



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ
ΚΑΙ ΔΙΟΙΚΗΣΗΣ

**Τεχνικές Μηχανικής και Βαθιάς Μάθησης για Βραχυπρόθεσμη
Πρόβλεψη Ζήτησης Ηλεκτρικής Ενέργειας:
Μια Συγκριτική Ανάλυση σε Χρονοσειρές Ευρωπαϊκών
Διαχειριστών Συστημάτων Μεταφοράς**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ιωάννης-Κωνσταντίνος Σεισόπουλος

Επιβλέπων : Ψαρράς Ιωάννης
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2022



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ
ΚΑΙ ΔΙΟΙΚΗΣΗΣ

**Τεχνικές Μηχανικής και Βαθιάς Μάθησης για Βραχυπρόθεσμη
Πρόβλεψη Ζήτησης Ηλεκτρικής Ενέργειας:
Μια Συγκριτική Ανάλυση σε Χρονοσειρές Ευρωπαϊκών
Διαχειριστών Συστημάτων Μεταφοράς**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ιωάννης-Κωνσταντίνος Σεισόπουλος

Επιβλέπων : Ψαρράς Ιωάννης
Καθηγητής Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 31^η Ιούλιος 2022.

.....
Ι. Ψαρράς

.....
Δ. Ασκούνης

.....
Χ. Δούκας

Αθήνα, Ιούλιος 2022

.....

Ιωάννης-Κωνσταντίνος Σεισόπουλος

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

Copyright © Ιωάννης-Κωνσταντίνος Σεισόπουλος, 2022

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

Η πρόβλεψη ηλεκτρικού φορτίου βραχυπρόθεσμου χρόνου (STLF) έχει αναδειχθεί τις τελευταίες δεκαετίες ως ένα από τα πιο βασικότερα πεδία έρευνας για την αποτελεσματική και αξιόπιστη λειτουργία των συστημάτων ηλεκτρικής ενέργειας (ΣΗΕ). Διαδραματίζει πολύ σημαντικό ρόλο στον τομέα του προγραμματισμού, της ανάλυσης ρίσκου, της ανάλυσης ροής φορτίου, του σχεδιασμού και της συντήρησης συστημάτων ηλεκτρικής ενέργειας, πόσο μάλλον στη συμμετοχή των ενεργειακών εταιρειών στις αγορές ενέργειας. Ειδικά με την εμφάνιση των έξυπνων δικτύων, η ανάγκη για ακριβή και αξιόπιστη εκτίμηση του ηλεκτρικού φορτίου είναι μεγαλύτερη από ποτέ. Ακόμα, ο προαναφερθέντας στόχος (STLF) έχει γίνει στις μέρες μας πιο δημοφιλής για τις ευρωπαϊκές χώρες καθώς η ενεργειακή κρίση έχει φτάσει σε πρωτοφανή επίπεδα με ποικίλες γεωπολιτικές προεκτάσεις. Οι αλγόριθμοι μηχανικής μάθησης (ML) μπορούν να προβλέψουν το ηλεκτρικό φορτίο, καθώς μπορούν να μοντελοποιήσουν μη γραμμικά προβλήματα πρόβλεψης χρονοσειρών μέσω μιας διαδικασίας εκπαίδευσης που αναλύει μοτίβα. Στην παρούσα Διπλωματική Εργασία διεξάγεται μια συγκριτική μελέτη τεχνικών Μηχανικής Μάθησης (ML) και Βαθιάς Μάθησης (DL), συγκεκριμένα, χρησιμοποιούνται τα Light Gradient Boosting Machine (LightGBM), Neural Basis Expansion Analysis for Time Series forecasting (N-Beats) and Χρονικά Συνελκτικά Δίκτυα (TCN), ενώ διερευνάται η επίδραση διαφόρων εξωτερικών μεταβλητών στη εν λόγω διαδικασία. Συγκεκριμένα εξετάζονται τέσσερις κατηγορίες χαρακτηριστικών, ηλεκτρικά φορτία, χρονικές μεταβλητές, τιμή ηλεκτρικού φορτίου και καιρός, τα οποία κωδικοποιούνται με διάφορους τρόπους. Στην πειραματική διαδικασία, χρησιμοποιούνται σύνολα δεδομένων φορτίου ωριαίας ανάλυσης για την Πορτογαλία, την Ισπανία και την Ελλάδα, οι αντίστοιχες τιμές του ηλεκτρικού φορτίου, καθώς και τοπικά καιρικά δεδομένα της Πορτογαλίας, σε μία προσπάθει πρόβλεψης του ηλεκτρικού φορτίου της επόμενης ημέρας. Τέλος, προτείνεται ένα υβριδικό μοντέλο συνδυάζοντας μοντέλα μηχανικής (ML) και βαθιάς μάθησης (DL), εφαρμόζοντας τεχνική Ensemble.

Λέξεις Κλειδιά: Μηχανική Μάθηση (ML), Βαθιά Μάθηση (DL), Πρόβλεψη Χρονοσειρών, Έξυπνα Δίκτυα, Πρόβλεψη Φορτίου Βραχυπρόθεσμου Χρόνου (STLF), N-Beats, LightGBM, Χρονικά Συνελκτικά Δίκτυα (TCN), Νευρωνικά Δίκτυα, Ensemble

Abstract

Electrical short-term load forecasting (STLF) has emerged as one of the most essential fields of research for efficient and reliable operation of power systems in the last few decades. It plays a very significant role in the field of scheduling, contingency analysis, load flow analysis, planning and maintenance of power systems, let alone in the participation of energy companies in the energy markets. Especially with the advent of smart grids, the need for fairly precise and highly reliable estimation of electricity load is greater than ever. Moreover, the aforementioned task (STLF) has become nowadays more popular for European countries as the energy crisis has reached an unprecedented peak with geopolitical extensions. Machine learning methods are well adapted to the nature of the electrical load, as they can model complicated nonlinear connections in Time Series Forecasting through a learning process containing historical data patterns. The presented Thesis conducts a comparative study of state-of-the-art Machine Learning (ML) and Deep Learning (DL) techniques, namely Light Gradient Boosting Machine (LightGBM), Neural Basis Expansion Analysis for Time Series forecasting and Temporal Convolutional Networks (TCN), while investigating the effect of various external variables in the process. Four categories of features, historical loads, temporal covariates, energy price and weather factors, are considered and utilized leveraging various encoding mechanisms. In the experimental studies, hourly-resolution load datasets from Portugal, Spain and Greece and the corresponding prices of load records, as well as the provincial weather data from Portugal are utilized, trying to effectively forecast the day-ahead load. Finally, a hybrid model is proposed, combining Machine (ML) and Deep Learning (DL) models, applying an Ensemble method.

Key Words: Machine Learning (ML), Deep Learning (DL), Time Series Forecasting, Smart Grid, Short Term Load Forecasting (STLF), N-Beats, LightGBM, Temporal Convolutional Networks (TCN), Neural Networks, Ensemble

Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω θερμά τον Καθηγητή Ιωάννη Ψαρρά που μου έδωσε την δυνατότητα να ασχοληθώ με το θέμα της πρόβλεψης ηλεκτρικού φορτίου. Πρόκειται για έναν ερευνητικό τομέα πολύ σημαντικό στον χώρο της ενέργειας αλλά και της οικονομίας μία χώρας. Ακόμα, θα ήθελα να ευχαριστήσω τα άλλα δύο μέλη της τριμελούς επιτροπής, τον Κ. Ασκούνη και Κ. Δούκα για την συμμετοχή τους σε αυτή.

Στην συνέχεια, θα ήθελα να ευχαριστήσω θερμά τον Κ. Σωτήρη Πελέκη , ο οποίος υπήρξε αρωγός κατά την εκπόνηση της παρούσας διπλωματικής εργασίας. Η καθοδήγηση του, οι συμβουλές του και ο χρόνος που αφιέρωσε για εμένα ήταν ανεκτίμητης αξίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου και τους φίλους μου για την στήριξη που μου παρείχαν κατά την διάρκεια των σπουδών μου.

Ιωάννης-Κων/νος Σεισόπουλος,
Αθήνα, 9^η Ιουλίου 2022

Περιεχόμενα

Περίληψη.....	6
Abstract	7
Κεφάλαιο 1 - Εισαγωγή.....	13
1.1 Επικαιρότητα και Κίνητρο Εργασίας	13
1.2 Σκοπός	14
1.3 Διάρθρωση Τόμου	15
Κεφάλαιο 2 - Θεωρητικό Μέρος.....	17
2.1 Υπόβαθρο	17
2.1.1 Σύστημα Ηλεκτρικής Ενέργειας (ΣΗΕ)	17
2.1.2 Smart Grid	17
2.1.3 Ανάγκη Πρόβλεψης Φορτίου	18
2.2 Πρόβλεψη Φορτίου	19
2.3 Βιβλιογραφική ανάλυση.....	20
2.4 LightGBM (Gradient Boosting Machine)	22
2.4.1 Gradient Boosting.....	22
2.4.2 Βασική Ιδέα του αλγορίθμου LightGBM.....	23
2.4.3 Χαρακτηριστικά αλγορίθμου LightGBM.....	23
2.5 N-BEATS (Neural Basis Expansion Analysis for Time Series forecasting)	24
2.5.1 Δομικά στοιχεία του N-BEATS	24
2.6 TCN (Temporal Convolutional Network).....	26
2.6.1 Διεσταλμένη Αιτιώδης Συνέλιξη.....	26
2.6.2 Πλεονεκτήματα και μειονεκτήματα των TCN	27
2.7 Μέθοδοι Ensemble	28
2.7.1 Averaging Ensemble	29
2.7.2 Bagging	29
2.7.3 Άλλες κατηγορίες Ensembling	29
Κεφάλαιο 3 - Πειραματική διαδικασία	31
3.1 Εισαγωγή	31
3.2 Συλλογή δεδομένων και εξαγωγή χαρακτηριστικών	31
3.2.1 Δεδομένα Φορτίου.....	31
3.2.2 Μετεωρολογικά Δεδομένα	32
3.2.3 Δεδομένα Τιμής Μεγαβατώρας (MWh).....	33
3.2.4 Άλλα Δεδομένα	33
3.2.5 Προεπεξεργασία Δεδομένων	34
3.2.5.1 Επεξεργασία Μετεωρολογικών Δεδομένων.....	35
3.2.5.2 Ελλειπούσες Τιμές (missing values)	35
3.2.5.3 Κανονικοποίηση.....	36
3.3 Ανάλυση δεδομένων.....	37
3.3.1 Χρονοσειρές Φορτίου.....	37
3.3.2 Χρονοσειρές Τιμής Μεγαβατώρας (MWh).....	43
3.3.3 Χρονοσειρές Καιρού Πορτογαλίας	44
3.4 Επιλογή Μοντέλων.....	46
3.5 Διαδικασία Εκπαίδευσης.....	46
3.5.1 LightGBM	47

3.5.2 N-Beats	47
3.5.3 TCN	48
3.5.4 Ensemble	49
3.6 Αξιολόγηση Μοντέλων	49
Κεφάλαιο 4 - Αποτελέσματα-Συμπεράσματα	51
4.1 Αξιολόγηση Μοντέλων	51
4.1.1 Αποτελέσματα του αλγορίθμου LightGBM.....	51
4.1.2 Αποτελέσματα της βαθιάς αρχιτεκτονικής N-BEATS.....	53
4.1.3 Αποτελέσματα της βαθιάς αρχιτεκτονικής TCN.....	55
4.2 Οπτικοποίηση αποτελεσμάτων, συγκριτική ανάλυση και υβριδικό Ensemble.....	57
4.2.1 Σύγκριση Αποτελεσμάτων στην Πορτογαλία	57
4.2.2 Σύγκριση Αποτελεσμάτων στην Ισπανία	59
4.2.3 Σύγκριση Αποτελεσμάτων στην Ελλάδα	61
4.3 Αποτελέσματα Συμμεταβλητής Καιρού.....	63
Κεφάλαιο 5	66
5.1 Σύνοψη Εργασίας και Επίλογος	66
5.2 Μελλοντικές Προεκτάσεις	67
Αναφορές.....	69

Λίστα Πινάκων

ΠΙΝΑΚΑΣ 1: ΠΑΡΑΓΟΜΕΝΕΣ ΧΡΟΝΟΣΕΙΡΕΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΧΡΟΝΟΥ ΚΑΙ ΕΡΓΑΣΙΜΗΣ ΗΜΕΡΑΣ (ΣΥΜΜΕΤΑΒΛΗΤΕΣ)	34
ΠΙΝΑΚΑΣ 2: ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΤΟΥ N-BEATS ΚΑΙ ΟΙ ΥΠΕΡΠΑΡΑΜΕΤΡΟΙ ΤΟΥ	48
ΠΙΝΑΚΑΣ 3: ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΤΟΥ TCN ΚΑΙ ΟΙ ΥΠΕΡΠΑΡΑΜΕΤΡΟΙ ΤΟΥ	48
ΠΙΝΑΚΑΣ 4: Η ΑΡΧΙΚΟΠΟΙΗΣΗ ΤΩΝ ΒΑΡΩΝ ΣΤΗ ΔΙΑΔΙΚΑΣΙΑ SOFT ENSEMBLE	49
ΠΙΝΑΚΑΣ 5: ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΕΙΡΑΜΑΤΩΝ LIGHTGBM	52
ΠΙΝΑΚΑΣ 6: ΑΠΟΤΕΛΕΣΜΑΤΑ N-BEATS	54
ΠΙΝΑΚΑΣ 7: ΑΠΟΤΕΛΕΣΜΑΤΑ ΜΟΝΤΕΛΟΥ TCN	56
ΠΙΝΑΚΑΣ 8: ΤΕΛΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ HYBRID AVERAGING ENSEMBLE (LIGHTGBM+N-BEATS)	63
ΠΙΝΑΚΑΣ 9: ΑΠΟΤΕΛΕΣΜΑΤΑ ΧΡΗΣΗΣ ΣΥΜΜΕΤΑΒΛΗΤΗ ΚΑΙΡΟΥ ΣΤΗΝ ΠΟΡΤΟΓΑΛΙΑ	64

Λίστα Εικόνων

ΕΙΚ. 1: ΑΛΛΑΓΕΣ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ ΗΛΕΚΤΡΙΚΗΣ ΕΝΕΡΓΕΙΑΣ	18
ΕΙΚ. 2: ΑΝΑΠΤΥΞΗ ΔΕΝΤΡΟΥ LEAF-WISE	24
ΕΙΚ. 3: ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΤΟΥ N-BEATS [36]	25
ΕΙΚ. 4: ΔΙΚΤΥΟ ΤΡΙΩΝ ΕΠΙΠΕΔΩΝ ΜΕ ΑΙΤΙΩΔΗ ΣΥΝΕΛΙΞΗ ΧΩΡΙΣ ΔΙΑΣΤΟΛΗ [39]	26
ΕΙΚ. 5: ΔΙΚΤΥΟ ΤΡΙΩΝ ΕΠΙΠΕΔΩΝ ΜΕ ΑΙΤΙΩΔΗ ΣΥΝΕΛΙΞΗ ΜΕ ΧΡΗΣΗ ΔΙΑΣΤΟΛΗΣ [39]	27
ΕΙΚ. 6: ΤΥΠΙΚΗ ΔΙΑΤΑΞΗ ENSEMBLE	28
ΕΙΚ. 7: ΠΟΡΤΟΓΑΛΙΑ - ΣΗΜΕΙΑ ΠΡΟΒΛΕΨΗΣ ΚΑΙΡΟΥ	33
ΕΙΚ. 8: ΙΣΤΟΓΡΑΜΜΑ ΩΡΙΑΙΟΥ ΦΟΡΤΙΟΥ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΠΟΡΤΟΓΑΛΙΑ (κWh)	38
ΕΙΚ. 9: ΙΣΤΟΓΡΑΜΜΑ ΩΡΙΑΙΟΥ ΦΟΡΤΙΟΥ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΙΣΠΑΝΙΑ (κWh)	38
ΕΙΚ. 10: ΙΣΤΟΓΡΑΜΜΑ ΩΡΙΑΙΟΥ ΦΟΡΤΙΟΥ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΕΛΛΑΔΑ (κWh)	39
ΕΙΚ. 11: ΜΗΝΙΑΙΟ ΦΟΡΤΙΟ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΠΟΡΤΟΓΑΛΙΑ (κWh)	40
ΕΙΚ. 12: ΜΗΝΙΑΙΟ ΦΟΡΤΙΟ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΙΣΠΑΝΙΑ (κWh)	40
ΕΙΚ. 13: ΜΗΝΙΑΙΟ ΦΟΡΤΙΟ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΕΛΛΑΔΑ (κWh)	41
ΕΙΚ. 14: ΜΕΣΟ ΗΜΕΡΗΣΙΟ ΠΡΟΦΙΛ ΦΟΡΤΙΟΥ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΠΟΡΤΟΓΑΛΙΑ (κWh)	41
ΕΙΚ. 15: ΜΕΣΟ ΗΜΕΡΗΣΙΟ ΠΡΟΦΙΛ ΦΟΡΤΙΟΥ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΙΣΠΑΝΙΑ (κWh)	42
ΕΙΚ. 16: ΜΕΣΟ ΗΜΕΡΗΣΙΟ ΠΡΟΦΙΛ ΦΟΡΤΙΟΥ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΕΛΛΑΔΑ (κWh)	42
ΕΙΚ. 17: ΜΕΣΗ ΜΗΝΙΑΙΑ ΤΙΜΗ ΜΕΓΑΒΑΤΩΡΑΣ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΠΟΡΤΟΓΑΛΙΑ (ΕΥΡΩ/MWh)	43
ΕΙΚ. 18: ΜΕΣΗ ΜΗΝΙΑΙΑ ΤΙΜΗ ΜΕΓΑΒΑΤΩΡΑΣ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΙΣΠΑΝΙΑ (ΕΥΡΩ/MWh)	44
ΕΙΚ. 19: ΜΕΣΗ ΜΗΝΙΑΙΑ ΤΙΜΗ ΜΕΓΑΒΑΤΩΡΑΣ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΕΛΛΑΔΑ (ΕΥΡΩ/MWh)	44
ΕΙΚ. 20: ΜΕΣΗ ΜΗΝΙΑΙΑ ΘΕΡΜΟΚΡΑΣΙΑ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΛΙΣΑΒΟΝΑ (C)	45
ΕΙΚ. 21: ΜΕΣΗ ΜΗΝΙΑΙΑ ΤΑΧΥΤΗΤΑ ΑΝΕΜΟΥ ΓΙΑ ΤΑ 2019-2021 ΣΤΗΝ ΛΙΣΑΒΟΝΑ (10 m/s)	45
ΕΙΚ. 22: ΣΥΓΚΡΙΣΗ ΜΟΝΤΕΛΩΝ ΣΤΗΝ ΧΡΟΝΟΣΕΙΡΑ ΤΗΣ ΠΟΡΤΟΓΑΛΙΑΣ	58
ΕΙΚ. 23: ΣΤΙΓΜΙΟΤΥΠΟ ΠΡΟΒΛΕΨΗΣ ΠΟΡΤΟΓΑΛΙΑ (1)	58

ΕΙΚ. 24: ΣΤΙΓΜΙΟΤΥΠΟ ΠΡΟΒΛΕΨΗΣ ΠΟΡΤΟΓΑΛΙΑ (2)	59
ΕΙΚ. 25: ΣΥΓΚΡΙΣΗ ΜΟΝΤΕΛΩΝ ΣΤΗΝ ΧΡΟΝΟΣΕΙΡΑ ΤΗΣ ΙΣΠΑΝΙΑΣ	60
ΕΙΚ. 26: ΣΤΙΓΜΙΟΤΥΠΟ ΠΡΟΒΛΕΨΗΣ ΙΣΠΑΝΙΑ (1).....	60
ΕΙΚ. 27: ΣΤΙΓΜΙΟΤΥΠΟ ΠΡΟΒΛΕΨΗΣ ΙΣΠΑΝΙΑ (2).....	61
ΕΙΚ. 28: ΣΥΓΚΡΙΣΗ ΜΟΝΤΕΛΩΝ ΣΤΗΝ ΧΡΟΝΟΣΕΙΡΑ ΤΗΣ ΕΛΛΑΔΑΣ	61
ΕΙΚ. 29: ΣΤΙΓΜΙΟΤΥΠΟ ΠΡΟΒΛΕΨΗΣ ΕΛΛΑΔΑ (1).....	62
ΕΙΚ. 30: ΣΤΙΓΜΙΟΤΥΠΟ ΠΡΟΒΛΕΨΗΣ ΕΛΛΑΔΑ (2).....	62
ΕΙΚ. 31: ΣΥΓΚΡΙΣΗ ΜΟΝΤΕΛΩΝ ΜΕ ΚΑΙ ΧΩΡΙΣ ΤΗΝ ΣΥΜΜΕΤΑΒΛΗΤΗ ΤΟΥ ΚΑΙΡΟΥ	65

Κεφάλαιο 1 - Εισαγωγή

1.1 Επικαιρότητα και Κίνητρο Εργασίας

Ο ενεργειακός τομέας βρίσκεται επί του παρόντος σε μια ταχεία μετάβαση που δημιουργεί πολλές προκλήσεις και ευκαιρίες, καθώς το μίγμα των χρησιμοποιούμενων καυσίμων αλλάζει και η ψηφιοποίηση και χρήση εξελιγμένων τεχνολογικών μέσων για όλο τον κύκλο ζωής του είναι επιτακτικής ανάγκης. Παρότι πρόκειται για έναν από τους παραδοσιακούς τομείς, ο ενεργειακός τομέας δείχνει να έχει μεγάλες δυνατότητες για ενσωμάτωση ψηφιακών τεχνολογιών. Δύο κύριοι παράγοντες δείχνουν ότι η ψηφιοποίηση είναι ένα ζωτικό βήμα για την αύξηση του ελέγχου και της αξιοπιστίας των ενεργειακών δικτύων. Ο πρώτος είναι η μετάβαση από τα ορυκτά καύσιμα σε ενέργεια από ανανεώσιμες πηγές (ΑΠΕ), κάτι που δημιουργεί ένα πιο δυναμικά μεταβαλλόμενο, στοχαστικό και ασταθές μοτίβο παραγωγής ηλεκτρικής ενέργειας. Ο δεύτερος παράγοντας είναι η αυξανόμενη ζήτηση των καταναλωτών λόγω του εξηλεκτισμού της καθημερινής ζωής, δηλαδή, της αύξησης των ηλεκτροκίνητων μεταφορών (κυρίως μέσω της αυξανόμενης διείσδυσης των ηλεκτρικών οχημάτων), του πλουραλισμού των οικιακών συσκευών αλλά και της αύξησης χρήσης της ηλεκτρικής ενέργειας για θέρμανση, ψύξη και κλιματισμό, κάτι που δημιουργεί μεταβαλλόμενα πρότυπα στην κατανάλωση. Το εν λόγω γεγονός εντάθηκε ακόμη περισσότερο με την έλευση της πανδημίας του κορονοϊού που προκάλεσε αρκετές μεταβολές στην κατανάλωση ενέργειας που επέφεραν τα κατά τόπους μέτρα περιορισμού της κυκλοφορίας αλλά και η αύξηση της τηλεργασίας.

Η κατανάλωση ηλεκτρικής ενέργειας επηρεάζεται ιστορικά από τις καιρικές συνθήκες, τις ανθρώπινες δραστηριότητες, την οικονομική ανάπτυξη καθώς και από διάφορους προσωρινούς και εποχικούς παράγοντες. Οι παράγοντες αυτοί καθιστούν την πρόβλεψή του φορτίου ηλεκτρικής ενέργειας μία άκρως πολύπλοκη διαδικασία. Οι αυξανόμενες ανάγκες των καταναλωτών συνεπάγονται υψηλότερη κατανάλωση ενέργειας που με την σειρά της οδηγεί σε υψηλότερες αιχμές κατανάλωσης. Στην Πορτογαλία και Ισπανία, η εκτεταμένη χρήση ηλεκτρικής ενέργειας κατά τους χειμερινούς μήνες λόγω ύπαρξης μηχανημάτων θέρμανσης συντελούν στην αλλοίωση του ημερήσιου προφίλ κατανάλωσης. Στην Ελλάδα ειδικότερα, η αυξημένη συχνότητα και διάρκεια των επεισοδίων καύσωνα οδηγεί σε συνεχή αύξηση της χρήσης ηλεκτρικών μηχανημάτων ψύξης με αποτέλεσμα αρκετά αλλοιωμένα προφίλ φορτίου κατά την θερινή περίοδο με όλο και πιο απότομες αιχμές κατανάλωσης. Ακόμα, η πλευρά της προσφοράς (παραγωγή) μπορεί να αντιμετωπίζει δυσκολίες στην κάλυψη της αιχμής της ζήτησης, ιδίως με την ολοένα αυξανόμενη διείσδυση ανανεώσιμων πηγών, εκτός εάν υπάρχει σύστημα που προβλέπει το επίπεδο, τον χρόνο εμφάνισης και την διάρκεια αυτής της αιχμής. Δεδομένου ότι είναι δαπανηρό για τις εταιρείες να αυξήσουν την ποσότητα παραγωγής, έχουν εισαχθεί μέθοδοι ανάπτυξης νέων υπολογιστικών εργαλείων για την επίτευξη ακριβούς προγραμματισμού παραγωγής ηλεκτρικής ενέργειας και διαχείρισης των ενεργειακών πόρων.

Ένας από τους συνηθέστερους τρόπους αντιμετώπισης αυτών των προκλήσεων είναι η δημιουργία της κατάλληλης ψηφιακής υποδομής που θα μπορούσε να χρησιμοποιηθεί για την ανάπτυξη προγνωστικών μοντέλων. Αυτά τα μοντέλα μπορούν να ωφελήσουν παραγωγούς και παρόχους ηλεκτρικής ενέργειας παρέχοντάς τους την δυνατότητα ακριβέστερου προγραμματισμού της παραγωγής και σχεδιασμού των συστημάτων. Η πρόβλεψη της ζήτησης ενέργειας είναι ένας από τους πιο δημοφιλείς τρόπους σχεδιασμού της παραγωγής και κατά συνέπεια εξάλειψης των διακοπών ηλεκτροδότησης. Επίσης η όσο το δυνατόν ακριβέστερη πρόβλεψη είναι επιθυμητή για χρήση στις αγοραπωλησίες ηλεκτρικής ενέργειας (χρηματιστήριο ενέργειας), στο εμπόριο ρύπων και στην βελτιστοποίηση των ηλεκτρικών διασυνδέσεων μεταξύ συστημάτων ή χωρών. Έτσι καθίσταται απαραίτητο να εντοπιστούν μοτίβα ζήτησης τα οποία να ικανοποιούνται με όσο μεγαλύτερη ακρίβεια γίνεται.

Συνοψίζοντας, σε μία χρονική περίοδο που οικονομικές και γεωπολιτικές αναταράξεις ταλανίζουν την εγγύς ευρωπαϊκή και όχι μόνο οικογένεια, ταυτόχρονα με τον όλο και αυξανόμενο εξηλεκτρισμό της καθημερινότητας και την διείδυση ανανεώσιμων πηγών ενέργειας, προσπάθειες ανάλυσης και πρόβλεψης φορτίου καθίστανται αναγκαίες τόσο για την εύρυθμη λειτουργία των συστημάτων διανομής ενέργειας όσο και για σχεδιασμό της επόμενης ημέρας στον κλάδο της ενέργειας.

1.2 Σκοπός

Σκοπός της παρούσα διπλωματική είναι η πρόβλεψη του ηλεκτρικού φορτίου βραχυπρόθεσμου χρόνου (STLF) ευρωπαϊκών χωρών, καθώς και η ανάλυση εξωγενών μεταβλητών που επηρεάζουν το φορτίο τους, αξιοποιώντας την τεχνητή νοημοσύνη και συγκεκριμένα την μηχανική (ML) και βαθιά μάθηση (DL). Συγκεκριμένα γίνεται εκτεταμένη χρήση του μοντέλου μηχανικής μάθησης (ML), LightGBM (Light Gradient Boosting Machine) της οικογένειας των δέντρων αποφάσεων, καθώς και των μοντέλων βαθιάς μάθησης (DL), N-Beats (Neural Basis Expansion Analysis for Time Series forecasting) και TCN (Temporal Convolutional Networks). Τα αποτελέσματα αυτής της διπλωματικής εργασίας μπορούν να χρησιμοποιηθούν από τον διαχειριστή συστήματος ή από εταιρίες παραγωγής ενέργειας που συμμετέχουν στην αγορά ενέργειας για την βελτιστοποίηση διάφορων εργασιών τους και την στοχοθεσία τους. Για την επίτευξη του σκοπού αυτού έγιναν συγκεκριμένα βήματα-στόχοι.

Πρώτος στόχος είναι η κατανόηση του προβλήματος πρόβλεψη του ηλεκτρικού φορτίου βραχυπρόθεσμου χρόνου (STLF), οι προεκτάσεις του και οι εφαρμογές του. Δεύτερος στόχος είναι η περάτωση του κύκλου ζωής της μηχανικής μάθησης (ML life-cycle), δηλαδή η συλλογή και επεξεργασία των δεδομένων, η επιλογή των μοντέλων πρόβλεψης, εκπαίδευση και η αξιολόγηση τους. Τρίτος στόχος είναι η διεξαγωγή πειραμάτων και η εξαγωγή συμπερασμάτων.

Στα πλαίσια του τρίτου στόχου και την διεξαγωγή πειραμάτων ο τελικός σκοπός είναι η πρόβλεψη φορτίου της επόμενης μέρας για τις Πορτογαλία, Ισπανία και Ελλάδα, τρεις Ευρωπαϊκές χώρες της Μεσογείου με παρόμοια χαρακτηριστικά κλίματος. Προς επίρρωση του στόχου αυτού, διευκρινίστηκαν οι αποδοτικότερες αρχιτεκτονικές των μοντέλων βαθιάς (DL) και μηχανικής μάθησης (ML) σύμφωνα με την επίδοσή τους και την βιωσιμότητά τους και έγινε βελτιστοποίηση των υπερ παραμέτρων τους. Επίσης, αναλύθηκε εάν και κατά πόσο εξωτερικές μεταβλητές, όπως μετεωρολογικά δεδομένα, δεδομένα τιμής μεγαβατώρας και χρονικές μεταβλητές, επηρεάζουν το μοντέλο και μπορούν να συνεισφέρουν στην απόδοση και βιωσιμότητα των μοντέλων.

Μέσω του πειραματισμού γίνεται προσπάθεια προσδιορισμού και ερμηνείας των:

- Η μορφή του συνόλου των δεδομένων φορτίου που τροφοδοτείται στα μοντέλα, κανονικοποιημένο ή στις πραγματικές τιμές του.
- Το μήκος εισόδου των παραθύρων προηγούμενης γνώσης (lookback window) των μοντέλων.
- Η επίδραση συμμεταβλητών στην επίδοση των μοντέλων. Συγκεκριμένα, δοκιμάζονται δεδομένα τιμής μεγαβατώρας και χρονικές μεταβλητές καθώς και διάφορες κωδικοποιήσεις τους. Επιπλέον, δοκιμάζονται μετεωρολογικά ιστορικά δεδομένα, όπως και δεδομένα προβλέψεων καιρού για την Πορτογαλία. Ακόμα προσδιορίζεται το παράθυρο προηγούμενης γνώσης ή γνώσης μέλλοντος για τις ανωτέρω εξωγενείς μεταβλητές.
- Για τα μοντέλα βαθιάς μάθησης (DL) δοκιμάζονται διάφορες αρχιτεκτονικές ανάλογα με το βάθος τους.
- Η εφαρμογή της τεχνικής Ensemble για το μοντέλο μηχανικής (ML) και βαθιάς μάθησης (DL).

Τα αρχεία κώδικα που έγιναν τα πειράματα της παρούσας εργασία μπορούν να βρεθούν στο αποθετήριο “Thesis_Short_Term_Load_Forecasting”¹.

1.3 Διάρθρωση Τόμου

Στο παρόν Κεφάλαιο 1 βρίσκεται η εισαγωγή, ο σκοπός της παρούσας εργασίας καθώς και η διάρθρωση του τόμου.

