



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

Βραχυπρόθεσμη πρόβλεψη φορτίου ηλεκτρικής ενέργειας με
χρήση τεχνητών νευρωνικών δικτύων τύπου Long-Short Term
Memory

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Του

Μαστοράκη Αλέξιου

Επιβλέπων : Νικόλαος Χατζηαργυρίου
Καθηγητής Ε.Μ.Π.

Αθήνα, Μάρτιος 2023



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

Βραχυπρόθεσμη πρόβλεψη φορτίου ηλεκτρικής ενέργειας με
χρήση τεχνητών νευρωνικών δικτύων τύπου Long-Short Term
Memory

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Του

Μαστοράκη Αλέξιου

Επιβλέπων : Νικόλαος Χατζηαργυρίου

Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 09^η Μαρτίου 2023.

.....

Νικόλαος Χατζηαργυρίου

Καθηγητής Ε.Μ.Π.

.....

Σταύρος Παπαθανασίου

Καθηγητής Ε.Μ.Π.

.....

Πάυλος Γεωργιάκης

Καθηγητής Ε.Μ.Π.

Αθήνα, Μάρτιος 2023

.....

Μαστοράκη Αλέξιου

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

Copyright © Αλέξιου Μαστοράκη, 2023

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 τον συγγραφέα.

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

Περίληψη

Στην παρούσα διπλωματική εργασία επιδιώκεται να πραγματοποιηθεί βραχυπρόθεσμη πρόβλεψη ορίζοντα μίας ώρας, 12 ωρών της ίδια ημέρας και των επόμενων 24 ωρών της επόμενης ημέρας. Η ανάγκη για τέτοια πρόβλεψη αναδείχτηκε με την απελευθέρωση της αγοράς και τη διεύθυνση πολλών συμμετεχόντων σε αυτήν. Στο περιβάλλον του ανταγωνισμού και της μεταβλητότητας που χαρακτηρίζει την αγορά, μια εύστοχη πρόγνωση του φορτίου της ηλεκτρικής ενέργειας της επόμενης ημέρας επωφελεί όλους τους συμμετέχοντες, για διαφορετικούς λόγους έκαστο.

Η προσέγγισή μας για αυτή την πρόγνωση γίνεται με δύο τρόπους, αρχικά με στατιστικές μεθόδους, δηλαδή με αυτοπαλιδρομικά μοντέλα κινητού μέσου όρου (ARIMA), και με τη χρήση ενός απλού νευρωνικού δικτύου (Artificial Neural Networks ή ANN) τύπου πολυεπίπεδου Perceptron (MLP ή Multilayer Perceptron) και στην συνέχεια με Αναδρομικά Νευρωνικά Δίκτυα που αποτελούνται από στρώματα νευρώνων Μακράς Βραχείας Μνήμης (Long Short Term Memory ή LSTM). Προκειμένου να μελετηθεί η αποδοτικότητα πρόβλεψης φορτίου μεταξύ των δύο κατηγοριών μοντέλων υλοποιούνται με χρήση δύο διαφορετικών διαστημάτων δεδομένων εισόδου: Το πρώτο διάστημα περιλαμβάνει τις προηγούμενες 24 ώρες και το δεύτερο διάστημα περιλαμβάνει τις προηγούμενες 168 ώρες, δηλαδή ωριαίες τιμές φορτίου της προηγούμενης ημέρας και εβδομάδας αντίστοιχα. Στη συνέχεια, εκτιμάται η απόδοση των μοντέλων, συγκρίνονται τα σφάλματα που προκύπτουν από τις διαφορετικές μεθόδους πρόβλεψης, διαπιστώνεται η υπεροχή των νευρωνικών έναντι των στατιστικών μοντέλων και αξιολογείται η βελτίωση που προσφέρει η χρήση μηχανικής μάθησης στην πρόγνωση της κατανάλωσης ηλεκτρικής ενέργειας.

Λέξεις κλειδιά

Πρόβλεψη Χρονοσειρών, βραχυπρόθεσμες προβλέψεις, Τεχνητά Νευρωνικά δίκτυα, φορτίο, κατανάλωση ηλεκτρικής ενέργειας, αναδρομικά νευρωνικά δίκτυα, LSTM.

Abstract

In this thesis, it is sought to perform a short-term horizon forecast of one hour, 12 hours same day and 24 hours for the day ahead. The need for such a forecast emerged with the liberalization of the energy market and the penetration of many participants into it. In the environment of competition and volatility that characterizes the market, an accurate forecast of the next day's electricity load benefits all participants, for different reasons.

Our approach for this forecast is done in two ways, first with statistical methods, i.e. with autoregressive moving average (ARIMA) models, and with the use of a feedforward neural network (Artificial Neural Networks or ANN) type MLP (Multilayer Perceptron) and then with Recursive Neural Networks consisting of Long Short Term Memory (LSTM) neuron layers. In order to study the efficiency of load forecasting between the two categories of models, two different intervals of input data are implemented: The first interval includes the previous 24 hours and the second interval includes the previous 168 hours, i.e. hourly load values of the previous day and week respectively. Then, the performance of the models is evaluated, the errors resulting from the different forecasting methods are compared, the superiority of neural over statistical models is established, and the improvement of using machine learning in electricity consumption forecasting is evaluated.

Keywords

Time Series Forecasting, short term forecasting, Artificial Neural Networks, electricity consumption, feedforward neural networks, recurrent neural networks, LSTM.

Ευχαριστίες

Η παρούσα διπλωματική εκπονήθηκε στο Εργαστήριο Συστημάτων Ηλεκτρικής Ενέργειας της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου. Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή κ. Χατζηαργυρίου, για την ευκαιρία που μου έδωσε, να εκπονήσω τη διπλωματική μου εργασία, σε ένα θέμα τόσο ενδιαφέρον και δημιουργικό. Επίσης ευχαριστώ ιδιαίτερα τον υποψήφιο διδάκτορα Θωδωρή Κωνσταντίνου για την εξαιρετική συνεργασία και τη συνεχή καθοδήγηση κατά τη συγγραφή της εργασίας αυτής.

Περιεχόμενα

Ευχαριστίες.....	9
Περίληψη.....	6
Abstract	8
Περιεχόμενα πινάκων και εικόνων.....	12
ΚΕΦΑΛΑΙΟ 1: Εισαγωγή.....	14
1.1 Σημασία της πρόβλεψης φορτίου	14
1.2 Παράγοντες που επηρεάζουν τη ζήτηση φορτίου ηλεκτρισμού	16
1.3 Το αντικείμενο της διπλωματικής	18
ΚΕΦΑΛΑΙΟ 2: Φορτίο.....	20
2.1 Γενικά	20
2.2 Φορτίο	20
ΚΕΦΑΛΑΙΟ 3: Μέθοδοι πρόβλεψης & Χρονοσειρές.....	28
3.1 Γενικά	28
3.2 Χρονοσειρές	28
3.2.1 Χαρακτηριστικά χρονοσειρών.....	28
3.2.2 Ολοκληρωμένα Αυτοπαλινδρομικά Μοντέλα Κινητού Μέσου Όρου - ARIMA.....	33
3.3 Μηχανική Μάθηση(Machine Learning) - Τεχνητά Νευρωνικά Δίκτυα(Artificial Neural Networks)	39
3.3.1 Εισαγωγή	39
3.3.2 Νευρώνας - Perceptron	42
3.3.3 Multilayer Perceptron	44
3.3.4 Συνελκτικό Νευρωνικό Δίκτυο - Convolutional Neural Networks.....	45
3.3.5 Αναδρομικά Νευρωνικά Δίκτυα - Recurrent Neural Networks	48
3.3.6 Νευρωνικό Δίκτυο Μακράς-Βραχείας Μνήμης - LSTM.....	50
3.3.7 Υπερπαράμετροι αλγορίθμων	51
3.3.8 Υπερπαράμετροι του μοντέλου LSTM.....	53
3.4 Μετρικές Σφαλμάτων	56
ΚΕΦΑΛΑΙΟ 4: Ανάλυση & Πρόβλεψη φορτίου χρονοσειράς ηλεκτρικής ενέργειας με χρήση νευρωνικού δικτύου LSTM.....	59
4.1 Εισαγωγή	59
4.2 Χρονοσειρά φορτίου	59
4.3 Αλγόριθμος LSTM	65
4.3.1 Παρουσίαση των μοντέλων ARIMA, MLP & LSTM	65
4.3.2 Σύγκριση	66

4.3.2.1 Πρόβλεψη επόμενης ώρας.....	66
4.3.2.2 Πρόβλεψη επόμενων 12 ωρών	70
4.3.2.3 Πρόβλεψη επόμενων 24 ωρών	76
ΚΕΦΑΛΑΙΟ 5: ΣΥΜΠΕΡΑΣΜΑΤΑ	82
5.1 Πρόβλεψη του φορτίου ηλεκτρικής ενέργειας της επόμενης ώρας (continuous 1h)	82
5.2 Πρόβλεψη του φορτίου ηλεκτρικής ενέργειας των επόμενων 12 ωρών της παρούσας ημέρας (intra day 12h)	83
5.3 Πρόβλεψη του φορτίου ηλεκτρικής ενέργειας των επόμενων 24 ωρών της επόμενης ημέρας (day ahead)	86
5.4 Γενικά συμπεράσματα	88
Βιβλιογραφία.....	91

Περιεχόμενα πινάκων και εικόνων

1-Σχήμα 3.1: Example of a Stationary and a Non-Stationary Time Series.....	29
2-Σχήμα 3.2: Trend, Cycle, Seasonality & Remainder of a Time Series.....	32
3-Σχήμα 3.3: Artificial Neuron model and its parts.....	43
4-Σχήμα 3.4: Activation functions.....	43
5-Σχήμα 3.5: Three layer multilayer perceptron (MLP) neural network.....	45
6-Σχήμα 3.6: CNN & MLP Architecture.....	46
7-Σχήμα 3.7: Convolutional Layer of a CNN.....	47
8-Σχήμα 3.8: CNN model architecture.....	48
9-Σχήμα 3.9: Recursive Neural Network Unfolding.....	48
10-Σχήμα 3.10: LSTM Model Architecture.....	50
11-Σχήμα 3.11: Παράδειγμα γραμμικής & λογιστικής παλινδρόμησης.....	51
12-Σχήμα 4.1: Αναπαράσταση του φορτίου σε MW για το έτος 2018.....	59
13-Σχήμα 4.2: Συσχέτιση μεταξύ προηγούμενων και επόμενων τιμών της χρονοσειράς.....	60
14-Σχήμα 4.3:Κατανομή τιμών χρονοσειράς.....	60
15-Σχήμα 4.4: Κατανομή εύρους τιμών φορτίου ανά μήνα.....	61
16-Σχήμα 4.5: Το διάγραμμα αυτοσυσχέτισης της χρονοσειράς.....	61
17-Σχήμα 4.6: Αναπαράσταση κατανάλωσης ηλεκτρικής ενέργειας στην αντίστοιχη ώρα της χρονοσειράς (πάνω) , και κατανομή τιμών φορτίου της χρονοσειράς (κάτω).....	62
18-Σχήμα 4.7: Παρόμοια με το σχήμα 4.6 αλλά με αναγωγή των τιμών μεταξύ -1 και 1 όπου 1 η μέγιστη τιμή κατανάλωσης της χρονοσειράς ανά ώρα.....	62
19-Σχήμα 4.8: Χρονοσειρά με Τάση.....	63
20-Σχήμα 4.9: Χρονοσειρά με απαλοιφή Τάσης.....	63
21-Σχήμα 4.10: Αναπαράσταση της εποχικότητας σε δείγμα 500 τιμών της χρονοσειράς.....	64
22-Σχήμα 4.11: Αναπαράσταση εποχικότητας σε δείγμα 500 τιμών της χρονοσειράς.....	64
23-Σχήμα 4.12: Συντελεστές του πολυωνύμου.....	65
24-Σχήμα 4.13: Διάγραμμα κατανομής σφάλματος ME μοντέλου LSTM.....	67
25-Σχήμα 4.14: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM.....	67
26-Σχήμα 4.15: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς...	68
27-Σχήμα 4.16: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM.....	68
28-Σχήμα 4.17: Διάγραμμα κατανομής σφάλματος ME μοντέλου LSTM.....	69
29-Σχήμα 4.18: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM.....	69
30-Σχήμα 4.19: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς...	70
31-Σχήμα 4.20: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM.....	70
32-Σχήμα 4.21: Διάγραμμα κατανομής σφάλματος ME μοντέλου LSTM.....	71
33-Σχήμα 4.22: Διάγραμμα σφάλματος MAPE μοντέλου LSTM.....	71
34-Σχήμα 4.23: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM.....	72
35-Σχήμα 4.24: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς...	72
36-Σχήμα 4.25: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM.....	73
37-Σχήμα 4.26: Διάγραμμα κατανομής σφάλματος ME μοντέλου LSTM.....	73
38-Σχήμα 4.27: Διάγραμμα σφάλματος MAPE μοντέλου LSTM.....	74
39-Σχήμα 4.28: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM.....	74
40-Σχήμα 4.29: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς...	75
41-Σχήμα 4.30: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM.....	75
42-Σχήμα 4.31: Διάγραμμα κατανομής σφάλματος ME μοντέλου LSTM.....	76
43-Σχήμα 4.32: Διάγραμμα σφάλματος MAPE μοντέλου LSTM.....	76
44-Σχήμα 4.33: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM.....	77
45-Σχήμα 4.34: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς...	77
46-Σχήμα 4.35: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM.....	78
47-Σχήμα 4.36: Διάγραμμα κατανομής σφάλματος ME μοντέλου LSTM.....	78
48-Σχήμα 4.37: Διάγραμμα σφάλματος MAPE μοντέλου LSTM.....	79

49-Σχήμα 4.38: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM	79
50-Σχήμα 4.39: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς...	80
51-Σχήμα 4.40: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM	80

ΚΕΦΑΛΑΙΟ 1: Εισαγωγή

1.1 Σημασία της πρόβλεψης φορτίου

Στην παρούσα διπλωματική εργασία καλούμαστε να απαντήσουμε τρία ερωτήματα:

- Για ποιο λόγο καλούμαστε να κάνουμε πρόβλεψη φορτίου?
- Πως αυτό αναλύεται μαθηματικά, είτε σαν δεδομένα είτε σαν αλγόριθμος (δηλαδή με ποιους στατιστικούς αλγορίθμους ή νευρωνικά δίκτυα)?
- Ποια είναι η βέλτιστη μοντελοποίηση των παραπάνω χρονοσειρών (εποχικότητα, κλπ) και αλγορίθμων (παράμετροι).

Η πρόβλεψη ζήτησης φορτίου ηλεκτρικής ενέργειας είναι ένας τομέας της επιστήμης των προβλέψεων, που προϋπάρχει εδώ και μερικές δεκαετίες, όμως έντονη ενασχόληση με αυτόν παρατηρείται μόλις τα τελευταία χρόνια. Η πρόβλεψη ζήτησης φορτίου είναι ιδιαίτερης σημασίας για τις εταιρείες παραγωγής ηλεκτρικής ενέργειας, καθώς έχει διάφορες εφαρμογές με βασικότερη τον οικονομικό προγραμματισμό του συστήματος παραγωγής. Κατέχοντας την πρόβλεψη ζήτησης ηλεκτρικού φορτίου για την επόμενη ημέρα έως και εβδομάδα, μπορεί η εταιρεία να διαχειριστεί τα υπάρχοντα συστήματα ηλεκτρικής ενέργειας (ΣΗΕ) και να προγραμματίσει τη βέλτιστη ένταξη μονάδων στο δίκτυο και την ανταλλαγή ηλεκτρικής ενέργειας με γειτονικά δίκτυα. Έχοντας υπόψη ότι βασικό χαρακτηριστικό της ηλεκτρικής ενέργειας είναι η δυσκολία αποθήκευσής της, καταλαβαίνει κανείς πόσο σημαντική είναι η διατήρηση της παραγωγής και της ζήτησης ηλεκτρικής ενέργειας στο ίδιο επίπεδο [1].

Η γνώση της προβλεπόμενης ζήτησης φορτίου διευκολύνει ακόμη το έργο των χειριστών του συστήματος ηλεκτρικής ενέργειας (ΣΗΕ). Με βάση την πρόβλεψη αυτή αποφασίζουν ποιες διαδικασίες θα ακολουθήσουν για τη διόρθωση της παραγωγής, από την ένταξη μονάδων, την αγορά επιπλέον ενέργειας από γειτονικά δίκτυα, τη ρύθμιση της τάσης των ζυγών για την επόμενη ημέρα -ως και εβδομάδα, ώστε να υπάρχει η απαιτούμενη εφεδρεία για την οικονομική λειτουργία και την εξασφάλιση της κάλυψης της ζήτησης, σε περιπτώσεις σφαλμάτων και ανωμαλιών του συστήματος. Σε διαφορετική περίπτωση, η εταιρεία παραγωγής ενέργειας αντιμετωπίζει ιδιαίτερα δαπανηρές επιλογές δράσης. Συγκεκριμένα, αν υποτιμηθεί αρκετά η μελλοντική ζήτηση φορτίου, η εταιρεία θα προβεί στην ενσωμάτωση μονάδων αιχμής στο δίκτυο για την κάλυψη της επιπλέον ζήτησης, εκτός του ότι κινδυνεύει η αξιοπιστία του συστήματος και ενδέχεται να παρατηρηθούν διακοπές ενέργειας. Ενώ σε περίπτωση έντονης υπερτίμησης της ζήτησης, θα έχει παραχθεί φορτίο χωρίς λόγο και αγοραστικό κοινό και έτσι οδηγείται και πάλι σε αύξηση του λειτουργικού κόστους. Επίσης, η κατοχή προβλέψεων ακριβείας για το συνδυασμό ζήτησης και τιμής του φορτίου δίνουν στις εταιρείες τη δύναμη να αναπτύξουν προσοδοφόρα πολιτική του προσφερόμενου φορτίου τους απέναντι στις ανταγωνίστριες εταιρείες και στον

καταναλωτή, να εξυπηρετούν μεγαλύτερο μέρος της αγοράς και να μεγιστοποιήσουν τα κέρδη τους [2].

Όπως αναφέρθηκε και παραπάνω, ο επιστημονικός κόσμος έχει παρουσιάσει ιδιαίτερο ενδιαφέρον για την πρόβλεψη ζήτησης φορτίου μόλις τα τελευταία χρόνια. Παλαιότερα δεν υπήρχε λόγος για έντονη ενασχόληση με τον τομέα αυτό των προβλέψεων, καθώς συγκεκριμένοι οργανισμοί αναλάμβαναν την παραγωγή ηλεκτρικής ενέργειας σε κάθε χώρα. Στην αρχή της δεκαετίας του '90 όμως ξεκίνησε η απελευθέρωση των αγορών ενέργειας σε πολλές χώρες, με αποτέλεσμα την κατάλυση των ισχυρών μονοπωλίων και την ύπαρξη πλέον ανταγωνιστικών αγορών με πολλές εταιρείες παραγωγής ενέργειας. Αυτή η νέα κατάσταση οδήγησε και στην έντονη ανάγκη για σωστή και όσο το δυνατόν πιο ακριβή πρόβλεψη της ζήτησης φορτίου ηλεκτρικής ενέργειας, για τον αποτελεσματικό προγραμματισμό και έλεγχο των ΣΗΕ (όπως προαναφέραμε), όπως και της τιμής της ενέργειας, ώστε να αποφασίζει η εκάστοτε εταιρεία την επόμενη κίνησή της για να επιβιώσει ή ακόμη και να ηγηθεί του ανταγωνισμού της αγοράς. Επιπλέον, η κατανάλωση ενέργειας με το πέρασμα του χρόνου αυξήθηκε σημαντικά, εφόσον στα νοικοκυριά όπως και στις υπηρεσίες ενσωματώθηκαν ηλεκτρικές συσκευές που δεν υπήρχαν στο παρελθόν (ηλεκτρονικός υπολογιστής, κλιματιστικό, κλπ) και η χρήση τους εξαρτάται από την χρονική περίοδο που βρισκόμαστε μέσα στη διάρκεια του ημερολογιακού έτους και άλλους παράγοντες. Στην Ελλάδα, η απελευθέρωση της αγοράς ηλεκτρικής ενέργειας οδήγησε στην κατάργηση του αποκλειστικού προνομίου που κατείχε η Δημόσια Εταιρεία Ηλεκτρισμού (ΔΕΗ) το Δεκέμβριο του 1999. Πλέον από το Φεβρουάριο του 2001 οι πελάτες που συνδέονται στην Υψηλή και τη Μέση Τάση έχουν το δικαίωμα επιλογής του προμηθευτή τους. Έτσι αποκτά σημασία και στη χώρα μας η πρόβλεψη ζήτησης φορτίου ηλεκτρικής ενέργειας και παρατηρείται ενασχόληση πολλών Ελλήνων ερευνητών με το αντικείμενο.

Συγκεκριμένα υπάρχουν τέσσερις κατηγορίες πρόβλεψης φορτίου, οι οποίες κατηγοριοποιούνται σύμφωνα με τον χρονικό ορίζοντά τους και είναι οι εξής:

- Η μακροπρόθεσμη πρόβλεψη φορτίου έχει ορίζοντα από ένα ως και αρκετά έτη, χρησιμεύει στο σχεδιασμό και τη συντήρηση των συστημάτων παραγωγής και διανομής ενέργειας. Είναι ακόμη βασικό κριτήριο, στο οποίο στηρίζονται οι εταιρείες για αποφάσεις σχετικά με μεγάλες επενδύσεις αγοράς εγκαταστάσεων και νέων μονάδων παραγωγής.
- Η μεσοπρόθεσμη πρόβλεψη φορτίου, με ορίζοντα από ένα μήνα ως και τρία έτη, χρειάζεται κατά κύριο λόγο στη ρύθμιση προγραμμάτων συντήρησης των διαθέσιμων ΣΗΕ, καθώς και στο χειρισμό και το σχεδιασμό τους.
- Η βραχυπρόθεσμη πρόβλεψη, με ορίζοντα από μερικές ώρες ως και μία εβδομάδα, απαιτείται για τον έλεγχο και τον προγραμματισμό των συστημάτων ηλεκτρικής ενέργειας, για την εκτίμηση της ασφάλειάς τους (για παράδειγμα καθορίζει τη σειρά ένταξης των μονάδων παραγωγής του ΣΗΕ). Η βραχυπρόθεσμη πρόβλεψη είναι επίσης απαραίτητη από εταιρείες μετάδοσης ενέργειας, στην περίπτωση λειτουργίας self-

dispatching αγορών. Σε ορισμένες τέτοιες αγορές χρησιμεύει στην εξισορρόπηση της παροχής με τη ζήτηση, έπειτα από την παράπλευρη συναλλαγή ανάμεσα στους παραγωγούς και τους προμηθευτές της ενέργειας, ώστε να εξασφαλίζεται η αξιόπιστη λειτουργία των συστημάτων ηλεκτρικής ενέργειας(ΣΗΕ).

- Η πολύ βραχυπρόθεσμη πρόβλεψη, με ορίζοντα από λίγα λεπτά ως και μία ώρα, χρησιμεύει για την κατανομή φορτίου στις μονάδες του ΣΗΕ κατά το βέλτιστο και πλέον οικονομικό τρόπο. Συμβάλλει επομένως στην κάλυψη των αναγκών του συστήματος των γεννητριών και με την άμεση παρακολούθηση της συχνότητας με το φορτίο, εξασφαλίζεται καλύτερα η ασφάλεια του συστήματος. Ακόμη, μπορεί να συμβάλει σημαντικά στη ρύθμιση της τιμής της ηλεκτρικής ενέργειας.

1.2 Παράγοντες που επηρεάζουν τη ζήτηση φορτίου ηλεκτρισμού

Στη μελέτη αυτή μας ενδιαφέρει η επίτευξη καλής προσέγγισης της ωριαίας ζήτησης φορτίου ηλεκτρισμού, με ορίζοντα μία ημέρα. Έχει επομένως σημασία να καταγράψουμε τους παράγοντες που επηρεάζουν την πορεία της ζήτησης φορτίου και να κρίνουμε ποιους θεωρούμε απαραίτητους να συμπεριληφθούν στις μεθόδους που θα εξετάσουμε στην πορεία. Παραθέτουμε λοιπόν τους σημαντικότερους παράγοντες που εντοπίσαμε και είναι εποχικοί, καιρικοί και οικονομικοί.

Ως εποχικούς παράγοντες θεωρούμε τις εναλλαγές των εποχών κατά τη διάρκεια του έτους, τις εποχικότητες που παρουσιάζουν τα δεδομένα και τις ειδικές ημέρες που μεταβάλλονται ανάλογα με τη χώρα και τις συνθήκες που επικρατούν σε αυτή. Πιο συγκεκριμένα, μέσα σε ένα ημερολογιακό έτος υπάρχει η διαδοχή των τεσσάρων εποχών (χειμώνας, άνοιξη, καλοκαίρι και φθινόπωρο). Η εναλλαγή των εποχών προκαλεί μεταβολή και στην ένταση των ανθρώπινων δραστηριοτήτων που απαιτούν ηλεκτρική ενέργεια, όπως για παράδειγμα στην Ελλάδα το καλοκαίρι αυξάνεται η ζήτηση ενέργειας λόγω της ύπαρξης μεγάλου πλήθους τουριστών, των οποίων οι ανάγκες πρέπει να ικανοποιηθούν. Παρατηρείται έτσι αλλαγή στο επίπεδο ζήτησης της ηλεκτρικής ενέργειας, η οποία παρουσιάζει διαφορετική μέση τιμή και τιμή αιχμής (πολύ σημαντικές για τη σωστή λειτουργία των ΣΗΕ) σε κάθε εποχή. Έχοντας μεγάλο πλήθος δεδομένων μπορεί να κατανοήσει κανείς την εποχικότητα που παρουσιάζουν τα δεδομένα ζήτησης φορτίου σε ετήσια βάση. Ακόμη, η ζήτηση ηλεκτρικού φορτίου παρουσιάζει έντονη εποχικότητα, τόσο ημερήσια όσο και εβδομαδιαία. Αυτή έχει ως αιτία τη διαφορετικότητα των ημερών της εβδομάδας, όπου έχουμε τις εργάσιμες ημέρες (καθημερινές) από Δευτέρα έως και Παρασκευή και το Σάββατο και την Κυριακή. Η πρώτη κατηγορία ημερών χαρακτηρίζεται από έντονη ενασχόληση όλων των εργαζομένων, το Σάββατο θεωρείται ως ημι-εργάσιμη ημέρα, ενώ η Κυριακή είναι θεωρητικά ημέρα ανάπαυσης για όλους. Αντίστοιχα διαμορφώνεται και η ζήτηση του φορτίου μέσα στην εβδομάδα. Επίσης, σε επίπεδο ημέρας, παρατηρείται αυξομείωση της ζήτησης του φορτίου, καθώς από τις 6:00 το πρωί ως και τις 22:00 το βράδυ υπάρχει έντονη δραστηριότητα του ανθρώπου και με ένα διάλειμμα το απόγευμα, τα επίπεδα ζήτησης είναι αυξημένα (λόγω λειτουργίας όλων των υπηρεσιών, εργοστασίων, κλπ την ημέρα και λόγω λοιπών δραστηριοτήτων

το απόγευμα και βράδυ, όπως τηλεόραση και ηλεκτρική κουζίνα). Τέλος, κάθε χώρα έχει τις δικές της ειδικές ημέρες, ημέρες εθνικών και μεγάλων θρησκευτικών εορτών, ιδιαίτερες αργίες και τις εκάστοτε απεργίες. Αυτές οι ημέρες προκαλούν ανωμαλίες στην πορεία της χρονοσειράς ζήτησης φορτίου είτε αυξάνοντας είτε μειώνοντας πολύ το επίπεδο (όπως Πρωτομαγιά).

Οι καιρικοί παράγοντες είναι άμεσα συνδεδεμένοι με την εναλλαγή των εποχών λόγω των μεταβολών της θερμοκρασίας. Για παράδειγμα, κατά τη διάρκεια του χειμώνα με τη μείωση της θερμοκρασίας είναι πιο έντονη η χρήση μέσων θέρμανσης, ενώ κατά τους καλοκαιρινούς μήνες αυξάνεται πολύ η θερμοκρασία και με αυτή η λειτουργία κλιματιστικών και άλλων μέσων ψύξης. Επιπλέον, ασυνήθιστες για την κάθε εποχή τιμές της θερμοκρασίας προκαλούν μεγάλες αποκλίσεις στη ζήτηση φορτίου, οι οποίες μάλιστα είναι έντονες και προκαλούν κινδύνους στα ΣΗΕ. Σε συνδυασμό με τη θερμοκρασία υπάρχει επιρροή και από το επίπεδο υγρασίας, τις βροχοπτώσεις, τη συννεφιά, την ύπαρξη έντονων ανέμων (ως και τυφώνων), το επίπεδο της ατμοσφαιρικής πίεσης και άλλους παρόμοιους παράγοντες.

Ως οικονομικούς παράγοντες αντιλαμβανόμαστε την οικονομική κατάσταση μιας χώρας (εφόσον αναφερόμαστε σε μελέτη ζήτησης φορτίου σε εθνικό επίπεδο) και το ρυθμό θετικής ή αρνητικής ανάπτυξης που την χαρακτηρίζει. Η σημασία τους είναι προφανής αφού με τη βελτίωση της οικονομικής κατάστασης μιας χώρας μεταβάλλεται και το βιοτικό επίπεδο των κατοίκων της, αυξάνεται η κατανάλωση σε αγαθά και υπηρεσίες, αναπτύσσονται νέες τεχνολογίες στον πρωτογενή και δευτερογενή τομέα για την κάλυψη των μεγαλύτερων πλέον αναγκών και υπάρχει γενικότερη εφορία. Προκαλείται με τον τρόπο αυτό αύξηση και της κατανάλωσης ηλεκτρικής ενέργειας, ώστε μεσοπρόθεσμα παρατηρείται άνοδος του επιπέδου των τιμών ζήτησης που μελετάμε. Σε κάποιο σημείο επέρχεται κορεσμός από την υπερκατανάλωση ενέργειας και οι καταναλωτές στρέφονται σε συσκευές με μικρότερη κατανάλωση ενέργειας στο πνεύμα ενός πιο οικολογικού τρόπου ζωής, με σκοπό τη μείωση του επιπέδου της ζητούμενης ενέργειας. Σε αντίθετη πορεία από μια χώρα ή μια περιοχή που βιώνει ανάπτυξη, όταν παρατηρείται αρνητική πορεία (ύφεση) σε αυτή και το βιοτικό επίπεδο των καταναλωτών υποβαθμίζεται, ο κόσμος απομακρύνεται από τα αστικά κέντρα, και υιοθετεί πιο εγκρατή νοοτροπία, η οποία αντικατοπτρίζεται και στη μείωση της ζήτησης ηλεκτρικής ενέργειας.

Υπάρχουν πολλοί ακόμη παράγοντες, εκτός από όσους αναφέρθηκαν, που δεν ανήκουν σε κάποια συγκεκριμένη κατηγορία. Πρόκειται για γεγονότα τυχαία ή απρόσμενα, όπως οι εκλογές, ένας πόλεμος ή μια πανδημία [3].

Από τους παράγοντες που καταγράφηκαν, ορισμένοι είναι πολύ δύσκολο να συνυπολογιστούν στα μοντέλα πρόβλεψης της ζήτησης φορτίου, όπως οι οικονομικοί παράγοντες (ο ρυθμός ανάπτυξης και η οικονομική κατάσταση της χώρας). Στα πλαίσια της εργασίας αυτής λαμβάνονται υπόψη οι εποχικοί παράγοντες. Με παρατήρηση της χρονοσειράς των διαθέσιμων δεδομένων ζήτησης φορτίου ηλεκτρισμού αντιλαμβάνεται κανείς ότι η εποχικότητα, ειδικά σε ημερήσια και εβδομαδιαία βάση, είναι το πιο έντονο φαινόμενο της ζήτησης φορτίου. Επομένως αποτελεί και το κύριο μέλημά μας, ενώ η ενσωμάτωση του καιρού δε θα μας

απασχολήσει προς το παρόν. Ο καιρός είναι όπως αναφέραμε σημαντικός παράγοντας στη διαμόρφωση της ζήτησης φορτίου, ωστόσο στη βραχυπρόθεσμη πρόβλεψη και μάλιστα με διασύνδεση με το δίκτυο, θεωρείται ιδιαίτερα πολύπλοκη η διαδικασία ενσωμάτωσης της επίδρασής του. Όπως δε γνωρίζουμε τις μελλοντικές τιμές για τη ζήτηση φορτίου και παράγουμε προβλέψεις τους, έτσι και με για τον καιρό χρησιμοποιούνται οι προβλεπόμενες τιμές της θερμοκρασίας (η θερμοκρασία είναι η βασικότερη μεταβλητή των καιρικών συνθηκών που επηρεάζει τη ζήτηση φορτίου). Αυτό το χαρακτηριστικό των τιμών θερμοκρασίας, υγρασίας και λοιπών παραγόντων (όποιων χαρακτηριστικών του καιρού αποφασίζεται να χρησιμοποιηθούν στην παραγωγή πρόβλεψης φορτίου ηλεκτρισμού) προσθέτει επιπλέον αβεβαιότητα στις τελικές προβλέψεις του ζητούμενου φορτίου.

