ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ
ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

Βραχυπρόθεσμη πρόβλεψη ενεργειακής ζήτησης Προσεγγίσεις βασισμένες στη Μηχανική Μάθηση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Παναγιώτης Γ. Λαδάς

Επιβλέπων: Ιωάννης Ψαρράς
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2014

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ
ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

Βραχυπρόθεσμη πρόβλεψη ενεργειακής ζήτησης Προσεγγίσεις βασισμένες στη Μηχανική Μάθηση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Παναγιώτης Γ. Λαδάς

Επιβλέπων: Ιωάννης Ψαρράς
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 29η Οκτωβρίου 2014

.....	.....	.....
Ιωάννης Ψαρράς	Δημήτριος Ασκούνης	Βασίλειος Ασημακόπουλος
Καθηγητής ΕΜΠ	Αν. Καθηγητής ΕΜΠ	Καθηγητής ΕΜΠ

Αθήνα, Οκτώβριος 2014

.....
Παναγιώτης Γ. Λαδάς

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Παναγιώτης Γ. Λαδάς, 2014
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν το συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

ΠΕΡΙΛΗΨΗ

Η πρόβλεψη της μελλοντικής ενεργειακής ζήτησης είναι πολύ σημαντική κατά τη διαδικασία λήψης αποφάσεων στον τομέας της ηλεκτρικής ενέργειας, ιδιαίτερα μέσα στο ανταγωνιστικό περιβάλλον που δημιουργήθηκε μετά την απελευθέρωση των αγορών ενέργειας. Η βραχυπρόθεσμη πρόβλεψη φορτίου ηλεκτρισμού συμβάλλει καθοριστικά στην αξιοπιστία ενός συστήματος ηλεκτρικής ενέργειας (ΣΗΕ). Επιπλέον είναι απαραίτητη για την ομαλή λειτουργία του ΣΗΕ, ενώ υψηλά σφάλματα στην πρόβλεψη φορτίου μπορούν να οδηγήσουν σε αύξηση του λειτουργικού κόστους.

Διάφορες τεχνικές έχουν αναπτυχθεί για τη βραχυπρόθεσμη πρόβλεψη της ηλεκτρικής κατανάλωσης (STLF). Οι τεχνικές αυτές διακρίνονται κυρίως σε στατιστικές μεθόδους πρόβλεψης και μοντέλα τεχνητής νοημοσύνης (AI). Το φορτίο ενός συστήματος αποτελεί μια στοχαστική, μη στάσιμη διαδικασία ενώ η συμπεριφορά του επηρεάζεται από πλήθος παραγόντων με πιο σημαντικές τις καιρικές συνθήκες, το είδος του τελικού χρήστη, τους εποχιακούς παράγοντες και τα τυχαία γεγονότα. Λόγω των μη γραμμικών σχέσεων που εμφανίζονται μεταξύ της ηλεκτρικής κατανάλωσης και των επιμέρους παραμέτρων, τα τελευταία χρόνια έχει δοθεί έμφαση σε μεθόδους μηχανικής μάθησης (Machine Learning) για την επίλυση τέτοιων προβλημάτων. Τα μοντέλα μηχανικής μάθησης βασίζονται στα διαθέσιμα δεδομένα και εκπαιδεύονται μέσα από αυτά. Σε αυτήν την κατηγορία ανήκουν τα τεχνητά νευρωνικά δίκτυα και οι μηχανές διανυσμάτων υποστήριξης (Support Vector Machines).

Αντικείμενο της διπλωματικής εργασίας είναι η ανάπτυξη μιας μεθοδολογίας για τη βραχυπρόθεσμη πρόβλεψη της ενεργειακής ζήτησης στον κτηριακό τομέα. Το πρόβλημα της εξοικονόμησης ενέργειας για οικονομικούς και περιβαλλοντικούς λόγους, καθιστά την ύπαρξη μιας τέτοιας μεθοδολογίας εξαιρετικά σημαντική. Σκοπός επομένως της παρούσας εργασίας είναι η πρόταση συγκεκριμένων λύσεων μέσω μιας πιλοτικής εφαρμογής. Σε αυτή γίνεται χρήση δεδομένων πραγματικού χρόνου, που συλλέχθηκαν από “έξυπνους μετρητές” εγκατεστημένους σε μια αλυσίδα δέκα εμπορικών καταστημάτων στην περιοχή της Αττικής.

Αρχικά επιχειρείται η ανάλυση συσχετίσεων μεταξύ της ηλεκτρικής κατανάλωσης και ορισμένων βασικών παραμέτρων που την επηρεάζουν (εσωτερική και εξωτερική θερμοκρασία, ώρες ηλιοφάνειας, είδος ημέρας κ.α.). Στη συνέχεια πραγματοποιείται η βραχυπρόθεσμη πρόβλεψη της ενεργειακής ζήτησης με τη χρήση διανυσμάτων υποστήριξης (SVM). Για την υλοποίηση των αλγορίθμων στην πιλοτική εφαρμογή γίνεται χρήση της γλώσσας Python και της βιβλιοθήκης Scikit-Learn, που αποτελεί δημοφιλή βιβλιοθήκη μηχανικής μάθησης. Τέλος προτείνεται η καταλληλότερη χρονική βάση για τα δεδομένα εισόδου και η βέλτιστη συνάρτηση πυρήνα (kernel) για τη βραχυπρόθεσμη πρόβλεψη με χρονικό ορίζοντα μια ημέρα.

Λέξεις κλειδιά: βραχυπρόθεσμη πρόβλεψη, ενεργειακή ζήτηση, μηχανική μάθηση, μηχανές διανυσμάτων υποστήριξης, συναρτήσεις πυρήνα, κτήριο.

ABSTRACT

Future energy demand forecasting is very important for the decision-making process in the electricity sector especially in the competitive environment created after the deregulation of the energy markets. The short-term load forecasting (STLF) contributes significantly to the reliability of an electric power system. In addition, it is necessary for the smooth operation of an electric power system, while high errors in load forecasting can increase operational cost.

Several methods have been developed for the short-term load forecasting (STLF). These methods are mainly divided in statistical forecasting methods and artificial intelligence (AI) methods. The load of a system is a stochastic, non-stationary process, whose behavior is affected by a number of factors. The most important among these are the weather conditions, the type of the end user, the seasonal factors and random events. Due to the non-linear relations between electric consumption and its components, particular emphasis has been given in the last years on machine learning methods to solve these problems. The machine learning methods are based on the available data and trained through these. In this category belong the artificial neural networks and the support vector machines (SVMs).

The objective of the diploma thesis is to develop a methodology for the short-term load forecasting in the building sector. Such a methodology is particularly important because of the problem of saving energy for environmental and economic reasons. This thesis proposes specific solutions through a pilot application. This application uses real time data gathered through smart meters placed in a chain of ten commercial shops in the Attica region.

Initially, it is attempted to analyze the correlation between electric consumption and some key parameters that affect it (internal and external temperature, daylight hours, types of day etc.). This is followed by the short-term prediction of energy demand (load forecasting) with the use of the SVMs. For the implementation of the algorithms in the pilot application, the Python language and the Scikit-Learn library, a popular library of machine learning, are used. At the end, the most appropriate time basis for the entry data is proposed as well as the optimal kernel function for the short-term forecasting with the horizon of one day.

Keywords: short-term forecasting, energy demand, machine learning, support vector machines, kernel functions, building.

ΠΡΟΛΟΓΟΣ

Η παρούσα διπλωματική εργασία εκπονήθηκε κατά το ακαδημαϊκό έτος 2013-2014 στον τομέα Ηλεκτρικών Βιομηχανικών Διατάξεων και Συστημάτων Αποφάσεων της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου.

Υπεύθυνος κατά την εκπόνηση της διπλωματικής ήταν ο Καθηγητής κ. Ιωάννης Ψαρράς στον οποίο οφείλω ιδιαίτερες ευχαριστίες για την ανάθεση αυτής και τη δυνατότητα που μου δόθηκε να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα. Θα ήθελα επίσης να ευχαριστήσω θερμά τον επιβλέποντα της διπλωματικής και Λέκτορα κ. Χάρη Δούκα για την υποστήριξη και την καθοδήγηση που μου παρείχε κατά τη συγγραφή της εργασίας. Ακόμα θα ήθελα να ευχαριστήσω τον Αναπλ. Καθηγητή κ. Δ. Ασκούνη και τον Καθηγητή κ. Β. Ασημακόπουλο για την τιμή που μου έκαναν να συμμετάσχουν στην επιτροπή εξέτασης της εργασίας.

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω το Διδάκτορα κ. Βασίλειο Νικολόπουλο, συνιδρυτή και διευθύνοντα σύμβουλο της εταιρείας Intelen όπου είχα την τύχη να κάνω την πρακτική μου εργασία. Οι πολύτιμες γνώσεις που αποκόμισα και η δυνατότητα χρήσης πραγματικών ενεργειακών δεδομένων για την πιλοτική εφαρμογή ήταν πολύ σημαντικά για την υλοποίηση της διπλωματικής εργασίας.

Κυρίως όμως θέλω να εκφράσω την ευγνωμοσύνη μου στην οικογένειά μου για την αμέριστη στήριξή τους όλα αυτά τα χρόνια ώστε να προσβλέπω δυναμικά στην υλοποίηση των επιδιωκόμενων στόχων μου. Τέλος θα ήταν παράλειψη να μην ευχαριστήσω όλους τους δικούς μου ανθρώπους για τη διάθεση και την κατανόηση που επέδειξαν σε όλη μου την πορεία.

Παναγιώτης Γ. Λαδάς

Οκτώβριος 2014

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ 5

ABSTRACT 7

ΠΡΟΛΟΓΟΣ 9

ΠΕΡΙΕΧΟΜΕΝΑ ΣΧΗΜΑΤΩΝ 15

ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ 16

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ 19

1.1 Αντικείμενο και Σκοπός 19

1.2 Δομή διπλωματικής εργασίας 21

ΚΕΦΑΛΑΙΟ 2: ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΠΡΟΒΛΕΨΗ 25

2.1 Εξόρυξη Δεδομένων 25

2.1.1 Εισαγωγή στην Εξόρυξη Δεδομένων 25

2.1.2 Σημασία της Εξόρυξης Δεδομένων 26

2.1.3 Μοντέλα και Μέθοδοι Εξόρυξης Δεδομένων 27

2.1.3.1 Ταξινόμηση 29

2.1.3.2 Παλινδρόμηση 29

2.1.3.3 Ανάλυση Χρονοσειρών 29

2.1.3.4 Πρόβλεψη 29

2.1.3.5 Συσταδοποίηση 30

2.1.3.6 Περίληψη 30

2.1.3.7 Κανόνες Συσχετίσεων 31

2.1.3.8 Ανακάλυψη Ακολουθιών 31

2.2 Πρόβλεψη Ζήτησης Φορτίου 31

2.2.1 Πρόβλεψη Ζήτησης Φορτίου 31

2.2.2 Παράγοντες που επηρεάζουν τη ζήτηση φορτίου 34

2.2.3 Βραχυπρόθεσμη πρόβλεψη φορτίου 35

ΚΕΦΑΛΑΙΟ 3: ΤΕΧΝΙΚΕΣ ΠΡΟΒΛΕΨΕΩΝ 39

3.1 Γενικά για τις Προβλέψεις 39

3.1.1 Κατηγορίες μεθόδων πρόβλεψης 40

3.1.2 Ορίζοντας πρόβλεψης 42

3.1.3 Διαδικασία Πρόβλεψης 43
3.2 Χρονοσειρές 44
3.2.1 Γενικά για τις Χρονοσειρές 44
3.2.2 Ποιοτικά Χαρακτηριστικά Χρονοσειρών 45
3.3 Μοντέλα Πρόβλεψης 47
3.3.1 Μοντέλο Χρονοσειρών 47
3.3.2 Αιτιοκρατικό ή Επεξηγηματικό μοντέλο 47
3.4 Κατηγορίες Μεθόδων Πρόβλεψης Φορτίου 48
3.4.1 Μέθοδοι Βραχυπρόθεσμης Πρόβλεψης Φορτίου 49
3.4.2 Μέθοδοι Μεσοπρόθεσμης και Μακροπρόθεσμης Πρόβλεψης Φορτίου 49
3.4.2.1 Μοντέλα Τελικής Χρήσης 50
3.4.2.2 Οικονομετρικά Μοντέλα 51
3.4.2.3 Στατιστικά Μοντέλα Μάθησης 51
3.4.3 Μέθοδοι Πολύ Βραχυπρόθεσμης Πρόβλεψης 52
3.4.4 Συνδυασμός Μεθόδων Πρόβλεψης 52
3.5 Δείκτες Αξιολόγησης των Μεθόδων Πρόβλεψης 53
3.6 Επιλογή Κατάλληλης Μεθόδου Πρόβλεψης 57
3.7 Γενικοί Τύποι Ενεργειακών Μοντέλων 59
3.7.1 Παραδοσιακά Ενεργειακά Μοντέλα 59
3.7.2 Ενεργειακά Μοντέλα με χρήση αισθητήρων 60
3.7.3 Σύγκριση Ενεργειακών Μοντέλων 60

ΚΕΦΑΛΑΙΟ 4: ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ 65
4.1 Στατιστικές Μέθοδοι Πρόβλεψης 65
4.1.1. Απλοϊκή Μέθοδος 65
4.1.2 Μέθοδοι Εξομάλυνσης 65
4.1.2.1 Μέθοδοι απλού ή κινητού μέσου όρου 66
4.1.2.2 Μέθοδοι εκθετικής εξομάλυνσης 67
1. Μοντέλο Σταθερού Επιπέδου 68
2. Μοντέλο Γραμμικής Τάσης 70
3. Μοντέλο μη Γραμμικής Τάσης 71
4.1.3 Μέθοδοι Αποσύνθεσης 73
4.1.4 Μέθοδος Theta 73
4.1.5 Μοντέλα Παλινδρόμησης 74
4.1.5.1 Απλή Γραμμική Παλινδρόμηση 75

4.1.5.2 Πολλαπλή Γραμμική Παλινδρόμηση 76
4.1.6 Οικονομετρική μέθοδος 77
4.1.7 Στοχαστικές Χρονοσειρές 77
4.1.7.1 Αυτοπαλινδρομικά Μοντέλα (AR) 78
4.1.7.2 Αυτοπαλινδρομικά Μοντέλα Κινητού Μέσου Όρου (ARMA) 78
4.1.7.3 Ολοκληρωμένα Αυτοπαλινδρομικά Μοντέλα Κ.Μ.Ο (ARIMA) 79
4.1.8 Προσέγγιση της παρόμοιας ημέρας 80
4.1.9 Λοιπές Παλινδρομικές και Μη Μέθοδοι 80
4.1.10 Μέθοδος Επαναληπτικών Επανασταθμισμένων Ελαχίστων Τετραγώνων 81
4.2 Μοντέλα Τεχνητής Νοημοσύνης 81
4.2.1 Τεχνητά Νευρωνικά Δίκτυα (ANN) 82
4.2.2 Μηχανές Διανυσμάτων Υποστήριξης (SVM) 83
4.2.3 Ασαφής Λογική 84
4.2.4 Έμπειρα Συστήματα που βασίζονται στη γνώση 86
4.2.5 Παλινδρόμηση Κ Κοντινότερων Γειτόνων (KNN Regression) 87
4.2.6 Δένδρα Ταξινόμησης και Παλινδρόμησης (CART) 87
4.2.7 Αυτοπαλινδρομικά Μοντέλα Κ.Μ.Ο με Εξωγενή Μεταβλητή (ARMAX) 89
4.2.8 Γκαουσιανή Μέθοδος (GP) 90

ΚΕΦΑΛΑΙΟ 5: ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ 93
5.1 Εισαγωγή στη Μηχανική Μάθηση 93
5.2 Τεχνητά Νευρωνικά Δίκτυα (ANN) 94
5.2.1 Πολυεπίπεδα Δίκτυα Prceptron (MLP) 95
5.2.2 Δίκτυα Πρόσθιας Τροφοδότησης (FFNN) 96
5.2.3 Δίκτυα Συνάρτησης Ακτινωτής Βάσης (RBFN) 96
5.2.4 Δίκτυα Γενικευμένης Παλινδρόμησης (GRNN) 97
5.2.5 Νευρωνικά Δίκτυα BNN 98
5.2.6 Δυνάμεις και Αδυναμίες των Νευρωνικών Δικτύων 99
5.3 Μηχανές Διανυσμάτων Υποστήριξης (SVM) 100
5.3.1 Βασική Ιδέα 100
5.3.2 Ελαχιστοποίηση Διαρθρωτικού Ρίσκου 101
5.3.3 Παλινδρόμηση Διανυσμάτων Υποστήριξης (SVR) 103
5.3.4 Δυϊκό Πρόβλημα και Τέχνασμα Πυρήνων 105
5.3.5 Δυνάμεις και Αδυναμίες των SVM 107

ΚΕΦΑΛΑΙΟ 6: ΠΕΙΡΑΜΑΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ 111
6.1 Εφαρμογή Μηχανικής Μάθησης 111
6.1.1 Μέθοδοι προεπεξεργασίας δεδομένων 112
6.1.2 Python 114
6.1.3 Βιβλιοθήκη Scikit-Learn 115
6.2 Πιλοτική Εφαρμογή 118
6.2.1 Διερεύνηση Δεδομένων Εισόδου 118
6.2.2 Ανάλυση Συσχετίσεων 123
6.2.3 Αλγόριθμοι SVR 130
6.2.4 Αποτελέσματα Πρόβλεψης 132

ΚΕΦΑΛΑΙΟ 7: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΟΠΤΙΚΕΣ 143
7.1 Συμπεράσματα 143
7.2 Προοπτικές και Μελλοντικές Προεκτάσεις 146

ΒΙΒΛΙΟΓΡΑΦΙΑ 149

ΠΕΡΙΕΧΟΜΕΝΑ ΣΧΗΜΑΤΩΝ

Σχήμα 2.1: Διαδικασία εξόρυξης δεδομένων [7] 26

Σχήμα 2.2: Τα βασικότερα βήματα της διαδικασίας Εξόρυξης Γνώσης [10] 27

Σχήμα 2.3: Ταξινόμηση των Τεχνικών Εξόρυξης Δεδομένων [12] 28

Σχήμα 2.4: Τα στάδια της διαδικασίας Συσταδοποίησης [21] 30

Σχήμα 3.1: Κατηγοριοποίηση μεθόδων πρόβλεψης με βάση τη διαδικασία παραγωγής ... 42

Σχήμα 3.2: Βασικές κατηγορίες μεθόδων πρόβλεψης 49

Σχήμα 4.1: Κατηγοριοποίηση μεθόδων εκθετικής εξομάλυνσης [63] 68

Σχήμα 5.1: Διαδικασία μετασχηματισμού ενός μοντέλου SVR [157] 103

Σχήμα 6.1: Διάγραμμα ροής του αλγόριθμου Scikit-Learn [178] 116

Σχήμα 6.2: Ακατέργαστα δεδομένα τετάρτου του μετρητή 4 118

Σχήμα 6.3: Ακατέργαστα δεδομένα τετάρτου του μετρητή 9 119

Σχήμα 6.4: Ακατέργαστα ωριαία δεδομένα του μετρητή 4 120

Σχήμα 6.5: Ακατέργαστα ωριαία δεδομένα του μετρητή 9 120

Σχήμα 6.6: Ακατέργαστα ημερήσια δεδομένα του μετρητή 4 121

Σχήμα 6.7: Ακατέργαστα ημερήσια δεδομένα του μετρητή 9 122

Σχήμα 6.8: Ημερήσια δεδομένα ηλεκτρικής κατανάλωσης του μετρητή 4 122

Σχήμα 6.9: Ημερήσια δεδομένα ηλεκτρικής κατανάλωσης του μετρητή 9 123

Σχήμα 6.10: Παραδείγματα συντελεστών συσχέτισης Pearson και MIC [181] 124

Σχήμα 6.11: Ημερήσια δεδομένα του μετρητή 4 125

Σχήμα 6.12: Ανάλυση συσχετίσεων του μετρητή 4 127

Σχήμα 6.13: Ανάλυση συσχετίσεων του μετρητή 9 127

Σχήμα 6.14: Ανάλυση συσχετίσεων του μετρητή 3 128

Σχήμα 6.15: Αποτελέσματα cross validation για ημερήσια δεδομένα του μετρητή 4, με τη μέθοδο Gaussian RBF 138

Σχήμα 6.16: Αποτελέσματα cross validation για ημερήσια δεδομένα του μετρητή 9, με τη μέθοδο Gaussian RBF 138

ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ

Πίνακας 5.1: Τύποι πυρήνων εσωτερικού γινομένου $K(x_i,x_j)$ [159] 107

Πίνακας 6.1: Αποτελέσματα συσχετίσεων των μετρητών 3, 4 και 9 129

Πίνακας 6.2: Αποτελέσματα συσχετίσεων των μετρητών 1, 2, 5, 6, 7, 8 και 10 130

Πίνακας 6.3: Εξισώσεις των εξεταζόμενων συναρτήσεων πυρήνα 131

Πίνακας 6.4: Αποτελέσματα μοντέλων SVR για δεδομένα τετάρτου 134

Πίνακας 6.5: Αποτελέσματα μοντέλων SVR για ωριαία δεδομένα..... 135

Πίνακας 6.6: Αποτελέσματα μοντέλων SVR για ημερήσια δεδομένα 137

Πίνακας 6.7: Μέση συσχέτιση παραμέτρων με την ηλεκτρική κατανάλωση 144

Πίνακας 6.8: Μέση τιμή σφαλμάτων των εξεταζόμενων πυρήνων 145

Πίνακας 6.9: Αποτελέσματα του Gaussian πυρήνα ακτινωτής βάσης 146

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

1.1 Αντικείμενο και Σκοπός

Η ακριβής πρόβλεψη της μελλοντικής ενεργειακής ζήτησης μετά την απελευθέρωση των αγορών ηλεκτρικής ενέργειας κατέχει πολύ σημαντικό ρόλο όσον αφορά στους τομείς των επενδύσεων σε συστήματα διανομής, τον προγραμματισμό του φορτίου ηλεκτρικής ενέργειας και τις στρατηγικές διαχείρισης σε περιφερειακά και εθνικά συστήματα. Το σφάλμα της πρόβλεψης του ηλεκτρικού φορτίου μπορεί να οδηγήσει στην αύξηση του λειτουργικού κόστους. Σε περίπτωση υποεκτίμησης των μελλοντικών φορτίων η αδυναμία παροχής επαρκών αποθεμάτων συνεπάγεται υψηλό κόστος ανά μονάδα αιχμής. Αντίθετα η υπερεκτίμηση των μελλοντικών φορτίων έχει ως αποτέλεσμα άσκοπη στρεφόμενη εφεδρεία και ένα πλεόνασμα προσφοράς που είναι ανεπιθύμητο από τα δίκτυα ενέργειας. Σύμφωνα με τις εφαρμογές της πρόβλεψης του φορτίου γίνεται εύκολα αντιληπτή η σημασία της καθώς συμμετέχει στον οικονομικό προγραμματισμό ενός συστήματος παραγωγής ενέργειας (ΣΗΕ) και συμβάλλει καθοριστικά στην αξιοπιστία του συστήματος και τη διατήρηση της ασφάλεια του ΣΗΕ.

Η ενεργειακή ζήτηση αποτελεί ένα σημαντικό και πολύπλευρο πρόβλημα, το οποίο βρίσκεται στο επίκεντρο της παγκόσμιας ερευνητικής κοινότητας. Τα οικιακά και εμπορικά κτήρια συνιστούν το μεγαλύτερο κομμάτι της πρωτογενούς κατανάλωσης ενέργειας σε παγκόσμιο επίπεδο. Καθώς η ενεργειακή αποδοτικότητα των κτηρίων περιγράφεται συχνά ως ο πιο εύκολος στόχος για τη μείωση της κατανάλωσης ενέργειας και των απαιτούμενων εκπομπών αερίων του θερμοκηπίου, αναδεικνύεται η κρισιμότητα για την πρόβλεψη της ζήτησης του ηλεκτρικού φορτίου.

Το φορτίο ενός συστήματος αποτελεί μια στοχαστική, μη στάσιμη διαδικασία η οποία απαρτίζεται από ένα σύνολο επιμέρους στοιχείων. Επομένως η συμπεριφορά του φορτίου ενός συστήματος επηρεάζεται από ένα πλήθος παραγόντων που μπορούν να ταξινομηθούν με βάση τη βιβλιογραφία στις ακόλουθες βασικές κατηγορίες: καιρικές συνθήκες, οικονομικοί παράγοντες, ημερολογιακοί παράγοντες, είδος τελικού χρήστη και τυχαία γεγονότα. Στο πλαίσιο της παρούσας διπλωματικής οι παράμετροι που συγκεντρώθηκαν για τη μοντελοποίηση της ενεργειακής ζήτησης είναι: η εσωτερική και εξωτερική θερμοκρασία, η πίεση, οι ώρες ηλιοφάνειας ανά ημέρα, το είδος της ημέρας (εργάσιμη ή μη) η κατανάλωση της προηγούμενης ημέρας και της αντίστοιχης ημέρας της προηγούμενης εβδομάδας για την ανάδειξη ημερών με παρόμοια συμπεριφορά.

Οι παράγοντες αυτοί καθιστούν τη διαδικασία πρόβλεψης του ηλεκτρικού φορτίου μια πολύπλοκη διαδικασία. Για την επιλογή της σωστής μεθόδου είναι σημαντική η κατηγοριοποίηση των προβλέψεων με βάση το χρονικό τους ορίζοντα. Έτσι διακρίνονται οι περιπτώσεις της πολύ βραχυπρόθεσμης, βραχυπρόθεσμης, μεσοπρόθεσμης και μακροπρόθεσμης πρόβλεψης. Στην παρούσα διπλωματική το ενδιαφέρον επικεντρώνεται στη βραχυπρόθεσμη πρόβλεψη της ενεργειακής ζήτησης.

Αρκετές μέθοδοι βραχυπρόθεσμης πρόβλεψης φορτίου (STLF – short term load forecasting) συμπεριλαμβανομένων των παραδοσιακών μοντέλων και των μεθόδων τεχνητής νοημοσύνης έχουν προταθεί τις τελευταίες δεκαετίες. Τα στατιστικά μοντέλα χρονοσειρών μπορούν να οδηγήσουν σε αριθμητική αστάθεια και έλλειψη ακρίβειας, εάν αγνοηθούν όπως αναφέραμε παραπάνω, παράγοντες υψηλής συσχέτισης όπως οι καιρικές μεταβλητές. Τα παραδοσιακά μοντέλα συνήθως αντικατοπτρίζουν σταθερές γραμμικές σχέσεις μεταξύ των φορτίων και των καιρικών μεταβλητών. Ωστόσο η πολύπλοκη και μη γραμμική σχέση μεταξύ του φορτίου και των καιρικών συνθηκών μπορεί να τα καταστήσει μη πρακτική επιλογή. Τα τελευταία χρόνια έχει δοθεί μεγάλη προσοχή σε μεθόδους μηχανικής μάθησης, δηλαδή μοντέλα που βασίζονται στα διαθέσιμα δεδομένα και εκπαιδεύονται μέσα από αυτά, όπως για παράδειγμα τα τεχνητά νευρωνικά δίκτυα (ANN – artificial neural networks) και οι μηχανές διανυσμάτων υποστήριξης (SVM – support vector machines). Αντίθετα με την ανάλυση της ζήτησης βασισμένης στα αναλυτικά μοντέλα, οι αλγόριθμοι εξόρυξης δεδομένων προσφέρουν ισχυρά εργαλεία για την ανάδειξη των προτύπων από μεγάλα σύνολα δεδομένων. Ένα ακόμα πλεονέκτημα των μοντέλων που στηρίζονται στα δεδομένα είναι ότι μπορούν εύκολα να ανανεώσουν τη βάση τους. Αυτό το στοιχείο είναι σημαντικό, καθώς η σχέση του φορτίου με τις καιρικές συνθήκες δεν είναι στατική.

Το ευρύτερο αντικείμενο της διπλωματικής εργασίας είναι η πρόβλεψη της ενεργειακής ζήτησης για βραχυπρόθεσμο χρονικό ορίζοντα στον κτηριακό τομέα. Σκοπός είναι η αποτύπωση των βασικών τεχνικών πρόβλεψης και η πρόταση συγκεκριμένων λύσεων μέσω μιας πιλοτικής εφαρμογής σε δέκα εμπορικά καταστήματα κατανεμημένα στην περιοχή της Αττικής. Η διαδικασία αυτή περιγράφεται στα κεφάλαια της παρούσας διπλωματικής όπως αυτά παρουσιάζονται στην επόμενη ενότητα.

Η προσέγγιση της ενεργειακής ζήτησης στη διπλωματική εργασία πραγματοποιήθηκε με τη χρήση των μηχανών διανυσμάτων υποστήριξης (SVM). Πρόκειται για ένα σύνολο μεθόδων επιβλεπόμενης μάθησης το οποίο ανήκει στην κατηγορία της μηχανικής μάθησης όπως αναφέραμε παραπάνω. Οι μέθοδοι SVM στηρίζονται στην αρχή της ελαχιστοποίησης του διαρθρωτικού ρίσκου και όχι στην ελαχιστοποίηση του εμπειρικού ρίσκου, η οποία εφαρμόζεται από τα περισσότερα παραδοσιακά μοντέλα νευρωνικών δικτύων. Αυτό σημαίνει ότι στόχος τους είναι να ελαχιστοποιήσουν την πιθανότητα το μοντέλο που κατασκευάστηκε σύμφωνα με τα δεδομένα εκπαίδευσης να κάνει λάθη σε νέα δεδομένα, με την εύρεση μιας λύσης η οποία γενικεύει με το βέλτιστο τρόπο τα δεδομένα εκπαίδευσης. Με την εισαγωγή της συνάρτησης απωλειών του Vapnik (ε-insensitive loss function), τα SVM επεκτάθηκαν και στην επίλυση μη γραμμικών προβλημάτων παλινδρόμησης. Για την υλοποίηση των αλγορίθμων στην πιλοτική εφαρμογή, έγινε χρήση της γλώσσας Python και της βιβλιοθήκης Scikit-Learn, που αποτελεί δημοφιλή βιβλιοθήκη μηχανικής μάθησης. Τέλος τα σύνολα δεδομένα που χρησιμοποιήθηκαν στη διπλωματική αποτελούνται από ενεργειακά δεδομένα πραγματικού χρόνου (real-time data), τα οποία συλλέχθηκαν από “έξυπνους μετρητές” εγκατεστημένους σε μια αλυσίδα δέκα εμπορικών καταστημάτων στην περιοχή της Αττικής.

1.2 Δομή διπλωματικής εργασίας

Αρχικά παρατίθεται η περίληψη της διπλωματικής εργασίας στην ελληνική και αγγλική γλώσσα, όπου παρουσιάζονται συνοπτικά τα κύρια σημεία της. Στη συνέχεια ακολουθεί ο πρόλογος και ο αναλυτικός πίνακας περιεχομένων. Τέλος ακολουθεί το κύριο μέρος της εργασίας, το οποίο αποτελείται από 7 κεφάλαια.

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

Πρόκειται για το παρόν κεφάλαιο, στο οποίο παρουσιάζεται συνοπτικά το αντικείμενο και ο σκοπός της εργασίας καθώς και το παρόν υποκεφάλαιο, στο οποίο παρουσιάζεται η δομή που υιοθετήθηκε.

ΚΕΦΑΛΑΙΟ 2: ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΠΡΟΒΛΕΨΗ

Στο κεφάλαιο αυτό εισάγεται αρχικά η έννοια της εξόρυξης γνώσης από ένα σύνολο δεδομένων. Περιγράφονται τα στάδια υλοποίησης που απαιτεί η διαδικασία για την εξαγωγή συσχετίσεων και προτύπων από μια βάση δεδομένων. Επίσης παρουσιάζονται οι βασικοί αλγόριθμοι εξόρυξης δεδομένων, οι οποίοι μπορούν να ομαδοποιηθούν σε δύο ευρύτερες κατηγορίες, την προβλεπτική και την περιγραφική. Στη συνέχεια αναδεικνύεται η σημασία της πρόβλεψης του ηλεκτρικού φορτίου και οι βασικοί παράγοντες που μπορούν να επηρεάσουν τη συμπεριφορά του. Τέλος αναφέρονται οι λόγοι που καθιστούν αναγκαία τη βραχυπρόθεσμη πρόβλεψη φορτίου, η οποία αποτελεί και το αντικείμενο της διπλωματικής εργασίας.

ΚΕΦΑΛΑΙΟ 3: ΤΕΧΝΙΚΕΣ ΠΡΟΒΛΕΨΕΩΝ

Το κεφάλαιο αυτό ξεκινά με την ταξινόμηση των τεχνικών προβλέψεων με βάση τη διαδικασία παραγωγής τους σε ποσοτικές, κριτικές και τεχνολογικές μεθόδους. Όσον αφορά στις ποσοτικές μεθόδους, παρουσιάζεται το αιτιοκρατικό μοντέλο και το μοντέλο των χρονοσειρών. Έπειτα περιγράφονται τα βήματα που απαιτούνται για την αποτελεσματική εξαγωγή και αξιολόγηση των προβλέψεων. Δίνονται συνοπτικά οι βασικές μέθοδοι που αντιπροσωπεύουν κάθε κατηγορία πρόβλεψης με βάση το χρονικό της ορίζοντα καθώς και τα διαθέσιμα σφάλματα για την αξιολόγηση των μεθόδων. Ακολουθεί η αναφορά των βασικών παραγόντων για τη σωστή επιλογή μιας μεθόδου πρόβλεψης. Τέλος περιγράφονται οι βασικοί τύποι των ενεργειακών μοντέλων, τα παραδοσιακά μοντέλα και τα μοντέλα που στηρίζονται στη χρήση αισθητήρων, επισημαίνοντας τα πλεονεκτήματα και τα μειονεκτήματα κάθε μοντέλου.

ΚΕΦΑΛΑΙΟ 4: ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ

Το κεφάλαιο αυτό ασχολείται με τη βραχυπρόθεσμη πρόβλεψη που αποτελεί τον άξονα της διπλωματικής. Γίνεται αρχικά μια βιβλιογραφική ανασκόπηση των στατιστικών μεθόδων, που στηρίζονται στη μαθηματική αναπαράσταση της καμπύλης του φορτίου. Βασικότεροι εκπρόσωποι αυτής της κατηγορίας είναι τα μοντέλα εξομάλυνσης, οι μέθοδοι παλινδρόμησης και τα αυτοπαλινδρομικά μοντέλα. Ακολουθεί στη συνέχεια βιβλιογραφική έρευνα σχετικά με τα μοντέλα τεχνητής νοημοσύνης (AI), όπου η μοντελοποίησή τους

μιμείται χαρακτηριστικά της ανθρώπινης συμπεριφοράς με αποτέλεσμα να παρουσιάζουν έστω και στοιχειώδη ευφυΐα. Τα πιο δημοφιλή συστήματα AI είναι η ασαφής λογική, τα τεχνητά νευρωνικά δίκτυα και οι μηχανές διανυσμάτων υποστήριξης.

ΚΕΦΑΛΑΙΟ 5: ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Το κεφάλαιο αυτό εισάγει αρχικά την έννοια της μηχανικής μάθησης που ανήκει στο πεδίο της τεχνητής νοημοσύνης. Η μηχανική μάθηση έγκειται στη δημιουργία μοντέλων ή προτύπων από ένα σύνολο δεδομένων με τη βοήθεια ενός υπολογιστικού συστήματος. Ανάλογα με τη φύση του προβλήματος τα δύο βασικότερα είδη είναι η επιβλεπόμενη μάθηση και η μάθηση χωρίς επίβλεψη. Στη συνέχεια γίνεται μια βιβλιογραφική ανασκόπηση των κυριότερων μοντέλων νευρωνικών δικτύων που αποτελούν δημοφιλή αλγόριθμο μηχανικής μάθησης. Τέλος πραγματοποιείται λεπτομερής περιγραφή των μηχανών διανυσμάτων υποστήριξης (SVM) που συνιστούν την προσέγγιση που επιλέχθηκε για την πιλοτική εφαρμογή. Περιγράφεται η βασική ιδέα των SVM και η αρχή ελαχιστοποίησης του διαρθρωτικού ρίσκου στην οποία στηρίζονται. Αναλύονται τα μοντέλα παλινδρόμησης διανυσμάτων υποστήριξης (SVR) και η χρήση συναρτήσεων πυρήνα για τον υπολογισμό του εσωτερικού γινομένου των διανυσμάτων εισόδου στο χώρο των χαρακτηριστικών.

ΚΕΦΑΛΑΙΟ 6: ΠΕΙΡΑΜΑΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ

Στο κεφάλαιο αυτό παρουσιάζονται αρχικά τα βήματα που απαιτούνται για την υλοποίηση μιας εφαρμογής μηχανικής μάθησης. Επισημαίνονται τα πλεονεκτήματα της Python, που επιλέχθηκε ως γλώσσα υλοποίησης των μοντέλων, ενώ εισάγεται μια νέα βιβλιοθήκη μηχανικής μάθησης, που ονομάζεται Scikit-Learn. Ακολουθεί στη συνέχεια η περιγραφή της πιλοτικής εφαρμογής που αναπτύχθηκε στα πλαίσια της παρούσας μελέτης. Ξεκινά με τη διερεύνηση των διαθέσιμων δεδομένων εισόδου και συνεχίζει με την ανάλυση των συσχετίσεων που εμφανίζονται μεταξύ των δεδομένων κατανάλωσης και των παραμέτρων που επιλέχθηκαν. Τέλος παρουσιάζονται αναλυτικά τα αποτελέσματα της πρόβλεψης για τους εξεταζόμενους τύπους πυρήνων με σκοπό την ανάδειξη της βέλτιστης μεθόδου και της χρονικής βάσης των δεδομένων εισόδου που προσέφεραν την υψηλότερη ακρίβεια πρόβλεψης.

ΚΕΦΑΛΑΙΟ 7: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΟΠΤΙΚΕΣ

Το κεφάλαιο αυτό αποτελεί τον επίλογο της διπλωματικής εργασίας όπου αναδεικνύονται τα κυριότερα συμπεράσματα που προέκυψαν από τα προηγούμενα κεφάλαια της παρούσας μελέτης. Επιπλέον καταγράφονται οι προοπτικές για περαιτέρω έρευνα στο αντικείμενο εργασίας και βελτιώσεις που θα καταστήσουν τη μεθοδολογία ακόμα πιο αποτελεσματική στο μέλλον.

ΚΕΦΑΛΑΙΟ 2

ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΠΡΟΒΛΕΨΗ

ΚΕΦΑΛΑΙΟ 2: ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΠΡΟΒΛΕΨΗ

2.1 Εξόρυξη Δεδομένων

2.1.1 Εισαγωγή στην Εξόρυξη Δεδομένων

Η εξόρυξη δεδομένων (data mining), που αναφέρεται και ως εξόρυξη γνώσης (knowledge discovery) σε βάσεις δεδομένων, είναι μια διαδικασία μη τετριμμένης εξαγωγής υπονοούμενων, προηγουμένως άγνωστων όπως συσχετίσεις, κανόνες γνώσης, πρότυπα (patterns), κανονικότητες, ανωμαλίες κ.α. και εύρεσης πληροφοριών από ένα μεγάλο όγκο δεδομένων ([1], [2], [3]). Οι βασικές τεχνικές data mining περιλαμβάνουν αλγόριθμους για κατηγοριοποίηση, παλινδρόμηση, πρόβλεψη συμβάντων, συσταδοποίηση, ανακάλυψη ακολουθιακών προτύπων και εύρεση συσχετίσεων.

Η εξόρυξη δεδομένων αποτελείται από μια σειρά βημάτων υλοποίησης. Τα πιο σημαντικά βήματα παρουσιάζονται συνοπτικά παρακάτω:

- Βήμα 1: Επιλογή Δεδομένων (Data Selection)

Η αρχική βάση δεδομένων αποθηκεύει τα δεδομένα που σχετίζονται με τους καταναλωτές ηλεκτρικής ενέργειας. Σε αυτό το βήμα γίνεται η ανάκτηση των δεδομένων που σχετίζονται με το αντικείμενο μελέτης.

- Βήμα 2: Καθαρισμός δεδομένων και Προεπεξεργασία δεδομένων (Data Cleaning and Preprocessing)

Στη φάση του καθαρισμού, γίνεται έλεγχος για διακυμάνσεις των δεδομένων και αφαιρούνται τυχόν ακραίες τιμές. Σε περίπτωση όπου υπάρχει ασυνέπεια στις τιμές κατανάλωσης, γίνεται χρήση των παρόμοιων ημερών για την αντιμετώπισή της. Στη φάση προεπεξεργασίας, συμπληρώνονται οι τιμές που λείπουν με τη βοήθεια τεχνικών παλινδρόμησης.

- Βήμα 3: Μείωση Δεδομένων (Data Reduction)

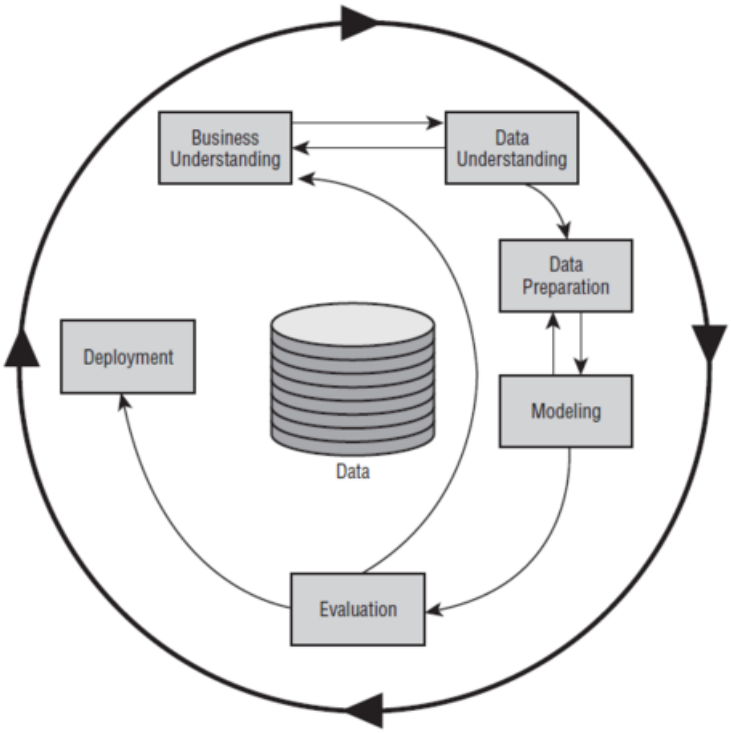
Η μείωση των δεδομένων επιτυγχάνεται με βάση προηγούμενες πληροφορίες που αφορούν στον τρόπο με τον οποίο οι παρόμοιες εποχές του χρόνου ή διάφορα είδη ημερών (εργάσιμες ημέρες, σαββατοκύριακα κ.α.) επηρεάζουν την κατανάλωση ηλεκτρικής ενέργειας.

- Βήμα 4: Εξόρυξη Δεδομένων (Data Mining)

Στο βήμα αυτό τα δεδομένα υφίστανται επεξεργασία με την εφαρμογή τεχνικών εξόρυξης δεδομένων. Στόχος είναι να αναδειχθούν χρήσιμα συμπεράσματα όπως συσχετίσεις, μοντέλα συμπεριφοράς, κανόνες γνώσης κ.α.

- Βήμα 5: Ερμηνεία της εξαγόμενης γνώσης (Interpretation of the discovered knowledge)

Οι πληροφορίες που προέκυψαν από την παραπάνω διαδικασία αξιολογούνται και παρουσιάζονται σε μια εύληπτη δομή, με τη βοήθεια ενός πίνακα ή μιας γραφικής παράστασης ([4], [5], [6]).



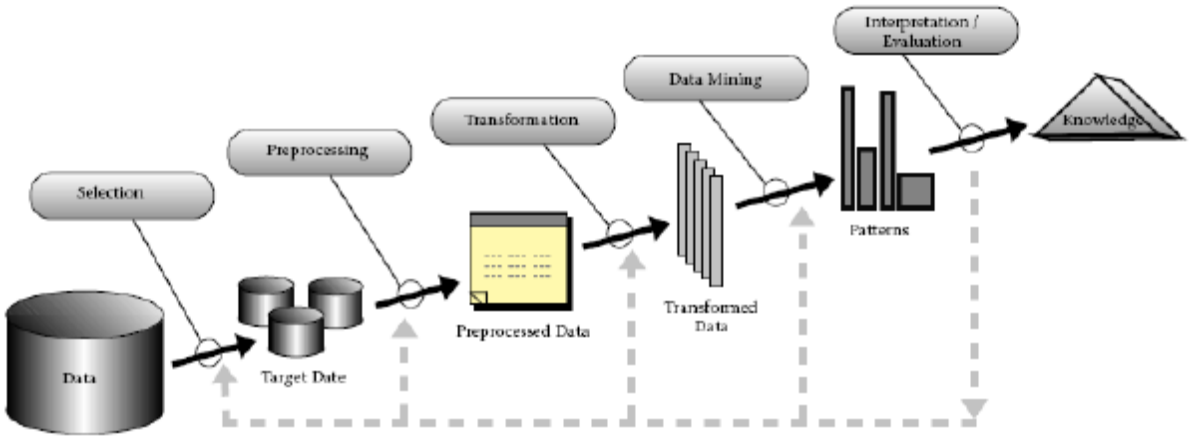
Σχήμα 2.1: Διαδικασία εξόρυξης δεδομένων [7]

2.1.2 Σημασία της Εξόρυξης Δεδομένων

Η εξόρυξη δεδομένων είναι η εξαγωγή "κρυμμένων" πληροφοριών πρόβλεψης μέσα από μεγάλες βάσεις δεδομένων. Με την ανάκτηση αυτής της γνώσης είναι δυνατό να προκύψουν σημαντικές περιοδικότητες και υψηλού επιπέδου πληροφορίες από τα εξεταζόμενα σύνολα δεδομένων. Η εξόρυξη γνώσης αποτελεί μια ισχυρή νέα τεχνολογία με πολλαπλές χρήσεις στον πραγματικό κόσμο σε οικονομικές, ιατρικές, επαγγελματικές εφαρμογές, εφαρμογές πρόβλεψης κ.α. Αξίζει να σημειώσουμε ότι συνδυάζει διάφορα επιστημονικά πεδία, όπως η στατιστική, η τεχνολογία βάσεων δεδομένων, η τεχνητή νοημοσύνη, τα νευρωνικά δίκτυα και η μηχανική μάθηση (machine learning). Το πρώτο σχετικό συνέδριο πραγματοποιήθηκε το 1995 [8].

Η σημασία της εξόρυξης δεδομένων προσδιορίζεται από την εφαρμογή που βρίσκουν τα αποτελέσματα της. Πιο συγκεκριμένα η γνώση που αποκτούμε σχετικά με τα δεδομένα μπορεί να χρησιμοποιηθεί στη διαχείριση της πληροφορίας, στη λήψη αποφάσεων και στην επεξεργασία ερωτημάτων για περαιτέρω απόκτηση πληροφοριών. Ένας άλλος τομέας

αφορά στα ενεργειακά δεδομένα και τη δυνατότητα δημιουργίας ενός προφίλ κατανάλωσης. Σε αυτήν την περίπτωση είναι δυνατή η εξαγωγή πληροφοριών σχετικά με τη συμπεριφορά ενός καταναλωτή, τις συσχετίσεις της κατανάλωσης με τις κλιματολογικές συνθήκες και την αποδοτικότητα του εξεταζόμενου κτηρίου (θέρμανση/ψύξη, κέλυφος κ.α.) αλλά και η αποφυγή λανθάνοντος φορτίου και κρίσιμων συμβάντων. Όλα τα παραπάνω οδηγούν στη βέλτιστη μοντελοποίηση του προφίλ κατανάλωσης και την ακριβέστερη πρόβλεψη μελλοντικών καταναλώσεων. Αξίζει να αναφέρουμε ότι τα τελευταία χρόνια έχουν κάνει την εμφάνισή τους εφαρμογές που προσφέρουν ενεργειακές υπηρεσίες μέσω διαδικτύου και στηρίζονται στην επεξεργασία ενεργειακών δεδομένων πραγματικού χρόνου (real-time data), Intelen [9].



Σχήμα 2.2: Τα βασικότερα βήματα της διαδικασίας Εξόρυξης Γνώσης [10]

2.1.3 Μοντέλα και Μέθοδοι Εξόρυξης Δεδομένων

Τα συστατικά στοιχεία ([10], [11]) που απαρτίζουν τους αλγόριθμους εξόρυξη δεδομένων είναι τα ακόλουθα:

1. Μέθοδος Αναζήτησης

Η αναζήτηση παραμέτρων ψάχνει για ελεύθερες παραμέτρους που βελτιστοποιούν την απόδοση ενός τελικού μοντέλου. Αντίστοιχα η αναζήτηση μοντέλου ψάχνει για το κατάλληλο μοντέλο ή την οικογένεια μοντέλων και για κάθε μια τέτοια δομή που βρίσκει, εφαρμόζει τις κατάλληλες παραμέτρους.

2. Παράσταση Μοντέλου

Είναι η γλώσσα που περιγράφει τις μεθόδους που έχουν ανακαλυφθεί. Για να παραχθεί ένα ικανοποιητικό μοντέλο είναι απαραίτητη η εκτενής αναπαράσταση σε συνδυασμό με παραδείγματα και επαρκή χρόνο εκπαίδευσης.

3. Αξιολόγηση Μοντέλου

Εδώ γίνεται η αξιολόγηση του μοντέλου με τις επιλεγμένες παραμέτρους. Είναι αναγκαίο η μέθοδος να πληροί τις προδιαγραφές της διαδικασίας KDD (Knowledge Discovery in Data) [10].

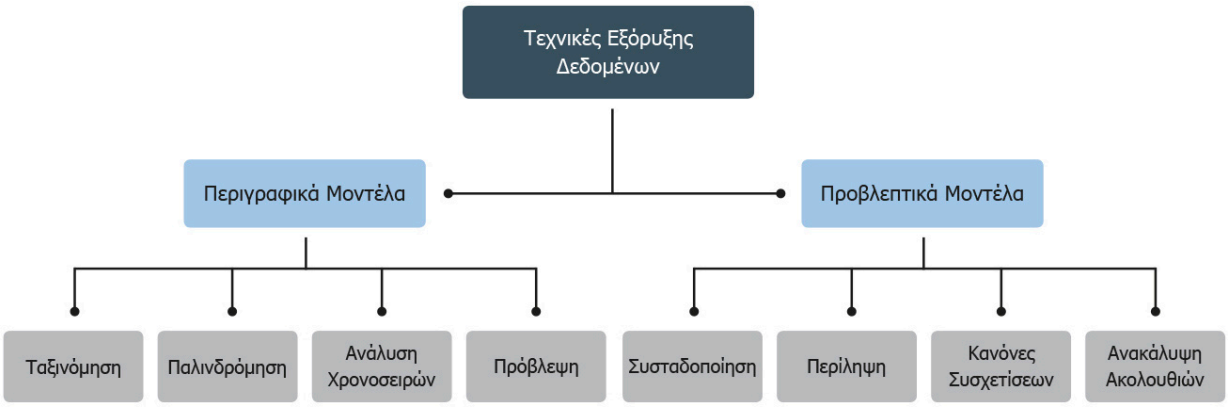
Οι τεχνικές Εξόρυξης Δεδομένων (Data Mining) που αναφέραμε στην παράγραφο 2.1.1 μπορούν να ομαδοποιηθούν σε δύο ευρύτερα μοντέλα. Έτσι έχουμε:

Το προβλεπτικό μοντέλο (predictive model) που έχει στόχο την πρόβλεψη των μελλοντικών δεδομένων, στηριζόμενο σε γνωστά αποτελέσματα που έχει βρει από άλλα δεδομένα. Σε αυτήν την περίπτωση η χρήση ιστορικών δεδομένων κρίνεται απαραίτητη κατά τη μοντελοποίηση της πρόβλεψης. Το προβλεπτικό μοντέλο απαρτίζεται από τις ακόλουθες μεθόδους data mining:

- Ταξινόμηση (Classification)
- Παλινδρόμηση (Regression)
- Ανάλυση Χρονοσειρών (Time Series Analysis)
- Πρόβλεψη (Prediction)

Το περιγραφικό μοντέλο (descriptive model) που έχει στόχο την αναγνώριση προτύπων και συσχετίσεων μεταξύ των δεδομένων. Με αυτόν τον τρόπο επιτυγχάνεται η διερεύνηση των ιδιοτήτων που φέρουν τα υπό εξέταση δεδομένα. Το περιγραφικό μοντέλο απαρτίζεται από τις ακόλουθες μεθόδους data mining:

- Συσταδοποίηση (Clustering)
- Περίληψη (Summarization)
- Κανόνες Συσχετίσεων (Association Rules)
- Ανακάλυψη Ακολουθιών (Sequential Pattern Discovery)



Σχήμα 2.3: Ταξινόμηση των Τεχνικών Εξόρυξης Δεδομένων [12]

2.1.3.1 Ταξινόμηση

Η ταξινόμηση (classification) είναι μια τεχνική data mining που κατατάσσει τα δεδομένα σε προκαθορισμένες κλάσεις. Συνήθως επιλύει προβλήματα όπως είναι η ανάλυση αποχωρήσεων (churn analysis), η διαχείριση κινδύνων και η στόχευση [13]. Είναι γνωστή και ως εποπτευόμενη μάθηση (Supervised Learning), καθώς σε αυτήν την περίπτωση οι κλάσεις είναι καθορισμένες πριν την ανάλυση των δεδομένων ([14], [15]). Γνωστές τεχνικές ταξινόμησης είναι:

- Naive Bayes μοντέλα
- Δένδρα αποφάσεων (Decision Trees)
- Κοντινότεροι γείτονες (K-Nearest Neighbor)
- Νευρωνικά Δίκτυα (Neural Networks)
- Μηχανές Διανυσμάτων Υποστήριξης (SVM)

2.1.3.2 Παλινδρόμηση

Η παλινδρόμηση (regression) εμφανίζει ομοιότητες με την ταξινόμηση με τη διαφορά ότι η μεταβλητή στόχευσης είναι συνεχής αριθμός. Αυτή η στατιστική τεχνική χρησιμοποιείται συχνά για την εξεύρεση συσχέτισης ανάμεσα σε μια εξαρτημένη μεταβλητή και μια ή περισσότερες ανεξάρτητες μεταβλητές. Στόχος της είναι η εκμάθηση μιας συνάρτησης $y = f(x)$ που να απεικονίζει ένα δεδομένο x σε μια μεταβλητή πρόβλεψης y ([12], [16], [17]). Προϋπόθεση για τη μοντελοποίηση των δεδομένων είναι να ταιριάζουν σε κάποιο γνωστό είδος συνάρτησης (γραμμική, πολυωνυμική κ.α.). Γνωστές μέθοδοι παλινδρόμησης είναι:

- Γραμμική παλινδρόμηση (Linear)
- Λογιστική παλινδρόμηση (Logistic)
- Δένδρα παλινδρόμησης (Regression Trees)
- Νευρωνικά Δίκτυα (Neural Networks)

2.1.3.3 Ανάλυση Χρονοσειρών

Στόχος της ανάλυσης χρονοσειρών (time series analysis) είναι η μελέτη της μεταβολής της τιμής ενός μεγέθους με το πέρασμα του χρόνου ([12], [16], [18]). Οι βασικές λειτουργίες αυτής της τεχνικής data mining είναι η εξέταση της δομής μιας χρονοσειράς, η εύρεση ομοιοτήτων μεταξύ χρονοσειρών και τέλος η χρήση διαγραμμάτων χρονοσειρών με στόχο την πρόβλεψη μελλοντικών τιμών.

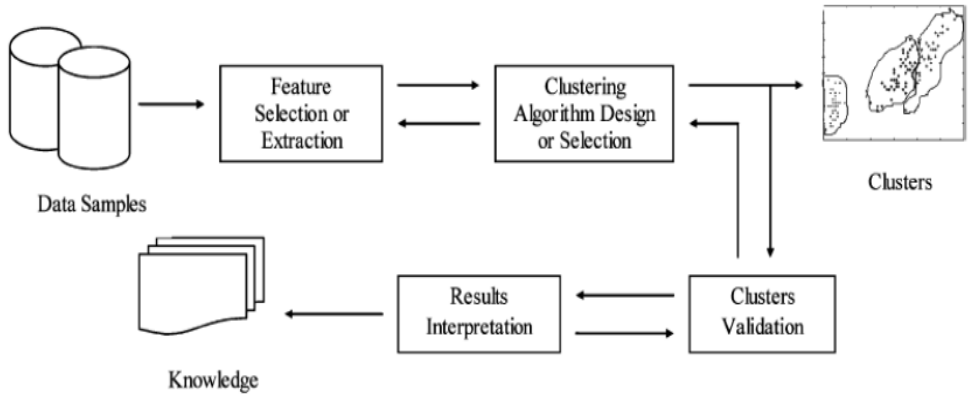
2.1.3.4 Πρόβλεψη

Η τεχνική αυτή, έχοντας διαθέσιμα ιστορικά και σημερινά δεδομένα χρησιμοποιείται σε εφαρμογές που μπορούν να θεωρηθούν πρόβλεψη μελλοντικών καταστάσεων [12].

Οποιεσδήποτε από τις τεχνικές που χρησιμοποιούνται για την ταξινόμηση μπορούν να προσαρμοστούν και στην πρόβλεψη με τη χρήση των παραδειγμάτων εκπαίδευσης όπου η τιμή της μεταβλητής που προβλέπεται είναι ήδη γνωστή, μαζί με τα ιστορικά στοιχεία. Στη συνέχεια χρησιμοποιείται το ιστορικό στοιχείο για να χτίσει ένα μοντέλο που εξηγεί την παρατηρηθείσα συμπεριφορά. Τέλος όταν αυτό το μοντέλο εφαρμόζεται στις τρέχουσες εισαγωγές, το αποτέλεσμα είναι μια πρόβλεψη της μελλοντικής συμπεριφοράς.

2.1.3.5 Συσταδοποίηση

Η συσταδοποίηση (clustering) είναι μια μέθοδος περιγραφικής διαδικασίας που έχει στόχο τον καθορισμό ενός συνόλου από κατηγορίες ή κλάσεις έτσι ώστε να περιγράφονται τα δεδομένα ([12], [16], [19], [20]). Στη συσταδοποίηση που είναι γνωστή και ως μη εποπτευόμενη μάθηση (Unsupervised Learning) οι κλάσεις δεν είναι προσδιορισμένες αλλά καθορίζονται από τα δεδομένα. Οι συστάδες (κλάσεις) μπορεί να είναι αμοιβαία αποκλειόμενες ή επικαλυπτόμενες κλάσεις. Η δημιουργία τους έγκειται στον καθορισμό της ομοιότητας μεταξύ των δεδομένων σε συγκεκριμένα γνωρίσματα.



Σχήμα 2.4: Τα στάδια της διαδικασίας Συσταδοποίησης [21]

2.1.3.6 Περίληψη

Η περίληψη (summarization) περιλαμβάνει μεθόδους με στόχο την εύρεση συμπαγούς περιγραφής υποσυνόλου των δεδομένων. Πιο αναλυτικά, χαρακτηρίζει τα δεδομένα παράγοντας αντιπροσωπευτικές πληροφορίες, γεγονός που συμβάλλει στην ανάδειξη και κατανόηση μερικών γνωρισμάτων τους [10]. Έννοιες της στατιστικής που εξυπηρετούν αυτόν το σκοπό είναι:

- Μέσος
- Διακύμανση
- Τυπική Απόκλιση
- Ιστόγραμμα
- Διάγραμμα Διασποράς

2.1.3.7 Κανόνες Συσχετίσεων

Οι κανόνες συσχετίσεων (association rules) αποτελούν μία από τις σημαντικότερες και νεότερες τεχνικές εξόρυξης γνώσης από μεγάλες βάσεις δεδομένων. Είναι ένα μοντέλο το οποίο με βάση τις συγκεντρωμένες πληροφορίες, παράγει ενδιαφέρουσες συσχετίσεις και πρότυπα. Αυτά μπορούν να βρουν εφαρμογή σε διάφορες εκφάνσεις της καθημερινής ζωής του ανθρώπου, στα τηλεπικοινωνιακά δίκτυα, την ανίχνευση εισβολών στο διαδίκτυο, τη διαχείριση ρίσκου, τη βιοπληροφορική, την ανάλυση ενεργειακών δεδομένων κ.α. Αξίζει να επισημάνουμε ότι η κατανόηση και ανάλυση του καλαθιού αγοράς (market basket analysis) ήταν αυτή που ανέδειξε σε σημαντικό βαθμό τους κανόνες συσχέτισης.

2.1.3.8 Ανακάλυψη Ακολουθιών

Στόχος της ακολουθιακής αυτής ανάλυσης (sequential pattern discovery) είναι ο καθορισμός προτύπων στα δεδομένα. Οι χρονοσειρές και οι ακολουθίες βασίζονται σε χρονικά συνεχόμενα δεδομένα, τα οποία με τη σειρά τους βασίζονται σε εξαρτημένες μεταξύ τους παρατηρήσεις. Το σημαντικό σε αυτήν την περίπτωση είναι ότι η συσχέτιση των προτύπων βασίζεται στο χρόνο.

2.2 Πρόβλεψη Ζήτησης Φορτίου

2.2.1 Πρόβλεψη Ζήτησης Φορτίου

Η πρόβλεψη της ζήτησης φορτίου ηλεκτρικής ενέργειας αποτελεί ένα τομέα της επιστήμης των προβλέψεων που παρουσιάζει ιδιαίτερο ενδιαφέρον τα τελευταία χρόνια. Αν και έκανε την εμφάνιση του πριν μερικές δεκαετίες, μόλις στο πρόσφατο παρελθόν παρατηρήθηκε έντονο ενδιαφέρον από την πλευρά των ερευνητών για την πρόβλεψη ζήτησης φορτίου. Το γεγονός αυτό οφείλεται σε ένα συνονθύλευμα παραγόντων με βασικότερο την ανάγκη προσαρμογής στα νέα δεδομένα των αγορών ηλεκτρικής ενέργειας.

Παλαιότερα όπου συγκεκριμένοι οργανισμοί αναλάμβαναν την παραγωγή ηλεκτρικής ενέργειας σε κάθε χώρα, το ενδιαφέρον για πρόβλεψη της ζήτησης φορτίου ήταν μικρό. Στις αρχές της δεκαετίας του 1990 όμως ξεκίνησε η απελευθέρωση των αγορών ενέργειας σε πολλές χώρες, γεγονός που οδήγησε στην κατάλυση των ισχυρών μονοπωλίων και τη δημιουργία ανταγωνιστικών αγορών. Πολλές νέες εταιρείες παραγωγής ενέργειας παρουσιάστηκαν στις επιμέρους αγορές ηλεκτρικής ενέργειας, με αποτέλεσμα όλες οι εταιρείες πλέον να λειτουργούν υπό κλίμα έντονου ανταγωνισμού.