Στο Κεφάλαιο 2 - Θεωρητικό Μέρος, αναλύεται διεξοδικά το πρόβλημα πρόβλεψης φορτίου βραχυπρόθεσμου χρόνου (STLF) το οποίο είναι κύριο αντικείμενο της παρούσας εργασίας. Συγκεκριμένα, στην σύγχρονη ευρωπαϊκή αγορά ενέργειας που το χρηματιστήριο ενέργειας είναι πραγματικότητα και στόχος της επόμενης μέρας, και η παραγωγή της επόμενης μέρας είναι κρίσιμης σημασίας παράγοντας για την εύρυθμη λειτουργία τόσο του συστήματος διανομής όσο και για τις ανάγκες παραγωγής, το πρόβλημα πρόβλεψης φορτίου της επόμενης μέρας έχει ποικίλες εφαρμογές τόσο από την πλευρά της στοχοθεσίας των εταιριών παραγωγής ενέργειας όσο και για τους διαχειριστές του συστήματος. Τα προαναφερθέντα σε συνδυασμό με τους στόχους της Ευρωπαϊκής Ένωσης για περαιτέρω διείσδυση ανανεώσιμων πηγών ενέργειας είτε μέσω net metering και την εκπροσώπηση των οικιακών παραγωγών στο χρηματιστήριο είτε μέσω εταιρειών παραγωγής με αιολικά πάρκα και την εκπροσώπηση τους, καθώς και συμβατικές μορφές ενέργειας όπως υδροηλεκτρικά και όχι μόνο εργοστάσια καθιστούν το πρόβλημα ζήτησης φορτίου γενικότερα, και ειδικότερα το πρόβλημα πρόβλεψης φορτίου βραχυπρόθεσμου χρόνου (STLF) ήσσονος σημασίας.

Στη συνέχεια γίνεται βιβλιογραφική ανασκόπηση για την επίλυση του προβλήματος πρόβλεψης φορτίου βραχυπρόθεσμου χρόνου (STLF) και παρουσιάζονται θεωρητικά τα μοντέλα μηχανικής (ML), LightGBM και βαθιάς μάθησης (DL), N-Beats και TCN που χρησιμοποιούνται για την εν λόγω πρόβλεψη.

Στο Κεφάλαιο 3 - Πειραματική διαδικασία περιγράφεται η μεθοδολογία που ακολουθήθηκε για την επίλυση του προβλήματος πρόβλεψης φορτίου βραχυπρόθεσμου χρόνου (STLF). Συγκεκριμένα αναλύθηκαν τα στάδια του κύκλου ζωής της μηχανικής μάθησης, δηλαδή, η συλλογή, η προ επεξεργασία και ανάλυση των δεδομένων, καθώς και η διαλογή, η εκπαίδευση και η αξιολόγηση των μοντέλων που επιλέχθηκαν. Η παραπάνω διαδικασία έγινε για την ωριαία πρόβλεψη του φορτίου τριών ευρωπαϊκών χωρών με μεσογειακό κλίμα, την Πορτογαλία, Ισπανία και Ελλάδα. Στα πλαίσια του σκοπού αυτού, χρησιμοποιήθηκαν εξωτερικές μεταβλητές όπως η τιμή της μεγαβατώρας, χρονικές μεταβλητές και μεταβλητές καιρού (μόνο στην Πορτογαλία). Χρησιμοποιήθηκαν δύο μοντέλα βαθιάς μάθησης (DL), N-Beats (Neural Basis Expansion Analysis for Time Series forecasting) και TCN (Temporal Convolutional Networks) και ένα μοντέλο μηχανικής μάθησης (ML) LightGBM (Gradient Boosting Machine).

Στο Κεφάλαιο 4 - Αποτελέσματα-Συμπεράσματα παρουσιάζονται τα αποτελέσματα και συμπεράσματα των πειραμάτων που διενεργήθηκαν. Αρχικά, αναλύονται τα αποτελέσματα των μοντέλων με τις επιδόσεις τους και γίνεται προσπάθεια ερμηνείας τους, ενώ στη συνέχεια γίνεται σύγκριση των αποτελεσμάτων των μοντελων για την κάθε χώρα. Τόσο τα σύνολα δεδομένων φορτίου της Πορτογαλίας, Ισπανία και Ελλάδας όσο και τα σύνολα των συμμεταβλητών παρελθοντικών ή μελλοντικών που χρησιμοποιήθηκαν (καιρός, τιμή μεγαβατώρας και χρονικές μεταβλητές), αποτελούνται από ωριαία δεδομένα από 01/03/2016 έως και 28/02/2022 (6 έτη).

Τέλος, στο Κεφάλαιο 5 γίνεται σύνοψη της εργασίας και παρουσιάζονται οι μελλοντικές προεκτάσεις της.

¹ https://github.com/giannisiso/Thesis_Short_Term_Load_Forecasting

Κεφάλαιο 2 - Θεωρητικό Μέρος

2.1 Υπόβαθρο

2.1.1 Σύστημα Ηλεκτρικής Ενέργειας (ΣΗΕ)

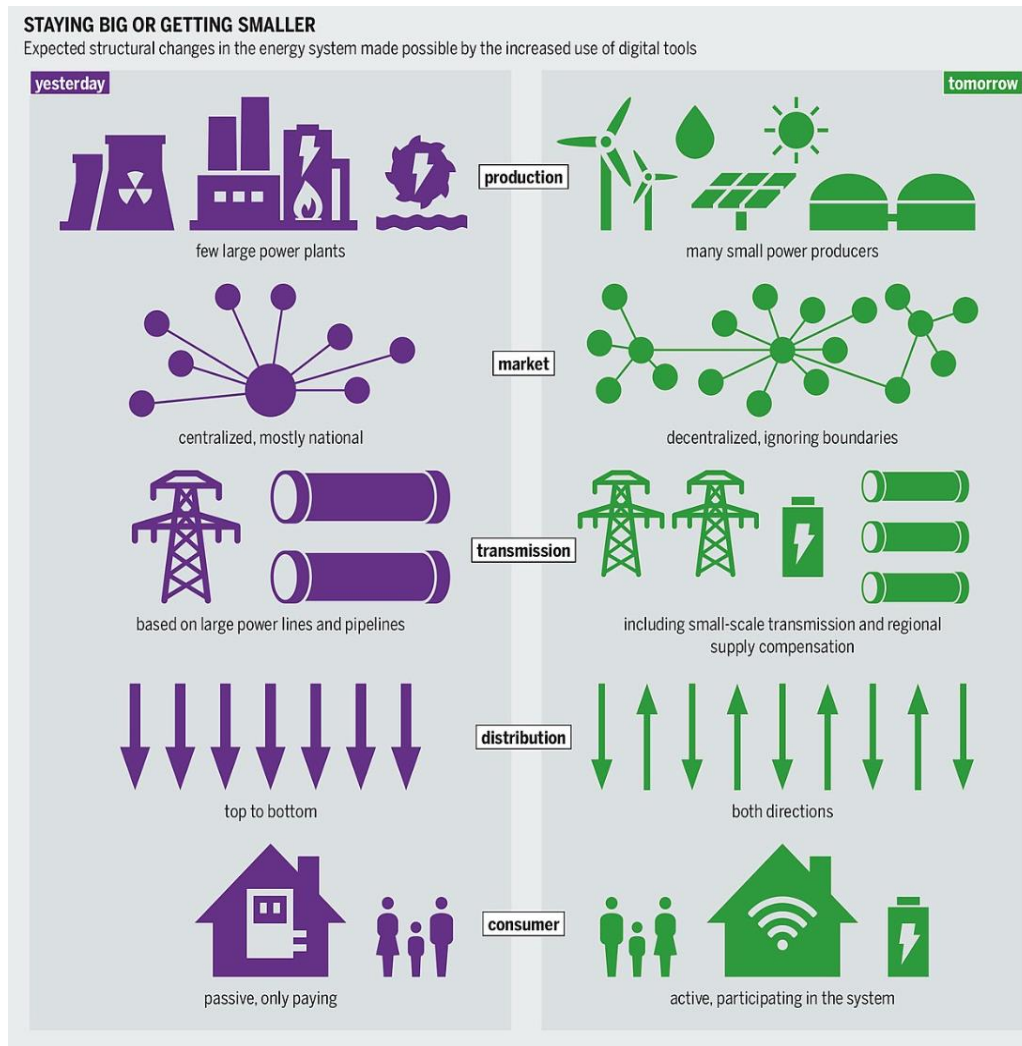
Σύγχρονο σύστημα ηλεκτρικής ενέργειας (ΣΗΕ) καλείται το σύνολο όλων εκείνων των εγκαταστάσεων που χρησιμοποιούνται για να τροφοδοτήσουν με ασφαλές, αξιόπιστο και φιλικό προς το περιβάλλον τρόπο, ηλεκτρική ενέργεια σε ένα σύνολο καταναλωτών. Οι βασικές λειτουργίες ενός τέτοιου συστήματος είναι η παραγωγή της ηλεκτρικής ενέργειας, η μεταφορά της μέσω καλωδίων υψηλής και μεσαίας τάσης και η διανομή της στους καταναλωτές. Η ηλεκτρική ενέργεια καθώς έχει φτάσει στους καταναλωτές μετατρέπεται σε άλλες μορφές προκειμένου να χρησιμοποιηθεί κατάλληλα για την κάλυψη των αναγκών της ηλεκτρικής εγκατάστασης.

Είναι προφανές πως στις μέρες μας οι απαιτήσεις στη βιομηχανία της ηλεκτρικής ενέργειας συνεχώς αλλάζουν και η ανάγκη για την δημιουργία και την εξέλιξη των συστημάτων εκείνων που θα ανταποκρίνονται σε αυτές είναι επιτακτική ανάγκη. Οι απαιτήσεις αυτές απορρέουν από το όραμα για ένα οικονομικό σύστημα, φιλικό προς το περιβάλλον, το οποίο θα χρησιμοποιεί στο μεγαλύτερο δυνατό τις ανανεώσιμες πηγές ενέργειας. Ο συνδυασμός των υφιστάμενων μέσων αλλά και των εξελίξεων που παρατηρούνται στον τομέα των υπολογιστών και των επικοινωνιών οδηγούν στην μετεξέλιξη του υπάρχοντος δικτύου σε ένα αποκεντρωμένο τοπικό δίκτυο στα πρότυπα ενός μικροδικτύου ή μιας ενεργειακής κοινότητας.

2.1.2 Smart Grid

Το Electric Power Research Institute (EPRI), ορίζει ως έξυπνο δίκτυο (smart grid [1]) εκείνο το σύστημα που υποστηρίζεται από τις τελευταίες τεχνολογίες στους τομείς της πληροφορίας, των υπολογιστών και της ηλεκτρονικής και το οποίο παράγει και μεταφέρει με αποτελεσματικούς, αξιόπιστους και φιλικούς προς το περιβάλλον τρόπους ενέργεια χαμηλού κόστους προκειμένου να ενισχύει την αγορά. Όπως είναι εύκολα αντιληπτό από τον παραπάνω ορισμό το έξυπνο δίκτυο (smart grid) είναι η μετεξέλιξη του υπάρχοντος συστήματος ηλεκτρικής ενέργειας στο οποίο η τεχνολογία της πληροφορίας και των υπολογιστών θα παίζουν πρωταγωνιστικό ρόλο. Ουσιαστικά με την χρήση αυτών των τεχνολογιών θα είναι δυνατόν να πραγματοποιείται σε πραγματικό χρόνο ο απομακρυσμένος έλεγχος του δικτύου αλλά και η καλύτερη διαχείριση της ενέργειας. Επιπρόσθετα με την αμφίδρομη επικοινωνία μεταξύ παραγωγού και καταναλωτή θα επιτυγχάνεται με μεγάλη ακρίβεια η εκτίμηση του ηλεκτρικού φορτίου που απαιτείται στο δίκτυο κάθε χρονική στιγμή με αποτέλεσμα η αποθήκευση της ενέργειας να γίνεται πιο αποδοτική αλλά και η εισαγωγή των σταθμών ανανεώσιμων πηγών ενέργειας (ΑΠΕ) στη παραγωγή ως διεσπαρμένη να πραγματοποιείται με αποτελεσματικότερο τρόπο.

Στην Εικ. 1 απεικονίζονται οι αλλαγές του συστήματος ηλεκτρικής ενέργειας τα επόμενα χρόνια.



Εικ. 1: Αλλαγές του συστήματος ηλεκτρικής ενέργειας

2.1.3 Ανάγκη Πρόβλεψης Φορτίου

Η πρόβλεψη του φορτίου ηλεκτρικής ενέργειας είναι ενέργεια ήσσονος σημασίας για κάθε χώρα καθώς συμβάλλει στην ορθή λειτουργία και ανάπτυξη του συστήματος ηλεκτρικής ενέργειας (ΣΗΕ). Οι προβλέψεις φορτίου τόσο της ζήτησης όσο και της παραγωγής οδηγεί σε προσδιορισμό της επάρκειας ενέργειας μιας χώρας και συντελεί στον προγραμματισμό της επόμενης μέρας, τον σχεδιασμό του δικτύου. Πρόσφατα, οι γεωπολιτικές αναταράξεις στην Ευρωπαϊκή ήπειρο που έχουν οδηγήσει σε ενεργειακή κρίση καθιστούν την πρόβλεψη φορτίου επιτακτική ανάγκη. Σύμφωνα με το ινστιτούτο ερευνών ασφαλείας της Ευρωπαϊκής Ένωσης, η ισορροπία ανάμεσα στην ασφάλεια και βιωσιμότητα του ενεργειακού τομέα έχει οδηγήσει το σύνολο των κρατών μελών σε λήψη μέτρων για την καταπολέμηση της κλιματικής αλλαγής και την εξασφάλιση της επάρκειας. Η ενσωμάτωση έξυπνων δικτύων (smart grids) και η διείσδυση ολοένα και αυξανόμενης ενέργειας από ανανεώσιμες πηγές (ΑΠΕ) είναι οδοδείκτης για την αντιμετώπιση της ενεργειακής κρίσης [2].

Πιο συγκεκριμένα, η πρόβλεψη φορτίου επιφέρει πλεονεκτήματα στους καταναλωτές. Οι καταναλωτές μέσω μιας ακριβούς πρόβλεψη φορτίου του κτιρίου τους, έχουν την δυνατότητα να ρυθμίζουν και να μειώνουν την κατανάλωση ενέργειας που δαπανούν. Ακόμα, έχοντας εντοπίσει τις περιόδους εκείνες της μέρας που παρουσιάζεται μεγάλη κατανάλωση, ο ιδιοκτήτης έχει την δυνατότητα να αναπροσαρμόσει την ενεργειακή συμπεριφορά του κτιρίου προγραμματίζοντας τις έξυπνες συσκευές της εγκατάστασης να λειτουργούν σε ώρες που η ζήτηση ενέργειας είναι χαμηλή, με πληθώρα οικονομικών οφελών. Παράλληλα, υπάρχουν σημαντικά πλεονεκτήματα και για το

δίκτυο. Δυο από τα σημαντικότερα είναι ότι το δίκτυο μπορεί σε πραγματικό χρόνο να έχει μια αξιόπιστη εκτίμηση του φορτίου που θα πρέπει να διαχειρίζεται ενώ παράλληλα έχει και την δυνατότητα να εντοπίζει τις ώρες εκείνες της μέρας όπου παρουσιάζεται η μεγαλύτερη κατανάλωση. Η δυνατότητα αυτή επιτρέπει στο δίκτυο να παράγει επαρκή ποσότητα ενέργειας χωρίς να υπάρχουν απώλειες. Επιπρόσθετα, ένα ακόμα σημαντικό στοιχείο άξιο αναφοράς είναι ότι το δίκτυο μπορεί, όταν αυτό είναι εφικτό να χρησιμοποιεί για την παραγωγή ενέργειας από ανανεώσιμες πηγές.

Τέλος, όπως γίνεται εύκολα αντιληπτό με όσα έχουν ήδη προαναφερθεί, η ανάγκη ύπαρξης αλγορίθμων που θα πετυχαίνουν ικανοποιητική πρόβλεψη ηλεκτρικού φορτίου είναι ένα ανοιχτό και συνεχώς εξελίξιμο επιστημονικό αντικείμενο.

2.2 Πρόβλεψη Φορτίου

Η πρόβλεψη φορτίου (Load Forecasting - LF) αποτελεί στόχος για οποιονδήποτε οργανισμό ασχολείται με το σύστημα ηλεκτρικής ενέργειας. Η πρόβλεψη της ζήτησης ηλεκτρικής ενέργειας διακρίνεται σε τέσσερις κύριες κατηγορίες ανάλογα με τον ορίζοντα πρόβλεψης:

- 1) Πρόβλεψη φορτίου Υπέρ-βραχυπρόθεσμου χρόνου (Very Short-Term Load Forecasting (VSTLF)). Αφορά ορίζοντα πρόβλεψης από λίγα λεπτά μέχρι λίγες ώρες. Η χρησιμότητα αυτού του είδους πρόβλεψης έγκειται σε real time εφαρμογές, όπως για παράδειγμα η λειτουργία και η συντήρηση, η διάγνωση σφαλμάτων και η αυτοθεραπεία (selfhealing) και γενικότερα η πιο ασφαλής και εύρωστη λειτουργία του ηλεκτρικού δικτύου. Σε συνδυασμό με το σύστημα SCADA ή και OnLine συστήματα παρακολούθησης και τηλεχειρισμών μπορούν να αποφευχθούν προβλήματα που προκύπτουν και να υπάρχει μια πιο ομαλή λειτουργία του ηλεκτρικού δικτύου. Κάποια από τα προβλήματα που συνεισφέρει στην επίλυσή τους είναι η πτώση τάσης και ρύθμιση συχνότητας
- 2) Πρόβλεψη φορτίου Βραχυπρόθεσμου χρόνου (Short-Term Load Forecasting (STLF)). Αναφέρεται σε ορίζοντες πρόβλεψης από μία μέρα έως μία εβδομάδα με δεδομένα από τετάρτου μέχρι ωριαίας ανάλυσης. Οι εφαρμογές εδώ αφορούν περισσότερο στον σχεδιασμό του συστήματος διανομής και μεταφοράς, στην χάραξη στρατηγικής όσον αφορά την προμήθεια ενέργειας αλλά και σε εφαρμογές που αγγίζουν τον καταναλωτή όπως απόκριση ζήτησης και έλεγχος της αποθήκευσης ενέργειας. Μία ακριβής τεχνική πρόβλεψη φορτίου βραχυπρόθεσμου χρόνου (STLF) μπορεί να μετριάσει το λειτουργικό κόστος, να κρατήσει αποτελεσματικές τις αγορές ενέργειας και να βοηθήσει στην καλύτερη κατανόηση της δυναμικής του συστήματος. Αντίθετα, μια λανθασμένη πρόβλεψη μπορεί να προκαλέσει είτε υπερεκτίμηση φορτίου, η οποία οδηγεί σε υπερβολική προσφορά και αποθεματικό ενέργειας και κατά συνέπεια, περισσότερες δαπάνες και απορρύθμιση της αγοράς ενέργειας είτε σε υπερεκτίμηση της παραγωγής ή οποία μπορεί να οδηγήσει και σε διακοπή της λειτουργίας του δικτύου
- 3) Πρόβλεψη φορτίου Μεσοπρόθεσμου χρόνου (Medium-Term Load Forecasting (MTLF)). Σχετίζεται με ορίζοντες πρόβλεψης από μία εβδομάδα μέχρι ένα χρόνο. Αυτή η πρόβλεψη φορτίου έχει πολλές εφαρμογές οικονομικού σχεδιασμού, όπως προγραμματισμό συντήρησης, μεσοπρόθεσμο υδροθερμικό συντονισμό, αξιολόγηση της επάρκειας,

διαχείριση περιορισμένων ενεργειακών μονάδων, διαπραγμάτευση προθεσμιακών συμβάσεων και ανάπτυξη στρατηγικών προμηθειών καυσίμων με αποτελεσματικό κόστος.

- 4) Πρόβλεψη φορτίου Μακροπρόθεσμου χρόνου (Long-Term Load Forecasting (LTLF)). Ο ορίζοντας πρόβλεψης σε αυτές τις εφαρμογές είναι έως είκοσι έτη και εφαρμόζεται κυρίως για στρατηγικό σχεδιασμό ανάπτυξης του δικτύου ενέργειας.

Η πρόβλεψη φορτίου Βραχυπρόθεσμου χρόνου με χρήση τεχνικών μηχανικής (ML) και βαθιάς μάθησης (DL) είναι ο σκοπός αυτής της εργασίας.

2.3 Βιβλιογραφική ανάλυση

Για την πρόβλεψη του ηλεκτρικού φορτίου (LF) και ιδιαίτερα για πρόβλεψη φορτίου βραχυπρόθεσμου χρόνου (STLF) έχει προταθεί ποικιλία μεθόδων ανά τα χρόνια. Ωστόσο, η επιστημονική κοινότητα δεν έχει καταλήξει σε σαφή συμπεράσματα για την ανωτερότητα κάποιων τεχνικών σε σχέση με άλλων, καθώς διαφορετικές οικογένειες μοντέλων, διαφορετικής πολυπλοκότητας, έχουν επιδείξει ανθεκτικότητα και καλή προβλεπτική ικανότητα. Τα προβλεπτικά μοντέλα μπορούν να χωριστούν σε δύο μεγάλες κατηγορίες. Η πρώτη κατηγορία αφορά συμβατικές στατιστικές μεθόδους και βασικά μοντέλα μηχανικής μάθησης (ML), ενώ η δεύτερη περιλαμβάνει state of the art τεχνικές μηχανικής (ML) και βαθιάς (DL) μάθησης, όπως και υβριδικά σχήματα.

Όσον αφορά στην πρώτη κατηγορία, πληθώρα στατιστικών μεθόδων έχουν δοκιμαστεί στο πρόβλημα πρόβλεψης φορτίου βραχυπρόθεσμου χρόνου (STLF). Από τις πρώτες οικογένειες μοντέλων που χρησιμοποιήθηκαν ήταν τα μοντέλα γραμμικής παλινδρόμησης [3], [4]. Ακόμα, έχουν χρησιμοποιηθεί μοντέλα Αυτοπαλινδρόμησης Κυλιόμενου Μέσου (ARMA) [5], τα οποία σύντομα αντικαταστάθηκαν από τα μοντέλα ARIMA [6] και SARIMA (Seasonal) [7], [8], για να αντιμετωπιστεί η χρονική και εποχιακή διακύμανση που συχνά παρουσιάζει το προφίλ του φορτίου. Παρόλα αυτά, δημιουργήθηκε η ανάγκη τροφοδοσίας των μοντέλων εξωτερικών μεταβλητών όπως καιρικά δεδομένα, τιμή μεγαβατώρας και χρονικές μεταβλητές, που τα προαναφερθέντα μοντέλα δεν έχουν την ικανότητα να συμπεριλάβουν. Για αυτό το λόγο, οι παραπάνω οικογένειες μοντέλων επεκτάθηκαν στις ARMAX [9] και ARIMAX [10]. Το κύριο χαρακτηριστικό των προβλεπτικών αυτών μοντέλων είναι η θεώρηση του παρατηρούμενου συστήματος ως γραμμικό, υπόθεση που δεν ισχύει γενικώς, καθιστώντας το κύριο μειονέκτημά του. Άλλες στατιστικές τεχνικές που έχουν επίσης χρησιμοποιηθεί για την πρόβλεψη φορτίου βραχυπρόθεσμου χρόνου (STLF) είναι η μέθοδος Box-Jenkins [11] και η μέθοδος της εκθετικής εξομάλυνσης [12].

Προκειμένου να επιλυθεί αυτός ο περιορισμός, προτάθηκαν μη γραμμικά μοντέλα όπως τα Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης (Feed-Forward Neural Networks) τα οποία βρήκαν σημαντική χρήση σε σενάρια που παρουσιάζουν σημαντική μη γραμμικότητα, όπως σε εργασίες πρόβλεψης φορτίων [13], η οποία εργασία παρουσιάζει λεπτομερή ανασκόπηση σχετικά με τις εφαρμογές τεχνητών νευρωνικών δικτύων σε βραχυπρόθεσμη πρόβλεψη φορτίου (STLF). Η χρήση των Νευρωνικών Δικτύων Πρόσθιας Τροφοδότησης σε βραχυπρόθεσμη πρόβλεψη φορτίου χρονολογείται από την δεκαετία του '90. Οι συγγραφείς στο [14] προτείνουν ένα "ρηγό" νευρωνικό δίκτυο με ένα μόνο κρυφό στρώμα το οποίο επιχείρησε 24ωρη πρόβλεψη χρησιμοποιώντας την πληροφορία τόσο του φορτίου όσο και της θερμοκρασίας. Στο [15] οι συγγραφείς παρουσιάζουν ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης για την πρόβλεψη ηλεκτρικών φορτίων σε εβδομαδιαία βάση. Η αραιά συνδεδεμένη αρχιτεκτονική πρόσθιας τροφοδότησης λαμβάνει τις χρονοσειρές του φορτίου, τις μετρήσεις της θερμοκρασίας καθώς και την ώρα και την ημέρα της εβδομάδας. Αποδεικνύεται ότι οι επιπλέον πληροφορίες βελτιώνουν την ακρίβεια των προβλέψεων σε σύγκριση

με ένα μοντέλο ARIMA που έχει εκπαιδευτεί για την ίδια εργασία. Στο [16] παρουσιάζεται ένα από τα πρώτα δίκτυα πολλών επιπέδων που προέβλεπαν το ωριαίο φορτίο ενός συστήματος ηλεκτρικής ενέργειας.

Επίσης, άλλες οικογένειες μοντέλων που έχουν χρησιμοποιηθεί για την επίλυση του προβλήματος πρόβλεψης φορτίου βραχυπρόθεσμου χρόνου (STLF) και ανήκουν στην πρώτη κατηγορία είναι οι μηχανές υποστήριξης διανυσμάτων (SVM) [17, 18], ή οι αυτοοργανωμένοι χάρτες (SOM) [19]. Ακόμα, εκτενής χρήση μοντέλων της οικογένειας δέντρων αποφάσεων έχει γίνει για την επίλυση του προβλήματος της παρούσας εργασίας. Συγκεκριμένα, έχουν χρησιμοποιηθεί δέντρα παλινδρόμησης [20], random forest [21], gradient boosting regression trees [22]. Τέλος, άλλες επιπλέον στατιστικές-πιθανοτικές τεχνικές που έχουν χρησιμοποιηθεί είναι οι Kalman Filtering [23], και Γκαουσιανές διαδικασίες (GP) [24].

Όσον αφορά στη δεύτερη κατηγορία που αναφέρθηκε, state-of-the-art τεχνικές, έχουν επιχειρηθεί για την επίλυση του προβλήματος πρόβλεψης φορτίου βραχυπρόθεσμου χρόνου (STLF), μοντέλα μηχανικής (ML) και βαθιάς μάθησης (DL) καθώς και συνδυασμός τους (Ensemble). Σε τέτοιες state of the art τεχνικές συμπεριλαμβάνονται προηγμένες αρχιτεκτονικές πρόσθιας τροφοδότησης με υπολειμματικές συνδέσεις έως συνελκτικές προσεγγίσεις και ανατροφοδοτούμενα νευρωνικά δίκτυα μαζί με πολλές παραλλαγές τους. Επιπλέον, έχουν προταθεί υβριδικές αρχιτεκτονικές με στόχο την καταγραφή των χρονικών εξαρτήσεων στα δεδομένα με ανατροφοδοτούμενα δίκτυα, ενώ εκτελούν και γενικότερη λειτουργία εξαγωγής χαρακτηριστικών με χρήση συνελκτικών επιπέδων. Στο [25] ένα ανατροφοδοτούμενο νευρωνικό δίκτυο χρησιμοποιήθηκε για να παρέχει ωριαίες προβλέψεις φορτίου. Η μελέτη συγκρίνει επίσης την απόδοση του δικτύου όταν τροφοδοτείται με πρόσθετες πληροφορίες που αφορούν τον καιρό, όπως η θερμοκρασία και η υγρασία. Οι συγγραφείς συμπεραίνουν, όπως αναμενόταν, πως το ανατροφοδοτούμενο δίκτυο επωφελείται από τα πολλαπλά δεδομένα εισόδου και, ειδικότερα, τα σχετιζόμενα με τον καιρό.

Επίσης, μία ακόμα κατηγορία state-of-the-art μοντέλων είναι τα Long Short Term Memory (LSTM), τα οποία υιοθετήθηκαν σε βραχυπρόθεσμες προβλέψεις φορτίου (STLF), αποδεικνύοντας ότι είναι πιο αποτελεσματικά από τις παραδοσιακές μεθόδους ανάλυσης χρονοσειρών. Στο [26] το LSTM φαίνεται να υπερέχει των παραδοσιακών μεθόδων πρόβλεψης ικανό να ανιχνεύσει τις μακροπρόθεσμες εξαρτήσεις που παρουσιάζει η χρονοσειρά προκειμένου να προβλεφθεί η ημερήσια κατανάλωση φορτίου. Ακόμα, στο [27] ένα δίκτυο που βασίζεται σε LSTM εκπαιδεύεται χρησιμοποιώντας μια είσοδο πολλαπλών χαρακτηριστικών που περιλαμβάνει πληροφορίες σχετικά με τη θερμοκρασία, τις αργίες / εργασίες ημέρες καθώς και τις πληροφορίες ημερομηνίας και ώρας, με ταυτόχρονη χρήση γενετικών αλγορίθμων για την εξακρίβωση των επιπέδων και παραθύρου προηγούμενης γνώσης του LSTM. Τέλος, η χρήση δικτύων αρχιτεκτονικής GRU αναφέρεται στο [28] χρησιμοποιώντας τα για την ημερήσια πρόβλεψη ατομικής κατανάλωσης (δηλαδή σε επίπεδο καταναλωτή).

Άλλη κατηγορία εξελιγμένων μοντέλων που έχουν προσφάτως χρησιμοποιηθεί για το πρόβλημα πρόβλεψης φορτίου βραχυπρόθεσμου χρόνου (STLF) είναι τα νευρωνικά δίκτυα Γενετικών Αλγορίθμων (GA-NN) [29], [30], όπου οι γενετικοί αλγόριθμοι μειώνουν τον χρόνο εκπαίδευσης και επιλέγουν τις παραμέτρους εκπαίδευσης του μοντέλου. Ακόμα, ακολουθιακά (sequence to sequence) μοντέλα, έχουν υιοθετηθεί σε βραχυπρόθεσμες προβλέψεις φορτίου. Στο [31] ένα βασισμένο σε αρχιτεκτονική LSTM μοντέλο κωδικοποιητή αποκωδικοποιητή παρουσιάζει ανώτερη απόδοση σε σύγκριση με την κλασική μορφή του LSTM. Στο [32] οι συγγραφείς παρουσιάζουν μια εκδοχή ακολουθιακού μοντέλου βασισμένου σε αρχιτεκτονική RNN για πρόβλεψη χρονοσειρών ηλεκτρικών φορτίων και αποδεικνύουν την καλύτερη απόδοσή τους σε σχέση με μια σειρά μοντέλων που κυμαίνονται από τυπικά RNN έως κλασικές στατιστικές αναλύσεις χρονοσειρών.

Όσον αφορά στα συνελκτικά δίκτυα, στο [33] παρουσιάζεται μια υβριδική προσέγγιση που περιλαμβάνει τόσο συνελκτικές όσο και αρχιτεκτονικές ανατροφοδότησης. Οι συγγραφείς ενσωματώνουν διαφορετικές πηγές εισόδου και χρησιμοποιούν συνελκτικά επίπεδα προκειμένου να εξάγουν σημαντικά χαρακτηριστικά από το ιστορικό του φορτίου ενώ το ανατροφοδοτούμενο δίκτυο μαθαίνει την δυναμική του συστήματος. Το μοντέλο αξιολογείται σε ένα μεγάλο σύνολο δεδομένων που περιέχει ωριαία φορτία από μια πόλη στην Βόρεια Κίνα και συγκρίνεται με ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης. Μια διαφορετική υβριδική προσέγγιση παρουσιάζεται στο [34], όπου οι συγγραφείς επεξεργάζονται τις πληροφορίες του φορτίου παράλληλα με ένα CNN και ένα LSTM. Τα χαρακτηριστικά που δημιουργούνται από τα δύο δίκτυα στη συνέχεια χρησιμοποιούνται ως είσοδος σε ένα πλήρως συνδεδεμένο δίκτυο για την τελική πρόβλεψη του φορτίου της επόμενης ημέρας.

2.4 LightGBM (Gradient Boosting Machine)

Ο αλγόριθμος LightGBM [35] ανήκει στην οικογένεια των Gradient Boosting αλγορίθμων, όπου η εκμάθηση του συνόλου δεδομένων βασίζεται σε δέντρα αποφάσεων. Σε αυτό το σημείο πρέπει να γίνει μία ανάλυση των βασικών χαρακτηριστικών αυτού του μοντέλου ώστε να γίνει κατανοητή η χρήση του στα πλαίσια αυτής της εργασίας.

