



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ
ΔΙΑΤΑΞΕΩΝ & ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

«Αξιολόγηση τεχνικών ενσωμάτωσης εξωγενών
μεταβλητών σε νευρωνικά δίκτυα πρόβλεψης
φορτίου επόμενης ημέρας»

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Κρεμάλας Δ. Ευάγγελος

Επιβλέπων : Ασημακόπουλος Βασίλειος, Καθηγητής Ε.Μ.Π.

Υπεύθυνος : Σεμένογλου Αρτέμης, Υποψήφιος Διδάκτωρ Ε.Μ.Π.

Αθήνα, Ιούνιος 2025



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ
ΔΙΑΤΑΞΕΩΝ & ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

«Αξιολόγηση τεχνικών ενσωμάτωσης εξωγενών
μεταβλητών σε νευρωνικά δίκτυα πρόβλεψης
φορτίου επόμενης ημέρας»

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Κρεμάλας Δ. Ευάγγελος

Επιβλέπων : Ασημακόπουλος Βασίλειος, Καθηγητής Ε.Μ.Π.

Υπεύθυνος : Σεμένογλου Αρτέμης, Υποψήφιος Διδάκτωρ Ε.Μ.Π.

Εγκρίθηκε από την τριμελή επιτροπή την 03/07/2025

.....
Βασίλειος Ασημακόπουλος
Ομότιμος Καθηγητής ΕΜΠ

.....
Ιωάννης Ψαρράς
Καθηγητής ΕΜΠ

.....
Δημήτριος Ασκούνης
Καθηγητής ΕΜΠ

Αθήνα, Ιούνιος 2025



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ
ΔΙΑΤΑΞΕΩΝ & ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

.....
Κρεμάλας Δ. Ευάγγελος

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Ηλεκτρονικών
Υπολογιστών

Copyright © Κρεμάλας Ευάγγελος, 2025.

Με την επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τους συγγραφείς.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου, του Επιβλέποντα, ή της επιτροπής που την ενέκριν

Περίληψη

Η παρούσα διπλωματική εργασία εστιάζει στη σημασία των εξωγενών μεταβλητών για την πρόβλεψη της ροής φορτίου της επόμενης ημέρας, με κύριο στόχο την εύρεση της βέλτιστης αρχιτεκτονικής για την ενσωμάτωσή τους σε πλήρως διασυνδεδεμένα νευρωνικά δίκτυα, όπως αναπτύχθηκε στην έρευνα «Short-term load forecasting with global models: A comparative analysis of neural network architectures» των Αρτέμιο-Ανάργυρου Σεμένογλου, Ευάγγελου Σπηλιώτη και Βασιλείου Ασημακόπουλου, η οποία παρουσιάστηκε στο 43rd International Symposium on Forecasting (ISF 2023, 25-28 Ιουνίου 2023, Charlottesville, ΗΠΑ). Τελικά επιτυγχάνεται βελτίωση του μοντέλου κατά 12% ως προς το MAPE και κατά 10.5% ως προς το RMSE.

Η πρόβλεψη της ροής φορτίου είναι κρίσιμη για την αποδοτική λειτουργία του ενεργειακού συστήματος, καθώς επιτρέπει τον προγραμματισμό παραγωγής και διανομής, εξασφαλίζοντας ισορροπία προσφοράς-ζήτησης, μειώνοντας διακοπές και βελτιστοποιώντας πόρους. Υποστηρίζει την ενσωμάτωση ΑΠΕ, αντιμετωπίζοντας τη στοχαστικότητα τους, μειώνει το κόστος λειτουργίας αποτρέποντας υπερπαραγωγή ή χρήση δαπανηρών εφεδρειών και σταθεροποιεί τις τιμές ενέργειας. Επιπρόσθετα, συμβάλλει στη μείωση περιβαλλοντικών επιπτώσεων μέσω βέλτιστης χρήσης καθαρών πηγών. Οι εξωγενείς μεταβλητές (καιρικές, ημερολογιακές, χαρακτηριστικά χώρας) βελτιώνουν σημαντικά την ακρίβεια των προβλέψεων, όπως επιβεβαιώνεται από τη βιβλιογραφία και τα πειράματα της εργασίας, με μείωση μετρικών σφάλματος (MAPE, RMSE). Ωστόσο, δεν υπάρχει καθολικά αποδεκτή μέθοδος ενσωμάτωσής τους, γεγονός που αποτελεί πρόκληση και ευκαιρία για έρευνα. Η εργασία εξετάζει διαφορετικές τεχνικές ενσωμάτωσης, υπογραμμίζοντας την ανάγκη εξατομικευμένων προσεγγίσεων.

Το πρώτο κεφάλαιο εισάγει τις βασικές έννοιες του ενεργειακού τομέα σε Ελλάδα και Ευρώπη, περιγράφει το περιβάλλον μετά την απελευθέρωση των αγορών, τους κύριους παίκτες και τη σημασία της πρόβλεψης φορτίου, καθώς και παραδοσιακές μεθόδους που χρησιμοποιούν διαχειριστές στην Ευρώπη. Το δεύτερο κεφάλαιο παρουσιάζει τη θεωρία των τεχνικών προβλέψεων, ενώ το τρίτο αναλύει τις έννοιες της μηχανικής μάθησης και των νευρωνικών δικτύων. Τέλος, περιγράφεται η μεθοδολογία, η πειραματική διαδικασία και τα αποτελέσματα των πειραμάτων.

Λέξεις κλειδιά

Πρόβλεψη φορτίου επόμενης ημέρας, Εξωγενείς μεταβλητές, Πλήρες διασυνδεδεμένο νευρωνικό δίκτυο, Νευρωνικά δίκτυα, Τεχνικές προβλέψεων, Ενεργειακές αγορές

Abstract

This diploma thesis focuses on the significance of exogenous variables for next-day load forecasting, with the primary objective of identifying the optimal architecture for their integration into fully connected neural networks, as developed in the research "Short-term load forecasting with global models: A comparative analysis of neural network architectures" by Artemios-Anargyros Semenoglou, Evangelos Spiliotis, and Vassilios Veskoukis, presented at the 43rd International Symposium on Forecasting (ISF 2023, June 25-28, 2023, Charlottesville, USA). Ultimately, it achieves a 12% and 10.5% improvement in the model's performance in terms of MAPE and RMSE.

Short-term load forecasting is critical for the efficient operation of the energy system, enabling the scheduling of production and distribution, ensuring supply-demand balance, reducing outages, and optimizing resource use. It supports the integration of Renewable Energy Sources (RES), addressing their stochastic nature, lowers operational costs by preventing overproduction or reliance on costly reserves, and stabilizes energy prices. Additionally, it contributes to reducing environmental impacts through the optimal use of clean energy sources. Exogenous variables (weather, calendar, country-specific characteristics) significantly enhance forecast accuracy, as confirmed by the literature and the experiments conducted in this thesis, with reductions in error metrics (MAPE, RMSE). However, there is no universally accepted method for their integration, posing both a challenge and an opportunity for further research. This thesis examines various integration techniques, highlighting the need for tailored approaches.

The first chapter introduces key concepts of the energy sector in Greece and Europe, describes the post-liberalization market environment, key stakeholders, and the importance of load forecasting, along with traditional methods used by European operators. The second chapter presents the theory of forecasting techniques, while the third analyzes the concepts of machine learning and neural networks. Finally, the methodology, experimental process, and results of the experiments are described.

Keywords

Short-term load forecasting, Exogenous variables, Fully connected neural network, Neural networks, Forecasting techniques, Energy markets

Ευχαριστίες

Ξεκινώντας θα ήθελα να ευχαριστήσω θερμά την ομάδα της μονάδας «Προβλέψεων και Στρατηγικής» του Μετσόβιου Πολυτεχνείου για την ευκαιρία που μου δώσανε να καταπιαστώ με τη παρούσα διπλωματική εργασία, η οποία αγγίζει θεματικές όπως οι τεχνικές προβλέψεων, η μηχανική μάθηση και η ενέργεια. Ειδικότερα θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή Ασημακόπουλο Βασίλειο, για την ανάθεση αυτής της διπλωματικής αλλά και τους καθηγητές Ψαρρά Ιωάννη και Ασκούνη Δημήτριο οι οποίοι αποτέλεσαν την εξεταστική επιτροπή αυτής της εργασίας. Ιδιαίτερα ευχαριστώ με τον πιο θερμό τρόπο τον υπεύθυνο αυτής της εργασίας τον Αρτέμη Σεμένογλου για την άριστη συνεργασία, η οποία χαρακτηρίστηκε απο όρεξη και μεράκι, καθώς ήταν πάντα διαθέσιμος να συνδράμει με τις πολύτιμες συμβουλές του.

Τέλος χρωστάω ένα μεγάλο ευχαριστώ σε όλους τους ανθρώπους που με στήριξαν, ο καθένας με το δικό του τρόπο, για να φτάσω σε αυτό το ορόσημο και να ολοκληρώσω τις σπουδές μου. Ακόμη , ευχαριστώ τους φίλους μου, με τους οποίους μοιράστηκα αμέτρητες ώρες μελέτης και υποστήριξης. Ιδιαίτερα ευχαριστώ βαθιά τους γονείς μου Δημήτρη και Δέσποινα για την αμέριστη στήριξη, τον αδερφό μου Ιωάννη και τέλος τη Μαρία που δεν σταματάει στιγμή να πιστεύει σε μένα.

Κρεμάλας Ευάγγελος,
Αθήνα, Ιούνιος 2025

Πίνακας Περιεχομένων

Περίληψη.....	5
Abstract.....	7
Ευχαριστίες.....	9
Κεφάλαιο 1 – ΕΝΕΡΓΕΙΑ.....	13
1.1 Ενιαία Ευρωπαϊκή Αγορά Ηλεκτρισμού.....	13
1.2 Σκοποί της ελεύθερης αγοράς Ηλεκτρικής ενέργειας.....	13
1.3 Συμμετέχοντες στην αγορά ηλεκτρικής ενέργειας.....	14
1.4 Δομή της αγοράς ενέργειας.....	20
1.5 Λειτουργία της αγοράς.....	28
1.6 Η αξία της πρόβλεψης φορτίου ενέργειας επόμενης ημέρας.....	29
1.7 Τεχνικές πρόβλεψης φορτίου επόμενης ημέρας στην Ευρώπη.....	31
Κεφάλαιο 2 – ΤΕΧΝΙΚΕΣ ΠΡΟΒΛΕΨΕΩΝ.....	36
2.1 Εισαγωγή.....	36
2.2 Ποιοτικές μέθοδοι.....	37
2.3 Ποσοτικές μέθοδοι.....	37
2.4 Χρονοσειρές.....	38
2.5 Παραδοσιακές μέθοδοι πρόβλεψης.....	40
2.6 Μέθοδοι αξιολόγησης προβλέψεων.....	46
Κεφάλαιο 3 – ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ.....	49
3.1 Εισαγωγή.....	49
3.2 Κατηγορίες μηχανικής μάθησης.....	50
3.3 Μοντέλα μηχανικής μάθησης.....	52
3.4 Νευρωνικά δίκτυα.....	54
3.4.1 Τεχνητός νευρώνας- Perceptron.....	54
3.4.2 Συνάρτηση ενεργοποίησης.....	56
3.4.3 Αρχιτεκτονικές νευρωνικών δικτύων - multilayer perceptrons.....	59
3.4.4 Εκπαίδευση νευρωνικού δικτύου – Επιβλεπόμενη μάθηση.....	62
3.5 Προεπεξεργασία μεταβλητών εισόδου μοντέλων για πρόβλεψη.....	69
3.6 Αρχιτεκτονικές δικτύων για πρόβλεψη φορτίου επόμενης ημέρας.....	73
3.7 Χρήση εξωγενών μεταβλητών για πρόβλεψη φορτίου επόμενης ημέρας.....	76
Κεφάλαιο 4 - ΠΕΙΡΑΜΑΤΙΚΗ ΔΙΑΔΙΚΑΣΙΑ.....	79
4.1 Στόχος.....	79
4.2 Περιγραφή της αρχιτεκτονικής του νευρωνικού δικτύου.....	80
4.3 Περιγραφή προτεινόμενης μεθοδολογίας.....	82
4.3.1 Αναπαραγωγή νευρωνικού δικτύου και αρχικών δεδομένων.....	82
4.3.2 Μορφές αναπαράστασης εξωγενών μεταβλητών.....	82
4.3.3 Συνδυασμός εξωγενών μεταβλητών.....	86
4.3.4 Αρχιτεκτονικές.....	86
Κεφάλαιο 5- ΑΠΟΤΕΛΕΣΜΑΤΑ.....	98
5.1 Περιβάλλον ανάπτυξης.....	98
5.2 Δεδομένα.....	98
5.3 Μετρικές πειραμάτων.....	100
5.4 Παρουσίαση αποτελεσμάτων αναπαράστασης αρχικών συνθηκών.....	101
5.5 Παρουσίαση αποτελεσμάτων μορφών των εξωγενών μεταβλητών.....	101
5.6 Παρουσίαση αποτελεσμάτων συνδυασμών των εξωγενών μεταβλητών.....	105
5.7 Παρουσίαση αποτελεσμάτων ενσωμάτωσης των εξωγενών μεταβλητών.....	106
Κεφάλαιο 6 – ΣΧΟΛΙΑΣΜΟΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ.....	110
6.1 Χρήση εξωγενών μεταβλητών.....	110
6.2 Αρχιτεκτονικές ενσωμάτωσης.....	112
6.3 Μελλοντικές προεκτάσεις.....	115
Βιβλιογραφία.....	118

Πίνακας Περιεχομένων Σχημάτων

Σχήμα 1.1 : Αναδρομή του Δ.ΑΠΕ.Ε.Π.....	21
Σχήμα 2.1 : Οπτικοποίηση των χαρακτηριστικών μίας χρονοσειράς.....	40
Σχήμα 2.2 : Παράδειγμα γραμμικής παλινδρόμησης.....	46
Σχήμα 3.1 : Αναπαράσταση perceptron.....	56
Σχήμα 3.2 : Παράδειγμα MLP.....	61
Σχήμα 4.1 : FCNN δίκτυο.....	82
Σχήμα 4.2 : Πείραμα 1.....	88
Σχήμα 4.3 : Πείραμα 2.....	89
Σχήμα 4.4 : Πείραμα 3.....	89
Σχήμα 4.5 : Πείραμα 4.....	90
Σχήμα 4.6 : Πείραμα 5.....	91
Σχήμα 4.7 : Πείραμα 6.....	91
Σχήμα 4.8 : Πείραμα 7.....	92
Σχήμα 4.9 : Πείραμα 8.....	92
Σχήμα 4.10 : Πείραμα 9.....	93
Σχήμα 4.11 : Πείραμα 10.....	94
Σχήμα 4.12 : Πείραμα 11.....	95
Σχήμα 4.13 : Πείραμα 12.....	96
Σχήμα 4.14 : Πείραμα 13.....	97
Σχήμα 4.15 : Πείραμα 14.....	97
Σχήμα 4.16 : Πείραμα 15.....	98
Σχήμα 5.1 : Αποτελέσματα MAPE – μορφές εξωγενών μεταβλητών.....	105
Σχήμα 5.2 : Αποτελέσματα RMSE – μορφές εξωγενών μεταβλητών.....	105
Σχήμα 5.3 : Αποτελέσματα MAPE – συνδυασμοί εξωγενών μεταβλητών.....	106
Σχήμα 5.4 : Αποτελέσματα RMSE – συνδυασμοί εξωγενών μεταβλητών.....	107
Σχήμα 5.5 : Αποτελέσματα MAPE – αρχιτεκτονικές ενσωμάτωσης εξωγενών.....	109
Σχήμα 5.6 : Αποτελέσματα RMSE – αρχιτεκτονικές ενσωμάτωσης εξωγενών.....	109

Πίνακας Περιεχομένων Πινάκων

Πίνακας 5.1 : Σφάλματα χωρίς χρήση εξωγενών μεταβλητών.....	102
Πίνακας 5.2 : Αποτελέσματα εξωγενών δεδομένων καιρού.....	103
Πίνακας 5.3 : Αποτελέσματα εξωγενών δεδομένων ημέρας.....	103
Πίνακας 5.4 : Αποτελέσματα εξωγενών δεδομένων χώρας.....	104
Πίνακας 5.5 : Αποτελέσματα συνδυασμών εξωγενών μεταβλητών..	106
Πίνακας 5.6 : Αποτελέσματα πειραμάτων ενσωμάτωσης των εξωγενών.....	108

ΚΕΦΑΛΑΙΟ 1 – ΕΝΕΡΓΕΙΑ

1.1 Ενιαία Ευρωπαϊκή Αγορά Ηλεκτρισμού

Η αξιόπιστη και προσιτή διάθεση ενέργειας, με παράλληλη ελαχιστοποίηση των περιβαλλοντικών επιπτώσεων, θεωρείται από την Ευρωπαϊκή Ένωση κρίσιμη για την εύρυθμη λειτουργία της ευρωπαϊκής οικονομίας [6]. Στο πλαίσιο αυτό, η θέσπιση μιας Ενιαίας Ευρωπαϊκής Αγοράς Ηλεκτρισμού (Internal Electricity Market) αποτέλεσε μια από τις κύριες προτεραιότητες της Ευρωπαϊκής ένωσης από τη δεκαετία του 1990 κιόλας [7]. Έτσι άρχισε να εφαρμόζεται μια πολιτική απελευθέρωσης της αγοράς ηλεκτρικής ενέργειας, η οποία μέχρι τότε ελεγχόταν σε μεγάλο βαθμό από κρατικά μονοπώλια σε όλες τις χώρες της Ένωσης [20].

Στόχοι ήταν η δημιουργία μιας ελεύθερης αγοράς ηλεκτρικής ενέργειας και ενός ευέλικτου ενεργειακού τομέα. Για την επίτευξή τους, θεσπίστηκαν ριζικές μεταρρυθμίσεις, με κεντρική αλλαγή τη διάκριση μεταξύ ανταγωνιστικών δραστηριοτήτων, όπως η παραγωγή, η προμήθεια και η εμπορία ηλεκτρικής ενέργειας, που λειτουργούν σε περιβάλλον ελεύθερης αγοράς, και μη ανταγωνιστικών δραστηριοτήτων, όπως η μεταφορά και η διανομή, που ρυθμίζονται από την πολιτεία ως φυσικά μονοπώλια [8]. Οι μεταρρυθμίσεις αυτές καθόρισαν σαφείς κανόνες για όλους τους συμμετέχοντες, περιορίζοντας τις μονοπωλιακές δομές των δικτύων και απελευθερώνοντας την αγορά προμήθειας ενέργειας.

Επιπλέον, θεσπίστηκε η υποχρέωση των διαχειριστών μη ανταγωνιστικών δραστηριοτήτων να παρέχουν πρόσβαση τρίτων στα δίκτυα τους [9]. Παράλληλα, καταργήθηκαν τα εμπόδια στην ελεύθερη δραστηριοποίηση εναλλακτικών προμηθευτών, ενισχύοντας τον ανταγωνισμό στην προμήθεια ενέργειας. Επίσης, η σταδιακή άρση των περιορισμών στην επιλογή προμηθευτή επέτρεψε στους καταναλωτές να επιλέγουν τον προμηθευτή με τις καλύτερες συνθήκες. Τέλος, ιδρύθηκαν ανεξάρτητες ρυθμιστικές αρχές για την εποπτεία της αγοράς, εξασφαλίζοντας τη διαφάνεια των συναλλαγών και ίσες ευκαιρίες για όλους τους συμμετέχοντες [14].

1.2 Σκοποί της ελεύθερης αγοράς Ηλεκτρικής ενέργειας

Σήμερα κάθε αγορά ηλεκτρικής ενέργειας έχει δύο κεντρικούς σκοπούς. Την διασφάλιση της φυσικής λειτουργίας του συστήματος για την αξιόπιστη παροχή ενέργειας στους καταναλωτές και την προστασία της εύρυθμης οικονομικής λειτουργίας της αγοράς [11]. Η ομαλή λειτουργία της αγοράς προϋποθέτει την ισορροπία μεταξύ προσφοράς και ζήτησης, με την κάλυψη της συνολικής ζήτησης ενέργειας από την προσφορά των παραγωγών κάθε χρονική στιγμή. Σε περιπτώσεις αιφνίδιας υψηλής ζήτησης ή διαταραχών στην παραγωγή, όπως απρόσμενες βλάβες σε μονάδες παραγωγής, οι ρυθμιστικές αρχές εξασφαλίζουν

αδιάκοπη παροχή ενέργειας μέσω μηχανισμών εφεδρείας, οι οποίοι παραμένουν σε ετοιμότητα [23]. Παράλληλα, βλάβες στο δίκτυο αντιμετωπίζονται μέσω διασυνδέσεων που επιτρέπουν τη διοχέτευση ενέργειας μέσω εναλλακτικών γραμμών και αυτοματοποιημένων συστημάτων αποκατάστασης ροής. Η αυξημένη χρήση ΑΠΕ δημιούργησε προκλήσεις λόγω της διαλείπουσας φύσης τους, καθώς η υπερπαραγωγή ενέργειας σε ώρες έντονου ηλιακού φωτός ή ισχυρών ανέμων μπορεί να υπερφορτώσει το δίκτυο. Για την αντιμετώπιση αυτών, εφαρμόζονται λύσεις όπως η αποθήκευση ενέργειας, η υδροηλεκτρική αντλησιοταμίευση και, μελλοντικά, η χρήση μπαταριών, ενώ η ψηφιακή τεχνολογία βελτιστοποιεί την πρόβλεψη και τη διαχείριση ροής [23]. Επιπλέον, μηχανισμοί διαχείρισης ζήτησης προσαρμόζουν την κατανάλωση των χρηστών στην προσφορά, με κίνητρα όπως ευέλικτα τιμολόγια για ώρες υψηλής παραγωγής ΑΠΕ [17].

Δεύτερος κεντρικός στόχος κάθε αγοράς ηλεκτρικής ενέργειας είναι η βελτιστοποίηση της οικονομικής λειτουργίας του συστήματος. Η προσπάθεια διαμόρφωσης σταθερού πλαισίου έχει ως αποτέλεσμα την μείωση των τιμών για τους καταναλωτές ενώ ταυτόχρονα η θωράκιση της βιωσιμότητας των ενεργειακών επενδύσεων δίνει επαρκή κίνητρα στους παραγωγούς, να χρηματοδοτούν ενεργειακές υποδομές και να παρέχουν την απαιτούμενη παραγωγική ικανότητα, ενισχύοντας έτσι την επάρκεια και την ποιότητα του ενεργειακού εφοδιασμού.

Αναλυτικότερα μηχανισμοί παρακολούθησης και ρύθμισης της αγοράς προάγουν την διαφάνεια των συναλλαγών και ενισχύουν τον ανταγωνισμό. Ανεξάρτητες ρυθμιστικές αρχές εποπτεύουν την αγορά και φροντίζουν για την συμμόρφωση των παικτών με τους κανονισμούς και αντιμετωπίζουν πρακτικές που περιορίζουν τον ανταγωνισμό. Ταυτόχρονα ρυθμίσεις διευκολύνουν την πρόσβαση σε νέους προμηθευτές, διατηρούν την ισορροπία μεταξύ μικρών και μεγάλων παικτών και περιορίζουν την κυριαρχία των ισχυρών στις κρίσιμες υποδομές όπως τα δίκτυα μεταφοράς. Η χρήση καινοτόμων τεχνολογιών επιτρέπει την παρακολούθηση της αγοράς σε πραγματικό χρόνο με αποτέλεσμα την άμεση ανίχνευση ανισορροπιών και την προσαρμογή της λειτουργίας της για την αποφυγή αθέμιτων πλεονεκτημάτων. Τέτοιο παράδειγμα αποτελεί ο εντοπισμός καταχρηστικής δεσπόζουσας θέσης. Γίνεται σαφές πως έτσι διαφυλάττονται τα συμφέροντα τόσο των καταναλωτών όσο και των προμηθευτών ενώ ταυτόχρονα ενισχύονται η ανάπτυξη και ο εκσυγχρονισμός του δικτύου.

1.3 Συμμετέχοντες στην αγορά ηλεκτρικής ενέργειας

Παραγωγοί

Ως παραγωγοί ορίζονται τα φυσικά ή νομικά πρόσωπα τα οποία δραστηριοποιούνται στην παραγωγή μικρότερων ή μεγαλύτερων ποσοτήτων ενέργειας από διάφορες πηγές, όπως θερμικούς σταθμούς, ανανεώσιμες πηγές ενέργειας (ΑΠΕ) και φυσικό αέριο οι οποίες εγχέονται στο δίκτυο. Κάθε

παραγωγός είναι κάτοχος αντίστοιχης άδειας από την ρυθμιστική αρχή ενέργειας (ΡΑΕ) [1, 24].

Προμηθευτές - πάροχοι

Ως προμηθευτές ορίζονται κυρίως νομικά πρόσωπα που αγοράζουν ενέργεια από την αγορά χονδρικής, είτε απευθείας από παραγωγούς ενέργειας είτε μέσω του χρηματιστηρίου ενέργειας και την πουλάν σε τελικούς καταναλωτές. Οι προμηθευτές είναι υπεύθυνοι για τη διαμόρφωση των τιμολογίων, την εξυπηρέτηση των πελατών, και τη διαχείριση των συμβάσεων παροχής ενέργειας. Κάθε προμηθευτής είναι κάτοχος αντίστοιχης άδειας από την ρυθμιστική αρχή ενέργειας [24].

Καταναλωτές

Ως τελικοί καταναλωτές ορίζονται φυσικά ή νομικά πρόσωπα τα οποία αγοράζουν ποσότητες ενέργειας, σε τιμές λιανικής, από τους προμηθευτές παρόχους, μετά την σύναψη συμβολαίων [24].

Ρυθμιστική Αρχή Αποβλήτων, Ενέργειας και Υδάτων (ΡΑΑΕΥ)

Η Ρυθμιστική Αρχή Αποβλήτων, Ενέργειας και Υδάτων (ΡΑΑΕΥ) αποτελεί την εθνική ρυθμιστική αρχή για την ενέργεια, τα ύδατα και τα απόβλητα. Ιδρύθηκε ως ανεξάρτητη αρχή με τον νόμο 2773/1999, στο πλαίσιο της εναρμόνισης με ευρωπαϊκές οδηγίες σχετικά με τους κοινούς κανόνες για την εσωτερική αγορά ηλεκτρικής ενέργειας, καθώς και οδηγίες για τον ηλεκτρισμό και το φυσικό αέριο. Αρχικά, οι αρμοδιότητές της επικεντρώνονταν στην εποπτεία της εγχώριας αγοράς ενέργειας, την προώθηση του ανταγωνισμού και την εισήγηση μέτρων προς την Πολιτεία για την απελευθέρωση των αγορών ηλεκτρικής ενέργειας και φυσικού αερίου. Με την πάροδο των ετών, ο ρόλος της ενισχύθηκε σημαντικά και το 2011 αναβαθμίστηκε σε «εγγυητή» της ομαλής λειτουργίας αυτών των αγορών. Σήμερα, η ΡΑΑΕΥ αποτελεί αυτόνομη νομική οντότητα με διοικητική και οικονομική αυτοτέλεια, υπόκειται δε μόνο σε κοινοβουλευτικό και δικαστικό έλεγχο, ενώ εποπτεύεται από τον Υπουργό Ανάπτυξης. Η σημασία της ως εθνικής ρυθμιστικής αρχής ενισχύθηκε περαιτέρω με την επέκταση των αρμοδιοτήτων της στους τομείς των υπηρεσιών ύδατος και της διαχείρισης αστικών αποβλήτων, οδηγώντας στη μετονομασία της από Ρυθμιστική Αρχή Ενέργειας (ΡΑΕ) σε ΡΑΑΕΥ [14, 1].

Στον τομέα της ενέργειας, στόχος της ΡΑΑΕΥ είναι η εποπτεία και η ρύθμιση της αγοράς ηλεκτρισμού και φυσικού αερίου, με ποικίλες αρμοδιότητες. Συγκεκριμένα, οι δραστηριότητές της περιλαμβάνουν την παρακολούθηση και εποπτεία της αγοράς, την έκδοση συστάσεων και την εισήγηση μέτρων για τη διασφάλιση του ανταγωνισμού, την προστασία του περιβάλλοντος μέσω της προώθησης της πράσινης ενέργειας και των ΑΠΕ, καθώς και την υποστήριξη της ενιαίας ευρωπαϊκής αγοράς ηλεκτρισμού. Αναλυτικότερα, η ΡΑΑΕΥ προστατεύει τους καταναλωτές, εξετάζοντας καταγγελίες που σχετίζονται αποκλειστικά με

ρυθμιστικά ζητήματα, ενώ έχει την εξουσία να χορηγεί και να ανακαλεί άδειες ενεργειακών δραστηριοτήτων, τηρώντας τις αρχές της ίσης μεταχείρισης. Στον τομέα των μη ανταγωνιστικών δραστηριοτήτων, πιστοποιεί εταιρείες ως διαχειριστές μεταφοράς σύμφωνα με την ενεργειακή νομοθεσία και εγκρίνει προγράμματα ανάπτυξης. Επίσης, θεσπίζει κανόνες πρόσβασης στο δίκτυο, καθορίζει την τιμολογιακή πολιτική και εξασφαλίζει ισότιμη πρόσβαση στο ενεργειακό σύστημα και τις σχετικές πληροφορίες. Παράλληλα, εγγυάται την ασφάλεια του ενεργειακού εφοδιασμού, εκτιμά τις μελλοντικές ενεργειακές ανάγκες και εφαρμόζει μέτρα για την αντιμετώπιση κρίσεων, όπως οι περίοδοι αιχμής ζήτησης. Τέλος, επιθεωρεί τη φυσική κατάσταση και τη συντήρηση των συστημάτων μεταφοράς και διανομής, καθώς και έργα που αποσκοπούν στην ενίσχυση της παραγωγικής ικανότητας.

Ανεξάρτητος Διαχειριστής Μεταφοράς Ηλεκτρικής Ενέργειας (Α.Δ.Μ.Η.Ε.)

Ως Ελληνικό Σύστημα Μεταφοράς Ηλεκτρικής Ενέργειας (ΕΣΜΗΕ) ορίζεται το δίκτυο υποδομών που χρησιμοποιείται για τη μεταφορά ηλεκτρικής ενέργειας από τους παραγωγούς στους διαχειριστές διανομής ή σε μεγάλους, βιομηχανικής κλίμακας, κυρίως τελικούς καταναλωτές. Χαρακτηρίζεται από μεγάλες αποστάσεις, καλύπτει εθνικά ή διασυνοριακά δίκτυα ενώ, αποτελείται από γραμμές υψηλής και υπερυψηλής τάσης.

Ο διαχειριστής του δικτύου μεταφοράς – (ΕΣΜΗΕ) είναι ο ΑΔΜΗΕ, ο οποίος συστάθηκε το 2011, κατά την ενσωμάτωση του τρίτου Ευρωπαϊκού πακέτου με την απόσχιση του ΔΕΣΜΗΕ στον ΛΑΓΗΕ, λειτουργού της αγοράς ενέργειας, και στον ΑΔΜΗΕ, διαχειριστή του δικτύου μεταφοράς. Εποπτεύεται από την ΡΑΑΕΥ αφού η λειτουργία και οι σκοποί του χαρακτηρίζονται ως μη ανταγωνιστική δραστηριότητα. Σκοπός του είναι η αδιάλειπτη παροχή ηλεκτρικής ενέργειας στην χώρα με την διαχείριση του δικτύου μεταφοράς καθώς και η λειτουργία της αγοράς εξισορρόπησης, όπως και η διαχείριση των εμπορικών συναλλαγών με τρίτες χώρες. Βασικές αρχές του ΑΔΜΗΕ είναι η ισότιμη πρόσβαση στο σύστημα από όλους τους χρήστες, η παροχή δεδομένων σε αυτούς όπως και η πλήρης διαφάνεια στις διαδικασίες λειτουργίας του. Είναι υπεύθυνος για την λειτουργία, συντήρηση και ανάπτυξη του δικτύου μεταφοράς. Στον τομέα της λειτουργίας του δικτύου είναι υπεύθυνος για την διαχείριση των ροών ενέργειας που σχετίζονται με την παραγωγή, την κατανάλωση και τις διάφορες εισαγωγές- εξαγωγές διατηρώντας την ισορροπία προσφοράς και ζήτησης ενεργειακού φορτίου σε πραγματικό χρόνο. Ακόμη είναι υπεύθυνος για την κάλυψη των απωλειών στο σύστημα οι οποίες οφείλονται σε φυσικά φαινόμενα, όπως η αντίσταση των γραμμών, αλλά και στις μεγάλες αποστάσεις στις οποίες αναπτύσσεται το δίκτυο. Εξαιτίας αυτών η ενέργεια που καταλήγει στους τελικούς χρήστες είναι μικρότερη από αυτήν που έχει παραχθεί. Για την αντιμετώπιση αυτού του φαινομένου, αφού υπολογιστούν τα μεγέθη των απωλειών σε πραγματικό χρόνο ο ΑΔΜΗΕ, μέσω εμπορικών συναλλαγών, καλύπτει την ποσότητα της ενέργειας που υπολείπεται. Ακόμη μία κομβική αρμοδιότητα είναι η διαχείριση των ηλεκτρικών διασυνδέσεων με τρίτες χώρες καθώς και με τα διασυνδεδεμένα νησιά της

Ελλάδας δηλαδή αυτά που έχουν συνδεθεί με το ηπειρωτικό ηλεκτρικό δίκτυο της Ελλάδας μέσω υποβρύχιων καλωδίων υψηλής τάσης [15, 2].

Στον τομέα της συντήρησης επιθεωρεί τη λειτουργία των εξαρτημάτων του δικτύου, όπως υποσταθμοί και γραμμές μεταφοράς, εφαρμόζει πρωτόκολλα αποφυγής βλαβών στο δίκτυο, επιδιορθώνει βλάβες και τέλος είναι υπεύθυνος για την διαχείριση της ενεργοποίησης των εφεδρειών σε έκτακτες περιπτώσεις. Τέλος στον τομέα της ανάπτυξης του δικτύου εξασφαλίζει μακροπρόθεσμα επαρκή χωρητικότητα ώστε να καλύπτονται οι ανάγκες ηλεκτρικής ενέργειας της χώρας. Χαρακτηριστικό παράδειγμα αποτελεί το σχέδιο δεκαετούς προγράμματος ανάπτυξης του ΕΣΜΗΕ, όπου προσδιορίζονται οι υποδομές μεταφοράς που θα αναβαθμιστούν καθώς και τα απαραίτητα έργα για την διείσδυση των ανανεώσιμων πηγών ενέργειας [25].

Διαχειριστής του Ελληνικού Δικτύου Διανομής Ηλεκτρικής Ενέργειας (ΔΕΔΔΗΕ)

Ως ελληνικό σύστημα διανομής ηλεκτρικής ενέργειας (Ε.Σ.Δ.Η.Ε.) ορίζεται το δίκτυο υποδομών που χρησιμοποιείται για την διανομή της ηλεκτρικής ενέργειας σε τελικούς καταναλωτές σε οικιακό, εμπορικό ή ακόμη και βιομηχανικό επίπεδο. Σε αντίθεση με το δίκτυο μεταφοράς περιλαμβάνει γραμμές μέσης και χαμηλής τάσης.

Διαχειριστής του δικτύου διανομής είναι ο ΔΕΔΔΗΕ. Συστάθηκε το 2011 μετά τον διαχωρισμό του από την Δ.Ε.Η. Α.Ε., σύμφωνα με ευρωπαϊκή οδηγία κατά την οποία καθετοποιημένες επιχειρήσεις, όπου συμμετέχουν σε πολλαπλά στάδια της παραγωγής, προμήθειας και μεταφοράς ηλεκτρικής ενέργειας, απαγορεύτηκε να διαχειρίζονται το δίκτυο ενέργειας. Πλέον σαν ανεξάρτητος φορέας ελέγχεται από την ΡΑΑΕΥ καθώς η διαχείριση του δικτύου αποτελεί φυσικό μονοπώλιο και εντάσσεται στις μη ανταγωνιστικές δραστηριότητες. Στο πλαίσιο αυτό κύριο σκοπό του αποτελεί η ολοκληρωτική διαχείριση και ανάπτυξη του δικτύου διανομής για την κατοχύρωση της πρόσβασης των τελικών χρηστών με τρόπο διάφανη και οικονομικά αποδοτικό. Υποχρεούται να λαμβάνει τα απαραίτητα μέτρα ώστε να τηρούνται τα όρια ποιότητας της ενέργειας. Αυτά είναι η αδιάλειπτη τροφοδότηση καθώς και η διασφάλιση ορισμένων τεχνικών ποιοτικών χαρακτηριστικών της τάσης, όπως το πλάτος, η συχνότητα, η κυματομορφή και η τριφασική συμμετρία. Ακόμη είναι υπεύθυνος για την συντήρηση του δικτύου καθώς και για την πρόβλεψη και την επιδιόρθωση βλαβών σε όσο το δυνατόν συντομότερο χρονικό διάστημα. Σε περιπτώσεις διακοπής της τροφοδότησης είναι υποχρεωμένος να εκδίδει ανακοινώσεις για το περιστατικό της βλάβης αλλά και για τα αίτια που οδήγησαν σε αυτήν προς γνώση των καταναλωτών. Επιπρόσθετα δεύτερη κύρια λειτουργία του είναι η διαχείριση των αγορών των μη διασυνδεδεμένων νησιών. Φροντίζει για την εύρυθμη λειτουργία του τοπικού δικτύου διανομής, όπως επίσης για τους τοπικούς αυτόνομους σταθμούς παραγωγής ηλεκτρισμού. Σε αυτά τα τοπικά δίκτυα, τα οποία είναι απομονωμένα από το κεντρικό δίκτυο της ηπειρωτικής Ελλάδας, υπεύθυνος για να καλύπτει τις ανάγκες εξισορρόπησης παραγωγής και της ζήτησης κάθε χρονική στιγμή έχει οριστεί ο ΔΕΔΔΗΕ [15, 3].

Ελληνικό χρηματιστήριο Ενέργειας Α.Ε. (Ε.Χ.Ε.)

Μέχρι το 2018, ο οργανισμός που ήταν υπεύθυνος για τη λειτουργία της κεντρικής αγοράς ηλεκτρισμού και τη διαχείριση των ανανεώσιμων πηγών ενέργειας (ΑΠΕ) ήταν ο Λειτουργός Αγοράς Ηλεκτρικής Ενέργειας (Λ.ΑΓ.Η.Ε.). Στο πλαίσιο του ευρύτερου σχεδίου της Ευρωπαϊκής Ένωσης για την απελευθέρωση της ελληνικής αγοράς ενέργειας, ο Λ.ΑΓ.Η.Ε. καταργήθηκε και αντικαταστάθηκε από δύο ανεξάρτητους φορείς: το Ελληνικό Χρηματιστήριο Ενέργειας (Ε.Χ.Ε.), το οποίο ανέλαβε τη λειτουργία της αγοράς, και τον Διαχειριστή ΑΠΕ και Εγγυήσεων Προέλευσης (Δ.ΑΠΕ.Ε.Π.), για τον οποίο γίνεται αναφορά στην επόμενη ενότητα. Το Ε.Χ.Ε. ιδρύθηκε ως θυγατρική του Χρηματιστηρίου Αθηνών και εποπτεύεται από τη Ρυθμιστική Αρχή Αποβλήτων, Ενέργειας και Υδάτων (ΡΑΑΕΥ). Σκοπός της δημιουργίας του ήταν η εναρμόνιση της ελληνικής αγοράς ενέργειας με το ευρωπαϊκό πλαίσιο, μέσω της περαιτέρω κατάργησης της μονοπωλιακής δομής της αγοράς.

Το Ε.Χ.Ε. λειτουργεί ως κεντρικός διαχειριστής της αγοράς ηλεκτρικής ενέργειας στην Ελλάδα, ρυθμίζοντας και διαχειριζόμενο τρεις από τις τέσσερις βασικές αγορές ενέργειας: την Αγορά Επόμενης Ημέρας (Day-Ahead Market), την Ενδοημερήσια Αγορά (Intraday Market) και την Αγορά Παραγώγων (Derivatives Market) [4, 18]. Η λειτουργία της Αγοράς Εξισορρόπησης (Balancing Market), όπως έχει ήδη αναφερθεί, αποτελεί αρμοδιότητα του Ανεξάρτητου Διαχειριστή Μεταφοράς Ηλεκτρικής Ενέργειας (Α.Δ.Μ.Η.Ε.). Στο πλαίσιο αυτό, το Ε.Χ.Ε. είναι υπεύθυνο για τη διαχείριση των συναλλαγών, τον συντονισμό και την εκκαθάριση των συμβολαίων μεταξύ των συμμετεχόντων, καθώς και για τη διασφάλιση της συμμόρφωσής τους με τους κανόνες λειτουργίας. Επιπλέον, παρέχει την πλατφόρμα λογισμικού μέσω της οποίας διεξάγονται όλες οι συναλλαγές της αγοράς ενέργειας. Τέλος, εξασφαλίζει ορθή τιμολόγηση, η οποία αντικατοπτρίζει τη δυναμική της αγοράς χωρίς εξωτερικές στρεβλώσεις για εμπορικούς σκοπούς. Αυτό είναι ιδιαίτερα σημαντικό, καθώς η ενέργεια αποτελεί κοινωνικό αγαθό, θεμελιώδες για τη βιωσιμότητα της κοινωνίας και της οικονομίας. Στις ευρύτερες αρμοδιότητές του περιλαμβάνεται η προώθηση του ανταγωνισμού μέσω της διατήρησης ενός διαφανούς περιβάλλοντος με ισότιμη μεταχείριση όλων των συμμετεχόντων, καθώς και η δημοσίευση δεδομένων, όπως τιμές και μεγέθη προσφοράς και ζήτησης, για ελεύθερη πρόσβαση όλων.

Ιδιαίτερα αξιοσημείωτος είναι ο ρόλος του Ε.Χ.Ε. στην υποστήριξη της ένταξης των ΑΠΕ, δημιουργώντας συνθήκες που ενισχύουν τη βιωσιμότητα των πράσινων επενδύσεων. Συγκεκριμένα, μέσω εξειδικευμένων μηχανισμών, εξασφαλίζει σταθερά έσοδα για τους παραγωγούς ΑΠΕ, μειώνοντας το επενδυτικό ρίσκο. Τέτοιοι μηχανισμοί περιλαμβάνουν την απευθείας συμμετοχή των παραγωγών ΑΠΕ στην Αγορά Επόμενης Ημέρας, τη δυνατότητα διόρθωσης των προβλέψεων παραγωγής τους στην Ενδοημερήσια Αγορά λόγω καιρικών συνθηκών, καθώς και τα συμβόλαια διαφοράς, τα οποία προστατεύουν τους επενδυτές από τις διακυμάνσεις της αγοράς.

Διαχειριστής Α.Π.Ε. και Εγγυήσεων Προέλευσης (Δ.ΑΠΕ.Ε.Π.)

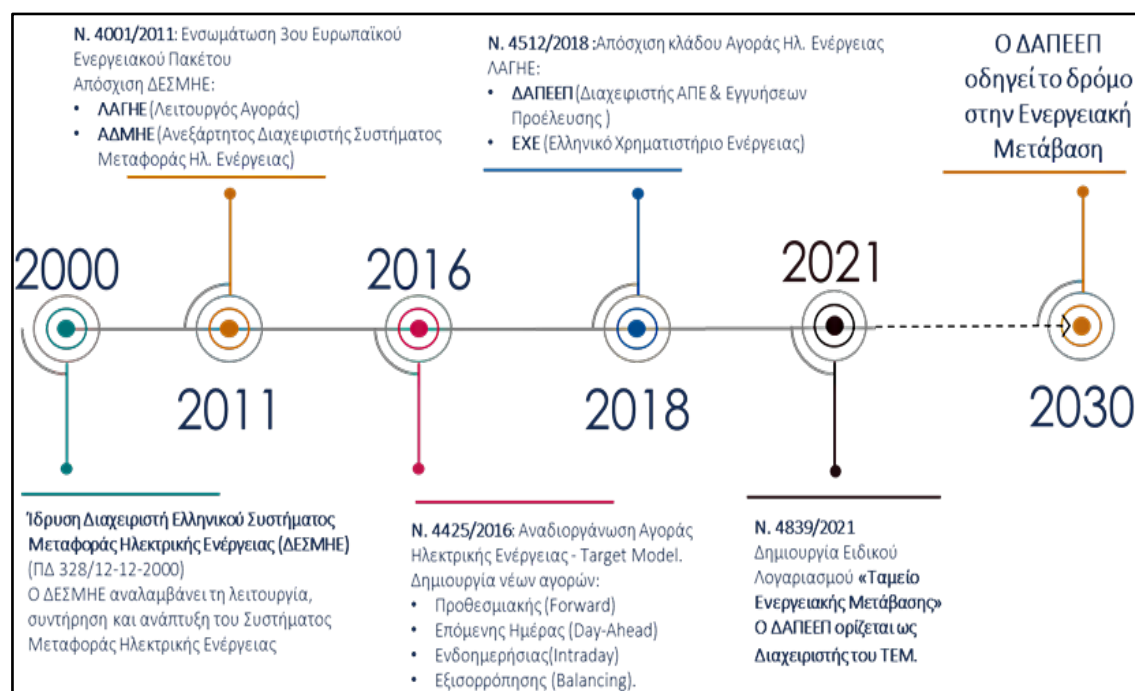
Η πράσινη ανάπτυξη έχει αναδειχθεί σε κεντρικό άξονα των πολιτικών της ευρωπαϊκής ένωσης, όχι μόνο για την προστασία του περιβάλλοντος αλλά και για την ενίσχυση της ενεργειακής ανεξαρτησίας και της οικονομικής ανθεκτικότητας. Στην Ελλάδα η συμμετοχή των Α.Π.Ε. στο ενεργειακό μείγμα παρουσιάζει σταθερή άνοδο τα τελευταία χρόνια ενώ το 2023 σημειώθηκε ιστορικό υψηλό, με το 57% της ηλεκτροπαραγωγής να προέρχεται από ΑΠΕ και υδροηλεκτρικές μονάδες [16, 19]. Το μείζον χαρακτηριστικό της παραγόμενης ενέργειας από Α.Π.Ε. είναι η στοχαστικότητα. Αυτή η στοχαστικότητα οφείλεται στην εξάρτησή της από φυσικούς παράγοντες όπως ο καιρός, με αποτέλεσμα απότομες αυξομειώσεις στην παραγωγή. Ακόμη δεν είναι δυνατή η απόλυτη πρόβλεψη του χρονικού ορίζοντα της παραγωγής, καθώς αυτή πραγματοποιείται μόνο όταν το επιτρέπουν οι καιρικές συνθήκες. Η ανάγκη για εξειδικευμένη διαχείριση της παραγόμενης ενέργειας από Α.Π.Ε. οδήγησε στην δημιουργία ενός αυτοτελή φορέα για την λειτουργία και τον συντονισμό των Α.Π.Ε.

Ο Δ.ΑΠΕ.Ε.Π. ιδρύθηκε το 2018 με την απόσχιση του Λ.Α.Γ.Η.Ε. και αποτελεί θυγατρική του Ελληνικού Δημοσίου και λειτουργεί υπό την εποπτεία του Υπουργείου Περιβάλλοντος και Ενέργειας ενώ έχει δύο κύριες αρμοδιότητες. Η πρώτη είναι η λειτουργία της αγοράς παραγωγής ενέργειας από Α.Π.Ε. και από μονάδες συμπαραγωγής ηλεκτρισμού και θερμότητας υψηλής απόδοσης (Σ.Η.Θ.Υ.Α.), ενώ δεύτερη είναι η διαχείριση των εγγυήσεων προέλευσης των παραπάνω «ειδικών» παραγωγών. Αναλυτικά γίνεται καταγραφή σε πραγματικό χρόνο της παραγόμενης ενέργειας, που ενσωματώνεται στο συνολικό ενεργειακό μείγμα, μέσω μετρήσεων που προέρχονται από συστήματα τηλεμετρίας και μετρητές των δικτύων μεταφοράς και διανομής. Τα δεδομένα αυτά παρέχονται από τον ΑΔΜΗΕ και τον ΔΕΔΔΗΕ αντίστοιχα. Τα έσοδα από την πώληση της ενέργειας στην αγορά χοντρικής συγκεντρώνονται σε ειδικό λογαριασμό του οποίου διαχειριστής είναι ο Δ.ΑΠΕ.Ε.Π. Στην συνέχεια αυτός προχωράει στην εκκαθάριση, δηλαδή στην κατανομή των εσόδων στους παραγωγούς με βάση τα συμφωνηθέντα τιμολόγια [5].

Αξίζει να σημειωθεί πως εξαιτίας της ανάμειξης του με τα οικονομικά, ο φορέας αποκτά και εποπτική λειτουργία καθώς εποπτεύει τις μονάδες ΑΠΕ και Σ.Η.Θ.Υ.Α, διασφαλίζοντας ότι τηρούν τους όρους της αδειοδότησης και σε περίπτωση μη συμμόρφωσης, δύναται να προχωρά σε διορθωτικές κινήσεις ή ακόμη και να αναστέλλει πληρωμές. Τέλος είναι διαχειριστής του ταμείου ενεργειακής μετάβασης, το οποίο χρηματοδοτείται από τα έσοδα από δημοπρασίες των δικαιωμάτων εκπομπών CO₂ και από κρατικές και ευρωπαϊκές επιχορηγήσεις. Σκοπός του ταμείου είναι η στήριξη στην πράσινη μετάβαση παρέχοντας οικονομικές διευκολύνσεις στους πολίτες. Τέτοιες είναι επιδότηση λογαριασμών, χρηματοδότηση προγραμμάτων ενεργειακής αναβάθμισης κτιρίων, και επιδοτήσεις στον τομέα της ηλεκτροκίνησης.

Συμπερασματικά ο Δ.ΑΠΕ.Ε.Π. λειτουργεί ως ο "συντονιστής" των ΑΠΕ αφού διαχειρίζεται τις οικονομικές και διοικητικές πτυχές που σχετίζονται με την

παραγωγή και πληρωμή. Δεν ελέγχει τη φυσική ένταξη της ενέργειας στο δίκτυο, αλλά ούτε και την τιμή πώλησης, αντιθέτως ο ρόλος του εξειδικεύεται στην αναλυτική καταγραφή της ενέργειας που έχει πωληθεί και στην εκκαθάριση των ποσών που προκύπτουν από τις δημοπρασίες.



Σχήμα 1.1 : Αναδρομή του Δ.ΑΠΕ.Ε.Π

1.4 Δομή της αγοράς ενέργειας

Η απελευθέρωση της αγοράς ενέργειας και παράλληλα η στόχευση για μία ενιαία Ευρωπαϊκή ενεργειακή ένωση αποτελεί μια σύνθετη διαδικασία που απαιτεί συνδυασμό ρυθμιστικών, τεχνολογικών και κοινωνικών μετασχηματισμών. Η διαδικασία της μετάβασης είναι χρονοβόρα εξαιτίας της ανάγκης προσαρμογής των υποδομών και αλλαγής της νοοτροπίας. Οι ζωές και οι άνθρωποι επηρεάζονται άμεσα, καθιστώντας απαραίτητη τη συνετή διαχείριση για την εξασφάλιση δίκαιης και βιώσιμης προόδου. Έτσι και στην Ελλάδα η δομή της αγοράς αλλάζει συνεχώς μέσα στα χρόνια.

Στον ελληνικό χώρο μέχρι και το 2020 επικρατούσε το μοντέλο κοινοπραξίας ισχύος (PoolCo Model) [17]. Γενικό χαρακτηριστικό του είναι η συγκεντρωτική λειτουργία της αγοράς κατά την οποία μέσω κοινοπρακτικού μηχανισμού πραγματοποιείται η γενική εκκαθάριση των συναλλαγών για παραγωγούς, προμηθευτές και καταναλωτές. Επιγραμματικά κάθε παραγωγός είναι ελεύθερος να υποβάλλει προσφορά της ποσότητας ενέργειας που είναι διατεθειμένος να παράξει αλλά και της τιμής της. Στη συνέχεια ο λειτουργός της αγοράς προβλέπει το φορτίο ζήτησης, ταξινομεί τις προσφορές των παραγωγών κατά αύξουσα σειρά

τιμής πώλησης και τις αποδέχεται μέχρι να καλυφθεί το μέγεθος της ζήτησης. Τελικά όλοι οι παραγωγοί θα πληρωθούν στην τιμή της εκκαθάρισης, Η οποία ορίζεται ως την υψηλότερη τιμή της προσφοράς ενέργειας που έγινε αποδεκτή από τον λειτουργό. Σε αυτό το πλαίσιο η τιμή της ενέργειας δεν μπορεί να είναι μικρότερη του κόστους παραγωγής αλλά ταυτόχρονα ο ανταγωνισμός αποκλείει φαινόμενα αισχροκέρδειας.

Με την απόφαση της ευρωπαϊκής ένωσης για την δημιουργία μίας έννοιας ευρωπαϊκής αγοράς ηλεκτρικής ενέργειας και φυσικού αερίου το μοντέλο κοινοπραξίας αντικαταστάθηκε από το μοντέλο στόχου (Target Model), το 2020 [10, 4]. Το μοντέλο αυτό στοχεύει στη διασυννοριακή ενοποίηση των αγορών, στη βελτίωση του ανταγωνισμού, στη περεταίρω μείωση του κόστους ενέργειας καθώς και στην προώθηση της βιώσιμης ανάπτυξης. Παρά το γεγονός ότι, αρκετά στοιχεία του παλαιού μοντέλου έχουν διατηρηθεί, έχουν υπάρξει ριζικές αλλαγές με βάση το ευρωπαϊκό ρυθμιστικό πλαίσιο. Στο μοντέλο στόχου η αγορά εξακολουθεί να βασίζεται σε έναν κεντρικό μηχανισμό εκκαθάρισης για παραγωγούς, προμηθευτές και καταναλωτές, ωστόσο, το Target Model λειτουργεί μέσω του χρηματιστήριου ενέργειας. Με αυτό τον τρόπο ενισχύεται ο ανταγωνισμός, καθώς, σε αντίθεση με το μοντέλο κοινοπραξίας όπου μόνο οι προμηθευτές είχαν δικαίωμα συμμετοχής στις δημοπρασίες υποβάλλοντας προσφορές, εδώ μπορούν να συμμετέχουν ενεργά και οι προμηθευτές-καταναλωτές με την κατάθεση προσφορών που περιλαμβάνουν την ποσότητα της ενέργειας που προβλέπουν ότι θα χρειαστούν καθώς και σε ποια τιμή είναι διατεθειμένοι να την αγοράσουν. Δημιουργούνται έτσι μία καμπύλη προσφοράς από τους παραγωγούς και μία καμπύλη ζήτησης από τους προμηθευτές – καταναλωτές. Το σημείο τομής των καμπυλών ορίζει την τιμή εκκαθάρισης της αγοράς. Την τιμή δηλαδή στην οποία θα διατεθεί το σύνολο της ενέργειας. Ακόμη εκτός από τους κεντρικούς μηχανισμούς για την επίτευξη συναλλαγών καθίστανται δυνατές και διμερείς συμφωνίες (Over the Counter - OTC), κατά τις οποίες δύο συμμετέχοντες συμφωνούν μεταξύ τους για την αγοροπωλησία συγκεκριμένων ποσοτήτων ενέργειας σε προκαθορισμένες τιμές. Ταυτόχρονα οι τιμές προκύπτουν από ανταγωνιστικές αγορές μέσω ανοιχτών και διαφανών διαδικασιών, ενώ στο προηγούμενο μοντέλο υπεύθυνος για την τιμή ήταν ένας κεντρικός λειτουργός της αγοράς, με όλες τις αποφάσεις για την κατανομή και την διαμόρφωση της τιμής να εξαρτάται από αυτόν. Ιδιαίτερη αναφορά πρέπει να γίνει επίσης στις Α.Π.Ε. Καθώς πλέον αυτές ενσωματώνονται πλήρως στο ενεργειακό μίγμα, αφού υπάρχουν συγκεκριμένοι μηχανισμοί που δίνουν προτεραιότητα στην απορρόφηση της παραγόμενης ενέργειας από αυτές.

Μία ακόμη μεγάλη καινοτομία του μοντέλου στόχου αποτέλεσε η διασύνδεση των αγορών διαφορετικών χωρών με το Market Coupling, όπου πρόκειται για έναν μηχανισμό που συνδέει τις αγορές ηλεκτρικής ενέργειας διαφορετικών χωρών προκειμένου να επιτευχθεί η αποτελεσματική κατανομή της διασυννοριακής δυναμικότητας ενισχύοντας την ασφάλεια εφοδιασμού καθώς οι χώρες μπορούν να επωφεληθούν από την παραγωγή άλλων αγορών σε περιόδους αυξημένης ζήτησης ή μειωμένης τοπικής παραγωγής. Ταυτόχρονα διαμορφώνονται κοινές τιμές ενέργειας καθώς όταν οι αγορές δύο χωρών παρουσιάζουν διαφορετικές

τιμές το Market Coupling επιτρέπει τη ροή ενέργειας από την αγορά με τη χαμηλότερη τιμή προς αυτήν με την υψηλότερη, εξισορροπώντας τις τιμές [21].

Τέλος η βασικότερη λειτουργία του μοντέλου στόχου είναι ο διαχωρισμός των αγορών και η οργάνωση τους σε τέσσερις εξειδικευμένους τομείς, σε αντίθεση με την έννοια της κεντρικής κοινοπραξίας που ήταν υπεύθυνη για το σύνολο της εκκαθάρισης. Αυτοί είναι η αγορά επόμενης ημέρας, η ενδοημερήσια αγορά, η αγορά εξισορρόπησης και η αγορά των παραγωγών.

Αγορά επόμενης ημέρας (day ahead market)

Η Αγορά Επόμενης Ημέρας (Day-Ahead Market) αποτελεί τον ακρογωνιαίο λίθο της αγοράς ηλεκτρικής ενέργειας, όπου οι συναλλαγές πραγματοποιούνται μία ημέρα πριν από τη φυσική παράδοση της ενέργειας. Στόχος της είναι να οριστεί κατά προσέγγιση ο προγραμματισμός για την παραγωγή και την κατανάλωση, καθώς και η τιμή εκκαθάρισης την επόμενη ημέρα σε ωριαία βάση [4, 18].

