



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ  
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ  
ΔΙΑΤΑΞΕΩΝ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

**ΠΡΟΒΛΕΨΕΙΣ ΕΝΕΡΓΕΙΑΚΗΣ ΠΑΡΑΓΩΓΗΣ  
ΑΝΕΜΟΓΕΝΝΗΤΡΙΩΝ ΜΕ ΤΗ ΧΡΗΣΗ ΜΕΘΟΔΩΝ  
ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ ΚΑΙ ΜΕΤΑ-  
ΕΚΠΑΙΔΕΥΣΗΣ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Μπουρνάζου Ηρακλή

**Επιβλέπων: Ευάγγελος Μαρινάκης**

Επίκουρος Καθηγητής Ε.Μ.Π

Αθήνα, Απρίλιος, 2024





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ  
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ  
ΔΙΑΤΑΞΕΩΝ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

**ΠΡΟΒΛΕΨΕΙΣ ΕΝΕΡΓΕΙΑΚΗΣ ΠΑΡΑΓΩΓΗΣ  
ΑΝΕΜΟΓΕΝΝΗΤΡΙΩΝ ΜΕ ΤΗ ΧΡΗΣΗ ΜΕΘΟΔΩΝ  
ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ ΚΑΙ ΜΕΤΑ-  
ΕΚΠΑΙΔΕΥΣΗΣ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

Μπουρνάζου Ηρακλή

**Επιβλέπων: Ευάγγελος Μαρινάκης**

Επίκουρος Καθηγητής Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 16/04/2024

.....  
**Δημήτριος Ασκούνης**

Καθηγητής Ε.Μ.Π

.....  
**Ιωάννης Ψαρράς**

Καθηγητής Ε.Μ.Π

.....  
**Ευάγγελος Μαρινάκης**

Επίκ. Καθηγητής Ε.Μ.Π

Αθήνα, Απρίλιος, 2024





.....  
Ηρακλής.Σ Μπουρνάζος

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

Copyright © Ηρακλής.Σ Μπουρνάζος, 2024.

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.



## ΠΕΡΙΛΗΨΗ

Η αιολική ενέργεια αποτελεί μια από τις πιο οικονομικές και δημοφιλείς ανανεώσιμες πηγές ενέργειας. Η μεταβλητότητα των χαρακτηριστικών του ανέμου δημιουργεί δυσκολίες για την βέλτιστη ενσωμάτωσή της στα δίκτυα ισχύος καθώς και για την συνολική διαχείρισή της. Για τον περιορισμό της μη σταθερότητας και των πιθανών μεταβλητών καταστάσεων ενεργειακής παραγωγής, απαιτούνται ευέλικτα μοντέλα πρόβλεψης παραγωγής ικανά να προσφέρουν αξιόπιστα αποτελέσματα σε εύρος χρονικού ορίζοντα και να συνεισφέρουν στην βελτιστοποίηση της ενεργειακής διαχείρισης και του προγραμματισμού. Επιπλέον, στην απελευθερωμένη αγορά ενέργειας η πρόβλεψη ενεργειακής παραγωγής από τις μονάδες αιολικής ενέργειας επωφελεί όλους τους συμμετέχοντες, προσφέροντας μεγαλύτερη ασφάλεια για τις βέλτιστες τιμές εκκαθάρισης των αγορών.

Στην παρούσα διπλωματική εργασία προτείνεται η χρήση μεθόδων μηχανικής μάθησης για τη δημιουργία προβλεπτικών μοντέλων, που βασίζονται σε αριθμητικές προβλέψεις καιρού. Τα μοντέλα αυτά αποσκοπούν στο να παρέχουν ακριβείς και έγκαιρες πληροφορίες σχετικά με την αναμενόμενη ενεργειακή παραγωγή ενός αιολικού πάρκου το οποίο αποτελείται από 4 ανεμογεννήτριες. Στη συγκεκριμένη προσέγγιση, μέσα από την επεξεργασία και ανάλυση ιστορικών δεδομένων, παρουσιάζονται τα αποτελέσματα δοκιμής, σε διαφορετικά χρονικά πλαίσια, από 5 βασικά μοντέλα μηχανικής μάθησης : το LSTM, το Bi-LSTM, το GRU, τον αλγόριθμο Random Forest και το XGBoost. Αυτή η μελέτη προτείνει τη βελτίωση των αποτελεσμάτων μέσω της συνδυαστικής μεθόδου μέτα-μάθησης για τη βελτίωση των προβλέψεων. Χρησιμοποιούνται δυο μέτα-εκπαιδευτές και τα συνολικά αποτελέσματα συγκρίνονται σε διαφορετικούς χρονικούς ορίζοντες. Τα διαφορετικά αποτελέσματα των βασικών μεθόδων συνθέτονται ώστε να υπάρξει ενίσχυση στις γενικευμένες λύσεις. Τα αποτελέσματα δείχνουν πως υπάρχει σημαντική ενίσχυση με τη χρήση των μεθόδων μέτα-εκπαίδευσης στην ικανότητα γενίκευσης και στην απόδοση των μοντέλων. Μέσω της χρήσης αριθμητικών δεδομένων καιρού μπορεί να πραγματοποιηθεί αξιόπιστη πρόβλεψη αιολικής παραγωγής με βήματα πρόβλεψης τόσο 10 λεπτών όσο και ημερησίων.

### ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Αριθμητικές Προβλέψεις Καιρού, Νευρωνικά Δίκτυα, Τυχαία Δάση, XGBoost, Μέτα-Εκπαίδευση

## ABSTRACT

Wind energy is one of the most economical and popular renewable energy sources. The variability of wind characteristics poses challenges for its optimal integration into power grids as well as for the overall energy management. To mitigate this variability and potential fluctuations in energy production, flexible forecasting models are essentials. These models should be capable of providing reliable results over various time horizons and contributing to the optimization of energy management and scheduling. Additionally, in the liberalized energy market, forecasting energy production from wind units benefits all participants by offering greater security and optimal market clearing prices.

This thesis proposes the use of machine learning methods to create predictive models based on numerical weather predictions. These models aim to provide accurate and timely information regarding the expected energy production of a wind farm consisting of 4 wind turbines. In this approach, the results of testing different machine learning models are presented across various time frames: LSTM, Bi-LSTM, GRU, Random Forest algorithm, and XGBoost. This study suggests improving results through ensemble meta-learning methods to enhance predictions. Two meta-trainers are used, and the overall results are compared across different time horizons. The different outcomes of the basic models are combined to reinforce generalized solutions. The results indicate a significant enhancement in the ability to generalize and the performance of the models. Through the use of numerical weather data, reliable wind production forecasts can be made for both 10-minute intervals and daily forecasts.

### KEYWORDS

Numerical weather predictions, Neural Networks, Random Forest, eXtreme Gradient Boosting (XGBoost), Meta-Learning

## ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διπλωματική εργασία ολοκληρώνει επιτυχώς τον κύκλο σπουδών μου στη Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών στο Εθνικό Μετσόβιο Πολυτεχνείο. Αυτό το ακαδημαϊκό ταξίδι με καθόρισε σαν άνθρωπο, με εφοδίασε με γνώσεις, εμπειρίες, αρχές, αξίες και με έκανε πιο ικανό και ανθεκτικό.

Θα ήθελα να ευχαριστήσω θερμά τον καθηγητή κ. Ευάγγελο Μαρινάκη για την δυνατότητα που μου έδωσε να εκπονήσω την συγκεκριμένη διπλωματική εργασία στο εργαστήριο των Ηλεκτρικών Βιομηχανικών Διατάξεων και Συστημάτων Αποφάσεων, καθώς και για την συνολική του στήριξη στους στόχους μου και στην πορεία των φοιτητικών μου χρόνων.

Ευχαριστώ ιδιαίτερα τους κ. Ευστάθιο Σταματόπουλο και κ. Ελισσαίο Σαρμά για τη συνολική τους καθοδήγηση και την εξαιρετική συνεργασία.

Θα ήθελα να ευχαριστήσω επίσης τους φίλους από τη σχολή για τη στήριξή τους, καθώς και τους ανθρώπους που όλα αυτά τα χρόνια ήταν δίπλα μου και με στήριζαν σε όλο αυτό το όμορφο ταξίδι.

Η διπλωματική εργασία αφιερώνεται στην οικογένεια μου και ιδιαίτερα στους γονείς μου και στην αδερφή μου. Τους ευχαριστώ για την ηθική συμπαράσταση, που είναι στο πλευρό μου και με στηρίζουν όλα αυτά τα χρόνια θυσιάζοντας πολλά πράγματα ώστε να επιτύχω τους στόχους μου.

Αθήνα, Μάρτιος 24  
Ηρακλής Μπουρνάζος



## ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

<b>ΠΕΡΙΛΗΨΗ.....</b>	<b>7</b>
<b>ABSTRACT.....</b>	<b>8</b>
<b>ΠΕΡΙΕΧΟΜΕΝΟ ΕΙΚΟΝΩΝ .....</b>	<b>13</b>
<b>ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ .....</b>	<b>17</b>
1.1. <b>ΑΝΤΙΚΕΙΜΕΝΟ ΕΡΓΑΣΙΑΣ .....</b>	<b>17</b>
1.2. <b>ΣΥΜΒΟΛΗ ΔΙΠΛΩΜΑΤΙΚΗΣ .....</b>	<b>18</b>
1.3. <b>ΟΡΓΑΝΩΣΗ ΤΟΜΟΥ.....</b>	<b>19</b>
<b>ΚΕΦΑΛΑΙΟ 2: ΑΙΟΛΙΚΗ ΕΝΕΡΓΕΙΑ .....</b>	<b>21</b>
2.1 <b>Ο ΡΟΛΟΣ ΤΩΝ ΑΝΑΝΕΩΣΙΜΩΝ ΠΗΓΩΝ ΕΝΕΡΓΕΙΑΣ ΣΤΗΝ ΕΝΕΡΓΕΙΑΚΗ ΜΕΤΑΒΑΣΗ .....</b>	<b>21</b>
2.2 <b>ΣΤΟΙΧΕΙΑ ΑΙΟΛΙΚΗΣ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΕΝΣΩΜΑΤΩΣΗΣ .....</b>	<b>22</b>
2.3 <b>Η ΒΑΣΙΚΗ ΛΕΙΤΟΥΡΓΙΑ ΤΩΝ ΑΝΕΜΟΓΕΝΝΗΤΡΙΩΝ .....</b>	<b>23</b>
2.4 <b>ΟΙΚΟΝΟΜΙΚΗ ΒΙΩΣΙΜΟΤΗΤΑ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΕΣ ΒΕΛΤΙΩΣΕΙΣ.....</b>	<b>25</b>
<b>ΚΕΦΑΛΑΙΟ 3: ΠΡΟΒΛΕΨΕΙΣ ΑΙΟΛΙΚΗΣ ΕΝΕΡΓΕΙΑΣ .....</b>	<b>27</b>
3.1 <b>ΑΡΙΘΜΗΤΙΚΕΣ ΠΡΟΒΛΕΨΕΙΣ ΚΑΙΡΟΥ .....</b>	<b>27</b>
3.1.1 <b>Ταχύτητα Ανέμου .....</b>	<b>27</b>
3.1.2 <b>Κατεύθυνση Ανέμου.....</b>	<b>28</b>
3.1.3 <b>Πυκνότητα Αέρα.....</b>	<b>28</b>
3.2 <b>Η ΣΗΜΑΣΙΑ ΤΗΣ ΠΡΟΒΛΕΨΗΣ .....</b>	<b>28</b>
3.3 <b>ΧΡΟΝΙΚΟΙ ΟΡΙΖΟΝΤΕΣ ΠΡΟΒΛΕΨΗΣ .....</b>	<b>30</b>
3.3.1 <b>Πολύ βραχυπρόθεσμη πρόβλεψη:.....</b>	<b>30</b>
3.3.2 <b>Βραχυπρόθεσμη πρόβλεψη:.....</b>	<b>30</b>
3.3.3 <b>Μεσοπρόθεσμη πρόβλεψη:.....</b>	<b>31</b>
3.3.4 <b>Μακροπρόθεσμη πρόβλεψη:.....</b>	<b>31</b>
<b>ΚΕΦΑΛΑΙΟ 4: ΜΟΝΤΕΛΑ ΠΡΟΒΛΕΨΗΣ.....</b>	<b>33</b>
4.1 <b>ΜΟΝΤΕΛΑ ΣΥΝΕΧΙΖΟΜΕΝΗΣ ΠΡΟΒΛΕΨΗΣ (PERSISTENCE MODELS) .....</b>	<b>33</b>
4.2 <b>ΦΥΣΙΚΕΣ ΜΕΘΟΔΟΙ .....</b>	<b>33</b>
4.3 <b>ΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ .....</b>	<b>34</b>
4.4 <b>ΜΕΘΟΔΟΙ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ.....</b>	<b>34</b>
4.4.1 <b>ΤΥΠΟΙ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ.....</b>	<b>35</b>
4.4.1.1 <b>Εποπτευόμενη Μάθηση (Supervised Learning) .....</b>	<b>36</b>
4.4.1.2 <b>Μη Εποπτευόμενη Μάθηση (Unsupervised Learning) .....</b>	<b>36</b>
4.4.1.3 <b>Ενισχυτική Μάθηση (Reinforcement Learning ) .....</b>	<b>36</b>
4.4.2 <b>ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΠΡΟΒΛΗΜΑΤΩΝ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ .....</b>	<b>37</b>
4.4.2.1 <b>Παλινδρόμηση (Regression) .....</b>	<b>37</b>
4.4.2.2 <b>Ταξινόμηση (Classification) .....</b>	<b>37</b>
4.4.2.3 <b>Συσταδοποίηση (Clustering) .....</b>	<b>37</b>
4.4.3 <b>ΑΛΓΟΡΙΘΜΟΙ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ ΣΥΧΝΑ ΕΜΦΑΝΙΖΟΜΕΝΟΙ ΣΤΗ ΒΙΒΛΙΟΓΡΑΦΙΑ             ΓΙΑ ΕΞΑΓΩΓΗ ΠΡΟΒΛΕΨΕΩΝ .....</b>	<b>37</b>

4.4.3.1	Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANN) .....	37
4.4.3.2	Τυχαία Δάση (Random Forest – RF) .....	38
4.4.3.3	Δέντρα Αποφάσεων (Decision Trees – DT) .....	39
4.4.3.4	Μοντέλα Ασαφούς Λογικής (Fuzzy Logic – FL).....	40
4.4.3.5	Μηχανή Διανυσματικής Υποστήριξης για Παλινδρόμηση (Support Vector Regression–SVR)	40
4.4.3.6	k Πλησιέστεροι Γείτονες (k Nearest Neighbors – kNN).....	41
4.4.3.7	XGBoost.....	41
4.4.3.8	GRU (Gated Recurrent Unit) .....	42
4.4.3.9	Categorical Boosting.....	43
4.4.3.10	LSTM (Long Short-Term Memory):.....	43
<b>4.4.4</b>	<b>ΣΥΝΔΥΑΣΤΙΚΟΙ ΜΕΘΟΔΟΙ (ENSEMBLE METHODS) .....</b>	<b>44</b>
4.4.4.1	Ανταγωνιστική (Competitive) Πρόβλεψη .....	44
4.4.4.2	Συνεργατική (Cooperative) Πρόβλεψη .....	44
<b>4.4.5</b>	<b>ΥΒΡΙΔΙΚΟΙ ΜΕΘΟΔΟΙ (Hybrid Methods) .....</b>	<b>44</b>
<b>ΚΕΦΑΛΑΙΟ 5: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΤΩΝ ΜΟΝΤΕΛΩΝ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ... 45</b>		
<b>5.1</b>	<b>ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ .....</b>	<b>45</b>
5.1.1	Συνάρτηση Ενεργοποίησης.....	46
5.1.2	Νευρωνικά Δίκτυα Εμπρόσθιας Τροφοδότησης-Προώθησης (Feedforward Neural Networks / FNNs)	49
5.1.3	Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Network / RNN).....	50
5.1.4	ΒΑΘΙΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ (Deep Neural Network / DNN) .....	51
5.1.5	ΣΥΝΕΛΙΚΤΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ (CNN) .....	52
<b>5.2</b>	<b>ΔΙΚΤΥΑ ΜΑΚΡΑΣ ΚΑΙ ΒΡΑΧΕΙΑΣ ΜΝΗΜΗΣ - LSTM .....</b>	<b>52</b>
5.2.1	Bidirectional LSTM (Bi-LSTM).....	53
<b>5.3</b>	<b>Gated Recurrent Unit (GRU).....</b>	<b>54</b>
<b>5.4</b>	<b>XGBOOST .....</b>	<b>55</b>
<b>5.5</b>	<b>ΤΥΧΑΙΑ ΔΑΣΗ-RANDOM FOREST.....</b>	<b>56</b>
<b>5.6</b>	<b>SVR – ΠΑΛΙΝΔΡΟΜΗΣΗ ΔΙΑΝΥΣΜΑΤΩΝ ΥΠΟΣΤΗΡΙΞΗΣ .....</b>	<b>58</b>
<b>5.7</b>	<b>ΜΕΤΡΙΚΕΣ ΑΞΙΟΛΟΓΗΣΗΣ ΜΟΝΤΕΛΩΝ ΠΡΟΒΛΕΨΗΣ .....</b>	<b>59</b>
<b>ΚΕΦΑΛΑΙΟ 6: ΠΡΑΚΤΙΚΟ ΜΕΡΟΣ - ΠΡΟΒΛΕΨΗ ΠΑΡΑΓΟΜΕΝΗΣ ΕΝΕΡΓΕΙΑΣ ΑΝΕΜΟΓΕΝΝΗΤΡΙΩΝ..... 61</b>		
<b>6.1</b>	<b>ΤΟ DATASET.....</b>	<b>62</b>
<b>6.2</b>	<b>ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ.....</b>	<b>64</b>
<b>6.3</b>	<b>ΜΕΛΕΤΗ ΕΠΟΧΙΑΚΟΤΗΤΑΣ - ΕΙΣΑΓΩΓΗ ΝΕΩΝ ΔΕΔΟΜΕΝΩΝ ΕΙΣΟΔΟΥ.....</b>	<b>75</b>
<b>6.4</b>	<b>ΣΥΣΧΕΤΙΣΕΙΣ.....</b>	<b>78</b>
<b>6.5</b>	<b>ΕΚΠΑΙΔΕΥΣΗ ΔΕΔΟΜΕΝΩΝ .....</b>	<b>81</b>
6.5.1	ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ ΔΙΑΣΤΗΜΑΤΟΣ 2 ΩΡΩΝ .....	86
6.5.2	ΗΜΕΡΗΣΙΕΣ ΠΡΟΒΛΕΨΕΙΣ .....	91
<b>ΚΕΦΑΛΑΙΟ 7: ΕΠΙΛΟΓΟΣ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ..... 97</b>		
<b>7.1</b>	<b>ΓΕΝΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ .....</b>	<b>97</b>
<b>7.2</b>	<b>ΠΕΡΙΟΡΙΣΜΟΙ.....</b>	<b>98</b>
<b>7.3</b>	<b>ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ .....</b>	<b>99</b>
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ.....</b>		<b>100</b>



## ΠΕΡΙΕΧΟΜΕΝΟ ΕΙΚΟΝΩΝ

Εικόνα 1: Προσθήκη νέας αιολικής ισχύος στην Ευρώπη ανά έτος.....	22
Εικόνα 2: Συνολική διαθέσιμη χωρητικότητα (MW) Ελληνικού δικτύου ανά έτος [5].....	23
Εικόνα 3: Καμπύλη αεροδυναμικού συντελεστή ισχύος συναρτήσει λόγου ταχύτητας ακροπτερυγίου $C_p(\lambda)$ [6].....	24
Εικόνα 4: Μεταβολή LCOE χειρσαίων αιολικών πάρκων 1983-2022 [8].....	26
Εικόνα 5: Μεταβολή LCOE χειρσαίων αιολικών στην Ελλάδα 2010-2022 [8].....	26
Εικόνα 6: Σχέση Τεχνητής Νοημοσύνης με Μηχανική Μάθηση και Βαθιά Μάθηση .....	35
Εικόνα 7: Εποπτευόμενη, Μη Εποπτευόμενη και Ενισχυτική Μάθηση.....	36
Εικόνα 8: Βασική Δομή Νευρωνικού Δικτύου .....	45
Εικόνα 9: Μοντέλο Τεχνητού Νευρωνικού Δικτύου .....	46
Εικόνα 10: Σιγμοειδής Συνάρτηση (Sigmoid Function) .....	47
Εικόνα 11: Συνάρτηση Εκθετικής Γραμμικής Μονάδας (ReLU - Rectified Linear Unit) ...	47
Εικόνα 12: Συνάρτηση Υπερβολικής Εφαπτομένης (Hyperbolic Tangent - Tanh).....	48
Εικόνα 13: Συνάρτηση Softmax.....	48
Εικόνα 14: Νευρωνικό Δίκτυο Εμπρόσθιας Τροφοδότησης με 2 κρυφά επίπεδα.....	49
Εικόνα 15: Αναδρομικό Νευρωνικό Δίκτυο - RNN .....	50
Εικόνα 16: Διαφορά μεταξύ ANN και DNN στα κρυφά επίπεδα [69].....	52
Εικόνα 17: Δομή LSTM [72] .....	53
Εικόνα 18: Δομή GRU [72].....	54
Εικόνα 19: Διαδικασία εκμάθησης με την τεχνική του Gradient Boosting .....	55
Εικόνα 20: Διαδικασία εκμάθησης με την τεχνική του Random Forest [78] .....	57
Εικόνα 21: Σχηματική απεικόνιση διαδικασίας εξαγωγής προβλέψεων αιολικής παραγωγής. .....	62
Εικόνα 22: Εκτύπωση κενών τιμών ανά στήλη δεδομένων που αφαιρέθηκαν για το έτος 2019 .....	65
Εικόνα 23: Εκτύπωση κενών τιμών ανά στήλη δεδομένων που αφαιρέθηκαν για το έτος 2020 .....	66
Εικόνα 24: Εκτύπωση κενών τιμών ανά στήλη δεδομένων που αφαιρέθηκαν για το έτος 2021 .....	66
Εικόνα 25: Εκτύπωση κενών τιμών ανά στήλη δεδομένων που αφαιρέθηκαν για το έτος 2022 .....	67
Εικόνα 26: Εκτύπωση μορφής δεδομένων ανά στήλη ενδιαφέροντος .....	68
Εικόνα 27: Αρνητικές τιμές ανά στήλες ενδιαφέροντος-Ανεμογεννήτρια 1 .....	68
Εικόνα 28: Ακραίες τιμές ανά στήλες ενδιαφέροντος Ανεμογεννήτρια 1 .....	69
Εικόνα 29: Αρνητικές τιμές ανά στήλες ενδιαφέροντος-Ανεμογεννήτρια 2 .....	69
Εικόνα 30: Ακραίες τιμές ανά στήλες ενδιαφέροντος Ανεμογεννήτρια 2 .....	69
Εικόνα 31: Αρνητικές τιμές ανά στήλες ενδιαφέροντος-Ανεμογεννήτρια 3 .....	70
Εικόνα 32: Ακραίες τιμές ανά στήλες ενδιαφέροντος Ανεμογεννήτρια 3 .....	70
Εικόνα 33: Αρνητικές τιμές ανά στήλες ενδιαφέροντος-Ανεμογεννήτρια 4 .....	70
Εικόνα 34: Ακραίες τιμές ανά στήλες ενδιαφέροντος Ανεμογεννήτρια 4 .....	71
Εικόνα 35: Μορφή δεδομένων, αρνητικές και ακραίες τιμές ύστερα από την προ- επεξεργασία για την Ανεμογεννήτρια 1 .....	72
Εικόνα 36: Μορφή δεδομένων, αρνητικές και ακραίες τιμές ύστερα από την προ- επεξεργασία για την Ανεμογεννήτρια 2 .....	73
Εικόνα 37: Μορφή δεδομένων, αρνητικές και ακραίες τιμές ύστερα από την προ- επεξεργασία για την Ανεμογεννήτρια 2 .....	74

Εικόνα 38: Μορφή δεδομένων, αρνητικές και ακραίες τιμές ύστερα από την προ-επεξεργασία για την Ανεμογεννήτρια 4 .....	75
Εικόνα 39: Μηνιαία παραγωγή ενέργειας και Μέση παραγωγή ενέργειας Ανεμογεννήτρια 1 .....	76
Εικόνα 40: Μηνιαία παραγωγή ενέργειας και Μέση παραγωγή ενέργειας Ανεμογεννήτρια 2 .....	76
Εικόνα 41: Μηνιαία παραγωγή ενέργειας και Μέση παραγωγή ενέργειας Ανεμογεννήτρια 3 .....	77
Εικόνα 42: Μηνιαία παραγωγή ενέργειας και Μέση παραγωγή ενέργειας Ανεμογεννήτρια 4 .....	77
Εικόνα 43: Πίνακας συσχέτισης χαρακτηριστικών Ανεμογεννήτρια 1 .....	79
Εικόνα 44: Πίνακας συσχέτισης χαρακτηριστικών Ανεμογεννήτρια 2 .....	79
Εικόνα 45: Πίνακας συσχέτισης χαρακτηριστικών Ανεμογεννήτρια 3 .....	80
Εικόνα 46: Πίνακας συσχέτισης χαρακτηριστικών Ανεμογεννήτρια 4 .....	80
Εικόνα 47: Σύγκριση συνολικών αποτελεσμάτων πρόβλεψης με πραγματικές τιμές.....	87
Εικόνα 48: Αποτέλεσμα πρόβλεψης βέλτιστου μοντέλου (Ανεμογεννήτρια 1).....	87
Εικόνα 49: Σύγκριση συνολικών αποτελεσμάτων πρόβλεψης με πραγματικές τιμές.....	88
Εικόνα 50: Αποτέλεσμα πρόβλεψης βέλτιστου μοντέλου (Ανεμογεννήτρια 2).....	89
Εικόνα 51: Σύγκριση συνολικών αποτελεσμάτων πρόβλεψης με πραγματικές τιμές.....	89
Εικόνα 52: Αποτέλεσμα πρόβλεψης βέλτιστου μοντέλου (Ανεμογεννήτρια 3).....	90
Εικόνα 53: Σύγκριση συνολικών αποτελεσμάτων πρόβλεψης με πραγματικές τιμές.....	90
Εικόνα 54: Αποτέλεσμα πρόβλεψης βέλτιστου μοντέλου (Ανεμογεννήτρια 4).....	91
Εικόνα 55: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-SVR με τις πραγματικές τιμές - Ανεμογεννήτρια 1 .....	92
Εικόνα 56: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-XGBoost με τις πραγματικές τιμές - Ανεμογεννήτρια 1 .....	93
Εικόνα 57: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-SVR με τις πραγματικές τιμές - Ανεμογεννήτρια 2 .....	93
Εικόνα 58: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-XGBoost με τις πραγματικές τιμές - Ανεμογεννήτρια 2 .....	94
Εικόνα 59: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-SVR με τις πραγματικές τιμές - Ανεμογεννήτρια 3 .....	94
Εικόνα 60: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-XGBoost με τις πραγματικές τιμές - Ανεμογεννήτρια 3 .....	95
Εικόνα 61: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-SVR με τις πραγματικές τιμές - Ανεμογεννήτρια 4 .....	95
Εικόνα 62: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-XGBoost με τις πραγματικές τιμές - Ανεμογεννήτρια 4 .....	96

## ΠΕΡΙΕΧΟΜΕΝΟ ΠΙΝΑΚΩΝ

Πίνακας 1: Χρονική ταξινόμηση προβλέψεων αιολικής παραγωγής [21].....	32
Πίνακας 2: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση ANN..	38
Πίνακας 3: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση Random Forest.....	39
Πίνακας 4: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση Decision Trees .....	40
Πίνακας 5: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση Fuzzy Logic.....	40
Πίνακας 6: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση SVM..	41
Πίνακας 7: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση kNN...	41
Πίνακας 8: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση XGBoost .....	42
Πίνακας 9: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση GRU ..	43
Πίνακας 10: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση CatBoost.....	43
Πίνακας 11:Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση LSTM	44
Πίνακας 12: Αποτελέσματα εκπαίδευσης LSTM .....	81
Πίνακας 13: Αποτελέσματα εκπαίδευσης Bi-LSTM .....	82
Πίνακας 14: Αποτελέσματα εκπαίδευσης GRU.....	83
Πίνακας 15: Αποτελέσματα εκπαίδευσης Random Forest.....	83
Πίνακας 16: Αποτελέσματα εκπαίδευσης XGBoost .....	84
Πίνακας 17: Αποτελέσματα εκπαίδευσης μεθόδων meta-learning.....	85
Πίνακας 18: Συγκεντρωτικά αποτελέσματα για την Ανεμογεννήτρια 1.....	85
Πίνακας 19: Συγκεντρωτικά αποτελέσματα για την Ανεμογεννήτρια 2.....	85
Πίνακας 20: Συγκεντρωτικά αποτελέσματα για την Ανεμογεννήτρια 3.....	85
Πίνακας 21: Συγκεντρωτικά αποτελέσματα για την Ανεμογεννήτρια 4.....	85
Πίνακας 22: Αποτελέσματα βραχυπρόθεσμης πρόβλεψης Ανεμογεννήτριας 1 .....	86
Πίνακας 23: Αποτελέσματα βραχυπρόθεσμης πρόβλεψης Ανεμογεννήτριας 2 .....	88
Πίνακας 24: Αποτελέσματα βραχυπρόθεσμης πρόβλεψης Ανεμογεννήτριας 3 .....	89
Πίνακας 25: Αποτελέσματα βραχυπρόθεσμης πρόβλεψης Ανεμογεννήτριας 4 .....	90
Πίνακας 26: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-SVR - Ανεμογεννήτρια 1.....	92
Πίνακας 27: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-XGBoost - Ανεμογεννήτρια 1 .....	92
Πίνακας 28: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-SVR - Ανεμογεννήτρια 2.....	93
Πίνακας 29: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-XGBoost - Ανεμογεννήτρια 2 .....	93
Πίνακας 30: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-SVR - Ανεμογεννήτρια 3.....	94
Πίνακας 31: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-XGBoost - Ανεμογεννήτρια 3 .....	94
Πίνακας 32: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-SVR - Ανεμογεννήτρια 4.....	95
Πίνακας 33: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-XGBoost - Ανεμογεννήτρια 4 .....	95
Πίνακας 34: Χρόνοι εκπαίδευσης μοντέλων για κάθε Ανεμογεννήτρια.....	99



# ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

## 1.1. ΑΝΤΙΚΕΙΜΕΝΟ ΕΡΓΑΣΙΑΣ

Ακραία φαινόμενα, μεγάλες οικολογικές καταστροφές, απόρροια των ολοένα αυξανόμενων τόσο σε ένταση όσο και σε συχνότητα φαινομένων κλιματικής αλλαγής, καθώς και οικονομικές αναταραχές των αγορών, έχουν δημιουργήσει νέες παγκόσμιες συνθήκες. Η βιωσιμότητα του πλανήτη τίθεται σε αμφισβήτηση και η ανάγκη αξιοποίησης των ανανεώσιμων πηγών ενέργειας κρίνεται επιτακτική. Η ανάγκη για μορφές ενέργειας που να είναι φιλικότερες στο περιβάλλον έχει αναδείξει την παραγωγή ενέργειας από ανανεώσιμες πηγές κομβικό παράγοντα για την Ελληνική οικονομία αλλά και την Ευρωπαϊκή αγορά ενέργειας. Μέχρι πρόσφατα που η ενέργεια προερχόταν κυρίως από συμβατικές μονάδες παραγωγής ορυκτών καυσίμων, η διαθέσιμη ηλεκτρική ισχύς στο δίκτυο είχε μια καθορισμένη τιμή που οριζόταν από την κατανάλωση του φορτίου. Η αυξανόμενη ενσωμάτωση ανανεώσιμων πηγών ενέργειας διαμορφώνει νέες συνθήκες καθώς και νέα προβλήματα για την κάλυψη των αναγκών ηλεκτρικής ενέργειας στο δίκτυο.

Για την ενσωμάτωση αυτών των πηγών με τρόπο που να προκαλεί τις λιγότερες δυνατές τεχνικές και οικονομικές συνέπειες σε όλους τους εμπλεκόμενους γύρω από την αγορά ηλεκτρικής ενέργειας, απαιτούνται αξιόπιστες τεχνικές που να μπορούν να προβλέπουν με ακρίβεια και αξιοπιστία την παραγόμενη ενέργεια. Η αξιοποίηση ορισμένων στοχαστικών μεταβλητών όπως η ταχύτητα και η διεύθυνση του ανέμου, απρόβλεπτες σε μεγάλο βαθμό κατά τη διάρκεια της ημέρας, σε συνδυασμό με τον τρόπο λειτουργίας των ανεμογεννητριών μπορούν να δώσουν αξιόπιστες λύσεις στο πρόβλημα πρόβλεψης παραγωγής ενέργειας, σε εύρος διαφορετικών χρονικών περιόδων, διαμορφώνοντας μεγαλύτερη ευελιξία για την βέλτιστη διείσδυσή τους στο δίκτυο.

Συνεπώς, είναι καταρχήν απρόβλεπτο πόση διαθέσιμη ισχύ θα έχει το ηλεκτρικό δίκτυο, και αν η ισχύς αυτή επαρκεί κάθε δεδομένη στιγμή για να καλύψει τη ζήτηση. Χωρίς κάποιο σύστημα υποστήριξης, η ηλεκτροπαραγωγή με Ανανεώσιμες Πηγές Ενέργειας αφήνει μια χώρα εκτεθειμένη σε σημαντικούς κινδύνους, όπως οι διαρκείς απώλειες παροχής ρεύματος, οι πτώσεις τάσης και η αδυναμία ενεργειακής κάλυψης. Η πρόβλεψη της παραγωγής εξασφαλίζει ότι το δίκτυο θα λειτουργεί με σταθερότερα και συνέπεια σε διαφορετικές συνθήκες ζήτησης.

Επιπλέον, έχοντας υπόψη ότι βασικό χαρακτηριστικό της ηλεκτρικής ενέργειας είναι η δυσκολία αποθήκευσής της σε συνδυασμό με την κοστοβόρα επιλογή ενεργειακών εφεδρειών, καταλαβαίνει κανείς πόσο σημαντική είναι η διατήρηση της παραγωγής και της ζήτησης ηλεκτρικής ενέργειας στο ίδιο επίπεδο με όσο το δυνατόν μικρότερες αποκλίσεις. Υπάρχει συνεπώς, ανάγκη να γίνονται εκ των προτέρων εκτιμήσεις για την ποσότητα της παραγόμενης ενέργειας. Η πρόβλεψη ζήτησης ηλεκτρικού φορτίου σε συνδυασμό με την πρόβλεψη παραγωγής ηλεκτρικής ενέργειας τόσο σε βραχυπρόθεσμους όσο και μακροπρόθεσμους ορίζοντες, είναι ικανή να βελτιστοποιήσει τη διαδικασία ένταξης μονάδων ανανεώσιμων πηγών ενέργειας στα συστήματα ηλεκτρικής ενέργειας (ΣΗΕ) καθώς και την ανταλλαγή ηλεκτρικής ενέργειας με γειτονικά δίκτυα.

Μέσω της διαδικασίας πρόβλεψης παραγωγής ενέργειας η βελτιστοποίηση λειτουργιών διαχείρισης προσφοράς φορτίου στο δίκτυο οδηγεί στην οικονομικότερη λειτουργία για την κάλυψη του ισοζυγίου προσφοράς και ζήτησης. Διαδικασίες όπως η ένταξη μονάδων, η αγορά επιπλέον ενέργειας από γειτονικά δίκτυα, η ρύθμιση της τάσης των ζυγών, η μείωση των τιμών εκκαθάρισης και η ταυτόχρονη διαδικασία της συνολικής οικονομικότερης

λειτουργίας του δικτύου που μεγιστοποιεί την ένταξη Α.Π.Ε, είναι τομείς που επηρεάζονται από την αξιοπιστία των αποτελεσμάτων πρόβλεψης.

Πιθανές αποκλίσεις μεταξύ προβλεπόμενης και παραγόμενης ισχύος εκτός των αποδεκτών ορίων, ενδέχεται να οδηγήσουν στη δημιουργία ποικίλων προβλημάτων στον ενεργειακό προγραμματισμό και στην οικονομική λειτουργία. Η ανεπιθύμητη εξάρτηση από συμβατικές μονάδες παραγωγής, η αύξηση χρήσης εφεδρειών, η πιθανή ύπαρξη προστίμων στους εμπλεκόμενους από τους φορείς της αγοράς ενέργειας καθώς και η μη αξιοποίηση περισσεύας παραγόμενης ανανεώσιμης ενέργειας αποτελούν ορισμένα από τα πιθανά αποτελέσματα των μη αποτελεσματικών προβλέψεων ηλεκτρικής παραγωγής. Απόρροια αυτών αποτελεί η συνολική αύξηση του κόστους παραγωγής της ηλεκτρικής ενέργειας και η διαμόρφωση αναξιοπιστίας στην αγορά ενέργειας με την ύπαρξη έντονων διακυμάνσεων στις τιμές κατανάλωσης.

Τόσο στο πρόβλημα της στοχαστικότητας των μετεωρολογικών δεδομένων όσο και στη συσχέτιση των δεδομένων εισόδου - εξόδου που βασίζονται σε μετρήσεις αισθητήρων και άλλων μετεωρολογικών δεδομένων, αλλά και στην ερμηνεία του μοναδικού τρόπου λειτουργίας των εκάστοτε ανεμογεννητριών, η ανάλυση δεδομένων και η μηχανική μάθηση καλείται να δώσει απαντήσεις. Η ευελιξία και οι δυνατότητες της μηχανικής μάθησης για την δημιουργία μοντέλων τα οποία μπορούν να αποτυπώνουν δυσνόητες σχέσεις αλληλεξάρτησης μεταξύ εισόδου και εξόδου με αποτελέσματα σε πολύ σύντομα χρονικά διαστήματα, έχουν δημιουργήσει τις ιδανικές συνθήκες για τον βέλτιστο χειρισμό και μπορούν να συνεισφέρουν στον ιδανικό τρόπο λειτουργίας και ενσωμάτωσης των Α.Π.Ε στο δίκτυο.

## 1.2. ΣΥΜΒΟΛΗ ΔΙΠΛΩΜΑΤΙΚΗΣ

Στην παρούσα διπλωματική, αναπτύσσονται διαφορετικά μοντέλα μηχανικής μάθησης, τα οποία εκπαιδεύονται από πραγματικά δεδομένα που έχουν συλλεχθεί από μετρήσεις αισθητήρων ενός πραγματικού αιολικού πάρκου, το οποίο αποτελείται συνολικά από 4 ανεμογεννήτριες. Στο θεωρητικό υπόβαθρο αναφέρονται βασικές έννοιες τόσο για την αιολική παραγωγή όσο και για την διαδικασία των προβλέψεων με βάση τις υφιστάμενες μεθόδους πρόβλεψης. Αναφέρονται οι αρχιτεκτονικές και ο τρόπος με τον οποίο διάφορα μοντέλα μηχανικής μάθησης μπορούν να εμφανίζουν μεγάλες δυνατότητες στο συγκεκριμένο κλάδο. Επιπλέον, παρουσιάζονται τεχνικές ανάλυσης δεδομένων και τα βήματα προεπεξεργασίας για την επιτυχημένη ενσωμάτωση των κατάλληλων δεδομένων στα προβλεπτικά μοντέλα και την ενίσχυση της απόδοσης πρόβλεψης. Τα δεδομένα προέρχονται από διάφορες μετρήσεις μεταβλητών με βήμα μέτρησης 10-λεπτών. Οι εμφανιζόμενες αλλαγές στις τιμές των συσχετίσεων, αναδεικνύουν τη σημασία της ενδεδειγμένης ανάλυσης και επεξεργασίας των δεδομένων εισόδου και έχουν κρίσιμες επιπτώσεις στα αποτελέσματα εξόδου για το εκάστοτε μοντέλο. Ειδικότερα, εκπαιδεύονται συνολικά 5 βασικά μοντέλα μηχανικής μάθησης για να επιτευχθεί η πρόβλεψη παραγωγής ηλεκτρικής ενέργειας από τις ανεμογεννήτριες, ενώ παρουσιάζονται και 2 συνδυαστικά μοντέλα μετά-εκπαίδευσης (meta-learning), τα οποία συνθέτουν χαρακτηριστικά από τα βασικά μοντέλα για την επίτευξη αποτελεσμάτων με το μικρότερο δυνατό σφάλμα. Σε αυτό το πραγματικό πρόβλημα τα προτεινόμενα βασικά μοντέλα είναι:

- 1) Long Short-Term Memory (LSTM)
- 2) Bidirectional-LSTM
- 3) eXtreme Gradient Boosting (XGBoost),
- 4) Gated Recurrent Unit (GRU)
- 5) Τυχαία Δέντρα (Random Forest)

Οι μέθοδοι μετά-εκπαίδευσης ενσωματώνουν αυτά τα αποτελέσματα και παρουσιάζουν ελαχιστοποίηση των σφαλμάτων συγκριτικά με τα αποτελέσματα των αυτόνομων μοντέλων πρόβλεψης. Πιο συγκεκριμένα ως μετά-εκπαιδευτές χρησιμοποιήθηκαν οι εξής:

- 1) META - Support Vector Machine (SVR)
- 2) META - XGBoost

Σκοπός αποτελεί η παρουσίαση των δυνατοτήτων του συγκεκριμένου κλάδου για την βελτιστοποίηση λειτουργιών της διαχείρισης ενέργειας και της αποδοτικότερης ενσωμάτωσης ΑΠΕ, η ανάδειξη της πρακτικής εφαρμογής των συγκεκριμένων μοντέλων μηχανικής μάθησης σε πραγματικά προβλήματα πρόβλεψης παραγωγής ενέργειας, αλλά και η πρακτική βελτιστοποίηση των αποτελεσμάτων με τη χρήση των μοντέλων μετά-εκπαίδευσης στο συγκεκριμένο dataset.

### 1.3. ΟΡΓΑΝΩΣΗ ΤΟΜΟΥ

- Στο πρώτο κεφάλαιο παρουσιάζεται συνοπτικά το αντικείμενο εργασίας, η συμβολή της διπλωματικής, καθώς και η δομή της.
- Στο δεύτερο κεφάλαιο γίνεται αναφορά σε στοιχεία τα οποία αναδεικνύουν τη σημαντικότητα των ανανεώσιμων πηγών ενέργειας στην πράσινη μετάβαση. Αναφέρονται στοιχεία ενσωμάτωσης αιολικής ενέργειας σε παγκόσμιο επίπεδο αλλά και στη χώρα μας, καθώς και οι λόγοι που καθιστούν την αιολική παραγωγή μια από τις πιο οικονομικά βιώσιμες και φιλικές προς το περιβάλλον μορφές ενέργειας. Επιπλέον, περιγράφεται και ο τρόπος λειτουργίας των ανεμογεννητριών
- Στο τρίτο κεφάλαιο αναφέρονται πληροφορίες για τις αριθμητικές προβλέψεις καιρού και της σημαντικότητάς τους για την εξαγωγή προβλέψεων αιολικής παραγωγής σε εκπαιδευμένα μοντέλα. Τονίζεται η σημασία ευέλικτων και αποδοτικών προβλέψεων ενεργειακής παραγωγής για την βελτιστοποίηση συστημάτων διαχείρισης ενέργειας, ενώ επισημαίνεται και η σημασία του συγκεκριμένου αντικειμένου ως προς τον τρόπο λειτουργίας της αγοράς ενέργειας. Επίσης, βασιζόμενη στην υφιστάμενη βιβλιογραφία, πραγματοποιείται και αναλυτική χρονική ταξινόμηση της πρόβλεψης ενεργειακής παραγωγής και παρουσιάζεται η χρησιμότητα των εφαρμογών πρόβλεψης σε διαφορετικούς χρονικούς ορίζοντες.
- Στο τέταρτο κεφάλαιο παρουσιάζονται οι διάφοροι μέθοδοι πρόβλεψης που υπάρχουν για το πρόβλημα της εκτίμησης ενεργειακής παραγωγής. Παρουσιάζονται οι προσεγγίσεις εξαγωγής αποτελεσμάτων με τη χρήση μοντέλων συνεχιζόμενης πρόβλεψης, στατιστικών μοντέλων, φυσικών μεθόδων και μοντέλων μηχανικής μάθησης. Με βάση τη βιβλιογραφία αναφέρονται τα πλεονεκτήματα και τα μειονεκτήματα της εκάστοτε προσέγγισης. Δίνεται έμφαση στην έννοια των μεθόδων τεχνητής νοημοσύνης, στα προβλήματα που μπορούν να αντιμετωπίσουν, ενώ παρουσιάζεται και αναλυτική έρευνα των κυριότερων και πιο δημοφιλών μοντέλων που είναι συχνά εμφανιζόμενα στην επιστημονική βιβλιογραφία.
- Στο πέμπτο κεφάλαιο πραγματοποιείται λεπτομερής περιγραφή τόσο του θεωρητικού υποβάθρου, όσο και του τρόπου λειτουργίας των βασικών μοντέλων που επιλέχθηκαν για την αντιμετώπιση του συγκεκριμένου προβλήματος. Ειδικότερα, περιγράφεται η αναλυτική λειτουργία των νευρωνικών δικτύων, των αναδρομικών νευρωνικών δικτύων, των LSTM, των GRU, των αλγορίθμων Random Forest, XGBoost και των μηχανών διανυσμάτων υποστήριξης για παλινδρόμηση (SVR).
- Στο έκτο κεφάλαιο περιγράφονται αναλυτικά τα βήματα που απαιτούνται για την υλοποίηση της δημιουργίας μοντέλων πρόβλεψης, τα οποία μέσω αριθμητικών

δεδομένων καιρού θα μπορούν να πραγματοποιούν αξιόπιστες προβλέψεις για ένα αιώλικό πάρκο το οποίο αποτελείται από 4 ανεμογεννήτριες. Παρουσιάζονται αναλυτικά τα βήματα υλοποίησης όπως η διαχείριση, η προ-επεξεργασία και η ανάλυση των δεδομένων για τη βελτιστοποίηση της απόδοσης πρόβλεψης των βασικών μοντέλων. Επιπλέον, παρουσιάζονται και οι μέτα-εκπαιδευτές και αναφέρεται εκτενώς ο τρόπος λειτουργίας τους για την αξιοποίηση των συνδυαστικών αποτελεσμάτων πρόβλεψης. Η αξιολόγηση τόσο των βασικών μοντέλων όσο και των μοντέλων μέτα-εκπαίδευσης εκτός από την διαδικασία εκπαίδευσης και δοκιμής βασίζεται και στην απόδοσή τους σε βραχυπρόθεσμη πρόβλεψη αλλά και σε αποτελέσματα με ημερήσια βήματα εκτίμησης. Μέσα από αυτές τις δοκιμές παρουσιάζονται αναλυτικά τα αποτελέσματα όλων των προβλέψεων, ώστε να αποσαφηνιστεί καλύτερα η απόδοση.

- Στο έβδομο κεφάλαιο αναδεικνύονται τα κυριότερα συμπεράσματα που προέκυψαν από τη συγκεκριμένη μελέτη. Επίσης, καταγράφονται περιορισμοί που εμφανίζονται στη συγκεκριμένη προσέγγιση, ενώ παρουσιάζονται πιθανές μελλοντικές προσεγγίσεις και προτάσεις εξέλιξης που σχετίζονται με το συγκεκριμένο αντικείμενο μελέτης για την ενίσχυση της αποτελεσματικότητας και για περισσότερη έρευνα.



## ΚΕΦΑΛΑΙΟ 2: ΑΙΟΛΙΚΗ ΕΝΕΡΓΕΙΑ

### 2.1 Ο ΡΟΛΟΣ ΤΩΝ ΑΝΑΝΕΩΣΙΜΩΝ ΠΗΓΩΝ ΕΝΕΡΓΕΙΑΣ ΣΤΗΝ ΕΝΕΡΓΕΙΑΚΗ ΜΕΤΑΒΑΣΗ

Διαχρονικά η ενεργειακή παραγωγή βασιζόταν κατά κύριο λόγο σε συμβατικές μονάδες, δηλαδή σε πηγές με μεγάλο οικολογικό αποτύπωμα που προέρχονται από ορυκτά καύσιμα και εξαντλούνται, αφού προϋπόθεση για τη μετατροπή της ενέργειας τους αποτελεί η καύση. Οι πιο συνηθισμένες μορφές συμβατικών πηγών ενέργειας είναι στερεάς μορφής (γαιάνθρακας, λιγνίτης, ανθρακίτης, τύρφη), υγρής (πετρέλαιο, μαζούτ, βενζίνη, κηροζίνη) και αέριας (φυσικό αέριο, υγραέριο). Η υπερεντατικοποίηση της εξόρυξης των ορυκτών καυσίμων, απόρροια των αυξανόμενων απαιτήσεων για παραγωγή ενέργειας καθώς και η κλιματική κρίση έχουν καταστήσει επίκαιρα τα ζητήματα της μελλοντικής εξάντλησής τους αλλά και της αναγκαιότητας περιορισμού τους.

