Εθνικο Μετσοβιο Πολυτεχνειο
Σχολη Ηλεκτρολογων Μηχανικων και Μηχανικων Υπολογιστων
Τομεας Ηλεκτρικων Βιομηχανικων Διαταξεων και Συστηματων Αποφασεων

Προβλέψεις παραγωγής από ανανεώσιμες πηγές ενέργειας: Προσεγγίσεις με επίκεντρο την Ομοσπονδιακή Μάθηση και την Ιδιωτικότητα

Διπλωματικη Εργασια

του

ΕΥΣΤΑΘΙΟΥ ΑΛΕΞΑΝΔΡΟΥ Κ. ΣΑΡΑΝΤΙΝΟΠΟΥΛΟΥ

Επιβλέπων: Ευάγγελος Μαρινάκης
Επίκουρος Καθηγητής

Αθήνα, Φεβρουάριος 2024

Εθνικο Μετσοβιο Πολυτεχνειο
Σχολη Ηλεκτρολογων Μηχανικων και Μηχανικων Υπολογιστων
Τομεας Ηλεκτρικων Βιομηχανικων Διαταξεων και Συστηματων Αποφασεων

Προβλέψεις παραγωγής από ανανεώσιμες πηγές ενέργειας: Προσεγγίσεις με επίκεντρο την Ομοσπονδιακή Μάθηση και την Ιδιωτικότητα

Διπλωματικη Εργασια

του

ΕΥΣΤΑΘΙΟΥ ΑΛΕΞΑΝΔΡΟΥ Κ. ΣΑΡΑΝΤΙΝΟΠΟΥΛΟΥ

Επιβλέπων: Ευάγγελος Μαρινάκης
Επίκουρος Καθηγητής

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 29η Φεβρουαρίου 2024.

(Υπογραφή)	*(Υπογραφή)*	*(Υπογραφή)*
.....	.....	.....
Ευάγγελος Μαρινάκης	Ιωάννης Ψαρράς	Δημήτριος Ασκούνης
Επίκουρος Καθηγητής	Καθηγητής	Καθηγητής

Αθήνα, Φεβρουάριος 2024

(Υπογραφή)

.....
Ευστάθιος Αλέξανδρος Κ.
Σαραντινόπουλος

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

Περίληψη

Η παρούσα διπλωματική εργασία διερευνά την εφαρμογή της ομοσπονδιακής μάθησης για την πρόβλεψη παραγωγής ενέργειας από ανανεώσιμες πηγές, με στόχο την επίτευξη ισορροπίας μεταξύ της ακρίβειας πρόβλεψης και της ιδιωτικότητας. Ξεκινά καλύπτοντας τις βασικές αρχές της ανάλυσης χρονοσειρών, της μηχανικής μάθησης και των νευρωνικών δικτύων που σχετίζονται με τις εργασίες πρόβλεψης. Ταυτόχρονα, εξετάζονται οι αρχές της ομοσπονδιακής μάθησης με εγγυήσεις διαφορικής ιδιωτικότητας, αναδεικνύοντας τις δυνατότητές τους για κατανεμημένη μάθηση με διατήρηση της ιδιωτικότητας.

Μέσω πειραμάτων σε ένα σύνολο δεδομένων 30 μικρής κλίμακας παραγωγών/καταναλωτών ηλεκτρικής ενέργειας, η εργασία συγκρίνει την ομοσπονδιακή μάθηση έναντι της συγκεντρωτικής, της τοπικής και της ενισχυμένης με διαφορική ιδιωτικότητα ομοσπονδιακής μάθησης. Τα αποτελέσματα δείχνουν ότι ενώ η συγκεντρωτική αποδίδει την υψηλότερη ακρίβεια, η ομοσπονδιακή παρέχει μια πειστική εναλλακτική λύση, διατηρώντας την ιδιωτικότητα χωρίς σημαντική υποβάθμιση της απόδοσης. Η υλοποίηση με διαφορική ιδιωτικότητα, αν και προσφέρει ισχυρότερες εγγυήσεις ιδιωτικότητας έχει χαμηλότερη επίδοση από τις άλλες συγκεντρωτικές μεθόδους ενώ υπερισχύει της τοπικής.

Ακόμη εξετάζεται μια πρωτοποριακή μέθοδος ομαδοποίησης των δεδομένων με βάση τις υπερπαραμέτρους των μοντέλων καθώς και η επίδραση της ομαδοποίησης στην απόδοση.

Συμπερασματικά, η παρούσα διπλωματική εργασία επιβεβαιώνει τη βιωσιμότητα της ομοσπονδιακής μάθησης για την πρόβλεψη παραγωγής, ιδίως σε σενάρια ευαίσθητων δεδομένων. Ακόμη, προσφέρει πληροφορίες σχετικά με τους συμβιβασμούς επιδόσεων που σχετίζονται με τη διαφορική ιδιωτικότητα και διερευνά τον ρόλο της ομαδοποίησης δεδομένων σε περιπτώσεις όπου υπάρχουν περιορισμοί πρόσβασης στα δεδομένα.

Λέξεις-κλειδιά: Πρόβλεψη παραγωγής, ομοσπονδιακή μάθηση, διαφορική ιδιωτικότητα, ανάλυση χρονοσειρών, νευρωνικά δίκτυα, ομαδοποίηση δεδομένων

Abstract

This thesis explores the application of federated learning (FL) for energy generation forecasting, with the goal of achieving a balance between forecast accuracy and privacy. It begins by covering the fundamentals of time series analysis, machine learning and neural networks related to forecasting tasks. At the same time, the principles of FL and differential privacy (DP) are reviewed, highlighting their potential for privacy-preserving distributed learning.

Through experiments on a dataset of 30 small-scale electricity prosumers, the thesis compares federated learning against centralized learning (CL), local learning (LL) and differential privacy-enhanced federated learning (FL-DP). Key findings indicate that while CL yields the highest accuracy, FL provides a convincing alternative, preserving privacy without significant performance degradation. The implementation with differential privacy, while offering stronger privacy guarantees, has lower performance than the other methods while outperforming LL.

Furthermore, the thesis explores an innovative method of clustering the data based on the hyperparameters of the models and the effect of clustering on performance.

In conclusion, this thesis confirms the viability of federated learning for production forecasting, especially in sensitive data scenarios. Furthermore, it offers insights into performance trade-offs associated with differential privacy and explores the role of data clustering in cases where the access to data is limited.

Keywords: Generation forecasting, federated learning, differential privacy, time series analysis, neural networks

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον καθηγητή μου κ. Ευάγγελο Μαρινάκη για την ευκαιρία που μου έδωσε να ασχοληθώ με ένα εξαιρετικά ενδιαφέρον θέμα, για την επίβλεψη, καθώς και την εμπιστοσύνη που μου έδειξε κατά την διάρκεια της εκπόνησης της εργασίας.

Ακόμη, ευχαριστώ ιδιαίτερα και τον υποψήφιο διδάκτορα κ. Βασίλειο Μιχαλακόπουλο για την καθοδήγηση, την προθυμία, καθώς και την εξαιρετική συνεργασία που είχαμε κατά την συγγραφή της παρούσας εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω από καρδιάς την οικογένειά μου, τους αγαπημένους μου φίλους και την Δανάη για την υποστήριξή τους στα χρόνια των σπουδών μου και όχι μόνο, καθώς και την Σελήνη για την συνεισφορά της στην συγγραφή αυτής της εργασίας.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ

1 Εισαγωγή

Η εξάπλωση των ανανεώσιμων πηγών ενέργειας, όπως η ηλιακή και η αιολική [1], έχει φέρει επανάσταση στο ενεργειακό τοπίο, επιτρέποντας στους καταναλωτές να γίνουν επαγγελματίες-καταναλωτές (prosumers), παράγοντας ενέργεια παράλληλα με τις παραδοσιακές κεντρικές πηγές. Αυτό το αποκεντρωμένο μοντέλο παραγωγής ενέργειας επιφέρει οφέλη όσον αφορά τη βιωσιμότητα και τη μειωμένη εξάρτηση από το δίκτυο. Ωστόσο, δημιουργεί επίσης προκλήσεις στη διαχείριση και βελτιστοποίηση του δικτύου, ιδίως όταν πρόκειται για την πρόβλεψη της μελλοντικής παραγωγής ενέργειας [2].

Οι έξυπνοι μετρητές και οι συναφείς τεχνολογίες που χρησιμοποιούνται από αυτούς τους prosumers παράγουν πληθώρα πολύτιμων δεδομένων. Τα δεδομένα αυτά μπορούν να αξιοποιηθούν για την ακριβή πρόβλεψη της παραγωγής ενέργειας, επιτρέποντας στους φορείς διαχείρισης του ηλεκτρικού δικτύου να προβλέπουν αποτελεσματικότερα την προσφορά και να λαμβάνουν τεκμηριωμένες αποφάσεις σχετικά με την κατανομή των πόρων και την εξισορρόπηση του φορτίου [3]. Ωστόσο, τα δεδομένα αυτά συχνά περιέχουν ευαίσθητες πληροφορίες σχετικά με τα ατομικά πρότυπα χρήσης ενέργειας, εγείροντας σημαντικές ανησυχίες για την προστασία της ιδιωτικότητας. Δυστυχώς, τα δεδομένα αυτά είναι επίσης εγγενώς ευαίσθητα, καθώς περιέχουν πληροφορίες σχετικά με τα πρότυπα κατανάλωσης ενέργειας ενός νοικοκυριού και ενδεχομένως αποκαλύπτουν λεπτομέρειες της καθημερινής ρουτίνας [4]. Τα τελευταία χρόνια έχουν προκύψει αυστηροί κανονισμοί για την προστασία της ιδιωτικότητας των δεδομένων, όπως ο Γενικός Κανονισμός για την Προστασία Δεδομένων (GDPR) [5] της Ευρωπαϊκής Ενωσης, υπογραμμίζοντας την ανάγκη για ισχυρές λύσεις διατήρησης της ιδιωτικότητας σε βιομηχανίες που βασίζονται στα δεδομένα.

Τα παραδοσιακά μοντέλα μηχανικής μάθησης βασίζονται συνήθως σε κεντρική συγκέντρωση δεδομένων. Αυτή η προσέγγιση, αν και αποτελεσματική για την επίτευξη υψηλής ακρίβειας, μπορεί να εκθέσει ευαίσθητες πληροφορίες σε πιθανές παραβιάσεις ή κατάχρηση. Η ομοσπονδιακή μάθηση προσφέρει μια εναλλακτική λύση, εκπαιδεύοντας τα μοντέλα συνεργατικά

σε κατανεμημένα σύνολα δεδομένων χωρίς να απαιτείται τα ακατέργαστα δεδομένα να εγκαταλείψουν ποτέ τη συσκευή του ιδιοκτήτη [6]. Αυτή η διαδικασία επιτρέπει τη μάθηση από διαφορετικές πηγές, διατηρώντας παράλληλα το απόρρητο των δεδομένων.

Επιπλέον, η διαφορική ιδιωτικότητα μπορεί να ενσωματωθεί με την ομοσπονδιακή μάθηση για να παρέχει ακόμη ισχυρότερες εγγυήσεις ιδιωτικότητας. Η διαφορική ιδιωτικότητα εισάγει προσεκτικά βαθμονομημένο θόρυβο στη διαδικασία μάθησης, αποκρύπτοντας τη συμβολή των δεδομένων κάθε μεμονωμένου χρήστη και καθιστώντας δυσκολότερη την εκ νέου ταυτοποίηση ευαίσθητων πληροφοριών [7].

1.1 Σκοπός της εργασίας

Ο σκοπός της παρούσας διπλωματικής είναι να διερευνηθεί η βιωσιμότητα της ομοσπονδιακής μάθησης ως ένα πλαίσιο διατήρησης της ιδιωτικότητας με υψηλή ακρίβεια προβλέψεων για εργασίες πρόβλεψης παραγωγής. Συγκεκριμένα, η παρούσα έρευνα αποσκοπεί στα εξής:

- Να διερευνήσει την απόδοση των μοντέλων ομοσπονδιακής μάθησης σε σύγκριση με την κεντρική μάθηση, την τοπική μάθηση και την ομοσπονδιακή μάθηση με εγγυήσεις διαφορικής ιδιωτικότητας.
- Να αξιολογήσει τον συμβιβασμό μεταξύ της ακρίβειας του μοντέλου και του επιπέδου προστασίας της ιδιωτικότητας που προσφέρει η διαφορική ιδιωτικότητα.
- Να εξετάσει τον αντίκτυπο των στρατηγικών ομαδοποίησης δεδομένων σε περιβάλλοντα με περιορισμένη πρόσβαση στα δεδομένα.

1.2 Δομή της εργασίας

Τα επόμενα κεφάλαια αυτής της διπλωματικής παρέχουν μια βάση στην ανάλυση χρονοσειρών, τη μηχανική μάθηση και τα νευρωνικά δίκτυα. Εμβαθύνουν στις αρχές της ομοσπονδιακής μάθησης και της διαφορικής ιδιωτικότητας και ακολουθούν την πειραματική διαδικασία και την ανάλυση των αποτελεσμάτων. Η εργασία ολοκληρώνεται με τα βασικά ευρήματα και τις ιδέες για μελλοντική έρευνα.

2 Χρονοσειρές

Σε αυτό το κεφάλαιο, εξετάζουμε τις χρονοσειρές, έναν βασικό τομέα μελέτης για την ανάλυση μοτίβων και την πρόβλεψη μελλοντικών τάσεων σε δεδομένα που συλλέγονται με την πάροδο του χρόνου. Ξεκινώντας με μια βασική κατανόηση του τι είναι οι χρονοσειρές και της σημασίας τους, προχωρούμε σταδιακά στην εξέταση διαφόρων μεθόδων πρόβλεψης και των προκλήσεων που συνεπάγεται η πρόβλεψη. Η συζήτηση επεκτείνεται για να καλύψει τα βασικά χαρακτηριστικά των χρονοσειρών, όπως οι τάσεις, η εποχικότητα και η στασιμότητα, και πώς αυτά τα στοιχεία επηρεάζουν την ανάλυση. Ασχολούμαστε επίσης με τα κρίσιμα βήματα της προεπεξεργασίας των δεδομένων χρονοσειρών, συμπεριλαμβανομένης της αποσύνθεσης, του χειρισμού του θορύβου και των ακραίων τιμών και της προετοιμασίας των δεδομένων μέσω της κλιμάκωσης και της κανονικοποίησης.

2.1 Ορισμός και φύση των Χρονοσειρών

Τα δεδομένα χρονοσειρών αποτελούνται από παρατηρήσεις που καταγράφονται διαδοχικά στο χρόνο και είναι ζωτικής σημασίας σε διάφορους τομείς όπως οι χρηματοοικονομικές αγορές, η μετεωρολογία και η πρόβλεψη ενεργειακών φορτίων. Το διακριτικό χαρακτηριστικό των δεδομένων χρονοσειρών είναι η χρονολογική τους σειρά, η οποία είναι ουσιώδης για την ανάλυση και την προγνωστική μοντελοποίηση. Οι χρονοσειρές μπορεί να είναι μονομεταβλητές, με μία παρατηρήσιμη μεταβλητή, ή πολυμεταβλητές, με πολλές αλληλεξαρτώμενες μεταβλητές. [8]

2.2 Προβλέψεις χρονοσειρών

Η πρόβλεψη χρονοσειρών είναι μια μεθοδολογική προσέγγιση που χρησιμοποιείται για την πρόβλεψη μελλοντικών γεγονότων ή τιμών με την ανάλυση μοτίβων σε ιστορικά δεδομένα που καταγράφονται με την πάροδο του χρόνου. Η ουσία της πρόβλεψης χρονοσειρών έγκειται στην ικανότητά της να μοντελοποιεί την εγγενή χρονική δυναμική ενός συνόλου δεδομένων, επιτρέποντας προβλέψεις για μελλοντικές τιμές με βάση παρελθοντικές και παρούσες παρατηρήσεις. [8]

Οπως αναφέρθηκε, τα δεδομένα χρονοσειρών έχουν μεγάλη χρησιμότητα σε διάφορους τομείς. Για παράδειγμα, η πρόβλεψη χρονοσειρών στα χρηματοοικονομικά, βοηθά στην πρόβλεψη των τιμών των μετοχών και των τάσεων της αγοράς. Στην πρόγνωση καιρού, προβλέπει τις καιρικές συνθήκες, βοηθώντας στη διαχείριση καταστροφών. Στον τομέα της ενέργειας, βοηθά στην πρόβλεψη της ζήτησης και της προσφοράς ενέργειας, απαραίτητη για τον αποτελεσματικό προγραμματισμό των πόρων.

2.2.1 Μέθοδοι Πρόβλεψης

Οι μέθοδοι πρόβλεψης χρονοσειρών κυμαίνονται από απλές στατιστικές τεχνικές έως πολύπλοκα μοντέλα μηχανικής μάθησης. Μερικές από τις παραδοσιακές μεθόδους είναι τα μοντέλα αυτοπαλίνδρομου ολοκληρωμένου κινητού μέσου (ARIMA) και η εκθετική εξομάλυνση. Με την έλευση των μεγάλων δεδομένων και των υπολογιστικών εξελίξεων, οι τεχνικές μηχανικής μάθησης, όπως τα νευρωνικά δίκτυα και η βαθιά μάθηση, έχουν γίνει όλο και πιο δημοφιλείς για πιο σύνθετες εργασίες πρόβλεψης και είναι και αυτές που χρησιμοποιήθηκαν στην παρούσα εργασία. [9]

2.2.2 Προκλήσεις προβλέψεων

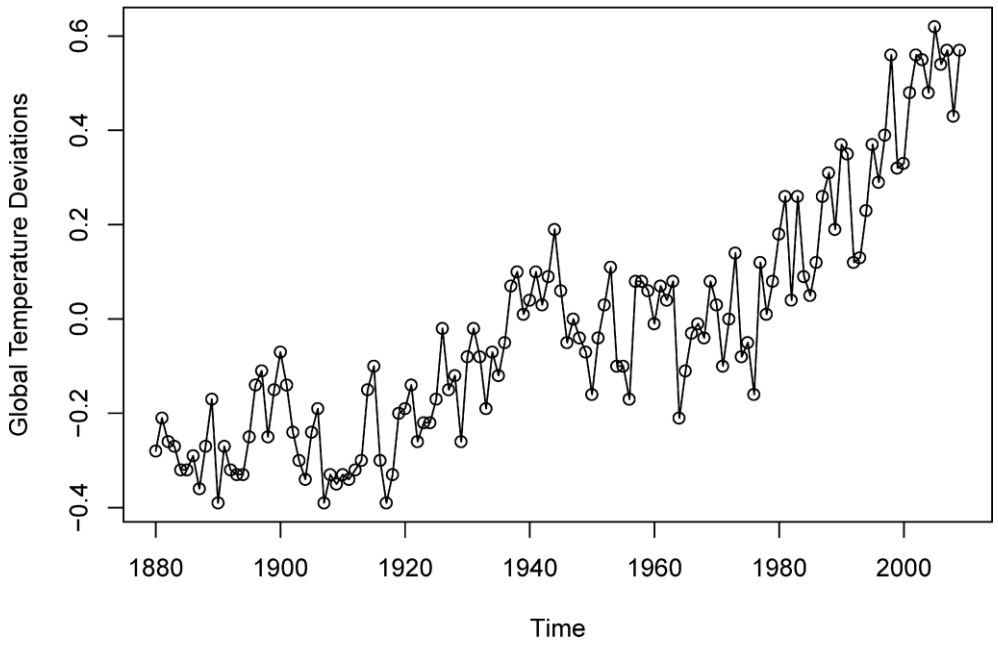
Οι προκλήσεις της πρόβλεψης περιλαμβάνουν την αντιμετώπιση ακανόνιστων προτύπων, τη διαχείριση ελλιπών δεδομένων και την κατανόηση και τον χειρισμό των επιπτώσεων εξωτερικών παραγόντων. Η ακρίβεια μιας πρόβλεψης μπορεί να επηρεαστεί σημαντικά από την ποιότητα και τη φύση των δεδομένων, την καταλληλόλητα του επιλεγμένου μοντέλου και την ικανότητα του ατόμου που πραγματοποιεί την πρόβλεψη να ερμηνεύει με ακρίβεια τα αποτελέσματα.

2.3 Χαρακτηριστικά χρονοσειρών

Οι περισσότερες χρονοσειρές επηρεάζονται από 4 βασικά ποιοτικά χαρακτηριστικά, τα οποία είναι διακριτά από τα παρατηρούμενα δεδομένα. Αυτά είναι η τάση, η κυκλικότητα, η εποχικότητα και η τυχαιότητα. Ακόμη, ιδιαίτερη σημασία δίνεται και σε μερικά ακόμα ποιοτικά χαρακτηριστικά που εμφανίζονται σε χρονοσειρές όπως η στασιμότητα και η οι ασυνέχειες. [8]

2.3.1 Τάση

Με τον όρο "Τάση" αναφερόμαστε στη συνολική κατεύθυνση προς την οποία κινούνται τα δεδομένα μιας χρονοσειράς κατά τη διάρκεια μιας εκτεταμένης περιόδου, υποδεικνύοντας αν αυξάνονται, μειώνονται ή παραμένουν σταθερά. Για παράδειγμα, οι χρονοσειρές που αφορούν την αύξηση του πληθυσμού, τον αριθμό των κατοικιών σε μια πόλη κ.λπ. παρουσιάζουν ανοδική τάση, ενώ πτωτική τάση μπορεί να παρατηρηθεί σε σειρές που αφορούν τα ποσοστά θνησιμότητας, τις επιδημίες κ.λπ. [10]

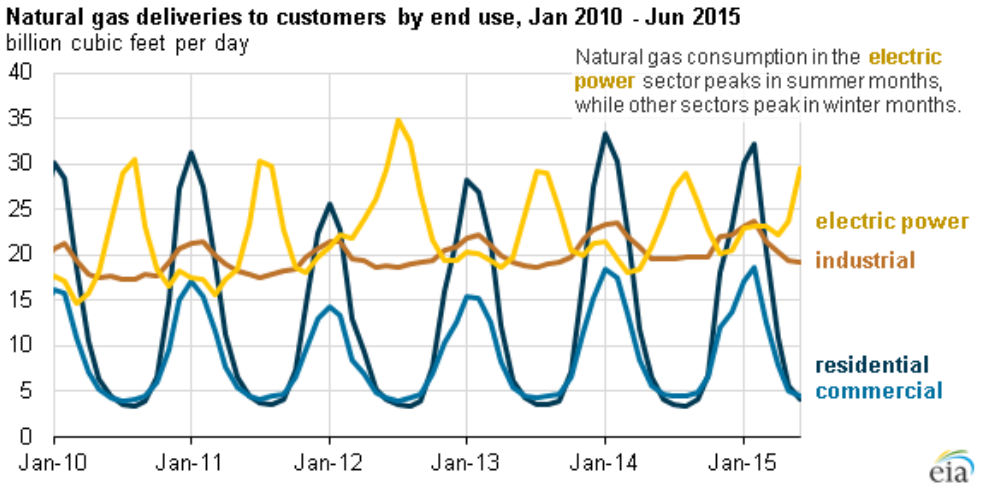


Εικόνα 2.1: *Ετήσιες μέσες αποκλίσεις της παγκόσμιας θερμοκρασίας (1880-2009) σε βαθμούς Κελσίου [11]*

Στην εικόνα 2.1 βλέπουμε μια χρονοσειρά, τα δεδομένα της οποίας είναι ο παγκόσμιος μέσος δείκτης θερμοκρασίας ξηράς-ωκεανού από το 1880 έως το 2009, με περίοδο βάσης 1951-1980. Πιο συγκεκριμένα, τα δεδομένα είναι αποκλίσεις, μετρούμενες σε βαθμούς Κελσίου, από τον μέσο όρο της περιόδου 1951-1980. Παρατηρούμε μια εμφανή ανοδική τάση στη χρονοσειρά κατά το τελευταίο μέρος του εικοστού αιώνα, η οποία έχει χρησιμοποιηθεί ως επιχείρημα υπέρ της υπόθεσης της υπερθέρμανσης του πλανήτη. [11]

2.3.2 Εποχικότητα

Η εποχικότητα σε δεδομένα χρονοσειρών αναφέρεται σε τακτικά και προβλέψιμα μοτίβα ή κινήσεις που επαναλαμβάνονται σε συγκεκριμένα χρονικά διαστήματα, τα οποία συνήθως επηρεάζονται από την εποχή του έτους, τον μήνα, την εβδομάδα ή ακόμη και την ώρα της ημέρας. Αυτά τα μοτίβα αποτελούν θεμελιώδη πτυχή πολλών χρονοσειρών και είναι ιδιαίτερα διαδεδομένα σε δεδομένα που σχετίζονται με τον καιρό, τις λιανικές πωλήσεις, την κατανάλωση ενέργειας κ.α., όπου αντικατοπτρίζουν την επιρροή εποχικών παραγόντων. Για παράδειγμα, οι επιχειρήσεις λιανικού εμπορίου συχνά παρουσιάζουν εποχιακές αιχμές κατά τη διάρκεια των περιόδων διακοπών και η χρήση ενέργειας μπορεί να αυξηθεί κατά τη διάρκεια του χειμώνα λόγω των αναγκών θέρμανσης ή του καλοκαιριού λόγω της χρήσης κλιματιστικών.

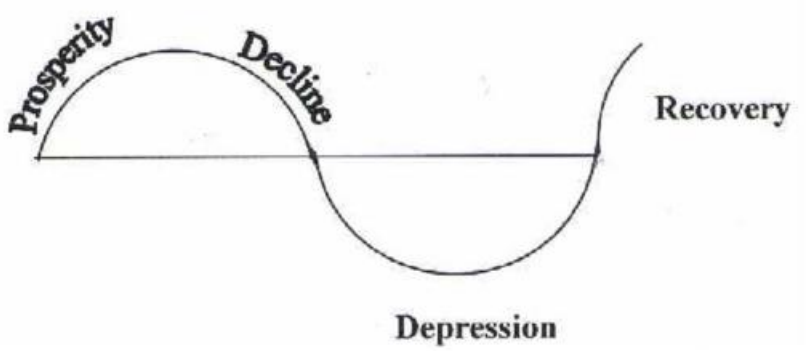


Εικόνα 2.2: *Παραδόσεις φυσικού αερίου σε καταναλωτές ανά τελική χρήση, Ιανουάριος 2010 – Ιούνιος 2015 [12]*

Στην εικόνα 2.2 παρατηρούμε πως η χρήση του φυσικού αερίου έχει δύο εποχιακές αιχμές, με τα μοντέλα κατανάλωσης να καθορίζονται κυρίως από τις καιρικές συνθήκες. Η μεγαλύτερη αιχμή παρατηρείται κατά τη διάρκεια του χειμώνα, όταν ο κρύος καιρός αυξάνει τη ζήτηση για θέρμανση χώρων με φυσικό αέριο στον οικιακό και εμπορικό τομέα. Μια δεύτερη, μικρότερη αιχμή παρατηρείται το καλοκαίρι, όταν η χρήση κλιματισμού αυξάνει τη ζήτηση ηλεκτρικής ενέργειας. [12]

2.3.3 Κυκλικότητα

Η κυκλικότητα σε δεδομένα χρονοσειρών αναφέρεται σε διακυμάνσεις που συμβαίνουν σε ακανόνιστα διαστήματα. Σε αντίθεση με την εποχικότητα, η οποία συνήθως παρατηρείται εντός μιας σταθερής και γνωστής περιόδου (όπως ημέρες, μήνες ή τρίμηνα), τα κυκλικά μοτίβα χαρακτηρίζονται από μεταβλητές συχνότητες και συχνά επηρεάζονται από ευρύτερους οικονομικούς, πολιτικούς ή περιβαλλοντικούς παράγοντες. Αυτοί οι κύκλοι μπορεί να καλύπτουν αρκετά χρόνια, γεγονός που τους καθιστά πιο πολύπλοκους και λιγότερο προβλέψιμους από τις εποχικές τάσεις. Για παράδειγμα, οι οικονομικές χρονοσειρές συχνά παρουσιάζουν κύκλους που αντιστοιχούν σε περιόδους επέκτασης και συρρίκνωσης, ανταναικλώντας τις φάσεις του επιχειρηματικού κύκλου. Για παράδειγμα, ένας επιχειρηματικός κύκλος αποτελείται από τέσσερις φάσεις, δηλαδή i) ευημερία, ii) πτώση, iii) ύφεση και iv) ανάκαμψη και μπορεί να αναπαρασταθεί σχηματικά ως εξής:



Εικόνα 2.3: *Ενας οικονομικός κύκλος τεσσάρων φάσεων [10]*

Ο εντοπισμός και η ανάλυση των κυκλικών προτύπων είναι ιδιαίτερα σημαντικά για την αποτελεσματική πρόβλεψη χρονοσειρών, ιδίως σε τομείς όπως η οικονομία και η χρηματοοικονομική, όπου τα πρότυπα αυτά μπορεί να έχουν σημαντικές επιπτώσεις. Ωστόσο, η διάκριση των κυκλικών μοτίβων από άλλες συνιστώσες της χρονοσειράς, όπως η συνολική τάση ή οι εποχικές επιδράσεις, μπορεί να αποτελέσει πρόκληση λόγω της ακανόνιστης φύσης και της μεταβλητής διάρκειάς τους.

2.3.4 Τυχαιότητα

Η τυχαιότητα στα δεδομένα χρονοσειρών, περιλαμβάνει τις απρόβλεπτες και μη συστηματικές διακυμάνσεις που δεν μπορούν να αποδοθούν σε τάση, εποχικότητα ή κυκλικά μοτίβα [10]. Αυτές οι ακανόνιστες διακυμάνσεις εκδηλώνονται ως στοχαστικές αποκλίσεις, που προκύπτουν από απρόβλεπτα ή τυχαία γεγονότα, τα οποία εισάγουν ένα επίπεδο πολυπλοκότητας στη διαδικασία πρόβλεψης. Στο πλαίσιο της πρόβλεψης φορτίου ενέργειας, αυτές οι διακυμάνσεις θα μπορούσαν να αναφερθούν ως παράδειγμα σε απότομες αλλαγές στην κατανάλωση ενέργειας λόγω απρόβλεπτων καιρικών φαινομένων ή σε ανωμαλίες στην παραγωγή ενέργειας που προέρχονται από απρογραμμάτιστες διακοπές ή βλάβες εξοπλισμού [13].

Ο ακριβής προσδιορισμός και η ποσοτικοποίηση των ακανόνιστων διακυμάνσεων είναι εξαιρετικά σημαντική για τη βελτίωση της ακρίβειας των μοντέλων πρόβλεψης ενέργειας. Λόγω της στοχαστικής τους φύσης, αυτές οι ακανόνιστες συνιστώσες συχνά υπερβαίνουν τις δυνατότητες των παραδοσιακών μεθόδων πρόβλεψης, καθιστώντας αναγκαία την εφαρμογή πιο εξελιγμένων προσεγγίσεων. Οι προηγμένες στατιστικές τεχνικές και τα μοντέλα μηχανικής μάθησης, ιδίως τα δίκτυα μακράς βραχυπρόθεσμης μνήμης (LSTM), τα οποία θα χρησιμοποιήσουμε και εμείς, έχουν αποδειχθεί ιδιαίτερα αποτελεσματικά στην καταγραφή αυτών των ακανόνιστων διακυμάνσεων. [14]

2.3.5 Στασιμότητα

Η στασιμότητα είναι μια θεμελιώδης έννοια στην ανάλυση χρονοσειρών που υποθέτει ότι τα στατιστικά χαρακτηριστικά των δεδομένων (όπως ο μέσος όρος τους και ο τρόπος με τον οποίο κυμαίνονται) παραμένουν σταθερά με την πάροδο του χρόνου [15]. Η στασιμότητα στα δεδομένα χρονοσειρών κατηγοριοποιείται σε δύο τύπους: ισχυρή και ασθενής στασιμότητα. Η ισχυρή στασιμότητα συνεπάγεται ότι οι στατιστικές ιδιότητες, συμπεριλαμβανομένων του μέσου όρου, και της διακύμανσης είναι συνεπείς σε ολόκληρη τη χρονοσειρά, ανεξάρτητα από τη χρονική στιγμή κατά την οποία μετριούνται. Αυτό σημαίνει ότι ολόκληρη η κατανομή των τιμών παραμένει σταθερή με την πάροδο του χρόνου. Από την άλλη πλευρά, η ασθενής στασιμότητα, η οποία είναι συχνά πιο πρακτική για δεδομένα του πραγματικού κόσμου, απαιτεί να παραμένουν σταθερές μόνο η μέση τιμή, η διακύμανση και η αυτοσυνδιακύμανση [11]. Η ασθενής

στασιμότητα είναι μια λιγότερο αυστηρή συνθήκη, γεγονός που την καθιστά πιο εφαρμόσιμη σε σενάρια όπως η πρόβλεψη ενέργειας, όπου τα δεδομένα μπορεί να παρουσιάζουν πολύπλοκες συμπεριφορές.