Λειτουργός της αγοράς επόμενης ημέρας είναι το Ελληνικό Χρηματιστήριο Ενέργειας. Οι παραγωγοί υποβάλλουν προσφορές για την ενέργεια που είναι διατεθειμένοι να παράξουν, καθορίζοντας τιμή και ποσότητα. Αντίστοιχα οι προμηθευτές και οι καταναλωτές υποβάλλουν προσφορές για την ποσότητα της ενέργειας που προβλέπουν να καταναλωθεί, καθορίζοντας επίσης την τιμή που είναι διατεθειμένοι να αγοράσουν. Έτσι δημιουργούνται οι καμπύλη της προσφοράς ενέργειας από τους προμηθευτές και η καμπύλη της ζήτησης ενέργειας από τους καταναλωτές. Στη συνέχεια ο λειτουργός διαμορφώνει την τιμή εκκαθάρισης αγοράς με βάση το σημείο τομής της καμπύλης προσφοράς και ζήτησης και τις ποσότητες ενέργειας που θα παραδοθούν από κάθε προμηθευτή. Παράλληλα η αγορά είναι συνδεδεμένη με τις αγορές άλλων ευρωπαϊκών χωρών μέσω του μηχανισμού Market Coupling, που επιτρέπει τη ροή ενέργειας μεταξύ διασυνδεδεμένων συστημάτων και την εξισορρόπηση των τιμών.

Η Αγορά Επόμενης Ημέρας έχει κυρίως προγραμματιστικό χαρακτήρα καθώς οι συναλλαγές ολοκληρώνονται μία ημέρα πριν, παρέχοντας επαρκή χρόνο για τον έγκαιρο σχεδιασμό της παραγωγής και της διανομής ενέργειας. Έτσι μειώνεται το κόστος, αφού οι συμμετέχοντες μπορούν να προβλέψουν καλύτερα τις ανάγκες τους ενώ ταυτόχρονα προωθείται ο ανταγωνισμός μέσα από τη διαφάνεια των συναλλαγών αλλά και του ορισμού της τιμής εκκαθάρισης. Τέλος αποτελεί την πιο σημαντική λειτουργία της αγοράς ενέργειας καθώς μέσω αυτής διακινείται το μεγαλύτερο μέρος της ενέργειας, ενώ επίσης οι τιμές που διαμορφώνονται λειτουργούν ως σήμα, επηρεάζοντας αποφάσεις για επενδύσεις στην παραγωγή ή τη μείωση του κόστους κατανάλωσης.

Ενδοημερήσια Αγορά Ενέργειας (Intraday Market)

Η Ενδοημερήσια Αγορά Ενέργειας (Intraday Market) παρέχει στους συμμετέχοντες τη δυνατότητα να πραγματοποιούν διορθωτικές κινήσεις στις συναλλαγές τους την ίδια ημέρα, πριν από τη φυσική παράδοση της ενέργειας. Σκοπός της είναι να προσφέρει μεγαλύτερη ευελιξία και ακρίβεια στη διαχείριση της παραγωγής και της ζήτησης, διασφαλίζοντας την αποδοτική λειτουργία του ενεργειακού συστήματος [4, 18].

Οι συναλλαγές διεξάγονται μέσω του Ελληνικού Χρηματιστηρίου Ενέργειας (Ε.Χ.Ε.), το οποίο λειτουργεί ως κεντρικός μηχανισμός εκκαθάρισης. Η τιμολόγηση είναι δυναμική κατά τη διάρκεια της ημέρας και επηρεάζεται από τις διακυμάνσεις στην ποσότητα ενέργειας που προορίζεται για αγορά ή πώληση, λόγω των διορθωμένων θέσεων των συμμετεχόντων στη δημοπρασία.

Έτσι, κατά τη διάρκεια αυτής της διαδικασίας, οι συμμετέχοντες μπορούν να υποβάλλουν νέες προσφορές ή να τροποποιούν τις υπάρχουσες καθ' όλη τη διάρκεια της ημέρας. Έτσι, προσαρμόζουν τη θέση τους με μεγαλύτερη ακρίβεια, με βάση τις πραγματικές συνθήκες παραγωγής, αντί να βασίζονται αποκλειστικά σε τεχνικές προβλέψεις, όπως συμβαίνει στην Αγορά Επόμενης Ημέρας. Παράλληλα, η Ενδοημερήσια Αγορά είναι διασυνδεδεμένη με άλλες ευρωπαϊκές αγορές μέσω του μηχανισμού Cross-Border Intraday Market Coupling. Ο μηχανισμός αυτός ενοποιεί τις ενδοημερήσιες αγορές διαφορετικών χωρών, επιτρέποντας τη διασυνοριακή ανταλλαγή ενέργειας κατά τη διάρκεια της ημέρας. Με αυτόν τον τρόπο, ενισχύεται η ευελιξία και η σταθερότητα του ενεργειακού συστήματος σε πραγματικό χρόνο. Οι διασυνοριακές συναλλαγές πραγματοποιούνται σε κοινές ευρωπαϊκές πλατφόρμες, όπως η Single Intraday Coupling (SIDC), εξασφαλίζοντας διαφάνεια στις διαδικασίες.

Γενικότερα, η Ενδοημερήσια Αγορά Ενέργειας αποτελεί μηχανισμό ζωτικής σημασίας για την ομαλή λειτουργία του ενεργειακού συστήματος. Επιτρέπει την προσαρμογή των συναλλαγών σε πραγματικό χρόνο, έως και λίγες ώρες πριν από τη φυσική παράδοση της ενέργειας, διορθώνοντας ανακρίβειες και αστοχίες που είναι αναπόφευκτες στις τεχνικές προβλέψεις. Επιπλέον, μειώνει τον κίνδυνο κυρώσεων για τους συμμετέχοντες, οι οποίοι ενδέχεται να μην μπορούν να εκπληρώσουν τις δημοπρατημένες προσφορές τους λόγω απρόβλεπτων συνθηκών. Τέλος, συμβάλλει σημαντικά στην ενσωμάτωση των Ανανεώσιμων Πηγών Ενέργειας (ΑΠΕ), καθώς, μέσω έγκυρων διορθώσεων, αντιμετωπίζει τη στοχαστικότητα της παραγωγής τους.

Αγορά Εξισορρόπησης (Balancing Market)

Η Αγορά Εξισορρόπησης (Balancing Market) αποτελεί το τελευταίο στάδιο στο σύστημα των αγορών ενέργειας και λειτουργεί σε πραγματικό χρόνο για να διατηρήσει την ισορροπία μεταξύ παραγωγής και κατανάλωσης ηλεκτρικής ενέργειας. Είναι ο μηχανισμός που ενεργοποιείται όταν προκύπτουν αποκλίσεις

από τα προγραμματισμένα επίπεδα ζήτησης ή παραγωγής, τα οποία δεν καλύπτονται από την Αγορά Επόμενης Ημέρας ή την Ενδοημερήσια Αγορά.

Λειτουργός της αγοράς εξισορρόπησης είναι ο διαχειριστής δικτύου μεταφοράς (ΑΔΜΗΕ), ο οποίος παρακολουθεί συνεχώς το δίκτυο για να διασφαλίσει την ισορροπία μεταξύ παραγωγής και η κατανάλωσης και στην περίπτωση που εντοπιστούν αποκλίσεις, ενεργοποιούνται εφεδρείες από μονάδες παραγωγής. Οι ενέργειες αυτές διασφαλίζουν την αξιοπιστία του συστήματος, αλλά κοστίζουν περισσότερο, καθώς πραγματοποιούνται άμεσα και σε συνθήκες πίεσης, με τις μονάδες παραγωγής που μπορούν να ανταπεξέλθουν σε αυτές τις συνθήκες να χρησιμοποιούν ακριβότερα και μη οικολογικά καύσιμα αυξάνοντας το κόστος παραγωγής. Ακόμη αρμοδιότητα του ΑΔΜΗΕ είναι να εκκαθαρίζει τις συναλλαγές και να επιβάλλει κόστη εξισορρόπησης σε όσους προκαλούν αποκλίσεις.

Η διαφορά της από την Αγορά Επόμενης Ημέρας και την Ενδοημερήσια Αγορά εντοπίζεται κυρίως ως προς το χρόνο και το σκοπό της. Η Αγορά Επόμενης Ημέρας προγραμματίζει συναλλαγές μία ημέρα πριν τη φυσική παράδοση, ενώ η Ενδοημερήσια Αγορά επιτρέπει διορθώσεις κατά τη διάρκεια της ημέρας. Αντίθετα, η Αγορά Εξισορρόπησης λειτουργεί αμέσως πριν ή κατά τη διάρκεια της φυσικής παράδοσης για να καλύψει τις αποκλίσεις που δεν προβλέφθηκαν αλλά και τις απώλειες του δικτύου. Τέλος είναι υποχρεωτική για όλους τους συμμετέχοντες, με τους παραγωγούς και τους καταναλωτές να υποχρεούνται να ανταποκρίνονται στις εντολές του ΑΔΜΗΕ. Γίνεται φανερό πως ο χαρακτήρας της Αγοράς Εξισορρόπησης είναι περισσότερο λειτουργικός για την σταθερότητα του δικτύου, διαφέροντας από τις υπόλοιπες οι οποίες έχουν καθαρά εμπορικό χαρακτήρα με έντονο το στοιχείο της ελεύθερης αγοράς στον σχεδιασμό τους.

Συμπερασματικά, η Αγορά Εξισορρόπησης διασφαλίζει τη εύρυθμη λειτουργία του ενεργειακού συστήματος, αντιμετωπίζοντας αποκλίσεις που θα μπορούσαν να οδηγήσουν σε αστάθεια ή διακοπές. Παρόλο που μπορεί να αυξηθεί το κόστος της ενέργειας λόγω της άμεσης ενεργοποίησης εφεδρειών, εξασφαλίζεται η αξιοπιστία του συστήματος με την αδιάκοπη παροχή ενέργειας στους πολίτες [4,18].

Αγορά Παραγώγων (Derivatives Market)

Η Αγορά Παραγώγων (Derivatives Market) αφορά τις συναλλαγές χρηματοοικονομικών προϊόντων που η αξία τους έχει άμεση σχέση με την αξία ενός βασικότερου προϊόντος, όπως η ηλεκτρική ενέργεια, το φυσικό αέριο, το πετρέλαιο ή άλλα εμπορεύματα. Στην περίπτωση της ενέργειας, η αγορά παραγώγων χρησιμοποιείται για τη διαχείριση του χρηματοοικονομικού κινδύνου που συνδέεται με τις διακυμάνσεις στις τιμές της ενέργειας. Ως παράγωγα προϊόντα στον τομέα του ηλεκτρισμού ορίζονται τα Συμβόλαια Μελλοντικής Εκπλήρωσης (Futures), τα οποία δεσμεύουν τα μέρη να αγοράσουν ή να πουλήσουν ενέργεια σε προκαθορισμένη τιμή και χρόνο καθώς και τα Δικαιώματα Προαίρεσης (Options), τα οποία δίνουν το δικαίωμα, χωρίς να υπάρχει κάποια άλλη δέσμευση, για αγορά ή πώληση ενέργειας σε συγκεκριμένη τιμή.

Οι συμμετέχοντες μπορούν να «κλειδώσουν» τιμές μέσω χρηματιστηριακών συμβάσεων, για να προστατευθούν από απρόβλεπτες αυξήσεις στις τιμές της ενέργειας. Οι συναλλαγές αυτές συνήθως διακανονίζονται χρηματοοικονομικά και αποτελούν πρόσφορο έδαφος για ανάπτυξη επιχειρηματικής δραστηριότητας καθώς οι επενδυτές συμμετέχουν στην αγορά παραγώγων για να κερδίσουν από τις μεταβολές στις τιμές της ενέργειας, χωρίς να εμπλέκονται στη φυσική παράδοση του προϊόντος και έτσι αυξάνεται η γενική ρευστότητα στην αγορά ενέργειας.

Συμπερασματικά η Αγορά Παραγώγων είναι άμεσο επακόλουθο της δημιουργίας χρηματιστηρίου ενέργειας και αποτελεί ένα πολύτιμο εργαλείο για τη διαχείριση κινδύνου και την ενίσχυση της οικονομικής σταθερότητας στην αγορά ενέργειας. Μέσω της χρήσης χρηματοοικονομικών συμβολαίων, οι συμμετέχοντες μπορούν να προστατεύσουν τα οικονομικά τους συμφέροντα σε μεγαλύτερο βαθμό.

Προθεσμιακή Αγορά

Η Προθεσμιακή Αγορά Ενέργειας είναι ένα από τα βασικά τμήματα της ελληνικής αγοράς ενέργειας, το οποίο προσφέρει στους συμμετέχοντες τη δυνατότητα να διαπραγματευτούν συμβόλαια για τη μελλοντική φυσική παράδοση ή αγορά ποσότητας ενέργειας σε προκαθορισμένες τιμές και όρους.

Οι συναλλαγές στην προθεσμιακή αγορά πραγματοποιούνται μέσω της πλατφόρμας του Ελληνικού Χρηματιστηρίου Ενέργειας (EXE) σε δίκαιο και ανοιχτό περιβάλλον. Αυτές ολοκληρώνονται είτε μέσω φυσικής παράδοσης της ενέργειας είτε μέσω χρηματοοικονομικού διακανονισμού, ανάλογα με τους όρους των συμβολαίων. Δίνεται λοιπόν η ευελιξία στους συμμετέχοντες να προσαρμόζουν τις στρατηγικές τους. Κύριο χαρακτηριστικό της προθεσμιακής αγοράς είναι η δυνατότητα δέσμευσης σταθερών ποσοτήτων ενέργειας σε προκαθορισμένη τιμή και με συγκεκριμένους όρους για μελλοντική χρήση. Αυτό πραγματοποιείται μέσω διμερών συμβολαίων Power Purchase Agreements (PPAs), τα οποία έχουν συνήθως διάρκεια 10 – 20 χρόνια. Τα PPAs μπορεί να είναι φυσικά ,όπου παραδίδεται φυσικά η ενέργεια μέσω του δικτύου ή χρηματοοικονομικά, όπου προκύπτει χρηματική αντιστάθμιση για την παραγόμενη ενέργεια, χωρίς φυσική παράδοση.

Τα οφέλη της προθεσμιακής αγοράς είναι πολλαπλά. Πρώτον, μειώνει την αβεβαιότητα και παρέχει σταθερότητα στις τιμές, προστατεύοντας τους συμμετέχοντες από τις απρόβλεπτες διακυμάνσεις της αγοράς. Δεύτερον, ενισχύει τη μακροπρόθεσμη ασφάλεια, επιτρέποντας τον στρατηγικό σχεδιασμό παραγωγής και κατανάλωσης, όπως παραδείγματος χάρη μεγάλες βιομηχανικές εγκαταστάσεις. Τέλος, υποστηρίζει τις επενδύσεις στον ενεργειακό τομέα, προσφέροντας ένα σταθερό οικονομικό πλαίσιο που ευνοεί τη βιώσιμη ανάπτυξη.

Σε σύγκριση με την αγορά παραγώγων που έχει καθαρά χρηματοοικονομικό χαρακτήρα, η προθεσμιακή αγορά εστιάζει περισσότερο στη φυσική διαχείριση της ενέργειας, ενώ αντίθετα η αγορά παραγώγων έχει κυρίως χρηματοοικονομικό

χαρακτήρα προσφέροντας εργαλεία για την αντιστάθμιση κινδύνου και την κερδοσκοπία, χωρίς να απαιτείται φυσική παράδοση ενέργειας. Έτσι, η προθεσμιακή αγορά εξυπηρετεί τις ανάγκες του ενεργειακού τομέα, ενώ η αγορά παραγώγων προσελκύει επενδυτές που ενδιαφέρονται για τη μεταβλητότητα των τιμών.

Αγορά Δικαιωμάτων Μεταφοράς Ηλεκτρικής Ενέργειας (Transmission Market)

Η Αγορά Δικαιωμάτων Μεταφοράς Ηλεκτρικής Ενέργειας (Transmission Market) αφορά τη διαχείριση και διαπραγμάτευση δικαιωμάτων πρόσβασης στο δίκτυο. Τέτοια είναι το δικαίωμα μεταφοράς ισχύος από ένα σημείο του δικτύου σε ένα άλλο δεσμεύοντας για εκείνο το χρονικό διάστημα κάποια χωρητικότητα, το δικαίωμα έγχυσης ισχύος στο δίκτυο και το δικαίωμα εξαγωγής ισχύος από αυτό. Η αξία κάθε δικαιώματος έγκειται στην εξασφάλιση που παρέχει στον κάτοχο του να κάνει χρήση του δικτύου, ιδίως σε περιπτώσεις συμφόρησης αλλά και στην δυνατότητα που του παρέχει να εμπορευτεί το συγκεκριμένο δικαίωμα πουλώντας το σε κάποιον τρίτο. Οι συναλλαγές πραγματοποιούνται στις δημοπρασίες δικαιωμάτων μεταφοράς οι οποίες χωρίζονται σε πρωτογενείς και δευτερογενείς ενώ, κάθε δικαίωμα που διαπραγματεύεται ορίζεται από τρία μεγέθη. Το ποσό της ενέργειας, την τιμή και τέλος τα σημεία έγχυσης και απορρόφησης στο δίκτυο.

Η πρωτογενής αγορά δικαιωμάτων μεταφοράς, πραγματοποιείται σε ετήσια βάση και αφορά την αρχική διάθεση των δικαιωμάτων από τον διαχειριστή του δικτύου μεταφοράς, που είναι ο ΑΔΜΗΕ. Ο διαχειριστής βασίζεται στα αναμενόμενα φορτία αιχμής, δηλαδή σε προβλέψεις φορτίου, για να καθορίσει τη διαθέσιμη χωρητικότητα προς πώληση. Έτσι, δημιουργείται μακροπρόθεσμη ασφάλεια, καθώς με την διάθεση ετήσιων δικαιωμάτων οι συμμετέχοντες εξασφαλίζουν πρόσβαση στο δίκτυο για μεγάλες χρονικές περιόδους και παράλληλα επιτρέπει σε μεγάλους καταναλωτές όπως μεγάλες βιομηχανικές εγκαταστάσεις όπου η πρόσβαση σε ενέργεια είναι ζωτικής σημασίας, να καταστρώσουν τον ενεργειακό τους σχεδιασμό με προβλεψιμότητα.

Η δευτερεύουσα αγορά αποτελεί μια πιο ευέλικτη και δυναμική αγορά για τη μεταπώληση των δικαιωμάτων μεταφοράς που έχουν ήδη διατεθεί στην πρωτογενή αγορά και τελικά δεν εξαργυρώνονται. Ταυτόχρονα προσφέρουν στους συμμετέχοντες τη δυνατότητα να αποκτήσουν δικαιώματα για σύντομο χρονικό διάστημα, ανάλογα με την άμεση ζήτηση που μπορεί να προκύψει από μεταβαλλόμενες συνθήκες του δικτύου. Οι συναλλαγές σε αυτή την αγορά μπορούν να γίνουν σε μηνιαία, ημερήσια ή ακόμα και ενδοημερήσια βάση.

Συμπερασματικά η πρωτογενής αγορά εξυπηρετεί μακροπρόθεσμες ανάγκες και παρέχει σταθερότητα σε μεγάλους καταναλωτές και παραγωγούς ενέργειας. Αντίθετα, η δευτερεύουσα αγορά είναι πιο δυναμική και επικεντρώνεται στην ευελιξία, διευκολύνοντας τις βραχυπρόθεσμες συναλλαγές και τη ρύθμιση των άμεσων αναγκών.

Και οι δύο αγορές λειτουργούν συμπληρωματικά, διασφαλίζοντας τη βέλτιστη χρήση της δυναμικότητας του δικτύου μεταφοράς και την οικονομική αποδοτικότητα του συστήματος.

Αγορά Επικουρικών Υπηρεσιών (Ancillary Services Market)

Η Αγορά Επικουρικών Υπηρεσιών συνδράμει στην αξιοπιστία του δικτύου εξασφαλίζοντας την ποιότητα της ενέργειας. Ο ΑΔΜΗΕ είναι υπεύθυνος για την εποπτεία της αγοράς, τη διαχείριση των προσφορών και την επιλογή των βέλτιστων λύσεων για το δίκτυο.

Σύμφωνα με τον «Κώδικα Διαχείρισης του Συστήματος και Συναλλαγών Ηλεκτρικής Ενέργειας» οι Επικουρικές Υπηρεσίες είναι οι εξής :

- Πρωτεύουσα Ρύθμιση και Εφεδρεία
- Δευτερεύουσα Ρύθμιση και Εύρος
- Τριτεύουσα Ρύθμιση και Στρεφόμενη Εφεδρεία
- Τριτεύουσα Μη Στρεφόμενη Εφεδρεία
- Στατή εφεδρεία
- Ρύθμιση Τάσης
- Επανεκκίνηση του Συστήματος

Οι πρώτες τέσσερις έχουν ρόλο στην ρύθμιση της συχνότητας, στην ονομαστική τιμή στα 50Hz και στην διαχείριση της ενεργού ισχύος του δικτύου σε περίπτωση που υπάρξουν αποκλίσεις στην ισορροπία μεταξύ παραγωγής και κατανάλωσης.

Η πρωτεύουσα ρύθμιση και εφεδρεία ενεργοποιείται αυτόματα μέσα σε δευτερόλεπτα όταν η συχνότητα του δικτύου αποκλίνει από την ονομαστική τιμή. Αφορά τη χρήση της παραγόμενης ενεργού ισχύος από στρεφόμενες μάζες και γεννήτριες για τη γρήγορη αποκατάσταση της ισορροπίας μεταξύ παραγωγής και κατανάλωσης. Ενεργοποιείται αυτόματα και δεν απαιτεί ανθρώπινη παρέμβαση.

Η δευτερεύουσα ρύθμιση ενεργοποιείται λιγότερο άμεσα συνήθως σε χρονικό διάστημα εντός λεπτών και μπορεί να διαρκέσει από δέκα δευτερόλεπτα έως δεκαπέντε λεπτά. Σκοπός της είναι η αντικατάσταση των πρωτευουσών εφεδρειών όταν δεν επαρκούν μετά από μεγαλύτερες αποκλίσεις που δεν μπορούν να διορθωθούν εξ ολοκλήρου από την πρωτεύουσα ρύθμιση. Ελέγχεται από κεντρικά συστήματα του διαχειριστή δικτύου ο οποίος ενεργοποιεί σταθμούς παραγωγής ή μονάδες αποθήκευσης ενέργειας που έχουν χαρακτηριστεί ως εφεδρικές δυνάμεις οι οποίες αποζημιώνονται και για τη διαθεσιμότητά τους αλλά και για την πραγματική ενεργοποίησή τους.

Η τριτεύουσα ρύθμιση αποτελεί το τελευταίο στάδιο στην διαδικασία εξισορρόπησης του ηλεκτρικού συστήματος και ενεργοποιείται μετά τη δευτερεύουσα ρύθμιση. Σκοπός της είναι διατήρησης της ισορροπίας του ηλεκτρικού συστήματος και εξασφαλίζει ισχύ για μεγαλύτερο χρονικό διάστημα με την αντικατάσταση των προσωρινών ρυθμίσεων από μόνιμες λύσεις και η

αποκατάσταση των αποθεμάτων εφεδρείας. Ο χρόνος απόκρισης της είναι μεγαλύτερος, ξεκινώντας από 15 λεπτά και άνω, καθιστώντας την απαραίτητη για τη διασφάλιση της μακροχρόνιας ευστάθειας του δικτύου. Κατά την εφαρμογή της Ενεργοποιούνται σταθμοί παραγωγής με πιο αργή απόκριση, όπως θερμικοί σταθμοί, και μονάδες φυσικού αερίου. Ακόμη μέσω συμβάσεων διακοπτόμενης ζήτησης ο διαχειριστής μπορεί να αναγκάσει μεγάλους καταναλωτές όπως βιομηχανίες να μειώσουν την ποσότητα της ενέργειας που χρησιμοποιούν, έναντι μίας αποζημίωσης.

Στην Τριτεύουσα Μη Στρεφόμενη Εφεδρεία ανήκουν μονάδες παραγωγής ενέργειας, οι οποίες δεν λειτουργούν σε συνεχή βάση αλλά μπορούν να ενεργοποιηθούν γρήγορα σε περίπτωση έκτακτης ανάγκης.

Τέτοιες είναι μονάδες φυσικού αερίου, υδροηλεκτρικοί σταθμοί με αντλησιοταμίευση, μονάδες που λειτουργούν με καύση λιγνίτη ή πετρελαίου, μονάδες ΑΠΕ με αποθήκευση, εφεδρικές μονάδες έκτακτης ανάγκης (π.χ. ντιζελοκινητήρες)

Ως στατική εφεδρεία (static reserve) ορίζεται το σύνολο της διαθέσιμης ισχύος που μπορεί να παρέχεται από μονάδες παραγωγής ή άλλες πηγές ενέργειας οι οποίες βρίσκονται σε εφεδρεία, αλλά δεν λειτουργούν ενεργά τη δεδομένη στιγμή.

Ως Ρύθμιση Τάσης ορίζεται η διαδικασία Σταθεροποίησης της τάσης για αποφυγή ζημιών στο δίκτυο και τις συσκευές. Αυτό πραγματοποιείται με παροχή άεργης ισχύος από μονάδες παραγωγής ή άλλες πηγές.

Κατά την Επανεκκίνηση Συστήματος (Black Start) σε περίπτωση πλήρους διακοπής ειδικές μονάδες μπορούν να επανεκκινήσουν το δίκτυο χωρίς εξωτερική παροχή ενέργειας [1,2, 3].

1.5 Λειτουργία της αγοράς

Η αγορά ενέργειας λειτουργεί σε καθημερινή βάση με βάση τις διαδικασίες που έχουν περιγραφεί προηγουμένως. Στην παρούσα ενότητα, παρουσιάζονται οι διαδικασίες αυτές με τη χρονική σειρά που λαμβάνουν χώρα, προκειμένου να αποσαφηνιστεί πλήρως η λειτουργία της αγοράς ενέργειας.

Την προηγούμενη ημέρα, η Αγορά Επόμενης Ημέρας καθορίζει τις ποσότητες και τις τιμές της ενέργειας που θα διακινηθούν την επόμενη ημέρα, σε ωριαία βάση. Οι συμμετέχοντες υποβάλλουν τις προσφορές τους έως το μεσημέρι της προηγούμενης ημέρας. Στο πλαίσιο αυτής της διαδικασίας, ο Ανεξάρτητος Διαχειριστής Μεταφοράς Ηλεκτρικής Ενέργειας (ΑΔΜΗΕ) εκτελεί το Πρόγραμμα Κατανομής, καθορίζοντας την ακριβή σειρά ενεργοποίησης των μονάδων παραγωγής για την κάλυψη της ζήτησης.

Παράλληλα, η διαχείριση των διασυννοριακών συναλλαγών μέσω του μηχανισμού Market Coupling συνδυάζεται με την Αγορά Επόμενης Ημέρας, ώστε να καθοριστούν οι διασυννοριακές ροές ενέργειας και η χρήση των διασυνδέσεων με τρίτες χώρες.

Κατά τη διάρκεια της ημέρας, η Ενδοημερήσια Αγορά παρέχει τη δυνατότητα προσαρμογών, διεξάγοντας διαδοχικές συνεδρίες που επιτρέπουν στους συμμετέχοντες να βελτιώνουν τις θέσεις τους. Ταυτόχρονα, η Αγορά Επικουρικών Υπηρεσιών λειτουργεί, με τις μονάδες παραγωγής να προσφέρουν υπηρεσίες, όπως εφεδρεία και ρύθμιση συχνότητας, οι οποίες ενδέχεται να ενεργοποιηθούν κατά τη διάρκεια της ημέρας. Επιπλέον, η Αγορά Εξισορρόπησης λειτουργεί σε πραγματικό χρόνο, εξασφαλίζοντας την ισορροπία μεταξύ παραγωγής και ζήτησης. Οι αποκλίσεις που προκύπτουν σε αυτήν την αγορά υποβάλλονται σε διαδικασία εκκαθάρισης, κατά την οποία υπολογίζονται τα κόστη και οι πιστώσεις για τους συμμετέχοντες.

Καθ' όλη τη διάρκεια της ημέρας, ο ΑΔΜΗΕ παρακολουθεί και ελέγχει το σύστημα, διασφαλίζοντας τη σταθερότητα του δικτύου. Μετά τη λήξη της ημέρας, διεξάγονται οι εκκαθαρίσεις των αγορών, προκειμένου να καθοριστούν οι χρηματικές υποχρεώσεις και απαιτήσεις των συμμετεχόντων σε όλες τις αγορές.

Μακροπρόθεσμα, ο μηχανισμός διασφάλισης επαρκούς ισχύος μέσω εφεδρειών εξασφαλίζει τη διαθεσιμότητα επαρκούς ισχύος στο σύστημα, καλύπτοντας περιόδους υψηλής ζήτησης ή έκτακτες ανάγκες. Επιπρόσθετα, η Προθεσμιακή Αγορά διευκολύνει τη διαπραγμάτευση μελλοντικών συμβολαίων, ενισχύοντας τη σταθερότητα και την προβλεψιμότητα της αγοράς.

1.6 Η αξία της πρόβλεψης φορτίου ενέργειας επόμενης ημέρας

Η ηλεκτρική ενέργεια δεν μπορεί να αποθηκευτεί σε μεγάλες ποσότητες και για αυτό η ικανότητα πρόβλεψης των αναγκών φορτίου ενέργειας για την επόμενη ημέρα σε ωριαία βάση, είναι κρίσιμη για τον σωστό προγραμματισμό. Μέσω της ακριβούς πρόβλεψης, διασφαλίζεται ο σωστός προγραμματισμός της παραγωγής και της κατανάλωσης, αποτρέποντας φαινόμενα ανισορροπίας στο δίκτυο. Η ακρίβεια στις προβλέψεις όχι μόνο ενισχύει την αποδοτική και αξιόπιστη παροχή ενέργειας, αλλά συμβάλλει και στη μείωση του κόστους για τους καταναλωτές.

Αξιοσημείωτο είναι πως Ολόκληρη η αγορά επόμενης ημέρας, η οποία αποτελεί τη ραχοκοκαλιά της αγοράς ενέργειας, βασίζεται στην δυνατότητα της πρόβλεψης φορτίων.

Οι προβλέψεις φορτίου της επόμενης ημέρας συμβάλλουν στη σταθερότητα του δικτύου μέσω της εκτίμησης της ζήτησης για ηλεκτρική ενέργεια και της αντίστοιχης κατανομής πόρων. Συγκεκριμένα, βελτιστοποιείται ο προγραμματισμός της παραγωγής και έτσι ο διαχειριστής του συστήματος μπορεί να προγραμματίσει ποια μονάδα παραγωγής θα ενεργοποιηθεί και σε ποιο χρονικό διάστημα, διασφαλίζοντας την ισορροπία μεταξύ παραγωγής και

ζήτησης. Σε περιπτώσεις υπερπαραγωγής, παρατηρείται σπατάλη πόρων και μείωση της αποδοτικότητας, ενώ σε περιπτώσεις υποπαραγωγής ενδέχεται να υπάρξουν ζημιές στο σύστημα με πτώση της συχνότητας, θερμικές βλάβες στις γραμμές εξαιτίας υπερβολικής φόρτισης ή ακόμη και διακοπές ρεύματος. Ακόμη, προβλέποντας περιόδους υψηλής ζήτησης, μπορούν να βρίσκονται σε αναμονή περισσότερες εφεδρικές μονάδες για να αντιμετωπιστούν ξαφνικές αιχμές στην ζήτηση. Με την ακριβή πρόβλεψη φορτίου αποφεύγονται αυτά τα προβλήματα και ενισχύεται η ευστάθεια και η ομαλή λειτουργία του δικτύου.

Ομοίως οι προβλέψεις φορτίου επόμενης ημέρας συμβάλλουν στη μείωση του κόστους ενέργειας μέσω του αποτελεσματικού σχεδιασμού και της βέλτιστης χρήσης των διαθέσιμων πόρων. Οι προβλέψεις φορτίου επιτρέπουν στον διαχειριστή να προτεραιοποιήσει σωστά τη λειτουργία των μονάδων παραγωγής με το χαμηλότερο κόστος, μειώνοντας έτσι το κόστος παραγωγής. Αντίστοιχα, η ακριβής πρόβλεψη μειώνει τις αποκλίσεις μεταξύ προβλεπόμενης και πραγματικής ζήτησης με συνέπεια την μειωμένη ανάγκη για ενεργοποίηση ακριβών υπηρεσιών εξισορρόπησης που αυξάνουν το συνολικό κόστος του συστήματος. Ένας ακριβής ενεργειακός προγραμματισμός, επιτρέπει στο σύστημα να λειτουργεί με μικρότερο απόθεμα εφεδρικής ισχύος, εξοικονομώντας πόρους και μειώνοντας το περιβαλλοντικό αποτύπωμα. Εξίσου σημαντική είναι η βελτιστοποίηση του σχεδιασμού των ροών μεταφοράς στο δίκτυο περιορίζοντας έτσι τις απώλειες μεταφοράς, καθώς μικρότερες απώλειες οδηγούν σε χαμηλότερο κόστος για την ίδια ποσότητα ενέργειας. Τέλος, η ακριβής πρόβλεψη συμβάλλει στην αποφυγή ακραίων φαινομένων, όπως οι ξαφνικές αυξήσεις τιμών. Τέτοιες αυξήσεις συχνά οφείλονται σε ανεπαρκή προγραμματισμό ή σε αιφνίδιες αλλαγές της ζήτησης. Με την καλύτερη πρόβλεψη των αναγκών φορτίου, οι διαχειριστές της αγοράς μειώνουν την πιθανότητα εμφάνισης τέτοιων διαταραχών διασφαλίζοντας την οικονομική σταθερότητα.

Ακόμη πρέπει να τονιστεί πως η μετάβαση στις Ανανεώσιμες Πηγές Ενέργειας, όπως η ηλιακή και η αιολική ενέργεια, αυξάνει τη σημασία της ακριβούς πρόβλεψης φορτίου. Οι πηγές αυτές, αν και φιλικές προς το περιβάλλον, παρουσιάζουν σημαντική μεταβλητότητα λόγω της εξάρτησής τους από τις καιρικές συνθήκες. Αυτή η διαλείπουσα και στοχαστική φύση της παραγωγής μπορεί να προκαλέσει διακυμάνσεις στην παροχή της συνολικής ενέργειας στο σύστημα [22]. Με την ακριβή πρόβλεψη του φορτίου οι διαχειριστές έχουν τη δυνατότητα να προγραμματίζουν με ακρίβεια την ενσωμάτωση της πράσινης ενέργειας στο ηλεκτρικό δίκτυο. Με αυτόν τον τρόπο, μειώνεται ο κίνδυνος ανισορροπίας μεταξύ παραγωγής και κατανάλωσης, εξασφαλίζοντας τη σταθερότητα του συστήματος. Επιπλέον, οι προβλέψεις βοηθούν στη βέλτιστη χρήση των θερμικών και εφεδρικών μονάδων, ενεργοποιώντας τις μόνο όταν είναι απαραίτητο, γεγονός που μειώνει το κόστος λειτουργίας και ενισχύει την αποδοτικότητα του ενεργειακού συστήματος. Η πρόβλεψη φορτίου, συνεπώς, καθίσταται κρίσιμη για τη βιώσιμη και οικονομική ενσωμάτωση των ΑΠΕ.

Κλείνοντας η βελτιστοποίηση της πρόβλεψης ενεργειακών αναγκών ενισχύει τον ανταγωνισμό και την διαφάνεια της αγοράς. Η πληροφόρηση σύμφωνα με τις

πραγματικές ανάγκες της αγοράς επιτρέπει στους συμμετέχοντες- παραγωγούς, προμηθευτές και διαχειριστές- να προσαρμόζουν τον στρατηγικό τους σχεδιασμό, να συμμετέχουν πιο αποτελεσματικά στις διάφορες αγορές και να κάνουν καλύτερες προσφορές.

Συνεπώς, εξασφαλίζεται σταθερότητα στην αγορά καθώς μειώνεται η μεταβλητότητα στις τιμές, προάγεται η αποδοτική κατανομή της ενέργειας και διασφαλίζεται η μακροπρόθεσμη βιωσιμότητα του ενεργειακού συστήματος με την δημιουργία ασφαλούς περιβάλλοντος για επενδύσεις.

Συνοψίζοντας, η πρόβλεψη των αναγκών φορτίου στην αγορά επόμενης ημέρας είναι ένας από τους σημαντικότερους παράγοντες για τη σωστή λειτουργία του ηλεκτρικού συστήματος. Βελτιώνει την ισορροπία προσφοράς και ζήτησης, ενισχύει την οικονομική αποδοτικότητα, υποστηρίζει την ενσωμάτωση ανανεώσιμων πηγών και μειώνει το κόστος για όλους τους εμπλεκόμενους. Παράλληλα, διασφαλίζει τη διαφάνεια, την αξιοπιστία και την προσαρμοστικότητα του συστήματος, καθιστώντας τη θεμέλιο λίθο για την αποτελεσματική διαχείριση της αγοράς ενέργειας.

1.7 Τεχνικές πρόβλεψης φορτίου επόμενης ημέρας στην Ευρώπη

Γενικότερα, πρέπει να τονιστεί ότι δεν υπάρχει ενιαία μέθοδος πρόβλεψης που να καλύπτει κάθε πιθανή εφαρμογή στον τομέα της ενέργειας, και συνεπώς εφαρμόζονται διαφορετικές προσεγγίσεις ανάλογα με τον σκοπό. Οι προβλέψεις στον τομέα της ηλεκτρικής ενέργειας μπορούν να ταξινομηθούν με βάση τρεις διαστάσεις. Πρώτον, το αντικείμενο της πρόβλεψης, όπως για παράδειγμα η πρόβλεψη παραγωγής, η πρόβλεψη καταναλισκόμενου φορτίου, η πρόβλεψη ζήτησης, η πρόβλεψη της τιμής ενέργειας ή άλλα. Δεύτερον, ο χρονικός ορίζοντας, ο οποίος ορίζει το χρονικό διάστημα για το οποίο γίνεται η πρόβλεψη και διακρίνεται σε τέσσερις κατηγορίες:

- Πολύ βραχυπρόθεσμη (VSTF): έως 1 ώρα
- Βραχυπρόθεσμη (STF): από 1 ημέρα έως 1–2 εβδομάδες
- Μεσοπρόθεσμη (MTF): από 1 έως 3 έτη
- Μακροπρόθεσμη (LTF): έως 30 έτη

Τρίτον, το επίπεδο και το μέγεθος του συστήματος ενέργειας για το οποίο πραγματοποιείται η πρόβλεψη, όπως για παράδειγμα σε διεθνές ολοκληρωμένο δίκτυο, στο ευρωπαϊκό δίκτυο ηλεκτρικής ενέργειας, σε εθνικό δίκτυο μεταφοράς μιας χώρας ή σε επίπεδο φορτίων υποσταθμών.

Στην παρούσα εργασία εξετάζεται η πρόβλεψη φορτίου ενέργειας της επόμενης ημέρας σε ευρωπαϊκό επίπεδο. Σε αυτό το πλαίσιο, η βραχυπρόθεσμη πρόβλεψη φορτίου αποτελεί βασικό εργαλείο για το σύνολο των εργασιών και λειτουργιών, από τη συντήρηση και τον προγραμματισμό εφεδρειών έως τη διασφάλιση της

αδιάλειπτης παροχής, την τιμολόγηση της ενέργειας και τη βελτιστοποίηση του κόστους. Για τον σκοπό αυτό, έχουν διεξαχθεί πολυάριθμες έρευνες για την ανάπτυξη μοντέλων πρόβλεψης, τα οποία περιλαμβάνουν κλασσικά στατιστικά μοντέλα, μοντέλα μηχανικής μάθησης ή υβριδικά μοντέλα που συνδυάζουν και τις δύο προσεγγίσεις.

Στην Ευρωπαϊκή Ένωση, την ευθύνη για τις προβλέψεις ηλεκτρικού φορτίου της επόμενης ημέρας φέρουν κυρίως οι Διαχειριστές Συστημάτων Μεταφοράς κάθε χώρας. Για τον σκοπό αυτό, καταρτίζουν καθημερινά προβλέψεις της συνολικής ζήτησης ηλεκτρικής ισχύος για την επόμενη ημέρα. Οι προβλέψεις αυτές καλύπτουν συνήθως το 24ωρο της επόμενης ημέρας με διακριτοποίηση ανά ώρα (αγοραία χρονική μονάδα), αντισταθμίζοντας το ωριαίο βήμα κατανομής του ευρωπαϊκού ημερήσιου ενεργειακού προγραμματισμού. Οι προβλέψεις ολοκληρώνονται και δημοσιεύονται συνήθως το βράδυ της προηγούμενης ημέρας, ώστε να είναι διαθέσιμες πριν από το κλείσιμο των αγορών της επόμενης ημέρας [26].

Σε διακρατικό επίπεδο, το Ευρωπαϊκό Δίκτυο Διαχειριστών Συστημάτων Μεταφοράς Ηλεκτρικής Ενέργειας (ENTSO-E) συντονίζει τη συλλογή και δημοσίευση αυτών των δεδομένων από όλες τις χώρες. Σύμφωνα με τον Κανονισμό (ΕΕ) 543/2013, οι διαχειριστές κάθε χώρας υποχρεούνται να υποβάλλουν τις προβλέψεις φορτίου για κάθε ζώνη αγοράς (bidding zone) της επόμενης ημέρας. Επιπρόσθετα, οι προβλέψεις αυτές, μαζί με τις αντίστοιχες μετρήσεις πραγματικού φορτίου, δημοσιεύονται στην κεντρική πανευρωπαϊκή πλατφόρμα διαφάνειας του ENTSO-E.

Όσον αφορά τις κλασικές μεθόδους που έχουν επικρατήσει μέχρι σήμερα, μεταξύ των πιο διαδεδομένων τεχνικών συγκαταλέγονται οι αυτοπαλινδρομικές μέθοδοι, όπως το μοντέλο ARMA, το οποίο απαιτεί η χρονοσειρά να είναι στάσιμη, δηλαδή να μην παρουσιάζει τάση ή εποχικότητα. Η στασιμότητα επιτυγχάνεται μέσω της διαφοροποίησης της μη στάσιμης χρονοσειράς. Επιπλέον, το μοντέλο ARIMA μπορεί να εφαρμοστεί σε μη στάσιμες χρονοσειρές, καθώς είναι ικανό να αναπαριστά αλλαγές σε διάφορα μοτίβα δεδομένων, ενώ απαιτεί λιγότερους παραμέτρους για την εκτίμησή του. Τέλος, το μοντέλο SARIMA εισήχθη για την αντιμετώπιση της εποχικότητας, όπως παρουσιάζεται στο. Αυτά τα μοντέλα είναι ευέλικτα και διαχειρίζονται αποτελεσματικά την εποχικότητα και τη μη στασιμότητα των δεδομένων, αλλά δεν ενσωματώνουν εύκολα εξωγενείς μεταβλητές, οι οποίες είναι κρίσιμες για τη βελτιστοποίηση της πρόβλεψης. Παρ' όλα αυτά, έχει αναπτυχθεί το μοντέλο ARIMAX, το οποίο επιτρέπει την εισαγωγή εξωγενών μεταβλητών [27].

Από την άλλη πλευρά, τα τελευταία χρόνια έχουν αναπτυχθεί και δοκιμάζονται μοντέλα μηχανικής μάθησης, τα οποία αξιοποιούν μεγάλους όγκους δεδομένων για την εξαγωγή προτύπων και σχέσεων που δεν είναι άμεσα εμφανείς. Σε αντίθεση με τα παραδοσιακά μοντέλα που βασίζονται σε προκαθορισμένες στατιστικές σχέσεις, αυτά τα βασιζόμενα σε δεδομένα μοντέλα «μαθαίνουν» από τα ιστορικά δεδομένα και βελτιώνουν συνεχώς την ακρίβειά τους.

Χαρακτηριστικές αρχιτεκτονικές περιλαμβάνουν τα τεχνητά νευρωνικά δίκτυα (ANN), και ιδιαίτερα τα Long Short-Term Memory (LSTM) νευρωνικά δίκτυα, τα οποία είναι εξειδικευμένα επαναληπτικά νευρωνικά δίκτυα, ιδανικά για χρονοσειρές, καθώς διατηρούν πληροφορίες από προηγούμενες καταστάσεις και είναι αποτελεσματικά σε προβλέψεις βραχυπρόθεσμων διακυμάνσεων. Ένα μοντέλο βασισμένο σε LSTM και μηχανισμούς προσοχής για οικιακή πρόβλεψη φορτίου παρουσίασε καλύτερη απόδοση σε σύγκριση με σύγχρονες μεθόδους. Επιπρόσθετα, έχουν δοκιμαστεί και άλλοι αλγόριθμοι μηχανικής μάθησης, όπως οι Υποστηρικτικοί Διανυσματικοί Μηχανισμοί (SVM), οι οποίοι έχουν χρησιμοποιηθεί για την εξαγωγή προβλέψεων με συνδυασμό δεδομένων φορτίου, θερμοκρασίας και ημερολογιακών πληροφοριών, αλλά και για την ανίχνευση ανωμαλιών. Άλλα παραδείγματα περιλαμβάνουν αλγορίθμους Random Forest, καθώς και εξελικτικές μεθόδους, όπως οι Γενετικοί Αλγόριθμοι (GA) [27].

Στην προσπάθεια βελτιστοποίησης της πρόγνωσης, συχνά υιοθετούνται υβριδικά μοντέλα που συνδυάζουν τα πλεονεκτήματα διαφορετικών μεθοδολογιών. Για παράδειγμα, μια κοινή προσέγγιση είναι ο συνδυασμός ενός στατιστικού μοντέλου, όπως το ARIMA, το οποίο αποδίδει καλά στη σύλληψη γραμμικών τάσεων και εποχικότητας, με ένα νευρωνικό δίκτυο που μαθαίνει μη γραμμικά υπολειπόμενα μοτίβα. Μέσω αυτού του διαμοιρασμού εργασιών, το υβριδικό σχήμα επιτυγχάνει καλύτερη ακρίβεια από κάθε μοντέλο ξεχωριστά. Μια άλλη προσέγγιση είναι τα μοντέλα ensemble, όπου πολλαπλοί προγνώστες εκπαιδεύονται και τα αποτελέσματά τους συνδυάζονται μέσω τεχνικών όπως bagging, boosting ή stacking. Ένα χαρακτηριστικό παράδειγμα είναι το Random Forest, το οποίο αποτελεί ensemble πολλαπλών δέντρων αποφάσεων. Τα μοντέλα ensemble συχνά μειώνουν τη διακύμανση του σφάλματος και βελτιώνουν τη γενίκευση. Η βιβλιογραφία υποδεικνύει ότι τα υβριδικά και τα μοντέλα ensemble συγκαταλέγονται μεταξύ των πιο αποδοτικών λύσεων για βραχυπρόθεσμες προβλέψεις. Ωστόσο, το κύριο μειονέκτημά τους είναι η αυξημένη υπολογιστική πολυπλοκότητα και το υψηλό υπολογιστικό κόστος. Παρ' όλα αυτά, στο πρόβλημα της πρόβλεψης φορτίου, όπου ακόμη και μια μικρή μείωση του σφάλματος συνεπάγεται σημαντική εξοικονόμηση κόστους για το σύστημα, οι σύνθετες προσεγγίσεις θεωρούνται δικαιολογημένες.

Στην Ευρωπαϊκή Ένωση σήμερα, οι Διαχειριστές Συστημάτων Μεταφοράς (TSOs) εφαρμόζουν ποικιλία μοντέλων για την πρόβλεψη φορτίου της επόμενης ημέρας, ανάλογα με τις ανάγκες και τις τεχνολογικές δυνατότητες κάθε χώρας. Αν και οι TSOs δεν δημοσιοποιούν αναλυτικά τις μεθόδους τους, καθώς οι ακριβείς τεχνικές πρόβλεψης αποτελούν εμπορικά απόρρητα και κρίσιμες επιχειρησιακές πληροφορίες, η βιβλιογραφία παρέχει ορισμένες ενδείξεις.

Στη Γαλλία, η RTE βασίζεται σε Generalized Additive Models (GAM) για τις επίσημες προβλέψεις φορτίου, καθώς αυτά συνδυάζουν ακρίβεια και διαφάνεια. Τα GAM επιτρέπουν την ευέλικτη, μη γραμμική μοντελοποίηση της σχέσης μεταξύ της ζήτησης και ανεξάρτητων μεταβλητών, όπως η θερμοκρασία, η ημέρα της εβδομάδας και οι τιμές ενέργειας, μέσω συναρτήσεων εξομάλυνσης. Επιπρόσθετα, σε ειδικές συνθήκες, όπως η πανδημία COVID-19, η RTE εφαρμόζει

προσαρμοστικές μεθόδους, όπως τα φίλτρα Kalman, ενώ δοκιμάζει νευρωνικά δίκτυα μέσω του προγράμματος Deep Forecast [28].

Η Terna της Ιταλίας χρησιμοποιεί ένα σύνθετο πολυπαραγοντικό μοντέλο, ενσωματώνοντας θερμοκρασία και κοινωνικοοικονομικούς δείκτες. Για την ωριαία κατανομή φορτίου, εφαρμόζει το Periodic ARMA, μια επέκταση του κλασικού ARMA, όπου οι παράμετροι του μοντέλου, όπως συντελεστές και μέσες τιμές, μεταβάλλονται περιοδικά, π.χ. ανά ώρα ή εποχή. Αυτό επιτρέπει την ακριβή κατανομή του φορτίου σε 24 ωριαίες τιμές, αναγνωρίζοντας ώρες αιχμής και χαμηλής κατανάλωσης [29].

Η TenneT στην Ολλανδία και τη Γερμανία χρησιμοποιεί στατιστικά μοντέλα, όπως το ARIMA, για προβλέψεις σε σταθερές συνθήκες, κυρίως κατά τους θερινούς μήνες. Κατά τη χειμερινή περίοδο, όπου οι διακυμάνσεις της ζήτησης είναι πιο έντονες λόγω χαμηλών θερμοκρασιών, εφαρμόζει το υβριδικό μοντέλο Prophet-LSTM. Το μοντέλο αυτό συνδυάζει το Facebook Prophet, το οποίο χειρίζεται τάσεις, εποχικότητα και ειδικά γεγονότα, με ένα LSTM νευρωνικό δίκτυο για την ανίχνευση μη γραμμικών μοτίβων. Αρχικά, το Prophet παράγει μια βασική πρόβλεψη, ενώ το LSTM εκπαιδεύεται στα υπολειπόμενα σφάλματα, μαθαίνοντας σύνθετες σχέσεις. Οι τελικές προβλέψεις προκύπτουν από το άθροισμα των προβλέψεων των δύο μοντέλων, προσφέροντας αξιόπιστες προβλέψεις ακόμη και σε συνθήκες υψηλής μεταβλητότητας [30].

Η REE της Ισπανίας εφαρμόζει Periodic ARMA για προβλέψεις της επόμενης ημέρας, ενώ για τη μη γραμμική επίδραση της θερμοκρασίας αξιοποιεί splines. Τα splines, μαθηματικές συναρτήσεις που αναπαριστούν μη γραμμικές σχέσεις μέσω πολυωνύμων και επιτρέπουν την ακριβή μοντελοποίηση της σχέσης μεταξύ θερμοκρασίας και ζήτησης, όπου η κατανάλωση παραμένει σταθερή σε θερμοκρασίες 18–22°C και αυξάνεται απότομα σε χαμηλές ή υψηλές τιμές λόγω χρήσης θέρμανσης ή κλιματισμού. Επιπρόσθετα, η REE έχει χωρίσει τη χώρα σε δεκαεννέα γεωγραφικές περιοχές για την πρόβλεψη φορτίου. Αυτή η γεωγραφική κατανομή λαμβάνει υπόψη τοπικές ιδιαιτερότητες, όπως το κλίμα, τον τουρισμό και τις βιομηχανικές δραστηριότητες, ενώ ειδικές προβλέψεις εφαρμόζονται για νησιωτικές περιοχές. Οι επιμέρους προβλέψεις συνδυάζονται για τη διαμόρφωση της συνολικής πρόβλεψης φορτίου της χώρας [27].

Στην Ελλάδα, ο ΑΔΜΗΕ πιθανότατα χρησιμοποιεί παραδοσιακές στατιστικές μεθόδους, όπως μοντέλα ARIMA, εκθετική εξομάλυνση και παλινδρόμηση, για την πρόβλεψη φορτίου της επόμενης ημέρας, αν και οι ακριβείς μέθοδοι δεν είναι δημοσίως γνωστές. Ωστόσο, πρόσφατες τάσεις στον κλάδο υποδεικνύουν στροφή προς προηγμένες τεχνικές μηχανικής μάθησης, όπως νευρωνικά δίκτυα και Gradient Boosting. Το Gradient Boosting, συγκεκριμένα, εκπαιδεύει διαδοχικά δέντρα αποφάσεων για να διορθώνουν τα σφάλματα των προηγούμενων μέσω gradient descent, επιτρέποντας την επεξεργασία μεγάλων συνόλων δεδομένων και την ανίχνευση πολύπλοκων σχέσεων μεταξύ εξωγενών μεταβλητών, όπως η θερμοκρασία ή η ώρα της ημέρας, και της ζήτησης ενέργειας, εξασφαλίζοντας πιο αξιόπιστες προβλέψεις για τη διαχείριση του ενεργειακού συστήματος [31].

Άλλοι διαχειριστές στην ΕΕ δεν δημοσιοποιούν αναλυτικά τις μεθοδολογίες τους, αλλά η γενική τάση είναι παρόμοια. Οι περισσότεροι βασίζονται σε στατιστικές τεχνικές, όπως ARIMA ή πολλαπλή παλινδρόμηση, ενώ τα τελευταία χρόνια δοκιμάζουν τεχνικές μηχανικής μάθησης μέσω ευρωπαϊκών ερευνητικών έργων ή πιλοτικών εφαρμογών. Παράλληλα, οι επιχειρησιακές προβλέψεις τους εξακολουθούν να βασίζονται κυρίως σε στατιστικές μεθόδους.

Εν κατακλείδι, διαφαίνεται μια σαφής τάση μεταξύ των διαχειριστών συστημάτων μεταφοράς της Ευρωπαϊκής Ένωσης να υιοθετούν υβριδικά μοντέλα, τα οποία συνδυάζουν την αξιοπιστία των παραδοσιακών στατιστικών μεθόδων, όπως τα μοντέλα ARIMA και Periodic ARMA, με την ευελιξία των σύγχρονων τεχνικών μηχανικής μάθησης, όπως τα νευρωνικά δίκτυα LSTM, τα Temporal Fusion Transformers και το Gradient Boosting. Αυτή η προσέγγιση επιτρέπει την αναγνώριση πολύπλοκων, μη γραμμικών μοτίβων και σχέσεων μεταξύ εξωγενών μεταβλητών, όπως η θερμοκρασία, οι ημερολογιακές συνθήκες και οι κοινωνικοοικονομικοί δείκτες, και της ζήτησης ενέργειας, ενισχύοντας σημαντικά την ακρίβεια των βραχυπρόθεσμων προβλέψεων φορτίου. Η μετάβαση αυτή αντανάκλα την αυξανόμενη ανάγκη για προσαρμοστικότητα και ακρίβεια στην ευρωπαϊκή ενεργειακή αγορά, ιδιαίτερα στο πλαίσιο της ενσωμάτωσης ανανεώσιμων πηγών ενέργειας και της αντιμετώπισης κρίσιμων προκλήσεων, όπως οι κλιματικές διακυμάνσεις και οι απρόβλεπτες διαταραχές, π.χ. κατά την πανδημία COVID-19. Επιπρόσθετα, η συνεχής ανάπτυξη ευρωπαϊκών ερευνητικών προγραμμάτων και πιλοτικών εφαρμογών υποδηλώνει ότι τα υβριδικά μοντέλα, παρά την υπολογιστική τους πολυπλοκότητα, θα αποτελέσουν τον πυρήνα των μελλοντικών συστημάτων πρόβλεψης, συμβάλλοντας στη βελτιστοποίηση της λειτουργίας των ενεργειακών δικτύων, τη μείωση του κόστους και την ενίσχυση της βιωσιμότητας του ευρωπαϊκού ενεργειακού συστήματος.

ΚΕΦΑΛΑΙΟ 2 – ΤΕΧΝΙΚΕΣ ΠΡΟΒΛΕΨΕΩΝ ^[32]

2.1 Εισαγωγή

Η πρόβλεψη αποτελεί μια θεμελιώδη διαδικασία λήψης αποφάσεων που συνδυάζει επιστημονικές μεθόδους και εμπειρικές γνώσεις, με στόχο την εκτίμηση μελλοντικών τιμών, γεγονότων ή τάσεων. Βασίζεται στην ανάλυση ιστορικών δεδομένων, παρατηρήσεων και τάσεων, ώστε να παράγει ακριβείς εκτιμήσεις για το μέλλον. Οι προβλέψεις αποτελούν το κυριότερο εργαλείο για την διαχείριση και προσπάθεια μείωσης των παραγόντων της αβεβαιότητας και του ρίσκου για το μέλλον. Το πεδίο εφαρμογής τους είναι ευρύ και αγγίζει πολλούς τομείς, όπως η ενέργεια, η οικονομία, η παραγωγή, το περιβάλλον και η επιχειρηματική στρατηγική.

Η χρησιμότητα των προβλέψεων έγκειται στη δυνατότητα που προσφέρουν να καθοδηγήσουν στρατηγικές αποφάσεις και να προετοιμάσουν οργανισμούς, άτομα ακόμα και ολόκληρες κοινωνίες για μελλοντικές προκλήσεις. Για παράδειγμα, στον τομέα της ενέργειας, οι προβλέψεις φορτίου επόμενης ημέρας αποτελούν απαραίτητο εργαλείο για τη διαχείριση ηλεκτρικών δικτύων με την εξασφάλιση της βέλτιστης ισορροπίας μεταξύ παραγωγής και ζήτησης, αλλά και για την διαμόρφωση της πιο ανταγωνιστικής τιμής εκκαθάρισης της αγοράς ενέργειας.

