

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

Τμήμα Πληροφορικής



- Πτυχιακή Εργασία –

Διαχείριση Μεγάλου Όγκου Δεδομένων Κατανάλωσης Ρεύματος με Τεχνικές
Μηχανικής Μάθησης

Μιχαήλ Καβαδίας

Π2019047

Επιβλέπων: Ανδρέας Καναβός

15 Σεπτεμβρίου 2024

Επιβλέπων

Ανδρέας Καναβός, Αναπληρωτής Καθηγητής
Ιόνιο Πανεπιστήμιο

Τριμελής Επιτροπή

Ανδρέας Καναβός, Αναπληρωτής Καθηγητής
Ιόνιο Πανεπιστήμιο

Κάτια – Λήδα Κερμανίδου, Αναπληρώτρια Καθηγήτρια
Ιόνιο Πανεπιστήμιο

Στέργιος Παλαμάς, Επίκουρος Καθηγητής
Ιόνιο Πανεπιστήμιο

Πρόλογος και Ευχαριστίες

Στο συνεχώς εξελισσόμενο τοπίο της λήψης αποφάσεων με γνώμονα τα δεδομένα, η αναζήτηση της γνώσης συχνά βρίσκει την γένεσή της στην περιέργεια του ανθρώπινου μυαλού. Είναι αυτή η έμφυτη επιθυμία να αποκαλυφθούν τα κρυμμένα μοτίβα, να αποκαλυφθεί η διασύνδεση διαφορετικών φαινομένων, που έχει τροφοδοτήσει την αδυσώπητη και συνεχώς εξελισσόμενη πορεία της επιστήμης και της τεχνολογίας. Τι γίνεται όμως όταν ο αριθμός των πληροφοριών που θέλουμε να αναλύσουμε είναι τόσο μεγάλος που ο ανθρώπινος νους δεν μπορεί να τις διαχειριστεί; Τότε παίρνει την σκυτάλη η επιστήμη της διαχείρισης Μεγάλου Όγκου Δεδομένων. Με την χρήση της Επιστήμης αυτής, επιδιώκουμε να ανακαλύψουμε την σχέση μεταξύ καιρού και κατανάλωσης ενέργειας, αλλά και την πρόβλεψη της κατανάλωσης αυτής με βάση την υπάρχουσα βάση δεδομένων.

Στις σελίδες που ακολουθούν, εμβαθύνουμε στον ιστό των δεδομένων αυτών προσπαθώντας να εξάγουμε γνώσεις που ξεπερνούν την απλή παρατήρηση και μας καθοδηγούν στην ακλόνητη και αδιαμφισβήτητη συνειδητοποίηση ότι η κατανάλωση ενέργειας και η θερμοκρασία συσχετίζονται σε άμεσο βαθμό.

Αυτή η εργασία δεν θα ήταν δυνατή χωρίς την καθοδήγηση και την αμέριστη υποστήριξη του καθηγητή Ανδρέα Καναβού, που δέχτηκε να με επιβλέψει και να με βοηθήσει στο ακαδημαϊκό αυτό ταξίδι, στο οποίο έχει υπάρξει φάρος έμπνευσης για πολλούς συναδέλφους με την αφοσίωσή του στην πρόοδο των μαθητών του.

Ξεκινώντας την εξερεύνηση της δυνατότητας της Ανάλυσης Μεγάλων Δεδομένων να διαμορφώσει την κατανόησή μας για τον κόσμο, στεκόμαστε στους ώμους ακαδημαϊκών τιτάνων και καθοδηγούμαστε από την συλλογική σοφία όσων ήρθαν πριν από εμάς. Με ευγνωμοσύνη λοιπόν προς αυτούς, προσπαθούμε να φωτίσουμε την πορεία μας προς ένα καλύτερο μέλλον για την ανάπτυξη της επιστήμης.

Κέρκυρα, Σεπτέμβριος 2024

Μιχαήλ Καβαδιάς

Περιεχόμενα

A' ΕΙΣΑΓΩΓΗ.....	1
A.1 Εισαγωγή	1
A.2 Ανάλυση Μεγάλου Όγκου δεδομένων και Ιστορική αναδρομή	2
A.3 Σύγχρονες Εφαρμογές της Μηχανικής Μάθησης στη Διαχείριση Ενέργειας	4
A.3.1 Εφαρμογές Μηχανικής Μάθησης στην Ενεργειακή Αγορά	5
A.4 Σύγχρονες Τεχνολογίες Ανάλυσης Δεδομένων	6
A.5 Machine Learning	8
A.6 Στόχος της Εργασίας	9
A.7 Περιορισμοί και Εμβέλεια	10
B' ΜΕΘΟΔΟΛΟΓΙΑ	11
B.1 Πηγή Δεδομένων και προ επεξεργασία	11
B.2 Εξερεύνηση και ανάλυση δεδομένων	11
B.3 Επιλογή Αναλυτικών Εργαλείων.....	12
B.4 Συλλογή δεδομένων και Μεθοδολογία.....	12
Γ' ΠΡΟΕΤΟΙΜΑΣΙΑ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ.....	14
Γ.1 Προετοιμασία του περιβάλλοντος έρευνας	14
Γ.2' Προετοιμασία δεδομένων	17
Γ.3' Λογική ανάπτυξης κώδικα.....	18
Γ.4' Ανάπτυξη και αποτελέσματα κώδικα	19
Γ.4.1 Energy.py	19
Γ.4.2 Temperature.py.....	20
Γ.4.3 forecast_whole.py	22
Γ.4.4 Forecast_per_LCLid	24
Γ.5' Περιορισμοί της έρευνας	31
Δ' Συζήτηση	33
Δ.1 Ανάλυση κώδικα.....	33
Δ.1.1 Pandas Library.....	34
Δ.2 Prophet Library	35
Δ.2.1 Εισαγωγή στη Βιβλιοθήκη Prophet	35
Δ.2.2 Πλεονεκτήματα της Χρήσης της Prophet στη Διαχείριση Ενέργειας	36

Δ.2.3 Εφαρμογή του Prophet στην Πτυχιακή Εργασία.....	37
Δ.2.4 Περιορισμοί της Χρήσης της Prophet.....	38
Δ.2.5 Συμπεράσματα για τη Βιβλιοθήκη Prophet	38
Δ.3 Ερμηνεία των αποτελεσμάτων.....	38
Δ.3.1 Σχέση κατανάλωσης ενέργειας και θερμοκρασίας	39
Δ.3.2 Προβλέψεις κατανάλωσης ενέργειας	41
Δ.4 Μελλοντικές Προοπτικές.....	47
Δ.4.1 Ενίσχυση της Μηχανικής Μάθησης στην Ενεργειακή Διαχείριση	47
Δ.4.2 Εξέλιξη των Έξυπνων Δικτύων (Smart Grids).....	47
Δ.4.3 Τεχνητή Νοημοσύνη (AI) και Βελτιστοποίηση Ενεργειακών Συστημάτων	48
Δ.4.4 Εφαρμογές Blockchain στην Ενεργειακή Αγορά	49
Δ.5 Συζήτηση.....	49
Δ.5.1 Προκλήσεις και πλαίσιο δεδομένων	51
Δ.5.2 Συμπεράσματα	51
BIBΛΙΟΓΡΑΦΙΑ.....	52

Περίληψη

Σκοπός της εργασίας αυτής είναι η χρήση ήδη υπαρχόντων αλγορίθμων σε συνδυασμό με το εξής [dataset](#) για την ανάλυση των δεδομένων του και καταγραφή παραλληλισμών ανάμεσα στα δεδομένα της κατανάλωσης ρεύματος και του καιρού, αλλά και την προσπάθεια χρήσης του prophet για πρόβλεψη μελλοντικής κατανάλωσης ρεύματος σύμφωνα με τα υπάρχοντα δεδομένα. Τέλος, θα προσπαθήσω να εμφανίσω με ακρίβεια την μέση κατανάλωση ρεύματος ανά οικοδομικό τετράγωνο και σε όλο το Λονδίνο κατά την περίοδο καταγραφής των δεδομένων.

Abstract

The purpose of this work is the use of already existing algorithms in combination with the following data set to analyze its data and record parallels between the data of power consumption and the weather, but also the attempt to use the prophet library to predict future power consumption according to with existing data. Finally, I will try to accurately display the average power consumption per block and across London over the period of data collection.

Κατάλογος Σχημάτων

Εικόνα 1 – επιφάνεια εργασίας	14
Εικόνα 2 – ανύψωση τερματικού σε διαχειριστή	15
Εικόνα 3 – εγκατάσταση Pandas	15
Εικόνα 4 – εγκατάσταση Pandas-doc	16
Εικόνα 5 – εγκατάσταση Matplotlib	16
Εικόνα 6 – εγκατάσταση python-pip	16
Εικόνα 7 – εγκατάσταση Seaborn	16
Εικόνα 8 – εγκατάσταση Numpy 1.22.1	16
Εικόνα 9 – εγκατάσταση rystan 2.19.1.1	17
Εικόνα 10 – εγκατάσταση python-dev	17
Εικόνα 11 – εγκατάσταση prophet	17

Πίνακας 1 – Συνολική μέση κατανάλωση ενέργειας ανά ημέρα	20
Πίνακας 2 – Μέση μέγιστη θερμοκρασία ανά ημέρα	21
Πίνακας 3 – Πρόβλεψη συνδυαστικής κατανάλωσης ενέργειας από prophet	23
Πίνακας 4 – πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000131	25
Πίνακας 5 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000132	26
Πίνακας 6 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000221	26
Πίνακας 7 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000228	27
Πίνακας 8 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000234	27
Πίνακας 9 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000235	28
Πίνακας 10 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC003451	28
Πίνακας 11 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC003506	29
Πίνακας 12 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC004356	29
Πίνακας 13 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC004538	30
Πίνακας 14 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC004548	30
Πίνακας 15 - Συνολική μέση κατανάλωση ενέργειας ανά ημέρα	40
Πίνακας 16 – Μέση μέγιστη θερμοκρασία ανά ημέρα	40
Πίνακας 17 – παραλληλισμός και πρόβλεψη συνδυαστικής κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία	42
Πίνακας 18 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC000131	42
Πίνακας 19 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC000132	43
Πίνακας 20 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC000221	43
Πίνακας 21 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC000228	44
Πίνακας 22 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC000234	44
Πίνακας 23 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC003451	45

Πίνακας 24 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC003506.....	45
Πίνακας 25 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC004356.....	46
Πίνακας 26 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC004548.....	46

Εξίσωση 1 - Κώδικας του αλγόριθμου Energy.py	19
Εξίσωση 2 – Κώδικας του αλγόριθμου Temperature.py.....	20
Εξίσωση 3 - Κώδικας του αλγόριθμου forecast_whole.py	22
Εξίσωση 4 - Κώδικας του αλγόριθμου Forecast_per_LCLid.....	24

Α' ΕΙΣΑΓΩΓΗ

A.1 Εισαγωγή

Τι είναι η πληροφορική; Πληροφορική ονομάζουμε την επιστήμη που έχει την δυνατότητα να μετατρέπει δεδομένα και πληροφορίες σε γνώση που μπορούν να χρησιμοποιήσουν οι άνθρωποι στην καθημερινότητά τους, μέσα από την μελέτη των αλγοριθμικών δομών και των αλληλεπιδράσεων φυσικών και ψηφιακών υπολογιστικών συστημάτων. Η χρήση της πληροφορικής εστιάζει κυρίως στην βέλτιστη επικοινωνία του ανθρώπου με την τεχνολογία, διαχωρίζοντας την πληροφορική από την *τεχνολογία της πληροφορίας* (IT) (Hastie, Tibshirani, & Friedman, 2009). Επιπλέον, η επιστήμη της πληροφορικής έχει απώτερο σκοπό την κατανόηση της “ζωής” σε μία ψηφιακή κοινωνία, όπου κάθε πτυχή της επηρεάζεται και καλυτερεύει από την πληροφορική. Για παράδειγμα, πολλοί επιστημονικοί και ιατρικοί τομείς επωφελούνται από την πληροφορική με φρενήρεις ρυθμούς, με την ανάπτυξη νέων αλγορίθμων για υπολογισμούς στην φυσική και την χημεία, ακόμα και στην δημιουργία καλύτερων και πιο αποδοτικών φαρμάκων. Ένας από τους τομείς αυτούς, είναι και η Διαχείριση Μεγάλου Όγκου Δεδομένων, όπου όταν αδυνατεί ο ανθρώπινος παράγοντας να διαχειριστεί τον μεγάλο όγκο των δεδομένων που μπορεί να πρέπει να χρησιμοποιήσει για να βγάλει κάποια αποτελέσματα, μπαίνει σε δράση η πληροφορική για την ανάλυση και την οργάνωση τεράστιων συνόλων δεδομένων (García, Luengo, & Herrera, 2015). Έτσι φτάνουμε στο θέμα της εργασίας αυτής, το οποίο είναι η ανάλυση της κατανάλωσης ρεύματος στο Λονδίνο του 2011-2014 ανά οικοδομικό τετράγωνο και σαν σύνολο σε ολόκληρη την πόλη, την πρόβλεψη της κατανάλωσης τον ακόλουθο χρόνο, αλλά και τον παραλληλισμό της κατανάλωσης με την θερμοκρασία την ίδια εποχή (2011-2014). Κατά την διάρκεια της εργασίας βρέθηκα αντιμέτωπος με κάποια προβλήματα σε σχέση με το σύνολο δεδομένων και την βέλτιστη οπτικοποίηση τους αλλά ευτυχώς η προσπάσή τους έγινε επιτυχώς. Χρησιμοποιήθηκε το λειτουργικό σύστημα Ubuntu 20.4 και η γλώσσα προγραμματισμού Python3. Επιπρόσθετα, για την ορθή υλοποίηση του

κώδικα εκμεταλλεύτηκα τις βιβλιοθήκες `pandas` για την επεξεργασία του συνόλου δεδομένων, `matplotlib.pyplot` για την οπτικοποίηση των ευρημάτων σε σχήμα, `matplotlib.backends.backend_pdf` για την αποθήκευση των σχημάτων σε pdf αρχείο, `seaborn` σε συνεργασία με την `pyplot` και `prophet` για τις προβλέψεις με βάση τα ήδη υπάρχοντα στοιχεία, καταγράφοντας την διαδικασία σε αρχεία και σχέδια ώστε να τα παραθέσω στην εργασία αυτή. Η υπόθεσή μου ήταν ότι η κατανάλωση ενέργειας ταυτίζεται αντιστρόφως ανάλογα και άμεσα με την θερμοκρασία, πράγμα το οποίο αποδείχτηκε κατά την εκτέλεση του κώδικα από τα σχέδια που παρήγαγε. Δημιούργησα δύο σχέδια για την σύγκριση, ένα για την θερμοκρασία και ένα για την κατανάλωση ρεύματος ανά ημέρα, ώστε να μπορούμε να τα συγκρίνουμε με ευκολία, για την πρόβλεψη πήρα ένα δείγμα 4000 (τεσσάρων χιλιάδων) καταχωρίσεων διότι το υλικό που διαθέτω δεν μπορούσε να διαχειριστεί κάτι μεγαλύτερο με ευκολία, όντας ένα μηχάνημα χαμηλής υπολογιστικής δύναμης με βάση τα δεδομένα του 2024. Επέλεξα αυτό το θέμα διότι δεν βρήκα κάποια έρευνα με σκοπό που να ταυτίζεται με αυτό και θέλησα να εξετάσω την σύνδεση ανάμεσα στις δύο βάσεις δεδομένων κάνοντας παραλληλισμούς μεταξύ τους αν αυτοί υπάρχουν.

A.2 Ανάλυση Μεγάλου Όγκου δεδομένων και Ιστορική αναδρομή

Αρχικά, καλό θα ήταν να κάνουμε μια ιστορική αναδρομή στην εξέλιξη της Ανάλυσης Μεγάλων Δεδομένων για την καλύτερη κατανόηση του τεχνολογικού αυτού επιτεύγματος. Την δεκαετία του '90 ξεκίνησε η αποθήκευση δεδομένων για την ευκολότερη ανάλυση δομημένων δεδομένων. Αργότερα, το 2006 δημιουργήθηκε το Hadoop και γνωστοποιήθηκε από τους Doug Cutting και Mike Cafarella. Το Hadoop είναι ένα πρόγραμμα ανοιχτού κώδικα που παρέχει επεξεργασία δεδομένων κατανεμημένου μέσου αποθήκευσης και μεγάλου μεγέθους (Zhou & Bialek, 2005). Το 2009 εισήχθησαν οι βάσεις δεδομένων NoSQL, οι οποίες προσέφεραν έναν ευέλικτο τρόπο για την αποθήκευση και την ανάκτηση μη δομημένων δεδομένων. Το υπολογιστικό νέφος έγινε πιο διαδεδομένο λίγο αργότερα, και βοηθάει τις εταιρείες αλλά και τους ιδιώτες να αποθηκεύουν τα σημαντικά τους δεδομένα, σε απομακρυσμένα κέντρα

δεδομένων εξοικονομώντας έτσι τα κόστη υποδομής και συντήρησης. Στην πιο σύγχρονη εποχή, λαμβάνει χώρα η Μηχανική Μάθηση, της οποίας οι αλγόριθμοι λειτουργούν σε μεγάλα δεδομένα για την ανάλυσή τους και την παραγωγή ουσιαστικών πληροφοριών από αυτά. Η ανάπτυξη της τεχνολογίας αυτής οδήγησε και στην δημιουργία των εφαρμογών τεχνητής νοημοσύνης, οι οποίες έχουν λάβει μεγάλη δημοσιότητα τα τελευταία χρόνια. Παράλληλα, υπάρχει και η τεχνολογία ροής δεδομένων, η οποία επεξεργάζεται δεδομένα μεγάλου όγκου σε πραγματικό χρόνο, για γρήγορες αποδόσεις (Dean & Ghemawat,, 2008). Σε γενικές γραμμές, η τεχνολογία μεγάλων δεδομένων έχει κάνει πολύ μεγάλη πρόοδο από τις πρώτες αποθήκες δεδομένων. Οι τεχνολογικές αναπτύξεις που καλύψαμε νωρίτερα έχουν επαναστατικοποιήσει τόσο τον τρόπο αποθήκευσης, όσο και τον τρόπο επεξεργασίας και ανάλυσης μεγάλων όγκων δεδομένων. Με την πάροδο των χρόνων θα ήταν ασφαλές να υποθέσουμε ότι η τεχνολογία αυτή θα παίξει καίριο ρόλο σε διάφορους επιστημονικούς τομείς.

Οι τεχνικές που αναπτύχθηκαν για τη διαχείριση μεγάλου όγκου δεδομένων (Big Data) έχουν διεισδύσει σε πολλούς τομείς της καθημερινότητας, με την ενέργεια να αποτελεί έναν από τους βασικούς τομείς εφαρμογής. Ωστόσο, υπάρχουν αρκετές προκλήσεις στην επεξεργασία αυτών των δεδομένων, ιδίως όταν πρόκειται για δεδομένα όπως η κατανάλωση ρεύματος, που συλλέγονται σε πραγματικό χρόνο και σε πολύ μεγάλες ποσότητες.

Μία από τις κύριες προκλήσεις είναι η αποθήκευση αυτών των δεδομένων. Παρά την ανάπτυξη τεχνολογιών όπως το Hadoop και το NoSQL, η διαχείριση της χωρητικότητας παραμένει ένα δύσκολο έργο για πολλές εταιρείες. Επιπλέον, τα δεδομένα που συλλέγονται δεν είναι πάντα πλήρως δομημένα, και η επεξεργασία τους απαιτεί σημαντική προετοιμασία και φιλτράρισμα.

Στον τομέα της ενέργειας, τα δεδομένα προέρχονται από πολλαπλές πηγές (π.χ. έξυπνοι μετρητές, αισθητήρες καιρού, συστήματα ελέγχου δικτύων) και συχνά οι μορφές και τα πρότυπα αυτών των δεδομένων διαφέρουν. Η ενοποίηση και συγχώνευση αυτών των διαφορετικών πηγών πληροφοριών απαιτεί την ανάπτυξη εξειδικευμένων εργαλείων και τεχνικών, που να μπορούν να αντιμετωπίσουν την ποικιλομορφία των δεδομένων.

Επιπλέον, οι προβλέψεις κατανάλωσης ενέργειας δεν εξαρτώνται μόνο από τα ιστορικά δεδομένα, αλλά και από εξωτερικούς παράγοντες, όπως οι καιρικές συνθήκες, η ώρα της ημέρας

και η οικονομική δραστηριότητα. Αυτό καθιστά δύσκολη τη δημιουργία ενός μοντέλου που να μπορεί να λαμβάνει υπόψη όλους αυτούς τους παράγοντες ταυτόχρονα, απαιτώντας τη χρήση πιο σύνθετων αλγορίθμων και μεγαλύτερη υπολογιστική ισχύ.

A.3 Σύγχρονες Εφαρμογές της Μηχανικής Μάθησης στη Διαχείριση Ενέργειας

Η χρήση των έξυπνων μετρητών (smart meters) έχει φέρει επανάσταση στη συλλογή δεδομένων κατανάλωσης ενέργειας. Μέσω αυτών των συσκευών, οι καταναλωτές έχουν πλέον τη δυνατότητα να παρακολουθούν τη χρήση τους σε πραγματικό χρόνο, ενώ οι πάροχοι ενέργειας αποκτούν λεπτομερείς πληροφορίες που μπορούν να οδηγήσουν σε βελτιστοποίηση των δικτύων διανομής. Αυτή η νέα μορφή δεδομένων επιτρέπει επίσης την εφαρμογή εξελιγμένων αλγορίθμων μηχανικής μάθησης, οι οποίοι μπορούν να αναγνωρίζουν πρότυπα κατανάλωσης και να προβλέπουν τη μελλοντική ζήτηση. (López, Valgañón, & Arroyo, 2021)

Η μηχανική μάθηση χρησιμοποιείται για τη βελτιστοποίηση των ενεργειακών συστημάτων με τρόπους που υπερβαίνουν την παραδοσιακή ανάλυση δεδομένων. Οι αλγόριθμοι Μηχανικής Μάθησης, όπως τα Νευρωνικά Δίκτυα (Neural Networks) και τα Δέντρα Απόφασης (Decision Trees) επιτρέπουν την ανάλυση μεγάλων συνόλων δεδομένων για την αναγνώριση προτύπων και τάσεων που είναι δύσκολο να εντοπιστούν με άλλες μεθόδους. Αυτά τα πρότυπα μπορούν να εφαρμοστούν σε διάφορους τομείς, όπως η πρόβλεψη της ζήτησης ενέργειας, η βελτιστοποίηση της αποδοτικότητας των έξυπνων δικτύων (smart grids) και η ανάλυση της επίδρασης εξωτερικών παραγόντων, όπως οι καιρικές συνθήκες, στην κατανάλωση.

