

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ



IONIAN
UNIVERSITY

Ανάλυση και Εξόρυξη Δεδομένων από Διαδικτυακές
Παραγγελίες Προϊόντων για την Ανάπτυξη Συστήματος
Πρόβλεψης Πωλήσεων

Σφηναρολάκης Στέφανος
Επιβλέπων: Ανδρέας Καναβός

October 8, 2025

Επιβλέπων(ουσα)

Ανδρέας Καναβός, Αναπληρωτής Καθηγητής,
Iονιο Πανεπιστήμιο

Τριμελής Επιτροπή

Κάτια - Λήδα Κερμανίδου, Καθηγήτρια,
Iονιο Πανεπιστήμιο

Στέργιος Παλαμάς, Επίκουρος Καθηγητής,
Iονιο Πανεπιστήμιο

Ανδρέας Καναβός, Αναπληρωτής Καθηγητής,
Iονιο Πανεπιστήμιο

Περίληψη

Η παρούσα πτυχιακή εργασία επικεντρώνεται στην ανάπτυξη και αξιολόγηση ενός μοντέλου πρόβλεψης πωλήσεων μέσω μεθόδων ανάλυσης χρονοσειρών και τεχνικών μηχανικής μάθησης. Η ακριβής πρόβλεψη πωλήσεων αποτελεί κρίσιμο παράγοντα για τον στρατηγικό σχεδιασμό και την επιχειρησιακή αποτελεσματικότητα, καθώς επηρεάζει άμεσα τη διαχείριση αποθεμάτων, τον οικονομικό προγραμματισμό και τη λήψη επιχειρηματικών αποφάσεων.

Στο θεωρητικό μέρος παρουσιάζεται εκτενής βιβλιογραφική ανασκόπηση, η οποία καλύπτει τόσο τις παραδοσιακές στατιστικές μεθόδους, όπως τα μοντέλα ARIMA και SARIMA, όσο και σύγχρονες προσεγγίσεις της μηχανικής και βαθιάς μάθησης, όπως τα νευρωνικά δίκτυα και οι αλγόριθμοι συνόλων. Μέσα από αυτή την ανάλυση αναδεικνύονται τα πλεονεκτήματα και οι περιορισμοί κάθε μεθοδολογίας, καθώς και οι κατευθύνσεις για μελλοντική έρευνα στον τομέα της πρόβλεψης χρονοσειρών.

Στο πρακτικό μέρος, η εργασία εστιάζει στην ανάπτυξη ενός προσαρμοσμένου μοντέλου πρόβλεψης, υλοποιημένου σε Python με χρήση βιβλιοθηκών μηχανικής μάθησης. Το μοντέλο εκπαιδεύεται σε πραγματικά δεδομένα πωλήσεων, με διαδικασίες προεπεξεργασίας που διασφαλίζουν την ποιότητα και την αξιοπιστία των εισόδων. Η απόδοσή του αξιολογείται με μετρικές όπως το MSE, MAE, RMSE και ο συντελεστής προσδιορισμού R^2 .

Τα αποτελέσματα δείχνουν ότι το προτεινόμενο μοντέλο μπορεί να επιτύχει υψηλή ακρίβεια στις προβλέψεις, ξεπερνώντας σε ορισμένες περιπτώσεις την απόδοση παραδοσιακών στατιστικών προσεγγίσεων. Η εργασία καταλήγει με συμπεράσματα για την πρακτική χρησιμότητα των τεχνικών μηχανικής μάθησης στην πρόβλεψη πωλήσεων.

Λέξεις-Κλειδιά: Πρόβλεψη Πωλήσεων, Χρονοσειρές, Στατιστικά Μοντέλα, ARIMA, SARIMA, Μηχανική Μάθηση, Βαθιά Μάθηση, Νευρωνικά Δίκτυα, Αλγόριθμοι Συνόλων, Python, MSE, MAE, RMSE, R^2 .

Abstract

This thesis focuses on the development and evaluation of a sales forecasting model by utilizing time series analysis methods and machine learning techniques. Accurate sales forecasting is a critical factor for strategic planning and operational efficiency, as it directly affects inventory management, financial planning, and business decision-making.

In the theoretical part, an extensive literature review is presented, covering both traditional statistical methods, such as ARIMA and SARIMA models, and modern approaches from machine learning and deep learning, including neural networks and ensemble algorithms. This analysis highlights the advantages and limitations of each methodology, as well as future research directions in the field of time series forecasting.

In the practical part, the study focuses on the development of a customized forecasting model implemented in Python, making use of machine learning libraries. The model is trained on real sales data, with preprocessing procedures ensuring the quality and reliability of the inputs. The evaluation of its performance is conducted using metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2).

The results demonstrate that the proposed model can achieve high forecasting accuracy, in some cases surpassing the performance of traditional statistical approaches. The thesis concludes with insights regarding the practical utility of machine learning techniques in sales forecasting and suggests directions for future research, aiming to enhance predictive accuracy and broaden the applicability of such methods in real-world business environments.

Keywords: Sales Forecasting, Time Series, Statistical Models, ARIMA, SARIMA, Machine Learning, Deep Learning, Neural Networks, Ensemble Algorithms, Python, Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R^2 .

Πρόλογος και Ευχαριστίες

Θα ήθελα να εκφράσω τις ειλικρινείς μου ευχαριστίες σε όλους όσοι συνέβαλαν με οποιονδήποτε τρόπο στην ολοκλήρωση της παρούσας πτυχιακής εργασίας.

Πρώτα απ' όλα, ευχαριστώ θερμά τον επιβλέποντα καθηγητή μου για την καθοδήγηση, την πολύτιμη υποστήριξη και τις χρήσιμες παρατηρήσεις του σε όλη τη διάρκεια της μελέτης. Η συμβολή του υπήρξε καθοριστική τόσο στην επιστημονική κατεύθυνση όσο και στην προσωπική μου πρόοδο.

Ιδιαίτερες ευχαριστίες οφείλω στους γονείς μου, για την αμέριστη αγάπη, την κατανόηση και την ηθική στήριξη που μου παρείχαν όλα αυτά τα χρόνια. Η υπομονή και η πίστη τους στις δυνατότητές μου αποτέλεσαν βασικό κίνητρο για την ολοκλήρωση των σπουδών μου.

Τέλος, θα ήθελα να ευχαριστήσω τους φίλους μου για την πολύτιμη βοήθεια, την εμφύγωση και τη συμπαράστασή τους σε όλη αυτήν τη διαδρομή. Η παρουσία τους αποτέλεσε σημαντικό στήριγμα στην προσπάθειά μου.

Περιεχόμενα

| | | |
|----------|--|-----------|
| A | Εισαγωγή | 1 |
| A.1 | ΕΙΣΑΓΩΓΗ | 1 |
| A.2 | ΟΡΙΣΜΟΣ ΠΡΟΒΛΕΨΗΣ ΠΩΛΗΣΕΩΝ | 2 |
| B | Βιβλιογραφική ανασκόπηση | 3 |
| B.1 | ΕΙΣΑΓΩΓΗ ΣΤΗ ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ | 3 |
| B.2 | ΠΑΡΑΔΟΣΙΑΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ ΣΤΗΝ ΠΡΟΒΛΕΨΗ ΠΩΛΗΣΕΩΝ | 4 |
| B.2.1 | Γενικό πλαίσιο παραδοσιακών μεθόδων | 4 |
| B.3 | ΑΝΑΓΚΗ ΓΙΑ ΣΤΓΧΡΟΝΕΣ ΜΕΘΟΔΟΤΥΣ | 11 |
| B.4 | ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ (MACHINE LEARNING) ΓΙΑ ΠΡΟΒΛΕΨΗ ΠΩΛΗΣΕΩΝ | 12 |
| B.5 | ΒΑΘΙΑ ΜΑΘΗΣΗ (DEEP LEARNING) ΚΑΙ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ | 17 |
| B.6 | ΟΡΙΖΟΝΤΙΕΣ ΜΕΘΟΔΟΛΟΓΙΚΕΣ ΑΡΧΕΣ (ΓΙΑ ΣΤΑΤΙΣΤΙΚΕΣ, ML και DL) | 18 |
| C | Σχετιζόμενη Μελέτη | 22 |
| C.1 | ΣΧΕΤΙΚΗ ΕΡΕΥΝΑ ΣΤΗΝ ΠΡΟΒΛΕΨΗ ΠΩΛΗΣΕΩΝ | 22 |
| D | Πλαίσιο Ανάπτυξης και Εκτέλεσης | 27 |
| D.1 | ΠΕΡΙΒΑΛΛΟΝ ΑΝΑΠΤΥΞΗΣ ΚΑΙ ΕΚΤΕΛΕΣΗΣ | 27 |
| D.1.1 | Γλώσσα Προγραμματισμού και Διερμηνευτής | 27 |
| D.1.2 | Περιβάλλον Εργασίας και Αναπαραγωγιμότητα | 28 |
| D.1.3 | Python | 28 |
| D.1.4 | Jupyter | 30 |
| D.1.5 | Πίνακας Περιβάλλοντος | 30 |

| | |
|--|-----------|
| E Δεδομένα | 32 |
| E.1 Το DATASET και η ΠΗΓΗ τΟΥ ΜΟΝΤΕΛΟΥ | 32 |
| E.2 ΠΛΗΡΟΦΟΡΙΕΣ ΓΙΑ ΤΟ DATASET | 32 |
| E.3 ΣΤΑΤΙΣΤΙΚΑ ΤΟΥ DATASET | 34 |
| E.4 ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ | 34 |
| E.5 ΑΝΑΛΥΣΗ ΚΑΤΗΓΟΡΙΩΝ ΠΕΛΑΤΩΝ ΚΑΙ ΠΕΡΙΟΧΩΝ | 34 |
| F Προεπεξεργασία και Προετοιμασία Δεδομένων | 36 |
| F.1 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ | 36 |
| F.2 ΣΥΝΟΛΙΚΗ ΔΙΑΔΙΚΑΣΙΑ ΠΡΟΕΠΕΞΕΡΓΑΣΙΑΣ | 40 |
| G Εκπαίδευση του Μοντέλου | 41 |
| G.1 ΕΚΠΑΙΔΕΥΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ ΚΑΙ ΡΥΘΜΙΣΗ ΥΠΕΡΠΑΡΑΜΕΤΡΩΝ | 41 |
| H Αξιολόγηση και Αποτελέσματα του Μοντέλου | 44 |
| H.1 ΑΞΙΟΛΟΓΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ ΜΕ ΜΕΤΡΙΚΕΣ ΑΠΟΔΟΣΗΣ | 44 |
| H.1.1 1. Test Loss (Απώλεια Δοκιμής) | 44 |
| H.1.2 2. R ² Score (Σκορ R ²) | 44 |
| H.1.3 3. Συνολική Αξιολόγηση με Μετρικές Απόδοσης | 45 |
| H.2 ΑΝΑΛΥΣΗ ΤΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΤΟΥ ΜΟΝΤΕΛΟΥ ΠΡΟΒΛΕΨΗΣ ΠΩΛΗΣΕΩΝ | 46 |
| H.3 ΣΥΜΠΕΡΑΣΜΑΤΑ | 51 |
| Συντμήσεις | 55 |
| Γλωσσάρι Εξενικών Όρων | 58 |

Κατάλογος Σχημάτων

| | | |
|------|--|----|
| F.1 | Καθαρισμός Δεδομένων: Αφαίρεση μη χρησίμων στηλών από το dataset. | 37 |
| F.2 | Εξαγωγή Χρονικών Στοιχείων: Επεξεργασία των ημερομηνιών σε χαρακτηριστικά έτους, μήνα και ημέρας για καλύτερη χρήση στα μοντέλα πρόβλεψης. | 37 |
| F.3 | One-Hot Encoding: Διαδικασία εφαρμογής One-Hot Encoding στις κατηγορικές στήλες του dataset. | 38 |
| F.4 | Διαχείριση Αχραίων Τιμών: Λογαριθμική μετατροπή της στήλης <i>Sales</i> για τη διαχείριση της ασυμμετρίας. | 38 |
| F.5 | Διαχωρισμός των Δεδομένων: 70% για εκπαίδευση και 30% για δοκιμή. | 39 |
| F.6 | Κανονικοποίηση των τιμών σε μια κλίμακα με μέσο όρο 0 και τυπική απόκλιση 1 | 39 |
| G.7 | Εκπαίδευση του Μοντέλου και Ρύθμιση Learning Rate. | 42 |
| G.8 | Εκπαιδευτικά Δεδομένα X_{train} Η εικόνα δείχνει τα κανονικοποιημένα δεδομένα εκπαίδευσης. | 43 |
| G.9 | Εκπαιδευτικά Δεδομένα (X_{train}) Η εικόνα δείχνει τα κανονικοποιημένα δεδομένα εκπαίδευσης. | 43 |
| H.10 | Test Loss για την αξιολόγηση του μοντέλου. | 45 |
| H.11 | R^2 Score | 45 |
| H.12 | Actual vs Predicted Sales | 47 |
| H.13 | Βαθμονομημένη σύγκριση των προβλέψεων σε διάφορα bins | 48 |
| H.14 | Learning Curve: Δείχνει την καμπύλη εκπαίδευσης με τη σύγκριση της Training Loss και της Validation Loss | 48 |
| H.15 | Απόδοση του μοντέλου με βάση το R^2 Score | 49 |
| H.16 | Στατιστικά του μοντέλου, όπως το συνολικό πλήθος παραμέτρων και το μέγεθος του αρχείου του μοντέλου | 50 |

| | |
|---|----|
| H.17 Στατιστικά της στήλης "Month" για την ανάλυση των πωλήσεων σε σχέση με τους μήνες | 50 |
|---|----|

Κατάλογος Πινάκων

| | |
|---|----|
| <i>D.1 Πίνακας περιβάλλοντος ανάπτυξης, πακέτων και ρόλων τους στο pipeline πρόβλεψης πωλήσεων.</i> | 30 |
| <i>E.1 Περιγραφή των στηλών του dataset και βασικά στατιστικά για κάθε χαρακτηριστικό.</i> | 33 |
| <i>H.1 Στατιστικά της Στίλης “Month”</i> | 50 |
| <i>H.2 Συνοπτικές Μετρικές Απόδοσης του Μοντέλου Πρόβλεψης Πωλήσεων</i> | 52 |

Κεφάλαιο Α

Εισαγωγή

A.1 Εισαγωγή

Η παρούσα πτυχιακή εργασία επικεντρώνεται στην πρόβλεψη πωλήσεων με τη χρήση μεθόδων ανάλυσης χρονοσειρών και τεχνικών μηχανικής μάθησης. Η ακριβής εκτίμηση της μελλοντικής ζήτησης αποτελεί χρίσιμο ζήτημα για τις επιχειρήσεις, καθώς συνδέεται με τον καλύτερο σχεδιασμό, την αποδοτικότερη διαχείριση πόρων και την ενίσχυση της ανταγωνιστικότητας. Στο πλαίσιο αυτό, η εργασία επιχειρεί να συνδυάσει τη θεωρητική προσέγγιση της σχετικής βιβλιογραφίας με την πρακτική υλοποίηση ενός προτεινόμενου μοντέλου.

Η δομή της εργασίας έχει σχεδιαστεί ώστε να καθοδηγεί τον αναγνώστη από το θεωρητικό υπόβαθρο έως την εφαρμογή και την αξιολόγηση. Αρχικά, παρουσιάζεται μια βιβλιογραφική ανασκόπηση που καλύπτει τόσο τις παραδοσιακές στατιστικές μεθόδους όσο και τις σύγχρονες υπολογιστικές τεχνικές, αναδεικνύοντας τα πλεονεκτήματα και τους περιορισμούς τους. Στη συνέχεια, περιγράφεται η μεθοδολογία που ακολουθήθηκε για την ανάπτυξη του προτεινόμενου μοντέλου, καθώς και τα δεδομένα που αξιοποιήθηκαν για την εκπαίδευσή του. Ακολουθεί η παρουσίαση των αποτελεσμάτων, με έμφαση στις μετρικές αξιολόγησης και στη συζήτηση γύρω από την ακρίβεια και την αξιοπιστία των προβλέψεων. Τέλος, η εργασία ολοκληρώνεται με τα συμπεράσματα και τις προτάσεις για μελλοντική έρευνα.

Με τον τρόπο αυτό, η εργασία φιλοδοξεί να προσφέρει μια ολοκληρωμένη εικόνα για το πεδίο της πρόβλεψης πωλήσεων, συνδυάζοντας την επιστημονική γνώση με την πρακτική εφαρμογή και συμβάλλοντας στη βαθύτερη κατανόηση και αξιοποίηση των τεχνικών μηχανικής μάθησης

στο συγκεκριμένο αντικείμενο.

A.2 Ορισμός πρόβλεψης πωλήσεων

Η πρόβλεψη πωλήσεων αποτελεί έναν από τους ψεμελιώδεις άξονες της στρατηγικής επιχειρηματικής διοίκησης, καθώς επιτρέπει την αξιόπιστη εκτίμηση της μελλοντικής ζήτησης αγαθών ή υπηρεσιών. Η ακρίβεια στη διαμόρφωση προβλέψεων συμβάλλει καθοριστικά στη βελτιστοποίηση της διαχείρισης αποθεμάτων, στον αποδοτικό προγραμματισμό της παραγωγής και στην πρόβλεψη των εσόδων (Singh & Aggarwal, 2020). Παραδοσιακές στατιστικές μέθοδοι, όπως η γραμμική παλινδρόμηση, τα μοντέλα Holt-Winters και οι συναρτήσεις S-καμπύλης (Logistic, Gompertz, Bass), έχουν εφαρμοστεί ευρέως στην επιχειρησιακή πρόβλεψη. Ωστόσο, παρουσιάζουν περιορισμούς στην αναγνώριση μη γραμμικών και δυναμικά εξελισσόμενων σχέσεων μεταξύ μεταβλητών (Alkhalfah et al., 2023). Η σύγχρονη εξέλιξη της Τεχνητής Νοημοσύνης και ειδικότερα της Μηχανικής Μάθησης έχει καταστήσει εφικτή την ανάλυση πολύπλοκων και εκτεταμένων συνόλων δεδομένων. Τα τεχνητά νευρωνικά δίκτυα (ANNs), ιδίως τα πολυεπίπεδα perceptron (MLP), επιτρέπουν την αυτόνομη εκμάθηση τάσεων από ιστορικά δεδομένα, χωρίς την ανάγκη ρητών κανόνων ή υπόθεσης γραμμικότητας (Padmaja et al., 2023). Ερευνητικά δεδομένα αποδεικνύουν την υπεροχή μοντέλων βαθιάς μάθησης έναντι των στατιστικών μεθόδων, χυρίως σε σύνθετα επιχειρησιακά περιβάλλοντα. Συγκεκριμένα, το μοντέλο ACNN πέτυχε ακρίβεια πρόβλεψης 93,9% σε σενάριο λιανικών πωλήσεων υπεραγοράς (Priscila et al., 2023), ενώ το XGBoost παρουσίασε τη χαμηλότερη τιμή σφάλματος σε σύγκριση με το Random Forest και τη Γραμμική Παλινδρόμηση στην πρόβλεψη πωλήσεων της Walmart (Walmart ML Report, 2023).

