



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

Πρόβλεψη Τιμών Ηλεκτρικής Ενέργειας με Χρήση της Αρχιτεκτονικής Transformer

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χουλάκης Ζαχαρίας

Επιβλέπων : Εμμανουήλ Βαρβαρίγος, Καθηγητής Ε.Μ.Π
Τομέας Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής

Αθήνα, Ιούνιος 2025



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Πρόβλεψη Τιμών Ηλεκτρικής Ενέργειας με Χρήση της Αρχιτεκτονικής Transformer

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χουλάκης Ζαχαρίας

Επιβλέπων : Εμμανουήλ Βαρβαρίγος, Καθηγητής Ε.Μ.Π.
Τομέας Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 20^η Ιουνίου 2025.

.....
κ. Βαρβαρίγος Εμμανουήλ
Καθηγητής

.....
κ. Βαρβαρίγου Θεοδώρα
Καθηγήτρια

.....
κ. Συμεών Παπαβασιλείου
Καθηγητής

Αθήνα, Ιούνιος 2025

.....
Χουλάκης Ζαχαρίας

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

Copyright © Χουλάκης Ζαχαρίας, 2025.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Η ενέργεια αποτελεί θεμέλιο για τη λειτουργία και την εξέλιξη της σύγχρονης κοινωνίας, καθώς είναι αναγκαία για την υποστήριξη του συνόλου σχεδόν των ανθρώπινων δραστηριοτήτων, από τη βιομηχανική παραγωγή και τις μεταφορές, έως την κατοικία και τις ψηφιακές υποδομές. Παράλληλα, αποτελεί βασικό μοχλό οικονομικής ανάπτυξης, με την προσβασιμότητα και την αποδοτική χρήση ενεργειακών πόρων να καθορίζουν σε μεγάλο βαθμό τη βιωσιμότητα και την ανταγωνιστικότητα των οικονομιών. Ωστόσο, η συνεχής αύξηση της ζήτησης, σε συνδυασμό με την πεπερασμένη φύση των παραδοσιακών ενεργειακών πόρων και τις αυξανόμενες περιβαλλοντικές πιέσεις, καθιστούν επιτακτική την ανάγκη για μετάβαση προς ένα πιο αποδοτικό, ευφυές και βιώσιμο ενεργειακό σύστημα.

Σε αυτό το πλαίσιο, οι τεχνολογίες της Τεχνητής Νοημοσύνης και πιο συγκεκριμένα η Μηχανική Μάθηση (Machine Learning), αναδεικνύονται ως κρίσιμα εργαλεία για την αντιμετώπιση της πολυπλοκότητας που χαρακτηρίζει την πρόβλεψη και διαχείριση της ενεργειακής αγοράς. Τα εξελιγμένα μοντέλα πρόβλεψης, όπως οι αρχιτεκτονικές τύπου Transformer, προσφέρουν τη δυνατότητα επεξεργασίας μεγάλων όγκων ιστορικών και χρονικά εξαρτώμενων δεδομένων, με στόχο την ακριβή εκτίμηση μελλοντικών τιμών και συνθηκών της αγοράς. Τα μοντέλα αυτά, βασισμένα σε μηχανισμούς αυτο-προσοχής (self-attention), έχουν ήδη αποδείξει την υπεροχή τους σε καθήκοντα όπως η φυσική γλώσσα και η ανάλυση ακολουθιακών δεδομένων, γεγονός που τα καθιστά ιδανικούς υποψήφιους και για προβλήματα πρόβλεψης τιμών ηλεκτρικής ενέργειας.

Η πρόβλεψη των τιμών ηλεκτρικής ενέργειας αποτελεί ένα ιδιαίτερα απαιτητικό πρόβλημα, καθώς επηρεάζεται από ένα ευρύ φάσμα μεταβλητών, όπως οι καιρικές συνθήκες (π.χ. θερμοκρασία, ηλιοφάνεια, άνεμος), οι διακυμάνσεις της ζήτησης, η διαθεσιμότητα και το κόστος των πηγών παραγωγής, καθώς και οι κανονιστικές και οικονομικές συνθήκες της αγοράς. Η δυνατότητα πρόβλεψης αυτών των τιμών με υψηλό βαθμό ακρίβειας έχει σημαντική πρακτική αξία, όχι μόνο για τους διαχειριστές δικτύου και τους παρόχους ενέργειας, αλλά και για τους επενδυτές, καθώς τους επιτρέπει να αξιολογούν ευκαιρίες, να σχεδιάζουν στρατηγικές τοποθετήσεις και να ελαχιστοποιούν τον επενδυτικό κίνδυνο.

Στο πλαίσιο της παρούσας διπλωματικής εργασίας, διερευνάται η αξιοποίηση ενός μοντέλου τύπου Transformer για την πρόβλεψη των τιμών ηλεκτρικής ενέργειας, με σκοπό την εξαγωγή στρατηγικών επένδυσης βασισμένων στα προβλεπόμενα δεδομένα. Η μεθοδολογική προσέγγιση περιλαμβάνει την ανάλυση ιστορικών δεδομένων, τη σχεδίαση και εκπαίδευση του μοντέλου, καθώς και την αξιολόγηση της αποτελεσματικότητάς του ως προς την ακρίβεια πρόβλεψης και τη χρηστική του αξία στο επενδυτικό πλαίσιο. Μέσα από την ανάλυση των αποτελεσμάτων, αναδεικνύεται η σημασία της ενσωμάτωσης της τεχνητής νοημοσύνης στη διαχείριση των ενεργειακών πόρων και η δυνατότητα αξιοποίησης των προβλέψεων για τη λήψη τεκμηριωμένων και αποδοτικών επενδυτικών αποφάσεων, συμβάλλοντας ταυτόχρονα στην προώθηση της ενεργειακής αποδοτικότητας και της βιωσιμότητας.

Λέξεις-κλειδιά: Ωριαία πρόβλεψη τιμών ηλεκτρικής ενέργειας, Transformer, Μηχανική μάθηση, Επενδυτική στρατηγική, Ανάλυση χρονοσειρών, Διαχείριση ρίσκου.

Abstract

Energy constitutes a foundational pillar for the functioning and advancement of modern society, as it is essential for supporting nearly all human activities from industrial production and transportation to housing and digital infrastructure. Simultaneously, it serves as a key driver of economic development, with accessibility to and efficient utilization of energy resources playing a critical role in determining the sustainability and competitiveness of economies. However, the continuous rise in demand, coupled with the finite nature of traditional energy sources and escalating environmental pressures, underscores the urgent need for a transition toward a more efficient, intelligent, and sustainable energy system.

Within this context, Artificial Intelligence technologies and more specifically Machine Learning are emerging as pivotal tools for addressing the inherent complexity of forecasting and managing energy markets. Advanced predictive models, such as Transformer-based architectures, offer the capability to process vast amounts of historical and temporally dependent data, aiming to accurately estimate future market prices and conditions. These models, grounded in self-attention mechanisms, have already demonstrated their superiority in tasks such as natural language processing and sequential data analysis, rendering them highly suitable candidates for the problem of electricity price forecasting.

Electricity price forecasting is a particularly challenging task, influenced by a wide array of variables, including weather conditions (e.g., temperature, solar radiation, wind), demand fluctuations, the availability and cost of production sources, as well as regulatory and economic market dynamics. The ability to predict these prices with a high degree of accuracy holds substantial practical value not only for grid operators and energy providers but also for investors, as it enables opportunity assessment, strategic positioning, and risk minimization.

This thesis explores the application of a Transformer-based model for forecasting electricity prices, with the objective of deriving investment strategies informed by the predicted data. The methodological approach encompasses the analysis of historical datasets, the design and training of the model, and the evaluation of its performance in terms of predictive accuracy and practical relevance within an investment framework. Through the analysis of results, the study highlights the importance of integrating Artificial Intelligence into energy resource management and demonstrates the potential of data-driven forecasts to support informed and effective investment decisions, thereby contributing to enhanced energy efficiency and sustainability.

Keywords: Hourly electricity price forecasting, Transformer, Machine learning, Investment strategy, Time series analysis, Risk management.

Ευχαριστίες

Με την ολοκλήρωση της Διπλωματικής μου εργασίας, θα ήθελα να ευχαριστήσω θερμά όσους στάθηκαν δίπλα μου κατά τη διάρκεια αυτής της διαδρομής. Αρχικά, τον καθηγητή κ. Εμμανουήλ (Μάνο) Βαρβαρίγο για την εμπιστοσύνη και την ευκαιρία να ασχοληθώ με ένα ενδιαφέρον και επίκαιρο θέμα, καθώς και για τις χρήσιμες παρατηρήσεις και κατευθύνσεις του. Ιδιαίτερες ευχαριστίες οφείλω επίσης στον Υποψήφιο Δρ. Κωνσταντίνο Κεχαγιά για τη συνεχή υποστήριξη, την τεχνική καθοδήγηση και την προθυμία του να απαντά σε κάθε μου απορία με σαφήνεια και υπομονή. Τέλος, ευχαριστώ από καρδιάς την οικογένειά μου και τους κοντινούς μου ανθρώπους για τη στήριξή τους, την κατανόηση και τη σταθερή παρουσία τους σε όλη τη διάρκεια των σπουδών μου.

Περιεχόμενα

Εισαγωγή.....	1
1.1 Διατύπωση του προβλήματος.....	1
1.2 Προσέγγιση της Μηχανικής Μάθησης.....	1
1.3 Ανάπτυξη Επενδυτικής Στρατηγικής.....	2
Θεωρητικό Υπόβαθρο.....	3
2.1 Νευρωνικά Δίκτυα.....	3
2.2 Δομή ενός Νευρώνα.....	3
2.3 Συναρτήσεις ενεργοποίησης.....	4
2.3.1 Η Σιγμοειδής συνάρτηση.....	4
2.3.2 Η συνάρτηση Υπερβολικής Εφαπτομένης.....	5
2.3.3 Η Συνάρτηση ReLU.....	6
2.4 Τύποι Νευρωνικών Δικτύων.....	7
2.4.1 Πλήρως Συνδεδεμένα Νευρωνικά Δίκτυα (FCNNs).....	7
2.4.2 Επαναλαμβανόμενα Νευρωνικά Δίκτυα (RNNs).....	8
2.5 Αρχιτεκτονική LSTM (Long Short-Term Memory).....	10
2.5.1 Το πρόβλημα εξασθένησης των βαθμίδων (Vanishing Gradient Problem).....	10
2.5.2 Δομή των LSTM.....	10
2.5.3 Πύλη Λήθης (Forget Gate).....	11
2.5.4 Πύλη Εισόδου (Input Gate).....	12
2.5.5 Πύλη Εξόδου (Output Gate).....	12
2.5.6 Λειτουργία του Κελιού LSTM.....	12
2.5.7 Πλεονεκτήματα και Εφαρμογές των LSTM.....	13
2.6 Αρχιτεκτονική των Συνελκτικών Νευρωνικών Δικτύων (CNN).....	14
2.6.1 Συνελκτικά Επίπεδα (Convolutional Layers).....	14
2.6.2 Επίπεδα Pooling (Pooling Layers).....	15
2.6.3 Πλήρως Συνδεδεμένα Επίπεδα (Fully Connected Layers).....	16
2.6.4 Εφαρμογή των CNNs σε Χρονοσειρές (Time Series).....	17
2.7 Vanilla transformer.....	17
2.7.1 Αρχιτεκτονική του Vanilla Transformer.....	18
2.7.1.1 Κωδικοποιητής (Encoder).....	18
2.7.1.2 Αποκωδικοποιητής (Decoder).....	20
2.7.1.3 Κωδικοποίηση Θέσης (Position Encoding).....	20
2.7.2 Παράδειγμα λειτουργίας Transformer.....	21
Πειραματικές Μέθοδοι και Υλοποίηση.....	23
3.1 Δεδομένα.....	23
3.2 Πειραματική Διαδικασία και Διαχωρισμός Δεδομένων.....	25
3.3 Αξιολόγηση Αποτελεσμάτων.....	26
3.3.1 Mean Absolute Error (MAE).....	26
3.3.2 Mean Absolute Percentage Error (MAPE).....	27

3.3.3 Root Mean Squared Error (RMSE).....	28
3.3.4 Σύγκριση των Μετρικών Αξιολόγησης.....	28
3.3.5 Αποτελέσματα Πειραμάτων.....	29
3.4 Σύγκριση αποτελεσμάτων με υπάρχουσα βιβλιογραφία.....	37
Επενδυτικές Στρατηγικές.....	39
4.1 Προσαρμοσμένη Επενδυτική Στρατηγική για Εμπορία Ηλεκτρικής Ενέργειας.....	39
4.2 Sharpe Ratio.....	39
4.3 Σενάριο 1.....	40
4.4 Σενάριο 2.....	45
4.5 Σενάριο 3.....	49
4.6 Ανανεώσιμες Πηγές Ενέργειας.....	56
4.7 Σενάριο 4.....	59
4.8 Σενάριο 5.....	60
4.9 Αποτελέσματα.....	63
Συμπεράσματα.....	65
Βιβλιογραφία.....	67

Κατάλογος Σχημάτων

1. Σχήμα 2.2-1 Δομή νευρώνα.
2. Σχήμα 2.3.1-1 Γραφική παράσταση Σιγμοειδούς συνάρτησης.
3. Σχήμα 2.3.2-1 Γραφική παράσταση συνάρτησης Υπερβολικής Εφαπτομένης.
4. Σχήμα 2.3.3-1 Γραφική παράσταση συνάρτησης ReLU.
5. Σχήμα 2.4.1-1 Δομή FCNN.
6. Σχήμα 2.4.2-1 Δομή RNN.
7. Σχήμα 2.5.2-1 Αρχιτεκτονική LSTM.
8. Σχήμα 2.6-1 Αρχιτεκτονική CNN.
9. Σχήμα 2.6.1-1 Διαδικασία Συνέλιξης.
10. Σχήμα 2.6.2-1 Pooling Layer.
11. Σχήμα 2.6.4-1 Η συνάρτηση Softmax.
12. Σχήμα 2.7.1-1 Αρχιτεκτονική Vanilla Transformer.
13. Σχήμα 3.3.4-1 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.
14. Σχήμα 3.3.4-2 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.
15. Σχήμα 3.3.4-3 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.
16. Σχήμα 3.3.4-4 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.
17. Σχήμα 3.3.4-5 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.
18. Σχήμα 3.3.4-6 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.
19. Σχήμα 3.3.4-7 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.
20. Σχήμα 4.3-1 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 1 στην Ουγγαρία.
21. Σχήμα 4.3-2 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 1 στην Ελλάδα.
22. Σχήμα 4.3-3 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 1 στην Αυστρία
23. Σχήμα 4.3-4 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 1 στη Γαλλία.
24. Σχήμα 4.4-1 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 2 στην Ουγγαρία.
25. Σχήμα 4.4-2 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 2 στην Ελλάδα.
26. Σχήμα 4.4-3 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 2 στην Αυστρία.
27. Σχήμα 4.4-4 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 2 στη Γαλλία.
28. Σχήμα 4.5-1 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 3 στην Ουγγαρία.
29. Σχήμα 4.5-2 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 3 στην Ελλάδα.
30. Σχήμα 4.5-3 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για

το Σενάριο 3 στην Αυστρία.

31. Σχήμα 4.5-4 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 3 στη Γαλλία.
32. Σχήμα 4.5-5 Σύγκριση Αξιολόγησης της Επένδυσης των τριών Σεναρίων για την Ουγγαρία.
33. Σχήμα 4.5-6 Σύγκριση Αξιολόγησης της Επένδυσης των τριών Σεναρίων για την Ελλάδα.
34. Σχήμα 4.5-7 Σύγκριση Αξιολόγησης της Επένδυσης των τριών Σεναρίων για την Αυστρία.
35. Σχήμα 4.5-8 Σύγκριση Αξιολόγησης της Επένδυσης των τριών Σεναρίων για τη Γαλλία.
36. Σχήμα 4.8-1 Μέση Μηνιαία Ηλιακή Ακτινοβολία στην Ελλάδα.
37. Σχήμα 4.8-2 Ώρες ηλιοφάνειας ημερησίως ανα μήνα στην Ελλάδα.
38. Σχήμα 4.8-3 Παράδειγμα ωριαίας κατανομής ηλιακής ακτινοβολίας.
39. Σχήμα 4.9-1 Σύγκριση Επένδυσης με και χωρίς μπαταρία.
40. Σχήμα 4.9-2 Σύγκριση των τεσσάρων σεναρίων που χρησιμοποιούν μπαταρία στην Ελλάδα.

Κατάλογος Πινάκων

1. Πίνακας 1 Δομή Δεδομένων Εργασίας.
2. Πίνακας 2 Δείκτες Ακρίβειας Προβλέψεων.
3. Πίνακας 3 Αξιολόγηση Επένδυσης Σεναρίου 1 στην Ουγγαρία.
4. Πίνακας 4 Αξιολόγηση Επένδυσης Σεναρίου 1 στην Ελλάδα.
5. Πίνακας 5 Αξιολόγηση Επένδυσης Σεναρίου 1 στην Αυστρία.
6. Πίνακας 6 Αξιολόγηση Επένδυσης Σεναρίου 1 στη Γαλλία.
7. Πίνακας 7 Αξιολόγηση Επένδυσης Σεναρίου 2 στην Ουγγαρία.
8. Πίνακας 8 Αξιολόγηση Επένδυσης Σεναρίου 2 στην Ελλάδα.
9. Πίνακας 9 Αξιολόγηση Επένδυσης Σεναρίου 2 στην Αυστρία.
10. Πίνακας 10 Αξιολόγηση Επένδυσης Σεναρίου 2 στη Γαλλία.
11. Πίνακας 11 Αξιολόγηση Επένδυσης Σεναρίου 3 στην Ουγγαρία.
12. Πίνακας 12 Αξιολόγηση Επένδυσης Σεναρίου 3 στην Ελλάδα.
13. Πίνακας 13 Αξιολόγηση Επένδυσης Σεναρίου 3 στην Αυστρία.
14. Πίνακας 14 Αξιολόγηση Επένδυσης Σεναρίου 3 στη Γαλλία.
15. Πίνακας 15 Αξιολόγηση Επένδυσης Σεναρίου 4.
16. Πίνακας 16 Αξιολόγηση Επένδυσης Σεναρίου 5.

Εισαγωγή

1.1 Διατύπωση του προβλήματος

Η παγκόσμια ενεργειακή αγορά αντιμετωπίζει σημαντικές προκλήσεις στον 21ο αιώνα, καθώς η ζήτηση για ηλεκτρική ενέργεια αυξάνεται και η μετάβαση σε ανανεώσιμες πηγές καθίσταται επιτακτική. Οι τιμές ηλεκτρικής ενέργειας διαμορφώνονται από ένα σύνθετο πλέγμα παραγόντων, όπως οι καιρικές συνθήκες, η εποχικότητα, η ζήτηση, η διαθεσιμότητα ενεργειακών πόρων, οι διακυμάνσεις στις τιμές καυσίμων, καθώς και οι γεωπολιτικές και πολιτικές εξελίξεις [11], [13], [16], [18]. Το αποτέλεσμα είναι ένα περιβάλλον υψηλής αβεβαιότητας, όπου η ακριβής πρόβλεψη των τιμών γίνεται ιδιαίτερα απαιτητική.

Η πρόβλεψη τιμών ηλεκτρικής ενέργειας έχει καθοριστική σημασία για διάφορους εμπλεκόμενους φορείς. Για τους παραγωγούς και παρόχους ενέργειας, αποτελεί εργαλείο βελτιστοποίησης της παραγωγής και της εμπορικής στρατηγικής. Για τους καταναλωτές, μπορεί να διευκολύνει τη στρατηγική κατανάλωση και τον περιορισμό του κόστους. Αντίστοιχα, για τους επενδυτές, η ακριβής εκτίμηση των τιμών είναι θεμελιώδης για την ανάπτυξη αποδοτικών στρατηγικών συναλλαγών που αξιοποιούν τις διακυμάνσεις της αγοράς [14], [37], [38], [39].

Παρά τη σημασία της, η πρόβλεψη των τιμών παραμένει ένα ιδιαίτερα πολύπλοκο πρόβλημα. Οι παραδοσιακές στατιστικές μέθοδοι αποδεικνύονται συχνά ανεπαρκείς στο να αποτυπώσουν τη μη γραμμική και δυναμική φύση της αγοράς [4]. Για τον λόγο αυτό, αποκτά ολοένα και μεγαλύτερη σημασία η αξιοποίηση προηγμένων τεχνικών μηχανικής μάθησης, οι οποίες προσφέρουν αυξημένη ακρίβεια, δυνατότητα ενσωμάτωσης πολύπλοκων σχέσεων και καλύτερη προσαρμοστικότητα στις διακυμάνσεις των δεδομένων [40].

1.2 Προσέγγιση της Μηχανικής Μάθησης

Η τεχνητή νοημοσύνη και η μηχανική μάθηση έχουν παρουσιάσει σημαντική πρόοδο τα τελευταία χρόνια στην ανάλυση χρονοσειρών και στην πρόβλεψη σύνθετων συστημάτων[6],[19],[12]. Σε αυτό το πλαίσιο, η αρχιτεκτονική Transformer, που αρχικά αναπτύχθηκε για εφαρμογές επεξεργασίας φυσικής γλώσσας (NLP), έχει προσελκύσει το ενδιαφέρον της ερευνητικής κοινότητας και για την ανάλυση χρονοσειρών[21], όπως οι τιμές ηλεκτρικής ενέργειας[10]. Ο Transformer ξεχωρίζει για τη δυνατότητά του να επεξεργάζεται χρονοσειρές πιο αποδοτικά και με μεγαλύτερη ακρίβεια σε σχέση με τις παραδοσιακές τεχνικές όπως τα RNNs και LSTMs.

Ένα από τα μεγαλύτερα πλεονεκτήματα του Transformer είναι ο μηχανισμός self-attention, ο οποίος επιτρέπει στο μοντέλο να αναγνωρίζει τις συσχετίσεις μεταξύ των τιμών σε διαφορετικά χρονικά σημεία και να ανιχνεύει σχέσεις τόσο βραχυπρόθεσμες όσο και μακροπρόθεσμες. Σε αντίθεση με τα παραδοσιακά μοντέλα χρονοσειρών που επεξεργάζονται τις τιμές σειριακά, ο Transformer μπορεί να επεξεργάζεται τα δεδομένα παράλληλα, προσφέροντας έτσι ταχύτερη εκπαίδευση και πιο ακριβείς προβλέψεις.

1.3 Ανάπτυξη Επενδυτικής Στρατηγικής

Η αξία της παρούσας εργασίας δεν περιορίζεται μόνο στην ανάπτυξη ενός ακριβούς μοντέλου πρόβλεψης τιμών ηλεκτρικής ενέργειας. Η καινοτομία της έρευνας έγκειται στην εφαρμογή αυτών των προβλέψεων για την ανάπτυξη μιας στρατηγικής επενδύσεων στην αγορά ενέργειας. Με τη χρήση των προβλέψεων του μοντέλου Transformer, αναπτύχθηκε μια επενδυτική στρατηγική που βασίζεται στην εκμετάλλευση των χαμηλών και υψηλών τιμών της αγοράς σε καθημερινή βάση. Αυτή η στρατηγική επιτρέπει την ελαχιστοποίηση των κινδύνων και την αύξηση των αποδόσεων για τους επενδυτές, αξιοποιώντας τις ακριβείς προβλέψεις του μοντέλου.

Η εργασία αυτή επιχειρεί να γεφυρώσει το χάσμα μεταξύ της ακαδημαϊκής έρευνας και της πρακτικής εφαρμογής των προβλέψεων στην αγορά ενέργειας. Με την ανάπτυξη και αξιολόγηση ενός μοντέλου Transformer και την εφαρμογή του στην επενδυτική στρατηγική, φιλοδοξούμε να προσφέρουμε μια σημαντική συνεισφορά στον τομέα της ενεργειακής αγοράς και των επενδύσεων.

Θεωρητικό Υπόβαθρο

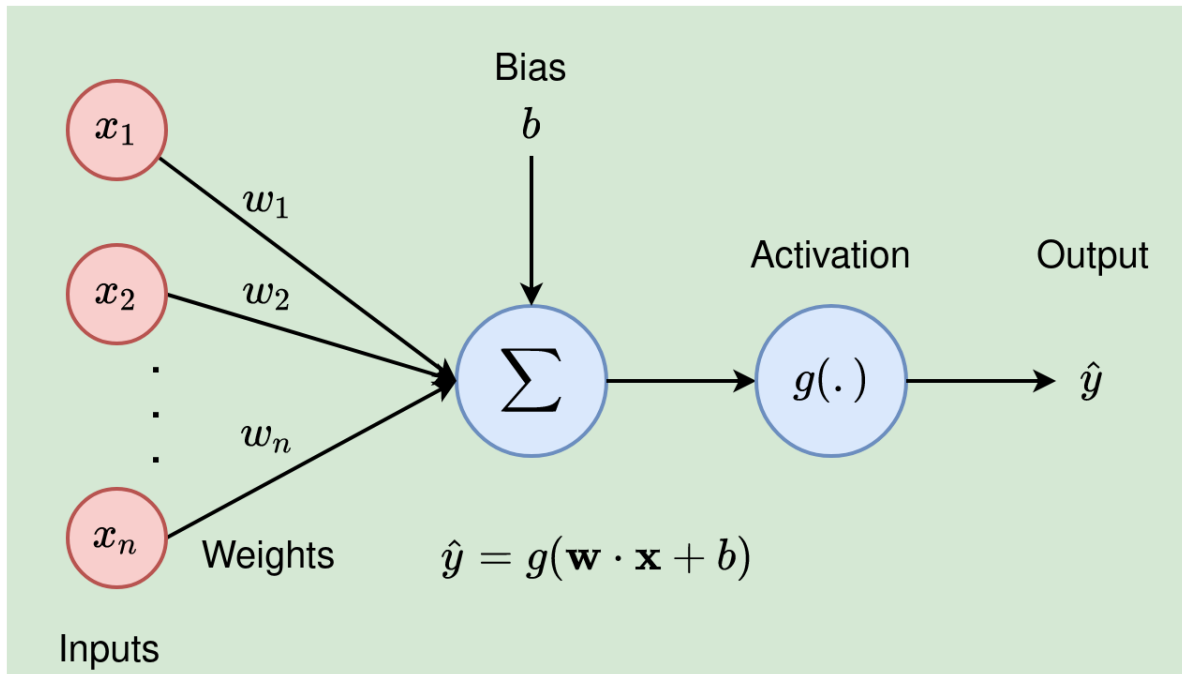
2.1 Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα (Neural Networks) αποτελούν τη βάση της τεχνητής νοημοσύνης και της μηχανικής μάθησης, μιμούμενα τη λειτουργία του ανθρώπινου εγκεφάλου για την επεξεργασία δεδομένων και την επίλυση σύνθετων προβλημάτων. Είναι μια υποκατηγορία των αλγορίθμων μηχανικής μάθησης που αποδεικνύονται εξαιρετικά ισχυρά σε ένα ευρύ φάσμα εφαρμογών, όπως η αναγνώριση εικόνας, η επεξεργασία φυσικής γλώσσας, η πρόβλεψη χρονοσειρών, και πολλά άλλα[36].

Τα νευρωνικά δίκτυα αποτελούνται από δομικές μονάδες που ονομάζονται νευρώνες, οι οποίοι είναι οργανωμένοι σε επίπεδα (layers). Οι νευρώνες σε κάθε επίπεδο λαμβάνουν σήματα από τους νευρώνες του προηγούμενου επιπέδου, επεξεργάζονται αυτές τις πληροφορίες και παράγουν σήματα που στέλνουν στο επόμενο επίπεδο.

2.2 Δομή ενός Νευρώνα

Κάθε νευρώνας είναι μια απλή μαθηματική μονάδα που εκτελεί μια σειρά από υπολογισμούς. Οι είσοδοι (inputs) του νευρώνα πολλαπλασιάζονται με συντελεστές που ονομάζονται βάρη (weights), και το άθροισμα των αποτελεσμάτων περνάει μέσα από μια συνάρτηση ενεργοποίησης (activation function) για να παράξει την έξοδο (output)[36].



Σχήμα 2.2-1 Δομή νευρώνα.

