ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

ΠΡΟΓΡΑΜΜΑ ΔΙΔΑΚΤΟΡΙΚΩΝ ΣΠΟΥΔΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Ανάπτυξη επιβλεπόμενων και μη επιβλεπόμενων μεθόδων μηχανικής μάθησης για την μοντελοποίηση χρονικά εξελισσόμενων συστημάτων με εφαρμογές σε έξυπνα ηλεκτρικά δίκτυα

Νικόλαος Α Γιαμαρέλος

ΑΘΗΝΑ-ΑΙΓΑΛΕΩ

ΣΕΠΤΕΜΒΡΙΟΣ 2023

Η εκπόνηση της διδακτορικής διατριβής συγχρηματοδοτήθηκε από την Ελλάδα και την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο – Ε.Κ.Τ.) μέσω του Επιχειρησιακού Προγράμματος «Ανάπτυξη Ανθρώπινου Δυναμικού, Εκπαίδευση και Δια Βίου Μάθηση» στο πλαίσιο της δράσης «Ενίσχυση του ανθρώπινου δυναμικού μέσω της υλοποίησης διδακτορικής έρευνας Υποδράση 2: Πρόγραμμα χορήγησης υποτροφιών ΙΚΥ σε υποψηφίους διδάκτορες των ΑΕΙ της Ελλάδας».

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

Επιχειρησιακό Πρόγραμμα
Ανάπτυξη Ανθρώπινου Δυναμικού,
Εκπαίδευση και Διά Βίου Μάθηση
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

ΕΣΠΑ
2014-2020
ανάπτυξη - εργασία - αλληλεγγύη

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Ανάπτυξη επιβλεπόμενων και μη επιβλεπόμενων μεθόδων μηχανικής μάθησης για την μοντελοποίηση χρονικά εξελισσόμενων συστημάτων με εφαρμογές σε έξυπνα ηλεκτρικά δίκτυα

Νικόλαος Α. Γιαμαρέλος

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: Δρ Ηλίας Ζώης, Αναπληρωτής Καθηγητής Τμήμα Ηλεκτρολόγων & Ηλεκτρονικών Μηχανικών, Πανεπιστήμιο Δυτικής Αττικής.

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ ΠΑΡΑΚΟΛΟΥΘΗΣΗΣ:

Ηλίας Ζώης, Αναπληρωτής Καθηγητής Τμήμα Ηλεκτρολόγων & Ηλεκτρονικών Μηχανικών, Πανεπιστήμιο Δυτικής Αττικής.

Γρηγόριος Κουλούρας, Αναπληρωτής Καθηγητής Τμήμα Ηλεκτρολόγων & Ηλεκτρονικών Μηχανικών, Πανεπιστήμιο Δυτικής Αττικής.

Αλέξανδρος Αλεξανδρίδης, Καθηγητής Τμήμα Ηλεκτρολόγων & Ηλεκτρονικών Μηχανικών, Πανεπιστήμιο Δυτικής Αττικής.

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

(Υπογραφή)

Ilias Zois Digitally signed by Ilias Zois Date: 2023.10.17 10:13:13 +03'00'

Ηλίας Ζώης,
Αναπληρωτής Καθηγητής ΠαΔΑ

(Υπογραφή)

Digitally signed by Grigorios Koulouras DN: c=GR, l=Athens, o=University of West Attica, ou=Class A - Private Key created and stored in hardware CSP, sn=Koulouras, givenName=Grigorios, serialNumber=6090016933, cn=Grigorios Koulouras, email=gregkoul@uniwa.gr Date: 2023.10.16 15:52:14 +03'00'

Γρηγόριος Κουλούρας,
Αναπληρωτής Καθηγητής ΠαΔΑ

(Υπογραφή)

Digitally signed by Alexandros Alexandridis Date: 2023.10.16 15:24:45 +03'00'

Αλέξανδρος Αλεξανδρίδης,
Καθηγητής ΠαΔΑ

(Υπογραφή)

Digitally signed by Konstantinos Psomopoulos Date: 2023.10.16 16:16:48 +03'00'

Κωνσταντίνος Ψωμόπουλος,
Καθηγητής ΠαΔΑ

(Υπογραφή)

Evangelos Zervas Digitally signed by Evangelos Zervas Date: 2023.10.16 13:41:29 +03'00'

Ευάγγελος Ζέρβας,
Καθηγητής ΠαΔΑ

(Υπογραφή)

Dimitrios Kalyvas Digitally signed by Dimitrios Kalyvas Date: 2023.10.16 15:09:05 +03'00'

Δημήτριος Καλύβας,
Καθηγητής ΠαΔΑ

(Υπογραφή)

EMMANOUIL KARAPIDAKIS
16/10/2023 17:56

Εμμανουήλ Καραπιδάκης,
Καθηγητής ΕΛΜΕΠΑ

Ημερομηνία εξέτασης 10/12/2023

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΔΑΚΤΟΡΙΚΗΣ ΔΙΑΤΡΙΒΗΣ

Ο κάτωθι υπογεγραμμένος Νικόλαος Γιαμαρέλος του Αντωνίου, υποψήφιος διδάκτορας του Τμήματος Ηλεκτρολόγων και Ηλεκτρονικών Μηχανικών της Σχολής Μηχανικών του Πανεπιστημίου Δυτικής Αττικής, δηλώνω ότι:

«Είμαι συγγραφέας και δικαιούχος των πνευματικών δικαιωμάτων επί της διατριβής και δεν προσβάλω τα πνευματικά δικαιώματα τρίτων. Για τη συγγραφή της διδακτορικής μου διατριβής δεν χρησιμοποίησα ολόκληρο ή μέρος έργου άλλου δημιουργού ή τις ιδέες και αντιλήψεις άλλου δημιουργού χωρίς να γίνεται αναφορά στην πηγή προέλευσης (βιβλίο, άρθρο από εφημερίδα ή περιοδικό, ιστοσελίδα κ.λπ.). Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Ο Δηλών

ΠΕΡΙΛΗΨΗ

Η παρούσα διδακτορική διατριβή (Δ.Δ.) άπτεται της μελέτης των χρονικά εξελισσόμενων συστημάτων με σκοπό την ανάπτυξη μεθοδολογιών για την μοντελοποίησή τους και κατ' επέκταση την χρήση τους στο πρόβλημα της πρόβλεψης της κατανάλωσης του ηλεκτρικού φορτίου από υποσταθμούς μέσης και υψηλής τάσης. Η λειτουργία τους βασίζεται στην αλγοριθμική προσαρμογή της σχέσης που περιγράφει την είσοδο και την έξοδο των δεδομένων αυτών σε μορφή χρονικών ακολουθιών ή χρονοσειρών. Βασικός σκοπός της διδακτορικής διατριβής είναι να προτείνει και να διερευνήσει την εφαρμογή νέων μεθόδων μηχανικής μάθησης με και χωρίς επίβλεψη για την πρόβλεψη χαρακτηριστικών από χρονοσειρές. Κατά συνέπεια, ένα σημαντικό τμήμα του θεωρητικού μέρους της εργασίας καταλαμβάνει η ανάλυση χρονοσειρών ως εργαλείο της μοντελοποίησης χρονικά εξελισσόμενων συστημάτων και η μελέτη της πρόβλεψης χρονοσειρών. Στη συνέχεια διεξάγεται μια ενδελεχής βιβλιογραφική ανασκόπηση σχετικά με τις μεθόδους που έχουν προταθεί για την πρόβλεψη του ηλεκτρικού φορτίου. Μέσα από αυτή αναδεικνύονται συγκεκριμένα ερευνητικά κενά που προκύπτουν κατά την διερεύνηση του ανωτέρω προβλήματος. Αρχικά, παρατηρείται ότι παρά το γεγονός ότι ένας μεγάλος αριθμός μεθοδολογιών μηχανικής μάθησης (Μ.Μ.) έχει προταθεί για την πρόβλεψη του ηλεκτρικού φορτίου, εντούτοις καμία από αυτές δεν δύναται να υπερέχει όλων των υπολοίπων. Αυτό οφείλεται κυρίως στη ύπαρξη μη γραμμικότητας και τις διαφορετικές στοχαστικές ιδιότητες που χαρακτηρίζουν τις χρονοσειρές κατανάλωσης του ηλεκτρικού φορτίου. Ένας επιπλέον παράγοντας που επιδεινώνει τα χαρακτηριστικά αυτά προέρχεται από το νέο μοντέλο δικτύου που επιβάλλει η μετάβαση από τα συμβατικά στα λεγόμενα έξυπνα δίκτυα (smart grids). Σαν αποτέλεσμα, το περιβάλλον του έξυπνου δικτύου εισάγει την αμφίδρομη ροή ισχύος ανάμεσα στους χρήστες και στο δίκτυο κοινής ωφέλειας, με αποτέλεσμα τα δεδομένα του φορτίου να περιέχουν μετρήσεις παραγωγής και κατανάλωσης, και οι χρονοσειρές του μικτού φορτίου όπως ονομάζεται, να παρουσιάζουν υψηλή αβεβαιότητα. Κρίνεται λοιπόν αναγκαία, η αναζήτηση πρωτοπόρων και εύρωστων μεθόδων Μ.Μ. ικανών να αντιμετωπίσουν αποτελεσματικά τα προαναφερθέντα φαινόμενα και να επιδείξουν καθολική υπεροχή έναντι άλλων ανταγωνιστικών. Μια ακόμη έλλειψη που εντοπίστηκε ήταν η απουσία ενασχόλησης με την πρόβλεψη του μικτού φορτίου, σε αντίθεση με την εκτεταμένη μελέτη της πρόβλεψης της παραγωγής ή της κατανάλωσης ηλεκτρικής ενέργειας μεμονωμένα. Με βάση τις επισημάνσεις αυτές, προτάθηκε η ανάπτυξη τριών νέων μεθόδων πρόβλεψης. Η πρώτη εξ αυτών βασίζεται στην θεωρία της αραιής αναπαράστασης (sparse representation - SR) και της εκμάθησης λεξικού (dictionary learning), με κύρια πρόταση την ενσωμάτωση των ατόμων και των αντιστοίχων αραιών συντελεστών σε μια ιεραρχική - δενδροειδή δομή. Η ανωτερότητα της προτεινόμενης μεθόδου έγκυται στη χαμηλή της πολυπλοκότητα και στην εγγενή της ικανότητα να αναπαριστά αποδοτικά τα μη γραμμικά χαρακτηριστικά του μικτού φορτίου, παραμένοντας ανθεκτική σε φαινόμενα υπερπροσαρμογής. Η επόμενη μεθοδολογία που προτείνεται αφορά σε ένα ανσάμπλ μοντέλο το οποίο εκπαιδεύεται από διάφορες μεθόδους Μ.Μ. Συγκεκριμένα συνδυάζονται τεχνικές από: τεχνητά νευρωνικά δίκτυα (Τ.Ν.Δ), γραμμική παλινδρόμηση (ΓΠ), παλινδρόμηση με χρήση διανυσμάτων υποστήριξης (support vector regression – SVR), τυχαία δάση (random forests – RF) και αραιή αναπαράσταση. Οι τελικές τιμές πρόβλεψης υπολογίζονται με τη βοήθεια ενός online μηχανισμού απόφασης που βασίζεται στη στάθμιση των υπο-μοντέλων σύμφωνα με την πρότερη απόδοσή τους. Το σύνολο των προτεινόμενων μεθοδολογιών πρόβλεψης ολοκληρώνεται με ένα μοντέλο νευρωνικών δικτύων Γράφων (Graph neural networks - GNN), το οποίο ενισχύεται με τις ιδιότητες των Γράφων ορατότητας (Visibility graphs). Σύμφωνα με την προσσέγγιση αυτή, η οποία αποκαλείται νευρωνικό δίκτυο Γράφων ορατότητας (Visibility graph neural network - VGNN), οι χρονοσειρές του μικτού φορτίου μετασχηματίζονται σε ένα μη κατευθυντικό Γράφο χρησιμοποιώντας το κριτήριο της φυσικής ορατότητας μεταξύ των κόμβων ενός Γράφου και στη συνέχεια χρησιμοποιείται ένας πίνακας γειτνίασης για την εκπαίδευση ενός GNN. Επόμενο βήμα

αποτέλεσε η αξιολόγηση των προτεινόμενων μεθόδων. Το πεδίο εφαρμογής για την πειραματική δοκιμή τους αντλείται από τον τομέα των έξυπνων δικτύων (smart grids). Συγκεκριμένα, προτάθηκε η πρόβλεψη του μικτού ηλεκτρικού φορτίου με χρήση πραγματικών δεδομένων από υποσταθμό (Υ/Σ) υψηλής και μέσης τάσης (ΥΤ/ΜΤ). Ύστερα από την αναλυτική εισαγωγή στα έξυπνα δίκτυα και την ουσιώδη συμβολή τους στην πρόβλεψη του φορτίου και κατ’ επέκταση στην αποδοτική λειτουργία τους, ακολουθεί η περιγραφή του πειραματικού πρωτοκόλλου των προαναφερθέντων προσομοιώσεων. Σύμφωνα με τα παραγόμενα αποτελέσματα οι προτεινόμενες μεθοδολογίες Μ.Μ. καταφέρνουν να παρέχουν αξιόπιστες, εύρωστες και ακριβείς προβλέψεις για περιπτώσεις διαφορετικών χρονικών οριζόντων πρόβλεψης που κυμαίνονται από 15 λεπτά έως 24 ώρες αλλά και διαφορετικών σεναρίων εκτέλεσης που αφορούν δεδομένα τα οποία εκτείνονται σε περιόδους ιδιότυπων μοτίβων κατανάλωσης.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Υπολογιστική Νοημοσύνη, Μηχανική Μάθηση, Έξυπνα Δίκτυα Ηλεκτρικής Ενέργειας.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Επιβλεπόμενη μάθηση, μη επιβλεπόμενη μάθηση, αραιή αναπαράσταση, πρόβλεψη μικτού φορτίου, νευρωνικά δίκτυα γράφων, γράφοι ορατότητας, εκμάθηση ανσάμπλ.

ABSTRACT

The present doctoral thesis concerns the study of time evolving systems with the aim of developing methodologies for their modeling and, their consequent use in the problem of forecasting the consumption of electrical load from medium and high voltage substations. Their operation is based on the algorithmic adaptation of the relationship that describes the input and output data in the form of time sequences or time series. The main purpose of this doctoral thesis is to propose and investigate the application of new supervised and unsupervised machine learning methods for prediction of features extracted from time series. Consequently, an important part of the theoretical section of this work is occupied by time series analysis as a tool for modeling of time evolving systems and the study of time series forecasting. A thorough literature review is then conducted on the methods that have been proposed for electric load forecasting, through which, specific research gaps that arise during the investigation of the above problem are highlighted. Initially, it is observed that despite the fact that a large number of machine learning methodologies have been proposed for the prediction of electrical load, however, none of them can outperform all the rest. This is mainly due to the existence of non-linearity and the different stochastic properties that characterize the time series of electric load consumption. An additional factor that worsens these characteristics stems from the new electrical grid paradigm imposed by the transition from conventional to so-called smart grids. As a result, the smart grid environment introduces two-way power flow between users and the utility grid, resulting in the load data containing both production and consumption measurements, and the so-called mixed load time series exhibiting high amount of uncertainty. It is therefore deemed necessary to search for novel and robust machine learning methods, able to effectively deal with the aforementioned phenomena and demonstrate global superiority over other competitors. Another shortcoming that has been identified is the scant consideration of mixed load forecasting, as opposed to the extensive study of forecasting electricity generation or consumption individually. Based on these observations, the development of three new forecasting methods is proposed. The first of these is based on the theory of sparse representation and dictionary learning, with its main proposition being the integration of the atoms and the corresponding sparse coefficients in a hierarchical tree structure. The superiority of the proposed method lies in its low complexity and its inherent ability to efficiently represent the nonlinear characteristics of the mixed load while remaining robust to overfitting effects. The next methodology that is introduced concerns an ensemble model which is trained by various machine learning methods. Specifically, techniques from: artificial neural networks, linear regression, support vector regression, random forests and sparse representation are combined. The final predicted values are calculated with the use of an online decision mechanism based on the weighting of the sub-models according to their past performance. The set of proposed forecasting methodologies is completed with a GNN model, which is enhanced with the properties of Visibility graphs. According to this approach, which is called a Visibility graph neural network (VGNN), the mixed load time series is transformed into an undirected graph using the criterion of natural visibility between the nodes of a graph, and then an adjacency matrix is employed for training a GNN. The next step was the evaluation of the proposed methods. The scope for their experimental testing is drawn from the field of smart grids. Specifically, the mixed electric load forecasting is taking place, using real data from a high and medium voltage substation. After the detailed introduction to smart grids and their essential contribution to load forecasting and, furtherly, to their efficient operation, follows the description of the experimental protocol of the aforementioned simulations. According to the obtained results, the proposed machine learning methodologies manage to provide reliable, robust and accurate forecasts for cases of different forecast time horizons ranging from 15 minutes to 24 hours ahead but also different execution scenarios involving data spanning periods of idiosyncratic consumption patterns.

SUBJECT AREA: Computational intelligence, Machine learning, Smart grids.

KEYWORDS: Supervised learning, unsupervised learning, sparse representation, mixed load forecasting, graph neural networks, visibility graphs, ensemble learning

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα εκφράσω τις ευχαριστίες μου πρωτίστως στους γονείς μου οι οποίοι με στήριζουν καθ' όλη τη διάρκεια των σπουδών μου με δυνατό κάθε τρόπο. Έπειτα ευχαριστώ τον επιβλέποντα της παρούσας Διδακτορικής Διατριβής, κ. Ζώη Ηλία για την αδάλειπτη παροχή γνώσεων και ερευνητικής καθοδήγησης αλλά και όλους τους συναδέλφους που μου πρόσφεραν αμέριστη βοήθεια και μια άψογη συνεργασία. Τέλος, θα ήθελα να ευχαριστήσω την σύντροφό μου για την συνεχή της στήριξη και κατανόηση.

ΛΙΣΤΑ ΔΗΜΟΣΙΕΥΣΕΩΝ

[1] M. Papadimitrakis, N. Giamarelos, M. Stogiannos, E.N. Zois, N.A. Livanos, A. Alexandridis, “*Metaheuristic search in smart grid: A review with emphasis on planning, scheduling and power flow optimization applications*”, Renewable and Sustainable Energy Reviews, 145 (2021), 111072.

[2] N. Giamarelos, E. N. Zois, M. Papadimitrakis, M. Stogiannos, N.-A. I. Livanos, A. Alexandridis, “*Short-Term Electric Load Forecasting with Sparse Coding Methods*”, IEEE Access, 9 (2021), pp. 102847-102861.

[3] N. A. I. Livanos, S. Hammal, N. Giamarelos, V. Alifragkis, C. S. Psomopoulos, E. N. Zois, “*OpenEdgePMU: An Open PMU Architecture with Edge Processing for Future Resilient Smart Grids”*, Energies, 16(6) (2023), 2756.

[4] N. Giamarelos, M. Papadimitrakis, M. Stogiannos, Elias N. Zois, Nikolaos A.-I. Livanos, Alex Alexandridis, “*Mixed power-load forecasting using a machine learning-based multi-model scheme and an online decision mechanism”,* Sensors, 23 (2023), 5436.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ

Ξενόγλωσσος όρος	**Ελληνικός Όρος**
Active Power	Ενεργή Ισχύς
Adaptive Functional-Coefficient Autoregression	Προσαρμοστική Παλινδρόμηση Συναρτησιακών Συντελεστών
Additive Models	Προσθετικά Μοντέλα
Adversarial Learning	Αντιπαραθετική Μάθηση
Aggregated load	Συγκεντρωμένο φορτίο
Aggregator	Φορέας Σωρευτικής Εκπροσώπησης
Alternating Direction Method of Multipliers	Μέθοδο εναλλασόμενων διευθύνσεων των πολλαπλασιαστών
Area networks	Δίκτυα περιοχής
Attentive transfer learning	Μάθηση προσθετικής μεταφοράς
Autoregression	Αυτοπαλινδρόμηση
Autoregressive Conditional Heteroscedasticity	Παλινδρομική Υπό Όρους Ετεροσκεδαστικότητα
Autoregressive Moving Average	Παλινδρόμηση με Κινούμενο Μέσο Όρο
Baseline model	Μοντέλο αναφοράς
Basis Pursuit	Αναζήτηση Βάσης
Batch Learning	Εκμάθηση σε Παρτίδες
Bilinear Models	Διγραμμικά Μοντέλα
Building Area Network	Δίκτυο κτιριακής περιοχής
Clustering	Κατηγοριοποίηση
Coefficient of Determination	Συντελεστής Συσχέτισης
Combined kernels	Συνδυασμένοι πυρήνες
Convolutional Neural Networks	Συνελικτικά Νευρωνικά Δίκτυα
Cross Validation	Διασταυρούμενη Αξιολόγηση
Deep Belief Networks	Δίκτυα Βαθιάς Πεποίθησης
Demand Response	Απόκριση ζήτησης
Demand Side Management	Διαχείριση ζήτησης
Dense average connection	Μέση πυκνή σύνδεση
Distributed Energy Management System	Σύστημα διαχείρισης διασπαρμένης ενέργειας
Distribution System Operator	Διαχειριστής Συστήματος Διανομής
Elastic net regularization	Ελαστική κανονικοποίηση
Electric Vehicles	Ηλεκτρικά οχήματα
Empirical Mode Decomposition	Μέθοδος εμπειρικής αποδόμησης
Energy Management System	Σύστημα διαχείρισης ενέργειας
Energy Storage System	Σύστημα αποθήκευσης ενέργειας
Ensemble	Ανσάμπλ
Event	Συμβάν
Exhaustive search	Εξαντλητική αναζήτηση
Expert Systems	Έμπειρα Συστήματα
Explainable Neural Networks	Ερμηνεύσιμα Νεωρωνικά Δίκτυα
Exploratory Data Analysis	Διερευνητική ανάλυση δεδομένων
Exponential Smoothing	Εκθετική Εξομάλυνση
Extrapolation	Παρεκβολή
Extreme learning machine	Μηχανή ακραίας εκμάθησης
Feedforward Neural Network	Νευρωνικό Δίκτυο Πρόσθιας Τροφοδότησης
Field Area Network	Δίκτυο πεδίου

Functional-coefficient Autoregression	Παλινδρόμηση Συναρτησιακών Συντελεστών
Fuzzy means algorithm	Αλγόριθμος ασαφών μέσων
Gaussian process	Γκαουσΐανή Διαδικασία
Global Positioning System	Παγκόσμια σύστημα στιγματοθέτησης
Gradient Boost Machine	Μηχανές Ενίσχυσης Κλίσης
Gradient boosting decision trees	Δένδρα αποφάσεων ενίσχυσης βάθμωσης
Graph Convolutional Networks	Συνελικτικά δίκτυα γράφων
Graph Neural Networks	Νευρωνικά Δίκτυα Γράφων
Greedy Algorithms	Άπληστοι Αλγόριθμοι
Hardware Accelerators	Επιταχυντές Υλικού
Home Management System	Σύστημα οικιακής διαχείρισης ενέργειας
Homotopy	Ομοτοπία
Household Area Network	Δίκτυο οικιστικής περιοχής
Image Visibility Graph	Γράφος Ορατότητας Εικόνας
Industrial Area Network	Δίκτυο βιομηχανικής περιοχής
Information and Communication Technologies	Τεχνολογίες πληροφοριών και επικοινωνιών
International Standards Organizations	Διεθνείς οργανισμοί προτύπων
Internet of Things	Διαδίκτυο των Πραγμάτων
Islanding	Νησιδοποίηση
K-nearest Neighbor Algorithm	Αλγόριθμος K-πλησιέστερων γειτόνων
Label	Ετικέτα
Labeled	Επισημασμένος
Latent variable	Λανθάνουσα μεταβλητή
Limited Penetrable Visibility Graph	Περιορισμένος Διαπερατός Γράφος Ορατότητας
Lines of Visibility	Γραμές Ορατότητας
Load curtailment	Περικοπή φορτίου
Load shifting	Μετακύλιση φορτίων
Local Functional-Coefficient Autoregression	Τοπική Παλινδρόμηση Συναρτησιακών Συντελεστών
Long Short-Term Memory	Δίκτυα Μακράς Βραχύχρονης Μνήμης
Markov Switching Autoregression	Παλινδρόμηση Μεταγωγής Markov
Mean Absolute Error	Μέσο Απόλυτο Σφάλμα
Mean Absolute Percentage Error	Μέσο Απόλυτο Ποσοστιαίο Σφάλμα
Mean Forecast Error	Μέσο Σφάλμα Πρόβλεψης
Mean Squared Error	Μέσο Τετραγωνικό Σφάλμα
Microgrid	Μικροδίκτυο
Moving Average	Κινούμενος Μέσος Όρος
Multi-agent System	Σύστημα πολλαπλών πρακτόρων
multi-layered	Πολλαπλών στοιβάδων
Multiple equation time-series	Χρονοσειρές πολλαπλών εξισώσεων
Multiplex Visibility Graph	Πολυμορφικός Γράφος Ορατότητας
Multivariate	Πολυμεταβλητό
Multivariate Local Polynomial Regression	Πολυμεταβλητή Τοπική Πολυωνυμική Παλινδρόμηση
Natural Visibility Graph	Γράφος Φυσικής Ορατότητας
Neighborhood Area Network	Δίκτυο γειτονιάς
Operator	Διαχειριστής
Outliers	Ακραίες τιμές
Overfitting	Υπερπροσαρμογή

Partial usage patterns	Μοτίβα μερικής χρήσης
Peak demand	Ζήτηση αιχμής
Peak load	Φορτίο αιχμής
Periodic Autoregressive with Exogenous Variables	Περιοδικη αυτοπαλινδόμιση με εξωγενείς μεταβλητές
Permutation	Μετάθεση
Phasor Measurement Unit	Μονάδων μέτρησης φασιθετών
Polynomial Autoregression	Πολυωνυμική Παλινδρόμηση
Principal Component Analysis	Ανάλυση Κύριων Συνιστοσών
Prosumers	Καταλανωτές και παραγωγοί
Purge	Εκκαθάριση
Quantile Regression	Παλινδρόμηση Ποσοστιαίων Σημείων
Quantiles	Ποσοστημόρια
Random Forests	Τυχαία Δάση
Rate of Change of Frequency	Ρυθμός αλλαγής συχνότητας
Realization	Πραγματοποίηση
Recommender Systems	Συστήματα Συστάσεων
Recurrent Neural Networks	Αναδρομικά Νευρωνικά Δίκτυα
Recursive feature elimination	Αναδρομική εξάλειψη χαρακτηριστικών
Regional Transmission Operators	Περιφερειακοί διαχειριστές μεταφοράς
Regression	Παλινδρόμηση
Reinforcement Learning	Ενισχυτική Μάθηση
Remote Terminal Unit	Απομακρυσμένη τερματική μονάδα
Residual Sum of Squares	Υπολειπόμενο Άθροίσμα Τετραγώνων
Resiliency	Ανθεκτικότητα
Resilient Backpropagation	Ανθεκτική Οπισθοδιάδοση
Root Mean Squared Error	Ρίζα Μέσου Τετραγωνικού Σφάλματος
Scalable Bayesian rule lists	Επεκτάσιμες μπεϋζιανές λίστες κανόνων
Seasonal Trend	Εποχική Τάση
Seasonal Variation	Εποχική Μεταβλητότητα
Self-healing	Αυτο-ίαση
Self-restoration	Αυτό-αποκατάσταση
Sequential Minimal Optimization	Διαδοχική Ελάχιστη Βελτιστοποίηση
Short-Term Load Forecasting	Βραχυπρόθεσμη πρόβλεψη φορτίου
Singular Value Decomposition	Παραγοντοποίηση Ιδιαζουσών Τιμών
Smart Grid	Εξυπνο Δίκτυο
Smooth Transition Autoregression	Παλινδρόμηση Ομαλής Μετάβασης
Sparse Bayesian Learning	Αραιή Μπεϋζιανή μάθηση
Stationarity	Στασιμότητα
Supervised Learning	Επιβλεπόμενη Μάθηση
Supervisory Control and Data Acquisition	Συστήματα εποπτικού ελέγχου και απόκτησης δεδομένων
Support Vector Regression	Παλινδρόμησης με χρήση μηχανών διανυσμάτων υποστήριξης
Synchrophasors	Συγχρονισμένες μετρήσεις φασιθέτη
Testing	Έλεγχος
Threshold Autoregression	Παλινδρόμηση Κατωφλίου
Time Evolving System	Χρονικά Εξελισσόμενο Σύστημα
Triple exponential smoothing	Τριπλή εκθετική εξομάλυνση
Tuple	Πλειάδα

Univariate	Μονομεταβλητό
Unlabeled	Μη Επισημασμένος
Unsupervised Learning	Μη Επιβλεπόμενη Μάθηση
Urban Intelligence	Αστική Νοημοσύνη
Visibility Graph	Γράφος Ορατότητας
Visibility Graph Neural Networks	Νευρωνικά Δίκτυα Γράφων Ορατότητας
Volatility	Μεταβλητότητα
Weighted	Σταθμισμένος
Wide Area Network	Δίκτυο ευρείας περιοχής

ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ

ADMM	Alternating Direction Method of Multipliers
AGG	AGGREGATE
AP	Active Power
AR	Autoregression
ARCH	Autoregressive Conditional Heteroscedasticity
ARIMA	Autoregressive Integrated Moving Average
ARIMAX	Autoregressive Integrated Moving Average with Exogenous variables
ARMA	Autoregressive Moving Average
BAN	Building Area Network
CNN	Convolutional Neural Networks
DBN	Deep Belief Networks
DERMS	Distributed Energy Management System
DR	Demand Response
DSM	Demand Side Management
DSO	Distribution System Operator
EMD	Empirical Mode Decomposition
EMS	Energy Management System
ESS	Energy Storage System
EVs	Electric Vehicles
FAN	Field Area Network
FAR	Functional-coefficient Autoregression
FFNN	Feedforward Neural Network
FFO	Fire-Fly Optimization
FISTA	Fast Iterative Shrinkage Thresholding
GBM	Gradient Boost Machine
GCN	Graph Convolutional Networks
GNN	Graph Neural Networks
GP	Gaussian Process
GPS	Global Positioning System
GWO	Grey Wolf Optimization

HAN	Household Area Network
HEMS	Home Management System
i.i.d.	Independent and identically distributed
IAN	Industrial Area Network
ICT	Information and Communication Technologies
IoT	Internet of Things
ISO	International Standards Organizations
ISTA	Iterative Shrinkage Thresholding Algorithm
IVG	Image Visibility Graph
KNN	K-nearest Neighbor
LARS	Least-Angle Regression
LASSO	Least Absolute Shrinkage and Selection Operator
LoV	Lines of Visibility
LPVG	Limited Penetrable Visibility Graph
LSTM	Long Short-Term Memory
MA	Moving Average
MAS	Multi-agent System
MLP	Multi-layer Perceptron
MSAR	Markov Switching Autoregression
MVG	Multiplex Visibility Graph
NAN	Neighborhood Area Network
NVG	Natural Visibility Graph
OMP	Orthogonal Matching Pursuit
PAR	Polynomial Autoregression
PARX	Periodic Autoregressive with Exogenous Variables
PCA	Principal Component Analysis
PMU	Phasor Measurement Unit
RBF	Radial Basis Function
RF	Random Forests
RNN	Recurrent Neural Networks
RoCoF	Rate of Change of Frequency
RPROP	Resilient Backpropagation
RSS	Residual Sum of Squares

RTO	Regional Transmission Operators
RTU	Remote Terminal Unit
SARIMA	Seasonal Autoregressive Integrated Moving Average
SBL	Sparse Bayesian Learning
SCADA	Supervisory Control and Data Acquisition
SMO	Sequential minimal optimization
STAR	Smooth Transition Autoregression
STLF	Short-Term Load Forecasting
SVD	Singular Value Decomposition
SVR	Support Vector Regression
TAR	Threshold Autoregression
UPD	UPDATE
VG	Visibility Graph
VGNN	Visibility Graph Neural Networks
WAN	Wide Area Network
XNN	Explainable Neural Networks
ΑΠΕ	Ανανεώσιμες Πηγές Ενέργειας
ΓΠ	Γραμμική Παλινδρόμηση
Δ.Δ.	Διδακτορική Διατριβή
ΔΣΔ	Διαχειριστής Συστήματος Διανομής
Ε.Ι.	Ενεργή Ισχύς
κ.α.	και άλλοι/άλλους
Μ.Μ.	Μηχανική Μάθηση
Τ.Ν.Δ.	Τεχνητά Νευρωνικά Δίκτυα
υ.τ.π.	Υπό τους περιορισμούς
Υ/Σ	Υποσταθμός
ΥΤ/ΜΤ	Υψηλή Τάση/Μέση Τάση
ΦΒ	Φωτοβολταϊκά συστήματα
ΦοΣΕ	Φορέας Σωρευτικής Εκπροσώπησης

1. ΕΙΣΑΓΩΓΗ

Μέσα από το παρόν εισαγωγικό κεφάλαιο, ο αναγνώστης πληροφορείται για το θέμα το οποίο πραγματεύεται η παρούσα μελέτη και τις βασικές ερευνητικές επιδιώξεις που αναμένωνται από αυτή. Συγκεκριμένα, οι δύο πρώτες ενότητες, ασχολούνται με τον καθορισμό του αντικειμένου της Δ.Δ. και των κινήτρων που οδήγησαν στην ερευνητική εμβάθυνση σε αυτό. Στην τρίτη ενότητα τίθονται οι ερευνητικοί στόχοι οι οποίοι θα αποτελέσουν τους βασικούς οδηγούς κατά την διεξαγωγής των ερευνητικών δραστηριοτήτων, ενώ τέλος παρουσιάζεται η δομή της Δ.Δ.

1.1 Αντικείμενο της διατριβής

Η ανάπτυξη μεθόδων μηχανικής μάθησης (Μ.Μ.) για την μοντελοποίηση χρονικά εξελισσόμενων συστημάτων αποτελεί το βασικό αντικείμενο της διδακτορικής διατριβής. Ως χρονικά εξελισσόμενα συστήματα νοούνται όσα τροφοδοτούνται και δημιουργούν σήματα τα οποία είναι συναρτήσεις του χρόνου. Η μαθηματική μοντελοποίηση ενός τέτοιου συστήματος συνίσταται στην εξαγωγή ενός συνόλου εξισώσεων που αναπαριστούν την συμπεριφορά του με την πάροδο του χρόνου. Η πλειοψηφία των συστημάτων που συναντώνται στον σύγχρονο κόσμο διαχειρίζονται δεδομένα σε μορφή χρονοσειρών. Ένας από τους βασικούς στόχους της ανάλυσης χρονοσειρών είναι η πρόβλεψη μελλοντικών τιμών τους, ένα πρόβλημα του οποίου η αντιμετώπιση αποτελεί πρωταρχική στόχευση της παρούσας έρευνας. Για το λόγο αυτό, στρεφόμαστε προς της αναζήτηση και υλοποίηση καινοτόμων μεθόδων Μ.Μ. με και χωρίς επίβλεψη, επιτυγχάνοντας την συγγραφή και την δημοσίευση τριών επιστημονικών άρθρων. Το πεδίο εφαρμογής των μεθόδων που αναπτύχθηκαν στα πλαίσια της εκπόνησης της Δ.Δ., αποτέλεσαν τα έξυπνα ηλεκτρικά δίκτυα και ειδικότερα η βραχυπρόθεσμη και μεσοπρόθεσμη πρόβλεψη του ηλεκτρικού φορτίου. Το πρόβλημα που επιχειρούμε να αντιμετωπίσουμε μέσα από την ανάπτυξη και την εφαρμογή των προτεινόμενων μεθόδων είναι αυτό της πρόβλεψης του ηλεκτρικού φορτίου στο επίπεδο της μέσης τάσης του δικτύου ηλεκτρικής ισχύος. Κύρια επιδίωξη αποτελεί η παροχή μοντέλων πρόβλεψης αξιόπιστων, εύρωστων και πιο αποδοτικών έναντι τυπικών μεθόδων Μ.Μ. Για το λόγο αυτό, η ανωτερότητα των προτεινόμενων μοντέλων αξιολογείται εξετάζοντας την ικανότητά τους να διαχειρίζονται ιδιότητες των χρονοσειρών του ηλεκτρικού φορτίου που δυσχεραίνουν την ακρίβεια των προβλέψεων, όπως η μη γραμμικότητα και η αβεβαιότητα. Τα πλεονεκτήματα και η σημασία της παροχής αξιόπιστων προβλέψεων του φορτίου για την βέλτιστη λειτουργία και διαχείριση των έξυπνων δικτύων, επιβεβαιώνεται πέρα από την πειραματική αξιολόγηση των προτεινόμενων μοντέλων, από δύο επιπλέον δημοσιεύσεις που πραγματοποιήθηκαν στα πλαίσια της ερευνητικής δραστηριότητας.

1.2 Κίνητρα

Είναι γεγονός πως οι επιστημονικοί κλάδοι της Μ.Μ. και της υπολογιστικής νοημοσύνης έχουν προσελκύσει τεράστιο ερευνητικό ενδιαφέρον κατά τις τελευταίες δεκαετίες, με εφαρμογές σε πάμπολλους τομείς μεταξύ των οποίων και αυτός της ηλεκτρικής ενέργειας. Η μετάβαση από τα συμβατικά στα έξυπνα δίκτυα αποτελεί αναμφισβήτητα μια από τις πιο ριζικές αλλαγές στον τομέα των ηλεκτρικών δικτύων, προσφέροντας πολλά πρεονεκτήματα και προοπτικές σε θέματα που αφορούν στην εξοικονόμηση ενέργειας, στην προστασία του περιβάλλοντος, στα οικονομικά ωφέλη των καταναλωτών, στην ομαλή και ασφαλή παροχή ισχύος και πολλά ακόμη. Οι παράγοντες αυτοί συνθέτουν ένα σημαντικό και επαρκές κίνητρο για την ενασχόληση σε ερευνητικό επίπεδο, με την πρόβλεψη του ηλεκτρικού φορτίου, η οποία αποτελεί μί από τις βασικές συνιστώσες για την αποδοτική λειτουργία και την πλήρη αξιοποίηση των δυνατοτήτων των έξυπνων δικτύων.

Από μεθοδολογικής άποψης, το πρόβλημα της πρόβλεψης του ηλεκτρικού φορτίου έχει αντιμετωπιστεί με μια πληθώρα μεθόδων μηχανικής μάθησης, που κυμαίνονται από απλούστερα μοντέλα που βασίζονται στη γραμμική παλινδρόμηση – ΓΠ (Linear Regression - LR) και την παλινδόμηση με χρήση διανυσμάτων υποστήριξης (Support Vector Regression - SVR), έως πιο προηγμένα που χρησιμοποιούν νευρωνικά δίκτυα και βαθιά μάθηση [1], [2]. Παρά την πολυετή ενδελεχή μελέτη, η πρόβλεψη του ηλεκτρικού φορτίου παραμένει ένα ανοιχτό πεδίο ενεργούς έρευνας. Ο κύριος λόγος πίσω από αυτό, είναι ότι πρόκειται για ένα σύνθετο πρόβλημα το οποίο είναι δύσκολο να αντιμετωπιστεί επαρκώς, λόγω της ιδιότυπης φύσης των χρονοσειρών του ηλεκτρικού φορτίου. Ως αποτέλεσμα, δεν υπάρχει συγκεκριμένη μέθοδος Μ.Μ. που να επιτυγχάνει καθολική υπεροχή όσον αφορά την προβλεπτική τους απόδοση. Για παράδειγμα, τα μοντέλα που βασίζονται σε νευρωνικά δίκτυα έχει φανεί ότι μπορούν να διαχειριστούν ικανοποιητικά τη μη γραμμική συμπεριφορά του φορτίου, αλλά ταυτόχρονα είναι ευαίσθητα σε δεδομένα εισόδου που ενέχουν θόρυβο. Από την άλλη πλευρά, τα απλά γραμμικά μοντέλα πρόβλεψης παρέχουν πολύ μεγαλύτερη ευρωστία, παρόλο που δεν είναι σε θέση να αποτυπώσουν πιο περίπλοκες, μη γραμμικές δυναμικές. Με βάση αυτή την παρατήρηση, η ανάπτυξη εναλλακτικών μέθοδων που να επιτυγχάνουν μια ισομερή αντιστάθμιση μεταξύ των προαναφερθέντων πλεονεκτημάτων, ανάγεται σε επιτακτική ανάγκη και αποτέλεσε μια βασική κινητήρια δύναμη για την εκπόνηση της παρούσας Δ.Δ.

1.3 Ερευνητικοί στόχοι

Ένα πρωταρχικό στάδιο κατά τη διεξαγωγή μιας ερευνητικής εργασίας είναι ο προσδιορισμός των ερευνητικών στόχων οι οποίοι θα απετελέσουν τις κατευθυντήριες γραμμές και θα κρίνουν το ερευνητικό αντίκτυπου του συνολικού έργου. Οι στόχοι που τέθηκαν στα πλάισια της παρούσας Δ.Δ. μπορούν κάλλιστα να προκύψουν από τα ακόλουθα ερευνητικά ερωτήματα:

- Μπορούν να χρησιμοποιηθούν με επιτυχία αποδοτικοί αλγόριθμοι μηχανικής μάθησης σε προβλήματα πρόβλεψης χρονοσειρών και να εφαρμοστούν σε συγκεκριμένα προβλήματα, όπως η πρόβλεψη του ηλεκτρικού φορτίου;
- Μπορούν να προταθούν νέοι αλγόριθμοι μηχανικής μάθησης σε προβλήματα πρόβλεψης χρονοσειρών και να εφαρμοστούν με επιτυχία σε συγκεκριμένα προβλήματα, όπως η πρόβλεψη του ηλεκτρικού φορτίου;

1.4 Δομή της διατριβής

Τα κεφάλαια της Δ.Δ. που ακολουθούν δομούνται ως εξής. Το 2^{ο} κεφάλαιο πραγματεύεται καταρχάς το θεωρητικό υπόβαθρο που πλαισιώνει την Δ.Δ. Επιπλέον, περιγράφονται ενδελεχώς το θεωρητικό μέρος των καινοτόμων μεθοδολογιών Μ.Μ. που αναπτύχθηκαν. Στο 3^{ο} κεφάλαιο παρέχεται μια εισαγωγή στα έξυπνα δίκτυα και περιγράφεται αναλυτικά το πειραματικό πρωτόκολλο των προσομοιώσεων που πραγματοποιήθηκαν με σκοπό την αξιολόγηση της απόδοσης των μεθόδων που υλοποιήθηκαν. Στη συνέχεια, στο κεφάλαιο 4, ακολουθεί η έκθεση των αποτελεσμάτων που παράχθηκαν από αυτές, ενώ στο 5^{ο} κεφάλαιο παρουσιάζονατι τα συμπεράσματα που εξάγονται από την έρευνα αλλά και οι περαιτέρω ερευνητικές κατευθύνσεις που μπορούν να προκύψουν από αυτήν.

2. ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΣΥΣΤΗΜΑΤΩΝ

Στο κεφάλαιο αυτό παρουσιάζεται, σε κάποια έκταση, το θεωρητικό υπόβαθρο που απαιτείται για τη διεξαγωγή της έρευνας και παρουσιάζονται λεπτομερώς τα θεωρητικά επιτεύγματα που προέκυψαν από την μελέτη του. Αρχικά, στην παράγραφο 2.1, παρέχονται οι ορισμοί των βασικών εννοιών, όπως των χρονικά εξελισσόμενων συστημάτων, και της μοντελοποίησης συστημάτων. Ακολουθεί η παρουσίαση των δύο κυρίων μεθόδων προσέγγισης της μοντελοποίησης: της θεωρητική και της πειραματικής. Ιδιαίτερη έμφαση δίνεται στην περίπτωση της πειραματικής προσέγγισης, η οποία αποτελεί το θεμέλιο λίθο στο οποίο εδράζει η παρούσα έρευνα. Η διερεύνηση των στόχων της ανάλυσης και της μοντελοποίησης των χρονοσειρών καθώς και η διαδικασία προ-επεξεργασίας τους παρατίθεται στην παράγραφο 2.2.1. Επίσης, η περιγραφή των βημάτων και των κύριων κατηγοριών μεθόδων πρόβλεψης χαρακτηριστικών από χρονοσειρές παρατίθεται στην παράγραφο 2.2.2. Στην παράγραφο 2.3, ακολουθεί μια σύντομη εισαγωγή στην έννοια της μηχανικής μάθησης και της εφαρμογής της στην πρόβλεψη χαρακτηριστικών από χρονοσειρές. Τέλος, το κεφάλαιο αυτό ολοκληρώνεται με την λεπτομερή παρουσίαση των μεθόδων επιβλεπόμενης και μη επιβλεπόμενης μάθησης που αναπτύχθηκαν στα πλαίσια της παρούσας έρευνας, (παράγραφοι 2.3.1 και 2.3.2 αντίστοιχα).

2.1 Χρονικά εξελισσόμενα συστήματα

Κάθε διεργασία η οποία δρα σε ένα σήμα $x(\cdot)$ και παράγει ως αποτέλεσμα ένα νέο σήμα $y(\cdot)$, καλείται σύστημα. Ως σήμα ορίζεται οποιαδήποτε μεταβολή ενός φυσικού μεγέθους σε συνάρτηση με κάποια ανεξάρτητη μεταβλητή, η οποία συνήθως είναι ο χρόνος (t) για την μονοδιάστατη περίπτωσης. Ένα σύστημα το οποίο δέχεται και παράγει σήματα τα οποία είναι συναρτήσεις του χρόνου, ονομάζεται χρονικά εξελισσόμενο σύστημα (time evolving system), αποκαλούμενο εφεξής και ως σύστημα. Από μαθηματικής σκοπιάς, ένα τέτοιο σύστημα μπορεί να θεωρηθεί ως ένας μετασχηματισμός $S[\cdot]$, που μετασχηματίζει ένα σήμα $x(t)$ σε ένα άλλο σήμα $y(t)$, σύμφωνα με τη σχέση $y(t) = S[x(t)]$. Με βάση αυτή την προσέγγιση, το σήμα $x(t)$ που διεγείρει το σύστημα καλείται είσοδος ή διέγερση του συστήματος, ενώ το αποτέλεσμα της διαδικασίας διέγερσης, δηλαδή το σήμα $y(t)$ λέγεται έξοδος ή απόκριση του συστήματος. Η λειτουργία των χρονικά εξελισσόμενων συστημάτων, φυσικών ή τεχνητών, τα οποία απαντώνται σε μια σειρά από διάφορα επιστημονικά πεδία όπως της ηλεκτρολογίας [3], [4], της μηχανολογίας [5], [6], της μηχανικής διεργασιών [7], [8] και των βιοεπιστημών [9]. Τα πεδία αυτά ενδέχεται να εμπίπτουν σε ποικίλους τομείς της επιστήμης όπως η βιολογία, η ιατρική, η χημεία, η φυσική, η οικονομία, μπορεί να γίνει ευρύτερα κατανοητή με τη βοήθεια μαθηματικών μοντέλων που προσπαθούν να προσεγγίσουν τους αντίστοιχους μετασχηματισμούς $S[\cdot]$.

Η ανάπτυξη τέτοιων μοντέλων αποτελεί αντικείμενο του επιστημονικού κλάδου της θεωρίας συστημάτων. Πρόκειται για μια μαθηματική περιγραφή η οποία βασίζεται στον απειροστικό λογισμό και έχει σαν στόχο την μελέτη και την εύρεση τρόπων αναφοράς και συσχέτισης ορισμένων χαρακτηριστικών ενός συστήματος, όπως για παράδειγμα οι μεταβλητές, η είσοδος και η έξοδός του. Τα συστήματα που συναντώνται στην καθημερινή ζωή και τα οποία τα μοντέλα καλούνται να περιγράψουν, ενδέχεται να διαφέρουν ως προς διάφορες παραμέτρους. Αρχικά, ανάλογα με το πλήθος των σημάτων εισόδου και εξόδου, τα συστήματα διακρίνονται σε (α) μονής εισόδου/μονής εξόδου (single input/single output – SISO), (β) πολλαπλών εισόδων/μονής εξόδου (multiple input/single output – MISO), (γ) μονής εισόδου/πολλαπλών εξόδων (single input/multiple output – SIMO) και (δ) πολλαπλών εισόδων/πολλαπλών εξόδων (multiple input/multiple output – MIMO). Ακόμη, τα σήματα εισόδου και εξόδου δύναται να είναι συνεχούς ή διακριτού χρόνου, ορίζοντας αντίστοιχα συστήματα συνεχούς (αναλογικά) ή διακριτού (ψηφιακά) χρόνου.

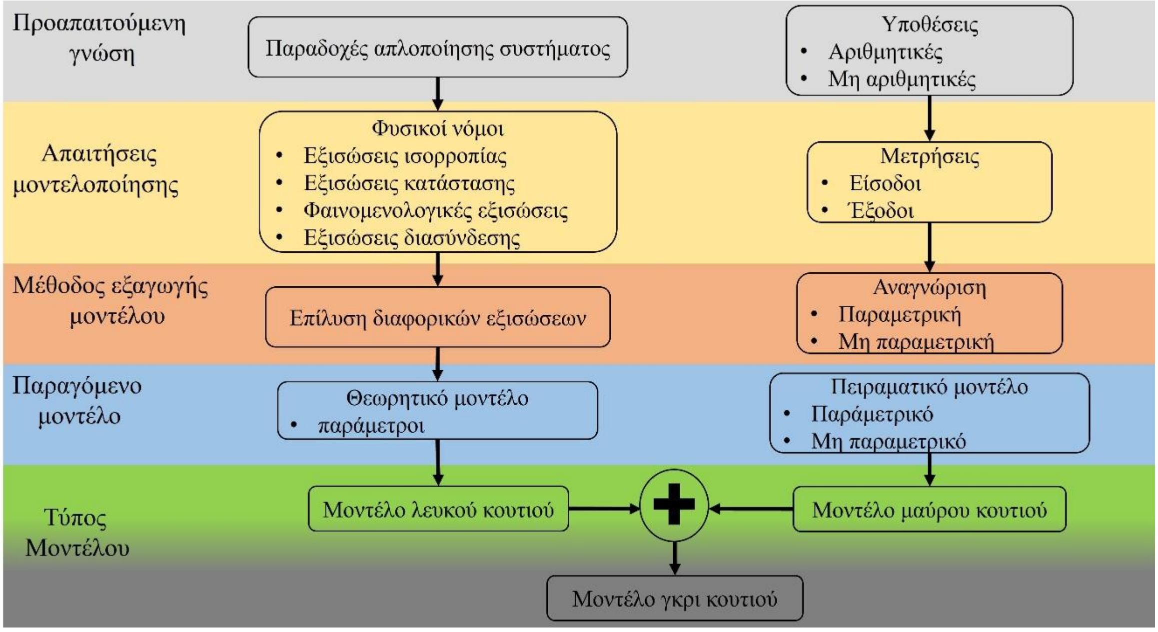
Έχει παρατηρηθεί ότι ο προσδιορισμός των τιμών σε κάθε χρονική στιγμή των σημάτων εισόδου και εξόδου ορισμένων συστημάτων υλοποιείται από μια αναλυτική μαθηματική έκφραση. Ένα τέτοιο σύστημα χαρακτηρίζεται ως αιτιοκρατικό. Από την άλλη, τα συστήματα στα οποία οι τιμές των σημάτων εισόδου και εξόδου εμφανίζονται με μια τυχαιότητα, δηλαδή ακολουθούν μια κατανομή, καλούνται στοχαστικά. Πέρα από το πλήθος και τη φύση των σημάτων εισόδου και εξόδου, τα συστήματα μπορούν να κατηγοριοποιηθούν και ως προς τον τρόπο συσχέτισης των σημάτων αυτών. Ένα σύστημα όπου η/οι έξοδος/οι για κάθε χρονική στιγμή εξαρτάται μόνο από την/τις τιμή/ές της/των εισόδου/ων την ίδια χρονική στιγμή και όχι από προηγούμενες τιμές της/τους, ονομάζεται στατικό ή σύστημα χωρίς μνήμη. Εάν ένα σύστημα δεν είναι στατικό τότε χαρακτηρίζεται ως δυναμικό ή σύστημα με μνήμη. Ανεξάρτητα από το είδος και την κατηγορία που ανήκει το οποιοδήποτε σύστημα, η ανάπτυξη ενός μοντέλου που το περιγράφει δεν μπορεί να πραγματοποιηθεί χωρίς την ύπαρξη επαρκούς μεγέθους σημάτων εισόδου και εξόδου.

Είναι γεγονός πως η πλειονότητα των σύγχρονων και προηγμένων τεχνολογιών και υπηρεσιών, με εφαρμογές μεταξύ άλλων στο Διαδίκτυο των Πραγμάτων (IoT) [10], [11] στο ψηφιακό μάρκετινγκ [12], [13], [14] και άλλων [15], [16] δημιουργούν, συλλέγουν, αποθηκεύουν και ανταλλάσσουν έναν τεράστιο όγκο χρονικά εξαρτώμενων δεδομένων. Μία από τις βασικότερες αξιώσεις της επιστήμης δεδομένων και της μηχανικής δεδομένων είναι η αποτελεσματική και αποδοτική διεξαγωγή εργασιών ανάλυσης ακολουθιών δεδομένων μεγάλου μήκους, όπως η εξαγωγή μακροπρόθεσμων προβλέψεων για τις τιμές εξόδου του συστήματος σε πραγματικό χρόνο χωρίς ανθρώπινη παρέμβαση [27], [28]. Στην πράξη, οι ακολουθίες πραγματικών δεδομένων φαίνονται να υπόκεινται σε πιθανοτικούς κανόνες οι οποίες δύνανται να μεταβάλλονται με το χρόνο. Αυτό υποδεικνύει ότι κάποιοι υποκείμενοι κανόνες συσχέτισης και μοτίβα συμπεριφορών που διέπουν τις ακολουθίες δεδομένων βρίσκονται σε άμεση εξάρτηση με το χρόνο. Γίνεται λοιπόν κατανοητό ότι προκειμένου να εκτιμήσουμε μελλοντικές αποκρίσεις ενός συστήματος με βάση τις ακολουθίες δεδομένων εισόδου και εξόδου, είναι απαραίτητο να επιχειρήσουμε την ανίχνευση αυτών των μοτίβων και των αλληλεπιδράσεών τους.

Η εφαρμογή της θεωρίας των συστημάτων απαιτεί την πρότερη γνώση για τα μοντέλα που περιγράφουν τη συμπεριφορά των συστημάτων και των στοιχείων τους. Η διαδικασία της κατασκευής ενός συνόλου εξισώσεων που αναπαριστούν με την υψηλότερη δυνατή ακρίβεια την εξέλιξη του υπό εξέταση συστήματος, ονομάζεται μαθηματική μοντελοποίηση συστημάτων. Όπως περιγράφεται στη συνέχεια (βλ. Σχήμα 1), υπάρχουν δύο γενικές προσεγγίσεις για τη μοντελοποίηση συστημάτων, η θεωρητική και η πειραματική μοντελοποίηση [17], καθεμία από τις οποίες έχει τα δικά της πλεονεκτήματα και μειονεκτήματα. Στην περίπτωση της θεωρητικής προσέγγισης, το μοντέλο προκύπτει από την εφαρμογή μεθόδων λογισμού στις χαρακτηριστικές εξισώσεις οι οποίες διέπουν την λειτουργία του [18]. Συνήθως εφαρμόζονται απλοποιητικές υποθέσεις (π.χ γραμμικότητα) σχετικά με το σύστημα και/ή τη διαδικασία, καθώς μόνο τότε είναι εφικτή η αντιμετώπιση του προβλήματος μέσω της μαθηματικής προτυποποίησης στις περισσότερες περιπτώσεις. Κατά κανόνα, για να δημιουργηθεί το μοντέλο συνδυάζονται εξισώσεις ισορροπίας, φυσικές ή χημικές εξισώσεις κατάστασης, φαινομενολογικές εξισώσεις και εξισώσεις διασύνδεσης [19]. Εφαρμόζοντας αυτές τις εξισώσεις, καταλήγει κανείς σε ένα σύνολο συνήθων ή μερικών διαφορικών εξισώσεων, το οποίο με την σειρά του οδηγεί σε ένα θεωρητικό μοντέλο με μια συγκεκριμένη δομή και ορισμένες παραμέτρους. Ακόμη όμως και στην περίπτωση που το επαγόμενο σύνολο εξισώσεων δεν μπορεί να λυθεί αναλυτικά, οι μεμονωμένες εξισώσεις εξακολουθούν να παρέχουν σημαντικές υποδείξεις σχετικά με τη δομή του μοντέλου.

Στην περίπτωση της πειραματικής προσέγγισης, η οποία ονομάζεται επίσης αναγνώριση, εξάγεται ένα μαθηματικό μοντέλο από τις παραγόμενες μετρήσεις [20]. Εδώ, συνήθως πρέπει κανείς να βασιστεί εκ των προτέρων σε κάποιες υποθέσεις, οι οποίες μπορεί να προέρχονται είτε από θεωρητική ανάλυση είτε από προηγούμενες (αρχικές) πειραματικές μετρήσεις. Οι υποθέσεις αυτές προσδιορίζουν τις ελλείπουσες πληροφορίες καθώς και τις αρχικές ή οριακές συνθήκες ή/και τους περιορισμούς που αφορούν τις μεταβλητές του συστήματος. Διακρίνονται δύο κατηγορίες υποθέσεων, οι αριθμητικές και οι μη αριθμητικές, εκ των οποίων οι μεν αναφέρονται αντίστοιχα στον καθορισμό τιμών άγνωστων ποσοτήτων και οι δε σε παραδοχές σχετικά με τη φύση του συστήματος, για παράδειγμα μια αναμενόμενη συμπεριφορά για δεδομένη είσοδο στο σύστημα. Το επόμενο βήμα είναι η λήψη των μετρήσεων για τις τιμές της εισόδου και της εξόδου και η εφαρμογή κάποιας μεθόδου αναγνώρισης προκειμένου να προσδιοριστούν οι παράμετροι του μοντέλου που περιγράφει τη σχέση μεταξύ τους. Τα σήματα εισόδου μπορεί να προέρχονται από πραγματικά δεδομένα τα οποία τροφοδοτούν το σύστημα, είτε μπορεί να είναι τεχνητά κατασκευασμένα σήματα με συγκεκριμένες προκαθορισμένες ιδιότητες που παράγονται με τεχνικές προσομοίωσης.

Ανάλογα με τον τύπο της εφαρμογής, μπορούν να αναπτυχθούν παραμετρικά ή μη παραμετρικά μοντέλα [21]. Τα παραμετρικά μοντέλα υποθέτουν την απλοποίηση της σχέσης μεταξύ εισόδου και εξόδου του συστήματος σε μια συνάρτηση με γνωστή μορφή, όπως για παράδειγμα η γραμμική πολυωνυμική συνάρτηση. Το πλήθος των παραμέτρων αυτών είναι συγκεκριμένο, δηλαδή το μοντέλο γνωρίζει ήδη τον αριθμό των παραμέτρων που απαιτεί, ανεξάρτητα από τα δεδομένα του. Οι παράμετροι είναι επίσης ανεξάρτητες από τον αριθμό των δεδομένων που εισάγονται κατά τη διαδικασία της αναγνώρισης. Αλγόριθμοι οι οποίοι δεν βασίζονται σε συγκεκριμένες υποθέσεις σχετικά με το είδος της συνάρτησης αντιστοίχισης μεταξύ των δεδομένων εισόδου και εξόδου, είναι γνωστοί ως μη παραμετρικοί αλγόριθμοι και τα μοντέλα που προκύπτουν χαρακτηρίζονται ως μη παραμετρικά μοντέλα. Ένα βασικό χαρακτηριστικό των αλγορίθμων αυτών είναι ότι απαιτούν σημαντικά μεγαλύτερο αριθμό δεδομένων σε σχέση με τους παραμετρικούς, για την εκτίμηση του στόχου, δηλαδή μιας συνάρτησης αντιστοίχισης. Μια επιπλέον διαφοροποίηση μεταξύ των δύο αναφερόμενων τεχνικών είναι ότι στην περίπτωση των μη παραμετρικών αλγορίθμων το πλήθος των παραμέτρων δεν είναι σταθερό αλλά μεταβάλλεται αναλογικά με το πλήθος των δεδομένων, πράγμα που σημαίνει ότι είναι υπολογιστικά πιο απαιτητικοί.



Σχήμα 1: Τα βήματα της διαδικασίας μοντελοποίησης συστημάτων.

Κατά την μοντελοποίηση συστημάτων, εκτός από την αποκλειστική χρήση της θεωρητικής ή της πειραματικής προσέγγισης, οδηγώντας στα λεγόμενα μοντέλα λευκού [22] ή μαύρου [23] κουτιού αντίστοιχα, συχνά ακολουθείται μια μίξη των δύο, καταλήγοντας στην απόκτηση των λεγόμενων μοντέλων γκρι κουτιού [24]. Η επιλογή της προσέγγισης μοντελοποίησης καθορίζεται από το πεδίο εφαρμογής του μοντέλου και από το ίδιο το σύστημα. Παρά το γεγονός ότι, υπό την προϋπόθεση ότι η εσωτερική συμπεριφορά είναι γνωστή και μπορεί να περιγραφεί μαθηματικά, η θεωρητική ανάλυση μπορεί εξ αρχής να παρέχει περισσότερες πληροφορίες σχετικά με το σύστημα, εντούτοις η πειραματική ανάλυση λαμβάνει αυξανόμενη προτίμηση την τελευταία πενήνταετία [25], [26], για τους κάτωθι λόγους:

- Η θεωρητική ανάλυση μπορεί να γίνει αρκετά περίπλοκη ακόμα και για απλά συστήματα.
- Συνήθως, οι συντελεστές του μοντέλου που προέρχονται από θεωρητικές αναλύσεις δεν είναι αρκετά ακριβείς, καθώς η εξαγωγή τους βασίζεται αποκλειστικά σε εξισώσεις που απορρέουν από φυσικούς νόμους αγνοώντας όμως τις εσωτερικές αλληλεπιδράσεις του συστήματος. Ως αποτέλεσμα, τα μοντέλα που προκύπτουν, αδυνατούν να αναπαραστήσουν τη δυναμική συμπεριφορά του συστήματος.
- Δεδομένου ότι η μοντελοποίηση στηρίζεται στην αξιοποίηση των δεδομένων εισόδου και εξόδου, οι διεργασίες που λαμβάνουν χώρα μέσα στο σύστημα δεν αναγνωρίζονται, μειώνοντας έτσι σε μεγάλο βαθμό την απαίτηση για βαθιά γνώση της θεωρίας που διέπει το εκάστοτε σύστημα.
- Οι ενέργειες που λαμβάνουν χώρα δεν μπορούν να περιγραφούν μαθηματικά με την απαιτούμενη ακρίβεια.
- Κάποια συστήματα είναι πολύπλοκα, καθιστώντας τη θεωρητική ανάλυση χρονοβόρα.
- Μοντέλα αναγνώρισης μπορούν να αποκτηθούν σε μικρότερο χρονικό διάστημα και με λιγότερη προσπάθεια σε σύγκριση με τη θεωρητική μοντελοποίηση.

Σε σχέση με την θεωρητική, η πειραματική ανάλυση επιτρέπει την ανάπτυξη υπολογιστικών μοντέλων μέσω της μέτρησης της εισόδου και της εξόδου των συστημάτων, ανεξάρτητα από τον τρόπο με τον οποίο αντιστοιχίζονται μεταξύ τους [27]. Ένα πλεονέκτημά της είναι το γεγονός ότι οι ίδιες μέθοδοι πειραματικής ανάλυσης μπορούν να εφαρμοστούν με την ίδια επιτυχία σε συστήματα ανεξαρτήτως του επιστημονικού κλάδου από τον οποίο προέρχονται και της πολυπλοκότητάς τους. Από την άλλη πλευρά, επειδή η μοντελοποίηση βασίζεται στην μέτρηση μόνο της εισόδου και της εξόδου, τα μοντέλα κατά κανόνα δεν είναι ικανά να περιγράφουν την ακριβή εσωτερική δομή του συστήματος. Τα προσεγγιστικά αυτά μοντέλα εξακολουθούν να θεωρούνται επαρκή σε διάφορους τομείς εφαρμογής για τους ακόλουθους λόγους:

- Μια μεγάλη ποικιλία σύγχρονων εφαρμογών, συμπεριλαμβανομένων οικονομικών εφαρμογών (π.χ. τιμές μετοχών [28]), της παρακολούθησης δικτύου (π.χ. καταγραφές κίνησης όγκου δεδομένων [29]), της κυβερνοασφάλειας (π.χ. καταχωρήσεις αυτοδιαγνωστικών δεδομένων των συσκευών του δικτύου [30]), της διαχείριση τηλεπικοινωνιακών δεδομένων (π.χ. εγγραφές κλήσεων [31]), των ηλεκτρικών δικτύων (π.χ. μετρήσεις ζήτησης ηλεκτρικού φορτίου [32]), τις ιστοσελίδες (π.χ. επισκέψεις σε ιστοσελίδες [33]), των δικτύων αισθητήρων (π.χ. αναγνώσεις αισθητήρων [34]) κ.λ.π., ενσωματώνουν συστήματα τα οποία δέχονται σαν είσοδο και εξάγουν ακολουθίες δεδομένων. Οι εξελίξεις των τεχνολογιών πληροφοριών και επικοινωνιών έχουν αλλάξει δραματικά τις μεθόδους συλλογής και επεξεργασίας δεδομένων. Αυτό που διακρίνει τα σύγχρονα συστήματα σε σχέση εκείνα των προηγούμενων δεκαετιών είναι η αυτόματη τροφοδοσία δεδομένων. Η μεταφορά της πληροφορίας δεν γίνεται πλέον με την εισαγωγή δεδομένων από ανθρώπους σε υπολογιστές αλλά με την ανταλλαγή δεδομένων μεταξύ υπολογιστών [35], [36]. Επιπλέον, η πρόοδος στην τεχνολογία αισθητήρων έχει οδηγήσει σε δίκτυα αισθητήρων, που συλλέγουν υψηλής λεπτομέρειας δεδομένα τόσο στον χώρο – τοπικά όσο και στον χρόνο [37], [38]. Σε αυτό το πλαίσιο οι ακολουθίες δεδομένων θέτουν νέους περιορισμούς κατά τη σχεδίαση αλγορίθμων εκμάθησης. Σύμφωνα με τους παραδοσιακούς αλγορίθμους εκμάθησης σε παρτίδες (batch learning), το τελικό μοντέλο προκύπτει έπειτα από επαναληπτική εκπαίδευση με διαφόρων, συνήθως μικρού μεγέθους, υποσυνόλων δεδομένων. Στην προσέγγιση αυτή, απαιτείται ολόκληρο το σύνολο δεδομένων εκπαίδευσης να είναι διαθέσιμο εξ αρχής και η εκπαίδευση διεξάγεται συνήθως σε offline περιβάλλον εξαιτίας της υψηλής υπολογιστικής πολυπλοκότητας τόσο σε χρόνο όσο και χώρο. Ένα επιπλέον μειονέκτημα των μεθόδων αυτών είναι ακόμη η αδυναμία επεκτασιμότητας για εφαρμογές οπούς απαιτείται η επανεκπαίδευση του μοντέλου με καινούρια δεδομένα από την αρχή. Ωστόσο, συνδυαστικές λύσεις μηχανικής μάθησης σε παρτίδες και online, έχουν προταθεί με μεγάλη επιτυχία σε πρακτικές εφαρμογές [39], [40], [41].
- Οι ακολουθίες δεδομένων χαρακτηρίζονται από ανομοιομορφία ως προς το μέγεθος και τις τιμές τους σε συνάρτηση με το χρόνο εμφάνισής τους, γεγονός που γεννά νέες απαιτήσεις από τις μεθόδους ανάλυσης δεδομένων. Όπως είναι γνωστό από την θεωρία πιθανοτήτων, ένα σύνολο τυχαίων μεταβλητών είναι ανεξάρτητες και ταυτόσημα κατανεμημένες (i.i.d) αν όλες οι τυχαίες μεταβλητές έχουν κοινή κατανομή πιθανότητας και είναι ανεξάρτητες μεταξύ τους. Ένα πολύ χαρακτηριστικό παράδειγμα i.i.d. τυχαίων μεταβλητών είναι αυτό της ακολουθίας των ρίψεων ενός δίκαιου νομίσματος. Το αποτέλεσμα του ενδεχομένου κάθε ρίψης (κορώνα ή γράμματα) δεν επηρεάζεται από το αποτέλεσμα καμιάς εκ των προηγούμενων ρίψεων. Συνεπώς, οι ρίψεις είναι ανεξάρτητες μεταξύ τους. Επιπλέον, κάθε ενδεχόμενο έχει πάντα 50% πιθανότητα να συμβεί, επομένως οι μεταβλητές είναι ταυτόσημα κατανεμημένες. Η τυπική υπόθεση ότι τα δεδομένα είναι ανεξάρτητα και ταυτόσημα κατανεμημένα (i.i.d.) δεν ισχύει απαραίτητα στην περίπτωση των ακολουθιών δεδομένων. Αναδεικνύεται επομένως η ανάγκη για ένα

νέο είδος ανάλυσης των χρονικά εξελισσόμενων συστημάτων η οποία δεν επικεντρώνεται σε ένα συγκεκριμένο πρόβλημα (π.χ. μόνο ταξινόμηση), αλλά περιλαμβάνει μια ποικιλία προβλημάτων προκειμένου να ανταποκριθεί στις ανάγκες των εφαρμογών (όπως η στατιστική περιγραφή, η πρόβλεψη χαρακτηριστικών, η αναγνώριση προτύπων κ.λ.π.) που εμπλέκουν ακολουθίες δεδομένων με ενδεχομένως διαφορετικές ιδιότητες.

