικό σφάλμα (MSE) έδωσε τιμή **0,60**, ενώ το μέσο τετραγωνικό σφάλμα ρίζας (RMSE) ήταν **0,78**, ποσοτικοποιώντας περαιτέρω την ακρίβεια του μοντέλου και την απόκλιση από τις πραγματικές τιμές.

Επιπλέον, ο συντελεστής προσδιορισμού (R^2) υπολογίστηκε σε **0,86**, υποδεικνύοντας το ποσοστό της διακύμανσης της μεταβλητής-στόχου που μπορεί να προβλεφθεί από τις ανεξάρτητες μεταβλητές. Αυτή η μέτρηση υποδηλώνει ότι το μοντέλο XGBoost εξηγεί ένα σημαντικό μέρος της μεταβλητότητας των τιμών των μετοχών, αναδεικνύοντας την προβλεπτική του ικανότητα. Επιπλέον, η βαθμολογία της Εξηγούμενης Διασποράς, που υπολογίστηκε σε **0,88**, ενισχύει περαιτέρω την ικανότητα του μοντέλου να εξηγεί τη διακύμανση της μεταβλητής-στόχου.

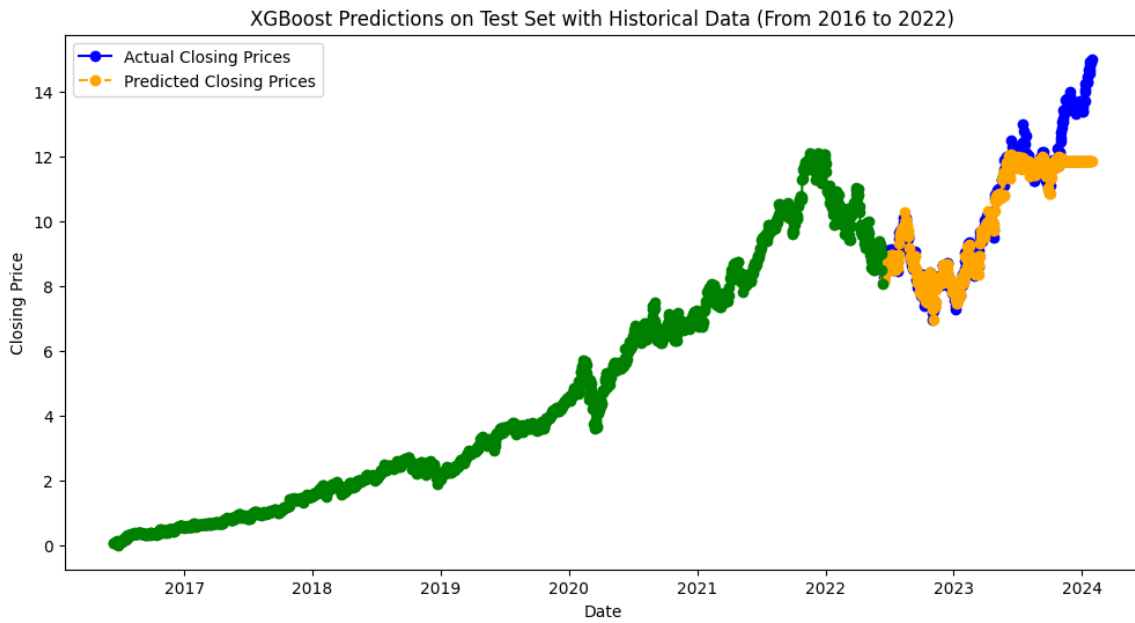
```
MAE: 0.44
MSE: 0.60
RMSE: 0.78
R^2: 0.86
Explained Variance: 0.88
```

Εικόνα 16. Μετρικές

Ωστόσο, είναι σημαντικό να σημειωθεί ότι, παρά την αρχική του επιτυχία, το μοντέλο XGBoost παρουσίασε ενδείξεις απόκλισης ή απόκλισης από τις πραγματικές τιμές των μετοχών κατά τους τελευταίους μήνες του 2023. Αυτή η απόκλιση υποδηλώνει ότι το μοντέλο μπορεί να δυσκολεύεται να συλλάβει απρόβλεπτα μοτίβα ή μετατοπίσεις στα δεδομένα, οδηγώντας ενδεχομένως σε λιγότερο αξιόπιστες προβλέψεις στο μέλλον. Επομένως, ενώ το μοντέλο XGBoost επιδεικνύει επάρκεια στην πρόβλεψη των τιμών των μετοχών βάσει ιστορικών δεδομένων, η απόδοσή του ενδέχεται να υπόκειται σε περιορισμούς όταν αντιμετωπίζει σημαντικές αλλαγές ή διακυμάνσεις στις συνθήκες της αγοράς.