Ο μαθηματικός τύπος που περιγράφει τη λειτουργία ενός νευρώνα είναι ο εξής:

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

όπου:

- x_i είναι οι είσοδοι.
- w_i είναι τα βάρη που πολλαπλασιάζουν τις εισόδους.
- b είναι το bias (προκατάληψη).
- z είναι το αποτέλεσμα πριν την εφαρμογή της συνάρτησης ενεργοποίησης

Στη συνέχεια, το αποτέλεσμα περνάει από μια συνάρτηση ενεργοποίησης:

$$y = f(z)$$

όπου $f(z)$ μπορεί να είναι μια μη γραμμική συνάρτηση, όπως το sigmoid, tanh, ή ReLU (Rectified Linear Unit)

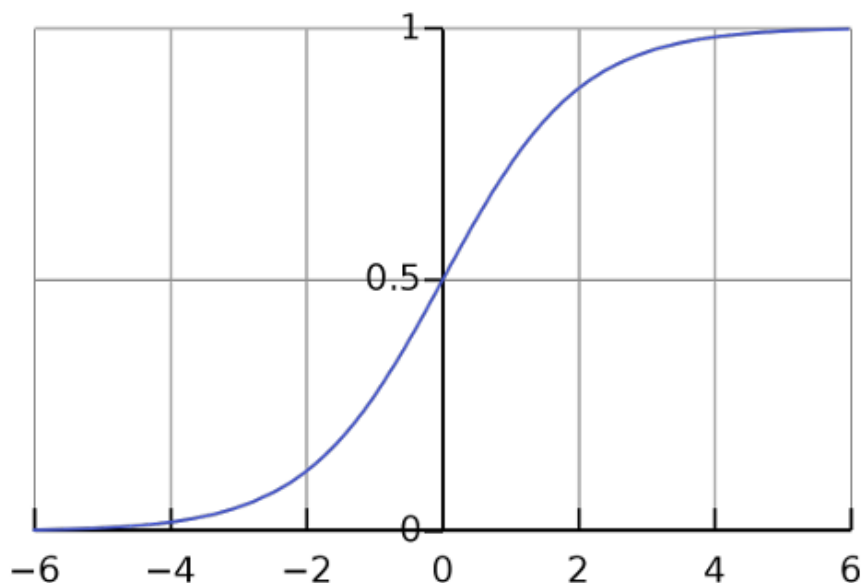
2.3 Συναρτήσεις ενεργοποίησης

2.3.1 Η Σιγμοειδής συνάρτηση:

Η σιγμοειδής συνάρτηση (sigmoid function) δίνεται από τον ακόλουθο τύπο:

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

Για είσοδο μία πραγματική τιμή απεικονίζει στην έξοδο έναν πραγματικό αριθμό στο διάστημα $[0,1]$. Χρησιμοποιείται ευρέως σε ταξινομητές καθώς μικρές μεταβολές της εισόδου x οδηγούν σε μεγάλες μεταβολές της εξόδου y , επιτρέποντας στο δίκτυο να αντιλαμβάνεται ευκολότερα μικρές μεταβολές των χαρακτηριστικών εισόδου[28]. Ωστόσο, σε κάθε ουρά στο 0 ή στο 1, οι τιμές της παραγώγου της είναι πολύ μικρές, συγκλίνοντας στο 0. Ως εκ τούτου, τα διανύσματα κλίσης "εξαφανίζονται" (φαινόμενο vanishing gradient), περιορίζοντας τις δυνατότητες μάθησης του μοντέλου.



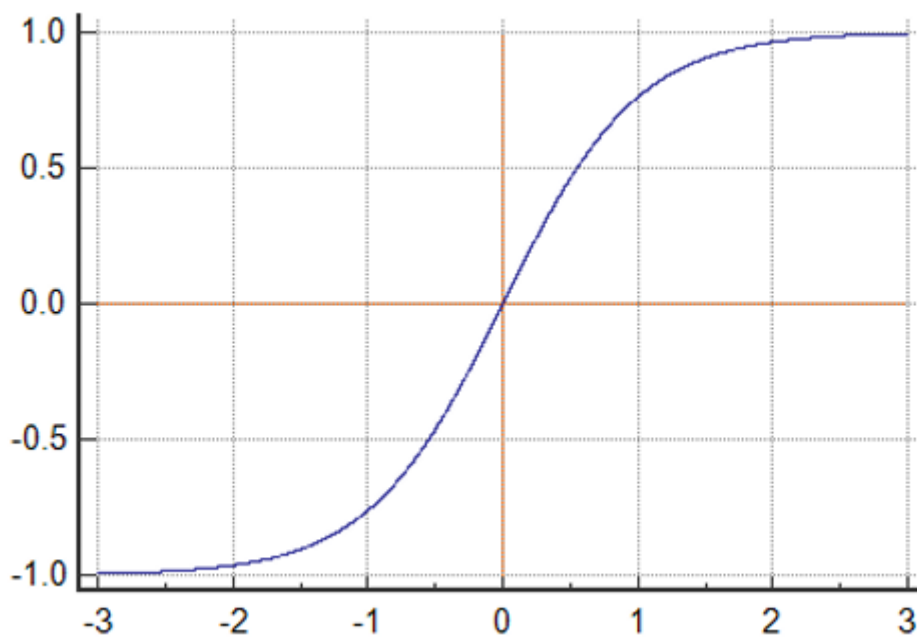
Σχήμα 2.3.1-1 Γραφική παράσταση Σιγμοειδούς συνάρτησης.

2.3.2 Η συνάρτηση Υπερβολικής Εφαπτομένης:

Η συνάρτηση υπερβολικής εφαπτομένης ($\tanh$) δίνεται από τον ακόλουθο μαθηματικό τύπο:

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

Οι τιμές εξόδου της $\tanh$ βρίσκονται στο διάστημα $[-1, 1]$. Βασικό πλεονέκτημα έναντι της σιγμοειδούς, είναι το γεγονός ότι οι παράγωγός της είναι περισσότερο απότομη, γεγονός που μπορεί να οδηγήσει σε μεγαλύτερες τιμές εξόδου[28]. Ως αποτέλεσμα, προσφέρει περισσότερες δυνατότητες για γρήγορη μάθηση και κατάβαση κλίσης. Παρόλα αυτά, το πρόβλημα σύγκλισης της κλίσης στο 0 κοντά στις δύο "ουρές" της συνάρτησης παραμένει.



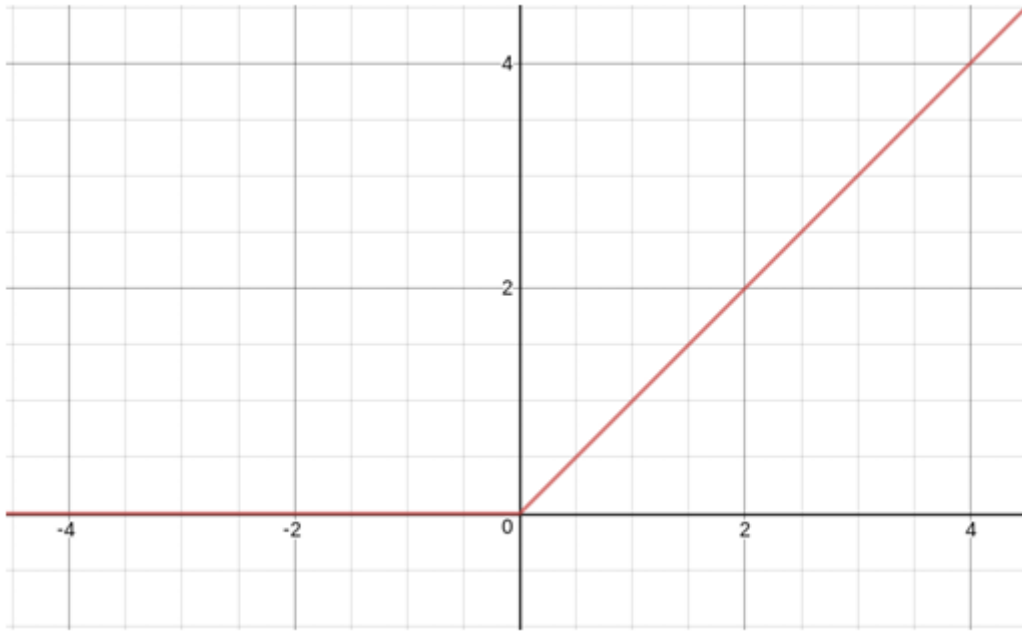
Σχήμα 2.3.2-1 Γραφική παράσταση συνάρτησης Υπερβολικής Εφαπτομένης.

2.3.3 Η Συνάρτηση ReLU:

Η συνάρτηση ενεργοποίησης ReLU (Rectified Linear Unit) δίνεται από τον ακόλουθο τύπο:

$$\text{ReLU}(z) = \max(0, z)$$

Ουσιαστικά αποτελεί ένα κατώφλι της εισόδου στο μηδέν. Σε αντίθεση με τις δύο προηγούμενες συναρτήσεις ενεργοποίησης, δεν περιλαμβάνει περίπλοκες υπολογιστικές πράξεις, με αποτέλεσμα να συγκλίνει πιο γρήγορα[28]. Επιπλέον, παρέχει πιο αραιή ενεργοποίηση, με αποτέλεσμα οι νευρώνες του δικτύου να μαθαίνει πιο σημαντικά χαρακτηριστικά του προβλήματος. Ωστόσο, η διαδικασία μάθησης μπορεί να οδηγηθεί σε τέλμα, όταν οι έξοδοι των νευρώνων δίνουν διαρκώς μηδέν, σε περίπτωση αρνητικής εισόδου. Τέλος, η συνάρτηση ReLU μπορεί να προκαλέσει φαινόμενα exploding gradients κατά τα οποία οι τιμές των παραγώγων αυξάνονται με ραγδαίο ρυθμό. Η συνάρτηση ενεργοποίησης είναι κρίσιμη γιατί εισάγει μη γραμμικότητα στο νευρωνικό δίκτυο, επιτρέποντάς του να μάθει και να εκπροσωπήσει πολύπλοκες σχέσεις μεταξύ των δεδομένων.



Σχήμα 2.3.3-1 Γραφική παράσταση συνάρτησης ReLU.

2.4 Τύποι Νευρωνικών Δικτύων

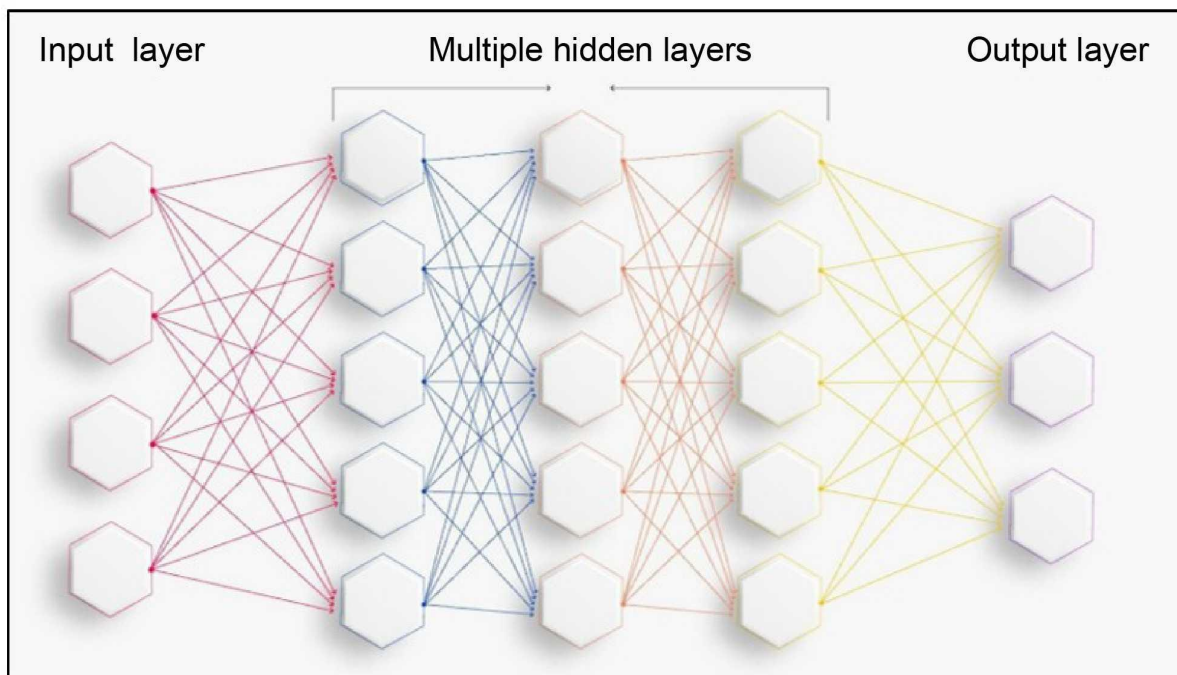
Τα νευρωνικά δίκτυα μπορούν να χωριστούν σε διάφορες κατηγορίες ανάλογα με την αρχιτεκτονική τους και τον τρόπο με τον οποίο συνδέονται οι νευρώνες τους.

2.4.1 Πλήρως Συνδεδεμένα Νευρωνικά Δίκτυα (FCNNs)

Τα πλήρως συνδεδεμένα νευρωνικά δίκτυα, γνωστά και ως πολυεπίπεδοι αντιληπτές (Multilayer Perceptrons - MLPs), αποτελούν μία από τις βασικότερες και πλέον κλασικές αρχιτεκτονικές στον χώρο της τεχνητής νοημοσύνης και της μηχανικής μάθησης. Η βασική τους δομή χαρακτηρίζεται από το γεγονός ότι κάθε νευρώνας ενός επιπέδου συνδέεται με κάθε νευρώνα του επόμενου επιπέδου, δημιουργώντας έτσι έναν πυκνό ιστό συνδέσεων, από τον οποίο προκύπτει και ο όρος "fully connected"[29].

Η αρχιτεκτονική ενός τέτοιου δικτύου αποτελείται από τρία βασικά μέρη: το επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα, και το επίπεδο εξόδου. Το επίπεδο εισόδου περιλαμβάνει τους νευρώνες που δέχονται τα αρχικά δεδομένα, τα οποία μπορεί να είναι χαρακτηριστικά από ένα σύνολο δεδομένων, εικονοστοιχεία από μια εικόνα ή άλλου τύπου αριθμητικές τιμές. Στη συνέχεια, τα δεδομένα μεταβιβάζονται στα κρυφά επίπεδα, όπου λαμβάνει χώρα η κύρια επεξεργασία της πληροφορίας. Κάθε κρυφό επίπεδο περιλαμβάνει πλήρως συνδεδεμένους νευρώνες, οι οποίοι εφαρμόζουν γραμμικούς μετασχηματισμούς στα εισερχόμενα σήματα, συνδυασμένους με μη γραμμικές συναρτήσεις ενεργοποίησης, όπως η ReLU ή η sigmoid. Ο αριθμός των κρυφών επιπέδων και των νευρώνων σε κάθε επίπεδο καθορίζει την πολυπλοκότητα και τη χωρητικότητα του δικτύου να προσαρμόζεται σε πολύπλοκα πρότυπα.

Τέλος, το επίπεδο εξόδου περιλαμβάνει τους νευρώνες που παράγουν το τελικό αποτέλεσμα του δικτύου, είτε αυτό αφορά μία κατηγορία στην περίπτωση ταξινόμησης είτε μία αριθμητική τιμή στην περίπτωση παλινδρόμησης. Παρότι τα FCNNs έχουν χρησιμοποιηθεί εκτενώς με επιτυχία σε πληθώρα εφαρμογών, όπως αναγνώριση προτύπων και πρόβλεψη τάσεων, παρουσιάζουν και περιορισμούς, κυρίως ως προς την αποδοτικότητα και την ικανότητα γενίκευσης σε δεδομένα υψηλής διαστασιμότητας, κάτι που οδήγησε στην ανάπτυξη πιο εξειδικευμένων αρχιτεκτονικών, όπως τα συνελκτικά (CNNs) και τα επαναληπτικά (RNNs) νευρωνικά δίκτυα.



Σχήμα 2.4.1-1 Δομή FCNN.

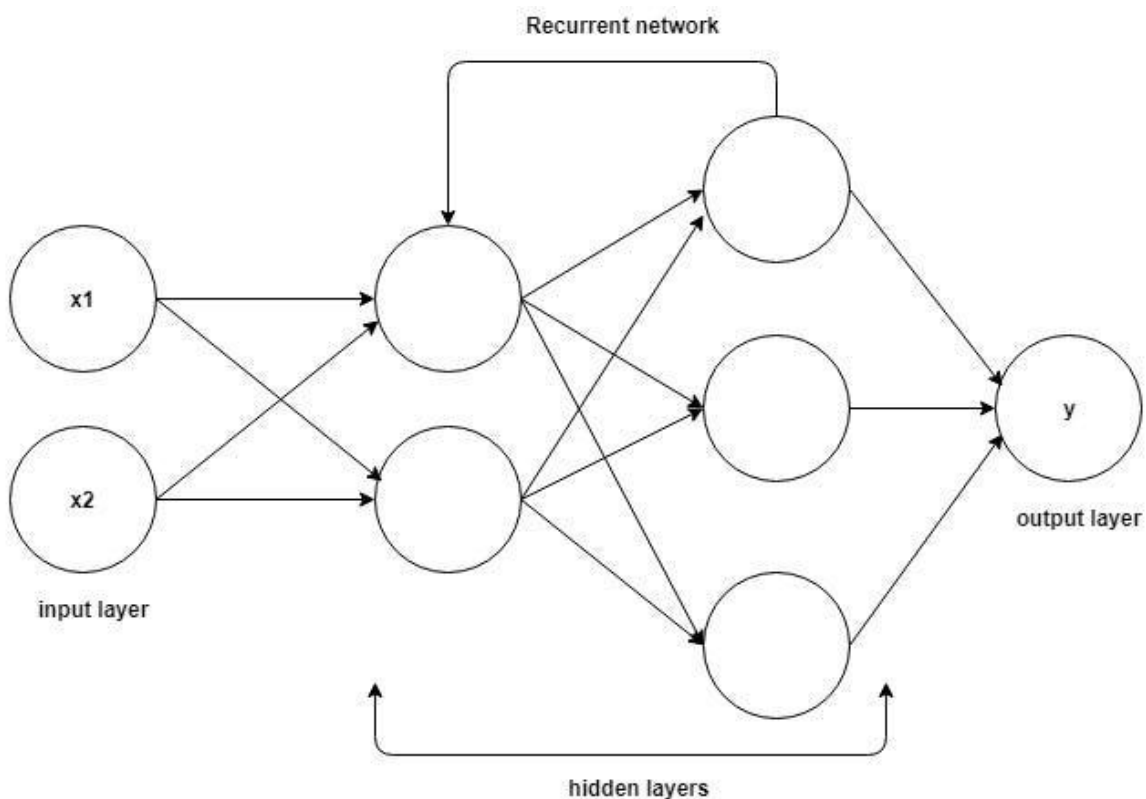
2.4.2 Επαναλαμβανόμενα Νευρωνικά Δίκτυα (RNNs)

Τα Επαναλαμβανόμενα Νευρωνικά Δίκτυα (RNNs) αποτελούν μια ιδιαίτερη κατηγορία τεχνητών νευρωνικών δικτύων, σχεδιασμένων ειδικά για την επεξεργασία ακολουθιακών δεδομένων, όπου η σειρά και η χρονική εξάρτηση των στοιχείων παίζουν καθοριστικό ρόλο. Παραδείγματα τέτοιων δεδομένων περιλαμβάνουν χρονοσειρές, μουσικά κομμάτια, σειρές γεγονότων, και κυρίως φυσική γλώσσα, όπου η σημασία κάθε λέξης εξαρτάται σε μεγάλο βαθμό από τις προηγούμενες.

Η βασική αρχιτεκτονική και καινοτομία των RNNs έγκειται στην ενσωμάτωση μιας εσωτερικής κατάστασης μνήμης, η οποία επιτρέπει στο δίκτυο να «θυμάται» πληροφορίες από προηγούμενα χρονικά βήματα της ακολουθίας. Αυτό επιτυγχάνεται μέσω της αναδρομικής σύνδεσης των νευρώνων: η έξοδος κάθε χρονικού βήματος δεν προωθείται μόνο προς τα έξω, αλλά τροφοδοτείται ξανά ως είσοδος στο ίδιο το δίκτυο στο επόμενο χρονικό βήμα. Έτσι, σε αντίθεση με τα πλήρως συνδεδεμένα δίκτυα, τα RNNs διαθέτουν έναν μηχανισμό χρονικής εξάρτησης, ο οποίος τα καθιστά ιδανικά για προβλήματα που απαιτούν κατανόηση συμφραζομένων ή εντοπισμό μοτίβων σε διαδοχικά δεδομένα.

Η ιδέα αυτή, αν και θεωρητικά ισχυρή, συνοδεύεται και από πρακτικές προκλήσεις. Ένα από τα σημαντικότερα προβλήματα που αντιμετωπίζουν τα βασικά RNNs είναι η εξασθένιση ή έκρηξη των βαθμίδων (vanishing/exploding gradients) κατά την εκπαίδευση μέσω της μεθόδου της ανάδρασης σφάλματος στο χρόνο (Backpropagation Through Time - BPTT)[5]. Το φαινόμενο αυτό περιορίζει την ικανότητα των RNNs να μαθαίνουν μακροπρόθεσμες εξαρτήσεις. Για την αντιμετώπιση του συγκεκριμένου ζητήματος έχουν προταθεί εξελιγμένες εκδοχές των RNNs, όπως τα LSTM (Long Short-Term Memory), τα οποία διαθέτουν πιο πολύπλοκους μηχανισμούς ελέγχου της μνήμης και της ροής της πληροφορίας.

Παρά τους περιορισμούς τους, τα RNNs υπήρξαν θεμέλιο για πολλές εφαρμογές στην επεξεργασία φυσικής γλώσσας, όπως μηχανική μετάφραση, αναγνώριση φωνής, παραγωγή κειμένου και ανάλυση συναισθήματος. Η δυνατότητά τους να μοντελοποιούν χρονικές ή λογικές εξαρτήσεις τα καθιστά ένα σημαντικό εργαλείο, ιδιαίτερα σε περιπτώσεις όπου το πλαίσιο και η ακολουθία των δεδομένων είναι κρίσιμα για την επιτυχία της εκάστοτε εφαρμογής[30].



Σχήμα 2.4.2-1 Δομή RNN.

2.5 Αρχιτεκτονική LSTM (Long Short-Term Memory)

Το LSTM (Long Short-Term Memory) είναι ένας τύπος επαναλαμβανόμενου νευρωνικού δικτύου (Recurrent Neural Network, RNN) που σχεδιάστηκε για να αντιμετωπίζει τα προβλήματα που παρουσιάζουν τα παραδοσιακά RNNs στην εκμάθηση μακροχρόνιων εξαρτήσεων. Τα LSTM έχουν την ικανότητα να διατηρούν πληροφορίες για μεγαλύτερα χρονικά διαστήματα, καθιστώντας τα ιδανικά για προβλήματα χρονοσειρών, όπως η πρόβλεψη τιμών ηλεκτρικής ενέργειας[31].

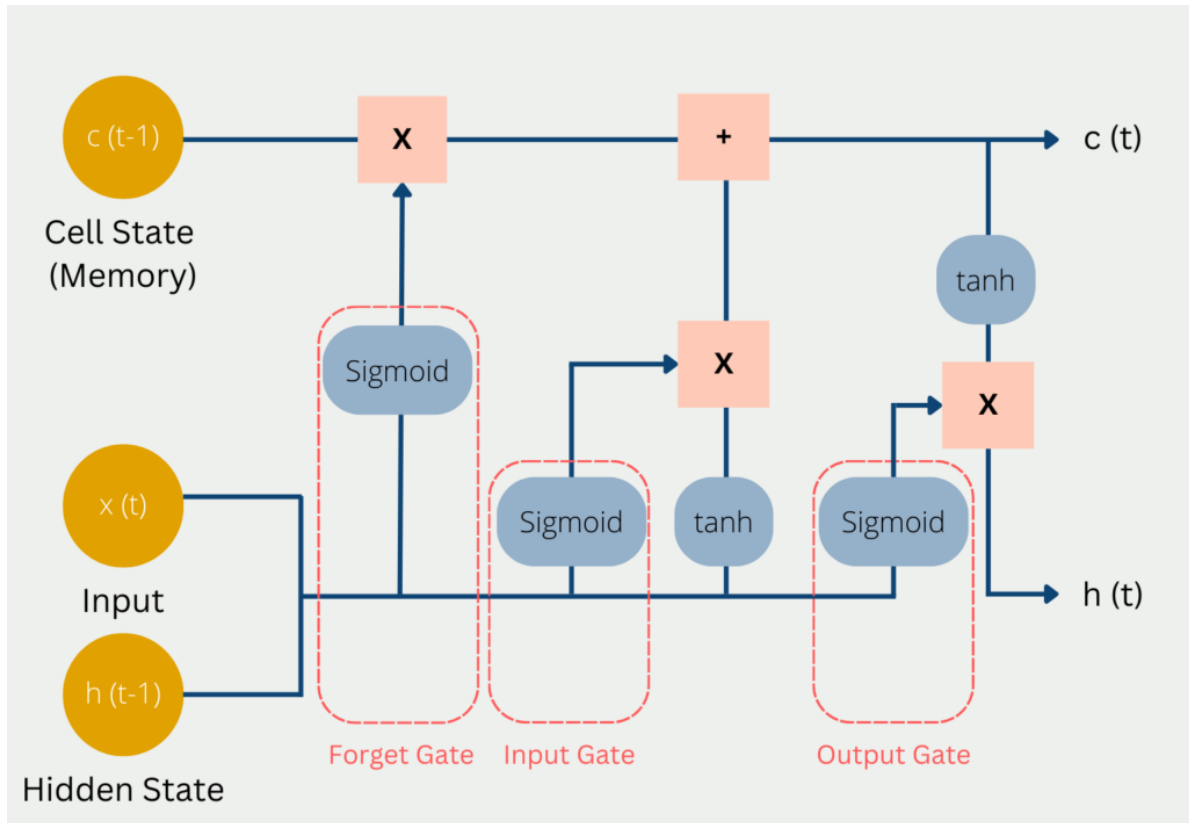
2.5.1 Το πρόβλημα εξασθένησης των βαθμίδων (Vanishing Gradient Problem)

Κατά την εκπαίδευση νευρωνικών δικτύων, εφαρμόζεται η μέθοδος της οπισθοδιάδοσης (backpropagation)[5], μέσω της οποίας υπολογίζεται το σφάλμα της εξόδου και αναπροσαρμόζονται τα βάρη του δικτύου. Η μέθοδος αυτή στηρίζεται στη χρήση του κανόνα της αλυσίδας για τον υπολογισμό των παραγώγων (βαθμίδων) των συναρτήσεων σφάλματος ως προς τα βάρη κάθε επιπέδου, με στόχο τη σταδιακή μείωση του σφάλματος μέσω της βελτιστοποίησης.

Ωστόσο, σε δίκτυα με μεγάλο βάθος ή σε αρχιτεκτονικές με χρονική διάσταση, όπως τα επαναλαμβανόμενα νευρωνικά δίκτυα (RNNs), παρατηρείται συχνά το φαινόμενο της εξασθένησης των βαθμίδων (vanishing gradient)[32]. Το φαινόμενο αυτό περιγράφει την τάση των παραγώγων να αποκτούν πολύ μικρές τιμές καθώς μεταδίδονται προς τα πίσω, από τα τελευταία επίπεδα του δικτύου προς τα πρώτα. Ως αποτέλεσμα, οι παράγωγοι που αντιστοιχούν σε πρώιμα επίπεδα ή πρώιμα χρονικά βήματα καθίστανται αμελητέοι, οδηγώντας σε ελάχιστες ή μηδενικές τροποποιήσεις των αντίστοιχων βαρών. Η συνέπεια είναι το δίκτυο να αδυνατεί να "μάθει" αποδοτικά, καθώς αγνοεί τις μακροπρόθεσμες εξαρτήσεις στην είσοδο, γεγονός που περιορίζει σοβαρά την ικανότητα γενίκευσης σε πολύπλοκα ακολουθιακά δεδομένα. Για να επιλυθεί αυτό το πρόβλημα, εισήχθη η αρχιτεκτονική LSTM από τον Hochreiter και τον Schmidhuber το 1997.

2.5.2 Δομή των LSTM

Το βασικό δομικό στοιχείο των LSTM είναι το "κελί μνήμης" (memory cell), το οποίο επιτρέπει την αποθήκευση πληροφοριών για μακροχρόνια χρήση. Κάθε κελί LSTM περιλαμβάνει τρεις πύλες (gates): την πύλη εισόδου (input gate), την πύλη εξόδου (output gate) και την πύλη λήθης (forget gate). Αυτές οι πύλες ελέγχουν την ροή των πληροφοριών μέσα και έξω από το κελί, επιτρέποντας στο μοντέλο να μαθαίνει ποιες πληροφορίες να αποθηκεύσει και ποιες να διαγράψει καθώς επεξεργάζεται τη χρονοσειρά[31].



Σχήμα 2.5.2-1 Αρχιτεκτονική LSTM.

Παρακάτω παρουσιάζεται κάθε πύλη αναλυτικά:

2.5.3 Πύλη Λήθης (Forget Gate):

Η πύλη λήθης καθορίζει τον βαθμό διατήρησης ή απόρριψης των προηγούμενων πληροφοριών, ελέγχοντας ποια δεδομένα θα παραμείνουν στη μνήμη του μοντέλου και ποια θα απορριφθούν ως μη χρήσιμα. Αυτό επιτυγχάνεται μέσω ενός sigmoid ενεργοποιητή, ο οποίος λαμβάνει ως είσοδο το προηγούμενο κρυφό επίπεδο και την τρέχουσα είσοδο της χρονοσειράς.

Υπολογισμός:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

όπου:

- f_t είναι η έξοδος της πύλης λήθης.
- W_f είναι οι συντελεστές που μαθαίνονται.
- h_{t-1} είναι το προηγούμενο κρυφό επίπεδο.
- x_t είναι η τρέχουσα είσοδος.
- b_f είναι το bias

2.5.4 Πύλη Εισόδου (Input Gate):

Η πύλη εισόδου ελέγχει πόση από την τρέχουσα πληροφορία θα αποθηκευτεί στο κελί μνήμης. Χρησιμοποιείται ένας sigmoid ενεργοποιητής για να καθορίσει πόση πληροφορία θα περάσει και ένας tanh ενεργοποιητής για την προσαρμογή της πληροφορίας.

Υπολογισμός:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

όπου:

- i_t είναι η έξοδος της πύλης εισόδου.
- C_t είναι η νέα πληροφορία που προτείνεται να αποθηκευτεί.
- W_i, W_C είναι οι συντελεστές που μαθαίνονται.
- b_i, b_C είναι τα bias.

2.5.5 Πύλη Εξόδου (Output Gate):

Η πύλη εξόδου καθορίζει την τελική έξοδο του κελιού, δηλαδή το νέο κρυφό επίπεδο, το οποίο χρησιμοποιείται στην επόμενη χρονική στιγμή και σε ενδεχόμενες προβλέψεις.