Κεφάλαιο Β

Βιβλιογραφική ανασκόπηση

B.1 Εισαγωγή στη βιβλιογραφική ανασκόπηση

Η ενότητα της Βιβλιογραφικής Ανασκόπησης αποσκοπεί στην παρουσίαση και ανάλυση των σημαντικότερων ερευνητικών προσεγγίσεων που έχουν διαμορφώσει το πεδίο της πρόβλεψης χρονοσειρών και ειδικότερα της πρόβλεψης πωλήσεων. Αρχικά, εξετάζονται οι παραδοσιακές στατιστικές μέθοδοι, όπως τα μοντέλα αυτοπαλίνδρομης ολοκληρωμένης κινούμενου μέσου (ARIMA) και οι εποχικές παραλλαγές τους (SARIMA), τα οποία αποτέλεσαν για δεκαετίες τη βάση της ανάλυσης και ερμηνείας χρονοσειρών. Η παρουσίασή τους καθιστά σαφές το θεωρητικό υπόβαθρο και τα πλεονεκτήματα αλλά και τους περιορισμούς που χαρακτηρίζουν τις κλασικές στατιστικές προσεγγίσεις.

Στη συνέχεια, η ανάλυση στρέφεται στις σύγχρονες υπολογιστικές μεθοδολογίες που ανήκουν στο πεδίο της μηχανικής μάθησης (MM) και της βαθιάς μάθησης (BM). Εξετάζονται μοντέλα παλινδρόμησης, αλγόριθμοι ensemble όπως τα Random Forests και οι τεχνικές gradient boosting, καθώς και πιο εξειδικευμένες προσεγγίσεις με τη χρήση νευρωνικών δικτύων, αναδρομικών αρχιτεκτονικών (RNN, LSTM, GRU) και συνελικτικών δικτύων (CNN) που έχουν προσαρμοστεί στην ανάλυση χρονοσειρών. Παράλληλα, παρουσιάζονται οι βασικές βιβλιοθήκες λογισμικού και τα εργαλεία που έχουν καταστήσει εφικτή την ευρεία εφαρμογή των τεχνικών αυτών στην πράξη.

Τέλος, αναλύονται μελέτες περίπτωσης που αφορούν την πρόβλεψη πωλήσεων και άλλων επιχειρησιακών μεγεθών, ώστε να αποτυπωθεί η πρακτική διάσταση της θεωρίας και να καταδειχθεί το πώς οι ερευνητικές εξελίξεις έχουν εφαρμοστεί σε πραγματικά σενάρια. Με

αυτόν τον τρόπο, η ανασκόπηση παρέχει ένα συνεκτικό πλαίσιο για την κατανόηση των μεθοδολογικών επιλογών που αξιοποιούνται στην παρούσα εργασία.

B.2 Παραδοσιακές Στατιστικές Προσεγγίσεις στην Πρόβλεψη Πωλήσεων

B.2.1 Γενικό πλαίσιο παραδοσιακών μεθόδων

Οι παραδοσιακές στατιστικές μέθοδοι αποτέλεσαν για δεκαετίες τον ακρογωνιαίο λίθο της ποσοτικής πρόβλεψης σε διοίκηση επιχειρήσεων και οικονομετρία, κυρίως επειδή συνδυάζουν χαμηλό υπολογιστικό κόστος, θεωρητική αριμότητα και υψηλή ερμηνευσιμότητα. Στηρίζονται σε ιστορικά δεδομένα και στην παρατήρηση δομών όπως τάση, εποχικότητα και αυτοσυσχέτιση, παρέχοντας αξιόπιστες προβλέψεις σε σχετικά σταθερά περιβάλλοντα. Όταν όμως το σύστημα είναι έντονα δυναμικό/μη γραμμικό (συχνές προωθήσεις, απότομα γεγονότα, δομικές τομές), η προγνωστική τους ικανότητα τείνει να περιορίζεται. Παρ' όλα αυτά, παραμένουν πολύτιμα baselines, χρήσιμα για αρχική διάγνωση χρονοσειρών και ως συγχριτικό σημείο έναντι πιο σύνθετων τεχνικών (Hyndman Athanasopoulos, 2021; Box, Jenkins, Reinsel, Ljung, 2016).

- ARIMA (AutoRegressive Integrated Moving Average) Το ARIMA αποτελεί ένα από τα πιο εμβληματικά μοντέλα χρονοσειρών, χαθώς συνδυάζει τρεις βασικές συνιστώσες για την πρόβλεψη της μελλοντικής συμπεριφοράς μιας σειράς. Η πρώτη συνιστώσα είναι η αυτοπολινδρόμηση AR(p), κατά την οποία η τρέχουσα τιμή y_t εξαρτάται γραμμικά από p προηγούμενες τιμές σύμφωνα με τη σχέση

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t$$

(B.1)

όπου οι ϕ_i είναι οι παράμετροι του μοντέλου και ϵ_t ο λευκός ίόρυβος (Box et al., 2016). Η δεύτερη συνιστώσα αφορά την ολοκλήρωση ή διαφοροποίηση I(d), η οποία εφαρμόζεται για να καταστεί η σειρά στάσιμη. Η πρώτη διαφορά ορίζεται ως

$$\Delta y_t = y_t - y_{t-1} \quad (\text{B.2})$$

ενώ γενικά η διαφοροποίηση τάξης d μπορεί να εκφραστεί με τον lag operator LLL ως

$$\Delta^d y_t = (1 - L)^d y_t \quad (\text{B.3})$$

(Hyndman & Athanasopoulos, 2021). Τέλος, η τρίτη συνιστώσα αφορά το κινητό μέσο (MA), στο οποίο η τρέχουσα τιμή εξαρτάται από τα προηγούμενα σφάλματα πρόβλεψης, δηλαδή

$$y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (\text{B.4})$$

με τα θ_j να αντιστοιχούν στα weights των σφαλμάτων. Συνδυάζοντας τις τρεις συνιστώσες, το ARIMA(p, d, q) μοντέλο μπορεί να εκφραστεί συνολικά ως

$$\Delta^d y_t = \phi_1 \Delta^d y_{t-1} + \dots + \phi_p \Delta^d y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (\text{B.5})$$

Η διαδικασία Box–Jenkins περιλαμβάνει κύκλο αναγνώρισης μέσω ACF/PACF, επιλογής των παραμέτρων (p, d, q) , εκτίμησης του μοντέλου, διάγνωσης των καταλοίπων με τεστ όπως το Ljung–Box, καθώς και σύγκρισης χριτηρίων πληροφορίας όπως AIC, AICc και BIC (Box et al., 2016). Το ARIMA αποδίδει ιδιαίτερα καλά σε σταθερές χρονοσειρές με περιορισμένο θόρυβο, αλλά περιορίζεται όταν εμφανίζονται απότομες αλλαγές, μεταβαλλόμενη εποχικότητα ή έντονος θόρυβος, συνθήκες συχνές σε εμπορικά δεδομένα (Box et al., 2016; Hyndman & Athanasopoulos, 2021).

- SARIMA (Seasonal ARIMA) Όταν παρατηρείται εμφανής εποχικότητα σε μια χρονοσειρά, το ARIMA μπορεί να επεκταθεί στο SARIMA προσθέτοντας εποχικές παραμέτρους (P, D, Q, s) , όπου s αντιπροσωπεύει την περίοδο της εποχικότητας, όπως $s = 12$ για μηνιαία δεδομένα με ετήσια περιοδικότητα ή $s = 7$ για ημερήσια δεδομένα με εβδομαδιαίο κύκλο (Hyndman & Athanasopoulos, 2021). Το SARIMA μοντελοποιεί ταυτόχρονα μη εποχικές και εποχικές δομές, συνδυάζοντας επίπεδο, τάση και επαναλαμβανόμενα εποχικά πρότυπα, επιτρέποντας καλύτερη προσαρμογή σε σειρές με ισχυρή και σταθερή περιοδικότητα.

Μαθηματικά, το μοντέλο $\text{SARIMA}(p, d, q)(P, D, Q)_s$ μπορεί να εκφραστεί ως:

$$\Phi_P(L^s) \phi_p(L) \Delta^d \Delta_s^D y_t = \Theta_Q(L^s) \theta_q(L) \epsilon_t, \quad (\text{B.6})$$

όπου:

- $\phi_p(L)$ και $\theta_q(L)$ είναι οι πολυωνυμικές μορφές των μη εποχικών AR και MA αντίστοιχα.
- $\Phi_P(L^s)$ και $\Theta_Q(L^s)$ οι εποχικές AR και MA συνιστώσες.
- Δ^d η μη εποχική διαφοροποίηση.
- $\Delta_s^D = (1 - L^s)^D$ η εποχική διαφοροποίηση τάξης D .

Η σωστή επιλογή της εποχικής διαφοροποίησης D είναι κρίσιμη, καθώς η υπερβολική διαφοροποίηση μπορεί να προσθέσει θόρυβο, ενώ η ανεπαρκής διαφοροποίηση αφήνει τη σειρά μη στάσιμη (Hyndman & Athanasopoulos, 2021).

Με αυτόν τον τρόπο, το SARIMA επεκτείνει την ευελιξία του ARIMA, επιτρέποντας την ταυτόχρονη εκτίμηση μη εποχικών και εποχικών μοτίβων, κάτι ιδιαίτερα χρήσιμο σε εμπορικά δεδομένα με σαφή περιοδικότητα.

- ARIMAX / SARIMAX (με εξωγενείς μεταβλητές) Σε πολλές εφαρμογές πωλήσεων, η ζήτηση επηρεάζεται από εξωγενείς παράγοντες όπως η τιμή, οι προωθήσεις, οι αργίες ή ο καιρός. Τα μοντέλα ARIMAX και SARIMAX επεκτείνουν τα αντίστοιχα ARIMA/SARIMA ώστε να ενσωματώνουν εξωγενείς ερμηνευτικές μεταβλητές X_t ως πρόσθετους γραμμικούς όρους.

Συγκεκριμένα, το ARIMAX(p, d, q) με εξωγενείς μεταβλητές μπορεί να γραφεί ως:

$$\Delta^d y_t = \phi_1 \Delta^d y_{t-1} + \dots + \phi_p \Delta^d y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} + \beta^\top X_t, \quad (\text{B.7})$$

όπου β είναι το διάνυσμα των συντελεστών που αντιστοιχούν στις εξωγενείς μεταβλητές X_t , και ϵ_t το λευκό θόρυβο (Hyndman & Athanasopoulos, 2021).

Στην περίπτωση του SARIMAX(p, d, q)(P, D, Q)_s, οι εποχικές συνιστώσες προστίθενται όπως στο SARIMA, ενώ οι εξωγενείς παράγοντες ενσωματώνονται με τον ίδιο γραμμικό τρόπο:

$$\Phi_P(L^s) \phi_p(L) \Delta^d \Delta_s^D y_t = \Theta_Q(L^s) \theta_q(L) \epsilon_t + \beta^\top X_t. \quad (\text{B.8})$$

Με αυτόν τον τρόπο, το μοντέλο συνδύαζει τη “εσωτερική” δυναμική της χρονοσειράς με τα “εξωτερικά” σήματα, κάτι που συχνά βελτιώνει την πρόβλεψη και ενισχύει την επιχειρησιακή ερμηνευσιμότητα, καθώς επιτρέπει την ποσοτικοποίηση της επίδρασης παραμέτρων όπως τιμή ή προωθήσεις.

Προϋπόθεση για την ορθή λειτουργία είναι ο σωστός χρονισμός και συγχρονισμός των εξωγενών μεταβλητών μέσω lags ή leads, ώστε να αποτυπώνονται ρεαλιστικά οι αιτιώδεις καθυστερήσεις (Hyndman & Athanasopoulos, 2021).

- Γραμμική Παλινδρόμηση (Simple/Multiple Linear Regression) Η γραμμική παλινδρόμηση αποτυπώνει μια γραμμική σχέση μεταξύ των πωλήσεων, που αποτελούν την εξαρτημένη μεταβλητή y , και ενός συνόλου ανεξάρτητων μεταβλητών $X = (x_1, x_2, \dots, x_p)$, όπως η τιμή, οι προωθητικές ενέργειες ή οι δείκτες εποχικότητας.

Το μοντέλο γράφεται μαθηματικά ως:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \epsilon_i, \quad i = 1, \dots, n, \quad (\text{B.9})$$

όπου β_0 είναι το intercept, β_j οι συντελεστές των ανεξάρτητων μεταβλητών και ϵ_i τα σφάλματα, τα οποία υποτίθεται ότι είναι ανεξάρτητα, κανονικά κατανεμημένα με μέση τιμή μηδέν και σταθερή διακύμανση (Bishop, 2006; Hyndman & Athanasopoulos, 2021).

Η εκτίμηση των παραμέτρων γίνεται συνήθως με τη μέθοδο των Ελαχίστων Τετραγώνων (OLS), δηλαδή μέσω της ελαχιστοποίησης του ανθροίσματος των τετραγώνων των καταλοίπων:

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2. \quad (\text{B.10})$$

Η βασική δύναμη της γραμμικής παλινδρόμησης έγκειται στην ερμηνευσιμότητα, καθώς οι συντελεστές β_j παρέχουν άμεση «օριακή» επίδραση κάθε παράγοντα στις πωλήσεις. Παράλληλα, η μέθοδος λειτουργεί ως baseline για πιο σύνθετα μοντέλα και ως εργαλείο διάγνωσης για την ανάλυση των σχέσεων μεταξύ των μεταβλητών.

Ωστόσο, βασίζεται σε υποθέσεις όπως η γραμμικότητα, η ομοσκεδαστικότητα, η μη αυτοσυχέτιση των σφαλμάτων και η απουσία ισχυρής συνεξάρτησης, οι οποίες σπάνια ικανοποιούνται πλήρως σε πραγματικά δεδομένα λιανεμπορίου. Συνεπώς, η γραμμική παλινδρόμηση μπορεί να περιορίζεται όταν οι σχέσεις είναι έντονα μη γραμμικές ή υπάρχουν πολύπλοκες αλληλεπιδράσεις (Bishop, 2006; Hyndman & Athanasopoulos, 2021).

- Εκθετική Εξομάλυνση Holt–Winters (Simple/Holt/Holt–Winters) Η εκθετική εξομάλυνση σταθμίζει περισσότερο τις πιο πρόσφατες παρατηρήσεις.
 - **Simple Exponential Smoothing:** για σειρές χωρίς σαφή τάση ή εποχικότητα.
 - **Holt’s Linear:** προσθέτει ρητή τάση (level + trend).
 - **Holt–Winters:** προσθέτει και εποχικότητα (προσθετική ή πολλαπλασιαστική), κατάλληλο όταν η εποχική μορφή είναι σχετικά σταθερή.

Πλεονεκτεί σε ταχύτητα και απλότητα και συχνά δίνει ισχυρά baselines σε πρακτικά σενάρια. Δεν ενσωματώνει εγγενές εξωγενείς μεταβλητές και μπορεί να υποχωρεί όταν

οι εποχικές δομές μεταβάλλονται ή όταν η αγορά οδηγείται από εξωτερικά σήματα (Holt, 1957; Winters, 1960; Hyndman & Athanasopoulos, 2021).

- Θεμελιώδεις έννοιες: Στασιμότητα, Τάση, Εποχικότητα **Στασιμότητα:** οι στατιστικές ιδιότητες (μέσος, διακύμανση, αυτοσυσχέτιση) δεν αλλάζουν με τον χρόνο.

Τάση: αργή συστηματική μεταβολή του επιπέδου (αύξουσα ή φθίνουσα).

Εποχικότητα: περιοδική δομή που επαναλαμβάνεται ανά σταθερή περίοδο (π.χ. 12 μήνες, 7 ημέρες).

Η επίτευξη στασιμότητας (μέσω διαφοροποίησης ή μετασχηματισμών) είναι κλειδί για ARIMA/SARIMA, ενώ η αξιόπιστη αναγνώριση εποχικότητας καθορίζει την επιλογή εποχικών όρων (Box et al., 2016).

- Διαφοροποίηση (Differencing) – επιλογή d/D Η διαφοροποίηση αφαιρεί την τάση/μη στασιμότητα. Επιλέγουμε το ελάχιστο αναγκαίο d (και εποχικό D) ώστε να μην εισάγουμε περιττό θόρυβο. Υπερδιαφοροποίηση οδηγεί σε «υπερ-θορυβώδη» κατάλοιπα και φτωχή πρόβλεψη, ενώ υποδιαφοροποίηση διατηρεί αυτοσυσχέτιση στα κατάλοιπα. Η διαφοροποίηση μπορεί να συνδυαστεί με μετασχηματισμούς (π.χ. log) όταν η διακύμανση αυξάνεται με το επίπεδο (Hyndman & Athanasopoulos, 2021).
- Επιλογή παραμέτρων (p, d, q) και (P, D, Q, s) Η αναγνώριση βασίζεται σε γραφή ματα ACF/PACF και σε αυτοματοποιημένους αλγόριθμους (π.χ. auto.arima) που ελαχιστοποιούν AIC/AICc/BIC. Ακολουθεί εκτίμηση και διαγνωστικός έλεγχος υπολοίπων, ώστε να προσεγγίζουν λευκό θόρυβο. Η διαδικασία είναι επαναληπτική: αν τα κατάλοιπα εμφανίζουν δομή, αναθεωρούμε παραμέτρους ή μοντέλο (Box et al., 2016).
- Διαγνωστικοί έλεγχοι τεστ στασιμότητας (Ljung–Box, ADF, KPSS) Μετά την εκτίμηση, ελέγχουμε ότι τα κατάλοιπα είναι μη συσχετισμένα (**Ljung–Box**) και οπτικά «λευκός θόρυβος» (ACF/PACF καταλοίπων).

Για στασιμότητα χρησιμοποιούμε:

- **ADF:** ελέγχει ύπαρξη μοναδιαίας ρίζας.
- **KPSS:** ελέγχει την αντίθετη υπόθεση — στασιμότητα γύρω από τάση/μέσο.

Ο συνδυασμός τους μειώνει τον κίνδυνο λανθασμένων συμπερασμάτων (Hyndman & Athanasopoulos, 2021).

- Μέτρα σφάλματος (MAE, RMSE, MAPE, sMAPE) **MAE (Mean Absolute Error)**: ανθεκτικό σε outliers και εκφράζεται σε «μονάδες» πωλήσεων.

RMSE (Root Mean Squared Error): τιμωρεί δυνατά τα μεγάλα λάθη, χρήσιμο όταν οι μεγάλες αστοχίες είναι επιχειρησιακά κρίσιμες.

MAPE (Mean Absolute Percentage Error): δίνει ποσοστό σφάλματος αλλά γίνεται ασταθές με μικρές ή μηδενικές πραγματικές τιμές; το **sMAPE** προσφέρει συμμετρία.

Στην πράξη, χρησιμοποιούνται συνδυαστικά και συχνά πλαισιώνονται από επιχειρησιακούς δείκτες (service level, out-of-stock rate) (Hyndman & Athanasopoulos, 2021).