Αν και η έννοια της πρόβλεψης υπήρχε από την αρχαιότητα με παραδείγματα τους χρησμούς των Δελφών στην αρχαία Ελλάδα και τις προβλέψεις αστρολογίας στη Μεσοποταμία, ο κλάδος των προβλέψεων άρχισε να αποκτά επιστημονική υπόσταση κατά τη διάρκεια του 20ού αιώνα. Η ανάπτυξη της στατιστικής επιστήμης, σε συνδυασμό με την ευρύτερη χρήση των μαθηματικών μοντέλων, αποτέλεσε το πρώτο βήμα για την αναλυτική μελέτη και τη δημιουργία συστηματικών μεθόδων δημιουργίας προβλέψεων. Κατά τη διάρκεια της δεκαετίας του 1950, η αύξηση της υπολογιστικής ισχύος και η εισαγωγή ηλεκτρονικών υπολογιστών επέτρεψαν την ανάλυση μεγάλων όγκων δεδομένων, δίνοντας ώθηση στην εφαρμογή στατιστικών και οικονομικών μοντέλων. Αργότερα, η έλευση της τεχνίτης νοημοσύνης και της μηχανικής μάθησης σηματοδότησε μια νέα εποχή για την επιστήμη των προβλέψεων, επιτυγχάνοντας μεγαλύτερη ακρίβεια και προσαρμοστικότητα. Σήμερα, ο τομέας των προβλέψεων αποτελεί έναν ταχέως αναπτυσσόμενο τομέα που η εφαρμογή του θεωρείται ζωτικής σημασίας στους περισσότερους κλάδους της παραγωγής και της οικονομίας.

Στην παρούσα ενότητα θα αναλυθεί Η θεωρία που περιβάλλει τις τεχνικές των προβλέψεων. Θα περιγράψουν εισαγωγικές έννοιες, τα σημαντικότερα χαρακτηριστικά τους, θα αναφερθούν παραδοσιακές μέθοδοι προβλέψεις καθώς και μέθοδοι αξιολόγησης.

2.2 Ποιοτικές μέθοδοι

Οι ποιοτικές μέθοδοι προβλέψεων είναι τεχνικές που βασίζονται σε μη αριθμητικές πληροφορίες αλλά και στην υποκειμενική κρίση, όπως η εμπειρία, η γνώμη ειδικών, και οι διαισθητικές εκτιμήσεις, για την πρόβλεψη μελλοντικών τάσεων ή γεγονότων. Χρησιμοποιούνται κυρίως όταν τα ιστορικά δεδομένα είναι περιορισμένα και αναξιόπιστα ή όταν οι τάσεις αλλάζουν συνεχώς. Παρά τον υποκειμενικό τους χαρακτήρα και την προσεγγιστική τους φύση αποτελούν χρήσιμες μεθόδους σε συνθήκες όπου το μέλλον βασίζεται σε απρόβλεπτους παράγοντες οι οποίοι δεν μπορούν να μοντελοποιηθούν σε επαρκή βαθμό. Τέτοιος παράγοντας είναι για παράδειγμα η ανθρώπινη ψυχολογία.

Παραδείγματα χαρακτηριστικών ποιοτικών μεθόδων αποτελούν :

- Η Δελφική Μέθοδος (Delphi Method), η οποία συνδυάζει τις προβλέψεις ειδικών μέσω επαναλαμβανόμενων γύρων αξιολόγησης για την επίτευξη συναίνεσης.
- Ομαδικές Συζητήσεις και Γνώμες Ειδικών, κατά τις οποίες το αποτέλεσμα της πρόβλεψης γεννάται μέσα από την αξιοποίηση της συνεργασίας και της εμπειρίας εξειδικευμένων ομάδων.
- Ανάλυση Σεναρίων: Κατά την οποία δημιουργούνται πιθανά σενάρια με βάση υπάρχουσες τάσεις και πιθανές εξελίξεις.
- Διαισθητικές Εκτιμήσεις, οι οποίες βασίζονται σε προσωπική κρίση ή εμπειρικές παρατηρήσεις.

2.3 Ποσοτικές μέθοδοι

Οι ποσοτικές μέθοδοι προβλέψεων είναι τεχνικές που βασίζονται στη χρήση μαθηματικών, στατιστικών και υπολογιστικών μοντέλων για την πρόβλεψη μελλοντικών τιμών ή τάσεων, αναλύοντας ιστορικά δεδομένα και αιτιατές σχέσεις. Δεδομένης της παγιωμένης χρήσης δεδομένων, σε αυτές τις μεθόδους απουσιάζει οποιοδήποτε δείγμα υποκειμενικότητας, ενώ αντίθετα χαρακτηρίζονται από ακριβείς και συστηματικές εκτιμήσεις. Δεδομένης της θεμελιώδους εξάρτησης των ποσοτικών μεθόδων από ιστορικά δεδομένα προκύπτει η ιδιομορφία πώς για να είναι πετυχημένες θα πρέπει ο χαρακτήρας των παρελθοντικών δεδομένων που χρησιμοποιήθηκαν για να πραγματοποιηθεί η πρόβλεψη να είναι ίδιος με τον χαρακτήρα των μελλοντικών δεδομένων. Ακόμη

Οι ποσοτικές μέθοδοι αποτελούνται από δύο μεγάλες κατηγορίες : την ανάλυση χρονολογικών σειρών που χρησιμοποιούνται για την ανάλυση παρελθοντικών μοτίβων και τα αιτιοκρατικά μοντέλα που βασίζονται σε σχέσεις αιτίας αποτελέσματος, δηλαδή στην υπόθεση της ύπαρξης μιας σταθερής σχέσης

ανάμεσα στο μέγεθος προς πρόβλεψη, ουσιαστικά μια εξαρτημένη μεταβλητή, και σε ορισμένες ανεξάρτητες παραμέτρους που το επηρεάζουν.

2.4 Χρονοσειρές

Ως χρονοσειρά ορίζεται μια ακολουθία αριθμητικών δεδομένων διακριτού χρόνου, που συλλέγονται και καταγράφονται σε ισαπέχουσες χρονικές στιγμές, σε διαστήματα τα οποία μπορεί να είναι ετήσια, μηνιαία, καθημερινά, ωριαία. Τα δεδομένα μιας χρονοσειράς απεικονίζουν την εξέλιξη μιας μεταβλητής στο χρόνο και συχνά χρησιμοποιούνται για την ανάλυση, τη μοντελοποίηση και την πρόβλεψη μελλοντικών τάσεων. Όσο πιο σύντομο είναι το διάστημα που μεσολαβεί ανάμεσα στα δεδομένα μιας χρονοσειράς, τόσο πιο ακριβής είναι η πληροφορία που αναπαριστά.

Επίσης, ανεξάρτητα από το βαθμό εξάρτησης των παρελθοντικών δεδομένων μεταξύ τους στο πραγματικό κόσμο οι μελλοντικές παρατηρήσεις επηρεάζονται πάντα από διάφορους παράγοντες. Έτσι δεν εξαρτώνται μόνο από τα παλιά τα δικά δεδομένα αλλά και από εξωγενείς παράγοντες όπως η τύχη. Ενδεικτικά στο πρόβλημα της πρόβλεψης ροής φορτίου επόμενης ημέρας κάποιες από τις εξωγενείς μεταβλητές που επηρεάζουν την ποσότητα κατανάλωσης φορτίου είναι οι καιρικές συνθήκες οι οποίες είτε ακολουθούν μία εποχικότητα είτε είναι έκτακτες, αν η ημέρα θεωρείται εργάσιμη ή αργία και ποια χώρα μελετάται.

Τέλος καθώς σε μια χρονοσειρά οι διαδοχικές τιμές είναι εξαρτημένες μεταξύ τους δεν αποτελούν απλώς τη συλλογή δεδομένων που παρατηρούνται σε διαδοχικές χρονικές στιγμές. Αντιθέτως, εκφράζουν τη δυναμική εξέλιξη ενός φαινομένου στο χρόνο και την αλληλεπίδραση των παραμέτρων που το διαμορφώνουν. Έτσι η ουσία των χρονοσειρών έγκειται στο γεγονός ότι επιτρέπουν την βαθύτερη κατανόηση του φαινομένου. Συγκεκριμένα, αντικατοπτρίζουν την αιτιώδη σχέση στο χρόνο, καθώς μέσα από τις χρονοσειρές μπορούμε να κατανοήσουμε τις σχέσεις αίτιου-αποτελέσματος που εξελίσσονται δυναμικά, ενώ επίσης περιγράφουν τη συμπεριφορά συστημάτων στο χρόνο καθώς μέσω αυτών μπορεί κανείς να δει πώς ένα φαινόμενο ανταποκρίνεται σε εξωτερικές αλλαγές.

Ως ανάλυση χρονοσειρών ορίζεται η διαδικασία μελέτης και ερμηνείας μιας χρονοσειράς με στόχο την ερμηνεία της και την αναγνώριση μοτίβων για μελλοντικές προβλέψεις. Τα κύρια χαρακτηριστικά στα οποία αναλύεται μια χρονοσειρά είναι η τάση, η εποχιακότητα, η κυκλικότητα και οι μη κανονικές διακυμάνσεις.

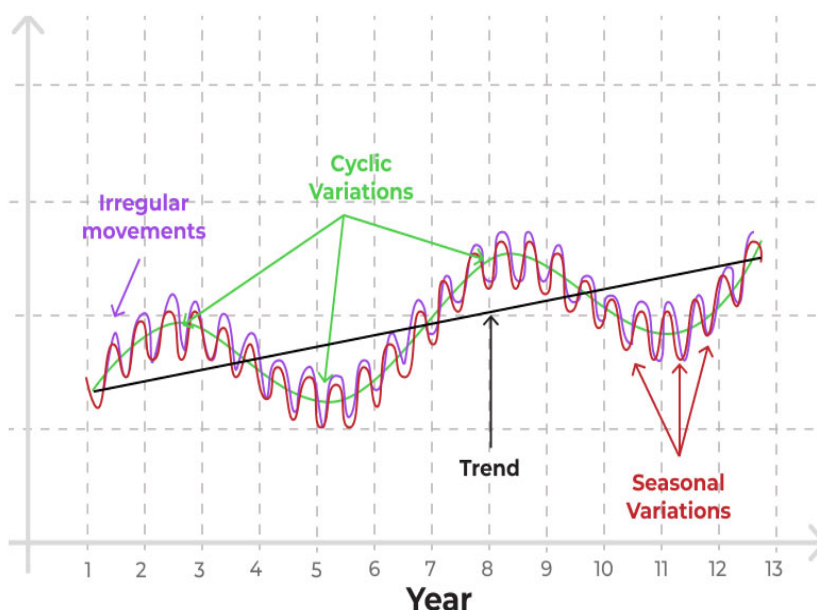
Η τάση είναι το γενικό μοτίβο μιας μακροχρόνιας μεταβολής του μέσου επιπέδου των τιμών της χρονοσειράς μέσα στο χρόνο. Η μεταβολή αυτή παρόλο που μπορεί να είναι γραμμική ή μη γραμμική παρέχει θεμελιώδεις πληροφορίες για τη γενική πορεία ενός φαινομένου.

Ως εποχιακότητα ορίζεται μια περιοδική και επαναλαμβανόμενη διακύμανση η οποία σχετίζεται με συγκεκριμένα χρονικά διαστήματα, μέσα σε μια περίοδο. Συνήθως αυτή η περιοδικότητα είναι κατανοητή και προβλέψιμη. Αυτές οι διακυμάνσεις προκαλούνται από επαναλαμβανόμενους παράγοντες, όπως οι καιρικές συνθήκες, οι πολιτιστικές συνήθειες ή τα οικονομικά γεγονότα.

Η κυκλικότητα αναφέρεται στις επαναλαμβανόμενες διακυμάνσεις που εμφανίζονται σε μια χρονοσειρά οι οποίες δεν ακολουθούν σταθερή χρονική περιοδικότητα, αλλά συνδέονται με ειδικές εξωγενείς συνθήκες. Σε αντίθεση με την εποχικότητα, η κυκλικότητα χαρακτηρίζεται από ακανόνιστη διάρκεια και ένταση. Τα πιο κλασσικά αίτια της κυκλικότητας των χρονοσειρών συνδέονται με οικονομικούς, κοινωνικούς ή φυσικούς κύκλους. Χαρακτηριστικό παράδειγμα ανάδειξης της κυκλικότητας αποτελούν οι οικονομικοί κύκλοι όπως φάσεις ανάπτυξης και ύφεσης, οι οποίες χαρακτηρίζονται από διαδοχικές ανόδους και υφέσεις.

Οι μη κανονικές διακυμάνσεις περιγράφονται επίσης από τον όρο «ασυνέχειες». Πρόκειται για μεμονωμένες και απότομες αλλαγές της γενικής συμπεριφοράς μίας χρονοσειράς οι οποίες είναι απρόβλεπτες καθώς δεν βασίζονται στα επαναλαμβανόμενα μοτίβα της χρονοσειράς. Αυτές οι διακυμάνσεις μπορεί να έχουν σύντομη ή μόνιμη επίδραση. Στην πρώτη περίπτωση περιγράφονται ως outliers ή special events και ουσιαστικά πρόκειται για μία ασυνήθιστη παρατήρηση ενώ, στη δεύτερη περίπτωση περιγράφονται ως level-shifts, αφού εμφανίζονται ως απότομες αλλαγές στο μέσο επίπεδο των τιμών της χρονοσειράς.

TIME SERIES DATA COMPONENTS



CODE IT
BECAUSE WE KNOW HOW

Σχήμα 2.1 : Οπτικοποίηση των χαρακτηριστικών μιας χρονοσειράς.

2.5 Παραδοσιακές μέθοδοι πρόβλεψης

Αφελής μέθοδος - Naïve

Η αφελής μέθοδος είναι μια από τις πιο απλές τεχνικές πρόβλεψης, βασισμένη στην υπόθεση ότι οι μελλοντικές τιμές μιας χρονοσειράς θα είναι ίδιες με τις πιο πρόσφατες παρατηρήσεις. Η βασική της μορφή προβλέπει ότι η τιμή της πρόβλεψης μιας χρονικής στιγμής θα είναι ίση με την τιμή της παρατήρησης την προηγούμενη χρονική στιγμή. Σε εποχικές χρονοσειρές, η πρόβλεψη βασίζεται στην τιμή της ίδιας χρονικής περιόδου του προηγούμενου κύκλου.

Το σημαντικότερο πλεονέκτημα της μεθόδου είναι το ελάχιστο υπολογιστικό κόστος και η απλοϊκή της μορφή. Χρησιμοποιείται ευρέως ως σημείο αναφοράς για τη σύγκριση με πιο πολύπλοκα μοντέλα πρόβλεψης.

$$\hat{Y}_i = Y_{i-1}$$

Μέθοδοι Εξομάλυνσης I - Μέθοδος Κινητού Μέσου Όρου

Οι μέθοδοι εξομάλυνσης δεδομένων αποτελούν σημαντικά εργαλεία για την ανάλυση χρονοσειρών και τη δημιουργία προβλέψεων καθώς αποτελούν διαδεδομένες τεχνικές πρόβλεψης. Χρησιμοποιούνται κυρίως για βραχυπρόθεσμες και μεσοπρόθεσμες προβλέψεις, καθώς επικεντρώνονται στη διαχείριση των τυχαίων διακυμάνσεων και στην ανάδειξη της γενικής συμπεριφοράς των δεδομένων, ενώ αποδίδουν καλύτερα σε δεδομένα που εμφανίζουν σχετική σταθερότητα, χωρίς έντονη εποχικότητα ή σύνθετες τάσεις.

Το σύνολο των μεθόδων εξομάλυνσής διακρίνονται σε δύο μεγάλες κατηγορίες. Πρώτων η μέθοδος κινητού μέσου όρου, κατά την οποία ένας σταθερός αριθμός δεδομένων με ίσα βαρύ χρησιμοποιείται για τον υπολογισμό της πρόβλεψης και δεύτερων η οικογένεια των μεθόδων εκθετικής εξομάλυνσής, κατά την οποία αποδίδονται άνισα βάρη στα ιστορικά δεδομένα, με τα πιο πρόσφατα χρονικά να φέρουν τα μεγαλύτερα.

Ως μέσος όρος στην στατιστική ορίζεται η μέση τιμή ενός συνόλου δεδομένων. Έτσι λοιπόν, η μέθοδος κινητού μέσου όρου στηρίζεται στην εφαρμογή του μέσου όρου ενός σταθερού αριθμού προηγούμενων παρατηρήσεων για τον υπολογισμό της επόμενης πρόβλεψης. Με αυτό τον τρόπο το μοντέλο υπολογίζει το βασικό μοτίβο συμπεριφοράς αλλά και τη γενική τάση των δεδομένων, ελαχιστοποιώντας τυχαίες διακυμάνσεις. Όσο μεγαλύτερη είναι η τιμή του k , δηλαδή του αριθμού των δεδομένων για τον υπολογισμό του μέσου όρου τόσο μεγαλύτερη είναι η επίδραση της εξομάλυνσης αφού μελετάται μία χρονικά γενικότερη εικόνα.

$$\hat{Y}_{t+1} = \frac{1}{k} \cdot \sum_{i=t-k+1}^t Y(i)$$

Η μέθοδος παρουσιάζει πλεονεκτήματα ως προς την απλότητά της και την ευκολία κατανόησης. Δεν αποτελεί χρηστική μέθοδο παραγωγής προβλέψεων στη πράξη καθώς αδυνατεί να συμπεριλάβει τη συνιστώσα της τάσης στις προβλέψεις. Ωστόσο χρησιμοποιείται κυρίως για την εκτίμηση της γραμμής τάσης-κύκλου αφού εξαλείφει την εποχιακότητα και τις τυχαίες μεταβολές (θόρυβος) από μια χρονοσειρά. Βρίσκει εφαρμογή σε πολλούς τομείς. Μερικά παραδείγματα είναι η επιστήμη των δεδομένων και της μηχανικής μάθησης, όπου χρησιμοποιείται για την εξομάλυνση θορύβου στα δεδομένα, ώστε τα μοντέλα μηχανικής μάθησης να λειτουργούν πιο αποτελεσματικά. Η τεχνική ανάλυση στις χρηματοοικονομικές αγορές όπου, χρησιμοποιείται για την εξομάλυνση τιμών μετοχών και την αναγνώριση τάσεων, παρέχοντας σήματα αγοράς, όπως αγορά ή πώληση μιας μετοχής και άλλα.

Υπάρχουν επεκτάσεις της μεθόδου που ενισχύουν την απόδοση της, όπως ο σταθμισμένος κινητός μέσος όρος, ο οποίος δίνει μεγαλύτερη βαρύτητα στις πιο πρόσφατες παρατηρήσεις, και ο εκθετικός κινητός μέσος όρος, ο οποίος λαμβάνει υπόψη τη χρονική σειρά με εκθετικά μειούμενα βάρη. Οι επεκτάσεις αυτές προσφέρουν μεγαλύτερη ευελιξία, ιδιαίτερα σε δεδομένα που αλλάζουν γρήγορα.

Μέθοδοι Εξομάλυνσης II - Μέθοδος Εκθετικής Εξομάλυνσης

Η διαφορά της μεθόδου εκθετικής εξομάλυνσής σχετικά με το προαναφερθέν μοντέλο έγκειται στη χρήση βαρών με φθίνουσα χρονική διάταξη για τον υπολογισμό του μέσου όρου. Με αυτό τον τρόπο οι πιο πρόσφατες παρατηρήσεις επηρεάζουν περισσότερο την πρόβλεψη μελλοντικών τιμών. Ο τρόπος με τον οποίον μαθηματικά δίνεται μεγαλύτερη βαρύτητα σε αυτές είναι μέσω ενός συντελεστή εξομάλυνσης (α).

Ο αναδρομικός τύπος υπολογισμού της είναι ο εξής:

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha) \hat{Y}_t$$

Η τιμή του συντελεστή εξομάλυνσης - α ελέγχει την επιρροή των πρόσφατων δεδομένων. Όταν αυτή είναι κοντά στο 1, δίνεται μεγαλύτερη βαρύτητα στις πιο πρόσφατες παρατηρήσεις, ενώ όταν είναι κοντά στο 0, το μοντέλο δίνει περισσότερη έμφαση στις παλαιότερες τιμές.

Η μέθοδος εκθετικής εξομάλυνσης χρησιμοποιείται ευρέως για την πρόβλεψη βραχυπρόθεσμων τάσεων, τη διαχείριση αποθεμάτων και τη μελέτη χρονοσειρών που εμφανίζουν σταθερή συμπεριφορά. Είναι απλή στην εφαρμογή και εύκολα κατανοητή, ενώ προσαρμόζεται γρήγορα σε αλλαγές στα δεδομένα μέσω της δυναμικής χρήσης του συντελεστή εξομάλυνσης. Ωστόσο, έχει ορισμένους περιορισμούς. Δεν αποδίδει καλά σε χρονοσειρές με έντονη εποχικότητα ή μακροπρόθεσμες τάσεις, και η απόδοση της εξαρτάται σημαντικά από τη σωστή επιλογή του συντελεστή. Επίσης, δεν προσαρμόζεται αποτελεσματικά σε ξαφνικές διακυμάνσεις των δεδομένων.

Για την αντιμετώπιση αυτών των περιορισμών, υπάρχουν επεκτάσεις της βασικής μεθόδου που χωρίζονται σε κατηγορίες αναλόγως του μοτίβου τάσης που ακολουθούν τα δεδομένα. Αναλυτικά, όταν η χρονοσειρά δεν παρουσιάζει κάποια τάση χρησιμοποιείται το μοντέλο σταθερού επιπέδου ή αλλιώς μοντέλο απλής εκθετικής εξομάλυνσης, που εφαρμόζεται σε δεδομένα χωρίς τάση και βασίζεται μόνο στην πρόβλεψη του επιπέδου, δηλαδή της μέσης τιμής των δεδομένων. Επιπρόσθετα, στην περίπτωση που τα δεδομένα της χρονοσειράς παρουσιάζουν γραμμική τάση χρησιμοποιείται το μοντέλο διπλής εκθετικής εξομάλυνσης - μοντέλο Holt, το οποίο εκτός από το επίπεδο λαμβάνει υπόψιν και την τάση, δηλαδή την κατεύθυνση και την κλίση των δεδομένων. Αυτό εφαρμόζεται κυρίως για μεσοπρόθεσμες προβλέψεις. Τέλος έχει παρατηρηθεί πως το γραμμικό μοντέλο δεν είναι ακριβές σε μακροχρόνιες περιόδους καθώς οι προβλέψεις του είναι υπερεκτιμημένες. Γι' αυτό δημιουργήθηκε το μοντέλο μη γραμμικής τάσης το οποίο προσαρμόζεται σε δεδομένα με πιο πολύπλοκη τάση. Εδώ παρατηρείται αύξων ή φθίνων ρυθμός μεταβολής, με την τάση σχεδιαστικά να μην ακολουθεί μια ευθεία γραμμή αλλά μια καμπύλη. Αυτό το μοντέλο χρησιμοποιείται κυρίως για μακροπρόθεσμες προβλέψεις.

Αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου (ARIMA)

Τα αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου ϕ αποτελούν μία πολύ δημοφιλή μέθοδο πρόβλεψης χρονοσειρών. Η βασική ιδέα λειτουργίας τους είναι η αναγνώριση μοτίβων στα δεδομένα, όπως τάσεις, εποχικές διακυμάνσεις ή τυχαίες αλλαγές, και η χρήση τους για την εξαγωγή αξιόπιστων προβλέψεων.

Σε αυτά τα μοντέλα συνδυάζονται τρεις βασικές ιδέες, από τις οποίες προέρχεται και η ονομασία AR-I-MA. Πρώτη είναι η αυτοπαλινδρόμηση (Autoregressive-AR), κατά την οποία η πρόβλεψη υπολογίζεται με βάση τον γραμμικό συνδυασμό ενός συγκεκριμένου αριθμού προηγούμενων τιμών της χρονοσειράς συνυπολογιζόμενου και ενός τυχαίου παράγοντα-θορύβου. Το πλήθος των προηγούμενων τιμών ορίζεται από την παράμετρο p . Δεύτερη συνιστώσα είναι η διαφορίση (I), κατά την οποία αφαιρείται η τάση και η εποχικότητα μετατρέποντας μια χρονοσειρά σε στάσιμη. Μαθηματικά αυτό εκφράζεται με τον υπολογισμό της διαφοράς μεταξύ κάθε τιμής με την αμέσως προηγούμενη της σειράς και έτσι υπολογίζεται η διαφορίση πρώτης τάξης. Αντίστοιχα στη διαφορίση δεύτερης τάξης επαναλαμβάνεται η ίδια διαδικασία στα δεδομένα που προέκυψαν κατά την πρώτη διαφορίση. Η παράμετρος d ορίζει τον βαθμό της διαφορίσης. Η τελευταία συνιστώσα της διαδικασίας είναι ο κινητός μέσος όρος (MA), όπου υποδεικνύει ότι το σφάλμα της πρόβλεψης στο παλινδρομικό μοντέλο είναι γραμμικός συνδυασμός των άλλων προηγούμενων σφαλμάτων πρόβλεψης. Η παράμετρος q ορίζει το πλήθος των προηγούμενων σφαλμάτων που λαμβάνονται υπόψιν από το μοντέλο.

Τα μοντέλα ARIMA είναι στοχαστικά μοντέλα καθώς λειτουργούν με την γενική ιδέα πως η τρέχουσα τιμή μιας σειράς επηρεάζεται όχι μόνο από τυχαίους παράγοντες, δηλαδή τον θόρυβο που συνυπολογίζεται στην διαδικασία της αυτοπαλινδρόμησης - AR, αλλά και από σφάλματα προηγούμενων προβλέψεων

τα οποία προσμετρώνται στο μοντέλο κινητού μέσου όρου -MA. Αυτό σημαίνει ότι, εκτός από μια κεντρική πρόβλεψη, παρέχουν και διαστήματα εμπιστοσύνης, τα οποία δείχνουν το εύρος πιθανών τιμών. Αυτή η ιδιότητα τα κάνει ιδιαίτερα χρήσιμα σε πραγματικά σενάρια, όπου απρόβλεπτοι εξωτερικοί παράγοντες όπως οι καιρικές συνθήκες μπορούν να επηρεάσουν δραστικά τα δεδομένα.

Πρόκειται για αρκετά αξιόπιστα και συγχρόνως ευέλικτα μοντέλα αν παραμετροποιηθούν κατάλληλα. Χρησιμοποιούνται κυρίως σε βραχυπρόθεσμες προβλέψεις στις οποίες είναι αρκετά ακριβή. Επίσης βασικό πλεονέκτημα τους είναι η ικανότητά τους να διαχειρίζονται σειρές που δεν είναι σταθερού επιπέδου, δηλαδή σειρές που εμφανίζουν τάση ή εποχικότητα. Σε αυτή τη περίπτωση απαιτείται προσεκτική προετοιμασία. Αρχικά, αφού η σειρά αναλυθεί ως προς την τάση και την εποχικότητα επιλέγεται ο κατάλληλος βαθμός διαφορίσης. Στη συνέχεια, προσαρμόζονται οι παράμετροι που καθορίζουν πόσες προηγούμενες τιμές και πόσα προηγούμενα σφάλματα από τις προβλέψεις θα ληφθούν υπόψη. Η διαδικασία αυτή, αν και χρονοβόρα, εξασφαλίζει ακριβείς προβλέψεις και μικρές αστοχίες.

Ωστόσο παρουσιάζουν και κάποια μειονεκτήματα τα οποία περιορίζουν και δυσχεραίνουν τη χρήση τους. Αρχικά απαιτούν μεγάλο όγκο ιστορικών δεδομένων για να αποδώσουν αποτελεσματικά και έτσι οι προβλέψεις τείνουν να είναι λιγότερο αξιόπιστες σε περιπτώσεις όπου τα δεδομένα είναι περιορισμένα σε αριθμό. Επιπλέον, η απόδοση τους βασίζεται στην υπόθεση ότι οι σχέσεις μεταξύ των παρατηρήσεων είναι γραμμικές. Σε πιο περίπλοκες χρονοσειρές με μη γραμμικές σχέσεις, τα ARIMA μοντέλα συχνά δεν μπορούν να αποδώσουν καλά, με αποτέλεσμα να υπολείπονται σε ακρίβεια. Η παραμετροποίηση τους μπορεί επίσης να είναι σύνθετη και απαιτεί εξειδικευμένη γνώση, καθώς ο προσδιορισμός των κατάλληλων παραμέτρων AR (p), I (d) και MA (q) ενδέχεται να μην είναι προφανής. Τέλος, για χρονοσειρές με μεγάλο όγκο δεδομένων ή πολύπλοκες διακυμάνσεις, η εκπαίδευση και η προσαρμογή τους μπορεί να απαιτήσει σημαντική υπολογιστική ισχύ, αυξάνοντας το κόστος και την πολυπλοκότητα της διαδικασίας.

Συνοψίζοντας, αυτά τα μοντέλα αποτελούν ένα ισχυρό εργαλείο για την πρόβλεψη χρονικών σειρών, ιδιαίτερα όταν υπάρχει ανάγκη για ευελιξία και ρεαλιστική απεικόνιση της αβεβαιότητας. Η ικανότητά τους να προσαρμόζονται σε διαφορετικά δεδομένα και οι ευρείες εφαρμογές τους τα καθιστούν απαραίτητα στη σύγχρονη ανάλυση δεδομένων.

Αιτιοκρατικό μοντέλο – μέθοδοι παλινδρόμησης

Το αιτιοκρατικό μοντέλο προβλέψεων είναι μια μέθοδος πρόβλεψης που βασίζεται στη σχέση αιτίου-αιτιατού. Αυτή αποτελεί την θεμελιώδη σύνδεση μεταξύ μιας αιτίας, που περιγράφεται από ένα σύνολο ανεξάρτητων μεταβλητών, και ενός αποτελέσματος, μιας εξαρτημένης μεταβλητής δηλαδή οι τιμές της οποίας καθορίζονται από τις ανεξάρτητες μεταβλητές. Σε ένα τέτοιο μοντέλο, οι προβλέψεις παράγονται με βάση προκαθορισμένες σχέσεις και κανόνες που

διέπουν το σύστημα, θεωρώντας ότι αν γνωρίζουμε τις αρχικές συνθήκες και τις δυναμικές του συστήματος, μπορούμε να προβλέψουμε με ακρίβεια το μέλλον. Η μαθηματική σχέση που περιγράφει τα παραπάνω, με Y_t την εξαρτημένη μεταβλητή και με $X_{1(t-1)}, X_{2(t-1)}, X_{3(t-1)}, X_{4(t-1)}$ τις ανεξάρτητες μεταβλητές την αμέσως προηγούμενη χρονική στιγμή είναι:

$$Y_t = f(X_{1(t-1)}, X_{2(t-1)}, X_{3(t-1)}, X_{4(t-1)})$$

Σύμφωνα με την παραπάνω σχέση η πρόβλεψη μία χρονική στιγμή εξαρτάται από ένα σύνολο μεταβλητών την αμέσως προηγούμενη χρονική στιγμή.

Η βασική διαφορά μεταξύ των δύο μοντέλων πρόβλεψης έγκειται στη φιλοσοφία και τη μέθοδο πρόβλεψης. Τα αιτιοκρατικά μοντέλα βασίζονται στη σχέση αιτίου-αιτιατού, χρησιμοποιώντας εξισώσεις και ανεξάρτητες μεταβλητές για να εξηγήσουν τις αλλαγές στην εξαρτημένη μεταβλητή, δίνοντας έμφαση στη βαθιά κατανόηση του συστήματος. Για αυτό το λόγο είναι αυστηρά ντετερμινιστικά και εφαρμόζονται σε σύνθετα συστήματα, όπως η μετεωρολογία ή η οικονομία. από την άλλη, τα μοντέλα χρονοσειρών επικεντρώνονται αποκλειστικά στα ιστορικά δεδομένα της ίδιας της μεταβλητής, χωρίς να εξετάζουν τις αιτίες πίσω από τις τάσεις. Μπορεί να είναι στοχαστικά, ενσωματώνουν την αβεβαιότητα και χρησιμοποιούνται ευρέως σε προβλέψεις όπως πωλήσεις ή τιμές μετοχών. Με λίγα λόγια ενώ τα αιτιοκρατικά μοντέλα αναζητούν την αιτία, τα μοντέλα χρονοσειρών ανιχνεύουν μοτίβα και τάσεις.

Η παλινδρόμηση είναι μια στατιστική μέθοδος που χρησιμοποιείται για τη μοντελοποίηση της σχέσης μεταξύ μιας εξαρτημένης μεταβλητής (αποτελέσματος) και μιας ή περισσότερων ανεξάρτητων μεταβλητών (προγνωστικών). Η βασική της λογική είναι η αναζήτηση μιας μαθηματικής συνάρτησης που περιγράφει πώς οι ανεξάρτητες μεταβλητές επηρεάζουν την εξαρτημένη. Η διαδικασία αυτή ελαχιστοποιεί τα σφάλματα πρόβλεψης, δηλαδή τη διαφορά μεταξύ των πραγματικών και των προβλεπόμενων τιμών, συχνά μέσω της μεθόδου των ελαχίστων τετραγώνων. Οι μέθοδοι παλινδρόμησης είναι εργαλεία που εξηγούν σχέσεις μεταξύ μεταβλητών και χρησιμοποιούνται σε προβλέψεις και βελτιστοποίηση πολύπλοκων συστημάτων. Ταυτόχρονα χρησιμοποιούνται και σαν βοηθητικό εργαλείο σε άλλες μεθόδους πρόβλεψης, όπως για παράδειγμα στη μέθοδο διπλής εκθετικής εξομάλυνσης – μέθοδος Holt.

Παρακάτω αναλύονται οι δύο κυριότερες μέθοδοι παλινδρόμησης. Αυτές είναι η γραμμική παλινδρόμηση και η πολλαπλή παλινδρόμηση.

Γραμμική Παλινδρόμηση

Η γραμμική παλινδρόμηση είναι μια στατιστική μέθοδος που χρησιμοποιείται για να μοντελοποιήσει και να αναλύσει τη σχέση μεταξύ μιας εξαρτημένης μεταβλητής (αποτελέσματος) και μιας ανεξάρτητης μεταβλητής (προγνωστικής). Ο κύριος στόχος είναι η εύρεση μιας γραμμικής συνάρτησης που περιγράφει αυτή τη σχέση, ώστε να γίνουν προβλέψεις ή να εξαχθούν συμπεράσματα.

Η βασική εξίσωση της γραμμικής παλινδρόμησης είναι:

$$y = \beta_0 + \beta_1 x_1 + \varepsilon_t$$

Όπου y είναι η εξαρτημένη μεταβλητή, x η ανεξάρτητη μεταβλητή, β_0 η τιμή του y όταν $x=0$, β_1 η κλίση της ευθείας που δείχνει το μέγεθος της επίδρασης του x στο y και ε_t το σφάλμα που αντιπροσωπεύει τη διαφορά μεταξύ της πραγματικής τιμής του y και της προβλεπόμενης τιμής από το μοντέλο. Ο πιο κοινός τρόπος σου υπολογισμού των συντελεστών β_0 και β_1 είναι η μέθοδος των ελαχίστων τετράγωνων. Στόχος της είναι να ελαχιστοποιήσει το άθροισμα των τετραγώνων των αποκλίσεων μεταξύ των πραγματικών τιμών y και των προβλεπόμενων τιμών $\hat{y}$, πρακτικά να ελαχιστοποιήσει το σφάλμα μεταξύ πραγματικής τιμής και τιμής πρόβλεψης.

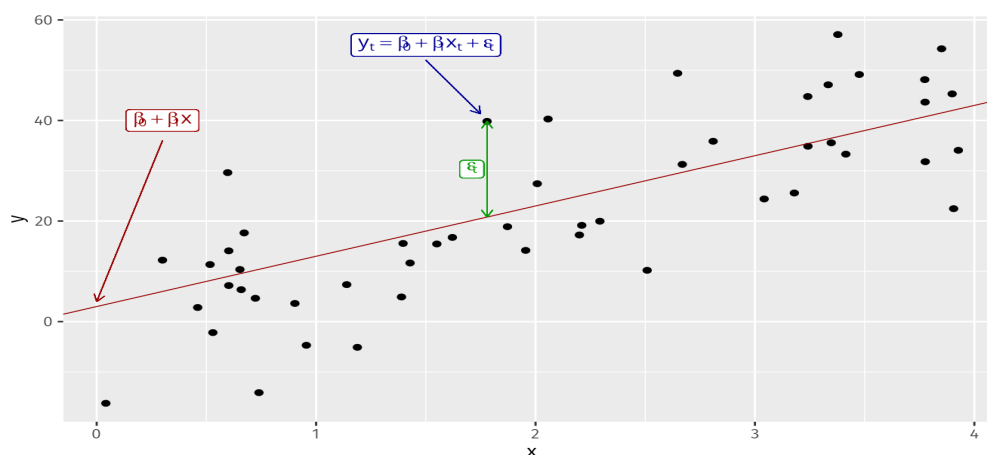
$$S = \sum (Y_i - \hat{Y}_i)^2$$

Με αυτόν τον τρόπο, η ευθεία πρόβλεψης που προκύπτει είναι αυτή που ταιριάζει καλύτερα στα δεδομένα, ελαχιστοποιώντας τις συνολικές αποκλίσεις.

Τελικά η πρόβλεψη μαθηματικά υπολογίζεται με τον τύπο : $\hat{y} = \beta_0 + \beta_1$

Η γραμμική παλινδρόμηση βασίζεται σε ορισμένες υποθέσεις για την αποτελεσματική λειτουργία της. Αυτές περιλαμβάνουν τη γραμμικότητα της σχέσης μεταξύ των μεταβλητών, την κανονική κατανομή των σφαλμάτων και τη σταθερή διακύμανση των σφαλμάτων. Εάν αυτές οι υποθέσεις ικανοποιούνται, η μέθοδος μπορεί να παρέχει αξιόπιστα αποτελέσματα και ακριβείς προβλέψεις.

Το μοντέλο της απλής γραμμικής παλινδρόμησης μπορεί να επεκταθεί και για σχέσεις όπου η ανεξάρτητη μεταβλητές είναι πολλές. Τότε γίνεται λόγος για το μοντέλο πολλαπλής γραμμικής παλινδρόμησης όπου μοιράζεται ακριβώς την ίδια φιλοσοφία με αυτό της απλής, απλά στους τύπος υπολογισμού περιλαμβάνει περισσότερες μεταβλητές. Ενδεικτικά η βασική εξίσωση περιγράφεται από τον τύπο:



Σχήμα 2.2 : Παράδειγμα γραμμικής παλινδρόμησης

2.6 Μέθοδοι αξιολόγησης προβλέψεων

Η στατιστική ακρίβεια προβλέψεων αναφέρεται στη διαδικασία αξιολόγησης της ποιότητας και της αξιοπιστίας των προβλέψεων που παράγονται από ένα μαθηματικό ή στατιστικό μοντέλο ενώ επίσης δεν αφορά μόνο την αξιολόγηση του τελικού αποτελέσματος, αλλά και την αποσαφήνιση του τρόπου με τον οποίο τα δεδομένα και οι υποθέσεις επηρεάζουν τη λειτουργία του μοντέλου. Είναι ουσιαστικά η ποσοτικοποίηση της ικανότητας του μοντέλου να προβλέπει με ακρίβεια τιμές που προσεγγίζουν τις πραγματικές παρατηρήσεις. Αυτή η αξιολόγηση πραγματοποιείται συγκρίνοντας τις προβλεπόμενες τιμές με τις πραγματικές τιμές ενός συνόλου δεδομένων, τα οποία έχουν ήδη παρατηρηθεί, αλλά δεν χρησιμοποιήθηκαν για την αρχική εκπαίδευση του μοντέλου. Η μέτρηση της ακρίβειας πρόβλεψης αποτελεί αναπόσπαστο κομμάτι κάθε διαδικασίας δημιουργίας προβλέψεων καθώς είναι το θεμέλιο για την μέτρηση της αξιοπιστίας και της δυνατότητας εφαρμογής του μοντέλου σε πραγματικές συνθήκες.

Το μέγεθος το οποίο περιγράφει μαθηματικά τον βαθμό προσεγγίσεις των προβλέψεων από τις πραγματικές παρατηρήσεις ονομάζεται σφάλμα και ορίζεται ως η διαφορά μεταξύ πραγματικής τιμής και πρόβλεψης.

$$e_i = Y_i - \hat{Y}_i, \text{ με } Y_i : \text{Πραγματική τιμή και } \hat{Y}_i : \text{Τιμή πρόβλεψης}$$

Στη συνέχεια ορίζονται οι κυριότεροι στατιστικοί δείκτες ακριβείας προβλέψεων.

Μέσο Σφάλμα (Mean Error - ME)

Το Μέσο Σφάλμα αντιπροσωπεύει τη μέση τιμή της διαφοράς (σφάλματος) μεταξύ των προβλεπόμενων τιμών ($\hat{Y}_i$) και των πραγματικών τιμών (Y_i) σε ένα σύνολο δεδομένων. Τιμές του μέσου σφάλματος κοντά στο μηδέν περιγράφουν τυχειότητα και όχι συστηματικότητα στα σφάλματα της προβλέψεως. Αντίστοιχα θετικές τιμές του μέσου σφάλματος δείχνουν πως οι προβλέψεις είναι κατά μέσο όρο μικρότερες των πραγματικών τιμών ενώ, αρνητικές τιμές το αντίθετο.

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

Μέσο Απόλυτο Σφάλμα (Mean Absolute Error – MAE)

Το μέσο απόλυτο σφάλμα υπολογίζει τη μέση τιμή της απόλυτης διαφοράς (απόκλισης) μεταξύ των πραγματικών τιμών και των προβλεπόμενων τιμών σε ένα σύνολο δεδομένων. Σχετίζεται άμεσα με την ακρίβεια τις προβλέψεις και περιγράφει το μέσο μετρό της αστοχίας χωρίς να δίνεται έμφαση στην κατεύθυνση.

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

Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error – MSE)

Το μέσο τετραγωνικό σφάλμα είναι κι αυτό ένα μέτρο της ακρίβειας της πρόβλεψης. Η διαφορά του με το μέσο απόλυτο σφάλμα είναι ότι εδώ λόγω του τετραγώνου δίνεται πολύ μεγαλύτερο βάρος στα μεγάλα σφάλματα. Καθώς το MSE εκφράζεται σε μονάδες που δεν είναι άμεσα κατανοητές, λόγω του τετραγώνου (π.χ., €² ή °C²), μπορεί να γίνει δύσκολη η κατανόηση του πόσο ακριβές είναι ένα μοντέλο σε πραγματικούς όρους. Χρησιμοποιείται όταν μεγαλύτερα σφάλματα είναι κρίσιμα για την αξιολόγηση του μοντέλου. Τέτοιες εφαρμογές είναι η πρόβλεψη καιρού, οι χρηματοοικονομικές προβλέψεις και η αξιολόγηση μοντέλων μηχανικής μάθησης.

Το μέσο τετραγωνικό σφάλμα χρησιμοποιείται συχνά σε συνδυασμό με τη Ρίζα Μέσου Τετραγωνικού Σφάλματος, το οποίο εκφράζεται στις ίδιες μονάδες με τα δεδομένα και παρέχει πιο διαισθητική κατανόηση της ακρίβειας του μοντέλου.

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

Ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Squared Error - RMSE)

Η ρίζα του μέσου τετραγωνικού σφάλματος υπολογίζεται λαμβάνοντας την ρίζα του MSE. Έτσι, το RMSE εκφράζεται στις ίδιες μονάδες με την χρονοσειρά, καθιστώντας το πιο διαισθητικό και κατανοητό σε σχέση με το μέσο τετραγωνικό σφάλμα.

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

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

Εκφράζει το μέσο απόλυτο σφάλμα ως ποσοστό των πραγματικών τιμών, καθιστώντας το ένα κανονικοποιημένο μέτρο που δεν εξαρτάται από τη μονάδα μέτρησης. Εκφράζει δηλαδή την απόκλιση μεταξύ των προβλεπόμενων και των πραγματικών τιμών ως ποσοστό των πραγματικών τιμών, γεγονός που το καθιστά εύκολα κατανοητό και συγκρίσιμο. Για παράδειγμα αν το MAPE είναι 10%, σημαίνει ότι, κατά μέσο όρο, οι προβλέψεις αποκλίνουν κατά 10% από τις πραγματικές τιμές. Παρόλο που το MAPE είναι ανεξάρτητο της μονάδας μέτρησης και έχει πολύ εύκολη ερμηνεία δεν μπορεί να εφαρμοστεί σε χρονοσειρές με πολλές μηδενικές τιμές ή τιμές που βρίσκονται κοντά στο μηδέν καθώς καταλήγει σε απροσδιοριστία. Ένα ακόμη μειονέκτημα του MAPE είναι πως τείνει να δίνει μεγαλύτερη έμφαση στις αποκλίσεις μικρών τιμών και μπορεί να υποεκτιμά τις μεγαλύτερες.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \cdot 100(\%)$$

Συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (Symmetric Mean Absolute Percentage Error – sMAPE)

Το sMAPE είναι μια τροποποιημένη εκδοχή του MAPE που αντιμετωπίζει την ασυμμετρία του, προσφέροντας έναν ισορροπημένο τρόπο μέτρησης των αποκλίσεων μεταξύ προβλεπόμενων και πραγματικών τιμών. Υπολογίζει την απόκλιση ως ποσοστό του μέσου όρου των πραγματικών και προβλεπόμενων τιμών, αποφεύγοντας την υπερβολική έμφαση σε μικρές τιμές.

$$SMAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{\frac{Y_i + \hat{Y}_i}{2}} \right| \cdot 100 (\%)$$

ΚΕΦΑΛΑΙΟ 3 – ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

3.1 Εισαγωγή

Η Τεχνητή Νοημοσύνη αποτελεί έναν πολυδιάστατο κλάδο της επιστήμης της πληροφορικής με στόχευση την δημιουργία συστημάτων ικανών να προσομοιάζουν ανθρώπινες νοητικές λειτουργίες. Οι κυριότερες από αυτές τις λειτουργίες είναι η αντίληψη, όπου για παράδειγμα ένα ρομποτικό σύστημα αντιλαμβάνεται τον φυσικό κόσμο μέσω ερεθισμάτων από αισθητήρες και αλληλοεπιδρά με αυτόν με αυτοματοποιημένους αλγόριθμους κίνησης. Ακόμη είναι η λήψη αποφάσεων, με αλγόριθμους αναζήτησης και βελτιστοποίησης όπως για παράδειγμα αυτοματοποιημένοι αλγόριθμοι για σκάκι. Ενώ τέλος η πιο διαδεδομένη από αυτές είναι η μάθηση, με κλασικό παράδειγμα τα νευρωνικά δίκτυα. Η μηχανική μάθηση αποτελεί υποκατηγορία του κλάδου της τεχνητής νοημοσύνης. Πρόκειται λοιπόν για συστήματα που εκτελούν λειτουργίες οι οποίες τυπικά απαιτούν ανθρώπινη νοημοσύνη.

Η ρίζες της τεχνητής νοημοσύνης ως ιδέα βρίσκονται σε αρκετά παλαιότερες εποχές. Η γενικευμένη ιδέα θεμελιώθηκε το 1950, από τον Alan Turing ο οποίος όχι μόνο εξέφρασε την άποψη ότι μία μηχανή μπορεί να σκέφτεται αλλά πρότεινε μια διαδικασία δοξολογία με μηχανή μπορεί να επιδείξει ανθρώπινη νοημοσύνη, το τεστ Turing [33]. Ωστόσο ως έτος της γέννησης της τεχνητής νοημοσύνης θεωρείται το 1956, όπου ο John McCarthy χρησιμοποίησε για πρώτη φορά τον όρο "Artificial Intelligence" [34]. Αυτή η εκδήλωση θεωρείται το επίσημο σημείο εκκίνησης της τεχνητής νοημοσύνης ως επιστημονικό πεδίο. Συνεχίζοντας κατά τη δεκαετία του 1960 πραγματοποιήθηκαν τα πρώτα βήματα, με την κατασκευή αλγορίθμων λογικής και αναζήτησης για την επίλυση προβλημάτων καθώς και πρώιμα συστήματα όπως το ELIZA, ένα σύστημα το οποίο προσομοίωνε συνομιλίες [36]. Ωστόσο, οι συνθήκες δεν ήταν ακόμη οι κατάλληλες για την διάνθηση του κλάδου και έτσι η δεκαετία του 1970 χαρακτηρίστηκε ως ο χειμώνας της τεχνητής νοημοσύνης. Η περιορισμένη υπολογιστική πόροι και η έλλειψη πρακτικής εφαρμογής της τεχνητής νοημοσύνης κατακρήμνισαν τις υπερβολικές προσδοκίες του παρελθόντος με αποτέλεσμα οι χρηματοδοτήσεις για έρευνα και ανάπτυξη να σταματήσουν. Τα επόμενα χρόνια με την πρόοδο της τεχνολογίας και την ανάπτυξη ισχυρότερων υπολογιστών ο κλάδος επανήλθε στο προσκήνιο και έτσι την δεκαετία του 1980 με την αναβίωση των νευρωνικών δικτύων με τη μέθοδο backpropagation τέθηκαν τα θεμέλια της τεχνητής νοημοσύνης όπως την ξέρουμε σήμερα [50]. Στα επόμενα χρόνια ακολούθησαν Επιτεύγματα που προκάλεσαν επανάσταση στον κόσμο της πληροφορικής. Σήμερα, είναι ενσωματωμένη σε πολλές πτυχές της καθημερινής μας ζωής καθώς και σε πάρα πολλά επίπεδα της παραγωγής, από τις εφαρμογές στις μεταφορές (αυτόνομα οχήματα) και την ιατρική διάγνωση, έως την ψυχαγωγία και τη διαχείριση δεδομένων, συνεχίζει να εξελίσσεται με ταχύ ρυθμό. Παράλληλα, παραμένουν έντονα ζητήματα που σχετίζονται με την ηθική, την ασφάλεια, την ιδιωτικότητα και την ενδεχόμενη προκατάληψη στους αλγορίθμους. Πλέον η τεχνητή νοημοσύνη έχει συνδεθεί άρρηκτα με την ζωή μας ενώ ταυτόχρονα αλλάζει στην καθημερινότητα μας σε όλα τα επίπεδα.

Η μηχανική μάθηση θεωρείται ο βασικότερος πυλώνας της τεχνητής νοημοσύνης και ορίζεται σαν τη διαδικασία κατά την οποία μοντέλα εκπαιδεύονται με δεδομένα για να παράξουν ένα αποτέλεσμα χωρίς να προγραμματίζονται ρητά [35].

3.2 Κατηγορίες μηχανικής μάθησης

Υπάρχουν τρεις γενικές κατηγορίες μηχανικής μάθησης. Καθεμία από τις οποίες αφορά σε προβλήματα διαφορετικής φύσης.

Επιβλεπόμενη μάθηση (Supervised Learning)

Πρόκειται για μία μέθοδο κατά την οποία το μοντέλο εκπαιδεύεται – μαθαίνει χρησιμοποιώντας ένα σύνολο δεδομένων, τα οποία ορίζονται πάντα σε ζεύγη εισόδου (features) και σωστής εξόδου (labels). Στόχος του μοντέλου είναι να μάθει τη σχέση μεταξύ εισόδων και εξόδων, ώστε να μπορεί να προβλέπει τις σωστές εξόδους διαβάζοντας νέες, άγνωστες εισόδους [38]. Ονομάζεται λοιπόν επιβλεπόμενη γιατί κατά τη διαδικασία της μάθησης το μοντέλο καθοδηγείται με την παροχή της σωστής εξόδου στην οποία πρέπει να προσαρμοστεί για δεδομένη είσοδο. Τα μοντέλα επιβλεπόμενης μάθησης επιτυγχάνουν πολύ μεγάλη ακρίβεια ενώ επίσης χρησιμοποιούνται σε ένα ευρύ φάσμα εφαρμογών. Εξαιτίας της φύσης της εκπαίδευσής τους η απόδοση εξαρτάται σε σημαντικό βαθμό από την ποσότητα και την ποιότητα των διαθέσιμων δεδομένων, ενώ ταυτόχρονα το υπολογιστικό κόστος δεν εξαρτάται μόνο από την διαδικασία της μάθησης αλλά και από το κόστος των δεδομένων.

Τις βασικότερες εφαρμογές επιβλεπόμενης μάθησης αποτελούν τα προβλήματα της ταξινόμησης (classification) και της πρόβλεψης-παλινδρόμησης (regression). Στα προβλήματα ταξινόμησης η έξοδος είναι ένα διακριτό σύνολο δύο ή περισσότερων τιμών και στόχος είναι το μοντέλο για κάθε είσοδο να επιλέγει την κατάλληλη έξοδο [39]. Στα προβλήματα ταξινόμησης, η αξιολόγηση της απόδοσης των μοντέλων γίνεται μέσω διάφορων μετρικών. Μία από τις βασικότερες είναι η Ακρίβεια (Accuracy), η οποία εκφράζει το ποσοστό των σωστών προβλέψεων σε σχέση με το σύνολο των παραδειγμάτων ενώ άλλες μετρικές όπως η Precision, η Recall και το F1-Score είναι ιδιαίτερα χρήσιμες σε περιπτώσεις όπου υπάρχει ανισορροπία δεδομένων ή όταν δίνεται έμφαση στην ανίχνευση θετικών κατηγοριών. Χαρακτηριστικό παράδειγμα είναι η αναγνώριση email ως "ανεπιθύμητο" ή "μη ανεπιθύμητο". Το μοντέλο εκπαιδεύεται με ένα σύνολο δεδομένων όπου κάθε email έχει ήδη επισημανθεί με την αντίστοιχη κατηγορία και στη συνέχεια καλείται να ταξινομήσει νέα, άγνωστα email. από την άλλη μεριά, τα προβλήματα παλινδρόμησης έχουν ως βασική στόχευση την πρόβλεψη μίας συνεχούς αριθμητικής τιμής, μέσω της συσχέτισης της μεταβλητής εξόδου με ένα σύνολο μεταβλητών εισόδου. Μετρικές για αυτού του είδους τα προβλήματα αποτελούν όλοι οι δείκτες σφαλμάτων που αναφέρθηκαν στην ενότητα των

προβλέψεων. Παράδειγμα αποτελεί η πρόβλεψη ροής φόρτου επόμενης ημέρας, το οποίο αποτελεί και το βασικό στοιχείο αυτής της εργασίας [10].

Μη επιβλεπόμενη μάθηση (Unsupervised Learning)

Η Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning) αποτελεί μέθοδο της μηχανικής μάθησης όπου τα μοντέλα εκπαιδεύονται σε δεδομένα που δεν σχετίζονται με προκαθορισμένες εξόδους - ετικέτες. Σε αντίθεση λοιπόν με την επιβλεπόμενη μάθηση, όπου υπάρχει "επιβλέπουσα καθοδήγηση" μέσω ετικετοποιημένων δεδομένων, σε αυτή τη μέθοδο το μοντέλο πρέπει αυτόνομα να ανακαλύψει μοτίβα, σχέσεις ή δομές μέσα στα δεδομένα, χωρίς καθοδήγηση. Το μοντέλο λειτουργεί σαν "εξερευνητής" που προσπαθεί να αποκαλύψει τη φυσική οργάνωση των δεδομένων χωρίς να γνωρίζει από πριν τι ακριβώς αναζητά.

Οι δύο κύριες κατηγορίες προβλημάτων που εμπίπτουν στη μη επιβλεπόμενη μάθηση είναι η ομαδοποίηση (clustering) και το πρόβλημα της μείωσης των διαστάσεων (dimensionality reduction). Αναλυτικά, αναφορικά με το πρόβλημα της ομαδοποίησης, πρόκειται για τη διαδικασία συσχετισμού δεδομένων βάσει κοινών χαρακτηριστικών τους, χωρίς να υπάρχουν προκαθορισμένες κατηγορίες. Χαρακτηριστικό παράδειγμα αποτελεί ο αλγόριθμος k-means, όπου είναι ένας δημοφιλής αλγόριθμος μη επιβλεπόμενης μάθησης που χρησιμοποιείται για την ομαδοποίηση δεδομένων, ελαχιστοποιώντας την απόσταση των σημείων από το κέντρο της ομάδας τους [46]. Μετρικές αξιολόγησης περιλαμβάνουν τον Silhouette Score που αξιολογεί τόσο τη συνοχή εντός των ομάδων όσο και τη διάκριση μεταξύ διαφορετικών ομάδων αλλά και τη μετρική Inertia που μετρά τη συνοχή εντός αυτών των ομάδων, υπολογίζοντας το άθροισμα των τετραγώνων των αποστάσεων των σημείων από τα κέντρα των ομάδων. Συνεχίζοντας δεύτερο πρόβλημα αποτελεί η μείωση διαστάσεων (dimensionality reduction), η οποία αφορά την μείωση των διαστάσεων των δεδομένων, διατηρώντας την μέγιστη δυνατή πληροφορία και παράλληλα μειώνοντας το θόρυβο. Παράδειγμα αποτελεί η ομαδοποίηση πελατών μιας επιχείρησης με βάση τις αγοραστικές τους συνήθειες. Το μοντέλο αναλύει δεδομένα αγορών χωρίς προκαθορισμένες ετικέτες και εντοπίζει μοτίβα που σχηματίζουν διαφορετικές ομάδες πελατών.

Ενισχυτική μάθηση (Reinforced learning)

Ενισχυτική μάθηση είναι η διαδικασία κατά την οποία το σύστημα αλληλοεπιδρά με ένα άγνωστο περιβάλλον εκτελώντας ενέργειες ενώ ταυτόχρονα λαμβάνει ανατροφοδότηση μέσω επιβράβευσης ή τιμωρίας και προσαρμόζει τη στρατηγική του με στόχο τη μεγιστοποίηση του ερεθίσματος της επιβράβευσης. Έτσι, το μοντέλο μπορεί να επιλύει όλο και πιο περίπλοκα προβλήματα. Η απόδοση της ενισχυτικής μάθησης είναι καλή σε προβλήματα που περιλαμβάνουν διαδοχικές αποφάσεις όπως και γενικότερα όταν οι κανόνες ή οι στόχοι είναι δυναμικοί και πολύπλοκοι από την άλλη όμως για την εκπαίδευση ενός μοντέλου

απαιτείται μεγάλος αριθμός δοκιμών και σφαλμάτων με αποτέλεσμα την χρήση μεγάλου όγκου υπολογιστικών πόρων. Χαρακτηριστικό παράδειγμα εφαρμογής αποτελεί η εκπαίδευση ενός συστήματος για αυτόνομη οδήγηση [40].

3.3 Μοντέλα μηχανικής μάθησης

Παρακάτω αναλύονται τα κυριότερα μοντέλα της μηχανικής μάθησης όπου το καθένα εκπαιδεύεται με μια από τις προαναφερθείσες μεθόδους εκπαίδευσης.