2.4 Προ-επεξεργασία χρονοσειρών

Η προεπεξεργασία χρονοσειρών είναι ένα κρίσιμο βήμα στον τομέα των προβλέψεων, όπου τα ακατέργαστα δεδομένα μετασχηματίζονται και προετοιμάζονται για την επακόλουθη ανάλυση και τη δημιουργία μοντέλων. Η διαδικασία αυτή περιλαμβάνει μια σειρά τεχνικών που αποσκοπούν στη βελτίωση της ποιότητας και της αποτελεσματικότητας των δεδομένων, διασφαλίζοντας ότι αυτά αντικατοπτρίζουν με ακρίβεια τα υποκείμενα πρότυπα και τάσεις.

Η προεπεξεργασία περιλαμβάνει εργασίες όπως η αποσύνθεση της χρονοσειράς στα συστατικά της στοιχεία, ο χειρισμός του θορύβου και των ακραίων τιμών και η διαχείριση των ελλιπών τιμών. Συνήθως, παράλληλα με την προεπεξεργασία, διενεργείται και η διερευνητική ανάλυση δεδομένων (ΔΑΔ)/exploratory data analysis (EDA) την οποία θα αναλύσουμε σε επόμενο κεφάλαιο.

2.4.1 Αποσύνθεση χρονοσειρών

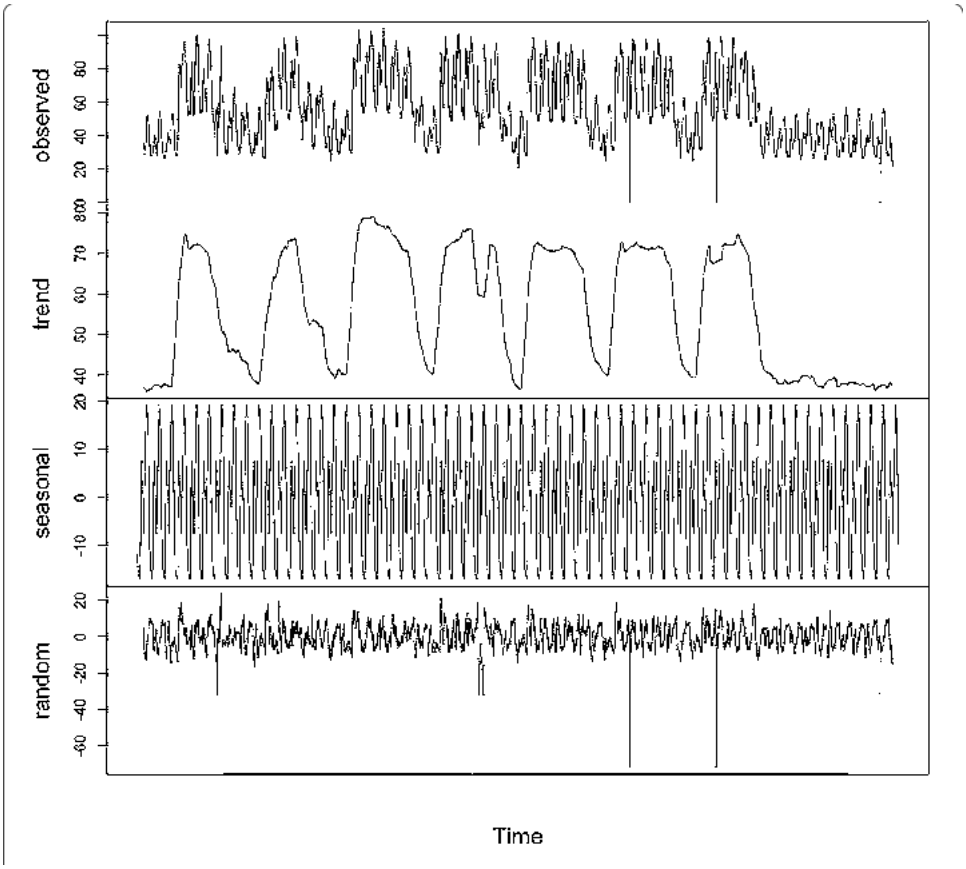
Η αποσύνθεση χρονοσειρών επιτρέπει στους υπεύθυνους πρόβλεψης να απομονώσουν και να κατανοήσουν τα υποκείμενα πρότυπα της ζήτησης ενέργειας μέσω του διαχωρισμού της χρονοσειράς σε τάση $T(t)$, εποχικότητα $S(t)$ και ανωμαλίες $I(t)$.

Η αποσύνθεση χωρίζεται σε δύο κύρια είδη:

- την προσθετική, όπου το μοντέλο της χρονοσειράς αναπαρίσταται ως:
$$X(t) = T(t) + S(t) + I(t)$$

- την πολλαπλασιαστική, όπου το μοντέλο της χρονοσειράς αναπαρίσταται ως:
$$X(t) = T(t) \times S(t) \times I(t)$$

Η επιλογή μεταξύ ενός προσθετικού και ενός πολλαπλασιαστικού μοντέλου εξαρτάται από τη φύση της χρονοσειράς. Για την πρόβλεψη παραγωγής, η επιλογή συχνά καθοδηγείται από το πρότυπο της εποχιακής διακύμανσης. Εάν η εποχική επίδραση μεταβάλλεται αναλογικά με το επίπεδο της χρονοσειράς, ένα πολλαπλασιαστικό μοντέλο είναι καταλληλότερο. Ωστόσο, εάν η εποχιακή επίδραση παραμένει σχετικά σταθερή με την πάροδο του χρόνου, ένα προσθετικό μοντέλο είναι προτιμότερο. Σε πολλά πρακτικά σενάρια πρόβλεψης φορτίου, ιδίως όταν υπάρχει σημαντική διακύμανση στη χρήση ενέργειας σε διάφορες εποχές του έτους, ένα πολλαπλασιαστικό μοντέλο είναι συχνά καλύτερο. Αυτό οφείλεται στο γεγονός ότι η ζήτηση ενέργειας τείνει να αυξάνεται αναλογικά με παράγοντες όπως οι μεταβολές της θερμοκρασίας ή οι μεταβολές της συμπεριφοράς των καταναλωτών, οι οποίες είναι εγγενώς εποχικές και μπορούν να ενταθούν ανάλογα με το επίπεδο της ζήτησης. [8]



Εικόνα 2.4: *Αποσύνθεση χρονοσειρών κατανάλωσης ενέργειας για μια κατοικημένη περιοχή [16]*

Στην παραπάνω εικόνα παρατηρούμε την αποσύνθεση μιας χρονοσειράς κατανάλωσης ενέργειας για μια κατοικημένη περιοχή. Ο άξονας x δείχνει το ηλεκτρικό ρεύμα που διαρρέει τις γραμμές διανομής και καταγράφεται κάθε 10 λεπτά. Ο άξονας y αναφέρεται στο ηλεκτρικό ρεύμα που διαρρέει μια συγκεκριμένη γραμμή μεταφοράς ηλεκτρικής ενέργειας μετρούμενο σε Ampère σε μια συγκεκριμένη χρονική στιγμή.

Η πρώτη χρονοσειρά είναι η παρατηρούμενη και στην συνέχεια αποσυντίθεται στα επι μέρους χαρακτηριστικά της. Η δεύτερη χρονοσειρά για παράδειγμα είναι η χρονοσειρά τάσης και παρατηρούμε το χαρακτηριστικό μοτίβο στις χρονοσειρές ενέργειας, με αυξητική τάση κατά την διάρκεια της μέρας που ακολουθείται από πτώση στις βραδινές ώρες. Η τρίτη και η τέταρτη χρονοσειρά δείχνουν την εποχικότητα και την τυχαιότητα αντίστοιχα.

Η αποσύνθεση χρονοσειρών είναι ιδιαίτερα χρήσιμη όταν χρησιμοποιούνται κλασικές στοχαστικές τεχνικές προβλέψεων. Στην περίπτωση χρήσης βαθιών νευρωνικών δικτύων, όπως τα μοντέλα μακράς βραχυπρόθεσμης μνήμης (LSTM), για πρόβλεψη δεν είναι απολύτως απαραίτητη, αλλά μπορεί να είναι επωφελής ανάλογα με το πλαίσιο και τους ειδικούς στόχους της ανάλυσής σας. Τα μοντέλα LSTM είναι ικανά να συλλάβουν πολύπλοκα μοτίβα σε ακατέργαστα δεδομένα χρονοσειρών, συμπεριλαμβανομένων των τάσεων, της εποχικότητας και των ανωμαλιών, χωρίς ρητή αποσύνθεση. Αυτά τα μοντέλα έχουν σχεδιαστεί για να μαθαίνουν από τη διαδοχική φύση των δεδομένων, καθιστώντας τα ικανά στο χειρισμό χρονοσειρών με περίπλοκη χρονική δυναμική.

2.4.2 Χειρισμός θορύβου και ακραίων τιμών

Στο πεδίο της ανάλυσης χρονοσειρών, ο θόρυβος και οι ακραίες τιμές αποτελούν σημαντικές προκλήσεις που μπορούν να αλλοιώσουν την ανάλυση και να οδηγήσουν σε ανακριβείς προβλέψεις. Ο θόρυβος αναφέρεται σε τυχαίες διακυμάνσεις στα δεδομένα που δεν αντικατοπτρίζουν τα πραγματικά υποκείμενα πρότυπα. Θα μπορούσε να οφείλεται σε σφάλματα μέτρησης, σε προβλήματα συλλογής δεδομένων ή σε εγγενή μεταβλητότητα του παρατηρούμενου συστήματος. Οι ακραίες τιμές είναι σημεία δεδομένων που αποκλίνουν

σημαντικά από το υπόλοιπο σύνολο δεδομένων, συχνά λόγω ανωμαλιών ή απροσδόκητων γεγονότων.

Για την μείωση του θορύβου συνήθως εφαρμόζονται τεχνικές όπως η εξομάλυνση (χρησιμοποιώντας κινητούς μέσους ή εκθετική εξομάλυνση) ή το φιλτράρισμα, με χρήση ψηφιακών φίλτρων. Οι τεχνικές αυτές βοηθούν στην εξομάλυνση των βραχυπρόθεσμων διακυμάνσεων και στην αποκάλυψη της υποκείμενης τάσης. [11]

Για την ανίχνευση και αντιμετώπιση ακραίων τιμών εφαρμόζονται τεχνικές όπως οι στατιστικές δοκιμές (Grubbs, Q Dixon tests), μέθοδοι οπτικοποίησης ή πιο εξελιγμένοι αλγόριθμοι μηχανικής μάθησης. Μόλις εντοπιστούν, οι ακραίες τιμές μπορούν να αντιμετωπιστούν με αφαίρεση, προσαρμογή ή με τη χρήση ισχυρών στατιστικών μεθόδων που είναι λιγότερο ευαίσθητες στις ακραίες τιμές.[8]

Στο πλαίσιο της πρόβλεψης φορτίου, ο χειρισμός του θορύβου και των ακραίων τιμών είναι απαραίτητος για τη διασφάλιση της ακρίβειας και της αξιοπιστίας των προβλέψεων. Τα δεδομένα κατανάλωσης ενέργειας μπορεί να υπόκεινται σε θόρυβο λόγω σφαλμάτων αισθητήρων ή προβλημάτων μετάδοσης δεδομένων. Οι ακραίες τιμές μπορεί να οφείλονται σε ασυνήθιστα γεγονότα, όπως διακοπές ρεύματος, ξαφνικές αλλαγές του καιρού ή μοναδικές συμπεριφορές των καταναλωτών.

2.4.3 Διαχείριση ελλιπών τιμών

Το ζήτημα των ελλιπών τιμών στα δεδομένα χρονολογικών σειρών επηρεάζει την ποιότητα και την ακρίβεια των αναλύσεων και των προβλέψεων. Οι ελλείψεις δεδομένων μπορεί να οφείλονται σε διάφορους λόγους, από σφάλματα στη συλλογή δεδομένων έως κενά στη διαβίβαση ή παραλείψεις στην καταγραφή. Οι στρατηγικές για την διαχείριση των ελλιπών τιμών ποικίλλουν. Μια συνήθης προσέγγιση είναι η μέθοδος του καταλογισμού (imputation), όπου οι ελλείπουσες τιμές αντικαθίστανται με εκτιμώμενες τιμές. Οι τεχνικές κυμαίνονται από απλές στρατηγικές, όπως η χρήση του μέσου όρου, της διάμεσου ή της τελευταίας παρατηρούμενης τιμής, έως πιο περίπλοκες μεθόδους, όπως η γραμμική παρεμβολή. Η μέθοδος της γραμμικής παρεμβολής υπολογίζει τις ελλείπουσες τιμές με βάση γραμμικά σταθμισμένους μέσους όρους των υφιστάμενων δεδομένων, λαμβάνοντας υπόψη την εγγενή δομή της χρονοσειράς. [17]

Στο πλαίσιο της πρόβλεψης φορτίου, η ακριβής πρόβλεψη εξαρτάται σε μεγάλο βαθμό από την ακεραιότητα και τη συνέχεια των δεδομένων. Η επιλεγείσα μέθοδος για το χειρισμό των ελλιπών τιμών θα πρέπει να ευθυγραμμίζεται με τη φύση και το πρότυπο των ελλιπών δεδομένων και της συνολικής χρονοσειράς. Για σποραδικές ελλείπουσες τιμές, μπορεί να αρκούν απλούστερες μέθοδοι υπολογισμού. Ωστόσο, για δεδομένα φορτίου με πιο εκτεταμένα κενά ή έντονα εποχιακά μοτίβα, προηγμένες τεχνικές, όπως ο υπολογισμός βάσει μοντέλου, είναι συχνά πιο κατάλληλες.

2.4.4 Μελέτη αυτοσυσχέτισης

Η αυτοσυσχέτιση (autocorrelation), γνωστή και ως σειριακή συσχέτιση, είναι ένα σημαντικό στατιστικό μέγεθος που ποσοτικοποιεί το βαθμό ομοιότητας μεταξύ μιας δεδομένης χρονοσειράς και μιας καθυστερημένης εκδοχής της σε διαδοχικά χρονικά διαστήματα. Αξιολογεί πόσο καλά η τρέχουσα τιμή μιας χρονοσειράς σχετίζεται με τις προηγούμενες τιμές της, παρέχοντας πληροφορίες για τα υποκείμενα πρότυπα και τις δομές των δεδομένων [9].

Μαθηματικά, η αυτοσυσχέτιση για μια χρονοσειρά εκφράζεται ως η συσχέτιση μεταξύ της χρονοσειράς σε δύο χρονικές στιγμές, οι οποίες απέχουν μεταξύ τους μια δεδομένη χρονική υστέρηση, k. Για μια σειρά y_t (όπου t συμβολίζει το χρόνο), η συνάρτηση αυτοσυσχέτισης (ACF) για υστέρηση k δίνεται από τον τύπο:

$$\rho_k = \frac{\sum_{t=1}^{N-k}(y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^{N}(y_t - \bar{y})^2}$$

Οπου ρ_k είναι ο συντελεστής αυτοσυσχέτισης στην υστέρηση k, y_t είναι η τιμή της χρονοσειράς τη χρονική στιγμή t, $\bar{y}$ είναι ο μέσος όρος της χρονοσειράς και N είναι ο συνολικός αριθμός των παρατηρήσεων.

Ο συντελεστής αυτοσυσχέτισης κυμαίνεται από -1 έως 1, όπου:

- Το 1 υποδηλώνει τέλεια θετική συσχέτιση: Καθώς αυξάνεται (ή μειώνεται) η χρονοσειρά, αυξάνεται (ή μειώνεται) και η χρονοσειρά με χρονική υστέρηση.

- Το -1 υποδηλώνει τέλεια αρνητική συσχέτιση: Καθώς η χρονοσειρά αυξάνεται (ή μειώνεται), η καθυστερημένη σειρά μειώνεται (ή αυξάνεται) και το αντίστροφο.
- Το 0 υποδηλώνει μηδενική συσχέτιση: η χρονοσειρά και η καθυστερημένη εκδοχή της δεν επηρεάζουν η μία την άλλη.

Η κατανόηση της αυτοσυσχέτισης είναι μεγάλης σημασίας για την ανάλυση χρονοσειρών καθώς βοηθά στον εντοπισμό της παρουσίας τάσεων ή εποχιακών προτύπων και στην επιλογή των κατάλληλων παραμέτρων για μοντέλα χρονοσειρών. Για παράδειγμα, στην περίπτωση μας έπαιξε καθοριστικό ρόλο στην επιλογή των οριζόντων πρόβλεψης του μοντέλου μας.

2.4.5 Κλιμάκωση και κανονικοποίηση δεδομένων

Η κλιμάκωση και η κανονικοποίηση των δεδομένων είναι κρίσιμα βήματα προεπεξεργασίας στην ανάλυση χρονοσειρών, ιδίως όταν εργάζονται με αλγορίθμους ευαίσθητους στην κλίμακα των δεδομένων, όπως τα νευρωνικά δίκτυα. Οι τεχνικές αυτές περιλαμβάνουν τον μετασχηματισμό των δεδομένων σε ένα συγκεκριμένο εύρος ή κατανομή, ενισχύοντας την αποτελεσματικότητα και τη σταθερότητα της εκπαίδευσης του μοντέλου. Υπάρχουν αρκετές μέθοδοι κλιμάκωσης δεδομένων όπως: [18]

- Standard Scaler (Z-score Normalization): Αυτή η μέθοδος μετασχηματίζει κάθε χαρακτηριστικό ώστε να έχει μέση τιμή 0 και τυπική απόκλιση 1. Ο τύπος που χρησιμοποιείται είναι ο εξής:

  $$Z = \frac{X - \mu}{\sigma}$$

  όπου μ είναι ο μέσος όρος και σ είναι η τυπική απόκλιση.
  Αυτή η κλίμακα είναι χρήσιμη όταν τα δεδομένα ακολουθούν κατανομή Gauss. [19]

- Robust Scaler: Αυτή η μέθοδος είναι ιδανική για δεδομένα με ακραίες τιμές, καθώς επιδιώκει να μετριάσει τις επιδράσεις τους κεντράροντας τα δεδομένα γύρω από τη διάμεσο (δεύτερο τεταρτημόριο του x, Q2(x)) και κλιμακώνοντάς τα σύμφωνα με το ενδοτεταρτημοριακό εύρος, το οποίο είναι το μέγεθος της διαφοράς μεταξύ του πρώτου τεταρτημόριου Q1(x) και του τρίτου τεταρτημόριου Q3(x) του x όπως φαίνεται στην παρακάτω εξίσωση:[19]

$$x'i = \frac{xi - Q2(x)}{Q3(x) - Q1(x)}$$

- Normalization (L1 & L2): Αυτή η μέθοδος κλιμακώνει τα μεμονωμένα δείγματα ώστε να έχουν μοναδιαία νόρμα. Η κανονικοποίηση L1, καθιστά το άθροισμα των απόλυτων τιμών κάθε γραμμής ίσο με 1, ενώ η κανονικοποίηση L2 εξασφαλίζει ότι το άθροισμα των τετραγώνων είναι 1. [20]

- MinMaxScaler: Η συγκεκριμένη μέθοδος είναι από τις πιο συνηθισμένες και είναι αυτή που χρησιμοποιήθηκε σε αυτήν την εργασία. Η μέθοδος μετασχηματίζει γραμμικά κάθε χαρακτηριστικό σε ένα δεδομένο εύρος, συνήθως μεταξύ 0 και 1. Ο μετασχηματισμός ακολουθεί τον τύπο:

$$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

  Οπου X είναι η αρχική τιμή, $X_{\min}$ και $X_{\max}$ η ελάχιστη και η μέγιστη τιμή και X_{scaled} η κλιμακούμενη τιμή. [19]

Η χρήση της κλίμακας MinMax στα νευρωνικά δίκτυα, ιδίως στο πλαίσιο της βελτιστοποίησης με κατάβαση κλίσης (Gradient Descent [21]), είναι πολύ σημαντική. Με την κλιμάκωση όλων των εισόδων σε ένα ομοιόμορφο εύρος (-1 έως 1), ο κλιμακωτής MinMax διασφαλίζει ότι τα χαρακτηριστικά συμβάλλουν εξίσου στην εκπαίδευση του μοντέλου, αποφεύγοντας τις προκαταλήψεις προς τα χαρακτηριστικά υψηλού μεγέθους. Οταν τα χαρακτηριστικά βρίσκονται σε παρόμοιες κλίμακες, η κάθοδος κλίσης μπορεί να κινηθεί ομαλά προς το ελάχιστο και ο ρυθμός μάθησης μπορεί να εφαρμοστεί ομοιόμορφα σε όλα τα χαρακτηριστικά. [18]

3 Τεχνητή Νοημοσύνη και Νευρωνικά Δίκτυα

Αυτή η ενότητα εξετάζει τις θεμελιώδεις αρχές της Τεχνητής Νοημοσύνης (ΤΝ) και τον καθοριστικό ρόλο των Νευρωνικών Δικτύων στην καινοτομία σε διάφορους τομείς. Σκιαγραφείται η πορεία της ΤΝ από την εννοιολογική της γέννηση έως την ανάπτυξη εξελιγμένων αρχιτεκτονικών νευρωνικών δικτύων. Μέσω αυτής της διερεύνησης, στοχεύουμε να αναδείξουμε πώς αυτές οι βασικές τεχνολογίες όχι μόνο αποτελούν τη ραχοκοκαλιά των δυνατοτήτων επίλυσης προβλημάτων της ΤΝ, αλλά και ανοίγουν το δρόμο για προηγμένες εφαρμογές σε τομείς που απαιτούν σύνθετη ανάλυση και πρόβλεψη δεδομένων όπως την διαχείριση ενέργειας.

3.1 Εισαγωγή στην Τεχνητή Νοημοσύνη

Η Τεχνητή Νοημοσύνη (Artificial Intelligence – AI) αποτελεί ένα επιστημονικό πεδίο με στόχο τη δημιουργία μηχανών ικανών για ευφυή συμπεριφορά που να αντικατοπτρίζει την ανθρώπινη νόηση. Ο τομέας της Τεχνητής Νοημοσύνης ενσωματώνει διάφορους επιστημονικούς κλάδους, όπως η επιστήμη των υπολογιστών, η γνωστική επιστήμη, η γλωσσολογία, η ψυχολογία κ.α., για να μπορέσουν οι μηχανές να εκτελέσουν εργασίες όπως η επίλυση προβλημάτων, η λήψη αποφάσεων και η κατανόηση της γλώσσας. Βασικός σκοπός της είναι να κατανοήσει και να αυτοματοποιήσει τις διανοητικές εργασίες με απώτερο στόχο να επιτύχει μια μηχανική μίμηση της ανθρώπινης νοημοσύνης. [22]

Οι πρόσφατες εξελίξεις στον τομέα της τεχνητής νοημοσύνης καθορίστηκαν από την εκθετική αύξηση της υπολογιστικής ισχύος, τη διαθεσιμότητα μεγάλων συνόλων δεδομένων (big data) και την πρόοδο των αλγορίθμων, ιδίως εκείνων που σχετίζονται με τα νευρωνικά δίκτυα και τη μηχανική μάθηση. Οι τεχνολογίες αυτές όχι μόνο έχουν διευρύνει το πεδίο των εργασιών που μπορούν να αυτοματοποιηθούν, αλλά έχουν επίσης βελτιώσει την αποτελεσματικότητα και την ακρίβεια αυτών των συστημάτων. [23]

Στο πλαίσιο της διαχείρισης της ενέργειας και της πρόβλεψης φορτίου, οι τεχνολογίες AI, ιδίως τα μοντέλα μηχανικής μάθησης, έχουν καθιερωθεί ως εργαλεία ζωτικής σημασίας καθώς αναλύουν ιστορικά δεδομένα κατανάλωσης, προβλέπουν τη μελλοντική ζήτηση και βοηθούν στη

βελτιστοποίηση των λειτουργιών του δικτύου. Η ικανότητά τους αυτή είναι κρίσιμη για τη μετάβαση προς πιο βιώσιμα και αποδοτικά ενεργειακά συστήματα.

3.2 Βασικά στοιχεία μηχανικής μάθησης

Η μηχανική μάθηση (Machine Learning - ML) είναι ένας κλάδος της Τεχνητής Νοημοσύνης που διδάσκει στους υπολογιστές πώς να μαθαίνουν και να λαμβάνουν αποφάσεις από δεδομένα χωρίς να προγραμματίζονται ρητά για κάθε εργασία. Η μηχανική μάθηση, χρησιμοποιεί δεδομένα για να βρίσκει αυτόματα μοτίβα και ιδέες, επιτρέποντας στους υπολογιστές να εκτελούν εργασίες όπως η αναγνώριση ομιλίας, η αναγνώριση εικόνων ή η πρόβλεψη τάσεων. [24]

Ο T. Mitchell έδωσε το 1997 τον εξής ορισμό για την Μηχανική Μάθηση *"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E"* [25]. Η εμπειρία E, η εργασία T, και η μετρική απόδοσης P ποικίλλουν. Στην περίπτωση της παρούσας εργασίας, η εμπειρία αποτελείται από τα παλαιότερα δεδομένα μετρήσεων φορτίου, η εργασία είναι η πρόβλεψη των μελλοντικών δεδομένων κατανάλωσης και η μετρική σχετίζεται με το ποσοστό σφάλματος των μετρήσεων σε σχέση με τις πραγματικές τιμές.

3.2.1 Τύποι Μηχανικής Μάθησης

Η μηχανική μάθηση μπορεί να κατηγοριοποιηθεί σε γενικές γραμμές σε τρεις κύριους τύπους, καθένας από τους οποίους έχει τη δική του μοναδική προσέγγιση και τομείς εφαρμογής: [24]

1) **Επιβλεπόμενη μάθηση**: Αυτή είναι η πιο διαδεδομένη μορφή μηχανικής μάθησης, όπου ο αλγόριθμος μαθαίνει από ένα σύνολο δεδομένων με ετικέτες. Περιλαμβάνει την αντιστοίχιση των δεδομένων εισόδου σε γνωστές ετικέτες εξόδου κατά τη διάρκεια της

εκπαίδευσης, επιτρέποντας στο μοντέλο να προβλέπει την έξοδο για αόρατα δεδομένα. Η μάθηση με επίβλεψη χρησιμοποιείται ευρέως σε εφαρμογές όπως η ανίχνευση ανεπιθύμητων μηνυμάτων, η ανάλυση συναισθήματος και, κυρίως, η πρόβλεψη παραγωγής, όπου είναι γνωστά τα ιστορικά δεδομένα (είσοδος) και η ζήτηση (έξοδος).

2) **Μη-επιβλεπόμενη μάθηση**: Σε αυτήν την μέθοδο, ο αλγόριθμος εντοπίζει μοτίβα και δομές από δεδομένα χωρίς ετικέτες. Χρησιμοποιείται κυρίως σε εργασίες ομαδοποίησης, μείωσης διαστάσεων και συσχέτισης, βοηθώντας στην αποκάλυψη κρυφών μοτίβων ή ομαδοποιήσεων στα δεδομένα. Τεχνικές όπως η ομαδοποίηση k-means, η ανάλυση κύριων συνιστωσών (PCA) και οι αυτοκωδικοποιητές είναι διαδεδομένες σε αυτή την κατηγορία.

3) **Ενισχυτική μάθηση:** Η συγκεκριμένη μέθοδος επικεντρώνεται στη λήψη μιας ακολουθίας αποφάσεων. Ο αλγόριθμος μαθαίνει να επιτυγχάνει έναν στόχο σε ένα αβέβαιο, ενδεχομένως πολύπλοκο περιβάλλον. Μέσω δοκιμών και σφαλμάτων, ανακαλύπτει ποιες ενέργειες αποφέρουν τη μεγαλύτερη ανταμοιβή. Οι εφαρμογές περιλαμβάνουν τη ρομποτική, τα παιχνίδια και την πλοήγηση, με αλγορίθμους όπως ο Q-learning και οι μέθοδοι πολιτικής κλίσης να έχουν κεντρικό ρόλο στην επιτυχία της.

3.2.2 Η μηχανική μάθηση στο πλαίσιο της πρόβλεψης φορτίου

Η σημασία της μηχανικής μάθησης στον τομέα τον προβλέψεων είναι πολύ μεγάλη. Αξιοποιώντας ιστορικά δεδομένα και μαθαίνοντας από πρότυπα, τα μοντέλα ML μπορούν να προβλέψουν μελλοντικές τάσεις, απαιτήσεις και συμπεριφορές με αξιοσημείωτη ακρίβεια. Ειδικότερα, στην πρόβλεψη παραγωγής τα μοντέλα αναλύουν δεδομένα κατανάλωσης του παρελθόντος, λαμβάνοντας υπόψη παράγοντες όπως οι καιρικές συνθήκες, η ώρα της ημέρας, οι αργίες και άλλα, για να προβλέψουν τη μελλοντική ζήτηση ενέργειας. Αυτή η ικανότητα επεξεργασίας και εκμάθησης από τεράστια σύνολα δεδομένων επιτρέπει συχνά ακριβέστερες και αποτελεσματικότερες προβλέψεις από τις παραδοσιακές στατιστικές μεθόδους, συμβάλλοντας στη βελτιστοποιημένη διαχείριση και τον προγραμματισμό της ενέργειας.

Διάφοροι αλγόριθμοι ML είναι ιδιαίτερα κατάλληλοι για την πρόβλεψη παραγωγής, ο καθένας με τα δυνατά του σημεία:

- **Γραμμική παλινδρόμηση**: Χρήσιμη για τον εντοπισμό γραμμικών σχέσεων μεταξύ μεταβλητών. Αποτελεί συχνά σημείο εκκίνησης για τα μοντέλα πρόβλεψης. [26]

- **Δέντρα αποφάσεων και τυχαία δάση**: Αυτά είναι ισχυρά για τη σύλληψη μη γραμμικών σχέσεων και αλληλεπιδράσεων μεταξύ πολλαπλών μεταβλητών, προσφέροντας πιο διαφοροποιημένες προβλέψεις από τα γραμμικά μοντέλα. [27]

- **Gradient Boosting Machines (GBM)**: Οι GBMs παρέχουν εξαιρετικά ακριβείς προβλέψεις διορθώνοντας διαδοχικά τα σφάλματα από προηγούμενα μοντέλα, καθιστώντας τα αποτελεσματικά για σύνθετα σύνολα δεδομένων. [28]

- **Νευρωνικά δίκτυα**: Τα νευρωνικά δίκτυα, και ιδίως τα LSTMs υπερέχουν στη σύλληψη μοτίβων σε δεδομένα χρονοσειρών, καθιστώντας τα ιδανικά για τις προβλέψεις όπου η χρονική δυναμική είναι κρίσιμη. [29] Αυτός είναι και ο βασικός λόγος που επιλέχθηκαν για την υλοποίηση της παρούσας εργασίας.

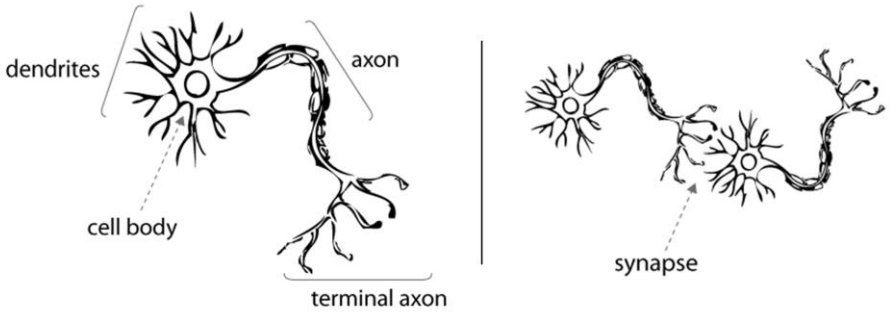
3.3 Νευρωνικά Δίκτυα

3.3.1 Νευρώνες & Τεχνητά Νευρωνικά Δίκτυα

Τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks - ANN) πιο γνωστά ως Νευρωνικά Δίκτυα (Neural Networks – NN) είναι υπολογιστικά μοντέλα εμπνευσμένα από τη δομή και τη λειτουργία του ανθρώπινου εγκεφάλου, ο οποίος έχει σχεδιαστεί για να αναγνωρίζει μοτίβα και να επιλύει σύνθετα προβλήματα.