2.4.1 Gradient Boosting

Η συγκεκριμένη τεχνική αποτελεί μία επέκταση της τεχνικής Boosting. Το όνομα της το πήρε από δύο μεθόδους, τον αλγόριθμο Gradient Descent και την τεχνική Boosting. Συγκεκριμένα η τεχνική αυτή προσθέτει διαδοχικά δέντρα σε κάθε χρονική στιγμή ανάλογα με την αρνητική παράγωγο της συνάρτησης κόστους (loss function), σε αντίθεση με την μέθοδο Boosting όπου η εκπαίδευση γινόταν στα σφάλματα (residuals) της προηγούμενης επανάληψης. Συνήθης συνάρτηση ζημίας είναι το ημίαθροισμα των τετραγώνων των σφαλμάτων (residuals).

$$L(y_i, \hat{y}_i) = \frac{1}{2} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (2.1)$$

και υπολογίζοντας την παράγωγο:

$$\frac{dL(y_i, \hat{y}_i)}{dy_i} = y_i - \hat{y}_i, \quad (2.2)$$

Άρα, η αρνητική παράγωγος της συνάρτησης κόστους ισούται με τα σφάλματα (residuals). Οπότε εκπαιδεύουμε ένα δέντρο στα σφάλματα και προσθέτουμε μία μειωμένη έκδοση του δέντρου, δημιουργώντας ένα νέο. Χρησιμοποιώντας τη συγκεκριμένη συνάρτηση ζημίας η τεχνική Gradient Boosting είναι ισοδύναμη με την Boosting. Ο λόγος που η μέθοδος Gradient Boosting είναι ισχυρότερη από αυτή του Boosting είναι ότι μας δίνει τη δυνατότητα επιλογής διαφορετικών συναρτήσεων κόστους. Ανάλογα με τη δομή των δεδομένων μας χρησιμοποιούνται διαφορετικές συναρτήσεις κόστους. Για παράδειγμα αν στα δεδομένα μας υπάρχουν ακραίες παρατηρήσεις επηρεάζεται πολύ περισσότερο το άθροισμα των τετραγώνων από ότι το άθροισμα των απόλυτων σφαλμάτων από αυτές κάνοντας την χρήση του δεύτερου καταλληλότερη. Οι παράμετροι που πρέπει να οριστούν είναι ίδιες με αυτές της τεχνικής Boosting με τη μόνη διαφορά ότι δίνεται η επιλογή της συνάρτησης ζημίας με την προϋπόθεση ότι η συνάρτηση που επιλέγεται είναι διαφορίσιμη.

2.4.2 Βασική Ιδέα του αλγορίθμου LightGBM

Η βασική ιδέα που ακολουθείται για την δημιουργία των δέντρων είναι η παρακάτω τόσο για το LightGBM όσο και για τον XGBoost και CATBoost:

Η συνάρτηση ζημιάς που ελαχιστοποιήσουμε σε κάθε χρονική στιγμή (t) έχει τύπο :

$$L^{(t)} = \sum_{i=1}^N l(y_i, \hat{y}_i^{(t)}) + \sum_{f=1}^T \Omega_f^{(t)}, \quad (2.3)$$

Όπου, Ο πρώτος όρος μετράει πόσο καλά προσαρμόζεται το μοντέλο στα δεδομένα εκπαίδευσης (μικρές τιμές υποδηλώνουν καλή προσαρμογή) και ο δεύτερος την πολυπλοκότητα του κάθε δέντρου. Στην πολυπλοκότητα του δέντρου εισάγεται και ένας νέος όρος εκτός από τον αριθμό φύλλων (T), πιο συγκεκριμένα γίνεται συρρίκνωση των βαρών των φύλλων.

$$\Omega_f^{(t)} = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2, \quad (2.4)$$

όπου, γ υποδηλώνει πόσο θα τιμωρείται το μέγεθος του δέντρου. Μεγάλες τιμές του γ οδηγούν σε μικρά δέντρα, αντίστοιχα μικρές τιμές του γ οδηγούν σε μεγάλα δέντρα και η παράμετρος λ ρυθμίζει κατά πόσο θα συρρικνώνονται τα βάρη του δέντρου. Όσο αυξάνεται η τιμή της τα βάρη (w_j) του δέντρου συρρικνώνονται.

Συνοπτικά, ένα δέντρο μεγαλώνει μέχρι ένα συγκεκριμένο βάθος που ορίζεται από τον χρήστη. Πραγματοποιείται κλάδεμα του δέντρου σε όσους διαχωρισμούς χρειάζεται ώστε να έχουμε μεγαλύτερο κέρδος. Τέλος προστίθεται στο μοντέλο περικομμένη έκδοση του νέου δέντρου και επαναλαμβάνουμε τη διαδικασία δημιουργώντας το μοντέλο.

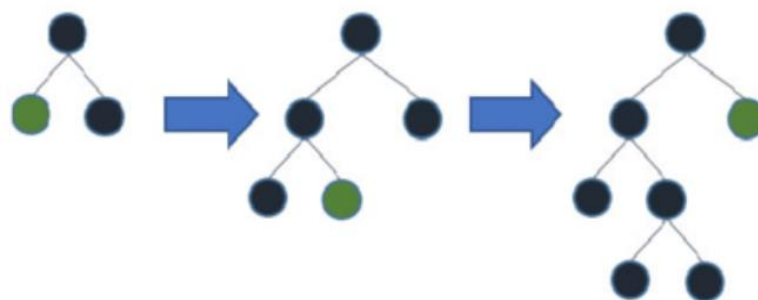
2.4.3 Χαρακτηριστικά αλγορίθμου LightGBM

Το πρώτο βασικό χαρακτηριστικό είναι στην διαδικασία που ακολουθείται ώστε να βρεθεί ο καλύτερος διαχωρισμός του συνόλου δεδομένων. Στον LightGBM δεν σαρώνεται το σύνολο των δεδομένων αναζητώντας τον βέλτιστο διαχωρισμό. Λαμβάνοντας υπόψη τον όγκο δεδομένων που είναι διαθέσιμος σήμερα μία τέτοια προσέγγιση πιθανότατα είναι υπολογιστικά ακριβή. Ο συγκεκριμένος αλγόριθμος δημιουργεί ιστογράμματα και χρησιμοποιεί τις κλάσεις που παράγονται, αντί για όλο το εύρος τιμών της κάθε μεταβλητής, επιτυγχάνοντας σημαντική μείωση στο χρόνο εκπαίδευσης.

Ακόμα, ο LightGBM δεν χρησιμοποιεί το σύνολο των δεδομένων εκπαίδευσης, αλλά ένα δείγμα του, το οποίο προκύπτει βάση της Gradient-based One Side Sampling technique (GOSS). Η ιδέα είναι ότι όλες οι παρατηρήσεις δεν συνεισφέρουν το ίδιο στην εκπαίδευση του αλγορίθμου, καθώς όσες έχουν μικρή πρώτη παράγωγο συνάρτησης ζημιάς είναι πιο καλά εκπαιδευμένες από όσες έχουν μεγάλη. Άρα επιλέγονται παρατηρήσεις με μεγάλη πρώτη παράγωγο ζημιάς και επιλέγονται επίσης τυχαία παρατηρήσεις με μικρή παράγωγο συνάρτησης ζημιάς (ικανοποιητικό δείγμα τους). Συγκεκριμένα, ταξινομούνται τα δεδομένα σύμφωνα με την απόλυτη τιμή της παραγωγού τους, επιλέγονται οι a (%) με την μεγαλύτερη παράγωγο και b (%) από τις υπόλοιπες με μικρή παράγωγο. Για τον υπολογισμό της συνάρτησης κέρδους οι παρατηρήσεις με μικρή παράγωγο πολλαπλασιάζονται με ένα συντελεστή, δίνοντας μεγαλύτερη σημασία στις μη καλά εκπαιδευμένες

χωρίς να διαφοροποιείται σημαντικά η κατανομή των δεδομένων. Έτσι, εκπαιδεύοντας μόνο ένα δείγμα σε κάθε επανάληψη επιτυγχάνεται σημαντική αύξηση στη διαδικασία εκμάθησης του αλγορίθμου, έχοντας ως αποτέλεσμα τη ταχύτερη σύγκληση του.

Επιπλέον, η τεχνική του LightGBM μεγαλώνει δέντρα ψάχνοντας τον βέλτιστο διαχωρισμό σε κάθε τερματικό κόμβο (φύλλο) ανεξαρτήτως επιπέδου (leaf-wise tree growth). Ουσιαστικά εξετάζει κάθε φορά ποιος τερματικός κόμβος αποφέρει τη μεγαλύτερη μείωση στη συνάρτηση ζημίας. Πρέπει να σημειωθεί ότι η συγκεκριμένη τεχνική είναι πολύ ευαίσθητη και μπορεί να οδηγήσει σε υπερπροσαρμογή του αλγορίθμου στο σύνολο δεδομένων που εκπαιδεύτηκε, κάνοντας τον μη αποτελεσματικό σε νέα σύνολα δεδομένων. Όμως, έτσι μειώνεται το υπολογιστικό κόστος του αλγορίθμου και η μνήμη που καταλαμβάνει. Η τεχνική αυτή παρουσιάζεται στην Εικ. 2.



Εικ. 2: Ανάπτυξη δέντρου leaf-wise

Τέλος, ο αλγόριθμος του LightGBM χρησιμοποιεί μία τεχνική μείωσης της διαστατικότητας των δεδομένων εισόδου, γνωστή ως Exclusive Feature Bundling technique (EFB). Αναλυτικά, τα δεδομένα υψηλών διαστάσεων είναι συνήθως πολύ αραιά και μας δίνουν τη δυνατότητα να σχεδιάσουμε μια προσέγγιση χωρίς απώλειες για να μειώσουμε τον αριθμό των χαρακτηριστικών. Συγκεκριμένα, σε έναν αραιό χώρο χαρακτηριστικών, πολλά χαρακτηριστικά είναι αμοιβαία αποκλειόμενα, δηλαδή δεν λαμβάνουν ποτέ μη μηδενικές τιμές ταυτόχρονα. Τα αμοιβαία αποκλειόμενα χαρακτηριστικά μπορούν να συνδυαστούν με ασφάλεια σε ένα μόνο χαρακτηριστικό. Ως εκ τούτου, η ταχύτητα εκπαίδευσης βελτιώνεται χωρίς να χειροτερεύσει η ακρίβεια.

2.5 N-BEATS (Neural Basis Expansion Analysis for Time Series forecasting)

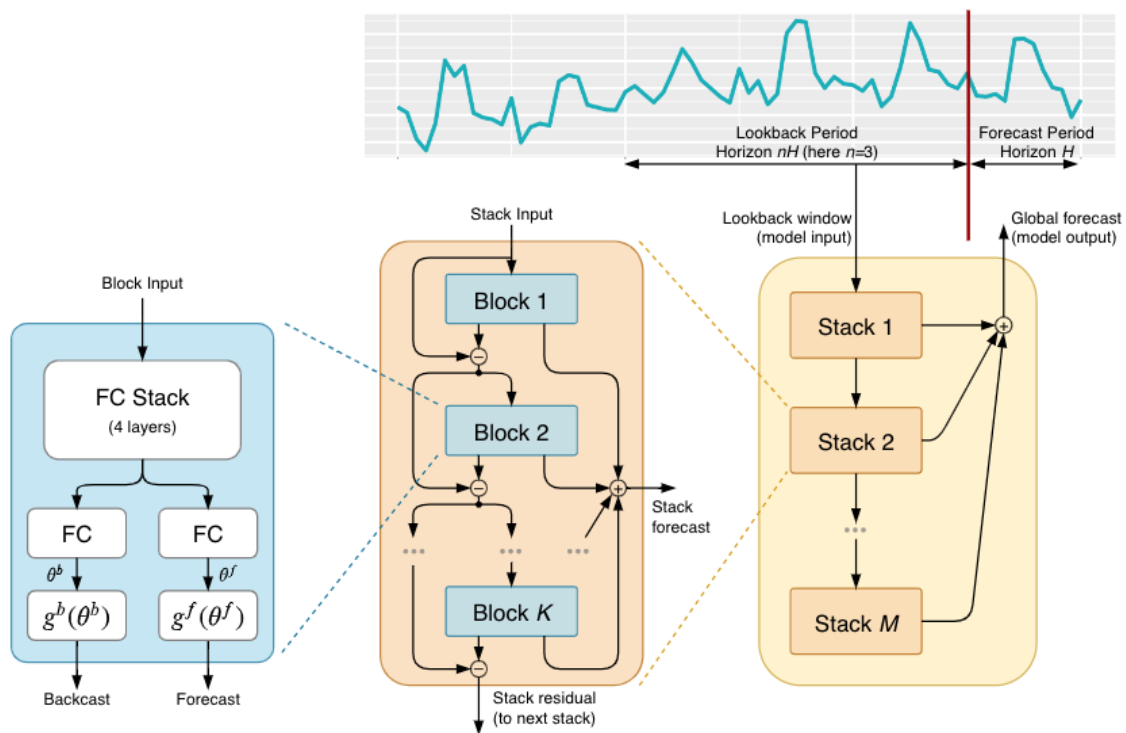
Το μοντέλο N-beats [36] είναι σχεδιασμένο με την αρχιτεκτονική των βαθιών νευρωνικών δικτύων. Βασίζεται σε μια παραλλαγή των υπολειμματικών δικτύων (Residual networks) και στα πλήρως συνδεδεμένα νευρωνικά δίκτυα πολλαπλών επιπέδων. Στο άρθρο αναφέρονται δύο διαφορετικές αρχιτεκτονικές, η γενική και η ερμηνεύσιμη. Παρακάτω αναλύεται η γενική αρχιτεκτονική, επειδή το άρθρο την προτείνει για δεδομένα φορτίου, καθ' όσον οδηγεί σε καλύτερα αποτελέσματα.

2.5.1 Δομικά στοιχεία του N-BEATS

Η βασική δομική μονάδα (Block) αποτελείται από 4 επίπεδα πλήρως συνδεδεμένων δικτύων με γραμμική μονάδα ανόρθωσης σε διάταξη στοιβάς (FC+ReLU stacked) και στη συνέχεια διαχωρίζεται σε δύο εξόδους περνώντας μέσα από ένα άλλο FC+ReLU επίπεδο. Η μία έξοδος αντιστοιχεί σε μια προσπάθεια πρόβλεψης μεγέθους ίσο με το ζητούμενο ορίζοντα πρόβλεψης (forecasting Horizon (H)), ενώ η δεύτερη έξοδος αντιστοιχεί σε μια προσπάθεια πρόβλεψης των δεδομένων εισόδου, δηλαδή του παραθύρου προηγούμενης γνώσης (backcast).

Η επόμενη δομή είναι μια στοίβα (Stack) που αποτελείται από αρκετές βασικές δομικές μονάδες (Blocks) σε διάταξη υπολειμματικής στοίβας. Ονομάζεται διπλή υπολειμματική στοίβα (Double Residual Stacking), γιατί υπάρχει για κάθε έξοδο της βασικής μονάδας (Block) μία διαφορετική αριθμητική πράξη. Η βασική ιδέα είναι ότι η πρόβλεψη των δεδομένων εισόδου (backcast) αφαιρείται κάθε φορά από την είσοδο της επόμενης δομικής μονάδας (Block), δηλαδή ουσιαστικά αφαιρείται το ποσοστό του σήματος που έχει “μάθει” η βασική μονάδα, με αποτέλεσμα να μειώνεται η δυσκολία πρόβλεψης των υπολειπόμενων δεδομένων εισόδου. Αντίστοιχα η πρόβλεψη (forecast) της δομικής μονάδας (Block) προστίθεται κάθε φορά (Ensemble) με αποτέλεσμα την συσσώρευση της προβλεπτικής ικανότητας. Αυτό είναι ένα σημαντικό πλεονέκτημα καθώς έχουμε καλύτερη εκτέλεση του αλγορίθμου κλισης οπισθοδιάδοσης (fluid gradient backpropagation).

Τέλος τοποθετείται ένα πλήθος από τις παραπάνω διπλές υπολειμματικές στοίβες σε μια μορφή υπερ στοίβα. Όπου όλα τα αποτελέσματα από τις στοίβες αθροίζονται (Ensemble) ώστε να προκύψει η τελική πρόβλεψη. Στην Εικ. 3 παρουσιάζεται η βασική αρχιτεκτονική του N-Beats.



Εικ. 3: Αρχιτεκτονική του N-Beats [36]

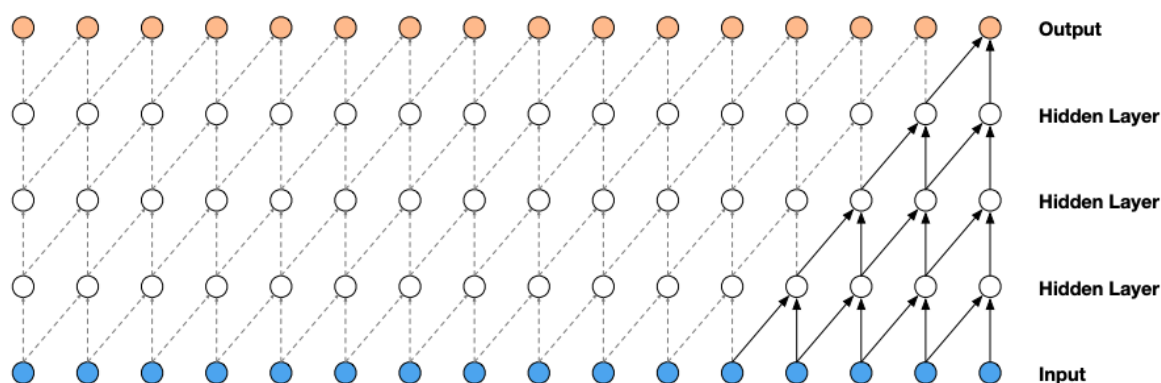
Τέλος, για την κανονικοποίηση (regularization) της προσαρμογής προτείνεται μία μέθοδος Ensemble, δηλαδή μία κατασκευή ενός πολυποικίλλου συνόλου μοντέλων N-beats, το κάθε μοντέλο με διαφορετική τυχαία αρχικοποίηση παραμέτρων (random state) και για την πρόβλεψη επιλέγεται η μέση έξοδος όλων των μοντέλων. Το σύνολο αυτό αποτελείται από μοντέλα τα οποία προσαρμόζονται με διαφορετικές συναρτήσεις κόστους (MAPE, sMAPE, MASE) και από μοντέλα με διαφορετικά παράθυρα προηγούμενης γνώσης πολλαπλάσια του ορίζοντα πρόβλεψης ($2H, 3H, \dots, 7H$). Προτείνεται να χρησιμοποιηθεί ο διάμεσος ως η μέση έξοδος αυτού του συνόλου. Το αποτέλεσμα είναι η καλύτερη ποιότητα προσαρμογής και καλύτερα αποτελέσματα στην

αξιολόγηση των προβλέψεων σε σχέση με άλλες μεθόδους κανονικοποίησης, όπως για παράδειγμα η μέθοδος αποβολής (dropout) και η κανονικοποίηση L2.

2.6 TCN (Temporal Convolutional Network)

Τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks CNNs) [37] είναι μια οικογένεια νευρωνικών δικτύων σχεδιασμένων να λειτουργούν με δεδομένα που μπορούν να αναπαρασταθούν σε μια τοπολογία τύπου πλέγματος. Τα CNN αρχικά χρησιμοποιήθηκαν σε δισδιάστατες και τρισδιάστατες εικόνες, αλλά αποδείχθηκαν επίσης κατάλληλα για μονοδιάστατα δεδομένα, όπως οι χρονοσειρές μιας μεταβλητής. Τα προβλήματα μοντελοποίησης ακολουθιών, όπως η βραχυπρόθεσμη πρόβλεψη ηλεκτρικών φορτίων (STLF), έχουν αντιμετωπιστεί κυρίως με την χρήση ανατροφοδοτούμενων νευρωνικών δικτύων, και άλλων τεχνικών που έχουν αναλυθεί, αλλά σχετικά πρόσφατες δείχθηκε πως τα συνελικτικά δίκτυα μπορούν επίσης να επιτύχουν κορυφαία απόδοση σε πολλές εφαρμογές, συμπεριλαμβανομένης της πρόβλεψη χρονοσειρών [38].

Τα συγκεκριμένα δίκτυα που χρησιμοποιούνται στην παρούσα εργασία ονομάζονται Χρονικά Συνελικτικά Δίκτυα (Temporal Convolutional Networks ή TCNs) για να καταδειχθεί ότι πρόκειται για συνελικτικά δίκτυα τα οποία είναι αυτοπαλινδρομητικά, ικανά να επεξεργαστούν ακολουθίες αυθαίρετου μήκους και να εξάγουν μια ακολουθία ίσου μήκους. Για να επιτευχθεί αυτό, το δίκτυο πρέπει να χρησιμοποιεί αιτιώδεις (δισταλμένες) συνελίξεις και οι υπολειπόμενες συνδέσεις θα πρέπει να χρησιμοποιούνται για τον χειρισμό πολύ μεγάλου μεγέθους ιστορικών τιμών. Στην Εικ. 4 παρουσιάζεται ένα συνελικτικό δίκτυο τριών επιπέδων με αιτιώδη συνέλιξη χωρίς διαστολή [39].



Εικ. 4: Δίκτυο τριών επιπέδων με αιτιώδη συνέλιξη χωρίς διαστολή [39]

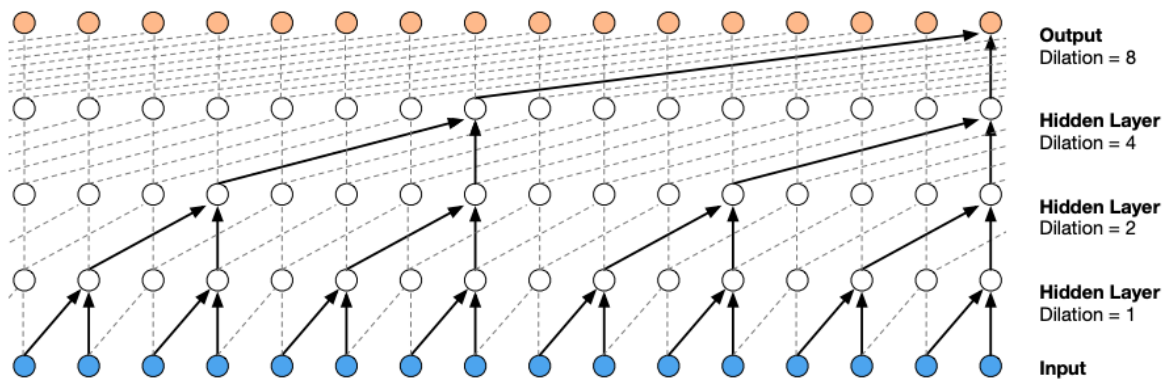
2.6.1 Δισταλμένη Αιτιώδης Συνέλιξη

Τα TCNs είναι μια οικογένεια αυτοπαλινδρομητικών μοντέλων, που η εκτιμώμενη τιμή την χρονική στιγμή t πρέπει να εξαρτάται μόνο από προηγούμενα δείγματα και όχι από μελλοντικά. Για να επιτευχθεί αυτή η συμπεριφορά σε ένα Συνελικτικό Νευρωνικό Δίκτυο, ο τυπικός χειριστής/σύμβολο της συνέλιξης αντικαθίσταται από την αιτιώδη συνέλιξη. Επιπλέον, προστίθεται παραγέμισμα με μηδενικά μήκους (μέγεθος φίλτρου = 1) προκειμένου να διασφαλιστεί ότι κάθε επίπεδο έχει το ίδιο μήκος με το επίπεδο εισόδου. Για την περαιτέρω ενίσχυση των δυνατοτήτων του δικτύου χρησιμοποιούνται δισταλμένες αιτιώδεις συνελίξεις, επιτρέποντας την αύξηση του δεκτικού πεδίου (receptive field) του δικτύου (δηλαδή, του αριθμού των νευρώνων εισόδου στους οποίους εφαρμόζεται το φίλτρο) και την ικανότητά του να μαθαίνει μακροχρόνιες εξαρτήσεις στις χρονοσειρές. Λαμβάνοντας υπόψη μια μονοδιάστατη είσοδο, x , και έναν πυρήνα, w , μια δισταλμένη έξοδος συνέλιξης με την χρήση ενός παράγοντα διαστολής d γίνεται:

$$f(i) = (x *_d w)(i) = \sum_{j=0}^{k-1} (i - dj)w(j), \quad (2.5)$$

Αυτό είναι ένα σημαντικό πλεονέκτημα σε σχέση με τις απλές αιτιώδεις συνελίξεις, καθώς στην τελευταία περίπτωση το δεκτικό πεδίο $r = k(L-1)$ (receptive field) αυξάνεται γραμμικά με το βάθος του δικτύου ενώ με την χρήση διεσταλμένων συνελίξεων η εξάρτηση είναι εκθετική $r = 2^{L-1}k$, διασφαλίζοντας ότι χρησιμοποιείται από το δίκτυο ένα πολύ μεγαλύτερο μέγεθος ιστορικού.

Στην Εικ. 5 παρουσιάζεται ένα συνελικτικό δίκτυο τριών επιπέδων με αιτιώδη συνέλιξη με χρήση διαστολής. Εδώ ο παράγοντας διαστολής d αυξάνεται με συντελεστή δύο (δηλαδή διπλασιάζεται) σε κάθε επίπεδο και το μέγεθος του πυρήνα k είναι 2, επομένως ο νευρώνας εξόδου επηρεάζεται από 8 νευρώνες εισόδου, δηλαδή το μέγεθος του ιστορικού είναι 8.



Εικ. 5: Δίκτυο τριών επιπέδων με αιτιώδη συνέλιξη με χρήση διαστολής [39]

Συγκεκριμένα, πρέπει να τονιστεί πως όλο το παράθυρο προηγούμενης γνώσης θα πρέπει να βρίσκεται εντός του πεδίου δεκτικότητας (receptive field). Για αυτό το λόγο τα συνελικτικά επίπεδα (layers, n) του δικτύου θα πρέπει να ικανοποιούν την σχέση:

$$n = \log_b \left(\frac{(l-1)(b-1)}{(k-1)} + 1 \right), \quad (2.6)$$

Όπου n είναι ο αριθμός των συνελικτικών επιπέδων, l το μήκος του παραθύρου προηγούμενης γνώσης, b η βάση διαστολής, και k το πλήθος των πυρήνων.

2.6.2 Πλεονεκτήματα και μειονεκτήματα των TCN

Τα χρονικά συνελικτικά νευρωνικά δίκτυα (TCN) έχουν διάφορα σημαντικά πλεονεκτήματα για αυτό χρησιμοποιούνται στην παρούσα εργασία. Συγκεκριμένα, μπορούν να εκπαιδευτούν σε παραλληλία, άρα όχι μόνο εκπαίδευσή τους γίνεται ταχύτερη αλλά μπορεί να γίνει χρήση της κάρτας γραφικών (GPU) με βέλτιστο τρόπο. Έτσι, δεν χρειάζεται μεγάλη χρήση μνήμης για την εκπαίδευσή τους. Ακόμα, δεν εμφανίζουν το πρόβλημα της εξαφάνισης της κλίσης (vanishing gradient) στο πεδίο του χρόνου, πρόβλημα που υπάρχει σε πληθώρα μοντέλων βαθιάς

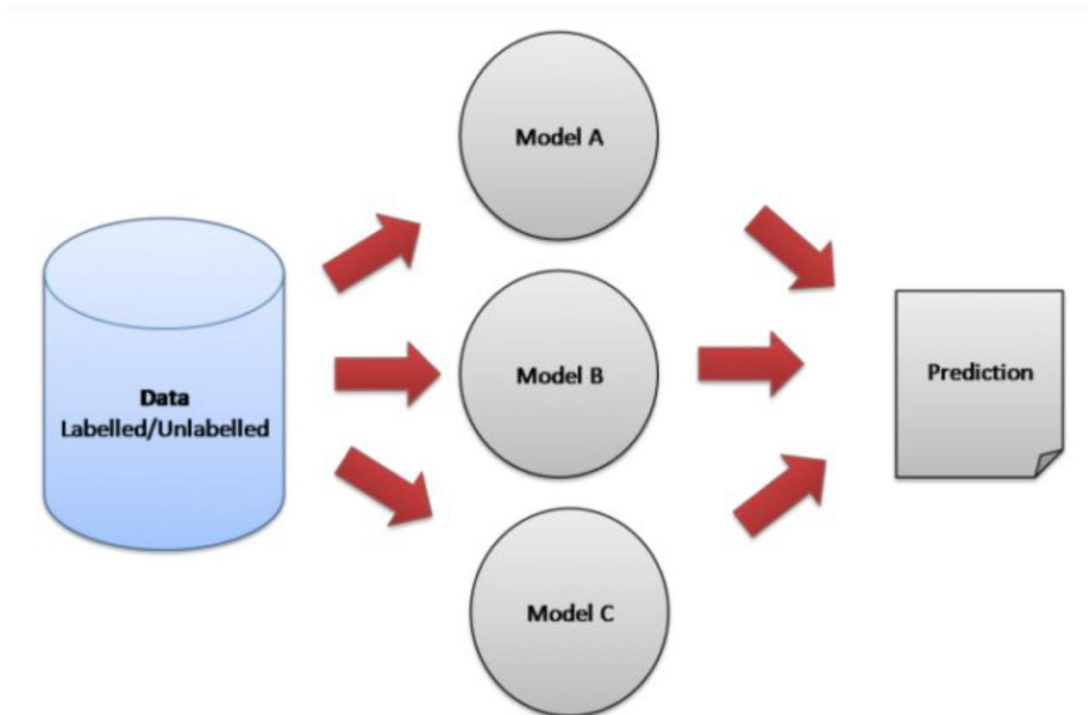
μάθησης όπως τα RNN. Έτσι, το χρονικό συνελκτικό δίκτυο (TCN) μπορεί να εκπαιδευτεί με την κατάλληλη διαστολή βάσης σε μεγάλα παράθυρα προηγούμενης γνώσης.

Παρόλα τα παραπάνω, τα δίκτυα αυτά παρουσιάζουν και διάφορα μειονεκτήματα. Το σημαντικότερο από αυτά είναι πως το πεδίο δεκτικότητας (receptive field) καθορίζεται a priori. Οπότε, δεδομένα εισόδου εκτός του πεδίου δεκτικότητας δεν μπορούν να συμπεριληφθούν στους υπολογισμούς της επικείμενης εξόδου. Έτσι, τα χρονικά συνελκτικά νευρωνικά δίκτυα (TCN) δεν μπορούν να έχουν αυθαίρετα μεγάλες εισόδους όπως και πεδίο δεκτικότητας (receptive field).

Η παραπάνω αρχιτεκτονική μπορεί να εξελιχθεί σε βαθιά χρησιμοποιώντας στοίβες (Stacks) από επίπεδα (layers). Σε αυτή την περίπτωση χρειάζονται υπολειπόμενες συνδέσεις (residual connections) ώστε να αποφευχθεί η εξαφάνιση κλίσης που μπορεί να εμφανιστεί ως πρόβλημα.

2.7 Μέθοδοι Ensemble

Πολλές φορές ένα μοντέλο δεν είναι αρκετό για να λύσει ένα πρόβλημα, καθώς μπορεί να εμφανίζει σφάλματα λόγω θορύβων, μεροληψίας ή διακύμανσης. Μια απλή ιδέα είναι η δημιουργία και εφαρμογή διαφόρων μοντέλων και η επίλυση να προέλθει από ένα συνδυασμό των απαντήσεων που δίνει κάθε μοντέλο, συνήθως μέσω κάποιας διαδικασίας εκλογής. Όταν έχουμε συγκερασμό μοντέλων, λέμε ότι έχουμε μια μέθοδο μάθησης συνόλων (ensemble) [39]. Η λογική είναι ότι η “γνώμη” της πλειοψηφίας υπερिशύει, ακόμα και αν αυτή αποτελείται από μη ισχυρά μοντέλα πρόβλεψης. Στην Εικ. 6 απεικονίζεται η τυπική διάταξη μιας Ensemble μεθόδου τριών μοντέλων.