Στις σύγχρονες εφαρμογές, οι αλγόριθμοι αυτοί μπορούν να βοηθήσουν τόσο στην εξοικονόμηση ενέργειας όσο και στη μείωση του ενεργειακού κόστους. Χρησιμοποιώντας τεχνικές όπως η ενισχυτική μάθηση (reinforcement learning), τα έξυπνα δίκτυα ενέργειας μπορούν να βελτιστοποιούν την κατανομή της ενέργειας, αυξάνοντας την αποδοτικότητα των πόρων. Επίσης, μέσω της πρόβλεψης της ζήτησης ενέργειας σε καθημερινό ή εβδομαδιαίο

επίπεδο, οι διαχειριστές δικτύων μπορούν να διασφαλίσουν την ομαλή λειτουργία των δικτύων και την αποτροπή πιθανών διακοπών λόγω υπερφόρτωσης. (Deb C. et al., 2017)

Η ενσωμάτωση τεχνολογιών IoT (Internet of Things) με τους αλγορίθμους μηχανικής μάθησης αποτελεί το επόμενο βήμα στην ενεργειακή διαχείριση. Μέσω των συστημάτων IoT, η παρακολούθηση της κατανάλωσης μπορεί να γίνει σε επίπεδο συσκευής, επιτρέποντας τη λεπτομερή ανάλυση της κατανάλωσης ενέργειας και τη λήψη πιο στοχευμένων μέτρων για εξοικονόμηση. (Fan, Xiao, & Zhao,, 2017)

A.3.1 Εφαρμογές Μηχανικής Μάθησης στην Ενεργειακή Αγορά

Η μηχανική μάθηση έχει αρχίσει να παίζει σημαντικό ρόλο στην ενεργειακή αγορά, τόσο στην πρόβλεψη της ζήτησης όσο και στη βελτιστοποίηση της χρήσης των πόρων (Ahmad T., Chen H. , & Wang, J., 2020). Ειδικότερα, η μηχανική μάθηση μπορεί να προσφέρει λύσεις για:

- **Πρόβλεψη ζήτησης ενέργειας:** Οι αλγόριθμοι πρόβλεψης, όπως το Prophet και τα δίκτυα νευρώνων, χρησιμοποιούνται για την πρόβλεψη της ενεργειακής ζήτησης σε διαφορετικές χρονικές περιόδους. Αυτές οι προβλέψεις βοηθούν τους διαχειριστές δικτύων να προγραμματίσουν τις παραγγελίες και την κατανομή των ενεργειακών πόρων.
- **Αναγνώριση προτύπων κατανάλωσης:** Οι αλγόριθμοι μηχανικής μάθησης μπορούν να εντοπίσουν μοτίβα στη χρήση ενέργειας σε επίπεδο νοικοκυριών, εταιρειών ή βιομηχανιών. Αυτά τα μοτίβα μπορούν να χρησιμοποιηθούν για τη βελτιστοποίηση της κατανάλωσης και τη μείωση της σπατάλης. (Liu, F., et al., 2020)
- **Διαχείριση Ανανεώσιμων Πηγών Ενέργειας (ΑΠΕ):** Οι ΑΠΕ, όπως η ηλιακή και η αιολική ενέργεια, είναι δύσκολο να προβλεφθούν λόγω της αβέβαιης φύσης των καιρικών συνθηκών. Η μηχανική μάθηση μπορεί να χρησιμοποιηθεί για την πρόβλεψη της απόδοσης αυτών των πηγών, επιτρέποντας τη βελτιστοποίηση της διαχείρισής τους.
- **Έξυπνα δίκτυα ενέργειας:** Τα έξυπνα δίκτυα χρησιμοποιούν αλγόριθμους μηχανικής μάθησης για να παρακολουθούν και να διαχειρίζονται την κατανομή ενέργειας σε

πραγματικό χρόνο. Αυτά τα δίκτυα είναι ικανά να ανταποκρίνονται δυναμικά στις αλλαγές της ζήτησης, βελτιώνοντας την αποδοτικότητα και μειώνοντας το κόστος παραγωγής. (Nazir, M. S., et al., 2021)

Οι εφαρμογές μηχανικής μάθησης στην ενεργειακή αγορά περιλαμβάνουν τη χρήση τεχνικών όπως τα Deep Learning και Reinforcement Learning (ενισχυτική μάθηση) για την πρόβλεψη της κατανάλωσης ενέργειας και τη βελτιστοποίηση της διαχείρισης των πόρων. Με τις συνεχώς μεταβαλλόμενες τιμές της ενέργειας, η ακριβής πρόβλεψη ζήτησης είναι κρίσιμη για τους διαχειριστές δικτύων και τις εταιρείες παραγωγής ενέργειας. Η χρήση μηχανικής μάθησης επιτρέπει στους διαχειριστές να αναλύουν ιστορικά δεδομένα και να προβλέπουν τις μελλοντικές ανάγκες με μεγαλύτερη ακρίβεια, μειώνοντας τα λειτουργικά κόστη και αυξάνοντας την αποδοτικότητα.

A.4 Σύγχρονες Τεχνολογίες Ανάλυσης Δεδομένων

Τα τελευταία χρόνια, η ανάλυση μεγάλου όγκου δεδομένων έχει επεκταθεί σημαντικά, υποστηριζόμενη από τεχνολογικές καινοτομίες και εργαλεία που επιτρέπουν τη διαχείριση και την ανάλυση τεράστιων ποσοτήτων δεδομένων. Ορισμένες από τις βασικές τεχνολογίες που χρησιμοποιούνται ευρέως περιλαμβάνουν:

- **Apache Hadoop:** Ένα σύστημα διαχείρισης μεγάλων δεδομένων που παρέχει την ικανότητα να επεξεργάζεται και να αποθηκεύει δεδομένα σε κατανεμημένα δίκτυα. Επιτρέπει την επεξεργασία δεδομένων μεγάλης κλίμακας σε πολλαπλά συστήματα, με χρήση παραλληλίας.
- **Apache Spark:** Ένα εξελιγμένο εργαλείο για την επεξεργασία μεγάλων δεδομένων, το οποίο προσφέρει ταχύτερη επεξεργασία σε σχέση με το Hadoop, αξιοποιώντας την ενσωματωμένη μνήμη του συστήματος (in-memory computing). Το Spark είναι ιδανικό για αναλύσεις πραγματικού χρόνου, όπως η παρακολούθηση ενεργειακής κατανάλωσης από έξυπνους μετρητές.

- **NoSQL βάσεις δεδομένων:** Οι βάσεις NoSQL, όπως οι Cassandra και MongoDB, προσφέρουν ευέλικτες λύσεις αποθήκευσης δεδομένων που δεν έχουν δομή πίνακα όπως οι παραδοσιακές βάσεις δεδομένων SQL. Αυτές οι βάσεις χρησιμοποιούνται ευρέως για δεδομένα μεγάλου όγκου και μη δομημένα δεδομένα.
- **Τεχνολογίες ροής δεδομένων (streaming):** Τεχνολογίες όπως το Apache Kafka επιτρέπουν την επεξεργασία και ανάλυση δεδομένων σε πραγματικό χρόνο. Στον τομέα της ενέργειας, η ανάλυση δεδομένων ροής μπορεί να χρησιμοποιηθεί για την παρακολούθηση της ζήτησης ενέργειας ανά πάσα στιγμή και την προσαρμογή της παραγωγής ανάλογα με τη ζήτηση.

Η διαχείριση και η ανάλυση των μεγάλων δεδομένων είναι καίριας σημασίας για πολλές επιχειρήσεις και τομείς, ειδικά στον ενεργειακό τομέα, όπου οι συνεχείς μεταβολές στη ζήτηση ενέργειας απαιτούν γρήγορη και ακριβή ανάλυση των δεδομένων. (Marjani,, 2017)

Οι σύγχρονες τεχνολογίες ανάλυσης δεδομένων, όπως τα Big Data Analytics, το Cloud Computing, και το Internet of Things (IoT), έχουν φέρει επανάσταση στη διαχείριση της ενέργειας. Τα συστήματα IoT παρέχουν συνεχή ροή δεδομένων από έξυπνους μετρητές και αισθητήρες, επιτρέποντας τη συλλογή μεγάλων ποσοτήτων δεδομένων σε πραγματικό χρόνο. Τα δεδομένα αυτά αναλύονται με τη χρήση τεχνολογιών cloud, επιτρέποντας στους διαχειριστές να λαμβάνουν αποφάσεις βάσει πραγματικών στοιχείων, βελτιώνοντας τη διανομή της ενέργειας και μειώνοντας τα κόστη.

Οι αλγόριθμοι ανάλυσης δεδομένων όπως τα Support Vector Machines (SVM) και οι αλγόριθμοι ομαδοποίησης (Clustering Algorithms) χρησιμοποιούνται ευρέως για την ανίχνευση ανωμαλιών στην κατανάλωση και την πρόβλεψη τάσεων. Η επεξεργασία δεδομένων IoT σε συνδυασμό με τη μηχανική μάθηση συμβάλλει στη βελτιστοποίηση των ενεργειακών πόρων (Hashem, I. A. T., et al., 2015).

A.5 Machine Learning

Η Μηχανική Μάθηση (ML) είναι ένας κλάδος της τεχνητής νοημοσύνης (AI) που επιτρέπει στα συστήματα να μαθαίνουν και να βελτιώνουν την απόδοσή τους αυτόνομα μέσω της εμπειρίας, χωρίς την χρήση ρητού προγραμματισμού. Αυτό επιτυγχάνεται με την ανάλυση μεγάλου όγκου δεδομένων και την εφαρμογή αλγορίθμων που μπορούν να εντοπίζουν πρότυπα και να κάνουν προβλέψεις. Στην Μηχανική Μάθηση, οι αλγόριθμοι χωρίζονται συνήθως σε τρεις βασικές κατηγορίες: μάθηση με επίβλεψη, όπου το μοντέλο εκπαιδεύεται σε δεδομένα εισόδου με γνωστές εξόδους, η μάθηση χωρίς επίβλεψη, κατά την οποία το σύστημα αναζητά πρότυπα, και τέλος την ενισχυτική μάθηση, κατά την οποία το σύστημα μαθαίνει μέσω αλληλεπιδράσεων με το περιβάλλον και προσαρμόζεται σε ανταπόκριση σε ανταμοιβές ή ποινές (Christopher, 2006).

Τα νευρωνικά δίκτυα και οι τεχνικές βαθιάς μάθησης έχουν φέρει επανάσταση στον τομέα της μηχανικής μάθησης, επιτρέποντας την ανάλυση πολύπλοκων δεδομένων, όπως εικόνες, ήχος και χρονοσειρές, με μεγάλη ακρίβεια. Τα νευρωνικά δίκτυα μιμούνται τη λειτουργία του ανθρώπινου εγκεφάλου και αποτελούνται από πολλαπλά στρώματα τεχνητών νευρώνων που αναλύουν και επεξεργάζονται δεδομένα σε πολλαπλά επίπεδα. Η βαθιά μάθηση που βασίζεται σε νευρωνικά δίκτυα πολλαπλών επιπέδων έχει επιτρέψει τη μοντελοποίηση πολύπλοκων συστημάτων όπως η πρόβλεψη της κατανάλωσης ενέργειας, η αναγνώριση ομιλίας και η αυτόνομη οδήγηση (Goodfellow I., Bengio Y., & Courville A., 2016).

Στον τομέα της διαχείρισης της ενέργειας, η μηχανική μάθηση έχει αποδειχθεί ιδιαίτερα χρήσιμη για την ανάλυση και την πρόβλεψη της κατανάλωσης ενέργειας, την αποτελεσματική κατανομή των πόρων και τη βελτιστοποίηση των έξυπνων δικτύων. Αλγόριθμοι πρόβλεψης χρονοσειρών, όπως τα επαναλαμβανόμενα νευρωνικά δίκτυα (RNN) και τα δίκτυα μακράς βραχυπρόθεσμης μνήμης (LSTM), έχουν χρησιμοποιηθεί για την ακριβή πρόβλεψη της ζήτησης ενέργειας, λαμβάνοντας υπόψη ιστορικά δεδομένα και εποχιακές διακυμάνσεις. Αυτοί οι αλγόριθμοι μπορούν να εντοπίσουν κρυφά μοτίβα χρήσης ενέργειας, βελτιώνοντας την ικανότητα των παρόχων κοινής ωφέλειας, να διαχειρίζονται καλύτερα τα αποθέματα ενέργειας και να μειώνουν τις απώλειες (Russell S. & Norvig P. , 2010).

Παράλληλα, αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται επίσης για τη βελτιστοποίηση της διαχείρισης ανανεώσιμων πηγών ενέργειας, όπως τα φωτοβολταϊκά και οι ανεμογεννήτριες. Οι καιρικές συνθήκες αποτελούν βασικό παράγοντα που καθορίζει την απόδοση των ανανεώσιμων πηγών ενέργειας και η μηχανική μάθηση καθιστά δυνατή την πρόβλεψη της παραγωγής ενέργειας με βάση τις προαναγγελίες του καιρού και τα ιστορικά δεδομένα. Επιπλέον, η μηχανική μάθηση συμβάλλει στη συντήρηση και τη βελτιστοποίηση των ενεργειακών υποδομών μέσω της προβλεπτικής συντήρησης, επιτρέποντας στους διαχειριστές να εντοπίζουν βλάβες πριν επηρεάσουν την παραγωγή ενέργειας ή την απόδοση του συστήματος (Deb C. & et al., 2017).

Συνολικά, η μηχανική μάθηση έχει τη δυνατότητα να αλλάξει τον τρόπο με τον οποίο οι εταιρείες ενέργειας διαχειρίζονται και προβλέπουν την κατανάλωση και τη διανομή ενέργειας, οδηγώντας σε ένα πιο αποτελεσματικό, ευέλικτο και βιώσιμο ενεργειακό σύστημα.

A.6 Στόχος της Εργασίας

Στόχος της εργασίας είναι η ανάλυση των δεδομένων κατανάλωσης ενέργειας, χρησιμοποιώντας τεχνικές μηχανικής μάθησης και ανάλυσης μεγάλου όγκου δεδομένων. Η ανάλυση αυτή αποσκοπεί στην πρόβλεψη των μελλοντικών αναγκών σε ενέργεια, καθώς και στην κατανόηση των παραγόντων που επηρεάζουν την κατανάλωση, μέσα σε έναν ημερολογιακό χρόνο. Για να επιτευχθεί το αποτέλεσμα, χρησιμοποιήθηκε ένα σύνολο αλγορίθμων και βιβλιοθηκών τα οποία εκτελέστηκαν με διαδοχική σειρά. Ονοματικά και με σειρά χρήσης, είναι τα εξής:

- Pandas
- Numpy
- Matplotlib.pyplot
- Seaborn
- Prophet

Ο αλγόριθμος pandas διαβάζει και διαμορφώνει τα δεδομένα για την βέλτιστη απόδοση των λοιπών αλγορίθμων, με την χρήση του Numpy και Matplotlib.pyplot εξασφαλίζουμε την

μαθηματική επάρκεια του προγράμματος και την οπτικοποίηση των αποτελεσμάτων σε γραφήματα ενώ με το Seaborn τα μορφοποιούμε. Τέλος, το Prophet μας βοηθάει στο κομμάτι της πρόβλεψης της κατανάλωσης ενέργειας σε μελλοντικό χρόνο σχετικά με το δείγμα του dataset.

Σχετικά με την εκτέλεση του προγράμματος, χώρισα τα εκτελέσιμα αρχεία σε 4. Αρχικά, το αρχείο με όνομα energy.py δημιουργεί ένα γράφο που οπτικοποιεί την κατανάλωση ενέργειας μέσα σε ένα ημερολογιακό έτος, ενώ το αρχείο temperature.py πραγματοποιεί το ίδιο για την θερμοκρασία στο ίδιο χρονικό διάστημα. Στην συνέχεια, υπάρχει το αρχείο Forecast_per_LCLid το οποίο πραγματοποιεί την πρόβλεψη και οπτικοποίηση της κατανάλωσης ρεύματος κατά οικοδομικό συγκρότημα, ενώ το forecast_whole.py κάνει την πρόβλεψη και την οπτικοποίηση για όλα τα δεδομένα μαζί.

Η παρούσα εργασία στοχεύει να συνεισφέρει στον τομέα της ενεργειακής διαχείρισης, ενσωματώνοντας τις πιο πρόσφατες εξελίξεις στις τεχνολογίες ανάλυσης δεδομένων και μηχανικής μάθησης. Μέσω της ανάλυσης των δεδομένων κατανάλωσης ενέργειας, θα δοθούν νέες γνώσεις που μπορούν να χρησιμοποιηθούν από διαχειριστές δικτύων και παρόχους ενέργειας για την καλύτερη κατανομή και εξοικονόμηση των ενεργειακών πόρων, ενώ ταυτόχρονα θα μειωθούν τα κόστη παραγωγής και θα ενισχυθεί η βιωσιμότητα.

A.7 Περιορισμοί και Εμβέλεια

Το χρονικό παράθυρο των δεδομένων που συμπεριλαμβάνονται μέσα στο dataset που χρησιμοποιήθηκε καλύπτουν ημερομηνίες από τον Νοέμβρη του 2011 μέχρι και τον Φεβρουάριο του 2014 από 5567 σπίτια που εγκατέστησαν «έξυπνους μετρητές» για την παροχή των δεδομένων κατανάλωσης ηλεκτρικού ρεύματος. Λόγω σχετικά αδύναμου μηχανήματος, όπως προανέφερα, ο μέγιστος αριθμός δεδομένων που λαμβάνει ο οποιοσδήποτε κώδικας που χρησιμοποιήθηκε ανέρχεται στα 4000.

Β' ΜΕΘΟΔΟΛΟΓΙΑ

Β.1 Πηγή Δεδομένων και προ επεξεργασία

Η Ακρίβεια των δεδομένων σε μία ερευνητική εργασία είναι από τα σημαντικότερα σημεία που κρίνουν την αξιοπιστία του αποτελέσματος. Σε αυτό το κεφάλαιο λοιπόν θα εξετάσουμε την πηγή των δεδομένων. Η περισυλλογή των δεδομένων ενέργειας έγινε μέσω της εταιρίας UK Power networks η οποία ισχυρίζεται πως το δείγμα των 5567 σπιτιών που συμμετείχαν στην έρευνα είναι αντιπροσωπευτικά του πληθυσμού του Λονδίνου. Αντίστοιχα, η συλλογή των δεδομένων καιρού έγινε με την χρήση του darksky api της Apple για τις ίδιες χρονικές περιόδους. Μετά την συλλογή δεδομένων, συνήθως ακολουθεί η προ επεξεργασία αυτών ως ένα απαραίτητο βήμα, για την διασφάλιση ότι έχουν αφαιρεθεί τυχών ατέλειες και έχουν τροποποιηθεί τα δεδομένα για την καλύτερη ανάλυση και χρήση τους. Το σύνολο των αρχείων, μας είναι διαθέσιμο στο Kaggle και βρίσκεται στην παρακάτω διεύθυνση:

https://www.kaggle.com/datasets/jeanmidev/smart-meters-in-london?select=acorn_details.csv

Το Kaggle είναι μια πλατφόρμα φιλοξενίας διαφόρων dataset κάνοντάς το έναν πολύ δημοφιλή προορισμό για ερευνητές που ασχολούνται με την μηχανική μάθηση και την ανάλυση μεγάλων δεδομένων. Στην περίπτωση του dataset που χρησιμοποίησα για την έρευνά μου, μου παρείχε όλες τις πληροφορίες που χρειάστηκα αλλά και κάποια επιπλέον δεδομένα όπως ωριαία και μισάωρη μέτρηση κατανάλωσης ρεύματος, και ωριαία μέτρηση καιρού καθώς και δεδομένα για τα σπιτικά που συμμετείχαν στην έρευνα.

Β.2 Εξερεύνηση και ανάλυση δεδομένων

Η εξερεύνηση και η ανάλυση των δεδομένων είναι καίριο κομμάτι κάθε έρευνας διότι χάρη στην βοήθειά τους μπορούμε να αναγνωρίσουμε διάφορα χαρακτηριστικά των δεδομένων και να εξερευνήσουμε διάφορες δυσκολίες που μπορεί να μας παρουσιάσουν.

Στην προκειμένη περίπτωση, με την εξερεύνηση και ανάλυση των δεδομένων, εξετάστηκε ο τύπος των δεδομένων και ο τρόπος διατύπωσής τους, πράγμα που βοήθησε στο να αναγνωριστούν ορισμένες εκχωρήσεις που είχαν ελλιπή δεδομένα και αργότερα να αντιμετωπιστούν στο κομμάτι του κώδικα.

B.3 Επιλογή Αναλυτικών Εργαλείων

Η επιλογή των εργαλείων που χρησιμοποιήθηκαν έγινε με γνώμονα την αποτελεσματικότητα και την αποδοτικότητα των εργαλείων. Ειδικότερα, Η γλώσσα Python επιλέχθηκε επειδή παρέχει πολλές βιβλιοθήκες για ανάλυση δεδομένων και το λεξιλόγιό της κάνει την διαδικασία αρκετά κατανοητή (McKinney, 2010). Το λειτουργικό σύστημα επιλέχθηκε λόγω της γηγενούς υποστήριξης της Python αλλά και την υποστήριξη όλων των υπόλοιπων εργαλείων. Το Pandas επιλέχθηκε διότι είναι γρήγορο και αποτελεσματικό στην διαχείριση πληροφοριών, αλλά προσφέρει και πολλές δυνατότητες μετατροπής και ανάλυσης δεδομένων. Το matplotlib σε συνεργασία με το seaborn προσφέρουν εύκολη και προσωποποιημένη απεικόνιση των δεδομένων με την χρήση γράφων ή διαγραμμάτων. Τέλος, επιλέχτηκε και η βιβλιοθήκη prophet παλαιότερα γνωστή και ως fbprophet για την πραγματοποίηση προβλέψεων με ακρίβεια και ευκολία, βάσει των δεδομένων του dataset.