Επιπλέον λόγω του μεγάλου όγκου αριθμητικών προβλέψεων καιρού χρειάζεται ειδική μεταχείριση των συγκεκριμένων δεδομένων χρειάζεται ειδική μεταχείριση για την αποδοτική ενσωμάτωσή τους στις μεθόδους πρόβλεψης. Πρέπει να τονίσουμε ακόμη, ότι σε κάθε περιοχή της χώρας επικρατούν διαφορετικές καιρικές συνθήκες, οι οποίες επηρεάζουν ανάλογα τη ζήτηση των καταναλωτών κατά τόπους. Η ξεχωριστή πρόβλεψη για τις διάφορες περιοχές της χώρας και ο συνδυασμός τους για την τελική προβλεπόμενη τιμή είναι χρονοβόρα και απαιτητική διαδικασία, η οποία δεν είναι πρακτική όταν στόχος είναι η πρόβλεψη των επόμενων είκοσι τεσσάρων ωριαίων τιμών του ζητούμενου ηλεκτρικού φορτίου. Τέλος, μπορεί να παρατηρήσει κανείς ότι η αλλαγή του καιρού δεν είναι τόσο απότομη όσο ένα απρόσμενο τυχαίο γεγονός. Έτσι, με την ομαλή μεταβολή των καιρικών συνθηκών, οι μέθοδοι μίας μεταβλητής (της εποχικότητας στην περίπτωση μας) θεωρούνται πολύ ικανοποιητικές, αφού η μεταβολή του καιρού αποτυπώνεται από μόνη της στη χρονοσειρά ζήτησης [4].

1.3 Το αντικείμενο της διπλωματικής

Η παρούσα διπλωματική εργασία αναλύει το πρόβλημα της βραχυπρόθεσμης πρόβλεψης ζήτησης φορτίου στα συστήματα ηλεκτρικής ενέργειας. Στόχος της είναι η χρήση νευρωνικών δικτύων για την πρόβλεψη της ζήτησης φορτίου του ηλεκτρικού δικτύου και τη σύγκρισή του με άλλες μεθόδους πρόβλεψης, καθώς και στην αξιολόγηση των παραμέτρων που επηρεάζουν την απόδοση του δικτύου. Η επίτευξη υψηλής ακρίβειας στη βραχυπρόθεσμη πρόβλεψη ζήτησης φορτίου είναι ζωτικής σημασίας για τα συστήματα ηλεκτρικής ενέργειας, καθώς απαιτείται για τον σχεδιασμό και τον συντονισμό των παραγωγικών μονάδων και του δικτύου. Επιπλέον, η βραχυπρόθεσμη πρόβλεψη ζήτησης φορτίου είναι ζωτικής σημασίας για την ανάπτυξη της διακυβέρνησης της ζήτησης και την προώθηση της ενεργειακής απόδοσης, καθώς μπορεί να βοηθήσει στην αποτροπή ακραίων καταστάσεων και την ανακούφιση του δικτύου σε περιόδους υψηλής ζήτησης. Συνεπώς, η βραχυπρόθεσμη πρόβλεψη ζήτησης φορτίου με χρήση νευρωνικών δικτύων είναι μια εξαιρετικά σημαντική τεχνική για την αξιολόγηση και βελτίωση της απόδοσης των συστημάτων ηλεκτρικής ενέργειας.

ΚΕΦΑΛΑΙΟ 2: Φορτίο

2.1 Γενικά

Παλιότερα, η κεντρική παραγωγή ήταν ικανή να προσφέρει εφεδρείες στο σύστημα ανάλογα με την ζήτηση, όπου η απώλεια παραγωγής και μεταφοράς δεν έπαιζαν τόσο σημαντικό τεχνικό και οικονομικό ρόλο καθώς θεωρούνταν απαραίτητη προϋπόθεση για την ανάπτυξη κάποιας περιοχής ή χώρας και είχαν καθαρά εθνικό χαρακτήρα. Πλέον, η είσοδος αποκεντρωμένης παραγωγής, ο διπλός ρόλος που καλείται να παίξει το δίκτυο μεταφοράς και διανομής, η ένταξη νέων κλάδων κατανάλωσης (Ηλεκτρικά αυτοκίνητα), η απελευθέρωση της αγοράς ενέργειας κ.α., κάνει αναγκαία την διεξόδυση τεχνολογιών πρόβλεψης ικανών να προσφέρουν οικονομικές (επενδυτές, καταναλωτές) και τεχνικές (διαχειριστή) λύσεις στα προβλήματα που έχουν προκύψει. Τα προβλήματα αυτά αφορούν την ευστάθεια του συστήματος, την τιμή της ηλεκτρικής ενέργειας, την επάρκεια του συστήματος, την ανθεκτικότητα κ.α. Η ψηφιοποίηση του συστήματος και οι τεχνικές προβλέψεων με χρήση νευρωνικών δικτύων έχουν την ικανότητα να προσφέρουν λύσεις τόσο σε οικονομικό όσο και σε τεχνικό επίπεδο.

2.2 Φορτίο

Γενικά, το φορτίο μίας ηλεκτρικής χρήσης απαρτίζεται από πολύ διαφορετικές καταναλωτικές μονάδες. Ένα μεγάλο μέρος του ηλεκτρισμού καταναλώνεται σε βιομηχανικές δραστηριότητες. Ένα άλλο μέρος φυσικά χρησιμοποιείται από ιδιώτες με σκοπό τη θέρμανση, το φωτισμό, το μαγείρεμα κ.ά. Επίσης, πολλές κοινωφελείς υπηρεσίες απαιτούν ηλεκτρισμό, π.χ. ο φωτισμός των δρόμων, ο ηλεκτρικός σιδηρόδρομος κ.ά.

Οι παράγοντες που επηρεάζουν το φορτίο εξαρτώνται από τις επιμέρους μονάδες κατανάλωσης. Το βιομηχανικό φορτίο συνήθως καθορίζεται περισσότερο από το επίπεδο της παραγωγής. Το φορτίο είναι συνήθως αρκετά σταθερό και είναι δυνατό να εκτιμηθεί η εξάρτηση του σε διαφορετικά επίπεδα παραγωγής. Παρ' όλα αυτά, από την σκοπιά της πώλησης ενέργειας, η βιομηχανική χρήση συχνά προσθέτει αβεβαιότητα στις προβλέψεις. Το πρόβλημα είναι η πιθανότητα απροσδόκητων συμβάντων, όπως μηχανικές βλάβες ή απεργίες, οι οποίες μπορούν να προκαλέσουν μεγάλες απρόβλεπτες διακυμάνσεις στο επίπεδο φορτίου. Στην περίπτωση των ιδιωτών, οι παράγοντες που επηρεάζουν το φορτίο είναι πολύ πιο δύσκολο να καθοριστούν. Κάθε άνθρωπος συμπεριφέρεται με τον δικό του ιδιαίτερο τρόπο και η ανθρώπινη ψυχολογία σχετίζεται με κάθε καταναλωτική απόφαση. Πολλοί κοινωνικοί παράγοντες αλλά και παράγοντες συμπεριφοράς μπορούν να παρατηρηθούν. Για παράδειγμα μεγάλα γεγονότα, αργίες ακόμη και τηλεοπτικά προγράμματα επηρεάζουν το φορτίο. Ο καιρός είναι ο πιο σημαντικός παράγοντας κι αυτό εξαιτίας της

ηλεκτρικής θέρμανσης των σπιτιών, η οποία γίνεται πιο έντονη όσο μειώνεται η θερμοκρασία .

Γενικά, θα λέγαμε ότι οι παράγοντες οι οποίοι επηρεάζουν το ηλεκτρικό φορτίο μπορούν να χωριστούν σε τέσσερις κατηγορίες:

- Οικονομικοί
- Χρονικοί
- Καιρικοί
- Τυχαίοι

Για τη δημιουργία ενός μοντέλου που περιγράφει το φορτίο θα πρέπει να γίνει κατανοητή η επίδραση όλων των παραγόντων στη διαμόρφωση της ζήτησης. Ακολουθεί μία σύντομη περιγραφή της επίδρασης των παραπάνω παραγόντων στη διαμόρφωση της ζήτησης της ηλεκτρικής ενέργειας.

Οικονομικοί Παράγοντες

Το οικονομικό περιβάλλον εντός του οποίου λειτουργεί μία ηλεκτρική εταιρεία επιδρά αναμφίβολα στη διαμόρφωση της ηλεκτρικής ζήτησης. Δημογραφικοί παράγοντες, επίπεδα βιομηχανικής δραστηριότητας, ανάπτυξη ή και κορεσμός της αγοράς νέων ηλεκτρικών συσκευών, αλλαγές στον αγροτικό τομέα, κρατικές παρεμβάσεις και γενικά η οικονομική δραστηριότητα, έχουν μεγάλη επίδραση στο ρυθμό αύξησης ή μείωσης του φορτίου του συστήματος. Αλλαγές στην τιμολογιακή πολιτική της επιχείρησης, καθώς και προγράμματα εξοικονόμησης ενέργειας και διαχείρισης της ζήτησης επιδρούν επίσης στη διαμόρφωση της ζήτησης. Τυπικά, οι οικονομικοί παράγοντες έχουν χρονικές σταθερές σημαντικά μεγαλύτερες από μία εβδομάδα και δεν επηρεάζουν βραχυπρόθεσμα τη ζήτηση.

Χρονικοί Παράγοντες

Τρεις βασικοί χρονικοί παράγοντες - εποχιακές μεταβολές, εβδομαδιαίος/ ημερήσιος κύκλος και εθνικές και θρησκευτικές εορτές/αργίες επηρεάζουν σημαντικά το ηλεκτρικό φορτίο. Οι εποχιακές μεταβολές επιδρούν στη μορφή της καμπύλης φορτίου. Στο σχήμα 1 δίνεται μία τυπική ημερήσια καμπύλη φορτίου χειμώνα και μία θέρους του εθνικού διασυνδεδεμένου συστήματος. Με απλή εποπτεία, παρατηρούμε τις διαφορές στην ώρα εμφάνισης της βραδινής αιχμής, καθώς και την τρίτη αιχμή που εμφανίζεται το χειμώνα λόγω της θερμοσυσσώρευσης. Ορισμένες μεταβολές στην καμπύλη φόρτισης γίνονται σταδιακά με την αλλαγή διαφόρων εποχιακών μεταβλητών, όπως η θερμοκρασία και οι ώρες ηλιοφάνειας. Άλλες εποχιακές μεταβολές δημιουργούν απότομες αλλαγές στην καμπύλη φόρτισης. Τέτοιες αλλαγές είναι η αλλαγή ώρας (Χειμερινή – Θερινή) (σε αυτήν την περίπτωση παρατηρείται μετατόπιση της καμπύλης φορτίου), η έναρξη της σχολικής χρονιάς, η πιθανή εποχιακή μεταβολή στις χρονικές ζώνες φθηνού τιμολογίου κ.ά.

Η εβδομαδιαία/ημερήσια περιοδικότητα του φορτίου είναι αποτέλεσμα της περιοδικότητας του κύκλου εργασίας - ανάπαυσης του πληθυσμού. Υπάρχουν προκαθορισμένες μορφές καμπυλών φορτίου για μία «τυπική» εβδομάδα κάθε εποχής.

Το Σαββατοκύριακο η ζήτηση είναι μειωμένη και το επίπεδο της ζήτησης, καθώς και η μορφή της καμπύλης φορτίου, της Δευτέρας και της Παρασκευής είναι κάπως διαφορετική από τις υπόλοιπες εργάσιμες μέρες λόγω γειννίας με το Σαββατοκύριακο. Η Τρίτη, η Τετάρτη και η Πέμπτη έχουν συνήθως παρόμοιες καμπύλες φορτίου «εργάσιμης μέρας». Τέλος, η ζήτηση φορτίου κατά τη διάρκεια των εθνικών ή των θρησκευτικών εορτών είναι σημαντικά χαμηλότερη από τα φυσιολογικά επίπεδα. Σημαντική μείωση της ζήτησης παρατηρείται επιπλέον και κατά τη διάρκεια των εθνικών ή των θρησκευτικών εορτών όπου είναι σημαντικά χαμηλότερη από τα φυσιολογικά επίπεδα. Σημαντική μείωση της ζήτησης παρατηρείται επιπλέον και κατά τη διάρκεια εορταστικών τριημέρων

Καιρικοί Παράγοντες

Οι καιρικές συνθήκες προκαλούν σημαντικές αλλαγές στην καμπύλη φορτίου. Αυτό συμβαίνει διότι η λειτουργία ενός μεγάλου ποσοστού ηλεκτρικών συσκευών (θέρμανση, κλιματισμός, αρδεύσεις) εξαρτάται από τις καιρικές συνθήκες. Σε πολλά συστήματα, η θερμοκρασία έχει τη σημαντικότερη επίδραση στη διαμόρφωση του φορτίου. Για συστήματα τα οποία καλύπτουν μεγάλη γεωγραφική έκταση χρειάζεται να ληφθούν υπόψη οι θερμοκρασίες σε διάφορες περιοχές για να υπολογιστεί η ακριβής επίδραση στο φορτίο. Άλλοι παράγοντες, όπως υγρασία, βροχόπτωση, άνεμος, νεφώσεις, επηρεάζουν επίσης τη ζήτηση [3].

Οι καιρικές συνθήκες είναι ένας από τους πιο σημαντικούς παράγοντες που επηρεάζουν τη ζήτηση ενέργειας, που κυμαίνονται από μεμονωμένους καταναλωτές έως ολόκληρες χώρες. Συγκεκριμένα, η θερμοκρασία, οι βροχοπτώσεις, οι χιονοπτώσεις και ο άνεμος επηρεάζουν σημαντικά τη ζήτηση ηλεκτρικής ενέργειας. Η θερμοκρασία είναι ένας από τους πιο κρίσιμους παράγοντες που επηρεάζουν τη ζήτηση ηλεκτρικής ενέργειας. Τους καλοκαιρινούς μήνες η ζήτηση για ηλεκτρική ενέργεια αυξάνεται λόγω των υψηλών θερμοκρασιών, καθώς οι καταναλωτές χρησιμοποιούν κλιματισμό για να δροσιστούν. Αντίθετα, κατά τους χειμερινούς μήνες, η ζήτηση ηλεκτρικής ενέργειας αυξάνεται λόγω του κρύου, καθώς οι καταναλωτές χρησιμοποιούν συστήματα θέρμανσης για να ζεσταθούν. Ως αποτέλεσμα, οι εποχιακές διακυμάνσεις της θερμοκρασίας έχουν άμεσο αντίκτυπο στη ζήτηση ενέργειας. Οι βροχοπτώσεις και οι χιονοπτώσεις επηρεάζουν επίσης τη ζήτηση ενέργειας. Κατά τη διάρκεια μιας καταιγίδας, η ζήτηση για ηλεκτρική ενέργεια μπορεί να αυξηθεί λόγω της ανάγκης για θέρμανση και φωτισμό, καθώς και την πιθανότητα διακοπής ρεύματος. Αυτό ισχύει ιδιαίτερα σε περιοχές που είναι επιρρεπείς σε πλημμύρες ή έντονες χιονοπτώσεις, όπου μπορεί να προκληθούν διακοπές ρεύματος λόγω ζημιών σε ηλεκτροφόρα καλώδια ή άλλες υποδομές. Ο άνεμος είναι ένας άλλος σημαντικός παράγοντας που επηρεάζει τη ζήτηση ηλεκτρικής ενέργειας, ειδικά σε περιοχές με υψηλές ταχύτητες ανέμου. Η αιολική ενέργεια είναι μια αυξανόμενη πηγή ανανεώσιμης ενέργειας και πολλές χώρες επενδύουν στην αιολική ενέργεια ως τρόπο να μειώσουν την εξάρτησή τους από τα ορυκτά καύσιμα. Ωστόσο, οι ισχυροί άνεμοι μπορούν επίσης να προκαλέσουν διακοπές ρεύματος, ειδικά εάν οι ανεμογεννήτριες δεν είναι σχεδιασμένες να αντέχουν σε ακραίες καιρικές συνθήκες.

Εκτός από αυτούς τους μεμονωμένους παράγοντες, τα καιρικά μοτίβα και τα ακραία καιρικά φαινόμενα μπορούν επίσης να επηρεάσουν τη ζήτηση ενέργειας σε εθνικό επίπεδο. Για παράδειγμα, κατά τη διάρκεια ενός καύσωνα ή ενός ψύχους, η ζήτηση για ηλεκτρική ενέργεια μπορεί να αυξηθεί σε μια ολόκληρη περιοχή ή χώρα. Σε ορισμένες περιπτώσεις, αυτή η αυξημένη ζήτηση μπορεί να υπερβεί τη διαθέσιμη προσφορά, οδηγώντας σε διακοπές ρεύματος ή άλλες διακοπές ρεύματος. Ομοίως, ακραία καιρικά φαινόμενα όπως τυφώνες, πλημμύρες ή πυρκαγιές μπορούν να έχουν σημαντικό αντίκτυπο στη ζήτηση ενέργειας. Εκτός από την πρόκληση φυσικής ζημιάς σε γραμμές ηλεκτροδότησης και άλλες υποδομές, αυτά τα γεγονότα μπορούν επίσης να διαταράξουν τις αλυσίδες εφοδιασμού και να προκαλέσουν αύξηση των τιμών της ενέργειας, οδηγώντας σε αυξημένη ζήτηση ενέργειας από άλλες πηγές. Συνολικά, οι καιρικές συνθήκες είναι ένας σημαντικός παράγοντας που επηρεάζει τη ζήτηση ενέργειας, από μεμονωμένους καταναλωτές έως μεγαλύτερα συστήματα όπως οι χώρες. Καθώς ο κόσμος συνεχίζει να αντιμετωπίζει πιο συχνά και έντονα καιρικά φαινόμενα λόγω της κλιματικής αλλαγής, η διαχείριση της ζήτησης ενέργειας θα γίνεται όλο και πιο κρίσιμη για τη διασφάλιση αξιόπιστου και βιώσιμου ενεργειακού εφοδιασμού.

Τυχαίοι Παράγοντες

Σε αυτήν την κατηγορία ανήκουν όλοι οι παράγοντες που επηρεάζουν το φορτίο και δεν ανήκουν σε καμία από τις προηγούμενες κατηγορίες. Η λειτουργία του συστήματος δέχεται συνεχώς διαταραχές οι οποίες οφείλονται στο γεγονός ότι ζήτηση του φορτίου συνίσταται από ένα μεγάλο αριθμό επί μέρους ζητήσεων των καταναλωτών. Εκτός από τους χιλιάδες μικρούς καταναλωτές, υπάρχουν και μεγάλοι καταναλωτές (χαλυβουργεία κ.ά.) που προκαλούν σχετικά μεγάλες τυχαίες διαταραχές στη ζήτηση φορτίου. Υπάρχουν επίσης ορισμένα άλλα γεγονότα, όπως μεγάλες απεργίες, εκλογές, ειδικά προγράμματα στην τηλεόραση, τα οποία, αν και είναι εκ των προτέρων γνωστά, είναι αρκετά δύσκολο να εκτιμηθεί η επίδρασή τους στη διαμόρφωση της ζήτησης.

Εκτός από τις καιρικές συνθήκες, υπάρχουν πολλοί άλλοι παράγοντες που μπορούν να επηρεάσουν τη ζήτηση ενέργειας, που κυμαίνονται από μεμονωμένους καταναλωτές έως ολόκληρες χώρες. Μερικοί από αυτούς τους παράγοντες είναι τυχαίοι και δύσκολο να προβλεφθούν, γεγονός που καθιστά δύσκολο για τους παρόχους ενέργειας να σχεδιάσουν και να διαχειριστούν τις διακυμάνσεις της ζήτησης. Ένας από τους πιο σημαντικούς τυχαίους παράγοντες που μπορεί να επηρεάσει τη ζήτηση ενέργειας είναι η ανθρώπινη συμπεριφορά. Οι καταναλωτές μπορούν να αλλάξουν τα πρότυπα κατανάλωσης ενέργειας με βάση διάφορους παράγοντες, συμπεριλαμβανομένων των προγραμμάτων εργασίας, των κοινωνικών εκδηλώσεων και των προσωπικών προτιμήσεων. Για παράδειγμα, κατά την περίοδο των διακοπών, η ζήτηση ενέργειας μπορεί να αυξηθεί λόγω των αυξημένων αναγκών φωτισμού και θέρμανσης, καθώς οι άνθρωποι περνούν περισσότερο χρόνο στο σπίτι. Ένας άλλος τυχαίος παράγοντας που μπορεί να επηρεάσει τη ζήτηση ενέργειας είναι η τεχνολογική καινοτομία. Η εισαγωγή νέων τεχνολογιών, όπως τα ηλεκτρικά οχήματα και τα έξυπνα

σπίτια, μπορεί να αλλάξει τον τρόπο με τον οποίο οι καταναλωτές χρησιμοποιούν και αλληλοεπιδρούν με την ενέργεια. Για παράδειγμα, καθώς περισσότεροι καταναλωτές στρέφονται στα ηλεκτρικά οχήματα, η ζήτηση ηλεκτρικής ενέργειας για την τροφοδοσία αυτών των οχημάτων θα αυξάνεται, οδηγώντας σε νέες προκλήσεις για τους παρόχους ενέργειας.

Οι οικονομικές συνθήκες μπορούν επίσης να έχουν σημαντικό αντίκτυπο στη ζήτηση ενέργειας. Σε περιόδους οικονομικής ανάπτυξης, η ζήτηση ενέργειας μπορεί να αυξηθεί καθώς οι καταναλωτές και οι επιχειρήσεις έχουν περισσότερα χρήματα να ξοδέψουν σε δραστηριότητες έντασης ενέργειας, όπως η κατασκευή και οι μεταφορές. Αντίθετα, σε περιόδους οικονομικής ύφεσης, η ζήτηση ενέργειας μπορεί να μειωθεί καθώς οι καταναλωτές και οι επιχειρήσεις περιορίζουν τις μη βασικές δαπάνες. Οι κυβερνητικές πολιτικές και κανονισμοί μπορούν επίσης να επηρεάσουν τη ζήτηση ενέργειας. Για παράδειγμα, οι επιδοτήσεις για ανανεώσιμες πηγές ενέργειας μπορούν να αυξήσουν τη ζήτηση για αυτές τις πηγές, ενώ οι κανονισμοί για την ενεργειακή απόδοση μπορούν να μειώσουν τη συνολική ζήτηση ενέργειας. Επιπλέον, γεωπολιτικά γεγονότα, όπως πόλεμοι ή φυσικές καταστροφές, μπορούν να διαταράξουν τις αλυσίδες εφοδιασμού ενέργειας και να οδηγήσουν σε ξαφνικές αυξήσεις της ζήτησης καθώς οι χώρες προσπαθούν να εξασφαλίσουν τον ενεργειακό τους εφοδιασμό. Σε εθνικό επίπεδο, οι δημογραφικές αλλαγές μπορούν επίσης να επηρεάσουν τη ζήτηση ενέργειας. Για παράδειγμα, καθώς οι πληθυσμοί γερνούν, η ζήτηση για υπηρεσίες υγειονομικής περίθαλψης μπορεί να αυξηθεί, οδηγώντας σε αντίστοιχη αύξηση της ζήτησης ενέργειας για νοσοκομεία και άλλες εγκαταστάσεις υγειονομικής περίθαλψης. Ομοίως, καθώς οι αστικοί πληθυσμοί αυξάνονται, η ζήτηση για υποδομές έντασης ενέργειας, όπως οι μεταφορές και τα κτίρια, μπορεί να αυξηθεί.

Συνολικά, υπάρχουν πολλοί τυχαίοι παράγοντες που μπορούν να επηρεάσουν τη ζήτηση ενέργειας, καθιστώντας δύσκολο για τους παρόχους ενέργειας να σχεδιάσουν και να διαχειριστούν τις διακυμάνσεις της ζήτησης. Για την αντιμετώπιση αυτών των προκλήσεων, πολλοί πάροχοι ενέργειας επενδύουν σε νέες τεχνολογίες όπως έξυπνα δίκτυα και συστήματα αποθήκευσης ενέργειας, καθώς και σε στενή συνεργασία με φορείς χάραξης πολιτικής και άλλους ενδιαφερόμενους φορείς για την ανάπτυξη στρατηγικών για τη διαχείριση της ζήτησης ενέργειας σε έναν ταχέως μεταβαλλόμενο κόσμο.

Ποιοι, όμως, από όλους αυτούς τους προαναφερθέντες παράγοντες λαμβάνονται υπόψη στη μελέτη μας; Ένα μεγάλο κομμάτι της κατανάλωσης οφείλεται σε ιδιώτες και άλλους μικρούς ηλεκτρικούς καταναλωτές. Μία συνήθης προσέγγιση της πρόβλεψης φορτίου είναι να επικεντρωθούμε στο συνολικό φορτίο. Αυτή είναι και η προσέγγιση που θα κάνουμε στη μελέτη μας. Το γεγονός αυτό μειώνει τον αριθμό των παραγόντων που πρέπει να ληφθούν υπόψη [5]. Οι πιο σημαντικοί είναι :

- Βραχυπρόθεσμα, οι μετεωρολογικές συνθήκες προκαλούν μεγάλες αποκλίσεις στο συνολικό φορτίο. Εκτός της θερμοκρασίας, και η ταχύτητα του ανέμου, τα σύννεφα και η υγρασία ασκούν κάποια επιρροή.
- Μακροπρόθεσμα, οικονομικοί και δημογραφικοί παράγοντες παίζουν τον πιο σημαντικό ρόλο στον καθορισμό της εξέλιξης της ηλεκτρικής ζήτησης.
- Από τη σκοπιά της πρόβλεψης, οι χρονικοί παράγοντες είναι βασικοί. Διάφορες περιοδικές επιδράσεις και επαναλαμβανόμενες καταστάσεις καθώς και νόμιμες και θρησκευτικές αργίες λαμβάνονται υπόψη.

Οι άλλοι παράγοντες που προκαλούν διαταραχές και μπορούν να καταταχθούν σαν τυχαίοι παράγοντες είναι συνήθως μικροί στην περίπτωση των ατομικών καταναλωτών. Παρόλ' αυτά μεγάλα κοινωνικά γεγονότα και δημοφιλή τηλεοπτικά προγράμματα προσθέτουν αβεβαιότητα στις προβλέψεις. Από την άλλη βιομηχανικές μονάδες μπορούν να προκαλέσουν σχετικά μεγάλες διαταραχές.

Η συγκέντρωση της ζήτησης ενέργειας σε επίπεδο χώρας μπορεί να συμβάλει στην ελαχιστοποίηση των διακυμάνσεων της συμπεριφοράς των καταναλωτών, δημιουργώντας μια πιο σταθερή και προβλέψιμη καμπύλη ζήτησης. Όταν η ζήτηση ενέργειας κατανέμεται σε μεγάλο αριθμό καταναλωτών, ο αντίκτυπος της ατομικής συμπεριφοράς στη συνολική ζήτηση μειώνεται. Για παράδειγμα, ενώ ένας μεμονωμένος καταναλωτής μπορεί να αλλάζει τα πρότυπα κατανάλωσης ενέργειας σε καθημερινή ή εβδομαδιαία βάση, η συνολική ενεργειακή ζήτηση για μια χώρα είναι πολύ πιο σταθερή και προβλέψιμη για μεγαλύτερες χρονικές περιόδους. Αυτή η σταθερότητα στη ζήτηση μπορεί να είναι χρήσιμη για τους παρόχους ενέργειας, καθώς τους επιτρέπει να σχεδιάζουν και να διαχειρίζονται καλύτερα τις διακυμάνσεις της προσφοράς και της ζήτησης. Έχοντας μια ακριβέστερη κατανόηση της συνολικής ζήτησης ενέργειας, οι πάροχοι μπορούν να λάβουν καλύτερες αποφάσεις σχετικά με το πότε και πού να επενδύσουν σε νέες υποδομές και τεχνολογίες και πώς να καταναείμουν τους πόρους τους πιο αποτελεσματικά. Επιπλέον, η συγκέντρωση της ζήτησης ενέργειας μπορεί επίσης να συμβάλει στη μείωση του συνολικού κόστους ενέργειας για τους καταναλωτές. Κατανέμοντας το πάγιο κόστος παραγωγής και διανομής ενέργειας σε μεγαλύτερο αριθμό καταναλωτών, το κόστος ανά μονάδα ενέργειας μπορεί να μειωθεί, καθιστώντας την ενέργεια πιο προσιτή για όλους. Συνολικά, η συγκέντρωση της ενεργειακής ζήτησης σε επίπεδο χώρας μπορεί να προσφέρει μια σειρά πλεονεκτημάτων τόσο για τους καταναλωτές όσο και για τους παρόχους ενέργειας, συμπεριλαμβανομένης μεγαλύτερης σταθερότητας στη ζήτηση, βελτιωμένου σχεδιασμού και διαχείρισης και χαμηλότερου κόστους ενέργειας.

Οι προβλέψεις της επόμενης ημέρας και οι βραχυπρόθεσμες προβλέψεις της ενεργειακής ζήτησης επηρεάζονται από μια σειρά παραγόντων, συμπεριλαμβανομένων των καιρικών συνθηκών, της ώρας της ημέρας, της ημέρας της εβδομάδας και τυχόν προγραμματισμένες ή απρογραμματίστες διακοπές στον ενεργειακό εφοδιασμό. Οι καιρικές συνθήκες, ειδικότερα, μπορούν να έχουν σημαντικό αντίκτυπο στη ζήτηση ενέργειας, καθώς οι άνθρωποι τείνουν να χρησιμοποιούν περισσότερη ενέργεια για θέρμανση ή ψύξη σε ακραίες θερμοκρασίες. Επιπλέον, οι μεγαλύτερες προβλέψεις για

τη ζήτηση ενέργειας επηρεάζονται από περισσότερους μακροοικονομικούς παράγοντες, όπως η πληθυσμιακή αύξηση, η οικονομική δραστηριότητα και οι αλλαγές στην τεχνολογία και τις ενεργειακές πολιτικές. Για παράδειγμα, μια αύξηση του πληθυσμού ή της οικονομικής δραστηριότητας μπορεί να οδηγήσει σε υψηλότερη ζήτηση ενέργειας με την πάροδο του χρόνου, ενώ η υιοθέτηση πιο ενεργειακά αποδοτικών τεχνολογιών ή πολιτικών μπορεί να οδηγήσει σε χαμηλότερη ζήτηση. Άλλοι παράγοντες που μπορούν να επηρεάσουν τις μακροπρόθεσμες προβλέψεις της ενεργειακής ζήτησης περιλαμβάνουν αλλαγές στη συμπεριφορά των καταναλωτών, όπως η αυξημένη υιοθέτηση ηλεκτρικών οχημάτων ή οι αλλαγές στα πρότυπα εργασίας από το σπίτι και η διαθεσιμότητα εναλλακτικών πηγών ενέργειας, όπως η ηλιακή ή η αιολική ενέργεια. Συνολικά, οι ακριβείς προβλέψεις της ενεργειακής ζήτησης είναι κρίσιμες για τους παρόχους ενέργειας προκειμένου να διασφαλίσουν ότι μπορούν να ανταποκριθούν στις ανάγκες των καταναλωτών και να διατηρήσουν έναν αξιόπιστο και σταθερό ενεργειακό εφοδιασμό. Λαμβάνοντας υπόψη μια σειρά παραγόντων, συμπεριλαμβανομένων τόσο των βραχυπρόθεσμων όσο και των μακροπρόθεσμων τάσεων, οι πάροχοι ενέργειας μπορούν να λάβουν πιο ενημερωμένες αποφάσεις σχετικά με το πότε και πού να επενδύσουν σε νέες υποδομές και τεχνολογίες και πώς να καταναείμουν τους πόρους τους πιο αποτελεσματικά.

ΚΕΦΑΛΑΙΟ 3: Μέθοδοι πρόβλεψης & Χρονοσειρές

3.1 Γενικά

Η πρόβλεψη όπως έχει τονιστεί είναι ιδιαίτερα σημαντική σε ένα Σύστημα Ηλεκτρικής Ενέργειας. Ο συνολικός χρόνος πρόβλεψης φτάνει από την μία βδομάδα έως και τον επόμενο μήνα με συνηθέστερες όμως εφαρμογές αυτές της πρόβλεψης της ζήτησης για την επόμενη μέρα. Παρακάτω αναλύονται τα βασικά χαρακτηριστικά των χρονοσειρών που είναι αναγκαία για την κατανόηση του τρόπου εργασίας. Στη συνέχεια παρουσιάζονται περιληπτικά οι τεχνικές προβλέψεων που θα χρησιμοποιήσει η διπλωματική. Σε αυτές γίνεται διάκριση μεταξύ των απλών στατιστικών μοντέλων και των τεχνικών μηχανική μάθησης όπως είναι τα νευρωνικά δίκτυα. Για τα νευρωνικά δίκτυα γίνεται αναφορά και στην συγκεκριμένη αρχιτεκτονική του κάθε μοντέλου [6] [7].