Εικ. 6: Τυπική διάταξη Ensemble

Συνήθως, η μέθοδος Ensemble αποτελείται από ένα σύνολο μοντέλων, που καλούνται μοντέλα βάσης (base learners) και συνήθως δημιουργούνται από έναν αλγόριθμο μάθησης (base learning algorithm), που μπορεί να είναι δέντρο αποφάσεων, νευρωνικό δίκτυο ή άλλα είδη αλγορίθμων μάθησης. Αν το είδος των μοντέλων βάσης είναι κοινό, τότε λέμε ότι έχουμε ομογενή μοντέλα βάσης, ενώ στην αντίθετη περίπτωση λέμε ότι έχουμε ανομοιογενή μοντέλα βάσης, όπως στην παρούσα εργασία.

2.7.1 Averaging Ensemble

Το Averaging Ensemble όπως υποδηλώνει και το όνομά του είναι μία μέθοδος που προκύπτει από το μέσο όρο των αποτελεσμάτων διάφορων μοντέλων για το ίδιο πρόβλημα. Συνήθως χρησιμοποιείται σε νευρωνικά δίκτυα βαθιάς μάθησης για την εξασφάλιση της αμεροληψίας των βαρών. Συγκεκριμένα, αρχικοποιείται (random state) ένα μοντέλο διαφορετικά και εκπαιδεύεται. Μετά την εκπαίδευση των μοντέλων, ως τελική πρόβλεψη θεωρείται ο μέσος όρος των αποτελεσμάτων των μοντέλων. Ανήκει στην κατηγορία Soft Voting. Η ίδια διαδικασία (Ensemble) αλλά παίρνοντας ένα σταθμισμένο μέσο όρο ή εφαρμόζοντας πολυωνυμική παλινδρόμηση θα μπορούσε να οδηγήσει σε ακόμα καλύτερα αποτελέσματα αλλά αυξάνοντας παράλληλα το υπολογιστικό κόστος.

2.7.2 Bagging

Το bagging (συντομογραφία για την διαδικασία bootstrap aggregating) είναι μια τεχνική που δειγματοληπτεί δείγματα επαναληπτικά με αντικατάσταση από ένα σύνολο δεδομένων σύμφωνα με την ομοιόμορφη κατανομή πιθανότητας. Κάθε δείγμα εκκίνησης έχει το ίδιο μέγεθος με τα αρχικά δεδομένα. Επειδή η δειγματοληψία γίνεται με αντικατάσταση, κάποια δεδομένα μπορεί να εμφανίζονται αρκετές φορές στο ίδιο σύνολο εκπαίδευσης, ενώ άλλα μπορεί να παραλειφθούν από αυτό. Κατά μέσο όρο, ένα δείγμα εκκίνησης περιέχει περίπου το 63% των αρχικών δεδομένων εκπαίδευσης, επειδή κάθε δείγμα έχει πιθανότητα να επιλεγεί ίση με $1 - (1 - 1/N)^N$, οπότε παίρνοντας το όριο καθώς το N τείνει στο άπειρο, η πιθανότητα συγκλίνει στο 0.632.

Αρχικά, επιλέγεται ένας αριθμός k που αντιστοιχεί στον αριθμό των ομάδων δειγμάτων εκκίνησης (bootstrap samples). Στην συνέχεια δημιουργούνται k ομάδες D_i μεγέθους N και ένα μοντέλο βάσης εκπαιδεύεται σε κάθε D_i αντίστοιχα. Τέλος, υπολογίζεται η μέση πρόβλεψη από όλα τα μοντέλα βάσης, η οποία δίνεται σαν τελική.

Η διαδικασία bagging βελτιώνει το σφάλμα γενίκευσης μειώνοντας τη διακύμανση των μοντέλων βάσης. Η απόδοση εξαρτάται από τη σταθερότητα του μοντέλου βάσης. Εάν ένας ταξινομητής βάσης είναι ασταθής, η σάρωση βοηθά στη μείωση των σφαλμάτων που σχετίζονται με τις τυχαίες διακυμάνσεις στα δεδομένα εκπαίδευσης. Εάν ένας ταξινομητής βάσης είναι σταθερός, δηλαδή ανθεκτικός σε μικρές διαταραχές στο σύνολο εκπαίδευσης, τότε το σφάλμα του συνόλου προκαλείται κυρίως από μεροληψία του μοντέλου βάσης. Τέλος, δεδομένου ότι κάθε δείγμα έχει την ίδια πιθανότητα επιλογής, η διαδικασία είναι λιγότερο επιρρεπής σε υπερπροσαρμογή του μοντέλου (overfitting) όταν εφαρμόζεται σε δεδομένα με μεγάλο θόρυβο.

2.7.3 Άλλες κατηγορίες Ensembling

Υπάρχουν πολλά είδη Ensemble όπως αναλύονται στο [39]. Μία διαδικασία που έχει αναλυθεί ήδη στην ενότητα 2.4.1 Gradient Boosting είναι η διαδικασία Boosting. Περιληπτικά, το Boosting είναι μια επαναληπτική διαδικασία που χρησιμοποιείται για αλλάζει σταδιακά την κατανομή του συνόλου εκπαίδευσης, έτσι ώστε τα μοντέλα βάσης να επικεντρωθούν σε δεδομένα που δεν μπορούν να προβλέψουν με καλή ακρίβεια. Η διαδικασία αναθέτει ένα σταθμισμένο βάρος σε κάθε παράδειγμα του συνόλου εκπαίδευσης, το οποίο μπορεί δυνητικά να αλλάζει τιμή σε κάθε γύρο. Σε αντίθεση με το Bagging που κάθε μοντέλο βάσης μοντέλο έτρεχε ανεξάρτητα και στη συνέχεια το αποτέλεσμα συγκεντρωνόταν στις εξόδους, χωρίς προτίμηση σε οποιοδήποτε μοντέλο, το Boosting σχετίζεται με την ομαδική εργασία. Κάθε μοντέλο που τρέχει, υπαγορεύει σε ποια χαρακτηριστικά να επικεντρωθεί το επόμενο μοντέλο.

Άλλες διαδικασίες Ensemble είναι τα τυχαία δάση (random forest), όπου γίνεται τυχαία δειγματοληψία από το σύνολο δεδομένων και επιλέγονται τυχαία υποσύνολα χαρακτηριστικών. Ακόμα, το Stacking Ensemble είναι μία άλλη γνωστή κατηγορία όπου ένα μοντέλο κάνει μία

διαδικασία μάθησης και τροφοδοτεί το αποτέλεσμά του σε ένα άλλο μοντέλο. Κλασικό παράδειγμα είναι η παραγωγή περιγραφών εικόνας (image captioning), όπου ένα συνελκτικό δίκτυο δημιουργεί Embeddings που δίνονται ως δεδομένα σε ένα LSTM για την παραγωγή caption.

Κεφάλαιο 3 - Πειραματική διαδικασία

3.1 Εισαγωγή

Σε αυτό το κεφάλαιο, περιγράφεται η μεθοδολογία που ακολουθήθηκε για την περάτωση του Machine Learning life cycle. Τα στάδια του κύκλου ζωής είναι η συλλογή των δεδομένων, η ανάλυση τους, η επιλογή των προβλεπτικών μοντέλων, η διαδικασία εκπαίδευσης τους και η αξιολόγηση τους. Συγκεκριμένα, χρησιμοποιήθηκαν δεδομένα φορτίου (target), καιρού (συμμεταβλητή), τιμή μεγαβατώρας (συμμεταβλητή) και διάφορες χρονικές μεταβλητές (συμμεταβλητή). Μετά την επεξεργασία των δεδομένων, διενεργήθηκε εις βάθος ανάλυση με σκοπό την κατανόηση τους. Στην συνέχεια, εκπαιδεύτηκαν τα μοντέλα μηχανικής (ML) και βαθιάς μάθησης (DL), όσο και συγκερασμός τους, όπου χρησιμοποιήθηκαν διαφορετικές αρχιτεκτονικές, βελτιστοποιήθηκαν οι παράμετροι τους και αξιολογήθηκαν τα μοντέλα σύμφωνα με τις ανάγκες του προβλήματος βραχυπρόθεσμου χρόνου πρόβλεψης φορτίου (STLF) της παρούσας εργασίας.

Σκοπός της διαδικασίας που αναλύεται παρακάτω είναι αφενός η ανάλυση και πρόβλεψη βραχυπρόθεσμου χρόνου φορτίου (STLF) χρησιμοποιώντας state of the art τεχνικές και η σύγκρισή τους σε επίπεδο αρχιτεκτονικών, επιδόσεων και βιωσιμότητας; αφετέρου η ανάλυση της επίδρασης εξωτερικών μεταβλητών στα μοντέλα, για την επίλυση του προβλήματος αυτού. Προς την επίτευξη των προαναφερθέντων χρησιμοποιήθηκαν χρονοσειρές φορτίου και τιμής της μεγαβατώρας, της Πορτογαλίας, Ισπανίας και Ελλάδας, δηλαδή μεσογειακών ευρωπαϊκών χωρών με παρόμοια χαρακτηριστικά. Επίσης, χρησιμοποιήθηκαν χρονοσειρές καιρού της Πορτογαλίας, η οποία αποτελείται τόσο από πραγματικά δεδομένα όσο και από προβλέψεις.

Τέλος, σε μία χρονική περίοδο που οικονομικές και γεωπολιτικές αναταράξεις ταλανίζουν την εγγύς ευρωπαϊκή και όχι μόνο οικογένεια, ταυτόχρονα με τον όλο και αυξανόμενο εξηλεκτρισμό της καθημερινότητας και την διείσδυση ανανεώσιμων πηγών ενέργειας, προσπάθειες ανάλυσης και πρόβλεψης φορτίου καθίστανται αναγκαίες τόσο για την εύρυθμη λειτουργία των συστημάτων διανομής ενέργειας όσο και για σχεδιασμό της επόμενης ημέρας στον κλάδο της ενέργειας.

3.2 Συλλογή δεδομένων και εξαγωγή χαρακτηριστικών

3.2.1 Δεδομένα Φορτίου

Κύριος σκοπός της παρούσας εργασίας είναι η μελέτη των δεδομένων φορτίου και η διεξαγωγή προβλέψεων για αυτά για βραχυπρόθεσμου χρόνου πρόβλεψης φορτίου (STLF). Οι προβλέψεις αυτές είναι σημαντικές για μια αξιόπιστη και οικονομικά σταθερή λειτουργία ενός συστήματος ισχύος, όπως αναλύθηκε παραπάνω στην ενότητα 1.2 Σκοπός. Τα δεδομένα αυτά δημοσιοποιούνται για όλες τις ευρωπαϊκές χώρες από το European Network of Transmission System Operators for Electricity (ENTSO-E)², το οποίο συλλέγει δεδομένα που αποστέλλουν οι κατά τόπους διαχειριστές συστημάτων μεταφοράς (Transmission System Operators - TSOs), στην Ελλάδα ο ΑΔΜΗΕ, στην Πορτογαλία και στην Ισπανία οι αντίστοιχοι. Αν και η ανάλυση πιο συγκεκριμένων τιμών, σε επίπεδο περιφέρειας ή ακόμα και πόλης για παράδειγμα, θα είχαν ενδιαφέρον, τέτοια δεδομένα δεν διατίθενται. Θα μπορούσαν να αναζητηθούν δεδομένα μετρήσεων SCADA που διαθέτει ο ίδιος ο ΑΔΜΗΕ, ή οι αντίστοιχοι διαχειριστές συστήματος, ωστόσο οι απαιτούμενες διορθώσεις τους (συμμετοχή ΑΠΕ) και άλλων απαραίτητων διεργασιών για να χρησιμοποιηθούν είναι εκτός της εμβέλειας της εργασίας αυτής. Οπότε οι πειραματισμοί έγιναν σε

² <https://www.entsoe.eu/>

δεδομένα πανελλαδικού, παν ισπανικού και παν πορτογαλικού επιπέδου παρεχόμενα από το ENTSO-E.

Συγκεκριμένα χρησιμοποιήθηκαν ωριαία δεδομένα φορτίου της Πορτογαλίας, της Ισπανίας και της Ελλάδας από το 2016 έως το 2022. Επίσης, χρησιμοποιήθηκαν ως σύνολα test (test sets) από τον Μάρτιο του 2021 έως και τον Φεβρουάριο του 2022 και για τις τρεις προαναφερθείσες χώρες, ενώ σύνολα εκπαίδευσης από τον Μάρτιο του 2016 έως και τον Φεβρουάριο του 2021.

3.2.2 Μετεωρολογικά Δεδομένα

Ο καιρός θεωρείται βασική κινητήρια δύναμη της κατανάλωσης ηλεκτρικής ενέργειας [40]. Μοντέλα βασισμένα σε μετεωρολογικά δεδομένα χρησιμοποιούνται συχνά στη πρόβλεψη ηλεκτρικού φορτίου, όπως επιχειρείται και στα [41]–[43].

Ιδιαίτερο ενδιαφέρον παρουσιάζει η περίπτωση τροφοδοσίας στα μοντέλα δεδομένων πρόβλεψης καιρού, καθώς αυτό θα ήταν μία πραγματική προσομοίωση της πραγματικότητας. Συγκεκριμένα για να μπορεί να δοθεί σε ένα μοντέλο πρόβλεψης φορτίου η δυνατότητα να ενσωματώσει την επίδραση του καιρού στην κατανάλωση του φορτίου την ημέρα που θέλουμε να προβλέψουμε, είναι απαραίτητη η χρήση δεδομένων πρόβλεψης καιρού.

Στην διάθεσή μας έχουμε δεδομένα πρόβλεψης καιρού από 1 Ιανουαρίου του 2021 έως και τον Φεβρουάριο του 2022. Αναλυτικά τα δεδομένα αφορούν ωριαίες προβλέψεις για δύο επόμενες μέρες (48 ώρες), από τα οποία απομονώθηκαν οι ωριαίες προβλέψεις της επόμενης μέρας, καθώς είναι και οι πιο αξιόπιστες προβλέψεις. Τα δεδομένα αυτά προήλθαν από το Εθνικό Αστεροσκοπείο Αθηνών (Meteo)³ για τις 9 μεγαλύτερες πόλεις σε πληθυσμό της Πορτογαλίας, τις: Λισαβόνα, Πόρτο, Μπράγκα, Αβέιρο, Λείρια, Σάντα Ρεμ, Φάρο, Κοΐμπρα και Βιζέου. Για την συμπλήρωση του συνόλου των δεδομένων δηλαδή για τα δεδομένα καιρού από τον Φεβρουάριο του 2016 μέχρι και το 2020, χρησιμοποιήθηκαν πραγματικά δεδομένα από το Copernicus⁴, θεωρώντας πως στο σύνολο εκπαίδευσης (train set) που αφορά το παρελθόν μπορούν να χρησιμοποιηθούν και γιατί είναι γνωστά αλλά και γιατί αποτυπώνουν την πραγματικότητα. Συγκεκριμένα, τα δεδομένα αυτά είναι αποτέλεσμα ανάλυσης (reanalysis), δηλαδή συνδιασμός δεδομένων από μεντέλα πρόβλεψης και παρατηρήσεων, χρησιμοποιώντας τους νόμους της φυσικής. Η τεχνική αυτή ονομάζεται αφομοίωση δεδομένων (data assimilation) και εφαρμόζεται ευρέως από κέντρα πρόβλεψης καιρού.⁵ Να σημειωθεί πως στο σύνολο εκπαίδευσης (train set) δόθηκαν προβλέψεις καιρού για δύο μήνες, από 01/01/2021 έως 28/02/2021, το οποίο είναι σημαντικό καθώς το εκάστοτε μοντέλο θα μπορεί κατά αυτό τον τρόπο να συμπεριλάβει στην εκπαίδευσή του τις ενδεχόμενες αποκλίσεις των προβλέψεων καιρού. Συγκεκριμένα, χρησιμοποιήθηκαν δεδομένα Θερμοκρασίας, δηλαδή η μέση ωριαία τιμή θερμοκρασίας του αέρα μετρούμενη στα δύο μέτρα και η μέση τιμή της ωριαίας ταχύτητας του ανέμου.

Στην παρακάτω Εικ. 7 παρουσιάζεται ένας χάρτης στον οποίο φαίνονται τα σημεία των δεδομένων πρόβλεψης (2021-2022) του καιρού όπως και ιστορικών δεδομένων (2016-2020) καιρού.

³ <https://www.meteo.gr/index.cfm>

⁴ <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form>

⁵ <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview>



Εικ. 7: Πορτογαλία - Σημεία πρόβλεψης καιρού

3.2.3 Δεδομένα Τιμής Μεγαβατώρας (MWh)

Η τιμή του ηλεκτρικού ρεύματος ήταν ανέκαθεν ένας παράγοντας καθορισμού της κατανάλωσης. Για παράδειγμα, στην Ελλάδα όπως και σε άλλες ευρωπαϊκές χώρες, δίνεται η δυνατότητα μειωμένης τιμής κιλοβατώρας την νύχτα ή κάποιες ώρες το μεσημέρι. Η πρακτική αυτή θα γίνει ακόμα πιο έντονη δεδομένης της μεγαλύτερης διείσδυσης ανανεώσιμων πηγών ενέργειας την προσεχή δεκαετία πανευρωπαϊκά. Η στοχαστικότητα της παραγωγής ενέργειας από ΑΠΕ θα καθορίζει την κατανάλωση δίνοντας οικονομικά κίνητρα και άλλα. Ως εκ τούτου, στην παρούσα εργασία γίνεται προσπάθεια χρήσης της τιμής μεγαβατώρας ως συμμεταβλητή με σκοπό την πρόβλεψη του φορτίου όπως και στο [44].

Τα δεδομένα τιμής μεγαβατώρας που είναι διαθέσιμα και για τις τρεις χώρες, Πορτογαλία, Ισπανία και Ελλάδα προερχόμενα από τις εκάστοτε αρχές σε ωριαία ανάλυση, για τα έτη 2016 έως και τον Φεβρουάριο του 2021.

3.2.4 Άλλα Δεδομένα

Τα αυτοπαλινοδρομούμενα μοντέλα προσπαθούν να προβλέψουν μελλοντικές τιμές χρονοσειρών βλέποντας τις παρελθοντικές τιμές αυτών. Εκτός από τους παραπάνω εξωγενείς

παράγοντες που χρησιμοποιήθηκαν, καιρός και τιμή, δόθηκαν στα μοντέλα κωδικοποιημένος ο χρόνος και οι εργάσιμες ημέρες για την εκάστοτε χώρα.

Στον Πίνακα 1 φαίνεται ο τρόπος κωδικοποίησης του χρόνου. Η συνήθης πρακτική που ακολουθήθηκε είναι ο ημιτονοειδής και συνημιτονοειδής μετασχηματισμός του χρόνου, καθώς έτσι μπορεί να αποδοθεί η περιοδικότητα του. Τόσο ο ημιτονοειδής όσο και ο συνημιτονοειδής μετασχηματισμός πρέπει να συνυπάρχει, επειδή η απουσία ενός θα σύγχυζε το μοντέλο λόγω της μη αμφιμονοσήμαντης φύσης των συναρτήσεων αυτών. Για τις εργάσιμες ή όχι μέρες δόθηκε δυαδική τιμή.

Τέλος επισημαίνεται πως η χρονοσειρές που προκύπτουν (temporal covariates) αποτελούν συμμεταβλητές που βοηθούν στην πρόβλεψη της χρονοσειράς στόχου, δηλαδή του φορτίου.

Χρόνος	Περιοδικότητα	Κωδικοποίηση	Παραγόμενες Χρονοσειρές
Έτος	-	Φυσικός Αριθμός	1
Μήνας	Ετήσια	sine & cosine	2
Μέρα	Ετήσια	sine & cosine	2
Ημέρα του Χρόνου	Ετήσια	sine & cosine	2
Ημέρα της Εβδομάδας	Εβδομαδιαία	sine & cosine	2
Εβδομάδα του χρόνου	Ετήσια	sine & cosine	2
Εργάσιμη Ημέρα	Ετήσια	Δυαδική	1

Πίνακας 1: Παραγόμενες χρονοσειρές κωδικοποίησης χρόνου και εργάσιμης ημέρας (συμμεταβλητές)

3.2.5 Προεπεξεργασία Δεδομένων

Πριν την εφαρμογή αλγορίθμων για την εξόρυξη γνώσης, το ερευνώμενο σύνολο δεδομένων (data set) πρέπει να τεθεί υπό κατάλληλη επεξεργασία για να βελτιστοποιηθεί η απόδοση του συστήματος. Η προεπεξεργασία των δεδομένων είναι απαραίτητη, καθώς τα αρχικά δεδομένα πάσχουν από διαφόρων ειδών προβλήματα. Σε αυτά συγκαταλέγονται η ύπαρξη αλληλοσυγκρουόμενων πληροφοριών, δηλαδή διπλό καταχωρημένες τιμές (duplicate entries), προερχόμενα για παράδειγμα από την αλλαγή από χειμερινή σε θερινή ώρα, η ύπαρξη ασυνέπειας ως προς την κωδικοποίηση, την ονοματοδοσία πεδίων και τις μονάδες μέτρησης, καθώς και η ύπαρξη χαμένων τιμών και θορύβου.

Τα προβληματικά αυτά δεδομένα καλούνται “ακάθαρτα” και η διαδικασία αντιμετώπισης των προβλημάτων τους καλείται “καθαρισμός δεδομένων” (data cleansing). Η προεπεξεργασία των δεδομένων περιλαμβάνει τον καθαρισμό τους, αλλά δεν περιορίζεται σε αυτόν. Ιδιαίτερα στον τομέα της μηχανικής μάθησης, συχνά επιβάλλεται ο μετασχηματισμός των δεδομένων και ιδιαίτερα η κανονικοποίηση τους. Η κανονικοποίηση είναι η μετατροπή αριθμητικών τιμών σε άλλες, πιο “κατάλληλες” για τον εκάστοτε αλγόριθμο, αριθμητικές τιμές (για παράδειγμα μετασχηματισμός στο διάστημα [0,1]), για το οποίο έγιναν πειράματα αν και κατα πόσο οδηγούν σε καλύτερα αποτελέσματα.

3.2.5.1 Επεξεργασία Μετεωρολογικών Δεδομένων

Τα πραγματικά δεδομένα προέρχονται από Copernicus⁶ όπως έχει προαναφερθεί. Αρχικά προσδιορίστηκε το γεωγραφικό πλέγμα (grid) που απαιτείται ώστε να περιέχει εντός του όλα τα σημεία ενδιαφέροντος (9 πόλεις της Πορτογαλίας). Μετά, για κάθε παράμετρο που μας ενδιαφέρει (θερμοκρασία, ταχύτητα ανέμου) γίνονται κλήσεις API (Application Programming Interface) που επιστρέφουν τα εν λόγω δεδομένα για το πλέγμα (grid) σε μορφή grid. Οι κλήσεις όπως και η επεξεργασία που έπεται γίνονται με παραλληλοποίηση (threads) (ένα thread για κάθε παράμετρο). Έπειτα, για κάθε πόλη δηλαδή σημείο προσδιορισμένο με latitude and longitude, εκτελείται η εντολή (grid_get), η οποία επιστρέφει σε μορφή csv τα εν λόγω δεδομένα που ζητούνται. Σε αυτά τα δεδομένα, αφού διαβαστούν, και προσδιοριστούν τα missing values τους, ονομάτοδοτούνται οι στήλες τους και το αρχείο αποθηκεύεται στην επιθυμητή κατάσταση (csv). Η διαδικασία αυτή υπάρχει στο φάκελο με όνομα “weather” στο αρχείο “get_weather_data.py”.⁷

Τα δεδομένα της θερμοκρασίας έχουν Kelvin ως μονάδα μέτρησης και η ταχύτητα του ανέμου έχει αναλυθεί σε 2 συνιστώσες, κάθετη και οριζόντια. Η τελική μορφή χρήσης της ταχύτητας του ανέμου προκύπτει από την νόρμα 2 των προαναφερθέντων συνιστωσών. Η διαδικασία αυτή υπάρχει στο αρχείο “Norm of wind (W) and T2.ipynb”⁸.

Τα δεδομένα προβλέψεων καιρού δόθηκαν σε αρχείο csv οπότε δεν χρειάστηκε περαιτέρω επεξεργασία. Η θερμοκρασία δίνεται σε βαθμούς Κελσίου και η ταχύτητα του ανέμου σε 10 m/s, όπως έχει αναλύεται στην ενότητα 3.3.3 Χρονοσειρές Καιρού Πορτογαλίας.

3.2.5.2 Ελλιπούσες Τιμές (missing values)

Η ύπαρξη χαμένων τιμών (missing values) είναι ένα από τα συνηθέστερα προβλήματα των δεδομένων του πραγματικού κόσμου. Οι λόγοι για αυτό το φαινόμενο είναι πολλοί. Κάποιες πληροφορίες μπορεί να μην ήταν διαθέσιμες την ώρα της καταχώρησης ή μπορεί να διαγράφηκαν αργότερα από ανθρώπινο λάθος. Η αστοχία υλικού (για παράδειγμα σένσορες) ή λογισμικού είναι ένας άλλος λόγος ύπαρξης χαμένων τιμών. Οι χαμένες τιμές είναι ένα σημαντικό πρόβλημα στην επεξεργασία δεδομένων γιατί μπορεί να οδηγήσουν σε πλήρη αποτυχία τους αλγορίθμους που εφαρμόζονται. Για τον λόγο αυτό, η επίλυση αυτού του προβλήματος από την αρχή καθίσταται επιτακτική.

Συγκεκριμένα, όσον αφορά στα ωριαία δεδομένα φορτίου των Πορτογαλία, Ισπανία και Ελλάδα, δεν υπάρχει καμία χαμένη τιμή στο σύνολο δεδομένων της Πορτογαλίας. Αντίθετα στο σύνολο δεδομένων φορτίου της Ισπανίας υπάρχουν 6 χαμένες τιμές και σε αυτό της Ελλάδας 39 τιμές. Το κάθε σύνολο δεδομένων έχει 52584 σημεία που αφορούν τις ωριαίες τιμές φορτίου των 6 χρόνων που έχουν προαναφερθεί.

Αντίστοιχα, στο σύνολο δεδομένων του καιρού για τις πραγματικές τιμές που χρησιμοποιήθηκαν μέχρι και το 2020 δεν υπάρχει καμία χαμένη τιμή, ενώ για τα δεδομένα πρόβλεψης καιρού υπάρχουν 96 τιμές που λείπουν που αντιστοιχούν σε 4 ημέρες. Για τα δεδομένα τιμής της μεγιστώρας υπάρχουν 21, 13, 11 τιμές χαμένες για την Πορτογαλία, την Ισπανία και την Ελλάδα αντίστοιχα.

Η αντικατάσταση αυτών των τιμών είναι προϋπόθεση για οποιαδήποτε περαιτέρω ενέργεια προεπεξεργασίας. Συνήθεις πρακτικές είναι η χρήση μιας σταθερής τιμής όπως η μέση τιμή ενός συνόλου παρατηρήσεων ή πρόβλεψη της τιμής αυτής. Στην περίπτωση των συνόλων δεδομένων του φορτίου, επειδή οι χαμένες τιμές ήταν λίγες σε αριθμό και δεν παρατηρήθηκε κάποιο μοτίβο σε αυτές χρησιμοποιήθηκε πολυωνυμική παρεμβολή (Interpolation) με πολυώνυμο πέμπτου βαθμού

⁶ <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form>

⁷ https://github.com/giannissiso/Thesis_Short_Term_Load_Forecasting/tree/main/weather

⁸ https://github.com/giannissiso/Thesis_Short_Term_Load_Forecasting/tree/main/weather

[45]. Με παρόμοιο τρόπο αντικαταστάθηκαν οι τιμές στα δεδομένα τιμής μεγαβατώρας, ενώ στη θερμοκρασία για τις μισές μέρες χρησιμοποιήθηκαν προβλέψεις της δεύτερης μέρας όπου αυτό ήταν εφικτό, ενώ για τις υπόλοιπες τιμές χρησιμοποιήθηκε η συμπλήρωσή τους με την αμέσως προηγούμενη διαθέσιμη. Τα παραπάνω έγιναν στα αρχεία “wether data preprocessing up until whole 21 (t2).ipynb” και “wether data preprocessing up until whole 21 (u,v).ipynb”.⁹ Εάν παρατηρούνταν μοτίβο στις χαμένες τιμές ή οι χαμένες τιμές ήταν σε συνεχές μεγάλο κομμάτι του συνόλου δεδομένων θα ήταν χρήσιμη μία Γκαουσιανή διαδικασία (Gaussian process), δηλαδή διαδικασία μηχανικής μάθησης όπως στο [46], κάτι που δεν υφίσταται στα σύνολα δεδομένων της παρούσας εργασίας και επομένως δεν έγινε.

3.2.5.3 Κανονικοποίηση

Η κανονικοποίηση (normalization) είναι μια διαδικασία μετασχηματισμού δεδομένων, κατά την οποία αριθμητικές τιμές αντικαθίστανται με άλλες, πιο κατάλληλες, αριθμητικές τιμές. Η κανονικοποίηση των δεδομένων γίνεται ώστε να αντιμετωπιστούν δυσκολίες ορισμένων μεθόδων ανάλυσης και εξόρυξης γνώσης. Για παράδειγμα, τα Νευρωνικά Δίκτυα λειτουργούν καλύτερα όταν οι τιμές εισόδου κυμαίνονται στην περιοχή [0,1], σε κάποιες περιπτώσεις. Πρακτικά το διάστημα διακύμανσης των δεδομένων προσαρμόζεται στο εύρος εξόδου των συναρτήσεων ενεργοποίησης (σιγμοειδής, Relu).

Μια επαρκής εξομάλυνση, όχι μόνο των μεταβλητών εξόδου του δικτύου (μέσω των συναρτήσεων ενεργοποίησης) αλλά και των εισόδων, πριν από τη διαδικασία εκπαίδευσης, είναι πολύ σημαντική για την επίτευξη καλών αποτελεσμάτων καθώς και για τη σημαντική μείωση του χρόνου υπολογισμών, άρα και εκπαίδευσης [47].

Υπάρχουν διάφορες μέθοδοι κανονικοποίησης των αριθμητικών τιμών. Ορισμένες από τις πλέον χρησιμοποιούμενες είναι οι ακόλουθες:

- 1) Z-score: Η στατιστική αυτή μέθοδος, πραγματοποιεί μετασχηματισμό των αριθμητικών τιμών, χρησιμοποιώντας τη μέση τιμή και την τυπική απόκλιση τους. Για μία μεταβλητή, με μέση τιμή μ και τυπική απόκλιση σ , ο μετασχηματισμός των τιμών γίνεται σύμφωνα με τη σχέση:

$$x' = \frac{x-\mu}{\sigma}, \quad (3.1),$$

όπου x η εκάστοτε παρατήρηση και x' η κανονικοποιημένη τιμή. Η μέθοδος αυτή έχει καλά αποτελέσματα όταν υπάρχουν πολλά outliers στο σύνολο δεδομένων.

- 2) Ελαχίστου-Μεγίστου (MinMaxScaler): Με αυτήν τη μέθοδο κανονικοποίησης, οι αριθμητικές τιμές αντιστοιχίζονται με άλλες, οι οποίες κυμαίνονται εντός μίας προκαθορισμένης περιοχής τιμών το διάστημα [0,1]. Η αντιστοίχιση γίνεται με τον γραμμικό μετασχηματισμό:

⁹ https://github.com/giannissiso/Thesis_Short_Term_Load_Forecasting/tree/main/weather

$$x' = \frac{x - \min_A(x)}{\max_A(x) - \min_A(x)}, \quad (3.2),$$

όπου $\max(x_i)$, $\min(x_i)$, το μέγιστο και το ελάχιστο των παρατηρήσεων σε ένα σύνολο δεδομένων A , x η εκάστοτε παρατήρηση και x' η κανονικοποιημένη τιμή. Η μέθοδος αυτή έχει το πλεονέκτημα ότι ο χρήστης προκαθορίζει την περιοχή τιμών, για παράδειγμα μπορεί να μετασχηματίσει τις τιμές έτσι ώστε να κυμαίνονται στην περιοχή $[0,1]$, ορίζοντας σαν $\text{new.max}(x_i)=1$, $\text{new.min}(x_i)=0$. Επίσης, με τη μέθοδο αυτή διατηρείται η αναλογία μεταξύ των τιμών που υπήρχε στα αρχικά δεδομένα και είναι πολύ απλός και ο αντίστροφος μετασχηματισμός των δεδομένων στο αρχικό διάστημα, ώστε να υπολογιστούν οι διάφορες μετρικές.