Πιο συγκεκριμένα στην Ελλάδα το Δεκέμβριο του 1999 η ΔΕΗ (Δημόσια Επιχείρηση Ηλεκτρισμού) έχασε το προνόμιο που κατείχε ως ο αποκλειστικός παραγωγός ηλεκτρικής ενέργειας για τη χώρα, ενώ έκαναν την εμφάνισή τους αρκετές ιδιωτικές εταιρείες στην παροχή μέσης (ΜΤ) και υψηλής τάσης (ΥΤ). Έτσι από τις αρχές του 2001 οι πελάτες που συνδέονται στη μέση και την υψηλή τάση έχουν δικαίωμα επιλογής όσον αφορά στον προμηθευτή τους.

Με βάση τη νέα αυτήν κατάσταση η ανάγκη της εκάστοτε εταιρείας να επιβιώσει ή ακόμα και να ηγηθεί του ανταγωνισμού της αγοράς, ανέδειξε το ενδιαφέρον για σωστή και ακριβή πρόβλεψη του φορτίου ζήτησης ηλεκτρικής ενέργειας αλλά και της τιμής της ενέργειας.

Προκειμένου να αναγνωρίσουμε και τους υπόλοιπους λόγους για τους οποίους η πρόβλεψη των τιμών του ζητούμενου φορτίου είναι αναγκαία, αρκεί να στηριχθούμε στις εφαρμογές της πρόβλεψης του φορτίου σε μια εταιρεία παραγωγής ηλεκτρικής ενέργειας.

Ο οικονομικός προγραμματισμός του συστήματος παραγωγής ενέργειας αποτελεί τη βασικότερη εφαρμογή της πρόβλεψης του ζητούμενου φορτίου. Η βραχυπρόθεσμη πρόβλεψη ζήτησης (για την επόμενη μέρα έως και την επόμενη εβδομάδα) μπορεί να αξιοποιηθεί από την εταιρεία στη διαχείριση των υπαρχόντων συστημάτων ηλεκτρικής ενέργειας (ΣΗΕ), στο σχεδιασμό της βέλτιστης ένταξης μονάδων παραγωγής στο δίκτυο και την επίτευξη προσοδοφόρων συμφωνιών ανταλλαγής ενέργειας με τα γειτονικά δίκτυα. Αξίζει να αναφέρουμε σε αυτό το σημείο ότι το βασικό χαρακτηριστικό της ηλεκτρικής ενέργειας έγκειται στη δυσκολία αποθήκευσής της, επομένως η διατήρηση της παραγωγής και της ζήτησης ηλεκτρικής ενέργειας στο ίδιο επίπεδο είναι βαρύνουσας σημασίας.

Επιπλέον η πρόβλεψη ζήτησης φορτίου λειτουργεί ως αρωγός στην ασφάλεια του συστήματος παραγωγής ενέργειας (ΣΗΕ). Στηριζόμενοι σε αυτήν την πρόβλεψη οι χειριστές του συστήματος μπορούν να αντιμετωπίσουν ενδεχόμενες διακυμάνσεις της ζήτησης και διαταραχές με στόχο την αποφυγή ανεπιθύμητων καταστάσεων, με ενέργειες όπως η ένταξη μονάδων αιχμής, η αγορά ενέργειας από τα γειτονικά δίκτυα και η ρύθμιση της τάσης των ζυγών για το συγκεκριμένο χρονικό ορίζοντα. Οι ενέργειες αυτές μπορούν να οδηγήσουν στην απαιτούμενη εφεδρεία για την εξασφάλιση της κάλυψης της ζήτησης, σε περιπτώσεις σφαλμάτων και ανωμαλιών του συστήματος καθώς και την οικονομική λειτουργία αυτού. Η κατοχή επομένως προβλέψεων ακριβείας όσον αφορά στο ζητούμενο φορτίο συμβάλλει άμεσα στην καλύτερη διαχείριση του συστήματος αλλά και τη μείωση του κόστους παραγωγής και λειτουργίας του ΣΗΕ. Αντίθετα η έλλειψη αξιόπιστων προβλέψεων μπορεί να οδηγήσει την εταιρεία παραγωγής σε δαπανηρές επιλογές δράσης. Πιο συγκεκριμένα στην περίπτωση της υπερεκτίμησης της μελλοντικής ζήτησης φορτίου θα υπάρχει παραγωγή φορτίου περισσότερη από το αναγκαίο με αποτέλεσμα την αύξηση του λειτουργικού κόστους. Ομοίως στην περίπτωση υποτίμησης της ζήτησης θα εμφανιστεί η ανάγκη ενσωμάτωσης μονάδων αιχμής στο δίκτυο με ενδεχόμενο κίνδυνο για την αξιοπιστία του συστήματος.

Τέλος η αύξηση της κατανάλωσης ενέργειας με το πέρασμα του χρόνου, αποτελεί εξίσου σημαντικό λόγο για την πρόβλεψη της ζήτησης φορτίου. Ένα σύστημα παραγωγής, μεταφοράς και διανομής ηλεκτρικής ενέργειας οφείλει να παρέχει στον καταναλωτή ηλεκτρική ενέργεια ανά πάσα στιγμή και σε οποιοδήποτε σημείο ζήτησης. Επίσης θα πρέπει να καλύπτει τις ζητούμενες απαιτήσεις του καταναλωτή όσον αφορά στην ποιότητα (επίπεδα τάσης, συχνότητας, συντελεστή παραμόρφωσης αρμονικών κ.α.), αλλά και την ποσότητα, δηλαδή πρέπει πάντα να υπάρχει επάρκεια ισχύος και ενέργειας. Για να επιτευχθεί αυτό θα πρέπει να ληφθούν υπόψη και πρόσθετες καταναλώσεις από τη μεριά των πελατών που προέκυψαν τα τελευταία χρόνια με την ενσωμάτωση ηλεκτρικών συσκευών που δεν υπήρχαν παλαιότερα (κλιματιστικό, ηλεκτρονικός υπολογιστής κ.α.)

όπως επίσης και τους παράγοντες που επηρεάζουν την κατανάλωση ανάλογα με την περίοδο του έτους που εξετάζουμε.

Απαραίτητη προϋπόθεση για την πλήρη κάλυψη της ζήτησης στο άμεσο αλλά και στο απώτερο μέλλον κρίνεται η ανάπτυξη κατάλληλων μεθοδολογιών πρόβλεψης του φορτίου.

Με κριτήριο το χρονικό ορίζοντα πρόβλεψης φορτίου μπορούμε να διακρίνουμε τις ακόλουθες τέσσερις κατηγορίες:

- Πολύ βραχυπρόθεσμη πρόβλεψη, που αφορά την πρόβλεψη της ζήτησης φορτίου για τα επόμενα 30 λεπτά ως και μία ώρα.
- Βραχυπρόθεσμη πρόβλεψη, που σχετίζεται με την πρόβλεψη του φορτίου για τις επόμενες ώρες ως και μία εβδομάδα.
- Μεσοπρόθεσμη πρόβλεψη, με χρονικό ορίζοντα πρόβλεψης από ένα μήνα ως και τρία έτη.
- Μακροπρόθεσμη πρόβλεψη, που αφορά την πρόβλεψη ζήτησης φορτίου με ορίζοντα από ένα ως και αρκετά έτη (10-20 έτη).

Η πολύ βραχυπρόθεσμη πρόβλεψη φορτίου πιο αναλυτικά έχει στόχο την κάλυψη των αναγκών του συστήματος των γεννητριών, παρέχοντας όσο το δυνατόν καλύτερη παρακολούθηση της σχέσης μεταξύ συχνότητας και φορτίου ώστε να εξασφαλιστεί η ασφάλεια του συστήματος. Έτσι είναι δυνατή η βέλτιστη κατανομή του φορτίου στις μονάδες του συστήματος ηλεκτρικής ενέργειας (ΣΗΕ) κατά τον οικονομικότερο τρόπο.

Η βραχυπρόθεσμη πρόβλεψη φορτίου συμβάλλει ουσιαστικά στην ενεργειακή διαχείριση του συστήματος. Έχει ως στόχο τον έλεγχο και τη ρύθμιση των συστημάτων ηλεκτρικής ενέργειας ενώ συμμετέχει και στη διαμόρφωση της σειράς με την οποία γίνεται η ένταξη των μονάδων παραγωγής των ΣΗΕ. Για την ένταξη των μονάδων λαμβάνονται υπόψη οικονομικά κριτήρια αλλά και τα αντίστοιχα προγράμματα συντήρησης.

Η μεσοπρόθεσμη πρόβλεψη φορτίου έχει ως στόχο τη ρύθμιση των προγραμμάτων συντήρησης των μονάδων παραγωγής, όπως επίσης συμβάλλει στο σχεδιασμό και το χειρισμό των διαθέσιμων ΣΗΕ.

Η μακροπρόθεσμη πρόβλεψη φορτίου είναι σημαντική για τον κατάλληλο σχεδιασμό και κατασκευή του συστήματος ηλεκτρικής ενέργειας αλλά και τη συντήρηση των μονάδων παραγωγής και διανομής. Επιπλέον λειτουργεί βοηθητικά για τις εταιρείες προκειμένου να αποφασίσουν σημαντικές δράσεις όπως επενδύσεις σε αγορά νέων εγκαταστάσεων και μονάδων παραγωγής.

Στην παρούσα μελέτη επικεντρωνόμαστε στη βραχυπρόθεσμη πρόβλεψη ενεργειακής ζήτησης. Στην επόμενη παράγραφο θα επιχειρήσουμε την καταγραφή των βασικών παραγόντων που επηρεάζουν τη ζήτηση του ηλεκτρικού φορτίου. Αυτό το βήμα είναι απαραίτητο πριν ξεκινήσουμε την ανάλυση και αξιολόγηση των ενεργειακών δεδομένων.

2.2.2 Παράγοντες που επηρεάζουν τη ζήτηση φορτίου

Όπως αναφέραμε στην προηγούμενη παράγραφο η πρόβλεψη ζήτησης φορτίου μπορεί να κατηγοριοποιηθεί σύμφωνα με τον ορίζοντα πρόβλεψης σε: πολύ βραχυπρόθεσμη, βραχυπρόθεσμη, μεσοπρόθεσμη και μακροπρόθεσμη πρόβλεψη. Σε κάθε περίπτωση υπάρχει ένα πλήθος παραγόντων που μπορούν να επηρεάσουν τη μορφή της ζήτησης φορτίου που εξετάζουμε. Οι παράγοντες αυτοί είναι κυρίως οικονομικής, εποχικής και μετεωρολογικής φύσης, ενώ σημαντική επίδραση μπορεί να έχουν και παράγοντες που οφείλονται σε τυχαία γεγονότα.

- Οικονομικοί Παράγοντες

Αφορούν την οικονομική κατάσταση, το οικονομικό περιβάλλον και τους ρυθμούς θετικής ή αρνητικής ανάπτυξης μιας περιοχής. Αυτοί οι παράγοντες συνδέονται άμεσα με την αλλαγή του βιοτικού επιπέδου των καταναλωτών και επομένως επιδρούν καθοριστικά στη διαμόρφωση της ζήτησης ηλεκτρικού φορτίου. Η οικονομική ανάπτυξη μιας περιοχής, η μεταβολή στο πλήθος των καταναλωτών, η αυτοματοποίηση των δραστηριοτήτων και η εξέλιξη της τεχνολογίας οδηγούν με τη σειρά τους σε αντίστοιχη μεταβολή της ζήτησης του φορτίου ηλεκτρικής ενέργειας.

- Εποχικοί Παράγοντες

Αντιπροσωπεύουν τις διακυμάνσεις του φορτίου κατά τη διάρκεια μια ημέρας, ενός έτους αλλά και την εναλλαγή των εποχών. Οι διακυμάνσεις αυτές μπορεί να οφείλονται είτε σε εποχιακές μεταβλητές κατά τη διάρκεια ενός ημερολογιακού έτους είτε σε μεταβολές των ανθρώπινων δραστηριοτήτων. Ένα παράδειγμα εποχιακής μεταβλητής είναι η διάρκεια της ημέρας, η οποία αυξάνεται κατά την καλοκαιρινή περίοδο και αντίθετα μειώνεται το χειμώνα, που μπορεί να επηρεάσει το επίπεδο ζήτησης φορτίου. Όσον αφορά στις ανθρώπινες δραστηριότητες, σημαντικό ρόλο παίζει και η γεωγραφική περιοχή, αφού για παράδειγμα κατά τη διάρκεια των καλοκαιρινών διακοπών είναι σύνηθες το φαινόμενο μείωσης της ζήτησης σε αστικά κέντρα και αντίστοιχα αύξηση της ηλεκτρικής κατανάλωσης σε παραθαλάσσιες περιοχές. Στην κατηγορία των εποχικών παραγόντων ανήκουν και οι ειδικές ημέρες (όπως οι αργίες των Χριστουγέννων και του Πάσχα) οι οποίες είναι δυνατό να οδηγήσουν σε απότομες αλλαγές του επιπέδου ζήτησης ενέργειας.

- Μετεωρολογικοί Παράγοντες

Είναι συνδεδεμένοι σε μεγάλο βαθμό με την εναλλαγή των εποχών κυρίως λόγω της μεταβολής που προκύπτει στη θερμοκρασία. Για παράδειγμα κατά τους καλοκαιρινούς μήνες όπου η θερμοκρασία κινείται σε πολύ υψηλά επίπεδα, παρατηρείται αύξηση της χρήσης κλιματιστικών και άλλων μέσων ψύξης, ενώ αντίθετα κατά τη διάρκεια του χειμώνα παρατηρείται αύξηση της χρήσης μέσων θέρμανσης εξαιτίας της χαμηλής θερμοκρασίας. Εκτός της θερμοκρασίας, που αποτελεί τον πιο σημαντικό παράγοντα αυτής της κατηγορίας, αξίζει να αναφέρουμε και άλλους παράγοντες που επιδρούν στη διαμόρφωση της ζήτησης όπως η υγρασία, το επίπεδο της ατμοσφαιρικής πίεσης, η ηλιοφάνεια, η ύπαρξη ισχυρών ανέμων και οι βροχοπτώσεις.

- Τυχαίοι Παράγοντες

Αφορούν ορισμένα τυχαία γεγονότα που ωστόσο μπορούν να μεταβάλλουν τη ζήτηση φορτίου, όπως για παράδειγμα απεργίες μεγάλης διάρκειας, εκλογές κτλ.

Τέλος θα μπορούσαμε να αναφέρουμε και μερικούς αβέβαιους παράγοντες που ενδεχομένως μπορούν να επηρεάζουν τα επίπεδα ζήτησης σε μακροπρόθεσμο ορίζοντα, όπως οι δυνατότητες ανάπτυξης μιας περιοχής, η δημογραφική εξέλιξη του πληθυσμού της και η οικονομική της πορεία.

Η επίδραση των παραγόντων που αναφέραμε παραπάνω είναι διαφορετική ανάλογα με τον ορίζοντα πρόβλεψης. Έτσι για παράδειγμα μπορούμε εύκολα να αντιληφθούμε ότι οι μετεωρολογικοί παράγοντες δεν επηρεάζουν σημαντικά τη μακροπρόθεσμη πρόβλεψη φορτίου. Αντίστοιχα και η βραχυπρόθεσμη πρόβλεψη είναι δύσκολο να επηρεαστεί από την εξέλιξη της οικονομίας μιας περιοχής σε τόσο στενό χρονικό διάστημα.

2.2.3 Βραχυπρόθεσμη πρόβλεψη φορτίου

Η βραχυπρόθεσμη πρόβλεψη φορτίου όπως αναφέραμε στην παράγραφο 2.2.1 σχετίζεται με την πρόβλεψη του φορτίου για τις επόμενες ώρες ως και μια εβδομάδα (24 ως 168 ώρες). Όπως θα προκύψει και στη συνέχεια από τη βιβλιογραφία, σημαντικοί παράγοντες που επηρεάζουν τη βραχυπρόθεσμη πρόβλεψη είναι η θερμοκρασία και άλλα καιρικά στοιχεία (υγρασία, άνεμοι κλπ), η διάρκεια της ημέρας και το φορτίο των προηγούμενων ημερών.

Η σπουδαιότητα της βραχυπρόθεσμης πρόβλεψης έγκειται στη δυνατότητα που προσφέρει για ενεργειακή διαχείριση του συστήματος, ώστε να επιτύχουμε την ομαλή λειτουργία του ΣΗΕ και να εξασφαλίσουμε την ασφάλειά του. Παραδείγματα μελετών που αναδεικνύουν τη σημασία της απότομης μεταβολής φορτίου ή της έλλειψης ακριβούς πρόβλεψης παρουσιάζονται στα ([22], [23]) με τη βοήθεια κατάλληλων μοντέλων.

Είναι ακόμα αναγκαία για τη μείωση του τελικού κόστους κατά το συνδυασμό των συσκευών αποθήκευσης ενέργειας με τα αντίστοιχα ηλεκτρικά τιμολόγια όπως φαίνεται στα ([24], [25]). Δηλαδή στα πλαίσια της απελευθέρωσης των αγορών ενέργειας, η ύπαρξη δεδομένων σχετικά με τις διάφορες κατηγορίες καταναλωτών κρίνεται απαραίτητη για την ομαλή διαχείριση του συστήματος και την εφαρμογή μέτρων εξοικονόμησης. Στο [26] παρουσιάζεται η χρονολογική καμπύλη φορτίου και τα αντίστοιχα μέτρα εξοικονόμησης έπειτα από τεχνοοικονομική μελέτη. Όπως αναφέρθηκε σε προηγούμενη παράγραφο η έλλειψη αξιόπιστων προβλέψεων μπορεί να επηρεάσει την ευστάθεια του συστήματος ηλεκτρικής ενέργειας ([27]) και να οδηγήσει σε αύξηση του κόστους λειτουργίας. Επιπλέον είναι δυνατό να επιδράσει στα αποτελέσματα από τα μοντέλα κατανομής της παραγωγής των γεννητριών και της απαιτούμενης εφεδρείας ([28], [29]). Για το λόγο αυτό έχουν γίνει προσπάθειες ([30], [31]) ανάπτυξης αλγόριθμων που βασίζονται σε μαθηματικά μοντέλα και τη λήψη δεδομένων πραγματικού χρόνου (real-time data) σε επιλεγμένα σημεία του δικτύου με στόχο τη διόρθωση των μετρητικών σφαλμάτων.

Λαμβάνοντας υπόψη όλα τα παραπάνω μπορούμε να αντιληφθούμε τη σημαντικότητα της βραχυπρόθεσμης πρόβλεψης φορτίου για την οποία έχει αναπτυχθεί πληθώρα μεθοδολογιών. Σε μια πρώτη κατηγοριοποίηση θα μπορούσαμε να διακρίνουμε τα κλασσικά μοντέλα (μέθοδοι εξομάλυνσης, αποσύνθεσης, γραμμικών παλινδρομήσεων και αυτοπαλινδρομήσεων κινητού μέσου όρου) και τα νεότερα μοντέλα (νευρωνικά δίκτυα, μηχανές διανυσμάτων υποστήριξης και ασαφής λογική). Περαιτέρω ανάλυση το διαθέσιμων μοντέλων λαμβάνει χώρα στο κεφάλαιο 4.

ΚΕΦΑΛΑΙΟ 3

ΤΕΧΝΙΚΕΣ ΠΡΟΒΛΕΨΕΩΝ

ΚΕΦΑΛΑΙΟ 3: ΤΕΧΝΙΚΕΣ ΠΡΟΒΛΕΨΕΩΝ

3.1 Γενικά για τις Προβλέψεις

Η πρόβλεψη είναι μια εκτίμηση των αβέβαιων μελλοντικών γεγονότων. Οι προβλέψεις χρησιμοποιούνται για τη βελτίωση της λήψης αποφάσεων και σχεδιασμού. Αν και εμπεριέχουν σχεδόν πάντα κάποιο ποσοστό σφάλματος, είναι προτιμότερο να έχουμε την περιορισμένη πληροφόρηση που μας παρέχουν οι προβλέψεις από το να λαμβάνουμε αποφάσεις έχοντας πλήρη άγνοια για το μέλλον.

Η επιθυμία του ανθρώπου για να ερμηνεύσει πράγματα και καταστάσεις που μέχρι σήμερα παραμένουν ανεξήγητα και η γενικότερη ανάγκη του για κατανόηση και μοντελοποίηση του σύμπαντος και της λειτουργίας του είχαν ως επακόλουθο τα δημιουργία επιστημών που ασχολούνται με τα μελλοντικά γεγονότα. Ο επιστημονικός κλάδος που ασχολείται με την παραγωγή προβλέψεων είναι αυτός των Τεχνικών Προβλέψεων. Αν και ο συγκεκριμένος κλάδος είναι σχετικά νέος στο χώρο της εκπαίδευσης, έχει ήδη αποδείξει τη χρησιμότητά του. Γνωρίζοντας εκ των προτέρων την πιθανή κατάληξη μιας δράσης ή ενός γεγονότος μας δίνεται η δυνατότητα να διαχειριστούμε καλύτερα το αποτέλεσμα ή ακόμα και να αποφύγουμε τυχόν κινδύνους.

Όσον αφορά στον τομέα της ενέργειας, η λήψη αποφάσεων θα πρέπει να βασίζεται σε ακριβείς προβλέψεις της ζήτησης φορτίου. Ως εκ τούτου, οι προβλέψεις φορτίου είναι σημαντικά εργαλεία στον τομέα της ενέργειας. Για τη λειτουργία του πολύπλοκου συστήματος ηλεκτρικής ενέργειας είναι απαραίτητες οι προβλέψεις για διαφορετικούς χρονικούς ορίζοντες αλλά και εύρη σφαλμάτων: “Η απόκριση του συστήματος παρακολουθεί στενά την απαίτηση φορτίου” [32]. Κατά τη λήψη αποφάσεων υπάρχει ένα πλήθος προβλημάτων που πρέπει να αντιμετωπιστούν και αφορούν σε διαφορετικές χρονικές κλίμακες αλλά και ιεραρχίες του συστήματος ηλεκτρικής ενέργειας: Αυτά τα προβλήματα περιλαμβάνουν για παράδειγμα το βέλτιστο και ασφαλή προσδιορισμό δέσμευσης μιας μονάδας και κατανομής της ενέργειας. Ωστόσο υπάρχει ανάγκη λήψης αποφάσεων όχι μόνο για τη βραχυπρόθεσμη λειτουργία του συστήματος ενέργειας αλλά και για επενδύσεις σε νέες εγκαταστάσεις, δράση που βασίζεται στην πρόβλεψη των μελλοντικών ενεργειακών αναγκών. Και στις δύο περιπτώσεις υπάρχει ανάγκη για αξιόπιστες προβλέψεις. Η απελευθέρωση των αγορών ενέργειας έχει αυξήσει με τη σειρά της την ανάγκη για πιο ακριβείς προβλέψεις [33]. Προκειμένου να συμμετάσχει ένας παίκτης στην αγορά, πρέπει να γνωρίζει με ακρίβεια την ενέργεια που χρειάζεται για ένα χρονικό διάστημα. Είναι πολύ σημαντικό να αποφύγει τυχόν υπερεκτίμηση της απαιτούμενης ενέργειας, γεγονός που θα οδηγήσει σε σπατάλη σημαντικών πόρων. Αντίστοιχα στην περίπτωση υποεκτίμησης της ηλεκτρικής ζήτησης, θα αυξηθεί σημαντικά το λειτουργικό κόστος για τον προμηθευτή στην προσπάθεια κάλυψης της πρόσθετης ζήτησης ενώ παράλληλα τίθεται σε κίνδυνο η αξιοπιστία του συστήματος. Τέλος η ζήτηση ηλεκτρικής ενέργειας συνιστά έναν από τους σημαντικότερους παράγοντες τιμολόγησης. Αντιλαμβανόμαστε επομένως ότι η πρόβλεψη φορτίου βρίσκεται στο επίκεντρο όλων των αποφάσεων που λαμβάνονται στις αγορές ενέργειας.

Μέχρι και σήμερα οι επιστήμονες που ασχολούνται με τις Τεχνικές Προβλέψεων έχουν εισάγει πλήθος προσεγγίσεων. Ορισμένες από αυτές διαθέτουν μόνο θεωρητικό υπόβαθρο ενώ άλλες προϋποθέτουν τη συμβολή της τεχνολογίας σε μεγάλο βαθμό.

Οι πιο σημαντικές παράμετροι βάσει των οποίων μπορούμε να κατηγοριοποιήσουμε τις προβλέψεις είναι οι ακόλουθες:

- ✓ Διαδικασία παραγωγής των προβλέψεων
- ✓ Χρονικός ορίζοντας της πρόβλεψης
- ✓ Χρήση ιστορικών δεδομένων και προτύπων συμπεριφοράς
- ✓ Κόστος παραγωγής προβλέψεων
- ✓ Ευκολία χρήσης της επιλεγμένης μεθόδου πρόβλεψης
- ✓ Αξιοπιστία της επιλεγμένης μεθόδου πρόβλεψης

3.1.1 Κατηγορίες μεθόδων πρόβλεψης

Οι μέθοδοι προβλέψεων διακρίνονται με βάση τη διαδικασία παραγωγής τους σε τρεις μεγάλες κατηγορίες:

- Ποσοτικές Μέθοδοι. Αυτές διαχωρίζονται περαιτέρω σε :
  - ✓ Μοντέλο Χρονοσειρών
  - ✓ Αιτιοκρατικό Μοντέλο
- Κριτικές Μέθοδοι. Αυτές διαχωρίζονται περαιτέρω σε :
  - ✓ Ατομικές μέθοδοι
  - ✓ Μέθοδοι επιτροπής
- Τεχνολογικές Μέθοδοι. Αυτές διαχωρίζονται περαιτέρω σε :
  - ✓ Διερευνητικές μέθοδοι
  - ✓ Κανονιστικές μέθοδοι

Στη συνέχεια θα παρουσιάσουμε συνοπτικά τα βασικότερα πλεονεκτήματα και μειονεκτήματα κάθε κατηγορίας πρόβλεψης:

- Ποσοτικές Μέθοδοι (Quantitative)

Η στατιστική πρόβλεψη προκύπτει από την εφαρμογή ποσοτικών μοντέλων πάνω σε μια σειρά δεδομένων. Με αυτόν τον τρόπο είναι δυνατό να παραχθεί με συστηματικό τρόπο η πρόβλεψη της πορείας της ακολουθίας. Οι δύο βασικές κατηγορίες μοντέλων που αφορούν

στις στατιστικές μεθόδους είναι το μοντέλο χρονοσειρών και το αιτιοκρατικό (επεξηγηματικό) μοντέλο. Κάθε μοντέλο στηρίζεται σε διαφορετικές υποθέσεις και επομένως είναι κατάλληλο σε διαφορετικές περιπτώσεις όπως θα περιγράφεται παρακάτω. Σημαντικά πλεονεκτήματα των στατιστικών μεθόδων είναι η ευκολία που εμφανίζουν στη χρήση και η παραγωγή αποτελεσμάτων με καλή ακρίβεια μέσα σε σύντομο χρονικό διάστημα και με τη χρήση ελάχιστων υπολογιστικών πόρων. Ωστόσο το σημαντικότερο μειονέκτημά τους είναι η αδυναμία πρόβλεψης ειδικών γεγονότων (special events) που μπορεί να μεταβάλλουν σε σημαντικό βαθμό τη συμπεριφορά της εξεταζόμενης χρονοσειράς. Σε περίπτωση επομένως όπου τα μελλοντικά δεδομένα δεν ακολουθούν το ίδιο πρότυπο συμπεριφοράς, η αδυναμία των μεθόδων να λάβουν υπόψη τους παράγοντες που τα επηρεάζουν είναι πιθανόν να οδηγήσει σε αστοχίες.

- Κριτικές Μέθοδοι (Judgmental)

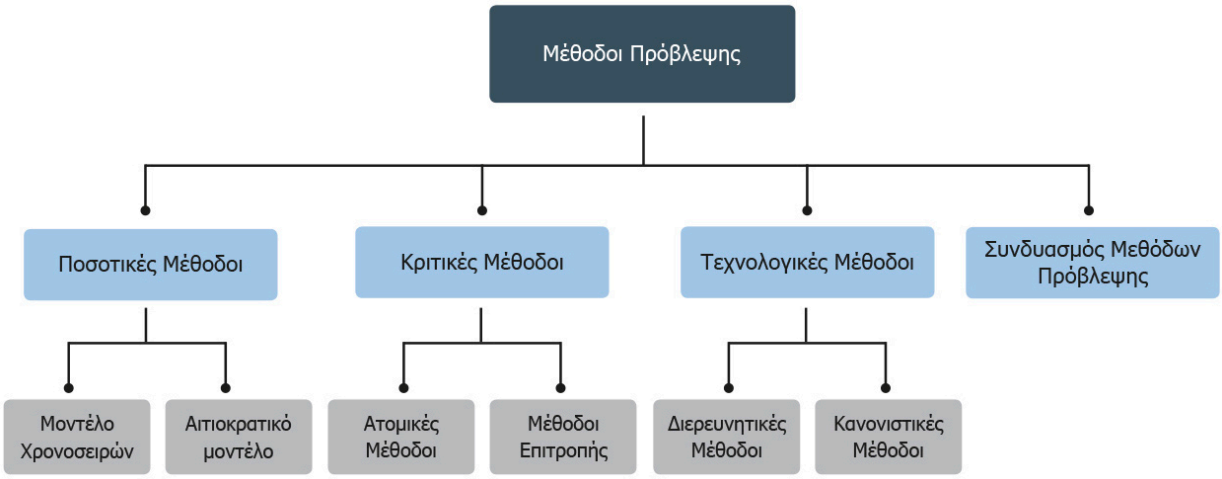
Οι κριτικές μέθοδοι πρόβλεψης μπορεί να βασίζονται στη διαίσθηση, την κρίση ή τη συσσωρευμένη γνώση ενός ατόμου ή μιας ομάδας ατόμων. Πρόκειται για μεθόδους που δε στηρίζονται στην ύπαρξη μεγάλου όγκου δεδομένων για τη χρονοσειρά και χρησιμοποιούνται πολύ συχνά σε επιχειρήσεις και οργανισμούς. Εδώ η πρόβλεψη μπορεί να βασίζεται είτε σε ένα άτομο (ατομικές μέθοδοι) είτε στο συνδυασμό απόψεων μιας ομάδας μελών (μέθοδοι επιτροπής). Συγκριτικά με τις στατιστικές μεθόδους έχουν το σημαντικό πλεονέκτημα ότι μπορούν να λάβουν υπόψη ειδικά γεγονότα και ενέργειες προκειμένου να αντισταθμίσουν ανεπάρκειες και να αντιμετωπίσουν ανομοιογένειες στα διαθέσιμα δεδομένα. Από την άλλη πλευρά αξιοσημείωτο μειονέκτημα τους είναι ότι χαρακτηρίζονται από προκατάληψη, στοιχείο που εξηγείται από την έμφυτη τάση των ανθρώπων να είναι αισιόδοξοι ή απαισιόδοξοι. Οι πιο γνωστές μέθοδοι κριτικής πρόβλεψης είναι η απλή κρίση, η μέθοδος Delphi και οι δομημένες αναλογίες.

- Τεχνολογικές Μέθοδοι (Technological)

Οι τεχνολογικές μέθοδοι πρόβλεψης αφορούν κυρίως μακροπρόθεσμα πλάνα τεχνολογικής, οικονομικής, κοινωνικής και πολιτικής φύσης και διακρίνονται σε διερευνητικές και κανονιστικές. Οι διερευνητικές (exploratory) μέθοδοι πρόβλεψης ξεκινώντας από το παρελθόν και το παρόν στοχεύουν στη διερεύνηση όλων των πιθανών περιπτώσεων στο μέλλον. Οι κανονιστικές (normative) μέθοδοι αφού καθοριστούν οι μελλοντικοί στόχοι, εξετάζουν τη δυνατότητα επίτευξης αυτών, σύμφωνα με τους περιορισμούς και τους διαθέσιμους πόρους.

- Συνδυασμός Μεθόδων Πρόβλεψης

Πολλές φορές η επιλογή μιας μεθόδου πρόβλεψης ως πιο κατάλληλη είναι δύσκολη, καθώς δεν επιτυγχάνεται πάντα η επιθυμητή ακρίβεια. Σε αυτήν την περίπτωση επιχειρείται συχνά ο συνδυασμός των προβλέψεων που προκύπτουν από τις μεθόδους που αναφέραμε (ποσοτικές, κριτικές, τεχνολογικές). Ακολουθώντας αυτήν τη διαδικασία είναι δυνατό να μειώσουμε σημαντικά το διάστημα διακύμανσης των σφαλμάτων και να επιτύχουμε ακρίβεια στα αποτελέσματά μας.



Σχήμα 3.1: Κατηγοριοποίηση μεθόδων πρόβλεψης με βάση τη διαδικασία παραγωγής

3.1.2 Ορίζοντας πρόβλεψης

Για τη λειτουργία και τον προγραμματισμό μιας εταιρείας κοινής ωφελείας, η ύπαρξη μοντέλων ακριβείας για την πρόβλεψη ηλεκτρικού φορτίου ηλεκτρικής ενέργειας είναι ουσιαστικής σημασίας [34]. Η πρόβλεψη φορτίου βοηθά μια ηλεκτρική εταιρεία να λάβει σημαντικές αποφάσεις, συμπεριλαμβανομένων των αποφάσεων για την αγορά και την παραγωγή ηλεκτρικού ρεύματος μεταγωγής φορτίου και τις υποδομές ανάπτυξης. Επίσης η πρόβλεψη φορτίου είναι εξαιρετικά σημαντική για τους προμηθευτές ενέργειας, τα χρηματοπιστωτικά ιδρύματα και άλλους συμμετέχοντες στην παραγωγή της ηλεκτρικής ενέργειας, τη μεταφορά, τη διανομή, και τις αγορές. Η πρόβλεψη φορτίου ήταν πάντα σημαντική για τον προγραμματισμό και τις επιχειρησιακές αποφάσεις που διεξάγονται από εταιρείες κοινής ωφέλειας. Ωστόσο με την απελευθέρωση των ενεργειακών βιομηχανιών, η πρόβλεψη φορτίου απέκτησε ακόμα μεγαλύτερη σημασία. Με τις διακυμάνσεις στην προσφορά και τη ζήτησης, τις αλλαγές των καιρικών συνθηκών και την αύξηση των τιμών ενέργειας κατά τη διάρκεια καταστάσεων αιχμής, η πρόβλεψη φορτίου απέκτησε ζωτική σημασία για τις υπηρεσίες κοινής ωφέλειας.

Όπως διαπιστώθηκε παραπάνω, οι προβλέψεις γίνονται για διάφορους σκοπούς. Για την παρακολούθηση της καθημερινής λειτουργίας του συστήματος ηλεκτρικής ενέργειας ([32]) είναι απαραίτητη η πρόβλεψη φορτίου για μια ημέρα μπροστά, ενώ η απόφαση για το αν θα προβεί η επιχείρηση σε μεγάλες διαρθρωτικές επενδύσεις χρειάζεται ένα πολύ μεγαλύτερο χρονικό ορίζοντα πρόβλεψης [33]. Μπορούμε να διακρίνουμε επομένως τις προβλέψεις σε:

- Πολύ βραχυπρόθεσμη πρόβλεψη (Very Short-term forecasting)

Αφορά στην πρόβλεψη της ζήτησης φορτίου για τα επόμενα 30 λεπτά ως και μια ώρα. Χρησιμοποιείται για την κατανομή φορτίου στις μονάδες του ΣΗΕ με τον αποδοτικότερο

τρόπο. Προσφέρει άμεση παρακολούθηση της συχνότητας του φορτίου, εξασφαλίζοντας έτσι την ασφάλεια του συστήματος. Τέλος συμμετέχει στη ρύθμιση της τιμής της ηλεκτρικής ενέργειας [35].

- Βραχυπρόθεσμη πρόβλεψη (Short-term forecasting)

Αφορά στην πρόβλεψη της ζήτησης φορτίου για τις επόμενες ώρες έως και μια εβδομάδα. Η βραχυπρόθεσμη πρόβλεψη μπορεί να βοηθήσει στην εκτίμηση των ροών φορτίου και στη λήψη αποφάσεων προκειμένου να αποφευχθούν καταστάσεις υπερφόρτωσης. Η έγκαιρη λήψη τέτοιων αποφάσεων μπορεί να οδηγήσει όχι μόνο στη βελτίωση της αξιοπιστίας του δικτύου αλλά και τη μείωση περιστατικών όπως βλάβες και διακοπές ρεύματος. Τέλος είναι σημαντική για την αξιολόγηση συμβάσεων και διάφορων χρηματοοικονομικών προϊόντων κατά την τιμολόγηση της ενέργειας που προσφέρει η αγορά.

- Μεσοπρόθεσμη Πρόβλεψη (Medium-term forecasting)

Αφορά στην πρόβλεψη της ζήτησης φορτίου για τον επόμενο μήνα έως και τρία έτη. Παρέχει στις εταιρείες εκτιμήσεις για τη ζήτηση φορτίου για μεγαλύτερο χρονικό διάστημα, γεγονός που βοηθά κατά τη διαπραγμάτευση των συμβάσεων με άλλες εταιρείες [36]. Επιπλέον συμμετέχει στο σχεδιασμό, το χειρισμό των διαθέσιμων ΣΗΕ και τη ρύθμιση των προγραμμάτων συντήρησης για τις μονάδες παραγωγής.

- Μακροπρόθεσμη Πρόβλεψη (Long-term forecasting)

Αφορά στην πρόβλεψη της ζήτησης φορτίου για τον επόμενο χρόνο ως και αρκετά έτη (10-20 έτη). Στην απελευθερωμένη πλέον οικονομία η λήψη αποφάσεων για τις κεφαλαιουχικές δαπάνες με βάση τη μακροπρόθεσμη πρόβλεψη είναι ακόμα πιο σημαντική συγκριτικά με πριν, όπου η αύξηση των επιτοκίων μπορούσε να δικαιολογηθεί με τις κεφαλαιουχικές δαπάνες έργων. Ακόμα είναι πολύ σημαντική για τον κατάλληλο σχεδιασμό του συστήματος ηλεκτρικής ενέργειας και τη συντήρηση των μονάδων διανομής. Τέλος αποτελεί τη βάση για το σχεδιασμό επενδύσεων και τη μακροχρόνια ανάπτυξη.

3.1.3 Διαδικασία Πρόβλεψης

Όπως αναφέρθηκε, έχουν διατυπωθεί διάφορες μέθοδοι πρόβλεψης με βάση τον τρόπο παραγωγής τους αλλά και το χρονικό τους ορίζοντα. Γίνεται αντιληπτό επομένως ότι η επιλογή της κατάλληλης μεθόδου για κάθε περίπτωση δεν είναι μια απλή διαδικασία αλλά χρειάζεται σχεδιασμό και μελέτη. Ωστόσο για την εξασφάλιση μιας επιτυχημένης πρόβλεψης χρειάζεται αρχικά να καθοριστεί το υπό εξέταση πρόβλημα, έπειτα να ελεγχθούν τα διαθέσιμα μοντέλα και τέλος να αποφασιστεί ποιο είναι το καταλληλότερο με βάση ορισμένα κριτήρια. Πιο αναλυτικά τα πέντε βήματα που απαιτούνται για τη σωστή εξαγωγή και αξιολόγηση των προβλέψεων είναι τα ακόλουθα [37]:

1. Ορισμός του προβλήματος

Αποτελεί συχνά το δυσκολότερο κομμάτι της διαδικασίας πρόβλεψης. Εδώ γίνεται η προσπάθεια να προσδιοριστούν τα εξεταζόμενα μεγέθη και οι αποδέκτες της πρόβλεψης λαμβάνοντας υπόψη τη χρήση και τη λειτουργία των προβλέψεων.

2. Συλλογή πληροφοριών και στοιχείων

Σε αυτό το βήμα στόχος είναι να συγκεντρωθούν όλα τα στατιστικά δεδομένα που απαιτούνται αλλά και όποια κριτική και εμπειρική πληροφορία είναι διαθέσιμη. Η διαδικασία συλλογής και συντήρησης των δεδομένων αποδεικνύεται συχνά χρονοβόρα.

3. Προεπεξεργασία Δεδομένων

Αρχικά επιχειρείται γραφική απεικόνιση των δεδομένων και υπολογισμός διάφορων στατιστικών μεγεθών. Στόχος είναι να αναγνωριστούν οι βασικές συνιστώσες τις χρονοσειράς, όπως τάση, κύκλος, εποχιακότητα και ασυνέχειες, για τις οποίες θα γίνει λόγος σε επόμενη παράγραφο. Με αυτόν τον τρόπο μπορεί να επιτευχθεί η εξομάλυνση των χρονοσειρών (μηδενικές, ακραίες τιμές κ.α.) που είναι απαραίτητη για το επόμενο βήμα.

4. Επιλογή της κατάλληλης μεθόδου πρόβλεψης

Με βάση την εξομαλυμένη χρονοσειρά επιχειρείται η αξιολόγηση των διαθέσιμων των μοντέλων πρόβλεψης. Στόχος είναι η εύρεση εκείνου του μοντέλου, που με τις κατάλληλες παραμέτρους θα οδηγήσει στα πιο ακριβή αποτελέσματα.

5. Χρήση και εφαρμογή της μεθόδου πρόβλεψης

Σε αυτό το βήμα χρησιμοποιείται το επιλεγμένο μοντέλο μαζί με τις κατάλληλες παραμέτρους για την παραγωγή των ζητούμενων προβλέψεων. Η αξιολόγηση του μοντέλου προκύπτει σύμφωνα με ορισμένους στατιστικούς δείκτες για τον υπολογισμό σφαλμάτων σε σύγκριση με τα πραγματικά δεδομένα της υπό εξέταση περιόδου.

3.2 Χρονοσειρές

3.2.1 Γενικά για τις Χρονοσειρές

Με τον όρο χρονοσειρές εννοούμε μια σειρά διαδοχικών παρατηρήσεων οι οποίες περιγράφουν την εξέλιξη ενός μεγέθους στο χρόνο. Οι παρατηρήσεις αυτές λαμβάνονται σε ορισμένες χρονικές στιγμές ή περιόδους που ισαπέχουν μεταξύ τους. Συμβολίζοντας με Xi τις n χρονικές στιγμές και με Yi τις αντίστοιχες παρατηρήσεις, προκύπτουν ζεύγη της μορφής $K(Xi, Yi)$. Αν ενώσουμε αυτά τα σημεία $(K1, \dots, Kn)$ μπορούμε να απεικονίσουμε γραφικά τη χρονοσειρά και να παρακολουθήσουμε τη χρονική της εξέλιξη.

Οι διαδοχικές παρατηρήσεις που αναφέρθηκαν παραπάνω σχετίζονται μεταξύ τους και μπορούν να χρησιμοποιηθούν στην προσπάθεια για πρόβλεψη των μελλοντικών τιμών της

χρονοσειράς. Με βάση τη διαθέσιμη γνώση σχετικά με τους παράγοντες που επηρεάζουν την εξέλιξη του εξεταζόμενου μεγέθους επιχειρείται η διάκριση δύο μοντέλων:

- Ντετερμινιστικά μοντέλα: Περιγράφουν την εξέλιξη ενός μεγέθους με την προϋπόθεση ότι είναι γνωστοί όλοι οι παράγοντες που το επηρεάζουν.

- Στοχαστικά μοντέλα: Λαμβάνουν υπόψη τους την επίδραση του τυχαίου παράγοντα στην εξέλιξη ενός μεγέθους. Όπως συμβαίνει και στις πραγματικές χρονοσειρές, η γνώση των ιστορικών δεδομένων δεν είναι αρκετή για να προσδιοριστούν οι μελλοντικές τιμές καθώς η εξέλιξη των χρονοσειρών επηρεάζεται σημαντικά και από τυχαίους παράγοντες.

3.2.2 Ποιοτικά Χαρακτηριστικά Χρονοσειρών

Για τη συστηματική μελέτη μιας χρονοσειράς είναι απαραίτητη η επισκόπηση της γραφικής της παράστασης στο πεδίο του χρόνου. Τα βασικά ποιοτικά χαρακτηριστικά μιας χρονοσειράς είναι τα εξής: Τάση, Κυκλικότητα, Εποχιακότητα, Ασυνέχειες και Τυχαιότητα. Στη συνέχεια δίνεται μια σύντομη περιγραφή αυτών των χαρακτηριστικών.

1. Τάση

Η τάση αντιπροσωπεύει τη γενική εικόνα της χρονοσειράς και μπορεί να ορισθεί ως η μακροπρόθεσμη μεταβολή του μέσου όρου των τιμών αυτής. Μπορεί να είναι ανοδική, σταθερή ή πτωτική και συνήθως εκτιμάται με μια ευθεία γραμμή ή εκθετική καμπύλη. Προκειμένου να εξαχθούν ασφαλή συμπεράσματα σχετικά με το αν μια χρονοσειρά παρουσιάζει ή όχι τάση, είναι απαραίτητη η διάθεση ενός ικανού αριθμού παρατηρήσεων σύμφωνα με το οποίο πρόκειται να εκτιμηθεί το κατάλληλο μήκος περιόδου όπου θα γίνει η αναζήτηση της τάσης. Στόχος είναι να αποφευχθεί η σύγχυση μεταξύ τάσης και κυκλικότητας που μπορεί να παρουσιάζει μια χρονοσειρά.

2. Κυκλικότητα

Η κυκλικότητα αντιπροσωπεύει μια “κυματοειδή” μεταβολή που εμφανίζεται κατά περιόδους. Η μεταβολή αυτή οφείλεται σε εξωγενείς παράγοντες και κυρίως στις οικονομικές συνθήκες. Σχετικά με τις περιόδους αξίζει να αναφερθεί ότι δεν είναι σταθερές ενώ το μήκος τους ξεπερνά πάντα το ένα έτος. Στις γραφικές παραστάσεις των χρονοσειρών ο κυκλικός παράγοντας εντοπίζεται από τις ανόδους και τις πτώσεις ανάμεσα στην υψηλότερη και τη χαμηλότερη στάθμη. Τέλος η κυκλικότητα συναντάται κυρίως σε χρονοσειρές που αφορούν δείκτες βιομηχανίας, τιμές μετοχών και άλλα μακροοικονομικά μεγέθη όπως το Ακαθάριστο Εγχώριο Προϊόν (ΑΕΠ).

3. Εποχιακότητα

Η εποχιακότητα αντιπροσωπεύει μια περιοδική διακύμανση της χρονοσειράς, με σταθερό μήκος και χρονική διάρκεια μικρότερη από ένα έτος. Η διακύμανση αυτή είναι εύκολα αναγνωρίσιμη γιατί επαναλαμβάνεται κατά τον ίδιο τρόπο με το χρόνο. Συναντάται κυρίως σε χρονοσειρές μεγεθών που επηρεάζονται σημαντικά από την εποχή. Η μέτρηση και απομόνωση της εποχιακότητας γίνεται εύκολα με διάφορες μεθόδους. Η διαφορά της από την κυκλικότητα έγκειται στο γεγονός ότι έχει σταθερή διάρκεια μικρότερη του ενός έτους, σε αντίθεση με την κυκλικότητα της οποίας η διάρκεια μπορεί να διαφέρει από κύκλο σε κύκλο.

4. Ασυνέχειες

Οι ασυνέχειες αντιπροσωπεύουν εκείνες τις τιμές μιας χρονοσειράς που δεν ακολουθούν το πρότυπο συμπεριφοράς της και με βάση τα ιστορικά δεδομένα δε θα ήταν δυνατό να έχουν προβλεφθεί. Με κριτήριο τη χρονική διάρκεια των ακραίων αυτών αλλαγών μπορούν να διακριθούν δύο περιπτώσεις:

- Ασυνήθιστες τιμές (outliers, special events):
  Έχουν μικρή διάρκεια και παροδικό χαρακτήρα. Η επίδρασή τους στη χρονοσειρά είναι περιορισμένη. Χρειάζεται ιδιαίτερη προσοχή στην ερμηνεία τους, ενώ απαιτείται καλή θεωρητική γνώση, κριτική ικανότητα και κοινή λογική. Χαρακτηριστικό παράδειγμα απρόβλεπτου γεγονότος είναι μια απεργία στον κλάδο της βιομηχανίας που θα μπορούσε να αλλάξει άρδην τα επίπεδα παραγωγής για ένα μικρό χρονικό διάστημα.

- Αλλαγές επιπέδου (level-shifts):
  Έχουν μεγάλη διάρκεια και μόνιμο χαρακτήρα. Η επίδρασή τους είναι σημαντική καθώς εμφανίζονται ως απότομες αλλαγές στο μέσο επίπεδο των τιμών της χρονοσειράς. Για παράδειγμα σε μια επιχείρηση η εμφάνιση ανταγωνιστικού προϊόντος στην αγορά θα μπορούσε να προκαλέσει πολύ σημαντική αλλαγή στο επίπεδο των πωλήσεών της.

5. Τυχαιότητα

Η τυχαιότητα αντιπροσωπεύει τις μη κανονικές διακυμάνσεις που μένουν όταν απομονώσουμε από τη χρονοσειρά τα τρία πρώτα χαρακτηριστικά (τάση, τυχαιότητα, κυκλικότητα). Αποτελεί το στοιχείο του σφάλματος και προκύπτει είτε από κάποια τυχαία μεταβλητή είτε από κάποια ασυνέχεια (outlier) λόγω ειδικού γεγονότος.

3.3 Μοντέλα Πρόβλεψης

Τα μοντέλα πρόβλεψης επεξηγούν τη διαδικασία που ακολουθείται κατά την παραγωγή προβλέψεων. Υπάρχουν διάφορες μέθοδοι προβλέψεων (ποσοτικές, κριτικές, τεχνολογικές) οι οποίες περιγράφηκαν στην παράγραφο 3.1.1 και σε κάθε μια αντιστοιχεί διαφορετικό μοντέλο. Μεγαλύτερο ενδιαφέρον δίνεται στις ποσοτικές μεθόδους πρόβλεψης οι οποίες στηρίζονται στην υπόθεση ότι τα ιστορικά δεδομένα διατηρούν ένα σταθερό πρότυπο συμπεριφοράς και προσπαθούν να ποσοτικοποιήσουν τις διαθέσιμες πληροφορίες. Οι ποσοτικές μέθοδοι μπορούν να ταξινομηθούν σε δύο κατηγορίες με βάση το μοντέλο που χρησιμοποιείται, έτσι προκύπτει το μοντέλο των Χρονοσειρών (Time series model) και το Αιτιοκρατικό ή Επεξηγηματικό (Causal relationship or explanatory). Για την κατανόηση των βασικών υποθέσεων πάνω στις οποίες στηρίζεται κάθε ποσοτική μέθοδος, μελετώνται αρχικά οι ιδιότητες και τα χαρακτηριστικά των παραπάνω μοντέλων.

3.3.1 Μοντέλο Χρονοσειρών

Αποτελεί το πιο διαδεδομένο μοντέλο όσον αφορά στις ποσοτικές μεθόδους πρόβλεψης. Είναι εύκολο στην εφαρμογή του, έχει μικρό κόστος και βασίζεται στην υπόθεση ότι η τιμή του εξεταζόμενου μεγέθους μεταβάλλεται με βάση ένα καθορισμένο και σταθερό πρότυπο που επαναλαμβάνεται στο χρόνο. Επομένως η πρόβλεψη επιτυγχάνεται με την αναγνώριση αυτού του προτύπου της χρονοσειράς και την επέκτασή του στο μέλλον. Επομένως είναι πολύ σημαντικό να έχουμε στη διάθεση μας ένα ικανό εύρος δεδομένων του υπό πρόβλεψη μεγέθους ώστε να επιλέξουμε το κατάλληλο πρότυπο συμπεριφοράς και να διασφαλίσουμε όσο το δυνατόν πιο ακριβείς προβλέψεις. Το μοντέλο των χρονοσειρών (time series model) είναι κατάλληλο στις περιπτώσεις όπου η εξέλιξη του εξεταζόμενου μεγέθους δεν επηρεάζεται από άλλες παραμέτρους αλλά βασίζεται στις προηγούμενες τιμές του. Το πιο σημαντικό μειονέκτημα αυτού του μοντέλου έγκειται δηλαδή στη μη συσχέτιση του προβλεπόμενου μεγέθους με άλλους παράγοντες που θα μπορούσαν να το επηρεάζουν.

Τα βασικότερα μοντέλα χρονοσειρών είναι η αποσύνθεση (decomposition), η εξομάλυνση (smoothing) και οι αυτοπαλινδρομικές μέθοδοι κινητού μέσου όρου (autoregressive moving average) για τις οποίες γίνεται λόγος στο επόμενο κεφάλαιο.

3.3.2 Αιτιοκρατικό ή Επεξηγηματικό μοντέλο

Το αιτιοκρατικό μοντέλο (causual relationship model) βασίζεται στην υπόθεση ότι υπάρχει μια σταθερή σχέση ανάμεσα στο προβλεπόμενο μέγεθος (εξαρτημένη μεταβλητή) και σε ορισμένες παραμέτρους που το επηρεάζουν (ανεξάρτητες μεταβλητές) χωρίς να είναι απαραίτητη η ύπαρξη χρονικής συνάρτησης. Σε αντίθεση με το μοντέλο χρονοσειρών όπου ορίζεται από την υπόθεση η συνάρτηση που περιγράφει το σύστημα, στο επεξηγηματικό μοντέλο προσδιορίζονται αρχικά τις τιμές των ανεξάρτητων μεταβλητών και στη συνέχεια

επιχειρείται η πρόβλεψη της ζητούμενης μεταβλητής. Το σύστημα περιγράφεται δηλαδή από μια εξίσωση της μορφής:

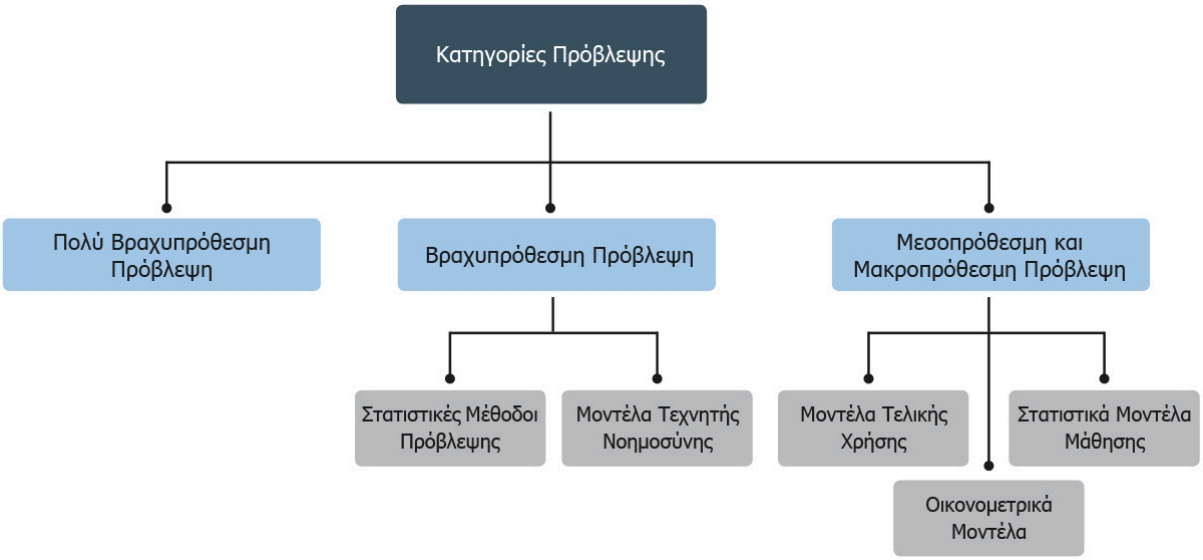
$$Y = f(X_i) \tag{3.1}$$

όπου
X_i: οι παράμετροι που επηρεάζουν τη μεταβολή του προβλεπόμενου μεγέθους
Y: η εξαρτημένη μεταβλητή

Σημαντικά πλεονεκτήματα του αιτιοκρατικού μοντέλου είναι ότι καθιστά πιο εύκολη την κατανόηση των συνθηκών και επιτρέπει στο χρήστη να προβλέψει τη μελλοντική τιμή ενός μεγέθους για διάφορους συνδυασμούς των μεταβλητών εισόδου. Ωστόσο παρουσιάζει και μερικά μειονεκτήματα καθώς απαιτείται μεγάλο πλήθος δεδομένων όχι μόνο για τη μεταβλητή που μελετάται αλλά και για το πλήθος των παραμέτρων που την επηρεάζουν. Επιπλέον μερικές φορές είναι απαραίτητη και η πρόβλεψη των ανεξάρτητων μεταβλητών, στοιχείο που μπορεί να αυξήσει σημαντικά το κόστος της εφαρμογής. Στο επεξηγηματικό μοντέλο ανήκουν οι μέθοδοι παλινδρόμησης (απλή γραμμική, πολλαπλή γραμμική παλινδρόμηση) και οι οικονομετρικές μέθοδοι.

3.4 Κατηγορίες Μεθόδων Πρόβλεψης Φορτίου

Όπως αναφέραμε στην παράγραφο 3.1.2 μπορεί να γίνει διάκριση των μεθόδων πρόβλεψης με βάση τον χρονικό τους ορίζοντα σε πολύ βραχυπρόθεσμη, βραχυπρόθεσμη, μεσοπρόθεσμη και μακροπρόθεσμη πρόβλεψη. Στα πλαίσια της παρούσας διπλωματικής δίνεται ιδιαίτερη έμφαση στην περίπτωση της βραχυπρόθεσμης πρόβλεψης του ηλεκτρικού φορτίου. Αφορά στην πρόβλεψη του φορτίου για τις επόμενες ώρες ως και μια εβδομάδα μπροστά και είναι απαραίτητη για την ενεργειακή διαχείριση ενός ΣΗΕ ενώ συμβάλλει και στην ασφάλεια του συστήματος (παράγραφος 2.2.3). Βάσει των τεχνικών που χρησιμοποιούν οι μέθοδοι βραχυπρόθεσμης πρόβλεψης μπορούμε να διακρίνουμε δύο ευρύτερες κατηγορίες: Τα μοντέλα που ακολουθούν μια πιο κλασσική προσέγγιση με τη χρήση μεθόδων στατιστικής (μέθοδοι εξομάλυνσης, παλινδρόμησης, στοχαστικά μοντέλα κ.α.) και τα νεότερα μοντέλα που ανήκουν στο επιστημονικό πεδίο της τεχνητής νοημοσύνης (νευρωνικά δίκτυα, μηχανές διανυσμάτων υποστήριξης, ασαφής λογική κ.α.).



Σχήμα 3.2: Βασικές κατηγορίες μεθόδων πρόβλεψης

3.4.1 Μέθοδοι Βραχυπρόθεσμης Πρόβλεψης Φορτίου

Η βραχυπρόθεσμη πρόβλεψη φορτίου (short-term load forecasting) έχει χρονικό ορίζοντα από μια ώρα ως και μια εβδομάδα μπροστά και είναι απαραίτητη για τη διαχείριση ενός συστήματος ηλεκτρικής ενέργειας. Η λήψη αποφάσεων που αφορούν το σχεδιασμό, τη λειτουργία σε πραγματικό χρόνο, καθώς και την ανάλυση της ασφάλειας του συστήματος εξαρτάται άμεσα από την ποιότητα της πρόβλεψης ζήτησης φορτίου. Πρόβλεψη με υψηλή ακρίβεια μας επιτρέπει να βελτιστοποιήσουμε την παραγωγή και τη λειτουργία του συστήματος [38]. Κατά τη διάρκεια των τελευταίων δεκαετιών έχουν αναπτυχθεί διάφορες μέθοδοι βραχυπρόθεσμης πρόβλεψης. Οι δύο πιο σημαντικές κατηγορίες μεθόδων βραχυπρόθεσμης πρόβλεψης είναι οι ακόλουθες:

- Στατιστικές Μέθοδοι (Statistical Methods)
- Μοντέλα Τεχνητής Νοημοσύνης (Artificial Intelligence Models)

Στο κεφάλαιο 4 που ακολουθεί επιχειρείται μια εκτενέστερη ανάλυση των μεθόδων βραχυπρόθεσμης πρόβλεψης.

3.4.2 Μέθοδοι Μεσοπρόθεσμης και Μακροπρόθεσμης Πρόβλεψης Φορτίου

Με την απελευθέρωση των αγορών ενέργειας, παρατηρούμε ότι εντείνεται το ενδιαφέρον και για προβλέψεις με μεγαλύτερο χρονικό ορίζοντα. Σύμφωνα με το [36] οι μεσοπρόθεσμες προβλέψεις φορτίου επιτρέπουν στις εταιρείες να εκτιμήσουν τη ζήτηση φορτίου για μεγαλύτερο χρονικό διάστημα, στοιχείο που είναι απαραίτητο κατά τη

διαπραγμάτευση των συμβάσεων με τις άλλες εταιρείες. Η μεσοπρόθεσμη πρόβλεψη φορτίου (medium-term load forecasting) έχει χρονικό ορίζοντα από μία εβδομάδα έως και ένα έτος. Συνήθως ενσωματώνει διάφορες παραμέτρους που επηρεάζουν την πρόβλεψη του φορτίου με πιο βασικές τις δημογραφικές και οικονομικές επιρροές. Κατά κύριο λόγο παρέχει την ημερήσια αιχμή και το μέσο φορτίο των υπό εξέταση δεδομένων, αλλά υπάρχουν και περιπτώσεις όπου προσεγγίζει ωριαία φορτία κατανάλωσης [39]. Η προσπάθεια πρόβλεψης του ηλεκτρικού φορτίου για περισσότερα έτη κατατάσσεται στη μακροπρόθεσμη πρόβλεψη φορτίου (long-term load forecasting) [34]. Σύμφωνα με το [32] ο χρονικό ορίζοντας των μακροπρόθεσμων προβλέψεων επεκτείνεται συνήθως έως τα 20 χρόνια μπροστά. Σε αυτήν την περίπτωση είναι απαραίτητη η λήψη περισσότερων δεικτών σχετικά με τη δημογραφική και οικονομική ανάπτυξη της περιόδου που εξετάζεται. Έτσι για παράδειγμα παράγοντες όπως η αύξηση του πληθυσμού, το ακαθάριστο εγχώριο προϊόν και η ανάπτυξη της τεχνολογίας λαμβάνονται υπόψη. Στόχος της μακροπρόθεσμης πρόβλεψης είναι να παρέχει εκτιμήσεις για το ετήσιο φορτίο και το φορτίο αιχμής.

Όπως αναφέρθηκε στην προηγούμενη ενότητα, για τη βραχυπρόθεσμη πρόβλεψη φορτίου χρησιμοποιούνται κυρίως στατιστικές τεχνικές (μέθοδοι εξομάλυνσης, παλινδρόμησης, στοχαστικά μοντέλα κ.α.) ή μέθοδοι τεχνητής νοημοσύνης (νευρωνικά δίκτυα, μηχανές διανυσμάτων υποστήριξης, ασαφής λογική κ.α.). Όσον αφορά στη μεσοπρόθεσμη και μακροπρόθεσμη πρόβλεψη, χρησιμοποιούνται κυρίως οι μέθοδοι τελικής χρήσης και οι οικονομετρικές προσεγγίσεις. Τα μοντέλα τελικής χρήσης απαιτούν μεγάλο όγκο δεδομένων προκειμένου να διασφαλίσουν καλή επίδοση ([40]), ενώ επηρεάζονται σημαντικά και από την ποιότητα των διαθέσιμων δεδομένων. Από την άλλη πλευρά οι οικονομετρικές προσεγγίσεις παρουσιάζουν έντονο ενδιαφέρον κυρίως για τις οικονομικές μεταβλητές. Πολλές φορές επιχειρείται συνδυασμός αυτών των δύο μεθόδων, ενώ η ανθρώπινη παρέμβαση κρίνεται σημαντική σε όλη τη διαδικασία. Διάφορες προσεγγίσεις για τις μεθόδου που αναφέραμε παρουσιάζονται στα ([41], [42]).