- Εκμεταλλευόμενα τους συσχετισμούς που απορρέουν από τις ακολουθίες δεδομένων πραγματοποιούν προβλέψεις τόσο με βραχυπρόθεσμους όσο και με μακροπρόθεσμους χρονικούς ορίζοντες.
- Λαμβάνουν υπόψιν παράγοντες αβεβαιότητας μέσα από την εξαγωγή πολλαπλών προβλέψεων και την αξιολόγηση της πιθανότητας κάθε μίας από αυτές να ανταποκρίνεται στην πραγματική μελλοντική τιμή.
- Προσαρμόζονται συνεχώς σε μεταβολές, μιας και σε ένα ασταθές σύστημα όπου τα δεδομένα συχνά έχουν μια στατιστική που εξελίσσεται η ικανότητα αγνόησης παλαιότερων δεδομένων αποδεικνύεται εξόχως σημαντική. Με τον τρόπο αυτό αποφεύγεται η μοντελοποίηση με δεδομένα που θα οδηγήσουν σε μη αξιόπιστες προβλέψεις. Για παράδειγμα, έστω ένα μοντέλο πρόβλεψης της τιμής μιας μετοχής που εκπαιδεύεται με δεδομένα που αντιστοιχούν στην προηγούμενη δεκαετία. Κατά πάσα πιθανότητα, οι προβλέψεις που θα παραχθούν, από το εν λόγω μοντέλο, για την τιμή της μετοχής στο άμεσο μέλλον ενδέχεται να παρουσιάσουν υψηλό σφάλμα πρόβλεψης, εξαιτίας των ραγδαίων μεταβολών των παραγόντων οι οποίοι επηρεάζουν την τιμή της μετοχής, κατά το διάστημα αυτών των ετών.
- Χαρακτηρίζονται από ικανότητα εκμάθησης, στο ενδεχόμενο απότομων αλλαγών, και δυνατότητα on-line ανίχνευσης μοτίβων χωρίς την ανάγκη αποθήκευσης όλων των δεδομένων και προσπέλασης .
- Δεν απαιτούν τον καθορισμό της δομής του μοντέλου εκ των προτέρων, αλλά επιτρέπουν στον αριθμό και τη φύση των παραμέτρων του μοντέλου να αποφασιστούν καθώς αυτό αναπτύσσεται με το χρόνο.

Οι χρονοσειρές είναι εξ ορισμού ακολουθίες παρατηρήσεων και, ως εκ τούτου, κάλλιστα μπορούν να θεωρηθούν ως ακολουθίες δεδομένων. Από μια άλλη οπτική, φαίνεται ότι οι τρέχουσες τιμές των ακολουθιών δεδομένων προέρχονται από διαδικασίες που έχουν μνήμη. Αυτή η συσχέτιση με το παρελθόν θεωρούμε ότι επιτρέπει τον χαρακτηρισμό των ακολουθιών και μεταξύ άλλων ως δυναμικές χρονοσειρές. Συνεπώς οι χρονοσειρές, η μελέτη των οποίων παρουσιάζεται στην επόμενη παράγραφο, μπορούν κάλλιστα να μελετηθούν χρησιμοποιώντας τα εργαλεία ανάλυσης των χρονικά εξελισσόμενων συστημάτων.

2.2 Μοντελοποίηση χρονοσειρών

Η παράγραφος αυτή πραγματεύεται τους κύριους στόχους της ανάλυσης χρονοσειρών περιγράφει αναλυτικά τα βήματα της προεπεξεργασίας αυτών και στη συνέχεια εστιάζει στην πρόβλεψη χαρακτηριστικών από χρονοσειρές, εξετάζοντας τα στάδια της διαδικασίας πρόβλεψης, τις κατηγορίες μεθόδων που χρησιμοποιούνται καθώς και τους παράγοντες που θα πρέπει να λαμβάνονται υπόψη κατά την ανάπτυξη ενός προβλεπτικού μοντέλου.

2.2.1 Ανάλυση χρονοσειρών και προ-επεξεργασία αυτών

Στην βιβλιογραφία, μία χρονοσειρά μπορεί να οριστεί μεταξύ άλλων και ως ένα σύνολο παρατηρήσεων-τιμών $\{x\} = \{x_0, x_1, x_2, \dots x_t, \dots, x_T\}$ με $x_t \in \mathbb{R}$ τις οποίας κάθε τιμή x_t καταγράφεται τη χρονική στιγμή t, ενώ το μήκος της χρονοσειράς δηλώνεται με T. Στην περίπτωση που οι χρονικές στιγμές κατά τις οποίες λαμβάνονται οι παρατηρήσεις αυτές ανήκουν σε ένα πεπερασμένο ή αριθμήσιμο σύνολο τιμών τότε η χρονοσειρά καλείται διακριτού χρόνου, ενώ στην περίπτωση που ανήκουν σε ένα μη αριθμήσιμο σύνολο, τότε χαρακτηρίζεται ως συνεχούς χρόνου.

Το πεδίο της ανάλυσης χρονοσειρών, δηλαδή η διερεύνηση της εφαρμογής μιας επιστημονικής μεθόδου για την μοντελοποίηση της διατεταγμένης ακολουθίας παρατηρήσεων συνήθως κάνει χρήση των θεμελιωδών αρχών της στατιστικής επιστήμης. Για παράδειγμα, αναφέρουμε την χρήση της παλινδρόμησης (regression) η οποία χρησιμοποιεί ως μεταβλητές εισόδου τις παρελθοντικές τιμές της χρονοσειράς και ως έξοδο, την τρέχουσα ή την μελλοντική τιμή της χρονοσειράς. Αξίζει να σημειωθεί πως όταν οι μεταβλητές εισόδου περιλαμβάνουν και τις τιμές από διαφορετικές χρονοσειρές, τότε αυτές καλούνται εξωγενείς. Για παράδειγμα, ένα μοντέλο πρόβλεψης της ποσότητας ζήτησης ενός προϊόντος μπορεί να δέχεται εκτός προηγούμενες τιμές της ποσότητας ζήτησης και τιμές εξωγενών μεταβλητών διαφορετικής φύσης που όμως επηρεάζουν την ποσότητα ζήτησής, όπως το κατά κεφαλήν εισόδημα ή το μέγεθος του πληθυσμού. Η εξέλιξη μιας χρονοσειράς επηρεάζεται συνήθως από την παρελθοντική συμπεριφορά της, και επομένως κατά την μοντελοποίηση των χρονοσειρών θα πρέπει να ληφθεί υπόψη η σειρά με την οποία συλλέγονται οι παρατηρήσεις.

Οι κύριοι στόχοι της ανάλυσης χρονοσειρών είναι:

- Η περιγραφή των δεδομένων χρησιμοποιώντας στατιστικές ή/και γραφικές μεθόδους [42]. Για το σκοπό αυτό, η γραφική αναπαράσταση των δεδομένων σε συνάρτηση με το χρόνο έχει προταθεί καθώς φαίνεται να είναι ιδιαίτερα βοηθητική [43].
- Η πρόβλεψη μελλοντικών τιμών της χρονοσειράς, η οποία εξαρτάται από την εύρεση ενός κατάλληλου στατιστικού μοντέλου κατάλληλου για την εκτίμηση μελλοντικών τιμών των δεδομένων μιας χρονοσειράς. Ανάλογα με το πλήθος των μεταβλητών που επηρεάζουν την εξέλιξη της χρονοσειράς ακολουθούνται διαφορετικές προσεγγίσεις μοντελοποίησης. Στα λεγόμενα μονομεταβλητά (univariate) μοντέλα υπάρχει μία μονοδιάστατη μεταβλητή εισόδου, η οποία περιέχει παρελθοντικούς όρους της χρονοσειράς που μελετάται, δηλαδή οι προβλέψεις βασίζονται αποκλειστικά σε τιμές ιστορικών δεδομένων της χρονοσειράς. Ένα μοντέλο που δέχεται σαν είσοδο προηγούμενες τιμές θερμοκρασίας ανά μία ώρα και παρέχει ωριαίες προβλέψεις θερμοκρασίας αποτελεί ένα απλό παράδειγμα ενός μονομεταβλητού μοντέλου. Από την άλλη πλευρά, στα πολυμεταβλητά (multivariate) μοντέλα προστίθενται επιπλέον μεταβλητές εισόδου που περιέχουν τιμές ιστορικών δεδομένων διαφορετικών χρονοσειρών οι οποίες σχετίζονται βέβαια με την μεταβλητή εξόδου. Αν στο προαναφερθέν μοντέλο πρόβλεψης της θερμοκρασίας προστεθούν στις εισόδους και οι προηγούμενες τιμές νεφοκάλυψης, υγρασίας, ταχύτητας ανέμου και σημείου δρόσου ανά μία ώρα, τότε προκύπτει ένα πολυμεταβλητό μοντέλο. Η πρόβλεψη χρονοσειρών απαντάται ως πρωταρχικής σημασίας αντικείμενο σε πολλούς τομείς της επιστήμης και της ζωής όπως στην οικονομία [44], στις ψηφιακές συναλλαγές [45], στον βιομηχανικό έλεγχο ποιότητας διεργασιών [46], στη διανομή ηλεκτρικής ενέργειας [47], στη μετεωρολογία [48], στη γενετική [49], στην επιδημιολογία [50] και πολλούς ακόμη.
- Οι ενέργειες ελέγχου και ρύθμισης ποικίλων διεργασιών, όπως χημικών ή βιολογικών. Σε αρκετές εφαρμογές συστημάτων ελέγχου, η αξιοποίηση αξιόπιστων προβλέψεων που αφορούν τις μεταβλητές του συστήματος, αποδεικνύεται ιδιαίτερα χρήσιμη για την ενίσχυση της ευρωστίας του [51], [52], [53].

- Η διεξαγωγή προσομοιώσεων. Δεδομένου ενός συνόλου εξισώσεων οι οποίες περιγράφουν την λειτουργία ενός συστήματος, δίνεται η δυνατότητα, μέσω προσομοιώσεων, της εκτίμησης των παραμέτρων που αφορούν την ομαλή λειτουργία του συνολικού συστήματος [54]. Για παράδειγμα, η κατασκευή ενός μοντέλου που προσομοιώνει τη λειτουργία μιας ανεμογεννήτριας αποτελεί ένα απαραίτητο στοιχείο για τη διενέργεια μελέτης ευστάθειας και τον καθορισμό των παραμέτρων ασφαλούς λειτουργίας της [55].
- Η ανίχνευση ακραίων τιμών (outliers) [56] για την κατανόηση της φυσιολογικής συμπεριφοράς της χρονοσειράς και τον εντοπισμό καταστάσεων που αποκλίνουν από αυτή.
- Η ανακάλυψη επαναλαμβανόμενων μοτίβων συμπεριφοράς στα δεδομένα, μέσω της διερεύνησης της συσχέτισης μεταξύ των εξαρτημένων μεταβλητών του μοντέλου της χρονοσειράς, η οποία αποτελεί ένα εξαιρετικά χρήσιμο εργαλείο για τα δυναμικά - online συστήματα συστάσεων (Recommender Systems) [57].
- Η κατάταξη των παρατηρήσεων των χρονοσειρών, η οποία αποτελεί μια από τις πιο σημαντικές κατηγορίες προβλημάτων μηχανικής μάθησης και σχετίζεται με εργασίες αναγνώρισης που εντάσσονται σε ποικίλους τομείς όπως η ανθρώπινη φυσική δραστηριότητα [58], η ιατρική διάγνωση [59], η ευφυής μετακίνηση [60], η βιομετρική επαλήθευση υπογραφών [61], η παρακολούθηση της κατανάλωσης ηλεκτρικής ενέργειας [62] και πολλοί άλλοι.

Ένα ιδιαίτερα σημαντικό στάδιο της ανάλυσης μιας χρονοσειράς είναι η επιλογή ενός κατάλληλου στατιστικού μοντέλου (ή κατηγορίας μοντέλων) για την αναπαράσταση των δεδομένων της. Προκειμένου να συμπεριλάβει κανείς την πιθανοφανή φύση των παρατηρήσεων, είναι δόκιμο να υποθέσει ότι οι παρατηρήσεις $x_t, t \in \mathbb{N}$ προκύπτουν από μία οικογένεια τυχαίων μεταβλητών $X = \{X_1, X_2, \dots\} = \{X_t\}_{t=1}^{\infty}$., η οποία ονομάζεται στοχαστική διαδικασία με από κοινού κατανομή πιθανότητας $\mathbb{P}(X_1, X_2, \dots)$. Επομένως, το σύνολο των παρατηρήσεων της χρονοσειράς $\{x_t | t \in \mathbb{N}\}$ αποτελεί ουσιαστικά ένα από τα πιθανά ενδεχόμενα της από κοινού κατανομής πιθανότητας και χαρακτηρίζεται ως μια πραγματοποίηση (realization) της στοχαστικής διαδικασίας X. Η αναπαράσταση μιας χρονοσειράς ως η πραγματοποίηση μιας στοχαστικής διαδικασίας αποτελεί μια ιδιαίτερα βοηθητική προσέγγιση η οποία βοηθάει τη μοντελοποίηση της. Η παραδοσιακή ανάλυση χρονοσειρών βασίζεται σε υποθέσεις περί γραμμικότητας και στασιμότητας (stationarity), έννοιες οι οποίες αναλύονται παρακάτω.

Η γραμμική μοντελοποίηση χρονοσειρών θεωρεί ότι κάθε τιμή της πραγματοποίησης της στοχαστικής διαδικασίας που αναπαριστά τις παρατηρήσεις μιας χρονοσειράς, μπορεί να εκφραστεί ως γραμμικός συνδυασμός των προηγούμενων τιμών της πραγματοποίησης. Μια επιπλέον απαίτηση των γραμμικών μοντέλων είναι η στασιμότητα - stationarity της στοχαστικής διαδικασίας. Μια στοχαστική διαδικασία X χαρακτηρίζεται ως στάσιμη εάν οι στατιστικές της ιδιότητες παραμένουν αμετάβλητες ανεξάρτητα από το χρονικό παράθυρο παρατήρησης της διαδικασίας. Στο σημείο αυτό θα πρέπει να επισημανθεί ο διαχωρισμός μεταξύ της αυστηρής και ασθενούς στασιμότητας.

Η διαδικασία X είναι αυστηρώς στάσιμη αν η κατανομή της ικανοποιεί την ακόλουθη ιδιότητα:

$$\mathbb{P}(X_1, X_2, \dots, X_T) = \mathbb{P}(X_{1+\tau}, X_{2+\tau}, \dots, X_{T+\tau}), \forall T, \tau \in \mathbb{N}^* \quad (2.1)$$

Η έννοια της αυστηρής στασιμότητας δηλώνει ότι η κατανομή της X, από την οποία προκύπτει η ακολουθία των παρατηρήσεων της χρονοσειράς, παραμένει ίδια με την πάροδο του χρόνου. Συγκεκριμένα, κάθε i.i.d στοχαστική διαδικασία είναι ισχυρά στάσιμη. Στην πράξη, είναι δύσκολο να ελεγχθεί η ισχύς της ιδιότητας (2.1), και για το λόγο αυτό έχει διατυπωθεί ο ορισμός της ασθενούς στασιμότητας.

Μια διαδικασία X ονομάζεται ασθενώς στάσιμη αν η μέση τιμή και η διακύμανσή της ικανοποιούν τις ακόλουθες ιδιότητες:

$$\begin{gathered}\mathbb{E}[\mathrm{X_t}] = \mu \\ \exists\, \gamma\colon \forall \mathrm{h} \in \mathbb{N}\colon \mathrm{Cov}(\mathrm{X_t}, \mathrm{X_{t+h}}) = \gamma(\mathrm{h}) \Rightarrow \mathrm{Var}(\mathrm{X_t}) = \gamma(0) < \infty\end{gathered} \tag{2.2}$$

όπου με Cov και Var δηλώνονται η συνδιακύμανση και η διακύμανση αντίστοιχα, ενώ η συνάρτηση γ είναι συμμετρική ($\gamma(-\mathrm{h}) = \gamma(\mathrm{h})$) και φραγμένη ($|\gamma(\mathrm{h})| \leq \gamma(0) < \infty$). Η ιδιότητα (2.2) εκφράζει ότι η μέση τιμή και η συνδιακύμανση της X είναι χρονικά ανεξάρτητες και η τιμή της διακύμανσή της πεπερασμένη.

Ένα γραμμικό μοντέλο παλινδρόμησης είναι σε θέση να προβλέπει την τιμή μιας χρονοσειράς τη χρονική στιγμή t, χρησιμοποιώντας έναν γραμμικό συνδυασμό των προηγούμενων τιμών της:

$$\mathrm{x_t} = \mathrm{f}(\mathrm{x_{t-1}}, \mathrm{x_{t-2}}, \dots) + \sigma(\mathrm{x_{t-1}}, \mathrm{x_{t-2}}, \dots)\varepsilon_\mathrm{t} \tag{2.3}$$

όπου ε_t είναι μια i.i.d $(0,\sigma^2)$ διαδικασία μέσω της οποίας μοντελοποιείται ο θόρυβος που υπάρχει στα δεδομένα, f και σ είναι αντίστοιχα η υπό-όρους μέση τιμή και η τυπική απόκλιση της μεταβλητής $\mathrm{X_t}$. Τα πιο δημοφιλή γραμμικά μοντέλα παλινδρόμησης που έχουν αναπτυχθεί για την ανάλυση χρονοσειρών είναι η εκθετική εξομάλυνση (exponential smoothing) [63], η αυτοπαλινδρόμηση (autoregression - AR) [64], ο κινούμενος μέσος όρος (moving average - MA) [65] και η παλινδρόμηση με κινούμενο μέσο όρο (autoregressive moving average - ARMA) [66]. Η εξίσωση (2.4) περιγράφει το γενικό μοντέλο γραμμικής παλινδρόμησης τάξης p:

$$\mathrm{x_t} = \varphi_0 + \sum_{\mathrm{i}=1}^{\mathrm{p}} \varphi_\mathrm{i} \mathrm{x_{t-i}} + \sigma \cdot \varepsilon_\mathrm{t} \tag{2.4}$$

όπου $\{\varphi_0, \varphi_1, \dots, \varphi_\mathrm{p}, \sigma^2\}$ είναι οι προς υπολογισμό παράμετροι του μοντέλου.

Τα γραμμικά μοντέλα χρονοσειρών έχουν μελετηθεί και χρησιμοποιηθεί ευρέως στην βιβλιογραφία [67], [68], [69] λόγω της απλότητας τους, της χαμηλής υπολογιστικής απαίτησης και της επαρκούς θεωρητικής τους θεμελίωσης, η οποία βασίζεται στην θεωρία πιθανοτήτων των κανονικών στοχαστικών διαδικασιών. Ωστόσο, υπάρχει ένα αυξανόμενο ενδιαφέρον για τη μελέτη μη γραμμικών και μη στάσιμων μοντέλων χρονοσειρών σε πολλά πρακτικά προβλήματα. Ο πρώτος και ο απλούστερος λόγος για αυτό είναι ότι πολλά προβλήματα του πραγματικού κόσμου δεν ικανοποιούν τις παραδοχές της γραμμικότητας ή/και της στασιμότητας [70].

Στην περίπτωση της μη γραμμικής μοντελοποίησης οι συναρτήσεις f και g είναι μη γραμμικές και μπορούν να μοντελοποιηθούν με παραμετρικές ή μη τεχνικές. Ανάμεσα στα πιο διαδεδομένα παραμετρικά μοντέλα περιλαμβάνονται η πολυωνυμική παλινδρόμηση (Polynomial Autoregression - PAR) [71], η παλινδρόμηση συναρτησιακών συντελεστών (Functional-coefficient Autoregression - FAR) [72], η παλινδρόμηση μεταγωγής Markov (Markov Switching Autoregression - MSAR) [73], η παλινδρόμηση ομαλής μετάβασης (Smooth Transition Autoregression - STAR) [74], η παλινδρομική υπό όρους έτερο-σκεδαστικότητα (Autoregressive Conditional Heteroscedasticity - ARCH) [75], η παλινδρόμηση κατωφλίου (Threshold Autoregression - TAR) [76] και τα διγραμμικά μοντέλα (Bilinear Models) [35]. Ο στόχος τέτοιων μεθόδων είναι ο υπολογισμός των παραμέτρων του μοντέλου που αναπαριστά μια στοχαστική διαδικασία [77]. Αντιθέτως, οι μη παραμετρικές μέθοδοι, αντί να απαιτούν μια καθορισμένη μορφή για τις συναρτήσεις f

και g, επιβάλουν ορισμένες ποιοτικές μόνο παραδοχές, όπως η προϋπόθεση να είναι λείες συναρτήσεις. Μια συνάρτηση είναι λεία αν οι παράγωγοι κάθε τάξης αυτής υπάρχουν και είναι συνεχείς συναρτήσεις στο πεδίο Στη συνέχεια, η προσέγγιση των συναρτήσεων γίνεται επαναληπτικά με χρήση της τεχνικής της δοκιμής και σφάλματος για το σύνολο των παρελθοντικών παρατηρήσεων, βελτιώνοντας με αυτό τον τρόπο την ακρίβειά της [78] . Σε αυτή την κατηγορία μεθόδων κατατάσσονται η πολυμεταβλητή τοπική πολυωνυμική παλινδρόμηση (Multivariate Local Polynomial Regression) [79], η τοπική παλινδρόμηση συναρτησιακών συντελεστών (Local Functional-Coefficient Autoregression) [80], η προσαρμοστική παλινδρόμηση συναρτησιακών συντελεστών (Adaptive Functional-Coefficient Autoregression) [81], τα προσθετικά μοντέλα (Additive Models) [82], καθώς και το μοντέλο Γκαουσιανής διαδικασίας (Gaussian process - GP) [42]. Κατά τη διάρκεια των τελευταίων δεκαετιών, μια καινούρια κλάση μεθόδων που έχει ευρύτερα εδραιωθεί ως μηχανική μάθηση [83], περιλαμβάνει μια σειρά από γνωστές πλέον προσεγγίσεις όπως: τα τεχνητά νευρωνικά δίκτυα [84], οι μηχανές διανυσμάτων υποστήριξης [85], η βαθιά και συνελικτική μάθηση [86], [87] κ.λ.π.

Σε όλες τις εφαρμογές του πραγματικού κόσμου, τα δεδομένα μιας χρονοσειράς είναι αδύνατο να χρησιμοποιηθούν στην μορφή που βρίσκονται όταν καταγράφονται για την ανάπτυξη ενός αποτελεσματικού μοντέλου. Αυτό μπορεί να οφείλεται εν μέρει σε παράγοντες που προέρχονται από τις μετρητικές διατάξεις όπως η εισαγωγή θορύβου οι εσφαλμένες καταγραφές η δυσλειτουργία αισθητήρων κ.λ.π., και εν μέρει στα φυσικά χαρακτηριστικά των δεδομένων.. Για το λόγο αυτό, κατά την διαδικασία της μοντελοποίησης χρονοσειρών, αρχικά λαμβάνει χώρα η προεπεξεργασία των δεδομένων, η οποία προβλέπει συνήθως τα ακόλουθα στάδια.

1. Εφαρμογή γραφικών τεχνικών για τη διερεύνηση των βασικών χαρακτηριστικών της, μερικά από τα οποία είναι:
    α. Εποχική μεταβλητότητα (seasonal variation). Αυτός ο τύπος μεταβλητότητας είναι περιοδικός με περίοδο που εξαρτάται από τη συχνότητα που λαμβάνονται οι μετρήσεις. Παρατηρείται για πολλές σειρές, όταν εμφανίζονται παρόμοια μοτίβα συμπεριφοράς ανά συγκεκριμένα χρονικά διαστήματα (συνήθως σε ετήσιο κύκλο). Σημειώνεται ότι εάν μια χρονοσειρά μετράται μόνο μία φορά το χρόνο, τότε δεν είναι δυνατό να διαπιστωθεί εάν υπάρχει εποχική διακύμανση.
    β. Εποχική τάση (Seasonal trend). Αυτός ο τύπος μεταβλητότητας υπάρχει όταν μια χρονοσειρά εμφανίζει σταθερή άνοδο ή κάθοδο των τιμών της, για αρκετές τουλάχιστον διαδοχικές χρονικές περιόδους.
    γ. Ακανόνιστες διακυμάνσεις. Ο όρος αυτός χρησιμοποιείται συχνά για να περιγράψει οποιοδήποτε τύπο μεταβλητότητας απομένει μετά την απαλοιφή μετά της τάσης, της εποχικότητας και άλλων συστηματικών φαινομένων. Ως εκ τούτου, μπορεί να είναι εντελώς τυχαίες, οπότε δεν μπορούν να προβλεφθούν. Ωστόσο, μπορεί να παρουσιάζουν βραχυπρόθεσμη συσχέτιση ή να μεμονωμένες ασυνέχειες.
    δ. Ακραίες παρατηρήσεις. Πρόκειται για τιμές οι οποίες διαφέρουν κατά πολύ σε σχέση με το εύρος τιμών όπου κινούνται τα δεδομένα κατά μέσο όρο. Η αναγνώριση και η αντιμετώπισή τους είναι ιδιαίτερης σημασίας καθώς στην αντίθετη περίπτωση ενδέχεται να καταστήσουν δυσμενή την αξιοπιστία και την ακρίβεια του τελικού μοντέλου.
2. Αφαίρεση της τάσης και της εποχικότητας ώστε να προκύψει μια στάσιμη χρονοσειρά. Για να επιτευχθεί αυτός ο στόχος, μερικές φορές μπορεί να είναι απαραίτητο να εφαρμοστεί ένας προκαταρκτικός μετασχηματισμός στα δεδομένα. Για παράδειγμα, εάν παρατηρείται μια ανοδική τάση στο γράφημα της χρονικής εξέλιξης της χρονοσειράς, και η διακύμανση φαίνεται να αυξάνεται, τότε ένας

μετασχηματισμός, όπως ο λογαριθμικός, τετραγωνικής ρίζας, Box-Cox και Yeo-Johnson, μπορεί να θεωρηθεί επιθυμητός προκειμένου να σταθεροποιηθεί η διακύμανση (ή να γίνει ομοσκεδαστική). Ακόμη, η ασυμμετρία που μπορεί να εμφανίζουν οι παρατηρήσεις ως προς το εύρος των θετικών και αρνητικών τιμών τους, μπορεί να εξομαλυνθεί με την εφαρμογή ενός κατάλληλου μετασχηματισμού. Τέλος, εάν το εποχικό φαινόμενο φαίνεται να είναι πολλαπλασιαστικό, δηλαδή το εύρος τιμών και η συχνότητα του φαινομένου μεταβάλλονται με τυχαίο τρόπο, τότε είναι θεμιτό να μετασχηματιστούν τα δεδομένα ώστε η μεταβολή των μεγεθών αυτών να συμβαίνει γραμμικά (προσθετική εποχικότητα), καθώς είναι γενικά πιο εύκολο να χειριστεί κανείς τα γραμμικά φαινόμενα. Στην βιβλιογραφία, περιγράφονται μέθοδοι με τους οποίους μπορεί να εκτιμηθεί και στην συνέχεια να αφαιρεθεί τόσο η τάση όσο και η εποχικότητα, όπως η τεχνική της αποδόμησης [88] και άλλες οι οποίες βασίζονται στη διαφόριση των δεδομένων [49]. Όποια μέθοδος και αν χρησιμοποιηθεί, ο τελικός διαχωρισμός των μη στάσιμων φαινομένων (εποχικότητα και τάση) από το υπολειπόμενο στοχαστικό μέρος της χρονοσειράς.

3. Εκκαθάριση (purge) των δεδομένων. Ένα σημαντικό μέρος της προεπεξεργασίας των δεδομένων είναι η αξιολόγηση της ποιότητας τους και το ενδεχόμενο τροποποίησης τους. Αυτή η διαδικασία ονομάζεται συχνά εκκαθάριση δεδομένων και αποτελεί ουσιώδης στάδιο στην προσπάθεια μοντελοποίησης. Το στάδιο αυτό μπορεί να περιλαμβάνει την αφαίρεση ακραίων τιμών, τον εντοπισμό και τη διόρθωση προφανών σφαλμάτων ή τη συμπλήρωση ελλειπουσών παρατηρήσεων. Ιδιαίτερα σημαντική είναι η κατάλληλη αντιμετώπιση των ακραίων τιμών της χρονοσειράς [89], [90]. Οι ακραίες τιμές ενδέχεται να επηρεάζουν είτε μόνο την τρέχουσα παρατήρηση είτε έναν αριθμό από παρατηρήσεις που ακολουθούν. Η περίπτωση αυτή υποδεικνύει κάποια μεταβολή στη δυναμική συμπεριφορά της χρονοσειράς η οποία είναι χρήσιμο να μοντελοποιηθεί. Αντίθετα, στην πρώτη περίπτωση πρόκειται κατά πάσα πιθανότητα για εσφαλμένη μέτρηση που μπορεί να επιδράσει αρνητικά στην αποδοτικότητα του τελικού μοντέλου. Παρόλο που η διάκριση μεταξύ των ακραίων τιμών που οφείλονται σε σφάλματα, στη μη γραμμικότητα ή στην ύπαρξη μη κανονικής κατανομής σφάλματος, αποτελεί μια πρόκληση, αναπτύσσονται και χρησιμοποιούνται όλο και περισσότερο εξελιγμένες μέθοδοι που μειώνουν τις αρνητικές επιπτώσεις των ακραίων παρατηρήσεων, συμπεριλαμβανομένης και της ολικής τους απομάκρυνσης [91], [92], [93].

2.2.2 Πρόβλεψη χαρακτηριστικών από χρονοσειρές

Ξεκινώντας, με τον όρο χαρακτηριστικό μιας χρονοσειράς νοείται μια μεταβλητή που καθορίζεται κατά την μοντελοποίηση της χρονοσειράς και μπορεί να αφορά είτε την ποσότητα που αντιπροσωπεύουν οι παρατηρήσεις της χρονοσειράς είτε κάποια παράγωγη ποσότητα που προκύπτει από αυτές (π.χ. μέσος όρος ανά χρονικό διάστημα Δt). Η πρόβλεψη χαρακτηριστικών από χρονοσειρές μπορεί να οριστεί ως η διαδικασία ανάπτυξης μιας προγνωστικής διαδικασίας[1] η οποία χρησιμοποιεί ιστορικά δεδομένα μιας ή περισσότερων χρονοσειρών και παράγει εκτιμήσεις για τις μελλοντικές τιμές ενός

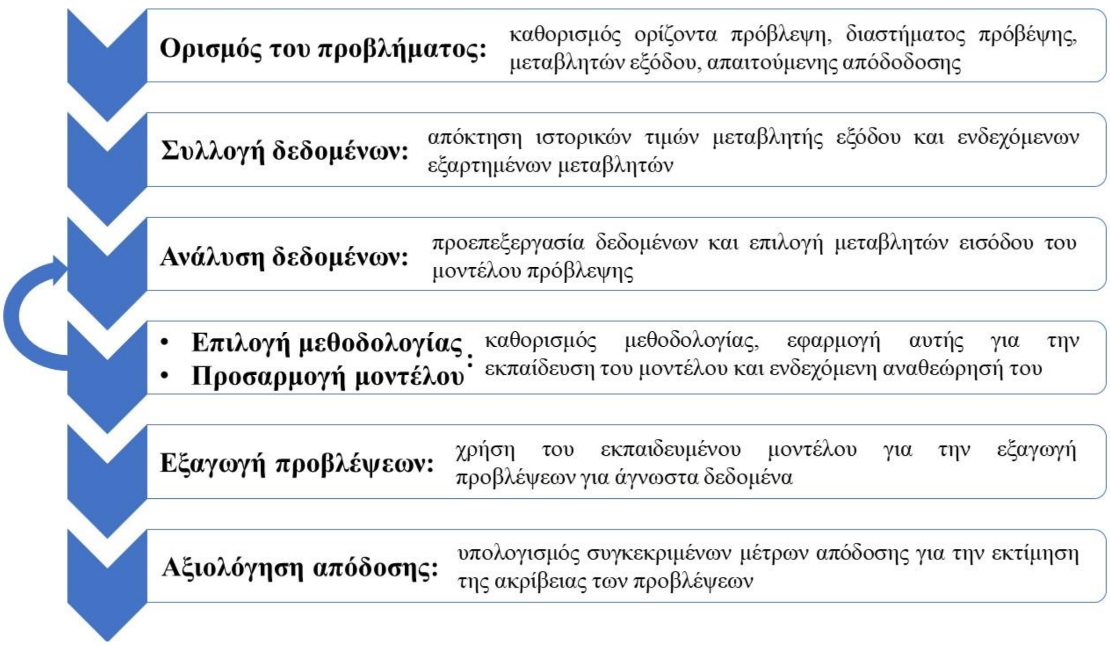
[1] Ή, κάνοντας χρήση του εκλαϊκευμένου όρου, μοντέλου.

χαρακτηριστικού μιας χρονοσειράς. Πλέον η μελέτη της πρόβλεψης χρονοσειρών έχει αποκτήσει εκτεταμένο ερευνητικό ενδιαφέρον με εφαρμογή σε ποικίλα πεδία που εκτείνονται από την ιατρική [94], [95] και τον οικονομικό σχεδιασμό [96] έως την κοινωνική δικτύωση [97] και την διαχείριση της ενέργειας [98].

Ας θεωρήσουμε την περίπτωση κατά την οποία διαθέτουμε μια συλλογή από παρατηρήσεις μιας χρονοσειράς $y_1, y_2, \dots, y_t$ και ζητείται η πρόβλεψη της τιμής της χρονοσειράς για την χρονική στιγμή $t+h$. Ο ακέραιος αριθμός h δηλώνει το πλήθος των μελλοντικών βημάτων από την τρέχουσα χρονική στιγμή και καλείται ορίζοντας ή παράθυρο πρόβλεψης, ενώ η εκτιμώμενη πρόβλεψη δηλώνεται ως $\hat{y}_{t+h}$. Ο ορίζοντας πρόβλεψης αποτελεί ένα βασικό χαρακτηριστικό ενός μοντέλου πρόβλεψης και εξαρτάται άμεσα από την εφαρμογή για την οποία προορίζεται το μοντέλο [99], [100]. Έτσι, ανάλογα με το χρονικό διάστημα στο οποίο αναφέρεται ο ορίζοντας πρόβλεψης, οι προβλέψεις χαρακτηριστικών από χρονοσειρές διακρίνονται σε[2]:

- Βραχυπρόθεσμες, οι οποίες δεν ξεπερνούν τον ορίζοντα των τριών μηνών,
- Μεσοπρόθεσμες, οι οποίες αναφέρονται σε ορίζοντες μεταξύ τριών μηνών και ενός έτους και
- Μακροπρόθεσμες, οι οποίες αφορούν ορίζοντες μεγαλύτερους του ενός έτους.

Στα πλαίσια της παρούσας διατριβής αναπτύσσονται μοντέλα βραχυπρόθεσμης πρόβλεψης της κατανάλωσης ηλεκτρικού φορτίου με ορίζοντα πρόβλεψης που εκτείνεται από μερικά λεπτά έως ένα 24ωρο. Εκτός από τον ορίζοντα, ένα μοντέλο πρόβλεψης χαρακτηρίζεται επίσης από τη συχνότητα με την οποία παρέχονται οι προβλέψεις, το οποίο καλείται διάστημα πρόβλεψης. Καθοριστικό ρόλο κατά το σχεδιασμό ενός μοντέλου παίζει επίσης η επιλογή των μεταβλητών εισόδου που δέχεται αλλά και η μεθοδολογία που ακολουθείται για την εκπαίδευσή του και την παραγωγή των προβλέψεων. Η ανάπτυξη ενός μοντέλου πρόβλεψης αποτελεί ένα από τα ενδιάμεσα στάδια της διαδικασίας πρόβλεψης (βλ. Σχήμα 2) τα οποία περιγράφονται παρακάτω:



Σχήμα 2: Τα βήματα της διαδικασίας πρόβλεψης χαρακτηριστικών από χρονοσειρές.

---

[2] Η παραπάνω κατηγοριοποίηση δεν είναι απόλυτη και μπορεί να διαφέρει ανάλογα με το πεδίο εφαρμογής.

1. Ορισμός του προβλήματος. Το στάδιο αυτό συνίσταται στην κατανόηση του τρόπου που θα χρησιμοποιηθούν οι προβλέψεις. Στο σημείο αυτό πραγματοποιείται ο καθορισμός του κατάλληλου ορίζοντα πρόβλεψης, της/των μεταβλητής/ων εξόδου του μοντέλου, του διαστήματος πρόβλεψης και της απαιτούμενης ακρίβειας των τελικών προβλέψεων.
2. Συλλογή δεδομένων. Σε αυτό το βήμα λαμβάνονται τα ιστορικά δεδομένα, δηλαδή οι παρελθοντικές τιμές της προς πρόβλεψη μεταβλητής και ενδεχομένως κάποιων επιπλέον εξαρτημένων μεταβλητών. Αναλόγως την εφαρμογή, τα δεδομένα αυτά συνήθως προέρχονται από μετρητές, αισθητήρες ή βάσεις δεδομένων [101]. Θα πρέπει να τονιστεί ότι αρκετά συχνά τα συστήματα συλλογής και αποθήκευσης δεδομένων ενδέχεται να αλλάζουν, συνεπώς ορισμένα από τα δεδομένα μπορεί να μην είναι πλέον χρήσιμα για ένα συγκεκριμένο πρόβλημα.
3. Ανάλυση χρονοσειρών. Πρόκειται για μια ιδιαίτερα σημαντική φάση της διαδικασίας πρόβλεψης χρονοσειρών, η οποία έχει περιγραφεί στο υποκεφάλαιο 2.2.1. Στην συνέχεια, σειρά έχει ο προσδιορισμός των μεταβλητών εισόδου του μοντέλου. Η επιλογή τους παίζει καθοριστικό ρόλο στην προβλεπτική ικανότητα του τελικού μοντέλου. Οι μεταβλητές εισόδου μπορεί να περιλαμβάνουν αναδρομικές παρατηρήσεις μίας ή περισσότερων χρονοσειρών καθώς και παράγωγες ποσότητες οι οποίες προκύπτουν από αυτές, για παράδειγμα αναφέρουμε γραμμικούς συνδυασμούς τους. Παρόλο που στη βιβλιογραφία συναντώνται μέθοδοι που υποδεικνύουν το βέλτιστο σύνολο μεταβλητών εισόδου, όπως η ανάλυση συσχέτισης, ή τεχνικές φιλτραρίσματος [102], ο τελικός καθορισμός τους επαφίεται στον σχεδιαστή.
4. Επιλογή μεθοδολογίας. Την ανάλυση και προεπεξεργασία των δεδομένων, ακολουθεί η επιλογή μεθοδολογίας πρόβλεψης. Σύμφωνα με το «No free lunch theorem» [103], δεν υπάρχει καμία μέθοδος πρόβλεψης που να υπερτερεί έναντι των υπολοίπων για κάθε είδους χρονοσειράς. Ωστόσο, μια κατάλληλη μεθοδολογία μπορεί να επιλεγεί λαμβάνοντας υπόψιν ορισμένες παραμέτρους όπως η απαιτούμενη ακρίβεια πρόβλεψης, οι διαθέσιμοι υπολογιστικοί πόροι, η απαιτούμενη συχνότητα παροχής προβλέψεων, τα στατιστικά χαρακτηριστικά της χρονοσειράς και από το είδος της εφαρμογής του προβλεπτικού μοντέλου (online/offline) [104], [105]. Η βιβλιογραφία βρίθει από μεθόδους πρόβλεψης που έχουν προταθεί και εφαρμοστεί και οι οποίες μπορούν να διακριθούν στις ακόλουθες κατηγορίες:
   α. Γραμμικές μέθοδοι
   Το βασικό χαρακτηριστικό των μεθόδων αυτών είναι η υπόθεση ότι υπάρχει γραμμική εξάρτηση μεταξύ των μεταβλητών εισόδου και της μεταβλητής εξόδου ακολουθώντας τη γενική μορφή:
   $$\hat{y}_t = \sum_{i=1}^{m} a_i x_i(t) + b \tag{2.5}$$
   όπου m είναι το πλήθος των μεταβλητών εισόδου, $x_i(t)$ είναι οι τιμές των μεταβλητών εισόδου για μία τυχαία χρονική στιγμή t, a_i και b είναι οι συντελεστές και ένας όρος μεροληψίας που θα πρέπει να εκτιμηθούν. Η πιο κοινή μέθοδος επίλυσης για την εκτίμηση των παραμέτρων αυτών συνήθως χρησιμοποιεί τεχνικές ελαχίστων τετραγώνων [106]. Οι πιο διαδεδομένες γραμμικές μέθοδοι είναι η απλή [107] και η πολλαπλή [108] γραμμική παλινδρόμηση, η εκθετική εξομάλυνση [109] και η μέθοδος Holt Winter

[110]. Όντας από τις πρώτες χρονολογικά προσεγγίσεις που προτάθηκαν για την πρόβλεψη χαρακτηριστικών από χρονοσειρές, οι μέθοδοι αυτοί έχουν μελετηθεί διεξοδικά ενώ χαρακτηρίζονται από χαμηλή πολυπλοκότητα και ευρωστία. Ωστόσο, η υπόθεση περί γραμμικότητας δυσχεραίνει την αποτελεσματικότητά τους, μιας και οι χρονοσειρές που προέρχονται από πραγματικά συστήματα χαρακτηρίζονται στην πλειοψηφία τους από μη γραμμική συμπεριφορά.

β. Στοχαστικές προσεγγίσεις
Δεδομένης της ακολουθιακής δομής των χρονοσειρών, δηλαδή της χρονολογικής διαδοχής των παρατηρήσεων, οι στοχαστικές προσεγγίσεις χρησιμοποιούνται ευρέως για την εξαγωγή συμπερασμάτων σχετικά με τα χαρακτηριστικά τους και κατ' επέκταση για την πρόβλεψη αυτών. Τα στοχαστικά μοντέλα περιλαμβάνουν μία ή και περισσότερες τυχαίες μεταβλητές εισόδου με κάποιο βαθμό αβεβαιότητας και παράγουν εκτιμήσεις για τις κατανομές πιθανότητας των δυνατών αποτελεσμάτων της μεταβλητής εξόδου. Οι κατανομές αυτές είναι μαθηματικές συναρτήσεις που εκφράζουν την πιθανοφάνεια των διαφορετικών αποτελεσμάτων, οι οποίες στη συνέχεια συνδυάζονται με κάποιο μηχανισμό, όπως ο μέσος όρος ή η διάμεσος των κατανομών [111], η παλινδρόμηση ποσοστιαίων σημείων (quantile regression) [18], η τεχνική linear opinion pool [112], παρέχοντας τις τελικές προβλέψεις. Στην κατηγορία των στοχαστικών προσεγγίσεων κατατάσσονται μοντέλα ARMA [113], Autoregressive Integrated Moving Average (ARIMA) [114], SARIMA [115], καθώς και μοντέλα που βασίζονται σε στοχαστικές διαφορικές εξισώσεις [116], [117].

γ. Τεχνικές soft computing
Οι τεχνικές soft computing, μπορούν να θεωρηθούν ως μια κλάση προσεγγιστικών υπολογιστικών μοντέλων, για την επίλυση πολύπλοκων μη γραμμικών προβλημάτων, τα οποία είναι αδύνατον να αντιμετωπιστούν με τα παραδοσιακά αναλυτικά μοντέλα (hard computing), είτε λόγω έλλειψης επαρκούς πληροφορίας είτε λόγω αυξημένης πολυπλοκότητας. Ο όρος soft computing προτάθηκε το 1992 από τον Lotfi Zadeh, σύμφωνα με τον οποίο αναφέρεται σε όλες τις μεθοδολογίες που μιμούνται την ικανότητα του ανθρώπου να μαθαίνει σε ένα περιβάλλον αβέβαιο και ασαφές [118]. Ουσιαστικά, κατά την εκπαίδευση ενός τέτοιου μοντέλου εισάγεται κάποια τυχαιότητα στους συντελεστές της προς προσέγγιση συνάρτησης, και επομένως η διαδικασία δεν είναι ντετερμινιστική. Τα νευρωνικά δίκτυα, οι γενετικοί αλγόριθμοι, η ασαφής λογική, η μηχανική μάθηση, τα έμπειρα συστήματα (expert systems) είναι μερικές από τις μεθόδους που ανήκουν στην κατηγορία των τεχνικών soft computing [119]. Μέθοδοι σαν αυτές χαρακτηρίζονται από ορισμένα πλεονεκτήματα που τις καθιστούν κατάλληλες για χρήση σε υπολογιστικά συστήματα υψηλής τεχνολογίας. Αρχικά, μπορούν να υποστηρίξουν την εκτέλεση παράλληλων υπολογισμών και να χειριστούν αποτελεσματικά δεδομένα με θόρυβο. Πέραν τούτου, δεν απαιτούν τη χρήση εξειδικευμένων μαθηματικών μορφοποιήσεων για το εκάστοτε πρόβλημα, γεγονός που βελτιώνει σημαντικά την επεκτασιμότητα και την προσαρμοστικότητά τους σε ποικίλες εφαρμογές. Επιπλέον, διακρίνονται για τη δυνατότητά τους να συνδυάζονται με άλλες μεθόδους αναπτύσσοντας υβριδικά μοντέλα με προηγμένες ιδιότητες [120].

5. Προσαρμογή μοντέλου. Η εκπαίδευση του μοντέλου με την μέθοδο που έχει επιλεχθεί, δηλαδή η εκτίμηση των παραμέτρων του, λαμβάνει χώρα στο στάδιο αυτό. Για να συμβεί αυτό πρέπει αρχικά το σύνολο των διαθέσιμων δεδομένων να χωριστεί σε δύο μέρη, το σύνολο δεδομένων εκπαίδευσης και το σύνολο δεδομένων

ελέγχου, ενώ ένα μέρος του συνόλου εκπαίδευσης που ονομάζεται σύνολο αξιολόγησης, διαχωρίζεται και χρησιμοποιείται για αξιολόγηση. Οι παρατηρήσεις που ανήκουν στο σύνολο εκπαίδευσης χρησιμοποιούνται για την εκπαίδευση του μοντέλου κατά την οποία, για κάθε δεδομένο, οι παράμετροι του μοντέλου, ξεκινώντας από κάποιες τυχαίες αρχικές τιμές, διορθώνονται έτσι ώστε να ελαχιστοποιείται το σφάλμα πρόβλεψης για τα δεδομένα του συνόλου αξιολόγησης. Ενσωματώνοντας το στάδιο της αξιολόγησης, μειώνεται ο κίνδυνος της υπέρ-προσαρμογής (overfitting) και αυξάνεται η αξιοπιστία του μοντέλου [121]. Στην περίπτωση που οι τιμές των μεταβλητών κυμαίνονται σε διαφορετικά δυναμικά όρια, καθίσταται αναγκαία η κανονικοποίηση των δεδομένων πριν από τη έναρξη της διαδικασίας της εκπαίδευσης.

6. Εξαγωγή προβλέψεων. Η εγκυρότητα του εκπαιδευμένου μοντέλου επικυρώνεται μέσα από την λειτουργία του και την παροχή των προβλέψεων για τα δεδομένα τα οποία ανήκουν στο σύνολο ελέγχου. Το γεγονός ότι οι παρατηρήσεις αυτές είναι άγνωστες στο μοντέλο καθώς δεν έχουν συμμετάσχει στο στάδιο της εκπαίδευσης χρησιμοποιούνται, ενισχύουν την προβλεπτική του ικανότητα. Στο σημείο αυτό ενδέχεται οι προβλέψεις να πρέπει να υποβληθούν σε από-κανονικοποίηση, ώστε να εκφραστούν στο εύρος των πραγματικών τιμών διευκολύνοντας την αξιολόγηση της ακρίβειας.
7. Αξιολόγηση απόδοσης. Είναι γεγονός πως ένα μοντέλο πρόβλεψης χαρακτηριστικών από χρονοσειρές μπορεί να αποτελέσει ένα πολύτιμο εργαλείο στα πλαίσια διάφορων εργασιών, όπως η λήψη αποφάσεων [122], η ασφάλεια και η συντήρηση συστημάτων [123], η έγκαιρη διάγνωση ασθενειών [124] και πολλές ακόμη. Για το λόγο αυτό, απαιτείται η χρήση ορισμένων μέτρων απόδοσης για την εκτίμηση της ακρίβειας των προβλέψεων και τη σύγκριση διαφορετικών μοντέλων [125]. Τα μέτρα αυτά είναι συναρτήσεις των πραγματικών τιμών της χρονοσειράς y_t και των παραγόμενων προβλέψεων $\hat{y}_t$ και τα πιο συχνά χρησιμοποιούμενα παρατίθονται στη συνέχεια. Όπως φαίνεται και παραπάνω, στους ακόλουθους ορισμούς, με y_t δηλώνεται η παραγματική τιμή της χρονοσειράς την χρονική στιγμή t, με $\hat{y}_t$ δηλώνεται η αντίστοιχη πρόβλεψη, με $e_t = y_t - \hat{y}_t$ το σφάλμα πρόβλεψης και με n το μέγεθος του συνόλου ελέγχου:
   α. Μέσο σφάλμα πρόβλεψης (mean forecast error)
   $$MFE = \frac{1}{n}\sum_{t=1}^{n} e_t \tag{2.6}$$
   β. Μέσο απόλυτο σφάλμα (mean absolute error)
   $$MAE = \frac{1}{n}\sum_{t=1}^{n} |e_t| \tag{2.7}$$
   γ. Μέσο απόλυτο ποσοστιαίο σφάλμα (mean absolute percentage error)
   $$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left|\frac{e_t}{y_t}\right| \times 100 \tag{2.8}$$
   δ. Μέσο τετραγωνικό σφάλμα (mean squared error)

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2 \tag{2.9}$$

ε. Ρίζα μέσου τετραγωνικού σφάλματος (root mean squared error)

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \tag{2.10}$$

στ. Συντελεστής συσχέτισης (coefficient of determination)

$$R^2 = 1 - \frac{SSE}{SST} \tag{2.11}$$

όπου $SSE = \sum_{t=1}^{n} e_t^2$ και $SST = \sum_{t=1}^{n}\left(y_t - \frac{1}{n}\sum_{t=1}^{n} y_t\right)^2$.

Με βάση τα παραπάνω, κατά την ανάπτυξη ενός προβλεπτικού μοντέλου είναι εφικτό σε κάποιο βαθμό να αξιολογηθεί και να βελτιωθεί η προβλεπτική του απόδοση. Εντούτοις, κατά την εφαρμογή του σε πραγματικό χρόνο προκύπτουν σοβαρές προκλήσεις οι οποίες θέτουν σε κίνδυνο την αξιοπιστία των προβλέψεων. Ένας από τους βασικότερους παράγοντες που μπορούν να υποβαθμίσουν την αποδοτικότητα των μοντέλων πρόβλεψης είναι η μη γραμμική δυναμική συμπεριφορά που χαρακτηρίζει τα περισσότερες χρονοσειρές. Από στατιστικής σκοπιάς, η μη γραμμικότητα των χρονοσειρών μπορεί γίνει αντιληπτή μέσα από ιδιότητες όπως η μη κανονική κατανομή των δεδομένων [126], η απεριοδικότητα, η ασυμμετρία, η μη γραμμική συσχέτιση μεταξύ της τρέχουσας και των παρελθοντικών παρατηρήσεων, η διακύμανση της προβλεπτικής απόδοσης με το χρόνο και η χρονική μη αντιστρεψιμότητα [127]. Εκτός από την πολυπλοκότητα της δυναμικής συμπεριφοράς, η δομή των μη γραμμικών σχέσεων μεταξύ των παρελθοντικών και των μελλοντικών παρατηρήσεων παραμένει άγνωστη αν όχι αδύνατο να προσδιοριστεί στις περισσότερες περιπτώσεις. Κατά συνέπεια, τα κλασικά μοντέλα χρονοσειρών τείνουν να παρουσιάζουν σοβαρά μειονεκτήματα στην πρόβλεψη χαρακτηριστικών από χρονοσειρές πολύπλοκων συστημάτων. Ως εκ τούτου, τις τελευταίες δεκαετίες έχουν σημειωθεί αρκετές εξελίξεις προς την διαχείριση των και μη γραμμικών φαινομένων στα δεδομένα χρονοσειρών μέσω της ανάπτυξης κατάλληλων μεθόδων [128], [129].

Μια παράμετρος που επηρεάζει την ακρίβεια των προβλέψεων και πρέπει απαραιτήτως να λαμβάνεται υπόψη, είναι η αβεβαιότητα η οποία προέρχεται από τρεις κυρίως παράγοντες:

- Αβεβαιότητα σχετικά με τη δομή του μοντέλου πρόβλεψης
- Αβεβαιότητα ως προς τις εκτιμήσεις των παραμέτρων του μοντέλου
- Αβεβαιότητα που πηγάζει από τα δεδομένα της χρονοσειράς, η οποία αποτυπώνεται ως τυχαία μεταβλητότητα στις παρατηρήσεις καθώς και εσφαλμένες μετρήσεις και καταγραφές.

Για την αντιμετώπιση της αβεβαιότητας, έχει προταθεί μια ποικιλία τεχνικών κατά την ανάπτυξη των μεθόδων πρόβλεψης χρονοσειρών, που στηρίζονται στις πιθανότητες, για την κατασκευή διαστημάτων πρόβλεψης, όπως η Μπεϋζιανή προσέγγιση, η παλινδρόμηση Γκαουσιανής διαδικασίας, η παλινδρόμηση ποσοστιαίων σημείων (quantile regression) και η μέθοδο bootstrap [130], [131]. Εξίσου διαδεδομένη είναι η ανάλυση με χρήση πολλαπλών σεναρίων μέσω προσομοίωσης Monte Carlo, δυναμικών μοντέλων ανάλυσης παραγόντων [132] ή άλλων τεχνικών. Μια ακόμη τακτική που συχνά ακολουθείται για την αντιμετώπιση της αβεβαιότητας είναι η ενσωμάτωση πολλαπλών μοντέλων πρόβλεψης [133], έτσι ώστε να ενισχυθεί η αξιοπιστία των τελικών προβλέψεων. Τέλος, ιδιαίτερα αποτελεσματική έχει αποδειχθεί η χρήση ασαφών συνόλων (U, μ) [134], με U ένα μη κενό σύνολο, όπου η αβεβαιότητα κάθε παρατήρησης $x \in U$, μοντελοποιείται μέσω μιας συνάρτησης συμμετοχής (membership function), $\mu: U \mapsto [0,1]$.

2.3 Η μηχανική μάθηση ως εργαλείο πρόβλεψης χαρακτηριστικών από χρονοσειρές που προκύπτουν από χρονικά εξελισσόμενα συστήματα

Αναγνωρίζοντας τόσο τα πλεονεκτήματα όσο και τις ελλείψεις-αδυναμίες των κλασσικών στατιστικών μεθόδων, η μηχανική μάθηση έχει καθιερωθεί πλέον ως ο κύριος ανταγωνιστής τους στην περιοχή της πρόβλεψης χαρακτηριστικών από χρονοσειρές. Τα μοντέλα μηχανικής μάθησης, αποκαλούμενα επίσης και μοντέλα μαύρου κουτιού (black box) ή οδηγούμενα από τα δεδομένα (data-driven) μοντέλα [135], είναι παραδείγματα μη παραμετρικών, μη γραμμικών μοντέλων τα οποία χρησιμοποιούν ιστορικά δεδομένα χρονοσειρών για την εκμάθηση της στοχαστικής εξάρτησης μεταξύ παρελθοντικής και μελλοντικής συμπεριφοράς των δεδομένων.

Η μηχανική μάθηση αποτελεί ένα επιμέρους πεδίο της υπολογιστικής νοημοσύνης μέσω του οποίου οι υπολογιστές καθίστανται ικανοί να τροποποιούν και να βελτιώνουν τις ενέργειές τους (εν προκειμένω τις προβλέψεις τους) για να επιτύχουν υψηλότερη ακρίβεια, η οποία μετράται δια μέσω του σφάλματος μεταξύ των προσεγγιστικών αποτελεσμάτων των επιλεγμένων μοντέλων και των πραγματικών [136]. Η υπολογιστική νοημοσύνη και η μηχανική μάθηση είναι δύο διαφορετικές έννοιες, που όμως συχνά συγχέονται. Η υπολογιστική νοημοσύνη είναι μια ευρεία έννοια που περικλείει μηχανές λήψης αποφάσεων, εκμάθηση νέων δεξιοτήτων και επίλυση προβλημάτων με τον τρόπο που μιμείται αυτόν των ανθρώπων, ενώ η μηχανική μάθηση είναι ένα υποσύνολο της τεχνητής νοημοσύνης που επιτρέπει στα ευφυή συστήματα να μαθαίνουν νέες πληροφορίες από τα δεδομένα. Ο στόχος της μηχανικής μάθησης είναι συχνά η ανακάλυψη της δομής των δεδομένων και η προσαρμογή τους σε μοντέλα τα οποία μπορούν να γίνουν κατανοητά και να χρησιμοποιηθούν από ανθρώπους. Είναι γεγονός πως τα τελευταία πενήντα χρόνια, έχει υπάρξει μια τεράστια αύξηση του όγκου πληροφοριών/δεδομένων που χρησιμοποιούν τα σύγχρονα πληροφοριακά συστήματα. Βέβαια, απουσία μεθόδων για την εξόρυξη χρήσιμων χαρακτηριστικών και ιδιοτήτων από τα δεδομένα, η πληροφορία αυτή είναι παραμένει ανώφελη. Την απαίτηση αυτή έρχεται να καλύψει η μηχανική μάθηση [137]. Έτσι τα κρυφά μοτίβα και οι συσχετίσεις των δεδομένων μπορούν να χρησιμοποιηθούν για την πρόβλεψη μελλοντικών τιμών μιας χρονοσειράς.

Σε ότι αφορά το αντικείμενο της παρούσας Δ.Δ. το βασικό καθήκον της μηχανικής μάθησης είναι η κατασκευή αλγορίθμων που μπορούν να λάβουν δεδομένα εισόδου και με τη χρήση μαθηματικών μοντέλων να εξάγουν προβλέψεις αναθεωρώντας τες καθώς γίνονται διαθέσιμα νέα δεδομένα. Η κύρια διαφοροποίηση των μεθόδων μηχανικής μάθησης σε σχέση με τις παραδοσιακές είναι ότι δεν απαιτείται ο πλήρης καθορισμός από το σχεδιαστή των κανόνων που διέπουν το μοντέλο πρόβλεψης [138]. Αντιθέτως, επιτρέπεται η εκμάθηση νέων κανόνων κατά τη διαδικασία της εκπαίδευσης.

Η ερευνητική δραστηριότητα στην περιοχή της πρόβλεψης χαρακτηριστικών από χρονοσειρές με μεθοδολογίες μηχανικής μάθησης ξεκίνησε κατά τη δεκαετία του 1980 με τη ανάπτυξη του μοντέλου νευρωνικού δικτύου [139]. Στη συνέχεια η έρευνα επεκτάθηκε με την υλοποίηση περαιτέρω μοντέλων όπως οι μηχανές διανυσμάτων υποστήριξης [140], τα δέντρα αποφάσεων [141], και πολλά ακόμη [142]. Τα τελευταία χρόνια ειδικότερα, έχει σημειωθεί εντυπωσιακή πρόοδος τόσο ως προς τη θεωρητική μελέτη όσο και ως προς την ποσότητα και την καινοτομία των προτεινόμενων μοντέλων. Η εμπειρική έρευνα έχει δείξει ότι οι αλγόριθμοι μηχανικής μάθησης για την πρόβλεψη χρονοσειρών παρέχουν πολύ ανταγωνιστικά αποτελέσματα, συχνά ξεπερνώντας τα στατιστικά μοντέλα [143], [144], [145].

Η υψηλή επίδοση των μεθόδων μηχανικής μάθησης στην πρόβλεψη χρονοσειρών οφείλεται σε μια σειρά από πλεονεκτήματα που παρουσιάζουν [146]:

1. Αναγνωρίζουν εύκολα και αποτελεσματικά τη συσχέτιση των δεδομένων ακόμη κι αν είναι μη γραμμική, διότι δεν απαιτούν την καθορισμένη δομή του μοντέλου πρόβλεψης.
2. Παρέχουν υψηλό βαθμό αυτοματοποίησης εφόσον η εκπαίδευση του μοντέλου και η εξαγωγή των προβλέψεων πραγματοποιούνται από τον ηλεκτρονικό υπολογιστή.
3. Διαχειρίζονται αποτελεσματικά δεδομένα κάθε είδους και ανεξαρτήτως ποσότητας.
4. Η απόδοσή τους βελτιώνεται συνεχώς μέσα από τη διαδικασία της εκπαίδευσης. Οι παράμετροι του μοντέλου ανανεώνονται έτσι ώστε αυτό να προσαρμόζεται σταδιακά καλύτερα στα δεδομένα αξιολόγησης.

Από την άλλη μεριά, οι μεθοδολογίες μηχανικής μάθησης συνοδεύονται και από ορισμένα μειονεκτήματα ή παράγοντες που ενδεχομένως μειώνουν την προβλεπτική του επίδοση. Παραθέτουμε παρακάτω μια λίστα από αυτά:

1. Μια από τις βασικότερες αιτίες εμφάνισης προβλέψεων χαμηλής ποιότητας από ένα μοντέλο μηχανικής μάθησης είναι η κακή ποιότητα των διαθέσιμων δεδομένων η οποία με τη σειρά της μπορεί να οφείλεται σε εσφαλμένες ή διακοπτόμενες εγγραφές δεδομένων ή σε θόρυβο. Αυτό έχει σαν αποτέλεσμα το μοντέλο να εκπαιδευτεί με μη έγκυρες παρατηρήσεις και οι μετέπειτα προβλέψεις που θα παρέχει να παρουσιάζουν υψηλό σφάλμα. Επομένως, είναι σημαντικό να χρησιμοποιούνται δεδομένα επαρκώς επικυρωμένα ως προς την ποιότητά τους. Ωστόσο, ελλείψει τέτοιων μετρήσεων, υπάρχουν πολλές στατιστικές μέθοδοι για τον μετριασμό του αντίκτυπου του σφάλματος μέτρησης, συμπεριλαμβανομένης της χρήσης μοντέλων μέτρησης λανθάνουσας μεταβλητής (latent variable) [147], ποσοτικής ανάλυσης μεροληψίας [148], τεχνικών που διαχειρίζονται τα σφάλματα ως ελλείπουσες τιμές [149], Μπεϋζιανών μεθόδων [150], και άλλων.
2. Απαιτούν μεγάλο όγκο δεδομένων κατά την εκπαίδευσή τους, προκειμένου να παρέχουν ακριβείς προβλέψεις. Μάλιστα όσο μεγαλύτερο είναι το σύνολο δεδομένων εκπαίδευσης και όσο καλύτερη η ποιότητά τους, τόσο πιο αξιόπιστο είναι το τελικό μοντέλο. Η διασταυρούμενη αξιολόγηση (cross validation) αποτελεί μια συνήθη προσέγγιση στην περίπτωση που διατίθεται περιορισμένος αριθμός δεδομένων.
3. Μια ακόμη πρόκληση είναι ο υπολογιστικός φόρτος των μεθόδων μηχανικής μάθησης, δηλαδή περιορισμοί μνήμης, χρονοβόρες εκτελέσεις κ.λ.π., ειδικά για μεγάλα σύνολα δεδομένων. Η εξέλιξη στην υπολογιστική ικανότητα των ηλεκτρονικών υπολογιστών, καθώς και η βελτιστοποίηση υλικού και λογισμικού μέσα από τη χρήση επιταχυντών υλικού (hardware accelerators) [151] και τον προγραμματισμό σε γλώσσες χαμηλού επιπέδου που να στηρίζουν τους παράλληλους υπολογισμούς, μπορούν να μετριάσουν σημαντικά τα εν λόγω ζητήματα.
4. Συχνά οι ειδικοί που ασχολούνται με προβλήματα πρόβλεψης χρησιμοποιούν μοντέλα μηχανικής μάθησης ανεξέλεγκτα και καταχρηστικά, χωρίς να γνωρίζουν την εσωτερική λειτουργία και δομή τους. Σαν αποτέλεσμα, η εφαρμογή των μοντέλων αυτών καθίσταται αναξιόπιστη για τους χρήστες, φαινόμενο που εγκυμονεί σοβαρούς κινδύνους [152], [153]. Η ερμηνευσιμότητα είναι ένα μέτρο που δείχνει το βαθμό στον οποίο ένας άνθρωπος μπορεί να κατανοήσει τον λόγο πίσω από τις αποφάσεις (δηλαδή τις προβλέψεις) που λαμβάνονται από τα μοντέλα μηχανικής μάθησης. Το πρωταρχικό καθήκον της δημιουργίας ενός ερμηνεύσιμου μοντέλου είναι η ενσωμάτωση τεχνικών λευκού κουτιού, όπως τα ερμηνεύσιμα

νευρωνικά δίκτυα (explainable neural networks - XNN) [154], οι επεκτάσιμες Μπεϋζιανές λίστες κανόνων (scalable Bayesian rule lists) [155], η μηχανές ενίσχυσης βάθμωσης (gradient boost machine - GBM) [156] κ.λπ. Ένα σημαντικό συμπέρασμα που προκύπτει είναι ότι η ερμηνευσιμότητα αυξάνει την αξιοπιστία των προβλέψεων ενός μοντέλου, καθώς ο υπολογισμός των μέτρων αξιολόγησης της ακρίβειας των προβλέψεων, δεν επαρκεί για να περιγράψει την αποτελεσματικότητα του μοντέλου στις περισσότερες εφαρμογές του πραγματικού κόσμου. Πληρώντας συγκεκριμένες προϋποθέσεις [157], [158], οι ερμηνεύσιμες μέθοδοι μηχανικής μάθησης παρέχουν πολλαπλά οφέλη, όπως:

α. βοηθούν τους χρήστες να εξάγουν ερμηνεύσιμα μοτίβα από εκπαιδευμένα μοντέλα μηχανικής μάθησης, δηλαδή οι χρήστες μπορούν να εντοπίσουν τους λόγους πίσω από μη έγκυρες προβλέψεις,
β. βοηθούν τους χρήστες να αυξήσουν την εμπιστοσύνη στις προβλέψεις των μοντέλων, δηλαδή οι χρήστες μπορούν να ανιχνεύουν την μεροληψία σε μοντέλα μηχανικής μάθησης,
γ. βοηθούν στην αντιμετώπιση της υπέρ-προσαρμογής κατά την εκπαίδευση ενός μοντέλου.

Οι μεθοδολογίες μηχανικής μάθησης ενδέχεται να διαφέρουν ως προς τον τρόπο με τον οποίο πραγματοποιείται η εκπαίδευσή τους. Έτσι εντοπίζονται τρεις κατηγορίες: α) Επιβλεπόμενη μάθηση (supervised learning), β) μη επιβλεπόμενη μάθηση (unsupervised learning) και γ) ενισχυτική μάθηση (reinforcement learning). Παρακάτω παρουσιάζονται οι μεθοδολογίες μηχανικής μάθησης που αναπτύχθηκαν στα πλαίσια της παρούσας διατριβής, και οι οποίες κατατάσσονται στις κατηγορίες μάθησης με και χωρίς επίβλεψη.

2.3.1 Επιβλεπόμενη μάθηση

Η εκμάθηση με επίβλεψη προϋποθέτει την πρότερη γνώση της αντιστοίχισης μεταξύ των τιμών των μεταβλητών εισόδου και των κατηγοριών της εξόδου. Συνεπώς, κατά την εκπαίδευση ενός μοντέλου με έναν αλγόριθμο επιβλεπόμενης μάθησης, τα δεδομένα εκπαίδευσης πρέπει να είναι επισημασμένα (labeled), που σημαίνει ότι το σύνολο δεδομένων, μεγέθους N, αποτελείται από N πλειάδες (tuples) όπου κάθε μια περιλαμβάνει τις τιμές των μεταβλητών εισόδου και εξόδου για τη χρονική στιγμή $t_i, i = 1, \ldots, N$, η οποία αναφέρεται στην παρατηρούμενη χρονοσειρά. Στόχος του αλγορίθμου είναι να αναζητήσει μοτίβα συσχέτισης μεταξύ των τιμών των μεταβλητών εισόδου και εξόδου του συνόλου δεδομένων εκπαίδευσης, και έπειτα, τροφοδοτούμενος με δεδομένα που περιέχουν νέες τιμές των μεταβλητών εισόδου, να είναι σε θέση να προβλέψει σωστά την τιμή της μεταβλητής εξόδου. Επομένως, οι αλγόριθμοι επιβλεπόμενης μάθησης είναι σχεδιασμένοι έτσι ώστε να μαθαίνουν μέσα από παραδείγματα ζευγών εισόδου – εξόδου. Ορισμένες από τις πιο διαδεδομένες μεθόδους μάθησης με επίβλεψη στο τομέα της πρόβλεψης χαρακτηριστικών από χρονοσειρές είναι η γραμμική παλινδρόμηση, η λογιστική παλινδρόμηση, τα τυχαία δάση, οι μηχανές διανυσμάτων υποστήριξης κ.α. [159], [160].

2.3.1.1 Μέθοδος ανσάμπλ

Η προσέγγιση που περιγράφεται στην παράγραφο αυτή δεν αποτελεί ουσιαστικά μια ξεχωριστή μέθοδο μηχανικής μάθησης αλλά μια μέθοδο ανσάμπλ, που συνδυάζει πολλαπλές επιμέρους μεθοδολογίες και συγκεκριμένα αυτές της γραμμικής παλινδρόμησης, της αραιής κωδικοποίησης, της παλινδρόμησης με χρήση διανυσμάτων υποστήριξης, των τυχαίων δασών, των νευρωνικών δικτύων radial basis function (RBF) και των νευρωνικών δικτύων multi-layer perceptron (MLP). Αναγνωρίζοντας τα επιμέρους πλεονεκτήματα και μειονεκτήματα αυτών των μεθόδων μηχανικής, το προτεινόμενο σχήμα πρόβλεψης επιδιώκει να δημιουργήσει ένα μοντέλο ανσάμπλ που θα συνδυάζει επιτυχώς τα πλεονεκτήματά τους σε μια ενιαία προσέγγιση. Για παράδειγμα, τα μοντέλα που βασίζονται σε νευρωνικά δίκτυα, RBF, παρουσιάζουν ανώτερη απόδοση πρόβλεψης, μόνο εφόσον το προς πρόβλεψη σημείο δεδομένων βρίσκεται εντός του πεδίου όπου βρίσκονται τα δεδομένα εκπαίδευσης, διαφορετικά αποτυγχάνει. Από την άλλη πλευρά, τα γραμμικά και αραιά μοντέλα πρόβλεψης εμφανίζουν γενικά πολύ καλύτερη ικανότητα παρεκβολής (extrapolation), παρόλο που δεν είναι σε θέση να αναπαραστήσουν πιο περίπλοκες, μη γραμμικές δυναμικές. Με άλλα λόγια, με την εναλλαγή μεταξύ των στιβαρών γραμμικών μοντέλων και των πιο ευαίσθητων αλλά και πιο αποτελεσματικών μη γραμμικών μοντέλων, μπορεί να κατασκευαστεί μια προσέγγιση ανώτερης απόδοσης για την πρόβλεψη χρονοσειρών ηλεκτρικού φορτίου.