Η μέθοδος που ακολουθείται είναι αυτή του ελαχίστου μεγίστου [48], τόσο στα δεδομένα καιρού και τιμής της μεγαβατώρας όσο και στα δεδομένα φορτίου με την διαφορά ότι σε αυτά δοκιμάστηκε και η εκπαίδευση των μοντέλων με τιμές στις φυσικές τους μονάδες. Η υλοποίηση έγινε με χρήση της βιβλιοθήκης `scikit-learn`¹⁰ της Python.

3.3 Ανάλυση δεδομένων

Οι γραφικές παραστάσεις της παρούσας ενότητας βρίσκονται στο αρχείο “Results .ipynb”¹¹

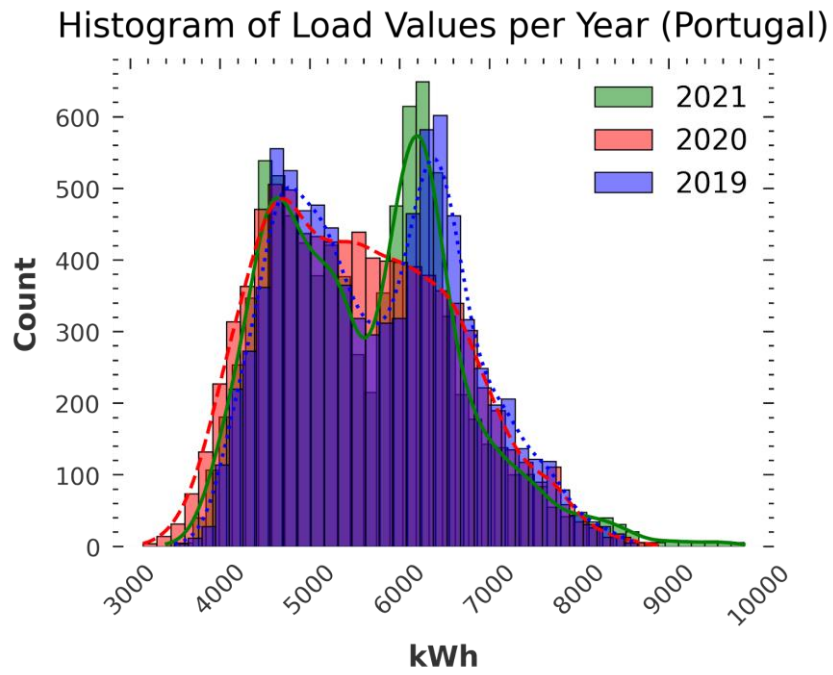
3.3.1 Χρονοσειρές Φορτίου

Αφού ολοκληρώθηκε η ανωτέρω διαδικασία καθαρισμού των δεδομένων, διεξήχθη ανάλυση με σκοπό την οπτικοποίηση τους και την εξαγωγή διαφορών χαρακτηριστικών. Κύριος σκοπός είναι η κατανόηση των δεδομένων και η παρατήρηση διαφορών προτύπων και μοτίβων. Η παρακάτω ανάλυση αφορά τα έτη 2019, 2020 που είναι το σύνολο αξιολόγησης (validation set) για τα μοντέλα βαθιάς μάθησης (DL) και το σύνολο δοκιμής (test set) 2021.

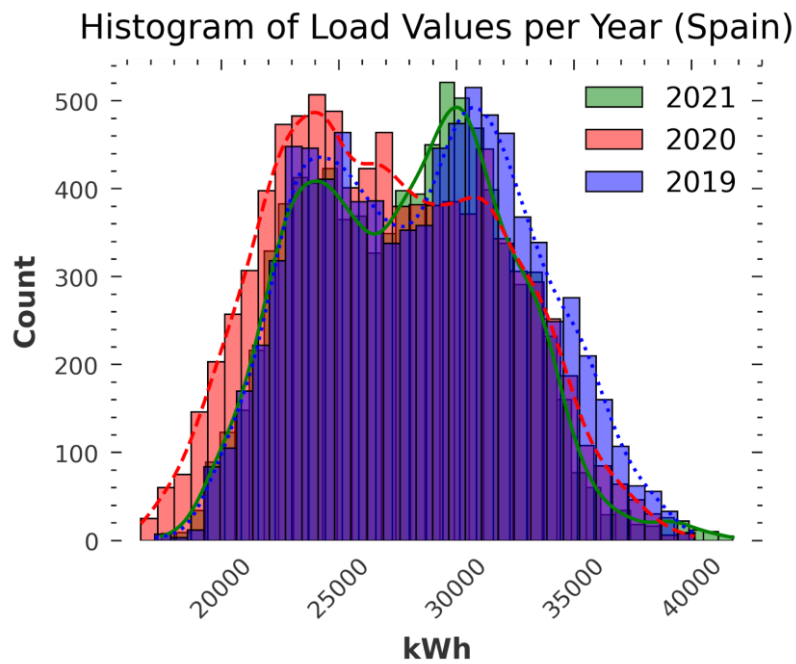
Στις Εικ. 8, 9, 10 απεικονίζονται τα ιστογράμματα των κατανομών φορτίου της Πορτογαλίας, Ισπανίας και Ελλάδας για το 2019, 2020 και 2021 αντίστοιχα σε kWh. Για τις Πορτογαλία και Ισπανία, παρατηρούνται για τα έτη εκτός του 2020 (COVID) ότι η κατανομή έχει 2 όροι (μίξη γκαουσιανών κατανομών) ενώ για το 2020 η κατανομή προσομοιάζει την Γκαουσιανή, δηλαδή έχει ένα όρος. Η αλλαγή αυτή οφείλεται στην περίοδο των περιορισμών λόγω της πανδημίας, όπου η κοινωνική και οικονομική ζωή περιορίστηκε οδηγώντας σε ραγδαία μείωση του καταναλισκόμενου φορτίου. Αντίθετα, η κατανομή φορτίου στην Ελλάδα έχει για όλα τα έτη ένα όρος προσομοιάζοντας κανονική κατανομή. Το έτος 2020 παρατηρείται όπως και στις άλλες χώρες μείωση του καταναλισκόμενου φορτίου.

¹⁰ <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html>

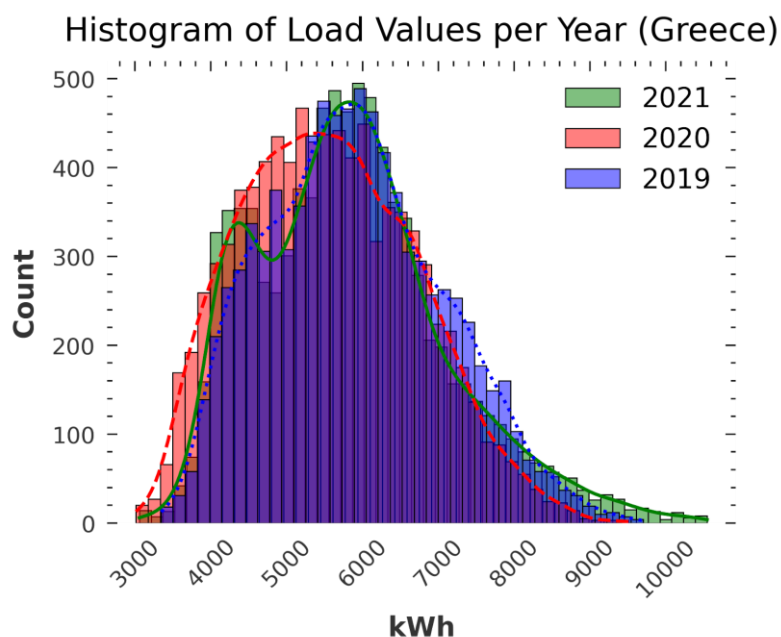
¹¹ https://github.com/giannisiso/Thesis_Short_Term_Load_Forecasting



Εικ. 8: Ιστόγραμμα ωριαίου φορτίου για τα 2019-2021 στην Πορτογαλία (kWh)

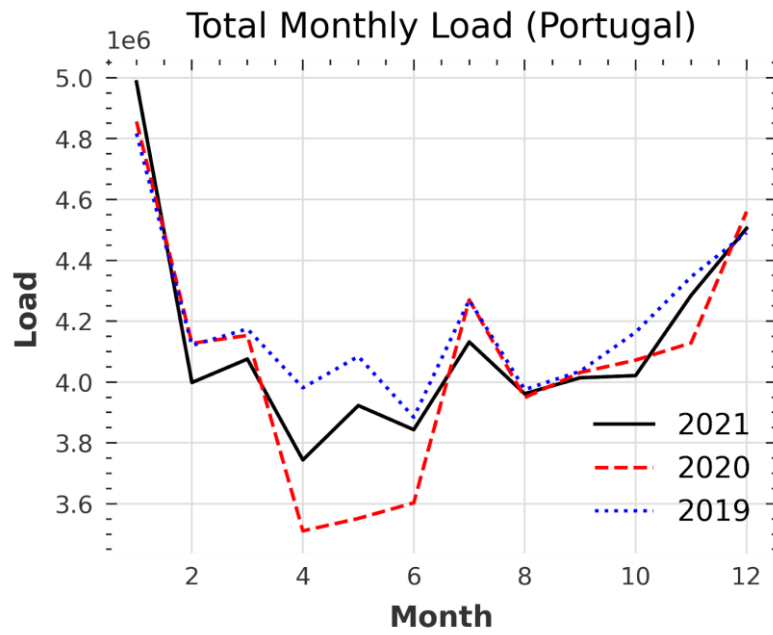


Εικ. 9: Ιστόγραμμα ωριαίου φορτίου για τα 2019-2021 στην Ισπανία (kWh)

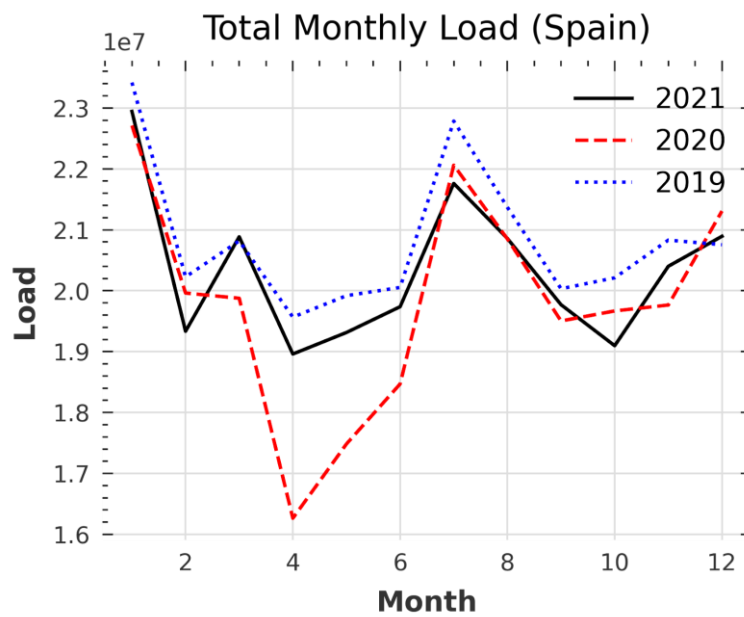


Εικ. 10: Ιστόγραμμα ωριαίου φορτίου για τα 2019-2021 στην Ελλάδα (kWh)

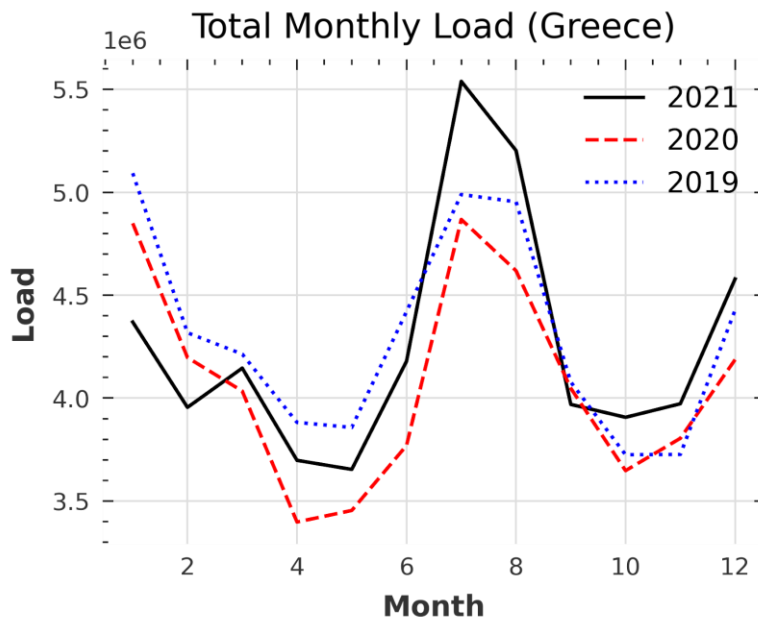
Στις Εικ. 11, 12, 13, απεικονίζεται το μηνιαίο φορτίο στις Πορτογαλία, Ισπανία, Ελλάδα, για τα έτη 2019-2021. Όσον αφορά την Πορτογαλία, τους χειμερινούς μήνες παρατηρείται η μεγαλύτερη κατανάλωση. Αντίστοιχα στην Ισπανία και Ελλάδα μεγάλη κατανάλωση παρατηρείται εκτός από τους χειμερινούς μήνες και τους καλοκαιρινούς λόγω της χρήσης κλιματιστικών. Λόγω του παρατεταμένου καύσωνα κατά τη διάρκεια του καλοκαιριού (Ιούλιο) στην Ελλάδα παρατηρείται εκτόξευση του καταναλισκόμενου φορτίου. Αυτή η αύξηση αναμένεται να μην προβλεφθεί ορθά από τα μοντέλα, οδηγώντας σε χειρότερες επιδόσεις. Ακόμα, κατά το έτος 2020 (validation set in DL models) λόγω των περιοριστικών μέτρων υπάρχει ραγδαία πτώση της καμπύλης φορτίου. Τέλος, τους μήνες του Φθινοπώρου και της Άνοιξης παρατηρείται και στις τρεις χώρες μείωση του καταναλισκόμενου φορτίου λόγω μειωμένης χρήσης συστημάτων θέρμανσης ή ψύξης.



Εικ. 11: Μηνιαίο φορτίο για τα 2019-2021 στην Πορτογαλία (kWh)

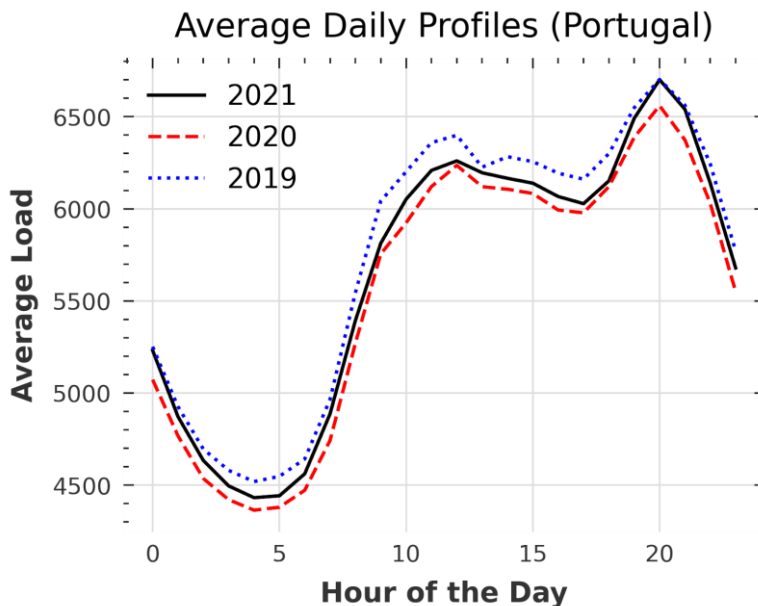


Εικ. 12: Μηνιαίο φορτίο για τα 2019-2021 στην Ισπανία (kWh)

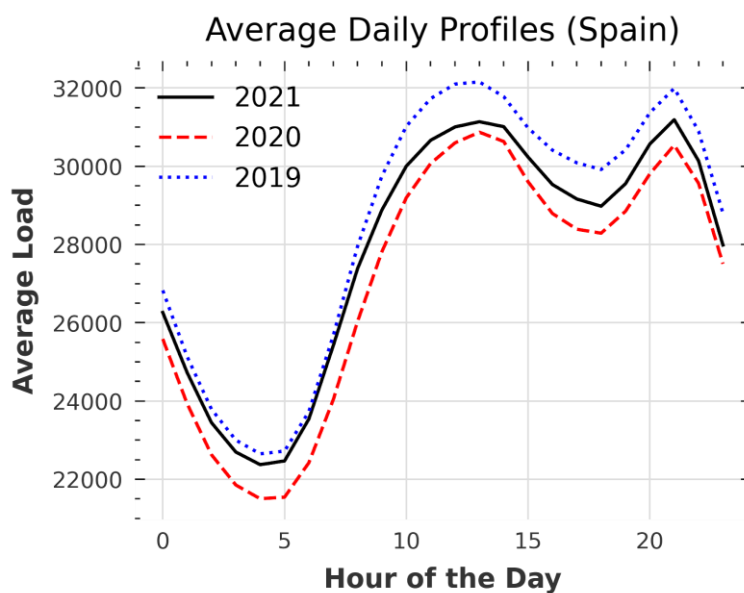


Εικ. 13: Μηνιαίο φορτίο για τα 2019-2021 στην Ελλάδα (kWh)

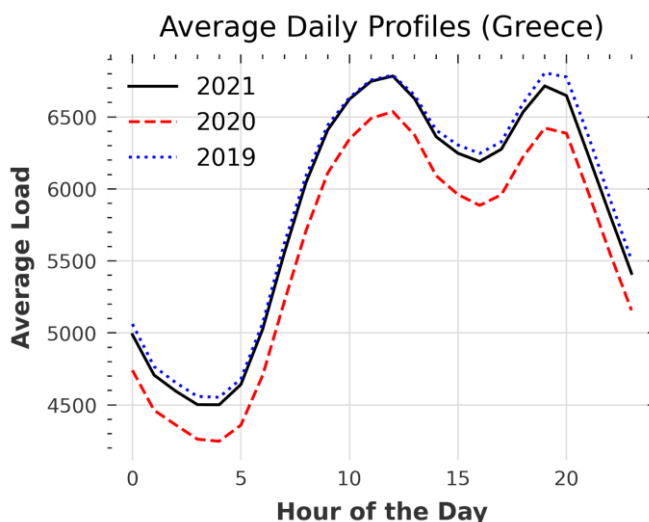
Στις Εικ. 14, 15, 16, απεικονίζεται το μέσο ημερήσιο προφίλ φορτίου στις, Πορτογαλία, Ισπανία και Ελλάδα για τα έτη 2019-2021. Επισημαίνεται πως κατά το έτος 2020, η εφαρμογή περιοριστικών μέτρων μέτρων για την ανακοπή εξάπλωσης του κοροναϊού τους μήνες της άνοιξης άλλαξε το προφίλ του καταναλισκόμενου φορτίου, μειώνοντας το εμφανώς και στις τρεις χώρες, ιδιαίτερα περισσότερο στην Ελλάδα όπου η περίοδος του εγκλεισμού ήταν μεγαλύτερη. Επιπροσθέτως, παρατηρείται και στις τρεις χώρες μείωση του φορτίου από τις 11.00 το βράδυ μέχρι και τις πρώτες πρωινές ώρες, ενώ από τις 10.00 το πρωί μέχρι τις 9.00 το βράδυ η κατανάλωση του φορτίου μεγιστοποιείται.



Εικ. 14: Μέσο ημερήσιο προφίλ φορτίου για τα 2019-2021 στην Πορτογαλία (kWh)



Εικ. 15: Μέσο ημερήσιο προφίλ φορτίου για τα 2019-2021 στην Ισπανία (kWh)



Εικ. 16: Μέσο ημερήσιο προφίλ φορτίου για τα 2019-2021 στην Ελλάδα (kWh)

Για την κατανόηση και ανάλυση των χρονοσειρών φορτίου των χωρών έγιναν επιπλέον διάφορες αναλύσεις ώστε να κατανοηθούν διάφορα βασικά χαρακτηριστικά. Αρχικά, βρέθηκε η εποχιακότητα (seasonality) τους η οποία προκύπτει για τα ωριαία δεδομένα, όλα τα πολλαπλάσια του 24 για όλες τις χρονοσειρές φορτίου. Στη συνέχεια, δεδομένου ότι ο μεγάλος όγκος των δεδομένων καθιστά γραφικές αναλύσεις όπως χάρτης αυτοσυσχέτισης (ACF) ή χάρτης μερικής αυτοσυσχέτισης (PACF) μη ερμηνευσιμα, έγιναν τα τεστ KPSS και ADF. Τα τεστ αυτά έδειξαν το μιν ADF στατικότητα (stationarity) ενώ το KPSS όχι. Συγκεκριμένα, για το KPSS τα p-values των χρονοσειρών φορτίου είναι μικρότερα από 5% (significance level) (KPSS = 0.01(PT), KPSS = 0.045(SP), KPSS = 0.01(GR)) άρα απορρίπτεται η μηδενική υπόθεση του τεστ ενάντια της εναλλακτικής της, οπότε η χρονοσειρές δεν έχουν στατικότητα. Αντίθετα, για το ADF τεστ τόσο οι κρίσιμες τιμές (όλων των χωρών αρνητικές) όσο και τα p-values (ADF = 0.0(PT), ADF = 0.01(SP), ADF = 0.0(GR)) είναι μικρότερα από το 5% (significance level), άρα οι χρονοσειρές έχουν στατικότητα. Ως εκ τούτου οι χρονοσειρές φορτίου χρειάζονται διαφόριση, η οποία έγινε και

διενεργήθηκαν εκ νέου τα τεστ ((KPSS = 0.1(PT), KPSS = 0.32(SP), KPSS = 0.09(GR)) τα οποία έδειξαν στατικότητα (στο ADF όλες οι τιμές των p-values ίδιες με πριν).

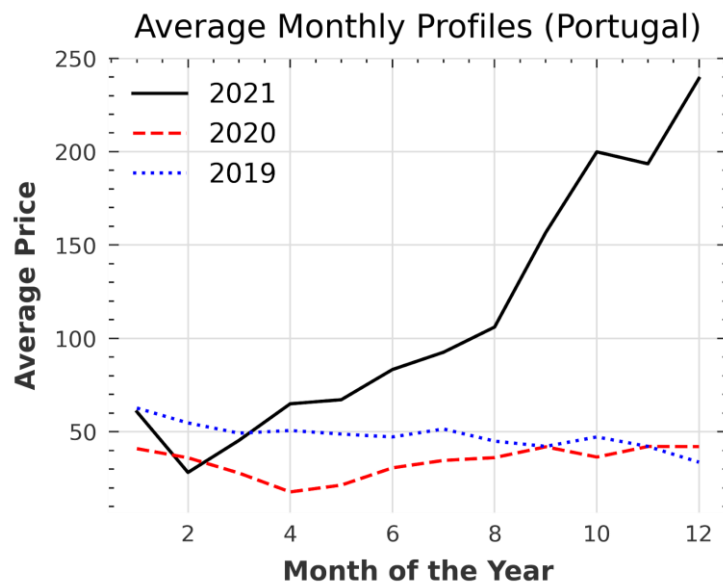
Ακόμα, διάφορα γενικά χαρακτηριστικά τα οποία επηρεάζουν την κατανάλωση φορτίου είναι ο πληθυσμός της εκάστοτε χώρας. Συγκεκριμένα, Πορτογαλία και Ελλάδα, με πληθυσμό 10.13 εκατομμύρια και 10.37 εκατομμύρια αντίστοιχα έχουν ένα μέσο ωριαίο προφίλ που δεν ξεπερνά τις 6500 kWh ενώ η Ισπανία με πληθυσμό 47.3 εκατομμύρια, σχεδόν 5 φορές μεγαλύτερο από τις άλλες χώρες έχει ημερήσιο προφίλ που δεν ξεπερνά τις 32000 kWh. Ο πληθυσμός είναι ανάλογος του καταναλισκόμενου φορτίου. Επιπροσθέτως, και οι τρεις χώρες είναι μεσογειακές με ανάλογα χαρακτηριστικά κλίματος και καιρού.

Τέλος επισημαίνεται πως η χρονοσειρά φορτίου αποτελεί την χρονοσειρά στόχο δηλαδή την χρονοσειρά που επιχειρείται η πρόβλεψη της.

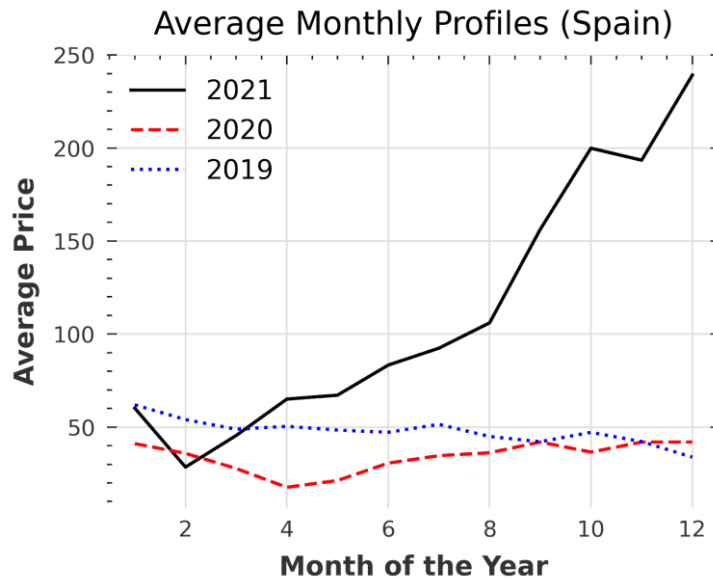
3.3.2 Χρονοσειρές Τιμής Μεγαβατώρας (MWh)

Στις Εικ. 17, 18, 19, φαίνεται η μέση μηνιαία τιμή μεγαβατώρας για τα έτη 2019, 2020, 2021 στις Πορτογαλία, Ισπανία και Ελλάδα αντίστοιχα. Όπως παρατηρείται, και στις τρεις χώρες αλλά και συνολικά σε όλη την ευρωζώνη η τιμή της μεγαβατώρας από τον Φεβρουάριο του 2021 και μέχρι σήμερα είναι στις υψηλότερες τιμές του ιστορικά. Με δεδομένο πως η χρήση του ηλεκτρισμού στις μέρες μας είναι απαραίτητη και δεν υπάρχει η δυνατότητα μη χρήσης η περιορισμού του, αναμένεται η ραγδαία αύξηση αυτή να μην μπορεί να βοηθήσει τα μοντέλα πρόβλεψης φορτίου. Αντίθετα αναμένεται η χειροτέρευση των μοντέλων καθώς δεν αποτυπώνεται αλλαγή στην συμπεριφορά του φορτίου με την αύξηση της τιμής της μεγαβατώρας.

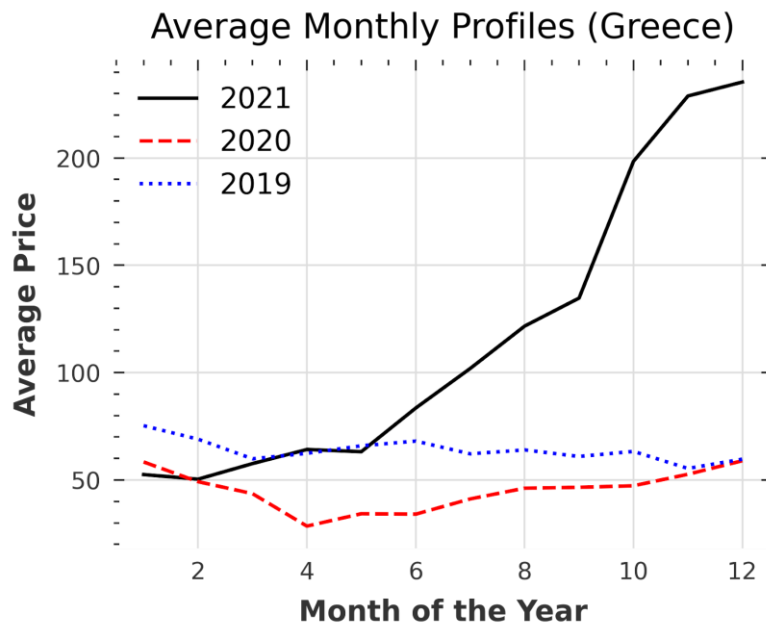
Τέλος επισημαίνεται πως η χρονοσειρά τιμής μεγαβατώρας αποτελεί την χρονοσειρά που βοηθά στην πρόβλεψη της χρονοσειράς στόχου, δηλαδή του φορτίου.



Εικ. 17: Μέση μηνιαία τιμή μεγαβατώρας για τα 2019-2021 στην Πορτογαλία (Ευρώ/MWh)



Εικ. 18: Μέση μηνιαία τιμή μεγαβατώρας για τα 2019-2021 στην Ισπανία (Ευρώ/MWh)



Εικ. 19: Μέση μηνιαία τιμή μεγαβατώρας για τα 2019-2021 στην Ελλάδα (Ευρώ/MWh)

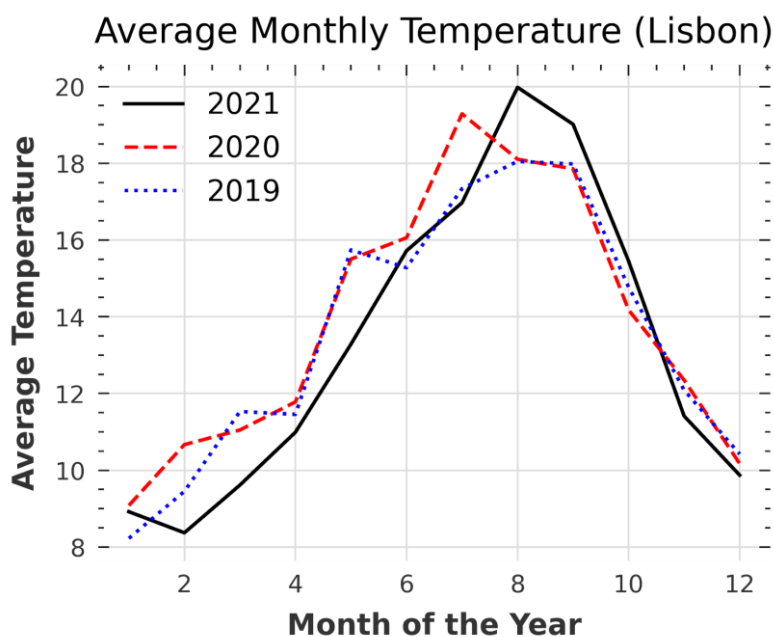
3.3.3 Χρονοσειρές Καιρού Πορτογαλίας

Αρκετές αναλύσεις έχουν γίνει στην συσχέτιση του καιρού με το φορτίο όπως στο άρθρο [41]. Στην ανάλυση που έγινε σε αυτή την εργασία χρησιμοποιήθηκε Granger Causality test, το οποίο είναι ενδεικνύμενο για το εάν μία χρονοσειρά (καιρός) είναι χρήσιμη για να προβλεφθεί μία άλλη (φορτίο). Από την διενέργεια του τεστ προέκυψε πως όλα τα p-values είναι κάτω του 0.05 για παράθυρο (lags) μέχρι 5, με διάστημα εμπιστοσύνης 95%. Επομένως, απορρίπτεται η Null Hypothesis, άρα ο καιρός χρησιμεύει στην πρόβλεψη του φορτίου. Επίσης η εποχικότητα του καιρού (seasonality) είναι ίδια με αυτή του φορτίου. Τα παραπάνω βρίσκονται στο αρχείο κώδικα “weatherloadcorrelation.ipynb”¹²

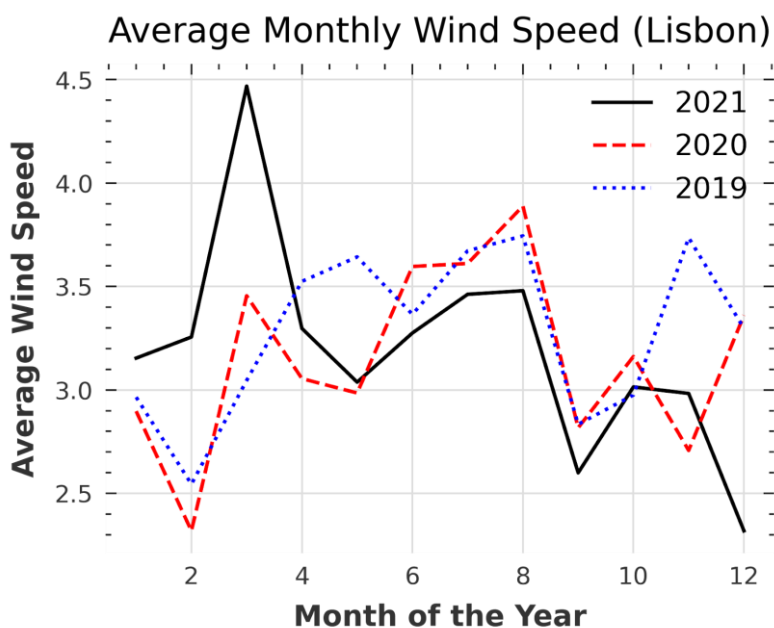
¹² https://github.com/giannisiso/Thesis_Short_Term_Load_Forecasting/tree/main/weather

Στις Εικ. 20, 21, παρουσιάζεται η μέση μηνιαία τιμή θερμοκρασίας σε βαθμούς Κελσίου και η μέση μηνιαία τιμή της ταχύτητας του ανέμου σε (10 m/s), στην Λισαβόνα, την πρωτεύουσα της Πορτογαλίας. Με βάση τις τιμές της θερμοκρασίας, αναμένεται στις πολύ χαμηλές τιμές της θερμοκρασίας να λειτουργούν ηλεκτρικά μηχανήματα θέρμανσης και αντίστοιχα στις υψηλές τιμές της θερμοκρασίας μηχανήματα ψύξης. Επομένως, αναμένεται αύξηση του καταναλισκόμενου φορτίου. Η ταχύτητα του ανέμου που δόθηκε είτε με τις πραγματικές τιμές είτε με δυαδική κωδικοποίηση πάνω και κάτω από κάποιο κατώφλι θερμοκρασίας θα μπορεί να δώσει αντίστοιχα χρήσιμη πληροφορία για εξόρυξη γνώσης.