Κατά την ανάπτυξη των μοντέλων τελικής χρήσης συλλέγονται συνήθως πληροφορίες που αφορούν σε περιγραφές των συσκευών που χρησιμοποιούν οι πελάτες, στο μέγεθος των εξεταζόμενων κτηρίων, στην ηλικία του εξοπλισμού, στη συμπεριφορά των πελατών και τη δυναμική του πληθυσμού [34]. Επιπλέον στα οικονομετρικά μοντέλα περιλαμβάνονται στοιχεία οικονομικών παραγόντων όπως οι τιμές της ηλεκτρικής ενέργειας, το κατά κεφαλήν εισόδημα, τα επίπεδα απασχόλησης κ.α.

3.4.2.1 Μοντέλα Τελικής Χρήσης

Οι μέθοδοι που προσεγγίζουν την τελική χρήση, προσπαθούν να υπολογίσουν άμεσα την κατανάλωση χρησιμοποιώντας εκτενείς πληροφορίες σχετικά με τους πελάτες και την τελική τους χρήση. Βασίζονται σε πληροφορίες σχετικά με τα μεγέθη των κτηρίων, τις συσκευές που περιλαμβάνονται, τον τρόπο χρήσης τους από τους καταναλωτές κ.α. Τα μοντέλα τελικής χρήσης εστιάζουν στις διάφορες καταναλώσεις της ηλεκτρικής ενέργειας τόσο στον οικιακό όσο και στον εμπορικό και βιομηχανικό τομέα. Πιο συγκεκριμένα στηρίζονται στην αρχή ότι η ζήτηση της ηλεκτρικής ενέργειας προκύπτει αθροιστικά από τη ζήτηση των πελατών για φωτισμό, θέρμανση, ψύξη κ.α. Προσπαθούν επομένως να

εξηγήσουν τη ζήτηση ενέργειας συναρτήσει του συνόλου των συσκευών, όπως προκύπτει από το [40].

Οι στατιστικές αυτές πληροφορίες σχετικά με τους χρήστες αποτελούν τη βάση για τη διαδικασία της πρόβλεψης. Σε ιδανικές συνθήκες η προσέγγιση αυτή θα μπορούσε να είναι πολύ ακριβής. Ωστόσο όπως αναφέρθηκε νωρίτερα η αποτελεσματικότητα της μεθόδου εξαρτάται σε μεγάλο βαθμό από το πλήθος και την ποιότητα των δεδομένων τελικής χρήσης. Αντιλαμβανόμαστε επομένως ότι εκτός από τα ιστορικά δεδομένα, το σύνολο των πληροφοριών που αφορούν τους καταναλωτές και την τελική τους χρήση παίζουν εξίσου σημαντικό ρόλο.

3.4.2.2 Οικονομετρικά Μοντέλα

Η οικονομετρική προσέγγιση προκύπτει από το συνδυασμό της οικονομικής θεωρίας με στατιστικές μεθόδους για την πρόβλεψη της ενεργειακής ζήτησης. Ως εξαρτημένη μεταβλητή παραμένει η καταναλισκόμενη ενέργεια ενώ γίνεται προσπάθεια να προσδιορισθούν οι βασικοί παράγοντες που επιδρούν σε αυτήν. Ο προσδιορισμός της σχέσης μεταξύ του φορτίου και των ανεξάρτητων παραμέτρων επιτυγχάνεται είτε με τη βοήθεια της μεθόδου ελαχίστων τετραγώνων είτε με τη χρήση χρονοσειρών.

Με βάση το πλαίσιο που περιγράψαμε, η διαδικασία ξεκινά αρχικά συγκεντρώνοντας την οικονομετρική προσέγγιση, όπου η κατανάλωση για τους διάφορους τομείς (οικιακό, εμπορικό, βιομηχανικό) έχει υπολογισθεί συναρτήσει παραγόντων όπως καιρικές συνθήκες, οικονομικές μεταβλητές κ.α. Στη συνέχεια οι εκτιμήσεις αυτές συνδυάζονται με τα ιστορικά δεδομένα που είναι διαθέσιμα. Τέλος είναι δυνατή η ένταξη του οικονομετρικού μοντέλου σε προσέγγιση τελικής χρήσης, στοιχείο που εισάγει στις εξισώσεις τελικής χρήσης συνιστώσες σχετικά με τη συμπεριφορά των χρηστών.

3.4.2.3 Στατιστικά Μοντέλα Μάθησης

Όπως περιγράφηκε πιο πάνω, τα μοντέλα τελικής χρήσης και η οικονομετρική μέθοδος μεγάλο όγκο πληροφοριών σχετικά με τους χρήστες, τις συσκευές, οικονομικές παραμέτρους κ.α. Συχνά αυτές οι πληροφορίες δεν είναι διαθέσιμες για συγκεκριμένους πελάτες, οπότε η προσέγγιση που επιλέγεται αφορά στη μοντελοποίηση ενός “μέσου” πελάτη ή ομάδες πελατών για τους διάφορους τομείς. Η εφαρμογή είναι περίπλοκη και απαιτεί την ανθρώπινη συμμετοχή. Σε ορισμένες περιπτώσεις μια εταιρεία ενδιαφέρεται να παράγει προβλέψεις για ορισμένες υπο-περιοχές της κατανάλωσης, στοιχείο που μπορεί να αυξήσει σημαντικά τις υπολογιστικές διαδικασίες. Επίσης δεν πρέπει να ξεχνάμε ότι τα χαρακτηριστικά αυτών των περιοχών μπορεί να διαφέρουν σημαντικά από το “μέσο” προφίλ που έχει υπολογιστεί. Για την αντιμετώπιση τέτοιων περιπτώσεων οι Genethliou, Feinberg κ.α. ([41], [42]) ανέπτυξαν ένα στατιστικό μοντέλο που προσδιορίζει τις παραμέτρους του μοντέλου με τη βοήθεια των ιστορικών δεδομένων. Πιο συγκεκριμένα προέκυψε ότι το μοντέλο που επιτυγχάνει τα καλύτερα αποτελέσματα είναι το πολλαπλασιαστικό, και περιγράφεται από τη σχέση:

$$L(t) = F\big(d(t), h(t)\big) \cdot f\big(w(t)\big) + R(t) \tag{3.2}$$

όπου,
$L(t)$: το πραγματικό φορτίο
$d(t)$: η ημέρα της εβδομάδας
$h(t)$: η ώρα της ημέρας
$F(d, h)$: η ημερήσια και ωριαία συνιστώσα
$w(t)$: τα δεδομένα καιρού (θερμοκρασία, υγρασία)
$f(w)$: ο καιρικός παράγοντας
$R(t)$: το τυχαίο σφάλμα

Επομένως σύμφωνα με τα παραπάνω, το λογισμικό που χρησιμοποιεί τη μέθοδο αυτή έχει την ικανότητα να μάθει τις παραμέτρους που απαιτούνται για την παραγωγή προβλέψεων, ενώ σε περίπτωση που δεν είναι όλα τα ιστορικά δεδομένα διαθέσιμα, μπορεί να στηριχθεί στα μοντέλα των προηγούμενων ετών για να παράγει εκτιμήσεις.

3.4.3 Μέθοδοι Πολύ Βραχυπρόθεσμης Πρόβλεψης

Όπως αναφέρθηκε στην παράγραφο 2.2.1 η πολύ βραχυπρόθεσμη πρόβλεψη (very short-term load forecasting) αφορά στην παραγωγή εκτιμήσεων για το ηλεκτρικό φορτίο τα επόμενα 30 λεπτά έως και μια ώρα, συνήθως με βήμα μερικών δευτερολέπτων έως λεπτού. Αυτές οι πληροφορίες είναι απαραίτητες για τη διατήρηση της ασφάλειας ενός συστήματος ηλεκτρικής ενέργειας (ΣΗΕ), ενώ είναι δυνατή και η παρακολούθηση της σχέσης μεταξύ της συχνότητας και του φορτίου. Με αυτόν τον τρόπο συμβάλλει στην ενεργειακή διαχείριση του ΣΗΕ και επιτρέπει τη βέλτιστη κατανομή του φορτίου στις μονάδες του συστήματος κατά τον οικονομικότερο τρόπο. Τέλος μπορεί να διαδραματίσει σημαντικό ρόλο και στη ρύθμιση της τιμής της ηλεκτρικής ενέργειας.

Σύμφωνα με το Liu,K κ.α. ([43]), τρία βασικά μοντέλα τα οποία μπορούν να προσεγγίσουν με καλά αποτελέσματα την πολύ βραχυπρόθεσμη πρόβλεψη φορτίου είναι οι τεχνικές ασαφούς λογικής (Fuzzy Logic), τα τεχνητά νευρωνικά δίκτυα (Neural Networks) και τα αυτοπαλινδρομικά μοντέλα (Autoregressive Models). Τα αποτελέσματα της έρευνας έδειξαν ότι είναι δυνατή η ανάπτυξη ενός ικανοποιητικού μοντέλου πρόβλεψης με πολύ βραχυπρόθεσμο ορίζοντα, ενώ κατάλληλοι υποψήφιοι αναδείχτηκαν τα νευρωνικά δίκτυα και οι τεχνικές ασαφούς λογικής.

3.4.4 Συνδυασμός Μεθόδων Πρόβλεψης

Η επιλογή της κατάλληλης μεθόδους πρόβλεψης για τα εκάστοτε δεδομένα μέσα από μια πληθώρα διαθέσιμων μοντέλων δεν είναι πάντα εύκολη. Συχνά προτιμάται ο συνδυασμός των προβλέψεων που παράγονται από διαφορετικές μεθόδους με ίσα ή άνισα βάρη. Κάθε μοντέλο πρόβλεψης με βάση τα χαρακτηριστικά του, εξετάζει με διαφορετικό τρόπο τις πληροφορίες μιας χρονοσειράς και επιστρέφει ορισμένα αποτελέσματα. Με το συνδυασμό

των μεθόδων υπάρχει η δυνατότητα αξιοποίησης αυτών των διαφορετικών αποτελεσμάτων για την παραγωγή τελικών προβλέψεων μεγαλύτερης ακρίβειας καθώς περιορίζεται σημαντικά το διάστημα διακύμανσης των σφαλμάτων. Σύμφωνα με το [44] η ακρίβεια της πρόβλεψης από το συνδυασμό διαφόρων μεθόδων υπερτερεί κατά μέσο όρο των μεμονωμένων μεθόδων που περιλαμβάνονται σε αυτόν και είναι αρκετά υψηλή σε σύγκριση με άλλες μεθόδους που δεν υπολογίζονται σε αυτόν. Επιπλέον η επιλογή ανάμεσα σε συνδυασμούς μεθόδων εμπεριέχει μικρότερο κίνδυνο σε σχέση με την επιλογή ανάμεσα σε μεμονωμένες μεθόδους πρόβλεψης [32].

Πολύ σημαντικό ρόλο για την αξιοποίηση μια τέτοιας συνδυαστικής προσέγγισης κατέχει η προσεκτική επιλογή των μεμονωμένων μεθόδων που συμμετέχουν. Ανάλογα με τα πρότυπα που ακολουθούν οι υπό εξέταση χρονοσειρές, ορισμένες μέθοδοι βρίσκουν καλή εφαρμογή και ανταποκρίνονται ικανοποιητικά στην εξέλιξη των δεδομένων μέσα στο χρόνο ενώ άλλες αδυνατούν να προσαρμοστούν. Η επιλογή των μεθόδων στηρίζεται κυρίως στην πρώτη ομάδα καθώς οι τελευταίες είναι πιο πιθανό να εισάγουν σημαντικά σφάλματα στη διαδικασία. Η διαμόρφωση της τελικής πρόβλεψης εξαρτάται και από τους συντελεστές βαρύτητας με τους οποίους συμμετέχουν οι διαφορετικές μέθοδοι. Συχνά χρησιμοποιείται ο μέσος όρος των μεμονωμένων μεθόδων ως σημείο αναφοράς για την αξιολόγηση της τελικής πρόβλεψης και ενδεχομένως την επαναπροσαρμογή των επιλεγμένων βαρών.

3.5 Δείκτες Αξιολόγησης των Μεθόδων Πρόβλεψης

Σημαντικό στάδιο για μια επιτυχημένη πρόβλεψη είναι η εύρεση των σφαλμάτων που προκύπτουν από τις διαθέσιμες μεθόδους πρόβλεψης και η αξιολόγηση αυτών. Η μέτρηση της ακρίβειας των παραγόμενων μεθόδων επιτυγχάνεται με τη βοήθεια στατιστικών δεικτών σφαλμάτων, όπου ο καθένας παρέχει διαφορετική πληροφορία [45].

Θεωρούμε ως Y_i την πραγματική τιμή της προς πρόβλεψη παρατήρησης, ως F_i την τιμή της πρόβλεψης και με τον αριθμό n αναφερόμαστε στον αριθμό των προς πρόβλεψη παρατηρήσεων. Το σφάλμα της πρόβλεψης περιγράφεται από την ακόλουθη σχέση:

$$e_t = Y_t - F_t \tag{3.3}$$

Στην περίπτωση όπου για την εξεταζόμενη περίοδο πρόβλεψης δεν είναι διαθέσιμες οι πραγματικές τιμές του μεγέθους, τότε είμαστε σε θέση να υπολογίσουμε μόνο το σφάλμα του μοντέλου πρόβλεψης, που είναι γνωστό με τον όρο in-sample error. Απαραίτητη προϋπόθεση για να υπολογίσουμε το πραγματικό σφάλμα (out-of-sample error) είναι να έχουμε στη διάθεσή μας τα πραγματικά δεδομένα της εξεταζόμενης περιόδου. Η χρησιμότητα και των δύο σφαλμάτων είναι μεγάλη, καθώς με τη βοήθεια του πρώτου μπορούμε να προσδιορίσουμε τις βέλτιστες τιμές των παραμέτρων για τις μεθόδους που εξετάζουμε, ενώ το δεύτερο μας παρέχει μια εικόνα για την ακρίβεια της μεθόδου, υποδεικνύοντας ενέργειες βελτιστοποίησης.

Στη συνέχεια παρουσιάζουμε τους βασικούς στατιστικούς δείκτες σφάλματος, ομαδοποιημένους με βάση τα χαρακτηριστικά τους.

- Μέσο Σφάλμα (Mean Error)

$$ME = \frac{1}{n}\sum_{i=1}^{n}(Y_i - F_i) \tag{3.4}$$

Προκύπτει ως ο μέσος όρος των σφαλμάτων και εκφράζει ένα μέτρο συστηματικότητας του σφάλματος. Τιμές του δείκτη κοντά στο μηδέν υποδηλώνουν ότι τα σφάλματα δεν είναι συστηματικά αλλά χαρακτηρίζονται από τυχαιότητα. Όταν το σφάλμα λαμβάνει θετικές τιμές αυτό μεταφράζεται σε απαισιοδοξία των προβλέψεων, ενώ αντίστοιχα οι αρνητικές τιμές υποδηλώνουν μια τάση αισιοδοξίας στις προβλέψεις.

Υπάρχουν μερικοί συχνά χρησιμοποιούμενοι δείκτες σφάλματος, η έκταση των οποίων εξαρτάται από την κλίμακα των δεδομένων. Οι δείκτες αυτοί είναι χρήσιμοι για να συγκρίνουμε διαφορετικές μεθόδους πρόβλεψης που εφαρμόζονται σε ένα σύνολο δεδομένων, ενώ η χρήση τους δε συνιστάται για τη σύγκριση συνόλων δεδομένων που έχουν διαφορετική κλίμακα. Σε αυτή την ομάδα σφαλμάτων ανήκουν τα ακόλουθα:

- Μέσο Τετραγωνικό Σφάλμα (Mean Square Error)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - F_i)^2 \tag{3.5}$$

Ο δείκτης αυτός αποτελεί ένα μέτρο της ακρίβειας των χρησιμοποιούμενων μοντέλων πρόβλεψης. Όπως αναφέραμε στην παράγραφο 4.1.2.2, μια σημαντική εφαρμογή του είναι στον υπολογισμό των βέλτιστων συντελεστών εξομάλυνσης. Λόγω του τετραγωνισμού των σφαλμάτων όπως φαίνεται στη σχέση (3.5), αντιλαμβανόμαστε ότι επηρεάζεται έντονα από τα μεγάλα σφάλματα, ενώ δίνει μικρό βάρος στα μικρά σφάλματα. Εφαρμογή του μέσου τετραγωνικού σφάλματος έχουμε από το Makridakis κ.α [46] στο διαγωνισμό Μ, ενώ στα ([47], [48]) παρουσιάζεται έρευνα σχετικά με τις περιπτώσεις όπου η χρήση του δείκτη δεν είναι κατάλληλη.

- Ρίζα του Μέσου Τετραγωνικού Σφάλματος (Root Mean Square Error)

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - F_i)^2} \tag{3.6}$$

Έχει τις ίδιες ιδιότητες με αυτές που αναφέρθηκαν για το μέσο τετραγωνικό σφάλμα (MSE), το οποίο όμως όπως φαίνεται στη σχέση (3.5) εκφράζεται σε μονάδες της αρχικής χρονοσειράς υψωμένες στο τετράγωνο. Έτσι προκειμένου να εκφραστεί το σφάλμα στις μονάδες της χρονοσειράς χρησιμοποιείται το $RMSE$.

- Μέσο Απόλυτο Σφάλμα (Mean Absolute Error)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - F_i| \tag{3.7}$$

Αποτελεί μέτρο αστοχίας της προβλεπόμενης τιμής ως προς την πραγματική, χωρίς ωστόσο να λαμβάνεται υπόψη η κατεύθυνση της πρόβλεψης. Επίσης διατηρεί τις μονάδες μέτρησης της αρχικής χρονοσειράς. Όπως φαίνεται στη σχέση (3.7) όσο μεγαλύτερη είναι η τιμή του δείκτη, αντιστοιχεί σε μικρότερη ακρίβεια του μοντέλου πρόβλεψης.

Συχνά το $RMSE$ προτιμάται έναντι του MSE γιατί τα αποτελέσματά του εκφράζονται στην ίδια κλίμακα με τα δεδομένα πρόβλεψης. Τα δύο αυτά σφάλματα έγιναν δημοφιλή κυρίως λόγω της θεωρητικής τους σημασίας στη στατιστική μοντελοποίηση. Ωστόσο παρουσιάζονται περισσότερο ευαίσθητα στις ακραίες τιμές συγκριτικά με το MAE, στοιχείο που αναγνωρίζουν ορισμένες έρευνες ως μειονέκτημα κατά την αξιολόγηση της ακρίβειας στην πρόβλεψη όπως προκύπτει από το [49]. Τα σφάλματα που ανήκουν σε αυτήν την ομάδα όπως αναφέραμε εκφράζονται σε μονάδες της αρχικής χρονοσειράς. Σε περιπτώσεις όπου οι πραγματικές τιμές του εξεταζόμενου μεγέθους είναι πολύ μεγάλες, ενδεχομένως να δημιουργείται πρόβλημα στη σύγκριση των διάφορων μεθόδων. Για το λόγο αυτό παρουσιάζουμε στη συνέχεια τα ποσοστιαία σφάλματα.

Το βασικό πλεονέκτημα των ποσοστιαίων σφαλμάτων είναι ότι δεν εξαρτώνται από την κλίμακα των δεδομένων, γεγονός που βοηθά τη σύγκριση της απόδοσης της πρόβλεψης ανάμεσα σε διαφορετικά σύνολα δεδομένων. Οι βασικότεροι δείκτες ποσοστιαίων σφαλμάτων είναι οι ακόλουθοι:

- Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (Mean Absolute Percentage Error)

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - F_i|}{|Y_i|}\cdot 100\% \tag{3.8}$$

Όπως προκύπτει από τη σχέση (3.8) ο δείκτης αυτός δε λαμβάνει αρνητικές τιμές. Χρησιμοποιείται συνήθως για την αξιολόγηση μιας μεθόδους πρόβλεψης, κυρίως όταν αυτή εφαρμόζεται σε σύνολα δεδομένων διαφορετικής κλίμακας. Επίσης λόγω της απλότητάς του συχνά προτιμάται σε χρονοσειρές όπου τα δεδομένα τους εμφανίζουν μεγάλες θετικές τιμές. Ωστόσο η χρήση τους δε συνιστάται σε περιπτώσεις χρονοσειρών διακοπτόμενης ζήτησης, καθώς οδηγείται σε απροσδιοριστία.

- Διάμεσος του Απόλυτου Ποσοστιαίου Σφάλματος (Media Absolute Percentage Error)

$$MdAPE = median(APE) = median\left(\sum_{i=1}^{n}\frac{|Y_i - F_i|}{|Y_i|}\cdot 100\%\right) \tag{3.9}$$

Ο δείκτης $MdAPE$ χαρακτηρίζεται από μεγαλύτερη σταθερότητα απέναντι σε ανωμαλίες των τιμών της χρονοσειράς συγκριτικά με το δείκτη σφάλματος $MAPE$ και συνήθως προτιμάται σε τέτοιες περιπτώσεις.

- Ρίζα του μέσου τετραγωνικού ποσοστιαίου σφάλματος (Root Mean Square Percentage Error)

$$RMSPE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i - F_i}{Y_i}\right)^2} \cdot 100\% \tag{3.10}$$

Σε αυτό το σημείο πρέπει να αναφέρουμε ένα σημαντικό μειονέκτημα των τριών παραπάνω δεικτών σφάλματος, που είναι η ύπαρξη απροσδιοριστίας όταν $\mathrm{Y_t} = 0$ για οποιαδήποτε χρονικής στιγμή της εξεταζόμενης περιόδου, αλλά και γενικότερα η εξαιρετικά ασύμμετρη κατανομή τους για τιμές $\mathrm{Y_t}$ πολύ κοντά στο μηδέν. Αυτό σημαίνει για παράδειγμα ότι ο δείκτης $MAPE$ συχνά είναι σημαντικά μεγαλύτερος από το σφάλμα $MdAPE$. Αντιλαμβανόμαστε επομένως την ακαταλληλότητα των δεικτών αυτών σε περιπτώσεις χρονοσειρών διακοπτόμενης ζήτησης [50]. Ένα επιπλέον μειονέκτημα των μεθόδων που βασίζονται σε ποσοστιαία σφάλματα είναι ότι υποθέτουν ένα ουσιαστικό μηδέν. Επομένως για παράδειγμα, δε βγάζουν κάποιο νόημα σε προβλέψεις σφάλματος για θερμοκρασίες σε κλίμακες Celsius ή Fahrenheit. Τέλος παρατηρείται ότι οι δείκτες σφάλματος $MAPE$ και $MdAPE$ δίνουν μεγαλύτερη βαρύτητα στα θετικά σφάλματα παρά στα αρνητικά. Το γεγονός αυτό οδήγησε στη χρήση των λεγόμενων συμμετρικών σφαλμάτων ([51]), τα βασικότερα των οποίων είναι τα ακόλουθα:

- Συμμετρικό Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (Symmetric Mean Absolute Percentage Error)

$$sMAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - F_i|}{\left|\frac{Y_i + F_i}{2}\right|} \cdot 100\% \tag{3.11}$$

Ο δείκτης $sMAPE$ μπορεί να αντιμετωπίσει καλύτερα την απροσδιοριστία συγκριτικά με το δείκτη $MAPE$ σε περίπτωση διακοπτόμενης ζήτησης. Επιπλέον όταν υπάρχει έντονη διαφορά μεταξύ της πρόβλεψης Y_i και τις πραγματικής τιμής F_i μπορεί να αποφύγει τα μεγάλα σφάλματα. Ωστόσο δε δικαιολογεί πλήρως το συμμετρικό χαρακτήρα του, καθώς αντιμετωπίζει με διαφορετικό τρόπο τις αισιόδοξες από τις απαισιόδοξες προβλέψεις [52].

- Διάμεσος του Συμμετρικού Απόλυτου Ποσοστιαίου Σφάλματος (Symmetric Median Absolute Percentage Error)

$$sMdAPE = median\left(\sum_{i=1}^{n}\frac{|Y_i - F_i|}{\left|\frac{Y_i + F_i}{2}\right|} \cdot 100\%\right) \tag{3.12}$$

Ο δείκτης $sMdAPE$ μπορεί να διαχειριστεί πιο αποτελεσματικά χρονοσειρές με πολύ μικρές τιμές δεδομένων. Όπως αναφέρθηκε και για το $MdAPE$, ούτε το $sMdAPE$ δικαιολογεί πλήρως το συμμετρικό του χαρακτήρα, αφού δίνει μεγαλύτερη βαρύτητα στα θετικά σφάλματα σε σχέση με τα αρνητικά, ωστόσο συμβαίνει σε μικρότερο βαθμό συγκριτικά με το $MdAPE$.

Γίνεται αντιληπτό επομένως ότι τα προβλήματα που προκύπτουν από πολύ μικρές τιμές Y_t, εμφανίζουν μικρότερες επιπτώσεις στους συμμετρικούς δείκτες σφαλμάτων. Ωστόσο ακόμα και στην περίπτωση όπου το Y_t είναι κοντά στο μηδέν, είναι επίσης πιθανό το F_t να είναι κοντά στο μηδέν, επομένως σύμφωνα με τη σχέση (3.11) εμπεριέχεται ο κίνδυνος διαίρεσης με αριθμό πλησίον του μηδέν. Τέλος αναφέραμε ότι οι δείκτες που βασίζονται στα ποσοστιαία σφάλματα δεν είναι πάντα τόσο συμμετρικοί και ίσως ένας ενδεχόμενος μετασχηματισμός (πχ. σε λογαρίθμους) να τους έκανε περισσότερο σταθερούς [53].

Ένας εναλλακτικός τρόπος αξιολόγησης μιας μεθόδου πρόβλεψης είναι διαιρώντας το σφάλμα e_t αυτής με το σφάλμα e_t^* μιας μεθόδου που θεωρείται ως βασική. Έτσι ορίζουμε το σχετικό σφάλμα όπως φαίνεται στην ακόλουθη σχέση:

$$r_t = \frac{e_t}{e_t^*} \tag{3.13}$$

Με αυτόν τον τρόπο επιτυγχάνεται η σύγκριση των εξεταζόμενων μεθόδων, έχοντας ως σημείο αναφοράς μια μέθοδο που έχει επιλεχθεί ως βασική. Από τη διαδικασία αυτή προκύπτει μια κατάταξη των μεθόδων με κριτήριο την απόδοσής τους. Ο πιο διαδεδομένος δείκτης σχετικού σφάλματος είναι ο ακόλουθος:

- Μέσο Σχετικό Απόλυτο Σφάλμα (Mean Relative Absolute Error)

$$MRAE = mean(|r_t|) \tag{3.14}$$

Περισσότερες πληροφορίες για τους δείκτες σχετικών σφαλμάτων και τις εφαρμογές τους στον τομέα των προβλέψεων παρουσιάζονται στο **Error! Reference source not found.**.

Στο πλαίσιο της παρούσας διπλωματικής επιλέξαμε το δείκτη σφάλματος $RMSPE$ για να συγκρίνουμε τα αποτελέσματα των μεθόδων για τους διαφορετικούς τύπους πυρήνων (linear, polynomial, RBF kernels). Η επιλογή του τετραγωνικού σφάλματος αιτιολογείται από την αρχή στην οποία στηρίζεται η εξεταζόμενη μέθοδος πρόβλεψης. Οι μηχανές διανυσμάτων υποστήριξης όπως θα αναλυθεί στη συνέχεια στηρίζονται στην ελαχιστοποίηση του δομικού σφάλματος. Επιπλέον με τη χρήση ποσοστιαίου σφάλματος διευκολύνεται η σύγκριση των τριών μοντέλων που αναπτύχθηκαν για τις τρεις χρονικές βάσεις των δεδομένων εισόδου.

3.6 Επιλογή Κατάλληλης Μεθόδου Πρόβλεψης

Η επιλογής της κατάλληλης μεθόδου πρόβλεψης, είναι μια διαδικασία που στηρίζεται σε πλήθος παραγόντων, οι βασικότεροι εκ των οποίων παρουσιάζονται στη συνέχεια. Στην

προηγούμενη παράγραφο αναδείχτηκε η σημαντικότητα των δεικτών σφάλματος για την αξιολόγηση των διαφόρων μεθόδων. Ωστόσο υπάρχουν ορισμένα ακόμα στοιχεία που επηρεάζουν τις δυνατότητες εφαρμογής μιας μεθόδου και επομένως τα αποτελέσματά της. Οι πιο σημαντικοί παράγοντες είναι:

- Τα χαρακτηριστικά των δεδομένων

Απαραίτητη προϋπόθεση για την επιλογή μιας μεθόδου πρόβλεψης είναι αρχικά η αναγνώριση ενός προτύπου συμπεριφοράς των υπό εξέταση δεδομένων. Όπως αναφέρθηκε στην παράγραφο 3.2.2 τα βασικά ποιοτικά χαρακτηριστικά μιας χρονοσειράς είναι η τάση, η κυκλικότητα, η εποχιακότητα και η τυχαιότητα. Η εποχιακότητα δεν παίζει τόσο σημαντικό ρόλο στην επιλογή της μεθόδου αφού, λόγω της κανονικότητας που εμφανίζει, περιγράφεται καλά από τα περισσότερα μοντέλα. Όσον αφορά στη τυχαιότητα, προτιμώνται απλές μέθοδοι σε χρονοσειρές που χαρακτηρίζονται έντονα από αυτήν. Τέλος για χρονοσειρές που εμφανίζουν έντονη τάση ή κυκλικότητα, τα αυτοπαλινδρομικά μοντέλα πρόβλεψης θεωρούνται μια καλή επιλογή. Αντιλαμβανόμαστε επομένως ότι διαφορετικές μέθοδοι πρόβλεψης εφαρμόζονται αποτελεσματικότερα για διαφορετικά πρότυπα δεδομένων.

- Ορίζοντας Πρόβλεψης

Το χρονικό διάστημα στο οποίο αναφέρεται μια πρόβλεψη επηρεάζει σημαντικά τις προτεινόμενες μεθόδους. Σύμφωνα με την παράγραφο 3.4, σε περιπτώσεις βραχυπρόθεσμης πρόβλεψης προτιμώνται στατιστικές μέθοδοι και μοντέλα τεχνητής νοημοσύνης, ενώ ο τύπος των δεδομένων είναι κυρίως ωριαίος ή ημερήσιος. Από την άλλη πλευρά ποιοτικές μέθοδοι και οικονομετρικά μοντέλα βρίσκουν μεγαλύτερη εφαρμογή σε περιπτώσεις μεσοπρόθεσμης και μακροπρόθεσμης πρόβλεψης, όπου τα δεδομένα είναι κυρίως ετήσια ή τριμηνιαία. Στην πρώτη περίπτωση συνήθως στα δεδομένα κυριαρχεί η τυχαιότητα έναντι της τάσης, ενώ στη δεύτερη περίπτωση εξετάζεται κυρίως η τάση των ετήσιων δεδομένων.

- Ευκολία εφαρμογής της μεθόδου

Σε πραγματικές συνθήκες η ευκολία και η απλότητα εφαρμογής μιας μεθόδου παίζει πολύ σημαντικό ρόλο, έτσι προτιμώνται εύληπτες μέθοδοι πρόβλεψης σε αντίθεση με πολύπλοκα μοντέλα που απαιτούν χρονοβόρες υπολογιστικές διαδικασίες.

- Κόστος

Το κόστος εφαρμογής μιας μεθόδου πρόβλεψης είναι ένας εξίσου σημαντικός παράγοντας για την επιλογή της. Το κόστος εξαρτάται κυρίως από τον όγκο των διαθέσιμων δεδομένων και την πολυπλοκότητα του εκάστοτε μοντέλου.

- Αξιοπιστία

Κάθε περίπτωση πρόβλεψης απαιτεί διαφορετικό επίπεδο λεπτομέρειας, το οποίο με τη σειρά του προσδιορίζει τα επιθυμητά όρια αξιοπιστίας. Σφάλμα της τάξης του ±20%

ορισμένες φορές κρίνεται ικανοποιητικό ενώ υπάρχουν περιπτώσεις όπου απαιτείται διακύμανση μικρότερη του ±5%.

3.7 Γενικοί Τύποι Ενεργειακών Μοντέλων

Το μεγαλύτερο κομμάτι της πρωτογενούς ενέργειας σε παγκόσμιο επίπεδο καταναλώνεται σε οικιακά και εμπορικά κτήρια. Η επίτευξη ενεργειακής αποδοτικότητας για ένα κτήριο θεωρείται συχνά ως ο πιο άμεσος τρόπος μείωσης της κατανάλωσης. Έτσι η ανάπτυξη ενός ενεργειακού μοντέλου αναδεικνύεται ως χρήσιμο εργαλείο για τη λήψη αποφάσεων που αφορούν σε υφιστάμενα ή νέα κτήρια. Ένα τέτοιο μοντέλο μπορεί να βοηθήσει στον εντοπισμό της ενεργειακής κατανάλωσης κατά τη διαδικασία του σχεδιασμού, τον προσδιορισμό άλλων παραμέτρων του κτηρίου (θέρμανση, ψύξη, εξαερισμός – HVAC) και τη βελτιστοποίηση των συστημάτων ελέγχου [55]. Μπορούμε να διακρίνουμε τους γενικούς τύπους ενεργειακών μοντέλων (general types of energy modeling) σε “παραδοσιακά” και σε μοντέλα με χρήση αισθητήρων.

3.7.1 Παραδοσιακά Ενεργειακά Μοντέλα

Όσον αφορά στον κτηριακό τομέα, τα “παραδοσιακά” ενεργειακά μοντέλα (traditional energy modeling) υποφέρουν από διάφορους παράγοντες, όπως το μεγάλο πλήθος των παραμέτρων εισόδου για το χαρακτηρισμό ενός κτηρίου, την απλούστευση ορισμένων παραδοχών κ.α. Σε αυτήν την κατηγορία οι δύο βασικότεροι τύποι είναι η παραδοσιακή “πρόσθια” μοντελοποίηση και η “αντίστροφη μοντελοποίηση”. Πιο αναλυτικά έχουμε:

- “Πρόσθια” Μοντελοποίηση (“Forward” Modeling)

Σε αυτήν την περίπτωση τα μοντέλα αρχικά χρησιμοποιούν ως παραμέτρους εισόδου στοιχεία όπως τα καιρικά δεδομένα, τη γεωμετρία του κτηρίου, το χρονοδιάγραμμα λειτουργίας κ.α. Στη συνέχεια με τη χρήση κατάλληλου λογισμικού αναπτύσσεται ένα μοντέλο το οποίο κινείται γρήγορα προς τα εμπρός προσομοιώνοντας το χρόνο και υπολογίζει την κατανάλωση ενέργειας για ένα κτήριο. Παρουσίαση και σύγκριση τέτοιων μοντέλων είναι διαθέσιμη στο [56].

- “Αντίστροφη” Μοντελοποίηση (“Inverse” Modeling)

Κατά την αντίστροφη μοντελοποίηση αρχικά υποθέτουμε μια γνωστή μαθηματική σχέση ανάμεσα σε ένα σύνολο εισόδων (θερμοκρασία, υγρασία κ.α.) και στην έξοδο, η οποία στη δική μας περίπτωση είναι η κατανάλωση ενέργειας. Έπειτα με βάση αυτή τη σχέση, γίνεται χρήση στατιστικών μεθόδων για την επίλυση του προβλήματος με τις βέλτιστες παραμέτρους. Στο [57] παρουσιάζονται εφαρμογές αντίστροφης μοντελοποίησης για διαφορετικές περιπτώσεις παραμέτρων.

3.7.2 Ενεργειακά Μοντέλα με χρήση αισθητήρων

Η ανάπτυξη ενός ενεργειακού μοντέλου με τη χρήση αισθητήρων (sensor-based energy modeling) μπορεί να θεωρηθεί ως υβριδική προσέγγιση των δύο παραδοσιακών μοντέλων που αναφέραμε παραπάνω. Η προσέγγιση αυτή υποθέτει ότι τα δεδομένα που λαμβάνονται με τη βοήθεια των αισθητήρων παρέχουν ένα βιώσιμο μοντέλο για το σύνολο του κτηρίου ("προς τα εμπρός" συνιστώσα). Επίσης τα δεδομένα εμπεριέχουν πληροφορίες για τις καιρικές συνθήκες, τον εξοπλισμό του κτηρίου και το χρονοδιάγραμμα του. Έτσι με τη χρήση του κατάλληλου λογισμικού παράγεται στατιστικά το τεχνολογικό μοντέλο ("αντίστροφη" συνιστώσα). Με βάση τα παραπάνω αντιλαμβανόμαστε ότι το ενεργειακό μοντέλο με τη χρήση αισθητήρων λειτουργεί ως εναλλακτική προσέγγιση των παραδοσιακών μοντέλων. Όσον αφορά στη μοντελοποίηση της ηλεκτρικής κατανάλωσης έχουν διεξαχθεί αρκετές μελέτες για εμπορικά κτήρια ([58], [59], [60]). Αντίθετα περιορισμένη έρευνα έχει γίνει σχετικά με τη μοντελοποίηση της κατανάλωσης στον οικιακό τομέα. Η απόκλιση αυτή μεταξύ των μελετών για οικιακά και εμπορικά κτήρια πηγάζει από το γεγονός ότι στην περίπτωση των οικιακών κτηρίων υπάρχει έλλειψη αναλυτικών δεδομένων καθώς συλλέγονται συνήθως σε μηναία βάση.

3.7.3 Σύγκριση Ενεργειακών Μοντέλων

Όσον αφορά στα παραδοσιακά ενεργειακά μοντέλα έχουν επισημανθεί ορισμένα μειονεκτήματα, τα περισσότερα εκ των οποίων στην προσέγγιση με τη χρήση αισθητήρων φαίνεται να μετριάζονται. Πρώτον, λίγες εταιρείες σχεδιασμού έχουν την απαραίτητη τεχνογνωσία για να προσδιορίσουν το σύνολο των παραμέτρων που χαρακτηρίζουν το κτήριο στη φάση του σχεδιασμού. Η προσέγγιση με τη χρήση αισθητήρων έχει ως προαπαιτούμενο την ύπαρξη τέτοιων δεδομένων, έτσι συχνά λαμβάνονται στοιχεία από παρόμοιο κτήριο το οποίο χρησιμοποιείται ως αναφορά και εκπαιδεύονται με τεχνικές μηχανικής μάθησης. Δεύτερον, συχνά υπάρχει απόσταση ανάμεσα στο σχέδιο και το τελικό κτήριο. Είναι δύσκολο να παρακολουθήσουμε τέτοιες αλλαγές, πρόβλημα που αντιμετωπίζεται όταν υπάρχουν δεδομένα από αισθητήρες, αφού μετρούν την πραγματική κατάσταση του κτηρίου. Τρίτον, σπάνια υπάρχει η διαθέσιμη γνώση για να ταξινομηθούν οι ιδιαιτερότητες του εξοπλισμού και συνήθως χρησιμοποιούνται πληροφορίες από το εγχειρίδιο της ASHRAE [61]. Από την άλλη πλευρά με τη βοήθεια των αισθητήρων δε μετράται μόνο η απόδοση των υλικών αλλά και η υποβάθμισή τους με την πάροδο του χρόνου. Τέταρτον, τα παραδοσιακά μοντέλα μπορεί να απαιτούν το χειροκίνητο καθορισμό του συνόλου των μεταβλητών, σε αντίθεση με τους αισθητήρες οι οποίοι μετρούν αξιόπιστα την εξέλιξη των δεδομένων αυτών μέσα στο χρόνο. Πέμπτον, και οι δύο προσεγγίσεις των παραδοσιακών μοντέλων προϋποθέτουν ορισμένες παραδοχές και στρογγυλοποιήσεις. Με τη χρήση αισθητήρων σε συνδυασμό με τη προσέγγιση μηχανικής μάθησης, το μοντέλο μπορεί να αντιμετωπίσει αυτές τις μαθηματικές προκλήσεις.

Στην προηγούμενη παράγραφο αναφέραμε αρκετά πλεονεκτήματα των μοντέλων που βασίζονται στα δεδομένα αισθητήρων, πρέπει ωστόσο να επισημάνουμε και ορισμένους περιορισμούς που εμφανίζουν. Πρώτον η απόκτηση και εγκατάσταση των αισθητήρων

μεταφράζεται σε πρόσθετο κόστος το οποίο δεν απαιτείται στη χρήση των παραδοσιακών μοντέλων. Έτσι αν και η μείωση του κόστους μαζί με την αύξηση των δυνατοτήτων για τους αισθητήρες συνεχίζεται, τις περισσότερες φορές αποτελεί μη αποδοτικό οικονομικά τρόπο. Δεύτερον, υπάρχει ένα ανοιχτό ερώτημα που αφορά στον αριθμό, το είδος και την τοποθέτηση των αισθητήρων που είναι αναγκαίοι ανά περίπτωση ώστε να λαμβάνονται τα απαιτούμενα δεδομένα. Ανάλογα με τη μετρική που αναζητάμε ίσως να χρειάζεται διαφορετικός τύπος, ενώ δεν υπάρχει καθορισμένη διαδικασία για την εκτίμηση του απαραίτητου συνόλου αισθητήρων. Τρίτον, οι αισθητήρες και τα συστήματα συλλογής δεδομένων χαρακτηρίζονται κάποιες φορές από αστάθεια με αποτέλεσμα την απώλεια μετρήσεων ή τη λανθασμένη καταγραφή των τιμών τους. Πρόκειται για τα λεγόμενα missing data, για τα οποία θα γίνει λόγος στη συνέχεια (κεφάλαιο 6). Τέτοιο πρόβλημα συναντάται συχνά σε πραγματικές εφαρμογές, ενώ γίνεται προσπάθεια για τη διόρθωση των τιμών αυτών με τη χρήση ευφυών αλγόριθμων ελέγχου και διασφάλισης ποιότητας. Ωστόσο αν και υπάρχουν ταξινομίες για την κατηγοριοποίηση του τύπου των προβλημάτων και των κατάλληλων αλγορίθμων μάθησης, σπάνια είναι εμφανής η βέλτιστη μέθοδος επίλυσης κάθε προβλήματος, [62].

Στο πλαίσιο της παρούσας διπλωματικής, η διαδικασία μοντελοποίησης βασίζεται σε ενεργειακά δεδομένα πραγματικού χρόνου , που συλλέχθηκαν από “έξυπνους μετρητές” εγκατεστημένους σε μια αλυσίδα δέκα εμπορικών καταστημάτων στην περιοχή της Αττικής. Στόχος της διαδικασίας είναι η εύρεση της βέλτιστης μεθόδου μηχανικής μάθησης, με τη χρήση διανυσμάτων υποστήριξης, για την παραγωγή βραχυπρόθεσμων προβλέψεων ηλεκτρικής κατανάλωσης με χρονικό ορίζοντα μια ημέρα.

ΚΕΦΑΛΑΙΟ 4

ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ

ΚΕΦΑΛΑΙΟ 4: ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ

4.1 Στατιστικές Μέθοδοι Πρόβλεψης

Οι στατιστικές προσεγγίσεις βασίζονται στη μαθηματική αναπαράσταση της καμπύλης του ηλεκτρικού φορτίου. Το ηλεκτρικό φορτίο μελετάται ως συνάρτηση διάφορων παραγόντων, όπως οι καιρικές συνθήκες, ο τύπος της ημέρας, η κατηγορία του πελάτη κ.α. Στη συνέχεια του κεφαλαίου παρουσιάζονται οι βασικότερες μέθοδοι στατιστικής πρόβλεψης.

4.1.1. Απλοϊκή Μέθοδος

Η απλοϊκή μέθοδος (Naive) είναι η πιο απλή στατιστική μέθοδος. Οι προβλέψεις της στις περισσότερες περιπτώσεις δεν είναι πολύ ακριβείς, ωστόσο χρησιμοποιούνται ως σημείο αναφοράς (benchmark) συγκριτικά με άλλα πιο εξελιγμένα μοντέλα. Με βάση τη Naive μέθοδο η προβλεπόμενη τιμή του εξεταζόμενου μεγέθους είναι ίση με την παρατήρηση της ακριβώς προηγούμενης χρονικής περιόδου. Δηλαδή η πρόβλεψη περιγράφεται από μια σχέση της μορφής:

$$F_t = Y_{t-1} \tag{4.1}$$

όπου
F_t: η πρόβλεψη για την περίοδο t
Y_{t-1}: η πρόβλεψη για την περίοδο $t-1$

Γίνεται εύκολα αντιληπτό ότι για να θεωρηθεί μια μέθοδος αποτελεσματική πρέπει να δίνει πιο ακριβή αποτελέσματα από την απλοϊκή μέθοδο. Η Naive μέθοδος είναι πολύ απλή στην εφαρμογή της και δίνει καλά αποτελέσματα σε σειρές που παρουσιάζουν σχετικά μικρή τάση ή εποχιακότητα.

4.1.2 Μέθοδοι Εξομάλυνσης

Οι μέθοδοι εξομάλυνσης (smoothing) μπορούν να εφαρμοστούν εύκολα και προσφέρουν συχνά βραχυπρόθεσμες προβλέψεις με ικανοποιητική ακρίβεια. Αρχικά στηρίζονται στις προηγούμενες παρατηρήσεις της χρονοσειράς προκειμένου να προσδιορίσουν μια εξομαλυμένη τιμή των δεδομένων. Στη συνέχεια προεκτείνουν αυτήν την εξομαλυμένη τιμή στο μέλλον και έτσι προκύπτει η πρόβλεψη. Η βασική παραδοχή που γίνεται κατά τη χρήση των μεθόδων εξομάλυνσης είναι ότι το υπό εξέταση μέγεθος ακολουθεί ένα λανθάνον πρότυπο συμπεριφοράς με κάποια διακύμανση. Στόχος είναι να εξομαλυνθεί αυτή η διακύμανση από τα διαθέσιμα δεδομένα έτσι ώστε προεκτείνοντας την τιμή που προκύπτει να επιτευχθεί η ζητούμενη πρόβλεψη. Μπορούμε να διακρίνουμε τις μεθόδους εξομάλυνσης στις ακόλουθες κατηγορίες:

- Μέθοδοι απλού ή κινητού μέσου όρου
- Μέθοδοι εκθετικής εξομάλυνσης

4.1.2.1 Μέθοδοι απλού ή κινητού μέσου όρου

Η ονομασία απλός μέσος όρος χρησιμοποιείται για να περιγράψει τη διαδικασία κατά την οποία λαμβάνεται ο μέσος όρος όλων των παρατηρήσεων ως πρόβλεψη [37]. Δηλαδή η πρόβλεψη για την περίοδο $t+1$ περιγράφεται από μια σχέση της μορφής:

$$F_{t+1} = \frac{1}{t}\sum_{i=1}^{t} Y_i \tag{4.2}$$

όπου
F_{t+1}: η πρόβλεψη για την περίοδο t
Y_i: τα δεδομένα των i προηγούμενων περιόδων

Ακολουθώντας την παραπάνω διαδικασία μπορούμε κάθε φορά να υπολογίσουμε την τιμή για την περίοδο $t+1$. Αυτή με τη σειρά της θα συμπεριλαμβάνεται ως δεδομένο στο μέσο όρο για την πρόβλεψη της επόμενης περιόδου. Ωστόσο η μέθοδος του απλού μέσου όρου σπάνια χρησιμοποιείται για την παραγωγή προβλέψεων. Στην περίπτωση όπου τα εξεταζόμενα δεδομένα έχουν εποχιακότητα ή τάση, κάθε επόμενη παρατήρηση που παράγουμε δεν επηρεάζει την πρόβλεψη μας η οποία θα συνεχίσει να κινείται κοντά στον προηγούμενο μέσο όρο. Παρατηρείται επομένως αδυναμία της μεθόδου όσον αφορά στην προσαρμογή της στις μεταβολές που μπορεί να παρουσιάσει μια χρονοσειρά. Το πρόβλημα αυτό οδήγησε στην τροποποίηση της αρχικής μεθόδου με στόχο καλύτερη προσαρμογή των προβλέψεων σε περιπτώσεις αλλαγών της χρονοσειράς. Η μέθοδος αυτή είναι γνωστή ως κινητός μέσος όρος και για την πρόβλεψη της λαμβάνει κάθε φορά ένα καθορισμένο και σταθερό πλήθος από τις προηγούμενες χρονικά παρατηρήσεις για να υπολογίσει το μέσο όρο. Έτσι η πρόβλεψη για την περίοδο $t+1$ περιγράφεται από μια σχέση της μορφής:

$$F_{t+1} = \frac{1}{k}\sum_{i=t-k+1}^{t} Y_i \tag{4.3}$$

Κάθε φορά η πρόβλεψη για την περίοδο F_{t+1} στηρίζεται στα πιο πρόσφατα δεδομένα λαμβάνοντας υπόψη την κυκλικότητα ή την τάση που μπορεί να εμφανίζουν. Σε αντίθεση με τη μέθοδο του απλού μέσου όρου, δίνει βαρύτητα σε κάθε νέα παρατήρηση που προκύπτει, ενώ επιτυγχάνει καλύτερα αποτελέσματα και σε δεδομένα που εμφανίζουν εποχιακότητα. Ανάλογα με το πλήθος τον παρατηρήσεων που λαμβάνονται υπόψη, η μέθοδος κινητού μέσου όρου μπορεί να χειριστεί διαφορετικά χαρακτηριστικά των χρονοσειρών. Πιο συγκεκριμένα αυξάνοντας τον αριθμό των δεδομένων επιτυγχάνεται μεγαλύτερη εξομάλυνση ενώ δίνεται μικρότερη σημασία στη διακύμανση των παρατηρήσεων. Αντίστοιχα όσο μικρότερος είναι ο αριθμός των δεδομένων δίνεται μεγαλύτερη σημασία στις διακυμάνσεις των παρατηρήσεων αφού κάθε νέα παρατήρηση

που παράγεται θα έχει μεγαλύτερο βάρος στην πρόβλεψη της επόμενης περιόδου. Τέλος για χρονοσειρές με έντονη εποχιακότητα είναι σημαντική η εφαρμογή της μέθοδο κινητού μέσου όρου με μήκος τουλάχιστον ίσο με αυτήν την περιοδικότητα.

Υπάρχουν μερικές ακόμα παραλλαγές της μεθόδου του μέσου όρου, η ανάλυση των οποίων ξεφεύγει από τα πλαίσια της παρούσας διπλωματικής. Συνοπτικά αυτές είναι:

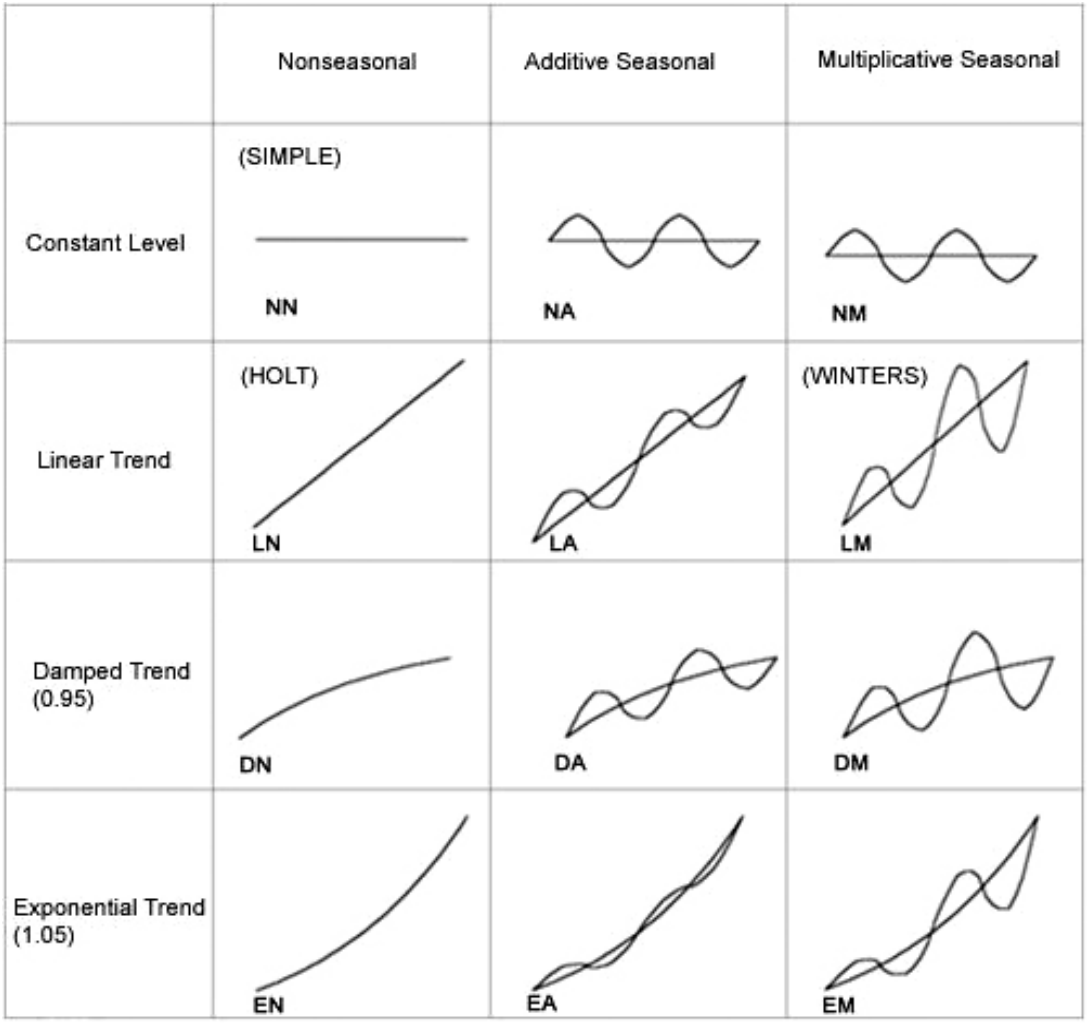
- Σταθμισμένος κινητός μέσος όρος, όπου δίνει έμφαση στις κοντινές παρατηρήσεις γύρω από την τρέχουσα παρατήρηση.
- Διπλός κινητός μέσος όρος, όπου εφαρμόζει τη μέθοδο του κινητού μέσου όρου που περιγράψαμε με ίσα ή άνισα μήκη.
- Κεντρικός κινητός μέσος όρος, όπου προκύπτει από το συνδυασμό του απλού και του διπλού κινητού μέσου όρου και εφαρμόζεται συχνά σε περιπτώσεις με άρτιο πλήθος παρατηρήσεων t.

4.1.2.2 Μέθοδοι εκθετικής εξομάλυνσης

Οι μέθοδοι της εκθετικής εξομάλυνσης (Exponential Smoothing) προκύπτουν με τον υπολογισμό του μέσου όρου δεδομένων όπως αναφέρθηκε παραπάνω, με διαφορετικά ωστόσο βάρη για κάθε παρατήρηση. Η σημαντική διαφορά σε αυτήν την περίπτωση είναι ότι δίνεται μεγαλύτερη βαρύτητα στις πιο πρόσφατες παρατηρήσεις του εξεταζόμενου μεγέθους, ενώ όσο προχωράμε σε δεδομένα από παλαιότερες περιόδους η βαρύτητα αυτή φθίνει.

Οι μέθοδοι εκθετικής εξομάλυνσης είναι ιδιαίτερα δημοφιλείς χάρις σε ορισμένα πλεονεκτήματα που παρουσιάζουν. Πιο συγκεκριμένα είναι ιδιαίτερα απλές στην εφαρμογή τους ενώ δεν εμφανίζουν υψηλές υπολογιστικές απαιτήσεις κατά την εκτέλεσή τους. Επιπλέον δεν έχουν μεγάλες απαιτήσεις σε αποθηκευτικό χώρο καθώς μπορούν να παράγουν προβλέψεις με περιορισμένο αριθμό παρατηρήσεων. Τα αποτελέσματα των μεθόδων εκθετικής εξομάλυνσης έχει αποδειχτεί ότι επιτυγχάνουν ικανοποιητική ακρίβεια ακόμα και σε σύγκριση με πιο πολύπλοκες μεθόδους που θα αναφερθούν στη συνέχεια. Με τη βοήθεια της εξομάλυνσης που λαμβάνει χώρα πριν την πρόβλεψη, δεν επηρεάζονται από ακραίες τιμές και ιδιομορφίες που ενδεχομένως παρουσιάζουν τα δεδομένα.

Με βάση τη γραφική παράσταση της εξεταζόμενης χρονοσειράς συναρτήσει του χρόνου μπορούμε να διακρίνουμε ορισμένες κατηγορίες στις μεθόδους εκθετικής εξομάλυνσης. Έτσι έχουμε τέσσερα μοντέλα τάσης (σταθερού επιπέδου, γραμμικής τάσης, φθίνουσας τάσης και εκθετικής τάσης) τα οποία συνδυάζονται με τρία εποχιακά μοντέλα (μη εποχιακό, προσθετικής εποχιακότητας και πολλαπλασιαστικής εποχιακότητας) και δίνουν τις δώδεκα βασικές κατηγορίες, όπως εμφανίζονται στο ακόλουθο σχήμα.



Σχήμα 4.1: Κατηγοριοποίηση μεθόδων εκθετικής εξομάλυνσης [63]

1. Μοντέλο Σταθερού Επιπέδου

Το μοντέλο σταθερού επιπέδου (Simple Exponential Smoothing), γνωστό και ως απλή εκθετική εξομάλυνση, χρησιμοποιείται σε δεδομένα χωρίς τάση και κυρίως σε προβλέψεις ενός βήματος, ενώ δεν ανταποκρίνεται με καλή ακρίβεια σε μακροχρόνιες προβλέψεις. Η εκάστοτε πρόβλεψη προκύπτει από την προέκταση μιας οριζόντιας ευθείας γραμμής. Έχει αποδειχτεί ότι εμφανίζει καλά αποτελέσματα κατά την εφαρμογή του σε χρονοσειρές με αυξημένο θόρυβο ή τυχαιότητα. Το μοντέλο σταθερού επιπέδου περιγράφεται από τις ακόλουθες εξισώσεις:

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + a \cdot e_t \tag{4.4}$$

$$F_{t+1} = S_t$$

όπου

t: η χρονική περίοδος της πρόβλεψης

Y_t: η πραγματική τιμή των δεδομένων τη χρονική στιγμή t
F_t: η προβλεπόμενη τιμή για τη χρονική στιγμή t
e_t: το σφάλμα της πρόβλεψης
S_t: το επίπεδο της χρονοσειράς
a: ο συντελεστής εξομάλυνσης του επιπέδου της μεθόδου

Ακόμα μπορούμε ενσωματώνοντας τις τρεις προηγούμενες εξισώσεις να συνθέσουμε την ακόλουθη εξίσωση για την πρόβλεψη της περιόδου t:

$$F_t = a \cdot Y_{t-1} + (1 - a) \cdot F_{t+1} \tag{4.5}$$

Αρχικά υπολογίζουμε το σφάλμα πρόβλεψης e_t που ορίζεται ως η διαφορά της πραγματικής τιμής από αυτήν της πρόβλεψης. Έπειτα προχωράμε στον καθορισμό του επιπέδου S_t της εξομαλυμένης χρονοσειράς , το οποίο προκύπτει αθροίζοντας το επίπεδο της προηγούμενης χρονικής στιγμής S_{t-1} με ένα ποσοστό του σφάλματος πρόβλεψης. Το ποσοστό αυτό εξαρτάται από το συντελεστή εξομάλυνσης a, ο οποίος λαμβάνει τιμές εντός του διαστήματος $[0,1]$. Τέλος υπολογίζουμε την πρόβλεψη F_{t+1} που πραγματοποιείται στο τέλος της χρονικής περιόδου t και αφορά σε m περιόδους μπροστά. Η παραπάνω διαδικασία έχει ως στόχο τον καθορισμό του πραγματικού επιπέδου της χρονοσειράς μέσω των σφαλμάτων κάθε πρόβλεψης.

Για την παραγωγή πρόβλεψης με τη μέθοδο της απλής εκθετικής εξομάλυνσης πρέπει πρώτα να ορίσουμε το αρχικό επίπεδο της χρονοσειράς S_0. Είναι σημαντικό να ορίσουμε ένα αντιπροσωπευτικό επίπεδο για τα δεδομένα, στηριζόμενοι στα ποιοτικά χαρακτηριστικά της εκάστοτε χρονοσειράς. Επίσης το αρχικό επίπεδο είναι δυνατό να επηρεάσει τις προβλέψεις όλων των επόμενων τιμών λόγω του επαναληπτικού χαρακτήρα που εμφανίζει η μέθοδος, ενώ συμβάλλει ουσιαστικά και στην επιλογή του συντελεστή εξομάλυνσης α για τον οποίο θα γίνει λόγος παρακάτω. Για την επιλογή του αρχικού επιπέδου της χρονοσειράς μπορεί να χρησιμοποιηθεί ένα από τα παρακάτω:

- Ο μέσος όρος όλων των παρατηρήσεων της χρονοσειράς
- Ο μέσος όρος ορισμένων αρχικών παρατηρήσεων της χρονοσειράς
- Η πρώτη παρατήρηση
- Το σταθερό επίπεδο από το μοντέλο γραμμικής παλινδρόμησης

Πέρα από την αρχική πρόβλεψη είναι σημαντικό να προσδιορίσουμε με επιτυχία και την τιμή του συντελεστή εξομάλυνσης α προκειμένου να παράγουμε πιο ακριβείς προβλέψεις. Οι δύο βασικοί παράγοντες που μπορούν να επηρεάσουν την τιμή του συντελεστή εξομάλυνσης είναι το ποσοστό του θορύβου στη χρονοσειρά και η σταθερότητα που εμφανίζει ο μέσος όρος της.

Όσο πιο μικρή είναι η τιμή του συντελεστή a, οι προβλέψεις μας κινούνται κοντά στο αρχικό επίπεδο τιμών αφού το σφάλμα δε συμμετέχει σημαντικά στη διαμόρφωσή τους. Αντίστοιχα, όσο πιο μεγάλη είναι η τιμή του a, το σφάλμα αποκτά σημαντικό ρόλο στην πρόβλεψη και έτσι επιτυγχάνεται μικρή εξομάλυνση της χρονοσειράς. Αντιλαμβανόμαστε επομένως ότι σε περιπτώσεις όπου υπάρχει αρκετός θόρυβος προτιμάται μικρό a για να

περιορίζεται η επίδρασή του στις προβλεπόμενες τιμές. Όσον αφορά στη σταθερότητα του μέσου όρου της χρονοσειράς, στην περίπτωση που μεταβάλλεται έντονα, προτιμάται μεγάλος συντελεστής εξομάλυνσης προκειμένου να παρακολουθεί τις μεταβολές των δεδομένων και αντίστροφα. Για την ακραία τιμή του $a = 0$ η πρόβλεψη είναι ίση με το αρχικό επίπεδο. Αντίστοιχα για τιμή $a = 1$, το μοντέλο ταυτίζεται πλέον με την απλοϊκή μέθοδο πρόβλεψης, με τα σφάλματα των τελευταίων περιόδων να επηρεάζουν σημαντικά την προβλεπόμενη τιμή.