B.4 Συλλογή δεδομένων και Μεθοδολογία

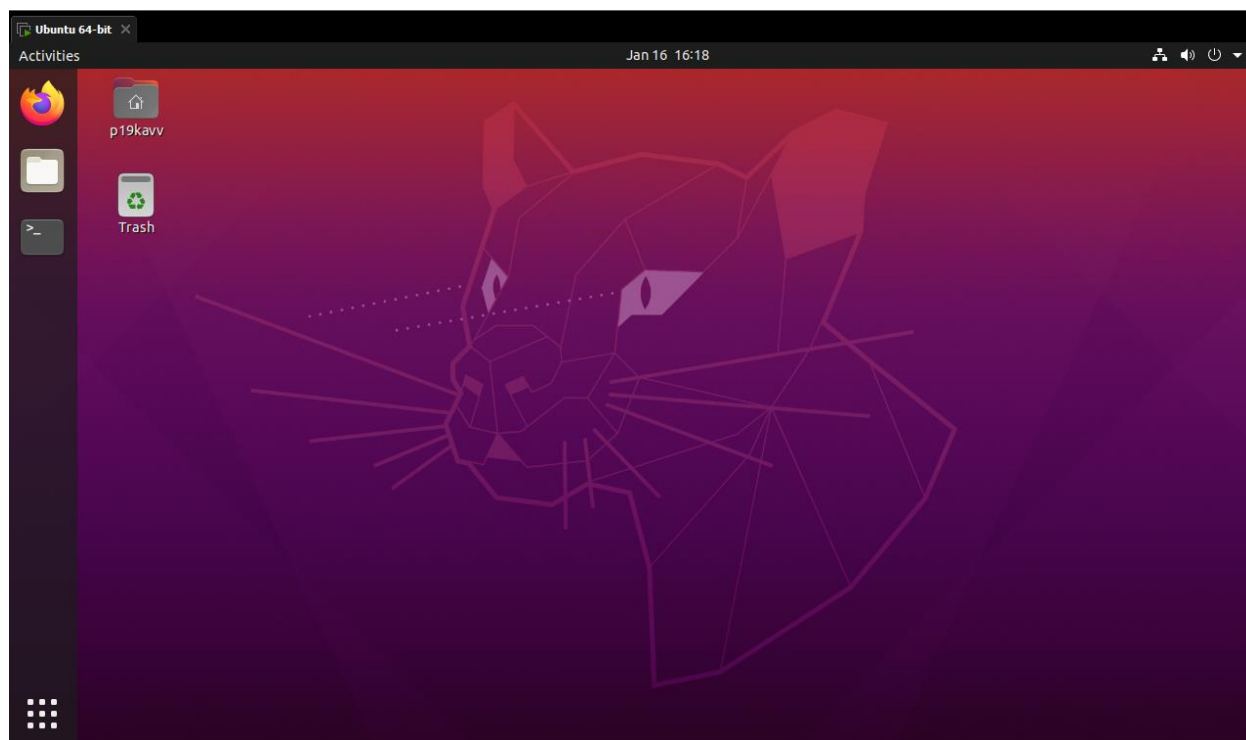
Όπως προαναφέρθηκε, τα δεδομένα που χρησιμοποιήθηκαν, λήφθηκαν από το Kaggle, και με βάση τις ανάγκες που θεώρησα ότι θα μου δημιουργήσει, ξεκίνησα την εγκατάσταση των συστημάτων που αναφέρω παρακάτω για την επεξεργασία των δεδομένων και την εξαγωγή συμπερασμάτων. Για την πραγματοποίηση της εργασίας και την παραγωγή ορθών αποτελεσμάτων, χρησιμοποιήθηκε λειτουργικό σύστημα Ubuntu 20.04 LTS με εγκατάσταση python3. Στην συνέχεια φόρτωσα το dataset που επέλεξα να αναλύσω. Τα δεδομένα είναι μορφοποιημένα και δεν έχουν άκυρες τιμές επί το πλείστον οπότε είναι έτοιμα για άμεση χρήση. Το επόμενο βήμα μου ήταν η προετοιμασία του λειτουργικού συστήματος με την εγκατάσταση

των αναγκαίων βιβλιοθηκών, ώστε να μπορεί να πραγματοποιηθεί η επεξεργασία των δεδομένων και η εξαγωγή συμπερασμάτων από αυτά. Αρχικά, έκανα εγκατάσταση την βιβλιοθήκη του `pandas` με την χρήση εντολής τερματικού, καθώς είναι από τις κύριες βιβλιοθήκες που χρειάστηκα για την επεξεργασία των δεδομένων και την εξαγωγή αποτελεσμάτων. Το αμέσως επόμενο βήμα, είναι η εγκατάσταση της `matplotlib` βιβλιοθήκης η οποία μας επιτρέπει να δημιουργήσουμε διάφορα σχέδια για την οπτικοποίηση δεδομένων, ώστε να γίνεται πιο κατανοητό το αποτέλεσμα μας και να κάνει πιο εύκολες διάφορες συγκρίσεις ανάμεσα στα δεδομένα μας. Στην συνέχεια, εγκαταστάθηκε το `pip` της `python` για την ευκολότερη εγκατάσταση των λοιπών βιβλιοθηκών. Η πρώτη βιβλιοθήκη που συμπεριλήφθηκε με την χρήση του `pip` είναι η `NumPy`, και συγκεκριμένα η έκδοση 1.22.1, επειδή οι πιο σύγχρονες εκδόσεις δεν λειτουργούν σωστά σε συνεργασία με τις υπόλοιπες βιβλιοθήκες διότι μετά από αυτή την έκδοση άλλαξαν κάποιες βασικές διεργασίες στην `NumPy`, άρα είμαστε αναγκασμένοι να χρησιμοποιήσουμε μία παλαιότερη έκδοση. Την `NumPy` ακολουθεί η βιβλιοθήκη `Seaborn`, χάρη στην οποία μπορούμε να δώσουμε οδηγίες στην `matplotlib` για την οπτικοποίηση των δεδομένων, στην προκειμένη περίπτωση, των γράφων που σκοπεύουμε να δημιουργήσουμε. Τέλος, οι βιβλιοθήκες `rystan` (συγκεκριμένα η έκδοση 2.19.1.1) και `prophet` εγκαταστάθηκαν για την πρόβλεψη της κατανάλωσης ρεύματος με βάση τα δεδομένα μας.

Γ' ΠΡΟΕΤΟΙΜΑΣΙΑ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

Γ.1 Προετοιμασία του περιβάλλοντος έρευνας

Aρχικά, έγινε η εγκατάσταση του λειτουργικού συστήματος ubuntu 20.04 lts σε μία εικονική μηχανή με την χρήση του iso που διατίθεται δωρεάν στον εξής [ιστότοπο](#) και ολοκλήρωσα την διαδικασία.

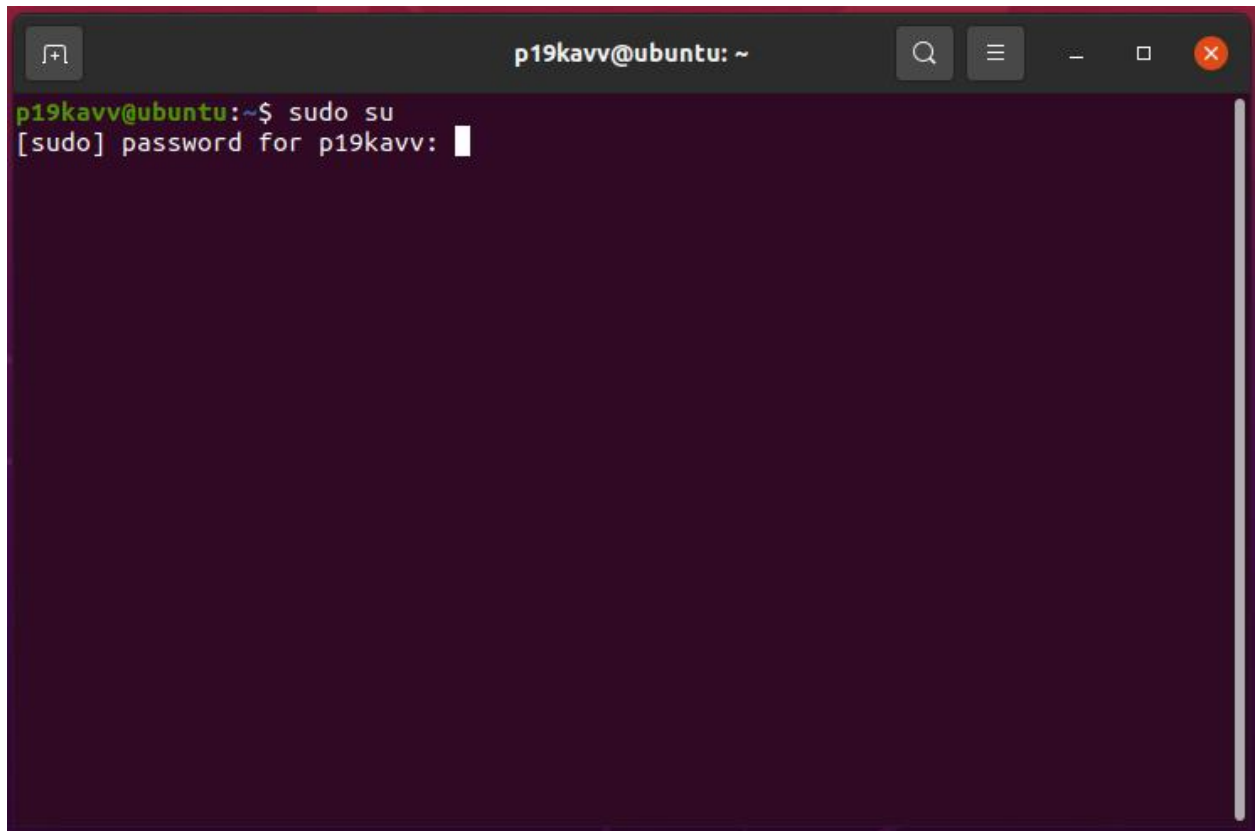


Εικόνα 1 – επιφάνεια εργασίας

Στην συνέχεια, δημιούργησα έναν φάκελο για την αποθήκευση του dataset και του κώδικα που έγραψα αργότερα με θέση στο `/home/p19kavn/thesis/ed`.

Το επόμενο βήμα μου ήταν να ανοίξω το τερματικό και να κάνω την εγκατάσταση όλων των προαναφερθέντων βιβλιοθηκών τις οποίες θα αναφέρω ονοματικά μαζί με τα στιγμιότυπά τους καθώς η επεξήγησή τους έγινε παραπάνω. Η διαδικασία ήταν η εξής:

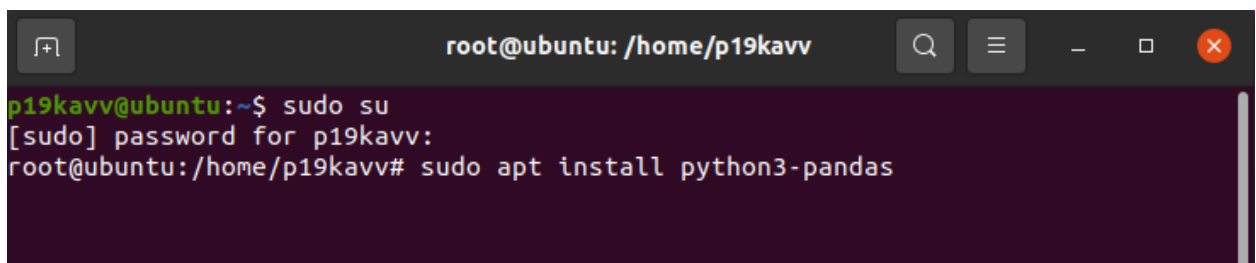
- 1) Πρώτα απ' όλα ζήτησα δικαιώματα διαχειριστή από το τερματικό για την ταχύτερη εγκατάσταση των βιβλιοθηκών:



```
p19kavv@ubuntu:~$ sudo su
[sudo] password for p19kavv: 
```

Εικόνα 2 – ανύψωση τερματικού σε διαχειριστή

- 2) Στην συνέχεια έγινε εγκατάσταση του Pandas και pandas-doc :



```
root@ubuntu: /home/p19kavv
p19kavv@ubuntu:~$ sudo su
[sudo] password for p19kavv:
root@ubuntu:/home/p19kavv# sudo apt install python3-pandas
```

Εικόνα 3 – εγκατάσταση Pandas

```
root@ubuntu:/home/p19kavv# sudo apt install python-pandas-doc
```

Εικόνα 4 – εγκατάσταση Pandas-doc

- 3) Το επόμενο βήμα ήταν η εγκατάσταση του matplotlib:

```
root@ubuntu:/home/p19kavv# sudo apt-get install python3-matpl
```

Εικόνα 5 – εγκατάσταση Matplotlib

- 4) Αμέσως μετά, χρειάστηκε να εγκαταστήσω το python pip ώστε να μπορέσω να εγκαταστήσω τις υπόλοιπες βιβλιοθήκες με ευκολία:

```
root@ubuntu:/home/p19kavv# sudo apt install python3-pip
```

Εικόνα 6 – εγκατάσταση python-pip

- 5) Με την βοήθεια του pip, συνέχισα στο seaborn και την έκδοση 1.22.1 του NumPy:

```
root@ubuntu:/home/p19kavv# pip install seaborn
```

Εικόνα 7 – εγκατάσταση Seaborn

```
root@ubuntu:/home/p19kavv# pip install numpy==1.22.1
```

Εικόνα 8 – εγκατάσταση Numpy 1.22.1

6) Τέλος, συνέχισα με την εγκατάσταση των προαπαιτούμενων και του ίδιου του prophet:

```
root@ubuntu:/home/p19kavv# python3 -m pip install pystan==2.19.1.1
```

Εικόνα 9 – εγκατάσταση pystan 2.19.1.1

```
root@ubuntu:/home/p19kavv# sudo apt install python3-dev
```

Εικόνα 10 – εγκατάσταση python-dev

```
root@ubuntu:/home/p19kavv# python3 -m pip install prophet
```

Εικόνα 11 – εγκατάσταση prophet

7) Αφού εγκατέστησα όλες τις βιβλιοθήκες που χρειαζόμουν, έκανα *update* και *upgrade* όλες τις βιβλιοθήκες με τις κατάλληλες εντολές και τοποθέτησα το *dataset* στον φάκελο */home/p19kavv/thesis/ed*

Γ.2' Προετοιμασία δεδομένων

Η προετοιμασία των δεδομένων είναι καθοριστικής σημασίας για τη σωστή εκτέλεση των αλγορίθμων ανάλυσης και πρόβλεψης. Στην παρούσα εργασία, τα δεδομένα που χρησιμοποιήθηκαν αφορούσαν ημερήσιες καταγραφές κατανάλωσης ενέργειας από έξυπνους μετρητές, οι οποίες συλλέχθηκαν σε επίπεδο νοικοκυριού και στο μεγαλύτερο επίπεδο όλης της πόλης.

Για την προετοιμασία των δεδομένων ακολούθησα τα παρακάτω βήματα:

1. Συλλογή και Καθαρισμός Δεδομένων:

Τα δεδομένα κατανάλωσης ήταν ήδη σε καλή μορφή, χωρίς ελλειπείς ή άκυρες τιμές, οπότε δεν χρειάστηκε εκτεταμένος καθαρισμός. Η διαδικασία επικεντρώθηκε στη φόρτωση των δεδομένων και τη σωστή διαχείρισή τους μέσω της βιβλιοθήκης Pandas.

2. Μορφοποίηση και Διαχείριση των Χρονοσειρών:

Οι χρονοσειρές που αντιπροσωπεύουν την κατανάλωση ενέργειας επεξεργάστηκαν με τη χρήση του NumPy και του Pandas, που επιτρέπουν τη δημιουργία πινάκων και τον χειρισμό μεγάλων δεδομένων. Στόχος ήταν η μετατροπή των ημερήσιων καταγραφών σε κατάλληλη μορφή για τη μετέπειτα ανάλυση με αλγορίθμους πρόβλεψης.

3. Οργάνωση Δεδομένων για Μοντέλα Πρόβλεψης:

Τα δεδομένα κατηγοριοποιήθηκαν με τέτοιο τρόπο ώστε να είναι έτοιμα να εισαχθούν στους αλγόριθμους πρόβλεψης, όπως το Prophet, το οποίο απαιτεί δομημένες χρονοσειρές για τη δημιουργία ακριβών προβλέψεων.

Γ.3' Λογική ανάπτυξης κώδικα

Με το περιβάλλον έτοιμο και το dataset τοποθετημένο στον φάκελο που θα γίνουν οι έρευνες, το επόμενο βήμα είναι η ανάπτυξη του κώδικα για την ανάλυση και την οπτικοποίηση της μέσης ημερήσιας κατανάλωσης ρεύματος. Αρχικά, προσπάθησα να συμπεριλάβω και την ανάλυση της μέσης ημερήσιας θερμοκρασίας στο ίδιο αρχείο αλλά δημιουργήθηκαν προβλήματα στην εκτέλεση και προκειμένου να τα αντιμετωπίσω αποφάσισα να δημιουργήσω διαφορετικά εκτελέσιμα αρχεία για το κάθε ζητούμενο. Η επιλογή αυτή αποδείχτηκε ευκολότερη και λιγότερο πολύπλοκη στην πράξη, αλλά παρόλα αυτά αρκετά αποτελεσματική.

Γ.4' Ανάπτυξη και αποτελέσματα κώδικα

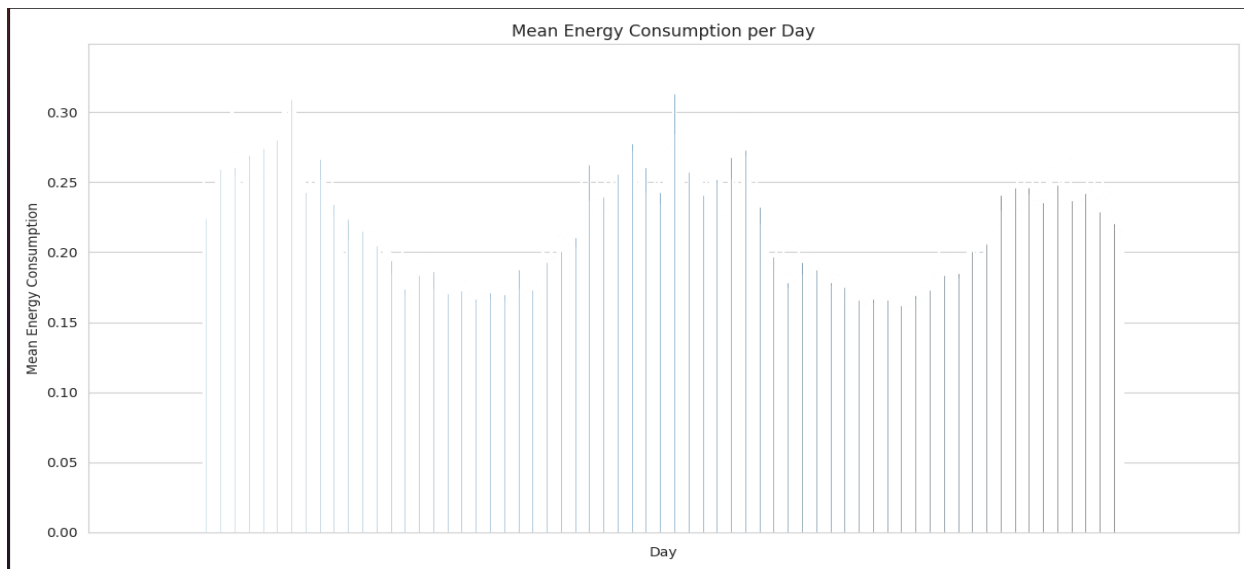
Γ.4.1 Energy.py

Εξίσωση 1 - Κώδικας του αλγόριθμου Energy.py

```
energy.py X
1 import pandas as pd
2 import matplotlib.pyplot as plt
3 import seaborn as sns
4
5 # Load data from daily_dataset.csv
6 power_df = pd.read_csv('daily_dataset.csv')
7
8 # Convert 'day' column to datetime
9 power_df['day'] = pd.to_datetime(power_df['day'])
10
11 # Limit the analysis to the first 60 entries
12 power_df = power_df.head(60)
13
14 # Group the data by day and calculate the mean energy consumption per day
15 mean_energy_per_day = power_df.groupby('day')['energy_mean'].mean().reset_index()
16
17 # Create a bar plot
18 plt.figure(figsize=(12, 6))
19 sns.set_style("whitegrid")
20 sns.barplot(x='day', y='energy_mean', data=mean_energy_per_day, palette="Blues_d")
21 plt.xticks(rotation=90) # Rotate x-axis labels for better readability
22
23 plt.xlabel('Day')
24 plt.ylabel('Mean Energy Consumption')
25 plt.title('Mean Energy Consumption per Day')
26 plt.tight_layout()
27
28 # Show the plot
29 plt.show()
30
```

Πίσω στην μέση ημερήσια κατανάλωση ενέργειας λοιπόν, σε αυτό το κομμάτι του κώδικα χρησιμοποιήθηκαν οι pandas, matplotlib και seaborn, πρώτα φόρτωσα τα αρχεία του dataset που χρειάζομαι με το pandas, μετέτρεψα τις μέρες σε μορφή ημερομηνίας/ώρας και υπολόγισα την μέση ημερήσια κατανάλωση ενέργειας με την συνάρτηση mean() ανά ημέρα. Τέλος, δημιούργησα έναν γράφο και τον μορφοποίησα με το seaborn ώστε να εμφανίζεται με την εκτέλεση, στον γράφο αυτό απεικονίζονται τα αποτελέσματα της παραπάνω διαδικασίας. Συγκεκριμένα η έξοδος είναι η εξής:

Πίνακας 1 – Συνολική μέση κατανάλωση ενέργειας ανά ημέρα



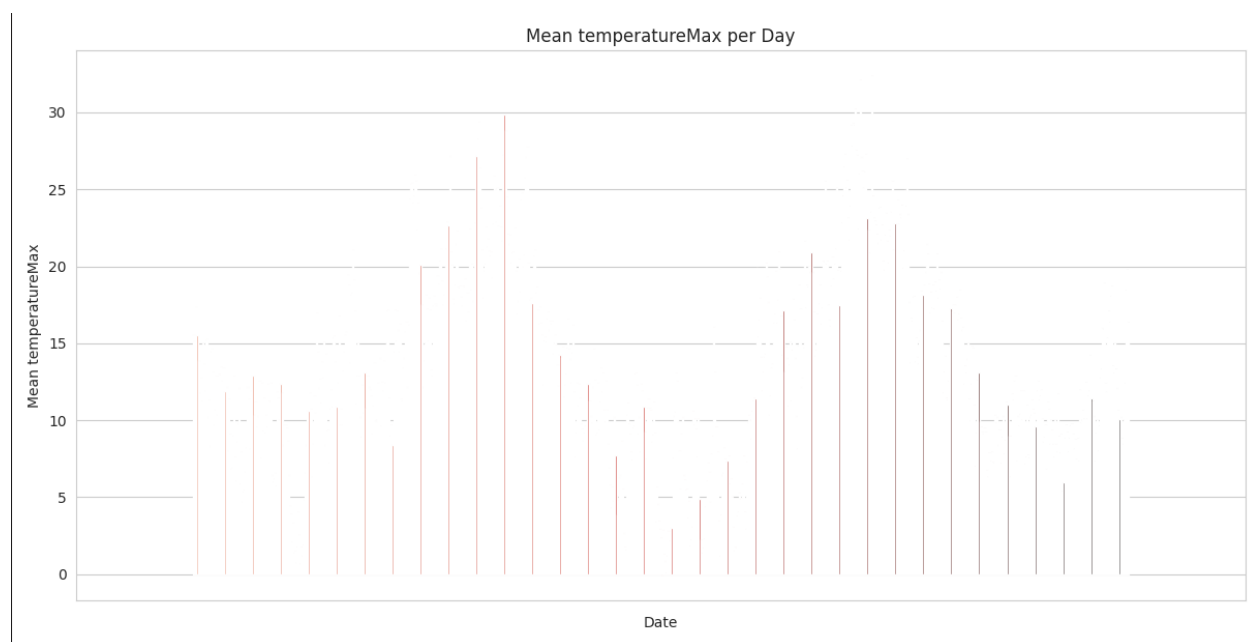
Γ.4.2 Temperature.py

Εξίσωση 2 – Κώδικας του αλγόριθμου Temperature.py

```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3 import seaborn as sns
4
5 # Load data from weather_daily_darksky.csv
6 weather_df = pd.read_csv('weather_daily_darksky.csv')
7
8 # Convert 'temperatureMaxTime' column to datetime
9 weather_df['temperatureMaxTime'] = pd.to_datetime(weather_df['temperatureMaxTime'])
10
11 # Ignore the first 43 entries and take the next 60 entries
12 #weather_df = weather_df.iloc[53:103] # Adjust the slice to select the desired range
13
14 # Group the data by date (ignoring time) and calculate the mean temperatureMax per date
15 weather_df['date'] = weather_df['temperatureMaxTime'].dt.date
16 mean_temp_max_per_day = weather_df.groupby('date')['temperatureMax'].mean().reset_index()
17
18 # Create a bar plot
19 plt.figure(figsize=(12, 6))
20 sns.set_style("whitegrid")
21 sns.barplot(x='date', y='temperatureMax', data=mean_temp_max_per_day, palette="Reds_d")
22 plt.xticks(rotation=90) # Rotate x-axis labels for better readability
23 plt.xlabel('Date')
24 plt.ylabel('Mean temperatureMax')
25 plt.title('Mean temperatureMax per Day (After Ignoring First 43 Entries and Taking Next 60)')
26 plt.tight_layout()
27
28 # Show the plot
29 plt.show()
30
```