Τέλος επισημαίνεται πως η χρονοσειρές καιρού αποτελούν χρονοσειρές που βοηθούν στην πρόβλεψη της χρονοσειράς στόχου, δηλαδή του φορτίου.



Εικ. 20: Μέση μηνιαία θερμοκρασία για τα 2019-2021 στην Λισαβόνα (C)



Εικ. 21: Μέση μηνιαία ταχύτητα ανέμου για τα 2019-2021 στην Λισαβόνα (10 m/s)

3.4 Επιλογή Μοντέλων

Σε αυτή την ενότητα παρουσιάζονται τα μοντέλα μηχανικής μάθησης και βαθιάς μάθησης που χρησιμοποιήθηκαν στην πειραματική διαδικασία. Τα μοντέλα αυτά έχουν παρουσιαστεί εκτενώς στις ενότητες 2.4 LightGBM (Gradient Boosting Machine) 2.5 N-BEATS (Neural Basis Expansion Analysis for Time Series forecasting) Στα παρακάτω μοντέλα χρησιμοποιήθηκαν οι υλοποιήσεις της βιβλιοθήκης Darts¹³

Θα γίνει χρήση των:

A) **LightGBM** (Gradient Boosting Machine)

Το LightGBM ανήκει στην κατηγορία των μοντέλων δέντρων αποφάσεων χρησιμοποιώντας όμως gradient boosting μέθοδο (ensemble) όπως προτάθηκε από το [35] και έχει χρησιμοποιηθεί σε προβλήματα βραχυπρόθεσμου χρόνου πρόβλεψης φορτίου (STLF) στα [49], [50]. Η επιλογή του μοντέλου αυτού μηχανικής μάθησης εδράζεται στο γεγονός ότι είναι state of the art για πρόβλεψη χρονοσειρών χρησιμοποιώντας έξυπνες τεχνικές. Επίσης, ο σύντομος χρόνος εκπαίδευσης του δίνει μία αξιόπιστη πρόβλεψη με λίγους πόρους (CPU).

B) **N-Beats** (Neural Basis Expansion Analysis for Time Series forecasting)

Το N-Beats είναι ένα νευρωνικό δίκτυο όπως έχει ήδη αναλυθεί που εισήχθη από τον [36] το 2019 και έχει χρησιμοποιηθεί έκτοτε σε προβλήματα βραχυπρόθεσμου χρόνου πρόβλεψης φορτίου (STLF) όπως στα [51]–[53]. Το μοντέλο αυτό επιλέχθηκε καθώς δεν έχει ευρέως χρησιμοποιηθεί για προβλήματα βραχυπρόθεσμου χρόνου πρόβλεψης φορτίου (STLF) και οι επιδόσεις του είναι εγνωσμένης αξίας καθώς έχει ξεπεράσει τον νικητή του διαγωνισμού πρόβλεψης M4 κατά 3%.

Γ) **TCN** (Temporal Convolutional Networks)

Το TCN είναι ένα συνελκτικό νευρωνικό δίκτυο με διεσταλμένα συνελκτικά επίπεδα μιας διάστασης το οποίο προτάθηκε στο [54] για πρόβλεψη χρονοσειρών. Έχει χρησιμοποιηθεί για παρόμοια προβλήματα βραχυπρόθεσμου χρόνου πρόβλεψης φορτίου στα [50], [53]. Η επιλογή του μοντέλου έγινε καθώς συνελκτικές τεχνικές είναι κατάλληλες για την εξαγωγή γνώσης από χρονοσειρές πολλών μεταβλητών, ανάλυση που είναι αντικείμενο της παρούσας εργασίας.

Δ) **Ensemble**

Η μέθοδος του Ensemble χρησιμοποιώντας τον μέσο όρο χρησιμοποιείται ευρέως σε νευρωνικά δίκτυα βαθιάς μάθησης και έχει αναλυθεί στο [55].

3.5 Διαδικασία Εκπαίδευσης

Στόχος της διαδικασίας εκπαίδευσης που θα αναλυθεί παρακάτω είναι η δημιουργία μοντέλων πρόβλεψης του ωριαίου φορτίου (1-hour resolution) σε kWh της επόμενης μέρας (day ahead forecast) για την ζήτηση ηλεκτρικής ενέργειας των διαχειριστών συστημάτων διανομής (TSOs) της Πορτογαλίας, της Ισπανίας και της Ελλάδας. Προς επίρρωση αυτού, χρησιμοποιήθηκαν τα επιλεγμένα μοντέλα της προηγούμενης ενότητας. Για τον διαχωρισμό του συνόλου δεδομένων (dataset split), στο μοντέλο μηχανικής μάθησης (ML), LightGBM δόθηκε ως σύνολο εκπαίδευσης (train set) τα πέντε έτη από 01/03/2016 έως 28/02/2021 ενώ στα μοντέλα βαθιάς μάθησης (DL), NBEATS και TCN δόθηκαν τα τέσσερα έτη από 01/03/2016 έως 29/02/2020 ως σύνολο εκπαίδευσης και το έτος από 01/03/2020 έως 28/02/2021 ως σύνολο αξιολόγησης (validation set). Ο λόγος που ο αλγόριθμος LightGBM δεν χρειάζεται validation set είναι η τεχνική Gradient-based One Side Sampling technique (GOSS) που αναλύεται στην Ενότητα 2.4.3 Χαρακτηριστικά αλγορίθμου LightGBM.

¹³ <https://unit8co.github.io/darts/>

Αρχικά, για το κάθε μοντέλο προσδιορίστηκε εάν αποδίδει καλύτερα με κανονικοποιημένα ή όχι δεδομένα φορτίου. Έπειτα, έγινε προσδιορισμός του καλύτερου παραθύρου προηγούμενης γνώσης (lookback window) για το κάθε μοντέλο. Για αυτό τον σκοπό δοκιμάστηκαν lookback windows από τέσσερις ημέρες ($4 \cdot 24 = 96$ παρατηρήσεις) έως δύο εβδομάδες ($14 \cdot 24 = 336$ παρατηρήσεις). Σαν μοντέλο βάσης για το εκάστοτε μοντέλο χρησιμοποιήθηκε το μοντέλο με είσοδο τη χρονοσειρά φορτίου με πραγματικές τιμές (kWh) και παράθυρο προηγούμενης γνώσης μία εβδομάδα ($7 \cdot 24 = 168$ παρατηρήσεις), χωρίς καμία εξωτερική μεταβλητή (univariate time series). Χρήσιμο θα ήταν σε αυτό το σημείο η διαφοροποίηση των συμμεταβλητών (covariates) που μπορούν να δεχτούν τα μοντέλα. Συγκεκριμένα υπάρχουν δύο είδη: οι παρελθοντικές (past covariates) και οι μελλοντικές συμμεταβλητές (future covariates). Η βασική τους διάκριση είναι πως οι μελλοντικές συμμεταβλητές είναι εις γνώση του μοντέλου για την ίδια στιγμή και για το μέλλον που θα γίνει πρόβλεψη ενώ οι παρελθοντικές δεν είναι, δηλαδή γνωρίζουμε μόνο τις παρελθοντικές τιμές. Ένα παράδειγμα που θα βοηθήσει στην κατανόηση είναι η πραγματική θερμοκρασία και οι προβλέψεις της θερμοκρασίας. Την μεν πραγματική θερμοκρασία δεν την γνωρίζουμε εκ των προτέρων οπότε αποτελεί παρελθοντική συμμεταβλητή ενώ τις προβλέψεις της θερμοκρασίας τις ξέρουμε και για την ίδια την τιμή πρόβλεψης και για τον ορίζοντα που μας δίνονται, οπότε αποτελεί μελλοντική συμμεταβλητή.

Η διαδικασία εκπαίδευσης για όλα τα μοντέλα έγιναν σε μηχάνημα με χαρακτηριστικά: 16 CPU cores, 64GB RAM, NVIDIA V100 32GB.

3.5.1 LightGBM

Όπως αναλύθηκε εκτενώς στο θεωρητικό μέρος ο τρόπος δημιουργίας πρόβλεψης των δέντρων αποφάσεων και πιο συγκεκριμένα του LightGBM, έχει παρουσιαστεί και η αρχιτεκτονική του είναι καθορισμένη.

Το μοντέλο αυτό είναι ικανό να δεχτεί τόσο παρελθοντικές όσο και μελλοντικές συμμεταβλητές (past and future covariates). Στον Πίνακα 1 φαίνονται οι συμμεταβλητές που δόθηκαν στο μοντέλο εξαρχής. Δεδομένου ότι όλες οι μεταβλητές του Πίνακα 1 είναι γνωστές εκ των προτέρων για όλο το σύνολο δεδομένων για κάθε χώρα επιλέχθηκε η χρήση τους σαν μελλοντικές. Στη συνέχεια δόθηκαν επίσης ως συμμεταβλητές οι χρονοσειρές τιμής της εκάστοτε χώρας. Όμως, καθώς το σύνολο δεδομένων του ηλεκτρικού ρεύματος αποτελεί γνώση που αποκτάται εκ των υστέρων και δεν υπάρχουν δεδομένα πρόβλεψης χρησιμοποιήθηκε ως παρελθοντική συμμεταβλητή. Τέλος, στην περίπτωση της Πορτογαλίας όπου υπάρχουν τα δεδομένα πρόβλεψης καιρού, δόθηκε στο μοντέλο ο καιρός σαν συμμεταβλητή μέλλοντος. Για τις μελλοντικές συμμεταβλητές του LightGBM δηλαδή τα temporal covariates και ο καιρός στην περίπτωση της Πορτογαλίας έγιναν πειράματα με διάφορα παράθυρα προηγούμενης γνώσης (look back window) με εύρος από 1 ώρα έως 48 ώρες και διάφορα παράθυρα μελλοντικής γνώσης (look forward window) από 1 έως 24 ώρες. Για τις παρελθοντικές συμμεταβλητές δηλαδή τις χρονοσειρές τιμής της μεγαβατώρας δοκιμάστηκαν επίσης look back window με εύρος 1 έως 24 ώρες. Τέλος, για την αρχικοποίηση των βαρών δοκιμάστηκε random state = 0 και για το LightGBM δεν χρησιμοποιήθηκε κάρτα γραφικών. Μέσος χρόνος εκπαίδευσης του μοντέλου είναι τα 3 λεπτά.

3.5.2 N-Beats

Αντίθετα με το προηγούμενο μοντέλο και γενικά σε οποιοδήποτε μοντέλο βαθιάς μάθησης (DL) η αρχιτεκτονική του μοντέλου παίζει ιδιαίτερο ρόλο στην απόδοση του ανάλογα με το εκάστοτε πρόβλημα. Στον Πίνακα 2 παρουσιάζονται οι διαφορετικές αρχιτεκτονικές σε αύξουσα σειρά πολυπλοκότητας.

N-Beats Architecture		
Δομικά στοιχεία N-Beats Έκδοσης	Έκδοση 1η	Έκδοση 2η
Number of stacks	20	30
Number of blocks	1	1
Number of layers	4	4
Layer widths	64	512
Expansion coefficient dimension	5	5

Πίνακας 2: Αρχιτεκτονικές του N-Beats και οι υπερπαραμέτροι του

Τα ανωτέρω δομικά στοιχεία έχουν αναλυθεί στην ενότητα 2.5.1 Δομικά στοιχεία του N-BEATS. Επίσης, προσδιορίστηκε για το εάν το μοντέλο αποδίδει καλύτερα με κανονικοποιημένα η δεδομένα σε φυσική τιμή.

Το μοντέλο αυτό είναι ικανό να δεχθεί μόνο παρελθοντικές συμμεταβλητές. Για αυτό το λόγο δόθηκαν ως τέτοιες, όπως και στο προηγούμενο μοντέλο οι:

- 1) Χρονικές συμμεταβλητές
- 2) Χρονοσειρά τιμής μεγαβατώρας
- 3) Χρονοσειρά καιρού (μόνο για Πορτογαλία)

Σε κάθε μία από αυτές δοκιμάστηκαν διαφορετικά παράθυρα προηγούμενης γνώσης και για τις τρεις χώρες. Μέσος χρόνος εκπαίδευσης του N-Beats είναι τα 10 λεπτά.

3.5.3 TCN

Και αυτό το μοντέλο αποτελεί μοντέλο βαθιάς μάθησης (DL). Παρακάτω, στον Πίνακα 3 παρατίθενται οι δύο διαφορετικές αρχιτεκτονικές που ελέγχθηκαν.

TCN Architecture		
Δομικά στοιχεία TCN Έκδοσης	Έκδοση 1η	Έκδοση 2η
Kernel size	3	5
Number of filters	3	5
Dilation base	2	3
Convolutional layers	auto	auto

Πίνακας 3: Αρχιτεκτονικές του TCN και οι υπερπαραμέτροι του

Όπως αναλύθηκε και στην ενότητα 2.6.1 Διεσταλμένη Αιτιώδης Συνέλιξη είναι μείζονος σημασίας τα συνελκτικά επίπεδα να είναι τόσο ώστε να καλύπτεται το δεκτικό πεδίο (receptive field). Συγκεκριμένα, όσο μεγαλώνει το παράθυρο προηγούμενης γνώσης τόσο μεγαλώνουν τα συνελκτικά επίπεδα. Επίσης, προσδιορίστηκε για το εάν το μοντέλο αποδίδει καλύτερα με κανονικοποιημένα η δεδομένα σε φυσική τιμή.

Το μοντέλο αυτό είναι ικανό να δεχθεί μόνο παρελθοντικές συμμεταβλητές. Για αυτό το λόγο δόθηκαν τα:

- 3) Χρονικές συμμεταβλητές
- 4) Χρονοσειρά τιμής μεγαβατώρας
- 5) Χρονοσειρά καιρού (μόνο για Πορτογαλία)

Σε κάθε μία από αυτές δοκιμάστηκαν διαφορετικά παράθυρα προηγούμενης γνώσης και για τις τρεις χώρες. Μέσος χρόνος εκπαίδευσης του TCN είναι τα 15 λεπτά.

Τέλος, στα μοντέλα βαθιάς μάθησης, N-Beats, TCN χρησιμοποιήθηκε Early Stopping της βιβλιοθήκης pytorch lightning. Συγκεκριμένα, ελεγχόταν κάθε 10 εποχές εάν το MAPE στο σύνολο αξιολόγησης (validation set) 2020-2021 μειωνόταν κατά 6·10⁻⁶, σε σχέση με όσο ήταν δέκα εποχές πριν. Εάν αυτό συνέβαινε τότε η διαδικασία εκπαίδευσης συνεχιζόταν, εάν όχι τότε σταμάτησε η διαδικασία εκπαίδευσης. Με αυτό τον τρόπο προσδιορίζεται ο αριθμός των εποχών εκπαίδευσης του Averaging Ensemble που έπεται.

3.5.4 Ensemble

Αφού προσδιορίστηκαν τα ακριβέστερα μοντέλα με τον καλύτερο συνδυασμό εξωτερικών μεταβλητών, για τα τρία μοντέλα που αναλύθηκαν, LightGBM (ML), N-Beats (DL) και TCN (DL), έγινε μία διαδικασία meta learning. Μέσω της χρήσης Early stopping της βιβλιοθήκης pytorch lightning προσδιορίστηκαν οι εποχές του εκάστοτε μοντέλου ώστε να μην γίνει overfitting για τα μοντέλα βαθιάς μάθησης που χρησιμοποιήθηκαν. Στη συνέχεια δόθηκαν στο μοντέλο soft ensemble ακριβώς οι εποχές που υπέδειξε η παραπάνω διαδικασία και εκπαιδεύτηκε το μοντέλο στα 5 έτη 2016-2020. Στη συνέχεια αξιολογήθηκε το μοντέλο αυτό στο έτος 2021 (test set).

Όπως αναλύθηκε και στην ενότητα 2.7.1 Averaging Ensemble, χρησιμοποιήθηκε η τεχνική Soft Ensemble με μέσο όρο (averaging) ώστε να εξαιρεθεί η μεροληψία των βαρών των μοντέλων. Αυτό επιτεύχθηκε με διαφορετική αρχικοποίηση τους (initialization of the weights) με διαφορετικά random state. Συγκεκριμένα, για το ensemble κάθε μοντέλου χρησιμοποιήθηκαν 20 μοντέλα, το καθένα με διαφορετική αρχικοποίηση βαρών και με σταθεροποιημένες τις εποχές του. παρακάτω στον Πίνακα 4 φαίνεται η αρχικοποίηση. Οι χρόνοι των μοντέλων φαίνεται στους Πίνακα 6: Αποτελέσματα N-Beats και Πίνακα 7: Αποτελέσματα μοντέλου TCN.

```
random_state_list= [0,42,35,2,37,38,41,43,40,50,36,39,45,12,3,1,17,7,44,45]
```

Πίνακας 4: η αρχικοποίηση των βαρών στη διαδικασία soft ensemble

3.6 Αξιολόγηση Μοντέλων

Η αξιολόγηση (evaluation) των μοντέλων όπως έχει ήδη προαναφερθεί, έγινε τόσο στο μοντέλο μηχανικής μάθησης (ML) όσο και στα μοντέλα βαθιάς μάθησης (DL) στο έτος από 01/03/2021 έως 28/02/2022, στο οποίο δεν έχουν εκπαιδευτεί, για το φορτίο ηλεκτρικής ενέργειας

των Πορτογαλία, Ισπανία και Ελλάδα. Για τα μοντέλα βαθιάς μάθησης εκτός του Ensemble χρησιμοποιήθηκε το έτος 2020 ως έτος επικύρωσης (validation set) ώστε να προσδιοριστούν οι υπερπαράμετροι του.

Για την αξιολόγηση της απόδοσης των μοντέλων υπάρχουν πολλές μετρικές που έχουν χρησιμοποιηθεί σε διάφορες εφαρμογές. Κάποιες από τις πλέον γνωστές είναι το Mean Absolute Error (MAE), το Mean Squared Error (MSE), το Mean Absolute Percentage Error (MAPE), το Root Mean Square Error (RMSE), το Symmetric Mean Absolute Percentage Error (SMAPE), και το Mean Absolute Scaled Error (MASE) όπως υποδεικνύεται από [56].

Στην παρούσα εργασία χρησιμοποιείται ως μετρική αξιολόγησης της απόδοσης των μοντέλων το MAPE.

$$MAPE = \frac{1}{m} \sum_{i=1}^m \left| \frac{Y_t - F_t}{Y_t} \right| 100\% , \quad (2.3)$$

όπου m είναι ο αριθμός των παρατηρήσεων προς αξιολόγηση, Y_t και F_t οι πραγματικές τιμές φορτίου και οι προβλέψεις των μοντέλων αντίστοιχα.

Όπως έχει ήδη αναφερθεί ο σκοπός της εργασίας είναι η πρόβλεψη του ωριαίου φορτίου της επόμενης μέρας (STLF). Για αυτό το λόγο χρησιμοποιήθηκε η μέθοδος backtesting. Η μέθοδος αυτή χρησιμοποιείται ώστε να ανανεώνονται ανά τακτά χρονικά διαστήματα οι πραγματικές τιμές του φορτίου εφόσον αυτές είναι γνωστές. Έτσι το παραγόμενο σφάλμα της προηγούμενης πρόβλεψης δεν μεταφέρεται και επομένως δεν διογκώνει το σφάλμα της επόμενης πρόβλεψη. Το χρονικό αυτό διάστημα καθορίζεται από την πρόβλεψη της επόμενης μέρας (day ahead forecast) δηλαδή από τον ορίζοντα πρόβλεψης, 24 τιμές στην περίπτωσή μας. Ένα παράδειγμα ώστε να γίνει σαφής η παραπάνω διαδικασία είναι: Έστω ότι ο διαχειριστής διανομής ενέργειας της εκάστοτε χώρας αποφασίζει να συμμετέχει στο χρηματιστήριο ενέργειας. Για αυτό το λόγο χρειάζεται η ωριαία πρόβλεψη φορτίου της επόμενης μέρας ώστε να καταρτιστεί η οικονομική πολιτική του. Όταν παρέλθει όμως μία μέρα, οι πραγματικές τιμές φορτίου (ground truth) γνωστοποιούνται, άρα μπορούν να χρησιμοποιηθούν για την εκπαίδευση των μοντέλων. Αυτή η μέθοδος εφαρμόστηκε σε όλα τα προαναφερθέντα μοντέλα DL και ML.

Κεφάλαιο 4 - Αποτελέσματα-Συμπεράσματα

Στο παρόν Κεφάλαιο παρουσιάζονται τα αποτελέσματα και συμπεράσματα των πειραμάτων που διενεργήθηκαν. Τόσο τα σύνολα δεδομένων φορτίου της Πορτογαλίας, Ισπανία και Ελλάδα όσο και τα σύνολα των συμμεταβλητών που χρησιμοποιήθηκαν αποτελούνται από ωριαία δεδομένα από 01/03/2016 έως και 28/02/2022 (6 έτη). Για τον διαχωρισμό του συνόλου δεδομένων (dataset split), στο μοντέλο μηχανικής μάθησης (ML), LightGBM, δόθηκε ως σύνολο εκπαίδευσης (train set) τα πέντε έτη από 01/03/2016 έως 28/02/2021 ενώ στα μοντέλα βαθιάς μάθησης (DL), NBEATS και TCN δόθηκαν τα τέσσερα έτη από 01/03/2016 έως 29/02/2020 ως σύνολο εκπαίδευσης και το έτος από 01/03/2020 έως 28/02/2021 ως σύνολο επικύρωσης (validation set). Τέλος, ως σύνολα τεστ (test sets) για όλα τα μοντέλα χρησιμοποιήθηκε το έτος από 01/03/2021 έως 28/02/2022. Οι γραφικές παραστάσεις που απεικονίζονται στο παρόν Κεφάλαιο βρίσκονται στο αρχείο “Results .ipynb”¹⁴.

4.1 Αξιολόγηση Μοντέλων

4.1.1 Αποτελέσματα του αλγορίθμου LightGBM

Παρακάτω, παρατίθεται ο Πίνακας 5 των αποτελεσμάτων των πειραμάτων με το μοντέλο μηχανικής μάθησης (ML), Lightgbm, με τις διάφορες υπερ παραμέτρους και συμμεταβλητές που χρησιμοποιήθηκαν.

¹⁴ https://github.com/giannisiso/Thesis_Short_Term_Load_Forecasting

LightGBM					
	Παράθυρο προηγούμενης γνώσης Φορτίου	Συμμεταβλητές Παρελθόντος	Συμμεταβλητές Μέλλοντος	Παράθυρο προηγούμενης γνώσης Συμμεταβλητών	MAPE (%)
Πορτογαλία	168	None	None	-	2.7
	120	Price	temporal	(4,2)	2.52
	120	None	temporal	(4,2)	2.01
Ισπανία	168	None	None	-	2.37
	120	Price	temporal	(4,2)	1.98
	120	None	temporal	(4,2)	1.79
Ελλάδα	168	None	None	-	3.45
	120	Price	temporal	(4,2)	3.02
	120	None	temporal	(4,2)	3.02

Πίνακας 5: Αποτελέσματα πειραμάτων LightGBM

Αρχικά, εξετάστηκε εάν το μοντέλο αποδίδει καλύτερα με κανονικοποιημένο φορτίο. Όπως αναμενόταν το μοντέλο αποδίδει καλύτερα με τις φυσικές του τιμές καθώς τα δέντρα αποφάσεων δεν χρησιμοποιούν συναρτήσεις ενεργοποίησης, επομένως στα παραπάνω πειράματα του Πίνακα 5 χρησιμοποιήθηκαν φυσικές τιμές φορτίου.

Ως μοντέλο βάσης για κάθε χώρα χρησιμοποιήθηκε το μοντέλο με 168 τιμές, μήκος παραθύρου προηγούμενης γνώσης (1 εβδομάδα). Δοκιμάστηκαν σε κάθε χώρα τα παράθυρα προηγούμενης γνώσης [72, 96, 120, 144, 168, 192, 216, 240, 264, 288, 312, 336]. Η καλύτερες αποδόσεις προέκυψαν με παράθυρο προηγούμενης γνώσης 120, δηλαδή 5 ημέρες. Πιθανός λόγος για αυτό είναι το τρόπος με τον οποίο το συγκεκριμένο μοντέλο δημιουργεί τα δέντρα. Μεγαλύτερο παράθυρο προηγούμενης γνώσης θα συνεπαγόταν μεγαλύτερα “σπασίματα” δέντρου (υποπροβλήματα), όπου το μοντέλο θα “χανόταν”.

Οι συμμεταβλητές χρόνου έχουν μεγάλη επίδραση στο LightGBM. Λόγω της έλλειψης μνήμης του μοντέλου που προέρχεται από την αρχιτεκτονική του δεν υπάρχει η δυνατότητα στο μοντέλο να λαμβάνει υπόψη του τον χρόνο. Οπότε η κωδικοποίηση του χρόνου η οποία φαίνεται στον Πίνακα 1 της ενότητας 3.2.4 Άλλα Δεδομένα δίνει την δυνατότητα στο μοντέλο να εκπαιδευτεί συμπεριλαμβάνοντας τον. Ιδιαίτερα η τελευταία μεταβλητή δηλαδή η δυαδική κωδικοποίηση της εργάσιμης ή όχι μέρας μειώνει αισθητά την μετρική απόδοσης MAPE. Όπως αναλύθηκε και στην ενότητα 2.4.3 Χαρακτηριστικά αλγορίθμου LightGBM, ο έξυπνος τρόπος

που ο αλγόριθμος του LightGBM πραγματοποιεί feature engineering (exclusive feature bundling), κωδικοποιώντας εκ νέου τις συμμεταβλητές μειώνοντας την διάσταση τους δίνει τη δυνατότητα στο μοντέλο να έχει υψηλές αποδόσεις με μεγάλο όγκο δεδομένων. Ακόμα, χρησιμοποιήθηκαν διάφορα παράθυρα για την εύρεση του βέλτιστου συνδυασμού προηγούμενων και επόμενων παρατηρήσεων (lags future covariates) που θα βλέπει το μοντέλο για την κάθε τιμή πρόβλεψης. Παρατηρήθηκε πως 4 παρελθοντικές και 2 μελλοντικές τιμές βοηθούν το μοντέλο να πετύχει καλύτερη επίδοση. Αυτό το αποτέλεσμα θα ερμηνευτεί αναλυτικά στην ενότητα 4.3 Αποτελέσματα Συμμεταβλητής Καιρού που χρησιμοποιήθηκε σαν συμμεταβλητή ο καιρός στα δεδομένα φορτίου της Πορτογαλίας.

Ακόμα, για τις παρελθοντικές συμμεταβλητές δηλαδή την τιμή της μεγαβατώρας χρησιμοποιήθηκαν ως τιμές που βλέπει το μοντέλο από μία έως τέσσερις. Παρατηρήθηκε πως η αύξηση της πολυπλοκότητας του ως προς αυτή την παράμετρο δεν αλλάζει την απόδοση του μοντέλου. Ακόμα, παρατηρήθηκε πως στις Πορτογαλία και Ισπανία χειροτερεύει αισθητά την απόδοση τους. Αυτό συμβαίνει διότι η ραγδαία αύξηση των τιμών του τελευταίου έτους δεν συνεπάγεται αλλαγή στο καταναλισκόμενο φορτίο. Αντίθετα στην περίπτωση πρόβλεψης του φορτίου της Ελλάδας, παρατηρείται ίδια απόδοση με και χωρίς τη χρήση της συμμεταβλητής της μεγαβατώρας. Αυτό εξηγείται καθώς στο σύνολο τεστ έχουμε ραγδαία αύξηση του καταναλισκόμενου φορτίου τους καλοκαιρινούς μήνες λόγω του μνημειώδους καύσιωνα το καλοκαίρι του 2021. Όμως ενώ η πρόβλεψη του καλοκαιρινού φορτίου γίνεται ακριβέστερη, χειροτερεύει η πρόβλεψη του φορτίου τους μήνες του φθινοπώρου και του χειμώνα.

Συμπερασματικά, το LightGBM προβλέπει καλύτερα τις εποχές όταν το καταναλισκόμενο φορτίο είναι χαμηλό δηλαδή την άνοιξη και το φθινόπωρο όπου τα ηλεκτρικά μηχανήματα ψύξης ή θέρμανσης δεν είναι σε λειτουργία.

Επιπλέον, αξίζει να σημειωθεί πως ο τρόπος με τον οποίο ο αλγόριθμος του LightGBM κάνει ελαχιστοποίηση του σφάλματος λαμβάνοντας υπόψη τα μεγαλύτερα σφάλματα κάθε φορά και κάνοντας δειγματοληψία στα μικρότερα, δίνει την δυνατότητα μιας πρόβλεψης με πολύ καλή απόδοση σε αρκετά μικρό χρόνο χωρίς την ανάγκη μεγάλης υπολογιστικής ισχύος.

Τέλος, όπως και στα παρακάτω μοντέλα βαθιάς μηχανικής μάθησης, έγινε πείραμα Ensemble του καλύτερου μοντέλου LightGBM με διαφορετική αρχικοποίηση (random state). Τα αποτελέσματα που προέκυψαν είναι ελαφρώς χειρότερα από αυτά που φαίνονται στον Πίνακας 5. Αυτό εδράζεται στο γεγονός πως ο αλγόριθμος του LightGBM κάνει Ensemble των διάφορων υποπροβλημάτων (δέντρων) που χωρίζει το πρόβλημα και στο τέλος συγκλίνει στη λύση.

4.1.2 Αποτελέσματα της βαθιάς αρχιτεκτονικής N-BEATS

Ο Πίνακας 6 απεικονίζει αποτελέσματα πειραμάτων με το μοντέλο βαθιάς μάθησης N-Beats, με τις διάφορες παραμέτρους και συμμεταβλητές που χρησιμοποιήθηκαν.

N-BEATS						
	Αρχιτεκτονική	Παράθυρο προηγούμενης γνώσης Φορτίου	Συμμεταβλητές Παρελθόντος	Epochs	GPU (min)	MAPE (%)
Πορτογαλία	1η	168	None	84	8	2.97
	2η	240	None	42	5	3.15
	1η	240	Price+hol	108	13	6.42
	1η	240	None	100	9	2.456
	Ensemble	240	None	100	146	2.036
Ισπανία	1η	168	None	73	11	2.33
	2η	240	None	44	6	2.41
	1η	240	Price+hol	104	14	8.34
	1η	240	hol	80	10	2.05
	Ensemble	240	hol	80	170	1.748
Ελλάδα	1η	168	None	88	12	3.61
	2η	240	None	26	5	3.798
	1η	240	Price+hol	108	15	5.01
	1η	240	None	100	15	3.39
	Ensemble	240	None	100	190	3.089

Πίνακας 6: Αποτελέσματα N-Beats

Αρχικά, εξετάστηκε εάν το μοντέλο βαθιάς μάθησης N-Beats (Neural Basis Expansion Analysis for Time Series forecasting) λειτουργεί καλύτερα με κανονικοποιημένο φορτίο. Στο συγκεκριμένο μοντέλο, η επίδοση βελτιώνεται αισθητά με κανονικοποιημένα δεδομένα λόγω ύπαρξης συνάρτησης ενεργοποίησης (ReLU). Επομένως σε όλα πειράματα του Πίνακας 6 χρησιμοποιήθηκαν κανονικοποιημένα δεδομένα.