Η επιλογή της τιμής για το συντελεστή εξομάλυνσης a μπορεί να γίνει είτε προσεγγιστικά είτε με τη χρήση υπολογιστικών εργαλείων. Ο πιο διαδεδομένος τρόπος εύρεσης του βέλτιστου a βασίζεται στη γραμμική αναζήτηση της τιμής που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα (Mean Square Error).

2. Μοντέλο Γραμμικής Τάσης

Το μοντέλο εξομάλυνσης για γραμμική τάση (Holt Exponential Smoothing) αποτελεί μια επέκταση του μοντέλου εκθετικής εξομάλυνσης, με την οποία μπορούμε πλέον να διαχειριστούμε τον παράγοντα της τάσης που ενδεχομένως υπάρχει σε μια χρονοσειρά (Holt 1957). Η προβλεπόμενη τιμή προκύπτει από την προέκταση μιας ευθείας γραμμής για οποιαδήποτε χρονική στιγμή στο μέλλον. Το μοντέλο γραμμικής τάσης περιγράφεται από τις ακόλουθες εξισώσεις:

$$
\begin{aligned}
e_t &= Y_t - F_t \\
S_t &= S_{t-1} + T_{t-1} + a \cdot e_t \\
T_t &= T_{t-1} + \beta \cdot e_t \\
F_{t+m} &= S_t + m \cdot T_t
\end{aligned}
\tag{4.6}
$$

όπου
t: η χρονική περίοδος της πρόβλεψης
Y_t: η πραγματική τιμή των δεδομένων τη χρονική στιγμή t
F_t: η προβλεπόμενη τιμή για τη χρονική στιγμή t
e_t: το σφάλμα της πρόβλεψης
S_t: το επίπεδο της χρονοσειράς
T_t: η τάση της χρονοσειράς
a: ο συντελεστής εξομάλυνσης του επιπέδου
β: ο συντελεστής εξομάλυνσης της τάσης
m: ο χρονικός ορίζοντας πρόβλεψης

Για την παραγωγή πρόβλεψης με τη μέθοδο γραμμικής τάσης είναι απαραίτητο να ορίσουμε το αρχικό επίπεδο S_0 και την αρχική τάση T_0 καθώς σύμφωνα με τις παραπάνω εξισώσεις αποτελεί μια αναδρομική διαδικασία. Όσον αφορά στο αρχικό επίπεδο, η επιλογή του γίνεται με έναν από τους τρόπους που περιγράψαμε στην απλή εκθετική

εξομάλυνση. Σχετικά με την αρχική τάση T_0 μπορεί να χρησιμοποιηθεί ένα από τα παρακάτω:

- Η διαφορά της δεύτερης από την πρώτη παρατήρηση $(Y_2 - Y_1)$
- Η διαφορά της ν −οστής παρατήρησης από την πρώτη, διαιρεμένη με $(\nu - 1)$
- Η σταθερά της κλίσης από το μοντέλο απλής γραμμικής παλινδρόμησης

Σημαντικό στοιχείο της μεθόδου αποτελεί και η εύρεση της βέλτιστης τιμής των συντελεστών εξομάλυνσης a και β, οι οποίοι λαμβάνουν τιμές εντός του διαστήματος $[0,1]$. Συνήθως η βέλτιστη τιμή του συντελεστή β για την τάση είναι μικρότερη από τη βέλτιστη τιμή του a για το επίπεδο. Αυτό εξηγείται εύκολα αν σκεφτούμε ότι συνήθως η τιμή της τάσης για κάθε περίοδο είναι μικρότερη από την τιμή του επιπέδου της χρονοσειράς. Όμοια με την προηγούμενη μέθοδο, η εύρεση των βέλτιστων τιμών προκύπτει μέσω της ελαχιστοποίησης του μέσου τετραγωνικού σφάλματος (MSE).

Παρατηρώντας την τελευταία εξίσωση της μεθόδου, γίνεται αντιληπτό ότι οι προβλεπόμενες τιμές σε αυτήν την περίπτωση εμφανίζουν τάση. Συχνά οι προβλέψεις που παράγονται από αυτήν τη μέθοδο κυρίως για μεσοπρόθεσμο και μακροπρόθεσμο ορίζοντα έχει παρατηρηθεί ότι παρουσιάζουν αισιοδοξία συγκριτικά με τις πραγματικές τιμές.

3. Μοντέλο μη Γραμμικής Τάσης

Το μοντέλο μη γραμμικής τάσης (Gardner και Mckenzie 1985) αποτελεί προσαρμογή του μοντέλου γραμμικής τάσης και έχει τη δυνατότητα να μεταβάλλει τη μορφή της χρονοσειράς σε περίπτωση μη γραμμικών τάσεων. Αυτό επιτυγχάνεται με την προσθήκη μιας επιπλέον παραμέτρου συγκριτικά με το μοντέλο γραμμικής τάσης. Η παράμετρος αυτή ονομάζεται παράμετρος διόρθωσης τάσης φ (trend-modification parameter) και ελέγχει το ρυθμό αύξησης των τιμών της τάσης σε μια χρονοσειρά. Το μοντέλο μη γραμμικής τάσης περιγράφεται από τις ακόλουθες εξισώσεις:

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + T_{t-1} + a \cdot e_t \tag{4.7}$$

$$T_t = T_{t-1} + \beta \cdot e_t$$

$$F_{t+m} = S_t + \sum_{i=1}^{m} \varphi^i \cdot T_t$$

όπου
t: η χρονική περίοδος της πρόβλεψης
Y_t: η πραγματική τιμή των δεδομένων τη χρονική στιγμή t
F_t: η προβλεπόμενη τιμή για τη χρονική στιγμή t
e_t: το σφάλμα της πρόβλεψης
S_t: το επίπεδο της χρονοσειράς
T_t: η τάση της χρονοσειράς

a: ο συντελεστής εξομάλυνσης του επιπέδου
β: ο συντελεστής εξομάλυνσης της τάσης
m: ο χρονικός ορίζοντας πρόβλεψης
φ: ο συντελεστής διόρθωσης της τάσης

Παρατηρούμε εύκολα ότι οι εξισώσεις είναι παρόμοιες με αυτές για το μοντέλο γραμμικής τάσης, με εξαίρεση την τελευταία όπου με τη βοήθεια του συντελεστή διόρθωσης τάσης φ, επιχειρείται μη γραμμικός υπολογισμός της τάσης . Η παράμετρος φ, σε αντίθεση με τους συντελεστές a και β, μπορεί να πάρει θετικές τιμές ακόμα και μεγαλύτερες της μονάδας. Σε αυτό το σημείο είναι σημαντικό να προσδιορίσουμε τη μορφή του μοντέλου μη γραμμικής τάσης με βάση την τιμή της παραμέτρου φ.

- Για $\varphi = 0$, η τάση δε συμμετέχει στον καθορισμό των στατιστικών προβλέψεων και έτσι προκύπτει το μοντέλο της απλής εκθετικής εξομάλυνσης.
- Για $0 < \varphi < 1$, προκύπτει το μοντέλο φθίνουσας τάσης (Damped Exponential Smoothing). Σε αυτήν την περίπτωση η επίδραση της τάσης στο αποτέλεσμα μειώνεται σταδιακά και για αυτό θεωρείται κατάλληλο μοντέλο για μεσοπρόθεσμες προβλέψεις.
- Για $\varphi = 1$, στην εξίσωση της πρόβλεψης έχουμε το γινόμενο $m \cdot T_t$ στη θέση του αθροίσματος και έτσι προκύπτει το μοντέλο γραμμικής τάσης.
- Για $\varphi > 1$, προκύπτει το μοντέλο εκθετικής εξομάλυνσης. Χαρακτηρίζεται από σχετική υπεραισιοδοξία και προτείνεται σε περιπτώσεις όπου μας ενδιαφέρει η πρόβλεψη ζήτησης ενός προϊόντος στην αρχή του κύκλου ζωής του.

Η επιλογή της βέλτιστης τιμής για την παράμετρο εξομάλυνσης φ είναι πολύ σημαντική για την ακρίβεια των προβλέψεων. Η τιμή της παραμέτρου περιορίζεται συνήθως στο διάστημα $[0,1]$ αποκλείοντας με αυτόν τον τρόπο το μοντέλο εκθετικής εξομάλυνσης που οδηγεί σε υπεραισιόδοξες προβλέψεις. Έτσι προκύπτει η μέθοδος φθίνουσας γραμμικής τάσης (Damped Exponential Smoothing). Ο υπολογισμός του βέλτιστου συνδυασμού των συντελεστών a, β και φ επιτυγχάνεται όπως αναφέραμε και στις προηγούμενες μεθόδους με τη βοήθεια του μέσου τετραγωνικού σφάλματος (MSE). Τέλος για την επιλογή του αρχικού επιπέδου S_0 και της αρχικής τάσης T_0 εφαρμόζεται συνήθως γραμμική παλινδρόμηση.

Έχει παρατηρηθεί ότι το μοντέλο μη γραμμικής τάσης παράγει πιο ακριβείς προβλέψεις σε σχέση με το μοντέλο γραμμικής τάσης αν και το τελευταίο είναι περισσότερο δημοφιλές. Πιο συγκεκριμένα το μοντέλο μη γραμμικής τάσης, που με τη βοήθεια της παραμέτρου φ μπορεί να προσαρμόζεται καλύτερα στις μεταβολές της τάσης που εμφανίζουν τα δεδομένα, υπερτερεί σε προβλέψεις μεγάλου χρονικού ορίζοντα πρόβλεψης συγκριτικά με τις υπόλοιπες μεθόδους εκθετικής εξομάλυνσης.

4.1.3 Μέθοδοι Αποσύνθεσης

Οι μέθοδοι αποσύνθεσης (Decomposition) ανήκουν στο μοντέλο των χρονοσειρών και έχουν στόχο την απομόνωση των βασικών συνιστωσών των χρονοσειρών με τη μεγαλύτερη ακρίβεια. Όπως αναφέραμε στην παράγραφο 3.2.2 τα τέσσερα βασικά χαρακτηριστικά των χρονοσειρών είναι η Τάση, η Κυκλικότητα, η Εποχιακότητα και η Τυχαιότητα [63].

Η Τάση T_t αντιπροσωπεύει τη γενική εξέλιξη της χρονοσειράς η οποία μπορεί να είναι ανοδική, σταθερή ή πτωτική. Η Κυκλικότητα C_t αφορά σε μια "κυματοειδή" μεταβολή που εμφανίζεται κατά περιόδους και οφείλεται σε εξωγενείς παράγοντες. Η Εποχιακότητα S_t αντιπροσωπεύει περιοδικές διακυμάνσεις που επαναλαμβάνονται κατά τον ίδιο τρόπο με το χρόνο. Η Τυχαιότητα R_t αφορά στη διαφορά που προκύπτει ανάμεσα στο συνδυασμό των τριών παραπάνω συνιστωσών (T_t, C_t, S_t) και των πραγματικών δεδομένων. Αν θεωρήσουμε ότι τα δεδομένα μας είναι συνάρτηση των τεσσάρων παραπάνω συνιστωσών, τότε η μαθηματική διατύπωση της αποσύνθεσης για την παρατήρηση Y_t είναι η ακόλουθη:

$$Y_t = f(S_t, T_t, C_t, R_t) \tag{4.8}$$

Τα απλούστερα μοντέλα αποσύνθεσης είναι το προσθετικό και το πολλαπλασιαστικό μοντέλο, όπου θεωρούμε ότι οι παρατηρήσεις προκύπτουν ως το άθροισμα ή το γινόμενο των τεσσάρων συνιστωσών αντίστοιχα. Η βασική μεθοδολογία προκύπτει με την απομόνωση διαδοχικά της εποχιακότητας, της τάσης και τέλος της κυκλικότητας. Με αυτόν τον τρόπο μπορούμε να αναγνωρίσουμε την τυχαιότητα της χρονοσειράς. Οι δημοφιλέστερες μέθοδοι αποσύνθεσης είναι η Σταθερή Πολλαπλασιαστική Μέθοδος η οποία είναι γνωστή και ως Κλασσική Μέθοδος Αποσύνθεσης και η μέθοδος Census II, που αποτελεί βελτίωση της πρώτης.

4.1.4 Μέθοδος Theta

Η μέθοδος Theta , Assimakopoulos και Nikolopoulos 2000 ([64]) αποτελεί μια μονοδιάστατη μέθοδο, που προσεγγίζει διαφορετικά την αποσύνθεση με στόχο την παραγωγή προβλέψεων με μεγαλύτερη ακρίβεια. Βασίζεται στη μεταβολή των τοπικών καμπυλοτήτων της χρονοσειράς μέσω της παραμέτρου θ (Theta), η οποία εφαρμόζεται πολλαπλασιαστικά στις διαφορές δεύτερης τάξης των δεδομένων. Η χρονοσειρά που προκύπτει έχει την ίδια μέση τιμή και κλίση με την αρχική χρονοσειρά αλλά διαφορετικές τοπικές καμπυλότητες και διακύμανση. Οι χρονοσειρές που δημιουργούνται μέσα από την παραπάνω διαδικασία ονομάζονται γραμμές Theta (Theta lines). Για διαφορετικές τιμές της παραμέτρου θ μπορούν είτε να αναδείξουν τα βραχυπρόθεσμα χαρακτηριστικά των δεδομένων είτε να προσεγγίσουν τη μακροπρόθεσμη συμπεριφορά αυτών. Όσο μικρότερη είναι η τιμή της παραμέτρου θ, τόσο μεγαλύτερος είναι ο βαθμός μείωσης των τοπικών καμπυλοτήτων, μέχρι που για τιμή $\theta = 0$, η μέθοδος παράγει την ευθεία ελαχίστων τετραγώνων. Πιο συγκεκριμένα για τιμές της παραμέτρου $\theta > 1$ τονίζονται τα βραχυπρόθεσμα χαρακτηριστικά της χρονοσειράς ενώ για τιμές $\theta < 1$ εστιάζουν περισσότερο στη

μακροπρόθεσμη συμπεριφορά. Η παραπάνω τροποποίηση των καμπυλών της χρονοσειράς γίνεται σύμφωνα με τη σχέση:

$$Y_t^{\theta} = \theta \cdot Y_t'' \tag{4.9}$$

$$Y_t'' = Y_t - 2 \cdot Y_{t-1} + Y_{t-2}$$

Με την εφαρμογή της μεθόδου Theta η αρχική χρονοσειρά αποσυντίθεται σε δύο ή περισσότερες τέτοιες γραμμές Theta. Στη συνέχεια οι γραμμές αυτές προεκτείνονται ως ξεχωριστές χρονοσειρές στο μέλλον και τέλος με το συνδυασμό των παραγόμενων προβλέψεων εξάγουμε την τελική πρόβλεψη. Με βάση τα χαρακτηριστικά της χρονοσειράς επιλέγονται τα κατάλληλα βάρη με τα οποία θα συνδυαστούν οι προβλέψεις από τις γραμμές Theta για να προκύψει η τελική πρόβλεψη.

Η κλασσική μέθοδος Theta (Theta Classic) αποσυνθέτει την αρχική χρονοσειρά σε δυο γραμμές Theta με παραμέτρους $\theta = 0$ (ευθεία γραμμή) και $\theta = 2$ (διπλασιασμός των τοπικών καμπυλοτήτων). Αυτή η μέθοδος έχει εφαρμοστεί για την παραγωγή προβλέψεων για τις 3003 χρονοσειρές του διαγωνισμού M3 ([44]) με πολύ καλά αποτελέσματα ιδιαίτερα όσον αφορά στις μηνιαίες χρονοσειρές και στα μακροοικονομικά δεδομένα.

Η απλοποιημένη εξίσωση που περιγράφει την κάθε γραμμή Theta είναι η ακόλουθη:

$$Y_t^{\theta} = \theta \cdot Y_t + (1 - \theta) \cdot LRL_t \tag{4.10}$$

όπου
LRL (Linear regression line): η ευθεία γραμμικής παλινδρόμησης (η γραμμή Theta με $\theta = 0$)

Επομένως η εξίσωση σύμφωνα με την οποία προκύπτει η τελική πρόβλεψη της κλασσικής μεθόδου Theta είναι η ακόλουθη:

$$Y_t = \frac{1}{2} \cdot (Y_t^{\theta=1+\alpha} + Y_t^{\theta=1-\alpha}) \tag{4.11}$$

Η μέθοδος αυτή έχει στόχο να διατηρήσει τα μακροπρόθεσμα χαρακτηριστικά μέσω της γραμμής Theta με $\theta = 0$ αλλά και να διατηρήσει τη βραχυπρόθεσμη πληροφορία που φέρουν τα δεδομένα μέσω της γραμμής Theta με $\theta = 2$. Μέσω αυτής της διαδικασίας επιτυγχάνεται καλύτερη αξιοποίηση των πληροφοριών που βρίσκονται στα δεδομένα πριν εφαρμοστεί κάποιο μοντέλο επέκτασης και αποδοτικότερος συνδυασμός των παραγόμενων προβλέψεων. Κλείνοντας, γίνεται αντιληπτό ότι η μέθοδος Theta με τη βοήθεια της παραμέτρου θ μπορεί να προσεγγίσει με επιτυχία τόσο τη μακροπρόθεσμη όσο και τη βραχυπρόθεσμη συμπεριφορά μιας χρονοσειράς.

4.1.5 Μοντέλα Παλινδρόμησης

Η κύρια χρήση των μοντέλων παλινδρόμησης (Regression Models) αφορά στην εύρεση συσχετίσεων ανάμεσα σε μια εξαρτημένη μεταβλητή (εξεταζόμενο μέγεθος) και μια ή

περισσότερες ανεξάρτητες μεταβλητές που την επηρεάζουν. Ανάλογα με τον αριθμό των ανεξάρτητων μεταβλητών προκύπτουν τα μοντέλα απλής ή πολλαπλής γραμμικής παλινδρόμησης. Σχετικά με τις προβλέψεις, αν θεωρήσουμε δεδομένες τις τιμές των ανεξάρτητων μεταβλητών μπορούμε με τη βοήθεια των μοντέλων παλινδρόμησης να κατανοήσουμε την εξέλιξη της εξαρτημένης μεταβλητής μέσω της εξίσωσης παλινδρόμησης. Ένα σημαντικό πλεονέκτημα αυτών των μεθόδων έγκειται στη δυνατότητα να προσεγγίσουμε την αλλαγή της εξαρτημένης μεταβλητής για οποιαδήποτε αλλαγή μιας από τις ανεξάρτητες μεταβλητές, εφόσον οι υπόλοιπες παραμένουν σταθερές. Επιπλέον σημαντικό ρόλο παίζει η διακύμανση της εξαρτημένης μεταβλητής γύρω από την εξίσωση παλινδρόμησης, η οποία μπορεί να περιγραφεί μέσω πιθανοτικής κατανομής. Η διαδικασία πρόβλεψης με τη βοήθεια μοντέλων παλινδρόμησης προϋποθέτει τη γνώση ή πρόβλεψη των ανεξάρτητων μεταβλητών, στοιχείο που συχνά καθιστά τη χρήση των μοντέλων δύσκολη.

Η παλινδρόμηση είναι η μία από τις πιο ευρέως χρησιμοποιούμενες στατιστικές τεχνικές. Μοντέλα της χρησιμοποιούνται πολύ συχνά στην πρόβλεψη ηλεκτρικού φορτίου ([32]) για τη διαμόρφωση της σχέσης μεταξύ της κατανάλωσης φορτίου και εξωτερικών παραγόντων όπως οι καιρικές συνθήκες, ο τύπος της ημέρας, η κατηγορία του πελάτη κ.α. [34]. Οι μέθοδοι παλινδρόμησης υποθέτουν ότι το φορτίο μπορεί να διαχωριστεί σε μια βασική τάση φορτίου και μια γραμμική τάση η οποία εξαρτάται από τους παραπάνω παράγοντες που επηρεάζουν το φορτίο [65]. Στο [66] παρουσιάζονται διάφορα μοντέλα πρόβλεψης της αιχμής του φορτίου για την επόμενη μέρα. Παρατηρούμε ότι τα μοντέλα αυτά ενσωματώνουν επιρροές από εξωγενείς παράγοντες όπως ο καιρός, ντετερμινιστικές επιδράσεις όπως οι περίοδοι διακοπών (holidays) και στοχαστικές επιδράσεις όπως η μέση τιμή των φορτίων. Πρόσθετες εφαρμογές των μοντέλων παλινδρόμησης σε προβλήματα πρόβλεψης φορτίου παρουσιάζονται στα ([67], [68], [69], [70]). Οι μέθοδοι παλινδρόμησης προσφέρουν αρκετά πλεονεκτήματα, [33]. Βοηθούν στην κατανόηση της σχέσης μεταξύ της εισόδου και των μεταβλητών εξόδου. Επιπλέον επιτρέπουν σχετικά εύκολα την αξιολόγηση της απόδοσης [39]. Ωστόσο σύμφωνα με το [32] μπορεί να υπάρχουν εγγενή προβλήματα στον εντοπισμό του σωστού μοντέλου, γεγονός που αιτιολογείται από τη μη γραμμική σχέση μεταξύ του φορτίου και των παραγόντων που το επηρεάζουν.

Οι δυο βασικές κατηγορίες των μοντέλων παλινδρόμησης που θα αναλυθούν παρακάτω είναι η απλή και η πολλαπλή γραμμική παλινδρόμηση.

4.1.5.1 Απλή Γραμμική Παλινδρόμηση

Στην περίπτωση της απλής γραμμικής παλινδρόμησης (Linear Regression) εξετάζουμε τη σχέση του μεγέθους που θέλουμε να προβλέψουμε (εξαρτημένη μεταβλητή) και μιας μόνο μεταβλητής που το επηρεάζει (ανεξάρτητη μεταβλητή). Υποθέτουμε ότι η σχέση των δύο μεταβλητών είναι γραμμική αν και στην πραγματικότητα αυτό δε συμβαίνει συχνά. Όσον αφορά στην πρόβλεψη χρονοσειρών συνήθως θεωρούμε το χρόνο ως ανεξάρτητη μεταβλητή. Οι μαθηματικές σχέσεις που περιγράφουν αυτό το μοντέλο πρόβλεψης είναι οι ακόλουθες:

$$\hat{Y}_i = \alpha + b \cdot X_i$$

$$b = \frac{\sum_{i=1}^{n}[(X_i - \bar{X}) \cdot (Y_i - \bar{Y})]}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{4.12}$$

$$\alpha = \bar{Y} - b \cdot \bar{X}$$

όπου
X_i: η ανεξάρτητη μεταβλητή
Y_i: η εξαρτημένη μεταβλητή
$\hat{Y}_i$: η εκτιμώμενη τιμή της εξαρτημένης μεταβλητής
α: η τεταγμένη του σημείου τομής με τον άξονα των εξαρτημένων μεταβλητών
b: η κλίση της ευθείας
$\bar{X}$: η μέση τιμή των ανεξάρτητων μεταβλητών
$\bar{Y}$: η μέση τιμή των εξαρτημένων μεταβλητών
n: ο αριθμός των διαθέσιμων παρατηρήσεων

Η πρώτη εξίσωση περιγράφει την ευθεία γραμμή που συνδέει τις δύο μεταβλητές. Για τον υπολογισμό των συντελεστών α και b γίνεται χρήση της μεθόδου ελαχίστων τετραγώνων. Δηλαδή επιλέγονται εκείνες οι τιμές των συντελεστών που ελαχιστοποιούν το άθροισμα των τετραγώνων των διαφορών ανάμεσα στις πραγματικές και τις προβλεπόμενες τιμές.

Τα αποτελέσματα που προκύπτουν από την παραπάνω μέθοδο συνήθως μας δίνουν μια καλή εικόνα για τη συμπεριφορά της χρονοσειράς σε μεσοπρόθεσμο και μακροπρόθεσμο ορίζοντα. Όπως αναφέραμε παραπάνω, η σχέση μεταξύ εξαρτημένης και ανεξάρτητης μεταβλητής τις περισσότερες φορές δεν είναι γραμμική. Σε αυτήν την περίπτωση απαιτείται αρχικά ο μετασχηματισμός της σχέσης των δύο μεταβλητών σε γραμμική προτού εφαρμοστεί η μέθοδος. Επιπλέον με τη χρήση του συντελεστή γραμμικής συσχέτισης των δύο μεταβλητών μπορούμε να αποφανθούμε σχετικά με το βαθμό εξάρτησης αυτών. Ο συντελεστής μπορεί να λάβει τιμές εντός του διαστήματος $[-1,1]$. Πιο συγκεκριμένα για τιμή ίση με το μηδέν προκύπτει ότι οι δυο μεταβλητές είναι ανεξάρτητες μεταξύ τους, ενώ για τιμές ± 1 υπάρχει απόλυτη συσχέτιση. Γενικά για θετικές τιμές του συντελεστή συσχέτισης οι αυξομειώσεις των μεταβλητών έχουν την ίδια κατεύθυνση. Το αντίθετο συμβαίνει για αρνητικές τιμές του συντελεστή. Η γραμμική παλινδρόμηση αν και απλή τεχνική, μπορεί να προσφέρει μια βασική εκτίμηση απόδοσης. Σύμφωνα με αυτό γίνεται αντιληπτό ότι αν μια τεχνική πρόβλεψης παράγει αποτελέσματα χειρότερα από το βασικό μοντέλο τότε είναι πιο πιθανό να μην είναι το κατάλληλο μοντέλο για το εξεταζόμενο σύνολο δεδομένων [55].

4.1.5.2 Πολλαπλή Γραμμική Παλινδρόμηση

Η μέθοδος της πολλαπλής γραμμικής παλινδρόμησης αποτελεί γενίκευση της παραπάνω μεθόδου και περιλαμβάνει περισσότερες από μια ανεξάρτητες μεταβλητές. Σκοπός είναι η πρόβλεψη μιας εξαρτημένης μεταβλητής λαμβάνοντας υπόψη όλες τις μεταβλητές που

μπορεί να το επηρεάζουν. Η μαθηματική σχέση που περιγράφει το μοντέλο πολλαπλής γραμμικής παλινδρόμησης είναι η ακόλουθη:

$$Y = b_0 + b_1 \cdot X_{1,i} + b_2 \cdot X_{2,i} + \cdots + b_k \cdot X_{k,i} + e_i \tag{4.13}$$

όπου
Y: η εξαρτημένη μεταβλητή
X: οι ανεξάρτητες μεταβλητές
b_i: συντελεστές (σταθερές παράμετροι)
e_i: απόκλιση από τις πραγματικές τιμές

Απαραίτητη προϋπόθεση για την παραγωγή του μοντέλου πολλαπλής παλινδρόμησης είναι η απουσία συσχέτισης μεταξύ των ανεξάρτητων μεταβλητών. Υποθέτουμε ότι η απόκλιση της εκτίμησης από την πραγματική τιμή είναι τυχαία. Η συνάρτηση που περιγράψαμε είναι γραμμική ως προς τους συντελεστές. Στόχος είναι ο υπολογισμός των σταθερών παραμέτρων b_i, ο οποίος επιτυγχάνεται με τη μέθοδο των ελαχίστων τετραγώνων. Η διαδικασία εύρεσης της μορφής για την παραπάνω συνάρτηση δεν είναι εύκολη, έτσι ασχολούμαστε αρχικά με τον υπολογισμό των συντελεστών b_i και του τυχαίου παράγοντα e_i. Πιο αναλυτικά για κάθε συντελεστή υπολογίζονται οι μερικές παράγωγοι του αθροίσματος του τετραγώνου των σφαλμάτων e_i, στη συνέχεια τίθενται ίσες με το μηδέν και προκύπτει ένα σύστημα εξισώσεων. Η επίλυση αυτού του γραμμικού συστήματος οδηγεί στην εύρεση των ζητούμενων συντελεστών.

4.1.6 Οικονομετρική μέθοδος

Η οικονομετρική μέθοδος εντάσσεται στο αιτιοκρατικό μοντέλο και επομένως προϋποθέτει την ύπαρξη σταθερής σχέσης ανάμεσα στο προβλεπόμενο μέγεθος (εξαρτημένη μεταβλητή) και σε ορισμένες παραμέτρους που το επηρεάζουν (ανεξάρτητες μεταβλητές) χωρίς να είναι απαραίτητη η ύπαρξη χρονικής συνάρτησης. Στόχος αρχικά είναι αξιοποίηση των υπαρχόντων πληροφοριών για την ανάπτυξη ενός μοντέλου. Στη συνέχεια έχουμε την επιλογή των δεδομένων για το εκάστοτε πρόβλημα και τη μεθοδική ανάλυση αυτών. Τέλος επιχειρείται η προσαρμογή του μοντέλου στα συμπεράσματα που προκύπτουν. Πολύ σημαντικό ρόλο σε αυτήν τη μέθοδο κατέχουν οι πληροφορίες σχετικά με τις αιτιοκρατικές σχέσεις των μεγεθών, βάσει των οποίων προβλέπουμε τις αλλαγές που επιφέρουν οι μεταβολές των παραμέτρων. Συνήθως τα οικονομετρικά μοντέλα αναπτύσσονται με τρόπο που περιγράφουν εξειδικευμένα ένα συγκεκριμένο πρόβλημα κάθε φορά, στοιχείο που αποτρέπει την ευρεία χρήση σε διάφορες περιπτώσεις.

4.1.7 Στοχαστικές Χρονοσειρές

Οι μέθοδοι χρονοσειρών (Stochastic Time Series) ξεκινούν με την παραδοχή ότι τα εξεταζόμενα δεδομένα παρουσιάζουν μια εσωτερική δομή, όπως αυτοσυσχέτιση, τάση ή εποχιακή διακύμανση, [34]. Στόχος αυτών των μεθόδων πρόβλεψης είναι να εντοπίσουν

και να διερευνήσουν τέτοιες δομές. Αρχικά επιχειρείται η προσέγγιση ενός μοτίβου που να ταιριάζει στα διαθέσιμα δεδομένα, με όσο το δυνατόν μεγαλύτερη ακρίβεια, και στη συνέχεια η παραγωγή της προβλεπόμενης τιμής σε σχέση με το χρόνο, χρησιμοποιώντας το πρότυπο μοντέλο, [65]. Οι χρονοσειρές έχουν χρησιμοποιηθεί για δεκαετίες σε διάφορους τομείς όπως η ψηφιακή επεξεργασία σήματος, η οικονομία καθώς η πρόβλεψη ηλεκτρικού φορτίου.

Οι προσεγγίσεις με τη βοήθεια χρονοσειρών, Box και Jekins ([71]) εντοπίζονται μεταξύ των παλαιότερων μεθόδων που εφαρμόζονται για την πρόβλεψη ηλεκτρικού φορτίου. Αν και τα μοτίβα κατανάλωσης που παρουσιάζει η ζήτηση ενέργειας είναι δύσκολο να αναλυθούν και να προβλεφθούν με την άμεση εφαρμογή τέτοιων μεθόδων, ωστόσο οι χρονοσειρές παραμένουν αρκετά δημοφιλείς όσον αφορά στη βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φορτίου (STLF) [72]. Οι μέθοδοι των χρονοσειρών μπορούν να διακριθούν σε διάφορα επίπεδα, έτσι έχουμε μονοπαραγοντικές και πολυμεταβλητές μεθόδους. Οι πρώτες χρησιμοποιούνται κατά κύριο λόγο στην πολύ βραχυπρόθεσμη πρόβλεψη φορτίου, ενώ οι τελευταίες βρίσκουν εφαρμογή ανεξαρτήτως χρονικού ορίζοντα. Τα δύο βασικά βήματα που απαρτίζουν τη διαδικασία πρόβλεψης με τη βοήθεια χρονοσειρών είναι αρχικά η ανάπτυξη ενός μοντέλου με βάση τα ιστορικά δεδομένα κατανάλωσης κι έπειτα η πρόβλεψη του μελλοντικού φορτίου σύμφωνα με αυτό το μοντέλο. Στη συνέχεια παρουσιάζουμε τα βασικά μοντέλα χρονοσειρών που χρησιμοποιούνται στην πρόβλεψη ηλεκτρικού φορτίου.

4.1.7.1 Αυτοπαλινδρομικά Μοντέλα (AR)

Το αυτοπαλινδρομικό μοντέλο (Autoregressive Model) μπορεί να χρησιμοποιηθεί για τη διαμόρφωση του προφίλ του φορτίου, εάν υποθέσουμε ότι το φορτίο προκύπτει ως γραμμικός συνδυασμός των προηγούμενων τιμών του. Σύμφωνα με το Liu ([73]) η μαθηματική σχέση που περιγράφει το μοντέλο είναι η ακόλουθη:

$$\hat{L}_k = -\sum_{i=1}^{m} \alpha_{ik} L_{k-i} + w_k \tag{4.14}$$

όπου
$\hat{L}_k$: η προβλεπόμενη τιμή του φορτίου τη χρονική στιγμή k
w_k: τυχαία διαταραχή του φορτίου
α_i: άγνωστοι συντελεστές

Για την εύρεση των αγνώστων συντελεστών της εξίσωσης γίνεται χρήση της μεθόδου ελαχίστων τετραγώνων. Εφαρμογές των αυτοπαλινδρομικών μοντέλων παρουσιάζονται στα ([74], [75]).

4.1.7.2 Αυτοπαλινδρομικά Μοντέλα Κινητού Μέσου Όρου (ARMA)

Τα αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου (Autoregressive Moving Average

Model) χρησιμοποιούνται συνήθως για την προσέγγιση στάσιμων χρονοσειρών φορτίου, δηλαδή χρονοσειρών των οποίων η μέση τιμή, η διακύμανση και η συνάρτηση αυτοσυσχέτισης διατηρούνται σταθερές σε όλη τη διάρκεια του χρόνου. Το γεγονός αυτό περιορίζει αρκετά το πεδίο εφαρμογής τους [72]. Η μαθηματική σχέση που περιγράφει τα ARMA μοντέλα (p, q) είναι η ακόλουθη:

$$y(t) = \varphi_1 y(t-1) + \cdots + \varphi_p y(t-p) + a(t) - \theta_1 \alpha(t-1) - \cdots - \theta_q \alpha(t-q) \quad (4.15)$$

όπου
$y(t)$: το ηλεκτρικό φορτίο
$a(t)$: ο λευκός θόρυβος για τη χρονική στιγμή t
θ_i: ο συντελεστής βάρους του θορύβου πριν i χρονικά βήματα
φ_i: ο συντελεστής βάρους του φορτίου πριν i χρονικά βήματα

Η τρέχουσα τιμή της χρονοσειράς εκφράζεται γραμμικά σε όρους των τιμών αυτής σε προηγούμενες περιόδους $[y(t-1), (y(t-2), \dots]$ και σε όρους των προηγούμενων τιμών του λευκού θορύβου $[\alpha(t), \alpha(t-1), \dots]$. Η διαδικασία αποτελείται από τρία βασικά βήματα, τον προσδιορισμό της p τάξης του αυτοπαλινδρομικού μέρους (AR −part), της τάξης q του κινητού μέσου όρου (MA −part) και την εκτίμηση των συντελεστών [76]. Για τον υπολογισμό των συντελεστών συνήθως χρησιμοποιείται κάποια παραλλαγή της μεθόδου ελαχίστων τετραγώνων είτε προσέγγιση της μέγιστης πιθανοφάνειας. Διάφορες περιπτώσεις πρόβλεψης του ηλεκτρικού φορτίου με τη χρήση παλινδρομικών μοντέλων κινητού μέσου όρου παρουσιάζονται στα ([77], [78], [79]).

4.1.7.3 Ολοκληρωμένα Αυτοπαλινδρομικά Μοντέλα Κ.Μ.Ο (ARIMA)

Οι ολοκληρωμένες αυτοπαλινδρομικές μέθοδοι κινητού μέσου όρου (Autoregressive Integrated Moving Average), Box και Jekins ([71]), αποτελούν επέκταση των μοντέλων ARMA, οι οποίες χρησιμοποιούν μεθόδους διαφορισμού προκειμένου να αντιμετωπίσουν περιπτώσεις μη-στασιμότητας. Όταν η διαδικασία είναι δυναμική (μη-στάσιμη) απαιτείται μετασχηματισμός της σειράς σε στάσιμη μορφή [72]. Αυτό επιτυγχάνεται μέσω διαδικασίας διαφορισμού. Οι μαθηματικές σχέσεις που περιγράφουν τα μοντέλα ARIMA (p, d, q) είναι οι ακόλουθες:

$$\varphi(B) \cdot \nabla^d y(t) = \theta(B) \cdot \alpha(t) \quad (4.16)$$

$$\nabla^d y(t) = (1-B)^d \cdot y(t)$$

όπου
$y(t)$: το ηλεκτρικό φορτίο
$\alpha(t)$: ο λευκός θόρυβος
θ_i: ο συντελεστής βάρους του θορύβου πριν i χρονικά βήματα
φ_i: ο συντελεστής βάρους του φορτίου πριν i χρονικά βήματα
d: ο βαθμός διαφορισμού

Τα μοντέλα ARIMA είναι μαθηματικά στοχαστικά μοντέλα τα οποία μπορούν να περιγράψουν τη διαχρονική εξέλιξη μιας χρονοσειράς που εξαρτάται από μη ντετερμινιστικούς παράγοντες. Για το λόγο αυτό είναι αρκετά δημοφιλή σε περιπτώσεις όπου εμπεριέχονται φυσικά μεγέθη τα οποία δε γνωρίζουμε απόλυτα. Βασίζονται στην παραδοχή της αλληλεξάρτησης μεταξύ των τιμών που λαμβάνει η χρονοσειρά για τις διάφορες χρονικές στιγμές.

Σύμφωνα με το [80] τα μονοδιάστατα μοντέλα ARIMA χρησιμοποιούνται συχνά ως μοντέλα αναφοράς για τη βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φορτίου (STLF). Στο [81] παρουσιάζεται μια βασική παραλλαγή των μοντέλων ARIMA που μας δίνει τη δυνατότητα να εκφράσουμε δύο ή και περισσότερες εποχικότητες. Επίσης στο [82] έχει αναπτυχθεί ένα μοντέλο ARIMA για την πρόβλεψη του φορτίου σε πραγματικό χρόνο, λαμβάνοντας υπόψη τη μετεωρολογική επίδραση ως επεξηγηματική μεταβλητή. Άλλες εφαρμογές αυτών των μοντέλων παρουσιάζονται στα ([83], [84]).

4.1.8 Προσέγγιση της παρόμοιας ημέρας

Η προσέγγιση αυτή (Similar Day Approach) βασίζεται στην αναζήτηση ιστορικών δεδομένων τα τελευταία ένα, δύο ή τρία χρόνια, για την εύρεση ημερών οι οποίες παρουσιάζουν παρόμοια χαρακτηριστικά με την υπό πρόβλεψη ημέρα [34]. Ως παρόμοια χαρακτηριστικά θεωρούνται κυρίως η ημερομηνία, το είδος της ημέρας (εργάσιμη ημέρα ή σαββατοκύριακο) και οι καιρικές συνθήκες. Σε αυτήν την περίπτωση η τιμή του φορτίου μιας παρόμοιας ημέρας θεωρείται ως πρόβλεψη. Συνήθως για καλύτερα αποτελέσματα αντί για το φορτίο μιας μεμονωμένης παρόμοιας ημέρας, η πρόβλεψη επιτυγχάνεται μέσω γραμμικού συνδυασμού ή διαδικασιών παλινδρόμησης σε ένα σύνολο παρόμοιων ημερών. Οι συντελεστές τάσης μπορούν να χρησιμοποιηθούν για παρόμοιες μέρες τα προηγούμενα χρόνια.

4.1.9 Λοιπές Παλινδρομικές και Μη Μέθοδοι

Εκτός από τις μεθόδους που αναφέρθηκαν παραπάνω, υπάρχει ένας πλήθος παραλλαγών αυτών, όπου για παράδειγμα χρησιμοποιούν διάφορες τεχνικές μετασχηματισμού για την παραγωγή προβλέψεων. Μια εφαρμογή παλινδρομικής μεθόδου πρόβλεψης συνδυασμένη με τεχνική μετασχηματισμού παρουσιάζεται στο [85] όπου μέσω πολυωνυμικών συναρτήσεων γίνεται η αναπαράσταση της μέγιστης και μέσης θερμοκρασίας των δύο τελευταίων ημερών μαζί με τη σχετική υγρασία που χρησιμοποιούνται ως είσοδοι στο μοντέλο. Επιπλέον μια προσέγγιση βραχυπρόθεσμης πρόβλεψης φορτίου σε μη στάσιμες χρονοσειρές παρουσιάζεται στο [86] όπου χρησιμοποιείται ένα συνεχές αυτοπαλινδρομικό μοντέλο με μετασχηματισμό τύπου Hilbert. Ακόμα υπάρχουν πλήθος μη παραμετρικών μεθόδων, οι οποίες χρησιμοποιούν εκτιμητές από των κλάδο των πιθανοτήτων, ντετερμινιστικές προσεγγίσεις προβλήματος ([87]) αλλά και διάφορες υβριδικές μορφές των κλασσικών μεθόδων [88].

4.1.10 Μέθοδος Επαναληπτικών Επανασταθμισμένων Ελαχίστων Τετραγώνων

Η μέθοδος αυτή (Iterative Reweighted Least-Squares) έγινε γνωστή από τους Mbamalu και El-Hawary (1992) όπου χρησιμοποίησαν μια διαδικασία για να προσδιορίσουν τη διάταξη του μοντέλου και τις παραμέτρους. Κατά τη διαδικασία αυτή χρησιμοποιούνται συναρτήσεις αυτοσυσχέτισης στα ιστορικά δεδομένα φορτίου με στόχο την εύρεση ενός αντίστοιχου μοντέλου της δυναμικής του φορτίου. Όσον αφορά στο πρόβλημα εκτίμησης παραμέτρων, η γραμμική σχέση που εκφράζει τη μέθοδο αυτή είναι η ακόλουθη:

$$Y = X \cdot \beta + \varepsilon \tag{4.17}$$

όπου
Y: είναι ένα διάνυσμα $(n * 1)$ των παρατηρήσεων
X: είναι μία μήτρα $(n * p)$ των γνωστών συντελεστών
β: είναι ένα διάνυσμα $(p * 1)$ των αγνώστων παραμέτρων
ε: είναι ένα διάνυσμα $(n * 1)$ των τυχαίων σφαλμάτων

Η τιμή του β μπορεί να υπολογιστεί με τη χρήση επαναληπτικών μεθόδων. Στο [89] παρουσιάζονται εναλλακτικές προσεγγίσεις με τη χρήση μεθόδων ελαχίστων τετραγώνων.

4.2 Μοντέλα Τεχνητής Νοημοσύνης

Με τον όρο τεχνητή νοημοσύνη (Artificial Intelligence) ορίζουμε γενικά μια αναλυτική διαδικασία η οποία βασίζεται σε υπολογιστικές μηχανές και εμφανίζει συμπεριφορά και ενέργειες που θεωρούνται "έξυπνες" από ανθρώπινους παρατηρητές. Προσπαθεί δηλαδή να μιμηθεί την ανθρώπινη πορεία σκέψης, συμπεριλαμβανομένων και στοιχείων λογικής και βελτιστοποίησης. Βασικό στόχος των μεθόδων τεχνητής νοημοσύνης είναι να βοηθήσουν στην οργάνωση και την παροχή πληροφοριών που απαιτούνται κατά τη διαδικασία λήψης αποφάσεων με τρόπο που να βελτιώνουν τη συνολική απόδοση. Στη συνέχεια θα ασχοληθούμε με τα πιο ευρέως χρησιμοποιούμενα συστήματα AI σε πρόβλεψη.

Τα τεχνητά νευρωνικά δίκτυα (Neural Networks) χρησιμοποιούνται στον τομέα της ηλεκτρικής ενέργειας ήδη από τη δεκαετία του 1990 και διαδραματίζουν σημαντικό ρόλο αφού επιτρέπουν την εισαγωγή μη γραμμικών σχέσεων [90]. Μια διαφορετική προσέγγιση που βρίσκει εφαρμογή στη βραχυπρόθεσμη πρόβλεψη βασίζεται στα εμπειρικά συστήματα (expert systems) τα οποία τροφοδοτούν τους κανόνες τους μέσα από μια διαδικασία πληροφορικής προκειμένου παράγουν προβλέψεις χωρίς την ανθρώπινη υποστήριξη. Πρόβλεψη του φορτίου ηλεκτρικής ενέργειας μπορεί ακόμα να επιτευχθεί με τη χρήση μεθόδων ασαφούς λογικής (fuzzy logic) που οδηγούν σε ικανοποιητικά αποτελέσματα χωρίς την ανάγκη μαθηματικής αναπαράστασης των παραμέτρων εισόδου και εξόδου ([91], [92], [93]). Τέλος αξίζει να αναφέρουμε τις μηχανές διανυσμάτων υποστήριξης (Support Vector Machines). Πρόκειται για μια πολύ ισχυρή τεχνική για την επίλυση προβλημάτων ταξινόμησης και παλινδρόμησης **Error! Reference source not found.**.

4.2.1 Τεχνητά Νευρωνικά Δίκτυα (ANN)

Τα τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks) είναι μαθηματικά ή υπολογιστικά μοντέλα βασισμένα στη δομή και τις αρχές λειτουργίας των βιολογικών νευρώνων. Αποτελούνται από διασυνδεδεμένες ομάδες τεχνητών νευρώνων και επεξεργάζονται τις πληροφορίες χρησιμοποιώντας μια συνδετική προσέγγιση υπολογισμού. Ένα ANN ως επί το πλείστον είναι ένα προσαρμοζόμενο σύστημα, το οποίο αλλάζει τη δομή του με βάση τις εσωτερικές και εξωτερικές πληροφορίες που εισέρχονται στο σύστημα κατά της φάση της εκπαίδευσής του. Ο τρόπος με τον οποίο το δίκτυο λαμβάνει τη γνώση από το περιβάλλον μέσω μιας διαδικασίας μάθησης καθώς και η ισχύς των συνδέσεων μεταξύ των νευρώνων (συναπτικό βάρος) που χρησιμοποιούνται για την αποθήκευση της γνώσης, αποτελούν δύο κοινά γνωρίσματα του δικτύου με τον ανθρώπινο εγκέφαλο.

Ο τρόπος λειτουργίας των νευρωνικών δικτύων δεν είναι σειριακός αλλά μοιάζει περισσότερο με παράλληλο καθώς μια εργασία μοιράζεται στους επιμέρους νευρώνες του δικτύου. Λόγω της κατανεμημένης αρχιτεκτονικής των δικτύων που αποτελούνται από απλές μονάδες επεξεργασίας, τα ANN χαρακτηρίζονται ως συστήματα παράλληλων κατανεμημένων διεργασιών (parallel distributed processing). Σε αυτό το σημείο πρέπει να τονιστεί ότι η αρχιτεκτονική των νευρωνικών δικτύων διαφέρει από αυτήν των παράλληλων επεξεργαστών. Οι τελευταίοι διαθέτουν συνήθως περισσότερους επεξεργαστές από τις διασυνδέσεις μεταξύ τους και όσον αφορά στην πολυπλοκότητα, ακολουθούν αυτή των μηχανών von Neumann. Αντίθετα οι απλοί επεξεργαστές των νευρωνικών δικτύων έχουν μεγαλύτερο αριθμό διασυνδέσεων συγκριτικά με τους νευρώνες, αφού κάθε νευρώνας διαθέτει πολλές συνδέσεις. Επίσης στην περίπτωση των ANN οι νευρώνες δε χρειάζονται συγχρονισμό και λειτουργούν ανεξάρτητα ο ένας από τον άλλο.

Χάρις στη μοντελοποίηση των τεχνητών νευρωνικών δικτύων με βάση των ανθρώπινο εγκέφαλο, είναι σε θέση να εμφανίζουν κάποια χρήσιμα χαρακτηριστικά τα οποία απουσιάζουν από αρχιτεκτονικές των υπολογιστών τύπου von Neumann . Η σημαντικότητά τους έγκειται στις παρακάτω ιδιότητες ([17], [95], [96]):

- Μη γραμμικότητα
- Προσαρμοστικότητα
- Ανοχή σε σφάλματα
- Κατανεμημένη επεξεργασία
- Ικανότητα μάθησης
- Ικανότητα γενίκευσης
- Αντιστοίχηση εισόδου-εξόδου

Όπως αναφέραμε παραπάνω, οι νευρώνες αποτελούν τις μονάδες επεξεργασίας της πληροφορίας του δικτύου. Το μοντέλο ενός νευρώνα αποτελείται από ένα σύνολο διασυνδέσεων όπου κάθε διασύνδεση χαρακτηρίζεται από το δικό της βάρος, γνωστό ως “συναπτικό βάρος”. Έχει πολλές εισόδους αλλά μια μόνο έξοδο. Τέλος η σχέση της εξόδου με τις εισόδους και τους συντελεστές βάρους καθορίζεται από μια συνάρτηση μεταφοράς.

Η εφαρμογή των νευρωνικών δικτύων σε προβλήματα πρόβλεψης ηλεκτρικού φορτίου προϋποθέτει την επιλογή της κατάλληλης μεθόδου ρύθμισης των βαρών (πχ. back propagation, Boltzmann machine κ.α.). Επίσης πρέπει να καθοριστεί ο αριθμός και η σύνδεση των στρωμάτων (layers) και η μορφή (δυαδική ή συνεχής) που θα χρησιμοποιηθεί στις εισόδους, τις εξόδους και εσωτερικά. Η πιο δημοφιλής επιλογή στη βραχυπρόθεσμη πρόβλεψη φορτίου είναι τα πολυεπίπεδα δίκτυα Perceptron (multilayer perceptron). Αυτό οφείλεται στην ικανότητα των δικτύων MLPs να αναγνωρίζουν πολύπλοκες σχέσεις μεταξύ των προτύπων εισόδου και εξόδου, στοιχείο που δεν επιτυγχάνεται εύκολα από συμβατικές μεθόδους [97]. Τα ιστορικά δεδομένα αποτελούν συνήθως την είσοδο των δικτύων, βάσει των οποίων γίνεται και η εκπαίδευση του συστήματος. Εκτενέστερη αναφορά στους βασικούς τύπους τον νευρωνικών δικτύων και στα πλεονεκτήματα τους, παρουσιάζεται στην παράγραφο 5.2.

4.2.2 Μηχανές Διανυσμάτων Υποστήριξης (SVM)

Οι μηχανές διανυσμάτων υποστήριξης (Support Vector Machines) προτάθηκαν από το Vapnik ([94]) και αποτελούν μια ισχυρή τεχνική για την επίλυση προβλημάτων ταξινόμησης και παλινδρόμησης. Ανήκουν στη μηχανική μάθηση (Machine Learning) και βασίζονται στη στατιστική θεωρία μάθησης (SLT – statistical learning theory). Σε αντίθεση με τα νευρωνικά δίκτυα, τα οποία προσπαθούν να καθορίσουν τις σύνθετες λειτουργίες του χώρου των παραμέτρων εισόδου, οι μηχανές διανυσμάτων υποστήριξης χρησιμοποιούν μη-γραμμική χαρτογράφηση του χώρου εισόδου σε ένα πολυδιάστατο χώρο χαρακτηριστικών (feature space), στον οποίο κατασκευάζεται ένα βέλτιστο υπερεπίπεδο (hyperplane). Καθώς αυτό το υπερεπίπεδο ανταποκρίνεται σε μη-γραμμική επιφάνεια στο χώρο εισόδου, οι μηχανές διανυσμάτων υποστήριξης χρησιμοποιούν απλές συναρτήσεις για να δημιουργήσουν γραμμικά όρια απόφασης στο νέο χώρο. Επομένως το πρόβλημα της επιλογής μιας αρχιτεκτονικής για το νευρωνικό δίκτυο, αντικαθίσταται σε αυτήν την περίπτωση με την επιλογή ενός κατάλληλου τύπου μηχανής εκμάθησης (kernel) για τα διανύσματα υποστήριξης [99].

Τα SVMs αντικαθιστούν την αρχή της ελαχιστοποίησης του εμπειρικού ρίσκου (ERM – empirical risk minimization) το οποίο εφαρμόζεται παραδοσιακά στα νευρωνικά δίκτυα, με την αρχή ελαχιστοποίησης του διαρθρωτικού ρίσκου (SRM – structural risk minimization) [100]. Η πιο σημαντική έννοια της μεθόδου SRM είναι η εφαρμογή της ελαχιστοποίησης ενός άνω ορίου στο σφάλμα γενίκευσης και όχι η ελαχιστοποίηση του σφάλματος εκπαίδευσης. Με βάση λοιπόν αυτήν την αρχή, ένα SVM είναι ισοδύναμο με την επίλυση ενός γραμμικά περιορισμένου προβλήματος δυαδικού προγραμματισμού. Αρχικά οι μηχανές διανυσμάτων υποστήριξης αναπτύχθηκαν για την επίλυση προβλημάτων ταξινόμησης με πολύ καλά αποτελέσματα ([101], [102], [103]). Έπειτα με την εισαγωγή της συνάρτησης απωλειών του Vapnik (ε-insensitive loss function), τα SVM επεκτάθηκαν και στην επίλυση προβλημάτων παλινδρόμησης, έτσι προέκυψε η Παλινδρόμηση Διανυσμάτων Υποστήριξης (SVR – support vector regression) [104].

Όπως περιγράψαμε, τα SVM αποτελούν μια επιτυχημένη μέθοδο πρόβλεψης η οποία βασίζεται στη χρήση πολυδιάστατων χώρων χαρακτηριστικών (feature space) και τιμωρεί την επακόλουθη πολυπλοκότητα με τη βοήθεια ενός όρου ποινής (penalty term) ο οποίος εμπεριέχεται στο χρησιμοποιούμενο σφάλμα. Το πιο γνωστό μοντέλο SVR είναι η γραμμική παλινδρόμηση. Η μαθηματική συνάρτηση εκτίμησης του γραμμικού μοντέλου σε αυτό το χώρο χαρακτηριστικών δίνεται από την ακόλουθη σχέση:

$$f(x) = \langle w, \varphi(x) \rangle + b \tag{4.18}$$

όπου
$\varphi(x)$: δηλώνει τον πολυδιάστατο χώρο χαρακτηριστικών, ο οποίος έχει χαρτογραφηθεί μη γραμμικά από το χώρο εισόδου
w: περιέχει τους συντελεστές που πρέπει να εκτιμηθούν από τα δεδομένα
b: είναι μια πραγματική σταθερά

Ο στόχος της μεθόδου είναι η ελαχιστοποίηση της συνάρτησης ρίσκου που περιγράφεται ως εξής, [105]:

$$\min_{w,b,\xi_i,\xi_i^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{4.19}$$

υπό τους περιορισμούς

$$\begin{cases} y_i - \langle w, \varphi(x_i) \rangle + b \leq \varepsilon + \xi_i^* \\ \langle w, \varphi(x_i) \rangle + b - y_i \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

όπου
x_i: έχει χαρτογραφηθεί σε έναν υψηλότερων διαστάσεων χώρο μέσω της συνάρτησης φ
ξ_i: το μικρότερο σφάλμα εκπαίδευσης εντός της παραμέτρου ε
ξ_i^*: το ανώτερο σφάλμα εκπαίδευσης εντός της παραμέτρου ε
$\frac{1}{2}\|w\|^2$: ο όρος τακτοποίησης (regularization term)
C: η σταθερά τακτοποίησης (regularization constant), η οποία καθορίζει την εξισορρόπηση (tradeoff) μεταξύ της μείωσης της κλίσης για την f και της ακρίβειας κατά τη λήψη των δεδομένων.

Στο πλαίσιο της παρούσας διπλωματικής στόχος είναι η ανάδειξη του βέλτιστου μοντέλου παλινδρόμησης SVR για τη βραχυπρόθεσμη πρόβλεψη του ηλεκτρικού φορτίου. Περαιτέρω ανάλυση σχετικά με τη μέθοδο SVR, τις αρχές στις οποίες βασίζεται και τους διαφορετικούς τύπους μηχανών εκμάθησης (kernel) παρουσιάζεται στην παράγραφο 5.3.

4.2.3 Ασαφής Λογική

Η ασαφής λογική (Fuzzy Logic) αποτελεί μια γενίκευση της δυαδικής λογικής (Boolean Logic) που χρησιμοποιείται για τη σχεδίαση ψηφιακών κυκλωμάτων. Η είσοδος σύμφωνα με τη δυαδική λογική λαμβάνει τιμές “0” ή “1”. Κατά την ασαφή λογική μια είσοδος είναι

συνδεδεμένη με ένα εύρος τιμών. Έτσι για παράδειγμα το φορτίο ενός μετασχηματιστή μπορεί να είναι "χαμηλό", "μέσο" ή "υψηλό". Η μέθοδος αυτή αποτελεί ουσιαστικά μια προσπάθεια αντιγραφής και αφομοίωσης του ανθρώπινου τρόπου σκέψης, ο ποίος δεν καθορίζεται με ένα απλό "ναι" ή "όχι" αλλά εμπεριέχει στοιχεία ασάφειας και αοριστίας. Η ασαφής λογική επιτρέπει στον ενδιαφερόμενο να συμπεράνει (με λογικό τρόπο) εξόδους από ασαφείς εισόδους και επομένως αποτελεί μέθοδο χαρτογράφησης των εισόδων σε εξόδους.

Είναι γνωστό ότι ένα σύστημα ασαφούς λογικής με κεντροειδή αποσαφοποίηση (όπου επιστρέφει δηλαδή το κέντρο της περιοχής κάτω από την καμπύλη) μπορεί να αναγνωρίσει και να προσεγγίσει οποιοδήποτε άγνωστο δυναμικό σύστημα, στην περίπτωσή μας είναι το ηλεκτρικό φορτίο. Σύμφωνα με το [73] ένα σύστημα ασαφούς λογικής έχει μεγάλη ικανότητα στην περιγραφή ομοιοτήτων από μεγάλο όγκο δεδομένων. Οι ομοιότητες αυτές στα δεδομένα εισόδου $(L_{-i} - L_0)$ μπορούν να εντοπιστούν από τις διαφορές πρώτης (V_k) και δεύτερης τάξης (A_k). Το μοντέλο περιγράφεται μαθηματικά από τις ακόλουθες σχέσεις:

$$V_k = \frac{(L_k - L_{k-1})}{T} \tag{4.20}$$

$$A_k = \frac{(V_k - V_{k-1})}{T}$$

Η διαδικασία πρόβλεψης με ασαφή λογική αποτελείται από δύο στάδια, την εκπαίδευση του συστήματος και την παραγωγή των προβλέψεων. Αρχικά στο στάδιο της κατάρτισης, με βάση τα ιστορικά δεδομένα γίνεται εκπαίδευση των $2m$ −εισόδων και $2n$ −εξόδων του συστήματος για την παραγωγή μιας βάσης προτύπων και ασαφών κανόνων που προκύπτουν από τις διαφορές πρώτης και δεύτερης τάξης των δεδομένων. Επόμενο βήμα είναι η σύνδεση με έναν ελεγκτή για την πρόβλεψη των αλλαγών του φορτίου. Εάν βρεθεί κάποιο μοτίβο με υψηλή πιθανότητα που να ταιριάζει, τότε δημιουργείται ένα μοτίβο εξόδου μέσω κεντροειδούς αποσαφοποίησης.

Βασικό πλεονέκτημα της πρόβλεψης με τη βοήθεια ασαφούς λογικής είναι η απουσία ανάγκης για ακριβείς εισόδους (χωρίς θόρυβο) ή ενός μαθηματικού μοντέλου χαρτογράφησης των εισόδων και εξόδων. Σύμφωνα με αυτούς τους κανόνες, ένα κατάλληλα σχεδιασμένο σύστημα ασαφούς λογικής θα μπορούσε να είναι πολύ ισχυρό κατά την παραγωγή προβλέψεων ηλεκτρικού φορτίου. Ωστόσο σε αρκετές περιπτώσεις είναι αναγκαία η ακριβής πρόβλεψη των εξόδων του συστήματος. Ύστερα από λογική επεξεργασία των εισόδων, γίνεται χρήση μιας μεθόδου αποσαφοποίησης για την παραγωγή εξόδων με ακρίβεια. Μερικές βασικές εφαρμογές αυτής της λογικής στην πρόβλεψη φορτίου παρουσιάζονται στα ([91], [92], [93]).

Έχουν αναπτυχθεί διάφορες προσεγγίσεις που αντιπροσωπεύουν τα μοντέλα ασαφούς λογικής στην πρόβλεψη φορτίου. Ο Hsu (1992) στο [106] παρουσίασε ένα εξειδικευμένο σύστημα ασαφούς λογικής για την ενημέρωση της συνάρτησης πρόβλεψης. Οι Al-Anbucky κ.α. (1995) στο [107] ασχολήθηκαν με την εφαρμογή μιας ασαφούς προσέγγισης προκειμένου να παρέχουν ένα διαρθρωτικό πλαίσιο για την αναπαράσταση και το χειρισμό δεδομένων σχετικά με την πρόβλεψη δεσμεύσεων ενέργειας. Οι Srinivasan κ.α. (1992) στο

[108] χρησιμοποίησαν μια υβριδική τεχνική ασαφούς λογικής και νευρωνικών δικτύων. Το μοντέλο αυτό ενισχύθηκε στη συνέχεια από τους Dash κ.α. (1995) στο [109] ώστε να προβλέπει με ακρίβεια τις καθημερινές ημέρες, τις γενικές αργίες, τις περιόδους διακοπών κ.α. Ακόμα οι Mori και Kobayashi (1996) χρησιμοποίησαν μέθοδο ασαφούς λογικής για να αναπτύξουν ένα μη-γραμμικό μοντέλο βελτιστοποίησης για βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φορτίου, με στόχο την ελαχιστοποίηση του σφάλματος [110]. Για την εύρεση της βέλτιστης αυτής λύσης έγινε χρήση της μεθόδου απότομης καθόδου. Στα ([111], [112], [113]) παρουσιάζονται μερικές ακόμα προσεγγίσεις με τη χρήση ασαφούς λογικής.

4.2.4 Έμπειρα Συστήματα που βασίζονται στη γνώση

Τα εμπειρικά αυτά συστήματα (Knowledge-based Expert Systems) είναι τεχνικές που έχουν προκύψει ως αποτέλεσμα της προόδου στον τομέα τεχνητής νοημοσύνης. Το έμπειρο σύστημα αναφέρεται σε ένα υπολογιστικό πρόγραμμα το οποίο έχει την ικανότητα να εξηγεί, να αιτιολογεί και να διευρύνει τη γνωστική του βάση όσο αυξάνονται οι διαθέσιμες πληροφορίες που κατέχει. Ένα εξειδικευμένο σύστημα ενσωματώνει κανόνες και διαδικασίες που χρησιμοποιούνται από εμπειρογνώμονες στο συγκεκριμένο τομέα για τη δημιουργία κατάλληλου λογισμικού το οποίο είναι ικανό να παράγει προβλέψεις αυτόματα, χωρίς την ανθρώπινη παρέμβαση. Για τη δημιουργία ενός τέτοιου μοντέλου, εξάγεται γνώση από εμπειρογνώμονες στον τομέα της πρόβλεψης ηλεκτρικού φορτίου, από τη συνιστώσα βασικών γνώσεων του ειδικού συστήματος. Με τη σειρά της η γνώση αυτή αναπαριστάται ως γεγονότα και κανόνες της λογικής “Εάν-Τότε” (If-Then rules) και αποτελείται από το πλήθος των σχέσεων που περιγράφουν τις αλλαγές στο φορτίο αλλά και τις παραμέτρους που επηρεάζουν τη χρήση της ηλεκτρικής ενέργειας. Μερικοί από τους κανόνες που προκύπτουν δεν αλλάζουν με την πάροδο του χρόνου, ενώ άλλοι απαιτούν διαρκή ενημέρωση.