Το ανθρώπινο νευρικό σύστημα περιέχει κύτταρα, τα οποία αναφέρονται ως νευρώνες. Οι νευρώνες συνδέονται μεταξύ τους με τη χρήση αξόνων και δενδριτών και οι περιοχές σύνδεσης μεταξύ αξόνων και δενδριτών ονομάζονται συνάψεις. Οι νευρώνες, δέχονται εξωτερικά

ερεθίσματα και τα μεταβιβάζουν μέσω των συνάψεων η ισχύς των οποίων αλλάζει ανάλογα το ερέθισμα. Αυτές οι αλλαγές στις συνάψεις είναι και ο βασικός μηχανισμός μάθησης στους ζωντανούς οργανισμούς. [30]



***Εικόνα 3.1**: Βιολογικοί νευρώνες και συνάψεις [31]*

Στον πυρήνα τους, τα NN αποτελούνται από επίπεδα τεχνητών νευρώνων, που μιμούνται την συμπεριφορά των βιολογικών σε μια απλοποιημένη μαθηματική εκδοχή. Αυτοί οι νευρώνες συνδέονται μεταξύ τους, με κάθε σύνδεση (σύναψη) να φέρει ένα βάρος που καθορίζει την ισχύ της επιρροής ενός νευρώνα σε έναν άλλο. Ο πρωταρχικός ρόλος ενός νευρώνα είναι να δέχεται εισόδους, να τις επεξεργάζεται και να παράγει μια έξοδο που μπορεί να μεταβιβαστεί σε άλλους νευρώνες ή να χρησιμεύσει ως μέρος της συνολικής εξόδου του δικτύου. [30]

Στην πράξη, η λειτουργία ενός νευρώνα μπορεί να περιγραφεί ως εξής: [32]

1) **Δέχεται εισόδους $(\boldsymbol{x^1}, \boldsymbol{x^2}, \ldots, \boldsymbol{x_n})$:** Κάθε νευρώνας λαμβάνει πολλαπλές εισόδους, οι οποίες είναι είτε ακατέργαστα σημεία δεδομένων από το σύνολο χαρακτηριστικών είτε οι έξοδοι από τους νευρώνες του προηγούμενου επιπέδου. Αυτές οι είσοδοι αντιπροσωπεύουν διαφορετικά χαρακτηριστικά των υπό επεξεργασία δεδομένων.

2) **Δέχεται βάρη $(\boldsymbol{w^1}, \boldsymbol{w^2}, \ldots, \boldsymbol{w_n})$:** Σε κάθε είσοδο στον νευρώνα αποδίδεται ένα βάρος που υποδηλώνει τη σημασία της. Τα βάρη είναι ρυθμιζόμενες παράμετροι που το νευρωνικό δίκτυο μαθαίνει κατά τη διαδικασία εκπαίδευσης. Αρχικά, τα βάρη αυτά ορίζονται σε τυχαίες τιμές και προσαρμόζονται μέσω της οπισθοδιάδοσης (backpropagation) καθώς το δίκτυο μαθαίνει.

3) **Δέχεται έναν όρο προκατάληψης (*b*):** Ο όρος προκατάληψης είναι μια πρόσθετη παράμετρος που χρησιμοποιείται στον υπολογισμό του νευρώνα, επιτρέποντας τη μετατόπιση της συνάρτησης ενεργοποίησης προς τα αριστερά ή προς τα δεξιά, η οποία μπορεί να είναι κρίσιμη για την εκμάθηση μοτίβων στα δεδομένα. Οπως και τα βάρη, η προκατάληψη μαθαίνεται κατά τη διάρκεια της εκπαίδευσης.

4) **Υπολογίζει το σταθμισμένο άθροισμα (z):** Ο νευρώνας υπολογίζει το σταθμισμένο άθροισμα των εισόδων του, το οποίο είναι το άθροισμα κάθε εισόδου πολλαπλασιασμένο με το αντίστοιχο βάρος, συν τον όρο προκατάληψης.

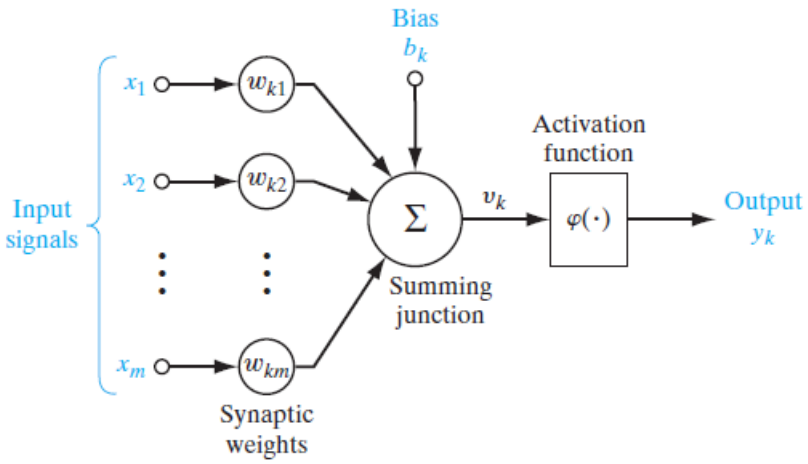
$$z = \sum_{i=1}^{n} w_i x_i + b$$

5) **Περνάει το σταθμισμένο άθροισμα μέσα από μια συνάρτηση ενεργοποίησης (*f*):** Αυτό το βήμα εισάγει μη γραμμικότητα στο μοντέλο, επιτρέποντας στο δίκτυο να μάθει πολύπλοκα μοτίβα και σχέσεις στα δεδομένα. Η επιλογή της συνάρτησης ενεργοποίησης επηρεάζει την έξοδο του νευρώνα και τη συνολική ικανότητα του δικτύου να συγκλίνει και να μαθαίνει αποτελεσματικά.

6) **Δίνει έξοδο y:** Η έξοδος y γίνεται η είσοδος στους νευρώνες του επόμενου επιπέδου ή μέρος της τελικής εξόδου του δικτύου αν βρίσκεται στο επίπεδο εξόδου.

Με βάση τα παραπάνω, η λειτουργία ενός νευρώνα μπορεί να αναπαρασταθεί μαθηματικά ως:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right)$$



Εικόνα 3.2: *Τεχνητός νευρώνας [32]*

Χωρίς συναρτήσεις ενεργοποίησης, ακόμη και τα βαθιά νευρωνικά δίκτυα με πολλά επίπεδα θα εξακολουθούσαν να λειτουργούν ως γραμμικά μοντέλα, περιορίζοντας σημαντικά την ικανότητά τους να επιλύουν πολύπλοκα προβλήματα. Οι συναρτήσεις ενεργοποίησης υπάρχουν σε διάφορες μορφές, καθεμία με συγκεκριμένα χαρακτηριστικά και εφαρμογές. Μεταξύ των πιο ευρέως χρησιμοποιούμενων συναρτήσεων ενεργοποίησης είναι η σιγμοειδής (sigmoid), η διορθωμένη γραμμική μονάδα (ReLU) και η υπερβολική εφαπτομένη (tanh), καθεμία από τις οποίες εξυπηρετεί διαφορετικούς σκοπούς στην αρχιτεκτονική των νευρωνικών δικτύων [32]. Μερικές από τις χρησιμοποιούμενες συναρτήσεις ενεργοποίησης είναι οι εξής:

- **Identity:** Η συνάρτηση ταυτότητας απλά επιστρέφει την είσοδο ως έξοδο, καθιστώντας την ως γραμμική ενεργοποίηση. Χρησιμοποιείται συνήθως στο επίπεδο εξόδου ενός μοντέλου παλινδρόμησης.

$$f(x) = x$$

- **Sign**: Η συνάρτηση προσήμου επιστρέφει -1 εάν η είσοδος είναι μικρότερη από 0, +1 εάν η είσοδος είναι μεγαλύτερη από 0 και 0 εάν η είσοδος είναι 0. Χρησιμοποιείται για εργασίες δυαδικής ταξινόμησης.

$$f(x) = \begin{cases} -1 \, if \;\; x \; < \; 0 \\ 0 \, if \; x \; = \; 0 \\ 1 \, if \; x \; > \; 0 \end{cases}$$

- **Sigmoid**: Η σιγμοειδής συνάρτηση αντιστοιχίζει οποιαδήποτε τιμή εισόδου σε μια έξοδο μεταξύ 0 και 1, καθιστώντας την χρήσιμη για προβλήματα δυαδικής ταξινόμησης.

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

- **ReLU**: Η ReLU είναι μια τμηματική συνάρτηση που εξάγει απευθείας την είσοδο εάν είναι θετική- διαφορετικά, εξάγει μηδέν. Εχει γίνει δημοφιλής λόγω της υπολογιστικής της αποδοτικότητας και της αποτελεσματικότητάς της σε μοντέλα βαθιάς μάθησης

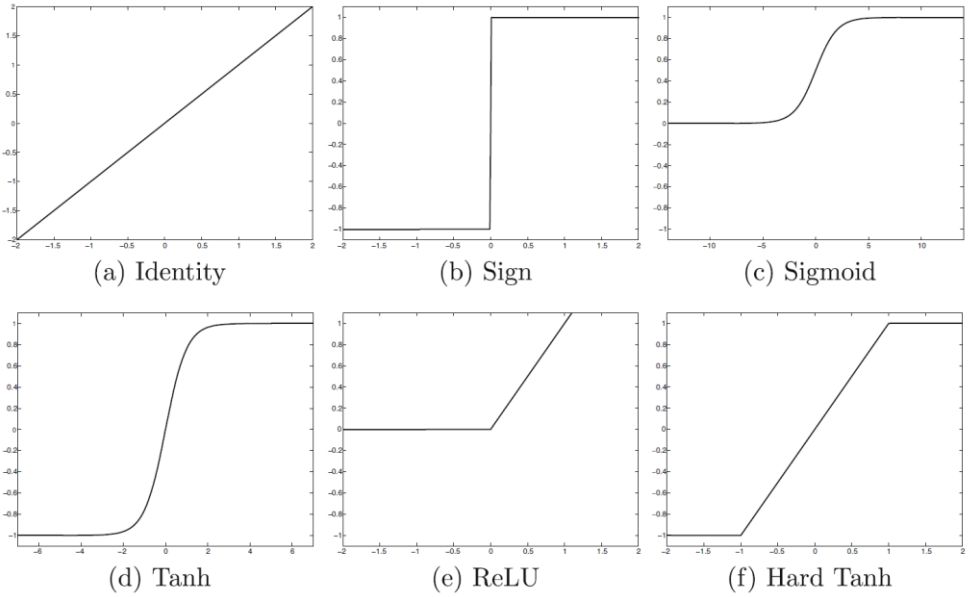
$$ReLU(x) = \max(0, x)$$

- **Συνάρτηση υπερβολικής εφαπτομένης:** Η συνάρτηση tanh δίνει τιμές μεταξύ -1 και 1. Είναι παρόμοια με τη σιγμοειδή, αλλά μπορεί να παρέχει καλύτερη απόδοση εκπαίδευσης για ορισμένα μοντέλα λόγω της συμμετρικής εξόδου της.

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

- **Συνάρτηση σκληρής εφαπτομένης:** Πρόκειται για μια γραμμική προσέγγιση της συνάρτησης tanh. Βγάζει -1 για εισόδους μικρότερες από -1, την τιμή εισόδου για εισόδους εντός του εύρους [-1, 1] και 1 για εισόδους μεγαλύτερες από 1.

$$f(x) = \begin{cases} -1 \quad if \;\; x \; < \; -1 \\ x \; if \;\; -1 \leq x \; \leq 1 \\ 1 \quad\;\; if \; x \; > \; 1 \end{cases}$$

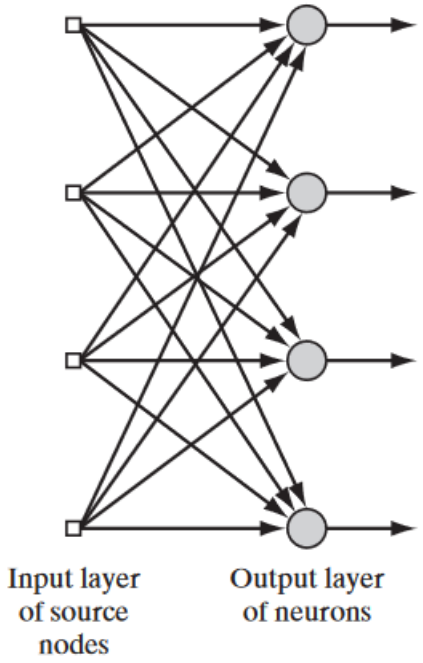


Εικόνα 3.3: *Συναρτήσεις ενεργοποίησης [30]*

3.3.2 Αρχιτεκτονική και είδη Νευρωνικών Δικτύων

Σύμφωνα με τον Haykin [32], τα νευρωνικά δίκτυα χωρίζονται σε 3 βασικές κατηγορίες με βάση τον τρόπο με τον οποίο δομούντα οι νευρώνες τους:

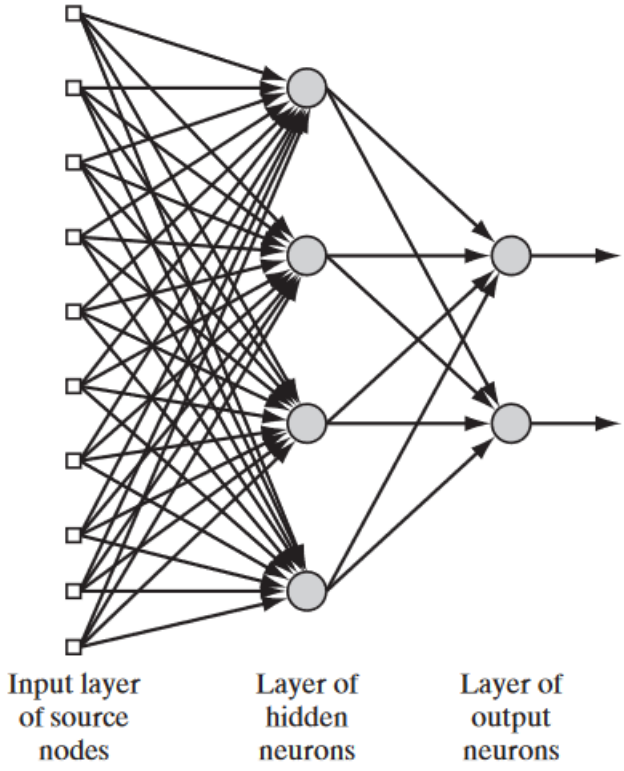
1) **Δίκτυα εμπρόσθιας τροφοδότησης ενός επιπέδου**: Σε ένα πολυεπίπεδο νευρωνικό δίκτυο, οι νευρώνες διατάσσονται σε επίπεδα. Στην πιο βασική του μορφή, αυτό το δίκτυο απαρτίζεται από ένα επίπεδο εισόδου, το οποίο αποτελείται από κόμβους πηγής και συνδέεται απευθείας με ένα επίπεδο εξόδου, το οποίο περιέχει νευρώνες που εκτελούν υπολογισμούς. Ωστόσο, η σύνδεση δεν λειτουργεί προς την αντίστροφη κατεύθυνση θέτοντας το δίκτυο ως αποκλειστικά εμπρόσθιας τροφοδότησης. Αυτή η διαμόρφωση είναι γνωστή ως δίκτυο ενός στρώματος , όπου ο χαρακτηρισμός "ενός στρώματος" αναφέρεται συγκεκριμένα στο στρώμα νευρώνων που εκτελεί υπολογισμούς, το στρώμα εξόδου. Το στρώμα εισόδου δεν λαμβάνεται υπόψη σε αυτή την ονομασία, καθώς δεν εκτελεί κανέναν υπολογισμό. Τέτοια δίκτυα χρησιμοποιούνται συχνά για απλές εργασίες αναγνώρισης προτύπων ή ταξινόμησης, όπου η σχέση μεταξύ της εισόδου και της εξόδου είναι σχετικά απλή.



Εικόνα 3.4: *Νευρωνικό δίκτυο εμπρόσθιας τροφοδοσίας με ένα επίπεδο νευρώνων [32]*

2) **Δίκτυα εμπρόσθιας τροφοδότησης πολλαπλών επιπέδων**: Η δεύτερη κατηγορία των νευρωνικών δικτύων εμπρόσθιας τροφοδότησης (Feedforward Neural Network – FNN) χαρακτηρίζεται από τη συμπερίληψη ενός ή περισσότερων κρυφών επιπέδων, τα οποία περιέχουν υπολογιστικούς κόμβους γνωστούς ως κρυφούς νευρώνες ή κρυφές μονάδες. Τα επίπεδα αυτά χαρακτηρίζονται ως "κρυφά" επειδή δεν είναι άμεσα παρατηρήσιμα από την είσοδο ή την έξοδο του δικτύου. Τα κρυφά επίπεδα σε αυτά τα δίκτυα είναι κρίσιμα για την ικανότητά τους να μαθαίνουν πολύπλοκα μοτίβα και σχέσεις στα δεδομένα. Επεξεργαζόμενα τις εισόδους μέσω διαδοχικών επιπέδων, τα πολυεπίπεδα νευρωνικά δίκτυα τροφοδότησης μπορούν να εξάγουν και να ενισχύουν σημαντικά χαρακτηριστικά και μοτίβα, καθιστώντας τα ιδιαίτερα αποτελεσματικά για εργασίες όπως η αναγνώριση

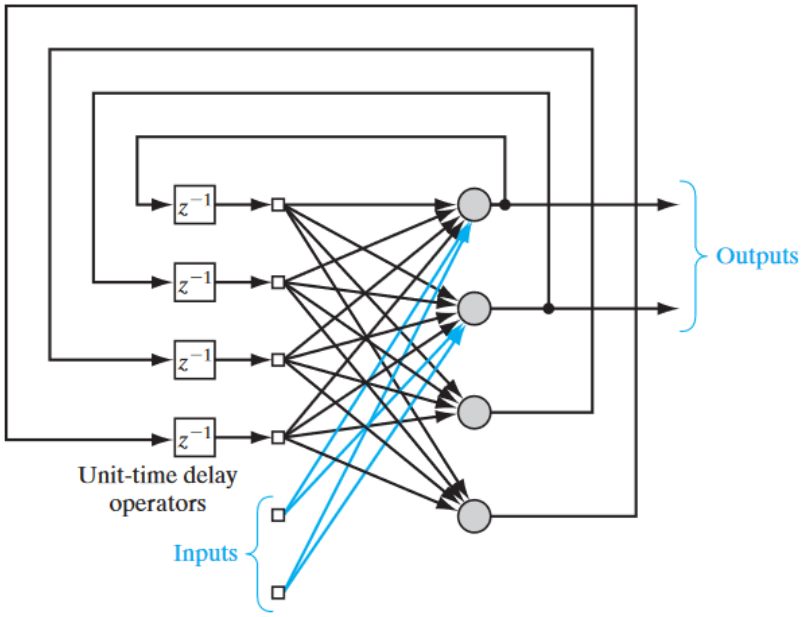
εικόνας, η αναγνώριση ομιλίας και πολλές άλλες μορφές μηχανικής μάθησης που απαιτούν τη μοντελοποίηση πολύπλοκων μη γραμμικών σχέσεων.



Εικόνα 3.5: *Πλήρως ενωμένο νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης με ένα κρυφό επίπεδο [32]*

3) Επαναλαμβανόμενα Νευρωνικά Δίκτυα: Ενα επαναλαμβανόμενο νευρωνικό δίκτυο (Recurrent Neural Network – RNN) διακρίνεται από ένα FNN στο ότι διαθέτει τουλάχιστον έναν βρόχο ανατροφοδότησης. Πιο συγκεκριμένα, σε ένα RNN, οι νευρώνες ενός επιπέδου μπορούν να είναι συνδεδεμένοι και με νευρώνες προηγούμενων επιπέδων, διαφορετικούς νευρώνες του ίδιου επιπέδου η ακόμη και με τον εαυτό τους σε έναν βρόγχο ανατροφοδότησης. Η παρουσία βρόχων ανατροφοδότησης, έχει βαθύτατο αντίκτυπο στην ικανότητα μάθησης του δικτύου και στην απόδοσή του. Επιπλέον, οι βρόχοι ανατροφοδότησης συνεπάγονται τη χρήση συγκεκριμένων κλάδων που αποτελούνται από στοιχεία μοναδιαίας χρονικής καθυστέρησης (συμβολίζονται με z^{-1}),

τα οποία οδηγούν σε μη γραμμική δυναμική συμπεριφορά [32]. Αυτή η αρχιτεκτονική επιτρέπει στα RNN να διατηρούν μια μορφή μνήμης, χρησιμοποιώντας την εσωτερική τους κατάσταση (κρυφά επίπεδα) για να επεξεργάζονται ακολουθίες εισόδων στο πέρασμα του χρόνου. Αυτό τα καθιστά ιδανικά για εργασίες που απαιτούν την κατανόηση του πλαισίου με την πάροδο του χρόνου, όπως η γλωσσική μετάφραση, η αναγνώριση ομιλίας και η πρόβλεψη χρονοσειρών.



Εικόνα 3.6: *Επαναλαμβανόμενο νευρωνικό δίκτυο με κρυφά επίπεδα [32]*

3.3.3 Βαθιά Νευρωνικά Δίκτυα

Τα διάφορα είδη νευρωνικών δικτύων έχουν έναν ακόμη διαχωρισμό. Αυτόν μεταξύ των βαθιών νευρωνικών δικτύων (Deep Neural Networks - DNN) και των αβαθών ή ρηχών νευρωνικών δικτύων (Shallow Neural Networks).

Τα DNNs αντιπροσωπεύουν μια κατηγορία προηγμένων νευρωνικών δικτύων που χαρακτηρίζονται από πολλαπλά κρυφά επίπεδα μεταξύ των επιπέδων εισόδου και εξόδου.

Αντίθετα, τα αβαθή νευρωνικά δίκτυα έχουν ένα η ελάχιστα κρυφά επίπεδα. Το βάθος των DNNs, με τα πολλαπλά επίπεδά τους, τους επιτρέπει να μαθαίνουν πολύπλοκα μοτίβα που τα αβαθή δίκτυα δεν μπορούν να μάθουν. [33]

Η βαθιά μάθηση, ένα υποσύνολο της μηχανικής μάθησης που περιλαμβάνει DNNs, αυτοματοποιεί τη διαδικασία εξαγωγής χαρακτηριστικών υψηλού επιπέδου από τα δεδομένα, ξεχωρίζοντας από τις παραδοσιακές προσεγγίσεις μηχανικής μάθησης που συχνά βασίζονται σε χειροκίνητη εξαγωγή χαρακτηριστικών. [33]

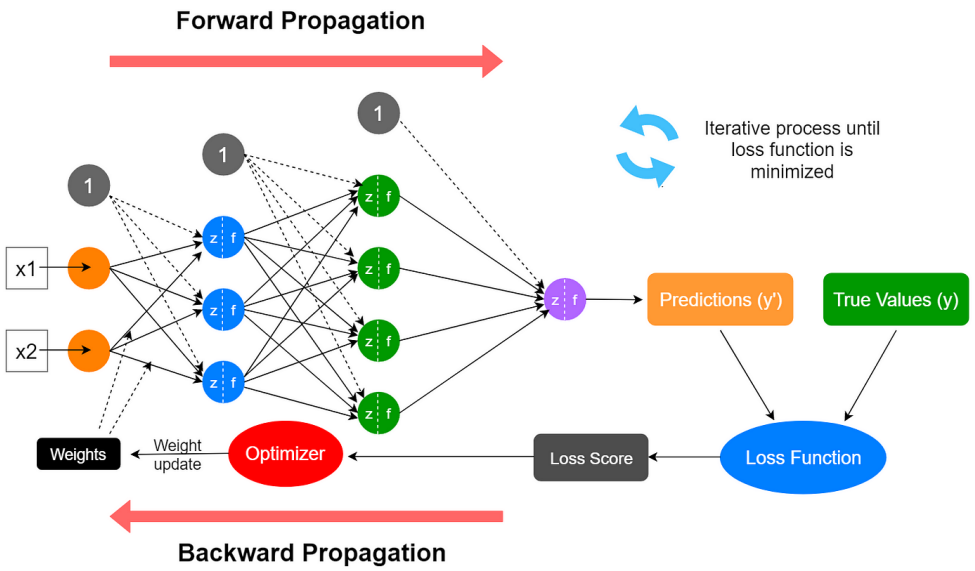
Κατά συνέπεια, τα DNNs συνήθως υπερτερούν των μη βαθιών δικτύων σε εργασίες που περιλαμβάνουν μεγάλο όγκο δεδομένων και πολύπλοκη επίλυση προβλημάτων, αν και απαιτούν σημαντικά περισσότερους υπολογιστικούς πόρους και δεδομένα για την εκπαίδευση.

3.3.4 Εκπαίδευση Νευρωνικών Δικτύων

Η εκπαίδευση ενός NN είναι μια θεμελιώδης πτυχή της ανάπτυξης μοντέλων μηχανικής μάθησης ικανών για εργασίες όπως η ταξινόμηση, η παλινδρόμηση και η πρόβλεψη. Η εκπαίδευση περιλαμβάνει την προσαρμογή των βαρών και των προκαταλήψεών του ώστε να ελαχιστοποιηθεί η διαφορά μεταξύ της προβλεπόμενης εξόδου και των πραγματικών τιμών-στόχων. Η διαδικασία αυτή είναι επαναληπτική και περιλαμβάνει διάφορα βασικά βήματα τα οποία θα περιγράψουμε παρακάτω, υπογραμμίζοντας τις σημαντικές έννοιες. [34]

1) **Αρχικοποίηση**: Κατά την αρχική φάση, τα βάρη και οι προκαταλήψεις του δικτύου λαμβάνουν τυχαίες τιμές. Αυτή η τυχαιότητα είναι σημαντικός παράγοντας για τη διάσπαση της συμμετρίας κατά τη διαδικασία μάθησης, διασφαλίζοντας ότι οι νευρώνες στο ίδιο επίπεδο μπορούν να μάθουν διαφορετικά μοτίβα.
2) **Εμπρόσθια διάδοση (forward propagation)**: Στην εμπρόσθια διάδοση, τα δεδομένα εισόδου τροφοδοτούνται στο δίκτυο, περνώντας διαδοχικά από κάθε επίπεδο. Σε κάθε επίπεδο, οι νευρώνες εκτελούν ένα σταθμισμένο άθροισμα των εισόδων, προσθέτουν μια προκατάληψη και εφαρμόζουν μια συνάρτηση ενεργοποίησης. Το αποτέλεσμα περνάει στο επόμενο επίπεδο, καταλήγοντας στην τελική έξοδο του δικτύου, όπως περιγράψαμε στο κεφάλαιο 2.3.1.

3) **Υπολογισμός απωλειών (loss calculation):** Μόλις το δίκτυο παράγει μια έξοδο, η διαφορά μεταξύ αυτής της εξόδου και των πραγματικών τιμών στόχου υπολογίζεται με τη χρήση μιας συνάρτησης απώλειας (loss function). Οι συνήθεις συναρτήσεις απωλειών περιλαμβάνουν το μέσο τετραγωνικό σφάλμα (MSE) για εργασίες παλινδρόμησης και την απώλεια διασταυρούμενης εντροπίας (Cross-Entropy Loss) για εργασίες ταξινόμησης. [35]
4) **Οπίσθια Διάδωση (backpropagation):** Η οπίσθια διάδοση είναι η διαδικασία υπολογισμού της κλίσης (gradient) της συνάρτησης απωλειών σε σχέση με κάθε βάρος και προκατάληψη στο δίκτυο, ξεκινώντας από το επίπεδο εξόδου και προχωρώντας προς τα πίσω μέσω του δικτύου. Αυτό το βήμα χρησιμοποιεί τον κανόνα της αλυσίδας του λογισμού για τον αποτελεσματικό υπολογισμό αυτών των κλίσεων. [30]
5) **Ενημέρωση Βάρους:** Με τον υπολογισμό των κλίσεων, τα βάρη και οι προκαταλήψεις ενημερώνονται για τη μείωση της απώλειας. Αυτή η ενημέρωση γίνεται συνήθως με τη χρήση ενός αλγορίθμου βελτιστοποίησης, όπως ο αλγόριθμος Gradient Descent [21] ή μία από τις παραλλαγές του (π.χ. Stochastic Gradient Descent, Adam) [36]. Το μέγεθος του βήματος που γίνεται προς την κατεύθυνση της αρνητικής κλίσης καθορίζεται από μια υπερπαράμετρο που ονομάζεται ρυθμός μάθησης.
6) **Επανάληψη:** Τα βήματα από την προς τα εμπρόσθια διάδοση έως την ενημέρωση των βαρών επαναλαμβάνονται σε πολλές επαναλήψεις, με το δίκτυο να επεξεργάζεται σύνολα δεδομένων (εποχές). Με κάθε επανάληψη, τα βάρη και οι προκαταλήψεις του δικτύου προσαρμόζονται ώστε να ελαχιστοποιείται η απώλεια, βελτιώνοντας την ακρίβεια και την προβλεπτική απόδοση του δικτύου.



Εικόνα 3.7: *Επισκόπηση της διαδικασίας εκπαίδευσης ενός νευρωνικού δικτύου [37]*

Δύο από τις βασικότερες προκλήσεις που συναντά κανείς κατά την εκπαίδευση ενός NN είναι τα προβλήματα της εξαφανιζόμενης κλίσης (vanishing gradient) και της εκρηγνυόμενης κλίσης (exploding gradient). [38]

Το πρόβλημα της **εξαφανιζόμενης κλίσης** εμφανίζεται όταν οι κλίσεις της συνάρτησης απώλειας πλησιάζουν το μηδέν καθώς διαδίδονται προς τα πίσω μέσω των επιπέδων του δικτύου κατά τη διάρκεια της εκπαίδευσης. Ως αποτέλεσμα, τα βάρη στα αρχικά στρώματα του δικτύου λαμβάνουν πολύ μικρές ενημερώσεις, οδηγώντας σε αργή ή στάσιμη πρόοδο μάθησης. Το πρόβλημα αυτό είναι ιδιαίτερα έντονο σε δίκτυα με πολλά επίπεδα και μπορεί να δυσχεράνει την αποτελεσματική εκπαίδευση βαθιών δικτύων. [39]

Αντίθετα, το πρόβλημα της **εκρηγνυόμενης κλίσης** συμβαίνει όταν οι κλίσεις της συνάρτησης απώλειας αυξάνονται εκθετικά καθώς διαδίδονται προς τα πίσω στο δίκτυο. Αυτό μπορεί να οδηγήσει στην ενημέρωση των βαρών με υπερβολικά μεγάλες τιμές, με αποτέλεσμα η διαδικασία μάθησης του δικτύου να γίνεται ασταθής και οι παράμετροι του μοντέλου να αποκλίνουν, καθιστώντας αδύνατη τη σύγκλιση του δικτύου σε μια λύση. [38]

Και τα δύο προβλήματα επηρεάζονται σε μεγάλο βαθμό από την επιλογή των συναρτήσεων ενεργοποίησης, τις μεθόδους αρχικοποίησης του δικτύου και την αρχιτεκτονική του ίδιου του

NN. Τεχνικές όπως η αποκοπή κλίσης, η κανονικοποίηση βάρους και η χρήση συγκεκριμένων τύπων στρωμάτων (όπως οι μονάδες LSTM στα αναδρομικά νευρωνικά δίκτυα) ή συναρτήσεων ενεργοποίησης (όπως η ReLU) έχουν αναπτυχθεί για να μετριάσουν αυτά τα προβλήματα και να διευκολύνουν την εκπαίδευση βαθιών νευρωνικών δικτύων.[39]

3.3.5 Δίκτυα Μακράς Βραχυπρόθεσμης Μνήμης

Τα δίκτυα μακράς βραχυπρόθεσμης μνήμης (Long Short-Term Memory - LSTM) είναι μια ειδική κατηγορία επαναλαμβανόμενων νευρωνικών δικτύων που έχουν σχεδιαστεί για να αντιμετωπίσουν τους περιορισμούς των παραδοσιακών RNNs, ιδίως τη δυσκολία τους να μάθουν μακροχρόνιες εξαρτήσεις σε δεδομένα ακολουθίας λόγω των προβλημάτων exploding και vanishing gradient που προαναφέραμε [30]. Παρουσιάστηκαν πρώτη φορά από τους Hochreiter και Schmidhuber το 1997 [40], και είναι θεμελιώδους σημασίας στον τομέα της μηχανικής μάθησης για εργασίες που περιλαμβάνουν ακολουθίες, όπως η πρόβλεψη χρονοσειρών, η επεξεργασία φυσικής γλώσσας και η αναγνώριση ομιλίας.