Οι Ανανεώσιμες Πηγές Ενέργειας (ΑΠΕ) αποτελούν θεμέλιο λίθο για την ενεργειακή μετάβαση προς ένα πιο βιώσιμο και καθαρό μέλλον. Οι ΑΠΕ αποτελούν πηγές ενέργειας οι οποίες ανανεώνονται με φυσική διαδικασία και δεν εξαντλούνται ποτέ, ενώ η εκμετάλλευσή τους αφήνει το μικρότερο δυνατό οικολογικό αποτύπωμα, μειώνοντας τις εκπομπές αερίων του θερμοκηπίου. Με τα φαινόμενα της κλιματικής κρίσης και την ανάγκη για μείωση των εκπομπών αερίων θερμοκηπίου, ο ρόλος των ΑΠΕ γίνεται ολοένα και πιο κρίσιμος. Κρίνεται ως επιτακτική η αναγκαιότητα για μια ενεργειακή παραγωγή που δεν βασίζεται σε ορυκτά καύσιμα και μπορεί να παρέχει σταθερότητα και αξιοπιστία στο ενεργειακό σύστημα επιφέροντας συνολικά μακροπρόθεσμα οικονομικά και περιβαλλοντικά οφέλη και ενισχύοντας την ενεργειακή ανεξαρτησία. Στις ανανεώσιμες πηγές ενέργειας συγκαταλέγονται οι:

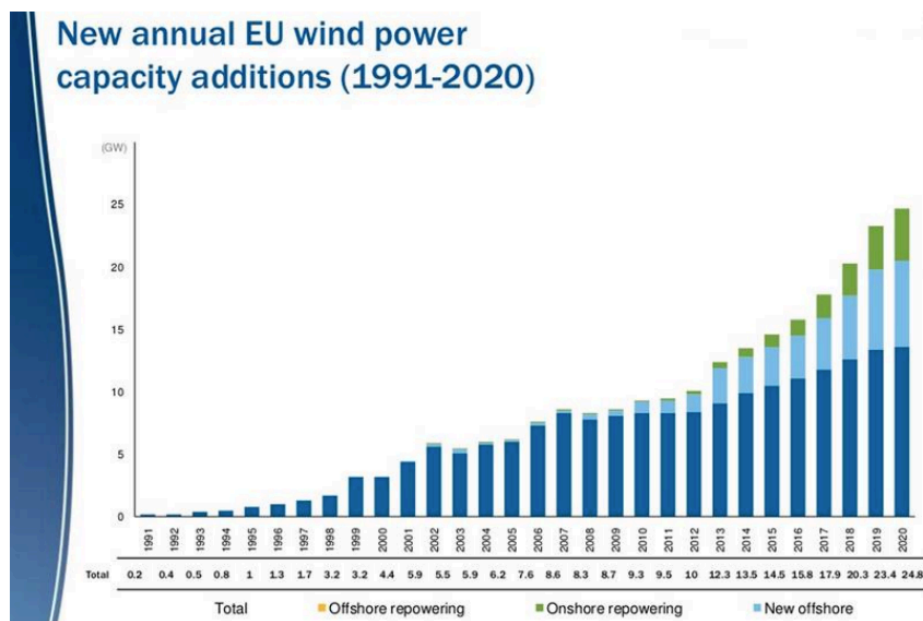
- Αιολική Ενέργεια: Η αιολική ενέργεια παράγεται από τη χρήση ανεμογεννητριών που μετατρέπουν την κινητική ενέργεια του αέρα σε ηλεκτρική ενέργεια.
- Ηλιακή Ενέργεια: Η ηλιακή ενέργεια παράγεται από την απευθείας χρήση της ηλιακής ακτινοβολίας ή από τη μετατροπή της σε θερμότητα μέσω ηλιακών συλλεκτών ή φωτοβολταϊκών πάνελ.
- Υδροηλεκτρική Ενέργεια: Η υδροηλεκτρική ενέργεια παράγεται από τη χρήση της κινητικής ενέργειας του νερού μέσω της εκμετάλλευσής φυσικών ροών όπως ποτάμια και καταρράκτες είτε από την αποθήκευσή τους σε φράγματα.
- Βιομάζα: Η βιομάζα προέρχεται από βιολογικά υλικά όπως ξύλο, απόβλητα, απορρίμματα τροφίμων και μπορεί να χρησιμοποιηθεί για την παραγωγή θερμότητας ή ηλεκτρικής ενέργειας.
- Γεωθερμική Ενέργεια: Η γεωθερμική ενέργεια προέρχεται από τη θερμότητα που εκπέμπεται από το εσωτερικό της Γης και μπορεί να χρησιμοποιηθεί για θέρμανση και ψύξη καθώς και την παραγωγή ηλεκτρικής ενέργειας.
- Ενέργεια κυψελών καυσίμου: Προέρχεται κατά κύριο λόγο από το υδρογόνο, το οποίο χρησιμοποιείται ως καύσιμο, και με την παρουσία καταλύτη μετατρέπεται η χημική του ενέργεια σε ηλεκτρική.

Από αρκετές μελέτες που βασίζονται σε μεθοδολογίες λήψης αποφάσεων πολλαπλών κριτηρίων με σκοπό την εύρεση της καλύτερης τεχνολογίας ανανεώσιμης ενέργειας για τον βιώσιμο σχεδιασμό υποδεικνύεται ότι η αιολική ενέργεια είναι η πιο κατάλληλη επιλογή [1].

## 2.2 ΣΤΟΙΧΕΙΑ ΑΙΟΛΙΚΗΣ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΕΝΣΩΜΑΤΩΣΗΣ

Σύμφωνα με τα δημοσιοποιημένα στοιχεία της Παγκόσμιας Ένωσης Αιολικής Ενέργειας (WWEA) από το 2021 η συνολική χωρητικότητα όλων των ανεμογεννητριών παγκοσμίως έχει ξεπεράσει τα 840 GW, αρκετή για να καλύψει πάνω από 7% της παγκόσμιας ζήτησης ενέργειας [2].

Ενδεικτικά, σύμφωνα με τα στοιχεία της Ευρωπαϊκής Ένωσης Αιολικής Ενέργειας (EWEA) στην Ευρώπη η αύξηση της αιολικής ισχύος παρουσίασε άνοδο περίπου κατά 300% σε ορίζοντα 10ετίας (2010-2020). Αναμένεται ότι η Ευρώπη θα εγκαταστήσει 129 GW νέων αιολικών πάρκων στην περίοδο 2023-2027, εκ των οποίων τα 27 κράτη-μέλη της Ευρωπαϊκής Ένωσης (27-Ε.Ε) θα εγκαταστήσουν τα 98 GW από αυτά. Οι περισσότερες από αυτές τις εγκαταστάσεις που θα πραγματοποιηθούν στον χρονικό ορίζοντα 2023-27 θα είναι εγκαταστάσεις εδάφους (onshore). Προβλέπεται ότι ο ρυθμός κατασκευής και διείσδυσης νέων αιολικών πάρκων στην Ευρωπαϊκή Ένωση για την περίοδο 2023-2027 θα κυμαίνεται κατά μέσο όρο στα 20 GW νέων αιολικών πάρκων ετησίως. Ωστόσο, για να επιτύχει τους στόχους της για το 2030, η Ε.Ε θα πρέπει να κατασκευάζει ετησίως πάνω από 30 GW νέων αιολικών πάρκων κατά μέσο όρο [3].



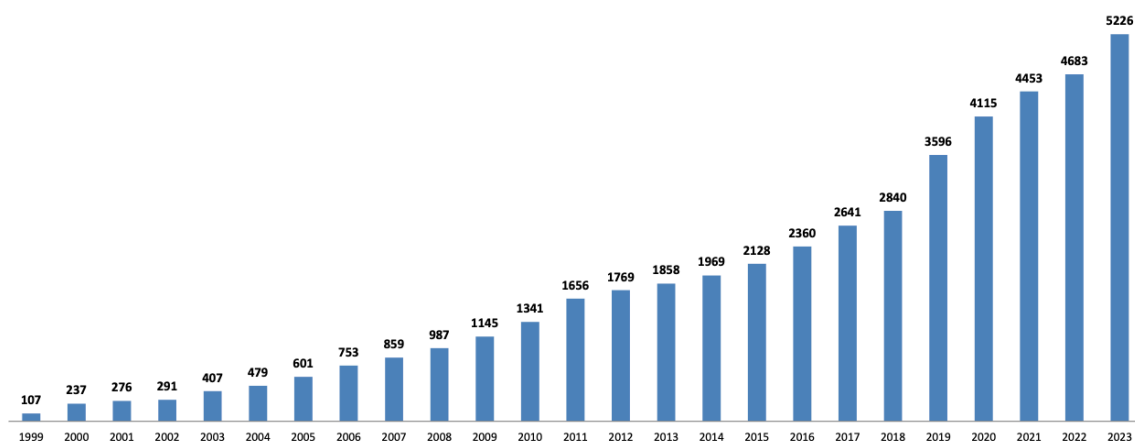
Εικόνα 1: Προσθήκη νέας αιολικής ισχύος στην Ευρώπη ανά έτος

Η παραγωγή αιολικής ενέργειας στην Ευρώπη αυξήθηκε σταθερά από 370 TWh το 2018 σε 489 TWh το 2022. Παράλληλα, η ζήτηση ηλεκτρικής ενέργειας μειώθηκε από 2.960 TWh το 2018 σε 2.830 TWh το 2022. Αυτό οφείλεται εν μέρει σε περιορισμούς λόγω της πανδημίας του COVID-19 το 2020 και του πολέμου στην Ουκρανία το 2022 [3]. Η Κίνα ως ο βασικός πρωταγωνιστής με την μεγαλύτερη διείσδυση αιολικής παραγωγής εισήγαγε 55,8 GW το 2021, καταφέροντας να ξεπεράσει το δικό της ρεκόρ του 2020, όταν εγκαταστάθηκαν 52 GW. Αυτό αντιστοιχεί σε αύξηση του 19,4%. Η Κίνα έχει πλέον εγκατεστημένα περισσότερα από 344 GW ανεμογεννητριών σε όλη την επικράτεια της. Οι Η.Π.Α αποτελούν τη δεύτερη μεγαλύτερη αγορά με εγκατεστημένη χωρητικότητα αιολικής ενέργειας να είναι τώρα περίπου στα 135 GW. Συνολικά, το 2022 παγκοσμίως, συνδέθηκαν στα δίκτυα ισχύος 77,6 GW νέας χωρητικότητας αιολικής ενέργειας, φέρνοντας τη συνολική εγκατεστημένη χωρητικότητα των ανεμογεννητριών σε 906 GW, με αύξηση 9% σε σχέση με το 2021. Οι πέντε κορυφαίες αγορές με νέες εγκαταστάσεις αιολικής παραγωγής για το 2022 ήταν:

- 1) Κίνα
- 2) ΗΠΑ
- 3) Βραζιλία
- 4) Γερμανία
- 5) Σουηδία

Συνολικά, αυτές οι χώρες αντιπροσώπευαν το 71% των παγκόσμιων εγκαταστάσεων του προηγούμενου έτους, με μείωση κατά 3,7% σε σχέση με το 2021. Αυτό οφείλεται κυρίως στις δύο μεγαλύτερες αγορές του κόσμου, την Κίνα και τις ΗΠΑ, που έχασαν συνολικά 5% του μεριδίου της αγοράς σε σύγκριση με το προηγούμενο έτος, το δεύτερο συνεχόμενο έτος που και οι δύο χώρες έχουν χάσει μερίδιο αγοράς [4].

Παρατηρείται αξιοσημείωτη προσπάθεια του αντίστοιχου μοντέλου παραγωγής ενέργειας και στην χώρα μας. Έτσι από τα 2.840 MW εγκατεστημένης ισχύος του έτους 2018, το επόμενο έτος εκτινάχθηκε στα 3.596 MW παρουσιάζοντας μεταβολή κατά 26,7% που αποτελεί το μεγαλύτερο ρυθμό αύξησης που έχει παρουσιαστεί στη χώρα μας. Η μεγάλη άνοδος συνεχίστηκε με σταδιακό ρυθμό μεταβολής και μέσα στην πενταετία έχει υπάρξει συνολική αύξηση 84% της συνολικής χωρητικότητας που είναι διαθέσιμη στο δίκτυο. Σύμφωνα με τα στοιχεία της Ελληνικής Επιστημονικής Ένωσης Αιολικής Ενέργειας το 2023 η Ελλάδα διαθέτει συνολική χωρητικότητα τροφοδοσίας στο δίκτυο 5.226 MW εκ των οποίων περισσότερα από τα μισά είναι κατανομημένα στην Στερεά Ελλάδα (2.293 MW) και την Πελοπόννησο (639 MW). Αξιοσημείωτο είναι ότι το 12.3% του ηλεκτρισμού προήλθε από αιολική ενέργεια που αντιστοιχεί στο 23.5% της συνολικής ηλεκτροπαραγωγής [5].



Εικόνα 2: Συνολική διαθέσιμη χωρητικότητα (MW) Ελληνικού δικτύου ανά έτος [5]

Κρίνεται επιτακτική η ανάγκη να αντιμετωπιστούν τα σημεία συμφόρησης καθώς και η βέλτιστη εισαγωγή ΑΠΕ στο δίκτυο ώστε να ελαχιστοποιηθούν φαινόμενα περικοπής και καθυστέρησης σύνδεσης. Οι πληροφορίες σχετικά με το δίκτυο για τη λειτουργία του συστήματος ηλεκτρικής ενέργειας είναι κεντρικές. Η δημοσίευση τακτικών δεδομένων σχετικά με τη συμφόρηση του δικτύου, την περικοπή και τη διαθέσιμη χωρητικότητα για νέες συνδέσεις και τη δυνατότητα παραγωγής ενέργειας από την αιολική παραγωγή κρίνεται ως επιτακτική.

## 2.3 Η ΒΑΣΙΚΗ ΛΕΙΤΟΥΡΓΙΑ ΤΩΝ ΑΝΕΜΟΓΕΝΝΗΤΡΙΩΝ

Η βασική λειτουργία των ανεμογεννητριών βασίζεται στην περιστροφική κίνηση του ρότορα για τον μετασχηματισμό της μηχανικής ενέργειας σε ηλεκτρική ενέργεια και την κατάλληλη

τροφοδοσία του δικτύου. Ένας από τους βασικούς παράγοντες για την περιστροφή είναι η ταχύτητα του ανέμου, η οποία σε ένα βαθμό καθορίζει την ισχύ του ρότορα. Η ένταση του ανέμου ουσιαστικά ρυθμίζει την ισχύ του τροχού, και επομένως η πρόβλεψη της ταχύτητας του ανέμου αποτελεί κρίσιμο θέμα για την βέλτιστη λειτουργία του συστήματος. Μέσω της πρόβλεψης της ταχύτητας και της κατεύθυνσης του ανέμου, τα συστήματα ελέγχου των ανεμογεννητριών μπορούν να προσαρμοστούν, ώστε να επιτευχθεί η αποδοτικότερη λειτουργία, μέσω της ρύθμισης των κατάλληλων γωνιών του ανεμοπτερυγίου (pitch control). Δυσμενείς καταστάσεις που προκύπτουν από την υψηλή ταχύτητα του ανέμου μπορούν να αποφευχθούν και το σύστημα μπορεί να τεθεί προληπτικά εκτός λειτουργίας, μετακινώντας τον ρότορα της ανεμογεννήτριας στις 90 μοίρες προς την κατεύθυνση του ανέμου.

Η ισχύς του ρότορα περιγράφεται από τον επόμενο τύπο:

$$P_A = \frac{1}{2} \rho C_p A v_w^3$$

Όπου :

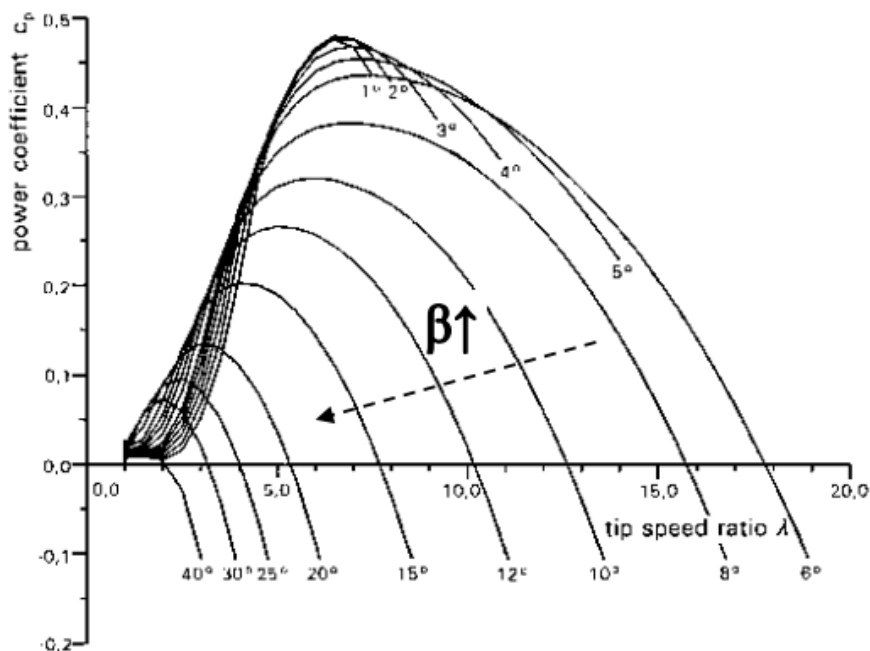
Πυκνότητα του αέρα:  $\rho$

Παράγοντας αεροδυναμικής ισχύος:  $C_p = C_p(\lambda, \beta) = C_p\left(\frac{R\omega_R}{v_w}\right)$  (όριο Betz:  $C_p < 0,59$ )

Επιφάνεια του ρότορα:  $A = \pi R^2$

Ταχύτητα του ανέμου:  $v_w$

Ο συντελεστής ισχύος  $C_p$  εξαρτάται από τις αεροδυναμικές χαρακτηριστικές της αιολικής εγκατάστασης. Η καμπύλη του συντελεστή ισχύος δείχνει τη σχέση αυτή:



Εικόνα 3: Καμπύλη αεροδυναμικού συντελεστή ισχύος συναρτήσει λόγου ταχύτητας ακροπτερυγίου  $C_p(\lambda)$  [6]

Η ακόλουθη εξίσωση μπορεί να χρησιμοποιηθεί για την εύρεση του παράγοντα αεροδυναμικής ισχύος μιας ανεμογεννήτριας :

$$C_p(\lambda, \beta) = 0.5 \left( \frac{116}{\lambda} - 0.4\beta - 5 \right) e^{-\frac{21}{\lambda}} \text{ με } \lambda = \frac{R\omega_R}{v_w}$$

Όπου:  $\omega_R$  η κυκλική ταχύτητα του ρότορα και  $R$  η ακτίνα του ρότορα [6].

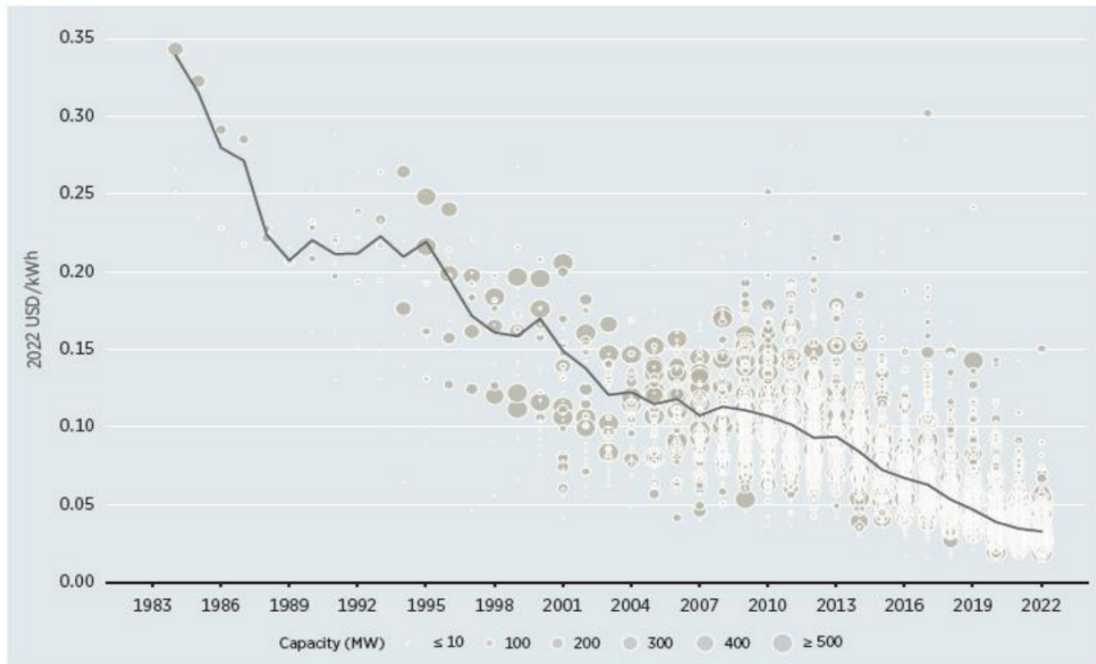
Σύμφωνα με αυτήν τη διατύπωση, ο παράγοντας ταχύτητας του ανέμου είναι πολύ σημαντικός και η πρόβλεψη του μπορεί να καθορίσει επίσης τη θεωρητική τιμή ενεργειακής παραγωγής της ανεμογεννήτριας. Ωστόσο, η πρόβλεψη παραγωγής σε πραγματικό χρόνο εμφανίζει διαφοροποίηση από τη θεωρητική προσέγγιση, η οποία εμφανίζεται με αυτό τον τρόπο συνήθως στις φυσικές μεθόδους. Οι παράγοντες των προκαθορισμένων τιμών των ανεμογεννητριών που στην πορεία του χρόνου μεταβάλλονται, σε συνδυασμό με την αβεβαιότητα των παραμέτρων και την ολοένα αυξανόμενη χαοτική συμπεριφορά των περιβαλλοντικών συνθηκών που καθορίζουν την λειτουργία των ΑΠΕ και μπορεί να εμφανίζουν διαφοροποίηση, καθιστούν την πρόβλεψη παραγόμενης ισχύος μια πιο απαιτητική διαδικασία.

Για ολόκληρο το αιολικό πάρκο (Α/Π), αυτή η σχέση εμφανίζει μεγαλύτερη πολυπλοκότητα, καθώς διάφορες ανεμογεννήτριες στο πάρκο χρησιμοποιούν πολλαπλές κατευθύνσεις ταχύτητας του ανέμου για να επιτύχουν την βέλτιστη παραγωγή ισχύος. Οποιοδήποτε σφάλμα στην πρόβλεψη της ταχύτητας του ανέμου θα οδηγήσει σε ένα μεγάλο σφάλμα στην εκτιμώμενη ισχύ. Επιπλέον, οι συνεχείς μετρήσεις και η προσαρμογή της εκάστοτε γεννήτριας καθιστούν τον συγκεκριμένο τρόπο εκτίμησης παραγωγής με την αποκλειστική χρήση δεδομένων ανέμου προβληματικό για έργα μεγάλης κλίμακας [7]. Συνολικά η πρόβλεψη ζήτησης και παραγωγής ΑΠΕ σε συνδυασμό με τα τεχνικοοικονομικά χαρακτηριστικά των μονάδων καθορίζουν τον ενεργειακό προγραμματισμό και τα συνολικά προγράμματα κατανομής. Επιπλέον, οι προβλέψεις είναι ιδιαίτερα σημαντικές για την εκτίμηση της ενεργειακής απόδοσης των Α/Π σε συγκεκριμένους χρονικούς ορίζοντες, ενώ καθορίζουν και την ενδεχόμενη δυνατότητα αποθήκευσης, καθώς και την ασφαλέστερη διαχείριση συντήρησης των συστημάτων. Οι μέθοδοι πρόβλεψης για την παραγωγή ηλεκτρικής ενέργειας από ανεμογεννήτριες μπορούν να κατηγοριοποιηθούν βάσει διαφόρων κριτηρίων, όπως ο χρονικός ορίζοντας πρόβλεψης (πολύ σύντομος, σύντομος, μεσοπρόθεσμος, μακροπρόθεσμος), και τα μοντέλα πρόβλεψης (φυσικά, στατιστικά, μηχανικής μάθησης).

## **2.4 ΟΙΚΟΝΟΜΙΚΗ ΒΙΩΣΙΜΟΤΗΤΑ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΕΣ ΒΕΛΤΙΩΣΕΙΣ**

Η συμβολή των ΑΠΕ στην ενεργειακή κρίση, στην σταδιακή απεξάρτηση από τα ορυκτά καύσιμα, στην ενεργειακή αυτονομία και στην μείωση εκπομπών CO<sub>2</sub>, καθιστούν τη βελτιστοποίηση των λειτουργιών τους επιτακτική για να πραγματοποιηθεί με μεγαλύτερο ρυθμό η ένταξή τους. Με βάση τα προαναφερθέντα στοιχεία αιολικής παραγωγής και ενσωμάτωσης, η αιολική ενέργεια τα τελευταία χρόνια αποτελεί μια από τις ταχύτερα αναδυόμενες τεχνολογίες ανανεώσιμων πηγών ενέργειας. Σύμφωνα με στοιχεία το επενδυτικό κόστος και το κόστος παραγωγής (LCOE) αιολικών πάρκων εμφανίζει σημαντικές βελτιώσεις παρά τις ανατιμήσεις υλικών που έχουν πραγματοποιηθεί. Το LCOE λαμβάνει υπόψη διάφορες παραμέτρους όπως το κόστος κατασκευής, τη διάρκεια ζωής του εγκατεστημένου εξοπλισμού, τα κόστη συντήρησης, την απόδοση της μονάδας και άλλες συναφείς μεταβλητές. Το LCOE εκφράζεται συνήθως σε κόστος ανά μονάδα ενέργειας και βοηθάει στη ανάλυση συνολικού κόστους που σχετίζεται με την παραγωγή ηλεκτρικής ενέργειας από μια συγκεκριμένη πηγή. Ειδικότερα, για την αιολική παραγωγή το συνολικό κόστος παραγωγής ήταν 710% πιο ακριβό από τη φθηνότερη λύση ορυκτών καυσίμων το 2010, ενώ οι τεχνολογικές βελτιώσεις οδήγησαν να είναι κατά 29% φθηνότερη επιλογή από την οικονομικότερη λύση ορυκτών καυσίμων το 2022, γεγονός που δείχνει την εξέλιξη που έχει υπάρξει τα τελευταία χρόνια στην τεχνολογία αιολικής παραγωγής [8]. Η οικονομική ανταγωνιστικότητα της συγκεκριμένης εναλλακτικής μεθόδου παραγωγής ενέργειας σε

συνδυασμό με την απόδοση παραγωγής της, την καθιστούν μια από τις πιο ανταγωνιστικές και ο στόχος της σταδιακής απεξάρτησης από την παραγωγή μέσω συμβατικών μονάδων ολοένα και πιο βιώσιμος.



Εικόνα 4: Μεταβολή LCOE χερσαίων αιολικών πάρκων 1983-2022 [8]



Εικόνα 5: Μεταβολή LCOE χερσαίων αιολικών στην Ελλάδα 2010-2022 [8]



# ΚΕΦΑΛΑΙΟ 3: ΠΡΟΒΛΕΨΕΙΣ ΑΙΟΛΙΚΗΣ ΕΝΕΡΓΕΙΑΣ

## 3.1 ΑΡΙΘΜΗΤΙΚΕΣ ΠΡΟΒΛΕΨΕΙΣ ΚΑΙΡΟΥ

Οι αριθμητικές προβλέψεις καιρού (Numerical Weather Predictions - NWP) χρησιμοποιούνται για να προβλέπουν τις καιρικές συνθήκες με τη χρήση ατμοσφαιρικών μοντέλων και υπολογιστικών τεχνικών. Καθώς η παραγωγή αιολικής ισχύος σχετίζεται άμεσα με τις καιρικές συνθήκες, το πρώτο βήμα για την πρόβλεψη της αιολικής ισχύος είναι η πρόβλεψη των μελλοντικών τιμών των καιρικών μεταβλητών τόσο στο επίπεδο του αιολικού πάρκου όσο και σε επίπεδο ανεμογεννητριών. Τα αποτελέσματα που προκύπτουν από τις αριθμητικές προβλέψεις καιρού είναι δεδομένα εισόδου για τα μοντέλα πρόβλεψης αιολικής ισχύος και αποτελούν σημείο εκκίνησης για την εξαγωγή προβλέψεων παραγωγής. Ένα μοντέλο αριθμητικής πρόβλεψης συνήθως αφορά υπολογιστικά προγράμματα που εξάγουν μετεωρολογικές πληροφορίες για μελλοντικές εποχές σε δεδομένες τοποθεσίες. Τα παγκόσμια μοντέλα αριθμητικών προβλέψεων καιρού αποτελούν τον πυρήνα των καιρικών προβλέψεων, αφού υποστηρίζουν στο μεγαλύτερο βαθμό την διαδικασία αφομοίωσης δεδομένων και παράγουν τις αρχικές και οριακές συνθήκες. Η χρήση, ως δεδομένα εισόδου, των πληροφοριών που εξάγονται από τα NWP's χρησιμοποιούνται κυρίως για μεσοπρόθεσμες ή μακροπρόθεσμες διαδικασίες πρόβλεψης, αφού χρειάζονται μεγάλους χρόνους εκτέλεσης και μεγάλους υπολογιστικούς πόρους [9]. Τα αντίστοιχα μοντέλα για άμεσες-βραχυπρόθεσμες προβλέψεις βασίζονται κατά κύριο λόγο σε στατιστικές προσεγγίσεις, ιδίως σε τεχνητά νευρωνικά δίκτυα (ANN), λόγω του χρόνου που απαιτείται για τη λειτουργία των αριθμητικών προβλέψεων καιρού (NWP). Ωστόσο, υπάρχουν και βραχυπρόθεσμα μοντέλα πρόβλεψης παραγωγής που βασίζονται σε NWP και παρουσιάζουν βελτιωμένα αποτελέσματα [9], [10], [11]. Για την πρόβλεψη της αιολικής ισχύος ενός αιολικού πάρκου βασιζόμενες σε NWP's, συνήθως, χρησιμοποιούνται τα αποτελέσματα ενός μοντέλου αριθμητικών προβλέψεων καιρού που παρέχονται από τις μετεωρολογικές υπηρεσίες της κάθε χώρας. Το πιο ευρέως διαδεδομένο μέσης κλίμακας μοντέλο είναι το MM5, ενώ το πιο γνωστό μικρής κλίμακας μοντέλο είναι το WAsP [12]. Τα μοντέλα των αριθμητικών προβλέψεων καιρού που είναι σε χρήση είναι το HIRLAM, το SKIRON, το ALADIN, GFS, WRF, NOGAPS και το ECMWF [13], [14].

Ορισμένα από τα πιο σημαντικά αριθμητικά δεδομένα πρόβλεψης καιρού που χρησιμοποιούνται ως είσοδοι στα εκπαιδευμένα μοντέλα για την εξαγωγή των επιθυμητών αποτελεσμάτων πρόβλεψης αιολικής παραγωγής είναι :

### 3.1.1 Ταχύτητα Ανέμου

Ο πιο σημαντικός παράγοντας είναι η ταχύτητα ανέμου. Με τον όρο ταχύτητα του ανέμου εννοούμε το μέτρο της συνισταμένης που προκύπτει από τις δύο κάθετες συνιστώσες του ανέμου. Πρέπει να σημειωθεί πως οι ταχύτητες αέρα που λειτουργεί μια ανεμογεννήτρια είναι τα 3,5-25m/s. Τα 3,5 m/s χαρακτηρίζονται ως ταχύτητα ένταξης (cut-in), ενώ τα 25m/s χαρακτηρίζονται ως ταχύτητα αποκοπής (cut-off) και πάνω από αυτή τη ταχύτητα ανέμου η ανεμογεννήτρια παύει να λειτουργεί. Συνεπώς η ωφέλιμη ταχύτητα ανέμου για μια ανεμογεννήτρια είναι τα 3,5 – 25 m/s, ενώ οι τυπικές ανεμογεννήτριες φτάνουν στην μέγιστη της απόδοση πάνω από τα 14m/s. Άρα η βέλτιστη παραγωγή επιτυγχάνεται στα 14 – 25 m/s ταχύτητας αέρα.

### 3.1.2 Κατεύθυνση Ανέμου

Ως κατεύθυνση του ανέμου ορίζεται η γωνία μεταξύ των δύο συνιστωσών του ανέμου η οποία μπορεί να πάρει όλο το εύρος τιμών από  $0^\circ$  έως  $360^\circ$ . Η επίδραση της κατεύθυνσης δεν είναι τόσο μεγάλη όσο αυτή της ταχύτητας, αλλά διαδραματίζει κι αυτή σημαντικό ρόλο στην παραγωγή και κατ' επέκταση στην πρόβλεψη, εφόσον η παραγωγή των ανεμογεννητριών διαφοροποιείται ανάλογα με την γωνία που σχηματίζουν με τον άνεμο. Επιπλέον, αξίζει να σημειωθεί πως βόρειες κατευθύνσεις ανέμου χαρακτηρίζονται πιο στρωτές με συνέπεια ο αέρας να έρχεται ομαλότερα στους έλικες της ανεμογεννήτριας και μπορούν να λειτουργούν με καλύτερο τρόπο σε σύγκριση με τις νότιες κατευθύνσεις ανέμων που εκεί παρατηρούνται πολλοί στροβιλισμοί (turbulences), με αποτέλεσμα ο άνεμος να χτυπάει με ακανόνιστο τρόπο την ανεμογεννήτρια.

### 3.1.3 Πυκνότητα Αέρα

Ένα χαρακτηριστικό το οποίο έχει επίσης σημαντική συμβολή στην παραγωγή αιολικής ενέργειας είναι και η πυκνότητα του αέρα. Η ενέργεια η οποία είναι διαθέσιμη από τον άνεμο είναι ευθέως ανάλογη με την πυκνότητα του αέρα. Καθώς η πυκνότητα του αέρα αυξάνεται η διαθέσιμη ενέργεια αυξάνεται κι αυτή. Σ' αυτό το σημείο υπεισέρχονται και δύο παράγοντες που επηρεάζουν την πυκνότητα του αέρα: η πίεση και η θερμοκρασία. Η πυκνότητα του αέρα αυξάνεται με την αύξηση της πίεσης καθώς επίσης και με την πτώση της θερμοκρασίας. Και οι δύο αυτοί παράγοντες μειώνονται καθώς αυξάνεται το υψόμετρο. Κατά συνέπεια γίνεται αντιληπτό ότι προκειμένου να υπολογιστεί η παραγωγή της αιολικής ενέργειας πρέπει να ληφθούν υπόψη τόσο η θερμοκρασία όσο και η πίεση του αέρα έτσι ώστε να υπάρξει ακόμα μεγαλύτερη αξιοπιστία στην πρόβλεψη.

## 3.2 Η ΣΗΜΑΣΙΑ ΤΗΣ ΠΡΟΒΛΕΨΗΣ

Το αυξανόμενο διεθνές ενδιαφέρον για τις ανανεώσιμες πηγές ενέργειας και η διεύρυνση της ενσωμάτωσής τους στο ηλεκτρικό δίκτυο έχει κεντρίσει το ενδιαφέρον πολλών ερευνητών σε παγκόσμιο επίπεδο. Η αξία της εξαγωγής ακριβών προβλέψεων ενεργειακής παραγωγής αφορά την βελτιστοποίηση συστημάτων διαχείρισης ενέργειας και αποτελεσμάτων σχετιζόμενων με διαδικασίες που αφορούν άμεσα την αγορά ενέργειας. Δημοφιλείς εφαρμογές έξυπνων συστημάτων ενέργειας που βασίζονται σε μεθόδους πρόβλεψης και περιλαμβάνουν ζητήματα προγνωστικού ελέγχου φορτίου, προγνωστικού ελέγχου της ενεργειακής παραγωγής και κατανομής ΑΠΕ, την ασφάλεια του δικτύου, την εκτίμηση ρίσκου, την προσδοκώμενη τιμολόγηση ενέργειας, τον εντοπισμό παρεμβολών στην ποιότητα της ισχύος, τον εντοπισμό βλαβών σε συστήματα ισχύος καθώς και την προβλεπτική συντήρηση έχουν άμεση συσχέτιση με το συγκεκριμένο αντικείμενο μελέτης. Η πρόβλεψη παραγωγής αφορά τόσο τους συμμετέχοντες στην απελευθερωμένη αγορά ενέργειας, όσο και τους λειτουργούς του συστήματος, λόγω της επιδραστικότητας στις αποφάσεις που αφορούν τη λειτουργία, τη διαχείριση του συνολικού δικτύου και την ενσωμάτωση μονάδων ΑΠΕ. Η ακριβής πρόβλεψη της παραγωγής ανανεώσιμων πηγών είναι ζωτικής σημασίας για τη διασφάλιση της αξιοπιστίας της μόνιμης λειτουργίας του δικτύου, καθώς και για τη ρύθμιση του κόστους της ενεργειακής αγοράς και των ενεργειακών συστημάτων. Όχι μόνο οι εγκαταστάσεις παραγωγής ενέργειας αλλά και οι χειριστές δικτύου, οι έμποροι ενέργειας και οι καταναλωτές επωφελούνται από την αποδοτικότητα των προβλέψεων παραγωγής. Επιπλέον, σε επίπεδο αγοράς η εκτίμηση της συνεισφοράς της αιολικής ενέργειας στη συνολικά παραγόμενη ενέργεια καθίσταται σημαντική για τον προσδιορισμό των ημερήσιων και ωριαίων τιμών, καθώς οι πιθανές διακυμάνσεις επηρεάζουν τις τιμές εκκαθάρισης ενεργειακών και λειτουργικών αποθεμάτων.



Ειδικότερα, η πρόβλεψη ενεργειακής παραγωγής παρουσιάζει ιδιαίτερη χρησιμότητα για μία σειρά φορέων του ενεργειακού τοπίου, οι οποίοι εξασφαλίζουν την εύρυθμη λειτουργία του συστήματος. Πιο αναλυτικά, τα τελευταία χρόνια καταβάλλεται προσπάθεια βελτίωσης και ανάπτυξης του Ελληνικού Συστήματος Μεταφοράς (ΕΣΜ) με τη δημιουργία του Χρηματιστηρίου Ενέργειας (Χ.Ε) και την υιοθέτηση του ευρωπαϊκού μοντέλου στόχου (Target Model) σε μία απόπειρα ευθυγράμμισης της ελληνικής και της ευρωπαϊκής αγοράς ενέργειας. Η ακριβής πρόβλεψη ηλεκτρικής παραγωγής από Α/Π είναι απαραίτητη για τους παραγωγούς και τους εμπόρους ηλεκτρικής ενέργειας, ώστε να συμμετέχουν αποτελεσματικά στον ημερήσιο προγραμματισμό. Αυξάνοντας την ακρίβεια πρόβλεψης, οι παραγωγοί και οι έμποροι είναι σε θέση να μειώσουν τα κόστη εξισορρόπησης και να αυξήσουν τα κέρδη τους. Με το «Μοντέλου Στόχου» επιδιώκεται η βελτιστοποίηση της χρήσης της δυναμικότητας του συστήματος μεταφοράς μέσα από συντονισμένες πρακτικές των διαχειριστών, για την επίτευξη αξιόπιστων τιμών και ρευστότητας στον καταμερισμό της δυναμικότητας των διασυνδέσεων για την αγορά της επόμενης ημέρας, την αποτελεσματική λειτουργία των προθεσμιακών αγορών και τον αποτελεσματικό σχεδιασμό των ενδοημερήσιων αγορών. Σύμφωνα με τον ν. 4512/2018, ορίζονται οι ακόλουθες αγορές ενεργειακών προϊόντων χονδρικής, κατά την έννοια του Κανονισμού (ΕΕ) αριθ. 1227/2011:

- *Ενεργειακή χρηματοπιστωτική αγορά - Forward Market*

Η Ενεργειακή Χρηματοπιστωτική Αγορά, γνωστή και ως Forward Market, είναι η αγορά στην οποία πραγματοποιούνται συναλλαγές για την αγοραπωλησία ηλεκτρικής ενέργειας σε μελλοντικές περιόδους. Οι συμφωνίες αυτές καθορίζονται εκ των προτέρων όσον αφορά την τιμή, την ποσότητα και τη χρονική στιγμή παράδοσης. Αποσκοπεί στην πρότερη κάλυψη πιθανών κινδύνων σχετιζόμενων με τις μεταβολές των τιμών ενέργειας.

- *Αγορά επόμενης ημέρας - (Day-ahead Market)*

Η Αγορά Επόμενης Ημέρας, ή Day-Ahead Market, είναι η αγορά όπου πραγματοποιούνται συναλλαγές ενέργειας για παράδοση την επόμενη ημέρα. Οι συμμετέχοντες υποβάλλουν προσφορές αγοράς ή πώλησης βάσει των προβλεπόμενων αναγκών ή της εκτιμώμενης παραγωγής τους. Η τιμή καθορίζεται από την ισορροπία μεταξύ προσφοράς και ζήτησης και η εξαγωγή ποιοτικών προβλέψεων αποτελεί κρίσιμο τομέα για την επιτυχημένη υποβολή προσφορών. Στην Ελληνική ημερήσια αγορά προτείνεται η χρήση απλών ωριαίων δηλώσεων (simple hourly orders), μπλοκ δηλώσεων (block orders) και σύνθετων δηλώσεων (complex orders) τόσο για τις προσφορές παραγωγής, όσο και για τις δηλώσεις φορτίου.

- *Ενδοημερήσια αγορά - Intra-day Market*

Η Ενδοημερήσια Αγορά, ή Intra-day Market, επιτρέπει στους συμμετέχοντες να προβαίνουν σε συναλλαγές ηλεκτρικής ενέργειας καθ' όλη τη διάρκεια της ημέρας παράδοσης, δίνοντάς τους τη δυνατότητα να προσαρμόζονται σε τροποποιήσεις της ζήτησης ή της προσφοράς, να διορθώνουν προβλέψεις ή να αντιμετωπίζουν αναπάντεχες αλλαγές στην παραγωγή.

- *Αγορά Εξισορρόπησης - Balancing Market*

Αγορά Εξισορρόπησης, ή Balancing Market, είναι η αγορά που λειτουργεί με σκοπό τη διατήρηση της ισορροπίας μεταξύ προσφοράς και ζήτησης σε πραγματικό χρόνο εντός του ηλεκτρικού συστήματος. Σε αυτή την αγορά, οι προσφορές για παροχή ή απορρόφηση ισχύος προέρχονται από παραγωγούς ή καταναλωτές που έχουν τη δυνατότητα να ανταποκριθούν σε στιγμιαίες ανάγκες του συστήματος για εξισορρόπηση.

Επιπλέον, η πρόβλεψη παραγωγής αποτελεί διαδικασία κρίσιμου χαρακτήρα για τον Ανεξάρτητο Διαχειριστή Μεταφοράς Ηλεκτρικής Ενέργειας (ΑΔΜΗΕ), τόσο σε μακροπρόθεσμο ορίζοντα, ώστε να συντελείται με επιτυχία ο απαραίτητος ενεργειακός προγραμματισμός, όσο και σε βραχυπρόθεσμο επίπεδο, για να διασφαλίζεται η αδιάλειπτη επάρκεια και η ορθή λειτουργία του δικτύου. Ο ρόλος των μεθόδων πρόβλεψης παραγωγής

ενέργειας είναι ιδιαίτερα σημαντικός και δίνεται μεγάλη βαρύτητα και σημασία, έχοντας άμεσο αντίκτυπο και σε οικονομικό επίπεδο για όλους τους εμπλεκόμενους. Πέραν των χρεώσεων που επιβάλλονται σε περίπτωση σημαντικής απόκλισης μεταξύ πραγματικής και προβλεπόμενης ενέργειας, μία αναντιστοιχία στην παραγωγή και ζήτηση είναι πιθανό να οδηγήσει είτε σε περιορισμό παραγωγής ΑΠΕ από τον ΑΔΜΗΕ (σε περίπτωση απρόσμενα υψηλότερης παραγωγής ΑΠΕ), άρα και σε διαφυγόντα κέρδη του παραγωγού, είτε στην ανάγκη άμεσης λειτουργίας θερμικών μονάδων και εφεδρειών (σε περίπτωση αισιόδοξης πρόβλεψης παραγωγής από ΑΠΕ που δεν ικανοποιήθηκε), άρα και σε αυξημένη τιμή εκκαθάρισης της αγοράς. Συνεπώς, όσο πιο ορθή είναι μια πρόβλεψη παραγωγής για την συμμετοχή μιας εταιρείας σε κάποια από τις προαναφερθείσες αγορές, τόσο μεγιστοποιείται το κέρδος, δεν υπάρχει κίνδυνος εφαρμογής προστίμου και δεν χρειάζεται περαιτέρω διόρθωση στην ενδοημερήσια αγορά και διαφυλάσσεται η σταθερότητα. Όλες οι εταιρείες λοιπόν για την καλύτερη και βέλτιστη συμμετοχή τους στο Χρηματιστήριο Ενέργειας προσπαθούν να αναπτύξουν ιδανικά και αξιόπιστα μοντέλα εξαγωγής βέλτιστων προβλέψεων ενεργειακής παραγωγής για τη διασφάλιση και τη μεγιστοποίηση των κερδών τους.

### 3.3 ΧΡΟΝΙΚΟΙ ΟΡΙΖΟΝΤΕΣ ΠΡΟΒΛΕΨΗΣ

Οι μέθοδοι πρόβλεψης παραγόμενης ισχύος από ΑΠΕ χωρίζονται σε διάφορες κατηγορίες ανάλογα με τον εξαγόμενο χρονικό ορίζοντα. Τα περισσότερα από τα συστήματα εμπορικής και ερευνητικής πρόβλεψης χρησιμοποιούνται για ορίζοντα χρόνου που κυμαίνεται έως και λίγους μήνες [7]. Η κατηγοριοποίηση των μεθόδων πρόβλεψης ανά χρονική κλίμακα δεν είναι σαφώς καθορισμένη επειδή αρκετοί συγγραφείς έχουν προτείνει διαφορετικά χρονικά όρια. Γενικά, η πρόβλεψη της παραγωγής αιολικής ενέργειας μπορεί να χωριστεί σε 4 βασικές κατηγορίες:

#### 3.3.1 Πολύ βραχυπρόθεσμη πρόβλεψη:

Παρέχουν προβλέψεις από λίγα δευτερόλεπτα έως αρκετά λεπτά μπροστά συνήθως έως και 30 λεπτά. Αυτές οι μέθοδοι χρησιμοποιούνται για τη διαχείριση ενέργειας σε έξυπνα δίκτυα, για παράδειγμα, για την άμεση και αποτελεσματική αντιστάθμιση ενεργού και άεργου ισχύος. Οι προβλέψεις αυτές συνεισφέρουν στη διαχείριση βοηθητικών υπηρεσιών του συστήματος ισχύος, όπως και στη δέσμευση μονάδων και την οικονομική κατανομή του φορτίου καθώς και σε δράσεις εκκαθάρισης και σε διαπραγματεύσεις για την ενδοημερήσια αγορά ενέργειας [15], [16], [17]. Γενικά, σύμφωνα με άλλες μελέτες και προτάσεις το εύρος του συγκεκριμένου χρονικού ορίζοντα μπορεί να είναι έως και μερικές ώρες, αλλά δεν υπάρχει ομοφωνία για τον αριθμό των ωρών. Σε διάφορες μελέτες προτείνεται ένα μέγιστο όριο καθορισμού των πολύ βραχυπρόθεσμων προβλέψεων από 4 έως 9 ώρες και συχνά ταυτίζονται με χρονικά πλαίσια βραχυπρόθεσμων προβλέψεων.

#### 3.3.2 Βραχυπρόθεσμη πρόβλεψη:

Παρέχουν προβλέψεις συνήθως από 30 λεπτά μπροστά έως και μερικές ώρες (συνήθως 6). Τέτοιες μέθοδοι είναι κατάλληλες για εφαρμογές σε επίπεδο έξυπνου δικτύου, εξασφαλίζοντας τη δέσμευση, το προγραμματισμό, τον έλεγχο των ανεμογεννητριών για την βέλτιστη αποστολή της ισχύος και το σχεδιασμό διανομής του φορτίου. Είναι επίσης χρήσιμες για τη διαδικασία της δυναμικής τιμολόγησης και την προβλεπτική συντήρησης [16], [18]. Στο συγκεκριμένο ορίζοντα πρόβλεψης αρκετές μελέτες που έχουν πραγματοποιηθεί οριοθετούν την βραχυπρόθεσμη πρόβλεψη και για χρονικό ορίζοντα 24 και 48 ωρών για την πραγματοποίηση διαδικασιών που αφορούν τη διαπραγμάτευση στην αγορά της επόμενης ημέρας. Για παράδειγμα, στην Ιβηρική Αγορά Ενέργειας (MIBEL), οι προσφορές πώλησης ηλεκτρικής ενέργειας για την επόμενη ημέρα πρέπει να υποβάλλονται

πριν τις 10:00 π.μ. και ένας χρονικός ορίζοντας πρόβλεψης 38 ωρών, που αναφέρεται ως βραχυπρόθεσμη πρόβλεψη καλύπτει ολόκληρη την επόμενη ημέρα [19].

### 3.3.3 Μεσοπρόθεσμη πρόβλεψη:

Παρέχουν προβλέψεις από κάποιες ώρες έως και ορισμένες ημέρες μπροστά (συνήθως 6 ώρες έως και 3 ημέρες). Αυτές οι μέθοδοι χρησιμοποιούνται συνήθως για τον βέλτιστο σχεδιασμό του ΣΗΕ, για διαδικασίες παροχής ασφαλείας στο δίκτυο (threat models), ενώ παρέχουν δεδομένα για την πιθανή ένταξη μονάδων παραγωγής και την προβλεπτική συντήρηση του δικτύου. Καθώς ο χρονικός ορίζοντας της πρόβλεψης αυξάνεται, το ίδιο ισχύει και για τα σφάλματα πρόβλεψης [16][18].

### 3.3.4 Μακροπρόθεσμη πρόβλεψη:

Ορισμένοι έχουν κατατάξει χρονικά διαστήματα που κυμαίνονται από ορισμένες ημέρες έως και αρκετούς μήνες μπροστά ως προγνωστικά μοντέλα μακροπρόθεσμης διάρκειας [18]. Η πλειονότητα κατηγοριοποιεί τα συγκεκριμένα μακροπρόθεσμα μοντέλα πρόβλεψης από ορισμένους μήνες έως και περισσότερο από ένα έτος, τα οποία βασίζονται σε αριθμητικά δεδομένα πρόβλεψης καιρού (NWP). Ο συγκεκριμένος ορίζοντας πρόβλεψης έχει ιδιαίτερη χρησιμότητα για τον μακροπρόθεσμο σχεδιασμό παραγωγής, για τον προγραμματισμό διαχείρισης εφεδρειών, τη διανομή ηλεκτρικής ενέργειας, καθώς και για την ανάλυση των εποχιακών τάσεων. Ωστόσο, αυτά τα μοντέλα έχουν μειωμένη ακρίβεια επειδή οι καιρικές διακυμάνσεις δεν μπορούν να προβλεφθούν με ιδιαίτερη επιτυχία χρησιμοποιώντας τόσο μακροπρόθεσμους ορίζοντες [16][20]. Επιπλέον τα μοντέλα μακροπρόθεσμης πρόβλεψης βασίζονται αποκλειστικά σε NWP για την εξαγωγή αποτελεσμάτων.

Στη συνέχεια με βάση τη βιβλιογραφία από το 2010 έως το 2014 παρουσιάζεται η χρονική ταξινόμηση που προτείνεται για την κατάταξη μαζί με τις εφαρμογές [21].