Δέντρα Απόφασης (Decision Trees)

Πρόκειται για έναν αλγόριθμο που εκπαιδεύεται με την μέθοδο της επιβλεπόμενης μάθησης και εφαρμόζεται σε προβλήματα ταξινόμησης και παλινδρόμησης. Ο αλγόριθμος περιγράφεται από δεντρικές δομές, όπου κάθε ενδιάμεσος κόμβος αντιπροσωπεύει ένα από τα χαρακτηριστικά του δείγματος και κάθε ένα από τα κλαδιά αυτού του κόμβου αναπαριστά τις πιθανές τιμές του χαρακτηριστικού. Στον πάτο βρίσκονται τα φύλλα τα οποία περιγράφουν την τελική πρόβλεψη ή απόφαση, ανάλογα με την αλυσίδα των χαρακτηριστικών που ακολουθήθηκε από την ρίζα ως κάτω. Στη πράξη εφαρμόζονται βελτιωμένες εκδοχές του αλγορίθμου οι οποίες εμφανίζουν σαφώς καλύτερη απόδοση όμως απαιτούν περισσότερο χρόνο για την εκπαίδευση. Αυτές είναι ο Random Forest που βασίζεται σε πολλαπλά δέντρα απόφασης, όπου καθένα εκπαιδεύεται σε ένα τυχαία κατακερματισμένο τμήμα του συνόλου των δεδομένων και στο τέλος συνδυάζονται για την τελική απόφαση και το Gradient Boosting Tree όπου εδώ τα δέντρα αποφάσεις εκπαιδεύεται διαδοχικά μαθαίνοντας το καθένα από τα λάθη του προηγούμενου [43].

Support Vector Machine (SVM)

Πρόκειται για μια τεχνική επιβλεπόμενης μάθησης που εφαρμόζεται κυρίως σε προβλήματα ταξινόμησης [44]. Στόχος του αλγορίθμου είναι να βρει την υπερεπιφάνεια που διαχωρίζει με τον βέλτιστο τρόπο δύο ή περισσότερες κατηγορίες δεδομένων. Για τον βέλτιστο βαθμό ταξινόμησης αυτή η υπερεπιφάνεια πρέπει να έχει ταυτόχρονα τη μέγιστη δυνατή απόσταση από τα πλησιέστερα σημεία όλων των κατηγοριών. Τα πλησιέστερα σημεία χαρακτηρίζονται ως υποστηρικτικά διανύσματα (support vectors). Σε περιπτώσεις όπου τα δεδομένα δεν είναι διαχωρίσιμα με γραμμικό τρόπο χρησιμοποιείται η τεχνική kernel trick, κατά την οποία τα δεδομένα μετασχηματίζονται σε υψηλότερες διαστάσεις στις οποίες είναι διαχωρίσιμα. Ο αλγόριθμος είναι αποδοτικός σε προβλήματα με δεδομένα υψηλών διαστάσεων, δηλαδή που περιγράφονται από μεγάλο αριθμό χαρακτηριστικών, όμως για πολυπληθή δεδομένα το υπολογιστικό κόστος είναι μεγάλο ενώ επίσης είναι ευαίσθητος ως προς την παραμετροποίηση.

Δίκτυα Bayes

Τα δίκτυα Bayes είναι πιθανοτικά μοντέλα τα οποία εκφράζουν σχέσεις αιτιότητας μεταξύ τυχαίων μεταβλητών, μέσω κατευθυνόμενων άκυκλων γράφων (DAG). Κάθε κόμβος του δικτύου αντιστοιχεί σε μια μεταβλητή, ενώ οι ακμές δείχνουν τις αιτιώδεις εξαρτήσεις μεταξύ τους, οι οποίες είναι υπό συνθήκη πιθανότητες [45].

Ειδική περίπτωση των δικτύων Bayes αποτελούν τα Γκαουσιανά Μαρκοβιανά Μοντέλα (GMMs), όπου οι τυχαίες μεταβλητές παίρνουν συνεχείς τιμές και ακολουθούν κανονική κατανομή. Σε αντίθεση με τα κλασσικά δίκτυα εδώ χρησιμοποιούνται μη κατευθυνόμενοι γράφοι ενώ οι ακμές περιγράφουν συσχετίσεις μεταξύ γειτονικών κόμβων, οι οποίες υπολογίζονται μέσω συνδιακύμανσης.

k-Means

Ο k-Means είναι ένας μη επιβλεπόμενος αλγόριθμος συσταδοποίησης που χρησιμοποιείται για την κατάταξη δεδομένων σε k αριθμό ομάδων, με την παράμετρο k να καθορίζεται από την αρχή. Ο αλγόριθμος ξεκινά επιλέγοντας k κεντροειδή σημεία και συνδέει κάθε σημείο των δεδομένων με το πλησιέστερο κεντροειδές, σχηματίζοντας έτσι προσωρινές ομάδες. Στη συνέχεια, υπολογίζονται νέα κεντροειδή ως ο μέσος όρος των σημείων της κάθε αρχικής ομάδας και ομαδοποιούνται ξανά όλα τα σημεία με βάση αυτά. Η διαδικασία αυτή επαναλαμβάνεται έως ότου τα κεντροειδή σταθεροποιηθούν. Ο k-Means είναι γρήγορος και εύκολος στην εφαρμογή, αλλά η απόδοση του εξαρτάται από την αρχική επιλογή των κεντροειδών. Για την σωστή επιλογή της παραμέτρου k χρησιμοποιούνται τεχνικές όπως η μέθοδος Elbow ή το Silhouette Score.

k-Nearest Neighbors (k-NN)

Ο αλγόριθμος βασίζεται στην αρχή ότι ένα νέο σημείο ταξινομείται με βάση την πλειοψηφία των k πλησιέστερων γειτόνων του. Υπολογίζει δηλαδή την απόσταση μεταξύ του νέου σημείου και των υπαρχόντων δεδομένων και βρίσκει τα k πιο κοντινά σημεία [47]. Στην ταξινόμηση, το νέο σημείο εντάσσεται στην πλειοψηφική κατηγορία, ενώ στην παλινδρόμηση, η τιμή προκύπτει από το μέσο όρο των τιμών των γειτόνων. Αν και κατηγοριοποιείται στην επιβλεπόμενη μάθηση, καθώς τα δεδομένα που του παρέχονται έχουν ετικέτες, χαρακτηρίζεται ως 'τεμπέλης' αφού δεν υφίσταται κάποια ουσιαστική εκπαίδευση αλλά αντίθετα απλώς αποθηκεύει τον σύνολο των δεδομένων. Επομένως η διαδικασία λήψης απόφασης συμβαίνει σε πραγματικό χρόνο χωρίς να έχει υπάρξει κάποια διαδικασία στο παρελθόν

Περί ενισχυτικής μάθησης

Οι αλγόριθμοι ενισχυτικής μάθησης χωρίζονται σε δύο κύριες κατηγορίες. Οι αλγόριθμοι που βασίζονται σε πίνακες και είναι κατάλληλοι για περιβάλλοντα με

περιορισμένο αριθμό καταστάσεων και ενεργειών, όπου οι πιθανότητες επιτυχίας κάθε ενέργειας μπορούν να καταγραφούν σε έναν πίνακα. Χαρακτηριστικό παράδειγμα αποτελεί ο Q-Learning, ο οποίος επιτρέπει σε έναν πράκτορα να μαθαίνει μέσω δοκιμής και σφάλματος ποια είναι η καλύτερη ενέργεια για κάθε κατάσταση, αποθηκεύοντας τις γνώσεις του σε έναν πίνακα τιμών [40]. Η δεύτερη μεγάλη κατηγορία περιλαμβάνει αλγόριθμους που χρησιμοποιούν προσεγγιστικά μοντέλα, όπως νευρωνικά δίκτυα, για να μάθουν πολύπλοκες στρατηγικές δράσης. Αυτοί οι αλγόριθμοι είναι κατάλληλοι για περιβάλλοντα όπου ο αριθμός των πιθανών καταστάσεων είναι πολύ μεγάλος, καθιστώντας ανέφικτη τη χρήση ενός απλού πίνακα. Ένα παράδειγμα είναι τα Deep Q-Networks (DQN), τα οποία χρησιμοποιούν νευρωνικά δίκτυα για να μάθουν την καλύτερη στρατηγική μέσα από εμπειρία, επιτρέποντας τη λήψη αποφάσεων ακόμα και σε δυναμικά περιβάλλοντα με πολλές μεταβλητές [41].

Νευρωνικά δίκτυα

Τέλος, ένα από τα πιο ισχυρά εργαλεία της μηχανικής μάθησης είναι τα νευρωνικά δίκτυα. Τα μοντέλα αυτά ξεχωρίζουν με διαφορά σε σχέση με τα υπόλοιπα σε εφαρμογές με πολύπλοκα δεδομένα, όπως εικόνες και φυσική γλώσσα, όμως απαιτούν μεγάλα σύνολα δεδομένων και υψηλή υπολογιστική ισχύ. Χαρακτηριστικό της ευρείας χρήσης τους αποτελεί το γεγονός ότι υπάρχει κλάδος της μηχανικής μάθησης κατά τον οποίο χρησιμοποιούνται πολύ επίπεδο νευρωνικά δίκτυα. Ο κλάδος αυτός ονομάζεται βαθιά μάθηση (Deep Learning). Στην επόμενη ενότητα γίνεται εκτενής ανάλυση

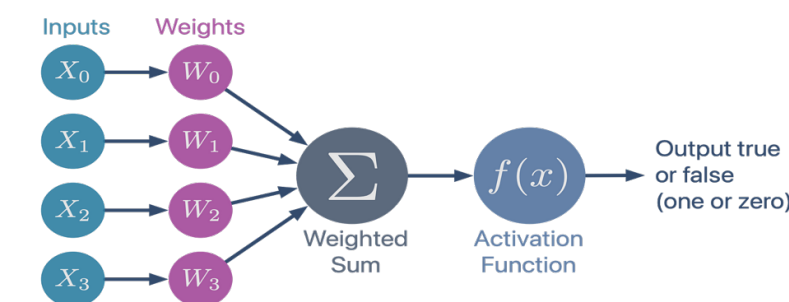
3.4 Νευρωνικά δίκτυα

Ως νευρωνικό δίκτυο ορίζεται ένα υπολογιστικό μοντέλο εμπνευσμένο από τον ανθρώπινο εγκέφαλο. Αποτελείται από τεχνητούς νευρώνες που μετασχηματίζουν και προσαρμόζουν πληροφορίες μέσω συνδεδεμένων βαρών. Τα νευρωνικά δίκτυα εκπαιδεύονται μαθαίνοντας μοτίβα από τα δεδομένα προσαρμόζοντάς τις παραμέτρους τους με στόχο να κάνουν ακριβείς προβλέψεις ή να παίρνουν αποφάσεις. Η ισχύς τους πηγάζει από την ικανότητα να μαθαίνουν πολύπλοκες μη γραμμικές σχέσεις από τα δεδομένα χωρίς να χρειάζονται ρητό προγραμματισμό[48].

3.4.1 Τεχνητός νευρώνας - Perceptron

Ο τεχνητός νευρώνας αποτελεί το βασικό δομικό στοιχείο των νευρωνικών δικτύων. Προτάθηκε το 1943 από τους McCulloch και Pitts ως δυαδικός ταξινομητής με μόνο δύο δυνατές καταστάσεις στην έξοδο on/off ή αλλιώς 1 και 0. Η βασική λειτουργία του είναι να επεξεργάζεται εισόδους, να τις ζυγίζει και να αποφασίζει αν πληρούν ένα συγκεκριμένο κριτήριο. Πρόκειται για μια διαδικασία

η οποία προσομοιάζει την ανθρώπινη ικανότητα αντίληψης και λήψης αποφάσεων. Λόγω αυτού ο τεχνητός νευρώνας ονομάζεται αλλιώς και αντίληπτρο (perceptron) [48].



Σχήμα 3.1 : Αναπαράσταση perceptron

Ο τεχνητός νευρώνας έχει εμπνευστεί από τον βιολογικό νευρώνα και αποτελείται από έναν αριθμό εισόδων, αντίστοιχο πλήθος βαρών όπου καθένα αντιστοιχεί σε μία είσοδο, ένα μέγεθος προκατάληψης, μια συνάρτηση ενεργοποίησης και μια έξοδο. Αρχικά, λαμβάνει από τις εισόδους το σύνολο των χαρακτηριστικών ή των παρατηρήσεων της πληροφορίας. Κάθε είσοδος πολλαπλασιάζεται με ένα αντίστοιχο βάρος το οποίο καθορίζει τη σημαντικότητα της κάθε εισόδου στην τελική απόφαση του νευρώνα. Συνεχίζοντας τα σταθμισμένα αυτά αποτελέσματα αθροίζονται και σε αυτά προστίθεται και η προκατάληψη, ένας παράγοντας που επιτρέπει στον νευρώνα να μετατοπίζει τη λειτουργία ενεργοποίησης, ανεξάρτητα από τις εισόδους. Τελικά στο αποτέλεσμα της συνάθροισης εφαρμόζεται μία συνάρτηση ενεργοποίησης για τον καθορισμό της τελικής εξόδου. Σημειώνεται πως στο μοντέλο των McCulloch και Pitts, ο νευρώνας που λειτουργούσε σαν δυαδικός ταξινομητής, είχε ορισμένη συνάρτηση ενεργοποίησης την βηματική συνάρτηση. Η μαθηματική διατύπωση της λειτουργίας του τεχνητού νευρώνα είναι :

$$y = f \left(\sum_{i=0}^n x_i w_i + b \right)$$

y : Η έξοδος του νευρώνα, f : Η συνάρτηση ενεργοποίησης

x_i : Μία συνιστώσα της εισόδου, με το σύνολο τους να είναι n

w_i : Μία συνιστώσα του διανύσματος των βαρών, b : Η προκατάληψη (bias)

Μεμονωμένα ένας τεχνητός νευρώνας έχει περιορισμένη λειτουργικότητα μόνο σε γραμμικά προβλήματα. Για παράδειγμα το μοντέλο που ορίζεται από έναν μόνο νευρώνα, με συνάρτηση ενεργοποίησης την βηματική επικρατεί στην βιβλιογραφία ως perceptron και όπως έχει ήδη αναφερθεί χρησιμοποιείται για δυαδική ταξινόμηση. Η εκπαίδευση πραγματοποιείται ως εξής : Αρχικά επιλέγονται τυχαία βάρη για δεδομένη είσοδο και γίνεται υπολογισμός της εξόδου. Όσο η πραγματική έξοδος είναι διαφορετική από την επιθυμητή έξοδο, τα

βάρη ενημερώνονται μέχρι αυτά να σταθεροποιηθούν, σύμφωνα με το κανόνα εκπαίδευσης :

$$w_i \leftarrow w_i + \eta(d - y)x_i$$

η : Ρυθμός μάθησης, d : Επιθυμητή έξοδος και y : Πραγματική έξοδος

Άλλο παράδειγμα αποτελεί το μοντέλο Adaline που ορίζεται πάλι από έναν νευρώνα με τη διαφορά ότι η συνάρτηση ενεργοποίησης είναι η ταυτοτική και το σφάλμα μεταξύ επιθυμητής και πραγματικής τιμής υπολογίζεται με τη μέθοδο των ελαχίστων τετραγώνων. Χρησιμοποιήθηκε κυρίως σε προσαρμοστικά φίλτρα και εφαρμογές ακύρωσης θορύβου όπου έχει σημασία το μέγεθος του σφάλματος και όχι απλά η ορθότητα του αποτελέσματος [49].

3.4.2 Συνάρτηση ενεργοποίησης

Η συνάρτηση ενεργοποίησης αποτελεί σημαντικό στοιχείο των τεχνητών νευρώνων, αφού από αυτές καθορίζεται η έξοδος, βάσει μιας σειράς εισόδων. Ταυτόχρονα, κάποιες συναρτήσεις ενεργοποίησης εισάγουν μη γραμμικότητα στους νευρώνες και κατ' επέκταση στο νευρωνικό δίκτυο, επιτρέποντας τους να μαθαίνουν και να αναπαριστούν πολύπλοκες, μη γραμμικές σχέσεις στα δεδομένα. Αυτό γίνεται μέσω της εφαρμογής μιας μη γραμμικής συνάρτησης στο σταθμισμένο άθροισμα των εισόδων κάθε νευρώνα. Παρακάτω αναλύονται οι βασικότερες συναρτήσεις μεταφοράς που συναντώνται στη βιβλιογραφία.

Γραμμική συνάρτηση

Η γραμμική συνάρτηση ενεργοποίησης είναι μια απλή συνάρτηση που επιστρέφει την είσοδο χωρίς καμία μετασχηματιστική επεξεργασία. Η κύρια ιδιότητα της συνάρτησης είναι η διατήρηση της γραμμικής φύσης του νευρώνα μεταξύ εισόδων και εξόδων. Χρησιμοποιείται κυρίως όταν δεν χρειάζεται να μεταβληθεί η έξοδος. Αυτό συμβαίνει κυρίως σε εφαρμογές προβλέψεις συνεχών τιμών καθώς επίσης και σε ειδικές περιπτώσεις σε κρυφά επίπεδα όπου δεν είναι απαραίτητες μη γραμμικές σχέσεις. Μαθηματικά εκφράζεται ως:

$$f(x) = x$$

Βηματική συνάρτηση (Step Activation Function)

Πρόκειται για συνάρτηση που επιστρέφει ένα αν η είσοδος υπερβαίνει ένα κατώφλι θ , διαφορετικά επιστρέφει μηδέν. Το κύριο χαρακτηριστικό της λοιπόν είναι πως είναι δυαδική, δηλαδή παράγει διακριτές εξόδους χωρίς βαθμιαία μετάβαση. Βρίσκει εφαρμογή σε απλούς ταξινομητές με στόχο την κατηγοριοποίηση δεδομένων σε δύο κλάσεις. Εξαιτίας της ασυνέχειας που έχει δεν είναι λειτουργική για το μεγαλύτερο εύρος των εφαρμογών, καθώς η μη

διαφορίσιμη φύση της εμποδίζει την εφαρμογή αλγορίθμων εκπαίδευσης, όπως για παράδειγμα τον backpropagation και γενικότερα δεν μπορεί να διαχειριστεί μη γραμμικά διαχωρίσιμα προβλήματα. Για τους παραπάνω λόγους, έχει αντικατασταθεί από πιο εξελιγμένες συναρτήσεις ενεργοποίησης, όπως η sigmoid και η ReLU. Παρόλα αυτά, παραμένει σταθμός στην ιστορική ανάπτυξη των νευρωνικών δικτύων και χρησιμοποιείται για διδακτικούς σκοπούς και σε πολύ απλά νευρωνικά μοντέλα.

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

Ειδική περίπτωση αποτελεί η βηματική διπολική συνάρτηση, με μόνη διαφορά πως η συνάρτηση επιστρέφει -1 , αντί για την μηδενική τιμή.

$$f(x) = \begin{cases} -1, & x \leq \theta \\ +1, & x \geq \theta \end{cases}$$

Ανορθωμένη Γραμμική Συνάρτηση (ReLU - Rectified Linear Unit)

Σύμφωνα και με την ονομασία της πρόκειται για μία συνάρτηση όπου για θετικές τιμές εισόδου συμπεριφέρεται όπως η γραμμική συνάρτηση ενώ, για αρνητικές τιμές εισόδου, η έξοδος μηδενίζεται. Χρησιμοποιείται κυρίως σε βαθιά νευρωνικά δίκτυα ενώ αποτελεί την προεπιλεγμένη συνάρτηση ενεργοποίησης σε συνεκτικά δίκτυα με εφαρμογές κυρίως στον τομέα της ανάλυσης εικόνας. Η γραμμικότητα που παρουσιάζει για θετικές τιμές εισόδου επιτρέπει την πολύ γρήγορη εκπαίδευση σε μεγάλα δίκτυα με χιλιάδες νευρώνες εξαιτίας της υπολογιστικής της απλότητας. Ταυτόχρονα η σταθερή της κλίση σε συνδυασμό με την έλλειψη κορεσμού για μεγάλες τιμές εισόδου εξασφαλίζουν την διατήρηση του gradient, αντιμετωπίζοντας το πρόβλημα της εξαφάνισης του. από την άλλη η ReLU για αρνητικές τιμές εισόδου συμπεριφέρεται μη γραμμικά αφού μηδενίζει την έξοδο. Έτσι αν μία είσοδος είναι αρνητική, αυτομάτως θεωρείται λιγότερο σημαντική και εξαλείφεται. Αυτό βοηθά το νευρωνικό δίκτυο επικεντρωθεί στις πιο κρίσιμες πληροφορίες αποφεύγοντας θόρυβο, καθώς πολλοί νευρώνες αποκλείονται δυναμικά μειώνοντας έτσι το κόστος εκπαίδευσης. Εν γένει τα βαθιά δίκτυα συχνά ευνοούνται από αραιές αναπαραστάσεις, όπου μόνο λίγοι νευρώνες ενεργοποιούνται κάθε φορά, επιτρέποντας καλύτερη γενίκευση στα δεδομένα. Το πιο σημαντικό όμως χαρακτηριστικό της είναι πως η γραμμική της συμπεριφορά συνδυαστικά με την ασυνέχεια που εμφανίζει για μη θετικές τιμές εισόδου επιτρέπουν στο νευρωνικό δίκτυο να προσεγγίζει μια μη γραμμική συνάρτηση μέσω πολλών τοπικών γραμμικών τμημάτων. Έτσι καταλήγει να εμφανίζει μη γραμμική συμπεριφορά με αποτέλεσμα να μπορεί να διαχειρίζεται πολύπλοκες μη γραμμικές σχέσεις, επιτρέποντάς του να μάθει και να αναπαραστήσει σύνθετα πρότυπα στα δεδομένα [52].

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

Σιγμοειδής συνάρτηση (Sigmoid)

Αυτή η συνάρτηση μετατρέπει κάθε πραγματικό αριθμό της εισόδου σε μια τιμή ανάμεσα στο μηδέν και το ένα. Για μεγάλες θετικές τιμές η συνάρτηση τείνει στο ένα ενώ, αντίστοιχα για εξαιρετικά αρνητικές τείνει στο μηδέν. Το κύριο χαρακτηριστικό της είναι λοιπόν πως η συμπεριφορά της προσομοιάζει το εύρος μίας πιθανότητας. Επιπρόσθετα, πρόκειται για μια συνεχή συνάρτηση, η οποία, όχι μόνο κατά κάποιο τρόπο προσεγγίζει τη συμπεριφορά της βηματικής συνάρτησης αλλά επίσης επιτρέπει την παραγωγή ως προς τα συναπτικά βάρη με στόχο την εκπαίδευση του δικτύου. Στην πράξη, χρησιμοποιείται συχνά στο τελικό επίπεδο ενός δικτύου για να προβλέψει την πιθανότητα ενός αποτελέσματος, για παράδειγμα στην δυαδική ταξινόμηση μπορεί να ερμηνεύσει απευθείας την πιθανότητα ένα δείγμα να ανήκει σε μία συγκεκριμένη κλάση. Η χρήση της όμως δεν περιορίζεται εκεί, καθώς αποτελεί βασικό μηχανισμό της αρχιτεκτονικής, των LSTM δικτύων (Long Short-Term Memory), όπου λειτουργεί ως πύλη ελέγχου της ροής πληροφορίας. Εντούτοις, σημαντικό μειονέκτημα της είναι ο κορεσμός των εξαιρετικά μεγάλων τιμών εισόδου, καθώς η έξοδος περιορίζεται ανάμεσα στο μηδέν και το ένα. Αυτό οδηγεί σε παραγωγούς σχεδόν μηδενικές σε σύγκριση με τις αρχικές τιμές εισόδου, οδηγώντας έτσι στο φαινόμενο της εξαφάνισης των παραγώγων ένα σημαντικό πρόβλημα της εκπαίδευσης βαθιών δικτύων.

$$f(u) = \frac{1}{1 + e^{-u}}$$

Υπερβολική εφάπτομένη συνάρτηση (tanh)

Πρόκειται για μια συνάρτηση μεταφοράς που μοιάζει με την σιγμοειδή. Αυτή μετατρέπει κάθε πραγματικό αριθμό της εισόδου σε μια τιμή ανάμεσα στο μείον ένα και το ένα. Είναι συνεχής και παραγωγίσιμη ενώ το βασικό χαρακτηριστικό της είναι η συμμετρικότητα που παρουσιάζει γύρω από το μηδέν. Αυτό σημαίνει ότι οι έξοδοι της μπορούν να πάρουν τόσο θετικές όσο και αρνητικές τιμές, σε αντίθεση με τη σιγμοειδή, η οποία επιστρέφει μόνο θετικές τιμές. Η συμμετρική της φύση βελτιώνει την απόδοση των νευρωνικών δικτύων, επιτρέποντας καλύτερη διάδοση της πληροφορίας και αποφεύγοντας φαινόμενα μετατόπισης τιμών σε θετικές περιοχές, που μπορούν να επηρεάσουν τη σύγκλιση της εκπαίδευσης. Τέλος, παρουσιάζει μεγαλύτερη ευαισθησία από την σιγμοειδή για μικρές τιμές εισόδου γεγονός που την καθιστά ιδιαίτερα χρήσιμη σε εφαρμογές όπου απαιτείται λεπτομερής διαφοροποίηση των δεδομένων εισόδου, όπως για παράδειγμα στην επεξεργασία φυσικής γλώσσας. Χρησιμοποιείται κυρίως όταν δεδομένα παρουσιάζουν ισορροπία ανάμεσα σε θετικές και αρνητικές τιμές καθώς η κατανομή των εξόδων της βοηθά στην ταχύτερη εκμάθηση των βαρών. Ακόμη είναι ιδιαίτερα χρήσιμη στα κρυφά επίπεδα των νευρωνικών δικτύων, όπου απαιτείται μια ομαλή, διαφορίσιμη συνάρτηση για την επεξεργασία των ενδιάμεσων ενεργοποιήσεων. Ιδιαίτερα εμφανίζεται κατά κανόνα σε ακολουθιακά μοντέλα, όπως τα LSTMs και τα GRUs.

$$f(u) = \tanh(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}}$$

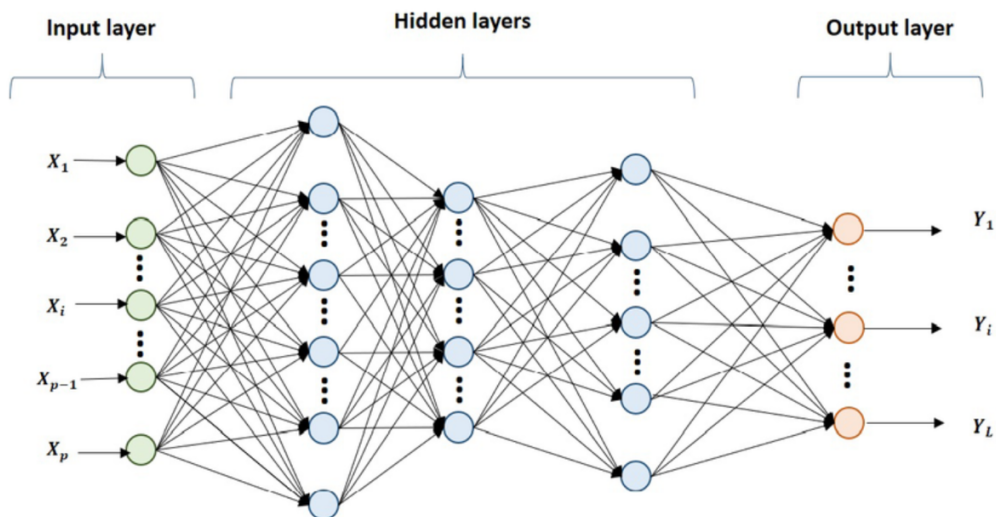
Κανονικοποιημένη Εκθετική Συνάρτηση (Normalized Exponential Function - Softmax)

Το βασικό χαρακτηριστικό της softmax, εκτός από το γεγονός ότι είναι και αυτή μια συνεχής και παραγωγίσιμη συνάρτηση, είναι ότι παίρνει ως είσοδο ένα διάνυσμα πραγματικών αριθμών και το μετατρέπει σε πιθανότητες, οι οποίες αθροίζουν στο ένα. Δηλαδή, μετατρέπει τις ακατέργαστες τιμές εισόδου σε μια κατανομή πιθανότητας, αποδίδοντας σε κάθε κατηγορία μια πιθανότητα που κυμαίνεται μεταξύ (0,1). Τελικά η κατηγορία με τη μεγαλύτερη πιθανότητα επιλέγεται ως έξοδος. Ένα από τα σημαντικά της πλεονεκτήματα, εκτός από τη μετατροπή των εξόδων σε πιθανότητες, είναι και η κανονικοποίηση των εισόδων, αφού τελικά όλες αθροίζονται σε ένα. Επιτρέπει έτσι την απευθείας εφαρμογή συναρτήσεων απώλειας. Η softmax χρησιμοποιείται κυρίως σε μοντέλα ταξινόμησης πολλαπλών κατηγοριών, με εφαρμογές για παράδειγμα στην αναγνώριση φυσικής γλώσσας, όπου το μοντέλο επιλέγει την πιο πιθανή επόμενη λέξη από ένα σύνολο διαφορετικών λέξεων. Ωστόσο, αντιμετωπίζει κάποια σημαντικά μειονεκτήματα. Αρχικά είναι ευαίσθητη σε μεγάλες τιμές εισόδου και έτσι αν οι είσοδοι έχουν μεγάλες διαφορές, υπάρχει το ενδεχόμενο να υπερτονιστεί μία κατηγορία μειώνοντας τις υπόλοιπες. Τελειώνοντας, συγκριτικά με άλλες συναρτήσεις ενεργοποίησης είναι υπολογιστικά απαιτητική, καθώς οι εκθετικοί υπολογισμοί, αυξάνουν την πολυπλοκότητα του μοντέλου. Παρακάτω παρατίθεται ο μαθηματικός τύπος της softmax, όπου όπου x_i είναι η συνιστώσα του διανύσματος εισόδου ενός νευρώνα και n το πλήθος των κατηγοριών.

$$f(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$$

3.4.3 Αρχιτεκτονικές νευρωνικών δικτύων - multilayer perceptrons

Ως μονοστρωματικά νευρωνικά δίκτυα ορίζονται αυτά που αποτελούνται από μόνο ένα επίπεδο παράλληλων νευρώνων, όπου όλοι λαμβάνουν τα ίδια σήματα εισόδου και καθένα έχει μια ανεξάρτητη έξοδο ενώ, κάθε νευρώνας διαθέτει τα δικά του βάρη. Αντίστοιχα, τα πολυστρωματικά νευρωνικά δίκτυα (MLP) αποτελούνται από πολλαπλά επίπεδα νευρώνων, τα οποία διακρίνονται στο επίπεδο της εισόδου, στο επίπεδο της εξόδου και σε όλα τα ενδιάμεσα επίπεδα τα οποία χαρακτηρίζονται ως κρυφά στρώματα. Η έξοδος κάθε νευρώνα μιας στρώσης είναι συνδεδεμένη με την είσοδο καθενός από τους νευρώνες της επόμενης στρώσης. Με αυτό το τρόπο η πληροφορία τροφοδοτείται από τα πρώτα στρώματα στα τελευταία συνθέτοντας έτσι ένα προσθιοδρομικό νευρωνικό δίκτυο. Μια στρώση όπου όλες οι εξοδοί της είναι συνδεδεμένες με τις εισόδους της επόμενης, ονομάζεται πυκνή στρώση και τα νευρωνικά δίκτυα που σχηματίζονται ονομάζονται πυκνά νευρωνικά δίκτυα.



Σχήμα 3.2 : Παράδειγμα MLP

Όπως έχει ήδη αναφερθεί τα πολυστρωματικά δίκτυα (MLP) είναι ένας από τους βασικούς τύπους τεχνητών νευρωνικών δικτύων όπου ανήκουν στην κατηγορία της επιβλεπόμενης μηχανικής μάθησης και χρησιμοποιούνται κυρίως σε προβλήματα παλινδρόμησης και ταξινόμησης. Το κύριο πλεονέκτημα τους είναι η ικανότητά τους να αντιλαμβάνονται τα μη γραμμικά μοτίβα που αναπτύσσονται μεταξύ των δεδομένων με τη χρήση πολλαπλών επιπέδων μη γραμμικών συναρτήσεων ενεργοποίησης. Επιπλέον, είναι ευέλικτα και μπορούν να καλύψουν ένα πολύ μεγάλο εύρος εφαρμογών με κατάλληλη επιλογή της ακριβούς αρχιτεκτονικής και σωστή παραμετροποίηση. Ωστόσο, ένα βασικό μειονέκτημα των MLP είναι ότι δεν εκμεταλλεύονται τη χωρική και χρονική δομή των δεδομένων. Αυτό σημαίνει ότι για παράδειγμα σε ένα πρόβλημα πρόβλεψης χρονοσειράς, η κάθε πρόβλεψη γίνεται αποκλειστικά με βάση την τρέχουσα είσοδο, χωρίς να λαμβάνεται υπόψη η επανεμφάνιση προτύπων που έχουν καταγραφεί στο παρελθόν ή η μακροπρόθεσμη εξάρτηση των δεδομένων. Με τα χρόνια λοιπόν αναπτύχθηκαν πιο περίπλοκες αρχιτεκτονικές.

Τα συνελκτικά δίκτυα (CNN) είναι ένας τύπος νευρωνικών δικτύων που έχει σχεδιαστεί για την επεξεργασία δεδομένων με χωρική δομή, όπως εικόνες και σήματα. Σε αντίθεση με την πυκνή δομή των κλασικών πυκνών δικτύων, αξιοποιούν συνελκτικές και συγκεντρωτικές διεργασίες για να ανιχνεύσουν χαρακτηριστικά στις εισόδους, διατηρώντας παράλληλα τη χωρική σχέση των δεδομένων. Αναλυτικά, αποτελούνται από συνελκτικά επίπεδα (convolutional layers), τα οποία χρησιμοποιούν φίλτρα για την εξαγωγή χαρακτηριστικών από τις εισόδους, επίπεδα συγκέντρωσης (pooling layers) που μειώνουν τη διάσταση των δεδομένων διατηρώντας τις πιο σημαντικές πληροφορίες και τέλος στην έξοδο πλήρως συνδεδεμένα επίπεδα (fully connected layers), όπου τα εξαγόμενα χαρακτηριστικά χρησιμοποιούνται για την τελική ταξινόμηση ή πρόβλεψη. Η χρήση τοπικών φίλτρων οδηγεί σε σταδιακή μείωση των παραμέτρων μειώνοντας το υπολογιστικό κόστος συγκριτικά με τα πολυστρωματικά δίκτυα. Αυτή η προσέγγιση τους επιτρέπει να ανιχνεύουν χωρικά χαρακτηριστικά, όπως άκρες, σχήματα και υφές, σε διαφορετικές περιοχές των δεδομένων

βελτιώνοντας την ικανότητά τους να αναγνωρίζουν μοτίβα ανεξάρτητα από το ακριβές σχήμα τους ή τη θέση τους στην εικόνα. Η ικανότητά τους λοιπόν να αναγνωρίζουν πολύπλοκα χαρακτηριστικά τα καθιστά ιδανικά για προβλήματα όρασης υπολογιστών και ανάλυσης εικόνων και για αυτό χρησιμοποιούνται ευρέως στην αναγνώριση εικόνων, την επεξεργασία βίντεο και την ανίχνευση αντικειμένων [55].

Αντίστοιχα για την αξιοποίηση της χρονικής εξάρτησης των δεδομένων, αναπτύχθηκαν τα αναδρομικά νευρωνικά δίκτυα (RNN). Κύριο χαρακτηριστικό τους αποτελούν οι αναδρομικές συνδέσεις μεταξύ των κρυφών στρωμάτων οι οποίες επιτρέπουν στους νευρώνες να διατηρούν σε εσωτερικές καταστάσεις πληροφορίες από προηγούμενα χρονικά βήματα και να τις χρησιμοποιούν για την επεξεργασία των επόμενων εισόδων προσομοιάζοντας έτσι την ικανότητα εσωτερικής μνήμης. Αυτό σημαίνει ότι η έξοδος του δικτύου επηρεάζεται όχι μόνο από την τρέχουσα είσοδο αλλά και από πληροφορίες του παρελθόντος, επιτρέποντας στα δίκτυα να μάθουν χρονικές εξαρτήσεις και επαναλαμβανόμενα μοτίβα. Στα αναδρομικά δίκτυα τα δεδομένα τροφοδοτούνται ως είσοδοι με ακολουθιακή σειρά και όχι μεμονωμένα όπως στα MLP. Αυτό επιτρέπει την επεξεργασία ευμετάβλητου μεγέθους δεδομένων.

Έτσι λοιπόν καθίστανται ιδανικά για εφαρμογές που απαιτούν την επεξεργασία διαδοχικών δεδομένων όπως πρόβλεψη χρονοσειρών. Ωστόσο, παρουσιάζουν μεγάλη αστάθεια κατά την εκπαίδευση σε πολύ μεγάλες ακολουθίες, αφού είναι τρομερά επιρρεπή στο πρόβλημα της εξαφάνισης των παραγώγων. Τα τελευταία χρόνια έχουν αναπτυχθεί εξελιγμένες αρχιτεκτονικές των αναδρομικών δικτύων όπως τα LSTM και GRU, οι οποίες καταφέρνουν να παραμένουν ευσταθείς για μεγάλο όγκο δεδομένων.

Τέλος αξίζει να αναφερθεί μία επαναστατική αρχιτεκτονική, ο Transformer. Πρόκειται για ένα νευρωνικό δίκτυο ειδικά σχεδιασμένο για την επεξεργασία ακολουθιών δεδομένων, όπως κείμενο, χρονικές σειρές και σήματα. Εισήχθη από τους Vaswani et al. (2017) στο άρθρο "Attention is All You Need" [53] και βασίζεται αποκλειστικά στον μηχανισμό προσοχής (Attention Mechanism) χωρίς τη χρήση επαναλαμβανόμενων (RNN) ή συνελκτικών (CNN) στρωμάτων. Ο πυρήνας της αρχιτεκτονικής είναι ο μηχανισμός Self-Attention, ο οποίος επιτρέπει στο μοντέλο να "εστιάζει" σε διαφορετικά τμήματα της ακολουθίας, αναγνωρίζοντας σχέσεις μεταξύ των στοιχείων της. Αυτό επιτυγχάνεται μέσω των Queries (Q), Keys (K) και Values (V), που χρησιμοποιούνται για τον υπολογισμό βαρών προσοχής. Πολλαπλοί μηχανισμοί προσοχής εφαρμόζονται ταυτόχρονα μέσω του Multi-Head Attention, προσφέροντας στο μοντέλο τη δυνατότητα να αναγνωρίζει πολλαπλά μοτίβα. Ένας Transformer αποτελείται από δύο κύρια μέρη: τον Encoder και τον Decoder. Ο Encoder επεξεργάζεται την είσοδο μέσω πολλαπλών επιπέδων Self-Attention και πλήρως συνδεδεμένων στρωμάτων (Feedforward Layers). Ο Decoder λαμβάνει την έξοδο του Encoder και την επεξεργάζεται για να παράγει την τελική πρόβλεψη, συνήθως μέσω μηχανισμών προσοχής και εξόδου. Επιπλέον, η χρήση Positional Encoding επιτρέπει στο μοντέλο να αναγνωρίζει τη θέση κάθε στοιχείου στην ακολουθία, κάτι που είναι κρίσιμο για χρονικές

εφαρμογές. Οι Transformers προσφέρουν υψηλή ακρίβεια και ευελιξία, επιτρέποντας την παράλληλη επεξεργασία δεδομένων και την αναγνώριση πολύπλοκων σχέσεων. Χρησιμοποιούνται ευρέως σε προβλήματα όπως η πρόβλεψη φορτίου ενέργειας, η μηχανική μετάφραση και η αναγνώριση ομιλίας [54].

Η επέκτασή τους σε Temporal Fusion Transformers (TFT) επιτρέπει την πρόβλεψη φορτίου σε χρονικές σειρές, ενσωματώνοντας πολλαπλές εξωγενείς μεταβλητές. Αυτά αξιοποιούν την αρχιτεκτονική των Transformers για να παρέχουν εξαιρετική απόδοση σε προβλήματα χρονικών σειρών, όπως η πρόβλεψη φορτίου. Το κύριο πλεονέκτημά τους είναι η δυνατότητα ενσωμάτωσης εξωγενών χαρακτηριστικών (όπως θερμοκρασία, ημερομηνία, ημέρα της εβδομάδας) και η ερμηνευσιμότητα μέσω μηχανισμών προσοχής, που αποκαλύπτουν τη σημασία κάθε χαρακτηριστικού για την πρόβλεψη. Τα TFT επιτρέπουν στο μοντέλο να αναγνωρίζει τόσο μακροχρόνιες όσο και βραχυχρόνιες εξαρτήσεις στα δεδομένα, ενώ χρησιμοποιούν προηγμένα συστατικά όπως gating mechanisms για την προσαρμογή της συνεισφοράς κάθε χαρακτηριστικού. Σε σύγκριση με τους κλασικούς Transformers, τα TFT είναι βελτιστοποιημένα για δεδομένα χρονοσειρών, καθιστώντας τα ιδανικά για την πρόβλεψη φορτίου, όπου η ιστορική πληροφορία και τα εξωτερικά δεδομένα επηρεάζουν άμεσα τη ζήτηση [57].

3.4.4 Εκπαίδευση νευρωνικού δικτύου – Επιβλεπόμενη μάθηση

Μέχρι στιγμής έχει πραγματοποιηθεί μια γενική ανάλυση των δομικών χαρακτηριστικών των νευρωνικών δικτύων. Αρχικά έγινε αναφορά στους τεχνητούς νευρώνες που αποτελούν το θεμελιώδες στοιχείο των νευρωνικών δικτύων, στη συνέχεια επεξηγήθηκε ο τρόπος με τον οποίο οι διάφορες συναρτήσεις ενεργοποίησης επιτρέπουν στους νευρώνες να ξεπερνούν τους περιορισμούς ως προς τη γραμμική τους φύση, ενώ τέλος διατυπώθηκαν διαφορετικές αρχιτεκτονικές νευρωνικών δικτύων. Σε αυτή την ενότητα λοιπόν αναλύεται ο τρόπος με τον οποίο τα νευρωνικά δίκτυα εκπαιδεύονται με την μέθοδο της επιβλεπόμενης μάθησης. Ο όρος εκπαίδευση περιγράφει τη διαδικασία προσαρμογής των διάφορων ενδιάμεσων παραμέτρων ενός δικτύου με βάση σύνολα δεδομένων ώστε να μπορεί να γενικεύει αποδοτικά εργασίες σχετικά με ευρύτερα σύνολα αυτών των δεδομένων. Παραδείγματα τέτοιων εργασιών είναι η πρόβλεψη, η ταξινόμηση, η ομαδοποίηση και η αναγνώριση. Η εκπαίδευση λοιπόν ενός νευρωνικού δικτύου αναφέρεται στη διαδικασία κατά την οποία το δίκτυο ελαχιστοποιεί το σφάλμα μεταξύ της πραγματικής και της επιθυμητής εξόδου, προσαρμόζοντας σταδιακά τις εσωτερικές του παραμέτρους, δηλαδή τα βάρη των εισόδων των νευρώνων και τα μεγέθη της προκατάληψης.

Η διαδικασία πραγματοποιείται σε τέσσερις διακριτές φάσεις, την πρόσθια διάδοση, τον υπολογισμό του σφάλματος, την οπίσθια διάδοση όπου εφαρμόζεται ο αλγόριθμος της οπισθοδιάδοσης (backpropagation) και τέλος την ενημέρωση των βαρών με στόχο την ελαχιστοποίηση του σφάλματος με χρήση

του αλγορίθμου της διαβάθμισης κλίσης (gradient descent). Η παραπάνω διαδικασία επαναλαμβάνεται μέχρι να επιτευχθεί μια συνθήκη τερματισμού. Κάθε βήμα της επανάληψης χαρακτηρίζεται στη βιβλιογραφία ως παρτίδα (batch) και σε κάθε παρτίδα γίνεται χρήση του συνόλου των δεδομένων [50].

Εμπρόσθια διάδοση (forward pass)

Αρχικά κατά τη φάση της εμπρόσθιας διάδοσης (forward pass) εισάγεται στο δίκτυο των σύνολο των δεδομένων προς εκπαίδευση και υπολογίζονται οι αντίστοιχες έξοδοι. Αν πρόκειται για την πρώτη παρτίδα της εκπαίδευσης, όπου δεν έχει προηγηθεί κάποιος υπολογισμός των παραμέτρων του δικτύου, δηλαδή τα βάρη και η προκατάληψη για κάθε νευρώνα, αυτές αρχικοποιούνται με βάση τυχαίες και μικρές τιμές.

Αναλυτικά, στο στρώμα της εισόδου εισάγεται το σύνολο των χαρακτηριστικών των δεδομένων, με κάθε χαρακτηριστικό να εισάγεται σε μια ξεχωριστή είσοδο. Κάθε χαρακτηριστικό αποτελεί στοιχείο του διανύσματος εισόδου. Στη συνέχεια το σύνολο των στοιχείων του διανύσματος εισόδου τροφοδοτούνται ως είσοδοι σε έναν αριθμό νευρώνων του επόμενου στρώματος του δικτύου. Εκεί πολλαπλασιάζονται με τα αντίστοιχα βαρύ και προστίθεται η σταθερά της προκατάληψης. Το αποτέλεσμα αυτού του γραμμικού μετασχηματισμού εισάγεται σε μια συνάρτηση ενεργοποίησης, που συνήθως είναι μη γραμμική και έτσι δημιουργείται η έξοδος του αντίστοιχου νευρώνα. Αυτή η έξοδος αποτελεί κομμάτι της κρυφής κατάστασης του δικτύου που αφορά πληροφορία για έναν συγκεκριμένο νευρώνα, η οποία αποθηκεύεται και χρησιμοποιείται ως είσοδος για τους νευρώνες του επόμενου στρώματος. Το ίδιο μοτίβο επαναλαμβάνεται μέχρι και το τελευταίο στρώμα όπου και παράγεται η τελική έξοδος του δικτύου. Αυτή μπορεί να είναι συνεχής ή διακριτή και μπορεί να είναι μοναδική ή να αποτελεί διάνυσμα τιμών εξόδου, αναλόγως την αρχιτεκτονική και το πλήθος εξόδων του δικτύου.

Στην παρούσα εργασία όλες οι αρχιτεκτονικές που αναπτύσσονται έχουν 128 εισόδους και 36 εξόδους συνεχών τιμών.

Έτσι, η πληροφορία διαδίδεται από το πρώτο στρώμα σταδιακά μέσω όλων των επιπέδων του δικτύου, με κάθε επίπεδο να εξάγει πιο περίπλοκα και αφηρημένα χαρακτηριστικά των αρχικών δεδομένων. Στο τελευταίο επίπεδο, οι έξοδοι των νευρώνων συνδυάζονται και παράγουν την τελική έξοδο του δικτύου. Αυτή η έξοδος μπορεί να είναι, για παράδειγμα, μια πρόβλεψη, μια ταξινόμηση ή κάποια άλλη επιθυμητή απάντηση. Με αυτόν τον τρόπο, κάθε στρώμα επεξεργάζεται την είσοδο που λαμβάνει, μετατρέπει την πληροφορία με βάση τα τρέχοντα βάρη και τις προκαταλήψεις του, και τη μεταβιβάζει στο επόμενο επίπεδο, οδηγώντας τελικά σε μια τελική έξοδο που αντανakλά την ολοκληρωμένη επεξεργασία των αρχικών δεδομένων. Καθ' όλη τη διάρκεια της προώθησης, αποθηκεύονται όλες οι ενδιάμεσες έξοδοι των νευρώνων ώστε να είναι διαθέσιμες για την προσαρμογή των βαρών.

Υπενθυμίζεται ότι μαθηματικά ο γραμμικός μετασχηματισμός ενός νευρώνα i που βρίσκεται στο στρώμα l περιγράφεται από τον παρακάτω τύπο :

$$z_i^l = \sum_k w_k^l \cdot a_k^{l-1} + b_i^l$$

Ενώ η έξοδος αυτού το νευρώνα είναι το αποτέλεσμα του γραμμικού μετασχηματισμού ο οποίος εισάγεται σε μια συνάρτηση ενεργοποίησης.

$$a_i^l = f(z_i^l)$$

Υπολογισμός του σφάλματος

Στη συνέχεια υπολογίζεται το σφάλμα της εκπαίδευσης, δηλαδή η διαφορά μεταξύ της επιθυμητής τιμής εξόδου y και της πραγματικής εξόδου $\hat{y}$. Αυτή η διαφορά υπολογίζεται με διάφορους τρόπους ανάλογα την εκάστοτε εφαρμογή. Έτσι για προβλήματα ταξινόμησης η συνάρτηση Cross-Entropy αποτελεί έναν βασικό μετρικό δείκτη σφάλματος. Σε προβλήματα δυαδικής ταξινόμησης χρησιμοποιείται η Binary Cross-Entropy ενώ για ταξινόμηση περισσότερων κατηγοριών επιλέγεται η Categorical Cross-Entropy σε συνδυασμό με softmax στην έξοδο. από την άλλη για εφαρμογές προβλέψεις η συνάρτηση σφάλματος ταυτίζεται με τις μεθόδους αξιολόγησης προβλέψεων όπως αναφέρθηκαν στο προηγούμενο κεφάλαιο. Για παράδειγμα με χρήση του μέσου τετραγωνικού σφάλματος για συνάρτηση απωλειών προκύπτει ο παρακάτω μαθηματικός τύπος με k να αντιπροσωπεύει το πλήθος δεδομένων της παρτίδας εκπαίδευσης.

$$J = cost(\hat{y}, y) = \frac{1}{n} \sum_k (\hat{y}_k - y_k)^2$$

Οπισθοδιάδοση – backpropagation

Ο αλγόριθμος της οπισθοδιάδοσης – backpropagation αποτελεί τον πιο βασικό αλγόριθμο κατά την διαδικασία εκπαίδευσης των νευρωνικών δικτύων με την διαδικασία της επιβλεπόμενης μάθησης καθώς αυτός ή παραλλαγές του, συναντώνται σε όλα τα παραδείγματα εκπαίδευσης. Σκοπός είναι η διάδοση του σφάλματος εκπαίδευσης προς τα πίσω, από την έξοδο του δικτύου ως τα πρώτα στρώματα των εισόδων, ώστε να υπολογιστούν οι παράγωγοι της συνάρτησης απωλειών ως προς κάθε εσωτερική παράμετρο. Αυτές οι παράγωγοι αποτυπώνουν το μέγεθος και την κατεύθυνση της επίδρασης κάθε ενδιάμεσης παραμέτρου στο τελικό σφάλμα και για τον υπολογισμό τους γίνεται χρήση του κανόνα της αλυσίδας του Leibnitz.

Αντίστοιχα ορίζεται ως τοπικό σφάλμα ενός νευρώνα, η μερική παράγωγος της συνάρτησης απωλειών ως προς τον γραμμικό μετασχηματισμό του συγκεκριμένου νευρώνα. Το τοπικό σφάλμα περιγράφει την επίδραση του μετασχηματισμού αυτού στο συνολικό σφάλμα.

$$\delta_i^l = \frac{\partial J}{\partial z_i^l}$$

Αναλυτικά, ξεκινώντας από το τελευταίο στρώμα (L), υπολογίζεται το τοπικό σφάλμα των νευρώνων της εξόδου. Πρόκειται για την παράγωγο της συνάρτησης απωλειών ως προς την έξοδο, που είναι γνωστό μέγεθος, πολλαπλασιασμένο με την αντίστοιχη παράγωγο της συνάρτησης ενεργοποίησης του νευρώνα.

$$\delta_i^L = \frac{\partial J}{\partial z_i^L} = \frac{\partial J}{\partial \alpha_i^L} \cdot \frac{\partial \alpha_i^L}{\partial z_i^L} = \frac{\partial J}{\partial \alpha_i^L} \cdot f'(z_i^L).$$

Στη συνέχεια, για κάθε νευρώνα σε κάθε κρυφό στρώμα (l), υπολογίζεται το αντίστοιχο τοπικό σφάλμα ως το σταθμισμένο άθροισμα του γινομένου όλων των τοπικών σφαλμάτων των νευρώνων του επόμενου στρώματος με τους οποίους είναι συνδεδεμένος και των αντίστοιχων βαρών, πολλαπλασιασμένο με τη παράγωγο της συνάρτησης ενεργοποίησης του νευρώνα. Αυτή η διαδικασία επαναλαμβάνεται αναδρομικά από το τελικό στρώμα προς το πρώτο στρώμα του δικτύου, αφού όπως γίνεται αντιληπτό τα τοπικά σφάλματα των νευρώνων σχετίζονται αναδρομικά

$$\frac{\partial J}{\partial z_i^l} = \delta_i^l = f'(z_i^l) \cdot \sum_k \delta_k^{l+1} \cdot w_k^{l+1}$$

Γνωρίζοντας τα τοπικά σφάλματα μπορεί να γίνει ο υπολογισμός των αντίστοιχων παραμέτρων των βαρών και της προκατάληψης. Έτσι για τον υπολογισμό της επίδρασης κάθε βάρους ενός τεχνητού νευρώνα ισχύει:

$$\frac{\partial J}{\partial w_i^l} = \frac{\partial J}{\partial z_i^l} \cdot \frac{\partial z_i^l}{\partial w_i^l} = \delta_i^l \cdot \alpha_i^{l-1}, \text{ αφού } \frac{\partial z_i^l}{\partial w_i^l} = \frac{\partial (w_i^l \cdot \alpha_j^{l-1} + b_i^l)}{\partial z_i^l} = \alpha_i^{l-1}$$

Όπου α_i^{l-1} η έξοδος των νευρώνων του προηγούμενου στρώματος που είναι συνδεδεμένοι ως είσοδοι με τον συγκεκριμένο νευρώνα. Σημειώνεται ότι η τιμή της εξόδου αυτής είναι αποθηκευμένη από το βήμα της πρόσθιας διάδοσης.

Αντίστοιχα για την επίδραση του μέγεθος της προκατάληψης ισχύει :

$$\frac{\partial J}{\partial b_i^l} = \frac{\partial J}{\partial z_i^l} \cdot \frac{\partial z_i^l}{\partial b_i^l} = \frac{\partial J}{\partial z_i^l} \cdot 1 = \delta_i^l, \text{ αφού } \frac{\partial z_i^l}{\partial b_i^l} = \frac{\partial (w_i^l \cdot \alpha_j^{l-1} + b_i^l)}{\partial b_i^l} = 1$$

Έτσι τελικά υπολογίζεται το σύνολο των Μερικών παραγωγών της συνάρτησης απωλειών και σχηματίζεται η συνολική παράγωγος της συνάρτησης ($\nabla J = \text{grad } J$)

Διαβάθμιση κλίσης - Gradient decent

Η μέθοδος διαβάθμιση κλίσης αποτελεί έναν αλγόριθμο βελτιστοποίησης που χρησιμοποιείται για την ελαχιστοποίηση της συνάρτησης κόστους. Η βασική ιδέα είναι ότι, δεδομένου ενός συνόλου παραμέτρων θ , μπορούμε να βελτιώσουμε το μοντέλο μειώνοντας σταδιακά την τιμή της συνάρτησης κόστους $J(\theta)$ [51]. Η λογική του gradient descent στηρίζεται στην παραδοχή ότι η συνάρτηση κόστους είναι τοπικά προσεγγίσιμη με γραμμικές προσεγγίσεις, δηλαδή μέσω της πρώτης παραγώγου, η οποία έχει υπολογιστεί προηγουμένως με τον αλγόριθμο της οπισθοδιάδοσης και η οποία δείχνει την κατεύθυνση της ταχύτερης αύξησης της συνάρτησης κόστους. Για την ελαχιστοποίηση της επομένως, πρέπει να προχωρήσουμε στην αντίθετη κατεύθυνση. Επομένως, με μικρά, συνεχόμενα βήματα μπορεί να μειωθεί το σφάλμα εκπαίδευσης μέχρι να προσεγγιστεί μια τιμή κοντά στο ολικό ελάχιστο της συνάρτησης απωλειών. Αξίζει να σημειωθεί πως δεν αποδεικνύεται ότι πρόκειται πάντα για την ολική ελάχιστη τιμή ή απλώς για μια τοπική ελάχιστη τιμή.

Η ενημέρωση των παραμέτρων σε κάθε παρτίδα γίνεται μέσω του τύπου:

$$\theta_{new} = \theta_{old} - \eta \nabla J(\theta_{old})$$

Η επιλογή του βήματος της μάθησης η αποτελεί έναν από τους σημαντικότερους παράγοντες κατά την εκπαίδευση καθώς μια βέλτιστη τιμή επιτρέπει την ομαλή ενημέρωση των βαρών, οδηγώντας στη σταθερή και ταχεία σύγκλιση της συνάρτησης κόστους, εξοικονομώντας υπολογιστικούς πόρους. Η σωστή ρύθμιση του εξαρτάται από τη δομή του δικτύου και τα χαρακτηριστικά των δεδομένων, και μπορεί να επιτευχθεί με τεχνικές προσαρμογής, όπως η χρήση adaptive learning rate optimizers.

Αντίθετα όταν το βήμα είναι μικρότερο ή μεγαλύτερο το δίκτυο είτε προσεγγίζει αργά είτε καθόλου το ολικό ελάχιστο της συνάρτησης απωλειών. Επιγραμματικά σε περιπτώσεις όπου το βήμα εκπαίδευσης είναι πολύ μικρό όχι μόνο καθυστερεί σημαντικά η διαδικασία μάθησης αλλά υπάρχει ο κίνδυνος να κολλήσει το δίκτυο σε τοπικά ελάχιστα της συνάρτησης απωλειών, θεωρώντας εσφαλμένα ότι έχει βρει τις βέλτιστες παραμέτρους ή ακόμη ένας πολύ μικρός ρυθμός μάθησης ενδέχεται να μην επιτρέπει την επαρκή προσαρμογή στις μεταβολές του σφάλματος, αποτρέποντας τη γενίκευση του μοντέλου σε άγνωστα δεδομένα. Από την άλλη πλευρά, όταν ο ρυθμός εκπαίδευσης είναι υπερβολικά μεγάλος, το δίκτυο οδηγείται σε αστάθεια, με αποτέλεσμα το σφάλμα να μην μειώνεται ομαλά και να παρουσιάζει ταλαντώσεις αφού υπερπηδάει το ελάχιστο εξαιτίας του μεγάλου βήματος. Σε τέτοιες περιπτώσεις, το μοντέλο μπορεί να μην συγκλίνει καθόλου ή ακόμη και να αυξάνει το σφάλμα αντί να το μειώνει.