Υπολογισμός:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$
$$h_t = o_t \cdot \tanh(C_t)$$

όπου:

- o_t είναι η έξοδος της πύλης εξόδου.
- h_t είναι το νέο κρυφό επίπεδο (output).

2.5.6 Λειτουργία του Κελιού LSTM

Η λειτουργία ενός κελιού LSTM μπορεί να περιγραφεί σε δύο φάσεις. Στη φάση εισόδου, η τρέχουσα είσοδος της χρονοσειράς (x_t) και η έξοδος της προηγούμενης στιγμής (h_{t-1}) χρησιμοποιούνται για τον υπολογισμό των πυλών (λήθης, εισόδου, εξόδου). Το κελί αποφασίζει ποιες πληροφορίες από την τρέχουσα είσοδο και τις προηγούμενες πληροφορίες θα αποθηκεύσει ή θα διαγράψει. Στη φάση εξόδου, η πύλη εξόδου αποφασίζει ποια θα είναι η τρέχουσα έξοδος του κελιού (κρυφό επίπεδο h_t) και η ενημέρωση του κελιού μνήμης[31].

2.5.7 Πλεονεκτήματα και Εφαρμογές των LSTM

Τα νευρωνικά δίκτυα τύπου Long Short-Term Memory (LSTM) αποτελούν μία από τις πιο σημαντικές εξελίξεις στον χώρο των επαναλαμβανόμενων νευρωνικών δικτύων (Recurrent Neural Networks - RNNs), καθώς επιλύουν βασικούς περιορισμούς των κλασικών RNNs, ιδιαίτερα σε ό,τι αφορά τη μάθηση και διαχείριση μακροπρόθεσμων εξαρτήσεων σε ακολουθιακά δεδομένα.

Ένα από τα βασικότερα πλεονεκτήματα των LSTM είναι η ικανότητά τους να μαθαίνουν και να διατηρούν μακροχρόνιες εξαρτήσεις. Σε αντίθεση με τα παραδοσιακά RNNs, τα οποία τείνουν να «ξεχνούν» σημαντικές πληροφορίες που σχετίζονται με προηγούμενα χρονικά βήματα, τα LSTM διαθέτουν μια πιο σύνθετη εσωτερική αρχιτεκτονική, που περιλαμβάνει κυψέλες μνήμης και πύλες ελέγχου (όπως input, forget και output gates)[31]. Αυτά τα δομικά στοιχεία τους επιτρέπουν να αποθηκεύουν, διατηρούν ή απορρίπτουν πληροφορία με στοχευμένο και δυναμικό τρόπο, εξασφαλίζοντας έτσι τη διατήρηση κρίσιμων δεδομένων για μεγάλα χρονικά διαστήματα. Η δυνατότητα αυτή είναι ιδιαίτερα πολύτιμη σε περιπτώσεις όπου η τρέχουσα απόφαση ή πρόβλεψη εξαρτάται από γεγονότα που συνέβησαν πολύ νωρίτερα στην ακολουθία.

Παράλληλα, τα LSTM παρουσιάζουν εξαιρετική ευελιξία στη διαχείριση τόσο βραχυπρόθεσμων όσο και μακροπρόθεσμων εξαρτήσεων μέσα σε χρονοσειρές, προσαρμόζόμενα δυναμικά στην εκάστοτε φύση των δεδομένων. Αυτό τα καθιστά ιδιαίτερα αποτελεσματικά σε εφαρμογές όπου η πληροφορία αναπτύσσεται σταδιακά και δεν έχει πάντα σταθερή διάρκεια ή ρυθμό. Ένα ακόμα σημαντικό πλεονέκτημα είναι η αντιμετώπιση του προβλήματος της εξασθένησης των βαθμίδων (vanishing gradient), το οποίο περιορίζει τη δυνατότητα εκμάθησης σε βαθιά ή εκτεταμένα χρονικά δίκτυα[32]. Η εσωτερική μνήμη των LSTM και οι μηχανισμοί πύλης που διαθέτουν επιτρέπουν τη σταθερή ροή πληροφορίας προς τα πίσω κατά την εκπαίδευση, διατηρώντας επαρκές σήμα για την ανανέωση των βαρών ακόμα και σε μακρινά χρονικά βήματα.

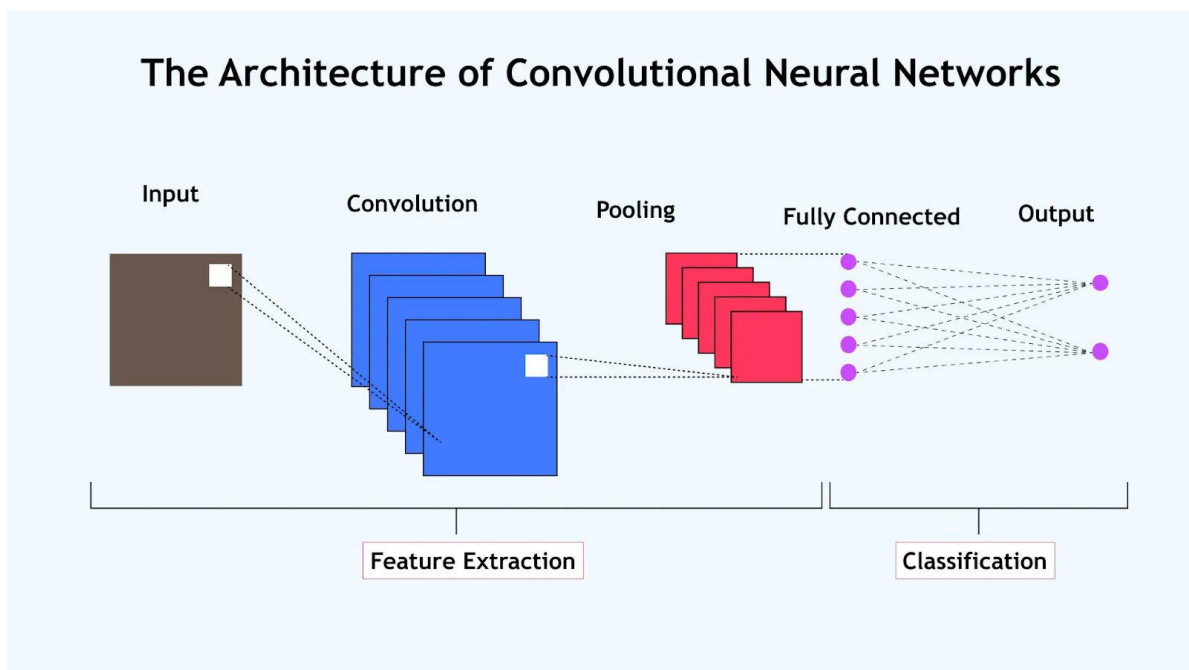
Η αποτελεσματικότητα των LSTM έχει οδηγήσει στην εκτεταμένη υιοθέτησή τους σε πληθώρα πρακτικών εφαρμογών, όπου απαιτείται μοντελοποίηση χρονικά εξαρτώμενων δεδομένων. Στον τομέα της πρόβλεψης χρονοσειρών, χρησιμοποιούνται για την πρόβλεψη πολύπλοκων και μη γραμμικών μοτίβων, όπως τιμές ενέργειας, καιρού, αποθεμάτων ή χρηματοοικονομικών δεικτών, προσφέροντας υψηλή ακρίβεια και αξιοπιστία[15]. Στον τομέα της επεξεργασίας φυσικής γλώσσας, τα LSTM εφαρμόζονται σε εργασίες όπως η μηχανική μετάφραση, η ανάλυση συναισθήματος, η αυτόματη παραγωγή κειμένου και η κατανόηση συμφραζομένων, όπου η αλληλουχία των λέξεων και το γραμματικό πλαίσιο είναι κρίσιμα. Επιπλέον, σε εφαρμογές αναγνώρισης ομιλίας, τα LSTM συνεισφέρουν στην αποδοτική μετατροπή φωνής σε κείμενο, επιτρέποντας την ακριβέστερη ερμηνεία των φωνολογικών και σημασιολογικών χαρακτηριστικών της ανθρώπινης ομιλίας[31].

Συνοψίζοντας, τα δίκτυα LSTM αποτελούν ένα εξαιρετικά ισχυρό εργαλείο για την ανάλυση και πρόβλεψη ακολουθιακών δεδομένων, προσφέροντας λύσεις σε προβλήματα που μέχρι πρότινος περιόριζαν σημαντικά τις επιδόσεις των νευρωνικών δικτύων σε χρονικά δυναμικά περιβάλλοντα. Η ευρεία τους εφαρμογή σε διαφορετικούς επιστημονικούς και τεχνολογικούς τομείς καταδεικνύει τη σημασία τους στην εξέλιξη της τεχνητής νοημοσύνης και της μηχανικής μάθησης.

2.6 Αρχιτεκτονική των Συνελικτικών Νευρωνικών Δικτύων (CNN)

Τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks, CNN) είναι ένας εξειδικευμένος τύπος νευρωνικών δικτύων, σχεδιασμένος για την αποτελεσματική επεξεργασία δεδομένων που έχουν χωρική ή χρονική διάρθρωση, όπως οι εικόνες, τα βίντεο και τα δεδομένα χρονοσειρών. Τα CNN έχουν αποδειχθεί ιδιαίτερα αποτελεσματικά στην ανάλυση και αναγνώριση εικόνων, λόγω της ικανότητάς τους να ανιχνεύουν σημαντικά χαρακτηριστικά και μοτίβα από τα δεδομένα εισόδου[33].

Τα CNN στηρίζονται στην έννοια της συνελικτικής πράξης (convolution operation), η οποία επιτρέπει στο δίκτυο να αναλύει δεδομένα με τη βοήθεια φίλτρων (kernels) που εντοπίζουν χαρακτηριστικά σε μικρότερες περιοχές της εικόνας ή των δεδομένων εισόδου. Το βασικό πλεονέκτημα των CNN είναι ότι εκμεταλλεύονται την τοπική συσχέτιση μεταξύ γειτονικών pixels, μειώνοντας τον αριθμό των παραμέτρων που χρειάζεται να μάθουν, σε σύγκριση με τα πλήρως συνδεδεμένα δίκτυα.



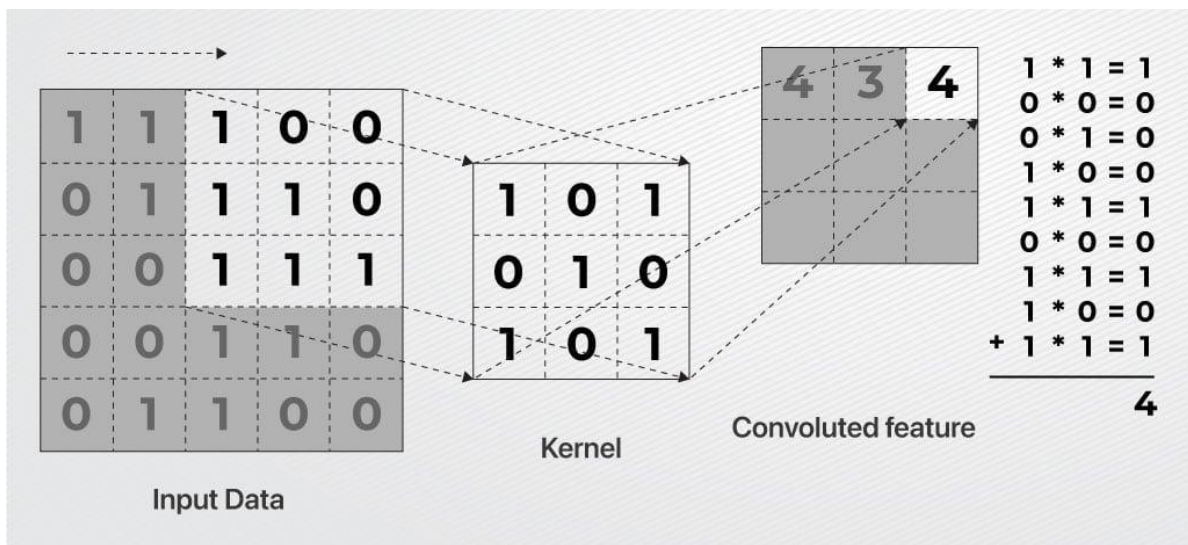
Σχήμα 2.6-1 Αρχιτεκτονική CNN.

Η τυπική αρχιτεκτονική των CNN αποτελείται από τα εξής κύρια επίπεδα:

- Συνελικτικά Επίπεδα (Convolutional Layers)
- Επίπεδα Pooling (Pooling Layers)
- Πλήρως Συνδεδεμένα Επίπεδα (Fully Connected Layers)

2.6.1 Συνελικτικά Επίπεδα (Convolutional Layers)

Τα συνελικτικά επίπεδα είναι η καρδιά ενός CNN και εκτελούν την συνελικτική πράξη (convolution), όπου ένα σύνολο φίλτρων (kernels) εφαρμόζεται στα δεδομένα εισόδου, παράγοντας χάρτες χαρακτηριστικών (feature maps). Τα φίλτρα αυτά "σαρώνουν" την είσοδο σε μικρά παράθυρα και ανιχνεύουν συγκεκριμένα χαρακτηριστικά, όπως άκρα, γωνίες, υφές ή άλλα μοτίβα[33].



Σχήμα 2.6.1-1 Διαδικασία Συνέλιξης.

Η συνελικτική πράξη μπορεί να περιγραφεί μαθηματικά ως:

$$y = f(W * x + b)$$

όπου:

- x είναι τα δεδομένα εισόδου (π.χ., μια εικόνα ή χρονοσειρά).
- W είναι τα φίλτρα (kernels).
- $*$ αντιπροσωπεύει τη συνελικτική πράξη.
- b είναι ο όρος προκατάληψης (bias).
- f είναι η μη γραμμική συνάρτηση ενεργοποίησης, συνήθως η ReLU.

Τα φίλτρα εφαρμόζονται με τρόπο συνελικτικό επί των δεδομένων εισόδου, σαρώνοντας τοπικές περιοχές της εικόνας ή του εκάστοτε εισερχόμενου σήματος. Μέσω αυτής της διαδικασίας εξάγονται χάρτες χαρακτηριστικών (feature maps), οι οποίοι αποτυπώνουν τα εντοπισμένα μοτίβα και δομές που είναι σημαντικά για την εκμάθηση του μοντέλου. Κάθε φίλτρο είναι υπεύθυνο για την ανίχνευση ενός συγκεκριμένου μοτίβου, και καθώς προχωρά η εκπαίδευση, τα φίλτρα μαθαίνουν να ανιχνεύουν όλο και πιο περίπλοκα χαρακτηριστικά[33].

Συχνά, η συνάρτηση ενεργοποίησης που χρησιμοποιείται μετά τη συνελικτική πράξη είναι η ReLU (Rectified Linear Unit). Η ReLU επιτρέπει μόνο τις θετικές τιμές να περάσουν, μηδενίζοντας τις αρνητικές τιμές, κάτι που προσδίδει μη γραμμικότητα στο δίκτυο και βοηθά στη μάθηση σύνθετων χαρακτηριστικών.

2.6.2 Επίπεδα Pooling (Pooling Layers)

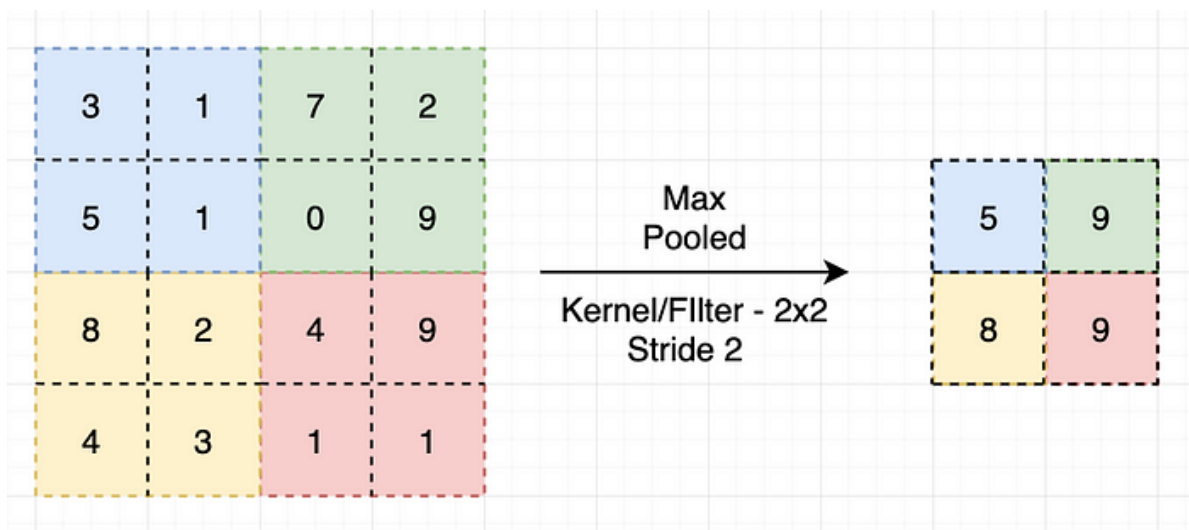
Τα επίπεδα pooling είναι υπεύθυνα για τη μείωση της διαστασιμότητας των χάρτων χαρακτηριστικών που παράγονται από τα συνελικτικά επίπεδα[33]. Αυτή η μείωση

επιτυγχάνεται με τη χρήση λειτουργιών όπως η max pooling, η οποία παίρνει το μέγιστο στοιχείο από μια μικρή περιοχή κάθε χάρτη χαρακτηριστικών.

Ο στόχος του pooling είναι να διατηρηθούν τα πιο σημαντικά χαρακτηριστικά ενώ μειώνονται οι διαστάσεις των δεδομένων, κάτι που οδηγεί σε μικρότερη υπολογιστική απαίτηση και λιγότερες παραμέτρους. Η τυπική πράξη max pooling εφαρμόζεται σε μικρές περιοχές (συνήθως 2x2 ή 3x3) και μπορεί να περιγραφεί μαθηματικά ως:

$$y = \max(x_1, x_2, \dots, x_n)$$

όπου $x_1, x_2, \dots, x_n$ είναι τα στοιχεία της περιοχής στην οποία εφαρμόζεται το pooling.



Σχήμα 2.6.2-1 Pooling Layer.

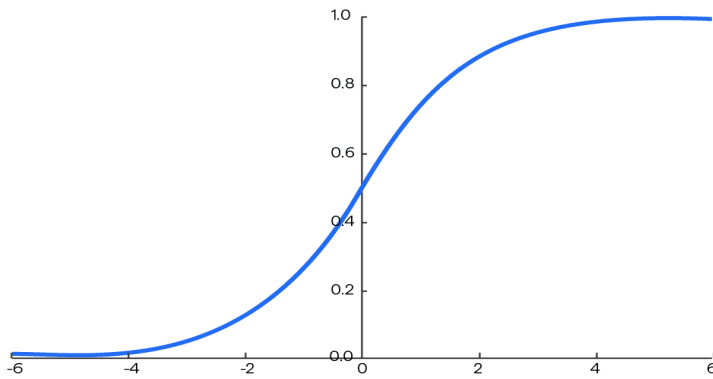
Η χρήση των pooling layers βοηθά επίσης στην αποφυγή υπερεκπαίδευσης, καθώς μειώνει την υπερβολική ευαισθησία του δικτύου σε μικρές μεταβολές των δεδομένων.

2.6.3 Πλήρως Συνδεδεμένα Επίπεδα (Fully Connected Layers)

Αφού τα δεδομένα περάσουν μέσα από τα συνελκτικά και τα επίπεδα pooling, στο τέλος της αρχιτεκτονικής ενός CNN προστίθενται ένα ή περισσότερα πλήρως συνδεδεμένα επίπεδα. Αυτά τα επίπεδα μοιάζουν με τα επίπεδα ενός τυπικού νευρωνικού δικτύου (Fully Connected Neural Network, FCNN), όπου κάθε νευρώνας συνδέεται με κάθε νευρώνα στο επόμενο επίπεδο[33].

Τα πλήρως συνδεδεμένα επίπεδα λαμβάνουν τα επίπεδα των χαρτών χαρακτηριστικών και τα μετασχηματίζουν σε μια γραμμική μορφή (flattening), για να δημιουργηθεί η τελική πρόβλεψη του δικτύου. Σε ένα πρόβλημα ταξινόμησης, για παράδειγμα, το τελευταίο πλήρως συνδεδεμένο επίπεδο μπορεί να έχει τόσους νευρώνες όσες οι κατηγορίες του προβλήματος και να χρησιμοποιεί μια συνάρτηση softmax[17] για την πρόβλεψη των πιθανοτήτων:

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$$



Σχήμα 2.6.4-1 Η συνάρτηση Softmax.

2.6.4 Εφαρμογή των CNNs σε Χρονοσειρές (Time Series)

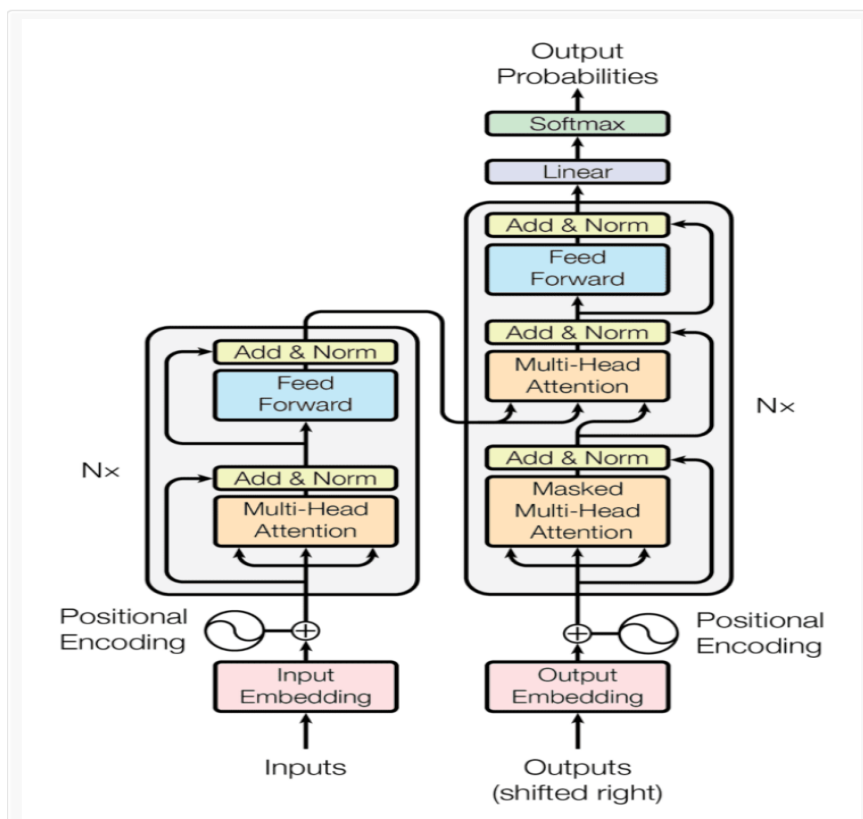
Η βασική λογική της εφαρμογής των CNNs σε χρονοσειρές στηρίζεται στη δυνατότητα των συνελκτικών φίλτρων (convolutional filters) να ανιχνεύουν μοτίβα μέσα στα δεδομένα. Σε χρονοσειρές, τα δεδομένα έχουν χρονική εξάρτηση, και ο στόχος είναι να εντοπιστούν τα μοτίβα ή οι τάσεις που σχετίζονται με την προβλεπόμενη τιμή σε επόμενες χρονικές στιγμές[34].

Σε μια εικόνα, τα φίλτρα ενός CNN ανιχνεύουν χωρικά μοτίβα (π.χ., άκρα, γωνίες) σε μικρές περιοχές της εικόνας. Παρομοίως, στις χρονοσειρές, τα φίλτρα ανιχνεύουν χρονικά μοτίβα σε μικρά χρονικά διαστήματα (π.χ., τάσεις ή επαναλαμβανόμενα μοτίβα) που μπορούν να χρησιμοποιηθούν για την πρόβλεψη μελλοντικών τιμών.

2.7 Vanilla transformer

Ο Vanilla Transformer είναι ένα από τα πιο ισχυρά και καινοτόμα μοντέλα νευρωνικών δικτύων που παρουσιάστηκαν από την ομάδα των Vaswani et al. το 2017. Έχει φέρει επανάσταση στη μηχανική μάθηση, ιδίως στον τομέα της επεξεργασίας φυσικής γλώσσας, της πρόβλεψης χρονοσειρών και της μετάφρασης κειμένου. Η δύναμή του έγκειται στην αρχιτεκτονική του, η οποία βασίζεται εξ ολοκλήρου στον μηχανισμό self-attention και όχι σε επαναλαμβανόμενες συνδέσεις όπως τα RNN ή LSTM[20].

Παρακάτω ακολουθεί μια αναλυτική εξήγηση της αρχιτεκτονικής του Vanilla Transformer, καλύπτοντας όλα τα βασικά στοιχεία του:



Σχήμα 2.7.1-1 Αρχιτεκτονική Vanilla Transformer.

2.7.1 Αρχιτεκτονική του Vanilla Transformer

Ο Vanilla Transformer αποτελείται από δύο κύρια μέρη, Encoder (Κωδικοποιητής) και Decoder (Αποκωδικοποιητής). Κάθε ένα από αυτά τα μέρη αποτελείται από πολλαπλά επίπεδα (layers), με τον Encoder και τον Decoder να διαδραματίζουν διαφορετικούς ρόλους. Στον πυρήνα της αρχιτεκτονικής βρίσκεται ο μηχανισμός multi-head self-attention, ο οποίος επιτρέπει στο Transformer να επικεντρώνεται σε διαφορετικά σημεία της ακολουθίας εισόδου κατά τη διάρκεια της επεξεργασίας[20].

2.7.1.1 Κωδικοποιητής (Encoder)

Ο Encoder είναι υπεύθυνος για τη λήψη των δεδομένων εισόδου (π.χ., μια πρόταση ή μια χρονοσειρά) και τη μετατροπή τους σε μια αναπαράσταση υψηλής διάστασης (κρυφή αναπαράσταση). Ο Encoder αποτελείται από N επαναλαμβανόμενα επίπεδα, και κάθε επίπεδο περιλαμβάνει δύο κύρια υποσυστήματα:

- **Μηχανισμός Πολυκεφαλικής Προσοχής (Multi-Head Self-Attention Mechanism)**
- **Πλήρως Συνδεδεμένο Νευρωνικό Δίκτυο (Feed-Forward Neural Network)**

Multi-Head Self-Attention

Ο μηχανισμός self-attention επιτρέπει στο μοντέλο να αναγνωρίζει τις συσχετίσεις μεταξύ των διαφορετικών στοιχείων της ακολουθίας εισόδου. Για παράδειγμα, αν το Transformer επεξεργάζεται μια πρόταση, το self-attention επιτρέπει στο μοντέλο να αντιληφθεί ποιες λέξεις σχετίζονται με άλλες και πώς επηρεάζουν το νόημα της πρότασης

Για κάθε στοιχείο της ακολουθίας εισόδου, το μοντέλο δημιουργεί τρία διανύσματα. Το διάνυσμα Query (Q) αντιπροσωπεύει τι αναζητά η τρέχουσα λέξη (ή στοιχείο της χρονοσειράς), το Key (K) όπου δείχνει πώς κάθε άλλο στοιχείο σχετίζεται με την τρέχουσα λέξη, ενώ το διάνυσμα Value (V) περιέχει την ίδια την πληροφορία του στοιχείου.

Το αποτέλεσμα της προσοχής (attention) υπολογίζεται με τον εξής τρόπο:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$

όπου:

- QK^T είναι το εσωτερικό γινόμενο των queries και των keys, το οποίο παράγει έναν πίνακα με τα σκορ προσοχής για κάθε ζεύγος στοιχείων.
- $\sqrt{d_k}$ είναι ένας παράγοντας κανονικοποίησης που εξισορροπεί τα σκορ, καθώς τα διανύσματα ενδέχεται να έχουν μεγάλο μέγεθος.
- Ο πίνακας των σκορ προσοχής κανονικοποιείται με το softmax, ώστε να μετατραπεί σε πιθανότητες.
- Τέλος, οι πιθανότητες αυτές πολλαπλασιάζονται με τα διανύσματα Value (V) για την παραγωγή του τελικού διανύσματος προσοχής.

Η προσθήκη του όρου Multi-Head Attention σημαίνει ότι το μοντέλο δημιουργεί πολλά διαφορετικά σύνολα από Queries, Keys και Values, τα οποία εκτελούν τον παραπάνω υπολογισμό ταυτόχρονα και ανεξάρτητα. Στο τέλος, συνδυάζονται τα αποτελέσματα από τις διάφορες κεφαλές προσοχής.

Δίκτυο Πρόσθιας Τροφοδότησης (Feed-Forward Neural Network)

Μετά τον μηχανισμό προσοχής, η έξοδος περνάει από ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο (FFN) που εφαρμόζεται ανεξάρτητα σε κάθε θέση της ακολουθίας. Το δίκτυο αποτελείται από δύο γραμμικά επίπεδα με μια μη γραμμική ενεργοποίηση (ReLU) μεταξύ τους.

Το FFN σε κάθε θέση εφαρμόζεται με την εξής μορφή:

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2$$

όπου:

- W_1 και W_2 είναι οι συντελεστές των γραμμικών επιπέδων που μαθαίνονται κατά την εκπαίδευση.

- b1 και b2 είναι τα αντίστοιχα bias

Υπολειμματικές Συνδέσεις (Residual Connections)

Για να βελτιωθεί η σταθερότητα της εκπαίδευσης και να διατηρηθεί η ροή πληροφορίας, χρησιμοποιούνται residual connections γύρω από κάθε υποσύστημα (δηλαδή τον μηχανισμό προσοχής και το FFN). Αυτό σημαίνει ότι η είσοδος σε κάθε υποσύστημα προστίθεται στην έξοδο του υποσυστήματος πριν προχωρήσει στο επόμενο.