Η βασική καινοτομία των δικτύων LSTM είναι η ικανότητά τους να διατηρούν μακροπρόθεσμη μνήμη μέσω ενός σύνθετου συστήματος πυλών: την πύλη εισόδου, την πύλη λήθης και την πύλη εξόδου. Αυτές οι πύλες αποφασίζουν συλλογικά ποιες πληροφορίες θα διατηρηθούν, θα απορριφθούν και θα περάσουν στην έξοδο σε κάθε βήμα της ακολουθίας, επιτρέποντας στα LSTM να διατηρούν πληροφορίες για μεγάλα χρονικά διαστήματα. [40]

Τα δομικά στοιχεία των δικτύων LSTM είναι τα κύτταρα (αντί για νευρώνες) LSTM . Ενα κύτταρο LSTM αποτελείται από την κατάσταση του κυττάρου **(cell state)** $\boldsymbol{c}$, την κρυφή κατάσταση **(hidden state)** $\boldsymbol{h}$, το βήμα ενημέρωσης (**update step**) $\boldsymbol{g}$, και τις 3 πύλες: εισόδου **(input gate)** $\boldsymbol{i}$, λήθης **(forget gate)** $\boldsymbol{f}$, και εξόδου **(output gate)** $\boldsymbol{o}$ [41].

Πιο αναλυτικά:

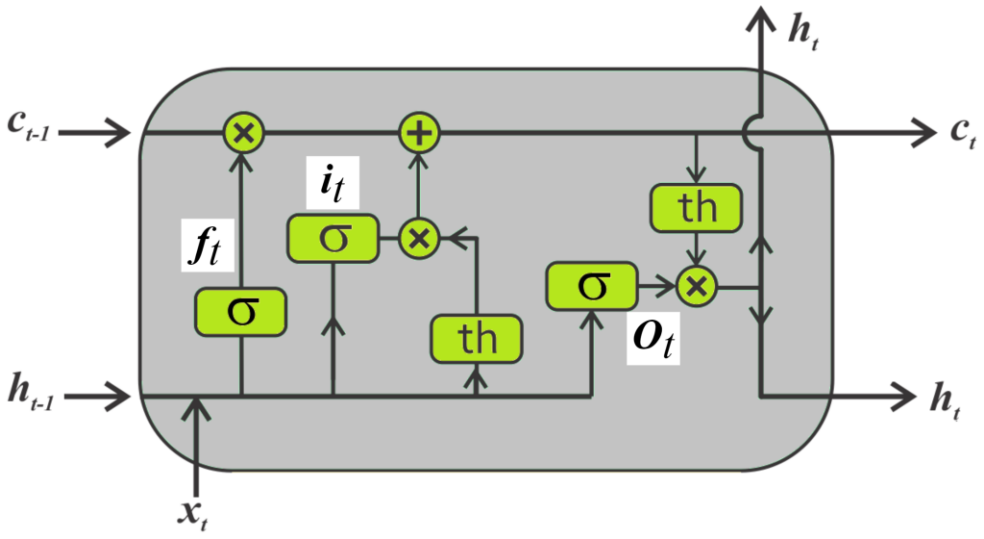
- **Η πύλη εισόδου:** Αποφασίζει ποια νέα πληροφορία θα αποθηκευτεί στην κατάσταση του κυττάρου
- **Η πύλη λήθης:** Καθορίζει ποιες πληροφορίες από την κατάσταση του κυττάρου δεν χρειάζονται πλέον και μπορούν να απορριφθούν.

- **Η πύλη εξόδου:** Καθορίζει την επόμενη κρυφή κατάσταση, η οποία περιέχει πληροφορίες για τις προηγούμενες εισόδους.

Ο υπολογισμός του κυττάρου LSTM τη χρονική στιγμή t, για την είσοδο x, δίνεται ως εξής: [42]

$$i_{[t]} = \sigma\left(W_{xi}x_{[t]} + b_{xi} + W_{hi}h_{[t-1]} + b_{hi}\right)$$
$$f_{[t]} = \sigma\left(W_{xf}x_{[t]} + b_{xf} + W_{hf}h_{[t-1]} + b_{hf}\right)$$
$$g_{[t]} = \tanh\left(W_{xg}x_{[t]} + b_{xg} + W_{hg}h_{[t-1]} + b_{hg}\right)$$
$$g_{[t]} = \left(W_{xo}x_{[t]} + b_{xo} + W_{ho}h_{[t-1]} + b_{ho}\right)$$
$$g_{[t]} = f_t \odot c_{[t-1]} + i_t \odot g_t$$
$$h_{[t]} = o_t \odot \tanh\left(c_{[t-1]}\right)$$

Οπου: το σ είναι η σιγμοειδής συνάρτηση ενεργοποίησης, το $tanh$ αντιπροσωπεύει την υπερβολική συνάρτηση ενεργοποίησης $tanh$ και το $\odot$ σημαίνει πολλαπλασιασμός κατά στοιχείο. Οι W_x είναι οι κρυφοί πίνακες βαρών εισόδου και οι W_h είναι οι παράμετροι των κρυφών πινάκων βαρών που μαθαίνονται κατά τη διάρκεια της εκπαίδευσης. Ομοίως, τα b_x και b_h είναι οι προκαταλήψεις που μαθαίνονται κατά την εκπαίδευση.



Εικόνα 3.8: *Κύτταρο LSTM [41]*

4 Ομοσπονδιακή Μάθηση

Τα τελευταία χρόνια, η ομοσπονδιακή μάθηση (Federated Learning - FL) έχει αναδειχθεί ως ένα πρωτοποριακό πρότυπο στον τομέα της μηχανικής μάθησης, προσφέροντας μια νέα προσέγγιση για την εκπαίδευση μοντέλων σε πολλαπλές αποκεντρωμένες συσκευές, διατηρώντας παράλληλα το απόρρητο και την ασφάλεια των δεδομένων. Αυτό το κεφάλαιο εισάγει την ομοσπονδιακή μάθηση εμβαθύνοντας στις βασικές αρχές της, την αρχιτεκτονική της και τη ιδιαίτερη διαδικασία που τη διακρίνει από τις παραδοσιακές συγκεντρωτικές μεθόδους μάθησης.

4.1 Εισαγωγή στην Ομοσπονδιακή Μάθηση

4.1.1 Ορισμός και βασικές αρχές

Η ομοσπονδιακή μάθηση είναι μια καινοτόμος τεχνική στη μηχανική μάθηση, όπου η διαδικασία εκπαίδευσης κατανέμεται σε πολλές συσκευές ή κόμβους, ο καθένας με το δικό του τοπικό σύνολο δεδομένων. Αυτή η μεθοδολογία έρχεται σε αντίθεση με τις παραδοσιακές συγκεντρωτικές τεχνικές μηχανικής μάθησης, όπου όλα τα δεδομένα μεταφορτώνονται σε έναν κεντρικό διακομιστή για την εκπαίδευση του μοντέλου. Στον πυρήνα της, η FL έχει σχεδιαστεί για να εκπαιδεύει αλγορίθμους συνεργατικά χωρίς να χρειάζεται να ανταλλάσσονται τα πραγματικά δεδομένα, διατηρώντας έτσι την ιδιωτικότητα και την ασφάλεια. [6]

Ενας τυπικός ορισμός της Ομοσπονδιακής μάθησης είναι ο παρακάτω: [43]

«Εστω N ιδιοκτήτες δεδομένων $\{\mathcal{F}_1, ...\mathcal{F}_N\}$, οι οποίοι επιθυμούν να εκπαιδεύσουν ένα μοντέλο μηχανικής μάθησης συγκεντρώνοντας τα αντίστοιχα δεδομένα τους $\{\mathcal{D}_1, ...\mathcal{D}_N\}$. Μια συμβατική μέθοδος είναι να συγκεντρώσουμε όλα τα δεδομένα και να χρησιμοποιήσουμε το $\mathcal{D} = \mathcal{D}_1 \cup ... \cup \mathcal{D}_N$ για να εκπαιδεύσουμε ένα μοντέλο $\mathcal{M}_{\mathcal{SUM}}$. Ενα σύστημα ομοσπονδιακής μάθησης είναι μια διαδικασία μάθησης στην οποία οι ιδιοκτήτες δεδομένων εκπαιδεύουν συνεργατικά ένα μοντέλο $\mathcal{M}_{\mathcal{FED}}$, στην οποία διαδικασία κάθε ιδιοκτήτης δεδομένων $\mathcal{F}_i$ δεν εκθέτει τα δεδομένα του $\mathcal{D}_i$ σε άλλους.

Επιπλέον, η ακρίβεια του $\mathcal{M}_{\mathcal{FED}}$, που συμβολίζεται ως $\mathcal{V}_{\mathcal{FED}}$ θα πρέπει να είναι πολύ κοντά στην απόδοση του $\mathcal{M}_{\mathcal{SUM}}$, $\mathcal{V}_{\mathcal{SUM}}$.

Τυπικά, έστω δ ένας μη αρνητικός πραγματικός αριθμός, εάν

$$|\mathcal{V}_{\mathcal{FED}} - \mathcal{V}_{\mathcal{SUM}}| < \delta$$

Τότε λέμε ότι ο αλγόριθμος ομοσπονδιακής μάθησης έχει απώλεια ακρίβειας δ»

4.1.2 Ιστορικό πλαίσιο και ανάπτυξη της ομοσπονδιακής μάθησης

Η ομοσπονδιακή μάθηση πρωτοεμφανίστηκε το 2016, όταν ερευνητές της Google την παρουσίασαν ως μια λύση για την εκπαίδευση μοντέλων σε αποκεντρωμένα δεδομένα που βρίσκονται στις συσκευές των χρηστών, όπως τα έξυπνα κινητά τηλέφωνα, χωρίς να διακυβεύεται το απόρρητό τους [6], [44]. Η προσέγγιση αυτή ήταν ιδιαίτερα ελκυστική για εφαρμογές όπου τα δεδομένα των χρηστών είναι ευαίσθητα ή όπου το κόστος μεταφοράς δεδομένων (τόσο σε χρηματικό όσο και σε χρονικό επίπεδο) είναι υψηλό.

Η ανάπτυξη της FL αντικατοπτρίζει μια ευρύτερη αλλαγή στον τομέα της ιδιωτικότητας των δεδομένων, η οποία χαρακτηρίζεται από τον αυξανόμενο ρυθμιστικό και δημόσιο έλεγχο σχετικά με τον τρόπο χρήσης και κοινής χρήσης των προσωπικών δεδομένων. Χαρακτηριστικό παράδειγμα αποτελούν νομοθεσίες όπως ο Γενικός Κανονισμός για την Προστασία Δεδομένων (GDPR) [5] στην Ευρώπη που έχουν υπογραμμίσει την ανάγκη για τεχνολογίες διατήρησης της ιδιωτικότητας στην ανάλυση δεδομένων.

Από το 2016 μέχρι σήμερα, η ομοσπονδιακή μάθηση έχει εξελιχθεί από μια εξειδικευμένη ιδέα σε μια βασική τεχνολογία στον τομέα της μηχανικής μάθησης με διαφύλαξη της ιδιωτικότητας (privacy-preserving machine learning). Η υιοθέτησή της οφείλεται τόσο στις τεχνολογικές εξελίξεις όσο και στην αυξανόμενη έμφαση στην προστασία της ιδιωτικής ζωής των χρηστών.

Η τεχνολογία έχει επίσης σημειώσει ταχεία εξέλιξη όσον αφορά την αντιμετώπιση των αρχικών περιορισμών, όπως ο χειρισμός δεδομένων που δεν είναι ανεξάρτητα και πανομοιότυπα κατανεμημένα (IID) σε όλες τις συσκευές, η βελτίωση της απόδοσης του μοντέλου υπό περιορισμένο εύρος ζώνης επικοινωνίας και η διασφάλιση της διαδικασίας FL από επιθέσεις. [44]

4.2 Αρχιτεκτονική και Διαδικασία Ομοσπονδιακής Μάθησης

4.2.1 Δομικά στοιχεία Ομοσπονδιακής Μάθησης

Η αρχιτεκτονική της ομοσπονδιακής μάθησης διευκολύνει την κατανεμημένη μηχανική μάθηση χωρίς συγκέντρωση δεδομένων, δίνοντας προτεραιότητα στην προστασία της ιδιωτικότητας και την αποδοτικότητα. Αποτελείται από: [45]

- **Πελάτες (clients):** Συσκευές που εκπαιδεύουν μοντέλα με τα τοπικά τους δεδομένα.
- **Κεντρικό διακομιστή (server):** Συντονίζει την διαδικασία μάθησης, συγκεντρώνοντας τις ενημερώσεις από τους clients
- **Παγκόσμιο Μοντέλο (global model):** Το μοντέλο που εκπαιδεύεται συνεργατικά από όλους τους πελάτες
- **Δίκτυο επικοινωνίας:** Το δίκτυο που επιτρέπει τις ασφαλείς συναλλαγές μεταξύ clients και server

Η αρχιτεκτονική client-server που εφαρμόζει η ομοσπονδιακή μάθηση πέρα από την αποκέντρωση, προσφέρει και αρκετά πλεονεκτήματα όπως:

- Διατήρηση του απορρήτου: Τα δεδομένα παραμένουν τοπικά, ελαχιστοποιώντας τους κινδύνους που μπορεί να προκύψουν κατά την μεταφορά τους. [46]
- Επεκτασιμότητα: Υποστηρίζει πολυάριθμους (έως και εκατομμύρια) πελάτες λόγω της διανομής των υπολογιστικών εργασιών [47].
- Μειωμένη επιβάρυνση δικτύου: Ελαχιστοποιεί την ανάγκη για μεταφορά μεγάλου όγκου δεδομένων βελτιστοποιώντας την χρήση των πόρων του δικτύου. [6]

4.2.2 Διαδικασία εκπαίδευσης Ομοσπονδιακής Μάθησης

Μια τυπική εκπαίδευση με βάση την παραπάνω αρχιτεκτονική περιλαμβάνει τα εξής βήματα: [43]

1) **Τοπική εκπαίδευση:** Κάθε πελάτης εκπαιδεύει ένα μοντέλο τοπικά χρησιμοποιώντας το δικό του σύνολο δεδομένων και στην συνέχεια στέλνει τις κλίσεις ή τα βάρη του

μοντέλου στον κεντρικό διακομιστή. Αυτό το βήμα αξιοποιεί τους υπολογιστικούς πόρους του πελάτη για να μάθει από τα τοπικά χαρακτηριστικά των δεδομένων χωρίς να εκθέσει ή να μεταφέρει τα δεδομένα

2) **Συγκέντρωση μοντέλων:** Ο διακομιστής συγκεντρώνει τις ενημερώσεις των τοπικών μοντέλων για να σχηματίσει ένα ενιαίο, βελτιωμένο καθολικό μοντέλο -χωρίς να μαθαίνει καθόλου πληροφορίες για τους πελάτες. Αυτό το βήμα συνάθροισης χρησιμοποιεί συχνά τον ομοσπονδιακό μέσο (FedAvg), ο οποίος υπολογίζει έναν σταθμισμένο μέσο όρο των ενημερώσεων
3) **Αποστολή καθολικού μοντέλου:** Το καθολικό μοντέλο στέλνεται πίσω στους πελάτες
4) **Ενημέρωση μοντέλου:**. Οι πελάτες χρησιμοποιούν το ενημερωμένο καθολικό μοντέλο ως σημείο εκκίνησης για τον επόμενο γύρο τοπικής εκπαίδευσης.

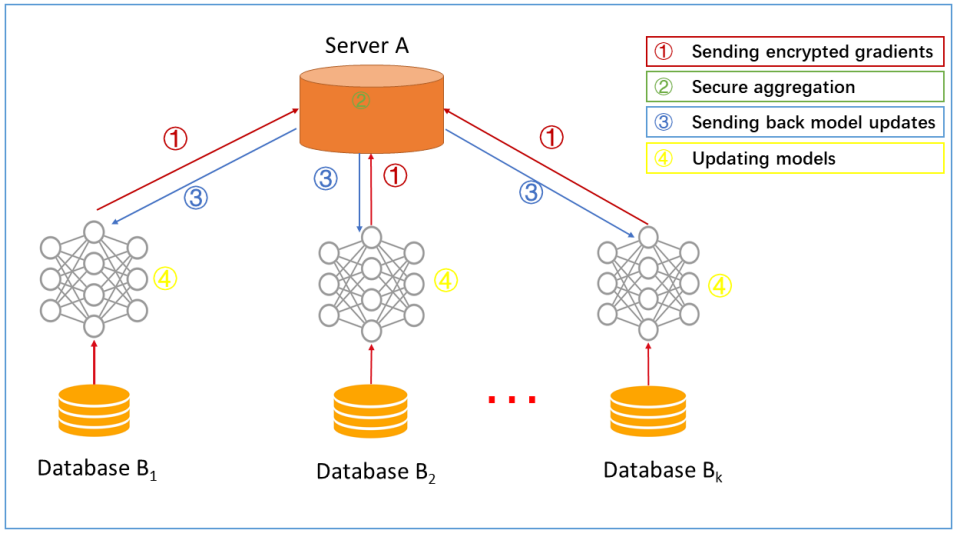
Algorithm 1 `FederatedAveraging`. The K clients are indexed by k; B is the local minibatch size, E is the number of local epochs, and η is the learning rate.

Server executes:
initialize w_0
for each round $t = 1, 2, \dots$ **do**
$m \leftarrow \max(C \cdot K, 1)$
$S_t \leftarrow$ (random set of m clients)
for each client $k \in S_t$ **in parallel do**
$w_{t+1}^k \leftarrow \text{ClientUpdate}(k, w_t)$
$m_t \leftarrow \sum_{k \in S_t} n_k$
$w_{t+1} \leftarrow \sum_{k \in S_t} \frac{n_k}{m_t} w_{t+1}^k$ *// Erratum*[4]

ClientUpdate(k, w): *// Run on client k*
$\mathcal{B} \leftarrow$ (split $\mathcal{P}_k$ into batches of size B)
for each local epoch i from 1 to E **do**
for batch $b \in \mathcal{B}$ **do**
$w \leftarrow w - \eta \nabla \ell(w; b)$
return w to server

Εικόνα 4.1: *Ο αλγόριθμος FedAvg [6]*

Οι επαναλήψεις των παραπάνω βημάτων συνεχίζονται έως ότου η συνάρτηση απώλειας να συγκλίνει, ολοκληρώνοντας έτσι τη διαδικασία εκπαίδευσης.



Εικόνα 4.2: *Διαδικασία εκπαίδευσης Ομοσπονδιακής Μάθησης [43]*

5 Διαφορική Ιδιωτικότητα

5.1 Εισαγωγή στην Διαφορική Ιδιωτικότητα

Η Διαφορική Ιδιωτικότητα (Differential Privacy - DP) αντιπροσωπεύει μια στροφή στον τρόπο με τον οποίο ορίζεται και εφαρμόζεται η ιδιωτικότητα των δεδομένων. Ο όρος προτάθηκε από την Cynthia Dwork το 2006 [48] και προέκυψε από τη διαπίστωση ότι οι παραδοσιακές τεχνικές ανωνυμοποίησης ήταν ανεπαρκείς μπροστά στα εξελιγμένα εργαλεία εξόρυξης δεδομένων και τους αλγορίθμους που είναι σε θέση να ταυτοποιήσουν εκ νέου άτομα από "ανωνυμοποιημένα" σύνολα δεδομένων [49]. Η DP εισάγει ένα ποσοτικοποιήσιμο μέτρο της απώλειας της ιδιωτικότητας, επιτρέποντας στους αναλυτές δεδομένων να παρέχουν ισχυρές εγγυήσεις ιδιωτικότητας, ενώ εξακολουθούν να εξάγουν πολύτιμες πληροφορίες από τα δεδομένα. Αυτή η ισορροπία είναι εξαιρετικά σημαντική στην σημερινή εποχή των μεγάλων δεδομένων και γι' αυτό θέλουμε να ερευνήσουμε την εφαρμογή της και στην εργασίας μας.

Η DP αντιμετωπίζει τις προκλήσεις της ιδιωτικότητας παρέχοντας ένα πλαίσιο που είναι μαθηματικά αυστηρό και προσαρμόσιμο σε διάφορους τύπους δεδομένων και τεχνικές ανάλυσης. Η υιοθέτησή του από μεγάλες εταιρείες τεχνολογίας και κυβερνητικές υπηρεσίες υπογραμμίζει

την πρακτική σημασία και αποτελεσματικότητά του. Χαρακτηριστικό παράδειγμα αποτελεί το γεγονός πως η απογραφή του πληθυσμού των Η.Π.Α. για το έτος 2020 πραγματοποιήθηκε με εγγυήσεις Διαφορικής Ιδιωτικότητας. [50]

Η βασική αρχή της DP έγκειται στην ιδέα ότι τα αποτελέσματα ενός υπολογισμού σε ένα σύνολο δεδομένων θα πρέπει να παραμένουν στατιστικά παρόμοια ανεξάρτητα από το αν τα δεδομένα ενός μεμονωμένου ατόμου περιλαμβάνονται ή αποκλείονται από το σύνολο δεδομένων [48]. Αυτό σημαίνει ότι ένας επιτιθέμενος δεν θα πρέπει να είναι σε θέση να διακρίνει, με μεγάλη πιθανότητα εάν ένα συγκεκριμένο άτομο συμμετείχε ή όχι σε μια μελέτη ή επηρέασε ένα αποτέλεσμα ερωτήματος. Η DP επιτυγχάνει αυτόν τον στόχο με τη στρατηγική εισαγωγή ελεγχόμενων ποσοτήτων τυχαιότητας (συνήθως θόρυβος) στους υπολογισμούς, αποκρύπτοντας έτσι τον αντίκτυπο κάθε μεμονωμένης εγγραφής, ενώ παράλληλα επιτρέπει την εξαγωγή σημαντικών πληροφοριών από τα δεδομένα συνολικά. [51]

5.2 Βασικές έννοιες Διαφορικής Ιδιωτικότητας

5.2.1 ε-Διαφορική Ιδιωτικότητα

Η ε-διαφορική ιδιωτικότητα ($\epsilon - DP$) τυποποιεί την κεντρική ιδέα της διαφορικής ιδιωτικότητας με μια ποσοτικοποιήσιμη παράμετρο ιδιωτικότητας, ε (έψιλον). Το ε λειτουργεί ως "προϋπολογισμός ιδιωτικότητας", ελέγχοντας τη μέγιστη διαφορά στην πιθανότητα εμφάνισης οποιουδήποτε συγκεκριμένου αποτελέσματος όταν ένας υπολογισμός εκτελείται σε δύο σύνολα δεδομένων που διαφέρουν κατά μία εγγραφή. Ενα μικρότερο ε υποδηλώνει ισχυρότερη προστασία, καθώς σημαίνει ότι η ικανότητα ενός επιτιθέμενου να συμπεράνει τη συμμετοχή ενός ατόμου με βάση το αποτέλεσμα περιορίζεται σημαντικά. Ωστόσο, η εισαγωγή του θορύβου που απαιτείται για την επίτευξη μικρότερου ε μπορεί επίσης να οδηγήσει σε κάποια απώλεια χρησιμότητας των αποτελεσμάτων.

Τυπικά, ένας τυχαιοποιημένος μηχανισμός $\mathcal{M}$ επιτυγχάνει $\epsilon - DP$ εάν για οποιαδήποτε δύο σύνολα δεδομένων D και D' που διαφέρουν σε ένα μόνο άτομο, για όλα τα σύνολα αποτελεσμάτων $\mathcal{S}$, η έξοδος του μηχανισμού υπακούει στην ακόλουθη ανισότητα: [51]

$$\Pr[\mathcal{M}(D) \in \mathcal{S}] \leq e^{\epsilon} \cdot \Pr[\mathcal{M}(D') \in \mathcal{S}]$$

5.2.2 Προσθήκη θορύβου (μηχανισμοί Laplace και Gauss)

Υπάρχουν αρκετοί μηχανισμοί θορύβου για την εισαγωγή της τυχαιότητας που απαιτείται για τη διαφορική ιδιωτικότητα. Παρακάτω θα εστιάσουμε στους βασικότερους:

- Μηχανισμός Laplace: Είναι ο βασικός μηχανισμός διαφορικής ιδιωτικότητας για τις αριθμητικές εξόδους. Προσθέτει θόρυβο που προέρχεται από μια κατανομή Laplace, η οποία έχει κέντρο το μηδέν και μια παράμετρο κλίμακας που καθορίζεται από το επιθυμητό επίπεδο ιδιωτικότητας (ε) και την ευαισθησία του ερωτήματος (πόσο μπορεί να αλλάξει με βάση μια μόνο εγγραφή). Οσο ευρύτερη είναι η κατανομή Laplace, τόσο ισχυρότερη είναι η προστασία της ιδιωτικότητας, αλλά ενδεχομένως με μειωμένη ακρίβεια. [51]
- Μηχανισμός Gauss: Ομοίως με τον μηχανισμό Laplace, ο μηχανισμός Gauss προσθέτει θόρυβο δειγματοληπτικά από μια γκαουσιανή (κανονική) κατανομή. Συχνά προτιμάται όταν έχει να κάνει με δεδομένα υψηλότερων διαστάσεων ή όταν πρέπει να εκτελεστούν διαδοχικά πολλαπλοί υπολογισμοί διαφορετικής ιδιωτικότητας. [51]

5.2.3 Καθολική & Τοπική Διαφορική Ιδιωτικότητα

Η Καθολική Διαφορική Ιδιωτικότητα (Global Differential Privacy - GDP) εφαρμόζει τους μηχανισμούς DP σε επίπεδο συνόλου δεδομένων, προσφέροντας εγγυήσεις απορρήτου για τα συγκεντρωτικά δεδομένα. Η προσέγγιση αυτή προϋποθέτει έναν έμπιστο επιμελητή που συλλέγει και ανωνυμοποιεί τα δεδομένα πριν από οποιαδήποτε ανάλυση. [52]

Αντίθετα, η Τοπική Διαφορική Ιδιωτικότητα (Local Differential Privacy - LDP) εφαρμόζει μηχανισμούς διατήρησης της ιδιωτικότητας σε επίπεδο μεμονωμένων καταχωρίσεων δεδομένων, πριν τα δεδομένα συγκεντρωθούν ή αναλυθούν. Αυτό το μοντέλο δεν απαιτεί έναν έμπιστο επιμελητή, καθώς τα δεδομένα ιδιωτικοποιούνται στην πηγή, προσφέροντας ισχυρότερες εγγυήσεις ιδιωτικότητας με κόστος τον δυνητικά υψηλότερο θόρυβο και τη μειωμένη χρησιμότητα των δεδομένων. Το LDP είναι ιδιαίτερα σημαντικό σε σενάρια όπου δεν μπορούμε να εμπιστευτούμε τον συλλέκτη δεδομένων. [52]

5.3 Διαφορική Ιδιωτικότητα στην Ομοσπονδιακή Μάθηση.

Η χρήση της διαφορικής ιδιωτικότητας στην ομόσπονδη μάθηση περιλαμβάνει την ενσωμάτωση μηχανισμών DP στη διαδικασία FL για να διασφαλιστεί ότι η εκπαίδευση του μοντέλου σε πολλαπλούς αποκεντρωμένους κόμβους ή πελάτες διατηρεί την ιδιωτικότητα των μεμονωμένων σημείων δεδομένων. Οι Abadi και McMahan που εισήγαν και την έννοια της FL έχουν πρωτοστατήσει στον συνδυασμό τους. Το 2016 δημοσίευσαν μαζί με τους συνεργάτες τους τον αλγόριθμο DP-SGD (Differentially Private Stochastic Gradient Descent) μια παραλλαγή του βασικού αλγορίθμου SGD της FL καθώς και την τεχνική Moments accountant για την ποσοτικοποίηση του privacy budget.[7]

5.3.1 Ο αλγόριθμος DP-SGD

Ο DP-SGD τροποποιεί την παραδοσιακή διαδικασία SGD εισάγοντας δύο βασικούς μηχανισμούς: την αποκοπή κλίσεων (gradient clipping) και την προσθήκη θορύβου (noise addition). Στην αποκοπή κλίσεων, οι κλίσεις που υπολογίζονται για κάθε παρτίδα δεδομένων αποκόπτονται για να εξασφαλιστεί ότι κανένα σημείο δεδομένων δεν μπορεί να επηρεάσει δυσανάλογα τις ενημερώσεις του μοντέλου. Αυτό επιτυγχάνεται με τη μείωση των κλίσεων που υπερβαίνουν ένα προκαθορισμένο κατώφλι, περιορίζοντας ουσιαστικά τη μέγιστη συνεισφορά κάθε σημείου δεδομένων στην κλίση.

Μετά την περικοπή, ο DP-SGD προσθέτει θόρυβο στις συγκεντρωτικές κλίσεις πριν από το βήμα ενημέρωσης του μοντέλου. Αυτός ο θόρυβος εισάγεται για να αποκρύψει τις συνεισφορές των μεμονωμένων σημείων δεδομένων, παρέχοντας μια πιθανολογική εγγύηση ότι η έξοδος του μοντέλου δεν μπορεί να χρησιμοποιηθεί για να εξάγει συμπεράσματα για οποιοδήποτε μεμονωμένο δεδομένο στο σύνολο εκπαίδευσης. [7]

Algorithm 1 Differentially private SGD (Outline)

Input: Examples $\{x_1, \ldots, x_N\}$, loss function $\mathcal{L}(\theta) = \frac{1}{N}\sum_i \mathcal{L}(\theta, x_i)$. Parameters: learning rate η_t, noise scale σ, group size L, gradient norm bound C.

Initialize θ_0 randomly

for $t \in [T]$ **do**

Take a random sample L_t with sampling probability L/N

Compute gradient

For each $i \in L_t$, compute $\mathbf{g}_t(x_i) \leftarrow \nabla_{\theta_t}\mathcal{L}(\theta_t, x_i)$

Clip gradient

$\bar{\mathbf{g}}_t(x_i) \leftarrow \mathbf{g}_t(x_i) / \max\left(1, \frac{\|\mathbf{g}_t(x_i)\|_2}{C}\right)$

Add noise

$\tilde{\mathbf{g}}_t \leftarrow \frac{1}{L}\left(\sum_i \bar{\mathbf{g}}_t(x_i) + \mathcal{N}(0, \sigma^2 C^2 \mathbf{I})\right)$

Descent

$\theta_{t+1} \leftarrow \theta_t - \eta_t \tilde{\mathbf{g}}_t$

Output θ_T and compute the overall privacy cost (ε, δ) using a privacy accounting method.

Αλγόριθμος 5.1: *DP-SGD [7]*

5.3.2 Η τεχνική Moments Accountant

Η μέθοδος Moments Accountant εισάγεται για την ακριβή ποσοτικοποίηση της συσσωρευτικής απώλειας ιδιωτικότητας σε αλγορίθμους που χρησιμοποιούν διαφορική ιδιωτικότητα, όπως ο DP-SGD. Προσφέρει σημαντική βελτίωση σε σχέση με τις παραδοσιακές μεθόδους μέτρησης απώλειας ιδιωτικότητας, παρέχοντας ένα αυστηρότερο όριο για τον αθροιστικό προϋπολογισμό ιδιωτικότητας σε πολλαπλούς υπολογισμούς ή αλγοριθμικές επαναλήψεις. Ουσιαστικά, ο λογιστής στιγμών παρακολουθεί το λογάριθμο της συνάρτησης δημιουργίας στιγμών της τυχαίας μεταβλητής απώλειας ιδιωτικότητας, επιτρέποντας μια πιο ακριβή ανάλυση του τρόπου με τον οποίο οι εγγυήσεις ιδιωτικότητας συσσωρεύονται όταν ένας αλγόριθμος εφαρμόζεται επανειλημμένα. [7]

Αυτό είναι ιδιαίτερα σημαντικό για σενάρια βαθιάς μάθησης και ομοσπονδιακής μάθησης, όπου τα μοντέλα υποβάλλονται σε πολυάριθμες επαναλήψεις ενημερώσεων. Με τη χρήση της moments accountant, μας δίνεται η δυνατότητα να διαχειριστούμε και να περιορίσουμε

αποτελεσματικότερα τον συνολικό προϋπολογισμό ιδιωτικότητας, διασφαλίζοντας ότι το μοντέλο παραμένει εντός αποδεκτών παραμέτρων απώλειας ιδιωτικότητας καθ' όλη τη διάρκεια της διαδικασίας εκπαίδευσης.