Προκειμένου να επιτευχθεί η καλύτερη δυνατή απόδοση κάθε υπο-μοντέλου, πρέπει να καθοριστεί η βέλτιστη διαμόρφωση για την εκπαίδευσή τους. Ξεκινώντας από τις απλούστερες μεθόδους που χρησιμοποιούνται, ένα γραμμικό και ένα μοντέλο αραιής αναπαράστασης εκπαιδεύονται με τον αλγόριθμο ελαχίστων τετραγώνων και τον αλγόριθμο fast iterative shrinkage thresholding (FISTA), αντίστοιχα, εκ των οποίων ο τελευταίος αποτελεί μια ταχύτερη υλοποίηση του αλγόριθμου iterative shrinkage thresholding algorithm (ISTA) [161]. Στην περίπτωση της προσέγγισης αραιής κωδικοποίησης, η αραιότητα επιβάλλεται από τη νόρμα ℓ_2 και η τιμή της παραμέτρου κανονικοποίησης ορίστηκε με δοκιμή και σφάλμα στο 0.01. Στη συνέχεια, χρησιμοποιείται ένα μοντέλο τυχαίου δάσους, όπου ο αριθμός των δέντρων απόφασης επιλέγεται να είναι 15, ώστε να διατηρείται ο απαιτούμενος χρόνος εκπαίδευσης του μοντέλου σε λογικά επίπεδα, χωρίς να μειώνεται η προγνωστική του ικανότητα. Όσον αφορά τις μη γραμμικές μεθόδους, αναπτύχθηκε ένα μοντέλο SVR με Γκαουσιανή συνάρτηση πυρήνα [162], χρησιμοποιώντας τις τενικές της διαδοχικής ελάχιστης βελτιστοποίησης (sequential minimal optimization - SMO) [163] για την εκπαίδευση του και της Μπεϋζιανής βελτιστοποίηση για τη βελτιστοποίηση των υπερπαραμέτρων του [164]. Στο σύνολο των υπο-μοντέλων εισάγονται επίσης δύο μοντέλα τεχνητών νευρωνικών δικτύων – Τ.Ν.Δ., τα οποία βασίζονται σε δύο διαφορετικές αρχιτεκτονικές. Το πρώτο είναι ένα δίκτυο MLP δύο κρυφών στοιβάδων που εκπαιδεύεται με τον αλγόριθμο οπισθοδιάδοσης Levenberg-Marquardt [165], σε συνδυασμό με 10-πλή διασταυρούμενη αξιολόγηση. Ο αριθμός των νευρώνων κάθε στοιβάδας επιλέγεται με δοκιμή και σφάλμα ως 20 και 10 αντίστοιχα. Σημειώνεται ότι, προκειμένου να μετριαστεί το μειονέκτημα της εξάρτησης της απόδοσης των μεθόδων εκπαίδευσης MLP από την αρχικοποίηση των βαρών, η διαδικασία της εκπαίδευσης διεξάγεται 10 φορές, με διαφορετικά τυχαία αρχικοποιημένα βάρη του δικτύου. Το δεύτερο μοντέλο Τ.Ν.Δ. ενσωματώνει μια αρχιτεκτονική RBF και εκπαιδεύεται χρησιμοποιώντας την τεχνική των ασαφών μέσων (fuzzy means) [166], ένας αλγόριθμος που έχει να επιδείξει πολλές επιτυχημένες εφαρμογές λόγω της αυξημένης ακρίβειας που παρέχει [167]–[169] σε συνδυασμό με χαμηλό χρόνο που απαιτεί για την εκπαίδευση ενός μοντέλου [170]. Στα πλαίσια των πειραμάτων που διεξάγονται για τις ανάγκες της Δ.Δ., ο αλγόριθμος ασαφών μέσων εκτελείται για ένα εύρος ασαφών συνόλων μεταξύ 4 και 15. Όταν εφαρμόζεται σε πραγματικό χρόνο, η προτεινόμενη προσέγγιση ανσάμπλ υπολογίζει την τιμή του δείκτη MAE σε ένα κυλιόμενο παράθυρο προηγούμενων προβλέψεων που προέρχονται από μια ομάδα εκπαιδευμένων μοντέλων, προκειμένου να δημιουργήσει ένα διάνυσμα που περιέχει τα βάρη των επιμέρους προβλέψεων και χρησιμοποιείται για την υπολογισμό της πρόβλεψης του επόμενου χρονικού βήματος. Μία σημαντική παράμετρος της προτεινόμενης μεθόδου που πρέπει να προσδιοριστεί είναι το μήκος του κυλιόμενου παραθύρου. Η τιμή της παραμέτρου αυτής εξαρτάται όχι μόνο από τον ορίζοντα πρόβλεψης, αλλά και από τις στατιστικές ιδιότητες της προβλεπόμενης μεταβλητής (μια μη στάσιμη χρονοσειρά που χαρακτηρίζεται από μεταβλητότητα θα απαιτούσε μικρότερο μήκος κυλιόμενου παραθύρου). Όταν η εκπαίδευση όλων των υπο-μοντέλων ολοκληρωθεί, το βέλτιστο μήκος του κυλιόμενου παραθύρου, το εύρος του οποίου τίθεται από 3 έως 15 χρονικά βήματα, υπολογίζεται μέσω εξαντλητικής αναζήτησης (exhaustive search) χρησιμοποιώντας τα ίδια δεδομένα αξιολόγησης. Η προτεινόμενη μέθοδος λειτουργεί ως εξής: για κάθε χρονικό βήμα k, η τρέχουσα απόδοση πρόβλεψης κάθε εκπαιδευμένου υπο-μοντέλου αξιολογείται υπολογίζοντας την τιμή του δείκτη MAE για έναν αριθμό προηγούμενων προβλέψεων αυτών, που αφορούν σε ένα κυλιόμενο χρονικό διάστημα h_w

$$MAE_i(k) = \frac{\sum_{j=0}^{h_w-1} |\hat{y}_i(k-j) - y(k-j)|}{h_w} \tag{2.12}$$

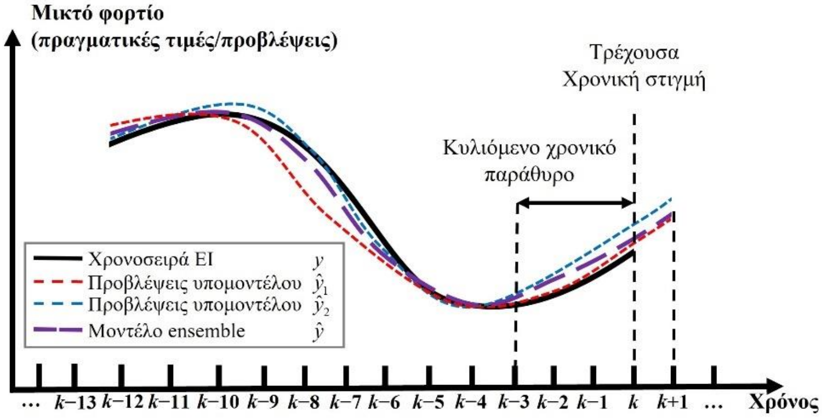
όπου με $\hat{y}_i(k)$ δηλώνονται οι προβλέψεις του i-οστού μοντέλου και με y οι πραγματικές τιμές της χρονοσειράς την χρονική στιγμή k. Έπειτα, ο δείκτης MAE χρησιμοποιείται για να υπολογιστεί το βάρος πρόβλεψης κάθε υπο-μοντέλου για το επόμενο χρονικό βήμα $k+1$

$$w_i(k+1) = \frac{MAE_i^{-1}(k)}{\sum_{i=1}^{N} MAE_i^{-1}(k)} \tag{2.13}$$

όπου MAE_i είναι η τιμή του MAE του i-οστού μοντέλου πρόβλεψης, N είναι ο συνολικός αριθμός υπο-μοντέλων και w_i είναι το βάρος πρόβλεψης για το επόμενο χρονικό βήμα. Η τελική έξοδος του προτεινόμενου μοντέλου υπολογίζεται ως το σταθμισμένο άθροισμα των προβλέψεων $\hat{y}_i$ που παράγονται από τα επιμέρους μοντέλα

$$\hat{y}(k+1) = \sum_{i=1}^{N} w_i(k+1)\,\hat{y}_i(k+1) \tag{2.14}$$

Ένα στιγμιότυπο της λειτουργίας μιας δοκιμαστικής έκδοσης της προτεινόμενης μεθόδου που περιλαμβάνει δύο υπο-μοντέλα, απεικονίζεται στο Σχήμα 3. Θα πρέπει να τονιστεί ότι η προτεινόμενη μέθοδος συνδυάζει τα πλεονεκτήματα των μεμονωμένων μοντέλων, τοποθετώντας μεγαλύτερο βάρος στο τρέχον μοντέλο με την καλύτερη απόδοση για το χρονικό παράθυρο μήκους h_w. Σύμφωνα με το σχήμα, οι προβλέψεις και των δύο μοντέλων $\hat{y}_1$ και $\hat{y}_2$ εμφανίζουν σχετικά υψηλό σφάλμα. Ωστόσο, με μια πιο προσεκτική παρατήρηση, φαίνεται πως οι προβλέψεις $\hat{y}_2$ είναι πιο ακριβείς κατά μήκος του πρώτου μισού της καμπύλης του y, και αντίστοιχα οι προβλέψεις $\hat{y}_1$ γαι το δεύτερο μισό. Αποδίδοντας μεγαλύτερο βαρος στο μοντέλο με την καλύτερη απόδοση προηγούμενων προβλέψεων εντός του χρονικού παραθύρου, μήκους h_w, η προτεινόμενη μέθοδος μπορεί να εναλλάσσει τις τιμές των υπολογιζόμενων βαρών, επιλέγοντας το καλύτερο διαθέσιμο μοντέλο για την τρέχουσα χρονική στιγμή. Με τον τρόπο αυτό επιτυγχάνεται μια συνολικά ανώτερη απόδοση πρόβλεψης.



Σχήμα 3: Σχηματική αναπαράσταση της δοκιμαστικής έκδοσης της προτεινόμενης μεθόδου που ενσωματώνει 2 υπο-μοντέλα. Η χρονοσειρά των πραγματικών δεδομένων του φορτίου δηλώνεται με y, οι προβλέψεις του i-οστού μοντέλου με $\hat{y}_i$, η σταθμισμένες προβλέψεις με $\hat{y}$ και η τρέχουσα χρονική στιγμή με k. Το μοντέλο ανσάμπλ αναγνωρίζοντας την υπεροχή των $\hat{y}_1$ έναντι των $\hat{y}_2$, εντός του κυλιόμενου παραθύρου, προσαρμόζει τα βάρη του έτσι ώστε να επιτυγχάνει μια πρόβλεψη με την υψηλότερη δυνατή ακρίβεια για το επόμενο χρονικό βήμα k+1.

2.3.1.2 Νευρωνικά δίκτυα γράφων ορατότητας

Κατά τη διάρκεια της πολύχρονης ερευνητικής δραστηριότητας στον τομέα της πρόβλεψης χαρακτηριστικών από χρονοσειρές, οι μεθοδολογίες μάθησης με επίβλεψη έχουν χρησιμοποιηθεί κατά κόρον. Ωστόσο, η πλειοψηφία των σύγχρονων εφαρμογών πρόβλεψης αναφέρονται σε χρονοσειρές με πολύπλοκη δυναμική συμπεριφορά, μη γραμμικότητα και υψηλή αβεβαιότητα, όπως αυτή του ηλεκτρικού φορτίου, χαρακτηριστικά τα οποία δεν μπορούν να αντιμετωπιστούν αποτελεσματικά από κάθε μέθοδο. Στις προσπάθειές της για αναβάθμιση των τωρινών τεχνικών, η ερευνητική κοινότητα έχει στραφεί στην βαθιά μάθηση με Γράφους [171]–[173]. Ανάμεσα στις μοντέλα εκμάθησης που έχουν αναπτυχθεί, ξεχωρίζουν τα νευρωνικά δίκτυα Γράφων (Graph Neural Networks - GNN) λόγω της ικανότητάς τους να δεδομένα με πολύπλοκες δομές [174]. Τα νευρωνικά δίκτυα γράφων είναι μέθοδοι βασιζόμενες στη βαθιά μάθηση που ενσωματώνουν θεωρία Γράφων για την αναπαράσταση της συσχέτισης των παρατηρήσεων της χρονοσειράς. Στο πρώτο στάδιο, τα δεδομένα εισόδου αναπαρίστανται ως δεδομένα με δομή γράφου. Οι Γράφοι που προκύπτουν χρησιμοποιούνται στη συνέχεια από τα νευρωνικά δίκτυα Γράφων για την εκμάθηση του τρόπου σύνδεσης των κόμβων και των ακμών τους, προκειμένου να εκτελεστεί στη συνέχεια η πρόβλεψη χαρακτηριστικών ή άλλες εργασίες όπως η ταξινόμηση κόμβων, η ομαδοποίηση γράφων. Η ισχύς τέτοιων μοντέλων έχει αποδειχθεί μέσα από την ανώτερη επίδοσή τους σε μια πληθώρα πεδίων που κυμαίνονται από την υπολογιστική όραση [175], [176] και την επεξεργασία φυσικής γλώσσας [177] έως την αστική νοημοσύνη (urban intelligence) [178], [179] και τα έξυπνα δίκτυα [180]–[184]. Αυτές οι εργασίες απαιτούν την διαχείριση του φαινομένου ότι τα δεδομένα των Γράφων δεν ορίζουν έναν Ευκλείδειο χώρο, και ως εκ τούτου δεν μπορούν να αντιμετωπιστούν σωστά από παραδοσιακά μοντέλα βαθιάς μάθησης (π.χ. συνελικτικά νευρωνικά δίκτυα (Convolutional Neural Networks - CNN) ή αναδρομικά νευρωνικά δίκτυα (Recurrent Neural Networks - RNN)) [185]–[187].

Παρά τα αξιοσημείωτα αποτελέσματα των νευρωνικών δικτύων γράφων, η αποδοτικότητά τους αμφισβητείται λόγω της υψηλής υπολογιστικής πολυπλοκότητας που παρουσιάζουν κατά το στάδιο εκμάθησης που προκύπτει από τη χρήση του πίνακα γειτνίασης που αντιπροσωπεύει τις ακμές ενός γράφου. Παρόλο που έχουν διεξαχθεί αρκετές μελέτες σχετικά με την ανάλυση της πολυπλοκότητας [188], [189] και έχουν παρουσιαστεί ορισμένες εξελίξεις [190], [191], αυτό παραμένει ένα ανοιχτό ζήτημα προς διερεύνηση. **Στα πλαίσια τις παρούσας διδακτορικής διατριβής, αυτό το ερευνητικό κενό επιδιώκεται να καλυφθεί με μια εναλλακτική προσέγγιση για την κατασκευή του γράφου που αναπαριστά τα δεδομένα εισόδου.**

Πέρα από την συνεισφορά των νευρωνικών δικτύων γράφων, η θεωρία Γράφων αποτελεί επίσης μια πολύτιμη πηγή γνώσης για την ανάλυση χρονοσειρών. Οι έννοιες της α) μετατροπής μιας χρονοσειράς σε δίκτυο β) της μελέτης των ιδιοτήτων της και γ) της εξαγωγής πληροφοριών για αυτή, έχουν προταθεί εδώ και πολλά χρόνια με τη θεωρία των Γράφων ορατότητας (Visibility Graph - VG) [192]. Ο γράφος ορατότητας διατηρεί αρκετές ιδιότητες της χρονοσειράς [193], [194] στη δομή του επιτρέποντας την διεξαγωγή εργασιών Γράφο-θεωρητικής ανάλυσης χρονοσειρών, όπως η εξαγωγή χαρακτηριστικών, ο τοπολογικός χαρακτηρισμός και η αναγνώριση δομικών ιδιοτήτων. Στη βιβλιογραφία συναντάται μια πληθώρα εφαρμογών των Γράφων ορατότητας που κυμαίνονται από την επεξεργασία εικόνας [195] και τη νευροεπιστήμη [196]–[198] μέχρι τα χρηματοοικονομικά [199], [200] και τις επικοινωνίες [201]. Ένας σημαντικός αριθμός δημοσιεύσεων έχει αφιερωθεί στην πρόβλεψη χρονοσειρών [202]. Οι περισσότερες από αυτές βασίζονται στη χρήση μέτρων ομοιότητας [203]–[206] μεταξύ των κόμβων για την εξαγωγή των προβλέψεων, ενώ σπάνια έχουν χρησιμοποιηθεί ευρετικές προσεγγίσεις [207]. Αν και οι προαναφερθείσες τεχνικές υπερτερούν των στατιστικών μεθόδους [208], [209], δεν μπορούν να συγκριθούν με τις μεθοδολογίες μηχανικής μάθησης, καθώς δεν διαθέτουν τη διαδικασία εκμάθησης ενός μοντέλου από ένα σύνολο δεδομένων εκπαίδευσης. Ωστόσο, πολύ λίγες προσπάθειες έχουν γίνει μέχρι στιγμής για την ενσωμάτωση Γράφων ορατότητας σε μοντέλα πρόβλεψης που βασίζονται σε αλγορίθμους μηχανικής μάθησης [210]–[215].

Λιγότερη ακόμα ερευνητική πρόοδος έχει σημειωθεί μέχρι στιγμής σχετικά με το συνδυασμό Γράφων ορατότητας και νευρωνικών δικτύων Γράφων. Αυτή αφορά αποκλειστικά και μόνο εφαρμογές αναγνώρισης προτύπων και ταξινόμησης [216]–[218]. Οι προσομοιώσεις που περιλαμβάνονται στα άρθρα αυτά, επιβεβαιώνουν ότι οι προτεινόμενες προσεγγίσεις ξεπερνούν σε απόδοση αυτές που στηρίζονται σε μεθόδους βαθιάς μάθησης. Επομένως φαίνεται να ενθαρρύνεται η περαιτέρω επιδίωξη της ανάπτυξης τέτοιων μοντέλων, που προορίζονται όχι μόνο για εφαρμογές ταξινόμησης αλλά και παλινδρόμησης. **Στην παρούσα διατριβή προτείνεται μια νέα μέθοδος για την βραχυπρόθεσμη πρόβλεψη του ηλεκτρικού φορτίου, η οποία θα αναφέρεται στη συνέχεια ως νευρωνικά δίκτυα Γράφων ορατότητας (Visibility Graph Neural Networks - VGNN) και η οποία αξιοποιεί αποτελεσματικά τα συνδυασμό των Γράφων ορατότητας και των νευρωνικών δικτύων γράφων. Αυτό περιγράφεται αναλυτικά στη συνέχεια.**

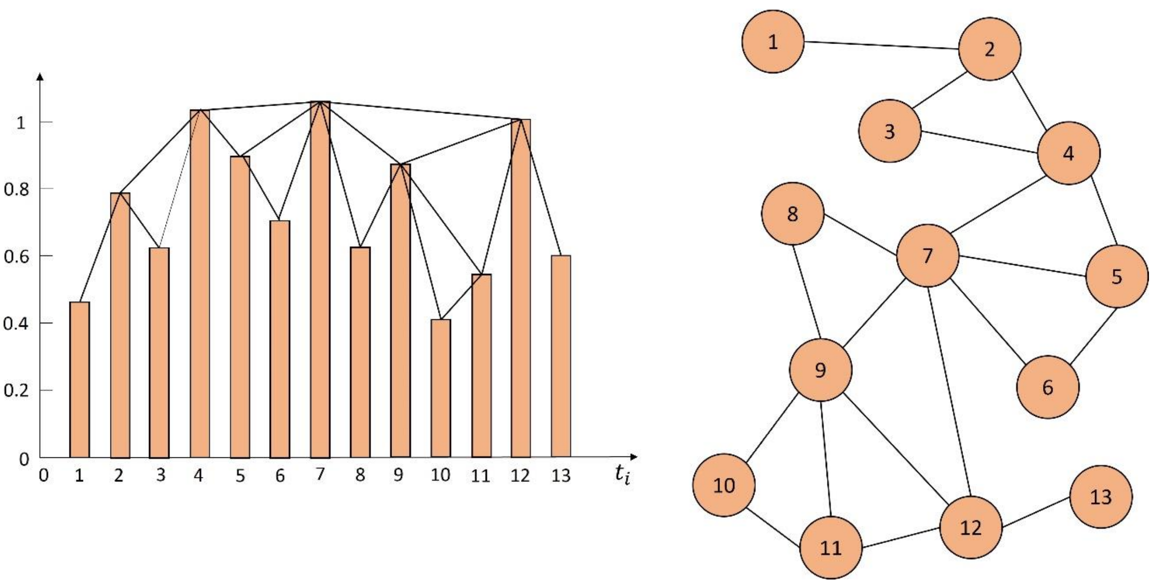
2.3.1.2.1 Στοιχεία θεωρίας Γράφων ορατότητας

Οι γράφοι ορατότητας προτάθηκαν ως εργαλείο για την ανάλυση χρονοσειρών αρχικά από τον Lacasa κ.α. [219]. Σύμφωνα με αυτούς, ο Γράφος ορατότητας είναι ένας αλγόριθμος που δημιουργεί ένα μη κατευθυνόμενο δίκτυο, έχοντας δεδομένη ως είσοδο μία χρονοσειρά. Το βασικό κίνητρο πίσω από την ανάπτυξη αυτού του αλγορίθμου ήταν η πρόθεση για το σχεδιασμού ένος εργαλείου για το στατιστικό χαρακτηρισμό των χρονοσειρών, βασισμένο στη θεωρία Γράφων. Η ανάλυσή του στα πλαίσια της παρούσας έρευνας αρχίζει με την παρουσίαση του απλούστερου Γράφου ορατότητας, ο οποίος αποκαλείται ως Γράφος φυσικής ορατότητας (natural visibility graph - NVG) και ο οποίος ορίζεται παρακάτω ως εξής:

Ας θεωρήσουμε ως $\{x(t_i)\}_{i=1,\ldots,N}$ μια χρονοσειρά που αποτελείται από N μονοδιάστατες παρατηρήσεις, καθεμιά καταχωρημένη την χρονική στιγμή t_i. Θεωρούμε ότι κάθε παρατήρηση της χρονοσειράς αντιστοιχεί σε έναν κόμβο του Γράφου φυσικής ορατότητας NVG. Δύο κόμβοι i και j του γράφου χαρακτηρίζονται ως συνδεδεμένοι εάν υπάρχει μία οπτική επαφή (line of sight - LoS) μεταξύ των τιμών (ύψη) των παρατηρήσεων $x(t_i)$ και $x(t_j)$ και η οποία δεν τέμνει καμία ενδιάμεση τιμή (ύψος) παρατήρησης $x(t_k)$. Το Σχήμα 4 απεικονίζει την παραπάνω πρόταση ή οποία εκφράζεται με μαθηματικούς όρους υπό το ακόλουθο γεωμετρικό κριτήριο:

$$x(t_k) < x(t_i) + x\left(x(t_j) - x(t_i)\right)\frac{t_k - t_i}{t_j - t_k} \tag{2.15}$$

(α) (β)



Σχήμα 4: α) Τυχαίο παράδειγμα μιας χρονοσειράς 13 σημειακών δεδομένων σε μορφή διαγράμματος με μπάρες με τις αντίστοιχες γραμμές ορατότητας (lines of visibility - LoV) κάθε σημειακού δεδομένου. β) Ο σχετικός με τη χρονοσειρά μη κατευθυνόμενος γράφος που δημιουργείται από τον Γράφο ορατότητας χρησιμοποιώντας το κριτήριο (2.15). Κάθε κόμβος του γράφου αντιστοιχεί σε ένα σημειακό δεδομένο ενώ κάθε ακμή μεταξύ δύο κόμβων αντιστοιχεί στην γραμμή ορατότητας ανάμεσα στα δεδομένα αυτά.

Ο αλγόριθμος που υλοποιεί τον παραπάνω κανόνα προκειμένου να απεικονίσει μία χρονοσειρά ως δομικά στοιχεία ενός Γράφου ορατότητας είναι απλός και περιγράφεται λεπτομερώς στην βιβλιογραφία [220]. Ο Γράφος ορατότητας που παράγεται από μία χρονοσειρά κατέχει μια σειρά από χρήσιμες ιδιότητες οι οποίες περιγράφονται παρακάτω:

1. Ο Γράφος είναι συνδεδεμένος: δηλαδή, κάθε κόμβος βλέπει τουλάχιστον τους δύο κοντινότερους σε αυτόν κόμβους (αριστερά και δεξιά).
2. Ο Γράφος είναι μη κατευθυνόμενος: δηλαδή, ο τρόπος με τον οποίο κατασκευάζει ο αλγόριθμος τον Γράφο, δεν επιβάλλει την έννοια της κατεύθυνσης στους συνδέσμους μεταξύ των κόμβων.
3. Ο Γράφος είναι αμετάβλητος σε ομοπαραλληλικούς μετασχηματισμούς των δεδομένων της χρονοσειράς: δηλαδή, η μορφή του γράφου ορατότητας παραμένει

ίδια σε κάθε πιθανή αλλαγή κλίμακας των δεδομένων ως προς τον οριζόντιο ή τον κατακόρυφο άξονα αλλά και σε κάθε πιθανή οριζόντια ή κάθετη μετατοπίσή τους. Λαμβάνοντας ως αφετηρία την αξιωματική θεμελίωση του Γράφου φυσικής ορατότητας NVG, ένας αριθμός από παραλλαγές του αρχικού αλγορίθμου έχει προταθεί όπως: ο οριζόντιος Γράφος ορατότητας HVG [221], ο περιορισμένος διαπερατός γράφος ορατότητας (limited penetrable visibility graph - LPVG) [222], ο πολυμορφικός γράφος ορατότητας (multiplex visibility graph - MVG) [223], και ο γράφος ορατότητας εικόνας (image visibility graph - IVG) [224], με στόχο τη μελέτη ειδικών περιπτώσεων χρονοσειρών και τον εντοπισμό διαφορετικών ιδιοτήτων [202]. **Στα πλαίσια της εργασίας αυτής χρησιμοποιείται μόνο ο γράφος φυσικής ορατότητας.**

2.3.1.2.2 Νευρωνικά δίκτυα γράφων

Η τοπολογία των νευρωνικών δικτύων Γράφων προτάθηκε αρχικά από τον Scarselli κ.α. [225]. Τα νευρωνικά δίκτυα Γράφων είναι ειδικά σχεδιασμένες αρχιτεκτονικές νευρωνικών δικτύων που λειτουργούν με δεδομένα με δικτυωτή δομή. Πριν από την αναλυτική περιγραφή της διαδικασίας εκμάθησης ενός νευρωνικού δικτύου γράφων, κρίνεται απαραίτητο να παρατεθούν ορισμένες χρήσιμες σημειώσεις. Έστω $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ ένας μη κατευθυνόμενος Γράφος, όπου $\mathcal{V}$ έναι ο χώρος που απεικονίζει το σύνολο των κόμβων του και $\mathcal{E}$ το σύνολο των ακμών που συνδέουν τους κόμβους του. Οι διαστάσεις των $\mathcal{V}$ και $\mathcal{E}$ σημειώνονται ως $|\mathcal{V}| = N$ και $|\mathcal{E}| = M$ αντίστοιχα, κάτι το οποίο σημαίνει ότι ο γράφος $\mathcal{G}$ απαρτίζεται από N κόμβους και M ακμές. Επιπλέον. η ποσότητα $A \in \mathbb{R}^{N\times N}$ είναι μια ιδιότητα των νευρωνικών δικτύων Γράφων. Αυτή, καλείται πίνακας γειτνίασης, ο οποίος ορίζεται ως ένας πίνακας λογικών τιμών με στοιχεία $a_{ij} = 1$ ή $a_{ij} = 0$ που δηλώνουν ότι οι κόμβοι i και j συνδέονται ή όχι. Στην συνέχεια, περιγράφουμε κάποιες επιπλέον παραμέτρους των νευρωνικών δικτύων Γράφων. Συγκεκριμένα, κάθε κόμβος n συνοδεύεται από μία ετικέτα (label) l_n, η οποία εκφράζει ένα χαρακτηριστικό του κόμβου n, εδώ την κατανάλωση του ηλεκτρικού φορτίο, και ένα διάνυσμα κατάστασης $\boldsymbol{x}_n \in \mathbb{R}^N$, που αναπαριστά την πληροφορία που περιέχεται στην γειτονιά, $ne[n]$, του κόμβου n. Πιο συγκεκριμένα, η εξάρτηση του κόμβου n από την γειτονιά του μπορεί να εκφραστεί από την ακόλουθη μη γραμμική, παραμετρική συνάρτηση h:

$$\boldsymbol{x}_n = \sum_{u \in ne[n]} h\big(l_n, \boldsymbol{l}_{(n,u)}, \boldsymbol{l}_u, \boldsymbol{x}_u\big) \tag{2.16}$$

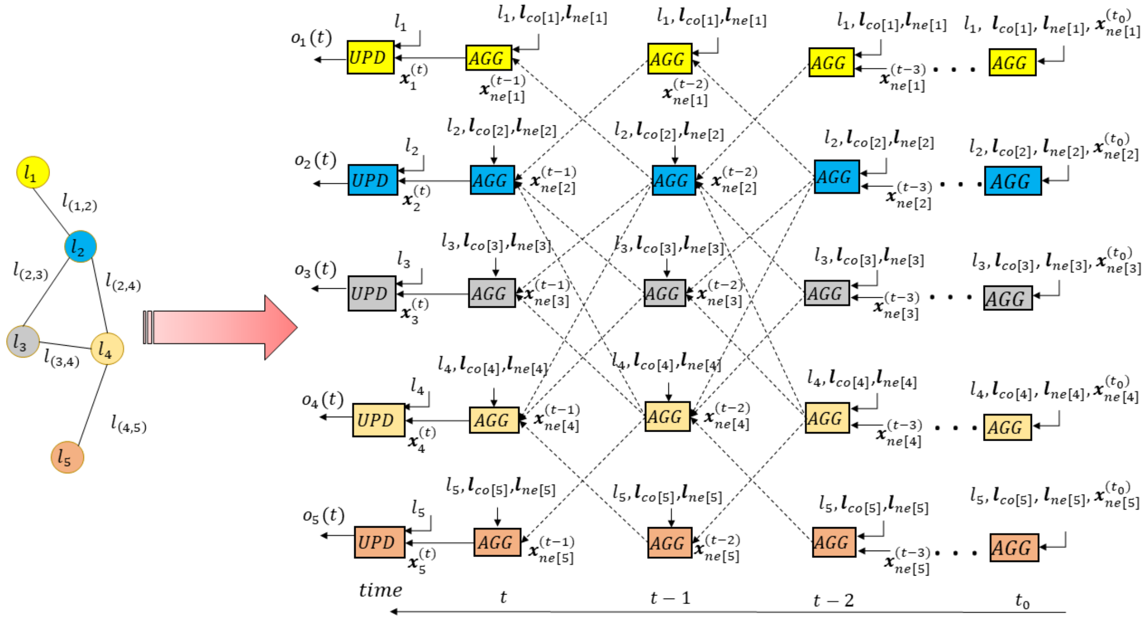
όπου με $\boldsymbol{l}_{(n,u)}, \boldsymbol{l}_u, \boldsymbol{x}_u$ δηλώνονται οι ετικέτες των ακμών, οι οποίες εκφράζουν κάποια χαρακτηριστικά που έχουν να κάνουν με τη σχέση των κόμβων που συνδέουν, οι ετικέτες των κόμβων και τα διανύσματα κατάστασης στην γειτονιά του n, αναφερόμενα εφεξής ως $\boldsymbol{x}_{ne[n]}$. Η συνάρτηση h είναι ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης (Feedforward Neural Network - FFNN), που έχει το ρόλο του να συγκεντρώνει τη πληροφορία από τη γειτονιά του κόμβου n, και ως εκ τούτου στο εξής θα αναφέρεται ως συνάρτηση $AGGREGATE$ ή AGG.

Η βασική λειτουργία ενός νευρωνικού δικτύου Γράφων είναι η ανανέωση των καταστάσεων όλων των κόμβων επαναληπτικά α) συγκεντρώνοντας την πληροφορία από τη γειτονιά του κάθε κόμβου και β) υπολογίζοντας την κατάσταση του κάθε κόμβου κατά την προηγούμενη επανάληψη ώστε να παραχθούν οι τιμές της μεταβλητής εξόδου, δηλαδή οι προβλέψεις της κατανάλωσης του ηλεκτρικού φορτίου. Η διαδικασία αυτή εκφράζεται από τις ακόλουθες σχέσεις:

$$\boldsymbol{x}_n(t+1) = \sum_{u \in ne[n]} AGGREGATE\left(l_n, \boldsymbol{l}_{(n,u)}, \boldsymbol{l}_u, \boldsymbol{x}_u(t)\right) \qquad , \forall n \in \mathcal{V} \tag{2.17}$$
$$\boldsymbol{o}_n(t) = UPDATE\left(\boldsymbol{x}_n(t), l_n\right)$$

όπου $UPDATE$ (ή UPD) είναι ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης μέσω του οποίου παράγονται οι τιμές εξόδου $\boldsymbol{o}_n$.

Για λόγους σαφήνειας στην παρουσίαση της μεθόδου, ένα παράδειγμα ενός νευρωνικού δικτύου Γράφων που εφαρμόζεται σε έναν τυχαίο γράφο 5 κόμβων απεικονίζεται στο Σχήμα 5. Η συγκεκριμένη απεικόνιση του μοντέλου αυτού δείχνει την μεταβίβαση των μηνυμάτων πληροφορίας κατά μήκος των t επιπέδων του δικτύου. Κάθε επίπεδο αντιστοιχεί σε μια χρονική στιγμή και περιέχει ένα αντίγραφο όλων των μονάδων του δικτύου κωδικοποίησης. Οι συνδέσεις μεταξύ των επιπέδων εξαρτώνται από τη συνδεσιμότητα του δικτύου κωδικοποίησης. Τις μονάδες του δικτύου αποτελούν τα AGG και UPD που αντιπροσωπεύουν τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης $AGGREGATE$ και $UPDATE$ αντίστοιχα.



Σχήμα 5: Επισκόπηση του Γράφου (αριστερά), και το εκτεταμένο δίκτυο, το οποίο απεικονίζει την αλληλουχία των επιπέδων του δικτύου κωδικοποίησης (δεξιά). Το μοντέλο συγκεντρώνει μηνύματα από τους τοπικούς γείτονες κάθε κόμβου, τα οποία μηνύματα με τη σειρά τους βασίζονται στην συγκεντρωμένη πληροφορία των δικών τους γειτόνων, κ.ο.κ. Κάθε χρονική στιγμή αναπαριστά ένα επίπεδο του νευρωνικού δικτύου Γράφων.

Μία λίστα που αποτελείται από α) τα σύνολα των ετικετών των ακμών, β) τις ετικέτες των κόμβων, και γ) τα διανύσματα κατάστασης σε κάθε χρονική στιγμή $\boldsymbol{t'}$, στην γειτονιά κάθε κόμβου δίνεται στον πίνακα Πίν.

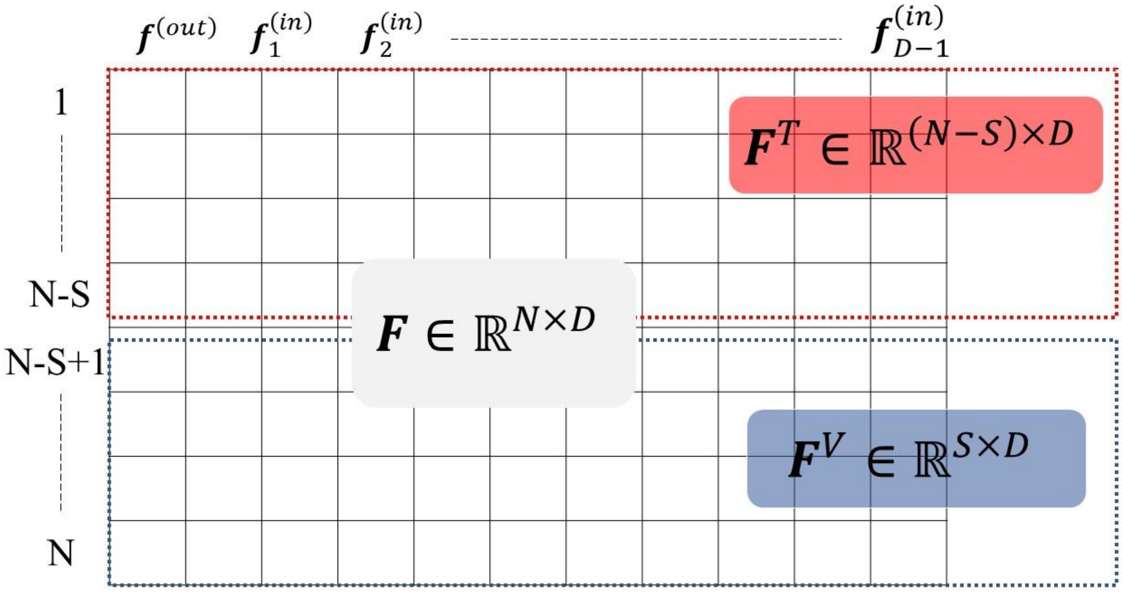
Πίνακας 1: Πληροφορία της γειτονιάς κάθε κόμβου ενός τυχαίου γράφου με 5 κόμβους

Πληροφορία	κόμβος 1	κόμβος 2	κόμβος 3	κόμβος 4	κόμβος 5
$\boldsymbol{l}_{co[n]}$	$\{l_{(1,2)}\}$	$\{l_{(1,2)}, l_{(2,3)}, l_{(2,4)}\}$	$\{l_{(2,3)}, l_{(3,4)}\}$	$\{l_{(2,3)}, l_{(3,4)}, l_{(4,5)}\}$	$\{l_{(4,5)}\}$
$\boldsymbol{l}_{ne\,[n]}$	$\{l_2\}$	$\{l_1, l_3, l_4\}$	$\{l_2, l_4\}$	$\{l_2, l_3, l_5\}$	$\{l_4\}$
$\boldsymbol{x}_{ne\,[n]}^{(t)}$	$\{\boldsymbol{x}_1^{(t)}, \boldsymbol{x}_2^{(t)}\}$	$\{\boldsymbol{x}_1^{(t)}, \boldsymbol{x}_2^{(t)}, \boldsymbol{x}_3^{(t)}, \boldsymbol{x}_4^{(t)}\}$	$\{\boldsymbol{x}_2^{(t)}, \boldsymbol{x}_3^{(t)}, \boldsymbol{x}_4^{(t)}\}$	$\{\boldsymbol{x}_2^{(t)}, \boldsymbol{x}_3^{(t)}, \boldsymbol{x}_4^{(t)}, \boldsymbol{x}_5^{(t)}\}$	$\{\boldsymbol{x}_4^{(t)}, \boldsymbol{x}_5^{(t)}\}$

Δοθέντος ενός συνόλου κόμβων με ετικέτες, η διαδικασία εκπαίδευσης έγκειται στην εκμάθηση των βαρών $\boldsymbol{W} = \{\boldsymbol{w}_i^j\} \in \mathbb{R}^{a_j \times c_j} \mid i \in \{AGG, UPD\}, j \in \{1, \ldots, p\}$ καθώς και της πόλωσης $\boldsymbol{b} = \{\boldsymbol{b}_i^j\} \in \mathbb{R}^{a_j} \mid i \in \{AGG, UPD\}, j \in \{1, \ldots, p\}$, όπου a_j είναι ο αριθμός των νευρώνων του j-οστού επιπέδου και c_j είναι ο αριθμός των νευρώνω του $j-1$-οστού επιπέδου. Όπως έχει ήδη αναφερθεί, οι συντμήσεις AGG and UPD αναφέρονται στα βάρη και τη μεροληψία των δικτύων $AGGREGATE$ και $UPDATE$ αντίστοιχα, ενώ οι εκθέτες $1, \ldots, p-1$ υποδεικνύουν τα κρυφά επίπεδα με p το επίπεδο εξόδου του κάθε δικτύου. Ο αλγόριθμος εκπαίδευσης που υιοθετείται στο προτεινόμενο μοντέλο θα παρουσιαστεί αμέσως μετά.

2.3.1.2.3 Νευρωνικά δίκτυα Γράφων ορατότητας

Τα δυο προαναφερθέντα θεωρητικά πλαίσια συγχωνεύονται για να δημιουργήσουν το προτεινόμενο μοντέλο, τα νευρωνικά δίκτυα Γράφων ορατότητας, το οποίο έχει σχεδιαστεί κατάλληλα ώστε να παρέχει προβλέψεις για τις μελλοντικές τιμές του ηλεκτρικού φορτίου. Το πρώτο στάδιο της διαδικασίας εκπαίδευσης των νευρωνικών δικτύων Γράφων ορατότητας είναι η κατασκευή του πίνακα $\boldsymbol{F} \in \mathbb{R}^{N \times D}$, ο οποίος φαίνεται στο Σχήμα 6.



Σχήμα 6. Γραφική απεικόνιση του πίνακα που περιέχει τα δεδομένα εκπαίδευσης του μοντέλου νευρωνικών δικτύων Γράφων ορατότητας.

Ο δείκτης κάθε μίας από τις N γραμμές του πίνακα F αναπαριστά μία χρονική στιγμή, ενώ οι D στήλες του αποτελούνται από το διάνυσμα της μεταβλητής εξόδου, δηλαδή του ηλεκτρικού φορτίου 24 ώρες μετά την τρέχουσα χρονική στιγμή, $\boldsymbol{f}^{(out)} \in \mathbb{R}^N$, τις τρέχουσες τιμές του φορτίου, $\boldsymbol{f}_1^{(in)} \in \mathbb{R}^N$ και τα διανύσματα των μεταβλητών εισόδου $\boldsymbol{f}_2^{(in)}, \ldots, \boldsymbol{f}_{D-1}^{(in)} \in \mathbb{R}^N$, που μπορεί να αναφέρονται σε τιμές φορτίου ή σε τιμές εξωγενών χαρακτηριστικών. Η μετατροπή του διανύσματος, $\boldsymbol{f}_1^{(in)} \in \mathbb{R}^N$, σε έναν γράφο φυσικής ορατότητας σύμφωνα με τον παραπάνω αλγόριθμο, είναι το πρώτο στάδιο της προτεινόμενης μεθόδου. Θα πρέπει να σημειωθεί ότι ένας προκαθορισμένος αριθμός δεικτών των γραμμών, έστω S, του πίνακα δεδομένων εκπαίδευσης, $\boldsymbol{F} = [\boldsymbol{F}^T; \boldsymbol{F}^V]$, χρησιμοποιείται για αξιολόγηση (validation) κατά το στάδιο εκπαίδευσης. Κατά συνέπεια προκύπτει μία λίστα με τις ακμές του γράφου $\boldsymbol{E}^T \in \mathbb{R}^{l_{E,T}}$, $\boldsymbol{E}^V \in \mathbb{R}^{l_{E,V}}$, όπου $l_{E,T}$ και $l_{E,V}$ που δηλώνουν το συνολικό αριθμό ακμών των Γράφων που αναφέρονται στα σύνολα δεδομένων εκπαίδευσης και αξιολόγησης αντίστοιχα, καθώς και οι αντίστοιχοι πίνακες γειτνίασης $\boldsymbol{A}^T \in \mathbb{R}^{(N-S)\times(N-S)}$ και $\boldsymbol{A}^V \in \mathbb{R}^{S\times S}$.

Κατά το δεύτερο στάδιο, τα μεγέθη αυτά χρησιμοποιούνται για την εκπαίδευση ενός νευρωνικού δικτύου γράφων, με βάση το μηχανισμό του συστήματος εξισώσεων (2.17). Ο αλγόριθμος εκμάθησης που περιγράφεται στον αλγόριθμο **Αλγόριθμος 1** (βλ. ΠΑΡΑΡΤΗΜΑ Ι), βασίζεται σε μια στρατηγική καθόδου βασισμένης στην κλίση και περιλαμβάνει τα ακόλουθα βήματα:

1) Οι καταστάσεις των κόμβων του γράφου εκπαίδευσης $\boldsymbol{x}_T(t) \in \mathbb{R}^{N\times(N-S)}$ ανανεώνονται επαναληπτικά με τον αλγόριθμο **Αλγόριθμος 2** (βλ. ΠΑΡΑΡΤΗΜΑ Ι) μέχρις ότου το σύστημα να φτάσει σε ένα σημείο ισορροπίας $\boldsymbol{x}(t^*)$, όπου t^* είναι η χρονική στιγμή κατά την οποία επιτυγχάνεται το σημείο αυτό. Πιο συγκεκριμένα, η διαδικασία αυτή επαναλαμβάνεται μέχρι η ποσότητα $|\,\boldsymbol{x}(t) - \boldsymbol{x}(t-1)|$ να είναι μικρότερη από ένα προκαθορισμένο κατώφλι ε_f. Τα στοιχεία που απαιτούνται για τον προσδιορισμό της κατάστασης ενός κόμβου $\boldsymbol{x}_n, n \in [1, \dots, N-S]$ καθορίζονται από τις εξής εκφράσεις:
$$\begin{cases} l_n = F^T_{n,2} \\ \boldsymbol{l}_{co[n]} = \left[F^T_{n,3}; \dots; F^T_{n,D}\right] * \left[A^T_{n,1}; A^T_{n,2}; \dots; A^T_{n,N-S}\right] \\ \boldsymbol{l}_{ne[n]} = \left[F^T_{1,2}; \dots; F^T_{N-S,2}\right] * \left[A^T_{n,1}; A^T_{n,2}; \dots; A^T_{n,N-S}\right] \\ \boldsymbol{x}_{ne[n]} = \left[\boldsymbol{x}_1; \dots; \boldsymbol{x}_{N-S}\right] * \left[A^T_{n,1}; A^T_{n,2}; \dots; A^T_{n,N-S}\right] \end{cases} \quad (2.18)$$
Όπου οι δείκτες $\cdot_{i,j}$ δηλώνουν την i-στή γραμμή και την j-οστή στήλη του εκάστοτε πίνακα.

2) Η κλίση $\frac{\partial \mathbf{e}}{\partial \mathbf{W}}$ υπολογίζεται σύμφωνα με τον αλγόριθμο **Αλγόριθμος 3** (βλ. ΠΑΡΑΡΤΗΜΑ Ι). Ειδικότερα, η διαδικασία οπισθοδιάδοσης επαναλαμβάνεται μέχρι να επιτευχθεί ο μέγιστος αριθμός βημάτων οπισθοδιάδοσης S_b.
3) Τα βάρη $\boldsymbol{W}$ ανανεώνονται χρησιμοποιώντας τη μέθοδο της ανθεκτικής οπισθοδιάδοσης (resilient backpropagation – RPROP) [99], που περιγράφεται στον αλγόριθμο **Αλγόριθμος 4** (βλ. ΠΑΡΑΡΤΗΜΑ Ι), έως ότου να επιτευχθεί ο μέγιστος αριθμός εποχών εκμάθησης S_l. Οι εσωτερικές παράμετροι του αλγορίθμου, δηλαδή ο παράγοντας ελάττωσης η^-, ο παράγοντας αύξησης η^+, η σταθερά μέγιστης ανανέωσης Δ_{max} και η σταθερά ελάχιστης ανανέωσης Δ_{min}, προσδιορίζονται από τον σχεδιαστή.
4) Το σφάλμα αξιολόγησης υπολογίζεται για κάθε S_v αριθμό εποχών και αν είναι ελάχιστο, τα τρέχοντα βάρη αποθηκεύονται.

Δοθέντος ενός πίνακα που περιέχει τα δεδομένα ελέγχου $\boldsymbol{F}^E \in \mathbb{R}^{Q\times D}$, όπου Q είναι το πλήθος των παρατηρήσεων του συνόλου ελέγχου, ο αντίστοιχος πίνακας γειτνίασης $\boldsymbol{A}^E \in \mathbb{R}^{Q\times Q}$ προκύπτει μετατρέποντας την χρονοσειρά δεδομένων ελέγχου, $\left[F^E_{1,2}; \dots; F^E_{Q,2}\right]$, η οποία αποτελεί τη 2η στήλη του πίνακα $\boldsymbol{F}^E$ σε έναν γράφο φυσικής ορατότητας. Αφού έχει ολοκληρωθεί η διαδικασία εκπαίδευσης, εφαρμόζοντας τον αλγόριθμο **Αλγόριθμος 5** (βλ. ΠΑΡΑΡΤΗΜΑ Ι), υπολογίζονται οι προβλέψεις.

2.3.2 Μη επιβλεπόμενη μάθηση

Η μη επιβλεπόμενη ή μη επιτηρούμενη μάθηση διαφέρει θεμελιωδώς από την επιβλεπόμενη καθώς μπορεί να διαχειριστεί μη επισημασμένα (unlabeled) δεδομένα, δηλαδή δεδομένα από τα οποία απουσιάζουν οι τιμές της εξαρτημένης μεταβλητής ή μεταβλητής εξόδου. Οι αλγόριθμοι που ακολουθούν αυτή την προσέγγιση, χρησιμοποιούν το σύνολο των διαθέσιμων δεδομένων. Ως στόχο έχουν την αναγνώριση μοτίβων ομοιότητας, και την συνακόλουθη ομαδοποίηση των δεδομένων βάσει των μοτίβων αυτών. Παρόλο που οι μέθοδοι μη επιβλεπόμενης μάθησης ενδείκνυνται κατά κύριο λόγο για εφαρμογές ομαδοποίησης (clustering) [226], εν τούτοις είναι δυνατόν, να αξιοποιηθεί η πληροφορία που αποκτάται σχετικά με το βαθμό ομοιότητας των δεδομένων για την ανάπτυξη ανταγωνιστικών μοντέλων πρόβλεψης χαρακτηριστικών από χρονοσειρές [227], [228]. Μεταξύ των αλγορίθμων μάθησης χωρίς επίβλεψη που ξεχωρίζουν για τη δημοφιλία και την επίδοσή τους, είναι η μέθοδος των Κ-πλησιέστερων γειτόνων (k-nearest neighbor - KNN), η ανάλυση σε κύριες συνιστώσες (Principal component analysis - PCA), και τα πιο πρόσφατα δίκτυα βαθιάς πεποίθησης (deep belief networks - DBN) καθώς και η αντιπαραθετική (adversarial) μηχανική μάθηση [229], [230].

2.3.2.1 Αραιή αναπαράσταση

Η αραιή κωδικοποίηση είναι μια κατηγορία μεθόδων χωρίς επίβλεψη για την εκμάθηση συνόλων από υπερπλήρεις βάσεις για την αποτελεσματική αναπαράσταση δεδομένων. Ο στόχος της αραιής κωδικοποίησης είναι να βρεθεί ένα σύνολο k διανυσμάτων βάσης φ_i έτσι ώστε να γίνει δυνατή η αναπαράσταση ενός οποιουδήποτε διανύσματος εισόδου $\boldsymbol{x}$ ως ένας γραμμικός συνδυασμός αυτών των διανυσμάτων βάσης:

$$\boldsymbol{x} = \sum_{i=1}^{k} a_i \varphi_i \tag{2.19}$$

Η αραιότητα, που υλοποιείται με τις αντίστοιχες μεθόδους αραιής αναπαράστασης, είναι μια έννοια (συνοδευόμενη από τα σχετικά ευρέως διαδεδομένα μοντέλα μηχανικής μάθησης), σκοπός της οποίας είναι η ανακατασκευή με αποτελεσματικό τρόπο, ενός ή ενός συνόλου σημάτων, από ένα γραμμικό συνδυασμό, λίγων το πλήθος, μη μηδενικών συντελεστών μιας κατάλληλης βάσης (ή λεξικού). Αυτή η βάση είναι συνήθως εξοπλισμένη με την ιδιότητα της υπερπληρότητας (overcomplete), η οποία δηλώνει ότι ο αριθμός των στοιχείων του λεξικού (ή ατόμων) πρέπει να είναι μεγαλύτερος από τη διάσταση του διανύσματος εισόδου. Συνεπώς, οι αραιές μέθοδοι κωδικοποιούν οποιαδήποτε παρατήρηση μιας χρονοσειράς, παρέχοντας μια λύση που ενεργοποιεί όσο το δυνατόν λιγότερα στοιχεία του λεξικού. Κατά τη διάρκεια των προηγούμενων ετών, μια σειρά από σημαντικές ερευνητικές εργασίες παρέχουν στοιχεία ότι η αραιή αναπαράσταση είναι μια διαδεδομένη μέθοδος για επιστημονικούς κλάδους όπως η επεξεργασία σήματος και εικόνας, η υπολογιστική όραση και η μηχανική μάθηση, η θεωρία πληροφοριών, η νευροεπιστήμη και συναφείς τομείς [231]–[235]. Έχει επίσης υποστηριχθεί ότι το τμήμα V1 του εγκεφάλου λειτουργεί με παρόμοιο τρόπο υπό τους αντικειμενικούς περιορισμούς αραιότητας [236].

Στο σημείο αυτό θα πρέπει να γίνει η επεξήγηση κάποιων απαιτούμενων όρων ορολογίας, προς διευκόλυνση της κατανόησης της μεθόδου από τον ενδιαφερόμενο αναγνώστη. Τα μονοδιάστατα διανύσματα με μηδενικά ή μοναδιαία στοιχεία αναπαρίστανται με $\mathbf{0}_d \in R^d$ και $\mathbf{1}_d \in R^d$, αντίστοιχα. Έστω επίσης ℓ_0 και ℓ_q (για $q \geq 1$) οι αντίστοιχες νόρμες του διανύσματος στήλης: $\ell_0(\boldsymbol{x}) =,\ \|\boldsymbol{x}\|_0 \triangleq \#\{j$ υ.π.τ.$[\boldsymbol{x}]_j \neq 0\}$, $\|\boldsymbol{x}\|_q \triangleq \left(\sum_{i=1}^{m}|[\boldsymbol{x}]_i|^q\right)^{\frac{1}{q}}$. Τέλος, η νόρμα Frobenius ενός πίνακα δηλώνεται ως $\|\boldsymbol{X}\|_F \triangleq \left(\sum_{i=1}^{m}\sum_{j=1}^{n}\left|([\boldsymbol{X}]_{i,j})^2\right|\right)^{\frac{1}{2}} = \sqrt{Tr(\boldsymbol{X}^T\boldsymbol{X})}$ με $Tr(\cdot)$ τον τελεστή ίχνους και T τον τελεστή αναστροφής.

Έχοντας εφοδιαστεί με την ανωτέρω ορολογία και ορίζοντας τα απαραίτητα μεγέθη, ακολουθεί η παρουσίαση της προτεινόμενη προσέγγισης αραιής κωδικοποίησης για την βραχυπρόθεσμη πρόβλεψη του ηλεκτρικού φορτίου. Έστω $\boldsymbol{x}^{(k)} \in \mathbb{R}^d$ ένα k-διάστατο διάνυσμα του οποίου τη τελευταία (πιο πρόσφατη) τιμή $[\boldsymbol{x}^{(k)}]_d$ αντιστοιχεί στην τιμή της ζήτησης της ηλεκτρικής ενέργειας $EL(k) = [\boldsymbol{x}^{(k)}]_d$ οποιασδήποτε ζητούμενης χρονικής στιγμής k, ενώ το $[\boldsymbol{x}^{(k)}]_i$, $i \in [1{:}\,d-1]$ δηλώνει κάθε προηγούμενη (δηλ. ιστορική) τιμή του ηλεκτρικού φορτίου. Στη συνέχεια ορίζεται ένα σύνολο ακολουθιών (δηλ. μία ομαδοποίηση δεδομένων σε μορφή πίνακα) διατάσσοντας σειριακά ομάδες από N-διάστατες ακολουθίες $\boldsymbol{x}^{(k)}$, $\boldsymbol{X} = [\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, ..\mathbf{x}^{(N)}]$: $\boldsymbol{X} \in \mathbb{R}^{d\times N}$.

Ιδανικά, εάν ο χώρος των d-διαστάσεων είναι εφοδιασμένος με ένα υπερπλήρες λεξικό $\boldsymbol{D} = [\boldsymbol{d}^{(1)}, \boldsymbol{d}^{(2)}, \dots, \boldsymbol{d}^{(K)}] \in \mathbb{R}^{d\times K}$ (με $K > d$), που προκύπτει από το $\boldsymbol{X}$ και μια ακολουθία ενδιαφέροντος $\boldsymbol{q} \in \mathbb{R}^d$, η μέθοδος της αραιής αναπαράστασης μορφοποιείται από τις ακόλουθες ισοδύναμες μεταξύ τους μορφές είτε κανονικοποιημένου είτε υπό περιορισμούς προβλήματος βελτιστοποίησης, όπως εκφράζονται από τις (2.20) και (2.21), αντίστοιχα:

$$\min_{\boldsymbol{\alpha}}\left(\frac{1}{2}\|\boldsymbol{q} - \boldsymbol{D\alpha}\|_2^2 + \lambda\psi(\boldsymbol{\alpha})\right) \tag{2.20}$$

$$\min_{\boldsymbol{\alpha}}\left(\frac{1}{2}\|\boldsymbol{q} - \boldsymbol{D\alpha}\|_2^2\right), \text{υ.τ.π: } \psi(\boldsymbol{\alpha}) \leq \rho \tag{2.21}$$

Παρακάτω παρατίθενται μερικές λεπτομέρειες για τους παραπάνω όρους. Το διάνυσμα συντελεστών $\boldsymbol{\alpha} \in \mathbb{R}^K$ των (2.20) και (2.21), αναπαριστά τους αραιά κατανεμημένα συντελεστές του σήματος εισόδου ενδιαφέροντος $\boldsymbol{q} \in \mathbb{R}^d$. $\|\boldsymbol{q} - \boldsymbol{D\alpha}\|_2^2$ είναι η αναπαράσταση του σφάλματος ανακατασκευής ανάμεσα στο σήμα εισόδου ενδιαφέροντος $\boldsymbol{q}$ και της αραιής του αναπαράστασης $\boldsymbol{D\alpha}$. Η παράμετρος λάμδα (λ) δηλώνει την παράμετρο κανονικοποίησης (ή τον πολλαπλασιαστή Lagrange). Τέλος, η συνάρτηση ποινής $\psi(\cdot)$ δηλώνει τον όρο επιβαλλόμενης αραιότητας των συντελεστών, η πυκνότητα της οποίας ρυθμίζεται από την παράμετρο λ. Πιο συγκεκριμένα, όταν $\lambda = 0$ προκύπτουν πυκνές λύσεις των συντελεστών, ενώ όσο το λ αυξάνεται τόσο υπερισχύει ο όρος της αραιότητας. Στη σχέση (2.21) η παράμετρος ρ χρησιμοποιείται για να αναπαραστήσει ένα μέτρο του επιπέδου αραιότητας (δηλ. του αριθμού των μη μηδενικών συντελεστών) του διανύσματος αναπαράστασης $\boldsymbol{\alpha}$.

Τυπικά, η συνάρτηση $\psi(\cdot)$ ενσωματώνει τη νόρμα ℓ_p, ορισμένη για $1 \leq p \leq \infty$, του διανύσματος συντελεστών $\boldsymbol{\alpha}$, δηλ. $\psi(\boldsymbol{\alpha}) = \ell_p(\boldsymbol{\alpha}) = \left(\sum_{j=1}^{K}([\boldsymbol{a}]_j)^p\right)^{\frac{1}{p}}$ για μια συγκεκριμένη τιμή του p. Μερικές από τις πιο δημοφιλείς μορφές που συναντώνται στην βιβλιογραφία είναι αυτές που ενσωματώνουν τις νόρμες ℓ_0 και ℓ_1, αντιστοιχα. Στα πλαίσια της παρούσας εργασίας, προτείνεται, επιπρόσθετα στις προαναφερθείσες τυπικές μορφές, η χρήση άλλων νορμών, συμπεριλαμβανομένων εκείνων οι οποίες επιλέγουν άτομα που βρίσκονται στον κυρτό φλοιό του χώρου των χαρακτηριστικών, συνθέτοντας την αρχετυπική ανάλυση, καθώς και μία νόρμα κανονικοποίησης με δομή ιεραρχικού δέντρου [237], η οποία έχει αποδειχθεί χρήσιμη σε μια πληθώρα περιπτώσεων διερεύνησης της συσχέτισης μεταξύ παρελθοντικών παρατηρήσεων της χρονοσειράς του ηλεκτρικού φορτίου. Παρόλο που οι προαναφερθείσες νόρμες έχουν περιγραφεί αναλυτικά στην βιβλιογραφία, για λόγους πληρότητας θα παρατεθούν στη συνέχεια κάποιες σημαντικές λεπτομέρειες σχετικά με την εφαρμογή τους.

Όπως ορίστηκε νωρίτερα, η αραιή αναπαράσταση χρησιμοποιεί ένα υπερπλήρες σύνολο από ακολουθίες (ή άτομα) ενός λεξικού $\mathbf{D}$. Ο σχηματισμός του λεξικού επιτυγχάνεται μέσα από μία συγκεκριμένη διαδικασία εκμάθησης η οποία προσπαθεί να προσαρμόσει τα άτομα του λεξικού στα δεδομένα εισόδου/εκπαίδευσης με τη μορφή γραμμικών συνδυασμών αυτών. Αυτές οι μέθοδοι μη επιβλεπόμενης μάθησης υπολογίζουν την συνάρτηση κόστους του σφάλματος ανακατασκευής με αξιοσημείωτα αποτελέσματα σε πολλούς τομείς μηχανικής μάθησης. Η εκμάθηση του λεξικού ορίζεται ως ένα από κοινού πρόβλημα βελτιστοποίησης ως προς το λεξικό $\boldsymbol{D} = \{\boldsymbol{d}^{(j)}\}, j = 1{:}K$ αλλά και τους συντελεστές $\boldsymbol{A} = \{\boldsymbol{\alpha}^{(k)}\}, k = 1{:}N$ και εκφράζεται συχνά ως εξής:

$$\min_{\boldsymbol{D}\in C, \boldsymbol{A}\in\mathbb{R}^{K\times N}} \left(\frac{1}{2}\|\boldsymbol{X}-\boldsymbol{D}\boldsymbol{A}\|_F^2 + \lambda\psi(\boldsymbol{A})\right) \tag{2.22}$$

$$\min_{\boldsymbol{D}\in C, \boldsymbol{A}\in R^{K\times N}} \left(\frac{1}{2}\|\boldsymbol{X}-\boldsymbol{D}\boldsymbol{A}\|_F^2\right) \text{ s.t: } \psi(\mathbf{A}) \le \rho \tag{2.23}$$

Όμοια με παραπάνω η συνάρτηση $\psi(\cdot)$ δηλώνει τον όρο της επιβαλλόμενης αραιότητας . Το σύνολο C των ατόμων του λεξικού ορίζεται συχνά ως το κυρτό σύνολο των πινάκων που ικανοποιεί τον ακόλουθο περιορισμό:

$$C \triangleq \{\boldsymbol{D} \in \mathbb{R}^{d\times K} \text{ υ.τ.π.: } \forall_{j=1:K}, (\boldsymbol{d}^{(j)})^T\boldsymbol{d}^{(j)} \le 1\} \tag{2.24}$$

Η εξίσωση (2.24) επιτρέπει τη χρήση κανονικοποιημένων ατόμων και όχι αυθαίρετων προκειμένου να αποφευχθούν οι περιπτώσεις υπολογισμού μικρών τιμών συντελεστών.

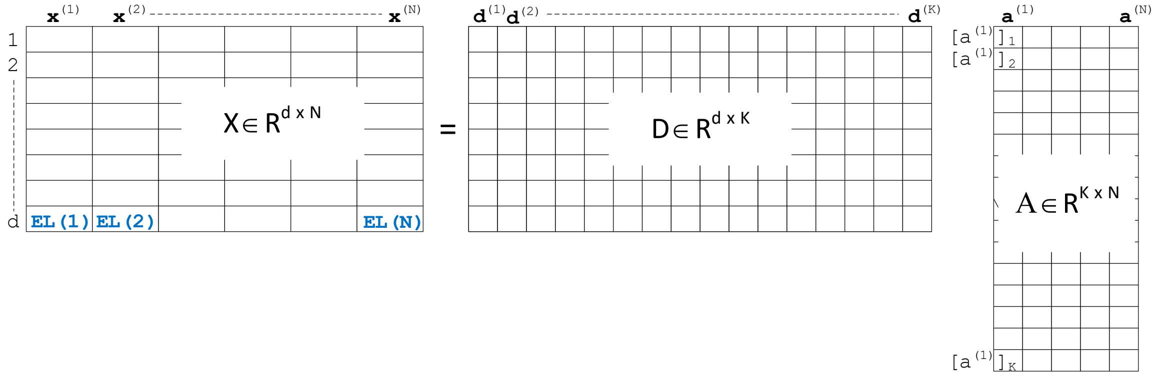
Η συνεισφορά της παρούσας έρευνας έγκειται επίσης στη διερεύνηση του σταδίου εκμάθησης του λεξικού, πέραν της τυπικής μορφοποίησης των νορμών $\psi(\cdot)$ και του περιορισμού C του λεξικού $\mathbf{D}$, α) τοποθετώντας τα άτομα του λεξικού σε ένα κατευθυνόμενο δέντρο και β) παρέχοντας μια λύση εμπνευσμένη από την αρχετυπική ανάλυση στην οποία το λεξικό $\mathbf{D}$ των αρχετύπων υπόκειται σε δύο δυικούς περιορισμούς: Ο πρώτος, υποστηρίζει ότι κάθε διάνυσμα $\boldsymbol{x}^{(i)}$ πρέπει να προσεγγίζεται από ένα κυρτό συνδυασμό κάποιων αρχετύπων $\boldsymbol{d}^{(j)}$, ενώ ο δεύτερος, διασφαλίζει ότι κάθε j-οστό αρχέτυπο πρέπει να προσεγγίζεται από ένα κυρτό συνδυασμό των διανυσμάτων $\boldsymbol{x}^{(i)}$.

Συνοψίζοντας, η εκμάθηση του λεξικού $\boldsymbol{D} \in \mathbb{R}^{d\times K}$ πραγματοποιείται μέσω των σχέσεων (2.22) και (2.23) σε συνδυασμό με α) τους τυπικούς πριορισμούς του λεξικού της εξίσωσης (2.24) καθώς και τους επιπρόσθετους που αναφέρθηκαν παραπάνω και β) κατάλληλες μορφές για τις νόρμες $\psi(\cdot)$. Η είσοδος των αλγορίθμων που επιλύουν τις (2.22) και (2.23) είναι το σύνολο $\boldsymbol{X} \in \mathbb{R}^{d\times N}$ των d-διάστατων ακολουθιών, μήκους N. Προκειμένου να παραχθεί μια εκτίμηση (δηλ. μία πρόβλεψη) για μια τιμή του φορτίου τη χρονική στιγμή k, $EL(k)$, δημιουργείται μία ακολουθία ενδιαφέροντος εισόδου $\boldsymbol{q} \in \mathbb{R}^{d-1}$ μιας και η ζητούμενη τιμή $EL(k)$ είναι άγνωστη και πρόκειται να προβλεφθεί. Για το λόγο αυτό, η επίλυση της (2.20) ή (2.21) γίνεται με τη βοήθεια ενός περικομμένου λεξικού $\boldsymbol{D}_{tr} \in \mathbb{R}^{(d-1)\times K}$ το οποίο περιέχει τις πρώτες (d-1) γραμμές του αρχικού λεξικού $\boldsymbol{D} \in \mathbb{R}^{d\times K}$. Γνωρίζοντας ότι η λύση της (2.20) ή (2.21) παρέχει ένα k-διάστατο αραιό διάνυσμα συντελεστών $\boldsymbol{\alpha}_{tr} \in \mathbb{R}^K$, εξάγεται μία τιμή πρόβλεψης του φορτίου για την χρονική στιγμή k, $EL(k)$ εφαρμόζοντας τον ακόλουθο κανόνα ανακατασκευής:

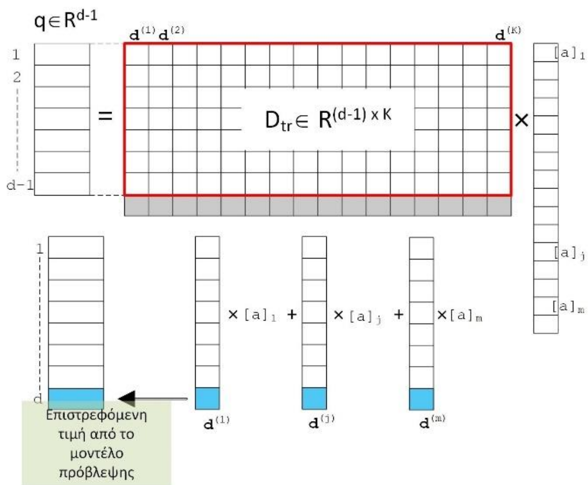
$$EL(k) = \boldsymbol{D}\alpha_{tr} \tag{2.25}$$

Θα πρέπει να αναφερθεί ότι για την εφαρμογή της (2.25), τα άτομα $\boldsymbol{d}_{tr}^{(j)}$ του περικομμένου λεξικού $\boldsymbol{D}_{tr}$ θα πρέπει να κανονικοποιηθούν εκ νέου ώστε να βρίσκονται στη μοναδιαία ℓ_2

σφαίρα. Το Σχήμα 7 παρέχει μια γραφική απεικόνιση της διαδικασίας εκμάθησης του λεξικού $\boldsymbol{D} \in \mathbb{R}^{d \times K}$, ενώ το Σχήμα 8 δείχνει την προτεινόμενη μέθοδο για την εξαγωγή προβλέψεων $EL(k)$ χρησιμοποιώντας ένα διάνυσμα $\boldsymbol{q} \in \mathbb{R}^{d-1}$ με παρελθοντικές τιμές της χρονοσειράς του φορτίου, το λεξικό $\boldsymbol{D}_{tr} \in \mathbb{R}^{(d-1) \times K}$, και το αντίστοιχο αραιό διάνυσμα συντελεστών $\boldsymbol{\alpha}_{tr} \in \mathbb{R}^{K}$.



Σχήμα 7: Γραφική απεικόνιση της διαδικασίας εκμάθησης του λεξικού D από ένα μητρώο X από N-διάστατες ακολουθίες ηλεκτρικού φορτίου.



Σχήμα 8: Προτεινόμενη μέθοδος για την πρόβλεψη του ηλεκτρικού φορτίου α) χρησιμοποιώντας ένα περικομμένο λεξικό D_{tr} για τον υπολογισμό των αραιών συντελεστών και β) εφαρμόζοντάς τους στο αρχικό λεξικό D.