Οι πρώτες εφαρμογές των εξειδικευμένων συστημάτων εντοπίζονται το 1960 και αφορούν σε γεωλογικές αναζητήσεις και σχεδιασμό με τη χρήση ηλεκτρονικών υπολογιστών. Απαραίτητη προϋπόθεση για την καλύτερη απόδοση αυτών των συστημάτων είναι η συνεργασία ανάμεσα σε ειδικούς στον τομέα της πρόβλεψης φορτίου και σε προγραμματιστές λογισμικού. Με αυτόν τον τρόπο μπορεί να υλοποιηθεί ένα λογισμικό που ενσωματώνει τις γνώσεις των ειδικών και τις κωδικοποιεί σε κανόνες. Η διερεύνηση των συντακτικών σχέσεων μεταξύ του ηλεκτρικού φορτίου και συνιστωσών όπως η θερμοκρασία, οι καιρικές συνθήκες, το είδος της ημέρας κ.α. οδήγησε στην ανάπτυξη διάφορων προσεγγίσεων.

Οι Ho et al. (1990) ανέπτυξαν ένα εξειδικευμένο σύστημα βασισμένο στη γνώση, όπου με τη χρήση ωριαίων παρατηρήσεων για την κατανάλωση φορτίου καθορίστηκαν έντεκα διαφορετικοί τύποι ημερών [114]. Ο αλγόριθμος αυτός που περιλάμβανε και παραμέτρους για τον καιρό αποδείχτηκε πιο αποτελεσματικός από τη συμβατική μέθοδο των Box-Jenkins. Σε μια άλλη προσέγγιση οι Brown κ.α. (1999) συνδύασαν την υπάρχουσα γνώση του συστήματος και τα πρότυπα του φορτίου για διαφορετικούς ορίζοντες με στόχο την ανάπτυξη διαφορετικών σεναρίων εξέλιξης της κατανάλωσης. Επιπλέον υπάρχουν και

υβριδικές μέθοδοι οι οποίες συνδυάζουν τα έμπειρα συστήματα με άλλα μοντέλα πρόβλεψης, όπως στο [115] με τη χρήση μεθόδου ασαφούς λογικής. Εναλλακτικές προσεγγίσεις των εξειδικευμένων συστημάτων παρουσιάζονται στα ([116], [117]).

4.2.5 Παλινδρόμηση Κ Κοντινότερων Γειτόνων (KNN Regression)

Η μέθοδος παλινδρόμησης των K κοντινότερων γειτόνων (K Nearest Neighbor Regression) είναι μια μη παραμετρική μέθοδος πρόβλεψης, η οποία έχει μη παραμετρικό, μικρό ποσοστό σφάλματος και καλή κατανομή σφάλματος [118]. Η αρχή πίσω από αυτήν τη μέθοδο, είναι η εύρεση ενός προκαθορισμένου αριθμού δειγμάτων εκπαίδευσης, σε πολύ κοντινή απόσταση από το νέο σημείο και ο προσδιορισμός της κατηγορίας αυτών. Ως πρόβλεψη του σημείου λαμβάνεται το αποτέλεσμα μιας διαδικασίας μέσου όρου, όπου η τιμή του προκύπτει από τη μέση τιμή των K κοντινότερων σημείων (γειτόνων). Ο αριθμός των δειγμάτων μπορεί να είναι σταθερά καθορισμένος από το χρήστη ή να διαφέρει ανάλογα με την τοπική πυκνότητα των σημείων.

Πιο αναλυτικά για ένα δοσμένο σημείο υπολογίζεται η Ευκλείδεια απόσταση του από τα υπόλοιπα σημεία εκπαίδευσης. Στη συνέχεια επιλέγονται τα πλησιέστερα δεδομένα κατάρτισης K και λαμβάνεται ως πρόβλεψη για το σημείο, ο μέσος όρος των εξόδων για αυτά τα δεδομένα [119]. Στην πράξη η διαδικασία αυτή μπορεί να βελτιωθεί με την προσθήκη βαρών στα Κ κοντινότερα δεδομένα κατάρτισης, όπου τα βάρη είναι ανάλογα της απόστασης από το ζητούμενο σημείο. Εάν υποθέσουμε το ζητούμενο σημείο x, η μαθηματική περιγραφή του μοντέλου δίνεται από την ακόλουθη σχέση:

$$\hat{Y} = \frac{1}{k} \sum_{m \in J(x)} Y_m \tag{4.21}$$

όπου
$J(x)$: το σύνολο των K κοντινότερο γειτόνων του σημείου x
Y_m: η έξοδος για κάθε σημείο κατάρτισης x_m

Όπως αντιλαμβανόμαστε από την παραπάνω σχέση, το K αποτελεί παράμετρο κλειδί για τη μέθοδο. Έτσι μια μεγάλη τιμή του K μπορεί να οδηγήσει σε πιο ομαλή προσαρμογή και ως εκ τούτου μικρότερη διακύμανση σε βάρος όμως της υψηλότερης προκατάληψης για τα δεδομένα. Παρά την απλότητα της μεθόδου K κοντινότερων γειτόνων, βρίσκει συχνά εφαρμογή σε προβλήματα ταξινόμησης και παλινδρόμησης. Ιδιαίτερα σε περιπτώσεις ταξινόμησης όπου τα όρια απόφασης είναι ακανόνιστα, η μη παραμετρική αυτή μέθοδος παρουσιάζει σημαντική επιτυχία.

4.2.6 Δένδρα Ταξινόμησης και Παλινδρόμησης (CART)

Η μέθοδος αυτή (Classification and Regression Trees) είναι ένα πρότυπο ταξινόμησης ή παλινδρόμησης το οποίο βασίζεται σε μια ιεραρχική κατάτμηση με τη μορφή δέντρου για

το χώρο εισόδου [120]. Το δέντρο απόφασης διευκρινίζει τη σχέση μεταξύ των μεταβλητών εισόδου και εξόδου σύμφωνα με τις συνθήκες διάσπασης και επιτρέπει την απεικόνιση του υπό μελέτη προβλήματος με τη μορφή δυαδικού δέντρου. Αυτό σημαίνει ότι το δέντρο απόφασης εντοπίζει κανόνες ενσωματωμένους στα δεδομένα εισόδου και εξόδου.

Τα δέντρα παλινδρόμησης χρησιμοποιούνται για συνεχή μεταβλητή εξόδου. Η διαδικασία περιλαμβάνει αρχικά τη διαίρεση του χώρου εισόδου σε τοπικές περιφέρειες, όπως αυτές προσδιορίζονται από μια σειρά αναδρομικών διασπάσεων. Κάθε δέντρο αποτελείται από εσωτερικούς κόμβους απόφασης και τερματικά φύλλα. Ο κόμβος ρίζας (root node) ορίζει την έναρξη του δέντρου παλινδρόμησης. Ο κόμβος διάσπασης (split node) βοηθά στο διαχωρισμό των μεταβλητών εισόδου σε δύο ομάδες σύμφωνα με κάποιο κανόνα και ο τερματικός κόμβος (terminal node) υποδεικνύει τα αποτελέσματα της ομαδοποίησης τους σε παρόμοιες ομάδες. Έτσι για ένα δοσμένο σημείο ξεκινά μια ακολουθία δοκιμών κατά μήκος των κόμβων απόφασης από τον κόμβο ρίζας, η οποία θα καθορίσει την πορεία μέσα στο δέντρο, μέχρι να φτάσει σε ένα κόμβο φύλλου. Σε αυτό το σημείο επιχειρείται η παραγωγή πρόβλεψης σύμφωνα με το σχετικό μοντέλο για αυτόν τον κόμβο.

Η κατασκευή ενός τέτοιου δένδρου στηρίζεται σε ένα σύνολο δεδομένων κατάρτισης και ξεκινά από τον κόμβο ρίζας. Στη συνέχεια γίνεται η επιλογή της μεταβλητής, η διάσπαση της οποίας θα οδηγήσει στη μεγαλύτερη μείωση του μέσου τετραγωνικού σφάλματος (MSE). Το σφάλμα ενός κόμβου είναι ο λόγος διακύμανσης όλων των δεδομένων εκμάθησης με τα δεδομένα αυτά που ανήκουν στον κόμβο [121]. Η μαθηματική εξίσωση που περιγράφει το σφάλμα είναι η ακόλουθη:

$$R(n) = \frac{V(n)}{V_0} \tag{4.22}$$

όπου
$R(n)$: το σφάλμα του κόμβου n
$V(n)$: η διακύμανση των δεδομένων εκπαίδευσης που ανήκουν στον κόμβο n
V_0: η διακύμανση όλων των δεδομένων εκπαίδευσης

Κατά τη διαδικασία εύρεσης των βέλτιστων συνθηκών διάσπασης για τα δεδομένα, το δένδρο παλινδρόμησης εξετάζει κάθε μεταβλητή εισόδου ως προσωρινή συνθήκη διάσπασης. Η μείωση του σφάλματος για κάθε συνθήκη υπολογίζεται σύμφωνα με την ακόλουθη σχέση.

$$\Delta R(s,n) = R(n_L) - R(n_R) \tag{4.23}$$

όπου
$\Delta R(s,n)$: η μείωση του σφάλματος σύμφωνα με την προσωρινή συνθήκη διάσπασης s
$R(n)$: το σφάλμα του μητρικού κόμβου
$R(n_L)$: το σφάλμα του αριστερού κόμβου
$R(n_R)$: το σφάλμα του δεξιού κόμβου

Οι καλύτερες συνθήκες διάσπασης είναι εκείνες που μεγιστοποιούν τη μείωση σφάλματος. Η διαδικασία αυτή επαναλαμβάνεται αναδρομικά έως ότου το σφάλμα φθάσει ένα αποδεκτό όριο. Μια τυπική πρακτική είναι η εκτέλεση ενός είδους “κλαδέματος” του

δέντρου μετά τη σχεδίασή του. Με αυτόν τον τρόπο επιτυγχάνεται η απομάκρυνση των αναποτελεσματικών κόμβων και η μείωση της πολυπλοκότητας του μοντέλου.

4.2.7 Αυτοπαλινδρομικά Μοντέλα Κ.Μ.Ο με Εξωγενή Μεταβλητή (ARMAX)

Ο προσδιορισμός ενός τέτοιου αυτοπαλινδρομικού μοντέλου (Autoregressive Moving Average with Exogenous variable) για την πρόβλεψη της ενεργειακής ζήτησης μπορεί να επιτευχθεί με δύο βασικές προσεγγίσεις. Είτε με τη βοήθεια ενός γενετικού αλγορίθμου (GA – Genetic Algorithm) είτε μέσω προσέγγισης εξελικτικού προγραμματισμού (EP – Evolutionary Programming). Η πρώτη τεχνική προσομοιώνει τη φυσική διαδικασία εξέλιξης και αποτελεί ένα στοχαστικό αλγόριθμο βελτιστοποίησης. Με τη βοήθεια του γενετικού αλγορίθμου (GA) είναι δυνατή η ταυτόχρονη αξιολόγηση πολλών σημείων στο χώρο αναζήτησης, ανεξάρτητα από το αν ο χώρος είναι διαφορίσιμος ή όχι, και η εύρεση ασυμπτώτων που συγκλίνουν στη βέλτιστη λύση. Αυτό έχει ως αποτέλεσμα τη βελτίωση ακρίβειας για το μοντέλο πρόβλεψης.

Πιο συγκεκριμένα το μοντέλο που στηρίζεται σε γενετικούς αλγόριθμους ακολουθεί την παρακάτω διαδικασία: Οι τιμές των εξεταζόμενων μεταβλητών στο γενετικό αλγόριθμο παριστάνονται ως ένα D −διάστατο διάνυσμα P. Ο αρχικός πληθυσμός των k μητρικών διανυσμάτων P_i, $i = 1, \dots, k$, παράγεται από ένα τυχαίο εύρος σε κάθε διάσταση. Έπειτα κάθε μητρικό διάνυσμα δημιουργεί έναν "απόγονο" συγχωνεύοντας ή τροποποιώντας τα διανύσματα του υπάρχοντος πληθυσμού. Με αυτόν τον τρόπο λαμβάνονται $2k$ νέα διανύσματα σε πληθυσμό. Η διαδικασία αυτή επαναλαμβάνεται μέχρις ότου η συνάρτηση f να μην επιδέχεται περαιτέρω βελτίωση είτε να έχει επιτευχθεί ο μέγιστος αριθμός γενεών. Η μαθηματική περιγραφή ενός μοντέλου ARMAX σύμφωνα με το [122] προκύπτει από την ακόλουθη σχέση:

$$A(q)y(t) = B(q)u(t) + C(q)e(t) \tag{4.24}$$

όπου
$y(t)$: το φορτίο για τη χρονική στιγμή t
$u(t)$: η θερμοκρασία για τη χρονική στιγμή t
$e(t)$: ο λευκός θόρυβος για τη χρονική στιγμή t
$A(q)$: η παράμετρος του αυτοπαλινδρομικού μέρους (AR)
$B(q)$: η παράμετρος του εξωγενούς μέρους (X)
$C(q)$: η παράμετρος του κινητού μέσου (MA)

Οι λύσεις της παραπάνω σχέσης με τη μέγιστη καταλληλότητα επιλέγονται ως δοκιμαστικά μοντέλα για την πρόβλεψη ηλεκτρικού φορτίου. Περαιτέρω εφαρμογές του μοντέλου ARMAX και των γενετικών αλγορίθμων περιγράφονται στα ([123], [124], [125]).

4.2.8 Γκαουσιανή Μέθοδος (GP)

Η διαδικασία Γκαουσιανής παλινδρόμησης (Gaussian Process) αποτελεί μη παραμετρική μέθοδο η οποία στηρίζεται στη μοντελοποίηση των παρατηρούμενων αποκρίσεων για τα διαφορετικά δεδομένα κατάρτισης ως μια πολυδιάστατη τυχαία μεταβλητή [126]. Υποθέτουμε μια a priori κατανομή για αυτές τις τιμές λειτουργίας, με στόχο την εξασφάλιση της ομαλής λειτουργίας. Πιο συγκεκριμένα όσο μικρότερη είναι η Ευκλείδεια απόσταση των διανυσμάτων εισόδου, η συσχέτιση μεταξύ των τιμών είναι υψηλή, ενώ αντίθετα φθίνει η συσχέτιση καθώς τα διανύσματα απομακρύνονται μεταξύ τους.

Θεωρούμε ως $V(X,X)$ τη μήτρα συνδιακύμανσης μεταξύ των τιμών λειτουργίας, όπου με X περιγράφονται τα διανύσματα εισόδου από το δεδομένα εκπαίδευσης. Αν θεωρήσουμε ως x_i το i-οστό διάνυσμα κατάρτισης, τότε η μαθηματική περιγραφή ενός τυπικού πίνακα διακύμανσης εκφράζεται από την ακόλουθη σχέση:

$$V(x_i, x_j) = \sigma_f^2 e^{-\frac{\|x_i - x_j\|^2}{2\gamma^2}} \tag{4.25}$$

Το αρχικό διάνυσμα των τιμών της συνάρτησης f ακολουθεί την πολυμεταβλητή Γκαουσιανή πυκνότητα, δηλαδή:

$$f \sim N_f\big(0, V(X,X)\big) \tag{4.26}$$

όπου
$N_f(\mu, \Sigma)$: η πολυμεταβλητή συνάρτηση πυκνότητας μιας μεταβλητής f με μέση τιμή μ και μήτρα συνδιακύμανσης Σ.

Στη συνέχεια προστίθενται στις τιμές λειτουργίας ορισμένες τιμές κατανεμημένου θορύβου με μηδενική μέση τιμή και τυπική απόκλιση σ_n, για να παραχθούν οι ζητούμενες αποκρίσεις. Έτσι η πρόβλεψη $\hat{f}_*$ για ένα δοσμένο διάνυσμα εισόδου προκύπτει από την ακόλουθη σχέση:

$$\hat{f}_* = E(f_*|X, y, x_*) = V(x, X)[V(X,X) + \sigma_n^2 I]^{-1} y \tag{4.27}$$

Η εξίσωση αυτή προκύπτει με τη χρήση του κανόνα Bayes στις συγκεκριμένες κανονικές πυκνότητες.

ΚΕΦΑΛΑΙΟ 5

ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

ΚΕΦΑΛΑΙΟ 5: ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

5.1 Εισαγωγή στη Μηχανική Μάθηση

Η μηχανική μάθηση εντάσσεται στο πεδίο της Τεχνητής Νοημοσύνης (Artificial Intelligence) που ασχολείται με την κατασκευή και τη μελέτη συστημάτων τα οποία μπορούν να «μαθαίνουν» από τα δεδομένα αντί να ακολουθούν ρητές προγραμματιστικές οδηγίες. Η διαδικασία της μηχανικής μάθησης είναι παρόμοια με αυτήν της εξόρυξης δεδομένων (data mining) που περιγράφηκε στο κεφάλαιο 2.1. Και τα δύο συστήματα αναλύουν δεδομένα και αναζητούν μοτίβα συμπεριφοράς. Ωστόσο αντί της εξαγωγής δεδομένων για την ανθρώπινη κατανόηση, όπως συμβαίνει σε εφαρμογές data mining, η μηχανική μάθηση χρησιμοποιεί τα δεδομένα για να βελτιώσει την κατανόηση του ίδιου του προγράμματος. Έτσι τα προγράμματα μηχανικής μάθησης ανιχνεύουν μοτίβα στα δεδομένα και προσαρμόζουν τις δράσεις τους με ανάλογο τρόπο.

Ένας σύγχρονος ορισμός για τη Μηχανική Μάθηση ο οποίος δόθηκε από τον Tom M. Mitchell ([127]) και χρησιμοποιείται ευρέως είναι: "Ένα πρόγραμμα υπολογιστή λέγεται ότι μαθαίνει από μια εμπειρία E σε σχέση με μια κλάση εργασιών T και μέτρο απόδοσης P, αν η απόδοσή του σε εργασίες από το T, όπως μετριέται από το P, βελτιώνεται μέσω της εμπειρίας E."

Οι αλγόριθμοι μηχανικής μάθησης μπορούν να κατηγοριοποιηθούν ανάλογα με το επιθυμητό αποτέλεσμα του αλγορίθμου. Οι δύο πιο βασικές κατηγορίες είναι οι ακόλουθες:

- Επιβλεπόμενη Μάθηση (Supervised Learning)

Στην επιβλεπόμενη μάθηση γνωρίζουμε ποια είναι η σωστή μορφή της εξόδου για το σύνολο των δεδομένων που εξετάζουμε, με την έννοια ότι υπάρχει μια σχέση μεταξύ της εισόδου και της εξόδου του συστήματος. Τα προβλήματα εποπτευόμενης μάθησης διακρίνονται σε προβλήματα παλινδρόμησης και ταξινόμησης. Στην πρώτη περίπτωση γίνεται προσπάθεια πρόβλεψης των αποτελεσμάτων για μια έξοδο συνεχούς τιμής, γεγονός που σημαίνει ότι προσπαθούμε να αντιστοιχήσουμε τις μεταβλητές εισόδου σε μια συνεχή συνάρτηση. Στη δεύτερη περίπτωση προσπαθούμε να προβλέψουμε τα αποτελέσματα για μια έξοδο διακριτής τιμής. Σκοπός επομένως είναι η αντιστοίχηση των μεταβλητών εισόδου σε διακριτές κατηγορίες.

- Μη επιβλεπόμενη Μάθηση (Unsupervised Learning)

Στη μη επιβλεπόμενη μάθηση γίνεται προσέγγιση του προβλήματος έχοντας λίγη σχετικά γνώση για τη σωστή μορφή των αποτελεσμάτων. Είναι δυνατό να αντλήσουμε πληροφορίες από τα δεδομένα, χωρίς να γνωρίζουμε ωστόσο την ακριβή επίδραση των μεταβλητών σε αυτά. Μπορούμε να εξάγουμε τη δομή ομαδοποιώντας τα δεδομένα βάσει των σχέσεων

που εμφανίζουν με τις μεταβλητές. Στη μη εποπτευόμενη μάθηση δεν υπάρχει ανατροφοδότηση (feedback) από τις τιμές πρόβλεψης.

Δύο από τους πιο δημοφιλείς αλγόριθμους μηχανικής μάθησης είναι τα τεχνητά νευρωνικά δίκτυα και οι μηχανές διανυσμάτων υποστήριξης ([119], [128]).

- Τεχνητά Νευρωνικά Δίκτυα: Αποτελούν έναν αλγόριθμο μάθησης εμπνευσμένο από τη δομή και τις λειτουργίες των βιολογικών νευρώνων. Μπορούν να εφαρμοστούν σε προβλήματα επιβλεπόμενης και μη μάθησης.
- Μηχανές Διανυσμάτων Υποστήριξης: Αποτελούν μοντέλα επιβλεπόμενης μηχανικής μάθησης. Στοχεύουν στην ανάλυση δεδομένων και την αναγνώριση προτύπων. Βρίσκουν εφαρμογές σε προβλήματα παλινδρόμησης και ταξινόμησης.

5.2 Τεχνητά Νευρωνικά Δίκτυα (ANN)

Τα τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks) αποτελούν προσπάθειες του ανθρώπου να μιμηθεί και να καταλάβει τι συμβαίνει στο νευρικό σύστημα, με την ελπίδα να αξιοποιήσει μέρος από τα πλεονεκτήματα αυτών των βιολογικών συστημάτων. Η λειτουργία των τεχνητών νευρώνων είναι παρόμοια με εκείνη των πραγματικών. Έτσι είναι σε θέση να επικοινωνούν στέλνοντας σήματα μεταξύ τους μέσω ενός μεγάλου αριθμού μεροληπτικών ή σταθμισμένων συνδέσεων. Κάθε νευρώνας δέχεται πληροφορίες στους κόμβους εισόδου και συγκεντρώνει δεδομένα. Η συσχετισμένη συνάρτηση μεταφοράς του, περιγράφει τον τρόπο με τον οποίο το σταθμισμένο άθροισμα των εισόδων του μετατρέπεται σε μία έξοδο. Η έξοδος αυτή μπορεί να αποτελέσει είσοδο για άλλους νευρώνες.

Η χρήση των τεχνητών νευρωνικών δικτύων στην πρόβλεψη ηλεκτρικού φορτίου εντοπίζεται στη δεκαετία του 1990 ([90], [129]). Τα τεχνητά νευρωνικά δίκτυα (ΤΝΔ) ουσιαστικά είναι μη γραμμικά δίκτυα τα οποία έχουν εμφανίσει σημαντική ικανότητα στη μη γραμμική προσαρμογή των καμπυλών του ηλεκτρικού φορτίου. Η βασική τους ιδέα βασίζεται στο φιλτράρισμα των ανεξάρτητων μεταβλητών εισόδου του μοντέλου για την παραγωγή προβλέψεων. Αυτό επιτυγχάνεται με τη βοήθεια κρυφών επιπέδων που αποτελούνται από κρυφούς κόμβους. Πρακτικά τα στοιχεία των δικτύων τοποθετούνται σε ένα σχετικά μικρό αριθμό διασυνδεδεμένων στρωμάτων μεταξύ εισόδων και εξόδου. Οι έξοδοι των νευρωνικών δικτύων είναι κάποιες μαθηματικές συναρτήσεις, γραμμικές ή μη, των εισόδων του συστήματος.

Η πιο δημοφιλής αρχιτεκτονική νευρωνικού δικτύου για την πρόβλεψη ηλεκτρικού φορτίου είναι αυτή της ανάστροφης διάδοσης (back propagation). Ο αλγόριθμος back propagation είναι μια βασική διαδικασία μάθησης ή ρύθμισης βαρών, η οποία διαδίδει το σφάλμα προς τα πίσω και προσαρμόζει τα βάρη αναλόγως ([17]). Ο συγκεκριμένος αλγόριθμος είναι αλγόριθμος επιβλεπόμενης μάθησης και χρησιμοποιείται κυρίως στα νευρωνικά δίκτυα MLP. Ο σκοπός της εισόδου με προκατάληψη ενός δικτύου back propagation είναι η σταθεροποίηση της συνάρτησης ενεργοποίησης για να προσφέρει καλύτερη μάθηση [130]. Στα συναπτικά βάρη εκχωρούνται αρχικά τυχαίες τιμές. Κατά τη διαδικασία εκπαίδευσης τα

δεδομένα επιλέγονται ομοιόμορφα. Ένα πέρασμα ολοκληρώνεται όταν όλα τα δεδομένα έχουν υποστεί επεξεργασία. Συνήθως απαιτούνται αρκετά περάσματα μέχρι να επιτευχθεί ένα επιθυμητό επίπεδο ακρίβειας. Το συνολικό σφάλμα προκύπτει από το μέσω τετραγωνικό σφάλμα (MSE) μεταξύ των πραγματικών δεδομένων εισόδου και των εξόδων που υπολογίζονται. Η ρύθμιση των βαρών διεξάγεται με τη χρήση του αλγορίθμου ανάστροφης διάδοσης, ο οποίος ελαχιστοποιεί το συνολικό σφάλμα ([131], [132]).

Η πιο δημοφιλής μορφή νευρωνικού δικτύου όσον αφορά στη βραχυπρόθεσμη πρόβλεψη είναι τα πολυεπίπεδα δίκτυα perceptron με τις διάφορες μορφές του αλγόριθμου ανάστροφης διάδοσης σφάλματος (back propagation). Στη συνέχεια παρουσιάζονται κάποιοι βασικοί τύποι νευρωνικών δικτύων οι οποίοι βρίσκουν εφαρμογή στην πρόβλεψη φορτίου.

5.2.1 Πολυεπίπεδα Δίκτυα Prceptron (MLP)

Το πολυεπίπεδο δίκτυο perceptron (Multilayer Perceptron) ανήκει στην κατηγορία των δικτύων πρόσθιας τροφοδότησης και αποτελεί γενίκευση του μονοστρωματικού perceptron. Απαρτίζεται από ένα επίπεδο εισόδου (input layer), ένα επίπεδο εξόδου (output layer) και ένα ή περισσότερα κρυφά επίπεδα (hidden layers) ενδιάμεσα. Κάθε ενδιάμεσο επίπεδο αποτελείται από κόμβους, τους οποίους ονομάζουμε νευρώνες. Κάθε νευρώνας με τη σειρά του συνδέεται μόνο με νευρώνες διαδοχικού επιπέδου μέσω παραμέτρων (βάρη). Η παρουσία αυτών των κρυφών στρωμάτων επιτρέπουν στο δίκτυο να υπολογίζει πιο περίπλοκες συσχετίσεις μεταξύ προτύπων. Η μαθηματική περιγραφή ενός δικτύου MLP δίνεται από την ακόλουθη σχέση:

$$y = v_0 + \sum_{j=1}^{NH} v_j g\left(w_j^T x'\right) \tag{5.1}$$

όπου
x': το διάνυσμα εισόδου x επαυξημένο με 1
w_j: το διάνυσμα βαρών του j −οστού κρυφού κόμβου
v_0, v_j: τα βάρη των κόμβων εξόδου
g: η έξοδος του κρυφού κόμβου
y: η έξοδος του δικτύου

Το MLP είναι ένα έντονα παραμετροποιήσιμο μοντέλο, όπου με την επιλογή του αριθμού των κρυφών κόμβων NH μπορεί να ελεγχθεί η πολυπλοκότητά του. Ένα δίκτυο MLP εκπαιδευμένο από τον αλγόριθμο ανάστροφης διάδοσης σφάλματος που περιγράψαμε παραπάνω, αποτελεί την πιο διαδεδομένη προσέγγιση του μοντέλου. Αυτό οφείλεται στη δυνατότητα χαρτογράφησης των σύνθετων σχέσεων μεταξύ εισόδου και εξόδου καθώς και των μαθηματικών τους ιδιοτήτων όπως παρουσιάζεται στα ([133], [134]).

5.2.2 Δίκτυα Πρόσθιας Τροφοδότησης (FFNN)

Στα πρόσθια δίκτυα τροφοδότησης (Feed Forward Neural Networks) η πληροφορία κινείται μόνο προς μία κατεύθυνση. Πιο συγκεκριμένα από τους νευρώνες εισόδου στα κρυφά επίπεδα, αν υπάρχουν, και από εκεί στους νευρώνες εξόδου. Οι νευρώνες συνδέονται χρησιμοποιώντας διανύσματα βαρών ενώ δεν υπάρχει ανατροφοδότηση ή σύνδεση μεταξύ των επιπέδων. Ένα δίκτυο FFNN με συνάρτηση προσέγγισης ενός κρυφού επιπέδου περιγράφεται σύμφωνα με την παρακάτω σχέση:

$$y = \sum_{j=1}^{N} w_j \Psi_j \left[\sum_{i=1}^{M} w_{ij} x_i + w_{i0} \right] + w_{j0} \tag{5.2}$$

όπου
Ψ_j: η συνάρτηση ενεργοποίησης για κάθε κρυφό κόμβο
N: ο συνολικός αριθμός των κρυφών κόμβων
M: ο συνολικός αριθμός των εισόδων
w_j: το διάνυσμα βαρών του j −οστού κρυφού κόμβου

Περισσότερες πληροφορίες σχετικά με τα δίκτυα πρόσθιας τροφοδότησης και τις εφαρμογές τους παρουσιάζονται στα ([36], [135], [136], [137]).

5.2.3 Δίκτυα Συνάρτησης Ακτινωτής Βάσης (RBFN)

Τα νευρωνικά δίκτυα RBF (Radial Basis Function Networks) εμφανίζουν παρόμοια αρχιτεκτονική με τα δίκτυα MLP, ωστόσο οι κόμβοι τους διαθέτουν τοπικές συναρτήσεις ενεργοποίησης, όπως περιγράφεται από τους Moody, Darken ([138]) και Powel ([139]). Η μαθηματική περιγραφή ενός δικτύου RBF δίνεται από την ακόλουθη σχέση:

$$y = \sum_{j=1}^{NB} w_j e^{-\frac{\|x - c_j\|^2}{\beta^2}} \tag{5.3}$$

όπου
w_j: το συνδυασμένο βάρος
c_j: το εύρος της συνάρτησης του κόμβου
β: το κέντρο της συνάρτησης του κόμβου για τη μονάδα j

Η έξοδος του δικτύου προκύπτει ως ο γραμμικός συνδυασμός των εξόδων από τους κόμβους. Οι κρυφοί κόμβοι του RBFN ακολουθούν την κανονικοποιημένη Gaussian συνάρτηση ενεργοποίησης. Σύμφωνα με το εύρος της μπορούμε ελέγξουμε την ομαλότητα της προσαρμοσμένης συνάρτησης. Επειδή ένα σταθερό κέντρο αντιστοιχεί σε μια δεδομένη παλινδρόμηση, όσον αφορά το μοντέλο γραμμικής παλινδρόμησης, η επιλογή των κέντρων RBF μπορεί να θεωρηθεί ως πρόβλημα επιλογής υποσυνόλου. Για την κατασκευή ενός δικτύου RBFN με ορθολογικό τρόπο εφαρμόζεται η μέθοδος ελαχίστων τετραγώνων ως διαδικασία πρόσθιας επιλογής. Έτσι ο αλγόριθμος επιλέγει τα κατάλληλα κέντρα RBF ένα

προς ένα από τα δεδομένα εκπαίδευσης μέχρι να προκύψει ένα ικανοποιητικό δίκτυο. Κάθε επιλεγμένο κέντρο ελαχιστοποιεί την αύξηση μέχρι την επιθυμητή διακύμανση εξόδου.

Λόγω της γραμμικής διαδικασίας που ακολουθείται στο επίπεδο εξόδου, το νευρωνικό δίκτυο RBFN απαιτεί μικρό χρόνο εκπαίδευσης. Ωστόσο ένα μειονέκτημα της μεθόδου σχετίζεται με τη διάσταση του χώρου εισόδου. Για μεγάλο πλήθος εισόδων, ο αριθμός των απαιτούμενων συναρτήσεων ακτινωτής βάσης αυξάνεται υπερβολικά. Έτσι με τη χρήση μεγάλου αριθμού κέντρων, το πλήθος των διαθέσιμων παραμέτρων κατά τη διαδικασία παλινδρόμησης θα έχει ως αποτέλεσμα ένα δίκτυο με μεγάλη ευαισθησία στα δεδομένα εκπαίδευσης και κακή απόδοση στη γενίκευση [140]. Εκτενέστερη ανάλυση του δικτύου RBFN παρουσιάζεται στα ([36], [141]).

5.2.4 Δίκτυα Γενικευμένης Παλινδρόμησης (GRNN)

Το νευρωνικό δίκτυο GRNN (Generalized Regression Neural Network) συχνά ονομάζεται εκτιμητής παλινδρόμησης πυρήνα και αναπτύχθηκε από τους Nadaraya και Watson ([142], [143]). Πρόκειται για ένα μη παραμετρικό μοντέλο όπου η πρόβλεψη ενός δεδομένου σημείου x προκύπτει ως ο μέσος όρος των εξόδων από τα σημεία εκπαίδευσης που είναι γειτονικά στο εξεταζόμενο σημείο x ([144]) . Η λειτουργία του δικτύου GRNN βασίζεται στη θεωρία της παλινδρόμησης πυρήνα. Έτσι ο υπολογισμός του τοπικού μέσου όρου προκύπτει σταθμίζοντας τα γειτονικά σημεία σύμφωνα με την απόστασή τους από το δεδομένο σημείο x, με τη βοήθεια κάποιας συνάρτησης πυρήνα. Το μοντέλο του δικτύου αποτελείται από τέσσερα επίπεδα. Κάθε επίπεδο εκτελεί μια συγκεκριμένη υπολογιστική λειτουργία για την εφαρμογή της μη γραμμικής συνάρτησης παλινδρόμησης. Το πρώτο επίπεδο του δικτύου λαμβάνει τις πληροφορίες και μέσω των νευρώνων εισόδου τροφοδοτεί τα δεδομένα στο δεύτερο στρώμα. Εκεί γίνεται η "απομνημόνευση" των σχέσεων μεταξύ των νευρώνων εισόδου και της κατάλληλης απόκρισης. Στο τρίτο επίπεδο, οι νευρώνες λαμβάνουν τις εξόδους του δεύτερου στρώματος και προσδιορίζουν το άθροισμα των μοτίβων μεταξύ των δεδομένων. Τέλος το άθροισμα αυτό τροφοδοτείται στο τέταρτο στρώμα για τον υπολογισμό της εξόδου του δικτύου.

Για τον υπολογισμό των βαρών στάθμισης χρησιμοποιείται η πολυμεταβλητή Gaussian συνάρτηση, η οποία περιγράφεται από την ακόλουθη σχέση:

$$K(u) = \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} \tag{5.4}$$

Επομένως τα σταθμισμένα βάρη w_m προκύπτουν σύμφωνα με τη σχέση:

$$w_m = \frac{K\left(\frac{\|x - x_m\|}{h}\right)}{\sum_{m'=1}^{M} K\left(\frac{\|x - x_{m'}\|}{h}\right)} \tag{5.5}$$

όπου
x: τα δεδομένα εκπαίδευσης
K: η συνάρτηση πυρήνα
h: το εύρος ζώνης

Η εκτίμηση που προκύπτει ως το σταθμισμένο άθροισμα των παρατηρούμενων αποκρίσεων, δίνεται από την ακόλουθη σχέση:

$$y = \sum_{m=1}^{M} w_m y_m \tag{5.6}$$

Η παράμετρος h είναι πολύ σημαντική αφού καθορίζει την ομαλότητα στην προσαρμογή. Με την αυξομείωση της τιμής της μπορούμε να ελέγξουμε το μέγεθος της περιοχής εξομάλυνσης.

5.2.5 Νευρωνικά Δίκτυα BNN

Το νευρωνικό δίκτυο BNN (Bayesian Neural Network) βασίζεται στην πιθανοτική διατύπωση κατά Bayesian ([145], [146]). Στα ([147], [148]) περιγράφονται οι αρχές του συλλογισμού Bayesian και η εφαρμογή τους για την εκτίμηση των NN παραμέτρων. Η προσέγγιση αυτή μπορεί να επιλύσει το πρόβλημα υπερπροσαρμογής (overfitting) των δεδομένων, ελέγχοντας την πολυπλοκότητα του μοντέλου. Η προσέγγιση Bayesian θεωρεί αρχικά μια συνάρτηση πιθανοτικής κατανομής (probability density function). Η βασική ιδέα των BNN είναι η διαχείριση των παραμέτρων του δικτύου και των βαρών ως τυχαίες μεταβλητές, οι οποίες ακολουθούν μια a priori κατανομή. Χάρις σε αυτήν την κατανομή είναι δυνατή η κατασκευή μοντέλων χαμηλής πολυπλοκότητας. Μετά την παρατήρηση των δεδομένων, αξιολογείται η a posteriori κατανομή των βαρών και υπολογίζεται η τελική πρόβλεψη. Όσον αφορά στο θέμα της τακτοποίησης (regularization), επιτυγχάνεται με την ελαχιστοποίηση μιας αντικειμενικής συνάρτησης η οποία περιγράφεται από την ακόλουθη σχέση:

$$J = aE_D + (1 - a)E_w \tag{5.7}$$

όπου
E_D: είναι το άθροισμα των τετραγωνικών σφαλμάτων στις εξόδους του δικτύου
E_W: είναι το άθροισμα των τετραγώνων των παραμέτρων
a: η παράμετρος τακτοποίησης (regularization)

Οι κατανομές a priori και a posteriori που αναφέραμε παραπάνω υπολογίζονται σύμφωνα με τις παρακάτω σχέσεις :

$$p(w) = \left(\frac{1-a}{\pi}\right)^{\frac{L}{2}} e^{-(1-a)E_W} \tag{5.8}$$

$$p(w|D,a) = \frac{p(D|w,a)p(w|a)}{p(D|a)} = ce^{-J}$$

όπου
L: ο αριθμός των παραμέτρων
D: τα δεδομένα που έχουν παρατηρηθεί
c: σταθερά ομαλοποίησης

Ο υπολογισμός της παραμέτρου a υπολογίζεται με τη χρήση εννοιών Bayesian από την ακόλουθη σχέση:

$$p(a|D) = \frac{p(D|a)p(a)}{p(D)} \tag{5.9}$$

Έτσι αντί για το "βέλτιστο" σύνολο βαρών που υπολογίζεται από την κλασσική προσέγγιση μέγιστης πιθανοφάνειας (μέσω ελαχιστοποίησης μιας συνάρτησης σφάλματος), η πιθανοτική κατανομή ορίζει αρχικά μια a priori κατανομή και τη μετατρέπει σε a posteriori κατανομή με τη χρήση του θεωρήματος Bayes μετά την παρατήρηση των δεδομένων. Αυτή η a posteriori κατανομή μπορεί να χρησιμοποιηθεί στη συνέχεια για να παραχθούν προβλέψεις του δικτύου σε νέες τιμές των μεταβλητών εισόδου.

5.2.6 Δυνάμεις και Αδυναμίες των Νευρωνικών Δικτύων

Όπως αναφέρθηκε στην παράγραφο 4.2.1, τα τεχνητά νευρωνικά δίκτυα, η μοντελοποίηση των οποίων βασίζεται στη λειτουργία του ανθρώπινου εγκεφάλου, εμφανίζουν ορισμένες πολύ σημαντικές ιδιότητες. Οι νευρώνες των δικτύων μπορεί να είναι είτε γραμμικοί, είτε μη γραμμικοί. Χάρις σε αυτήν την ιδιότητα αποτελούν μια καλή προσέγγιση σύνθετων προβλημάτων που δεν ακολουθούν γραμμικά πρότυπα. Επίσης η δυνατότητα γενίκευσης που διαθέτουν, επιτρέπει την εφαρμογή των νευρωνικών δικτύων στον τομέα της πρόβλεψης ενώ χάρις στην προσαρμοστικότητά τους μπορούν να ανταποκριθούν καλά σε ενδεχόμενες αλλαγές των μεταβλητών. Ένα ακόμα πολύ βασικό τους πλεονέκτημα είναι η ανοχή στα σφάλματα. Αυτό σημαίνει ότι αν ένα μέρος των νευρώνων του συστήματος ή των συνδέσεων μεταξύ τους καταστραφεί, το δίκτυο συνεχίζει να λειτουργεί έστω και με ένα μικρό σφάλμα. Τα νευρωνικά δίκτυα τέλος αποτελούν συστήματα παράλληλων κατανεμημένων διεργασιών, στοιχείο που συμβάλει σημαντικά στην ταχύτητά τους.

Ωστόσο αξίζει να τονιστεί ότι απαιτείται ιδιαίτερη προσοχή στην επιλογή του αριθμού των επιπέδων (layers) και των νευρώνων ανά επίπεδο. Ο βέλτιστος αριθμός τους δεν ίδιος για κάθε πρόβλημα και μπορεί να επηρεάσει σημαντικά την επίδοση του δικτύου. Έτσι για παράδειγμα η επιλογή πολύ μεγάλου αριθμού νευρώνων στα κρυφά επίπεδα μπορεί να οδηγήσει σε απομνημόνευση των δεδομένων εκπαίδευσης με αποτέλεσμα την αδυναμία γενίκευσης των δεδομένων. Αντίθετα η ανάπτυξη ενός πολύπλοκου δικτύου με μικρό αριθμό νευρώνων ενδέχεται να μην επιτυγχάνει καλή ταξινόμηση προτύπων.

5.3 Μηχανές Διανυσμάτων Υποστήριξης (SVM)

Ο αλγόριθμος SV (support vector) αποτελεί μια μη γραμμική γενίκευση του γενικευμένου αλγόριθμου "πορτρέτου", ο οποίος αναπτύχθηκε στη Ρωσία το 1960 ([149], [150]). Έχει τα θεμέλιά του στο πλαίσιο της στατιστικής θεωρίας μάθησης ή VC θεωρίας, η οποία αναπτύχθηκε από τους Vapnik και Chervonenkis ([94], [151]). Με λίγα λόγια η VC θεωρία χαρακτηρίζει τις ιδιότητες των μηχανών εκμάθησης, οι οποίες τους επιτρέπουν να γενικεύουν με επιτυχία τα δεδομένα.

5.3.1 Βασική Ιδέα

Τα SVM χρησιμοποιούν ένα γραμμικό μοντέλο για την εφαρμογή μη-γραμμικών ορίων μέσω μη-γραμμικής χαρτογράφησης των διανυσμάτων εισόδου x σε πολυδιάστατο χώρο χαρακτηριστικών. Με την κατασκευή ενός γραμμικού μοντέλου στο νέο χώρο προσδιορίζεται ένα μη γραμμικό όριο απόφασης στον αρχικό χώρο (original space). Στο νέο χώρο κατασκευάζεται ένα βέλτιστο διαχωριστικό υπερεπίπεδο. Για το λόγο αυτό η μέθοδος SVM είναι γνωστή ως ο αλγόριθμος που βρίσκει ένα γραμμικό μοντέλο ειδικού τύπου, το λεγόμενο υπερεπίπεδο μέγιστου περιθωρίου (maximum margin hyperplane). Αυτό το υπερεπίπεδο δίνει το μέγιστο όριο διαχωρισμού μεταξύ των κατηγοριών απόφασης. Ως διανύσματα υποστήριξης ορίζονται τα παραδείγματα εκπαίδευσης που βρίσκονται πιο κοντά στο υπερεπίπεδο μέγιστου περιθωρίου. Τα υπόλοιπα δεδομένα εκπαίδευσης δε σχετίζονται με τον καθορισμό των εκάστοτε ορίων δυαδικής τάξης [152].

Όσον αφορά στο γραμμικό διαχωρισμό, το υπερεπίπεδο που διαχωρίζει τα όρια απόφασης για την περίπτωση των τριών γνωρισμάτων (three-attribute case) μπορεί να περιγραφεί μαθηματικά από την ακόλουθη σχέση:

$$y = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3 \tag{5.10}$$

όπου
y: η τιμή του αποτελέσματος
x_i: οι τιμές των τριών γνωρισμάτων
w_i: τα τέσσερα βάρη που πρέπει να προσδιοριστούν από τον αλγόριθμο μάθησης

Τα βάρη w_i της σχέσης (5.10) είναι οι παράμετροι με βάση τις οποίες καθορίζεται το υπερεπίπεδο. Έτσι η μαθηματική αναπαράσταση του υπερεπιπέδου μέγιστου περιθωρίου σε όρους των διανυσμάτων υποστήριξης δίνεται από την ακόλουθη εξίσωση:

$$y = b + \sum \alpha_i y_i x(i) \cdot x \tag{5.11}$$

όπου
y_i: η τιμή της κατηγορίας των δεδομένων εκπαίδευσης $x(i)$
x: ένα παράδειγμα δοκιμής
$x(i)$: τα διανύσματα υποστήριξης
b, α_i: οι παράμετροι που καθορίζουν το υπερεπίπεδο

Από άποψη υλοποίησης, ο καθορισμός των παραμέτρων b και a και η εύρεση των διανυσμάτων υποστήριξης ισοδυναμεί με την επίλυση ενός γραμμικά περιορισμένου δευτεροβάθμιου προβλήματος (QP – quadratic programming). Αντίστοιχα για την περίπτωση του μη-γραμμικού διαχωρισμού, μια πολυδιάστατη εκδοχή της σχέσης (5.11) περιγράφεται απλοποιημένα ως ακολούθως:

$$y = b + \sum \alpha_i y_i K(x(i) \cdot x) \tag{5.12}$$

όπου
$K(x(i) \cdot x)$: η επιλεγμένη συνάρτηση πυρήνα

Υπάρχουν ορισμένοι τύποι μηχανών εκμάθησης (πυρήνες), οι οποίοι περιγράφονται πιο αναλυτικά στην παράγραφο 5.3.4. Αυτά τα είδη πυρήνων (kernels) χρησιμοποιούνται για τη γενίκευση του εσωτερικού γινομένου στις παραπάνω σχέσεις. Έτσι κατασκευάζονται μηχανές εκμάθησης με διαφορετικούς τύπους μη γραμμικών επιφανειών απόφασης στο χώρο εισόδου. Η επιλογή του κατάλληλου πυρήνα για την ελαχιστοποίηση του σφάλματος είναι απαραίτητη για τη δημιουργία του βέλτιστου μοντέλου πρόβλεψης.

5.3.2 Ελαχιστοποίηση Διαρθρωτικού Ρίσκου

Οι προσεγγίσεις τεχνητής νοημοσύνης έχουν την τάση να βασίζονται στη εύρεση συναρτήσεων που αντιστοιχούν στα σφάλματα εκπαίδευσης επί ενός συνόλου δεδομένων εκπαίδευσης. Αυτή η μέθοδος είναι γνωστή ως ελαχιστοποίηση του εμπειρικού ρίσκου (ERM – empirical risk minimization), δηλαδή ελαχιστοποίηση του λάθους κατηγοριοποίησης της εκπαίδευσης. Ωστόσο η μέθοδος ERM δεν εγγυάται πάντα κατάλληλη γενίκευση του νέου συνόλου δεδομένων δοκιμής. Έτσι τα SVM εφαρμόζουν τη μέθοδο ελαχιστοποίησης του διαρθρωτικού ρίσκου (SRM – structural risk minimization) για να διαχωρίσουν τις κλάσεις με μια επιφάνεια (υπερεπίπεδο) η οποία μεγιστοποιεί το περιθώριο ανάμεσα στα σύνολα των δεδομένων εκπαίδευσης. Με αυτόν τον τρόπο στοχεύουν στη δέσμευση του σφάλματος γενίκευσης και όχι στην ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (MSE) για τα δεδομένα κατάρτισης. Τα SVM παρέχουν μια καλά καθορισμένη ποσοτική μέτρηση σχετικά με την ικανότητα μιας συνάρτησης να προσδιορίσει την πραγματική δομή των δεδομένων διανομής και την ικανότητα γενίκευσης πάνω σε άγνωστα δεδομένα δοκιμής. Με βάση την αρχή της VC θεωρίας, που αναπτύχθηκε από τους Vladimir Vapnik και Alexey Chervonenkis, όπως περιγράφεται στα ([94], [105]), με την επιλογή της κατάλληλης συνάρτησης και την ελαχιστοποίηση του εμπειρικού ρίσκου για τα δεδομένα εκπαίδευσης, η μέθοδος SRM μπορεί να εγγυηθεί ένα κατώτατο όριο στα δεδομένα δοκιμής.

Υποθέτουμε ένα σύνολο δεδομένων εκπαίδευσης N στοιχείων $\{(x_i, y_i), i = 1,2, \dots, N\}$ με

$$\begin{cases} x_i = [x_{1i}, \dots, x_{ni}] \in R^n \\ \quad y_i \in \{-1, +1\} \end{cases}$$

όπου
x_i: το i −οστό στοιχείου του n −διάστατου χώρου
y_i: η ετικέτα του x_i

Στη συνέχεια ορίζουμε μια ντετερμινιστική συνάρτηση $f: x \to \{-1, +1\}$ για μια δεδομένη είσοδο x και ρυθμιζόμενα βάρη $w (w \in R^n)$ για σταθερή πυκνότητα κατανομής $P(x, y)$. Η ρύθμιση των βαρών w λαμβάνει χώρα στη φάση εκπαίδευσης των δεδομένων. Καθώς η πιθανότητα κατανομής είναι άγνωστη, η ελαχιστοποίηση του άνω ορίου για την πιθανότητα των σφαλμάτων ταξινόμησης δεν μπορεί να επιτευχθεί άμεσα. Για το λόγο αυτό επιχειρείται αρχικά η εκτίμηση μιας συνάρτησης $R(f)$ κατά προσέγγιση, δηλαδή το εμπειρικό ρίσκο ($R_{emp}(f)$), η οποία είναι κοντά στη βέλτιστη συνάρτηση για τα δεδομένα εκπαίδευσης. Ο υπολογισμός των συναρτήσεων αυτών σύμφωνα με τη μέθοδο SRM επιτυγχάνεται σύμφωνα με τις ακόλουθες σχέσεις ([153]):

$$R(f) \leq R_{emp}(f) + \varepsilon_1\left(N, h, \eta, R_{emp}\right)$$

$$R_{emp}(f) = \frac{1}{N}\sum_{i=1}^{N} |y_i - f(x_i)|_{loss\ function} \tag{5.13}$$

$$\varepsilon_1\left(N, h, \eta, R_{emp}\right) = 2\varepsilon_0^2(N, h, \eta)\left(1 + \sqrt{1 + \frac{R_{emp}(f)}{\varepsilon_0^2(N, h, \eta)}}\right)$$

$$\varepsilon_0(N, h, \eta) = \sqrt{\frac{h\left(\ln\left(\frac{2N}{h}\right) + 1\right) - \ln\left(\frac{\eta}{4}\right)}{N}}$$

όπου
h: ένα ακέραιος που ονομάζεται VC −διάσταση και αποτελεί μέτρο για την πολυπλοκότητα της συνάρτησης f
η: η τιμή της πιθανότητας
N: ο αριθμός των δεδομένων εκπαίδευσης
$\varepsilon_0(N, h, \eta)$: το VC διάστημα εμπιστοσύνης

Όσο πιο πολύπλοκο είναι το σύστημα τόσο καλύτερα μπορεί να ανταπεξέλθει στα μοτίβα που εμφανίζουν τα δεδομένα κατάρτισης και αντιστρόφως τόσο δυσκολότερα μπορεί να γενικευθεί όσον αφορά μεγαλύτερα δείγματα δεδομένων. Αξίζει να επισημανθεί ότι το προς αναζήτηση όριο δε σχετίζεται με την πιθανότητα κατανομής P αλλά εξαρτάται από τη VC διάσταση, η οποία δεν μπορεί να υπολογιστεί με άμεσο τρόπο, [154].

Σύμφωνα με τις παραπάνω σχέσεις, από την πρώτη εξίσωση προκύπτει ότι για ακραίες τιμές εμπειρικού ρίσκου, το συνολικό ρίσκο λαμβάνει τις ακόλουθες τιμές:

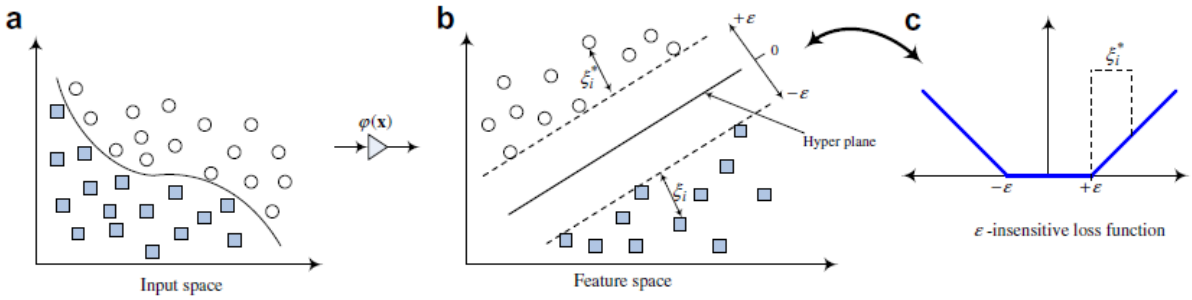
$$\begin{cases} R_{emp}(f) \approx 0 \ \ \to \ \ R(f) = R_{emp}(f) + 4\varepsilon_0^2(N, h, \eta) \\ R_{emp}(f) \approx 1 \ \ \to \ \ R(f) = R_{emp}(f) + \varepsilon_0(N, h, \eta) \end{cases} \tag{5.14}$$

Αντιλαμβανόμαστε επομένως ότι υπάρχουν δύο στρατηγικές για την ελαχιστοποίηση του άνω ορίου $R(f)$. Σύμφωνα με την πρώτη στρατηγική διατηρούμε σταθερό το VC διάστημα

εμπιστοσύνης και ελαχιστοποιούμε το εμπειρικό ρίσκο. Αυτή είναι η μέθοδος που εφαρμόζουν συνήθως τα νευρωνικά δίκτυα. Ωστόσο αυτός ο τρόπος δεν είναι πάντα αποτελεσματικός γιατί επηρεάζοντας μόνο τον όρο $R_{emp}(f)$ δε διασφαλίζεται η μείωση της VC εμπιστοσύνης. Σύμφωνα με τη δεύτερη στρατηγική, διαμορφώνουμε μια μικρή τιμή για το εμπειρικό ρίσκο και ελαχιστοποιούμε το VC διάστημα εμπιστοσύνης. Η μέθοδος αυτή αποτελεί την αρχή SRM που αναφέρθηκε παραπάνω. Τα SVM στηρίζονται σε αυτήν την αρχή, ωστόσο ο αλγόριθμος εκπαίδευσης που χρησιμοποιούν εξακολουθεί να βασίζεται σε μια ιεραρχία που εξαρτάται από τα δεδομένα [155].

5.3.3 Παλινδρόμηση Διανυσμάτων Υποστήριξης (SVR)

Οι συναρτήσεις πυρήνα (kernels) και τα προβλήματα δευτεροβάθμιου προγραμματισμού (quadratic programming) αποτελούν τα δύο βασικά χαρακτηριστικά κατά την παλινδρόμησης διανυσμάτων υποστήριξης (Support Vector Regression). Ο προσδιορισμός των παραμέτρων ενός μοντέλου SVR επιτυγχάνεται με την επίλυση τέτοιων προβλημάτων δευτεροβάθμιου προγραμματισμού, που έχουν ως περιορισμούς γραμμικές ισότητες και ανισότητες. Επίσης η ευελιξία των συναρτήσεων πυρήνα επιτρέπουν στο μοντέλο την αναζήτηση των χώρων επίλυσης μέσα από ένα ευρύ φάσμα [156].



Σχήμα 5.1: Διαδικασία μετασχηματισμού ενός μοντέλου SVR [157]

Οι μηχανές SVM αρχικά χρησιμοποιήθηκαν για την επίλυση προβλημάτων ταξινόμησης. Ωστόσο οι αρχές τους επεκτάθηκαν στη συνέχεια και στον τομέα της παλινδρόμησης και της πρόβλεψης χρονοσειρών. Η κεντρική ιδέα των μηχανών διανυσμάτων υποστήριξης για την παλινδρόμηση έγκειται στην αντιστοίχηση ενός δεδομένου εισόδου x_i σε ένα μεγαλύτερων διαστάσεων R^{n_h} χώρο χαρακτηριστικών με τη βοήθεια μιας μη γραμμικής συνάρτησης χαρτογράφησης φ.

$$\varphi(x): R^n \rightarrow R^{n_h}$$

$$x \in R^n \rightarrow \varphi(x) = [\varphi_1(x)\ \varphi_2(x) \dots \varphi_n(x)]^T \in R^{n_h}$$

Με τη χαρτογράφηση των δεδομένων εισόδου σε χώρο μεγαλύτερων διαστάσεων, το πρόβλημα περνά από πλαίσιο της μη-γραμμικής παλινδρόμησης στη γραμμική,

επιτρέποντας έτσι τη διατύπωσή του ως πρόβλημα δευτεροβάθμιου προγραμματισμού. Έτσι στον πολυδιάστατο πλέον χώρο χαρακτηριστικών R^{n_h}, υπάρχει θεωρητικά μια γραμμική συνάρτηση f που είναι ικανή να διατυπώσει τη μη-γραμμική σχέση μεταξύ των δεδομένων εισόδου και εξόδου, όπως φαίνεται στο σχήμα 8 (a, b). Η μαθηματική περιγραφή της συνάρτησης SVR δίνεται από την παρακάτω σχέση:

$$f(x) = w^T \varphi(x) + b \tag{5.15}$$

όπου
$f(x)$: η τιμή της πρόβλεψης
$\varphi(x)$: το χαρακτηριστικό γνώρισμα που έχει χαρτογραφηθεί μη-γραμμικά από το χώρο εισόδου x
w, b: ρυθμιζόμενοι συντελεστές

Οι συντελεστές w $(w \in R^{n_h})$ και b $(b \in R)$ προσδιορίζονται μέσω της διαδικασίας ελαχιστοποίησης του εμπειρικού ρίσκου, η οποία περιγράφεται από την ακόλουθη σχέση:

$$R_{emp}(f) = \frac{1}{N} \sum_{1}^{N} \Theta_\varepsilon(y_i, w^T \varphi(x_i) + b) \tag{5.16}$$

όπου
$\Theta_\varepsilon\big(y, f(x)\big)$ είναι η συνάρτηση απωλειών (ε-insensitive loss function), η γραφική αναπαράσταση της οποίας φαίνεται στο σχήμα 8 (c) με έντονη υπογράμμιση. Η μαθηματική περιγραφή της συνάρτησης δίνεται από την ακόλουθη σχέση:

$$\Theta_\varepsilon\big(y, f(x)\big) = \begin{cases} |f(x) - y| - \varepsilon, & \text{όταν } |f(x) - y| \geq \varepsilon \\ 0, & \text{διαφορετικά} \end{cases} \tag{5.17}$$

Η σχέση (5.17) βοηθά στην εύρεση ενός βέλτιστου υπερεπιπέδου στο νέο χώρο για τη μεγιστοποίηση της απόστασης που διαχωρίζει τα δεδομένα εκπαίδευσης σε δύο υποσύνολα. Ο όρος ε, αναπαριστά τη συνάρτηση του Vapnik (ε −insensitive loss function),η οποία ορίζει την ε −ζώνη (ε −zone) όπως φαίνεται στο σχήμα 8 (b). ΑΝ η προβλεπόμενη τιμή είναι εντός της περιοχής, τότε το σφάλμα προσέγγισης θα ισούται με μηδέν. Για όλες τις υπόλοιπες τιμές που βρίσκονται εκτός της περιοχής, η απώλεια ισούται με την απόσταση μεταξύ του δεδομένου (σημείο ξ_i) και της ακτίνα της ζώνης ε −insensitive zone. Η μέθοδος SVR εστιάζει επομένως στην εύρεση ενός κατάλληλου υπερεπιπέδου και την ελαχιστοποίηση του σφάλματος εκπαίδευσης μεταξύ των δεδομένων και της συνάρτησης απωλειών. Αυτό το πρόβλημα βελτιστοποίησης περιγράφεται από την ακόλουθη σχέση:

$$\min_{w,b,\xi,\xi^*} R_\varepsilon(w, \xi, \xi^*) = \frac{1}{2} w^T w + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) \tag{5.18}$$

υπό τους περιορισμούς

$$y_i - w^T \varphi(x_i) - b \leq \varepsilon + \xi_i^*, \qquad i = 1,2, \dots, N$$

$$-y_i + w^T \varphi(x_i) + b \leq \varepsilon + \xi_i^*, \qquad i = 1,2, \dots, N$$

$$\xi_i^* \geq 0, \qquad i = 1,2,\dots,N$$

$$\xi_i \geq 0, \qquad i = 1,2,\dots,N$$

όπου
ε: ο επιθυμητός βαθμός ανεκτικότητας στο σφάλμα για όλα τα σημεία
C: ένα όρος "ποινής" για την εξισορρόπηση μεταξύ της προσαρμογής στα δεδομένα και την εξομάλυνση
ξ_i, ξ_i^*: ρυθμιζόμενες μεταβλητές που εξασφαλίζουν ότι υπάρχει λύση για κάθε ε
w: τα βάρη που χρησιμοποιούνται στην παλινδρόμηση
φ: αντιπροσωπεύει μια συνάρτηση πυρήνα για την αντιστοίχηση του χώρου εισόδου σε ένα πολυδιάστατο χώρο χαρακτηριστικών

Ο πρώτος όρος $\left(\frac{1}{2}w^T w\right)$ της σχέσης (5.18) εφαρμόζει την έννοια της μεγιστοποίησης της απόστασης μεταξύ δύο ξεχωριστών δεδομένων εκπαίδευσης, ενώ χρησιμοποιείται για να ρυθμίσει τις τιμές των βαρών, να επιβάλλει "ποινές" όταν οι τιμές τους είναι αυξημένες και να διατηρήσει την ομαλότητα της συνάρτησης παλινδρόμησης. Ο δεύτερος όρος $\left(C\sum_{i=1}^{N}(\xi_i + \xi_i^*)\right)$ επιβάλλει "ποινή" στα σφάλματα εκπαίδευσης των $f(x)$ και y μέσω της συνάρτησης ε −insensitive loss function. Η παράμετρος κόστους C καθορίζει την εξισορρόπηση ανάμεσα στο εμπειρικό ρίσκο και το είδος ομαλοποίησης που περιλαμβάνει το μοντέλο. Εάν ένα σημείο x βρίσκεται εκτός της περιοχής ε −zone, τότε υπάρχει ένα σφάλμα ξ_i (εάν το σημείο βρίσκεται κάτω από $-\varepsilon$) ή ξ_i^* (εάν το σημείο βρίσκεται πάνω από $+\varepsilon$) το οποίο προσπαθούμε να ελαχιστοποιήσουμε στην αντικειμενική συνάρτηση (5.18). Το SVR ελαχιστοποιώντας τους δύο όρους της σχέσης προσπαθεί να αποφύγει την υποπροσαρμογή ή υπερπροσαρμογή των δεδομένων εκπαίδευσης.