Κανονικοποίηση Επιπέδου (Layer Normalization)

Για να διατηρηθεί η σταθερότητα κατά την εκπαίδευση, κάθε επίπεδο ακολουθείται από layer normalization, το οποίο διασφαλίζει ότι τα δεδομένα που περνούν στα επόμενα επίπεδα έχουν κατάλληλη κλίμακα.

2.7.1.2 Αποκωδικοποιητής (Decoder)

Ο Decoder είναι παρόμοιος με τον Encoder, με τη βασική διαφορά ότι, εκτός από τα ίδια επίπεδα προσοχής και feed-forward, περιλαμβάνει και ένα επιπλέον επίπεδο προσοχής, το οποίο συνδέεται με την έξοδο του Encoder.

Ο Decoder αποτελείται από τα εξής μέρη:

- Masked Multi-Head Attention: Όπως ο μηχανισμός self-attention στον Encoder, αλλά με μια σημαντική διαφορά: χρησιμοποιείται "μάσκα" ("masking") για να αποτρέψει το μοντέλο από το να κοιτάζει τις μελλοντικές λέξεις κατά την εκπαίδευση.
- Multi-Head Attention (με προσοχή στην έξοδο του Encoder): Ο Decoder χρησιμοποιεί την έξοδο του Encoder και δημιουργεί συσχετίσεις μεταξύ των θέσεων εισόδου και εξόδου.
- Feed-Forward Neural Network (FFN): Το πλήρως συνδεδεμένο δίκτυο λειτουργεί όπως και στον Encoder.

Ο Decoder παίρνει ως είσοδο τις κρυφές αναπαραστάσεις από τον Encoder και τις πληροφορίες της προηγούμενης κατάστασης εξόδου (π.χ., τα στοιχεία που έχουν ήδη προβλεφθεί). Ο μηχανισμός masking βοηθάει να μην χρησιμοποιούνται μελλοντικές πληροφορίες, διατηρώντας έτσι την αιτιότητα στη μοντελοποίηση των ακολουθιών.

2.7.1.3 Κωδικοποίηση Θέσης (Position Encoding)

Ο Transformer, σε αντίθεση με τα RNN, δεν έχει εγγενή σειρά στα δεδομένα που εισάγονται. Για να αντιληφθεί τη σειρά των δεδομένων (π.χ., τη χρονική ακολουθία σε χρονοσειρές), χρησιμοποιεί position encodings. Αυτές οι κωδικοποιήσεις θέσης προστίθενται στην είσοδο και παρέχουν στο μοντέλο πληροφορίες για τη σειρά των στοιχείων.

Η θέση κωδικοποιείται με τη μορφή συνημιτόνου (sin) και ημιτόνου (cos) με βάση την εξής συνάρτηση:

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000^{\frac{2i}{d}}}\right)$$
$$PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{\frac{2i}{d}}}\right)$$

όπου:

- pos είναι η θέση στην ακολουθία.
- i είναι ο δείκτης της διάστασης.
- d είναι η διάσταση της εισόδου.

2.7.2 Παράδειγμα λειτουργίας Transformer

Έστω περίπτωση στην οποία με χρήση Transformer γίνεται πρόβλεψη τιμών ηλεκτρικής ενέργειας για τις επόμενες 24 ώρες κάνοντας χρήση των ιστορικών δεδομένων τιμών ενέργειας των προηγούμενων 24 ωρών.

Αντιπροσώπευση των Δεδομένων Εισόδου

Η είσοδος στο μοντέλο αποτελείται από μια χρονοσειρά 24 διαδοχικών τιμών, οι οποίες αντιστοιχούν στις τιμές της ηλεκτρικής ενέργειας κατά τις τελευταίες 24 ώρες. Αυτές οι τιμές μορφοποιούνται ως αριθμητικό διάνυσμα εισόδου, το οποίο μετατρέπεται περαιτέρω ώστε να καταστεί κατάλληλο για επεξεργασία από το μοντέλο.

Κωδικοποίηση Θέσης

Καθώς το Transformer δεν διαθέτει ενσωματωμένη έννοια χρονικής διαδοχής, είναι απαραίτητο να προστεθούν πληροφορίες θέσης σε κάθε στοιχείο της εισόδου. Αυτό επιτυγχάνεται μέσω της χρήσης Positional Encoding, το οποίο ενισχύει την ακολουθία των τιμών με πληροφορία σχετικά με τη χρονική τους σειρά. Η προσθήκη αυτή επιτρέπει στο μοντέλο να λαμβάνει υπόψη τη χρονική εξάρτηση μεταξύ των σημείων της χρονοσειράς.

Δομή του Encoder

Το υποσύστημα του Encoder λαμβάνει ως είσοδο την κωδικοποιημένη χρονοσειρά και παράγει μια αναπαράσταση υψηλότερου επιπέδου, η οποία ενσωματώνει τις συσχετίσεις μεταξύ των τιμών στις διαφορετικές χρονικές στιγμές. Η διαδικασία αυτή περιλαμβάνει τον μηχανισμό αυτοπροσοχής (self-attention), μέσω του οποίου κάθε τιμή της σειράς συσχετίζεται με τις υπόλοιπες, επιτρέποντας την ανάδειξη σημαντικών χρονικών σχέσεων και μοτίβων. Ακολουθεί ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο (Feed-Forward Network), το οποίο μετασχηματίζει περαιτέρω τα ενδιάμεσα χαρακτηριστικά.

Δομή του Decoder και Παραγωγή Προβλέψεων

Ο Decoder έχει ως κύριο ρόλο την παραγωγή των προβλεπόμενων τιμών. Η διαδικασία είναι αυτοπαλίνδρομη (autoregressive), καθώς η κάθε προβλεπόμενη τιμή χρησιμοποιείται ως είσοδος για την πρόβλεψη της επόμενης. Η παραγωγή ξεκινά με βάση την έξοδο του Encoder και συνεχίζεται διαδοχικά: για την πρόβλεψη της 25ης ώρας, το μοντέλο βασίζεται αποκλειστικά στα δεδομένα εισόδου. Η πρόβλεψη αυτή στη συνέχεια ενσωματώνεται ως νέα είσοδος για την παραγωγή της τιμής της 26ης ώρας, και ούτω καθεξής.

Η διαδικασία συνεχίζεται έως ότου παραχθούν οι τιμές για ολόκληρο το μελλοντικό χρονικό παράθυρο των 24 ωρών. Κάθε νέα πρόβλεψη εξαρτάται τόσο από τις προηγούμενες όσο και από την κωδικοποιημένη αναπαράσταση της αρχικής χρονοσειράς.

Χρήση Masked Multi-Head Self-Attention στον Decoder

Κατά τη φάση εκπαίδευσης, εφαρμόζεται τεχνική masked attention στον Decoder, με σκοπό την αποτροπή της πρόσβασης του μοντέλου σε μελλοντικές χρονικές στιγμές κατά την πρόβλεψη κάθε επόμενης τιμής. Έτσι, διασφαλίζεται ότι η διαδικασία παραγωγής της κάθε πρόβλεψης βασίζεται αποκλειστικά σε παλαιότερες και όχι σε μελλοντικές πληροφορίες, κάτι που είναι κρίσιμο για την εγκυρότητα της χρονοσειριακής πρόβλεψης.

Τελική Έξοδος

Το αποτέλεσμα της διαδικασίας είναι μια ακολουθία 24 προβλεπόμενων τιμών, οι οποίες παριστούν τις εκτιμώμενες τιμές ενέργειας για τις επόμενες 24 ώρες. Η σειρά αυτή παράγεται βήμα προς βήμα, με κάθε πρόβλεψη να αξιοποιείται ως είσοδος για την επόμενη χρονική στιγμή. Αυτή η προσέγγιση αποδεικνύει την ικανότητα του μοντέλου Transformer να διαχειρίζεται προβλήματα πρόβλεψης χρονοσειρών πολλαπλών βημάτων (multi-step forecasting) με υψηλή ακρίβεια, αξιοποιώντας τις ισχυρές ιδιότητες του μηχανισμού προσοχής.

Πειραματικές Μέθοδοι και Υλοποίηση

3.1 Δεδομένα

Στην περίπτωση της ανάλυσης και πρόβλεψης της αγοράς ενέργειας στην Ουγγαρία, χρησιμοποιούνται δεδομένα από την περίοδο 2010 έως 2019[24]. Τα δεδομένα περιλαμβάνουν τόσο τις ωριαίες τιμές ενέργειας όσο και τις ωριαίες τιμές θερμοκρασίας. Παρακάτω αναλύονται τα χαρακτηριστικά των δεδομένων, οι πηγές και οι πληροφορίες που περιέχουν.

Τα δεδομένα αποτελούνται από τρία βασικά μέρη:

- Δεδομένα Τιμών Ενέργειας: Περιλαμβάνουν τις ωριαίες τιμές της ενέργειας για την αγορά της Ουγγαρίας από το 2010 έως το 2019.
- Δεδομένα χρονικής στιγμής
- Δεδομένα Θερμοκρασίας: Περιλαμβάνουν τις ωριαίες τιμές θερμοκρασίας κατά την ίδια χρονική περίοδο, από το 2010 έως το 2019.

Ακολουθεί αναλυτική περιγραφή των χαρακτηριστικών:

Τιμή Ενέργειας (Hourly Energy Price)

Η Τιμή Ενέργειας είναι η ωριαία τιμή της ηλεκτρικής ενέργειας στην αγορά της Ουγγαρίας ευρώ ανά μεγαβατώρα (€/MWh). Καταγράφεται ανά ώρα και αντανακλά τη διακύμανση της τιμής κατά τη διάρκεια της ημέρας. Τα δεδομένα αυτά είναι ιδιαίτερα σημαντικά για την κατανόηση των διακυμάνσεων της αγοράς ενέργειας και των παραγόντων που επηρεάζουν τις τιμές σε βραχυπρόθεσμο και μακροπρόθεσμο επίπεδο.

Χρονική στιγμή (Timestamp)

Πεδία που περιλαμβάνονται στα δεδομένα χρονικής στιγμής:

- Έτος: Δείχνει το έτος καταγραφής της τιμής (π.χ., 2010, 2011, ..., 2019).
- Μήνας: Δείχνει τον μήνα του έτους (π.χ., Ιανουάριος, Φεβρουάριος, ..., Δεκέμβριος).
- Ημέρα: Δείχνει την ημέρα του μήνα (π.χ., 1, 2, ..., 31).
- Ώρα: Δείχνει την ώρα της ημέρας (π.χ., 00:00, 01:00, ..., 23:00).

Θερμοκρασία (Hourly Temperature)

Η Θερμοκρασία αποτελεί κρίσιμο χαρακτηριστικό για την πρόβλεψη της τιμής ενέργειας, καθώς η κατανάλωση ενέργειας συνδέεται στενά με τις θερμοκρασίες του

περιβάλλοντος. Σε χαμηλές θερμοκρασίες για παράδειγμα, η ζήτηση ενέργειας αυξάνεται λόγω της θέρμανσης, ενώ σε υψηλές θερμοκρασίες αυξάνεται λόγω της χρήσης κλιματισμού. Επομένως, η θερμοκρασία παρέχει μια βασική ένδειξη για την πρόβλεψη των μελλοντικών τιμών ενέργειας. Στα δεδομένα μας για κάθε ώρα της ημέρας αναγράφεται και η αντιστοιχη θερμοκρασία.

Τα δεδομένα οργανώνονται ως χρονοσειρές, όπου κάθε γραμμή του συνόλου δεδομένων αντιπροσωπεύει μια συγκεκριμένη ώρα της ημέρας και περιλαμβάνει τις πληροφορίες ημερομηνίας, τιμής ενέργειας και θερμοκρασίας για την συγκεκριμένη ώρα. Αυτή η δομή επιτρέπει τη χρήση τους σε προβλήματα πρόβλεψης χρονοσειρών, καθώς και τη μοντελοποίηση της αλληλεπίδρασης μεταξύ της τιμής ενέργειας και της θερμοκρασίας σε διαφορετικές χρονικές περιόδους.

Timestamp	Hourly Price (Euro/MWh)	Temperature (C)
2020-01-01 0:00:00	33.55	-1.2
2020-01-01 1:00:00	30.4	-1.3
2020-01-01 2:00:00	28.53	-1.2
2020-01-01 3:00:00	27.37	-1.5
2020-01-01 4:00:00	26.38	-1.6
2020-01-01 5:00:00	26.5	-1.8
2020-01-01 6:00:00	27.01	-1.6
2020-01-01 7:00:00	27.4	-1.6
2020-01-01 8:00:00	27.99	-1.6
2020-01-01 9:00:00	27.71	-1.1
2020-01-01 10:00:00	28.4	0.9

Πίνακας 1 Δομή Δεδομένων Εργασίας.

Για την αξιολόγηση της επίδοσης του μοντέλου στη πρόβλεψη των τιμών ηλεκτρικής ενέργειας, διεξήχθησαν δύο πειράματα με διαφορετικές προσεγγίσεις ως προς τα δεδομένα εισόδου. Στόχος των πειραμάτων ήταν να εκτιμηθεί κατά πόσο η προσθήκη της ωριαίας θερμοκρασίας ως εισόδου ενισχύει την ακρίβεια της πρόβλεψης, συγκριτικά με την απλή χρήση των ιστορικών τιμών ενέργειας και των στοιχείων χρονικής στιγμής. Ακολουθεί μια αναλυτική περιγραφή της μεθοδολογίας και των δύο πειραμάτων, καθώς και της διαδικασίας διαχωρισμού των δεδομένων σε σύνολα εκπαίδευσης και δοκιμής.

3.2 Πειραματική Διαδικασία και Διαχωρισμός Δεδομένων

Για τη σωστή αξιολόγηση της επίδοσης του μοντέλου, το σύνολο δεδομένων χωρίστηκε σε δύο υποσύνολα:

- Training Set (Σύνολο Εκπαίδευσης): Το 92% του συνόλου δεδομένων χρησιμοποιήθηκε για την εκπαίδευση του μοντέλου, επιτρέποντάς του να μάθει τα μοτίβα και τις σχέσεις που υπάρχουν στις τιμές ενέργειας.
- Test Set (Σύνολο Δοκιμής): Το 8% των δεδομένων χρησιμοποιήθηκε για την αξιολόγηση της απόδοσης του μοντέλου σε άγνωστα δεδομένα, ελέγχοντας πόσο καλά μπορεί να γενικεύσει τις προβλέψεις του σε νέα δεδομένα.

Ο διαχωρισμός αυτός διασφαλίζει ότι το μοντέλο αξιολογείται ανεξάρτητα από τα δεδομένα εκπαίδευσης, και έτσι επιτρέπει τη σωστή εκτίμηση της ικανότητας του μοντέλου να προβλέπει μελλοντικές τιμές.

Πείραμα 1: Πρόβλεψη με Ιστορικές Τιμές και Δεδομένα Χρονικής Στιγμής

Στο πρώτο πείραμα, το μοντέλο εκπαιδεύτηκε και αξιολογήθηκε χρησιμοποιώντας μόνο τις ιστορικές τιμές ενέργειας και τις χρονικές πληροφορίες (έτος, μήνας, ημέρα και ώρα) ως είσοδο. Συγκεκριμένα, τα δεδομένα εισόδου για κάθε χρονική στιγμή περιλάμβαναν:

- Ιστορικές Τιμές Ενέργειας για τις προηγούμενες 48 ώρες, οι οποίες δίνουν στο μοντέλο πληροφορίες σχετικά με τις πρόσφατες διακυμάνσεις στην τιμή της ενέργειας.
- Χρονικές Στιγμές (έτος, μήνας, ημέρα, ώρα), οι οποίες βοηθούν το μοντέλο να αναγνωρίζει εποχιακά μοτίβα και τις διακυμάνσεις ανά ώρα της ημέρας.

Η επιλογή των ιστορικών τιμών και των χρονικών στιγμών έχει στόχο να εντοπίσει τις βασικές τάσεις στις τιμές ενέργειας και να εκμεταλλευτεί τις πληροφορίες που σχετίζονται με την εποχικότητα και τους ημερήσιους κύκλους της αγοράς ενέργειας. Αυτή η προσέγγιση βασίζεται στη λογική ότι οι τιμές ενέργειας παρουσιάζουν κυκλικότητα και μοτίβα που επαναλαμβάνονται ανάλογα με τη χρονική περίοδο (π.χ., αυξημένη ζήτηση τις ώρες αιχμής ή τους χειμερινούς μήνες).

Ωστόσο, αυτή η μέθοδος μπορεί να περιορίζει την ακρίβεια, καθώς δεν λαμβάνει υπόψη εξωτερικούς παράγοντες, όπως η θερμοκρασία, που επηρεάζουν άμεσα τη ζήτηση και κατά συνέπεια την τιμή της ενέργειας.

Πείραμα 2: Πρόβλεψη με Ιστορικές Τιμές, Δεδομένα Χρονικής Στιγμής και Θερμοκρασία

Στο δεύτερο πείραμα, το μοντέλο εκπαιδεύτηκε και αξιολογήθηκε χρησιμοποιώντας, εκτός από τις ιστορικές τιμές ενέργειας και τα χρονικά δεδομένα, τις τιμές θερμοκρασίας για τις δύο προηγούμενες ημέρες και τις θερμοκρασίες για την ημέρα πρόβλεψης. Συγκεκριμένα, τα δεδομένα εισόδου περιλάμβαναν:

- Ιστορικές Τιμές Ενέργειας για τις προηγούμενες 48 ώρες, που παρέχουν στο μοντέλο γνώση των διακυμάνσεων στην αγορά σε κοντινό χρονικό διάστημα.
- Χρονικές Στιγμές: Έτος, μήνας, ημέρα και ώρα κάθε τιμής, που βοηθούν το μοντέλο να αναγνωρίσει εποχιακά και ημερήσια μοτίβα.
- Τιμές Θερμοκρασίας:
 - Οι ωριαίες θερμοκρασίες των δύο προηγούμενων ημερών, οι οποίες αντανακλούν τη ζήτηση ενέργειας τις προηγούμενες ημέρες λόγω καιρικών συνθηκών.
 - Οι θερμοκρασίες της επόμενης ημέρας. Αυτές παρέχουν στο μοντέλο μια εικόνα της ζήτησης για την περίοδο πρόβλεψης, επιτρέποντάς του να προσαρμόσει τις τιμές ανάλογα με τις καιρικές συνθήκες.

Η προσθήκη των θερμοκρασιών βελτιώνει την πρόβλεψη του μοντέλου, καθώς η θερμοκρασία αποτελεί κρίσιμο παράγοντα που επηρεάζει άμεσα τη ζήτηση ενέργειας. Για παράδειγμα, σε πολύ χαμηλές θερμοκρασίες, η ζήτηση για θέρμανση αυξάνεται, επηρεάζοντας την τιμή της ενέργειας, ενώ σε υψηλές θερμοκρασίες, η χρήση κλιματισμού αυξάνει τη ζήτηση ενέργειας, προκαλώντας αύξηση των τιμών.

Συνεπώς, με τη χρήση των δεδομένων θερμοκρασίας, το μοντέλο μπορεί να προβλέψει πιο αποτελεσματικά τις τιμές ενέργειας, καθώς λαμβάνει υπόψη εξωγενείς παράγοντες που επιδρούν στη ζήτηση και προσφορά.

3.3 Αξιολόγηση Αποτελεσμάτων

Μετά την εκπαίδευση των δύο μοντέλων (ένα ανά πειραματική διάταξη), ακολούθησε η διαδικασία αξιολόγησης της απόδοσής τους, η οποία πραγματοποιήθηκε επί του συνόλου δοκιμής, αντιστοιχώντας στο 8% των συνολικών δεδομένων. Στόχος της αξιολόγησης ήταν η σύγκριση των δύο μοντέλων, με σκοπό να διερευνηθεί ο βαθμός επίδρασης της ενσωμάτωσης των θερμοκρασιακών δεδομένων στην ακρίβεια της πρόβλεψης. Τα αποτελέσματα της σύγκρισης καταδεικνύουν τη σημασία της χρήσης πολυδιάστατων χαρακτηριστικών εισόδου για την ενίσχυση της ακρίβειας των προβλέψεων, επιβεβαιώνοντας τον κρίσιμο ρόλο παραμέτρων όπως η θερμοκρασία στην πρόβλεψη των τιμών ενέργειας.

Για την ποσοτική αξιολόγηση της απόδοσης των μοντέλων εφαρμόστηκαν τα εξής στατιστικά μέτρα σφάλματος: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) και Root Mean Squared Error (RMSE). Η επιλογή αυτών των μετρικών αποσκοπεί στην κάλυψη διαφορετικών διαστάσεων της ακρίβειας, προσφέροντας μια σφαιρική και αξιόπιστη εικόνα της προβλεπτικής ικανότητας κάθε μοντέλου.

Ακολουθεί αναλυτική περιγραφή των μέτρων και η σημασία τους:

3.3.1 Mean Absolute Error (MAE)

Το Mean Absolute Error (MAE) μετρά την μέση απόλυτη διαφορά μεταξύ των προβλεπόμενων τιμών και των πραγματικών τιμών. Το MAE υπολογίζεται ως ο μέσος όρος των απόλυτων διαφορών μεταξύ της πραγματικής τιμής και της τιμής που προβλέφθηκε από το μοντέλο[35].

Η μαθηματική έκφραση για το MAE είναι:

$$\text{MAE} = \frac{\sum_{i=0}^n |y_i - \hat{y}_i|}{n}$$

όπου:

- n είναι ο αριθμός των παρατηρήσεων.
- y_i είναι η πραγματική τιμή.
- $\hat{y}_i$ είναι η προβλεπόμενη τιμή.

Το MAE παρέχει ένα εύκολο στην κατανόηση μέτρο ακρίβειας, καθώς δείχνει την μέση απόσταση μεταξύ των προβλέψεων και των πραγματικών τιμών σε πραγματικές μονάδες (π.χ., €/MWh). Όσο μικρότερο είναι το MAE, τόσο πιο κοντά είναι οι προβλέψεις στις πραγματικές τιμές. Είναι χρήσιμο καθώς δίνει τη μέση απόσταση της πρόβλεψης από την πραγματική τιμή χωρίς να δίνει περισσότερη σημασία σε ακραία σφάλματα, καθώς χρησιμοποιεί την απόλυτη διαφορά. Είναι ιδανικό για την κατανόηση της απόλυτης απόκλισης των προβλέψεων, ανεξάρτητα από το πρόσημο των σφαλμάτων.

3.3.2 Mean Absolute Percentage Error (MAPE)

Το Mean Absolute Percentage Error (MAPE) εκφράζει το σφάλμα ως ποσοστό των πραγματικών τιμών, διευκολύνοντας έτσι την κατανόηση του σφάλματος ως προς το μέγεθος των τιμών της χρονοσειράς[35].

Η μαθηματική έκφραση για το MAPE είναι:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$$

όπου:

- n είναι ο αριθμός των παρατηρήσεων.
- y_i είναι η πραγματική τιμή.
- $\hat{y}_i$ είναι η προβλεπόμενη τιμή.

Το MAPE εκφράζεται σε ποσοστό, επομένως είναι εύκολα ερμηνεύσιμο και χρήσιμο για συγκρίσεις μεταξύ διαφορετικών χρονοσειρών ή διαφορετικών μοντέλων. Ένα MAPE της τάξης του 10% σημαίνει ότι, κατά μέσο όρο, η πρόβλεψη αποκλίνει από την πραγματική τιμή κατά 10%. Είναι ιδιαίτερα χρήσιμο για την εκτίμηση της σχετικής ακρίβειας των προβλέψεων, καθώς παρέχει την απόκλιση σε ποσοστιαία βάση. Ωστόσο, μπορεί να είναι ευαίσθητο σε πολύ χαμηλές πραγματικές τιμές, όπου μικρά σφάλματα μπορεί να οδηγούν σε μεγάλα ποσοστιαία σφάλματα.

3.3.3 Root Mean Squared Error (RMSE)

Το Root Mean Squared Error (RMSE) είναι ένα metric που μετρά την τετραγωνική απόκλιση μεταξύ των πραγματικών και των προβλεπόμενων τιμών. Το RMSE υπολογίζεται ως η τετραγωνική ρίζα του μέσου όρου των τετραγωνικών σφαλμάτων[35].

Η μαθηματική έκφραση για το RMSE είναι:

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

όπου:

- n είναι ο αριθμός των παρατηρήσεων.
- y_i είναι η πραγματική τιμή.
- $\hat{y}_i$ είναι η προβλεπόμενη τιμή.

Το RMSE δίνει μεγαλύτερη βαρύτητα σε μεγαλύτερα σφάλματα, καθώς χρησιμοποιεί το τετράγωνο της διαφοράς. Έτσι, το RMSE είναι ευαίσθητο σε ακραίες τιμές σφάλματος (outliers) και παρέχει μια πιο αυστηρή εκτίμηση της ακρίβειας του μοντέλου σε σχέση με το MAE. Θεωρείται ιδιαίτερα χρήσιμο σε εφαρμογές όπου τα μεγάλα σφάλματα θεωρούνται πιο σοβαρά από τα μικρά σφάλματα, καθώς αυτά τα σφάλματα "τιμωρούνται" περισσότερο. Το RMSE επιτρέπει στο μοντέλο να βελτιώνει τις προβλέψεις του σε περιπτώσεις μεγάλων αποκλίσεων, γεγονός που το καθιστά πολύτιμο σε προβλήματα όπου η ακρίβεια στις υψηλές αποκλίσεις είναι κρίσιμη.

3.3.4 Σύγκριση των Μετρικών Αξιολόγησης

Η μέση απόλυτη απόκλιση (Mean Absolute Error – MAE) υπολογίζει τη μέση τιμή της απόλυτης διαφοράς μεταξύ των προβλεπόμενων και των πραγματικών τιμών. Αποτελεί μια απλή και εύκολα κατανοήσιμη μέτρηση, κατάλληλη για την αποτύπωση της γενικής ακρίβειας ενός μοντέλου. Η μέση απόλυτη ποσοστιαία απόκλιση (Mean Absolute Percentage Error – MAPE) εκφράζει το σφάλμα πρόβλεψης σε ποσοστιαία μορφή, γεγονός που την καθιστά ιδιαίτερα χρήσιμη σε συγκρίσεις μεταξύ διαφορετικών συνόλων δεδομένων ή περιπτώσεων με διαφορετικά μεγέθη. Αντιθέτως, η ριζική μέση τετραγωνική απόκλιση (Root Mean Squared Error – RMSE) δίνει μεγαλύτερη βαρύτητα σε μεγαλύτερα σφάλματα λόγω της τετραγωνικής τιμής τους. Ως αποτέλεσμα, η RMSE είναι περισσότερο ευαίσθητη σε ακραίες αποκλίσεις, προσφέροντας μια πιο αυστηρή εκτίμηση της απόδοσης, ιδίως σε περιπτώσεις όπου τα μεγάλα σφάλματα θεωρούνται ιδιαίτερης σημασίας.

Χρησιμοποιώντας αυτά τα metrics, έχουμε μια σφαιρική αξιολόγηση της απόδοσης του μοντέλου και την ικανότητα του να προβλέπει ακριβώς τόσο σε επίπεδο απόλυτων τιμών όσο και σε ποσοστιαία βάση, ενώ παράλληλα αναγνωρίζονται και διορθώνονται ακραίες αποκλίσεις.

3.3.5 Αποτελέσματα Πειραμάτων

Τα αποτελέσματα των μετρικών MAE, MAPE και RMSE παρουσιάζονται στον παρακάτω πίνακα:

Μοντέλο Transformer	MAE	MAPE	RMSE
<u>Χωρίς θερμοκρασία</u>	8.29	19.84	10.21
<u>Με θερμοκρασία</u>	7.40	16.15	9.37

Πίνακας 2 Δείκτες Ακρίβειας Προβλέψεων.

1. Mean Absolute Error (MAE)

Ο δείκτης Mean Absolute Error (MAE) αποτυπώνει τη μέση απόλυτη διαφορά μεταξύ των προβλεπόμενων και των πραγματικών τιμών της χρονοσειράς. Στο μοντέλο που δεν περιελάμβανε δεδομένα θερμοκρασίας, το MAE ανήλθε σε 8.29. Με την ενσωμάτωση των θερμοκρασιακών δεδομένων, ο δείκτης μειώθηκε σε 7.40. Η διαφορά αυτή καταδεικνύει τη βελτίωση της απόδοσης του μοντέλου, καθώς οι προβλέψεις του βρίσκονται εγγύτερα στις πραγματικές τιμές. Η μείωση του MAE αποδεικνύει ότι η θερμοκρασία ως εξωγενής μεταβλητή προσφέρει επιπρόσθετη πληροφορία που ενισχύει την ικανότητα πρόβλεψης της γενικής τάσης των τιμών ενέργειας, οδηγώντας σε μικρότερα σφάλματα σε απόλυτη τιμή.

2. Mean Absolute Percentage Error (MAPE)

Ο δείκτης Mean Absolute Percentage Error (MAPE) υπολογίζει τη μέση ποσοστιαία απόκλιση των προβλέψεων σε σχέση με τις πραγματικές τιμές, προσφέροντας μια κλιμακούμενη εκτίμηση του σφάλματος. Στην περίπτωση που δεν χρησιμοποιήθηκαν θερμοκρασιακά δεδομένα, το MAPE ήταν 19.84%. Με την προσθήκη της μεταβλητής της θερμοκρασίας, ο δείκτης μειώθηκε σε 16.15%, γεγονός που καταδεικνύει σημαντική βελτίωση στη σχετική ακρίβεια των προβλέψεων. Το χαμηλότερο MAPE επισημαίνει ότι το μοντέλο με θερμοκρασία αποδίδει προβλέψεις με μικρότερη αναλογική απόκλιση, ανεξαρτήτως του μεγέθους των τιμών, γεγονός που το καθιστά πιο σταθερό και αξιόπιστο για δυναμικά περιβάλλοντα τιμολόγησης.