Ως μοντέλο βάσης για κάθε χώρα χρησιμοποιήθηκε το μοντέλο με 168 τιμές, μήκος παραθύρου προηγούμενης γνώσης (1 εβδομάδα) και καμία συμμεταβλητή. Δοκιμάστηκαν σε κάθε χώρα τα παράθυρα προηγούμενης γνώσης [72, 96, 120, 144, 168, 192, 216, 240, 264, 288, 312, 336]. Η καλύτερες αποδόσεις προέκυψαν με παράθυρο προηγούμενης γνώσης 240, δηλαδή 10 ημέρες. Στο ίδιο αποτέλεσμα όσον αφορά το παράθυρο προηγούμενης γνώσης κατέληξε και το άρθρο [36] αφού για δεδομένα ωριαίας ανάλυσης το παράθυρο προηγούμενης γνώσης προτείνεται να είναι δεκαπλάσιο του ορίζοντα πρόβλεψης ($10 \cdot 24 = 240$). Επιπλέον παρατηρήθηκε πως με την αύξηση του παραθύρου προηγούμενης γνώσης έως και 14 μέρες τα αποτελέσματα παραμένουν το ίδιο ικανοποιητικά. Όμως τα μοντέλα αυτά έχουν χειρότερες αποδόσεις όσον αφορά στο χρόνο εκπαίδευσης και την πολυπλοκότητα τους.

Το N-Beats, έχει δυνατότητα χρήσης μόνο παρελθοντικών συμμεταβλητών όπως έχει προαναφερθεί. Οι συμμεταβλητές χρόνου έχουν αρνητική επίδραση στο N-Beats, καθώς το μοντέλο μπορεί να αναγνωρίσει τις χρονικές συσχετίσεις μέσω της διαδικασίας εκπαίδευσης και του τρόπου που κάνει “πίσω” πρόβλεψη (backcast). Η μόνη χρονική συμμεταβλητή που δεν χειροτερεύει εμφανώς τα αποτελέσματα είναι αυτή της εργάσιμης μέρας (hol). Συγκεκριμένα, στην χρονοσειρά φορτίου της Ισπανίας καλυτερεύει τα αποτελέσματα κατά 0.11% (MAPE = 2.149%) σε σχέση με το μοντέλο χωρίς καμία συμμεταβλητή. Για το γεγονός αυτό μάλλον ευθύνεται η στατικότητα της χρονοσειράς της Ισπανίας. Ακόμα, η συμμεταβλητή της τιμής της μεγαβατώρας έχει εξαιρετικά αρνητική επίδραση στην απόδοση του μοντέλου, επειδή το μοντέλο εκπαιδεύεται σε δεδομένα που έχουν μέση τιμή (Ευρώ/MWh) πολύ μικρότερη από αυτή του συνόλου τεστ.

Όσον αφορά την αρχιτεκτονική του μοντέλου, το πιο ρηχό μοντέλο (1η αρχιτεκτονική) παράγει καλύτερα αποτελέσματα. Παρότι οι χρόνοι εκπαίδευσης του είναι μεγαλύτεροι (μεγαλύτερος χρόνος χρήσης της GPU) και η προσαρμογή (fit) στο σύνολο δεδομένων δυσκολότερη (περισσότερες εποχές), το μοντέλο αποδίδει καλύτερα. Με τη χρήση Early Stopping αποφεύχθηκε η υπερπροσαρμογή (overfitting), και προσδιορίστηκαν οι εποχές εκπαίδευσης του Average Ensemble. Με τη σειρά της η τεχνική αυτή παράγει εμφανώς καλύτερα αποτελέσματα όπως φαίνεται στον Πίνακα 6 μειώνοντας την μετρική απόδοσης (MAPE) κατά 0.3% - 0.4%, προσεγγίζοντας έτσι τις επιδόσεις του LightGBM .

4.1.3 Αποτελέσματα της βαθιάς αρχιτεκτονικής TCN

Στον Πίνακα 7 παρατίθενται τα αποτελέσματα των πειραμάτων με το μοντέλο βαθιάς μάθησης (DL), TCN, καθώς και οι διάφορες υπερ παράμετροι και συμμεταβλητές που χρησιμοποιήθηκαν.

TCN						
	Αρχιτεκτονική	Παράθυρο προηγούμενης γνώσης Φορτίου	Συμμεταβλητές Παρελθόντος	Epochs	GPU (min)	MAPE (%)
Πορτογαλία	2η	168	None	116	14	3.51
	1η	240	None	232	36	3.48
	2η	288	Price+hol	148	17	4.24
	2η	288	None	194	23	2.94
	Ensemble	288	None	194	250	2.63
Ισπανία	2η	168	None	104	15	3.95
	1η	240	None	238	41	4.59
	2η	288	Price+hol	152	20	5.68
	2η	288	None	192	15	2.91
	Ensemble	288	None	192	225	2.6
Ελλάδα	2η	168	None	128	16	4.8
	1η	288	None	276	37	4.34
	2η	288	Price+hol	140	21	6.45
	2η	288	None	198	23	4.2
	Ensemble	288	None	198	290	3.88

Πίνακας 7: Αποτελέσματα μοντέλου TCN

Αρχικά, διαπιστώθηκε εάν το μοντέλο βαθιάς μάθησης TCN (Temporal Convolutional Neural Network) λειτουργεί καλύτερα με κανονικοποιημένο φορτίο. Στο συγκεκριμένο μοντέλο, η απόδοση αυξάνεται εμφανώς με τις φυσικές τιμές των δεδομένων. Συγκεκριμένα, η μετρική αξιολόγησης MAPE μειώνεται κατά 0.42%. Μία πιθανή εξήγηση για αυτό το φαινόμενο είναι και ταυτόχρονα ένα βασικό χαρακτηριστικό του μοντέλου αυτού, η ικανότητα του να βλέπει το

παράθυρο προηγούμενης γνώσης (εντός του receptive field) και μόνο αυτό για να κάνει πρόβλεψη. Επομένως σε όλα πειράματα του Πίνακα 7 χρησιμοποιήθηκαν δεδομένα στις φυσικές τους τιμές.

Ως μοντέλο βάσης για κάθε χώρα χρησιμοποιήθηκε το μοντέλο με 168 τιμές (lookback window) (1 εβδομάδα). Δοκιμάστηκαν σε κάθε χώρα τα παράθυρα προηγούμενης γνώσης [72, 96, 120, 144, 168, 192, 216, 240, 264, 288, 312, 336]. Η καλύτερες αποδόσεις προέκυψαν με παράθυρο προηγούμενης γνώσης 288, δηλαδή 12 ημέρες, το μεγαλύτερο lookback window από τα μοντέλα που χρησιμοποιήθηκαν σε αυτή την εργασία. Ο λόγος είναι το χαρακτηριστικό που αναλύθηκε παραπάνω. Αναλυτικότερα, η έλλειψη μνήμης του μοντέλου, δηλαδή η παραγωγή πρόβλεψης μόνο από δεδομένα που ανήκουν στο receptive field, καθιστούν την επιλογή του παραθύρου προηγούμενης γνώσης κρίσιμη. Επομένως, πρέπει το receptive field να έχει τέτοιο μέγεθος που να περιλαμβάνει την απαιτούμενη πληροφορία, όμως ταυτόχρονα να μην είναι τόσο μεγάλο ώστε να μην μπορεί να την αφομοιώσει.

Το TCN, όπως και το N-Beats, μπορεί να εκπαιδευτεί μόνο με παρελθοντικές συμμεταβλητές. Επομένως, όπως και προηγουμένως οι συμμεταβλητές χρόνου επιδρούν αρνητικά, χειροτερεύοντας την απόδοση του μοντέλου. Το ίδιο ισχύει και για την χρονοσειρά της τιμής της μεγαβατώρας, όπου τα μοντέλα που εκπαιδεύτηκαν με αυτή παρουσιάζουν άνοδο ίδια με αυτή της τιμής.

Παρότι το TCN, έχει την δυνατότητα παραλληλοποίησης διεργασιών με χρήση GPU, ο χρόνος εκπαίδευσης είναι αισθητά μεγαλύτερος από το N-Beats. Ακόμα, η βαθύτερη αρχιτεκτονική του TCN, δηλαδή με μέγεθος πυρήνα 5, βάση διαστολής 3, 5 φίλτρα και 4 επίπεδα (όπως αυτά ορίστηκαν για να είναι όλο το παράθυρο προηγούμενης γνώσης εντός του receptive field), οδηγεί σε καλύτερα αποτελέσματα από την πιο “ρηχή”. Οι εποχές που χρειάζονται, όπως αυτές καθορίστηκαν από το Early Stopping, είναι λιγότερες για πιο βαθιές αρχιτεκτονικές και ο χρόνος εκπαίδευσης μικρότερος, καθότι το πιο βαθύ μοντέλο μπορεί να προσαρμοστεί καλύτερα στο σύνολο δεδομένων και να συγκλίνει ταχύτερα.

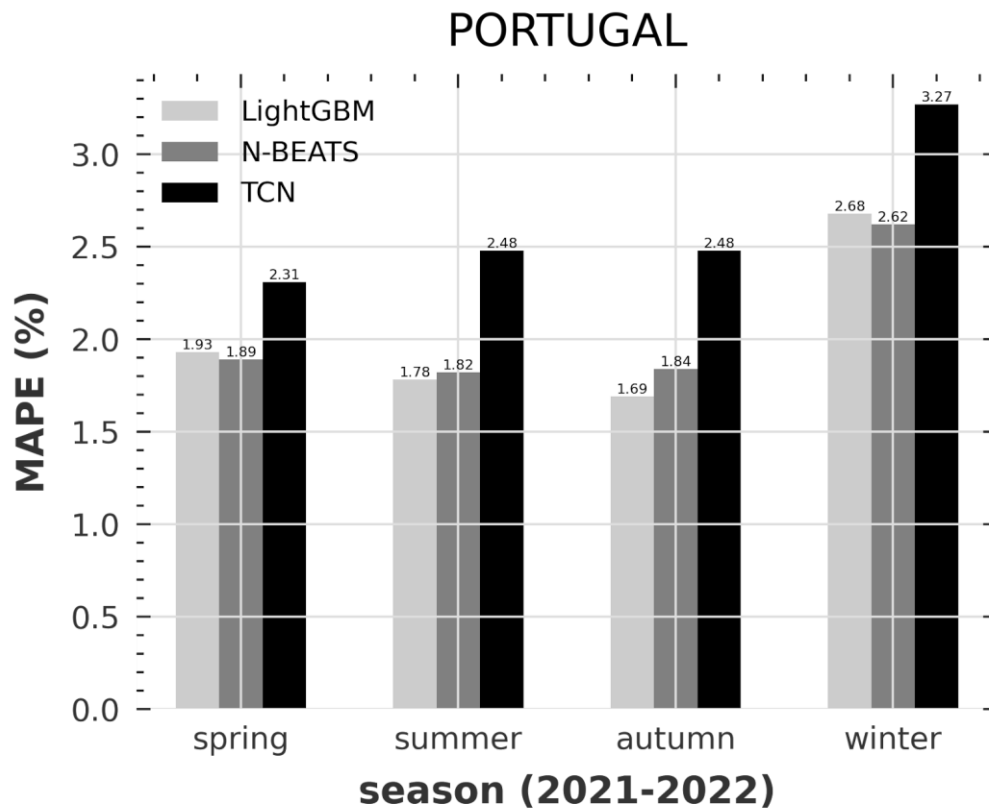
Τέλος, όπως και στο N-Beats, προσδιορίστηκαν οι εποχές εκπαίδευσης του καλύτερου μοντέλου για το Average Ensemble. Με τη σειρά της η τεχνική αυτή οδηγεί σε καλύτερα αποτελέσματα μειώνοντας την μετρική απόδοσης (MAPE) κατά 0.3%.

Γενικότερα, παρατηρείται χειρότερη αντιμετώπιση του μοντέλου αυτού σε χρονικές περιόδους της χρονοσειράς που έχουμε είτε αισθητά αυξημένο είτε μειωμένο καταναλισκόμενο φορτίο δηλαδή τον χειμώνα ή το καλοκαίρι. Όπως αναλύεται παρακάτω, το TCN έχει τα χειρότερα αποτελέσματα σε σχέση με τα προηγούμενα μοντέλα.

4.2 Οπτικοποίηση αποτελεσμάτων, συγκριτική ανάλυση και υβριδικό Ensemble

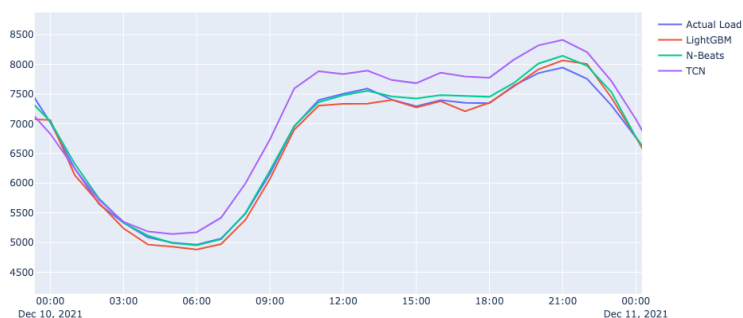
4.2.1 Σύγκριση Αποτελεσμάτων στην Πορτογαλία

Στην Εικ. 22 απεικονίζονται το σφάλμα των μοντέλων ανάλογα με τις εποχές του χρόνου, άνοιξη, καλοκαίρι, φθινόπωρο, χειμώνας στην Πορτογαλία.



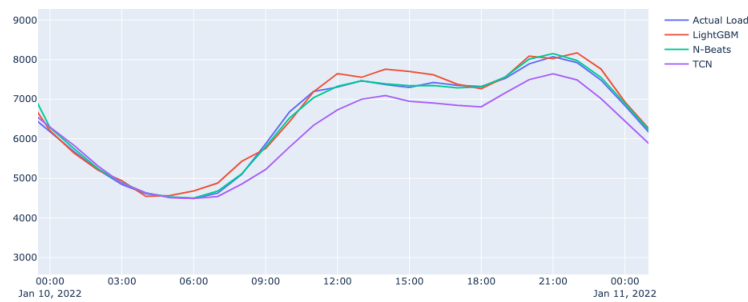
Εικ. 22: Σύγκριση μοντέλων στην χρονοσειρά της Πορτογαλίας

Όπως φαίνεται από τα σφάλματα των μοντέλων, το φορτίο κατά τους χειμερινούς μήνες στην Πορτογαλία δεν μπορεί να προβλεφθεί τόσο καλά όσο τις άλλες εποχές. Αυτό μπορεί να οφείλεται σε πολλούς παράγοντες. Ένας πιθανός λόγος είναι ότι σε σχέση με τις υπόλοιπες εποχές, τον χειμώνα το καταναλισκόμενο φορτίο της Πορτογαλίας παίρνει τις μέγιστες τιμές του. Επομένως, από τα παραπάνω αποτελέσματα προέκυψε η ανάγκη περαιτέρω διερεύνησης. Για αυτό τον λόγο αναλύθηκαν γραφικά τα αποτελέσματα με χρήση της βιβλιοθήκης `plotly.express`¹⁵.



Εικ. 23: Στιγμιότυπο πρόβλεψης στην Πορτογαλία (1)

¹⁵ <https://plotly.com/python/line-charts/>



Εικ. 24: Στιγμιότυπο πρόβλεψης στην Πορτογαλία (2)

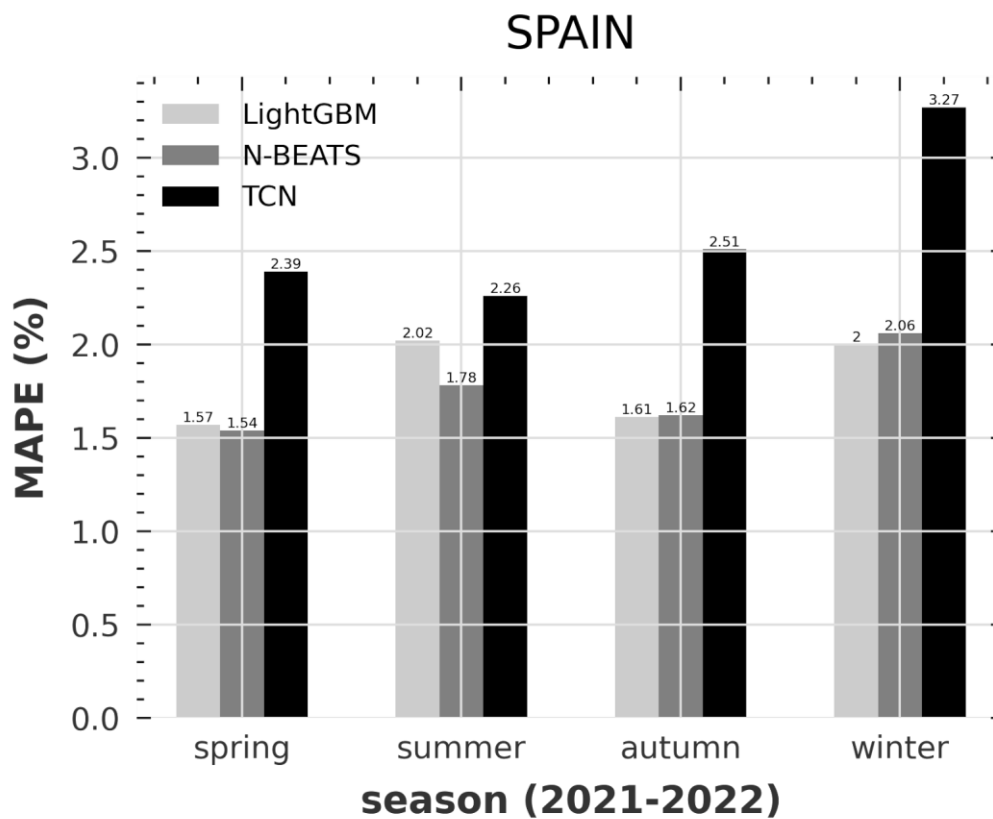
Στις Εικ. 23, 24 και φαίνονται δύο παραδείγματα ημερών, 10/12/21 και 10/01/22 όπου η καμπύλη φορτίου έχει τυπική εικόνα σύμφωνα με το μέσο καταναλισκόμενο φορτίο το οποίο απεικονίζεται στην Εικ. 14. Όπως φαίνεται στα παραπάνω παράδειγματα η πρόβλεψη του TCN έχει εμφανώς χειρότερα αποτελέσματα από τα άλλα δύο μοντέλα, όπως απεικονίζεται αυτό και στην Εικ. 22. Αντίθετα, τα μοντέλα, μηχανικής μάθησης LightGBM και βαθιάς μάθησης N-Beats έχουν αρκετά καλύτερη απόδοση με μικρότερα σφάλματα σε όλες τις εποχές. Επίσης παρατηρείται πως το μεν LightGBM κάνει καλύτερη πρόβλεψη χρονικές περιόδους που το φορτίο έχει ομαλές τιμές ενώ το N-Beats έχει καλύτερη απόδοση σε χρονικές περιόδους με αυξημένο ή μειωμένο αισθητά φορτίο. Τέλος, το μεν LightGBM τήνει να έχει σφάλματα υπερεκτίμησης ενώ το N-Beats υποεκτίμησης. Από τα παραπάνω αποφασίστηκε η χρήση Averaging Ensemble του LightGBM και N-Beats του οποίου το αποτέλεσμα παρουσιάζεται στον Πίνακας 8.

Η διαδικασία εκπαίδευσης και αξιολόγησης των μοντέλων για τις χρονοσειρές φορτίου της Πορτογαλίας υπάρχουν στο αρχείο “Portugal_with_weather_forecasts.ipynb”¹⁶.

4.2.2 Σύγκριση Αποτελεσμάτων στην Ισπανία

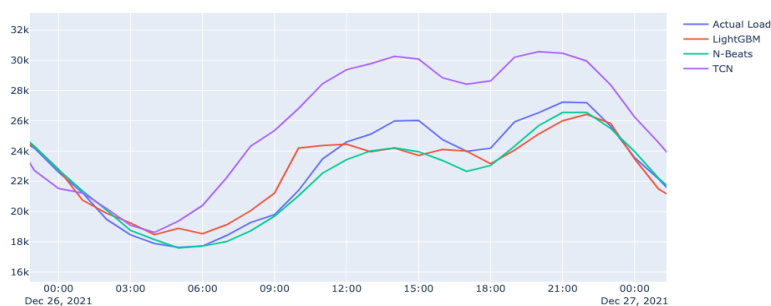
Στην Εικ. 25 απεικονίζονται το σφάλμα των μοντέλων ανάλογα με τις εποχές του χρόνου, άνοιξη, καλοκαίρι, φθινόπωρο, χειμώνας στην Ισπανία.

¹⁶ https://github.com/giannisiso/Thesis_Short_Term_Load_Forecasting

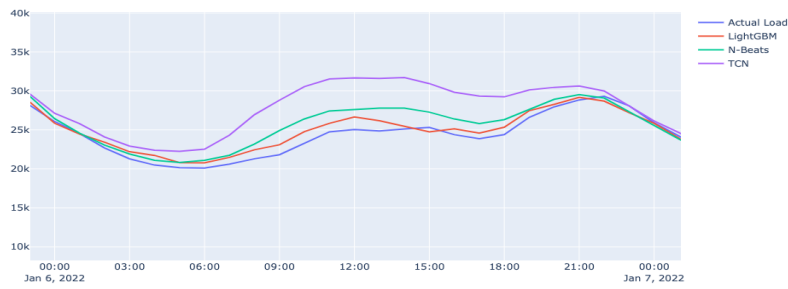


Εικ. 25: Σύγκριση μοντέλων στην χρονοσειρά της Ισπανίας

Και στην Ισπανία τα σφάλματα των μοντέλων κατά τους χειμερινούς μήνες είναι αυξημένα. Επιπλέον, οι καλοκαιρινοί μήνες προβλέπονται με χειρότερη ακρίβεια. Σε αυτό το γεγονός συνηγορεί η αύξηση του φορτίου που καταναλώνεται τις εποχές αυτές, και οι περίεργες αυξομειώσεις φορτίου όπως για παράδειγμα κατά τις μέρες αργίας των Χριστουγέννων και της Πρωτοχρονιάς.



Εικ. 26: Στιγμιότυπο πρόβλεψης στην Ισπανία (1)



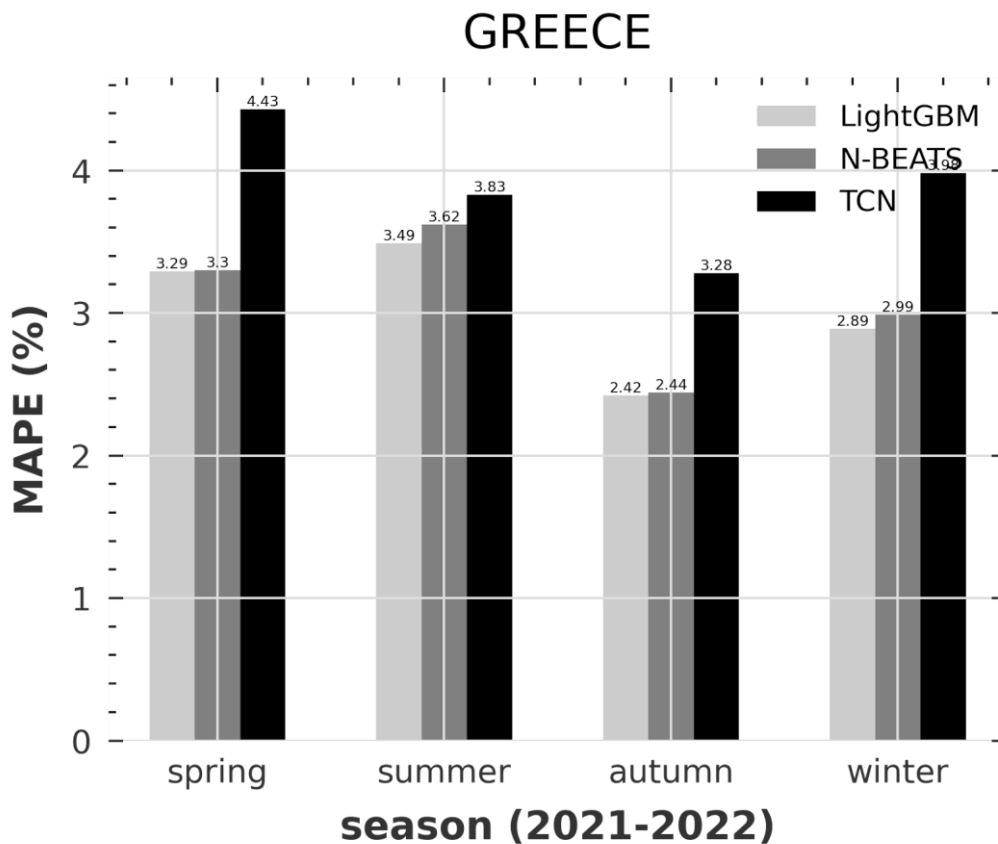
Εικ. 27: Στιγμιότυπο πρόβλεψης στην Ισπανία (2)

Στις Εικ. 26, 27 παρουσιάζονται δύο παραδείγματα ημερών, 26/12/2021 και 06/01/2022 του συνόλου τεστ της Ισπανίας μαζί με τις προβλέψεις των μοντέλων. Όπως και στην χρονοσειρά της Πορτογαλίας, το LightGBM και το N-Beats έχουν καλύτερες προβλέψεις από το TCN με διαφορά. Οπότε στο Averaging Ensemble του Πίνακα 8 συμπεριλήφθηκαν τα δύο αυτά μοντέλα.

Η διαδικασία εκπαίδευσης και αξιολόγησης των μοντέλων για τις χρονοσειρές φορτίου της Ισπανίας υπάρχουν στο αρχείο “SPAIN .ipynb”¹⁷

4.2.3 Σύγκριση Αποτελεσμάτων στην Ελλάδα

Τέλος, στην Εικ. 28 παρουσιάζεται το ραβδόγραμμα των σφαλμάτων των μοντέλων στην χρονοσειρά της Ελλάδας ανάλογα με τις εποχές του χρόνου.



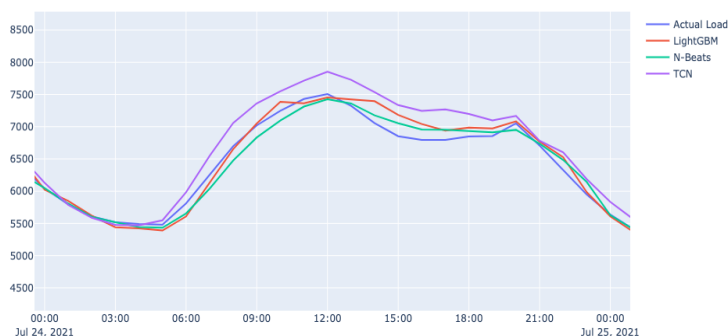
Εικ. 28: Σύγκριση μοντέλων στην χρονοσειρά της Ελλάδας

¹⁷ https://github.com/giannisiso/Thesis_Short_Term_Load_Forecasting

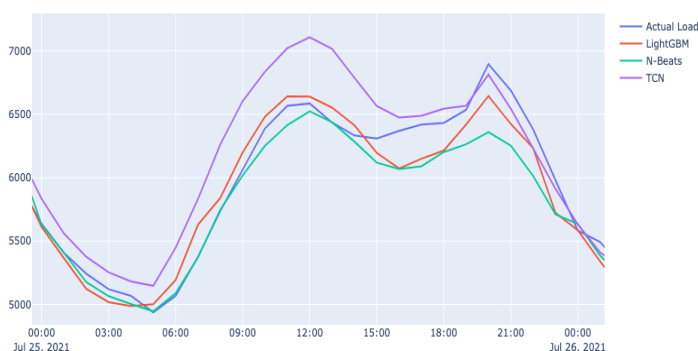
Η χρονοσειρά καταναλισκόμενου φορτίου της Ελλάδας έχει εμφανώς χειρότερα αποτελέσματα από τις άλλες δύο χώρες. Η προβλεπτική ικανότητα των μοντέλων είναι μειωμένη οδηγώντας σε μεγαλύτερα σφάλματα. Αυτό το γεγονός οφείλεται σε διάφορους παράγοντες όπως η ποιότητα των δεδομένων και η επεξεργασία τους αλλά και παράγοντες φύσης του πρότυπου της κατανομής του καταναλισκόμενου φορτίου, όπως η αυξημένη κατανάλωση ηλεκτρικής ενέργειας κατά τους καλοκαιρινούς μήνες. Τα μεγαλύτερα σφάλματα των μοντέλων παρατηρούνται κατά τους καλοκαιρινούς μήνες, όπου παρατηρείται πως το TCN έχει μεν χειρότερη απόδοση από τα άλλα μοντέλα αλλά το χάσμα τους είναι μικρότερο. Αντίθετα, οι προβλέψεις του TCN τις υπόλοιπες εποχές είναι αισθητά χειρότερες από τα άλλα δύο μοντέλα. Για αυτό διενεργήθηκε Averaging Ensemble των LightGBM και N-Beats και στην χρονοσειρά φορτίου της Ελλάδας.

Η διαδικασία εκπαίδευσης και αξιολόγησης των μοντέλων για τις χρονοσειρές φορτίου της Ελλάδας υπάρχουν στο αρχείο “GREECE .ipynb”¹⁸

Στις Εικ. 29, 30 παρουσιάζονται δύο παραδείγματα αποτυχίας σωστής πρόβλεψης του TCN σε σχέση με τα LightGBM και N-Beats τις ημέρες 24/07/2021 και 25/07/2021.



Εικ. 29: Στιγμιότυπο πρόβλεψης στην Ελλάδα (1)



Εικ. 30: Στιγμιότυπο πρόβλεψης στην Ελλάδα (2)

¹⁸ https://github.com/giannisiso/Thesis_Short_Term_Load_Forecasting

Στον Πίνακα 8 απεικονίζονται τα τελικά αποτελέσματα με το Ensemble των LightGBM και N-Beats της εκάστοτε χώρας, που αντιστοιχούν στα συγκεντρωτικά καλύτερα αποτελέσματα για τις χρονοσειρές φορτίου των Πορτογαλία, Ισπανία και Ελλάδα.

Hybrid Averaging Ensemble			
	Best LightGBM MAPE (%)	Best N-Beats MAPE (%)	Averaging Ensemble (LightGBM+N-Beats) MAPE (%)
Πορτογαλία	2.01	2.036	1.899
Ισπανία	1.79	1.748	1.671
Ελλάδα	3.02	3.089	2.892

Πίνακας 8: Τελικά αποτελέσματα Hybrid Averaging Ensemble (LightGBM+N-Beats)

4.3 Αποτελέσματα Συμμεταβλητής Καιρού

Τα δεδομένα καιρού όπως έχει προαναφερθεί ήταν διαθέσιμα μόνο για την Πορτογαλία. Αρχικά, διευκρινίζεται πως στο μοντέλο μηχανικής μάθησης LightGBM η ενσωμάτωση του καιρού ως παρελθοντική συμμεταβλητή οδηγεί σε χειρότερα αποτελέσματα. Αντίθετα, η χρήση του καιρού ως συμμεταβλητή μέλλοντος μειώνει ελάχιστα το σφάλμα. Πιο συγκεκριμένα, η χρήση της ταχύτητας του αέρα είτε με τις φυσικές του τιμές, είτε σε μορφή δυαδικής κωδικοποίησης πάνω και κάτω από κάποιο κατώφλι θερμοκρασίας χειροτερεύει την πρόβλεψη του μοντέλου, ενώ η θερμοκρασία με κανονικοποιημένες τιμές την καλύτερεύει.