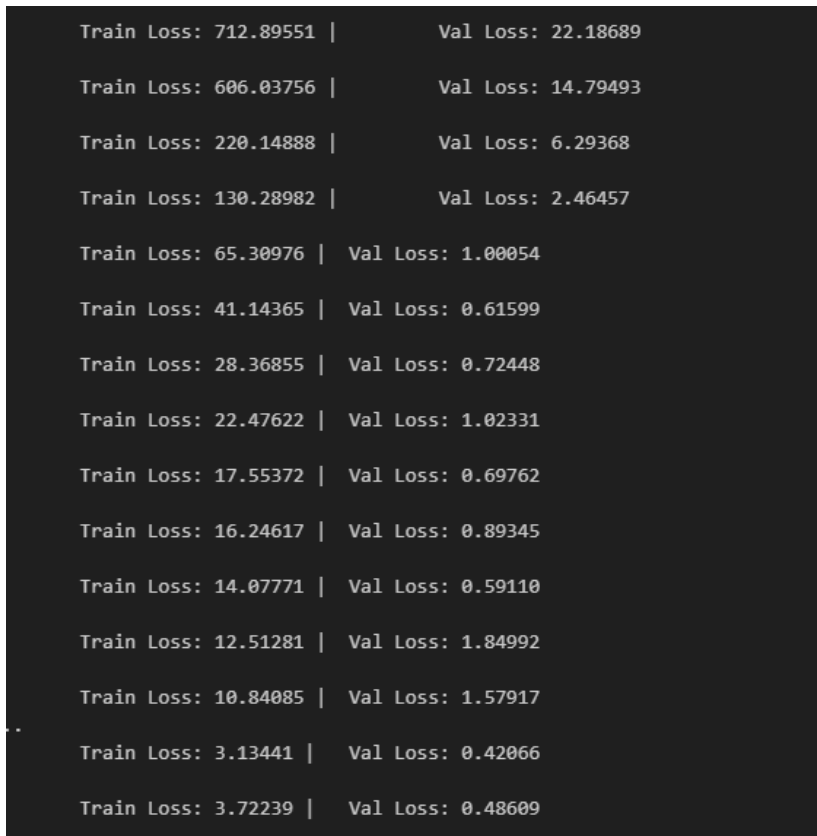
Εικόνα 17 Απόκλιση στους τελευταίους μήνες του prediction



Εικόνα 18 Απόκλιση στους τελευταίους μήνες του prediction

5.4. Αποτελέσματα LSTM

Το μοντέλο LSTM επέδειξε προοδευτική μείωση των απωλειών τόσο στην εκπαίδευση όσο και στην επικύρωση κατά τη διάρκεια της εκπαίδευσης, υποδεικνύοντας την ικανότητά του να μαθαίνει αποτελεσματικά από τα δεδομένα και να γενικεύει σε αθέατα παραδείγματα. Αρχικά, οι απώλειες εκπαίδευσης ανέρχονταν σε 712,89551, οι οποίες μειώθηκαν σταδιακά σε 10,84085 μέχρι την τελευταία εποχή, αναδεικνύοντας την τελειοποίηση του μοντέλου στην καταγραφή των υποκείμενων προτύπων στα δεδομένα εκπαίδευσης. Ομοίως, η απώλεια επικύρωσης παρουσίασε φθίνουσα τάση, ξεκινώντας από 22,18689 και συγκλίνοντας στο 1,57917, γεγονός που υποδηλώνει ότι η απόδοση του μοντέλου επεκτείνεται σε νέα δείγματα δεδομένων.