- **Χρονικά συνεπής επικύρωση (Rolling / Walk-Forward Validation)** Σε χρονοσειρές αποφεύγουμε την τυχαία διαιρεση τύπου k-fold για αξιολόγηση μοντέλων, καθώς η τυχαιότητα παραβιάζει τη χρονική σειρά των δεδομένων. Αντίθετα, χρησιμοποιούμε προσεγγίσεις τύπου *rolling origin* ή *walk-forward*: εκπαιδεύουμε το μοντέλο μέχρι τον χρόνο t , προβλέπουμε για τα βήματα $t + 1$ έως $t + h$, αξιολογούμε την ακρίβεια των προβλέψεων και στη συνέχεια «κυλάμε» το cutoff προς τα εμπρός για να επαναλάβουμε τη διαδικασία. Με αυτόν τον τρόπο προσομοιώνεται ρεαλιστικά η μελλοντική χρήση του μοντέλου και αποφεύγονται αισιόδοξες εκτιμήσεις απόδοσης που θα προέκυπταν από τυχαία διαιρεση (Bergmeir & Benítez, 2012; Hyndman & Athanasopoulos, 2021).
- **Ακραίες τιμές, ελλείψεις και προσθετική vs πολλαπλασιαστική εποχικότητα Outliers** (π.χ. αιχμές από προωθήσεις, stock-outs ή ασυνέχειες καταγραφής) πρέπει να επισημαίνονται και να αντιμετωπίζονται, για παράδειγμα μέσω σήμανσης γεγονότων ή συντηρητικού imputation. Σε σειρές όπου η διακύμανση αυξάνεται με το επίπεδο, προτιμώνται πολλαπλασιαστικά μοντέλα ή λογαριθμικοί μετασχηματισμοί, ενώ όταν η διακύμανση παραμένει περίπου σταθερή, τα προσθετικά μοντέλα ταιριάζουν καλύτερα (Hyndman & Athanasopoulos, 2021).
- **Συνοπτικά πλεονεκτήματα και περιορισμοί**

Τα πλεονεκτήματα περιλαμβάνουν ταχύτητα, θεωρητική κατοχύρωση, ευκολία εφαρμογής και διάγνωσης, ερμηνευσιμότητα, καθώς και ιδανική χρήση ως baseline, ειδικά για σταθερά εποχικά μοτίβα.

Οι περιορισμοί αφορούν τη δυσκολία αντιμετώπισης έντονων μη γραμμικοτήτων, μεταβαλλόμενων δομών, πλούσιων εξωγενών αλληλεπιδράσεων και προβλέψεων υψηλής granular (π.χ. SKU × Κατάστημα × Ημέρα). Σε τέτοιες περιπτώσεις απαιτείται εμπλουτισμός μέσω ARIMAX/SARIMAX ή/και μετάβαση σε σύγχρονες μεθόδους.

B.3 Ανάγκη για σύγχρονες μεθόδους

Η συνεχής αύξηση της πολυπλοκότητας στα δεδομένα πωλήσεων, σε συνδυασμό με τον αυξανόμενο όγκο και τη δυναμική φύση της πληροφορίας, καθιστά τις παραδοσιακές στατιστικές μεθόδους ανεπαρκείς για την κάλυψη των σύγχρονων επιχειρησιακών αναγκών. Στο σημερινό ανταγωνιστικό περιβάλλον, οι οργανισμοί απαιτούν προγνωστικά μοντέλα που να μπορούν να ανιχνεύουν μη γραμμικές σχέσεις, να προσαρμόζονται σε νέες μεταβλητές και να παρέχουν προβλέψεις σε πραγματικό ή σχεδόν πραγματικό χρόνο (Alkhalifah et al., 2023). Η εξέλιξη της Τεχνητής Νοημοσύνης (AI), και ειδικότερα των μεθόδων Μηχανικής Μάθησης (Machine Learning) και Βαθιάς Μάθησης (Deep Learning), προσφέρει λύσεις με σημαντικά πλεονεκτήματα: αυτονομία μάθησης, ευελιξία προσαρμογής σε πολύπλοκα δεδομένα και ικανότητα ενσωμάτωσης εξωτερικών παραγόντων (Priscila et al., 2023,. Τα τεχνητά νευρωνικά δίκτυα (ANN), για παράδειγμα, επιτρέπουν την αυτόματη ανίχνευση και αφομοίωση προτύπων, χωρίς να απαιτούν αυστηρές παραδοχές στατιστικής φύσης. Σε πρόσφατες εφαρμογές, σύγχρονα νευρωνικά μοντέλα όπως τα LSTM και ACNN έχουν επιδείξει εξαιρετικά αποτελέσματα στην πρόβλεψη πωλήσεων, ακόμα και σε datasets με θόρυβο, εποχικότητα ή ακανόνιστη συμπεριφορά καταναλωτών (Walmart ML Report, 2023, σ. 91· Chan et al., 2025). Η χρήση αυτών των μεθόδων έχει αποδειχθεί ιδιαιτέρως χρήσιμη και στον τομέα της ενεργειακής πρόβλεψης, όπου τα μοντέλα αυτά ανταποχρίνονται καλύτερα σε απότομες αυξομειώσεις της ζήτησης, σε σχέση με τα παραδοσιακά στατιστικά μοντέλα (Papitha et al., 2024). Η ανάγκη για τέτοια ευέλικτα, προσαρμόσιμα και «ευφυή» μοντέλα πρόβλεψης είναι πλέον επιτακτική σε όλους τους τομείς της αγοράς — από τη βιομηχανία και το λιανεμπόριο έως την ενέργεια και τις χρηματοπιστωτικές υπηρεσίες. Συνοψίζοντας, καθίσταται σαφές ότι η αδυναμία των συμβατικών προσεγγίσεων να ανταποχριθούν στις απαιτήσεις ενός διαρκώς μεταβαλλόμενου επιχειρησιακού περιβάλλοντος έχει δημιουργήσει ένα ευνοϊκό πεδίο για την ενσωμάτωση προηγμένων μεθόδων τεχνητής νοημοσύνης. Η παρούσα εργασία εστιάζει στην αξιοποίηση των τεχνητών νευρωνικών δικτύων, τα οποία προσφέρουν δυνατότητες προσαρμογής, μάθησης και πρόβλεψης με ακρίβεια και ταχύτητα. Πριν την παρουσίαση της προτεινόμενης μεθοδολογίας,

χρίνεται αναγκαία μια συστηματική ανασκόπηση της υπάρχουσας βιβλιογραφίας, προκειμένου να αποσαφηνιστούν τα ερευνητικά πεδία, οι εφαρμογές και τα αποτελέσματα προηγούμενων μελετών.

B.4 Μηχανική Μάθηση (Machine Learning) για Πρόβλεψη Πωλήσεων

Random Forest (Σύνολο Δέντρων Απόφασης)

To Random Forest αποτελεί μία από τις σημαντικότερες μεθόδους συνόλων (ensemble methods), η οποία εισήχθη από τον Breiman (2001). Βασίζεται στον συνδυασμό μεγάλου αριθμού δέντρων απόφασης, τα οποία εκπαιδεύονται σε τυχαία δείγματα των δεδομένων και σε τυχαία υποσύνολα χαρακτηριστικών. Η διαδικασία αυτή στηρίζεται σε δύο θεμελιώδεις τεχνικές: (α) *bagging* (bootstrap aggregating), δηλαδή την εκπαίδευση κάθε δέντρου σε ένα δείγμα με αντικατάσταση από το αρχικό σύνολο δεδομένων, και (β) *feature subsampling*, δηλαδή την τυχαία επιλογή υποσυνόλων χαρακτηριστικών σε κάθε κόμβο του δέντρου για την αναζήτηση του βέλτιστου διαχωρισμού.

Έστω ότι έχουμε ένα σύνολο εκπαίδευσης

$$\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\},$$

(B.11)

όπου $x_i \in \mathbb{R}^p$ είναι τα διανύσματα χαρακτηριστικών και y_i οι αντίστοιχες ετικέτες. To Random Forest κατασκευάζει B δέντρα απόφασης $T_b(x)$, $b = 1, 2, \dots, B$, το καθένα εκπαιδευμένο σε ένα bootstrap δείγμα \mathcal{D}_b και σε τυχαία υποσύνολα χαρακτηριστικών.

Η τελική πρόβλεψη δίνεται από τον μέσο όρο για παλινδρόμηση:

$$\hat{f}(x) = \frac{1}{B} \sum_{b=1}^B T_b(x),$$

(B.12)

ή από την πλειοψηφία ψήφων για ταξινόμηση:

$$\hat{C}(x) = \arg \max_k \sum_{b=1}^B I\{T_b(x) = k\}, \quad (\text{B.13})$$

όπου $I\{\cdot\}$ είναι η συνάρτηση ένδειξης.

Η τεχνική αυτή μειώνει σημαντικά το overfitting που χαρακτηρίζει τα μεμονωμένα δέντρα, αυξάνοντας τη γενίκευση του μοντέλου. To Random Forest παρουσιάζει ιδιαίτερη ανθεκτικότητα σε υθροβάδη δεδομένα, μπορεί να χειριστεί μείγματα κατηγοριακών και αριθμητικών χαρακτηριστικών, είναι ανεκτικό στην ύπαρξη ελλιπών δεδομένων και παρέχει μέτρα σημαντικότητας χαρακτηριστικών (feature importance), τα οποία βελτιώνουν την ερμηνευσιμότητα και τη χρησιμότητα του μοντέλου σε επιχειρησιακό επίπεδο (Breiman, 2001; Hastie, Tibshirani, & Friedman, 2009). Ωστόσο, μειονεκτεί σε περιπτώσεις όπου απαιτείται λεπτομερής μοντελοποίηση σύνθετων αλληλεπιδράσεων ή όταν ο στόχος είναι η μέγιστη δυνατή ακρίβεια σε δύσκολα tabular datasets, καθώς σε τέτοια πλαίσια άλλες τεχνικές συνόλων όπως το Gradient Boosting συχνά υπερτερούν (Breiman, 2001· Prokhorenkova et al., 2018).

Support Vector Regression (SVR)

To Support Vector Regression (SVR) αποτελεί τη γενίκευση των Support Vector Machines (SVM) στο πεδίο της παλινδρόμησης, όπως αρχικά διατυπώθηκε από τους Drucker, Burges, Kaufman, Smola και Vapnik (1997). Η κεντρική ιδέα του SVR είναι η αναζήτηση μιας συνάρτησης:

$$f(x) = w^\top \phi(x) + b,$$

(B.14)

όπου $\phi(x)$ είναι ένας μη γραμμικός μετασχηματισμός σε χώρο χαρακτηριστικών υψηλής διάστασης, τέτοιος ώστε η συνάρτηση να βρίσκεται όσο το δυνατόν πιο «επίπεδα» (flat), δηλαδή με ελεγχόμενη πολυπλοκότητα (Vapnik, 1995).

Η καινοτομία του SVR εισάγεται μέσω της έννοιας του ε -insensitive tube: τα σφάλματα που βρίσκονται εντός του σωλήνα ανοχής ε δεν τιμωρούνται, ενώ αποκλίσεις μεγαλύτερες από αυτό το όριο υφίστανται ποινή. Η βελτιστοποίηση διατυπώνεται ως εξής:

$$\min_{w,b,\xi,\xi^*} \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*),$$

(B.15)

υπό τους περιορισμούς:

$$\{ y_i - w^\top \phi(x_i) - b \leq \varepsilon + \xi_i, w^\top \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^*, \xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, n.$$

Οι μεταβλητές ξ_i, ξ_i^* απορροφούν τις αποκλίσεις εκτός του σωλήνα και η παράμετρος C ελέγχει το trade-off μεταξύ πολυπλοκότητας του μοντέλου και αποδεκτών σφαλμάτων (Smola & Schölkopf, 2004).

Με την εισαγωγή της δυϊκής μορφής και του *kernel trick*, η τελική συνάρτηση πρόβλεψης λαμβάνει τη μορφή:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b,$$

(B.16)

όπου $K(x_i, x_j) = \phi(x_i)^\top \phi(x_j)$ είναι η kernel συνάρτηση (συχνά RBF, πολυωνυμική ή γραμμική). Η προσέγγιση αυτή επιτρέπει στο SVR να αποτυπώσει πολύπλοκες μη γραμμικές σχέσεις (Smola & Schölkopf, 2004).

Στην πράξη, το SVR είναι ιδιαίτερα ισχυρό σε μικρά έως μεσαία σύνολα δεδομένων με καθαρές δομές, αλλά παρουσιάζει προκλήσεις στην κλιμάκωση σε πολύ μεγάλα σύνολα, καθώς η πολυπλοκότητα του αλγορίθμου εξαρτάται από το πλήθος των δειγμάτων εκπαίδευσης. Επιπλέον, η απόδοσή του απαιτεί προσεκτική ρύθμιση των υπερπαραμέτρων C , ε και γ (όταν χρησιμοποιείται RBF kernel). Η ερμηνευσιμότητα παραμένει μέτρια σε σύγχριση με γραμμικά μοντέλα, αλλά η στιβαρή θεωρητική του βάση και η ικανότητά του να χειρίζεται ακραίες τιμές καθιστούν το SVR πολύτιμο εργαλείο «μεσαίας κλίμακας» (Drucker, Burges, Kaufman, Smola, & Vapnik, 1997).

Gradient Boosting / XGBoost

To Gradient Boosting εκπαιδεύει διαδοχικά «αδύναμα» δέντρα, καθένα από τα οποία στοχεύει να διορθώσει τα σφάλματα των προηγούμενων μοντέλων. Θεωρητικά, το μοντέλο δημιουργεί μια συνάρτηση πρόβλεψης ως άθροισμα M δέντρων:

$$F_M(x) = \sum_{m=1}^M \gamma_m h_m(x), \quad (\text{B.17})$$

όπου $h_m(x)$ είναι το δέντρο στο βήμα m και γ_m ο συντελεστής step size που προκύπτει από την ελαχιστοποίηση της συνάρτησης απώλειας $L(y, F(x))$.

Η διαδικασία επαναληπτικής εκπαίδευσης εκφράζεται μαθηματικά ως:

$$r_{im} = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)}, \quad i = 1, \dots, n, \quad (\text{B.18})$$

όπου τα r_{im} αποτελούν τα residuals (δηλαδή τα gradients της απώλειας) που προσπαθεί να προσεγγίσει το νέο δέντρο $h_m(x)$. Στη συνέχεια, η ενημέρωση του μοντέλου γίνεται με:

$$F_m(x) = F_{m-1}(x) + \nu \gamma_m h_m(x), \quad (\text{B.19})$$

όπου $\nu \in (0, 1]$ είναι η learning rate, που ελέγχει την ταχύτητα μάθησης και αποτρέπει υπερβολική προσαρμογή στα residuals (Friedman, 2001).

Η υλοποίηση XGBoost προσθέτει επιπλέον δυνατότητες κανονικοποίησης της πολυπλοκότητας κάθε δέντρου:

$$\Omega(h) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2,$$

(B.20)

όπου T ο αριθμός των φύλλων, w_j τα weights των φύλλων, και γ, λ οι παράμετροι τωχτικοποίησης L1/L2.

Επιπλέον χρησιμοποιεί δεύτερης τάξης Taylor expansion για πιο ακριβή υπολογισμό των step sizes γ_m , αποδοτική διαχείριση αραιών ή ελλιπών τιμών, παραλληλοποίηση και πλούσια επιλογή υπερπαραμέτρων για βελτιστοποίηση της απόδοσης (Chen & Guestrin, 2016).

Σε προβλήματα πωλήσεων με πλούσια χαρακτηριστικά (π.χ. τιμή, προωθητικές ενέργειες, ημερολογιακά, καιρικά δεδομένα, κατηγορία προϊόντος), το XGBoost συχνά προσφέρει κορυφαίες επιδόσεις, ειδικά όταν συνοδεύεται από σχολαστική προεπεξεργασία και time-aware validation. Το trade-off είναι η αυξημένη πολυπλοκότητα ρύθμισης υπερπαραμέτρων και η περιορισμένη ερμηνευσιμότητα, που απαιτεί post-hoc εργαλεία όπως feature importance ή SHAP values (Chen & Guestrin, 2016; Friedman, 2001).

Προεπεξεργασία & Feature Engineering (για ML)

Η απόδοση των ML μοντέλων εξαρτάται καθοριστικά από την ποιότητα των χαρακτηριστικών: ημερολογιακές μεταβλητές (μήνας, εβδομάδα, ημέρα εβδομάδας, αργίες), κυλιόμενα στατιστικά (rolling mean/median/std), υστερήσεις (lags) για να «περάσει» η χρονική εξάρτηση, target/impact encoding για κατηγορικές υψηλής καρδινάλιοτητας (SKU, κατάστημα), και ενίστε λογαριθμικοί ή Box–Cox μετασχηματισμοί για σταθεροποίηση διακύμανσης. Η συνεπής κανονικοποίηση/τυποποίηση βελτιώνει τη σταθερότητα εκπαίδευσης, ιδίως για SVR και νευρωνικά δίκτυα.

Χρονικά Συνεπής Επικύρωση (Time-Series CV)

Τα ML μοντέλα σε χρονοσειρές πρέπει να αξιολογούνται με rolling ή walk-forward validation, όχι με τυχαία k-fold splits. Τα παράθυρα εκπαίδευσης «κυλούν» προς το μέλλον, τα σετ αξιολόγησης αντιπροσωπεύουν πραγματικές μελλοντικές περιόδους, και οι μετρικές υπολογίζονται ανά ορίζοντα ($h = 1, h = 7, h = 28$) για πλήρη εικόνα απόδοσης (Bergmeir & Benítez, 2012; Hyndman & Athanasopoulos, 2021).

B.5 Βαθιά Μάθηση (Deep Learning) και Νευρωνικά Δίκτυα

MLP (Multilayer Perceptron)

Το MLP είναι δίκτυο προώθησης με διαδοχικά πλήρως συνδεδεμένα στρώματα που «μαθαίνει» μη γραμμικές σχέσεις μέσω συναρτήσεων ενεργοποίησης (ReLU/LeakyReLU/GELU) και εκπαιδεύεται με οπισθοδιάδοση και βελτιστοποίηση (Adam). Στην πρόβλεψη πωλήσεων, λειτουργεί άριστα σε «ταμπελοποιημένα» δεδομένα όταν προηγηθεί σωστή μηχανική χαρακτηριστικών (lags, rolling, ημερολόγιο, promos) και εφαρμοστούν τακτικοποίηση (L2, dropout), κανονικοποίηση inputs και early stopping. Μπορεί να ενσωματώσει embeddings για κατηγορίες (SKU, κατάστημα), μετατρέποντας υψηλή καρδινάλιοτητα σε συμπαγείς αναπαραστάσεις και ενισχύοντας την «global» μάθηση σε πολλές σειρές (Goodfellow, Bengio, Courville, 2016; Kingma Ba, 2015; Srivastava et al., 2014).