Μια συλλογή από τις μεθόδους αραιής αναπαράστασης για την πρόβλεψη του ηλεκτρικού φορτίου που διερευνώνται στα πλαίσια της διατριβής παρουσιάζεται στον πίνακα Πίνακας 2. Συγκεκριμένα, εφαρμόζονται οι ακόλουθες συναρτήσεις κόστους για αραιή αναπαράσταση και εκμάθηση λεξικού: α) η νόρμα ℓ_0 που υλοποιείται με τους αλγορίθμους K-SVD και orthogonal matching pursuit (OMP), β) η νόρμα ℓ_1 που υλοποιείται με τους αλγορίθμους online και least-angle regression-least absolute shrinkage and selection operator (LARS-LASSO), γ) η νόρμα ℓ_2 που συνδέεται με περιορισμούς της αρχετυπικής ανάλυσης για τα άτομα και δ) μία νόρμα ℓ_0^{tree} όπου τα άτομα ενσωματώνονται σε μια δενδροειδή δομή $g(T)$. Παρακάτω περιγράφεται σύντομα η κάθε μία από αυτές ως εξής:

Πίνακας 2: Λίστα μεθόδων αραιής αναπαράστασης για την πρόβλεψη ηλεκτρικού φορτίου

Νόρμα	Τύποι συνάρτησης $\psi(\cdot)$, σύνολα λεξικών C και πρόσθετες ιδιότητες του $\boldsymbol{\alpha}$
ℓ_0	$\min_{\boldsymbol{D}\in C,\boldsymbol{A}\in\mathbb{R}^{K\times N}}(0.5\lVert\boldsymbol{X}-\boldsymbol{DA}\rVert_F^2)$, υ.τ.π.: $\lVert\boldsymbol{A}\rVert_0 \le \rho$ $C \triangleq \{\boldsymbol{D}\in\mathbb{R}^{d\times K}$ υ.τ.π.: $\forall_{j=1:K}, (\boldsymbol{d}^{(j)})^T\boldsymbol{d}^{(j)} = 1\}$
ℓ_1	$\min_{\boldsymbol{D}\in C,\boldsymbol{A}\in\mathbb{R}^{K\times N}}(0.5\lVert\boldsymbol{X}-\boldsymbol{DA}\rVert_F^2 + \lambda\lVert\boldsymbol{A}\rVert_1)$ $C \triangleq \{\boldsymbol{D}\in\mathbb{R}^{d\times K}$ υ.τ.π.: $\forall_{j=1:K}, (\boldsymbol{d}^{(j)})^T\boldsymbol{d}^{(j)} \le 1\}$
ℓ_2	$\min(\lVert\boldsymbol{X}-\boldsymbol{DA}\rVert_F^2 = \lVert\boldsymbol{X}-\boldsymbol{XBA}\rVert_F^2)$ υ.τ.π. $\boldsymbol{a}^{(i)}\in\Delta_K, \boldsymbol{b}^{(j)}\in\Delta_N$ $\Delta_K \triangleq \left\{\boldsymbol{a}^{(i)}\in\mathbb{R}^K \text{ υ.τ.π. } \boldsymbol{a}^{(i)}\ge 0 \; \kappa\alpha\iota \sum_{j=1}^{K}[\boldsymbol{a}^{(i)}]_j = 1\right\}$ $\Delta_N \triangleq \left\{\boldsymbol{b}^{(j)}\in\mathbb{R}^N \text{ υ.τ.π. } \mathbf{b}^{(j)}\ge 0 \; \kappa\alpha\iota \sum_{i=1}^{N}[\boldsymbol{b}^{(j)}]_i = 1\right\} \forall i\in[1:N], \& \forall j \in [1:K]$
ℓ_0^{tree}	$\min_{\boldsymbol{D}\in C,\boldsymbol{A}\in\mathbb{R}^{K\times N}}(0.5\lVert\boldsymbol{X}-\boldsymbol{DA}\rVert_F^2 + \lambda\ell_o^{tree}(\boldsymbol{A}))$ $C \triangleq \{\boldsymbol{D}\in\mathbb{R}^{d\times K}$ υ.τ.π.: $\forall_{j=1:K}, (\boldsymbol{d}^{(j)})^T\boldsymbol{d}^{(j)} \le 1\}$

1. Νόρμα ℓ_0:K-SVD/OMP. Είναι γνωστό ότι η εφαρμογή της νόρμας ℓ_0 οδηγεί σε ένα συνδυαστικό NP-σκληρό πρόβλημα βελτιστοποίηση, η λύση του οποίου μπορεί να επιτευχθεί μόνο μέσω προσεγγιστικών μεθόδων. Για το σκοπό αυτό, εφαρμόζονται άπληστοι αλγόριθμοι (greedy algorithms) για την εύρεση του διανύσματος βέλτιστης λύσης. Η στρατηγική αυτή έγκειται στην αναζήτηση του ατόμου που αντιστοιχεί στο δείγμα εισόδου με την ισχυρότερη συσχέτιση με την έννοια των ελαχίστων τετραγώνων. Στην περίπτωση αυτή, χρησιμοποιείται η δημοφιλής μέθοδος OMP [238]. Δοθέντος ενός λεξικού $\mathbf{D}$ και ενός τυχαίου δείγματος $\boldsymbol{x}^{(i)}$, η μέθοδος OMP επιλέγει διαδοχικά τα άτομα με την υψηλότερη συσχέτιση με το αντίστοιχο υπόλοιπο του δείγματος. Σε κάποιο βήμα s: $0 < s \le \rho$, με ρ το επίπεδο αραιότητας, ο αλγόριθμος επιλογής του καταλληλότερου ατόμου εκφράζεται ως $k_s = argmax_j\left|(\boldsymbol{d}^{(j)})^T\boldsymbol{r}_{s-1}\right|$, όπου $\boldsymbol{r}_{s-1}$ είναι το τρέχον υπόλοιπο. Έπειτα από την επιλογή ενός ατόμου, το σήμα $\boldsymbol{x}^{(i)}$ προβάλλεται στο εύρος των επί του παρόντος επιλεγμένων ατόμων ως: $\widehat{\boldsymbol{a}}_s = \left(\boldsymbol{D}_{V_s}\right)^{+}\boldsymbol{x}^{(i)}$, όπου $V_s = V_{s-1} \cup k_s$ αναπαριστά το σύνολο των δεικτών που δηλώνουν τα τρέχοντα επιλεγόμενα άτομα του λεξικού και $\boldsymbol{D}_{V_s}$ είναι το υποσύνολο του λεξικού με τα στοιχεία που καθορίζει το σύνολο των δεικτών V_s. Η τιμή του νέου υπολοίπου είναι $\boldsymbol{r}_s = \boldsymbol{x}^{(i)} - \boldsymbol{D}_{V_s}\widehat{\boldsymbol{a}}_s$ και η διαδικασία επαναλαμβάνεται μέχρι είτε να επιλεχθούν ρ άτομα, είτε να ελαχιστοποιηθεί η τιμή

του υπολοίπου. Στα πλαίσια της παρούσας έρευνας επιλέγεται ο αλγόριθμος batch-OMP [239], με την αποδόμηση Cholesky. Η εκμάθηση του λεξικού, κανονικοποιημένου κατά τη νόρμα ℓ_0 επιτυγχάνεται με τον αλγόριθμο KSVD/OMP. Δεδομένου ενός συνόλου δεδομένων εκπαίδευσης X και του αντίστοιχου συνόλου αραιών συντελεστών $A^{(t)}$ υπολογισμένων με τον αλγόριθμο OMP για το λεξικό $\boldsymbol{D}^{(t)} = \left[\boldsymbol{d}_1^{(t)}, \boldsymbol{d}_2^{(t)}, \dots, \boldsymbol{d}_K^{(t)}\right]$ κατά το βήμα επανάληψης t, ο στόχος είναι η απόκτηση του ανανεωμένου λεξικού $D^{(t+1)}$:

$$\min_{\boldsymbol{D}^{(t+1)}} \left\{ \left\| \boldsymbol{X} - \boldsymbol{D}^{(t+1)} \boldsymbol{A}^{(t)} \right\|_F^2 \right\} \quad \upsilon.\tau.\pi. \quad \left\| \boldsymbol{a}^{(i),(t)} \right\|_0 \leq \rho \tag{2.26}$$

Έστω ω_k το σύνολο των δεικτών που αφορούν τα δείγματα εκπαίδευσης τα οποία χρησιμοποιούν το άτομο $d^{(k),(t)}$: $\omega_k = \{i | 1 \leq i \leq N, [\boldsymbol{A}]_{k,i} \neq 0\}$, με $[\boldsymbol{A}]_{(k,:)}$ να δηλώνει την k-οστή γραμμή του πίνακα συντελεστών $\boldsymbol{A} \in \mathbb{R}^{d \times N}$. Έστω, επίσης $\Omega_k \in R^{K \times |\omega_k|}$ ένας πίνακας με μονάδες όπου $(\omega_k(i), i)$ μηδενικά αλλού. Ορίζοντας τον πίνακα σφαλμάτων ως $\boldsymbol{E}_k = \boldsymbol{X} - \sum_{j \neq k} \boldsymbol{d}^{(j),(t)}$, ο πίνακας $[\boldsymbol{A}]_{j,:}$ εκφράζει την πληροφορία που δεν μπορεί να εξηγηθεί από την αναπαράσταση αυτή. Επομένως, το k-οστό άτομο μπορεί να βελτιστοποιηθεί έτσι ώστε να αναπαριστά ακριβέστερα την πληροφορία αυτή. Συνοψίζοντας, ο αλγόριθμος K-SVD ανανεώνει το k-οστό άτομο εφαρμόζοντας παραγοντοποίηση ιδιαζουσών τιμών (singular value decomposition - SVD) στον πίνακα $\boldsymbol{E}_k^{\omega}$, όπου $\boldsymbol{E}_k^{\omega} = \boldsymbol{E}_k \Omega_k$, αναζητώντας τον κοντινότερο πίνακα βαθμού 1 έτσι ώστε $\boldsymbol{E}_k^{\omega} = \boldsymbol{U\Delta V}^T$. Η ανανεωμένη τιμή $\boldsymbol{d}_k^{(t+1)}$του k-οστού ατόμου ορίζεται ως η πρώτη στήλη του $\boldsymbol{U}$, και το νέο διάνυσμα συντελεστών $\boldsymbol{a}_{\omega}^{(k)}$ ως η πρώτη στήλη του $\boldsymbol{V}$ πολλαπλασιασμένη με $[\boldsymbol{\Delta}]_{1,1}$. Η διαδικασία αυτή επαναλαμβάνεται για κάθε άτομο, δημιουργώντας ένα ανανεωμένο λεξικό, το οποίο χρησιμοποιείται από τον αλγόριθμο OMP ώστε να προκύψουν οι ανανεωμένοι αραιοί συντελεστές $\boldsymbol{A}^{(t+1)}$.

2. Νόρμα ℓ_1: ONLINE/LARS-LASSO. Η νόρμα ℓ_1, χρησιμοποιείται ευρέως για την επίλυση προβλήματα αραιής αναπαράστασης δεδομένου ότι είναι σε θέση να παρέχει μια αναλυτική λύση σε πολυωνυμικό χρόνο. Παραδείγματα διαδεδομένων αλγορίθμων είναι η αναζήτηση βάσης (basis pursuit) καθώς και ο LASSO [240]. Σε αντίθεση με τις μεθόδους συντεταγμένης καθόδου, οι οποίες υποθέτουν ότι τα άτομα του λεξικού παρουσιάζουν χαμηλή συσχέτιση, εδώ εφαρμόζονται μέθοδοι που βασίζονται σε ομοτοπία (homotopy) όπως ο αλγόριθμος LARS-LASSO [241]. Μία ομοτοπία ανάμεσα σε δύο συνεχείς συναρτήσεις f και g από έναν τοπολογικό χώρο X σε έναν τοπολογικό χώρο Y, ορίζεται ως μια συνεχής συνάρτηση $H: X \times [0,1] \rightarrow Y$, δηλαδή μια απεικόνιση από το χώρο που ορίζεται από το γινόμενου του X με το μοναδιαίο διάστημα $[0,1]$ στον χώρο Y, έτσι ώστε $H(x,0) = f(x)$ και $H(x,1) = g(x)$ για κάθε x $\in X$. Ο αλγόριθμος LARS-LASSO υπολογίζει την αλληλουχία των λύσεων μειώνοντας επαναληπτικά την τιμή του λ και χρησιμοποιώντας την προηγούμενη λύση σαν θερμική επανεκκίνηση. Έχει αναφερθεί ότι η μοναδικότητα της λύσης για μια συγκεκριμένη τιμή της παραμέτρου λ, είναι εγγυημένη και κατ' επέκταση μπορεί να αποδειχθεί ότι η αλληλουχία των λύσεων είναι τμηματικά γραμμική [242]. Η ιδιότητα αυτή είναι πολύ σημαντική καθώς ο αλγόριθμος ακολουθεί την κατεύθυνση κάθε λύσης εωσότου να φτάσει ένα κρίσιμο σημείο, όπου είτε ένα μη μηδενικό στοιχείο γίνεται μηδενικό (κι έτσι αφαιρείται από το ενεργό σύνολο συντελεστών), είτε ένα νέο μη μηδενικό στοιχείο προστίθεται στο ενεργό σύνολο συντελεστών. Ως αποτέλεσμα, η μέθοδος ομοτοπίας αρχικοποιείται με ένα κενό σύνολο συντελεστών και ανανεώνει επαναληπτικά κατά ένα στοιχείο κάθε φορά. Η πολυπλοκότητα της μεθόδου βασίζεται στην αντιστροφή του πίνακα συνδιακύμανσης των επιλεγμένων ατόμων σε κάθε οριακό σημείο προκειμένου να ενημερωθεί το ενεργό σύνολο

συντελεστών, η οποία εκτελείται από την αποδόμηση Cholesky ή τον τύπο Woodbury. Στην παρούσα προσέγγιση υιοθετείται η μέθοδος online μάθησης σε συνδυασμό με τον αλγόριθμο LARS-LASSO για την εκμάθηση του λεξικού.

3. Νόρμα ℓ_2: Αρχετυπική ανάλυση. Η αρχετυπική ανάλυση είναι μια καινοτόμος μέθοδος μάθησης χωρίς επίβλεψη που σχετίζεται με μεθόδους ανάλυσης γενετικών δεδομένων όπως η αραιή κωδικοποίηση [243]. Τα αρχέτυπα αποτελούν ειδική περίπτωση στοιχείων λεξικού, λόγω του ότι τοποθετούνται κυρίως στον κυρτό φλοιό του χώρου των χαρακτηριστικών. Η αρχετυπική ανάλυση έχει να κάνει με την εκμάθηση μιας παραγοντικής αναπαράστασης του $\boldsymbol{X}$ μέσω της επίλυσης ενός αντίστοιχου αρχετυπικού προβλήματος. Αναζητά ένα σύνολο αρχετύπων $\boldsymbol{D}$ όπου κάθε διάνυσμα $\boldsymbol{d}^{(j)}$ ανήκει επίσης στο χώρο των χαρακτηριστικών, αλλά ο σχηματισμός τους περιορίζεται από δύο δυϊκούς γεωμετρικούς περιορισμούς. Ο πρώτος δηλώνει ότι οποιαδήποτε ακολουθία $\boldsymbol{x}^{(i)}$ πρέπει να προσεγγιστεί επαρκώς από έναν κυρτό συνδυασμό μερικών αρχετύπων $\boldsymbol{d}^{(j)}$, ενώ ο δεύτερος δηλώνει ότι κάθε αρχέτυπο $\boldsymbol{d}^{(j)}$ πρέπει επίσης να προσεγγιστεί με έναν κυρτό συνδυασμό του διανύσματος ακολουθίας $\boldsymbol{x}^{(i)}$. Έτσι, δεδομένου ενός συνόλου αρχετύπων $\mathbf{D}$, κάθε ακολουθία $\boldsymbol{x}^{(i)}$ θα πρέπει να προσεγγίζεται ως $\boldsymbol{x}_{apprx}^{(i)} = \boldsymbol{D}\boldsymbol{a}^{(i)}$, όπου $\boldsymbol{a}^{(i)} \in \mathbb{R}^K$, είναι ένα διάνυσμα στήλης με τους συντελεστές που βρίσκεται στον κυρτό φλοιό Δ_K [244]:

$$\Delta_K \triangleq \left\{ \boldsymbol{a}^{(i)} \in \mathbb{R}^K \ \upsilon.\tau.\pi.\ \boldsymbol{a}^{(i)} \geq 0 \ \kappa\alpha\iota \ \sum_{j=1}^{K} [\boldsymbol{a}^{(i)}]_j = 1 \right\} \tag{2.27}$$

Και κάθε αρχέτυπο $\boldsymbol{d}^{(j)}$ πρέπει να προσεγγιστεί από το γινόμενο $\boldsymbol{X} \cdot \boldsymbol{b}^{(j)}$, με $\boldsymbol{b}^{(j)} \in \mathbb{R}^N$ να δηλώνει ένα άλλο διάνυσμα στήλη με συντελεστές το οποίο βρίσκεται στον κυρτό φλοιό Δ_N:

$$\Delta_N \triangleq \left\{ \boldsymbol{b}^{(j)} \in \mathbb{R}^N \ \upsilon.\tau.\pi.\ \mathbf{b}^{(j)} \geq 0 \ \kappa\alpha\iota \ \sum_{i=1}^{N} [\boldsymbol{b}^{(j)}]_i = 1 \right\} \tag{2.28}$$

Τότε το παραπάνω πρόβλημα εκφράζεται με την ελαχιστοποίηση του υπολειπόμενου αθροίσματος τετραγώνων (Residual Sum of Squares - RSS) ως εξής:

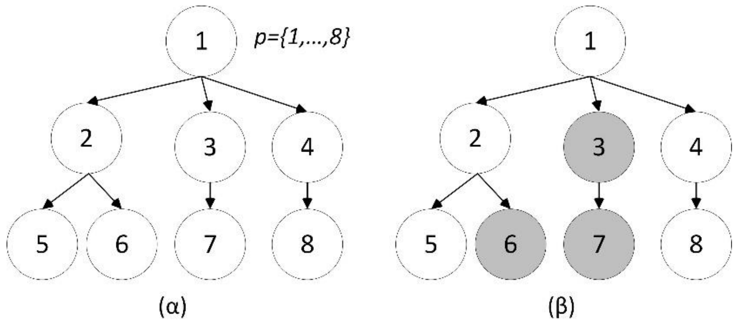
$$\min_{\boldsymbol{a}_i \in \Delta_p \ \text{για}\ i=1:N} \sum_{i=1}^{N} \left\| \boldsymbol{x}^{(i)} - \boldsymbol{D}\boldsymbol{a}^{(i)} \right\|_2^2 \tag{2.29}$$

όπου $\forall j, \boldsymbol{d}^{(j)} = \boldsymbol{X} \cdot \boldsymbol{b}^{(j)}$, το οποίο ισοδυναμεί με το:

$$\min_{\substack{\boldsymbol{a}^{(i)} \in \Delta_K \ \text{για}\ i=1:N \\ \boldsymbol{b}^{(j)} \in \Delta_N \ \text{για}\ \mathrm{j}=1:K}} \|\boldsymbol{X} - \boldsymbol{X} \cdot \boldsymbol{B} \cdot \boldsymbol{A}\|_F^2 \tag{2.30}$$

με $\boldsymbol{A} \in \mathbb{R}^{K \times N}$, $\mathbf{B} \in \mathbb{R}^{N \times K}$ και $\boldsymbol{D} = \boldsymbol{X} \cdot \boldsymbol{B}$. Η σχέση (2.29) αποτελεί ένα μη κυρτό πρόβλημα βελτιστοποίησης, ωστόσο είναι κυρτό ως προς μια από τις μεταβλητές $\mathbf{A}$ ή $\mathbf{B}$, όταν η άλλη μεταβλητή παραμένει σταθερή. Η παρατήρηση αυτή επιτρέπει την ανάπτυξη μια τεχνικής συντεταγμένης καθόδου που εγγυάται με ασυμπτωτικό τρόπο την ύπαρξη ενός στάσιμου σημείου του προβλήματος. Θα πρέπει να σημειωθεί ότι η κύρια διαφορά μεταξύ της αραιής κωδικοποίησης και της αρχετυπικής ανάλυσης, εκτός από τη μη αρνητικότητα του $\boldsymbol{a}^{(i)}$, είναι ότι τα αρχέτυπα πρέπει να είναι κυρτοί συνδυασμοί των δεδομένων $\mathbf{X}$, και με παρόμοιο τρόπο , το $\boldsymbol{b}^{(j)}$ πρέπει επίσης να βρίσκεται στον κυρτό φλοιό Δ_N, ιδιότητες που τους αποδίδουν εγγενώς το χαρακτηριστικό της αραιότητας. Η λύση του προβλήματος της σχέσης (2.29) μπορεί να αντιμετωπιστεί αποτελεσματικά από έναν αλγόριθμο ενεργών συνόλων.

4. Νόρμα ℓ_0^{tree}: Ιεραρχική αραιότητα. Στην προτεινόμενη εργασία η τυπική προσέγγιση αραιής αναπαράστασης επεκτείνεται εισάγοντας επιπλέον μια νόρμα $\psi\ (\cdot)$ που επιβάλλει αραιότητα ο ρόλος της οποίας είναι να ενσωματώσει τα άτομα του $\boldsymbol{D}$ και συνεπώς τους συντελεστές $\boldsymbol{\alpha}$ σε μία σταθερή ιεραρχική δενδροειδή δομή. Αν υποτεθεί ένα δένδρο T με K κόμβους: $T = \{j\}, j = \{1, \dots, K\}$ όπως ορίζεται στην [232], τότε ο στόχος είναι η εύρεση συγκεκριμένων συνδυασμών των συντελεστών $\boldsymbol{\alpha}$ οι οποίοι θα πρέπει υποχρεωτικά να είναι μέρος ενός συνδεδεμένου υποδένδρου της αρχικής δενδροειδούς δομής T. Προκειμένου να γίνει κατανοητή η πρόταση αυτή, δίνεται ένα παράδειγμα στο Σχήμα 9. Δεδομένης μιας λύσης $\boldsymbol{\alpha} \in \mathbb{R}^8$, τα σύνολα των προγόνων - $ancestors(j)$, ενός κόμβου j ορίζονται από το υποσύνολο των δεικτών που αντιστοιχεί στους προγόνους του κόμβου j.



Σχήμα 9: (α) Ένα παράδειγμα ενός διανύσματος λύσης που ενσωματώνεται σε μια δενδροειδή δομή *T*. (β) Λύση με κόμβους μη μηδενικών τιμών (μη σκιασμένοι κόμβοι) όπου ισχύει η ιδιότητα των προγόνων: $\boldsymbol{\alpha}_8 \neq \mathbf{0} \Rightarrow \boldsymbol{\alpha}_{(\text{ancestors}(8))} \neq \mathbf{0}$.

Έπειτα, το διάνυσμα λύσης $\boldsymbol{\alpha}$ θα πρέπει να τηρεί την συνθήκη: $\boldsymbol{a}_j \neq 0, \Rightarrow [\boldsymbol{a}_k \neq 0]$ για όλα τα k που ανήκουν στους προγόνους - $ancestors(j)$. Η συνθήκη αυτή ουσιαστικά επιβάλλει τον ακόλουθο κανόνα: η συνεισφορά του ατόμου $\boldsymbol{d}^{(j)} \in \mathbb{R}^d$ στην ανακατασκευή του σήματος $\boldsymbol{x}$, είναι επιτρεπτή μόνο αν τα στοιχεία του συνόλου των προγόνων-ancestors(j) ανήκουν επίσης στο σύνολο των δεικτών των κόμβων που συνεισφέρουν. Οι Jennaton κ.α. [245], έχουν προτείνει μια συμπληρωματική περιγραφή ορίζοντας τα σύνολα απογόνων-descendants(j) ως εξής: θέτουμε $\boldsymbol{a}_j = 0, \Rightarrow [\boldsymbol{a} = 0]$, για όλα εκείνα τα k τα οποία ανήκουν στους απογόνους-descendants(j). Αυτό διαισθητικά δηλώνει ότι, αν ένα άτομο $\boldsymbol{d}^{(j)} \in \mathbb{R}^d$ δεν συνεισφέρει ενεργά στην αναπαράσταση του $\boldsymbol{x}$, τότε δεν θα πρέπει να συνεισφέρουν ούτε οι απόγονοί του στο δένδρο T. Ακόμη, ο συμβολισμός $T \triangleq \{\text{descendants}(j)\}$, με $j = 1{:}\,p$, εννοεί το σύνολο των στοιχείων των απογόνων-descendants(j) για κάθε κόμβο. Τέλος, κάθε στοιχείο g του T αναφέρεται ως μια ομάδα. Τότε ο αριθμός των ομάδων g στο $\mathbf{G}$, που συνεισφέρουν στην αναπαράσταση $\boldsymbol{x}$, μπορεί να ρυθμιστεί με τη χρήση μιας νόρμας δένδρου $\psi(\boldsymbol{\alpha}) \triangleq l_0^{tree}(\boldsymbol{\alpha})$, η οποία εντοπίζει τουλάχιστον έναν μη μηδενικό συντελεστή του $\boldsymbol{\alpha}$:

$$\psi(\boldsymbol{\alpha}) \triangleq l_0^{tree}(\boldsymbol{\alpha}) = \sum_{g \in T} \delta^g(\boldsymbol{\alpha}), \tag{2.31}$$

$$\delta^g = \begin{cases} 1 \text{ αν υπάρχει } j \in g \text{ τέτοιο ώστε } \boldsymbol{\alpha}_j \neq 0 \\ \qquad\qquad 0 \text{ αλλιώς} \end{cases}$$

Έτσι, η (2.20) σε συνδυασμό με τον περιορισμό της (2.31), οδηγεί στην ακόλουθη μη κυρτό πρόβλημα βελτιστοποίησης:

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^K}\left(\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{D\alpha}\|_2^2+\lambda\sum_{g\in T}\delta^g\right. \tag{2.32}$$

Η εξίσωση (2.32) επιλύεται προσεγγιστικά με δυαδικούς διαχωρισμούς παρέχοντας μία λύση της μορφής:

$$\min_{\boldsymbol{\alpha}\in R^K}\left(\frac{1}{2}\|\boldsymbol{u}-\boldsymbol{a}\|_2^2+\lambda\sum_{g\in T}\delta^g(\boldsymbol{a})\right) \tag{2.33}$$

όπου $\boldsymbol{u}\in\mathbb{R}^K$ είναι ένα σταθερό σήμα. Το πρόβλημα αυτό αντιμετωπίζεται με την εισαγωγή του προσεγγιστικού τελεστή για τον όρο μη κυρτής κανονικοποίησης $\lambda\cdot\ell_0^{tree}(\boldsymbol{\alpha})$. Στη βιβλιογραφία έχει αποδειχθεί ότι η (2.33) μπορεί να λυθεί αποτελεσματικά με τον αλγόριθμο ISTA [232]. Η εκμάθηση λεξικού, στο πλαίσιο της προτεινόμενης ιεραρχικής αραιής κωδικοποίησης, υλοποιείται ως:

$$\begin{aligned}&\min_{\boldsymbol{D}\in C,\boldsymbol{A}\in\mathbb{R}^{K\times N}}\left(0.5\|\boldsymbol{X}-\boldsymbol{DA}\|_F^2+\lambda\ell_o^{tree}(\boldsymbol{A})\right)=\\&\min_{\boldsymbol{D}\in C,\boldsymbol{A}\in\mathbb{R}^{K\times N}}\frac{1}{N}\sum_{i=1}^{N}\left(0.5\left\|\boldsymbol{x}^{(i)}-\boldsymbol{D}\boldsymbol{a}^{(i)}\right\|_2^2+\lambda\ell_o^{tree}\left(\boldsymbol{a}^{(i)}\right)\right)\end{aligned} \tag{2.34}$$

όπου C είναι το κυρτό σύνολο των πινάκων $C\triangleq\boldsymbol{D}\in\mathbb{R}^{N\times K}$υ.π.τ. $\forall_{j=1:K},(\boldsymbol{d}^{(j)})^T\boldsymbol{d}^{(j)}\leq 1$.

3. ΕΦΑΡΜΟΓΕΣ ΣΕ ΕΞΥΠΝΑ ΔΙΚΤΥΑ ΗΛΕΚΤΡΙΚΗΣ ΕΝΕΡΓΕΙΑΣ

Το κεφάλαιο αυτό πραγματεύεται την αξιοποίηση των προβλεπτικών μεθόδων Μ.Μ στα πλαίσια των έξυπνων δικτύων ηλεκτρικής ενέργειας. Ανακεφαλαιώνοντας, στην παράγραφο 3.1 παρουσιάζεται μια αναλυτική εισαγωγή στα έξυπνα δίκτυα κατά την οποία περιγράφονται τα βασικά χαρακτηριστικά, η αρχιτεκτονική, οι τεχνολογίες και οι προδιαγραφές που ενσωματώνουν. Στην επόμενη παράγραφο 3.2, εξετάζεται από ερευνητικής σκοπιάς, ο ρόλος της πρόβλεψης του ηλεκτρικού φορτίου σε σχέση με τις προδιαγραφές του έξυπνου δικτύου και ταυτόχρονα παρατίθεται μια εκτενής βιβλιογραφική ανασκόπηση της πρόβλεψης του ηλεκτρικού φορτίου με μεθόδους Μ.Μ. Το κεφάλαιο ολοκληρώνεται στην παράγραφο 3.3, με την διεξοδική περιγραφή μιας σειράς από μελέτες περίπτωσης, με στόχο την αξιολόγηση των μεθοδολογιών πρόβλεψης που αναπτύχθηκαν στα πλαίσια της Δ.Δ.

3.1 Εισαγωγή στα έξυπνα ηλεκτρικά δίκτυα

Είναι γεγονός πως ο τομέας της ηλεκτρικής ενέργειας υφίσταται διαρκείς μεταβάσεις σε παγκόσμιο επίπεδο, οι οποίες έχουν ενταθεί ιδιαίτερα κατά τις τελευταίες δεκαετίες. Το παραδοσιακό δίκτυο ηλεκτρικής ενέργειας που είναι υπεύθυνο για την παροχή ισχύος στους καταναλωτές, βασιζεται σε ηλεκτρομηχανική υποδομή που επιβάλλει την κεντροποιημένη παραγωγή ενέργειας και την μονόδρομη ροή ενέγειας από το επίπεδο παραγωγής προς το επίπεδο διανομής, ενώ αποκλείει την δυνατότητα επικοιωνίας μεταξύ των επιμέρους στοιχείων του δικτύου. Ωστόσο, το συμβατικό ηλεκτρικό δίκτυο έχει υποβαθμιστεί και δεν είναι σε θέση να ανταποκριθεί στις απαιτήσεις που διαμορφώνουν οι ενεργειακές ανάγκες και οι τεχνολογικές και οικονομικές εξελίξεις [246]. Οι ευρέως γνωστές ελλείψεις του συστήματος ισχύος προκαλούνται από αστάθεια τάσης, διακοπτόμενη παροχής ενέργειας, περικοπές φορτίων (load curtailments), διακοπές ρεύματος και καταστάσεις ανισορροπίας στο ισοζύγιο ενέργειας ή υπερφόρτωσης [247]. Εκτός από αυτά τα ανεπιθύμητα ζητήματα, υπάρχουν πολλοί άλλοι παράγοντες που συντελούν στην ανάπτυξη ενός νέου παραδείγματος, του έξυπνου δικτύου (smart grid). Το μοντέλο αυτό παρουσιάζει βελτιωμένα χαρακτηριστικά παρακολούθησης, εκτέλεσης υπολογισμών και ελέγχου, χρησιμοποιώντας σύγχρονες τεχνολογίες σε αυτούς τους τομείς. Μερικοί από τους σημαντικότερους παράγοντες συνοψίζονται παρακάτω [248]:

1. Η απορρύθμιση της αγοράς ηλεκτρικής ενέργειας έχει θέσει νέους κανόνες στην εμπορία ενέργειας με αποτέλεσμα να προκύπτουν σενάρια ροής ισχύος και αβεβαιότητες για τη διαχείριση των οποίων το συμβατικό δίκτυο δεν είναι κατάλληλο.
2. Η αυξανόμενη διείσδυση των ανανεώσιμων πηγών ενέργειας έχει αλλάξει τη μορφή σύστημα διανομής από ένα ακτινικό σε ένα δικτυωτό σύστημα εντείνοντας τις απαιτήσεις για παρακολούθηση και έλεγχο. Η διαλείπουσα φύση αυτών των ανανεώσιμων πηγών ενέργειας αποτελεί επίσης πρόκληση λόγω της μεταβαλλόμενης δυναμικής του συστήματος ηλεκτρικής ενέργειας.
3. Η νέα εποχή ψηφιακής οικονομίας που διανύουμε απαιτεί παροχή ισχύος υψηλής ποιότητας και αξιοπιστίας.
4. Καθώς επεκτείνεται η δικτυωτή φύση του δικτύου ηλεκτρικής ενέργειας με την αύξηση της χρήσης επικοινωνιών ευρείας περιοχής, η απειλή φυσικών επιθέσεων

και κυβερνοεπιθέσεων εισάγει περαιτέρω πολυπλοκότητα κατά την παρακολούθηση και τη διασύνδεση των στοιχείων του ηλεκτρικού δικτύου.

5. Μια από τις σοβαρότερες ανησυχίες της σημερινής εποχής είναι η διατήρηση του περιβάλλοντος και η προώθηση της βιώσιμης ανάπτυξης. Οι στόχοι αυτοί επιτυγχάνονται μέσα από την ενεργειακή αποδοτικότητα του δικτύου, τη μείωση της ζήτησης αιχμής (peak demand) και τη μεγιστοποίηση της ενσωμάτωσης των ανανεώσιμων πηγών ηλεκτρικής ενέργειας στο δίκτυο.

Ένας ευρέως αποδεκτός ορισμός του έξυπνου δικτύου διατυπώνεται ως εξής [249]: Το έξυπνο δίκτυο περιλαμβάνει την ενοποίηση τεχνολογιών ισχύος, επικοινωνιών και πληροφοριών, στοχεύοντας σε μια βελτιωμένη υποδομή ηλεκτρικής ενέργειας που εξυπηρετεί επαρκώς τις ανάγκες των καταναλωτών ενώ παρέχει μια συνεχή εξέλιξη των εφαρμογών τελικής χρήσης. Παρόλο που η κύρια καινοτομία του έξυπνου δικτύου είναι η δυνατότητα αμφίδρομης επικοινωνίας και ροής ισχύος [250], έχουν εντοπιστεί ορισμένα θεμελιώδη χαρακτηριστικά τα οποία περιλαμβάνουν [251]:

1. Την ικανότητα αυτο-ίασης (self-healing), που σημαίνει ότι το έξυπνο δίκτυο έχει επίγνωση της κατάστασης λειτουργίας του που προκύπτει από την ευρεία παρακολούθηση των στοιχείων του σε πραγματικό χρόνο. Επιπλέον, έχει τη δυνατότητα νησιδοποίησης (islanding), απομονώνοντας τα ελαττωματικά στοιχεία από το υπόλοιπο σύστημα.
2. Την ανθεκτικότητα (resiliency) σε διαταραχές, φυσικές επιθέσεις και κυβερνοεπιθέσεις καθώς και φυσικές καταστροφές.
3. Την ανάμιξη των καταναλωτών. Σε αντίθεση με το συμβατικό δίκτυο, οι καταναλωτές συμμετέχουν ενεργά στο έξυπνο δίκτυο επιτρέποντάς τους να είναι συγχρόνως και παραγωγοί (prosumers). Αυτός ο ενεργός ρόλος των χρηστών θα βοηθήσει στην εξισορρόπηση της προσφοράς και της ζήτησης και θα διασφαλίσει την αξιοπιστία του δικτύου.
4. Την παροχή υψηλής ποιότητας ισχύος. Το έξυπνο δίκτυο χαρακτηρίζεται από πολύ υψηλής ποιότητας παροχή ισχύος. Αυτό θα περιλαμβάνει τη χρήση τεχνικών ελέγχου και μετρήσεων μέσω αισθητήρων σε πραγματικό χρόνο, επιτρέποντας ταχεία διάγνωση και λύσεις σε συμβάντα που επηρεάζουν την ποιότητα ισχύος, όπως υπερτάσεις μεταγωγής, βλάβες γραμμών και αρμονικές.
5. Την υψηλή διείσδυση ανανεώσιμων πηγών ενέργειας (ΑΠΕ) φιλικών προς το περιβάλλον.
6. Την αλληλεπίδραση με την αγορά. Το έξυπνο δίκτυο δημιουργεί τις προϋποθέσεις για ένα περιβάλλον αποδεσμευμένης αγοράς ηλεκτρικής ενέργειας. Το έξυπνο δίκτυο σχεδιάζεται έτσι ώστε να παρέχει την ευελιξία σε επιχειρήσεις κοινής ωφελείας, ρυθμιστικούς φορείς και καταναλωτές να επινοήσουν κατάλληλους κανόνες εργασίας και ανταλλαγής υπηρεσιών σύμφωνα με τις απαιτήσεις μιας συγκεκριμένης περιοχής.
7. Την ανάπτυξη «έξυπνων» (δηλ. αυτοματοποιημένων, διαδραστικών και πργματικού χρόνου) τεχνολογιών που βελτιστοποιούν τη λειτουργία του εξοπλισμού και των συσκευών των καταναλωτών και προορίζονται για μετρήσεις, όπως οι έξυπνοι μετρητές (smart meters), επικοινωνία σχετικά με τις λειτουργίες και την κατάσταση του δικτύου και αυτοματισμό της διανομής.
8. Την ανάπτυξη προτύπων επικοινωνίας και διαλειτουργικότητας του εξοπλισμού και των συσκευών που συνδέονται με το ηλεκτρικό δίκτυο, συμπεριλαμβανομένης της υποδομής που πλαισιώνει τη λειτουργία του δικτύου.

Τα ανωτέρω ιδιαίτερα χαρακτηριστικά ενός έξυπνου δικτύου ηλεκτρικής ενέργειας έχουν βελτιωθεί σημαντικά τα τελευταία χρόνια με τη χρήση εξελιγμένων τεχνολογιών πληροφοριών και επικοινωνιών (information and communication technologies - ICT) [252], επιτρέποντας την αμφίδρομη μετάδοση δεδομένων και πληροφοριών μέσω μιας ασφαλούς από κυβερνοεπιθέσεις διεπαφής. Ταυτόχρονα, η χρήση αναβαθμισμένων τεχνολογιών στα ηλεκτρονικά ισχύος έχει συντελέσει στην εκτέλεση των απαραίτητων εργασιών παρακολούθησης, μέτρησης και ελέγχου με αξιόπιστο τρόπο. Αναδεικνύεται λοιπόν, σε αυτό το πλαίσιο, η προφανής σημασία της υπολογιστικής νοημοσύνης ως ένα πολύτιμο συμπληρωματικό εργαλείο του σύγχρονου δικτύου διανομής και διαχείρισης της ηλεκτρικής ενέργειας για την αποτελεσματική διάγνωση και επίλυση προβλημάτων σε κάθε επίπεδο του συστήματος [253], δηλαδή την παραγωγή, την μεταφορά, την διανομή και την κατανάλωση ηλεκτρική ενέργειας.

Η ύπαρξη των ψηφιακών δεδομένων αποτελεί μια ακόμη ιδιαίτερη αναβάθμιση που προσφέρει το έξυπνο δίκτυο και η οποία επιτυγχάνεται με την εκτεταμένη χρήση έξυπνων αισθητήρων και δικτύων αισθητήρων [254]. Η ενσωμάτωση τέτοιων τεχνολογιών επιτρέπει την ανίχνευση διακοπών ρεύματος, προβλημάτων ποιότητας ισχύος, καθώς και την παρακολούθηση των ρυθμών παραγωγής και κατανάλωσης σε πραγματικό χρόνο. Παράλληλα, οι έξυπνοι υποσταθμοί που αποτελούνται από μετασχηματιστές, διακόπτες κυκλώματος, πυκνωτές και μεταγωγείς, είναι τοποθετημένοι στο δίκτυο διανομής για την παρακολούθηση της απόδοσης του συστήματος και την εξυπηρέτηση των τοπικών φορτίων.

Επιπλέον, η αποδοτικότητα του έξυπνου δικτύου οφείλεται σε μεγάλο βαθμό στη βελτίωση των συστημάτων εποπτικού ελέγχου και απόκτησης δεδομένων (supervisory control and data acquisition - SCADA) και διαχείρισης ενέργειας (energy management systems - EMS), μέσω της εισαγωγής των συγχρονισμένων μετρήσεων φασιθέτη (synchrophasors) και των μονάδων μέτρησης φασιθετών (phasor measurement units - PMUs) [255], [256]. Τα στοιχεία αυτά παρέχουν διαδραστική και άμεση παρακολούθηση του συνολικού συστήματος μέσα από αναλύσεις συμβάντων (events). Επίσης, είναι εφοδιασμένα με μια σειρά από ρουτίνες ανίχνευσης, για τον εντοπισμό του ρυθμού αλλαγής συχνότητας (rate of change of frequency - RoCoF), νησιδοποίησης και σταθερότητας της τάσης, ενώ ταυτόχρονα βασίζονται στην απόκτηση δεδομένων από παγκόσμια συστήματα στιγματοθέτησης (global positioning systems - GPS) και την μετάδοσή τους σε ένα κέντρο παρακολούθησης. Με τον τρόπο αυτό, παρέχονται προβλέψεις και προειδοποιήσεις για τυχόν δυσλειτουργίες οι οποίες μπορεί να οφείλονται στο εύρος, την φάση ή τη συχνότητα ισχύος και τάσης.

Έτσι, η νέα δομή πλέγματος (smart grid) που έχει επικρατήσει, σε συνδυασμό με χρήση των τεχνολογιών που αναφέρθηκαν και την εφαρμογή τεχνικών που βασίζονται σε τεχνητή νοημοσύνη έχουν προσδώσει στο έξυπνο δίκτυο δυνατότητες αυτό-αποκατάστασης (self-restoration) και αυτό-ίασης, τη στιγμή που η αποκατάσταση βλαβών του συμβατικού δικτύου πραγματοποιείται χειροκίνητα ή τοπικά [257], [258].

Η υλοποίηση των βασικών χαρακτηριστικών του έξυπνου δικτύου απαιτεί την ενσωμάτωση ποικίλων τεχνολογιών οι οποίες αναφέρονται στη συνέχεια [259], [260].

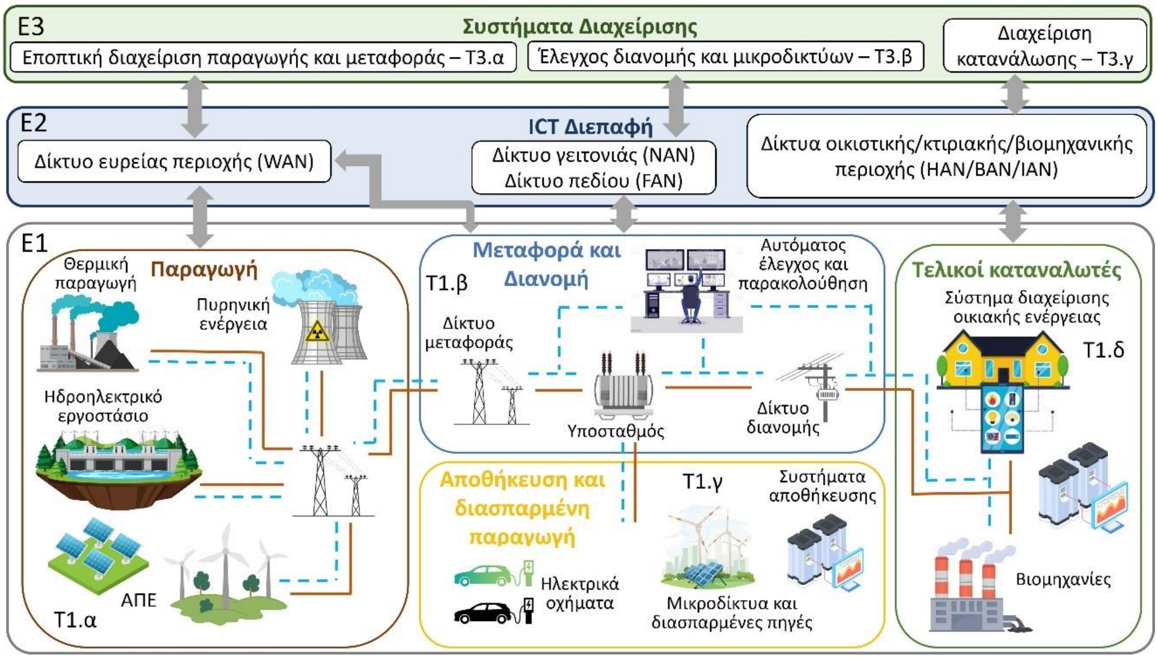
- Διασπαρμένη παραγωγή και αποθήκευση:
  Αναφέρεται στην ευρεία χρήση διασπαρμένων πηγών ενέργειας, με έμφαση στις φιλικές προς το περιβάλλον ΑΠΕ, όπως οι ανεμογεννήτριες, τα φωτοβολταϊκά, η γεωθερμία και τα μικρό-υδροηλεκτρικά εργοστάσια, με στόχο την αύξηση της ποιότητας ισχύος. Η βασική ιδέα είναι ότι η ενέργεια παράγεται και διανέμεται όχι από μεγάλες μονάδες αλλά από μικρότερες πηγές εγκατεστημένες κοντά στους καταναλωτές. Η ενσωμάτωση της διασπαρμένης παραγωγής όχι μόνο μειώνει το λειτουργικό κόστος, αλλά προσφέρει επίσης ορισμένα οφέλη στους σχεδιαστές

δικτύου, όπως η μείωση της ζήτησης αιχμής και η παροχή ενέργειας σε απομακρυσμένες και δυσπρόσιτες περιοχές, ειδικά στις αναπτυσσόμενες χώρες.

- Παρακολούθηση και έλεγχος σε πραγματικό χρόνο:
  Σε αντίθεση με το περιβάλλον του συμβατικού δικτύου, ο καταναλωτής αναμένεται να συμμετέχει ενεργά στις λειτουργίες του έξυπνου δικτύου μέσα από την εγκατάσταση ενός μεγάλου αριθμού αισθητήρων και ατομικών έξυπνων μετρητών που παρέχουν πληροφορίες για την κατάσταση των στοιχείων και της απόδοσης του συστήματος σε πραγματικό χρόνο.
- Διανεμημένη ευφυΐα
  Η δικτυωτή δομή και η αμφίδρομη ροή ισχύος και πληροφορίας που χαρακτηρίζουν το έξυπνο δίκτυο, επιβάλλουν την ύπαρξη της κατάλληλης υποδομής για την αποτελεσματική λειτουργία και διαχείρισή του. Αυτή μπορεί να επιτευχθεί μέσα από την αξιοποίηση μεθόδων και τεχνολογιών διανεμημένης ευφυΐας και ελέγχου σε όλα τα επίπεδα του δικτύου, όπως οι έξυπνες οικιακές συσκευές, οι συσκευές edge computing, και οι απομακρυσμένες τερματικές μονάδες (remote terminal units - RTUs).
- Διαχείριση ζήτησης (demand-side management – DSM)
  Η διαχείριση ζήτησης περιλαμβάνει την εφαρμογή διάφορων λύσεων από την πλευρά των καταναλωτών οι οποίες οδηγούν σε μια επιθυμητή μεταβολή της καμπύλης κατανάλωσης και παραγωγή φορτίου. Οι λύσεις αυτές επηρεάζουν κυρίως τη συμπεριφορά των καταναλωτών που σχετίζεται με τη χρήση της ηλεκτρικής ενέργειας. Ανάλογα με τον τομέα όπου στοχεύει η κάθε μία από την πλευρά του πελάτη, οι προσεγγίσεις DSM μπορούν να κατηγοριοποιηθούν σε (α) ενεργειακή απόδοση, (β) χρόνοπρογραματισμό κατανάλωσης και (γ) απόκριση ζήτησης (demand response - DR).
- Απόκριση ζήτησης
  Είναι γεγονός ότι η ζήτηση της ηλεκτρικής ενέργειας εμφανίζει μια διακύμανση κατά τη διάρκεια της ημέρας με μια μικρή περίοδο αιχμής, τουτέστιν αυξημένης ζήτησης, διάρκειας λίγων ωρών. Ένα σημαντικό μέλημα για τους διαχειριστές και τους σχεδιαστές του δικτύου είναι να μειώσουν τη διαφορά μεταξύ του φορτίου αιχμής (peak load) και των τιμών του φορτίου κατά τη διάρκεια της υπόλοιπης ημέρας, για τη βέλτιστη χρήση του δικτύου ηλεκτρικής ενέργειας. Αυτή η εξομάλυνση της καμπύλης του φορτίου μπορεί να επιτευχθεί με την μετακύλιση των φορτίων (load shifting) σε διαφορετικές ώρες της ημέρας.

Τα πεδία λειτουργίας του έξυπνου δικτύου αποτελούνται από τη μαζική παραγωγή, τη μεταφορά, τη διανομή και την κατανάλωση στο επίπεδο της ροής ισχύος και από τις αγορές, τους διαχειριστές και τους παρόχους υπηρεσιών ηλεκτρικής ενέργειας στο επίπεδο της ροής πληροφοριών [261]. Τα επίπεδα ισχύος και επικοινωνίας αλληλοεπιδρούν μεταξύ τους χρησιμοποιώντας αρχιτεκτονικές επικοινωνίας, όπως δίκτυα περιοχής (area networks), αρχιτεκτονικές συστημάτων ισχύος και αρχιτεκτονικές ICT.

Η απεικόνιση της ολοκληρωμένης υποδομής ενός έξυπνου δικτύου φαίνεται στο Σχήμα 10. Το κατώτερο επίπεδο (επίπεδο 1) αναπαριστά το σύστημα ισχύος του έξυπνου δικτύου το οποίο περιλαμβάνει τη μαζική παραγωγή (τμήμα 1.α), την μεταφορά, την διανομή (τμήμα 1.β), την αποθήκευση, την διασπαρμένη παραγωγή ενέργειας (τμήμα 1.γ) και τους τελικούς καταναλωτές (τμήμα 1.δ). Η μαζική παραγωγή προκύπτει κυρίως από συμβατικές πηγές όπως μονάδες θερμικής παραγωγής ενέργειας, παραγωγής πυρηνικής ενέργειας, υδροηλεκτρικά εργοστάσια και ΑΠΕ όπως αιολικά και ηλιακά πάρκα. Οι γραμμές ροής ισχύος του δικτύου απεικονίζονται με κόκκινο χρώμα, ενώ οι διακεκομμένες μπλε γραμμές είναι γραμμές επικοινωνίας. Οι διαδικασίες ελέγχου και παρακολούθησης της παραγωγής εκτελούνται στο τμήμα 1.α από τους περιφερειακούς διαχειριστές μεταφοράς (regional transmission operators - RTOs) και τους διεθνείς οργανισμούς προτύπων (international standards organizations - ISOs).



Σχήμα 10: Τα βασικά στοιχεία στην υποδομή του έξυπνου δικτύου.

3.2 Ερευνητικό υπόβαθρο

3.2.1 Η πρόβλεψη του ηλεκτρικού φορτίου σε σχέση με τις προδιαγραφές των έξυπνων δικτύων

Προκειμένου να είναι επιτυχής και αποτελεσματική η ενσωμάτωση των προαναφερθέντων τεχνολογιών στο πλαίσιο των έξυπνων δικτύων και να επιτυγχάνονται οι στόχοι του, θα πρέπει να τηρούνται μια σειρά από προδιαγραφές σε σχέση με τις αλληλεπιδράσεις ανάμεσα στις εμπλεκόμενες οντότητες, τις τεχνολογικές εξελίξεις και τους κανόνες της αγοράς ενέργειας. Η πρόβλεψη του ηλεκτρικού φορτίου αποτελεί ένα συμπληρωματικό στοιχείο κατά την καθορισμό των προδιαγραφών αυτών όπως καταδεικνύεται παρακάτω.

Όπως έχει ήδη αναφερθεί το παραδοσιακό δίκτυο ηλεκτρικής ενέργειας έχει υποστεί ραγδαίες μεταβολές, με αποτέλεσμα την υπάρχουσα κατάσταση η οποία χαρακτηρίζεται από τα εξής: Από τη μια μεριά το μοντέλο της ελεύθερης αγοράς ενέργειας [262] έχει προκαλέσει την αύξηση του ανταγωνισμού ανάμεσα στους διαφορετικές οντότητες που συμμετέχουν σε αυτήν [263], ενώ από την άλλη μεριά, οι συνεχείς αναβαθμίσεις στον εξοπλισμό, την ασφάλεια, την ποιότητα ισχύος και την τεχνολογία επικοινωνίας και πληροφοριών [264], [265] έχουν πυροδοτήσει την εξάπλωση των έξυπνων δικτύων τα οποία διασυνδέουν διαχειριστές (operators), φορείς σωρευτικής εκπροσώπησης (aggregators) και prosumers [266], [267]. Παράλληλα, η ενεργειακή κρίση που κλονίζει τον πλανήτη ασκεί έντονη πίεση για εξυπνότερη και αποδοτικότερη διαχείριση την παραγωγής και διανομής ενέργειας [268], [269]. Συνεπώς, ο συνδυασμός των ανωτέρω παραγόντων συνθέτει ένα περιβάλλον υψηλής αβεβαιότητας και αστάθειας για τα μέλη του δικτύου, είτε αυτοί είναι διαχειριστές είτε παραγωγοί/καταναλωτές.

Μια από τις προτεινόμενες λύσεις μεταξύ άλλων για την αντιμετώπιση των ανωτέρω θεμάτων είναι η ψηφιοποίηση της υπάρχουσας τεχνολογίας [270] καθώς και η ενσωμάτωση ευφυών συστημάτων για την υποστήριξη των αναγκών του δικτύου στο επίπεδο της κατανάλωσης [271]. Τέτοιου είδους λύσεις αναφέρονται ως συστήματα διαχείρισης διασπαρμένης ενέργειας (distributed energy management systems - DERMS) και παρουσιάζουν μια σειρά από πλεονεκτήματα όπως, η άμεση επεκτασιμότητα του δικτύου, η συμβατότητα με τεχνολογίες αυτοματισμού, η εύκολη ενσωμάτωση των συναλλαγών ενέργειας σε αυτά, η ανθεκτικότητα σε βλάβες του δικτύου, και η περιβαλλοντική βιωσιμότητα [272], [273].

Αναπόσπαστο μέρος μιας τέτοιας ψηφιακής πλατφόρμας αποτελεί η πρόβλεψη της κατανάλωσης ή/και ζήτησης του ηλεκτρικού φορτίου που συμβάλλει σε μια πληθώρα διαφορετικών λειτουργιών της τοπολογίας ενός έξυπνου δικτύου [274], [275]. Αρχικά, από τη σκοπιά του Διαχειριστή Συστήματος Διανομής - ΔΣΔ (Distribution System Operator - DSO), ένα ουσιαστικό στάδιο για τη διασφάλιση της αποδοτικότητας, τόσο της οικονομικής όσο και στο επίπεδο των διαθέσιμων πόρων, είναι ο χρονοπρογραμματισμός των μονάδων παραγωγής ενέργειας σε έναν ορίζοντα τουλάχιστον είκοσι τεσσάρων ωρών, δηλαδή της επόμενης ημέρας [276]. Είναι προφανές ότι για να επιτευχθεί αυτό θα πρέπει κανείς να έχει στη διάθεσή του ακριβείς προβλέψεις ζήτησης ηλεκτρικού φορτίου οδηγώντας σε ευέλικτα, διαλειτουργικά και βιώσιμα έξυπνα δίκτυα. Ειδικότερα, η δημοσίευση [253] που ήταν αποτέλεσμα της παρούσας έρευνας περιλαμβάνει μια ενδελεχή ανασκόπηση σχετικά με την εφαρμογή μεταευρετικών μεθόδων βελτιστοποίησης στο πλαίσιο του έξυπνου δικτύου, όπου γίνεται φανερή η συνεισφορά της πρόβλεψης του ηλεκτρικού φορτίου σε μια πληθώρα εργασιών βελτιστοποίησης που εκτείνονται από τον βραχυπρόθεσμο έλεγχο ροής ισχύος έως τον μακροπρόθεσμο σχεδιασμό του μεγέθους και της σύστασης του δικτύου.

Από την άλλη πλευρά, ίσως το πιο θεμελιώδες χαρακτηριστικό των παραδοσιακών δικτύων ηλεκτρικής ενέργειας αποτελεί η μονόδρομη ροή της ενέργειας από τις μονάδες παραγωγής στις μονάδες κατανάλωσης, όπως εργοστάσια, σπίτια, δημόσιος φωτισμός και ηλεκτρικά οχήματα. Ωστόσο, όπως έχει ήδη αναφερθεί, στα πλαίσια των έξυπνων δικτύων, οι καταναλωτές επιτρέπεται να παράγουν τη δική τους ενέργεια από ανανεώσιμες ή από άλλους τύπους πηγών ενέργειας και να πωλούν στο δίκτυο την ποσότητα παραγόμενης ενέργειας η οποία δεν καταναλώνεται [277]. Ως εκ τούτου, το νέο μοντέλο της αμφίδρομης ροής ισχύος θα έχει (και έχει) ως αποτέλεσμα την αύξηση της διείσδυσης των ΑΠΕ [278]. Έτσι, οι μετρήσεις της ενεργούς ισχύος – Ε.Ι. (active power - AP) που αποκτώνται από οποιονδήποτε υποσταθμό υψηλής τάσης/μέσης τάσης (ΥΤ/ΜΤ) είναι το αποτέλεσμα δεδομένων τόσο παραγόμενης όσο και ζητούμενης ενέργειας, αναφερόμενες εφεξής ως «μεικτό φορτίο». Λαμβάνοντας επίσης υπόψη τη στοχαστική φύση των ΑΠΕ λόγω διαλείπουσας παραγωγής [279], το έργο της πρόβλεψης του μικτού φορτίου γίνεται αρκετά δύσκολο και απαιτεί τη χρήση προηγμένων μεθοδολογιών Μ.Μ. Η αστάθεια της παραγωγής από ΑΠΕ στα σύγχρονα δίκτυα ηλεκτρικής ενέργειας αντιμετωπίζεται με την εγκατάσταση συστημάτων αποθήκευσης ενέργειας [280] και την εφαρμογή ενεργειών μετακύλισης και εξισορρόπησης φορτίου [281], οι οποίες απαιτούν, όπως κι αν έχει το πράγμα, αξιόπιστους και εύρωστους αλγόριθμους πρόβλεψης φορτίου.

Μια άλλη θεμελιώδης αλλαγή που λαμβάνει χώρα αυτή τη στιγμή στην απορυθμισμένη αγορά ενέργειας είναι η αποδέσμευση των ΔΣΔ [282], επιτρέποντας την εισαγωγή ολοένα και περισσότερων φορέων και κατ' ακολουθίαν αυξάνοντας τον ανταγωνισμό μεταξύ τους. Προκειμένου οι παραγωγοί - καταναλωτές να προετοιμάσουν τις αντίστοιχες στρατηγικές υποβολής προσφορών ώστε να μεγιστοποιήσουν τα κέρδη τους, η πρόβλεψη της τιμής εκκαθάρισης της αγοράς σε καθημερινή βάση αναμένεται να είναι ένα ισχυρό εργαλείο [283], το οποίο μπορεί να ενισχυθεί σημαντικά με την εφαρμογή μεθόδων βραχυπρόθεσμης πρόβλεψης φορτίου (short-term load forecasting - STLF) [284]. Από επιχειρησιακή άποψη, η STLF μπορεί επίσης να είναι ένας παράγοντας βελτίωσης της ποιότητας ισχύος, της ασφάλειας, της προστασίας του εξοπλισμού και της σταθερότητας των δικτύων ισχύος [285], [286].

3.2.2 Βιβλιογραφική έρευνα πρόβλεψης ηλεκτρικού φορτίου

Μεταξύ των μεθοδολογιών που έχουν προταθεί για την αντιμετώπιση του προβλήματος της πρόβλεψης του ηλεκτρικού φορτίου, το πεδίο της υπολογιστικής νοημοσύνης και συγκεκριμένα ο κλάδος της μηχανικής μάθησης [287] έχει φανεί αναπάντεχα ικανός στο να παρέχει μια πληθώρα προσεγγίσεων με αξιόλογα αποτελέσματα. Για τον λόγο αυτό, παρατίθεται παρακάτω μια βιβλιογραφική επισκόπηση μεταξύ των πιο διαδεδομένων και καθιερωμένων μεθοδολογιών μηχανικής μάθησης, οι οποίες χρησιμοποιούνται για την πρόβλεψη της κατανάλωσης του ηλεκτρικού φορτίου.

Η Γραμμική Παλινδρόμηση (ΓΠ) ανήκει στις μεθόδους που πρωταρχικά εφαρμόστηκαν για την πρόβλεψη της κατανάλωσης του ηλεκτρικού φορτίου. Μία προσέγγιση που βασίζεται στην γραμμική παλινδρόμηση χρονοσειρών αποδομημένων σε επιμέρους στοιχεία που αναπαριστούν την εποχή του έτους και την ημέρα της εβδομάδας, παρουσιάζεται στη [288] για τη μοντελοποίηση τόσο της εποχικής τάσης όσο και της μεταβλητότητας (volatility) της ζήτησης ηλεκτρικής ενέργειας. Μία πιθανοκρατική προσέγγιση για την αντιμετώπιση της αβεβαιότητας που εμφανίζεται στα δεδομένα του φορτίου έχει προταθεί και παρουσιαστεί στην [289], εφαρμόζοντας παλινδρόμηση ποσοστιαίων σημείων, ενώ ένα μοντέλο που βασίζεται στην ίδια μέθοδο αλλά λαμβάνει σαν εισόδους επίσης τις καιρικές συνθήκες και της τιμή της ηλεκτρικής ενέργειας παρουσιάζεται στην [290]. Η επίδραση των μεταβλητών που περιγράφουν τις κλιματικές συνθήκες, στην πρόβλεψη της ζήτησης της ηλεκτρικής ενέργειας μελετάται επίσης διεξοδικά στις αναφορές [291] και [292]. Το ζήτημα της βελτίωσης της πρόβλεψης δεδομένων κατανάλωσης ηλεκτρικού φορτίου από ένα σύνολο έξυπνων μετρητών διερευνάται στη [293], προτείνοντας ένα περιοδικό αυτοπαλινδρομικό μοντέλο με εξωγενείς μεταβλητές (periodic autoregressive model with exogenous variables - PARX), οι οποίες περιλαμβάνουν την θερμοκρασία και ημερολογιακές πληροφορίες. Ένα σημαντικό ερευνητικό θέμα που ανακύπτει κατά την διαδικασία της πρόβλεψης του ηλεκτρικού φορτίου αποτελεί η παρουσία των εποχικών φαινομένων, οι οποίες προκύπτουν τόσο από την μεταβλητότητα της κατανάλωσης κατά τη διάρκεια της ημέρας και για διαφορετικές πειόδους του έτους, όσο και από τις μεταβολές των καιρικών συνθηκών. Για παράδειγμα, για έναν υποτιθέμενο Υ/Σ, εγκατεστημένο σε μια περιοχή της Ελλάδας, συνήθως αναμένεται μια άνοδος της παραγωγής αλλά και της ζήτησης, κατά τους καλοκαιρινούς μήνες λόγω της μεγάλης διάρκειας ηλιοφάνειας και της υψηλής χρήσης κλιματιστικών, αντίστοιχα. Προς αυτή την κατεύθυνση, έχουν προταθεί διάφορες τεχνικές, μεταξύ των οποίων η τριπλή εκθετική εξομάλυνση (triple exponential smoothing) [294], η μέθοδος της αποδόμησης [295], [296] ή η χρήση χρονοσειρών πολλαπλών εξισώσεων (multiple equation time-series) [297]. Ακολουθώντας μια παρόμοια προσέγγιση, η πρόβλεψη της μέγιστης τιμής της εβδομαδιαίας κατανάλωσης φορτίου με ορίζοντα ενός έτους επιχειρείται στη [298], όπου το φορτίο αποδομείται σε τρία διαφορετικά συστατικά, την δυναμική τάση αύξησης της κατανάλωσης, την εποχική τάση που σχετίζεται με τις καιρικές συνθήκες και την περιοδικότητα της κατανάλωσης που δεν οφείλεται στις καιρικές συνθήκες, οι οποίες μοντελοποιούνται με την χρήση τεχνικών Autoregressive Integrated Moving Average with Exogenous variables (ARIMAX) και ARIMA. Συγκεκριμένα, με τα μοντέλα αυτά ενσωματώνουμε τα σφάλματα προηγούμενων προβλέψεων στην εξίσωση παλινδρόμησης. Έτσι, επιτυγχάνεται η υπεροχή στην απόδοση της πρόβλεψης σε σχέση με αυτά που προκύπτουν από τα απλά AR μοντέλα. Οι συγγραφείς της [114] προτείνουν και κάνουν χρήση ενός μοντέλου ARIMA για την πρόβλεψη της ζήτησης της ενέργειας για τη φόρτιση ηλεκτρικών οχημάτων, οι έξοδοι του οποίου αξιοποιούνται για τη μορφοποίηση ενός προβλήματος στοχαστικού χρονοπρογραμματισμού των μονάδων παραγωγής ενέργειας για το επόμενο 24ωρο.

Στα τελευταία έτη, η αραιότητα αποκτά ολένα και περισσότερη δημοφιλία στον τομέα της μοντελοποίησης σημάτων. Παρόλο που αρχικά είχε εφαρμοστεί στην επεξεργασία σήματος και μετέπειτα εικόνας, κυρίως ως εργαλείο αποθορυβοποίησης, πρόσφατα έχει παρουσιαστεί ένα αυξανόμενο ερευνητικό ενδιαφέρον για την χρήση της σε ηλεκτρικά σήματα. Όπως έχει ήδη αναφερθεί, αναπαράσταση συνίσταται στην αναζήτηση της αραιότερης λύσης του διανύσματος των αραιών συντελεστών, δηλαδή εκείνης με τα λιγότερα μη μηδενικά στοιχεία, δεδομένου ενός λεξικού. Επίσης, ένα καθοριστικό στάδιο κατά την προαναφερθεία διαδικασία είναι η εκμάθηση του λεξικού, η οποία έγκειται στην εύρεση των ατόμων και μορφοποιείται σαν ένα απο κοινού πρόβλημα βελτιστοποίησης όπως περιγράφεται από τις σχέσεις (2.20) και (2.21). Η ανάπτυξη μεθόδων αραιής αναπαράστασης για την μοντελοποίηση και την πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας νοικοκυριών μελετάται στο [238] προτείνοντας έναν αλγόριθμο που βασίζεται στη μέθοδο εναλλασόμενων διευθύνσεων των πολλαπλασιαστών (alternating direction method of multipliers - ADMM) για την επίλυση του προβλήματος της εκμάθησης του λεξικού. Ένας αριθμός ερευνητικών εργασιών που βασίζονται στην αραιή Μπεϋζιανή μάθηση (sparse Bayesian learning - SBL) έχουν επίσης δημοσιευθεί κατά την τελευταία δεκαετία, χρησιμοποιώντας τη μέθοδο σταθμισμένης (weighted) SBL [299], [300] ή της SBL με συνδυασμένους πυρήνες (combined kernels) [301]. Πέρα από την πρόβλεψη τουορτίου, η προσέγγιση της αραιής κωδικοποίησης έχει χρησιμοποιηθεί για την αντιμετώπιση πολλών ακόμη προβλημάτων που αφορούν τα δίκτυα ισχύος. Στις δημοσίευσεις [302], [303], η πρόβλεψη της παραγώμενης αιολικής ενέργειας σε πραγματικό χρόνο επιτυγχάνεται με τη μέθοδο της αραιής κωδικοποίησης με χρήση μη αρνητικών συντελεστών και με ελαστική κανονικοποίηση (elastic net regularization). Ο γνωστός για την αποτελεσματικότητά του αλγόριθμος K-SVD [238], [303], έχει επίσης χρησιμοποιθεί πρόσφατα στο πλαίσιο του έξυπνου δικτύου, είτε για το διαχωρισμό της καταναλισκώμενης ενέργειας ενός κτιρίου στην ενέργεια που δαπανάται από κάθε συσκευή του κτιρίου ξεχωριστά [304], είτε για την συμπίεση δεδομένων από ατομικούς έξυπνους μετρητές και την εξαγωγή μοτίβων μερικής χρήσης (partial usage patterns) που εντοπίζονται στο προφίλ του φορτίου των καταναλωτών [305].

Η πρόβλεψη του φορτίου με τη μέθοδο της παλινδρόμησης με χρήση μηχανών διανυσμάτων υποστήριξης (support vector regression - SVR), έχει κερδίσει επίσης δημοτικότητα τα τελευταία χρόνια, λόγω της ικανότητας αυτής της μεθόδου να μοντελοποιεί τις μη γραμμικότητες που δυνητικά εντοπίζονται στα δεδομένα των χρονοσειρών του ηλεκτρικού φορτίου. Το μοντέλο SVR που παρουσιάζεται στην [306] αφορά στην πρόβλεψη του φορτίου αιχμής, δηλαδή της μέγιστης τιμής κατανάλωσης, διαμέσου μιας πολυωνυμικής συνάρτηση πυρήνα, οι παράμετροι της οποίας βελτιστοποιούνται μέσω της πολλαπλής διασταυρούμενης αξιολόγησης. Σε αντίθεση με τα συμβατικά μοντέλα SVR, η ιδέα της συγκέντρωσης προβλέψεων από πολλαπλά διαφορετικά μοντέλα με διαφορετικές συναρτήσεις πυρήνα εισάγεται στις [307] και [308]. Η πρόβλεψη της ωριαίας κατανάλωσης φορτίου αποτελεί το αντικείμενο μελέτης της [309], υλοποιώντας ένα μοντέλο SVR που χρησιμοποιεί τη μέθοδο της εμπειρικής αποδόμησης (empirical mode decomposition – EMD) για το διαχωρισμό μιας χρονοσειράς σε δύο συστατικά, τα οποία περιγράφουν την εποχική τάση και τις τοπικές διαταραχές στις μετρήσεις κατανάλωσης αντίστοιχα. Ένα υβριδικό μοντέλο [310] έχει προταθεί πρόσφατα για την πρόβλεψη της ζήτησης με ορίζοντα μισής ώρας, το οποίο χρησιμοποιεί έναν τροποποιημένο αλγόριθμο fire-fly optimization (FFO) για τη βαθμονόμηση των παραμέτρων του μοντέλου.