Αλγόριθμοι βελτιστοποίησης

Ο αλγόριθμος κατάβασης κλίσης στην κλασική του μορφή περιγράφεται ως batch gradient descent και όπως έχει ήδη αναφερθεί εντάσσει στην παρτίδα

εκπαίδευσης (batch) ολόκληρο το σύνολο των δεδομένων εκπαίδευσης. Εξαιτίας λοιπόν της χρήσης ολόκληρου του συνόλου δεδομένων, οι ενημερώσεις των βαρών είναι πιο σταθερές και δεν επηρεάζονται από τον θόρυβο των μεμονωμένων δειγμάτων, με αποτέλεσμα την ομαλή και σταθερή σύγκλιση. Ενώ, ταυτόχρονα καθώς η κλίση που υπολογίζεται βασίζεται σε όλα τα δεδομένα, δεν εμφανίζει μεγάλες διακυμάνσεις και επομένως η σύγκλιση είναι περισσότερο προβλέψιμη. Ωστόσο ο κλασικός αλγόριθμος έχει αρκετά μειονεκτήματα, ειδικά σε εφαρμογές όπου υπάρχουν μεγάλα σύνολα δεδομένων. Δεδομένου ότι οι ενημερώσεις των βαρών γίνονται μόνο αφού γίνει χρήση όλων των δεδομένων η διαδικασία της εκπαίδευσης καταλήγει να είναι πολύ αργή ενώ ταυτόχρονα χρειάζονται μεγάλοι υπολογιστικοί πόροι και τεράστια διαθέσιμη μνήμη για την αποθήκευση όλων των ενδιαμέσων παραμέτρων του δικτύου με είσοδο όλα τα δεδομένα προκειμένου να γίνει η ενημέρωση. Τέλος, άμεση συνέπεια του σταθερού χαρακτήρα της σύγκλισης του αλγορίθμου είναι οι αυξημένες πιθανότητες να παγιδευτεί σε τοπικά ελάχιστα ή επίπεδες περιοχές της συνάρτησης κόστους καθώς οι ενημερώσεις παρουσιάζουν μεγαλύτερη σταθερότητα.

Οι παραπάνω περιορισμοί οδήγησαν στην ανάπτυξη βελτιστοποιημένων αλγορίθμων εκμάθησης], οι οποίοι διαφοροποιούνται είτε ως προς το σύνολο των δεδομένων που συνυπολογίζουν για την ενημέρωση των παραμέτρων, είτε ενσωματώνουν στο κύριο μηχανισμό πρόσθετα χαρακτηριστικά όπως την αδράνεια και την έξυπνη διαχείριση των κατευθύνσεων ενημέρωσης, είτε τέλος πρόσθετες πληροφορίες σαν την ακριβέστερη ρύθμιση του βήματος εκπαίδευσης μέσω προσαρμοστικότητας. Αυτές οι τεχνικές επιταχύνουν τη σύγκλιση με μεγαλύτερη πιθανότητα και ταχύτητα ενώ, επιτρέπουν καλύτερη απόδοση σε πολύπλοκα προβλήματα βελτιστοποίησης. Παρακάτω αναφέρονται οι επικρατέστεροι αλγόριθμοι βελτιστοποίησης που υπάρχουν στη βιβλιογραφία.

Αρχικά υπάρχουν παραλλαγές του αλγορίθμου που σχετίζονται με τον αριθμό των δειγμάτων που χρησιμοποιούνται για την ενημέρωση των παραμέτρων (batch). από τη μια πλευρά ο Στοχαστικός Αλγόριθμος Κατάβασης Κλίσης (Stochastic Gradient Descent - SGD) ενημερώνει τις παραμέτρους του μοντέλου μετά από κάθε μεμονωμένο δείγμα εκπαίδευσης. Αυτό τον καθιστά πολύ πιο αποδοτικό υπολογιστικά, ιδιαίτερα σε μεγάλα σύνολα δεδομένων, καθώς μειώνεται η απαιτούμενη μνήμη που απαιτείται για την αποθήκευση των παραμέτρων ενώ ταυτόχρονα επιταχύνεται η διαδικασία της εκπαίδευσης, καθώς αποδεικνύεται πως η διακύμανση της κλίσης οδηγεί σε σύγκλιση με λιγότερα βήματα. Ωστόσο, η στοχαστική φύση του προκαλεί μεγαλύτερη διακύμανση στις ενημερώσεις, οδηγώντας σε πιο ασταθή διαδρομή προς το ελάχιστο. Παρά την εισαγωγή θορύβου από τις διακυμάνσεις, αυτές παρουσιάζουν κάποια πλεονεκτήματα, όπως την υπερπήδηση τοπικών ελαχίστων και επίσης την καλύτερη γενίκευση του μοντέλου. από την άλλη πλευρά, έχει αναπτυχθεί ακόμη μια παραλλαγή που αποτελεί έναν συνδυασμό του κλασσικού και του στοχαστικού αλγορίθμου, και ονομάζεται Mini-batch Gradient Descent. Σε αυτή τη περίπτωση οι παράμετροι ενημερώνονται με βάση μικρά υποσύνολα (mini-batches) των δεδομένων αντί για ένα μόνο δείγμα ή ολόκληρο το σύνολο. Αυτή η προσέγγιση εξισορροπεί την

υπολογιστική αποδοτικότητα και τη σταθερότητα της εκπαίδευσης, καθώς περιορίζει το θορυβώδη χαρακτήρα των ενημερώσεων του στοχαστικού αλγορίθμου, ενώ παράλληλα βελτιώνει την ταχύτητα σύγκλισης και μειώνει το υπολογιστικό κόστος συγκριτικά με τον κλασσικό αλγόριθμο.

Συνεχίζοντας, υπάρχουν πρόσθετες παραλλαγές οι οποίες εισάγουν πρόσθετες πληροφορίες κατά την ενημέρωση των παραμέτρων εκτός από την κλίση και το βήμα της μάθησης. Έτσι ο Momentum Gradient Descent εισάγει την έννοια της ορμής (momentum). Η ενημέρωση των παραμέτρων δεν εξαρτάται μόνο από την τρέχουσα κλίση και το βήμα εκπαίδευσης, αλλά και από τις προηγούμενες υπολογισμένες κλίσεις προσομοιώνοντας έτσι την συμπεριφορά της ορμής κατά την κατάβαση στο ελάχιστο της συνάρτησης απωλειών. Αυτή η διατήρηση της ορμής επιτρέπει στον αλγόριθμο να υπερπηδάει τοπικά ελάχιστα ή επίπεδες επιφάνειες της συνάρτησης κόστους ενώ ταυτόχρονα μπορεί να προσαρμόζεται στο χαρακτήρα της σύγκλισης. Έτσι όταν η κατεύθυνση της κίνησης δεν αλλάζει, δηλαδή η μεγαλύτερη πλειοψηφία των προηγούμενων παραγώγων έχουν το ίδιο πρόσημο, εξαιτίας της ορμής οι παράμετροι ενημερώνονται με μεγαλύτερο ρυθμό από το βήμα εκπαίδευσης με αποτέλεσμα την ταχύτερη σύγκλιση ενώ σε περιπτώσεις όπου παρουσιάζονται πολλές αλλαγές στη κατεύθυνση, δηλαδή οι προηγούμενες παραγώγους έχουν και θετικό και αρνητικό πρόσημο, ο ρυθμός ενημέρωσης μειώνεται. Μια βελτιωμένη εκδοχή του παραπάνω αποτελεί ο Nesterov Accelerated Gradient Descent που στοχεύει σε ταχύτερη και πιο σταθερή σύγκλιση καθώς προβλέπει το επόμενο σημείο ενημέρωσης και υπολογίζει την κλίση εκεί, αντί για το τρέχον σημείο. Αυτό επιτρέπει στο μοντέλο να "βλέπει μπροστά", διορθώνοντας έγκαιρα την κατεύθυνση της κίνησης και αποφεύγοντας περιττές ταλαντώσεις.

$$\theta_{k+1} = \theta_k + u_k \quad \text{με} \quad u_k = \eta \cdot \sum_{i=1}^k \beta^{k-i} \nabla J(\theta_i)$$

Τέλος μια τελευταία ομάδα βελτιστοποιήσεων αφορούν στην προσαρμογή του βήματος εκπαίδευσης για κάθε παράμετρο ξεχωριστά. Πρώτα ο αλγόριθμος AdaGrad εξομαλύνει το βήμα εκπαίδευσης σύμφωνα με το άθροισμα τετραγώνων όλων των προηγούμενων παραγώγων της αντίστοιχης παραμέτρου. Αυτό σημαίνει ότι παράμετροι που ενημερώνονται συχνά αποκτούν μικρότερο βήμα, καθώς το σύνολο των όρων του αθροίσματος αυξάνεται, ενώ οι λιγότερο συχνά ενημερωμένες παράμετροι έχουν πιο υψηλό ρυθμό εκπαίδευσης. Ωστόσο, ένα βασικό μειονέκτημα του είναι ότι ο ρυθμός μάθησης μειώνεται μονοτονικά σε κάθε επανάληψη και μπορεί να γίνει υπερβολικά μικρός, εμποδίζοντας την περαιτέρω εκπαίδευση του μοντέλου μετά από κάποιο σημείο. Για την αντιμετώπιση του προβλήματος της συνεχούς μείωσης του ρυθμού μάθησης εξαιτίας της νομοτελειακής αύξησης των όρων του αθροίσματος τετραγώνων με κάθε ενημέρωση προέκυψε ο αλγόριθμος RMSProp (Root Mean Square Propagation) όπου κατά την ενημέρωση του βήματος χρησιμοποιεί τον εκθετικά σταθμισμένο μέσο όρο των τετραγώνων των προηγούμενων παραγώγων της συνάρτησης απωλειών ως προς τη συγκεκριμένη παράμετρο. Έτσι αποτρέπεται η υπερβολική μείωση του βήματος εκπαίδευσης.

$$\theta_{new} = \theta_{old} - \frac{\eta}{\sqrt{G_t + \varepsilon}} \nabla J(\theta_{old}) \text{ όπου :}$$

$$G_t = \sum_{i=1}^t \nabla J(\theta_i)^2 \text{ για AdaGrad και}$$

$$G_t = \sum_{i=0}^t \beta(1 - \beta)^i \cdot \nabla J(\theta_{t-i})^2 \text{ για RMSProp}$$

Ιδιαίτερη αναφορά ωστόσο πρέπει να γίνει στον αλγόριθμο Adam (Adaptive Moment Estimation), όπου συνδυάζει τα πλεονεκτήματα του Momentum GD και του RMSProp. Η κύρια ιδέα λειτουργίας του είναι πως χρησιμοποιεί δύο εκθετικά κινούμενους μέσους όρους για να συνδυάσει την ιδέα της ορμής για τις ενημερώσεις των βαρών και την ιδέα της προσαρμοστικότητας του βήματος εκπαίδευσης[56,57].

$$\theta_{new} = \theta_{old} - \frac{\eta}{\sqrt{G_t + \varepsilon}} \cdot \sum_{t=1}^{\kappa} \beta^{\kappa-t} \cdot \nabla J(\theta_t), \quad \mu\epsilon \quad G_t = \sum_{i=0}^t \beta(1 - \beta)^i \cdot \nabla J(\theta_{t-i})^2$$

3.5 Προεπεξεργασία μεταβλητών εισόδου για νευρωνικά δίκτυα σε εφαρμογές πρόβλεψης

Η προεπεξεργασία των μεταβλητών εισόδου είναι κρίσιμη για τη βελτίωση της απόδοσης των νευρωνικών δικτύων στις προβλέψεις φορτίου ενέργειας επόμενης ημέρας. Η κατάλληλη προεπεξεργασία διασφαλίζει ότι οι πληροφορίες εισάγονται στο μοντέλο με τρόπο που μεγιστοποιεί τη συνεισφορά τους ενώ, ταυτόχρονα μειώνεται η πολυπλοκότητα του μοντέλου και ο κίνδυνος υπερεκπαίδευσης. Η προεπεξεργασία σαν διαδικασία περιλαμβάνει διάφορα στάδια μετατροπής των δεδομένων και παρακάτω περιγράφονται οι πιο μείζονες από αυτές.

Feature Scaling με Κανονικοποίηση ή Τυποποίηση

Η κλιμάκωση των αριθμητικών εξωγενών μεταβλητών μέσω κανονικοποίησης (Min–Max Scaling) ή τυποποίησης (Z-Score Normalization) αποτελεί καίριο βήμα της προεπεξεργασίας. Στην κανονικοποίηση, κάθε τιμή ενός χαρακτηριστικού μετασχηματίζεται ώστε να κυμαίνεται σε προκαθορισμένο εύρος, συνήθως μεταξύ 0 και 1. Έτσι ενώ διατηρείται η κατανομή των δεδομένων, τυχών σημεία με ακραίες μέγιστες ή ελάχιστες τιμές συμπιέζονται. από την άλλη, με τη μέθοδο τυποποίησης επιτυγχάνεται κατανομή με μηδενικό μέσο όρο και μοναδιαία διασπορά. Η μέθοδος αυτή αντιμετωπίζει καλύτερα τα ακραία σημεία σε σχέση με την κανονικοποίηση αλλά δεν εγγυάται σταθερό εύρος τιμών. Και οι δύο τεχνικές βελτιώνουν τη σταθερότητα και την ταχύτητα σύγκλισης κατά τη διαδικασία της μάθησης, αποτρέποντας τα χαρακτηριστικά με αφύσικα μεγάλες τιμές να επηρεάζουν την ενημέρωση των βαρών. Η επιλογή τους εξαρτάται από τη φύση των δεδομένων. Συνήθως αν υπάρχει ανάγκη όλα τα χαρακτηριστικά να βρίσκονται σε κοινή, προκαθορισμένη κλίμακα, για παράδειγμα εικονοστοιχεία

από εικόνες, προτιμάται η κανονικοποίηση ενώ αν υπάρχουν έντονα outliers ή η κατανομή των δεδομένων παρουσιάζει ιδιομορφίες και συνάδει με την κανονική κατανομή, όπως για παράδειγμα ασυμμετρίες ή ασυνέχειες, καλύτερη επιλογή είναι η τυποποίηση καθώς με αυτή τη μέθοδο δεν αλλάζει το σχήμα της κατανομής

Κωδικοποίηση περιοδικών χρονικών πληροφοριών

Οι μεταβλητές που σχετίζονται περιοδικά με το χρόνο, όπως η ημέρα της εβδομάδας, η ώρα της ημέρας ή ο μήνας του έτους, κωδικοποιούνται με ημιτονοειδείς και συνημιτονοειδείς συναρτήσεις. Η βασική ιδέα αυτών των κωδικοποιήσεων για περιοδικές χρονικές μεταβλητές είναι να μετατρέπουμε την περιοδικότητα-κυκλικότητα του χρόνου σε συνεχείς διανυσματικές αναπαραστάσεις, ώστε ένα μοντέλο να κατανοεί ότι το “τέλος” του κύκλου συνδέεται ξανά με την αρχή του. Πρακτικά για μια μεταβλητή x με περίοδο P , υπολογίζονται δύο τιμές:

$$\sin\left(2\pi \frac{x}{P}\right) \text{ και } \cos\left(2\pi \frac{x}{P}\right)$$

Αυτές οι δύο τιμές, που περιγράφονται και σαν συντεταγμένες σε έναν κύκλο, αποδίδουν μοναδική θέση για κάθε σημείο του κύκλου, διασφαλίζοντας ότι τα γειτονικά σημεία (π.χ. 23:00 και 0:00) βρίσκονται κοντά στον χώρο των χαρακτηριστικών, ενώ τα αντίθετα (π.χ. 6:00 και 18:00) απέχουν πολύ. Με αυτόν τον τρόπο ένα μοντέλο κατανοεί τόσο την απόλυτη θέση στη χρονική περίοδο όσο και τη σχετική γειτνίαση των ωρών ή των ημερών, διατηρώντας παράλληλα τη φυσική επαναληπτικότητα και τις εποχιακές μεταβολές των δεδομένων.

Κωδικοποίηση κατηγορικών μεταβλητών

Ως κατηγορικές μεταβλητές ορίζονται οι μεταβλητές που περιγράφουν δεδομένα τα οποία ανήκουν σε διακριτές κατηγορίες, χωρίς να έχουν αριθμητική ή ποσοτική σημασία, και επομένως δεν μπορούν να χρησιμοποιηθούν για μαθηματικές πράξεις, όπως η πρόσθεση ή η αφαίρεση. Κατά την κωδικοποίηση, οι μεταβλητές αυτές μετασχηματίζονται σε αριθμητικές μορφές, ώστε να μπορούν να χρησιμοποιηθούν από νευρωνικά δίκτυα που απαιτούν αριθμητικές εισόδους.

Μια τέτοια μέθοδος κωδικοποίησης είναι η one-hot encoding, η οποία μετατρέπει κάθε κατηγορία μιας κατηγορικής μεταβλητής σε έναν δυαδικό διανυσματικό κώδικα. Έτσι, για κάθε μοναδική τιμή της κατηγορικής μεταβλητής δημιουργείται μια νέα στήλη που περιέχει 0 ή 1. Το one-hot encoding διατηρεί τη διάκριση μεταξύ κατηγοριών χωρίς να εισάγει ψευδή σειρά, είναι εύκολα κατανοητή μέθοδος και συμβατή με πολλούς αλγορίθμους μηχανικής μάθησης. Ωστόσο, αυξάνει το αριθμό των διαστάσεων που απαιτούνται για περιγραφή των δεδομένων, καθώς οι διαστάσεις ισούνται με το σύνολο των διαφορετικών κατηγοριών, με αποτέλεσμα σε εφαρμογές με πολλές κατηγορίες να προκύπτουν προβλήματα υπολογιστικής αποδοτικότητας.

Επιπρόσθετα, άλλες εναλλακτικές μέθοδοι κωδικοποίησης περιλαμβάνουν την Label Encoding, όπου κάθε κατηγορία αντιστοιχίζεται σε έναν ακέραιο, και την Binary Encoding, η οποία μειώνει τον αριθμό των διαστάσεων κωδικοποιώντας τον ακέραιο δείκτη κάθε κατηγορίας σε δυαδική μορφή. Επιπλέον, η Target (Mean) Encoding αντικαθιστά κάθε κατηγορία με τον μέσο όρο του στόχου (target) των δειγμάτων της, ενσωματώνοντας άμεσα σήμα προς τον στόχο και διατηρώντας χαμηλό αριθμό διαστάσεων. Ωστόσο, απαιτείται smoothing ή cross-validation για να αποφευχθεί η υπερπροσαρμογή και η διαφυγή δεδομένων (data leakage). Τέλος, η κωδικοποίηση με βάση την συχνότητα απεικονίζει κάθε κατηγορία με τη συχνότητα εμφάνισής της στο dataset, προσφέροντας απλότητα και σταθερό εύρος χωρίς τη δημιουργία νέων στηλών, αλλά χάνει λεπτομέρειες σχετικότητας και μπορεί να εισάγει θόρυβο όταν οι συχνότητες είναι παρόμοιες. Η σωστή επιλογή μεταξύ αυτών των μεθόδων εξαρτάται από το μέγεθος, την κατανομή και τη φύση των κατηγοριών, καθώς και από τον τύπο του μοντέλου που θα εκπαιδευτεί.

Γενικότερα, η κωδικοποίηση των δεδομένων αποτελεί κρίσιμο βήμα για την αποτελεσματική λειτουργία των νευρωνικών δικτύων, καθώς εξασφαλίζει ότι τα δεδομένα βρίσκονται σε ομογενοποιημένη μορφή, ανεξαρτήτως του τι περιγράφουν, επιτρέποντας τη χρήση πολλών διαφορετικών τύπων δεδομένων.

Εντοπισμός και διαχείριση ακραίων τιμών (outliers)

Οι ακραίες τιμές, όπως απότομες μεταβολές της θερμοκρασίας ή σφάλματα καταγραφής, μπορούν να επηρεάσουν αρνητικά την ακρίβεια των μοντέλων πρόβλεψης, καθώς αποτελούν μεμονωμένες περιπτώσεις που, υπό κανονικές συνθήκες, δεν θα έπρεπε να επηρεάζουν τη σχέση εξάρτησης με τη μεταβλητή-στόχο. Πρακτικά, αυτές οι τιμές εισάγουν θόρυβο στο μοντέλο. Για την αντιμετώπιση τέτοιων καταστάσεων έχουν αναπτυχθεί διάφορες τεχνικές, μεταξύ των οποίων οι πιο διαδεδομένες είναι οι εξής:

- Z-score: Όπως αναλύθηκε προηγουμένως, η μέθοδος αυτή υπολογίζει την απόσταση μιας τιμής από τη μέση τιμή σε μονάδες τυπικής απόκλισης, επιτρέποντας τον εντοπισμό ακραίων τιμών.
- Isolation Forest: Αυτός ο αλγόριθμος χρησιμοποιεί τυχαία δέντρα απομόνωσης για να απομονώσει ανώμαλες τιμές, αξιοποιώντας την ιδιότητά τους να είναι ευκολότερα διαχωρίσιμες από τις φυσιολογικές τιμές.
- Στατιστική μέθοδος IQR (Interquartile Range): Η μέθοδος αυτή εφαρμόζεται σε μονοδιάστατα χαρακτηριστικά και ορίζει ένα φυσιολογικό εύρος τιμών με βάση το εύρος τεταρτημόριου. Συγκεκριμένα, το Q1 ορίζεται ως η τιμή κάτω από την οποία βρίσκεται το 25% των ταξινομημένων δεδομένων, ενώ το Q2 και το Q3 αντιστοιχούν στο 50% και 75%, αντίστοιχα. Το IQR υπολογίζεται ως $Q3 - Q1$, και οποιαδήποτε τιμή βρίσκεται εκτός του διαστήματος $[Q1 - 1.5 \times IQR, Q3 + 1.5 \times IQR]$ θεωρείται ανωμαλία.

Οι παραπάνω μέθοδοι επιτρέπουν τον αποτελεσματικό εντοπισμό και τη διαχείριση των ανωμαλιών, συμβάλλοντας στη βελτίωση της αξιοπιστίας και της ακρίβειας των μοντέλων πρόβλεψης.

Διαχείριση ελλειπουσών τιμών (missing values)

Για την αντιμετώπιση ελλειπουσών τιμών χρησιμοποιούνται τρεις βασικές στρατηγικές. Πρώτα κατά την απαλοιφή γραμμών ή στηλών αφαιρούνται εντελώς οι κενές τιμές όταν το ποσοστό τους είναι μικρό. Η μέθοδος είναι απλή και δεν αλλοιώνει τη στατιστική κατανομή των υπαρχόντων τιμών, αλλά μειώνει το σύνολο των διαθέσιμων δεδομένων. Συνεχίζοντας μια ακόμη μέθοδος είναι η μέθοδος υποκατάστασης με τη μέση τιμή όπου εδώ συμπληρώνεται κάθε κενό με την αντίστοιχη μέση τιμή της στήλης. Έτσι διατηρείται άθικτο το πλήθος των δειγμάτων, ωστόσο μειώνεται η φυσική μεταβλητότητα και τυχών συσχετίσεις μεταξύ των χαρακτηριστικών μπορεί να αλλοιωθούν. Τέλος, ο αλγόριθμος *k-Nearest Neighbours* εντοπίζει, για κάθε κενή τιμή, τα *k* πλησιέστερα δείγματα στο χώρο των υπολοίπων χαρακτηριστικών και υποκαθιστά το κενό με τον μέσο ή τη διάμεσο των αντίστοιχων τιμών. Έτσι διατηρείται η τοπική δομή των δεδομένων και οι στατιστικές συσχετίσεις, παρά το αυξημένο υπολογιστικό κόστος συγκριτικά με τις άλλες δύο μεθόδους, σε μεγάλα ή αραιά σύνολα.

Χρονική παραθυροποίηση (sliding window)

Η μέθοδος των κινούμενων παραθύρων είναι μια τεχνική δημιουργίας ακολουθιών από χρονοσειρές, ώστε να μπορούν να τροφοδοτηθούν σε μοντέλα μηχανικής μάθησης όπως τα LSTM. Αυτή η μέθοδος ομαδοποιεί το σύνολο των δεδομένων με βάση ένα εύρος και στη συνέχεια αυτές οι ομάδες των δεδομένων χρησιμοποιούνται ως εισόδοι στο μοντέλο. Έτσι δημιουργούνται ακολουθίες εισόδου – εξόδου και στο μοντέλο δεν εισάγεται ένα δείγμα τη φορά αλλά αντίθετα ένα σύνολο δειγμάτων επιτρέποντας του να κατανοεί χρονικές εξαρτήσεις [55]. Για παράδειγμα αν η ζήτηση ηλεκτρικού φορτίου εξαρτάται από τη θερμοκρασία και την υγρασία τις τελευταίες 12 ώρες, δημιουργείται ένα παράθυρο 12 χρονικών βημάτων για κάθε μεταβλητή. Το μοντέλο εκπαιδεύεται να προβλέπει τη ζήτηση με βάση αυτά τα 12 τελευταία δείγματα.

Επιλογή χαρακτηριστικών (Feature selection)

Η ποιότητα των χαρακτηριστικών που τροφοδοτούνται σε ένα μοντέλο είναι συχνά πιο καθοριστική από την ίδια την αρχιτεκτονική. Κατά το στάδιο της επιλογής των χαρακτηριστικών αποφασίζεται ποια από τα διαθέσιμα δεδομένα θα χρησιμοποιηθούν τελικά στο μοντέλο. Στόχος αυτής της διαδικασίας είναι η επιλογή του μικρότερου δυνατού πλήθους δεδομένων χωρίς όμως να χάνεται πληροφορία, με αποτέλεσμα την βέλτιστη ακρίβεια με ταυτόχρονη μείωση τους υπολογιστικού κόστους. Για την επιλογή αυτών των δεδομένων που είναι πλούσια σε πληροφορία και την απόρριψη άλλων στην πράξη χρησιμοποιείται μια κλιμακωτή διαδικασία που συνδυάζει διάφορες τεχνικές. Πρώτα μέσω ενός ελέγχου της συσχέτισης με τη μεταβλητή στόχο γίνεται ένα αρχικό φιλτράρισμα.

Τέτοιο παράδειγμα αποτελεί ο συντελεστής Pearson, ένας αδιάστατος δείκτης που ποσοτικοποιεί τη γραμμική σχέση μεταξύ δύο συνεχών τυχαίων μεταβλητών.

Επίσης, έχουν αναπτυχθεί αλγόριθμοι που αξιολογούν τα χαρακτηριστικά κατά την ίδια τη φάση μάθησης. Επειδή αυτές οι μέθοδοι είναι ενσωματωμένες στο μοντέλο χαρακτηρίζονται ως ενσωματωμένες. Τέτοιο παράδειγμα αποτελεί ο LASSO, όπου είναι μια παραλλαγή γραμμικής παλινδρόμησης που, πέρα από τη μείωση του σφάλματος πρόβλεψης, επιβάλλει ποινή στο άθροισμα των απόλυτων τιμών των συντελεστών (L1-penalty). Αυτή η ποινή «συρρικνώνει» όλους τους συντελεστές και ωθεί πολλούς να γίνουν ακριβώς μηδέν, ώστε το μοντέλο να κρατά μόνο τα χαρακτηριστικά με τη μεγαλύτερη πληροφοριακή αξία έτσι το LASSO λειτουργεί ταυτόχρονα ως μέθοδος κανονικοποίησης **και** αυτόματης επιλογής χαρακτηριστικών.

Τέλος χρησιμοποιούνται και άλλες πιο εξειδικευμένες επαναληπτικές μέθοδοι επιλογής χαρακτηριστικών όπως η Recursive Feature Elimination η οποία βασίζεται σε μια αναδρομική διαδικασία κατά την οποία εκπαιδεύονται απλά μοντέλα (όπως SVM και παλινδρόμηση) στο σύνολο των δεδομένων και έπειτα κάθε γνώρισμα βαθμολογείται βάσει της συμβολής του. Τελικά διατηρούνται μόνο εκείνα που συμβάλλουν περισσότερο στην απόδοση του μοντέλου. Αυτές οι μέθοδοι αν και πιο ακριβείς είναι αρκετά ακριβές υπολογιστικά και για αυτό εφαρμόζονται μόνο αφού το σύνολο των δεδομένων είναι ήδη περιορισμένο.

Παράλληλα με τις τεχνικές μεθόδους, η εμπειρική γνώση είναι σημαντική για την επιλογή των χαρακτηριστικών. Για παράδειγμα είναι γνωστό εκ των προτέρων ότι κατά την πρόβλεψη φορτίου ενέργειας η θερμοκρασία είναι ένας από τους σημαντικότερους εξωγενείς παράγοντες.

3.6 Αρχιτεκτονικές δικτύων για πρόβλεψη φορτίου επόμενης ημέρας

Η πρόβλεψη φορτίου επόμενης ημέρας αποτελεί κρίσιμο πυλώνα για τη διαχείριση των ενεργειακών συστημάτων, επιτρέποντας στους παρόχους ενέργειας να προβλέπουν με ακρίβεια τη ζήτηση ηλεκτρικής ενέργειας, να βελτιστοποιούν τη διανομή και να μειώνουν το λειτουργικό κόστος. Στο σύγχρονο πεδίο της τεχνητής νοημοσύνης, οι αρχιτεκτονικές νευρωνικών δικτύων έχουν αναδειχθεί σε ισχυρά εργαλεία, χάρη στην ικανότητά τους να μοντελοποιούν πολύπλοκες χρονικές εξαρτήσεις και μη γραμμικές σχέσεις στα δεδομένα φορτίου. Η συνεχής εξέλιξη των τεχνικών βαθιάς μάθησης έχει οδηγήσει σε σημαντικές βελτιώσεις στην ακρίβεια των προβλέψεων. Το παρόν κείμενο παρέχει μια ολοκληρωμένη επισκόπηση των state-of-the-art αρχιτεκτονικών νευρωνικών δικτύων που χρησιμοποιούνται για την πρόβλεψη φορτίου επόμενης ημέρας, εξετάζοντας την απόδοσή τους, τις προκλήσεις και τους περιορισμούς που προκύπτουν από την εφαρμογή τους. Αναλυτικά, οι αρχιτεκτονικές που εφαρμόζονται περιλαμβάνουν μια ποικιλία προσεγγίσεων, καθεμία με μοναδικά χαρακτηριστικά, όπως τα δίκτυα Long Short-Term Memory (LSTM), τα Gated

Recurrent Unit (GRU), τα Temporal Convolutional Networks (TCN), τα υβριδικά μοντέλα, τα Graph Neural Networks (GNNs), τα Temporal Fusion Transformers (TFT) και τα μη γραμμικά αυτοπαλινδρομικά νευρωνικά δίκτυα με εξωγενείς εισροές (NARX). Κάθε αρχιτεκτονική έχει σχεδιαστεί για να αντιμετωπίζει διαφορετικές πτυχές της πρόβλεψης χρονοσειρών, προσφέροντας προσαρμοσμένες λύσεις για τις ανάγκες των ενεργειακών συστημάτων.

Τα δίκτυα LSTM είναι ειδικά σχεδιασμένα για τη διαχείριση μακροχρόνιων εξαρτήσεων σε ακολουθιακά δεδομένα, καθιστώντας τα ιδανικά για δεδομένα φορτίου με κυκλικές ή εποχιακές τάσεις. Μια μελέτη του 2024 κατέδειξε ότι τα LSTM επιτυγχάνουν εξαιρετικά χαμηλά σφάλματα σε εβδομαδιαίες και μηνιαίες προβλέψεις, με RMSE 0,13% και MAE 0,74 για το σύνολο δεδομένων Chandigarh UT, υπερτερώντας παραδοσιακών μεθόδων, όπως τα Recurrent Neural Networks (RNN) και τα Support Vector Machines (SVM) [58]. Παράλληλα, τα δίκτυα GRU αποτελούν μια απλοποιημένη παραλλαγή των LSTM, προσφέροντας παρόμοια απόδοση με μειωμένη υπολογιστική πολυπλοκότητα. Μια δημοσίευση του 2024 παρουσίασε ένα υβριδικό μοντέλο GRU-TCN, το οποίο πέτυχε ακρίβεια άνω του 95% σε σύνολα δεδομένων, όπως τα GEFCom2014 και ERCOT, με RMSE μικρότερο από 4 και MAPE μικρότερο από 6% [59]. Η αποδοτικότητα των GRU τα καθιστά ιδιαίτερα κατάλληλα για υβριδικές διαμορφώσεις. Επιπρόσθετα, τα TCN χρησιμοποιούν διασταλμένες συνελίξεις για την αιχμαλώτιση χρονικών προτύπων, προσφέροντας πλεονεκτήματα, όπως η παράλληλη επεξεργασία. Μια μελέτη του 2022 ανέδειξε ότι τα TCN μπορούν να υπερτερούν των LSTM σε περιπτώσεις με πολύπλοκα χρονικά πρότυπα, επιτυγχάνοντας MSE 0,0648 και MAE 14,7869 MW για το σύνολο δεδομένων Montenegro Load [60].

Τα υβριδικά μοντέλα συνδυάζουν στοιχεία από διαφορετικές αρχιτεκτονικές, εκμεταλλευόμενα τα πλεονεκτήματα κάθε μιας. Για παράδειγμα, τα CNN-LSTM χρησιμοποιούν συνελκτικά δίκτυα για την εξαγωγή χαρακτηριστικών και LSTM για την πρόβλεψη ακολουθιών, ενώ ο μηχανισμός προσοχής ενισχύει την εστίαση σε σημαντικά χρονικά βήματα. Μια ανασκόπηση του 2021 υπογράμμισε την ανωτερότητα των υβριδικών μοντέλων έναντι μονοσήμαντων δικτύων [61]. Το υβριδικό μοντέλο GRU-TCN του 2024 επιβεβαίωσε την εξαιρετική απόδοση, συνδυάζοντας την αποδοτικότητα των GRU με την ικανότητα των TCN να χειρίζονται χρονικά πρότυπα [59]. Επιπλέον, τα GNNs αποτελούν μια αναδυόμενη κατηγορία, ιδανική για τη μοντελοποίηση χωρικών και χρονικών εξαρτήσεων. Μια δημοσίευση του 2024 πρότεινε τα Visibility Graph Neural Networks (VGNN), τα οποία βελτίωσαν την ακρίβεια πρόβλεψης κατά 2-19% στο MAE για δεδομένα από υποσταθμούς υψηλής και μεσαίας τάσης [62]. Τα GNNs είναι ιδιαίτερα υποσχόμενα για περιφερειακές προβλέψεις, όπου οι χωρικές σχέσεις είναι κρίσιμες.

Οι Temporal Fusion Transformers (TFT), που παρουσιάστηκαν το 2021, αποτελούν μια αρχιτεκτονική που συνδυάζει τα πλεονεκτήματα των LSTM, των μηχανισμών προσοχής και της επιλογής χαρακτηριστικών για την πρόβλεψη χρονοσειρών πολλαπλών οριζόντων [63]. Η δομή τους περιλαμβάνει στρώματα πυλών για την καταστολή περιττών πληροφοριών, επαναλαμβανόμενα στρώματα για τοπική

επεξεργασία και αυτο-προσοχή για τη μοντελοποίηση μακροχρόνιων εξαρτήσεων. Μια μελέτη του 2022 εξέτασε την εφαρμογή των TFT στην πρόβλεψη φορτίου του Hanoi, επιτυγχάνοντας σφάλμα πρόβλεψης 3,38%, σημαντικά χαμηλότερο από το 4,64% των LSTM [64]. Οι TFT υπερέχουν σε εφαρμογές με πολλαπλές εξωγενείς μεταβλητές, όπως καιρικά δεδομένα και ημερολογιακές πληροφορίες, χάρη στην ικανότητά τους να ζυγίζουν τη σημασία κάθε χαρακτηριστικού. Πρόσφατες εξελίξεις περιλαμβάνουν το Improved Temporal Fusion Transformer (ITFT), το οποίο αντικαθιστά τα LSTM με GRU για αποδοτικότερη εκμάθηση. Μια μελέτη του 2023 παρουσίασε το ITFT για προβλέψεις μεσαίου ορίζοντα, επιτυγχάνοντας καλύτερη απόδοση σε σχέση με το αρχικό μοντέλο TFT [65].

Το μοντέλο NARX, που παρουσιάστηκε το 2017, σχεδιάστηκε για τη βραχυπρόθεσμη πρόβλεψη φορτίου. Αποτελείται από ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο με ένα κρυφό επίπεδο, χρησιμοποιώντας σιγμοειδείς συναρτήσεις ενεργοποίησης στους κρυφούς κόμβους και γραμμική συνάρτηση στους κόμβους εξόδου. Το δίκτυο εκπαιδεύεται σε ανοιχτό βρόχο με δεδομένα φορτίου και καιρού, ενώ στη φάση πρόβλεψης λειτουργεί σε κλειστό βρόχο, χρησιμοποιώντας την προβλεπόμενη έξοδο ως είσοδο ανάδρασης. Οι εξωγενείς μεταβλητές περιλαμβάνουν χρονικές παραμέτρους (μήνας, ημέρα, ώρα, εργάσιμη ημέρα) και καιρικές μεταβλητές (θερμοκρασία, σημείο δρόσου), επιλεγμένες μέσω ανάλυσης συσχέτισης. Μια μελέτη ανέφερε MAPE 0,85% για το δίκτυο της Νέας Αγγλίας, υπερτερώντας του ARMAX (MAPE 1,09%) [66]. Η επαναλαμβανόμενη δομή του NARX μειώνει την εξάρτηση από εξωτερικά δεδομένα, ενισχύοντας την αυτονομία του.

Παρά την πρόοδο, υπάρχουν σημαντικές προκλήσεις. Η απόδοση των αρχιτεκτονικών ποικίλλει ανάλογα με το σύνολο δεδομένων, με τα TCN να υπερέχουν σε ορισμένες περιπτώσεις και τα LSTM σε άλλες, λόγω διαφορών στην εποχικότητα ή τον θόρυβο [60, 58]. Τα υβριδικά μοντέλα και τα GNNs παρουσιάζουν υψηλή υπολογιστική πολυπλοκότητα, δυσχεραίνοντας την εφαρμογή τους σε συστήματα με περιορισμένους πόρους. Επιπλέον, η ανάγκη για μεγάλα σύνολα δεδομένων μπορεί να αποτελεί εμπόδιο, ιδιαίτερα σε περιοχές με περιορισμένη πρόσβαση σε δεδομένα. Η γενίκευση των μοντέλων σε διαφορετικά γεωγραφικά ή κλιματικά περιβάλλοντα παραμένει πρόκληση, καθώς τα δεδομένα φορτίου διαφέρουν σημαντικά [67]. Εν κατακλείδι, η πρόβλεψη φορτίου επόμενης ημέρας βασίζεται σε προηγμένες αρχιτεκτονικές νευρωνικών δικτύων, με τα LSTM, GRU, TCN, υβριδικά μοντέλα, GNNs, TFT και NARX να κατέχουν πρωταγωνιστικό ρόλο. Η επιλογή της κατάλληλης αρχιτεκτονικής εξαρτάται από το πλαίσιο και τα δεδομένα, με τις μελλοντικές εξελίξεις να αναμένεται να εστιάσουν στην περαιτέρω ανάπτυξη των GNNs και τη βελτιστοποίηση των υβριδικών μοντέλων, ενισχύοντας τη βιωσιμότητα των ενεργειακών συστημάτων.

3.7 Χρήση εξωγενών μεταβλητών για πρόβλεψη φορτίου επόμενης ημέρας

Η πρόβλεψη φορτίου ηλεκτρικής ενέργειας αποτελεί θεμελιώδες στοιχείο για την αποτελεσματική διαχείριση έξυπνων δικτύων, την υποστήριξη στρατηγικών αποφάσεων των παραγωγών ενέργειας και την ενίσχυση της ενεργού συμμετοχής ιδιωτών σε απορυθμισμένες αγορές ενέργειας [68]. Η ενσωμάτωση εξωγενών μεταβλητών, δηλαδή δεδομένων που υπερβαίνουν τα ιστορικά δεδομένα κατανάλωσης, ενισχύει σημαντικά την ακρίβεια των προβλέψεων, αποτυπώνοντας τη μεταβλητότητα της ζήτησης που προκύπτει από εξωτερικούς παράγοντες [69, 70]. Βασισμένη στις ανασκοπήσεις των Christen et al. (2020) και Vanting et al. (2021), η παρούσα ανάλυση διερευνά τη σημασία, την ποικιλία, την πληροφοριακή αξία και τις προκλήσεις που συνδέονται με τη χρήση εξωγενών μεταβλητών, εστιάζοντας τόσο σε συνολικά όσο και σε ατομικά επίπεδα, ενώ εξετάζει τη συμβολή τους σε διαφορετικά πλαίσια, όπως τα νοικοκυριά, τα κτίρια και τα έξυπνα δίκτυα.

Οι εξωγενείς μεταβλητές παρέχουν κρίσιμες πληροφορίες για συνθήκες υψηλότερου επιπέδου που επηρεάζουν άμεσα την κατανάλωση και την παραγωγή ενέργειας. Συγκεκριμένα, η ζήτηση εξαρτάται από τη λειτουργία ηλεκτρικών συσκευών, η οποία καθορίζεται από εξωτερικούς παράγοντες, όπως οι καιρικές συνθήκες ή οι συνήθειες των χρηστών [71]. Παράλληλα, η παραγωγή από ανανεώσιμες πηγές, όπως τα φωτοβολταϊκά συστήματα, επηρεάζεται από παράγοντες όπως η ηλιακή ακτινοβολία ή η ταχύτητα του ανέμου [72]. Έρευνες, όπως αυτές των Heinemann et al. (1966), López et al. (2017), Selvi & Mishra (2018, 2020), Torabi & Hashemi (2012), και Eseye et al. (2019), έχουν καταδείξει ότι η συμπερίληψη εξωγενών δεδομένων, ιδιαίτερα των καιρικών, μειώνει την αβεβαιότητα των προβλέψεων σε σύγκριση με την αποκλειστική αξιοποίηση ιστορικών δεδομένων [73–78]. Αυτή η βελτίωση είναι ιδιαίτερα κρίσιμη σε απορυθμισμένες αγορές, όπου οι καταναλωτές απαιτούν ακριβείς εκτιμήσεις για να εξισορροπήσουν την παραγωγή και την κατανάλωσή τους, ελαχιστοποιώντας το κόστος και μεγιστοποιώντας την αυτοδυναμία [79].

Επιπρόσθετα, οι εξωγενείς μεταβλητές προσφέρουν πρόσθετο πλαίσιο που διευκολύνει την κατανόηση των διακυμάνσεων στη ζήτηση ενέργειας, ιδιαίτερα όταν εφαρμόζονται βαθιά νευρωνικά δίκτυα [80]. Αυτά τα μοντέλα επωφελούνται από την ικανότητά τους να επεξεργάζονται πολλαπλές εισόδους, αξιοποιώντας εξωγενείς μεταβλητές για να αποτυπώσουν καλύτερα την εποχικότητα, τις διακυμάνσεις και τα πρότυπα κατανάλωσης [81]. Σε περιβάλλοντα υψηλής μεταβλητότητας, όπως τα νοικοκυριά ή τα έξυπνα δίκτυα, οι μεταβλητές αυτές παρέχουν πληροφορίες που δεν μπορούν να αντληθούν αποκλειστικά από ιστορικά δεδομένα, διευκολύνοντας τη διαχείριση της ζήτησης και τον έλεγχο του δικτύου [82].

Η κατηγοριοποίηση των εξωγενών μεταβλητών αποκαλύπτει την ποικιλία και τη σημασία τους. Περισσότερες από 50 διαφορετικές μεταβλητές έχουν εντοπιστεί, με τα καιρικά δεδομένα να χρησιμοποιούνται στο 63% των προσεγγίσεων,

κυριαρχώντας λόγω της ισχυρής συσχέτισής τους με την κατανάλωση, όπως στη χρήση κλιματιστικών ή συστημάτων HVAC [83]. Τα ημερολογιακά δεδομένα (55%) και οι πληροφορίες για τον τύπο της ημέρας (53%), όπως οι αργίες, οι εργάσιμες ημέρες ή η εποχή, επηρεάζουν καθοριστικά τη συμπεριφορά κατανάλωσης, διευκολύνοντας την αναγνώριση επαναλαμβανόμενων προτύπων, όπως η αυξημένη ζήτηση σε ώρες αιχμής [84]. Τα κοινωνικοοικονομικά (8%) και δημογραφικά δεδομένα (5%) βρίσκουν εφαρμογή κυρίως σε μακροπρόθεσμες προβλέψεις, αντικατοπτρίζοντας τάσεις όπως η αύξηση του πληθυσμού ή η οικονομική ανάπτυξη [85, 86]. Δεδομένα που αφορούν τα χαρακτηριστικά κτιρίων και νοικοκυριών, όπως η συμπεριφορά των ενοίκων ή η μόνωση, αποκτούν ιδιαίτερη σημασία σε ατομικό επίπεδο [87, 88]. Λιγότερο διαδεδομένες πηγές, όπως δεδομένα από αισθητήρες, κοινωνικά μέσα ή συσκευές IoT, προσφέρουν δυνητικά πολύτιμες πληροφορίες, ιδιαίτερα για έξυπνα δίκτυα, αν και η πολυπλοκότητα της ανάλυσής τους περιορίζει τη χρήση τους [89, 90].

Η πληροφοριακή αξία των εξωγενών μεταβλητών εξαρτάται από παράγοντες όπως το χρονικό παράθυρο, το επίπεδο συσσώρευσης και η γεωγραφική θέση. Στις βραχυπρόθεσμες προβλέψεις, μεταβλητές όπως η θερμοκρασία ή η νεφοκάλυψη είναι καθοριστικές, ενώ στις μακροπρόθεσμες κυριαρχούν πιο αργά εξελισσόμενες μεταβλητές, όπως οι δημογραφικές ή οικονομικές [91]. Σε επίπεδο κοινής ωφέλειας, όπου το 82% των μελετών εστιάζει, η επίδραση μεμονωμένων συσκευών εξουδετερώνεται λόγω του φαινομένου του μέσου όρου, καθιστώντας τα καιρικά δεδομένα πιο σημαντικά [92]. Αντιθέτως, σε ατομικό επίπεδο, που αντιστοιχεί στο 13% των μελετών, δεδομένα όπως η κίνηση σε κτίρια ή η παρουσία ενοίκων αποκτούν μεγαλύτερη βαρύτητα λόγω της υψηλής μεταβλητότητας της κατανάλωσης [93]. Στα νοικοκυριά, η πρόβλεψη φορτίου παρουσιάζει προκλήσεις λόγω αυτής της μεταβλητότητας, με εξωγενείς μεταβλητές να συμβάλλουν στην κατανόησή της [94]. Η αθροιστική ανάλυση δεδομένων ή η ομαδοποίηση νοικοκυριών με παρόμοια προφίλ κατανάλωσης προτείνεται για τη μείωση αυτής της μεταβλητότητας [95, 96]. Στα κτίρια, ιδιαίτερα σε δημόσια ή γραφειακά, μεταβλητές όπως τα δεδομένα HVAC, ο φωτισμός και η παρουσία ενοίκων είναι απαραίτητες για τη διαχείριση της ζήτησης, υποστηρίζοντας στρατηγικές όπως η μετατόπιση φορτίων [97, 98]. Στα έξυπνα δίκτυα, οι εξωγενείς μεταβλητές διευκολύνουν τους διαχειριστές συστημάτων διανομής να προγραμματίσουν την παροχή ενέργειας και να εντοπίσουν περιοχές πιθανής αστάθειας, ενισχύοντας μέτρα απόκρισης ζήτησης [99, 100].

Η αποτελεσματική αξιοποίηση εξωγενών μεταβλητών προϋποθέτει προσεκτική επιλογή χαρακτηριστικών, καθώς η υπερβολική ή περιττή πληροφορία μπορεί να υποβαθμίσει την ακρίβεια λόγω θορύβου [101]. Μέθοδοι όπως τα Random Forests χρησιμοποιούνται για την επιλογή χαρακτηριστικών υψηλής συνάφειας [102, 103]. Η ποιότητα των δεδομένων είναι εξίσου κρίσιμη, με την ανάγκη φιλτραρίσματος ανωμαλιών να αποτρέπει παραμορφώσεις [104, 105]. Η αβεβαιότητα σε προβλέψεις εξωγενών μεταβλητών, όπως ο καιρός, μπορεί να εισαγάγει σφάλματα, τα οποία μετριάζονται μέσω τεχνικών όπως οι συνδυαστικές προβλέψεις ή η αποσύνθεση δεδομένων [106]. Παράλληλα, η χρήση εξωγενών

μεταβλητών συνοδεύεται από προκλήσεις, όπως η περιορισμένη διαθεσιμότητα ή η χαμηλή ποιότητα δεδομένων. Η ακρίβεια των προβλέψεων καιρού ή η δυσκολία συλλογής δεδομένων, όπως η παρουσία ενοίκων, μπορεί να περιορίσει την απόδοση των μοντέλων. Η προσθήκη πολλαπλών μεταβλητών αυξάνει την πολυπλοκότητα των μοντέλων, οδηγώντας σε μεγαλύτερους χρόνους εκπαίδευσης και υπολογιστική επιβάρυνση. Επιπλέον, η υψηλή μεταβλητότητα των δεδομένων, ιδιαίτερα στα νοικοκυριά, απαιτεί προσεκτική επιλογή μεταβλητών για την αποφυγή υπερπροσαρμογής. Τέλος, η γενίκευση των μοντέλων σε διαφορετικά γεωγραφικά ή κοινωνικοοικονομικά περιβάλλοντα παραμένει δύσκολη λόγω των διαφορών στα πρότυπα κατανάλωσης, ενώ η ικανότητα σύγκρισης των αποτελεσμάτων περιορίζεται από την εξάρτηση από το πλαίσιο και την πηγή δεδομένων.

Εν κατακλείδι, η ενσωμάτωση εξωγενών μεταβλητών στην πρόβλεψη φορτίου ηλεκτρικής ενέργειας αποτελεί δυναμικό πεδίο έρευνας με σημαντικές προοπτικές για την ενίσχυση της ενεργειακής αποδοτικότητας και βιωσιμότητας. Παρά την πρόοδο, η πληροφοριακή αξία αυτών των δεδομένων, ιδιαίτερα σε ατομικό επίπεδο, παραμένει ανεπαρκώς διερευνημένη. Νέες πηγές, όπως τα δεδομένα από κοινωνικά μέσα ή συσκευές IoT, προσφέρουν δυνατότητες για πιο λεπτομερείς προβλέψεις, αλλά απαιτούν περαιτέρω έρευνα για την κατανόηση της συμβολής τους. Η ανάπτυξη μετρικών για την ποσοτικοποίηση της αξίας των δεδομένων και η βελτιστοποίηση της επιλογής χαρακτηριστικών μέσω προηγμένων τεχνικών αποτελούν κρίσιμες κατευθύνσεις για τη βελτίωση της ακρίβειας. Με τη συνεχή εξέλιξη των τεχνολογιών βαθιάς μάθησης και τη διεύρυνση των διαθέσιμων δεδομένων, η πρόβλεψη φορτίου αναμένεται να καταστεί ακόμη πιο αξιόπιστη, υποστηρίζοντας τη μετάβαση προς αποδοτικότερα και βιώσιμα ενεργειακά συστήματα.

ΚΕΦΑΛΑΙΟ 4 – ΠΕΙΡΑΜΑΤΙΚΗ ΔΙΑΔΙΚΑΣΙΑ

4.1 Στόχος

Όπως έχει ήδη αναλυθεί, η σημασία της πρόβλεψης φορτίου επόμενης ημέρας είναι κρίσιμη για τη βιωσιμότητα του δικτύου της ενέργειας παγκοσμίως ειδικά με την ένταξη των ανανεώσιμων πηγών ενέργειας, αλλά και για τη διατήρηση των τιμών της ενέργειας σε προσιτά επίπεδα.

Τα τελευταία χρόνια έχει σημειωθεί σημαντική πρόοδος στον τομέα της πρόβλεψης φορτίου επόμενης ημέρας, με την χρήση νευρωνικών δικτύων και την σταδιακή αποφυγή της απόλυτης υιοθέτησης καθαρά στατιστικών μεθόδων. Οι μελέτες επικεντρώνονται στην ανάπτυξη διαφορετικών αρχιτεκτονικών νευρωνικών δικτύων αλλά και στην κατάλληλη ενσωμάτωση των δεδομένων σε αυτές τις αρχιτεκτονικές, με στόχο αφενός την βελτιστοποίηση της πρόβλεψης και αφετέρου την μείωση του υπολογιστικού κόστους. Άλλωστε τη σύγχρονη εποχή της ραγδαίας ανάπτυξης των νευρωνικών δικτύων και της αξιοποίησης μεγάλων δεδομένων, το υπολογιστικό κόστος έχει καταστεί ένας από τους πλέον καθοριστικούς παράγοντες για την επιτυχία των εφαρμογών τεχνητής νοημοσύνης. Καθώς τα μοντέλα γίνονται ολοένα και πιο πολύπλοκα, με δισεκατομμύρια παραμέτρους και τεράστιες απαιτήσεις επεξεργασίας, η ανάγκη για αποτελεσματική διαχείριση των υπολογιστικών πόρων καθίσταται επιτακτική. Για τον λόγο αυτό, η επιλογή αρχιτεκτονικών νευρωνικών δικτύων, η βελτιστοποίηση των υπάρχοντων μοντέλων και η προσεκτική αξιοποίηση εξωγενών μεταβλητών πρέπει να γίνονται με κύριο γνώμονα την αποδοτικότητα.

Ένας από τους βασικούς άξονες έρευνας στον τομέα της πρόβλεψης φορτίου επόμενης ημέρας αφορά την ενσωμάτωση εξωγενών μεταβλητών, δηλαδή δεδομένων που δεν προέρχονται άμεσα από το ιστορικό των τιμών φορτίου, όμως επηρεάζουν σημαντικά τη ζήτηση ηλεκτρικής ενέργειας. Οι εξωγενείς μεταβλητές αντιπροσωπεύουν μια πλούσια πηγή πληροφορίας που μπορεί να βελτιώσει την ακρίβεια των προβλέψεων, καθώς επιτρέπουν στο μοντέλο να λαμβάνει υπόψη παράγοντες που σχετίζονται με την πραγματική κατανάλωση ενέργειας. Η επιλογή του βέλτιστου τρόπου ενσωμάτωσης των εξωγενών μεταβλητών αποτελεί ακόμη ανοιχτό πεδίο έρευνας ενώ, η εκτίμηση της συμβολής τους στην ακρίβεια της πρόβλεψης αν και σημαντική παραμένει ενεργό ερευνητικό ερώτημα.

Στόχος της παρούσας διπλωματικής εργασίας είναι να ερευνήσει τον πιο αποδοτικό τρόπο ενσωμάτωσης εξωγενών μεταβλητών σε αρχιτεκτονική πλήρους συνδεδεμένου νευρωνικού δικτύου (FCNN) που αναπτύχθηκε και δοκιμάστηκε σε δεδομένα από εικοσιτέσσερις χώρες της Ευρωπαϊκής Ένωσης. Αρχικά αναζητείται με πειράματα η βέλτιστη δομή των εξωγενών μεταβλητών και έπειτα μελετάται ο τρόπος ενσωμάτωσης τους στο νευρωνικό μοντέλο με σκοπό την βελτίωση της ακρίβειας των προβλέψεων για το φορτίο ενέργειας επόμενης

ημέρας και για τις εικοσιτέσσερις αυτές χώρες. Ακόμη δευτερογενή στόχο αποτέλεσε ο γενικότερος πειραματισμός με την προεπεξεργασία και τις αρχιτεκτονικές ενσωμάτωσης των εξωγενών μεταβλητών προκειμένου να γίνει μια γενική και έγκυρη καταγραφή της συμπεριφοράς τους.

4.2 Περιγραφή της αρχιτεκτονικής του νευρωνικού δικτύου

Αρχικά πρέπει να αναφερθεί πως η συγκεκριμένη αρχιτεκτονική αναπτύχθηκε στα πλαίσια της ερευνητικής εργασίας «Short-term load forecasting with global models: A comparative analysis of neural network architectures» με συγγραφείς τους «Αρτέμιο-Ανάργυρο Σεμένογλου, Ευάγγελο Σπηλιώτη και Βασίλειο Ασημακόπουλο».

Σε αυτήν την εργασία, η αρχιτεκτονική που επιλέχθηκε ως βέλτιστη για την εφαρμογή της πρόβλεψης φορτίου επόμενης ημέρας ήταν μια Fully-Connected Neural Network - Deep (FCNN-Deep), λόγω της υψηλής ακρίβειας που πέτυχε στις προβλέψεις, παρά την φαινομενικά απλότητα της και συνάμα το μικρό υπολογιστικό κόστος σε σχέση με πιο προηγμένες αρχιτεκτονικές. Η δομή της αρχιτεκτονικής, χωρίς ενσωμάτωση εξωγενών μεταβλητών, είναι η εξής :

Επίπεδο Εισόδου (Input Layer)

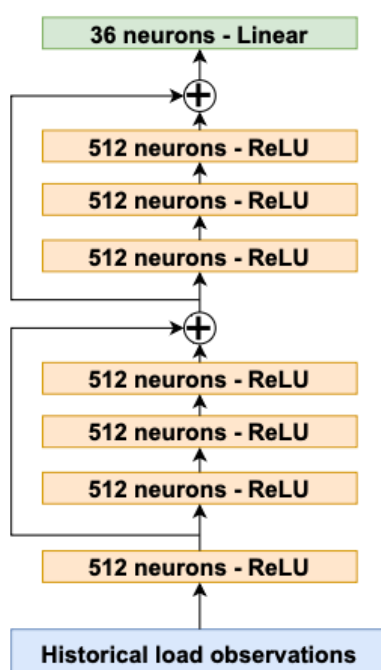
Σε αυτό το επίπεδο τροφοδοτούνται τα δεδομένα ιστορικού φορτίου για μια καθορισμένη περίοδο των προηγούμενων επτά ημερών ή αλλιώς των προηγούμενων 168 ωρών. Και για αυτό ο αριθμός των εισόδων είναι 168. Παρατηρούμε ότι γίνεται χρήση της τεχνικής του κινούμενου παραθύρου (sliding window), όπου τα δεδομένα εισάγονται σε παρτίδες των 168 δειγμάτων. Αυτός ο σχεδιασμός αποσκοπεί στο μοντέλο να λαμβάνει υπόψη του τις τάσεις και τα μοτίβα της προηγούμενης εβδομάδας για να προβλέψει το φορτίο της επόμενης ημέρας.