<b>ΧΡΟΝΙΚΗ ΤΑΞΙΝΟΜΗΣΗ</b>	<b>ΧΡΟΝΙΚΟΣ ΟΡΙΖΟΝΤΑΣ</b>	<b>ΕΦΑΡΜΟΓΕΣ</b>
Πολύ Βραχυπρόθεσμη	Δευτερόλεπτα έως 1 ώρα	<ul style="list-style-type: none"> <li>• Καθαρισμός αγοράς ηλεκτρισμού</li> <li>• Παρακολούθηση φορτίου</li> <li>• Έλεγχος συχνότητας συστήματος ηλεκτρικής ενέργειας</li> <li>• Λειτουργίες πραγματικού χρόνου δικτύου</li> <li>• Ρυθμιστικές ενέργειες</li> <li>• Έλεγχος ανεμογεννήτριας</li> </ul>
Βραχυπρόθεσμη	1-6 ώρες	<ul style="list-style-type: none"> <li>• Σχεδιασμός οικονομικής διανομής φορτίου</li> <li>• Αποφάσεις αύξησης/μείωσης φορτίου</li> <li>• Λειτουργική ασφάλεια στην αγορά ηλεκτρισμού</li> <li>• Λειτουργίες δικτύου πραγματικού χρόνου</li> <li>• Ρυθμιστικές ενέργειες</li> </ul>
Μεσοπρόθεσμη	6-72 ώρες	<ul style="list-style-type: none"> <li>• Αγορά ηλεκτρικής ενέργειας προηγούμενης ημέρας</li> <li>• Οικονομική διανομή</li> <li>• Εμπόριο ηλεκτρικής ενέργειας</li> </ul>

		<ul style="list-style-type: none"> <li>• Αποφάσεις λειτουργίας/διακοπής γεννήτριας</li> <li>• Προγραμματισμός συντήρησης για βέλτιστο κόστος λειτουργίας</li> <li>• Διαχείριση λειτουργίας</li> <li>• Λειτουργική ασφάλεια στην αγορά ηλεκτρικής ενέργειας</li> <li>• Λειτουργική ασφάλεια στην αγορά ηλεκτρικής ενέργειας επόμενης ημέρας</li> <li>• Αποφάσεις απαιτήσεων αποθεμάτων</li> <li>• Δέσμευση μονάδων</li> </ul>
Μακροπρόθεσμη	72 ώρες-αρκετούς μήνες	<ul style="list-style-type: none"> <li>• Μελέτη σκοπιμότητας αιολικού πάρκου</li> <li>• Σχεδιασμός συντήρησης</li> <li>• Διαχείριση λειτουργίας</li> <li>• Βέλτιστο κόστος λειτουργίας</li> <li>• Σχεδιασμός ενεργειακού συστήματος</li> <li>• Αποφάσεις για τις απαιτήσεις αποθεματικών</li> <li>• Αποφάσεις δέσμευσης μονάδων</li> <li>• Σχεδιασμός κατανομής αιολικής ενέργειας Α/Π</li> </ul>

Πίνακας 1: Χρονική ταξινόμηση προβλέψεων αιολικής παραγωγής [21]

## ΚΕΦΑΛΑΙΟ 4: ΜΟΝΤΕΛΑ ΠΡΟΒΛΕΨΗΣ

### 4.1 ΜΟΝΤΕΛΑ ΣΥΝΕΧΙΖΟΜΕΝΗΣ ΠΡΟΒΛΕΨΗΣ (PERSISTENCE MODELS)

Ο απλούστερος τρόπος για να επιτευχθεί η επιθυμητή πρόβλεψη είναι τα μοντέλα συνεχιζόμενης πρόβλεψης (persistence models). Η συγκεκριμένη μέθοδος βασίζεται στην απλή υπόθεση ότι η προβλεπόμενη τιμή στο χρόνο  $t + x$  θα ισούται με εκείνη τη χρονική στιγμή  $t$ . Αυτή η μέθοδος αναπτύχθηκε ως ένα εργαλείο σύγκρισης και ως σημείο αφετηρίας για να συμπληρώσει τα μοντέλα πρόβλεψης καιρού NWP. Η συγκεκριμένη μέθοδος βασίζεται στη λογική των αμελητέων μετεωρολογικών αλλαγών σε σύντομα χρονικά διαστήματα και επομένως η επόμενη πρόβλεψη της παραγωγής θεωρείται πάντα ίση με τη μετρημένη παραγωγή της τρέχουσας χρονικής στιγμής. Αυτή η απλοποιημένη μέθοδος στην οποία δεν εισέρχεται πολυπλοκότητα και η ταχύτητα υπολογισμού αποτελεί πλεονέκτημα μπορεί να είναι αποτελεσματικότερη από ένα μοντέλο που βασίζεται σε NWP σε ορισμένες πολύ σύντομες προβλέψεις (έως και λίγα λεπτά) [22]. Ωστόσο, η ακρίβεια αυτού του μοντέλου υποβαθμίζεται γρήγορα με την αύξηση του χρόνου πρόβλεψης και η αξιοπιστία του αμφισβητείται ιδίως σε περιοχές στις οποίες υπάρχουν συχνές και έντονες ριπές ανέμου χωρίς ενσωματωμένο έλεγχο λειτουργίας [7].

### 4.2 ΦΥΣΙΚΕΣ ΜΕΘΟΔΟΙ

Η ανάπτυξη ενός φυσικού μοντέλου πρόβλεψης της αιολικής ισχύος συνεπάγεται στην χρήση εξόδου δεδομένων κάποιου μετεωρολογικού μοντέλου. Για την πρόβλεψη της αιολικής ισχύος, συνήθως, χρησιμοποιούνται τα αποτελέσματα ενός μοντέλου αριθμητικών προβλέψεων καιρού που παρέχονται από τις μετεωρολογικές υπηρεσίες της κάθε χώρας. Οι φυσικές μέθοδοι βασίζονται σε αριθμητικά μοντέλα πρόβλεψης καιρού (NWP) για να εκτιμήσουν την παραγωγή ηλεκτρικής ενέργειας κυρίως μέσα από χαρακτηριστικά του ανέμου. Εκτός από τα προαναφερθέντα δεδομένα που προκύπτουν από την επιστήμη της μετεωρολογίας η χρήση επιπρόσθετων γεωγραφικών πληροφοριών, όπως θερμοκρασία, πίεση, χαρακτηριστικά επιφάνειας, εμπόδια και άλλες φυσικές μετρήσεις απαιτούνται ως είσοδοι για να εκτιμηθεί η παραγωγή. Οι φυσικές μέθοδοι απαιτούν πολλούς υπολογισμούς και αντιμετωπίζουν προβλήματα συνήθως με τον καθορισμό των τιμών πολλών παραμέτρων με συχνή υψηλή πολυπλοκότητα και αβεβαιότητα στην προσέγγισή τους. Η μοντελοποίηση του ανέμου πάνω από ανώμαλα εδάφη χωρίζεται συνήθως σε δύο τύπους: τα δυναμικά μοντέλα (προγνωστικά) και τα κινηματικά μοντέλα (διαγνωστικά) [23] [24]. Επιπλέον, συνήθως εισάγονται και οι μεταβλητές του εμπλεκόμενου εξοπλισμού, οι οποίες όμως εξαρτώνται σε μεγάλο βαθμό από τα τεχνικά χαρακτηριστικά των κατασκευαστών. Αξίζει να επισημανθεί ότι τα δεδομένα αυτά είναι πιθανόν να μην είναι διαθέσιμα, να είναι μειωμένης ακρίβειας, ενώ συνήθως προκύπτουν σε εργαστηριακό περιβάλλον, εισάγοντας έτσι τον ανθρώπινο εμπειρικό παράγοντα και δυσχεραίνοντας την αποτελεσματικότητα της μεθόδου [25].

Γενικά, αυτή η μέθοδος παρέχει αποτελεσματικότητα, αλλά υπάρχουν ορισμένα βασικά προβλήματα, όπως ο μεγάλος όγκος των δεδομένων εισόδου. Τα δεδομένα εισόδου κάθε στιγμής είναι διαφορετικά και η αναποτελεσματικότητα χειρισμού απροσδόκων σφαλμάτων κατά την πρόβλεψη τους δεν τις καθιστούν κατάλληλες για τον προβλεπτικό ορίζοντα μικρού χρονικού διαστήματος του αιολικού πάρκου, όπου η ακρίβεια της ταχύτητας του ανέμου είναι κρίσιμη για τον προβλεπτικό υπολογισμό της παρεχόμενης ισχύος του δικτύου. Συνεπώς, αν και μπορεί να παρουσιάζουν ικανοποιητικά αποτελέσματα πρόβλεψης,

η εξειδίκευση που χρειάζονται σε συνδυασμό με την έλλειψη ευελιξίας τους και την εισαγωγή πολλών και σύνθετων δεδομένων για τον υπολογισμό τους, αποτελούν ορισμένα από τα μειονεκτήματά τους [24], [26], [27]. Ως εκ τούτου, η απόδοση των φυσικών μοντέλων είναι συχνά ικανοποιητική συνήθως για μεγάλους χρονικούς ορίζοντες πρόβλεψης (περισσότερο από 6 ώρες) και εμφανίζει μειονεκτήματα για την βραχυπρόθεσμη πρόβλεψη λόγω της δυσκολίας σύντομης απόκτησης έγκυρων πληροφοριών και της πολύπλοκης υπολογιστικής διαδικασίας που απαιτούν.

### 4.3 ΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ

Σε αντίθεση με τις φυσικές μεθόδους πρόβλεψης που βασίζονται σε αριθμητικά μοντέλα πρόβλεψης και απαιτούν μετεωρολογικά δεδομένα για την εκτίμηση της παραγωγής ηλεκτρικής ενέργειας, οι στατιστικές μέθοδοι για την πρόβλεψη της ενεργειακής παραγωγής χρησιμοποιούν ιστορικά δεδομένα από τις ανεμογεννήτριες. Αυτές οι μέθοδοι προσπαθούν, μέσω της πρότερης συμπεριφοράς των μονάδων παραγωγής, να προβλέψουν τα μελλοντικά αποτελέσματα. Έχουν αναπτυχθεί πολύπλοκα μοντέλα χρησιμοποιώντας σύνθετες τεχνικές ανάλυσης δεδομένων, με δυνατότητα πολλών επιπέδων πρόβλεψης. Ωστόσο, οι στατιστικές μέθοδοι είναι ευαίσθητες στον όγκο και την ποιότητα των εισαγόμενων δεδομένων τα οποία πρέπει να είναι αντιπροσωπευτικά και να μην περιλαμβάνουν ακραίες τιμές. Επιπλέον, δυσκολεύονται να αντιληφθούν σύνθετες μη γραμμικές σχέσεις, ιδίως στην παραγωγή ενέργειας ανεμογεννητριών, ενώ η απόκτηση και επεξεργασία υψηλής ποιότητας ιστορικών δεδομένων για να επιτευχθεί το επιθυμητό αποτέλεσμα αποτελεί πρόκληση. Η στατιστική προσέγγιση αντιμετωπίζει έντονες προκλήσεις στις μακροπρόθεσμες προβλέψεις ωστόσο είναι γενικά μια οικονομική και γρηγορότερη επιλογή.

Κάποια δεδομένα χρονοσειρών αλλάζουν εποχιακά, αργά και σχετικά ομαλά, όπως η μηνιαία ζήτηση ηλεκτρικής ενέργειας. Άλλες χρονοσειρές μπορεί να εκδηλώνουν σχετικά χαοτική συμπεριφορά, διαμορφώνοντας δυσκολίες στην πρόβλεψη με την αποκλειστική χρήση στατιστικών μεθόδων και συχνά εμφανίζουν αδυναμία αποδοτικής ενσωμάτωσης των μη γραμμικών παραμέτρων. Μερικές από τις τεχνικές που χρησιμοποιούνται για την εξαγωγή προβλέψεων περιλαμβάνουν τη χρήση μοντέλων αυτοπαλινδρόμησης AR, κινητούς μέσους όρους MA, ARIMA, ARMA, SARIMA, SARIMAX, Box-Jenkins και Kalman filter. Αυτές οι μέθοδοι βασίζονται σε παρελθοντικά δεδομένα παραγωγής και γενικά υπερέχουν του μοντέλου συνεχιζόμενης πρόβλεψης. Εμφανίζουν συχνά ικανοποιητικά αποτελέσματα για χρονικούς ορίζοντες πρόβλεψης έως 6 ώρες. Στην περίπτωση των ανεμογεννητριών, είναι σύνθετες να χρησιμοποιούνται και δεδομένα από την πρόβλεψη των μετεωρολογικών συνθηκών (ταχύτητα και διεύθυνση του ανέμου) για την βελτιστοποίηση των αποτελεσμάτων και στη συνέχεια να δημιουργείται η τελική πρόβλεψη παραγωγής. Αυτή η διαδικασία συχνά καθιστά τη μέθοδο πιο εφαρμόσιμη και αξιόπιστη. Η μέγιστη βελτίωση στους σύντομους χρονικούς ορίζοντες κυμαίνεται στο εύρος του 15% έως 20%. Για χρονικούς ορίζοντες μεγαλύτερους από 3-6 ώρες, πρέπει να χρησιμοποιούνται εκτενώς οι Αριθμητικές Προβλέψεις Καιρού (NWP) ως είσοδοι στα μοντέλα πρόβλεψης [19], [28].

### 4.4 ΜΕΘΟΔΟΙ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ

#### ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

Η τεχνητή νοημοσύνη (Artificial Intelligence – AI) προέκυψε από τη σύνθεση στοιχείων πολλών επιστημών. Πρόκειται για την «επιστήμη και μηχανική της κατασκευής ευφυών μηχανών» (McCarthy, 1958). Η Μηχανική Μάθηση (Machine Learning – ML) αποτελεί υποσύνολο της Τεχνητής Νοημοσύνης, που αναπτύχθηκε από τη μελέτη της αναγνώρισης



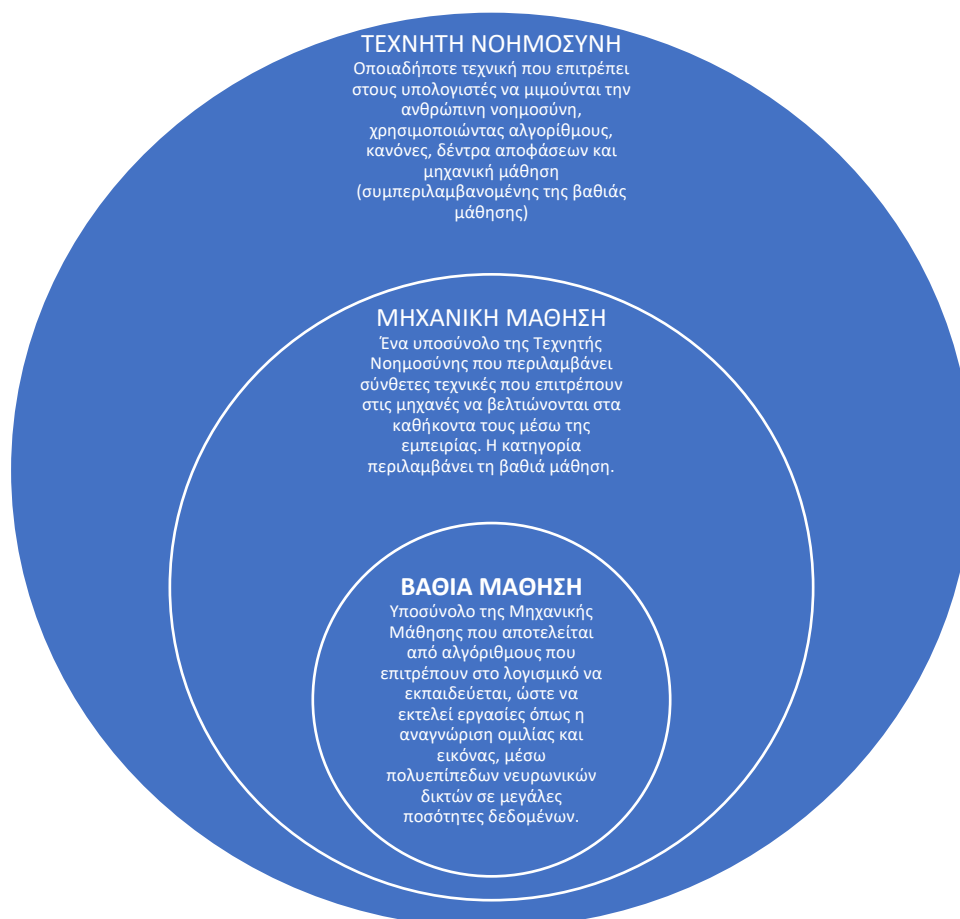
προτύπων και της υπολογιστικής θεωρίας μάθησης στην τεχνητή νοημοσύνη. Το 1959, ο Άρθουρ Σάμουελ ορίζει την μηχανική μάθηση ως Πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί.

### ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Η μηχανική μάθηση διερευνά τη μελέτη αλγορίθμων που μπορούν να μαθαίνουν από τα δεδομένα και να κάνουν προβλέψεις σχετικά με αυτά. Τέτοιοι αλγόριθμοι λειτουργούν κατασκευάζοντας μοντέλα από πειραματικά δεδομένα, προκειμένου να κάνουν προβλέψεις βασισόμενες στα ίδια τα δεδομένα ή να εξάγουν αποφάσεις που εκφράζονται ως το βέλτιστο αποτέλεσμα.

### ΒΑΘΙΑ ΜΑΘΗΣΗ

Η βαθιά μάθηση (Deep Learning) είναι ένα υποσύνολο της μηχανικής μάθησης που επικεντρώνεται κυρίως στη χρήση νευρωνικών δικτύων με πολλά επίπεδα (γνωστά ως βαθιά νευρωνικά δίκτυα) για την εκμάθηση και την αναγνώριση χαρακτηριστικών από τα δεδομένα. Η βαθιά μάθηση επιτρέπει την αυτόματη εξαγωγή χαρακτηριστικών από τα δεδομένα, χωρίς την ανάγκη για εκτεταμένη προ-επεξεργασία.



Εικόνα 6: Σχέση Τεχνητής Νοημοσύνης με Μηχανική Μάθηση και Βαθιά Μάθηση

## 4.4.1 ΤΥΠΟΙ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

Η μηχανική μάθηση διακρίνεται σε τρεις βασικούς τύπους, ανάλογα με τον τρόπο εκπαίδευσης του αλγορίθμου επιλογής, σε σχέση με τα δεδομένα που έχει στη διάθεσή του. Αυτοί οι τρεις τύποι είναι:

#### 4.4.1.1 Εποπτευόμενη Μάθηση (Supervised Learning)

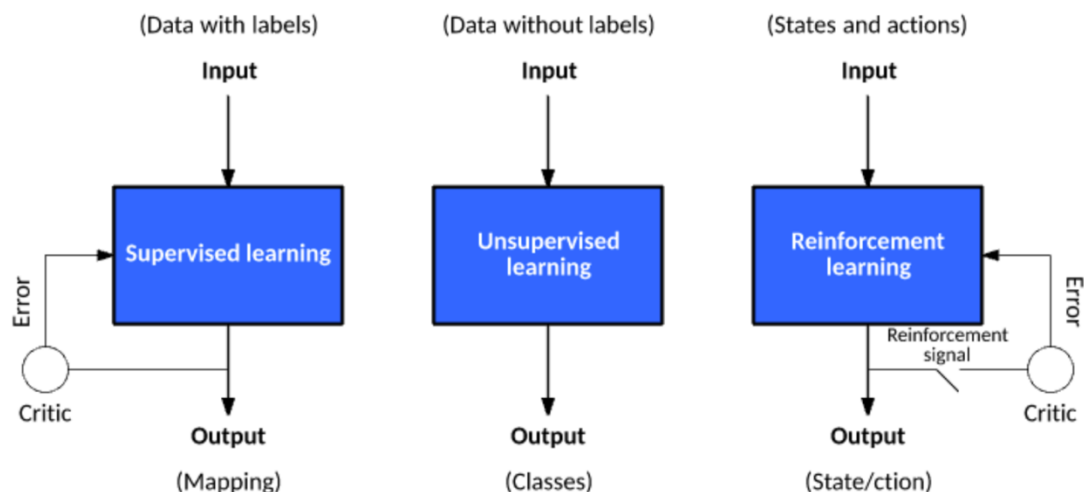
Η εποπτευόμενη μάθηση αποτελεί μια υποκατηγορία μηχανικής μάθησης όπου ο αλγόριθμος εκπαιδεύεται σε ένα επισημασμένο σύνολο δεδομένων. Τα επισημασμένα δεδομένα υποδηλώνουν ότι κάποια από τα δεδομένα εισόδου ακολουθούν τα επιθυμητά αποτελέσματα εξόδου. Ο αλγόριθμος εκπαιδεύεται να κάνει προβλέψεις και να εξάγει συμπεράσματα χρησιμοποιώντας αυτά τα ζευγάρια εισόδου-ετικέτας (label) εξόδου. Συνεπώς, οι αλγόριθμοι εκπαιδεύονται από τα υπάρχοντα δεδομένα καθώς και οι παράμετροι του μοντέλου προσαρμόζονται έτσι ώστε να βελτιστοποιηθεί η ακρίβεια των προβλέψεων και με βάση τα μελλοντικά χαρακτηριστικά των δεδομένων εισόδου προκύπτει η πρόβλεψη εξόδου. Η προσέγγιση που παρουσιάζεται στη συγκεκριμένη διπλωματική εργασία για την εξαγωγή προβλέψεων ενεργειακής παραγωγής αποτελεί ένα χαρακτηριστικό παράδειγμα εποπτευόμενης μάθησης.

#### 4.4.1.2 Μη Εποπτευόμενη Μάθηση (Unsupervised Learning)

Αποτελεί μια κατηγορία μηχανικής μάθησης όπου το μοντέλο εκπαιδεύεται χωρίς να χρησιμοποιεί δεδομένα ετικέτας. Αυτό σημαίνει ότι τα δεδομένα εκπαίδευσης δεν περιλαμβάνουν προκαθορισμένες ετικέτες ή απαντήσεις. Στην μη επιβλεπόμενη μάθηση, το μοντέλο προσπαθεί να ανακαλύψει αυτόνομα τα δομικά πρότυπα, ώστε να βρεθούν μοτίβα ή χαρακτηριστικά από τα δεδομένα χωρίς προηγούμενη καθοδήγηση. Οι κύριες εφαρμογές περιλαμβάνουν τη συσταδοποίηση (clustering) των δεδομένων σε ομάδες, τη μείωση της διαστατικότητας των δεδομένων για ευκολότερη ανάλυση, και την ανίχνευση κρυμμένων προτύπων ή συσχετίσεων.

#### 4.4.1.3 Ενισχυτική Μάθηση (Reinforcement Learning )

Στην ενισχυτική μάθηση ο αλγόριθμος μαθαίνει πώς να λαμβάνει αποφάσεις σε ένα περιβάλλον με σκοπό να μεγιστοποιεί μια ανταμοιβή (ή ενίσχυση). Η παρατήρηση της επιθυμητής ανταμοιβής επιτυγχάνεται μέσω της λεγόμενης συνάρτησης ανταμοιβής η οποία αποτελεί το βασικό κριτήριο εκπαίδευσης του μοντέλου και κρίνεται με βάση την επίδοση του για τα επιθυμητά αποτελέσματα εξόδου.



Εικόνα 7: Εποπτευόμενη, Μη Εποπτευόμενη και Ενισχυτική Μάθηση



## 4.4.2 ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΠΡΟΒΛΗΜΑΤΩΝ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

Στη συνέχεια πραγματοποιείται διαχωρισμός των προβλημάτων. Τα κυριότερα προβλήματα στα οποία η μηχανική μάθηση καλείται να δώσει απαντήσεις είναι:

### 4.4.2.1 Παλινδρόμηση (Regression)

Η παλινδρόμηση αναφέρεται στη μέθοδο που χρησιμοποιείται για τον προσδιορισμό της σχέσης μεταξύ μιας μεταβλητής στόχου, που αποτελεί την εξαρτημένη μεταβλητή, και μιας ή περισσότερων μεταβλητών εισόδου, οι οποίες λειτουργούν ως ανεξάρτητες μεταβλητές. Η παλινδρόμηση είναι γραμμική όταν χρησιμοποιείται μια ανεξάρτητη μεταβλητή, ενώ χαρακτηρίζεται ως πολλαπλή όταν έχουμε δύο ή περισσότερες ανεξάρτητες μεταβλητές. Αυτή η τεχνική εντάσσεται στην εποπτευόμενη μάθηση, όπου το μοντέλο εκπαιδεύεται με βάση τα δεδομένα και τη σχέση μεταξύ των μεταβλητών. Τα προβλήματα εξαγωγής προβλέψεων με βάση ιστορικά δεδομένα αποτελούν ένα χαρακτηριστικό παράδειγμα προβλήματος παλινδρόμησης.

### 4.4.2.2 Ταξινόμηση (Classification)

Στη μηχανική μάθηση, η διαδικασία ταξινόμησης αναφέρεται στον προσδιορισμό και την ομαδοποίηση δεδομένων σε διάφορες κατηγορίες. Οι αλγόριθμοι ταξινόμησης χρησιμοποιούν εκπαιδευμένα και κατηγοριοποιημένα δεδομένα εισόδου για να οδηγήσουν στην ομαδοποίηση μελλοντικών δεδομένων σε αντίστοιχες κατηγορίες. Αυτή η διαδικασία ανήκει στην κατηγορία της εποπτευόμενης μάθησης, καθώς ο αλγόριθμος εκπαιδεύεται με βάση τα προηγούμενα δεδομένα και τις γνωστές κατηγορίες.

### 4.4.2.3 Συσταδοποίηση (Clustering)

Η συσταδοποίηση αποτελεί μια μέθοδο διαχωρισμού μη κωδικοποιημένων σημείων δεδομένων που εμφανίζουν παρόμοια χαρακτηριστικά σε ομάδες, που αποκαλούνται συστάδες. Στις συστάδες, τα σημεία δεδομένων έχουν υψηλή ομοιότητα μεταξύ τους και χαμηλή ομοιότητα με σημεία δεδομένων από άλλες διαφορετικές συστάδες. Η συσταδοποίηση αποτελεί ένα περιγραφικό μοντέλο μη εποπτευόμενης μάθησης, όπου το σύστημα ανακαλύπτει αυτόματα τις ομοιότητες και τα πρότυπα μεταξύ των δεδομένων, χωρίς προηγούμενη επιβλεπόμενη καθοδήγηση.

## 4.4.3 ΑΛΓΟΡΙΘΜΟΙ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ ΣΥΧΝΑ ΕΜΦΑΝΙΖΟΜΕΝΟΙ ΣΤΗ ΒΙΒΛΙΟΓΡΑΦΙΑ ΓΙΑ ΕΞΑΓΩΓΗ ΠΡΟΒΛΕΨΕΩΝ

Στην πρόβλεψη της παραγωγής ηλεκτρικής ενέργειας, χρησιμοποιούνται ποικίλες τεχνικές μηχανικής μάθησης για τη δημιουργία ακριβών προβλέψεων. Ορισμένοι από τους κύριους αλγόριθμους είναι οι εξής:

### 4.4.3.1 Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANN)

Εφαρμογή προηγμένων μοντέλων εμπνευσμένων από τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου. Αποτελούνται από διασυνδεδεμένους νευρώνες (ή κόμβους) ταξινομημένους σε στρώματα. Κάθε νευρώνας δέχεται δεδομένα εισόδου και υπολογίζει μια έξοδο βασισμένη σε μια μη γραμμική συνάρτηση ενεργοποίησης.

Μοντέλο	Παράμετροι εισόδου	Χρονικός Ορίζοντας Πρόβλεψης	Μετρικές	Πηγή
ANN	Ημερήσια μέση ταχύτητα ανέμου (WS), ημερήσια τυπική απόκλιση ταχύτητας ανέμου (STD), ημερήσια μέγιστη και ελάχιστη ταχύτητα ανέμου (MAX/MIN), ωριαία NWP's.	Ημερήσια με βήμα 1 ώρας	R <sup>2</sup>	[29]
RBFNN-CNN	Δεδομένα παραγωγής αιολικής ενέργειας	1 ημέρα	RMSE, NMSE, MAPE, R <sup>2</sup>	[30]
ANN-ARIMA	Ιστορικά δεδομένα ανέμου και ισχύος μέσω συστήματος SCADA	Βραχυπρόθεσμος (λίγα λεπτά έως λίγες ώρες)	MAE, RMSE, MRE	[31]
ANN	Μηνιαία μέση ταχύτητα ανέμου, μηνιαία μέση σχετική υγρασία, ώρες παραγωγής	Μηνιαίος	MSE, MAE, RMSE	[32]
ANN-ARIMA	Δεδομένα ταχύτητας ανέμου	1, 6, 12 12 ώρες	MAE, ME, MAPE, MSE	[33]
ANN, ANN-ARIMA	Ημερήσια μέση/μέγιστη/ελάχιστη ταχύτητα ανέμου, ωριαία NWP's, ιστορικά δεδομένα ανέμου και ισχύος μέσω SCADA, μηνιαία μέση ταχύτητα ανέμου, μηνιαία μέση σχετική υγρασία, ώρες παραγωγής	Ημερήσια, μηνιαία, βραχυπρόθεσμη (λίγα λεπτά έως λίγες ώρες)	MAE, RMSE, MXE, NMAE, MAPE, MASE, R <sup>2</sup>	[34]

Πίνακας 2: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση ANN

#### 4.4.3.2 Τυχαία Δάση (Random Forest – RF)

Δημιουργούν μια συλλογή από δέντρα αποφάσεων κατά τη διάρκεια της εκπαίδευσης. Κάθε δέντρο εκπαιδεύεται σε διαφορετικά υποσύνολα δεδομένων. Στην εξαγωγή πρόβλεψης, τα αποτελέσματα από όλα τα δέντρα συνδυάζονται, για να παραχθεί ένα τελικό αποτέλεσμα.

Μοντέλο	Παράμετροι εισόδου	Χρονικός Ορίζοντας Πρόβλεψης	Μετρικές	Πηγή
RF	Ταχύτητα ανέμου, διεύθυνση ανέμου, παρελθοντικές τιμές ισχύος	1 ώρα	MAE, RMSE, MXE, NMAE, MAPE, MASE, CC	[35]

RF	Θερμοκρασία του αέρα, ατμοσφαιρική πίεση, υγρασία, κατεύθυνση ανέμου, ταχύτητα ανέμου στα 10, 30, 50, 70, 100m, ιστορική παραγωγής αιολικής ενέργειας	1, 2, 4 ώρες και βραχυπρόθεσμες προβλέψεις	MAE, RMSE, MXE, NMAE, MAPE, MASE, CC, R2	[34]
RF	Ταχύτητα ανέμου, διεύθυνση ανέμου, υγρασία, θερμοκρασία, ατμοσφαιρική πίεση	4 ώρες	MAE, RMSE	[36]
RF	Ταχύτητα ανέμου, κατεύθυνση ανέμου, ενεργός ισχύς,	Βραχυπρόθεσμη	MAE, MAPE, RMSE, MSE, R <sup>2</sup>	[37]
RF	16 παράμετροι εισόδου όπως: Θερμοκρασία, γωνία πτερυγίων, στροφές ανά λεπτό, ταχύτητα ανέμου, κατεύθυνση ανέμου	Πολύ βραχυπρόθεσμη	MAE, MSE, RMSE, R <sup>2</sup>	[38]
RF	Ημερήσια μέση ταχύτητα ανέμου (WS) ημερήσια τυπική απόκλιση ταχύτητας ανέμου (STD), ημερήσια μέγιστη και ελάχιστη ταχύτητα ανέμου (MAX/MIN), ωριαία NWP	Ημερήσια με βήμα 1 ώρας	R <sup>2</sup>	[29]
RF	Ημερήσια μέση ταχύτητα ανέμου, τυπική απόκλιση της ταχύτητας ανέμου, συνολική παραγόμενη αιολική ενέργεια	1 έτος	R <sup>2</sup> , MAE, RMSE	[39]

Πίνακας 3: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση Random Forest

#### 4.4.3.3 Δέντρα Αποφάσεων (Decision Trees – DT)

Δημιουργία δομημένων κανόνων για την πρόβλεψη. Λειτουργεί με την ανάπτυξη ενός μοντέλου που προβλέπει την τιμή ενός μεταβλητού-στόχου με την μάθηση απλών κανόνων απόφασης που προκύπτουν από τα δεδομένα των χαρακτηριστικών.

Μοντέλο	Παράμετροι εισόδου	Χρονικός Ορίζοντας Πρόβλεψης	Μετρικές	Πηγή
Decision Tree regression	Ταχύτητα ανέμου, κατεύθυνση ανέμου, ενεργός ισχύς, θεωρητική καμπύλη ισχύος	Βραχυπρόθεσμος	MAE, MAPE, RMSE, MSE, R <sup>2</sup>	[37]
Decision Trees	Δυναμικά χαρακτηριστικά ανέμου, χαρακτηριστικά ανεμογεννήτριας	Ωριαία, εβδομαδιαία και ετήσια πρόβλεψη	MAE, RMSE	[40]
Decision Trees	Ταχύτητα ανέμου, κατεύθυνση ανέμου, ενεργός	1, 2, 4 ώρες	MAE, RMSE,	[34]

	ισχύς, θεωρητική καμπύλη ισχύος	Και βραχυπρόθεσμες προβλέψεις	MSE, MAPE, R <sup>2</sup>	
--	---------------------------------	-------------------------------	---------------------------	--

Πίνακας 4: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση Decision Trees

#### 4.4.3.4 Μοντέλα Ασαφούς Λογικής (Fuzzy Logic – FL)

Χρήση ασαφών μεταβλητών για την αντιμετώπιση της αβεβαιότητας. Αυτά τα μοντέλα μπορούν να χρησιμοποιηθούν για την εκτίμηση της παραγωγής ενέργειας λαμβάνοντας υπόψη ασαφείς παραμέτρους όπως "υψηλή" ή "χαμηλή" ταχύτητα ανέμου, ή "θερμή" ή "κρύα" θερμοκρασία, παρέχοντας έναν πιο ευέλικτο και προσαρμοσμένο τρόπο αντιμετώπισης της πρόβλεψης.

Μοντέλο	Παράμετροι εισόδου	Χρονικός Ορίζοντας Πρόβλεψης	Μετρικές	Πηγή
Fuzzy logic	Πυκνότητα αέρα, ταχύτητα ανέμου	Μηνιαία	RMSE, MAD, MSE	[41]
Fuzzy Neural Network	Δεδομένα χαρακτηριστικά αέρα	2 ώρες με βήμα 15 λεπτών	RMSE, sRMSE, NMAE, IA	[42]
ANFIS	Δυναμικά χαρακτηριστικά ανεμογεννητριών, χωρητικότητα, συντελεστής χωρητικότητας, η ταχύτητα και διεύθυνση του ανέμου	Βραχυπρόθεσμη, ωριαία πρόβλεψη	MAE, MAPE, RMSE, MSE	[43]
ANFIS	Θερμοκρασία, πυκνότητα, πίεση και ταχύτητα του ανέμου	1 ώρα	ARMSE, AReErr	[44]
FL-ANFIS	Ταχύτητα ανέμου, κατεύθυνση ανέμου, ενεργός ισχύς, πυκνότητα αέρα	Μηνιαία, βραχυπρόθεσμη, ωριαία.	RMSE, MAD, MSE, MAE, MAPE	[34]

Πίνακας 5: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση Fuzzy Logic

#### 4.4.3.5 Μηχανή Διανυσματικής Υποστήριξης για Παλινδρόμηση (Support Vector Regression– SVR)

Παραλλαγή του τυπικού μοντέλου SVM που χρησιμοποιείται για την επίλυση προβλημάτων παλινδρόμησης.

Μοντέλο	Παράμετροι εισόδου	Χρονικός Ορίζοντας Πρόβλεψης	Μετρικές	Πηγές
SVR	Ταχύτητα ανέμου, παραγόμενη ισχύς	30 λεπτά	MSE	[45]
SVR	Ταχύτητα ανέμου, παραγόμενη ισχύς	1,2,3,8 ώρες	MAE, std, MAPE	[46]
SVR	Δεδομένα ταχύτητας ανέμου	1 έως 24 ώρες	MAE, RMSE, MASE	[47]

SVR	Δεδομένα ταχύτητας ανέμου	10 λεπτά έως 6 ώρες	MAE, RMSE, ε-loss	[48]
SVR	Ημερήσια Μέση Ταχύτητα Ανέμου, Τυπική Απόκλιση της Ταχύτητας Ανέμου, Συνολική Παραγόμενη Αιολική Ενέργεια	1 έτος	R <sup>2</sup> , MAE, RMSE	[39]
SVR	Παραγωγή αιολικής ενέργειας	Βραχυπρόθεσμη	RMSE, MAE	[49]
SVR	Ιστορικά δεδομένα παραγωγής ηλεκτρικής ενέργειας, δεδομένα καιρού όπως: θερμοκρασία, ατμοσφαιρική πίεση, κατεύθυνση και ταχύτητα του ανέμου	1 ώρα	RMSE	[6]
SVR	Ημερήσια μέση ταχύτητα ανέμου (WS), ημερήσια τυπική απόκλιση ταχύτητας ανέμου (STD), ημερήσια μέγιστη και ελάχιστη ταχύτητα ανέμου (MAX/MIN), ωριαία NWP	Ημερήσια με βήμα 1 ώρας	R <sup>2</sup>	[29]

Πίνακας 6: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση SVM

#### 4.4.3.6 k Πλησιέστεροι Γείτονες (k Nearest Neighbors – kNN)

Βασίζεται στην ιδέα ότι παρόμοια δεδομένα έχουν παρόμοιες προβλέψεις.

Μοντέλο	Παράμετροι εισόδου	Χρονικός Ορίζοντας Πρόβλεψης	Μετρικές	Πηγές
kNN	Ημερήσια μέση ταχύτητα ανέμου, τυπική απόκλιση της ταχύτητας ανέμου, συνολική παραγόμενη αιολική ενέργεια	1 έτος	R <sup>2</sup> , MAE, RMSE	[39]
kNN	Ημερήσια μέση ταχύτητα ανέμου (WS) ημερήσια τυπική απόκλιση ταχύτητας ανέμου (STD), ημερήσια μέγιστη και ελάχιστη ταχύτητα ανέμου (MAX/MIN), ωριαία NWP	Ημερήσια με βήμα 1 ώρας	R <sup>2</sup>	[29]

Πίνακας 7: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση kNN

#### 4.4.3.7 XGBoost

Ένας αλγόριθμος βελτιστοποίησης των δέντρων αποφάσεων που χρησιμοποιεί ένα σύνολο από δέντρα αποφάσεων για την πρόβλεψη.

Μοντέλο	Παράμετροι εισόδου	Χρονικός Ορίζοντας Πρόβλεψης	Μετρικές	Πηγή
XGBoost	Ταχύτητα ανέμου, κατεύθυνση ανέμου, θερμοκρασία, πίεση αέρα, και υγρασία.	Λίγες ώρες έως 7 ημέρες	MSE, RMSE, MAE, RMAE	[50]
XGBoost	Ιστορικές μετρήσεις ισχύος, NWP	3 ώρες	RMSE, PICP, PINAW, CWC	[51]
XGBoost-TCN	NWP	1 ώρα	RMSE, MAE	[52]
XGBoost	Ιστορικά δεδομένα ισχύος ανέμου και ταχύτητας ανέμου, NWP	βραχυπρόθεσμη	RMSE, MAE	[53]
XGBoost	Ημερήσια μέση ταχύτητα ανέμου, τυπική απόκλιση της ταχύτητας ανέμου, συνολική παραγόμενη αιολική ενέργεια	1 έτος	R <sup>2</sup> , MAE, RMSE	[39]
XGBoost	Προβλεπόμενη ταχύτητα ανέμου (GFS), ιστορική ταχύτητα ανέμου, ιστορική παραγωγή ενέργειας	4 ώρες με διάστημα πρόβλεψης 15 λεπτών	RMSE, MAE	[54]

Πίνακας 8: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση XGBoost

#### 4.4.3.8 GRU (Gated Recurrent Unit)

Ένα είδος αναδρομικού νευρωνικού δικτύου (RNN) που χρησιμοποιείται για εξαγωγή προβλέψεων.

Μοντέλο	Παράμετροι εισόδου	Χρονικός Ορίζοντας Πρόβλεψης	Μετρικές	Πηγή
GRU	Ταχύτητα και κατεύθυνση του ανέμου σε διαφορετικά ύψη (10 m, 30 m, 50 m, 70 m), ιστορικά δεδομένα παραγωγής ισχύος.	βραχυπρόθεσμος	RMSE, MAE, R <sup>2</sup>	[55]
GRU	Δεδομένα ανέμου (ιστορική ισχύς, ταχύτητα, κατεύθυνση), καιρικές πληροφορίες (θερμοκρασία, ατμοσφαιρική πίεση, πυκνότητα του αέρα)	Έως αρκετές ημέρες	RMSE, MAE, R <sup>2</sup>	[56]
GRU	Ταχύτητα ανέμου, θερμοκρασία, πίεση	Έως 48 ώρες	MAE, MAPE, RMSE	[57]
GRU	Δεδομένα ανέμου	Μεσοπρόθεσμος και βραχυπρόθεσμος	PICP, PINRW,	[58]

			PINC, CWC	
--	--	--	--------------	--

Πίνακας 9: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση GRU

#### 4.4.3.9 Categorical Boosting

Ένας αλγόριθμος μηχανικής μάθησης που βασίζεται σε δέντρα αποφάσεων, σχεδιασμένος

Μοντέλο	Παράμετροι εισόδου	Χρονικός Ορίζοντας Πρόβλεψης	Μετρικές	Πηγές
CatBoost	Ιστορικά δεδομένα	30 έως 90 λεπτά με διαστήματα 15 λεπτών	MAPE, RMSE	[59]
CatBoost	Προβλεπόμενη ταχύτητα ανέμου (GFS), ιστορική ταχύτητα ανέμου, ιστορική παραγωγή ενέργειας	4 ώρες με διάστημα πρόβλεψης 15 λεπτών	RMSE, MAE	[54]
CatBoost	16 παραμέτρους εισόδου όπως: Θερμοκρασία, γωνία πτερυγίων, στροφές ανά λεπτό, ταχύτητα ανέμου, κατεύθυνση ανέμου	Πολύ βραχυπρόθεσμος	MAE, MSE, RMSE, R <sup>2</sup>	[38]
CatBoost	Ιστορικά δεδομένα παραγωγής ενέργειας, δεδομένα καιρού όπως θερμοκρασία, ατμοσφαιρική πίεση, την κατεύθυνση και την ταχύτητα του ανέμου	1 ώρα	RMSE	[6]

Πίνακας 10: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση CatBoost

#### 4.4.3.10 LSTM (Long Short-Term Memory):

Ένα είδος RNN που χρησιμοποιείται ευρέως για την εξαγωγή προβλέψεων

Μοντέλο	Παράμετροι εισόδου	Χρονικός Ορίζοντας Πρόβλεψης	Μετρικές	Πηγή
LSTM-EFG	Δεδομένα από ανεμογεννήτριες, ταχύτητα ανέμου, παραγωγή ενέργειας	90 λεπτά	MSE	[60]
Bi-LSTM	Δεδομένα παραγωγής αιολικής ενέργειας, ταχύτητα ανέμου	1 ώρα	MSE, MAPE, MAE	[61]
LSTM-GA	Ταχύτητα ανέμου, κατεύθυνση ανέμου, πίεση	Ωριαίες προβλέψεις	MSE, MAE, RMSE, EV, R <sup>2</sup>	[62]
LSTM-CNN	Ταχύτητα ανέμου, κατεύθυνση ανέμου, θερμοκρασία, ατμοσφαιρική πίεση, υγρασία	1 ώρα με 5λεπτο βήμα μέτρησης	MAE, MAPE, RMSE, NRMSE	[63]



Bi-LSTM	Δεδομένα ανέμου, δεδομένα ισχύος, NWP	1 ημέρα	NRMSE, MAE, MAPE, PA	[64]
---------	---------------------------------------	---------	----------------------	------

Πίνακας 11: Βιβλιογραφικές αναφορές για την εξαγωγή προβλέψεων με την χρήση LSTM

#### 4.4.4 ΣΥΝΔΥΑΣΤΙΚΟΙ ΜΕΘΟΔΟΙ (ENSEMBLE METHODS)

Οι συνδυαστικοί μέθοδοι βασίζονται στη χρήση πολλαπλών προγνωστικών μοντέλων για την εξαγωγή μιας συνολικής απόφασης, η οποία είναι καλύτερη από κάθε ένα από τα μεμονωμένα βασικά προγνωστικά μοντέλα [65]. Ιδιαίτερη αποδοτικότητα έχει και η τεχνική της μετά-εκπαίδευσης (meta-learning) και πιο συγκεκριμένα η τεχνική stacking, όπου διάφορα μοντέλα, τα επονομαζόμενα ως βασικά μοντέλα ή προγνωστές παράγουν προβλέψεις που στη συνέχεια χρησιμοποιούνται ως είσοδοι σε ένα άλλο μέτα-μοντέλο (meta-model) για την παραγωγή της τελικής πρόβλεψης. Γενικά οι συνδυαστικοί μέθοδοι διακρίνονται σε ανταγωνιστικούς και συνεργατικούς τύπους:

##### 4.4.4.1 Ανταγωνιστική (Competitive) Πρόβλεψη

Χρησιμοποιεί διάφορα μοντέλα με μικρές αρχικές διαφοροποιήσεις ή διαφορετικές παραμέτρους για να κατασκευάσει ανεξάρτητα προγνωστικά μοντέλα. Η τελική πρόβλεψη προκύπτει από τον συνδυασμό των προβλέψεων όλων ή επιλεγμένων μοντέλων, μέσω διαφόρων μεθόδων όπως ο μέσος όρος. Η ποικιλομορφία των αποφάσεων των βασικών μοντέλων - προγνωστών αποτελεί κεντρικό στοιχείο για την επίτευξη σημαντικής βελτίωσης της ακρίβειας της τελικής πρόβλεψης [65].

##### 4.4.4.2 Συνεργατική (Cooperative) Πρόβλεψη

Η συνεργατική συνδυαστική πρόβλεψη διαιρεί την εργασία πρόβλεψης σε αρκετές υπο-εργασίες και επιλέγει κατάλληλους προγνωστές για κάθε υπο-εργασία με βάση τα χαρακτηριστικά τους. Η τελική απόφαση είναι το άθροισμα όλων των εξόδων των βασικών προγνωστικών μοντέλων [65].

#### 4.4.5 ΥΒΡΙΔΙΚΟΙ ΜΕΘΟΔΟΙ (Hybrid Methods)

Οι υβριδικοί μέθοδοι πρόβλεψης αποτελούν μια συνδυασμένη προσέγγιση, η οποία ενσωματώνει διάφορες τεχνικές πρόβλεψης για να εκμεταλλευτεί τα πλεονεκτήματα της εκάστοτε μεθόδου. Αυτά τα μοντέλα μπορεί να συνδυάζουν στοιχεία από στατιστικά μοντέλα, μηχανική μάθηση, βαθιά μάθηση, και φυσικά μοντέλα για να επιτύχουν μια πιο ολοκληρωμένη και αξιόπιστη πρόβλεψη. Οι υβριδικές μέθοδοι μπορεί επίσης να χρησιμοποιήσουν τεχνικές βελτιστοποίησης ώστε να προσαρμόζουν τα βάρη και τις παραμέτρους των μοντέλων. Αυτή η προσέγγιση επιδιώκει να βελτιώσει την ακρίβεια και την αξιοπιστία των προβλέψεων, συνδυάζοντας διαφορετικούς τύπους πληροφορίας και αναλυτικές τεχνικές. Στην πράξη, οι υβριδικοί μέθοδοι μπορούν να προσαρμόζουν ευέλικτα την προσέγγισή τους ανάλογα με τις ανάγκες του προβλήματος πρόβλεψης, παρέχοντας έτσι μια πιο ολοκληρωμένη και εξατομικευμένη λύση.



# ΚΕΦΑΛΑΙΟ 5: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΤΩΝ ΜΟΝΤΕΛΩΝ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ

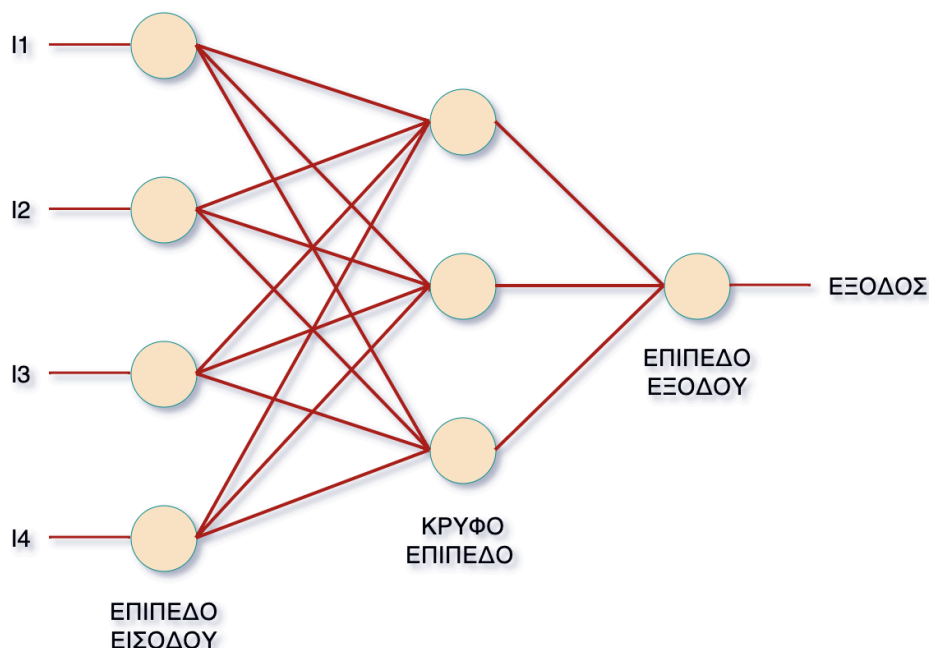
Στη συνέχεια θα υπάρξει εκτεταμένη θεωρητική αναφορά στα νευρωνικά δίκτυα που αποτελούν δομικό στοιχείο στη σύγχρονη εξαγωγή προβλέψεων παραγωγής ενέργειας καθώς και στα υπόλοιπα μοντέλα μηχανικής μάθησης και βαθιάς μάθησης τα οποία επιλέχθηκαν για την εξαγωγή των προβλέψεων σε ένα πραγματικό πρόβλημα 4 ανεμογεννητριών όπως τα Τυχαία Δάση (Random Forest), τα δίκτυα μακράς και βραχυχρόνιας μνήμης (LSTM), τα XGBoost και τα SVM.

## 5.1 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Μια υποκατηγορία μοντέλων μηχανικής μάθησης αποτελούν τα νευρωνικά ή αλλιώς τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks - ANN), τα οποία προσπαθούν να μιμηθούν την λειτουργία του ανθρώπινου εγκεφάλου. Τα βασικά επίπεδα των νευρωνικών δικτύων είναι το επίπεδο εισόδου, εξόδου καθώς και τα κρυφά επίπεδα (hidden layers). Αυτά τα επίπεδα αποτελούνται από νευρώνες, βασικά στοιχεία των οποίων είναι οι εισοδοί, τα βάρη, το bias και η έξοδος. Η έξοδος κάθε νευρώνα προκύπτει από τον τύπο:

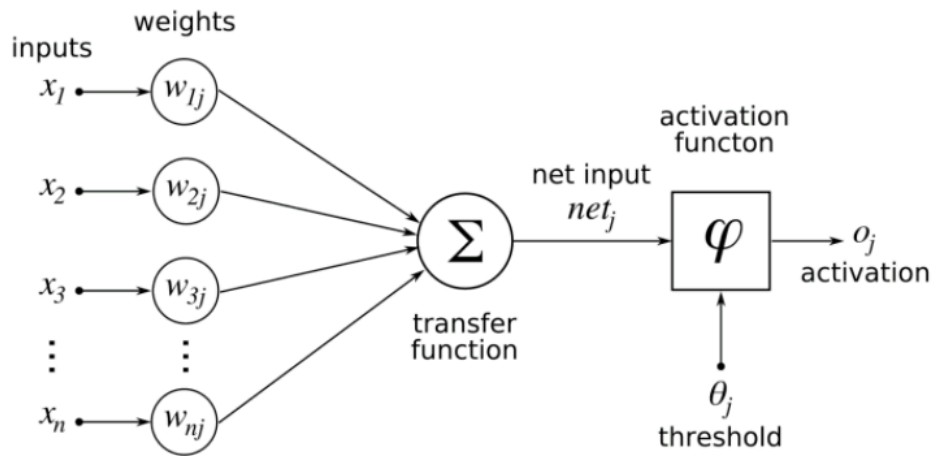
$$y = \sum_{i=1}^n w_i \cdot x_i - bias$$

Όπου  $w_i$  είναι τα βάρη του νευρώνα και  $x_i$  οι εισοδοί του νευρώνα. Πιο συγκεκριμένα, ένας νευρώνας ενεργοποιείται μέσω της συνάρτησης ενεργοποίησης  $\phi$  (activation function), όταν το εσωτερικό γινόμενο των βαρών και των εισόδων ξεπεράσει το κατώφλι (threshold ή bias). Το bias επιτρέπει την ρύθμιση της συνάρτησης ενεργοποίησης, προσφέροντας έτσι σημαντική αναλυτική υποστήριξη για την επιτυχία της εκπαίδευσης του νευρωνικού. Η έξοδος αυτών των δικτύων μπορεί να μοντελοποιηθεί, ώστε να αντιστοιχεί σε μια συγκεκριμένη κατηγορία.