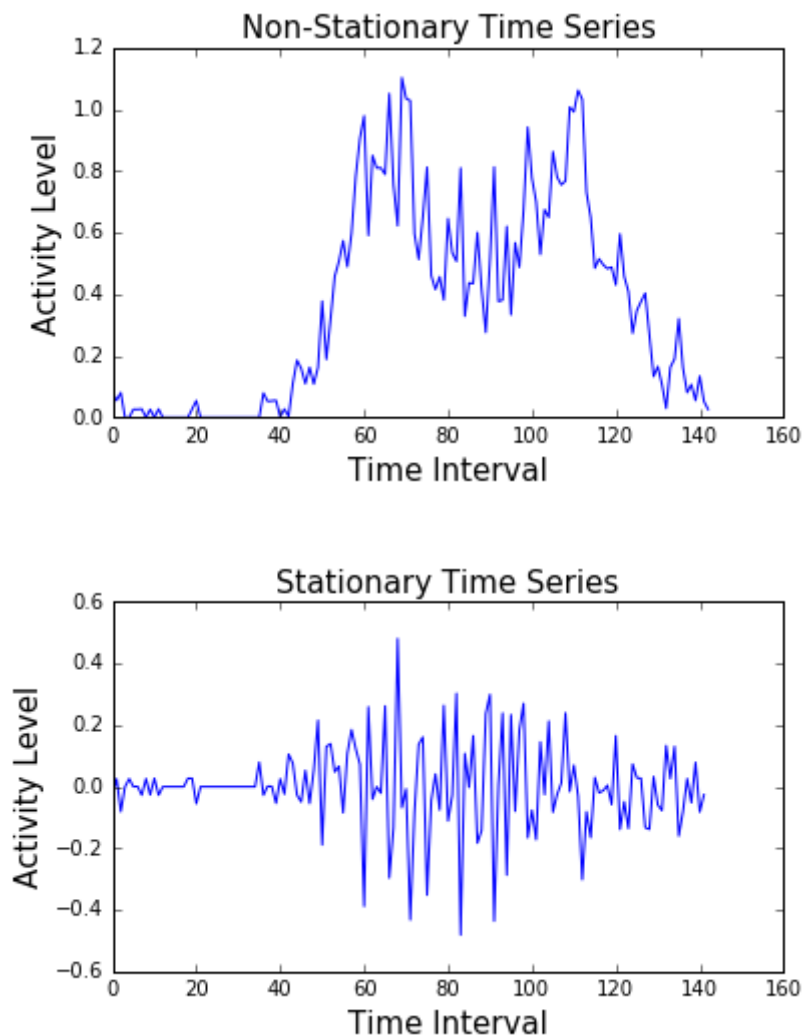
3.2 Χρονοσειρές

3.2.1 Χαρακτηριστικά χρονοσειρών

Η ανάλυση χρονοσειρών και η πρόβλεψη δεδομένων είναι ένα από τα πλέον χρήσιμα και διαδεδομένα εργαλεία προκειμένου να στηριχθεί η λήψη αποφάσεων από διοικητικά στελέχη, επιχειρήσεις, πολυεθνικές εταιρίες και κυβερνήσεις. Με τον όρο χρονοσειρά (time series) εννοούμε μία σειρά από δεδομένα τα οποία συλλέγονται διαχρονικά και εκφράζουν την εξέλιξη μιας μεταβλητής κατά τη διάρκεια ίσων χρονικών στιγμών ή περιόδων. Πιο συγκεκριμένα, η χρονοσειρά αποτελείται από ένα σύνολο παρατηρήσεων μιας μεταβλητής της οποίας οι τιμές αναφέρονται με βάση κάποια χρονική περίοδο π.χ. μέρα, εβδομάδα, μήνας, τρίμηνο κ.α. Μαθηματικά η χρονοσειρά ορίζεται από ένα δείγμα τιμών Y_1, Y_2, \dots, Y_t της μεταβλητής Y για κάθε χρονική στιγμή t . Επομένως, το Y είναι μία συνάρτηση του t , και αυτό συμβολίζεται ως $Y = f(t)$ ενώ η γραφική παράσταση της εν λόγω συνάρτησης παρουσιάζει την εξέλιξη της μεταβλητής Y στο χρόνο.

Στασιμότητα: Η στασιμότητα (stationarity) είναι μία από τις σημαντικότερες έννοιες, όσον αφορά τις χρονοσειρές, διότι είναι απαραίτητη προϋπόθεση για τα περισσότερα εργαλεία ανάλυσης. Γενικά, θα λέγαμε ότι μία χρονοσειρά είναι στάσιμη, εφόσον τα δεδομένα της παραμένουν διαχρονικά σε ένα σταθερό μέσο επίπεδο, ενώ αντίθετα στην περίπτωση που τα χαρακτηριστικά της μεταβάλλονται διαχρονικά, τότε αυτή θεωρείται μη στάσιμη. Πιο συγκεκριμένα, μία χρονοσειρά ορίζεται ως στάσιμη όταν η συνάρτηση κατανομής των μετρήσεων είναι ανεξάρτητη του χρόνου. Μία μη στάσιμη χρονοσειρά παρουσιάζει διακυμάνσεις των τιμών της με την πάροδο του χρόνου και μπορεί να

αναλυθεί σε τέσσερις συνιστώσες, οι οποίες είναι η τάση, η κυκλική συνιστώσα, η εποχικότητα και τα κατάλοιπα.



1-Σχήμα 3.1: Example of a Stationary and a Non-Stationary Time Series

Πιο συγκεκριμένα:

- Η **τάση** (Trend T_t) αντιπροσωπεύει τη γενική μακροχρόνια συμπεριφορά της χρονοσειράς, η οποία μπορεί να είναι είτε αυξητική είτε με τάση μείωσης. Επίσης, μπορεί να αποτελείται είτε από μία γραμμική ή μη-γραμμική διαδικασία. Η τάση είναι ένα από τα πιο σημαντικά χαρακτηριστικά μιας χρονοσειράς. Αναφέρεται στη γενική κατεύθυνση των δεδομένων με την πάροδο του χρόνου, αντικατοπτρίζοντας τις μακροπρόθεσμες αλλαγές ή την ανάπτυξη στη σειρά. Με άλλα λόγια, η τάση είναι η συστηματική συνιστώσα της χρονοσειράς που αποτυπώνει τα υποκείμενα μοτίβα ή τις δομικές αλλαγές στα δεδομένα. Οι τάσεις μπορεί να είναι είτε αυξανόμενες είτε φθίνουσες ή μπορεί να είναι επίπεδες ή γραμμικές. Για παράδειγμα, μια χρονοσειρά δεδομένων μηνιαίων πωλήσεων για ένα προϊόν μπορεί να παρουσιάζει

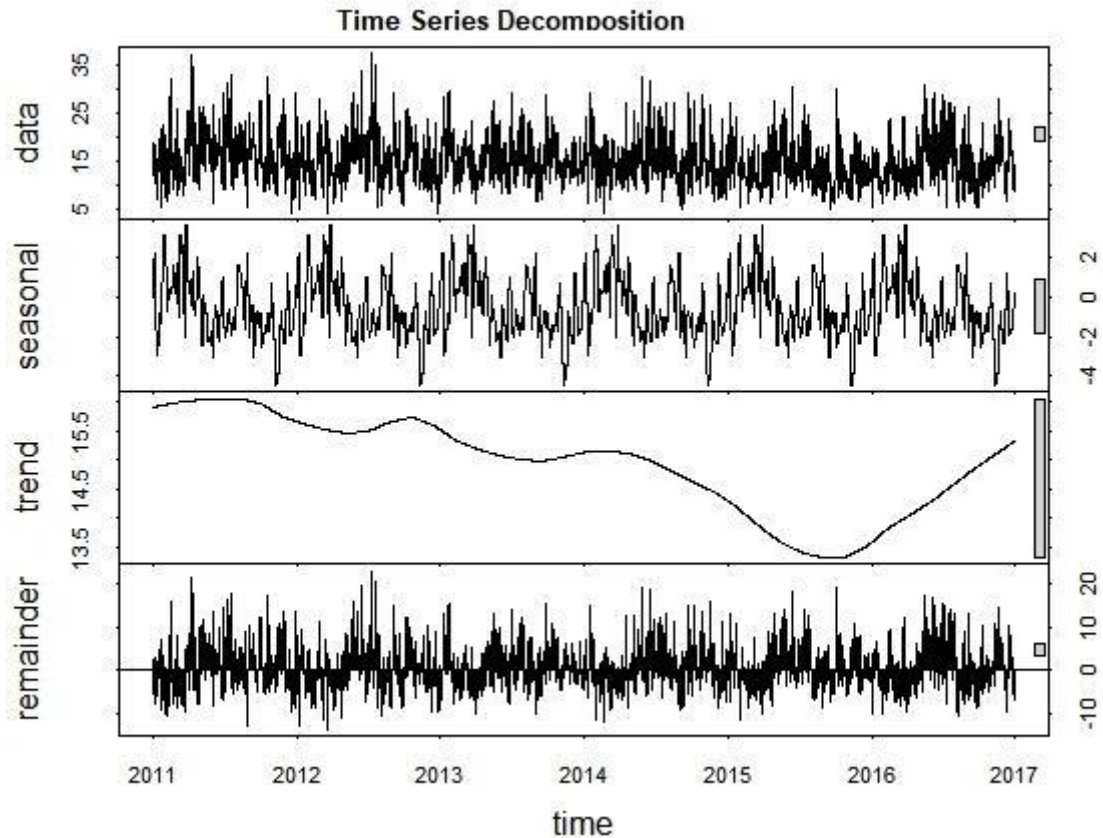
αυξητική τάση εάν οι πωλήσεις αυξάνονται σταθερά με την πάροδο του χρόνου, υποδηλώνοντας αυξανόμενη ζήτηση για το προϊόν. Από την άλλη πλευρά, μια χρονική σειρά μηνιαίων ποσοστών ανεργίας μπορεί να παρουσιάσει πτωτική τάση εάν τα ποσοστά ανεργίας μειώνονται σταθερά με την πάροδο του χρόνου, υποδηλώνοντας μια βελτίωση της αγοράς εργασίας. Οι τάσεις μπορούν να επηρεαστούν από διάφορους παράγοντες όπως η πληθυσμιακή αύξηση, οι τεχνολογικές εξελίξεις, οι οικονομικοί κύκλοι και οι δημογραφικές αλλαγές. Η ανίχνευση και η μοντελοποίηση τάσεων στα δεδομένα χρονοσειρών είναι ζωτικής σημασίας για πολλές εφαρμογές, συμπεριλαμβανομένων των προβλέψεων, της ανίχνευσης ανωμαλιών και της λήψης αποφάσεων. Μια δημοφιλής μέθοδος για τον εντοπισμό τάσεων είναι η χρήση ανάλυσης παλινδρόμησης για την προσαρμογή ενός γραμμικού ή μη γραμμικού μοντέλου στα δεδομένα. Μια άλλη μέθοδος είναι η χρήση τεχνικών αποσύνθεσης όπως το φίλτρο Hodrick-Prescott ή η αποσύνθεση εποχικής τάσης (STL) για την απομόνωση και την εκτίμηση της συνιστώσας τάσης των δεδομένων.

- Η **κυκλική διακύμανση** (Cycle C_t) αντιπροσωπεύει και αυτή τη μακροχρόνια συμπεριφορά της χρονοσειράς και αποτελείται από περιοδικές διακυμάνσεις οι οποίες διαρκούν περισσότερο της μίας χρονικής περιόδου. Το εύρος και η διάρκεια των διακυμάνσεων αυτών δεν είναι σταθερή. Ο κύκλος είναι ένα άλλο σημαντικό χαρακτηριστικό μιας χρονοσειράς, που αντιπροσωπεύει τις διακυμάνσεις ή τις ταλαντώσεις στα δεδομένα γύρω από τη μακροπρόθεσμη τάση της. Αναφέρεται στα επαναλαμβανόμενα μοτίβα ή περιοδικές διακυμάνσεις στα δεδομένα που εμφανίζονται σε μια χρονική περίοδο, αλλά όχι απαραίτητα σε τακτά χρονικά διαστήματα. Η διάρκεια του κύκλου μπορεί να ποικίλλει ευρέως ανάλογα με τη φύση των δεδομένων και τα υποκείμενα φαινόμενα που αντιπροσωπεύει. Για παράδειγμα, σε πολλές χώρες, η ζήτηση ηλεκτρικής ενέργειας τείνει να είναι μεγαλύτερη κατά τους καλοκαιρινούς μήνες λόγω της αυξημένης χρήσης κλιματιστικών μονάδων, ενώ η ζήτηση κατά τους χειμερινούς μήνες τείνει να είναι μεγαλύτερη λόγω της αυξημένης χρήσης συστημάτων θέρμανσης. Αυτό το μοτίβο επαναλαμβάνεται χρόνο με το χρόνο και μπορεί να θεωρηθεί κύκλος στη χρονοσειρά της ζήτησης ενέργειας. Η ανίχνευση και η μοντελοποίηση αυτών των κύκλων σε χρονοσειρές ζήτησης ενέργειας είναι σημαντική για την πρόβλεψη της μελλοντικής ζήτησης και τη διασφάλιση ότι υπάρχει αρκετός ενεργειακός εφοδιασμός για την κάλυψη αυτής της ζήτησης. Μπορεί επίσης να βοηθήσει τις ενεργειακές εταιρείες και τους υπεύθυνους χάραξης πολιτικής στη λήψη αποφάσεων σχετικά με επενδύσεις σε υποδομές και στρατηγικές τιμολόγησης.
- Η **εποχικότητα** (Seasonal effect I_t) αντιπροσωπεύει τη βραχυχρόνια συμπεριφορά της χρονοσειράς και αποτελείται από περιοδικές ταλαντώσεις οι οποίες επαναλαμβάνονται με την ίδια περίοδο και σε επόμενα χρονικά διαστήματα. Οι τιμές της χρονοσειράς αλλάζουν συμπεριφορά σε διαστήματα μικρότερα του έτους π.χ. βδομάδα, μήνα. Το εποχικό χαρακτηριστικό μιας

χρονοσειράς αναφέρεται στα επαναλαμβανόμενα μοτίβα που παρατηρούνται μέσα σε κάθε ημερολογιακό έτος. Για τις χρονοσειρές ζήτησης ενέργειας, τα εποχιακά μοτίβα μπορούν να αποδοθούν σε αλλαγές του καιρού και σε άλλους παράγοντες που επηρεάζουν τη ζήτηση ενέργειας σε σταθερή βάση κάθε χρόνο. Ένα παράδειγμα εποχιακού προτύπου στη ζήτηση ενέργειας μπορεί να παρατηρηθεί στη ζήτηση ηλεκτρικής ενέργειας κατά τους χειμερινούς μήνες. Σε πολλές χώρες, η χειμερινή περίοδος φέρνει χαμηλότερες θερμοκρασίες, γεγονός που οδηγεί σε αυξημένη ζήτηση ενέργειας για σκοπούς θέρμανσης. Αυτό έχει ως αποτέλεσμα ένα εποχιακό μοτίβο όπου η ζήτηση ηλεκτρικής ενέργειας αυξάνεται κατά τους χειμερινούς μήνες και μειώνεται κατά τους καλοκαιρινούς μήνες. Για την ανάλυση και την πρόβλεψη της ζήτησης ενέργειας με εποχιακά χαρακτηριστικά, μπορούν να χρησιμοποιηθούν τεχνικές χρονοσειρών όπως εποχιακή αποσύνθεση, εποχιακά μοντέλα ARIMA ή μοντέλα εποχιακής εκθετικής εξομάλυνσης. Αυτές οι μέθοδοι επιτρέπουν τον διαχωρισμό της εποχικής συνιστώσας από άλλους παράγοντες όπως η τάση και ο τυχαίος θόρυβος στα δεδομένα. Αυτό επιτρέπει την καλύτερη αναγνώριση και μοντελοποίηση των εποχιακών προτύπων, τα οποία με τη σειρά τους μπορούν να οδηγήσουν σε ακριβέστερη πρόβλεψη της ζήτησης ενέργειας. Η κατανόηση και η μοντελοποίηση εποχιακών προτύπων στη ζήτηση ενέργειας μπορεί να είναι πολύτιμη για τις ενεργειακές εταιρείες και τους υπεύθυνους χάραξης πολιτικής στη λήψη τεκμηριωμένων αποφάσεων σχετικά με την κατανομή πόρων, τις στρατηγικές τιμολόγησης και τις επενδύσεις υποδομής, τα οποία μπορούν τελικά να οδηγήσουν σε πιο αποτελεσματική και βιώσιμη διαχείριση ενέργειας.

- Τα **κατάλοιπα** (Remainder E_t) αντιπροσωπεύουν τις βραχυχρόνιες διακυμάνσεις των τιμών της χρονοσειράς και οφείλονται σε τυχαίους παράγοντες. Συχνά, τα κατάλοιπα περιέχουν χρήσιμες πληροφορίες για την υπό εξέταση χρονοσειρά. Το υπολειπόμενο χαρακτηριστικό μιας χρονοσειράς είναι το τμήμα που παραμένει μετά την αφαίρεση των τάσεων, των εποχιακών και των κυκλικών στοιχείων από την αρχική χρονοσειρά. Αντιπροσωπεύει τις τυχαίες και απρόβλεπτες διακυμάνσεις στα δεδομένα. Το υπόλειμμα χρησιμοποιείται συχνά για την αξιολόγηση της καλής προσαρμογής ενός μοντέλου. Εάν το μοντέλο ταιριάζει καλά, τα υπολείμματα θα πρέπει να είναι τυχαία και να μην παρουσιάζουν σημαντικά μοτίβα ή τάσεις. Για παράδειγμα, ας εξετάσουμε μια χρονοσειρά της ημερήσιας ζήτησης ενέργειας σε μια πόλη. Μετά την εκτέλεση μιας εποχιακής αποσύνθεσης, προσδιορίσαμε την τάση, την εποχιακή και την κυκλική συνιστώσα της χρονοσειράς. Η υπολειπόμενη συνιστώσα αντιπροσωπεύει τις τυχαίες διακυμάνσεις στη ζήτηση ενέργειας που δεν μπορούν να εξηγηθούν από την τάση, τα εποχιακά ή κυκλικά μοτίβα. Εάν σχεδιάσουμε το υπολειπόμενο στοιχείο με την πάροδο του χρόνου, μπορεί να παρατηρήσουμε κάποιες τυχαίες διακυμάνσεις ή θόρυβο. Ωστόσο, εάν υπάρχει κάποιο σημαντικό μοτίβο ή τάση στο υπόλοιπο, υποδηλώνει ότι το μοντέλο δεν είναι σε θέση να συλλάβει όλη τη διακύμανση των δεδομένων. Σε αυτήν την

περίπτωση, ίσως χρειαστεί να εξετάσουμε ένα διαφορετικό μοντέλο ή να τροποποιήσουμε το υπάρχον για να βελτιώσουμε την εφαρμογή. Επομένως, η υπολειπόμενη συνιστώσα είναι μια ουσιαστική πτυχή της ανάλυσης χρονοσειρών, καθώς μας επιτρέπει να αξιολογήσουμε την επάρκεια του μοντέλου και να εντοπίσουμε τυχόν πιθανές πηγές σφάλματος.



2-Σχήμα 3.2: Trend, Cycle, Seasonality & Remainder of a Time Series

- **Αυτοσυσχέτιση** :Η εμφάνιση αυτοσυσχέτισης (autocorrelation) στις χρονοσειρές αναφέρεται συχνά ως γραμμική ή σειριακή αυτοσυσχέτιση (linear autocorrelation). Η αυτοσυσχέτιση ρ_k για κάποια υστέρηση k είναι ο συντελεστής συσχέτισης δύο στοιχείων της χρονοσειράς που έχουν μετατοπιστεί χρονικά k βήματα μεταξύ τους. Η συνάρτηση αυτοσυσχέτισης (ACF) μίας στοχαστικής διαδικασίας είναι η εξής:

$$\rho_k = \frac{\sum_{t=k+1}^n (Y_t - \mu)(Y_{t-k} - \mu)}{\sum_{t=1}^n (Y_t - \mu)^2}$$

Η σημασία της συνάρτησης αυτοσυσχέτισης είναι ιδιαίτερα σημαντική διότι δείχνει το βαθμό συσχέτισης μεταξύ των παρατηρήσεων, οι οποίες

απέχουν κατά χρονικό διάστημα k μεταξύ τους. Επιπλέον η εν λόγω συνάρτηση αντιπροσωπεύει το βαθμό (ένταση) και το μήκος (χρονική διάρκεια) της μνήμης της στοχαστικής διαδικασίας. Η έννοια της αυτοσυσχέτισης έχει νόημα όταν η χρονοσειρά είναι περιοδική. Αν επιχειρήσουμε να υπολογίζουμε τη συνάρτηση αυτοσυσχέτισης σε μία μη περιοδική χρονοσειρά με τάσεις, τότε θα παρατηρήσουμε ότι υπάρχουν πολύ υψηλές τιμές για μικρές μετατοπίσεις και ότι φθίνει πολύ αργά. Αυτό αποδεικνύει ότι υπάρχουν ισχυρές συσχετίσεις μεταξύ των κοντινών χρονικών σημείων, λόγω τάσης. Στις παραπάνω περιπτώσεις η πληροφορία για τις συσχετίσεις δεν περιέχεται σε κάποιες υστερήσεις αλλά ενδεχομένως να μας δοθούν από τη χρονοσειρά των υπολοίπων. Θεωρητικά, η αυτοσυσχέτιση της χρονοσειράς λευκού θορύβου είναι μηδενική για $s=0$, πρακτικά όμως η αυτοσυσχέτιση υπολογίζεται από μία πεπερασμένη χρονοσειρά και άρα οι διακυμάνσεις της κινούνται γύρω από το μηδέν. Η υπόθεση ότι μία περιοδική χρονοσειρά, απαλλαγμένη από τάσεις και περιοδικότητα, δεν περιέχει αυτοσυσχέτιση είναι πολύ σημαντική για την ανάλυση των χρονοσειρών. Η έλλειψη αυτοσυσχέτισης είναι μία σημαντική υπόθεση που πρέπει να αποδειχθεί για μία χρονοσειρά προκειμένου να προχωρήσει κάποιος στην ανάλυση της και στη δημιουργία μοντέλου με σκοπό την περαιτέρω ανάλυση ή τη διενέργεια προβλέψεων.

3.2.2 Ολοκληρωμένα Αυτοπαλινδρομικά Μοντέλα Κινητού Μέσου Όρου - ARIMA

Για την πρόβλεψη χρονοσειρών χρησιμοποιούνται ευρέως τα ARIMA [8] [9]. Πρόκειται για στοχαστικά μοντέλα τα οποία μας βοηθάνε να περιγράψουμε το μηχανισμό εξέλιξης ενός μεγέθους ενδιαφέροντος και ως εκ τούτου να προβλέψουμε την τιμή του στο μέλλον. Όπως και άλλα στατιστικά μοντέλα πρόβλεψης χρονοσειρών, π.χ. αυτά της εκθετικής εξομάλυνσης, τα μοντέλα ARIMA παράγουν προβλέψεις βασισμένα αποκλειστικά στις ιστορικές παρατηρήσεις του υπό εξέταση μεγέθους. Αυτό αποτελεί από μόνο του σημαντικό πλεονέκτημα, ειδικά σε περιπτώσεις που δεν γνωρίζουμε τις εξωτερικές μεταβλητές που το επηρεάζουν ή δεν μπορούμε να τις αξιοποιήσουμε άμεσα (βλ. ντετερμινιστικά μοντέλα). Στην πραγματικότητα, το κάθε μοντέλο ARIMA εκφράζει ένα διαφορετικό μηχανισμό εξέλιξης και η επιλογή του καταλληλότερου για την προέκταση μιας χρονοσειράς γίνεται εξετάζοντας παράγοντες όπως η σχέση μεταξύ k διαδοχικών παρατηρήσεων (αυτοσυσχέτιση), η ύπαρξη τάσης, η ύπαρξη εποχικότητας και το σφάλμα πρόβλεψης.

Στην γενική τους μορφή τα μοντέλα ARIMA αποτελούνται από τον τυχαίο παράγοντα (παράγοντας MA), τις τιμές που εμφανίστηκαν σε προηγούμενες περιόδους (παράγοντας AR και I), και άλλες στοχαστικές μεταβλητές. Πιο συγκεκριμένα, κάθε μοντέλο ARIMA μπορεί να εκφραστεί ως γραμμικός συνδυασμός των παραπάνω παραγόντων και στόχος μας είναι να ανακαλύψουμε εκείνον που παράγει τις καλύτερες προβλέψεις. Έτσι, αν το μοντέλο περιλαμβάνει αποκλειστικά παράγοντες αυτοπαλινδρόμησης αναφέρεται ως AR(p), αν περιλαμβάνει αποκλειστικά παράγοντες

κινητών μέσω όρων ως $MA(q)$, και αν περιλαμβάνει και τους δύο ως $ARMA(p,q)$, όπου τα p και q δηλώνουν την τάξη του μοντέλου ανά παράγοντα. Ο παράγοντας $I(d)$ αναφέρεται στη διαφόριση της χρονοσειράς πριν την εφαρμογή ενός μοντέλου $ARMA(p,q)$ και έχει ως στόχο την αφαίρεση της τάσης από τα δεδομένα. Σε περίπτωση που η χρονοσειρά είναι εποχιακή, τα μοντέλα $ARIMA$ μπορούν να επεκταθούν κατάλληλα προκειμένου να προσομοιώσουν και την εποχιακή συμπεριφορά των δεδομένων. Σε αυτή την περίπτωση η έκφρασή τους έχει τη μορφή $ARIMA(p,d,q)(P,D,Q)$, όπου τα P, D & Q αναφέρονται αντίστοιχα στην τάξη των εποχιακών παραγόντων $ARIMA$. Στην πράξη βέβαια δεν μπορούμε να είμαστε ποτέ σίγουροι για το ποιος είναι ο βέλτιστος συνδυασμός παραγόντων ή καλύτερα για το αν καταφέραμε να αποκαλύψουμε πλήρως το μοτίβο της χρονοσειράς. Μπορούμε ωστόσο να δίνουμε μία ικανοποιητική προσέγγιση στο εν λόγω ερώτημα εντοπίζοντας ένα μοντέλο που πληροί συγκεκριμένες προϋποθέσεις και παραμετροποιώντας το κατάλληλα.

Η εφαρμογή των μοντέλων $ARMA$ προϋποθέτει επίσης ότι πληρούνται ορισμένες απαιτήσεις. Αρχικά, η χρονοσειρά μήκους n πρέπει να είναι διακριτή, δηλαδή οι παρατηρήσεις της Y_t να αναφέρονται σε ισαπέχουσες χρονικές στιγμές $Y_t, Y_{t+T}, \dots, Y_{t+kT}$, όπου t ακέραιος μεγαλύτερος του μηδέν. Η παραπάνω απαίτηση είναι απαραίτητη καθώς, δεδομένου ότι τα μοντέλα συσχετίζουν χρονικά τις παρατηρήσεις της χρονοσειράς, η μη χρονική συνέχειά τους αναιρεί οποιαδήποτε υπόθεση έχει γίνει για τις σχέσεις διασύνδεσής τους. Εκτός αυτού, η χρονοσειρά οφείλει να είναι στάσιμη. Σε περίπτωση που η χρονοσειρά δεν είναι στάσιμη, κάτι τέτοιο μπορεί να επιτευχθεί με χρήση μετασχηματισμών ή/και διαφόρισης, με εφαρμογή δηλαδή κάποιου ολοκληρωμένου εποχιακού ή μη εποχιακού μοντέλου $ARIMA$ αντί για $ARMA$. Τέλος, η εφαρμογή των μοντέλων $ARIMA$ προϋποθέτει την παραγωγή βραχυπρόθεσμων προβλέψεων, καθώς, η αξιοπιστία και η ακρίβεια πρόβλεψής ενός μοντέλου $ARIMA$ αναμένεται να μειώνεται σημαντικά καθώς αυξάνει ο ορίζοντας πρόβλεψης.

Θέλοντας λοιπόν να αυτοματοποιήσουμε τη διαδικασία επιλογής, είθισται να ακολουθούμε την παρακάτω διαδικασία η οποία περιλαμβάνει τρία στάδια: την Αναγνώριση, την Εκτίμηση και τη Διάγνωση.

- **Αναγνώριση:** Στο στάδιο της αναγνώρισης επιλέγονται ένα ή περισσότερα μοντέλα $ARIMA$ τα οποία θεωρούμε βάσει κάποιων ενδείξεων, όπως π.χ. οι γραφικές παραστάσεις της αυτοσυσχέτισης και της μερικής αυτοσυσχέτισης, ότι μπορούν να περιγράψουν ικανοποιητικά τη χρονοσειρά.
- **Εκτίμηση:** Στο στάδιο της εκτίμησης εντοπίζονται για κάθε ένα από τα μοντέλα που αναγνωρίστηκαν οι παράμετροι p, d, q, P, D και Q , και βάσει αυτών παραμετροποιούνται κατάλληλα. Η εν λόγω διαδικασία μπορεί να γίνει με αρκετούς τρόπους, ο πιο διαδεδομένος εκ των οποίων είναι ο υπολογισμός της προσδοκώμενης πιθανοφάνειας. Η προσδοκώμενη πιθανοφάνεια δείχνει επί της ουσίας κατά πόσο το μοντέλο με τις παραμέτρους που επιλέχθηκαν έχει τη δυνατότητα να αναπαράγει τις πραγματικές τιμές της χρονοσειράς. Οι

παράμετροι υπολογίζονται λοιπόν με κριτήριο την μεγιστοποίηση της πιθανοφάνειας ή την ελαχιστοποίηση του λογάριθμου αυτής ως εξής :

$$L = \prod_{t=1}^n (2\pi\sigma^2)^{-\frac{1}{2}} e^{-\frac{(Y_t - F_t)^2}{2\sigma^2}} \rightarrow$$

$$L = (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{t=1}^n (Y_t - F_t)^2} \rightarrow$$

$$-2\log L = n \log(2\pi) + n \log(\sigma^2) + \frac{\sum_{t=1}^n e_t^2}{\sigma^2}$$

όπου \mathbf{L} η προσδοκώμενη πιθανοφάνεια ταύτισης του μοντέλου με τα αρχικά δεδομένα, F_t η προβλεπόμενη από το μοντέλο τιμή την περίοδο t , n ο αριθμός των ιστορικών δεδομένων, e_t το σφάλμα πρόβλεψης και η διακύμανση των σφαλμάτων του μοντέλου. Σημειώνεται πως η παραπάνω σχέση ισχύει αποκλειστικά για μοντέλα ARMA, δηλαδή δεν λαμβάνει υπόψη της τυχόν διαφορίσεις. Το κριτήριο της πιθανοφάνειας λειτουργεί πρακτικά όπως η μέθοδος ελαχίστων τετραγώνων στην περίπτωση της απλής γραμμικής παλινδρόμησης για την επιλογή των παραμέτρων α και β , το οποίο και ελαχιστοποιεί το άθροισμα των τετραγώνων των υπολειπόμενων σφαλμάτων μέσω της απαίτησης:

$$\min \left(\sum_{t=1}^n e_t^2 \right), \text{ όπου } e_t = Y_t - F_t$$

Μάλιστα, στην περίπτωση της γραμμικής παλινδρόμησης το κριτήριο της πιθανοφάνειας δίνει τις ίδιες παραμέτρους με αυτό των ελαχίστων τετραγώνων. Για την επιλογή των παραμέτρων συχνά συνδυάζεται η τεχνική των ελαχίστων τετραγώνων με αυτή της προσδοκώμενης πιθανοφάνειας για ακόμα καλύτερα αποτελέσματα. Εναλλακτικά, μπορεί κανείς να χρησιμοποιήσει άλλα κλασικά κριτήρια ελαχιστοποίησης σφαλμάτων (ME, MAPE, sMAPE κ.ο.κ.) όπως συμβαίνει π.χ. κατά τον υπολογισμό των παραμέτρων των μοντέλων εκθετικής εξομάλυνσης. Η πιθανοφάνεια αποτελεί ωστόσο όπως αναφέρθηκε την πιο διαδεδομένη βιβλιογραφικά αντικειμενική συνάρτηση βελτιστοποίησης παραμέτρων.

- **Διάγνωση:** Στο στάδιο του διαγνωστικού ελέγχου εφαρμόζονται στατιστικοί έλεγχοι προκειμένου να εξακριβωθεί αν τα μοντέλα που αναγνωρίστηκαν και εκτιμήθηκαν είναι προβλεπτικά άρτια. Ο διαγνωστικός έλεγχος γίνεται μελετώντας κυρίως την κατανομή των σφαλμάτων πρόβλεψης e_t των υποψήφιων μοντέλων. Αν το μοντέλο είναι άρτιο, τότε τα σφάλματα που αυτό παράγει θα πρέπει να οφείλονται αποκλειστικά στην τυχαιότητα της χρονοσειράς και συνεπώς να μην συσχετίζονται μεταξύ τους χρονικά.

Παρουσιάζουμε τα βασικά μοντέλα που συνθέτουν ένα ARIMA:

- **Autoregressive (AR) Model :**

Σε ένα αυτοπαλινδρομικό μοντέλο [9] , η μεταβλητή εξόδου y_t εξαρτάται γραμμικά από τις προηγούμενες τιμές της (y_{t-1}, \dots, y_{t-p}) και κάποιου λευκού θορύβου e_t . Εξορισμού ένα υπόδειγμα y_t λέγεται ότι είναι autoregressive p -τάξης και συμβολίζεται με AR(p), εάν το y_t περιγράφεται από την εξής σχέση:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p} + e_t$$

όπου:

e_t :λευκός θόρυβος (white noise) με μηδενικό μέσο (zero mean) και σταθερή διακύμανση

a_1, \dots, a_p : οι παράμετροι του υποδείγματος

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

όπου N το πλήθος των δεδομένων, X η εκάστοτε παρατήρηση και μ η μέση τιμή. Επιπρόσθετα, αν θεωρήσουμε ως τυχαία τα διαδοχικά στοιχεία μιας χρονοσειράς, τότε αυτή η χρονοσειρά λέγεται ότι αποτελείται από ανεξάρτητες τυχαίες μεταβλητές με ίδια κατανομή, για x_t, x_{t+1}, x_{t+t} για $\tau > 1$. Μια τέτοια χρονολογική σειρά είναι εντελώς τυχαία και δεν περιέχει αυτόσυσχετίσεις, γραμμικές ή μη-γραμμικές. Μια τέτοια χρονολογική σειρά, ονομάζεται και λευκός θόρυβος (white noise).

○ **Moving Average (MA) Model :**

Υποθέτοντας ότι t είναι μια καθαρά τυχαία διαδικασία με μηδενικό μέσο και διακύμανση σ^2 τότε ένα υπόδειγμα y_t λέγεται κινητού μέσου τάξης q MA(q) εάν το y_t περιγράφεται από την εξής σχέση:

$$y_t = e_t + \beta_1 e_{t-1} + \beta_2 e_{t-2} + \dots + \beta_p e_{t-p}$$

όπου:

e_t :λευκός θόρυβος (white noise)

β_1, \dots, β_p : οι παράμετροι του υποδείγματος

Στη διαδικασία κινητού μέσου το υπόδειγμα y_t θεωρείται ότι δημιουργείται ως ένας σταθμικός μέσος (weighted average) τυχαίων σφαλμάτων των q προηγούμενων περιόδων.