Λόγω της ικανότητάς τους να ανιχνεύουν μη γραμμικές συσχετίσεις, τα νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης [311] έχουν εφαρμοστεί εκτενώς για την πρόβλεψη τόσο της κατανάλωσης του ηλεκτρικού φορτίου όσο και της ζήτησης. Οι πιο αντιπροσωπευτικές αρχιτεκτονικές FFNN είναι τα δίκτυα MLP και RBF. Οι δημοσιεύσεις [312], [313] και [314] προτείνουν την ανάπτυξη δικτύων πρόσθιας τροφοδότησης για την πρόβλεψη της ωριαίας κατανάλωσης. Εκτός από τα ιστορικά δεδομένα φορτίου, στις μεταβλητές εισόδου των δύο πρώτων μοντέλων συμπεριλαμβάνονται ορισμένοι ημερολογιακοί δείκτες, όπως η μέρα της εβδομάδας, η ώρα της ημέρας και η εποχή του έτους, ενώ η θερμοκρασία και η υγρασία εισάγονται στο τρίτο μοντέλο. Η χρήση ημερολογιακών και κλιματικών μεταβλητών εισόδου εφαρμόζεται στα περισσότερα μοντέλα πρόβλεψης, καθώς η θετική επίδραση των εξωγενών μεταβλητών κατά την πρόβλεψη του ηλεκτρικού φορτίου έχει επιβεβαιωθεί από μια πληθώρα εργασιών στη βιβλιογραφία, [315]–[317], [318]. Σε πολλές περιπτώσεις, ο αλγόριθμος Levenberg-Marquardt επιλέγεται για την εκπαίδευση των νευρωνικών δικτύων σημειώνοντας ανώτερη απόδοση σε σχέση με άλλους αλγορίθμους [319], [320]. Σύμφωνα με την αναφορά [321], ένα δίκτυο με πολλαπλές στοιβάδες (multi-layered) βελτιστοποιείται με την χρήση του αλγορίθμου grey wolf optimization - GWO. Τα πειραματικά αποτελέσματα δείχνουν ότι το προτεινόμενο μοντέλο παρουσιάζει υψηλότερη προβλεπτική ακρίβεια σε σύγκριση τόσο με την απλή γραμμική παλινδρόμηση αλλά και με τα νευρωνικά δίκτυα MLP σε συνδυασμό με μεταευρετικές μεθόδους. Επιπλέον, ένα τεχνητό νευρωνικό δίκτυο που ακολουθεί μια καινοτόμα προσέγγιση για τη σύνδεση μεταξύ των στοιβάδων, η οποία αποκαλείται μέση πυκνή σύνδεση (dense average connection), έχει προταθεί πρόσφατα [322], εμφανίζοντας ικανοποιητικά αποτελέσματα. Έχοντας σαν στόχο τη βελτίωση της διαδικασίας της εκπαίδευσης ενός FFNN τριών στοιβάδων, οι συγγραφείς της [323] υλοποιούν έναν αναβαθμισμένο αλγόριθμο οπισθοδιάδοσης, σημειώνοντας ικανοποιητικά αποτελέσματα πρόβλεψης, με βάση τον δείκτη MAPE. Μία μεθοδολογία βραχυπρόθεσμης πρόβλεψης του φορτίου των κόμβων ενός δικτύου περιγράφεται επίσης στην [324], όπου ένα νευρωνικό δίκτυο MLP συνδυάζεται με έναν αλγόριθμο ομαδοποίησης των κόμβων, επιτυγχάνοντας έτσι μειωμένο υπολογιστικό χρόνο και παρόμοια προβλεπτική ικανότητα σε σύγκριση με τα απλά δίκτυα MLP. Ένα σύστημα πολλαπλών πρακτόρων (multi-agent system – MAS) σχεδιάζεται στην [325] συμπεριλαμβανομένης α) της αναγνώρισης μοτίβων ζήτησης ηλεκτρικού φορτίου σε διάστημα 24 ωρών, β) της κατάταξης των ημερήσιων προφίλ ζήτησης σε κατηγορίες με βάση τα μοτίβα αυτά και τέλος γ) ένα σύνολο από δίκτυα MLP για την πρόβλεψη της ζήτησης. Η εφαρμογή των δικτύων RBF για την πρόβλεψη του ηλεκτρικού φορτίου παρατηρείται επίσης σε μικρότερη κλίμακα με βάση τη βιβλιογραφία, τόσο σε βραχυπρόθεσμους [326] όσο και σε μακροπρόθεσμους [327] ορίζοντες πρόβλεψης.

Τα Τυχαία Δάση (Random forests - RF) αποτελούν μια μέθοδο μηχανικής μάθησης που βασίζεται σε πολλαπλές επιμέρους προβλέψεις παρουσιάζοντας πολύ καλή προβλεπτική ακρίβεια ενώ έχει χρησιμοποιηθεί σε ένα μεγάλο αριθμό εφαρμογών, συμπεριλαμβανομένης και της πρόβλεψης του ηλεκτρικού φορτίου. Διάφορα παλινδρομικά μοντέλα αξιολογούνται στη [328], με το μοντέλο τυχαίων δασών να παρέχει καλύτερες βραχυπρόθεσμες προβλέψης ζήτησης απ' ότι τα μοντέλα KNN και γραμμικής παλινδρόμησης σύμφωνα με τον δείκτη MAPE. Ένα πιθανοκρατικό μοντέλο πρόβλεψης του φορτίου βασισμένο σε παλινδρόμηση ποσοστιαίων σημείων αναπτύσσεται στη [329] και επιπλέον χαρακτηριστικό την αναδρομικής εξάλειψης χαρακτηριστικών (recursive feature elimination) για την επιλογή των μεταβλητών εισόδου. Επιπλέον, οι συγγραφείς προτείνουν μία εναλλακτική μέθοδο για τον προσδιορισμό των ποσοστημορίων (quantiles) για την αντιμετώπιση του ζητήματος της αξιοπιστίας που προκύπτει κατά την κατασκευή των διαστημάτων πρόβλεψης.

Λόγω της ικανότητάς τους να διαχειρίζονται χωρο-χρονικά δεδομένα αποτελεσματικά, τα Νευρωνικά Δίκτυα Γράφων έχουν εφαρμοστεί επιτυχώς για την πρόβλεψη της κατανάλωσης του οικιακού ηλεκτρικού φορτίου. Η διερεύνηση της χρήσης του συγκεκριμένου πεδίου έχει ξεκινήσει αρκετά πρόσφατα και έτσι η σχετική βιβλιογραφία είναι περιορισμένη. Ένα αυτο-προσαρμοζόμενο δίκτυο, που ονομάζεται Graph WaveNet, προτείνεται στην [330] ως ένας αποτελεσματικός τρόπος για την ανίχνευση χωρικών εξαρτήσεων μεταξύ διαφορετικών κατοικιών, παρουσιάζοντας σημαντική βελτίωση στην ακρίβεια των προβλέψεων σε σχέση με διάφορα μοντέλα αναφοράς (baseline models). Η πρόβλεψη του συγκεντρωμένου φορτίου (aggregated load) πραγματοποιείται με χωρο-χρονικά νευρωνικά δίκτυα γράφων στη [331]. Το προτεινόμενο μοντέλο GNN ενσωματώνει ένα σύνολο από εξατομικευμένους γράφους όπου κάθε κόμβος αντιστοιχεί στο συγκεντρωμένο φορτίο μιας κλάσης έξυπνων μετρητών και έναν κεντρικό γράφο που αναπαριστά την χρονοσειρά του συνολικού φορτίου. Ένα παρόμοιο μοντέλο για την πρόβλεψη της κατανάλωσης της ηλεκτρικής ενέργειας από νοικοκυριά, το οποίο δεν απαιτεί γεωγραφικές ή περιβαλλοντικές μεταβλητές υλοποιείται στην [332], συνδυάζοντας συνελικτικά δίκτυα γράφων (Graph Convolutional Networks - GCN) και Δίκτυα Μακράς Βραχύχρονης Μνήμης (Long Short-Term Memory - LSTM). Το πρόβλημα της ανεπαρκούς ποσότητας δεδομένων εκπαίδευσης μοντέλων GNN για την βραχυπρόθεσμη πρόβλεψη του ηλεκτρικού φορτίου αντιμετωπίζεται στην αναφορά [333] μέσω μιας προσέγγισης μάθησης προσθετικής μεταφοράς (attentive transfer learning). Η προτεινόμενη μεθοδολογία μπορεί δυναμικά να αναθέτει βάρη σε διαφορετικά μοντέλα GNN ανάλογα με την επάρκεια των δεδομένων εισόδου.

Είναι προφανές από τη βιβλιογραφία, ότι αν και το πρόβλημα της πρόβλεψης του ηλεκτρικού φορτίου έχει αντιμετωπιστεί με πολλαπλές τεχνικές μηχανικής μάθησης, εντούτοις καμια από αυτές δεν επιτυγχάνει καθολική υπεροχή σε ότι αφορά την ακρίβεια των προβλέψεων. Η παρατήρηση αυτή επιβεβαιώνεται μελετώντας τα ερευνητικά αποτελέσματα, αλλά και αξιολογώντας συγκρίσεις αναφοράς από τη βιβλιογραφία [334]–[336]. Η αδυναμία εξαγωγής καθολικά ανώτερης επίδοσης από τις μεθόδους της βιβλιογραφικής ανασκόπησης που προηγήθηκε είναι αναμενόμενη αν λάβει κανείς υπόψη τα ανεπιθύμητα χαρακτηριστικά του προβλήματος της πρόβλεψης του φορτίου, το οποίο περιλαμβάνει μη γραμμικότητες και υψηλά επίπεδα θορύβου στα δεδομένα. Επιπλέον, η στατιστική της χρονοσειράς του φορτίου δεν χαρακτηρίζεται από στατικότητα [337], εξαιτίας της τυχαιότητας και της μεταβλητότητας που χαρακτηρίζουν τις καιρικές συνθήκες που επηρεάζουν την παραγωγή ενέργειας. Διαφορετικές κλάσεις μεθόδων μηχανικής μάθησης ενδέχεται να διαχειρίζονται επαρκώς ορισμένα από αυτά τα ζητήματα, όμως συνήθως αποτυγχάνουν σε σχέση με την αντιμετώπιση κάποιων άλλων, π.χ., τα γραμμικά μοντέλα είναι πιο εύρωστα στο θόρυβο, αλλά δεν μπορούν χειριστούν τις μη γραμμικότητες. Παρόλο που οι προαναφερθείσες προκλήσεις είναι εγγενείς στην πρόβλεψη του φορτίου, εξαρτώνται άμεσα από τον χρονικό ορίζοντα πρόβλεψης, καθιστώντας αδύνατο να προσδιορίσει κανείς μία και μόνο μέθοδο μηχανικής μάθησης που να υπερτερεί έναντι των υπολοίπων για διαφορετικούς ορίζοντες πρόβλεψης. Για παράδειγμα, τα γραμμικά μοντέλα συνήθως παρουσιάζουν αξιόπιστες προβλέψεις για βραχυπρόθεσμους ορίζοντες, ενώ η απόδοσή τους μειώνεται καθώς ο ορίζοντας πρόβλεψης αυξάνεται λόγω της επικράτησης μη γραμμικών συσχετίσεων στα δεδομένα. Θα πρέπει να σημειωθεί ότι η προηγούμενη παρατήρηση σχετικά με την αδυναμία μιας μεμονωμένης μεθόδου να επικρατήσει έναντι όλων των άλλων δεν αφορά μόνο στο πλαίσιο της πρόβλεψης φορτίου, αλλά αποτελεί μια πιο γενική έννοια στη μηχανική μάθηση και τη βελτιστοποίηση, όπως εκφράζεται από το θεώρημα “no free lunch” [338].

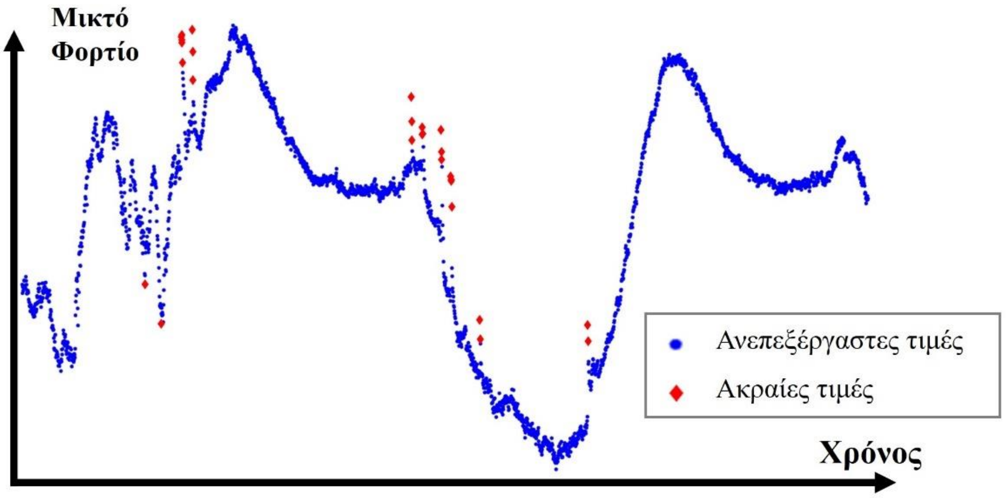
Για να αντιμετωπιστεί αυτό το πρόβλημα, θα μπορούσε κανείς να καταφύγει στη χρήση μιας προσέγγισης πολλαπλών μοντέλων [339], συνδυάζοντας διαφορετικές μεθοδολογίες μηχανικής μάθησης. Δυστυχώς, σε ένα σενάριο υλοποίησης σε πραγματικό χρόνο, προκύπτει μια σημαντική πρακτική δυσκολία για τις προσεγγίσεις πολλαπλών μοντέλων: πώς επιλέγει κανείς το καταλληλότερο μοντέλο από μια ομάδα εκπαιδευμένων μοντέλων για το επόμενο χρονικό βήμα πρόβλεψης; Μια λύση είναι η αξιοποίηση ενός συστήματος αποφάσεων βασισμένου σε κανόνες που χρησιμοποιεί πρότερη διαθέσιμη γνώση, όπως η ώρα της ημέρας και οι μετρήσεις που αφορούν τις καιρικές συνθήκες σε επίπεδο υποσταθμού. Ωστόσο, η τεχνική αυτή συνοδεύεται από ένα σημαντικό εμπόδιο, καθώς όχι μόνο οι κανόνες ενός τέτοιου συστήματος είναι δύσκολο να κατανοηθούν, αλλά δεν προσφέρουν επίσης καμία εγγύηση για τη διαρκή βέλτιστη επιλογή μοντέλου. Από την άλλη, η πλήρης απόρριψη ενός συστήματος αποφάσεων είναι επίσης κάτι που δεν συστήνεται, καθώς οι μεμονωμένες προβλέψεις από διαφορετικές μεθόδους δεν προσφέρουν από μόνες τους καμία αξιοποιήσιμη πληροφορία. Πιο συγκεκριμένα, ένας επαγγελματίας, είτε πρόκειται για έναν τεχνικό διαχειριστή του συστήματος διανομής είτε για φορέα σωρευτικής εκπροσώπησης (ΦοΣΕ) ΑΠΕ, χρειάζεται μια τελική τιμή πρόβλεψης για να αναπτύξει τη στρατηγική λειτουργίας του. Μια πρακτική λύση είναι να απορρίψει κανείς τέτοιους κανόνες επιλογής και αντ 'αυτού να χρησιμοποιήσει ένα σύστημα στάθμισης που αξιολογεί μοντέλα μόνο χρησιμοποιώντας τις προηγούμενες επιδόσεις πρόβλεψής τους [340]. Η στάθμιση των αποτελεσμάτων εξόδου των μοντέλων πρόβλεψης LSTM που περιγράφεται στην [341] βασίζεται στον βαθμό ομοιότητας μεταξύ των πραγματικών και αναγνωρισμένων τυπικών τιμών κατανάλωσης φορτίου. Δύο διαφορετικές προσεγγίσεις για τον προσδιορισμό των βαρών μοντέλων πολλαπλών προβλέψεων ακολουθούνται στις αναφορές [342], [343], χρησιμοποιώντας την καινοτόμα στρατηγική της σταδιακά αυξητικής ενημέρωσης των βαρών. Εναλλακτικά, μια μηχανή ακραίας εκμάθησης (extreme learning machine) μπορεί να χρησιμοποιηθεί για το συνδυασμό των αποτελεσμάτων μιας δεξαμενής προβλέψεων όπως αυτό περιγράφεται στη [344]. Ένα ευφυές σχήμα υποστήριξης λήψης αποφάσεων που περιλαμβάνει την προγνωστική αξιολόγηση απόδοσης, την ανάλυση ιδιοτήτων μοντέλου, την βελτιστοποίηση της δομής και της επιλογής μοντέλου, παρουσιάζεται με μια εξελικτική μέθοδο εκμάθησης πολλαπλών μοντέλων για την βραχυπρόθεσμη πρόβλεψη του ηλεκτρικού φορτίου [345]. Τέλος, το αυτοματοποιημένο σύστημα που περιγράφεται στη [346], βασίζεται σε κρυφές Μαρκοβιανές αλυσίδες για την εξαγωγή παρόμοιων ημερήσιων μοτίβων κατανάλωσης ώστε να αποκτηθεί το καλύτερο μοντέλο από μια βιβλιοθήκη διαθέσιμων μοντέλων πρόβλεψης. Τα ακόλουθα μοντέλα νευρωνικών δικτύων ακολουθούν τεχνικές που διαφοροποιούνται από τις προηγούμενες εργασίες, και προκύπτουν από πολλαπλούς κύκλους εκπαίδευσης που βασίζονται σε στιγμιότυπα [347] ή στα κρυφά χαρακτηριστικά ενός δικτύου Random Vector Functional Link [348].

3.3 Ανάπτυξη μεθόδων πρόβλεψης ηλεκτρικού φορτίου

3.3.1 Μη επιβλεπόμενες μέθοδοι: Αραιή αναπαράσταση και εκμάθηση λεξικού

Τα δεδομένα της μελέτης περίπτωσης έχουν συλλεχθεί από έναν υποσταθμό ΥΤ/ΜΤ στην ηπειρωτική Ευρώπη και εκτείνονται στο χρονικό διάστημα Σεπτέμβριος 2017-Δεκέμβριος 2018. Η ακριβής προέλευση και η χρονολογία των δεδομένων δεν είναι δημόσια διαθέσιμες για λόγους εμπιστευτικότητας και απορρήτου. Για τον ίδιο λόγο, τα αποτελέσματα που προκύπτουν και παρουσιάζονται στο επόμενο κεφάλαιο περιλαμβάνουν τις κανονικοποιημένες τιμές των προβλέψεων. Οι καταγεγραμμένες μετρήσεις αναφέρονται σε τιμές μικτού φορτίου παραγωγής-ζήτησης, δεδομένης της ύπαρξης πολλαπλών φωτοβολταϊκών (ΦΒ) συστημάτων, οικιακά ή μη, τα οποία συμμετέχουν στο δίκτυο διανομής και παραγωγής ηλεκτρικής ενέργειας ΜΤ. Οι εν λόγω μετρήσεις λαμβάνονται ανά διαστήματα του ενός λεπτού και περιέχουν αφενός τις τιμές μικτού φορτίου, καθώς και δεδομένα καιρού από τον μετεωρολογικό σταθμό του υποσταθμού. Συγκεκριμένα παρέχονται τιμές νεφοκάλυψης, ταχύτητας ανέμου, υγρασίας και θερμοκρασίας. Σημειώνεται ότι, οι χρονοσειρές δεδομένων έχουν υποβληθεί σε αναδειγματοληψία εφαρμόζοντας μια μέθοδο παρεμβολής, όπου κάθε δείγμα ορίζεται ως ο μέσος όρος ενός κινούμενου παραθύρου 15 μετρήσεων, εάν τουλάχιστον 8 από αυτές είναι έγκυρες και ως απροσδιόριστη τιμή (null), διαφορετικά. Στη συνέχεια ακολουθεί η κατάλληλη προεπεξεργασία των ανωτέρω μετρήσεων προκειμένου να δημιουργηθεί ένα μοντέλο πρόβλεψης του μικτού φορτίου του δικτύου, για ένα σύντομο ορίζοντα πρόβλεψης όπως αυτός της μίας ώρας.

Ένα πρωταρχικό ζήτημα κατά την προεπεξεργασία των δεδομένων είναι η αντιμετώπιση δεδομένων που λείπουν ή είναι κατεστραμμένα, δηλαδή είναι αδύνατη η ανάγνωση ή η χρήση τους. Τέτοιου είδους δεδομένα μπορούν να προκύψουν είτε από διακοπές της λειτουργίας των αισθητήρων, κυρίως λόγω δυσλειτουργίας, είτε λόγω εργασιών συντήρησης, είτε εξαιτίας κάποιου κακόβουλου λογισμικού. Στο πλαίσιο της εκπόνησης της διδακτορικής διατριβής, το στάδιο της προεπεξεργασίας των δεδομένων, αναλαμβάνει την αφαίρεση των κατεστραμμένων δεδομένων και ακραίων τιμών, ενώ τα δεδομένα που λείπουν αγνοούνται. Λόγω του μεγέθους του συνόλου των διαθέσιμων δεδομένων, η χειροκίνητη εξαγωγή των προβληατικών δεδομένων καθίσταται εξαιρετικά δύσκολη αν όχι αδύνατη κι ετσι προτιμήθηκε η δημιουργία μιας ρουτίνας αυτόματης ανίχνευσης των δεδομένων αυτών. Σε αντίθεση με τις τιμές κατεστραμμένων δεδομένων, οι οποίες υποδεικνύονται από τις εκτός των επιτρεπόμενων ορίων τιμές του μικτού φορτίου ή από την αδυναμία προσπέλασής τους, ο εντοπισμός των ακραίων τιμών αποτελεί μια δυσκολότερη επιστημονική πρόκληση. Μια σειρά από αποτελεσματικές τεχνικές διαχείρισης δεδομένων μπορεί να βρει κανείς στην αναφορά [349]. Για τις ανάγκες των πειραμάτων που πραγματοποιήθηκαν στα πλαίσια της Δ.Δ, εφαρμόζεται μια προσέγγιση όπου η ανίχνευση βασίζεται στην υπέρβαση μιας ανώτερης τιμής της διαμέσου τιμής επί του συνόλου των τιμών του μικτού φορτίου που περιέχονται σε ένα κυλιόμενο χρονικό παράθυρο. Έτσι επιτυγχάνεται η αποτελεσματική αναγνώριση των ακραίων τιμών, αποφεύγοντας παράλληλα την λανθασμένη κατηγοριοποίηση φυσιολογικών μετρήσεων ως ακραίες τιμές Σχήμα 11. Οι τιμές του μήκους του κυλιόμενου παραθύρου και του κατωφλίου της διαμέσου επιλέγονται ως 60 και 0.15 αντίστοιχα.



Σχήμα 11: Λειτουργία της ρουτίνας ανίχνευσης ακραίων τιμών που βασίζεται στο κατώφλι κυλιόμενης διαμέσου.

Ένα σημαντικό στάδιο στη δημιουργία ενός μοντέλου πρόβλεψης του ηλεκτρικού φορτίου αποτελεί αυτό της επιλογής των μεταβλητών εισόδου. Όλα τα μοντέλα που αναπτύχθηκαν στο πλαίσιο αυτής της μελέτης θεωρούνται ως αυτοπαλινδρομικά με εξωγενείς μεταβλητές με την έννοια ότι οι είσοδοί τους αποτελούνται αποκλειστικά από προηγούμενες τιμές της εξόδου και τιμές των καιρικών δεδομένων. Δεδομένου ότι ο στόχος είναι η δημιουργία μοντέλων πρόβλεψης βραχυπρόθεσμου ορίζοντα, η χρήση ιστορικών δεδομένων τόσο ηλεκτρικού φορτίου όσο και δεδομένων καιρού μπορεί να θεωρηθεί επαρκής. Παρακάτω παρατίθεται, για την πληρέστερη περιγραφή των μεταβλητών εισόδου, οι ακόλουθοι συμβολισμοί. Έτσο, ένα d-διάστατο διανύσμα-στήλη δηλώνεται με πεζά έντονα γράμματα (δλδ., $\boldsymbol{x}$ ή $\boldsymbol{x}^{(k)} \in R^d$) και ο όρος $[\cdot]_i$ χρησιμοποιείται για το συμβολισμό των δεικτών του στοιχείου ενός διανύσματος-στήλης. Οι τιμές των συνιστωσών του διανύσματος μετρήσεων $\boldsymbol{x}^{(k)}$ περιέχουν πραγματικούς αριθμούς οι οποίοι απεικονίζουν τις τρέχουσες και ιστορικές τιμές του μικτού φορτίου $\boldsymbol{p}$ και τις καιρικές μεταβλητές $\boldsymbol{w}$. Κατά συνέπεια, ως $\boldsymbol{p}^{(t)}$ και $\boldsymbol{w}^{(t)}$ δηλώνονται οι τρέχουσες και ιστορικές τιμές του μικτού φορτίου καθώς και οι μετρήσεις των καιρικών μεταβλητών, καταχωρημένες την τρέχουσα χρονική στιγμή *t*. Αξίζει να διευκρινιστεί ότι οι δείκτες t+*4* και *t-4* αναφέρονται σε διάστημα μιας ώρας μετά και πριν από την τρέχουσα χρονική στιγμή αντίστοιχα, η τιμή του d ισούται με το συνολικό αριθμό των μεταβλητών των μοντέλων και η τιμή του N ισούται με το πλήθος των δεδομένων εκπαίδευσης.

Οι επιλεχθείσες μεταβλητές εισόδου, οι οποίες είναι κοινές για όλα τα μοντέλα που υλοποιούνται στα πλαίσια της μελέτης περίπτωσης, περιγράφονται αναλυτικά στον Πίν. Τα στοιχεία του διανύσματος $[\boldsymbol{p}^{(t)}]_n$, $n \in [1{:}\,5]$ περιέχουν τις τιμές της n-οστής μεταβλητής εισόδου που σχετίζεται με το μικτό φορτίο, ενώ τα στοιχεία του διανύσματος $[\boldsymbol{w}^{(t)}]_m$, $m \in [1{:}\,4]$ περιέχουν τις τιμές της m-οστής καιρικής μεταβλητής εισόδου. Οι συμβολισμοί $[c^{(t)}]_1$, $[s^{(t)}]_1$, $[h^{(t)}]_1$ και $[t^{(t)}]_1$ χρησιμοποιούνται για τη δήλωση των τρέχοντων τιμών της νεφοκάλυψης, της ταχύτητας ανέμου, της υγρασίας και της θερμοκρασίας αντίστοιχα. Προκειμένου να αποκτηθούν οι προβλέψεις του μικτού φορτίου της επόμενης ώρας, οι οποίες δηλώνονται ως $[\boldsymbol{p}^{(t+4)}]_1$, οι είσοδοι που φαίνονται στον πίνακα Πίνακας 3, εισάγονται σε κάθε μοντέλο.

Πίνακας 3: Μεταβλητές εισόδου των μοντέλων πρόβλεψης αραιής αναπαράστασης

Αύξοντας αριθμός μεταβλητής εισόδου	Διάνυσμα μεταβλητής εισόδου	Τύπος μεταβλητής εισόδου	Μαθηματική έκφραση	Περιγραφή
1	$[\boldsymbol{x}^{(t)}]_9$	$[\boldsymbol{p}^{(t)}]_1$	$[p^{(t)}]_1$	Τρέχουσα τιμή Ε.Ι.
2	$[\boldsymbol{x}^{(t)}]_8$	$[\boldsymbol{p}^{(t)}]_2$	$\frac{\sum_{n=0}^{3}[p^{(t-n)}]_1}{4}$	Μέση τιμή Ε.Ι. για την τελευταία ώρα
3	$[\boldsymbol{x}^{(t)}]_7$	$[\boldsymbol{p}^{(t)}]_3$	$[p^{(t)}]_1 - [p^{(t-1)}]_1$	Διαφορά μεταξύ τρέχουσας τιμής Ε.Ι. και τιμής πριν 15 λεπτά
4	$[\boldsymbol{x}^{(t)}]_6$	$[\boldsymbol{p}^{(t)}]_4$	$[p^{(t)}]_1 - [p^{(t-4)}]_1$	Διαφορά μεταξύ τρέχουσας τιμής Ε.Ι. και τιμής πριν 1 ώρα
5	$[\boldsymbol{x}^{(t)}]_5$	$[\boldsymbol{p}^{(t)}]_5$	$[p^{(t)}]_1 - [p^{(t-8)}]_1$	Διαφορά μεταξύ τρέχουσας τιμής Ε.Ι. και τιμής πριν 2 ώρες
6	$[\boldsymbol{x}^{(t)}]_4$	$[\boldsymbol{w}^{(t)}]_1$	$[c^{(t+4)}]_1 - [c^{(t)}]_1$	Διαφορά μεταξύ τρέχουσας τιμής νεφοκάλυψης και τιμής μετά από 1 ώρα
7	$[\boldsymbol{x}^{(t)}]_3$	$[\boldsymbol{w}^{(t)}]_2$	$[s^{(t+4)}]_1 - [s^{(t)}]_1$	Διαφορά μεταξύ τρέχουσας τιμής ταχύτητας ανέμου και τιμής μετά από 1 ώρα
8	$[\boldsymbol{x}^{(t)}]_2$	$[\boldsymbol{w}^{(t)}]_3$	$[h^{(t+4)}]_1 - [h^{(t)}]_1$	Διαφορά μεταξύ τρέχουσας τιμής υγρασίας και τιμής μετά από 1 ώρα
9	$[\boldsymbol{x}^{(t)}]_1$	$[\boldsymbol{w}^{(t)}]_4$	$[t^{(t+4)}]_1 - [t^{(t)}]_1$	Διαφορά μεταξύ τρέχουσας τιμής θερμοκρασίας και τιμής μετά από 1 ώρα

Η χρήση των μεταβλητών αυτών κρίνεται δικαιολογημένη καθώς υπάρχει ισχυρή εξάρτηση της χρονοσειράς του ηλεκτρικού φορτίου από την τρέχουσα τιμή [316]. Μια συνιθισμένη τακτική [114] αποτελεί επίσης η αξιοποίηση της εποχικής τάσης του ηλεκτρικού φορτίου, όπως αυτή αντικατοπτρίζεται στις διαφορές μεταξύ της τρέχουσας και προηγούμενων τιμών του μικτού φορτίου σύμφωνα με τους όρους $[\boldsymbol{p}^{(t)}]_{3:5}$, ενώ η εισαγωγή του μέσου όρου των προηγούμενων τιμών, $[\boldsymbol{p}^{(t)}]_2$ σύμφωνα με τη βιβλιογραφία, είναι εξίσου σημαντική [317]. Τέλος, η ενσωμάτωση των τιμών που απεικονίζουν το διάνυσμα των καιρικών χαρακτηριστικών, $[\boldsymbol{w}^{(t)}]_{1:4}$ αναμένεται να βελτιώσει την ακρίβεια των προβλέψεων με βάση τα διάφορα πειραματικά αποτελέσματα [291], [292]. Οι τιμές των μεταβλητών αυτών αναμένεται να αποτελέσουν το σύνολο δεδομένων εκπαίδευσης με βάση το οποίο θα γίνει η εκμάθηση του λεξικού και των αραιών συντελεστών.

Ολοκληρώνοντας την προεπεξεργασία των δεδομένων και την επιλογή των μεταβλητών εισόδου, ολοκληρώνουμε την προετοιμασία του συνόλου δεδομένων εκπαίδευσης. Μετά από προσεκτική εξέταση των διαθέσιμων δεδομένων, κρίθηκε ότι τα δεδομένα που καλύπτουν το διάστημα του φθινοπώρου του 2017 είναι τα καταλληλότερα για την εκπαίδευση των μοντέλων, αφήνοντας διαθέσιμα τα υπόλοιπα δεδομένα, δηλαδή εκείνα που εκτείνονται από τον χειμώνα του 2017 έως τον χειμώνα του 2018 για τη διαδικασία του ελέγχου (testing) των μοντέλων. Η επιλογή των δεδομένων εκπαίδευσης βασίστηκε στο γεγονός ότι τα δεδομένα σε αυτό το διάστημα παρουσίαζαν τον ελάχιστο δυνατό αριθμό ελλειπουσών και ακραίων τιμών σε σύγκριση με άλλα διαστήματα. Επιπρόσθετα, αυτός ο διαχωρισμός των δεδομένων επιτρέπει τη διατήρηση ενός ολόκληρου έτους διαθέσιμου ως σύνολο δεδομένων ελέγχου, κάτι που βοηθά στην καλύτερη αξιολόγηση των γενικευμένων προγνωστικών ικανοτήτων του μοντέλου σε διαφορετικές εποχές του έτους. Αυτό είναι ένα σημαντικό ζήτημα, καθώς η χρονοσειρά αποτελείται από μια συνιστώσα της ζητούμενης και μια συνιστώσα της παραγόμενης ενέργειας, οι στατιστικές ιδιότητες των οποίων δεν είναι στατικές κατά τη διάρκεια του έτους, κυρίως λόγω της μεταβλητότητας του καιρού. Τέλος, ένα σημείο που χρήζει ιδιαίτερης προσοχής είναι ότι δεν εφαρμόζεται μετάθεση (permutation) των δεδομένων πριν από τον διαχωρισμό των διαθέσιμων δεδομένων στα σύνολα δεδομένων εκπαίδευσης και ελέγχου. Έτσι, τα δεδομένα που χρησιμοποιούνται κατά τη διαδικασία ελέγχου είναι εντελώς άγνωστα στο εκάστοτε προτεινόμενο μοντέλο, ενισχύοντας την αξιοπιστία των προβλέψεων.

Η ανάλυση των προδιαγραφών με βάση τις οποίες διεξάγεται η παρούσα μελέτη περίπτωσης, ολοκληρώνεται με την περιγραφή του πειραματικού πρωτοκόλλου. Ο πρώτος κατά σειρά αλγόριθμος αραιής κωδικοποίησης που εφαρμόστηκε ήταν ο K-SVD. Όπως έχει ήδη αναφερθεί, ο K-SVD αποτελεί μια επαναληπτική μέθοδο που εναλλάσσει τη λειτουργία του μεταξύ της αραιής κωδικοποίησης του διανύσματος ενδιαφέροντος με βάση το τρέχον λεξικό και μιας διαδικασίας ανανέωσης των ατόμων του λεξικού για την καλύτερη προσαρμογή στα δεδομένα εκπαίδευσης. Τηρώντας την ιδιότητα του υπερπλήρους λεξικού, ο αριθμός των ατόμων τίθεται ίσος με 60, δηλαδή 6 φορές μεγαλύτερος από τη διάσταση του χώρου των μεταβλητών εισόδου. Μία επιπλέον παράμετρος που θα πρέπει να προσδιοριστεί, είναι ο αριθμός των μη μηδενικών στοιχείων που περιέχονται στα άτομα, ή αλλιώς το επίπεδο αραιότητας. Με βάση τον κανόνα του αντίχειρα (rule of thumb), η τιμή αυτή, δεν ξεπερνά το 10% του πλήθους των ατόμων και κατά συνέπεια, επιλέγεται να είναι ίση με 6. Εφόσον, ο αλγόριθμος K-SVD χαρακτηρίζεται από ευελιξία σε σχέση με τη συμβατότητά του με την τεχνική αραιής αποδόμησης που θα χρησιμοποιηθεί, επιλέγεται η μέθοδος OMP. Τέλος, η ανανέωση του λεξικού πραγματοποιείται ανά μέγιστο αριθμό 100 επαναλήψεων.

Ο online αλγόριθμος εκμάθησης λεξικού που περιγράφεται στην [350], εφαρμόζεται στη συνέχεια σε συνδυασμό με η μέθοδο LARS-LASSO για να επιτευχθεί η αραιή αποδόμηση. Στα πλαίσια της προσέγγισης αυτής, η αραιότητα επιβάλλεται μέσω της νόρμας ℓ_1, σύμφωνα με την σχέση (2.22). Η παράμετρος κανονικοποίησης λ τίθεται ίση με 0.1. Για λόγους πληρότητας παρέχονται επίσης αποτελέσματα με τη χρήση της κανονικοποίησης elastic net [351], η οποία συνδυάζει τις νόρμες ℓ_1 και ℓ_2. Στην περίπτωση αυτή, οι συντελεστές κανονικοποίησης λ_1 και λ_2 επιλέγεται να είναι ίσοι με 0.1 και 0.01 αντίστοιχα. Ένα επιπλέον μοντέλο υλοποιείται, το οποίο χρησιμοποιεί τον online αλγόριθμο για την εκμάθηση του λεξικού και τη μέθοδο OMP για το στάδιο της αραιής κωδικοποίησης [352] αντί της LASSO. Όμοια με την προηγούμενη υλοποίηση η τιμή του λ επιλέγεται να είναι ίση με 0.1, ακολουθώντας μια διαδικασία δοκιμής και σφάλματος. Προκειμένου να εξασφαλιστεί η δίκαιη σύγκριση μεταξύ των μοντέλων, το μέγεθος του λεξικού παραμένει ίδιο για όλα τα πειράματα που λαμβάνουν χώρα στα πλαίσια της μελέτης περίπτωσης. Ένας ακόμη παράγοντας που επηρεάζει την απόδοση του μοντέλου είναι η μέγιστη διάρκεια εκτέλεσης του αλγορίθμου, η οποία ισούται με 20 δευτερόλεπτα, καθώς δεν παρατηρείται περαιτέρω βελτίωση πέραν του χρόνου αυτού.

Η επόμενη μέθοδος που αναπτύχθηκε ενσωματώνει την αρχετυπική ανάλυση. Σύμφωνα με την προσέγγιση αυτή, η εκμάθηση του λεξικού συνθέτει το μη κυρτό πρόβλημα βελτιστοποίησης όπως εκφράζεται από τις σχέσεις (2.29) και (2.30) και το οποίο επιλύεται με τον αλγόριθμο ενεργών συνόλων που προτείνεται στο άρθρο [244]. Η νόρμα ℓ_2 επιστρατεύεται για την επιβολή της αραιότητας τοποθετώντας τα άτομα στην γειτονιά του κυρτού φλοιού του χώρου των χαρακτηριστικών.

Μέσω της παρούσας μελέτης περίπτωσης, εκτός από την ανάδειξη της καταλληλότερης μεθοδολογίας αραιής κωδικοποίησης για το πρόβλημά της βραχυπρόθεσμης πρόβλεψης του μικτού φορτίου, εξετάζεται επίσης η επίδραση των διάφορων νορμών νομιμοποίησης. Προς αυτή την κατεύθυνση, υλοποιείται μια επιπλέον στρατηγική, αυτή της ιεραρχικής αραιότητας η οποία υπαγορεύει την ενσωμάτωση των ατόμων του λεξικού σε ένα κατευθυνόμενο δένδρο τόσο κατά την εκμάθηση του λεξικού όσο και κατά το στάδιο της αραιής αναπαράστασης. Στα πλαίσια των πειραματικών προσομοιώσεων του μοντέλου που βασίζεται στην ιεραρχική αραιότητα, χρησιμοποιείται μια νόρμα με δενδροειδή δομή ℓ_0^{tree}, όπου ο αριθμός των ατόμων καθορίζεται από τη διαμόρφωση του δένδρου. Η δομή διακλάδωσης που επιλέγεται, προσδιορίζεται από ένα δένδρο 3 επιπέδων με 4 κλάδους σε κάθε ένα από τα δύο πρώτα επίπεδα και 2 κλάδους στο τελευταίο επίπεδο. Η διαμόρφωση αυτή οδηγεί σε ένα λεξικό με διάσταση 53, γεγονός που εξασφαλίζει μέγεθος λεξικού συγκρίσιμο με τις υπόλοιπες μεθόδους καθώς και εύλογο χρόνο εκτέλεσης. Μέσα από μια διαδικασία δοκιμής και σφάλματος η παράμετρος λ τέθηκε ίση με 0.01.

Προκειμένου να τονιστεί η υπεροχή της προτεινόμενης μεθοδολογίας αραιής κωδικοποίησης, δοκιμάζονται επίσης τυπικές γραμμικές ή μη γραμμικές μέθοδοι μηχανικής μάθησης, συμπεριλαμβανομένης της απλής γραμμικής παλινδρόμησης, της παλινδρόμησης με χρήση διανυσμάτων υποστήριξης και των νευρωνικών δικτύων MLP. Όλα τα συγκριτικά μοντέλα χρησιμοποιούν τις ίδιες μεταβλητές εισόδου που περιγράφονται παραπάνω, αλλά και κοινά σύνολα δεδομένων εκπαίδευσης και ελέγχου. Πιθανώς η πιο κοινή και απλή μέθοδος πρόβλεψης είναι η γραμμική παλινδρόμηση [108]. Έτσι, ένα μοντέλο παλινδρόμησης ελαχίστων τετραγώνων δημιουργήθηκε και χρησιμοποιήθηκε ως μοντέλο αναφοράς. Στη συνέχεια, εφαρμόστηκε ένα μοντέλο SVR με Γκαουσιανή συνάρτηση πυρήνα [162]. Η εκπαίδευση του μοντέλου πραγματοποιήθηκε με τον αλγόριθμο SMO, ενώ για τη βελτιστοποίηση των υπερπαραμέτρων του χρησιμοποιήθηκε η Μπεϋζιανή βελτιστοποίηση. Η πιο αντιπροσωπευτική κλάση μη γραμμικών μεθόδων για την βραχυπρόθεσμη πρόβλεψη του ηλεκτρικού φορτίου είναι τα νευρωνικά δίκτυα MLP [353]. Έτσι, το σύνολο των συγκριτικών μοντέλων που αναπτύσσονται ολοκληρώνεται με την εισαγωγή ενός δικτύου MLP με δύο στοιβάδες το οποίο εκπαιδεύεται με τον αλγόριθμο οπισθοδιάδοσης Levenberg-Marquardt [165]. Θα πρέπει να σημειωθεί ότι, λόγω της ευαισθησίας των μεθόδων εκπαίδευσης των δικτύων MLP στην αρχικοποίηση των βαρών, η διαδικασία εκπαίδευσης επαναλήφθηκε 20 φορές, χρησιμοποιώντας κάθε φορά διαφορετική τυχαία προετοιμασία για τα βάρη του νευρωνικού δικτύου.

3.3.2 Επιβλεπόμενες μέθοδοι: Μέθοδος ανσάμπλ

Η πρώτη μέθοδος που αναπτύχθηκε στα πλαίσια της της παρούσας Δ.Δ. ακολουθεί τις αρχές της επιβλεπόμενης μάθησης. Συγκεκριμένα, βασίζεται στη μέθοδο ανσάμπλ και είναι σε θέση να παρέχει αξιόπιστες προβλέψεις, για πολλαπλούς ορίζοντες πρόβλεψης, ειδικότερα δε για χρονικά παράθυρα των 15 λεπτών, 1 ώρας και 2, 3, 6 έως και 24 ωρών. Για την αξιολόγηση την απόδοσης της προτεινόμενης μεθόδου δημιουργήθηκαν μια σειρά από πειράματα τα οποία χρησιμοποιούν τα δεδομένα που αναφέρθηκαν στην παράγραφο 3.3.1, με τη διαφορά ότι στην περίπτωση αυτή τα δεδομένα ολόκληρου του έτους 2017 είναι διαθέσιμα. Οι συνθήκες αναδειγματοληψίας καθώς και η διαδικασία προεπεξεργασίας των δεδομένων συμπεριλαμβανομένου του αλγορίθμου ανίχνευσης ακραίων τιμών, παραμένουν ίδιες με αυτές που περιγράφονται στην προηγούμενη μελέτη περίπτωσης.

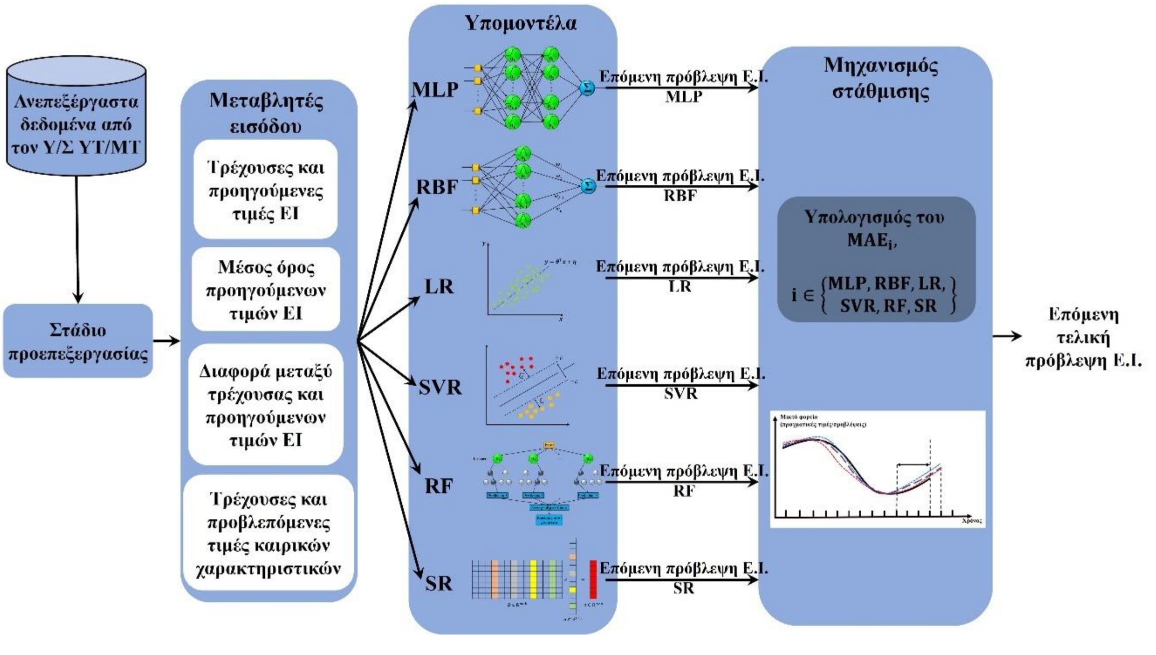
Η επιλογή των μεταβλητών εισόδου συνδέεται άμεσα με τον ορίζοντα πρόβλεψης. Για το λόγο αυτό και με βάση τη βιβλιογραφία, ένα σύνολο χαρακτηριστικών διανυσμάτων - εισόδων δημιουργήθηκε για κάθε ορίζοντα πρόβλεψης, αποτελούμενο από ιστορικά δεδομένα του φορτίου και δεδομένα καιρού. Στη συνέχεια, η συνεισφορά των μεταβλητών αυτών στην βελτίωση την ακρίβειας των προβλέψεων εξετάστηκε με την μέθοδο της δοκιμής και του σφάλματος, οδηγώντας στην τελικά επιλογή των μεταβλητών εισόδου του μοντέλου για κάθε ορίζοντα πρόβλεψης. Εναλλακτικά, άλλες προσεγγίσεις όπως τα δένδρα αποφάσεων ενίσχυσης βάθμωσης (gradient boosting decision tree), ο συντελεστής συσχέτισης Pearson [354], ο μηχανισμός προσοχής (attention mechanism) [355] ή η διερευνητική ανάλυση δεδομένων (Exploratory Data Analysis) [356], θεωρούνται επίσης χρήσιμες για τη μείωση και επιλογή των χαρακτηριστικών εισόδου. Είναι σημαντικό να σημειωθεί ότι, για κάθε ορίζοντα, οι μεταβλητές εισόδου είναι κοινές για όλες τις μεθόδους μηχανικής μάθησης που εξετάζονται στην παρούσα Δ.Δ. Οι μεταβλητές που δέχονται όλα τα μοντέλα μπορούν να χωριστούν σε 4 διαφορετικές κατηγορίες, κάτι το οποίο απεικονίζεται στον πίνακα Πίνακας 4, όπου παρατηρούμε τα εξής: (α) τρέχουσες και προηγούμενες τιμές Ε.Ι., (β) διαφορά μεταξύ τρέχουσας και προηγούμενων τιμών Ε.Ι., (γ) μέσος όρος προηγούμενων τιμών Ε.Ι. και (δ) μετρήσεις καιρικών συνθηκών. Προς διευκόλυνση του αναγνώστη, αναφέρεται ότι οι τρέχουσες και προηγούμενες τιμές Ε.Ι. δηλώνονται ως $\boldsymbol{p}^{(t)}$, η μεταβλητή εξόδου, δηλαδή η τιμή του μικτού φορτίου μετά από s το πλήθος 15-λεπτα διαστήματα ως $\widehat{\boldsymbol{p}}^{(t+s)}$, ενώ οι καιρικές μεταβλητές, οι οποίες είναι η νεφοκάλυψη, η ταχύτητα ανέμου, η υγρασία και η θερμοκρασία, ως $\boldsymbol{w}^{(t)}$. Είναι σημαντικό να αναφερθεί ότι, κατά το στάδιο της εκπαίδευσης του μοντέλου πρόβλεψης, οι μεταβλητές εισόδου $\boldsymbol{w}^{(t)}$ εισάγονται ως μετρούμενες τιμές πραγματικών δεδομένων καιρού την χρονική στιγμή t. Αντιθέτως, κατά την εφαρμογή του μοντέλου σε πραγματικό χρόνο, οι μελλοντικές τιμές των καιρικών δεδομένων $\boldsymbol{w}^{(t+s)}$ όντας άγνωστες, θα πρέπει να αντικατασταθούν από καιρικές προβλέψεις, εισάγοντας επιπλέον αβεβαιότητα.

Πίνακας 4. Περιγραφή των μεταβλητών εισόδου των μοντέλων που υλοποιούνται για κάθε ορίζοντα πρόβλεψης. Κάθε γραμμή του πίνακα αναφέρεται στις διαφορετικές κατηγορίες εισόδων, ενώ η τελευταία γραμμή αναφέρεται στην μεταβλητή εξόδου.

Ορίζοντας πρόβλεψης	**15 λεπτά** $(\mathbf{t+1})$	**1 ώρα** $(\mathbf{t+4})$	**2 ώρες** $(\mathbf{t+8})$	**3 ώρες** $(\mathbf{t+12})$	**6 ώρες** $(\mathbf{t+24})$	**24 ώρες** $(\mathbf{t+96})$
Τρέχουσες και προηγούμενες τιμές Ε.Ι.	$\boldsymbol{p}^{(t-i)}$, $i = 0{,}95{,}671$	$\boldsymbol{p}^{(t-i)}$, $i = 0{,}4{,}92{,}668$	$\boldsymbol{p}^{(t-i)}$, $i = 0{,}8{,}88{,}664$	$\boldsymbol{p}^{(t-i)}$, $i = 0{,}12{,}84{,}660$	$\boldsymbol{p}^{(t-i)}$, $i = 0{,}24{,}72{,}648$	$\boldsymbol{p}^{(t-i)}$, $i = 0{,}96{,}576$
Μέσος όρος προηγούμενων τιμών Ε.Ι.	$\frac{\sum_{n=0}^{3} \boldsymbol{p}^{(t-n)}}{4}$	$\frac{\sum_{n=0}^{3} \boldsymbol{p}^{(t-n)}}{4}$	$\frac{\sum_{n=0}^{7} \boldsymbol{p}^{(t-n)}}{8}$	$\frac{\sum_{n=0}^{11} \boldsymbol{p}^{(t-n)}}{12}$	$\frac{\sum_{n=0}^{23} \boldsymbol{p}^{(t-n)}}{24}$	$\frac{\sum_{n=0}^{95} \boldsymbol{p}^{(t-n)}}{96}$
Διαφορά μεταξύ τιμών Ε.Ι.	$\boldsymbol{p}^{(t)} - \boldsymbol{p}^{(t-i)}$, $i = 1$	$\boldsymbol{p}^{(t)} - \boldsymbol{p}^{(t-i)}$, $i = 4$	$\boldsymbol{p}^{(t)} - \boldsymbol{p}^{(t-i)}$, $i = 8$	$\boldsymbol{p}^{(t)} - \boldsymbol{p}^{(t-i)}$, $i = 12$	$\boldsymbol{p}^{(t)} - \boldsymbol{p}^{(t-i)}$, $i = 24$	$\boldsymbol{p}^{(t)} - \boldsymbol{p}^{(t-i)}$, $i = 96$
Τιμές καιρικών συνθηκών	$\boldsymbol{w}^{(t+i)}$, $i = 0$	$\boldsymbol{w}^{(t+i)}$, $i = 4$	$\boldsymbol{w}^{(t+i)}$, $i = 4{,}8$	$\boldsymbol{w}^{(t+i)}$, $i = 4{,}8{,}12$	$\boldsymbol{w}^{(t+i)}$, $i = 16{,}20{,}24$	$\boldsymbol{w}^{(t+i)}$, $i = 88{,}92{,}96$
Τιμές πρόβλεψης Ε.Ι. (μεταβλητή εξόδου)	$\widehat{\boldsymbol{p}}^{(t+1)}$	$\widehat{\boldsymbol{p}}^{(t+4)}$	$\widehat{\boldsymbol{p}}^{(t+8)}$	$\widehat{\boldsymbol{p}}^{(t+12)}$	$\widehat{\boldsymbol{p}}^{(t+24)}$	$\widehat{\boldsymbol{p}}^{(t+96)}$

Κατά τη διαδικασία του διαχωρισμού των διαθέσιμων δεδομενών σε συνόλα δεδομένων εκπαίδευσης και ελέγχου θα πρέπει να ληφθεί υπόψην η εξής παρατήρηση. Το ηλεκτρικό δίκτυο επεκτείνεται συνεχώς, εντάσσοντας περισσότερους καταναλωτές και παραγωγούς ΑΠΕ, καθένας με διαφορετικά προφίλ κατανάλωσης και παραγωγής. Συνεπώς, είναι επιθυμητό το σύνολο δεδομένων εκπαίδευσης να απέχει όσο το δυνατόν λιγότερο χρονικά από το διάστημα για το οποίο θα πραγματοποιηθούν οι προβλέψεις. Εφόσον τα διαθέσιμα δεδομένα καλύπτουν δύο διαδοχικά έτη, εκείνα που αντιστοιχούν στο έτος 2017 επιλέχθηκαν ως σύνολο δεδομένων εκπαίδευσης και εκείνα που αντιστοιχούν στο έτος 2018, επιλέχθηκαν ως σύνολο δεδομένων ελέγχου.

Για λόγους πληρότητας υπογραμμίζεται το γεγονός ότι πριν από την εκπαίδευση του μοντέλου δεν έχει προηγηθεί ανακάτεμα των δεδομένων, ενώ επιπλέον οι πραγματικές τιμές και οι προβλέψεις του μικτού φορτίου υπόκεινται σε κανονικοποίηση πριν παρουσιαστούν καθώς συντρέχουν δεσμεύσεις απορρήτου. Επίσης, τα μοντέλα που κατά την εκπαίδευσή τους απαιτούν ένα στάδιο αξιολόγησης, δηλαδή τα μοντέλα που βασίζονται σε νευρωνικά δίκτυα MLP και RBF, επιστρατεύουν την 10-πλή διασταυρούμενη αξιολόγηση, ενώ στην περίπτωση των μοντέλων που απαιτούν πολλαπλές εκτελέσεις εκπαίδευσης για κάθε διαφορετική αρχικοποίηση των βαρών, όπως τα δίκτυα MLP, επιλέγεται το μοντέλο που επιτυγχάνει την υψηλότερη επίδοση κατά το στάδιο της διασταυρούμενης αξιολόγησης. Μια επισκόπηση του προτεινόμενου μοντέλου παρέχεται στο Σχήμα 12, το οποίο απεικονίζει την αλληλουχία των βημάτων που λαμβάνουν χώρα κατά την εφαρμογή του μοντέλου, ξεκινώντας από την απόκτηση των ανεπεξέργαστων δεδομένων Ε.Ι. από τον Υ/Σ έως την παραγωγή των τελικών προβλέψεων. Θα πρέπει να τονιστεί ότι πρόκειται για ένα γενικευμένο διάγραμμα το ποιο δεν αναφέρεται σε κάποιο συγκεκριμένο χρονικό ορίζοντα πρόβλεψης.



Σχήμα 12: Διάγραμμα του προτεινόμενου μοντέλου ανσάμπλ που περιγράφει την εφαρμογή του για την πρόβλεψη του μικτού φορτίου διενεργώντας τα ακόλουθα βήματα: (α) απόκτηση των ανεπεξέργαστων δεδομένων, (β) προεπεξεργασία δεδομένων, (γ) κατασκευή των μεταβλητών εισόδου και των συνόλων δεδομένων εκπαίδευσης και ελέγχου, (δ) εκπαίδευση των υπο-μοντέλων και εξαγωγή των αντίστοιχων προβλέψεων, και (ε) εφαρμογή του μηχανισμού στάθμισης των μεμονωμένων προβλέψεων των υπο-μοντέλων και υπολογισμός της τελικής πρόβλεψης της επόμενης τιμής Ε.Ι..

Προκειμένου να παρέχουμε ένα ολοκληρωμένο πλαίσιο αξιολόγησης της ακρίβειας των προβλέψεων της προτεινόμενης μεθόδου, κρίθηκε σκόπιμο να πραγματοποιηθεί η σύγκρισή της με ένα αντίστοιχο μοντέλο από τη βιβλιογραφία που ενσωματώνει πολλαπλές τοπολογίες νευρωνικών δικτύων MLP [357] και εφαρμόζεται για την πρόβλεψη του ηλεκτρικού φορτίου. Έτσι, ακολουθώντας το πειραματικό πρωτόκολλο που περιγράφεται σε αυτή την εργασία, ένα σύνολο από FFNN, με μία κρυφή στοιβάδα, εκπαιδεύεται με 14 διαφορετικές τυχαίες αρχικοποιήσεις των βαρών. Για κάθε αρχικοποίηση, το πλήθος των νευρώνων της κρυφής στοιβάδας κυμαίνεται από 3 έως 50. Η σιγμοειδής συνάρτηση υπερβολικής εφαπτομένης επιλέγεται ως συνάρτηση ενεργοποίησης των στοιβάδων των δικτύων, ενώ όλα τα επιμέρους νευρωνικά δίκτυα εκπαιδεύονται με τον αλγόριθμο ανθεκτικής οπισθοδιάδοσης. Κατά τη διαδικασία της εκπαίδευσης, τα νευρωνικά δίκτυα διατάσσονται σε αύξουσα σειρά ανάλογα με την τιμή του δείκτη MAPE που επιτυγχάνουν σε ένα κοινό σύνολο αξιολόγησης, το οποίο στην περίπτωση αυτή ορίζεται ως το 20% του συνόλου δεδομένων εκπαίδευσης. Στη συνέχεια, επιλέγονται τα δίκτυα που αντιστοιχούν στις πρώτες 5 τιμές του MAPE και οι τελικές προβλέψεις υπολογίζονται από τον μέσο όρο των επιμέρους προβλέψεων αυτών των 5 μοντέλων. Αξίζει να τονιστεί ότι τα συγκρινόμενα μοντέλα εκπαίδευονται και αξιολογούνται σε κοινά σύνολα δεδομένων εκπαίδευσης και ελέγχου αντίστοιχα.

3.3.3 Επιβλεπόμενες μέθοδοι: Νευρωνικά δίκτυα γράφων ορατότητας

Η απόδοση και η αποτελεσματικότητα της προτεινόμενης μεθοδολογίας για τη βραχυπρόθεσμη πρόβλεψη του μικτού φορτίου αξιολογείται μέσα από τη διεξαγωγή μιας μελέτης περίπτωσης στα πραγματικά δεδομένα για τα οποία αποδόθηκε περιγραφή στην παράγραφο 3.3.2. Με βάση τις πειραματικές προσομοιώσεις που προηγήθηκαν και τα αντίστοιχα αποτελέσματα, τα οποία παρατίθενται στο επόμενο κεφάλαιο, η διαδικασία της προεπεξεργασίας των δεδομένων και η ρουτίνα ανίχνευσης ακραίων τιμών που περιγράφονται σε αυτές, κρίνονται απολύτως επαρκείς και αποτελεσματικές. Για το λόγο αυτό, υιοθετούνται και εφαρμόζονται αυτούσιες κατά την εκτέλεση των ακόλουθων πειραματικών δοκιμών. Αν και η πρόθεση να χρησιμοποιηθούν όσο το δυνατόν πιο πρόσφατα δεδομένα στο πλαίσιο των πειραμάτων ήταν πρωταρχικής σημασίας, το ελάχιστο όριο των 4 ετών για τον αποχαρακτηρισμό των ιστορικών δεδομένων που ορίστηκε από τον φορέα παροχής των δεδομένων, επέβαλλε την δεδομένων δεδομένων προηγούμενων ετών. Ωστόσο, η αξιοπιστία και η ευρωστία του προτεινόμενου μοντέλου δεν τίθονται υπό αμφισβήτηση καθώς η συμπεριφορά των καταναλωτών, η διείσδυση των ΑΠΕ και η τεχνική υποδομή του δικτύου κοινής ωφέλειας δεν έχουν αλλάξει σημαντικά από τα έτη αναφοράς των συνόλων δεδομένων εκπαίδευσης και ελέγχου της μεθόδου, μέχρι σήμερα. Ο υπό εξέταση υποσταθμός ΥΤ/ΜΤ τροφοδοτεί μια αρκετά μεγάλη περιοχή με διαφορετικές κατηγορίες καταναλωτών, συμπεριλαμβανομένων βιομηχανικών και εμπορικών μονάδων καθώς και οικιακών φορτίων. Υπολογίζεται ότι εξυπηρετεί περίπου πεντακόσιες χιλιάδες πελάτες, καλύπτοντας πολλαπλές διαφορετικές συμπεριφορές ζήτησης, επομένως δύναται να αποτελέσει την κατάλληλη πηγή δεδομένων προκειμένου να εξεταστούν διάφορα πιθανά σενάρια που θα παρουσιαστούν στη συνέχεια. Θα πρέπει να σημειωθεί ότι για τον καθορισμό των συνόλων δεδομένων εκπαίδευσης και ελέγχου ακολουθείται η ίδια προσέγγιση με εκείνη που αναφέρεται στην παράγραφο 3.3.2.

Δεδομένου ότι: α) στην περίπτωση που απασχόλησε την παρούσα Δ.Δ, ο επιθυμητός ορίζοντας πρόβλεψης είναι 24 ώρες και β) συστήνεται από τη σχετική βιβλιογραφία [358]–[360], ένα σύνολο εισόδων που αποτελείται από ένα μείγμα ιστορικών τιμών μικτού φορτίου και μετεωρολογικών μετρήσεων που αναφέρονται σε διάφορες παρελθοντικές χρονικές στιγμές, κατασκευάστηκε και αναθεωρήθηκε με δοκιμή και σφάλμα ένα βέλτιστο σύνολο μεταβλητών εισόδου.Οι επιλεγμένες μεταβλητές εισόδου που τροφοδοτούν το μοντέλο VGNN μπορούν να χωριστούν σε τέσσερις κατηγορίες σύμφωνα με τον πίνακα Πίνακας 5.

Κατά την προσπάθεια ενίσχυσης της ευρωστίας του προτεινόμενου μοντέλου, διεξάγονται προσομοιώσεις που αφορούν ασυνήθιστες συνθήκες κατανάλωσης που δεν συναντώνται στην βιβλιογραφία. Σύμφωνα με τις αυτές, αντί το μοντέλο να τροφοδοτηθεί με τα δεδομένα ολόκληρου του έτους, επιλέγονται διάφορα υποσύνολα, που καλύπτουν διαφορετικές χρονικές περιόδους του έτους, και θα καλούνται εφεξής σενάρια. Ακολουθώντας τη στρατηγική αυτή λοιπόν, ορίζονται 4 διαφορετικά σενάρια, τα οποία περιγράφονται στον πίνακα Πίνακας 6. Κάθε σενάριο χρησιμοποιεί ως σύνολο δεδομένων εκπαίδευσης (α) τα δεδομένα που αντιστοιχούν στον Χειμώνα του 2017, (β) τα δεδομένα που αντιστοιχούν στην Άνοιξη του 2017, (γ) τα δεδομένα που αντιστοιχούν στο Καλοκαίρι του 2017, (δ) τα δεδομένα που αντιστοιχούν στο Φθινόπωρο του 2017, (ε) τα δεδομένα που αντιστοιχούν στις εργάσιμες ημέρες (δηλαδή ολόκληρα 24ωρα) του 2017, (στ) τα δεδομένα που αντιστοιχούν στις χρονικές περιόδους του 2017 με φως του ήλιου (δηλαδή στις ώρες από την ανατολή έως τη δύση του ήλιου για ολόκληρο το έτος) και (ζ) τα δεδομένα που αντιστοιχούν στις χρονικές περιόδους των εργάσιμων ημερών του 2017 με φως του ήλιου (δηλαδή στις ώρες από την ανατολή μέχρι τη δύση του ήλιου μόνο για τις εργάσιμες μέρες του έτους). Κατ’ αντιστοιχία, το σύνολο δεδομένων ελέγχου κάθε σεναρίου περιέχει τα ανάλογα δεδομένα του έτους 2018. Αυτά τα σενάρια αντιπροσωπεύουν ουσιαστικά περιόδους αυξημένης ζήτησης για τις οποίες η παροχή έγκυρων προβλέψεων, αν και δύσκολη, είναι ιδιαίτερα πολύτιμη τόσο για τον διαχειριστή όσο και για τους παραγωγούς. Ένα σημείο που αξίζει να αναφερθεί είναι ότι δεν πραγματοποιείται ανακάτεμα των δεδομένων πριν από την εκπαίδευση του μοντέλου. Λόγω συμφωνίας περί απορρήτου, οι πραγματικές τιμές και οι τιμές πρόβλεψης μικτού φορτίου έχουν κανονικοποιηθεί προκειμένου να παρουσιαστούν.

Πίνακας 5: Περιγραφή των μεταβλητών εισόδου του προτεινόμενου μοντέλου.

Μεταβλητή	**Μελλοντικές προβλέψεις Ε.Ι (μεταβλητή εξόδου)**	**Τρέχουσα και προηγούμενες τιμές Ε.Ι**	**Μέσος όρος προηγούμενων τιμών Ε.Ι**	**Διαφορά μεταξύ τρέχουσας και προηγούμενων τιμών Ε.Ι**	**Τιμές καιρικών συνθηκών**
Μαθηματικός τύπος	$\hat{p}(t+96)$	$p(t-i)$, $i=0{,}96{,}576$	$\frac{\sum_{i=0}^{95} p(t-i)}{96}$	$p(t)-p(t-i)$, $i=96$	$w(t+i)$, $i=88, 92, 96$
Περιγραφή	Τιμή πρόβλεψης Ε.Ι 24 ώρες μετά την τρέχουσα χρονική στιγμή t	Τιμές Ε.Ι κατά την τρέχουσα χρονική στιμή t, πριν 24 ώρες και 1 εβδομάδα πριν από την τρέχουσα χρονική στιγμή t	Μέσος όρος τιμών Ε.Ι μετρημένων κατά τη διάρκεια του τελευταίου 24ώρου	Διαφορά μεταξύ των τιμών Ε.Ι την τρέχουσα χρονική στιμή t και 24 ώρες πριν από την τρέχουσα χρονική στιγμή t	Τιμές καιρικών συνθηκών 22, 23 και 24 ώρες μετά από την τρέχουσα χρονική στιγμή t

οι δείκτες $(t \pm i)$ δηλώνουν i 15-λεπτα διαστήματα πριν/μετά από την τρέχουσα χρονική στιγμή t, δηλ., ο δείκτης $t+4$ αναφέρεται σε διάστημα 1 ώρας μετά από την τρέχουσα χρονική στιγμή

Πίνακας 6: Περιγραφή των σεναρίων προσομοίωσης. Το έτος αναφοράς του συνόλου δεδομένων εκπαίδευσης είναι το 2017 και του συνόλου δεδομένων ελέγχου το 2018. Όλα τα σύνολα δεδομένων αποτελούνται από δεδομένα Ε.Ι. και καιρικών συνθηκών.

Τίτλος σεναρίου	**Σ1**	**Σ2**	**Σ3**	**Σ4**	**Σ5**	**Σ6**	**Σ7**
Χρονική περίοδο στην όποια υπάρχουν τα δεδομένα	Η εποχή του χειμώνα	Η εποχή της άνοιξης	Η εποχή του καλοκαιριού	Η εποχή του φθινοπώρου	Εργάσιμες ημέρες ολόκληρου του έτους	Χρονικές περίοδοι ολόκληρου του έτους με φώς του ήλιου	Χρονικές περίοδοι των εργάσιμων ημερών με φως του ήλιου

Κάθε πείραμα αναφέρεται στην εκπαίδευση του μοντέλου VGNN και στην επακόλουθη εξαγωγή προβλέψεων, χρησιμοποιώντας τα δεδομένα των διαφορετικών σεναρίων, που περιγράφονται παραπάνω. Όλες οι πειραματικές δοκιμές ακολουθούν ένα κοινό πρωτόκολλο το οποίο περιγράφεται αναλυτικά στη συνέχεια. Αρχικά, κατασκευάζεται ένας γράφος φυσικής ορατότητας που αντιστοιχεί στη χρονοσειρά εισόδου του φορτίου. Έπειτα χρησιμοποιείται ένα μη γραμμικό και non-positional νευρωνικό δίκτυο γράφων, όπως ορίζεται στην εργασία [225]. Τα δύο δίκτυα πρόσθιας τροφοδότησης, AGGREGATE και UPDATE, που ενσωματώνει το νευρωνικό δίκτυο γράφων, περιέχουν μία κρυφή στοιβάδα με 2 κρυφούς νευρώνες το καθένα, ενώ ως συναρτήσεις ενεργοποίησης του κάθε νευρώνα επιλέγονται η υπερβολική εφαπτομένη και μια γραμμική συνάρτηση αντίστοιχα. Κατά τη διαδικασία της εκπαίδευσης ο δείκτης MSE χρησιμοποιείται ως μέτρο του σφάλματος εκπαίδευσης και αξιολόγησης. Προκειμένου να αποφευχθεί η ενδεχόμενη υπερπροσαρμογή του μοντέλου εκπαίδευσης, κάνουμε χρήση του 70% των συνολικών δεδομένων του συνόλου εκπαίδευσης για την φάση της εκμάθησης (training) του μοντέλου ενώ το 30% αποτελεί το σύνολο δεδομένων αξιολόγησης (validation). Οι τιμές των παραμέτρων της ανθεκτικής οπισθοδιάδοσης, Δ_{max}, Δ_{min}, η^{-}, και η^{+} τέθηκαν ίσες με 50, $1e-6$, 0.5 και 1.2 αντίστοιχα ενώ ο μέγιστος αριθμός βημάτων οπισθοδιάδοσης S_b ανέρχεται σε 50. Το μοντέλο GNN που επιτυγχάνει τη μικρότερη τιμή του MSE στο σύνολο των δεδομένων αξιολόγησης θεωρείται ως το βέλτιστο μοντέλο. Τελικά, η απόδοση του μοντέλου VGNN αξιλογείται με βάση τους δείκτες MAE και R^2 κατά το στάδιο του ελέγχου (testing). Οι πειραματικές δοκιμές που έλαβαν χώρα, δεν περιορίζονται σε προσομοιώσεις του προτεινόμενου μοντέλου, αντιθέτως για σκοπούς συγκρίσεων αναφοράς, περιλαμβάνουν προσομοιώσεις με ορισμένες από τις ευρύτερα διαδεδομένες μεθοδολογίες πρόβλεψης. Παρακάτω παρατίθενται οι μεθολογίες αυτές καθώς και οι τεχνικές λεπτομέρειες και τις διαμορφώσεις των αντίστοιχων μοντέλων πρόβλεψης που υλοποιήθηκαν:

- *Γραμμική παλινδρόμηση*. Οι μοναδικές παράμετροι που συνδέονται με τη μέθοδο αυτή, είναι οι συντελεστές του μοντέλου γραμμικής παλινδρόμησης οι οποίοι υπολογίζονται με τη βοήθεια του αλγορίθμου ελαχίστων τετραγώνων [108].
- *Παλινδρόμηση με χρήση διανυσμάτων υποστήριξης*. Ένα μοντέλο SVR με Γκαουσιανή συνάρτηση πυρήνα υλοποιείται [162] ενώ ο αλγόριθμος της διαδοχικής ελάχιστης βελτιστοποίησης εφαρμόζεται για την εκπάιδευσή του. Οι υπερπαράμετροι του μοντέλου, δηλαδή ο παράγοντας ε, ο παράγοντας κλιμάκωσης πυρήνα (kernel scale) sigma και ο όρος box constraint [361], βελτιστοποιούνται με τη βοήθεια του αλγορίθμου της Μπεϋζιανής βελτιοστοποίησης.
- *Νευρωνικά δίκτυα MLP*. Ένα μοντέλο νευρωνικών δικτύων MLP αναπτύσσεται και εκπαιδεύεται με γνωστό αλγόριθμο Levenberg-Marquardt [362]. Προκειμένου η υπολογιστική πολυπλοκότητα του μοντέλου να διατηρηθεί σε συγκρίσιμα επίπεδα με αυτή των υπόλοιπων μεθόδων, το νευρωνικό δίκτυο περιλαμβάνει 2 κρυφές στοιβάδες. Για τον προσδιορισμό των υπερπαραμέτρων του, ακολουθείται μια διαδικασία εξαντλητικής αναζήτησης, δοκιμάζοντας όλους του δυνατούς συνδυασμούς νευρώνων στο εύρος μεταξύ 5 και 20 για κάθε στοιβάδα και υπολογίζοντας την απόδοση του μοντέλου μέσω 10πλής διασταυρούμενης αξιολόγησης. Η εφαρμογή της προσέγγισης αυτής οδηγεί σε μια αρχιτεκτονική 15 και 5 νευρώνων σε κάθε στοιβάδα. Στο σημείο αυτό υπογραμμίζεται ότι για την αντιμετώπιση της ευαισθησίας που εμφανίζουν οι αλγόριθμοι εκπαίδευσης των νευρωνικών δικτύων MLP λόγω της αρχικοποίησης των βαρών, υιοθετείται η τακτική που χρησιμοποιείται κατά τις μελέτες περίπτωσης που αναλύονται στις παραγράφους 3.3.1 και 3.3.2.