Εικόνα 8: Βασική Δομή Νευρωνικού Δικτύου

Για να χαρακτηριστεί μια υπολογιστική δομή ως ANN θα πρέπει να δομείται σε μορφή κατευθυνόμενου γράφου και για τον λόγο αυτό στη διεθνή βιβλιογραφία ο όρος νευρώνας (neuron) αντικαθίσταται συχνά από τον όρο κόμβος (node) [66].



Εικόνα 9: Μοντέλο Τεχνητού Νευρωνικού Δικτύου

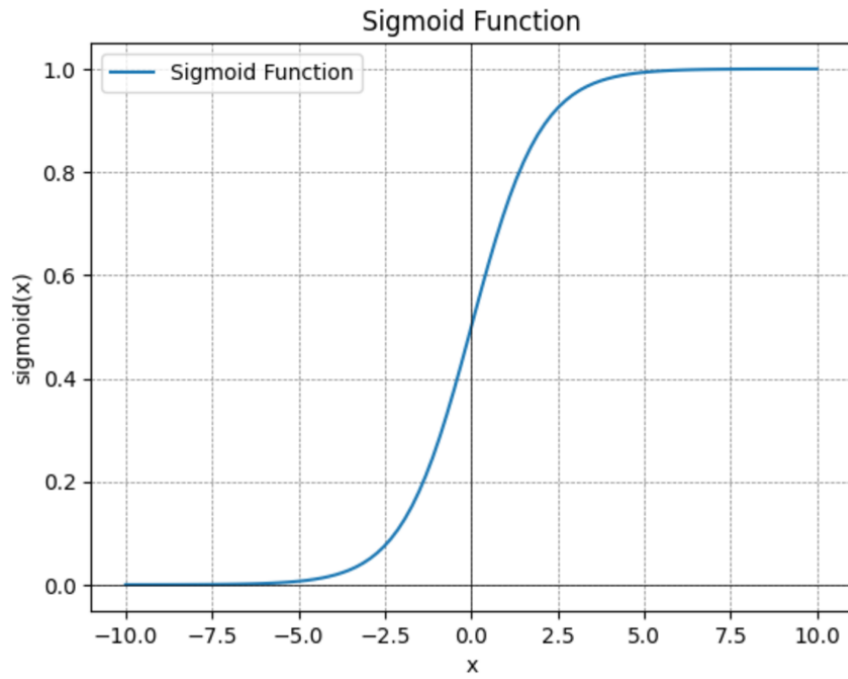
Οι δύο βασικές λειτουργίες ενός νευρωνικού δικτύου είναι η εκπαίδευση (train) και ο έλεγχος (test). Κατά τη διάρκεια της εκπαίδευσης, ο αλγόριθμος μάθησης προσπαθεί να ανακαλύψει τη σχέση μεταξύ της εισόδου και της εξόδου, αλλάζοντας τις τιμές των βαρών των συνάψεων. Μετά τη σύγκριση της πραγματικής εξόδου με την προβλεπόμενη έξοδο, υπολογίζεται το σφάλμα, και οι τιμές των βαρών και των πόλων του νευρωνικού δικτύου ενημερώνονται με βάση αυτό το σφάλμα. Αυτή η διαδικασία επαναλαμβάνεται μέχρι η προβλεπόμενη και πραγματική έξοδος να ταυτιστούν. Η τελική έξοδος υπολογίζεται με βάση τις τελικές τιμές των βαρών και των δεδομένων εισόδου που χρησιμοποιήθηκαν για τον έλεγχο. Τα αποτελέσματα επηρεάζονται σημαντικά από τα δεδομένα εισόδου, τη συνάρτηση ενεργοποίησης και την αρχιτεκτονική του.

### 5.1.1 Συνάρτηση Ενεργοποίησης

Ο σκοπός της συνάρτησης ενεργοποίησης είναι να εισάγει μη γραμμικότητα στην έξοδο ενός νευρώνα. Αυτό είναι σημαντικό καθώς στα προβλήματα πρόβλεψης παρουσιάζονται και μη γραμμικές αλληλουχίες. Ορισμένες από τις συχνότερες συναρτήσεις ενεργοποίησης που εμφανίζονται στα Νευρωνικά Δίκτυα είναι [67]:

A. Σιγμοειδής Συνάρτηση (Sigmoid Function):

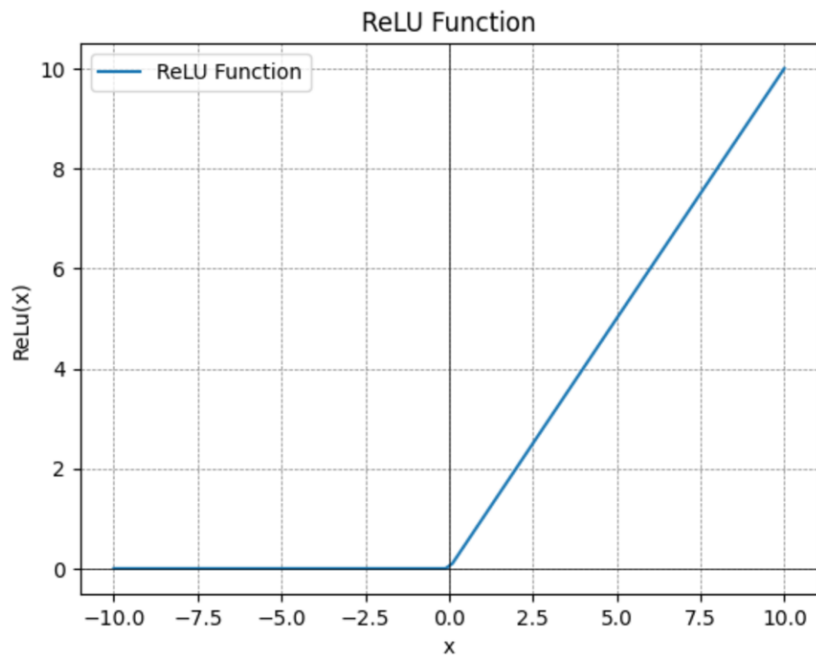
$$\varphi(x) = \frac{1}{1 + e^{-x}}$$



Εικόνα 10: Σιγμοειδής Συνάρτηση (Sigmoid Function)

B. Συνάρτηση Εκθετικής Γραμμικής Μονάδας (ReLU - Rectified Linear Unit)

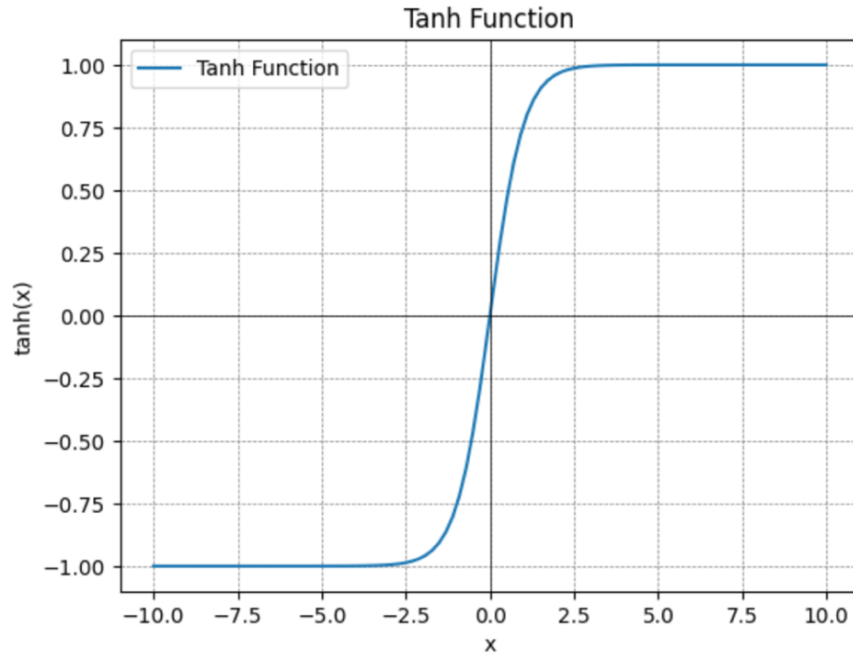
$$\varphi(x) = \max(0, x)$$



Εικόνα 11: Συνάρτηση Εκθετικής Γραμμικής Μονάδας (ReLU - Rectified Linear Unit)

C. Συνάρτηση Υπερβολικής Εφαπτομένης (Hyperbolic Tangent - Tanh)

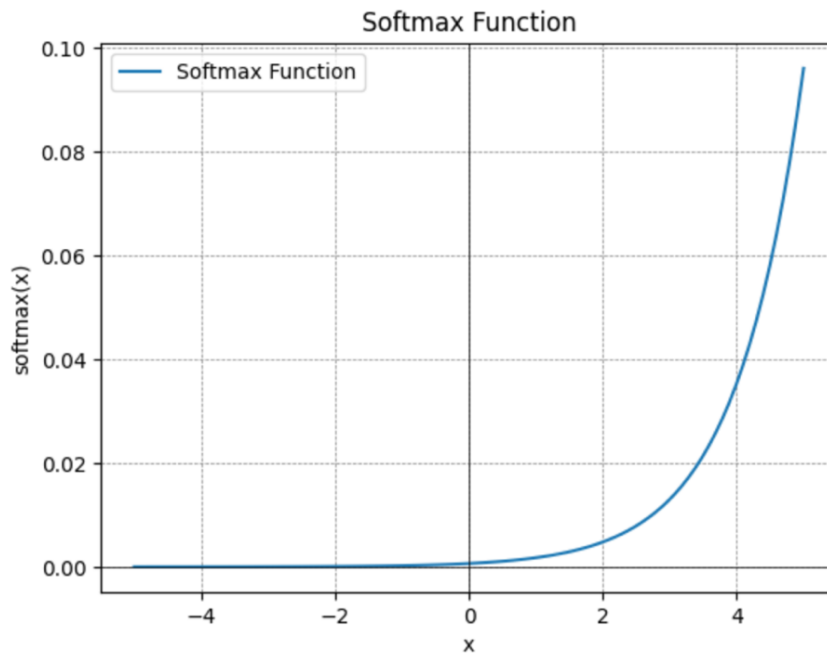
$$\varphi(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}$$



Εικόνα 12: Συνάρτηση Υπερβολικής Εφαπτομένης (Hyperbolic Tangent - Tanh)

D. Softmax Function:

$$\varphi(x)_i = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

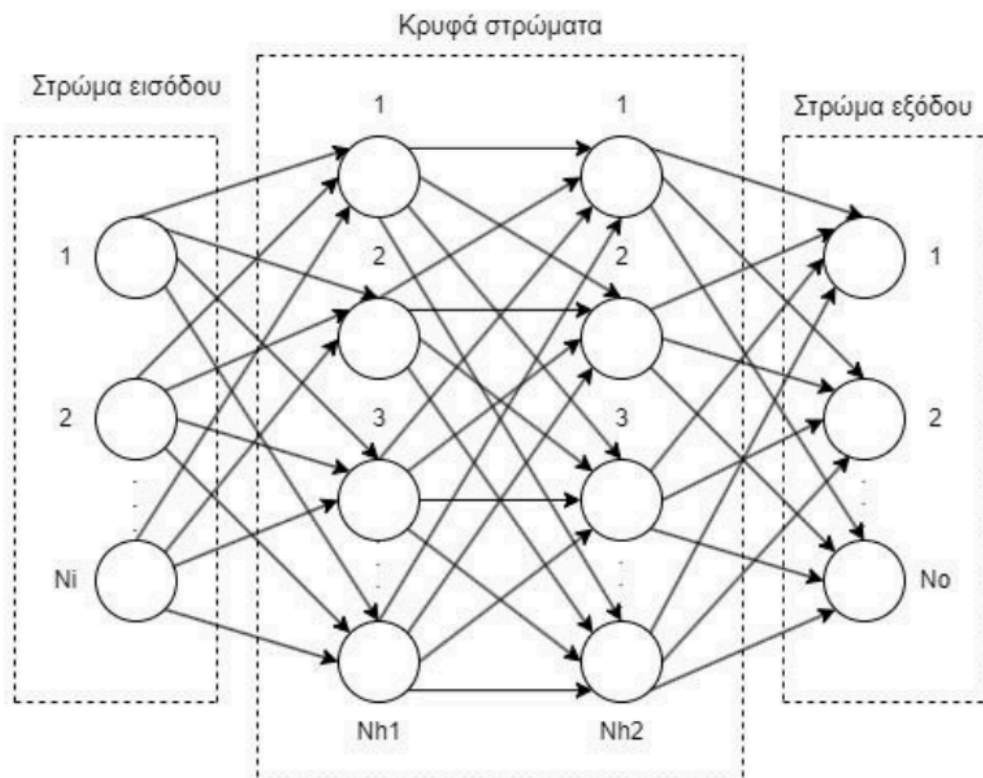


Εικόνα 13: Συνάρτηση Softmax

Η απόδοση ενός νευρωνικού δικτύου, επηρεάζεται αφενός από την ποιότητα των δεδομένων τα οποία καλείται να διαχειριστεί και αφετέρου από την αρχιτεκτονική και τη συνολική δομή των εσωτερικών του συνδέσεων. Χωρίζονται συνήθως σε δύο βασικές κατηγορίες, ανάλογα με την κατεύθυνση της ροής της πληροφορίας εντός της δομής τους:

### 5.1.2 Νευρωνικά Δίκτυα Εμπρόσθιας Τροφοδότησης-Προώθησης (Feedforward Neural Networks / FNNs)

Στο συγκεκριμένο είδος νευρωνικού δικτύου η ροή της πληροφορίας διακυμαίνεται μόνο προς τα εμπρός, από την είσοδο προς την έξοδο, χωρίς αναδρομικές συνδέσεις. Συνεπώς, κάθε επίπεδο επεξεργάζεται την είσοδο και παράγει έξοδο χωρίς επίδραση από προηγούμενες εξόδους



Εικόνα 14: Νευρωνικό Δίκτυο Εμπρόσθιας Τροφοδότησης με 2 κρυφά επίπεδα

Η έξοδος του νευρώνα στα Νευρωνικά Δίκτυα Εμπροσθιας Τροφοδότησης υπολογίζεται με βάση τις συναρτήσεις ενεργοποίησης και τα βάρη των συνδέσεων μεταξύ των νευρώνων των διάφορων στρωμάτων. Η εξίσωση αυτή λαμβάνει υπόψη τη συνολική αρχιτεκτονική του δικτύου και τη ροή της πληροφορίας μεταξύ των στρωμάτων. Παρακάτω ακολουθεί η εξίσωση εξόδου-πρόβλεψης για ένα νευρωνικό δίκτυο με ένα κρυφό επίπεδο (hidden layer) [68].

$$y_1 = f_1 \left[ \sum_{k=1}^K w_{1k} f_2 \left( \sum_{j=1}^J w_{kj} x_j + b_k \right) + b_l \right]$$

Όπου:

- $y_1$ : Η έξοδος, πρόβλεψη
- $f_1$ : Η συνάρτηση ενεργοποίησης του κρυφού επιπέδου
- $w_{1k}$ : Το βάρος της σύνδεσης μεταξύ της εισόδου και του κρυφού νευρώνα

- $f_2$ : Η συνάρτηση ενεργοποίησης του νευρώνα εξόδου
- $w_{kj}$ : Το βάρος της σύνδεσης μεταξύ της εισόδου και του κρυφού νευρώνα
- $b_k$ : Η πόλωση (bias) του κρυφού νευρώνα
- $b_l$ : Η πόλωση του νευρώνα εξόδου

### 5.1.3 Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Network / RNN)

Έχουν αναδρομικές συνδέσεις, δηλαδή υπάρχει αναδρομή της πληροφορίας στη δομή τους, επιτρέποντας τη ροή της πληροφορίας μέσω της μνήμης να διακυμαίνεται προς τα εμπρός και προς τα πίσω.

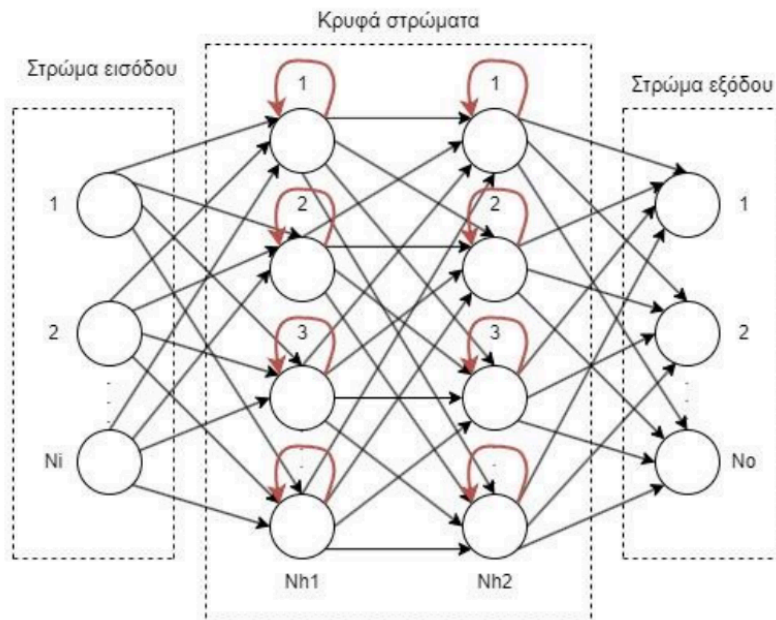
$$h(t) = \sigma(W_h h(t-1) + W_x x(t) + b)$$

$$y(t) = h(t)$$

Όπου :

- $h(t)$  είναι η κατάσταση του νευρώνα στη χρονική στιγμή  $t$ .
- $\sigma$  είναι η συνάρτηση ενεργοποίησης.
- $W_h$  είναι ο πίνακας βαρών για τη σύνδεση μεταξύ των καταστάσεων.
- $h(t-1)$  είναι η προηγούμενη κατάσταση του νευρώνα.
- $W_x$  είναι ο πίνακας βαρών για τη σύνδεση μεταξύ της εισόδου.
- $x(t)$  είναι η είσοδος στη χρονική στιγμή  $t$ .
- $b$  είναι ο όρος πόλωσης.

Ο νευρώνας εξόδου  $y(t)$  λαμβάνει την κατάσταση  $h(t)$  ως είσοδο, προκειμένου να παράγει την έξοδο του δικτύου σε κάθε χρονική στιγμή  $t$ .



Εικόνα 15: Αναδρομικό Νευρωνικό Δίκτυο - RNN

Η διαδικασία για την αναδρομικότητα του νευρωνικού δικτύου επιτυγχάνεται μέσω του backpropagation. Η ιδέα πίσω από το backpropagation είναι ότι υπολογίζουμε τις κλίσεις της τελικής απώλειας ως προς τα βάρη του δικτύου για να καταλάβουμε την κατεύθυνση μείωσης

της απώλειας. Κατά τη διάρκεια της βελτιστοποίησης, κινούμαστε κατά μήκος αυτής της κατεύθυνσης και ενημερώνουμε τα βάρη, προκειμένου να ελαχιστοποιήσουμε την απώλεια.

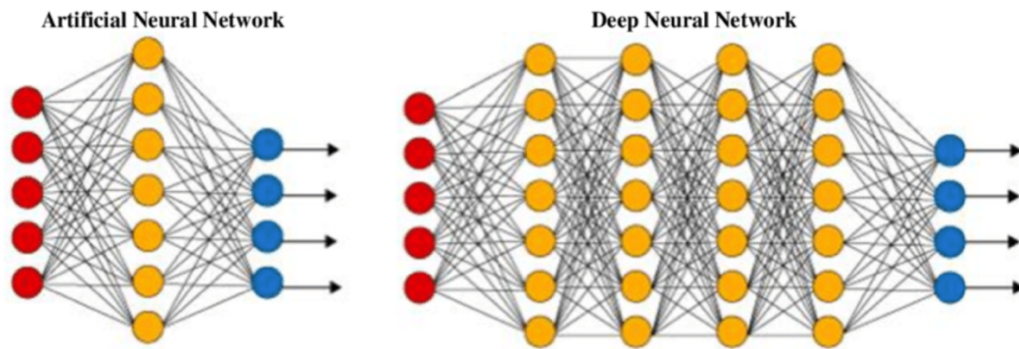
Λόγω αυτού του μηχανισμού, τα αναδρομικά νευρωνικά δίκτυα έχουν τη δυνατότητα να μοντελοποιούν χρονικές εξαρτήσεις, ακόμα και μεταξύ μη συνεχών παρατηρήσεων, εφόσον χρησιμοποιούν την αναδρομικότητα ως έναν μηχανισμό μνήμης. Ωστόσο, στη γενική τους μορφή, τα αναδρομικά νευρωνικά δίκτυα δυσκολεύονται να μοντελοποιήσουν αποτελεσματικά εξαρτήσεις μακράς διάρκειας (όπου η έννοια της "μακράς" είναι σχετική και εξαρτάται από το συγκεκριμένο πρόβλημα).

Η δυσκολία αυτή πηγάζει από τον τρόπο λειτουργίας του αλγορίθμου Back Propagation στο χρόνο, καθώς δεν διατηρεί μεγάλες τιμές κλίσεων όταν εξετάζονται παλαιότερα τμήματα της ακολουθίας. Κατά την εκπαίδευση, ο υπολογισμός της κλίσης του κόστους ως προς παρελθοντικές εισόδους περιλαμβάνει την παραγωγή ενός συνδυασμού συναρτήσεων που οδηγεί σε ένα γινόμενο με αυξανόμενο αριθμό παραγόντων καθώς κινούμαστε προς το παρελθόν. Αυτό μπορεί να οδηγήσει σε αστάθεια, καθώς οι ανανεώσεις των παραμέτρων βασίζονται στον υπολογισμό της κλίσης, και υπερβολικά μεγάλες ή μικρές τιμές μπορεί να αποτρέψουν την αποτελεσματική μάθηση. Αυτό το πρόβλημα συνήθως αναφέρεται ως "εξαφάνιση" ή "εκτόξευση" της κλίσης (vanishing or exploding gradient problem) και μπορεί να οδηγήσει σε αδυναμία εύρεσης ικανοποιητικών τιμών για τα βάρη του μοντέλου κατά την εκπαίδευση. Συνεπώς τα κλασσικά μοντέλα των αναδρομικών νευρωνικών δικτύων εμφανίζουν προβλήματα ως προς την αποτελεσματικότητα μεγάλων ακολουθιακών δεδομένων με μακροχρόνιες εξαρτήσεις.

#### 5.1.4 ΒΑΘΙΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ (Deep Neural Network / DNN)

Τα μοντέλα βαθιάς μάθησης (DNN) αποτελούν υποσύνολο των νευρωνικών δικτύων και διαθέτουν μεγαλύτερο πλήθος (περισσότερα από 3) ενδιάμεσων επιπέδων (κρυφά επίπεδα), γεγονός που τα καθιστά πιο "βαθιά". Υπάρχει μια ποικιλία από διάφορα κρυφά ενδιάμεσα επίπεδα, όπως το επίπεδο συνέλιξης (convolutional layer), το επίπεδο μέγιστης συσσώρευσης (max-pooling layer), το πυκνό επίπεδο (dense layer) και άλλα μοναδικά επίπεδα που το καθένα επιτελεί συγκεκριμένες λειτουργίες. Η βασική διαφορά των βαθιών νευρωνικών δικτύων με τα συμβατικά μοντέλα μηχανικής μάθησης είναι ότι τα πρώτα μπορούν να επεξεργάζονται δεδομένα στην αρχική τους μορφή, χωρίς περαιτέρω προ-επεξεργασία, περιορίζοντας με αυτό τον τρόπο την ανθρώπινη παρέμβαση. Συνεπώς, τα συγκεκριμένα μοντέλα ενδείκνυται να χρησιμοποιούνται για φαινόμενα τα οποία εμφανίζουν μεγάλη πολυπλοκότητα, μη γραμμικότητα, πιθανώς να έχουν μεγάλα σύνολα δεδομένων (big data) και η ύπαρξη πολυπληθών κρυφών στρωμάτων μπορεί να εμφανίσει βελτιωμένα αποτελέσματα στην απόδοση. Ένα χαρακτηριστικό παράδειγμα DNN στο οποίο η ύπαρξη πολλών κρυφών στρωμάτων δίνει τη δυνατότητα για διαχείριση προβλημάτων με μεγαλύτερη πολυπλοκότητα, είναι ο Πολυστρωματικός Ταξινομητής (MLP). Οι MLP αποτελούν μια βασική αρχιτεκτονική ενός Δικτύου Νευρώνων Εμπρόσθιας Τροφοδότησης (Feedforward Neural Network - FFNN) και είναι ιδιαίτερα χρήσιμοι για προβλήματα ταξινόμησης και πρόβλεψης.





Εικόνα 16: Διαφορά μεταξύ ANN και DNN στα κρυφά επίπεδα [69]

### 5.1.5 ΣΥΝΕΛΙΚΤΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ (CNN)

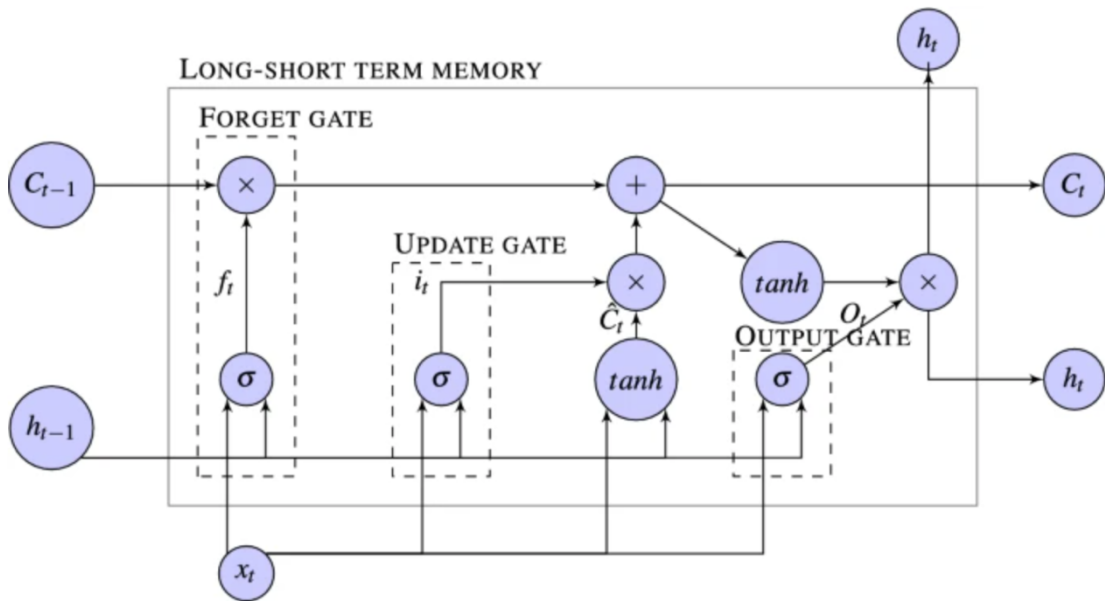
Το Συνελικτικό Νευρωνικό Δίκτυο (CNN) αναφέρεται σε ένα Δίκτυο Νευρώνων που αποτελείται από διάφορα στρώματα, και συγκεκριμένα, το στρώμα εισόδου, τα στρώματα συνέλιξης ή φίλτρα (Convolutional Layers), τα στρώματα συγκέντρωσης (pooling layers), τα πλήρως συνδεδεμένα στρώματα (fully connected layers) και το στρώμα εξόδου. Οι δομές των υλοποιήσεων CNN συνήθως περιλαμβάνουν 5 κρυμμένα στρώματα μεταξύ του στρώματος εισόδου και του στρώματος εξόδου. Ειδικότερα, τα Στρώματα Συνέλιξης περιέχουν πολυδιάστατους πίνακες βαρών, γνωστούς και ως φίλτρα, τα οποία μπορούν να εντοπίζουν συγκεκριμένα χαρακτηριστικά.

Τα φίλτρα αυτά υλοποιούν τη λειτουργία της συνέλιξης μεταξύ των βαρών, ενώ τα στρώματα συγκέντρωσης μειώνουν τις διαστάσεις των στρωμάτων συνέλιξης συνδυάζοντας τις εξόδους των προηγούμενων στρωμάτων προς τους επόμενους νευρώνες. Τα πλήρως συνδεδεμένα στρώματα χρησιμοποιούνται συνήθως στο τελευταίο στάδιο των CNN για την τελική απόφαση ή κατηγοριοποίηση και συνδέουν κάθε νευρώνα του ενός στρώματος με κάθε νευρώνα του επόμενου στρώματος, δημιουργώντας έτσι μια πλήρη σύνδεση. Συνελικτικά αναδρομικά Νευρωνικά Δίκτυα με διαφορετικές αρχιτεκτονικές εμφανίζουν ικανοποιητική απόδοση σε προβλήματα παλινδρόμησης, ενώ χρησιμοποιούνται συχνά για ταξινόμηση και ομαδοποίηση εικόνων, καθώς και για τον εντοπισμό αντικειμένων σε μια εικόνα [70], [71].

## 5.2 ΔΙΚΤΥΑ ΜΑΚΡΑΣ ΚΑΙ ΒΡΑΧΕΙΑΣ ΜΝΗΜΗΣ - LSTM

Τα δίκτυα μακράς και βραχείας μνήμης (Long Short-Term Memory / LSTM) είναι ένα είδος νευρωνικών δικτύων που ανήκουν στην ευρύτερη κατηγορία των αναδρομικών νευρωνικών δικτύων (RNN). Προτάθηκαν πρώτη φορά από τους Hochreiter και Schmidhuber (1997). Τα αναδρομικά νευρωνικά δίκτυα (RNN), παρά την αποτελεσματικότητά τους σε σχέση με τα παραδοσιακά νευρωνικά δίκτυα, εμφανίζουν προβλήματα όπως η "εξαφάνιση διαβαθμίσεων" - "Vanishing Gradient" ή "έκρηξη διαβαθμίσεων" - "Exploding Gradient". Αυτό συμβαίνει λόγω του χειρισμού σφαλματικών σημάτων που κατευθύνονται ανάστροφα. Ως αποτέλεσμα, δημιουργούνται βάρη που ταλαντώνονται ή χάνονται σε μακροπρόθεσμες εξαρτήσεις. Για να αντιμετωπιστεί αυτό το πρόβλημα, χρησιμοποιούνται μονάδες όπως τα LSTM, τα οποία αντιμετωπίζουν το παραπάνω πρόβλημα, εισάγοντας μηχανισμούς πύλης ώστε να ελέγχεται η κατεύθυνση των πληροφοριών μεταξύ των επιπέδων. Συνεπώς, το LSTM διατηρεί μια εσωτερική κατάσταση (cell state) που μπορεί να ενημερώνεται και να χρησιμοποιείται με έναν ελεγκτή για τη λήψη αποφάσεων σχετικά με την πληροφορία που πρέπει να διατηρηθεί ή να απορριφθεί [72].





Εικόνα 17: Δομή LSTM [72]

Στο πλαίσιο των αρχιτεκτονικών των μονάδων LSTM, υπάρχουν τρεις βασικές πύλες που ελέγχουν τη ροή της πληροφορίας: η πύλη εισόδου (input gate), η πύλη εξόδου (output gate) και η πύλη λήψης αποφάσεων (forget gate). Η συνάρτηση ενεργοποίησης (activation functions) που χρησιμοποιείται στις πύλες είναι κυρίως η σιγμοειδής συνάρτηση (sigmoid function) [62].

Οι εξισώσεις που περιγράφουν τις λειτουργίες των πυλών είναι οι εξής:

1. **Πύλη Λήψης Αποφάσεων (Forget Gate):**  $f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$
2. **Πύλη Εισόδου (Input Gate):**  $i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$
3. **Πύλη Εξόδου (Output Gate):**  $o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$
4. **Ενημέρωση της Κυψέλης (Cell Update):**  $g_t = \tan h(W_g x_t + U_g h_{t-1} + b_g)$   
 $c_t = f_t c_{t-1} + i_t g_t$
5. **Παραγωγή της Εξόδου:**  $h_t = o_t \tan h(c_t)$

Όπου:

- $x_t$  είναι η είσοδος στο χρονικό βήμα  $t$ .
- $h_{t-1}$  είναι η προηγούμενη κρυφή κατάσταση (hidden state).
- Το  $\sigma$  αναπαριστά τη σιγμοειδή συνάρτηση (sigmoid function) και αποφασίζει ποιες τιμές επιτρέπονται από το 0 έως το 1
- $\tan h$  είναι η συνάρτηση υπερβολικής εφαιτομένης  $\tan h$  που δίνει βάρη στις τιμές που περνούν καθορίζοντας το επίπεδο σημαντικότητάς τους το οποίο κυμαίνεται από -1 έως 1.
- $W$  και  $U$  είναι πίνακες βαρών.
- $b$  είναι οι παράμετροι πόλωσης.

Αυτές οι πύλες επιτρέπουν στο LSTM να ελέγχει τη ροή των πληροφοριών, επιτρέποντας την αποθήκευση, την ενημέρωση και την ανάκληση πληροφοριών ανάλογα με τις ανάγκες της εκπαίδευσης [73].

### 5.2.1 Bidirectional LSTM (Bi-LSTM)

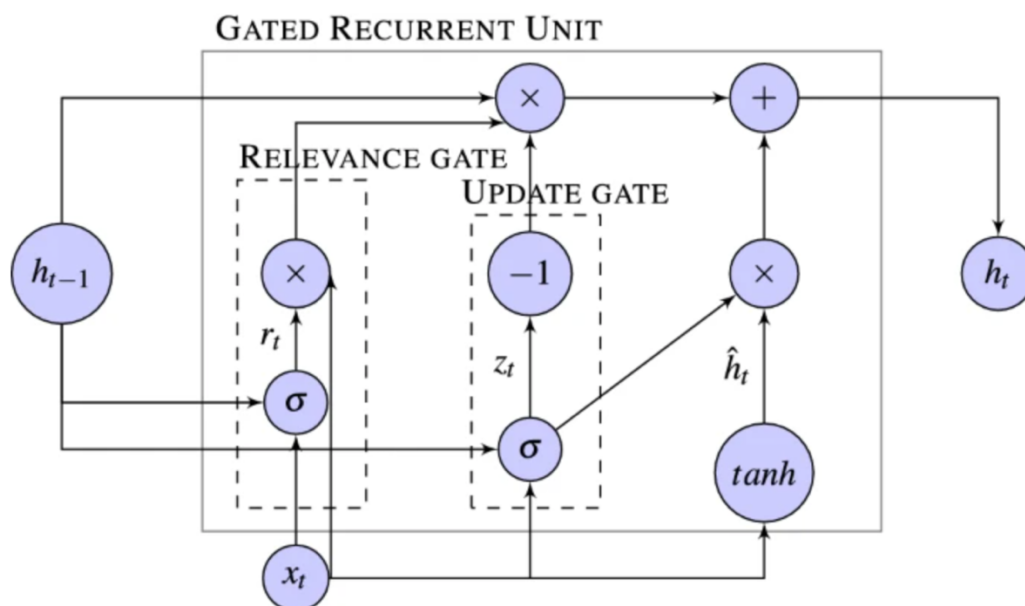
Τα Bi-Birectional Long Short-Term Memory (Bi-LSTM) είναι μια προηγμένη μορφή των LSTM που ενσωματώνουν την ικανότητα επεξεργασίας δεδομένων και από τις δύο

κατευθύνσεις, δηλαδή από το παρελθόν προς το μέλλον και από το μέλλον προς το παρελθόν. Αυτή η διπλή προσέγγιση επιτρέπει στα Bi-LSTM να αντλούν πληροφορίες από μια ευρύτερη περιοχή του χρονικού πλαισίου των δεδομένων, παρέχοντας ένα πιο πλούσιο πλαίσιο για τη λήψη αποφάσεων και την παραγωγή προβλέψεων. Στα Bi-LSTM, τα δεδομένα περνούν από δύο ξεχωριστές αλληλουχίες LSTM και συνδυάζουν πληροφορίες και από τις δύο κατευθύνσεις. Η πρώτη αλληλουχία επεξεργάζεται τη σειρά των δεδομένων από την αρχή προς το τέλος (forward pass) και στη συνέχεια πραγματοποιείται η επεξεργασία της σειράς από το τέλος προς την αρχή (backward pass). Στη συνέχεια οι έξοδοι από τις δύο αυτές αλληλουχίες συνδυάζονται για να παραχθεί η τελική έξοδος για κάθε σημείο της σειράς [74], [75].

### 5.3 Gated Recurrent Unit (GRU)

Ένα άλλο είδος νευρωνικών δικτύων που ανήκει στην ευρύτερη κατηγορία των αναδρομικών νευρωνικών δικτύων (RNN) είναι τα δίκτυα με πύλες ανανέωσης (Gated Recurrent Unit ή GRU). Τα GRU αποτελούν μια παραλλαγή των LSTM και προτάθηκαν για πρώτη φορά από τον Cho, Kyunghyun (2014).

Τα GRU χρησιμοποιούν παρόμοιους μηχανισμούς με τα LSTM, αλλά έχουν μια πιο απλοποιημένη δομή. Συγκεκριμένα, έχουν δύο βασικές πύλες: την πύλη ενημέρωσης (Update Gate) και την πύλη επαναφοράς (Reset Gate). Όπως και τα LSTM, τα GRU σχεδιάστηκαν για να αντιμετωπίσουν προβλήματα όπως η εξαφάνιση και η έκρηξη των διαβαθμίσεων. Αυτές οι πύλες επιτρέπουν στα GRU να ρυθμίζουν τον τρόπο με τον οποίο ενημερώνουν την κατάστασή τους.



Εικόνα 18: Δομή GRU [72]

Οι εξισώσεις που χαρακτηρίζουν τις λειτουργίες των πυλών σε ένα GRU είναι οι εξής:

- **Πύλη Ενημέρωσης (Update Gate):**  $z_t = \sigma(W_z \cdot x_t + U_z \cdot h_{t-1} + b_z)$
- **Πύλη Επαναφοράς (Reset Gate):**  $r_t = \sigma(W_r \cdot x_t + U_r \cdot h_{t-1} + b_r)$
- **Ενημέρωση Κατάστασης (Update State):**  $\tilde{h}_t = \tan h(W \cdot x_t + U \cdot (r_t \odot h_{t-1}) + b)$
- **Ενημέρωση Κατάστασης Βασισμένη στην Πύλη Ενημέρωσης:**  $h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$

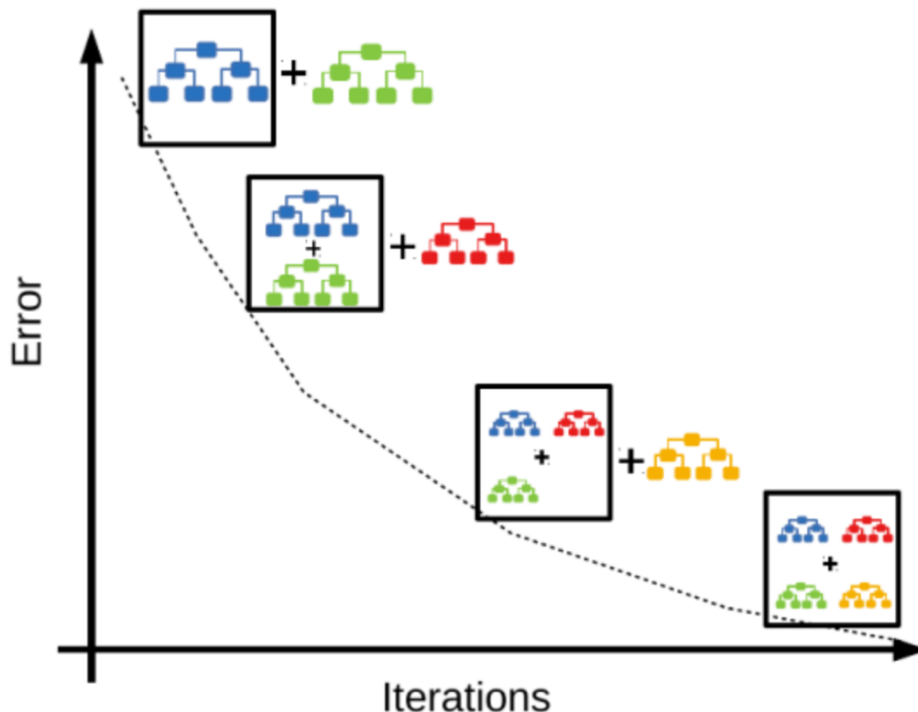
Όπου:

- $x_t$  είναι η είσοδος στο χρονικό βήμα  $t$ .
- $h_{t-1}$  είναι η προηγούμενη κρυφή κατάσταση (hidden state).
- $\sigma$  αναπαριστά τη σιγμοειδή συνάρτηση (sigmoid function).
- $\odot$  αναπαριστά το element-wise product

Με αυτόν τον τρόπο, τα GRU επιτυγχάνουν μια ισορροπία μεταξύ αποτελεσματικής μάθησης με μακροπρόθεσμες εξαρτήσεις και απλότητας στη δομή τους. Τα GRU έχουν λιγότερες παραμέτρους από τα LSTM και δεν διαθέτουν ξεχωριστό "cell state," πράγμα που μπορεί να οδηγήσει σε χαμηλότερη χρήση μνήμης και σε γρηγορότερη εκπαίδευση [72].

## 5.4 XGBOOST

Το XGBoost αποτελεί έναν πολύ δημοφιλή αλγόριθμο μηχανικής μάθησης που προτάθηκε από τον T. Chen (2011) και αναπτύχθηκε με σκοπό την βέλτιστη απόδοση στη μέγιστη υπολογιστική ταχύτητα [76]. Αποτελεί μια πολύ γρήγορη, κλιμακούμενη και βελτιωμένη εφαρμογή του αλγορίθμου Gradient Boosting σύμφωνα με τον οποίο δημιουργούνται πολλά αδύναμα μοντέλα τα οποία συνδυάζονται για να επιτύχουν συνολικά βελτιωμένη απόδοση. Επιλέγει τα δέντρα απόφασης ως βασικούς του μαθητές και προσθέτει νέους μαθητές για τη μείωση του σφάλματος μεταξύ των προβλεπόμενων και των πραγματικών τιμών, ενώ οι τελικές προβλεπόμενες τιμές προκύπτουν από όλα τα δέντρα απόφασης που έχουν χρησιμοποιηθεί. Δηλαδή, η ιδέα είναι να ελαχιστοποιήσει την αντικειμενική συνάρτηση κατασκευάζοντας καλύτερα δέντρα απόφασης και αυτό επιτυγχάνεται μέσω της σταδιακής βελτίωσης, προσθέτοντας νέα δέντρα και προσαρμόζοντας τα βάρη των φύλλων για ελαχιστοποίηση του σφάλματος.



Εικόνα 19: Διαδικασία εκμάθησης με την τεχνική του Gradient Boosting

Οι εξισώσεις που χαρακτηρίζουν τις λειτουργίες του αλγορίθμου XGBoost Regression είναι οι εξής:

$$\text{Αντικειμενική Συνάρτηση: } F_{obj}(\theta) = L(\theta) + \Omega(\theta)$$

Όπου :

$L(\theta) = l(\hat{y}_i, y_i)$  : Η Συνάρτηση απώλειας μεταξύ προβλεπόμενης  $\hat{y}_i$  και πραγματικής  $y_i$  τιμής

$\Omega(\theta) = \gamma T + \frac{1}{2} \lambda \|w\|^2$  : Η συνάρτηση κανονικοποίησης

- $\gamma$  είναι ο ρυθμός κανονικοποίησης (regularization rate),
- $T$  είναι ο αριθμός των φύλλων στο δέντρο
- $\lambda$  είναι η παράμετρος κανονικοποίησης (regularization parameter),
- $w$  είναι τα βάρη των φύλλων

Ο αλγόριθμος XGBoost είναι γρήγορος, κλιμακούμενος, ευέλικτος και εμφανίζει μεγάλη υπολογιστική ταχύτητα για την εξαγωγή αποτελεσμάτων. Συναντά χρήση σε προβλήματα παλινδρόμησης και ταξινόμησης, μοιράζεται χαρακτηριστικά με το Random Forest, αλλά τα δέντρα κατασκευάζονται διαδοχικά, βελτιώνοντας συνεχώς τις προβλέψεις. Το πλεονέκτημα του είναι η χαμηλή μεροληψία ως προς τις σχέσεις μεταξύ των χαρακτηριστικών εκπαίδευσης και των στόχων εξόδου [50], [77].

Στον συγκεκριμένο αλγόριθμο που εφαρμόζεται συχνά για προβλήματα πρόβλεψης παρατηρείται και η βελτιστοποίηση με διάφορες τεχνικές ορισμένων κρίσιμων παραμέτρων ενώ προτείνεται και για ορισμένες εύρος τιμών δοκιμών. Μερικές από αυτές και η χρησιμότητα τους στον αλγόριθμο παρουσιάζονται παρακάτω:

- **n\_estimators**: Καθορίζει τον αριθμό των δέντρων. Μπορεί να επηρεάσει την ακρίβεια του μοντέλου, βελτιώνει τη γενίκευση αλλά αυξάνει τον υπολογιστικό φόρτο.
- **min\_child\_weight**: Καθορίζει το ελάχιστο άθροισμα βαρών των παρατηρήσεων που απαιτείται σε έναν κόμβο. Η ρύθμισή της συγκεκριμένης παραμέτρου βοηθάει στον περιορισμό της υπερεκπαίδευσης.
- **max\_depth**: Καθορίζει το μέγιστο βάθος του κάθε δέντρου και η ρύθμισή της βοηθάει στον έλεγχο της πολυπλοκότητας.
- **Gamma**: Καθορίζει την ελάχιστη μείωση της απώλειας που απαιτείται για την εκτέλεση μιας διαίρεσης
- **Subsample**: Ορίζει το ποσοστό των δειγμάτων που χρησιμοποιούνται για την κατασκευή κάθε δέντρου και βοηθάει στην αύξηση της γενίκευσης και τον περιορισμό υπερεκπαίδευσης
- **colsample\_bytree**: Καθορίζει το ποσοστό των χαρακτηριστικών που επιλέγονται τυχαία για κάθε δέντρο και για κάθε διαίρεση

## 5.5 ΤΥΧΑΙΑ ΔΑΣΗ-RANDOM FOREST

Ο αλγόριθμος Random Forest (RF) συνδυάζει έναν αριθμό δέντρων απόφασης, δημιουργώντας πολλά δέντρα απόφασης που επιλέγονται τυχαία από τα δεδομένα. Ο συγκεκριμένος αλγόριθμος είναι αρκετά δημοφιλής και εντάσσεται στην ευρύτερη κατηγορία των μεθόδων δέντρων αποφάσεων, οι οποίοι κατασκευάζουν πολλά δέντρα απόφασης από ένα σύνολο δεδομένων εισόδου. Ο RF διαιρεί τα χαρακτηριστικά του συνόλου δεδομένων σε διάφορα μέρη και κατασκευάζει δέντρα απόφασης για κάθε μέρος των χαρακτηριστικών. Στη συνέχεια, τα αποτελέσματα κάθε δέντρου απόφασης χρησιμοποιούνται για μια τελική απόφαση και συχνά παρατηρούνται ισχυρές προβλέψεις/κατηγοριοποιήσεις. Συνήθως, ο μέσος όρος των αποτελεσμάτων χρησιμοποιείται για να παραχθούν νέα αποτελέσματα. Με

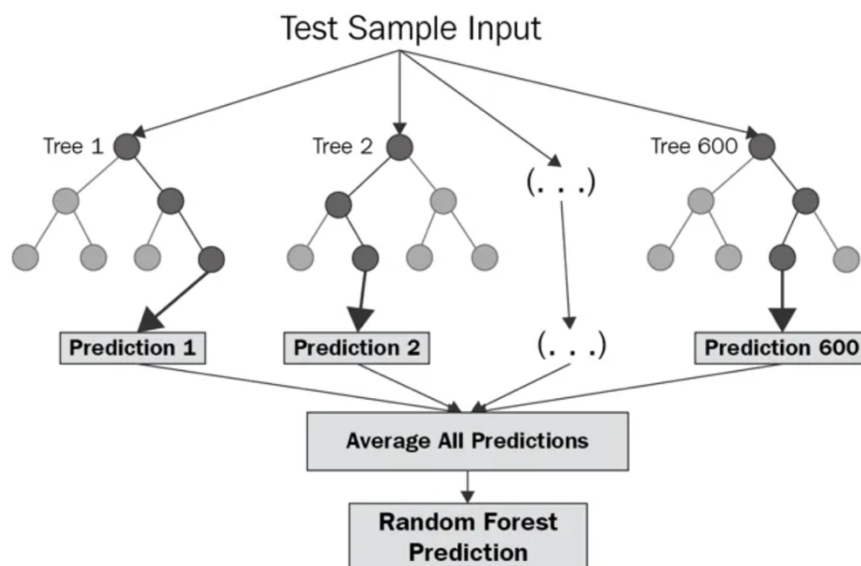
αυτήν τη μεθοδολογία, το πολύπλοκο πρόβλημα με πολλούς χώρους χαρακτηριστικών διαιρείται σε απλούστερα και πιο ερμηνεύσιμα μέρη.

Στο Random Forest, δημιουργείται ένα τυχαίο διάνυσμα  $k$ , που αποτελεί υποσύνολο του χώρου χαρακτηριστικών του συνόλου δεδομένων. Κάθε δέντρο κατασκευάζεται χρησιμοποιώντας αυτό το υποσύνολο  $k$  και το σύνολο εκπαίδευσης. Το σφάλμα γενίκευσης και η συνάρτηση περιθωρίου στο Random Forest δίνονται από την εξίσωση:

$$PE = P_{X,Y}(mg(X,Y) < 0)$$

Όπου:  $mg(X,Y) = \text{argmax}_k I(h_k(X) = Y) - \text{max}_{j \neq Y} \text{argmax}_k I(h_k(X) = j)$

$X$  και  $Y$  είναι τυχαία διανύσματα,  $mg$  (margin function) είναι η συνάρτηση περιθωρίου που ελέγχει τα τυχαία διανύσματα για τη σωστή έξοδο σε σύγκριση με οποιαδήποτε άλλη έξοδο,  $I(\cdot)$  είναι η συνάρτηση δείκτη, και  $h_k$  είναι οι ταξινομητές. Γενικά το σφάλμα γενίκευσης αναφέρεται στο πόσο καλά το μοντέλο μπορεί να γενικεύσει σε νέα δεδομένα που δεν έχει αντιμετωπίσει κατά τη διάρκεια της εκπαίδευσης. Η συνάρτηση περιθωρίου μετρά τη διαφορά μεταξύ της σωστής εξόδου και οποιασδήποτε άλλης πιθανής εξόδου.