LSTM / GRU (Αναδρομικά Νευρωνικά με Μνήμη)

Τα LSTM εισάγουν «πύλες» (input/forget/output) που επιτρέπουν ροή/απόρριψη πληροφορίας στο χρόνο, συλλαμβάνοντας μακροχρόνιες εξαρτήσεις—ιδιαίτερο όταν η σημερινή ζήτηση επηρεάζεται από επιδράσεις εβδομάδων/μηνών (εποχικά μοτίβα, linger effect μετά από προωθητική). Τα GRU απλοποιούν τη δομή με συγκρίσιμες επιδόσεις. Προσφέρουν ισχύ σε καθαρές σειρές, ενώ συχνά ωφελούνται από συνδυασμό με εξωγενείς «διαδρομές» (π.χ. concatenation exogenous features) ώστε να αξιοποιούν πλήρως τιμή/προωθήσεις/ημερολόγιο (Hochreiter Schmidhuber, 1997).

CNN 1D σε Χρονοσειρές

Τα συνελικτικά δίκτυα με 1D πυρήνες ανιχνεύουν τοπικά χρονικά μοτίβα (π.χ. εβδομαδιαίους κύκλους, αιχμές promos). Με κατάλληλα kernel sizes και dilations, ένα CNN «βλέπει» παραθύρια πολλαπλών κλιμάκων. Σε πωλήσεις, τα CNNs ξεχωρίζουν όταν τα σήματα είναι «αιχμηρά» και εντοπίζονται σε γειτονιές χρόνου (π.χ. πριν/κατά τη διάρκεια μιας καμπάνιας). Συχνά συνδυάζονται με dense layers για τελική παλινδρόμηση.

Attention / Self-Attention (έννοια εστίασης)

Οι μηχανισμοί attention επιτρέπουν στο δίκτυο να «σταθμίζει» δυναμικά ποια χρονικά βήματα ή ποια χαρακτηριστικά είναι σημαντικότερα για την παρούσα πρόβλεψη. Σε δεδομένα πωλήσεων, αυτό μεταφράζεται σε αυξημένη ικανότητα ανάδειξης «χρίσματων εβδομάδων» (π.χ. Black Friday, εορτές) ή ισχυρών αλληλεπιδράσεων (τιμή×promo×ημέρα). Η προσθήκη attention πάνω από LSTM/CNN ή σε καθαρό self-attention πλαίσιο ενισχύει την εξήγηση και

συχνά την ακρίβεια.

Global vs Local μοντέλα

Local: ένα μοντέλο ανά σειρά (SKU×Κατάστημα). Global: ένα ενιαίο μοντέλο που μαθαίνει από όλες τις σειρές μαζί (με embeddings/ID). Σε περιβάλλοντα με πολλές βραχείες/θορυβώδεις σειρές, τα global δίκτυα συχνά υπερέχουν επειδή «δανείζονται» πληροφορία από ομοειδείς σειρές και γενικεύουν μοτίβα (ιδανική χρήση για MLP/LSTM/CNN με embeddings).

Τακτικοποίηση/Σταθεροποίηση στην Εκπαίδευση

Dropout, L2 (weight decay), early stopping, learning rate schedules (cosine, plateau), batch normalization/layer normalization όπου αρμόζει: όλα στοχεύουν στον περιορισμό του overfitting και στη σταθερότητα της εκπαίδευσης. Για πωλήσεις με πολλές εξωγενείς μεταβλητές, το συνδυαστικό τακτικοποιημένο MLP είναι συχνά «γλυκιά» ισορροπία ισχύος/χόστους.

B.6 Οριζόντιες Μεθοδολογικές Αρχές (για Στατιστικές, ML και DL)

Feature Engineering για Πωλήσεις

Η ποιότητα των χαρακτηριστικών καθορίζει το ταβάνι επίδοσης:

- Ημερολόγιο (ημέρα εβδομάδας, μήνας, αργίες, εποχικά dummies)
- Τιμή, εκπτώσεις, promos (τύπος προώθησης, ένταση, διάρκεια, lead/lag effects)
- Κυλιόμενα παράθυρα (rolling mean/median/std/min/max) και υστερήσεις (lag 1, 7, 28 κ.λπ.)
- Συναθροίσεις ανά επίπεδο (SKU→Κατηγορία→Κατάστημα→Περιοχή) για «ομαλή»
- Καιρός/εξωτερικοί δείκτες όπου σχετικό

Η συνεπής δημιουργία και τεκμηρίωση των features είναι επένδυση που αποδίδει περισσότερο από το «κυνηγητό» υπερπαραμέτρων.

Επικύρωση Χρονοσειρών (Rolling / Walk-Forward)

Η αξιολόγηση πρέπει να σέβεται τον χρόνο: train έως t , predict $t + 1..t + h$, καταγραφή σφαλμάτων, κύλιση cutoff και επανάληψη. Παράγονται καμπύλες επίδοσης ανά ορίζοντα

(short/medium/long) και ανά ιεραρχικό επίπεδο (SKU, κατηγορία, σύνολο), ώστε να εντοπίζονται συστηματικές αδυναμίες (Bergmeir Benítez, 2012; Hyndman Athanasopoulos, 2021).

Mean Absolute Error (MAE)

Το Μέσο Απόλυτο Σφάλμα (MAE) αποτελεί μία από τις πιο απλές αλλά και ευρέως χρησιμοποιούμενες μετρικές για την αξιολόγηση της ακρίβειας προβλέψεων. Ορίζεται μαθηματικά ως

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

(B.21)

όπου y_t είναι η πραγματική τιμή, \hat{y}_t η πρόβλεψη και n το πλήθος των παρατηρήσεων. Το MAE εκφράζει τη μέση απόλυτη απόκλιση σε ίδιες μονάδες με τα δεδομένα (π.χ. τεμάχια, ευρώ), γεγονός που διευκολύνει την ερμηνεία του (Chatfield, 2000). Η βασική του αρετή είναι ότι είναι σχετικά ανθεκτικό στις ακραίες τιμές, καθώς όλα τα σφάλματα συνεισφέρουν γραμμικά στον μέσο όρο. Ωστόσο, μειονεκτεί στο ότι δεν διαφοροποιεί τις μεγάλες αποκλίσεις από τις μικρές, κάτι που μπορεί να είναι πρόβλημα σε επιχειρησιακά περιβάλλοντα όπου οι μεγάλες αστοχίες έχουν δυσανάλογο κόστος (Makridakis, Wheelwright, Hyndman, 1998).

Root Mean Squared Error (RMSE)

Η Ριζική Μέση Τετραγωνική Απόκλιση (RMSE) ορίζεται ως

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

(B.22)

Σε αντίθεση με το MAE, η RMSE υψώνει στο τετράγωνο τις αποκλίσεις, δίνοντας πολύ μεγαλύτερη βαρύτητα σε μεγάλα σφάλματα. Αυτό σημαίνει ότι λίγες σοβαρές αποκλίσεις μπορούν να επηρεάσουν δυσανάλογα το τελικό αποτέλεσμα. Για τον λόγο αυτόν, η RMSE θεωρείται πιο κατάλληλη όταν οι μεγάλες αστοχίες έχουν σημαντικό επιχειρησιακό αντίκτυπο, όπως στην πρόβλεψη ζήτησης αποθεμάτων όπου ένα μεγάλο stock-out μπορεί να είναι πολύ

πιο δαπανηρό από πολλά μικρά (Armstrong Collopy, 1992). Ωστόσο, η ευαισθησία της σε outliers είναι και το βασικό της μειονέκτημα, καθώς ενδέχεται να παρουσιάσει υπερβολικά υψηλές τιμές όταν υπάρχουν λίγες ακραίες παρατηρήσεις (Flores, 1986).

Mean Absolute Percentage Error (MAPE)

Το Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (MAPE) εκφράζει το σφάλμα σε ποσοστιαία μορφή και ορίζεται ως

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|.$$

(B.23)

Η μετρική αυτή είναι ιδιαίτερα δημοφιλής, κυρίως λόγω της ευκολίας με την οποία μπορεί να ερμηνευθεί από μη τεχνικό κοινό και να χρησιμοποιηθεί για συγκρίσεις μεταξύ διαφορετικών προϊόντων ή αγορών (Makridakis, Spiliotis, Assimakopoulos, 2018). Ωστόσο, το MAPE παρουσιάζει σοβαρά προβλήματα όταν οι πραγματικές τιμές είναι μικρές ή μηδενικές, καθώς ο παρονομαστής μπορεί να οδηγήσει σε υπερβολικά μεγάλες ή ακόμη και άπειρες τιμές (Hyndman Koehler, 2006). Επιπλέον, έχει την τάση να υπερτονίζει τα σφάλματα σε περιπτώσεις χαμηλής ζήτησης, γεγονός που το καθιστά λιγότερο αξιόπιστο σε περιβάλλοντα με πολλά προϊόντα μικρού όγκου.

Symmetric Mean Absolute Percentage Error (sMAPE)

Για να ξεπεραστούν τα προβλήματα του MAPE, έχει προταθεί η συμμετρική εκδοχή του, γνωστή ως Symmetric Mean Absolute Percentage Error (sMAPE). Ορίζεται ως

$$sMAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{\frac{|y_t| + |\hat{y}_t|}{2}}.$$

(B.24)

Σε αυτήν την περίπτωση, ο παρονομαστής αποτελείται από τον μέσο όρο της απόλυτης πραγματικής και της προβλεπόμενης τιμής, μειώνοντας έτσι την υπερβολική διόγκωση του σφάλματος σε χαμηλές τιμές (Hyndman Koehler, 2006). Το sMAPE είναι επομένως πιο ισορροπημένο και κατάλληλο για datasets που περιέχουν πολλές μικρές ή μηδενικές παρατηρήσεις,

όπως στην περίπτωση των προϊόντων χαμηλής ή ακανόνιστης ζήτησης. Παράλληλα, η συμμετρική φύση του δείκτη δυσχεραίνει κάπως την ερμηνεία του και ενδέχεται να μην αποτυπώνει πάντοτε με σαφήνεια τη σημασία μιας απόκλισης σε επιχειρησιακό επίπεδο (Makridakis et al., 1998).

Τακτικοποίηση (Regularization) & Generalization

Σε ML: depth/learning rate/γαντζωμένοι υπερπαράμετροι (XGBoost), subsampling κ.ά.

Σε DL: dropout, L2, early stopping, data scaling, σωστά batch sizes.

Η τακτικοποίηση στοχεύει στον περιορισμό της υπερπροσαρμογής, ιδίως όταν υπάρχουν πολλά features, λίγα σημεία ανά σειρά ή υψηλή ύφορυβώδης διακύμανση.

Ερμηνευσιμότητα (Explainability) – SHAP

Τα ισχυρά μοντέλα (XGBoost/MLP) συχνά είναι «μαύρα κουτιά». Με SHAP values μπορούμε να αποδώσουμε σε κάθε χαρακτηριστικό τη συνεισφορά του στην εκάστοτε πρόβλεψη (global & local explainability). Επιχειρησιακά, αυτό μεταφράζεται σε εξηγήσεις τύπου «η μείωση τιμής και η ύπαρξη promo αύξησαν τη ζητούμενη ποσότητα κατά X». Έτσι, γεφυρώνεται η ανάγκη ακρίβειας με την ανάγκη εμπιστοσύνης και ελέγχου (Lundberg & Lee, 2017).

Ακραίες Τιμές, Ελλείψεις, Stock-outs

Προωθητικές περίοδοι, stock-outs, ασυνέχειες καταγραφής κ.λπ. μπορούν να «διαστρεβλώσουν» μάθηση και αξιολόγηση. Επιβάλλεται εντοπισμός/σήμανση γεγονότων (event flags), κατάλληλη απαλοιφή «τεχνητών» outliers, συντηρητική συμπλήρωση κενών (imputation) και—όπου είναι δυνατό—χρήση εξωγενών μεταβλητών που εξηγούν τις αιχμές αντί να τις αφήνουν στα κατάλοιπα.

Ιεραρχικές και Ομαδοποιημένες Προβλέψεις

Πολλές επιχειρήσεις απαιτούν συνέπεια (reconciliation) από το λεπτό επίπεδο (SKU×Κατάστημα) έως το συνολικό (εταιρεία). Μεθοδολογίες ιεραρχικής πρόβλεψης και συγχρονισμού επιτρέπουν οι προβλέψεις να «κουμπώνουν» αριθμητικά σε όλα τα επίπεδα ανάλυσης, κάτι που διευκολύνει τον εφοδιασμό και τον οικονομικό προγραμματισμό.

Κεφάλαιο C

Σχετιζόμενη Μελέτη

C.1 Σχετική Έρευνα στην Πρόβλεψη Πωλήσεων

Η ερευνητική παραγωγή γύρω από την πρόβλεψη πωλήσεων είναι τεράστια και διατρέχει ολόκληρο το φάσμα των τεχνικών: από τα κλασικά στατιστικά μοντέλα χρονοσειρών, στις οικονομετρικές προσεγγίσεις με εξωγενείς μεταβλητές, από τα μοντέλα μηχανικής μάθησης για «ταμπελοποιημένα» δεδομένα (tabular data), έως τις νευρωνικές αρχιτεκτονικές της βαθιάς μάθησης που μαθαίνουν περίπλοκες, μη γραμμικές εξαρτήσεις. Στην καρδιά αυτού του σώματος γνώσης βρίσκονται τόσο μεγάλες, διεθνώς αναγνωρισμένες μελέτες αξιολόγησης μεθόδων (όπως οι διαγωνισμοί M-competitions), όσο και θεμελιωτικές εργασίες που καθόρισαν βέλτιστες πρακτικές, μετρικές αξιολόγησης και διαδικασίες επικύρωσης για χρονοσειρές. Το παρόν τμήμα επιχειρεί μια συνεκτική αφήγηση αυτής της πορείας, εστιάζοντας σε εκείνα τα ευρήματα που είναι πιο χρήσιμα για τον σχεδιασμό και την τεκμηρίωση ενός σύγχρονου συστήματος πρόβλεψης πωλήσεων.

Κλασικές μελέτες και θεμέλια: ARIMA/ETS, Theta και ιεραρχική πρόβλεψη

Η παράδοση των Box–Jenkins (Box, Jenkins, Reinsel, & Ljung, 2016) θεμελίωσε την αναλυτική προσέγγιση στη μοντελοποίηση χρονοσειρών μέσω των ARIMA και των εποχικών τους επεκτάσεων (SARIMA), εισάγοντας μια πειθαρχημένη διαδικασία αναγνώρισης, εκτίμησης και διάγνωσης (ACF/PACF, AIC/BIC, έλεγχοι καταλοίπων).

Παράλληλα, η οικογένεια της εκθετικής εξομάλυνσης εξελίχθηκε από τις κλασικές μορφές Holt-Winters σε state space παραμετροποιήσεις (γνωστές ως ETS), που παρέχουν στιβαρό θεωρητικό υπόβαθρο στα πρακτικά σχήματα εξομάλυνσης και εξηγούν γιατί και πότε αυτά λειτουργούν τόσο καλά σε σταθερές εποχικές δομές (Hyndman & Athanasopoulos, 2021).

Ένα σημαντικό ελληνικό απούπωμα στη διεύνη βιβλιογραφία είναι η μέθοδος Theta, η οποία, αποδομώντας ένα σήμα σε «θάτα-γραμμές» και ανασυνθέτοντας προβλέψεις, κατέγραψε ιδιαίτερα καλή απόδοση στον παγκόσμιο διαγωνισμό M3, καθιερώνοντας την ιδέα ότι απλούστερες, σωστά ρυθμισμένες μέθοδοι μπορούν να ανταγωνιστούν σύνθετα σχήματα σε μεγάλες συλλογές ετερογενών σειρών (Assimakopoulos & Nikolopoulos, 2000).

Σε οργανισμούς με πολλαπλά επίπεδα αναφοράς (SKU → κατηγορία → κατάστημα → περιοχή → σύνολο), οι τεχνικές iεραρχικής πρόβλεψης και «συμφιλίωσης» (reconciliation) των προβλέψεων, ώστε οι ανθροίσεις να «χλείνουν» αριθμητικά σε όλα τα επίπεδα, προσέφεραν μια συνεπή μεθοδολογία που γεφυρώνει την αναλυτική ακρίβεια με τις επιχειρησιακές ανάγκες (Hyndman, Ahmed, Athanasopoulos, & Shang, 2011).

Συμπληρωματικά προς τις «συνεχείς» χρονοσειρές, ένας κομβικός κλάδος έρευνας αφορά τη διαλλειπτική ζήτηση (intermittent demand), συχνή σε ανταλλακτικά, βιομηχανικά είδη και αργοκίνητα SKU. Η κλασική προσέγγιση του Croston και οι μεταγενέστερες διορθώσεις της (Syntetos & Boylan) έδειξαν ότι όταν οι μηδενικές περίοδοι είναι συχνές, οι συνήθεις μέθοδοι παρεκτιμούν ή υπερεκτιμούν συστηματικά τη ζήτηση, και χρειάζονται ειδικά προσαρμοσμένες τεχνικές που μοντελοποιούν χωριστά το μέγεθος και το διάστημα μεταξύ των ανανεώσεων (Croston, 1972; Syntetos & Boylan, 2005). Αυτές οι εργασίες είναι κρίσιμες για ρεαλιστική διαχείριση αποθεμάτων σε περιβάλλοντα λιανικής και after-sales υπηρεσιών.

Τα M-competitions ως «καθρέφτης» του πεδίου: M3, M4 και M5

Κομβικό ρόλο στη διαμόρφωση συναίνεσης για το τι «δουλεύει» στην πράξη έπαιξαν οι μεγάλοι διαγωνισμοί πρόβλεψης M-competitions. Ο διαγωνισμός M3 κατέδειξε την ανθεκτικότητα απλών, καλά βαθμονομημένων μεθόδων (όπως η Theta και η ETS) σε ευρύ φάσμα σειρών (Makridakis & Hibon, 2000).

Ο M4 (2018) ανέδειξε μια υβριδική πραγματικότητα: συνδυασμοί κλασικών και νευρωνικών μεθόδων, συνοδευόμενοι από προσεγμένες διαδικασίες εξομάλυνσης/συνένωσης, έδωσαν κορυφαίες επιδόσεις, με τη νικητήρια προσέγγιση (Smyl) να υιοθετεί ένα υβρίδιο εξομάλυνσης

και RNN (Smyl, 2020).

Ο M5 (2020), που εστίασε σε λιανικές πωλήσεις σε iεραρχική δομή (Walmart), σηματοδότησε την έμφαση στα tabular δεδομένα με πλούσιες εξωγενείς μεταβλητές: τα κορυφαία συστήματα αξιοποίησαν gradient boosting (LightGBM/XGBoost) με εκτεταμένο feature engineering (ημερολόγιο, promos, υστερήσεις/κυλιόμενα) και time-aware validation, δείχνοντας ότι σε προβλήματα λιανικής η σωστή αξιοποίηση εξωτερικής πληροφορίας μπορεί να υπερκαλύψει τη σημασία της «καθαρής» αυτοσυσχέτισης (Makridakis et al., 2020).

Η πρακτική επίγνωση που προκύπτει από τα M-competitions είναι διπλή: πρώτον, δεν υπάρχει μία μέθοδος που να υπερέχει παντού· δεύτερον, η μεθοδολογία (features, validation, iεραρχική συμφιλίωση) είναι συχνά τόσο σημαντική όσο και η ίδια η επιλογή του μοντέλου.