3. Root Mean Squared Error (RMSE)

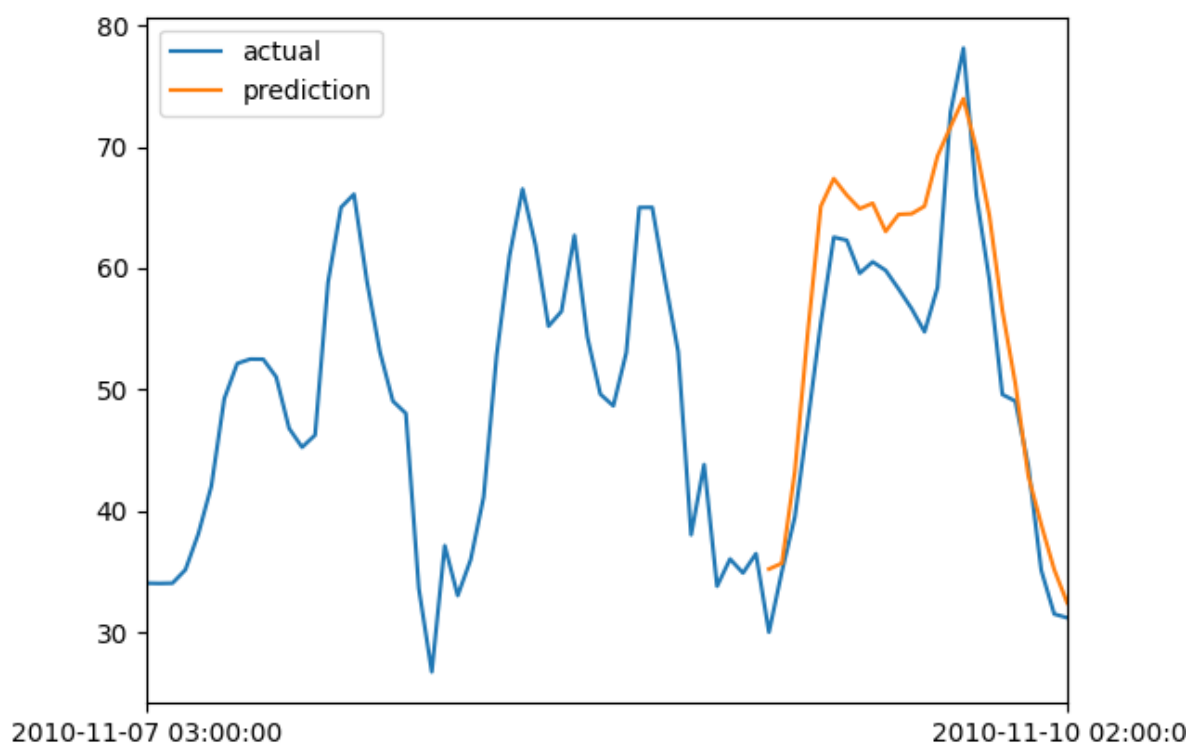
Η μετρική Root Mean Squared Error (RMSE) αποτιμά τη μέση τετραγωνική απόκλιση μεταξύ των προβλεπόμενων και των πραγματικών τιμών, προσδίδοντας μεγαλύτερη βαρύτητα στα μεγάλα σφάλματα λόγω της τετραγωνικής φύσης του υπολογισμού της. Στην περίπτωση του μοντέλου χωρίς τη χρήση θερμοκρασιακών δεδομένων, το RMSE ανήλθε σε 10.21. Αντιθέτως, με την προσθήκη της θερμοκρασίας ως εξωτερικής μεταβλητής, η τιμή μειώθηκε σε 9.37, γεγονός που υποδηλώνει αυξημένη ανθεκτικότητα του μοντέλου σε μεγάλες αποκλίσεις και βελτιωμένη σταθερότητα υπό ακραίες συνθήκες. Η συγκεκριμένη μείωση ενισχύει την άποψη ότι οι μεταβολές

θερμοκρασίας ενσωματώνουν κρίσιμη πληροφορία για την εξέλιξη των τιμών ενέργειας, επιτρέποντας στο μοντέλο να προβλέπει με μεγαλύτερη ακρίβεια ακόμη και σε περιόδους υψηλής μεταβλητότητας.

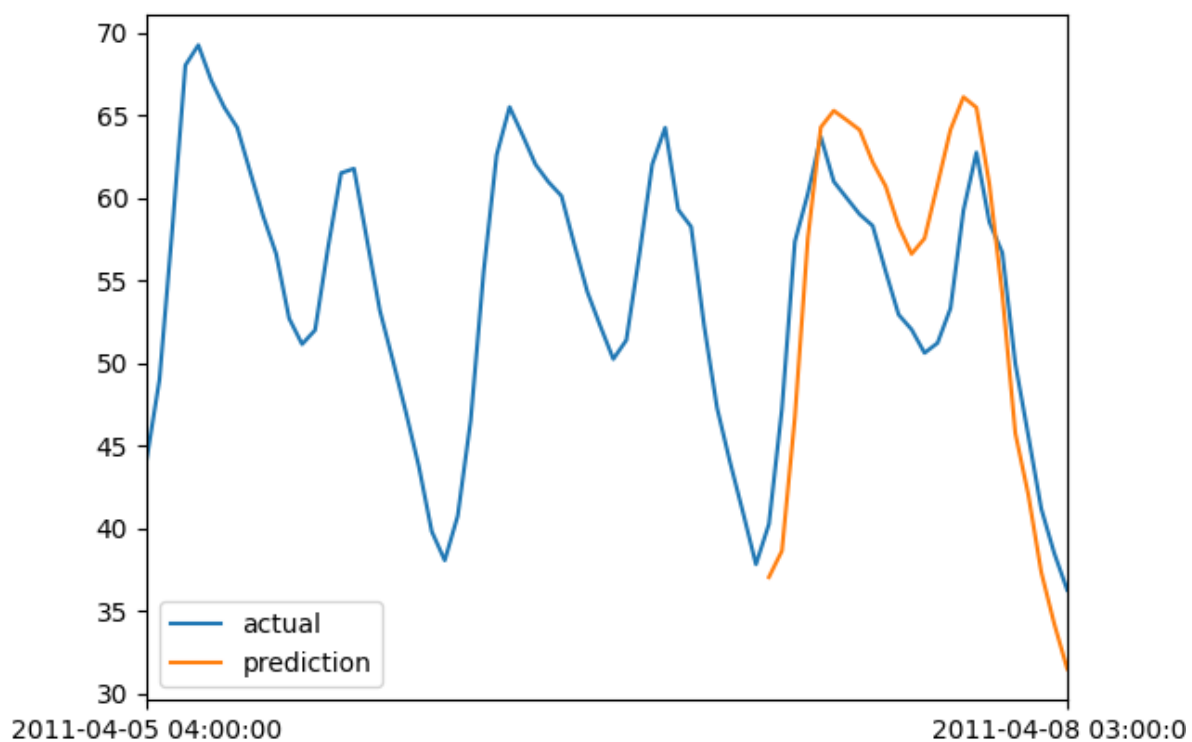
Συνολική Αποτίμηση Απόδοσης Μοντέλου

Η αξιολόγηση της απόδοσης του μοντέλου με τη χρήση των μετρικών MAE, MAPE και RMSE ανέδειξε τη σημαντική συμβολή της ενσωμάτωσης θερμοκρασιακών δεδομένων στην πρόβλεψη των τιμών ηλεκτρικής ενέργειας. Και οι τρεις μετρικές κατέγραψαν αισθητή μείωση όταν προστέθηκε η εξωγενής μεταβλητή της θερμοκρασίας, υποδεικνύοντας βελτίωση τόσο σε απόλυτη όσο και σε σχετική ακρίβεια, αλλά και σε σταθερότητα έναντι ακραίων αποκλίσεων. Συγκεκριμένα, η μείωση του MAE επιβεβαιώνει την καλύτερη προσέγγιση των μέσων τιμών, το χαμηλότερο MAPE αναδεικνύει τη μεγαλύτερη αξιοπιστία των προβλέψεων σε αναλογική βάση, ενώ η πτώση του RMSE καταδεικνύει την αποτελεσματικότερη διαχείριση των μεγάλων σφαλμάτων. Τα αποτελέσματα αυτά επιβεβαιώνουν ότι η ενσωμάτωση εξωτερικών παραμέτρων, όπως η θερμοκρασία, βελτιστοποιεί τη συμπεριφορά του μοντέλου σε πραγματικές συνθήκες αγοράς, καθιστώντας το κατάλληλο για εφαρμογές πρόβλεψης σε πολύπλοκα ενεργειακά περιβάλλοντα.

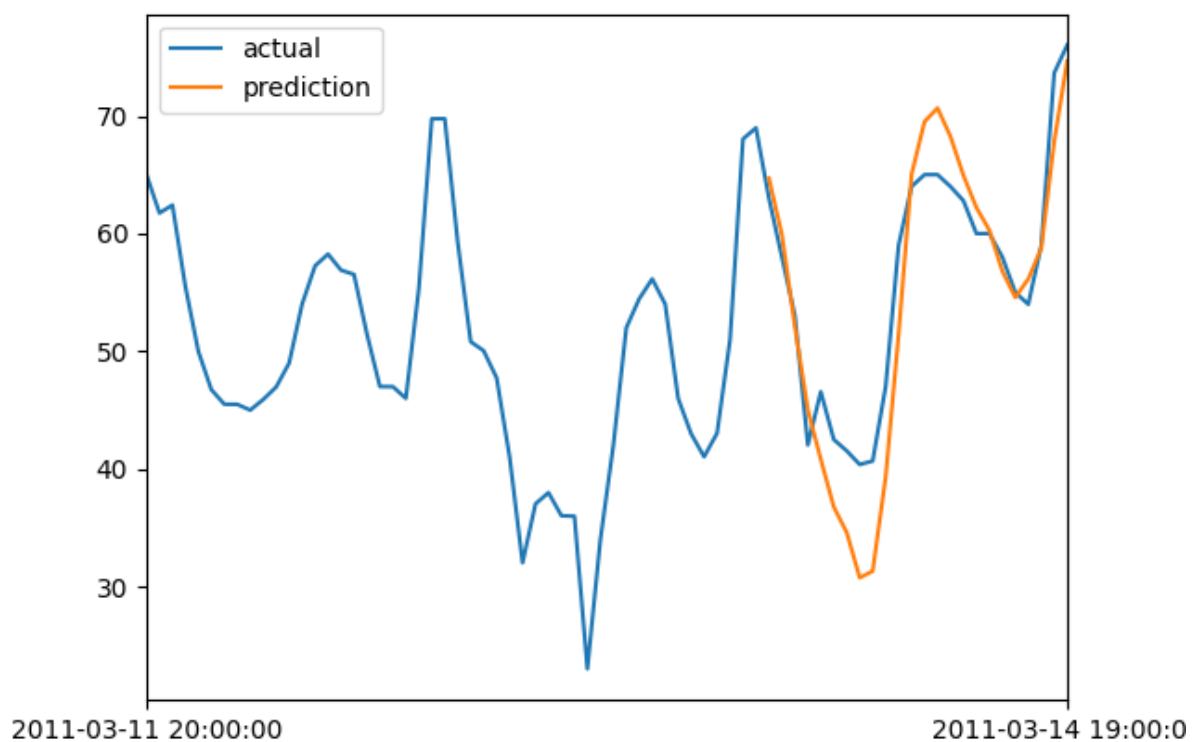
Παραδείγματα Προβλέψεων με Αξιοποίηση Δεδομένων Θερμοκρασίας



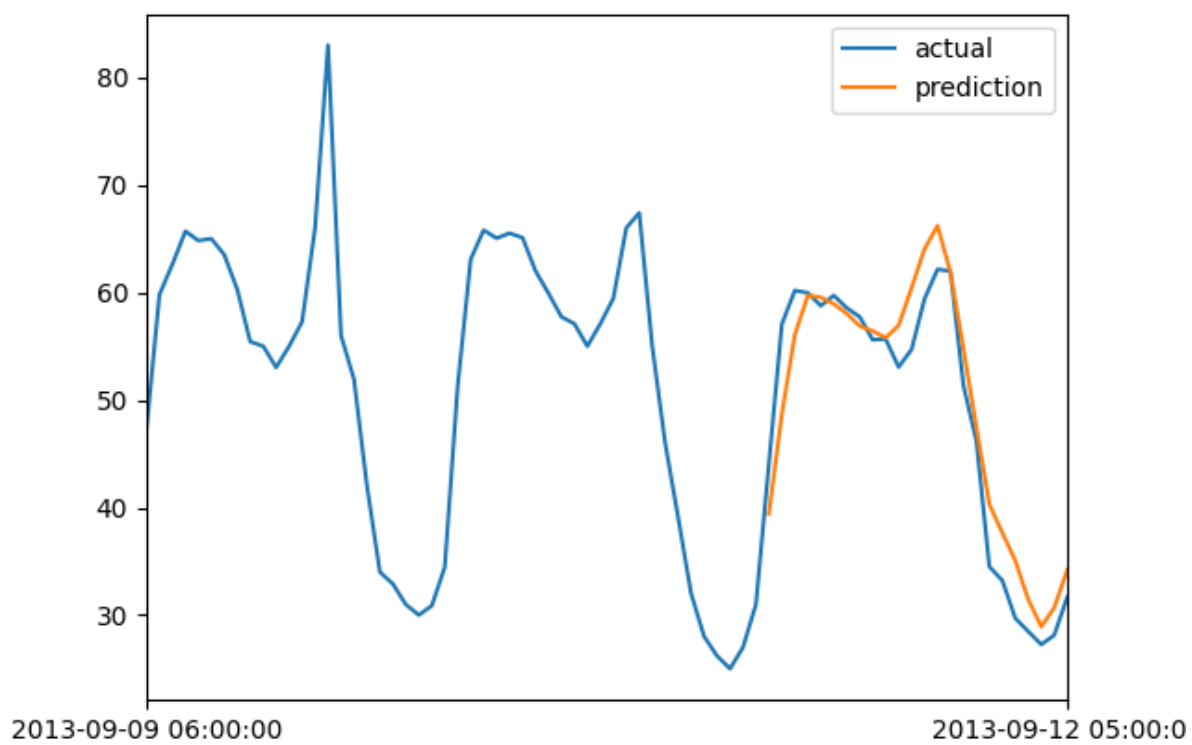
Σχήμα 3.3.4-1 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.



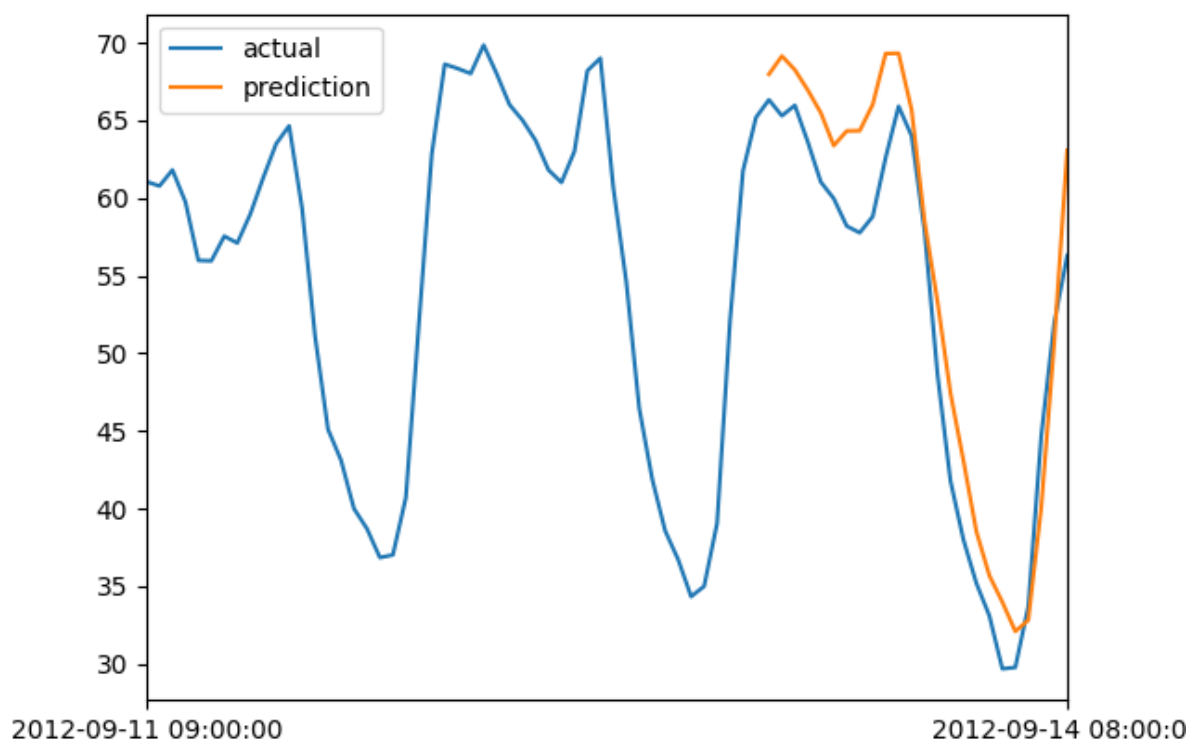
Σχήμα 3.3.4-2 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.



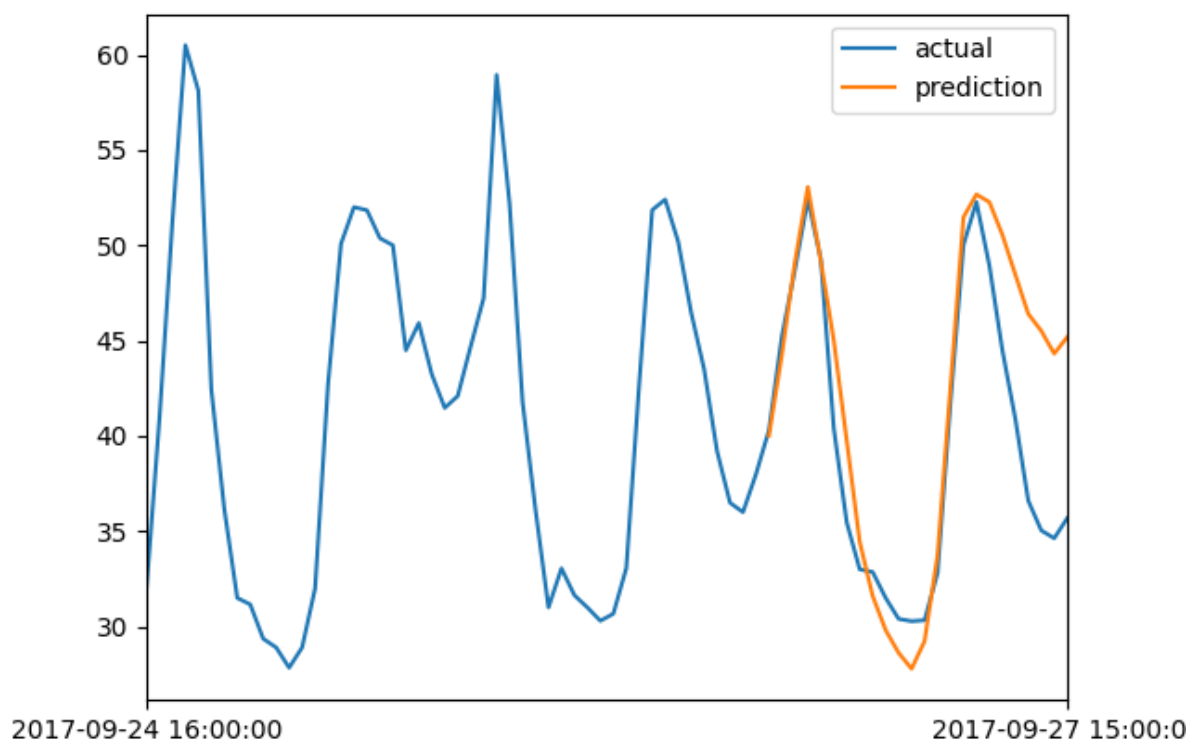
Σχήμα 3.3.4-3 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.



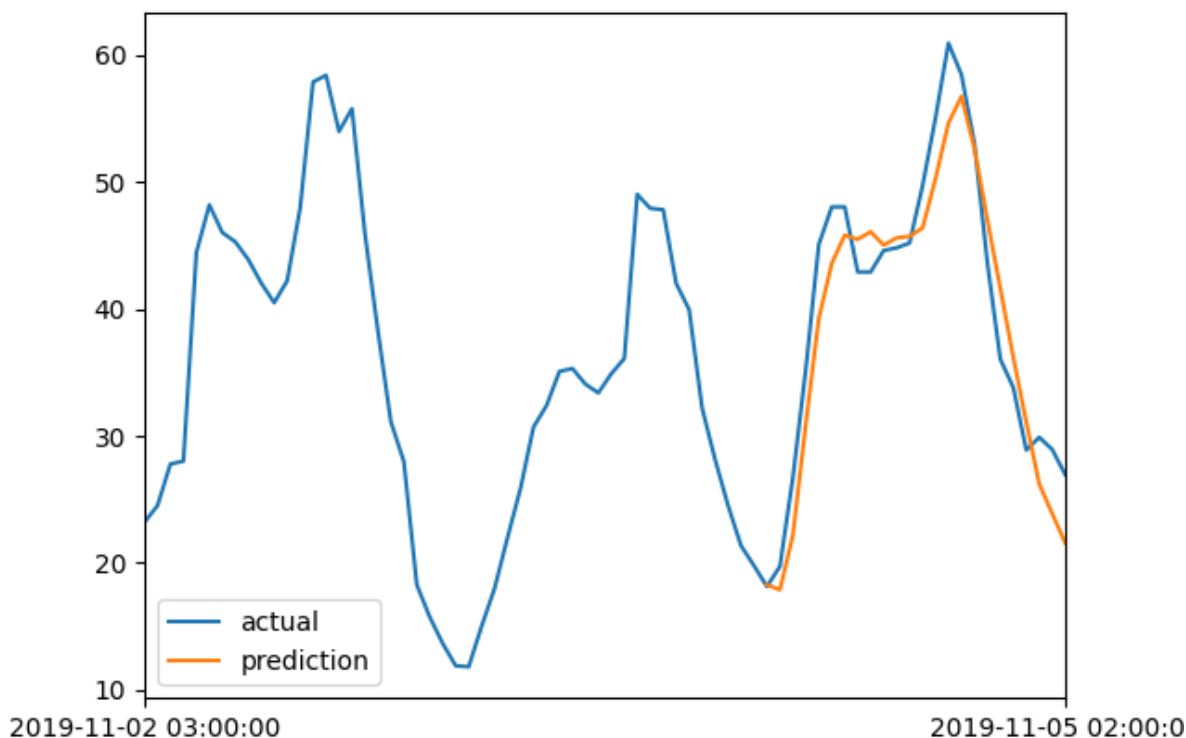
Σχήμα 3.3.4-4 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.



Σχήμα 3.3.4-5 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.



Σχήμα 3.3.4-6 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.



Σχήμα 3.3.4-7 Παράδειγμα προβλέψεων - Πραγματικών Τιμών.

3.4 Σύγκριση αποτελεσμάτων με υπάρχουσα βιβλιογραφία

Για τη σύγκριση των αποτελεσμάτων της παρούσας διπλωματικής εργασίας με εκείνα της μελέτης [7], δίνεται έμφαση σε δύο κύρια σημεία: την ικανότητα του Transformer να εντοπίζει τάσεις και μοτίβα στα δεδομένα και τη σημασία της ενσωμάτωσης των δεδομένων θερμοκρασίας στην πρόβλεψη τιμών ενέργειας.

Στη μελέτη [7], χρησιμοποιήθηκαν μοντέλα: LSTM, CNN και ένας υβριδικός συνδυασμός LSTM-CNN. Από αυτά, το υβριδικό δίκτυο LSTM-CNN παρουσίασε τη μεγαλύτερη ακρίβεια, επιτυγχάνοντας RMSE 15.023, σε σύγκριση με RMSE 15.868 για το LSTM και 18.527 για το CNN-based encoder. Αξιοσημείωτο είναι το γεγονός ότι η προσθήκη μετεωρολογικών δεδομένων, όπως η θερμοκρασία, και δεδομένων χρονικής στιγμής οδήγησε σε χειρότερη ακρίβεια, επιδεινώνοντας τις προβλέψεις με RMSE 17.856. Η προσθήκη αυτών των επιπλέον δεδομένων δεν βελτίωσε τις προβλέψεις, γεγονός που υποδηλώνει ότι τα δίκτυα LSTM και CNN δεν αξιοποιούν πάντα αποτελεσματικά πρόσθετες εξωτερικές μεταβλητές και ενδέχεται να δυσκολεύονται να αποτυπώσουν σχέσεις ή μοτίβα που προέρχονται από διαφορετικές πηγές δεδομένων.

Αντίθετα, στην παρούσα μελέτη, η αρχιτεκτονική Transformer εμφάνισε σημαντική βελτίωση όταν προστέθηκαν δεδομένα θερμοκρασίας στις εισόδους του μοντέλου. Το μοντέλο που χρησιμοποιήθηκε στην παρούσα εργασία σημείωσε σημαντικά καλύτερο

δείκτη RMSE σε σχέση με της μελέτης [7] και στις δύο περιπτώσεις, 9.37 και 10.21 αντίστοιχα κάνοντας χρήση και όχι δεδομένων θερμοκρασίας, αναδεικνύοντας την αρχιτεκτονική που χρησιμοποιήθηκε έναντι αυτές των LSTM και CNN. Συγκεκριμένα, η χρήση της θερμοκρασίας βελτίωσε την ακρίβεια των προβλέψεων, με μείωση στο MAE, MAPE και RMSE σε σχέση με τα αποτελέσματα που προέκυψαν χωρίς τα δεδομένα θερμοκρασίας. Ο Transformer κατόρθωσε να ενσωματώσει τη θερμοκρασία, προσφέροντας έτσι μια βαθύτερη κατανόηση της επίδρασης των καιρικών συνθηκών στη ζήτηση ενέργειας και στις τιμές. Αυτή η διαφορά μπορεί να αποδοθεί στον μηχανισμό self-attention του Transformer, ο οποίος δίνει τη δυνατότητα στο μοντέλο να αναγνωρίζει σχέσεις μεταξύ των μεταβλητών και να εστιάζει σε εκείνες τις παραμέτρους που είναι σημαντικές για την πρόβλεψη κάθε χρονικής στιγμής.

Επενδυτικές Στρατηγικές

4.1 Προσαρμοσμένη Επενδυτική Στρατηγική για Εμπορία Ηλεκτρικής Ενέργειας

Για τις ανάγκες της επενδυτικής στρατηγικής εκπαιδεύτηκε το παραπάνω μοντέλο Transformer σε δεδομένα από το 2020 έως και το 2024 και δοκιμάστηκε σε 375 ημέρες του 2023 και 2024[27]. Η αρχιτεκτονική και οι επενδυτικές στρατηγικές που αναπτύχθηκαν εφαρμόστηκαν στα δεδομένα πέντε Ευρωπαϊκών χωρών και συγκεκριμένα στα δεδομένα της Ουγγαρίας, της Ελλάδας, της Αυστρίας και της Γαλλίας. Αξιοποιώντας τις προηγούμενες 48 ώρες, προβλέπονται οι ωριαίες τιμές ηλεκτρικής ενέργειας και για τις 24 ώρες της εκάστοτε επόμενης ημέρας. Το μοντέλο αξιοποιεί δεδομένα χρονικής στιγμής καθώς και δεδομένα ωριαίας θερμοκρασίας, αντίστοιχα με τα αρχικά πειράματα.

Αναπτύχθηκε στρατηγική η οποία βασίζεται στον αλγοριθμικό εντοπισμό ωρών αιχμής και κοιλάδων στην ημερήσια καμπύλη τιμών ηλεκτρικής ενέργειας με βάση τις προβλέψεις του μοντέλου Transformer σε ημερήσια βάση. Η μπαταρία συνδέεται στο δίκτυο με σκοπό την αγορά και πώληση ενέργειας στην αγορά Επόμενης Ημέρας - Day Ahead Market(DAM). Στόχος είναι η μεγιστοποίηση του κέρδους μέσω αγοράς ενέργειας σε ώρες ελάχιστης τιμής (κοιλάδα) και στη συνέχεια πώληση της ενέργειας σε ώρες μέγιστης τιμής (αιχμή), αξιοποιώντας συστήματα μπαταριών. Ο αλγόριθμος επιλέγει αυτόματα ζεύξεις αγοράς-πώλησης ανά ημέρα, με βάση τη μεγαλύτερη διαφορά τιμών στις ωριαίες προβλέψεις.