Train Loss: 712.89551	Val Loss: 22.18689
Train Loss: 606.03756	Val Loss: 14.79493
Train Loss: 220.14888	Val Loss: 6.29368
Train Loss: 130.28982	Val Loss: 2.46457
Train Loss: 65.30976	Val Loss: 1.00054
Train Loss: 41.14365	Val Loss: 0.61599
Train Loss: 28.36855	Val Loss: 0.72448
Train Loss: 22.47622	Val Loss: 1.02331
Train Loss: 17.55372	Val Loss: 0.69762
Train Loss: 16.24617	Val Loss: 0.89345
Train Loss: 14.07771	Val Loss: 0.59110
Train Loss: 12.51281	Val Loss: 1.84992
Train Loss: 10.84085	Val Loss: 1.57917
Train Loss: 3.13441	Val Loss: 0.42066
Train Loss: 3.72239	Val Loss: 0.48609

Εικόνα 19. Τιμές απωλειών στην διαδικασία εκπαίδευσης

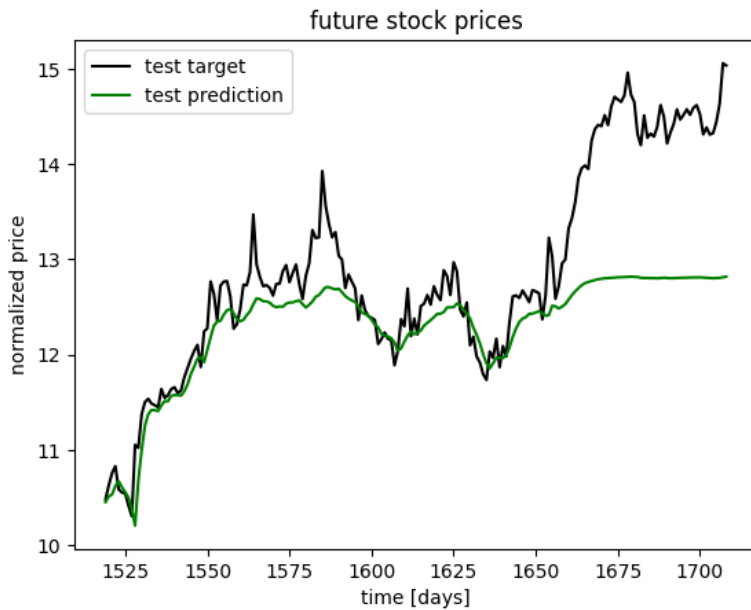
Ωστόσο, κατά την αξιολόγηση των επιδόσεων του μοντέλου LSTM στα δεδομένα δοκιμής, τα αποτελέσματα ήταν ανάμεικτα. Το μέσο απόλυτο σφάλμα (**MAE**) υπολογίστηκε σε **0,6259**, υποδεικνύοντας τη μέση απόλυτη διαφορά μεταξύ των προβλέψεων του μοντέλου και των πραγματικών τιμών των μετοχών. Η Εξηγούμενη διακύμανση (**explained variance**), που υπολογίστηκε ως **0,6203**, ποσοτικοποιεί το ποσοστό της διακύμανσης της μεταβλητής-στόχου που το μοντέλο εξηγεί. Ενώ αυτές οι μετρήσεις υποδηλώνουν ένα λογικό επίπεδο ακρίβειας πρόβλεψης, το μέσο τετραγωνικό σφάλμα (**RMSE**) της ρίζας **0,9049** και το μέσο τετραγωνικό σφάλμα (**MSE**) **0,8188** αναδεικνύουν τα υπολειπόμενα σφάλματα μεταξύ των

προβλεπόμενων και των πραγματικών τιμών, υποδηλώνοντας περιοχές όπου το μοντέλο μπορεί να εξακολουθεί να δυσκολεύεται να προβλέψει με ακρίβεια τις τιμές των μετοχών.

```
mae is: 0.6259327574147383
explained_variance is: 0.6202971105450784
rmse is: 0.9048821686756193
mse is: 0.818811739187092
r2 is: 0.3191615019239372
```

Εικόνα 20.Μετρικές

Επιπλέον, ο συντελεστής προσδιορισμού (R^2) εκτιμήθηκε σε **0,3192**, υποδεικνύοντας το ποσοστό της διακύμανσης της μεταβλητής-στόχου που εξηγείται από τις ανεξάρτητες μεταβλητές. Αυτή η μέτρηση υποδηλώνει ότι το μοντέλο LSTM συλλαμβάνει μόνο ένα μικρό μέρος της μεταβλητότητας των τιμών των μετοχών, περιορίζοντας ενδεχομένως την αποτελεσματικότητά του στην ακριβή πρόβλεψη μελλοντικών τάσεων.