Εικόνα 20: Διαδικασία εκμάθησης με την τεχνική του Random Forest [78]

Η απλότητα, η ευελιξία, αλλά και ο μικρός χρόνος που απαιτείται για να εκπαιδευτεί ο συγκεκριμένος αλγόριθμος τον καθιστούν ένα αρκετά χρήσιμο μοντέλο πρόβλεψης. Ωστόσο, η αύξηση της ακρίβειας πρόβλεψης προϋποθέτει τη χρήση περισσότερων δέντρων αποφάσεων, κάτι που καθιστά τον αλγόριθμο πιο αργό [77], [78], [79], [80]. Στον συγκεκριμένο αλγόριθμο που εφαρμόζεται συχνά για προβλήματα πρόβλεψης παρατηρείται και η βελτιστοποίηση με διάφορες τεχνικές ορισμένων κρίσιμων παραμέτρων. Μερικές από αυτές και η χρησιμότητα τους στον αλγόριθμο παρουσιάζονται παρακάτω [81]:

- $n\_estimators$ : Ο αριθμός των δέντρων είναι κομβικός αφού μπορεί να βελτιώσει μέχρι ένα σημείο την ακρίβεια του μοντέλου χωρίς τον κίνδυνο υπερεκπαίδευσης. Η αύξηση των δέντρων βελτιώνει τη γενίκευση αλλά αυξάνει και τον υπολογιστικό φόρτο.
- $max\_depth$ : Η ρύθμιση του μέγιστου βάθους των δέντρων μπορεί να βοηθήσει στην αποφυγή υπερεκπαίδευσης και επηρεάζει την πολυπλοκότητα του μοντέλου.
- $min\_samples\_split$ : Ο ελάχιστος αριθμός δειγμάτων που απαιτούνται για να διασπαστεί ένας εσωτερικός κόμβος.

- `min_samples_leaf`: Ο ελάχιστος αριθμός δειγμάτων που απαιτούνται σε ένα φύλλο-κόμβο. Συχνά, το  $m = 5$  συνιστάται ως ένα καλό σημείο αφετηρίας.
- `max_features`: Ο αριθμός των χαρακτηριστικών για κάθε διαίρεση μπορεί να μειώσει τη συσχέτιση μεταξύ των δέντρων και να βελτιώσει την ακρίβεια του μοντέλου. Συνιστάται η χρήση  $F = n/3$ , όπου  $n$  είναι ο συνολικός αριθμός των χαρακτηριστικών για προβλήματα παλινδρόμησης και η τετραγωνική ρίζα - "sqrt" στα προβλήματα ταξινόμησης.
- `bootstrap`: Η τεχνική αυτή συμβάλλει στην αύξηση της αντοχής του μοντέλου στην υπερεκπαίδευση.

## 5.6 SVR – ΠΑΛΙΝΔΡΟΜΗΣΗ ΔΙΑΝΥΣΜΑΤΩΝ ΥΠΟΣΤΗΡΙΞΗΣ

Η Παλινδρόμηση Διανυσμάτων Υποστήριξης (Support Vector Regression - SVR) αποτελεί μια προηγμένη τεχνική στο πεδίο της μηχανικής μάθησης, η οποία επεκτείνει τις δυνατότητες των Μηχανών Διανυσμάτων Υποστήριξης (Support Vector Machines - SVM) από την ταξινόμηση στην παλινδρόμηση. Η SVR που προτάθηκε από τον Vapnik και την ομάδα του διαφέρει από την τυπική παλινδρόμηση στο ότι δεν επικεντρώνεται απλώς στην ελαχιστοποίηση του σφάλματος πρόβλεψης, αλλά επιδιώκει την εύρεση του βέλτιστου κατωφλίου (margin) μέσα στο οποίο τα σφάλματα είναι αποδεκτά, παρουσιάζοντας μεγαλύτερη γενίκευση και ανθεκτικότητα σε νέα δεδομένα. Η ουσιαστική αρχή της μεθόδου SVM για παλινδρόμηση είναι να απεικονίσει τα δεδομένα σε έναν χώρο υψηλής διάστασης μέσω μη-γραμμικής απεικόνισης και στη συνέχεια να εκτελέσει μια γραμμική παλινδρόμηση σε αυτόν τον χώρο χαρακτηριστικών.

Η συνάρτηση παλινδρόμησης ορίζεται ως :

$$f(x) = \sum_{i=1}^n w_i \varphi_i + b$$

Όπου:

$\varphi_i(x)$  είναι τα ονομαζόμενα χαρακτηριστικά. Αναφέρονται στη μη-γραμμική απεικόνιση των δεδομένων εισόδου σε έναν χώρο υψηλότερης διάστασης. Η ακριβής φύση αυτών των χαρακτηριστικών εξαρτάται από την επιλογή της συνάρτησης πυρήνα. Ο όρος  $b$  αντιπροσωπεύει τον όρο πόλωσης και το  $w_i$  μπορεί να επιτευχθεί με την εξής βελτιστοποίηση:  $\min \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^n (\xi_i + \xi_i^*)$

$$s. t \begin{cases} y_i - \langle \omega, \varphi(x_i) \rangle - b \leq \varepsilon + \xi_i \\ \langle \omega, \varphi(x_i) \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, n \end{cases}$$

Όπου  $\xi_i, \xi_i^*$  είναι μεταβλητές χαλάρωσης και  $C > 0$  είναι μια σταθερά που καθορίζει τις ποινές. Οι μεταβλητές χαλάρωσης επιτρέπουν την ελαστικότητα στο μοντέλο, δίνοντας τη δυνατότητα για εκτός ορίων προβλέψεις. Η σταθερά  $C$  ελέγχει το μέγεθος της ποινής για αυτές τις εκτός ορίων προβλέψεις, ισορροπώντας έτσι μεταξύ της επιθυμίας για ένα μοντέλο με μικρό σφάλμα στα δεδομένα εκπαίδευσης και την ανάγκη να αποφεύγεται η υπερπροσαρμογή. Βάσει αυτών των συνθηκών, το ζήτημα βελτιστοποίησης μπορεί να επιλυθεί και η συνάρτηση εκτίμησης εκφράζεται ως [82]:



$$f(x, \alpha, \alpha^*) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \langle \varphi(x_i), \varphi(x) \rangle + b$$

$$\text{Όπου: } \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0, 0 \leq \alpha_i, \alpha_i^* \leq c$$

## 5.7 ΜΕΤΡΙΚΕΣ ΑΞΙΟΛΟΓΗΣΗΣ ΜΟΝΤΕΛΩΝ ΠΡΟΒΛΕΨΗΣ

Για να προκύψουν συμπεράσματα για την επιλογή του καλύτερου μοντέλου πρόβλεψης πρέπει να γίνει αξιολόγηση των σφαλμάτων που προκύπτουν κατά τη διαδικασία της πρόβλεψης. Η ελαχιστοποίηση του σφάλματος έχει ιδιαίτερα μεγάλη σημασία, αφού η απόκλιση από τις πραγματικές τιμές συνεπάγεται σε οικονομικές απώλειες για του τελικούς χρήστες των προβλέψεων, ενώ μπορεί να δημιουργήσει προβλήματα στο ΣΗΕ. Συνεπώς, η αποτίμηση των προβλέψεων είναι ύψιστης σημασίας για τον καθορισμό της λειτουργικότητας των επιλεγμένων προσεγγίσεων, αλλά επίσης και για την απόκτηση βαθύτερης διορατικότητας σε ό,τι αφορά την αβεβαιότητα της πρόβλεψης.

Για την αξιολόγηση των διαφόρων μεθόδων και της ακρίβειας με την οποία αυτές προβλέπουν την αιολική παραγωγή, χρησιμοποιούνται ορισμένα κριτήρια. Με βάση τη βιβλιογραφία υπάρχουν ορισμένα κριτήρια για την αξιολόγηση των μοντέλων που χρησιμοποιούνται συχνότερα και υπερτερούν έναντι άλλων. Με βάση αυτά τα κριτήρια θα βασιστεί και η αξιολόγηση των μοντέλων πρόβλεψης στο πειραματικό μέρος.

**Ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Square Error–RMSE):**

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

Όπου  $\hat{y}_i$  είναι η προβλεπόμενη τιμή,  $y_i$  είναι η πραγματική τιμή, και  $N$  είναι το πλήθος των σημείων πρόβλεψης. Ένα μικρότερο RMSE σημαίνει ότι το προτεινόμενο μοντέλο λειτουργεί καλύτερα.

**Μέσο απόλυτο σφάλμα (Mean Absolute Error–MAE):**

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

Όπου  $\hat{y}_i$  είναι η προβλεπόμενη τιμή,  $y_i$  είναι η πραγματική τιμή, και  $N$  είναι το πλήθος των σημείων πρόβλεψης. Στον υπολογισμό του λαμβάνονται υπ' όψιν μόνο οι απόλυτες τιμές των σφαλμάτων και όχι οι πραγματικές τιμές τους. Αυτό σημαίνει ότι το MAE είναι ανεξάρτητο από θετικές και αρνητικές τιμές του σφάλματος, δηλαδή είναι ανεξάρτητο από το αν οι τιμές των προβλέψεων είναι μικρότερες (υποεκτίμηση) ή μεγαλύτερες (υπερεκτίμηση) των πραγματικών τιμών.

**Μέσο τετραγωνικό σφάλμα (Mean Squared Error–MSE):**

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Όπου  $\hat{y}_i$  είναι η προβλεπόμενη τιμή,  $y_i$  είναι η πραγματική τιμή, και  $N$  είναι το πλήθος των σημείων πρόβλεψης. Υπολογίζει το μέσο τετραγωνικό σφάλμα μεταξύ των εκτιμώμενων και των αρχικών παραμέτρων, το οποίο μπορεί να αποφύγει το πρόβλημα ότι τα σφάλματα ακυρώνουν το ένα το άλλο, και αντανακλά με ακρίβεια το πραγματικό σφάλμα πρόβλεψης.

### Συντελεστής Προσδιορισμού ( $R^2$ ):

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \text{ για } \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$

Όπου  $\bar{y}$  είναι ο μέσος όρος των πραγματικών τιμών,  $\hat{y}_i$  είναι η προβλεπόμενη τιμή,  $y_i$  είναι η πραγματική τιμή, και  $N$  είναι ο αριθμός των δειγμάτων πρόβλεψης. Ο συντελεστής προσδιορισμού ( $R^2$ ), εκφράζει την αναλογία της διακύμανσης στην εξαρτημένη μεταβλητή που προβλέπεται από τις ανεξάρτητες μεταβλητές. Δείχνει το επίπεδο συσχέτισης μεταξύ της προβλεπόμενης τιμής και της πραγματικής τιμής, και βοηθά στην επιλογή του καλύτερου μοντέλου με την υψηλότερη ακρίβεια πρόβλεψης.

Η σχέση μεταξύ των κριτηρίων RMSE και MSE είναι ανάλογη με τη σχέση μεταξύ τυπικής απόκλισης και διακύμανσης. Αντίστοιχα, το MAE υποδεικνύει την απόλυτη απόκλιση της πρόβλεψης. Ο συντελεστής προσδιορισμού εκφράζει την επιτυχία συσχέτισης εξαρτημένης μεταβλητής και ανεξάρτητων μεταβλητών [83].



## ΚΕΦΑΛΑΙΟ 6: ΠΡΑΚΤΙΚΟ ΜΕΡΟΣ - ΠΡΟΒΛΕΨΗ ΠΑΡΑΓΟΜΕΝΗΣ ΕΝΕΡΓΕΙΑΣ ΑΝΕΜΟΓΕΝΝΗΤΡΙΩΝ

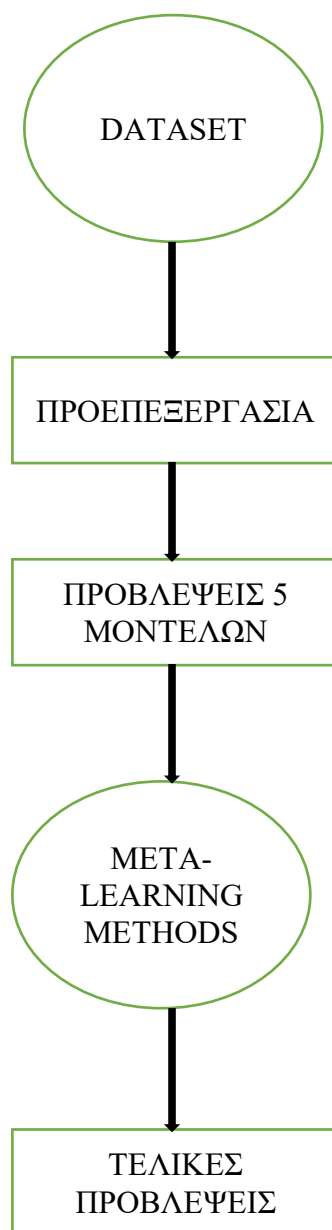
Η μηχανική μάθηση αποτελεί έναν τομέα στην επιστήμη των υπολογιστών που προσδίδει στους υπολογιστές και άλλες συσκευές τη δυνατότητα να μαθαίνουν χωρίς να χρειάζεται η λειτουργία τους να καθορίζεται ρητά. Στοχεύει στην ανάπτυξη μεθόδων και αλγορίθμων που επιτρέπουν τη διαδικασία άντλησης γνώσης από δεδομένα, ώστε να μπορούν προβλέπουν με βάση την εκπαίδευση τους σε ένα σύνολο δεδομένων. Οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται με επιτυχία για να περιγράψουν τη συμπεριφορά δεδομένων, να μοντελοποιήσουν τα χαρακτηριστικά εισόδου σε σχέση με την αναμενόμενη έξοδο, ώστε να προβλέπουν τα χαρακτηριστικά εξόδου με βάση τα ιστορικά τους δεδομένα.

Οι αλγόριθμοι μηχανικής μάθησης είναι αποτελεσματικοί όταν η σχέση μεταξύ των δεδομένων εισόδου και της εξόδου δεν είναι σαφής ή δεν υπάρχει μαθηματικό μοντέλο για το ζεύγος εισόδου-εξόδου. Μπορούν να αντιληφθούν τις αλλαγές στο περιβάλλον και να προσαρμόζουν τα αποτελέσματα σε νέα περιβάλλοντα. Μπορούν επίσης να αντιμετωπίσουν πολύπλοκα συστήματα που αποτελούνται από πολλά μέρη και διαθέτουν ροή δεδομένων μεταξύ τους.

Για την πρόβλεψη, χρησιμοποιούνται ευρέως αλγόριθμοι ταξινόμησης και παλινδρόμησης. Πολλοί από αυτούς τους αλγορίθμους αρχικά ανήκουν στην κατηγορία της ταξινόμησης και τροποποιούνται για να παράγουν πραγματικές τιμές από τα δεδομένα εισόδου. Αυτό το υποσύνολο της ταξινόμησης ονομάζεται ανάλυση παλινδρόμησης.

Σε αυτήν την ενότητα η οποία περιλαμβάνει το πειραματικό μέρος, παρουσιάζονται στοιχεία του dataset, οι διαδικασίες ανάλυσης και προ-επεξεργασίας των δεδομένων του dataset, οι αλγόριθμοι που χρησιμοποιήθηκαν, τα αποτελέσματα καθώς και οι meta-learning προσεγγίσεις που περιλαμβάνουν την ενσωμάτωση όλων των αποτελεσμάτων πρόβλεψης από τις υπόλοιπες μεθόδους. Επιλέχθηκαν να συνδυαστούν για την εξαγωγή προβλέψεων οι συγκεκριμένοι μέθοδοι λόγω της ευρείας χρήσης και της υψηλής απόδοσής τους στη βιβλιογραφία για αντίστοιχα προβλήματα πρόβλεψης αιολικής παραγωγής.

Υπάρχουν πολλές παράμετροι για κάθε αλγόριθμο που επηρεάζουν την απόδοση και τον χρόνο εκτέλεσης των αλγορίθμων. Για την επιλογή των καλύτερων παραμέτρων για κάθε αλγόριθμο, το πειραματικό μέρος βασίστηκε σε μια προσέγγιση δοκιμής και σφάλματος. Συνολικά οι αλγόριθμοι εκτελέστηκαν με διάφορες ρυθμίσεις παραμέτρων και χρησιμοποιήθηκαν τα καλύτερα αποτελέσματα τα οποία παρουσιάζονται στη συνέχεια. Η βασική λογική για την αντιμετώπιση του προβλήματος παρουσιάζεται στο επόμενο σχήμα:



Εικόνα 21: Σχηματική απεικόνιση διαδικασίας εξαγωγής προβλέψεων αιολικής παραγωγής.

## 6.1 TO DATASET

Στα πλαίσια της παρούσας διπλωματικής εργασίας, όλες οι προσομοιώσεις και οι προβλέψεις πραγματοποιούνται για 4 ανεμογεννήτριες. Ειδικότερα τα αρχικά δεδομένα του dataset ήταν κατοχυρωμένα σε μηνιαία βάση δομώντας συνολικά 37 διαφορετικά αρχεία, από τα οποία το καθένα περιέχει μετρήσεις για τις 4 διαφορετικές ανεμογεννήτριες. Τα δεδομένα έχουν ως χρονικό διάστημα μετρήσεων το εξής διάστημα: 01/11/2019 έως και 31/12/2022, ενώ το βήμα μέτρησης και κατοχύρωσης των χαρακτηριστικών είναι ανά δέκα λεπτά. Όλα τα αρχεία CSV πάνω στα οποία βασίστηκε η εξαγωγή προβλέψεων παρουσίαζαν κοινά features που περιλαμβάνουν τόσο αριθμητικά δεδομένα καιρού όσο και μετρήσεις που αφορούν τον τρόπο λειτουργίας των ανεμογεννητριών. Ειδικότερα συνολικά 37 χαρακτηριστικά δομούν τα μηνιαία αρχεία για τις 4 ανεμογεννήτριες εκ των οποίων πολλά από αυτά αντιμετώπιζαν προβλήματα, όπως κενές τιμές, ακραίες τιμές, αρνητικές καταχωρήσεις και λανθασμένες εισαγωγές δεδομένων. Στη συνέχεια παρουσιάζονται τα χαρακτηριστικά τα οποία δομούν τα

αρχικά δεδομένα μετρήσεων του dataset και στη συνέχεια θα αναλυθεί η λογική προεπεξεργασίας ώστε τα δεδομένα να μπορέσουν να εισαχθούν με επιτυχία στα μοντέλα.

Ειδικότερα σε όλα τα αρχεία υπάρχουν οι συγκεκριμένες στήλες:

- **Plant:** Αναφέρει τον αριθμό της ανεμογεννήτριας (1 έως 4) στην οποία παρουσιάζονται οι μετρήσεις
- **Serial no.:** Αναφέρει το σειριακό αριθμό της ανεμογεννήτριας που χρησιμεύει για την ταυτοποίησή της
- **Alias:** Εναλλακτικό ή συντομευμένο όνομα εξεταζόμενης ανεμογεννήτριας
- **Time:** Η χρονική στιγμή της καταγραφής των δεδομένων που ήταν στη μορφή DD/MM/YEAR HOUR AM/PM
- **Wind  $\bar{O}$  [m/s]:** Η μέση ταχύτητα του ανέμου στο εξεταζόμενο διάστημα των 10 λεπτών
- **Wind max. [m/s]:** Η μέγιστη ταχύτητα του ανέμου στο εξεταζόμενο διάστημα των 10 λεπτών
- **Wind min. [m/s]:** Η ελάχιστη ταχύτητα του ανέμου στο εξεταζόμενο διάστημα των 10 λεπτών
- **Rotation speed  $\bar{O}$  [1/min]:** Η μέση ταχύτητα περιστροφής της ανεμογεννήτριας σε περιστροφές ανά λεπτό
- **Rotation speed max. [1/min]:** Η μέγιστη ταχύτητα περιστροφής στο εξεταζόμενο διάστημα των 10 λεπτών
- **Rotation speed min. [1/min]:** Η ελάχιστη ταχύτητα περιστροφής στο εξεταζόμενο διάστημα των 10 λεπτών
- **Power  $\bar{O}$  [kW]:** Η μέση παραγωγή ισχύος της ανεμογεννήτριας
- **Power max. [kW]:** Η μέγιστη παραγωγή ισχύος της ανεμογεννήτριας στο διάστημα παρατήρησης
- **Power min. [kW]:** Η ελάχιστη παραγωγή ισχύος της ανεμογεννήτριας στο διάστημα παρατήρησης
- **Power Avail. wind.  $\bar{O}$  [kW]:** Η μέση διαθέσιμη ισχύς από τον άνεμο
- **Power Avail. techn.  $\bar{O}$  [kW]:** Η μέση διαθέσιμη τεχνική ισχύς
- **Power Avail. force maj.  $\bar{O}$  [kW]:** Η μέση διαθέσιμη ισχύς λόγω μεγάλης ανάγκης
- **Power Avail. ext.  $\bar{O}$  [kW]:** Η μέση διαθέσιμη εξωτερική ισχύς
- **Reactive power  $\bar{O}$  [kvar]:** Η μέση άεργος ισχύς
- **Reactive power max. [kvar]:** Η μέγιστη τιμή της αέργου ισχύος
- **Reactive power min. [kvar]:** Η ελάχιστη τιμή της αέργου ισχύος
- **Energy prod. [kWh]:** Η συνολική ενεργειακή παραγωγή από την πρώτη λειτουργία της ανεμογεννήτριας
- **Energy prod. [kWh].I:** Η παραγωγή ενέργειας μέσα στο διάστημα παρατήρησης
- **Blade angle  $\bar{O}$  [°]:** Η μέση γωνία των πτερυγίων της ανεμογεννήτριας.
- **Precipitate  $\bar{O}$  [mm/min]:** Μέση τιμή της βροχόπτωσης σε χιλιοστά ανά λεπτό
- **Precipitate max [mm/min]:** Η μέγιστη βροχόπτωση σε χιλιοστά ανά λεπτό κατά τη διάρκεια του διαστήματος παρατήρησης
- **Precipitate min [mm/min]:** Η ελάχιστη βροχόπτωση σε χιλιοστά ανά λεπτό κατά τη διάρκεια του διαστήματος παρατήρησης
- **Visibility  $\bar{O}$  [km]:** Η μέση ορατότητα
- **Visibility max [km]:** Η μέγιστη ορατότητα σε χιλιόμετρα κατά τη διάρκεια του διαστήματος παρατήρησης
- **Visibility min [km]:** Η ελάχιστη ορατότητα σε χιλιόμετρα κατά τη διάρκεια του διαστήματος παρατήρησης

- **Illumination  $\theta$  [Lux]**: Η μέση φωτεινότητα
- **Labko ice amplitude  $\theta$  [%]**: Μέση τιμή μέτρησης συσσώρευσης παγετού (ανιχνεύεται από συσκευές Labko, ενσωματωμένες στις ανεμογεννήτριες)
- **Ice acc. timer  $\theta$  [°Cmin]**: Η μέση τιμή της συσσώρευσης παγετού
- **Air pressure  $\theta$  [mBar]**: Η μέση πίεση του αέρα
- **Air humidity  $\theta$  [%]**: Η μέση υγρασία του αέρα
- **Operating hours**: Ώρες λειτουργίας της ανεμογεννήτριας
- **Nacelle position [°]**: Η γωνιακή θέση της καμπίνας της ανεμογεννήτριας

Όταν αναφερόμαστε σε μέση τιμή στο dataset με βήμα 10 λεπτών, εννοούμε την μέση τιμή μιας μετρήσιμης παραμέτρου μέσα σε αυτό το διάστημα. Αυτό σημαίνει ότι παίρνουμε τις μετρήσεις που έχουν γίνει καθ' όλη τη διάρκεια των 10 λεπτών και υπολογίζουμε την μέση τιμή τους για να πάρουμε ένα ενιαίο αποτέλεσμα που αντιπροσωπεύει τη συμπεριφορά ή την κατάσταση της μεταβλητής σε εκείνο το διάστημα. Αυτή η προσέγγιση βοηθά στην απλούστευση των δεδομένων και στην κατανόηση των γενικών τάσεων, χωρίς να επικεντρώνεται σε κάθε μεμονωμένη μέτρηση.

## 6.2 ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ

Η λογική εκμετάλλευσης των συγκεκριμένων προβλεπτικών μοντέλων βασίζεται κατά κύριο λόγο στην προσέγγιση των προβλέψεων μέσα από τα μοντέλα που έχουν εκπαιδευτεί σε ένα σετ δεδομένων μεγάλης διάρκειας και βασίζεται στην απόκριση της ενεργειακής παραγωγής της ανεμογεννήτριας σε σχέση με τις μετεωρολογικές συνθήκες που επικρατούν σε κάθε βήμα μέτρησης. Συνεπώς τα δεδομένα με τα οποία θα εκπαιδευτεί το μοντέλο, θέτοντας ως στήλη στόχου την ενεργειακή παραγωγή και τα οποία θα μπορούν μελλοντικά να εισαχθούν στο μοντέλο είναι κατά κύριο λόγο NWP's και μπορούν να αναζητηθούν από τις μετεωρολογικές υπηρεσίες. Για να οριστούν τα κατάλληλα χαρακτηριστικά και για την ενίσχυση της απόδοσης ήταν αναγκαία η ανάλυση δεδομένων ορισμένων χαρακτηριστικών και ιδιαίτερα της στήλης στόχου, καθώς και η προ-επεξεργασία τους ώστε να είναι στην κατάλληλη μορφή.

Τονίζεται, ότι κομβικός παράγοντας για την πρόβλεψη κατόπιν δοκιμών φάνηκε να αποτελεί και η ταχύτητα περιστροφής της ανεμογεννήτριας, η οποία δεν εμφανίζει μεγάλες αποκλίσεις τιμών κυρίως σε σύντομα χρονικά διαστήματα, ενώ μπορεί να προβλεφθεί και εκείνη με αυτόνομη πρόβλεψη με βάση τα NWP's και στη συνέχεια οι έξοδοι πρόβλεψης να εισαχθούν σαν είσοδοι ενίσχυσης για την τελική εξαγωγή πρόβλεψης παραγωγής για την βελτίωση της απόδοσης. Ωστόσο, τα συγκεκριμένα μοντέλα εμφανίζουν ικανοποιητικά αποτελέσματα και με την αποκλειστική χρήση NWP's αν και η εισαγωγή σαν δεδομένα εισόδου της ταχύτητας περιστροφής της ανεμογεννήτριας βελτιώνει την απόδοσή τους λόγω της υψηλής γραμμικής συσχέτισης.

Η προ-επεξεργασία περιλαμβάνει ως βασικό τμήμα τη δημιουργία συνδυασμένων και χρονολογικά ταξινομημένων αρχείων CSV για κάθε ανεμογεννήτρια και για κάθε έτος ξεχωριστά. Πρώτο βήμα προ-επεξεργασίας είναι η δημιουργία και η επεξεργασία των κατάλληλων ενιαίων αρχείων csv για κάθε μια από τις ανεμογεννήτριες. Ειδικότερα, πραγματοποιήθηκε η Λήψη Λίστας με Αρχεία CSV που αντιστοιχούν σε ένα συγκεκριμένο πρότυπο στον φάκελο για όλα τα έτη ξεχωριστά. Αφού δημιουργηθούν 4 νέα CSVs (*merged\_sorted\_data\_2019*, *merged\_sorted\_data\_2020*, *merged\_sorted\_data\_2021*, *merged\_sorted\_data\_2022*) στα οποία έχουν συγχωνευθεί τα μηνιαία αρχεία και έχουν ταξινομηθεί χρονικά, εκτυπώνονται οι κενές τιμές για κάθε μια στήλη ξεχωριστά.

```

-----2019 dropped columns-----
All values
34828
2019 null values:
Alias                                34828
Precipitate Ø [mm/min]              34378
Precipitate max. [mm/min]           34378
Precipitate min. [mm/min]           34378
Visibility Ø [km]                   34378
Visibility max. [km]                34378
Visibility min. [km]                34378
Labko ice amplitude Ø [%]           34378
Air pressure Ø [mBar]               34378
Power Avail. wind. Ø [kW]           0
Power Avail. techn. Ø [kW]          0
Power Avail. force maj. Ø [kW]      0
Power Avail. ext. Ø [kW]            0
Reactive power Ø [kvar]             0
Energy prod. [kWh]                  0
Reactive power max. [kvar]          0
Reactive power min. [kvar]          0
Serial no.                           0
Illumination Ø [Lux]                25779
Power max. [kW]                     0
Power min. [kW]                     0
Operating hours                      0

```

Εικόνα 22: Εκτύπωση κενών τιμών ανά στήλη δεδομένων που αφαιρέθηκαν για το έτος 2019

Όπως φαίνεται και στην εικόνα για το έτος 2019 επιλέχθηκε να αφαιρεθούν ορισμένες στήλες οι οποίες αφορούν στοιχεία τα οποία δεν μπορούν να αντληθούν μελλοντικά για να εξαχθεί πρόβλεψη παραγωγής ενέργειας όπως: ‘Power Avail. Wind. Ø [kW]’, ‘Power Avail. Techn. Ø [kW]’, ‘Power Avail. Force maj. Ø [kW]’, ‘Power Avail. Ext. Ø [kW]’, ‘Reactive power Ø [kvar]’, ‘Reactive power max. [kvar]’, ‘Reactive power min. [kvar]’, ‘Energy prod. [kWh]’. Επιπλέον, έχουν αφαιρεθεί και στήλες όπως: ‘Precipitate Ø [mm/min]’, ‘Visibility Ø [km]’, ‘Illumination Ø [Lux]’, ‘Labko ice amplitude Ø [%]’, ‘Air pressure Ø [mBar]’, οι οποίες αν και περιέχουν στοιχεία μετεωρολογικών δεδομένων, εμφανίζουν μεγάλο ποσοστό κενών τιμών το οποίο δεν είναι διαχειρίσιμο, ώστε να εισαχθούν στα προβλεπτικά μοντέλα. Επιπλέον ορισμένες στήλες όπως ‘Serial no.’, ‘Alias’, ‘Energy prod. [kWh]’ και ‘Operating hours’ κρίθηκαν ότι δεν χρειάζονται να εισαχθούν στα μοντέλα για την απλούστευση της διαδικασίας καθώς δεν εισάγουν κάποια ουσιαστική πληροφορία για την ενίσχυση της εκμάθησης και της απόδοσης. Οι αντίστοιχες πληροφορίες που φανερώνουν την έλλειψη δεδομένων στα χαρακτηριστικά που αναφέρθηκαν παρουσιάζονται και για τα επόμενα έτη από τα συγχωνευμένα αρχεία.

```

-----2020 dropped columns-----
All values
209836
2020 null values:
Alias                                209836
Precipitate Ø [mm/min]              209547
Precipitate max. [mm/min]           209547
Precipitate min. [mm/min]           209547
Visibility Ø [km]                    209547
Visibility max. [km]                 209547
Visibility min. [km]                 209547
Labko ice amplitude Ø [%]           209547
Air pressure Ø [mBar]                209547
Power Avail. wind. Ø [kW]            0
Power Avail. techn. Ø [kW]          0
Power Avail. force maj. Ø [kW]      0
Power Avail. ext. Ø [kW]             0
Reactive power Ø [kvar]              0
Energy prod. [kWh]                  0
Reactive power max. [kvar]          0
Reactive power min. [kvar]          0
Serial no.                           0
Illumination Ø [Lux]                157150
Power max. [kW]                      0
Power min. [kW]                      0
Operating hours                       0

```

Εικόνα 23: Εκτίπωση κενών τιμών ανά στήλη δεδομένων που αφαιρέθηκαν για το έτος 2020

```

-----2021 dropped columns-----
All values
206710
2021 null values:
Alias                                206710
Precipitate Ø [mm/min]              206180
Precipitate max. [mm/min]           206180
Precipitate min. [mm/min]           206180
Visibility Ø [km]                    206180
Visibility max. [km]                 206180
Visibility min. [km]                 206180
Labko ice amplitude Ø [%]           206180
Air pressure Ø [mBar]                206180
Power Avail. wind. Ø [kW]            0
Power Avail. techn. Ø [kW]          0
Power Avail. force maj. Ø [kW]      0
Power Avail. ext. Ø [kW]             0
Reactive power Ø [kvar]              0
Energy prod. [kWh]                  0
Reactive power max. [kvar]          0
Reactive power min. [kvar]          0
Serial no.                           0
Illumination Ø [Lux]                154597
Power max. [kW]                      0
Power min. [kW]                      0
Operating hours                       0

```

Εικόνα 24: Εκτίπωση κενών τιμών ανά στήλη δεδομένων που αφαιρέθηκαν για το έτος 2021

```

-----2022 dropped columns-----
All values
208496
2022 null values:
Alias                                208496
Precipitate Ø [mm/min]                208339
Precipitate max. [mm/min]             208339
Precipitate min. [mm/min]             208339
Visibility Ø [km]                      208339
Visibility max. [km]                  208339
Visibility min. [km]                  208339
Labko ice amplitude Ø [%]             208339
Air pressure Ø [mBar]                 208339
Power Avail. wind. Ø [kW]              0
Power Avail. techn. Ø [kW]            0
Power Avail. force maj. Ø [kW]        0
Power Avail. ext. Ø [kW]              0
Reactive power Ø [kvar]                0
Energy prod. [kWh]                    0
Reactive power max. [kvar]            0
Reactive power min. [kvar]            0
Serial no.                             0
Illumination Ø [Lux]                  156255
Power max. [kW]                       0
Power min. [kW]                       0
Operating hours                        0

```

Εικόνα 25: Εκτύπωση κενών τιμών ανά στήλη δεδομένων που αφαιρέθηκαν για το έτος 2022

Επόμενο βήμα της προ-επεξεργασίας είναι η δημιουργία της συνάρτησης ‘prepare\_csv(df)’ η οποία θα εφαρμοστεί στα ετήσια DataFrames που έχουν δημιουργηθεί και αποσκοπεί στην κατάλληλη μετατροπή σε αριθμητικό τύπο δεδομένων των στηλών που μας ενδιαφέρουν. Ειδικότερα, ορισμένα από τα δεδομένα είχαν καταγραφεί με τέτοιο τρόπο ώστε το δεκαδικό μέρος χωριζόταν από το ακέραιο τμήμα με τη χρήση υποδιαστολής (.), ενώ η διαχώριση των χιλιάδων γινόταν με τη χρήση τελείας (.). Αυτή η μορφοποίηση δεν αντιστοιχεί στη επιθυμητή μορφή για την εισαγωγή των δεδομένων στα μοντέλα. Αρχικά, η συγκεκριμένη συνάρτηση μετατρέπει τη στήλη time από συμβολοσειρά σε μορφή datetime. Αυτό είναι αναγκαίο, διότι τα datetime objects μπορούν να ταξινομηθούν χρονολογικά και να διευκολύνουν την χρονική ανάλυση. Επιπλέον, η συνάρτηση επαναλαμβάνει μια σειρά από βήματα καθαρισμού και μετατροπής για κάθε μία από τις στήλες ενδιαφέροντος που ορίζονται στη λίστα. Ειδικότερα, αφαιρούνται οι τελείες από τις τιμές, οι οποίες μπορεί να υποδεικνύουν διαχωρισμό χιλιάδων σε κάποιες περιπτώσεις, για να αποφευχθεί η σύγχυση στην ανάγνωση των αριθμών. Στη συνέχεια, η συνάρτηση αντικαθιστά τα κόμματα με τελείες, διότι στην Python το δεκαδικό διαχωριστικό είναι η τελεία. Η συνάρτηση μετατρέπει τις τιμές των στηλών σε αριθμητικούς τύπους δεδομένων (float, int) που είναι οι κατάλληλοι τύποι για να εκτελεστούν αριθμητικοί υπολογισμοί. Τέλος, τα DataFrames ταξινομούνται με βάση την επεξεργασμένη στήλη ‘Time’. Συνεπώς, με την εφαρμογή της συνάρτησης ‘prepare\_csv(df)’ στα ετήσια DataFrames παρατηρούμε πως οι στήλες ενδιαφέροντος είναι στην επιθυμητή μορφή αριθμητικού τύπου δεδομένων:



Plant	int64
Time	datetime64[ns]
Wind Ø [m/s]	float64
Wind max. [m/s]	float64
Wind min. [m/s]	float64
Rotation speed Ø [1/min]	float64
Rotation speed max. [1/min]	float64
Rotation speed min. [1/min]	float64
Power Ø [kW]	float64
Energy prod. [kWh].1	float64
Blade angle Ø [°]	float64
Ice acc. timer Ø [°Cmin]	float64
Air humidity Ø [%]	float64
Nacelle position [°]	int64

Εικόνα 26: Εκτύπωση μορφής δεδομένων ανά στήλη ενδιαφέροντος

Σημαντικό βήμα για την εξαγωγή της πρόβλεψης παραγωγής της κάθε ανεμογεννήτριας ξεχωριστά είναι η δημιουργία ενός ενιαίου DataFrame που θα περιλαμβάνει τα χρονολογικά ταξινομημένα και επεξεργασμένα δεδομένα των μετρήσεων από το 2019 έως και το 2022 για κάθε μία από τις 4 ανεμογεννήτριες. Αυτό, πραγματοποιείται μέσω της αναγνωριστικής στήλης Plant που με τους ακέραιους αριθμούς να κυμαίνονται από 1 έως 4 διαχωρίζονται οι ανεμογεννήτριες σε ξεχωριστά DataFrames για κάθε μια από τις 4 ανεμογεννήτριες. Για παράδειγμα, για το έτος 2019, δημιουργήθηκαν ξεχωριστά DataFrames για τις ανεμογεννήτριες (1 - 4) και το ίδιο πραγματοποιήθηκε και για τα υπόλοιπα έτη. Στη συνέχεια συγχωνεύθηκαν τα ετήσια DataFrames για κάθε μία από τις ανεμογεννήτριες και δημιουργήθηκαν τα νέα αρχεία PLANT1\_2019\_2022\_values.csv, PLANT2\_2019\_2022\_values.csv, PLANT3\_2019\_2022\_values.csv και PLANT4\_2019\_2022\_values.csv. Αυτή η διαδικασία διαχωρισμού δεδομένων ανά ανεμογεννήτρια είναι αναγκαία για την διευκόλυνση περαιτέρω ανάλυση για κάθε μια από τις ανεμογεννήτριες, καθώς και για την ξεχωριστή μοντελοποίηση αυτόνομων προβλέψεων για κάθε ανεμογεννήτρια.

Για τον εντοπισμό λαθών και ατελειών στα δεδομένα τα οποία απαιτούν εξειδικευμένη διαχείριση πριν την περαιτέρω ανάλυση και εκπαίδευση των προβλεπτικών μοντέλων, πραγματοποιήθηκε η απεικόνιση των αρνητικών τιμών και των ακραίων τιμών. Παρατηρήθηκαν ορισμένα παράδοξα, τα οποία σηματοδοτούν την ύπαρξη λαθών στην καταγραφή και καταχώρηση δεδομένων. Στη συνέχεια παρουσιάζονται οι αναλυτικές πληροφορίες των αρνητικών και ακραίων τιμών για κάθε μια από τις ανεμογεννήτριες.

### PLANT 1

Αρνητικές τιμές ανά στήλη:	
Plant	0
Wind Ø [m/s]	0
Wind max. [m/s]	0
Wind min. [m/s]	0
Rotation speed Ø [1/min]	0
Rotation speed max. [1/min]	0
Rotation speed min. [1/min]	0
Power Ø [kW]	14858
Energy prod. [kWh].1	84
Blade angle Ø [°]	0
Ice acc. timer Ø [°Cmin]	0
Air humidity Ø [%]	0
Nacelle position [°]	0
Ice acc. timer Ø [°C/min]	0
dtype:	int64

Εικόνα 27: Αρνητικές τιμές ανά στήλες ενδιαφέροντος-Ανεμογεννήτρια 1



```

Ακραίες τιμές (Μέγιστες και Ελάχιστες) ανά στήλη:
Plant Wind Ø [m/s] Wind max. [m/s] Wind min. [m/s] \
min 1 0.0 0.0 0.0
max 1 30.9 40.1 6553.5

Rotation speed Ø [1/min] Rotation speed max. [1/min] \
min 0.00 0.00
max 20.95 28.28

Rotation speed min. [1/min] Power Ø [kW] Energy prod. [kWh].1
min 0.00 -21.0 -15044260.0
max 655.35 999.0 15044260.0

Blade angle Ø [°] Ice acc. timer Ø [°Cmin] Air humidity Ø [%]
min 0.0 0.0 0.0
max 92.2 360.0 100.0

Nacelle position [°] Ice acc. timer Ø [°C/min]
min 0 0.0
max 359 360.0

```

Εικόνα 28: Ακραίες τιμές ανά στήλες ενδιαφέροντος Ανεμογεννήτρια 1

## PLANT2

Αρνητικές τιμές ανά στήλη:

```

Plant 0
Wind Ø [m/s] 0
Wind max. [m/s] 0
Wind min. [m/s] 0
Rotation speed Ø [1/min] 0
Rotation speed max. [1/min] 0
Rotation speed min. [1/min] 0
Power Ø [kW] 15387
Energy prod. [kWh].1 85
Blade angle Ø [°] 0
Ice acc. timer Ø [°Cmin] 0
Air humidity Ø [%] 0
Nacelle position [°] 0
Ice acc. timer Ø [°C/min] 0
dtype: int64

```

Εικόνα 29: Αρνητικές τιμές ανά στήλες ενδιαφέροντος-Ανεμογεννήτρια 2

```

Ακραίες τιμές (Μέγιστες και Ελάχιστες) ανά στήλη:
Plant Wind Ø [m/s] Wind max. [m/s] Wind min. [m/s] \
min 2 0.0 0.0 0.0
max 2 26.4 36.4 6553.5

Rotation speed Ø [1/min] Rotation speed max. [1/min] \
min 0.00 0.00
max 20.98 28.44

Rotation speed min. [1/min] Power Ø [kW] Energy prod. [kWh].1 \
min 0.00 -24.0 -12132134.0
max 655.35 999.0 12132134.0

Blade angle Ø [°] Ice acc. timer Ø [°Cmin] Air humidity Ø [%] \
min 0.0 0.0 0.0
max 92.2 360.0 100.0

Nacelle position [°] Ice acc. timer Ø [°C/min]
min 0 0.0
max 359 360.0

```

Εικόνα 30: Ακραίες τιμές ανά στήλες ενδιαφέροντος Ανεμογεννήτρια 2

### PLANT3

Αρνητικές τιμές ανά στήλη:

Plant	0
Wind Ø [m/s]	0
Wind max. [m/s]	0
Wind min. [m/s]	0
Rotation speed Ø [1/min]	0
Rotation speed max. [1/min]	0
Rotation speed min. [1/min]	0
Power Ø [kW]	16929
Energy prod. [kWh].1	86
Blade angle Ø [°]	0
Ice acc. timer Ø [°Cmin]	0
Air humidity Ø [%]	0
Nacelle position [°]	0
Ice acc. timer Ø [°C/min]	0

dtype: int64

Εικόνα 31: Αρνητικές τιμές ανά στήλες ενδιαφέροντος-Ανεμογεννήτρια 3

Ακραίες τιμές (Μέγιστες και Ελάχιστες) ανά στήλη:

	Plant	Wind Ø [m/s]	Wind max. [m/s]	Wind min. [m/s]	\
min	3	0.0	0.0	0.0	
max	3	30.2	42.2	6553.5	
		Rotation speed Ø [1/min]	Rotation speed max. [1/min]		\
min		0.00	0.00		
max		20.95	27.02		
		Rotation speed min. [1/min]	Power Ø [kW]	Energy prod. [kWh].1	\
min		0.00	-25.0	-11802859.0	
max		655.35	999.0	11802859.0	
		Blade angle Ø [°]	Ice acc. timer Ø [°Cmin]	Air humidity Ø [%]	\
min		0.0	0.0	0.0	
max		92.1	360.0	100.0	
		Nacelle position [°]	Ice acc. timer Ø [°C/min]		
min		0	0.0		
max		359	360.0		

Εικόνα 32: Ακραίες τιμές ανά στήλες ενδιαφέροντος Ανεμογεννήτρια 3

### PLANT 4

Αρνητικές τιμές ανά στήλη:

Plant	0
Wind Ø [m/s]	0
Wind max. [m/s]	0
Wind min. [m/s]	0
Rotation speed Ø [1/min]	0
Rotation speed max. [1/min]	0
Rotation speed min. [1/min]	0
Power Ø [kW]	16077
Energy prod. [kWh].1	95
Blade angle Ø [°]	0
Ice acc. timer Ø [°Cmin]	0
Air humidity Ø [%]	0
Nacelle position [°]	0
Ice acc. timer Ø [°C/min]	0

dtype: int64

Εικόνα 33: Αρνητικές τιμές ανά στήλες ενδιαφέροντος-Ανεμογεννήτρια 4

Ακραίες τιμές (Μέγιστες και Ελάχιστες) ανά στήλη:					
	Plant	Wind Ø [m/s]	Wind max. [m/s]	Wind min. [m/s]	\
min	4	0.0	0.0	0.0	
max	4	30.5	39.7	6553.5	
		Rotation speed Ø [1/min]	Rotation speed max. [1/min]		\
min		0.00	0.00		
max		21.02	28.33		
		Rotation speed min. [1/min]	Power Ø [kW]	Energy prod. [kWh].1	\
min		0.00	-25.0	-14360511.0	
max		655.35	999.0	14360511.0	
		Blade angle Ø [°]	Ice acc. timer Ø [°Cmin]	Air humidity Ø [%]	\
min		0.0	0.0	0.0	
max		94.0	360.0	100.0	
		Nacelle position [°]	Ice acc. timer Ø [°C/min]		
min		0	0.0		
max		359	360.0		

Εικόνα 34: Ακραίες τιμές ανά στήλες ενδιαφέροντος Ανεμογεννήτρια 4

Συνοπτικά, οι περισσότερες τιμές φαίνονται λογικές εκτός από τις περιπτώσεις: ‘Energy prod. [kWh].1’, ‘Power Ø [kW]’ και ‘Wind min. [m/s]’.

Ειδικότερα:

- Οι ανεμογεννήτριες βρίσκονται υπό κανονικές συνθήκες λειτουργίας, και η ύπαρξη αρνητικών τιμών ισχύος υποδηλώνει ότι η ανεμογεννήτρια καταναλώνει ενέργεια αντί να την παράγει, γεγονός που εγείρει προβληματισμούς. Επιπλέον, δεδομένου ότι η παραγωγή ενέργειας δεν εμφανίζει μηδενικές τιμές στα αντίστοιχα τμήματα αρνητικής ισχύος σηματοδοτεί ότι η ανεμογεννήτρια δεν βρίσκεται υπό τεχνική συντήρηση ή εκτός λειτουργίας για να δικαιολογήσει τις συγκεκριμένες καταγραφές. Συνεπώς η ύπαρξη αυτών των λανθασμένων μετρήσεων πρέπει να αντιμετωπιστεί.
- Η ενεργειακή παραγωγή εμφανίζει ελάχιστες αρνητικές τιμές οι οποίες θα αντικατασταθούν για την ομοιομορφία των δεδομένων της στήλης στόχου, δεδομένου ότι οι προβλέψεις θέλουν να αντιμετωπίσουν το πρόβλημα εύρεσης μη αρνητικών τιμών υπό συγκεκριμένες συνθήκες μετεωρολογικών εισόδων και οι αρνητικές τιμές επηρεάζουν την απόδοση των θετικών τιμών πρόβλεψης.
- Μέσω της απεικόνισης ακραίων τιμών παρατηρήθηκαν μη ρεαλιστικές τιμές μετρήσεων για ορισμένα χαρακτηριστικά. Ειδικότερα μη ρεαλιστικές τιμές στην ελάχιστη ταχύτητα του ανέμου αλλά και στην ενεργειακή παραγωγή σηματοδοτούν την ανάγκη επανεξέτασης και εξισορρόπησης αυτών των τιμών.

Για την αντιμετώπιση των προαναφερθέντων ζητημάτων δημιουργήθηκε η συνάρτηση ‘preprocessing(df)’ η οποία θα εφαρμόζεται στα DataFrames των ανεμογεννητριών. Πιο συγκεκριμένα η συγκεκριμένη συνάρτηση:

- Καθαρίζει την στήλη στόχου ‘Energy prod. [kWh].1’ αντικαθιστώντας με 0 όλες τις μη αριθμητικές τιμές ή τις τιμές που είναι μεγαλύτερες του 1000. Αν κάποια τιμή λείπει (NaN), αυτή επίσης αντικαθίσταται με 0.
- Εφαρμόζει γραμμική παρεμβολή για τη στήλη ‘Air humidity Ø [%]’ αντικαθιστώντας τα κενά δεδομένα (NaN) στη συγκεκριμένη στήλη βοηθώντας στη διατήρηση της συνοχής των δεδομένων.
- Πραγματοποιεί αντικατάσταση των αρνητικών τιμών στη στήλη ‘Power Ø [kW]’ με τον μέσο όρο (mean) των τιμών της στήλης, εξαιρώντας τις αρνητικές τιμές.
- Πραγματοποιεί αντικατάσταση των τιμών που ξεπερνούν οριοθετημένα κατώφλια με εκείνων των μέσων όρων για τις στήλες ‘Rotation speed min. [1/min]’ και ‘Wind min. [m/s]’

Τονίζεται ότι στη συγκεκριμένη συνάρτηση αποφασίστηκε η διαγραφή των στηλών ‘Ice acc. timer Ø [°Cmin]’, ‘Ice acc. timer Ø [°C/min]’ αφού παρατηρήθηκε η ύπαρξη αρκετών Null τιμών και η συσχέτιση με την στήλη στόχου δεν ήταν ιδιαίτερη κρίσιμη για την εξαγωγή προβλέψεων. Επιπλέον, δεδομένου ότι ως στήλη στόχου ορίστηκε η ενεργειακή παραγωγή αποφασίστηκε και η διαγραφή της στήλης ‘Power Ø [kW]’ αν και η προ-επεξεργασία για την συγκεκριμένη στήλη έχει πραγματοποιηθεί και θα μπορούσε και εκείνη αντίστοιχα να αποτελέσει στήλη-στόχο για την εξαγωγή της εκτιμώμενης ισχύος.