Με βάση τη βιβλιογραφία [22],[23] χρησιμοποιούνται οι εξής παραδοχές:

- Κόστος αγοράς μπαταρίας **84 EUR/kWh**
- Κόστος ηλεκτρονικών ισχύος **80 EUR/kW**
- Κόστος ηλεκτρολογικού εξοπλισμού **30 EUR/kW**
- Κόστος εγκατάστασης **30 EUR/kW**
- Ετήσια μείωση χωρητικότητας μπαταρίας **2%**
- Ζωή μπαταρίας **20 έτη**

Για τη είσοδο στο Χρηματιστήριο Ενέργειας [24]:

- Τέλος εισόδου στην αγορά: **18,000 ευρώ.**
- Μηνιαία τέλη μέλους: **1200 ευρώ**
- Τέλη συναλλαγής: **0.07 EUR/MWh**
- Πρόσθετες τεχνικές χρεώσεις: **550 ευρώ/μήνα**

4.2 Sharpe Ratio

Ο δείκτης Sharpe Ratio αποτελεί έναν από τους πλέον ευρέως χρησιμοποιούμενους δείκτες στην ποσοτική αξιολόγηση επενδυτικών στρατηγικών και χαρτοφυλακίων. Εισήχθη από τον William F. Sharpe το 1966 και στοχεύει στην ποσοτικοποίηση της απόδοσης μιας επένδυσης σε σχέση με τον κίνδυνο που αναλαμβάνεται. Συγκεκριμένα, ο δείκτης αυτός υπολογίζεται ως ο λόγος της διαφοράς μεταξύ της αναμενόμενης απόδοσης του χαρτοφυλακίου και του επιτοκίου χωρίς κίνδυνο (risk-free rate), προς την τυπική απόκλιση

(standard deviation) των αποδόσεων του χαρτοφυλακίου, η οποία λειτουργεί ως μέτρο κινδύνου[2]. Η μαθηματική του έκφραση είναι:

$$\text{Sharpe Ratio} = \frac{E(R_p) - R_f}{\sigma_p}$$

όπου $E(R_p)$ είναι η αναμενόμενη απόδοση του χαρτοφυλακίου, R_f το επιτόκιο χωρίς κίνδυνο, και σ_p η τυπική απόκλιση των αποδόσεων του χαρτοφυλακίου.

Ο όρος risk-free rate αναφέρεται στο θεωρητικά ασφαλές επιτόκιο απόδοσης μιας επένδυσης μηδενικού κινδύνου. Στην πράξη, χρησιμοποιούνται συχνά οι αποδόσεις κρατικών ομολόγων υψηλής πιστοληπτικής αξιολόγησης, όπως τα δεκαετή ομόλογα των ΗΠΑ ή της Γερμανίας, ως προσέγγιση για το risk-free rate. Το επιτόκιο αυτό λειτουργεί ως σημείο αναφοράς: αποδόσεις πάνω από αυτό θεωρούνται «ανταμοιβή» για τον αναληφθέντα επενδυτικό κίνδυνο.

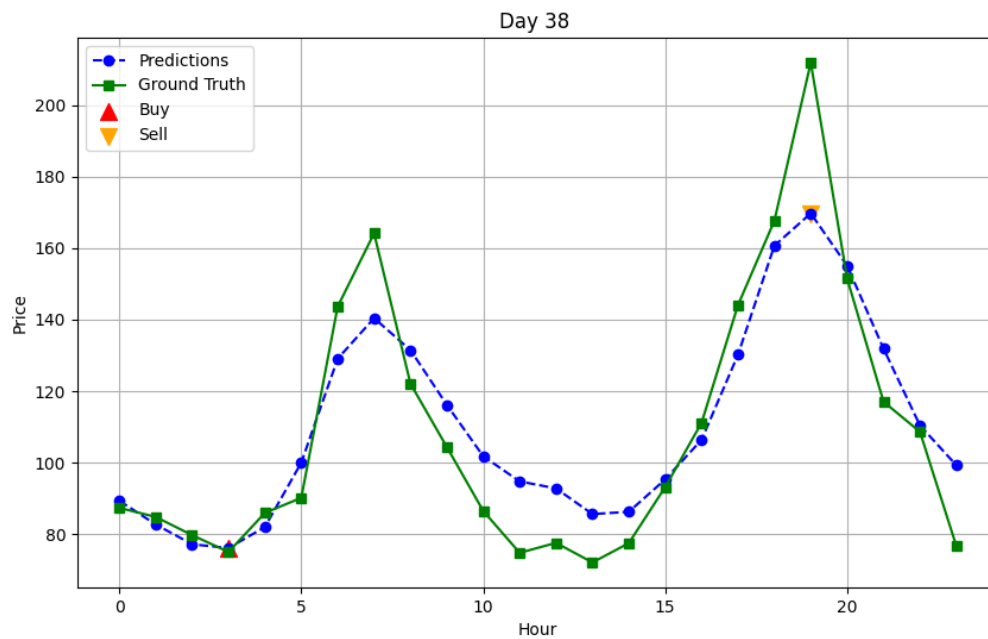
Η αξία του Sharpe Ratio έγκειται στην ικανότητά του να κανονικοποιεί την απόδοση ως προς τον κίνδυνο, καθιστώντας δυνατή τη σύγκριση μεταξύ διαφορετικών επενδυτικών στρατηγικών, ανεξαρτήτως του μεγέθους ή του τύπου τους. Όσο υψηλότερη είναι η τιμή του δείκτη, τόσο καλύτερη θεωρείται η επένδυση, καθώς παράγει μεγαλύτερη «καθαρή» απόδοση ανά μονάδα κινδύνου. Ενδεικτικά, μια τιμή του Sharpe Ratio πάνω από 1 θεωρείται ικανοποιητική, τιμές άνω του 2 υποδηλώνουν εξαιρετική απόδοση σε σχέση με τον κίνδυνο, ενώ τιμές κάτω του 1 ενδέχεται να υποδεικνύουν μη αποδοτική διαχείριση κινδύνου. Σε κάθε περίπτωση, το Sharpe Ratio αποτελεί ουσιώδες εργαλείο για τη λήψη τεκμηριωμένων επενδυτικών αποφάσεων, την αξιολόγηση αλγοριθμικών μοντέλων πρόβλεψης, καθώς και τη βελτιστοποίηση χαρτοφυλακίων με βάση την απόδοση και τη σταθερότητα.

Για τα πειράματα υπολογίζεται ετήσιος δείκτης sharpe ratio θεωρώντας ως χωρητικότητα της μπαταρίας τη μέση χωρητικότητα που θα έχει στη διάρκεια της επένδυσης ώστε τα αποτελέσματα να είναι αντιπροσωπευτικά.

4.3 Σενάριο 1

Στην παρούσα ανάλυση εξετάζεται το επενδυτικό σενάριο που αφορά την προμήθεια συσσωρευτή με ρυθμό φόρτισης και εκφόρτισης ίσο με 1 (1 MW / 1 MWh). Η στρατηγική λειτουργίας βασίζεται στην καθημερινή αγορά ηλεκτρικής ενέργειας κατά τη χρονική στιγμή με τη χαμηλότερη προβλεπόμενη τιμή, σε ποσότητα ίση με τη χωρητικότητα του συσσωρευτή, και στην πώληση του συνόλου της αποθηκευμένης ενέργειας κατά τη χρονική στιγμή με τη μέγιστη προβλεπόμενη τιμή της ίδιας ημέρας. Τα αποτελέσματα που παρουσιάζονται αφορούν τόσο την ακρίβεια των προβλέψεων τιμών, όσο και τους βασικούς οικονομικούς δείκτες, για πέντε ευρωπαϊκές χώρες και για διαφορετικές τιμές χωρητικότητας της μπαταρίας.

Ουγγαρία:

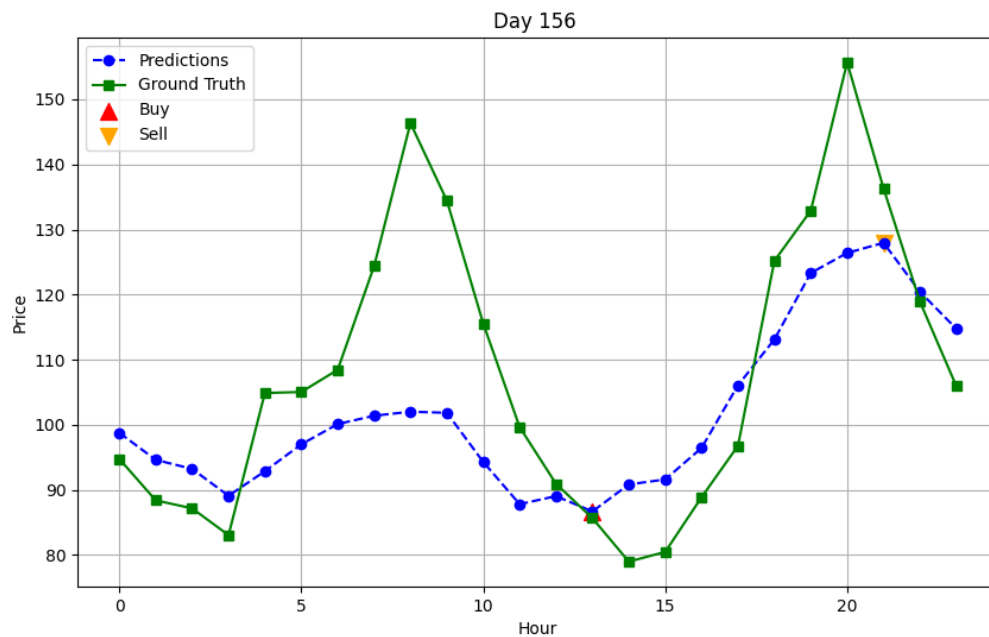


Σχήμα 4.3-1 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 1 στην Ουγγαρία.

Hungary			
MAE	RMSE		
22.15	27.20		
Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.22	0.49	-0.52
2	0.16	0.73	0.17
5	0.64	1.03	0.59
10	0.91	1.19	0.73
20	1.07	1.30	0.80
50	1.18	1.37	0.84
100	1.23	1.39	0.85

Πίνακας 3 Αξιολόγηση Επένδυσης Σεναρίου 1 στην Ουγγαρία.

Ελλάδα:

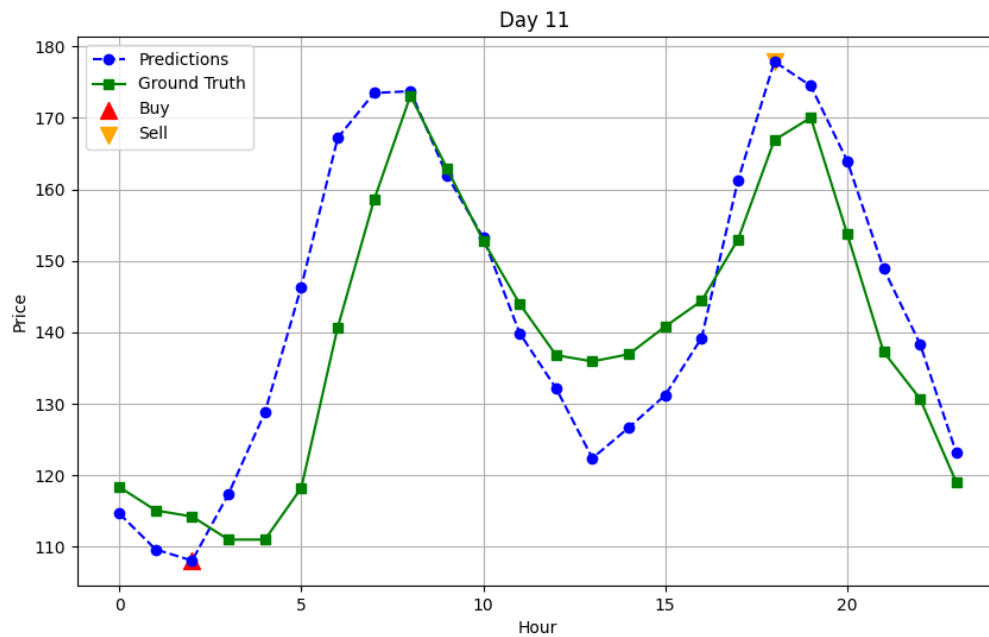


Σχήμα 4.3-2 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 1 στην Ελλάδα.

Greece			
MAE	RMSE		
23.39	29.95		
Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.37	0.44	-0.95
2	-0.07	0.65	-0.17
5	0.31	0.92	0.30
10	0.52	1.06	0.45
20	0.65	1.15	0.53
50	0.74	1.22	0.58
100	0.78	1.24	0.59

Πίνακας 4 Αξιολόγηση Επένδυσης Σεναρίου 1 στην Ελλάδα.

Αυστρία:

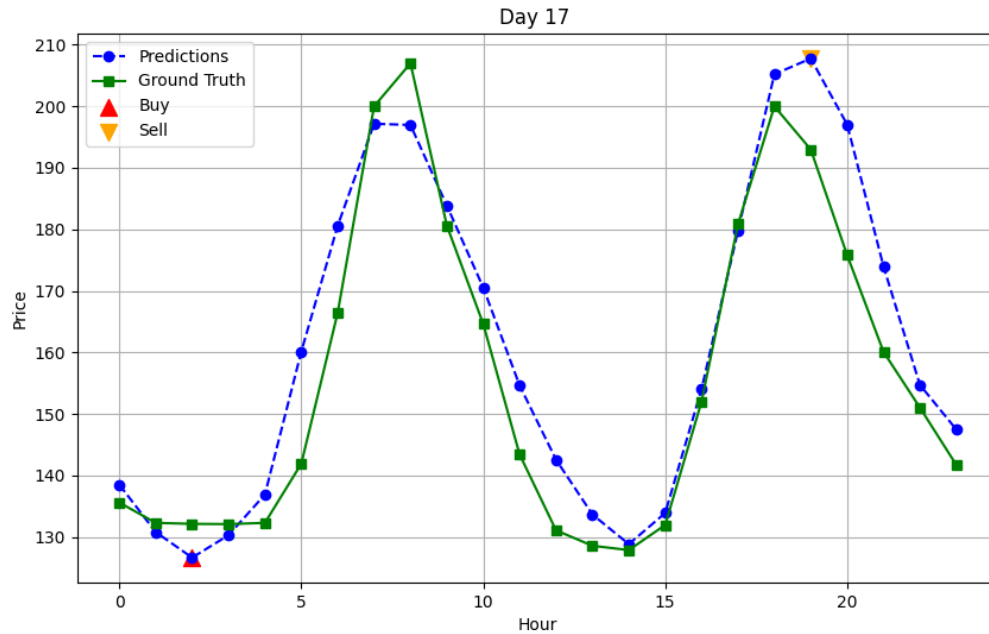


Σχήμα 4.3-3 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 1 στην Αυστρία.

Austria			
MAE	RMSE		
19.03	23.68		
Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.53	0.41	-1.38
2	-0.30	0.61	-0.55
5	-0.01	0.86	-0.05
10	0.15	1.00	0.11
20	0.25	1.09	0.19
50	0.32	1.14	0.24
100	0.34	1.17	0.26

Πίνακας 5 Αξιολόγηση Επένδυσης Σεναρίου 1 στην Αυστρία.

Γαλλία:



Σχήμα 4.3-4 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 1 στη Γαλλία.

France			
MAE	RMSE		
18.90	22.70		
Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.57	0.35	-1.73
2	-0.35	0.52	-0.76
5	-0.09	0.74	-0.17
10	0.06	0.85	0.02
20	0.15	0.93	0.12
50	0.21	0.98	0.18
100	0.23	1.00	0.19

Πίνακας 6 Αξιολόγηση Επένδυσης Σεναρίου 1 στη Γαλλία.

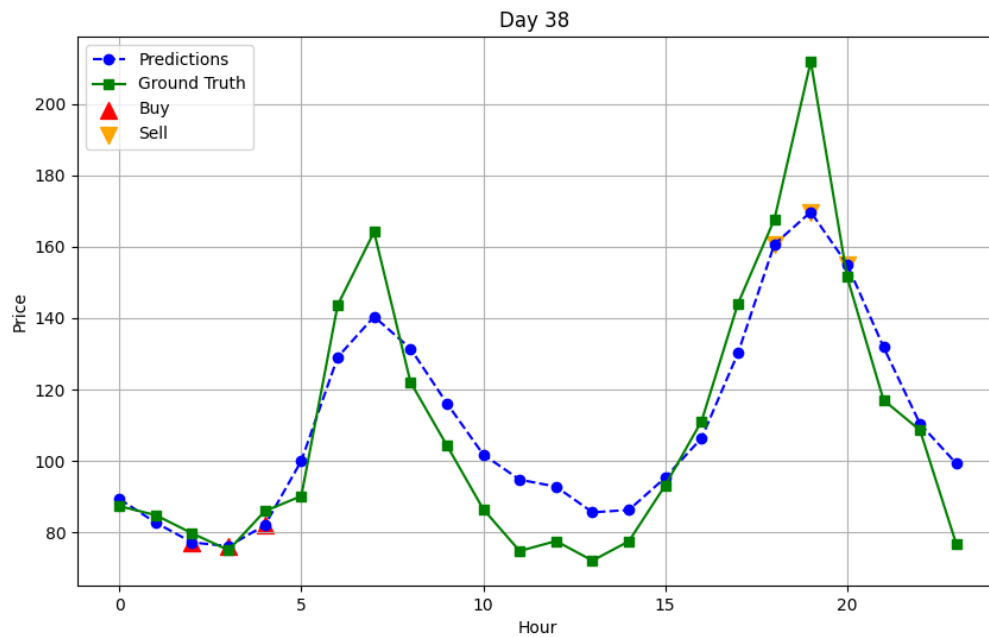
4.4 Σενάριο 2

Λαμβάνοντας υπόψη τις υφιστάμενες οικονομικές παραδοχές, η μείωση της ισχύος της μπαταρίας συνεπάγεται σημαντική μείωση του συνολικού κόστους του συστήματος. Για τον λόγο αυτό, εξετάζεται ένα εναλλακτικό σενάριο στο οποίο γίνεται προμήθεια συσσωρευτή με ρυθμό φόρτισης ίσο με το $\frac{1}{3}$ της χωρητικότητάς του (3 MWh / 1 MW). Η συγκεκριμένη διαστασιολόγηση συνεπάγεται χρόνο πλήρους φόρτισης και εκφόρτισης ίσο με τρεις ώρες.

Στο πλαίσιο αυτού του σεναρίου, η στρατηγική λειτουργίας περιλαμβάνει την αγορά ηλεκτρικής ενέργειας όχι μόνο κατά την ώρα της χαμηλότερης προβλεπόμενης τιμής, αλλά και κατά τις προηγούμενες και επόμενες ώρες αυτής, ενώ αντίστοιχα η πώληση πραγματοποιείται κατά την ώρα της υψηλότερης προβλεπόμενης τιμής και κατά τις γειτονικές ώρες, ώστε να καλυφθεί ο απαιτούμενος χρόνος φόρτισης και εκφόρτισης.

Τα αποτελέσματα που προκύπτουν από την εφαρμογή του συγκεκριμένου σεναρίου παρουσιάζονται στη συνέχεια, με στόχο την αξιολόγηση τόσο της οικονομικής απόδοσης όσο και της αποδοτικότητας της στρατηγικής πρόβλεψης.

Ουγγαρία:

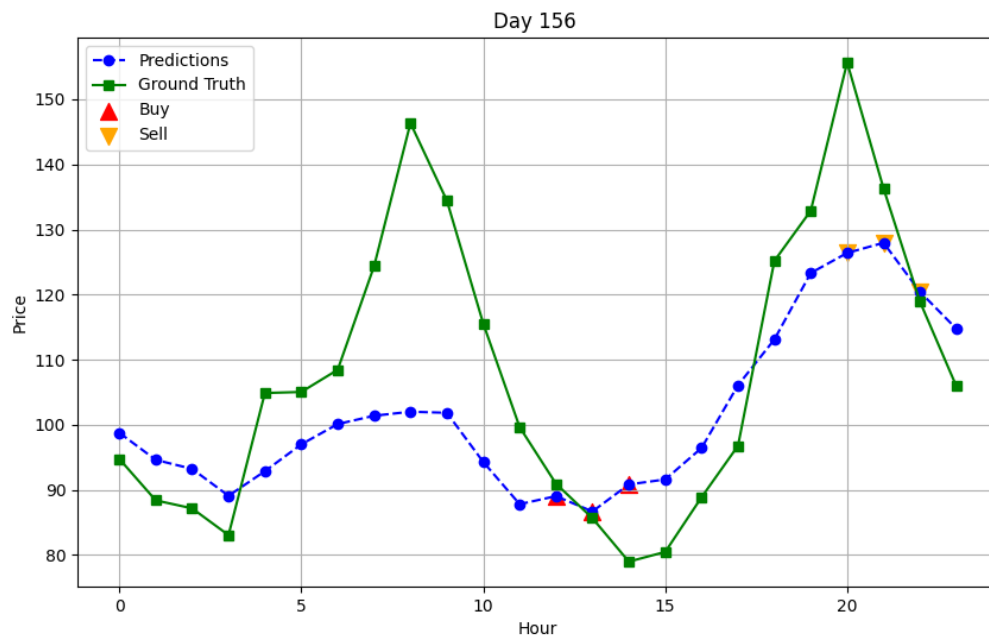


Σχήμα 4.4-1 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 2 στην Ουγγαρία.

Battery Capacity MWh	Return	Standard Deviation	Sharpe Ratio
1	-0.17	0.53	-0.40
2	0.34	0.85	0.35
5	1.13	1.35	0.81
10	1.65	1.68	0.96
20	2.02	1.92	1.03
50	2.30	2.09	1.08
100	2.40	2.16	1.09

Πίνακας 7 Αξιολόγηση Επένδυσης Σεναρίου 2 στην Ουγγαρία.

Ελλάδα:

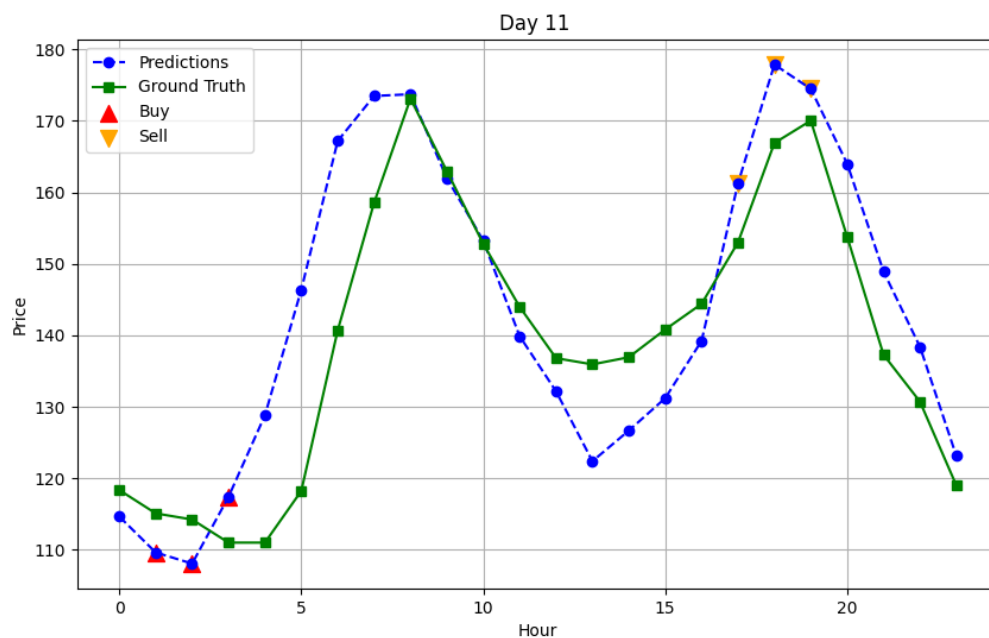


Σχήμα 4.4-2 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 2 στην Ελλάδα.

Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.35	0.44	-0.88
2	0.06	0.71	0.03
5	0.68	1.13	0.57
10	1.10	1.40	0.75
20	1.39	1.60	0.84
50	1.61	1.74	0.90
100	1.69	1.80	0.92

Πίνακας 8 Αξιολόγηση Επένδυσης Σεναρίου 2 στην Ελλάδα.

Αυστρία:

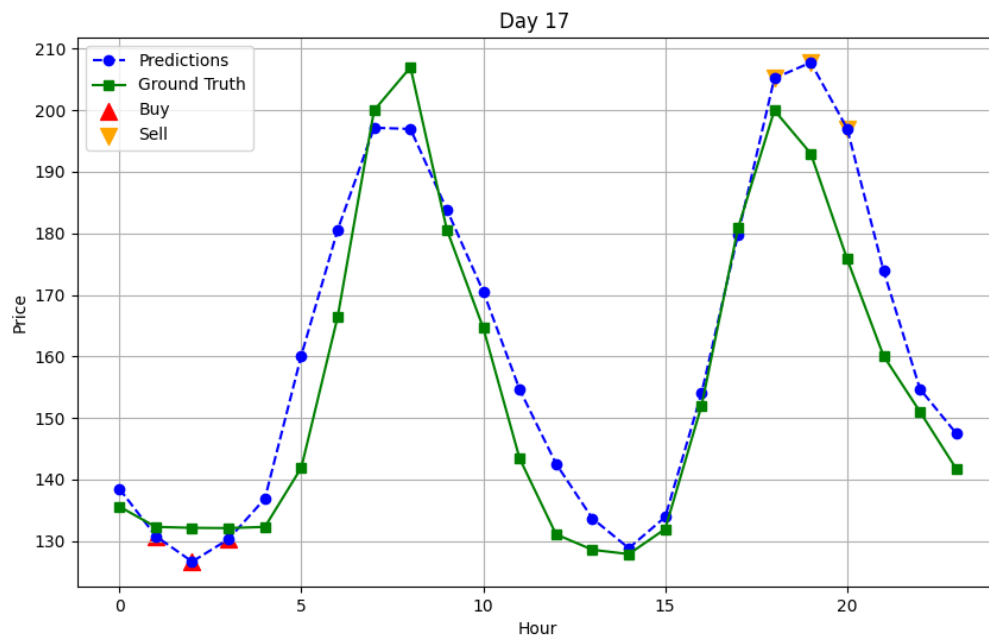


Σχήμα 4.4-3 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 2 στην Αυστρία.

Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.50	0.44	-1.23
2	-0.19	0.71	-0.33
5	0.28	1.14	0.21
10	0.60	1.41	0.39
20	0.82	1.61	0.48
50	0.99	1.76	0.54
100	1.05	1.81	0.56

Πίνακας 9 Αξιολόγηση Επένδυσης Σεναρίου 2 στην Αυστρία.

Γαλλία:



Σχήμα 4.4-4 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 2 στη Γαλλία.

Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.54	0.38	-1.52
2	-0.25	0.61	-0.47
5	0.19	0.98	0.16
10	0.48	1.22	0.37
20	0.69	1.39	0.47
50	0.85	1.51	0.53
100	0.90	1.56	0.55

Πίνακας 10 Αξιολόγηση Επένδυσης Σεναρίου 2 στη Γαλλία.

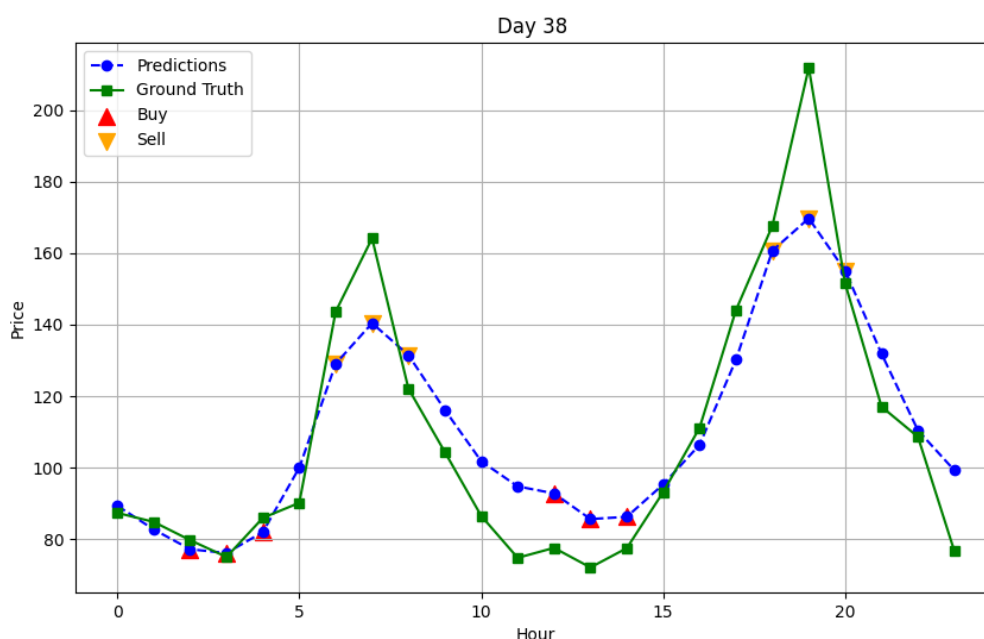
4.5 Σενάριο 3

Διαπιστώθηκε ότι οι ωριαίες καμπύλες τιμών ηλεκτρικής ενέργειας παρουσιάζουν, κατά κανόνα, δύο τοπικά μέγιστα και δύο τοπικά ελάχιστα εντός κάθε ημέρας. Με βάση

αυτήν την παρατήρηση, εξετάζεται ένα τροποποιημένο λειτουργικό σενάριο, στο οποίο η διαδικασία αγοράς και πώλησης ενέργειας πραγματοποιείται δύο φορές ημερησίως.

Λόγω του διπλασιασμού των κύκλων φόρτισης και εκφόρτισης, θεωρείται ότι η ετήσια υποβάθμιση της χωρητικότητας της μπαταρίας αυξάνεται αντιστοίχως κατά δύο φορές. Η μπαταρία που χρησιμοποιείται στο συγκεκριμένο σενάριο είναι ίδια με εκείνη που χρησιμοποιήθηκε στο δεύτερο σενάριο (3 MWh / 1 MW), και τα αποτελέσματα που προκύπτουν από αυτήν τη στρατηγική παρουσιάζονται στη συνέχεια.