○ **ARMA:**

Το συγκεκριμένο μοντέλο είναι ένα από τα πιο ευρέως χρησιμοποιούμενα μοντέλα καθώς συνδυάζει τα πλεονεκτήματα των δυο προαναφερθέντων. Το συγκεκριμένο υπόδειγμα τάξης (p, q) ορίζεται ως εξής:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p} + e_t + \beta_1 e_{t-1} + \beta_2 e_{t-2} + \dots + \beta_p e_{t-p}$$

όπου

e_t : λευκός θόρυβος (white noise)

a_1, \dots, a_p : οι παράμετροι του υποδείγματος

β_1, \dots, β_p : οι παράμετροι του υποδείγματος για τον MA

○ **Auto Regressive Integrated Moving Average (ARIMA) Model:**

Τα μοντέλα που παρουσιάστηκαν προηγουμένως, AR – AM – ARMA, χρησιμοποιούνται σε στάσιμες διαδικασίες. Δηλαδή, ο μέσος και η διακύμανση δεν εξαρτώνται από τον χρόνο t αλλά παραμένουν σταθερά. Στην πράξη οι περισσότερες χρονοσειρές είναι μη-στάσιμες οπότε για να εφαρμοστεί ένα στάσιμο υπόδειγμα θα πρέπει να αφαιρεθούν τα μη-στάσιμα χαρακτηριστικά. Μια λύση σε αυτό παρουσιάστηκε με το μοντέλο ARIMA τάξεως (p,d,q) . Η ικανότητα του ARIMA μοντέλου να ανταπεξέρχεται σε μη-στάσιμες διαδικασίες, το καθιστά μια από τις πιο χρησιμοποιούμενες προσεγγίσεις στη πρόβλεψη χρονοσειρών. Υπερνικά αυτόν τον περιορισμό με τη χρήση διαφορών και αυτό επιτυγχάνεται αφαιρώντας την παρατήρηση της τρέχουσας περιόδου από την προηγούμενη. Η διαφορίση αυτή μπορεί να είναι διαφόρων τάξεων. Για παράδειγμα η πρώτη τάξης διαφορίση πραγματοποιείται αντικαθιστώντας το y_t με την εξής σχέση:

$$y'_t = y_t - y_{t-1}$$

Η γενική μορφή του υποδείγματος ARIMA (p,d,q) περιγράφεται από τη παρακάτω σχέση:

$$y'_t = a_1 y'_{t-1} + a_2 y'_{t-2} + \dots + a_p y'_{t-p} + e_t + \beta_1 e_{t-1} + \beta_2 e_{t-2} + \dots + \beta_p e_{t-p}$$

Όπου,

p : οι παράμετροι της αυτοπαλίνδρομης διαδικασίας

d : ο αριθμός των διαφορών προκειμένου η χρονοσειρά να γίνει στάσιμη

q : οι παράμετροι της διαδικασίας του κινητού μέσου

○ **Seasonal ARIMA (SARIMA) Model:**

Το μοντέλο SARIMA είναι μια επέκταση του ARIMA και χρησιμοποιείται για δεδομένα όπου παρουσιάζουν περιοδικότητα. Το SARIMA διατυπώνεται ως SARIMA $(p,d,q)(P,D,Q)m$ όπου P το πλήθος των εποχικών autoregressive όρων, D το πλήθος των εποχικών διαφορών και Q το πλήθος των εποχικών όρων κινητού μέσου, m ο αριθμός των χρονικών βημάτων στη χρονική περίοδο μια εποχής.

○ **Seasonal ARIMA (SARIMAX) Model with exogenous variables:**

Ο εποχιακός αυτοπαλινδρομικός ολοκληρωμένος κινούμενος μέσος με εξωγενείς μεταβλητές (SARIMAX) είναι μια επέκταση του παραδοσιακού μοντέλου ARIMA που λαμβάνει υπόψη τα εποχιακά πρότυπα και ενσωματώνει εξωτερικές μεταβλητές για τη βελτίωση της ακρίβειας της πρόβλεψης χρονοσειρών. Όπως το ARIMA, το SARIMAX μοντελοποιεί τις μελλοντικές τιμές μιας χρονοσειράς με βάση τις προηγούμενες τιμές της, αλλά περιλαμβάνει επίσης εποχιακά στοιχεία που μπορούν να συλλάβουν περιοδικά μοτίβα στα δεδομένα. Επιπλέον, επιτρέπει τη συμπερίληψη εξωγενών μεταβλητών, οι οποίες είναι εξωτερικοί παράγοντες που μπορεί να επηρεάσουν τις χρονοσειρές που προβλέπονται. Τα μοντέλα SARIMAX χαρακτηρίζονται από τέσσερις παραμέτρους: $(p, d, q) \times (P, D, Q, m)$, όπου p, d και q είναι οι παράμετροι για τα μη εποχικά στοιχεία του μοντέλου, P, D , και Q είναι οι παράμετροι για τα εποχιακά στοιχεία και m είναι ο αριθμός των περιόδων σε κάθε εποχή. Οι μη εποχικές συνιστώσες είναι παρόμοιες με αυτές του ARIMA, ενώ οι εποχιακές συνιστώσες αποτελούνται από αυτοπαλινδρομικούς και κινούμενους μέσους όρους που περιγράφουν τα εποχιακά μοτίβα στα δεδομένα. Η συμπερίληψη εξωγενών μεταβλητών στα μοντέλα SARIMAX μπορεί να βελτιώσει την ακρίβεια πρόβλεψής τους, καθώς αυτές οι μεταβλητές μπορεί να παρέχουν πρόσθετες πληροφορίες που μπορούν να βοηθήσουν στην εξήγηση της συμπεριφοράς των χρονοσειρών που προβλέπονται. Παραδείγματα εξωγενών μεταβλητών περιλαμβάνουν οικονομικούς δείκτες, δεδομένα καιρού ή δημογραφικές πληροφορίες. Τα μοντέλα SARIMAX έχουν ένα ευρύ φάσμα εφαρμογών, συμπεριλαμβανομένων προβλέψεων πωλήσεων, ζήτησης προϊόντων ή κατανάλωσης ενέργειας. Είναι ιδιαίτερα χρήσιμα όταν ασχολούμαστε με δεδομένα χρονοσειρών που εμφανίζουν εποχιακά μοτίβα και επηρεάζονται από εξωτερικούς παράγοντες. Συνοπτικά, το SARIMAX είναι ένα ισχυρό μοντέλο πρόβλεψης χρονοσειρών που μπορεί να καταγράψει τόσο εποχιακά μοτίβα όσο και εξωτερικούς παράγοντες για τη βελτίωση της ακρίβειας. Η ικανότητά του να ενσωματώνει εξωγενείς μεταβλητές το καθιστά πολύτιμο εργαλείο για τις προβλέψεις σε ένα ευρύ φάσμα τομέων, από τη χρηματοδότηση και την οικονομία μέχρι την ενέργεια και το λιανικό εμπόριο.

3.3 Μηχανική Μάθηση(Machine Learning) - Τεχνητά Νευρωνικά Δίκτυα(Artificial Neural Networks)

3.3.1 Εισαγωγή

Ο Arthur Samuel ορίζει ως μηχανική μάθηση «Το πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί» [10]. Η μηχανική μάθηση επικεντρώνεται στην ανάπτυξη αλγορίθμων που μπορούν να έχουν πρόσβαση σε δεδομένα και να τα χρησιμοποιούν για να μάθουν και να βελτιώνονται. Η μηχανική μάθηση χωρίζεται σε τρεις κύριες κατηγορίες [11]:

- **Επιβλεπόμενη Μάθηση (Supervised Learning).** Ο αλγόριθμος εκπαιδεύεται σε ένα σύνολο παραδειγμάτων με ζευγάρια εισόδων και επιθυμητών εξόδων. Μερικοί από τους αλγορίθμους που χρησιμοποιούνται είναι: Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines - SVM), Γραμμική Παλινδρόμηση (Linear Regression), Λογιστική Παλινδρόμηση (Logistic Regression), Δέντρα Αποφάσεων (Decision Trees), Νευρωνικά Δίκτυα (Neural Networks).

Η εποπτευόμενη μάθηση είναι ένας τύπος μηχανικής μάθησης στην οποία ένας αλγόριθμος εκπαιδεύεται σε δεδομένα με ετικέτα, που σημαίνει ότι τα δεδομένα εισόδου συνοδεύονται από γνωστές τιμές εξόδου. Ο στόχος είναι να μάθουμε μια συνάρτηση χαρτογράφησης από την είσοδο στην έξοδο, η οποία μπορεί στη συνέχεια να χρησιμοποιηθεί για να κάνει προβλέψεις σε νέα, αόρατα δεδομένα. Υπάρχουν πολλά δημοφιλή μοντέλα που χρησιμοποιούνται στην εποπτευόμενη μάθηση, όπως:

Γραμμική παλινδρόμηση: Ένα απλό αλλά ισχυρό μοντέλο που χρησιμοποιείται για την πρόβλεψη συνεχών τιμών εξόδου. Υποθέτει μια γραμμική σχέση μεταξύ των μεταβλητών εισόδου και εξόδου.

Λογιστική παλινδρόμηση: Ένας τύπος παλινδρόμησης που χρησιμοποιείται για προβλήματα δυαδικής ταξινόμησης, όπου η μεταβλητή εξόδου παίρνει μόνο δύο πιθανές τιμές.

Δέντρα αποφάσεων: Ένα μη παραμετρικό μοντέλο που χρησιμοποιεί μια δομή που μοιάζει με δέντρο για να αναπαραστήσει αποφάσεις και τις πιθανές συνέπειές τους.

Τυχαία δάση: Ένας τύπος μοντέλου εκμάθησης συνόλου που συνδυάζει πολλαπλά δέντρα αποφάσεων για να βελτιώσει την ακρίβεια και την ευρωστία των προβλέψεων.

Μηχανές Διανυσμάτων Υποστήριξης (SVMs): Ένα μοντέλο που προσπαθεί να βρει ένα υπερεπίπεδο που διαχωρίζει τα δεδομένα εισόδου σε διαφορετικές κλάσεις.

Νευρωνικά δίκτυα: Μια ισχυρή κατηγορία μοντέλων εμπνευσμένων από τη δομή και τη λειτουργία του ανθρώπινου εγκεφάλου. Μπορούν να χρησιμοποιηθούν τόσο για εργασίες παλινδρόμησης όσο και για εργασίες ταξινόμησης και είναι ιδιαίτερα χρήσιμα όταν ασχολούμαστε με πολύπλοκα, υψηλών διαστάσεων δεδομένα.

Αυτά τα μοντέλα μπορούν να χρησιμοποιηθούν για ένα ευρύ φάσμα εποπτευόμενων εργασιών μάθησης, συμπεριλαμβανομένης της παλινδρόμησης, της ταξινόμησης και της πρόβλεψης. Η επιλογή του μοντέλου εξαρτάται από το συγκεκριμένο πρόβλημα και τα χαρακτηριστικά των δεδομένων που αναλύονται. Γενικά, τα πιο απλά μοντέλα όπως η γραμμική παλινδρόμηση και η λογιστική παλινδρόμηση προτιμώνται όταν τα δεδομένα είναι χαμηλών διαστάσεων και η σχέση μεταξύ των μεταβλητών εισόδου και εξόδου είναι καλά κατανοητή. Πιο πολύπλοκα μοντέλα όπως τα νευρωνικά δίκτυα είναι καλύτερα κατάλληλα για δεδομένα υψηλών διαστάσεων με πολύπλοκα μοτίβα και σχέσεις.

- **Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning).** Ο αλγόριθμος εκπαιδεύεται χρησιμοποιώντας πληροφορίες που δεν είναι ούτε ταξινομημένες αλλά ούτε επισημασμένες και επιτρέπουν το αλγόριθμο να ενεργεί χωρίς καθοδήγηση. Αλγόριθμοι για αυτή τη κατηγορία είναι: Συσταδοποίηση k-means, Ιεραρχική Συσταδοποίηση (Hierarchical Clustering), Ανάλυση Κύριων Συνιστωσών (Principle Component Analysis - PCA).

Η μάθηση χωρίς επίβλεψη είναι ένας τύπος μηχανικής μάθησης στην οποία ο στόχος είναι η εύρεση μοτίβων ή σχέσεων στα δεδομένα χωρίς τη χρήση επισημασμένης εξόδου. Τα δεδομένα εισόδου δεν φέρουν ετικέτα και ο αλγόριθμος πρέπει να μάθει μόνος του την υποκείμενη δομή ή τις σχέσεις μεταξύ των μεταβλητών. Η μάθηση χωρίς επίβλεψη χρησιμοποιείται συνήθως για ομαδοποίηση, μείωση διαστάσεων και ανίχνευση ανωμαλιών. Υπάρχουν πολλά δημοφιλή μοντέλα που χρησιμοποιούνται στη μάθηση χωρίς επίβλεψη, όπως:

Ομαδοποίηση K-means: Ένας απλός αλλά αποτελεσματικός αλγόριθμος για την κατάτμηση ενός συνόλου δεδομένων σε K διακριτές συστάδες με βάση την ομοιότητα. Κάθε σύμπλεγμα αντιπροσωπεύεται από το κέντρο του και ο στόχος είναι να ελαχιστοποιηθεί η απόσταση μεταξύ κάθε σημείου δεδομένων και του αντίστοιχου κέντρου συστάδας.

Ανάλυση κύριου συστατικού (PCA): Μια τεχνική για τη μείωση της διάστασης ενός συνόλου δεδομένων προβάλλοντας τα δεδομένα σε χώρο χαμηλότερης διάστασης διατηρώντας όσο το δυνατόν μεγαλύτερη διακύμανση.

Autoencoders: Ένας τύπος νευρωνικού δικτύου που μαθαίνει να κωδικοποιεί τα δεδομένα εισόδου σε μια συμπιεσμένη αναπαράσταση και στη συνέχεια να τα αποκωδικοποιεί ξανά στην αρχική τους μορφή. Οι αυτόματες κωδικοποιητές είναι χρήσιμοι για μείωση διαστάσεων, ανίχνευση ανωμαλιών και εξαγωγή χαρακτηριστικών.

Μοντέλα μειγμάτων Gaussian (GMMs): Ένα πιθανοτικό μοντέλο που αναπαριστά τα δεδομένα με ένα μείγμα κανονικών κατανομών. Τα GMM μπορούν να χρησιμοποιηθούν για ομαδοποίηση, εκτίμηση πυκνότητας και ανίχνευση ανωμαλιών.

Ιεραρχική ομαδοποίηση: Μια μέθοδος για την ομαδοποίηση δεδομένων σε μια ιεραρχία ένθετων συστάδων. Υπάρχουν δύο κύριοι τύποι ιεραρχικής ομαδοποίησης: η αθροιστική, η οποία ξεκινά με κάθε σημείο δεδομένων ως τη δική της συστάδα και στη συνέχεια συγχωνεύει επαναληπτικά τις πλησιέστερες συστάδες και η διαιρετική, η οποία ξεκινά με όλα τα σημεία δεδομένων σε μια ενιαία συστάδα και στη συνέχεια τα διαιρεί αναδρομικά σε μικρότερα συμπλέγματα .

Αυτά τα μοντέλα μπορούν να χρησιμοποιηθούν για μια ποικιλία εργασιών μάθησης χωρίς επίβλεψη, όπως η ομαδοποίηση παρόμοιων σημείων δεδομένων, η ανίχνευση ακραίων σημείων ή ανωμαλιών στα δεδομένα και η μείωση της διάστασης των συνόλων δεδομένων υψηλών διαστάσεων. Η επιλογή του μοντέλου εξαρτάται από το συγκεκριμένο πρόβλημα και τα χαρακτηριστικά των δεδομένων που αναλύονται. Γενικά, τα πιο απλά μοντέλα όπως η ομαδοποίηση **k-means** ή το PCA προτιμώνται όταν τα δεδομένα είναι χαμηλών διαστάσεων και η υποκείμενη δομή είναι καλά κατανοητή, ενώ πιο πολύπλοκα μοντέλα όπως οι αυτόματες κωδικοποιητές ή τα GMM είναι καλύτερα κατάλληλα για δεδομένα υψηλών διαστάσεων με πολύπλοκες σχέσεις.

- **Ενισχυτική Μάθηση (Reinforcement Learning).** Η ενισχυτική μάθηση (RL) είναι ένα υποπεδίο της μηχανικής μάθησης που εστιάζει στη διδασκαλία των πρακτόρων πώς να λαμβάνουν αποφάσεις σε ένα περιβάλλον λαμβάνοντας ανατροφοδότηση με τη μορφή ανταμοιβών ή τιμωριών. Ο στόχος της ενισχυτικής μάθησης είναι να μεγιστοποιήσει τη σωρευτική ανταμοιβή με την πάροδο του χρόνου, μαθαίνοντας μια πολιτική που αντιστοιχίζει καταστάσεις σε ενέργειες. Τα πιο δημοφιλή μοντέλα που χρησιμοποιούνται στην ενισχυτική μάθηση είναι τα Q-learning, SARSA και Deep Q-Networks.

Το Q-learning είναι ένας αλγόριθμος RL χωρίς μοντέλα που χρησιμοποιεί έναν πίνακα για να αποθηκεύσει την αναμενόμενη ανταμοιβή για κάθε ζεύγος κατάστασης-ενέργειας. Ο αλγόριθμος ενημερώνει αυτόν τον πίνακα μαθαίνοντας από τις ανταμοιβές που λαμβάνει από το περιβάλλον. Η Q-learning είναι αποτελεσματική σε προβλήματα μικρού έως μεσαίου μεγέθους όπου ο αριθμός των καταστάσεων και των ενεργειών είναι περιορισμένος.

Το SARSA είναι ένας άλλος αλγόριθμος RL χωρίς μοντέλα που μαθαίνει μια βέλτιστη πολιτική ενημερώνοντας τις εκτιμήσεις του με βάση την τρέχουσα πολιτική. Σε αντίθεση με το Q-learning, το SARSA λαμβάνει υπόψη την τρέχουσα κατάσταση, την τρέχουσα ενέργεια, την ανταμοιβή και την επόμενη κατάσταση, την επόμενη ενέργεια κατά την ενημέρωση των εκτιμήσεών του. Το SARSA είναι αποτελεσματικό σε προβλήματα όπου η πολιτική είναι στοχαστική ή όταν το περιβάλλον είναι θορυβώδες.

Τα Deep Q-Networks (DQN) είναι ένας τύπος αλγορίθμου RL που χρησιμοποιεί βαθιά νευρωνικά δίκτυα για να προσεγγίσει τη συνάρτηση Q-value. Τα DQN έχουν επιτύχει στην επίλυση προβλημάτων με χώρους καταστάσεων υψηλών διαστάσεων, όπως η αναγνώριση εικόνας και τα βιντεοπαιχνίδια. Ο αλγόριθμος χρησιμοποιεί επανάληψη εμπειρίας και στοχεύει δίκτυα για να βελτιώσει τη σταθερότητα της μαθησιακής διαδικασίας.

Εκτός από αυτά τα μοντέλα, υπάρχουν και άλλοι δημοφιλείς αλγόριθμοι RL, όπως μέθοδοι διαβάθμισης πολιτικής, μέθοδοι κριτικής ηθοποιών και αναζήτηση δέντρου Monte Carlo. Κάθε αλγόριθμος έχει τα δυνατά και τα αδύνατα σημεία του και είναι κατάλληλος για διαφορετικούς τύπους προβλημάτων. Η ενισχυτική μάθηση έχει πολλές πρακτικές εφαρμογές, συμπεριλαμβανομένης της ρομποτικής, του παιχνιδιού και της διαχείρισης πόρων.

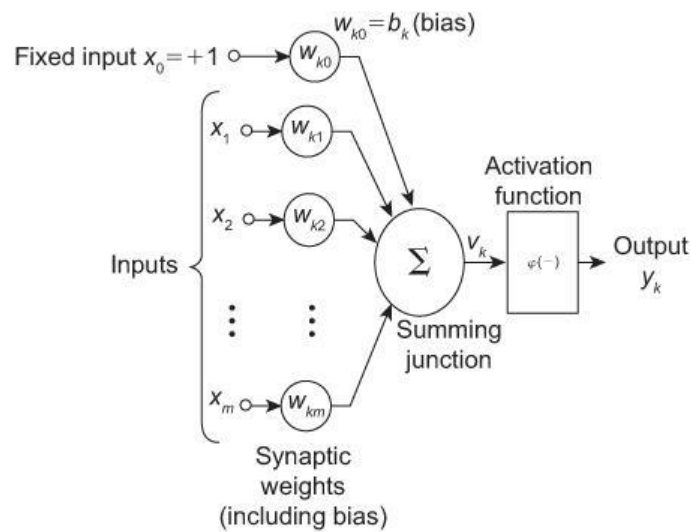
3.3.2 Νευρώνας - Perceptron

Η βασική μονάδα υπολογισμού σε ένα ANN είναι ο νευρώνας (Neuron) , συχνά ονομαζόμενος και ως κόμβος. Το μοντέλο του Perceptron, περιγράφηκε από τον Rosenblatt [12]. Λαμβάνει εισόδους από άλλους κόμβους ή από μια εξωτερική πηγή και υπολογίζει μια έξοδο. Κάθε είσοδος πολλαπλασιάζεται με το αντίστοιχο βάρος (Weight) και υπολογίζεται το ολικό άθροισμα των γινομένων. Ο κόμβος εφαρμόζει μια συνάρτηση ενεργοποίησης σε αυτό το άθροισμα και υπολογίζεται η έξοδος του νευρώνα. Παρακάτω φαίνεται η αναπαράσταση ενός νευρώνα, καθώς και η εξίσωση της εξόδου.

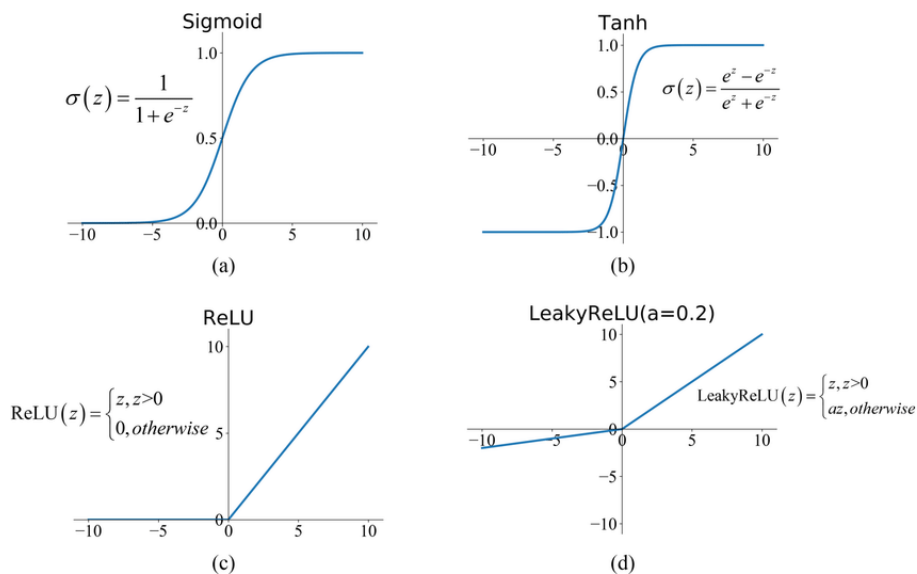
$$k = f((\sum_{i=1}^n w_i * x_i) + b_i) \quad (3.3)$$

Το παραπάνω δίκτυο παίρνει ως εισόδους τα $X_1, X_2 \dots X_n$ που έχουν για βάρη τα $W_1, W_2 \dots W_n$ αντίστοιχα. Επιπλέον υπάρχει ακόμα μια είσοδος $\mathbf{1}$ με βάρος \mathbf{b} η οποία ονομάζεται πόλωση (bias). Η συνάρτηση \mathbf{f} είναι μη γραμμική και ονομάζεται

συνάρτηση ενεργοποίησης (activation function). Ο σκοπός της συνάρτησης ενεργοποίησης είναι να εισάγει μη γραμμικότητα στην έξοδο ενός νευρώνα. Αυτό είναι σημαντικό καθώς σχεδόν όλα τα πραγματικά δεδομένα είναι μη γραμμικά. Παραδείγματα τέτοιων συναρτήσεων είναι η σιγμοειδής συνάρτηση, η υπερβολική εφαπτομένη και η ReLU. Η ReLU αποδίδει συχνά καλύτερα από άλλες συναρτήσεις ενεργοποίησης για κρυφά επίπεδα. Ο βασικός λόγος της αυξημένης απόδοσης οφείλεται στο γεγονός ότι η ReLU είναι μια γραμμική συνάρτηση μη κορεσμού. Ο κορεσμός είναι το μεγαλύτερο πρόβλημα των δυο προηγούμενων σιγμοειδών συναρτήσεων. Σε αντίθεση λοιπόν με την **logistic ή tanh**, η ReLU δεν κορέζεται στο **-1, 0 ή 1**. Στην πράξη η συχνότερα εμφανιζόμενη συνάρτηση ενεργοποίησης είναι η ReLU.



3-Σχήμα 3.3: Artificial Neuron model and its parts



4-Σχήμα 3.4: Activation functions

Η χρήση της ReLU έχει επικρατήσει διότι έχει χαρακτηριστικά που διευκολύνουν τη μάθηση μέσω μεθόδων gradient descent [11]. Συγκεκριμένα, το γεγονός ότι στο ήμισυ του πεδίου ορισμού της, η τιμή της ReLU είναι μηδέν, διευκολύνει την ύπαρξη μεγαλύτερων τιμών κλίσης (gradient) της εξόδου ενός νευρώνα ως προς τα βάρη των συνάψεών του. Αντίθετα, οι σιγμοειδείς συναρτήσεις ενεργοποίησης παρουσιάζουν κορεσμό είτε σε μία θετική είτε σε μία αρνητική τιμή για το μεγαλύτερο μέρος του πεδίου ορισμού τους, γεγονός που οδηγεί σε πολύ μικρές τιμές κλίσης, καθιστώντας τη μάθηση αργή ή ακόμα και αδύνατη. Η συνάρτηση υπερβολικής εφαιπτομένης πολλές φορές εμφανίζεται σαν συνάρτηση ενεργοποίησης σε νευρώνες αναδρομικών νευρωνικών μοντέλων όπως Long-Term Short-Term νευρώνες.

3.3.3 Multilayer Perceptron

Σύνδεση μεταξύ δύο νευρώνων υφίσταται, όταν η έξοδος του ενός χρησιμοποιείται σαν είσοδος του άλλου. Η σύνδεση πολλών νευρώνων μεταξύ τους μπορεί να δημιουργήσει μια συνάρτηση που να έχει τη δυνατότητα να λύσει μη γραμμικά διαχωρίσιμα προβλήματα. Μια από τις συνηθισμένες κατασκευές είναι το MLP. Στην τοπολογία αυτού του τύπου, οι νευρώνες διαμερίζονται σε επίπεδα (layers), και τα γειτονικά επίπεδα συνδέονται με συνδέσεις μονής κατεύθυνσης, χωρίς αναδρομή της πληροφορίας. Ουσιαστικά, κάθε νευρώνας του \mathbf{i} επιπέδου δέχεται σαν είσοδο τις εξόδους από κάθε νευρώνα του $\mathbf{i} - 1$ επιπέδου. Η έξοδος του τροφοδοτείται στους νευρώνες του επόμενου επιπέδου, ή στην έξοδο του δικτύου αν ο νευρώνας ανήκει στο τελευταίο κρυφό επίπεδο.

Για ένα δίκτυο με \mathbf{m} αριθμό επιπέδων, για την έξοδο του \mathbf{i} -οστού επιπέδου, όπου: $\mathbf{0} < \mathbf{i} < \mathbf{m}$, ισχύει:

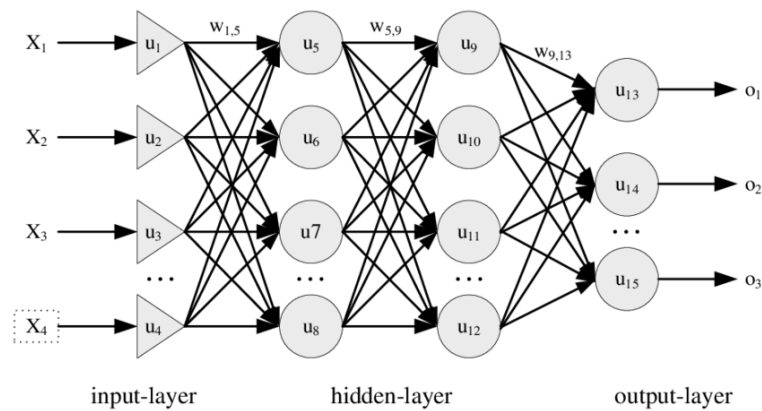
$$h^{(i)} = g^{(i)}(W^{(i)}h^{(i-1)} + b^{(i)}) = g^{(i)}(h^{(i-1)}; j^{(i)})$$

όπου: $h^{(i)}$ η έξοδος του επιπέδου, $g^{(i)}$ η συνάρτηση ενεργοποίησης των νευρώνων του επιπέδου, $W^{(i)}$ ο πίνακας των βαρών των συνδέσεων και $b^{(i)}$ το διάνυσμα της προκατάληψης.

Όμοια,

$$h^{(0)} = g^{(0)}(x; \theta^{(0)}) \text{ και } h^{(m-1)} = g^{(0)}(h^{(m-2)}; \theta^{(m-1)})$$

Τα επίπεδα $\mathbf{0} < \mathbf{i} < \mathbf{m} - 1$ ονομάζονται κρυφά επίπεδα (hidden layers) αφού η συμβολή τους στο δίκτυο δεν είναι εμφανής εκτός του μοντέλου. Τα επίπεδα $\mathbf{i} = \mathbf{0}$ και $\mathbf{i} = \mathbf{m} - 1$ ονομάζονται επίπεδο εισόδου και εξόδου, αντίστοιχα. Γενικά, ένα δίκτυο ονομάζεται βαθύ (deep) αν $\mathbf{m} > \mathbf{3}$, δηλαδή έχει τουλάχιστον ένα κρυφό επίπεδο.

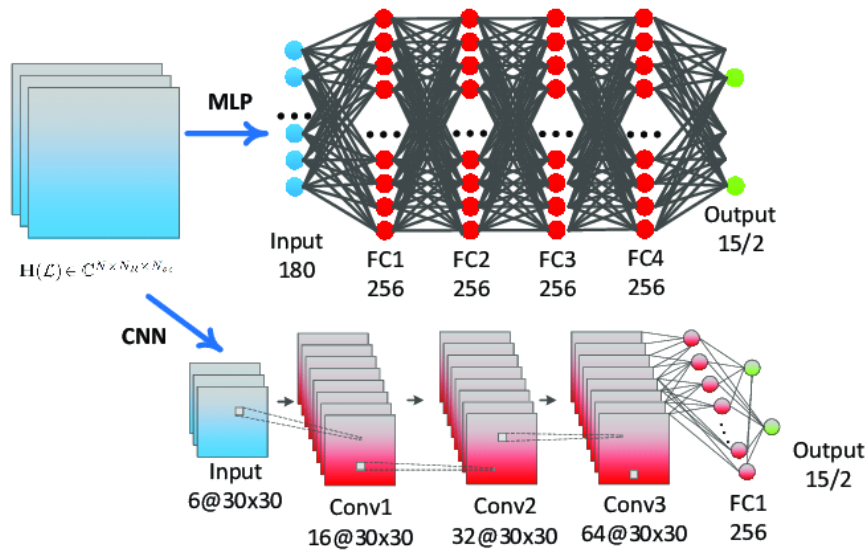


5-Σχήμα 3.5: Three layer multilayer perceptron (MLP) neural network

3.3.4 Συνελκτικό Νευρωνικό Δίκτυο - Convolutional Neural Networks

Στη βαθιά μηχανική μάθηση, τα Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks, CNN) είναι μια κατηγορία νευρωνικών δικτύων που χρησιμοποιούνται συνήθως στον κλάδο της ανάλυσης εικόνας [13]. Τα CNN είναι κανονικοποιημένες εκδόσεις των MLP. Τα MLP είναι πλήρως συνδεδεμένα δίκτυα. Αυτή η συνδεσιμότητα τα καθιστά επιρρεπή σε υπερεκπαίδευση. Τα CNNs αξιοποιούν τα ιεραρχικά μοτίβα που υπάρχουν στα δεδομένα και κατασκευάζουν περίπλοκα πρότυπα χρησιμοποιώντας μικρότερα και απλούστερα. Αυτός ο τύπος δικτύου είναι εμπνευσμένος από τον οπτικό φλοιό (visual cortex) του εγκεφάλου.

Η λέξη "συνελκτικό" στην ονομασία του δικτύου υποδεικνύει ότι αξιοποιεί τη μαθηματική πράξη της συνέλιξης πινάκων (element-wise matrix product ή Hadamard product). Η συνέλιξη είναι ένας ειδικός τύπος γραμμικού υπολογισμού. Τα CNN χρησιμοποιούν τη συνέλιξη στη θέση του πολλαπλασιασμού πινάκων, σε τουλάχιστον ένα από τα επίπεδά τους. Τα κρυφά επίπεδα ενός CNN, συνήθως, αποτελούνται από μια σειρά από συνελκτικά επίπεδα (convolutional layers) που πραγματοποιούν συνέλιξεις. Συνήθως, η συνάρτηση ενεργοποίησης αυτών των επιπέδων είναι ένα ReLu layer που ακολουθείται από επιπλέον επίπεδα, που ονομάζονται επίπεδα υπερδειγματοληψίας (pooling layers).