ΠΡΑΚΤΙΚΟ ΜΕΡΟΣ

6 Διερευνητική Ανάλυση Δεδομένων

Το σύνολο δεδομένων στο οποίο εργαζόμαστε είναι πραγματικό και προέρχεται από το ηλεκτρικό δίκτυο της πόλης Terni της Ιταλίας. Πρόκειται για ένα σύνολο 30 παραγωγών-καταναλωτών (prosumers) ηλεκτρικής ενέργειας μικρής κλίμακας, οι οποίοι χρησιμοποιούν Φωτοβολταικά συστήματα για παραγωγή ενέργειας [53].

Τα δεδομένα, τα οποία δεν είναι διαθέσιμα στο διαδίκτυο, καλύπτουν την περίοδο από 01/05/2015 έως 28/02/2019 και καταγράφουν την παραγόμενη ηλεκτρική ενέργεια σε kW/h για κάθε prosumer ανά 15 λεπτά.

Χρησιμοποιώντας μια ποικιλία τεχνικών διερευνητικής ανάλυσης, εξετάζουμε λεπτομερώς το σύνολο των δεδομένων, εντοπίζοντας τις βασικές μεταβολές και αξιολογώντας τον πιθανό αντίκτυπό τους στα αποτελέσματα της μελέτης μας. Αυτή η αρχική διερεύνηση είναι καθοριστικής σημασίας, καθώς ενημερώνει τις επακόλουθες στρατηγικές μοντελοποίησης και τον έλεγχο υποθέσεων.

Πιο συγκεκριμένα, θα αναζητήσουμε τις ελλιπείς τιμές για να διασφαλίσουμε την πληρότητα των δεδομένων, θα ανασυνθέσουμε τις παρατηρήσεις (resampling) για να ενισχύσουμε την αντιπροσωπευτικότητά τους, θα τα εξετάσουμε για ακραίες τιμές για να διατηρήσουμε την ακεραιότητα τους, θα υπολογίσουμε την αυτοσυσχέτιση (autocorrelation) για να κατανοήσουμε τις χρονικές εξαρτήσεις και θα εκτελέσουμε ομαδοποίηση (clustering) για να εντοπίσουμε εγγενείς ομοιότητες. Καθένα από αυτά τα βήματα είναι κρίσιμο για τη δημιουργία μιας στέρεης βάσης για την ανάλυσή μας, διασφαλίζοντας ότι η προσέγγισή μας είναι τόσο εμπεριστατωμένη όσο και ακριβής.

6.1 Διαχείριση ελλιπών τιμών

Πριν κάνουμε οποιαδήποτε ενέργεια με τα δεδομένα μας πρέπει να σιγουρευτούμε ότι η χρονοσειρά μας είναι ακέραια και δεν έχει ελλιπείς τιμές. Πραγματοποιώντας τον σχετικό

έλεγχο διαπιστώσαμε ότι ορισμένοι prosumers είχαν ελλιπή δεδομένα για το χρονικό διάστημα 02:00-03:00 τις ημέρες 29-03-2015 και 27-03-2016 το οποίο πιθανώς να οφείλεται σε εργασίες συντήρησης ή κάποιο άλλο τεχνικό πρόβλημα. Πέραν αυτών υπήρχαν κάποιες σποραδικές ελλείψεις χωρίς όμως να παρατηρείται κάποιο μοτίβο.

Οπως αναφέραμε και στο θεωρητικό μέρος οι τρόποι διαχείρισης των ελλιπών τιμών ποικίλλουν. Σε μια αντίστοιχη περίπτωση πρόβλεψης φορτίου [54], οι ερευνητές επέλεξαν να θέσουν την τιμή ίδια με αυτήν της προηγούμενης μέτρησης. Εξετάζοντας λοιπόν τις αμέσως προηγούμενες και επόμενες μετρήσεις από τις ελλιπείς τιμές παρατηρήσαμε πως διέφεραν ελάχιστα μεταξύ τους. Ετσι αποφασίσαμε να συμπληρώσουμε τις τιμές με τον μέσο όρο της προηγούμενης και επόμενης μέτρησης. Η επίπτωση της αλλαγής είναι στατιστικά μη σημαντική καθώς χρειάστηκε να αλλάξουμε το πολύ 2 από τις 67208 παρατηρήσεις (0.003%).

6.2 Ανασύνθεση Δειγμάτων

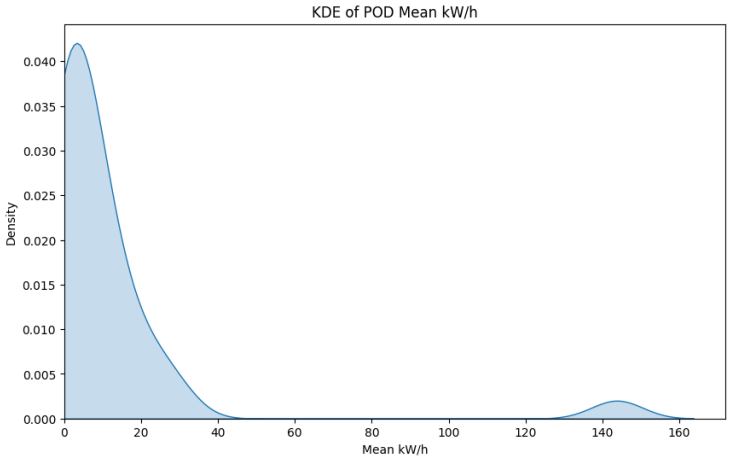
Η ανασύνθεση των παρατηρήσεων (resampling) είναι σημαντική για την εξισορρόπηση μεταξύ της λεπτομέρειας και των αναλυτικών στόχων. Η αρχική καταγραφή του συνόλου δεδομένων με συχνότητα παρατηρήσεων των 15 λεπτών αποδείχθηκε πολύ υψηλή κατά την διάρκεια των πειραμάτων χωρίς να προσφέρει ουσιαστικό πλεονέκτημα. Ακόμη, κατά την μελέτη της σχετικής βιβλιογραφίας, παρατηρήσαμε ότι στις περισσότερες περιπτώσεις τα δεδομένα προς επεξεργασία είχαν παρατηρήσεις χρονικού εύρους 30 ή 60 λεπτών. [3], [14], [41], [54], [55]

Για τους παραπάνω λόγους επιλέξαμε να κάνουμε resample τις χρονοσειρές σε παρατηρήσεις ανά 30 λεπτά. Με αυτή τη στρατηγική resampling καταφέραμε να διασφαλίσουμε τη συνοχή με τα ευρύτερα ερευνητικά πρότυπα, αλλά και να μειώσουμε αποδοτικά τον όγκο των δεδομένων, με αποτέλεσμα τον μετριασμό των κινδύνων υπερπροσαρμογής (overfitting) και την επιτάχυνση της σύγκλισης (convergence) του μοντέλου.

6.3 Διαχείριση ακραίων τιμών

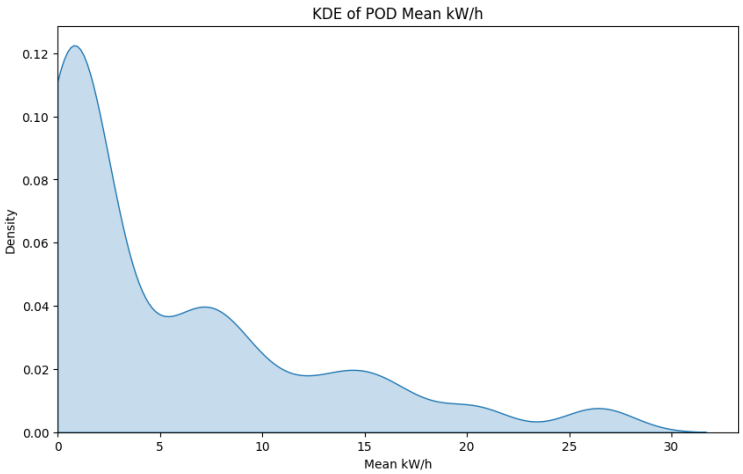
Η μη αντιμετώπιση των ακραίων τιμών μπορεί να οδηγήσει σε ανακριβείς προβλέψεις, παραπλανητικές ερμηνείες και λανθασμένη λήψη αποφάσεων [56]. Για την εύρεση των παραγωγών με ακραίες τιμές χρησιμοποιήσαμε μέθοδο Εκτίμησης Πιθανότητας Πυρήνα (Kernel

Density Estimation - KDE) [57]. Η KDE είναι μια μη παραμετρική μέθοδος που χρησιμοποιείται για την εκτίμηση της συνάρτησης πυκνότητας πιθανότητας μιας συνεχούς τυχαίας μεταβλητής. Στο πλαίσιο της εργασίας την χρησιμοποιήσαμε για την ανίχνευση ακραίων τιμών με την εκτίμηση της συνάρτησης πυκνότητας πιθανότητας των σημείων δεδομένων με την πάροδο του χρόνου. Με την εξομάλυνση των δεδομένων της χρονοσειράς εντοπίσαμε σημεία αποκλίσεων από την τυπική κατανομή των δεδομένων, όπως φαίνεται στο διάγραμμα 6.1.



Διάγραμμα 6.1: *KDE πριν την αφαίρεση ακραίων στοιχείων*

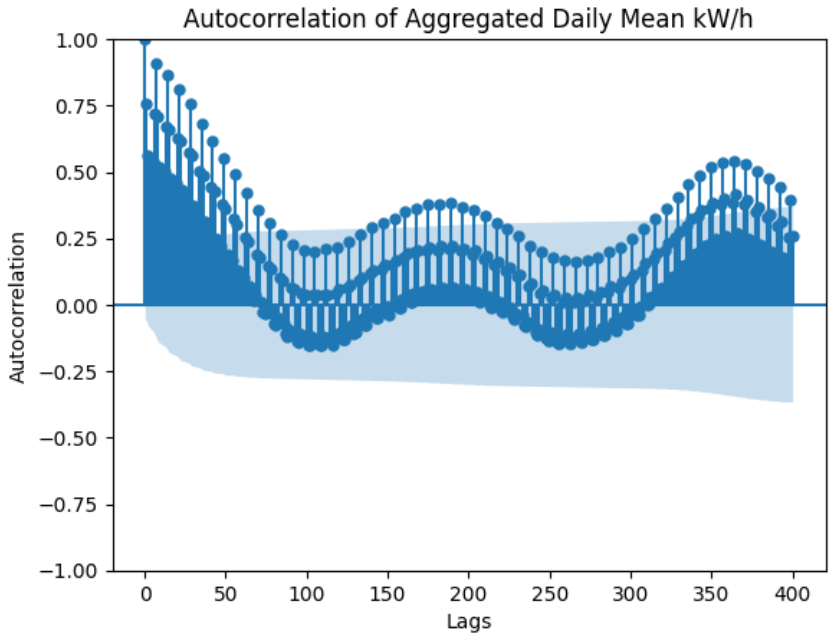
Παρατηρούμε το την τιμή στο 140 που οφείλεται στον παραγωγό 'P_00506E' τον οποίο και αφαιρέσαμε από το σύνολο δεδομένων. Η καμπύλη μετά την αφαίρεσή του φαίνεται στο διάγραμμα 6.2.



Διάγραμμα 6.2: *KDE μετά την αφαίρεση ακραίων στοιχείων*

6.4 Αυτοσυσχέτιση

Αφού αφαιρέσαμε τις ακραίες τιμές και έχουμε μια περισσότερο ομογενή χρονοσειρά, αποφασίσαμε να μελετήσουμε την αυτοσυσχέτιση της χρονοσειράς. Οπως αναφέραμε και στο θεωρητικό μέρος, η αυτοσυσχέτιση, είναι ένα στατιστικό μέγεθος που ποσοτικοποιεί το βαθμό ομοιότητας μεταξύ μιας δεδομένης χρονοσειράς και μιας καθυστερημένης εκδοχής της σε διαδοχικά χρονικά διαστήματα [9]. Για να την υπολογίσουμε χρησιμοποιούμε την συνάρτηση αυτοσυσχέτισης που περιέχεται στην βιβλιοθήκη statsmodels [58] της python και δημιουργούμε την γραφική της παράσταση. Στο διάγραμμα 6.3 παρατηρούμε την αυτοσυσχέτιση για την μέση ημερήσια παραγωγή ενέργειας όλων των παραγωγών.



Διάγραμμα 6.3: *Αυτοσυσχέτιση μέσης ημερήσιας παραγωγής ενέργειας για το σύνολο των παραγωγών*

Ο άξονας Y δείχνει τις τιμές της αυτοσυσχέτισης και η τιμή του είναι μεταξύ -1 και 1

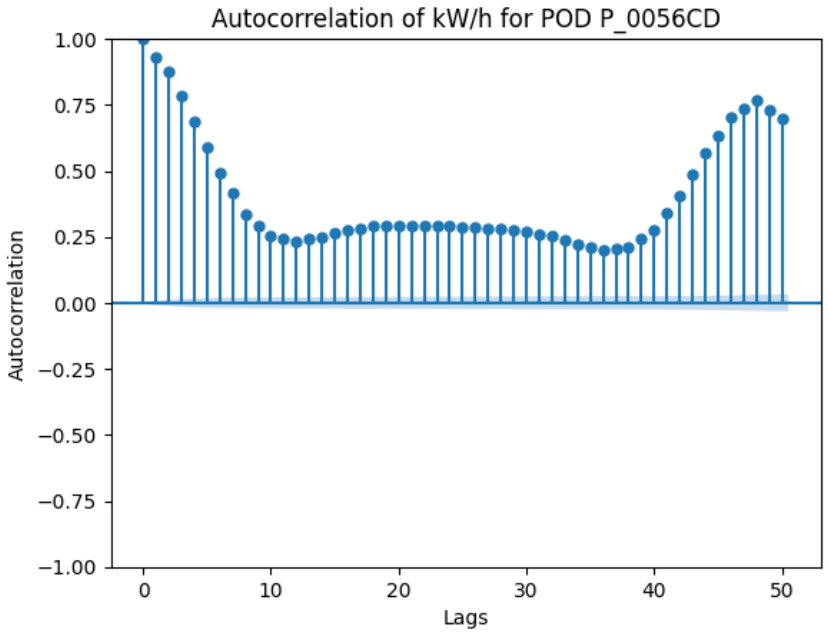
- Το 1 υποδηλώνει τέλεια θετική συσχέτιση: καθώς η χρονοσειρά κινείται από το ένα χρονικό σημείο στο επόμενο, το κάνει με απόλυτα προβλέψιμο τρόπο.
- Το -1 υποδηλώνει τέλεια αρνητική συσχέτιση: καθώς η χρονοσειρά κινείται από το ένα χρονικό σημείο στο επόμενο, το κάνει με απόλυτα αντίστροφο προβλέψιμο τρόπο.

- Το 0 υποδηλώνει πως δεν υπάρχει καμία συσχέτιση: η γνώση της τιμής σε ένα χρονικό σημείο δεν παρέχει καμία πληροφορία για την τιμή στο επόμενο χρονικό σημείο.

Ο άξονας *X* δείχνει τον αριθμό των χρονικών βημάτων που χωρίζουν τις παρατηρήσεις της σειράς (υστέρηση). Η υστέρηση 0 έχει πάντα αυτοσυσχέτιση 1, επειδή είναι η συσχέτιση της σειράς με τον εαυτό της.
Η υστέρηση 1 δείχνει τη συσχέτιση της σειράς με την προηγούμενη τιμή της, η υστέρηση 2 με δύο χρονικά βήματα πίσω, κ.ο.κ.

Διαστήματα εμπιστοσύνης: Το διάγραμμα περιλαμβάνει και μια σκιασμένη περιοχή ή γραμμές που επισημαίνουν το διάστημα εμπιστοσύνης (συνήθως 95%). Εάν η αυτοσυσχέτιση σε μια συγκεκριμένη υστέρηση βρίσκεται εκτός αυτής της περιοχής, είναι στατιστικά σημαντική.

Στην περίπτωση μας χρησιμοποιούμε το διάγραμμα για να μας βοηθήσει στην επιλογή του παρελθοντικού χρονικού ορίζοντα των παρατηρήσεων για την εκπαίδευση των μοντέλων μας. Οπως θα δούμε παρακάτω, για την προετοιμασία των δεδομένων εκπαίδευσης χρησιμοποιούμε την τεχνική του κυλιόμενου παραθύρου. Κοιτάμε δηλαδή τις X προηγούμενες τιμές για προβλέψουμε την επόμενη. Στο διάγραμμα 6.4 βλέπουμε πως ο συντελεστής αυτοσυσχέτισης είναι εκτός του διαστήματος εμπιστοσύνης και κοντά στο 1 στην τιμή 48.
Η ένδειξη αυτή δεν είναι τυχαία καθώς έχουμε παρατηρήσεις ανά μισή ώρα, συνεπώς μια ημέρα θα έχει 48 παρατηρήσεις, τον αριθμό που επιλέξαμε για την εκπαίδευση του LSTM μοντέλου μας.



***Διάγραμμα 6.4**: Συνάρτηση αυτοσυσχέτισης για τον παραγωγό P_0056CD*

Στο διάγραμμα 6.3 όπου υπολογίσαμε την εμβέλεια μέτρησης μέχρι τις 400 παρατηρήσεις, βλέπουμε ότι υπάρχει άλλη μια περιοχή σημαντικής αυτοσυσχέτισης γύρω από την τιμή 350. Και πάλι η τιμή είναι σημαντική καθώς 336 παρατηρήσεις σηματοδοτούν μια εβδομάδα. Στις δοκιμές του μοντέλου μας ωστόσο, το παράθυρο 336 παρατηρήσεων ήταν χειρότερο από των 48.

7 Πειραματική Διαδικασία

Στο παρόν κεφάλαιο παρουσιάζεται η ολοκληρωμένη μεθοδολογία που ακολουθήσαμε για τη διερεύνηση των δυνατοτήτων της ομοσπονδιακής μάθησης για την πρόβλεψη παραγωγής με διατήρηση της ιδιωτικότητας. Αρχικά, παρουσιάζονται τα πλαίσια λογισμικού που υποστηρίζουν τα πειραματικά σενάρια που σχεδιάσαμε και στην συνέχεια περιγράφονται οι επιλεγμένες αρχιτεκτονικές νευρωνικών δικτύων. Στην συνέχεια, εμβαθύνουμε στις αρχές της FL και της DP, τονίζοντας τους ρόλους τους στη διασφάλιση των δεδομένων των χρηστών. Το κεφάλαιο προχωρά στη σκιαγράφηση των συγκριτικών σεναρίων τονίζοντας τις βασικές διαφορές τους όσον αφορά τον χειρισμό των δεδομένων και την προστασία της ιδιωτικότητας. Για να

διασφαλιστεί μια δίκαιη αξιολόγηση, περιγράφονται λεπτομερώς οι διαδικασίες εκπαίδευσης που αφορούν το κάθε παράδειγμα μάθησης. Τέλος, το κεφάλαιο εισάγει το σύνολο των μετρικών αξιολόγησης που θα χρησιμοποιηθούν και τους λόγους επιλογής τους.

7.1 Εισαγωγή στα σενάρια, και τα εργαλεία.

Οπως αναφέραμε στην αρχή, στόχος αυτής της εργασίας είναι να διερευνήσουμε τις προοπτικές και τα πλεονεκτήματα της ομοσπονδιακής μάθησης στον τομέα των προβλέψεων φορτίου. Η μελέτη μας επικεντρώνεται στη σύγκριση της αποτελεσματικότητας και των συμβιβασμών της ομοσπονδιακής μάθησης, ιδίως όταν συνδυάζεται με τη διαφορική ιδιωτικότητα έναντι πιο παραδοσιακών μεθόδων μάθησης, όπως η συγκεντρωτική και η τοπική. Για την μελέτη της αποτελεσματικότητας θα εκπαιδεύσουμε ένα δίκτυο μακράς βραχυπρόθεσμης μνήμης. Η εκπαίδευση θα γίνει στα ίδια δεδομένα και στην συνέχεια θα πραγματοποιήσουμε συγκρίσεις μεταξύ των τεσσάρων διαφορετικών μεθόδων μάθησης – τοπική, συγκεντρωτική, ομοσπονδιακή με και χωρίς διαφορική ιδιωτικότητα.

Η μελέτη μας χρησιμοποιεί το TensorFlow (TF) [59], ένα από τα βασικότερα ολοκληρωμένα εργαλεία για μηχανική μάθηση. Είναι ιδανικό για τις ανάγκες μας λόγω της αποδοτικότητας και της ικανότητάς του να χειρίζεται λεπτομερείς αριθμητικές εργασίες αλλά και λόγω του TensorFlow Federated (TFF) που αποτέλεσε και τον βασικό πυλώνα τον πειραμάτων μας στην ομοσπονδιακή μάθηση. Το TFF είναι ένα πλαίσιο λογισμικού ανοιχτού κώδικα, ειδικά σχεδιασμένο για την υποστήριξη εργασιών FL, που επεκτείνει τις δυνατότητες του TF επιτρέποντας την ανάπτυξη και προσομοίωση μοντέλων σε αποκεντρωμένα δεδομένα [60].

Για την κατασκευή μοντέλων και τον πειραματισμό, χρησιμοποιήσαμε το Keras [61], ένα υψηλού επιπέδου API νευρωνικών δικτύων που τρέχει πάνω στο TensorFlow και το TFF. Το Keras απλοποιεί τη δημιουργία μοντέλων βαθιάς μάθησης με τη φιλική προς το χρήστη διεπαφή του, επιτρέποντας την ταχεία δημιουργία πρωτοτύπων και τον πειραματισμό.

Τέλος, για τις διαδικασίες DP χρησιμοποιήσαμε ένα ακόμα πλαίσιο του TF, το TensorFlow Privacy.

Τα πειράματα πραγματοποιήθηκαν στο ακόλουθο περιβάλλον:

Ubuntu 22.04, Python v3.10, TF v2.14, TFF v0.68, Keras v.2.14.1, 4.2GHz CPU, 48GB RAM.

7.2 Αρχιτεκτονική μοντέλων & βελτιστοποίηση παραμέτρων

Το μοντέλο που αποφασίσουμε να εκπαιδεύσουμε με τα δεδομένα μας είναι ένα δίκτυο LSTM. Σε ένα ρεαλιστικό σενάριο ομοσπονδιακής μάθησης για πρόβλεψη παραγωγής, η συλλογή των δεδομένων καθώς και η εκπαίδευση των τοπικών μοντέλων θα γινόταν τοπικά με Edge συσκευές, όπως λ.χ. έξυπνους μετρητές [4]. Η υπολογιστική ισχύ τέτοιων συσκευών είναι αναμενόμενο να είναι περιορισμένη και ένα «ελαφρύ» μοντέλο πιθανώς να είναι προτιμότερο αν δεν παρουσιάζει σημαντική μείωση απόδοσης. Το ενδεχόμενο αυτό αξίζει να διερευνηθεί περεταίρω.

Από την άλλη, τα δίκτυα LSTM αποτελούν την βασική επιλογή νευρωνικού δικτύου σε πληθώρα εργασιών πρόβλεψης φορτίου [3], [14], [54], [55], [62] λόγω της ικανότητάς τους να μοντελοποιούν και να προβλέπουν με ακρίβεια δεδομένα χρονοσειρών, ειδικά όταν πρόκειται για μακροχρόνιες εξαρτήσεις και εποχιακές διακυμάνσεις [33]. Επιπλέον, υπάρχουν τρόποι να πραγματοποιηθεί η ομοσπονδιακή μάθηση με edge τεχνολογίες πέρα από τοπικούς έξυπνους μετρητές όπως με τοπικούς σταθμούς επεξεργασίας [62] οι οποίοι είναι ικανοί να υποστηρίξουν υπολογιστικά σύνθετα μοντέλα.

Για την επιλογή της αρχιτεκτονικής του μοντέλου, πραγματοποιήσαμε βελτιστοποίηση υπερπαραμέτρων κάνοντας χρήση του Optuna Framework [63]. Η διαδικασία που ακολουθήσαμε ήταν να δώσουμε τις επιλογές και το εύρος τιμών που επιθυμούσαμε γι' αυτές και το Optuna έκανε δοκιμές με όλους τους συνδυασμούς έως ότου να καταλήξει στην καλύτερη εναλλακτική, την οποία και χρησιμοποιήσαμε. Στον πίνακα 7.1 έχουμε τους συνδυασμούς και την βέλτιστη επιλογή υπερπαραμέτρων. Η διαδικασία βελτιστοποίησης υπερπαραμέτρων διαδραμάτισε σημαντικό ρόλο και στην διαδικασία της ομαδοποίησης όπως θα δούμε στο κεφάλαιο 7.3.

Hyperparameter	Description	Range of Values	Optimal Choice
n_{past}	Number of past time steps in input sequence	[12, 48]	48
LSTM_units	Number of units in the LSTM layer	[32, 128]	64
dropout_rate	Dropout rate for the first LSTM layer	[0.1, 0.5]	0.1
dense_units	Number of units in the first dense layer	[32, 64]	32
dense_2_units	Number of units in the second dense layer	[32, 64]	32
learning_rate	Learning rate for model training	[0.00001, 0.1]	0.001
batch_size	Number of samples processed per training iteration	[16, 128]	32
epochs	Number of training epochs	[20, 60]	50

Πίνακας 7.1: *Εύρος υπερπαραμέτρων και ιδανική επιλογή για το μοντέλο LSTM*

Η παραπάνω βασική αρχιτεκτονική του μοντέλου είναι ίδια για όλες τις μεθόδους μάθησης. Κατά την διάρκεια της εκπαίδευσης θα ρυθμίσουμε και μερικές ακόμα υπερπαραμέτρους όπως τον αλγόριθμο βελτιστοποίησης, το μέγεθος παρτίδας κ.α.

7.3 Ομαδοποίηση δεδομένων:

Η ομαδοποίηση δεδομένων εφαρμόζεται συχνά σε περιπτώσεις όπου έχουμε να εκπαιδεύσουμε μοντέλα σε μεγάλο όγκο δεδομένων που παρουσιάζουν σημαντικές διακυμάνσεις. Ειδικότερα στην περίπτωση της ομοσπονδιακής μάθησης το ζήτημα της ομαδοποίησης είναι θεμελιώδους σημασίας καθώς με την μέθοδο FedAvg , ο κεντρικός διακομιστής υπολογίζει των μέσο όρο των τοπικών βαρών των μοντέλων. Στην περίπτωση που το προφίλ κάποιου πελάτη διαφέρει αρκετά, είναι πολύ πιθανό να επηρεάσει σε μεγάλο βαθμό το μοντέλο στο δικό του προφίλ και να το απομακρύνει από τους υπόλοιπους. [64]

Στις περισσότερες περιπτώσεις πρόβλεψης φορτίου, η ομαδοποίηση γίνεται με βάση τα ενεργειακά δεδομένα[3], [55], όπως μέση, μέγιστη η ελάχιστη κατανάλωση καθώς πράγματι, τα μοτίβα κατανάλωσης ή παραγωγής ενέργειας αντικατοπτρίζουν άμεσα τα χαρακτηριστικά κίνησης του φορτίου με την πάροδο του χρόνου.

Αλλες παραλλαγές της ομαδοποίησης στην πρόβλεψη παραγωγής μπορεί να περιλαμβάνουν τη χρήση πρόσθετων χαρακτηριστικών ή διαστάσεων πέραν της κατανάλωσης ενέργειας. Για παράδειγμα, η ομαδοποίηση θα μπορούσε να βασίζεται σε δημογραφικές πληροφορίες όπως κοινωνική ή οικονομική κατάσταση [54], [55], καιρικές συνθήκες, γεωγραφικές τοποθεσίες ή τον τύπο των καταναλωτών (όπως οικιστικοί, εμπορικοί ή βιομηχανικοί) [16]. Χαρακτηριστικό παράδειγμα αποτελεί η κατηγοριοποίηση ACORN της CACI όπου κατηγοριοποιεί τον πληθυσμό της Μ. Βρετανίας σε 7 κατηγορίες και 22 δήμο-γεωγραφικές κατηγορίες. [65]

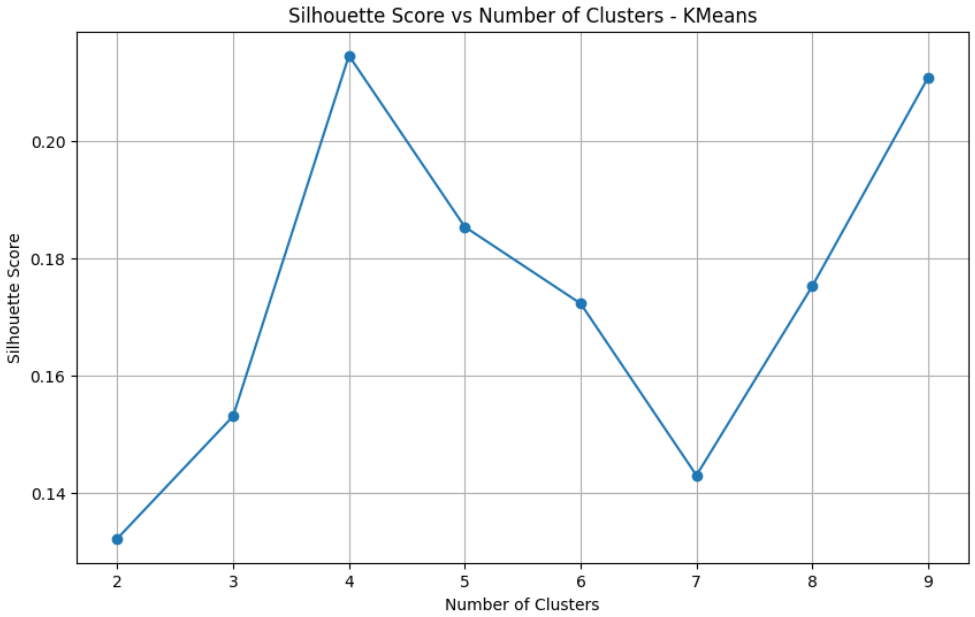
Η ομαδοποίηση με βάση τα ενεργειακά δεδομένα δίνει τα καλύτερα αποτελέσματα στην περίπτωσή μας. Ωστόσο, ενώ δουλεύει καλά στην θεωρία, σε ένα ρεαλιστικό σενάριο όπου θα χρειαζόταν να εφαρμοστεί ομοσπονδιακή μάθηση, θα ήταν αρκετά πιθανό να μην είχαμε πρόσβαση στα δεδομένα παραγωγής.

Στην παρούσα εργασία επιλέξαμε να εφαρμόσουμε μια πρωτοποριακή μέθοδο ομαδοποίησης και να ομαδοποιήσουμε τους παραγωγούς μας με βάση την ομοιότητα των προφίλ παραγωγής τους, βασιζόμενοι στην συμπεριφορά των μοντέλων. Η διαδικασία είναι η εξής:

1. Οι παραγωγοί εκτελούν τοπικά τα μοντέλα πρόβλεψης, συγκεντρώνοντας τις υπερπαραμέτρους του μοντέλου που χρησιμοποιούνται κατά τη διάρκεια της εκπαίδευσης.
2. Αυτές οι υπερπαράμετροι, μαζί με τις εκπαιδευμένες παραμέτρους του μοντέλου, διαβιβάζονται στη συνέχεια πίσω στον κεντρικό διακομιστή, σχηματίζοντας ένα αποθετήριο διαφορετικών προφίλ υπερπαραμέτρων.
3. Αξιοποιώντας αλγορίθμους ομαδοποίησης, ο διακομιστής ομαδοποιεί παρόμοια προφίλ υπερπαραμέτρων, αποκαλύπτοντας μοτίβα και προτιμήσεις μεταξύ των πελατών.

Πιο συγκεκριμένα εφαρμόσαμε τον αλγόριθμο K-means [66] για να ομαδοποιήσουμε τα προφίλ των υπερπαραμέτρων χρησιμοποιώντας την μέθοδο Silhouette Score [67]. Δώσαμε εύρος clusters 2 έως 9 και ο αλγόριθμος έκανε δοκιμές για διαφορετικούς συνδυασμούς clusters

με βάση το μεγαλύτερο Silhouette Score. Στο διάγραμμα 7.1 βλέπουμε τις διαφορετικές επιλογές με βάση το σκορ τους.



Διάγραμμα 7.1: *Silhouetter score ανά αριθμό clusters*

Εφαρμόζοντας λοιπόν την παραπάνω ομαδοποίηση καταλήξαμε στα παρακάτω 4 clusters.