Ουγγαρία:

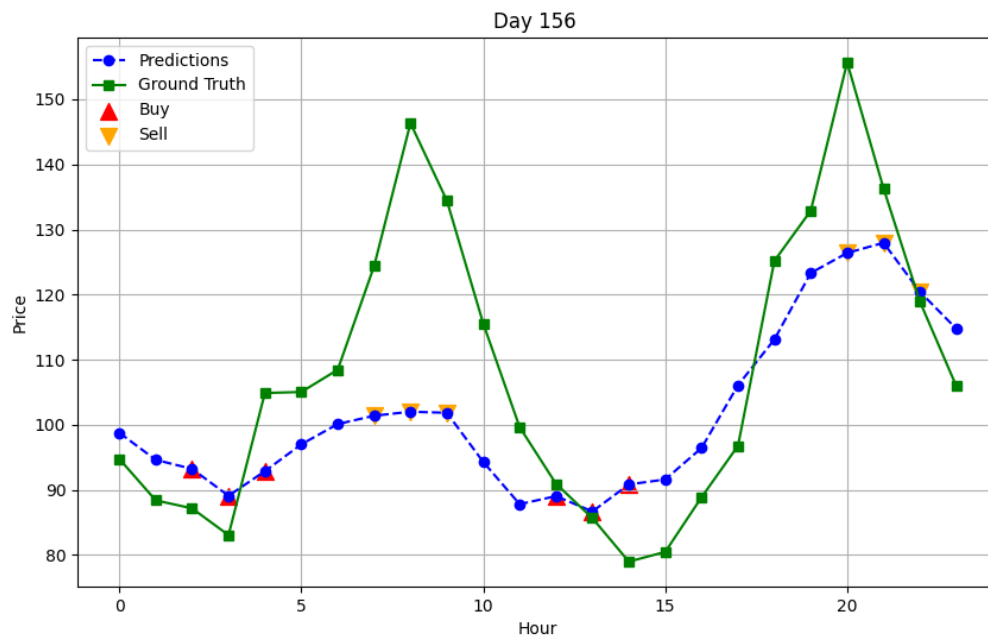


Σχήμα 4.5-1 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 3 στην Ουγγαρία.

Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.09	0.41	-0.31
2	0.49	0.67	0.67
5	1.38	1.07	1.25
10	1.98	1.34	1.45
20	2.41	1.53	1.55
50	2.73	1.68	1.60
100	2.85	1.73	1.62

Πίνακας 11 Αξιολόγηση Επένδυσης Σεναρίου 3 στην Ουγγαρία.

Ελλάδα:

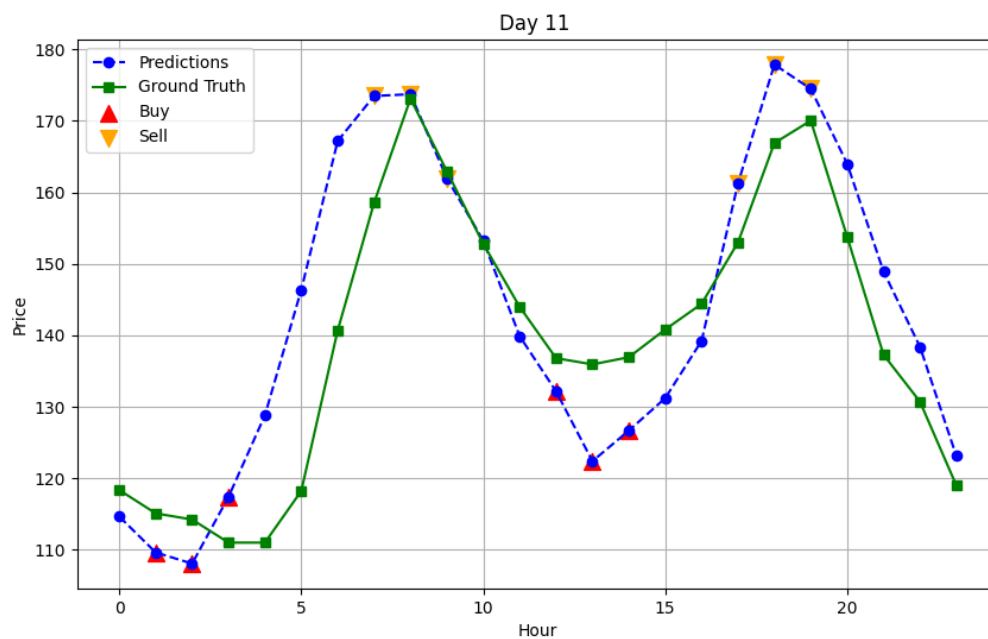


Σχήμα 4.5-2 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 3 στην Ελλάδα.

Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.29	0.43	-0.76
2	0.16	0.70	0.18
5	0.86	1.12	0.74
10	1.33	1.40	0.92
20	1.67	1.60	1.02
50	1.92	1.75	1.07
100	2.02	1.81	1.09

Πίνακας 12 Αξιολόγηση Επένδυσης Σεναρίου 3 στην Ελλάδα.

Αυστρία:

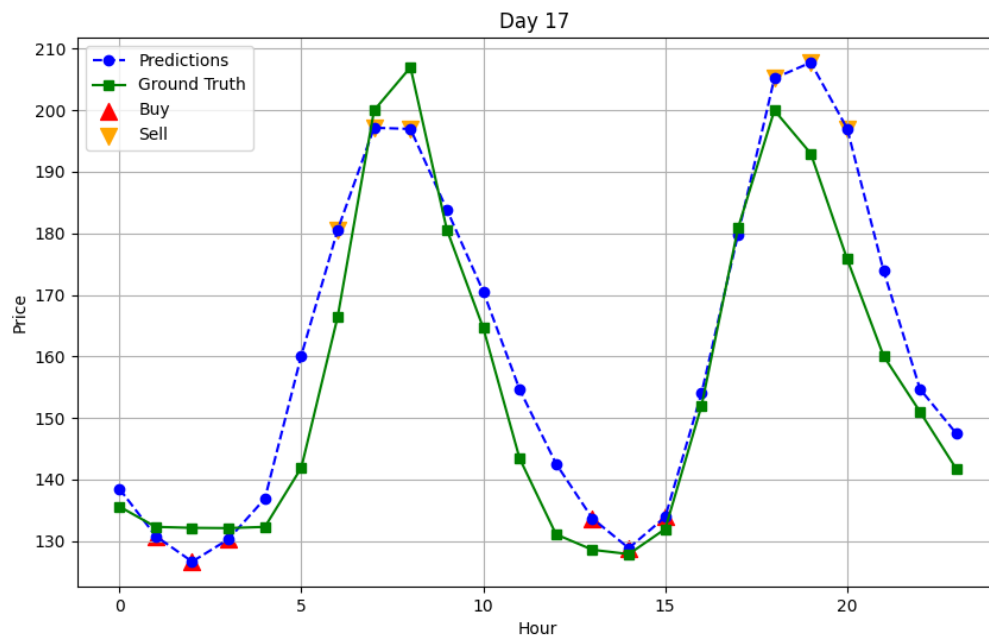


Σχήμα 4.5-3 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 3 στην Αυστρία.

Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.17	0.40	-0.54
2	0.34	0.64	0.46
5	1.12	1.02	1.06
10	1.64	1.27	1.26
20	2.00	1.44	1.36
50	2.28	1.57	1.42
100	2.38	1.62	1.44

Πίνακας 13 Αξιολόγηση Επένδυσης Σεναρίου 3 στην Αυστρία.

Γαλλία:



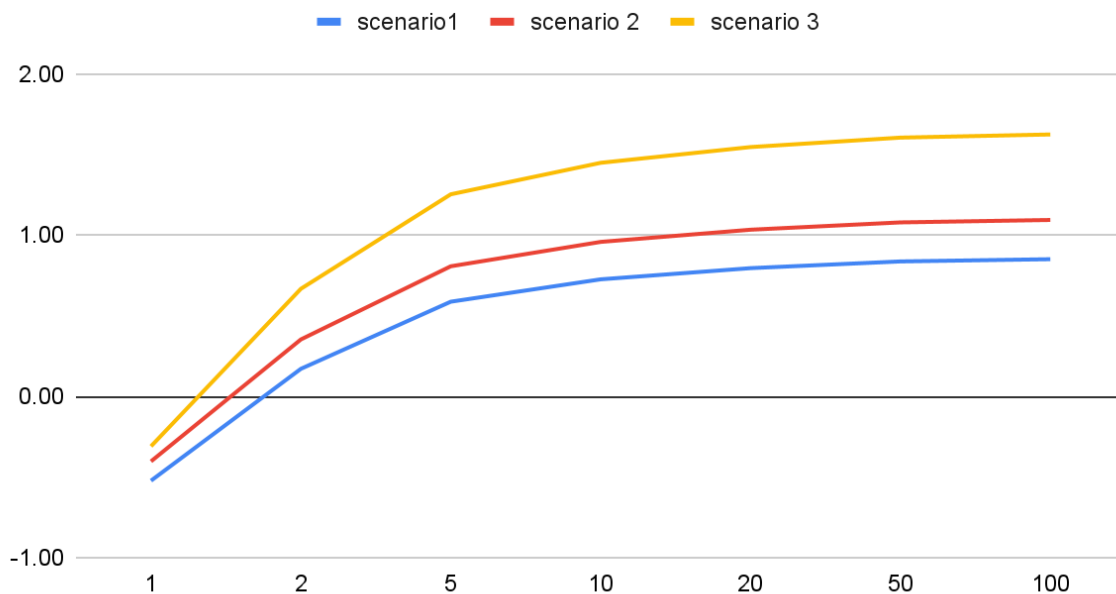
Σχήμα 4.5-4 Ενδεικτική Αναπαράσταση Ωρών - Τιμών Αγοράς και Πώλησης για το Σενάριο 3 στη Γαλλία.

Battery Capacity (MWh)	Return	Standard Deviation	Sharpe Ratio
1	-0.25	0.38	-0.77
2	0.21	0.61	0.27
5	0.92	0.97	0.90
10	1.39	1.21	1.11
20	1.72	1.38	1.22
50	1.96	1.50	1.28
100	2.05	1.55	1.30

Πίνακας 14 Αξιολόγηση Επένδυσης Σεναρίου 3 στη Γαλλία.

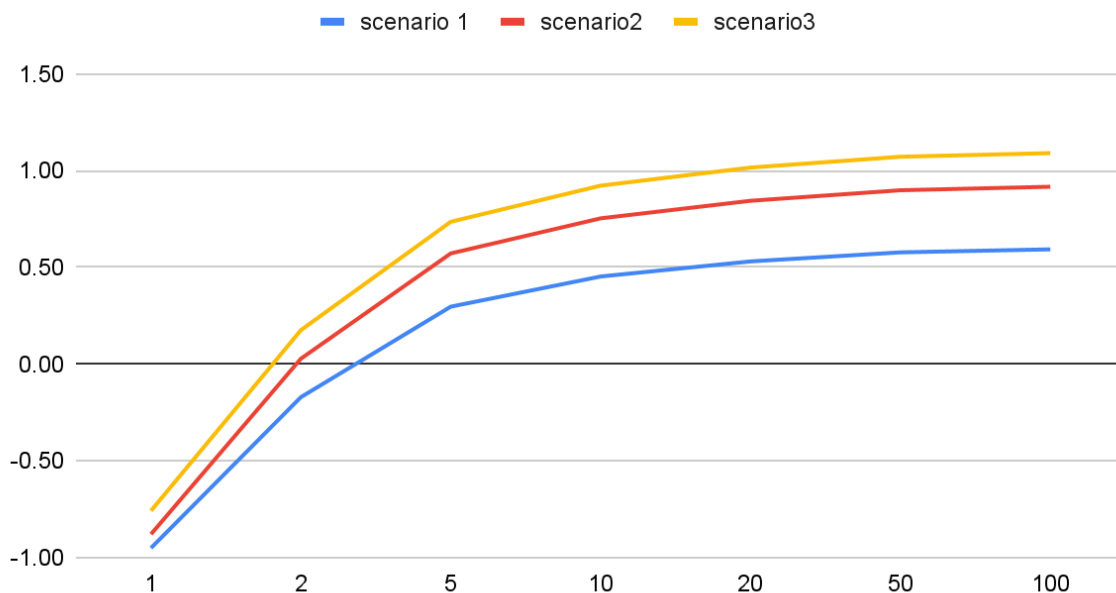
Παρουσιάζονται συγκεντρωτικά διαγράμματα και για τα τρία σενάρια ώστε να γίνει η σύγκριση τους ως προς την οικονομική απόδοση τους:

Hungary Sharpe Ratio/Battery capacity (MWh)



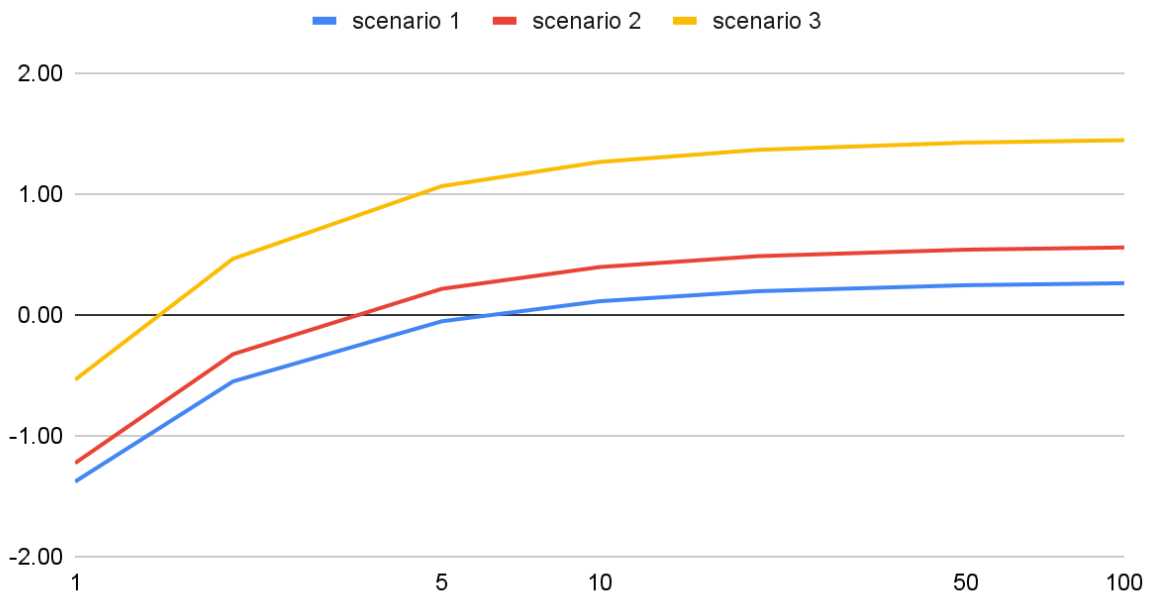
Σχήμα 4.5-5 Σύγκριση Αξιολόγησης της Επένδυσης των τριών Σεναρίων στην Ουγγαρία.

Greece Sharpe Ratio/Battery capacity (MWh)



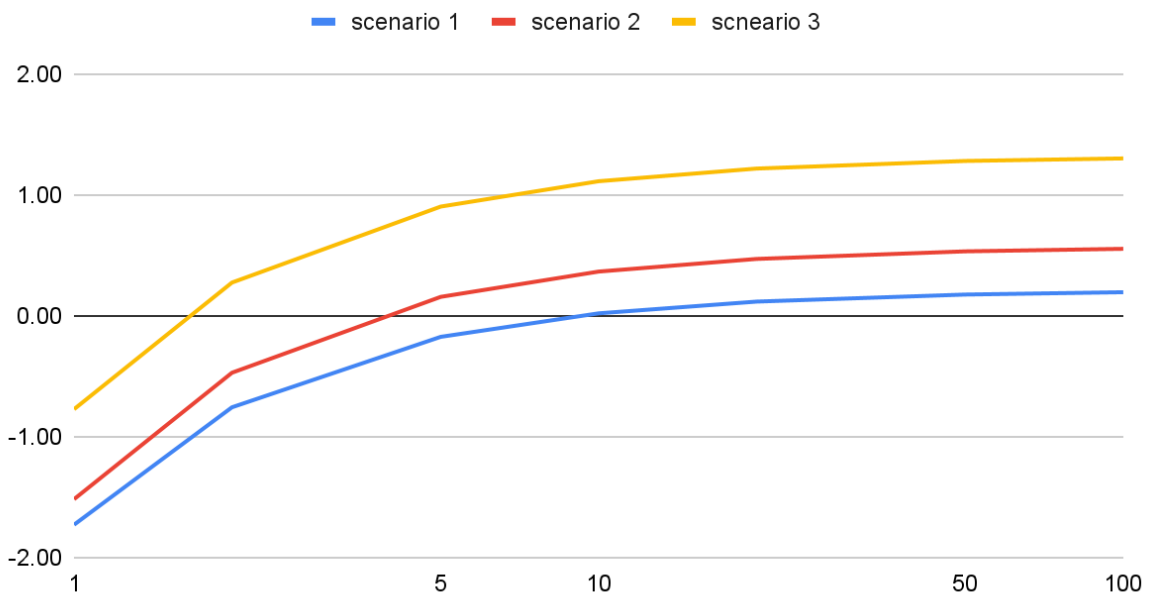
Σχήμα 4.5-6 Σύγκριση Αξιολόγησης της Επένδυσης των τριών Σεναρίων στην Ελλάδα.

Austria Sharpe Ratio/Battery capacity (MWh)



Σχήμα 4.5-7 Σύγκριση Αξιολόγησης της Επένδυσης των τριών Σεναρίων στην Αυστρία.

France Sharpe Ratio/Battery capacity (MWh)



Σχήμα 4.5-8 Σύγκριση Αξιολόγησης της Επένδυσης των τριών Σεναρίων στη Γαλλία.

Όπως προκύπτει από τα αποτελέσματα, για όλες τις υπό μελέτη χώρες παρατηρείται αύξηση του δείκτη Sharpe ratio με την αύξηση της χωρητικότητας της μπαταρίας. Η συμπεριφορά αυτή είναι αναμενόμενη, δεδομένου ότι το υψηλό, αλλά σταθερό, κόστος

πρόσβασης στο δίκτυο καθίσταται αναλογικά μικρότερο όσο αυξάνεται η πωλούμενη ποσότητα ενέργειας, βελτιώνοντας έτσι την αποδοτικότητα της επένδυσης.

Ιδιαίτερο ενδιαφέρον παρουσιάζει, επίσης, η σχετική αύξηση του δείκτη Sharpe ratio μεταξύ των εξεταζόμενων σεναρίων. Το Σενάριο 3 εμφανίζει τη μεγαλύτερη τιμή του δείκτη, υποδεικνύοντας ότι, από τις τρεις στρατηγικές, προσφέρει την υψηλότερη αναλογία αναμενόμενης απόδοσης προς επενδυτικό κίνδυνο, καθιστώντας το την πλέον ασφαλή επενδυτική επιλογή.

4.6 Ανανεώσιμες Πηγές Ενέργειας

Για την περίπτωση της Ελλάδας εξετάζεται ακόμη μία περίπτωση, αυτή της αξιοποίησης της ηλιακής ενέργειας. Δοκιμάζεται η εγκατάσταση φωτοβολταϊκών πάνελ και γίνεται σύγκριση με τα προηγούμενα τρία σενάρια.

Η αξιοποίηση της ηλιακής ενέργειας μέσω φωτοβολταϊκών συστημάτων αποτελεί μια ιδιαίτερα ελκυστική λύση για την Ελλάδα, δεδομένης της υψηλής ηλιοφάνειας που χαρακτηρίζει τη χώρα. Η εγκατάσταση φωτοβολταϊκών παρουσιάζει συγκριτικά χαμηλό λειτουργικό κόστος, ενώ η τεχνολογική πρόοδος έχει βελτιώσει σημαντικά την απόδοσή τους, καθιστώντας την επένδυση πιο αποδοτική σε βάθος χρόνου[8].

Με βάση τις παραδοχές που προκύπτουν από τη βιβλιογραφία, το συνολικό κόστος εγκατάστασης και λειτουργίας ενός φωτοβολταϊκού συστήματος εξαρτάται από την κλίμακα του έργου. Όσο μεγαλύτερη είναι η εγκατάσταση, τόσο μειώνεται το κόστος ανά μονάδα ισχύος, γεγονός που καθιστά τις μεγάλες επενδύσεις πιο συμφέρουσες από τις μικρής κλίμακας εγκαταστάσεις.

Ένα κρίσιμο στοιχείο που πρέπει να ληφθεί υπόψη είναι η σταδιακή μείωση της απόδοσης των φωτοβολταϊκών πάνελ με την πάροδο του χρόνου. Σύμφωνα με τους υπολογισμούς, η μέση μείωση της απόδοσης ανέρχεται σε περίπου 1,2% ετησίως[25]. Αυτό σημαίνει ότι στο τέλος της 20ετούς διάρκειας ζωής τους, η συνολική αποδοτικότητα των πάνελ μπορεί να έχει μειωθεί κατά περίπου 22-25% σε σύγκριση με την αρχική ισχύ. Αυτή η μείωση μπορεί να επηρεάσει το συνολικό οικονομικό όφελος του έργου, γι' αυτό και είναι σημαντική η σωστή διαστασιολόγηση του συστήματος καθώς και η τακτική συντήρηση, ώστε να διασφαλιστεί η μέγιστη δυνατή ενεργειακή παραγωγή καθ' όλη τη διάρκεια ζωής του συστήματος.

Παραδοχές με βάση τη βιβλιογραφία [25]:

- Κόστος φωτοβολταϊκών πάνελ **400 - 600 Ευρώ/kWp**
 - 600 Ευρώ/kWp για εγκατάσταση έως 500 kwp
 - 500 Ευρώ/kWp για εγκατάσταση από 500 έως 1000 kwp
 - 400 Ευρώ/kWp για εγκατάσταση > 1000 kwp
- Inverter **350 Ευρώ/kWp**
- Εγκατάσταση **180 Ευρώ/kWp**
- Λειτουργία και Συντήρηση **3% του συνόλου**
- Διάρκεια ζωής **20 χρόνια**

Η παραγωγή ηλεκτρικής ενέργειας από ένα φωτοβολταϊκό σύστημα[26] υπολογίζεται με βάση τη σχέση:

$$P_v \text{ Production} = (P_v \text{ Area}) \times n \times PR \times (\text{Solar Radiation})$$

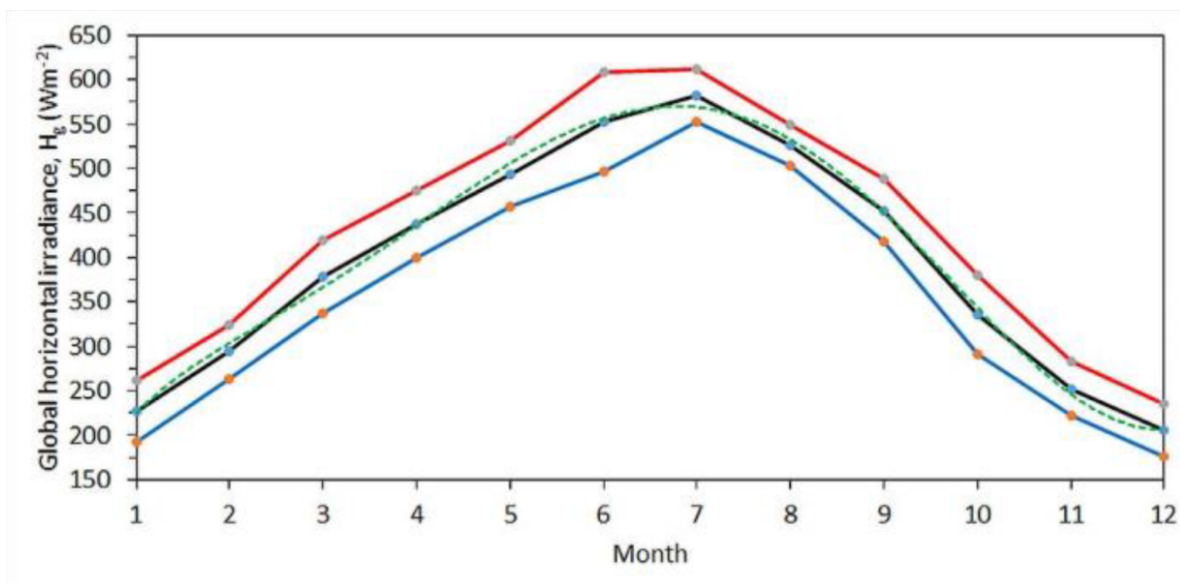
Όπου:

- **P_v Area:** Το συνολικό εμβαδόν των φωτοβολταϊκών πάνελ (m²).
- **n:** Ο βαθμός απόδοσης των φωτοβολταϊκών κυψελών, που εξαρτάται από την τεχνολογία του πάνελ και τυπικά κυμαίνεται μεταξύ 15%-22%.
- **PR (Performance Ratio):** Ο λόγος απόδοσης του συστήματος, ο οποίος λαμβάνει υπόψη απώλειες λόγω θερμοκρασίας, ρύπανσης, σκίασης και λοιπών τεχνικών παραμέτρων. Συνήθως έχει τιμή 0.85.
- **Solar Radiation:** Η συνολική ηλιακή ακτινοβολία που λαμβάνει το πάνελ (kWh/m²), η οποία εξαρτάται από τη γεωγραφική θέση και την εποχή.

Για τα πειράματα θεωρείται **n** = 0.15 και **PR** = 0.85. Ο βαθμός απόδοσης θεωρείται ότι μειώνεται ετησίως με το ρυθμό που αναφέρεται παραπάνω και στα πειράματα χρησιμοποιείται η μέση τιμή του κατά τη διάρκεια της επένδυσης ώστε οι οικονομικοί δείκτες να είναι το δυνατόν αντιπροσωπευτικοί.

Η παραπάνω σχέση επιτρέπει την εκτίμηση της ημερήσιας παραγωγής ενέργειας λαμβάνοντας υπόψη τόσο τα φυσικά χαρακτηριστικά της εγκατάστασης όσο και τις περιβαλλοντικές συνθήκες. Στο πλαίσιο της παρούσας μελέτης, η ηλιακή ακτινοβολία υπολογίζεται με βάση τα δεδομένα ηλιακής ακτινοβολίας που παρουσιάζονται προηγουμένως, ενώ η κατανομή της σε ωριαία βάση επιτυγχάνεται μέσω της ημιτονοειδούς συνάρτησης, επιτρέποντας την ανάλυση της παραγωγής σε επίπεδο Day-Ahead Market.

Για τις ανάγκες της οικονομικής μελέτης, χρησιμοποιήθηκαν τα διαθέσιμα δεδομένα σχετικά με τις ώρες ηλιοφάνειας στην Ελλάδα, τα οποία παρουσιάζονται αναλυτικά σε σχετικό πίνακα. Επιπλέον, αξιοποιήθηκαν δεδομένα από τη βιβλιογραφία σχετικά με τη μέση ηλιακή ακτινοβολία στη χώρα ανα μήνα[9], προκειμένου να υπάρχει μια αξιόπιστη εκτίμηση της διαθέσιμης ηλιακής ενέργειας.



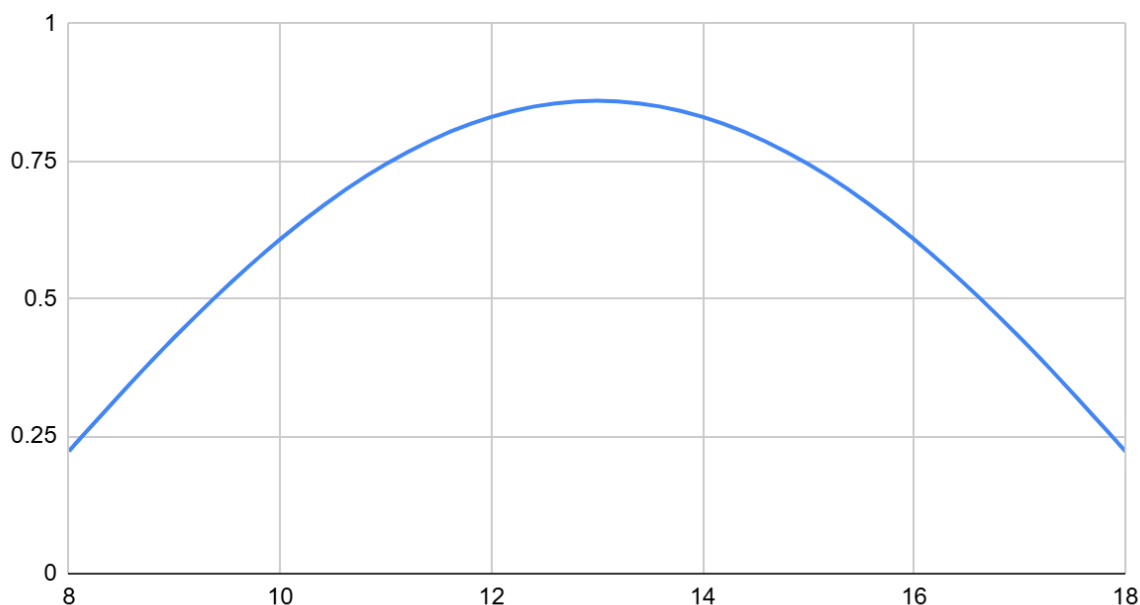
Σχήμα 4.8-1 Μέση Μηνιαία Ηλιακή Ακτινοβολία στην Ελλάδα.