Παρόμοια διαδικασία ακολούθησα και για το δεύτερο εκτελέσιμο αρχείο, που υπολογίζει και οπτικοποιεί σε γράφο την μέση ημερήσια θερμοκρασία, από τα δεδομένα, με σκοπό την εύκολη σύγκριση της σχέσης της θερμοκρασίας και της κατανάλωσης ηλεκτρικού ρεύματος ανά ημέρα, οπότε έγινε και η χρήση των ίδιων βιβλιοθηκών. Η ανάπτυξη ξεκίνησε φορτώνοντας τα δεδομένα της θερμοκρασίας και μορφοποιώντας την ημερομηνία της μέγιστης ανά ημέρα θερμοκρασίας στην σωστή μορφή. Συνεχίζοντας, υπολόγισα την μέση μέγιστη θερμοκρασία ανά ημέρα για την χρήση της στην οπτικοποίηση. Τέλος, όπως και με το energy.py δημιούργησα ένα γράφο με ίδιες προδιαγραφές μορφοποίησης για να μπορέσω να συγκρίνω την κατανάλωση ενέργειας με τις αυξομειώσεις της θερμοκρασίας. Τα αποτελέσματα του γράφου παραθέτονται παρακάτω:

Πίνακας 2 – Μέση μέγιστη θερμοκρασία ανά ημέρα



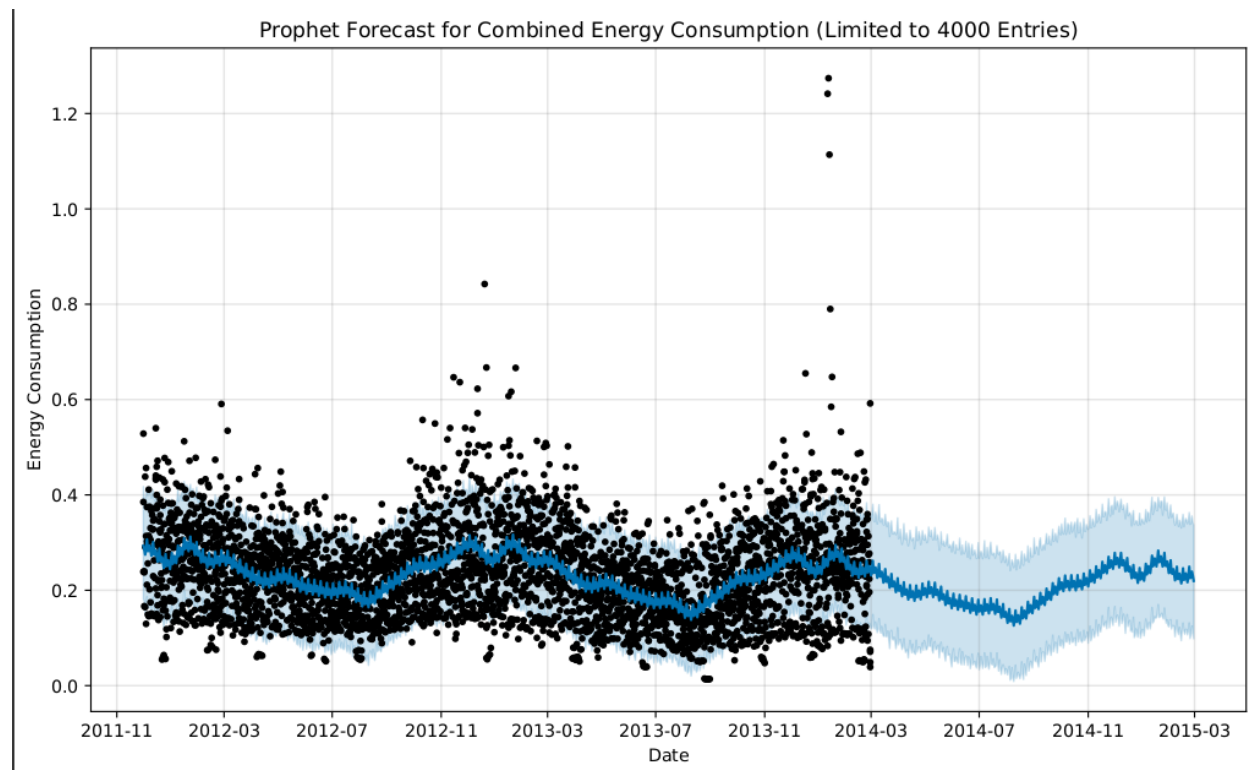
Γ.4.3 forecast_whole.py

Εξίσωση 3 - Κώδικας του αλγόριθμου forecast_whole.py

```
1 import pandas as pd
2 from prophet import Prophet
3 import matplotlib.pyplot as plt
4 from matplotlib.backends.backend_pdf import PdfPages
5
6 # Load the daily dataset
7 df = pd.read_csv('daily_dataset.csv')
8
9 # Get all unique LCLids
10 unique_LCLids = df['LCLid'].unique()
11
12 # Initialize a DataFrame to store aggregated data
13 aggregated_df = pd.DataFrame(columns=['ds', 'y'])
14
15 # Initialize a counter for the number of entries used
16 entry_count = 0
17
18 # Loop through the unique LCLids
19 for LCLid in unique_LCLids:
20     # Filter data for the current LCLid
21     df_selected = df[df['LCLid'] == LCLid]
22
23     # Check if there are at least 2 non-NaN rows for this 'LCLid'
24     if df_selected.dropna().shape[0] < 2:
25         print(f'Skipping LCLid: {LCLid} - Insufficient non-NaN data.')
26         continue
27
28     # Rename columns as required by Prophet
29     df_selected = df_selected.rename(columns={'day': 'ds', 'energy_mean': 'y'})
30
31     # Append the data for this 'LCLid' to the aggregated DataFrame
32     aggregated_df = pd.concat([aggregated_df, df_selected[['ds', 'y']]], ignore_index=True)
33
34     # Increment the entry count
35     entry_count += df_selected.shape[0]
36
37     # Check if we have reached the limit of required entries
38     if entry_count >= 4000:
39         break # Exit the loop if we have enough entries
40
41 # Initialize and fit Prophet model on the aggregated data
42 model = Prophet()
43 model.fit(aggregated_df)
44
45 # Create a dataframe for future dates (forecasting for 1 year into the future)
46 future = model.make_future_dataframe(periods=365)
47
48 # Make forecasts
49 forecast = model.predict(future)
50
51 # Plot the forecast
52 fig = model.plot(forecast)
53 plt.title('Prophet Forecast for Combined Energy Consumption (Limited to 4000 Entries)')
54 plt.xlabel('Date')
55 plt.ylabel('Energy Consumption')
56
57 # Save the plot to a PDF file
58 pdf_pages = PdfPages('forecast_combined_limited.pdf')
59 pdf_pages.savefig(fig, bbox_inches='tight')
60 pdf_pages.close()
61
62 # You can access the combined forecast data in the 'forecast' DataFrame.
```

Αφού παρήγαγα τους γράφους για την μέση ημερήσια θερμοκρασία και την μέση κατανάλωση ρεύματος, επόμενο βήμα ήταν η συγγραφή κώδικα που εκμεταλλεύεται τις βιβλιοθήκες `pandas`, `prophet` και `matplotlib` για την δημιουργία προβλέψεων κατανάλωσης ενέργειας μετά το πέρας του χρονικού περιθωρίου των δεδομένων, και την οπτικοποίηση των προβλέψεων αυτών με βάση το `dataset`. Η διαδικασία αυτή πραγματοποιήθηκε αρχικά φορτώνοντας τα αρχεία της ημερήσιας κατανάλωσης και τα αναγνωριστικά ανά οικοδομικό τετράγωνο. Αμέσως μετά δημιούργησα ένα πλαίσιο δεδομένων για να μπορούν να αποθηκευτούν όλα τα αναγκαία δεδομένα και με την βοήθεια ενός βρόγχου προσπέρασα τα 4000 (τέσσερις χιλιάδες) στοιχεία. Μετά από την διαδικασία αυτή ο κώδικας σε συνεργασία με το `prophet` επεξεργάζεται τα δεδομένα και παράγει ένα οπτικό αποτέλεσμα με την βοήθεια του `matplotlib` και του `seaborn`, αποθηκεύοντας το αποτέλεσμα σε ένα `pdf` αρχείο.

Πίνακας 3 – Πρόβλεψη συνδυαστικής κατανάλωσης ενέργειας από `prophet`



Γ.4.4 Forecast_per_LCLid

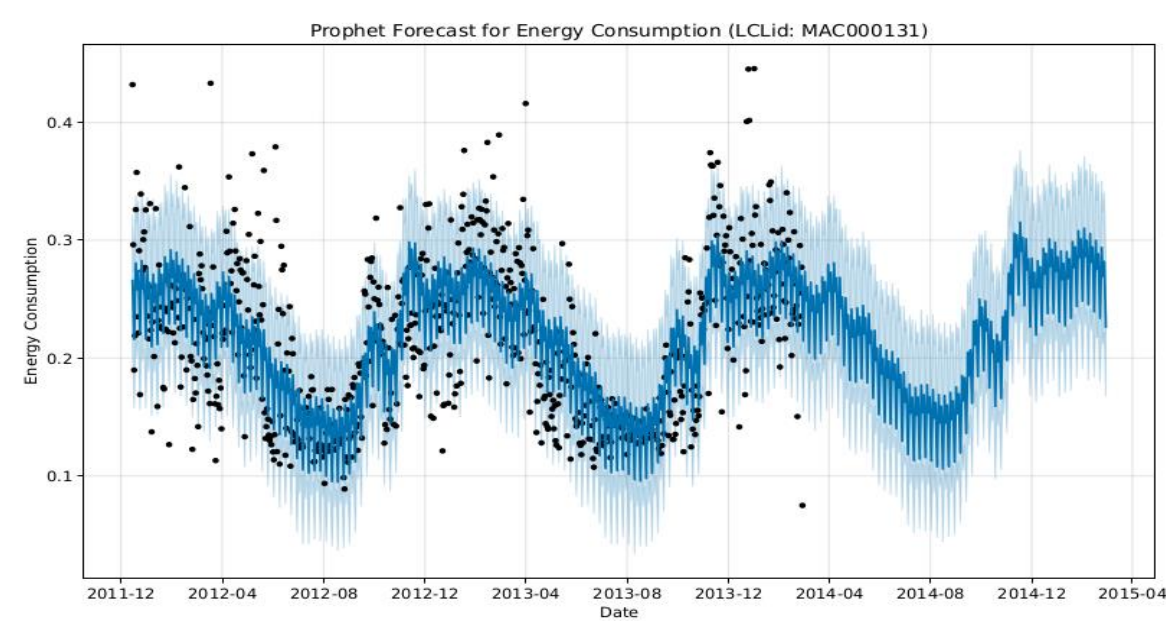
Εξίσωση 4 - Κώδικας του αλγόριθμου Forecast_per_LCLid

```
1  import pandas as pd
2  from prophet import Prophet
3  import matplotlib.pyplot as plt
4  from matplotlib.backends.backend_pdf import PdfPages
5
6  # Load the daily dataset
7  df = pd.read_csv('daily_dataset.csv')
8
9  # Get the first 10 unique LCLids for demonstration
10 unique_LCLids = df['LCLid'].unique()[:10]
11
12 # Initialize a dictionary to store forecasts
13 forecasts = {}
14
15 # Loop through the unique LCLids
16 for LCLid in unique_LCLids:
17     # Filter data for the current LCLid
18     df_selected = df[df['LCLid'] == LCLid]
19
20     # Rename columns as required by Prophet
21     df_selected = df_selected.rename(columns={'day': 'ds', 'energy_mean': 'y'})
22
23     # Initialize and fit Prophet model
24     model = Prophet()
25     model.fit(df_selected)
26
27     # Create a dataframe for future dates (forecasting for 1 year into the future)
28     future = model.make_future_dataframe(periods=365)
29
30     # Make forecasts
31     forecast = model.predict(future)
32
33     # Store the forecast in the dictionary
34     forecasts[LCLid] = forecast
35
36     # Create a PDF file for saving the plots
37     pdf_pages = PdfPages(f'forecast_LCLid_{LCLid}.pdf')
38
39     # Plot the forecast for demonstration
40     fig = model.plot(forecast)
41     plt.title(f'Prophet Forecast for Energy Consumption (LCLid: {LCLid})')
42     plt.xlabel('Date')
43     plt.ylabel('Energy Consumption')
44
45     # Save the plot to the PDF file
46     pdf_pages.savefig(fig, bbox_inches='tight')
47
48     # Close the PDF file
49     pdf_pages.close()
50
51 # You can access forecasts[LCLid] to retrieve the forecast data for a specific LCLid.
52
```

Το τελευταίο εκτελέσιμο που δημιούργησα, είχε ως σκοπό την ανάλυση της κατανάλωσης και την παραγωγή προβλέψεων ανά νοικοκυριό. Αυτό πραγματοποιήθηκε με παρόμοιο τρόπο σε σχέση με το προηγούμενο αρχείο που αναφέρθηκε, τα βήματα που ακολουθήθηκαν ήταν η φόρτωση των δεδομένων ηλεκτρικής κατανάλωσης, μεταφορά σε ένα πλαίσιο δεδομένων των πρώτων 11 αναγνωριστικών, καθώς ο σκοπός μου είναι να επιδείξω ότι ο κώδικας λειτουργεί, αλλά με το μηχάνημα που διαθέτω η επανάληψη της διαδικασίας για πάνω από 5000 εισχωρήσεις είναι πρακτικά αδύνατη. Οπότε, για την προσωπική μου ευκολία αλλά και για χάρη της εργασίας, περιόρισα τα αποτελέσματα στα πρώτα 11. Λέγοντας αυτό όμως ο κώδικας είναι προσαρμοσμένος ώστε να αλλάζει με ευκολία για να συμπεριλάβει επιπλέον δεδομένα αλλά και συγκεκριμένα αναγνωριστικά αν θέλουμε να δούμε τα αποτελέσματα για κάποιο σπίτι συγκεκριμένα. Στην συνέχεια, έγινε επεξεργασία στηλών όπως απαιτεί το prophet και δημιουργήθηκε πλαίσιο δεδομένων για την πρόβλεψη, αλλά και προετοιμάστηκε η μορφοποίηση του pdf. Λόγω ευχρηστίας αποφάσισα η αποθήκευση εκάστοτε πρόβλεψης να γίνεται σε ένα αρχείο pdf που συμπεριλαμβάνει το αναγνωριστικό του σπιτιού στο όνομά του. Παρακάτω παραθέτω τα 11 αποτελέσματα:

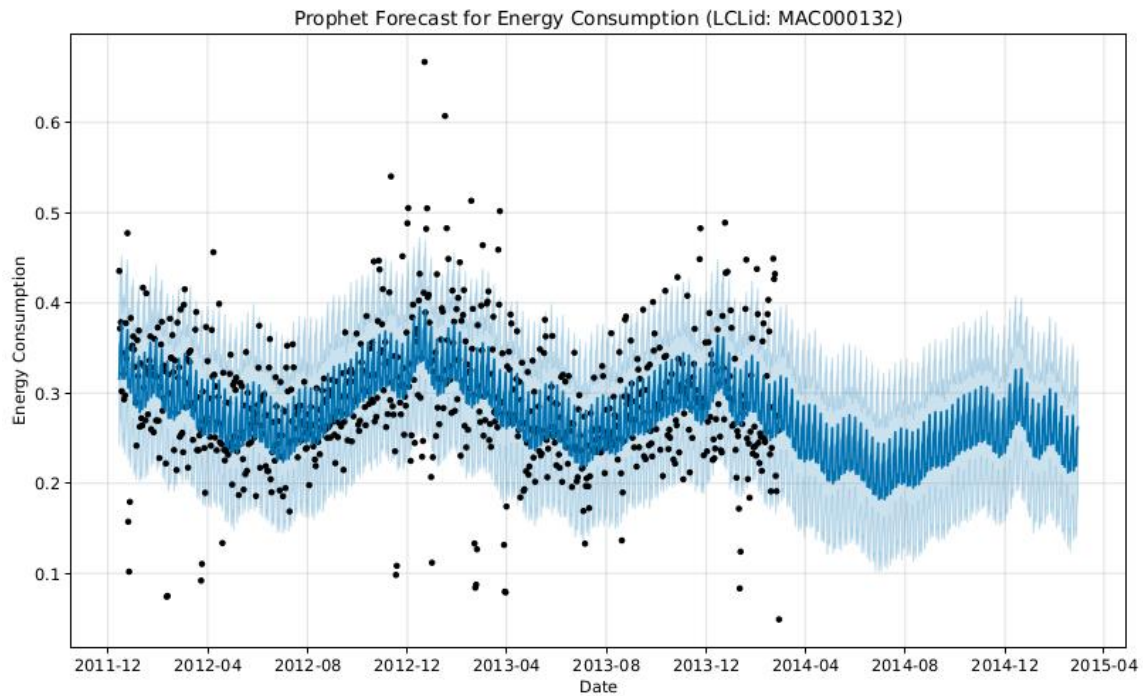
1) MAC000131:

Πίνακας 4 – πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000131



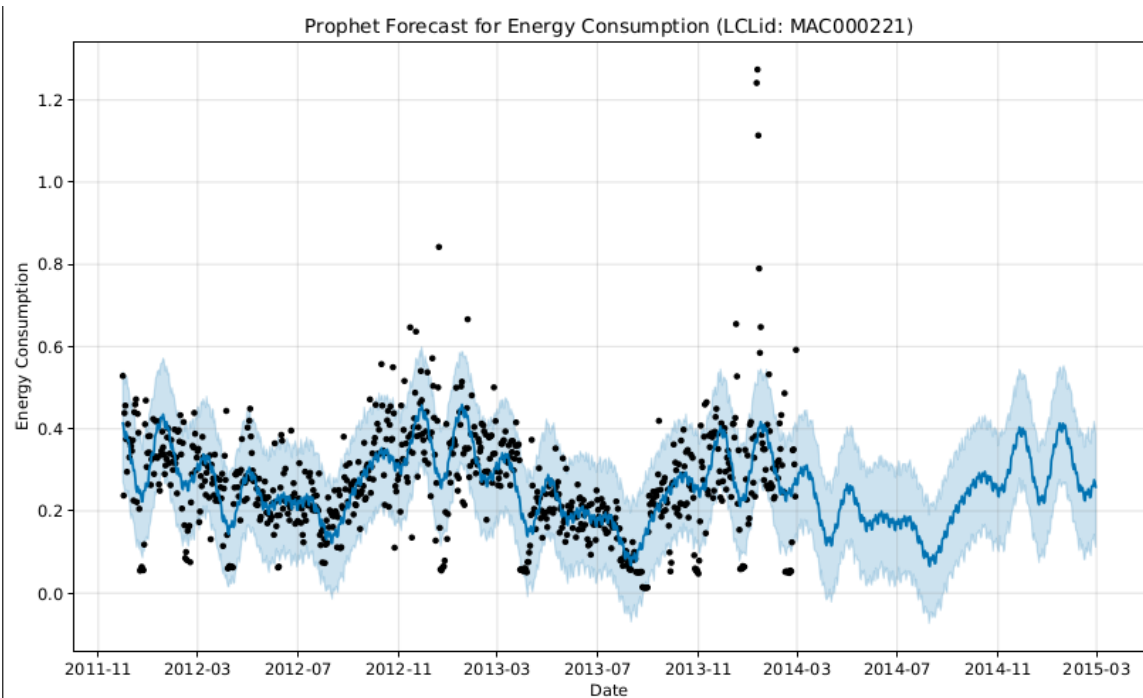
2) MAC000132:

Πίνακας 5 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000132



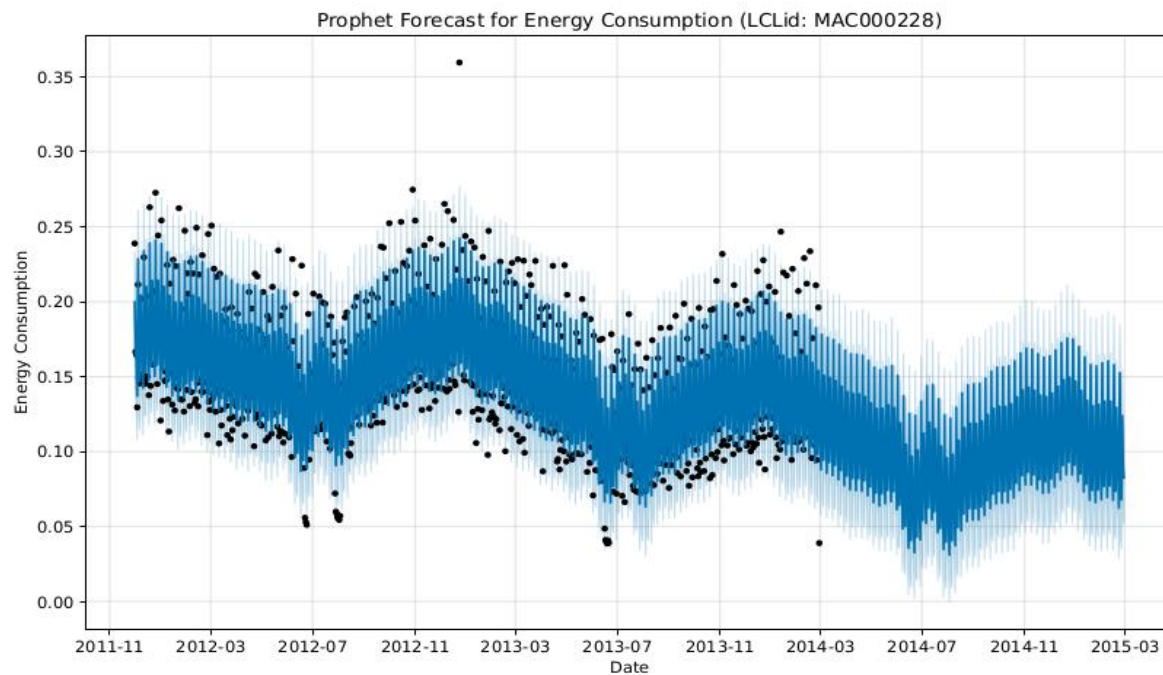
3) MAC000221:

Πίνακας 6 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000221



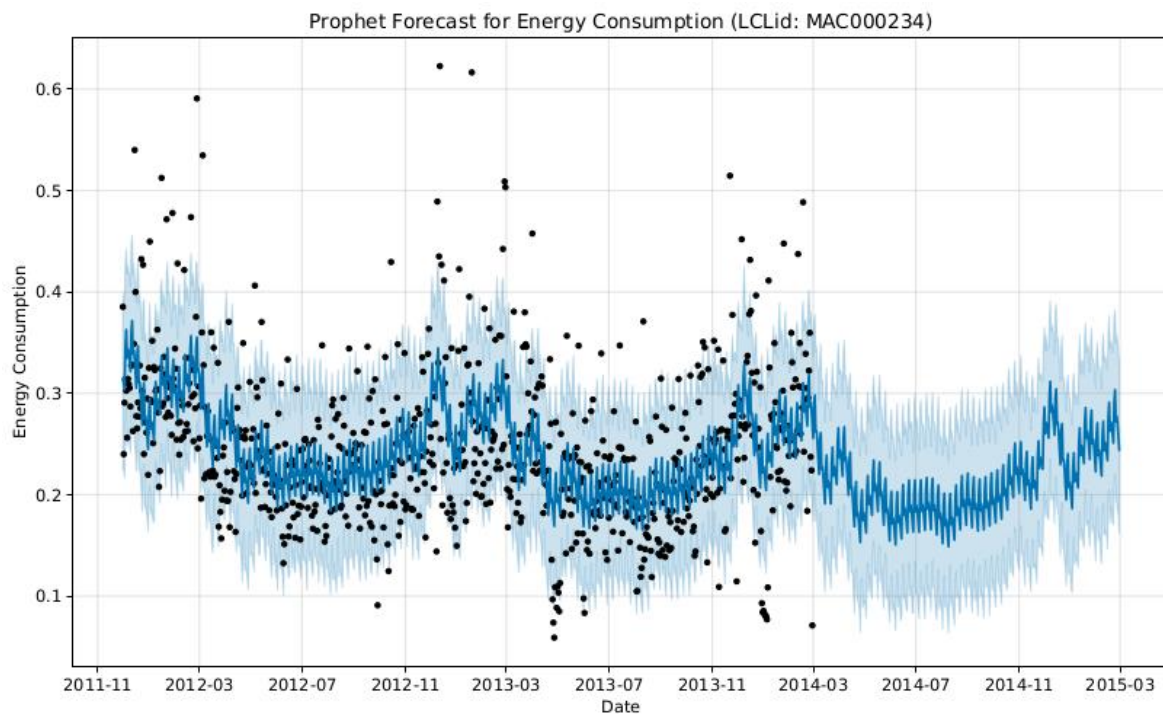
4) MAC000222:

Πίνακας 7 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000228



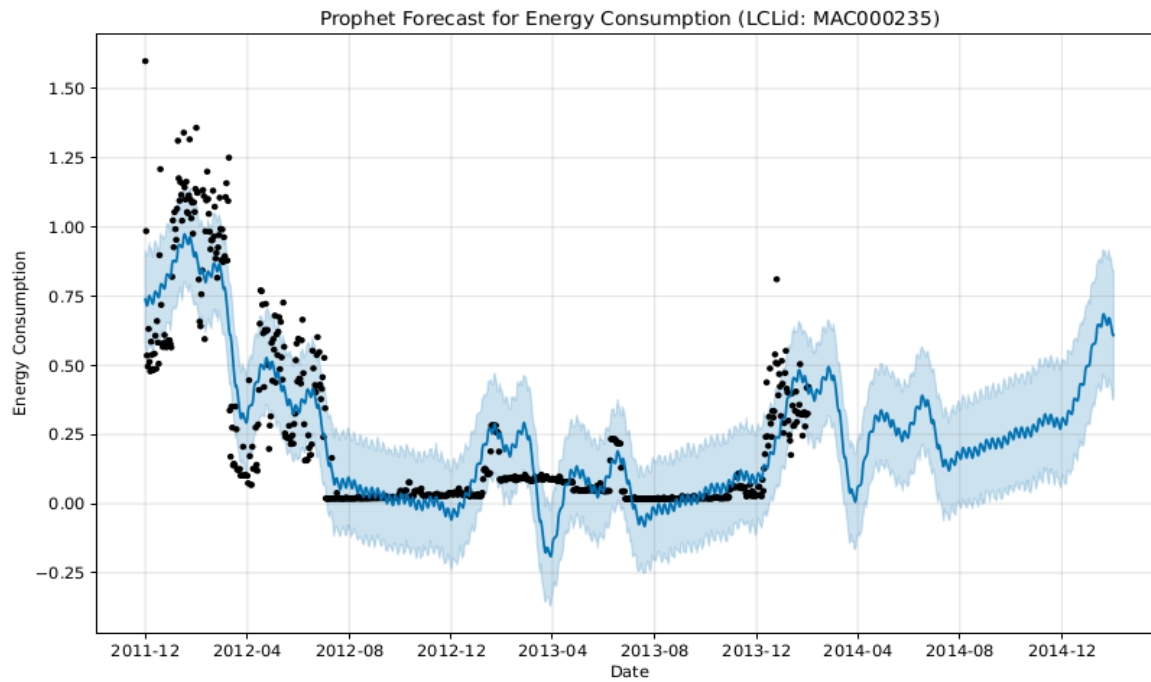
5) MAC000234:

Πίνακας 8 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000234



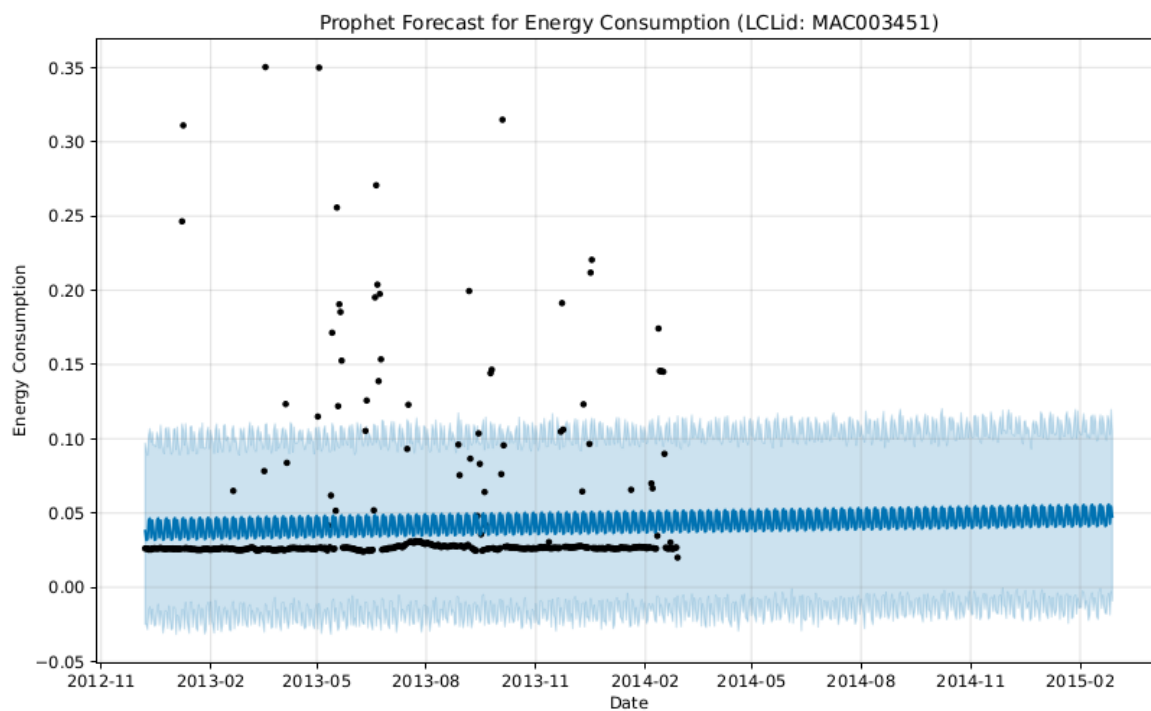
6) MAC000235:

Πίνακας 9 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC000235



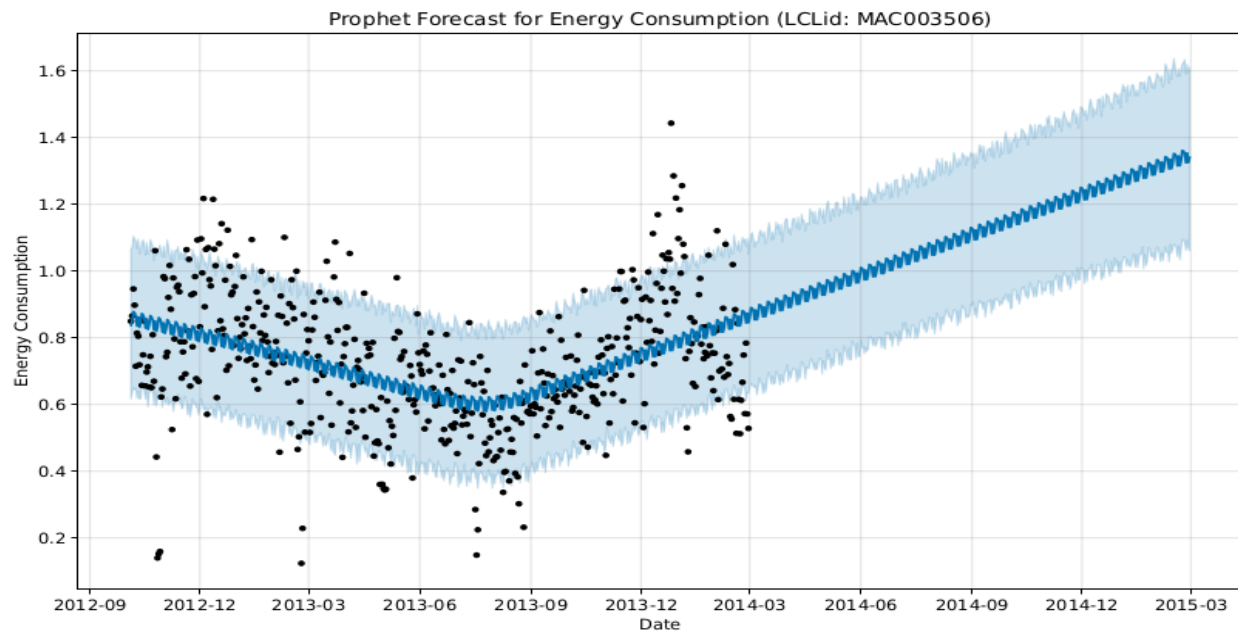
7) MAC003451:

Πίνακας 10 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC003451



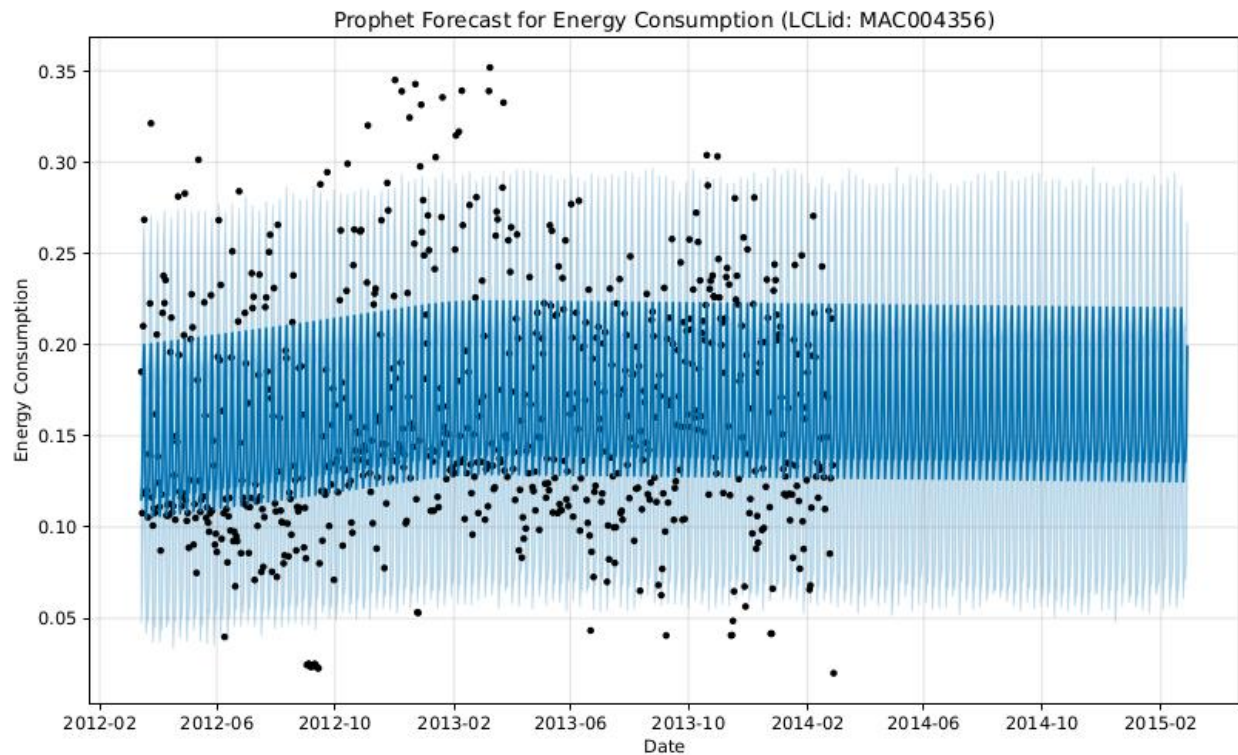
8) MAC003506:

Πίνακας 11 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC003506



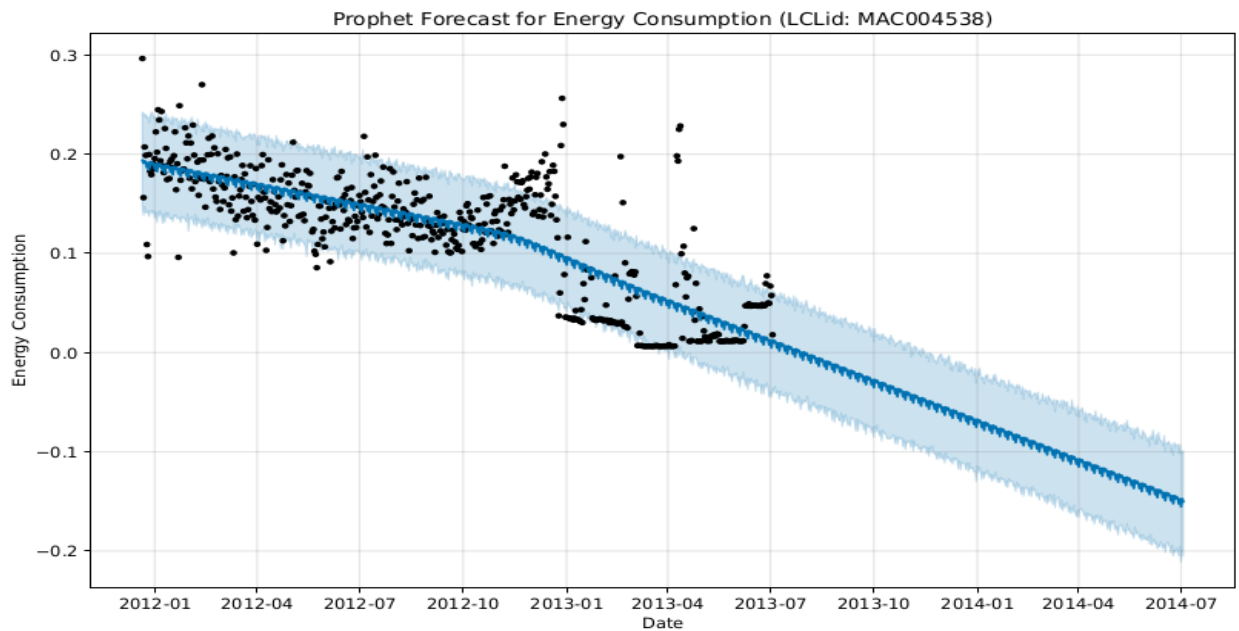
9) MAC004356:

Πίνακας 12 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC004356



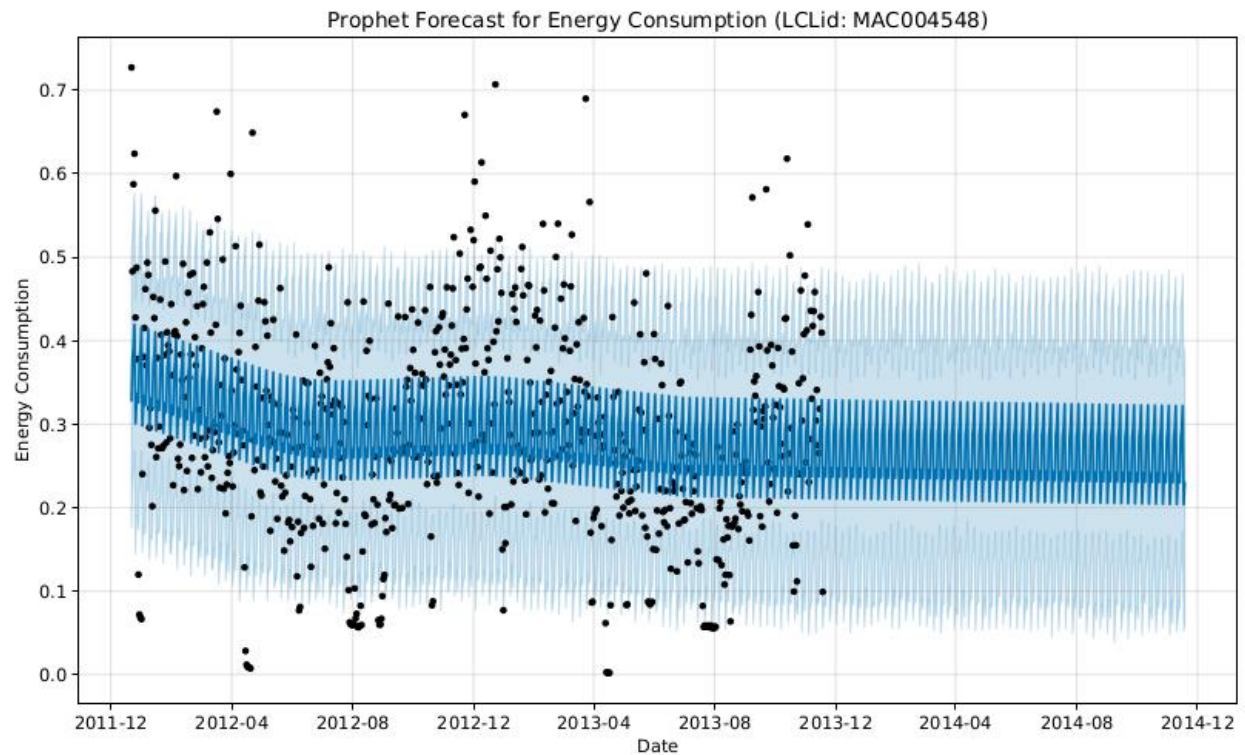
10) MAC004538:

Πίνακας 13 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC004538



11) MAC004548:

Πίνακας 14 - πρόβλεψη κατανάλωσης ενέργειας για LCLid MAC004548



Γ.5' Περιορισμοί της έρευνας

Οι περιορισμοί που μπορεί να προκύψουν σε μια έρευνα είναι ποικίλοι και ιδιαίτεροι πολλές φορές, ανάλογα με την διαδικασία εξαγωγής τους. Στην περίπτωση της επεξεργασίας μεγάλου όγκου δεδομένων και ειδικότερα τούτης της ερευνητικής εργασίας, ο βασικότερος περιορισμός που αντιμετωπίστηκε ήταν η παλαιότητα του μηχανήματός μου αλλά με μία μικρή μετατροπή στον αριθμό των δεδομένων που λαμβάνει ο κώδικας ήταν δυνατή η περάτωση της έρευνας με επιτυχία. Αναλυτικότερα, κάποια από τα προβλήματα που αντιμετωπίστηκαν ή μπορεί να πρέπει να προσπελάσει κάποιος που θα αποφασίσει να κάνει μια παρόμοια έρευνα είναι τα ακόλουθα:

Αρχικά, όσον αφορά την διαθεσιμότητα και την ποιότητα των δεδομένων, οι χρονοσειρές δεδομένων μπορεί να περιέχουν σφάλματα ή κενά, τα οποία μπορεί να περιορίσουν την αποτελεσματικότητα των μοντέλων μηχανικής μάθησης, καθώς η ακρίβεια των αποτελεσμάτων εξαρτάται άμεσα από την ποιότητα των δεδομένων εισόδου.

Επιπλέον, Η ποικιλομορφία των δεδομένων κατανάλωσης ενέργειας, ανάλογα με την περιοχή, την ώρα, ή άλλους παράγοντες, ενδέχεται να επηρεάσει την ανάλυση. Δεδομένα από διαφορετικές γεωγραφικές περιοχές, με διαφορετικά πρότυπα κατανάλωσης, ενδέχεται να παρουσίασαν αποκλίσεις που δεν μπορούσαν να καλυφθούν πλήρως από τα επιλεγμένα μοντέλα.

Επιπρόσθετα, οι αλγόριθμοι που χρησιμοποιήθηκαν για την ανάλυση, όπως το Prophet, αν και είναι ιδανικοί για την πρόβλεψη χρονοσειρών, ενδέχεται να έχουν περιορισμούς ως προς την προσαρμογή σε **ξαφνικές** αλλαγές της κατανάλωσης ενέργειας λόγω απρόβλεπτων γεγονότων, όπως απότομες αλλαγές στον καιρό ή καταναλωτικά πρότυπα που μεταβάλλονται λόγω κοινωνικών ή οικονομικών παραγόντων.

Ειδικότερα για αυτήν την έρευνα, το μεγαλύτερο πρόβλημα που αντιμετωπίστηκε ήταν αυτό της υπολογιστικής δύναμης. Η εφαρμογή αλγορίθμων μηχανικής μάθησης σε μεγάλα σύνολα δεδομένων απαιτεί σημαντικούς υπολογιστικούς πόρους. Ο περιορισμός σε υπολογιστική ισχύ μπορεί να έχει επηρεάσει την ανάλυση, καθώς πιο πολύπλοκοι αλγόριθμοι ή πιο εκτεταμένες προβλέψεις ενδέχεται να μην ήταν εφικτές λόγω τεχνικών περιορισμών.

Τέλος, με αφορμή τα δεδομένα πριν από λίγα χρόνια, πιστεύω είναι σημαντικό να αναφερθούν στα προβλήματα και οι απρόβλεπτες συνθήκες, όπως ξαφνικές αλλαγές στις τιμές ενέργειας ή εξωγενείς

παράγοντες (π.χ. πανδημίες, φυσικές καταστροφές), οι οποίες δεν μπορούν να ληφθούν υπόψη στα μοντέλα πρόβλεψης, με αποτέλεσμα τα αποτελέσματα να επηρεαστούν σε περιπτώσεις όπου τέτοιοι παράγοντες επηρέασαν την κατανάλωση ενέργειας κατά την υπό εξέταση περίοδο.

Δ' Συζήτηση

Δ.1 Ανάλυση κώδικα

Ο πηγαίος κώδικας έχει αναπτυχθεί σε Python3 και έχει σκοπό την σύγκριση σχέσεων θερμοκρασίας και κατανάλωσης ηλεκτρικής ενέργειας αλλά και την προσπάθεια πρόβλεψης της κατανάλωσης αυτής με την χρήση του prophet.

Με γνώμονα τα παραπάνω λοιπόν ο κώδικας είναι χωρισμένος σε τέσσερα κομμάτια για την απλούστερη χρήση και ανάπτυξή του.