Από τη στατιστική στη μηχανική μάθηση: ensembles, meta-learning και αξιολόγηση

Στο μεταίχμιο στατιστικής και μηχανικής μάθησης, οι ensembles και τα boosting μοντέλα καθιερώθηκαν ως «πρώτης γραμμής» λύσεις για πωλήσεις με πλούσια χαρακτηριστικά. Το XGBoost (Chen & Guestrin, 2016) και τα παράγωγά του (LightGBM, CatBoost) ενσωμάτωσαν αποδοτικές διαδικασίες δόμησης δέντρων, τακτικοποίηση, υποδειγματική διαχείριση ελλιπών/αραιών τιμών και ευελιξία υπερπαραμέτρων, προσφέροντας άριστη σχέση απόδοσης/ταχύτητας σε μεγάλα tabular σύνολα.

Παράλληλα, η ιδέα του meta-learning για συνδυασμό προβλέψεων πολλών μεθόδων («τις βάζουμε να συναγωνιστούν και να συνεργαστούν») κωδικοποιήθηκε σε σχήματα όπως το FFORMA, που μαθαίνουν βάρη συνδυασμού ανάλογα με τα χαρακτηριστικά της σειράς, αποσπώντας κέρδος από την ετερογένεια (Montero-Manso, Athanasopoulos, Hyndman, & Talagala, 2020).

Στο μεθοδολογικό επίπεδο, η σωστή αξιολόγηση παραμένει θεμέλιο: η κλασική εργασία των Bergmeir & Benítez (2012) εξηγεί γιατί ο τυχαίος k-fold cross-validation είναι παραπλανητικός για χρονοσειρές και προτείνει κυλιόμενες (rolling) διασπάσεις που σέβονται τον χρόνο, ενώ οι Hyndman & Koehler (2006) επανεξετάζουν τα μέτρα σφάλματος, εισάγοντας το MASE ως πιο σταθερό εναλλακτικό του MAPE, ιδίως όταν υπάρχουν μηδενικές ή πολύ μικρές πραγματικές τιμές.

Από τη στατιστική στη μηχανική μάθηση: ensembles, meta-learning και αξιολόγηση

Στο μεταίχμιο στατιστικής και μηχανικής μάθησης, τα ensembles και τα boosting μοντέλα καθιερώθηκαν ως «πρώτης γραμμής» λύσεις για πωλήσεις με πλούσια χαρακτηριστικά. Το XGBoost (Chen & Guestrin, 2016) και τα παράγωγά του (LightGBM, CatBoost) ενσωμάτωσαν αποδοτικές διαδικασίες δόμησης δέντρων, ταχτικοποίηση, υποδειγματική διαχείριση ελλιπών/αραιών τιμών και ευελιξία υπερπαραμέτρων, προσφέροντας άριστη σχέση απόδοσης/ταχύτητας σε μεγάλα tabular σύνολα.

Παράλληλα, η ιδέα του meta-learning για συνδυασμό προβλέψεων πολλών μεθόδων («τις βάζουμε να συναγωνιστούν και να συνεργαστούν») υλοποιήθηκε σε σχήματα όπως το FFOMA, που μαθαίνουν βάρη συνδυασμού ανάλογα με τα χαρακτηριστικά της σειράς, αποσπώντας κέρδος από την ετερογένεια (Montero-Manso, Athanasopoulos, Hyndman, & Talagala, 2020).

Σε μεθοδολογικό επίπεδο, η σωστή αξιολόγηση παραμένει θεμέλιο: η κλασική εργασία των Bergmeir & Benítez (2012) εξηγεί γιατί ο τυχαίος k-fold cross-validation είναι παραπλανητικός για χρονοσειρές και προτείνει κυλιόμενες (rolling) διασπάσεις που σέβονται τον χρόνο, ενώ οι Hyndman & Koehler (2006) επανεξετάζουν τα μέτρα σφάλματος, εισάγοντας το MASE ως πιο σταθερό εναλλακτικό του MAPE, ιδιαίτερα όταν υπάρχουν μηδενικές ή πολύ μικρές προγνωστικές τιμές.

Εξηγησιμότητα και επιχειρησιακή ενσωμάτωση: από τις τιμές SHAP στις αποφάσεις

Όσο η έρευνα μετακινείται προς «ισχυρά αλλά αδιαφανή» μοντέλα, αυξάνει η ανάγκη για εξηγησιμότητα. Η εργασία των Lundberg & Lee (2017) εισήγαγε τις τιμές SHAP, που προσφέρουν μια συνεπή μέθοδο απόδοσης «συνεισφοράς» κάθε χαρακτηριστικού σε μια πρόβλεψη, τόσο τοπικά όσο και συνολικά. Στην πράξη λιανικής, αυτό μεταφράζεται σε πολύτιμες αφηγήσεις: ποιο είναι το ποσοτικό αποτύπωμα της τιμής, του τύπου προώθησης, ή της χρονικής στιγμής (ημέρα/εβδομάδα/εποχή) πάνω στην αναμενόμενη ζήτηση;

Σε συνδυασμό με σωστή επικύρωση (walk-forward), ιεραρχική συμφιλίωση και μετρικές κατάλληλες για πωλήσεις (MAE/RMSE/sMAPE/MASE), το σκέλος της εξήγησης καθιστά τις προχωρημένες μεθόδους μεταφράσιμες σε πολιτικές αποφάσεις, όπως προϋπολογισμός pro-

mos, διαχείριση αποθεμάτων και σχεδιασμός καναλιών.

Τι «μένει» από τη σχετική έρευνα για μια σύγχρονη εφαρμογή πωλήσεων

Η συνολική εικόνα που αναδύεται από τις μελέτες είναι ρεαλιστική και ισορροπημένη. Πρώτον, δεν υπάρχει πανάκεια: η απόδοση εξαρτάται από τη φύση των δεδομένων, το επίπεδο granular πρόβλεψης, τη διαθεσιμότητα εξωγενών μεταβλητών και τον ορίζοντα πρόβλεψης.

Δεύτερον, η μεθοδολογία — feature engineering, time-aware validation, ιεραρχική συμφιλίωση, τακτικοποίηση και εξήγηση — είναι συχνά καθοριστικότερη από τη «μάχη αλγορίθμων».

Τρίτον, για tabular δεδομένα λιανικής με πλούσιες εξωτερικές μεταβλητές, τα gradient boosting και τα MLP με σωστή προεπεξεργασία αποτελούν «ασφαλείς» κορμούς, ενώ όπου η καθαρή χρονοσειρά κρύβει κρίσιμη πληροφορία, τα LSTM/GRU/CNN και τα TFT/N-BEATS αποδίδουν ιδιαίτερα καλά σε global εκπαίδευση.

Τέταρτον, η ερμηνευσιμότητα μέσω SHAP/attention και η αξιοπιστία αξιολόγησης μέσω rolling backtesting αποτελούν αναγκαίες προϋποθέσεις για υιοθέτηση σε παραγωγή.

Αυτές οι κατευθύνσεις συνάδουν με τις ανάγκες μιας πτυχιακής που αναπτύσσει ένα MLP ή συγγενές νευρωνικό σε Python/TensorFlow: η βιβλιογραφία υποστηρίζει τεχνικά τις σχεδιαστικές επιλογές (embedding κατηγοριών, υστερήσεις/κυλιόμενα παράθυρα, dropout/L2/early stopping, walk-forward validation) και παρέχει τον ψεωρητικό «σκελετό» που δικαιολογεί γιατί αυτές οι επιλογές είναι κατάλληλες για πρόβλεψη πωλήσεων σε σύγχρονο επιχειρησιακό περιβάλλον.

Κεφάλαιο D

Πλαίσιο Ανάπτυξης και Εκτέλεσης

D.1 Περιβάλλον Ανάπτυξης και Εκτέλεσης

Η ανάπτυξη, ο πειραματισμός και η εκτέλεση του πηγαίου κώδικα της παρούσας εργασίας πραγματοποιήθηκαν σε περιβάλλον Jupyter Notebook (JupyterLab/Notebook), και όχι ως αυτόνομο εκτελέσιμο αρχείο Python (.py). Η επιλογή του Jupyter δεν είναι τυχαία· συνάδει με τις διεθνείς βέλτιστες πρακτικές για διαδραστική, επαναλήψιμη και τεκμηριωμένη υπολογιστική ανάλυση: επιτρέπει εκτέλεση κώδικα σε κελιά, άμεση οπτικοποίηση αποτελεσμάτων, ενσωμάτωση κειμένου (Markdown) για σχολιασμό/μεθοδολογικές σημειώσεις και, χυρίως, διευκολύνει τον κύκλο πειραματισμού (ταχεία αλλαγή υπερπαραμέτρων, έλεγχος σφαλμάτων, καταγραφή ευρημάτων) χωρίς να διαρρηγνύεται η ροή της ανάλυσης (Kluyver et al., 2016; Pérez Granger, 2007; Rule, Tabard, Hollan, 2019). Έτσι, το notebook λειτούργησε ως «εργαστηριακός πάγκος», όπου το pipeline προεπεξεργασίας, εκπαίδευσης, αξιολόγησης και οπτικοποίησης συνετέθη, δοκιμάστηκε και τεκμηριώθηκε σε ένα ενιαίο, αναπαραγώγιμο έγγραφο.

D.1.1 Γλώσσα Προγραμματισμού και Διερμηνευτής

Η εργασία υλοποιήθηκε στη γλώσσα Python (έκδοση 3.x), η οποία αποτελεί de facto πρότυπο για επιστημονικό υπολογισμό, ανάλυση δεδομένων και μηχανική μάθηση, χάρη στο ώριμο οικοσύστημα βιβλιοθηκών, την καθαρότητα σύνταξης και την εκφραστικότητα που επιτρέπει γρήγορη πρωτότυποποίηση (Van Rossum Drake, 2009). Ο διερμηνευτής Python ήταν

ενσωματωμένος στο kernel του Jupyter, εξασφαλίζοντας την ομαλή εναλλαγή ανάμεσα σε κώδικα, κείμενο και γραφήματα.

D.1.2 Περιβάλλον Εργασίας και Αναπαραγωγιμότητα

Για την αναπαραγωγιμότητα των αποτελεσμάτων, χρησιμοποιήθηκε απομόνωση εξαρτήσεων (π.χ. `venv` ή `conda` περιβάλλον) και καταγράφηκαν οι εκδόσεις βιβλιοθηκών στο τέλος του πειραματικού κύκλου (π.χ. με `pip freeze`), ώστε να μπορεί να ανακατασκευαστεί το ίδιο software stack σε δεύτερο μηχάνημα. Επιπλέον, ορίστηκαν τυχαίοι σπόροι (random seeds) σε NumPy/TensorFlow/OS, για να μειωθεί η στοχαστικότητα στην εκπαίδευση και να επιτευχθεί σχεδόν ντετερμινιστική συμπεριφορά όπου αυτό είναι εφικτό (π.χ. σε MLP χωρίς μη-ντετερμινιστικούς πυρήνες). Η λογική αυτή ευθυγραμμίζεται με τις οδηγίες της διευθούντος κοινότητας για επαναλήψιμη υπολογιστική έρευνα σε notebooks (Rule et al., 2019). Για έλεγχο εκδόσεων κώδικα και τεκμηρίωση αλλαγών, χρησιμοποιήθηκε σύστημα ελέγχου εκδόσεων (π.χ. Git), ιδιαίτερα χρήσιμο όταν γίνονται πολλαπλές δοκιμές με διαφορετικές αρχιτεκτονικές/υπερπαραμέτρους. Η παραχολούθηση αλλαγών στον κώδικα και η τεκμηρίωση κάθε πειραματικού βήματος διευκολύνουν τη συνεργασία και την αναπαραγωγή αποτελεσμάτων σε άλλα μηχανήματα.

D.1.3 Python

Η υλοποίηση βασίστηκε σε ένα συνεκτικό οικοσύστημα βιβλιοθηκών Python:

- **NumPy:** Παρέχει πολυδιάστατους πίνακες (arrays) υψηλής απόδοσης και αποδοτικές πράξεις γραμμικής άλγεβρας, οι οποίες αποτελούν τη βάση για σχεδόν όλες τις λειτουργίες επιστημονικού υπολογισμού σε Python (Harris et al., 2020). Στα πειράματα πρόβλεψης πωλήσεων, χρησιμοποιήθηκε για τον υπολογισμό υστερήσεων (lags) σε χρονοσειρές, για τη δημιουργία κυλιόμενων παραθύρων (rolling windows) που επιτρέπουν την εκτίμηση στατιστικών χαρακτηριστικών, καθώς και για την ταχεία εκτέλεση στοιχειο-πράξεων σε μεγάλους πίνακες δεδομένων. Επιπλέον, το NumPy διευκολύνει την προετοιμασία δεδομένων εισόδου για νευρωνικά δίκτυα, καθώς οι TensorFlow / Keras tensors βασίζονται στην ίδια δομή δεδομένων.
- **pandas:** Προσφέρει εύχρηστες δομές δεδομένων, όπως DataFrame και Series, οι οποίες υποστηρίζουν πλήθος λειτουργιών για καθαρισμό, μετασχηματισμό και συγ-

χώνευση δεδομένων (McKinney, 2010). Στο πλαίσιο της πρόβλεψης πωλήσεων, το pandas χρησιμοποιήθηκε για την οργάνωση χρονοσειρών πωλήσεων, τη διαχείριση ελλιπών ή ακραίων τιμών, την εφαρμογή rolling και expanding συναρτήσεων, και τη σύνδεση των εσωτερικών δεδομένων πωλήσεων με εξωγενείς πίνακες, όπως τιμές, προωθήσεις ή αργίες. Η δυνατότητα χρονικής ευθυγράμμισης των δεδομένων καθιστά το pandas εργαλείο απαραίτητο για την προετοιμασία datasets πριν την εκπαίδευση μοντέλων.

- **scikit-learn:** Παρέχει ολοκληρωμένα εργαλεία για την κατασκευή pipelines προεπεξεργασίας, διάσπαση δεδομένων σε train/validation/test sets, κανονικοποίηση ή scaling, καθώς και ρύθμιση υπερπαραμέτρων μέσω grid ή random search (Pedregosa et al., 2011). Στα πειράματα, χρησιμοποιήθηκε για τη δημιουργία time-aware splits ώστε να διατηρείται η χρονική συνοχή των χρονοσειρών, για την εφαρμογή StandardScaler ή MinMaxScaler σε χαρακτηριστικά, και για την αξιολόγηση μοντέλων με μετρικές όπως MAE, RMSE και R^2 . Επιπλέον, η scikit-learn διευκόλυνε τη σύνδεση με TensorFlow / Keras, παρέχοντας σταθερές εισόδους και προκαθορισμένες διαδικασίες preprocessing.
- **TensorFlow / Keras:** Αποτελούν το βασικό εργαλείο για τη δημιουργία και εκπαίδευση MLP και άλλων νευρωνικών δικτύων (Abadi et al., 2016; Chollet, 2015; Kingma & Ba, 2015). Η Keras επιτρέπει τον υψηλού επιπέδου ορισμό μοντέλων μέσω Sequential ή Functional API, ενώ το TensorFlow αναλαμβάνει την υπολογιστική βελτιστοποίηση, την εκτέλεση gradient descent, την αξιοποίηση GPU και την παρακολούθηση callbacks όπως early stopping και model checkpointing. Στα πειράματα, εφαρμόστηκαν επίσης κανονικοποίηση εισόδων, L2 regularization και dropout, για βελτίωση της γενίκευσης, καθώς και έλεγχος επιδόσεων σε σετ επικύρωσης. Η χρήση αυτών των βιβλιοθηκών διευκόλυνε την ταχεία κατασκευή, εκπαίδευση και αξιολόγηση πολλαπλών μοντέλων πρόβλεψης πωλήσεων.
- **Matplotlib / Seaborn:** Χρησιμοποιούνται για την οπτικοποίηση δεδομένων, καμπυλών εκπαίδευσης/επικύρωσης (learning curves), κατανομών σφαλμάτων (residuals), εποχικών μοτίβων και στατιστικών διαγραμμάτων (Hunter, 2007; Waskom, 2021). Η Matplotlib λειτουργεί ως βασικό 2D γραφικό περιβάλλον, ενώ η Seaborn προσφέρει εύχρηστα εργαλεία για στατιστική απεικόνιση υψηλού επιπέδου. Στα πειράματα, οι βιβλιοθήκες αυτές χρησιμοποιήθηκαν για την ανάλυση της συμπεριφοράς των μοντέλων, τον εντοπισμό overfitting ή underfitting, και τη διευκόλυνση διαγνωστικών αναλύσεων, επιτρέποντας άμεση ανατροφοδότηση κατά την εκπαίδευση των μοντέλων.

Σε περίπτωση ανάγκης για GPU, το TensorFlow μπορεί να αξιοποιήσει CUDA/cuDNN όπου διατίθενται (Abadi et al., 2016). Στην παρούσα εργασία, δεν απαιτείται ειδική GPU υποδομή, αλλά το MLP είναι συμβατό με επιτάχυνση όπου υπάρχει.

D.1.4 Jupyter

Τα πειράματα εκτελέστηκαν αποκλειστικά σε Jupyter Notebook / JupyterLab, δηλαδή σε ένα διαδραστικό περιβάλλον όπου ο κώδικας, το επεζηγηματικό κείμενο και τα γραφήματα συνυπάρχουν σε «κελιά». Το Jupyter επιτρέπει την εκτέλεση κώδικα «on the fly», την ενσωμάτωση εξισώσεων, οπτικοποιήσεων και σχολίων στο ίδιο έγγραφο (Project Jupyter, n.d.).

Αυτό το περιβάλλον προσφέρει:

- Άμεση δοκιμή και πειραματισμό με υπερπαραμέτρους.
- Οπτικοποίηση απωλειών και μετρικών σε πραγματικό χρόνο.
- Τεχμηρίωση βήμα-βήμα κάθε πειράματος, ενισχύοντας την αναπαραγωγιμότητα.
- Εύκολη επικοινωνία αποτελεσμάτων σε συνεργάτες και αξιολογητές.