5.3.4 Δυϊκό Πρόβλημα και Τέχνασμα Πυρήνων

Η βασική ιδέα είναι να κατασκευαστεί μια συνάρτηση Lagrange από την αντικειμενική συνάρτηση (5.18) και να προσδιοριστούν οι αντίστοιχοι περιορισμοί με τη βοήθεια ενός διπλού συνόλου μεταβλητών (Dual Problem). Εισάγοντας τους πολλαπλασιαστές Lagrange a_i και a_i^* με $a_i a_i^* = 0$ και $a_i, a_i^* \geq 0$ για $i = 1,\dots,N$ και σύμφωνα με τις συνθήκες βελτιστοποίησης Karush-Kuhn-Tucker ([104]), η διαδικασία εκπαίδευσης του SVR καλείται να επιλύσει ένα δευτεροβάθμιο πρόβλημα βελτιστοποίησης που περιγράφεται από τις ακόλουθες σχέσεις:

$$\max_{a,a^*}\begin{cases} -\frac{1}{2}\sum_{i,j=1}^{N}(a_i - a_i^*)(a_j - a_j^*)\langle\varphi(x_i),\varphi(x_j)\rangle \\ \quad -\varepsilon\sum_{i=1}^{N}(a_i + a_i^*) + \sum_{i=1}^{N} y_i(a_i - a_i^*) \end{cases} \tag{5.19}$$

υπό τους περιορισμούς

$$\begin{cases} \sum_{i=1}^{N}(a_i - a_i^*) = 0 \\ a_i, a_i^* \in [0, C] \end{cases}$$

όπου
$\langle \varphi(x_i), \varphi(x_j) \rangle$: το εσωτερικό γινόμενο των σημείων $\varphi(x_i)$ και $\varphi(x_j)$ όπως έχουν αντιστοιχηθεί στο χώρο χαρακτηριστικών.

Η επίλυση του δευτεροβάθμιου προβλήματος προκύπτει συναρτήσει των διπλών μεταβλητών a_i και a_i^*. Το διάνυσμα βαρών w του βέλτιστου υπερεπιπέδου μπορεί να βρεθεί με τη βοήθεια των πολλαπλασιαστών Lagrange σύμφωνα με τις ακόλουθη σχέση:

$$w = \sum_{i=1}^{N}(a_i - a_i^*)x_i \tag{5.20}$$

Επομένως η συνάρτηση παλινδρόμησης SVR στο νέο χώρο λαμβάνει την παρακάτω μορφή:

$$f(x) = \sum_{i=1}^{N}(a_i - a_i^*)K(x_i, x_j) + b \tag{5.21}$$

όπου
$K(x_i, x_j)$: η συνάρτηση πυρήνα

Η τιμή του πυρήνα (Kernel) είναι ίση με το εσωτερικό γινόμενο των διανυσμάτων x_i και x_j, στο χώρο χαρακτηριστικών (feature space), $\varphi(x_i)$ και $\varphi(x_j)$ αντίστοιχα. Χρησιμοποιώντας τον πυρήνα εσωτερικού γινομένου $K(x_i, x_j)$, όλοι οι απαραίτητοι υπολογισμοί μπορούν να εκτελεστούν απευθείας στο χώρο εισόδου χωρίς τον προσδιορισμό της συνάρτησης χαρτογράφησης $\varphi(x)$. Η συνάρτηση πυρήνα ακολουθεί αυστηρά τους περιορισμούς που επιβάλλονται από το θεώρημα Mercer και αυτό παρέχει τον υπολογισμό, με ένα βήμα, του γινομένου μεταξύ των $\varphi(x_i)$ και $\varphi(x_j)$ [158]. Η επέκταση του πυρήνα εσωτερικού γινομένου επιτρέπει την κατασκευή μιας επιφάνειας απόφασης που είναι μη γραμμική στο χώρο εισόδου, αλλά έχει γραμμική εικόνα στο χώρο χαρακτηριστικών. Σε αυτήν την επέκταση βασίζεται και η διπλή (dual) μορφή για τη βελτιστοποίηση ενός SVM. Με βάση το θεώρημα του Mercer μπορεί να διαπιστωθεί εάν ένας υποψήφιος πυρήνας είναι ή όχι πυρήνας εσωτερικού γινομένου σε κάποιο χώρο και επομένως αποδεκτός για χρήση σε ένα support vector machine.

Ανάλογα με το πώς παράγεται ο πυρήνας, είναι εφικτό να κατασκευαστούν διαφορετικά είδη μηχανών εκμάθησης τα οποία χαρακτηρίζονται από διαφορετικές μη γραμμικές επιφάνειες απόφασης. Στον ακόλουθο πίνακα παρουσιάζονται τα βασικά είδη πυρήνων και οι παράμετροι που τα χαρακτηρίζουν.

Η θεμελιώδης αρχή ενός SVM δε στηρίζεται στη χρήση κάποιας ευριστικής μεθόδου (heuristics), όπως συμβαίνει συχνά στα νευρωνικά δίκτυα RBF και MLP (παράγραφος 5.2). Για παράδειγμα στον τύπο Gaussian RBF (συνάρτηση ακτινωτής βάσης) ενός support vector machine το πλήθος των συναρτήσεων RBF και τα κέντρα τους καθορίζονται αυτόματα από

το πλήθος των SVM και των τιμών τους αντίστοιχα. Η διάσταση του χώρου χαρακτηριστικών (feature space) καθορίζεται από το πλήθος των SVM που εξήχθησαν από τα δεδομένα εκπαίδευσης κατά τη διαδικασία επίλυσης του προβλήματος βελτιστοποίησης. Περισσότερες πληροφορίες σχετικά με τα είδη των πυρήνων και τις εφαρμογές τους παρουσιάζονται στα ([160], [161], [162]).

Kernel Type	Expression	Parameter
Linear	$K(x_i, x_j) = x_i \cdot x_j$	—
Polynomial	$K(x_i, x_j) = (x_i \cdot x_j + 1)^d$	d
Gaussian Radial Basis Function	$K(x_i, x_j) = exp\left(-\frac{\lVert x_i - x_j \rVert^2}{2\sigma^2}\right)$	σ
Exponential Radial Basis Function	$K(x_i, x_j) = exp\left(-\frac{\lVert x_i - x_j \rVert}{2\sigma^2}\right)$	σ
Hyperbolic Tangent	$K(x_i, x_j) = \tanh(b(x_i \cdot x_j) + c)$	b, c
Fourier Series	$K(x_i, x_j) = \frac{\sin\left(\delta + \frac{1}{2}\right)(x_i - x_j)}{\sin\left(\frac{1}{2}(x_i - x_j)\right)}$	δ
Linear Splines	$K(x_i, x_j) = 1 + xy + xy\min(x, y) - \frac{(x+y)}{2}(\min(x, y))^2 + \frac{1}{3}(\max(x, y))^3$	—

Πίνακας 5.1: Τύποι πυρήνων εσωτερικού γινομένου $K(x_i,x_j)$ [159]

5.3.5 Δυνάμεις και Αδυναμίες των SVM

Οι μηχανές διανυσμάτων υποστήριξης (SVMs) αποτελούν ένα σύνολο μεθόδων επιβλεπόμενης μάθησης (supervised learning) που χρησιμοποιούνται σε περιπτώσεις ταξινόμησης, παλινδρόμησης και ανίχνευσης ακραίων τιμών.

Τα SVM βασίζονται στη στατιστική θεωρία μάθησης και μπορούν να χρησιμοποιηθούν για την πρόβλεψη μελλοντικών δεδομένων. Η εκπαίδευση ενός SVM επιτυγχάνεται από τη επίλυση ενός τετραγωνικού προβλήματος βελτιστοποίησης. Το σημαντικό πλεονέκτημα αυτής της μοντελοποίησης είναι η εύρεση μιας μοναδικής βέλτιστης λύσης για κάθε επιλογή των SVM παραμέτρων. Ωστόσο αυτή η βέλτιστη λύση εξαρτάται σημαντικά από την παράμετρο C αλλά και τις παραμέτρους του επιλεγμένου τύπου πυρήνα για τη συνάρτηση

χαρτογράφησης $\varphi(x)$. Υπάρχουν πολλές τεχνικές για την επιλογή των κατάλληλων παραμέτρων όπως η αναζήτηση πλέγματος (grid search) σε συνδυασμό με σταυρωτή επικύρωση (cross-validation). Λεπτομερής επισκόπηση των τεχνικών για τη ρύθμιση των παραμέτρων παρουσιάζεται στο [104].

Τα τέσσερα πιο σημαντικά χαρακτηριστικά των SVM είναι οι πυρήνες, η δυαδικότητα, η κυρτότητα και η σποραδικότητα. Η μέθοδος παρουσιάζει ευελιξία χάρις στους διάφορους τύπους πυρήνων που χρησιμοποιεί για να καθορίσει τις συναρτήσεις απόφασης. Εκτός από τους βασικούς τύπους που φαίνονται στον πίνακα 5.1, υπάρχει δυνατότητα καθορισμού προσαρμοσμένων πυρήνων ανάλογα με το πρόβλημα. Επίσης οι μηχανές διανυσμάτων υποστήριξης είναι αποτελεσματικές σε χώρους υψηλών διαστάσεων ακόμα και σε περιπτώσεις όπου ο αριθμός των διαστάσεων είναι μεγαλύτερος από τον αριθμό των δειγμάτων. Τέλος, καθώς χρησιμοποιούν ένα υποσύνολο των δεδομένων εκπαίδευσης στη συνάρτηση απόφασης (το λεγόμενο διάνυσμα υποστήριξης), εμφανίζουν χαμηλές απαιτήσεις σε αποθηκευτικούς πόρους.

Ένα εν δυνάμει μειονέκτημα της μεθόδου SVR έγκειται στην επεκτασιμότητα (scalability). Αυτή αφορά στην κλιμάκωση των δεδομένων σε υψηλές διαστάσεις και την εξισορρόπηση μεταξύ της πολυπλοκότητας και του εμπειρικού ρίσκου. Εάν ο αριθμός των διαστάσεων του χώρου χαρακτηριστικών είναι εξαιρετικά μεγαλύτερος του αριθμό των δειγμάτων, η μέθοδος είναι πιθανό να εμφανίσει μέτρια απόδοση.

Μια εφαρμογή της μεθόδου SVR σε βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φορτίου παρουσιάζεται στο [163]. Ο συγγραφέας συνέκρινε την επίδοση των SVM με την αυτοπαλινδρομική μέθοδο. Σύμφωνα με τις τελικές προβλέψεις οι μηχανές διανυσμάτων υποστήριξης ήταν πιο αποτελεσματικές. Περισσότερες πληροφορίες για τη χρήση της μεθόδου SVR σε βραχυπρόθεσμη πρόβλεψη παρουσιάζονται στα ([164], [165], [166], [167], [168]). Εφαρμογές της μεθόδου σε μεσοπρόθεσμη πρόβλεψη φορτίου περιγράφονται στο [169]. Στο [170] οι Chen et al. ανέπτυξαν ένα μοντέλο SVM για την πρόβλεψη της ημερήσιας ζήτησης για ένα μήνα. Το πρόγραμμά τους ήταν η νικητήρια συμμετοχή στο διαγωνισμό που οργανώθηκε από το δίκτυο EUNITE 2001.

Στο πλαίσιο της παρούσας διπλωματικής στόχος είναι η βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φορτίου με τη χρήση μηχανών υποστήριξης SVM. Οι πυρήνες εσωτερικού γινομένου που θα εξεταστούν είναι ο γραμμικός (linear), ο πολυωνυμικός (polynomial) και η Γκαουσιανή συνάρτηση ακτινωτής βάσης (Gaussian RBF). Εκτενέστερη περιγραφή των πειραματικών διαδικασιών και σχολιασμός των αποτελεσμάτων παρουσιάζονται στο κεφάλαιο 6.

ΚΕΦΑΛΑΙΟ 6

ΠΕΙΡΑΜΑΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ

ΚΕΦΑΛΑΙΟ 6: ΠΕΙΡΑΜΑΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ

6.1 Εφαρμογή Μηχανικής Μάθησης

Αντικείμενο της διπλωματικής εργασίας, όπως αναφέραμε στο κεφάλαιο 1, είναι η πρόβλεψη της ενεργειακής ζήτησης για βραχυπρόθεσμο χρονικό ορίζοντα στον κτηριακό τομέα. Στόχος είναι η πρόταση συγκεκριμένων λύσεων μέσω της πιλοτικής εφαρμογής που παρουσιάζεται στην παράγραφο 6.2. Για την υλοποίηση των προβλεπτικών μοντέλων έγινε χρήση των μηχανών διανυσμάτων υποστήριξης (SVM), μέθοδος που ανήκει στη μηχανική μάθηση (παράγραφος 5.3).

Η προσέγγιση που ακολουθείτε για την ανάπτυξη μιας εφαρμογής με τη χρήση μεθόδων μηχανικής μάθησης, διαρθρώνεται συνοπτικά από τα παρακάτω βήματα:

- ΒΗΜΑ 1: Συλλογή δεδομένων

Η συλλογή δεδομένων προκύπτει συνήθως μέσω κάποιας ιστοσελίδας με την εξαγωγή δεδομένων, είτε από πληροφορίες που προέρχονται από ροές δεδομένων RSS είτε μέσω μιας διεπαφής προγραμματισμού εφαρμογών (API). Συχνά για τη μελέτη ενεργειακών προβλημάτων πρόβλεψης γίνεται χρήση “ανοιχτών” συνόλων δεδομένων που προκύπτουν από σχετικούς διαγωνισμούς. Τέλος υπάρχει η δυνατότητα απόκτησης δεδομένων μέσω κάποιας μετρητικής συσκευής, όπως είναι οι αισθητήρες κατανάλωσης στον κτηριακό τομέα.

- ΒΗΜΑ 2: Προετοιμασία των δεδομένων εισόδου

Πολύ σημαντικό βήμα μετά τη συλλογή δεδομένων είναι η μορφοποίησή τους κατά τρόπο που να είναι αξιοποιήσιμα. Είναι αναγκαία η επιλογή της κατάλληλης γλώσσας προγραμματισμού και ενδεχομένως κάποιου αλγόριθμου για την προεπεξεργασία των δεδομένων προκειμένου να ελεγχθούν διακυμάνσεις και ακραίες τιμές των δεδομένων και να συμπληρωθούν τιμές που λείπουν με τη βοήθεια τεχνικών παλινδρόμησης. Η γλώσσα Python αποτελεί δημοφιλή επιλογή σε εφαρμογές μηχανικής μάθησης.

- ΒΗΜΑ 3: Ανάλυση των δεδομένων εισόδου

Σε αυτό το σημείο γίνεται η εξέταση των δεδομένων που προέκυψαν από το προηγούμενο βήμα. Αυτό θα μπορούσε να σημαίνει απλά επισκόπηση των δεδομένων που έχουν αναλυθεί σε ένα πρόγραμμα επεξεργασίας κειμένου ή μια βάση δεδομένων ώστε να επιβεβαιωθεί η σωστή υλοποίηση των δύο πρώτων βημάτων. Στις περισσότερες περιπτώσεις επεκτείνεται σε μια προσπάθεια αναγνώρισης προτύπων μέσα από τα δεδομένα. Για το σκοπό αυτό επιλέγεται κάποια μέθοδος συσχέτισης μεταξύ των δεδομένων που εξετάζουμε και παραμέτρων που το επηρεάζουν. Σε ενεργειακές εφαρμογές η σχέση μεταξύ του φορτίου κατανάλωσης και των μεταβλητών που το επηρεάζουν είναι συνήθως μη γραμμική. Μια μέθοδος που προτείνεται για την ανίχνευση τέτοιων συσχετίσεων ονομάζεται MIC (maximal information coefficient).

- ΒΗΜΑ 4: Εκπαίδευση του αλγόριθμου

Σε αυτό το επίπεδο λαμβάνει χώρα η μηχανική μάθηση. Αυτό και το επόμενο βήμα αποτελούν το ουσιαστικό κομμάτι της μοντελοποίησης, σύμφωνα με τον επιλεγμένο αλγόριθμο. Ο αλγόριθμος τροφοδοτείται από τα "καθαρά" πλέον δεδομένα των δύο πρώτων βημάτων και εξάγει γνώσεις ή πληροφορίες. Η αποθήκευση αυτής της γνώσης γίνεται συνήθως σε μορφή που είναι εύχρηστη για τα δύο επόμενα βήματα. Στη περίπτωση της μη επιβλεπόμενης μάθησης, παράγραφος 5.1, για την πρόβλεψη μίας συνεχούς μεταβλητής επιλέγεται κάποια μέθοδος παλινδρόμησης. Στο πλαίσιο της μηχανικής μάθησης, οι μηχανές διανυσμάτων υποστήριξης έχουν αναδειχτεί σε αποτελεσματική μέθοδο τα τελευταία χρόνια. Για την υλοποίησή της σε γλώσσα Python, χρησιμοποιείται κυρίως η βιβλιοθήκη Scikit-Learn, η οποία αναλύεται στην παράγραφο 6.1.3.

- ΒΗΜΑ 5: Έλεγχος του αλγόριθμου

Σε αυτό το επίπεδο γίνεται χρήση των πληροφοριών που προέκυψαν από το προηγούμενο βήμα. Η αξιολόγηση της αποτελεσματικότητας ενός αλγορίθμου, προϋποθέτει τη δοκιμή του στα υπό εξέταση δεδομένα. Όσον αφορά στην επιβλεπόμενη μάθηση, καθώς οι τιμές των εξόδων είναι γνωστές εκ των προτέρων, η διαδικασία της αξιολόγησης είναι εύκολη. Στην περίπτωση της μη επιβλεπόμενης μάθησης, υπάρχουν διάφοροι δείκτες για την αξιολόγηση του αλγόριθμου, οι οποίοι περιγράφονται στην παράγραφο 3.5. Και στις δύο περιπτώσεις αν τα αποτελέσματα δεν εμπίπτουν στα αποδεκτά όρια σφαλμάτων, μια λύση είναι η επιστροφή στο τέταρτο βήμα και η αλλαγή των επιλεγμένων παραμέτρων. Συχνά όμως το πρόβλημα εντοπίζεται στη διαδικασία συλλογής και προετοιμασίας των δεδομένων και απαιτείται η επιστροφή στο πρώτο βήμα.

- ΒΗΜΑ 6: Χρήση του προγράμματος

Στο τελευταίο βήμα λαμβάνει χώρα η εφαρμογή του προγράμματος μηχανικής μάθησης, όπως αυτό προέκυψε από την παραπάνω διαδικασία. Στην περίπτωση όπου προκύψουν επιπλέον δεδομένα εισόδου, απαιτείται ανανέωση της βάσης δεδομένων και ενδεχομένως επανεξέταση των προηγούμενων βημάτων.

6.1.1 Μέθοδοι προεπεξεργασίας δεδομένων

Τα αρχικά (ακατέργαστα) δεδομένα είναι πολύ ευαίσθητα στο θόρυβο ενώ συνήθως διακρίνονται από ελλιπείς τιμές και ανακολουθίες. Η ποιότητα των δεδομένων μπορεί να επηρεάσει σημαντικά τη διαδικασία πρόβλεψης, ανεξάρτητα από το επιλεγμένο μοντέλο. Προκειμένου να βελτιωθεί η ποιότητα τους και κατά συνέπεια η ποιότητα των αποτελεσμάτων πρόβλεψης, τα ακατέργαστα δεδομένα (raw data) υφίστανται μία διαδικασία προεπεξεργασίας με στόχο τη βελτίωση της απόδοσης και τη διευκόλυνση της διαδικασίας πρόβλεψης. Η προεπεξεργασία αυτή των δεδομένων αποτελεί ένα πολύ σημαντικό βήμα σε μια εφαρμογή μηχανικής μάθησης, όπως περιγράφεται παραπάνω, και

ασχολείται με την προετοιμασία και το μετασχηματισμό του αρχικού συνόλου δεδομένων [171]. Το πρώτο και πιο σημαντικό στάδιο είναι η διεξοδική ανάλυση των δεδομένων για την καλύτερη κατανόησή τους. Το υπόβαθρο πίσω από την ανίχνευση και τη συλλογή των δεδομένων λειτουργεί ως αρωγός στην αντίληψη των τιμών που λείπουν και τα επίπεδα θορύβου που διακρίνουν τα δεδομένα. Στη συνέχεια παρουσιάζονται οι βασικές μέθοδοι προεπεξεργασίας των δεδομένων.

- **Κανονικοποίηση (Normalization)**

Η κανονικοποίηση όλων των χαρακτηριστικών (features) δεν είναι πάντα απαραίτητη, αλλά σε γενικές γραμμές λειτουργεί βοηθητικά στην τελική πρόβλεψη. Μπορεί να πραγματοποιηθεί είτε σε επίπεδο δεδομένων, είτε σε επίπεδο πυρήνων, ή και στα δύο. Όταν συλλέγονται δεδομένα σε διαφορετικές κλίμακες, γίνεται προσπάθεια να καθοριστούν οι τιμές έτσι ώστε να εκφράζονται με την ίδια κλίμακα. Η πιο συνηθισμένη μέθοδος κανονικοποίησης είναι η γραμμική, όπου οι τιμές των χαρακτηριστικών χαρτογραφούνται σε ένα εύρος [0,1] και μετατρέπονται με αυτόν τον τρόπο σε μοναδιαία διανύσματα. Μια εναλλακτική μέθοδος είναι η z-κανονικοποίηση, όπου σε κάθε μέτρηση αφαιρείται η μέση τιμή και το αποτέλεσμα διαιρείται με την τυπική απόκλιση. Η κανονικοποίηση συχνά προσφέρει βελτιωμένη απόδοση σε γραμμικούς και μη γραμμικούς πυρήνες, ενώ μπορεί να οδηγήσει και σε ταχύτερη σύγκλιση.

- **Μη διαθέσιμες τιμές (missing values)**

Στα εξεταζόμενα δεδομένα η διαχείριση των τιμών που λείπουν αποτελεί σημαντικό κομμάτι της προετοιμασίας. Στην περίπτωση όπου ο αριθμός των διαθέσιμων δειγμάτων είναι μεγάλος, υπάρχει η δυνατότητα παράλειψης εκείνων των δειγμάτων που παρουσιάζουν ελλείψεις, ή αποκοπής μιας συγκεκριμένης περιόδου με στόχο την απομάκρυνση του θορύβου. Η συμπλήρωση των τιμών που λείπουν με το χέρι, αποτελεί τον πιο απλό τρόπο, ωστόσο μπορεί να αποδειχθεί χρονοβόρος ειδικά για ένα μεγάλο σύνολο δεδομένων. Συχνά χρησιμοποιείται η μέση τιμή του υπό εξέταση χαρακτηριστικού είτε η μέση τιμή όλων των χαρακτηριστικών που ανήκουν στην ίδια κλάση για τη συμπλήρωση των μη διαθέσιμων τιμών. Τέλος η πιο δημοφιλής μέθοδος για την εκτίμηση των τιμών που λείπουν είναι εκείνη της γραμμικής παρεμβολής.

- **Διαχείριση ασυνήθιστων τιμών (Outliers)**

Η ύπαρξη ακραίων τιμών που έχουν συνήθως μικρή διάρκεια και οφείλονται σε κάποιο μεμονωμένο γεγονός αποτελούν ένα είδος ασυνέχειας στα δεδομένα. Τις περισσότερες φορές πρόκειται απλά για θόρυβο, λάθος στα δεδομένα, ή ένδειξη κάποιας ειδικής συμπεριφοράς. Κατά τη διαδικασία προετοιμασίας των δεδομένων είναι καλή ιδέα η αφαίρεση των ακραίων τιμών, οι οποίες συνήθως αν και λίγες σε αριθμό, μπορεί να έχουν μεγάλο αντίκτυπο στα αποτελέσματα της πρόβλεψης.

- **Συνάθροιση (Aggregation)**

Ο μετασχηματισμός των εξεταζόμενων δεδομένων σε κατάλληλη μορφή πριν την είσοδό τους σε μια εφαρμογή πρόβλεψης, αποτελεί μία τακτική συχνά χρησιμοποιούμενη. Για

παράδειγμα οι μετρήσεις ανά τέταρτο του ηλεκτρικού φορτίου μπορούν να συγκεντρωθούν σε ωριαίες ή ημερήσιες μετρήσεις. Με αυτόν τον τρόπο υπάρχει δυνατότητα αποφυγής θορύβου λόγω μεγάλης λεπτομέρειας στα δεδομένα αλλά και αξιολόγησης ενός συνόλου δεδομένων για διαφορετικά χρονικά επίπεδα ανάλυσης.

6.1.2 Python

Η επιλογή της προγραμματιστικής γλώσσας Python για την υλοποίηση της πιλοτικής εφαρμογής έγκειται στις ιδιότητες που τη διακρίνουν. Πρόκειται για μια δημοφιλή γλώσσα σε προβλήματα μηχανικής μάθησης, γεγονός που στηρίζεται σε ένα μεγάλο αριθμό λόγων. Η γλώσσα Python διαθέτει σαφή σύνταξη. Επιπλέον καθιστά το χειρισμό κειμένου μια εξαιρετικά εύκολη διαδικασία. Χάρις στη σαφή της σύνταξη έχει κερδίσει το χαρακτηρισμό του “εκτελέσιμου ψευδοκώδικα”. Η προεπιλεγμένη εγκατάσταση της Python διαθέτει εξ αρχής τύπους δεδομένων υψηλού επιπέδου, όπως λίστες, ακολουθίες, λεξικά, data frames κ.α. Επίσης υποστηρίζει πολλαπλά πρότυπα προγραμματισμού, όπως αντικειμενοστραφή, διαδικαστικό, λειτουργικό και ούτω καθεξής. Όπως αναφέραμε, με την Python είναι εύκολη η επεξεργασία και ο χειρισμός κειμένου, στοιχείο που την καθιστά ιδανική στην επεξεργασία μη αριθμητικών δεδομένων [172].

Αξίζει να αναφερθεί ωστόσο ότι η Python είναι εξίσου δημοφιλής στις επιστημονικές και οικονομικές κοινότητες. Διαθέτει προγραμματιστικές βιβλιοθήκες (SciPy, NumPy κ.α.), οι οποίες επιτρέπουν εργασίες με τη χρήση διανυσμάτων και πινάκων. Με αυτόν τον τρόπο, ο κώδικας που προκύπτει είναι ευανάγνωστος και η μορφή του εμφανίζει ομοιότητες με τη γραμμική άλγεβρα. Καθώς οι παραπάνω βιβλιοθήκες καταρτίζονται με τη χρήση γλωσσών χαμηλότερου επιπέδου (C και Fortran), αυτό έχει ως αποτέλεσμα τη μείωση του απαιτούμενου χρόνου για τους υπολογισμούς. Τα επιστημονικά εργαλεία στην Python συνεργάζονται αποτελεσματικά με ένα εργαλείο σχεδίασης το οποίο ονομάζεται Matplotlib. Με τη χρήση του Matplotlib είναι δυνατός ο σχεδιασμός σε δύο και τρεις διαστάσεις, ενώ είναι σε θέση να χειριστεί τους περισσότερους από τους τύπους σχεδίων που χρησιμοποιούνται στον επιστημονικό κόσμο. Τέλος η Python διαθέτει ένα διαδραστικό περιβάλλον όπου ο χρήστης μπορεί να προβάλει και να επιθεωρήσει στοιχεία του προγράμματος κατά τη διαδικασία της ανάπτυξής του.

Πέρα από την Python υπάρχουν κι άλλες γλώσσες υψηλού επιπέδου, οι οποίες επιτρέπουν τους μαθηματικούς υπολογισμούς με τη χρήση πινάκων, όπως τα προγράμματα Matlab και Mathematica. Πιο συγκεκριμένα το Matlab διαθέτει πλήθος ενσωματωμένων χαρακτηριστικών που διευκολύνουν τις διαδικασίες μηχανικής μάθησης. Το πρόβλημα ωστόσο εντοπίζεται στο υψηλό κόστος που απαιτείται για την αγορά της άδειας χρήσης, σε αντίθεση με την Python, που αποτελεί ένα δωρεάν λογισμικό ανοιχτού κώδικα. Μαθηματικές βιβλιοθήκες για τη χρήση πινάκων είναι διαθέσιμες και σε γλώσσες χαμηλού επιπέδου όπως Java και C, ωστόσο τις περισσότερες φορές αυτό μεταφράζεται σε πολλές γραμμές κώδικα για τη διεκπεραίωση απλών υπολογισμών. Σε αυτό το σημείο υπερτερεί η Python, αφού χαρακτηρίζεται από καθαρό, συνοπτικό και εύληπτο κώδικα. Το βασικότερο μειονέκτημα της έγκειται σε θέματα ταχύτητας. Λόγω της φύσης της γλώσσας δεν είναι

εξίσου γρήγορη με γλώσσες όπως η Java και C. Στα πλαίσια της παρούσας διπλωματικής για τα σύνολα των δεδομένων που εξετάζονται αυτό δεν αποτελεί πρόβλημα. Ωστόσο σε περίπτωση που ο χρόνος εκτέλεσης εμφανίζεται αρκετά αυξημένος, υπάρχει η δυνατότητα κλήσης προγραμμάτων που έχουν καταρτιστεί σε C από την Python, προκειμένου να αυξηθεί η ταχύτητα των υπολογισμών. Εν κατακλείδι πρόκειται για μία γλώσσα υψηλού επιπέδου, η οποία επιτρέπει την ενασχόληση με την ποιοτική σημασία των δεδομένων, μειώνοντας την ανησυχία σχετικά με τον τρόπο που η μηχανή προσεγγίζει τα δεδομένα.

6.1.3 Βιβλιοθήκη Scikit-Learn

Το Scikit-Learn είναι μία δημοφιλής βιβλιοθήκη μηχανικής μάθησης ανοιχτού κώδικα, το συντακτικό της οποίας είναι γραμμένο σε γλώσσα Python [173]. Ξεκίνησε ως έργο το 2007 από το David Cournapeau στα πλαίσια ενός καλοκαιρινού προγράμματος της Google και αργότερα το ίδιο έτος συνεχίστηκε από το Matthieu Bruch. Το 2010 οι Pedregosa, Varoquaux, Gramfort και Michel από το ερευνητικό ίδρυμα INRIA επιχείρησαν την πρώτη δημόσια έκδοσή του. Από τότε ακολούθησαν αρκετές ακόμα κυκλοφορίες που οδήγησαν στην ανάπτυξη του Scikit-Learn.

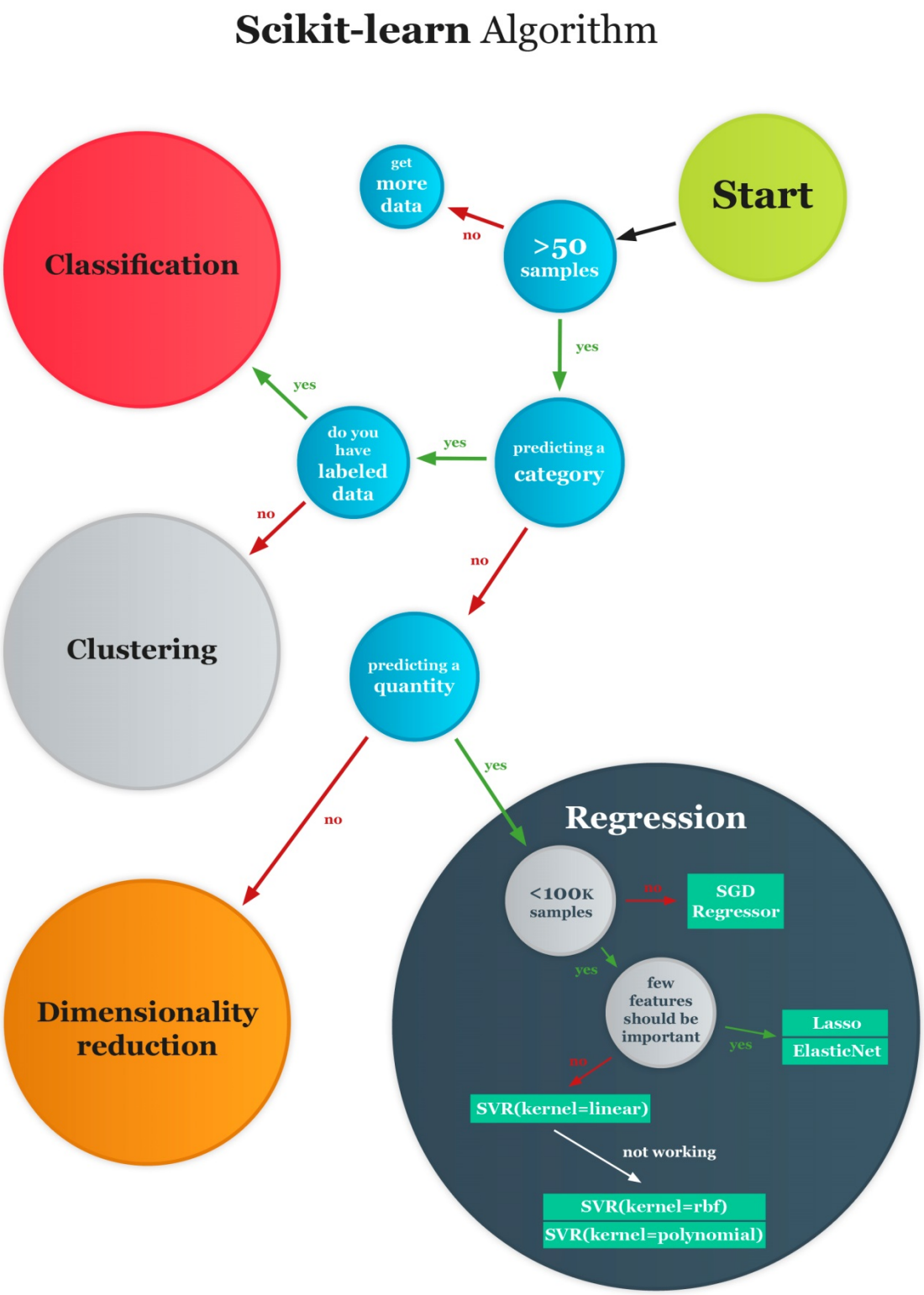
Ο πρωταρχικός στόχος της βιβλιοθήκης είναι να παρέχει μια συλλογή αποτελεσματικών και καθιερωμένων αλγόριθμων μηχανικής μάθησης. Μεταξύ των άλλων περιλαμβάνει αλγόριθμους επιβλεπόμενης και μη επιβλεπόμενης μάθησης, εργαλεία για αξιολόγηση και επιλογή μοντέλων καθώς και εργαλεία για προεπεξεργασία δεδομένων. Χρησιμοποιώντας ένα συνεκτικό περιβάλλον με επίκεντρο την ολοκλήρωση των απαιτούμενων εργασιών, επιτρέπει την εύκολη σύγκριση των εξεταζόμενων μεθόδων για μια δεδομένη εφαρμογή. Στηρίζεται στο επιστημονικό οικοσύστημα της Python, επομένως μπορεί εύκολα να ενσωματωθεί σε εφαρμογές έξω από το παραδοσιακό φάσμα της στατιστικής ανάλυσης δεδομένων. Το σημαντικότερο είναι ότι οι αλγόριθμοι υλοποιούνται σε μια γλώσσα υψηλού επιπέδου κα μπορούν να χρησιμοποιηθούν ως δομικά στοιχεία σε ευρύτερες χρήσεις, όπως η πρόβλεψη κατανάλωσης του ηλεκτρικού φορτίου.

Όπως αναφέραμε το Scikit-Learn παρέχει ένα εύρος αλγόριθμων επιβλεπόμενης και μη μάθησης μέσω μιας συνεκτικής διεπαφής σε Python. Η βιβλιοθήκη είναι βασισμένη στη Scipy (Scientific Python), η οποία περιλαμβάνει τα ακόλουθα πακέτα:

- Numpy: Βασικό πακέτο n −διάστατων πινάκων, για αποτελεσματική αποθήκευση και χειρισμό πολυδιάστατων πινάκων [174]
- SciPy: Θεμελιώδης βιβλιοθήκη για επιστημονικούς υπολογισμούς [175]
- Matplotlib: Ολοκληρωμένη 2D/3D σχεδίαση [176]
- Pandas: Δομές δεδομένων και ανάλυση
- IPython: Ενισχυμένη διαδραστική κονσόλα [177]

Συνήθως το πιο δύσκολο κομμάτι στη διαδικασία επίλυσης ενός προβλήματος μηχανικής μάθησης είναι η εύρεση του κατάλληλου εκτιμητή για τη συγκεκριμένη εργασία. Διαφορετικοί εκτιμητές προσαρμόζονται καλύτερα σε διαφορετικούς τύπους δεδομένων. Το διάγραμμα ροής που παρουσιάζεται στο σχήμα 10 αποτελεί ένα χρηστικό οδηγό για την

επιλογή του κατάλληλου εκτιμητή ανάλογα με το είδος του προβλήματος που καλείται να προσεγγίσει.



Σχήμα 6.1: Διάγραμμα ροής του αλγόριθμου Scikit-Learn [178]

Οι πιο δημοφιλείς ομάδες μοντέλων που παρέχονται από το Scikit-Learn περιλαμβάνουν:

- **Classification**: Τεχνική κατάταξης των δεδομένων σε προκαθορισμένες κλάσεις. Δεν περιορίζεται σε γραμμικά μοντέλα, αλλά διακρίνει την ανάλυση σε μοντέλα Bayes, μηχανές SVM κ.α.

- **Regression**: Μοντέλο επιβλεπόμενης μάθησης σε συνεχείς μεταβλητές. Χρησιμοποιείται για την εύρεση συσχετίσεων μεταξύ μιας εξαρτημένης μεταβλητής και μιας ή περισσότερων ανεξάρτητων μεταβλητών. Βασικό μοντέλο είναι η παλινδρόμηση SVR με τη χρήση πυρήνων εσωτερικού γινομένου.

- **Clustering**: Για την ομαδοποίηση μη κατηγοριοποιημένων δεδομένων με μεθόδους όπως το KMeans κ.α.

- **Dimensionality reduction**: Για τη μείωση του αριθμού των χαρακτηριστικών στα δεδομένα με σκοπό την περιληπτική παρουσίαση, την απεικόνιση και την επιλογή γνωρισμάτων όπως η ανάλυση PCA.

- **Cross validation**: Για την εκτίμηση της απόδοσης των επιβλεπόμενων μοντέλων σε νέα δεδομένα.

- **Feature selection**: Για τον εντοπισμό σημαντικών χαρακτηριστικών, βάσει τον οποίων μπορούν αναπτυχθούν επιβλεπόμενα μοντέλα.

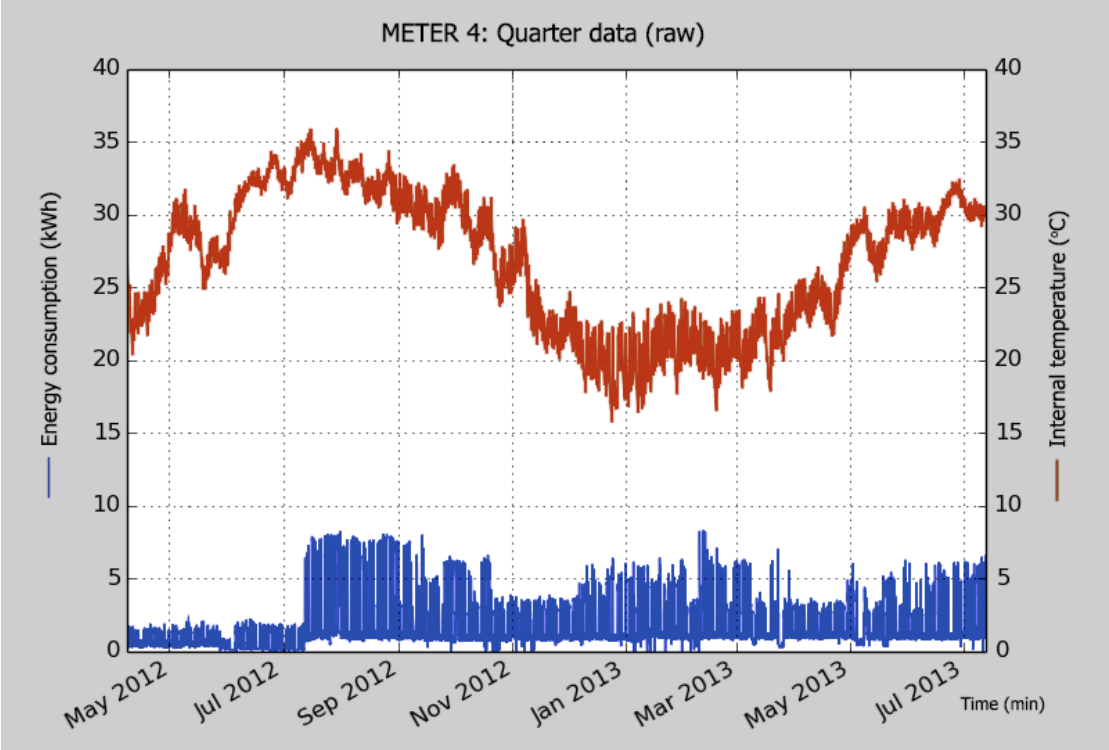
- **Parameter tuning**: Για την εύρεση των βέλτιστων παραμέτρων και την καλύτερη προσαρμογή των μοντέλων.

6.2 Πιλοτική Εφαρμογή

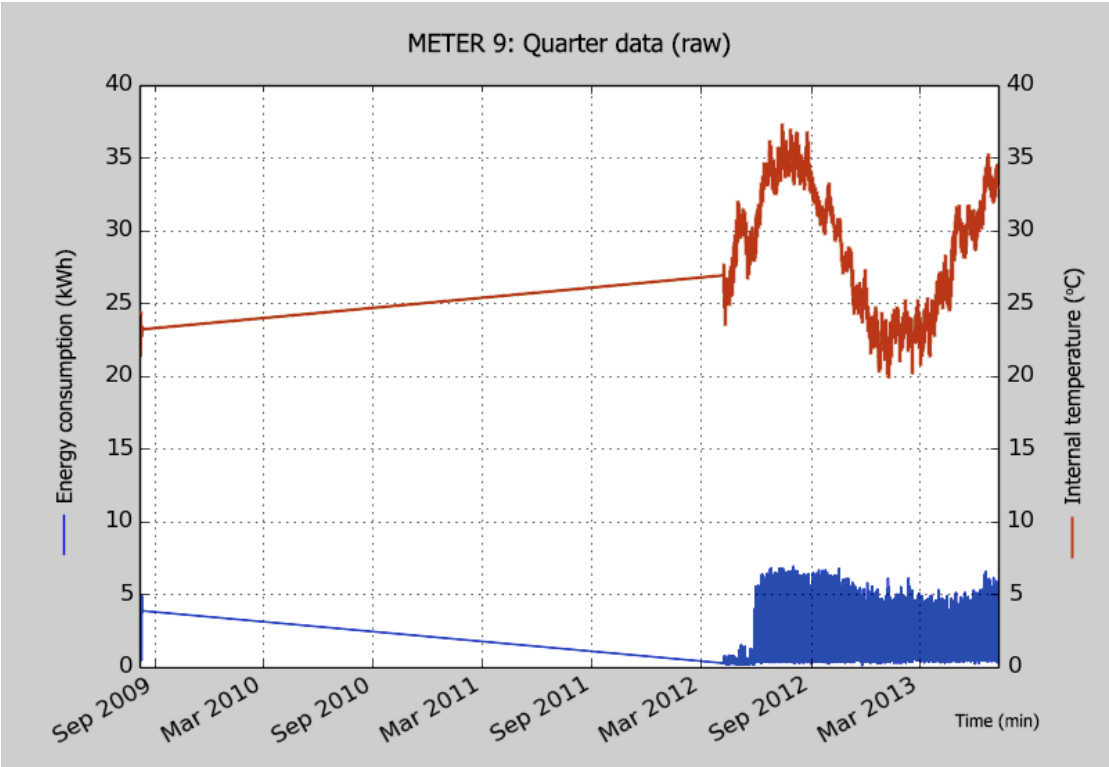
6.2.1 Διερεύνηση Δεδομένων Εισόδου

Για την υλοποίηση της πιλοτικής εφαρμογής χρησιμοποιήθηκαν πραγματικά ενεργειακά δεδομένα πραγματικού χρόνου (real-time energy data). Τα δεδομένα αυτά προέρχονται από "έξυπνους" μετρητές , οι οποίοι είναι εγκατεστημένοι σε δέκα εμπορικά καταστήματα, ένας ανά κατάστημα, στην περιοχή της Αττικής. Με τη βοήθεια των μετρητών είναι δυνατή η συλλογή δεδομένων για την κατανάλωση της ηλεκτρικής ενέργειας (kWh) και την εσωτερική θερμοκρασία (Celsius) ανά τέταρτο.

Τα δεδομένα αυτά παρέχονται σε μορφή ενεργειακών μετρήσεων σε δεκαπεντάλεπτα χρονικά παράθυρα, ακριβώς όπως στέλνονται από τους μετρητές και διατίθενται σε αρχεία CSV για την περαιτέρω επεξεργασία τους με τη βοήθεια υπολογιστικών εργαλείων. Καθώς τα ενεργειακά δεδομένα της πιλοτικής εφαρμογής ανήκουν σε πραγματική επιχείρηση, δεσμεύονται από το εμπορικό απόρρητο και γι' αυτό δεν διατίθενται οι ακριβείς τιμές των δεδομένων εντός της διπλωματικής εργασίας. Διατίθεται ωστόσο η γραφική τους απεικόνιση για την κατανόηση της μεθοδολογίας που ακολουθήθηκε. Επιλέχθηκαν οι μετρητές 4 και 9 για την ανάλυση της μεθοδολογίας, ενώ παρέχονται παράλληλα σχόλια και για τη συμπεριφορά των υπόλοιπων μετρητών. Αρχικά παρουσιάζουμε μια αναπαράσταση των δεδομένων τετάρτου για το ηλεκτρικού φορτίο και την εσωτερική θερμοκρασία των δύο μετρητών.



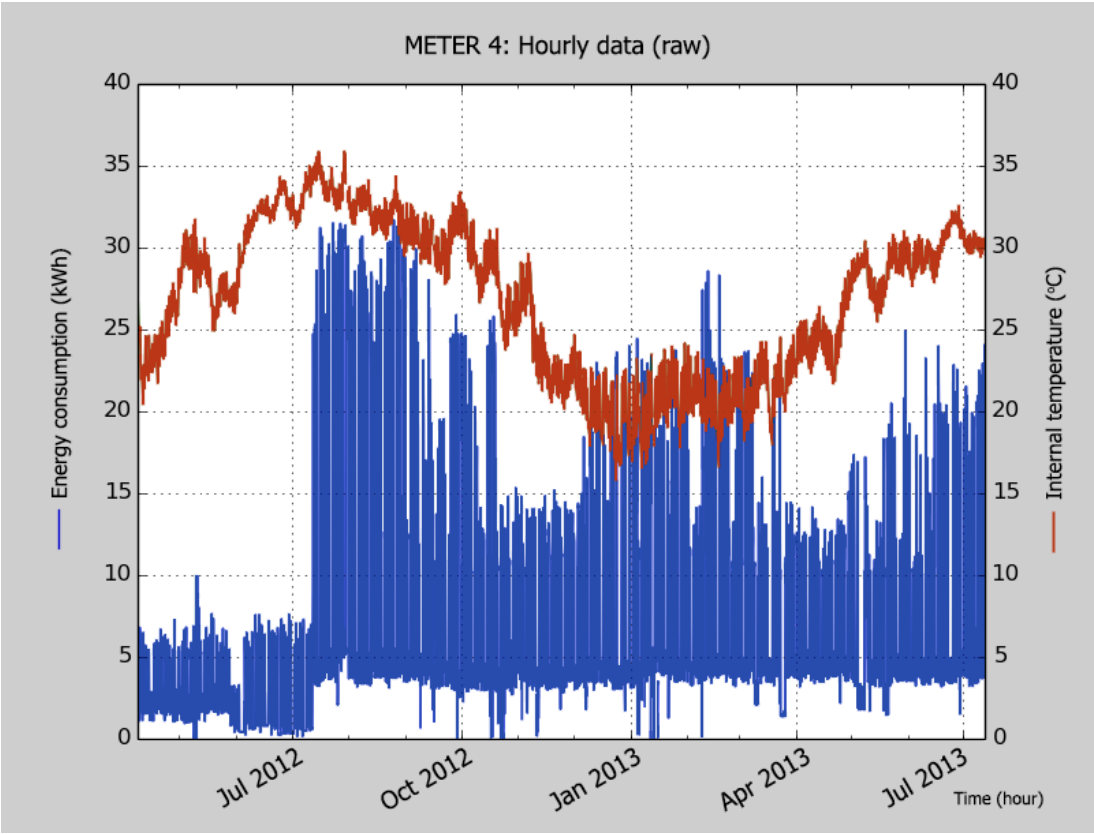
Σχήμα 6.2: Ακατέργαστα δεδομένα τετάρτου του μετρητή 4



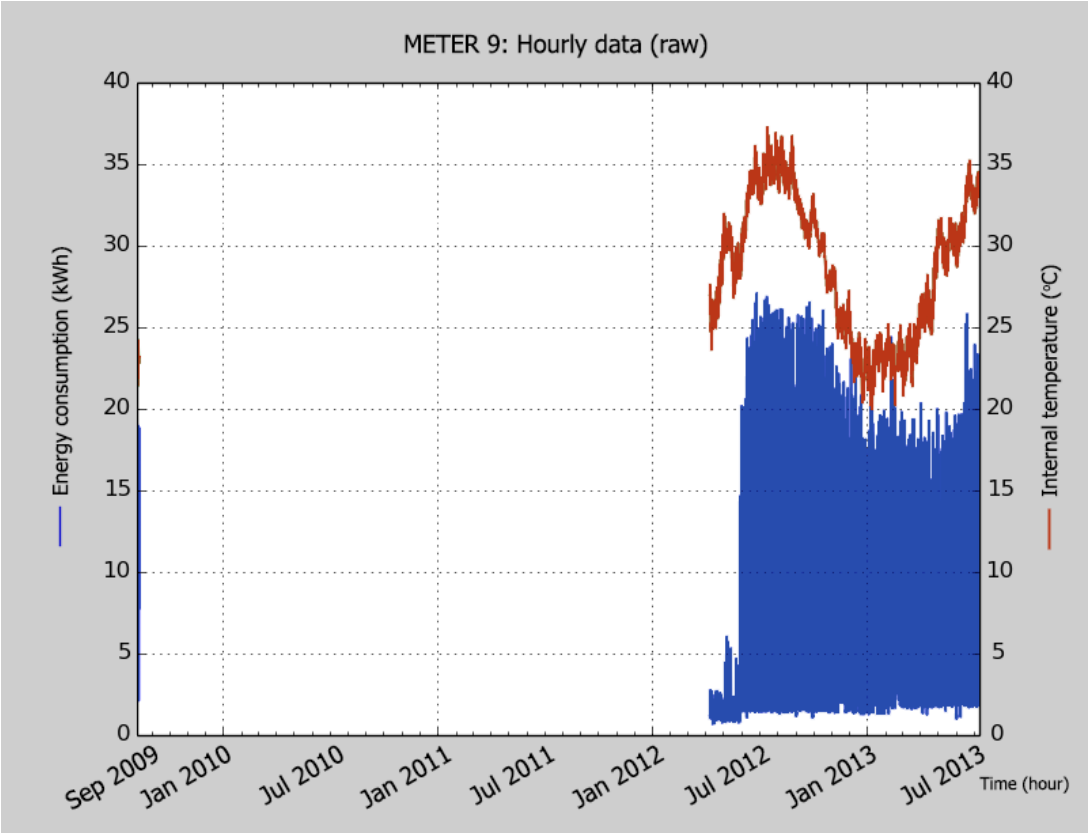
Σχήμα 6.3: Ακατέργαστα δεδομένα τετάρτου του μετρητή 9

Με τη βοήθεια των παραπάνω σχημάτων παρατηρούμε ότι οι δύο μετρητές δεν εγκαταστάθηκαν την ίδια περίοδο. Πιο συγκεκριμένα ο μετρητής 9 εμφανίζει μεμονωμένα δεδομένα για το μήνα Αύγουστο του 2009 ενώ στη συνέχεια ακολουθεί μια μεγάλη περίοδος χωρίς διαθέσιμες μετρήσεις, καθώς ο μετρητής βρισκόταν εκτός λειτουργίας. Τα δεδομένα και των δέκα μετρητών συγκεντρώνονται στο διάστημα (Απρίλιος 2012 – Ιούλιος 2013). Στα πλαίσια της παρούσας ανάλυσης, αυτή είναι η περίοδος που χρησιμοποιείται για τη δημιουργία των μοντέλων πρόβλεψης με τη χρήση μηχανών διανυσμάτων υποστήριξης (SVM). Ο μετασχηματισμός των δεδομένων σε άλλη χρονική βάση αποτελεί μία συνήθη τακτική κατά την είσοδό τους σε μια εφαρμογή πρόβλεψης. Όπως αναφέρθηκε, τα εξεταζόμενα δεδομένα προέρχονται από πραγματικές μετρήσεις ανά τέταρτο σε δέκα εμπορικά καταστήματα. Οι μετρήσεις αυτές μπορούν να συγκεντρωθούν σε ωριαίες και ημερήσιες. Για λόγους πληρότητας της μελέτης, τα μοντέλα παλινδρόμησης με τη χρήση πυρήνων (Linear kernel, Polynomial kernel, Gaussian RBF kernel) πρόκειται να εφαρμοστούν στα δεδομένα και για τα τρία επίπεδα ανάλυσης.

Στη συνέχεια παρατίθενται οι γραφικές απεικονίσεις των μετρητών για ωριαία δεδομένα. Αυτά προέκυψαν με την άθροιση (aggregation) των καταναλώσεων του φορτίου για τα τέσσερα δεκαπεντάλεπτα ανά ώρα και τον υπολογισμό της ωριαίας τιμής για την εσωτερική θερμοκρασία, ως μέσο όρο των τεσσάρων επιμέρους μετρήσεων.

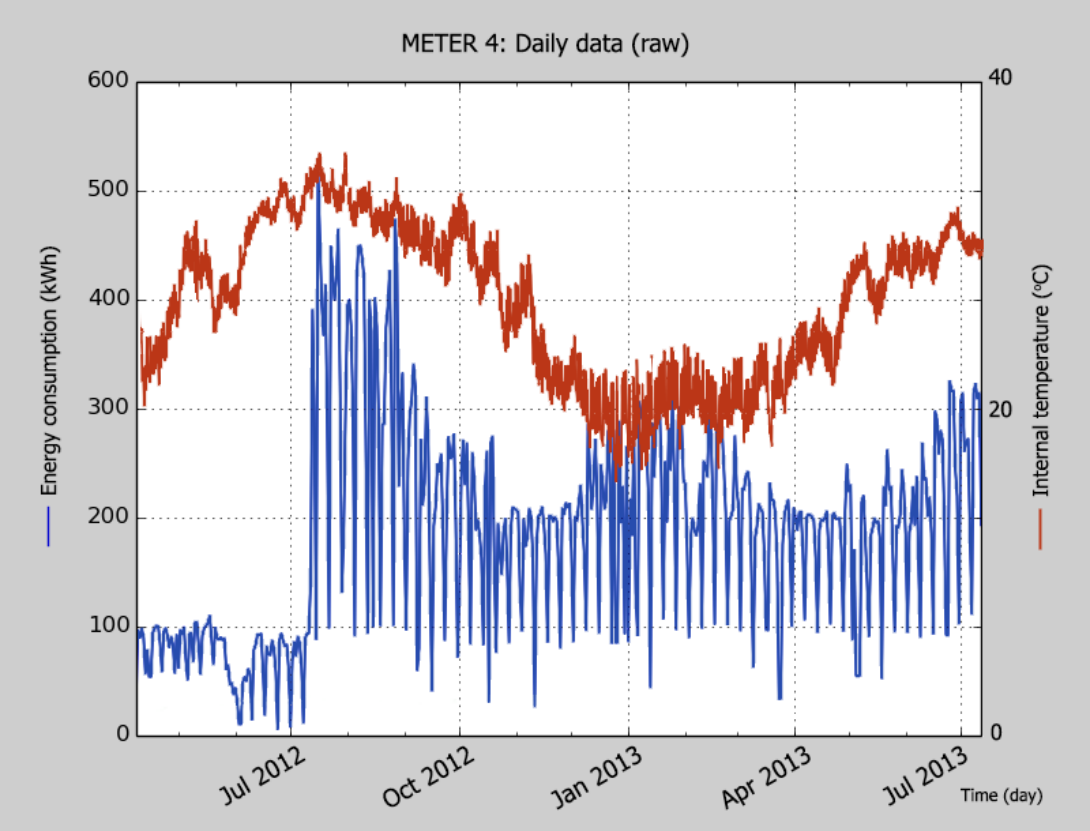


Σχήμα 6.4: Ακατέργαστα ωριαία δεδομένα του μετρητή 4



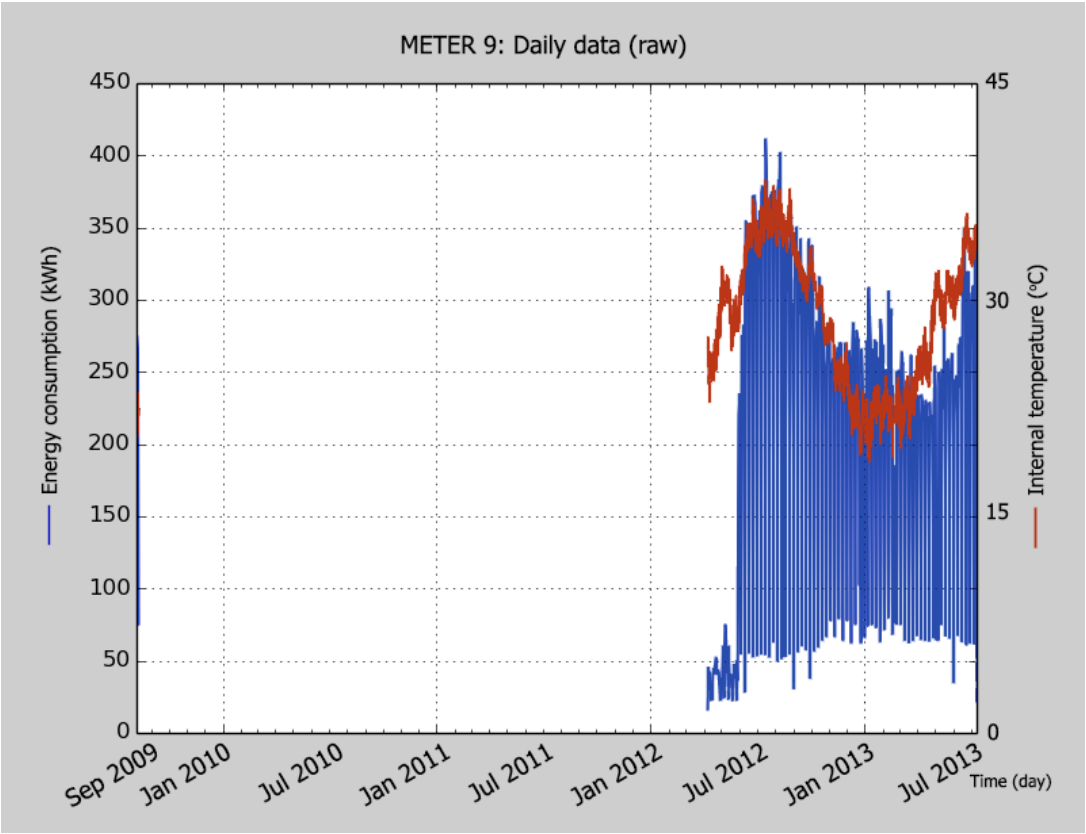
Σχήμα 6.5: Ακατέργαστα ωριαία δεδομένα του μετρητή 9

Παρόμοια διαδικασία ακολουθείται για τη διαμόρφωση των ημερήσιων δεδομένων. Έτσι όσον αφορά στην κατανάλωση φορτίου αθροίζονται (aggregate) οι μετρήσεις των 24*4=96 δεκαπεντάλεπτων που απαρτίζουν κάθε ημέρα της εξεταζόμενης περιόδου. Αντίστοιχα η ημερήσια τιμή της εσωτερικής θερμοκρασίας προκύπτει από το μέσο όρο των 96 μετρήσεων ανά ημέρα. Στις γραφικές παραστάσεις που ακολουθούν, η εσωτερική θερμοκρασία διατηρείται σε ωριαία βάση για εποπτικούς λόγους. Ωστόσο σε επόμενο βήμα παρατίθεται απεικόνιση της εσωτερικής θερμοκρασίας σε ημερήσια βάση, αποσκοπώντας στην ανάδειξη των θετικών και αρνητικών αποτελεσμάτων που προέρχονται από μια τέτοια απλοποίηση.

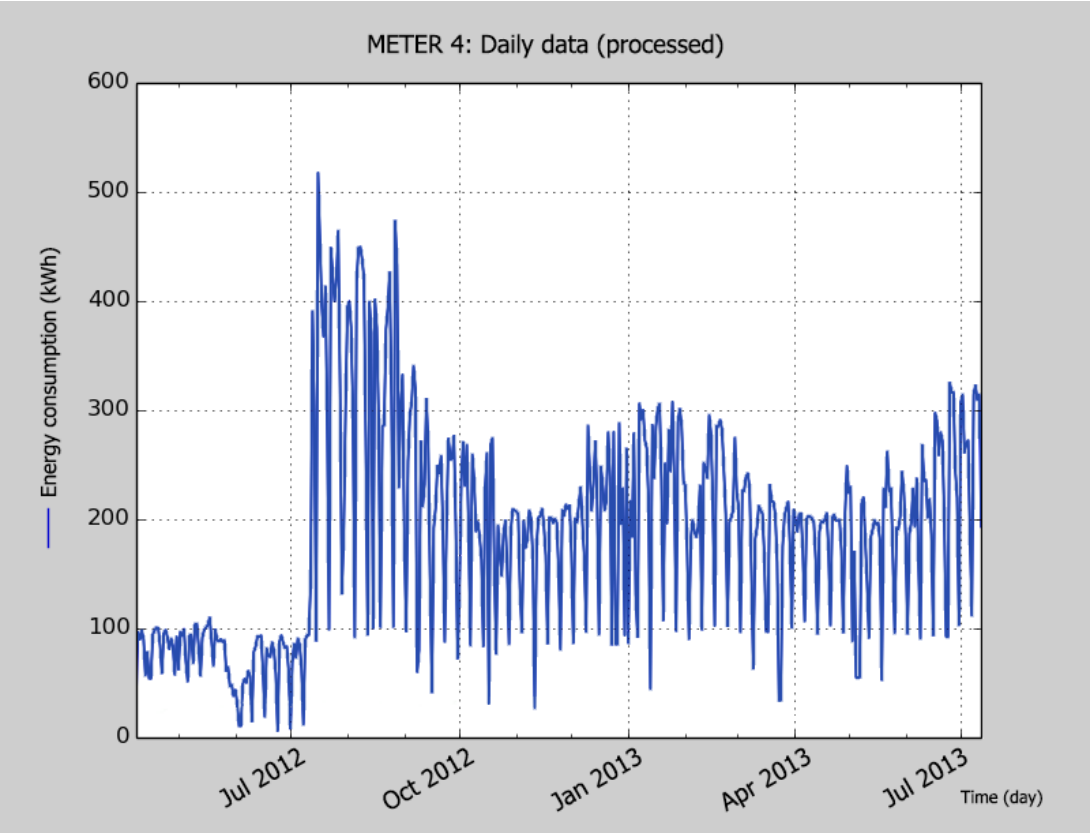


Σχήμα 6.6: Ακατέργαστα ημερήσια δεδομένα του μετρητή 4

Στις διαθέσιμες γραφικές για το μετρητή 9, είναι εμφανής μια περίοδος χωρίς διαθέσιμη πληροφορία, καθώς ο μετρητής βρισκόταν εκτός λειτουργίας. Αντίστοιχη πορεία εμφανίζουν και οι μετρητές 1, 3, 6 και 8 . Καθώς η περίοδος αυτή είναι πολύ μεγάλη, δεν κατατάσσεται στην περίπτωση των μη διαθέσιμων τιμών (missing values) που περιγράφηκαν στην παράγραφο 6.1.1. Η λύση στην παρούσα κατάσταση είναι η αποκοπή του πρώτου διαστήματος μέχρι το Μάρτιο του 2012. Με αυτόν τον τρόπο δημιουργείται μια κοινή περίοδος ανάλυσης όπου οι δέκα μετρητές έχουν στη διάθεσή τους αναλυτικά δεδομένα για την κατανάλωση ηλεκτρικού φορτίου και την εσωτερική θερμοκρασία για κάθε εμπορικό κατάστημα. Η περίοδος (4/2012 – 7/2013) κρίνεται επαρκώς μεγάλη για να δοθεί ιδιαίτερη έμφαση στις ημερήσιες μετρήσεις χάρις στις οποίες μπορεί να αποφευχθεί ανεπιθύμητος θόρυβος που εμπεριέχεται λόγω μεγάλης λεπτομέρειας σε ωριαία δεδομένα και δεδομένα τετάρτου.