Αρχικά, θα ήθελα να ξεκινήσω από τους αλγόριθμους energy.py και temperature.py καθώς δουλεύουν με τον ίδιο σχεδόν τρόπο και στην συνέχεια να επεκτείνω για τους Forecast_per_LCLid.py και forecast_whole.py. Τα βήματα αναγράφονται παρακάτω:

- 1) Εισαγωγή δεδομένων:** Πρώτα απ' όλα, ο αλγόριθμος είναι προγραμματισμένος να φορτώνει τα κατάλληλα δεδομένα προς χρήση από το εκάστοτε CSV αρχείο όπου εμπεριέχονται. Στην προκείμενη περίπτωση, ο energy.py, forecast_whole.py και Forecast_per_LCLid.py από το αρχείο daily_dataset.csv.
Ενώ ο temperature.py από το αρχείο weather_daily_darksky.csv.
- 2) Προ-επεξεργασία δεδομένων:** Αυτό το κομμάτι της έρευνας είναι ίδιο για όλους τους διαφορετικούς αλγόριθμους, αλλά παραμένει σημαντικό να εκτελεστεί σωστά ο καθένας τους. Ο κώδικας εκτελεί πολύ σημαντικά βήματα στην προ-επεξεργασία των δεδομένων με σκοπό την προετοιμασία τους. Αρχικά, επιλέγει τις σχετικές στήλες προς επεξεργασία από το σύνολο των δεδομένων. Ακόμα διαχειρίζεται τις τιμές που δεν ενδείκνυνται για χρήση με τους συγκεκριμένους αλγόριθμους, αγνοώντας τις τελείως αν δεν πληρούν τα κριτήρια για την ορθή λειτουργία του προγράμματος.
- 3) Εξαγωγή αποτελεσμάτων:** Με την ολοκλήρωση του προηγούμενου βήματος, ο forecast_whole.py και ο Forecast_per_LCLid.py, ξεκινούν τις διαδικασίες για την

παραγωγή σχημάτων, αλλά πριν από αυτό, ορίζουν ο καθένας τον αριθμό των στοιχείων που θα εξεταστούν, αυτός ο αριθμός για λόγους προσαρμοστικότητας, είναι εύκολα μεταβλητός από οποιονδήποτε με πρόσβαση στον πηγαίο κώδικα. Στην συνέχεια τα στοιχεία που είναι κατάλληλα για επεξεργασία φορτώνονται σε ένα δεύτερο πλαίσιο δεδομένων για την ευκολότερη διαχείρισή τους καθώς το επόμενο βήμα είναι η εφαρμογή του prophet πριν την αποθήκευση των παραγόμενων γράφων. Τέλος, ακολουθεί η αποθήκευση των αποτελεσμάτων σε PDF αρχείο.

Αντίστοιχα οι energy.py και temperature.py εκτελούν τις πράξεις για την παραγωγή μέσου όρου κατανάλωσης ρεύματος και μέσου όρου Μέγιστης θερμοκρασίας ανά ημέρα, επίσης εκτυπώνοντας τα αποτελέσματα σε αντίστοιχα pdf αρχεία.

Δ.1.1 Pandas Library

Το Pandas Library αποτελεί μια ανοικτού κώδικα βιβλιοθήκη για την διαχείριση και την ανάλυση δεδομένων στη γλώσσα προγραμματισμού Python. Παρέχει εύχρηστες δομές δεδομένων και διάφορες λειτουργίες για την επεξεργασία μεγάλου όγκου δεδομένων, όπως πίνακες ή υπολογιστικά φύλλα. Η βιβλιοθήκη Pandas χρησιμοποιείται ευρέως στην επιστήμη δεδομένων, την ανάλυση δεδομένων και τη μηχανική μάθηση. Η βιβλιοθήκη χρησιμεύει για εργασίες όπως ο καθαρισμός, η εξερεύνηση και η προετοιμασία δεδομένων. Οι δύο κύριες δομές της βιβλιοθήκης Pandas είναι οι εξής:

- 1) **Πλαίσια δεδομένων:** πρόκειται για δισδιάστατους πίνακες με γραμμές και στήλες, οι οποίοι είναι δομημένοι ως εξής: Κάθε στήλη μπορεί να περιέχει πολλές στήλες, παρόμοια με ένα λογιστικό φύλλο ή έναν πίνακα SQL.
Μπορεί να διαχειριστεί διαφορετικούς τύπους δεδομένων, όπως αριθμούς, συμβολοσειρές, ημερομηνίες κ.λπ.
- 2) **Σειρές:** μονοδιάστατος πίνακας με ετικέτες και στήλες από ένα πλαίσιο δεδομένων.

- 3) **Οπτικοποίηση δεδομένων:** Οπτικοποίηση δεδομένων. Η βιβλιοθήκη Pandas μπορεί να ενσωματωθεί με πολλές βιβλιοθήκες οπτικοποίησης δεδομένων. Το γεγονός αυτό έχει την δυνατότητα απρόσκοπτης ενσωμάτωσης με κοινές βιβλιοθήκες απλοποιεί την εξερεύνηση και την ανάλυση. Επιπλέον, οι εξελιγμένες εφαρμογές της βιβλιοθήκης επιτρέπουν την εύκολη εξερεύνηση και την παρουσίαση δεδομένων.

Συμπερασματικά, η βιβλιοθήκη Pandas είναι ένα ευέλικτο εργαλείο για την ανάλυση και τον χειρισμό δεδομένων. Είναι ιδιαίτερα κατάλληλη για τον καθαρισμό, τη διερεύνηση και την ενσωμάτωση δεδομένων, την προετοιμασία δεδομένων για μηχανική μάθηση και άλλες εφαρμογές όπως η επεξεργασία και η οπτικοποίηση δεδομένων.

Δ.2 Prophet Library

Δ.2.1 Εισαγωγή στη Βιβλιοθήκη Prophet

Η βιβλιοθήκη **Prophet**, που αναπτύχθηκε από το Facebook, είναι ένα ισχυρό εργαλείο πρόβλεψης χρονοσειρών το οποίο σχεδιάστηκε για δεδομένα που παρουσιάζουν εποχιακές διακυμάνσεις και ανωμαλίες. Η βιβλιοθήκη αυτή είναι ειδικά προσαρμοσμένη για τη δημιουργία προβλέψεων που βασίζονται σε χρονοσειρές, επιτρέποντας στους χρήστες να δημιουργούν μοντέλα που προσαρμόζονται στις διακυμάνσεις των δεδομένων με υψηλή ακρίβεια.

Το Prophet είναι ιδανικό για την ανάλυση δεδομένων που επηρεάζονται από καθημερινές, εβδομαδιαίες ή ετήσιες τάσεις, όπως τα δεδομένα κατανάλωσης ενέργειας. Έχει σχεδιαστεί με στόχο να παρέχει προβλέψεις που είναι εύκολο να εφαρμοστούν και να τροποποιηθούν, ακόμα και από χρήστες με περιορισμένη εμπειρία στη μοντελοποίηση χρονοσειρών (Kang, C., et al., 2015).

Χαρακτηριστικά της Βιβλιοθήκης Prophet:

- **Απλότητα χρήσης:** Το Prophet παρέχει εύχρηστα εργαλεία για την εισαγωγή δεδομένων χρονοσειρών και την άμεση δημιουργία προβλέψεων.

- **Αυτόματη ανίχνευση εποχιακών τάσεων:** Η βιβλιοθήκη μπορεί να εντοπίσει αυτόματα εποχιακές διακυμάνσεις, επιτρέποντας την εύκολη ανάλυση δεδομένων με έντονες εποχιακές διακυμάνσεις, όπως η κατανάλωση ενέργειας το καλοκαίρι και τον χειμώνα.
- **Ευελιξία στην προσαρμογή παραμέτρων:** Ο χρήστης μπορεί να προσαρμόσει τις παραμέτρους για να βελτιώσει την ακρίβεια της πρόβλεψης και να τροποποιήσει το μοντέλο με βάση ειδικές απαιτήσεις.

Δ.2.2 Πλεονεκτήματα της Χρήσης της Prophet στη Διαχείριση Ενέργειας

Η βιβλιοθήκη Prophet χρησιμοποιείται εκτενώς σε εφαρμογές όπου τα δεδομένα είναι ευαίσθητα σε εποχιακές τάσεις και διακυμάνσεις. Στην παρούσα πτυχιακή εργασία, η βιβλιοθήκη Prophet εφαρμόστηκε για την πρόβλεψη της κατανάλωσης ενέργειας, όπου οι τάσεις στην κατανάλωση ρεύματος επηρεάζονται από διάφορους παράγοντες όπως οι καιρικές συνθήκες, οι καθημερινές διακυμάνσεις και οι εποχιακές αλλαγές.

Κύρια Πλεονεκτήματα:

1. **Αντιμετώπιση εποχιακών και καθημερινών τάσεων:**
 - ο Στην κατανάλωση ενέργειας, παρατηρούνται συχνά έντονες διαφοροποιήσεις ανάλογα με την εποχή (π.χ., αυξημένη χρήση ρεύματος το καλοκαίρι λόγω κλιματιστικών). Η βιβλιοθήκη Prophet μπορεί να ανιχνεύσει και να μοντελοποιήσει αυτές τις τάσεις, επιτρέποντας ακριβέστερες προβλέψεις.
2. **Ανθεκτικότητα σε ελλιπή ή ακανόνιστα δεδομένα:**
 - ο Το Prophet είναι ανθεκτικό σε ελλιπή ή μη κανονικά κατανεμημένα δεδομένα, κάτι που το καθιστά κατάλληλο για τη διαχείριση δεδομένων κατανάλωσης ενέργειας, όπου μπορεί να υπάρχουν κενά ή ακραίες τιμές.
3. **Γρήγορη δημιουργία προγνωστικών μοντέλων:**
 - ο Σε αντίθεση με πιο σύνθετα μοντέλα πρόβλεψης, το Prophet επιτρέπει την ταχεία δημιουργία μοντέλων, καθιστώντας την ιδανική για την ταχύτατη ανάλυση

μεγάλων συνόλων δεδομένων, όπως αυτά που χρησιμοποιούνται για την ανάλυση της κατανάλωσης ενέργειας.

4. Ανάλυση των ανωμαλιών:

- ο Το Prophet μπορεί να ανιχνεύσει ανωμαλίες (outliers) στα δεδομένα και να τις λάβει υπόψη χωρίς να επηρεαστεί η ακρίβεια του μοντέλου. Αυτό είναι ιδιαίτερα χρήσιμο για δεδομένα κατανάλωσης ενέργειας, που μπορεί να παρουσιάζουν αιχμές λόγω ειδικών γεγονότων.

Δ.2.3 Εφαρμογή του Prophet στην Πτυχιακή Εργασία

Στην παρούσα εργασία, η βιβλιοθήκη Prophet χρησιμοποιήθηκε για την πρόβλεψη της κατανάλωσης ενέργειας με βάση δεδομένα που συλλέχθηκαν από έξυπνους μετρητές. Τα δεδομένα αυτά αναλύθηκαν για να εντοπιστούν καθημερινές και εποχιακές τάσεις, καθώς και οι επιπτώσεις εξωτερικών παραγόντων, όπως οι αλλαγές στις καιρικές συνθήκες. Ακολουθούν τα βασικά βήματα που εφαρμόστηκαν:

Προετοιμασία των δεδομένων:

- Τα δεδομένα χρονοσειρών εισήχθησαν στο Prophet, αφού πρώτα προεπεξεργάστηκαν μέσω της βιβλιοθήκης Pandas. Τα δεδομένα περιλάμβαναν ημερήσια κατανάλωση ενέργειας, με καταγραφή για αρκετούς μήνες.

Δημιουργία προβλέψεων:

- Το μοντέλο χρησιμοποιήθηκε για την πρόβλεψη της κατανάλωσης ενέργειας για τους επόμενους μήνες. Οι προβλέψεις έδειξαν πως η κατανάλωση αυξάνεται κατά τους καλοκαιρινούς μήνες, κάτι που συμφωνεί με τις προηγούμενες παρατηρήσεις για τις διακυμάνσεις στη χρήση ενέργειας.

Οπτικοποίηση αποτελεσμάτων:

- Τα αποτελέσματα παρουσιάστηκαν μέσω γραφημάτων, με τη χρήση της βιβλιοθήκης Matplotlib και Seaborn.

Δ.2.4 Περιορισμοί της Χρήσης της Prophet

Αν και η βιβλιοθήκη Prophet προσφέρει σημαντικά πλεονεκτήματα, υπάρχουν ορισμένοι περιορισμοί που πρέπει να ληφθούν υπόψη:

1. Περιορισμοί στην προσαρμογή σε απότομες μεταβολές:

- Το Prophet βασίζεται σε ιστορικά δεδομένα για να ανιχνεύσει τάσεις, γεγονός που το καθιστά λιγότερο αποτελεσματικό σε περιπτώσεις ξαφνικών μεταβολών ή απρόβλεπτων γεγονότων που δεν έχουν καταγραφεί στην ιστορία των δεδομένων.

2. Απαιτεί μεγάλη ποσότητα δεδομένων:

- Η βιβλιοθήκη Prophet αποδίδει καλύτερα όταν υπάρχουν πολλά δεδομένα για την εκπαίδευση του μοντέλου. Σε περιπτώσεις όπου τα δεδομένα είναι περιορισμένα ή λείπουν σημαντικές καταγραφές, η ακρίβεια των προβλέψεων μπορεί να μειωθεί.

Δ.2.5 Συμπεράσματα για τη Βιβλιοθήκη Prophet

Η βιβλιοθήκη Prophet αποδεικνύεται ένα ισχυρό εργαλείο για την πρόβλεψη δεδομένων κατανάλωσης ενέργειας, με τη δυνατότητα να αντιμετωπίσει εποχιακές διακυμάνσεις και ανωμαλίες στα δεδομένα. Στην παρούσα εργασία, το Prophet βοήθησε στη δημιουργία ακριβών προβλέψεων για τη μελλοντική κατανάλωση ενέργειας, παρέχοντας χρήσιμα δεδομένα που μπορούν να χρησιμοποιηθούν από διαχειριστές δικτύων και παρόχους ενέργειας για τη βελτίωση της ενεργειακής διαχείρισης.

Δ.3 Ερμηνεία των αποτελεσμάτων

Σε αυτό το κεφάλαιο της εργασίας αναπτύσσονται και εξηγούνται λεπτομερώς τα βασικά ευρήματα της έρευνας, προσπαθώντας να ερμηνεύσουμε τη σημασία των συλλεγμένων δεδομένων. Επίσης, είναι σημαντικό να αναγνωρίσουμε τις περιορισμένες πτυχές της μελέτης, καθώς αυτό συμβάλλει στην κατανόηση των περιορισμών που υπάρχουν στα ευρήματά μας. Η ανάλυση των αποτελεσμάτων μιας έρευνας αποτελεί κρίσιμη διαδικασία για την πλήρη κατανόηση των ευρημάτων και την αποτελεσματική τους εφαρμογή στην πράξη. Σε αυτή την

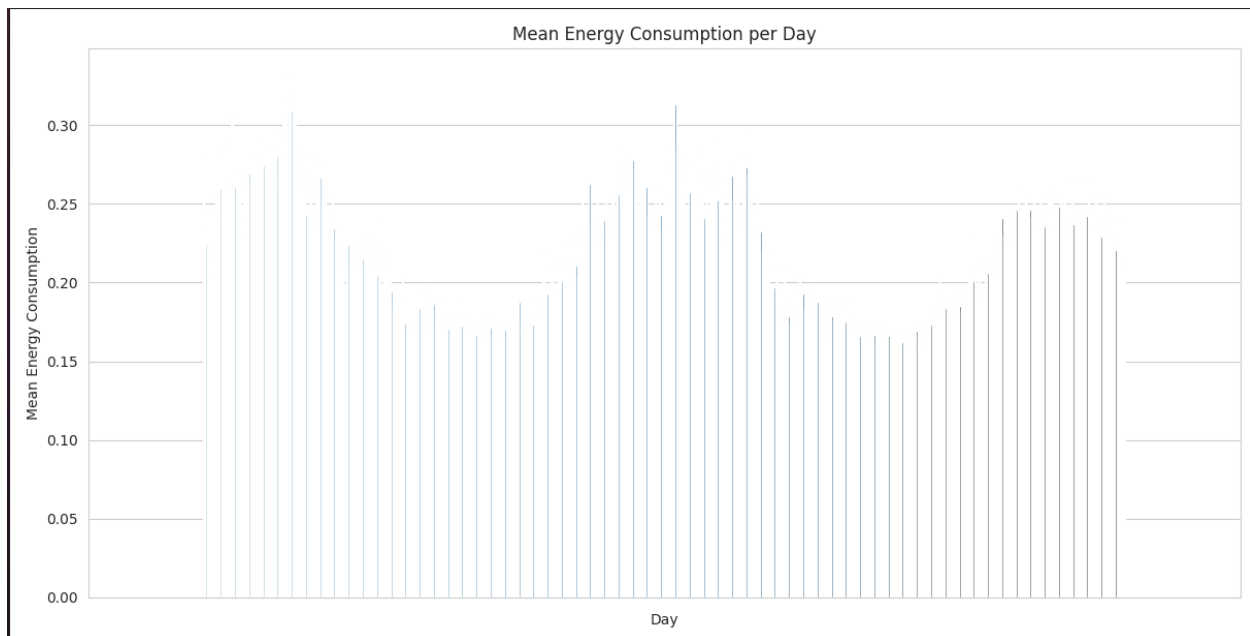
μελέτη, εξετάζεται η σχέση μεταξύ του καιρού και της κατανάλωσης ηλεκτρικής ενέργειας, δίνοντας μας μία εικόνα για τις περιόδους υψηλής κατανάλωσης στην κοινωνία του Λονδίνου. Οι αλγόριθμοι που αναπτύχθηκαν έχουν σκοπό την ανακάλυψη τέτοιων σχέσεων αλλά και την πρόβλεψη κατανάλωσης ρεύματος την ακόλουθη χρονική περίοδο, μετά το πέρας των δεδομένων.

Δ.3.1 Σχέση κατανάλωσης ενέργειας και θερμοκρασίας

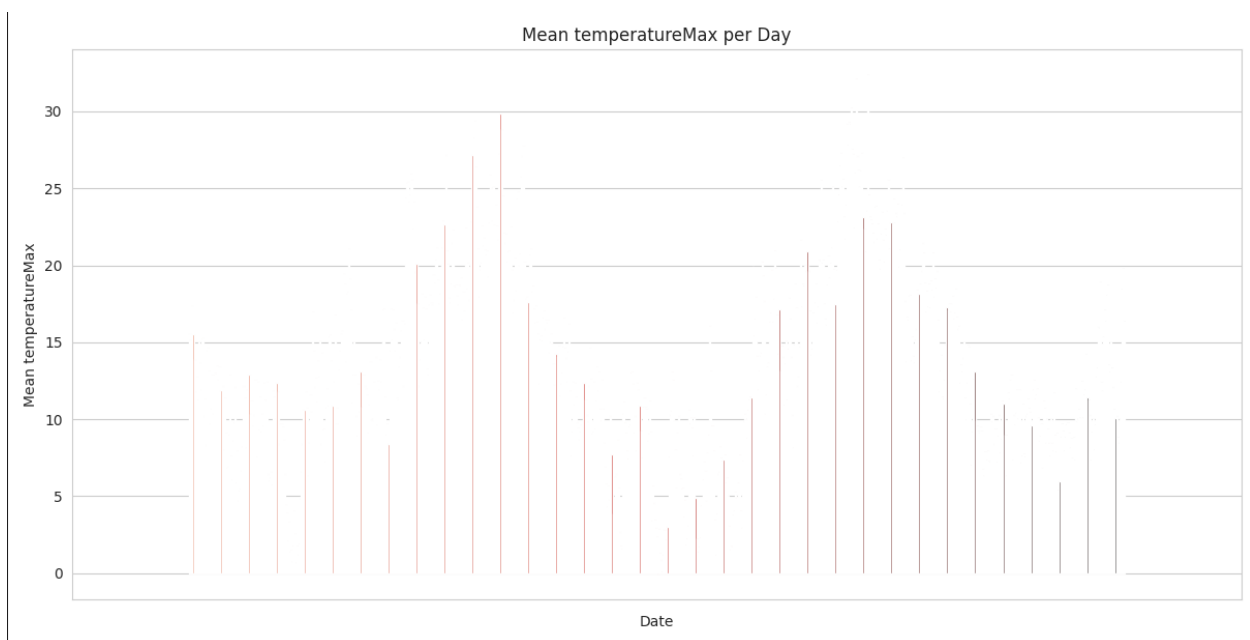
Μετά την εκτέλεση και των δύο ομώνυμων με το κεφάλαιο αλγορίθμων, πιο συγκεκριμένα, των `energy.py` και `temperature.py` το τελικό αποτέλεσμα που μας δίνουν ως έξοδο είναι άμεσα συγκρίσιμο, ώστε να εξάγουμε αποτελέσματα με γνώμονα αυτό.

Πιο συγκεκριμένα, τα προγράμματα παράγουν τους παρακάτω γράφους:

Πίνακας 15 - Συνολική μέση κατανάλωση ενέργειας ανά ημέρα



Πίνακας 16 – Μέση μέγιστη θερμοκρασία ανά ημέρα



Από τους ανωτέρω γράφους, μπορούμε να εξαγάγουμε το συμπέρασμα ότι σε αντίστοιχες χρονικές περιόδους είναι *αντιστρόφως ανάλογοι*, δηλαδή, όπου ο γράφος της μέσης μέγιστης θερμοκρασίας έχει υψηλές τιμές, αυτές του γράφου μέσης ημερήσιας ηλεκτρικής κατανάλωσης μειώνονται με παρόμοιο τρόπο. Το γεγονός ότι αυτό το φαινόμενο επικρατεί σε ολόκληρη την

χρονική έκταση των δεδομένων, μας αποδεικνύει ότι υπάρχει όντως αντιστρόφως ανάλογη σχέση μεταξύ των δύο πλαισίων δεδομένων.

Δ.3.2 Προβλέψεις κατανάλωσης ενέργειας

Το Prophet είναι μια ισχυρή βιβλιοθήκη πρόβλεψης χρονοσειρών που αναπτύχθηκε από την ομάδα του Facebook Research. Χρησιμοποιώντας αυτήν τη βιβλιοθήκη, κατάφερα να δημιουργήσω μια πρόβλεψη για την κατανάλωση ρεύματος για έναν χρόνο, βασιζόμενος στα δεδομένα του dataset μου. Η πρόβλεψη βασίστηκε σε δεδομένα κατανάλωσης ρεύματος που συλλέχθηκαν από έξυπνους μετρητές στο Λονδίνο. Το Prophet επιλέχθηκε για την ακρίβειά του στην πρόβλεψη χρονοσειρών που επηρεάζονται από περιοδικές τάσεις.

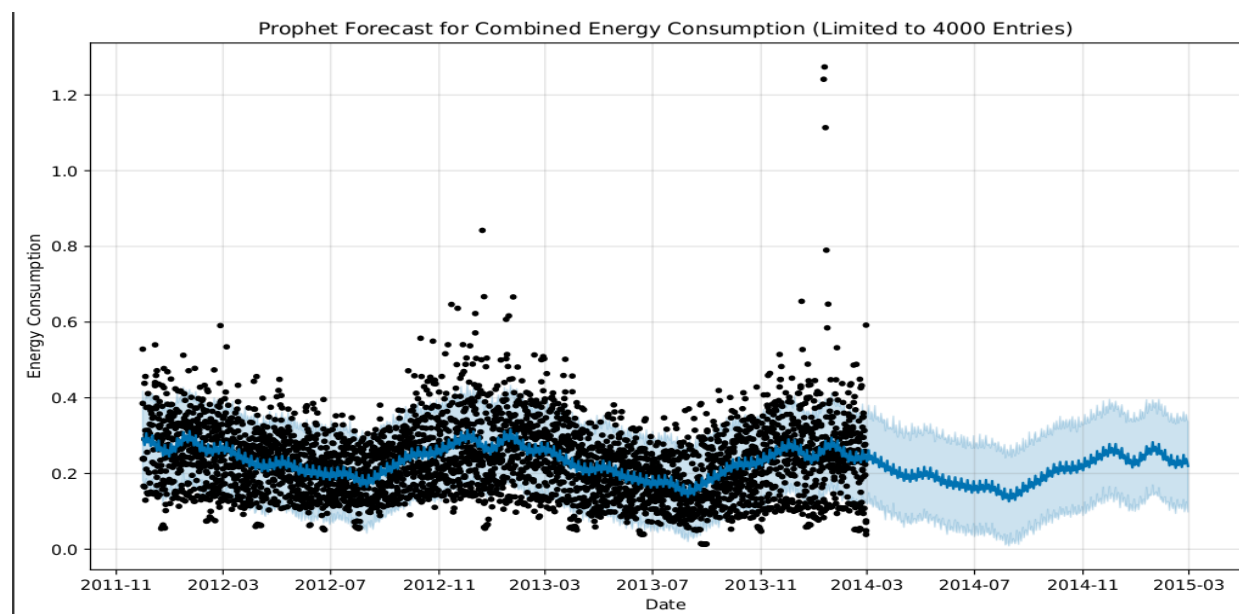
Αρχικά, τα δεδομένα κατανάλωσης εισήχθησαν και προεπεξεργάστηκαν με τη χρήση της βιβλιοθήκης Pandas για την προσαρμογή τους στο απαιτούμενο σχήμα χρονοσειράς. Η βιβλιοθήκη Prophet απαιτεί καθαρά και δομημένα δεδομένα, γι' αυτό και προηγήθηκε μια διαδικασία διαμόρφωσης των ημερομηνιών και τιμών. Μετά την επεξεργασία, τα δεδομένα εισήχθησαν στο Prophet για να προβλεφθεί η κατανάλωση για το επόμενο έτος μετά την τελευταία διαθέσιμη καταγραφή. Το μοντέλο δημιούργησε προβλέψεις τόσο για την κατανάλωση ανά οικοδομικό τετράγωνο όσο και για το σύνολο της πόλης. Η ανάλυση περιλαμβάνει τη δημιουργία γραφημάτων που οπτικοποιούν την αναμενόμενη κατανάλωση ενέργειας.