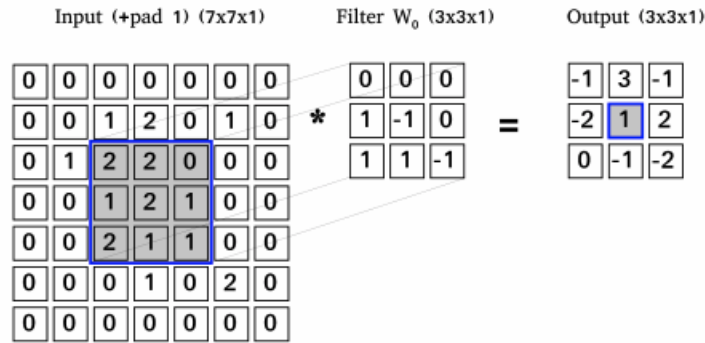


6-Σχήμα 3.6: CNN & MLP Architecture

Κατά τον σχεδιασμό ενός CNN, κάθε convolutional layer πρέπει να έχει τα παρακάτω γνωρίσματα:

- Η είσοδος του πρέπει να είναι ένα διάνυσμα ή ένας πίνακας με διαστάσεις (αριθμός εικόνων) x (πλάτος εικόνας) x (ύψος εικόνας) x (αριθμός χρωμάτων (image depth))
- Συνελκτικούς πυρήνες (convolutional kernels) το πλάτος και το ύψος των οποίων είναι υπερπαράμετροι και το βάθος τους ίσο ή μικρότερο με το βάθος των εικόνων που επεξεργάζεται.

Τα συνελκτικά επίπεδα διασχίζουν την είσοδο σε μέρη και περνάνε το αποτέλεσμα τους στο επόμενο επίπεδο. Αυτή η λειτουργία είναι παρόμοια με τον τρόπο που ανταποκρίνονται οι νευρώνες στον οπτικό φλοιό για ένα συγκεκριμένο ερέθισμα. Κάθε συνελκτικός νευρώνας επεξεργάζεται δεδομένα που βρίσκονται αποκλειστικά στο δεκτικό του πεδίο (receptive field). Αν και για τις ίδιες λειτουργίες θα μπορούσε να χρησιμοποιηθεί ένα πλήρες συνελκτικό MLP, δεν είναι πρακτικό να χρησιμοποιηθεί αυτή η αρχιτεκτονική για την επεξεργασία εικόνων. Θα έπρεπε να χρησιμοποιηθεί μεγάλο πλήθος νευρώνων, ακόμα και σε αρχιτεκτονικές λίγων επιπέδων, εξαιτίας του μεγάλου μεγέθους της εισόδου.



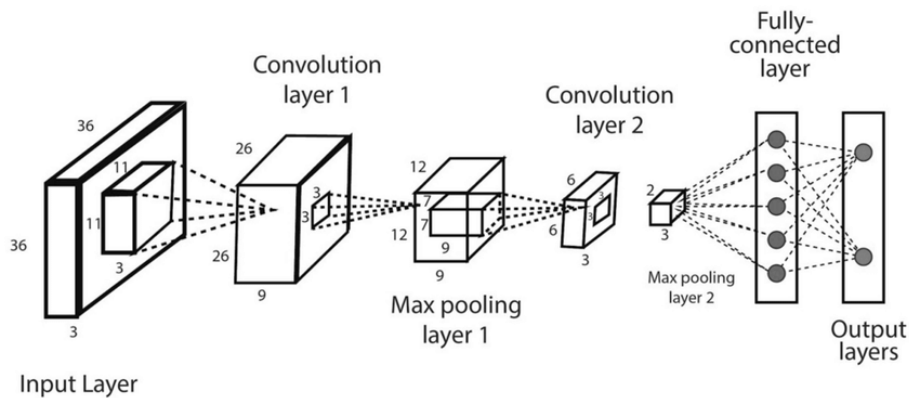
7-Σχήμα 3.7: Convolutional Layer of a CNN

Για παράδειγμα, ένα MLP, για μια μικρή εικόνα διαστάσεων 100 x 100, έχει 10000 παραμέτρους για κάθε νευρώνα του δευτέρου επιπέδου. Η συνέλιξη δίνει λύσει σε αυτό το πρόβλημα, αφού μειώνει τον αριθμό των εκπαιδευσιμων παραμέτρων, επιτρέποντας στο δίκτυο να έχει περισσότερα επίπεδα και λιγότερες παραμέτρους. Παραδείγματος χάριν, ανεξάρτητα από το μέγεθος της εικόνας, η χρήση πυρήνων 5 x 5, που μοιράζονται τα ίδια βάρη, χρειάζεται μόνο 25 εκπαιδευσιμες παραμέτρους.

Ειδικότερα, στη περίπτωση της δισδιάστατης συνέλιξης, αν η είσοδος είναι μία εικόνα ή ένας τρισδιάστατος τανυστής, η μαθηματική περιγραφή της λειτουργίας του συνελκτικού στρώματός για κάθε φίλτρο είναι:

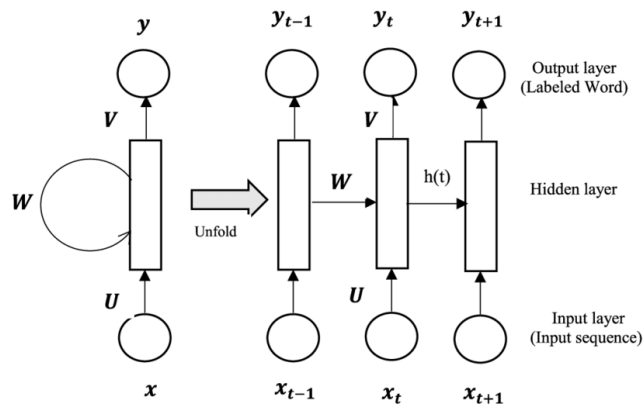
$$S_l[i, j] = \text{ReLu}(b_l + \sum_{m=0}^{M_l-1} \sum_{n=0}^{N_l-1} W_l[m, n, :] oI[i - m, j - n, :]) \quad (3.4)$$

όπου l ο δείκτης που αντιστοιχεί στο φίλτρο με τρισδιάστατο τανυστή βαρών W_l , M_l το ύψος και N_l , b_l ένα διάνυσμα πόλωσης διάστασης ίσης με την τρίτη διάσταση του φίλτρου. Στη θέση της μη γραμμικής συνάρτησης ReLU θα μπορούσε να τοποθετηθεί οποιαδήποτε άλλη μη γραμμική συνάρτηση ενεργοποίησης. Στην πράξη, μέσω πειραματισμού έχει φανεί πως η ReLU δίνει τα καλύτερα αποτελέσματα σε πολλά προβλήματα [26, 24]. Μία σημαντική λεπτομέρεια που αποτελεί σχεδιαστική επιλογή, στην συντριπτική πλειοψηφία τέτοιων μοντέλων, είναι ότι ο αριθμός καναλιών της εισόδου ισούται με την τρίτη διάσταση του τανυστή των φίλτρων. Αυτό υποδηλώνεται από την παραπάνω εξίσωση μέσω του συμβόλου "o" που υπονοεί πως όλα τα στοιχεία του τανυστή κατά μήκος αυτής της τρίτης διάστασης, για σταθερή γραμμή m και στήλη n , χρησιμοποιούνται κατά τον υπολογισμό του κάθε γινομένου. Συνεπώς, τα γινόμενα της εξίσωσης εκτελούνται στοιχείο προς στοιχείο ανάμεσα στα βάρη του φίλτρου και το δεκτικό πεδίο του επί της εισόδου.



8-Σχήμα 3.8: CNN model architecture

Τα CNN ορισμένες φορές περιλαμβάνουν τα επίπεδα υπερδειγματοληψίας τα οποία μειώνουν τις διαστάσεις των δεδομένων από προηγούμενα επίπεδα συγκεντρώνοντας τις εξόδους των νευρώνων σε ομάδες. Το τοπικό επίπεδο υπερδειγματοληψίας (local pooling) κάνει μικρές ομαδοποιήσεις, συνήθως 2 x 2. Αντίθετα, το καθολικό επίπεδο υπερδειγματοληψίας (global pooling) επιδρά σε όλους τους νευρώνες του convolutional layer. Η τεχνική της υπερδειγματοληψίας, μπορεί να υπολογίζει τη μέγιστη τιμή (max pooling) ή τον μέσο όρο (average pooling) της κάθε ομάδας από νευρώνες.



9-Σχήμα 3.9: Recurrent Neural Network Unfolding

3.3.5 Αναδρομικά Νευρωνικά Δίκτυα - Recurrent Neural Networks

Τα αναδρομικά νευρωνικά δίκτυα είναι μία ειδική κατηγορία νευρωνικών δικτύων που επεξεργάζονται αποτελεσματικά κάθε είδους ακολουθιακά δεδομένα. Παραδείγματα, αποτελούν η φωνή, η γραφή, η οπτική πληροφορία που προκύπτει από μία κίνηση ή μία δράση ή ακόμα και τα pixels μίας εικόνας αν τα διατρέξουμε με κάποιο δομημένο τρόπο [14].

Ένα αναδρομικό δίκτυο μετασχηματίζει κάθε νέα είσοδο με τρόπο που εξαρτάται τόσο από την ίδια την είσοδο όσο και από τις προηγούμενες εισόδους που έχει δεχτεί. Φορμαλιστικά αυτή η βασική αρχή περιγράφεται ως εξής: Αν \mathbf{X}_t είναι η είσοδος την χρονική στιγμή t , $f()$ η συνάρτηση που περιγράφει την επίδραση του αναδρομικού δικτύου πάνω στην είσοδο και \mathbf{h}_t η έξοδος του, τότε:

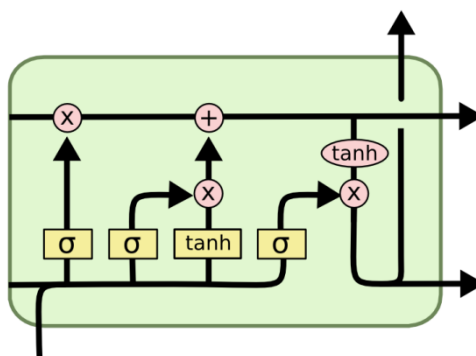
$$h_t = f(x_t, h_{t-1}) = f(x_t, f(x_{t-1}, h_{t-2})) = \dots = f(x_t, f(x_{t-1}, \dots, (f(x_1, h_0), \dots))) \quad (3.5)$$

Είναι φανερό πως υπάρχει στενή σχέση ανάμεσα στα γραφικά μοντέλα (graphical models) και τα αναδρομικά δίκτυα. Αυτή η σύνδεση ανάμεσα στις δύο κατηγορίες μοντέλων μπορεί να φανεί καλύτερα αν, όπως στο Σχήμα 2.8, το δίκτυο "ξεδιπλωθεί", σχεδιάζοντας "εικονικές" υπολογιστικές μονάδες για να αναπαρασταθούν οι αναδρομικές εκτελέσεις της συνάρτησης f πάνω στις εισόδους \mathbf{x}_t , \mathbf{x}_{t+1} , ..., \mathbf{x}_T . Σημειώνεται πως η έξοδος \mathbf{h}_t συχνά αναφέρεται και ως κατάσταση του δικτύου. Στις παραπάνω εξισώσεις δεν εισήχθη για λόγους απλότητας η ύπαρξη ενός συνόλου παραμέτρων θ ή βαρών που όπως και στις άλλες περιπτώσεις νευρωνικών μοντέλων καθορίζουν την επίδραση του μοντέλου πάνω στην είσοδο τους. Οι παράμετροι αυτοί, και για τα αναδρομικά νευρωνικά δίκτυα είναι υπό μάθηση. Τα βάρη σε ένα αναδρομικό μοντέλο μπορούν να διαχωριστούν σε δύο κατηγορίες: αυτά που επιδρούν πάνω στην είσοδο (W) και αυτά που καθορίζουν την σημασία που δίνεται στην προηγούμενη κατάσταση του μοντέλου για τον υπολογισμό της επόμενης κατάστασης (U). Η γενική μορφή των εξισώσεων ενός αναδρομικού δικτύου είναι η εξής:

$$h_t = \varphi(Wx_t + Uh_{t-1} + b) \quad (3.6)$$

όπου (\cdot) μία μη γραμμική συνάρτηση ενεργοποίησης, W ο πίνακας παραμέτρων που επιδρούν πάνω στην είσοδο \mathbf{Q}_t , U ο πίνακας παραμέτρων που επιδρούν πάνω στην έξοδο του δικτύου την προηγούμενη χρονική στιγμή και b ένα διάνυσμα πόλωσης. Λόγω αυτού του μηχανισμού τα αναδρομικά δίκτυα έχουν τη δυνατότητα να μοντελοποιούν χρονικές εξαρτήσεις ακόμα και ανάμεσα σε μη συνεχόμενες παρατηρήσεις, αφού μέσω της αναδρομικότητας υλοποιούν έναν μηχανισμό μνήμης. Όμως, στη γενική μορφή τους τα νευρωνικά αναδρομικά δίκτυα δεν καταφέρνουν να μοντελοποιήσουν αποτελεσματικά εξαρτήσεις μακράς διάρκειας (όπου η χρήση του όρου μακράς είναι σχετική και εξαρτάται από το πρόβλημα). Η δυσκολία αυτή πηγάζει από το γεγονός ότι ο αλγόριθμος Back Propagation στο χρόνο, δεν διατηρεί μεγάλες τιμές κλίσεων όσο εξετάζουμε όλο και παλαιότερα τμήματα της ακολουθίας. Όπως φαίνεται και από την εξίσωση 2.6, ο υπολογισμός της κλίσης του κόστους ως προς κάποια παρελθοντική είσοδο περιλαμβάνει την παραγωγή μιας σύνθεσης συναρτήσεων που οδηγεί σε ένα όλο και αυξανόμενο αριθμού παραγόντων γινόμενου, καθώς κινούμαστε προς το παρελθόν. Κάποιο από τους παράγοντες αυτού του γινομένου είναι πιθανό να οδηγήσουν σε αστάθεια το συνολικό αποτέλεσμα αυξάνοντας ή μειώνοντας υπερβολικά την κλίση του κόστους ως προς τις παραμέτρους. Οι ανανεώσεις των παραμέτρων αυτών που περιγράφει ο αλγόριθμος

Stochastic Gradient Descend βασίζονται στον υπολογισμό της κλίσης και συνεπώς αν σε αυτόν παρουσιαστούν υπερβολικά μεγάλες ή μικρές τιμές η μάθηση θα αποτύχει. Το πρόβλημα αυτό συνήθως αναφέρεται ως εξαφάνιση ή εκτόξευση της κλίσης (the vanishing or exploding gradient problem) και έχει σαν αποτέλεσμα η διαδικασία εκπαίδευσης να μην εντοπίζει ικανοποιητικά καλές τιμές για τα βάρη του μοντέλου.



10-Σχήμα 3.10: LSTM Model Architecture

3.3.6 Νευρωνικό Δίκτυο Μακράς-Βραχείας Μνήμης - LSTM

Η παραπάνω δυσλειτουργία των αναδρομικών δικτύων, αντιμετωπίστηκε μέσω της εισαγωγής των νευρώνων Μακράς-Βραχείας Μνήμης (L.S.T.M) που προτάθηκαν πρώτη φορά από τους Hochreiter και Schmidhuber[15].

Αυτό το πιο εύρωστο είδος αναδρομικών δικτύων έχει φανεί πολύ αποτελεσματικό σε εφαρμογές όπως η επεξεργασία και μετάφραση φυσικής γλώσσας ή η αναγνώριση συνεχούς χειρόγραφου[16]. Αυτός ο τύπος δικτύου, μπορεί να "συγκρατεί" ακολουθίες εισόδου μεγάλου μήκους. Μια συνήθης αρχιτεκτονική των μονάδων του LSTM περιλαμβάνει: μια πύλη εισόδου (input gate), μια πύλη εξόδου (output gate) και μια πύλη (forget gate). Το input gate ελέγχει τη ροή ενός νέου στοιχείου, το forget gate ελέγχει αν θα παραμείνει το στοιχείο αυτό στη μονάδα του LSTM και το output gate ελέγχει τον βαθμό με τον οποίο, το στοιχείο που βρίσκεται στη μονάδα χρησιμοποιείται για να υπολογιστεί το αποτέλεσμα της συνάρτησης ενεργοποίησης της μονάδας. Τα activation functions του πυλών μιας μονάδας LSTM είναι sigmoid functions. Οι υπολογισμοί που πραγματοποιεί μια LSTM μονάδα κατά το εμπρόσθιο πέρασμα, είναι οι παρακάτω:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (3.7)$$

$$i_t = \sigma(W_i x_t + U_f h_{t-1} + b_i) \quad (3.8)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (3.9)$$

$$c_t = f_t c_{t-1} + i_t \sigma_c (W_c x_t + U_c h_{t-1} + b_c) \quad (3.10)$$

$$h_t = o_t \sigma_h (c_t) \quad (3.11)$$

3.3.7 Υπερπαράμετροι αλγορίθμων

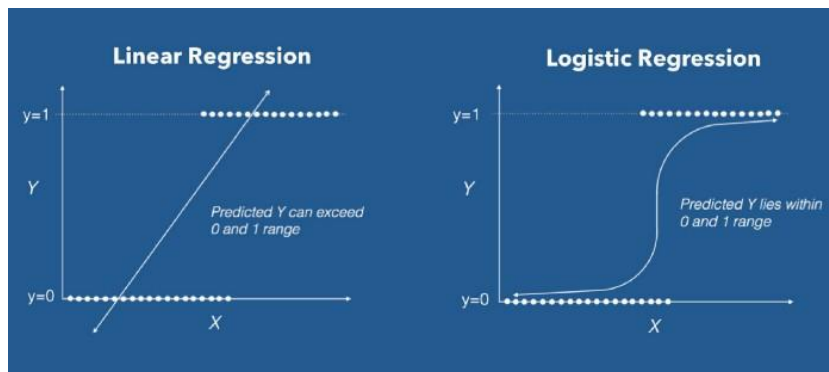
Μια υπερπαράμετρος είναι μια παράμετρος της οποίας η τιμή ορίζεται πριν ξεκινήσει η διαδικασία μηχανικής εκμάθησης. Αντίθετα, οι τιμές άλλων παραμέτρων προέρχονται από την εκπαίδευση. Οι υπερπαράμετροι του αλγορίθμου επηρεάζουν την ταχύτητα και την ποιότητα της μαθησιακής διαδικασίας.

Οι υπερπαράμετροι είναι σημαντικές επειδή μπορούν να έχουν άμεσο αντίκτυπο στη συμπεριφορά του αλγορίθμου εκπαίδευσης και να έχουν σημαντικό αντίκτυπο στην απόδοση του μοντέλου που εκπαιδεύεται. Η επιλογή των κατάλληλων υπερπαραμέτρων διαδραματίζει κρίσιμο ρόλο στην επιτυχία της αρχιτεκτονικής των νευρωνικών δικτύων και έχει τεράστιο αντίκτυπο στο μοντέλο που μαθαίνεται. Για παράδειγμα, αν ο ρυθμός εκμάθησης είναι πολύ χαμηλός, το μοντέλο θα χάσει σημαντικά μοτίβα στα δεδομένα. Αν είναι ψηλά, ενδέχεται να έχει συγκρούσεις.

Μερικά παραδείγματα μοντέλων και υπερπαραμέτρων αναφέρονται παρακάτω:

Γραμμική παλινδρόμηση: Ένα απλό αλλά ισχυρό μοντέλο που χρησιμοποιείται για την πρόβλεψη συνεχών τιμών εξόδου. Υποθέτει μια γραμμική σχέση μεταξύ των μεταβλητών εισόδου και εξόδου. Οι υπερπαράμετροι του μοντέλου είναι ο αριθμός των επαναλήψεων και το ποσοστό μάθησης.

Λογιστική παλινδρόμηση: Ένας τύπος παλινδρόμησης που χρησιμοποιείται για προβλήματα δυαδικής ταξινόμησης, όπου η μεταβλητή εξόδου παίρνει μόνο δύο πιθανές τιμές. Οι υπερπαράμετροι του εν λόγω μοντέλου είναι η ποινή σε λογιστική παλινδρόμηση και η απώλεια σε στοχαστική κλίση κάθοδος



11-Σχήμα 3.11: Παράδειγμα γραμμικής & λογιστικής παλινδρόμησης

Δέντρα αποφάσεων: Ένα μη παραμετρικό μοντέλο που χρησιμοποιεί μια δομή που μοιάζει με δέντρο για να αναπαραστήσει αποφάσεις και τις πιθανές συνέπειές τους. Οι βασικοί υπερπαραμέτροι του μοντέλου αυτού είναι οι παρακάτω:

- κριτήριο («gini» ή «εντροπία») – η συνάρτηση («gini» ή «εντροπία») που χρησιμοποιείται για τον υπολογισμό της αβεβαιότητας στον επιλεγμένο κανόνα διάκρισης.
- splitter ("καλύτερο" ή "τυχαίο") – η στρατηγική που χρησιμοποιείται για την επιλογή του στοιχείου βάσει του οποίου θα δημιουργηθεί ένας κανόνας διάκρισης. Η προεπιλεγμένη τιμή είναι "best". Δηλαδή, για κάθε κόμβο, ο αλγόριθμος λαμβάνει υπόψη όλα τα χαρακτηριστικά και επιλέγει αυτόν με το καλύτερο κριτήριο. Εάν "τυχαία", θα ληφθεί ένα τυχαίο χαρακτηριστικό. Η τεχνική "τυχαία" χρησιμοποιείται για να αποφευχθεί η υπερβολική προσαρμογή.
- max_features (ακέραιος) – Ο αριθμός των χαρακτηριστικών που πρέπει να ληφθούν υπόψη κατά την αναζήτηση του καλύτερου διαχωρισμού (κανόνας βέλτιστης διάκρισης).
- random_state (ακέραιος) – ελέγχει την τυχαιότητα στον αλγόριθμο. Εάν δεν έχει καθοριστεί αριθμός, ο αλγόριθμος θα του δώσει έναν τυχαίο αριθμό για κάθε νέα εκπαίδευση. Αυτό θα αλλάζει το αποτέλεσμα για κάθε νέα προπόνηση. Εάν καθοριστεί ένας αριθμός, ο αλγόριθμος θα έχει πάντα το ίδιο αποτέλεσμα ακόμα και μετά από μια νέα εκπαίδευση.

Τυχαία δάση: Ένας τύπος μοντέλου εκμάθησης συνόλου που συνδυάζει πολλαπλά δέντρα αποφάσεων για να βελτιώσει την ακρίβεια και την ευρωστία των προβλέψεων. Στο τυχαίο δάσος, οι υπερπαραμέτροι είναι ο αριθμός των δέντρων, ο αριθμός των χαρακτηριστικών και ο τύπος των δέντρων (όπως GBM ή M5)

Υποστήριξη διανυσματικών μηχανών (SVMs): Ένα μοντέλο που προσπαθεί να βρει ένα υπερεπίπεδο που διαχωρίζει τα δεδομένα εισόδου σε διαφορετικές κλάσεις. Τα SVMs κάνουν χρήση πολλών υπερπαραμέτρων αλλά οι πιο σημαντικοί είναι οι υπερπαραμέτροι C και kernel.

Η τιμή του C καθορίζει την ποινή για τον ταξινομητή. Εάν το C είναι πολύ μεγάλο, τότε θα υπάρξει μεγάλη ποινή για λάθος κατάταξη στην προπόνηση, άρα μικρό περιθώριο. Και αν το C είναι μικρό, τότε μικρό πέναλτι και μεγάλο περιθώριο.

Ο πυρήνας είναι μια συνάρτηση που χρησιμοποιείται στο SVM για να βοηθήσει στην επίλυση προβλημάτων. Παρέχουν συντομεύσεις για την αποφυγή περίπλοκων υπολογισμών. Το εκπληκτικό με τον πυρήνα είναι ότι μπορούμε να πάμε σε υψηλότερες διαστάσεις και να εκτελέσουμε ομαλούς υπολογισμούς με τη βοήθειά του. Μπορούμε να ανεβούμε σε άπειρο αριθμό διαστάσεων χρησιμοποιώντας πυρήνες.

3.3.8 Υπερπαράμετροι του μοντέλου LSTM

Θα παρουσιάσουμε στην συνέχεια τις υπερπαράμετρους του μοντέλου LSTM.

- Αριθμός κόμβων και κρυφών επιπέδων (Number of nodes & hidden layers):

Τα επίπεδα μεταξύ των επιπέδων εισόδου και εξόδου ονομάζονται κρυφά επίπεδα. Αυτή η θεμελιώδης ιδέα είναι που κάνει τα δίκτυα βαθιάς μάθησης να ονομάζονται «μαύρο κουτί», που συχνά επικρίνονται επειδή δεν είναι διαφανή και ότι οι προβλέψεις τους δεν είναι ανιχνεύσιμες από τον άνθρωπο. Δεν υπάρχει ακριβής αριθμός για το πόσους κόμβους (κρυφούς νευρώνες) ή κρυφά στρώματα πρέπει να χρησιμοποιήσει κάποιος, οπότε ο βέλτιστος αριθμός αυτών των στοιχείων προσδιορίζεται για κάθε μεμονωμένο πρόβλημα με μια προσέγγιση δοκιμής και λάθους.

Ως ένας απλός εμπειρικός κανόνας, ένα κρυφό επίπεδο θα δώσει για τα περισσότερα απλά προβλήματα καλά αποτελέσματα και η χρήση δύο και παραπάνω επιπέδων θα δώσει καλά αποτελέσματα για πιο πολύπλοκα. Επίσης, ενώ πολλοί κόμβοι (με τεχνικές τακτοποίησης) μέσα σε ένα επίπεδο μπορούν να αυξήσουν την ακρίβεια, λιγότερος αριθμός κόμβων μπορεί να προκαλέσει υποπροσαρμογή.

- Αριθμός μονάδων σε πυκνό στρώμα (Number of units on a dense layer)

Μέθοδος: `model.add(Dense(10,...))`

Ένα πυκνό στρώμα είναι το πιο συχνά χρησιμοποιούμενο στρώμα, στο οποίο κάθε νευρώνας λαμβάνει είσοδο από όλους τους νευρώνες του προηγούμενου στρώματος — επομένως, "πυκνά συνδεδεμένο". Τα πυκνά στρώματα βελτιώνουν τη συνολική ακρίβεια του μοντέλου και ενδεικτικά 5–10 κόμβοι ανά επίπεδο είναι μια καλή επιλογή. Έτσι το σχήμα εξόδου του τελικού πυκνού στρώματος θα επηρεαστεί από τον αριθμό των νευρώνων/μονάδων που καθορίζονται.

- Απόσυρση (Dropout)

Μέθοδος: `model.add(LSTM(..., dropout=0.5))`

Κάθε επίπεδο LSTM θα πρέπει να συνοδεύεται από ένα στρώμα απόσυρσης. Ένα τέτοιο στρώμα βοηθά στην αποφυγή υπερβολικής προσαρμογής στην προπόνηση παρακάμπτοντας τυχαία επιλεγμένους νευρώνες, μειώνοντας έτσι την ευαισθησία σε συγκεκριμένα βάρη των μεμονωμένων νευρώνων. Ενώ τα επίπεδα απόσυρσης/εγκατάλειψης μπορούν

να χρησιμοποιηθούν με επίπεδα εισόδου, δεν θα πρέπει να χρησιμοποιούνται με τα επίπεδα εξόδου, καθώς αυτό μπορεί να αναστατώσει την έξοδο από το μοντέλο και τον υπολογισμό του σφάλματος. Ενώ η προσθήκη μεγαλύτερης πολυπλοκότητας μπορεί να κινδυνεύσει να γίνει υπερβολική προσαρμογή (αυξάνοντας τους κόμβους σε πυκνά στρώματα ή προσθέτοντας περισσότερο αριθμό πυκνών στρωμάτων και έχουν χαμηλή ακρίβεια επικύρωσης), αυτό μπορεί να αντιμετωπιστεί με την προσθήκη εγκατάλειψης.

Ένα καλό σημείο εκκίνησης είναι το 20%, αλλά η τιμή απόσυρσης πρέπει να διατηρείται μικρή (έως 50%). Η τιμή του 20% είναι ευρέως αποδεκτή ως ο καλύτερος συμβιβασμός μεταξύ της αποτροπής της υπερβολικής προσαρμογής του μοντέλου και της διατήρησης της ακρίβειας του μοντέλου.

- Αρχικοποίηση βάρους (Weight initialization):

Στην ιδανική περίπτωση, είναι καλύτερο να χρησιμοποιείτε διαφορετικά σχήματα αρχικοποίησης βάρους ανάλογα με τη συνάρτηση ενεργοποίησης που χρησιμοποιείται. Ωστόσο, πιο συχνά χρησιμοποιείται ομοιόμορφη κατανομή κατά την επιλογή αρχικών τιμών βάρους. Δεν είναι δυνατό να τεθούν όλα τα βάρη στο 0,0 καθώς η ασυμμετρία στη διαβάθμιση σφάλματος προκύπτει από τον αλγόριθμο βελτιστοποίησης. για να ξεκινήσετε την αποτελεσματική αναζήτηση. Διαφορετικό σύνολο βαρών έχει ως αποτέλεσμα διαφορετικά σημεία εκκίνησης της διαδικασίας βελτιστοποίησης, οδηγώντας ενδεχομένως σε διαφορετικά τελικά σετ με διαφορετικά χαρακτηριστικά απόδοσης. Τα βάρη θα πρέπει τελικά να αρχικοποιηθούν τυχαία σε μικρούς αριθμούς (μια προσδοκία του αλγόριθμου στοχαστικής βελτιστοποίησης, αλλιώς γνωστή ως στοχαστική κλίση κάθοδος) για να αξιοποιηθεί η τυχαιότητα στη διαδικασία αναζήτησης.

- Ποσοστό απόσπασης (Decay rate):

Η μείωση βάρους μπορεί να προστεθεί στον κανόνα ενημέρωσης βάρους που κάνει τα βάρη να μηδενίζονται εκθετικά, εάν δεν έχει προγραμματιστεί άλλη ενημέρωση βάρους. Μετά από κάθε ενημέρωση, τα βάρη πολλαπλασιάζονται με έναν παράγοντα ελαφρώς μικρότερο από 1, εμποδίζοντας τα έτσι να γίνουν τεράστια. Αυτό καθορίζει την τακτοποίηση στο δίκτυο.

Η προεπιλεγμένη τιμή 0,97 θα πρέπει να είναι αρκετή.

- Λειτουργία ενεργοποίησης(Activation function):

Οι συναρτήσεις ενεργοποίησης είναι αυτές που ορίζουν την έξοδο ενός κόμβου είτε ως ON είτε ως OFF. Αυτές οι συναρτήσεις χρησιμοποιούνται για την εισαγωγή μη γραμμικότητας στα μοντέλα, επιτρέποντας στα μοντέλα βαθιάς μάθησης να μαθαίνουν μη γραμμικά όρια πρόβλεψης. Τεχνικά, οι λειτουργίες ενεργοποίησης μπορούν να συμπεριληφθούν στα πυκνά στρώματα,

αλλά ο διαχωρισμός τους σε διαφορετικά στρώματα καθιστά δυνατή την ανάκτηση της μειωμένης εξόδου του στρώματος πυκνότητας.

Και πάλι, η επιλογή του επιπέδου ενεργοποίησης εξαρτάται από την εφαρμογή, ωστόσο, η λειτουργία ενεργοποίησης ανορθωτή είναι πιο δημοφιλής. Οι συγκεκριμένες καταστάσεις συνεπάγονται συγκεκριμένες λειτουργίες. Για παράδειγμα, η ενεργοποίηση σιγμοειδούς χρησιμοποιείται στο επίπεδο εξόδου για δυαδικές προβλέψεις και το softmax χρησιμοποιείται για την πραγματοποίηση προβλέψεων πολλαπλών κλάσεων (το softmax δίνει την δυνατότητα να ερμηνεύονται οι έξοδοι ως πιθανότητες).

Μέθοδος: Η διαδικασία είναι να δημιουργηθούν συναρτήσεις που ορίζονται από το χρήστη και να επιστρέψει την έξοδο που σχετίζεται με οποιαδήποτε συγκεκριμένη λειτουργία ενεργοποίησης. Για παράδειγμα, εδώ είναι μια συνάρτηση ενεργοποίησης σιγμοειδούς:

```
Def sigmoid(x):  
    return 1/(1+np.exp(-x))
```

Το σιγμοειδές (log-sigmoid) και η υπερβολική εφαπτομένη είναι μερικές από τις πιο δημοφιλείς συναρτήσεις ενεργοποίησης που υιοθετούνται στα μπλοκ LSTM.

- Ρυθμός μάθησης (Learning rate):

Αυτή η υπερπαραμέτρος καθορίζει πόσο γρήγορα το δίκτυο ενημερώνει τις παραμέτρους του. Ο καθορισμός υψηλότερου ρυθμού μάθησης επιταχύνει τη μάθηση, αλλά το μοντέλο μπορεί να μην συγκλίνει (μια κατάσταση κατά τη διάρκεια της εκπαίδευσης όπου η απώλεια ρυθμίζεται εντός ενός εύρους σφάλματος γύρω από την τελική τιμή) ή ακόμη και να αποκλίνει. Αντίθετα, ένας χαμηλότερος ρυθμός θα επιβραδύνει δραστικά τη μάθηση καθώς τα βήματα προς τη συνάρτηση ελάχιστης απώλειας θα είναι μικροσκοπικά, αλλά θα επιτρέψουν στο μοντέλο να συγκλίνει ομαλά.

Συνήθως προτιμάται ένας μειωμένος ρυθμός μάθησης και αυτή η υπερπαραμέτρος χρησιμοποιείται στη φάση της εκπαίδευσης και έχει μια μικρή θετική τιμή, κυρίως μεταξύ 0,0 και 0,1.

- Ορμή (Momentum):

Η υπερπαραμέτρος ορμής έχει ερευνηθεί για να ενσωματωθεί με τα RNN και LSTM. Η ορμή είναι μια μοναδική υπερπαραμέτρος που επιτρέπει τη συσσώρευση των διαβαθμίσεων των προηγούμενων βημάτων για τον καθορισμό της κατεύθυνσης που πρέπει να ακολουθηθεί, αντί να χρησιμοποιείται η κλίση μόνο του τρέχοντος βήματος για να καθοδηγήσει την αναζήτηση.