Cluster_0	**Cluster_1**	**Cluster_2**	**Cluster_3**
P_000006	P_000004	P_000013	P_000001
P_001A05	P_000010	P_000107	P_00119D
P_003F4D	P_000410	P_00410C	P_004CB7
P_004613	P_0008D9	P_0048AB	P_005169
P_0080F1	P_00280E	P_004A72	P_007209
P_009FA7	P_005D9A	P_00509B	P_007C29
	P_00680A	P_0056CD	P_0091D7
	P_00B211	P_006A55	

Πίνακας 7.2: Τελικά clusters

7.4 Περιγραφή σεναρίων

Αυτό το κεφάλαιο εισάγει τον πυρήνα της πειραματικής μας διερεύνησης, περιγράφοντας τα ξεχωριστά σενάρια βάσει των οποίων διεξάγεται η συγκριτική μελέτη των μοντέλων και των τρόπων πρόβλεψης φορτίου. Κάθε σενάριο, από την τοπική μάθηση έως την ομοσπονδιακή μάθηση με και χωρίς διαφορική ιδιωτικότητα, είναι σχεδιασμένο για να καταδείξει τις λεπτομέρειες της απόδοσης των μοντέλων, τις επιπτώσεις στην ιδιωτικότητα και την εφαρμοσιμότητά τους σε πραγματικές συνθήκες. Με την ανάλυση αυτών των σεναρίων, στοχεύουμε να παράσχουμε μια ολοκληρωμένη κατανόηση του τρόπου με τον οποίο τα διαφορετικά πρότυπα μάθησης επηρεάζουν την ακρίβεια και την ιδιωτικότητα της πρόβλεψης φορτίου ηλεκτρικής ενέργειας.

7.4.1 Σενάριο Πρώτο: Τοπική Μάθηση

Η τοπική μάθηση, ως το βασικό σενάριο στη μελέτη μας για την πρόβλεψη παραγωγής, επικεντρώνεται στην εκπαίδευση μοντέλων σε δεδομένα από μεμονωμένους κόμβους, αντικατοπτρίζοντας μια παραδοσιακή προσέγγιση μηχανικής μάθησης χωρίς κοινή χρήση δεδομένων σε ολόκληρο το δίκτυο. Αυτή η μέθοδος παρέχει θεμελιώδη σημεία αναφοράς (είναι ουσιαστικά ένα benchmark) για την ακρίβεια, αναδεικνύοντας τους περιορισμούς και τα πλεονεκτήματα των μεμονωμένων μοντέλων στην πρόβλεψη χρονοσειρών σε σχέση με πιο προηγμένες μεθόδους όπως η συγκεντρωτική και η ομοσπονδιακή μάθηση.

7.4.2 Σενάριο Δεύτερο: Συγκεντρωτική Μάθηση

Το σενάριο της συγκεντρωτικής μάθηση είναι ιδιαίτερα σημαντικό για την συγκριτική μας μελέτη, καθώς αντιπροσωπεύει την παραδοσιακή προσέγγιση όπου όλα τα δεδομένα συγκεντρώνονται και υποβάλλονται σε επεξεργασία σε μια ενιαία, κεντρική μονάδα. Η εξέταση αυτού του σεναρίου είναι ουσιώδης για την αξιολόγηση της αποτελεσματικότητας και της επεκτασιμότητας των μοντέλων πρόβλεψης φορτίου όταν υπάρχει απεριόριστη πρόσβαση στα δεδομένα. Η συγκεντρωτική μάθηση αναμένεται να προσφέρει υψηλή ακρίβεια στις προβλέψεις λόγω του ολοκληρωμένου συνόλου δεδομένων που χρησιμοποιεί. Ωστόσο, αναδεικνύει επίσης τις επιπτώσεις που σχετίζονται με την μεταφορά μεγάλου όγκου δεδομένων, όπως το κόστος σε πόρους, ασφάλεια και ιδιωτικότητα [68]. Οπως έχει αποδειχθεί στο παρελθόν, η ασφαλής

διαφύλαξη των συγκεντρωμένων δεδομένων είναι πρόκληση και μέτρα προστασίας όπως η ανωνυμοποίηση έχουν περιορισμένη αποτελεσματικότητα [49], [69].

Συνεπώς, το σενάριο από την μία υπογραμμίζει τα κέρδη αποδοτικότητας που είναι εφικτά με τα συγκεντρωτικά μοντέλα, και από την άλλη φέρνει στο φως σημαντικές ανησυχίες για την προστασία των δεδομένων. Η ουσιαστική συνεισφορά του είναι πως θέτει τις βάσεις για την αξιολόγηση του τρόπου με τον οποίο η ομοσπονδιακή μάθηση και η διαφορική ιδιωτικότητα προσφέρουν λύσεις σε αυτές τις προκλήσεις, παρέχοντας μια βάση για τη σύγκριση των συμβιβασμών μεταξύ απόδοσης και ιδιωτικότητας στην πρόβλεψη παραγωγής

7.4.3 Σενάριο Τρίτο: Ομοσπονδιακή Μάθηση

Το σενάριο της ομοσπονδιακής μάθησης εισάγει την καινοτομία της αποκέντρωσης της διαδικασίας εκπαίδευσης, όπου τα μοντέλα εκπαιδεύονται σε πολλαπλούς κόμβους χωρίς άμεση ανταλλαγή δεδομένων. Η προσέγγιση αυτή αποσκοπεί στην αξιοποίηση της συλλογικής πληροφορίας και των πόρων των κατανεμημένων πηγών δεδομένων, ενώ παράλληλα δίνει προτεραιότητα στην προστασία της ιδιωτικότητας και της ασφάλειας των δεδομένων.

Το παρόν σενάριο διερευνά τη δυνατότητα της FL να επιτύχει ανταγωνιστική ακρίβεια πρόβλεψης φορτίου, προσφέροντας παράλληλα σημαντικά βελτιωμένες εγγυήσεις απορρήτου σε σύγκριση με τα συγκεντρωτικά πρότυπα μάθησης. Πέρα από την προστασία της ιδιωτικότητας, η FL υπόσχεται διάφορα πλεονεκτήματα, όπως τη μείωση των ευπαθειών που συνδέονται με την κοινή χρήση δεδομένων [70], τα δυνητικά χαμηλότερα γενικά έξοδα επικοινωνίας και την αυξημένη ανθεκτικότητα σε σφάλματα που μπορούν να παρουσιάσουν τα συγκεντρωτικά συστήματα [68]. Τα πειράματα αξιολογούν την ικανότητα της FL να ανταποκρίνεται στις εξελισσόμενες απαιτήσεις των εφαρμογών με γνώμονα την προστασία της ιδιωτικής ζωής και την ικανότητά της να επιτυγχάνει μια ισχυρή ισορροπία μεταξύ της προβλεπτικής απόδοσης και της προστασίας των ευαίσθητων δεδομένων των χρηστών.

7.4.4 Σενάριο Τέταρτο: Ομοσπονδιακή Μάθηση ενισχυμένη με Διαφορική Ιδιωτικότητα

Χτίζοντας πάνω στο πλαίσιο της ομοσπονδιακής μάθησης, η εισαγωγή της Διαφορικής Ιδιωτικότητας σηματοδοτεί μια σημαντική εξέλιξη στην προσπάθειά μας για αυξημένη ασφάλεια δεδομένων. Αυτή η διαμόρφωση εισάγει παραμέτρους απορρήτου που προσθέτουν ελεγχόμενο θόρυβο στη διαδικασία εκπαίδευσης του μοντέλου, διασφαλίζοντας ότι οι πληροφορίες που μοιράζονται κατά τη φάση της συνάθροισης δεν θέτουν σε κίνδυνο την ιδιωτικότητα των μεμονωμένων χρηστών [7]. Το σενάριο αποσκοπεί στη διερεύνηση της βέλτιστης ισορροπίας μεταξύ της διατήρησης υψηλής ακρίβειας πρόβλεψης και της επίτευξης ισχυρών εγγυήσεων απορρήτου.

Μέσω αυτού του διευρυμένου σεναρίου, εμβαθύνουμε στις πρακτικές προκλήσεις και τις ευκαιρίες εφαρμογής της διαφορικής ιδιωτικότητας στην ομοσπονδιακή μάθηση για την πρόβλεψη παραγωγής, αξιολογώντας τον αντίκτυπό της στην ισορροπία μεταξύ συλλογικής ευφυΐας και προστασίας της ιδιωτικότητας.

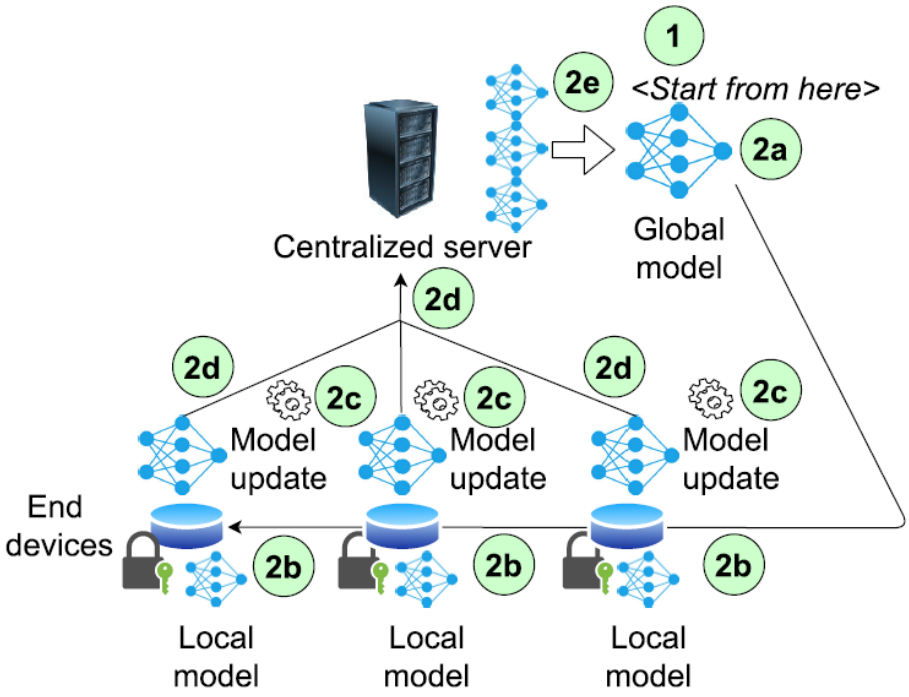
Για την αξιολόγηση του σεναρίου θα συγκρίνουμε την απόδοση του σε σχέση με την απλή ομοσπονδιακή μάθηση, για διαφορετικά επίπεδα θορύβου.

7.5 Υλοποίηση Ομοσπονδιακής Μάθησης

Οπως προαναφέραμε για την υλοποίηση της ομοσπονδιακής μάθησης θα χρησιμοποιήσουμε το framework TFF της Google. Το περιβάλλον της ομοσπονδιακής μάθησης θα αποτελείται από μια προσομοίωση ενός ρεαλιστικού προτύπου. Θα υπάρχει ένας εικονικός κεντρικός διακομιστής ο οποίος θα αναλάβει τον συντονισμό των εικονικών πελατών του δικτύου οι οποίοι θα αντιπροσωπεύονται από τα ιστορικά δεδομένα που έχουμε στην διάθεσή μας. Πιο συγκεκριμένα τα δεδομένα κάθε cluster θα χωριστούν σε ένα ξεχωριστό dataset για κάθε σημείο διανομής που θα αντιπροσωπεύει έναν πελάτη του δικτύου.

Σε ένα ρεαλιστικό σενάριο ομοσπονδιακής μάθησης με εκατοντάδες πελάτες πριν ξεκινήσει κάποιος γύρος μάθησης ο server θα επέλεγε τυχαία ένα υποσύνολο των συνολικών πελατών για να συμμετάσχουν στην εκπαίδευση για εξοικονόμηση πόρων και βελτίωση αποδοτικότητας [71]. Στην περίπτωση μας, τα δεδομένα μας δεν χρειάζεται να διαχωριστούν περαιτέρω οπότε επιλέγουμε να συνεχίσουμε την εκπαίδευση με ολόκληρο το cluster.

Ακολουθεί περιγραφή της διαδικασίας εκπαίδευσης τα βήματα της οποίας παριστάνονται και στην εικόνα 7.1. Εχοντας επιλέξει όλους τους συμμετέχοντες (2a), ο Server τους ενημερώνει για την επιλογή τους και τους αποστέλλει το καθολικό μοντέλο για να ξεκινήσουν την εκπαίδευση (2b). Επιλέξαμε κάθε γύρος μάθησης να περιλαμβάνει μια εποχή εκπαίδευσης για τα τοπικά μοντέλα. Αφού τα μοντέλα εκτελέσουν την εκπαίδευση (2c) θα στείλουν στον Server τα βάρη τους (2d) και θα εκτελεστεί ο αλγόριθμος FedAvg για το aggregation των βαρών (2e). Με βάση την παρακάτω διαδικασία. Θα υπολογίσουμε τις μετρικές εκπαίδευσης, θα ανανεωθεί το καθολικό μοντέλο και θα ξεκινήσει ο επόμενος γύρος από την αρχή.



Εικόνα 7.1: *Περιγραφή διαδικασίας εκπαίδευσης FL [55]*

Για την υλοποίηση της διαδικασίας της ομοσπονδιακής μάθησης ενισχυμένης με διαφορική ιδιωτικότητα, ακολουθήσαμε τα ίδια βήματα με δύο μικρές διαφορές στα βήματα 2c και 2e. Πιο συγκεκριμένα, στην κατασκευή του βρόγχου εκπαίδευσης επιλέγουμε να χρησιμοποιήσουμε έναν dp-aggregator ο οποίος εφαρμόζει τον αλγόριθμο DP-SGD [7]. Ετσι στο βήμα 2c θα έχουμε αποκοπή κλίσεων. Οι κλίσεις δηλαδή που υπολογίζονται για κάθε παρτίδα δεδομένων

αποκόπτονται για να εξασφαλιστεί ότι κανένα σημείο δεδομένων δεν μπορεί να επηρεάσει δυσανάλογα τις ενημερώσεις του μοντέλου. Τέλος στο βήμα 2e όπου γίνεται το aggregation, ο server θα προσθέσει θόρυβο ανάλογο του noise modifier που ορίζουμε.

7.6 Προετοιμασία Εκπαίδευσης.

7.6.1 Κανονικοποίηση δεδομένων και χωρισμός συνόλων

Πριν ξεκινήσουμε την διαδικασία εκπαίδευσης, πραγματοποιούμε την τελευταία επεξεργασία των δεδομένων εκπαίδευσης, την κανονικοποίηση (scaling) και τον χωρισμό σε σύνολα εκπαίδευσης και δοκιμής (train-test split). Στην περίπτωσή μας, to scaling είναι εξαιρετικά σημαντικό καθώς θα εργαστούμε με δεδομένα ανανεώσιμων πηγών ενέργειας οι καμπύλες των οποίων δεν ακολουθούν Γκαουσιανές κατανομές. Η απόκλιση από την εν λόγω κατανομή μπορεί να αποδοθεί σε διάφορους παράγοντες, όπως ο δείκτης καθαρού ουρανού, η αζιμουθιακή γωνία, η σκίαση, η τοποθέτηση των ηλιακών συλλεκτών κ.α. [3]. Με την κανονικοποίηση πετυχαίνουμε μειώσουμε την επίδραση των ακραίων τιμών και να βελτιώσουμε την σύγκλιση του μοντέλου. Ακόμη, το μοντέλο μας θα μπορεί να γενικευτεί καλύτερα σε νέα δεδομένα.

Για την κανονικοποίηση θα χρησιμοποιήσουμε την υλοποίηση του MinMaxScaler [19] από την βιβλιοθήκη Scikit-learn [72], μετασχηματίζοντας τα δεδομένα μας, με αναδιαβάθμιση της τιμής ενέργειάς τους σε ένα εύρος από 0 έως 1.

Μετά την κλιμάκωση, χρησιμοποιώ τη συνάρτηση train_test_split (επίσης από την Scikit-learn) για να χωρίσω το σύνολο δεδομένων σε σύνολα εκπαίδευσης και δοκιμής. Αυτός ο διαχωρισμός επιτρέπει στα μοντέλα να μαθαίνουν από το σύνολο εκπαίδευσης και να αξιολογούνται δίκαια σε αθέατα δεδομένα στο σύνολο δοκιμής. Το train set αποτελούσε το 80% των συνολικών παρατηρήσεων για κάθε παραγωγό ενώ το test set το υπόλοιπο 20%.

7.6.2 Προετοιμασία δεδομένων κινούμενου παραθύρου.

Εχοντας ορίσει το μοντέλο μας μένει να ετοιμάσουμε τα δεδομένα για την εκπαίδευσή του. Καθώς εργαζόμαστε στο framework TensorFlow, πρέπει να μετατρέψουμε τα δεδομένα μας σε τανυστές (tensors). Οι tensors είναι μια μορφή πινάκων πολλών διαστάσεων. Για την πρόβλεψη

χρονοσειρών θα χρειαστεί να μετατρέψουμε τα δεδομένα μας σε πίνακες 3 διαστάσεων με την μέθοδο κυλιόμενου παραθύρου που περιγράψαμε στο κεφάλαιο 2.3.4.

Πιο αναλυτικά, θα δημιουργήσουμε 2 πίνακες, τον πίνακα X ή πίνακα παραθύρου για την είσοδο και τον πίνακα Y ή πίνακα στόχου για τον στόχο πρόβλεψης. Ο πίνακας X έχει 3 διαστάσεις: x_1, x_2, x_3.

Η διάσταση x_1 αντιπροσωπεύει τον συνολικό αριθμό των παραθύρων πρόβλεψης που δημιουργήσαμε, όσα δηλαδή και οι παρατηρήσεις μας.

Η διάσταση x_2 είναι το μέγεθος παραθύρου, που στην περίπτωσή μας έχει οριστεί σε 48. Αυτό σημαίνει ότι το μοντέλο θα χρησιμοποιήσει τις 48 πιο πρόσφατες παρατηρήσεις (δηλαδή μια ημέρα) για να κάνει την επόμενη πρόβλεψη του.

Η διάσταση x_3 είναι ο αριθμός τον στόχων πρόβλεψης, στην περίπτωσή μας 1, η παραγόμενη ενέργεια της χρονικής περιόδου.

Ο αριθμός των διαστάσεων του πίνακα Y εξαρτάται από τον αριθμό των στόχων. Στην περίπτωση μας 1. Η τιμή του θα είναι η αμέσως επόμενη τιμή του ¨παραθύρου". Με απλούστερους όρους, οργανώνουμε τα δεδομένα μας έτσι ώστε το μοντέλο να εξετάσει τις μετρήσεις ολόκληρης της μέρας ώστε να προβλέψει την ενέργεια για την επόμενη παρατήρηση.

7.7 Διαδικασία Εκπαίδευσης

Κατά την εκτέλεση των πειραμάτων εστιάσαμε σε μια διεξοδική διαδικασία εκπαίδευσης και δοκιμής μοντέλων για την επαλήθευση των υποθέσεών μας σε διαφορετικά σενάρια μάθησης. Για κάθε μοντέλο, προσπαθήσαμε να διατηρήσουμε τις παραμέτρους εκπαίδευσης και επαλήθευσης σταθερές ανεξαρτήτως τύπου μάθησης ώστε να εξασφαλίσουμε δίκαιες συγκρίσεις.

Η φάση εκπαίδευσης, που διεξήχθη σε 50 εποχές, είχε ως στόχο τη βελτιστοποίηση των μοντέλων για ακριβή πρόβλεψη παραγωγής ηλεκτρικής ενέργειας την επόμενη ημέρα. Δοκιμάζοντας άμεσα αυτά τα μοντέλα σε αθέατα δεδομένα, αξιολογήσαμε την ακρίβεια πρόβλεψης και τη γενικευσιμότητά τους, εξασφαλίζοντας μια ισχυρή αξιολόγηση των επιδόσεών τους.

Παρακάτω θα αναλύσουμε τις παραμέτρους των μοντέλων που χρησιμοποιήθηκαν και τον λόγο επιλογής τους για κάθε σενάριο, ξεκινώντας από την τοπική μάθηση.

Στην εκπαίδευση προσθέσαμε έναν μηχανισμό early stopping για αποφυγή της υπερπροσαρμογής (overfitting). Με αυτόν τον μηχανισμό θα παρακολουθούμε την συνάρτηση απώλειας με μια περίοδο υπομονής 10 εποχών. Σε περίπτωση που η συνάρτηση απώλειας σταματήσει να μειώνεται για 10 συνεχόμενες εποχές, η εκπαίδευση σταματάει και το μοντέλο επιστρέφει στην καλύτερη προηγούμενη κατάσταση. Το μέγεθος παρτίδας (batch size) είναι 32 λόγους συμβατότητας με τους περιορισμούς της ομοσπονδιακής μάθησης που θα δούμε παρακάτω. Αφού ολοκληρωθεί η εκπαίδευση, υλοποιούμε τις προβλέψεις μας, με την μέθοδο predict του Keras. Τέλος, αξιολογούμε τα αποτελέσματά μας, συγκρίνοντας τις προβλέψεις μας με τις πραγματικές τιμές του συνόλου δοκιμής.

Η παραπάνω βασική διαδικασία έχει μερικές μεταβολές στον τρόπο διαχείρισης δεδομένων και εκπαίδευσης ανάλογα με το είδος μάθησης. Πιο συγκεκριμένα:

7.7.1 Τοπική & Συγκεντρωτική Εκπαίδευση

Για Τοπική Μάθηση με $\mathcal{N}$ παραγωγούς:

- Δεδομένα $\mathcal{D}$: Τα ιστορικά μας δεδομένα χωρίζονται ανά παραγωγό. Το scaling και ο διαχωρισμός των train & test sets γίνονται σε τοπικό επίπεδο Εστω σύνολο $\mathcal{D} = \{\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n\}$ που συμβολίζει τη συλλογή δεδομένων από n παραγωγούς. Κάθε D_i χωρίζεται σε ένα υποσύνολο εκπαίδευσης $D_{i,train}$ και ένα υποσύνολο δοκιμής $D_{i,test}$. Δηλαδή:

$$D_i = D_{i,\text{train}} \cup D_{i,\text{test}}, \quad \forall i \in \{1,2,\dots,n\}$$

- Εκπαίδευση: Κάθε παραγωγός εκπαιδεύει ένα μοντέλο $\mathcal{M}_i$ χρησιμοποιώντας τα δεδομένα εκπαίδευσής του $D_{i,train}$:

$$M_i = \text{Train}\big(D_{i,\text{train}}\big), \quad \forall i \in \{1,2,\dots,n\}$$

- Αξιολόγηση: Κάθε μοντέλο εφαρμόζεται στο αντίστοιχο σύνολο δοκιμής $D_{i,\text{test}}$, παράγοντας τις προβλέψεις $\mathcal{P}$:

$$P_i = M_i(D_{i,\text{test}}), \quad \forall i \in \{1,2,\dots,n\}$$

Για Συγκεντρωτική Μάθηση με $\mathcal{N}$ παραγωγούς:

- Δεδομένα $\mathcal{D}$: Τα ιστορικά μας δεδομένα συγκεντρώνονται στον κεντρικό διακομιστή. Το scaling και ο διαχωρισμός των train & test sets γίνονται σε κεντρικό επίπεδο. Εστω σύνολο $\mathcal{D} = \{\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n\}$ που συμβολίζει τη συλλογή δεδομένων από n παραγωγούς. Κάθε D_i χωρίζεται σε ένα υποσύνολο εκπαίδευσης $D_{i,train}$ και ένα υποσύνολο δοκιμής $D_{i,test}$. Δηλαδή:

  $$D_i = D_{i,\text{train}} \cup D_{i,\text{test}}, \quad \forall i \in \{1,2,\dots,n\}$$

  Στην συνέχεια, δημιουργείται το συγκεντρωτικό σύνολο εκπαίδευσης $D_{\text{central,train}}$ με τα δεδομένα όλων των παραγωγών:

  $$D_{\text{central,train}} = \bigcup_{i=1}^{n} D_{i,\text{train}}$$

- Εκπαίδευση: Ενα μόνο συγκεντρωτικό μοντέλο M_{central} εκπαιδεύεται στο συγκεντρωτικό σύνολο εκπαίδευσης:

  $$M_{\text{central}} = \text{Train}(D_{\text{central,train}})$$

- Αξιολόγηση: Κάθε μοντέλο εφαρμόζεται στο αντίστοιχο σύνολο δοκιμής $D_{i,\text{test}}$, παράγοντας τις προβλέψεις $\mathcal{P}$:

  $$P_i = M_{\text{central}}(D_{i,\text{test}}), \quad \forall i \in \{1,2,\dots,n\}$$

7.7.2 Ομοσπονδιακή Εκπαίδευση

Η ομοσπονδιακή μάθηση, λόγω του επιλεχθέντος framework αλλά και της ιδιαίτερης φύσης της έχει ελαφρώς διαφορετική διαδικασία τόσο στην διαχείριση των δεδομένων όσο και στην εκπαίδευση του μοντέλου.

Αρχικά, για την διαχείριση των δεδομένων χρειάζεται ένα παραπάνω βήμα προεπεξεργασίας. Τα train και test data πρέπει να μετασχηματιστούν σε Tensorflow Datasets. Αφού εκτελέσουμε αυτό το βήμα με χρήση των βιβλιοθηκών του TF, έχουμε να επιλέξουμε τα δεδομένα προς εκπαίδευση. Υπό ρεαλιστικές συνθήκες, σε κάθε γύρο εκπαίδευσης ο server θα επέλεγε τυχαία ένα υποσύνολο των διαθέσιμων πελατών για να εκπαιδεύσει το καθολικό μοντέλο, αφού οι πελάτες θα είναι πολυάριθμοι [71]. Τα πειραματικά δεδομένα που έχουμε στην διάθεσή μας ωστόσο είναι διαχειρίσιμα, οπότε επιλέγουμε να συμμετάσχουν όλοι οι πελάτες στην εκπαίδευση.

Η επόμενη διαφορά σε σχέση με την υλοποίηση της εκπαίδευσης είναι ότι έχουμε έναν optimizer για τους clients και έναν optimizer για τον server. Ο client optimizer χρησιμοποιείται μόνο για τον υπολογισμό τοπικών ενημερώσεων του μοντέλου σε κάθε πελάτη. Ο server optimizer εφαρμόζει τη μέση ενημέρωση στο παγκόσμιο μοντέλο στο διακομιστή. Σε αντίθεση με τα υπόλοιπα μοντέλα που χρησιμοποιούσαν την συνάρτηση βελτιστοποίησης Adam, για την ομοσπονδιακή μάθηση θα χρησιμοποιήσουμε την συνάρτηση βελτιστοποίησης SGD με learning rate 0.001 για τους clients και 0.99 για τον server αντίστοιχα.

Στην συνέχεια έχουμε να επιλέξουμε τον aggregator με βάση τον οποίο θα γίνει η συγκέντρωση των βαρών στον server. Επιλέγουμε να χρησιμοποιήσουμε τον κλασικό FedAvg [6] και χρησιμοποιούμε την υλοποίηση του TFF.

Τέλος, έχουμε να επιλέξουμε το πόσες εποχές τοπικής εκπαίδευσης θα περιέχει κάθε γύρος federated training. Για ευκολία στις συγκρίσεις με τις άλλες μεθόδους, θα επιλέξουμε μια εποχή για κάθε γύρο.

Με το πέρας της εκπαίδευσης η υπόλοιπη διαδικασία του evaluation είναι αντίστοιχη με την συγκεντρωτική μάθηση.

Συνοπτική παρουσίαση για N παραγωγούς:

- Δεδομένα:

$$D_i = D_{i,\text{train}} \cup D_{i,\text{test}}, \forall i$$

- Εκπαίδευση:

$$M_i = \text{Train}\left(D_{i,\text{train}}\right), \forall i$$

- Συγκέντρωση βαρών:

$$M_{\text{global}} = \text{Aggregate}(\{M_1, M_2, \dots, M_n\})$$

- Διαμοιρασμός καθολικού μοντέλου και αξιολόγηση:

$$P_i = M_{\text{global}}(D_{i,\text{test}}), \forall i$$

7.7.3 Ομοσπονδιακή εκπαίδευση με Διαφορική Ιδιωτικότητα

Στην περίπτωση της ομοσπονδιακής μάθησης ενισχυμένης με διαφορική ιδιωτικότητα, η διαδικασία της εκπαίδευσης αλλάζει ελαφρώς. Για να επιτύχουμε τις απαραίτητες εγγυήσεις ασφάλειας θα πρέπει να αλλάξουμε τον αλγόριθμο FedAvg σε 2 σημεία. Η πρώτη αλλαγή έχει να κάνει με το στάδιο ενημέρωσης του κεντρικού server για τα τοπικά βάρη καθώς θα πρέπει να γίνει clipping των βαρών. Η δεύτερη έχει να κάνει με την διαδικασία συγκέντρωσης των βαρών στον τοπικό server [73].

Και οι δύο αλλαγές επιτυγχάνονται περνώντας τις κατάλληλες παραμέτρους στο χτίσιμο της Federated training διαδικασίας του TFF. Πιο συγκεκριμένα, επιλέγουμε σαν aggregator να περάσουμε τον dp_aggregator του TFF ο οποίος υλοποιεί τον αλγόριθμο FedAvg με adaptive clipping [74] και προσθέτοντας θόρυβο στο aggregation.

Η λειτουργία του dp_aggregator επηρεάζεται από δύο παραμέτρους που ορίζει ο χρήστης. Τον αριθμό των clients που συμμετέχουν σε κάθε γύρο εκπαίδευσης (στην περίπτωσή μας όλοι) και τον συντελεστή θορύβου.

Ο συντελεστής θορύβου έχει να κάνει με το ποσοστό θορύβου που θέλουμε να προσθέσουμε στα δεδομένα μας. Υψηλός θόρυβος σημαίνει μεγαλύτερη ασφάλεια αλλά πτώση στην ποιότητα των προβλέψεων. Κάναμε διάφορες δοκιμές για τον ιδανικό τροποποιητή θορύβου που θα χρησιμοποιήσουμε και καταλήξαμε πως το ανώτερο όριο που μπορούμε να φτάσουμε είναι το 0.2 καθώς από εκείνο το σημείο και έπειτα τα δεδομένα χαλάνε πολύ έντονα.

Οι Wang et al. [75] έχουν υλοποιήσει μια μέθοδο για να ρυθμίσεις με ακρίβεια τις παραμέτρους του dp_aggregator ή οποία συμπεριλαμβάνεται στο TF στην βιβλιοθήκη DP_accounting [59].

Η μέθοδος λαμβάνει ως ορίσματα τον αριθμό των clients που συμμετέχουν στην εκπαίδευση και το επιθυμητό επίπεδο διαφορικής ασφάλειας και επιστρέφει τον ιδανικό συνδυασμό συντελεστή θορύβου και πελατών ανά γύρο για να επιτευχθεί με την μικρότερη δυνατή απώλεια.

Ωστόσο είναι σχεδιασμένη για ρεαλιστικές εφαρμογές με χιλιάδες clients και η χρήση της δεν έχει νόημα στα πλαίσια αυτής της εργασίας. Δεδομένου ότι ο αριθμός τον πελατών ανά

γύρο είναι σταθερός (αφού τους χρησιμοποιούμε όλους) θα πραγματοποιήσουμε εκπαιδεύσεις με διαφορετικούς συντελεστές θορύβου ώστε να αναλύσουμε τον αντίκτυπό τους στις προβλέψεις.

Τέλος, ακολουθεί η διαδικασία της εκπαίδευσης για N παραγωγούς συνοπτικά:

- Δεδομένα:
$$D_i = D_{i,\text{train}} \cup D_{i,\text{test}}, \forall i$$
- Εκπαίδευση:
$$M_i^{DP} = \text{Train}_{DP}\left(D_{i,\text{train}}\right), \forall i$$
- Συγκέντρωση βαρών και καθολικό μοντέλο:
$$M_{\text{global}}^{DP} = \text{Aggregate}_{DP}(\{M_1^{DP}, M_2^{DP}, \dots, M_n^{DP}\})$$
- Διαμοιρασμός καθολικού μοντέλου και αξιολόγηση:
$$P_i^{DP} = M_{\text{global}}^{DP}\left(D_{i,\text{test}}\right), \forall i$$

7.8 Μετρικές αξιολόγησης μοντέλων

Η επιλογή των κατάλληλων μετρικών απόδοσης είναι καθοριστική για την αξιολόγηση της αποτελεσματικότητας των προγνωστικών μοντέλων στην πρόβλεψη παραγωγής. Η παρούσα μελέτη χρησιμοποιεί το μέσο απόλυτο σφάλμα (MAE) ως κύρια συνάρτηση απωλειών (loss function) και το μέσο τετραγωνικό σφάλμα (RMSE) για την αξιολόγηση των επιδόσεων, που επιλέχθηκαν για τα πλεονεκτήματά τους στην αποτύπωση της ακρίβειας και της αξιοπιστίας των προβλέψεων.