Τέλος, η οπτικοποίηση των προβλέψεων έγινε με τη βοήθεια των βιβλιοθηκών Matplotlib και Seaborn, όπου δημιουργήθηκαν γραφήματα για την αναπαράσταση των προβλέψεων. Τα γραφήματα αυτά έδειξαν αύξηση της κατανάλωσης ενέργειας κατά τους πιο ψυχρούς μήνες, γεγονός που συνάδει με τις παρατηρήσεις για τις εποχιακές ανάγκες σε ενέργεια. Αυτά τα γραφήματα μπορούν να βοηθήσουν στην κατανόηση των αναμενόμενων τάσεων και τα μοτίβα

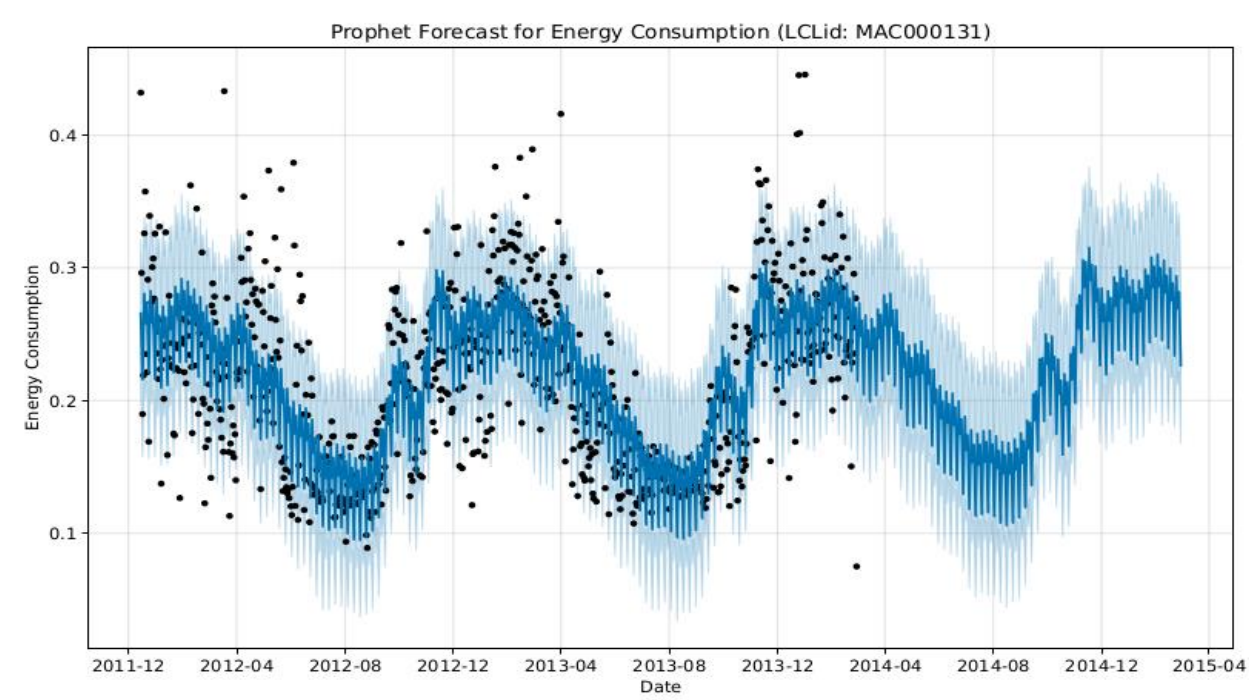
στην κατανάλωση ενέργειας για τον επόμενο χρόνο, πράγμα που είναι χρήσιμο για προγραμματισμό και λήψη αποφάσεων σχετικά με την ενέργεια.

Ακολουθούν κάποια από τα αποτελέσματα παραδειγματικά:

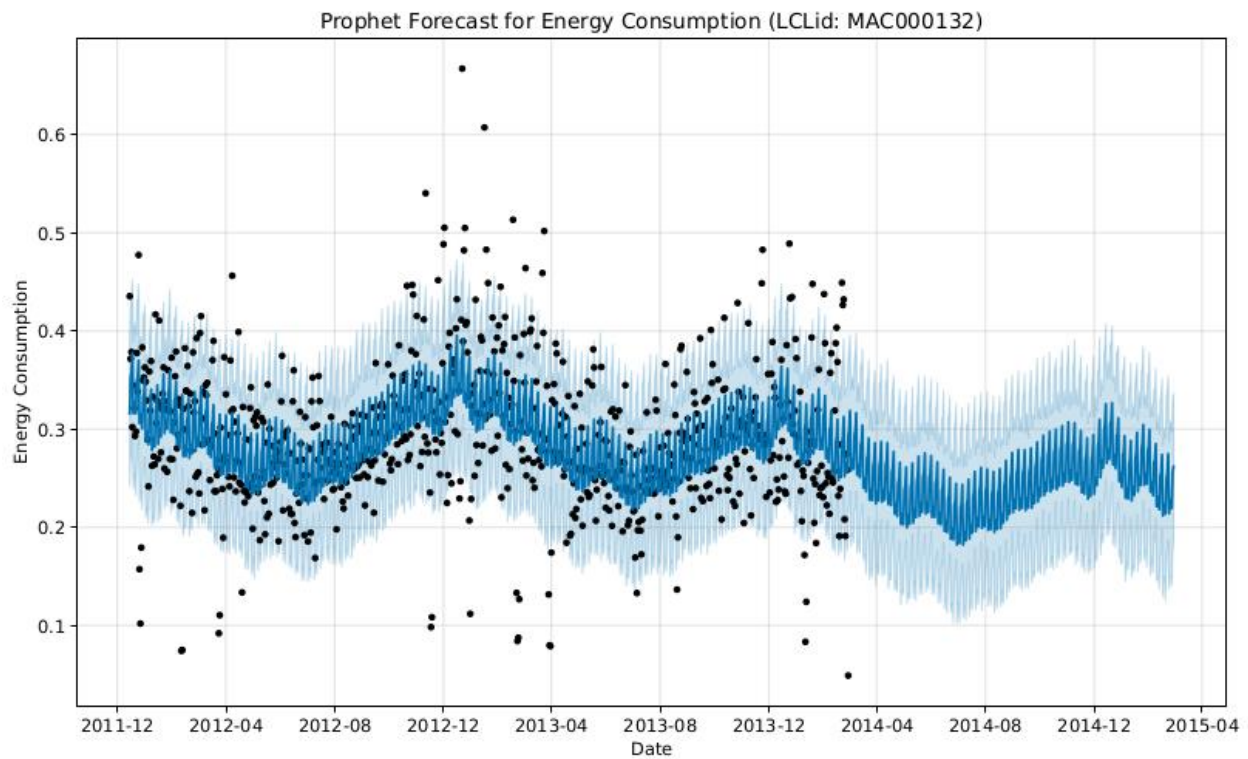
Πίνακας 17 – παραλληλισμός και πρόβλεψη συνδυαστικής κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία



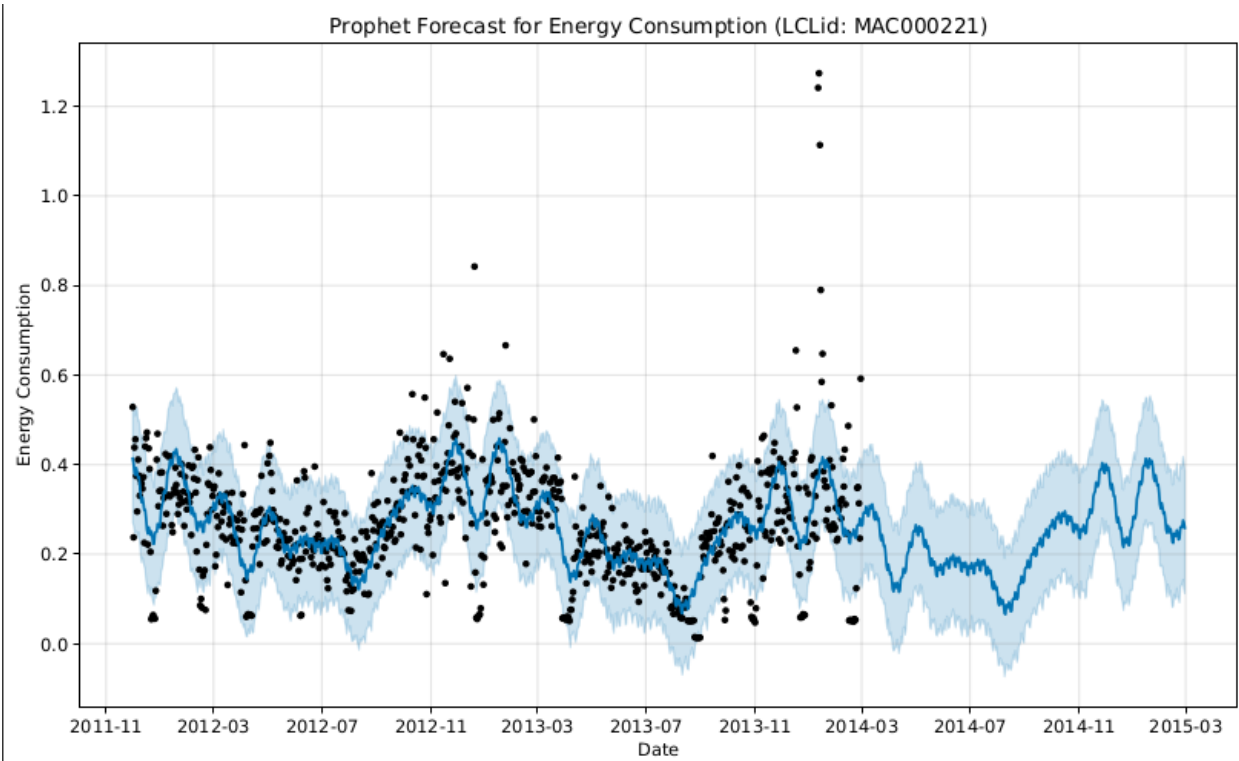
Πίνακας 18 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC000131



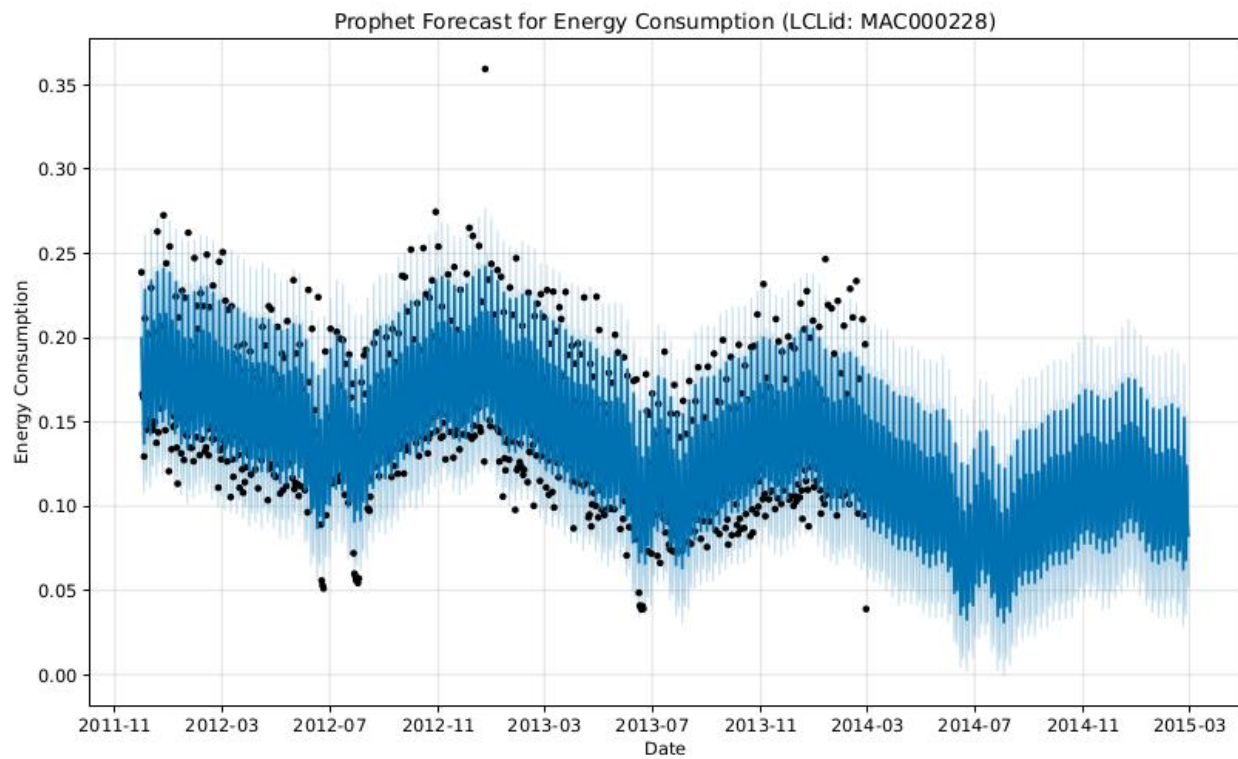
Πίνακας 19 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC000132



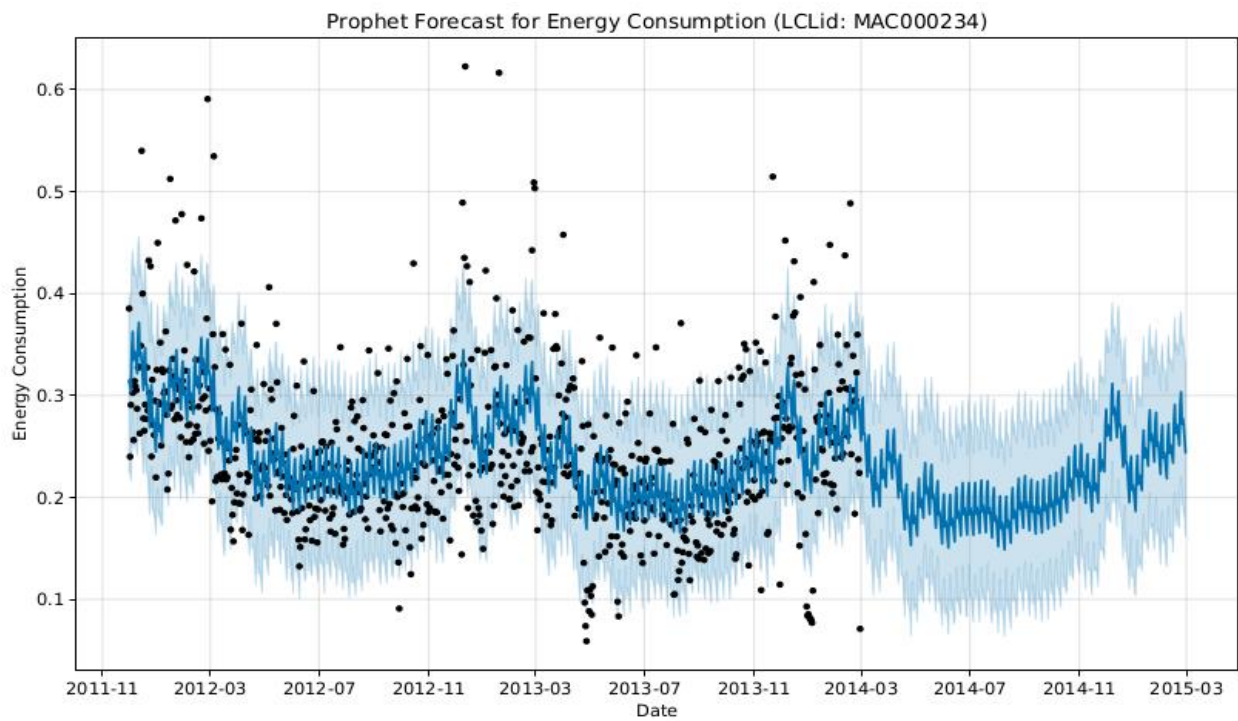
Πίνακας 20 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC000221



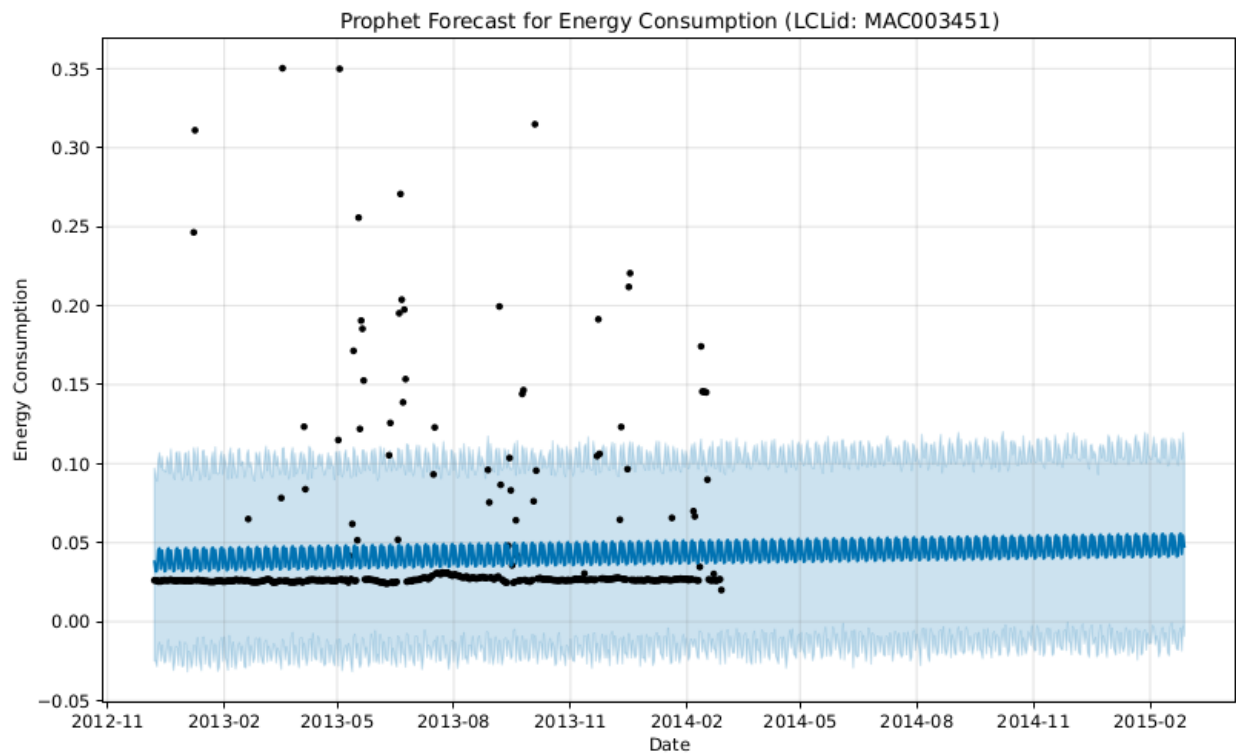
Πίνακας 21 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC000228



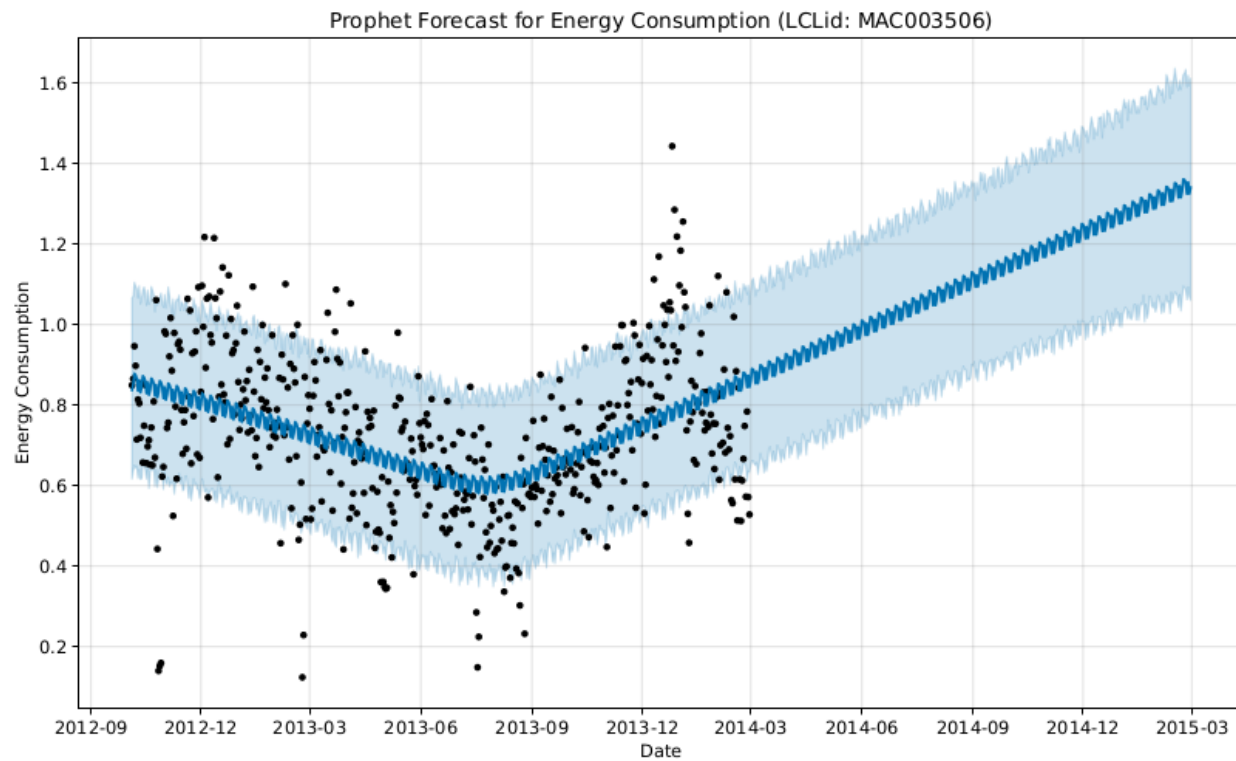
Πίνακας 22 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC000234