Συνήθως, η τιμή είναι μεταξύ 0,5 και 0,9.

- Αριθμός επαναλήψεων (Number of epochs):

Αυτή η υπερπαράμετρος ορίζει πόσες πλήρεις επαναλήψεις του συνόλου δεδομένων θα εκτελεστούν. Ενώ θεωρητικά, αυτός ο αριθμός μπορεί να οριστεί σε μια ακέραια τιμή μεταξύ ενός και άπειρου, αυτή θα πρέπει να αυξάνεται μέχρι να αρχίσει να μειώνεται η ακρίβεια επικύρωσης, παρόλο που αυξάνεται η ακρίβεια της προπόνησης (και επομένως υπάρχει κίνδυνος υπερπροσαρμογής).

Μια πλεονεκτική κίνηση είναι να χρησιμοποιηθεί η μέθοδος έγκαιρης διακοπής για να καθοριστεί πρώτα ένας μεγάλος αριθμός εποχών εκπαίδευσης και να σταματήσει η εκπαίδευση μόλις σταματήσει να βελτιώνεται η απόδοση του μοντέλου κατά ένα προκαθορισμένο όριο στο σύνολο δεδομένων επικύρωσης.

- Μέγεθος παρτίδας (Batch size):

Αυτή η υπερπαράμετρος καθορίζει τον αριθμό των δειγμάτων που πρέπει να εργαστούν πριν από την ενημέρωση των εσωτερικών παραμέτρων του μοντέλου. Τα μεγάλα μεγέθη κάνουν μεγάλα βήματα κλίσης σε σύγκριση με τα μικρότερα για τον ίδιο αριθμό δειγμάτων που «βλέπονται».

Ευρέως αποδεκτή, μια καλή προεπιλεγμένη τιμή για το μέγεθος παρτίδας είναι το 32 και πολλαπλάσια της.

3.4 Μετρικές Σφαλμάτων

Τα μοντέλα που θα υλοποιηθούν στην παρούσα διπλωματική θα χρησιμοποιηθούν για βραχυπρόθεσμες προβλέψεις μίας ώρας, 12 ωρών της ίδια ημέρας και των επόμενων 24 ωρών της επόμενης ημέρας. Διατηρώντας ένα μέρος των αρχικών παρατηρήσεων ως «δείγμα επαλήθευσης» θα αξιολογήσουμε πόσο ικανοποιητικές είναι οι παραχθείσες προβλέψεις.

Η αξιολόγηση των μεθόδων πρόβλεψης που χρησιμοποιήθηκαν γίνεται με τα παρακάτω μέτρα ακρίβειας που στηρίζονται στο σφάλμα πρόβλεψης, που εκφράζει στην απόκλιση της προβλεπόμενης τιμής από την πραγματική τιμή [9][17][18]. Συνήθως συμβολίζεται με e_i και δίνεται από τον τύπο $Y - Y_i$, όπου Y η πραγματική τιμή και Y_i η προβλεπόμενη τιμή που μας παράγει η μέθοδος πρόβλεψης που επιλέξαμε. Παρακάτω θα αναφέρουμε τα πλέον συνηθισμένα μέτρα ακρίβειας εκφρασμένα με το σφάλμα πρόβλεψης όπως ορίστηκε πιο πάνω και το N που είναι το πλήθος των δεδομένων της χρονοσειράς. Τα πλέον χρησιμοποιούμενα μέτρα ακρίβειας

ενός εφαρμοσμένου μοντέλου αναφέρονται παρακάτω. Για καλύτερη εικόνα των αποτελεσμάτων συνήθως χρησιμοποιούνται ταυτόχρονα, παραπάνω από ένα μέτρα ακρίβειας στις παραχθείσες προβλέψεις.

- **Το μέσο σφάλμα πρόβλεψης - Mean Forecast Error (MFE) ή απλά ME.**

Το μέσο σφάλμα αποτελεί ένα μέτρο της μέσης απόκλισης των προβλεπόμενων τιμών από τις πραγματικές και υπολογίζεται ως $ME = \frac{1}{N} e_t$. Δείχνει την κατεύθυνση του σφάλματος και για αυτό ονομάζεται και αμεροληψία πρόβλεψης. Για μια καλή πρόβλεψη είναι επιθυμητό το ME να είναι όσο το δυνατόν πιο κοντά στο μηδέν. Στο ME, οι επιπτώσεις των θετικών και αρνητικών σφαλμάτων ακυρώνονται και δεν πάχρει τρόπος να γνωρίζουμε το ακριβές ποσό τους, ενώ ένα μηδενικό ME δεν σημαίνει ότι οι προβλέψεις είναι τέλειες, δηλαδή ότι δεν περιέχουν σφάλματα, παρά μόνο δείχνει ότι οι προβλέψεις είναι προς τη σωστή κατεύθυνση. Ακόμη, δεν περικλείει ακραία σφάλματα και εξαρτάται από την κλίμακα μέτρησης των δεδομένων καθώς επίσης επηρεάζεται από τους μετασχηματισμούς των δεδομένων.

- **Το μέσο απόλυτο σφάλμα Mean Absolute Error (MAE) ή Μέση Απόλυτη Απόκλιση Mean Absolute Deviation (MAD)**

Αποτελεί το πιο απλό κριτήριο καθώς υπολογίζει την μέση απόλυτη απόκλιση των προβλέψεων από τις πραγματικές τιμές και γι' αυτό ονομάζεται και μέση απόλυτη απόκλιση και δίνεται από τον τύπο $MAE = \frac{1}{N} * \sum_{t=1}^N |e_t|$. Σε αντίθεση με το ME στο MAE τα θετικά και αρνητικά σφάλματα δεν εξουδετερώνονται οπότε παράλληλα δεν δείχνει και την κατεύθυνση των σφαλμάτων, δείχνει όμως όλο το μέγεθος των σφαλμάτων. Επίσης εξαρτάται από την κλίμακα μέτρησης και επηρεάζεται από τους μετασχηματισμούς δεδομένων. Για να είναι καλές οι προβλέψεις θα πρέπει το MAE να είναι όσο το δυνατόν μικρότερο.

- **Το μέσο απόλυτο ποσοστιαίο σφάλμα - Mean Absolute Percentage Error (MAPE)**

Αντανακλά το ποσοστό του μέσου απόλυτου σφάλματος, δεν δείχνει την κατεύθυνση των σφαλμάτων αλλά ούτε και μεγάλες αποκλίσεις που πιθανόν να υπάρχουν. Το μέσο απόλυτο ποσοστιαίο σφάλμα δίνεται από τον τύπο $MAPE = \frac{1}{N} * \sum_{t=1}^N \frac{|e_t|}{Y_t}$. Επειδή υπολογίζει τα απόλυτα, πάλι τυχόν ετερόσημα σφάλματα δεν εξουδετερώνονται. Σε αντίθεση με τα παραπάνω σφάλματα το MAPE είναι ανεξάρτητο από την κλίμακα μεγέθους, επηρεάζεται όμως από τυχόν μετασχηματισμούς στα δεδομένα. Επειδή το MAPE είναι απαλλαγμένο από μονάδες μέτρησης το χρησιμοποιούμε για να συγκρίνουμε την ακρίβεια ανάμεσα σε διαφορετικές μεθόδους πρόβλεψης.

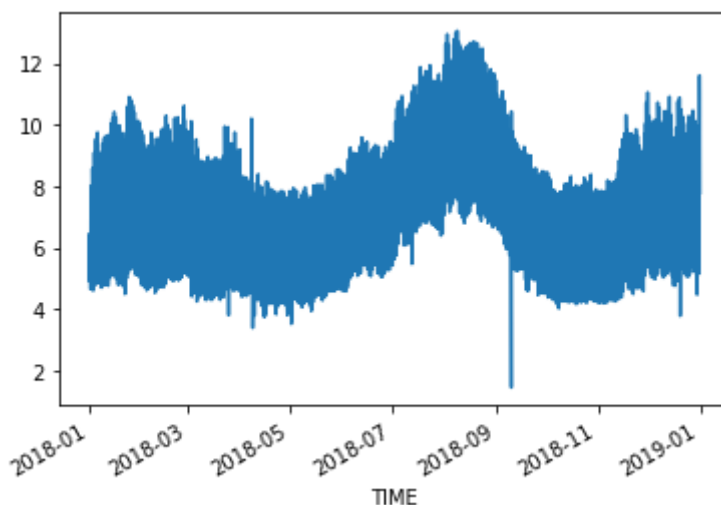
ΚΕΦΑΛΑΙΟ 4: Ανάλυση & Πρόβλεψη φορτίου χρονοσειράς ηλεκτρικής ενέργειας με χρήση νευρωνικού δικτύου LSTM

4.1 Εισαγωγή

Στο παρόν κεφάλαιο θα παρουσιάσουμε την ανάλυση της χρονοσειράς με βάση την τάση, την εποχικότητα, την κυκλική διακύμανση και την αυτοσυσχέτιση μεταξύ των στοιχείων της. Στην συνέχεια, με χρήση στατιστικών αλγορίθμων και νευρωνικών δικτύων LSTM θα προβλέψουμε το φορτίο ηλεκτρικής ενέργειας [17] [18]. Τέλος, για κάθε ένα από τους στατιστικούς αλγορίθμους καθώς και το νευρωνικό δίκτυο LSTM θα παρουσιάσουμε τα αποτελέσματα και θα συγκρίνουμε την ποιότητα τους με βάση το μέσο απόλυτο ποσοστιαίο σφάλμα. Για την εκτέλεση των ανωτέρω διαδικασιών έγινε χρήση του περιβάλλοντος Colab της εταιρείας Google και ο κώδικας στον οποίο αποτυπώθηκαν οι αλγόριθμοι γράφτηκε σε γλώσσα Python.

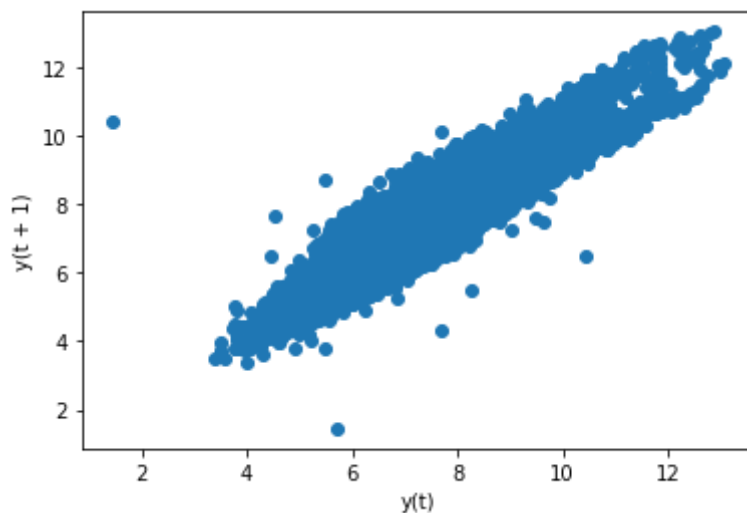
4.2 Χρονοσειρά φορτίου

Η χρονοσειρά μας περιλαμβάνει το ωριαίο φορτίο ηλεκτρικής ενέργειας της χώρας για την χρονιά 2018-2019. Σκοπός την εν λόγω ενότητας είναι η ανάλυση της χρονοσειράς βάσει της τάσης, της εποχικότητας, της αυτοσυσχέτισης, της κυκλικής διακύμανσης και άλλων παραμέτρων που θα μας ορίσουν την συμπεριφορά του φορτίου στον συγκεκριμένο χρόνο.



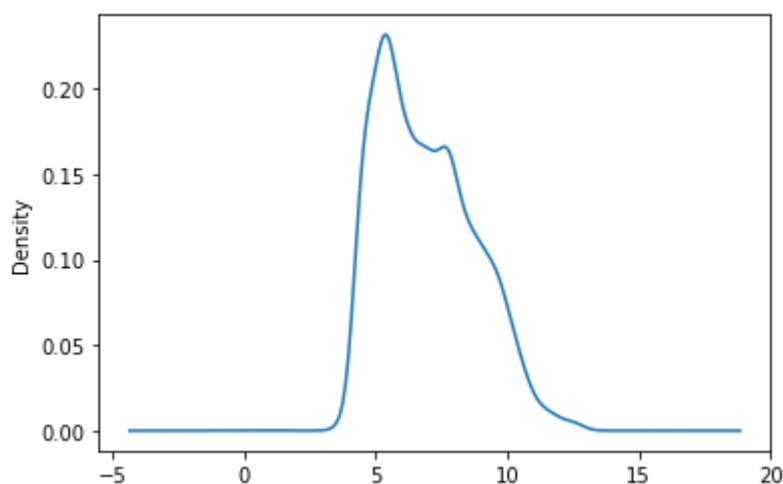
12-Σχήμα 4.1: Αναπαράσταση του φορτίου σε MW για το έτος 2018

Παρακάτω κάνοντας χρήση της διαδικασίας καθυστέρησης (lag_plot) παρατηρούμε μια συσχέτιση μεταξύ επομένου και προηγούμενου στοιχείου της χρονοσειράς.



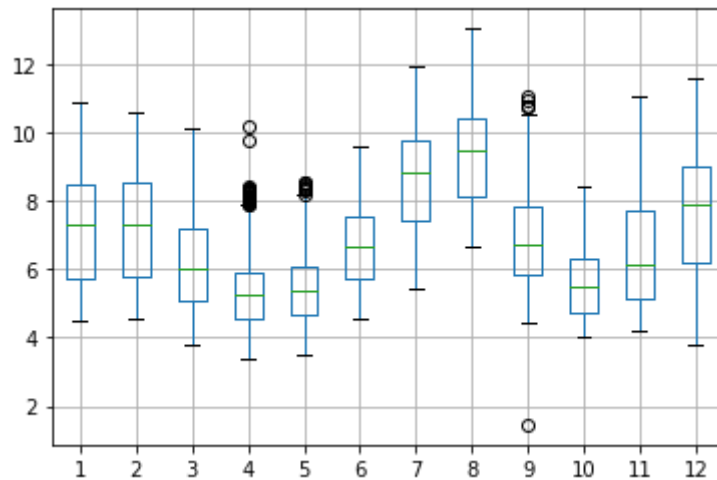
13-Σχήμα 4.2: Συσχέτιση μεταξύ προηγούμενων και επόμενων τιμών της χρονοσειράς

Η ποσότητα του φορτίου της προς εξέταση περιοχής προσδιορίζεται στο επόμενο γράφημα :



14-Σχήμα 4.3: Κατανομή τιμών χρονοσειράς

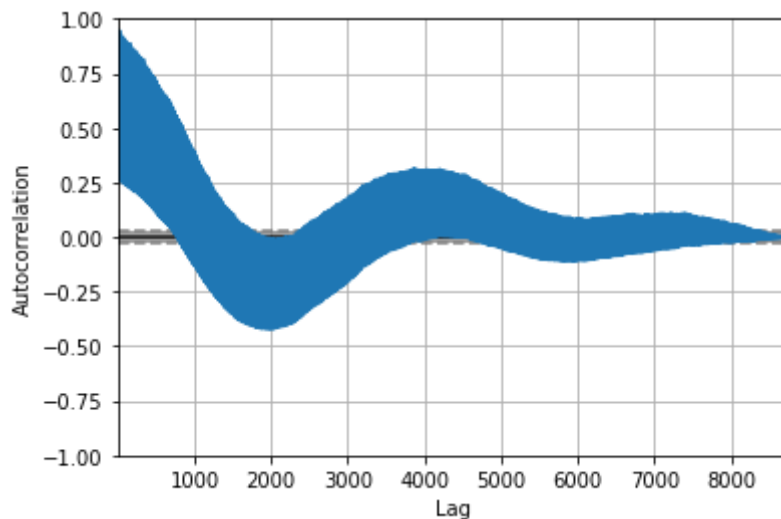
, και υποδηλώνει μια κατανάλωση εύρους από 4.5 μέχρι 12MW ανά ώρα με μία μέση κατανάλωση 6MW ανά ώρα . Στην συνέχεια, παρουσιάζονται τα φορτία ανά μήνα:



15-Σχήμα 4.4: Κατανομή εύρους τιμών φορτίου ανά μήνα

Όπου ο οριζόντιος άξονας αναφέρεται στους μήνες του έτους και ο κάθετος στην κατανάλωση ηλεκτρικής ενέργειας.

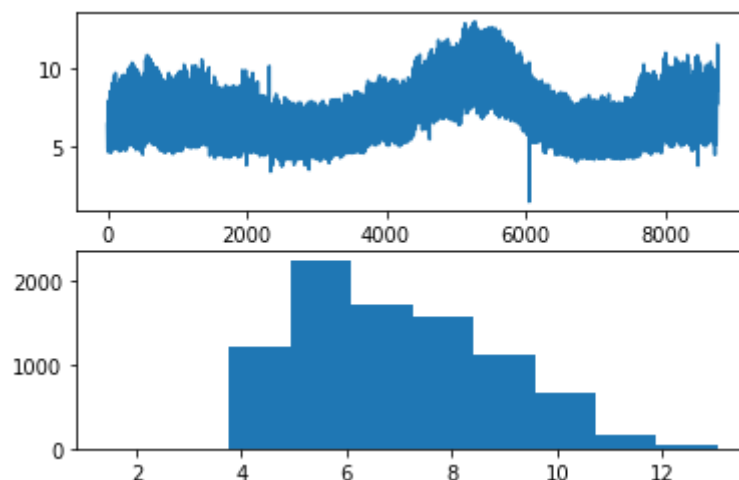
Η παρακάτω αναπαράσταση των τιμών της χρονοσειράς :



16-Σχήμα 4.5: Το διάγραμμα αυτοσυσχέτισης της χρονοσειράς

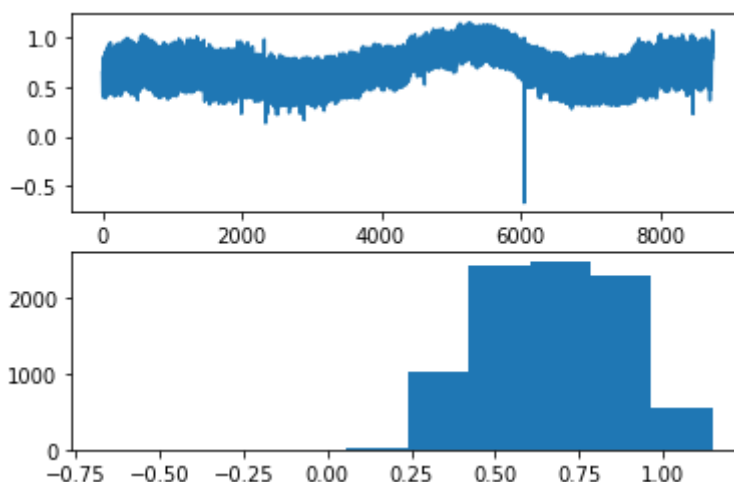
παρουσιάζει την αυτοσυσχέτιση μεταξύ των δεδομένων.

Αναλύοντας τα δεδομένα της χρονοσειράς πρέπει να λάβουμε υπόψιν μας την τάση , την εποχικότητα και τα μη γραμμικά στοιχεία αυτής.



17-Σχήμα 4.6: Αναπαράσταση κατανάλωσης ηλεκτρικής ενέργειας στην αντίστοιχη ώρα της χρονοσειράς (πάνω) , και κατανομή τιμών φορτίου της χρονοσειράς (κάτω)

Κάνοντας χρήση της διαδικασίας Box-Cox ώστε να σταθεροποιήσουμε την διακύμανση



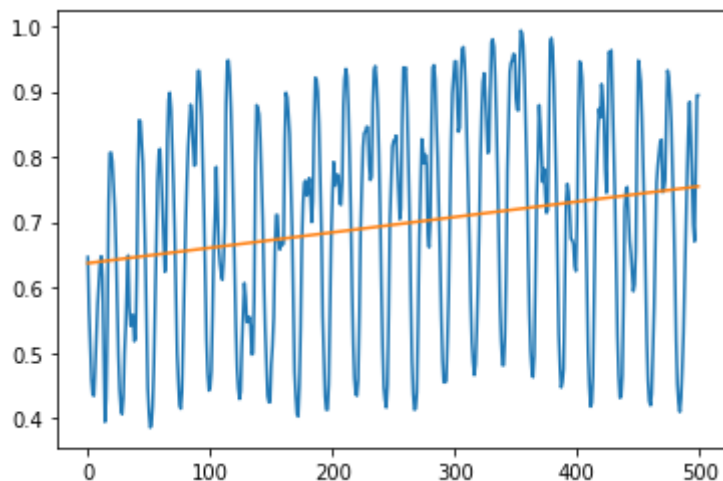
18-Σχήμα 4.7: Παρόμοια με το σχήμα 4.6 αλλά με αναγωγή των τιμών μεταξύ -1 και 1 όπου 1 η μέγιστη τιμή κατανάλωσης της χρονοσειράς ανά ώρα.

Καταλήγουμε στην τιμή της παραμέτρου:

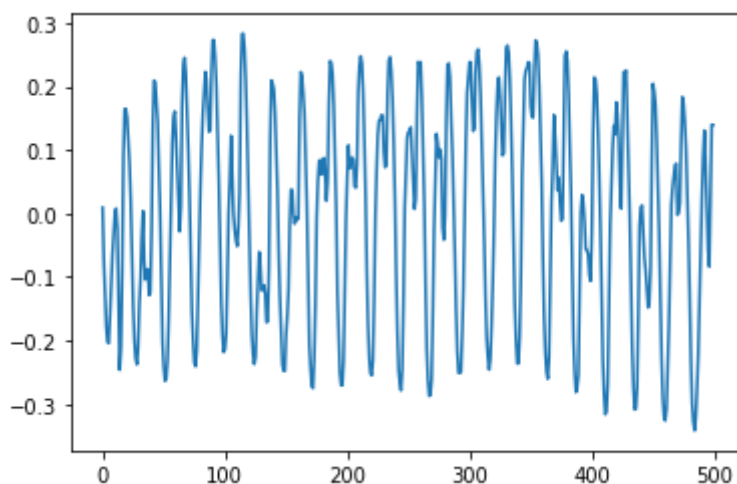
$$\lambda=0.882788$$

η οποία μας καθορίζει κατά πόσο η χρονοσειρά μας έχει σταθερή διακύμανση ή όχι και πιο μοτίβο ακολουθεί. Δεδομένου ότι η παράμετρος $\lambda \neq 1$ μας επιβάλει να την αποσυνθέσουμε ως αναφορά την τάση και εποχικότητα της με αντίστοιχα μοντέλα.

Για να σταθεροποιήσουμε την τάση θα κάνουμε χρήση του μοντέλου της γραμμικής παλινδρόμησης όπως φαίνεται παρακάτω:

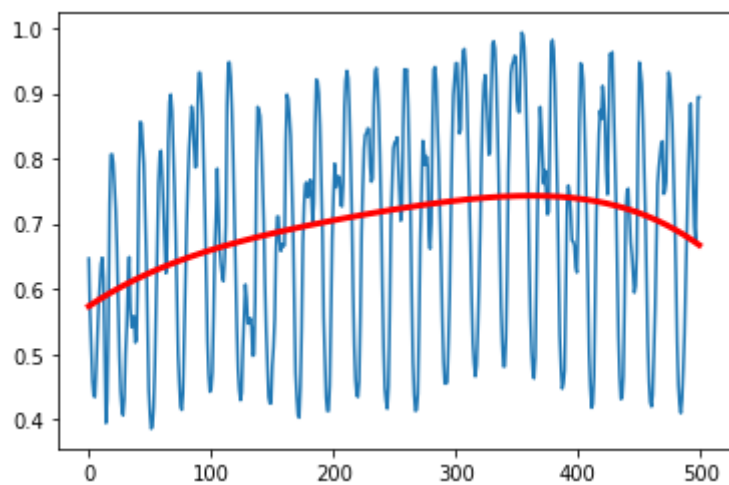


19-Σχήμα 4.8: Χρονοσειρά με Τάση

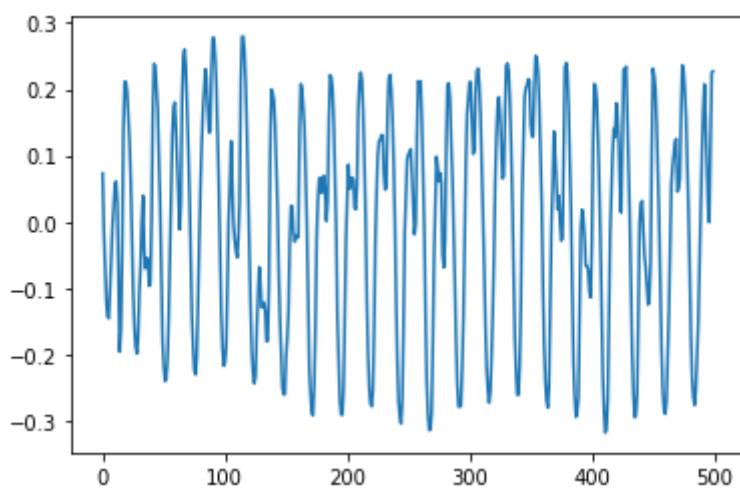


20-Σχήμα 4.9: Χρονοσειρά με απαλοιφή Τάσης

Για την απαλοιφή της εποχικότητας θα κάνουμε χρήση της διαδικασίας Polyfit όπου υπολογίζει ένα πολυώνυμο ελαχίστων τετραγώνων για ένα δεδομένο σύνολο δεδομένων. Το Polyfit δημιουργεί τους συντελεστές του πολυωνύμου, οι οποίοι μπορούν να χρησιμοποιηθούν για τη μοντελοποίηση μιας καμπύλης που ταιριάζει στα δεδομένα. Στην προκειμένη περίπτωση έχουμε:



21-Σχήμα 4.10: Αναπαράσταση της εποχικότητας σε δείγμα 500 τιμών της χρονοσειράς



22-Σχήμα 4.11: Αναπαράσταση εποχικότητας σε δείγμα 500 τιμών της χρονοσειράς

Coefficients
-1.56024662e-11
1.36513265e-08
-5.04892889e-06

1.24680548e-03

5.73489721e-01

23-Σχήμα 4.12: Συντελεστές του πολυωνόμου

Συμπεραίνουμε ότι η σωστή αποτύπωση και ανάλυση της χρονοσειράς των δεδομένων μας βάση της περιοδικότητας, της τάσης, της εποχικότητας και των μη γραμμικών στοιχείων αυτής μπορεί με την χρήση σωστών μοντέλων όπως box-cox αποσύνθεση και ορισμό της παραμέτρου λ να μας καθορίσει την ταυτότητα της και πως θα εισαχθεί καλύτερα στα μοντέλα αλγορίθμων ανάλογα με τον ορίζοντα πρόβλεψης.

4.3 Αλγόριθμος LSTM

4.3.1 Παρουσίαση των μοντέλων ARIMA, MLP & LSTM

Τα μοντέλα ARIMA είναι στοχαστικά μαθηματικά μοντέλα τα οποία μας βοηθάνε να αναλύσουμε και να προβλέψουμε την εξέλιξη μεγεθών. Στην γενική τους μορφή τα μοντέλα ARIMA περιέχουν τον τυχαίο παράγοντα (σφάλμα πρόβλεψης), τιμές του μεγέθους που εμφανίστηκαν σε προηγούμενες περιόδους και σχετικούς στοχαστικούς παράγοντες. Πιο συγκεκριμένα, κάθε μοντέλο ARIMA μπορεί να εκφραστεί ως γραμμικός συνδυασμός των παραπάνω παραγόντων και στόχος μας είναι να ανακαλύψουμε εκείνον που παράγει τις καλύτερες προβλέψεις. Η εφαρμογή των μοντέλων ARIMA προϋποθέτει να πληρούνται ορισμένες απαιτήσεις. Αρχικά, η χρονοσειρά πρέπει να είναι διακριτή, δηλαδή οι παρατηρήσεις της y_t να αναφέρονται σε ισαπέχουσες χρονικές στιγμές $y_t, y_{t+1}, y_{t+2}, \dots$. Εκτός αυτού, η χρονοσειρά πρέπει να είναι στάσιμη. Αυτό σημαίνει πως η μέση τιμή (μ), η διακύμανση (σ^2) και η συνάρτηση αυτοσυσχέτισης (ACF) της χρονοσειράς πρέπει να είναι σταθερές καθ' όλη τη διάρκεια του χρόνου. Έτσι, τα χαρακτηριστικά της δεν εξαρτώνται από τη χρονική στιγμή την οποία αυτή εξετάζεται (βλέπε λευκός θόρυβος) και οποιοδήποτε δείγμα της κατανομής της y_1, y_2, \dots, y_m ταυτίζεται εν γένει με αυτό της $y_{1+t}, y_{2+t}, \dots, y_{m+t}$. Με αυτό τον τρόπο η χρονοσειρά αποδεσμεύεται από την έννοια του χρόνου και μπορεί να μελετηθεί στοχαστικά. Τέλος, η εφαρμογή μοντέλων ARIMA προϋποθέτει την εξαγωγή βραχυπρόθεσμων προβλέψεων.

Ένα MLP είναι μια πλήρως συνδεδεμένη κατηγορία τεχνητού νευρωνικού δικτύου (ANN). Ένα MLP έχει στρώματα εισόδου και εξόδου, και ένα ή περισσότερα κρυφά στρώματα με πολλούς νευρώνες στοιβαγμένους μεταξύ τους. Και ενώ στο Perceptron ο νευρώνας πρέπει να έχει μια συνάρτηση ενεργοποίησης που επιβάλλει ένα όριο, όπως το ReLU ή το σιγμοειδές, οι νευρώνες σε ένα Perceptron πολλαπλών επιπέδων μπορούν να χρησιμοποιήσουν οποιαδήποτε αυθαίρετη συνάρτηση ενεργοποίησης. Το Multilayer Perceptron εμπίπτει στην κατηγορία των αλγορίθμων

feedforward, επειδή οι εισοδοί συνδυάζονται με τα αρχικά βάρη σε ένα σταθμισμένο άθροισμα και υπόκεινται στη συνάρτηση ενεργοποίησης, όπως ακριβώς στο Perceptron. Αλλά η διαφορά είναι ότι κάθε γραμμικός συνδυασμός διαδίδεται στο επόμενο στρώμα. Κάθε στρώμα τροφοδοτεί το επόμενο με το αποτέλεσμα του υπολογισμού του, την εσωτερική του αναπαράσταση των δεδομένων. Αυτό περνάει από τα κρυφά επίπεδα μέχρι το επίπεδο εξόδου. Ένα MLP αποτελείται από τουλάχιστον τρία επίπεδα κόμβων: ένα στρώμα εισόδου, ένα κρυφό στρώμα και ένα στρώμα εξόδου. Εκτός από τους κόμβους εισόδου, κάθε κόμβος είναι ένας νευρώνας που χρησιμοποιεί μια μη γραμμική συνάρτηση ενεργοποίησης. Το MLP χρησιμοποιεί μια εποπτευόμενη τεχνική εκμάθησης που ονομάζεται backpropagation για εκπαίδευση. Τα πολλαπλά στρώματα και η μη γραμμική ενεργοποίησή του διακρίνουν το MLP από ένα γραμμικό perceptron. Μπορεί να διακρίνει δεδομένα που δεν διαχωρίζονται γραμμικά.

Το LSTM είναι ένα RNN που έχει και ανατροφοδοτούμενες συνδέσεις που μπορεί να επεξεργαστεί εκτός από σημειακά δεδομένα όπως εικόνες, και σειρές δεδομένων όπως η χρονοσειρές. Αυτή η ιδιότητα είναι πολύ χρήσιμη σε εφαρμογές για την επεξεργασία και πρόβλεψη φορτίου ηλεκτρικής ενέργειας.

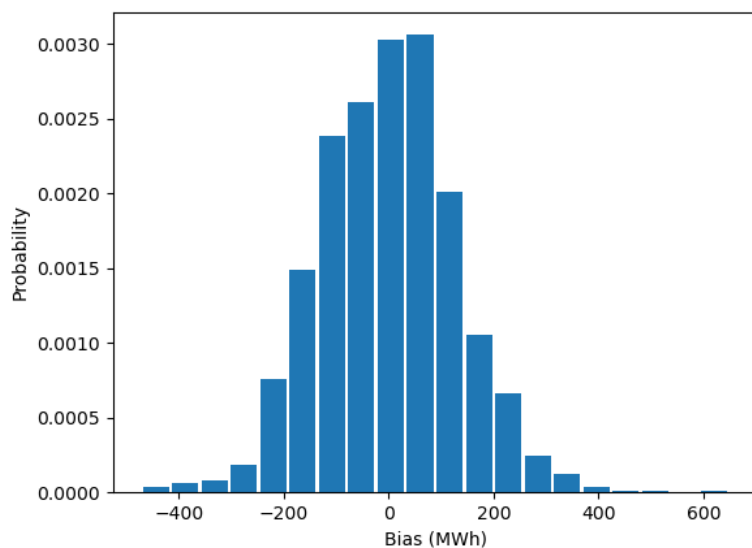
4.3.2 Σύγκριση

Σκοπός της παρούσας ενότητας είναι να παρουσιαστεί η σύγκριση μεταξύ των τριών αλγορίθμων με βάση τα σφάλματα MAPE, MAE & ME, δηλαδή την ακρίβεια των προβλέψεων κατανάλωσης ηλεκτρικής ενέργειας σε διαφορετικά διαστήματα και για διαφορετικό ορίζοντα πρόβλεψης. Επομένως, η χρονοσειρά έχει χωριστεί σε διαστήματα των 24 ωρών και 168 ωρών για την πρόβλεψη του φορτίου ηλεκτρικής ενέργειας της επόμενης ώρας (continuous 1h), των επόμενων 12 ωρών της παρούσας ημέρας (intra day 12h) και των επόμενων 24 ωρών (day ahead).