Πολλαπλά Κρυφά Επίπεδα (Hidden Layers)

Τα κρυφά επίπεδα είναι επτά πλήρως συνδεδεμένα επίπεδα, το καθένα από τα οποία έχει 512 νευρώνες και αυτοί οι νευρώνες κάθε επιπέδου είναι συνδεδεμένοι με όλους τους νευρώνες του επόμενου επιπέδου. Ως συνάρτηση ενεργοποίησης για κάθε επίπεδο χρησιμοποιείται η ReLU. Ακόμη χρησιμοποιούνται παρακάμψεις (skip connections). Πρόκειται για συνδέσεις σε βαθιά νευρωνικά δίκτυα που επιτρέπουν στην έξοδο ενός επιπέδου να παρακάμπτει ένα ή περισσότερα ενδιάμεσα επίπεδα και να συνδέεται απευθείας σε επόμενο επίπεδο. Με αυτό το τρόπο σημαντικές πληροφορίες από τα πρώτα επίπεδα διατηρούνται χωρίς να παραμορφώνονται πλήρως από τα ενδιάμεσα επίπεδα και έτσι η αρχική πληροφορία μπορεί να προστεθεί απευθείας στην έξοδο του δικτύου, βελτιώνοντας την ικανότητα μάθησης.

Επίπεδο Εξόδου (Output Layer)

Αυτό το επίπεδο αποτελείται από 36 νευρώνες, όπου κάθε νευρώνας εκφράζει την πρόβλεψη φορτίου για μια αντίστοιχη ώρα. Δηλαδή η πρόβλεψη γίνεται σε ορίζοντα 36 ωρών, όπου οι πρώτες 12 ώρες αφορούν την ίδια ημέρα (12:00 – 24:00) και οι υπόλοιπες 24 ώρες αφορούν το σύνολο της επόμενης ημέρας. Αυτό συμβαίνει καθώς σε πραγματικές εφαρμογές πρόβλεψης φορτίου επόμενης ημέρας, πραγματοποιείται πρόβλεψη και για τις δώδεκα ώρες της ίδιας ημέρας. Τέλος ως συνάρτηση ενεργοποίησης για το επίπεδο εξόδου χρησιμοποιείται η γραμμική συνάρτηση η οποία είναι και η συνήθης σε εφαρμογές πρόβλεψης.



Σχήμα 4.1 : FCNN δίκτυο

Τέλος αναφορικά με τις υπερπαραμέτρους και την μέθοδο εκπαίδευσης, αυτή πραγματοποιήθηκε χρησιμοποιώντας τον αλγόριθμο βελτιστοποίησης για την εκπαίδευση *Adam* με αρχικό ρυθμό εκμάθησης *0,001* και μέγεθος παρτίδας *512*. Ως συνάρτηση απώλειας επιλέχθηκε η Mean Absolute Error (MAE), η οποία είναι κατάλληλη για προβλήματα παλινδρόμησης. Επίσης για την αύξηση της ακρίβειας των προβλέψεων χρησιμοποιήθηκε μια προσέγγιση *ensemble learning*, όπου εκπαιδεύτηκαν 30 ίδια μοντέλα, το καθένα με διαφορετικές αρχικές συνθήκες. Οι τελικές προβλέψεις προέκυψαν από τον μέσο όρο των προβλέψεων των 30 μοντέλων, κάτι που ενίσχυσε την αξιοπιστία των αποτελεσμάτων.

4.3 Περιγραφή προτεινόμενης μεθοδολογίας

4.3.1 Αναπαραγωγή νευρωνικού δικτύου και αρχικών δεδομένων

Αρχικό βήμα αποτελεί η αναπαραγωγή της πρώτης εργασίας με την εκτύπωση του σχήματος των δεδομένων, δηλαδή του αριθμού στηλών και γραμμών καθενός συνόλου καθώς επίσης και η κατασκευή του πλήρως διασυνδεδεμένου νευρωνικού δικτύου. Στην συνέχεια λαμβάνει χώρα η εκπαίδευση του πάνω στα ιστορικά δεδομένα φορτίου, χρησιμοποιώντας ίδια σύνολα εκπαίδευσης και αξιολόγησης αλλά και ίδιες τιμές υπερπαραμέτρων. Στόχος είναι να υπάρξει σημείο αναφοράς ως προς τις τιμές του σφάλματος προκειμένου στη πορεία να είναι μετρήσιμη τυχόν βελτίωση από τα πειράματα. Ακόμη τα αποτελέσματα αυτά συγκρίνονται ως προς την εγγύτητα τους με τα αποτελέσματα της αρχικής εργασίας προκειμένου να διασφαλιστεί η ορθότητα των πειραμάτων δεδομένο ότι οι αρχικές συνθήκες είναι περίπου ίδιες.

Με την εξασφάλιση παρόμοιων συνθηκών με τα αποτελέσματα της αρχικής εργασίας τα πειράματα μπορούν να ξεκινήσουν. Αυτά διαχωρίστηκαν σε τρεις διακριτές κατηγορίες. Στην εύρεση της βέλτιστης μορφής κάθε εξωγενούς μεταβλητής, στην εύρεση του πιο αποδοτικού συνδυασμού τους και τέλος στην εύρεση της καλύτερης αρχιτεκτονικής τροφοδοσίας τους στο νευρωνικό μοντέλο.

4.3.2 Μορφές αναπαράστασης εξωγενών μεταβλητών

Για την εύρεση της πιο αποδοτικής μορφής κάθε εξωγενούς μεταβλητής, αφού σχηματίστηκαν οι πιθανές μορφές για κάθε μία μεταβλητή αυτές τροφοδοτήθηκαν μια προς μια στο νευρωνικό δίκτυο, συνδυαστικά με τα ιστορικά δεδομένα φορτίου και μετρήθηκε πως μεταβλήθηκε η απόδοση των προβλέψεων.

Η αρχιτεκτονική με την οποία έγινε η ενσωμάτωση των εξωγενών μεταβλητών στο σύστημα χαρακτηρίζεται από απλότητα, καθώς σε αυτό το στάδιο κύριος στόχος είναι η στείρα σύγκριση των διαφορετικών μορφών των εξωγενών μεταξύ τους και όχι η εν γένει βελτίωση της ακρίβειας των προβλέψεων του συστήματος. Έτσι οι εξωγενείς μεταβλητές τροφοδοτούνται ως είσοδοι σε ένα μόνο πυκνό στρώμα με συνάρτηση ενεργοποίησης την γραμμική, σε παράλληλη ροή με το νευρωνικό δίκτυο των ιστορικών δεδομένων φορτίου. Τελικά οι δύο έξοδοι, του κύριου δικτύου και της δευτερογενούς ροής των εξωγενών, προστίθενται στοιχείο προς στοιχείο για να παραχθεί η τελική έξοδος.

Αυτή η προσέγγιση εξασφαλίζει μια φτηνή υπολογιστικά και αποδοτική ροή, που είναι ιδανική για σύγκριση των εξωγενών. Η επιλογή της γραμμικής συνάρτησης ενεργοποίησης βασίζεται κυρίως στην απλότητα και στην ανάγκη σύγκρισης της επίδρασης των εξωγενών χαρακτηριστικών, αφού η ακρίβεια της τελικής πρόβλεψης είναι ασήμαντη. Η γραμμική ενεργοποίηση επιτρέπει στις εξωγενείς μεταβλητές να περνούν απευθείας στην έξοδο χωρίς μη γραμμικές παραμορφώσεις, διατηρώντας την ταυτότητα των επιδράσεών τους. Έτσι,

διασφαλίζεται ότι κάθε εξωγενής μεταβλητή συνεισφέρει με τον ίδιο τρόπο, καθιστώντας ευκολότερη τη σύγκριση της συνεισφοράς τους. Επιπλέον, η χρήση ενός μόνο πυκνού στρώματος μειώνει σημαντικά το υπολογιστικό κόστος, αυξάνοντας την ταχύτητα εκτέλεσης. Το μοντέλο παραμένει απλό, καθώς οι παράμετροι που πρέπει να εκπαιδευτούν είναι ελάχιστες.

Μορφή εξωγενούς μεταβλητής καιρικών συνθηκών

Όπως ήδη έχει περιγραφεί η μεταβλητή που περιγράφει τις καιρικές συνθήκες περιγράφεται από ένα σύνολο 6 στηλών, με τις πρώτες 3 να περιγράφουν την μέγιστη, ελάχιστη και μέση τιμή θερμοκρασίας σε μια ημέρα και τις υπόλοιπες 3 την απόκλιση αυτών των τιμών από τη μέση θερμοκρασία των 18°C. Άλλες κωδικοποιήσεις που δοκιμάστηκαν είναι πρώτα η χρήση μόνο των τιμών της ελάχιστης, μέσης και μέγιστης θερμοκρασίας και δεύτερον η χρυσή μόνο της απόκλισης των αντίστοιχων τιμών. Έτσι δημιουργήθηκαν τρεις διαφορετικές αναπαραστάσεις για τη θερμοκρασία, με την πρώτη να αποτελείται από 6 στήλες (τιμές - αποκλίσεις), η δεύτερη από 3 στήλες (τιμές) και τέλος η τρίτη που και αυτή αποτελείται από 3 στήλες (αποκλίσεις).

Τελικά οι τρεις αναπαραστάσεις συνοψίζονται ως εξής :

- Δείγματα 6 διαστάσεων, με τιμές θερμοκρασίας και τις αποκλίσεις τους.
- Δείγματα 3 διαστάσεων, μόνο με τιμές θερμοκρασίας.
- Δείγματα 3 διαστάσεων, μόνο με αποκλίσεις των τιμών θερμοκρασίας.

Μορφή εξωγενούς μεταβλητής ημερολογιακής ημέρας

Για την μεταβλητή αυτή επίσης κατασκευάστηκαν τρεις διαφορετικές αναπαραστάσεις. Πρώτη είναι η αναπαράσταση που ήδη περιεγράφηκε με 21 στήλες και boolean δυαδικές τιμές σε στήλες 2, 7, 12.

Συνεχίζοντας η δεύτερη αναπαράσταση που δοκιμάστηκε βασίστηκε στο γεγονός ότι οι ημερολογιακές ημέρες αποτελούν χρονικά περιοδικές μεταβλητές και έχουν επαναληπτική φύση. Η εναλλαγή των επτά ημερών της εβδομάδας καθώς επίσης και των δώδεκα μηνών του έτους δημιουργούν ένα σταθερό κύκλο όπου το τέλος συνδέεται με την αρχή. Μια πιο προηγμένη προσέγγιση λοιπόν περιλαμβάνει τη χρήση ημιτόνων και συνημίτονων, ως περιοδικές συναρτήσεις, προκειμένου να μοντελοποιηθεί η παραπάνω συμπεριφορά. Κάθε ημέρα της εβδομάδας περιγράφεται από δύο τιμές x και y οι οποίες ορίζονται ως :

$$x = \sin\left(\frac{2\pi d}{7}\right) \text{ και } y = \cos\left(\frac{2\pi d}{7}\right)$$

Αντίστοιχα οι μήνες του έτους περιγράφονται από τις τιμές :

$$x = \sin\left(\frac{2\pi d}{12}\right) \text{ και } y = \cos\left(\frac{2\pi d}{12}\right)$$

Τέλος η τρίτη αναπαράσταση της μεταβλητής των ημερών που δοκιμάστηκε είναι ο διαχωρισμός των ημερών μόνο σε εργάσιμες και Σαββατοκύριακα. Η επιλογή αυτής της αναπαράστασης βασίζεται στην παρατήρηση ότι οι ανθρώπινες δραστηριότητες και τα μοτίβα ζήτησης παρουσιάζουν σημαντικές διαφορές από εργάσιμες ημέρες σε Σαββατοκύριακα και ταυτόχρονα πολλές ομοιότητες μεταξύ των εργάσιμων ημερών. Οι εργάσιμες ημέρες χαρακτηρίζονται από σταθερά και σχεδόν ομοιόμορφα μοτίβα κατανάλωσης ενέργειας, λόγω της επαναλαμβανόμενης φύσης των ωρών εργασίας, της κυκλοφορίας και των βιομηχανικών δραστηριοτήτων ενώ, αντίθετα τα Σαββατοκύριακα παρουσιάζουν διαφορετική κατανομή δραστηριοτήτων, χαλάρωση και μεταβολή στη ζήτηση. Αυτή η αναπαράσταση επιτρέπει στο μοντέλο να καταγράψει μόνο τις σημαντικές διαφοροποιήσεις μεταξύ των ημερολογιακών ημερών με πολύ μικρότερο αριθμό διαστάσεων των δεδομένων με αποτέλεσμα την μείωση της πολυπλοκότητας και συνάμα του υπολογιστικού κόστους.

Τελικά οι τρεις αναπαραστάσεις συνοψίζονται ως εξής :

- Πρώτα η αρχική boolean αναπαράσταση, όπου κάθε δείγμα αναπαρίσταται με 21 διαστάσεις.
- Δεύτερη η περιοδική αναπαράσταση όπου κάθε δείγμα έχει 6 διαστάσεις, 2 τιμές boolean για την ώρα, 2 τιμές για την ημέρα και 2 τιμές για τον μήνα.
- Η τρίτη αναπαράσταση αποτελείται από δείγματα με 16 διαστάσεις, 2 που σχετίζονται με την boolean τιμή, 2 για εργάσιμες ή μη ημέρες και τέλος 12 διαστάσεις για τους μήνες του έτους.

Μορφή εξωγενούς μεταβλητής χώρας

Αναφορικά με την μεταβλητή της χώρας δοκιμάστηκαν συνολικά τρεις συνδυασμοί. Πρώτα χρησιμοποιήθηκε η αρχική αναπαράσταση με την κλασσική one-hot κωδικοποίηση η οποία περιγράφεται από δείγματα με 24 στήλες. Για τις υπόλοιπες δυο αναπαραστάσεις δοκιμάστηκε η ομαδοποίηση των εικοσιτεσσάρων χωρών της Ευρωπαϊκής Ένωσης με δυο διαφορετικά κριτήρια. Τονίζεται πως η αναπαράσταση των ομάδων έγινε πάλι με one-hot κωδικοποίηση.

Έτσι στη μία περίπτωση οι χώρες ομαδοποιήθηκαν με βάση το ενεργειακό προφίλ τους. Αναλυτικότερα πραγματοποιήθηκε διαχωρισμός σε έξι ομάδες με βάση τον κοινό τόπο του ενεργειακού προφίλ κάθε χώρας. Έτσι κάθε ομάδα μοιράζεται παρόμοια χαρακτηριστικά παραγωγής και κατανάλωσης ενέργειας. Στη Βόρεια Ευρώπη, Δανία, Φινλανδία, Νορβηγία, Σουηδία, η υψηλή διείσδυση υδροηλεκτρικής, αιολικής και γεωθερμικής ενέργειας σε συνδυασμό με πολιτικές ενεργειακής αποδοτικότητας εξασφαλίζουν χαμηλή χρήση ορυκτών και τεράστιες εξαγωγές ηλεκτρισμού. Στη Δυτική Ευρώπη, Αυστρία, Βέλγιο, Γαλλία, Γερμανία, Λουξεμβούργο, Ολλανδία, Ελβετία, η μετάβαση στις ΑΠΕ συμπληρώνεται από σημαντική χρήση πυρηνικής ενέργειας. Ταυτόχρονα πραγματοποιείται εισαγωγή μεγάλων ποσοτήτων φυσικού αερίου για να καλυφθούν οι μεγάλες βιομηχανικές

ανάγκες. Στη Νότια Ευρώπη, Ελλάδα, Ιταλία, Πορτογαλία, Ισπανία, κυριαρχεί το ηλιακό δυναμικό και συνεχώς αυξάνεται και η αιολική παραγωγή, αλλά ταυτόχρονα υπάρχει υψηλή εισαγωγή φυσικού αερίου και πετρελαίου. Στην Κεντρική Ευρώπη, Τσεχία, Ουγγαρία, Πολωνία, Σλοβακία, Σλοβενία, η βαριά βιομηχανία βασίζεται ακόμη σε λιγνίτη και φυσικό αέριο, ενώ πρόσφατα άρχισαν να ενσωματώνονται οι ανανεώσιμες πηγές. Στη Νοτιοανατολική Ευρώπη, Βουλγαρία, Κροατία, Ρουμανία, Σερβία, Μαυροβούνιο, το ενεργειακό μίγμα αποτελείται κυρίως από μίξη ενέργειας παραγόμενης από υδροηλεκτρικά έργα και ορυκτών καυσίμων, ενώ ταυτόχρονα η μετάβαση σε ΑΠΕ ενισχύεται με ευρωπαϊκή στήριξη. Τέλος, οι χώρες της Βαλτικής, Εσθονία, Λετονία, Λιθουανία, βρίσκονται σε ταχεία ενεργειακή μετάβαση, μειώνοντας την εξάρτηση από εισαγωγές ενώ ταυτόχρονα ενισχύουν την παραγωγή από βιομάζα και αιολικά πάρκα.

Στην άλλη περίπτωση η ομαδοποίηση έγινε με βάση την εγγύτητα των ιστορικών δεδομένων φορτίου, με χρήση του αλγόριθμου ομαδοποίησης *k-Means*. Ο αλγόριθμος ξεκινά επιλέγοντας τυχαία k κέντρα (centroids) και στη συνέχεια αναθέτει κάθε δεδομένο στο πλησιέστερο κέντρο βάσει της ευκλείδειας απόστασης. Μετά την πρώτη ομαδοποίηση, υπολογίζει εκ νέου τα κέντρα ως τον μέσο όρο των σημείων που ανήκουν στη κάθε ομάδα. Έπειτα επαναλαμβάνει διαδοχικά τα δύο αυτά βήματα έως ότου τα κέντρα σταθεροποιηθούν, ελαχιστοποιώντας έτσι την εσωτερική διασπορά των ομάδων. Ο αριθμός k είναι υπερπαράμετρος, δηλαδή ο χρήστης επιλέγει σε πόσες ομάδες θα γίνει ο διαχωρισμός των δεδομένων. Για την εύρεση της τιμής του k έγινε πειραματισμός και εν τέλει βρέθηκε πως ο βέλτιστος αριθμός ήταν οι πέντε ομάδες ($k = 5$).

Τελικά οι τρεις αναπαραστάσεις συνοψίζονται ως εξής :

- One-Hot κωδικοποίηση των 24 χωρών με κάθε χώρα να θεωρείται μια διάσταση και επομένως τα δείγματα περιγράφονται με 24 διαστάσεις.

Οι 24 αυτές χώρες είναι οι : Ολλανδία, Σουηδία, Λουξεμβούργο, Τσεχία, Λετονία, Ιταλία, Ισπανία, Γαλλία, Σλοβακία, Δανία, Ελβετία, Ελλάδα, Γερμανία, Πολωνία, Νορβηγία, Ουγγαρία, Μαυροβούνιο, Σλοβενία, Κροατία, Πορτογαλία, Λιθουανία, Φινλανδία, Ρουμανία, Βουλγαρία, Σερβία, Αυστρία, Εσθονία, Βέλγιο.

- One-Hot κωδικοποίηση των 6 ομάδων χωρών με βάση το ενεργειακό τους προφίλ. Κάθε δείγμα περιγράφεται με 6 διαστάσεις.
 1. Δανία, Φινλανδία, Νορβηγία, Σουηδία
 2. Αυστρία, Βέλγιο, Γαλλία, Γερμανία, Λουξεμβούργο, Ολλανδία, Ελβετία
 3. Ελλάδα, Ιταλία, Πορτογαλία, Ισπανία
 4. Τσεχία, Ουγγαρία, Πολωνία, Σλοβακία, Σλοβενία
 5. Βουλγαρία, Κροατία, Ρουμανία, Σερβία, Μαυροβούνιο
 6. Εσθονία, Λετονία, Λιθουανία

- One-Hot κωδικοποίηση των 5 ομάδων χωρών με βάση τον αλγόριθμο k-Means. Κάθε δείγμα περιγράφεται με 5 διαστάσεις
 1. Λουξεμβούργο, Λετονία, Σλοβακία, Δανία, Μαυροβούνιο, Σλοβενία, Κροατία, Βουλγαρία, Σερβία
 2. Γαλλία, Γερμανία
 3. Ολλανδία, Σουηδία, Πολωνία, Νορβηγία
 4. Ιταλία, Ισπανία
 5. Τσεχία, Ελβετία, Ελλάδα, Ουγγαρία, Πορτογαλία, Φινλανδία, Ρουμανία, Αυστρία, Βέλγιο

4.3.3 Συνδυασμός εξωγενών μεταβλητών

Σε αυτό το στάδιο της πειραματικής διαδικασίας, αφού έχουν βρεθεί οι πιο αποδοτικές μορφές των εξωγενών μεταβλητών, τροφοδοτούνται στο δίκτυο με την ίδια βασική αρχιτεκτονική όπως και προηγουμένως, οι εξής συνδυασμοί :

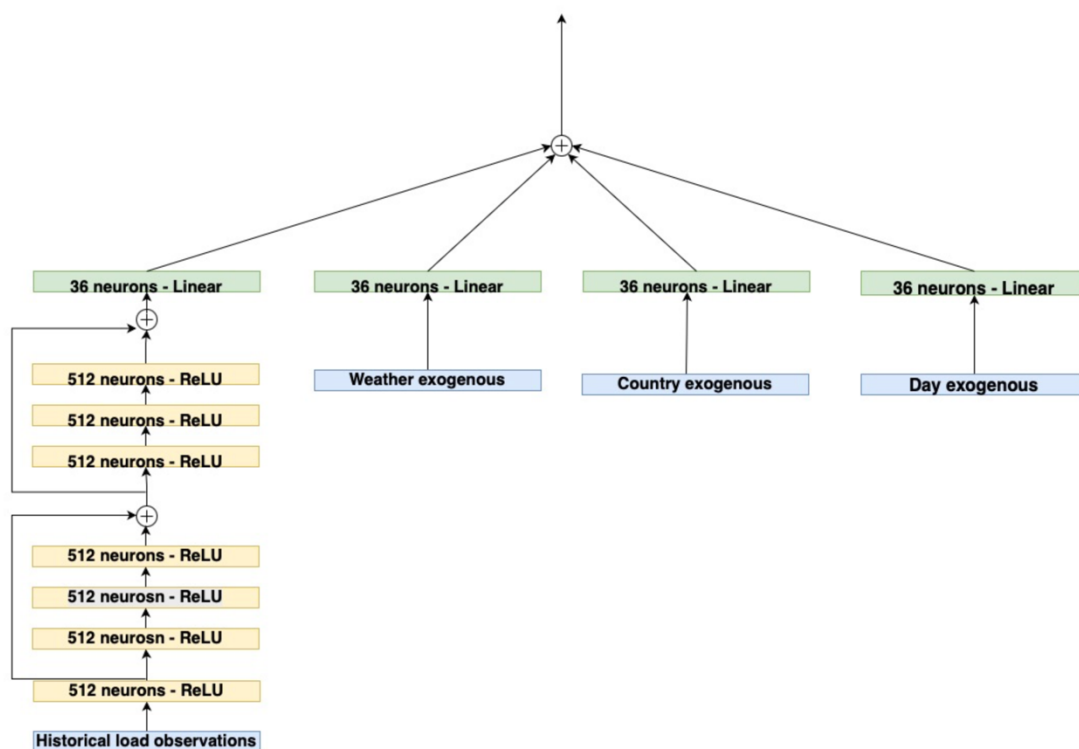
- Εξωγενής μεταβλητή καιρού και ημερολογιακής ημέρας.
- Εξωγενής μεταβλητή ημερολογιακής ημέρας και χώρας.
- Εξωγενής μεταβλητή χώρας και καιρού.
- Εξωγενής μεταβλητή χώρας και καιρού και ημερολογιακής ημέρας.

4.3.4 Αρχιτεκτονικές

Αφού πλέον έχουν εντοπιστεί για κάθε εξωγενή μεταβλητή η πιο αποδοτική της αναπαράσταση καθώς επίσης και ο καλύτερος συνδυασμός του, ξεκίνησε ο πειραματισμός και η ανάπτυξη νευρωνικών μοντέλων προκειμένου να εντοπιστεί η βέλτιστη αρχιτεκτονική ενσωμάτωσης των εξωγενών δεδομένων στο αρχικό πλήρως διασυνδεδεμένο νευρωνικό δίκτυο. Αντίστοιχα, για κάθε μία από τις υποψήφιες αρχιτεκτονικές πραγματοποιήθηκε εκπαίδευση του μοντέλου και αξιολόγηση των αποτελεσμάτων σε σύγκριση με τις αρχικές τιμές προκειμένου να ααρχιτεκτονικές μοντέλων.

Στην μεγαλύτερη πλειοψηφία αυτών των πειραμάτων κάθε εξωγενής μεταβλητή αποτελεί μια ανεξάρτητη ροή παράλληλη με την κύρια ροή του κεντρικού νευρωνικού και τελικά αυτές οι τέσσερις παράλληλες και ανεξάρτητες ροές συγχωνεύονται με διαφορετικούς τρόπους προκειμένου να σχηματιστεί ένα σύστημα προβλέψεων το οποίο χρησιμοποιεί ιστορικές τιμές φορτίου αλλά και εξωγενείς μεταβλητές.

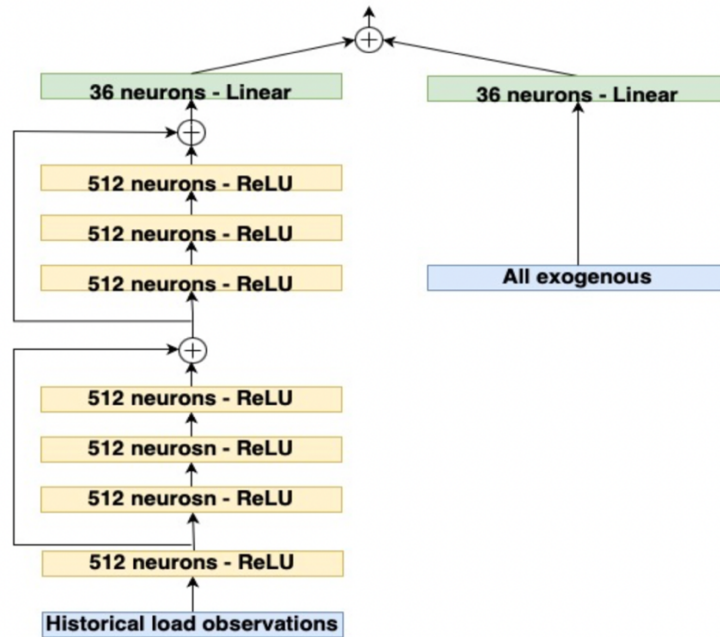
Αναλυτικά στο πρώτο πείραμα που υπάρχουν ανεξάρτητες εξωγενείς ροές, όπου κάθε εξωγενείς μεταβλητή τροφοδοτείται σε ένα πυκνό στρώμα με 512 νευρώνες και συνάρτηση ενεργοποίησης την γραμμική. Τελικά οι συνολικά τέσσερις τανιστές, των εξωγενών ροών και ενός της εξόδου του πλήρως διασυνδεδεμένου μοντέλου των ιστορικών δεδομένων του φορτίου προστίθενται στοιχείο προς στοιχείο. Αυτή θεωρείται και η έξοδος του συστήματος πρόβλεψης.



Σχήμα 4.2 : Πείραμα 1

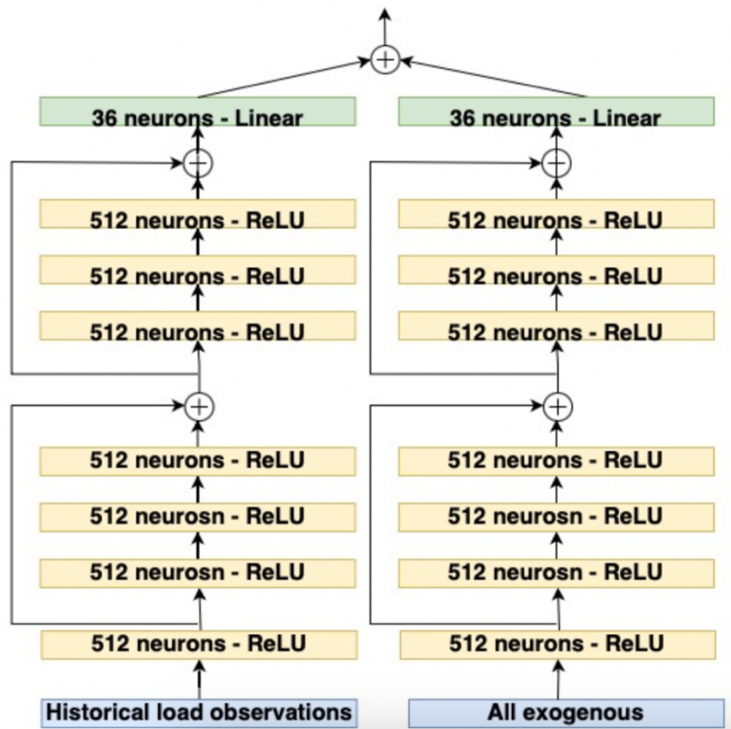
Στα πειράματα δύο, τρία και τέσσερα οι εξωγενείς μεταβλητές συνενώνονται σε μια ενιαία ροή, αντί για τρεις ανεξάρτητες και αυτόνομες. Αυτό μπορεί να προσφέρει σημαντικά πλεονεκτήματα σε ένα πλήρως διασυνδεδεμένο νευρωνικό δίκτυο, καθώς η ενοποίηση της πληροφορίας επιτρέπει στο δίκτυο να αντιλαμβάνεται τις μεταβλητές ως μια συνδυασμένη "μεγάλη μεταβλητή", βοηθώντας το να μαθαίνει πιο σύνθετες αλληλεπιδράσεις μεταξύ τους. Τέτοιο παράδειγμα αποτελεί το πώς η θερμοκρασία και το αν είναι εργάσιμη ή όχι ημέρα μαζί επηρεάζουν την κατανάλωση. Δεύτερον, η συνένωση απλοποιεί την αρχιτεκτονική του δικτύου, μειώνοντας την πολυπλοκότητα και τον υπολογιστικό φόρτο, καθώς αποφεύγεται η ανάγκη για ξεχωριστά υπο-δίκτυα ή επίπεδα για κάθε ροή. Τρίτον, βελτιώνει τη γενίκευση του μοντέλου, καθώς η ενιαία ροή μειώνει τον κίνδυνο υπερ-προσαρμογής σε μεμονωμένες μεταβλητές, βοηθώντας το δίκτυο να "βλέπει" τη συνολική εικόνα και να αποδίδει καλύτερα σε νέα δεδομένα. Τέλος, ενισχύει τη μη γραμμική μάθηση, καθώς το δίκτυο μπορεί να εκμεταλλευτεί την ικανότητά του να συνδυάζει τις μεταβλητές με τρόπους που αποκαλύπτουν κρυφές σχέσεις, όπως για παράδειγμα η σημασία ενός συγκεκριμένου συνδυασμού τιμών για την πρόβλεψη.

Έτσι στο δεύτερο πείραμα οι εξωγενείς μεταβλητές συνενώνονται και τροφοδοτούνται σε μια ενιαία ροή παράλληλη με τη ροή των ιστορικών δεδομένων. Η ροή των εξωγενών είναι αποτελείται από ένα πυκνό γραμμικό στρώμα με έξοδο 36 νευρώνες.



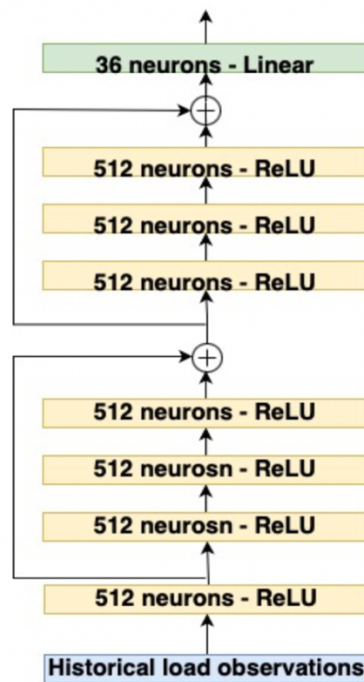
Σχήμα 4.3 : Πείραμα 2

Στο τρίτο πείραμα η εξωγενείς μεταβλητές αφού συνενωθούν τροφοδοτούνται ως είσοδοι σε μια ενιαία ροή, η οποία αποτελείται από την ίδια αρχιτεκτονική του πλήρους δε συνδεδεμένο νευρωνικού δικτύου που χρησιμοποιείται και στα ιστορικά δεδομένα.



Σχήμα 4.4 : Πείραμα 3

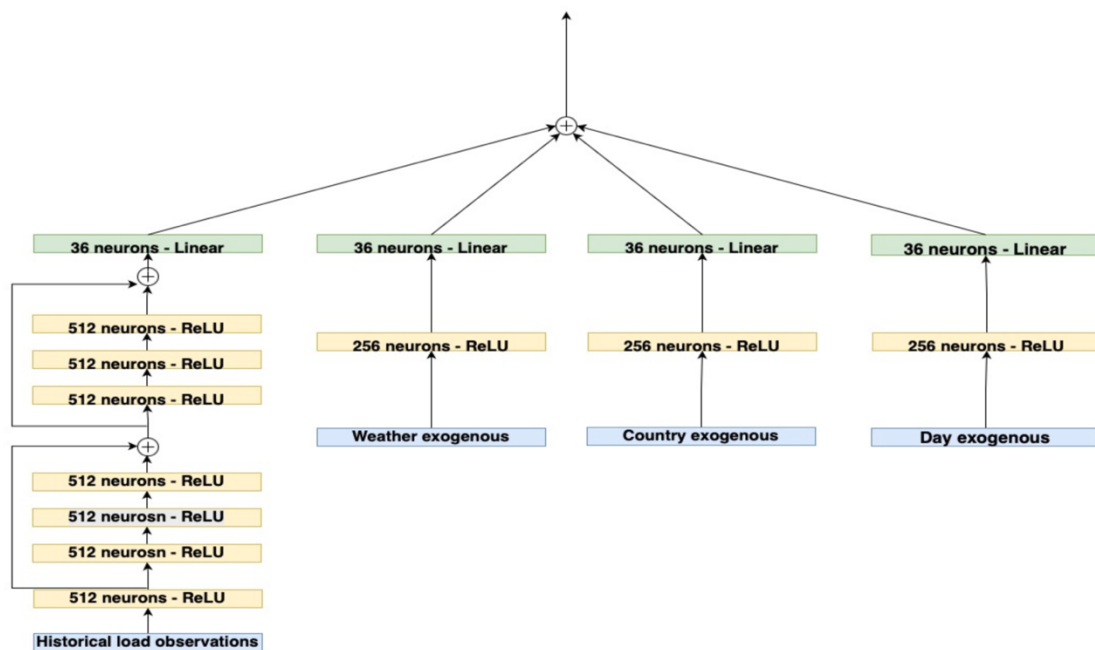
Στο τέταρτο πείραμα όλα τα δεδομένα, τα ιστορικά και τα εξωγενή, συνενώνονται και έπειτα τροφοδοτούνται ως μία έννοια είσοδος στο πλήρες διασυνδεδεμένο νευρωνικό δίκτυο.



Σχήμα 4.5 : Πείραμα 4

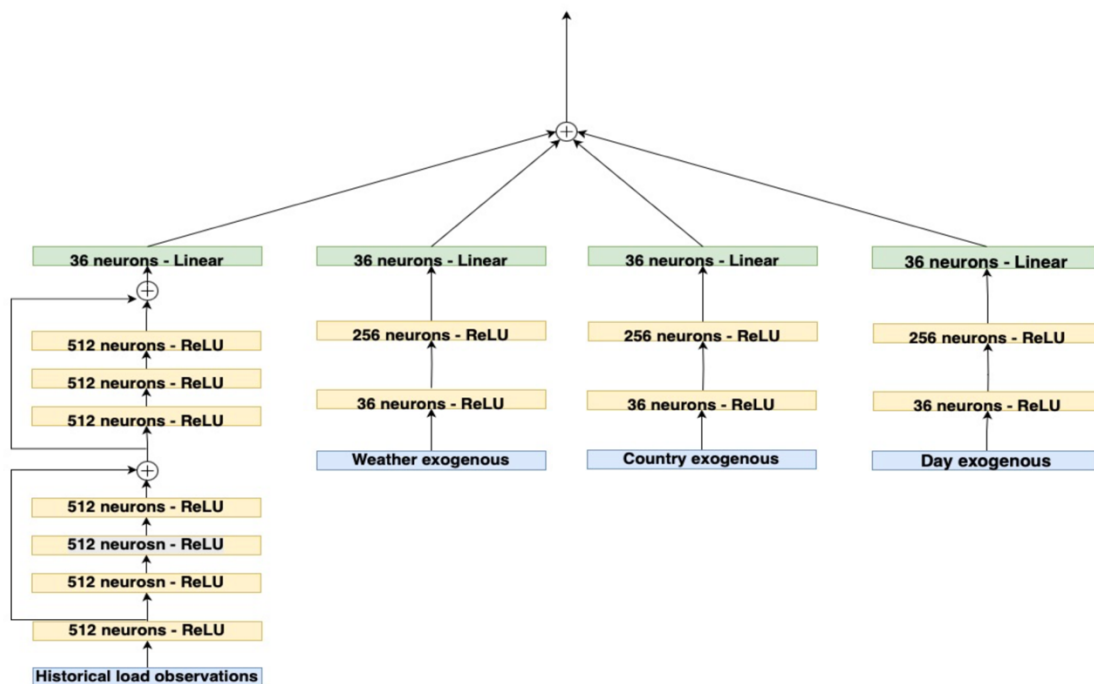
Σε αυτό το σημείο σημειώνεται πως σε όλα τα επόμενα πειράματα οι εξωγενείς μεταβλητές κατά κανόνα τροφοδοτούνται σε ανεξάρτητες ροές η καθεμία, οι οποίες είναι παράλληλες με την ροή των ιστορικών δεδομένων. Ακόμη η ροή των ιστορικών δεδομένων δεν μεταβάλλεται και είναι ίδια με αυτή που συζητείται από την αρχή της πειραματικής μεθόδου. Έτσι όλοι οι πειραματισμοί σχετίζονται με μεθόδους ενσωμάτωσης αυτών των τεσσάρων παράλληλων ροών καθώς επίσης και με την εσωτερική δομή καθεμιάς από τις εξωγενείς ροές, με τον κανόνα όμως πως όλες οι εξωγενείς ροές περιγράφονται από την ίδια ακριβώς αρχιτεκτονική σε κάθε πείραμα.

Συνεχίζοντας στο πέμπτο πείραμα οι εξωγενείς ροές αποτελούνται από ένα «κρυφό» πυκνό στρώμα 256 νευρώνων με συνάρτηση ενεργοποίησης την ReLU και από ένα γραμμικό στρώμα 36 νευρώνων. Για την βέλτιστη επιλογή του μεγέθους του «κρυφού» πυκνού στρώματος έγιναν πάλι πειραματισμοί ο πίνακας των οποίων επισυνάπτεται στο τέλος. Τελικά οι τέσσερις ροές προστίθεται στοιχείο προς στοιχείο και σχηματίζουν την τελική έξοδο του συστήματος.



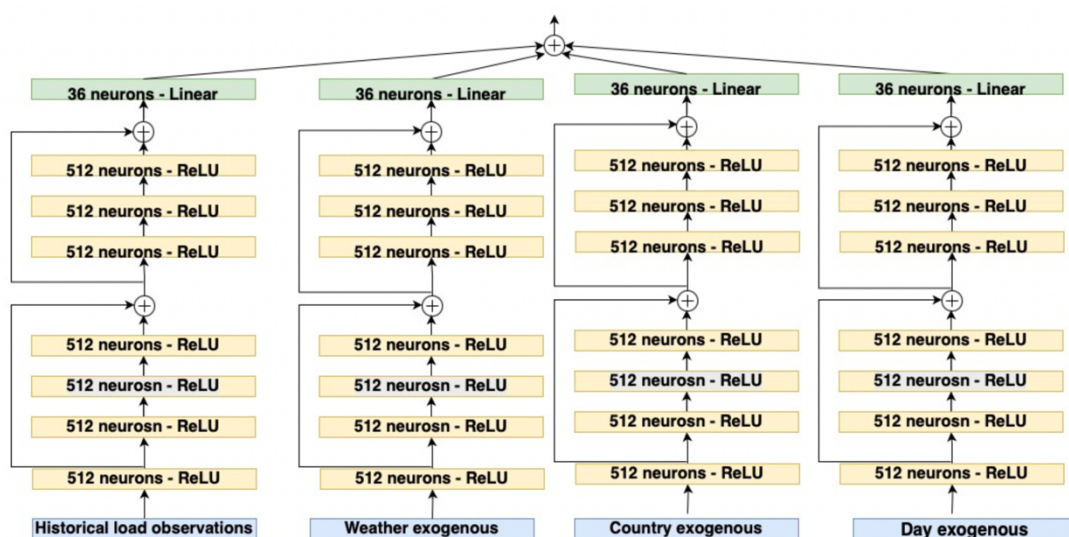
Σχήμα 4.6 : Πείραμα 5

Στο έκτο πείραμα ακολουθείται ίδια λογική με το προηγούμενο, με τη διαφορά πως οι εξωγενείς ροές αποτελούνται από δύο «κρυφά» στρώματα και έπειτα από το ένα γραμμικό. Πρόκειται για δυο στρώματα με μέγεθος 36 νευρώνες και 256 νευρώνες αντίστοιχα και συνάρτηση ενεργοποίησης της ReLU. Το γραμμικό στρώμα παραμένει το ίδιο με προηγουμένως. Σημειώνεται πως και πάλι για τον εντοπισμό των βέλτιστων παραμέτρων των δύο κρυφών στρωμάτων πάλι έγινε πειραματισμοί ο πίνακας με τα αποτελέσματα των οποίων επισυνάπτεται στο τέλος.



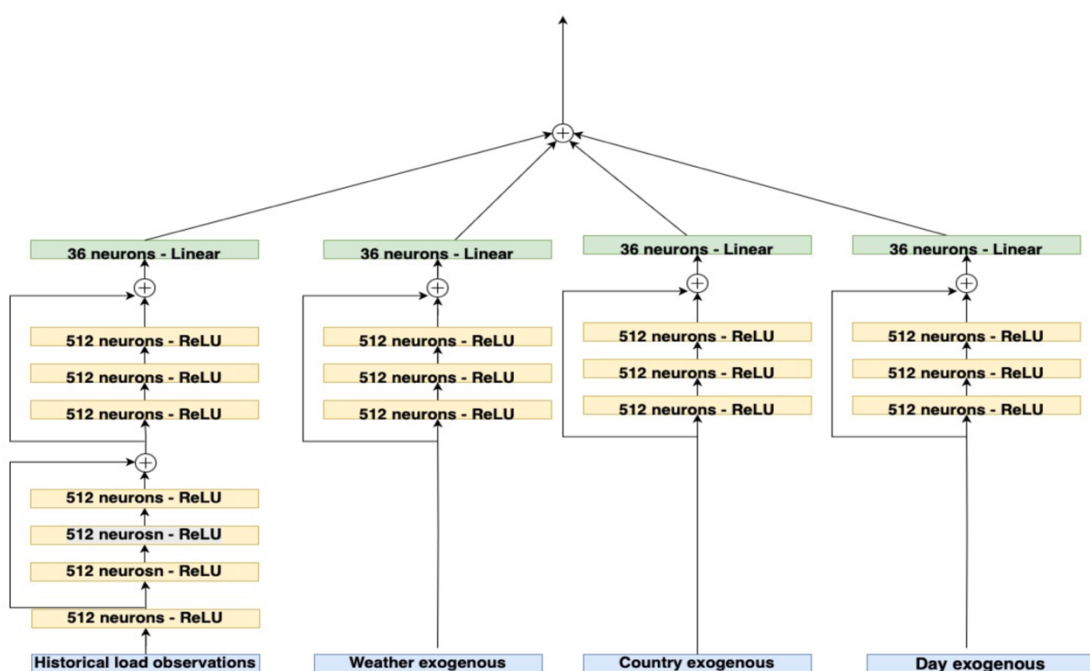
Σχήμα 4.7 : Πείραμα 6

Στο έβδομο πείραμα οι εξωγενείς ροές έχουν την ίδια ακριβώς αρχιτεκτονική με αυτή των ιστορικών δεδομένων. Τελικά οι τέσσερις ροές προστίθενται κατά στοιχείο και σχηματίζουν την έξοδο του συστήματος. (ΠΑΛΙΟ 9)



Σχήμα 4.8 : Πείραμα 7

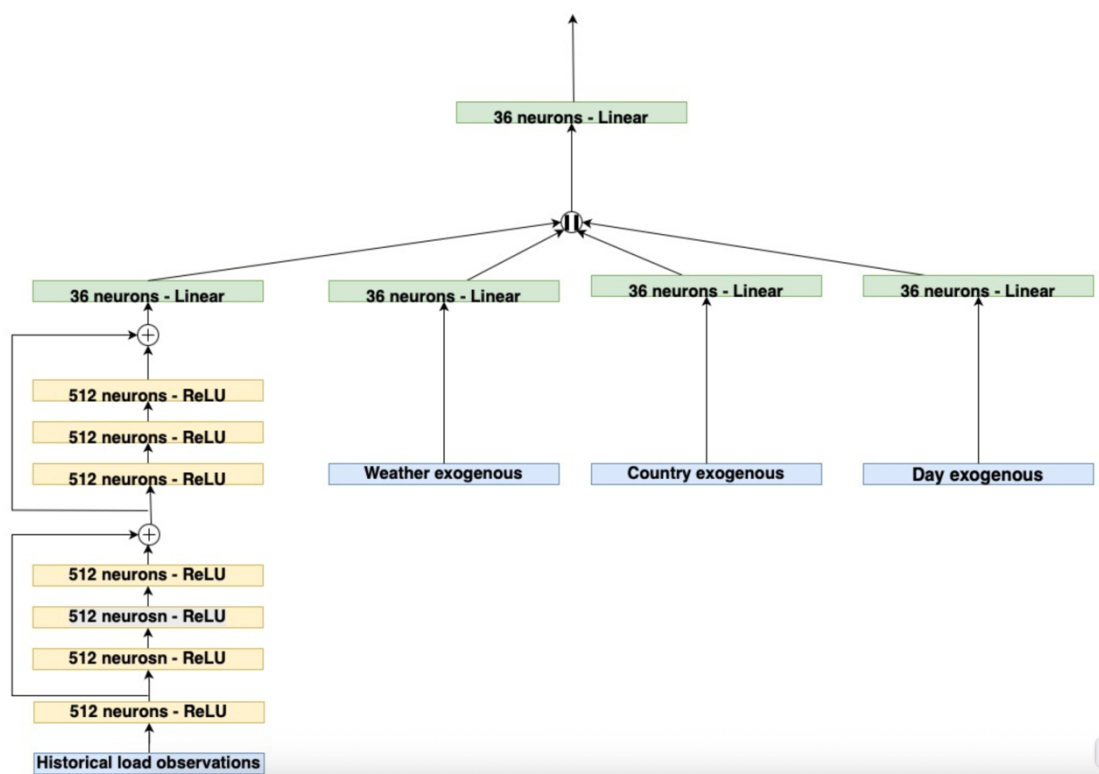
Το όγδοο πείραμα είναι παρόμοιο με το προηγούμενο με την διαφορά ότι η εξωγενείς ροές δεν περνούν την ίδια ακριβώς επεξεργασία με τα ιστορικά δεδομένα, παρά μόνο την μισή. Αυτό επιτυγχάνεται καθώς οι εξωγενείς ροές αποτελούνται μόνο από το τελευταίο μισό του συνόλου των στρωμάτων σε σχέση με την ροή των ιστορικών δεδομένων. Δηλαδή από τρία διαδοχικά πυκνά στρώματα μεγέθους 512 νευρώνων και συνάρτησης ενεργοποίησης ReLU, ενός τελικού γραμμικού στρώματος 36 νευρώνων και ενός skip connection.



Σχήμα 4.9 : Πείραμα 8

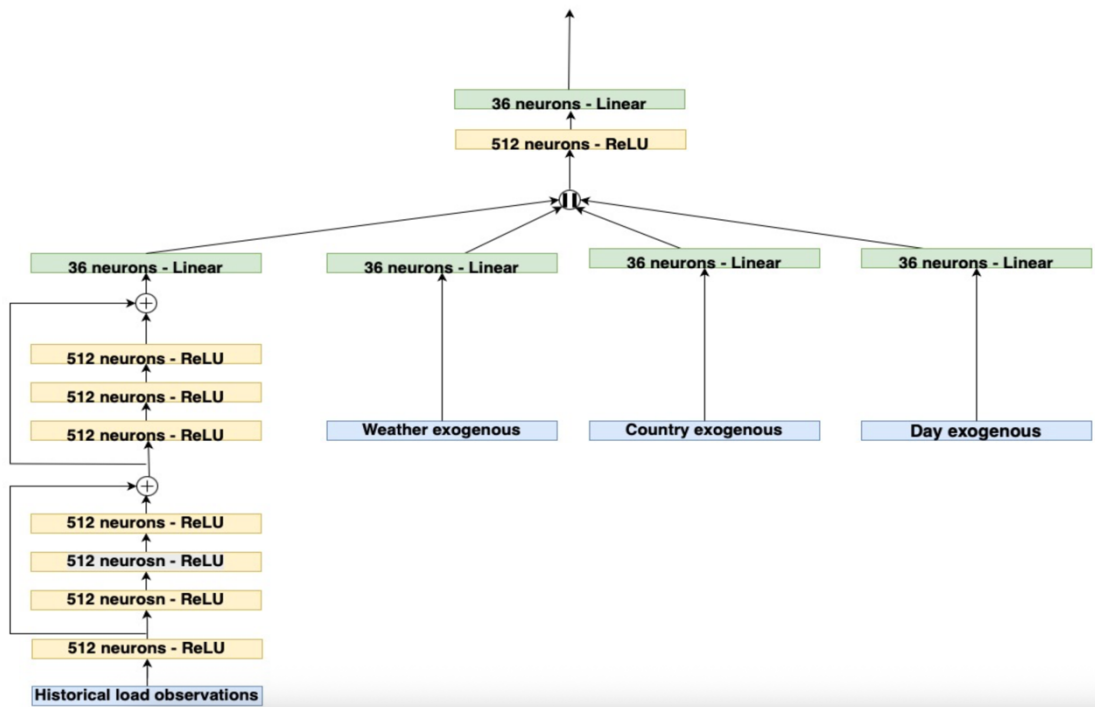
Απο αυτό το σημείο και έπειτα οι τέσσερις ροές των ιστορικών και των εξωγενών δεδομένων, συνενώνονται και δεν προστίθενται στοιχείο κατά στοιχείο. Στη συνέχεια το αποτέλεσμα της συνένωσης τροφοδοτείται σε ένα ή περισσότερα νευρωνικά στρώματα και από εκεί προκύπτει η έξοδος του συστήματος. Με τον όρο συνένωση εννοείται πως τα διανύσματα χαρακτηριστικών της κάθε μεταβλητής συνενώνονται και σχηματίζουν μια ενιαία μεταβλητή με ένα μεγάλο διάνυσμα χαρακτηριστικών.

Στο ένατο πείραμα λοιπόν οι εξωγενείς ροές αποτελούνται από ένα απλό γραμμικό στρώμα όπου τελικά συνενώνονται με την έξοδο της ροής των ιστορικών δεδομένων. Το αποτέλεσμα της συνένωσης τροφοδοτείται σε ένα πυκνό γραμμικό στρώμα 36 νευρώνων από το οποίο προκύπτει και η έξοδος του συστήματος.



Σχήμα 4.10 : Πείραμα 9

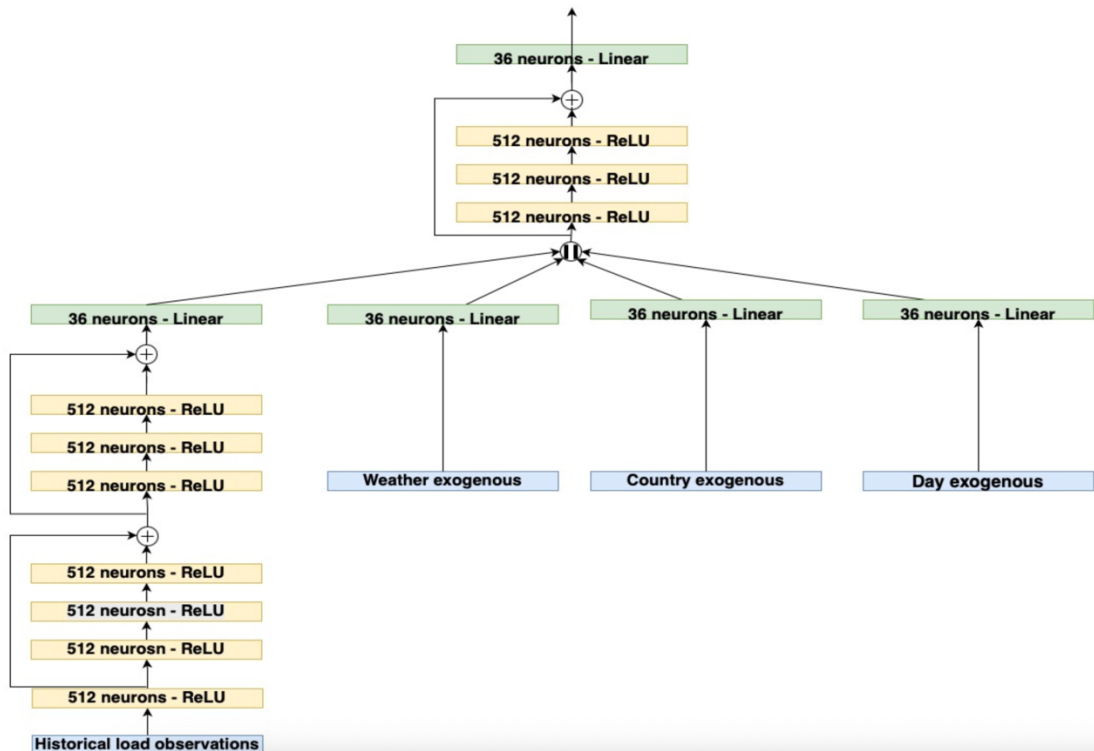
Στο δέκατο πείραμα επαναλαμβάνεται η ίδια λογική με το προηγούμενο πείραμα με τη διαφορά όμως ότι το αποτέλεσμα της συνένωσης τροφοδοτείται σε ένα πυκνό στρώμα μεγέθους 512 νευρώνων με ReLU συνάρτηση ενεργοποίησης, το οποίο διαδέχεται ένα πυκνό γραμμικό στρώμα 36 νευρώνων. Αυτό αποτελεί και την έξοδο του συστήματος.



Σχήμα 4.11 : Πείραμα 10

Στα επόμενα πειράματα που ακολουθούν η μέθοδος της συνένωσης των τεσσάρων ροών είναι η ίδια και πρόκειται για το σύνολο των διαδοχικών στρωμάτων του δεύτερου νοητού μέρους της αρχικής αρχιτεκτονικής. Αναλυτικότερα οι τέσσερις ροές συνενώνονται σε ένα ενιαίο διάνυσμα χαρακτηριστικών και έτσι όλες οι πληροφορίες συγκεντρώνονται σε μία ενιαία αναπαράσταση. Στη συνέχεια αυτό διάνυσμα περνά διαδοχικά από τρία πυκνά στρώματα μεγέθους 512 νευρώνων με ReLU συνάρτηση ενεργοποίησης, τα οποία επιτρέπουν στο δίκτυο να μάθει μη-γραμμικά μοτίβα και πολύπλοκες σχέσεις μεταξύ της ιστορικής ροής και των εξωγενών παραγόντων. Η χρήση πολλαπλών κρυφών στρωμάτων ενισχύει τη δυνατότητα του μοντέλου να αντιληφθεί κοινά μοτίβα για τις τέσσερις αυτές ροές. Για την ενίσχυση της σταθερότητας της μάθησης χρησιμοποιείται επίσης ένα skip connection το οποίο εφαρμόζεται στην έξοδο του τρίτου κρυφού στρώματος, προσθέτοντας το αποτέλεσμα στοιχείο προς στοιχείο με την έξοδο του πρώτου στρώματος της βαθιάς επεξεργασίας. Τέλος, το προσαρμοσμένο διάνυσμα διέρχεται από ένα γραμμικό πυκνό επίπεδο, το οποίο μειώνει τη διάσταση στο εύρος του ορίζοντα πρόβλεψης (π.χ. 36 ώρες). Με αυτόν τον τρόπο, το μοντέλο καταφέρνει να αξιοποιήσει πλήρως τόσο τα ιστορικά δεδομένα όσο και τους εξωγενείς παράγοντες.

Έτσι στο εντέκατο πείραμα οι εξωγενείς ροές αποτελούνται από ένα πυκνό γραμμικό στρώμα 36 νευρώνων και τελικά αυτές συνενώνονται με τη ροή των ιστορικών δεδομένων με τον τρόπο που αναλύθηκε μόλις.



Σχήμα 4.12 : Πείραμα 11

Στο δωδέκατο πείραμα επαναλαμβάνεται ακριβώς το ίδιο πείραμα με τη διαφορά ότι στα στρώματα που ακολουθούν την συνένωση δεν περιλαμβάνεται το skip connection.