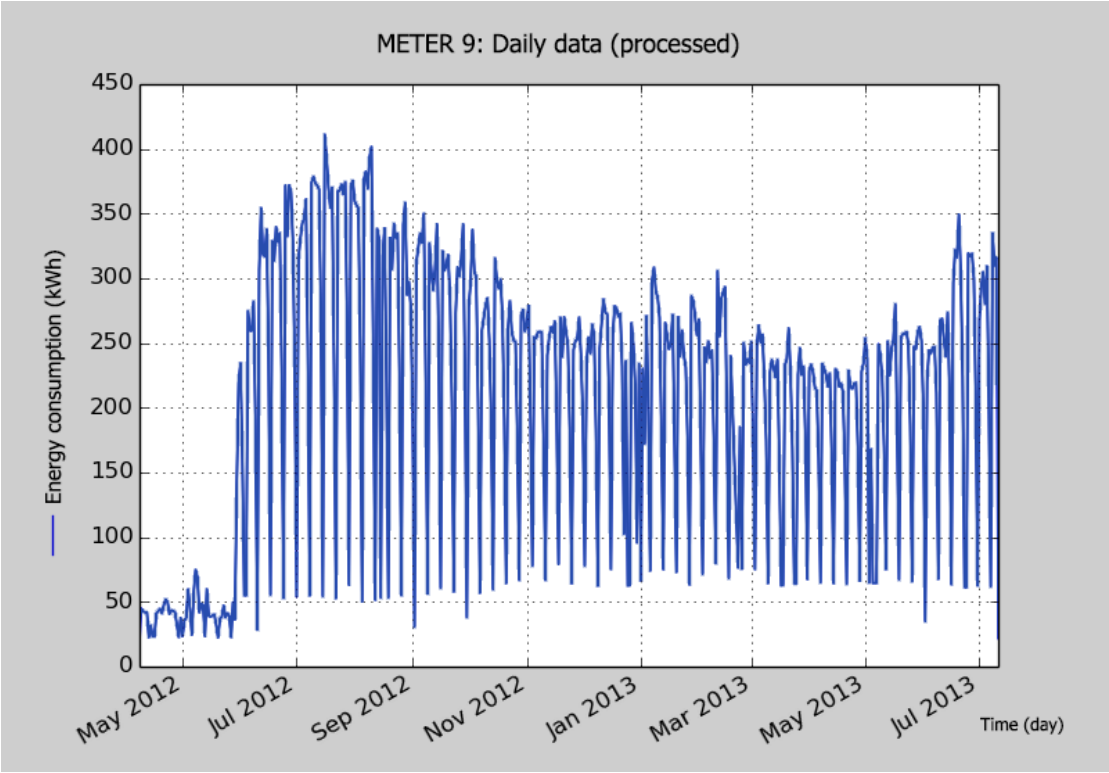


Σχήμα 6.7: Ακατέργαστα ημερήσια δεδομένα του μετρητή 9



Σχήμα 6.8: Ημερήσια δεδομένα ηλεκτρικής κατανάλωσης του μετρητή 4

Στα σχήματα 6.8 και 6.9 παρουσιάζεται η ημερήσια ηλεκτρική κατανάλωση για τους μετρητές 4 και 9. Παρατηρείται αύξηση των ενεργειακών απαιτήσεων κατά τους καλοκαιρινούς μήνες, γεγονός που μπορεί να εντοπιστεί στην κατανάλωση των συστημάτων ψύξης λόγω των υψηλών θερμοκρασιών που εμφανίζονται κατά την περίοδο του καλοκαιριού. Αξίζει ωστόσο να επισημανθεί η διαφορά μεταξύ της μέγιστης κατανάλωσης για τα δύο καταστήματα. Σύμφωνα με τα δεδομένα του μετρητή 4 η κατανάλωση φορτίου για μεμονωμένες ημέρες φτάνει τις 500kWh, ενώ η μέγιστη τιμή για το κατάστημα 9 περιορίζεται στις 400kWh. Αυτό έχει ως αποτέλεσμα και την υψηλότερη διακύμανση στις τιμές κατανάλωσης του καταστήματος 4 που δείχνει να επηρεάζεται πιο έντονα από την εναλλαγή των εποχών του έτους. Παρόμοια συμπεριφορά σύμφωνα με τους μετρητές, εντοπίζεται στα καταστήματα 2, 5 και 10.



Σχήμα 6.9: Ημερήσια δεδομένα ηλεκτρικής κατανάλωσης του μετρητή 9

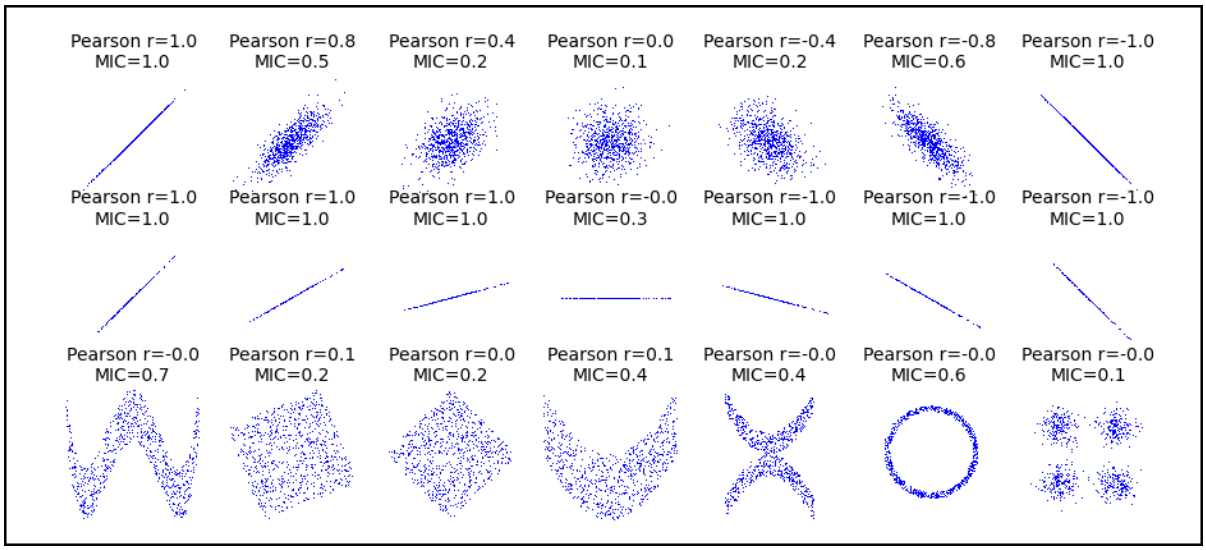
6.2.2 Ανάλυση Συσχετίσεων

Κατά τη διαδικασία ανάλυσης μεγάλων συνόλων δεδομένων, η ανίχνευση νέων συσχετίσεων ή συσχετίσεων (association analysis) που δεν είναι προφανείς, αποτελεί σημαντικό βήμα για την αύξηση της αποτελεσματικότητας των συστημάτων απόφασης. Το αντικείμενο της παρούσας διπλωματικής έγκειται στην ανάλυση ενεργειακών δεδομένων. Επομένως η εύρεση συσχετίσεων και προτύπων που εμπεριέχονται σε αυτά θα μπορούσε να συνεισφέρει σημαντικά στη λήψη αποφάσεων για την αποτελεσματική μοντελοποίηση των δεδομένων. Σύμφωνα με τη βιβλιογραφία έχει αναπτυχθεί πλήθος μεθόδων για το σκοπό αυτό, η πλειοψηφία των οποίων παρουσιάζει ωστόσο περιορισμένη εμβέλεια και δυνατότητες εφαρμογής. Η ανίχνευση γραμμικών συσχετίσεων αποτελεί τον εύκολο στόχο

αυτών των μεθόδων, ενώ αδυνατούν συχνά να ανταπεξέλθουν σε πιο σύνθετους τύπους συσχετίσεων.

Όπως αναφέρθηκε στην προηγούμενη παράγραφο, τα υπό εξέταση δεδομένα προέρχονται από μετρήσεις, σε πραγματικό χρόνο, της κατανάλωσης δέκα εμπορικών καταστημάτων. Σε αυτήν την περίπτωση οι σχέσεις μεταξύ του ηλεκτρικού φορτίου και άλλων μεταβλητών που το επηρεάζουν αποδεικνύονται πολύπλοκες και δεν προκύπτουν διαισθητικά. Γνωστές μετρικές όπως οι συντελεστές Pearson και Spearman δυσκολεύονται να ανιχνεύσουν αυτές τις συσχετίσεις με επιτυχία [179]. Έτσι κρίνεται αναγκαία η αναζήτηση νέων μεθόδων, ικανών να ανιχνεύσουν το σύνολο των γραμμικών ή μη γραμμικών σχέσεων που μπορεί να χαρακτηρίζουν τα δεδομένα.

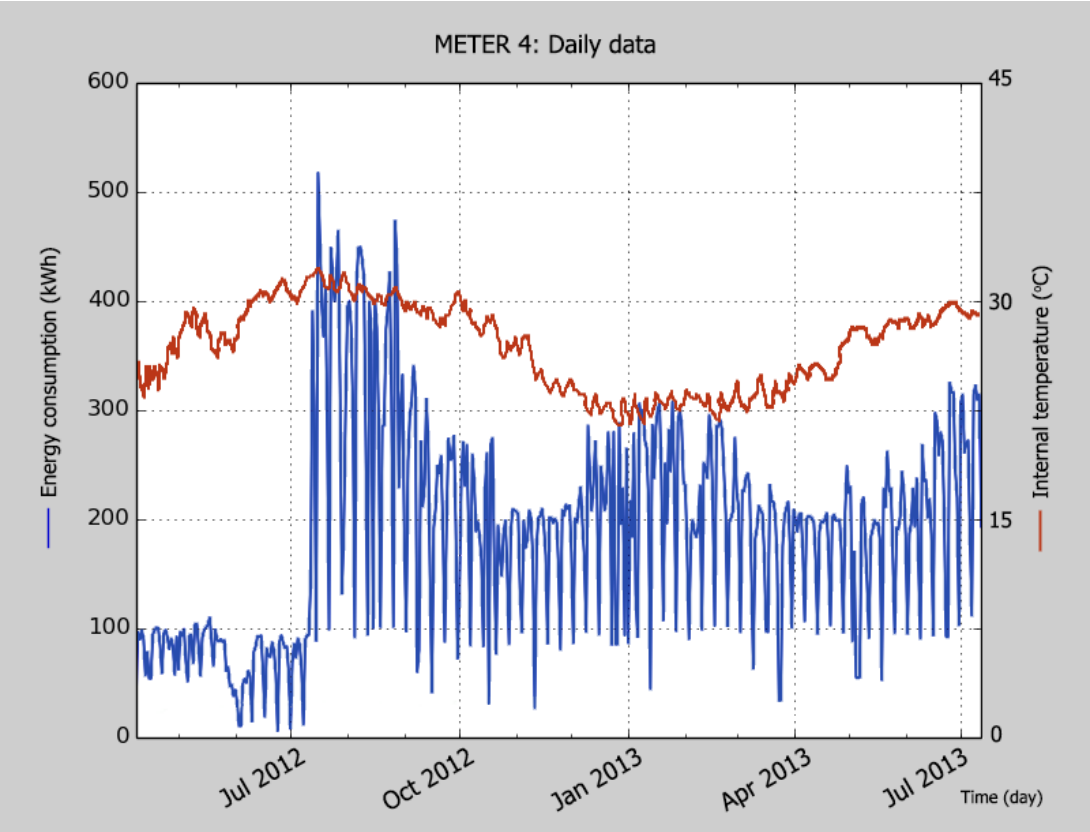
Η μέθοδος MIC, που προτάθηκε από τους Reshef κ.α. [180], είναι μια ενδιαφέρουσα νέα προσέγγιση για την ανακάλυψη μη γραμμικών εξαρτήσεων μεταξύ των ζευγών των μετρήσεων κατά τη διερευνητική εξόρυξη δεδομένων. Ανήκει στην ευρύτερη κατηγορία των στατιστικών στοιχείων MINE (maximal information-based nonparametric exploration), τα οποία χρησιμοποιούνται για τη γρήγορη διερεύνηση των σχέσεων ανάμεσα σε δύο μεταβλητές σε πολυδιάστατα σύνολα δεδομένων. Η μέθοδος MIC (maximal information coefficient) δεν είναι απλά μια στατιστική μέθοδος διάταξης που μετρά την εξάρτηση μεταξύ δύο μεταβλητών, αλλά είναι ικανή να ανιχνεύσει μια μεγάλη ποικιλία λειτουργικών και μη λειτουργικών σχέσεων. Όσον αφορά στην πρώτη περίπτωση, η συνάρτηση θα μπορούσε να είναι οποιουδήποτε τύπου (γραμμική, πολυωνυμική, εκθετική κ.α.).



Σχήμα 6.10: Παραδείγματα συντελεστών συσχέτισης Pearson και MIC [181]

Στο σχήμα 6.10 παρουσιάζονται διάφορα ζεύγη (X, Y) σημείων και οι συντελεστές συσχέτισης Pearson και MIC. Ο συντελεστής Pearson αντανακλά τη μη γραμμικότητα και την κατεύθυνση της γραμμικής σχέσης (1η σειρά), όχι όμως την κλίση της σχέσης (2η σειρά), ούτε μερικές περιπτώσεις μη γραμμικών σχέσεων (3η σειρά), όπως εκθετικές συναρτήσεις, παραβολικές κ.α. Αντίθετα η μέθοδος MIC είναι σε θέση να ανιχνεύσει τέτοιες πολύπλοκες σχέσεις μεταξύ των δεδομένων.

Στο πλαίσιο της παρούσας ανάλυσης, επιχειρείται η αναζήτηση συσχετίσεων μεταξύ του ηλεκτρικού φορτίου και ορισμένων παραγόντων που το επηρεάζουν για ημερήσια βάση. Στο παρακάτω σχήμα παρουσιάζονται τα δεδομένα του μετρητή 4 για την ημερήσια κατανάλωση ενέργειας. Η διαφορά με το σχήμα 6.6 αφορά στην εσωτερική θερμοκρασία, οι τιμές της οποίας προέκυψαν ανά ημέρα από το μέσω όρο των 96 επιμέρους τιμών (ανά δεκαπεντάλεπτο).



Σχήμα 6.11: Ημερήσια δεδομένα του μετρητή 4

Συγκρίνοντας τα σχήματα 6.2, 6.4 και 6.11 παρατηρούμε ότι όσο αυξάνεται το χρονικό παράθυρο της ανάλυσης από ένα τέταρτο, σε μία ώρα και στη συνέχεια σε μία ημέρα, η καμπύλη της εσωτερικής θερμοκρασίας χάνει την έντονη διακύμανσή της. Αυτό οφείλεται στην απώλεια θορύβου που εντοπίζεται αρχικά στη λεπτομερή απεικόνιση των μετρήσεών της ανά τέταρτο έναντι των 480 ημερήσιων μετρήσεων (16 μήνες $*$ 30 τιμές). Μαζί με τον ανεπιθύμητο θόρυβο ωστόσο χάνει και ένα μέρος της περιοδικότητάς της, επομένως μειώνεται η πληροφορία που εμπεριείχε στην αρχική της φάση.

Σε αυτό το σημείο επιχειρείται η ανάλυση συσχετίσεων για τα διαθέσιμα σύνολα δεδομένων από τα δέκα εμπορικά καταστήματα με τη μέθοδο MIC που παρουσιάσαμε παραπάνω. Η ανάλυση αυτή θα λάβει χώρα για τα δεδομένα σε ημερήσια βάση με σκοπό μόνο την εποπτική κατανόηση των εξαρτήσεων ανάμεσα στην εξεταζόμενη μεταβλητή Y (κατανάλωση ενέργειας) και επιλεγμένων χαρακτηριστικών X (παράμετροι) που την επηρεάζουν. Περαιτέρω ανάλυση των παραμέτρων κρίνεται αναγκαία στην περίπτωση όπου οι διαθέσιμες παράμετροι είναι πολυάριθμες (συνήθως πάνω από πενήντα) με σκοπό

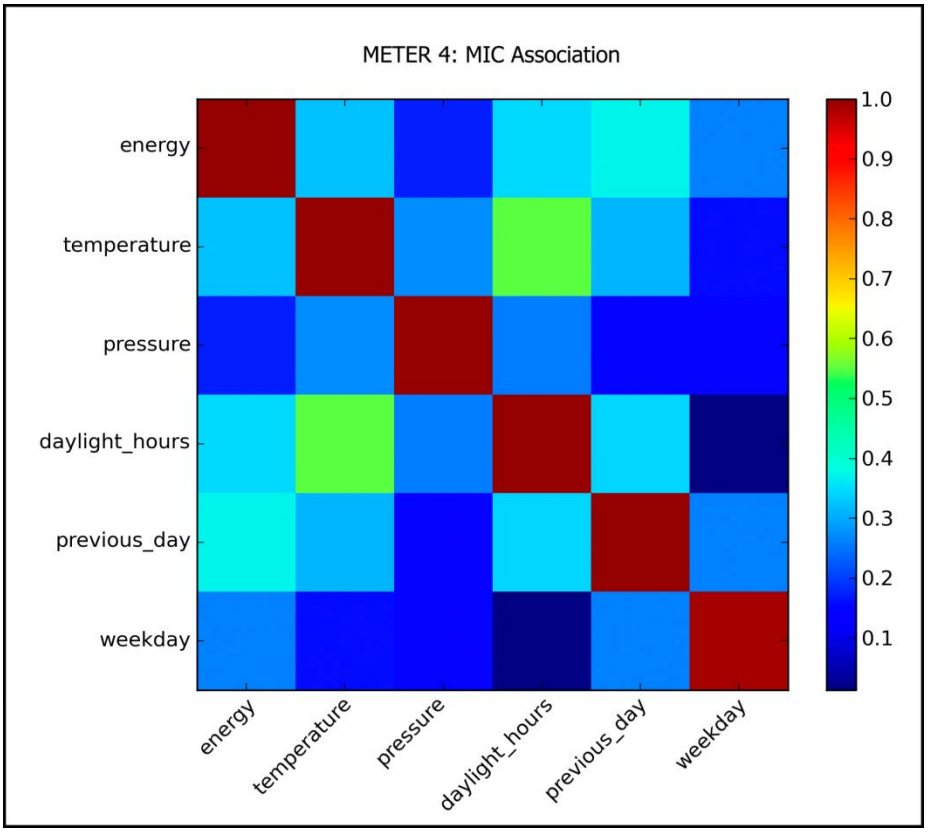
την περιληπτική παρουσίαση, την απεικόνιση ή την επιλογή των κατάλληλων γνωρισμάτων. Αν λάβουμε υπόψη την ύπαρξη συσχετίσεων και μεταξύ των παραμέτρων, γίνεται εύκολα αντιληπτή η ανάγκη αντικατάστασής τους με περιορισμένα σε αριθμό διανύσματα, τα οποία είναι ανεξάρτητα μεταξύ τους. Τα διανύσματα αυτά προκύπτουν με τη χρήση κάποιας μεθόδου για τη μείωση των διαστάσεων, όπως η μέθοδος PCA (principal component analysis) και είναι ικανά να αντιπροσωπεύσουν το πλήθος τον αρχικών συσχετίσεων.

Στα πλαίσια της παρούσας μελέτης, τα δεδομένα που συλλέξαμε από τους δέκα μετρητές αφορούν σε μετρήσεις για την κατανάλωση ενέργειας και την εσωτερική θερμοκρασία στο εκάστοτε κατάστημα. Έπειτα με τη βοήθεια της διαδικτυακής μετεωρολογικής υπηρεσίας Weather Underground (www.wunderground.com), συλλέξαμε πληροφορίες σχετικά με τις καιρικές συνθήκες για την εξεταζόμενη περίοδο. Έτσι μέσω αρχείων CSV εξάγαμε πληροφορίες για την εξωτερική θερμοκρασία και την ατμοσφαιρική πίεση. Πρέπει να επισημανθεί ότι οι μετρήσεις αυτές είναι διαθέσιμες σε ωριαία βάση, οπότε για την περίπτωση της ανάλυσης δεδομένων ανά τέταρτο, έγινε κατάλληλη συμπλήρωση των επιμέρους τιμών. Σε αυτά τα δεδομένα μέσω επαναληπτικής διαδικασίας προστέθηκαν πληροφορίες σχετικά με τις ώρες ηλιοφάνειας ανά ημέρα, το είδος της ημέρας (εργάσιμη ή σαββατοκύριακο) και τις ημέρες εορτών, αργιών κ.α. Τέλος έγινε χρήση της κατανάλωσης που εμφάνισε η εκάστοτε προηγούμενη ημέρα και η αντίστοιχη ημέρα του προηγούμενου μήνα για την ανάδειξη ημερών με παρόμοια συμπεριφορά. Η παραπάνω διαδικασία είχε ως αποτέλεσμα τη δημιουργία εμπλουτισμένων πινάκων με δεδομένα για την εξέταση των τριών χρονικών βάσεων (δεδομένα τετάρτου, ωριαία και ημερήσια δεδομένα). Λόγω ύπαρξης έντονου θορύβου στις δύο πρώτες χρονικές βάσεις, επιλέχθηκε η απεικόνιση συσχετίσεων μόνο για τα ημερήσια δεδομένα.

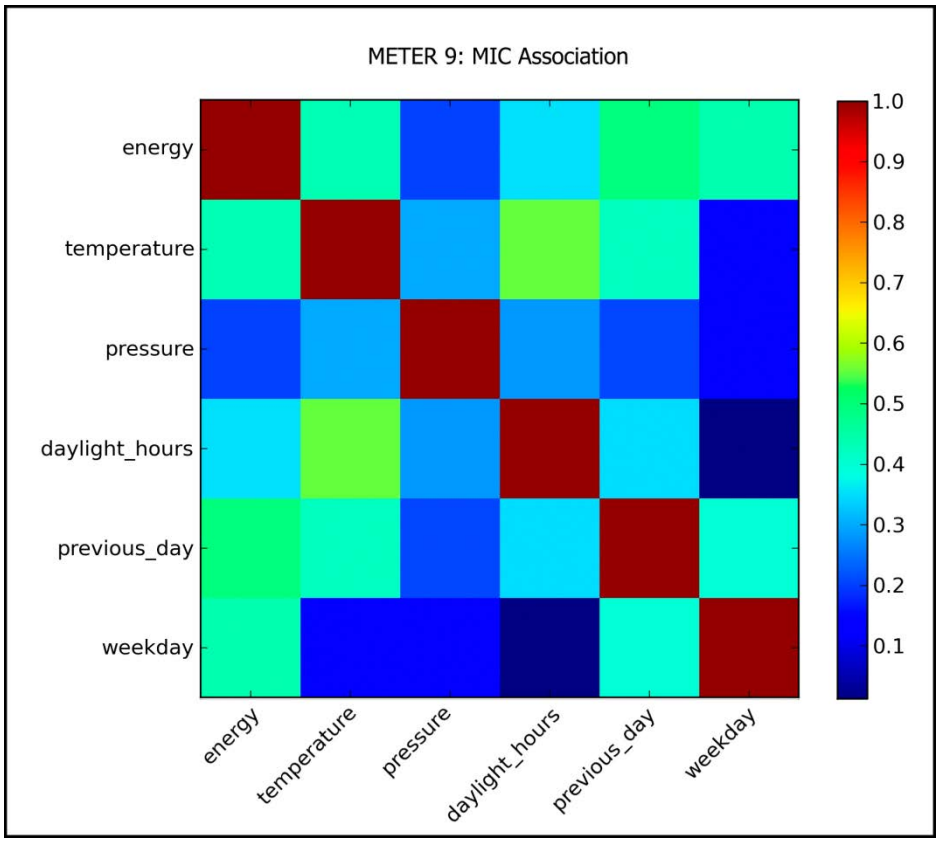
Ως υπό εξέταση τιμή (target value) παραμένει η κατανάλωση ηλεκτρικής ενέργειας (kWh) ενώ οι παράμετροι που συμπεριλάβαμε στη μέθοδο MIC είναι οι ακόλουθες:

- Εξωτερική Θερμοκρασία (Celsius)
- Ατμοσφαιρική Πίεση (hPa)
- Ώρες ηλιοφάνειας / ημέρα (ακέραιος αριθμός)
- Είδος ημέρας (δυαδική μεταβλητή)
- Κατανάλωση προηγούμενης ημέρας (kWh)

Τα αποτελέσματα της ανάλυσης συσχετίσεων μπορούν να χρησιμεύσουν για την εξαγωγή γνώσης σχετικά με τις εξαρτήσεις των εξεταζόμενων κτηρίων. Στη συνέχεια παρουσιάζονται γραφικά με τη βοήθεια του χάρτη θερμότητας (heat map) οι συσχετίσεις για τους μετρητές 4 και 9 που επιλέξαμε στην αρχή της ανάλυσης.



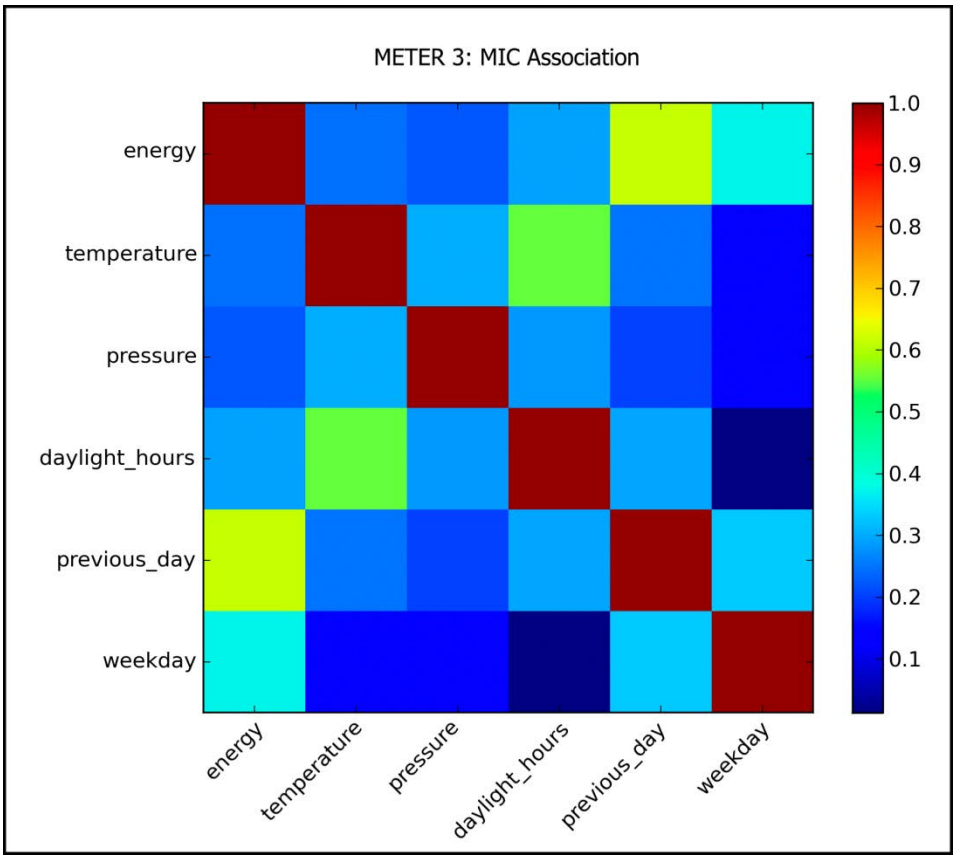
Σχήμα 6.12: Ανάλυση συσχετίσεων του μετρητή 4



Σχήμα 6.13: Ανάλυση συσχετίσεων του μετρητή 9

Ο χάρτης θερμότητας μπορεί να απεικονίσει χρωματικά την ένταση της συσχέτισης για κάθε ζεύγος του πίνακα. Τα αποτελέσματα της μεθόδου MIC εκφράζονται με τιμές που κινούνται σε εύρος $[-1,1]$. Σύμφωνα με την κλίμακα που βρίσκεται στα αριστερά, γίνεται αντιληπτό ότι το σκούρο μπλε χρώμα αντιστοιχεί σε μηδενική τιμή συσχέτισης. Η μηδενική τιμή μεταφράζεται σε απουσία συσχέτισης για ένα ζεύγος μεταβλητών, αφού τα διανύσματά τους είναι ανεξάρτητα, δηλαδή κάθετα μεταξύ τους. Αντίστοιχα παρατηρούμε ότι το σκούρο κόκκινο χρώμα εκφράζει πολύ ισχυρή θετική συσχέτιση. Έτσι εξηγείται η μοναδιαία συσχέτιση που εμφανίζεται στη διαγώνιο του πίνακα.

Στους χάρτες θερμότητας, το ενδιαφέρον επικεντρώνεται στην πρώτη στήλη (ή γραμμή αντίστοιχα), η οποία εκφράζει τη συσχέτιση της ενέργειας με τις παραμέτρους που επιλέξαμε. Στο στάδιο της ανάλυσης συσχετίσεων ο μετρητής 3 παρουσιάζει επίσης ενδιαφέρον, για το λόγο αυτό προσθέσαμε παρακάτω τα αποτελέσματα του μετά τη χρήση της μεθόδου MIC.



Σχήμα 6.14: Ανάλυση συσχετίσεων του μετρητή 3

Οι χρωματικές απεικονίσεις λειτουργούν βοηθητικά σε μια πρώτη επισκόπηση των συσχετίσεων. Είναι σημαντικό να επισημανθεί η ύπαρξη συσχετίσεων μεταξύ των επιλεγμένων παραμέτρων. Όπως αναφέρθηκε, η δημιουργία ανεξάρτητων παραμέτρων απαιτεί μεγαλύτερο πλήθος χαρακτηριστικών (features) και πολύπλοκες μεθόδους ανάλυσης (PCA), οι οποίες ξεφεύγουν από τα όρια της παρούσας μελέτης. Προκειμένου να εξαχθούν ασφαλή συμπεράσματα, παραθέτουμε στη συνέχεια ένα συνοπτικό πίνακα με αριθμητικές τιμές για την πρώτη στήλη των σχημάτων 6.12, 6.13 και 6.14.

	Energy of Meter 3	Energy of Meter 4	Energy of Meter 9
energy	1.00	1.00	1.00
temperature	0.22	0.32	0.44
pressure	0.21	0.17	0.20
daylight_hours	0.29	0.35	0.35
previous_day	0.61	0.38	0.49
weekday	0.37	0.26	0.42

Πίνακας 6.1: Αποτελέσματα συσχετίσεων των μετρητών 3, 4 και 9

Σύμφωνα με τον πίνακα 6.1, προκύπτει ότι η πιο ισχυρή συσχέτιση και στους τρεις μετρητές αφορά στην κατανάλωση της προηγούμενης ημέρας. Ειδικά στην περίπτωση του μετρητή 3 η τιμή φτάνει το 0.61, αριθμός που δηλώνει πολύ έντονη θετική συσχέτιση. Καθώς οι μετρητές είναι εγκατεστημένοι σε δέκα εμπορικά καταστήματα ίδιου τύπου, γίνεται αντιληπτό ότι τα μοτίβα κατανάλωσης της ενέργειας είναι σχετικά σταθερά. Οι ομοιότητες στο μέγεθος των καταστημάτων και στα ωράρια λειτουργίας εξηγούν την έντονη συσχέτιση της κατανάλωσης μεταξύ των ημερών. Αντίθετα η πίεση αποτελεί την παράμετρο με την ασθενέστερη συσχέτιση και στις τρεις περιπτώσεις. Φαίνεται ότι οι μετρήσεις της πίεσης που συλλέχθηκαν μέσω του Wunderground API δεν εισάγουν πληροφορία που να επηρεάζει έντονα την κατανάλωση της ενέργειας. Για το λόγο αυτό, όπως φαίνεται στο επόμενο στάδιο της ανάλυσης, η πίεση δεν περιλαμβάνεται στο τελικό σύνολο δεδομένων κατά την υλοποίηση των προβλεπτικών μοντέλων. Πιο αναλυτικά για τους τρεις εξεταζόμενους μετρητές παρατηρούμε:

- Μετρητής 3

Η παράμετρος της θερμοκρασίας σε αντίθεση με τις δύο επόμενες περιπτώσεις, εμφανίζει ασθενή σχετικά συσχέτιση με τιμή 0.22. Η κατανάλωση ενέργειας φαίνεται έντονα συσχετισμένη με το είδος της ημέρας, ενώ η συσχέτιση με τις ώρες ηλιοφάνειας περιορίζεται στο 0.29. Συγκριτικά με το κατάστημα 9, παρατηρούμε ότι η επίδραση των παραμέτρων, εκτός της προηγούμενης ημέρας, είναι αισθητά μειωμένη στο σύνολό της.

- Μετρητής 4

Η εξωτερική θερμοκρασία του καταστήματος και οι ώρες ηλιοφάνειας που επηρεάζονται άμεσα από την εναλλαγή των εποχών, αποτελούν τις επόμενες πιο έντονα συσχετισμένες παραμέτρους με τιμές 0.32 και 0.35 αντίστοιχα. Το είδος της ημέρας, πρόκειται για μια δυαδική μεταβλητή που δεν εμφανίζει έντονη επίδραση σε αυτήν την περίπτωση.

- Μετρητής 9

Στην περίπτωση του μετρητή 9, η εξωτερική θερμοκρασία και το είδος της ημέρας εμφανίζουν εξίσου δυνατή συσχέτιση με την κατανάλωση της προηγούμενης ημέρας με τιμές 0.44 και 0.42 αντίστοιχα. Έντονη επίδραση διατηρεί ωστόσο και η παράμετρος της

ηλιοφάνειας. Γίνεται επομένως αντιληπτό ότι η κατανάλωση ενέργειας για το κατάστημα 9 είναι άμεσα συσχετισμένη με όλα τα διαθέσιμα χαρακτηριστικά (features) εκτός από την πίεση.

Για λόγους πληρότητας παρατίθεται στη συνέχεια ο συνοπτικός πίνακας με τα αποτελέσματα των συσχετίσεων για τα δεδομένα των υπόλοιπων μετρητών.

	Energy of						
	Meter 1	Meter 2	Meter 5	Meter 6	Meter 7	Meter 8	Meter 10
energy	1.00	1.00	1.00	1.00	1.00	1.00	1.00
temperature	0.39	0.34	0.37	0.33	0.41	0.35	0.39
pressure	0.18	0.21	0.19	0.17	0.21	0.20	0.18
no of daylight hours	0.28	0.34	0.28	0.33	0.25	0.29	0.29
previous day	0.40	0.56	0.49	0.48	0.57	0.38	0.52
weekday	0.31	0.21	0.27	0.23	0.27	0.21	0.28

Πίνακας 6.2: Αποτελέσματα συσχετίσεων των μετρητών 1, 2, 5, 6, 7, 8 και 10

6.2.3 Αλγόριθμοι SVR

Τα αποτελέσματα της προηγούμενης ενότητας δίνουν μια καλή εικόνα σχετικά με το βαθμό συσχέτισης των παραμέτρων με την εξαρτημένη μεταβλητή, δηλαδή την κατανάλωση ηλεκτρικής ενέργειας. Σε αυτό το στάδιο επιχειρείται η ανάπτυξη μοντέλων πρόβλεψης για τα διαθέσιμα σύνολα δεδομένων. Όπως αναφέρθηκε στην εισαγωγή, τα μοτίβα συμπεριφοράς που ακολουθεί η ηλεκτρική κατανάλωση και η πολύπλοκη σχέση που εμφανίζει με τις παραμέτρους που την επηρεάζουν είναι οι βασικότεροι λόγοι που μας οδήγησαν στην αναζήτηση κατάλληλων μεθόδων πρόβλεψης από το πεδίο της μηχανικής μάθησης.

Τα μοντέλα μηχανικής μάθησης εκπαιδεύονται από τα διαθέσιμα δεδομένα με στόχο την εύρεση μιας γενικευμένης λύσης. Με αυτόν τον τρόπο στην πλειοψηφία των περιπτώσεων μπορούν να προσαρμοστούν με επιτυχία στα νέα δεδομένα. Πιο συγκεκριμένα η προσέγγιση που χρησιμοποιείται στα πλαίσια της παρούσας μελέτης, βασίζεται στη χρήση των μηχανών διανυσμάτων υποστήριξης (SVM). Η τεχνική μάθησης SVM στηρίζεται στην ελαχιστοποίηση του διαρθρωτικού σφάλματος. Με την εισαγωγή της συνάρτησης απωλειών του Vapnik (ε-insensitive loss function), επεκτείνεται και στην επίλυση μη γραμμικών προβλημάτων παλινδρόμησης.

Για την υλοποίηση επομένως των προβλεπτικών μοντέλων για τα δέκα εξεταζόμενα εμπορικά καταστήματα, γίνεται χρήση της μεθόδου παλινδρόμησης διανυσμάτων υποστήριξης (SVR). Η μέθοδος αυτή ανήκει στη γενική κατηγορία των μεθόδων πυρήνων [158], [182]. Μια μέθοδος πυρήνα (kernel) είναι ένας αλγόριθμος που εξαρτάται από τα δεδομένα μέσω βαθμωτών γινομένων (scalar or inner products). Σε αυτήν την περίπτωση το

εσωτερικό γινόμενο μπορεί να αντικατασταθεί από μια συνάρτηση πυρήνα (kernel function), η οποία υπολογίζει το βαθμωτό γινόμενο σε ένα χώρο χαρακτηριστικών με υψηλότερες διαστάσεις (υπερεπίπεδο). Με αυτόν τον τρόπο είναι δυνατόν να γενικευθούν μη γραμμικά όρια αποφάσεων με τη χρήση μεθόδων που έχουν σχεδιαστεί για γραμμική παλινδρόμηση. Η παραπάνω διαδικασία είναι γνωστή ως "Τέχνασμα Πυρήνα" (Kernel Trick). Τα βαθμωτά dot προϊόντα είναι το μόνο μέρος της υπολογιστικής διαδικασίας που εξαρτάται από τις διαστάσεις του υπερεπιπέδου. Με την επιλογή μιας συνάρτησης χαρτογράφησης για την οποία μπορεί να δουλέψει το "τέχνασμα πυρήνα", δεν είμαστε αναγκασμένοι να πληρώσουμε το υψηλό υπολογιστικό τίμημα από την επίλυση ενός προβλήματος μάθησης σε ένα χώρο υψηλών διαστάσεων.

Το βασικό ερώτημα που τίθεται σε αυτό το σημείο αφορά στην ύπαρξη γενικού κανόνα για την επιλογή της κατάλληλης συνάρτησης πυρήνα. Πολλές απαντήσεις έχουν δοθεί σε αυτό το ερώτημα, ωστόσο όπως και στα περισσότερα πρακτικά ζητήματα στον τομέα της μηχανικής μάθησης, η επιλογή είναι άμεσα εξαρτημένη από τα δεδομένα. Επομένως απαιτείται η εφαρμογή διαφόρων πυρήνων και η αξιολόγηση των σφαλμάτων στα σύνολα δοκιμής. Η διαδικασία που ακολουθείται συνήθως είναι η παρακάτω: Αρχικά γίνεται η δοκιμή του γραμμικού πυρήνα και στη συνέχεια γίνεται προσπάθεια βελτίωσης της απόδοσης με τη χρήση μη γραμμικών πυρήνων. Ο γραμμικός πυρήνας (linear kernel) ορίζεται ως ο χαμηλότερος βαθμός του πολυωνυμικού πυρήνα και δεν είναι επαρκής όταν εμφανίζονται μη γραμμικές σχέσεις μεταξύ των εξεταζόμενων παραμέτρων. Τα αποτελέσματά του χρησιμοποιούνται συνήθως ως σημείο αναφοράς (benchmark) συγκριτικά με πιο πολύπλοκους τύπους συναρτήσεων. Όσον αφορά στις μη γραμμικές συναρτήσεις, επιλέχθηκαν ο πολυωνυμικός πυρήνας (polynomial kernel) και ο πυρήνας ακτινωτής βάσης (Gaussian RBF kenel). Ο λόγος έγκειται στην ευελιξία που διαθέτουν, η οποία τους επιτρέπει την καλή προσαρμογή στα δεδομένα μέσα σε ένα χώρο υψηλών διαστάσεων. Μεταξύ αυτών των δύο συναρτήσεων, οι μελέτες έχουν αποδείξει ότι ο Gaussian πυρήνας υπερτερεί συνήθως του πολυωνυμικού τόσο σε ακρίβεια όσο και στον απαιτούμενο χρόνο σύγκλισης. Σύμφωνα με την ανάλυση που έγινε στην παράγραφο 5.3.4, οι συναρτήσεις που περιγράφουν τους επιλεγμένους πυρήνες είναι οι ακόλουθες:

Kernel Type	Expression
Linear	$K(x_i, x_j) = x_i \cdot x_j$
Polynomial	$K(x_i, x_j) = (x_i \cdot x_j + 1)^d$
Gaussian Radial Basis Function	$K(x_i, x_j) = exp\left(-\frac{\lVert x_i - x_j \rVert^2}{2\sigma^2}\right)$

Πίνακας 6.3: Εξισώσεις των εξεταζόμενων συναρτήσεων πυρήνα

Κάθε μοντέλο παλινδρόμησης SVR χαρακτηρίζεται από δύο βασικές παραμέτρους, οι οποίες ελέγχουν το βαθμό πολυπλοκότητας των μοντέλων. Η πιο σημαντική παράμετρος είναι η C (trade-off parameter), όπου εκφράζει την ανοχή στο σφάλμα όσον αφορά τα δεδομένα εκπαίδευσης έναντι της απλότητας της επιφάνειας απόφασης. Μια χαμηλή τιμή της παραμέτρου καθιστά την επιφάνεια απόφασης ομαλή, ενώ μια υψηλή τιμής της C στοχεύει στην πολύ καλή προσαρμογή του μοντέλου στα διαθέσιμα δεδομένα αυξάνοντας την πολυπλοκότητα της επιφάνειας απόφασης. Βαρύνουσας σημασίας είναι και η παράμετρος ε, όπου εκφράζει τον επιθυμητό βαθμό ανεκτικότητας στο σφάλμα για όλα τα σημεία εκπαίδευσης. Για την εύρεση των τιμών των δύο παραμέτρων γίνεται χρήση της μεθόδου αναζήτηση πλέγματος (grid search). Σύμφωνα με αυτήν, ορίζεται ένα δισδιάστατο πλέγμα των δύο παραμέτρων και υπολογίζεται η απόδοση του μοντέλου για διάφορους συνδυασμούς τιμών. Τέλος επιλέγεται εκείνο το ζεύγος τιμών για το οποίο εμφανίστηκε η καλύτερη απόδοση. Στη επόμενη παράγραφο παρατίθενται πίνακες με τα αποτελέσματα όλων των μεθόδων και τα βέλτιστα ζεύγη παραμέτρων ανά περίπτωση. Η τιμή της παραμέτρου C φαίνεται να επιδρά πιο έντονα στο τελικό αποτέλεσμα, ενώ η τιμή της παραμέτρου ε κινείται σε ένα μικρό εύρος τιμών και πιο συγκεκριμένα:

$$\varepsilon \in \{0.1, 0.2, \dots, 0.9\}$$

Ως επιθυμητό μέγεθος πρόβλεψης (target value) παραμένει η κατανάλωση ηλεκτρικής ενέργειας (kWh) ενώ οι παράμετροι που συμπεριλάβαμε κατά τη μοντελοποίηση των αλγορίθμων παλινδρόμησης SVR είναι οι ακόλουθες:

- Εσωτερική Θερμοκρασία (Celsius)
- Εξωτερική Θερμοκρασία (Celsius)
- Ώρες ηλιοφάνειας / ημέρα (ακέραιος αριθμός)
- Είδος ημέρας (δυαδική μεταβλητή)
- Κατανάλωση προηγούμενης ημέρας (kWh)
- Κατανάλωσης αντίστοιχης ημέρας της προηγούμενης εβδομάδας (kWh)

Τα μοντέλα SVR υλοποιήθηκαν με βοήθεια της βιβλιοθήκης Scikit-Learn (παράγραφος 6.1.3) σε γλώσσα Python.

6.2.4 Αποτελέσματα Πρόβλεψης

Κυρίαρχη προσέγγιση σύμφωνα με τη βιβλιογραφία της μηχανικής μάθησης για την επιλογή του κατάλληλου μοντέλου έχει αναδειχτεί η μέθοδος διασταυρωμένης επικύρωσης (cross validation). Είναι μια εναλλακτική μέθοδος αξιολόγησης, η οποία εκτιμά πόσο καλά το εκπαιδευμένο μοντέλο μπορεί να λειτουργήσει σε άγνωστα δεδομένα, δηλαδή πόσο καλή είναι η ικανότητα γενίκευσής του. Βασίζεται στην ιδέα ότι η προσδοκία του σφάλματος σε ένα υποσύνολο του δείγματος κατάρτισης που δε χρησιμοποιείται κατά τη φάση εκπαίδευσης, είναι ταυτόσημη με το ίδιο το αναμενόμενο σφάλμα [104]. Είναι μία τεχνική η οποία ελέγχει τον τρόπο με τον οποίο θα χρησιμοποιηθούν τα δεδομένα σε μία στατιστική ανάλυση. Με τη βοήθεια της μεθόδου cross validation γίνεται η αξιολόγηση των

μεθόδων πρόβλεψης για τους τρεις τύπους πυρήνων. Πιο αναλυτικά το εξεταζόμενο σύνολο δεδομένων χωρίζεται σε δύο ή περισσότερα (n) υποσύνολα, όπου η παράμετρος n ορίζεται από το χρήστη. Το ένα υποσύνολο χρησιμοποιείται για την εκπαίδευση του αλγορίθμου (training dataset) και το άλλο για την αξιολόγηση του (testing dataset). Για πιο ασφαλή αποτελέσματα η διαδικασία του cross validation επαναλαμβάνεται για διαφορετικά υποσύνολα κάθε φορά και στο τέλος εκτιμάται ο μέσος όρος από όλες τις επαναλήψεις. Στην παρούσα μελέτη για τη βραχυπρόθεσμη πρόβλεψη της ηλεκτρικής κατανάλωσης στο διάστημα (Απρίλιος 2012 – Ιούλιος 2013) επιλέγεται ημερήσιο χρονικό παράθυρο για την αξιολόγηση των μεθόδων. Στη συνέχεια παρουσιάζονται οι πιο βασικές μέθοδοι διασταυρωμένης επικύρωσης:

- k-fold cross-validation

Σύμφωνα με αυτή τη μέθοδο το διαθέσιμο σύνολο δεδομένων χωρίζεται αρχικά σε k υποσύνολα. Έπειτα κάθε υποσύνολο k χρησιμοποιείται για την επαλήθευση του αλγόριθμου ενώ τα υπόλοιπα $(k-1)$ υποσύνολα αποτελούν τα δεδομένα εκπαίδευσης. Η διαδικασία αυτή επαναλαμβάνεται k φορές, όσες και ο αριθμός των folds. Έτσι κάθε ένα από τα υποσύνολα χρησιμοποιείται μία φορά ως σύνολο αξιολόγησης. Σημαντικό πλεονέκτημα της μεθόδου είναι πως όλα τα υποσύνολα χρησιμοποιούνται και για εκπαίδευση αλλά και για επαλήθευση.

- 2-fold cross-validation

Αποτελεί την απλοποιημένη μορφή της προηγούμενης μεθόδου. Εδώ το διαθέσιμο σύνολο δεδομένων διαιρείται σε δύο υποσύνολα. Κάθε fold απαρτίζεται από τον ίδιο αριθμό αντικειμένων, τα οποία λαμβάνονται με τυχαίο τρόπο. Η εκπαίδευση του αλγόριθμου βασίζεται στο ένα fold ενώ το άλλο χρησιμοποιείται για την αξιολόγηση του. Το πλεονέκτημα αυτής της μεθόδου είναι ότι τα σύνολα εκπαίδευσης και αξιολόγησης είναι αρκετά μεγάλα, καθώς και ότι κάθε διαθέσιμο δεδομένο χρησιμοποιείται τόσο για την κατάρτιση όσο και για την επικύρωση κάθε fold.

- Leave-one-out cross-validation

Σύμφωνα με αυτήν τη μέθοδο η επαλήθευση του αλγόριθμου βασίζεται σε ένα μόνο αντικείμενο του αρχικού συνόλου δεδομένων, ενώ τα υπόλοιπα αντικείμενα χρησιμοποιούνται στο στάδιο της εκπαίδευσης. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να χρησιμοποιηθούν όλα τα αντικείμενα από μία φορά για την αξιολόγηση του αλγόριθμου. Προκύπτει με βάση τη μέθοδο k-fold cross-validation, αν θέσουμε την παράμετρο k ίση με τον αριθμό των αντικειμένων. Έχει αποδειχθεί ότι αυτή η μέθοδος δίνει τα καλύτερα αποτελέσματα. Το κόστος που απαιτεί σε υπολογιστική ισχύ ωστόσο μπορεί να αποτελέσει μειονέκτημα σε περίπτωση αυξημένου αριθμού επαναλήψεων.

Στα πλαίσια της παρούσας μελέτης επιχειρείται η αξιολόγηση των μοντέλων πρόβλεψης με βάση δεδομένα τετάρτου, ωριαία και ημερήσια δεδομένα (παράγραφος 6.2.1). Για τις δύο πρώτες περιπτώσεις χρησιμοποιήθηκε η τεχνική k-fold cross-validation. Το k υποσύνολο αποτελείται από τον απαραίτητο αριθμό δεδομένων που απαρτίζουν μια ημέρα της εξεταζόμενης περιόδου, δηλαδή είτε 96 μετρήσεις τετάρτου είτε 24 ωριαίες μετρήσεις. Στην

περίπτωση των ημερήσιων δεδομένων, η παραπάνω τεχνική μετατρέπεται σε Leave-one-out cross-validation, αφού τα αντικείμενα του αρχικού συνόλου είναι οι ημερήσιες τιμές των δεδομένων.

Στη συνέχεια παραθέτουμε τρεις πίνακες με τα αποτελέσματα των μοντέλων παλινδρόμησης SVR που εξετάσαμε. Για κάθε χρονική βάση υπολογίζεται το σφάλμα πρόβλεψης ανά πυρήνα. Προκειμένου να αξιολογηθεί και να συγκριθεί η απόδοση των εξεταζόμενων μοντέλων υιοθετείται η ρίζα του μέσου τετραγωνικού ποσοστιαίου σφάλματος RMSPE, όπως ορίζεται από τη σχέση (3.10). Το τελικό σφάλμα κάθε μοντέλου από τη μέθοδο σταυρωτής επικύρωσης, προκύπτει ως ο μέσος όρος όλων των επιμέρους σφαλμάτων (Average RMSPE). Αυτή η προσέγγιση αν και αυξάνει το υπολογιστικό κόστος, αξιοποιεί όλα τα διαθέσιμα δεδομένα. Τα τελικά αποτελέσματα σφαλμάτων συνοδεύονται και από τις αντίστοιχες τιμές για τις παραμέτρους C και ε.

Average RMSPE % - Quarter Data

Meter	Linear kernel		Polynomial kernel		Gaussian RBF kernel	
1	18,9625	1,7E+12	15,17	3,7E+12	11,3775	8,3E+12
		0,4		0,4		0,4
2	25,5225	2,1E+12	20,418	9,1E+12	15,3135	2,9E+12
		0,8		0,8		0,9
3	34,03	3,7E+12	27,224	4,7E+12	20,418	1,6E+11
		0,2		0,5		0,7
4	15,63125	4,5E+12	12,505	2,7E+12	9,37875	1,74E+12
		0,2		0,6		0,7
5	33,415	3,6E+12	26,732	9,1E+12	20,049	5,1E+12
		0,1		0,6		0,3
6	36,695	6,7E+12	29,356	3,3E+12	22,017	8,8E+12
		0,9		0,6		0,7
7	19,71075	3,2E+12	15,7686	8,4E+13	11,82645	3,3E+12
		0,2		0,5		0,3
8	20,1925	1,6E+12	16,154	6,5E+12	12,1155	8,1E+12
		0,7		0,9		0,9
9	29,4175	6,6E+12	23,534	7,5E+12	17,6505	3,4E+12
		0,7		0,2		0,2
10	24,2925	6,1E+12	19,434	9,1E+12	14,5755	4,6E+12
		0,7		0,5		0,5

*στα πεδία με πράσινο χρώμα αναγράφονται οι τιμές της παραμέτρου **C** για κάθε μοντέλο
*στα πεδία με γκρι χρώμα αναγράφονται οι τιμές της παραμέτρου **ε** για κάθε μοντέλο

Πίνακας 6.4: Αποτελέσματα μοντέλων SVR για δεδομένα τετάρτου

Στον πίνακα 6.4 παρουσιάζονται οι μέσοι όροι των σφαλμάτων για τις τρεις συναρτήσεις πυρήνων με βάση τα διαθέσιμα δεδομένα τετάρτου από τους μετρητές και τις παραμέτρους που επιλέχθηκαν. Με μία πρώτη επισκόπηση των αποτελεσμάτων προκύπτει εύκολα η διάταξη των μεθόδων ως προς την ακρίβεια πρόβλεψης για κάθε μετρητή. Η συνάρτηση γραμμικού πυρήνα εμφανίζει τα υψηλότερα σφάλματα, όπως ήταν

αναμενόμενο, αφού δυσκολεύεται να προσαρμοστεί στις μη γραμμικές σχέσεις των δεδομένων. Καλύτερα αποτελέσματα δίνει ο πολυωνυμικός πυρήνας, όπου χάρις στην ευελιξία του μειώνει αισθητά τα τελικά σφάλματα. Ωστόσο βέλτιστη μέθοδος αναδεικνύεται η πρόβλεψη με τη χρήση της Gaussian συνάρτησης ακτινωτής βάσης. Στην πλειοψηφία των μετρητών επιτυγχάνει αποδεκτές τιμές σφάλματος ενώ στο σύνολό της εμφανίζεται περισσότερο αποτελεσματική από τους δύο προηγούμενους πυρήνες.

Στην προσπάθεια να ομαδοποιήσουμε τους μετρητές με βάση τα παραπάνω αποτελέσματα για τη μέθοδο Gaussian RBF, διακρίνουμε αρχικά τους μετρητές 1, 4, 7 και 8 με σφάλματα που κινούνται σε ικανοποιητικά όρια, κάτω από το 12%. Ακολουθούν οι μετρητές 2, 9 και 10 με σφάλματα της τάξης 14-17%. Τέλος οι μετρητές 3, 5 και 6 εμφανίζουν πολύ υψηλά σφάλματα που ξεπερνούν το 20% και άρα κινούνται σε μη αποδεκτά όρια. Στην παράγραφο 6.2.2 όπου αναλύθηκαν οι συσχετίσεις για το μετρητή 3, αποδείχτηκε μικρή συσχέτιση της ηλεκτρικής του κατανάλωσης με βασικές παραμέτρους όπως η εξωτερική θερμοκρασία. Αυτό σημαίνει ότι η συνεισφορά των παραμέτρων κατά τη διαδικασία της πρόβλεψης ήταν περιορισμένη. Στοιχείο που εξηγεί ως ένα βαθμό την αδυναμία προσαρμογής των μοντέλων SVR στα δεδομένα του μετρητή και άρα το υψηλό σφάλμα στην πρόβλεψη.

Average RMSPE % - Hourly Data

Meter	Linear kernel		Polynomial kernel		Gaussian RBF kernel	
1	18,5	6,2E+12	14,8	7E+12	11,1	6,8E+12
		0,5		0,8		0,4
2	24,9	1,25E+11	19,92	2,1E+12	14,94	2,7E+12
		0,8		0,6		0,8
3	33,2	6,7E+12	26,56	6,7E+12	19,92	3,1E+12
		0,2		0,5		0,5
4	15,25	4,7E+12	12,2	6,7E+12	9,15	3,6E+12
		0,2		0,4		0,7
5	32,6	3,6E+12	26,08	3E+12	19,56	3,7E+12
		0,6		0,6		0,4
6	35,8	1,07E+13	28,64	1,03E+13	21,48	6,8E+12
		0,8		0,7		0,6
7	19,23	3,2E+12	15,384	3E+12	11,538	3,7E+12
		0,2		0,9		0,3
8	19,7	3,8E+12	15,76	6,6E+12	11,82	7,1E+12
		0,7		0,2		0,9
9	28,7	9,4E+12	22,96	6E+12	17,22	1,8E+12
		0,7		0,7		0,2
10	23,7	3,5E+12	18,96	4,6E+12	14,22	3,2E+12
		0,7		0,3		0,5

*στα πεδία με πράσινο χρώμα ■ αναγράφονται οι τιμές της παραμέτρου **C** για κάθε μοντέλο
*στα πεδία με γκρι χρώμα ■ αναγράφονται οι τιμές της παραμέτρου **ε** για κάθε μοντέλο

Πίνακας 6.5: Αποτελέσματα μοντέλων SVR για ωριαία δεδομένα

Στον πίνακα 6.5 παρουσιάζονται τα σφάλματα για τις τρεις συναρτήσεις πυρήνων με βάση τα ωριαία δεδομένα. Όσον αφορά τα δεδομένα για την ηλεκτρική κατανάλωση της ζητούμενης ημέρας, της προηγούμενης και της αντίστοιχης ημέρας του προηγούμενου μήνα, αυτά προέκυψαν αθροίζοντας (aggregate) τις τέσσερις επιμέρους μετρήσεις ανά ώρα. Αντίστοιχα τα δεδομένα για την εσωτερική και εξωτερική θερμοκρασία υπολογίστηκαν ως ο μέσος όρος των επιμέρους τιμών.

Τα αποτελέσματα στην περίπτωση της ωριαίας χρονικής βάσης εμφανίζουν μια πολύ μικρή βελτίωση συγκριτικά με τα δεδομένα τετάρτου. Βέλτιστη μέθοδος παραμένει η Gaussian συνάρτηση ακτινωτής βάσης ενώ ο γραμμικός πυρήνας δεν βελτιώνει την επίδοση του σε αποδεκτά όρια.

Η ομαδοποίηση των μετρητών σύμφωνα με την ακρίβεια της πρόβλεψης είναι παρόμοια με την προηγούμενη ανάλυση. Μετρητές όπως ο 4, όπου η συσχέτιση της ηλεκτρικής κατανάλωσης με τις υπόλοιπες παραμέτρους είναι ισχυρή, επιτυγχάνουν καλά αποτελέσματα. Αντίθετα μετρητές με ασθενείς συσχετίσεις όπως ο 3 χαρακτηρίζονται από υψηλή τιμή σφάλματος πρόβλεψης κοντά στο 20%.

Το πέρασμα από τα δεδομένα τετάρτου σε ωριαία οδηγεί στη μείωση του θορύβου και επιτρέπει την καλύτερη προσαρμογή των μοντέλων πρόβλεψης στα δεδομένα. Ωστόσο η μείωση των σφαλμάτων δεν είναι αρκούντως μεγάλη για να εισέλθουν όλοι οι μετρητές εντός αποδεκτών ορίων πρόβλεψης.

Στον πίνακα 6.6 που ακολουθεί, παρουσιάζονται τα σφάλματα για τις τρεις συναρτήσεις πυρήνων με βάση τα ημερήσια δεδομένα. Το εξεταζόμενο σύνολο δεδομένων σε αυτήν την περίπτωση προέκυψε με την άθροιση (aggregation) των 96 τιμών ανά ημέρα όσον αφορά στην ηλεκτρική κατανάλωση και το μέσο όρο των αντίστοιχων τιμών όσον αφορά στις παραμέτρους της θερμοκρασίας.

Παρατηρείται αισθητή μείωση των τελικών σφαλμάτων και για τα τρία μοντέλα παλινδρόμησης SVR. Βέλτιστη μέθοδος πρόβλεψης παραμένει ο Gaussian πυρήνας ακτινωτής βάσης. Αξίζει να σημειωθεί ωστόσο πως τόσο ο πολυωνυμικός όσο και ο γραμμικός πυρήνας εμφανίζουν βελτιωμένη επίδοση και πλησιάζουν στα αποδεκτά όρια πρόβλεψης για την πλειοψηφία των μετρητών.

Η ομαδοποίηση των μετρητών σύμφωνα με την ακρίβεια πρόβλεψης της μεθόδου Gaussian RBF διαφοροποιείται ως ένα βαθμό στα ημερήσια δεδομένα. Έτσι από τους εξεταζόμενους μετρητές, περισσότεροι από τους μισούς (μετρητές 1, 2, 4, 7, 8 και 10) εμφανίζουν ικανοποιητικά σφάλματα που κινούνται κοντά στο 10% ή χαμηλότερα. Σημαντική βελτίωση στην ακρίβεια της πρόβλεψης επισημαίνεται για τους μετρητές 3, 5 και 6, όπου ενώ στις δύο προηγούμενες αναλύσεις χαρακτηρίζονταν από υψηλά σφάλματα (20%), εδώ περιορίζονται σε τιμές κοντά στο 15%.