D.1.5 Πίνακας Περιβάλλοντος

Πίνακας D.1: Πίνακας περιβάλλοντος ανάπτυξης, πακέτων και ρόλων τους στο pipeline πρόβλεψης πωλήσεων.

| Πακέτο (pip) | Import name | Ρόλος στο Pipeline |
|-----------------|--------------------|---|
| numpy | import numpy as np | Θεμέλιο αριθμητικών πινάκων/διανυσμάτων, υστερήσεις/rolling, γρήγορες πράξεις |

Συνέχεια από την προηγούμενη σελίδα

| Πακέτο (pip) | Import name | Ρόλος στο Pipeline |
|-----------------|--|--|
| pandas | <code>import pandas as pd</code> | DataFrames: καθαρισμός, μετασχηματισμοί, συγχωνεύσεις (π.χ. promos/τιμές) |
| scipy | <code>import scipy ...</code> | Στατιστικές/βοηθητικές ρουτίνες, φίλτρα/μετασχηματισμοί |
| scikit-learn | <code>from sklearn ...</code> | Preprocessing (scalers), splits (time-aware), μετρικές, pipelines |
| tensorflow | <code>import tensorflow as tf</code> | Εκπαίδευση MLP (Keras API), βελτιστοποιητές, callbacks |
| matplotlib | <code>import matplotlib.pyplot as plt</code> | Οπτικοποιήσεις χρονοσειρών, residuals, learning curves |
| xgboost | <code>import xgboost as xgb</code> | Gradient boosting για συγκριτική αξιολόγηση |

Κεφάλαιο Ε

Δεδομένα

E.1 Το Dataset και η Πηγή του Μοντέλου

Για τη δημιουργία του προγνωστικού μοντέλου, χρησιμοποιήθηκε το Superstore Sales Dataset που προέρχεται από την Kaggle και περιλαμβάνει δεδομένα από πωλήσεις προϊόντων σε μία διεθνή αλυσίδα καταστημάτων λιανικής. Το σύνολο δεδομένων χρησιμοποιείται για χρονική ανάλυση και πρόβλεψη πωλήσεων για τις επόμενες ημέρες, χρησιμοποιώντας δεδομένα που καταγράφηκαν από προηγούμενες παραγγελίες. Η πηγή του dataset είναι η εξής:

- Πηγή : Superstore Sales Dataset - Kaggle
<https://www.kaggle.com/datasets/rohitsahoo/sales-forecasting/data>

Το dataset περιλαμβάνει 9800 παρατηρήσεις και 18 στήλες, οι οποίες περιγράφονται αναλυτικά παρακάτω. Κάθε παρατήρηση αντιστοιχεί σε μία παραγγελία και περιλαμβάνει διάφορες πληροφορίες για την πώληση (όπως την ημερομηνία, το κέρδος, την κατηγορία του προϊόντος, και την τοποθεσία).

E.2 Πληροφορίες για το Dataset

Το Superstore Sales Dataset περιέχει 9.800 γραμμές (observations) και 18 στήλες (attributes), οι οποίες παρέχουν δεδομένα σχετικά με τις παραγγελίες, τις περιοχές, τις κατηγορίες προϊόντων, την ημερομηνία των παραγγελιών, καθώς και το κέρδος και την έκπτωση

για κάθε παραγγελία. Ακολουθεί ο πίνακας με τις 18 στήλες του dataset:

| Στήλη | Περιγραφή | Τιμές / Αριθμοί |
|---------------|--|---|
| Row ID | Μοναδικός αριθμός για κάθε σειρά δεδομένων | 9800 (γραμμές) |
| Order ID | Μοναδικός αριθμός παραγγελίας | 4922 μοναδικές τιμές |
| Order Date | Ημερομηνία παραγγελίας | 1230 μοναδικές τιμές |
| Ship Date | Ημερομηνία αποστολής της παραγγελίας | 1326 μοναδικές τιμές |
| Ship Mode | Τρόπος αποστολής (Standard Class, Second Class κ.λπ.) | 3 τύποι (Standard Class, Second Class, Other) |
| Customer ID | Μοναδικός αριθμός πελάτη | 793 μοναδικές τιμές |
| Customer Name | Όνομα πελάτη | 793 μοναδικές τιμές |
| Segment | Τμήμα πελάτη (Consumer, Corporate, Home Office) | 3 κατηγορίες (Consumer: 52%, Corporate: 30%, Other: 18%) |
| Country | Χώρα του πελάτη | 1 χώρα (United States) |
| City | Πόλη του πελάτη | 4 κύριες πόλεις (New York City: 9%, Los Angeles: 7%, άλλες πόλεις: 83%) |
| Product | Κατηγορία προϊόντος (π.χ. Office Supplies, Technology) | 3 κατηγορίες |
| Category | | |
| Sales | Συνολικές πωλήσεις για την παραγγελία | - |
| Quantity | Ποσότητα των προϊόντων που πουλήθηκαν | - |
| Discount | Ποσοστό έκπτωσης που εφαρμόστηκε στην παραγγελία | - |

Πίνακας E.1: Περιγραφή των στηλών του dataset και βασικά στατιστικά για κάθε χαρακτηριστικό.

E.3 Στατιστικά του Dataset

Ακολουθεί η ανάλυση μερικών χαρακτηριστικών και των μοναδικών τιμών τους:

- Order ID (Αριθμός Παραγγελίας): Το πεδίο αυτό περιέχει 4922 μοναδικές παραγγελίες.
- Order Date (Ημερομηνία Παραγγελίας): Υπάρχουν 1230 μοναδικές ημερομηνίες.
- Ship Mode (Τρόπος Αποστολής): Υπάρχουν 3 τύποι αποστολής, με την πλειοψηφία των παραγγελιών (60)
- Customer ID (Αριθμός Πελάτη): Υπάρχουν 793 μοναδικοί πελάτες στο σύνολο των παραγγελιών.
- Segment (Τμήμα Πελάτη): Τα τμήματα των πελατών κατανέμονται ως εξής: 52

E.4 Σύνολο Δεδομένων

Το σύνολο δεδομένων Superstore Sales Dataset περιέχει:

- 9800 παρατηρήσεις (γραμμές), που αντιστοιχούν σε παραγγελίες λιανικής.
- 18 στήλες (attributes) που περιλαμβάνουν διάφορες πτυχές κάθε παραγγελίας, όπως το προϊόν, την κατηγορία, το κέρδος και την τοποθεσία του πελάτη.

Αυτές οι πληροφορίες θα χρησιμοποιηθούν για να πραγματοποιηθεί η προεπεξεργασία των δεδομένων, που είναι το επόμενο βήμα πριν την εφαρμογή του προγνωστικού μοντέλου.

E.5 Ανάλυση Κατηγοριών Πελατών και Περιοχών

Η στήλη Segment (Τμήμα Πελάτη) περιλαμβάνει 3 κατηγορίες:

- Consumer (Καταναλωτής): 52
- Corporate (Επαρχικός): 30
- Other ('Άλλοι): 18

Η κατανομή αυτών των κατηγοριών είναι σημαντική για την ανάλυση και κατηγοριοποίηση των δεδομένων. Μπορεί να επηρεάσει τη στρατηγική πρόβλεψης, καθώς οι καταναλωτές και οι εταιρικοί πελάτες ενδέχεται να έχουν διαφορετικά μοτίβα αγοράς. Η στήλη City (Πόλη) δείχνει τις πόλεις με τη μεγαλύτερη παρουσία πελατών. New York City (9 Αυτές οι αναλύσεις θα χρησιμοποιηθούν για να ενσωματωθούν επιπλέον εξωγενείς μεταβλητές (όπως η πόλη ή το τμήμα του πελάτη) στις προβλέψεις μας.

Κεφάλαιο F

Προεπεξεργασία και Προετοιμασία Δεδομένων

F.1 Προεπεξεργασία Δεδομένων

Η προεπεξεργασία των δεδομένων είναι μία από τις πιο χρήσιμες διαδικασίες πριν την εφαρμογή οποιουδήποτε μοντέλου μηχανικής μάθησης ή βαθιάς μάθησης. Στο πλαίσιο της πρόβλεψης πωλήσεων, το dataset περιλαμβάνει διάφορους τύπους δεδομένων (συνέχεια, κατηγορικές τιμές, ημερομηνίες) που απαιτούν ειδική μεταχείριση ώστε να καταστούν κατάλληλα για εκπαίδευση. Η διαδικασία αυτή περιλαμβάνει καθαρισμό δεδομένων, κωδικοποίηση κατηγορικών μεταβλητών, διαχείριση ακραίων τιμών και μετασχηματισμούς που είναι απαραίτητοι για την αποτελεσματική εκπαίδευση του μοντέλου. Αναλυτικά, τα βήματα που ακολουθήθηκαν είναι τα εξής:

1. Καθαρισμός Δεδομένων και Αφαίρεση Μη Χρησίμων Στηλών

Αρχικά, το dataset περιέχει στήλες που δεν είναι χρήσιμες για την ανάλυση ή την πρόβλεψη πωλήσεων και επομένως πρέπει να αφαιρεθούν. Αυτές οι στήλες περιλαμβάνουν πληροφορίες που δεν επηρεάζουν τη συμπεριφορά των πωλήσεων, όπως το *Row ID*, το *Customer Name*, και το *Product Name*. Η αφαίρεση αυτών των δεδομένων είναι σημαντική για να επικεντρωθούμε

στις παραμέτρους που πραγματικά επηρεάζουν την πρόβλεψη (όπως το *Order ID*, *Order Date*, *Ship Date*, κ.λπ.).

```
# Drop unnecessary columns
df = df.drop(['Row ID', 'Customer Name', 'Country', 'Product Name'], axis=1)
df = df.drop(['Order ID', 'Customer ID'], axis=1)
```

Σχήμα F.1: Καθαρισμός Δεδομένων: Αφαίρεση μη χρησίμων στηλών από το dataset.

2. Εξαγωγή Χρονικών Στοιχείων από Ημερομηνίες

Στο dataset περιλαμβάνονται οι στήλες *Order Date* και *Ship Date*, οι οποίες περιέχουν πληροφορίες για την ημερομηνία παραγγελίας και αποστολής. Αυτές οι ημερομηνίες είναι ιδιαίτερα χρήσιμες, καθώς επηρεάζουν τις πωλήσεις λόγω εποχικότητας, ειδικών γεγονότων ή προωθητικών ενεργειών. Αντί να χρησιμοποιούμε ολόκληρη την ημερομηνία, τις αναλύουμε σε επιμέρους χαρακτηριστικά, όπως το έτος, τον μήνα και την ημέρα, καθώς αυτές οι μεταβλητές μπορεί να συμβάλλουν στην κατανόηση του κύκλου πωλήσεων και της ζήτησης.

Για παράδειγμα, ο μήνας μπορεί να επηρεάζει τις πωλήσεις λόγω εποχικότητας, όπως οι γιορτές ή οι καλοκαιρινές εκπώσεις. Με τη μετατροπή αυτών των ημερομηνιών σε επιμέρους χαρακτηριστικά, δημιουργούμε νέες δυνατότητες για το μοντέλο, επιτρέποντας του να μάθει τις εξαρτήσεις με βάση το χρόνο.

```
# Encode dates (Order Date and Ship Date) into year, month, and day
df = encode_dates(df, column='Order Date')
df = encode_dates(df, column='Ship Date')
```

Σχήμα F.2: Εξαγωγή Χρονικών Στοιχείων: Επεξεργασία των ημερομηνιών σε χαρακτηριστικά έτους, μήνα και ημέρας για καλύτερη χρήση στα μοντέλα πρόβλεψης.

3. One-Hot Encoding για Κατηγορικές Στήλες

Η επόμενη διαδικασία είναι η κωδικοποίηση κατηγορικών μεταβλητών. Στο dataset υπάρχουν πολλές στήλες που περιέχουν κατηγορικές τιμές, όπως το *Ship Mode*, το *Segment*, το *City*, το *State*, το *Product Category*, κ.λπ. Τα μη αριθμητικά δεδομένα πρέπει να μετατραπούν σε αριθμητική μορφή για να είναι συμβατά με τα μοντέλα μηχανικής μάθησης.

Η μέθοδος που χρησιμοποιείται για αυτό είναι η *One-Hot Encoding*, η οποία δημιουργεί μια νέα στήλη για κάθε κατηγορία μιας κατηγορηματικής μεταβλητής και αναθέτει σε κάθε παρατήρηση την τιμή 1 για τη στήλη της κατηγορίας στην οποία ανήκει και 0 για τις υπόλοιπες στήλες. Αυτό επιτρέπει στο μοντέλο να κατανοήσει τις σχέσεις μεταξύ των διαφορετικών κατηγοριών, χωρίς να εισάγεται οποιαδήποτε ιεραρχία ή τάξη μεταξύ τους.

```
# Function to OneHot encode categorical columns
def onehot_encode(df, column):
    df = df.copy()
    dummies = pd.get_dummies(df[column], prefix=column)
    df = pd.concat([df, dummies], axis=1)
    df = df.drop(column, axis=1)
    return df
```

Σχήμα F.3: One-Hot Encoding: Διαδικασία εφαρμογής One-Hot Encoding στις κατηγορικές στήλες του dataset.

4. Διαχείριση Ακραίων Τιμών και Ανισομετρίας

Στη συνέχεια, εξετάζουμε τη διαχείριση των ακραίων τιμών και την ανισομετρία των δεδομένων, ιδιαίτερα στη στήλη *Sales* (πωλήσεις). Στα δεδομένα πωλήσεων, τα οποία συνήθως έχουν σημαντική διακύμανση, είναι σημαντικό να εφαρμοστεί ένας μετασχηματισμός για να μειωθεί η επίδραση των ακραίων τιμών και να γίνει η κατανομή των δεδομένων πιο συμμετρική.

Η λογαριθμική μετατροπή είναι η κατάλληλη μέθοδος για τη μείωση αυτών των ακραίων τιμών και την εξισορρόπηση της κατανομής. Εφαρμόζοντας τον λογάριθμο στη στήλη *Sales*, η κατανομή γίνεται πιο ομαλή, βοηθώντας το μοντέλο να μάθει καλύτερα τη σχέση μεταξύ των χαρακτηριστικών και των πωλήσεων.

```
# Handle outliers and skewness
df['Sales'] = np.log1p(df['Sales']) # Apply log transformation to Sales to manage skewness
```

Σχήμα F.4: Διαχείριση Ακραίων Τιμών: Λογαριθμική μετατροπή της στήλης *Sales* για τη διαχείριση της ασυμμετρίας.

5. Διαχωρισμός Δεδομένων σε Εκπαιδευτικά και Δοκιμαστικά Σετ

Ένα άλλο σημαντικό βήμα είναι ο διαχωρισμός των δεδομένων σε εκπαιδευτικά (*training*) και δοκιμαστικά (*testing*) σετ. Για να αποφύγουμε την υπερπροσαρμογή του μοντέλου (*overfitting*), το dataset διαχωρίζεται σε δύο μέρη: το 70% χρησιμοποιείται για την εκπαίδευση του μοντέλου και το υπόλοιπο 30% χρησιμοποιείται για τη δοκιμή του μοντέλου και την αξιολόγηση της απόδοσης του.

Αυτό επιτρέπει να εκτιμήσουμε την ικανότητα του μοντέλου να γενικεύει σε δεδομένα που δεν έχει δει κατά τη διάρκεια της εκπαίδευσης, γεγονός που είναι καθοριστικό για την αξιοπιστία της πρόβλεψης.

```
# Train-test split: 70% for training, 30% for testing
X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, shuffle=True, random_state=1)
```

Σχήμα F.5: Διαχωρισμός των Δεδομένων: 70% για εκπαίδευση και 30% για δοκιμή.

6. Κλιμάκωση των Χαρακτηριστικών (Feature Scaling)

Η κλιμάκωση των χαρακτηριστικών είναι επίσης απαραίτητη για τη βελτίωση της απόδοσης του μοντέλου. Επειδή τα χαρακτηριστικά του dataset μπορεί να έχουν διαφορετικές μονάδες μέτρησης (π.χ. οι πωλήσεις σε ευρώ και η ποσότητα σε μονάδες), πρέπει να κλιμακωθούν σε μία κοινή κλίμακα, ώστε να μην υπερισχύουν τα χαρακτηριστικά με μεγάλες τιμές έναντι αυτών με μικρότερες τιμές.

Χρησιμοποιήθηκε ο `StandardScaler` για να κανονικοποιηθούν όλα τα χαρακτηριστικά σε μια κλίμακα με μέσο όρο 0 και τυπική απόκλιση 1.

```
# Scale X (features) using StandardScaler
scaler = StandardScaler()
scaler.fit(X_train)
X_train = pd.DataFrame(scaler.transform(X_train), columns=X.columns)
X_test = pd.DataFrame(scaler.transform(X_test), columns=X.columns)
```

Σχήμα F.6: Κανονικοποίηση των τιμών σε μια κλίμακα με μέσο όρο 0 και τυπική απόκλιση 1

F.2 Συνολική Διαδικασία Προεπεξεργασίας

Η προεπεξεργασία των δεδομένων που πραγματοποιήθηκε περιλαμβάνει όλα τα παραπάνω βήματα, τα οποία εξασφαλίζουν ότι το dataset είναι έτοιμο για τη δημιουργία του προγνωστικού μοντέλου. Η σωστή διαχείριση των κατηγορικών μεταβλητών, η διαχείριση των ακραίων τιμών και η κανονικοποίηση των χαρακτηριστικών είναι χρίσιμα για την επιτυχία του μοντέλου και την απόδοσή του.

Ας προχωρήσουμε με την επόμενη ενότητα, αναλύοντας το επόμενο βήμα στην προεπεξεργασία των δεδομένων. Στο σημείο αυτό, η διαδικασία αφορά την εκπαίδευση του μοντέλου και τις τεχνικές ρύθμισης των υπερπαραμέτρων, με στόχο την καλύτερη απόδοση του μοντέλου πρόβλεψης πωλήσεων. Εδώ, θα εξηγήσουμε την εκπαίδευση του μοντέλου μέσω του Keras/TensorFlow και τη χρήση callbacks για την πρόληψη της υπερπροσαρμογής (*overfitting*), όπως EarlyStopping και ReduceLROnPlateau.

Κεφάλαιο G

Εκπαίδευση του Μοντέλου

G.1 Εκπαίδευση του Μοντέλου και Ρύθμιση Υπερπαραμέτρων

Το μοντέλο εκπαιδεύεται με Keras χρησιμοποιώντας τον βελτιστοποιητή Adam, ο οποίος προσαρμόζει τη μάθηση (learning rate) για να επιτύχει καλύτερη σύγκλιση κατά τη διάρκεια της εκπαίδευσης. Ο ρυθμός μάθησης έχει οριστεί σε 0.001, προκειμένου να επιτευχθεί αργή αλλά σταθερή σύγκλιση κατά την εκπαίδευση του μοντέλου. Ο βελτιστοποιητής χρησιμοποιεί την προσαρμογή της μάθησης για να εξασφαλίσει ότι το μοντέλο εκπαιδεύεται σωστά, αποφεύγοντας μεγάλες αποκλίσεις στις παραμέτρους που θα μπορούσαν να οδηγήσουν σε υπερπροσαρμογή ή υποεκπαίδευση.

Η εκπαίδευση πραγματοποιείται με την εντολή `model.fit()`, όπου τα `X_train` και `y_train` αποτελούν τα δεδομένα εκπαίδευσης. Το `X_train` περιλαμβάνει τα χαρακτηριστικά (features) και το `y_train` είναι η εξαρτημένη μεταβλητή των πωλήσεων (Sales). Το 20% των δεδομένων χρησιμοποιείται για επικύρωση (validation split), ώστε να διασφαλιστεί ότι το μοντέλο δεν υπερπροσαρμόζεται. Η εκπαίδευση γίνεται σε batch μεγέθους 256 παρατηρήσεων και για 200 εποχές (epochs), ή έως ότου το μοντέλο σταματήσει να βελτιώνεται.

Για την αποφυγή υπερπροσαρμογής και τη βελτίωση της σύγκλισης χρησιμοποιούνται δύο callbacks:

- EarlyStopping: Σταματά την εκπαίδευση όταν η απόδοση στο validation set δεν βελτιώνεται για ένα καθορισμένο αριθμό εποχών (patience), αποτρέποντας την υπερπροσαρμογή.

Παράλληλα, επαναφέρει τα καλύτερα βάρη του μοντέλου.