Εικόνα 21.Πρόβλεψη πάνω στο target προβλέψεων..

Συνολικά, ενώ το μοντέλο LSTM επιδεικνύει πολλά υποσχόμενες ικανότητες στη μάθηση των χρονικών εξαρτήσεων και στη σύλληψη μοτίβων σε διαδοχικά δεδομένα, οι επιδόσεις του στα δεδομένα δοκιμής υποδηλώνουν περιθώρια βελτίωσης, ιδίως όσον αφορά τη μείωση των σφαλμάτων πρόβλεψης και την ενίσχυση της ακρίβειας πρόβλεψης.

Κεφάλαιο 6 . Πόρισμα

Η αξιολόγηση πολλών μοντέλων, συμπεριλαμβανομένων των LSTM, MLP και XGBoost, δείχνει ότι η πρόβλεψη μετοχών με τη χρήση νευρωνικών δικτύων έχει πλεονεκτήματα και μειονεκτήματα. Πολύπλοκες χρονικές σχέσεις μπορούν να αποτυπωθούν από τα νευρωνικά δίκτυα, ιδίως τα μοντέλα LSTM, τα οποία μπορούν επίσης να χρησιμοποιηθούν για την ανακάλυψη μοτίβων από παρελθοντικά δεδομένα μετοχών. Ωστόσο, μια σειρά από μεταβλητές, όπως η αρχιτεκτονική του μοντέλου, η ρύθμιση των υπερπαραμέτρων και η ποιότητα των δεδομένων, επηρεάζουν το πόσο καλά αυτά τα μοντέλα προβλέπουν τις μελλοντικές τιμές των μετοχών.

Η μείωση των απωλειών εκπαίδευσης και επικύρωσης κατά τη διάρκεια της εκπαίδευσης είναι μια πολλά υποσχόμενη επίδοση του μοντέλου LSTM, το οποίο είναι γνωστό για την ικανότητά του να χειρίζεται διαδοχικά δεδομένα. Η απόδοσή του σε δεδομένα δοκιμής παρουσιάζει ωστόσο ορισμένους περιορισμούς, καθώς μετριοπαθείς μετρικές όπως το MAE, η Εξηγούμενη Διακύμανση και το R2 δείχνουν ότι η μεταβλητότητα των τιμών των μετοχών καταγράφεται μόνο εν μέρει. Τα RMSE και MSE του μοντέλου υποδηλώνουν υπολειμματικά σφάλματα, αναδεικνύοντας περιοχές για βελτίωση της ακρίβειας πρόβλεψης. Τα RMSE και MSE του μοντέλου υποδηλώνουν υπολειμματικά σφάλματα, αναδεικνύοντας περιοχές για βελτίωση της ακρίβειας πρόβλεψης.

Παρόμοια, πολλά πλήρως συνδεδεμένα επίπεδα χρησιμοποιούνται από τα μοντέλα MLP για την ανακάλυψη σύνθετων συσχετίσεων στα δεδομένα, καθιστώντας τα μια ευέλικτη μέθοδο για την πρόβλεψη αποθεμάτων. Παρόλο που το μοντέλο MLP παρήγαγε ανταγωνιστικούς δείκτες επιδόσεων, όπως η Εξηγούμενη Διακύμανση και το MAE, εξακολουθούν να υπάρχουν ερωτήματα σχετικά με το πόσο καλά καταγράφει όλο το εύρος των μεταβολών των τιμών των μετοχών.

Τα χαμηλά MAE και RMSE του αλγορίθμου βαθμωτής ενίσχυσης XGBoost δείχνουν ότι αποδίδει πολύ καλά στην πρόβλεψη. Αλλά η ευπάθειά του στην υπερπροσαρμογή -η οποία γίνεται εμφανής στο τέλος της περιόδου αξιολόγησης, όταν οι προβλέψεις του μοντέλου αρχίζουν να αποκλίνουν από τις πραγματικές τιμές- δημιουργεί ανησυχίες σχετικά με το πόσο καλά συλλαμβάνει τις μακροπρόθεσμες τάσεις.