7.8.1 Μέσο Απόλυτο Σφάλμα

Το μέσο απόλυτο σφάλμα (mean absolute error - MAE) παρέχει ένα απλό μέτρο του μέσου μεγέθους του σφάλματος, μετρώντας την απόλυτη διαφορά μεταξύ της πραγματικής και της προβλεπόμενης εξόδου. Αυτό το καθιστά ιδιαίτερα χρήσιμο για την κατανόηση του μέσου αντίκτυπου των σφαλμάτων στην πρόβλεψη. Πιο συγκεκριμένα, ορίζεται ως ο μέσος όρος των απόλυτων διαφορών μεταξύ των προβλεπόμενων τιμών και των πραγματικών τιμών και δίνεται από τον παρακάτω τύπο:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Οπου n, είναι ο αριθμός των παρατηρήσεων, y_i η πραγματική τιμή, $\hat{y}_i$ η προβλεπόμενη τιμή και $|y_i - \hat{y}_i|$ το απόλυτο σφάλμα μεταξύ τους.

Το MAE έχει το πλεονέκτημα πως αντιστοιχεί άμεσα το μέσο σφάλμα στις ίδιες μονάδες με τα δεδομένα διευκολύνοντας τη βελτιστοποίηση του μοντέλου με έμφαση στην ελαχιστοποίηση των μέσων σφαλμάτων πρόβλεψης. Επίσης είναι αρκετά ανθεκτικό στις ακραίες τιμές ένα χαρακτηριστικό του συνόλου δεδομένων μας [3].

7.8.2 Μέσο Τετραγωνικό Σφάλμα

Το μέσο τετραγωνικό σφάλμα (RMSE) χρησιμοποιείται ως μέτρο απόδοσης για την αξιολόγηση της ακρίβειας των μοντέλων. Το RMSE είναι η τετραγωνική ρίζα του μέσου όρου των τετραγωνικών διαφορών μεταξύ πρόβλεψης και πραγματικής παρατήρησης και δίνεται από τον παρακάτω τύπο.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Οπου n, είναι ο αριθμός των παρατηρήσεων, y_i η πραγματική τιμή, $\hat{y}_i$ η προβλεπόμενη τιμή και $(y_i - \hat{y}_i)^2$ το τετραγωνικό σφάλμα μεταξύ τους.

Η διαδικασία τετραγωνισμού τιμωρεί αυστηρότερα τα μεγαλύτερα σφάλματα, παρέχοντας ένα πιο ευαίσθητο μέτρο απόδοσης όταν τα μεγάλα σφάλματα είναι ιδιαίτερα ανεπιθύμητα [76]. Στο πλαίσιο της εργασίας, όπου υποεκτιμήσεις ή υπερεκτιμήσεις ενέργειας μπορούν να οδηγήσουν σε σημαντικές λειτουργικές ανεπάρκειες ή οικονομικές απώλειες, το RMSE προσφέρει ένα πολύτιμο μέτρο του πόσο σημαντικά αποκλίνουν οι προβλέψεις του μοντέλου από τις πραγματικές τιμές.

8 Ανάλυση Αποτελεσμάτων

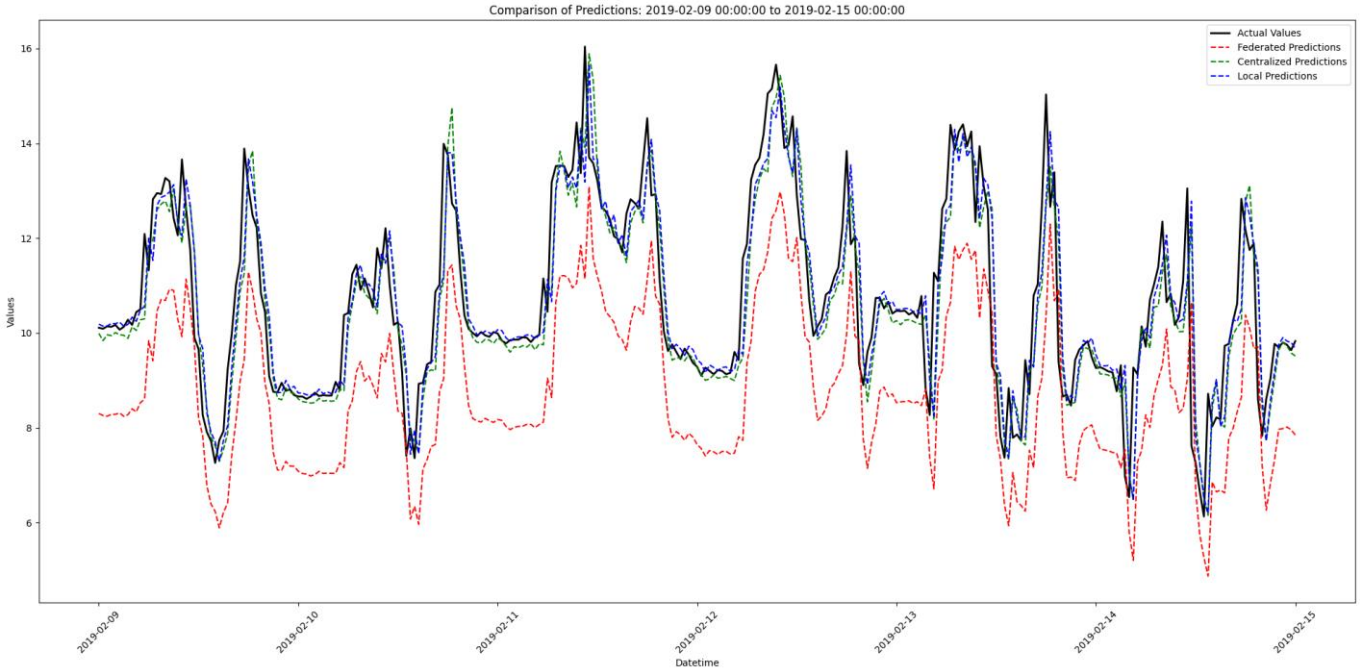
Με την ολοκλήρωση της εκπαίδευσης προχωράμε στην αξιολόγηση των αποτελεσμάτων κάνοντας προβλέψεις στα train & test set. Οι πιο σημαντικές ενδείξεις προκύπτουν από τα αποτελέσματα που προκύπτουν από τα δεδομένα του test set τα οποία και είναι άγνωστα για το μοντέλο. Ωστόσο, η διαφορά στην ποιότητα των προβλέψεων μεταξύ των δύο συνόλων μας δίνει στοιχεία για την ποιότητα της εκπαίδευσής μας.

Για την αξιολόγηση των αποτελεσμάτων μας θα συγκρίνουμε τις μετρικές απόδοσης για όλες τις μεθόδους μάθησης για κάθε cluster ξεχωριστά και στην συνέχεια για τον μέσο όρο όλων των clusters. Ακόμη, θα εξετάσουμε τον αντίκτυπο της ομαδοποίησης στην συγκεντρωτική και την ομοσπονδιακή μάθηση, συγκρίνοντας τα αποτελέσματα των εκπαιδευμένων σε ομάδες μοντέλα με δύο μοντέλα εκπαιδευμένα σε ολόκληρο το σύνολο δεδομένων.

8.1 Συγκριτική απόδοση μεθόδων μάθησης

Σε αυτή την ενότητα, πραγματοποιούμε μια λεπτομερή σύγκριση των τοπικών, συγκεντρωτικών και ομοσπονδιακών μεθόδων μάθησης με βάση την απόδοσή τους στην πρόβλεψη παραγωγής. Εστιάζοντας στο μέσο απόλυτο σφάλμα (MAE) και στο μέσο τετραγωνικό σφάλμα (RMSE), στοχεύουμε στην αξιολόγηση της ακρίβειας και της αξιοπιστίας κάθε προσέγγισης. Μέσω πινάκων και γραφικών παραστάσεων, θα παρουσιάσουμε το μέσο MAE και RMSE για κάθε cluster, παρέχοντας μια σαφή εικόνα για το πώς κάθε μέθοδος μάθησης ανταπεξέρχεται στην πρακτική εφαρμογή.

Αντιπαραβάλλοντας αυτά τα παραδείγματα μάθησης, στόχος μας είναι να αναδείξουμε τα αντίστοιχα πλεονεκτήματα και αδυναμίες τους, προσφέροντας πληροφορίες για τις δυνατότητες της ομοσπονδιακής μάθησης σε σύγκριση με τις παραδοσιακές μεθόδους.

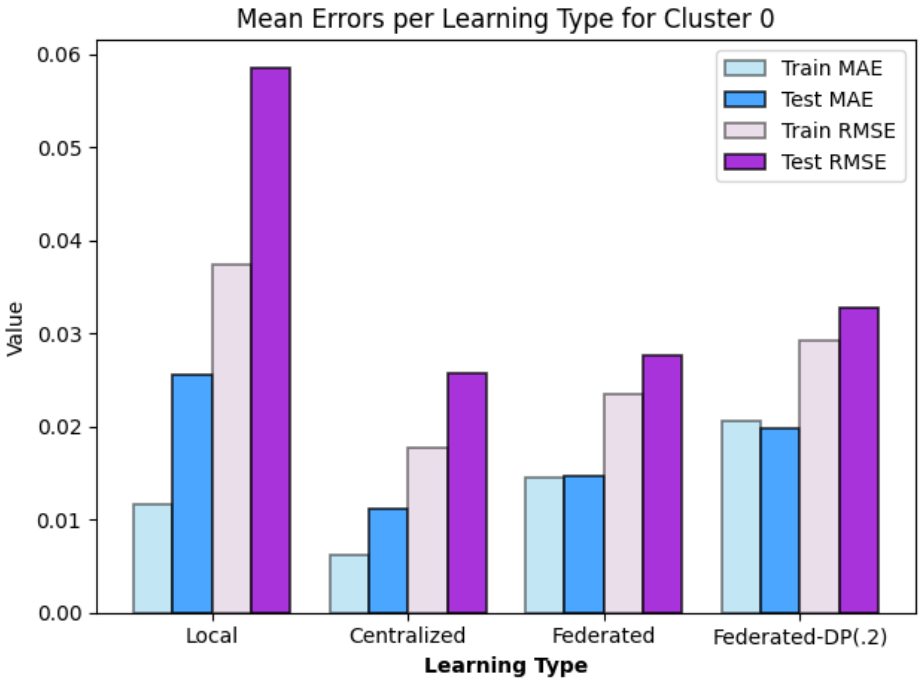


Διάγραμμα 8.1: *Προβλέψεις παραγωγής Τοπικής, Συγκεντρωτικής και Ομοσπονδιακής μάθησης για τον παραγωγό P_0091D7*

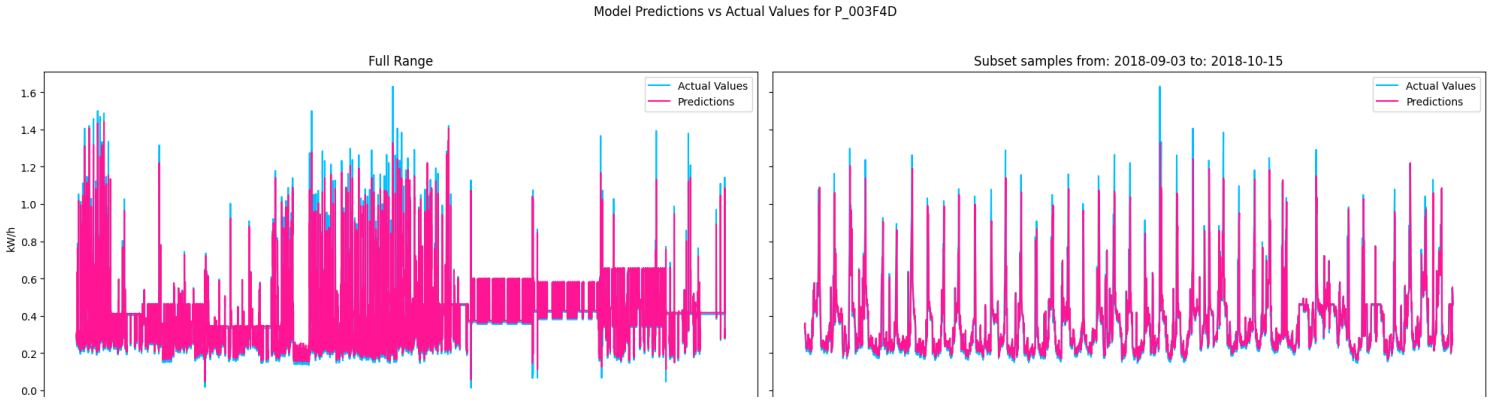
Στον πίνακα 8.1 παρατηρούμε το μέσο σφάλμα για όλους τους παραγωγούς του cluster0 ανά μέθοδο μάθησης για με αξιολόγηση στο σύνολο εκπαίδευσης και στο σύνολο δοκιμής. Για καλύτερη οπτικοποίηση των αποτελεσμάτων κατασκευάζουμε και την γραφική απεικόνιση των παραπάνω δεδομένων στο διάγραμμα 8.1.

	Local		Centralized		Federated		Federated-DP(.2)	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Train	0.0116	0.0374	0.0061	0.0177	0.0145	0.0235	0.0206	0.0292
Test	0.0255	0.0586	0.0111	0.0257	0.0147	0.0277	0.0198	0.0327

Πίνακας 8.1: *Μέσο σφάλμα για κάθε παραγωγό του Cluster 0.*



Διάγραμμα 8.2: *Μέσο σφάλμα για κάθε παραγωγό του Cluster 0*

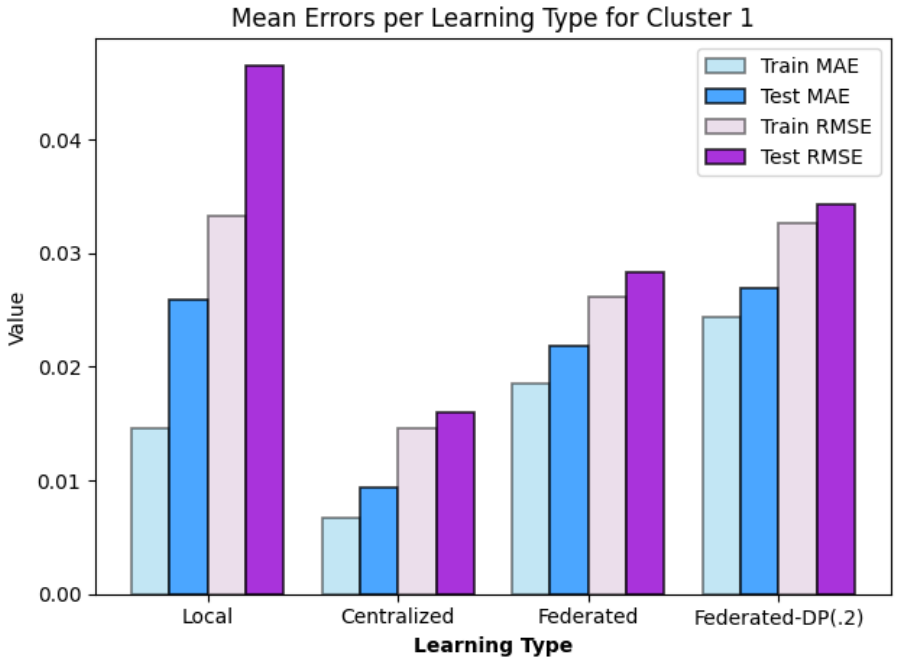


Διάγραμμα 8.3: *Προβλέψεις παραγωγής Τοπικής Μάθησης για τον παραγωγό P_003F4D από το Cluster 0*

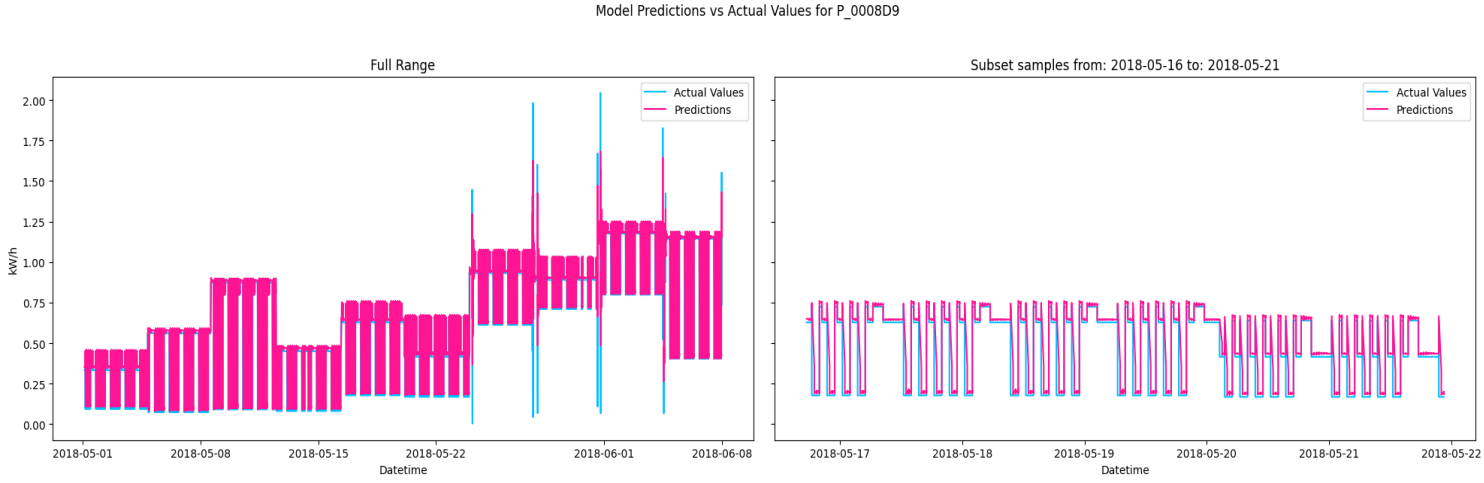
Ακολουθούμε την ίδια διαδικασία και για τα υπόλοιπα 3 clusters.

	Local		Centralized		Federated		Federated-DP(.2)	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Train	0.0146	0.0333	0.0067	0.0146	0.0185	0.0261	0.0243	0.0326
Test	0.0259	0.0466	0.0094	0.0159	0.0218	0.0283	0.0269	0.0343

Πίνακας 8.2: Μέσο σφάλμα για κάθε παραγωγό του Cluster 1



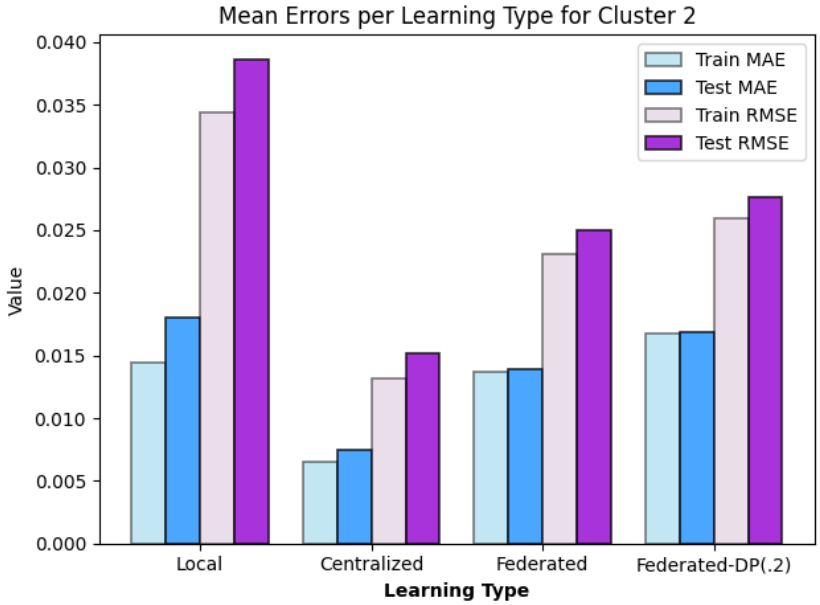
Διάγραμμα 8.4: *Μέσο σφάλμα για κάθε παραγωγό του Cluster 1*



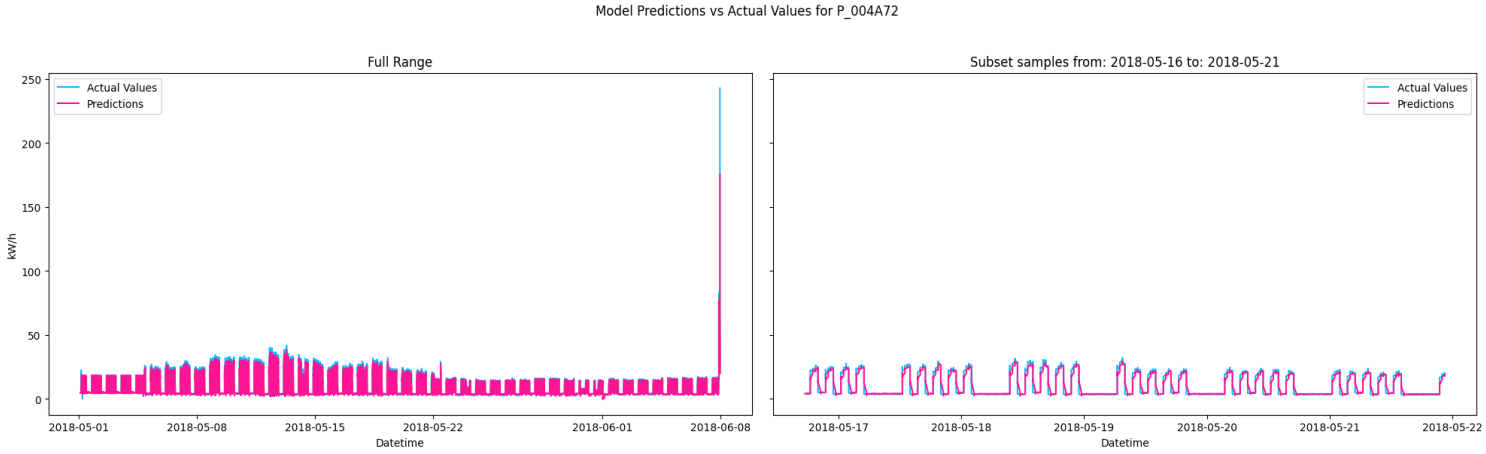
Διάγραμμα 8.5: *Προβλέψεις παραγωγής με συγκεντρωτική μάθηση για τον παραγωγό P_0008D9 από το Cluster 1*

	Local		Centralized		Federated		Federated-DP(.2)	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Train	0.0144	0.0344	0.0065	0.0132	0.0137	0.0231	0.0167	0.0259
Test	0.0180	0.0387	0.0075	0.0152	0.0139	0.0250	0.0169	0.0276

Πίνακας 8.3: Μέσο σφάλμα για κάθε παραγωγό του Cluster 2



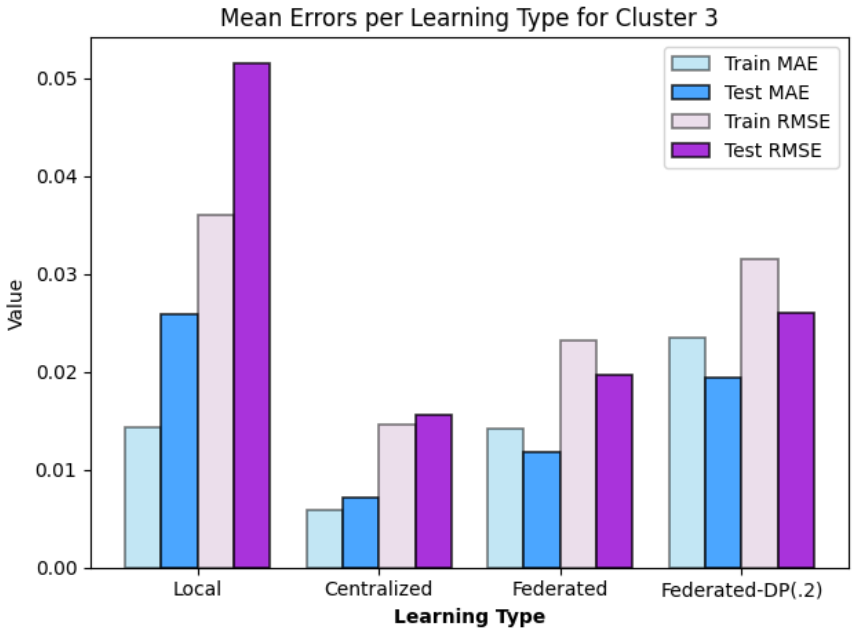
Διάγραμμα 8.6: *Μέσο σφάλμα για κάθε παραγωγό του Cluster 2*



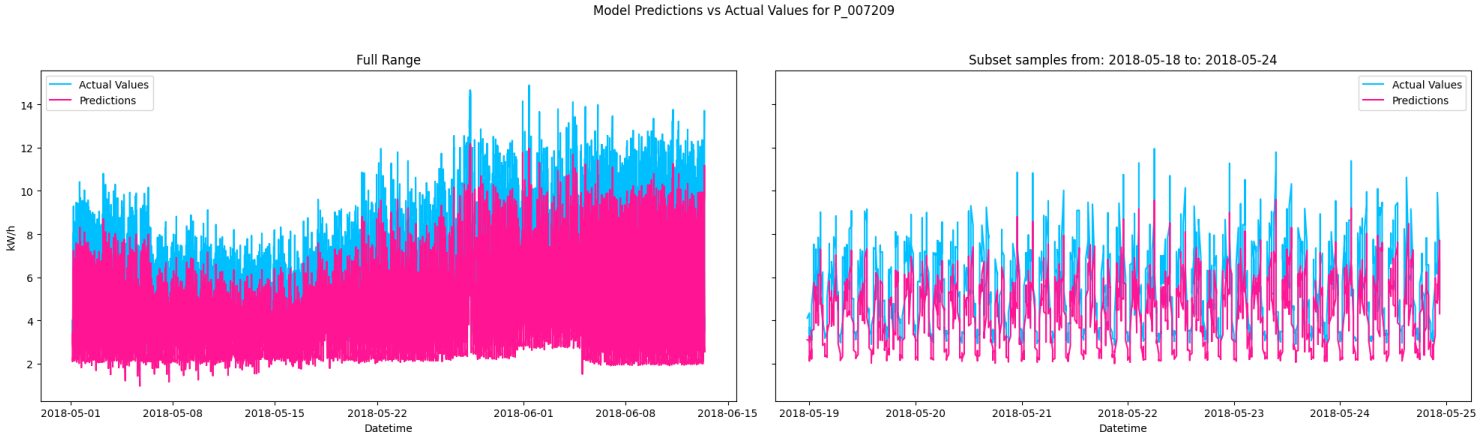
Διάγραμμα 8.7: *Προβλέψεις παραγωγής για τον παραγωγό P_004A72 από το Cluster 2*

	Local		Centralized		Federated		Federated-DP(.2)	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Train	0.0143	0.0361	0.0058	0.0146	0.0141	0.0232	0.0235	0.0315
Test	0.0260	0.0516	0.0072	0.0155	0.0118	0.0196	0.0194	0.0261

Πίνακας 8.4: *Μέσο σφάλμα για κάθε παραγωγό του Cluster 3*



Διάγραμμα 8.8: *Μέσο σφάλμα για κάθε παραγωγό του Cluster*

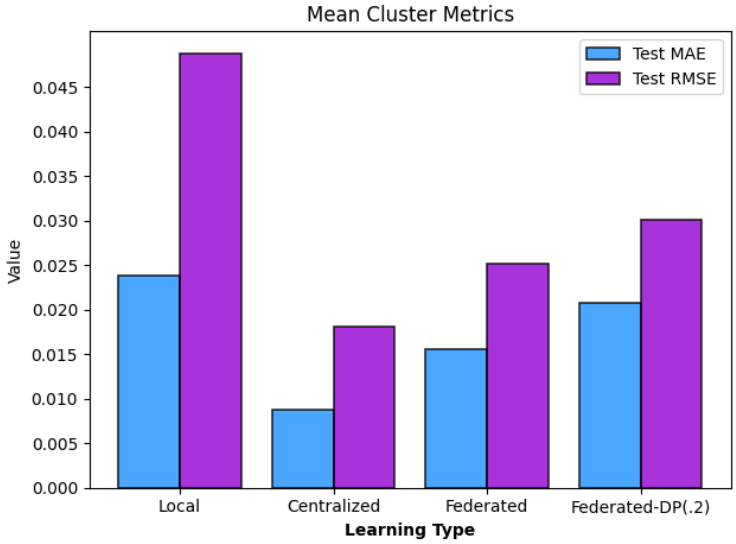


Διάγραμμα 8.9: *Προβλέψεις παραγωγής για τον παραγωγό P_007209 από το Cluster 3*

Και τα 4 clusters ακολουθούν παρόμοια κατανομή οπότε είναι ασφαλές να υπολογίσουμε τον μέσο όρο και των τεσσάρων ώστε να έχουμε μια συνολική εικόνα για να αντλήσουμε συμπεράσματα.

Model	MAE	RMSE
Local	0.0238	0.0488
Centralized	0.0088	0.0181
Federated	0.0155	0.0252
Federated-DP(.2)	0.0208	0.0302

Πίνακας 8.5: *Σφάλματα μέσου όρου clusters*



Διάγραμμα 8.10: *Σφάλματα μέσου όρου clusters*

Ο πίνακας 8.5 και το διάγραμμα 8.10 συνοψίζουν την απόδοση των μεθόδων κατά την διάρκεια της εκπαίδευσης. Με βάση αυτά μπορούμε να εξάγουμε μερικά ασφαλή συμπεράσματα.

Αρχικά, είναι εμφανές πως οι μέθοδοι μάθησης που χρησιμοποιούν δεδομένα συνεργατικά έχουν πολύ καλύτερη απόδοση από την τοπική με τα μεμονωμένα δεδομένα. Πιο συγκεκριμένα, η υπεροχή της ομοσπονδιακής μάθησης έναντι της τοπικής μάθησης υποδηλώνει την αξία της συνεργασίας και της ανταλλαγής γνώσεων μεταξύ κατανεμημένων μοντέλων, ακόμη και χωρίς την πλήρη συγκέντρωση των δεδομένων.

Επιπλέον, το γεγονός πως η συγκεντρωτική μάθηση (CL) υπερτερεί έναντι της FL αναδεικνύει τα πλεονεκτήματα ενός μοντέλου που εκπαιδεύεται σε ένα ενοποιημένο αποθετήριο

δεδομένων. Αυτή η υπεροχή μπορεί να αιτιολογηθεί από την ικανότητα του συγκεντρωτικού μοντέλου να αξιοποιεί άμεσα τα συγκεντρωτικά δεδομένα από όλους τους κόμβους, σε αντίθεση με την FL, όπου τα μοντέλα εκπαιδεύονται τοπικά σε κόμβους με δυνητικά μη ανεξάρτητα και πανομοιότυπα κατανεμημένα δεδομένα (non i.i.d) και στη συνέχεια υπολογίζεται ο μέσος όρος, αποδυναμώνοντας τις ιδιαιτερότητες που καταγράφονται από τα μεμονωμένα μοντέλα.

Ωστόσο, η επιτυχία της FL υπογραμμίζει τα πλεονεκτήματά της σε σενάρια όπου οι ανησυχίες για την προστασία της ιδιωτικότητας των δεδομένων ή την κλιμάκωση σε μεγάλο αριθμό ενδέχεται να αποκλείουν πλήρως συγκεντρωτικές προσεγγίσεις. Σχετικά με την απόδοση της FL&DP παρατηρούμε ότι έχει μια αισθητή πτώση σε σχέση με την απλή FL. Η επιλογή ανάγεται και πάλι στο πόση απόδοση είμαστε πρόθυμοι να θυσιάσουμε για την διαφύλαξη των δεδομένων. Στην προκειμένη περίπτωση φαίνεται να έχει μεγάλο αντίκτυπο στην ακρίβεια, ωστόσο σε ένα σενάριο με περισσότερους συμμετέχοντες η απόδοση δύναται να φτάσει αυτήν της απλής FL.