Στη συνέχεια παρουσιάζονται όλα τα αποτελέσματα με τις αριθμητικές εισχωρήσεις, τις κενές τιμές και τις ακραίες τιμές όλων των στηλών στα τελικά DataFrames των ανεμογεννητριών ύστερα από την εφαρμογή της συνάρτησης ‘preprocessing(df)’:

```

                                PLANT 1
#      Column                                Non-Null Count  Dtype
---  -
0      Time                                164968 non-null  datetime64[ns]
1      Wind Ø [m/s]                        164968 non-null  float64
2      Wind max. [m/s]                     164968 non-null  float64
3      Wind min. [m/s]                     164968 non-null  float64
4      Rotation speed Ø [1/min]            164968 non-null  float64
5      Rotation speed max. [1/min]         164968 non-null  float64
6      Rotation speed min. [1/min]         164968 non-null  float64
7      Energy prod. [kWh].1                164968 non-null  float64
8      Blade angle Ø [°]                   164968 non-null  float64
9      Air humidity Ø [%]                  164968 non-null  float64
10     Nacelle position [°]                 164968 non-null  int64
dtypes: datetime64[ns](1), float64(9), int64(1)
memory usage: 15.1 MB
None

```

Αρνητικές τιμές ανά στήλη:

```

Wind Ø [m/s]                0
Wind max. [m/s]             0
Wind min. [m/s]             0
Rotation speed Ø [1/min]    0
Rotation speed max. [1/min] 0
Rotation speed min. [1/min] 0
Energy prod. [kWh].1        0
Blade angle Ø [°]           0
Air humidity Ø [%]          0
Nacelle position [°]        0
dtype: int64

```

Ακραίες τιμές (Μέγιστες και Ελάχιστες) ανά στήλη:

```

      Wind Ø [m/s]  Wind max. [m/s]  Wind min. [m/s]  \
min                0.0              0.0              0.0
max               30.9              40.1              23.0

      Rotation speed Ø [1/min]  Rotation speed max. [1/min]  \
min                          0.00                          0.00
max                          20.95                          28.28

      Rotation speed min. [1/min]  Energy prod. [kWh].1  Blade angle Ø [°]  \
min                               0.00                       0.0              0.0
max                               20.71                       557.0            92.2

      Air humidity Ø [%]  Nacelle position [°]
min                      0.0              0
max                      100.0            359

```

Εικόνα 35: Μορφή δεδομένων, αρνητικές και ακραίες τιμές ύστερα από την προ-επεξεργασία για την Ανεμογεννήτρια 1

## PLANT 2

#	Column	Non-Null Count	Dtype
0	Time	164968 non-null	datetime64[ns]
1	Wind Ø [m/s]	164968 non-null	float64
2	Wind max. [m/s]	164968 non-null	float64
3	Wind min. [m/s]	164968 non-null	float64
4	Rotation speed Ø [1/min]	164968 non-null	float64
5	Rotation speed max. [1/min]	164968 non-null	float64
6	Rotation speed min. [1/min]	164968 non-null	float64
7	Energy prod. [kWh].1	164968 non-null	float64
8	Blade angle Ø [°]	164968 non-null	float64
9	Air humidity Ø [%]	164968 non-null	float64
10	Nacelle position [°]	164968 non-null	int64

dtypes: datetime64[ns](1), float64(9), int64(1)

memory usage: 15.1 MB

None

Αρνητικές τιμές ανά στήλη:

Wind Ø [m/s]	0
Wind max. [m/s]	0
Wind min. [m/s]	0
Rotation speed Ø [1/min]	0
Rotation speed max. [1/min]	0
Rotation speed min. [1/min]	0
Energy prod. [kWh].1	0
Blade angle Ø [°]	0
Air humidity Ø [%]	0
Nacelle position [°]	0

dtype: int64

Ακραίες τιμές (Μέγιστες και Ελάχιστες) ανά στήλη:

	Wind Ø [m/s]	Wind max. [m/s]	Wind min. [m/s]	\
min	0.0	0.0	0.0	
max	26.4	36.4	18.2	

	Rotation speed Ø [1/min]	Rotation speed max. [1/min]	\
min	0.00	0.00	
max	20.98	28.44	

	Rotation speed min. [1/min]	Energy prod. [kWh].1	Blade angle Ø [°]	\
min	0.00	0.0	0.0	
max	20.45	695.0	92.2	

	Air humidity Ø [%]	Nacelle position [°]
min	0.0	0
max	100.0	359

Εικόνα 36: Μορφή δεδομένων, αρνητικές και ακραίες τιμές ύστερα από την προ-επεξεργασία για την Ανεμογεννήτρια 2

### PLANT 3

#	Column	Non-Null Count	Dtype
0	Time	164967 non-null	datetime64[ns]
1	Wind Ø [m/s]	164967 non-null	float64
2	Wind max. [m/s]	164967 non-null	float64
3	Wind min. [m/s]	164967 non-null	float64
4	Rotation speed Ø [1/min]	164967 non-null	float64
5	Rotation speed max. [1/min]	164967 non-null	float64
6	Rotation speed min. [1/min]	164967 non-null	float64
7	Energy prod. [kWh].1	164967 non-null	float64
8	Blade angle Ø [°]	164967 non-null	float64
9	Air humidity Ø [%]	164967 non-null	float64
10	Nacelle position [°]	164967 non-null	int64

dtypes: datetime64[ns](1), float64(9), int64(1)

memory usage: 15.1 MB

None

Αρνητικές τιμές ανά στήλη:

Wind Ø [m/s]	0
Wind max. [m/s]	0
Wind min. [m/s]	0
Rotation speed Ø [1/min]	0
Rotation speed max. [1/min]	0
Rotation speed min. [1/min]	0
Energy prod. [kWh].1	0
Blade angle Ø [°]	0
Air humidity Ø [%]	0
Nacelle position [°]	0

dtype: int64

Ακραίες τιμές (Μέγιστες και Ελάχιστες) ανά στήλη:

	Wind Ø [m/s]	Wind max. [m/s]	Wind min. [m/s]	\
min	0.0	0.0	0.0	
max	30.2	42.2	21.6	
	Rotation speed Ø [1/min]	Rotation speed max. [1/min]	\	
min	0.00	0.00		
max	20.95	27.02		
	Rotation speed min. [1/min]	Energy prod. [kWh].1	Blade angle Ø [°]	\
min	0.00	0.0	0.0	
max	20.67	484.0	92.1	
	Air humidity Ø [%]	Nacelle position [°]		
min	0.0	0		
max	100.0	359		

Εικόνα 37: Μορφή δεδομένων, αρνητικές και ακραίες τιμές ύστερα από την προ-επεξεργασία για την Ανεμογεννήτρια 2

## PLANT 4

```

#      Column                                Non-Null Count  Dtype
---  -
0      Time                                164967 non-null  datetime64[ns]
1      Wind Ø [m/s]                       164967 non-null  float64
2      Wind max. [m/s]                    164967 non-null  float64
3      Wind min. [m/s]                    164967 non-null  float64
4      Rotation speed Ø [1/min]           164967 non-null  float64
5      Rotation speed max. [1/min]         164967 non-null  float64
6      Rotation speed min. [1/min]         164967 non-null  float64
7      Energy prod. [kWh].1                164967 non-null  float64
8      Blade angle Ø [°]                   164967 non-null  float64
9      Air humidity Ø [%]                  164967 non-null  float64
10     Nacelle position [°]                 164967 non-null  int64
dtypes: datetime64[ns](1), float64(9), int64(1)
memory usage: 15.1 MB
None

```

Αρνητικές τιμές ανά στήλη:

```

Wind Ø [m/s]                0
Wind max. [m/s]             0
Wind min. [m/s]             0
Rotation speed Ø [1/min]    0
Rotation speed max. [1/min] 0
Rotation speed min. [1/min] 0
Energy prod. [kWh].1        0
Blade angle Ø [°]          0
Air humidity Ø [%]         0
Nacelle position [°]        0
dtype: int64

```

Ακραίες τιμές (Μέγιστες και Ελάχιστες) ανά στήλη:

```

      Wind Ø [m/s]  Wind max. [m/s]  Wind min. [m/s]  \
min           0.0           0.0           0.0
max          30.5           39.7           22.3

      Rotation speed Ø [1/min]  Rotation speed max. [1/min]  \
min                0.00                0.00
max               21.02                28.33

      Rotation speed min. [1/min]  Energy prod. [kWh].1  Blade angle Ø [°]  \
min                0.00                0.0                0.0
max               20.67               463.0                94.0

      Air humidity Ø [%]  Nacelle position [°]
min                0.0                0
max               100.0               359

```

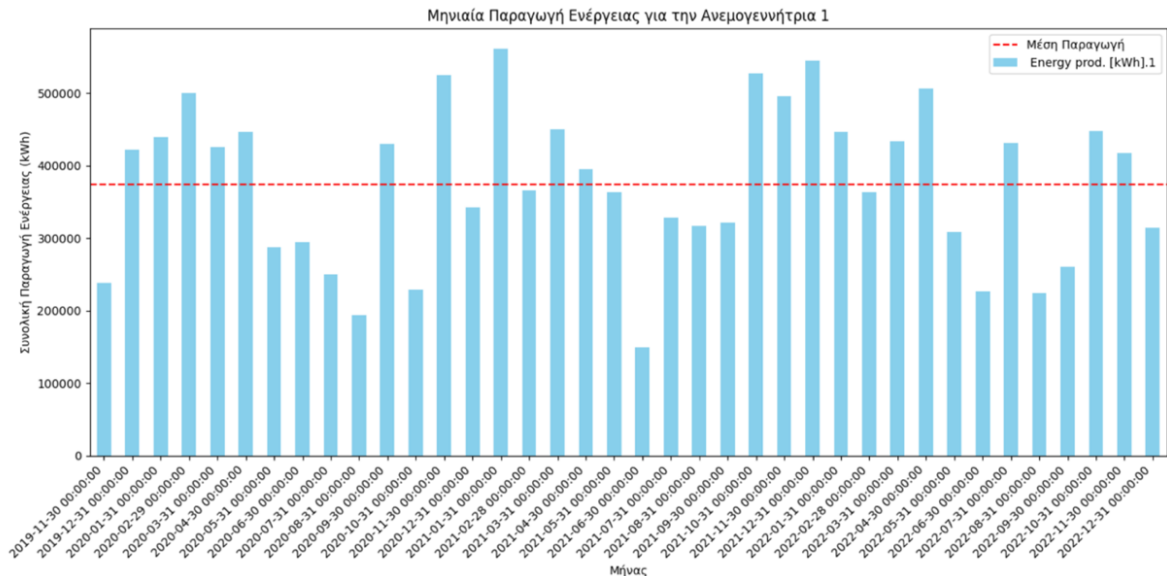
Εικόνα 38: Μορφή δεδομένων, αρνητικές και ακραίες τιμές ύστερα από την προ-επεξεργασία για την Ανεμογεννήτρια 4

### 6.3 ΜΕΛΕΤΗ ΕΠΟΧΙΑΚΟΤΗΤΑΣ - ΕΙΣΑΓΩΓΗ ΝΕΩΝ ΔΕΔΟΜΕΝΩΝ ΕΙΣΟΔΟΥ

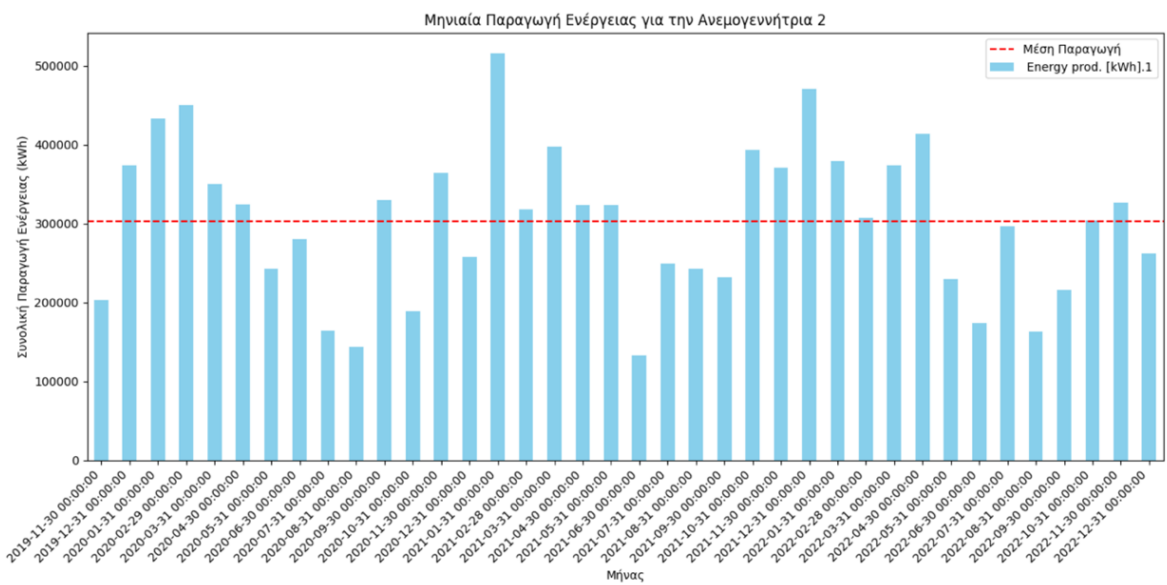
Σημαντικό κομμάτι για την ανάλυση δεδομένων στην συγκεκριμένη εργασία αποτελεί η μελέτη εποχιακότητας της ενεργειακής παραγωγής από τις ανεμογεννήτριες, καθώς και η προσθήκη ορισμένων νέων χαρακτηριστικών που βασίζονται σε αριθμητικά δεδομένα καιρού και μπορούν να ενισχύσουν την απόδοση πρόβλεψης. Συγκεκριμένα, μέσα από την ανάλυση της στήλης στόχου για κάθε ανεμογεννήτρια παρατηρήθηκε ότι υπάρχουν ορισμένα διαστήματα στα οποία η παραγωγή ενέργειας φαίνεται να εμφανίζει μεγαλύτερες διακυμάνσεις. Ειδικότερα, για κάθε ανεμογεννήτρια παρουσιάζεται η συνολική μηνιαία παραγωγή, ώστε να προκύψουν οι τιμές παραγωγής σε συγκεκριμένες εποχές του χρόνου που μπορεί να σχετίζονται με πιο έντονους ανέμους. Ταυτόχρονα



πραγματοποιήθηκε σύγκριση με την μηνιαία μέση παραγωγή, ώστε να διαπιστωθεί αν υπάρχουν περιόδοι με σταθερά υψηλότερη ή χαμηλότερη παραγωγή με ετήσια συνέπεια για κάθε μία από τις ανεμογεννήτριες.

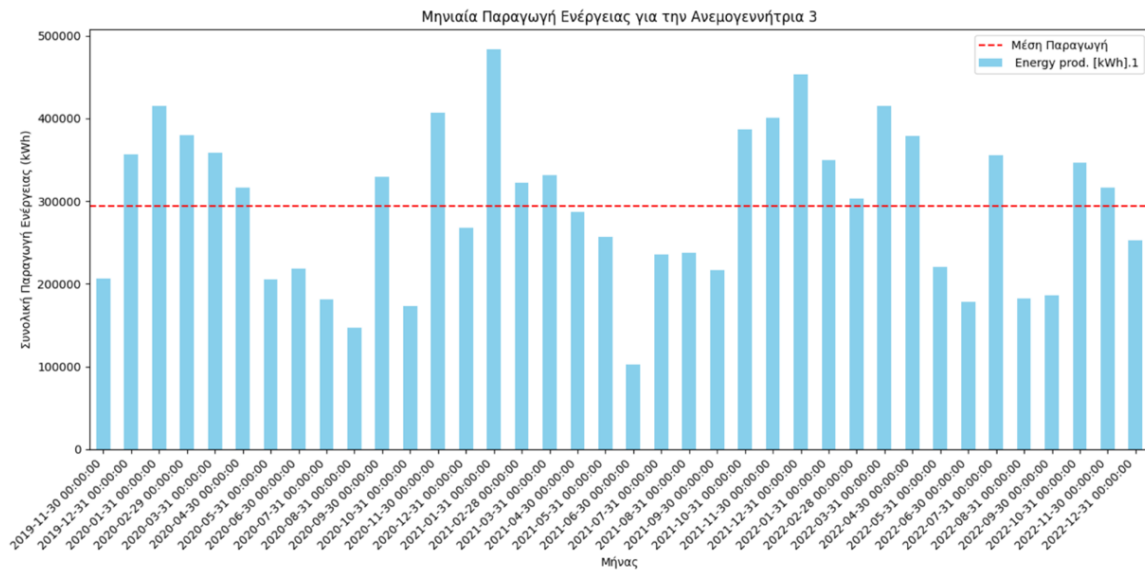


Εικόνα 39: Μηνιαία παραγωγή ενέργειας και Μέση παραγωγή ενέργειας Ανεμογεννήτρια 1

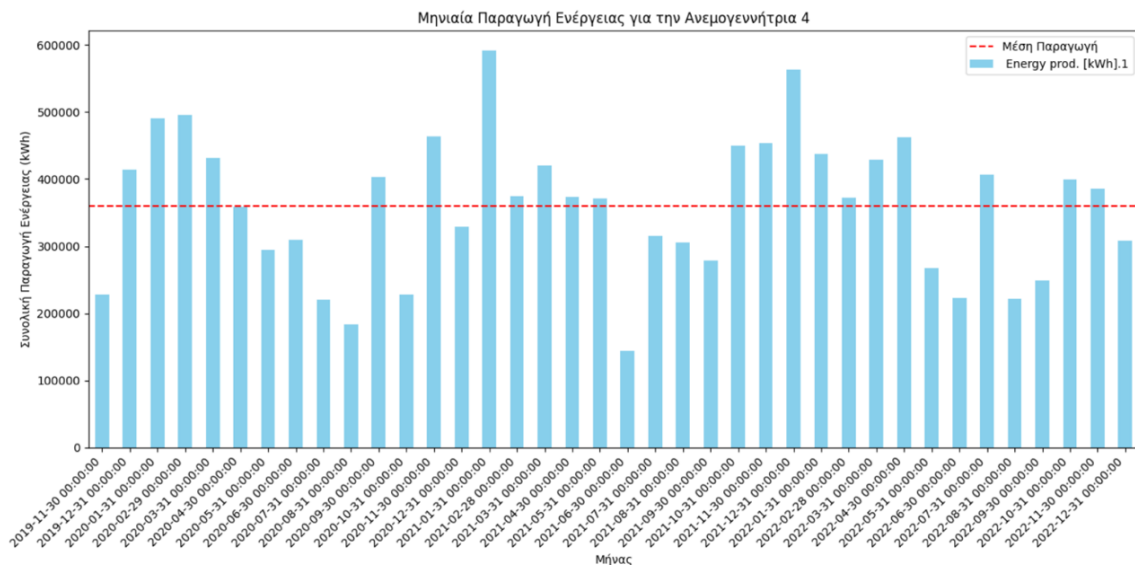


Εικόνα 40: Μηνιαία παραγωγή ενέργειας και Μέση παραγωγή ενέργειας Ανεμογεννήτρια 2





Εικόνα 41: Μηνιαία παραγωγή ενέργειας και Μέση παραγωγή ενέργειας Ανεμογεννήτρια 3



Εικόνα 42: Μηνιαία παραγωγή ενέργειας και Μέση παραγωγή ενέργειας Ανεμογεννήτρια 4

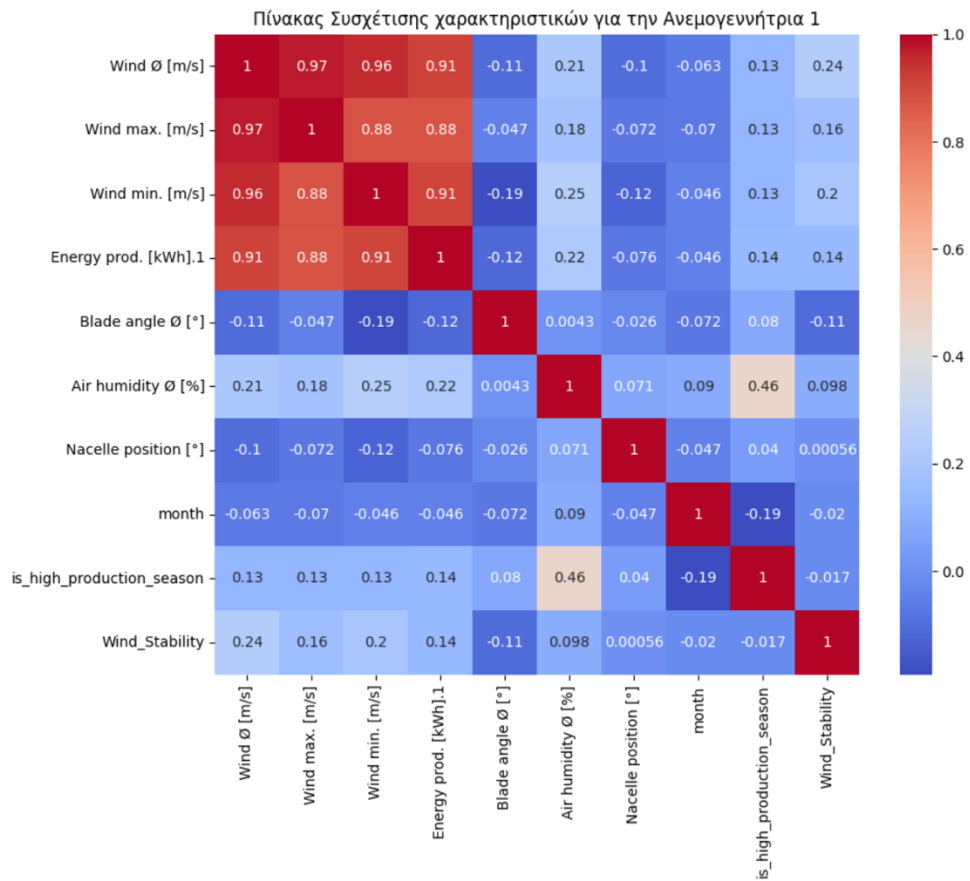
Μέσα από την ανάλυση των συγκεκριμένων διαγραμμάτων για κάθε ανεμογεννήτρια παρατηρούμε ορισμένα μοτίβα σαφούς εποχικότητας για συγκεκριμένους μήνες κάθε έτους που η παραγωγή είναι σταθερά υψηλή, ενώ παρατηρούνται και μήνες που παρουσιάζουν σταθερά μειωμένη παραγωγή. Πιο συγκεκριμένα, το Νοέμβριο, τους χειμερινούς μήνες, καθώς και για τους μήνες του Μαρτίου και του Απριλίου φαίνεται να υπάρχει μεγαλύτερη μηνιαία παραγωγή σε ετήσια βάση. Για να εισαχθούν τα συγκεκριμένα στοιχεία παρατηρούμενης εποχικότητας στο σύνολο των δεδομένων επιλέχθηκε να χρησιμοποιηθούν εποχιακές μεταβλητές στο σύνολο δεδομένων που υποδεικνύει τις εποχές υψηλής παραγωγής ενέργειας από τις ανεμογεννήτριες. Αυτές οι μεταβλητές μπορούν να βοηθήσουν το μοντέλο να "καταλάβει" και να λάβει υπόψιν τις εποχιακές διακυμάνσεις. Συνεπώς, ορίστηκε ως νέα στήλη στο σύνολο δεδομένων η δυαδική (binary) μεταβλητή 'is\_high\_production\_season' η οποία παίρνει την τιμή 1 για τους μήνες που έχουν καθοριστεί ως μήνες υψηλής παραγωγής και 0 για τους υπόλοιπους μήνες. Με αυτόν τον τρόπο παρέχεται στα μοντέλα ως είσοδος μια σαφής ένδειξη των περιόδων που αναμένεται να υπάρχει αυξημένη παραγωγή, γεγονός το

οποίο συνεισφέρει στην αποδοτικότητα των μοντέλων πρόβλεψης να λαμβάνουν υπόψη την εποχιακότητα παραγωγής και να κάνουν πιο ακριβείς προβλέψεις.

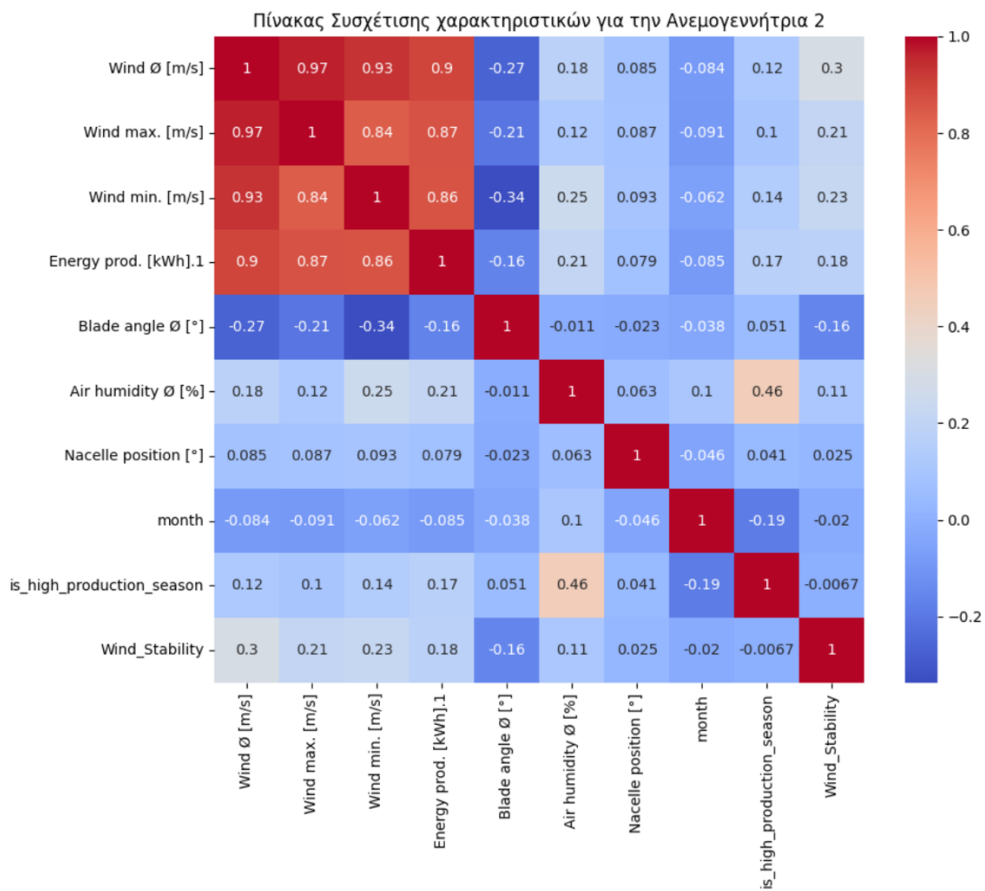
Επιπλέον, για την ενίσχυση των αποτελεσμάτων πρόβλεψης δημιουργήθηκε η στήλη δείκτη ευστάθειας ανέμου (*Wind Stability*), η οποία υπολογίζεται από τη σχέση της μέσης ταχύτητας ανέμου προς την μέγιστη και την ελάχιστη παρατηρούμενη ταχύτητα στο εκάστοτε διάστημα μέτρησης. Οι υψηλότερες τιμές υποδηλώνουν πιο σταθερές συνθήκες ανέμου, οι οποίες είναι ευνοϊκές για την παραγωγή ενέργειας. Ύστερα από δοκιμές εκπαίδευσης παρατηρήθηκε η ενίσχυση της απόδοσης και η μείωση των σφαλμάτων με τη χρήση των νέων χαρακτηριστικών. Συνεπώς τα δεδομένα εισόδου που χρησιμοποιήθηκαν για την εκπαίδευση των μοντέλων και την εξαγωγή προβλέψεων της στήλης στόχου *Energy prod. [kWh].1* είναι τα εξής: *Wind Ø [m/s]*, *Wind max. [m/s]*, *Wind min. [m/s]*, *Blade angle Ø [°]*, *Air humidity Ø [%]*, *Nacelle position [°]*, *month, is\_high\_production\_season* και *Wind Stability*.

## 6.4 ΣΥΣΧΕΤΙΣΕΙΣ

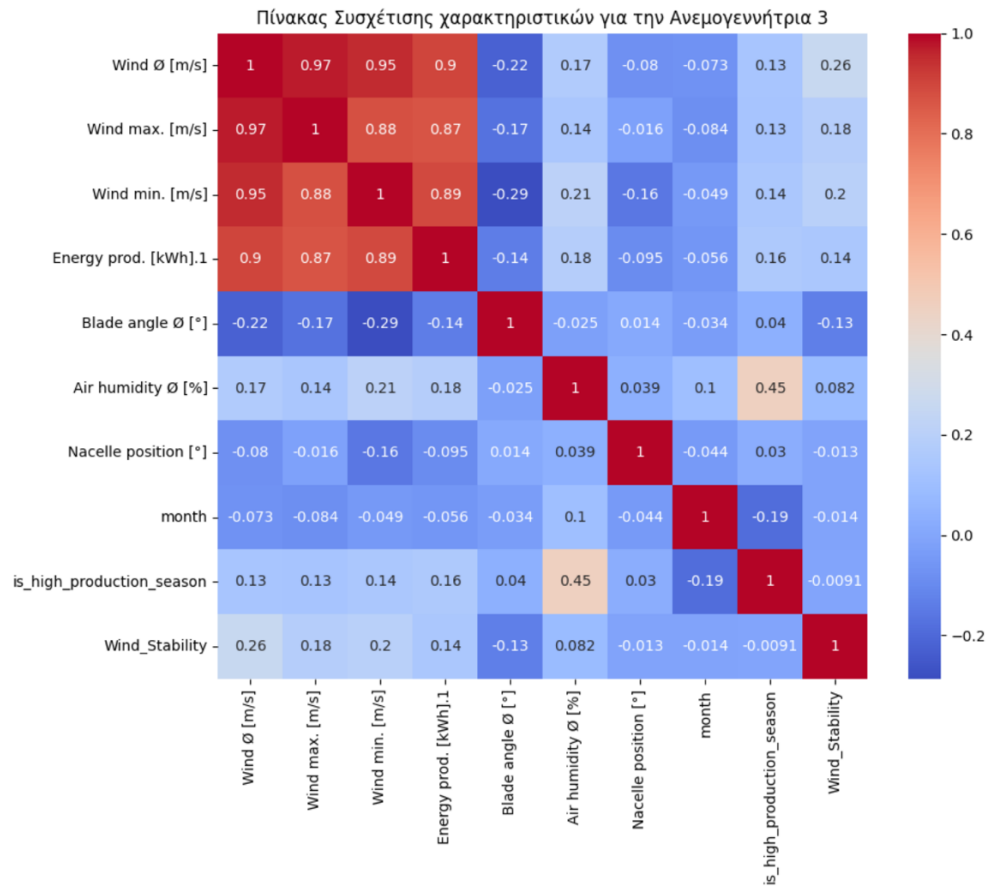
Στη συνέχεια πραγματοποιήθηκε εύρεση των γραμμικών συσχετίσεων των δεδομένων εισόδου. Από την ανάλυση των συσχετίσεων των χαρακτηριστικών με την παραγωγή ενέργειας είναι αντιληπτό ότι τα χαρακτηριστικά *Wind Ø [m/s]*, *Wind max. [m/s]*, *Wind min. [m/s]* εμφανίζουν τις ισχυρότερες θετικές συσχετίσεις αποσαφηνίζοντας ότι δεδομένα που αφορούν την ταχύτητα του ανέμου έχουν τον κρισιμότερο ρόλο στην ποσότητα της παραγόμενης ενέργειας. Επιπλέον, επιδραστικό ρόλο έχει και η ταχύτητα περιστροφής της ανεμογεννήτριας παρουσιάζοντας ισχυρές θετικές συσχετίσεις, ωστόσο επειδή δεν αποτελεί τιμή που προκύπτει από αριθμητικά δεδομένα καιρού δεν αποτελεί είσοδο για τα μοντέλα εξαγωγής πρόβλεψης. Αντιθέτως, τα χαρακτηριστικά *Blade angle Ø [°]*, *Air humidity Ø [%]*, *Nacelle position [°]*, *month, is\_high\_production\_season* και *Wind Stability* δείχνουν να έχουν χαμηλές συσχετίσεις με την παραγωγή ενέργειας. Εντούτοις, είναι σημαντικό να σημειωθεί ότι, παρά τις χαμηλές αυτές τιμές συσχέτισης, η αφαίρεση αυτών των χαρακτηριστικών από τα μοντέλα πρόβλεψης οδήγησε σε υποβάθμιση της ακρίβειας για αυτό από την στιγμή που υπάγονται σε μετεωρολογικά δεδομένα προβλέψεων και ενισχύουν τα μοντέλα πρόβλεψης επιλέχθηκε να χρησιμοποιηθούν και αυτά ως δεδομένα εισόδου.



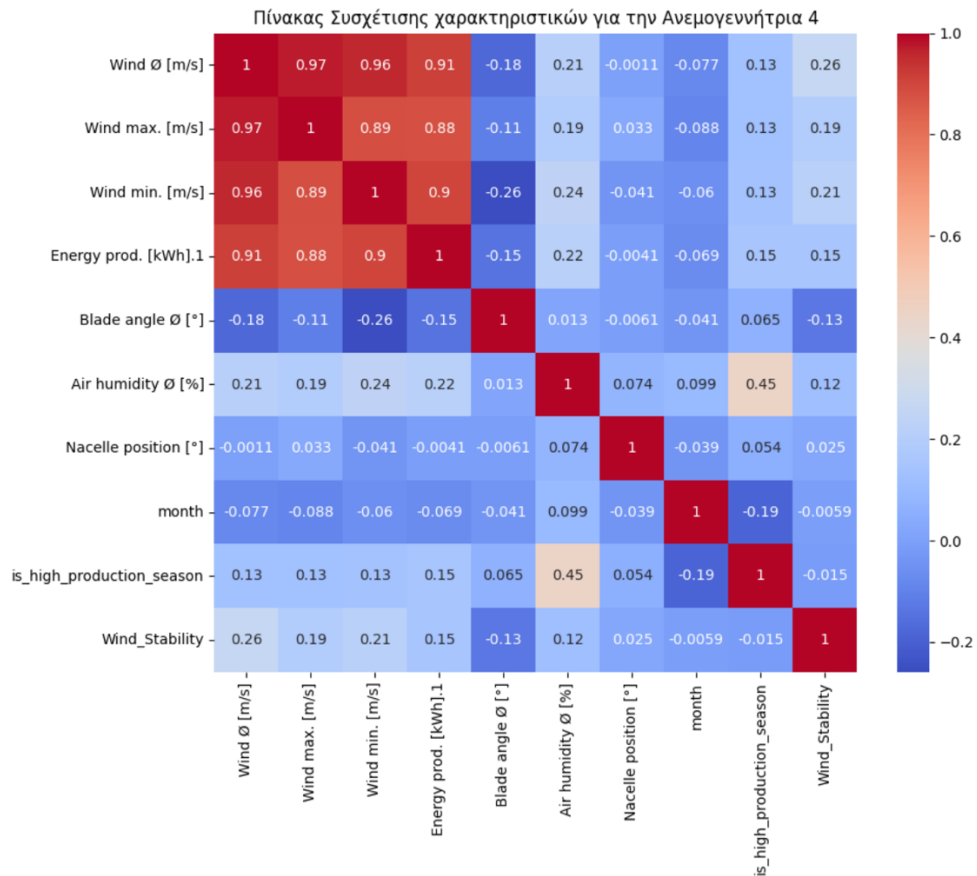
Εικόνα 43: Πίνακας συσχέτισης χαρακτηριστικών Ανεμογεννήτρια 1



Εικόνα 44: Πίνακας συσχέτισης χαρακτηριστικών Ανεμογεννήτρια 2



Εικόνα 45: Πίνακας συσχέτισης χαρακτηριστικών Ανεμογεννήτρια 3



Εικόνα 46: Πίνακας συσχέτισης χαρακτηριστικών Ανεμογεννήτρια 4

## 6.5 ΕΚΠΑΙΔΕΥΣΗ ΔΕΔΟΜΕΝΩΝ

Ακολουθεί η περιγραφή των αρχιτεκτονικών και των παραμέτρων των μοντέλων που χρησιμοποιήθηκαν μέσα από την χρήση της βιβλιοθήκης Keras, καθώς και τα αποτελέσματα εκπαίδευσης μέσω των επιλεγμένων μετρικών. Τα δεδομένα διαχωρίζονται σε σετ εκπαίδευσης και δοκιμής με αναλογία 70% - 30% του συνολικού dataset. Αυτό επιλέχθηκε ώστε το set δοκιμής να περιλαμβάνει το 30% του συνόλου δεδομένων που αντιστοιχεί σε ένα έτος. Επιπλέον, για όλα τα μοντέλα χρησιμοποιήθηκε κλιμάκωση δεδομένων MinMaxScaler για να κλιμακώσει τα δεδομένα στο εύρος [0, 1], το οποίο βοηθά στην βελτίωση της απόδοσης των μοντέλων.

### LSTM

Για το συγκεκριμένο αλγόριθμο πρώτο βήμα στη διαδικασία ήταν η αναδιαμόρφωση των δεδομένων σε μια μορφή συμβατή με τις απαιτήσεις εισόδου του LSTM, δηλαδή στη μορφή (samples, time steps, features). Αυτό επιτρέπει στο μοντέλο να αναλύσει τη χρονική σειρά των δεδομένων με τρόπο που αντικατοπτρίζει την χρονική τους συνέχεια, κρίσιμη για την πρόβλεψη των μελλοντικών τιμών. Συνολικά στο LSTM χρησιμοποιήθηκαν δύο στρώματα τα οποία βελτιώνουν την ακρίβεια του μοντέλου, καθώς επιτρέπουν μια πιο βαθιά εκμάθηση των χρονικών δυναμικών στα δεδομένα μέσα σε λογικά χρονικά πλαίσια εκπαίδευσης με συμβατούς υπολογιστικούς πόρους. Επιπλέον με την ύπαρξη 2 layers πραγματοποιείται και βελτίωση της ακρίβειας του μοντέλου, καθώς παρέχεται δυνατότητα για πιο βαθιά εκμάθηση των δεδομένων. Το πρώτο εισαγωγικό LSTM layer αποτελείται από 50 κόμβους και χρησιμοποιεί την συνάρτηση ενεργοποίησης ReLU ενισχύοντας την ικανότητα του μοντέλου να αντιμετωπίζει το πρόβλημα της εξαφάνισης των gradients, ενώ ταυτόχρονα ενισχύει την εκμάθηση από μεγάλες σειρές δεδομένων. Το δεύτερο LSTM layer, επίσης με 50 κόμβους και συνάρτηση ενεργοποίησης ReLU, δεν επιστρέφει ακολουθίες και προετοιμάζει τα δεδομένα για την τελική πρόβλεψη της ενεργειακής παραγωγής. Για τη βελτιστοποίηση χρησιμοποιήθηκε το Adam optimizer με ρυθμό μάθησης 0.01. Ο ρυθμός μάθησης με τιμή 0.01 επιλέχθηκε μετά από δοκιμές και αποδείχθηκε αποτελεσματικός στην επίτευξη γρήγορης σύγκλισης χωρίς να θυσιάζεται η σταθερότητα της εκπαίδευσης. Χαμηλότερες τιμές θα μπορούσαν να καταστήσουν την εκπαίδευση πιο αργή, ενώ υψηλότερες τιμές θα μπορούσαν να οδηγήσουν σε αστάθεια. Ως συνάρτηση απώλειας επιλέχθηκε το Mean Squared Error (MSE), που είναι κατάλληλο για προβλέψεις συνεχών τιμών. Η εκπαίδευση διεξήχθη για 50 εποχές με batch size 64, διατηρώντας τη χρονική σειρά των δεδομένων (shuffle=False), γεγονός που είναι κρίσιμο για τη διατήρηση της χρονικής συνέχειας στην εκπαίδευση χρονοσειρών.

Στην συνέχεια παρουσιάζονται τα αποτελέσματα εκπαίδευσης για εξαγωγή προβλέψεων του συγκεκριμένου μοντέλου για τις 4 ανεμογεννήτριες :

LSTM				
	<i>PLANT 1</i>	<i>PLANT 2</i>	<i>PLANT 3</i>	<i>PLANT 4</i>
MSE	247.6	182.818	123.542	162.739
RMSE	15.735	13.521	11.114	12.756
MAE	10.468	8.395	5.762	7.332
R <sup>2</sup>	0.977	0.975	0.985	0.983

Πίνακας 12: Αποτελέσματα εκπαίδευσης LSTM

## Bi-LSTM

Όπως έχει αναφερθεί και στο θεωρητικό υπόβαθρο τα μοντέλα Bi-LSTM επεκτείνουν την ικανότητα των κλασικών LSTM επιτρέποντας στο νευρωνικό δίκτυο να λάβει υπόψη τόσο παρελθοντικές όσο και μελλοντικές πληροφορίες για κάθε σημείο της ακολουθίας. Αυτό μπορεί να αυξήσει την ακρίβεια του μοντέλου σε πολλές εφαρμογές. Αντίστοιχη αναδιαμόρφωση δεδομένων πραγματοποιείται ώστε να υπάρχει συμβατότητα στη μορφή εισόδου που απαιτείται. Η δομή του μοντέλου Bi-LSTM κατά αντιστοιχία περιλαμβάνει δύο bidirectional LSTM layers με 50 κόμβους το καθένα και χρησιμοποιώντας τη συνάρτηση ενεργοποίησης ReLU. Το Bi-LSTM είναι διαμορφωμένο να επιστρέφει ακολουθίες, επιτρέποντας στην πληροφορία να ρέει σε και από τα δύο κατευθύνσεις (προς το παρελθόν και το μέλλον) για κάθε σημείο της ακολουθίας. Η βελτιστοποίηση γίνεται με τη χρήση του Adam optimizer και με ρυθμό μάθησης 0.01, ενώ η συνάρτηση απώλειας είναι το Mean Squared Error (MSE). Το μοντέλο εκπαιδεύεται για 50 εποχές με batch size 64, με την επιλογή shuffle=False για τη διατήρηση της χρονικής σειράς των δεδομένων. Το συγκεκριμένο μοντέλο φαίνεται να προσφέρει σημαντική βελτίωση στην ικανότητα πρόβλεψης, επωφελούμενο από την πληροφορία που προέρχεται και από τις δύο κατευθύνσεις της χρονικής ακολουθίας.

Στην συνέχεια παρουσιάζονται τα αποτελέσματα εκπαίδευσης για εξαγωγή προβλέψεων του συγκεκριμένου μοντέλου για τις 4 ανεμογεννήτριες :

<b>Bi - LSTM</b>				
	<i>PLANT 1</i>	<i>PLANT 2</i>	<i>PLANT 3</i>	<i>PLANT 4</i>
MSE	178.366	172.222	118.053	182.21
RMSE	13.355	13.123	10.865	13.498
MAE	8.046	8.163	5.608	7.777
R <sup>2</sup>	0.983	0.977	0.985	0.981

Πίνακας 13: Αποτελέσματα εκπαίδευσης Bi-LSTM

## GRU

Η αναδιαμόρφωση των δεδομένων στην κατάλληλη μορφή είναι μια απαιτούμενη διαδικασία και για το μοντέλο GRU. Η δομή του μοντέλου GRU περιλαμβάνει ένα στρώμα GRU με 50 κόμβους, χρησιμοποιώντας τη συνάρτηση ενεργοποίησης ReLU, και ένα Dense layer που πραγματοποιεί την τελική πρόβλεψη της παραγωγής ενέργειας. Η επιλογή του GRU στρώματος αντικατοπτρίζει την προσπάθεια για αποδοτική εκμάθηση των χρονικών δυναμικών των δεδομένων, προσφέροντας μια πιο απλουστευμένη δομή σε σύγκριση με τα αντίστοιχα νευρωνικά δίκτυα LSTM και Bi-LSTM, αλλά διατηρώντας την ικανότητα να αντιμετωπίζει το πρόβλημα της εξαφάνισης των gradients. Η βελτιστοποίηση του μοντέλου γίνεται μέσω του Adam optimizer με ρυθμό μάθησης 0.01, που αποδείχθηκε αποτελεσματικός στην επίτευξη γρήγορης σύγκλισης ενώ χρησιμοποιήθηκε το Mean Squared Error (MSE) ως συνάρτηση απώλειας. Η εκπαίδευση διεξάγεται σε 50 εποχές με batch size 64. Αξίζει να τονιστεί η ταχύτητα εκμάθησης του μοντέλου, συγκριτικά με τα αντίστοιχα LSTM και Bi-LSTM, η οποία παρατηρήθηκε να είναι σημαντικά ταχύτερη.

Στην συνέχεια παρουσιάζονται τα αποτελέσματα εκπαίδευσης για εξαγωγή προβλέψεων του συγκεκριμένου μοντέλου για τις 4 ανεμογεννήτριες :



GRU				
	<i>PLANT 1</i>	<i>PLANT 2</i>	<i>PLANT 3</i>	<i>PLANT 4</i>
MSE	297.124	159.587	121.649	170.914
RMSE	17.237	12.632	11.029	13.073
MAE	10.717	7.72	5.698	7.5
R <sup>2</sup>	0.973	0.978	0.985	0.982

Πίνακας 14: Αποτελέσματα εκπαίδευσης GRU

## **RANDOM FOREST**

Το μοντέλο Random Forest χρησιμοποιεί πολλαπλά δέντρα αποφάσεων για να προβλέψει την τιμή της παραγωγής ενέργειας. Στο συγκεκριμένο μοντέλο πραγματοποιήθηκαν πολλές δοκιμές στις διαφορετικές παραμέτρους ώστε να μεγιστοποιηθεί η απόδοσή του. Η δομή του συγκεκριμένου μοντέλου Random Forest στο οποίο εφαρμόστηκε tuning parameter χαρακτηρίζεται από την επιλογή των παραμέτρων, όπως ο αριθμός των δένδρων ( $n\_estimators=100$ ), το μέγιστο βάθος των δένδρων ( $max\_depth=30$ ), τον ελάχιστο αριθμό δειγμάτων που είναι απαιτούμενος για να διασπαστεί ένας κόμβος ( $min\_samples\_split=15$ ), τον ελάχιστο αριθμό δειγμάτων που είναι απαιτούμενος σε έναν τερματικό κόμβο φύλλου ( $min\_samples\_leaf=4$ ), και ο τρόπος επιλογής των χαρακτηριστικών σε κάθε διαίρεση ( $max\_features=N/3$ ). Ειδικότερα για την παράμετρο  $max\_features$  προτιμήθηκε η επιλογή του  $N/3$  όπου  $N$  ο αριθμός των χαρακτηριστικών. Η επιλογή του  $bootstrap=true$  επιτρέπει την ανάπτυξη κάθε δέντρου σε διαφορετικά υποσύνολα των δεδομένων, αυξάνοντας τη γενικευτική ικανότητα του μοντέλου.

Στην συνέχεια παρουσιάζονται τα αποτελέσματα εκπαίδευσης για εξαγωγή προβλέψεων του συγκεκριμένου μοντέλου για τις 4 ανεμογεννήτριες :

RANDOM FOREST				
	<i>PLANT 1</i>	<i>PLANT 2</i>	<i>PLANT 3</i>	<i>PLANT 4</i>
MSE	88.733	270.228	107.356	77.446
RMSE	9.419	16.438	10.361	8.8
MAE	5.244	8.251	4.269	4.912
R <sup>2</sup>	0.991	0.964	0.987	0.992

Πίνακας 15: Αποτελέσματα εκπαίδευσης Random Forest

## **XGBOOST**

Το μοντέλο XGBoost (eXtreme Gradient Boosting) αποτελεί μια προηγμένη υλοποίηση των αλγορίθμων ενισχυτικής μάθησης (boosting algorithms), ο οποίος χρησιμοποιεί πολλαπλά δέντρα αποφάσεων για να παράγει πιο ακριβείς προβλέψεις. Το XGBoost βελτιστοποιεί συνεχώς τις προβλέψεις του μέσω της διαδικασίας του gradient boosting, δηλαδή προσθέτει δέντρα ένα προς ένα προσπαθώντας να μειώσει το σφάλμα των προηγούμενων δέντρων. Στο συγκεκριμένο μοντέλο πραγματοποιήθηκαν πολλές δοκιμές σε διαφορετικές κρίσιμες παραμέτρους ώστε να μεγιστοποιηθεί η απόδοσή του. Η επιλογή των τελικών τιμών για τις παραμέτρους πραγματοποιήθηκε ύστερα από δοκιμές αλλά και μέσω προκαθορισμένων τιμών. Ειδικότερα σε παραμέτρους όπως ο αριθμός των δένδρων ( $n\_estimators=100$ ), το μέγιστο βάθος δέντρου ( $max\_depth=30$ ), ο ελάχιστος αριθμός δειγμάτων που χρειάζεται για να δημιουργηθεί ένας νέος κόμβος ( $min\_child\_weight=1$ ) και ο ρυθμός μάθησης ( $learning\_rate=0.1$ ) δοκιμάστηκαν διαφορετικές τιμές που ελαχιστοποιούν τα σφάλματα πρόβλεψης [84]. Επιπλέον, αντίστοιχη διαδικασία δοκιμάστηκε και στο εύρος τιμών δοκιμής

που προτείνεται για τις παραμέτρους subsample και colsample\_bytree που ορίζουν το ποσοστό των δειγμάτων και χαρακτηριστικών που χρησιμοποιούνται για κάθε δέντρο. Επιπλέον, χρησιμοποιήθηκε και η παράμετρος gamma, η οποία ορίζει την ελάχιστη μείωση στην απώλεια που απαιτείται για να γίνει περαιτέρω διαίρεση ενός κόμβου στο δέντρο και βοηθάει στην αποφυγή της υπερεκπαίδευσης.

Στην συνέχεια παρουσιάζονται τα αποτελέσματα εκπαίδευσης για εξαγωγή προβλέψεων του συγκεκριμένου μοντέλου για τις 4 ανεμογεννήτριες :

<b>XGBOOST</b>				
	<i>PLANT 1</i>	<i>PLANT 2</i>	<i>PLANT 3</i>	<i>PLANT 4</i>
MSE	82.815	231.892	81.198	66.067
RMSE	9.1	15.228	9.011	8.128
MAE	5.24	8.11	4.197	4.852
R <sup>2</sup>	0.992	0.969	0.99	0.993

Πίνακας 16: Αποτελέσματα εκπαίδευσης XGBoost

## **META-LEARNING METHODS**

Η μέθοδος μέτα-μάθησης, γνωστή και ως stacking ensemble method είναι η διαδικασία σύμφωνα με την οποία ένα μοντέλο μηχανικής μάθησης καλείται να βελτιώσει την ακρίβεια εξαγωγής συγκεκριμένων βασικών προβλέψεων, χρησιμοποιώντας ως δεδομένα (meta-data) τις προβλέψεις άλλων μοντέλων μηχανικής μάθησης. Αυτή η διαδικασία παρουσιάζει ιδιαίτερο ενδιαφέρον σε προβλήματα εξαγωγής προβλέψεων, όπου διαφορετικά μοντέλα πρόβλεψης με διάφορα χαρακτηριστικά, δομές και αρχιτεκτονικές εμφανίζουν ετερότητα στις επιδόσεις. Ορισμένα μοντέλα μπορεί να έχουν καλύτερη απόδοση σε ορισμένους εποχιακούς δείκτες, καιρικές συνθήκες ή και στις ώρες αιχμής, ωστόσο μπορεί να μην αποδίδουν το ίδιο καλά σε ευμετάβλητες ή μη γραμμικές συσχετίσεις ή και σε ώρες χαμηλής παραγωγής. Ο συνδυασμός των βασικών προβλέψεων που συχνά αναφέρονται ως προβλέψεις επιπέδου-0 και εκείνων που παράγονται από τον μέτα-εκπαιδευτή και καλούνται ως προβλέψεις επιπέδου-1, κατανοούν τις συγκεκριμένες συνθήκες κάτω από τις οποίες ο συνδυασμός των βασικών μοντέλων εμφανίζει καλύτερα αποτελέσματα, δίνοντας έμφαση συχνά σε γραμμικές και μη γραμμικές συσχετίσεις.