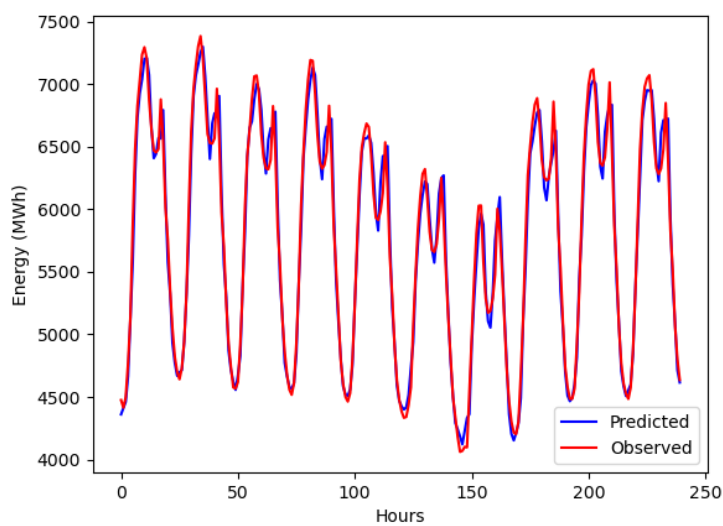
4.3.2.1 Πρόβλεψη επόμενης ώρας

Η πρόβλεψη επόμενης ώρας μας βοηθάει να κατανοήσουμε στην αμέσως επόμενη (very short term) χρονική περίοδο την κατανάλωση ηλεκτρικής ενέργειας. Σε αυτήν την ενότητα, αξιολογείται η ακρίβεια του μοντέλου LSTM και ύστερα γίνεται σύγκριση με τα μοντέλα αναφοράς, ARIMA & MLP. Μέσω της σύγκρισης, η παρούσα εργασία εξετάζει εάν το νευρωνικό δίκτυο LSTM είναι καταλληλότερο για τη δημιουργία βραχυπρόθεσμων προβλέψεων που αφορούν χρονοσειρές κατανάλωσης ηλεκτρικής ενέργειας. Παρακάτω βλέπουμε τα διαγράμματα ανάλυσης της χρονοσειράς και την πρόβλεψη του νευρωνικού LSTM.

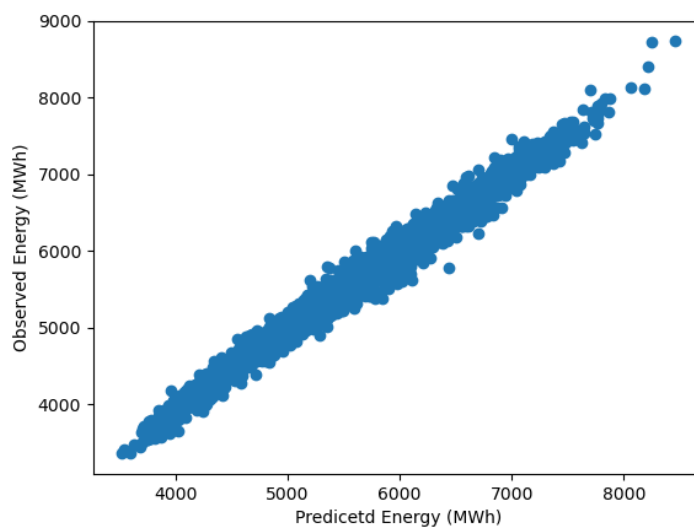
Για χρονικό διάστημα δεδομένων εισόδου στο μοντέλο πρόβλεψης 24 ώρες:



24-Σχήμα 4.13: Διάγραμμα κατανομής σφάλματος ΜΕ μοντέλου LSTM



25-Σχήμα 4.14: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM



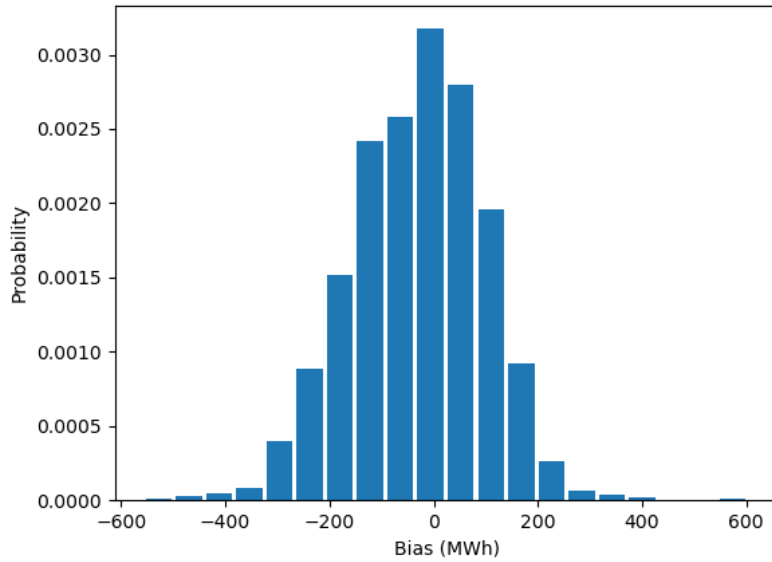
26-Σχήμα 4.15: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς

Στον παρακάτω πίνακα αναγράφονται τα αποτελέσματα και η σύγκριση και για τα τρία μοντέλα πρόβλεψης για διάστημα πρόβλεψης μίας ώρα και διάστημα χρήσης για την είσοδο στο μοντέλο πρόβλεψης 24 ώρες :

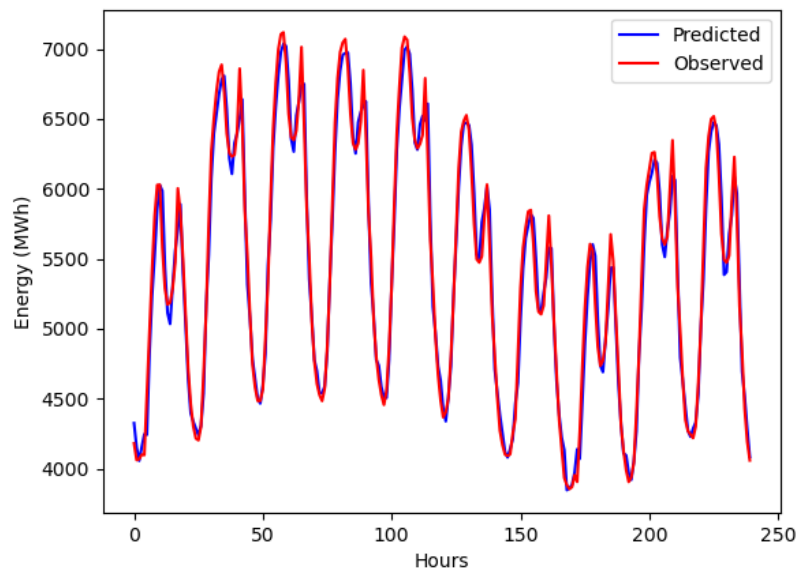
Α/Α	ΤΥΠΟΣ ΣΦΑΛΜΑΤΟΣ		
	ΜΑΡΕ (%)	ΜΑΕ(MWh)	ΜΕ (MWh)
ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ			
ARIMA	4,13	194,54	-45,15
MLP	2,89	150,64	30,05
LSTM	1,89	102,41	-0,46
ΟΡΙΖΟΝΤΑΣ ΠΡΟΒΛΕΨΗΣ	1h		
ΔΙΑΣΤΗΜΑ ΔΕΔΟΜΕΝΩΝ ΕΙΣΟΔΟΥ	24h		

27-Σχήμα 4.16: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM

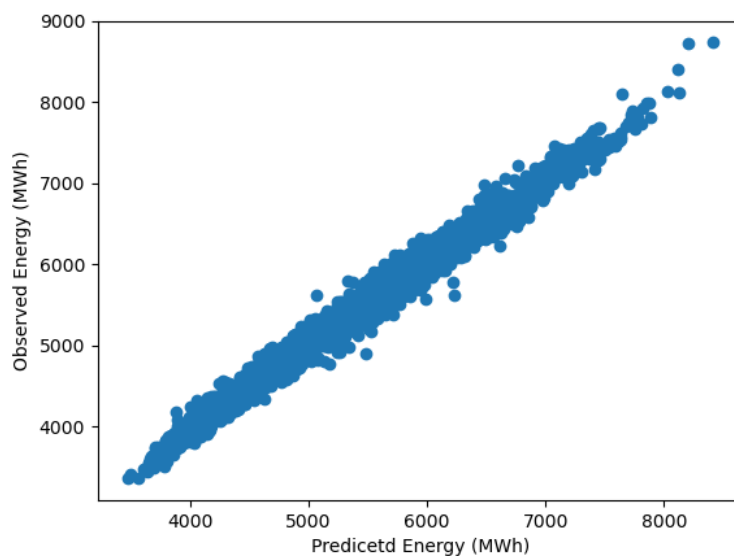
Για χρονικό διάστημα δεδομένων εισόδου στο μοντέλο πρόβλεψης 168 ώρες:



28-Σχήμα 4.17: Διάγραμμα κατανομής σφάλματος ME μοντέλου LSTM



29-Σχήμα 4.18: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM



30-Σχήμα 4.19: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς

Ομοίως, παρακάτω αναγράφονται τα αποτελέσματα και η σύγκριση και για τα τρία μοντέλα πρόβλεψης για διάστημα πρόβλεψης μίας ώρα και διάστημα χρήσης για την είσοδο στο μοντέλο πρόβλεψης 168 ώρες :

A/A ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ	ΤΥΠΟΣ ΣΦΑΛΜΑΤΟΣ		
	MAPE (%)	MAE(MWh)	ME (MWh)
ARIMA	3,87	166,26	50,81
MLP	2,54	134,57	-40,15
LSTM	1,92	102,70	-26,33
ΟΡΙΖΟΝΤΑΣ ΠΡΟΒΛΕΨΗΣ	1h		
ΔΙΑΣΤΗΜΑ ΔΕΔΟΜΕΝΩΝ ΕΙΣΟΔΟΥ	168h		

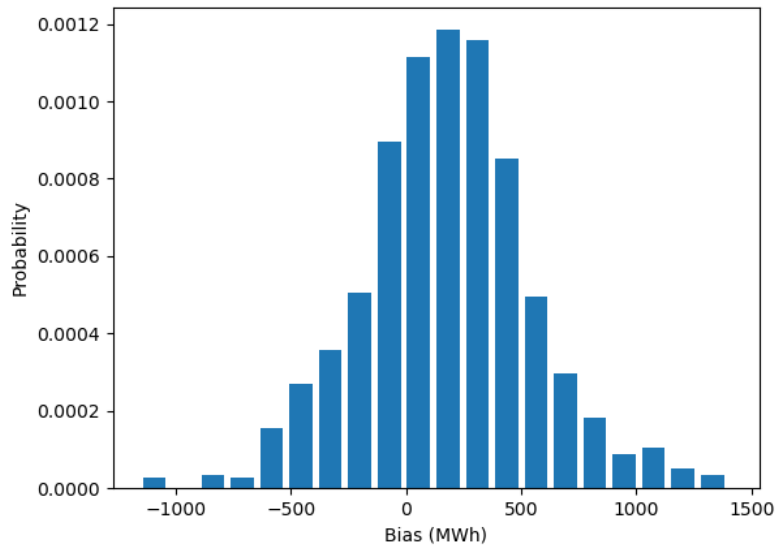
31-Σχήμα 4.20: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM

4.3.2.2 Πρόβλεψη επόμενων 12 ωρών

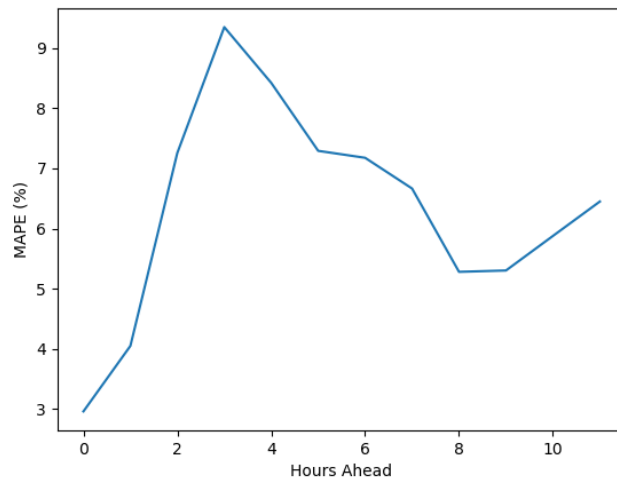
Η πρόβλεψη των επόμενων 12 ωρών μέσα στην ημέρα μας μας βοηθάει να κατανοήσουμε για την κοντινή (short term) περίοδο την κατανάλωση ηλεκτρικής ενέργειας . Παρόμοια ,όπως στο 4.3.2.1, αξιολογείται η ακρίβεια του μοντέλου LSTM και ύστερα γίνεται σύγκριση με τα μοντέλα αναφοράς, ARIMA & MLP. Μέσω της σύγκρισης, η παρούσα εργασία εξετάζει εάν το νευρωνικό δίκτυο LSTM είναι καταλληλότερο για τη δημιουργία βραχυπρόθεσμων προβλέψεων που αφορούν

χρονοσειρές κατανάλωσης ηλεκτρικής ενέργειας. Παρακάτω βλέπουμε τα διαγράμματα ανάλυσης της χρονοσειράς και την πρόβλεψη του νευρωνικού LSTM.

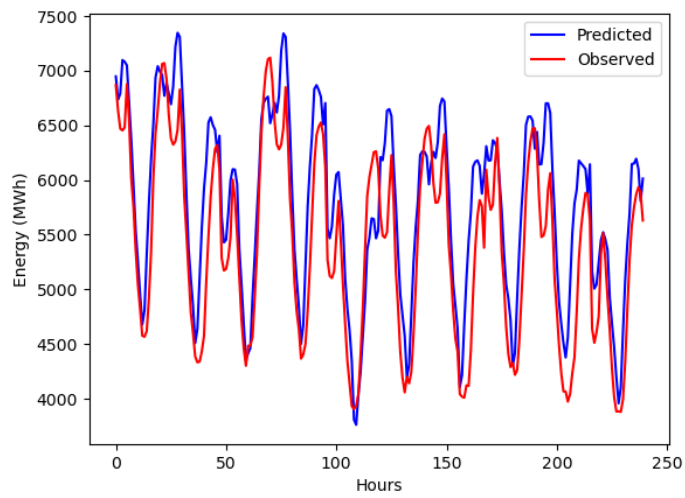
Για χρονικό διάστημα δεδομένων εισόδου στο μοντέλο πρόβλεψης 24 ώρες:



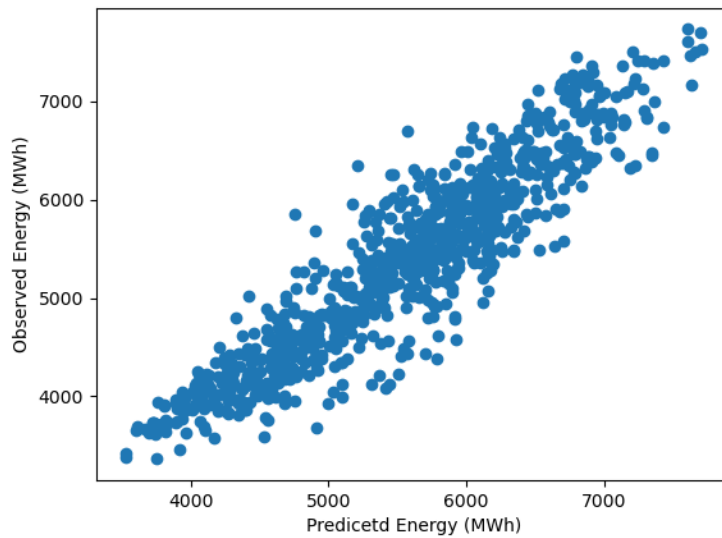
32-Σχήμα 4.21: Διάγραμμα κατανομής σφάλματος ME μοντέλου LSTM



33-Σχήμα 4.22: Διάγραμμα σφάλματος MAPE μοντέλου LSTM



34-Σχήμα 4.23: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM

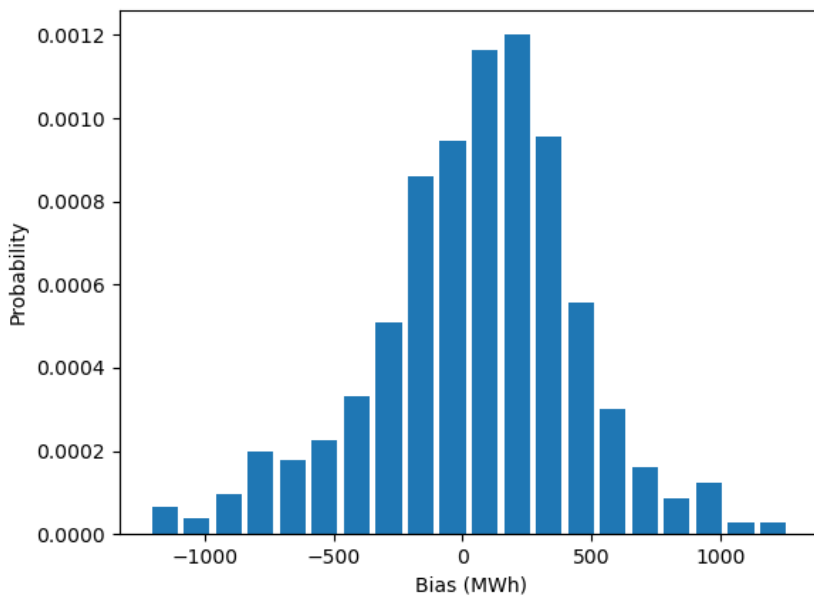


35-Σχήμα 4.24: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς

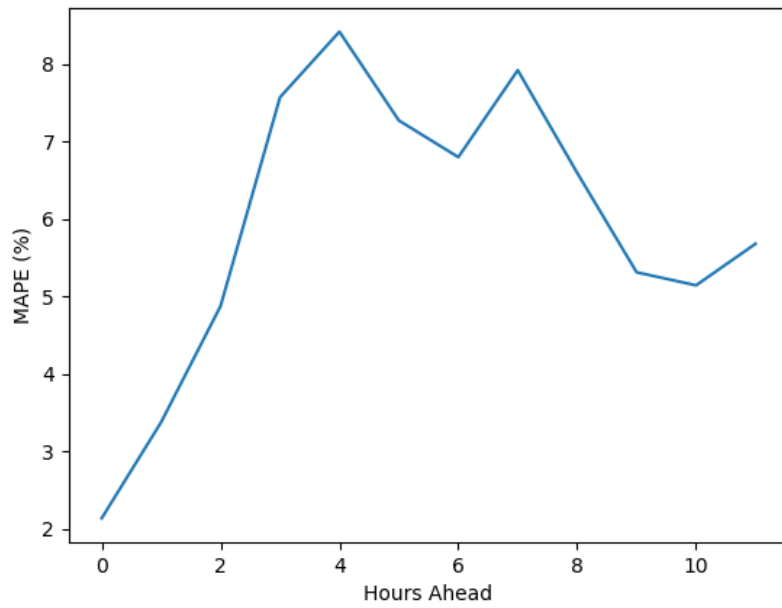
Α/Α ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ	ΤΥΠΟΣ ΣΦΑΛΜΑΤΟΣ		
	ΜΑΡΕ (%)	ΜΑΕ(MWh)	ΜΕ (MWh)
ARIMA	9,73	570,48	-120,90
MLP	8,25	421,24	150,38
LSTM	6,34	325,74	190,26
ΟΡΙΖΟΝΤΑΣ ΠΡΟΒΛΕΨΗΣ	12h		
ΔΙΑΣΤΗΜΑ ΔΕΔΟΜΕΝΩΝ ΕΙΣΟΔΟΥ	24h		

36-Σχήμα 4.25: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM

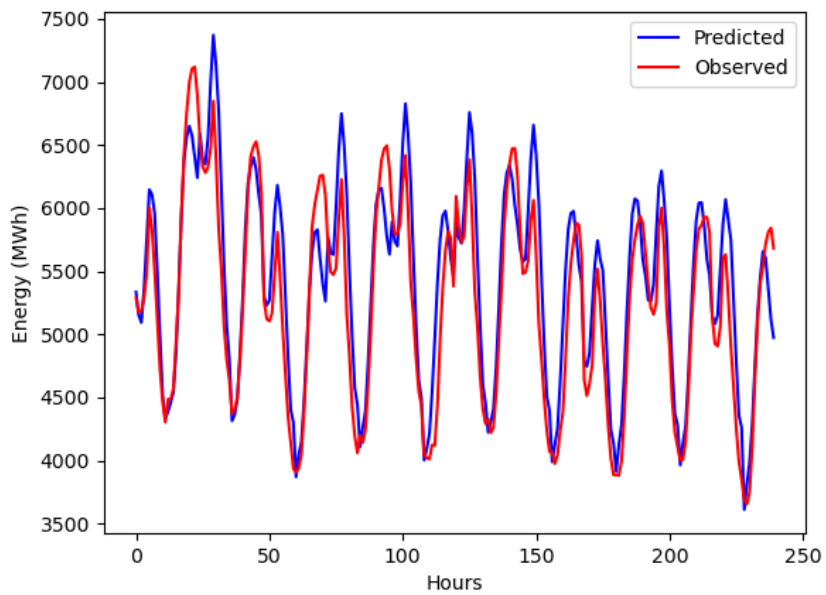
Για χρονικό διάστημα δεδομένων εισόδου στο μοντέλο πρόβλεψης 168 ώρες:



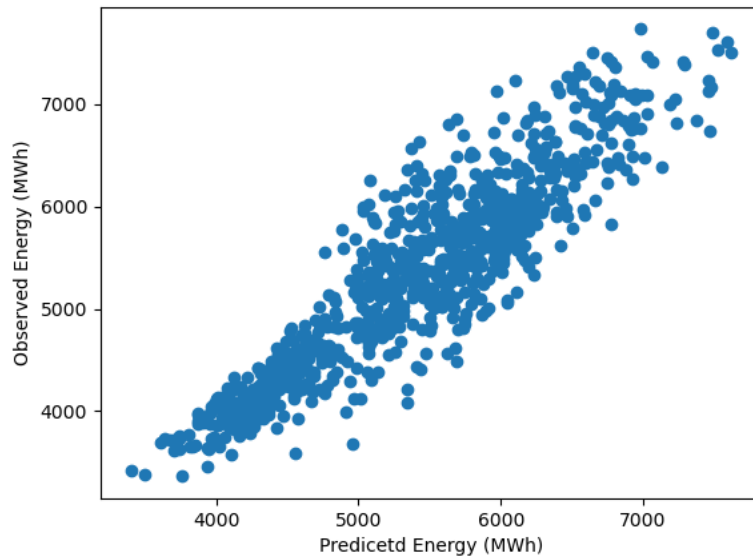
37-Σχήμα 4.26: Διάγραμμα κατανομής σφάλματος ΜΕ μοντέλου LSTM



38-Σχήμα 4.27: Διάγραμμα σφάλματος MAPE μοντέλου LSTM



39-Σχήμα 4.28: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM



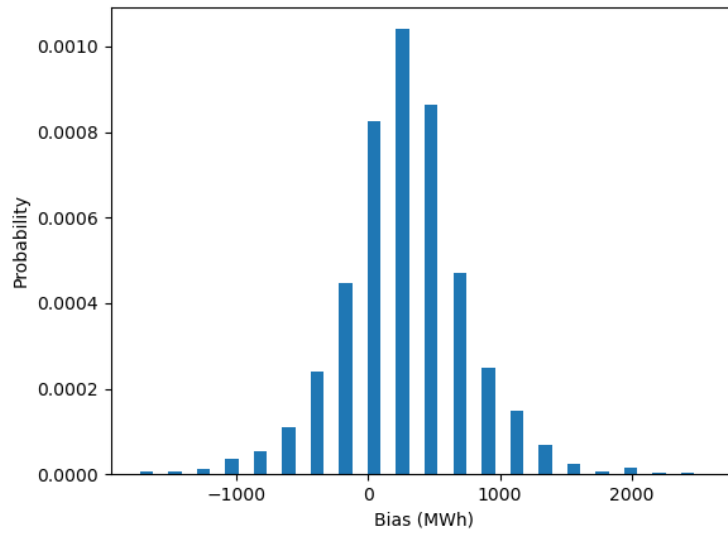
40-Σχήμα 4.29: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς

Α/Α ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ	ΤΥΠΟΣ ΣΦΑΛΜΑΤΟΣ		
	ΜΑΡΕ (%)	ΜΑΕ(MWh)	ΜΕ (MWh)
ARIMA	9,65	511,84	105,95
MLP	7,68	423,15	89,31
LSTM	5,92	314,77	64,64
ΟΡΙΖΟΝΤΑΣ ΠΡΟΒΛΕΨΗΣ	12h		
ΔΙΑΣΤΗΜΑ ΔΕΔΟΜΕΝΩΝ ΕΙΣΟΔΟΥ	168h		

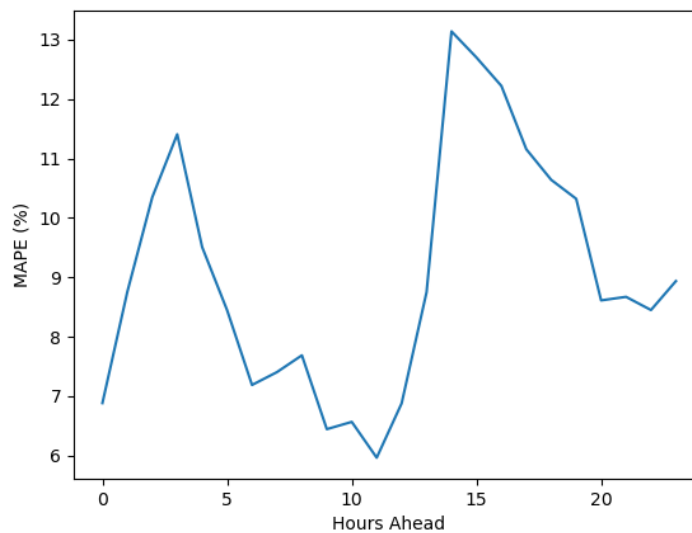
41-Σχήμα 4.30: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM

4.3.2.3 Πρόβλεψη επόμενων 24 ωρών

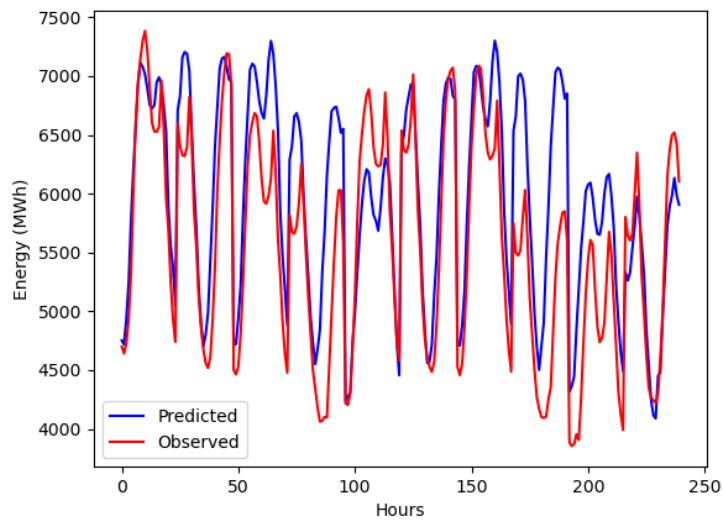
Για χρονικό διάστημα δεδομένων εισόδου στο μοντέλο πρόβλεψης 24 ώρες:



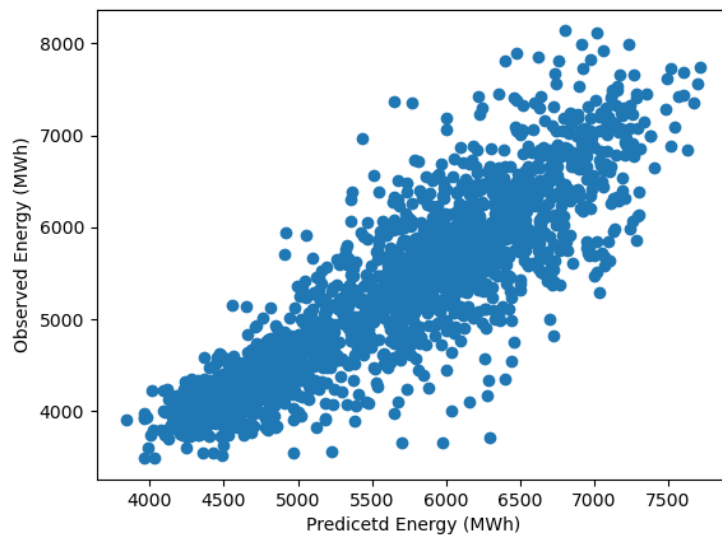
42-Σχήμα 4.31: Διάγραμμα κατανομής σφάλματος ME μοντέλου LSTM



43-Σχήμα 4.32: Διάγραμμα σφάλματος MAPE μοντέλου LSTM



44-Σχήμα 4.33: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM

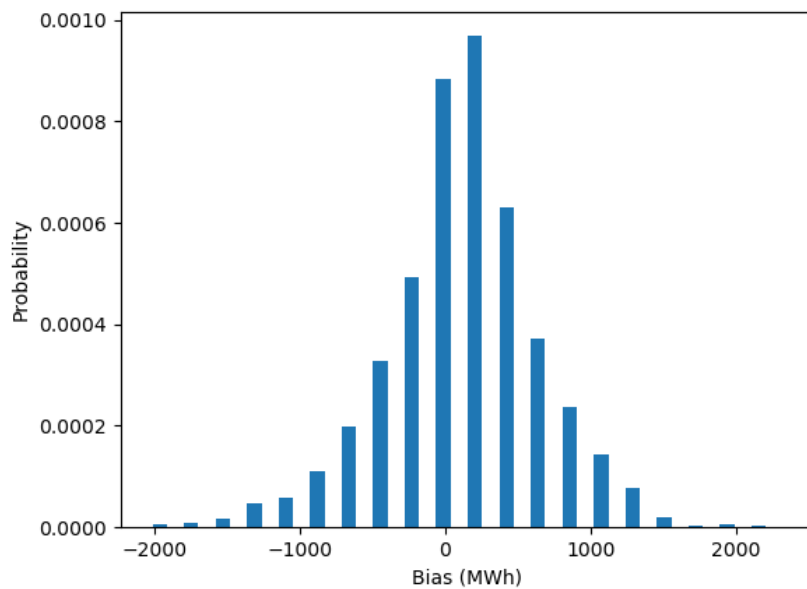


45-Σχήμα 4.34: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς

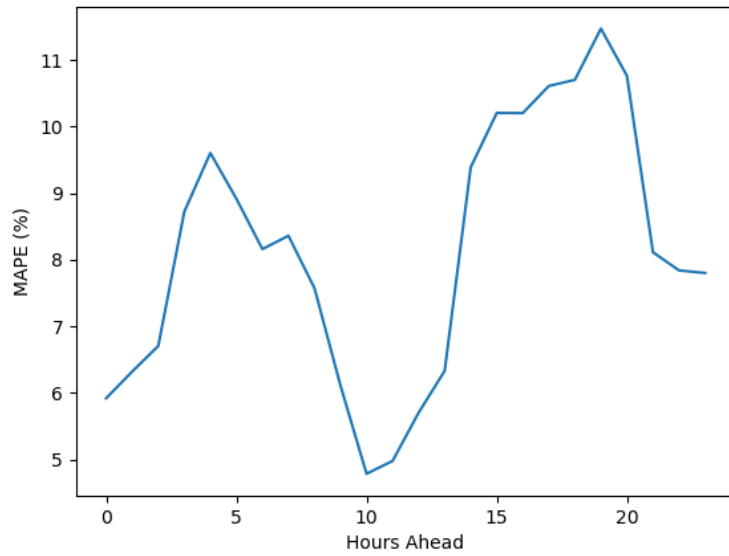
Α/Α ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ	ΤΥΠΟΣ ΣΦΑΛΜΑΤΟΣ		
	ΜΑΡΕ (%)	ΜΑΕ(MWh)	ΜΕ (MWh)
ARIMA	13,89	704,47	511,09
MLP	11,63	670,89	370,12
LSTM	9,05	459,69	333,11
ΟΡΙΖΟΝΤΑΣ ΠΡΟΒΛΕΨΗΣ	24h		
ΔΙΑΣΤΗΜΑ ΔΕΔΟΜΕΝΩΝ ΕΙΣΟΔΟΥ	24h		

46-Σχήμα 4.35: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM

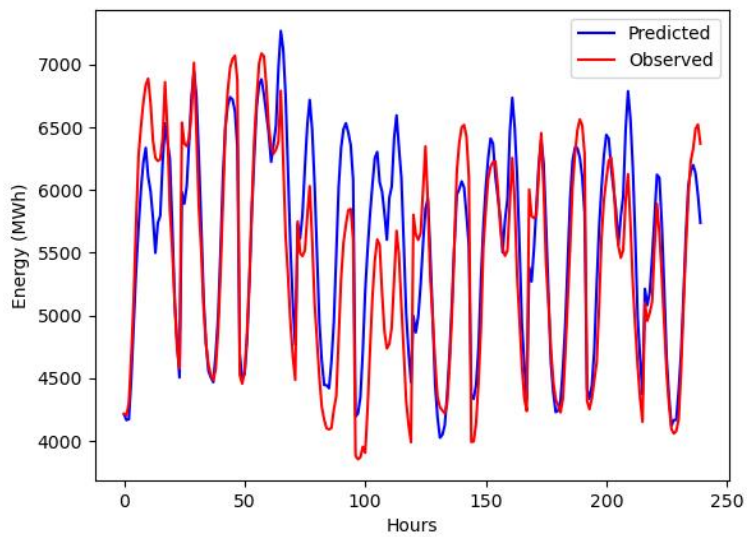
Για χρονικό διάστημα δεδομένων εισόδου στο μοντέλο πρόβλεψης 168 ώρες:



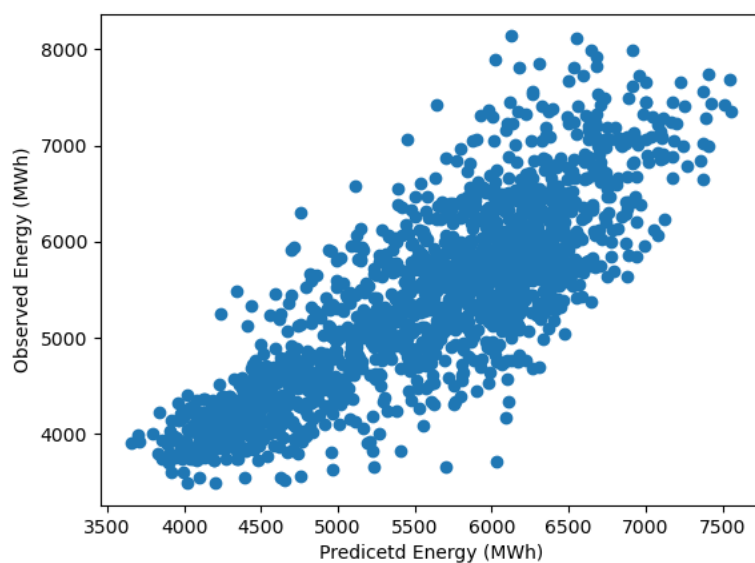
47-Σχήμα 4.36: Διάγραμμα κατανομής σφάλματος ΜΕ μοντέλου LSTM



48-Σχήμα 4.37: Διάγραμμα σφάλματος MAPE μοντέλου LSTM



49-Σχήμα 4.38: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας βάση μοντέλου LSTM



50-Σχήμα 4.39: Συσχετισμός μεταξύ παρατηρούμενων και προβλεπόμενων τιμών της χρονοσειράς

Α/Α	ΤΥΠΟΣ ΣΦΑΛΜΑΤΟΣ		
	ΜΑΡΕ (%)	ΜΑΕ(MWh)	ΜΕ (MWh)
ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ			
ARIMA	12,34	639,00	259,50
MLP	10,25	564,48	201,02
LSTM	8,14	426,16	172,99
ΟΡΙΖΟΝΤΑΣ ΠΡΟΒΛΕΨΗΣ	24h		
ΔΙΑΣΤΗΜΑ ΔΕΔΟΜΕΝΩΝ ΕΙΣΟΔΟΥ	168h		

51-Σχήμα 4.40: Διάγραμμα σύγκρισης αποτελεσμάτων μεταξύ βασικών μοντέλων και LSTM

ΚΕΦΑΛΑΙΟ 5: ΣΥΜΠΕΡΑΣΜΑΤΑ

5.1 Πρόβλεψη του φορτίου ηλεκτρικής ενέργειας της επόμενης ώρας (continuous 1h)

Παρατηρούμε από τα διαγράμματα της ενότητας 4.3.2.1 ότι η χρήση νευρωνικών δικτύων LSTM για την βραχυπρόθεσμη πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας φέρει καλύτερα αποτελέσματα από τους στατικούς αλγόριθμους τύπου ARIMA ή απλών νευρωνικών δικτύων όπως τα MLP για ορίζοντα πρόβλεψης της επόμενης ώρας. Το LSTM μοντέλο φέρει εξαιρετικές προβλέψεις για αυτόν τον ορίζοντα.