Για την εξασφάλιση ενός κοινού πλαισίου συγκρίσεων μεταξύ της προτεινόμενης μεθόδου και των μεθόδων αναφοράς, όλα τα μοντέλα τροφοδοτούνται και αξιολογούνται με κοινά δεδομένα. Επιπλέον, όλα τα πειράματα υλοποιούνται με τη βοήθεια του υπολογιστικού πακέτου MathWorks MATLAB R2022a και διεξάγονται στο ίδιο υπολογιστικό σύστημα, που είναι εξοπλισμένο με επεξεργαστή Intel(R) Core(TM) i7-10700K CPU @ 3.80GHz, μνήμη RAM 32GB και κάρτα γραφικών NVIDIA GeForce GTX 1080.

4. ΑΠΟΤΕΛΕΣΜΑΤΑ

Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα που προέκυψαν από τις πειραματικές προσομοιώσεις πρόβλεψης με χρήση πραγματικών δεδομένων μικτού φορτίου για κάθε μια από τις μεθοδολογίες Μ.Μ. που αναπτύχθηκαν στα πλαίσια της Δ.Δ., δηλαδή του ιεραρχικού μοντέλου αραιής αναπαράστασης (παράγραφος 4.1), του μοντέλου ανσάμπλ (παράγραφος 4.2) και του μοντέλου Νευρωνικών Δικτύων Γράφων (παράγραφος 4.3). Η παράγραφος 4.4 που ακολουθεί, εκθέτει τα αποτελέσματα πρόβλεψης που συμπεριλαμβάνονται σε δημοσίευση η οποία πραγματοποιήθηκε στα πλαίσια της Δ.Δ. Σε όλη την έκταση του κεφαλαίου, ιδιαίτερη έμφαση αποδίδεται στην οπτική αναπαράσταση των αποτελεσμάτων, έτσι ώστε να είναι ευανάγνωστα και κατατοπιστικά.

4.1 Μοντέλο αραιής αναπαράστασης και εκμάθησης λεξικού

Ένας πολύ σημαντικός παράγοντας κατά την αξιολόγηση ενός μοντέλου πρόβλεψης αποτελεί η ακρίβεια που επιτυγχάνει στην εκτίμηση των μελλοντικών τιμών κατανάλωσης ηλεκτρικού φορτίου. Για να είναι δυνατή η σύγκριση και η αξιολόγηση των πειραματικών αποτελεσμάτων, κρίθηκε σκόπιμο να χρησιμοποιηθεί ένας συνοπτικός πίνακας Έτσι, ο Πίνακας 7 περιέχει τις τιμές των δεικτών απόδοσης για την προτεινόμενη προσέγγιση, μαζί με τις υπόλοιπες μεθόδους που βασίζονται σε αραιή κωδικοποίηση. Για λόγους καλύτερης αναγνωσιμότητας, κάθε μέθοδος του πίνακα συνοδεύεται από την αντίστοιχη νόρμα του σφάλματος που ενσωματώνει. Όπως φαίνεται, το ιεραρχικό μοντέλο αποδεικνύεται ότι είναι το πιο αποτελεσματικό μεταξύ όλων των υλοποιήσεων αραιότητας, συνεπώς κρίνεται απαραίτητο η απόδοσή του να συγκριθεί με εκείνες που παράγονται από τις γνωστές γραμμικές και μη γραμμικές μεθόδους. Τα αντίστοιχα αποτελέσματα παρατίθενται στον πίνακα Πίνακας 8. Για την μέτρηση της αποδοτικότητας των μοντέλων χρησιμοποιούνται οι δείκτες MAE και R^2, καθώς αποτελούν τους πιο κοινώς χρησιμοποιούμενους στη βιβλιογραφία. Στο πλαίσιο της διερεύνησης της προβλεπτικής ακρίβειας του προτεινόμενου μοντέλου, οι δείκτες απόδοσης υπολογίστηκαν μεμονωμένα για σενάρια ημέρας και νύχτας και συμπεριλαμβάνονται στον πίνακα Πίνακας 9.

Πίνακας 7: Απόδοση της προτεινόμενης και των υπολοίπων μεθόδων αραιότητας για προβλέψεις με ορίζοντα 1 ώρας.

Μέθοδος Αραιότητας	**Νόρμα**	**R^2**	**MAE**
K-SVD/OMP	ℓ_0	0.89	0.81
Online / LARS-LASSO	ℓ_1	0.88	0.74
Elastic Net	$\ell_1 + \ell_2$	0.90	0.86
Online-OMP / LARS-LASSO	ℓ_0	0.92	0.90
Αρχετυπική Ανάλυση	ℓ_2	0.80	1.02
Ιεραρχική Αραιότητα	ℓ_0^{tree}	**0.94**	**0.61**

Πίνακας 8: Απόδοση της προτεινόμενης μεθόδου αραιότητας και των μεθόδων αναφοράς για προβλέψεις με ορίζοντα 1 ώρας.

Μέθοδος	**R^2**	**MAE**
LR	0.92	0.76
SVR	0.93	0.73
MLP[3]	0.92 (0.89 ± 0.11)	0.79 (0.82 ± 0.16)
Ιεραρχική Αραιότητα	**0.94**	**0.61**

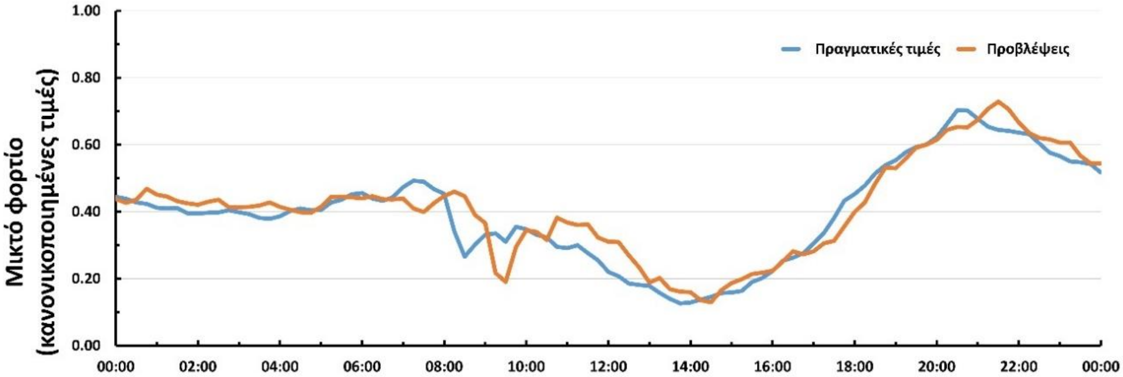
Πίνακας 9: Απόδοση της προτεινόμενης μεθόδου αραιότητας και των μεθόδων αναφοράς για προβλέψεις περιόδων ημέρας και νύχτας με ορίζοντα 1 ώρας.

Μέθοδος	**Σενάριο ημέρας**		**Σενάριο νύχτας**	
	R^2	**MAE**	**R^2**	**MAE**
LR	0.89	0.98	0.94	0.58
SVR	0.90	0.94	0.95	0.53
MLP[3]	0.89 (0.86 ± 0.13)	0.99 (1.02 ± 0.20)	0.94 (0.91 ± 0.08)	0.59 (0.61 ± 0.13)
Ιεραρχική Αραιότητα	**0.90**	**0.88**	**0.98**	**0.34**

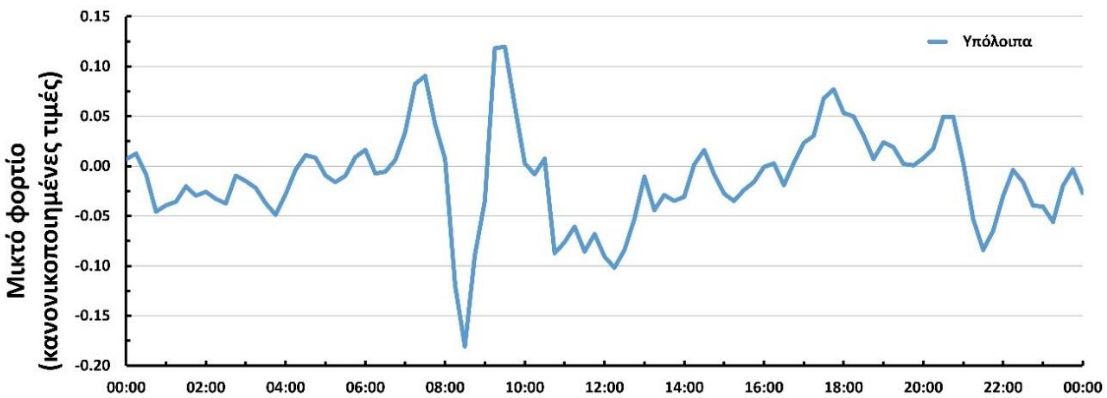
Η απόδοση πρόβλεψης του ιεραρχικά δομημένου μοντέλου αραιής αναπαράστασης παρουσιάζεται γραφικά στα σχήματα: Σχήμα 13, Σχήμα 14 και Σχήμα 15. Τα δεδομένα που εμφανίζονται σε όλα τα γραφήματα οφείλουν να είναι κανονικοποιημένα στο εύρος [-1, 1] για λόγους εμπιστευτικότητας. Ένα στιγμιότυπο των ωριαίων πραγματικών τιμών και τιμών πρόβλεψης που εκτείνεται σε ένα τυχαία επιλεγμένο διάστημα 24 ωρών, απεικονίζεται στο Σχήμα 13, ενώ τα αντίστοιχα υπόλοιπα (residuals) παρουσιάζονται στο Σχήμα 14. Σημειώνεται στο σημείο αυτό, ότι οι προβλέψεις για την επόμενη ώρα παρέχονται κάθε 15 λεπτά. Αυτό σημαίνει ότι κατά την εφαρμογή του σε πραγματικό χρόνο, το μοντέλο παρέχει ανά 15 λεπτά μια νέα πρόβλεψη για τιμή του μικτού φορτίου 1 ώρα μετά από την τρέχουσα χρονική στιγμή. Ο κατακόρυφος άξονας και στα δύο προαναφερθέντα γραφήματα (Σχήμα 13, Σχήμα 14 και Σχήμα 15) αντιπροσωπεύει τις κανονικοποιημένες τιμές του μικτού φορτίου. Μια εναλλακτική οπτική απεικόνιση των προβλέψεων απεικονίζεται στο Σχήμα 15, όπου οι πραγματικές τιμές Ε.Ι έναντι των αντίστοιχων προβλέψεων παρουσιάζονται χρησιμοποιώντας ένα διάγραμμα διασποράς

[3] Για το μοντέλο MLP, η υψηλότερη απόδοση του δικτύου παρουσιάζεται καθώς και ο μέσος όρος και η τυπική απόκλιση 50 εκτελέσεων σε παρένθεση

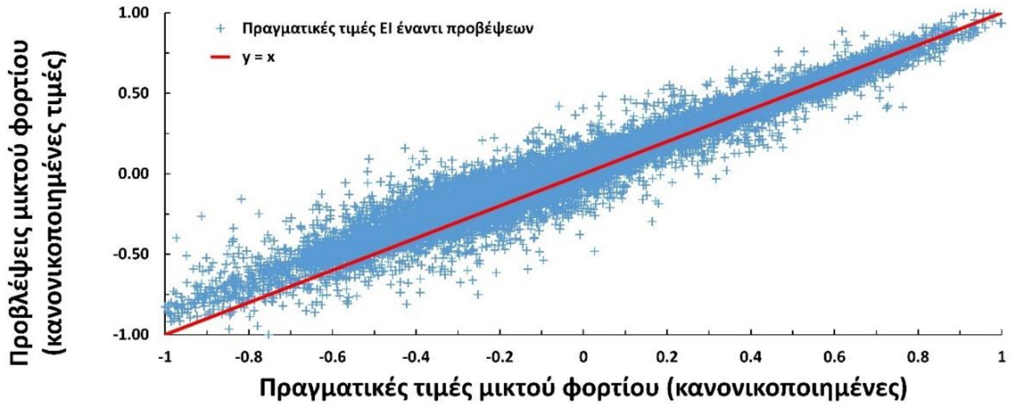
(scatterplot). Η συνεχής διαγώνια γραμμή που φαίνεται στο διάγραμμα, αναπαριστά την ευθεία αναφοράς (reference line), κατά μήκος της οποίας σημειώνεται μηδενικό σφάλμα μεταξύ πραγματικών τιμών και των προβλέψεων AI.



Σχήμα 13: Ωριαίες πραγματικές τιμές και τιμές πρόβλεψης μικτού φορτίου σύμφωνα με το μοντέλο ιεραρχικής αραιότητας.



Σχήμα 14: Τιμές υπολοίπων (ή σφάλματος) ανάμεσα στις ωριαίες πραγματικές τιμές και τιμές πρόβλεψης μικτού φορτίου σύμφωνα με το μοντέλο ιεραρχικής αραιότητας.



Σχήμα 15: Διάγραμμα διασποράς των ωριαίων πραγματικών τιμών μικτού φορτίου έναντι των αντίστοιχων προβλέψεων που προκύπτουν από το μοντέλο ιεραρχικής αραιότητας.

4.2 Μοντέλο ανσάμπλ

Στην παράγραφο αυτή παρουσιάζονται τα αποτελέσματα των εκτεταμένων προσομοιώσεων που έλαβαν χώρα με σκοπό την αξιολόγηση της προβλεπτικής ικανότητας της προτεινόμενης μεθόδου ανσάμπλ. Στα σχήματα: Σχήμα 16.α-Σχήμα 16.στ, απεικονίζονται μια σειρά από διαγράμματα διασποράς. Κάθε ένα από αυτά αναπαριστά τις πραγματικές τιμές μικτού φορτίου σε αντιπαράθεση με τις αντίστοιχες τιμές πρόβλεψης για χρονικό ορίζοντας πρόβλεψης 15 λεπτών, 1, 2, 3, 6 και 24 ωρών, αντίστοιχα. Οι διαγώνιες γραμμές υποδηλώνουν ταύτιση των προβλέψεων με τις πραγματικές τιμές, ενώ οι άξονες των γραφημάτων περιέχουν κανονικοποιημένες τιμές Ε.Ι.

Στον πίνακα Πίνακας 10 παρέχονται τα αριθμητικά αποτελέσματα των επιδόσεων της πρόβλεψης της προτεινόμενης μεθόδου σε αντιπαραβολή με τις επιμέρους μεθόδους μηχανικής εκμάθησης που περιλαμβάνουν το σύνολο των υπο-μοντέλων. Προκειμένου τα αποτελέσματα των προβλέψεων για τους διαφορετικούς χρονικούς ορίζοντες να είναι ευδιάκριτα, ο πίνακας χωρίζεται σε αντίστοιχες ενότητες (I-VI). Η ακρίβεια των προβλέψεων του μοντέλου αξιολογείται μέσω του συντελεστή συσχέτισης (R^2), του δείκτη RMSE και του δείκτη MAE, καθώς τα μέτρα αυτά είναι τόσο αντιπροσωπευτικά όσο και αποτελεσματικά [363]. Για λόγους σύγκρισης, ο πίνακας περιέχει επίσης τις τιμές των δεικτών για όλα τα υπο-μοντέλα, καθώς και το ποσοστό κατάταξής τους στην πρώτη θέση αυτή. Η ποσότητα αυτή, επισημαίνεται ως «Κατάταξη #1» στον πίνακα Πίνακας 10, και δηλώνει πόσες φορές κάθε υπο-μοντέλο σημείωσε την 1η θέση μεταξύ όλων των υπο-μοντέλων, δηλαδή πέτυχε το χαμηλότερο MAE.

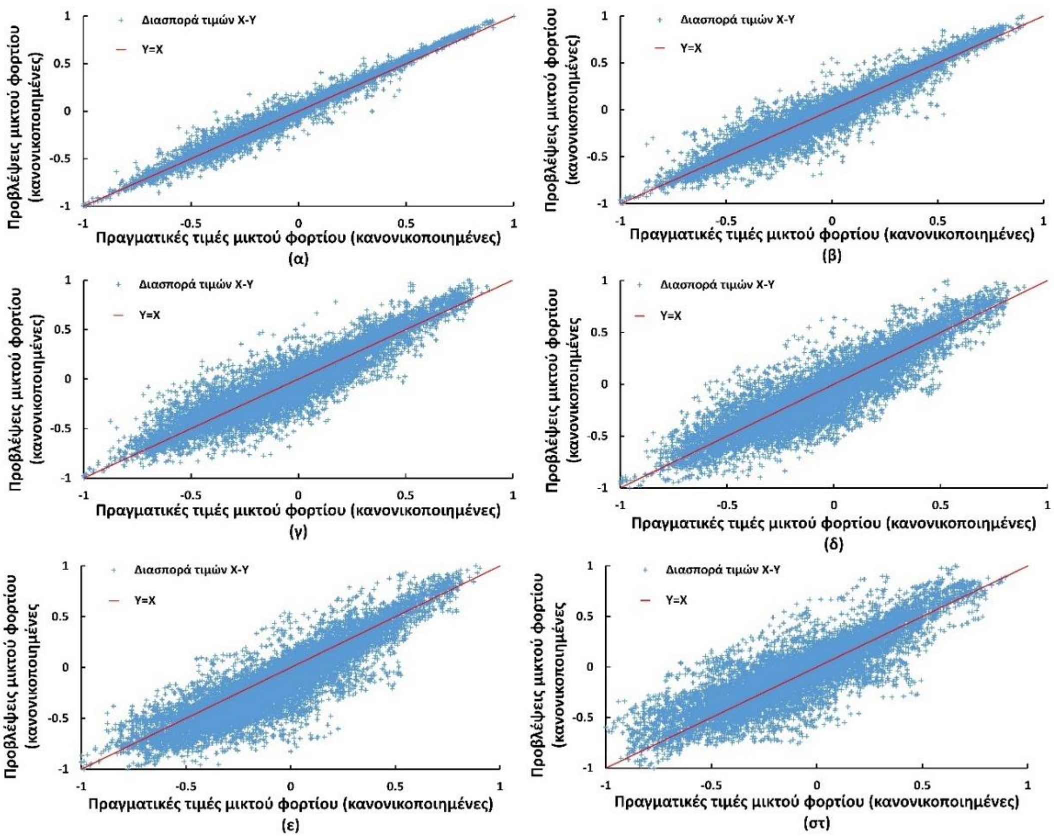
Πίνακας 10: Απόδοση της προτεινόμενης μεθοδολογίας ανσάμπλ, του μοντέλου ανσάμπλ MLP της αναφοράς [357] και των μεμονωμένων μεθόδων Μ.Μ. για κάθε ορίζοντα πρόβλεψης. Στον πίνακα περιλαμβάνονται οι τιμές των δεικτών MAE, RMSE και R^2 που επιτυγχάνει κάθε μοντέλο καθώς και το ποσοστό σύμφωνα με το οποίο κάθε υπο-μοντέλο σημειώνει την χαμηλότερη τιμή του MAE.

Μέθοδος	R^2	MAE	RMSE	Κατάταξη #1
Ενότητα I - 15 λεπτά				
Προτεινόμενη	0.99	0.26	0.47	-
MLP ανσάμπλ	0.99	0.28	0.49	-
MLP	0.99	0.27	0.48	19.25%
RBF	0.99	0.27	0.48	19.94%
LR	0.99	0.27	0.48	10.49%
SVR	0.99	0.27	0.48	14.23%
RF	0.98	0.30	0.51	24.27%
SR	0.99	0.27	0.48	11.82%
Ενότητα II - 1 ώρα				
Προτεινόμενη	0.94	0.60	1.00	-
MLP ανσάμπλ	0.93	0.63	1.02	-
MLP	0.92	0.67	1.15	21.56%
RBF	0.93	0.64	1.04	20.77%
LR	0.93	0.64	1.04	8.91%
SVR	0.93	0.65	1.06	10.70%
RF	0.93	0.67	1.06	20.79%

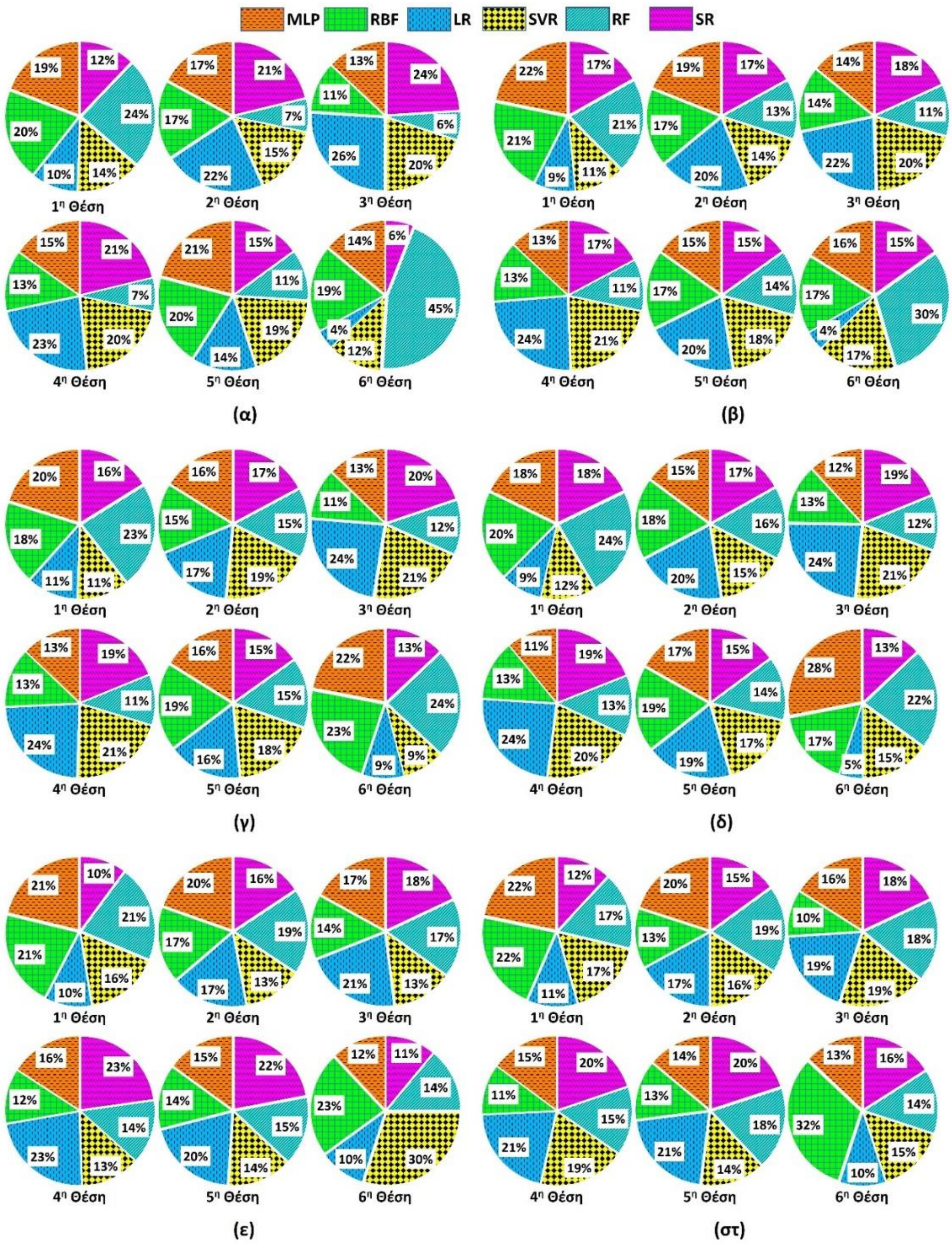
SR	0.93	0.64	1.05	17.27%
Ενότητα III - 2 ώρες				
Προτεινόμενη	0.88	0.88	1.38	-
MLP ανσάμπλ	0.88	0.90	1.37	-
MLP	0.85	1.00	1.59	20.32%
RBF	0.87	0.96	1.45	18.32%
LR	0.87	0.93	1.44	11.20%
SVR	0.87	0.94	1.46	11.38%
RF	0.87	0.97	1.47	22.94%
SR	0.87	0.93	1.45	15.92%
Ενότητα IV - 3 ώρες				
Προτεινόμενη	0.84	1.06	1.59	-
MLP ανσάμπλ	0.84	1.09	1.62	-
MLP	0.79	1.25	1.854	18.08%
RBF	0.83	1.14	1.67	20.14%
LR	0.82	1.13	1.68	8.74%
SVR	0.82	1.15	1.67	11.82%
RF	0.82	1.14	1.70	23.67%
SR	0.82	1.12	1.70	17.54%
Ενότητα V - 6 ώρες				
Προτεινόμενη	0.83	1.11	1.65	-
MLP ανσάμπλ	0.83	1.20	1.70	-
MLP	0.83	1.18	1.67	20.77%
RBF	0.80	1.29	1.80	21.31%
LR	0.78	1.33	1.91	10.34%
SVR	0.75	1.43	2.02	16.42%
RF	0.81	1.21	1.76	21.08%
SR	0.77	1.34	1.93	10.08%
Ενότητα VI - 24 ώρες				
Προτεινόμενη	0.79	1.18	1.82	-
MLP ανσάμπλ	0.78	1.24	1.85	-
MLP	0.78	1.23	1.86	21.93%
RBF	0.74	1.41	2.04	21.83%
LR	0.76	1.32	1.95	11.16%

SVR	0.74	1.30	2.03	16.82%
RF	0.77	1.27	1.92	16.71%
SR	0.75	1.34	1.99	11.56%

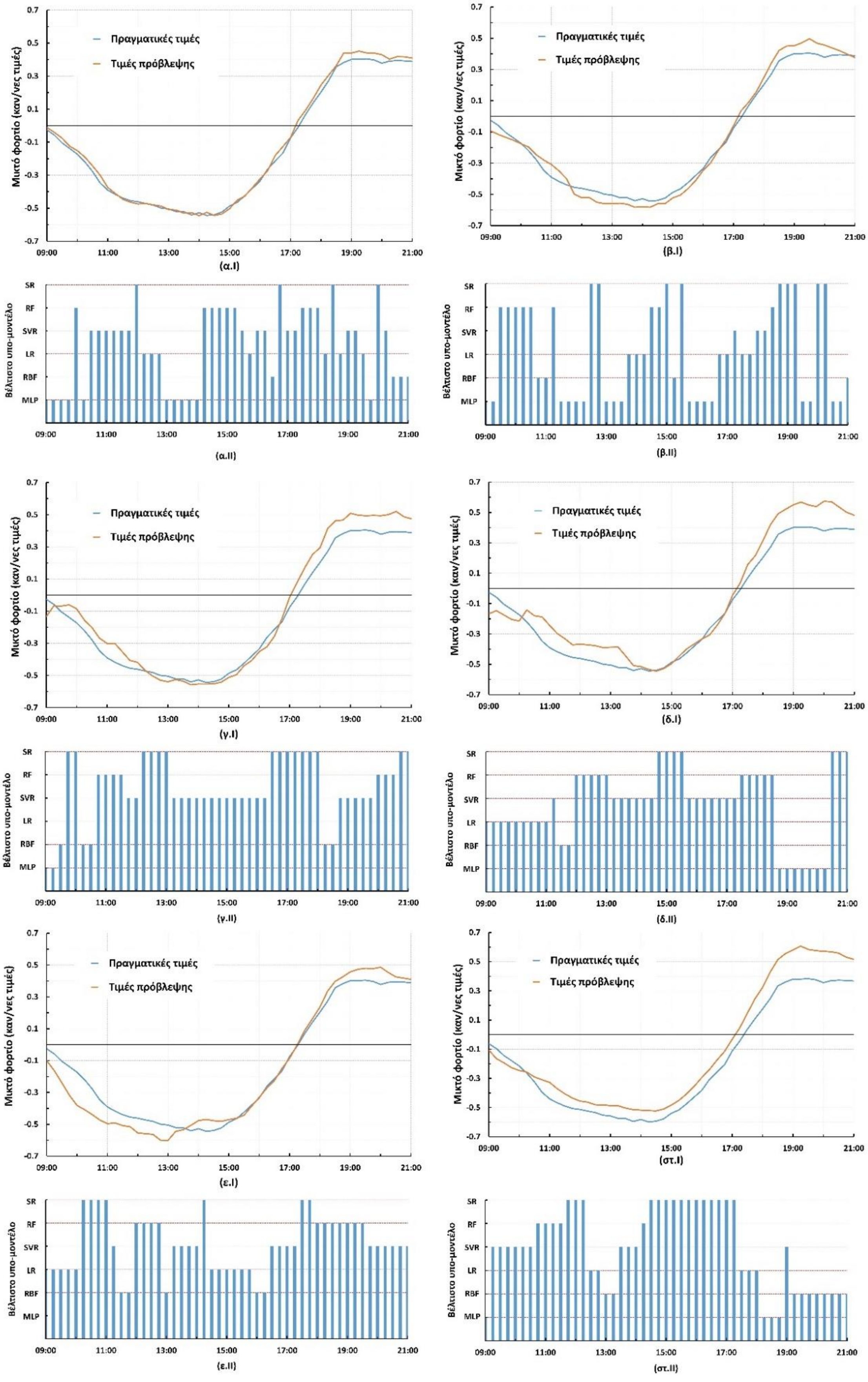
Η προαναφερθείσα μορφή κατάταξης των υπό-μοντέλων απεικονίζεται στο Σχήμα 17. Κάθε ένα από τα σχήματα Σχήμα 17.α-Σχήμα 17.στ, αναφέρεται σε ορίζοντες πρόβλεψης 15 λεπτών, 1, 2, 3, 6 και 24 ωρών, αντίστοιχα. Κάθε ένα από τα επιμέρους σχήματα περιέχει 6 γραφήματα πίτας, τα οποία δείχνουν τα ποσοστά που αντιστοιχούν στο πλήθος των εμφανίσεων κατά τις οποίες κάθε υπό-μοντέλο κατατάχθηκε στην οποιαδήποτε θέση, από την $1^{η}$ έως την $6^{η}$, με βάση το σταθμισμένο MAE του. Για παράδειγμα η $2^{η}$ πίτα του σχήματος Σχήμα 17.α υποδηλώνει ότι για 15-λεπτες προβλέψεις, το υπο-μοντέλο MLP κατατάχθηκε στην $2^{η}$ θέση ανάμεσα στο σύνολο των υπό-μοντέλων σημειώνοντας ποσοστό 17%, το υπο-μοντέλο SR σημειώνοντας ποσοστό 21% και ούτω καθεξής. Τέλος, παρέχονται αναλυτικά γραφήματα για κάθε χρονικό ορίζοντα πρόβλεψης σύμφωνα με τα σχήματα: Σχήμα 18.α.I-Σχήμα 18.στ.I. Σε κάθε ένα από αυτά απεικονίζονται τόσο οι πραγματικές τιμές όσο και οι αντίστοιχες προβλέψεις μικτού φορτίου για ένα τυχαία επιλεγμένο χρονικό παράθυρο διάρκειας 12 ωρών (από τις 09:00 έως τις 21:00), μιας αυθαίρετα επιλεγμένη ημέρας του υποσυνόλου δεδομένων ελέγχου. Αυτά τα γραφήματα συνοδεύονται από τα σχήματα: Σχήμα 18.α.II-Σχήμα 18.στ.II, τα οποία υποδεικνύουν ποιο υπο-μοντέλο έχει το βάρος με την υψηλότερη τιμή για κάθε χρονικό βήμα πρόβλεψης, χρησιμοποιώντας μια γραφική παράσταση ράβδων.



Σχήμα 16: Διαγράμματα διασποράς των πραγματικών τιμών έναντι των τιμών πρόβλεψης μικτού φορτίου για διάφορους ορίζοντες: (α) 15 λεπτών, (β) 1 ώρας, (γ) 2 ωρών, (δ) 3 ωρών, (ε) 6 ωρών, και (στ) 24 ωρών. Οι τιμές πρόβλεψης που βρίσκονται πάνω στην διαγώνια γραμμή ταυτίζονται με τις πραγματικές τιμές του φορτίου. Κάθε σημειακό σύμβολο (+) αναφέρεται σε ένα δεδομένο και δηλώνει την απόκλιση της πρόβλεψης από την πραγματική του τιμή.



Σχήμα 17: Διαγράμματα πίτας που αναπαριστούν την κατάταξη των υπό-μοντέλων που ενσωματώνονται στο προτεινόμενο μοντέλο ανσάμπλ για διάφορους ορίζοντες πρόβλεψης: (α) 15-λεπτών, (β) 1 ώρας, (γ) 2 ωρών, (δ) 3 ωρών, (ε) 6 ωρών, και (στ) 24 ωρών. Κάθε διάγραμμα πίτας αναφέρεται σε μία θέση κατάταξης και δείχνει το ποσοστό σύμφωνα με το οποίο κάθε υπο-μοντέλο κατατάχθηκε στη θέση αυτή. Κάθε υπο-μοντέλο παρουσιάζεται με διαφορετικό χρώμα και μοτίβο.



Σχήμα 18: Αποτελέσματα προβλέψεων για ένα τυχαία επιλεγμένο διάστημα διάρκειας 12 ωρών με ορίζοντες: (α) 15-λεπτών, (β) 1 ώρας, (γ) 2 ωρών, (δ) 3 ωρών, (ε) 6 ωρών, και (στ) 24 ωρών. Τα επιμέρους γραφήματα με την επισήμανση «Ι» απεικονίζουν τις πραγματικές τιμές και τις προβλέψεις, ενώ εκείνα με την επισήμανση «ΙΙ» απεικονίζουν τα αποτελέσματα της απόδοσης του βέλτιστου υπο-μοντέλου.

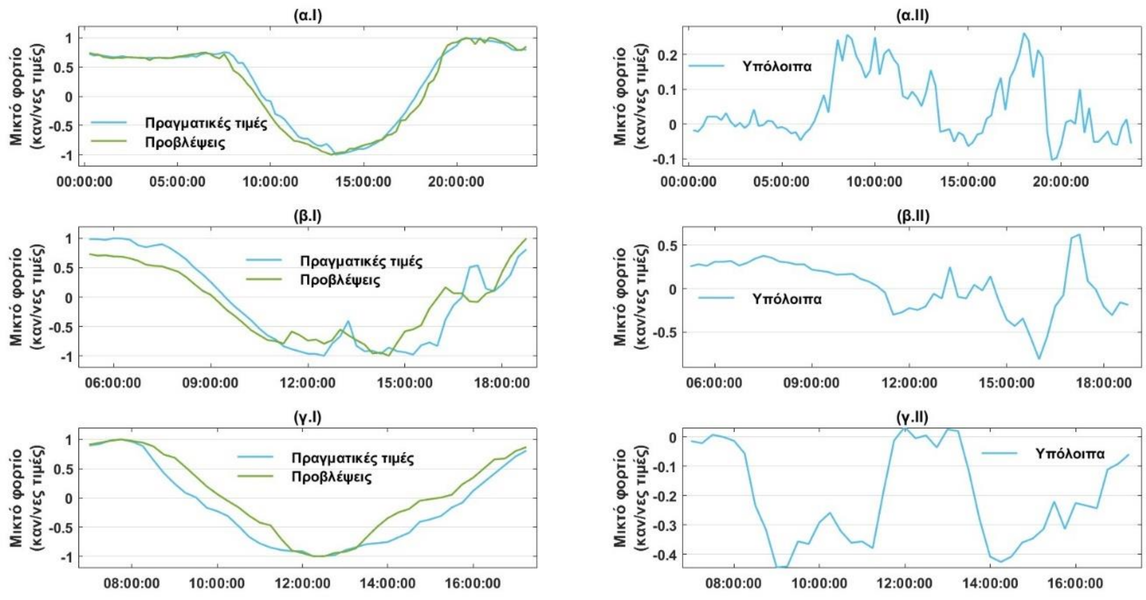
4.3 Μοντέλο νευρωνικών δικτύων γράφων ορατότητας

Τα αποτελέσματα που προέκυψαν κατά την διεξαγωγή των προσομοιώσεων του μοντέλου νευρωνικών δικτύων γράφων παρατίθενται στην παράγραφο αυτή. Αρχικά, εξετάζεται η ακρίβεια των προβλέψεων του μοντέλου η οποία ποσοτικοποιείται με τη βοήθεια των δεικτών απόδοσης R^2 και MAE. Οι τιμές των δεικτών αυτών τόσο για το προτεινόμενο μοντέλο όσο και για διάφορες γραμμικές και μη γραμμικές διαδεδομένες μεθόδους πρόβλεψης, δηλαδή της ΓΠ, των νευρωνικών δικτύων MLP και της SVR, φαίνονται στον πίνακα Πίνακας 11. Ο χρόνος που απαιτεί κάθε μεθοδολογία για την εκτέλεση κάθε σεναρίου προσομοίωσης περιλαμβάνεται στον πίνακα έτσι ώστε να παρέχεται ένα ολοκληρωμένο πλαίσιο σύγκρισης ανάμεσά τους.

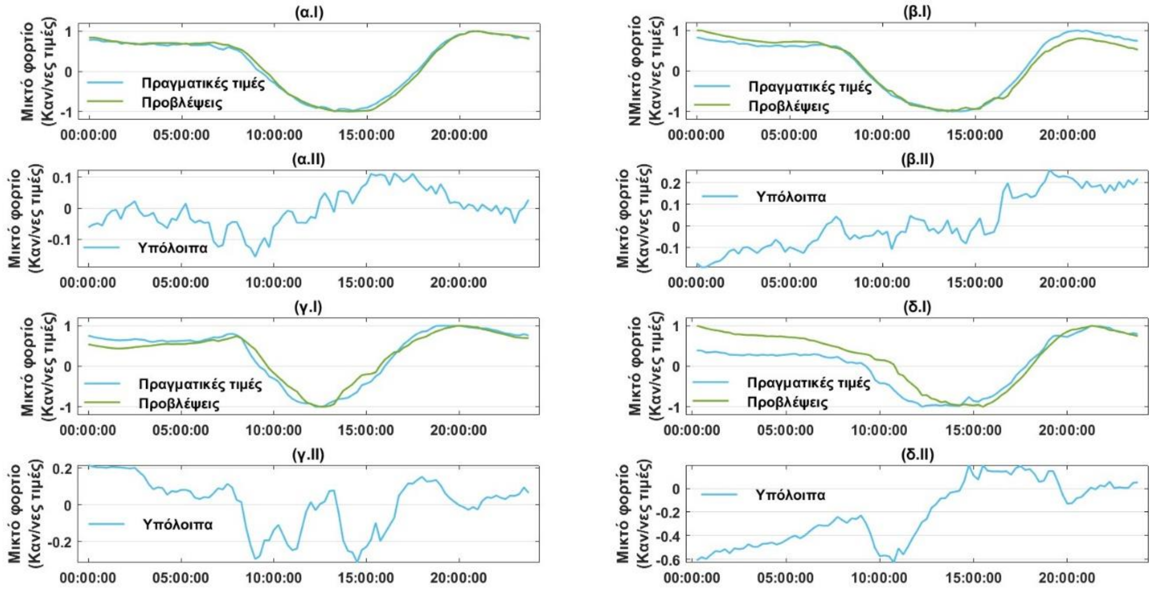
Πίνακας 11: Απόδοση του μοντέλου VGNN και των μοντέλων αναφοράς για την πρόβλεψη του μικτού φορτίου με ορίζοντα 24 ωρών ανά σενάριο προσομοίωσης, όπως αυτά περιγράφονται στον πίνακα Πίνακας 6.

Μέθοδος / Σενάριο	LR			MLP			SVR			VGNN		
	MAE	R^2	Χρόνος εκτέλεσης (s)	MAE	R^2	Χρόνος εκτέλεσης (s)	MAE	R^2	Χρόνος εκτέλεσης (s)	MAE	R^2	Χρόνος εκτέλεσης (s)
Σ1	3.61	0.45	1.40	3.70	0.39	186.27	4.13	0.34	415.08	**3.57**	**0.50**	194.37
Σ2	3.85	0.64	1.71	3.67	0.65	164.68	4.45	0.53	425.28	**3.61**	**0.68**	208.62
Σ3	2.80	0.84	1.764	2.79	0.81	252.20	3.25	0.77	421.63	**2.72**	**0.85**	207.97
Σ4	3.00	0.68	1.61	3.05	0.65	164.96	3.46	0.62	422.02	**2.97**	**0.69**	185.26
Σ5	3.25	0.69	1.34	3.12	0.69	170.90	3.67	0.63	412.85	**3.15**	**0.71**	310.30
Σ6	4.44	0.52	1.44	4.41	0.50	165.03	5.01	0.43	417.08	**4.31**	**0.55**	260.52
Σ7	4.42	0.52	1.27	4.33	0.53	167.21	4.84	0.47	411.18	**4.25**	**0.57**	219.44

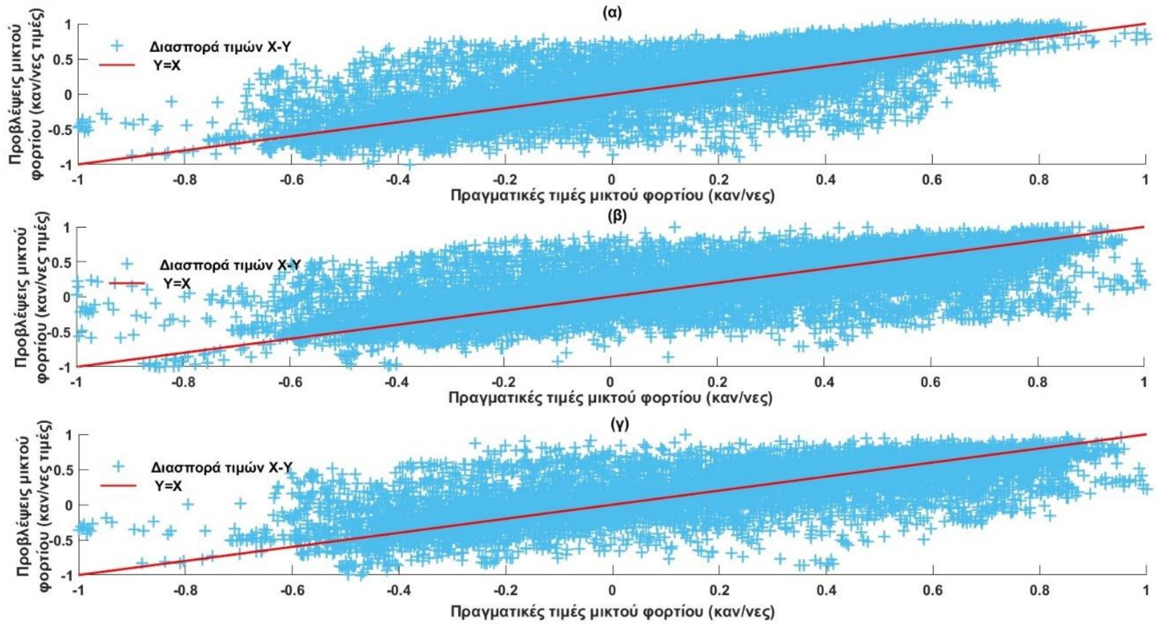
Η προβλεπτική απόδοση του προτεινόμενου μοντέλου VGNN παρουσιάζεται γραφικά στα σχήματα: Σχήμα 19-Σχήμα 22. Μια συλλογή από στιγμιότυπα των πραγματικών τιμών και των προβλέψεων που αφορούν σε δεδομένα (α) μίας εργάσιμης ημέρας (σενάριο Σ5), (β) των ωρών μιας οποιασδήποτε μέρας του έτους μέρας με φως του ήλιου (σενάριο Σ6) και (γ) των ωρών μιας εργάσιμης μέρας με φως του ήλιου (σενάριο Σ7), συνοδευόμενα από αντίστοιχα γραφήματα με τις τιμές των υπολοίπων (ή σφάλματος), απεικονίζονται στο Σχήμα 19. Ανάλογα διαγράμματα για δεδομένα που αντιστοιχούν σε (α) μία μέρα του καλοκαιριού (σενάριο Σ3), (β) μία μέρα του φθινοπώρου (σενάριο Σ4), (γ) μία μέρα του χειμώνα (σενάριο Σ1) και (δ) μία μέρα της άνοιξης (σενάριο Σ2), παρέχονται στο Σχήμα 20. Θα πρέπει να αναφερθεί ότι αυτές οι γραφικές παραστάσεις αναφέρονται σε διαφορετικές τυχαία επιλεγμένες ημέρες και ότι οι προβλέψεις παρέχονται κάθε 15 λεπτά. Τέλος, οι πραγματικές τιμές Ε.Ι έναντι των προβλέψεων για όλα τα σενάρια προσομοίωσης παρουσιάζονται με τη μορφή διαγραμμάτων διασποράς (Σχήμα 21-Σχήμα 22). Λόγω των συμφωνιών μη γνωστοποίησης, όλα τα δεδομένα μικτού φορτίου πρέπει να παρουσιαστούν με κανονικοποιημένα εύρη [-1, 1].



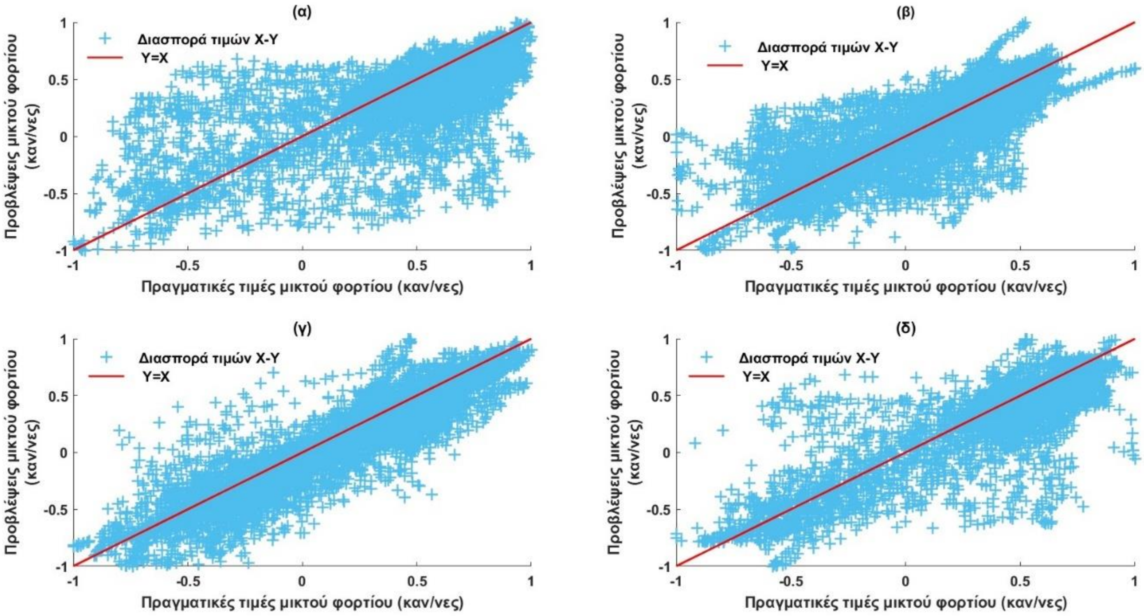
Σχήμα 19: Πραγματικές τιμές και τιμές προβλέψεων που προκύπτουν από τις πειραματικές δοκιμές του μοντέλου VGNN για τα σενάρια (α) Σ5, (β) Σ6, και (γ) Σ7, όπως αυτά περιγράφονται στον πίνακα Πίνακας 6.



Σχήμα 20: Πραγματικές τιμές και τιμές προβλέψεων που προκύπτουν από τις πειραματικές δοκιμές του μοντέλου VGNN για τα σενάρια (α) Σ3, (β) Σ4, (γ) Σ1, και (δ) Σ2, όπως αυτά περιγράφονται στον πίνακα Πίνακας 6.



Σχήμα 21: Διαγράμματα διασποράς των πραγματικών τιμών και των τιμών προβλέψεων που προκύπτουν από τις πειραματικές δοκιμές του μοντέλου VGNN για τα σενάρια (α) Σ5, (β) Σ6, και (γ) Σ7, όπως αυτά περιγράφονται στον πίνακα Πίνακας 6.



Σχήμα 22: Διαγράμματα διασποράς των πραγματικών τιμών και των τιμών προβλέψεων που προκύπτουν από τις πειραματικές δοκιμές του μοντέλου VGNN για τα σενάρια (α) Σ3, (β) Σ4, (γ) Σ1, και (δ) Σ2, όπως αυτά περιγράφονται στον πίνακα Πίνακας 6.

4.4 Αρχιτεκτονική OpenEdgePMU

Το κεφάλαιο ολοκληρώνεται με την παρουσίαση των αποτελεσμάτων μια σειράς από προσομοιώσεις πρόβλεψης του ηλεκτρικού φορτίου οι οποίες πραγματοποιούνται στο επίπεδο της χαμηλές τάσης του ηλεκτρικού δικτύου και όχι σε κάποιο απομακρυσμένο υπολογιστικό σύστημα διακομιστή (server). Οι εν λόγω πειραματικές δοκιμές έλαβαν χώρα στα πλαίσια της δημοσίευσης [256] που παράχθηκε κατά τη διάρκεια κατά τη διάρκεια εκπόνησης της Δ.Δ. Στο άρθρο αυτό, παρουσιάζεται η ανάπτυξη μια πλατφόρμας PMU που ενσωματώνει λογισμικό ανοιχτού κώδικα, έχει χαμηλό κόστος, υψηλή επίδοση σε εργασίες επεξεργασίας σήματος, παρέχει δυνατότητα άμεσης επέκτασης των εφαρμογών της, ενώ αποτελεί ένα κατάλληλο εργαλείο για δραστηριότητες έρευνας και ανάπτυξης. Βασικός στόχος της πλατφόρμας OpenEdgePMU, όπως ονομάζεται, είναι η ενίσχυση της ανθεκτικότητας των έξυπνων δικτύων. Προς αυτή την κατεύθυνση, η προτεινόμενη αρχιτεκτονική PMU είναι εφοδιασμένη με ένα ολοκληρωμένο αλγοριθμικό πλαίσιο παροχής προβλέψεων ηλεκτρικού φορτίου με πολλαπλούς χρονικούς ορίζοντες, που επιτρέπει την εύκολη και άμεση οποιουδήποτε νέου μοντέλου πρόβλεψης. Παράλληλα, παρέχει τη δυνατότητα δημιουργίας διαφορετικών σεναρίων εκτέλεσης, όπου χρησιμοποιούνται σύνολα δεδομένων εκπαίδευσης και ελέγχου τα οποία αντιστοιχούν σε χρονικές περιόδους επιλεγμένες από το χρήστη. Λόγω του υψηλού υπολογιστικού κόστους που απαιτείται συχνά για την εκπαίδευση των μοντέλων πρόβλεψης που βασίζονται σε μεθόδους ΜΜ, αυτή η διαδικασία λαμβάνει χώρα σε ένα κέντρο επιχειρήσεων (εφαρμογή σε νέφος - cloud), ενώ η παραγωγή προβλέψεων από τα ήδη εκπαιδευμένα μοντέλα εκτελείται απρόσκοπτα από το σύστημα OpenEdgePMU, και συγκεκριμένα χρησιμοποιώντας τη μονάδα edge processing.

Για σκοπούς επίδειξης και αξιολόγησης της λειτουργίας της προτεινόμενης πλατφόρμας διεξάγεται ένα αριθμός πειράματα με χρήση πραγματικών δεδομένων ηλεκτρικού φορτίου.

Η αξιοποίηση των δεδομένων υπόκειται σε κανόνες εμπιστευτικότητας και απορρήτου και έτσι η προέλευση και το χρονικό εύρος που καλύπτουν δεν γνωστοποιείται. Ωστόσο, είναι διαθέσιμη η πληροφορία ότι τα σύνολα δεδομένων εκπαίδευσης και ελέγχου εκτείνονται στο χρονικό εύρος ενός έτους το καθένα και τα έτη αυτά είναι διαδοχικά. Οι επιδόσεις ορισμένων τυπικών μεθόδων Μ.Μ. για προβλέψεις με διαφορετικούς ορίζοντες και διαφορετικά σενάρια εκτέλεσης, αξιολογούνται με βάση οι τιμές των δεικτών MSE και R2 όπως φαίνονται στον πίνακα Πίνακας 12. Τέλος, αξίζει να αναφερθεί ότι οι χρόνοι εκτέλεσης για τα μοντέλα πρόβλεψης φορτίου στο σύστημα OpenEdgePMU δεν υπερβαίνουν στη διάρκεια τα λίγα δευτερόλεπτα.

Πίνακας 12: Προβλεπτική απόδοση τυπικών μεθόδων Μ.Μ. που ενσωματώνει το σύστημα OpenEdgePMU για διαφορετικούς συνδυασμούς χρονικών οριζόντων και σεναρίων εκτέλεσης.

Μέθοδος	Μέτρα απόδοσης	Ορίζοντας: 15 λεπτά Σενάριο: Σ7	Ορίζοντας: 1 ώρα Σενάριο: Σ6	Ορίζοντας: 6 ώρες Σενάριο: Σ2	Ορίζοντας: 12 ώρες Σενάριο: Σ5	Ορίζοντας: 24 ώρες Σενάριο: Σ3
LR	MAE	0.92	2.25	3.48	2.95	2.63
	R^2	0.97	0.88	0.73	0.76	0.86
CART	MAE	0.93	2.10	2.92	2.66	2.83
	R^2	0.97	0.90	0.77	0.78	0.83
RF	MAE	0.93	2.05	2.91	2.66	2.52
	R^2	0.97	0.90	0.77	0.78	0.87
MLP	MAE	0.90	2.08	3.15	2.82	2.60
	R^2	0.97	0.89	0.75	0.75	0.86
LSTM	MAE	0.96	2.20	3.19	2.85	2.64
	R^2	0.97	0.88	0.76	0.78	0.86
SVR	MAE	0.96	2.22	3.40	2.99	2.55
	R^2	0.97	0.88	0.72	0.74	0.85

5. ΣΥΜΠΕΡΑΣΜΑΤΑ - ΘΕΜΑΤΑ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΕΡΕΥΝΑ

Στο τελευταίο κεφάλαιο παρατίθενται οι παρατηρήσεις και τα συμπεράσματα που εξάγονται από τα πειραματικά αποτελέσματα κάθε μιας από τις μελέτες περίπτωσης που παρουσιάστηκαν στο προηγούμενο κεφάλαιο (παράγραφος 5.1). Στην συνέχεια, γίνεται αναφορά στα γενικά συμπεράσματα που απορρέουν από την ερευνητική μελέτη που έλαβε χώρα στο σύνολό της. Κατ' ακολουθίαν, στην παράγραφο 5.2, δηλαδή στην τελευταία ενότητα του παρόντος κειμένου, αναφέρονται θέματα που χρήζουν περαιτέρω έρευνας και που έχουν αναδειχθεί μέσα από την εκπόνηση της παρούσης διδακτορικής διατριβής.

5.1 Συμπεράσματα

5.1.1 Μοντέλο αραιής αναπαράστασης και εκμάθησης λεξικού

Τα πρωταρχικά συμπεράσματα που προκύπτουν από τα αποτελέσματα των πειραματικών δοκιμών των μεθόδων αραιότητας που αναπτύχθηκαν στα πλαίσια της Δ.Δ., αφορούν στην αξιοπιστία των παραγόμενων προβλέψεων η οποία εξετάζεται με τη βοήθεια των δεικτών R^2 και MAE. Όπως διαπιστώνεται από τον πίνακα Πίνακας 7 το μοντέλο που βασίζεται στην ιεραρχική αραιότητα, υπερτερεί έναντι των υπόλοιπων προσεγγίσεων αραιής αναπαράστασης. Προσπαθώντας να αντιπαραβάλει κανείς τις αραιές μεθοδολογίες μεταξύ τους, παρατηρεί ότι η αρχετυπική ανάλυση δεν είναι μια ανταγωνιστική προσέγγιση σε σύγκριση με την προτεινόμενη. Αυτό πιθανότατα οφείλεται στους περιορισμούς της κυρτότητας στους οποίους υπόκειται το πρόβλημα εξαγωγής του λεξικού και των αντιστοίχων αραιών συντελεστών. Έτσι, το μοντέλο αυτό φαίνεται να μην επιτυγχάνει την επαρκή αναπαράσταση του σήματος, κάτι το οποίο απεικονίζεται στην αδυναμία του να παρέχει ακριβείς και αξιόπιστες προβλέψεις. Πέραν τούτου, η συγκεκριμένη νόρμα ενδέχεται να μην ευνοεί τη δημιουργία ενός αποτελεσματικού μοντέλου πρόβλεψης. Ωστόσο, σημαντική βελτίωση στην απόδοση σημειώνεται με τη χρήση των αλγορίθμων K-SVD/OMP, Online-OMP/LARS-LASSO και Online/LARS-LASSO, οι οποίοι ενσωματώνουν τις νόρμες ℓ_0 ,ℓ_1 ή $\ell_1 + \ell_2$ για την επιβολή του βαθμού αραιότητας. Σε αυτήν την περίπτωση, η εύρεση των ατόμων του λεξικού δεν υπόκειται σε περιορισμούς κυρτότητας ή μη αρνητικότητας, κάτι που πιθανώς καθιστά αυτά τα μοντέλα ανώτερα από την ανάλυση αρχετύπων. Επιπλέον, η μικρή διαφορά του 6% υπέρ της προσέγγισης ONLINE/LARS-LASSO, μας οδηγεί στο συμπέρασμα ότι η νόρμα ℓ_1 φαίνεται να είναι πιο κατάλληλη για το πρόβλημα που αντιμετωπίστηκε κατά την διάρκεια της παρούσας Δ.Δ.

Ένα σημείο άξιο σχολιασμού, αφορά τη μέθοδο με την οποία δημιουργείται μαθαίνεται το λεξικό, δηλαδή τα άτομα του. Παρά την αρκετά μεγάλη διαφορά στην απόδοσή τους, τα μοντέλα αραιής αναπαράστασης που συνδυάζονται με τις νόρμες ℓ_1 και ℓ_0^{tree} μοιράζονται τον ίδιο αλγόριθμο εκμάθησης λεξικού, αυτόν της online παραγοντοποίησης πινάκων (online matrix factorization). Αυτή η παρατήρηση μας οδηγεί στο συμπέρασμα ότι η υπεροχή του ιεραρχικού μοντέλου πηγάζει από την αρχή λειτουργίας του, η οποία επιβάλλει την επιλογή μη μηδενικών συντελεστών που ανήκουν απαραίτητα σε ένα υποσύνολο του αρχικού δέντρου. Η υποδεέστερη προγνωστική συμπεριφορά των μοντέλων που ενσωματώνουν είτε την τεχνική K-SVD/OMP είτε τα αρχέτυπα μπορεί να ερμηνευτεί παρόμοια. Ένα επιπλέον πλεονέκτημα της προτεινόμενης μεθοδολογίας έγκειται στην ευελιξία που παρέχει στην επιλογή των ατόμων, κατά τη διαδικασία εκμάθησης του λεξικού, μέσω διαφορετικών διαμορφώσεων διακλάδωσης.

Τα μέχρι τώρα συμπεράσματα μπορούν να επιβεβαιωθούν και γραφικά μέσω του σχήματος Σχήμα 13, όπου γίνεται εμφανής ισχυρή προβλεπτική ικανότητα του ιεραρχικού μοντέλου, ειδικά κατά τις βραδινές ώρες, σύμφωνα με τον πίνακα Πίνακας 9. Αυτά τα αποτελέσματα είναι απολύτως λογικά αν λάβουμε υπόψη την υψηλή διείσδυση των ΦΒ που χαρακτηρίζει τα δεδομένα του υποσταθμού. Προφανώς, η λειτουργία των ΦΒ κατά τις ώρες που χαρακτηρίζονται από ηλιοφάνεια, καθιστά την πραγματοποίηση έγκυρων προβλέψεων ως ένα σαφώς πιο δύσκολο αντικείμενο. Προκειμένου να εξεταστεί περαιτέρω η αξιοπιστία του ιεραρχικού μοντέλου, κατασκευάστηκε το διάγραμμα διασποράς του σχήματος Σχήμα 15. Με βάση αυτό το γράφημα συμπεραίνουμε ότι η προτεινόμενη μεθοδολογία καταφέρνει να παρέχει αρκετά ακριβείς προβλέψεις για το σύνολο των δεδομένων ελέγχου. Ωστόσο, όπως υποδεικνύεται από την ευθεία αναφοράς, οι προβλέψεις είναι λιγότερο ακριβείς όταν οι πραγματικές τιμές του μικτού φορτίου είναι μικρές.

Επιπλέον, η προτεινόμενη μέθοδος υπερισχύει όχι μόνο έναντι των ανταγωνιστικών μοντέλων αραιής αναπαράστασης, αλλά και έναντι ορισμένων καθιερωμένων μεθοδολογιών πρόβλεψης. Σύμφωνα με τον πίνακα

Πίνακας **8**, είναι προφανές ότι η αραιή αναπαράσταση με χρήση μιας δενδροειδούς δομής έχει θετική επίδραση στις προβλέψεις του μικτού φορτίου, καθώς η τιμή του MAE που επιτυγχάνεται είναι 17% χαμηλότερη, ενώ του συντελεστή R^2 2.3% υψηλότερη σε σχέση με τις αντίστοιχες τιμές για το μοντέλο LR. Το μοντέλο SVR εμφανίζει παρομοίως κατώτερη απόδοση σε σύγκριση με το προτεινόμενο μοντέλο, με μια μικρή βελτίωση σε σχέση με το LR. Αυτή η παρατήρηση δικαιολογείται, καθώς λαμβάνοντας υπόψη τη γενική μη γραμμική συμπεριφορά του ηλεκτρικού φορτίου, αναμένουμε υψηλότερες επιδόσεις στην ακρίβεια της πρόβλεψης από μη γραμμικές μεθοδολογίες. Ωστόσο, κάτι ανάλογο δεν συμβαίνει με τα νευρωνικά δίκτυα. Η υποβαθμισμένη απόδοση του μοντέλου MLP, σε σύγκριση με τα LR και SVR, που αποδεικνύεται από τις τιμές των μέτρων αξιολόγησης οφείλεται πιθανώς στο φαινόμενο της υπερ-προσαρμογής, παρόλο που εφαρμόστηκε η τεχνική της διασταυρούμενης αξιολόγησης κατά την εκπαίδευση του μοντέλου. Συμπερασματικά, το ιεραρχικό μοντέλο αραιής αναπαράστασης επιτυγχάνει μια αξιοσημείωτη υπεροχή έναντι τόσο των γραμμικών όσο και των μη γραμμικών μεθόδων. Με τον τρόπο αυτό επικυρώνεται πειραματικά η πεποίθηση ότι η εκμάθηση ενός λεξικού που βασίζεται σε δενδροειδείς διαμορφώσεις της αραιής αναπαράστασης, αποτελεί ένα ισχυρό και εύρωστο εργαλείο εκπαίδευσης, που οδηγεί σε ένα μοντέλο υψηλής προβλεπτικής ακρίβειας, αποφεύγοντας το μειονέκτημα της ευαισθησίας στο θόρυβο των δεδομένων που χαρακτηρίζει τις μη γραμμικές μεθόδους.

5.1.2 Μοντέλο ανσάμπλ

Η επόμενη μελέτη περίπτωσης που πραγματοποιήθηκε κατά την εκπόνηση της Δ.Δ. αφορούσε την προτεινόμενη μέθοδο ανσάμπλ. Τα αποτελέσματα που προέκυψαν εξηγούνται και συζητούνται στο σημείο αυτό. Από τα σχήματα: Σχήμα 16.α-Σχήμα 16.στ μπορεί κανείς να εξάγει συμπεράσματα για την ποιότητα των προβλέψεων. Όταν ο χρονικός ορίζοντας πρόβλεψης είναι πολύ μικρός (Σχήμα 16.α), τα σφάλματα πρόβλεψης κατανέμονται κοντά στη διαγώνια γραμμή, πράγμα που συνεπάγεται αρκετά ακριβείς προβλέψεις. Αυξάνοντας όμως τον ορίζοντα πρόβλεψης, οι προβλέψεις γίνονται λιγότερο ακριβείς (Σχήμα 16.β, 16.γ, 16.δ, 16.ε, 16.στ), καθώς όπως φαίνεται και οπτικά, τα ζεύγη πραγματικών τιμών και προβλέψεων είναι διάσπαρτα πιο μακριά από την ευθεία αναφοράς.

Η εξέταση του πίνακα Πίνακας 10, δείχνει ότι το προτεινόμενο μοντέλο ανσάμπλ υπερτερεί όλων των μεμονωμένων υπο-μοντέλων και του ανταγωνιστικού μοντέλου ανσάμπλ MLP σύμφωνα με τους δείκτες MAE, R^2 και RMSE. Επιπλέον, αυτό το συμπέρασμα ισχύει για όλους τους χρονικούς ορίζοντες πρόβλεψης. Καθώς ο ορίζοντας πρόβλεψης μεγαλώνει, το σφάλμα πρόβλεψης αυξάνεται, κάτι που είναι απολύτως εύλογο. Οι μόνες εξαιρέσεις είναι οι τιμές R^2 και RMSE που λαμβάνονται από το μοντέλο ανσάμπλ MLP για ορίζοντα πρόβλεψης 2 ωρών, οι οποίες υπερβαίνουν ελάχιστα αυτές του προτεινόμενου μοντέλου. Ωστόσο, οι διαφορές αυτές δεν μπορούν να θεωρηθούν σημαντικές καθώς είναι οριακές, ενώ από την άλλη η αντίστοιχη τιμή του δείκτη MAE ευνοεί σαφώς την προτεινόμενη μέθοδο. Ένα αποτέλεσμα που αξίζει να σχολιαστεί είναι ότι η αύξηση στην διαφορά της απόδοσης μεταξύ του προτεινόμενου μοντέλου ανσάμπλ και του τρέχοντος καλύτερου υπο-μοντέλου, καθώς ο ορίζοντας μεγαλώνει. Πιο συγκεκριμένα, η μείωση του MAE που επιτυγχάνει η προτεινόμενη προσέγγιση ανσάμπλ σε σχέση με το καλύτερο από τα επιμέρους μοντέλα, κυμαίνεται από 0.0341 έως 0.316. Μια τέτοια βελτίωση της απόδοσης θα μπορούσε εν μέρει να εξηγηθεί από την εμφάνιση αβεβαιότητας στις χρονοσειρές του μικτού φορτίου. Καθώς ο ορίζοντας πρόβλεψης μεγαλώνει, το επίπεδο αβεβαιότητας επίσης αυξάνεται, ένα φαινόμενο το οποίο αντιμετωπίζεται πιο αποτελεσματικά από το μοντέλο ανσάμπλ.

Όσον αφορά την αποτελεσματικότητα των επιμέρους μοντέλων του ανσάμπλ των υπο-μοντέλων, οι τιμές των MAE, RMSE και R^2 δείχνουν ότι δεν υπάρχει μόνο ένα μοντέλο που να υπερισχύει των άλλων σε όλες τις περιπτώσεις. Για τους μικρότερους ορίζοντες πρόβλεψης και πιο συγκεκριμένα μέχρι 3 ώρες, τα μοντέλα LR και SR φαίνεται να επιτυγχάνουν οριακά μικρότερα σφάλματα πρόβλεψης από τα μη γραμμικά μοντέλα. Αν και οι μη γραμμικότητες είναι ένα εγγενές χαρακτηριστικό του μικτού φορτίου [364], αυτή η ιδιότητα γίνεται πιο εμφανής καθώς ο ορίζοντας πρόβλεψης μεγαλώνει. Ως εκ τούτου, τα μοντέλα που βασίζονται στην μέθοδο LR, είναι σε θέση να παρέχουν εύρωστα αποτελέσματα για πολύ βραχυπρόθεσμες προβλέψεις.

Από την άλλη πλευρά, ένα σημαντικό πλεονέκτημα των νευρωνικών δικτύων είναι η ικανότητά τους να μοντελοποιούν μη γραμμικά συστήματα. Μια σημαντική παρατήρηση είναι ότι τα νευρωνικά δίκτυα φαίνεται να αποδίδουν καλύτερα για μεγαλύτερους ορίζοντες πρόβλεψης κάτι που μπορεί να αποδοθεί στο γεγονός ότι, καθώς ο ορίζοντας πρόβλεψης μεγαλώνει, οι μη γραμμικές ιδιότητες του φορτίου γίνονται πιο ισχυρές. Σαν αποτέλεσμα, τα νευρωνικά δίκτυα MLP αποκτούν προβάδισμα στην προβλεπτική συμπεριφορά, όταν απαιτούνται προβλέψεις για μεγαλύτερους ορίζοντες. Ωστόσο, το ίδιο δεν ισχύει για τα δίκτυα RBF. Όπως αναφέρθηκε παραπάνω, για να έχουν καλή απόδοση τα δίκτυα RBF, απαιτούνται πυκνά και καλής ποιότητας δεδομένα. Κατά συνέπεια, η απόδοσή τους μειώνεται για ορίζοντες προβλέψεων 24 ωρών, όπου οι πληροφορίες των δεδομένων εισόδου είναι λιγότερο επαρκείς λόγω της διαδικασίας επαναδειγματοληψίας. Αν και τα υπόλοιπα υπο-μοντέλα, SVR και RF παρουσιάζουν μέτρια ικανότητα πρόβλεψης, συμβάλλουν θετικά στη συνολική απόδοση του προτεινόμενου μοντέλου. Αυτό το συμπέρασμα επιβεβαιώνει τον ισχυρισμό μας για την ανάγκη χρήσης πολλαπλών μοντέλων προκειμένου να ενισχυθεί η αξιοπιστία των προβλέψεων του φορτίου.