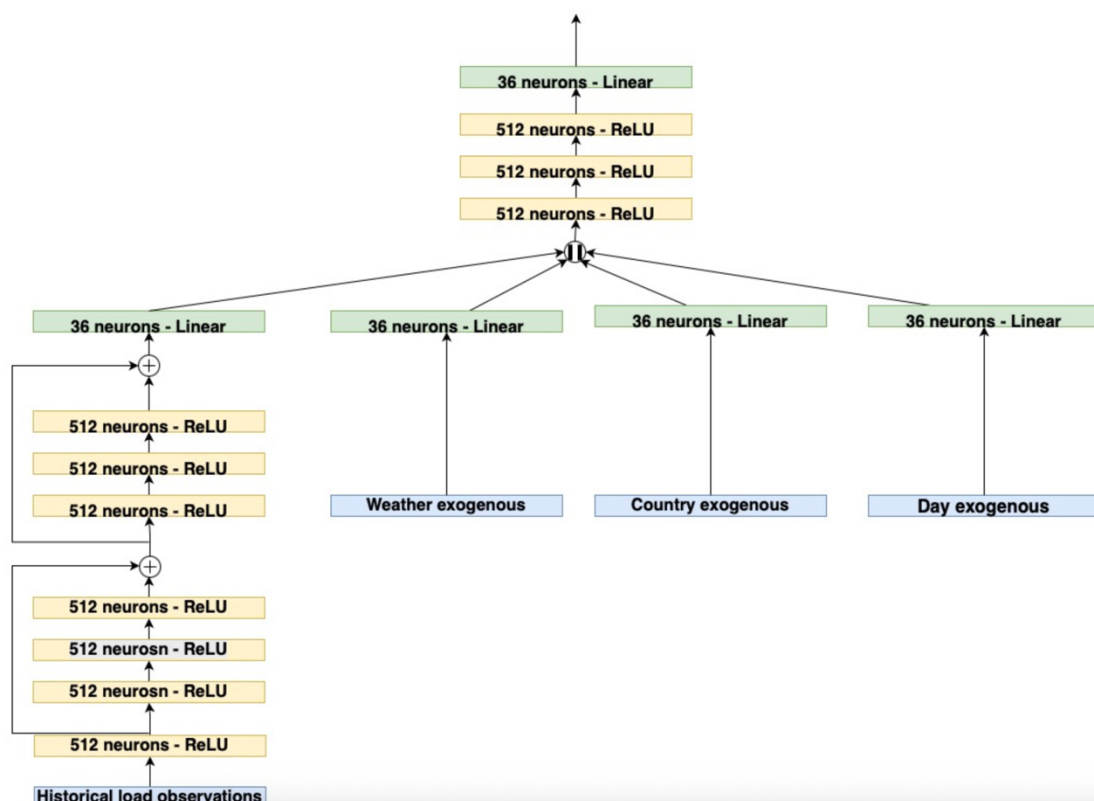
Αν και μια skip connection κατά γενική πρακτική παρουσιάζει θετικά αποτελέσματα, η αφαίρεση της μπορεί να αποδειχθεί επωφελής για διάφορους λόγους, οι οποίοι σχετίζονται με τη λειτουργία και τη δυναμική του δικτύου. Αρχικά, η skip connection επιτρέπει στα συνδυασμένα χαρακτηριστικά, που προκύπτουν από την ένωση των μεταβλητών να περάσουν απευθείας στην τελική έξοδο χωρίς σημαντική επεξεργασία από τα τρία ενδιάμεσα επίπεδα. Αυτό μπορεί να οδηγήσει το δίκτυο σε υπερβολική εξάρτηση από αυτά τα αρχικά χαρακτηριστικά, μειώνοντας την επίδραση των ενδιάμεσων επιπέδων στη μάθηση πιο αφηρημένων και χρήσιμων αναπαραστάσεων που θα μπορούσαν να βελτιώσουν την απόδοση του μοντέλου. Κατά συνέπεια, η αφαίρεση της skip connection αναγκάζει αυτά τα επίπεδα να επεξεργάζονται την πληροφορία πιο ενδελεχώς, ενισχύοντας τη μάθηση και τη συμβολή τους στη συνολική αρχιτεκτονική. Χωρίς τη skip connection, τα ενδιάμεσα επίπεδα αναγκάζονται να φιλτράρουν και να επεξεργάζονται την πληροφορία πιο προσεκτικά, λειτουργώντας σαν ένα είδος κανονικοποίησης που μειώνει τον κίνδυνο υπερπροσαρμογής και βοηθά το δίκτυο να γενικεύει καλύτερα, βελτιώνοντας την απόδοσή του σε πραγματικές συνθήκες.

Παράλληλα, το συγκεκριμένο δίκτυο δεν είναι υπερβολικά βαθύ μετά τη συνένωση, καθώς περιλαμβάνει μόνο τρία επίπεδα των 512 νευρώνων πριν την τελική έξοδο των 36 νευρώνων. Σε τέτοιες περιπτώσεις ρηχών συνόλων

στρωμάτων το πρόβλημα της εξαφάνισης της παραγώγου δεν αποτελεί σημαντικό ζήτημα.

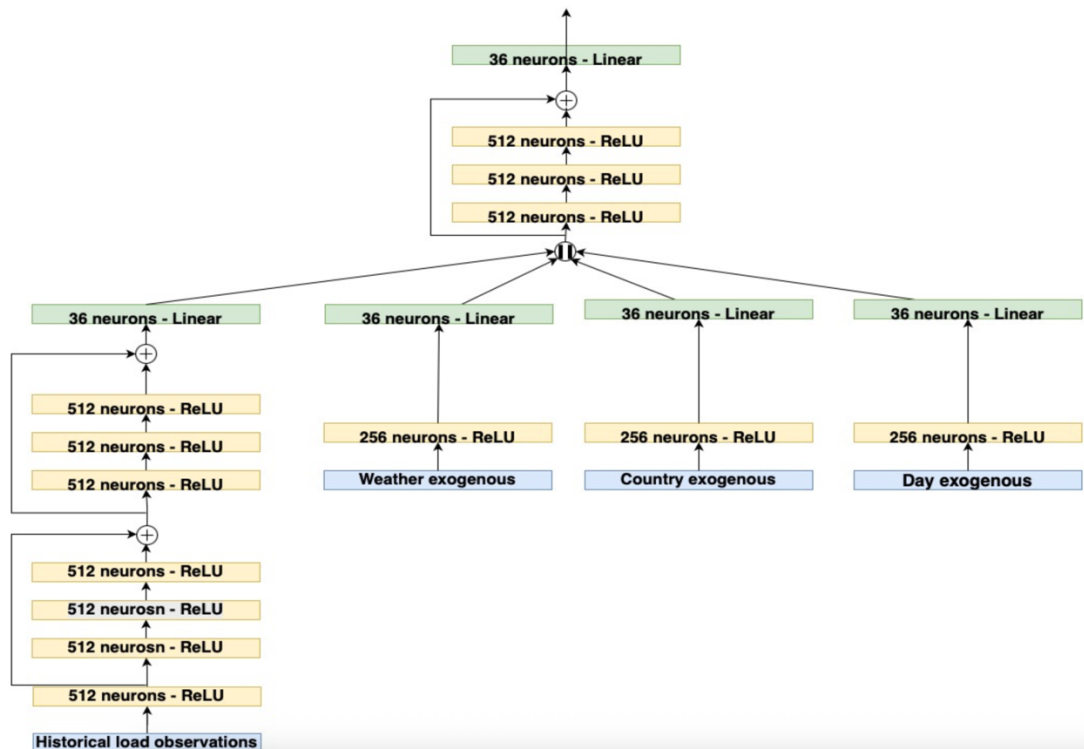
Τέλος, ο πειραματισμός με την αφαίρεση της skip connection μπορεί να προσφέρει μια βαθύτερη κατανόηση της δυναμικής του δικτύου και της συνεισφοράς κάθε επιπέδου. Δοκιμάζοντας το μοντέλο χωρίς αυτήν, είναι δυνατόν να παρατηρηθεί πώς τα επίπεδα μετά τη συνένωση επεξεργάζονται την πληροφορία και αν συνεισφέρουν ουσιαστικά στην τελική πρόβλεψη ή αν απλώς «περνούν» την πληροφορία στην έξοδο χωρίς να προσθέτουν αξία. Αυτό μπορεί να αποκαλύψει ευκαιρίες για βελτιώσεις στην αρχιτεκτονική, όπως η μείωση του αριθμού των επιπέδων, η αλλαγή των συναρτήσεων ενεργοποίησης ή η προσαρμογή άλλων υπερπαραμέτρων, οδηγώντας σε ένα πιο αποδοτικό και αποτελεσματικό μοντέλο.

Έτσι με βάση όλα τα παραπάνω, η αφαίρεση της τελευταίας skip connection αποτελεί μια στρατηγική που αξίζει να μελετηθεί, παρά το γεγονός ότι οι skip connections γενικά είναι μια καλή πρακτική που ακολουθείται ειδικά σε βαθιά νευρωνικά δίκτυα.



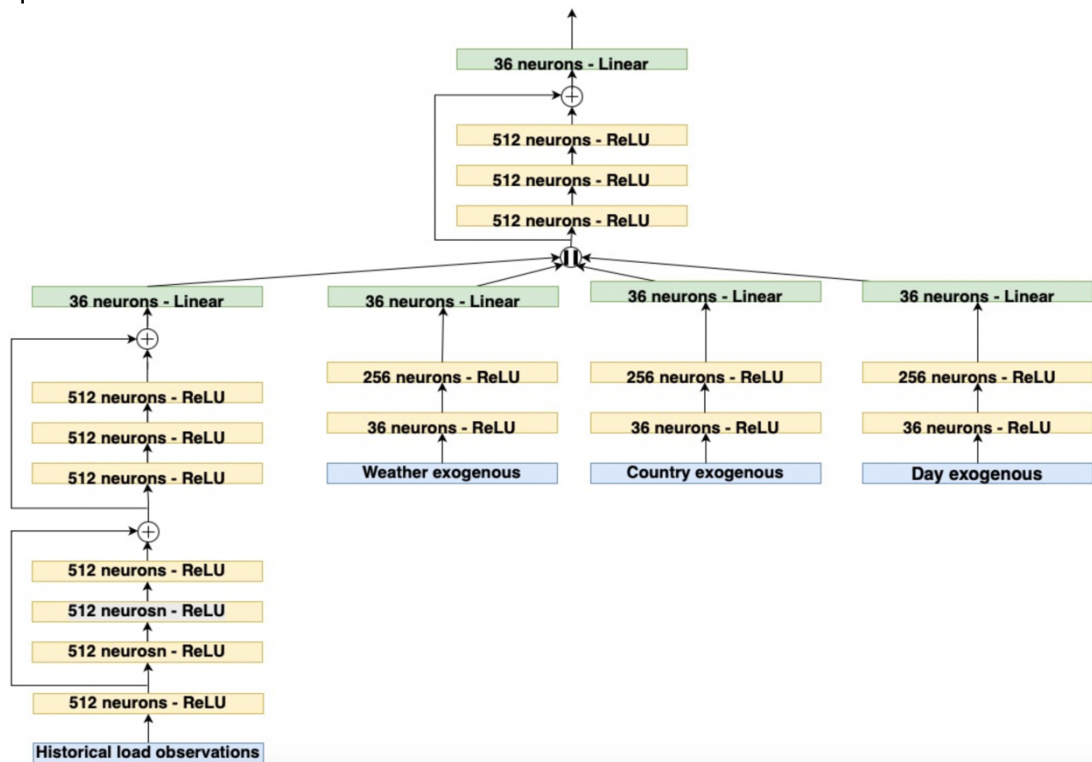
Σχήμα 4.13 : Πείραμα 12

Στο δέκατο τρίτο πείραμα οι εξωγενείς ροές πριν την συνένωση αποτελούνται από ένα κρυφό πυκνό στρώμα 256 νευρώνων με συνάρτηση ενεργοποίησης ReLU όπως και στο πείραμα πέντε.



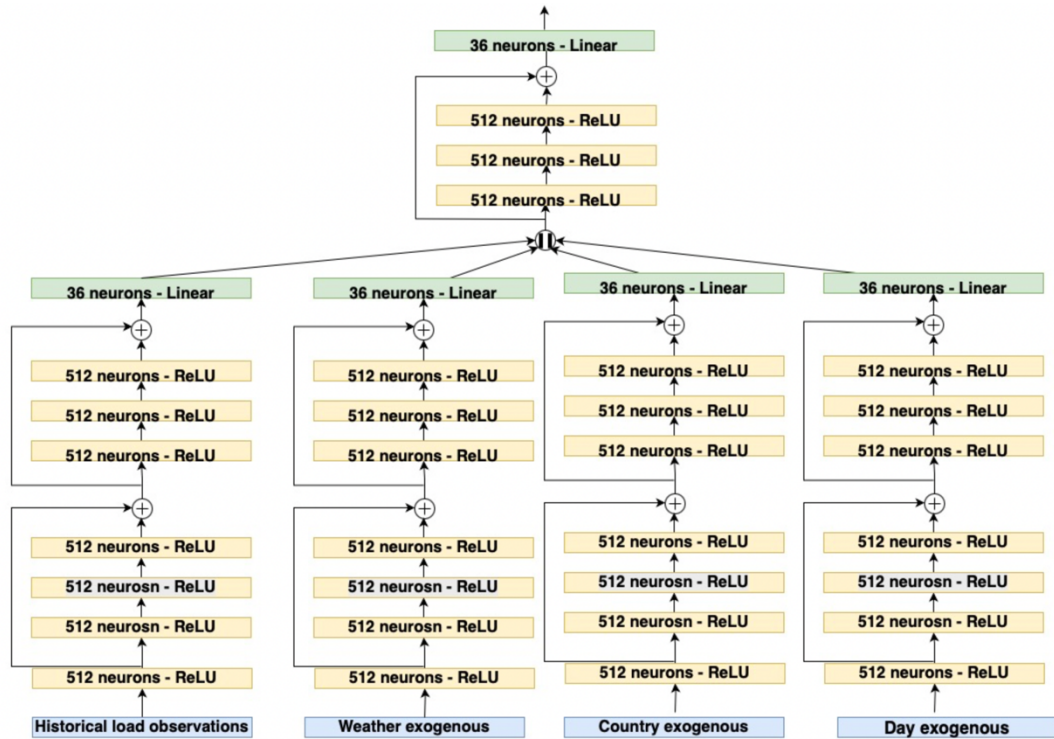
Σχήμα 4.14 : Πείραμα 13

Στο δέκατο τέταρτο πείραμα οι εξωγενείς ροές αποτελούνται από δύο κρυφά στρώματα τα οποία διαδέχεται το γραμμικό στρώμα 36 νευρώνων. Η αρχιτεκτονική είναι ίδια με του πειράματος έξι όπου τα δύο κρυφά δίκτυα αποτελούνται από 36 και 256 νευρώνες αντίστοιχα με συνάρτηση ενεργοποίησης τη ReLU.



Σχήμα 4.15 : Πείραμα 14

Στο τελευταίο πείραμα που διεξάγεται, το δέκατο πέμπτο, οι εξωγενείς ροές έχουν την ίδια ακριβώς αρχιτεκτονική με την ροή των ιστορικών δεδομένων και τελικά οι τέσσερις αυτές συνενώνονται σε μια ενιαία είσοδο η οποία τροφοδοτείται στο σύνολο των στρωμάτων όπως το έχουμε περιγράψει.



Σχήμα 4.16 : Πείραμα 15

ΚΕΦΑΛΑΙΟ 5 – ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1 Περιβάλλον ανάπτυξης

Ολόκληρη η πειραματική διαδικασία όπως η συλλογή και προεπεξεργασία των δεδομένων, καθώς και η ανάπτυξη, εκπαίδευση και αξιολόγηση των διαφορετικών αρχιτεκτονικών των νευρωνικών δικτύων πραγματοποιήθηκαν σε γλώσσα προγραμματισμού *Python 3.11.7* και επίσης έγινε εκτενής χρήση της βιβλιοθήκης *Tensorflow 2.17.0*. Η συγκέντρωση και η επεξεργασία των αποτελεσμάτων όπως επίσης και η απεικόνιση τους σε διαγράμματα και σχήματα έγινε με χρήση του λογισμικού *excel* αλλά και της *Python*. Ακόμη αξίζει να σημειωθεί πως όλη η ανάπτυξη και τα πειράματα πραγματοποιήθηκαν σε περιβάλλον *macOS* με επεξεργαστή *ARM M3*.

5.2 Δεδομένα

Τα δεδομένα που χρησιμοποιήθηκαν στη παρούσα διπλωματική εργασία είναι τα ίδια που χρησιμοποιήθηκαν για την ανάπτυξη και αξιολόγηση του αρχικού μοντέλου. Πιο συγκεκριμένα χρησιμοποιήθηκαν δεδομένα ιστορικού φορτίου από 24 ευρωπαϊκές χώρες, τα οποία συλλέχθηκαν από την πλατφόρμα διαφάνειας του ENTSO-E (European Network of Transmission System Operators for Electricity). Το σύνολο δεδομένων περιλαμβάνει ωριαίες τιμές φορτίου για περίοδο με αφετηρία τον Ιανουάριο του 2016 έως και τον Δεκέμβριο του 2022.

Το σετ εκπαίδευσης περιλαμβάνει τις ωριαίες μετρήσεις από τον Ιανουάριο του 2016 έως τον Δεκέμβριο του 2021. Χρησιμοποιήθηκε τεχνική κινούμενου παραθύρου (sliding window) με εύρος 204 ωρών, όπου οι πρώτες 168 ώρες χρησιμοποιούνται ως είσοδος στο δίκτυο και οι επόμενες 36 ώρες ως επιθυμητή έξοδος, δημιουργώντας συνολικά 101.562 δείγματα. Αντίστοιχα το σετ αξιολόγησης καλύπτει ολόκληρο το έτος 2022 και κατασκευάστηκε με την ίδια τεχνική κινούμενου παραθύρου με τη διαφορά ότι στο σετ εκπαίδευσης τα παράθυρα επικαλύπτονται με βήμα 12 ωρών, ώστε να αυξηθεί ο αριθμός των δειγμάτων και να καταγραφούν λεπτομερέστερα οι διακυμάνσεις της χρονοσειράς ενώ, αντίθετα, στο σετ αξιολόγησης το βήμα έχει εύρος 24 ώρες, οπότε τα παράθυρα δεν επικαλύπτονται μεταξύ τους. Έτσι τελικά το σετ αξιολόγησης παράγει 8.760 δείγματα και η τελική αναλογία είναι 92% σύνολο εκπαίδευσης και 8% αξιολόγησης.

Η διαδικασία προεπεξεργασίας των δεδομένων περιελάμβανε την αναπλήρωση των ελλειπών τιμών, μέχρι και πέντε συνεχόμενων ωρών με γραμμική παρεμβολή, ενώ τα κενά μεγαλύτερης έκτασης απορρίφθηκαν για την αποφυγή στρεβλώσεων στα δεδομένα. Επίσης τυχών ακραίες τιμές αφαιρέθηκαν με βάση τη τεχνική Winsorization, όπου τιμές κάτω από 0,1% και πάνω από 99,9% αφαιρέθηκαν, ώστε να διατηρηθεί η σταθερότητα της σειράς. Τέλος για την διασφάλιση της ομοιομορφίας των δεδομένων, κάθε σειρά φορτίου

κανονικοποιήθηκε με βάση τον μηδενικό μέσο όρο και τυπική απόκλιση ίση με ένα.

Επιπρόσθετα ως εξωγενείς μεταβλητές επιλέχθηκαν από την ίδια περίοδο πληροφορίες σχετικές με τον καιρό, την ημέρα της εβδομάδας και την χώρα την οποία αφορούν τα ιστορικά δεδομένα του φορτίου. Αντίστοιχα, όπως και οι τιμές φορτίου έτσι και οι εξωγενείς μεταβλητές χωρίστηκαν σε δεδομένα εκπαίδευσης και αξιολόγησης. Αναλυτικότερα για την ενσωμάτωση εξωγενών δεδομένων που αφορούν καιρικές συνθήκες, χρησιμοποιήθηκαν μετεωρολογικά δεδομένα που περιλαμβάνουν τη τιμή της μέσης, ελάχιστης και μέγιστης θερμοκρασίας της ημέρας για κάθε πρωτεύουσα των 24 ευρωπαϊκών χωρών που συμμετείχαν στη μελέτη. Τα δεδομένα αυτά συλλέχθηκαν από την πλατφόρμα *rp5.ru*. Επιπλέον για να διασφαλιστεί ότι τα δεδομένα της θερμοκρασίας σχετίζονται με τις ενεργειακές ανάγκες των χωρών, υπολογίστηκαν επιπλέον οι αποκλίσεις κάθε τιμής της θερμοκρασίας από τη μέση βασική θερμοκρασία των 18°C, η οποία θεωρείται ως ένα σημείο άνεσης, δηλαδή το όριο στο οποίο η θερμοκρασία αρχίζει να επηρεάζει την κατανάλωση ενέργειας από τους τελικούς χρήστες μέσω θέρμανσης ή ψύξης. Αυτές οι αποκλίσεις επιτρέπουν στο μοντέλο να λαμβάνει υπόψη την επίδραση των ακραίων θερμοκρασιών στη ζήτηση ενέργειας. Έτσι λοιπόν διαμορφώνονται δύο σύνολα, με το σύνολο εκπαίδευσης να αποτελείται από 101.562 δείγματα με 6 τιμές το καθένα και αντίστοιχα το σύνολο αξιολόγησης από 8.760 δείγματα και 6 τιμές το καθένα.

Συνεχίζοντας, τα εξωγενή δεδομένα της ημερολογιακής ημέρας ενσωματώθηκαν στο μοντέλο πρόβλεψης φορτίου για να επιτρέψουν την αναγνώριση των εποχικών μοτίβων που χαρακτηρίζουν τη ζήτηση ενέργειας. Τα δεδομένα αυτά κατασκευάστηκαν με τη χρήση τριών ομάδων χαρακτηριστικών: την ώρα, την ημέρα της εβδομάδας και τον μήνα του έτους

Επιπρόσθετα, για την αναπαράσταση της ημέρας της εβδομάδας χρησιμοποιήθηκαν επτά δυαδικές μεταβλητές, καθεμία από τις οποίες αντιστοιχεί σε μία ημέρα της εβδομάδας. Κάθε μεταβλητή λαμβάνει την τιμή 1 αν η αντίστοιχη ημέρα της εβδομάδας αντιστοιχούσε στην ημερομηνία της πρόβλεψης και 0 διαφορετικά. Με αυτόν τον τρόπο, το μοντέλο μπορούσε να αναγνωρίσει τις διαφορές στη ζήτηση ενέργειας μεταξύ των διάφορων ημερών της εβδομάδας. Αντίστοιχα, για την αναπαράσταση του μήνα του έτους χρησιμοποιήθηκαν δώδεκα δυαδικές μεταβλητές, μία για κάθε μήνα του έτους. Όπως και στην περίπτωση των ημερών της εβδομάδας, κάθε μεταβλητή λαμβάνει την τιμή 1 αν ο μήνας αντιστοιχεί στην ημερομηνία της πρόβλεψης και 0 διαφορετικά. Αυτή η προσέγγιση επέτρεψε στο μοντέλο να λαμβάνει υπόψη του την εποχικότητα της ζήτησης, καθώς ορισμένοι μήνες εμφανίζουν αυξημένες ενεργειακές ανάγκες λόγω θέρμανσης ή ψύξης.

Τέλος η μεταβλητή που σχετίζεται με τα εξωγενή δεδομένα χώρας συγκροτήθηκε χρησιμοποιώντας μια κωδικοποίηση one-hot encoding. Για την αναπαράσταση της χώρας στην οποία ανήκει κάθε δείγμα, δημιουργήθηκε ένας διανυσματικός χώρος με 24 δυαδικές μεταβλητές, όπου καθεμία αντιστοιχεί σε μία από τις 24

ευρωπαϊκές χώρες που περιλαμβάνονται στο σύνολο των δεδομένων. Σε κάθε δείγμα, μόνο η μεταβλητή που αντιστοιχεί στη χώρα προέλευσης λαμβάνει την τιμή 1, ενώ όλες οι υπόλοιπες λαμβάνουν την τιμή 0. Αυτή η προσέγγιση επιτρέπει στο μοντέλο να αναγνωρίζει την προέλευση κάθε δείγματος, διασφαλίζοντας ότι η πληροφορία της χώρας ενσωματώνεται στο μοντέλο χωρίς να εισάγεται άμεση σχέση τάξης ή μεγέθους μεταξύ των χωρών. Καθώς το προφίλ ζήτησης ενέργειας διαφέρει σημαντικά μεταξύ των χωρών λόγω γεωγραφικών, κλιματικών και κοινωνικών παραγόντων η ενσωμάτωση εξωγενών δεδομένων της χώρας στο μοντέλο μοιάζει είναι εύλογη στρατηγική.

5.3 Μετρικές πειραμάτων

Για την αξιολόγηση των αρχιτεκτονικών νευρωνικών δικτύων που αναπτύχθηκαν αναφορικά με την ενσωμάτωση εξωγενών μεταβλητών σε πυκνό δίκτυο για την εφαρμογή της πρόβλεψης ροής φορτίου επόμενης ημέρας, επιλέχθηκαν δύο μετρικές το Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (MAPE) και η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (RMSE). Αυτές οι μετρικές είναι ευρέως διαδεδομένες γενικότερα σε εφαρμογές πρόβλεψης χρονοσειρών και συμπληρωματικά προσφέρουν μια πλήρη οπτική στην απόδοση των μοντέλων προβλέψεων.

Στην μία περίπτωση το MAPE υπολογίζει ως ποσοστό τη μέση απόλυτη απόκλιση μεταξύ προβλεπόμενων και πραγματικών τιμών για κάθε σημείο. Το πλεονέκτημά του έγκειται στο γεγονός ότι προσφέρει μια εύκολα κατανοητή πληροφορία καθώς το σφάλμα, λόγω της μετατροπής σε ποσοστιαίο, γίνεται συγκρίσιμο μεταξύ διαφορετικών συνόλων δεδομένων ή μοντέλων, ανεξάρτητα από την κλίμακα των δεδομένων. Έτσι επιτρέπεται η σύγκριση διαφορετικών μοντέλων πρόβλεψης. Ωστόσο, παρουσιάζει μια σημαντική αδυναμία καθώς επηρεάζεται έντονα όταν οι πραγματικές τιμές ενδέχεται να περιλαμβάνουν μηδενικές ή εξαιρετικά μικρές τιμές. Στις περιπτώσεις όπου οι πραγματικές τιμές είναι πολύ μικρές (κοντά στο μηδέν), το MAPE μπορεί να διογκωθεί υπερβολικά, εμφανίζοντας δυσανάλογα μεγάλες τιμές σφάλματος. Αυτό οδηγεί σε μια παραπλανητική εικόνα της ακρίβειας του μοντέλου, καθώς ένα σχετικά μικρό απόλυτο σφάλμα μπορεί να καταλήξει να εκφράζεται ως ένα πολύ υψηλό ποσοστό σφάλματος. Το αποτέλεσμα είναι η δημιουργία μιας λανθασμένης εντύπωσης χαμηλής απόδοσης του μοντέλου, ακόμη και όταν οι προβλέψεις του μοντέλου είναι ακριβείς.

Για αυτό το λόγο συνήθως η διαδικασία αξιολόγησης των μοντέλων συμπληρώνεται από το RMSE. Αυτό μετρά το απόλυτο σφάλμα, δίνοντας έμφαση σε μεγάλες αποκλίσεις λόγω της τετραγωνικής του φύσης. Έτσι δίνεται έμφαση κυρίως στις μεγάλες αποκλίσεις, οι οποίες και τιμωρούνται πιο αυστηρά. Αυτό είναι αρκετά χρήσιμο σε εφαρμογές όπως η πρόβλεψη ροής φορτίου, όπου μεγάλες αποκλίσεις μπορεί να προκαλέσουν λειτουργικά προβλήματα στο δίκτυο, επιβαρύνοντας τη σταθερότητα και την αξιοπιστία των υποδομών.

Ο συνδυασμός των δύο μετρικών εξασφαλίζει μια ολοκληρωμένη εικόνα της απόδοσης των μοντέλων. Το MAPE προσφέρει μια κατανόηση του σφάλματος σε

ποσοστιαία βάση, ιδανική για τη γενική αξιολόγηση της ακρίβειας των προβλέψεων, ενώ το RMSE δίνει έμφαση στη συνολική επίδραση των αποκλίσεων, διασφαλίζοντας ότι το μοντέλο αποφεύγει μεγάλες αποκλίσεις που θα μπορούσαν να επηρεάσουν τη σταθερότητα του δικτύου. Με αυτόν τον τρόπο, οι δύο μετρικές παρέχουν μια ισορροπημένη αποτίμηση της απόδοσης των μοντέλων, εξασφαλίζοντας τόσο τη γενική ακρίβεια όσο και τη σταθερότητα των προβλέψεων.

5.4 Παρουσίαση αποτελεσμάτων αναπαράστασης αρχικών συνθηκών

Αρχικά θα παρουσιαστούν τα αποτελέσματα της αναπαραγωγής του αρχικού νευρωνικού δικτύου. Σε αυτή τη περίπτωση στα δεδομένα δεν περιλαμβάνονται εξωγενείς μεταβλητές, παρά μόνο τα ιστορικά δεδομένα φορτίου όπως αυτά έχουν αναλυθεί παραπάνω. Τα αποτελέσματα που προέκυψαν μετά από εκπαίδευση με τις ίδιες συνθήκες και στα ίδια δεδομένα με την αρχική εργασία ήταν αρκετά κοντά. Έτσι οι τιμές που επετεύχθησαν είναι :

Μοντέλο	MAPE	RMSE
Δική μου	3,203	741,86
Αρχικής	2,788	358

Πίνακας 5.1 : Σφάλματα χωρίς χρήση εξωγενών μεταβλητών

Παρατηρείται πως κατά την αναπαραγωγή του πυκνού νευρωνικού δικτύου και την αξιολόγηση του στο ίδιο σετ δεδομένων δεν έφτασαν τα νούμερα ακριβώς σε εκείνα τα επίπεδα. Το MAPE είναι αυξημένο κατά 0,415 μονάδες όμως το RMSE είναι αυξημένο κατά 383,86, δηλαδή περίπου διπλάσιο από της αρχικής εργασίας. Παρότι τα νέα σφάλματα είναι μεγαλύτερα, ειδικά στο RMSE, τα αποτελέσματα κρίνονται ικανοποιητικά για τον σκοπό της αναπαραγωγής και θα χρησιμοποιηθούν ως σημείο αναφοράς για την αξιολόγηση των επόμενων πειραμάτων.

5.5 Παρουσίαση αποτελεσμάτων μορφών των εξωγενών μεταβλητών

Σε αυτή την ενότητα παρουσιάζονται τα αποτελέσματα των πειραμάτων που αξιολογούν την απόδοση διαφορετικών μορφών εξωγενών δεδομένων ως εισόδους στο σύστημα πρόβλεψης μεμονωμένα. Για κάθε μια κατηγορία εξωγενών μεταβλητών, μετεωρολογικές συνθήκες, ημέρες και χώρες, δοκιμάστηκαν εναλλακτικές μορφές αναπαράστασης, οι οποίες τροφοδοτήθηκαν ως είσοδοι σε μια εξωγενή ροή παράλληλη με τη ροή του κεντρικού νευρωνικού δικτύου. Η συγκεκριμένη ροή, αποτελείται από ένα απλό πυκνό γραμμικό στρώμα και αντιπροσωπεύει την πιο λιτή και απλή αρχιτεκτονική ενσωμάτωσης εξωγενών χαρακτηριστικών στο σύστημα.

Εξωγενείς μεταβλητές καιρικών συνθηκών

Αρχικά για την μεταβλητή των καιρικών συνθηκών πραγματοποιήθηκαν τρία πειράματα. Το πρώτο θεώρησε δεδομένα έξι διαστάσεων, τιμές ελάχιστης, μέγιστης και μέσης θερμοκρασίας όπως και τις αντίστοιχες αποκλίσεις τους. Το δεύτερο θεώρησε δεδομένα τριών διαστάσεων, μόνο τις τιμές θερμοκρασίας και τέλος το τρίτο πείραμα θεώρησε πάλι δεδομένα τριών διαστάσεων, μόνο τις αποκλίσεις των αντίστοιχων τιμών. Τα αποτελέσματα παρουσιάζονται στον επόμενο πίνακα.

Μορφή	MAPE	RMSE
Καιρός - αριθμητικές τιμές	3,201	742,16
Καιρός - τιμές απόκλισης	3,113	728,16
Καιρός – συνδυαστικά	3,086	724,71

Πίνακας 5.2 : Αποτελέσματα εξωγενών δεδομένων καιρού

Η εισαγωγή μετεωρολογικών μεταβλητών που αναπαρίστανται με περιορισμένη πληροφορία, δηλαδή μόνο αριθμητικές τιμές ή μόνο τιμές απόκλισης, φαίνεται να βελτιώνει μετρίως την ακρίβεια των προβλέψεων, ενώ το όφελος αυξάνεται καθώς ενσωματώνονται οι αριθμητικές τιμές και οι τιμές απόκλισης. Συγκεκριμένα, στο σενάριο μόνο των αριθμητικών τιμών, το MAPE μειώνεται οριακά κατά 0,06 %, ενώ το RMSE αυξάνεται κατά 0,04 %, γεγονός που δείχνει ελάχιστη επίδραση. Επιπρόσθετα, στο σενάριο των τιμών απόκλισης, παρατηρείται σημαντικότερη βελτίωση, με το MAPE να μειώνεται κατά 2,81 % και το RMSE κατά 1,85 %, υποδεικνύοντας ότι οι συγκεκριμένες μεταβλητές προσφέρουν μεγαλύτερη πληροφορία. Τέλος, στο σενάριο που περιλαμβάνονται όλες οι τιμές, όπου περιλαμβάνονται όλες οι μετεωρολογικές μεταβλητές, επιτυγχάνεται η μεγαλύτερη βελτίωση, με μείωση του MAPE κατά 3,65 % και του RMSE κατά 2,31 %. Συνολικά, τα αποτελέσματα δείχνουν ότι οι εξωγενείς καιρικές μεταβλητές συμβάλλουν στη μείωση τόσο του σχετικού (MAPE) όσο και του απόλυτου (RMSE) σφάλματος, καθιστώντας τις χρήσιμες για το τελικό μοντέλο.

Εξωγενείς μεταβλητές ημερολογιακής ημέρας

Συνεχίζοντας, τα αποτελέσματα που αφορούν τα πειράματα που έγιναν για την εύρεση της βέλτιστης μορφής της μεταβλητής που αφορά τις ημερολογιακές ημέρες παρατίθενται στον επόμενο πίνακα.

Μορφή	MAPE	RMSE
Ημέρες - One Hot	3,214	747,21
Ημέρες - Κυκλικές	3,199	746,52
Ημέρες – Αργίες και μη	3,222	749,48

Πίνακας 5.3 : Αποτελέσματα εξωγενών δεδομένων ημέρας

Η αξιολόγηση των διαφορετικών μορφών αναπαράστασης ημερολογιακών δεδομένων δείχνει μικρές αποκλίσεις στην απόδοση του μοντέλου. Η προσέγγιση κυκλικής μορφής μεταβλητών που αναπαριστά τις ημερολογιακές πληροφορίες κυκλικά, επιτυγχάνει τη χαμηλότερη τιμή σφάλματος MAPE και το μικρότερο RMSE, σημειώνοντας μείωση 0,12 % και 0,63 % αντίστοιχα σε σχέση με το σημείο αναφοράς. Ακόμη στην περίπτωση της one-hot κωδικοποίησης καταγράφεται ελαφρώς υψηλότερο σφάλμα, ενώ η κατηγοριοποίηση σε Σαββατοκύριακα και εργάσιμες ημέρες αυξάνει το σφάλμα κατά 0,59 % στο MAPE και 1,03 % στο RMSE. Συμπερασματικά, τα αποτελέσματα υποδεικνύουν ότι η κυκλική αναπαράσταση είναι η πιο αποδοτική επιλογή για την ενσωμάτωση ημερολογιακών πληροφοριών.

Εξωγενείς μεταβλητές χώρας

Τέλος, παρατίθενται τα αποτελέσματα των πειραμάτων αναφορικά με την εξωγενή μεταβλητή που σχετίζεται με την κωδικοποίηση ή και ομαδοποίηση του συνόλου των χωρών.

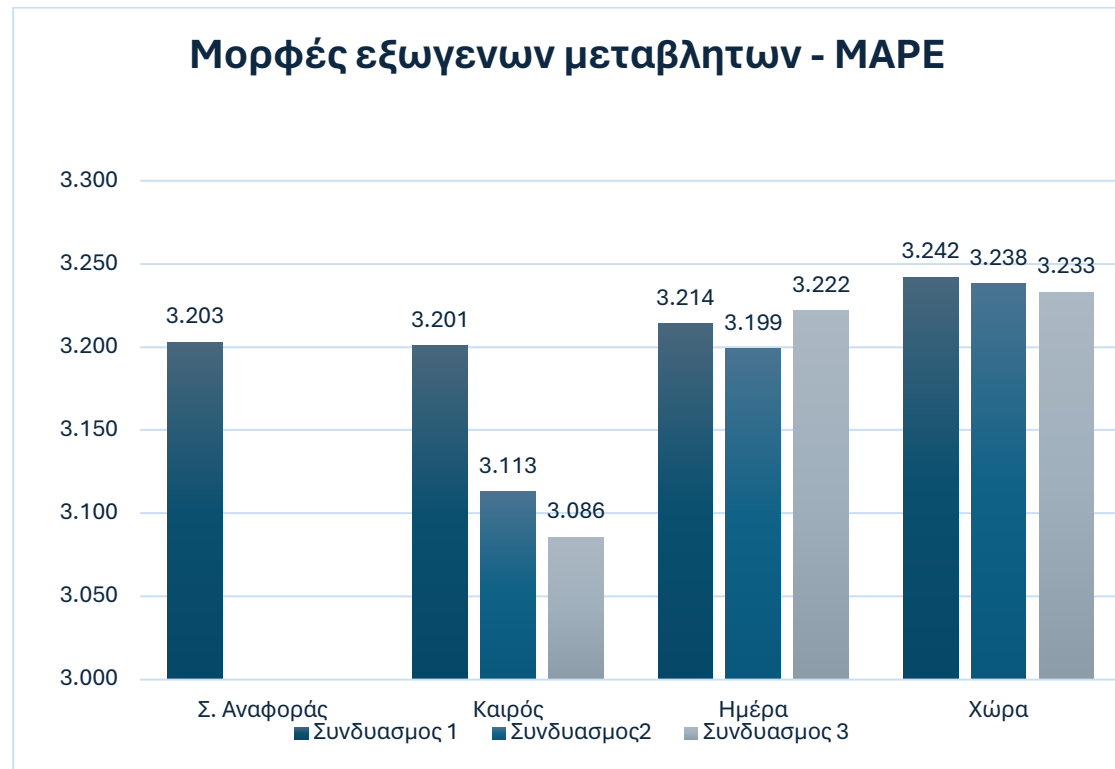
Μορφή	MAPE	RMSE
Χώρες – One Hot	3,242	749,88
Χώρες – Περιοχές	3,238	752,29
Χώρες – Περιοχές K-Means	3,233	759,19

Πίνακας 5.4 : Αποτελέσματα εξωγενών δεδομένων χώρας

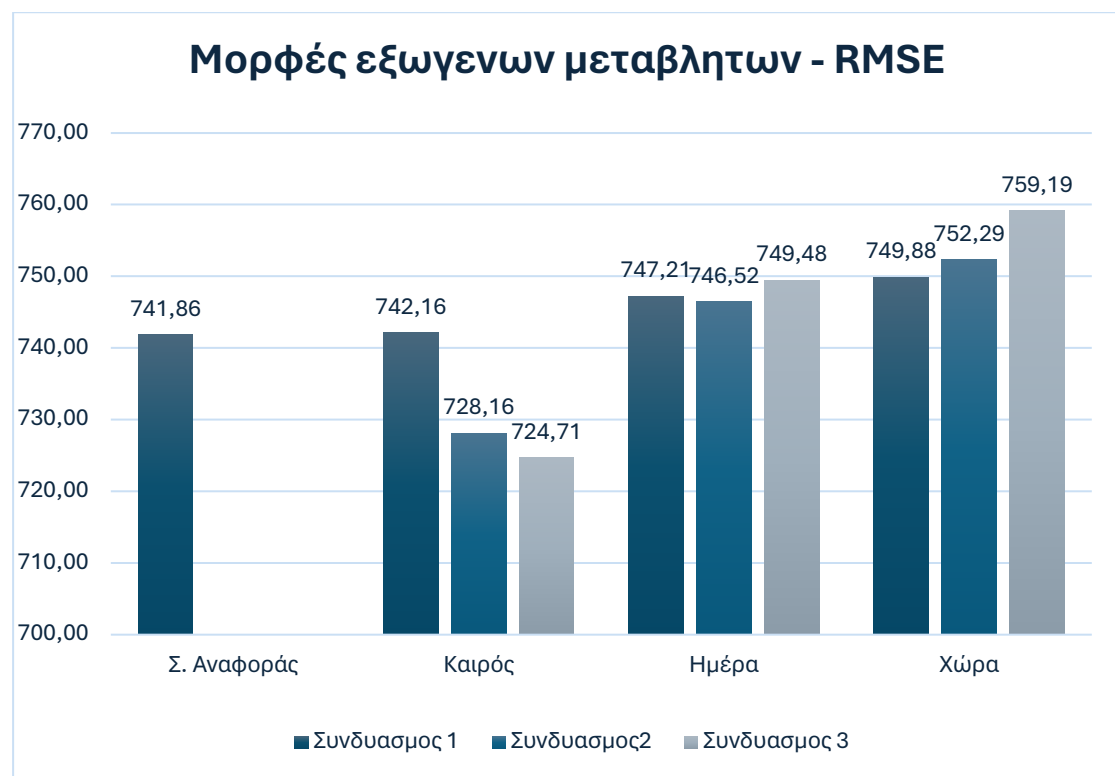
Η αξιολόγηση των διαφορετικών αναπαραστάσεων χωρικών δεδομένων δείχνει ότι η χρήση της μεθόδου ομαδοποίησης χωρών με βάση τον αλγόριθμο K-Means οδηγεί στο χαμηλότερο σφάλμα MAPE με τη μικρότερη αύξηση 0,94 % σε σχέση με το σημείο αναφοράς, αλλά στο υψηλότερο RMSE, σημειώνοντας αύξηση 2,34 %. Η απλή κωδικοποίηση των χωρών με αναπαράσταση one-hot παρουσιάζει τη μεγαλύτερη αύξηση στο MAPE, 1,22 %, αλλά περιορισμένη αύξηση στο RMSE, 1,08 %, υποδεικνύοντας ότι διατηρεί καλύτερη σταθερότητα. Τέλος ο διαμοιρασμός των χωρών σε περιοχές με βάση τα ενεργειακά τους προφίλ εμφανίζει ενδιάμεση συμπεριφορά, με αύξηση 1,09 % στο MAPE και 1,41 % στο RMSE. Συμπερασματικά, τα αποτελέσματα δείχνουν ότι οι χωρικές πληροφορίες από μόνες τους δεν βελτιώνουν την ακρίβεια του μοντέλου, ενώ η επιλογή του K-Means οδηγεί σε καλύτερη σχετική ακρίβεια εις βάρος του απόλυτου σφάλματος (RMSE).

Συμπερασματικά με βάση τα αποτελέσματα, ως βέλτιστες μορφές εξωγενών δεδομένων επιλέγονται η κυκλική αναπαράσταση της μεταβλητής των ημερολογιακών ημερών για τις ημερολογιακές πληροφορίες, καθώς πέτυχε το χαμηλότερο MAPE και RMSE, και η αναπαράσταση των καιρικών δεδομένων με αριθμητικές τιμές και τις αποκλίσεις τους, που παρουσίασε τη μεγαλύτερη βελτίωση και στους δύο δείκτες σφάλματος. Για τα εξωγενή δεδομένα των χωρών, επιλέγεται ο διαχωρισμός με βάση τον αλγόριθμο K-Means καθώς

προσφέρει τη χαμηλότερη τιμή MAPE, παρά την ελαφρώς αυξημένη τιμή RMSE, αντανακλώντας την καλύτερη σχετική ακρίβεια του μοντέλου.



Σχήμα 5.1 : Αποτελέσματα MAPE – μορφές εξωγενών μεταβλητών



Σχήμα 5.2 : Αποτελέσματα RMSE – μορφές εξωγενών μεταβλητών

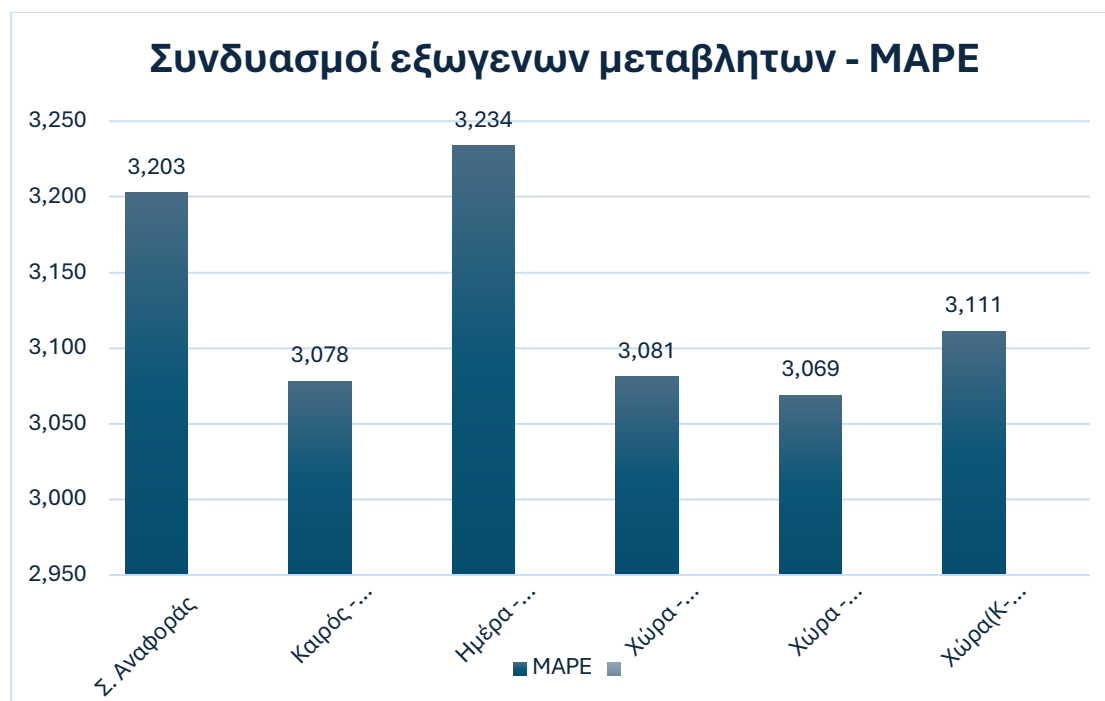
5.6 Παρουσίαση αποτελεσμάτων συνδυασμών των εξωγενών μεταβλητών

Σε αυτό το στάδιο αξιολογείται η χρήση συνδυασμών εξωγενών μεταβλητών. Κάθε μια τροφοδοτείται σε ανεξάρτητη παράλληλη ροή, με αυτή των ιστορικών. Αναλόγως των τιμών των μετρικών υπολογίζεται ποιος συνδυασμός βελτιώνει σε μεγαλύτερο βαθμό την απόδοση των προβλέψεων. Σημειώνεται πως κάθε μεταβλητή αναπαρίσταται με την βέλτιστη μορφή η οποία έχει εντοπιστεί στην προηγούμενη ενότητα. Παρακάτω παρουσιάζονται τα αποτελέσματα :

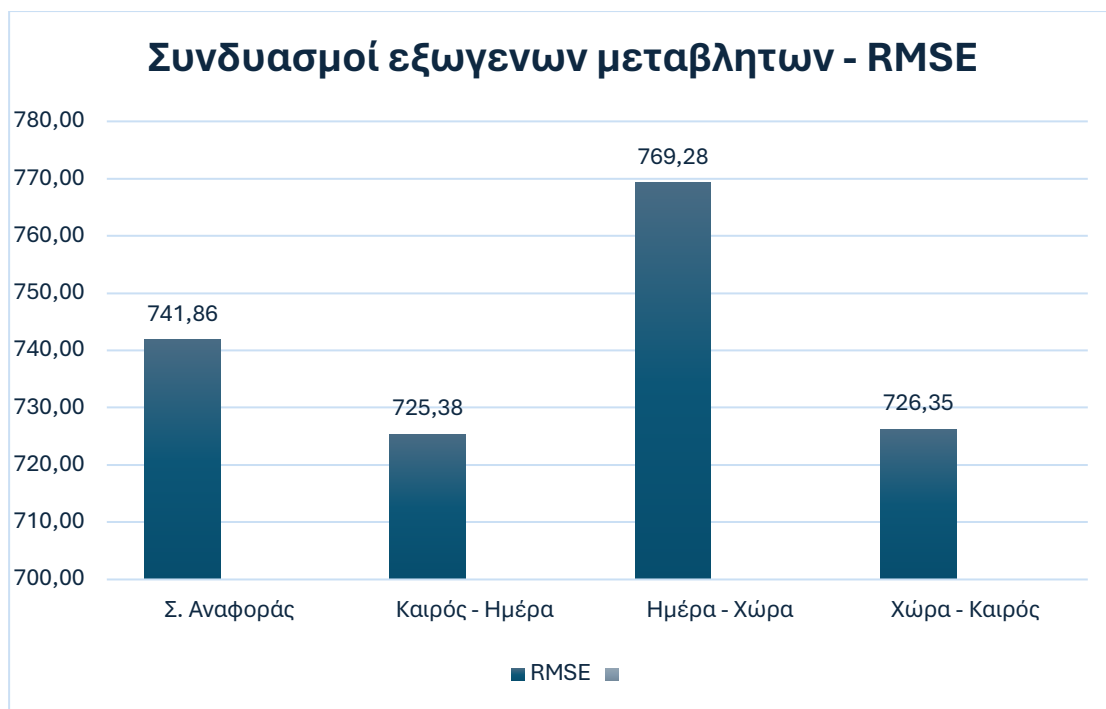
Συνδυασμός	MAPE	RMSE
Καιρός – Ημέρα	3,078	725,38
Ημέρα – Χώρα	3,234	769,28
Χώρα - Καιρός	3,081	726,35
Καιρός – Ημέρα - Χώρα	3,069	711,52

Πίνακας 5.5 : Αποτελέσματα συνδυασμών εξωγενών μεταβλητών

Τα αποτελέσματα δείχνουν ότι ο συνδυασμός καιρικών και ημερολογιακών μεταβλητών βελτιώνει σημαντικά την πρόβλεψη, μειώνοντας το MAPE κατά 3,90% και το RMSE κατά 2,22%. Παρόμοια απόδοση καταγράφεται με τις μεταβλητές χώρας και καιρού, με μείωση 3,81% στο MAPE και 2,09% στο RMSE. Αντίθετα, ο συνδυασμός ημερολογιακών μεταβλητών και χώρας επιδεινώνει τις μετρικές, αυξάνοντας το MAPE κατά 0,97% και το RMSE κατά 3,70%. Η χρήση όλων των εξωγενών μεταβλητών αποτελεί την βέλτιστη επιλογή και επιτυγχάνει τη μεγαλύτερη βελτίωση, με μείωση 4,18% στο MAPE και 4,09% στο RMSE.



Σχήμα 5.3 : Αποτελέσματα MAPE – συνδυασμοί εξωγενών μεταβλητών



Σχήμα 5.4 : Αποτελέσματα RMSE – συνδυασμοί εξωγενών μεταβλητών

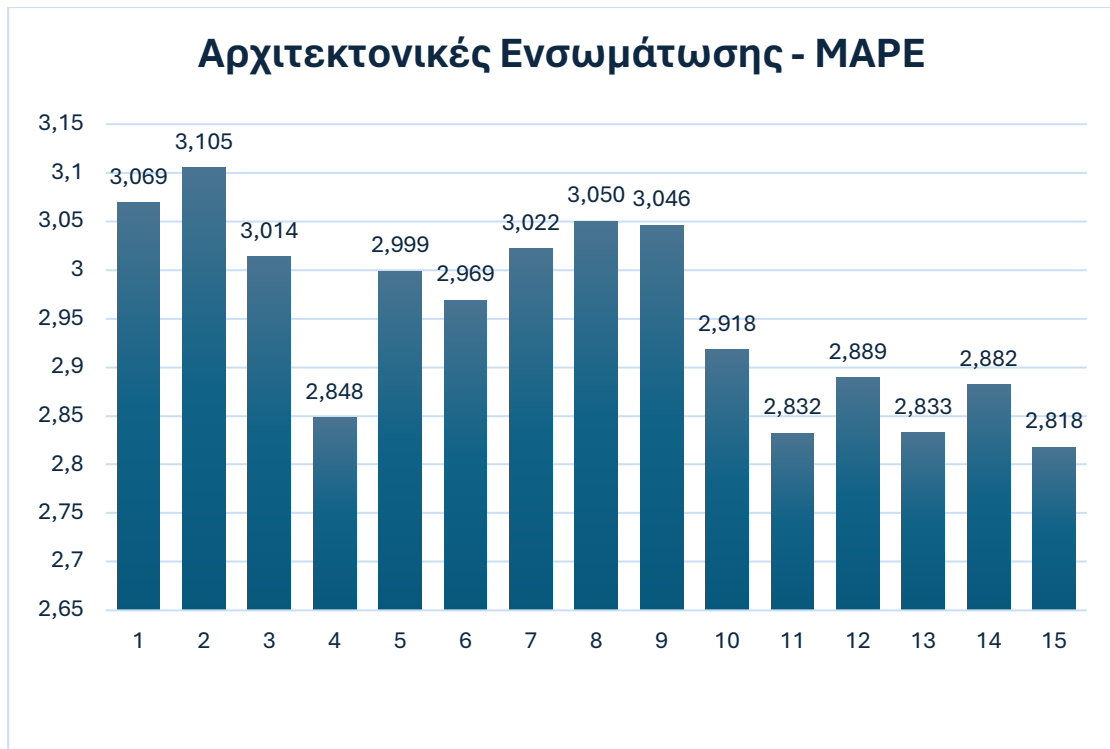
5.7 Παρουσίαση αποτελεσμάτων ενσωμάτωσης των εξωγενών μεταβλητών

Σε αυτή την ενότητα δοκιμάζονται και αξιολογούνται διαφορετικές αρχιτεκτονικές ενσωμάτωσης του βέλτιστου συνδυασμού καθώς και μορφής των εξωγενών μεταβλητών. Θα γίνει λοιπόν χρήση και των τριών μεταβλητών και συγκεκριμένα η μεταβλητή του καιρού με χρήση αριθμητικών τιμών και τιμών απόκλισης, η μεταβλητή των χωρών με ομαδοποίηση κατά K-Means και τέλος η μεταβλητή των ημερών με κυκλική περιγραφή.

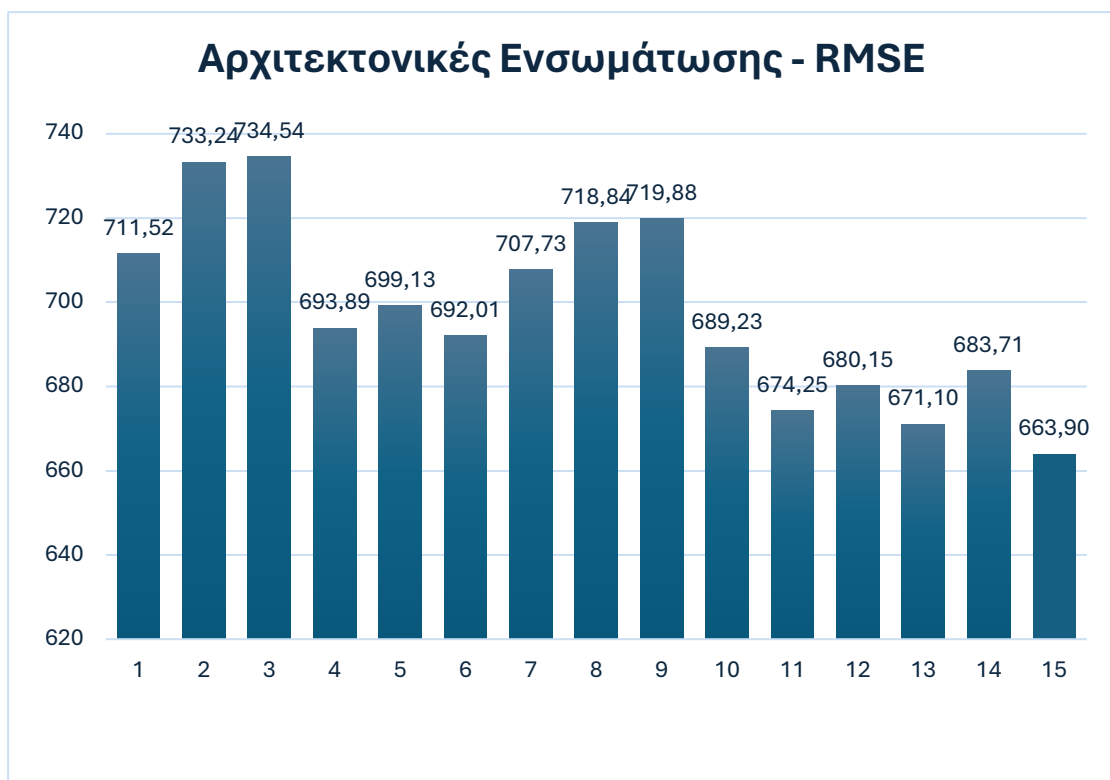
Δοκιμάστηκαν ποικίλοι συνδυασμοί στη φιλοσοφία του αρχικού πλήρους διασυνδεδεμένου νευρωνικού δικτύου με στόχο την διατήρηση του υπολογιστικού κόστους σε χαμηλά επίπεδα.