Month	Sunrise	Sunset	Hours of daylight	Solar noon
January	07:38 am	05:30 pm	9:53 h	12:34 pm
February	07:13 am	06:05 pm	10:52 h	12:39 pm
March	06:34 am	06:33 pm	11:59 h	12:33 pm
April	06:47 am	08:02 pm	13:14 h	01:25 pm
May	06:13 am	08:29 pm	14:17 h	01:21 pm
June	06:00 am	08:50 pm	14:50 h	01:25 pm
July	06:13 am	08:48 pm	14:35 h	01:31 pm
August	06:39 am	08:19 pm	13:41 h	01:29 pm
September	07:05 am	07:34 pm	12:29 h	01:20 pm
October	07:32 am	06:49 pm	11:17 h	01:10 pm
November	07:04 am	05:14 pm	10:10 h	12:09 pm
December	07:32 am	05:08 pm	9:36 h	12:20 pm

Σχήμα 4.8-2 Ώρες ηλιοφάνειας ημερησίως ανα μήνα στην Ελλάδα.

Για τη βελτιστοποίηση της ανάλυσης, η ημερήσια κατανομή της ηλιακής ακτινοβολίας προσεγγίζεται μέσω μιας ημιτονοειδούς συνάρτησης, η οποία επιτρέπει την εκτίμηση της έντασης της ηλιακής ακτινοβολίας σε επιμέρους ώρες κάθε ημέρας. Με τον τρόπο αυτό, αποκτούμε μια πιο λεπτομερή εικόνα της διακύμανσης της ηλιακής παραγωγής σε ωριαία βάση, στοιχείο που είναι κρίσιμο για τη διαμόρφωση στρατηγικών στο Day-Ahead Market. Παρουσιάζεται ένα παράδειγμα από την ωριαία παραγωγή ενέργειας κατά τη διάρκεια μίας ημέρας ενός συστήματος με την παραπάνω κατανομή:

Pv Energy (Mwh) / Hour



Σχήμα 4.8-3 Παράδειγμα ωριαίας κατανομής ηλιακής ακτινοβολίας.

Η συγκεκριμένη προσέγγιση επιτρέπει την καλύτερη προσαρμογή του μοντέλου στην πραγματική αγορά ηλεκτρικής ενέργειας, καθώς η πρόβλεψη της ωριαίας παραγωγής των φωτοβολταϊκών είναι καθοριστικός παράγοντας για τη βελτιστοποίηση της οικονομικής αποδοτικότητας και την αποτελεσματική ένταξή τους στο ενεργειακό σύστημα.

4.7 Σενάριο 4

Στο παρόν σενάριο εξετάζεται η εγκατάσταση φωτοβολταϊκών πάνελ, τα οποία συνδέονται απευθείας στο δίκτυο ηλεκτρικής ενέργειας[1],[3]. Η παραγόμενη ενέργεια από το φωτοβολταϊκό σύστημα διοχετεύεται εξ ολοκλήρου και άμεσα στο δίκτυο, χωρίς τη μεσολάβηση μονάδας αποθήκευσης, όπως συσσωρευτής ή άλλη τεχνολογία αποθήκευσης ενέργειας. Συνεπώς, η συνολική παραγόμενη ενέργεια πωλείται τη στιγμή της παραγωγής της, χωρίς δυνατότητα αποθήκευσης για αξιοποίηση σε χρονικές στιγμές υψηλότερης τιμής.

PV				
To the grid				
Pv Area (m2)	Return	Stdev	Sharpe Ratio	Mwp
7844.00	1.00	1.20	0.80	1
15688.00	1.38	1.43	0.94	2
19610.00	1.47	1.48	0.97	2.5
39220.00	1.68	1.61	1.02	5
62752.00	1.77	1.66	1.04	8
78440.00	1.80	1.68	1.05	10
392200.00	1.90	1.74	1.07	50
784400.00	1.91	1.75	1.07	100

Πίνακας 15 Αξιολόγηση Επένδυσης Σεναρίου 4.

4.8 Σενάριο 5

Στη συνέχεια, μελετάται η ενσωμάτωση συστήματος αποθήκευσης ενέργειας, χρησιμοποιώντας μπαταρία με χαρακτηριστικά αντίστοιχα του Σεναρίου 3, για τη βελτιστοποίηση της διαχείρισης της παραγόμενης ενέργειας από το φωτοβολταϊκό σύστημα. Η στρατηγική που ακολουθείται περιλαμβάνει την αποθήκευση της παραγόμενης ενέργειας και τη στοχευμένη πώλησή της κατά τις ώρες αιχμής κάθε ημέρας, όταν οι τιμές ενέργειας είναι υψηλότερες.

Στο πλαίσιο της ανάλυσης, δοκιμάζονται διάφοροι συνδυασμοί φωτοβολταϊκών εγκαταστάσεων και χωρητικότητας μπαταρίας, με στόχο την εύρεση του βέλτιστου λόγου μεταξύ αυτών. Από τα αποτελέσματα προκύπτει ότι η βέλτιστη αναλογία προκύπτει όταν ο λόγος (Χωρητικότητα Μπαταρίας) / (Εγκατεστημένη Ισχύς Φωτοβολταϊκών, σε MWp) είναι ίσος με 4, γεγονός που οδηγεί στην εξής απόδοση του συστήματος:

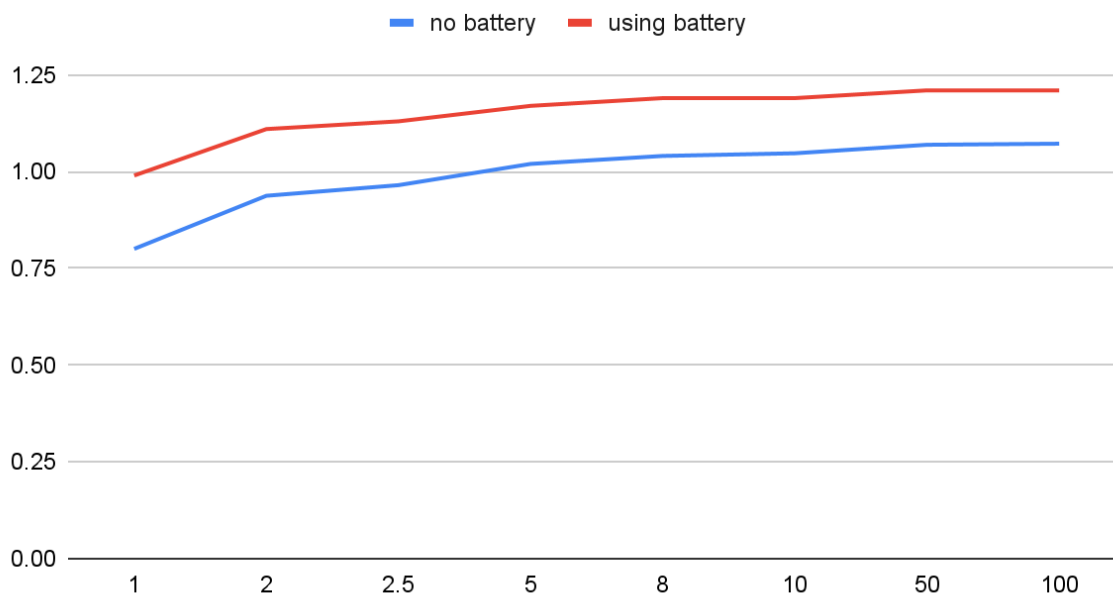
Using battery						
Battery Capacity (MWh)	PV Area (m2)	Average Return Percent	Annualized Standard Deviation	Sharpe Ratio	MWp	ratio
1	1961.00	0.18	0.6	0.23	0.25	4
2	3922.00	0.65	0.84	0.72	0.5	4
4	7844.00	1.12	1.08	0.99	1	4
5	9805.00	1.22	1.13	1.04	1.25	4
8	15688.00	1.39	1.22	1.11	2	4
10	19610.00	1.46	1.26	1.13	2.5	4
20	39220.00	1.59	1.33	1.17	5	4
32	62752.00	1.65	1.36	1.19	8	4
40	78440.00	1.67	1.37	1.19	10	4
50	98050.00	1.69	1.37	1.2	12.5	4
80	156880.00	1.71	1.39	1.2	20	4
100	196100.00	1.72	1.39	1.21	25	4
200	392200.00	1.73	1.4	1.21	50	4
400	784400.00	1.74	1.4	1.21	100	4
800	1568800.00	1.75	1.4	1.21	200	4

Πίνακας 16 Αξιολόγηση Επένδυσης Σεναρίου 5.

Στο παρόν στάδιο, παρουσιάζονται διαγράμματα τα οποία αποτυπώνουν την οικονομική απόδοση του σεναρίου εγκατάστασης φωτοβολταϊκών πάνελ, τόσο με όσο και χωρίς τη χρήση συστήματος αποθήκευσης ενέργειας (μπαταρίας). Η εν λόγω απεικόνιση επιτρέπει την αποτίμηση της οικονομικής σκοπιμότητας ενσωμάτωσης συστήματος αποθήκευσης, αναδεικνύοντας τη δυνητική συνεισφορά του στη βελτίωση της χρηματοοικονομικής απόδοσης της συνολικής επένδυσης.

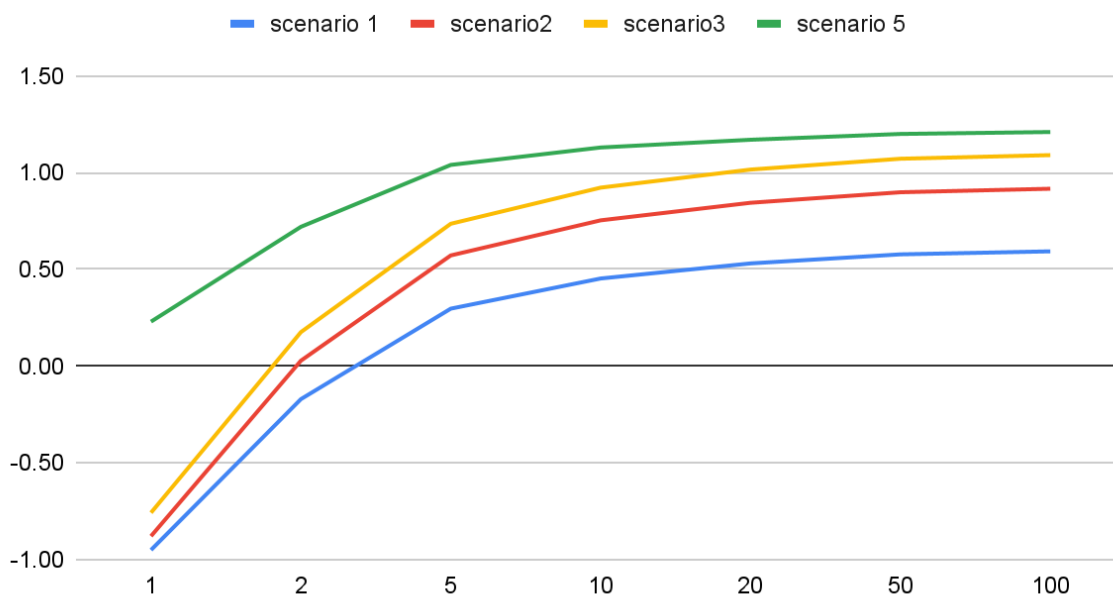
Παράλληλα, παρουσιάζεται συγκριτική οικονομική ανάλυση όλων των σεναρίων στα οποία γίνεται χρήση μπαταρίας, με στόχο την εξαγωγή συμπερασμάτων σχετικά με το αν η ιδιοπαραγωγή ενέργειας μέσω φωτοβολταϊκών με αποθήκευση και στοχευμένη μεταγενέστερη πώληση αποδεικνύεται αποδοτικότερη στρατηγική έναντι της αγοράς ηλεκτρικής ενέργειας από το δίκτυο με σκοπό την εκμετάλλευση διαφορών τιμής (arbitrage).

Sharpe Ratio / MWp



Σχήμα 4.9-1 Σύγκριση Επένδυσης με και χωρίς μπαταρία.

Greece Sharpe Ratio/Battery capacity (MWh)



Σχήμα 4.9-2 Σύγκριση των τεσσάρων σεναρίων που χρησιμοποιούν μπαταρία στην Ελλάδα.

4.9 Αποτελέσματα

Όπως φαίνεται στο Σχήμα 4.9-1, η ενσωμάτωση συστήματος αποθήκευσης ενέργειας μέσω μπαταρίας οδηγεί σε μείωση της απόδοσης της επένδυσης (Return on Investment), κυρίως λόγω του αυξημένου κόστους που συνεπάγεται η εγκατάσταση και λειτουργία της μπαταρίας. Ωστόσο, η χρήση αποθήκευσης συμβάλλει σημαντικά στη μείωση του επενδυτικού κινδύνου, γεγονός που αποτυπώνεται στην αύξηση του δείκτη Sharpe ratio.

Η κόκκινη καμπύλη, η οποία αντιστοιχεί στο Σενάριο 5 (συνδυασμός φωτοβολταϊκών με αποθήκευση και στρατηγική πώλησης βάσει προβλέψεων τιμών), εμφανίζει σαφώς υψηλότερες τιμές Sharpe ratio σε όλο το εύρος των εξεταζόμενων εγκατεστημένων ισχύων. Αντιθέτως, η μπλε καμπύλη, η οποία αναφέρεται στο Σενάριο 4 (άμεση πώληση της παραγόμενης ενέργειας χωρίς αποθήκευση), παρουσιάζει αισθητά χαμηλότερη επενδυτική αποδοτικότητα.

Η υπεροχή του Σχεδίου 5 είναι ιδιαίτερα εμφανής σε μικρές και μεσαίες εγκαταστάσεις. Σε πολύ μεγάλες εγκαταστάσεις (άνω των 50 MWp), παρατηρείται τάση σύγκλισης των δύο καμπυλών, καθώς το σταθερό κόστος συμμετοχής στην αγορά ενέργειας καθίσταται αμελητέο σε σχέση με το συνολικό μέγεθος της επένδυσης.

Συνολικά, η προσθήκη συστήματος αποθήκευσης σε συνδυασμό με φωτοβολταϊκή παραγωγή ενισχύει ουσιαστικά τη σταθερότητα και την αξιοπιστία της επένδυσης, καθιστώντας το Σενάριο 5 επενδυτικά πιο ελκυστικό σε σύγκριση με το Σενάριο 4.

Όπως απεικονίζεται στο Σχήμα 4.9-2, η αντιστοίχιση των χωρητικότητων των μπαταριών στα διάφορα σενάρια επενδύσεων υποδεικνύει ότι η στρατηγική κατά την οποία η πωλούμενη ενέργεια στο δίκτυο προέρχεται απευθείας από φωτοβολταϊκά πάνελ (Σενάριο 5) θεωρείται πιο ασφαλής με βάση τον δείκτη Sharpe ratio, σε σύγκριση με το σενάριο όπου χρησιμοποιείται αντίστοιχη μπαταρία για την αγορά ενέργειας από το δίκτυο, η οποία στη συνέχεια πωλείται σε υψηλότερη τιμή.

Σενάριο 1: Η απλή καθημερινή στρατηγική αγοράς στην ώρα με τη χαμηλότερη προβλεπόμενη τιμή και πώλησης στην ώρα με τη μέγιστη τιμή δεν αποδίδει ικανοποιητικά για μικρές χωρητικότητες μπαταρίας, με τον δείκτη Sharpe ratio να λαμβάνει αρνητικές τιμές κάτω από 2 MWh. Η στρατηγική βελτιώνεται με την αύξηση της χωρητικότητας, ωστόσο παραμένει λιγότερο αποδοτική συγκριτικά με τις άλλες προσεγγίσεις.

Σενάριο 2: Η μείωση του κόστους λόγω της χαμηλότερης ισχύος φόρτισης, σε συνδυασμό με την αξιοποίηση περιόδου φόρτισης-εκφόρτισης τριών ωρών, οδηγεί σε σημαντική βελτίωση της απόδοσης και θετικό Sharpe ratio από τις πρώτες μονάδες χωρητικότητας.

Σενάριο 3: Με διπλή ημερήσια πράξη αγοράς και πώλησης και εντατικότερη χρήση

της μπαταρίας, το σενάριο παρουσιάζει ακόμη υψηλότερη απόδοση, παρά τη γρηγορότερη υποβάθμιση της χωρητικότητας που σχετίζεται με την αυξημένη χρήση (degradation).

Σενάριο 5: Αποτελεί σαφώς την πλέον αποδοτική στρατηγική, καθώς η παραγόμενη ενέργεια προέρχεται από το ίδιο φωτοβολταϊκό σύστημα και αποθηκεύεται για πώληση κατά τις βέλτιστες χρονικές στιγμές. Το σενάριο επιτυγχάνει δείκτη Sharpe ratio άνω του 1,2 για υψηλές χωρητικότητες μπαταρίας και διατηρεί υπεροχή σε όλο το φάσμα των χωρητικοτήτων που εξετάστηκαν.

Συμπεράσματα

Η παρούσα διπλωματική εργασία ανέδειξε τη σημαντική συμβολή της τεχνητής νοημοσύνης στην αποτελεσματική διαχείριση και εμπορία ηλεκτρικής ενέργειας, μέσω της αξιοποίησης προγνωστικών μοντέλων αιχμής και της τεχνοοικονομικής ανάλυσης στρατηγικών αποθήκευσης. Συγκεκριμένα, αναπτύχθηκε και εκπαιδεύτηκε ένα μοντέλο βασισμένο στην αρχιτεκτονική Transformer, το οποίο κατέδειξε σημαντικά ανώτερη ακρίβεια στην πρόβλεψη των ωριαίων τιμών ενέργειας σε σύγκριση με παραδοσιακές προσεγγίσεις, όπως τα LSTM και τα υβριδικά LSTM-CNN. Επιπλέον, το μοντέλο Transformer επέδειξε μεγαλύτερη ικανότητα στην ενσωμάτωση εξωγενών μεταβλητών, όπως οι θερμοκρασιακές παράμετροι, ενισχύοντας περαιτέρω την προβλεπτική του απόδοση, σε αντίθεση με τα συγκριτικά μοντέλα που δεν κατάφεραν να εκμεταλλευτούν με ανάλογη αποτελεσματικότητα αυτές τις πληροφορίες.

Οι προβλέψεις αυτές αξιοποιήθηκαν σε πέντε διαφορετικά επενδυτικά σενάρια, με στόχο την εκτίμηση της οικονομικής βιωσιμότητας των συστημάτων αποθήκευσης και παραγωγής ενέργειας. Μέσω αυτής της μεθοδολογίας, αναδείχθηκε ότι η χρήση ακριβών προβλέψεων δύναται να μεγιστοποιήσει το χρηματοοικονομικό όφελος που προκύπτει από στρατηγικές arbitrage στην αγορά ενέργειας, ιδιαίτερα όταν αυτές συνδυάζονται με τεχνολογίες αποθήκευσης και ιδιοπαραγωγής μέσω φωτοβολταϊκών συστημάτων. Ιδιαίτερα, το σενάριο που συνδυάζει φωτοβολταϊκή παραγωγή με αποθήκευση και δυναμική πώληση βασισμένη σε προβλέψεις (Σενάριο 5) κατέδειξε την υψηλότερη απόδοση, επιτυγχάνοντας τον μεγαλύτερο λόγο απόδοσης προς κίνδυνο (Sharpe ratio) σε όλα τα εξεταζόμενα μεγέθη και περιπτώσεις.

Η συνολική μεθοδολογία που αναπτύχθηκε αναδεικνύει την πολυδιάστατη συμβολή της τεχνητής νοημοσύνης, η οποία εκτείνεται από τη βελτιστοποίηση της πρόβλεψης μέχρι τη στοχευμένη αξιοποίηση ενεργειακών συστημάτων σε ένα δυναμικό και συνεχώς μεταβαλλόμενο περιβάλλον. Η παρούσα εργασία υπογραμμίζει τον ρόλο των εξελιγμένων μοντέλων τεχνητής νοημοσύνης ως ουσιαστικών εργαλείων στρατηγικής λήψης αποφάσεων, ενισχύοντας τη βιωσιμότητα και την οικονομική απόδοση επενδύσεων στον τομέα των ανανεώσιμων πηγών ενέργειας και της αποθήκευσης.

Ως μελλοντική προοπτική, προτείνεται η δοκιμή και αξιολόγηση νεότερων αλγορίθμων τεχνητής νοημοσύνης, λαμβάνοντας υπόψη την ταχεία εξέλιξη του τομέα. Επιπλέον, κρίνεται αναγκαία η επικαιροποίηση και διεύρυνση της τεχνοοικονομικής ανάλυσης με βάση τις πρόσφατες εξελίξεις της αγοράς, ιδιαίτερα όσον αφορά τη σημαντική μείωση του κόστους αποθήκευσης και την εμφάνιση νέων τεχνολογιών μπαταριών. Δεδομένης της ταχύτατης προόδου στον ενεργειακό τομέα, μια πιο δυναμική και διαχρονικά ευέλικτη προσέγγιση αξιολόγησης δύναται να προσφέρει μια ακριβέστερη και πιο ρεαλιστική εικόνα της πραγματικής αποδοτικότητας τέτοιων επενδύσεων.

Βιβλιογραφία

- [1] Abo-Khalil, A. G., Sayed, K., Radwan, A., & El-Sharkawy, Ibrahim I. A. (2023). Analysis of the PV system sizing and economic feasibility study in a grid-connected PV system. *Case Studies in Thermal Engineering*, 45, 102903. <https://doi.org/10.1016/j.csite.2023.102903>
- [2] Atmaca, M. E. (2022). Portfolio management and performance improvement with Sharpe and Treynor ratios in electricity markets. *Energy Reports*, 8, 192–201. <https://doi.org/10.1016/j.egy.2021.11.287>
- [3] Atsu, D., Seres, I., & Farkas, I. (2021). The state of solar PV and performance analysis of different PV technologies grid-connected installations in Hungary. *Renewable and Sustainable Energy Reviews*, 141, 110808. <https://doi.org/10.1016/j.rser.2021.110808>
- [4] Beigaite, R., Krilavicius, T., & Man, K. L. (2018). Electricity Price Forecasting for Nord Pool Data. *2018 International Conference on Platform Technology and Service (PlatCon)*, 1–6. <https://doi.org/10.1109/PlatCon.2018.8472762>
- [5] Sekhar, C., & Meghana, P. S. (2020). A Study on Backpropagation in Artificial Neural Networks. *Asia-Pacific Journal of Neural Networks and Its Applications*, 4(1), 21–28. <https://doi.org/10.21742/AJNNIA.2020.4.1.03>
- [6] Chowdary, K. L., Krishna, C. N., Manaswini, K. S., & Jithendra, B. (2023). Electricity Price Prediction using Machine Learning. *Proceedings of the 3rd International Conference on Artificial Intelligence and Smart Energy, ICAIS 2023*. <https://doi.org/10.1109/ICAIS56108.2023.10073777>
- [7] Dombi, G., & Dulai, T. (2022). Hourly electricity price forecast for short-and long-term, using deep neural networks. *Acta Universitatis Sapientiae, Informatica*, 14(2), 208–222. <https://doi.org/10.2478/ausi-2022-0013>
- [8] Fantidis, J. G., Bandekas, D. V., Potolias, C., & Vordos, N. (2013). Cost of PV electricity – Case study of Greece. *Solar Energy*, 91, 120–130. <https://doi.org/10.1016/j.solener.2013.02.001>
- [9] Kambezidis, H. D. (2021). The Solar Radiation Climate of Greece. *Climate*, 9(12), 183. <https://doi.org/10.3390/cli9120183>
- [10] Kaya, M., Karan, M. B., & Telatar, E. (2023). Electricity price estimation using deep learning approaches: An empirical study on Turkish markets in normal and Covid-19 periods. *Expert Systems with Applications*, 224, 120026. <https://doi.org/10.1016/j.eswa.2023.120026>

- [11] Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4), 1030–1081.
<https://doi.org/10.1016/j.ijforecast.2014.08.008>
- [12] Li, W., & Becker, D. M. (2021). Day-ahead electricity price prediction applying hybrid models of LSTM-based deep learning methods and feature selection algorithms under consideration of market coupling. *Energy*, 237, 121543.
<https://doi.org/10.1016/j.energy.2021.121543>
- [13] Lu, H., Ma, X., Ma, M., & Zhu, S. (2021). Energy price prediction using data-driven models: A decade review. *Computer Science Review*, 39, 100356.
<https://doi.org/10.1016/j.cosrev.2020.100356>
- [14] Matos, M., Almeida, J., Gonçalves, P., Baldo, F., Braz, F. J., & Bartolomeu, P. C. (2024). A Machine Learning-Based Electricity Consumption Forecast and Management System for Renewable Energy Communities. *Energies*, 17(3), 630.
<https://doi.org/10.3390/en17030630>
- [15] Mehtab, S., & Sen, J. (2020). *Stock Price Prediction Using CNN and LSTM-Based Deep Learning Models*. <https://doi.org/10.1109/DASA51403.2020.9317207>
- [16] Poggi, A., di Persio, L., & Ehrhardt, M. (2023). Electricity Price Forecasting via Statistical and Deep Learning Approaches: The German Case. *AppliedMath*, 3(2), 316–342.
<https://doi.org/10.3390/appliedmath3020018>
- [17] Mehra, S., Raut, G., Purkayastha, R. das, Vishvakarma, S. K., & Biasizzo, A. (2023). An Empirical Evaluation of Enhanced Performance Softmax Function in Deep Learning. *IEEE Access*, 11, 34912–34924. <https://doi.org/10.1109/ACCESS.2023.3265327>
- [18] Romero, Á., Dorronsoro, J. R., & Díaz, J. (2019). Day-Ahead Price Forecasting for the Spanish Electricity Market. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(4), 42. <https://doi.org/10.9781/ijimai.2018.04.008>
- [19] Singhal, D., & Swarup, K. S. (2011). Electricity price forecasting using artificial neural networks. *International Journal of Electrical Power & Energy Systems*, 33(3), 550–555.
<https://doi.org/10.1016/j.ijepes.2010.12.009>
- [20] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*.
- [21] Wang, C., Chen, Y., Zhang, S., & Zhang, Q. (2022). Stock market index prediction using deep Transformer model. *Expert Systems with Applications*, 208, 118128.
<https://doi.org/10.1016/j.eswa.2022.118128>

- [22] Wrålsen, B., & Faessler, B. (2022). Multiple Scenario Analysis of Battery Energy Storage System Investment: Measuring Economic and Circular Viability. *Batteries*, 8(2), 7. <https://doi.org/10.3390/batteries8020007>
- [23] Vermeer, W., Chandra Mouli, G. R., & Bauer, P. (2022). A Comprehensive Review on the Characteristics and Modeling of Lithium-Ion Battery Aging. *IEEE Transactions on Transportation Electrification*, 8(2), 2205–2232. <https://doi.org/10.1109/TTE.2021.3138357>
- [24] HUPX. (n.d.). Hungarian Power Exchange. <https://hupx.hu>
- [25] Pippi, K. D., Papadopoulos, T. A., Kryptonidis, G. C., & Kyriakopoulos, E. D. (2023). The net-metering practice in medium-voltage PV-BES prosumers: A techno-economic analysis of the Greek case. *Sustainable Energy, Grids and Networks*, 36, 101156. <https://doi.org/10.1016/j.segan.2023.101156>
- [26] Rotar, R., Jurj, S. L., Susany, R., Opritoiu, F., & Vladutiu, M. (2021). Global Energy Production Computation of a Solar-Powered Smart Home Automation System Using Reliability-Oriented Metrics. *Energies*, 14(9), 2541. <https://doi.org/10.3390/en14092541>
- [27] Ember, *Ember – Data, analysis and insights on clean electricity*. <https://ember-energy.org/>
- [28] Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2022). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, 503, 92–108. <https://doi.org/10.1016/j.neucom.2022.06.111>
- [29] Scabini, L. F. S., & Bruno, O. M. (2021). Structure and Performance of Fully Connected Neural Networks: Emerging Complex Network Properties.
- [30] Schmidt, R. M. (2019). Recurrent Neural Networks (RNNs): A gentle Introduction and Overview.
- [31] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Computation*, 31(7), 1235–1270. https://doi.org/10.1162/neco_a_01199
- [32] Hu, Z., Zhang, J., & Ge, Y. (2021). Handling Vanishing Gradient Problem Using Artificial Derivative. *IEEE Access*, 9, 22371–22377. <https://doi.org/10.1109/ACCESS.2021.3054915>
- [33] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1), 53. <https://doi.org/10.1186/s40537-021-00444-8>

- [34] Borovykh, A., Bohte, S., & Oosterlee, C. W. (2017). Conditional Time Series Forecasting with Convolutional Neural Networks.
- [35] Jierula, A., Wang, S., OH, T.-M., & Wang, P. (2021). Study on Accuracy Metrics for Evaluating the Predictions of Damage Locations in Deep Piles Using Artificial Neural Networks with Acoustic Emission Data. *Applied Sciences*, 11(5), 2314. <https://doi.org/10.3390/app11052314>
- [36] MAKOSSO, T. L., ALMAKTOOF, A., & Abo-Al-Ez, K. (2025). Review of Different Types of Neural Network Architectures. *International Journal of Electrical Engineering and Applied Sciences (IJEAS)*, 7(2). <https://doi.org/10.54554/ijeeas.2024.7.02.006>
- [37] Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, 32(3), 914–938. <https://doi.org/10.1016/j.ijforecast.2015.11.011>
- [38] Lago, J., de Ridder, F., & de Schutter, B. (2018). Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy*, 221, 386–405. <https://doi.org/10.1016/j.apenergy.2018.02.069>
- [39] Nowotarski, J., & Weron, R. (2018). Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable and Sustainable Energy Reviews*, 81, 1548–1568. <https://doi.org/10.1016/j.rser.2017.05.234>
- [40] Lago, J., Marcjasz, G., de Schutter, B., & Weron, R. (2021). Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Applied Energy*, 293, 116983. <https://doi.org/10.1016/j.apenergy.2021.116983>