Παρατηρώντας τα συγκριτικά αποτελέσματα των πραγματικών τιμών και των τιμών προβλέψεων στο Σχήμα 18, επιβεβαιώνεται ότι, όσο αυξάνεται ο χρονικός ορίζοντας, η ακριβής πρόβλεψη του μικτού φορτίου γίνεται όλο και πιο δύσκολη. Συνεχίζοντας με τα επιμέρους γραφήματα του σχήματος Σχήμα 18, (επισήμανση «II») που απεικονίζουν την εναλλαγή μεταξύ των υπο-μοντέλων με στόχο την βελτίωση της ακρίβειας των τελικών προβλέψεων, εικάζουμε ότι ο μηχανισμός στάθμισης του προτεινόμενου μοντέλου φαίνεται να αποδίδει επαρκώς ανεξάρτητα από το χρονικό ορίζοντα. Μάλιστα, είναι εμφανές ότι λαμβάνονται αρκετά αξιόπιστες προβλέψεις κατά τη διάρκεια της σταθερά ανοδικής ή καθοδικής μετάβασης των πραγματικών τιμών Ε.Ι.. Αντίθετα, οι προβλέψεις γίνονται λιγότερο ακριβείς όταν η καμπύλη Ε.Ι. παρουσιάζει έντονες ταλαντώσεις (fluctuations).

Αρκετά ενδιαφέροντα συμπεράσματα μπορούν επίσης να εξαχθούν από τα διαγράμματα πίτας του σχήματος Σχήμα 17. Κάθε ποσοστό στις πίτες αντιπροσωπεύει τον βαθμό στον οποίο το αντίστοιχο μοντέλο απέδωσε το βάρος με την υψηλότερη τιμή, ή ισοδύναμα το MAE με τη χαμηλότερη τιμή. Τα υψηλότερα ποσοστά κατάταξης στην πρώτη θέση (πάνω από 18%) ανήκουν στα μοντέλα MLP, RBF και RF και αυτό ισχύει για όλους τους ορίζοντες εκτός από αυτόν των 24 ωρών, όπου το μοντέλο SVR παίρνει τη θέση του RF. Ειδικότερα, η μέθοδος RF επιτυγχάνει χαμηλότερο MAE σε μεγαλύτερο ποσοστό από τις υπόλοιπες όταν ο ορίζοντας πρόβλεψης δεν υπερβαίνει τις 3 ώρες. Για μεγαλύτερους ορίζοντες, τα νευρωνικά δίκτυα RBF ξεπερνούν τα υπόλοιπα υπο-μοντέλα. Μια ενδιαφέρουσα παρατήρηση είναι ότι τα προαναφερθέντα μοντέλα έχουν εξίσου υψηλά ποσοστά στην έκτη θέση. Έτσι, αυτές οι μέθοδοι φαίνεται να επιτυγχάνουν είτε πολύ υψηλή είτε χαμηλή απόδοση. Αυτή η επισήμανση είναι αρκετά σημαντική και ενισχύει έντονα τη χρησιμότητα και την αποτελεσματικότητα της προτεινόμενης μεθόδου. Τα ποσοστά των υπολοίπων υπο-μοντέλων στις περισσότερες περιπτώσεις κατανέμονται στις ενδιάμεσες βαθμολογίες, με εξαίρεση το υψηλό ποσοστό του SVR στην κατάταξη στην έκτη θέση για τον ορίζοντα των 6 ωρών.

5.1.3 Μοντέλο νευρωνικών δικτύων γράφων

Τα αποτελέσματα που προέκυψαν από τις προσομοιώσεις εντός της μελέτης περίπτωσης που πραγματοποιήθηκε εξετάζονται και επεξηγούνται σε αυτή την ενότητα ώστε να εξαχθούν συμπεράσματα σχετικά με την απόδοση αυτής της προτεινόμενης μεθοδολογίας. Όπως αποδεικνύεται από τις τιμές των δεικτών R^2 και MAE του πίνακα Πίνακας 11, τα νευρωνικά δίκτυα γράφων ορατότητας ξεπερνούν όλα τα υπόλοιπα μοντέλα που δοκιμάζονται, επιτυγχάνοντας 2 έως 10 % χαμηλότερη τιμή MAE και 1 έως 5 % υψηλότερη τιμή R^2 σε σύγκριση με την πιο αποτελεσματική ανταγωνιστική μεθοδολογία και ανάλογα με το σενάριο προσομοίωσης. Ένα άμεσο συμπέρασμα επίσης προκύπτει αξιολογώντας τα αποτελέσματα για τα σενάρια που αναφέρονται στις εποχές του έτους. Κρίνοντας από τις αντίστοιχα αναφερόμενες τιμές του MAE, παρατηρείται ότι η ακρίβεια πρόβλεψης της πλειονότητας των εξεταζόμενων μοντέλων χειροτερεύει, καθώς οι εποχές κατατάσσονται από το καλοκαίρι προς την άνοιξη. Αυτό μπορεί να αποδοθεί στις κλιματολογικές συνθήκες της συγκεκριμένης περιοχής στην οποία βρίσκεται και εξυπηρετεί ο υποσταθμός, καθιστώντας την άνοιξη μια εποχή με έντονες διακυμάνσεις στις καιρικές μεταβλητές καθώς και στην παραγωγή και ζήτηση ηλεκτρικής ενέργειας. Ένας επιπλέον παράγοντας ίσως να είναι η παγκόσμια κλιματική αλλαγή, η οποία οδήγησε σε μια μετατόπιση των εποχικών συνθηκών. Κατά την εξέταση της απόδοσης των μοντέλων για τα υπόλοιπα σενάρια προσομοίωσης, φαίνεται ότι, πιο ακριβείς προβλέψεις λαμβάνονται στην περίπτωση του σεναρίου Σ5, ακολουθούμενες από τα σενάρια Σ7 και Σ6. Το αποτέλεσμα αυτό μπορεί να θεωρηθεί αναμενόμενο, καθώς οι εργάσιμες ώρες συνοδεύονται τόσο από αυξημένη ζήτηση όσο και παραγωγή, από συμβατικές και από ανανεώσιμες πηγές ενέργειας και προφανώς από φωτοβολταϊκά. Συνεπώς, εισάγεται ένα σημαντικό ποσοστό αβεβαιότητας και εγγενούς μη γραμμικότητας. Αξίζει να σημειωθεί ότι η εναλλασσόμενη προγνωστική συμπεριφορά σε σχέση με τα διαφορετικά σενάρια δεν είναι τυχαία, καθώς διέπει όλα τα εξεταζόμενα μοντέλα με τον ίδιο τρόπο, γεγονός που επικυρώνει την αξία της διερεύνησης του ηλεκτρικού φορτίου σε ασυνήθιστες συνθήκες κατανάλωσης, όπως αυτών που περιγράφονται από τα σενάρια Σ5, Σ6, Σ7.

Αξιολογώντας τα αποτελέσματα του πίνακα Πίνακας 11 ως προς τις ανταγωνιστικές μεθοδολογίες, παρατηρείται ότι για όλα τα σενάρια προσομοίωσης επικρατούν τα νευρωνικά δίκτυα MLP, εκτός από τα σενάρια του φθινοπώρου και του χειμώνα όπου υπερέχει η γραμμική παλινδρόμηση. Όπως αναμενόταν, τα σύνολα δεδομένων που αντιστοιχούν σε ιδιότυπες συνθήκες κατανάλωσης, παρουσιάζουν ισχυρές μη γραμμικότητες και επομένως είναι επόμενο το μοντέλο MLP να ανταποκρίνεται πιο αποτελεσματικά από τη γραμμική παλινδρόμηση, η οποία, αν και παρέχει ευρωστία, δεν είναι σε θέση να χειριστεί μη γραμμικούς συσχετισμούς. Παραδόξως, η μέθοδος SVR, αν και είναι ικανή να χειρίζεται δεδομένα που διέπονται από γραμμική συμπεριφορά, ωστόσο παρουσιάζει την πιο αδύναμη απόδοση από όλες τις προσεγγίσεις. Η αύξηση του αριθμού των επαναλήψεων κατά τη βελτιστοποίηση των υπερπαραμέτρων του μοντέλου πιθανότατα θα βελτίωνε την απόδοσή της, με το αντάλλαγμα όμως του αυξημένου χρόνου εκτέλεσης, κάτι που θα αμφισβητούσε τη δίκαιη σύγκριση μεταξύ των εξεταζόμενων μεθόδων.

Το θέμα της υπολογιστικής πολυπλοκότητας διαδραματίζει ιδιαίτερο ρόλο στην παρούσα έρευνα, καθώς αποτελεί μια από τις επιζήμιες ιδιότητες της βαθιάς μάθησης. Ωστόσο, όπως δείχνουν οι αντίστοιχες τιμές του πίνακα Πίνακας 11, το προτεινόμενο μοντέλο VGNN επιτυγχάνει ανώτερη απόδοση, διατηρώντας τον υπολογιστικό φόρτο σε συγκρίσιμα επίπεδα με τις άλλες μεθόδους, εκτός φυσικά από τη γραμμική παλινδρόμηση, η οποία απαιτεί ελάχιστους υπολογισμούς. Είναι χρήσιμο να αναφερθεί ότι το μοντέλο GNN απουσιάζει από τα αποτελέσματα των πειραμάτων καθώς ήταν υπολογιστικά αδύνατο να εκτελεστεί από το διαθέσιμο υπολογιστικό σύστημα. Αν και δεν μπορούμε να αποφανθούμε με σιγουριά για την υπεροχή του VGNN έναντι του GNN όσον αφορά την ακρίβεια πρόβλεψης, το απαγορευτικό υπολογιστικό κόστος που απαιτεί το GNN, αναμφίβολα θέτει το προτεινόμενο μοντέλο σε πλεονεκτική θέση.

Προκειμένου να παρέχουμε μια ολοκληρωμένη επισκόπηση της απόδοσης της προτεινόμενης μεθόδου πρόβλεψης, γίνεται μια προσπάθεια αξιολόγησης της ποιότητας των προβλέψεων σύμφωνα με τα σχήματα: Σχήμα 19-Σχήμα 22. Η υψηλή προβλεπτική ικανότητα του μοντέλου VGNN επιβεβαιώνεται από τα σχήματα: Σχήμα 19-Σχήμα 20, τα οποία καταδεικνύουν ότι οι προβλέψεις προσαρμόζονται με αρκετά ικανοποιητική ακρίβεια στη χρονοσειρά του πραγματικού μικτού φορτίου. Επιπλέον, παρατηρώντας κάθε επιμέρους γραφική παράσταση των Σχημάτων 19-20, επιβεβαιώνεται και γραφικά η εγκυρότητα των παραπάνω συμπερασμάτων σχετικά με τις προβλέψεις μικτών φορτίων σε διαφορετικά σενάρια. Για παράδειγμα, είναι προφανές ότι η ακρίβεια των προβλέψεων επιδεινώνεται από εποχή σε εποχή καθώς το μέγιστο απόλυτο σφάλμα αυξάνεται από 0.156 για το σενάριο της εποχής του καλοκαιριού (Σχήμα 20(α.ΙΙ)) σε 0.6273 για το σενάριο της εποχής της άνοιξης (Σχήμα 20((δ .ΙΙ)). Ομοίως, σύμφωνα με το Σχήμα 19(β), τα δεδομένα που αντιστοιχούν στο σενάριο των περιόδων με φως του ήλιου είναι πράγματι τα πιο δύσκολα να μοντελοποιηθούν και αντίστοιχα να προβλεφθούν. Μια γενική παρατήρηση, είναι ότι αρκετά αξιόπιστες προβλέψεις λαμβάνονται κατά τη διάρκεια της σταθερής ανόδου ή πτώσης των τιμών Ε.Ι. Αντίθετα, οι προβλέψεις γίνονται λιγότερο ακριβείς όταν η χρονοσειρά των πραγματικών τιμών Ε.Ι. παρουσιάζει έντονες ταλαντώσεις.

Για την εξαγωγή χρήσιμων συμπερασμάτων σχετικά την αξιοπιστία της προτεινόμενης μεθόδου κατά την πρόβλεψη του μικτού φορτίου με χρονικό ορίζοντα 24 ωρών, χρησιμοποιούμε τα διαγράμματα διασποράς που απεικονίζονται στα σχήματα: Σχήμα 21-Σχήμα 22. Οι προβλέψεις που αντιστοιχούν στο σενάριο που αναφέρεται σε δεδομένα εκπαίδευσης και ελέγχου που καλύπτουν την εποχή του καλοκαιριού (Σχήμα 22.γ) παρουσιάζουν την υψηλότερη ακρίβεια ενώ η μετάβαση μεταξύ σεναρίων που αναφέρονται σε δεδομένα εκπαίδευσης και ελέγχου που καλύπτουν τις επόμενες εποχές του έτους, δείχνει ότι η διασπορά των σφαλμάτων πρόβλεψης γύρω από την ευθεία αναφοράς αυξάνεται, όπως αναμενόταν. Συγκεκριμένα, για τα σενάρια που αναφέρονται σε δεδομένα εκπαίδευσης και ελέγχου που καλύπτουν την εποχή του χειμώνα (Σχήμα 22.α) και άνοιξης (Σχήμα 22.β), επισημαίνεται ότι η μείωση της αποτελεσματικότητας του μοντέλου αποδίδεται στην ύπαρξη μεγάλου αριθμού ακραίων τιμών, πιθανότατα λόγω θορύβου στα σύνολα δεδομένων εκπαίδευσης και ελέγχου. Η πιο φτωχή απόδοση του VGNN σημειώνεται για το σενάριο που αναφέρεται σε δεδομένα χρονικών περιόδων με φως του ήλιου και αντικατοπτρίζεται στο αντίστοιχο διάγραμμα διασποράς (Σχήμα 21.β), όπου παρατηρείται αυξημένο σφάλμα πρόβλεψης τόσο για τις περιοχές μικρών όσο και μεγάλων τιμών Ε.Ι.

5.1.4 Γενικά συμπεράσματα

Στην παράγραφο αυτή θα πραγματοποιηθεί μια αποτίμηση της συνεισφοράς του έργου που παράχθηκε στα πλαίσια της Δ.Δ., στο τομέα της έρευνας του προβλήματος της πρόβλεψης του ηλεκτρικού φορτίου. Όπως αποδεικνύουν τα αποτελέσματα που προέκυψαν από τις μελέτες περίπτωσης που έλαβαν χώρα στα πλαίσια της Δ.Δ., αλλά και η ανάλυση και ο σχολιασμός αυτών, στις πραραγργράφους του παρόντος κεφαλαίου που προηγήθηκαν, οι μέθοδοι που προτείνονται για την πρόβλεψη του μικτού ηλεκτρικού φορτίου είναι πράγματι αποτελεσματικές και καινοτόμες. Πιο συγκεκριμένα, μέσα από τις προσομοιώσεις που διεξάγονται, αντιμετωπίζονται με επιτυχία ορισμένες δύσκολες προκλήσεις, όπως η πρόβλεψη του φορτίου με χρήση δεδομένων που αναφέρονται σε ιδιότυπες συνθήκες κατανάλωσης (βλ. πίνακα Πίνακας 6), και η παροχή αξιόπιστων προβλέψεων για πολλαπλούς χρονικούς ορίζοντες μέσω μιας ενιαίας προσέγγισης. Στο σημείο αυτό είναι χρήσιμο να διευκρινιστεί ότι οι πειραματικές προσομοιώσεις που αφορούν σε κάθε μία από τις προτεινόμενες μεθόδους, χρησιμοποιούν διαφορετικά σύνολα δεδομένων εκπαίδευσης και ελέγχου. Κάτι τέτοιο προκύπτει από τη διαθεσιμότητα των δεδομένων κατά την διεξαγωγή των προσομοιώσεων και το σκοπό αξιοποίησης του εκάστοτε μοντέλου. Για παράδειγμα, μια από τις βασικές συνεισφορές του μοντέλου VGNN, είναι η παροχή αξιόπιστων προβλέψεων για σενάρια που αναφέρονται σε δεδομένα που ενέχουν υψηλά επίπεδα μη γραμμικότητας και αβεβαιότητας. Κατά συνέπεια, εφαρμόστηκαν σύνολα δεδομένων εκπαίδευσης και ελέγχου τα οποία χαρακτηρίζονται από αυτές τις ιδιότητες. Η υψηλή προβλεπτική απόδοση των προτεινόμενων μεθόδων, οι οποίες παρουσιάζονται και εφαρμόζονται για πρώτη φορά για την πρόβλεψη του μικτού φορτίου, επιβεβαιώνεται από την υπεροχή τους έναντι καθιερωμένων γραμμικών και μη γραμμικών μεθόδων Μ.Μ. Αξιολογώντας συνολικά τα αποτελέσματα των πειραματικών προσομοιώσεων, καταλήγουμε στο συμπέρασμα ότι οι προτεινόμενες μέθοδοι είναι σε θέση να παρέχουν ακριβείς προβλέψεις μικτού φορτίου ακόμα και σε περιπτώσεις αξιολόγησης σε δεδομένα με υψηλό ποσοστό αβεβαιότητας και μη γραμμικών συσχετίσεων. Παρόλο που οι τιμές του δείκτη MAE καταδεικνύουν την υψηλή επίδοση των προτεινόμενων μοντέλων, εντούτοις δεν είναι δυνατόν να παρέχουν ένα μέτρο της ακρίβειας των προβλέψεων σε κλίμακα πραγματικού μεγέθους, λόγω της κανονικοποίησης των τιμών Ε.Ι. Ωστόσο, η αξιοπιστία των προτεινόμενων μοντέλων δύναται να επιβεβαιωθεί με βάση τις τιμές του συντελεστή R^2, ο οποίος εκφράζεται ως καθαρός αριθμός και επομένως δεν επηρρεάζεται από την κανονικοποίηση των δεδομένων. Συγκεκριμένα, το μοντέλο ιεραρχικής αραιότητας με ορίζοντα πρόβλεψης 1 ώρας, επιτυγχάνει τιμή του δείκτη R^2 ίση με 0.94 για δεδομένα ελέγχου που καλύπτουν ένα ολόκληρο έτος. Οι αντίστοιχες τιμές για το μοντέλο ανσάμπλ κυμαίνονται από 0.79 έως 0.99 αναφερόμενες σε διαφορετικούς χρονικούς ορίζοντες πρόβλεψης, στο εύρος μεταξύ 15 λεπτών και 24 ωρών. Θα πρέπει να σημειωθεί ότι στην περίπτωση των προσομοιώσεων του μοντέλου ανσάμπλ, το χρονικό διάστημα το οποίο καλύπτει το σύνολο δεδομένων ελέγχου είναι ένα έτος, δεν ταυτίζεται όμως με το αντίστοιχο διάστημα για την περίπτωση του μοντέλου ιεραρχικής αραιότητας. Τέλος, η ακρίβεια του μοντέλου VGNN κατά την πρόβλεψη του φορτίου για σενάρια που αναφέρονται σε χρονικές περιόδους με ιδιότυπες συνθήκες κατανάλωσης, ως προς τον δείκτη R^2, είναι σε κάθε περίπτωση μεγαλύτερη από 0.50. Η αξία της επίτευξης προβλέψεων υψηλής ακρίβειας σε εύρος πολαπλών οριζόντων, κρίνεται ως καθοριστικής σημασίας τόσο για τον διαχειριστή όσο και τους ΦοΣΕ του δικτύου. Συνεπώς, μπορούμε να συμπαιράνουμε ότι οι ερευνητικοί στόχοι της Δ.Δ. επιτυγχάνονται σε ικανοποιητικό βαθμό.

5.2 Θέματα για περαιτέρω έρευνα

Η ενότητα αυτή πραγματεύεται την ανάδειξη μελλοντικών ερευνητικών προτάσεων και προοπτικών που προκύπτουν ως προς τα άρθρα που δημοσιεύτηκαν στα πλαίσια της Δ.Δ. αλλά και σε σχέση με το αντικείμενο της διδακτορικής έρευνας.

Με αφετηρία την υπεροχή του ιεραρχικού μοντέλου αραιότητας για την πρόβλεψη του μικτού ηλεκτρικού φορτίου, μια προτεινόμενη μελλοντική ερευνητική κατεύθυνση θα ήταν η περαιτέρω αξιοποίηση του τρόπου λειτουργίας της μεθόδου, όπως η ανάθεση της πρόβλεψης διαφορετικών μοτίβων φορτίου σε διαφορετικούς κλάδους της δενδροειδούς δομής. Δεδομένου των ενθαρρυντικών αποτελεσμάτων που δημοσιεύθηκαν από την ερευνητική μας ομάδα [161], για την βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φορτίου σε επίπεδο υποσταθμού, η χρήση της αραιής κωδικοποίησης συνιστάται σαφώς για την ανάπτυξη παρόμοιων μοντέλων πρόβλεψης. Μια επερχόμενη πρόκληση των σημερινών δικτύων ηλεκτρικής ενέργειας είναι η ενσωμάτωση συστημάτων αποθήκευσης ενέργειας, προκειμένου να καλυφθούν οι διαρκώς αυξανόμενες ενεργειακές ανάγκες. Για το λόγο αυτό, είναι βέβαιο ότι αξιόπιστα μοντέλα πρόβλεψης σαν το προτεινόμενο μπορούν να αξιοποιηθούν για τη διαχείριση της συμφόρησης φορτίου και τον εντοπισμό των σφαλμάτων (fault detection) του δικτύου, ενισχύοντας σημαντικά την ευστάθεια και την ασφάλεια των μελλοντικών δικτύων διανομής. Λαμβάνοντας υπόψη επίσης, το περιορισμένο πεδίο εφαρμογής των μεθόδων αραιής κωδικοποίησης στον τομέα της ηλεκτρικής ενέργειας, η επέκταση της χρήσης τους σε πρόσθετους τομείς, όπως η πρόβλεψη παραγωγής ενέργειας από ΑΠΕ, καθώς και η μακροπρόθεσμη πρόβλεψη του μικτού φορτίου και της άεργου ισχύος, θεωρείται ως μια υποσχόμενη ερευνητική πρόταση. Μια καρποφόρα αναβάθμιση από μεθοδολογικής σκοπιάς, θα ήταν η πιθανή επέκταση των αρχών αραιής κωδικοποίησης σε μεθόδους βαθιάς μάθησης. Για την υλοποίηση αυτής της πρότασης, απαιτείται η ενσωμάτωση δύο σχετικών παραμετρικών και γενικευμένων μοντέλων για τη διαχείριση σημάτων: α) την Συνελικτική Αραιή Κωδικοποίηση (Convolutional Sparse Coding - CSC) και β) την Πολυεπίπεδη Συνελικτική Αραιή Κωδικοποίηση (multi-layered Convolutional Sparse Coding - ML-CSC) [240], η οποία συνδυάζεται αποτελεσματικά με μια σειρά από κοινές αρχιτεκτονικές βαθιάς μάθησης.

Η διερεύνηση της βραχυπρόθεσμης πρόβλεψης του μικτού φορτίου στα πλαίσια της Δ.Δ. επεκτείνεται με την ανάπτυξη της προτεινόμενης μεθοδολογίας ανσάμπλ. Παρόλο που η εξέταση χρονικών οριζόντων πρόβλεψης μεγαλύτερων των 24 ωρών ξεπερνούσε τους σκοπούς της αντίστοιχης εργασίας ως μέρος ενός ερευνητικού προγράμματος που προσανατολιζόταν στην αξιοποίηση των προβλέψεων για τον ημερήσιο χρονοπρογραμματισμό των πηγών παραγωγής του δικτύου, πιστεύουμε ότι το προτεινόμενο μοντέλο θα μπορούσε να χρησιμεύσει ως βάση για το σχεδιασμό ενός τέτοιου εργαλείου πρόβλεψης. Από την άλλη πλευρά, είναι πολύ πιθανό ότι σε αυτή την περίπτωση θα χρειαζόταν ένα διαφορετικό σύνολο μεταβλητών εισόδου, που παρουσιάζουν υψηλότερη συσχέτιση με τη μακροπρόθεσμη εξέλιξη του μικτού φορτίου. Με γνώμονα την αξιοσημείωτη απόδοση της προτεινόμενης μεθοδολογίας στην πρόβλεψη μικτού φορτίου, η εφαρμογή της θα μπορούσε να επεκταθεί και σε άλλους κρίσιμους τομείς του έξυπνου δικτύου, όπως η πρόβλεψη της τιμής της ηλεκτρικής ενέργειας και της παραγωγή από ΑΠΕ, προκειμένου να προγραμματιστεί πιο αποτελεσματικά η παραγωγή ενέργειας από τις συμβατικές πηγές.

Αναγνωρίζοντας τα αξιοσημείωτα αποτελέσματα του προτεινόμενου μοντέλου VGNN και λαμβάνοντας υπόψη την περιορισμένη βιβλιογραφική έκταση των νευρωνικών δικτύων γράφων στον τομέα της πρόβλεψης φορτίου, μια άμεση ερευνητική πρόταση θα ήταν η εντατικοποίηση της εφαρμογής τους στον συγκεκριμένο τομέα αλλά και για τη διερεύνηση εργασιών βελτιστοποίησης του έξυπνου δικτύου. Επίσης, αξιοποιώντας το πλεονέκτημα της σημαντικά μειωμένης υπολογιστικής πολυπλοκότητας που παρέχει η προτεινόμενη μεθοδολογία, η χρήση της σε εφαρμογές όπου εμπλέκονται ζητήματα καθυστέρησης επικοινωνίας, θα μπορούσε σίγουρα να αποτελέσει μια μελλοντική κατεύθυνση για έρευνα. Εφαρμογές όπως η παροχή προβλέψεων για την κατανάλωση ηλεκτρικής ενέργειας [365] ή η ταξινόμηση ανωμαλιών ποιότητας ισχύος [366], στο πεδίο καθώς και εκείνες που σχετίζονται με edge computing [256], [367], [368], αναμένεται να μονοπωλήσουν το ενδιαφέρον για τη διαχείριση κατανεμημένων έξυπνων δικτύων κατά τα επόμενα έτη. Η απόδοση της προτεινόμενης μεθόδου αποδεικνύει επίσης ότι οι γράφοι ορατότητας αποτελούν μια αποτελεσματική και υποσχόμενη τεχνική η οποία, σε συνδυασμό με μεθόδους μηχανικής μάθησης, είναι ικανή να οδηγήσει σε αναβαθμισμένα μοντέλα. Πρόβλεψης της κατανάλωσης της ηλεκτρικής ενέργειας. Επίσης, λαμβάνοντας υπόψη τον περιορισμένο αριθμό παρόμοιων προσεγγίσεων μέχρι στιγμής, συνιστάται ιδιαίτερα η στόχευση στην ανάπτυξη τέτοιου είδους υβριδικών μοντέλων.

Παρότι το έργο που παρουσιάζεται στην παρούσα έρευνα προσφέρει μια αρκετά σημαντική πηγή μιας σειράς αξιόπιστων και αποδοτικών μεθόδων Μ.Μ. για την πρόβλεψη του ηλεκτρικού φορτίου, το εν λόγω πεδίο θεωρείται και παραμένει ανοικτό προς διερεύνηση. Συνεπώς, τα αποτελέσματα και τα συμπεράσματα της Δ.Δ. θα μπορούσαν να αποτελέσουν το έναυσμα για περαιτέρω ερευνητική δραστηριότητα. Ένα πρώτο βήμα θα μπορούσε να είναι ο συνδυασμός των προτεινόμενων μεθόδων. Συγκεκριμένα, το μοντέλο ανσάμπλ θα μπορούσε να αναβαθμιστεί σημαντικά με την ενσωμάτωση των νευρωνικών δικτύων γράφων ορατότητας στο σύνολο των υπο-μοντέλων. Λόγω της ικανότητας των GNN να αναπαριστούν επιτυχώς τόσο τα χωρικά όσο και τα χρονικά χαρακτηριστικά των δεδομένων εισόδου, η αξία ενός τέτοιου μοντέλου θα επεκτεινόταν στην πρόβλεψη της οικιακής ζήτησης ή του συγκεντρωμένου φορτίου που αντιστοιχεί σε αρκετούς υποσταθμούς. Αν και οι στόχοι της Δ.Δ. περιορίζονται στην διερεύνηση μεθόδων επιβλεπόμενης και μη επιβλεπόμενης μάθησης, μια σειρά από πλεονεκτήματα που χαρακτηρίζουν τους αλγόριθμους εκμάθησης με ενίσχυση όπως η αποδοτική λειτουργία τους ακόμη και με λίγα δεδομένα και η προσαρμοστικότητά τους σε σχέση με εκείνους που βασίζονται στη μάθηση με επίβλεψη, εγείρουν το έντονο ενδιαφέρον για την ανάπτυξη τέτοιων μοντέλων με σκοπό την πρόβλεψη του ηλεκτρικού φορτίου. Στην προσπάθεια ανάπτυξης καινοτόμων προσεγγίσεων για την αντιμετώπιση προβλημάτων παλινδρόμησης, όπως η πρόβλεψη του ηλεκτρικού φορτίου, θα μπορούσε κανείς να στραφεί σε εναλλακτικές τεχνικές όπως η εκμάθηση σε μη διανυσματικούς χώρους πολυπτύχων (manifold learning). Πολλές από τις κλασσικές μεθοδολογίες πρόβλεψης χρονοσειρών όπως ARMA [369], GARCH [370], Τ.Ν.Δ [371], SVR [372] βασίζονται στην παραμετρική μοντελοποίηση, η οποία υποθέτει η σχέση μεταξύ μεταβητών εισόδου και εξόδου αναπαρίσταται από μια παραμετρική συνάρτηση με γνωστή μορφή. Τα κύρια μειονεκτήματα και περιορισμοί αυτής της μοντελοποίησης είναι ότι απαιτεί πολύ αυστηρές παραμετρικές υποθέσεις και χρονική ομοιογένεια κατά την εξέλιξη της της χρονοσειράς. Οι σύγχρονες τεχνικές πρόβλεψης χρονοσειρών χρησιμοποιούν πιο ευέλικτα μη παραμετρικά ή ημι-παραμετρικά μοντέλα, όπως: οι Μπεϋζιανές μέθοδοι [373], η πρόβλεψη με βάση τον πυρήνα (kernel-based) [374], η βαθιά μάθηση (deep learning) [375] και η ενισχυτική μάθηση (reinforcement learning) [376]. Ωστόσο, για πολυμεταβλητές χρονοσειρές, όλες οι αναφερόμενες προσεγγίσεις υποφέρουν από την κακόφημη «*κατάρα* *της διαστατικότητας*» (curse of dimensionality). Για το σκοπό αυτό, μπορούν να χρησιμοποιηθούν διάφορες μέθοδοι εκμάθησης πολυπτύχων (manifold learning) και τεχνικές μείωσης των διαστάσεων όπως η ανάλυση κύριων συσνιστωσών (principal components analysis – PCA) [377] και η multidimensional scaling (MDS) [378]. Δυστυχώς, οι προαναφερθείσες μέθοδοι μείωσης των διαστάσεων υποθέτουν ότι οι παρατηρήσεις βρίσκονται ακριβώς πάνω σε μια γραμμική πολλαπλότητα, όπως μια ευθεία, ένα επίπεδο ή ένα υπερεπίπεδο ανάλογα με τη διάσταση του χώρου των δεδομένων εισόδου, και συχνά παρουσιάζουν χαμηλή απόδοση όταν υπάρχει θόρυβος στα δεδομένα εισόδου. Επιπλέον, τα δεδομένα των πραγματικών εφαρμογών σπανίως βρίσκονται πάνω σε οποιοδήποτε είδους πολλαπλότητα. Η μη γραμμική εκμάθηση πολυπτύχων περιλαμβάνει τεχνικές όπως οι Χάρτες διάχυσης (Diffusion Maps) [379], η ISOMAP [380] κ.α. που χαλαρώνουν την παραπάνω απαίτηση, επιδιώκοντας την εκμάθηση της απεικόνισης μιας μη γραμμικής πολλαπλότητας από έναν πολυδιάστατο χώρο σε ένα χώρο λίγων διαστάσεων, και μπορούν να εξάγουν από τα δεδομένα ποιοτικές πληροφορίες [381], [382] που αναμένεται να οδηγήσουν σε πιο αξιόπιστες προβλέψεις.

ΑΝΑΦΟΡΕΣ

[1] A. K. Srivastava, A. S. Pandey, and D. Singh, "Short-term load forecasting methods: A review," in *International Conference on Emerging Trends in Electrical, Electronics and Sustainable Energy Systems, ICETEESES 2016*, 2016. doi: 10.1109/ICETEESES.2016.7581373.

[2] K. Zor, O. Timur, and A. Teke, "A state-of-the-art review of artificial intelligence techniques for short-term electric load forecasting," in *2017 6th International Youth Conference on Energy (IYCE)*, IEEE, Jun. 2017, pp. 1–7. doi: 10.1109/IYCE.2017.8003734.

[3] A. S. Semenov, M. N. Semenova, and Y. V. Bebikhov, "Development of Universal Mathematical Model of Electrical Power Supply System of Area of Industrial Enterprise," in *Proceedings - 2019 International Russian Automation Conference, RusAutoCon 2019*, 2019. doi: 10.1109/RUSAUTOCON.2019.8867704.

[4] S. Cieślik, "Mathematical modeling of the dynamics of linear electrical systems with parallel calculations," *Energies (Basel)*, 2021, doi: 10.3390/en14102930.

[5] M. S. Rizvi, P. Kumar, D. S. Katti, and A. Pal, "Mathematical model of mechanical behavior of micro/nanofibrous materials designed for extracellular matrix substitutes," *Acta Biomater*, 2012, doi: 10.1016/j.actbio.2012.07.025.

[6] I. Pavlenko, J. Trojanowska, V. Ivanov, and O. Liaposhchenko, "Scientific and methodological approach for the identification of mathematical models of mechanical systems by using artificial neural networks," in *Lecture Notes in Electrical Engineering*, 2019. doi: 10.1007/978-3-319-91334-6_41.

[7] S. V. Trigorly and V. V. Zakharov, "Mathematical Modeling of the Processes of Microwave Thermal Processing of Extended Dielectrics with the Use of Mobile Radiation Systems," in *2018 International Conference on Actual Problems of Electron Devices Engineering, APEDE 2018*, 2018. doi: 10.1109/APEDE.2018.8542453.

[8] R. Busqué, R. Torres, J. Grau, V. Roda, and A. Husar, "Mathematical modeling, numerical simulation and experimental comparison of the desorption process in a metal hydride hydrogen storage system," *Int J Hydrogen Energy*, 2018, doi: 10.1016/j.ijhydene.2017.12.172.

[9] N. V. Torres and G. Santos, “The (mathematical) modeling process in biosciences,” *Frontiers in Genetics*. 2015. doi: 10.3389/fgene.2015.00354.

[10] A. El Khaouat and L. Benhlima, “Big data based management for smart grids,” in *Proceedings of 2016 International Renewable and Sustainable Energy Conference, IRSEC 2016*, 2017, pp. 1044–1047. doi: 10.1109/IRSEC.2016.7983902.

[11] G. D. F. Morales, A. Bifet, L. Khan, J. Gama, and W. Fan, “IoT big data stream mining,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016. doi: 10.1145/2939672.2945385.

[12] W. Y. C. Wang and Y. Wang, “Analytics in the era of big data: The digital transformations and value creation in industrial marketing,” *Industrial Marketing Management*. 2020. doi: 10.1016/j.indmarman.2020.01.005.

[13] C. Brewis, S. Dibb, and M. Meadows, “Leveraging big data for strategic marketing: A dynamic capabilities model for incumbent firms,” *Technol Forecast Soc Change*, 2023, doi: 10.1016/j.techfore.2023.122402.

[14] S. Gupta, T. Justy, S. Kamboj, A. Kumar, and E. Kristoffersen, “Big data and firm marketing performance: Findings from knowledge-based view,” *Technol Forecast Soc Change*, 2021, doi: 10.1016/j.techfore.2021.120986.

[15] Y. Sakurai, Y. Matsubara, and C. Faloutsos, “Mining Big Time-series Data on the Web,” in *WWW 2016 Companion - Proceedings of the 25th International Conference on World Wide Web*, 2016. doi: 10.1145/2872518.2891061.

[16] M. Z. Kastouni and A. Ait Lahcen, “Big data analytics in telecommunications: Governance, architecture and use cases,” *Journal of King Saud University - Computer and Information Sciences*. 2022. doi: 10.1016/j.jksuci.2020.11.024.

[17] J. P. de Kock and L. Tarassenko, “Pulse oximetry: Theoretical and experimental models,” *Med Biol Eng Comput*, 1993, doi: 10.1007/BF02458049.

[18] R. Laubscher and R. T. Dobson, “Theoretical and experimental modelling of a heat pipe heat exchanger for high temperature nuclear reactor technology,” *Appl Therm Eng*, 2013, doi: 10.1016/j.applthermaleng.2013.06.063.

[19] R. Isermann, *Mechatronic systems: Fundamentals*. 2005. doi: 10.1007/1-84628-259-4.

[20] P. D. Moncarz and H. Krawinkler, "THEORY AND APPLICATION OF EXPERIMENTAL MODEL ANALYSIS IN EARTHQUAKE ENGINEERING.," *Report - Stanford University, John A. Blume Earthquake Engineering Center*, 1981.

[21] Z. Zhao, "Parametric and nonparametric models and methods in financial econometrics," *Statistics Surveys*. 2007. doi: 10.1214/08-SS034.

[22] N. Wei, L. Yin, C. Li, C. Li, C. Chan, and F. Zeng, "Forecasting the daily natural gas consumption with an accurate white-box model," *Energy*, 2021, doi: 10.1016/j.energy.2021.121036.

[23] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," *ACM Comput Surv*, 2018, doi: 10.1145/3236009.

[24] E. Pintelas, I. E. Livieris, and P. Pintelas, "A Grey-Box ensemble model exploiting Black-Box accuracy and White-Box intrinsic interpretability," *Algorithms*, 2020, doi: 10.3390/a13010017.

[25] L. Mariani, M. Pezzè, and D. Zuddas, "Recent Advances in Automatic Black-Box Testing," in *Advances in Computers*, 2015. doi: 10.1016/bs.adcom.2015.04.002.

[26] C. B. Azodi, J. Tang, and S. H. Shiu, "Opening the Black Box: Interpretable Machine Learning for Geneticists," *Trends in Genetics*. 2020. doi: 10.1016/j.tig.2020.03.005.

[27] L. Ljung, "Black-box models from input-output measurements," in *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, 2001. doi: 10.1109/imtc.2001.928802.

[28] A. Behrend, C. Dorau, R. Manthey, and G. Schüeller, "Incremental view-based analysis of stock market data streams," in *ACM International Conference Proceeding Series*, 2008. doi: 10.1145/1451940.1451978.

[29] P. I. Politopoulos, E. P. Markatos, and S. Ioannidis, "Evaluation of compression of remote network monitoring data streams," in *2008 IEEE Network Operations and Management Symposium Workshops - NOMS 08*, 2008. doi: 10.1109/NOMSW.2007.20.

[30] N. Miloslavskaya, "Stream Data Analytics for Network Attacks' Prediction," in *Procedia Computer Science*, 2020. doi: 10.1016/j.procs.2020.02.114.

[31] I. H. Sarker, A. Colman, M. A. Kabir, and J. Han, "Individualized Time-Series Segmentation for Mining Mobile Phone User Behavior," *Computer Journal*, 2018, doi: 10.1093/comjnl/bxx082.

[32] M. Couceiro, R. Ferrando, D. Manzano, and L. Lafuente, "Stream analytics for utilities. Predicting power supply and demand in a smart grid," in *2012 3rd International Workshop on Cognitive Information Processing, CIP 2012*, 2012. doi: 10.1109/CIP.2012.6232904.

[33] S. Gupta, U. Ojha, and V. S. DIxit, "Personalized web recommendations analyzing sequential behaviour using implicit data streams: A survey," in *8th International Conference on Computing, Communications and Networking Technologies, ICCCNT 2017*, 2017. doi: 10.1109/ICCCNT.2017.8203976.

[34] M. Halatchev and L. Gruenwald, "Estimating Missing Values in Related Sensor Data Streams," *Data Management*, 2005.

[35] N. Chen, J. Gong, L. Di, and G. Yu, "Automatic On-Demand Data Feed Service for AutoChem Based on Reusable Geo-Processing Workflow," *IEEE J Sel Top Appl Earth Obs Remote Sens*, 2010, doi: 10.1109/JSTARS.2010.2049094.

[36] J. Liu, F. Zhu, C. Chai, Y. Luo, and N. Tang, "Automatic data acquisition for deep learning," in *Proceedings of the VLDB Endowment*, 2021. doi: 10.14778/3476311.3476333.

[37] J. Gama and M. M. Gaber, *Learning from data streams: Processing techniques in sensor networks*. 2007. doi: 10.1007/3-540-73679-4.

[38] S. Nittel, "Real-time sensor data streams," *SIGSPATIAL Special*, 2015, doi: 10.1145/2826686.2826691.

[39] D. Crankshaw, X. Wang, G. Zhou, M. J. Franklin, J. E. Gonzalez, and I. Stoica, "Clipper: A low-latency online prediction serving system," in *Proceedings of the 14th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2017*, 2017.

[40] E. Frigó, R. Pálovics, D. Kelen, L. Kocsis, and A. A. Benczúr, "Online ranking prediction in non-stationary environments," in *CEUR Workshop Proceedings*, 2017.

[41] L. Li, J. Silva, M. Zhou, and L. Carin, "Online Bayesian dictionary learning for large datasets," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2012. doi: 10.1109/ICASSP.2012.6288339.

[42] K. Bhaskaran, A. Gasparrini, S. Hajat, L. Smeeth, and B. Armstrong, "Time series regression studies in environmental epidemiology," *Int J Epidemiol*, 2013, doi: 10.1093/ije/dyt092.

[43] D. C. Montgomery, C. L. Jennings, and M. Kulahci, *Introduction to time series analysis and forecasting*, 2nd ed. Hoboken, New Jersey: John Wiley & Sons, Inc., 2015.

[44] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," *Applied Soft Computing Journal*, 2020, doi: 10.1016/j.asoc.2020.106181.

[45] M. Mudassir, S. Bennbaia, D. Unal, and M. Hammoudeh, "Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach," *Neural Comput Appl*, 2020, doi: 10.1007/s00521-020-05129-6.

[46] K. Yang, H. Jin, X. Chen, J. Dai, L. Wang, and D. Zhang, "Soft sensor development for online quality prediction of industrial batch rubber mixing process using ensemble just-in-time Gaussian process regression models," *Chemometrics and Intelligent Laboratory Systems*, 2016, doi: 10.1016/j.chemolab.2016.04.009.

[47] M. Talaat, M. A. Farahat, N. Mansour, and A. Y. Hatata, "Load forecasting based on grasshopper optimization and a multilayer feed-forward neural network using regressive approach," *Energy*, vol. 196, p. 117087, Apr. 2020, doi: 10.1016/J.ENERGY.2020.117087.

[48] M. Murat, I. Malinowska, M. Gos, and J. Krzyszczak, "Forecasting daily meteorological time series using ARIMA and regression models," *Int Agrophys*, 2018, doi: 10.1515/intag-2017-0007.

[49] M. Monti, J. Fiorentino, E. Milanetti, G. Gosti, and G. G. Tartaglia, "Prediction of Time Series Gene Expression and Structural Analysis of Gene Regulatory Networks Using Recurrent Neural Networks," *Entropy*, 2022, doi: 10.3390/e24020141.

[50] R. Kumar, A. Jain, A. K. Tripathi, and S. Tyagi, "Covid-19 outbreak: An epidemic analysis using time series prediction model," in *Proceedings of the Confluence 2021: 11th International Conference on Cloud Computing, Data Science and Engineering*, 2021. doi: 10.1109/Confluence51648.2021.9377075.

[51] S. Bisgaard and M. Kulahci, "Quality quandaries: Using a time series model for process adjustment and control," *Qual Eng*, 2008, doi: 10.1080/08982110701771042.

[52] D. Ferreira, A. Braga Reis, C. Senna, and S. Sargento, "A Forecasting Approach to Improve Control and Management for 5G Networks," *IEEE Transactions on Network and Service Management*, 2021, doi: 10.1109/TNSM.2021.3056222.

[53] C. Bennett, M. Moghimi, M. J. Hossain, J. Lu, and R. A. Stewart, "Applicability of load forecasting techniques for customer energy storage control systems," in *Asia-Pacific Power and Energy Engineering Conference, APPEEC*, 2016. doi: 10.1109/APPEEC.2015.7380906.

[54] V. F. Barabanov, O. J. Kravets, I. N. Kryuchkova, O. Y. Makarov, A. K. Pogodayev, and O. N. Choporov, "Discrete processes dynamics neural network simulation based on multivariate time series analysis with significant factors delayed influence consideration," *World Appl Sci J*, 2013, doi: 10.5829/idosi.wasj.2013.23.09.13136.

[55] A. Honrubia-Escribano, F. Jiménez-Buendía, A. Molina-García, J. A. Fuentes-Moreno, E. Muljadi, and E. Gómez-Lázaro, "Analysis of Wind Turbine Simulation Models: Assessment of Simplified versus Complete Methodologies Preprint ANALYSIS OF WIND TURBINE SIMULATION MODELS: ASSESSMENT OF SIMPLIFIED VERSUS COMPLETE METHODOLOGIES," 2015. [Online]. Available: http:/www.osti.gov/scitech

[56] M. Gupta, J. Gao, C. C. Aggarwal, and J. Han, "Outlier Detection for Temporal Data: A Survey," *IEEE Trans Knowl Data Eng*, vol. 26, no. 9, pp. 2250–2267, 2014, doi: 10.1109/TKDE.2013.184.

[57] W. Zhou *et al.*, "Abnormal group user detection in recommender systems using multi-dimension time series," in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, 2017. doi: 10.1007/978-3-319-59288-6_34.

[58] S. Ando and E. Suzuki, "Minimizing response time in time series classification," *Knowl Inf Syst*, 2016, doi: 10.1007/s10115-015-0826-7.

[59] G. He, W. Zhao, and X. Xia, "Confidence-based early classification of multivariate time series with multiple interpretable rules," *Pattern Analysis and Applications*, 2020, doi: 10.1007/s10044-019-00782-7.

[60] A. Gupta, H. P. Gupta, B. Biswas, and T. Dutta, "An Early Classification Approach for Multivariate Time Series of On-Vehicle Sensors in Transportation," *IEEE Transactions on Intelligent Transportation Systems*, 2020, doi: 10.1109/TITS.2019.2957325.

[61] D. Tsourounis, I. Theodorakopoulos, E. N. Zois, and G. Economou, "From text to signatures: Knowledge transfer for efficient deep feature learning in offline signature verification," *Expert Syst Appl*, 2022, doi: 10.1016/j.eswa.2021.116136.

[62] R. Tavenard and S. Malinowski, "Cost-aware early classification of time series," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016. doi: 10.1007/978-3-319-46128-1_40.

[63] R. P. Kristianto, "Modeling of time series data prediction using fruit fly optimization algorithm and triple exponential smoothing," in *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering, ICITISEE 2019*, 2019. doi: 10.1109/ICITISEE48480.2019.9003895.

[64] S. Kaur and M. Rakshit, "Predictability and Analysis of Weekly Temperature using Autoregressive Time Series Models," in *2019 International Conference on Computing, Power and Communication Technologies, GUCON 2019*, 2019.

[65] S. Hansun, "A new approach of moving average method in time series analysis," in *2013 International Conference on New Media Studies, CoNMedia 2013*, 2013. doi: 10.1109/conmedia.2013.6708545.

[66] L. O. F. Sales, A. P. Alencar, and L. L. Ho, "The BerG generalized autoregressive moving average model for count time series," *Comput Ind Eng*, 2022, doi: 10.1016/j.cie.2022.108104.

[67] N. Bosowski, V. Ingle, and D. Manolakis, "Generalized Linear Models for count time series," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2017. doi: 10.1109/ICASSP.2017.7952962.

[68] E. J. Hannan and L. Kavalieris, "Multivariate linear time series models," *Adv Appl Probab*, 1984, doi: 10.2307/1427286.

[69] C. A. Sims, J. H. Stock, and M. W. Watson, "Inference in Linear Time Series Models with some Unit Roots," *Econometrica*, 1990, doi: 10.2307/2938337.

[70] T. Schreiber, "Interdisciplinary application of nonlinear time series methods," *Physics Report*, 1999, doi: 10.1016/S0370-1573(98)00035-0.

[71] O. Karakuş, E. E. Kuruoǐlu, and M. A. Altinkaya, "Long term wind speed prediction with polynomial autoregressive model," in *2015 23rd Signal Processing and Communications Applications Conference, SIU 2015 - Proceedings*, 2015. doi: 10.1109/SIU.2015.7129907.

[72] P. Geng, "Estimation of functional-coefficient autoregressive models with measurement error," *J Multivar Anal*, 2022, doi: 10.1016/j.jmva.2022.105077.

[73] Y. A. Shiferaw, "Regime shifts in the COVID-19 case fatality rate dynamics: A Markov-switching autoregressive model analysis," *Chaos, Solitons and Fractals: X*, 2021, doi: 10.1016/j.csfx.2021.100059.

[74] U. M. Umer, T. Sevil, and G. Sevil, "Forecasting performance of smooth transition autoregressive (STAR) model on travel and leisure stock index," *Journal of Finance and Data Science*, 2018, doi: 10.1016/j.jfds.2017.11.006.

[75] D. K. Huriso, B. B. Anjullo, Y. T. Haile, and D. A. Teni, "Multivariate Generalized Autoregressive Conditional Heteroscedasticity Modeling of the Relationship Between Major Economic Indicators in Ethiopia," *Journal of the Knowledge Economy*, 2023, doi: 10.1007/S13132-023-01422-6.

[76] Y. Sun, A. Han, Y. Hong, and S. Wang, "Threshold autoregressive models for interval-valued time series data," *J Econom*, 2018, doi: 10.1016/j.jeconom.2018.06.009.

[77] L. C. Tsai, E. P. Macalalad, and C. H. Liu, "Taiwan Ionospheric Model (TWIM) prediction based on time series autoregressive analysis," *Radio Sci*, 2014, doi: 10.1002/2014RS005448.

[78] A. Gautam and V. Singh, "Parametric versus non-parametric time series forecasting methods: A review," *Journal of Engineering Science and Technology Review*. 2020. doi: 10.25103/JESTR.133.18.

[79] H. Han, X. Wu, L. Ge, and J. Qiao, "A sludge volume index (SVI) model based on the multivariate local quadratic polynomial regression method," *Chin J Chem Eng*, 2018, doi: 10.1016/j.cjche.2017.08.007.

[80] L. Su and C. Li, "Local functional coefficient autoregressive model for multistep prediction of chaotic time series," *Discrete Dyn Nat Soc*, 2015, doi: 10.1155/2015/329487.

[81] P. Chen, L. Dong, W. Chen, and J. G. Lin, "Outlier detection in adaptive functional-coefficient autoregressive models based on extreme value theory," *Math Probl Eng*, 2013, doi: 10.1155/2013/910828.

[82] G. L. Simpson, "Modelling palaeoecological time series using generalised additive models," *Front Ecol Evol*, 2018, doi: 10.3389/fevo.2018.00149.

[83] B. C. P. Nugroho, J. F. Chan, V. Vananda, H. Lucky, and D. Suhartono, "Detecting Phishing Websites with Non-Parametric Machine Learning," in *2022 International Conference on Informatics Electrical and Electronics, ICIEE 2022 - Proceedings*, 2022. doi: 10.1109/ICIEE55596.2022.10009939.

[84] M. Akhoondzadeh, "A MLP neural network as an investigator of TEC time series to detect seismo-ionospheric anomalies," *Advances in Space Research*, 2013, doi: 10.1016/j.asr.2013.01.012.

[85] I. B. Tijani and R. Akmeliawati, "Support vector regression based friction modeling and compensation in motion control system," *Eng Appl Artif Intell*, vol. 25, no. 5, pp. 1043–1052, Aug. 2012, doi: 10.1016/J.ENGAPPAI.2012.03.018.

[86] B. Bauer and M. Kohler, "On deep learning as a remedy for the curse of dimensionality in nonparametric regression," *Ann Stat*, 2019, doi: 10.1214/18-AOS1747.

[87] J. Schmidt-Hieber, "Nonparametric regression using deep neural networks with relu activation function," *Ann Stat*, 2020, doi: 10.1214/19-AOS1875.

[88] M. Z. Bahri and S. Vahidnia, "Time Series Forecasting Using Smoothing Ensemble Empirical Mode Decomposition and Machine Learning Techniques," in *International Conference on Electrical, Computer, Communications and Mechatronics Engineering, ICECCME 2022*, 2022. doi: 10.1109/ICECCME55909.2022.9988336.

[89] K. Singh and S. Upadhyaya, "Outlier Detection: Applications And Techniques.," *International Journal of Computer …*, 2012.

[90] G. K. Vishwakarma, C. Paul, and A. M. Elsawah, "An algorithm for outlier detection in a time series model using backpropagation neural network," *J King Saud Univ Sci*, 2020, doi: 10.1016/j.jksus.2020.09.018.

[91] A. Blázquez-García, A. Conde, U. Mori, and J. A. Lozano, "A Review on Outlier/Anomaly Detection in Time Series Data," *ACM Computing Surveys*. 2021. doi: 10.1145/3444690.

[92] K. Shaukat *et al.*, "A Review of Time-Series Anomaly Detection Techniques: A Step to Future Perspectives," in *Advances in Intelligent Systems and Computing*, 2021. doi: 10.1007/978-3-030-73100-7_60.

[93] D. Peña and V. J. Yohai, "A Review of Outlier Detection and Robust Estimation Methods for High Dimensional Time Series Data," *Econom Stat*, 2023, doi: 10.1016/j.ecosta.2023.02.001.

[94] M. Krishnamoorthy, M. S. A. Hameed, T. Kopinski, and A. Schwung, "Disease Prediction Based on Individual's Medical History Using CNN," in *Proceedings - 20th IEEE International Conference on Machine Learning and Applications, ICMLA 2021*, 2021. doi: 10.1109/ICMLA52953.2021.00022.

[95] H. D. Park, Y. Han, and J. H. Choi, "Medical Time-series Prediction with LSTM-MDN-ATTN," in *ICTC 2019 - 10th International Conference on ICT Convergence: ICT Convergence Leading the Autonomous Future*, 2019. doi: 10.1109/ICTC46691.2019.8939761.

[96] J. V. Hansen and R. D. Nelson, "Neural networks and traditional time series methods: A synergistic combination in state economic forecasts," *IEEE Trans Neural Netw*, 1997, doi: 10.1109/72.595884.

[97] F. C. Yuan, "Intelligent forecasting of inbound tourist arrivals by social networking analysis," *Physica A: Statistical Mechanics and its Applications*, 2020, doi: 10.1016/j.physa.2020.124944.

[98] S. Chapaloglou *et al.*, "Smart energy management algorithm for load smoothing and peak shaving based on load forecasting of an island's power system," *Appl Energy*, 2019, doi: 10.1016/j.apenergy.2019.01.102.

[99] K. Pun, S. M. S. Basnet, and W. Jewell, "Solar Power Prediction in Different Forecasting Horizons Using Machine Learning and Time Series Techniques," in *2021 IEEE Conference on Technologies for Sustainability, SusTech 2021*, 2021. doi: 10.1109/SusTech51236.2021.9467464.

[100] R. Cui, J. Wang, and Z. Wang, "Multiplicative Attention Mechanism for Multi-horizon Time Series Forecasting," in *Proceedings of the International Joint Conference on Neural Networks*, 2021. doi: 10.1109/IJCNN52387.2021.9533598.

[101] T. W. Wlodarczyk, "Overview of Time Series Storage and Processing in a Cloud Environment," in *CloudCom 2012 - Proceedings: 2012 4th IEEE International Conference on Cloud Computing Technology and Science*, 2012. doi: 10.1109/CloudCom.2012.6427510.

[102] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," in *Proceedings of 2014 Science and Information Conference, SAI 2014*, 2014. doi: 10.1109/SAI.2014.6918213.

[103] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67–82, 1997, doi: 10.1109/4235.585893.

[104] J. S. Armstrong, "Selecting Forecasting Methods," *SSRN Electronic Journal*, 2012, doi: 10.2139/ssrn.1941247.

[105] M. Zuefle *et al.*, "Autonomic Forecasting Method Selection: Examination and Ways Ahead," *Proceedings - 2019 IEEE International Conference on Autonomic Computing, ICAC 2019*, pp. 167–176, 2019, doi: 10.1109/ICAC.2019.00028.

[106] V. Skala, "Least square method robustness of computations: What is not usually considered and taught," in *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, FedCSIS 2017*, Institute of Electrical and Electronics Engineers Inc., Nov. 2017, pp. 537–541. doi: 10.15439/2017F7.

[107] G. G. V. Douglas C. Montgomery, Elizabeth A. Peck, *Introduction to Linear Regression Analysis*, 5th ed. 2012.

[108] B. Dhaval and A. Deshpande, "Short-term load forecasting with using multiple linear regression," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 4, pp. 3911–3917, 2020, doi: 10.11591/ijece.v10i4.pp3911-3917.

[109] Z. Dong, D. Yang, T. Reindl, and W. M. Walsh, "Short-term solar irradiance forecasting using exponential smoothing state space model," *Energy*, 2013, doi: 10.1016/j.energy.2013.04.027.

[110] R. S. Santos and D. Pour Yousefian Barfeh, "Time Series Forecasting for Issuance of Alien Employment Permits by Nationality in the Philippines Using Holt-Winters Method," in *Proceedings of 2019 International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2019*, 2019. doi: 10.1109/ICCIKE47802.2019.9004243.

[111] J. W. Taylor and K. S. Taylor, "Combining probabilistic forecasts of COVID-19 mortality in the United States," *Eur J Oper Res*, 2023, doi: 10.1016/j.ejor.2021.06.044.

[112] M. P. Clements and D. I. Harvey, "Combining probability forecasts," *Int J Forecast*, 2011, doi: 10.1016/j.ijforecast.2009.12.016.

[113] P. J. Brockwell, "On continuous-time threshold ARMA processes," *J Stat Plan Inference*, 1994, doi: 10.1016/0378-3758(94)90210-0.

[114] M. H. Amini, A. Kargarian, and O. Karabasoglu, "ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation," *Electric Power Systems Research*, vol. 140, pp. 378–390, Nov. 2016, doi: 10.1016/j.epsr.2016.06.003.

[115] D. Chikobvu and C. Sigauke, "Regression-SARIMA modelling of daily peak electricity demand in South Africa," *Journal of Energy in Southern Africa*, 2012, doi: 10.17159/2413-3051/2012/v23i3a3169.

[116] P. Cui, Z. Deng, W. Hu, and J. Zhu, "Accurate and Reliable Forecasting using Stochastic Differential Equations," 2021.

[117] S. R. Straja, "Stochastic modeling of stock prices," *Montgomery Investment Technology, Inc*, 1997.

[118] Tanvi, N. Goel, and M. Kaur, "A review of soft computing techniques in biometrics," in *2015 2nd International Conference on Recent Advances in Engineering and Computational Sciences, RAECS 2015*, 2016. doi: 10.1109/RAECS.2015.7453282.

[119] D. Ibrahim, "An Overview of Soft Computing," in *Procedia Computer Science*, 2016. doi: 10.1016/j.procs.2016.09.366.

[120] H. Bhaumik, S. Bhattacharyya, M. Das Nath, and S. Chakraborty, "Hybrid soft computing approaches to content based video retrieval: A brief review," *Appl Soft Comput*, 2016, doi: 10.1016/j.asoc.2016.03.022.

[121] X. Ying, "An Overview of Overfitting and its Solutions," in *Journal of Physics: Conference Series*, 2019. doi: 10.1088/1742-6596/1168/2/022022.

[122] O. Nooruldeen, S. Alturki, M. R. Baker, and A. Ghareeb, "Time Series Forecasting for Decision Making on City-Wide Energy Demand: A Comparative Study," in *2022 International Conference on Decision Aid Sciences and Applications, DASA 2022*, 2022. doi: 10.1109/DASA54658.2022.9765193.

[123] G. Makridis, D. Kyriazis, and S. Plitsos, "Predictive maintenance leveraging machine learning for time-series forecasting in the maritime industry," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*, 2020. doi: 10.1109/ITSC45102.2020.9294450.

[124] U. Ahmed, J. C. W. Lin, and G. Srivastava, "Multivariate time-series sensor vital sign forecasting of cardiovascular and chronic respiratory diseases," *Sustainable Computing:*

Informatics and Systems, vol. 38, no. November 2022, p. 100868, 2023, doi: 10.1016/j.suscom.2023.100868.

[125] P. St-Aubin and B. Agard, “Precision and Reliability of Forecasts Performance Metrics,” *Forecasting*, 2022, doi: 10.3390/forecast4040048.

[126] K. L. Sainani, “Dealing With Non-normal Data,” *PM and R*, vol. 4, no. 12, pp. 1001–1005, 2012, doi: 10.1016/j.pmrj.2012.10.013.

[127] J. F. Donges, R. V. Donner, and J. Kurths, “Testing time series irreversibility using complex network methods,” *Epl*, vol. 102, no. 1, pp. 0–6, 2013, doi: 10.1209/0295-5075/102/10004.

[128] Y. Zou, R. V. Donner, N. Marwan, J. F. Donges, and J. Kurths, “Complex network approaches to nonlinear time series analysis,” *Phys Rep*, vol. 787, pp. 1–97, 2019, doi: 10.1016/j.physrep.2018.10.005.

[129] C. Cheng *et al.*, “Time series forecasting for nonlinear and non-stationary processes: a review and comparative study,” *https://doi.org/10.1080/0740817X.2014.999180*, vol. 47, no. 10, pp. 1053–1071, Oct. 2015, doi: 10.1080/0740817X.2014.999180.

[130] T. Kono, S. Yamaguchi, and T. Nagao, “Time Series Prediction with Dual Reliability: Uncertainty and Explainability,” in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, 2020. doi: 10.1109/SMC42975.2020.9283170.

[131] N. Sun, S. Zhang, T. Peng, J. Zhou, and X. Sun, “A Composite Uncertainty Forecasting Model for Unstable Time Series: Application of Wind Speed and Streamflow Forecasting,” *IEEE Access*, 2020, doi: 10.1109/ACCESS.2020.3034127.

[132] W. Bai, D. Lee, and K. Y. Lee, “A multivariate time series forecast model for wind and storage integrated system operation,” in *IEEE Power and Energy Society General Meeting*, 2018. doi: 10.1109/PESGM.2017.8274436.

[133] K. Wickstrom, K. O. Mikalsen, M. Kampffmeyer, A. Revhaug, and R. Jenssen, “Uncertainty-Aware Deep Ensembles for Reliable and Explainable Predictions of Clinical Time Series,” *IEEE J Biomed Health Inform*, 2021, doi: 10.1109/JBHI.2020.3042637.

[134] V. Borisov and V. Luferov, “Forecasting of Multidimensional Time Series Basing on Fuzzy Rule-Based Models,” in *Proceedings - 2019 21st International Conference "Complex Systems: Control and Modeling Problems", CSCMP 2019*, 2019. doi: 10.1109/CSCMP45713.2019.8976821.

[135] M. Dong, D. Wiebe, and J. Shi, "An Accelerated and Risk-free AC Power Flow Method with Machine Learning based Initiation," in *2022 IEEE Electrical Power and Energy Conference, EPEC 2022*, 2022. doi: 10.1109/EPEC56903.2022.10000110.

[136] J. Alzubi, A. Nayyar, and A. Kumar, "Machine Learning from Theory to Algorithms: An Overview," in *Journal of Physics: Conference Series*, 2018. doi: 10.1088/1742-6596/1142/1/012012.

[137] R. Rajendran, P. Sharma, N. K. Saran, S. Ray, J. Alanya-Beltran, and K. Tongkachok, "An Exploratory analysis of Machine Learning adaptability in Big Data Analytics Environments: A Data Aggregation in the age of Big Data and the Internet of Things," in *Proceedings of 2nd International Conference on Innovative Practices in Technology and Management, ICIPTM 2022*, 2022. doi: 10.1109/ICIPTM54933.2022.9753921.

[138] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward," *PLoS One*, 2018, doi: 10.1371/journal.pone.0194889.

[139] A. Lapedes and R. Farber, "How Neural Nets Work," in *Evolution, Learning and Cognition*, 1989. doi: 10.1142/9789814434102_0012.

[140] N. Sapankevych and R. Sankar, "Time series prediction using support vector machines: A survey," *IEEE Comput Intell Mag*, 2009, doi: 10.1109/MCI.2009.932254.

[141] E. Spiliotis, "Decision Trees for Time-Series Forecasting," *Foresight: The International Journal of Applied Forecasting*, vol. 1, no. 64, pp. 30–44, 2022.

[142] K. Das and R. N. Behera, "A Survey on Machine Learning: Concept, Algorithms and Applications," *International Journal of Innovative Research in Computer and Communication Engineering*, 2017.

[143] X. Song, A. Mitnitski, J. Cox, and K. Rockwood, "Comparison of machine learning techniques with classical statistical models in predicting health outcomes," *Stud Health Technol Inform*, 2004, doi: 10.3233/978-1-60750-949-3-736.

[144] J. W. Shin, J. H. Chang, and N. S. Kim, "Voice activity detection based on statistical models and machine learning approaches," *Comput Speech Lang*, 2010, doi: 10.1016/j.csl.2009.02.003.

[145] E. Fayyoumi, S. Idwan, and H. Aboshindi, "Machine learning and statistical modelling for prediction of Novel COVID-19 patients case study: Jordan," *International Journal of*

Advanced Computer Science and Applications, 2020, doi: 10.14569/IJACSA.2020.0110518.

[146] K. Chhaya, A. Khanzode, and R. D. Sarode, "Advantages and disadvantages of artificial intelligence and machine learning: A literature review," *International Journal of Library & Information Science (IJLIS)* , vol. 9, no. 1, pp. 30–36, 2020.

[147] C. R. Chen, C. C. Yang, L. T. Tsai, and C. C. Yang, "Latent variable models for assessing interaction effects in cognitive neuroscience," in *Proceedings - 2010 3rd International Conference on Biomedical Engineering and Informatics, BMEI 2010*, 2010. doi: 10.1109/BMEI.2010.5639504.

[148] T. L. Lash, M. P. Fox, R. F. Maclehose, G. Maldonado, L. C. Mccandless, and S. Greenland, "Good practices for quantitative bias analysis," *Int J Epidemiol*, 2014, doi: 10.1093/ije/dyu149.

[149] M. Valaste and R. Lehtonen, "Multiple imputation for measurement error correction in survey data," *Q2010 European Conference*, no. July 2015, pp. 1–9, 2010.

[150] J. T. Chien and P. Liu, "Bayesian latent variable models for speech recognition," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2013. doi: 10.1109/ICASSP.2013.6639099.

[151] R. Zhao, W. Luk, X. Niu, H. Shi, and H. Wang, "Hardware Acceleration for Machine Learning," *Proceedings of IEEE Computer Society Annual Symposium on VLSI, ISVLSI*, vol. 2017-July, pp. 645–650, 2017, doi: 10.1109/ISVLSI.2017.127.

[152] G. Kim and S. Lim, "Development of an Interpretable Maritime Accident Prediction System Using Machine Learning Techniques," *IEEE Access*, vol. 10, pp. 41313–41329, 2022, doi: 10.1109/ACCESS.2022.3168302.

[153] X. Zhen, Y. Ning, W. Du, Y. Huang, and J. E. Vinnem, "An interpretable and augmented machine-learning approach for causation analysis of major accident risk indicators in the offshore petroleum industry," *Process Safety and Environmental Protection*, vol. 173, no. October 2022, pp. 922–933, 2023, doi: 10.1016/j.psep.2023.03.063.

[154] Z. Yang, A. Zhang, and A. Sudjianto, "GAMI-Net: An explainable neural network based on generalized additive models with structured interactions," *Pattern Recognit*, 2021, doi: 10.1016/j.patcog.2021.108192.

[155] H. Yang, C. Rudin, and M. Seltzer, "Scalable Bayesian rule lists," in *34th International Conference on Machine Learning, ICML 2017*, 2017.

[156] A. V. Konstantinov and L. V. Utkin, "Interpretable machine learning with an ensemble of gradient boosting machines," *Knowl Based Syst*, 2021, doi: 10.1016/j.knosys.2021.106993.

[157] S. Das, N. Agarwal, D. Venugopal, F. T. Sheldon, and S. Shiva, "Taxonomy and Survey of Interpretable Machine Learning Method," in *2020 IEEE Symposium Series on Computational Intelligence, SSCI 2020*, 2020. doi: 10.1109/SSCI47803.2020.9308404.

[158] M. A. Ahmad, A. Teredesai, and C. Eckert, "Interpretable machine learning in healthcare," in *Proceedings - 2018 IEEE International Conference on Healthcare Informatics, ICHI 2018*, 2018. doi: 10.1109/ICHI.2018.00095.

[159] V. Nasteski, "An overview of the supervised machine learning methods," *HORIZONS.B*, 2017, doi: 10.20544/horizons.b.04.1.17.p05.

[160] S. H. Shetty, S. Shetty, C. Singh, and A. Rao, "Supervised machine learning: Algorithms and applications," in *Fundamentals and Methods of Machine and Deep Learning: Algorithms, Tools, and Applications*, 2022. doi: 10.1002/9781119821908.ch1.

[161] N. Giamarelos, E. N. Zois, M. Papadimitrakis, M. Stogiannos, N. A.-I. Livanos, and A. Alexandridis, "Short-Term Electric Load Forecasting with Sparse Coding Methods," *IEEE Access*, 2021, doi: 10.1109/ACCESS.2021.3098121.

[162] Z. Tan, J. Zhang, Y. He, Y. Zhang, G. Xiong, and Y. Liu, "Short-Term Load Forecasting Based on Integration of SVR and Stacking," *IEEE Access*, vol. 8, pp. 227719–227728, 2020, doi: 10.1109/ACCESS.2020.3041779.