ΑΑ	ΑΡΧΙΤΕΚΤΟΝΙΚΗ	MAPE	RMSE
1	Εξωγενείς ως ανεξάρτητες παράλληλες ροές (Γραμμικό στρώμα) (+ / στοιχείο)	3,069	711,52
2	Εξωγενείς συνενωμένες ως είσοδοι σε μια ροή (Γραμμικό στρώμα) (+ / στοιχείο)	3,105	733,24
3	Εξωγενείς συνενωμένες ως είσοδοι σε μια ροή (FCNN)	3,014	734,54
4	Εξωγενείς και ιστορικές μεταβλητές συνενωμένες ως είσοδοι σε μια ενιαία ροή	2,848	693,89
5	Εξωγενείς ως ανεξάρτητες παράλληλες ροές (Κρυφό + Γραμμικό στρώματα) (+ / στοιχείο)	2,999	699,13
6	Εξωγενείς ως ανεξάρτητες παράλληλες ροές (Δύο Κρυφά + Γραμμικό στρώματα) (+ / στοιχείο)	2,969	692,01
7	Εξωγενείς ως ανεξάρτητες παράλληλες ροές (FCNN) (+ / στοιχείο)	3,022	707,73
8	Εξωγενείς ως ανεξάρτητες παράλληλες ροές (μέρος FCNN) (+ / στοιχείο)	3,050	718,84
9	Εξωγενείς ως ανεξάρτητες παράλληλες ροές (Γραμμικό στρώμα) (Συνένωση + γραμμικό)	3,046	719,88
10	Εξωγενείς ως ανεξάρτητες παράλληλες ροές (Γραμμικό στρώμα) (Συνένωση + κρυφό + γραμμικό)	2,918	689,23
11	Εξωγενείς ως ανεξάρτητες ροές (γραμμικό στρώμα) που τελικά συνενώνονται με τις ιστορικές και εισάγονται σε στρώματα. Χρήση skip connection	2,832	674,25
12	Εξωγενείς ως ανεξάρτητες ροές (γραμμικό στρώμα) που τελικά συνενώνονται με τις ιστορικές και εισάγονται σε στρώματα. Μη χρήση skip connection	2,889	680,15
13	Εξωγενείς ως ανεξάρτητες ροές (κρυφό + γραμμικό στρώμα) που τελικά συνενώνονται με τις ιστορικές και εισάγονται σε στρώματα. Χρήση skip connection	2,833	671,10
14	Εξωγενείς ως ανεξάρτητες ροές (2 κρυφά + γραμμικό στρώμα) που τελικά συνενώνονται με τις ιστορικές και εισάγονται σε στρώματα. Χρήση skip connection	2,882	683,71
15	Εξωγενείς ως ανεξάρτητες ροές (FCNN) που τελικά συνενώνονται με τις ιστορικές και εισάγονται σε στρώματα. Χρήση skip connection	2,818	663,90

Πίνακας 5.6 : Αποτελέσματα πειραμάτων ενσωμάτωσης των εξωγενών



Σχήμα 5.5 : Αποτελέσματα MAPE – αρχιτεκτονικές ενσωμάτωσης εξωγενών



Σχήμα 5.6 : Αποτελέσματα RMSE – αρχιτεκτονικές ενσωμάτωσης εξωγενών

Ως μετρική – σημείο αναφοράς για την αξιολόγηση των αρχιτεκτονικών ορίζονται οι τιμές MAPE 3,203 και RMSE 741,86, οι οποίες αντιστοιχούν στην απόδοση του συστήματος προβλέψεων χωρίς τη χρήση εξωγενών μεταβλητών, όπως έγινε η ανακατασκευή του. Γενικά η αρχιτεκτονική 15 ξεχωρίζει ως η καλύτερη, επιτυγχάνοντας μείωση κατά 12.02% στο MAPE και μείωση κατά 10.51% στο RMSE. Αυτή η αρχιτεκτονική είναι η πιο σύνθετη από όλες, καθώς οι τέσσερις ροές αποτελούνται από το πλήρες διασυνδεδεμένο μοντέλο και, αφού συνενωθούν, τροφοδοτούνται ως ενιαία είσοδος σε τέσσερα στρώματα. Στη συνέχεια έπεται η αρχιτεκτονική 11 με μείωση 11.58% στο MAPE και 9.11% στο RMSE, η οποία κατατάσσεται δεύτερη καλύτερη. Η αρχιτεκτονική 13, με μείωση 11.55% στο MAPE και 9.54% στο RMSE, κατατάσσεται τρίτη και αποτελείται από εξωγενείς ροές με ένα γραμμικό στρώμα, οι οποίες συνενώνονται με την κύρια ροή και τροφοδοτούνται σε ένα στρώμα με ReLU και ένα γραμμικό, αποδεικνύοντας ότι η αυξημένη χρήση στρωμάτων μετά τη συνένωση, παρά την φτωχή επεξεργασία σε κάθε ανεξάρτητη εξωγενή ροή, μπορεί να οδηγήσει σε βελτιωμένη απόδοση, αν και υπολείπεται ελαφρώς της αρχιτεκτονικής 15. Ακόμη η αρχιτεκτονική 12 με MAPE που μεταβλήθηκε κατά 9.80% και RMSE κατά 8.32% παρουσιάζει και αυτή καλή ακρίβεια παρά την έλλειψη skip connection.

Επιπρόσθετα, αρχιτεκτονικές όπως η 4, η 5 και η 6 παρουσιάζουν μέτρια απόδοση, με ποσοστιαία διαφορά MAPE που κυμαίνεται από 11.08% έως 6.37% και RMSE από 6.72% έως 5.76%. Παρά την αξιόλογη απόδοσή τους, δεν φτάνουν την ακρίβεια των κορυφαίων αρχιτεκτονικών.

Αντίθετα, οι αρχιτεκτονικές 1, 2, 3, 7, 8 και 9 εμφανίζουν τις χειρότερες επιδόσεις, με MAPE που μεταβλήθηκε από 5.90% έως 3.06% και RMSE από 4.60% έως 0.99%, υποδεικνύοντας ότι οι απλούστερες δομές ή η χρήση στοιχείων αντί για συνδυασμένες ροές δεν μπορούν να αποτυπώσουν αποτελεσματικά τις πολύπλοκες σχέσεις στα δεδομένα. Η αρχιτεκτονική 2, με μείωση MAPE κατά 3.06% και μείωση RMSE κατά 1.16%, είναι από τις χειρότερες, πιθανώς λόγω της μειωμένης χωρητικότητας του μικρού δικτύου που χρησιμοποιεί.

Συμπερασματικά, η αρχιτεκτονική 15 αποτελεί την καλύτερη επιλογή και πετυχαίνει βελτίωση κατά 12 % ως προς MAPE και 10.5% ως προς RMSE. Στη συνέχεια ακολουθούν οι 11, 12, 13, 14 χάρη στη χρήση αποτελεσματικού συνδυασμού ροών. Μια μέση τιμή επιτυγχάνουν οι αρχιτεκτονικές 10, 4, 5, 6 και τέλος οι αρχιτεκτονικές 1, 2, 3, 7, 8, 9 δεν παρουσιάζουν ενθαρρυντικά αποτελέσματα.

ΚΕΦΑΛΑΙΟ 6 – ΣΧΟΛΙΑΣΜΟΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

6.1 Χρήση εξωγενών μεταβλητών

Η ανάλυση των αποτελεσμάτων των πειραμάτων που παρουσιάστηκαν στις ενότητες 6.5 και 6.6 παρέχει σημαντικές πληροφορίες για την επίδραση των εξωγενών μεταβλητών στην ακρίβεια των προβλέψεων φορτίου της επόμενης ημέρας, καθώς και για τη βέλτιστη μορφή και τον συνδυασμό τους. Ο σχολιασμός που ακολουθεί επικεντρώνεται στη συμβολή κάθε εξωγενούς μεταβλητής, στην αποτελεσματικότητα των διαφορετικών μορφών αναπαράστασης, στους συνδυασμούς μεταβλητών, καθώς και στη γενικότερη σημασία των ευρημάτων για την έρευνα και την πρακτική εφαρμογή.

Τα αποτελέσματα των πειραμάτων καταδεικνύουν ότι οι εξωγενείς μεταβλητές αποτελούν κρίσιμο στοιχείο για την ενίσχυση της προβλεπτικής ικανότητας του νευρωνικού δικτύου. Μεταξύ των τριών κατηγοριών εξωγενών μεταβλητών (καιρικές συνθήκες, ημερολογιακές ημέρες, χώρες), οι καιρικές συνθήκες αποδείχθηκαν οι πιο σημαντικές, με τη μορφή που συνδυάζει τιμές μέγιστης, ελάχιστης και μέσης θερμοκρασίας με τις αποκλίσεις τους από τη βασική θερμοκρασία των 18°C να επιτυγχάνει τη μεγαλύτερη μείωση των μετρικών σφαλμάτων (MAPE -3,65% και RMSE -2,31%). Αυτό το εύρημα είναι λογικό, καθώς η θερμοκρασία επηρεάζει άμεσα τη ζήτηση ενέργειας μέσω της χρήσης συστημάτων θέρμανσης και κλιματισμού. Επιπλέον, η συμπερίληψη των αποκλίσεων από τη μέση θερμοκρασία ενισχύει την ικανότητα του μοντέλου να συλλαμβάνει τη μη γραμμική σχέση μεταξύ θερμοκρασίας και κατανάλωσης ενέργειας, καθώς αυτές οι αποκλίσεις αποτελούν πιο άμεσο δείκτη των ενεργειακών αναγκών.

Συνεχίζοντας, οι ημερολογιακές ημέρες, με τη βέλτιστη κυκλική αναπαράσταση, προσφέρουν μικρότερη αλλά αξιοσημείωτη βελτίωση. Η κυκλική αναπαράσταση, που χρησιμοποιεί ημίτονα και συνημίτονα για την κωδικοποίηση της περιοδικότητας των ημερών και των μηνών, αποδείχθηκε πιο αποτελεσματική από την παραδοσιακή boolean κωδικοποίηση ή την απλοποιημένη κατηγοριοποίηση σε εργάσιμες ημέρες και Σαββατοκύριακα. Αυτό υποδηλώνει ότι η περιοδική φύση των ημερολογιακών δεδομένων, όπως οι εβδομαδιαίοι και μηνιαίοι κύκλοι, είναι σημαντική για την αποτύπωση των εποχικών μοτίβων κατανάλωσης ενέργειας. Ωστόσο, η περιορισμένη συμβολή των ημερολογιακών μεταβλητών σε σχέση με τις καιρικές συνθήκες μπορεί να οφείλεται στο γεγονός ότι τα ιστορικά δεδομένα φορτίου περιλαμβάνουν ήδη πληροφορίες για τις χρονικές τάσεις, μειώνοντας τη συμπληρωματική αξία των ημερολογιακών δεδομένων.

Οι μεταβλητές που σχετίζονται με τις χώρες, με τη βέλτιστη μορφή να είναι η ομαδοποίηση K-Means, δεν βελτίωσαν την ακρίβεια του μοντέλου όταν

χρησιμοποιήθηκαν μεμονωμένα. Αν και το αποτέλεσμα αυτό φαίνεται παράδοξο, είναι λογικό, καθώς η πληροφορία της χώρας από μόνη της δεν παρέχει δυναμικά χαρακτηριστικά που σχετίζονται άμεσα με τις ωριαίες διακυμάνσεις της ζήτησης ενέργειας. Η ομαδοποίηση K-Means, που βασίζεται στην ομοιότητα των ιστορικών δεδομένων φορτίου, προσφέρει καλύτερη σχετική ακρίβεια σε σχέση με την one-hot κωδικοποίηση ή την ομαδοποίηση βάσει ενεργειακού προφίλ, πιθανώς επειδή συλλαμβάνει πιο αποτελεσματικά τα πρότυπα κατανάλωσης, καθώς βασίζεται στην καθαρή ομοιότητα των δεδομένων κατανάλωσης κάθε χώρας. Ωστόσο, η αύξηση του RMSE κατά την ομαδοποίηση υποδεικνύει ότι η μείωση των διαστάσεων μέσω ομαδοποίησης μπορεί να οδηγεί σε απώλεια λεπτομερών πληροφοριών, περιορίζοντας τη συνολική χρησιμότητα της μεταβλητής.

Εξίσου σημαντικός είναι ο τομέας του συνδυασμού των εξωγενών μεταβλητών, καθώς η αξιολόγηση των συνδυασμών επιβεβαιώνει ότι η συνδυαστική χρήση όλων των μεταβλητών οδηγεί στη μεγαλύτερη βελτίωση της ακρίβειας, με μείωση του MAPE κατά 4,18% και του RMSE κατά 4,09%. Αυτό το αποτέλεσμα υπογραμμίζει τη συμπληρωματική φύση των εξωγενών μεταβλητών, όπου καθεμία συμβάλλει μοναδικές πληροφορίες που ενισχύουν τη συνολική κατανόηση των παραγόντων που επηρεάζουν τη ζήτηση ενέργειας. Για παράδειγμα, οι καιρικές συνθήκες παρέχουν πληροφορίες για περιβαλλοντικές επιδράσεις, οι ημερολογιακές ημέρες αποτυπώνουν κοινωνικές και εποχικές συμπεριφορές, ενώ οι μεταβλητές της χώρας προσθέτουν γεωγραφικό και ενεργειακό πλαίσιο. Αναλυτικότερα, οι συνδυασμοί που περιλαμβάνουν καιρικές μεταβλητές παρουσιάζουν σημαντική βελτίωση, επιβεβαιώνοντας την κυριαρχία της θερμοκρασίας ως βασικού παράγοντα. Αντίθετα, ο συνδυασμός ημερολογιακής ημέρας και χώρας χωρίς καιρικές μεταβλητές επιδεινώνει την απόδοση, γεγονός που υποδηλώνει ότι αυτές οι μεταβλητές από μόνες τους δεν επαρκούν για να συλλάβουν τις πολύπλοκες δυναμικές της ζήτησης ενέργειας. Η επιτυχία του συνδυασμού όλων των μεταβλητών μπορεί να αποδοθεί στην ικανότητά του να καλύπτει ένα ευρύτερο φάσμα επιδράσεων, επιτρέποντας στο νευρωνικό δίκτυο να μαθαίνει πιο σύνθετες και μη γραμμικές σχέσεις μεταξύ των δεδομένων. Επιπρόσθετα, παρόλο που η μεταβλητή της χώρας δεν φαίνεται χρήσιμη μεμονωμένα, ο συνδυασμός της με τις υπόλοιπες βελτιώνει τις μετρικές αξιολόγησης, γεγονός που υποδεικνύει τη σωστή της αξιοποίησης.

Συμπερασματικά, από ερευνητικής πλευράς, η μελέτη υπογραμμίζει τη σημασία της προσεκτικής επιλογής του είδους, της αναπαράστασης και του συνδυασμού των εξωγενών μεταβλητών. Η ανώτερη απόδοση της μορφής των καιρικών δεδομένων με τιμές και αποκλίσεις υποδεικνύει ότι οι μεταβλητές που σχετίζονται άμεσα με φυσικούς ή περιβαλλοντικούς παράγοντες πρέπει να κωδικοποιούνται με τρόπο που να αποτυπώνει τη μη γραμμική τους επίδραση. Ομοίως, η επιτυχία της κυκλικής αναπαράστασης για τις ημερολογιακές ημέρες υποστηρίζει τη χρήση προηγμένων τεχνικών κωδικοποίησης για δεδομένα με περιοδική φύση, αντί για παραδοσιακές μεθόδους, όπως η one-hot κωδικοποίηση. Αντίθετα, η περιορισμένη συμβολή των μεταβλητών της χώρας όταν χρησιμοποιούνται μεμονωμένα, αλλά η σημαντική τους επίδραση όταν συνδυάζονται με άλλες

μεταβλητές, υποδηλώνει ότι οι γεωγραφικές πληροφορίες λειτουργούν κυρίως ως συμπληρωματικό πλαίσιο. Το εύρημα αυτό συνάδει με τη βιβλιογραφία, όπως η μελέτη των Zolfaghari και Golabi (2022), που τονίζει ότι οι μετεωρολογικοί παράγοντες είναι οι πιο επιδραστικοί, ενώ άλλοι παράγοντες, όπως οι γεωγραφικοί, ενισχύουν την απόδοση μόνο όταν συνδυάζονται με πιο δυναμικές μεταβλητές.

Παρά τα θετικά αποτελέσματα της έρευνας, αξίζει να επισημανθούν ορισμένοι περιορισμοί αναφορικά με τα εξωγενή δεδομένα. Πρώτον, η περιορισμένη συμβολή των μεταβλητών της χώρας υποδηλώνει ότι η γεωγραφική πληροφορία ενδέχεται να μην προσφέρει νέες πληροφορίες, καθώς τα ιστορικά δεδομένα φορτίου περιλαμβάνουν ήδη σχετικές τάσεις. Δεύτερον, η επιτυχία της κυκλικής αναπαράστασης για τις ημερολογιακές ημέρες μπορεί να εξαρτάται από την ποιότητα των δεδομένων και την απουσία θορύβου, καθώς η περιοδικότητα μπορεί να επηρεάζεται από μη τυπικές συνθήκες, όπως εθνικές αργίες που δεν κωδικοποιήθηκαν επαρκώς. Τέλος, η εστίαση σε μετεωρολογικά δεδομένα μόνο από τις πρωτεύουσες ενδέχεται να περιορίζει την ακρίβεια, καθώς οι καιρικές συνθήκες ποικίλλουν εντός κάθε χώρας. Η χρήση δεδομένων από περισσότερες γεωγραφικές τοποθεσίες θα μπορούσε να ενισχύσει περαιτέρω την απόδοση.

Συνολικά, τα αποτελέσματα των πειραμάτων επιβεβαιώνουν ότι οι εξωγενείς μεταβλητές αποτελούν κρίσιμο στοιχείο για την ενίσχυση της απόδοσης των προβλέψεων φορτίου της επόμενης ημέρας. Η βέλτιστη μορφή κάθε μεταβλητής και η συνδυαστική τους χρήση οδηγούν σε σημαντική μείωση των σφαλμάτων πρόβλεψης, ενισχύοντας την ικανότητα του νευρωνικού δικτύου να αποτυπώνει τις πολύπλοκες δυναμικές της ζήτησης ενέργειας και τις εξαρτήσεις της. Επιπρόσθετα, επιβεβαιώνεται ότι οι καιρικές συνθήκες αποτελούν τον πιο σημαντικό παράγοντα που επηρεάζει την κατανάλωση ενέργειας, ενώ οι ημερολογιακές ημέρες και οι μεταβλητές της χώρας, όταν χρησιμοποιούνται συνδυαστικά, παρέχουν τις σημαντικότερες πληροφορίες που χρειάζεται το σύστημα για να επιτύχει αποδοτικές προβλέψεις.

6.2 Αρχιτεκτονικές ενσωμάτωσης

Η ανάλυση της γενικής εικόνας αποδεικνύει ότι το κλειδί για ακριβέστερες προβλέψεις έγκειται στην αυτόνομη και παράλληλη επεξεργασία κάθε εξωγενούς μεταβλητής, τόσο μεταξύ τους όσο και σε συνδυασμό με τη ροή των ιστορικών δεδομένων. Στη συνέχεια, οι έξοδοι από αυτές τις τέσσερις ανεξάρτητες ροές συνενώνονται και τροφοδοτούνται ως είσοδοι σε μια σειρά διαδοχικών πυκνών νευρωνικών στρωμάτων. Αυτό το μοτίβο, το οποίο ακολουθείται από την δέκατη αρχιτεκτονική και εξής, υποδεικνύει ότι η απλή πρόσθεση στοιχείων κατά στοιχείο όλων των μεταβλητών δεν είναι αποτελεσματική σε σύγκριση με τη συνένωση και την περαιτέρω επεξεργασία τους σε επόμενα στρώματα. Τα διαδοχικά στρώματα επιτρέπουν στο νευρωνικό δίκτυο να αναπτύσσει πολύπλοκες, μη γραμμικές σχέσεις μεταξύ των εξωγενών και ιστορικών δεδομένων, διευκολύνοντας την

αναγνώριση μοτίβων που δεν θα μπορούσαν να εντοπιστούν μόνο με την πρόσθεση, χωρίς περαιτέρω επεξεργασία.

Επιπρόσθετα, είναι σαφές ότι η ομαδοποίηση εξωγενών και ιστορικών δεδομένων πριν από οποιαδήποτε επεξεργασία δεν ευνοεί την απόδοση του δικτύου. Αντιθέτως φαίνεται πως η αυτόνομη και ανεξάρτητη επεξεργασία των εξωγενών μεταβλητών σε παράλληλες ροές πριν από τη συνένωσή τους σε ένα νευρωνικό δίκτυο αποτελεί μια αποτελεσματική προσέγγιση που βελτιώνει σημαντικά την απόδοση του μοντέλου, ιδιαίτερα σε προβλήματα με πολύπλοκα και ετερογενή δεδομένα. Πρώτον, η διατήρηση της ανεξαρτησίας των χαρακτηριστικών κάθε μεταβλητής είναι σημαντική, καθώς οι εξωγενείς μεταβλητές συχνά περιέχουν μοναδικά μοτίβα που δεν σχετίζονται άμεσα μεταξύ τους ή με την κύρια μεταβλητή των ιστορικών δεδομένων. Έτσι η ανεξάρτητη επεξεργασία επιτρέπει την εξαγωγή εξειδικευμένων χαρακτηριστικών, όπως χρονικές εξαρτήσεις ή τοπικά μοτίβα, χωρίς παρεμβολές από άλλες πηγές. Δεύτερον, η ανεξάρτητη επεξεργασία μειώνει τον κίνδυνο υπερπροσαρμογής, καθώς αποτρέπει τη μάθηση ψευδών συσχετίσεων μεταξύ μεταβλητών, ιδιαίτερα σε θορυβώδη δεδομένα, επιτρέποντας στο δίκτυο να εστιάσει στις εγγενείς ιδιότητες κάθε μεταβλητής πριν τη συνένωση. Τέλος, η σχεδίαση των παράλληλων ροών ως ανεξάρτητες ενισχύει την ερμηνευσιμότητα του μοντέλου, καθώς κάθε ροή μπορεί να αναλυθεί ξεχωριστά, διευκολύνοντας τη διάγνωση προβλημάτων και την κατανόηση της συμβολής κάθε μεταβλητής στην τελική πρόβλεψη. Συμπερασματικά, η προσέγγιση αυτή εξασφαλίζει πιο ακριβείς και γενικεύσιμες προβλέψεις, καθιστώντας την ιδανική για εφαρμογές με ετερογενείς εξωγενείς μεταβλητές κάτι το οποίο επιβεβαιώνεται πλήρως και από τα πειράματα καθώς σε όσες περιπτώσεις τα εξωγενή δεδομένα συνενώθηκαν με τα ιστορικά ή ακόμη τα εξωγενή συνενώθηκαν μεταξύ τους χωρίς να υπάρξει πρώιμη επεξεργασία τα αποτελέσματα δεν ήταν ενθαρρυντικά.

Όσον αφορά τη χρήση συνδέσεων παράκαμψης (skip connections), είναι εμφανές ότι συμβάλλουν σημαντικά στη βελτίωση των προβλέψεων και στη μείωση των σφαλμάτων. Συγκρίνοντας τις αρχιτεκτονικές 10 και 11, οι οποίες είναι πανομοιότυπες με μοναδική διαφορά την απουσία skip connection στην πρώτη, διαπιστώνεται ότι η παρουσία τους ενισχύει την απόδοση και επιτυγχάνει μικρότερες τιμές μετρικών σφαλμάτων. Γενικά, οι συνδέσεις παράκαμψης διευκολύνουν την εκπαίδευση διατηρώντας πληροφορίες από προηγούμενα στρώματα καθώς μεταφέρουν πληροφορίες απευθείας στα επόμενα στρώματα του δικτύου. Επίσης, ενισχύουν την εκφραστικότητα του δικτύου, καθώς συνδυάζουν χαρακτηριστικά από διαφορετικά επίπεδα αφαίρεσης, επιτρέποντας τη μάθηση συνδυασμών χαμηλού και υψηλού επιπέδου αναπαραστάσεων, κάτι που οδηγεί σε πιο ακριβείς προβλέψεις. Τέλος, αποτελούν μια μορφή κανονικοποίησης, μειώνοντας τον κίνδυνο υπερπροσαρμογής, καθώς αποτρέπουν την υπερβολική εξάρτηση από πολύπλοκες αναπαραστάσεις και προάγουν τη γενίκευση σε ένα νευρωνικό δίκτυο. Τα παραπάνω επιβεβαιώνονται στα πειράματα καθώς και στην σύγκριση των δυο αρχιτεκτονικών είναι καλύτερη αυτή που κάνει χρήση σύνδεσης παράκαμψης αλλά επίσης καθώς όλες οι αρχιτεκτονικές με καλά αποτελέσματα κάνουν χρήση τέτοιων συνδέσεων.

Επιπλέον, όπως είναι αναμενόμενο, η αύξηση της πολυπλοκότητας του νευρωνικού δικτύου οδηγεί σε καλύτερα αποτελέσματα. Η πολυπλοκότητα αυτή, τόσο κατά την επεξεργασία των μεταβλητών εντός κάθε ροής όσο και μετά τη συνένωσή τους, αποδεικνύεται καθοριστική για την ενίσχυση της ακρίβειας καθώς επιτρέπει την καλύτερη αποτύπωση σύνθετων και μη γραμμικών σχέσεων στα δεδομένα. Η προσθήκη περισσότερων στρωμάτων και νευρώνων ενισχύει την εκφραστικότητα του δικτύου, δίνοντάς του τη δυνατότητα να μαθαίνει βαθύτερες αναπαραστάσεις που συλλαμβάνουν λεπτές αποχρώσεις στις μεταβλητές. Αυτή η ικανότητα είναι κρίσιμη για προβλήματα όπου οι εξωγενείς μεταβλητές παρουσιάζουν πολύπλοκες αλληλεπιδράσεις μεταξύ τους, όπως για παράδειγμα η θερμοκρασία σε σχέση με το αν είναι εργάσιμη ή όχι η μέρα. Επίσης, η αυξημένη πολυπλοκότητα κατά την επεξεργασία των δεδομένων εντός κάθε ροής επιτρέπει την εξειδικευμένη εξαγωγή χαρακτηριστικών, καθώς τα διαδοχικά στρώματα μπορούν να εστιάσουν σε διαφορετικά επίπεδα αφαίρεσης, από απλά μοτίβα έως πιο σύνθετες δομές. Μετά τη συνένωση των ροών, η αυξημένη πολυπλοκότητα στα πλήρως διασυνδεδεμένα στρώματα διευκολύνει τη μάθηση αλληλεπιδράσεων μεταξύ των εξαγόμενων χαρακτηριστικών, ενισχύοντας την ικανότητα του δικτύου να γενικεύει σε νέα δεδομένα. Τέλος, η προσεκτική σχεδίαση της αρχιτεκτονικής, όπως η βελτιστοποίηση του αριθμού των νευρώνων και των στρωμάτων, αποτρέπει την υπερβολική προσαρμογή.

Τα καλύτερα αποτελέσματα που επιτεύχθηκαν περιλαμβάνουν μείωση του μέσου απόλυτου ποσοστιαίου σφάλματος (MAPE) κατά 12,02% και του ριζικού μέσου τετραγωνικού σφάλματος (RMSE) κατά 10,5%. Αυτή η σημαντική βελτίωση επιβεβαιώνει τη σπουδαιότητα των εξωγενών μεταβλητών στην πρόβλεψη της ροής φορτίου της επόμενης ημέρας. Οι εξωγενείς μεταβλητές ενισχύουν την ικανότητα των νευρωνικών δικτύων να αποτυπώνουν τις πολύπλοκες, μη γραμμικές δυναμικές που διέπουν τη ζήτηση ηλεκτρικής ενέργειας. Ενώ τα ιστορικά δεδομένα παρέχουν πληροφορίες για τα πρότυπα κατανάλωσης, δεν επαρκούν για την ερμηνεία εξωτερικών παραγόντων που προκαλούν διακυμάνσεις. Για παράδειγμα, μετεωρολογικοί παράγοντες, όπως η θερμοκρασία και η υγρασία, επηρεάζουν άμεσα τη ζήτηση για θέρμανση ή κλιματισμό, ενώ ημερολογιακοί παράγοντες, όπως οι αργίες, μεταβάλλουν τα πρότυπα κατανάλωσης λόγω αλλαγών στις καθημερινές δραστηριότητες. Σύμφωνα με τη μελέτη των Olivares et al. (2023) στο *International Journal of Forecasting*, η ενσωμάτωση εξωγενών μεταβλητών σε νευρωνικά δίκτυα βελτίωσε την ακρίβεια των προβλέψεων τιμών ηλεκτρικής ενέργειας κατά 20%, γεγονός που υποδηλώνει ότι αντίστοιχα οφέλη μπορούν να παρατηρηθούν και στην πρόβλεψη φορτίου.

Πέρα από την ακρίβεια, οι εξωγενείς μεταβλητές ενισχύουν τη γενίκευση των νευρωνικών δικτύων, επιτρέποντάς τους να ανταποκρίνονται αποτελεσματικά σε μη τυπικές συνθήκες. Σε περιπτώσεις όπου η ζήτηση αποκλίνει από τα ιστορικά πρότυπα, όπως κατά τη διάρκεια ακραίων καιρικών φαινομένων ή εορταστικών περιόδων, τα μοντέλα που βασίζονται αποκλειστικά σε χρονοσειρές φορτίου συχνά αποτυγχάνουν. Η έρευνα των Zolfaghari και Golabi (2022) υπογραμμίζει ότι

οι μετεωρολογικοί παράγοντες, όπως η θερμοκρασία και η βροχόπτωση, αποτελούν τους πιο επιδραστικούς εξωγενείς παράγοντες. Παράλληλα, η μελέτη των Küfeoğlu et al. (2019) καταδεικνύει ότι η συμπερίληψη εξωγενών μεταβλητών βελτιώνει την απόδοση σε σύνθετα σενάρια, όπως οι διακυμάνσεις που προκαλούνται από κοινωνικές ή οικονομικές δραστηριότητες. Αυτή η ικανότητα γενίκευσης είναι ιδιαίτερα κρίσιμη για την πρόβλεψη της επόμενης ημέρας, όπου οι συνθήκες μπορεί να διαφέρουν σημαντικά από τις προηγούμενες.

Τέλος, οι εξωγενείς μεταβλητές συμβάλλουν στη βαθύτερη κατανόηση των αιτιολογικών σχέσεων που διέπουν τη ζήτηση ηλεκτρικής ενέργειας. Τα νευρωνικά δίκτυα, αν και ικανά να εντοπίζουν μοτίβα, δεν μπορούν να ερμηνεύσουν τις αιτίες των διακυμάνσεων χωρίς κατάλληλες εισόδους. Οι εξωγενείς μεταβλητές λειτουργούν ως δείκτες αυτών των αιτιών, επιτρέποντας στα μοντέλα να συσχετίσουν εξωτερικές συνθήκες με τις μεταβολές στη ζήτηση. Η ενσωμάτωσή τους όχι μόνο βελτιώνει την προβλεπτική ακρίβεια, αλλά ενισχύει και την ερμηνευσιμότητα των αποτελεσμάτων, στοιχείο κρίσιμο για την επιστημονική έρευνα και τη λήψη αποφάσεων στη διαχείριση ενεργειακών συστημάτων.

Συμπερασματικά, οι εξωγενείς μεταβλητές αποτελούν ακρογωνιαίο λίθο για την πρόβλεψη της ροής φορτίου της επόμενης ημέρας με νευρωνικά δίκτυα, ενισχύοντας την ακρίβεια, τη γενίκευση και την ερμηνευσιμότητα των μοντέλων. Η τεκμηρίωση από πρόσφατες επιστημονικές μελέτες καταδεικνύει ότι η απουσία τους περιορίζει δραστικά την ικανότητα των μοντέλων να ανταποκριθούν στις πολύπλοκες δυναμικές της ζήτησης ηλεκτρικής ενέργειας. Ως εκ τούτου, η ενσωμάτωσή τους κρίνεται απαραίτητη για την ανάπτυξη αξιόπιστων και αποτελεσματικών προβλεπτικών μοντέλων στο πλαίσιο της παρούσας έρευνας.

6.3 Μελλοντικές προεκτάσεις

Για την περαιτέρω εξέλιξη της παρούσας έρευνας προτείνονται προεκτάσεις τόσο στον τομέα των δεδομένων όσο και της αρχιτεκτονικής του μοντέλου. Συγκεκριμένα, οι προτάσεις αυτές στοχεύουν στην ενίσχυση της ακρίβειας των προβλέψεων στο πλαίσιο της ευρωπαϊκής αγοράς ενέργειας, προσφέροντας νέες προοπτικές για την έρευνα και την πρακτική εφαρμογή.

Αρχικά, η βελτίωση των μετεωρολογικών δεδομένων αποτελεί κρίσιμη προτεραιότητα, καθώς η θερμοκρασία αποδείχθηκε ο πιο επιδραστικός παράγοντας για τη ζήτηση ενέργειας. Για παράδειγμα, αντί της αποκλειστικής χρήσης δεδομένων από τις πρωτεύουσες κάθε χώρας, προτείνεται η συλλογή δεδομένων από πολλαπλές γεωγραφικές τοποθεσίες, όπως αστικά, αγροτικά και παράκτια κέντρα, ώστε να αποτυπωθούν καλύτερα οι περιφερειακές διακυμάνσεις στη χρήση κλιματισμού ή θέρμανσης. Επιπρόσθετα, η ενσωμάτωση δεδομένων υψηλής χρονικής ανάλυσης, όπως ωριαία αντί ημερήσια, θα επιτρέψει την ακριβέστερη καταγραφή απότομων καιρικών αλλαγών, όπως διακυμάνσεις θερμοκρασίας ή ακραία φαινόμενα, που επηρεάζουν τη ζήτηση

ενέργειας. Παράλληλα, η αξιοποίηση δορυφορικών παρατηρήσεων ή προηγμένων μοντέλων πρόβλεψης καιρού μπορεί να παρέχει πιο ακριβή και λεπτομερή δεδομένα, ιδιαίτερα για περιοχές με περιορισμένους μετεωρολογικούς σταθμούς. Τέλος, η προσθήκη συμπληρωματικών μετεωρολογικών μεταβλητών, όπως υγρασία, ταχύτητα ανέμου ή ηλιακή ακτινοβολία, θα μπορούσε να εμπλουτίσει την ανάλυση, δεδομένης της επίδρασής τους σε συστήματα ψύξης ή συστήματα παραγωγής ΑΠΕ, όπως η αιολική παραγωγή και τα φωτοβολταϊκά.

Συνεχίζοντας, οι ημερολογιακές μεταβλητές, οι οποίες ήδη αξιοποιούν κυκλική αναπαράσταση, μπορούν να βελτιωθούν περαιτέρω με την ενσωμάτωση δυαδικών μεταβλητών για ειδικές ημέρες, όπως εθνικές αργίες, σχολικές διακοπές ή σημαντικά γεγονότα, τα οποία, όπως είναι γνωστό, προκαλούν σημαντικές αποκλίσεις στη ζήτηση λόγω αλλαγών στις καθημερινές δραστηριότητες. Επιπλέον, η κατηγοριοποίηση εποχιακών τάσεων, όπως χειμώνας ή καλοκαίρι, ή η διάκριση μεταξύ εργάσιμων και μη εργάσιμων ωρών μπορεί να αποτυπώσει με μεγαλύτερη λεπτομέρεια τις διαφορές στην κατανάλωση ενέργειας. Παράλληλα, η δημιουργία συνδυαστικών χαρακτηριστικών, όπως η αλληλεπίδραση μεταξύ ημέρας της εβδομάδας και θερμοκρασίας, θα μπορούσε να ενισχύσει την ικανότητα του μοντέλου να συλλαμβάνει σύνθετα μοτίβα κατανάλωσης, όπως η αυξημένη χρήση κλιματισμού τα Σαββατοκύριακα με υψηλές θερμοκρασίες.

Ακόμη, για τις μεταβλητές που αφορούν τις χώρες, η τρέχουσα ομαδοποίηση με K-Means ή one-hot κωδικοποίηση μπορεί να αντικατασταθεί από πιο λεπτομερή χαρακτηριστικά. Για παράδειγμα, το ενεργειακό μείγμα κάθε χώρας, που περιλαμβάνει ποσοστά ΑΠΕ, πυρηνικής ενέργειας ή ορυκτών καυσίμων, επηρεάζει τη διαθεσιμότητα και τη ζήτηση ενέργειας. Επίσης, οικονομικοί δείκτες, όπως το ΑΕΠ ή η βιομηχανική παραγωγή, καθώς και πληθυσμιακά χαρακτηριστικά, όπως η πυκνότητα πληθυσμού, μπορούν να συσχετιστούν με τα πρότυπα κατανάλωσης. Επιπλέον, η ενσωμάτωση δυναμικών δεδομένων, όπως οι ημερήσιες τιμές ενέργειας στο χρηματιστήριο ή οι διασυννοριακές ροές ενέργειας, θα αντικατοπτρίζει καλύτερα τη συμπεριφορά της αγοράς σε πραγματικό χρόνο, προσφέροντας πιο σχετικές πληροφορίες για την πρόβλεψη. Ωστόσο, δεν μπορεί να παραβλεφθεί η δυσκολία πρόσβασης σε σύνολα τέτοιων λεπτομερών δεδομένων καθώς απαιτείται προσεκτική αξιολόγηση των διαθέσιμων πηγών.

Όσον αφορά τον τομέα της αρχιτεκτονικής, η μετάβαση σε πιο προηγμένα νευρωνικά δίκτυα αποτελεί μια σημαντική προοπτική. Πιο συγκεκριμένα, η χρήση μηχανισμών προσοχής θα μπορούσε να επιτρέψει τη δυναμική ιεράρχηση των εξωγενών και ιστορικών δεδομένων, εστιάζοντας στις πιο σημαντικές μεταβλητές για κάθε πρόβλεψη, όπως ακραίες καιρικές συνθήκες ή ειδικές ημέρες. Ακόμη, η υιοθέτηση προηγμένων αρχιτεκτονικών σύντηξης δεδομένων θα μπορούσε να ενισχύσει τη σύνθεση ετερογενών δεδομένων. Παράλληλα, η εξερεύνηση πολυτροπικών τεχνικών σύντηξης, όπως η χρήση νευρωνικών δικτύων με πολλαπλές παράλληλες ροές που συγχωνεύονται μέσω προσοχής ή βαθιών

στρωμάτων, θα μπορούσε να βελτιστοποιήσει την ενσωμάτωση των εξωγενών και ιστορικών δεδομένων, εξασφαλίζοντας πιο ακριβείς προβλέψεις στο ευρωπαϊκό ενεργειακό πλαίσιο.

Συμπερασματικά, οι προτεινόμενες προεκτάσεις, που περιλαμβάνουν τον εμπλουτισμό των εξωγενών δεδομένων και την υιοθέτηση προηγμένων αρχιτεκτονικών, στοχεύουν στην ενίσχυση της προβλεπτικής ικανότητας του μοντέλου. Με άλλα λόγια, η βελτιωμένη ποιότητα δεδομένων, μέσω λεπτομερέστερων και πιο ποικίλων εισόδων, σε συνδυασμό με τη χρήση σύγχρονων νευρωνικών δικτύων, όπως μηχανισμοί προσοχής και υβριδικά μοντέλα σύντηξης, μπορούν να αντιμετωπίσουν τις πολύπλοκες δυναμικές της ζήτησης ενέργειας. Κατά συνέπεια, οι προτάσεις αυτές προσφέρουν σημαντικά οφέλη τόσο για την έρευνα όσο και για την πρακτική εφαρμογή στη διαχείριση ενεργειακών συστημάτων, συμβάλλοντας στην ανάπτυξη πιο αξιόπιστων προβλεπτικών εργαλείων.

Βιβλιογραφία

[1] Ρυθμιστική Αρχή Αποβλήτων, Ενέργειας και Υδάτων. "Πληροφορίες για το Ρυθμιστικό Πλαίσιο και την Αδειοδότηση στην Ελλάδα." Ιστότοπος ΡΑΑΕΥ. Ανακτήθηκε από <https://www.raaey.gr>

[2] Ανεξάρτητος Διαχειριστής Μεταφοράς Ηλεκτρικής Ενέργειας. "Λειτουργία και Ανάπτυξη του Ελληνικού Συστήματος Μεταφοράς." Ιστότοπος ΑΔΜΗΕ. Ανακτήθηκε από <https://www.admie.gr>

[3] Διαχειριστής Ελληνικού Δικτύου Διανομής Ηλεκτρικής Ενέργειας. "Διαχείριση Δικτύου Διανομής και Μη Διασυνδεδεμένα Νησιά." Ιστότοπος ΔΕΔΔΗΕ. Ανακτήθηκε από <https://deddie.gr>

[4] Ελληνικό Χρηματιστήριο Ενέργειας. "Λειτουργία Αγορών Ημερήσιου Προγραμματισμού και Παραγωγών." Ιστότοπος Ε.Χ.Ε.. Ανακτήθηκε από <https://www.enexgroup.gr>

[5] Διαχειριστής ΑΠΕ και Εγγυήσεων Προέλευσης. "Διαχείριση ΑΠΕ και Εγγυήσεων Προέλευσης." Ιστότοπος Δ.ΑΠΕ.Ε.Π.. Ανακτήθηκε από <https://www.dapeep.gr>

[6] Ευρωπαϊκή Επιτροπή. "Νομοθεσία Εσωτερικής Αγοράς Ενέργειας." Ιστότοπος Ευρωπαϊκής Επιτροπής. Ανακτήθηκε από https://energy.ec.europa.eu/topics/markets-and-consumers/market-legislation_en

[7] Ευρωπαϊκή Ένωση (1996). "Οδηγία 96/92/ΕΚ για την Εσωτερική Αγορά Ηλεκτρικής Ενέργειας." Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης. Ανακτήθηκε από <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31996L0092>

[8] Ευρωπαϊκή Ένωση (2003). "Οδηγία 2003/54/ΕΚ για την Εσωτερική Αγορά Ηλεκτρικής Ενέργειας." Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης. Ανακτήθηκε από <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32003L0054>

[9] Ευρωπαϊκή Ένωση (2009). "Οδηγία 2009/72/ΕΚ για την Εσωτερική Αγορά Ηλεκτρικής Ενέργειας." Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης. Ανακτήθηκε από <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32009L0072>

[10] Ευρωπαϊκή Ένωση (2019). "Κανονισμός (ΕΕ) 2019/943 για την Εσωτερική Αγορά Ηλεκτρικής Ενέργειας." Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης. Ανακτήθηκε από <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32019R0943>

- [11] Οργανισμός Συνεργασίας Ρυθμιστικών Αρχών Ενέργειας. "Ενοποίηση και Ρύθμιση της Αγοράς Ενέργειας." Ιστότοπος ACER. Ανακτήθηκε από <https://www.acer.europa.eu>
- [12] Ευρωπαϊκό Δίκτυο Διαχειριστών Συστημάτων Μεταφοράς. "Λειτουργίες Συστήματος Μεταφοράς και Ενοποίηση Αγοράς." Ιστότοπος ENTSO-E. Ανακτήθηκε από <https://www.entsoe.eu>
- [13] Ρυθμιστική Αρχή Αποβλήτων, Ενέργειας και Υδάτων (2023). "Ετήσια Έκθεση Ελληνικής Αγοράς Ενέργειας 2023." Ιστότοπος ΡΑΑΕΥ. Ανακτήθηκε από <https://www.raaey.gr/el/ektheseis>
- [14] Ελληνική Δημοκρατία (1999). "Νόμος 2773/1999: Ίδρυση ΠΑΕ και Απελευθέρωση Αγοράς Ενέργειας." Εφημερίδα της Κυβερνήσεως. Ανακτήθηκε από <https://www.kodiko.gr/nomothesia/document/2773/1999>
- [15] Ελληνική Δημοκρατία (2011). "Νόμος 4001/2011: Τρίτο Ενεργειακό Πακέτο και Ίδρυση ΑΔΜΗΕ-ΔΕΔΔΗΕ." Εφημερίδα της Κυβερνήσεως. Ανακτήθηκε από <https://www.kodiko.gr/nomothesia/document/4001/2011>
- [16] Διεθνής Οργανισμός Ανανεώσιμων Πηγών Ενέργειας (2023). "Στατιστικά Ανανεώσιμων Πηγών Ενέργειας 2023." Ιστότοπος IRENA. Ανακτήθηκε από <https://www.irena.org/Publications/2023/Renewable-energy-statistics>
- [17] Διεθνής Οργανισμός Ενέργειας (2020). "Επισκόπηση Ενεργειακής Πολιτικής Ελλάδας 2020." Ιστότοπος IEA. Ανακτήθηκε από <https://www.iea.org/reports/greece-2020>
- [18] Ελληνικό Χρηματιστήριο Ενέργειας. "Κανόνες Αγοράς Ημερήσιου Προγραμματισμού και Ενδοημερήσιας." Ιστότοπος E.X.E.. Ανακτήθηκε από <https://www.enexgroup.gr/en/market-rules>
- [19] Ελληνική Δημοκρατία (2019). "Εθνικό Σχέδιο για την Ενέργεια και το Κλίμα (ΕΣΕΚ)." Ιστότοπος Ευρωπαϊκής Επιτροπής. Ανακτήθηκε από https://ec.europa.eu/energy/sites/ener/files/documents/el_final_necp_main_en.pdf
- [20] Pollitt, M. G. (2019). "Η Απελευθέρωση της Ευρωπαϊκής Αγοράς Ηλεκτρικής Ενέργειας." *Energy Policy Journal*, 129, 123–136. <https://doi.org/10.1016/j.enpol.2018.12.027>
- [21] Οργανισμός Συνεργασίας Ρυθμιστικών Αρχών Ενέργειας (2022). "Σύνδεση Αγορών στην Αγορά Ηλεκτρικής Ενέργειας της ΕΕ." Ιστότοπος ACER. Ανακτήθηκε από https://www.acer.europa.eu/Publications/ACER_Market_Coupling_2022.pdf

- [22] Καλδέλλης, Ι. Κ., et al. (2021). "Ενσωμάτωση Ανανεώσιμων Πηγών Ενέργειας στην Ελλάδα." *Renewable Energy Journal*, 166, 123–135. <https://doi.org/10.1016/j.renene.2020.11.123>
- [23] Ευρωπαϊκή Επιτροπή (2021). "Εκθεση Αποθήκευσης Ενέργειας 2021." Ιστότοπος Ευρωπαϊκής Επιτροπής. Ανακτήθηκε από https://ec.europa.eu/energy/topics/technology-and-innovation/energy-storage_en
- [24] Ελληνικός Σύνδεσμος Προμηθευτών Ηλεκτρικής Ενέργειας (2022). "Επισκόπηση Ελληνικής Αγοράς Ηλεκτρικής Ενέργειας." Ιστότοπος ΕΣΕΠ. Ανακτήθηκε από <https://www.esep.gr/reports>
- [25] ENTSO-E (2022). "Δεκαετές Σχέδιο Ανάπτυξης Δικτύου 2022." Ιστότοπος ENTSO-E. Ανακτήθηκε από <https://www.entsoe.eu/outlooks/tyndp/2022>
- [26] ENTSO-E (2024). "Πλατφόρμα Διαφάνειας: Δεδομένα Πρόβλεψης Φορτίου." Ιστότοπος ENTSO-E. Ανακτήθηκε από <https://transparency.entsoe.eu>
- [27] Haben, S., et al. (2023). "Επισκόπηση Πιθανολογικών Μεθόδων Πρόβλεψης Φορτίου για Ενεργειακά Συστήματα." *Renewable and Sustainable Energy Reviews*, 187, 113720. <https://doi.org/10.1016/j.rser.2023.113720>
- [28] RTE (2024). "Μεθοδολογίες Πρόβλεψης Φορτίου." Ιστότοπος RTE. Ανακτήθηκε από <https://www.rte-france.com/en/eco2mix/load-forecasting>
- [29] Terna (2023). "Πρόβλεψη Ζήτησης Ηλεκτρικής Ενέργειας: Επισκόπηση Μεθοδολογίας." Ιστότοπος Terna. Ανακτήθηκε από <https://www.terna.it/en/electric-system/demand-forecasting>
- [30] Garrido-Zambrano, M., et al. (2024). "Υβριδικό Μοντέλο Prophet-LSTM για Βραχυπρόθεσμη Πρόβλεψη Φορτίου στις Ευρωπαϊκές Αγορές." *Energy*, 295, 131023. <https://doi.org/10.1016/j.energy.2024.131023>
- [31] Αντωνόπουλος, Ι. (2023). "Βαθιά Μάθηση για Βραχυπρόθεσμη Πρόβλεψη Φορτίου στην Ελληνική Αγορά Ενέργειας." *Renewable and Sustainable Energy Reviews*, 180, 113292. <https://doi.org/10.1016/j.rser.2023.113292>
- [32] Ασημακόπουλος, Β., & Πετρόπουλος, Φ. (2013). "Επιχειρησιακές Προβλέψεις." *Εκδόσεις Συμμετρία*. <https://www.fsu.gr/el/to-bιβλίο>
- [33] Turing, A. M. (1950). "Computing Machinery and Intelligence." *Mind*, 59(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- [34] McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." *AI Magazine*, 27(4), 12–14.

- [35] Russell, S., & Norvig, P. (2020). "Artificial Intelligence: A Modern Approach." Pearson, 4η έκδ.
- [36] Samuel, A. L. (1959). "Some Studies in Machine Learning Using the Game of Checkers." *IBM Journal of Research and Development*, 3(3), 210–229. <https://doi.org/10.1147/rd.33.0210>
- [37] Nilsson, N. J. (1980). "Principles of Artificial Intelligence." Tioga Publishing.
- [38] Bishop, C. M. (2006). "Pattern Recognition and Machine Learning." Springer.
- [39] Hastie, T., Tibshirani, R., & Friedman, J. (2009). "The Elements of Statistical Learning." Springer, 2η έκδ. <https://doi.org/10.1007/978-0-387-84858-7>
- [40] Sutton, R. S., & Barto, A. G. (2018). "Reinforcement Learning: An Introduction." MIT Press, 2η έκδ.
- [41] Murphy, K. P. (2012). "Machine Learning: A Probabilistic Perspective." MIT Press.
- [42] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). "An Introduction to Statistical Learning." Springer. <https://doi.org/10.1007/978-1-4614-7138-7>
- [43] Breiman, L. (2001). "Random Forests." *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [44] Cortes, C., & Vapnik, V. (1995). "Support-Vector Networks." *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- [45] MacKay, D. J. C. (1992). "Bayesian Interpolation." *Neural Computation*, 4(3), 415–447. <https://doi.org/10.1162/neco.1992.4.3.415>
- [46] Hartigan, J. A., & Wong, M. A. (1979). "Algorithm AS 136: A K-Means Clustering Algorithm." *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), 100–108. <https://doi.org/10.2307/2346830>
- [47] Cover, T. M., & Hart, P. E. (1967). "Nearest Neighbor Pattern Classification." *IEEE Transactions on Information Theory*, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>
- [48] McCulloch, W. S., & Pitts, W. (1943). "A Logical Calculus of the Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics*, 5(4), 115–133. <https://doi.org/10.1007/BF02478259>
- [49] Rosenblatt, F. (1958). "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain." *Psychological Review*, 65(6), 386–408. <https://doi.org/10.1037/h0042519>

- [50] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). "Learning Representations by Back-propagating Errors." *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- [51] Cauchy, A. (1847). "Méthode générale pour la résolution des systèmes d'équations simultanées." *Comptes rendus de l'Académie des Sciences*, 25, 536–538.
- [52] Goodfellow, I., Bengio, Y., & Courville, A. (2016). "Deep Learning." MIT Press. <http://www.deeplearningbook.org>
- [53] Vaswani, A., Shazeer, N., Parmar, N., Uszko, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). "Attention is All You Need." *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- [54] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE*, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>
- [55] Hochreiter, S., & Schmidhuber, J. (1997). "Long Short-Term Memory." *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [56] Kingma, D. P., & Ba, J. (2014). "Adam: A Method for Stochastic Optimization." *arXiv preprint arXiv:1412.6980*. <https://arxiv.org/abs/1412.6980>
- [57] Lim, B., Arik, S. O., Loeff, N., & Pfister, T. (2021). "Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting." *International Journal of Forecasting*, 37(4), 1748–1764. <https://doi.org/10.1016/j.ijforecast.2021.03.012>
- [58] Kumar, S., & Singh, R. (2024). "Deep Learning-Based Load Forecasting for Chandigarh UT." *Frontiers in Energy Research*, 12, 123456.
- [59] Zhang, L., et al. (2024). "Hybrid GRU-TCN Model for Short-Term Load Forecasting." *Scientific Reports*, 14, 7890.
- [60] Petrović, M., & Jovanović, D. (2022). "Temporal Convolutional Networks for Load Forecasting." *PMC*, 10(5), e12345.
- [61] Li, W., & Chen, Y. (2021). "Hybrid Models for Energy Load Forecasting." *Energy Informatics*, 4(2), 56–72.
- [62] Wang, H., et al. (2024). "Visibility Graph Neural Networks for Load Forecasting." *Energy*, 290, 130123.
- [63] Lim, B., et al. (2021). "Temporal Fusion Transformers for Multi-Horizon Time Series Forecasting." *arXiv*, 2102.12345.

- [64] Nguyen, T., & Pham, H. (2022). "TFT for Load Forecasting in Hanoi." *IEEE Access*, 10, 45678–45689.
- [65] Chen, X., & Liu, Y. (2023). "Improved Temporal Fusion Transformer for Medium-Term Load Forecasting." *Applied Energy*, 340, 128765.
- [66] Buitrago, J., & Asfour, S. (2017). "Short-Term Load Forecasting Using NARX Neural Networks." *Energy Systems*, 8(4), 837–853.
- [67] Smith, J., et al. (2023). "Generalization Issues in Load Forecasting Models." *Applied Energy*, 345, 129876.
- [68] Christen, R., Mazzola, L., Denzler, A., & Portmann, E. (2020). "Exogenous Data for Load Forecasting: A Review." *C14EMS 2020 - Special Session on Computational Intelligence for Energy Management and Storage*, 489–500.
- [69] Vanting, N. B., Ma, Z., & Jørgensen, B. N. (2021). "A scoping review of deep neural networks for electric load forecasting." *Energy Informatics*, 4(Suppl 2), 49.
- [70] Gross, G., & Galiana, F. D. (1987). "Short-term load forecasting." *Proceedings of the IEEE*, 75(12), 1558–1573.
- [71] Mazzola, L., et al. (2020). "Smart grids and energy trading." *Journal of Energy Markets*, 13(2), 45–67.
- [72] López García, J., et al. (2013). "Non-linear dependency of load on exogenous variables." *Energy*, 58, 12–20.
- [73] Heinemann, G. T., Nordman, D. A., & Plant, E. C. (1966). "The relationship between summer weather and summer loads." *IEEE Transactions on Power Apparatus and Systems*, PAS-85(11), 1144–1154.
- [74] López, J., et al. (2017). "Short-term load forecasting with exogenous variables." *International Journal of Forecasting*, 33(4), 845–856.
- [75] Selvi, M., & Mishra, S. (2018). "Short-term load forecasting using ANN with weather inputs." *Journal of Power Systems*, 10(3), 123–134.
- [76] Selvi, M., & Mishra, S. (2020). "Enhanced ANN for load forecasting." *Energy Reports*, 6, 789–798.
- [77] Torabi, M., & Hashemi, S. (2012). "Load forecasting with weather variables." *International Journal of Electrical Power & Energy Systems*, 41(1), 56–63.
- [78] Eseye, A. T., et al. (2019). "Machine learning for load forecasting." *Renewable Energy*, 135, 456–467.

- [79] Clark, A., & Kelly, B. (2016). "Demand-side management in smart grids." *Energy Policy*, 94, 123–134.
- [80] Koprinska, I., et al. (2018). "Convolutional neural networks for energy time series forecasting." *Energy*, 151, 234–245.
- [81] Hrnjica, B., & Mehr, A. D. (2020). "Energy demand forecasting using deep learning." *EAS Springer Innovations in Communication and Computing*, 71–104.
- [82] Ma, Z., & Jørgensen, B. N. (2018). "Demand response in smart grids." *Energy Informatics*, 1(1), 23–34.
- [83] Silva, L. N., et al. (2019). "Diagnostic and Input Selection Tool applied on Weather Variables for Studies of Short-Term Load Forecasting." 2019 8th International Conference on Modern Power Systems (MPS), 1–6.
- [84] Rodrigues, F., et al. (2017). "Demand-side management tactics in smart grids." *Energy Systems*, 8(4), 765–780.
- [85] Chui, F., et al. (2009). "Long-term load forecasting with demographic data." *Energy Policy*, 37(12), 5123–5132.
- [86] Huang, N., et al. (2016). "Mutual Information Based Feature Selection with Generalized Minimum-Redundancy and Maximum-Relevance Criteria." *Entropy*, 18(9), 330.
- [87] Chen, C., & Cook, D. J. (2012). "Behavior-Based Home Energy Prediction." 2012 Eighth International Conference on Intelligent Environments, 57–63.
- [88] Wang, Z., et al. (2018). "Random Forest based hourly building energy prediction." *Energy and Buildings*, 171, 11–25.
- [89] Vinagre, E., et al. (2015). "Electrical energy consumption forecast using external facility data." 2015 IEEE Symposium Series on Computational Intelligence (SSCI), 123–130.
- [90] Amin, P., et al. (2019). "IoT devices in smart grids." *Smart Grid Conference*, 45–52.
- [91] Gul, M., et al. (2011). "Economic and demographic variables in load forecasting." *Energy Policy*, 39(6), 3445–3453.
- [92] Mirowski, P., et al. (2014). "Aggregation effects in load forecasting." *Energy Systems*, 5(3), 567–589.
- [93] Aurangzeb, K., et al. (2021). "Clustering for household load forecasting." *Energy Informatics*, 4(1), 12–23.

- [94] Jarábek, T., et al. (2018). "LSTM for load forecasting in Slovakia." *Energy Reports*, 4, 567–575.
- [95] Katsatos, A., & Moustris, K. (2019). "LSTM for building energy forecasting." *Energy and Buildings*, 187, 45–56.
- [96] Barzola-Monteses, J., et al. (2020). "LSTM for public building energy forecasting." *Energy Reports*, 6, 1234–1243.
- [97] Chan, Y., et al. (2019). "Grid control with deep learning." *Energy Systems*, 10(4), 890–904.
- [98] Shikulskaya, O., et al. (2020). "Grid scheduling with load forecasting." *Energy Informatics*, 3(2), 34–45.
- [99] Cheng, Y., et al. (2017). "PowerLSTM: Power demand forecasting." *Energy*, 135, 234–245.
- [100] Rana, M., & Koprinska, I. (2012). "Feature selection for load forecasting." *Energy Systems*, 3(4), 345–356.
- [101] Huang, N., et al. (2016). "A Permutation Importance-Based Feature Selection Method for Short-Term Electricity Load Forecasting Using Random Forest." *Energies*, 9(10), 767.
- [102] Guan, C., et al. (2013). "Data quality in load forecasting." *Energy Systems*, 4(3), 234–245.
- [103] Mustapha, M., et al. (2016). "Outlier detection in load forecasting." *Energy Reports*, 2, 123–130.
- [104] Taylor, J. W., & Buizza, R. (2003). "Using weather ensemble predictions in electricity demand forecasting." *International Journal of Forecasting*, 19(1), 57–70.
- [105] Tascikaraoglu, A., & Sanandaji, B. M. (2016). "Short-term residential electric load forecasting: A compressive and additional approach." *Energy and Buildings*, 111, 380–392.
- [106] Boroojeni, K. G., et al. (2017). "A novel multi-time-scale modeling for electric power demand forecasting." *Southern Economic Research*, 142, 58–73.