Στους συνοπτικούς πίνακες των αποτελεσμάτων που παρουσιάστηκαν, εκτός από το σφάλμα της πρόβλεψης αν περίπτωση, παρατίθενται και τα ζεύγη των δύο βασικών παραμέτρων που χαρακτηρίζουν τα μοντέλα SVR. Η παράμετρος ε, όπου εκφράζει την ανοχή στο σφάλμα για τα δεδομένα εκπαίδευσης κινείται σε περιορισμένο εύρος τιμών

Average RMSPE % - Daily Data

Meter	Linear kernel		Polynomial kernel		Gaussian RBF kernel	
1	13,468	4,2E+12	10,7744	2,6E+12	8,0808	1,8E+12
		0,4		0,4		0,5
2	18,1272	1,8E+12	14,50176	2E+12	10,87632	2,2E+12
		0,8		0,8		0,8
3	24,1696	5,5E+12	19,33568	2,9E+12	14,50176	2,3E+12
		0,5		0,5		0,1
4	11,102	2,1E+12	8,8816	2,6E+12	6,6612	3,9E+12
		0,6		0,7		0,2
5	23,7328	3,7E+12	18,98624	2,1E+12	14,23968	3,8E+12
		0,4		0,4		0,4
6	26,0624	7,7E+12	20,84992	9,8E+12	15,63744	4,5E+12
		0,6		0,6		0,8
7	13,99944	4,3E+12	11,199552	3,4E+12	8,399664	2,9E+12
		0,3		0,3		0,6
8	14,3416	5,7E+12	11,47328	2,1E+12	8,60496	1,6E+12
		0,9		0,9		0,7
9	20,8936	4,4E+12	16,71488	7,8E+12	12,53616	2,3E+12
		0,2		0,2		0,7
10	17,2536	2,1E+12	13,80288	2,5E+12	10,35216	9,2E+12
		0,5		0,5		0,9

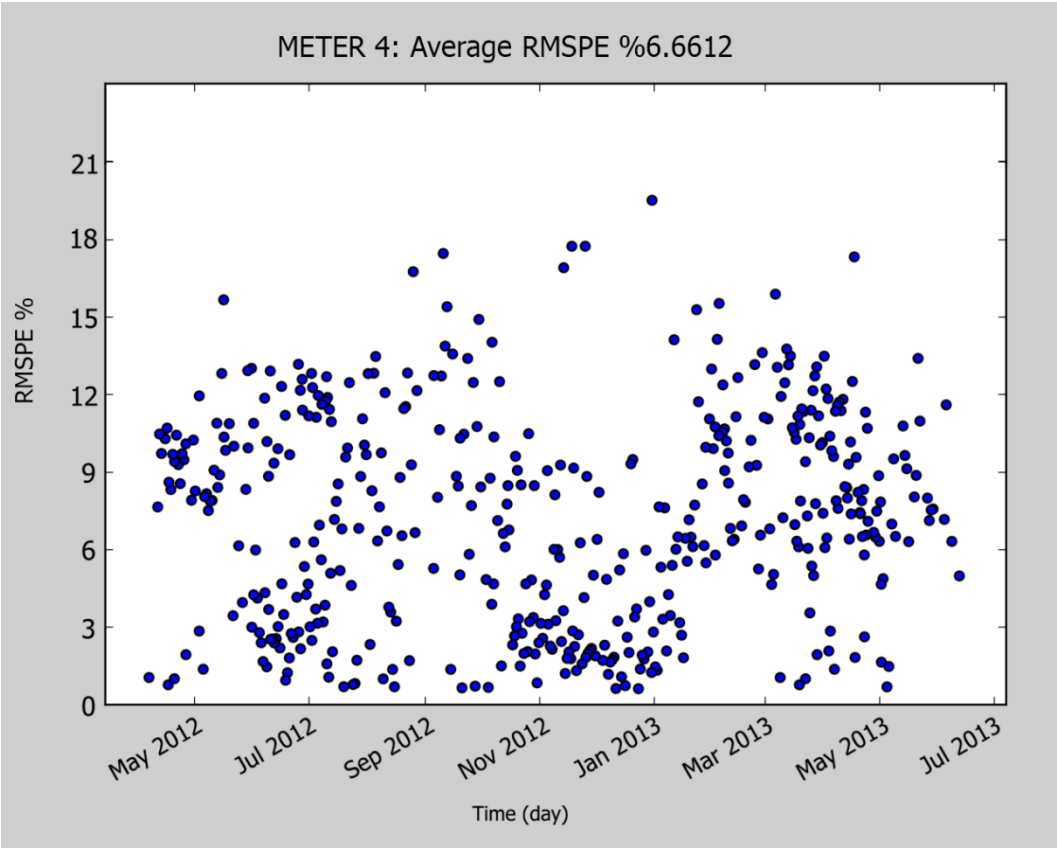
*στα πεδία με πράσινο χρώμα ■ αναγράφονται οι τιμές της παραμέτρου **C** για κάθε μοντέλο
*στα πεδία με γκρι χρώμα ■ αναγράφονται οι τιμές της παραμέτρου **ε** για κάθε μοντέλο

Πίνακας 6.6: Αποτελέσματα μοντέλων SVR για ημερήσια δεδομένα

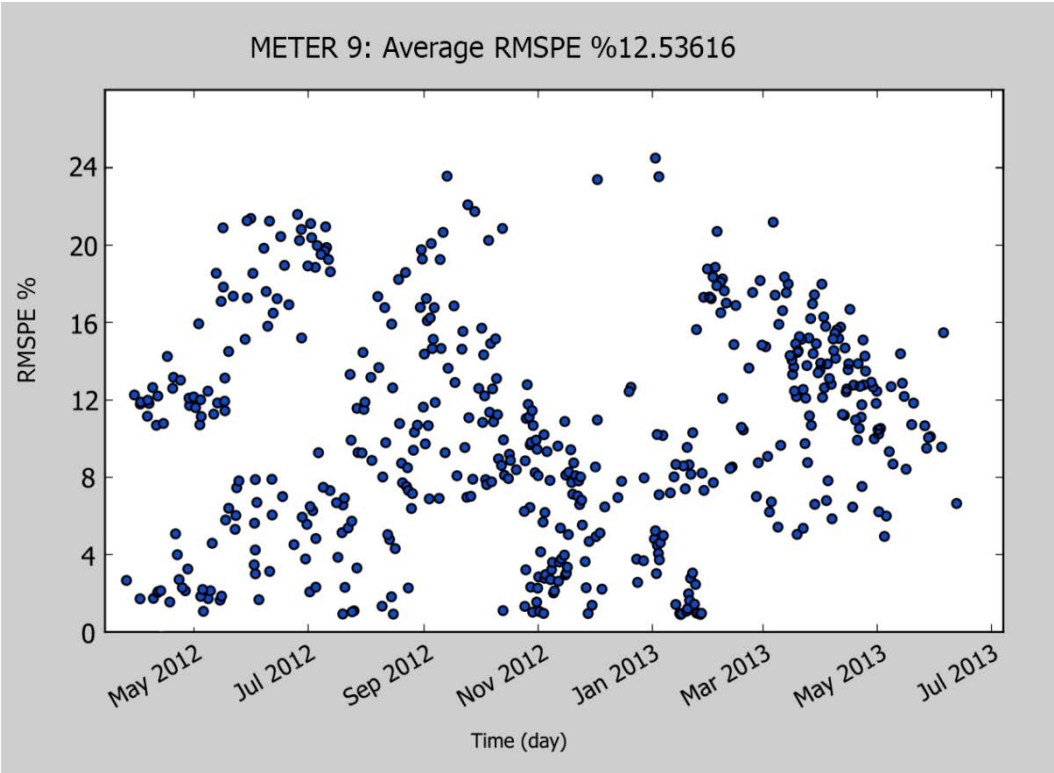
μικρότερο της μονάδας. Αντιλαμβανόμαστε επομένως ότι σημαντικότερο ρόλο κατά τη μοντελοποίηση των επιλεγμένων πυρήνων διαδραματίζει η παράμετρος C. Με τη βοήθεια της επιλέγεται η πολυπλοκότητα της επιφάνειας απόφασης σε σχέση με τον επιθυμητό βαθμό προσαρμογής στα εξεταζόμενα δεδομένα. Παρατηρούμε ότι κινείται σε υψηλά επίπεδα, με την ελάχιστη τιμή της να ισούται με $C_{min} = 1.25 * 10^{11}$ και τη μέγιστη τιμή της να φτάνει στο $C_{max} = 1.03 * 10^{13}$.

Σύμφωνα με τις αναλύσεις της πρόβλεψης για δεδομένα τετάρτου, ωριαία και ημερήσια δεδομένα προκύπτει ότι η τελευταία περίπτωση οδηγεί σε μικρότερα σφάλματα πρόβλεψης. Κατά τη μετατροπή των δεδομένων σε ημερήσια βάση, μειώνεται η λεπτομέρεια λόγω της συνάθροισης των τιμών και κατ' επέκταση ο θόρυβος που υπεισέρχεται στα δεδομένα από τις διακυμάνσεις της κατανάλωσης της ηλεκτρικής ενέργειας. Στο επόμενο κεφάλαιο παρουσιάζονται συνοπτικά τα κυριότερα συμπεράσματα που προέκυψαν από την παρούσα μελέτη και προτείνονται ορισμένες βελτιώσεις που θα μπορούσαν να οδηγήσουν ενδεχομένως σε περαιτέρω μείωση των σφαλμάτων πρόβλεψης.

Στη συνέχεια παρουσιάζονται γραφικά τα αποτελέσματα της τεχνικής cross validation για τους μετρητές 4 και 9, με βάση τα ημερήσια δεδομένα. Τα αποτελέσματα αφορούν το Gaussian πυρήνα ακτινωτής βάσης.



Σχήμα 6.15: Αποτελέσματα cross validation για ημερήσια δεδομένα του μετρητή 4, με τη μέθοδο Gaussian RBF



Σχήμα 6.16: Αποτελέσματα cross validation για ημερήσια δεδομένα του μετρητή 9, με τη μέθοδο Gaussian RBF

Κάθε σημείο της γραφικής αντιπροσωπεύει το σφάλμα RMSPE για μία ημέρα της εξεταζόμενης περιόδου. Το τελικό σφάλμα κάθε μοντέλου, σύμφωνα με το οποίο γίνεται η αξιολόγησή του, προκύπτει ως ο μέσος όρος των επιμέρους σφαλμάτων. Έτσι υπολογίζεται ότι η μέθοδος Gaussian RBF επιτυγχάνει πρόβλεψη για τους μετρητές 4 και 9 με σφάλματα $\mathrm{Av.\,RMSPE}_4 = 6.6612\%$ και $\mathrm{Av.\,RMSPE}_9 = 12.53616\%$ αντίστοιχα.

ΚΕΦΑΛΑΙΟ 7

ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΟΠΤΙΚΕΣ

ΚΕΦΑΛΑΙΟ 7: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΟΠΤΙΚΕΣ

7.1 Συμπεράσματα

Η πρόβλεψη του ηλεκτρικού φορτίου είναι πολύ σημαντική κατά τη διαδικασία λήψης αποφάσεων στον τομέα της ηλεκτρικής ενέργειας, ιδιαίτερα μέσα στο ανταγωνιστικό περιβάλλον που δημιουργήθηκε μετά την απελευθέρωση των αγορών ενέργειας. Επιπλέον η ύπαρξη εκτιμήσεων με ακρίβεια είναι απαραίτητη για τη λειτουργία του συστήματος ηλεκτρικής ενέργειας, για το λόγο αυτό έχουν αναπτυχθεί διάφορες τεχνικές για τη βραχυπρόθεσμη πρόβλεψη της ηλεκτρικής κατανάλωσης. Από τη μέχρι τώρα έρευνα μπορεί να συναχθεί το συμπέρασμα ότι τεχνικές που βασίζονται σε υπολογιστικές μεθόδους κερδίζουν σημαντικά πλεονεκτήματα για την αποτελεσματική χρήση τους έναντι πιο παραδοσιακών μεθόδων. Έτσι το αντικείμενο της διπλωματικής εργασίας εστιάζει στην ανάπτυξη προσεγγίσεων για τα βραχυπρόθεσμη πρόβλεψη της ενεργειακής ζήτησης με τη χρήση μεθόδων μηχανικής μάθησης.

Η παρούσα μελέτη αποτελείται από δύο μέρη. Στο πρώτο μέρος εισάγεται η έννοια της εξόρυξης δεδομένων που αφορά στην εξαγωγή γνώσης από μεγάλα σύνολα δεδομένων σχετικά με πρότυπα, μοτίβα και συσχετίσεις που μπορεί να εμφανίζουν. Στη συνέχεια γίνεται μια βιβλιογραφική ανασκόπηση των βασικών μεθόδων πρόβλεψης με έμφαση στις στατιστικές μεθόδους και τα μοντέλα τεχνητής νοημοσύνης που έχουν αναπτυχθεί για τη βραχυπρόθεσμη πρόβλεψη φορτίου. Επιπλέον παρουσιάζονται οι πιο σημαντικοί παράγοντες που επιδρούν στη συμπεριφορά της ηλεκτρικής κατανάλωσης, όπως οι καιρικές συνθήκες, οι οικονομικοί παράγοντες, το είδος του τελικού χρήστη και τα τυχαία γεγονότα. Τέλος εισάγεται η έννοια της μηχανικής μάθησης που χρησιμοποιεί τα διαθέσιμα δεδομένα για την εκπαίδευση των μοντέλων πρόβλεψης, με πιο γνωστά τα τεχνητά νευρωνικά δίκτυα (ANN) και τις μηχανές διανυσμάτων υποστήριξης (SVM). Οι τελευταίες έχουν αναδειχθεί σε μια νέα και ανταγωνιστική μέθοδο για την πρόβλεψη του ηλεκτρικού φορτίου.

Στο δεύτερο μέρος της μελέτης αναπτύσσεται και εφαρμόζεται μια προσέγγιση για τη βραχυπρόθεσμη πρόβλεψη ενεργειακής ζήτησης με τη χρήση των μηχανών διανυσμάτων υποστήριξης για τρεις συναρτήσεις πυρήνων εσωτερικού γινομένου. Η πιλοτική εφαρμογή βασίζεται σε ενεργειακά δεδομένα πραγματικού χρόνου (real-time data) όπως αυτά συλλέχθηκαν από “έξυπνους μετρητές” εγκατεστημένους σε δέκα εμπορικά καταστήματα στην περιοχή της Αττικής.

Αρχικά επιχειρείται η ανίχνευση συσχετίσεων μεταξύ της ηλεκτρικής κατανάλωσης και ορισμένων βασικών παραμέτρων που την επηρεάζουν. Σύμφωνα με τους πίνακες 6.1 και 6.2 συμπεραίνουμε ότι:

- Η υψηλότερη συσχέτιση για όλα τα καταστήματα αφορά στην ηλεκτρική κατανάλωση της προηγούμενης ημέρας. Αντιλαμβανόμαστε επομένως ότι η συμπεριφορά της κατανάλωσης αν και μη γραμμική, εμφανίζει ωστόσο κάποια σταθερά μοτίβα. Πρόκειται για μια αλυσίδα δέκα εμπορικών καταστημάτων με κοινό ωράριο, συγκεκριμένες βασικές

καταναλώσεις (φωτισμός, θέρμανση/ψύξη, αποθήκες κ.α.), παρόμοιες διαστάσεις κτηρίων και εύρος πελατών χωρίς μεγάλες διακυμάνσεις. Τα παραπάνω στοιχεία αιτιολογούν την ύπαρξη παρόμοιων ενεργειακών απαιτήσεων ανά ημέρα με εξαίρεση κάποιο ειδικό γεγονός.

- Η εξωτερική θερμοκρασία αποτελεί την επόμενη πιο έντονα συσχετισμένη παράμετρο. Καθώς αλλάζει επίπεδα με την εναλλαγή των εποχών φαίνεται ότι επηρεάζει σημαντικά την ηλεκτρική κατανάλωση. Πολύ κοντά βρίσκεται και η παράμετρος της ηλιοφάνειας, η οποία επίσης εξαρτάται από εποχικούς παράγοντες.

- Η παράμετρος που αφορά το είδος της ημέρας κινείται λίγο πιο χαμηλά από τις δύο παραπάνω παραμέτρους. Ωστόσο στην περίπτωση των μετρητών 3 και 9 εμφανίζει ανεβασμένη συσχέτιση συγκριτικά με τη μέση τιμή της.

- Η ατμοσφαιρική πίεσης τέλος εμφανίζει την ασθενέστερη συσχέτιση με την ηλεκτρική κατανάλωση. Οι πληροφορίες που εμπεριέχει η παράμετρος αυτή, δεν επηρεάζουν άμεσα την κατανάλωση. Αυτός είναι και ο λόγος που δεν επιλέχθηκε ως παράμετρος του προβλεπτικού μοντέλου στη συνέχεια.

Στον παρακάτω πίνακα παρουσιάζονται οι παράμετροι που αναλύσαμε και η μέση συσχέτιση που εμφανίζουν για τα εξεταζόμενα καταστήματα.

	Energy of all meters
temperature	0,36
pressure	0,19
daylight_hours	0,31
previous_day	0,49
weekday	0,28

Πίνακας 6.7: Μέση συσχέτιση παραμέτρων με την ηλεκτρική κατανάλωση

Ως χρονική βάση της ανάλυσης επιλέχθηκαν αρχικά τα δεδομένα τετάρτου, στη συνέχεια τα ωριαία δεδομένα και τέλος τα ημερήσια δεδομένα. Στόχος είναι να βρεθεί η κατάλληλη βάση δεδομένων για την αντιμετώπιση προβλημάτων βραχυπρόθεσμης πρόβλεψης φορτίου για την επόμενη ημέρα. Για κάθε χρονική βάση εξετάστηκε η ανάλυση της επίδρασης τριών συναρτήσεων πυρήνα για τη χαρτογράφηση των μεταβλητών εισόδου σε ένα χώρο χαρακτηριστικών με υψηλότερες διαστάσεις. Αρχικά έγινε δοκιμή του γραμμικού πυρήνα (linear kernel), στην συνέχεια του πολυωνυμικού πυρήνα (polynomial kernel) και τέλος χρησιμοποιήθηκε ο Gaussian πυρήνας ακτινωτής βάσης (Gaussian RBF kernel). Αναλυτικά τα αποτελέσματα όλων των μεθόδων για τις τρεις χρονικές βάσεις παρουσιάζονται στους πίνακες 6.4, 6.5 και 6.6. Αυτά συνοψίζονται στον επόμενο πίνακα, από όπου προκύπτουν τα ακόλουθα συμπεράσματα:

Data	Linear kernel	Polynomial kernel	Gaussian RBF kernel
quarter	25,78695	20,62956	15,47217
hourly	25,158	20,1264	15,0948
daily	18,315024	14,6520192	10,9890144

Πίνακας 6.8: Μέση τιμή σφαλμάτων των εξεταζόμενων πυρήνων

- Η παλινδρόμηση SVR με γραμμικό πυρήνα χαρακτηρίζεται από τα υψηλότερα σφάλματα και για τις τρεις χρονικές βάσεις. Αυτό εξηγείται εύκολα λόγω της αδυναμίας της μεθόδου να προσαρμοστεί στις μη γραμμικές σχέσεις που εμφανίζει η κατανάλωση της ηλεκτρικής ενέργειας και ορισμένες παράμετροι του μοντέλου.

- Η συνάρτηση του πολυωνυμικού πυρήνα εμφανίζει σαφώς βελτιωμένα αποτελέσματα (περίπου 5%). Ουσιαστικά ο γραμμικός πυρήνας αποτελεί απλοποίηση του πολυωνυμικού για μηδενική τάξη συνάρτησης. Έτσι ο τελευταίος χάρις στην ευελιξία του προσαρμόζεται καλύτερα στα διαθέσιμα δεδομένα.

- Βέλτιστη μέθοδος πρόβλεψης και για τις τρεις χρονικές βάσεις αναδεικνύεται ο Gaussian πυρήνας ακτινωτής βάσης. Σε σύγκριση με τον πολυωνυμικό και το γραμμικό πυρήνα εμφανίζει βελτίωση της τάξης του 5% και 10% αντίστοιχα.

- Οι βασικές παράμετροι που χαρακτηρίζουν τα τρία μοντέλα SVR είναι οι παράμετροι C και ε. Η πρώτη εμφανίζει ιδιαίτερα αυξημένες τιμές και όπως φαίνεται επηρεάζει σημαντικά την απόδοση των εξεταζόμενων πυρήνων. Αντίθετα η δεύτερη δε φαίνεται να έχει μεγάλη επιρροή στα αποτελέσματα ενώ η τιμή είναι της θετική αλλά παραμένει κάτω από τη μονάδα.

Σε αυτό το σημείο επιχειρείται η κατάταξη των αποτελεσμάτων σύμφωνα με τη χρονική βάση των δεδομένων εισόδου. Στον πίνακα 6.9 παρουσιάζονται τα αποτελέσματα της βέλτιστης μεθόδου πρόβλεψης (Gaussian RBF) για το σύνολο των δέκα μετρητών. Τα συμπεράσματα που προκύπτουν είναι τα ακόλουθα:

- Η πρόβλεψη με βάση τα δεδομένα τετάρτου παρουσιάζει τα υψηλότερα σφάλματα. Αυτό εξηγείται από την ύπαρξη έντονου θορύβου στα δεδομένα λόγω του μικρού χρονικού περιθωρίου μεταξύ τους.

- Τα δεδομένα ωριαίας βάσης οδηγούν σε πολύ μικρή βελτίωση της ακρίβειας πρόβλεψης, περίπου 1% σε σχέση με τα προηγούμενα. Παραμένει μεγάλο το πλήθος των μετρήσεων για την εξεταζόμενη περίοδο που φτάνει τους 16 μήνες.

- Αισθητή είναι η μείωση των σφαλμάτων κατά την πρόβλεψη σύμφωνα με τα δεδομένα ημερήσιας βάσης. Λόγω της συνάθροισης των επιμέρους τιμών σε 480 ημερήσιες μετρήσεις αποφεύγεται ο θόρυβος που εμπεριέχουν τα αρχικά δεδομένα.

Gaussian RBF kernel

Meter	Quarter data		Hourly data		Daily data	
1	11,3775	8,3E+12 0,4	11,1	6,8E+12 0,4	8,0808	1,8E+12 0,5
2	15,3135	2,9E+12 0,9	14,94	2,7E+12 0,8	10,87632	2,2E+12 0,8
3	20,418	1,6E+11 0,7	19,92	3,1E+12 0,5	14,50176	2,3E+12 0,1
4	9,37875	1,74E+12 0,7	9,15	3,6E+12 0,7	6,6612	3,9E+12 0,2
5	20,049	5,1E+12 0,3	19,56	3,7E+12 0,4	14,23968	3,8E+12 0,4
6	22,017	8,8E+12 0,7	21,48	6,8E+12 0,6	15,63744	4,5E+12 0,8
7	11,82645	3,3E+12 0,3	11,538	3,7E+12 0,3	8,399664	2,9E+12 0,6
8	12,1155	8,1E+12 0,9	11,82	7,1E+12 0,9	8,60496	1,6E+12 0,7
9	17,6505	3,4E+12 0,2	17,22	1,8E+12 0,2	12,53616	2,3E+12 0,7
10	14,5755	4,6E+12 0,5	14,22	3,2E+12 0,5	10,35216	9,2E+12 0,9

Πίνακας 6.9: Αποτελέσματα του Gaussian πυρήνα ακτινωτής βάσης

Προκύπτει επομένως ότι η προτεινόμενη χρονική βάση δεδομένων για βραχυπρόθεσμη πρόβλεψη της επόμενης ημέρας είναι η ημερήσια. Όσον αφορά στην προτεινόμενη συνάρτηση πυρήνα για μεγαλύτερη ακρίβεια πρόβλεψης, προτείνεται ο Gaussian πυρήνας ακτινωτής βάσης.

7.2 Προοπτικές και Μελλοντικές Προεκτάσεις

Όπως αναφέρθηκε στην παράγραφο 5.3 οι μηχανές διανυσμάτων υποστήριξης (SVM) βασίζονται στη στατιστική θεωρίας μάθησης. Στόχος τους είναι η ελαχιστοποίηση του διαρθρωτικού ρίσκου που εμφανίζει το προβλεπτικό μοντέλο. Με αυτόν τον τρόπο μπορούν γενικεύουν τα δεδομένα εκπαίδευσης κατά το βέλτιστο τρόπο. Στα πλαίσια της παρούσας ανάλυσης οι τύποι των συναρτήσεων πυρήνα που εξετάστηκαν είναι ο γραμμικός, ο πολυωνυμικός και ο Gaussian πυρήνας ακτινωτής βάσης. Στα πλαίσια επέκτασης θα μπορούσαν να εξεταστούν και τα υπόλοιπα είδη διαθέσιμων πυρήνων (πίνακας 5.1) ενώ ιδιαίτερο ενδιαφέρον θα παρουσίαζε η χρήση συνδυαστικών μοντέλων πρόβλεψης. Σύμφωνα με τη βιβλιογραφία είναι συνηθισμένη η συνδυαστική χρήση των μηχανών διανυσμάτων υποστήριξης με τεχνικές ασαφούς λογικής, γενετικούς αλγόριθμους και τεχνητά νευρωνικά δίκτυα. Σε πολλές εφαρμογές έχει αποδειχθεί ότι η χρήση

συνδυασμού μεθόδων προσφέρει μεγαλύτερη ακρίβεια πρόβλεψης αφού μπορεί να αξιοποιήσει τα πλεονεκτήματα των επιμέρους μοντέλων για να επιτύχει καλύτερη προσαρμογή στα διαθέσιμα δεδομένα.

Τα δεδομένα της πιλοτικής εφαρμογής προέρχονται από "έξυπνους μετρητές" οι οποίοι είναι εγκατεστημένοι σε μια αλυσίδα δέκα εμπορικών καταστημάτων. Αυτοί καταγράφουν τις μετρήσεις τετάρτου για την ηλεκτρική κατανάλωση και την εσωτερική θερμοκρασία κάθε καταστήματος. Όπως αναφέρθηκε παραπάνω, τα μοτίβα χρήσης που εμφανίζει η εμπορική ηλεκτρική κατανάλωση είναι σχετικά σταθερά χωρίς ξαφνικές αλλαγές. Αντίθετα η οικιακή κατανάλωση λόγω της μεταβλητότητας του ανθρώπινου παράγοντα είναι δυνατόν να εμφανίσει περισσότερες διακυμάνσεις. Σε ένα οικιακό κτήριο οι κάτοικοι έχουν τη δυνατότητα να αλλάξουν για παράδειγμα το σημείο λειτουργίας του θερμοστάτη, να χρησιμοποιήσουν όλες τις ηλεκτρονικές συσκευές και πολλά άλλα. Αυτή η απεριόριστη αυτονομία οδηγεί την ηλεκτρική κατανάλωση μιας οικίας σε σποραδικές διακυμάνσεις, τις περισσότερες φορές κατά έναν απρόβλεπτο τρόπο. Επομένως η προσπάθεια μοντελοποίησης και άλλης κλάσης πελατών θα ήταν μια ενδιαφέρουσα προοπτική εφαρμογής της μεθοδολογίας που αναπτύχθηκε.

Όσον αφορά στο κομμάτι της ανάλυσης συσχετίσεων έγινε χρήση ορισμένων παραμέτρων για την καλύτερη μοντελοποίηση της ηλεκτρικής κατανάλωσης. Αυτή έδειξε ότι πληροφορίες που αφορούν στην προηγούμενη κατανάλωση και στη θερμοκρασία για το εξεταζόμενο κτήριο εμφανίζουν σημαντική συσχέτιση με την κατανάλωση του φορτίου. Η προοπτική της χρήσης μιας μεθόδου για τη μείωση διαστάσεων όπως η μέθοδος PCA (principal component analysis) αποτελεί σημαντικό βήμα εξέλιξης της παρούσας εργασίας. Αυτό προϋποθέτει τη συλλογή μεγάλου πλήθους παραμέτρων για τα διαθέσιμα δεδομένα. Οι παράμετροι θα μπορούσαν να σχετίζονται με τις διαστάσεις του κτηρίου, τη γεωγραφική του θέση, την πληρότητα του κτηρίου σε προσωπικό και πελάτες ανά ημέρα κ.α. Με αυτόν τον τρόπο θα μπορούσαν να υπολογιστούν ανεξάρτητα διανύσματα παραμέτρων για την καλύτερη μοντελοποίηση των δεδομένων κατανάλωσης.

Σκόπιμη κρίνεται σε αυτό το σημείο η εξέταση της ακρίβειας των μοντέλων που αναπτύχθηκαν στην παρούσα εργασία για μεγαλύτερο ορίζοντας πρόβλεψης (μεσοπρόθεσμο και μακροπρόθεσμο ορίζοντα). Επιπλέον θα μπορούσε να τροποποιηθούν τα προτεινόμενα μοντέλα ώστε να λαμβάνουν υπόψη και άλλες καταναλώσεις πέρα από την ηλεκτρική, όπως το φυσικό αέριο, το πετρέλαιο κ.α. Με αυτόν τον τρόπο θα ήταν δυνατή μια μοντελοποίηση περισσότερο πλήρης όσον αφορά στις καταναλώσεις ενός εμπορικού κτηρίου. Εξάλλου σε περιπτώσεις κτηρίων που ανήκουν στον πρωτογενή και δευτερογενή τομέα κυριαρχούν οι παραπάνω πηγές ενέργειας.

Τέλος η δημιουργία ενός πληροφοριακού συστήματος που θα συλλέγει τις μετρήσεις της ηλεκτρικής κατανάλωσης από "έξυπνους" μετρητές που είναι εγκατεστημένοι στο κατάστημα και θα κάνει χρήση των προτεινόμενων προσεγγίσεων για την παραγωγή βραχυπρόθεσμων προβλέψεων, αποτελεί μια πολύ σημαντική και πρακτική εφαρμογή της παρούσας εργασίας. Με τη βοήθεια του διαδικτύου θα υπάρχει η δυνατότητα ειδοποίησης του διαχειριστή σε ενδεχόμενα σφάλματα λειτουργίας και γενικότερα η ευαισθητοποίησή του σχετικά με επίπεδα των καταναλώσεών του.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1]. G. Piatetsky-Shapiro and W. J. Frawley, "*Knowledge Discovery in Databases", AAAI (American Association for Artificial Intelligence)/MIT Press, 1991*

[2]. Κ. Χαραλάμπους, Διπλωματική Εργασία με Θέμα «Οικονομικές Επιπτώσεις Ολυμπιακών Αγώνων και Τεχνικές Εξόρυξης Πληροφορίας: Απόπειρες Εξαγωγής Αποτελεσμάτων και Συμπεράσματα», Αθήνα, Ιούλιος 2009

[3]. Γ. Θεοδωρίδη, Ν. Πελέκης, «Αποθήκες Δεδομένων και Εξόρυξης Γνώσης, Εισαγωγή» http://infolab.cs.unipi.gr/pre-eclass/courses/dwdm/slides/9-spatial.pdf

[4]. M. R. Devi, R. Manonmani "Electricity forecasting using data mining techniques in Tamil Nadu and other countries: a survey". International Journal of Emerging Trends in Engineering and Development, 2012

[5]. G. J. Tsekouras, Ch. N. Elias, S. Kavatza, G. C. Contaxis: "A Hybrid Non-Linear Regression Midterm Energy Forecasting Method Using Data Mining", 2003 IEEE Bologna Power Tech Conference, June 23th-26th 2003, Bologna, Italy (με πλήρη κρίση στο κείμενο–paper 0-7803-7967-5/03 -2003 IEEE)

[6]. V. Figueiredo, F. Rodrigues, Z. Vale, "An Electric Energy Consumer Characterization Framework Based On Data Mining Techniques". IEEE Transactions On Power Systems, Vol. 20, No. 2, May 2005.

[7]. C. Shearer, The CRISP-DM model: the new blueprint for data mining, J Data Warehousing (2000); 5:13—22.

[8]. Γ. Τζιράλης. «Αλγόριθμοι Εξόρυξης Πληροφορίας, Διάλεξη 1, Εισαγωγή» https://sites.google.com/site/gtziralis/datamining

[9]. V. Nikolopoulos, G. Mpardis, I. Giannoukos, I. Lykourentzou, and V. Loumos, "Web-based decision-support system methodology for smart provision of adaptive digital energy services over cloud technologies," Software, IET, vol. 5, no. 5, pp. 454–465, 2011.

[10]. U. Fayyad, G. Piatetsky-Shaprio, P. Smyth and R. Uthurusamy, (1996), "Advances in Knowledge Discovery and Data Mining", MIT Press, Cambridge.

[11]. Π. Θεόδωρος –Τεχνικές Εξόρυξης Δεδομένων (Data Mining) Μελέτη Εφαρμογής σε Ελληνική εταιρεία του κλάδου της ένδυσης με τη χρήση του λογισμικού Weka.

[12]. Dunham M.H., (2004)"Data Mining introductory and advanced topics", Prentice Hall.

[13]. Α. Μακρής Εξόρυξη Δεδομένων – Data Mining,: http://amacris.ode.unipi.gr/DST/07_DST_DM.pdf

[14]. Haykin, S., (1999) "Neural Networks: A Comprehensive Foundation", Upper Saddle River, NJ: Prentice Hall, 2nd ed.

[15]. Ripley, B. D. (1994), "Neural Networks and Related Methods for Classification", Journal of the Royal Statistics Society, 56(3), 409-437.

[16]. J. and Kamber M., (2006) "Data Mining: Concepts and Techniques", MorganKaufmann Publishers.

[17]. Bishop, C.M., (1995) "Neural Network for Pattern Recognition", Oxford University Press Inc., New York.

[18]. Balkin, S. & Ord, J., (2000) "Automatic Neural Network Modeling for Univariate Times Series", International Journal of Forecasting, 16, 509-515.

[19]. Zaiane O, (2005) "Machine Learning and Data Mining", Dunham University, Lecture notes.

[20]. Xu R., and Wunsch C. D. II, (2005) "Survey of Clustering Algorithms", IEEE Transactions on Neural Networks, vol. 16, no. 3, pp 645-678.

[21]. Kaur et al., International Journal of Advanced Research in Computer Science and Software Engineering 4(7), July - 2014, pp. 1040-1044

[22]. Roy Billinton, Mahmud Fotuhi-Firuzabad: "A basic framework for generating system operating health analysis", IEEE Transactions on Power Systems, Vol. 9, No.3, August 1994, pp. 1610 – 1617

[23]. Saleh Aboreshaid, Roy Billinton, Mahmud Fotuhi-Firuzabad: "Probalistic transient stability studies using th method of bisection", IEEE Transactions on Power Systems, Vol. 11, No.4, November 1996, pp. 1990 – 1995

[24]. M.H. Nehrir , P.S. Dolan , V.Gerez , W.J. Jameson : "Development and Validation of a Physically-based Computer Model for Predicting Winter Electric Heating Loads", IEEE Transactions on Power Systems, Vol. 10, No.1, February 1995, pp. 266 – 272

[25]. B.Rautenbach, I.E.Lane: "The multi-objective controller: a novel approach to domestic hot water load control", IEEE Transactions on Power Systems, Vol. 11, No.4, November 1996, pp. 1832 – 1837

[26]. Γ.Τσεκούρας: «Καταγραφή & Αναγνώριση Φορτίου, Μελέτη Μέτρων Εξοικονόμησης Ενέργειας & Τιμολογιακών Μέτρων σε μία Χημική Βιομηχανία», Μεταπτυχιακή Εργασία στο μάθημα «Ενεργειακός Προγραμματισμός Ελαχίστου Κόστους», Ε.Μ.Π., Ιούλιος 2000

[27]. Mark H. Gravener, Chika Nwanpka: "Available transfer capability and first order sensitivity", IEEE Transactions on Power Systems, Vol.14, No.2, May 1999, pp. 512 – 518

[28]. H.B.Gooi, D.P.Mendes, K.R.W.Bell, D.S.Kirschen: "Optimal scheduling of spinning reserve", IEEE Transactions on Power Systems, Vol. 14, No.4, November 1999, pp. 1485- 1492

[29]. Δ. Παπαχρήστου: «Συμβολή στην Προσομοίωση Συστημάτων Παραγωγής Ηλεκτρικής Ενέργειας», Διδακτορική Διατριβή, Αθήνα, Ε.Μ.Π., 2000

[30]. Mesut E. Baran, Arthur W.Kelley: "State estimation for real-time monitoring of distribution systems", IEEE Transactions on Power Systems, Vol. 9, No.3, August 1994, pp. 1601 – 1609

[31]. Σπ. Κιαρτζής: «Εφαρμογές τεχνητής νοημοσύνης στη βραχυπρόθεσμη πρόβλεψη φορτίου», Διδακτορική Διατριβή, Θεσσαλονίκη, Α.Π.Θ., 1998

[32]. Kyriakides, E., Polycarpou, M., 2007. Short term electric load forecasting: A tutorial. In: Chen, K., Wang, L. (Eds.), Trends in Neural Computation, Studies in Computational Intelligence, vol. 35. Springer, pp. 391–418 (Chapter 16).

[33]. H. Hahn, S. Meyer-Nieberg, and S. Pickl, "Electric load forecasting methods: Tools for decision making," European Journal of Operational Research, vol. 199, no. 3, pp. 902–907, 2009.

[34]. E. A. Feinberg and D. Genethliou, "Load forecasting," in Applied mathematics for restructured electric power systems. Springer, 2005, pp. 269–285.

[35]. Yang, J., 2006. Power system short-term load forecasting. Ph.D. Thesis, TU Darmstadt.

[36]. Gonzalez-Romera, E., Jaramillo-Moran, M.A., Carmona-Fernandez, D., 2006. Monthly electric energy demand forecasting based on trend extraction. IEEE Transactions on Power Systems 21 (4), 1946–1953.

[37]. Makridakis, Spyros, Steven C. Wheelwright, and Rob J. Hyndman, Forecasting: Methods and Applications, Third edition. John Wiley and Sons, 1998, p.13-15.

[38]. Bakirtzis, A. G., Petridis, V., Kiartzis, S. J., Alexiadis, M. C., Maissis A. H. (1995). A Neural Network Short Term Load Forecasting Model for the Greek Power System. IEEE Transac. on Power Systems, 11(2), 858-863

[39]. Bruhns, A., Deurveilher, G., Roy, J.-S., 2005. A non-linear regression model for mid-term load forecasting and improvements in seasonality. In: 15th Power Systems Computation Conference

[40]. C.W. Gellings. Demand Forecasting for Electric Utilities. The Fairmont Press, Lilburn, GA, 1996.

[41]. E.A. Feinberg, J.T. Hajagos, and D. Genethliou, "Load pocket modeling", Proceedings of the 2nd IASTED International Conference: Power and Energy Systems, p.p. 50-54, Crete, 2002.

[42]. E.A. Feinberg, J.T. Hajagos, and D. Genethliou, "Statistical load modeling", Proceedings of the 7th IASTED International Conference: Power and Energy Systems, p.p. 88-91, Palm Springs, CA, 2003.

[43]. Liu, K., Subbarayan, S., Shoults, R.R., Manry, M.T. "Comparison of very short-term load forecasting techniques", Autom. & Robotics Res. Inst., Texas Univ., Arlington, TX, USA

[44]. Makridakis, S. & Hibon, M. (2000), 'The M3-competition: Results, conclusions and implications', International Journal of Forecasting 16, 451–476.

[45]. R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," International Journal of Forecasting, vol. 22, no. 4, pp. 679–688, 2006.

[46]. Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., & Lewandowski, R., et al. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. Journal of Forecasting, 1, 111 – 153.

[47]. Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. International Journal of Forecasting, 8, 69– 80.

[48]. Chatfield, C. (1988). Apples, oranges and mean square error. International Journal of Forecasting, 4, 515–518.

[49]. Armstrong, J. S. (2001). Evaluating forecasting methods, Chapter 14. In J. S. Armstrong (Ed.), Principles of forecasting: A handbook for researchers and practitioners. Norwell, MA7 Kluwer Academic Publishers.

[50]. Gardner, E. (1990). Evaluating forecast performance in an inventory control system. Management Science, 36, 490–499.

[51]. Makridakis, S. (1993). Accuracy measures: Theoretical and practical concerns. International Journal of Forecasting, 9, 527– 529.

[52]. Goodwin, P., & Lawton, R. (1999). On the asymmetry of the symmetric MAPE. International Journal of Forecasting, 14, 405– 408.

[53]. Swanson, D. A., Tayman, J., & Barr, C. F. (2000). A note on the measurement of accuracy for subnational demographic estimates. Demography, 2, 193–201.

[54]. D. Ahlburg (1992), "A commentary on error measures: Error measures and choice of a forecast method," International Journal of Forecasting, 8, 99-111.

[55]. R. E. Edwards, J. New, and L. E. Parker, “Predicting future hourly residential electrical consumption: A machine learning case study,” Energy and Buildings, vol. 49, pp. 591–603, 2012.

[56]. D. Crawley, J. Hand, M. Kummert, B. Griffith, Contrasting the capabilities of building energy performance simulation programs, Building and Environment 43 (4) (2008) 661–673

[57]. J. Kissock, J. Haberl, D. Claridge, Inverse modeling toolkit (1050rp): numerical algorithms, ASHRAE Transactions 109 (Part 2) (2003) 425–434.

[58]. J. Kreider, J. Haberl, Predicting hourly building energy use: the great energy predictor shootout – overview and discussion of results, ASHRAE Transactions 100 (2) (1994) 1104–1118.

[59]. S. Karatasou, M. Santamouris, V. Geros, Modeling and predicting building’s energy use with artificial neural networks: methods and results, Energy and Buildings 38 (8) (2006) 949–958.

[60]. K. Li, H. Su, J. Chu, Forecasting building energy consumption using neural networks and hybrid neuro-fuzzy system: a comparative study, Energy and Buildings 43 (10) (2011) 2893–2899, ISSN 0378-7788, doi:10.1016/j.enbuild.2011.07.010. [7] J. Kolter, J. Ferreira Jr., A large-scale study

[61]. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., 2009 ASHRAE Handbook – Fundamentals, D&R International, Ltd, 2009. [9] R. Schaller, Moore’s law: past, present and future, IEEE Spectrum 34 (6) (1997)

[62]. S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Pren-tice Hall Series in Artificial Intelligence, Prentice Hall, 2010, URL: http://books.google.com/books?id=8jZBksh-bUMC

[63]. Φ. Πετρόπουλος, Β. Ασημακόπουλος, Επιχειρησιακές προβλέψεις -1η έκδ.- Αθήνα, Συμμετρία, 2011

[64]. Assimakopoulos, V. And Nikolopoulos ,N. (200), The Theta modelQ a decomposition approach to forecasting, International Journal of Forecasting, Vol.16, No4, pp. 521-530

[65]. Arunesh Kumar Singh Ibraheem,Shahida Khatoon, MD Muazzam and D K Chaturvedi, “ An Overview of Electricity Demand Forecasting Techniques”, National Conference on “Emerging Trends in Electrical, Instrumentation & Communication Engineering (ETEIC-2012), pp. 539-549, April 6-7, 2012.

[66]. R.F. Engle, C. Mustafa, and J. Rice. Modeling Peak Electricity

[67]. W. Charytoniuk, M.S. Chen, and P. Van Olinda. Nonparametric Regression Based Short-Term Load Forecasting. IEEE Transactions on Power Systems, 13:725‑730, 1998.

[68]. T. Haida and S. Muto. Regression Based Peak Load Forecasting using a Transformation Technique. IEEE Transactions on Power Systems, 9:1788‑1794, 1994.

[69]. O. Hyde and P.F. Hodnett. An Adaptable Automated Procedure for Short-Term Electricity Load Forecasting. IEEE Transactions on Power Systems, 12:84‑93, 1997.

[70]. S. Ruzic, A. Vuckovic, and N. Nikolic. Weather Sensitive Method for Short-Term Load Forecasting in Electric Power Utility of Serbia. IEEE Transactions on Power Systems, 18:1581 1586, 2003.

[71]. Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis. Forecasting and Control. Holden-Day, San Francisco.

[72]. H. K. Alfares and M. Nazeeruddin, “Electric load forecasting: literature survey and classification of methods,” International Journal of Systems Science, vol. 33, no. 1, pp. 23–34, 2002.

[73]. LIU, K., SUBBARAYAN, S., SHOULTS, R. R.,MANRY,M. T., KWAN, C.,LEWIS, F. L., and NACCARINO, J., 1996, Comparison of very short-term load forecasting. IEEE Transactions on Power Systems, 11, 877±882.

[74]. HUANG, S. R., 1997, Short-term load forecasting using threshold autoregressive models. IEE Proceedings: Generation, Transmission, and Distribution, 144, 477±481.

[75]. ZHAO,H., REN, Z., and HUANG,W., 1997, Short-term load forecasting considering weekly period based on periodical auto regression. Proceedings of the Chinese Society of Electrical Engineers, 17, pp. 211±213.

[76]. Brockwell, P.J., Davis, R.A., 1991. Time Series: Theory and Methods, 2nd ed. Springer-Verlag.

[77]. BARAKAT, E. H., AL-QASSIM, J. M., and AL-RASHED, S. A., 1992, New model for peak demand forecasting applied to highly complex load characteristics of a fast developing area. IEE Proceedings - C, 139, 136±149.

[78]. CHEN, J.-F.,WANG,W.-M., and HUANG, C.-M., 1995, Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting. Electric Power Systems Research, 34, 187±196.

[79]. FAN, J.Y., and MCDONALD, J. D., 1994, A Real time implementation of short-term load forecasting for distribution power system. IEEE Transactions on Power Systems, 9, 988±993.

[80]. Taylor, J.W., De Menezes, L.M., McSharry, P.E., 2006. A comparison of univariate methods for forecasting electricity demand up to a day ahead. International Journal of Forecasting 22, 1–16.

[81]. Ibrahim Moghram, Saifur Rahman: “Analysis and evaluation of five short-term load forecasting techiques”, IEEE Transactions on Power Systems, Vol. 4, No.4, November 1989, pp. 1484 – 1491

[82]. JUBERIAS, G., YUNTA, R., GARCIA MORINO, J., and MENDIVIL, C., 1999, A new ARIMA model for hourly load forecasting. IEEE Transmission and Distribution Conference Proceedings, 1, 314±319.

[83]. ELRAZAZ, Z. S., and MAZI, A. A., 1989, Unified weekly peak load forecasting for fast growing power system. IEE Proceedings Ð C, 136, 29±41.

[84]. BARAKAT, E. H., QAYYUM, M. A., HAMED, M. N., and AL-RASHED, S. A., 1990, Short-term peak demand forecasting in fast developing utility with inherent dynamic load characteristics. IEEE Transactions on Power Systems, 5, 813±824.

[85]. Takeshi Haidda “Regression based peak load using a transformation technique”, IEEE Transactions on Power Systems, Vol. 9, No.4, November 1994, pp.1788-1794.

[86]. Alizera Khotanzad, Rey-Chue Hwang, Alireza Abaye, Dominic Maratukulam: “An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities”, IEEE Transactions on Power Systems, Vol. 10, No.3, August 1995, pp. 1716 – 1722.

[87]. Abdolhosein S.Dehdashti, James R.Tudor, Michael C. Smith: “Forecasting of hourly load by pattern recognition – a deterministic approach”, IEEE Transactions on Power Apparatus and Systems, Vol. 101, No.9, September 1982, pp. 3290 – 3294.

[88]. Alex D.Papalexopoulos, Timothy C.Hesterberg: “A regression-based approach to short-term system load forecasting”, IEEE Transactions on Power Systems, Vol. 5, No.4, November 1990, pp. 1535 – 1547

[89]. G. A. N. Mbamalu and M. E. El-Hawary, “Load Forecasting Via Suboptimal Seasonal Autoregressive Models And Iteratively Reweighted Least Squares Estimation," IEEE Transaction on Power System, Vol.8, pp.343-348, 1992.

[90]. M. Peng, N.F. Hubele, and G.G. Karadi, “Advancement in the application of neural networks for short-term load forecasting”, IEEE Transactions on Power Systems, vol. 7, p.p. 250-257, Issue: 1, 1992.

[91]. S.J. Kiartzis, A.G. Bakirtzis, J.B. Theocharis, and G. Tsagas, “A fuzzy expert system for peak load forecasting: application to the Greek Power System”, Proceedings of the 10th Mediterraean Electrotechnical Conference, vol. 3, p.p. 1097-1100, 2000.

[92]. V. Miranda, and C. Monteiro, ”Fuzzy interference in spatial load forecasting”, Proceedings of IEEE Power Engineering Winter Meeting, vol. 2, p.p. 1063-1068, 2000.

[93]. S.E. Skarman, and M. Georgiopoulos, “Short-term electrical load forecasting using a fuzzy ARTMAP neural network”, Proceedings of the SPIE, p.p. 181-191, 1998.

[94]. V.N. Vapnik, The Nature of Statistical Learnig Theory, New York: Springer Verlag, 1995.

[95]. Haykin, S., (1999) “Neural Networks: A Comprehensive Foundation”, Upper Saddle River, NJ: Prentice Hall, 2nd ed.

[96]. Ripley B.D., (1996) “Pattern Recognition and Neural Networks”, New York: Cambridge University Press.

[97]. Lee, K.Y., Cha, Y.T., and Park, J.H. 1992. “Short Term Load Forecasting using an Artificial Neural Network”. IEEE Transactions on Power Systems. 7(1):124-132.

[98]. M. Vidyasagar, A Theory of Learning and Generalization, Springer-Verlag, 1997.

[99]. N. Christiani and J.S. Taylor. An Itroduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, Cambridge, 2000.

[100]. E. E. Elattar, J. Goulermas, and Q. H. Wu, “Electric load forecasting based on locally weighted support vector regression,” Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 40, no. 4, pp. 438–447, 2010.

[101]. B. Scholkopf, C. Burges, and V. Vapnik, “Extracting support data for a given task,” in Proc. 1st Int. Conf. Knowl. Discovery Data Min., Menlo Park. Menlo Park, CA: AAAI Press, 1995, pp. 252–257.

[102]. B. Ribeiro, “Support vector machines for quality monitoring in a plastic injection molding process,” IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 35, no. 3, pp. 401–410, Aug. 2005

[103]. J. I. Park, S. H. Baek,M. K. Jeong, and S. J. Bae, “Dual features functional support vector machines for fault detection of rechargeable batteries,” IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 39, no. 4, pp. 480–485, Jul. 2009.

[104]. A. J. Smola and B. Schölkopf, “A tutorial on support vector regression,” Statistics and computing, vol. 14, no. 3, pp. 199– 222, 2004.

[105]. V. N. Vapnik, Statistical Learning Theory. New York: Wiley, 1998.
[106]. HSU, Y. Y., and YANG, L.-C., 1992, Design of artificial neural networks for short term load forecasting. IEEE Proceedings Ð C, 138, 407±413.
[107]. AL-ANBUKY, A., BATAINEH, S., and AL-AQTASH, S., 1995, Power demand prediction using fuzzy logic. Control Engineering Practice, 3, 1291±1298.
[108]. SRINIVASAN, D., CHANG, C. S., and LIEW, A. C., 1992, Demand forecasting using fuzzy neural computation, with special emphasis on weekend and public holiday forecasting. IEEE Transactions on Power Systems, 8, 343±348.
[109]. DASH, P. K., LIEW, A. C., and RAHMAN, S., 1995a, Comparison of fuzzy neural networks for the generation of daily average and peak load profiles. International Journal of System Science, 26, 2091±2106.
[110]. MORI, H., and KOBAYASHI, H., 1996, Optimal fuzzy inference for short-term load forecasting. IEEE Transactions on Power Systems, 11, 390±396.
[111]. RAMIREZ-ROSADO, I. J., and DOMINGUEZ-NAVARRO, J. A., 1996, Distribution planning of electric energy using fuzzy models. International Journal of Power and Energy Systems, 16, 49±55.
[112]. CHOW, M., and TRAM,H., 1997, Application of fuzzy logic technology for spatial load forecasting. IEEE Transactions on Power Systems, 12, 1360±1366.
[113]. MASTOROCOSTAS , P. A. THEOCHARIS, J. B., and BAKIRTZIS, A. G., 1999, Fuzzy modeling for short term load forecasting using the orthogonal least squares method. IEEE Transactions on Power Systems, 14, 29±36.
[114]. HO, K., HSU, Y., CHEN, C., LEE, T., LIANG, C., LAI, T., and CHEN, K., 1990, Short term load forecasting of Taiwan power system using a knowledge based expert system. IEEE Transactions on Power Systems, 5, 1214±1221.
[115]. DASH, P. K., DASH, S., RAMA KRISHNA, G., and RAHMAN, S., 1993, Forecasting of a load time series using a fuzzy expert system and fuzzy neural networks. International Journal of Engineering Intelligent Systems, 1, 103±118.
[116]. S. Rahman and O. Hazim. Load Forecasting for Multiple Sites: Development of an Expert System-Based Technique. Electric Power Systems Research, 39:161‑169, 1996.
[117]. BATAINEH, S., AL-ANBUKY, A., and AL-AQTASH, S., 1996, Expert system for unit commitment and power demand prediction using fuzzy logic and neural networks. Expert Systems, 13, 29±40.
[118]. Yiannis, K. & Poulicos, P. (1857). Forecasting traffic flow conditions in an urban Network-comparison of multivariate and univariate approaches. Transportation Research Record 1857, 2003:74-84.
[119]. Hastie, T. Tibshirani, R., Friedman, J. (2001). The Elements of Statistical Learning. Springer Series in Statistics. Springer-Verlag.
[120]. Breiman, L. (1993). Classification and Regression Trees. Boca Raton, FL: Chapman & Hall.
[121]. H. Mori, N. Kosemura, K. Ishiguro and T. Kondo, "Short-term Load Forecasting with Fuzzy Regression Tree in Power Systems," Proc. of 2001 IEEE Systems, Man, and Cybernetics(SMC) Meeting(CD-ROM), Tucson, AZ, USA, Oct. 2001.

[122]. YANG, H.-T., HUANG, C.-M., and HUANG, C.-L., 1996, Identification of ARMAX model for short term load forecasting: an evolutionary programming approach. IEEE Transactions on Power Systems, 11, 403±408.

[123]. LEE,D.G.,LEE,B.W., and CHANG, S.H., 1997, Genetic programming model for long-term forecasting of electric power demand. Electric Power Systems Research, 40, 17±22.

[124]. MA, X., EL-KEIB, A. A., SMITH, R. E., and MA, H., 1995, Genetic algorithm based approach to thermal unit commitment of electric power systems. Electric Power Systems Research, 34, 29±36.

[125]. YANG, H.-T., and HUANG, C.-M., 1998, New short term load-forecasting approach using self-organizing fuzzy ARMAX models. IEEE Transactions on Power Systems, 13, 217±225.

[126]. N. K. Ahmed, A. F. Atiya, N. E. Gayar, and H. El-Shishiny, “An empirical comparison of machine learning models for time series forecasting,” Econometric Reviews, vol. 29, no. 5-6, pp. 594–621, 2010.

[127]. Mitchell, T. (1997). Machine Learning, McGraw Hill. ISBN 0-07-042807-7, p.2.

[128]. Alpaydin, E. (2004). Introduction to Machine Learning. Cambridge, MA: MIT Press.

[129]. DAMBORG, M. J., EL-SHARKAWI, M. A., AGGOUNE, M. E., and MARKS, R. J., II, 1990, Potential of artificial neural network to power system operation. Proceedings of the IEEE International Symposium on Circuits and Systems, New Orleans, LA, pp. 2933± 2937.

[130]. Tasgetiren MF. Neural networks: multivariate. Sakarya University. http://elektroteknoloji.com/Elektrik_Elektronik/PLC_Sistemleri/Yapay_Sinir_Aglar_Cok_Katmanli.html.

[131]. Fausett L. Fundamentals of neural networks. New Jersey: Prentice-Hall; 1994.

[132]. Ozsahin TS, Birinci A, Cakırog˘lu AO. Prediction of contact lengths between an elastic layer and two elastic circular punches with neural networks. Struct Eng Mech 2004;18:441–59.

[133]. M. El-Sharkawi, D. Niebur, Artificial neural networks with applications to power systems, IEEE Catalog No 96TP112-0, IEEE Service Center, NJ, 1996.

[134]. C.-T. Lin, G.S.G. Lee, Neural Fuzzy Systems, Prentice Hall, NJ, 1996.

[135]. Hippert, H.S., Pedreira, C.E., Souza, R.C., 2001. Neural networks for short-term load forecasting: A review and evaluation. IEEE Transactions on Power Systems 16 (1), 44–55.

[136]. Becalli, M., Cellura, M., Lo Brano, V., Marvuglia, A., 2004. Forecasting daily urban electric load profiles using artificial neural networks. Energy Comversion and Management 45, 2879–2900.

[137]. Ringwood, J.V., Bofelli, D., Murray, F.T., 2001. Forecasting electricity demand on short, medium, and long-time scales using neural networks. Journal of Intelligent and Robotic Systems 31, 129–147.

[138]. Moody, J. E., Darken, C. (1989). Fast learning in networks of locally-tuned processing units. Neural Computation 1:281–294.

[139]. Powell, M. J. D. (1987). Radial basis functions for multivariable interpolation: In: Mason, J. C., Cox, M. G., eds. A Review in Algorithms for Approximation. Oxford: Clarendon, pp. 143–168.

[140]. S. Chen, E. Chng, Regularized orthogonal least squares algorithm for constructing radial basis function networks, Int. J. Control 64 (5) (1996) 829_/837.

[141]. Ranaweera, D.K., Hubele, N.F., Papalexopoulos, A.D., 1995. Application of radial basis function neural network model for short-term load forecasting. IEE Proceedings Generation, Transmission and Distribution 142, 45–50.

[142]. Nadaraya, E. A. (1964). On estimating regression. Theory of Probability and Its Applications 10:186–190.

[143]. Watson, G. S. (1964). Smooth regression analysis. Sankhy Series A 26:359–372.

[144]. Hardle, W. (1990). Applied Nonparametric Regression, Econometric Society Monographs, 19. Cambridge, UK: Cambridge University Press.

[145]. MacKay, D. J. C. (1992a). Bayesian interpolation. Neural Computation 4:415–447.

[146]. MacKay, D. J. C. (1992b). A practical Bayesian framework for backpropagation networks. Neural Computation 4:448–472.

[147]. Malakoff DM. Bayes offer 'new' way to make sense of numbers. Science 1999;286:1460–4.

[148]. Jaynes ET. Probability theory – the logic of science. Cambridge: Cambridge University Press; 2003

[149]. Vapnik V. and Lerner A. 1963. Pattern recognition using generalized portrait method. Automation and Remote Control, 24: 774–780.

[150]. Vapnik V. and Chervonenkis A. 1964. A note on one class of perceptrons. Automation and Remote Control, 25.

[151]. Vapnik V. and Chervonenkis A. 1974. Theory of Pattern Recognition [in Russian]. Nauka, Moscow. (German Translation: Wapnik W. & Tscherwonenkis A., Theorie der Zeichenerkennung, Akademie-Verlag, Berlin, 1979).

[152]. K.-j. Kim, "Financial time series forecasting using support vector machines," Neurocomputing, vol. 55, no. 1, pp. 307–319, 2003.

[153]. Cortes C, Vapnik V (1995) Support vector networks. Mach Learn 20:273–297. doi:10.1023/ A:1022627411411

[154]. A Survey of Time Series Prediction Using SVM,Yongning Ma, October, 1 2012

[155]. Shawe-Taylor J, Bartlett PL, Williamson RC, Anthony M (1998) Structural risk minimization over data-dependent hierarchies. IEEE Trans Inf Theory 44:1926–1940. doi:10.1109/ 18.705570

[156]. O. A. Omitaomu, M. K. Jeong, A. B. Badiru, and J. W. Hines, "Online support vector regression approach for the monitoring of motor shaft misalignment and feedwater flow rate," IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 37, no. 5, pp. 962–970, Sep. 2007.

[157]. W.-C. Hong, "Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model," Energy Conversion and Management, vol. 50, no. 1, pp. 105– 117, 2009.

[158]. Schölkopf, B, A. Smola. (2002). Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond, MIT Press, Cambridge, MA.

[159]. Short-Term Load Forecasting Using Support Vector Machines, JJF Pereira, RM Salgado, T Ohishi, R Ballini - IEEE PES Transmission and Distribution Conference, 2006

[160]. Gunn, S. (1998). Support Vector Machine for Classification and Regression. Image Speech & Intelligent Systems Group, Technical Report ISIS-1-98, University of Southampton.
[161]. Amari S, Wu S (1999) Improving support vector machine classifiers by modifying kernel functions. Neural Netw 12:783–789. doi:10.1016/S0893- 6080(99)00032-5
[162]. Vojislav K (2001) Learning and soft computing—support vector machines, neural networks and fuzzy logic models. The MIT Press, Cambridge, MA, 2001
[163]. M. Mohandes. Support Vector Machines for Short-Term Electrical Load Forecasting. International Journal of Energy Research, 26:335‑345, 2002.
[164]. Y. Li and T. Fang. Wavelet and Support Vector Machines for Short-Term Electrical Load Forecasting. Proceedings of International Conference on Wavelet Analysis and its Applications, 1:399‑ 404, 2003.
[165]. Niu, D., Li, J., Li, J., Wang, Q., 2007c. Daily load forecasting using support-vector machine and case-based reasoning. In: Second IEEE Conference on Industrial Electronics and Applications, 2007. ICIEA 2007. IEEE, pp. 1271–1274.
[166]. Hsu, C.-C., Wu, C.-H., Chen, S.-C., Peng, K.-L., 2006. Dynamically optimizing parameters in support vector regression: An application of electricity load forecasting. In: Proceedings of the 39th Hawaii International Conference on System Sciences – 2006 (HICSS 2006), pp. 1–8.
[167]. Afshin, M., Sadeghian, A., 2007. PCA-based least squares support vector machines in week-ahead load forecasting. In: Industrial and Commercial Power System Technical Conference, 2007 (ICPS 2007). IEEE, pp. 1–6.
[168]. Li, G., Cheng, C.-T., Lin, J.-Y., Zeng, Y., 2007. Short-term load forecasting using support vector machine with SCE-UA algorithm. In: Third International Conference on Natural Computation (ICNC 2007). IEEE, pp. 290–294.
[169]. Niu, D., Li, J., Li, J., 2007b. Middle-long electric power load forecasting based on cointegration and support vector machine. In: Third International Conference on Natural Computation (ICNC 2007). IEEE, pp. 596–600.
[170]. B.-J. Chen, M.-W. Chang, et al., “Load forecasting using support vector machines: A study on eunite competition 2001,” Power Systems, IEEE Transactions on, vol. 19, no. 4, pp. 1821– 1830, 2004.
[171]. Winter School on "Data Mining Techniques and Tools for Knowledge Discovery in Agricultural Datasets” 144
[172]. P.Harrington, “Machine Learning in Action’, Manning Publications, April 19, 2012
[173]. F. Pedregosa et al. Scikit-learn: Machine learning in Python. JMLR, 12:2825-2830, 2011
[174]. S. van der Walt et al. The NumPy array: a structure for efficient numerical computation. CiSE, 13:22, 2011.
[175]. T. E. Oliphant. Python for scienti_c computing. CiSE, 9:10, 2007.
[176]. J. D. Hunter. Matplotlib: A 2d graphics environment. CiSE, 9: 90, 2007.
[177]. F. Perez and B. E. Granger. IPython: a system for interactive scienti_c computing. CiSE, 9:21, 2007.
[178]. Scikit Learn tutorial: http://scikit-learn.org/stable/
[179]. T. Kossowski, “Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data”, 2011, Quaestiones Geographicae, 30, 87-93

[180]. D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, “Detecting novel associations in large data sets" science, vol. 334, no. 6062, pp. 1518-1524, 2011.
[181]. Mine tutorial: http://minepy.sourceforge.net/
[182]. J. Shawe-Taylor and N. Cristianini. Kernel Methods for Pattern Analysis. Cambridge UP, Cambridge, UK, 2004.