Στο συγκεκριμένο πρόβλημα ένας από τους μέτα-εκπαιδευτές που χρησιμοποιήθηκε βασίστηκε στη μέθοδο SVR. Ειδικότερα, ο συγκεκριμένος είναι ένας ευρέως διαδεδομένος αλγόριθμος μηχανικής μάθησης που εκτελεί παλινδρόμηση με τρόπο παρόμοιο με τη συμβατική γραμμική παλινδρόμηση, στον βαθμό που η γραμμή που χρησιμοποιείται για τις προβλέψεις είναι της μορφής :  $y = wX + b$ , όπου  $y$  είναι η μεταβλητή-στόχος,  $X$  είναι το σύνολο των χαρακτηριστικών που χρησιμοποιούνται ως μεταβλητές, το  $w$  αφορά το διάνυσμα βαρών και  $b$  είναι μια σταθερά που λαμβάνει υπόψη το bias. Ωστόσο, εμφανίζει διαφοροποίηση διότι η γραμμή πρόβλεψης αποτελεί ένα υπερεπίπεδο και τα δεδομένα σε κάθε πλευρά του που εμφανίζουν τη μικρότερη απόσταση από το υπερεπίπεδο ονομάζονται διανύσματα υποστήριξης. Η συγκεκριμένη μέθοδος στοχεύει στην εύρεση και προσαρμογή της ιδανικής γραμμής εντός μιας οριακής τιμής  $\epsilon$ , η οποία περιγράφει την απόσταση μεταξύ του υπερεπίπεδου και των διανυσμάτων υποστήριξης. Ικανοποιεί τη συνθήκη  $|y - (wX + b)| < \epsilon$ . Στη συγκεκριμένη περίπτωση ορίστηκε ως συνάρτηση πυρήνα η Radial Basis Function (RBF) για να μπορέσουν να χαρτογραφηθούν τα δείγματα δεδομένων σε έναν χώρο υψηλής διάστασης και να επιτρέπεται η μη γραμμική παλινδρόμηση. Επιπλέον, στη μελέτη αυτή δοκιμάστηκε και ο XGBoost Regressor που αξιοποιεί την αρχιτεκτονική των gradient boosted trees και λειτουργεί αντίστοιχα με το βασικό μοντέλο που χρησιμοποιήθηκε για την εξαγωγή βασικών προβλέψεων. Επιλέχθηκε λόγω της υψηλής υπολογιστικής ταχύτητας



συγκριτικά με τα υπόλοιπα μοντέλα μηχανικής μάθησης. Ο συγκεκριμένος μετα-εκπαιδευτής προσφέρει ενίσχυση της προβλεπτικής ικανότητας μέσω της σταδιακής μείωσης του σφάλματος, παρέχοντας έτσι μια εναλλακτική προσέγγιση στη σύνθεση προβλέψεων από διάφορα βασικά μοντέλα. Ο SVR και ο XGBoost regressor προτιμήθηκαν σε σχέση με άλλα μοντέλα meta-learning λόγω της ανίχνευσης μη γραμμικών σχέσεων μεταξύ των χαρακτηριστικών. Παρατηρήθηκε σημαντική βελτίωση των επιδόσεων και εξαιρετική ικανότητα γενίκευσης στα δεδομένα εκπαίδευσης. Συνολικά οι μέθοδοι μετα-εκπαίδευσης φαίνεται να υπερτερούν και να μειώνουν σημαντικά τα σφάλματα πρόβλεψης των βασικών μοντέλων.

Στην συνέχεια παρουσιάζονται τα αποτελέσματα εκπαίδευσης για εξαγωγή προβλέψεων των συγκεκριμένων μοντέλων για τις 4 ανεμογεννήτριες :

META-LEARNING METHODS								
	PLANT 1		PLANT 2		PLANT 3		PLANT 4	
	SVR	XGBOOST	SVR	XGBOOST	SVR	XGBOOST	SVR	XGBOOST
MSE	65.495	49.764	107.258	65.191	75.92	33.314	66.31	48.629
RMSE	8.092	7.054	10.356	8.074	8.713	5.771	8.143	6.973
MAE	4.628	4.116	5.67	4.76	4.05	3.432	4.801	4.169
R <sup>2</sup>	0.994	0.995	0.985	0.991	0.99	0.996	0.993	0.995

Πίνακας 17: Αποτελέσματα εκπαίδευσης μεθόδων meta-learning

### ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΚΑΘΕ ΑΝΕΜΟΓΕΝΝΗΤΡΙΑ

Συγκεντρωτικά παρουσιάζονται τα αποτελέσματα των μετρικών για κάθε ένα από τα βασικά μοντέλα αλλά και από τα μοντέλα meta-learning που χρησιμοποιήθηκαν σε κάθε μια από τις 4 ανεμογεννήτριες στο σύνολο δεδομένων με σετ εκπαίδευσης 70% και σετ δοκιμής 30%.

PLANT 1	LSTM	Bi - LSTM	GRU	RF	XGBOOST	META - SVR	META - XGBOOST
MSE	247.6	178.366	297.124	88.733	82.815	65.495	49.764
RMSE	15.735	13.355	17.237	9.419	9.1	8.092	7.054
MAE	10.468	8.046	10.717	5.244	5.24	4.628	4.116
R <sup>2</sup>	0.977	0.983	0.973	0.991	0.992	0.994	0.995

Πίνακας 18: Συγκεντρωτικά αποτελέσματα για την Ανεμογεννήτρια 1

PLANT 2	LSTM	Bi - LSTM	GRU	RF	XGBOOST	META - SVR	META - XGBOOST
MSE	182.818	172.222	159.587	270.228	231.892	107.258	65.191
RMSE	13.521	13.123	12.632	16.438	15.228	10.356	8.074
MAE	8.395	8.163	7.72	8.251	8.11	5.67	4.76
R <sup>2</sup>	0.975	0.977	0.978	0.964	0.969	0.985	0.991

Πίνακας 19: Συγκεντρωτικά αποτελέσματα για την Ανεμογεννήτρια 2

PLANT 3	LSTM	Bi - LSTM	GRU	RF	XGBOOST	META - SVR	META - XGBOOST
MSE	123.542	118.053	121.649	107.356	81.198	75.92	33.314
RMSE	11.114	10.865	11.029	10.361	9.011	8.713	5.771
MAE	5.762	5.608	5.698	4.269	4.197	4.05	3.432
R <sup>2</sup>	0.985	0.985	0.985	0.987	0.99	0.99	0.996

Πίνακας 20: Συγκεντρωτικά αποτελέσματα για την Ανεμογεννήτρια 3

PLANT 4	LSTM	Bi - LSTM	GRU	RF	XGBOOST	META - SVR	META - XGBOOST
MSE	162.739	182.21	170.914	77.446	66.067	66.31	48.629
RMSE	12.756	13.498	13.073	8.8	8.128	8.143	6.973
MAE	7.332	7.777	7.5	4.912	4.852	4.801	4.169
R <sup>2</sup>	0.983	0.981	0.982	0.992	0.993	0.993	0.995

Πίνακας 21: Συγκεντρωτικά αποτελέσματα για την Ανεμογεννήτρια 4

Με βάση τα αποτελέσματα της εκπαίδευσης η οποία παρέχει μια γενική εικόνα της απόδοσης των μοντέλων σε ολόκληρο το σετ δεδομένων ελέγχου οι μέθοδοι meta-learning εμφανίζουν μεγάλη βελτίωση στα αποτελέσματα των μετρικών. Η συγκεκριμένη ανάλυση των συνολικών αποτελεσμάτων ελέγχου αποκαλύπτει το πόσο καλά ένα από τα μοντέλα μπορεί

να γενικεύει με επιτυχία σε ολόκληρο το σετ δεδομένων. Οι προτεινόμενοι μέθοδοι meta-learning φαίνεται να έχουν πολύ καλή γενίκευση και εξαιρετική προσαρμογή στα δεδομένα γεγονός που τα καθιστά ιδιαίτερα αποτελεσματικά για την πρόβλεψη παραγωγής ενέργειας. Επιπλέον, τα μοντέλα που βασίζονται σε νευρωνικά δίκτυα φαίνεται να έχουν καλή γενίκευση και προσαρμογή στα δεδομένα αν και εμφανίζουν μεγαλύτερα σφάλματα συγκριτικά με τα μοντέλα μετά-εκπαίδευσης. Το Bi-LSTM φαίνεται να έχει συχνά την καλύτερη απόδοση με μικρή διαφορά από το μοντέλο LSTM, ενώ εντυπωσιακά αποτελέσματα φαίνεται να έχει και το GRU το οποίο υπολογιστικά είναι ταχύτερη επιλογή συγκριτικά με τα υπόλοιπα μοντέλα νευρωνικών δικτύων. Τέλος, οι μέθοδοι των Random Forests και XGBoost παρουσιάζουν μικρά σφάλματα εκτίμησης παραγωγής στο συνολικό σετ εκπαίδευσης. Επιπλέον, σε ορισμένες ανεμογεννήτριες τα αποτελέσματα τους φαίνεται να είναι ανταγωνιστικά και με τις μεθόδους μετά-εκπαίδευσης, ωστόσο παρουσιάζονται και αποτελέσματα στα οποία έχουν την μεγαλύτερη απόκλιση συγκριτικά με τις πραγματικές τιμές παραγωγής, όπως στην Ανεμογεννήτρια 2. Αξίζει να τονιστεί οι βασικοί μέθοδοι πρόβλεψης ανεξάρτητα από τα σφάλματα που αναδεικνύονται στη διαδικασία της εκπαίδευσης, φαίνεται να ενισχύουν την απόδοση αποτελεσμάτων των συνδυαστικών μεθόδων. Ύστερα από δοκιμές με την ενσωμάτωση και μη των συγκεκριμένων μεθόδων στα συνδυαστικά μοντέλα μετά-εκπαίδευσης, παρατηρήθηκε η μείωση των επιδόσεων πρόβλεψης μέσω της μη ενσωμάτωσης ακόμη και των πιο ‘αδύναμων’ μοντέλων. Συνεπώς, φαίνεται ότι οι αυτόνομοι μέθοδοι παρόλο που δεν είχαν την καλύτερη απόδοση, μπορούν να συνεισφέρουν σημαντικά στη βελτίωση των συνδυαστικών μοντέλων μετά-εκπαίδευσης, κάτι το οποίο αποτελεί ένα ενδιαφέρον εύρημα.

### 6.5.1 ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ ΔΙΑΣΤΗΜΑΤΟΣ 2 ΩΡΩΝ

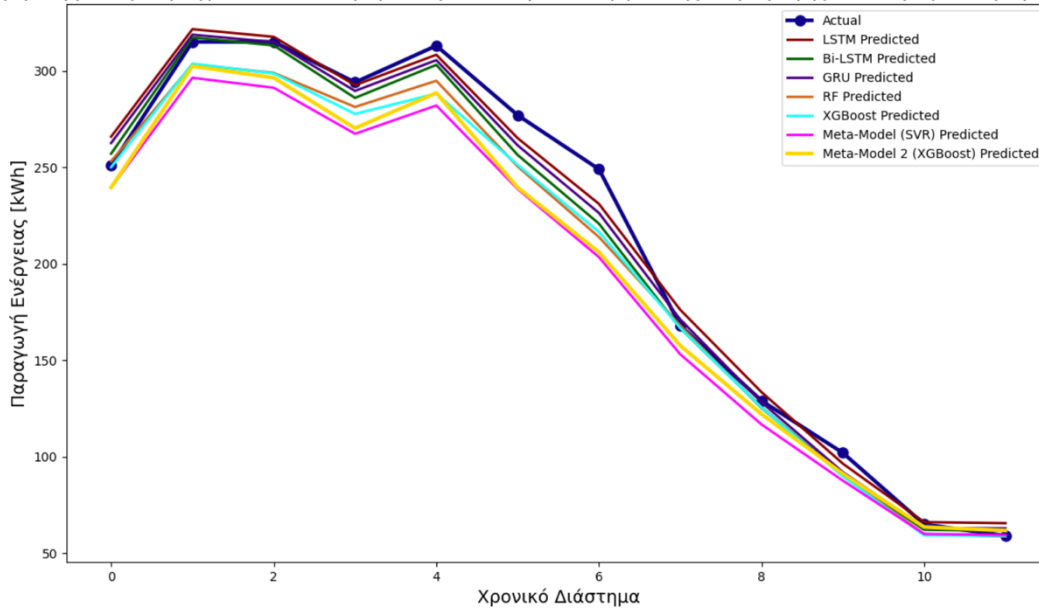
Για την περαιτέρω ανάλυση της συμπεριφοράς των συγκεκριμένων μοντέλων πρόβλεψης πραγματοποιήθηκε και η εξαγωγή βραχυπρόθεσμων αποτελεσμάτων πρόβλεψης. Στην συνέχεια θα παρουσιαστούν τα αποτελέσματα δοκιμής σε βραχυπρόθεσμη πρόβλεψη της ενεργειακής παραγωγής για κάθε μια από τις ανεμογεννήτριες χρησιμοποιώντας τις υφιστάμενες μεθόδους. Επισημαίνεται ότι το βήμα μέτρησης και καταγραφής αφορά βήμα 10 λεπτών και τα αποτελέσματα παρουσιάζουν την εκτιμώμενη ενεργειακή παραγωγή σε βάθος 12 βημάτων δηλαδή 2 ωρών (120 λεπτών). Παρουσιάζονται τα αποτελέσματα από τις μετρικές για να εξαχθούν οι κατάλληλες πληροφορίες απόδοσης των μεθόδων για το πρόβλημα της βραχυπρόθεσμης εκτιμώμενης ενεργειακής παραγωγής. Επιπλέον, το διάγραμμα πρόβλεψης παρέχεται για την ταυτόχρονη οπτικοποίηση όλων των αποτελεσμάτων που εξάχθηκαν από τα μοντέλα σε σχέση με τις πραγματικές τιμές ενεργειακής παραγωγής. Τέλος, παρουσιάζεται και το αποτέλεσμα πρόβλεψης της μεθόδου που εμφάνισε τα καλύτερα αποτελέσματα πρόβλεψης σε κάθε μια από τις ανεμογεννήτριες.

ΑΝΕΜΟΓΕΝΝΗΤΡΙΑ 1

PLANT 1	LSTM	Bi - LSTM	GRU	RF	XGBOOST	META - SVR	META - XGBOOST
MSE	76.743	132.146	92.434	251.042	266.342	571.884	444.6178
RMSE	8.76	11.495	9.614	15.844	16.319	23.914	21.085
MAE	7.159	8.202	7.101	11.961	12.571	20.196	16.977
R <sup>2</sup>	0.991	0.985	0.990	0.972	0.971	0.938	0.952

Πίνακας 22: Αποτελέσματα βραχυπρόθεσμης πρόβλεψης Ανεμογεννήτριας 1

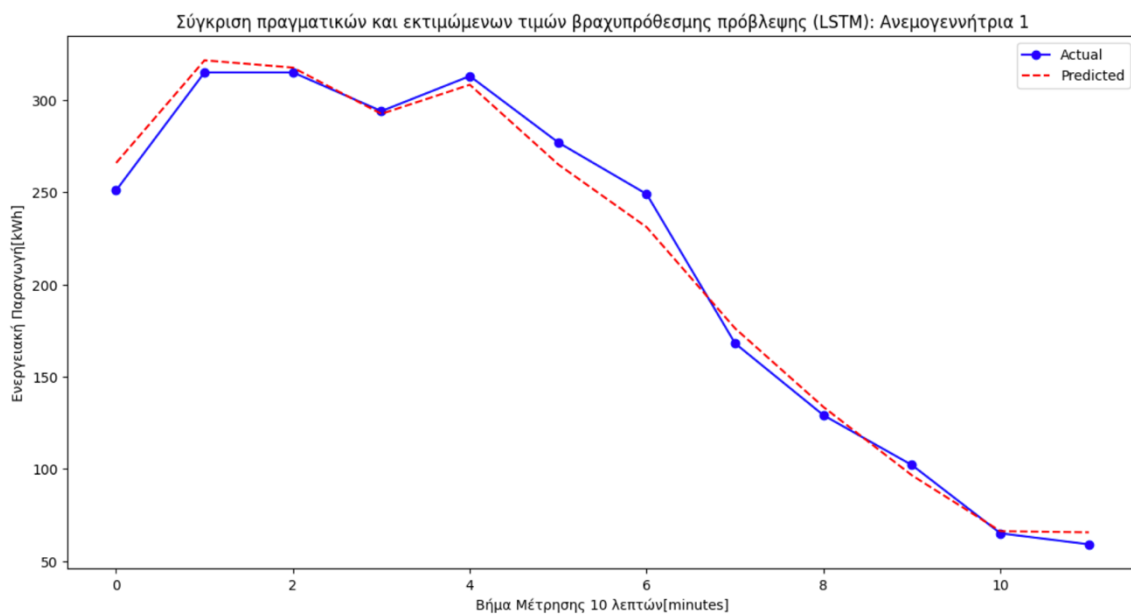
Σύγκριση μεταξύ Πραγματικών και Προβλεπόμενων Τιμών Ενεργειακής Παραγωγής σε 2 ώρες : Ανεμογεννήτρια 1



Εικόνα 47: Σύγκριση συνολικών αποτελεσμάτων πρόβλεψης με πραγματικές τιμές

Τα μοντέλα LSTM , και GRU έχουν την καλύτερη απόδοση σύμφωνα με όλες τις μετρικές και προσεγγίζουν με ιδιαίτερη επιτυχία τις πραγματικές τιμές ενεργειακής παραγωγής. Ειδικότερα στην συγκεκριμένη περίπτωση το LSTM αποτελεί το ιδανικότερο μοντέλο για την εξαγωγή της βραχυπρόθεσμης πρόβλεψης και με μικρή διαφορά ακολουθούν τα υπόλοιπα μοντέλα νευρωνικών δικτύων. Τα μοντέλα Random Forest (RF) και XGBoost εμφανίζουν την μεγαλύτερη υστέρηση στην απόδοση πρόβλεψης, ωστόσο φαίνεται να βελτιώνουν τα αποτελέσματα των συνδυαστικών μεθόδων όταν εισάγονται τα αποτελέσματά τους.

Στην συνέχεια παρουσιάζεται το αποτέλεσμα εκτίμησης του μοντέλου LSTM, που εμφανίζει τα καλύτερα αποτελέσματα βραχυπρόθεσμης πρόβλεψης για την ανεμογεννήτρια 1, σε σχέση με τις πραγματικές τιμές ενεργειακής παραγωγής.



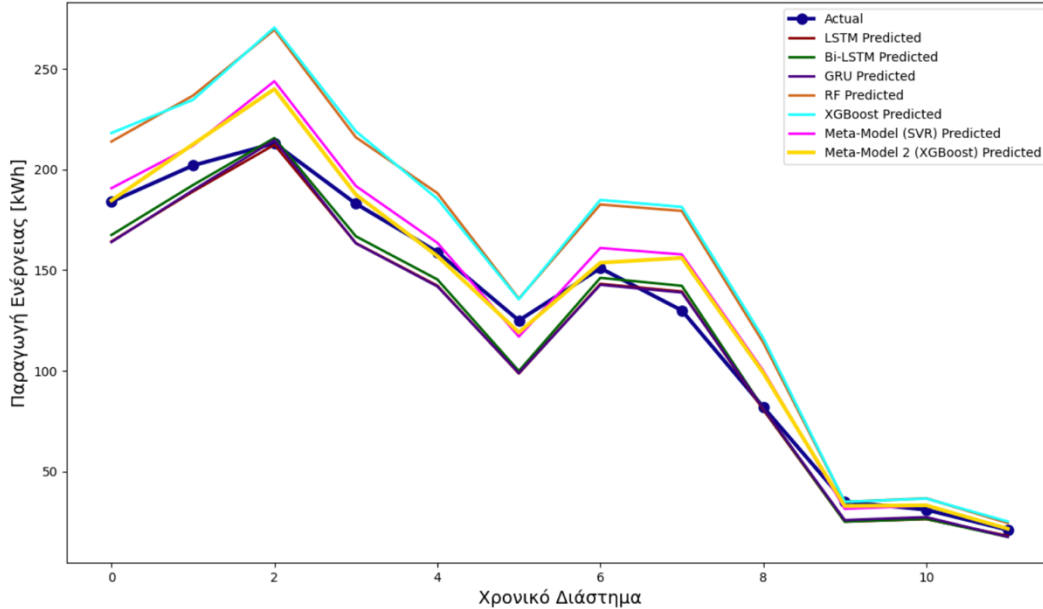
Εικόνα 48: Αποτέλεσμα πρόβλεψης βέλτιστου μοντέλου (Ανεμογεννήτρια 1)

## ΑΝΕΜΟΓΕΝΝΗΤΡΙΑ 2

PLANT 2	LSTM	Bi-LSTM	GRU	RF	XGBOOST	META - SVR	META - XGBOOST
MSE	182.689	146.397	181.335	987.312	1054.461	205.948	156.87
RMSE	13.516	12.099	13.466	31.421	32.472	14.350	12.524
MAE	11.005	9.954	10.908	26.35	27.211	10.902	8.430
R <sup>2</sup>	0.957	0.966	0.958	0.771	0.756	0.952	0.963

*Πίνακας 23: Αποτελέσματα βραχυπρόθεσμης πρόβλεψης Ανεμογεννήτριας 2*

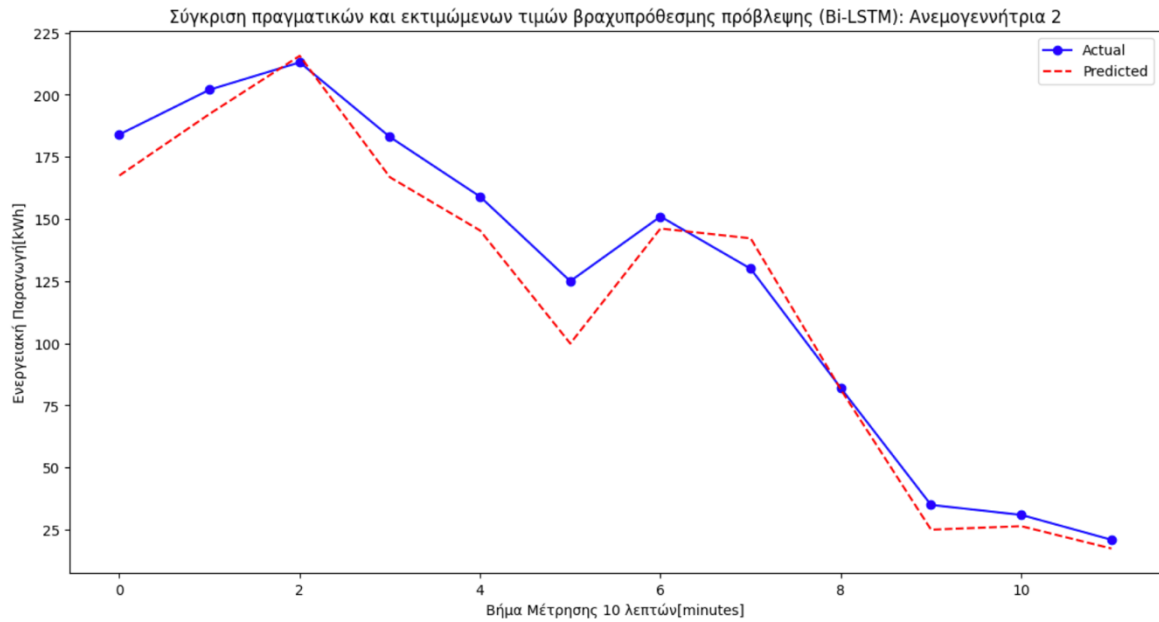
Σύγκριση μεταξύ Πραγματικών και Προβλεπόμενων Τιμών Ενεργειακής Παραγωγής σε 2 ώρες : Ανεμογεννήτρια 2



*Εικόνα 49: Σύγκριση συνολικών αποτελεσμάτων πρόβλεψης με πραγματικές τιμές*

Για την ανεμογεννήτρια 2 το Bi-LSTM εμφανίζει εξαιρετικά αποτελέσματα πρόβλεψης, ενώ και η απόδοση του μέτα-εκπαιδευτή-XGBoost καθώς και των νευρωνικών δικτύων είναι ικανοποιητική και φαίνεται να συγκλίνει στην πραγματική ενεργειακή παραγωγή. Αντίστοιχη συμπεριφορά φαίνεται να έχει και ο μέτα-εκπαιδευτής-SVR. Στην συγκεκριμένη περίπτωση τα μοντέλα Random Forest (RF) και XGBoost αποτελούν τις επιλογές με τις λιγότερο ικανοποιητικές αποδόσεις πρόβλεψης.

Επιπλέον, το αποτέλεσμα εκτίμησης του βέλτιστου μοντέλου, που εμφανίζει τα καλύτερα αποτελέσματα βραχυπρόθεσμης πρόβλεψης για την ανεμογεννήτρια 2, παρουσιάζεται σε αυτόνομο διάγραμμα σε σχέση με τις πραγματικές τιμές ενεργειακής παραγωγής.



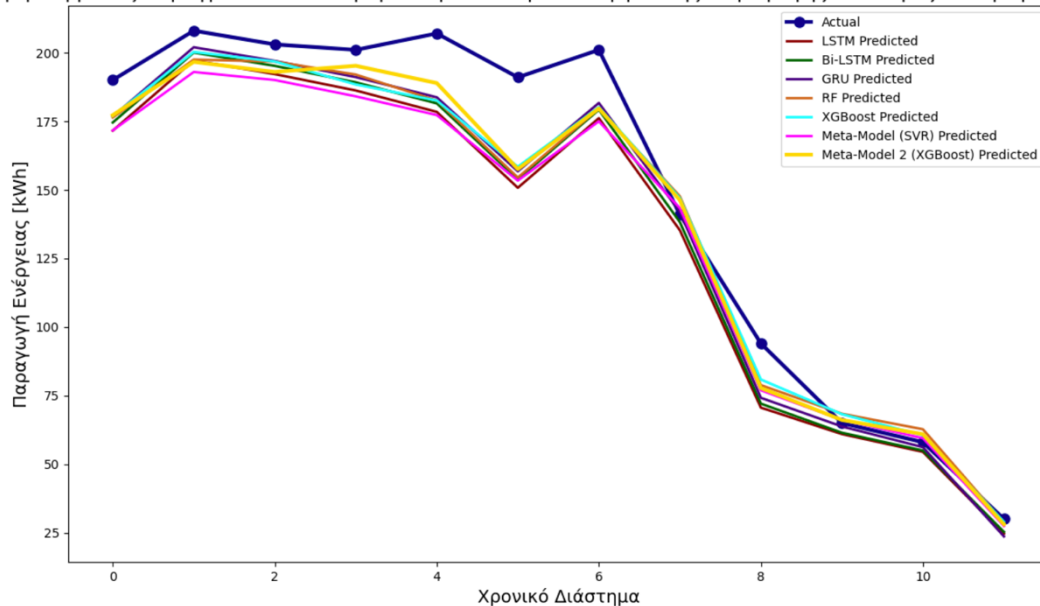
Εικόνα 50: Αποτέλεσμα πρόβλεψης βέλτιστου μοντέλου (Ανεμογεννήτρια 2)

### ΑΝΕΜΟΓΕΝΝΗΤΡΙΑ 3

PLANT 3	LSTM	Bi - LSTM	GRU	RF	XGBOOST	META - SVR	META - XGBOOST
MSE	374.232	295.929	239.719	260.059	229.429	359.164	219.564
RMSE	19.345	17.202	15.482	16.126	15.146	18.951	14.817
MAE	15.896	13.64	11.863	12.756	11.96	15.117	11.695
R <sup>2</sup>	0.912	0.93	0.943	0.939	0.94	0.915	0.948

Πίνακας 24: Αποτελέσματα βραχυπρόθεσμης πρόβλεψης Ανεμογεννήτριας 3

Σύγκριση μεταξύ Πραγματικών και Προβλεπόμενων Τιμών Ενεργειακής Παραγωγής σε 2 ώρες : Ανεμογεννήτρια 3

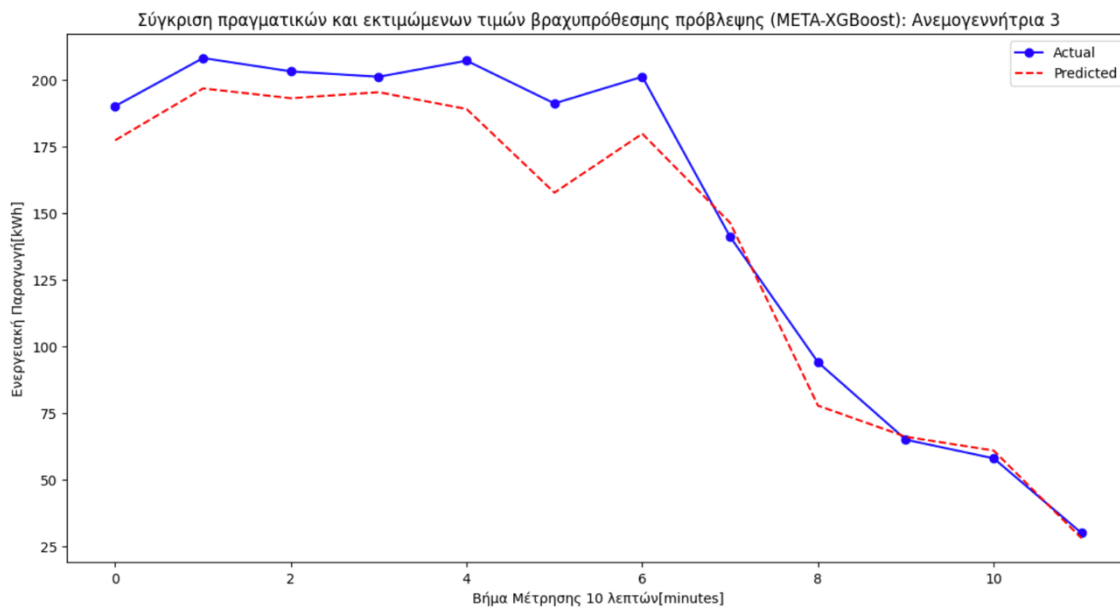


Εικόνα 51: Σύγκριση συνολικών αποτελεσμάτων πρόβλεψης με πραγματικές τιμές

Για την ανεμογεννήτρια 3 η επιλογή του συνδυαστικού μοντέλου meta-XGBoost φαίνεται να είναι η καλύτερη επιλογή για την προσέγγιση των πραγματικών τιμών ενεργειακής παραγωγής από τη συγκεκριμένη ανεμογεννήτρια. Αυτό έρχεται σε συμφωνία με το σετ δοκιμής του συνολικού dataset. Επιπλέον, ικανοποιητικά αποτελέσματα και σε αυτή την περίπτωση φαίνεται να εμφανίζει και ο εναλλακτικός μετά-εκπαιδευτής καθώς και το LSTM. Τα μοντέλα Random Forest (RF) και XGBoost εμφανίζουν σε αυτή την περίπτωση καλύτερη

απόδοση πρόβλεψης, αν και υποεκτιμούν την ενεργειακή παραγωγή. Το LSTM παρότι εμφανίζει υστέρηση απόδοσης, συγκριτικά με τις άλλες περιπτώσεις, φαίνεται να συνεισφέρει στη βελτίωση των αποτελεσμάτων μετά-εκπαίδευσης

Επιπλέον, το αποτέλεσμα βραχυπρόθεσμης εκτίμησης του συγκεκριμένου μετά-εκπαιδευτή σε σχέση με τις πραγματικές τιμές ενεργειακής παραγωγής για την ανεμογεννήτρια 3 παρουσιάζονται στο ακόλουθο διάγραμμα:



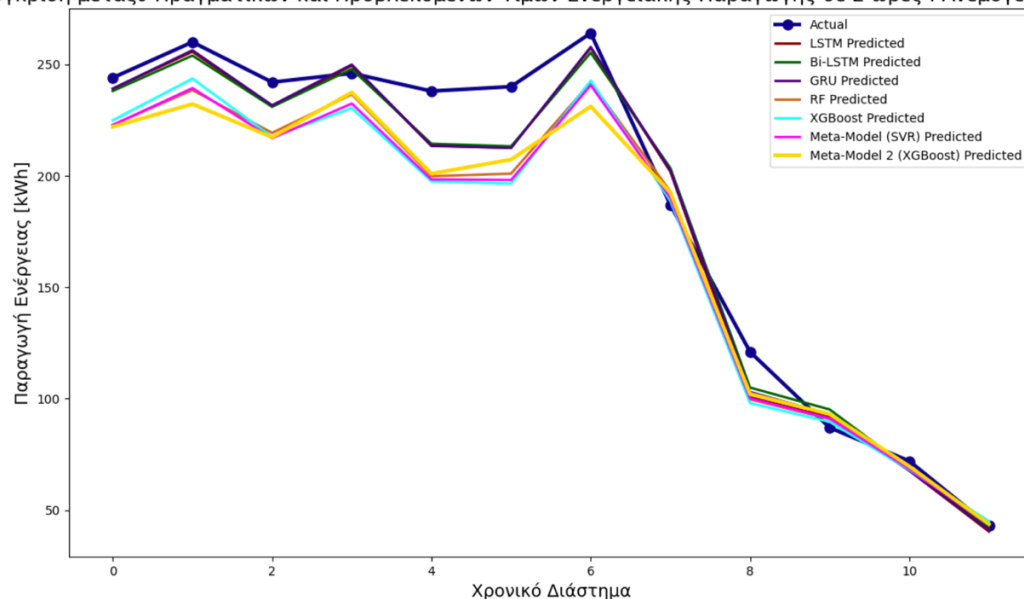
Εικόνα 52: Αποτέλεσμα πρόβλεψης βέλτιστου μοντέλου (Ανεμογεννήτρια 3)

#### ΑΝΕΜΟΓΕΝΝΗΤΡΙΑ 4

PLANT 4	LSTM	Bi-LSTM	GRU	RF	XGBOOST	META - SVR	META - XGBOOST
MSE	188.372	179.87	181.035	458.419	502.261	503.336	490.811
RMSE	13.724	13.411	13.454	21.41	22.411	22.435	22.154
MAE	10.765	10.795	10.539	17.436	17.851	18.128	18.257
R <sup>2</sup>	0.969	0.971	0.97	0.926	0.919	0.919	0.921

Πίνακας 25: Αποτελέσματα βραχυπρόθεσμης πρόβλεψης Ανεμογεννήτριας 4

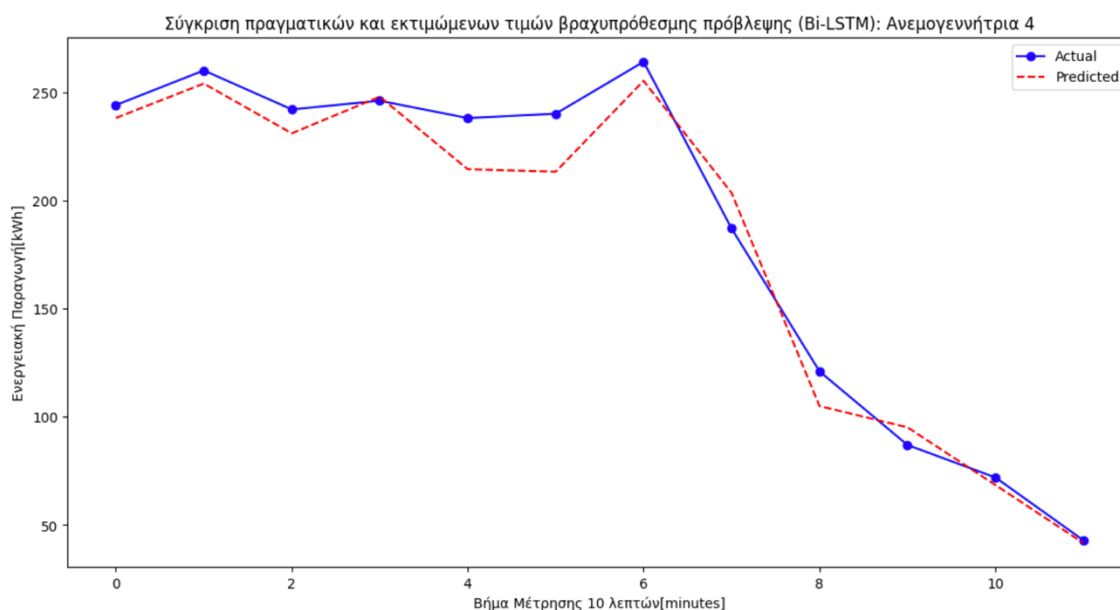
Σύγκριση μεταξύ Πραγματικών και Προβλεπόμενων Τιμών Ενεργειακής Παραγωγής σε 2 ώρες : Ανεμογεννήτρια 4



Εικόνα 53: Σύγκριση συνολικών αποτελεσμάτων πρόβλεψης με πραγματικές τιμές

Η ανεμογεννήτρια 4 φαίνεται να παρουσιάζει τη μεγαλύτερη δυσκολία στη συγκεκριμένη βραχυπρόθεσμη πρόβλεψη σε σχέση με τις υπόλοιπες ανεμογεννήτριες και η επιλογή του Bi-LSTM φαίνεται να εμφανίζει την καλύτερη απόδοση. Αξίζει να αναφερθεί ότι τα μοντέλα μετά-εκπαίδευσης εμφανίζουν ικανοποιητική σύγκλιση στις πραγματικές τιμές για τα πρώτα βήματα (80 λεπτά) αλλά δυσκολεύονται συνολικά να ερμηνεύσουν τα αποτελέσματα τους για τα επόμενα βήματα στη συγκεκριμένη πρόβλεψη. Επιπλέον, οι υπόλοιπες επιλογές νευρωνικών δικτύων φαίνεται να συγκλίνουν με την ίδια επιτυχία όπως και στις υπόλοιπες ανεμογεννήτριες και αποτελούν μια ικανοποιητική εναλλακτική επιλογή, ενώ τα μοντέλα Random Forest (RF) και XGBoost εμφανίζουν και σε αυτή την περίπτωση τις χειρότερες προβλέψεις.

Ακολουθεί το αποτέλεσμα πρόβλεψης του Bi-LSTM σε σχέση με τις πραγματικές τιμές ενεργειακής παραγωγής για την ανεμογεννήτρια 4.



Εικόνα 54: Αποτέλεσμα πρόβλεψης βέλτιστου μοντέλου (Ανεμογεννήτρια 4)

## 6.5.2 ΗΜΕΡΗΣΙΕΣ ΠΡΟΒΛΕΨΕΙΣ

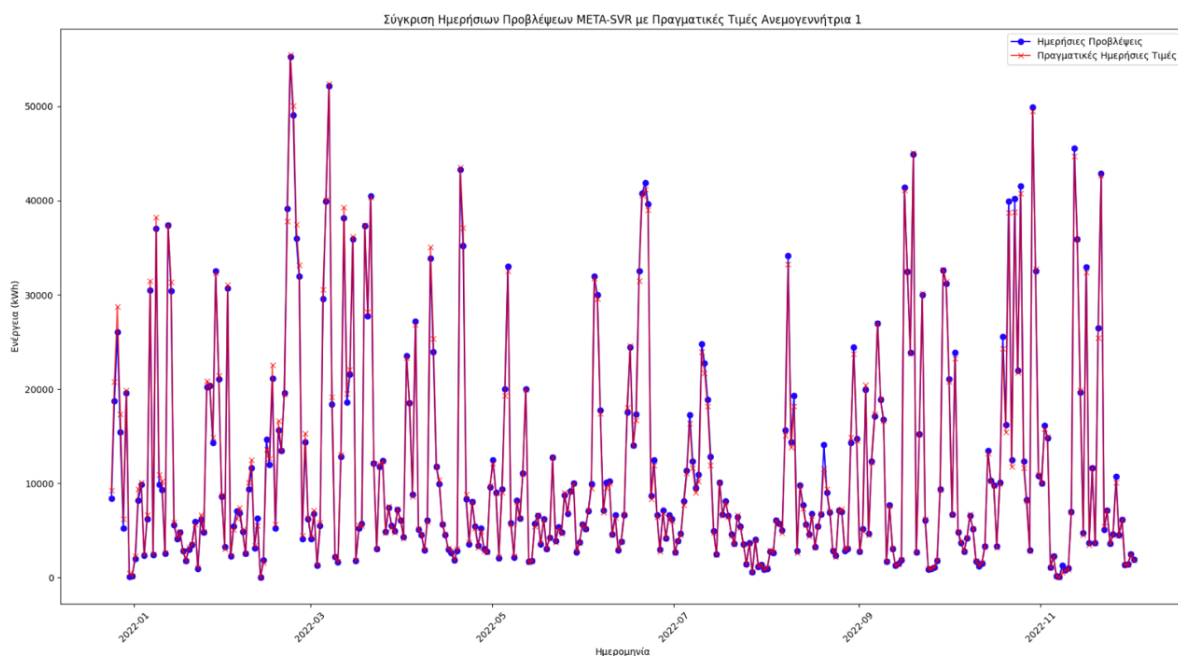
Για την περαιτέρω ανάλυση των αποτελεσμάτων πρόβλεψης, ώστε να επιτευχθεί μια ολιστική και καλύτερη κατανόηση των αποτελεσμάτων πραγματοποιήθηκε και η ημερήσια πρόβλεψη στο σετ δοκιμής για κάθε μια από τις ανεμογεννήτριες. Τα ημερήσια αποτελέσματα προβλέψεων μπορούν να προσφέρουν αξιολόγηση της απόδοσης των μοντέλων σε διαφορετικές περιόδους του χρόνου. Για την αξιολόγηση των ημερήσιων αποτελεσμάτων χρησιμοποιήθηκαν το ημερήσιο Root Mean Squared Error (RMSE), το οποίο υποδεικνύει το μέσο όρο της τυπικής απόκλισης πραγματικής και προβλεπόμενης τιμής, το ημερήσιο Mean Absolute Error (MAE), το οποίο υποδεικνύει μέσο απόλυτο σφάλμα μεταξύ πραγματικής και προβλεπόμενης τιμής και το Ημερήσιο  $R^2$  δηλαδή το ποσοστό προσαρμογής στην διακύμανση. Στην περίπτωση των ημερήσιων διακυμάνσεων στην πρόβλεψη παραγωγής οι μέθοδοι μετά-εκπαίδευσης παρουσιάζουν την καλύτερη απόδοση με τα μικρότερα σφάλματα.

Για την καλύτερη ανάλυση των αποτελεσμάτων πρόβλεψης αναφέρεται ότι η συνολική μέση τιμή της ενέργειας παραγωγής με βήμα μέτρησης 10 λεπτών στο συνολικό σετ δεδομένων είναι 86.213 kWh. Στην περίπτωση των ημερήσιων διακυμάνσεων στην πρόβλεψη παραγωγής οι μέθοδοι μετά-εκπαίδευσης παρουσιάζουν την καλύτερη απόδοση με τα μικρότερα δυνατά σφάλματα.



<b>META-SVR</b>	<b>ANEMOΓΕΝΝΗΤΡΙΑ 1</b>
Ημερήσιο RMSE	506.102
Ημερήσιο MAE	309.367
Ημερήσιο R <sup>2</sup>	0.998

Πίνακας 26: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-SVR - Ανεμογεννήτρια 1

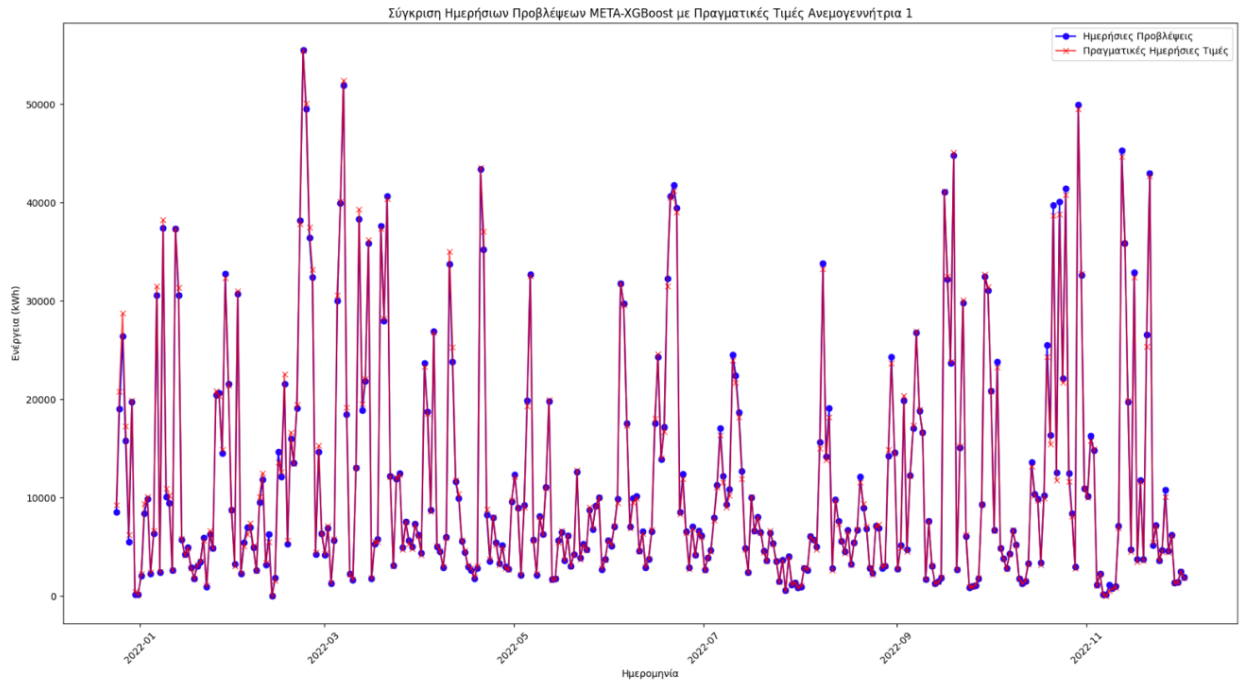


Εικόνα 55: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-SVR με τις πραγματικές τιμές - Ανεμογεννήτρια 1

<b>META-XGBoost</b>	<b>ANEMOΓΕΝΝΗΤΡΙΑ 1</b>
Ημερήσιο RMSE	417.530
Ημερήσιο MAE	265.467
Ημερήσιο R <sup>2</sup>	0.998

Πίνακας 27: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-XGBoost - Ανεμογεννήτρια 1

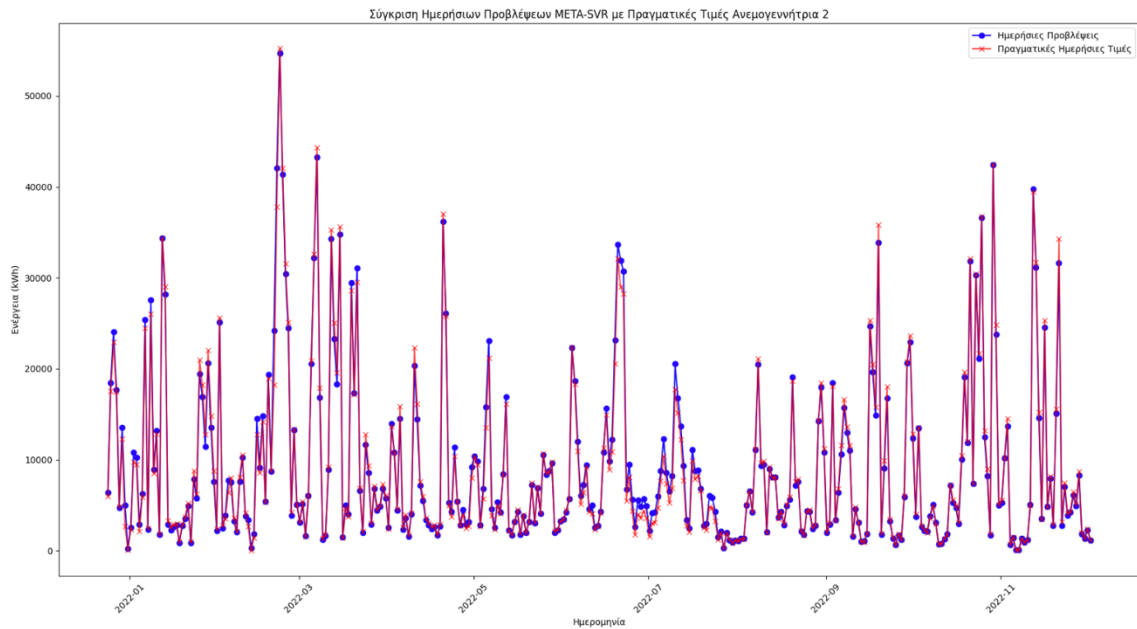




Εικόνα 56: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-XGBoost με τις πραγματικές τιμές - Ανεμογεννήτρια 1

META-SVR	ANEMOΓΕΝΝΗΤΡΙΑ 2
Ημερήσιο RMSE	838.470
Ημερήσιο MAE	519.798
Ημερήσιο R <sup>2</sup>	0.992

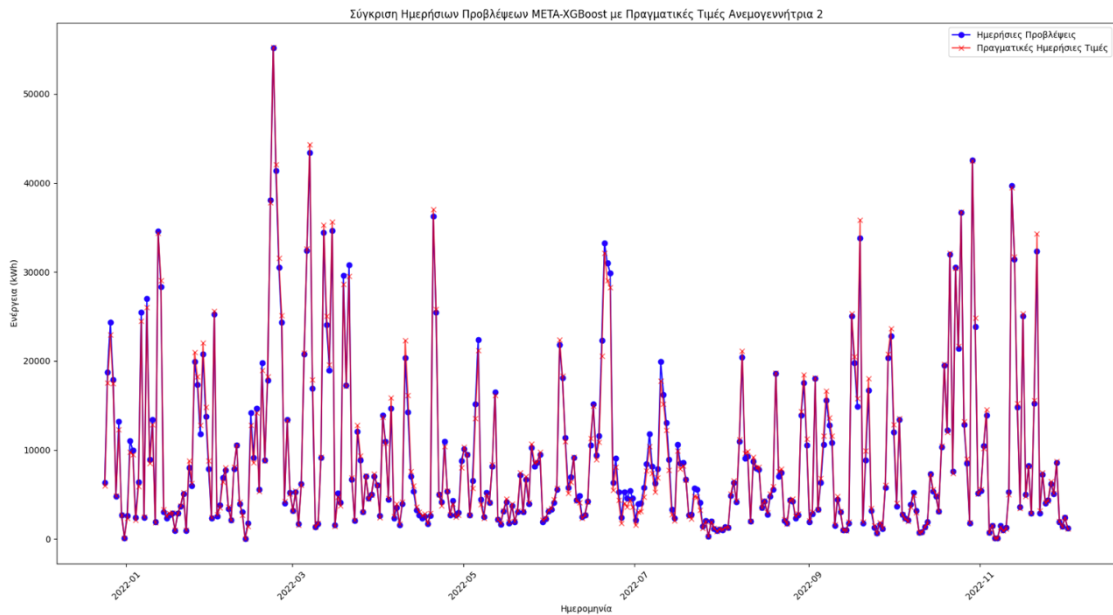
Πίνακας 28: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-SVR - Ανεμογεννήτρια 2



Εικόνα 57: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-SVR με τις πραγματικές τιμές - Ανεμογεννήτρια 2