[163] K. M. Nakanishi, K. Fujii, and S. Todo, "Sequential minimal optimization for quantum-classical hybrid algorithms," *Phys Rev Res*, vol. 2, no. 4, Oct. 2020, doi: 10.1103/PhysRevResearch.2.043158.

[164] M. Zulfiqar, K. A. A. Gamage, M. Kamran, and M. B. Rasheed, "Hyperparameter Optimization of Bayesian Neural Network Using Bayesian Optimization and Intelligent Feature Engineering for Load Forecasting," *Sensors 2022, Vol. 22, Page 4446*, vol. 22, no. 12, p. 4446, Jun. 2022, doi: 10.3390/S22124446.

[165] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal of the Society for Industrial and Applied Mathematics*, 1963, doi: 10.1137/0111030.

[166] A. Alexandridis, E. Chondrodima, and H. Sarimveis, "Radial basis function network training using a nonsymmetric partition of the input space and particle swarm optimization," *IEEE Trans Neural Netw Learn Syst*, vol. 24, no. 2, pp. 219–230, 2013, doi: 10.1109/TNNLS.2012.2227794.

[167] M. Papadimitrakis and A. Alexandridis, "Active vehicle suspension control using road preview model predictive control and radial basis function networks," *Appl Soft Comput*, 2022, doi: 10.1016/j.asoc.2022.108646.

[168] D. Karamichailidou, A. Alexandridis, G. Anagnostopoulos, G. Syriopoulos, and O. Sekkas, "Modeling biogas production from anaerobic wastewater treatment plants using radial basis function networks and differential evolution," *Comput Chem Eng*, 2022, doi: 10.1016/j.compchemeng.2021.107629.

[169] E. Chondrodima, H. Georgiou, N. Pelekis, and Y. Theodoridis, "Particle swarm optimization and RBF neural networks for public transport arrival time prediction using GTFS data," *International Journal of Information Management Data Insights*, 2022, doi: 10.1016/j.jjimei.2022.100086.

[170] D. Karamichailidou, V. Kaloutsa, and A. Alexandridis, "Wind turbine power curve modeling using radial basis function neural networks and tabu search," *Renew Energy*, 2021, doi: 10.1016/j.renene.2020.10.020.

[171] V. A. C. Horta, I. Tiddi, S. Little, and A. Mileo, "Extracting knowledge from Deep Neural Networks through graph analysis," *Future Generation Computer Systems*, vol. 120, pp. 109–118, Jul. 2021, doi: 10.1016/J.FUTURE.2021.02.009.

[172] D. Bacciu, F. Errica, A. Micheli, and M. Podda, "A gentle introduction to deep learning for graphs," *Neural Networks*, vol. 129, pp. 203–221, Sep. 2020, doi: 10.1016/J.NEUNET.2020.06.006.

[173] D. Ahmedt-Aristizabal, M. A. Armin, S. Denman, C. Fookes, and L. Petersson, "A survey on graph-based deep learning for computational histopathology," *Computerized Medical Imaging and Graphics*, vol. 95, p. 102027, Jan. 2022, doi: 10.1016/J.COMPMEDIMAG.2021.102027.

[174] Z. Liu and J. Zhou, "Introduction to Graph Neural Networks," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 14, no. 2, pp. 1–127, 2020, doi: 10.2200/S00980ED1V01Y202001AIM045.

[175] W. Zhao, C. Persello, and A. Stein, "Extracting planar roof structures from very high resolution images using graph neural networks," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 187, pp. 34–45, May 2022, doi: 10.1016/J.ISPRSJPRS.2022.02.022.

[176] L. Xie *et al.*, "Graph neural network-enabled manufacturing method classification from engineering drawings," *Comput Ind*, vol. 142, p. 103697, Nov. 2022, doi: 10.1016/J.COMPIND.2022.103697.

[177] W. Liao, B. Zeng, J. Liu, P. Wei, X. Cheng, and W. Zhang, "Multi-level graph neural network for text sentiment analysis," *Computers & Electrical Engineering*, vol. 92, p. 107096, Jun. 2021, doi: 10.1016/J.COMPELECENG.2021.107096.

[178] Q. Zhang *et al.*, "Graph Neural Network-Driven Traffic Forecasting for the Connected Internet of Vehicles," *IEEE Trans Netw Sci Eng*, vol. 9, no. 5, pp. 3015–3027, 2022, doi: 10.1109/TNSE.2021.3126830.

[179] F. Zhou, Q. Yang, T. Zhong, D. Chen, and N. Zhang, "Variational Graph Neural Networks for Road Traffic Prediction in Intelligent Transportation Systems," *IEEE Trans Industr Inform*, vol. 17, no. 4, pp. 2802–2812, Apr. 2021, doi: 10.1109/TII.2020.3009280.

[180] F. Fusco, B. Eck, R. Gormally, M. Purcell, and S. Tirupathi, "Knowledge- And Data-driven Services for Energy Systems using Graph Neural Networks," *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, pp. 1301–1308, Dec. 2020, doi: 10.1109/BIGDATA50022.2020.9377845.

[181] O. Boyaci *et al.*, "Graph Neural Networks Based Detection of Stealth False Data Injection Attacks in Smart Grids," *IEEE Syst J*, vol. 16, no. 2, pp. 2946–2957, Jun. 2022, doi: 10.1109/JSYST.2021.3109082.

[182] M. A. Islam, M. Ismail, O. Boyaci, R. Atat, and S. Shannigrahi, "Graph Neural Network Based Prediction of Data Traffic in Cyber-Physical Smart Power Grids," *2022 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, SmartGridComm 2022*, pp. 322–327, 2022, doi: 10.1109/SMARTGRIDCOMM52983.2022.9960963.

[183] F. Gama, E. Tolstaya, and A. Ribeiro, "Graph neural networks for decentralized controllers," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2021-June, pp. 5260–5264, 2021, doi: 10.1109/ICASSP39728.2021.9414563.

[184] A. Yaniv, P. Kumar, and Y. Beck, "Towards adoption of GNNs for power flow applications in distribution systems," *Electric Power Systems Research*, vol. 216, p. 109005, Mar. 2023, doi: 10.1016/J.EPSR.2022.109005.

[185] J. Moe, K. Pogorelov, D. T. Schroeder, and J. Langguth, "Implementating Spatio-Temporal Graph Convolutional Networks on Graphcore IPUs," in *Proceedings - 2022 IEEE 36th International Parallel and Distributed Processing Symposium Workshops, IPDPSW 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 45–54. doi: 10.1109/IPDPSW55747.2022.00016.

[186] T.-L. H. Y. L. T. Zhang, "Encrypted Network Traffic Classification Using a Geometric Learning Model," in *IEEE International Symposium on Integrated Network Management (IM)*, Bordeaux, France: IEEE, 2021, pp. 376–383. Accessed: Mar. 10, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9463930

[187] J. Lu, Y. Tian, S. Wang, M. Sheng, and X. Zheng, "PearNet: A Pearson Correlation-based Graph Attention Network for Sleep Stage Recognition," Sep. 2022, doi: 10.1109/DSAA54385.2022.10032354.

[188] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "Computational capabilities of graph neural networks," *IEEE Trans Neural Netw*, 2009, doi: 10.1109/TNN.2008.2005141.

[189] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A Comprehensive Survey on Graph Neural Networks," *IEEE Trans Neural Netw Learn Syst*, 2021, doi: 10.1109/TNNLS.2020.2978386.

[190] S. Suresh, V. Budde, J. Neville, P. Li, and J. Ma, "Breaking the Limit of Graph Neural Networks by Improving the Assortativity of Graphs with Local Mixing Patterns," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2021. doi: 10.1145/3447548.3467373.

[191] G. Li, M. Müller, B. Ghanem, and V. Koltun, "Training Graph Neural Networks with 1000 Layers." PMLR, pp. 6437–6449, Jul. 01, 2021. Accessed: Mar. 10, 2023. [Online]. Available: https://proceedings.mlr.press/v139/li21o.html

[192] M. Stephen, C. Gu, and H. Yang, "Visibility graph based time series analysis," *PLoS One*, 2015, doi: 10.1371/journal.pone.0143015.

[193] L. Lacasa and J. Iacovacci, "Visibility graphs of random scalar fields and spatial data," *Phys Rev E*, vol. 96, no. 1, Feb. 2017, doi: 10.1103/PhysRevE.96.012318.

[194] A. M., L. Lacasa, J. Patricio, and B. Luque, "Visibility Algorithms: A Short Review," in *New Frontiers in Graph Theory*, 2012. doi: 10.5772/34810.

[195] E. N. Zois, E. Zervas, D. Tsourounis, and G. Economou, "Sequential motif profiles and topological plots for offline signature verification," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 13245–13255, 2020, doi: 10.1109/CVPR42600.2020.01326.

[196] A. Olamat, P. Shams, and A. Akan, "State transfer network of time series based on visibility graph analysis for classifying and prediction of epilepsy seizures," *2017 Medical Technologies National Conference, TIPTEKNO 2017*, vol. 2017-Janua, pp. 1–4, Dec. 2017, doi: 10.1109/TIPTEKNO.2017.8238041.

[197] C. Hao, Z. Chen, and Z. Zhao, "Analysis and Prediction of Epilepsy Based on Visibility Graph," *Proceedings - 2016 3rd International Conference on Information Science and Control Engineering, ICISCE 2016*, pp. 1271–1274, Oct. 2016, doi: 10.1109/ICISCE.2016.272.

[198] J. Zhao, Y. Hu, and C. Hao, "Prediction of epileptic disease based on complex network," *Proceedings - 6th International Symposium on Computational Intelligence and Design, ISCID 2013*, vol. 2, pp. 395–398, 2013, doi: 10.1109/ISCID.2013.211.

[199] Y. Yang, J. Wang, H. Yang, and J. Mang, "Visibility graph approach to exchange rate series," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 20, pp. 4431–4437, Oct. 2009, doi: 10.1016/J.PHYSA.2009.07.016.

[200] Y. Huang, X. Mao, and Y. Deng, "Natural visibility encoding for time series and its application in stock trend prediction," *Knowl Based Syst*, vol. 232, p. 107478, Nov. 2021, doi: 10.1016/J.KNOSYS.2021.107478.

[201] T. Alwajeeh, P. Combeau, and L. Aveneau, "An Efficient Ray-Tracing Based Model Dedicated to Wireless Sensor Network Simulators for Smart Cities Environments," *IEEE Access*, vol. 8, pp. 206528–206547, 2020, doi: 10.1109/ACCESS.2020.3037135.

[202] T. Wen, H. Chen, and K. H. Cheong, "Visibility graph for time series prediction and image classification: a review," *Nonlinear Dyn*, 2022, doi: 10.1007/s11071-022-08002-4.

[203] F. Liu and Y. Deng, “A Fast Algorithm for Network Forecasting Time Series,” *IEEE Access*, vol. 7, pp. 102554–102560, 2019, doi: 10.1109/ACCESS.2019.2926986.

[204] T. Zhan and F. Xiao, “A novel weighted approach for time series forecasting based on visibility graph,” Mar. 2021, Accessed: Jan. 10, 2022. [Online]. Available: http://arxiv.org/abs/2103.13870

[205] S. Mao and F. Xiao, “Time Series Forecasting Based on Complex Network Analysis,” *IEEE Access*, vol. 7, pp. 40220–40229, 2019, doi: 10.1109/ACCESS.2019.2906268.

[206] S. Chen, X. Lan, Y. Hu, Q. Liu, and Y. Deng, “The time series forecasting: from the aspect of network,” Mar. 2014, Accessed: Jul. 05, 2021. [Online]. Available: http://arxiv.org/abs/1403.1713

[207] J. Zhao, H. Mo, and Y. Deng, “An Efficient Network Method for Time Series Forecasting Based on the DC Algorithm and Visibility Relation,” *IEEE Access*, vol. 8, pp. 7598–7608, 2020, doi: 10.1109/ACCESS.2020.2964067.

[208] R. Zhang, B. Ashuri, and Y. Deng, “A novel method for forecasting time series based on fuzzy logic and visibility graph,” *Adv Data Anal Classif*, vol. 11, no. 4, pp. 759–783, Dec. 2017, doi: 10.1007/S11634-017-0300-3.

[209] Y. Hu and F. Xiao, “A novel method for forecasting time series based on directed visibility graph and improved random walk,” *Physica A: Statistical Mechanics and its Applications*, vol. 594, p. 127029, 2022, doi: 10.1016/j.physa.2022.127029.

[210] M. Bristone, R. Prasad, and A. A. Abubakar, “CPPCNDL: Crude oil price prediction using complex network and deep learning algorithms,” *Petroleum*, vol. 6, no. 4, pp. 353–361, Dec. 2020, doi: 10.1016/J.PETLM.2019.11.009.

[211] Q. Li, Y. Chen, N. Ao, X. Han, and Z. Wu, “Echo state network-based visibility graph method for nonlinear time series prediction,” *Proceedings of the 30th Chinese Control and Decision Conference, CCDC 2018*, pp. 1854–1859, Jul. 2018, doi: 10.1109/CCDC.2018.8407428.

[212] S. Modak *et al.*, “Detection of Focal EEG Signals Employing Weighted Visibility Graph,” *2020 International Conference on Computer, Electrical and Communication Engineering, ICCECE 2020*, Jan. 2020, doi: 10.1109/ICCECE48148.2020.9223096.

[213] T. G. Altundogan and M. Karakose, “Cracked Wall Image Classification Based on Deep Neural Network Using Visibility Graph Features,” *2021 International Conference on*

Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2021, pp. 36–39, Sep. 2021, doi: 10.1109/3ICT53449.2021.9581830.

[214] T. Wan, X. Fu, K. Jiang, Y. Zhao, and B. Tang, "Radar Antenna Scan Pattern Intelligent Recognition Using Visibility Graph," *IEEE Access*, vol. 7, pp. 175628–175641, 2019, doi: 10.1109/ACCESS.2019.2957769.

[215] S. S. Roy *et al.*, "Accurate Detection of Bearing Faults Using Difference Visibility Graph and Bi-Directional Long Short-Term Memory Network Classifier," *IEEE Trans Ind Appl*, vol. 58, no. 4, pp. 4542–4551, 2022, doi: 10.1109/TIA.2022.3167658.

[216] C. Li, L. Mo, and R. Yan, "Rolling Bearing Fault Diagnosis Based on Horizontal Visibility Graph and Graph Neural Networks," *International Conference on Sensing, Measurement and Data Analytics in the Era of Artificial Intelligence, ICSMD 2020 - Proceedings*, pp. 275–279, Oct. 2020, doi: 10.1109/ICSMD50554.2020.9261687.

[217] Q. Xuan *et al.*, "AvgNet: Adaptive Visibility Graph Neural Network and Its Application in Modulation Classification," *IEEE Trans Netw Sci Eng*, vol. 9, no. 3, pp. 1516–1526, 2022, doi: 10.1109/TNSE.2022.3146836.

[218] R. Ying, C. Morris, W. L. Hamilton, J. You, X. Ren, and J. Leskovec, "Hierarchical Graph Representation Learning with Differentiable Pooling," *Adv Neural Inf Process Syst*, vol. 2018-Decem, pp. 4800–4810, Jun. 2018, doi: 10.48550/arxiv.1806.08804.

[219] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, "From time series to complex networks: The visibility graph," *Proceedings of the National Academy of Sciences*, vol. 105, no. 13, pp. 4972–4975, Apr. 2008, doi: 10.1073/PNAS.0709247105.

[220] X. Lan, H. Mo, S. Chen, Q. Liu, and Y. Deng, "Fast transformation from time series to visibility graphs," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 8, p. 083105, Aug. 2015, doi: 10.1063/1.4927835.

[221] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, "Horizontal visibility graphs: exact results for random time series," *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 80, no. 4, Feb. 2010, doi: 10.48550/arxiv.1002.4526.

[222] Z. Ting-Ting *et al.*, "Limited penetrable visibility graph for establishing complex network from time series," *Acta Physica Sinica, 2012, Vol. 61, Issue 3, Pages: 030506-030506*, vol. 61, no. 3, pp. 030506–030506, Mar. 2012, doi: 10.7498/APS.61.030506.

[223] L. Lacasa, V. Nicosia, and V. Latora, "Network structure of multivariate time series," *Sci Rep*, 2015, doi: 10.1038/srep15508.

[224] J. Iacovacci and L. Lacasa, "Visibility graphs for image processing," *IEEE Trans Pattern Anal Mach Intell*, vol. 42, no. 4, pp. 974–987, Apr. 2020, doi: 10.1109/TPAMI.2019.2891742.

[225] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans Neural Netw*, 2009, doi: 10.1109/TNN.2008.2005605.

[226] J. Wang and F. Biljecki, "Unsupervised machine learning in urban studies: A systematic review of applications," *Cities*, 2022, doi: 10.1016/j.cities.2022.103925.

[227] N. G. Pavlidis, D. K. Tasoulis, and M. N. Vrahatis, "Financial forecasting through unsupervised clustering and evolutionary trained neural networks," in *2003 Congress on Evolutionary Computation, CEC 2003 - Proceedings*, 2003. doi: 10.1109/CEC.2003.1299377.

[228] C. Feng, M. Cui, B. M. Hodge, S. Lu, H. F. Hamann, and J. Zhang, "Unsupervised Clustering-Based Short-Term Solar Forecasting," *IEEE Trans Sustain Energy*, 2019, doi: 10.1109/TSTE.2018.2881531.

[229] T. Chauhan, S. Rawat, S. Malik, and P. Singh, "Supervised and Unsupervised Machine Learning based Review on Diabetes Care," in *2021 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021*, 2021. doi: 10.1109/ICACCS51430.2021.9442021.

[230] A. Almalaq and G. Edwards, "A review of deep learning methods applied on load forecasting," in *Proceedings - 16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017*, 2017. doi: 10.1109/ICMLA.2017.0-110.

[231] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online Learning for Matrix Factorization and Sparse Coding," *J. Mach. Learn. Res.*, vol. 11, pp. 19–60, 2010.

[232] R. Jenatton, I. Julien Mairal, G. Obozinski, and F. Bach, "Proximal Methods for Sparse Hierarchical Dictionary Learning," in *Proceedings of the 27th International Conference on Machine Learning*, 2010.

[233] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Supervised dictionary learning," in *Advances in Neural Information Processing Systems 21 - Proceedings of the 2008 Conference*, 2009.

[234] J. Mairal, M. Elad, and G. Sapiro, "Sparse Representation for Color Image Restoration," *IEEE Transactions on Image Processing*, vol. 17, no. 1, pp. 53–69, Jan. 2008, doi: 10.1109/TIP.2007.911828.

[235] E. N. Zois, D. Tsourounis, I. Theodorakopoulos, A. L. Kesidis, and G. Economou, "A comprehensive study of sparse representation techniques for offline signature verification," *IEEE Trans Biom Behav Identity Sci*, vol. 1, no. 1, pp. 68–81, Jul. 2019, doi: 10.1109/tbiom.2019.2897802.

[236] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, 1996, doi: 10.1038/381607a0.

[237] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Optimization with Sparsity-Inducing Penalties," *Foundations and Trends in Machine Learning*, vol. 4, no. 1, pp. 1–106, Aug. 2011, Accessed: May 11, 2021. [Online]. Available: http://arxiv.org/abs/1108.0775

[238] C. N. Yu, P. Mirowski, and T. K. Ho, "A Sparse Coding Approach to Household Electricity Demand Forecasting in Smart Grids," *IEEE Trans Smart Grid*, vol. 8, no. 2, pp. 738–748, Mar. 2017, doi: 10.1109/TSG.2015.2513900.

[239] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proceedings of the IEEE*, 2010, doi: 10.1109/JPROC.2010.2040551.

[240] V. Papyan, Y. Romano, J. Sulam, and M. Elad, "Theoretical foundations of deep learning via sparse representations: A multilayer sparse model and its connection to convolutional neural networks," *IEEE Signal Process Mag*, vol. 35, no. 4, pp. 72–89, Jul. 2018, doi: 10.1109/MSP.2018.2820224.

[241] J. Sulam, B. Ophir, M. Zibulevsky, and M. Elad, "Trainlets: Dictionary Learning in High Dimensions," *IEEE Transactions on Signal Processing*, vol. 64, no. 12, pp. 3180–3193, Jun. 2016, doi: 10.1109/TSP.2016.2540599.

[242] E. Elhamifar, G. Sapiro, and S. S. Sastry, "Dissimilarity-based Sparse Subset Selection," *IEEE Trans Pattern Anal Mach Intell*, vol. 38, no. 11, pp. 2182–2197, Jul. 2014, Accessed: May 06, 2021. [Online]. Available: http://arxiv.org/abs/1407.6810

[243] A. Cutler and L. Breiman, "Archetypal Analysis," *Technometrics*, vol. 36, no. 4, p. 338, Nov. 1994, doi: 10.2307/1269949.

[244] Y. Chen, J. Mairal, and Z. Harchaoui, "Fast and Robust Archetypal Analysis for Representation Learning," *Proceedings of the IEEE Computer Society Conference on*

Computer Vision and Pattern Recognition, pp. 1478–1485, May 2014, Accessed: Apr. 28, 2021. [Online]. Available: http://arxiv.org/abs/1405.6472

[245] R. Jenatton, I. Julien Mairal, G. Obozinski, and F. Bach, "Proximal Methods for Hierarchical Sparse Coding," *Journal of Machine Learning Research*, vol. 12, pp. 2297–2334, 2011.

[246] H. Farhangi, "The path of the smart grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18–28, Jan. 2010, doi: 10.1109/MPE.2009.934876.

[247] N. S. Nafi, K. Ahmed, M. A. Gregory, and M. Datta, "A survey of smart grid architectures, applications, benefits and standardization," *Journal of Network and Computer Applications*, vol. 76. Academic Press, pp. 23–36, Dec. 01, 2016. doi: 10.1016/j.jnca.2016.10.003.

[248] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid - The new and improved power grid: A survey," *IEEE Communications Surveys and Tutorials*. 2012. doi: 10.1109/SURV.2011.101911.00087.

[249] IEEE Standards Committee, "IEEE Guide for Smart Grid Interoperability of Energy Technology and Information Technology Operation with the Electric Power System (EPS), End-Use Applications, and Loads," New York, 2011.

[250] R. Morello, C. De Capua, G. Fulco, and S. C. Mukhopadhyay, "A smart power meter to monitor energy flow in smart grids: The role of advanced sensing and iot in the electric grid of the future," *IEEE Sens J*, vol. 17, no. 23, pp. 7828–7837, Dec. 2017, doi: 10.1109/JSEN.2017.2760014.

[251] T. Ahmad, "Non-technical loss analysis and prevention using smart meters," *Renewable and Sustainable Energy Reviews*, vol. 72. Elsevier Ltd, pp. 573–589, 2017. doi: 10.1016/j.rser.2017.01.100.

[252] G. Lobaccaro, S. Carlucci, and E. Löfström, "A review of systems and technologies for smart homes and smart grids," *Energies*. 2016. doi: 10.3390/en9050348.

[253] M. Papadimitrakis, N. Giamarelos, M. Stogiannos, E. N. Zois, N. A. I. Livanos, and A. Alexandridis, "Metaheuristic search in smart grid: A review with emphasis on planning, scheduling and power flow optimization applications," *Renewable and Sustainable Energy Reviews*, vol. 145. Elsevier Ltd, p. 111072, Jul. 01, 2021. doi: 10.1016/j.rser.2021.111072.

[254] S. Hussain, C. Z. El-Bayeh, C. Lai, and U. Eicker, "Multi-Level Energy Management Systems Toward a Smarter Grid: A Review," *IEEE Access*, 2021, doi: 10.1109/ACCESS.2021.3078082.

[255] Y. Seyedi, H. Karimi, and S. Grijalva, "Irregularity Detection in Output Power of Distributed Energy Resources Using PMU Data Analytics in Smart Grids," *IEEE Trans Industr Inform*, vol. 15, no. 4, pp. 2222–2232, 2019, doi: 10.1109/TII.2018.2865765.

[256] N.-A. I. Livanos, S. Hammal, N. Giamarelos, V. Alifragkis, C. S. Psomopoulos, and E. N. Zois, "OpenEdgePMU: An Open PMU Architecture with Edge Processing for Future Resilient Smart Grids," *Energies (Basel)*, 2023, doi: 10.3390/en16062756.

[257] F. Li *et al.*, "Smart transmission grid: Vision and framework," *IEEE Trans Smart Grid*, vol. 1, no. 2, pp. 168–177, 2010, doi: 10.1109/TSG.2010.2053726.

[258] Takele Ferede Agajie and Mesfin Fanuel, "Reliability improvement of distribution using distributed generation in smart grid: A case study," in *Active Electrical Distribution Network*, 1st ed., Sanjeevikumar Padmanaban, Baseem Khan, Om Prakash Mahela, Hassan Haes Alhelou, and S. Rajkumar, Eds., Academic Press, 2022, pp. 305–326.

[259] M. C. Falvo, L. Martirano, D. Sbordone, and E. Bocci, "Technologies for smart grids: A brief review," in *12th International Conference on Environment and Electrical Engineering, EEEIC 2013*, 2013. doi: 10.1109/EEEIC.2013.6549544.

[260] C. Cecati, G. Mokryani, A. Piccolo, and P. Siano, "An overview on the Smart Grid concept," *IECON Proceedings (Industrial Electronics Conference)*, pp. 3322–3327, 2010, doi: 10.1109/IECON.2010.5675310.

[261] P. Palensky and F. Kupzog, "Smart Grids," *Annu Rev Environ Resour*, vol. 38, pp. 201–226, Oct. 2013, doi: 10.1146/ANNUREV-ENVIRON-031312-102947.

[262] K. Steriotis, G. Tsaousoglou, N. Efthymiopoulos, P. Makris, and E. (Manos) Varvarigos, "A novel behavioral real time pricing scheme for the active energy consumers' participation in emerging flexibility markets," *Sustainable Energy, Grids and Networks*, vol. 16, pp. 14–27, Dec. 2018, doi: 10.1016/J.SEGAN.2018.05.002.

[263] C. Defeuilley, "Retail competition in electricity markets," *Energy Policy*, vol. 37, no. 2, pp. 377–386, 2009, doi: 10.1016/j.enpol.2008.07.025.

[264] G. Dileep, "A survey on smart grid technologies and applications," *Renew Energy*, vol. 146, pp. 2589–2625, Feb. 2020, doi: 10.1016/J.RENENE.2019.08.092.

[265] T. Karthick, S. Charles Raja, J. Jeslin Drusila Nesamalar, and K. Chandrasekaran, "Design of IoT based smart compact energy meter for monitoring and controlling the usage of energy and power quality issues with demand side management for a commercial building," *Sustainable Energy, Grids and Networks*, 2021, doi: 10.1016/j.segan.2021.100454.

[266] J. C. Bedoya, C. C. Liu, G. Krishnamoorthy, and A. Dubey, "Bilateral Electricity Market in a Distribution System Environment," *IEEE Trans Smart Grid*, vol. 10, no. 6, pp. 6701–6713, 2019, doi: 10.1109/TSG.2019.2910216.

[267] Y. K. Renani, M. Ehsan, and M. Shahidehpour, "Optimal transactive market operations with distribution system operators," *IEEE Trans Smart Grid*, vol. 9, no. 6, pp. 6692–6701, 2018, doi: 10.1109/TSG.2017.2718546.

[268] A. Q. Huang, M. L. Crow, G. T. Heydt, J. P. Zheng, and S. J. Dale, "The future renewable electric energy delivery and management (FREEDM) system: The energy internet," *Proceedings of the IEEE*, 2011, doi: 10.1109/JPROC.2010.2081330.

[269] J. Zheng, H. Okamura, T. Pang, and T. Dohi, "Availability importance measures of components in smart electric power grid systems," *Reliab Eng Syst Saf*, 2021, doi: 10.1016/j.ress.2020.107164.

[270] D. Çelik, M. E. Meral, and M. Waseem, "Investigation and analysis of effective approaches, opportunities, bottlenecks and future potential capabilities for digitalization of energy systems and sustainable development goals," *Electric Power Systems Research*, vol. 211, p. 108251, Oct. 2022, doi: 10.1016/J.EPSR.2022.108251.

[271] R. Singh, S. V. Akram, A. Gehlot, D. Buddhi, N. Priyadarshi, and B. Twala, "Energy System 4.0: Digitalization of the Energy Sector with Inclination towards Sustainability," *Sensors 2022, Vol. 22, Page 6619*, vol. 22, no. 17, p. 6619, Sep. 2022, doi: 10.3390/S22176619.

[272] M. Papadimitrakis, N. Giamarelos, M. Stogiannos, E. N. Zois, N. A. Livanos, and A. Alexandridis, "Metaheuristic search in smart grid : A review with emphasis on planning , scheduling and power flow optimization applications," *Renewable and Sustainable Energy Reviews*, vol. 145, no. April, p. 111072, 2021, doi: 10.1016/j.rser.2021.111072.

[273] N. Petrovic, L. Strezoski, and B. Dumnic, "Overview of software tools for integration and active management of high penetration of DERs in emerging distribution networks," in

EUROCON 2019 - 18th International Conference on Smart Technologies, 2019. doi: 10.1109/EUROCON.2019.8861765.

[274] M. Bashari and A. Rahimi-Kian, "Forecasting electric load by aggregating meteorological and history-based deep learning modules," in *IEEE Power and Energy Society General Meeting*, 2020. doi: 10.1109/PESGM41954.2020.9282124.

[275] D. Putz, M. Gumhalter, and H. Auer, "The true value of a forecast: Assessing the impact of accuracy on local energy communities," *Sustainable Energy, Grids and Networks*, vol. 33, p. 100983, Mar. 2023, doi: 10.1016/J.SEGAN.2022.100983.

[276] R. Fonteijn, P. H. Nguyen, J. Morren, and J. G. Slootweg, "Demonstrating a generic four-step approach for applying flexibility for congestion management in daily operation," *Sustainable Energy, Grids and Networks*, vol. 23, p. 100378, Sep. 2020, doi: 10.1016/J.SEGAN.2020.100378.

[277] I. Colak, R. Bayindir, and S. Sagiroglu, "The Effects of the Smart Grid System on the National Grids," *8th International Conference on Smart Grid, icSmartGrid 2020*, pp. 122–126, Jun. 2020, doi: 10.1109/ICSMARTGRID49881.2020.9144891.

[278] I. Worighi, A. Maach, A. Hafid, O. Hegazy, and J. Van Mierlo, "Integrating renewable energy in smart grid system: Architecture, virtualization and analysis," *Sustainable Energy, Grids and Networks*, vol. 18, p. 100226, Jun. 2019, doi: 10.1016/J.SEGAN.2019.100226.

[279] N. Eghbali, S. M. Hakimi, A. Hasankhani, G. Derakhshan, and B. Abdi, "Stochastic energy management for a renewable energy based microgrid considering battery, hydrogen storage, and demand response," *Sustainable Energy, Grids and Networks*, vol. 30, p. 100652, Jun. 2022, doi: 10.1016/J.SEGAN.2022.100652.

[280] M. Numan, M. Asif, M. W. Khan, S. M. Malik, and F. U. Khilji, "Impact of Dynamic Thermal Rating on optimal siting and sizing of energy storage systems under Renewable Portfolio Standards requirements," *Sustainable Energy, Grids and Networks*, vol. 32, p. 100881, Dec. 2022, doi: 10.1016/J.SEGAN.2022.100881.

[281] D. Putz, M. Gumhalter, and H. Auer, "The true value of a forecast: Assessing the impact of accuracy on local energy communities," *Sustainable Energy, Grids and Networks*, vol. 33, p. 100983, Mar. 2023, doi: 10.1016/J.SEGAN.2022.100983.

[282] S. Kouhi and F. Keynia, “A new cascade NN based method to short-term load forecast in deregulated electricity market,” *Energy Convers Manag*, vol. 71, pp. 76–83, Jul. 2013, doi: 10.1016/j.enconman.2013.03.014.

[283] P. S. Georgilakis, “Market clearing price forecasting in deregulated electricity markets using adaptively trained neural networks,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2006. doi: 10.1007/11752912_8.

[284] S. K. Dash and D. Patel, “Short-term electric load forecasting using Extreme Learning Machine - A case study of Indian power market,” in *2015 IEEE Power, Communication and Information Technology Conference, PCITC 2015 - Proceedings*, IEEE, Oct. 2016, pp. 961–966. doi: 10.1109/PCITC.2015.7438135.

[285] A. J. P. P. J. Santos, A. G. Martins, “Short-term Load Forecasting Based on ANN Applied to Electrical Distribution Substations,” *Proc. of the 39th International Universities Power Engineering Conference*, vol. 1, pp. 427–432, 2004.

[286] E. Mohammadi, M. Alizadeh, M. Asgarimoghaddam, X. Wang, and M. G. Simoes, “A Review on Application of Artificial Intelligence Techniques in Microgrids,” *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, vol. 3, no. 4, pp. 878–890, Aug. 2022, doi: 10.1109/JESTIE.2022.3198504.

[287] S. Theodoridis, *Machine Learning: A Bayesian and Optimization Perspective*. 2015. doi: 10.1016/C2013-0-19102-7.

[288] C. Wang, G. Grozev, and S. Seo, “Decomposition and statistical analysis for regional electricity demand forecasting,” *Energy*, vol. 41, no. 1, pp. 313–325, May 2012, doi: 10.1016/J.ENERGY.2012.03.011.

[289] A. Gerossier, R. Girard, G. Kariniotakis, and A. Michiorri, “Probabilistic day-ahead forecasting of household electricity demand,” *CIRED - Open Access Proceedings Journal*, vol. 2017, no. 1, pp. 2500–2504, Oct. 2017, doi: 10.1049/oap-cired.2017.0625.

[290] Y. Yang, S. Li, W. Li, and M. Qu, “Power load probability density forecasting using Gaussian process quantile regression,” *Appl Energy*, 2018, doi: 10.1016/j.apenergy.2017.11.035.

[291] D. H. Vu, K. M. Muttaqi, and A. P. Agalgaonkar, "Assessing the influence of climatic variables on electricity demand," in *2014 IEEE PES General Meeting | Conference & Exposition*, IEEE, Jul. 2014, pp. 1–5. doi: 10.1109/PESGM.2014.6939377.

[292] J. Xie, Y. Chen, T. Hong, and T. D. Laing, "Relative humidity for load forecasting models," *IEEE Trans Smart Grid*, 2018, doi: 10.1109/TSG.2016.2547964.

[293] C. Alzate and M. Sinn, "Improved electricity load forecasting via kernel spectral clustering of smart meters," in *Proceedings - IEEE International Conference on Data Mining, ICDM*, 2013, pp. 943–948. doi: 10.1109/ICDM.2013.144.

[294] M. Rossi and D. Brunelli, "Electricity demand forecasting of single residential units," in *2013 IEEE Workshop on Environmental, Energy and Structural Monitoring Systems, EESMS 2013 - Proceedings*, 2013. doi: 10.1109/EESMS.2013.6661693.

[295] P. Damrongkulkamjorn and P. Churueang, "Monthly energy forecasting using decomposition method with application of seasonal ARIMA," *7th International Power Engineering Conference, IPEC2005*, vol. 2005, 2005, doi: 10.1109/IPEC.2005.206911.

[296] M. Ould Mohamed Mahmoud, F. Mhamdi, and M. Jaïdane-Saïdane, "Long term multi-scale analysis of the daily peak load based on the empirical mode decomposition," *2009 IEEE Bucharest PowerTech: Innovative Ideas Toward the Electrical Grid of the Future*, 2009, doi: 10.1109/PTC.2009.5281805.

[297] A. E. Clements, A. S. Hurn, and Z. Li, "Forecasting day-ahead electricity load using a multiple equation time series approach," *Eur J Oper Res*, vol. 251, no. 2, pp. 522–530, Jun. 2016, doi: 10.1016/j.ejor.2015.12.030.

[298] Z. S. Elrazaz and A. A. Mazi, "Unified weekly peak load forecasting for fast growing power system," *IEE Proceedings C Generation, Transmission and Distribution*, vol. 136, no. 1, p. 29, 2010, doi: 10.1049/ip-c.1989.0005.

[299] D. Yang, L. Xu, S. Gong, H. Li, G. D. Peterson, and Z. Zhang, "Joint electrical load modeling and forecasting based on sparse Bayesian Learning for the smart grid," in *2011 45th Annual Conference on Information Sciences and Systems, CISS 2011*, 2011. doi: 10.1109/CISS.2011.5766184.

[300] X. Sun, X. Wang, J. Wu, and Y. Liu, "Hierarchical sparse learning for load forecasting in cyber-physical energy systems," in *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, 2013. doi: 10.1109/I2MTC.2013.6555474.

[301] Q. Duan, W. X. Sheng, Y. Ma, and K. Ma, "Sparse Bayesian Learning using combined kernels for medium term load forecasting," in *IET Conference Publications*, 2013. doi: 10.1049/cp.2013.1740.

[302] Y. Zhang, S. J. Kim, and G. B. Giannakis, "Short-term wind power forecasting using nonnegative sparse coding," in *2015 49th Annual Conference on Information Sciences and Systems, CISS 2015*, Institute of Electrical and Electronics Engineers Inc., Apr. 2015. doi: 10.1109/CISS.2015.7086873.

[303] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006, doi: 10.1109/TSP.2006.881199.

[304] S. Singh and A. Majumdar, "Deep sparse coding for non-intrusive load monitoring," *IEEE Trans Smart Grid*, vol. 9, no. 5, pp. 4669–4678, Sep. 2018, doi: 10.1109/TSG.2017.2666220.

[305] Y. Wang, Q. Chen, C. Kang, Q. Xia, and M. Luo, "Sparse and Redundant Representation-Based Smart Meter Data Compression and Pattern Extraction," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2142–2151, May 2017, doi: 10.1109/TPWRS.2016.2604389.

[306] M. K. Azad, S. Uddin, and M. Takruri, "Support vector regression based electricity peak load forecasting," in *11th International Symposium on Mechatronics and its Applications, ISMA 2018*, IEEE, Mar. 2018, pp. 1–5. doi: 10.1109/ISMA.2018.8330143.

[307] J. Che and J. Wang, "Short-term load forecasting using a kernel-based support vector regression combination model," *Appl Energy*, vol. 132, pp. 602–609, Nov. 2014, doi: 10.1016/J.APENERGY.2014.07.064.

[308] Y. Li, J. Che, and Y. Yang, "Subsampled support vector regression ensemble for short term electric load forecasting," *Energy*, 2018, doi: 10.1016/j.energy.2018.08.169.

[309] L. Ghelardoni, A. Ghio, and D. Anguita, "Energy load forecasting using empirical mode decomposition and support vector regression," *IEEE Trans Smart Grid*, 2013, doi: 10.1109/TSG.2012.2235089.

[310] G. Hafeez, I. Khan, S. Jan, I. A. Shah, F. A. Khan, and A. Derhab, "A novel hybrid load forecasting framework with intelligent feature engineering and optimization algorithm in

smart grid," *Appl Energy*, vol. 299, p. 117178, Oct. 2021, doi: 10.1016/J.APENERGY.2021.117178.

[311] N. Korovesis, D. Kandris, G. Koulouras, and A. Alexandridis, "Robot motion control via an eeg-based brain–computer interface by using neural networks and alpha brainwaves," *Electronics (Switzerland)*, 2019, doi: 10.3390/electronics8121387.

[312] E. Akarslan and F. O. Hocaoglu, "Electricity demand forecasting of a micro grid using ANN," in *2018 9th International Renewable Energy Congress, IREC 2018*, IEEE, Mar. 2018, pp. 1–5. doi: 10.1109/IREC.2018.8362471.

[313] A. K. Pandey, K. B. Sahay, D. Chandra, and M. M. Tripathi, "Day Ahead Load Forecast in ISO New England Market and Ontario Market using a Novel ANN," 2015. Accessed: Apr. 17, 2019. [Online]. Available: https://ijrest.net/downloads/volume-2/issue-4/pid-ijrest-24201517.pdf

[314] A. Bala and N. K. Yadav, "Load Forecasting For Weekend Load Using ANN Technique in Deregulated Environment," 2014. doi: 10.9790/1676-09420106.

[315] M. H. M. R. S. Dilhani and C. Jeenanunta, "Daily electric load forecasting: Case of Thailand," in *7th International Conference on Information Communication Technology for Embedded Systems 2016, IC-ICTES 2016*, IEEE, Mar. 2016, pp. 25–29. doi: 10.1109/ICTEmSys.2016.7467116.

[316] M. Q. Raza, Z. Baharudin, Badar-Ul-Islam, M. Azman Zakariya, and M. H. M. Khir, "Neural network based STLF model to study the seasonal impact of weather and exogenous variables," *Research Journal of Applied Sciences, Engineering and Technology*, 2013, doi: 10.19026/rjaset.6.3583.

[317] K. B. Sahay and M. M. Tripathi, "Day ahead hourly load forecast of PJM electricity market and iso new england market by using artificial neural network," in *ISGT 2014*, IEEE, Feb. 2014, pp. 1–5. doi: 10.1109/ISGT.2014.6816486.

[318] C. Roldán-Blay, G. Escrivá-Escrivá, C. Álvarez-Bel, C. Roldán-Porta, and J. Rodríguez-García, "Upgrade of an artificial neural network prediction method for electrical consumption forecasting using an hourly temperature curve model," *Energy Build*, 2013, doi: 10.1016/j.enbuild.2012.12.009.

[319] F. Rodrigues, C. Cardeira, and J. M. F. Calado, "The Daily and Hourly Energy Consumption and Load Forecasting Using Artificial Neural Network Method: A Case Study Using a Set

of 93 Households in Portugal," *Energy Procedia*, vol. 62, pp. 220–229, Jan. 2014, doi: 10.1016/J.EGYPRO.2014.12.383.

[320] K. B. Sahay, S. Sahu, and P. Singh, "Short-term load forecasting of Toronto Canada by using different ANN algorithms," in *2016 IEEE 6th International Conference on Power Systems (ICPS)*, IEEE, Mar. 2016, pp. 1–6. doi: 10.1109/ICPES.2016.7584044.

[321] L. Alhmoud, R. Abu Khurma, A. M. Al-Zoubi, and I. Aljarah, "A Real-Time Electrical Load Forecasting in Jordan Using an Enhanced Evolutionary Feedforward Neural Network," *Sensors 2021, Vol. 21, Page 6240*, vol. 21, no. 18, p. 6240, Sep. 2021, doi: 10.3390/S21186240.

[322] Z. Liao *et al.*, "Short-term load forecasting with dense average network," *Expert Syst Appl*, 2021, doi: 10.1016/j.eswa.2021.115748.

[323] M. Ding, L. Wang, and R. Bi, "An ANN-based approach for forecasting the power output of photovoltaic system," in *Procedia Environmental Sciences*, 2011, pp. 1308–1315. doi: 10.1016/j.proenv.2011.12.196.

[324] R. M. Salgado, T. Ohishi, and R. Ballini, "A short-term bus load forecasting system," in *2010 10th International Conference on Hybrid Intelligent Systems*, Atlanta: IEEE, Aug. 2010, pp. 55–60. doi: 10.1109/HIS.2010.5600075.

[325] L. Hernandez *et al.*, "A multi-agent system architecture for smart grid management and forecasting of energy demand in virtual power plants," *IEEE Communications Magazine*, vol. 51, no. 1, pp. 106–113, 2013, doi: 10.1109/MCOM.2013.6400446.

[326] Y. I. Alamin, J. D. Álvarez, M. del Mar Castilla, and A. Ruano, "An Artificial Neural Network (ANN) model to predict the electric load profile for an HVAC system," *IFAC-PapersOnLine*, vol. 51, no. 10, pp. 26–31, Jan. 2018, doi: 10.1016/j.ifacol.2018.06.231.

[327] E. González-Romera, M. Á. Jaramillo-Morán, and D. Carmona-Fernández, "Monthly electric energy demand forecasting based on trend extraction," *IEEE Transactions on Power Systems*, 2006, doi: 10.1109/TPWRS.2006.883666.

[328] N. J. Johannesen, M. Kolhe, and M. Goodwin, "Relative evaluation of regression tools for urban area electrical energy demand forecasting," *J Clean Prod*, Jan. 2019, doi: 10.1016/J.JCLEPRO.2019.01.108.

[329] W. Zhang, H. Quan, and D. Srinivasan, "Parallel and reliable probabilistic load forecasting via quantile regression forest and quantile determination," *Energy*, 2018, doi: 10.1016/j.energy.2018.07.019.

[330] W. Lin, D. Wu, and B. Boulet, "Spatial-Temporal Residential Short-Term Load Forecasting via Graph Neural Networks," *IEEE Trans Smart Grid*, 2021, doi: 10.1109/TSG.2021.3093515.

[331] S. Eandi, A. Cini, S. Lukovic, and C. Alippi, "Spatio-Temporal Graph Neural Networks for Aggregate Load Forecasting," *Proceedings of the International Joint Conference on Neural Networks*, vol. 2022-July, 2022, doi: 10.1109/IJCNN55064.2022.9892780.

[332] S. Arastehfar, M. Matinkia, and M. R. Jabbarpour, "Short-term residential load forecasting using Graph Convolutional Recurrent Neural Networks," *Eng Appl Artif Intell*, vol. 116, no. January, p. 105358, 2022, doi: 10.1016/j.engappai.2022.105358.

[333] D. Wu and W. Lin, "Efficient Residential Electric Load Forecasting via Transfer Learning and Graph Neural Networks," *IEEE Trans Smart Grid*, vol. XX, no. X, pp. 1–9, 2022, doi: 10.1109/TSG.2022.3208211.

[334] P. Lusis, K. R. Khalilpour, L. Andrew, and A. Liebman, "Short-term residential load forecasting: Impact of calendar effects and forecast granularity," *Appl Energy*, vol. 205, pp. 654–669, Nov. 2017, doi: 10.1016/J.APENERGY.2017.07.114.

[335] T. Vantuch, A. G. Vidal, A. P. Ramallo-Gonzalez, A. F. Skarmeta, and S. Misak, "Machine learning based electric load forecasting for short and long-term period," in *2018 IEEE 4th World Forum on Internet of Things (WF-IoT)*, IEEE, Feb. 2018, pp. 511–516. doi: 10.1109/WF-IoT.2018.8355123.

[336] M. Bessec and J. Fouquau, "Short-run electricity load forecasting with combinations of stationary wavelet transforms," *Eur J Oper Res*, vol. 264, no. 1, pp. 149–164, 2018, doi: 10.1016/j.ejor.2017.05.037.

[337] N. Mughees, S. A. Mohsin, A. Mughees, and A. Mughees, "Deep sequence to sequence Bi-LSTM neural networks for day-ahead peak load forecasting," *Expert Syst Appl*, 2021, doi: 10.1016/j.eswa.2021.114844.

[338] S. P. Adam, S. A. N. Alexandropoulos, P. M. Pardalos, and M. N. Vrahatis, "No free lunch theorem: A review," *Springer Optimization and Its Applications*, vol. 145, pp. 57–82, 2019, doi: 10.1007/978-3-030-12767-1_5.

[339] O. Ahmia and N. Farah, “Multi-model approach for electrical load forecasting,” *IntelliSys 2015 - Proceedings of 2015 SAI Intelligent Systems Conference*, pp. 87–92, Dec. 2015, doi: 10.1109/INTELLISYS.2015.7361089.

[340] P. Zeng, M. Jin, and Md. F. Elahe, “Short-Term Power Load Forecasting Based on Cross Multi-Model and Second Decision Mechanism,” *IEEE Access*, vol. 8, pp. 184061–184072, 2020, doi: 10.1109/access.2020.3028649.

[341] C. Peng, Y. Tao, Z. Chen, Y. Zhang, and X. Sun, “Multi-source transfer learning guided ensemble LSTM for building multi-load forecasting,” *Expert Syst Appl*, vol. 202, p. 117194, Sep. 2022, doi: 10.1016/J.ESWA.2022.117194.

[342] S. Li, Y. Zhong, and J. Lin, “AWS-DAIE: Incremental Ensemble Short-Term Electricity Load Forecasting Based on Sample Domain Adaptation,” *Sustainability 2022, Vol. 14, Page 14205*, vol. 14, no. 21, p. 14205, Oct. 2022, doi: 10.3390/SU142114205.

[343] L. Yu, S. Wang, and K. K. Lai, “A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates,” *Comput Oper Res*, vol. 32, no. 10, pp. 2523–2541, Oct. 2005, doi: 10.1016/J.COR.2004.06.024.

[344] P. S. G. De Mattos Neto *et al.*, “Energy Consumption Forecasting for Smart Meters Using Extreme Learning Machine Ensemble,” *Sensors 2021, Vol. 21, Page 8096*, vol. 21, no. 23, p. 8096, Dec. 2021, doi: 10.3390/S21238096.

[345] Y. Hu *et al.*, “Short-term load forecasting using multimodal evolutionary algorithm and random vector functional link network based ensemble learning,” *Appl Energy*, vol. 285, p. 116415, Mar. 2021, doi: 10.1016/J.APENERGY.2020.116415.

[346] Q. Duan, J. Liu, and D. Zhao, “Short term electric load forecasting using an automated system of model choice,” *International Journal of Electrical Power and Energy Systems*, vol. 91, pp. 92–100, Oct. 2017, doi: 10.1016/J.IJEPES.2017.03.006.

[347] Q. Chen, W. Zhang, K. Zhu, D. Zhou, H. Dai, and Q. Wu, “A novel trilinear deep residual network with self-adaptive Dropout method for short-term load forecasting,” *Expert Syst Appl*, 2021, doi: 10.1016/j.eswa.2021.115272.

[348] R. Gao, L. Du, P. N. Suganthan, Q. Zhou, and K. F. Yuen, “Random vector functional link neural network based ensemble deep learning for short-term load forecasting,” *Expert Syst Appl*, vol. 206, p. 117784, Nov. 2022, doi: 10.1016/J.ESWA.2022.117784.

[349] L. Sun, K. Zhou, X. Zhang, and S. Yang, "Outlier Data Treatment Methods Toward Smart Grid Applications," *IEEE Access*, vol. 6, no. c, pp. 39849–39859, 2018, doi: 10.1109/ACCESS.2018.2852759.

[350] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *Journal of Machine Learning Research*, vol. 11, pp. 19–60, 2010, doi: 10.1145/1756006.1756008.

[351] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J. R. Statist. Soc. B*, vol. 67, no. 2, pp. 301–320, 2005.

[352] S. G. Mallat and Z. Zhang, "Matching Pursuits With Time-Frequency Dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993, doi: 10.1109/78.258082.

[353] G. Dudek, "Multilayer perceptron for short-term load forecasting: from global to local approach," *Neural Comput Appl*, vol. 32, doi: 10.1007/s00521-019-04130-y.

[354] Y. Wang *et al.*, "Short-term load forecasting of industrial customers based on SVMD and XGBoost," *International Journal of Electrical Power & Energy Systems*, vol. 129, p. 106830, Jul. 2021, doi: 10.1016/J.IJEPES.2021.106830.

[355] S. Wang, X. Wang, S. Wang, and D. Wang, "Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting," *International Journal of Electrical Power & Energy Systems*, vol. 109, pp. 470–479, Jul. 2019, doi: 10.1016/J.IJEPES.2019.02.022.

[356] U. Javed *et al.*, "A novel short receptive field based dilated causal convolutional network integrated with Bidirectional LSTM for short-term load forecasting," *Expert Syst Appl*, vol. 205, p. 117689, Nov. 2022, doi: 10.1016/J.ESWA.2022.117689.

[357] M. Saviozzi, S. Massucco, and F. Silvestro, "Implementation of advanced functionalities for Distribution Management Systems: Load forecasting and modeling through Artificial Neural Networks ensembles," *Electric Power Systems Research*, vol. 167, pp. 230–239, Feb. 2019.

[358] I. Panapakidis, M. Katsivelakis, and D. Bargiotas, "A Metaheuristics-Based Inputs Selection and Training Set Formation Method for Load Forecasting," *Symmetry 2022, Vol. 14, Page 1733*, vol. 14, no. 8, p. 1733, Aug. 2022, doi: 10.3390/SYM14081733.

[359] B. ul Islam, U. Ijaz, and M. T. Raza, “Hybrid Input Data Selection Technique for Neural Network Based Load Forecast Model,” *NFC-IEFR Journal of Engineering and Scientific Research*, vol. 4, no. 1, pp. 74–80, Dec. 2016, doi: 10.24081/NIJESR.2016.1.0014.

[360] M. R. Arahal, A. Cepeda, and E. F. Camacho, “INPUT VARIABLE SELECTION FOR FORECASTING MODELS,” *IFAC Proceedings Volumes*, vol. 35, no. 1, pp. 463–468, Jan. 2002, doi: 10.3182/20020721-6-ES-1901.00730.

[361] M. S. Alam, N. Sultana, and S. M. Z. Hossain, “Bayesian optimization algorithm based support vector regression analysis for estimation of shear capacity of FRP reinforced concrete members,” *Appl Soft Comput*, vol. 105, p. 107281, Jul. 2021, doi: 10.1016/j.asoc.2021.107281.

[362] M. T. Hagan and M. B. Menhaj, “Training Feedforward Networks with the Marquardt Algorithm,” *IEEE Trans Neural Netw*, vol. 5, no. 6, pp. 989–993, 1994, doi: 10.1109/72.329697.

[363] M. M. Rashidi, M. Alhuyi Nazari, I. Mahariq, and N. Ali, “Modeling and Sensitivity Analysis of Thermal Conductivity of Ethylene Glycol-Water Based Nanofluids with Alumina Nanoparticles,” *Exp Tech*, 2023, doi: 10.1007/s40799-022-00567-4.

[364] B. Zhou *et al.*, “Multi-energy net load forecasting for integrated local energy systems with heterogeneous prosumers,” *International Journal of Electrical Power & Energy Systems*, vol. 126, p. 106542, Mar. 2021, doi: 10.1016/J.IJEPES.2020.106542.

[365] A. Taik and S. Cherkaoui, “Electrical Load Forecasting Using Edge Computing and Federated Learning,” *IEEE International Conference on Communications*, vol. 2020-June, Jun. 2020, doi: 10.1109/ICC40277.2020.9148937.

[366] Qiyue Li; Yadong Zhu; Jinjin Ding; Weitao Li; Wei Sun; Lijian Ding, “Deep reinforcement learning based resource allocation for cloud edge collaboration fault detection in smart grid,” *CSEE Journal of Power and Energy Systems*, pp. 1–10, 2022, doi: 10.17775/CSEEJPES.2021.02390.

[367] Prajeesha and M. Anuradha, “EDGE Computing Application in SMART GRID-A Review,” *Proceedings of the 2nd International Conference on Electronics and Sustainable Communication Systems, ICESC 2021*, pp. 397–402, Aug. 2021, doi: 10.1109/ICESC51422.2021.9532792.

[368] J. Li, C. Gu, Y. Xiang, and F. Li, "Edge-cloud Computing Systems for Smart Grid: State-of-the-art, Architecture, and Applications," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 4, pp. 805–817, Jul. 2022, doi: 10.35833/MPCE.2021.000161.

[369] P. Whittle, "On Stationary Processes in the Plane," *Biometrika*, 1954, doi: 10.2307/2332724.

[370] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *J Econom*, 1986, doi: 10.1016/0304-4076(86)90063-1.

[371] W. Remus and M. O'Connor, "Neural Networks for Time-Series Forecasting," pp. 245–256, 2001, doi: 10.1007/978-0-306-47630-3_12.

[372] P. F. Pai, K. P. Lin, C. S. Lin, and P. T. Chang, "Time series forecasting by a seasonal support vector regression model," *Expert Syst Appl*, vol. 37, no. 6, pp. 4261–4265, Jun. 2010, doi: 10.1016/J.ESWA.2009.11.076.

[373] E. De Alba and M. Mendoza, "Bayesian Forecasting Methods for Short Time Series," *The International Journal of Applied Forecasting*, 2007.

[374] F. Gilani, D. Giannakis, and J. Harlim, "Kernel-based prediction of non-Markovian time series," *Physica D*, 2021, doi: 10.1016/j.physd.2020.132829.

[375] B. Lim and S. Zohren, "Time-series forecasting with deep learning: A survey," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 379, no. 2194, 2021, doi: 10.1098/rsta.2020.0209.

[376] S. K. Perepu, B. S. Balaji, H. K. Tanneru, S. Kathari, and V. S. Pinnamaraju, "Reinforcement Learning based dynamic weighing of Ensemble Models for Time Series Forecasting," Aug. 2020, Accessed: Dec. 21, 2021. [Online]. Available: https://www.researchgate.net/publication/343786503_Reinforcement_Learning_based_dynamic_weighing_of_Ensemble_Models_for_Time_Series_Forecasting

[377] C. Lam, Q. Yao, and N. Bathia, "Estimation of latent factors for high-dimensional time series," *Biometrika*, 2011, doi: 10.1093/biomet/asr048.

[378] M. C. Hout, M. H. Papesh, and S. D. Goldinger, "Multidimensional scaling," *Wiley Interdiscip Rev Cogn Sci*, vol. 4, no. 1, pp. 93–103, Jan. 2013, doi: 10.1002/wcs.1203.

[379] R. R. Coifman *et al.*, "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps," *Proc Natl Acad Sci U S A*, 2005, doi: 10.1073/pnas.0500334102.

[380] J. B. Tenenbaum, V. De Silva, and J. C. Langford, “A global geometric framework for nonlinear dimensionality reduction,” *Science (1979)*, 2000, doi: 10.1126/science.290.5500.2319.

[381] P. Papaioannou, R. Talmon, I. G. Kevrekidis, and C. Siettos, “Time Series Forecasting Using Manifold Learning,” Oct. 2021, Accessed: Dec. 19, 2021. [Online]. Available: https://arxiv.org/abs/2110.03625v4

[382] N. Puchkin, A. Timofeev, and V. Spokoiny, “Manifold-based time series forecasting,” Dec. 2020, Accessed: Dec. 19, 2021. [Online]. Available: https://arxiv.org/abs/2012.08244v1

ΠΑΡΑΡΤΗΜΑ Ι: Ορισμοί και Αλγόριθμοι

- Διαδικασία $i.i.d(0, \sigma^2)$: οι τυχαίες μεταβλητές της διαδικασίας είναι ανεξάρτητα και ταυτόσημα κατανεμημένες και έχουν κοινή κατανομή πιθανότητας με μηδενικό μέσο όρο και διακύμανση σ^2.

Αλγόριθμος 1: Αλγόριθμος εκπαίδευσης.

Είσοδος: Το σύνολο δεδομένων εκπαίδευσης $\boldsymbol{F}^T$, το σύνολο δεδομένων αξιολόγησης $\boldsymbol{F}^V$, τους πίνακες γειτνίασης $\boldsymbol{A}^T$, $\boldsymbol{A}^V$

Έξοδος: Τα βάρη $\boldsymbol{W}$ των νευρωνικών δικτύων *AGGREGATE* και *UPDATE*

ΕΚΚΙΝΗΣΗ

ΘΕΣΕ: Αυθαίρετες τιμές στα προς εκμάθηση βάρη $\boldsymbol{W}$.
ΥΠΟΛΟΓΙΣΕ: $\boldsymbol{x} = TRANSITION\,(\boldsymbol{W}, \boldsymbol{F}^T, \boldsymbol{A}^T\,)$
ΘΕΣΕ: $K = 1$
ΟΣΟ: $K \leq Μέγιστο\ αριθμό\ εποχών\ εκμάθησης$
    ΥΠΟΛΟΓΙΣΕ: $\frac{\partial \mathbf{e}}{\partial \mathbf{W}}(K) = GRADIENT\big(\boldsymbol{x}, \boldsymbol{W}(\boldsymbol{K})\big)$
        ΑΝ: $K > 1$
            ΥΠΟΛΟΓΙΣΕ:$\boldsymbol{W}(K) = RPROP\left(\frac{\partial \mathbf{e}}{\partial \mathbf{W}}(K), \frac{\partial \mathbf{e}}{\partial \mathbf{W}}(K-1), \boldsymbol{W}(K-1)\right)$
    ΤΕΛΟΣ_ΑΝ
    ΑΝ: $Βήματα\ αξιολόγησης\ MOD\ K = 0$
            ΥΠΟΛΟΓΙΣΕ: $Τρέχων\ βήμα\ =\ Βήματα\ αξιολόγησης\ MOD\ K$
            ΘΕΣΕ: $M\ =\ 1$
            ΟΣΟ: $M \leq Μέγιστος\ αριθμός\ βημάτων\ αξιολόγησης$
                ΥΠΟΛΟΓΙΣΕ: $\boldsymbol{x}_V\ =\ TRANSITION(\boldsymbol{W}, \boldsymbol{F}^V, \boldsymbol{A}^V)$
            ΤΕΛΟΣ_ΟΣΟ
            ΥΠΟΛΟΓΙΣΕ: $\boldsymbol{o}_V\ =\ UPDATE\ (\boldsymbol{x}_V, \boldsymbol{l}_N)$
            ΥΠΟΛΟΓΙΣΕ: $\boldsymbol{e}_V(Τρέχον\ βήμα) =\ \boldsymbol{t}_V - \boldsymbol{o}_V$
            ΑΝ: $\boldsymbol{e}_V(Τρέχον\ βήμα) < \boldsymbol{e}_V(Τρέχον\ βήμα - 1)$
                ΕΠΙΣΤΡΕΨΕ: $\mathbf{W}(K)$
            ΤΕΛΟΣ_ΑΝ
    ΤΕΛΟΣ_ΑΝ
ΤΕΛΟΣ_ΟΣΟ

ΤΕΡΜΑΤΙΣΜΟΣ

Αλγόριθμος 2: Αλγόριθμος TRANSITION.

Είσοδος: Τα βάρη $\boldsymbol{W}$ που έχουν προκύψει από τον αλγόριθμο **Αλγόριθμος 4**, το σύνολο δεδομένων αναφοράς $\boldsymbol{F}$, τον πίνακα γειτνίασης $\boldsymbol{A}$.

Έξοδος: Τις καταστάσεις $\boldsymbol{x}_n$ των κόμβων ενός γράφου $\mathcal{G}$.

ΕΚΚΙΝΗΣΗ

1: **ΘΕΣΕ**: Βήμα επανάληψης $t = 0$.

2: **ΘΕΣΕ**: Αυθαίρετες τιμές στο διάνυσμα καταστασης $\boldsymbol{x}(0)$.

3: **ΥΠΟΛΟΓΙΣΕ**: $\boldsymbol{l}_{co[n]} = \left[F_{n,3}; \dots; F_{n,D}\right] * \left[A_{n,1}; A_{n,2}; \dots; A_{n,N}\right]$

4: **ΥΠΟΛΟΓΙΣΕ**: $\boldsymbol{l}_{ne[n]} = \left[F_{1,2,}; \dots; F_{N,2}\right] * \left[A_{n,1}; A_{n,2}; \dots; A_{n,N}\right]$

5: **ΥΠΟΛΟΓΙΣΕ**: $\boldsymbol{x}_{ne[n]} = \left[\boldsymbol{x}_{1,}; \dots; \boldsymbol{x}_N\right] * \left[A_{n,1}; A_{n,2}; \dots; A_{n,N}\right]$

6: **ΟΣΟ**: $\|\boldsymbol{x}(t) - \boldsymbol{x}(t-1)\| > \varepsilon_f$

7: **ΥΠΟΛΟΓΙΣΕ**: $h_n(t) = AGGREGATE\left(l_n, \boldsymbol{l}_{co\,[n]}, \boldsymbol{l}_{ne\,[n]}, \boldsymbol{x}_{ne[n]}\right)$

8: **ΥΠΟΛΟΓΙΣΕ**: $\boldsymbol{x}_n(t+1) = \sum_u(\{h_u(t), \forall u \in ne[n]\})$.

9: **ΘΕΣΕ**: $t = t + 1$.

10: **ΤΕΛΟΣ_ΟΣΟ**

11: **ΕΠΕΣΤΡΕΨΕ**: Το διάνυσμα κατάστασης $\boldsymbol{x}_n$.

ΤΕΡΜΑΤΙΣΜΟΣ

Αλγόριθμος 3: Αλγόριθμος GRADIENT .

Είσοδος: Τα βάρη $\boldsymbol{W}$ που έχουν προκύψει από τον αλγόριθμο **Αλγόριθμος 4.**

Έξοδος: Την κλίση του σφάλματος εκπαίδευσης $\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{W}}$.

ΕΚΚΙΝΗΣΗ

ΥΠΟΛΟΓΙΣΕ: $\boldsymbol{o}(\boldsymbol{x}, \boldsymbol{l}_N) = UPDATE(\boldsymbol{x}, \boldsymbol{l}_N)$.

ΥΠΟΛΟΓΙΣΕ: $\boldsymbol{e}(\boldsymbol{x}, \boldsymbol{l}_N) = \sum_{i=1}^{N}(t_i - o_i)$

ΥΠΟΛΟΓΙΣΕ: $h_n(\boldsymbol{x}, \boldsymbol{l}) = AGGREGATE\left(l_n, \boldsymbol{l}_{co[n]}, \boldsymbol{l}_{ne[n]}, \boldsymbol{x}_{ne[n]}\right)$

ΥΠΟΛΟΓΙΣΕ: $H(\boldsymbol{x}, \boldsymbol{l}) = \sum_u(\{h_u, \forall u \in ne[n]\})$.

ΥΠΟΛΟΓΙΣΕ: $C = \frac{\partial H}{\partial \boldsymbol{x}}(\boldsymbol{x}, \boldsymbol{l})$.

ΥΠΟΛΟΓΙΣΕ: $B = \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{o}} \cdot \frac{\partial \boldsymbol{o}}{\partial \boldsymbol{x}}(\boldsymbol{x}, \boldsymbol{l}_N)$.

ΘΕΣΕ: Βήμα επανάληψης $t = 0$.

ΘΕΣΕ: Αυθαίρετες τιμές στο δίανυσμα σύγκλισης $\boldsymbol{z}(0)$.

ΘΕΣΕ: $I = 1$.

ΟΣΟ: $I < Μέγιστος\ αριθμός\ βημάτων\ οπισθοδιάδοσης$

  ΥΠΟΛΟΓΙΣΕ: $\boldsymbol{z}(t) = \boldsymbol{z}(t+1) \cdot C + B$.

  ΘΕΣΕ: $t = t - 1$.

  ΘΕΣΕ: $I = I + 1$

ΤΕΛΟΣ_ΟΣΟ

ΥΠΟΛΟΓΙΣΕ: $P = \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{o}} \cdot \frac{\partial \boldsymbol{o}}{\partial \boldsymbol{W}}(\boldsymbol{x}, \boldsymbol{l}_N)$.

ΥΠΟΛΟΓΙΣΕ: $Q = \boldsymbol{z}(t) \cdot \frac{\partial H}{\partial \boldsymbol{W}}(\boldsymbol{x}, \boldsymbol{l})$.

ΥΠΟΛΟΓΙΣΕ: $\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{w}} = P + Q$.

ΕΠΕΣΤΡΕΨΕ: Την κλίση $\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{W}}$.

ΤΕΡΜΑΤΙΣΜΟΣ

Αλγόριθμος 4: Αλγόριθμος RPROP.

Είσοδος: Τα βάρη που έχουν προκύψει από το προηγούμενο χρονικό βήμα $\boldsymbol{W}(K-1)$, την κλίση του σφάλματος εκπαίδευσης του προηγούμενου και του τρέχοντος χρονικού βήματος, $\frac{\partial \mathbf{e}}{\partial \mathbf{w}}(K)$ και ,$\frac{\partial \mathbf{e}}{\partial \mathbf{w}}(K-1)$, αντίστοιχα που έχουν προκύψει από τον αλγόριθμο **Αλγόριθμος 3**, και το τρέχον βήμα επανάληψης K.

Έξοδος: Τα ανανεωμένα βάρη των νευρωνικών δικτύων *AGGREGATE* και *UPDATE*.

ΕΚΚΙΝΗΣΗ

1: **ΑΝ**: $K = 2$

2: **ΘΕΣΕ**: Κάποια αρχική τιμή $\boldsymbol{\Delta}(1)$ για τους συντελεστές ανανέωσης.

3: **ΤΕΛΟΣ_ΑΝ**

4: **ΥΠΟΛΟΓΙΣΕ**:

$$\boldsymbol{\Delta}(K) = \begin{cases} \eta^{+} * \boldsymbol{\Delta}(K-1), if\ \frac{\partial \mathbf{e}}{\partial \mathbf{w}}(K-1) * \frac{\partial \mathbf{e}}{\partial \mathbf{w}}(K) > 0 \\ \eta^{-} * \boldsymbol{\Delta}(K-1), if\ \frac{\partial \mathbf{e}}{\partial \mathbf{w}}(K-1) * \frac{\partial \mathbf{e}}{\partial \mathbf{w}}(K) < 0 \\ 0, αλλιώς \end{cases}$$

5: **ΥΠΟΛΟΓΙΣΕ**: $\boldsymbol{\Delta W}(K) = \begin{cases} -\boldsymbol{\Delta}(K), if\ \frac{\partial \mathbf{e}}{\partial \mathbf{w}}(K) > 0 \\ \boldsymbol{\Delta}(K), if\ \frac{\partial \mathbf{e}}{\partial \mathbf{w}}(K) < 0 \\ 0, αλλιώς \end{cases}$

6: **ΥΠΟΛΟΓΙΣΕ**: $\boldsymbol{W}(K) = \boldsymbol{W}(K-1) + \boldsymbol{\Delta W}(K)$

7: **ΕΠΕΣΤΡΕΨΕ**: Τα ανανεωμένα βάρη $\boldsymbol{W}(K)$

ΤΕΡΜΑΤΙΣΜΟΣ

Αλγόριθμος 5: Αλγόριθμος ελέγχου.

Είσοδος: Τα βάρη $\boldsymbol{W}$ που έχουν προκύψει από τον αλγόριθμο **Αλγόριθμος 1**, το σύνολο δεδομένων ελέγχου $\boldsymbol{F}^E$, τον πίνακα γειτνίασης $\boldsymbol{A}^E$

Έξοδος: Τις τιμές των προβλέψεων του ηλεκτρικού φορτίου, δοθέντος ενός συνόλου δεδομένων ελέγχου, άγνωστων για το νευρωνικό δίκτυο Γράφων ορατότητας.

ΕΚΚΙΝΗΣΗ

1: **ΥΠΟΛΟΓΙΣΕ:** $\boldsymbol{x} = TRANSITION\,(\boldsymbol{W}, \boldsymbol{F}^E, \boldsymbol{A}^E)$

2: **ΥΠΟΛΟΓΙΣΕ**: $\boldsymbol{o}(\boldsymbol{x}, \boldsymbol{l}_N) = \ UPDATE\ (\boldsymbol{x}, \boldsymbol{l}_N)$

3: **ΕΠΙΣΤΡΕΨΕ**: Τις τιμές των προβλέψεων του ηλεκτρικού φορτίου $\mathbf{o}(\boldsymbol{x}, \boldsymbol{l}_N)$

ΤΕΡΜΑΤΙΣΜΟΣ