Επιπρόσθετα, δόθηκαν στο μοντέλο οι θερμοκρασίες των μεγαλύτερων σε πληθυσμό πόλεων της Πορτογαλίας: Λισαβόνα, Πόρτο, Μπράγκα, Αβέιρο, Λείρια, Σάντα Ρεμ, Φάρο, Κοΐμπρα και Βιζέου. Όσον αφορά το παράθυρο προηγούμενης γνώσης της θερμοκρασίας έγιναν αρκετοί πειραματισμοί με καλύτερο μοντέλο αυτό των 4 ωρών (lags future covariates). Αυτό το αποτέλεσμα είναι σημαντικό καθότι υποδηλώνει πως η αλλαγή της θερμοκρασίας έχει επίδραση στο φορτίο με κάποια χρονοκαθυστέρηση. Η ερμηνεία του χαρακτηριστικού αυτού είναι αξιοσημείωτη καθώς σε εσωτερικό χώρο η αλλαγή της θερμοκρασίας γίνεται αντιληπτή κάποια ώρα μεταγενέστερα. Για παράδειγμα, το χειμώνα, η μεγάλη πτώση της θερμοκρασίας τις απογευματινές ώρες με την δύση του ηλίου οδηγεί σε αυξημένη λειτουργία μηχανήματα θέρμανσης και φωταγώγησης νωρίς το βράδυ.

Στη συνέχεια, δόθηκε στα μοντέλα Βαθιάς μάθησης N-Beats και TCN ο καιρός ως παρελθοντική συμμεταβλητή. Και στα δύο μοντέλα, η αύξηση της τάξης εισόδου των συμμεταβλητών οδηγεί το μοντέλο σε χειρότερα αποτελέσματα, άρα μειώθηκε η διαστατικότητα τους. Αναλυτικά η ταχύτητα του ανέμου χειροτερεύει αισθητά την απόδοση των μοντέλων όπως και στο LightGBM. Οι καλύτερες επιδόσεις προέκυψαν με τη θερμοκρασία της Λισσαβόνας ως συμμεταβλητή. Το μεν TCN έχει καλύτερες επιδόσεις με την μεταβλητή της θερμοκρασίας, όμως οι εποχές και ο χρόνος εκπαίδευσης είναι σημαντικά αυξημένος. Αντίθετα, το N-Beats έχει συστηματικά χειρότερες επιδόσεις.

Στον Πίνακα 9 φαίνονται αναλυτικά τα αποτελέσματα. Στην στήλη Καλύτερο μοντέλο (Best Model MAPE (%) without weather) υπάρχουν τα καλύτερα αποτελέσματα ως προς τη μετρική MAPE για το LightGBM, N-Beats και TCN αντίστοιχα χωρίς χρήση της συμμεταβλητής του

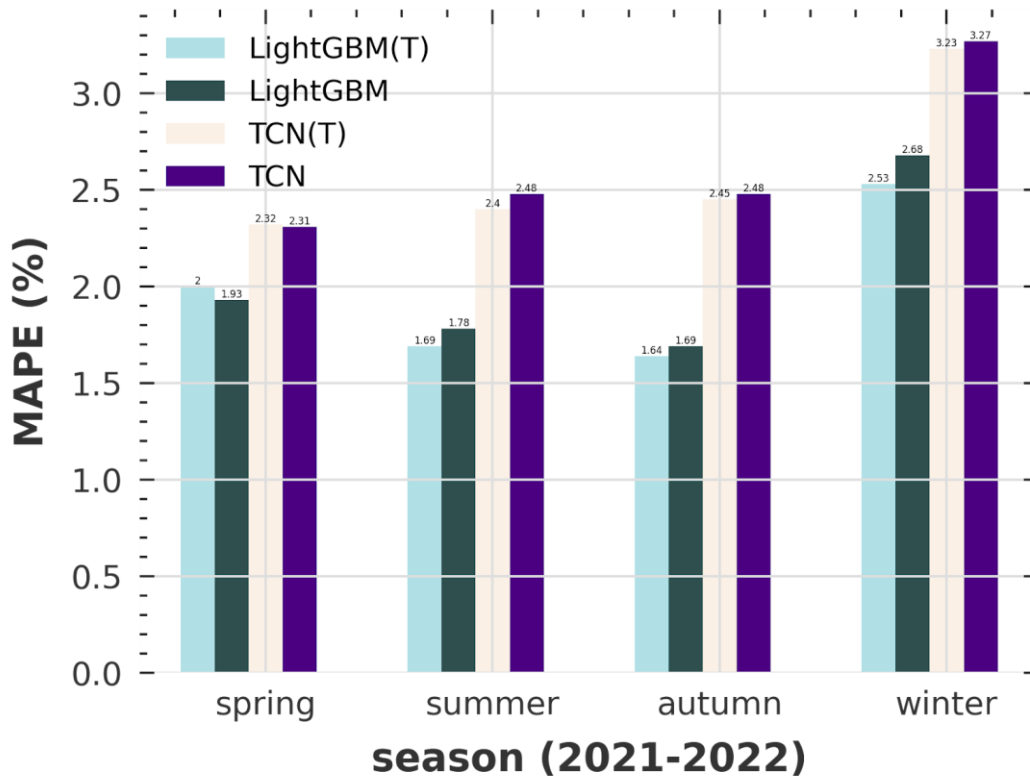
καιρού, όπως αυτά απεικονίζονται στους Πίνακες 5: Αποτελέσματα πειραμάτων LightGBM, Πίνακας 6: Αποτελέσματα N-Beats, Πίνακας 7: Αποτελέσματα μοντέλου TCN για την Πορτογαλία, ενώ στην στήλη (MAPE(%)) είναι τα αποτελέσματα των μοντέλων με χρήση της συμμεταβλητής του καιρού.

Weather					
	Best Model MAPE (%) without weather	Μεταβλητή Καιρού	Εποχές	GPU time (min)	MAPE (%)
LightGBM	2.01	9 πόλεις (T)	-	-	1.96
N-Beats	2.456	Λισαβόνα (T)	58	7	2.57
Ensemble (N-Beats)	2.036	Λισαβόνα (T)	58	115	2.19
TCN	2.94	Λισαβόνα (T)	226	27	2.85
Ensemble (TCN)	2.63	Λισαβόνα (T)	226	400	2.59

Πίνακας 9: Αποτελέσματα χρήσης συμμεταβλητή καιρού στην Πορτογαλία

Στην Εικ. 31 απεικονίζεται η επίδοση των μοντέλων LightGBM και TCN, που πέτυχαν καλύτερα αποτελέσματα με τη συμμεταβλητή του καιρού, συγκεκριμένα της θερμοκρασίας (LightGBM(T), TCN(T)) στις εποχές του χρόνου σε σύγκριση με τα μοντέλα που πέτυχαν τα καλύτερα αποτελέσματα χωρίς τη συμμεταβλητή του καιρού. Όπως παρατηρείται, η επίδοση των μοντέλων σε όλες τις εποχές εκτός από την άνοιξη έχουν καλύτερα αποτελέσματα. Ιδιαίτερο ενδιαφέρον παρουσιάζει η εποχή του χειμώνα, όπου το καταναλισκόμενο φορτίο έχει τις μέγιστες τιμές του και καθοριστικός παράγοντας για αυτό είναι η χαμηλή θερμοκρασία. Κατά τους χειμερινούς λοιπόν, μήνες το LightGBM έχει καλύτερη επίδοση εμφανώς με τη θερμοκρασία. Ίδια συμπεριφορά και για το TCN παρατηρείται κατά τους μήνες του καλοκαιριού. Το N-Beats, έχει χειρότερα αποτελέσματα χρησιμοποιώντας τη συμμεταβλητή του καιρού οπότε εξαιρείται από την συγκεκριμένη ανάλυση

PORTUGAL (WEATHER)



Εικ. 31: Σύγκριση μοντέλων με και χωρίς την συμμεταβλητή του καιρού

Η διαδικασία εκπαίδευσης και αξιολόγησης των μοντέλων για τις χρονοσειρές φορτίου της Πορτογαλίας με τη συμμεταβλητή των μετεωρολογικών δεδομένων υπάρχουν στο αρχείο “Portugal_with_weather_forecasts.ipynb”¹⁹.

¹⁹ https://github.com/giannisiso/Thesis_Short_Term_Load_Forecasting

Κεφάλαιο 5

5.1 Σύνοψη Εργασίας και Επίλογος

Στην παρούσα διπλωματική εργασία έγινε προσπάθεια πρόβλεψης του ηλεκτρικού φορτίου βραχυπρόθεσμου χρόνου (STLF) ευρωπαϊκών χωρών, καθώς και ανάλυση εξωγενών μεταβλητών που επηρεάζουν το φορτίο της Πορτογαλίας, της Ισπανίας και της Ελλάδας, αξιοποιώντας την τεχνητή νοημοσύνη και συγκεκριμένα την μηχανική (ML) και βαθιά μάθηση (DL). Για την επίτευξη του σκοπού αυτού χρησιμοποιούνται state-of-the-art αλγόριθμοι πρόβλεψης. Συγκεκριμένα γίνεται εκτεταμένη χρήση του μοντέλου μηχανικής μάθησης (ML), LightGBM (Light Gradient Boosting Machine) της οικογένειας των δέντρων αποφάσεων, καθώς και των μοντέλων βαθιάς μάθησης (DL), N-Beats (Neural Basis Expansion Analysis for Time Series forecasting) και TCN (Temporal Convolutional Networks). Μετά την εκτενή αναφορά στην βιβλιογραφία και στο θεωρητικό υπόβαθρο του προαναφερθέντος προβλήματος ακολουθεί η θεωρητική θεμελίωση των τριών αλγορίθμων μηχανικής (ML) και βαθιάς μάθησης (DL) που επιχειρούνται στην παρούσα διπλωματική εργασία. Στην συνέχεια, παρουσιάζονται τα σύνολα δεδομένων που χρησιμοποιήθηκαν και αναλύεται η μεθοδολογία που ακολουθήθηκε για την περάτωση του κύκλου ζωής της μηχανικής μάθησης (ML life-cycle). Τα αποτελέσματα αυτής της εργασίας παρουσιάστηκαν, αναλύθηκαν και εξήχθησαν σημαντικά συμπεράσματα για την συνολική αντιμετώπιση του προβλήματος πρόβλεψης του φορτίου της επόμενης ημέρας των Μεσογειακών χωρών Πορτογαλία, Ισπανία και Ελλάδα.

Τόσο τα σύνολα δεδομένων φορτίου της Πορτογαλίας, Ισπανία και Ελλάδας όσο και τα σύνολα των συμμεταβλητών παρελθοντικών ή μελλοντικών που χρησιμοποιήθηκαν (καιρός, τιμή μεγαβατώρας και χρονικές μεταβλητές), αποτελούνται από ωριαία δεδομένα από 01/03/2016 έως και 28/02/2022 (6 έτη). Για τον διαχωρισμό του συνόλου δεδομένων (dataset split), στο μοντέλο μηχανικής μάθησης (ML), LightGBM δόθηκε ως σύνολο εκπαίδευσης (train set) τα πέντε έτη από 01/03/2016 έως 28/02/2021 ενώ στα μοντέλα βαθιάς μάθησης (DL), NBEATS και TCN δόθηκαν τα τέσσερα έτη από 01/03/2016 έως 29/02/2020 ως σύνολο εκπαίδευσης και το έτος από 01/03/2020 έως 28/02/2021 ως σύνολο επικύρωσης (validation set). Ενώ ως σύνολα τεστ (test sets) για όλα τα μοντέλα χρησιμοποιήθηκε το έτος από 01/03/2021 έως 28/02/2022.

Μέσω του πειραματισμού διευκρινίστηκαν τα εξής:

- Το μοντέλο μηχανικής μάθησης LightGBM όπως και για το μοντέλο βαθιάς μάθησης TCN έχει καλύτερες επιδόσεις όταν εκπαιδεύονται με σύνολο δεδομένων φορτίου σε πραγματικές τιμές, λόγω μη ύπαρξης συναρτήσεων ενεργοποίησης στο LightGBM και εξαφάνισης κλίσης στο TCN. Αντίθετα, η ύπαρξη συναρτήσεων ενεργοποίησης στο N-Beats οδηγεί το μοντέλο σε καλύτερες επιδόσεις με κανονικοποιημένο σύνολο δεδομένων φορτίου.
- Το LightGBM έχει καλύτερη επίδοση με παράθυρο προηγούμενης γνώσης 5 ημέρες (120 τιμές), καθώς μπορεί να διαχειριστεί την διαδικασία εκπαίδευσης του αποδοτικότερα. Το N-Beats έχει καλύτερη επίδοση με παράθυρο προηγούμενης γνώσης 10 ημέρες (240 τιμές) όπως αποδεικνύεται μέσω του πειραματισμού αφού το σύνολο δεδομένων είναι ωριαίας ανάλυσης. Το TCN χρειάζεται για την εκπαίδευσή του παράθυρο προηγούμενης γνώσης 12 ημέρες (288 τιμές), το μεγαλύτερο από τα υπόλοιπα μοντέλα καθώς η εξόρυξη γνώσεις μπορεί να γίνει μόνο στο μήκος του πεδίου δεκτικότητάς του.
- Το LightGBM επωφελείται τόσο από χρονικές μεταβλητές όσο και από τα μετεωρολογικά δεδομένα, καθώς χρησιμοποιεί την τεχνική Exclusive Feature Bundling (EFB). Το TCN

επωφελείται μόνο μετεωρολογικά δεδομένα, ενώ το N-Beats χωρίς τη χρήση καμίας συμμεταβλητής.

- Το N-Beats έχει καλύτερες επιδόσεις με την λιγότερο βαθιά αρχιτεκτονική (1η Πίνακας 2), παρότι οι χρόνοι εκπαίδευσης είναι αυξημένοι λόγω μη εύκολης προσαρμογής του μοντέλου, ενώ το TCN με τη βαθύτερη αρχιτεκτονική (2η Πίνακας 3) για τον αντίθετο λόγο.
- Η τεχνική Average Ensemble με διαφορετική αρχικοποίηση βαρών οδηγεί σε ακριβέστερες προβλέψεις τα μοντέλα βαθιάς μάθησης N-Beats και TCN, λόγω της επιτυχούς αμεροληψίας των βαρών, ενώ οδηγεί σε λίγο χειρότερες επιδόσεις το μοντέλο μηχανικής μάθησης LightGBM λόγω της τεχνικής Gradient-based One Side Sampling (GOSS) που χρησιμοποιεί για την αποδοτικότερη εκπαίδευσή του.

Αφού έγιναν τα παραπάνω πειράματα, χρησιμοποιήθηκε υβριδικό Averaging Ensemble αποτελούμενο από το μοντέλο μηχανικής μάθησης LightGBM και το μοντέλο βαθιάς μάθησης N-Beats, τα οποία πέτυχαν τις καλύτερες επιδόσεις όπως φαίνεται στον Πίνακας 8: Τελικά αποτελέσματα Hybrid Averaging Ensemble (LightGBM+N-Beats).

Συνοψίζοντας, για την επίλυση του προβλήματος πρόβλεψης βραχυπρόθεσμου χρόνου (STLF) προτείνεται η χρήση του αλγορίθμου LightGBM για την αποτελεσματικότητα του τόσο σε επίπεδο αξιολόγησης (MAPE) όσο και σε υπολογιστική ισχύ, καθώς δεν είναι αναγκαία η χρήση GPU και οι χρόνοι εκπαίδευσής του είναι μικροί. Παρόλα αυτά, είναι αναγκαία η χρήση εξωγενών μεταβλητών και ιδιαίτερα η κωδικοποίηση του χρόνου ώστε να επιτευχθούν υψηλές επιδόσεις. Αντίθετα, στην περίπτωση δυνατότητας χρήσης GPU και ιδιαίτερα αν δεν υπάρχει πρόσβαση σε εξωγενείς μεταβλητές προτείνεται η χρήση του μοντέλου βαθιάς μάθησης N-Beats, το οποίο σε αποδεκτούς χρόνους εκπαίδευσης πετυχαίνει εξαιρετικές επιδόσεις. Το N-Beats και ιδιαίτερα το TCN επηρεάζεται σημαντικά από τον όγκο των δεδομένων. Οι επιδόσεις του TCN είναι χειρότερες σε σχέση με τα άλλα δύο προαναφερθέντα μοντέλα λόγω μη επαρκούς όγκου δεδομένων που επιφέρει underfitting. Τέλος, για την όσο το δυνατόν αντικειμενικότερη πρόβλεψη προτείνεται η χρήση της τεχνικής Ensemble.

Εν κατακλείδι, αξίζει να τονιστεί πως πάντα υπάρχουν περιθώρια για βελτίωση τόσο στην βελτιστοποίηση των υπερπαραμέτρων όσο και σε τεχνικές επεξεργασίας των δεδομένων. Αν μη τι άλλο, όπως και σε οποιοδήποτε πρόβλημα που απαιτεί διαχείριση και χρήση μεγάλου όγκου δεδομένων, η ποιότητα τους και ο όγκος τους είναι άρρηκτα συνδεδεμένα με την τελική επίδοση των τεχνικών προβλέψεων.

5.2 Μελλοντικές Προεκτάσεις

Τέλος, αξίζει να σημειωθούν πιθανές επεκτάσεις της παρούσας εργασίας. Αρχικά, η επέκταση των αποτελεσμάτων με χρήση καιρικών δεδομένων άλλων χωρών θα ήταν μία εφαρμογή που θα είχε ιδιαίτερο νόημα. Ειδικά στην περίπτωση του ελληνικού συνόλου δεδομένων φορτίου που παρουσιάζει μέγιστες τιμές κατά τους καλοκαιρινούς μήνες λόγω της εκτεταμένης χρήσης μηχανημάτων ψύξης, η ενσωμάτωση και εξόρυξη γνώσης από δεδομένα θερμοκρασίας, ιδιαίτερα προβλέψεων που θα μπορούν να χρησιμοποιηθούν σαν μελλοντικές συμμεταβλητές, εκτιμάται ότι θα οδηγήσουν σε καλύτερες επιδόσεις. Προς επίρρωση του στόχου αυτού μπορούν να χρησιμοποιηθούν οι δυνατότητες του διαδικτύου των πραγμάτων (Internet of things - IoT) ώστε να συλλέγονται καθημερινά τα εν λόγω δεδομένα προβλέψεων καιρού.

Ακόμα, εφόσον αποδείχθηκε από τον πειραματισμό πως τεχνικές Ensemble μπορούν να οδηγήσουν σε καλύτερα αποτελέσματα, τότε η χρήση ενός διαφορετικού Ensemble θα μπορούσε να δώσει όχι μόνο καλύτερες επιδόσεις αλλά και ερμηνεία, διορατικότητα των αποτελεσμάτων και των μοντέλων. Συγκεκριμένα ένα πολυστρωματικό Perceptron, ή απλώς πολυωνυμική παλινδρόμηση, που θα δέχεται σαν εισόδους τις προβλέψεις των μοντέλων και θα δημιουργεί την τελική πρόβλεψη θα είχε αρκετό ενδιαφέρον σαν μελλοντική επέκταση της παρούσας εργασίας.

Τέλος, μία ακόλουθη εφαρμογή είναι η εκπαίδευση μοντέλου με πολλαπλά σύνολα δεδομένων, δηλαδή χρονοσειρών φορτίου διαφορετικών χωρών και η εκτέλεση μεταφοράς μάθησης (Transfer Learning) σε άλλα δεδομένα. Μία καλή επιλογή μοντέλου για αυτή την εφαρμογή είναι το N-Beats που αφενός έχει ήδη χρησιμοποιηθεί για παρόμοια προβλήματα, αφετέρου έχει εξαιρετικές επιδόσεις χωρίς την χρήση συμμεταβλητών, με σχετικά λίγους πόρους, δηλαδή μικρούς χρόνους χρήσης GPU.

Αναφορές

- [1] M. E. El-Hawary, “The smart grid - State-of-the-art and future trends,” *Electr. Power Components Syst.*, vol. 42, no. 3–4, pp. 239–250, Mar. 2014, doi: 10.1080/15325008.2013.868558.
- [2] C. W. Su, K. Khan, M. Umar, and W. Zhang, “Does renewable energy redefine geopolitical risks?,” *Energy Policy*, vol. 158, Nov. 2021, doi: 10.1016/j.enpol.2021.112566.
- [3] S. M. Moghaddas-Tafreshi and M. Farhadi, “A linear regression-based study for temperature sensi.wity analysis of iran electrical load,” 2008. doi: 10.1109/ICIT.2008.4608590.
- [4] B. Yildiz, J. I. Bilbao, and A. B. Sproul, “A review and analysis of regression and machine learning models on commercial building electricity load forecasting,” *Renewable and Sustainable Energy Reviews*, vol. 73. Elsevier Ltd, pp. 1104–1122, 2017. doi: 10.1016/j.rser.2017.02.023.
- [5] J.-F. Chen, W.-M. Wang, and C.-M. Huang, “Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting,” 1995.
- [6] W. Li and Z. G. Zhang, “Based on time sequence of ARIMA model in the application of short-term electricity load forecasting,” in *ICRCCS 2009 - 2009 International Conference on Research Challenges in Computer Science*, 2009, pp. 11–14. doi: 10.1109/ICRCCS.2009.12.
- [7] M. T. Hagan, “THE TIME SERIES APPROACH TO SHORT TERM LOAD FORECASTING,” 1987.
- [8] H. Musbah and M. El-Hawary, “SARIMA Model Forecasting of Short-Term Electrical Load Data Augmented by Fast Fourier Transform Seasonality Detection,” 2019. doi: 10.1109/CCECE.2019.8861542.
- [9] H. T. Yang and Chao-Ming, “Identification of ARMAX model for short term load forecasting: An evolutionary programming approach,” *IEEE Trans. Power Syst.*, vol. 11, no. 1, pp. 403–408, 1996, doi: 10.1109/59.486125.
- [10] A. Ruzzelli, “BuildSys’10 - Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings: Message from the general chair,” *BuildSys’10 - Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*. Association for Computing Machinery, p. 86, 2010.
- [11] P. Vähäkylä, E. Hakonen, and P. Léman, “Short-term forecasting of grid load using Box-Jenkins techniques,” *Int. J. Electr. Power Energy Syst.*, vol. 2, no. 1, pp. 29–34, 1980, doi: 10.1016/0142-0615(80)90004-6.
- [12] W. R. Christiaanse, “Short-term load forecasting using general exponential smoothing,” *IEEE Trans. Power Appar. Syst.*, vol. PAS-90, no. 2, pp. 900–911, 1971, doi: 10.1109/TPAS.1971.293123.
- [13] H. S. Hippert, C. E. Pedreira, and R. C. Souza, “Neural Networks for Short-Term Load Forecasting: A Review and Evaluation,” 2001.
- [14] D. C. Park, M. A. El-Sharkawi, R. J. Marks, L. E. Atlas, and M. J. Damborg, “Electric Load Forecasting Using An Artificial Neural Network,” 1991.
- [15] “Weather Sensitlve Short-Term Load Forecasting Using Nonfully Connected Artificial Neural Network,” 1992.
- [16] D. Srinivasan, A. C. Liew, and C. S. Chang, “A neural network short-term load forecaster,” *Electr. Power Syst. Res.*, vol. 28, no. 3, pp. 227–234, 1994, doi: 10.1016/0378-7796(94)90037-X.
- [17] W. C. Hong, “Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm,” *Energy*, vol. 36, no. 9, pp. 5568–5578, 2011, doi: 10.1016/j.energy.2011.07.015.
- [18] J. X. Che and J. Z. Wang, “Short-term load forecasting using a kernel-based support vector regression combination model,” *Appl. Energy*, vol. 132, pp. 602–609, Nov. 2014, doi: 10.1016/j.apenergy.2014.07.064.
- [19] S. Fan and L. Chen, “Short-term load forecasting based on an adaptive hybrid method,” *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 392–401, Feb. 2006, doi: 10.1109/TPWRS.2005.860944.
- [20] H. Mori and N. Kosemura, “Optimal regression tree based rule discovery for short-term load

- forecasting,” *Proc. IEEE Power Eng. Soc. Transm. Distrib. Conf.*, vol. 2, no. WINTER MEETING, pp. 421–426, 2001, doi: 10.1109/PESW.2001.916878.
- [21] G. Dudek, “Short-term load forecasting using random forests,” *Adv. Intell. Syst. Comput.*, vol. 323, pp. 821–828, 2015, doi: 10.1007/978-3-319-11310-4_71.
- [22] S. Ben Taieb and R. J. Hyndman, “A gradient boosting approach to the Kaggle load forecasting competition,” *Int. J. Forecast.*, vol. 30, no. 2, pp. 382–394, Apr. 2014, doi: 10.1016/j.ijforecast.2013.07.005.
- [23] H. M. Al-Hamadi and S. A. Soliman, “Fuzzy short-term electric load forecasting using Kalman filter,” *IEE Proc. Gener. Transm. Distrib.*, vol. 153, no. 2, pp. 217–227, Mar. 2006, doi: 10.1049/ip-gtd:20050088.
- [24] H. Mori and M. Ohmi, “Probabilistic short-term load forecasting with Gaussian processes,” in *Proceedings of the 13th International Conference on Intelligent Systems Application to Power Systems, ISAP’05*, 2005, vol. 2005, pp. 452–457. doi: 10.1109/ISAP.2005.1599306.
- [25] N. Siddarameshwara, A. Yelamali, and K. Byahatti, “Electricity short term load forecasting using Elman recurrent neural network,” in *Proceedings - 2nd International Conference on Advances in Recent Technologies in Communication and Computing, ARTCom 2010*, 2010, pp. 351–354. doi: 10.1109/ARTCom.2010.44.
- [26] IEEE Staff and IEEE., “2022 56th Annual Conference on Information Sciences and Systems, CISS 2022,” *2022 56th Annual Conference on Information Sciences and Systems, CISS 2022*. IEEE, pp. 1–6, 2022.
- [27] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, “Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches,” *Energies*, vol. 11, no. 7, 2018, doi: 10.3390/en11071636.
- [28] Y. Wang, M. Liu, Z. Bao, and S. Zhang, “Short-term load forecasting with multi-source data using gated recurrent unit neural networks,” *Energies*, vol. 11, no. 5, 2018, doi: 10.3390/en11051138.
- [29] A. V. Vesa *et al.*, “Energy flexibility prediction for data center engagement in demand response programs,” *Sustain.*, vol. 12, no. 4, Feb. 2020, doi: 10.3390/su12041417.
- [30] S. Mishra and S. K. Patra, “Short term load forecasting using neural network trained with genetic algorithm & particle swarm optimization,” in *Proceedings - 1st International Conference on Emerging Trends in Engineering and Technology, ICETET 2008*, 2008, pp. 606–611. doi: 10.1109/ICETET.2008.94.
- [31] I. IEEE Industrial Electronics Society. Annual Conference (42nd: 2016: Florence *et al.*, “Proceedings of the IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society: Florence (Italy), October 24-27, 2016,” in *IECON 2016 - The 42nd Annual Conference of IEEE Industrial Electronics Society*, 2016, pp. 7046–7051. [Online]. Available: <https://dSPACE.lboro.ac.uk/dSPACE-jspui/handle/2134/23301>
- [32] H. Wilms, M. Cupelli, and A. Monti, “Combining auto-regression with exogenous variables in sequence-to-sequence recurrent neural networks for short-term load forecasting,” in *Proceedings - IEEE 16th International Conference on Industrial Informatics, INDIN 2018*, 2018, pp. 673–679. doi: 10.1109/INDIN.2018.8471953.
- [33] W. He, “Load Forecasting via Deep Neural Networks,” in *Procedia Computer Science*, 2017, vol. 122, pp. 308–314. doi: 10.1016/j.procs.2017.11.374.
- [34] C. Tian, J. Ma, C. Zhang, and P. Zhan, “A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network,” *Energies*, vol. 11, no. 12, Dec. 2018, doi: 10.3390/en11123493.
- [35] G. Ke *et al.*, “LightGBM: A highly efficient gradient boosting decision tree,” in *Advances in Neural Information Processing Systems*, 2017, vol. 2017-Decem, pp. 3147–3155. [Online]. Available: <https://github.com/Microsoft/LightGBM>.
- [36] B. N. Oreshkin, D. Carпов, N. Chapados, and Y. Bengio, “N-BEATS: Neural basis expansion analysis for interpretable time series forecasting,” May , [Online]. Available: <http://arxiv.org/abs/1905.10437>
- [37] K. S. Choi, J. S. Shin, J. J. Lee, Y. S. Kim, S. B. Kim, and C. W. Kim, “In vitro trans-differentiation of rat mesenchymal cells into insulin-producing cells by rat pancreatic extract,” *Biochem. Biophys. Res. Commun.*, vol. 330, no. 4, pp. 1299–1305, 2005, doi:

- 10.1016/j.bbrc.2005.03.111.
- [38] A. Borovykh, S. Bohte, and C. W. Oosterlee, “Conditional time series forecasting with convolutional neural networks,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Mar. 2017, vol. 10614 LNCS, pp. 729–730. [Online]. Available: <http://arxiv.org/abs/1703.04691>
- [39] O. Sagi and L. Rokach, “Ensemble learning: A survey,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4. Wiley-Blackwell, Jul. 01, 2018. doi: 10.1002/widm.1249.
- [40] J. A. Dutton, “American Meteorological Society OPPORTUNITIES AND PRIORITIES IN A NEW ERA FOR WEATHER AND CLIMATE SERVICES,” *Source Bull. Am. Meteorol. Soc.*, vol. 83, no. 9, pp. 1303–1311, 2002, doi: 10.2307/26215393.
- [41] K. S. L. Madhavi *et al.*, “Advanced electricity load forecasting combining electricity and transportation network,” 2017. doi: 10.1109/NAPS.2017.8107312.
- [42] Institute of Electrical and Electronics Engineers, “ICPRE 2016 : 2016 IEEE International Conference on Power and Renewable Energy : October 21-23, 2016, Shanghai, China.,” pp. 388–393, 2016.
- [43] L. Li, K. Ota, and M. Dong, “When Weather Matters: IoT-Based Electrical Load Forecasting for Smart Grid,” *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 46–51, Oct. 2017, doi: 10.1109/MCOM.2017.1700168.
- [44] D. Solyali, “A comparative analysis of machine learning approaches for short-/long-term electricity load forecasting in Cyprus,” *Sustain.*, vol. 12, no. 9, May 2020, doi: 10.3390/SU12093612.
- [45] G. Huang, “Missing data filling method based on linear interpolation and lightgbm,” in *Journal of Physics: Conference Series*, Feb. 2021, vol. 1754, no. 1. doi: 10.1088/1742-6596/1754/1/012187.
- [46] S. Kanwal, *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2018.
- [47] J. Sola and J. Sevilla, “Importance of Input Data Normalization for the Application of Neural Networks to Complex Industrial Problems,” 1997. Accessed: Jun. 09, 2022. [Online]. Available: <https://sci-hub.se/10.1109/23.589532>
- [48] S. G. K. Patro and K. K. Sahu, “Normalization: A Preprocessing Stage,” *IARJSET*, pp. 20–22, 2015, doi: 10.17148/iarjset.2015.2305.
- [49] J. Park and E. Hwang, “A two-stage multistep-ahead electricity load forecasting scheme based on lightgbm and attention-bilstm,” *Sensors*, vol. 21, no. 22, Nov. 2021, doi: 10.3390/s21227697.
- [50] Y. Wang *et al.*, “Short-term load forecasting for industrial customers based on TCN-LightGBM,” *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 1984–1997, May 2021, doi: 10.1109/TPWRS.2020.3028133.
- [51] B. N. Oreshkin, G. Dudek, P. Pelka, and E. Turkina, “N-BEATS neural network for mid-term electricity load forecasting,” *Appl. Energy*, vol. 293, Jul. 2021, doi: 10.1016/j.apenergy.2021.116918.
- [52] H. Wen, J. Gu, J. Ma, L. Yuan, and Z. Jin, “Probabilistic Load Forecasting via Neural Basis Expansion Model Based Prediction Intervals,” *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3648–3660, Jul. 2021, doi: 10.1109/TSG.2021.3066567.
- [53] S. Pelekis *et al.*, “In Search of Deep Learning Architectures for Load Forecasting: A Comparative Analysis and the Impact of the Covid-19 Pandemic on Model Performance,” 2022.
- [54] S. Bai, J. Z. Kolter, and V. Koltun, “An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling,” Mar. 2018, [Online]. Available: <http://arxiv.org/abs/1803.01271>
- [55] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, “Ensemble deep learning: A review,” vol. 2, pp. 1–49, Apr. 2021, doi: <https://doi.org/10.48550/arXiv.2104.02395>.
- [56] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, Oct. 2006, doi: 10.1016/j.ijforecast.2006.03.001.