Εν κατακλείδι, τα νευρωνικά δίκτυα -όπως το XGBoost, το MLP και το LSTM- προσφέρουν χρήσιμα εργαλεία για την πρόβλεψη μετοχών, αλλά η αποτελεσματικότητά τους εξαρτάται από τη συνεχή βελτίωση και βελτιστοποίηση. Για τη βελτίωση της ακρίβειας πρόβλεψης και τη μείωση προβλημάτων όπως η υπερπροσαρμογή, οι βελτιώσεις στη μηχανική των χαρακτηριστικών, την προετοιμασία των δεδομένων και τη ρύθμιση του μοντέλου είναι ζωτικής σημασίας. Επιπλέον, ο συνδυασμός διαφορετικών μοντέλων ή μεθόδων συνόλου μπορεί να παρέχει μια πιο αξιόπιστη μέθοδο πρόβλεψης αποθεμάτων, αξιοποιώντας τα πλεονεκτήματα κάθε μοντέλου και αντισταθμίζοντας τα μειονεκτήματά του. Η ερευνά αυτή υποδεικνύει την σημασία της ερμηνείας των χαρακτηριστικών (πχ

.πόλεμος, οικονομική κρίση housing market,covid) και υπογραμμίζει πως ο ακριβής υπολογισμός πρόβλεψης τιμών είναι ακόμα αδύνατος καθώς οι προβλέψεις βασίζονται σε παρελθοντικές τάσεις ανεξάρτητες από τις περιστασιακές μεταβλητές . Συνολικά, τα νευρωνικά δίκτυα έχουν δυνατότητες για την πρόβλεψη μετοχών, αλλά απαιτείται περισσότερη μελέτη και πρόοδος για την πλήρη αξιοποίηση αυτού του πεδίου.

Βιβλιογραφία

- [1] Khashei, M., & Hajirahimi, Z. (2019). A comparative study of series arima/mlp hybrid models for stock price forecasting. *Communications in Statistics-Simulation and Computation*, 48(9), 2625-2640. <https://ikee.lib.auth.gr/record/320586/files/Kazaklari%20Fourkiotis.pdf>
- [2] Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175.
- [3] Yan, X., Weihan, W., & Chang, M. (2021). Research on financial assets transaction prediction model based on LSTM neural network. *Neural Computing and Applications*, 33, 257-270.
- [4] Yun, K. K., Yoon, S. W., & Won, D. (2021). Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process. *Expert Systems with Applications*, 186, 115716.
- [5] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [7] Μέξης, Φ. Δ. (2020). Ανάπτυξη τεχνητών νευρωνικών δικτύων για πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας.
- [8] O. I. Abiodun et al., "Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition," in *IEEE Access*, vol. 7, pp. 158820-158846, 2019, doi: 10.1109/ACCESS.2019.2945545.
- [9] Συμεωνίδης, Π. (2023). Μηχανική Μάθηση με Νευρωνικά Δίκτυα.
- [10] Βλαχοπούλου, Α. (2022). Ανάλυση διάδοσης κύματος με καθοδηγούμενα από τη φυσική τεχνητά νευρωνικά δίκτυα.
- [11] Σχινάς, Γ. (2023). Πρόβλεψη φωτοβολταϊκής παραγωγής με μεθόδους μηχανικής μάθησης και τεχνητά νευρωνικά δίκτυα.
- [12] Κονταξή, Α. (2023). Εφαρμογές της τεχνητής νοημοσύνης στα χρηματοοικονομικά μέσω της χρήσης των νευρωνικών δικτύων.
- [6] https://www.youtube.com/watch?v=6fXNiJXUheI&ab_channel=DATAcated
- [13] <https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/History/history1.html>
- [13] Islam, S., Elmekki, H., Elsebai, A., Bentahar, J., Drawel, N., Rjoub, G., & Pedrycz, W. (2023). A comprehensive survey on applications of transformers for deep learning tasks. *Expert Systems with Applications*, 122666.