- ReduceLROnPlateau: Μειώνει τον ρυθμό μάθησης όταν η απόδοση του μοντέλου φτάσει σε ένα πλατό και δεν παρουσιάζει περαιτέρω βελτίωση, ώστε η εκπαίδευση να συνεχίζεται με μικρότερο learning rate για καλύτερη σύγκλιση.

Η παρακολούθηση της εκπαίδευσης δείχνει τη μείωση της απώλειας (loss) και της επικύρωσης απώλειας (val_loss) καθώς το μοντέλο εκπαιδεύεται, ενώ ο ρυθμός μάθησης μειώνεται όταν το μοντέλο συναντά επίπεδο (plateau).

```
# Compiling the model with adjusted learning rate for Adam
model.compile(
    optimizer=tf.keras.optimizers.Adam(learning_rate=0.001), # Experiment with a smaller learning rate
    loss='mse',
)
# Training the Model with adjusted Epochs and Dynamic Learning Rate
history = model.fit(
    x_train,
    y_train,
    validation_split=0.2,
    batch_size=256,
    epochs=200,
    callbacks=[
        tf.keras.callbacks.EarlyStopping(
            monitor='val_loss',
            patience=10,
            restore_best_weights=True,
        ),
        tf.keras.callbacks.ReduceLROnPlateau() # Reduce learning rate on plateau
    ],
)
Q 16.7s
Epoch 1/200
22/22 ━━━━━━━━━━ 4s 37ms/step - loss: 20.4293 - val_loss: 16.8455 - learning_rate: 0.0010
Epoch 2/200
22/22 ━━━━━━━━━━ 1s 21ms/step - loss: 16.2311 - val_loss: 14.8876 - learning_rate: 0.0010
Epoch 3/200
22/22 ━━━━━━━━━━ 1s 21ms/step - loss: 13.6994 - val_loss: 12.6609 - learning_rate: 0.0010
Epoch 4/200
22/22 ━━━━━━━━━━ 1s 21ms/step - loss: 10.9257 - val_loss: 10.0671 - learning_rate: 0.0010
Epoch 5/200
```

Σχήμα G.7: Εκπαίδευση του Μοντέλου και Ρύθμιση Learning Rate.

Ο εκπαιδευτικός πίνακας X_{train} περιλαμβάνει τις εξωγενείς μεταβλητές που χρησιμοποιούνται για την πρόβλεψη των πωλήσεων. Κάθε γραμμή αντιπροσωπεύει μια παρατήρηση με τα χαρακτηριστικά του Order Date, Ship Mode, Product Category και άλλα, τα οποία έχουν κανονικοποιηθεί και κωδικοποιηθεί για καλύτερη απόδοση του μοντέλου.

Τα στατιστικά στοιχεία των δεδομένων εκπαίδευσης (mean, std, 25%, 50%, 75%) επιτρέπουν την κατανόηση της κατανομής των χαρακτηριστικών και τον εντοπισμό ενδεχόμενων προβλημάτων, όπως ανισομετρία ή ακραίες τιμές. Η εξαρτημένη μεταβλητή y_{train} περιλαμβάνει τις τιμές των πωλήσεων που χρησιμοποιούνται για την εκπαίδευση του μοντέλου, ενώ η κατανομή τους εμφανίζεται στην παρακάτω εικόνα:

Αφού ολοκληρωθεί η εκπαίδευση, το μοντέλο είναι έτοιμο για αξιολόγηση στα δεδομένα

| | X_train | | | | | | | | | | | | | Pyth | | |
|------|-----------------|------------------|----------------|----------------|-----------------|---------------|----------------------|--------------------|------------------------|--------------------------|------------------------|-------------|----|------|----|----|
| | 0s | | | | | | | | | | | | | | | |
| | Order Date_year | Order Date_month | Order Date_day | Ship Date_year | Ship Date_month | Ship Date_day | Ship Mode_Fast Class | Ship Mode_Same Day | Ship Mode_Second Class | Ship Mode_Standard Class | Sub-Category_Envelopes | Category_Fa | | | | |
| 0 | 1.153020 | 0.551120 | 0.061954 | 1.138089 | 0.522895 | 0.348732 | 0.429885 | -0.249166 | 0.484918 | 0.824704 | - | 0.162742 | - | | | |
| 1 | -0.676744 | 0.665107 | -0.290535 | -0.389970 | 0.674488 | 0.191900 | -0.429885 | -0.249166 | -0.484918 | 0.824704 | - | -0.162742 | - | | | |
| 2 | 1.153020 | 0.246913 | 1.652236 | 1.138089 | 0.223459 | 0.501046 | 0.429885 | 0.249166 | 0.484918 | 0.824704 | - | 0.162742 | - | | | |
| 3 | -0.676744 | 0.964914 | 0.509494 | -0.638620 | 0.944094 | 1.040424 | -0.429885 | -0.249166 | -0.484918 | 0.824704 | - | -0.162742 | - | | | |
| 4 | 1.153020 | 0.361501 | 1.205557 | 1.138089 | 0.375252 | 0.576700 | 0.429885 | 0.249166 | 0.484918 | 0.824704 | - | 0.162742 | - | | | |
| .. | .. | .. | .. | .. | .. | .. | .. | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| 5855 | 1.153020 | 0.665707 | 0.852528 | 1.138089 | 0.671688 | 1.484894 | 0.429885 | 0.249166 | 0.484918 | 0.824704 | - | 0.162742 | - | | | |
| 6057 | 0.263188 | -1.161947 | -0.176164 | 0.249584 | -1.204049 | -0.105733 | 1.121502 | -0.249166 | -0.484918 | -1.212556 | - | -0.162742 | - | | | |
| 6058 | 0.263188 | -1.159534 | 0.662126 | 0.249584 | 1.121687 | 0.446581 | 0.429885 | -0.249166 | 0.484918 | 0.824704 | - | 0.162742 | - | | | |
| 6059 | -1.515475 | 0.964914 | 1.424390 | -1.527625 | 1.271480 | -1.646309 | -0.429885 | -0.249166 | 2.062203 | -1.212556 | - | -0.162742 | - | | | |
| 6859 | 1.153020 | -1.159534 | 0.976437 | 1.138089 | -1.121687 | 0.446581 | 0.429885 | 0.249166 | 0.484918 | 0.824704 | - | 0.162742 | - | | | |

Σχήμα G.8: Εκπαιδευτικά Δεδομένα X_{train} Η εικόνα δείχνει τα κανονικοποιημένα δεδομένα εκπαίδευσης.

| y_train | | |
|---|--|----------|
| ✓ 0.0s | | |
| 8788 | | 6.550154 |
| 847 | | 5.666219 |
| 3397 | | 3.292870 |
| 6244 | | 4.787825 |
| 8198 | | 2.806507 |
| ... | | |
| 2895 | | 3.607669 |
| 7813 | | 2.447551 |
| 905 | | 2.636196 |
| 5192 | | 5.987958 |
| 235 | | 6.428066 |
| Name: Sales, Length: 6860, dtype: float64 | | |

Σχήμα G.9: Εκπαιδευτικά Δεδομένα (X_{train}) Η εικόνα δείχνει τα κανονικοποιημένα δεδομένα εκπαίδευσης.

δοκιμής (test set). Μετρικές απόδοσης όπως R^2 , MAE και RMSE χρησιμοποιούνται για την εκτίμηση της ικανότητας του μοντέλου να προβλέψει τις πωλήσεις σε νέα, αιθέατα δεδομένα.

Κεφάλαιο Η

Αξιολόγηση και Αποτελέσματα του Μοντέλου

H.1 Αξιολόγηση του Μοντέλου με Μετρικές Απόδοσης

Η αξιολόγηση του μοντέλου περιλαμβάνει τον υπολογισμό του Test Loss και άλλων μετρικών, όπως το R^2 Score, για να διαπιστώσουμε πόσο καλά το μοντέλο τα κατάφερε με τα δεδομένα που δεν έχει δει προηγουμένως.

H.1.1 1. Test Loss (Απώλεια Δοκιμής)

Η Test Loss δείχνει πόσο καλά το μοντέλο γενικεύει τα αποτελέσματά του σε δεδομένα που δεν έχει δει πριν (δοκιμαστικό σύνολο δεδομένων). Μικρή Test Loss σημαίνει ότι το μοντέλο γενικά τα πήγε καλά στην πρόβλεψη των πωλήσεων. Αν η απώλεια είναι υψηλή, αυτό μπορεί να υποδηλώνει υπερπροσαρμογή (overfitting) ή ότι το μοντέλο δεν είναι αρκετά ισχυρό για να κατανοήσει τα μοτίβα των δεδομένων.

Όπως φαίνεται στην παρακάτω εικόνα, η Test Loss του μοντέλου είναι 0.925468.

H.1.2 2. R^2 Score (Σκορ R^2)

Το R^2 Score ή Coefficient of Determination είναι μια άλλη κρίσιμη μετρική για την αξιολόγηση της απόδοσης του μοντέλου. Το R^2 μετράει πόσο καλά το μοντέλο κατάφερε να

```

RESULTS

# Evaluate the model on the test set
test_loss = model.evaluate(X_test, y_test, verbose=0)
print(f'Test Loss: {test_loss:.6f}')
✓ 0.4s
Test Loss: 0.925468

```

Σχήμα H.10: Test Loss για την αξιολόγηση του μοντέλου.

εξηγήσει τη διακύμανση των δεδομένων. Αν το R^2 είναι 1, το μοντέλο εξηγεί πλήρως τα δεδομένα, ενώ αν είναι 0, το μοντέλο δεν μπορεί να εξηγήσει καθόλου τα δεδομένα.

Στην παρακάτω εικόνα, το R^2 Score του μοντέλου είναι 0.63000, που σημαίνει ότι το μοντέλο καταφέρνει να εξηγήσει περίπου το 63% της διακύμανσης των δεδομένων στις πωλήσεις.

```

# Predict and calculate R^2 score
y_pred = np.squeeze(model.predict(X_test))
test_r2 = r2_score(y_test, y_pred)

print(f'R^2 Score: {test_r2:.5f}')
✓ 0.8s
92/92 ━━━━━━━━━━ 1s 5ms/step
R^2 Score: 0.63000

```

Σχήμα H.11: R^2 Score

H.1.3 3. Συνολική Αξιολόγηση με Μετρικές Απόδοσης

Με βάση τις μετρικές απόδοσης, μπορούμε να κατανοήσουμε την ποιότητα του μοντέλου μας:

- **Test Loss:** Ένα σημαντικό μέτρο για το πόσο καλά το μοντέλο κάνει τις προβλέψεις του. Στην περίπτωσή μας, το Test Loss είναι 0.925468.
- **R^2 Score:** Αυτή η μετρική δείχνει την ποσοτική εξήγηση του μοντέλου για τις παρατηρούμενες πωλήσεις. Το R^2 Score είναι 0.63000, δηλαδή το μοντέλο καταφέρνει να εξηγήσει το 63% της διακύμανσης στις πωλήσεις.

Αυτές οι μετρικές δείχνουν ότι το μοντέλο έχει ικανοποιητική απόδοση, αλλά υπάρχει περιθώριο βελτίωσης, το οποίο μπορεί να επιτευχθεί είτε με τη χρήση άλλων τεχνικών μηχανικής μάθησης (π.χ. XGBoost), είτε με περισσότερη βελτιστοποίηση του μοντέλου (ρυθμίσεις υπερπαραμέτρων, περισσότερα δεδομένα εκπαίδευσης κ.λπ.).

H.2 Ανάλυση των Αποτελεσμάτων του Μοντέλου Πρόβλεψης Πωλήσεων

Η διαδικασία ανάπτυξης και αξιολόγησης του μοντέλου πρόβλεψης πωλήσεων περιλαμβανε τα εξής βήματα:

- **Προεπεξεργασία Δεδομένων:** Καθαρίστηκαν τα δεδομένα, εφαρμόστηκε One-Hot Encoding στις κατηγορικές στήλες, και μετατράπηκαν οι ημερομηνίες σε χρήσιμα χαρακτηριστικά.
- **Εκπαίδευση του Μοντέλου:** Το μοντέλο εκπαίδευτηκε χρησιμοποιώντας τεχνικές Backpropagation και Adam optimizer, με ρυθμό μάθησης 0.001 και 200 εποχές.
- **Αξιολόγηση του Μοντέλου:** Χρησιμοποιήθηκαν οι μετρικές Test Loss και R² Score για να αξιολογηθεί η απόδοση του μοντέλου. Η απόδοση του μοντέλου επιβεβαιώθηκε με τη χρήση πραγματικών και προβλεπόμενων πωλήσεων σε διάφορους μήνες.

1. Test Loss και R² Score

Test Loss: Η Test Loss του μοντέλου είναι **0.925468**, η οποία δείχνει την απόκλιση του μοντέλου από τις πραγματικές πωλήσεις στο δοκιμαστικό σύνολο. Μια μικρή τιμή Test Loss υποδεικνύει ικανοποιητική απόδοση.

R² Score: Η τιμή του R² Score είναι **0.63000**, υποδεικνύοντας ότι το μοντέλο εξηγεί το 63% της διακύμανσης των πωλήσεων. Αν και το μοντέλο έχει ικανοποιητική απόδοση, υπάρχει περιθώριο για βελτίωση.

2. Actual vs Predicted Sales (Πραγματικές vs Προβλεπόμενες Πωλήσεις)

Η παρακάτω εικόνα συγχρίνει τις πραγματικές και προβλεπόμενες πωλήσεις του μοντέλου. Ο άξονας x αντιπροσωπεύει τις πραγματικές πωλήσεις, ενώ ο άξονας y τις προβλεπόμενες πωλήσεις από το μοντέλο. Η κόκκινη διακεκομένη γραμμή δείχνει την ιδανική περίπτωση όπου οι προβλέψεις αντιστοιχούν ακριβώς στις πραγματικές πωλήσεις.



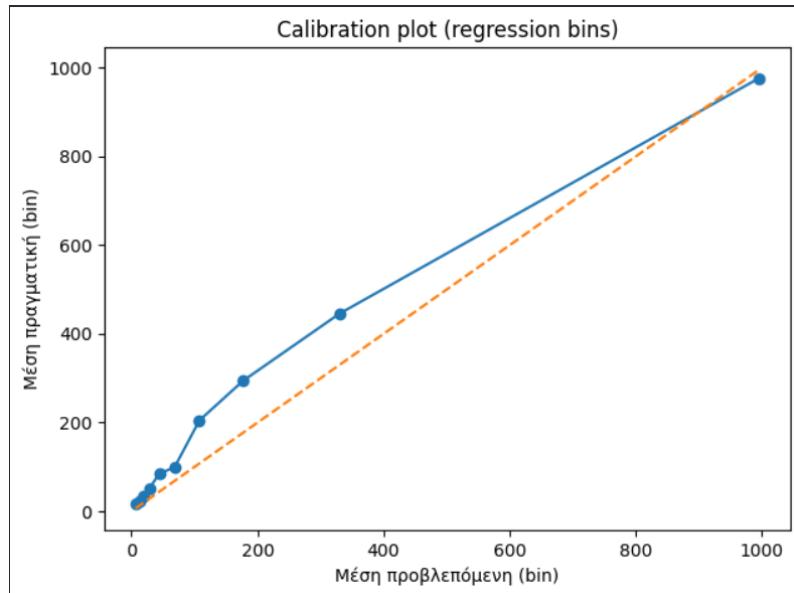
Σχήμα H.12: Actual vs Predicted Sales

3. Calibration Plot (Βαθμονομημένο Διάγραμμα)

Η παρακάτω εικόνα δείχνει το βαθμονομημένο διάγραμμα (calibration plot), το οποίο συγχρίνει τις προβλεπόμενες πωλήσεις με τις πραγματικές πωλήσεις. Το διάγραμμα αυτό βοηθά στην αξιολόγηση της ακρίβειας των προβλέψεων του μοντέλου. Όσο πιο κοντά είναι η γραμμή στη διαγώνιο (δηλαδή την ιδανική γραμμή), τόσο πιο ακριβείς είναι οι προβλέψεις.

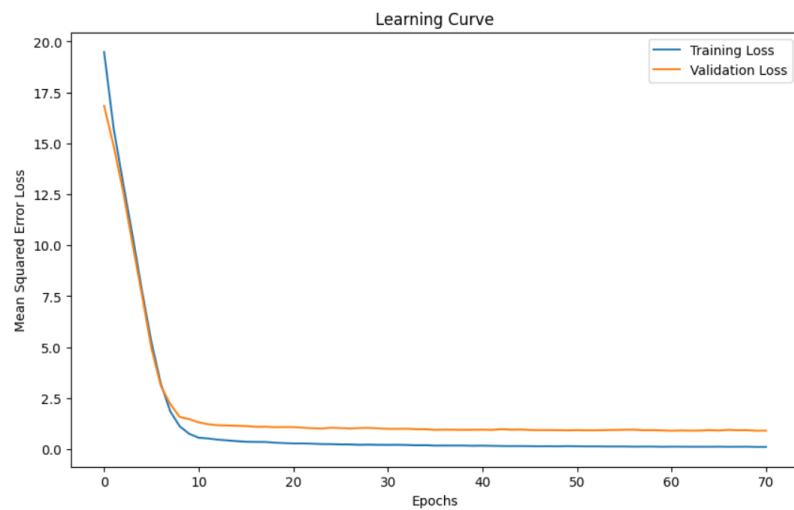
4. Learning Curve (Καμπύλη Εκπαίδευσης)

Η Learning Curve δείχνει την πρόοδο της εκπαίδευσης του μοντέλου. Στην παρακάτω εικόνα, βλέπουμε τις καμπύλες *Training Loss* και *Validation Loss* κατά τη διάρκεια των 70 εποχών εκπαίδευσης. Η *Training Loss* μειώνεται σημαντικά στην αρχή, δείχνοντας ότι το μοντέλο



Σχήμα H.13: Βαθμονομημένη σύγκριση των προβλέψεων σε διάφορα bins

μαθαίνει από τα δεδομένα, ενώ η *Validation Loss* σταθεροποιείται, αποδεικνύοντας ότι το μοντέλο είναι αρκετά καλό και στα δεδομένα που δεν έχει δει.



Σχήμα H.14: Learning Curve: Δείχνει την καμπύλη εκπαίδευσης με τη σύγκριση της Training Loss και της Validation Loss

5. Αξιολόγηση Απόδοσης με R² Score

Η παρακάτω εικόνα δείχνει την απόδοση του μοντέλου με βάση το R² Score. Το γράφημα δείχνει ότι το μοντέλο έχει ικανοποιητική απόδοση ($R^2 \approx 0.63$). Αυτό σημαίνει ότι το μοντέλο μπορεί να εξηγήσει περίπου το 63% της διακύμανσης στις πωλήσεις, χάτι που είναι καλό, αλλά υπάρχει περιθώριο βελτίωσης για το μοντέλο.



Σχήμα Η.15: Απόδοση του μοντέλου με βάση το R² Score

6. Μετρικά του Μοντέλου

- Σύνολο Παραμέτρων: 837,377
- Μέγεθος Αρχείου Μοντέλου: 9.63 MB
- Χρόνος Πρόβλεψης για 1024 Δείγματα: 0.4610s (περίπου 0.45 ms/δείγμα)

Αυτά τα στατιστικά μας δίνουν μια εικόνα για το μέγεθος και την ταχύτητα του μοντέλου, δείχνοντας ότι το μοντέλο είναι αρκετά ελαφρύ και αποδοτικό σε επίπεδο πρόβλεψης.