Για διάστημα δεδομένων εισόδου 24h:

Όπως φαίνεται στο σχήμα 4.13 η κατανομή του σφάλματος είναι συμμετρική γύρω από το μηδέν και επομένως, δεν υπάρχει θετική ή αρνητική ασυμμετρία στην κατανομή που να επηρεάζει το σφάλμα πρόβλεψης για το νευρωνικό LSTM.

Στα σχήματα 4.13 & 4.15 παρατηρούμε ότι οι τιμές πρόβλεψης σχεδόν ταυτίζονται με τις αντίστοιχες τιμές της χρονοσειράς, γεγονός που μας οδηγεί στο συμπέρασμα ότι το διάστημα δεδομένων εισόδου με τον αντίστοιχο ορίζοντα πρόβλεψης αποφέρουν άριστα αποτελέσματα.

Στο σχήμα 4.16 καλούμαστε να συγκρίνουμε για ίδιο διάστημα δεδομένων εισόδου και ορίζοντα πρόβλεψης τα δύο βασικά μοντέλα με το μοντέλο του LSTM, όπου:

- Το μοντέλο ARIMA ως στατιστικό μοντέλο προέβλεψε με σφάλμα $MAPE=4.13\%$, που εάν και αρκετά καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υποπρόβλεψης σαν μοντέλο καθώς η μέση τιμή της απόκλισης πραγματικών και προβλεπόμενων τιμών είναι αρνητική. $ME(MWh) = -45.15$
- Το μοντέλο MLP ως νευρωνικό μοντέλο (ANN) πρόβλεψε με σφάλμα $MAPE=2.89\%$, που εάν και αρκετά καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υπερπρόβλεψης καθώς το μέσο σφάλμα $ME(MWh) = 30.05$ είναι θετικό.
- Το μοντέλο LSTM ως νευρωνικό μοντέλο (RNN) πρόβλεψε με σφάλμα $MAPE=1.89\%$, που θεωρείται άριστη πρόβλεψη και παρουσιάζει μια συμπεριφορά συμμετρικής κατανομής όπως αναφέρθηκε παραπάνω καθώς η μέση τιμή της απόκλισης πραγματικών και προβλεπόμενων τιμών είναι σχεδόν μηδενική. $ME(MWh) = -0.46$

Για διάστημα δεδομένων εισόδου 168h:

Όπως φαίνεται στο σχήμα 4.17 η κατανομή του σφάλματος είναι συμμετρική γύρω από την μέση τιμή σφάλματος $ME = -30MWh$ και όχι στο μηδέν όπως για το διάστημα δεδομένων εισόδου 24h όμως το εύρος τιμών σφάλματος δεν υπερβαίνει τα +-

400MWh και στις δύο περιπτώσεις. Επομένως, υπάρχει αρνητική ασυμμετρία στην κατανομή που επηρεάζει το σφάλμα πρόβλεψης για το νευρωνικό LSTM.

Στα σχήματα 4.18 & 4.19 παρατηρούμε ότι οι τιμές πρόβλεψης σχεδόν ταυτίζονται με τις αντίστοιχες τιμές της χρονοσειράς, γεγονός που μας οδηγεί στο συμπέρασμα ότι το διάστημα δεδομένων εισόδου με τον αντίστοιχο ορίζοντα πρόβλεψης αποφέρουν άριστα αποτελέσματα.

Στο σχήμα 4.20 καλούμαστε να συγκρίνουμε για ίδιο διάστημα δεδομένων εισόδου και ορίζοντα πρόβλεψης τα δύο βασικά μοντέλα με το μοντέλο του LSTM, όπου:

- Το μοντέλο ARIMA ως στατιστικό μοντέλο πρόβλεψε με σφάλμα $MAPE=3.87\%$, που εάν και αρκετά καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υπερπρόβλεψης καθώς το μέσο σφάλμα είναι θετικό. $ME(MWh) = 50.81$
- Το μοντέλο MLP ως νευρωνικό μοντέλο (ANN) πρόβλεψε με σφάλμα $MAPE=2.54\%$, που εάν και αρκετά καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υποπρόβλεψης καθώς το μέσο σφάλμα $ME(MWh) = -40.15$ είναι αρνητικό.
- Το μοντέλο LSTM ως νευρωνικό μοντέλο (RNN) πρόβλεψε με σφάλμα $MAPE=1.92\%$, που θεωρείται άριστη πρόβλεψη και παρουσιάζει μια συμπεριφορά υποπρόβλεψης καθώς το μέσο σφάλμα $ME(MWh) = -26.33$ είναι αρνητικό.

Συμπεραίνουμε λοιπόν για την πρόβλεψη επόμενης ώρας ότι η αύξηση του διαστήματος δεδομένων εισόδου στα βασικά μας μοντέλα οδήγησε σε βελτίωση των αποτελεσμάτων καθώς είχαμε μικρότερα σφάλματα (MAPE) αλλά αυξήσεις στις αντίστοιχες μέσες τιμές σφάλματος (ME). Για το LSTM μοντέλο, με την αύξηση του διαστήματος δεδομένων εισόδου παρατηρήθηκε αύξηση του μέσου σφάλματος και μετατόπιση της κατανομής των σφαλμάτων αλλά το MAPE παρέμεινε το ίδιο εξαιρετικό. Επομένως η κατανομή σφαλμάτων πρέπει να είναι όσο το δυνατόν πιο συμμετρική και η μέση τιμή των σφαλμάτων να είναι μηδέν για να επιτύχουμε τις καλύτερες δυνατές προβλέψεις.

5.2 Πρόβλεψη του φορτίου ηλεκτρικής ενέργειας των επόμενων 12 ωρών της παρούσας ημέρας (intra day 12h)

Και εδώ παρατηρούμε από τα διαγράμματα της ενότητας 4.3.2.2 ότι η χρήση νευρωνικών δικτύων LSTM για την βραχυπρόθεσμη πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας φέρει καλύτερα αποτελέσματα από τους στατικούς αλγόριθμους

τύπου ARIMA ή απλών νευρωνικών δικτύων όπως τα MLP για ορίζοντα πρόβλεψης των επόμενων 12 ωρών της τρέχουσας ημέρας. Το LSTM μοντέλο φέρει αρκετά καλές προβλέψεις για αυτόν τον ορίζοντα.

Για διάστημα δεδομένων εισόδου 24h:

Όπως φαίνεται στο σχήμα 4.21 η κατανομή του σφάλματος είναι συμμετρική γύρω από την τιμή σφάλματος $ME=190\text{MWh}$ και επομένως, υπάρχει θετική ασυμμετρία στην κατανομή που επηρεάζει το σφάλμα πρόβλεψης για το νευρωνικό LSTM.

Στο σχήμα 4.22 παρατηρούμε ότι το σφάλμα MAPE είναι πολύ μικρότερο ,και επομένως έχουμε καλύτερη πρόβλεψη, για τις αρχικές δύο ώρες του 12ωρου και για τις 6η-12η ώρες του 12ωρου πρόβλεψης. Αυτό συμβαίνει διότι ο ορίζοντας πρόβλεψης 12ωρου αποτυπώνεται μεταξύ 12 το μεσημέρι και 12 τα μεσάνυχτα. Επομένως το σφάλμα αλλάζει ανάλογα των διακυμάνσεων του φορτίου , οι οποίες είναι απότομες και λιγότερο προβλέψιμες κατά τις μεσημεριανές και απογευματινές ώρες όπου τα φορτία από τις περιοχές εργασίας μετατοπίζονται στις περιοχές κατοικίας.

Στο σχήμα 4.23 παρατηρούμε ότι οι τιμές πρόβλεψης έχουν μια καθυστέρηση σε σχέση με τις πραγματικές τιμές της χρονοσειράς, γεγονός που μας οδηγεί στο συμπέρασμα ότι το διάστημα δεδομένων εισόδου με τον αντίστοιχο ορίζοντα πρόβλεψης αποφέρουν ικανοποιητικά αποτελέσματα όσον αφορά την περιοδικότητα αλλά όχι ταύτιση, ειδικά στα σημεία αλλαγής της τάσης όπως φαίνεται από την κυματομορφή.

Στο σχήμα 4.24 παρατηρούμε ότι η διασπορά μεταξύ των προβλέψεων και των πραγματικών τιμών της χρονοσειράς δεν ταυτίζεται με τον άξονα 45 μοιρών αλλά έχει μικρή διασπορά για μικρά φορτία και μεγαλύτερη όσο τα φορτία αυξάνονται.

Στο σχήμα 4.25 καλούμαστε να συγκρίνουμε για ίδιο διάστημα δεδομένων εισόδου και ορίζοντα πρόβλεψης τα δύο βασικά μοντέλα με το μοντέλο του LSTM, όπου:

- Το μοντέλο ARIMA ως στατιστικό μοντέλο πρόβλεψε με σφάλμα $MAPE=9.73\%$, που εάν και καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υποπρόβλεψης σαν μοντέλο καθώς το μέσο σφάλμα $ME(\text{MWh}) = -120.9$ είναι αρνητικό.
- Το μοντέλο MLP ως νευρωνικό μοντέλο (ANN) πρόβλεψε με σφάλμα $MAPE=8.25\%$, που εάν και καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υπερπρόβλεψης καθώς το μέσο σφάλμα $ME(\text{MWh}) = 150.38$ είναι θετικό.
- Το μοντέλο LSTM ως νευρωνικό μοντέλο (RNN) πρόβλεψε με σφάλμα $MAPE=6.34\%$, που θεωρείται αρκετά καλή πρόβλεψη και παρουσιάζει μια συμπεριφορά υπερπρόβλεψης καθώς το μέσο σφάλμα , $ME(\text{MWh}) = 190.26$, είναι θετικό.

Για διάστημα δεδομένων εισόδου 168h:

Όπως φαίνεται στο σχήμα 4.26 η κατανομή του σφάλματος είναι συμμετρική γύρω από την τιμή σφάλματος $ME=64.64$ MWh και επομένως, υπάρχει θετική ασυμμετρία στην κατανομή που επηρεάζει το σφάλμα πρόβλεψης για το νευρωνικό LSTM.

Στο σχήμα 4.27 παρατηρούμε ότι το σφάλμα MAPE είναι πολύ μικρότερο ,και επομένως έχουμε καλύτερη πρόβλεψη, για τις αρχικές δύο ώρες του 12ωρου και από την 9^η έως και την 12η ώρα του 12ωρου πρόβλεψης. Αυτό συμβαίνει διότι ο ορίζοντας πρόβλεψης 12ωρου αποτυπώνεται μεταξύ 12 το μεσημέρι και 12 τα μεσάνυχτα. Επομένως το σφάλμα αλλάζει ανάλογα των διακυμάνσεων του φορτίου , οι οποίες είναι απότομες και λιγότερο προβλέψιμες κατά τις μεσημεριανές και απογευματινές ώρες όπου τα φορτία από τις περιοχές εργασίας μετατοπίζονται στις περιοχές κατοικίας. Δεδομένου ότι εδώ είχαμε λάβει υπόψιν μας ώρες σε μέρες αργίας όπως Κυριακή και Σάββατο το σφάλμα επηρεάστηκε σε μεγαλύτερο διάστημα του 12ωρου.

Στο σχήμα 4.28 παρατηρούμε ότι οι τιμές πρόβλεψης έχουν μια καθυστέρηση σε σχέση με τις πραγματικές τιμές της χρονοσειράς, γεγονός που μας οδηγεί στο συμπέρασμα ότι το διάστημα δεδομένων εισόδου με τον αντίστοιχο ορίζοντα πρόβλεψης αποφέρουν ικανοποιητικά αποτελέσματα όσον αφορά την περιοδικότητα αλλά όχι ταύτιση, ειδικά στα σημεία αλλαγής της τάσης όπως φαίνεται από την κυματομορφή.

Στο σχήμα 4.29 παρατηρούμε ότι η διασπορά μεταξύ των προβλέψεων και των πραγματικών τιμών της χρονοσειράς δεν ταυτίζεται με τον άξονα 45 μοιρών αλλά έχει μικρή διασπορά για μικρά φορτία και μεγαλύτερη όσο τα φορτία αυξάνονται.

Στο σχήμα 4.30 καλούμαστε να συγκρίνουμε για ίδιο διάστημα δεδομένων εισόδου και ορίζοντα πρόβλεψης τα δύο βασικά μοντέλα με το μοντέλο του LSTM, όπου:

- Το μοντέλο ARIMA ως στατιστικό μοντέλο πρόβλεψε με σφάλμα $MAPE=9.65\%$, που εάν και καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υπερπρόβλεψης σαν μοντέλο καθώς το μέσο σφάλμα $ME(MWh) = 105.95$ είναι θετικό.
- Το μοντέλο MLP ως νευρωνικό μοντέλο (ANN) πρόβλεψε με σφάλμα $MAPE=8.25\%$, που εάν και καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υπερπρόβλεψης καθώς το μέσο σφάλμα $ME(MWh) = 89.31$ είναι θετικό.
- Το μοντέλο LSTM ως νευρωνικό μοντέλο (RNN) πρόβλεψε με σφάλμα $MAPE=5.92\%$, που θεωρείται αρκετά καλή πρόβλεψη και παρουσιάζει μια συμπεριφορά υπερπρόβλεψης και το μέσο σφάλμα είναι θετικό. $ME(MWh) = 64.64$.

Συμπεραίνουμε λοιπόν για την πρόβλεψη επόμενων 12 ωρών της ίδιας ημέρας ότι η αύξηση του διαστήματος δεδομένων εισόδου στα βασικά μας μοντέλα οδήγησε σε οριακή βελτιστοποίηση των αποτελεσμάτων καθώς είχαμε μικρότερα σφάλματα (MAPE) αλλά και μείωση στις αντίστοιχες μέσες τιμές σφάλματος (ME). Το LSTM μοντέλο σημείωσε οριακή βελτίωση του σφάλματος MAPE άρα και της πρόβλεψης αλλά μπόρεσε να μειώσει και την μέση τιμή σφάλματος (ME) που μας δείχνει ότι έκανε την κατανομή των μέσων σφαλμάτων πιο συμμετρική και με μικρότερη ασυμμετρία όπως φαίνεται κι από τα σχήματα. Επομένως η κατανομή σφαλμάτων πρέπει να είναι

όσο το δυνατόν πιο συμμετρική και η ασυμμετρία της κατανομής να είναι μηδέν για να επιτύχουμε τις καλύτερες δυνατές προβλέψεις.

5.3 Πρόβλεψη του φορτίου ηλεκτρικής ενέργειας των επόμενων 24 ωρών της επόμενης ημέρας (day ahead)

Τέλος παρατηρούμε από τα διαγράμματα της ενότητας 4.3.2.3 ότι η χρήση νευρωνικών δικτύων LSTM για την βραχυπρόθεσμη πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας φέρει καλύτερα αποτελέσματα από τους στατικούς αλγόριθμους τύπου ARIMA ή απλών νευρωνικών δικτύων όπως τα MLP για ορίζοντα πρόβλεψης των επόμενων 24 ωρών της επόμενης ημέρας (day ahead). Το LSTM μοντέλο φέρει αρκετά καλές προβλέψεις για αυτόν τον ορίζοντα.

Για διάστημα δεδομένων εισόδου 24h:

Όπως φαίνεται στο σχήμα 4.31 η κατανομή του σφάλματος είναι συμμετρική γύρω από το βάρος της κατανομής σφάλματος $ME=333.11MWh$ και επομένως, υπάρχει θετική «κλίση» στην κατανομή που επηρεάζει το σφάλμα πρόβλεψης για το νευρωνικό LSTM.

Στο σχήμα 4.32 παρατηρούμε ότι το σφάλμα MAPE είναι πολύ μικρότερο ,και επομένως έχουμε καλύτερη πρόβλεψη, για τις αρχικές δύο ώρες του 24ωρου, για τις 6^η έως 12^η ώρες του 24ωρου και για τις τελευταίες 20^η έως 24^η ώρες του 24ωρου πρόβλεψης. Αυτό συμβαίνει διότι ο ορίζοντας πρόβλεψης 24ωρου αποτυπώνεται μεταξύ 12 τα μεσάνυχτα της τρέχουσας ημέρας και 12 τα μεσάνυχτα της επόμενης ημέρας. Επομένως το σφάλμα αλλάζει ανάλογα των διακυμάνσεων του φορτίου , οι οποίες είναι απότομες και λιγότερο προβλέψιμες κατά τις μεσημεριανές, απογευματινές και τις πρώτες πρωινές ώρες όπου τα φορτία από τις περιοχές εργασίας μετατοπίζονται στις περιοχές κατοικίας και αντίστροφα.

Στο σχήμα 4.33 παρατηρούμε ότι οι τιμές πρόβλεψης έχουν μια καθυστέρηση σε σχέση με τις πραγματικές τιμές της χρονοσειράς, γεγονός που μας οδηγεί στο συμπέρασμα ότι το διάστημα δεδομένων εισόδου με τον αντίστοιχο ορίζοντα πρόβλεψης αποφέρουν ικανοποιητικά αποτελέσματα όσον αφορά την περιοδικότητα αλλά όχι ταύτιση αυτής, και αυτό εκδηλώνεται έντονα στα σημεία αλλαγής της τάσης-κλίσης, δηλαδή μέγιστα-ελάχιστα, όπως φαίνεται από την κυματομορφή.

Στο σχήμα 4.34 παρατηρούμε ότι η διασπορά μεταξύ των προβλέψεων και των πραγματικών τιμών της χρονοσειράς δεν ταυτίζεται με τον άξονα 45 μοιρών αλλά έχει μεγάλη διασπορά σε όλο το εύρος του φορτίου.

Στο σχήμα 4.35 καλούμαστε να συγκρίνουμε για ίδιο διάστημα δεδομένων εισόδου και ορίζοντα πρόβλεψης τα δύο βασικά μοντέλα με το μοντέλο του LSTM, όπου:

- Το μοντέλο ARIMA ως στατιστικό μοντέλο πρόβλεψε με σφάλμα MAPE=13.89% , που εάν και καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υπερπρόβλεψης σαν μοντέλο καθώς το μέσο σφάλμα ME(MWh) = 511.09 είναι θετικό.
- Το μοντέλο MLP ως νευρωνικό μοντέλο (ANN) πρόβλεψε με σφάλμα MAPE=11.63% , που εάν και καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υπερπρόβλεψης καθώς το μέσο σφάλμα ME(MWh) = 370.12 είναι θετικό.
- Το μοντέλο LSTM ως νευρωνικό μοντέλο (RNN) πρόβλεψε με σφάλμα MAPE=9.05% , που θεωρείται αρκετά καλή πρόβλεψη και παρουσιάζει μια συμπεριφορά υπερπρόβλεψης με μέσο σφάλμα ME(MWh) = 333.11 .

Για διάστημα δεδομένων εισόδου 168h:

Όπως φαίνεται στο σχήμα 4.36 η κατανομή του σφάλματος είναι συμμετρική γύρω από την τιμή σφάλματος ME=64.64 MWh και επομένως, υπάρχει θετική «κλίση»-βάρος στην κατανομή που επηρεάζει το σφάλμα πρόβλεψης για το νευρωνικό LSTM.

Στο σχήμα 4.37 παρατηρούμε ότι το σφάλμα MAPE είναι πολύ μικρότερο ,και επομένως έχουμε καλύτερη πρόβλεψη, για τις αρχικές δύο ώρες του 24ωρου, για τις 6η-12η ώρες του 24ωρου και για τις τελευταίες 20^η έως 24^η ώρες του 24ωρου πρόβλεψης. Αυτό συμβαίνει διότι ο ορίζοντας πρόβλεψης 24ωρου αποτυπώνεται μεταξύ 12 τα μεσάνυχτα της τρέχουσας ημέρας και 12 τα μεσάνυχτα της επόμενης ημέρας. Επομένως το σφάλμα αλλάζει ανάλογα των διακυμάνσεων του φορτίου , οι οποίες είναι απότομες και λιγότερο προβλέψιμες κατά τις μεσημεριανές, απογευματινές και τις πρώτες πρωινές ώρες όπου τα φορτία από τις περιοχές εργασίας μετατοπίζονται στις περιοχές κατοικίας και αντίστροφα.

Στο σχήμα 4.38 παρατηρούμε ότι οι τιμές πρόβλεψης έχουν μια καθυστέρηση σε σχέση με τις πραγματικές τιμές της χρονοσειράς, γεγονός που μας οδηγεί στο συμπέρασμα ότι το διάστημα δεδομένων εισόδου με τον αντίστοιχο ορίζοντα πρόβλεψης αποφέρουν ικανοποιητικά αποτελέσματα όσον αφορά την περιοδικότητα αλλά όχι ταύτιση αυτής, και αυτό εκδηλώνεται έντονα στα σημεία αλλαγής της τάσης-κλίσης, δηλαδή μέγιστα-ελάχιστα, όπως φαίνεται από την κυματομορφή.

Στο σχήμα 4.39 παρατηρούμε ότι η διασπορά μεταξύ των προβλέψεων και των πραγματικών τιμών της χρονοσειράς δεν ταυτίζεται με τον άξονα 45 μοιρών αλλά έχει μεγάλη διασπορά σε όλο το εύρος του φορτίου.

Στο σχήμα 4.40 καλούμαστε να συγκρίνουμε για ίδιο διάστημα δεδομένων εισόδου και ορίζοντα πρόβλεψης τα δύο βασικά μοντέλα με το μοντέλο του LSTM, όπου:

- Το μοντέλο ARIMA ως στατιστικό μοντέλο πρόβλεψε με σφάλμα MAPE=12.34% , που εάν και καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υπερπρόβλεψης σαν μοντέλο καθώς το μέσο σφάλμα ME(MWh) = 259.5 είναι θετικό.

- Το μοντέλο MLP ως νευρωνικό μοντέλο (ANN) πρόβλεψε με σφάλμα MAPE=10.25% , που εάν και καλή πρόβλεψη παρουσιάζει μια συμπεριφορά υπερπρόβλεψης καθώς το μέσο σφάλμα ME(MWh) = 201.02 είναι θετικό.
- Το μοντέλο LSTM ως νευρωνικό μοντέλο (RNN) πρόβλεψε με σφάλμα MAPE=8.14% , που θεωρείται αρκετά καλή πρόβλεψη και παρουσιάζει μια συμπεριφορά υπερπρόβλεψης με το μέσο σφάλμα να είναι θετικό ME(MWh) = 172.99 .

Συμπεραίνουμε λοιπόν για την πρόβλεψη επόμενων 24 ωρών της ίδιας ημέρας ότι η αύξηση του διαστήματος δεδομένων εισόδου στα βασικά μας μοντέλα οδήγησε σε οριακή βελτιστοποίηση των αποτελεσμάτων καθώς είχαμε μικρότερα σφάλματα (MAPE) αλλά και μείωση στις αντίστοιχες μέσες τιμές σφάλματος (ME). Το LSTM μοντέλο σημείωσε οριακή βελτίωση του σφάλματος MAPE άρα και της πρόβλεψης αλλά μπόρεσε να μειώσει και την μέση τιμή σφάλματος (ME) που μας δείχνει ότι έκανε την κατανομή των μέσων σφαλμάτων πιο συμμετρική και με μικρότερο βάρος όπως φαίνεται κι από τα σχήματα. Επομένως η κατανομή σφαλμάτων πρέπει να είναι όσο το δυνατόν πιο συμμετρική και η ασυμμετρία της κατανομής να είναι όσο το δυνατόν πιο κοντά στο μηδέν για να επιτύχουμε τις καλύτερες δυνατές προβλέψεις.

5.4 Γενικά συμπεράσματα

Παρατηρούμε από τα παραπάνω διαγράμματα ότι η χρήση νευρωνικών δικτύων LSTM για την βραχυπρόθεσμη πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας φέρει καλύτερα αποτελέσματα από τους στατικούς αλγόριθμους τύπου ARIMA ή απλών νευρωνικών δικτύων όπως τα MLP για ορίζοντα πρόβλεψης μια ώρα μπροστά καθώς όσο μικρότερος είναι ο ορίζοντας πρόβλεψης τόσο καλύτερα αποτελέσματα έχουμε. Δηλαδή, για ορίζοντες πρόβλεψης της τάξης της μία ώρας τα αποτελέσματα ήταν εξαιρετικά, και απείχαν πολύ από τα βασικά μοντέλα σύγκρισης. Για τις intraday προβλεψης , τις επόμενες 12 ώρες εντός της ημέρας, παρατηρήθηκε ότι η ακρίβεια των δύο βασικών μοντέλων και του LSTM ήταν επίσης αρκετά καλή με το LSTM να έχει 1-2% καλύτερη απόδοση από το MLP και 3-5% καλύτερη απόδοση από το ARIMA. Για την πρόβλεψη της επόμενης ημέρας παρατηρήθηκε ότι το LSTM δεν απέδωσε το ίδιο καλά αποτελέσματα με τα αποτελέσματα που είχε στους άλλους δύο ορίζοντες δεδομένου ότι στις day ahead προβλέψεις υπάρχει μια χρονική ασυνέχεια μεταξύ των διαθέσιμων δεδομένων εισόδου που συγκεντρώνονται μέχρι τις δώδεκα το μεσημέρι και τις πρώτης ώρας πρόβλεψης της επόμενης ημέρας που είναι 12 το βράδυ και οι επόμενες 24 ώρες.

Αξίζει να σημειωθεί ότι η παράμετρος των καιρικών συνθηκών δεν λήφθηκε υπόψιν καθώς το πρόβλημα που κληθήκαμε να επιλύσουμε, επικεντρώθηκε στην δυνατότητα πρόβλεψης του φορτίου ηλεκτρικής ενέργειας στην περίπτωση μην διαθεσιμότητας αριθμητικών προβλέψεων καιρού. Επομένως το πρόβλημα αντιμετωπίστηκε ως ένα

αμιγώς πρόβλημα πρόβλεψης χρονοσειράς χωρίς την χρήση εξωγενών μεταβλητών εισόδου.

Αποδείχθηκε λοιπόν ότι το νευρωνικό LSTM μπορεί να αποδώσει άριστες προβλέψεις για την επόμενη ώρα και πολύ καλές για τις ενδιάμεσες ώρες και για την επόμενη ημέρα. Επιπρόσθετα παρατηρήσαμε ότι το διάστημα δεδομένων εισόδου που εισάχθηκε κατά την εκπαίδευση του LSTM επηρέασε την ασυμμετρία των σφαλμάτων, το ME και το MAPE, τα οποία βελτιώθηκαν με την αύξηση του διαστήματος δεδομένων εισόδου.

Τέλος, ιδιαίτερη προσοχή αξίζει η ανάλυση της χρονοσειράς πριν την είσοδο της στο νευρωνικό μας δίκτυο. Καθώς η τάση, η εποχικότητα, η περιοδικότητα και τα μη γραμμικά στοιχεία παίζουν σημαντικό ρόλο στην κατανομή και το εύρος των σφαλμάτων πρόβλεψης καθώς ιδανικά χρειαζόμαστε κατανομή συμμετρική με το μικρότερο εύρος και μηδενικό βάρος για να επιτύχουμε την καλύτερη πρόβλεψη όπως φάνηκε και στα αποτελέσματά μας.

Βιβλιογραφία

- [1]L. Suganthi, Anand A. Samuel,Energy models for demand forecasting—A review,Renewable and Sustainable Energy Reviews,Volume 16, Issue 2,2012,Pages1223-1240,ISSN 1364-0321
- [2]Zhen Shao, Fu Chao, Shan-Lin Yang, Kai-Le Zhou,A review of the decomposition methodology for extracting and identifying the fluctuation characteristics in electricity demand forecasting,Renewable and Sustainable Energy Reviews,Volume 75,2017,Pages 123-136,ISSN 1364-0321,
- [3]Moncef Krarti, Mohammad Aldubyan,Review analysis of COVID-19 impact on electricity demand for residential buildings,Renewable and Sustainable Energy Reviews,Volume 143,2021,110888,ISSN 1364-0321,
- [4]Ching-Lai Hor, S. J. Watson and S. Majithia, "Analyzing the impact of weather variables on monthly electricity demand," in IEEE Transactions on Power Systems, vol. 20, no. 4, pp. 2078-2085, Nov. 2005,.
- [5]A. González-Briones, G. Hernández, J. M. Corchado, S. Omatu and M. S. Mohamad, "Machine Learning Models for Electricity Consumption Forecasting: A Review," 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 2019, pp. 1-6, doi: 10.1109/CAIS.2019.8769508.
- [6]Alvarez, Francisco Martinez, et al. "Energy time series forecasting based on pattern sequence similarity." IEEE Transactions on Knowledge and Data Engineering 23.8 (2010): 1230-1243.
- [7]A. González-Briones, G. Hernández, J. M. Corchado, S. Omatu and M. S. Mohamad, "Machine Learning Models for Electricity Consumption Forecasting: A Review," 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 2019, pp. 1-6.
- [8]M. Elsaraiti, G. Ali, H. Musbah, A. Merabet and T. Little, "Time Series Analysis of Electricity Consumption Forecasting Using ARIMA Model," 2021 IEEE Green Technologies Conference (GreenTech), Denver, CO, USA, 2021, pp. 259-262.
- [9]Mahia, Farjana, et al. "Forecasting electricity consumption using ARIMA model." 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI). IEEE, 2019.
- [10]Samuel, Arthur L. "Machine learning." The Technology Review 62.1 (1959): 42-45.

- [11]Mahesh, Batta. "Machine learning algorithms-a review." International Journal of Science and Research (IJSR).[Internet] 9 (2020): 381-386.
- [12]Rosenblatt, Frank. "Perceptron simulation experiments." Proceedings of the IRE 48.3 (1960): 301-309.
- [13]Ajit, Arohan, Koustav Acharya, and Abhishek Samanta. "A review of convolutional neural networks." 2020 international conference on emerging trends in information technology and engineering (ic-ETITE). IEEE, 2020.
- [14]Medsker, Larry R., and L. C. Jain. "Recurrent neural networks." Design and Applications 5 (2001): 64-67.
- [15]Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
- [16]Yao, Lirong, and Yazhuo Guan. "An improved LSTM structure for natural language processing." 2018 IEEE International Conference of Safety Produce Informatization (IICSPI). IEEE, 2018.
- [17]Kim, Nakyoung, Minkyung Kim, and Jun Kyun Choi. "LSTM based short-term electricity consumption forecast with daily load profile sequences." 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE). IEEE, 2018.
- [18]Bodur, Idris, Emre Celik, and Nihat Ozturk. "A Short-Term Load Demand Forecasting based on the Method of LSTM." 2021 10th International Conference on Renewable Energy Research and Application (ICRERA). IEEE, 2021.