8.2 Αντίκτυπος Ομαδοποίησης

Σε συνέχεια της συγκριτικής ανάλυσης των μεθόδων μάθησης επιλέξαμε να πραγματοποιήσουμε μια αξιολόγηση του αντίκτυπου της ομαδοποίησης των δεδομένων στις προβλέψεις των μοντέλων. Για την υλοποίηση αυτής της αξιολόγησης θα εκτελέσουμε από μια εκπαίδευση CL και FL σε ολόκληρο το σύνολο δεδομένων και στην συνέχεια θα συγκρίνουμε τα αποτελέσματα με τον μέσο όρο όλων των clusters.

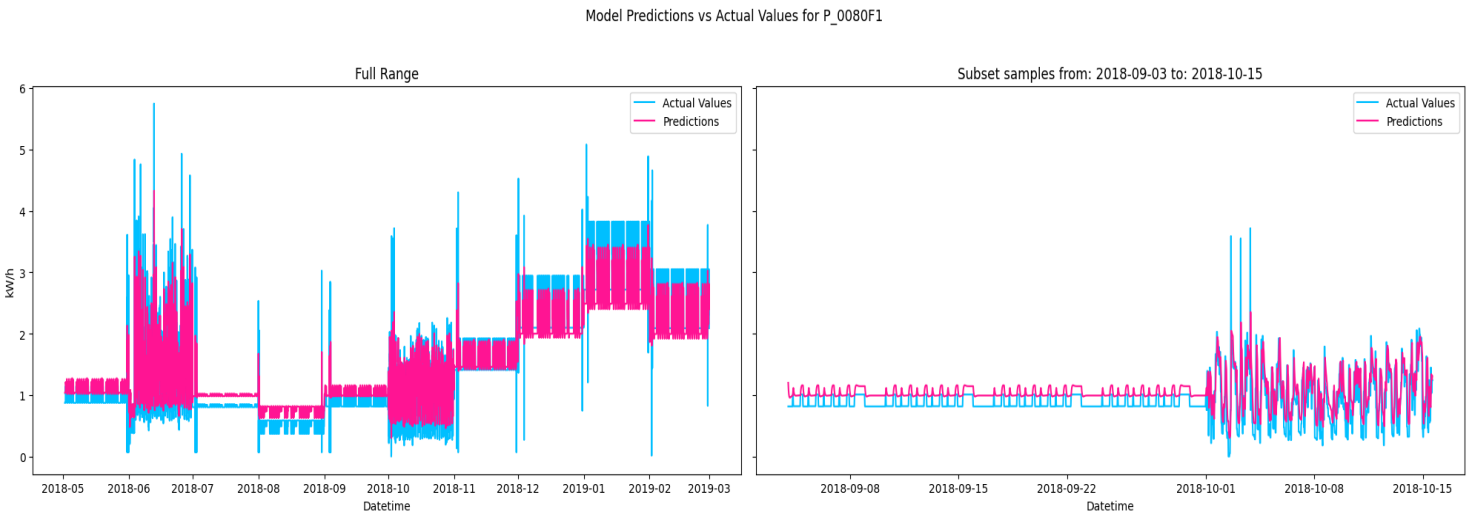
8.2.1 Ομαδοποίηση στην Ομοσπονδιακή Μάθηση

Στον πίνακα 8.6 βλέπουμε το μέσο σφάλμα ομαδοποιημένης και μη Ομοσπονδιακής μάθησης.

Model	MAE	RMSE
Unclustered Federated Learning	0.01641	0.02249
Clustered Federated Learning	0.0155	0.0252

Πίνακας 8.6: *Μέσο σφάλμα ομαδοποιημένης και μη Ομοσπονδιακής Μάθησης*

Σύμφωνα με τα ευρήματά μας, τα μοντέλα ομαδοποιημένης FL επέδειξαν ένα ενδιαφέρον προφίλ επιδόσεων με καλύτερο MAE αλλά χειρότερο RMSE σε σύγκριση με τα μη ομαδοποιημένα μοντέλα. Το φαινόμενο αυτό μπορεί να αποδοθεί στα εγγενή χαρακτηριστικά αυτών των μετρικών σφάλματος και στην ευαισθησία τους στην κατανομή των δεδομένων. Το καλύτερο MAE με τα ομαδοποιημένα μοντέλα υποδηλώνει ότι, κατά μέσο όρο, τα μοντέλα αυτά επιτυγχάνουν πιο κοντινές προβλέψεις στις πραγματικές τιμές, πράγμα που σημαίνει μεγαλύτερη συνέπεια ή σταθερότητα στο χειρισμό της πλειοψηφίας των προβλέψεων. Ωστόσο, το αυξημένο RMSE υποδεικνύει μια ευπάθεια στην αντιμετώπιση ακραίων τιμών, οι οποίες σταθμίζονται περισσότερο σε αυτή τη μετρική. Αυτή η ασυμφωνία υποδηλώνει ότι ενώ η ομαδοποίηση μπορεί να βελτιώσει την απόδοση του μοντέλου σε πιο ομοιογενή υποσύνολα δεδομένων, οδηγώντας σε χαμηλότερο μέσο σφάλμα, μπορεί επίσης να θέσει σε κίνδυνο την ικανότητα του μοντέλου να προβλέπει με ακρίβεια σενάρια ακραίων τιμών.



Διάγραμμα 8.11: *Πρόβλεψη παραγωγής μη-ομαδοποιημένης ομοσπονδιακής μάθησης για τον παραγωγό P_0080F1*

8.2.2 Ομαδοποίηση στην Συγκεντρωτική Μάθηση

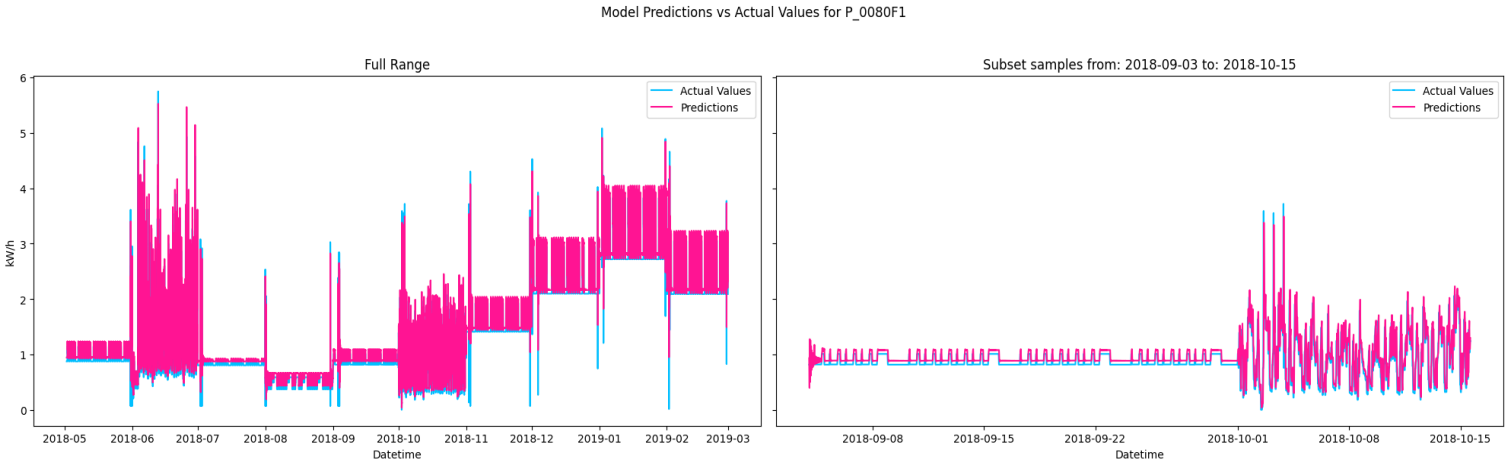
Στον πίνακα 8.7 βλέπουμε το μέσο σφάλμα ομαδοποιημένης και μη Συγκεντρωτικής Μάθησης και παρατηρούμε πως η μη ομαδοποιημένη απόδοση είναι αισθητά καλύτερη από την ομαδοποιημένη.

Model	MAE	RMSE
Unclustered Centralized Learning	0.006402	0.012083
Clustered Centralized Learning	0.0088	0.0181

Πίνακας 8.7: *Μέσο σφάλμα ομαδοποιημένης και μη Συγκεντρωτικής Μάθησης.*

Η συγκεκριμένη απόδοση μπορεί να αποδοθεί κυρίως στην ολοκληρωμένη πρόσβαση στα δεδομένα, η οποία περιλαμβάνει ένα ευρύτερο φάσμα παραλλαγών και προτύπων δεδομένων. Αυτή η άμεση έκθεση σε διαφορετικά δεδομένα επιτρέπει στο μοντέλο χωρίς clusters να αναπτύξει μια πιο γενικευμένη και ισχυρή αναπαράσταση. Από την άλλη πλευρά, ενδέχεται η ομαδοποίηση των λίγων πελατών του συνόλου να οδήγησε σε κατάτμηση των δεδομένων και απώλεια κρίσιμων πληροφοριών. Επιπλέον, το μη ομαδοποιημένο μοντέλο επωφελείται από μια καλύτερη αντιστάθμιση προκατάληψης-διακύμανσης, αξιοποιώντας το πλήρες σύνολο δεδομένων, μειώνοντας έτσι την προκατάληψη και εξασφαλίζοντας ότι το μοντέλο συλλαμβάνει τα πολύπλοκα καθολικά μοτίβα πιο αποτελεσματικά από τα ομαδοποιημένα μοντέλα, τα οποία μπορεί να είναι υπερβολικά εξειδικευμένα και λιγότερο ικανά να γενικεύσουν στο ευρύτερο σύνολο δεδομένων.

Στο διάγραμμα 8.11 βλέπουμε την αισθητή διαφορά στο επίπεδο των προβλέψεων της ομοσπονδιακής μάθησης σε σχέση με την συγκεντρωτική στο διάγραμμα 8.12. Αξίζει βέβαια να σημειωθεί πως η ολοκλήρωση της εκπαίδευσης σε όλο το σύνολο δεδομένων έγινε σε πολύ λιγότερο χρόνο στην περίπτωση της ομοσπονδιακής μάθησης.



Διάγραμμα 8.12: Πρόβλεψη παραγωγής μη-ομαδοποιημένης συγκεντρωτικής μάθησης για τον παραγωγό P_0080F1

9 Συμπεράσματα

Η παρούσα διπλωματική εργασία διερεύνησε τις δυνατότητες της ομοσπονδιακής μάθησης για την πρόβλεψη παραγωγής σε σύγκριση με την κεντρική μάθηση, την τοπική μάθηση και την ομοσπονδιακή μάθηση με διαφορική ιδιωτικότητα. Τα ευρήματα μας βοηθούν να κατανοήσουμε καλύτερα τα πλεονεκτήματα, τους περιορισμούς και τα αντισταθμιστικά οφέλη αυτών των προσεγγίσεων κατανεμημένης μάθησης και μπορούμε να τα χωρίσουμε σε 4 κατηγορίες:

Ιεραρχία επιδόσεων: Τα πειραματικά αποτελέσματα καθόρισαν μια σαφή ιεραρχία επιδόσεων: CL > FL > FLDP > LL. Η κεντρική μάθηση, με πρόσβαση σε ολόκληρο το σύνολο των δεδομένων, επέδειξε ανώτερη ακρίβεια. Η ομοσπονδιακή μάθηση πέτυχε καλή ακρίβεια, ενώ τα μέτρα προστασίας της ιδιωτικότητας της διαφορικής ιδιωτικότητας στο FLDP οδήγησαν σε κάποια μείωση της απόδοσης. Η τοπική μάθηση, με μοντέλα εκπαιδευμένα μόνο σε μεμονωμένα σύνολα δεδομένων, παρουσίασε τη χαμηλότερη απόδοση.

Επίδραση της ομαδοποίησης: Η επίδραση της ομαδοποίησης στις επιδόσεις παρουσίασε ενδιαφέρουσες πτυχές. Ενώ τα μη ομαδοποιημένα δεδομένα είχαν καλύτερες επιδόσεις στο

σενάριο κεντρικής μάθησης, στο πλαίσιο της ομοσπονδιακής μάθησης είχαν καλύτερο RMSE αλλά χειρότερο MAE.. Αυτό υποδηλώνει ότι η ομαδοποίηση μπορεί να προσφέρει πλεονεκτήματα στην ομοσπονδιακή μάθηση, επιτρέποντας καλύτερη μοντελοποίηση της συνολικής κατανομής σφαλμάτων, όμως ταυτόχρονα μειώνει την ανθεκτικότητα του μοντέλου στην επίδραση των ακραίων τιμών.

FL για την πρόβλεψη παραγωγής: Τα αποτελέσματα επιβεβαιώνουν τη βιωσιμότητα της ομοσπονδιακής μάθησης για εργασίες πρόβλεψης παραγωγής. Η FL προσφέρει μια ισορροπία μεταξύ ακρίβειας και διατήρησης της ιδιωτικότητας, καθιστώντας την ιδιαίτερα ελκυστική σε σενάρια όπου η ευαισθησία των δεδομένων ή οι κανονιστικοί περιορισμοί επιβάλλουν την προστασία των ατομικών δεδομένων των χρηστών.

Ο συμβιβασμός ακρίβειας-ιδιωτικότητας: Τα πειραματικά ευρήματα αναδεικνύουν τον συμβιβασμό μεταξύ ακρίβειας και ιδιωτικότητας που είναι εγγενής στη διαφορική ιδιωτικότητα. Ενώ η FL-DP προστατεύει την εμπιστευτικότητα των δεδομένων, επιφέρει κάποια μείωση στην απόδοση του μοντέλου. Η συγκεκριμένη απόφαση σχετικά με το κατά πόσον αυτός ο συμβιβασμός είναι αποδεκτός εξαρτάται από την ευαισθησία των δεδομένων παραγωγής και τις απαιτήσεις απορρήτου σε μια συγκεκριμένη περίπτωση χρήσης.

9.1 Μελλοντικές κατευθύνσεις

Με την παρούσα διπλωματική θέσαμε το υπόβαθρο για την κατανόηση των μεθόδων πρόβλεψης παραγωγής σε κατανεμημένες συνθήκες με επίκεντρο την διαφύλαξη της ιδιωτικότητας. Ωστόσο το περιορισμένο σύνολο δεδομένων αφήνει ανοιχτό το ενδεχόμενο περεταίρω διερεύνησης. Μια μελέτη σε ένα σύνολο μεγάλης κλίμακας θα παρουσίαζε ιδιαίτερο ενδιαφέρον. Εξίσου ενδιαφέρουσα θα ήταν μια μελέτη για την υλοποίηση ενός δικτύου ομοσπονδιακής μάθησης σε πραγματικές συνθήκες για παράδειγμα με την χρήση έξυπνων μετρητών.

Σε ένα πιο θεωρητικό επίπεδο, μελλοντικές εργασίες θα μπορούσαν επίσης να διερευνήσουν εξειδικευμένες τεχνικές εφαρμογής διαφορικής ιδιωτικότητας ώστε να μειωθεί ακόμα περισσότερο ο συμβιβασμός ακρίβειας-ιδιωτικότητας. Ακόμη, ιδιαίτερο ενδιαφέρον παρουσιάζει

και ο συνδυασμός ομοσπονδιακής μάθησης με άλλες τεχνικές διατήρησης της ιδιωτικότητας όπως οι secure aggregation, homomorphic encryption

Συντομογραφίες

AI	artificial intelligence; τεχνητή νοημοσύνη
ANN	artificial neural networks; τεχνητά νευρωνικά δίκτυα
ARIMA	autoregressive integrated moving average; μοντέλα αυτοπαλίνδρομου ολοκληρωμένου κινητού μέσου
CL	centralized learning; συγκεντρωτική μάθηση
DNN	deep neural networks; βαθιά νευρωνικά δίκτύα
DP	διαφορική ιδιωτικότητα
DP-SGD	dp-sgd (differentially private stochastic gradient descent)
EDA	exploratory data analysis; διερευνητική ανάλυση δεδομένων
FedAvg	federated average - ομοσπονδιακό μέσος
FL	federated learning; ομοσπονδιακή μάθηση
FL-DP	differential privacy-enhanced federated learning; ομοσπονδιακή μάθηση ενισχυμένη με διαφορική ιδιωτικότητα
FNN	feedforward neural network; νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης
GBM	gradient boosting machines; μηχανές ενίσχυσης κλίσης
GDP	global gifferential privacy; καθολική διαφορική ιδιωτικότητα
GDPR	general data protection regulation γενικός κανονισμός για την προστασία δεδομένων

KDE	kernel density estimation; εκτίμησης πιθανότητας πυρήνα
LDP	local differential privacy; τοπική διαφορική ιδιωτικότητα
LL	local learning; τοπική μάθηση
LSTM	long short-term memory; δίκτυα μακράς βραχυπρόθεσμης μνήμης
MAE	mean absolute error; μέσο απόλυτο σφάλμα
ML	machine learning; μηχανική μάθηση
MSE	mean squared error; μέσο τετραγωνικό σφάλμα
NN	neural networks; νευρωνικά δίκτυα
PCA	principal component analysis; ανάλυση κύριων συνιστωσών
RMSE	root mean squared error; μέσο τετραγωνικό σφάλμα
RNN	recurrent neural network; επαναλαμβανόμενο νευρωνικό δίκτυο
SNN	shallow neural networks; αβαθή νευρωνικά δίκτύα
TF	tensorflow
TFF	tensorflow federated
$\epsilon-DP$	ε-διαφορική ιδιωτικότητα ($\epsilon-DP$)

Βιβλιογραφία

[1] H. Ritchie, M. Roser, and P. Rosado, 'Renewable Energy', *Our World Data*, 2020.

[2] Y. Parag and B. K. Sovacool, 'Electricity market design for the prosumer era', *Nat. Energy*, vol. 1, no. 4, pp. 1–6, Mar. 2016, doi: 10.1038/nenergy.2016.32.

[3] V. Chifu, T. Cioara, C. Anitiei, C. Pop, and I. Anghel, 'FedWOA: A Federated Learning Model that uses the Whale Optimization Algorithm for Renewable Energy Prediction'. arXiv, Sep. 19, 2023. doi: 10.48550/arXiv.2309.10337.

[4] A. Taïk and S. Cherkaoui, 'Electrical Load Forecasting Using Edge Computing and Federated Learning', in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, Jun. 2020, pp. 1–6. doi: 10.1109/ICC40277.2020.9148937.

[5] *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. 2016. Accessed: Feb. 13, 2024. [Online]. Available: http://data.europa.eu/eli/reg/2016/679/2016-05-04/eng

[6] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, 'Communication-Efficient Learning of Deep Networks from Decentralized Data'. arXiv, Oct. 21, 2016. doi: 10.48550/arXiv.1602.05629.

[7] M. Abadi *et al.*, 'Deep Learning with Differential Privacy', in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, Oct. 2016, pp. 308–318. doi: 10.1145/2976749.2978318.

[8] R.J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice (3rd ed)*. OTexts: Melbourne, Australia. Accessed: Feb. 04, 2024. [Online]. Available: https://otexts.com/fpp3/

[9] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung, *Time Series Analysis: Forecasting and Control, 5th Edition | Wiley*, 5th ed. Wiley, 2015. Accessed: Feb. 05, 2024. [Online]. Available: https://www.wiley.com/en-ca/Time+Series+Analysis%3A+Forecasting+and+Control%2C+5th+Edition-p-9781118675021

[10] R. Adhikari and R. K. Agrawal, 'An Introductory Study on Time Series Modeling and Forecasting'. arXiv, Feb. 26, 2013. doi: 10.48550/arXiv.1302.6613.

[11] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications: With R Examples*. in Springer Texts in Statistics. Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-52452-8.

[12] 'Natural gas use features two seasonal peaks per year - U.S. Energy Information Administration (EIA)'. Accessed: Feb. 06, 2024. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=22892

[13] H. S. Hippert, C. E. Pedreira, and R. C. Souza, 'Neural networks for short-term load forecasting: a review and evaluation', *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44–55, Feb. 2001, doi: 10.1109/59.910780.

[14] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, 'Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network', *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan. 2019, doi: 10.1109/TSG.2017.2753802.

[15] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*. in Springer Texts in Statistics. Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-29854-2.

[16] A. Bogomolov, B. Lepri, R. Larcher, F. Antonelli, F. Pianesi, and A. Pentland, 'Energy consumption prediction using people dynamics derived from cellular network data', *EPJ Data Sci.*, vol. 5, Mar. 2016, doi: 10.1140/epjds/s13688-016-0075-3.

[17] X. Tang, H. Yao, Y. Sun, C. Aggarwal, P. Mitra, and S. Wang, 'Joint Modeling of Local and Global Temporal Dynamics for Multivariate Time Series Forecasting with Missing Values', *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 04, Art. no. 04, Apr. 2020, doi: 10.1609/aaai.v34i04.6056.

[18] M. M. Ahsan, M. A. P. Mahmud, P. K. Saha, K. D. Gupta, and Z. Siddique, 'Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance', *Technologies*, vol. 9, no. 3, Art. no. 3, Sep. 2021, doi: 10.3390/technologies9030052.

[19] L. B. V. de Amorim, G. D. C. Cavalcanti, and R. M. O. Cruz, 'The choice of scaling technique matters for classification performance', *Appl. Soft Comput.*, vol. 133, p. 109924, Jan. 2023, doi: 10.1016/j.asoc.2022.109924.

[20] A. Y. Ng, 'Feature selection, L1 vs. L2 regularization, and rotational invariance', in *Proceedings of the twenty-first international conference on Machine learning*, in ICML '04. New York, NY, USA: Association for Computing Machinery, Jul. 2004, p. 78. doi: 10.1145/1015330.1015435.

[21] J. Zhang, 'Gradient Descent based Optimization Algorithms for Deep Learning Models Training'. arXiv, Mar. 11, 2019. doi: 10.48550/arXiv.1903.03614.

[22] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd edition. Pearson, 2009.

[23] M. I. Jordan and T. M. Mitchell, 'Machine learning: Trends, perspectives, and prospects', *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015, doi: 10.1126/science.aaa8415.

[24] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, Illustrated edition. Cambridge, MA: The MIT Press, 2012.

[25] T. M. Mitchell, *Machine Learning*, 1st edition. New York: McGraw-Hill Education, 1997.

[26] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 5th edition. Hoboken, NJ: Wiley, 2012.

[27] L. Breiman, 'Random Forests', *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[28] J. H. Friedman, 'Greedy Function Approximation: A Gradient Boosting Machine', *Ann. Stat.*, vol. 29, no. 5, pp. 1189–1232, 2001.

[29] A. Y. Barrera-Animas, L. O. Oyedele, M. Bilal, T. D. Akinosho, J. M. D. Delgado, and L. A. Akanbi, 'Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting', *Mach. Learn. Appl.*, vol. 7, p. 100204, Mar. 2022, doi: 10.1016/j.mlwa.2021.100204.

[30] C. C. Aggarwal, *Neural Networks and Deep Learning: A Textbook*, 1st ed. 2018 edition. Springer, 2018.

[31] Z. Meng, Y. Hu, and C. Ancey, 'Using a Data Driven Approach to Predict Waves Generated by Gravity Driven Mass Flows', *Water*, vol. 12, Feb. 2020, doi: 10.3390/w12020600.

[32] S. Haykin, *Neural Networks and Learning Machines*, 3rd edition. New York Munich: Pearson, 2008.

[33] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[34] Y. LeCun, Y. Bengio, and G. Hinton, 'Deep learning', *Nature*, vol. 521, no. 7553, Art. no. 7553, May 2015, doi: 10.1038/nature14539.

[35] L. Ciampiconi, A. Elwood, M. Leonardi, A. Mohamed, and A. Rozza, 'A survey and taxonomy of loss functions in machine learning'. arXiv, Jan. 13, 2023. doi: 10.48550/arXiv.2301.05579.

[36] D. P. Kingma and J. Ba, 'Adam: A Method for Stochastic Optimization'. arXiv, Jan. 29, 2017. doi: 10.48550/arXiv.1412.6980.

[37] R. Pramoditha, 'Overview of a Neural Network's Learning Process', Data Science 365. Accessed: Feb. 12, 2024. [Online]. Available: https://medium.com/data-science-365/overview-of-a-neural-networks-learning-process-61690a502fa

[38] R. Pascanu, T. Mikolov, and Y. Bengio, 'On the difficulty of training Recurrent Neural Networks'. arXiv, Feb. 15, 2013. doi: 10.48550/arXiv.1211.5063.

[39] 'A Field Guide to Dynamical Recurrent Networks | IEEE eBooks | IEEE Xplore'. Accessed: Feb. 12, 2024. [Online]. Available: https://ieeexplore.ieee.org/book/5263132

[40] S. Hochreiter and J. Schmidhuber, 'Long Short-Term Memory', *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[41] M. N. Fekri, H. Patel, K. Grolinger, and V. Sharma, 'Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network', *Appl. Energy*, vol. 282, p. 116177, Jan. 2021, doi: 10.1016/j.apenergy.2020.116177.

[42] L. Sehovac and K. Grolinger, 'Deep Learning for Load Forecasting: Sequence to Sequence Recurrent Neural Networks With Attention', *IEEE Access*, vol. 8, pp. 36411–36426, 2020, doi: 10.1109/ACCESS.2020.2975738.

[43] Q. Yang, Y. Liu, T. Chen, and Y. Tong, 'Federated Machine Learning: Concept and Applications'. arXiv, Feb. 13, 2019. doi: 10.48550/arXiv.1902.04885.

[44] J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, 'Federated Optimization: Distributed Machine Learning for On-Device Intelligence'. arXiv, Oct. 08, 2016. doi: 10.48550/arXiv.1610.02527.

[45] K. Bonawitz *et al.*, 'Towards Federated Learning at Scale: System Design'. arXiv, Mar. 22, 2019. doi: 10.48550/arXiv.1902.01046.

[46] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, 'Federated Learning: Challenges, Methods, and Future Directions', *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 50–60, May 2020, doi: 10.1109/MSP.2020.2975749.

[47] V. Smith, C.-K. Chiang, M. Sanjabi, and A. S. Talwalkar, 'Federated Multi-Task Learning', in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017. Accessed: Feb. 13, 2024. [Online]. Available: https://papers.nips.cc/paper_files/paper/2017/hash/6211080fa89981f66b1a0c9d55c61d0f-Abstract.html

[48] C. Dwork, 'Differential Privacy', in *Automata, Languages and Programming*, M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2006, pp. 1–12. doi: 10.1007/11787006_1.

[49] L. Rocher, J. M. Hendrickx, and Y.-A. de Montjoye, 'Estimating the success of re-identifications in incomplete datasets using generative models', *Nat. Commun.*, vol. 10, no. 1, Art. no. 1, Jul. 2019, doi: 10.1038/s41467-019-10933-3.

[50] 'Differential Privacy and the 2020 US Census · Winter 2022'. Accessed: Feb. 27, 2024. [Online]. Available: https://mit-serc.pubpub.org/pub/differential-privacy-2020-us-census/release/1

[51] C. Dwork and A. Roth, 'The Algorithmic Foundations of Differential Privacy', *Found. Trends® Theor. Comput. Sci.*, vol. 9, no. 3–4, pp. 211–407, 2013, doi: 10.1561/0400000042.

[52] S. P. Kasiviswanathan and A. Smith, 'On the "Semantics" of Differential Privacy: A Bayesian Formulation', *J. Priv. Confidentiality*, vol. 6, no. 1, Jun. 2014, doi: 10.29012/jpc.v6i1.634.

[53] C. Antal *et al.*, 'Blockchain based decentralized local energy flexibility market', *Energy Rep.*, vol. 7, pp. 5269–5288, Nov. 2021, doi: 10.1016/j.egyr.2021.08.118.

[54] A. M. Alonso, F. J. Nogales, and C. Ruiz, 'A Single Scalable LSTM Model for Short-Term Forecasting of Disaggregated Electricity Loads'. arXiv, Mar. 06, 2020. doi: 10.48550/arXiv.1910.06640.

[55] M. Savi and F. Olivadese, 'Short-Term Energy Consumption Forecasting at the Edge: A Federated Learning Approach', *IEEE Access*, vol. 9, pp. 95949–95969, 2021, doi: 10.1109/ACCESS.2021.3094089.

[56] T. Proietti, 'Missing data in time series: A note on the equivalence of the dummy variable and the skipping approaches', *Stat. Probab. Lett.*, vol. 78, no. 3, pp. 257–264, Feb. 2008, doi: 10.1016/j.spl.2007.05.031.

[57] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*, 2nd edition. Hoboken, New Jersey: Wiley, 2015.

[58] Josef Perktold *et al.*, 'statsmodels/statsmodels: Release 0.14.1'. Zenodo, Dec. 14, 2023. doi: 10.5281/ZENODO.593847.

[59] Martín Abadi *et al.*, 'TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems'. 2015. [Online]. Available: https://www.tensorflow.org/

[60] The Tensorflow Federated Authors and Google, 'TensorFlow Federated'. Google, Dec. 12, 2018. [Online]. Available: https://github.com/tensorflow/federated

[61] F. Chollet and others, 'Keras'. 2015. [Online]. Available: https://keras.io

[62] H. Wang, Y. Zhao, S. He, Y. Xiao, J. Tang, and Z. Cai, 'Federated learning-based privacy-preserving electricity load forecasting scheme in edge computing scenario', *Int. J. Commun. Syst.*, vol. 37, no. 5, p. e5670, 2024, doi: 10.1002/dac.5670.

[63] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, 'Optuna: A Next-generation Hyperparameter Optimization Framework'. arXiv, Jul. 25, 2019. doi: 10.48550/arXiv.1907.10902.

[64] N. Gholizadeh and P. Musilek, 'Federated learning with hyperparameter-based clustering for electrical load forecasting', *Internet Things*, vol. 17, p. 100470, Mar. 2022, doi: 10.1016/j.iot.2021.100470.

[65] 'Acorn | Geodemographic Segmentation | Acorn Data | CACI', Acorn. Accessed: Feb. 20, 2024. [Online]. Available: https://acorn.caci.co.uk/

[66] S. Lloyd, 'Least squares quantization in PCM', *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982, doi: 10.1109/TIT.1982.1056489.

[67] P. J. Rousseeuw, 'Silhouettes: A graphical aid to the interpretation and validation of cluster analysis', *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, Nov. 1987, doi: 10.1016/0377-0427(87)90125-7.

[68] G. Drainakis, K. V. Katsaros, P. Pantazopoulos, V. Sourlas, and A. Amditis, 'Federated vs. Centralized Machine Learning under Privacy-elastic Users: A Comparative Analysis', in *2020 IEEE 19th International Symposium on Network Computing and Applications (NCA)*, Nov. 2020, pp. 1–8. doi: 10.1109/NCA51143.2020.9306745.

[69] C. Culnane, B. I. P. Rubinstein, and V. Teague, 'Health Data in an Open World'. arXiv, Dec. 15, 2017. doi: 10.48550/arXiv.1712.05627.

[70] J. C. Liu, J. Goetz, S. Sen, and A. Tewari, 'Learning From Others Without Sacrificing Privacy: Simulation Comparing Centralized and Federated Machine Learning on Mobile Health Data', *JMIR MHealth UHealth*, vol. 9, no. 3, p. e23728, Mar. 2021, doi: 10.2196/23728.

[71] V. Perifanis, N. Pavlidis, R.-A. Koutsiamanis, and P. S. Efraimidis, 'Federated Learning for 5G Base Station Traffic Forecasting', *Comput. Netw.*, vol. 235, p. 109950, Nov. 2023, doi: 10.1016/j.comnet.2023.109950.

[72] F. Pedregosa *et al.*, 'Scikit-learn: Machine Learning in Python', *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[73] H. B. McMahan, D. Ramage, K. Talwar, and L. Zhang, 'Learning Differentially Private Recurrent Language Models'. arXiv, Feb. 23, 2018. doi: 10.48550/arXiv.1710.06963.

[74] G. Andrew, O. Thakkar, H. B. McMahan, and S. Ramaswamy, 'Differentially Private Learning with Adaptive Clipping'. arXiv, May 09, 2022. doi: 10.48550/arXiv.1905.03871.

[75] Y.-X. Wang, B. Balle, and S. Kasiviswanathan, 'Subsampled R\'enyi Differential Privacy and Analytical Moments Accountant'. arXiv, Dec. 04, 2018. doi: 10.48550/arXiv.1808.00087.

[76] T. Chai and R. R. Draxler, 'Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature', *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014, doi: 10.5194/gmd-7-1247-2014.