META-XGBoost	ANEMOΓΕΝΝΗΤΡΙΑ 2
Ημερήσιο RMSE	587.932
Ημερήσιο MAE	404.975
Ημερήσιο R <sup>2</sup>	0.996

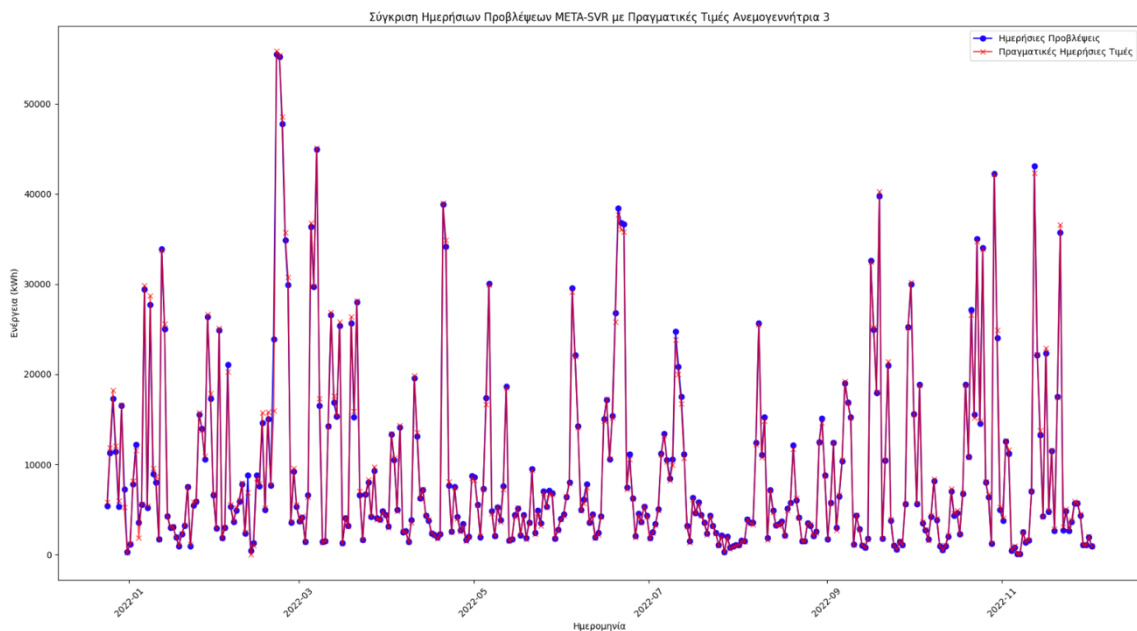
Πίνακας 29: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-XGBoost - Ανεμογεννήτρια 2



Εικόνα 58: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-XGBoost με τις πραγματικές τιμές - Ανεμογεννήτρια 2

META-SVR	ANEMOΓΕΝΝΗΤΡΙΑ 3
Ημερήσιο RMSE	548.220
Ημερήσιο MAE	239.270
Ημερήσιο R <sup>2</sup>	0.997

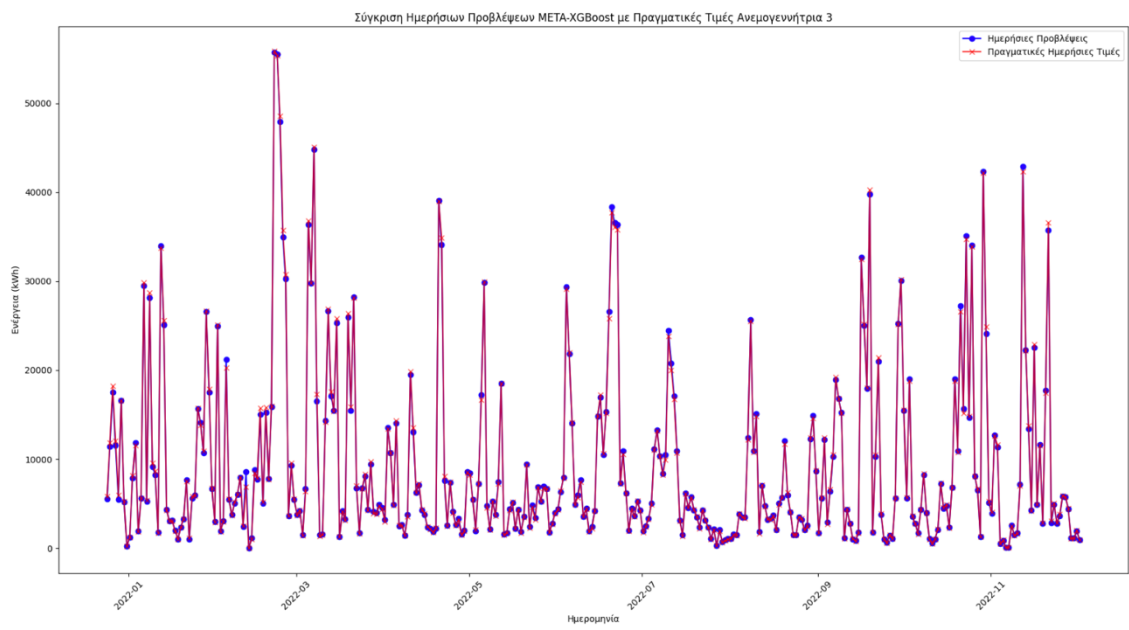
Πίνακας 30: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-SVR - Ανεμογεννήτρια 3



Εικόνα 59: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-SVR με τις πραγματικές τιμές - Ανεμογεννήτρια 3

META-XGBoost	ANEMOΓΕΝΝΗΤΡΙΑ 3
Ημερήσιο RMSE	251.871
Ημερήσιο MAE	160.452
Ημερήσιο R <sup>2</sup>	0.999

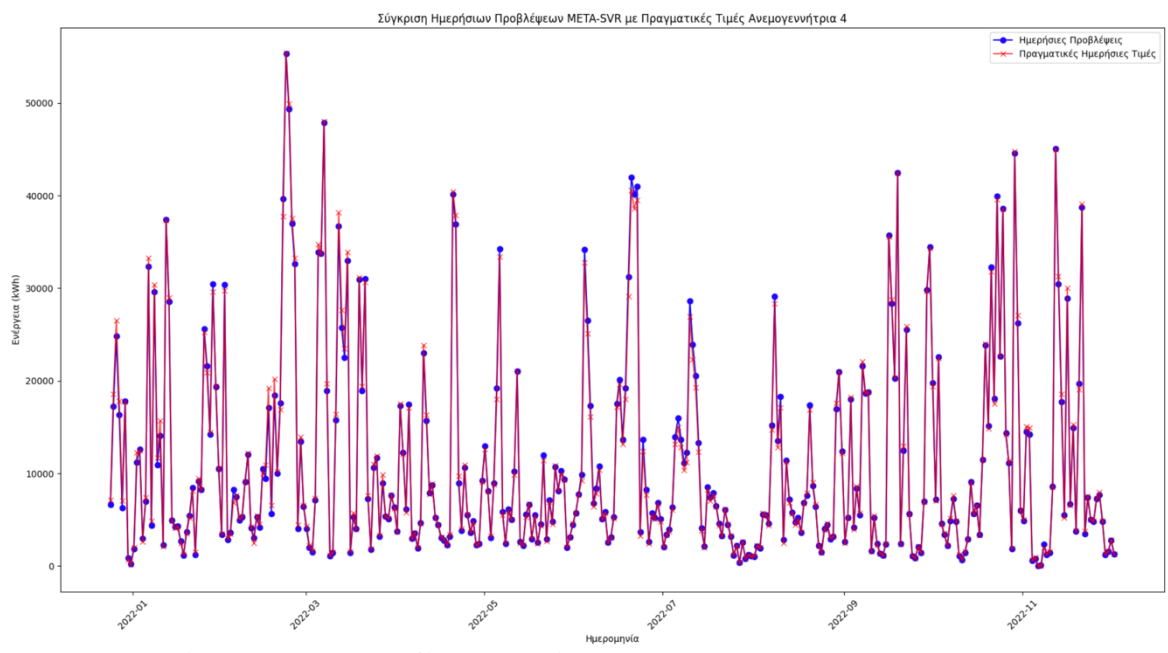
Πίνακας 31: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-XGBoost - Ανεμογεννήτρια 3



Εικόνα 60: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-XGBoost με τις πραγματικές τιμές - Ανεμογεννήτρια 3

META-SVR	ΑΝΕΜΟΓΕΝΝΗΤΡΙΑ 4
Ημερήσιο RMSE	537.107
Ημερήσιο MAE	340.870
Ημερήσιο R <sup>2</sup>	0.997

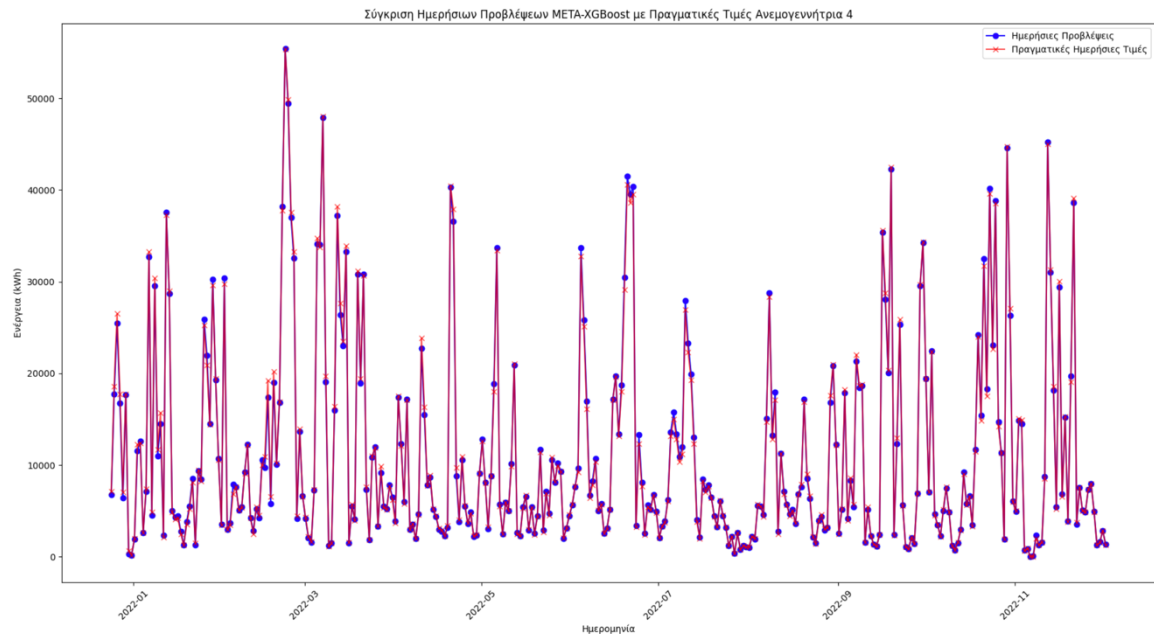
Πίνακας 32: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-SVR - Ανεμογεννήτρια 4



Εικόνα 61: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-SVR με τις πραγματικές τιμές - Ανεμογεννήτρια 4

META-XGBoost	ΑΝΕΜΟΓΕΝΝΗΤΡΙΑ 4
Ημερήσιο RMSE	402.312
Ημερήσιο MAE	265.630
Ημερήσιο R <sup>2</sup>	0.998

Πίνακας 33: Αποτελέσματα απόδοσης ημερήσιων προβλέψεων μοντέλου META-XGBoost - Ανεμογεννήτρια 4



Εικόνα 62: Αποτελέσματα ημερήσιων προβλέψεων μοντέλου META-XGBoost με τις πραγματικές τιμές - Ανεμογεννήτρια 4

Για κάθε μια από τις ανεμογεννήτριες οι μέθοδοι μέτα-εκπαίδευσης που αποτελούν την βέλτιστη επιλογή έχουν χαμηλούς δείκτες τόσο στον ημερήσιο μέσο όρο τυπικής απόκλισης όσο και στο ημερήσιο μέσο απόλυτο σφάλμα των πραγματικών τιμών από τις προβλεπόμενες. Ειδικότερα, αν αναλογιστούμε ότι η μέση τιμή ενεργειακής παραγωγής ανά 10 λεπτά συνεχόμενης λειτουργίας είναι 86.213 kWh, τα αποτελέσματα των σφαλμάτων βρίσκονται πολύ κοντά σε διορθώσεις λειτουργίας υποεκτίμησης ή υπερεκτίμησης λιγότερο από 1 ώρα λειτουργίας. Τέλος, μια τόσο υψηλή τιμή του  $R^2$  υποδεικνύει ότι το μοντέλο είναι πολύ καλό στην εξήγηση της διακύμανσης των πραγματικών τιμών, προσεγγίζοντας την τέλεια προσαρμογή του μοντέλου στα δεδομένα.

# ΚΕΦΑΛΑΙΟ 7: ΕΠΙΛΟΓΟΣ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ

## 7.1 ΓΕΝΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ

Η διαρκής αύξηση των ΑΠΕ και συγκεκριμένα της Αιολικής Ενέργειας δημιουργεί νέες ανάγκες στους επενδυτές. Μια πολύ σημαντική ανάγκη είναι η πρόβλεψη της παραγωγής μέσω της χρήσης μετεωρολογικών δεδομένων και της αξιοποίησης των δυνατοτήτων της τεχνητής νοημοσύνης. Η εμφάνιση πολλών εταιρειών τα τελευταία χρόνια που προσφέρουν τέτοιου είδους υπηρεσίες, μέσω της ανάπτυξης διαφόρων μεθόδων, ώστε να προσφέρονται βέλτιστες προβλέψεις παραγωγής, αναδεικνύει την σημαντικότητα του κλάδου. Η διαθεσιμότητα αξιόπιστων ιστορικών δεδομένων, η άρτια επεξεργασία τους και η επιλογή του κατάλληλου μοντέλου, αποτελούν κρίσιμους παράγοντες για την αντιμετώπιση του συγκεκριμένου προβλήματος.

Η συγκεκριμένη εργασία παρουσιάζει τις δυνατότητες πρόβλεψης ενεργειακής παραγωγής με τη χρήση βασικών και συνδυαστικών μοντέλων μηχανικής μάθησης. Οι προβλέψεις των βασικών μοντέλων αποτελούν τη βάση για τα μέτα-μαθησιακά μοντέλα, τα οποία προσπαθούν να βελτιστοποιήσουν τα αποτελέσματα πρόβλεψης και να συνδυάσουν τα ισχυρά σημεία των εκάστοτε μοντέλων δημιουργώντας πιο αποδοτικές προβλέψεις και μεγαλύτερη ασφάλεια.

Παρατηρούμε από τα παραπάνω αποτελέσματα της εκπαίδευσης των δεδομένων, αλλά και από τα αποτελέσματα της βραχυπρόθεσμης καθώς και ημερήσιας πρόβλεψης ενεργειακής παραγωγής, ότι η χρήση των μεθόδων μέτα-εκπαίδευσης κατά κύριο λόγο εμφανίζουν βελτιωμένα αποτελέσματα γενίκευσης και απόδοσης. Τα αποτελέσματα των νευρωνικών δικτύων, ιδίως για την βραχυπρόθεσμη πρόβλεψη της ενεργειακής παραγωγής, εμφανίζουν ικανοποιητικά αποτελέσματα. Επιπλέον, τα βασικά μοντέλα, Random Forest και XGBoost σε ορισμένες περιπτώσεις εμφανίζουν αποτελέσματα ανταγωνιστικά με τις μεθόδους μέτα-εκπαίδευσης, ενώ σε άλλες περιπτώσεις εμφανίζουν μεγάλες αποκλίσεις φανερώνοντας την έλλειψη σταθερότητας στην απόδοσή τους ανάλογα με τα δεδομένα. Τα προτεινόμενα μοντέλα αναμειγνύουν τις βασικές προβλέψεις εκμεταλλευόμενα τα πλεονεκτήματα του εκάστοτε βασικού μοντέλου, βελτιώνοντας τη συνολική ακρίβεια και γενίκευση των λύσεων στις περισσότερες περιπτώσεις.

Η εκπαίδευση των συγκεκριμένων μοντέλων που πραγματοποιήθηκε σε μια μεγάλη χρονική περίοδο με βάση κυρίως τα αριθμητικά δεδομένα καιρού, εμφανίζουν πολύ καλή σύγκλιση στις πραγματικές τιμές πρόβλεψης τόσο σε βήμα 10 λεπτών όσο και σε ημερήσιο βήμα. Το πλεονέκτημα της συγκεκριμένης πρόβλεψης, όπως διαφαίνεται και από τη συγκεκριμένη μελέτη, είναι η ευελιξία στην απόδοση σε διάφορους χρονικούς ορίζοντες. Η αξιοποίηση της προτεινόμενης μεθόδου για εξαγωγή προβλέψεων σε διαφορετικούς χρονικούς ορίζοντες, με βάση την εκμάθηση από τα πρότυπα που ορίζονται από τα βασικά μοντέλα και κατ' επέκταση από τους μέτα-εκπαιδευτές, βελτιώνει συνολικά την απόδοση. Αξίζει να τονιστεί ότι ακόμα και τα βασικά μοντέλα που εμφανίζουν τις χειρότερες αποδόσεις φαίνεται μέσα από τις δοκιμές να ενισχύουν τα αποτελέσματα μέτα-εκπαίδευσης.

Μέσα από την κατάλληλη επιλογή παραμέτρων για το εκάστοτε μοντέλο, με βάση τη βιβλιογραφία και τη διαδικασία δοκιμής - σφάλματος (trial and error), πραγματοποιήθηκε η βελτιστοποίηση των αποτελεσμάτων εκπαίδευσης και ο περιορισμός του κινδύνου μεγάλων σφαλμάτων πρόβλεψης. Η εισροή δεδομένων, που βασίζονται σε αριθμητικά δεδομένα καιρού στα συγκεκριμένα μοντέλα πρόβλεψης, μπορούν να αποτελέσουν τη βάση για

εξαγωγή προβλέψεων με μικρό σφάλμα για διαφορετικούς χρονικούς ορίζοντες, αφού τα αποτελέσματα ελέγχου για τα μοντέλα μετά-εκπαίδευσης εμφανίζουν εξαιρετική απόδοση και ελαχιστοποιούν τα σφάλματα. Ειδικότερα για την μετρική MAE που προσφέρει εκτίμηση για το μέσο σφάλμα μεταξύ προβλεπόμενης και πραγματικής τιμής, φαίνεται ότι οι μέθοδοι μετά-εκπαίδευσης βελτιώνουν τα αποτελέσματα από την απόδοση του καλύτερου βασικού μοντέλου έως και 23.66% για την ανεμογεννήτρια 1, έως και 39% για την ανεμογεννήτρια 2, έως και 20% για την ανεμογεννήτρια 3 και έως και 17% για την ανεμογεννήτρια 4. Συγκεντρωτικά, η μέση βελτίωση στο μέσο σφάλμα για το σύνολο του αιολικού πάρκου είναι 25%, μέσω της χρήσης μετά-εκπαίδευσης.

Σε επίπεδο αγοράς τα αποτελέσματα βελτιωμένης πρόβλεψης καθίστανται σημαντικά για την εκτίμηση της συνεισφοράς της αιολικής παραγωγής στη συνολικά παραγόμενη ενέργεια και τον προσδιορισμό των ημερήσιων και ωριαίων τιμών. Η βελτιστοποίηση της ενεργειακής διαχείρισης, ο καθορισμός των τιμών εκκαθάρισης ενεργειακών και λειτουργικών αποθεμάτων είναι άμεσα συνδεδεμένος με το συγκεκριμένο κλάδο.

## 7.2 ΠΕΡΙΟΡΙΣΜΟΙ

Ως απώτερος σκοπός της συγκεκριμένης εργασίας είναι η δημιουργία εργαλείου εξαγωγής προβλέψεων παραγωγής αιολικής ενέργειας που θα προσφέρει χρονική ευελιξία, ασφάλεια και επιδόσεις πρόβλεψης με μικρές σφαλματικές διακυμάνσεις. Παρόλο που η συγκεκριμένη υλοποίηση εμφανίζει ικανοποιητικά αποτελέσματα, αποδεκτά σφάλματα και οι τιμές εκτιμώμενης με την πραγματική ενεργειακή παραγωγή είναι στα επιθυμητά όρια, παρατηρούνται ορισμένοι περιορισμοί και δυσκολίες στη συνολική τους εφαρμογή.

- Οι προβλέψεις αριθμητικών δεδομένων καιρού είναι ύψιστης σημασίας για την απόδοση των μοντέλων για το συγκεκριμένο αιολικό πάρκο και τις μελλοντικές προβλέψεις. Τα αποτελέσματα των μοντέλων βασίζονται σε μεγάλο βαθμό στην αξιοπιστία των τιμών πρόβλεψης αριθμητικών δεδομένων καιρού, τα οποία αποτελούν και δεδομένα εισόδου σε όλα τα βασικά μοντέλα πρόβλεψης. Το γεγονός της υψηλής ευαισθησίας των προβλέψεων ως προς συγκεκριμένες αριθμητικές τιμές προκύπτει λόγω της υψηλής συσχέτισης των δεδομένων εισόδου με τη στήλη στόχου. Θα πρέπει να υπάρξει προσεκτική επιλογή των νέων δεδομένων εισόδου για το συγκεκριμένο αιολικό πάρκο, ώστε να διασφαλιστεί η αξιοπιστία των προβλέψεων στα ίδια επίπεδα με εκείνα των παρουσιαζόμενων αποτελεσμάτων.
- Με τις συγκεκριμένες μεθόδους σε περίπτωση ανάγκης ανανέωσης της εκπαίδευσης, είτε στην περίπτωση αξιοποίησης των συγκεκριμένων μεθόδων με δεδομένα αντίστοιχης φύσεως για ανεμογεννήτριες που έχουν άλλα χαρακτηριστικά επίδοσης και βρίσκονται σε διαφορετικές τοποθεσίες, με άλλες καιρικές συνθήκες, θα πρέπει να αναλογιστούμε και τον παράγοντα του χρόνου εκπαίδευσης των μεθόδων. Τονίζεται ότι η αξιοποίηση των μεθόδων μετά-εκπαίδευσης προϋποθέτει την εκπαίδευση όλων των βασικών μοντέλων. Η ορθή διαχείριση του συγκεκριμένου περιορισμού έχει ιδιαίτερη σημασία για την ασφαλή χρονική διασφάλιση του ορίζοντα εκτίμησης. Ειδικότερα, με όγκο δεδομένων 2 ετών, με βήμα μέτρησης 10 λεπτών, με τα προαναφερθέντα χαρακτηριστικά εισόδου και με τη χρήση συμβατικών πόρων, παρατηρήθηκαν οι παρακάτω χρόνοι εκπαίδευσης σε δευτερόλεπτα για κάθε μια από τις μεθόδους με συνολικό χρόνο 124 λεπτά :



Μοντέλο	Ανεμογεννήτρια 1	Ανεμογεννήτρια 2	Ανεμογεννήτρια 3	Ανεμογεννήτρια 4
LSTM	454	639	447	368
Bi-LSTM	648	955	651	573
GRU	333	442	328	252
RF	15	33	11	14
XGBoost	2	7	9	4
Meta-SVR	323	274	270	384
Meta-XGBoost	1	2	2	3
Συνολικά	1776	2352	1718	1598

Πίνακας 34: Χρόνοι εκπαίδευσης μοντέλων για κάθε Ανεμογεννήτρια

### 7.3 ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ

Μερικές από τις βασικές μελλοντικές προσεγγίσεις θα ήταν η δοκιμή των παραπάνω μοντέλων σε σύνολα δεδομένων διαφορετικών ανεμογεννητριών, προκειμένου να αποσαφηνιστεί αν έχουν την ίδια αποδοτικότητα. Μπορούν να γίνουν περισσότερες εκτεταμένες δοκιμές, με τη χρήση μικρότερου μεγέθους δεδομένων εισόδου, ώστε να δοκιμαστεί περαιτέρω η αξιοπιστία των μοντέλων όταν υπάρχουν ακόμα λιγότερες πληροφορίες.

Επιπλέον, παρότι πραγματοποιήθηκε εκτενής επεξεργασία για την εφαρμογή των καλύτερων παραμέτρων μέσα από αρκετές εκπαιδεύσεις δεδομένων, η εφαρμογή προηγμένων τεχνικών για την εξαγωγή νέων δεδομένων και την ανακάλυψη κρυμμένων προτύπων, αποτελεί μελλοντική προσέγγιση που μπορεί να βελτιώσει περαιτέρω την απόδοση. Συνεπώς, μπορούν να γίνουν δοκιμές στα συνδυαστικά μοντέλα με άλλες μεταβλητές εισόδου.

Με βάση το υπάρχον σύνολο δεδομένων των ανεμογεννητριών, μπορεί να πραγματοποιηθεί και στατιστική ανάλυση χρονοσειρών για την εξαγωγή προβλέψεων ενεργειακής παραγωγής μη βασιζόμενη σε αριθμητικά δεδομένα καιρού. Η απαιτούμενη στατιστική ανάλυση μπορεί να χρησιμοποιηθεί για να εισάγει περαιτέρω στατιστικά στοιχεία ως νέα δεδομένα εισόδου στα μοντέλα και να βελτιώσει περαιτέρω τα αποτελέσματα.

Βασιζόμενοι στις συγκεκριμένες υλοποιήσεις και σε πιο προηγμένες εφαρμογές μπορεί να επιτευχθεί η δημιουργία υπολογιστικού εργαλείου εξαγωγής προβλέψεων παραγωγής αιολικής ενέργειας με ενσωματωμένο έλεγχο επανεκπαίδευσης, βασιζόμενο στις μετρήσεις του αιολικού πάρκου. Με αυτόν τον τρόπο θα παρέχεται χρονική ευελιξία, μεγαλύτερη ασφάλεια και ενίσχυση της απόδοσης των επιδόσεων πρόβλεψης με μικρές σφαλματικές διακυμάνσεις, λαμβάνοντας υπόψιν τις πιθανές αλλαγές στις συνθήκες και τον τρόπο λειτουργίας των ανεμογεννητριών.

Εν κατακλείδι, η παρούσα διπλωματική παρέχει τις γνώσεις που χρειάζεται να έχει κάποιος που θέλει να ασχοληθεί με την εξαγωγή ενεργειακών προβλέψεων βασιζόμενη σε ιστορικά δεδομένα, αριθμητικά δεδομένα καιρού και τεχνικές μηχανικής μάθησης, ώστε να αναπτυχθούν μοντέλα πρόβλεψης αιολικής ενεργειακής παραγωγής με μικρό ποσοστό σφάλματος. Η απόδοση αποτελεσμάτων σε ήδη υπάρχουσες εφαρμογές για τις απαιτήσεις της ενεργειακής αγοράς σε επίπεδα διαχείρισης, συντήρησης και χρηματιστηρίου ενέργειας αποτελεί τον πιο αποδοτικό τρόπο αξιολόγησης των προτεινόμενων μεθόδων. Η δοκιμή των μοντέλων με τη χρήση διαφορετικών NWP σε περιβάλλον κανονικής λειτουργίας αποτελεί μελλοντική προσέγγιση για την δοκιμή της υπολογιστικής ταχύτητας και απόδοσης του εκάστοτε μοντέλου.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] O. DEMİRTAS, “Evaluating the Best Renewable Energy Technology For Sustainable Energy Planning,” *International Journal of Energy Economics and Policy*, vol. 3, no. 4, pp. 23–33, 2013.
- [2] World Wind Energy Association, “World Market for Wind Power Saw Another Record Year in 2021: 97,3 Gigawatt of New Capacity Added. .”
- [3] “Wind energy in Europe : 2022 Statistics and the outlook for 2023-2027.” Accessed: Feb. 08, 2024. [Online]. Available: <https://windeurope.org/intelligence-platform/product/wind-energy-in-europe-2022-statistics-and-the-outlook-for-2023-2027/>
- [4] GLOBAL WIND ENERGY COUNCIL, “Global Wind Report 2023.” Accessed: Feb. 08, 2024. [Online]. Available: <https://gwec.net/globalwindreport2023/>
- [5] Hellenic Wind Energy Association, “Wind Energy Statistics – 2023.” Accessed: Feb. 08, 2024. [Online]. Available: [https://eletaen.gr/wp-content/uploads/2024/01/2024-01-18-2023-HWEA\\_Statistics-Greece.pdf](https://eletaen.gr/wp-content/uploads/2024/01/2024-01-18-2023-HWEA_Statistics-Greece.pdf)
- [6] A. Malakhov, F. Goncharov, and E. Gryazina, “Testing machine learning approaches for wind plants power output,” in *2019 International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)*, 2019, pp. 1–6. doi: 10.1109/REEPE.2019.8708815.
- [7] Y.-K. Wu and J.-S. Hong, “A literature review of wind forecasting technology in the world,” in *2007 IEEE Lausanne Power Tech*, 2007, pp. 504–509. doi: 10.1109/PCT.2007.4538368.
- [8] IRENA, “IRENA (2023), Renewable power generation costs in 2022, International Renewable Energy Agency, Abu Dhabi.” [Online]. Available: <https://www.irena.org/Publications/2023/Aug/Renewable-Power-Generation-Costs-in-2022>
- [9] X. Wang, P. Guo, and X. Huang, “A Review of Wind Power Forecasting Models,” *Energy Procedia*, vol. 12, pp. 770–778, 2011, doi: <https://doi.org/10.1016/j.egypro.2011.10.103>.
- [10] Y. Zhang, Y. Li, and G. Zhang, “Short-term wind power forecasting approach based on Seq2Seq model using NWP data,” *Energy*, vol. 213, p. 118371, 2020, doi: <https://doi.org/10.1016/j.energy.2020.118371>.
- [11] Q. Xu *et al.*, “A Short-Term Wind Power Forecasting Approach With Adjustment of Numerical Weather Prediction Input by Data Mining,” *IEEE Trans Sustain Energy*, vol. 6, no. 4, pp. 1283–1291, 2015, doi: 10.1109/TSTE.2015.2429586.
- [12] L. Landberg and S. J. Watson, “Short-term prediction of local wind conditions,” *Boundary Layer Meteorol*, vol. 70, no. 1, pp. 171–195, 1994, doi: 10.1007/BF00712528.
- [13] M. I. I. M. T. S. N. G. G. J. T. I. W. J. U. R. B. G. K. U. F. I. S. N. H. A. M. P. P. F. G. N. Kariniotakis, “Advanced Short-term Forecasting of Wind Generation - Anemos.,” 2006. Accessed: Feb. 11, 2024. [Online]. Available: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=f2dfdc5af6a87ff40d09e7db15d3f1e90a5324c9>
- [14] D. Carvalho, A. Rocha, M. Gómez-Gesteira, and C. Silva Santos, “WRF wind simulation and wind energy production estimates forced by different reanalyses: Comparison with observed data for Portugal,” *Appl Energy*, vol. 117, pp. 116–126, 2014, doi: <https://doi.org/10.1016/j.apenergy.2013.12.001>.
- [15] S. Aslam, H. Herodotou, S. M. Mohsin, N. Javaid, N. Ashraf, and S. Aslam, “A survey on deep learning methods for power load and renewable energy forecasting in



- smart microgrids,” *Renewable and Sustainable Energy Reviews*, vol. 144, p. 110992, 2021, doi: <https://doi.org/10.1016/j.rser.2021.110992>.
- [16] E. Sarmas, E. Spiliotis, E. Stamatopoulos, V. Marinakis, and H. Doukas, “Short-term photovoltaic power forecasting using meta-learning and numerical weather prediction independent Long Short-Term Memory models,” *Renew Energy*, vol. 216, p. 118997, 2023, doi: <https://doi.org/10.1016/j.renene.2023.118997>.
- [17] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, “A review of wind power and wind speed forecasting methods with different time horizons,” in *North American Power Symposium 2010*, 2010, pp. 1–8. doi: 10.1109/NAPS.2010.5619586.
- [18] S. Hanifi, X. Liu, Z. Lin, and S. Lotfian, “A Critical Review of Wind Power Forecasting Methods—Past, Present and Future,” *Energies (Basel)*, vol. 13, no. 15, 2020, doi: 10.3390/en13153764.
- [19] C. Monteiro *et al.*, “Wind power forecasting : state-of-the-art 2009.” United States, 2009. doi: 10.2172/968212.
- [20] R. Ahmed, V. Sreeram, Y. Mishra, and M. D. Arif, “A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization,” *Renewable and Sustainable Energy Reviews*, vol. 124, p. 109792, 2020, doi: <https://doi.org/10.1016/j.rser.2020.109792>.
- [21] J. Rounkvist and P. Enevoldsen, “Time-scale classification in wind forecasting: A review of the state-of-the-art,” *J Forecast*, vol. 39, Feb. 2020, doi: 10.1002/for.2657.
- [22] T. S. Nielsen, A. Joensen, H. Madsen, L. Landberg, and G. Giebel, “A new reference for wind power forecasting,” *Wind Energy*, vol. 1, no. 1, pp. 29–34, 1998, doi: [https://doi.org/10.1002/\(SICI\)1099-1824\(199809\)1:1<29::AID-WE10>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1099-1824(199809)1:1<29::AID-WE10>3.0.CO;2-B).
- [23] D. P. LALAS, “Wind Energy Estimation and Siting in Complex Terrain,” *International Journal of Solar Energy*, vol. 3, no. 2, pp. 43–71, 1985, doi: 10.1080/01425918508914383.
- [24] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, “Current methods and advances in forecasting of wind power generation,” *Renew Energy*, vol. 37, no. 1, pp. 1–8, 2012, doi: <https://doi.org/10.1016/j.renene.2011.05.033>.
- [25] D. Markovics and M. J. Mayer, “Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction,” *Renewable and Sustainable Energy Reviews*, vol. 161, p. 112364, 2022, doi: <https://doi.org/10.1016/j.rser.2022.112364>.
- [26] Y. Hao and C. Tian, “A novel two-stage forecasting model based on error factor and ensemble method for multi-step wind power forecasting,” *Appl Energy*, vol. 238, pp. 368–383, 2019, doi: <https://doi.org/10.1016/j.apenergy.2019.01.063>.
- [27] H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, “A review of deep learning for renewable energy forecasting,” *Energy Convers Manag*, vol. 198, p. 111799, 2019, doi: <https://doi.org/10.1016/j.enconman.2019.111799>.
- [28] Y. Kim and J. Hur, “An Ensemble Forecasting Model of Wind Power Outputs Based on Improved Statistical Approaches,” *Energies (Basel)*, vol. 13, no. 5, 2020, doi: 10.3390/en13051071.
- [29] J. Kim, A. Afzal, H.-G. Kim, C. T. Dinh, and S. G. Park, “Wind power forecasting based on hourly wind speed data in South Korea using machine learning algorithms,” *Journal of Mechanical Science and Technology*, vol. 36, no. 12, pp. 6107–6113, 2022, doi: 10.1007/s12206-022-1125-3.
- [30] Y.-Y. Hong and C. L. P. P. Rioflorido, “A hybrid deep learning-based neural network for 24-h ahead wind power forecasting,” *Appl Energy*, vol. 250, pp. 530–539, 2019, doi: <https://doi.org/10.1016/j.apenergy.2019.05.044>.

- [31] P. , & C. R. Gomes, “Wind speed and wind power forecasting using statistical models: autoregressive moving average (ARMA) and artificial neural networks (ANN). *International Journal of Sustainable Energy Development*, 1(1/2).,” *International Journal of Sustainable Energy Development*, 2012, Accessed: Feb. 13, 2024. [Online]. Available: <https://infonomics-society.org/wp-content/uploads/ijsted/published-papers/volume-1-2012/Wind-Speed-and-Wind-Power-Forecasting-using-Statistical-Models-AutoRegressive-Moving-Average-ARMA-and-Artificial-Neural-Networks-ANN.pdf>
- [32] M. Carolin Mabel and E. Fernandez, “Analysis of wind power generation and prediction using ANN: A case study,” *Renew Energy*, vol. 33, no. 5, pp. 986–992, 2008, doi: <https://doi.org/10.1016/j.renene.2007.06.013>.
- [33] N. and N. R. Singh Pavan Kumar and Singh, “Wind Power Forecasting Using Hybrid ARIMA-ANN Technique,” in *Ambient Communications and Computer Systems*, S. and M. K. K. and T. M. C. Hu Yu-Chen and Tiwari, Ed., Singapore: Springer Singapore, 2019, pp. 209–220.
- [34] K. Shi, Y. Qiao, W. Zhao, Q. Wang, M. Liu, and Z. Lu, “An improved random forest model of short-term wind-power forecasting to enhance accuracy, efficiency, and robustness,” *Wind Energy*, vol. 21, no. 12, pp. 1383–1394, 2018, doi: <https://doi.org/10.1002/we.2261>.
- [35] A. Lahouar and J. Ben Hadj Slama, “Hour-ahead wind power forecast based on random forests,” *Renew Energy*, vol. 109, pp. 529–541, 2017, doi: <https://doi.org/10.1016/j.renene.2017.03.064>.
- [36] Z. Wenting, L. Jie, L. Yalong, and L. Yingyu, “Short-term wind power forecasting model based on random forest algorithm and TCN,” in *2021 40th Chinese Control Conference (CCC)*, 2021, pp. 5776–5781. doi: 10.23919/CCC52363.2021.9550052.
- [37] U. Singh, M. Rizwan, M. Alaraj, and I. Alsaidan, “A Machine Learning-Based Gradient Boosting Regression Approach for Wind Power Production Forecasting: A Step towards Smart Grid Environments,” *Energies (Basel)*, vol. 14, no. 16, 2021, doi: 10.3390/en14165196.
- [38] G. Ponkumar, S. Jayaprakash, and K. Kanagarathinam, “Advanced Machine Learning Techniques for Accurate Very-Short-Term Wind Power Forecasting in Wind Energy Systems Using Historical Data Analysis,” *Energies (Basel)*, vol. 16, no. 14, 2023, doi: 10.3390/en16145459.
- [39] H. Demolli, A. S. Dokuz, A. Ecemis, and M. Gokcek, “Wind power forecasting based on daily wind speed data using machine learning algorithms,” *Energy Convers Manag*, vol. 198, p. 111823, 2019, doi: <https://doi.org/10.1016/j.enconman.2019.111823>.
- [40] M. Khan, C. He, T. Liu, and F. Ullah, “A New Hybrid Approach of Clustering Based Probabilistic Decision Tree to Forecast Wind Power on Large Scales,” *Journal of Electrical Engineering & Technology*, vol. 16, no. 2, pp. 697–710, 2021, doi: 10.1007/s42835-020-00616-1.
- [41] I. Akhtar, S. Kirmani, M. Ahmad, and S. Ahmad, “Average Monthly Wind Power Forecasting Using Fuzzy Approach,” *IEEE Access*, vol. 9, pp. 30426–30440, 2021, doi: 10.1109/ACCESS.2021.3056562.
- [42] Q. Dong, Y. Sun, and P. Li, “A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: A case study of wind farms in China,” *Renew Energy*, vol. 102, pp. 241–257, 2017, doi: <https://doi.org/10.1016/j.renene.2016.10.030>.

- [43] Y. Kassa, J. H. Zhang, D. H. Zheng, and D. Wei, "Short term wind power prediction using ANFIS," in *2016 IEEE International Conference on Power and Renewable Energy (ICPRE)*, 2016, pp. 388–393. doi: 10.1109/ICPRE.2016.7871238.
- [44] M. , & A. A. Vosoogha, "An intelligent power prediction method for wind energy generation based on optimized fuzzy system," 2019. Accessed: Feb. 12, 2024. [Online]. Available: [https://jms.procedia.org/archive/CRPASE\\_169/CRPASE\\_procedia\\_2019\\_5\\_2\\_2.pdf](https://jms.procedia.org/archive/CRPASE_169/CRPASE_procedia_2019_5_2_2.pdf)
- [45] O. Heinermann Justin and Kramer, "Precise Wind Power Prediction with SVM Ensemble Regression," in *Artificial Neural Networks and Machine Learning – ICANN 2014*, C. and D. W. and H. T. and K.-H. P. and M. S. and P. G. and V. A. E. P. Wermter Stefan and Weber, Ed., Cham: Springer International Publishing, 2014, pp. 797–804.
- [46] J. Zeng and W. Qiao, "Support vector machine-based short-term wind power forecasting," in *2011 IEEE/PES Power Systems Conference and Exposition*, 2011, pp. 1–8. doi: 10.1109/PSCE.2011.5772573.
- [47] G. Santamaría-Bonfil, A. Reyes-Ballesteros, and C. Gershenson, "Wind speed forecasting for wind farms: A method based on support vector regression," *Renew Energy*, vol. 85, pp. 790–809, 2016, doi: <https://doi.org/10.1016/j.renene.2015.07.004>.
- [48] F. Kramer Oliver and Gieseke, "Short-Term Wind Energy Forecasting Using Support Vector Regression," in *Soft Computing Models in Industrial and Environmental Applications, 6th International Conference SOCO 2011*, V. and S. J. and H. A. E. and C. J. L. and S. D. Corchado Emilio and Snášel, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 271–280.
- [49] J. Wang, J. Sun, and H. Zhang, "Short-term wind power forecasting based on support vector machine," in *2013 5th International Conference on Power Electronics Systems and Applications (PESA)*, 2013, pp. 1–5. doi: 10.1109/PESA.2013.6828211.
- [50] H. Zheng and Y. Wu, "A XGBoost Model with Weather Similarity Analysis and Feature Engineering for Short-Term Wind Power Forecasting," *Applied Sciences*, vol. 9, no. 15, 2019, doi: 10.3390/app9153019.
- [51] H. Zheng and Y. Wu, "A XGBoost Model with Weather Similarity Analysis and Feature Engineering for Short-Term Wind Power Forecasting," *Applied Sciences*, vol. 9, no. 15, 2019, doi: 10.3390/app9153019.
- [52] Q.-T. Phan, Y.-K. Wu, and Q.-D. Phan, "A Comparative Analysis of XGBoost and Temporal Convolutional Network Models for Wind Power Forecasting," in *2020 International Symposium on Computer, Consumer and Control (IS3C)*, 2020, pp. 416–419. doi: 10.1109/IS3C50286.2020.00113.
- [53] Y. Li *et al.*, "Numerical Weather Prediction Correction Strategy for Short-Term Wind Power Forecasting Based on Bidirectional Gated Recurrent Unit and XGBoost," *Front Energy Res*, vol. 9, 2022, doi: 10.3389/fenrg.2021.836144.
- [54] X. Liu, L. Zhang, Z. Zhang, T. Zhao, and L. Zou, "Ultra Short Term Wind Power Prediction Model Based on WRF Wind Speed Prediction and CatBoost," *IOP Conf Ser Earth Environ Sci*, vol. 838, no. 1, p. 12001, Aug. 2021, doi: 10.1088/1755-1315/838/1/012001.
- [55] Y. Xiao, C. Zou, H. Chi, and R. Fang, "Boosted GRU model for short-term forecasting of wind power with feature-weighted principal component analysis," *Energy*, vol. 267, p. 126503, 2023, doi: <https://doi.org/10.1016/j.energy.2022.126503>.
- [56] Z. Niu, Z. Yu, W. Tang, Q. Wu, and M. Reformat, "Wind power forecasting using attention-based gated recurrent unit network," *Energy*, vol. 196, p. 117081, 2020, doi: <https://doi.org/10.1016/j.energy.2020.117081>.

- [57] V. K. Saini, B. Bhardwaj, V. Gupta, R. Kumar, and A. Mathur, "Gated Recurrent Unit (GRU) Based Short Term Forecasting for Wind Energy Estimation," in *2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)*, 2020, pp. 1–6. doi: 10.1109/ICPECTS49113.2020.9336973.
- [58] C. Li, G. Tang, X. Xue, A. Saeed, and X. Hu, "Short-Term Wind Speed Interval Prediction Based on Ensemble GRU Model," *IEEE Trans Sustain Energy*, vol. 11, no. 3, pp. 1370–1380, 2020, doi: 10.1109/TSTE.2019.2926147.
- [59] R. Zhang *et al.*, "Hybrid Model Based on Isolated Forest and WT-Catboost for Wind Power Forecasting," in *2022 2nd International Conference on Computation, Communication and Engineering (ICCCE)*, 2022, pp. 102–105. doi: 10.1109/ICCCE55785.2022.10036227.
- [60] R. Yu *et al.*, "LSTM-EFG for wind power forecasting based on sequential correlation features," *Future Generation Computer Systems*, vol. 93, pp. 33–42, 2019, doi: <https://doi.org/10.1016/j.future.2018.09.054>.
- [61] M.-S. Ko, K. Lee, J.-K. Kim, C. W. Hong, Z. Y. Dong, and K. Hur, "Deep Concatenated Residual Network With Bidirectional LSTM for One-Hour-Ahead Wind Power Forecasting," *IEEE Trans Sustain Energy*, vol. 12, no. 2, pp. 1321–1335, 2021, doi: 10.1109/TSTE.2020.3043884.
- [62] F. Shahid, A. Zameer, and M. Muneeb, "A novel genetic LSTM model for wind power forecast," *Energy*, vol. 223, p. 120069, 2021, doi: <https://doi.org/10.1016/j.energy.2021.120069>.
- [63] Q. Wu, F. Guan, C. Lv, and Y. Huang, "Ultra-short-term multi-step wind power forecasting based on CNN-LSTM," *IET Renewable Power Generation*, vol. 15, no. 5, pp. 1019–1029, 2021, doi: <https://doi.org/10.1049/rpg2.12085>.
- [64] P. Xu, M. Zhang, Z. Chen, B. Wang, C. Cheng, and R. Liu, "A Deep Learning Framework for Day Ahead Wind Power Short-Term Prediction," *Applied Sciences*, vol. 13, no. 6, 2023, doi: 10.3390/app13064042.
- [65] Y. Ren, P. N. Suganthan, and N. Srikanth, "Ensemble methods for wind and solar power forecasting—A state-of-the-art review," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 82–91, 2015, doi: <https://doi.org/10.1016/j.rser.2015.04.081>.
- [66] N. Majee and A. Roy, "Asymptotic Behavior of an Artificial Neural Network Defined on Multipartite Directed Graph," *Online J Biol Sci*, vol. 10, Feb. 2010.
- [67] S. S. S. S. , & A. A. Sharma, "Activation functions in neural networks.," 2017.
- [68] Z. M. Yaseen, A. El-Shafie, H. A. Afan, M. Hameed, W. H. M. W. Mohtar, and A. Hussain, "RBFNN versus FFNN for daily river flow forecasting at Johor River, Malaysia," *Neural Comput Appl*, vol. 27, no. 6, pp. 1533–1542, 2016, doi: 10.1007/s00521-015-1952-6.
- [69] B. Mostafa, N. El-Attar, S. Abd-Elhafeez, and W. Awad, "Machine and Deep Learning Approaches in Genome: Review Article," *Alfarama Journal of Basic & Applied Sciences*, Apr. 2020, doi: 10.21608/ajbas.2020.34160.1023.
- [70] A. Zhu, X. Li, Z. Mo, and R. Wu, "Wind power prediction based on a convolutional neural network," in *2017 International Conference on Circuits, Devices and Systems (ICCDs)*, 2017, pp. 131–135. doi: 10.1109/ICCDs.2017.8120465.
- [71] S. Abirami and P. Chitra, "Chapter Fourteen - Energy-efficient edge based real-time healthcare support system," in *The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases*, vol. 117, no. 1, P. Raj and P. Evangeline, Eds., in *Advances in Computers*, vol. 117. , Elsevier, 2020, pp. 339–368. doi: <https://doi.org/10.1016/bs.adcom.2019.09.007>.
- [72] J. de J. Serrano-Pérez, G. Fernández-Anaya, S. Carrillo-Moreno, and W. Yu, "New Results for Prediction of Chaotic Systems Using Deep Recurrent Neural Networks,"

- Neural Process Lett*, vol. 53, no. 2, pp. 1579–1596, 2021, doi: 10.1007/s11063-021-10466-1.
- [73] G. Van Houdt, C. Mosquera, and G. Nápoles, “A review on the long short-term memory model,” *Artif Intell Rev*, vol. 53, no. 8, pp. 5929–5955, 2020, doi: 10.1007/s10462-020-09838-1.
- [74] A. A. Alhussan, E.-S. M. El-Kenawy, A. A. Abdelhamid, A. Ibrahim, M. M. Eid, and D. S. Khafaga, “Wind speed forecasting using optimized bidirectional LSTM based on dipper throated and genetic optimization algorithms,” *Front Energy Res*, vol. 11, 2023, doi: 10.3389/fenrg.2023.1172176.
- [75] L. Huang, L. Li, X. Wei, and D. Zhang, “Short-term prediction of wind power based on BiLSTM–CNN–WGAN–GP,” *Soft comput*, vol. 26, no. 20, pp. 10607–10621, 2022, doi: 10.1007/s00500-021-06725-x.
- [76] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD ’16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
- [77] A. Torres-Barrán, Á. Alonso, and J. R. Dorronsoro, “Regression tree ensembles for wind energy and solar radiation prediction,” *Neurocomputing*, vol. 326–327, pp. 151–160, 2019, doi: <https://doi.org/10.1016/j.neucom.2017.05.104>.
- [78] H. Demolli, A. S. Dokuz, A. Ecemis, and M. Gokcek, “Wind power forecasting based on daily wind speed data using machine learning algorithms,” *Energy Convers Manag*, vol. 198, p. 111823, 2019, doi: <https://doi.org/10.1016/j.enconman.2019.111823>.
- [79] A. Torres-Barrán, Á. Alonso, and J. R. Dorronsoro, “Regression tree ensembles for wind energy and solar radiation prediction,” *Neurocomputing*, vol. 326–327, pp. 151–160, 2019, doi: <https://doi.org/10.1016/j.neucom.2017.05.104>.
- [80] L. Wang, X. Zhou, X. Zhu, Z. Dong, and W. Guo, “Estimation of biomass in wheat using random forest regression algorithm and remote sensing data,” *Crop J*, vol. 4, no. 3, pp. 212–219, 2016, doi: <https://doi.org/10.1016/j.cj.2016.01.008>.
- [81] G. Dudek, “Short-Term Load Forecasting Using Random Forests,” in *Intelligent Systems ’2014*, J. and K. J. and K. M. and P. I. and R. L. and S. V. and S. E. and S. P. and Z. S. Filev D. and Jabłkowski, Ed., Cham: Springer International Publishing, 2015, pp. 821–828.
- [82] J. Hu, J. Heng, J. Tang, and M. Guo, “Research and application of a hybrid model based on Meta learning strategy for wind power deterministic and probabilistic forecasting,” *Energy Convers Manag*, vol. 173, pp. 197–209, 2018, doi: <https://doi.org/10.1016/j.enconman.2018.07.052>.
- [83] P. Piotrowski, I. Rutyna, D. Baczyński, and M. Kopyt, “Evaluation Metrics for Wind Power Forecasts: A Comprehensive Review and Statistical Analysis of Errors,” *Energies (Basel)*, vol. 15, no. 24, 2022, doi: 10.3390/en15249657.
- [84] X. Xiong, X. Guo, P. Zeng, R. Zou, and X. Wang, “A Short-Term Wind Power Forecast Method via XGBoost Hyper-Parameters Optimization,” *Front Energy Res*, vol. 10, 2022, doi: 10.3389/fenrg.2022.905155.