Σύνολο παραμέτρων: 837,377
 Μέγεθος αρχείου μοντέλου: 9.63 MB
 Χρόνος πρόβλεψης για 1024 δείγματα: 0.4610s (~0.450 ms/δείγμα)

Σχήμα H.16: Στατιστικά του μοντέλου, όπως το συνολικό πλήθος παραμέτρων και το μέγεθος του αρχείου του μοντέλου

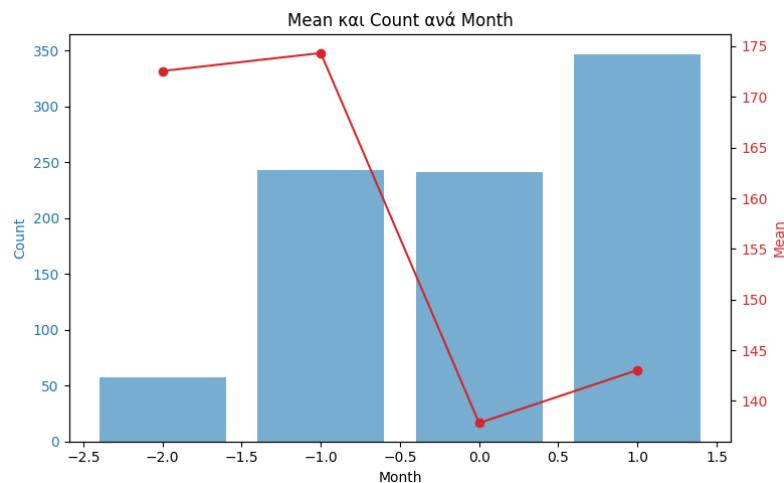
7. Στατιστικά της Στήλης “Month”

Η ανάλυση της στήλης month μας παρέχει χρήσιμα στατιστικά για τις πωλήσεις κατά τις διάφορες χρονικές περιόδους, δείχνοντας πώς οι μήνες πριν και μετά την παραγγελία επηρεάζουν τις πωλήσεις.

Πίνακας H.1: Στατιστικά της Στήλης “Month”

| Month | Mean | Median | Count |
|-------|------------|-----------|-------|
| -2.0 | 172.552832 | 42.456679 | 57 |
| -1.0 | 174.326536 | 28.941210 | 243 |
| 0.0 | 137.824516 | 29.082532 | 241 |
| 1.0 | 143.039369 | 33.983384 | 347 |

Αυτή η ανάλυση δείχνει τις σημαντικές διακυμάνσεις στις πωλήσεις ανάλογα με τον μήνα.



Σχήμα H.17: Στατιστικά της στήλης ”Month” για την ανάλυση των πωλήσεων σε σχέση με τους μήνες

H.3 Συμπεράσματα

Η συνολική ανάλυση δείχνει ότι το προτεινόμενο μοντέλο πρόβλεψης πωλήσεων πέτυχε ικανοποιητική απόδοση, με **Test Loss = 0.925468** και **R² Score = 0.63000**. Αυτό σημαίνει ότι το μοντέλο μπορεί να εξηγήσει περίπου το 63% της διακύμανσης στις πωλήσεις, προσφέροντας μια αξιόπιστη βάση για προβλέψεις σε επιχειρησιακά δεδομένα. Αν και το ποσοστό αυτό δεν είναι άριστο, επιβεβαιώνει τη χρησιμότητα της προτεινόμενης μεθοδολογίας σε ένα πεδίο όπου η πλήρης ακρίβεια είναι δύσκολα επιτεύξιμη λόγω πολυπλοκότητας και αβεβαιότητας στα δεδομένα.

Η ανάλυση των πραγματικών και προβλεπόμενων τιμών, σε συνδυασμό με τα διαγράμματα καλιμπραρίσματος, έδειξε ότι το μοντέλο τείνει να ακολουθεί ικανοποιητικά τις γενικές τάσεις των πωλήσεων, αν και παρουσιάζει αποκλίσεις σε ορισμένες περιόδους με έντονη εποχικότητα ή απότομες διακυμάνσεις. Η καμπύλη εκπαίδευσης (Learning Curve) κατέδειξε ότι η εκπαίδευση ήταν σταθερή, χωρίς υπερεφαρμογή (overfitting), γεγονός που ενισχύει την αξιοπιστία των αποτελεσμάτων. Παράλληλα, η αξιολόγηση της στήλης month ανέδειξε την ύπαρξη εποχικότητας, στοιχείο που επιβεβαιώνει την ανάγκη ενσωμάτωσης χρονικών χαρακτηριστικών στη διαδικασία πρόβλεψης.

Από άποψη απόδοσης, το μοντέλο διαθέτει σχετικά μικρό αριθμό παραμέτρων και μικρό μέγεθος αρχείου, γεγονός που το καθιστά ελαφρύ και αποδοτικό σε επίπεδο πρόβλεψης. Ο χρόνος εκτέλεσης για μεγάλα δείγματα είναι χαμηλός, επιτρέποντας την πρακτική αξιοποίησή του σε πραγματικά επιχειρησιακά σενάρια. Αυτό το χαρακτηριστικό το καθιστά κατάλληλο για εφαρμογές όπου απαιτούνται γρήγορες προβλέψεις, όπως η διαχείριση αποθεμάτων ή η βραχυπρόθεσμη εκτίμηση ζήτησης.

Παρά τα θετικά αυτά στοιχεία, υπάρχουν και ορισμένοι περιορισμοί. Η τιμή του R² δείχνει ότι παραμένει ένα ποσοστό διακύμανσης ανεξήγητο από το μοντέλο, κάτι που υποδεικνύει περιθώρια βελτίωσης. Επιπλέον, η εποχικότητα και οι εξωγενείς παράγοντες (π.χ. οικονομικές συγκυρίες, προσφορές, αλλαγές στη ζήτηση) δεν έχουν αποτυπωθεί πλήρως. Η αξιοποίηση πιο σύνθετων αρχιτεκτονικών, όπως LSTM ή GRU για χρονοσειρές, ή ακόμα και τεχνικών ensemble που συνδυάζουν διαφορετικά μοντέλα, θα μπορούσε να βελτιώσει την ακρίβεια. Τέλος, η ενσωμάτωση πρόσθετων χαρακτηριστικών, όπως εξωτερικοί δείκτες αγοράς ή δημογραφικά δεδομένα, θα μπορούσε να εμπλουτίσει τις προβλέψεις.

Συνοψίζοντας, η εργασία έδειξε ότι τα μοντέλα μηχανικής μάθησης μπορούν να προσφέρουν αξιόπιστες και πρακτικά χρήσιμες λύσεις για την πρόβλεψη πωλήσεων. Τα αποτελέσματα

καταδεικνύουν τη σημασία της σωστής προεπεξεργασίας δεδομένων, της επιλογής κατάλληλων μετρικών και της ανάλυσης εποχικότητας. Η συμβολή της παρούσας μελέτης βρίσκεται τόσο στην εφαρμογή των μεθοδολογιών όσο και στην ανάδειξη των δυνατοτήτων και των ορίων τους, προσφέροντας ένα σταθερό υπόβαθρο για περαιτέρω έρευνα και ανάπτυξη πιο εξελιγμένων μοντέλων πρόβλεψης.

Πίνακας Η.2: Συνοπτικές Μετρικές Απόδοσης του Μοντέλου Πρόβλεψης Πωλήσεων

| Μετρική | Αποτέλεσμα |
|------------------------------------|------------|
| Test Loss | 0.925468 |
| R ² Score | 0.63000 |
| Σύνολο Παραμέτρων | 837,377 |
| Μέγεθος Αρχείου Μοντέλου | 9.63 MB |
| Χρόνος Πρόβλεψης για 1024 Δείγματα | 0.4610 s |

Βιβλιογραφία

- Alkhalifah, e. a., A. (2023). Sales forecasting using machine learning: A comprehensive review. *International Journal of Forecasting*, 39(1), 1–15.
- Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 8(1), 69–80.
- Assimakopoulos, V., & Nikolopoulos, K. (2000). The theta model: A decomposition approach to forecasting. *International Journal of Forecasting*, 16(4), 521–530.
- Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation [Information Sciences, 191, 192–213].
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Breiman, L. (2001). Random forests [Machine Learning, 45(1), 5–32].
- Chan, e. a., Y. (2025). Advanced neural networks for sales prediction in retail [Retail Technology Review, 8(2), 45–59].
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD Conference* (pp. 785–794), 785–794.
- Chollet, F. (2015). Keras [computer software].
- Croston, J. D. (1972). Forecasting and stock control for intermittent demands [Operational Research Quarterly, 23(3), 289–303].
- Flores, B. E. (1986). A pragmatic view of accuracy measurement in forecasting.
- Haykin, S. (2009). *Neural networks and learning machines* (3rd ed. Pearson).
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory [Neural Computation, 9(8), 1735–1780].
- Holt, C. C. (1957). Forecasting seasonal and trends by exponentially weighted moving averages [O.N.R. Research Memorandum].

- Hunter, J. D. (2007). Matplotlib: A 2d graphics environment [Computing in Science Engineering, 9(3), 90–95].
- Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed []). OTexts].
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688.
- Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*.
- Lim, B. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748–1764.
- McKinney, W. (2010). Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference (SciPy 2010)*, 51–56, 51–56.
- Padmaja, e. a., S. (2023). A systematic review of sales forecasting methods. *International Journal of Business Forecasting*, 12(1), 1–20.
- Papitha, e. a., M. (2024). Energy demand forecasting using deep learning models [Energy Systems, 10(3), 123–135].
- Priscila, e. a., L. (2023). Deep learning models for sales forecasting: A comparative study. *Journal of Retail Analytics*, 5(4), 22–34.
- Project Jupyter. (n.d.). Project jupyter [Retrieved October 3, 2025].
- Report, W. M. (2023). Machine learning for sales prediction.
- Singh, A., & Aggarwal, R. (2020). The role of predictive analytics in inventory management. *Journal of Business Analytics*, 3(2), 1–10.
- Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting*, 36(1), 75–85.
- Syntetos, A. A., & Boylan, J. E. (2005). The accuracy of intermittent demand estimates. *International Journal of Forecasting*, 21(2), 303–314.
- Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual*. CreateSpace.
- Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages [Management Science, 6(3), 324–342].

Συντμήσεις

| | |
|----------|--|
| AI | Artificial Intelligence – Τεχνητή Νοημοσύνη |
| ML | Machine Learning – Μηχανική Μάθηση |
| DL | Deep Learning – Βαθιά Μάθηση |
| ANN | Artificial Neural Network – Τεχνητό Νευρωνικό Δίκτυο |
| CNN | Convolutional Neural Network – Συνελικτικό Νευρωνικό Δίκτυο |
| RNN | Recurrent Neural Network – Αναδρομικό Νευρωνικό Δίκτυο |
| LSTM | Long Short-Term Memory – Νευρωνικό Δίκτυο Μακράς και Βραχείας Μνήμης |
| GRU | Gated Recurrent Unit – Αναδρομική Μονάδα με Πύλες |
| ARIMA | Autoregressive Integrated Moving Average – Αυτοπαλινδρομο Ολοκληρωμένο Κινούμενο Μέσου |
| SARIMA | Seasonal ARIMA – Εποχικό ARIMA |
| SVR | Support Vector Regression – Παλινδρόμηση με Υποστηρικτικά Διανύσματα |
| RF | Random Forest – Τυχαίο Δάσος |
| XGBoost | Extreme Gradient Boosting – Ακραία Ενίσχυση Κλίσεων |
| CatBoost | Categorical Boosting – Ενίσχυση Κατηγοριών Δεδομένων |
| MSE | Mean Squared Error – Μέσο Τετραγωνικό Σφάλμα |
| MAE | Mean Absolute Error – Μέσο Απόλυτο Σφάλμα |

| | |
|----------------|---|
| RMSE | Root Mean Squared Error – Τετραγωνική Πίζα Μέσου Τετραγωνικού Σφάλματος |
| R ² | Coefficient of Determination – Συντελεστής Προσδιορισμού |
| API | Application Programming Interface – Διεπαφή Προγραμματισμού Εφαρμογών |
| CSV | Comma-Separated Values – Αρχείο Διαχωρισμένο με Κόμματα |
| JSON | JavaScript Object Notation – Μορφή Ανταλλαγής Δεδομένων |
| SQL | Structured Query Language – Δομημένη Γλώσσα Ερωτημάτων |
| NoSQL | Not Only SQL – Βάσεις Δεδομένων χωρίς αυστηρή δομή SQL |
| DB | Database – Βάση Δεδομένων |
| GPU | Graphics Processing Unit – Μονάδα Επεξεργασίας Γραφικών |
| CPU | Central Processing Unit – Κεντρική Μονάδα Επεξεργασίας |
| AI | Artificial Intelligence – Τεχνητή Νοημοσύνη |
| ML | Machine Learning – Μηχανική Μάθηση |
| DL | Deep Learning – Βαθιά Μάθηση |
| MLP | Multi-Layer Perceptron – Πολυεπίπεδος Αντιληπτικός Νευρώνας |
| MSE | Mean Squared Error – Μέσο Τετραγωνικό Σφάλμα |
| MAE | Mean Absolute Error – Μέσο Απόλυτο Σφάλμα |

| | |
|----------------|--|
| RMSE | Root Mean Squared Error – Τετραγωνική Πίζα Μέσου Τετραγωνικού Σφάλματος |
| R ² | Coefficient of Determination – Συντελεστής Προσδιορισμού |
| ReLU | Rectified Linear Unit – Συνάρτηση Ενεργοποίησης ReLU |
| CSV | Comma-Separated Values – Μορφή Αρχείου CSV |

Γλωσσάρι Εντικών θρων

| | |
|---|---|
| Accuracy (Ακρίβεια) | Μετρική που δείχνει το ποσοστό σωστών προβλέψεων σε σχέση με το σύνολο των δεδομένων. |
| Precision (Ακρίβεια θετικών προβλέψεων) | Μετρική που δείχνει το ποσοστό των σωστών θετικών προβλέψεων. |
| Recall (Ανάκληση) | Μετρική που δείχνει το ποσοστό των πραγματικών θετικών που αναγνωρίστηκαν σωστά. |
| F1-Score | Ο αρμονικός μέσος του Precision και του Recall, χρήσιμος όταν υπάρχει ανισορροπία στις κατηγορίες. |
| Epoch | Ένας πλήρης κύκλος εκπαίδευσης του μοντέλου πάνω σε όλα τα δεδομένα εκπαίδευσης. |
| Batch | Τυποσύνολο του συνόλου δεδομένων που χρησιμοποιείται σε μία επανάληψη εκπαίδευσης. |
| Overfitting (Υπερεφαρμογή) | Όταν το μοντέλο μαθαίνει υπερβολικά καλά τα δεδομένα εκπαίδευσης και αποτυγχάνει να γενικεύσει σε νέα. |
| Underfitting (Υπεκπαίδευση) | Όταν το μοντέλο αποτυγχάνει να μάθει τις σχέσεις των δεδομένων, οδηγώντας σε χαμηλή απόδοση. |
| Dropout | Τεχνική regularization που απενεργοποιεί τυχαία νευρώνες κατά την εκπαίδευση για να μειώσει την υπερεφαρμογή. |
| Regularization (Κανονικοποίηση Μοντέλου) | Μέθοδοι που μειώνουν την πολυπλοκότητα του μοντέλου για να αποφευχθεί η υπερεφαρμογή. |
| Activation Function (Συνάρτηση Ενεργοποίησης) | Μαθηματική συνάρτηση που εισάγει μη-γραμμικότητα στο δίκτυο. |

| | |
|--|---|
| Loss Function (Συνάρτηση Κόστους) | Συνάρτηση που μετράει το σφάλμα των προβλέψεων και καθιστάει τη διαδικασία εκπαίδευσης. |
| Optimizer (Βελτιστοποιητής) | Αλγόριθμος που προσαρμόζει τα βάρη του μοντέλου ώστε να ελαχιστοποιήσει τη συνάρτηση κόστους. |
| Gradient Descent (Κατιούσα Κλίση) | Μέθοδος βελτιστοποίησης που ενημερώνει τα βάρη του μοντέλου με βάση την κλίση της συνάρτησης κόστους. |
| Hyperparameter (Τυπερπαράμετρος) | Παράμετρος που ορίζεται πριν την εκπαίδευση και ρυθμίζει τη συμπεριφορά του αλγορίθμου. |
| Cross-Validation (Διασταυρούμενη Επικύρωση) | Μέθοδος αξιολόγησης που χωρίζει τα δεδομένα σε πολλά τμήματα για πιο αξιόπιστη μέτρηση της απόδοσης. |
| Training Set | Το σύνολο δεδομένων που χρησιμοποιείται για την εκπαίδευση του μοντέλου. |
| Validation Set | Το σύνολο δεδομένων που χρησιμοποιείται για ρύθμιση υπερπαραμέτρων και αποφυγή υπερεφαρμογής. |
| Test Set | Το σύνολο δεδομένων που χρησιμοποιείται για την τελική αξιολόγηση του μοντέλου. |
| Normalization (Κανονικοποίηση) | Μετασχηματισμός των δεδομένων σε κοινή κλίμακα για βελτίωση της εκπαίδευσης. |
| Standardization (Τυποποίηση) | Μετατροπή δεδομένων ώστε να έχουν μέση τιμή 0 και τυπική απόκλιση 1. |
| One-Hot Encoding | Τεχνική μετατροπής κατηγοριακών μεταβλητών σε δυαδικούς διανύσματα. |
| Feature (Χαρακτηριστικό) | Μεταβλητή εισόδου που χρησιμοποιείται από το μοντέλο για πρόβλεψη. |
| Label (Ετικέτα) | Η τιμή-στόχος που το μοντέλο καλείται να προβλέψει. |
| Stationarity (Στασιμότητα) | Ιδιότητα μιας χρονοσειράς όπου οι στατιστικές της δεν αλλάζουν με τον χρόνο. |
| Seasonality (Εποχικότητα) | Περιοδικές διακυμάνσεις μιας χρονοσειράς που επαναλαμβάνονται σε τακτά διαστήματα. |
| Sliding Window (Κινούμενο Παράθυρο) | Μέθοδος δειγματοληψίας δεδομένων από χρονοσειρά με σταθερό μήκος παραθύρου. |

| | |
|--|--|
| Forecast Horizon (Οριζόντας Πρόβλεψης) | Ο αριθμός μελλοντικών χρονικών σημείων που το μοντέλο καλείται να προβλέψει. |
| Pandas | Βιβλιοθήκη Python για ανάλυση και διαχείριση δεδομένων. |
| NumPy | Βιβλιοθήκη Python για αριθμητικές υπολογισμούς και πράξεις σε πίνακες. |
| Scikit-learn | Βιβλιοθήκη Python για μηχανική μάθηση. |
| TensorFlow | Βιβλιοθήκη Python για deep learning. |
| PyTorch | Βιβλιοθήκη Python για deep learning. |