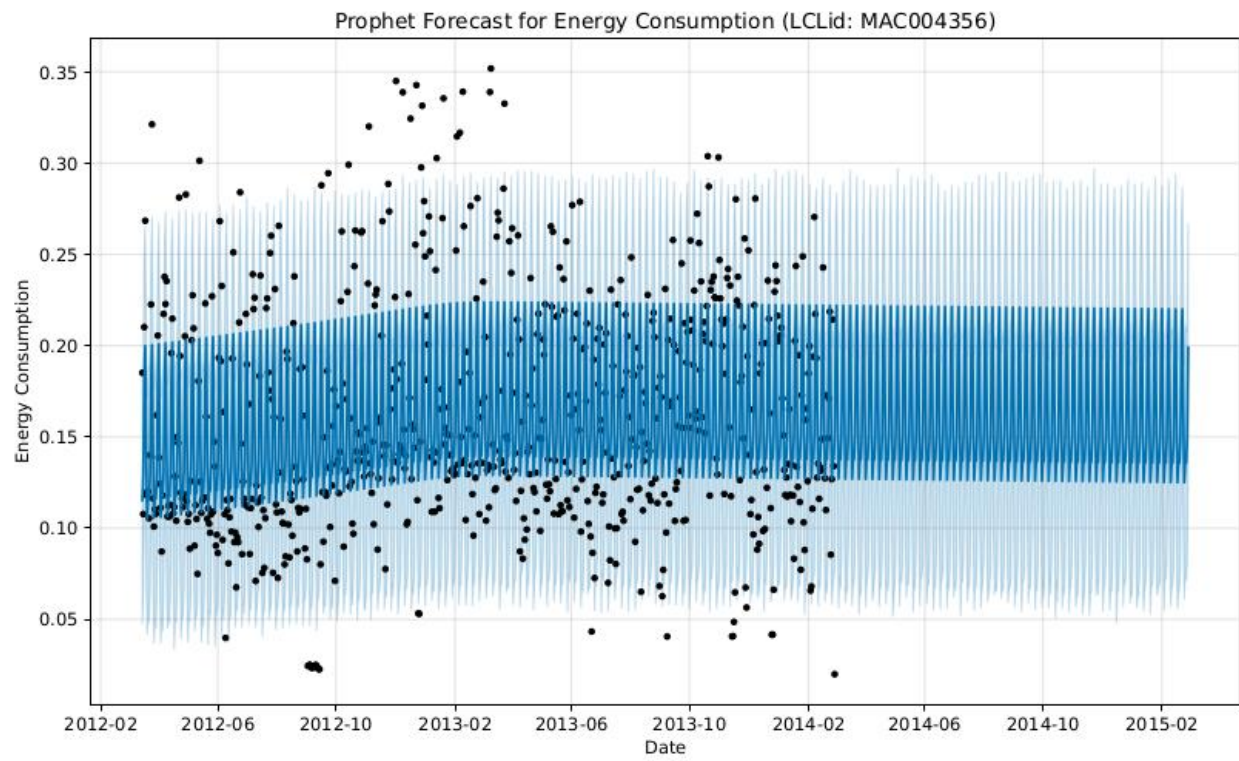
Πίνακας 23 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC003451



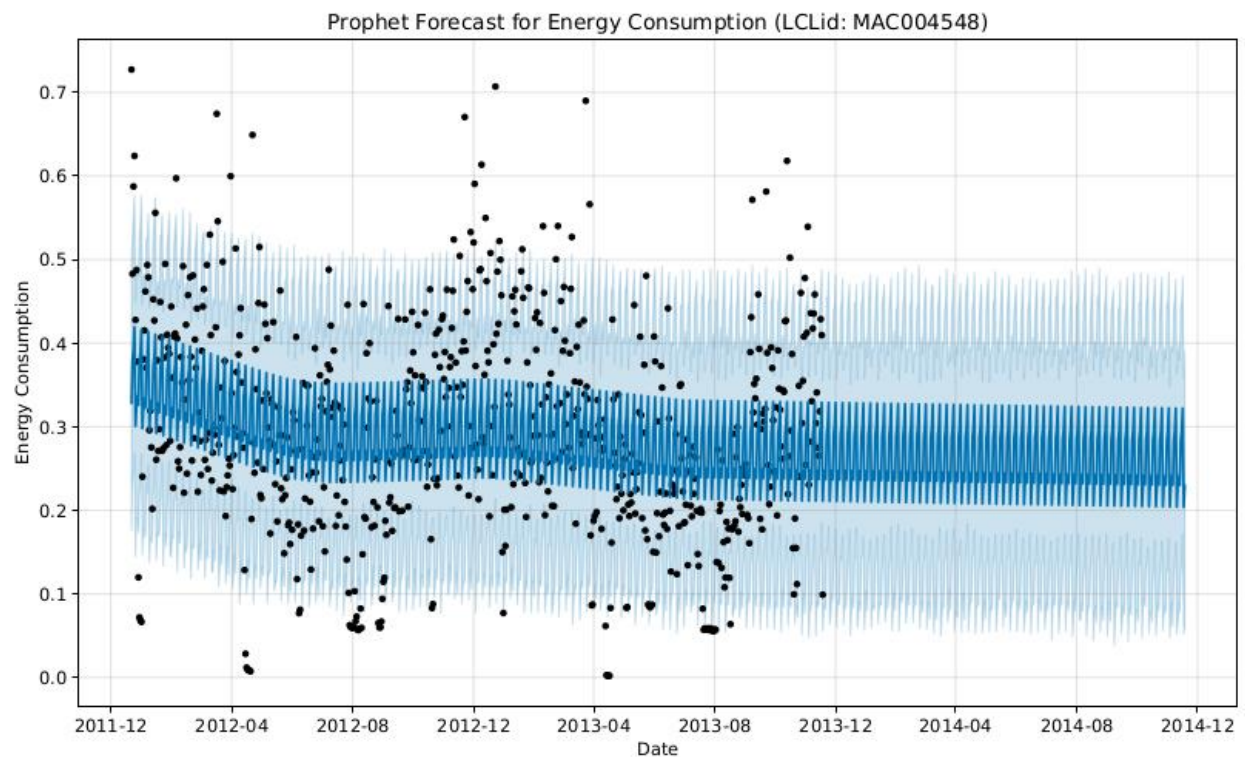
Πίνακας 24 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC003506



Πίνακας 25 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC004356



Πίνακας 26 - παραλληλισμός και πρόβλεψη κατανάλωσης ενέργειας σε σχέση με την θερμοκρασία LCLid MAC004548



Όπως μπορούμε να παρατηρήσουμε και παραπάνω οι εισαγωγές που μπορεί να λάβει ο αλγόριθμος Prophet όσον αφορά την πρόβλεψη του συνόλου είναι 4000 λόγω μη αρκετής υπολογιστικής ισχύς, όμως τα γραφήματα είναι αντιπροσωπευτικά των δυνατοτήτων που μπορεί να προσφέρει ο αλγόριθμος.

Δ.4 Μελλοντικές Προοπτικές

Δ.4.1 Ενίσχυση της Μηχανικής Μάθησης στην Ενεργειακή Διαχείριση

Η χρήση μηχανικής μάθησης στον τομέα της ενέργειας αναμένεται να ενισχυθεί περαιτέρω τα επόμενα χρόνια, με νέα μοντέλα και αλγόριθμους να εφαρμόζονται σε ολοένα και μεγαλύτερα σύνολα δεδομένων. Οι μελλοντικές προοπτικές περιλαμβάνουν την εφαρμογή πιο σύνθετων αλγορίθμων, όπως τα Νευρωνικά Δίκτυα (Neural Networks) και τους Γενετικούς Αλγόριθμους (Genetic Algorithms), για την επίλυση περίπλοκων προβλημάτων που σχετίζονται με την ενεργειακή κατανάλωση, τη διαχείριση των πόρων και την πρόβλεψη των αναγκών.

Πιθανές Εξελίξεις:

- **Βελτίωση της πρόβλεψης με πιο εξελιγμένα μοντέλα:** Τα μελλοντικά μοντέλα μηχανικής μάθησης θα ενσωματώνουν περισσότερες μεταβλητές, λαμβάνοντας υπόψη εξωτερικούς παράγοντες, όπως κλιματικές αλλαγές, κοινωνικές συνθήκες και τεχνολογικές εξελίξεις, για την ακριβέστερη πρόβλεψη της κατανάλωσης ενέργειας.
- **Αυτοματοποιημένα συστήματα μάθησης:** Οι αλγόριθμοι μηχανικής μάθησης αναμένεται να ενσωματώσουν πιο αυτοματοποιημένες λειτουργίες μάθησης, επιτρέποντας στα ενεργειακά συστήματα να προσαρμόζονται αυτόματα σε μεταβαλλόμενες συνθήκες χωρίς ανθρώπινη παρέμβαση (Bishop, 2006).

Δ.4.2 Εξέλιξη των Έξυπνων Δικτύων (Smart Grids)

Τα έξυπνα δίκτυα ενέργειας αποτελούν ήδη μια βασική τεχνολογία για τη βελτιστοποίηση της διαχείρισης της ενέργειας. Στο μέλλον, τα smart grids αναμένεται να ενσωματώσουν ακόμα πιο εξελιγμένες τεχνολογίες, όπως το **Internet of Things (IoT)** και το **Artificial Intelligence (AI)**, με στόχο τη συνεχή παρακολούθηση και ανάλυση της ενεργειακής κατανάλωσης σε πραγματικό χρόνο.

Πιθανές Εξελίξεις:

- **Προσαρμοστική διαχείριση σε πραγματικό χρόνο:** Τα έξυπνα δίκτυα θα χρησιμοποιούν τεχνητή νοημοσύνη για την άμεση προσαρμογή της παροχής ενέργειας ανάλογα με τη ζήτηση, μειώνοντας τις απώλειες και αυξάνοντας την απόδοση. Αυτό θα επιτρέψει στους διαχειριστές να αντιμετωπίζουν τις αλλαγές στη ζήτηση πιο αποτελεσματικά, αποφεύγοντας την υπερφόρτωση του δικτύου.
- **Ενσωμάτωση κατανεμημένων πηγών ενέργειας (DER):** Στο μέλλον, οι ανανεώσιμες πηγές ενέργειας, όπως τα ηλιακά πάνελ και οι ανεμογεννήτριες, θα συνδεθούν πιο στενά με τα έξυπνα δίκτυα μέσω συστημάτων IoT. Αυτό θα επιτρέψει την κατανομή της ενέργειας σε περιόδους αιχμής ή την αποθήκευση ενέργειας όταν υπάρχει υπερπαραγωγή (Mengelkamp, 2018).

Δ.4.3 Τεχνητή Νοημοσύνη (AI) και Βελτιστοποίηση Ενεργειακών Συστημάτων

Η τεχνητή νοημοσύνη (AI) αναμένεται να φέρει επανάσταση στη διαχείριση ενέργειας, επιτρέποντας την ανάλυση τεράστιων συνόλων δεδομένων και τη δημιουργία προγνωστικών μοντέλων που θα μπορούν να προσαρμόζονται σε πραγματικό χρόνο. Η ενισχυτική μάθηση (reinforcement learning) θα εφαρμοστεί για την αυτόματη λήψη αποφάσεων σε έξυπνα δίκτυα, επιτρέποντας την αποδοτικότερη διαχείριση πόρων και την αυτοματοποίηση της διανομής ενέργειας.

Πιθανές Εξελίξεις:

- **Αυτοματοποίηση της διαχείρισης ενέργειας:** Τα συστήματα AI θα μπορούν να διαχειρίζονται τη διανομή και την αποθήκευση ενέργειας αυτόνομα, λαμβάνοντας υπόψη τις συνθήκες της ζήτησης και της προσφοράς, αυξάνοντας την αποδοτικότητα των δικτύων.
- **Βελτιστοποίηση των ανανεώσιμων πηγών ενέργειας:** Τα συστήματα AI θα χρησιμοποιούνται για τη βελτίωση της πρόβλεψης της παραγωγής ενέργειας από ανανεώσιμες πηγές, λαμβάνοντας υπόψη τις καιρικές συνθήκες και άλλους παράγοντες, επιτρέποντας στους διαχειριστές να ενσωματώσουν καλύτερα την ανανεώσιμη ενέργεια στα δίκτυα (Russell & Norvig, 2010).

Δ.4.4 Εφαρμογές Blockchain στην Ενεργειακή Αγορά

Η τεχνολογία blockchain αναμένεται να αλλάξει τον τρόπο με τον οποίο διαχειριζόμαστε και συναλλασσόμαστε ενέργεια, επιτρέποντας ασφαλείς και διαφανείς συναλλαγές μεταξύ παραγωγών και καταναλωτών. Οι τοπικές αγορές ενέργειας θα γίνουν πιο βιώσιμες και αποκεντρωμένες μέσω της χρήσης blockchain για τη διαχείριση των ενεργειακών συναλλαγών.

Πιθανές Εξελίξεις:

- Αποκεντρωμένες ενεργειακές αγορές: Το blockchain θα επιτρέπει σε καταναλωτές και παραγωγούς ενέργειας να ανταλλάσσουν ενέργεια απευθείας, χωρίς τη μεσολάβηση τρίτων, μειώνοντας το κόστος και ενισχύοντας τη διαφάνεια.
- Ασφαλείς συναλλαγές: Η τεχνολογία blockchain θα εξασφαλίσει τη διαφάνεια και την ασφάλεια των ενεργειακών συναλλαγών, μειώνοντας τον κίνδυνο απάτης και βελτιώνοντας την εμπιστοσύνη των συμμετεχόντων (Mengelkamp, 2018).

Δ.5 Συζήτηση

Οι ιδέες για τη μελλοντική ανάπτυξη αυτής της έρευνας περιλαμβάνουν τη χρήση τεχνολογιών τεχνητής νοημοσύνης (AI) και μηχανικής μάθησης (ML) σε μεγαλύτερη κλίμακα για την παροχή ακριβέστερων και έγκαιρων προβλέψεων κατανάλωσης ενέργειας. Ειδικότερα, η ενσωμάτωση πιο εξελιγμένων αλγορίθμων, όπως τα Βαθιά Νευρωνικά Δίκτυα και τα συστήματα ενισχυτικής μάθησης, θα επιτρέψει την ανάλυση ακόμη πιο σύνθετων σχέσεων μεταξύ δεδομένων και εξωτερικών παραγόντων (όπως η κλιματική αλλαγή και οι κοινωνικές συνθήκες).

Η εξέλιξη αυτή θα μπορούσε να αποσκοπεί σε μια σειρά σημαντικών βελτιώσεων, όπως οι παρακάτω:

- **Ακριβέστερη πρόβλεψη της κατανάλωσης:**
τα προηγμένα μοντέλα τεχνητής νοημοσύνης θα μπορούσαν να καταστήσουν τις προβλέψεις κατανάλωσης ακριβέστερες, λαμβάνοντας υπόψη παράγοντες που επηρεάζουν δυναμικά την

κατανάλωση ενέργειας, όπως οι καιρικές μεταβολές, οι κοινωνικοί και οικονομικοί παράγοντες ή τυχαία γεγονότα όπως οι φυσικές καταστροφές (Taylor & Letham, 2018).

- **Διαχείριση και βελτιστοποίηση των ενεργειακών πόρων:**

οι προμηθευτές ενέργειας μπορούν να χρησιμοποιούν αυτές τις προβλέψεις για να διαχειρίζονται τους ενεργειακούς πόρους τους πιο αποτελεσματικά. Οι προβλέψεις, τους επιτρέπουν να προσαρμόζουν την παραγωγή και τη διανομή των ενεργειακών πόρων ώστε να ανταποκρίνονται στη ζήτηση, να εξοικονομούν ενέργεια και να μειώνουν το λειτουργικό κόστος. Ταυτόχρονα, η σωστή πρόβλεψη μπορεί να αποτρέψει προβλήματα όπως η υπερφόρτωση του δικτύου (Deb, Zhang, Yang, Lee, & Shah, 2017).

- **Εξοικονόμηση ενέργειας για τους καταναλωτές:**

Η δυνατότητα πρόβλεψης της μελλοντικής κατανάλωσης επιτρέπει στους καταναλωτές να γνωρίζουν εκ των προτέρων την αναμενόμενη ζήτηση και τις τάσεις. Αυτό επιτρέπει στους καταναλωτές να προσαρμόσουν τη χρήση ενέργειας και να μειώσουν το κόστος. Για παράδειγμα, μπορούν να προγραμματίσουν την κατανάλωση ενέργειας σε περιόδους χαμηλής ζήτησης και να επωφεληθούν από καλύτερες τιμές (Deb, Zhang, Yang, Lee, & Shah, 2017).

- **Ανάλυση κρίσεων και διαχείριση κρίσεων:**

σε περίπτωση απρόβλεπτων κρίσεων, όπως ακραία καιρικά φαινόμενα ή διακοπές ρεύματος, τα μοντέλα TN μπορούν να προσαρμοστούν για την πρόβλεψη τέτοιων καταστάσεων. Αυτό επιτρέπει στους διαχειριστές να λαμβάνουν προληπτικά μέτρα και να ελαχιστοποιούν τις επιπτώσεις της κρίσης.

- **Αυτόματη εκμάθηση και βελτίωση των μοντέλων:**

τα μελλοντικά συστήματα μπορεί να βασίζονται σε τεχνικές ενισχυτικής εκμάθησης και να βελτιώνουν αυτόματα τα μοντέλα τους με την πάροδο του χρόνου. Όσο περισσότερα δεδομένα συλλέγονται, τόσο πιο αποτελεσματικό θα είναι το σύστημα πρόβλεψης, με μεγαλύτερη ακρίβεια και μικρότερα σφάλματα.

- **Διαχείριση μεγάλων δεδομένων:**

η επεξεργασία τεράστιων συνόλων δεδομένων μπορεί να γίνει πιο αποτελεσματική με την εφαρμογή της TN, επιτρέποντας την ανάλυση δεδομένων από περισσότερες πηγές, όπως έξυπνοι μετρητές, αισθητήρες και πληροφορίες καιρού σε πραγματικό χρόνο (M. M., Al-Fuqaha, Sorour, & M., 2018).

Δ.5.1 Προκλήσεις και πλαίσιο δεδομένων

Για να υλοποιηθεί αυτό το όραμα του μέλλοντος, είναι απαραίτητη η διαθεσιμότητα κατάλληλων δεδομένων σε πραγματικό χρόνο. Θα απαιτηθεί μια πλήρης βάση δεδομένων που θα περιέχει ετήσια δεδομένα, κατανάλωση ανά περιοχή, καιρικές συνθήκες και άλλες βασικές παραμέτρους. Η συνεχής ενημέρωση αυτών των δεδομένων θα επιτρέψει τη δημιουργία ενός δυναμικού και συνεχώς αναπτυσσόμενου μοντέλου πρόβλεψης. Συμπερασματικά, η χρήση τεχνολογιών τεχνητής νοημοσύνης στη διαχείριση της ενέργειας μπορεί να επιφέρει θεμελιώδεις αλλαγές τόσο στη βελτιστοποίηση χρήσης των ενεργειακών πόρων, όσο και στη μείωση του κόστους για τους καταναλωτές και να προωθήσει μια πιο βιώσιμη και αποδοτική χρήση της ενέργειας.

Δ.5.2 Συμπεράσματα

Η διαχείριση μεγάλου όγκου δεδομένων αποτελεί πλέον βασικό τομέα σε πολλούς κλάδους, και η διαχείριση της ενέργειας δεν αποτελεί εξαίρεση. Τα δεδομένα που παράγονται από έξυπνους μετρητές, αισθητήρες και άλλα συστήματα απαιτούν εξελιγμένη ανάλυση για να αποκαλύψουν πολύτιμες πληροφορίες σχετικά με τα πρότυπα κατανάλωσης ενέργειας. Αξιοποιώντας τεχνολογίες όπως η μηχανική μάθηση, η τεχνητή νοημοσύνη και η ανάλυση μεγάλων δεδομένων, οι οργανισμοί μπορούν να προβλέπουν με ακρίβεια τη ζήτηση ενέργειας, να βελτιστοποιούν τη χρήση των πόρων και να αυξάνουν τη βιωσιμότητα. Παρά τις προκλήσεις της επεξεργασίας μεγάλων δεδομένων, όπως η υψηλή υπολογιστική ισχύς και η διαχείριση δεδομένων σε πραγματικό χρόνο, τα οφέλη από τη σωστή διαχείρισή τους είναι σημαντικά. Συνολικά, η διαχείριση μεγάλων δεδομένων αποτελεί πολύτιμο εργαλείο για τη λήψη στρατηγικών αποφάσεων, την αύξηση της ενεργειακής απόδοσης και την προώθηση πιο βιώσιμων ενεργειακών πολιτικών στο μέλλον.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Ahmad T., Chen H. , & Wang, J. (2020). *A review on applications of artificial intelligence in smart grids: Future directions and challenges*. *Sustainable Cities and Society*, 61.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Christopher, B. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- D. C., Z. F., Y. J., L. S., & S. K. (2017). A Review on Time Series Forecasting Techniques for Building Energy Consumption. *Renewable and Sustainable Energy Reviews*, σσ. 74, 902-924.
- D. J., & G. S. (2008). *MapReduce: Simplified Data Processing on Large Clusters*. *Communications of the ACM*, 51(1), 107-113.
- Deb C. et al. (2017). *A review on time series forecasting techniques for building energy consumption*. *Renewable and Sustainable Energy Reviews*, 74.
- Deb C., & et al. (2017). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews*.
- F. C., X. F., & Z. Y. (2017). *A short-term building cooling load prediction method using deep learning algorithms*. *Applied Energy*, 195, 222-233.
- G. S., L. J., & H. F. (2015). *Data Preprocessing in Data Mining*. Springer.
- Goodfellow I., Bengio Y., & Courville A. (2016). *Deep Learning*. MIT Press.
- H. T., T. R., & F. J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- Hashem, I. A. T., et al. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, σσ. 98-115.
- Kang, C., et al. (2015). Big data analytics for grid modernization. *In IEEE Power & Energy Society General Meeting*, σσ. 1-5.
- L. M., V. E., & A. J. (2021). *Energy demand forecasting using deep learning methods: A case study of Spain's energy market*. *Energy Reports*, 7, 1-8.
- Liu, F., et al. (2020). *Artificial intelligence-based techniques for emerging power systems: State-of-the-art, challenges, and future research trends*. *Electric Power Systems Research*.
- M. E. (2018). A blockchain-based smart grid: towards sustainable local energy markets. *Computer Science - Research and Development*, σσ. 207-214.
- M. M. (2017). *Big IoT data analytics: Architecture, opportunities, and open research challenges*. IEEE Access.
- M. M., A.-F. A., S. S., & M., G. (2018). Deep Learning for IoT Big Data and Streaming Analytics: A Survey. *IEEE Communications Surveys & Tutorials*, σσ. 17-23.

- McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. Proceedings of the 9th Python in Science Conference, 51-56.
- Nazir, M. S., et al. (2021). *Smart grid edge analytics with deep learning: A systematic review for sustainable energy systems*. Sustainable Energy, Grids and Networks.
- R. S., & N. P. (2010). Artificial Intelligence: A Modern Approach. *Prentice Hall*.
- Russell S., & Norvig P. . (2010). Artificial Intelligence: A Modern Approach. *Prentice Hall*.
- T. S., & L. B. (2018). Forecasting at Scale. *The American Statistician*, σσ. 72(1), 37-45.
- Z. Q., & B. J. (2005). *Generation curtailment to manage voltage constraints in distribution networks*. IEEE Transactions on Power Systems, 20(1), 273-281.