



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών
Τομέας Ηλεκτρικής Ισχύος
Εργαστήριο Συστημάτων Ηλεκτρικής ενέργειας

Σύστημα Πρόβλεψης Εισροών-Εκροών Φυσικού Αερίου στην Ελλάδα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ευστράτιος Τσιτούρας

Επιβλέπων: Άρης-Ευάγγελος Δημέας
Επ. Καθηγητής Ε.Μ.Π.

Αθήνα, Φεβρουάριος 2025



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών
Τομέας Ηλεκτρικής Ισχύος
Εργαστήριο Συστημάτων Ηλεκτρικής Ενέργειας

Σύστημα Πρόβλεψης Εισροών-Εκροών Φυσικού Αερίου στην Ελλάδα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ευστράτιος Τσιτούρας

Επιβλέπων: Άρης-Ευάγγελος Δημέας
Επ. Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 25^η Φεβρουαρίου
2025.

.....
.....
Α. Δημέας
Επ. Καθηγητής Ε.Μ.Π.

.....
.....
Γ. Κορρές
Καθηγητής Ε.Μ.Π.

.....
.....
Β. Νικολαΐδης
Επ. Καθηγητής Ε.Μ.Π.

Αθήνα, Φεβρουάριος 2025

.....

Ευστράτιος Τσιπούρας

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

Copyright © (2025) Ευστράτιος Τσιπούρας
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Ένα από τα κρίσιμότερα ζητήματα στη σύγχρονη αγορά ενέργειας είναι η πρόβλεψη της παραγωγής, της ζήτησης και της τιμής εκκαθάρισης της επόμενης μέρας πάνω στο δίκτυο πηγών και διανομής ενέργειας. Πρόκειται για ένα πολυπαραγοντικό πρόβλημα που επηρεάζεται από πληθώρα παραμέτρων σε κάθε στάδιο της διαδρομής του ενεργειακού φορτίου από τους παραγωγούς προς τους καταναλωτές. Ενδεικτικά οι περιβαλλοντικές και τεχνολογικές συνθήκες (μεταξύ άλλων) επηρεάζουν τους κόμβους παραγωγής, οι παγκόσμιες γεωπολιτικές συνθήκες το δίκτυο και τον τρόπο διανομής ενώ οι οικονομικές συνθήκες τη συμπεριφορά των κόμβων ζήτησης πάνω στο δίκτυο της ενέργειας. Δεδομένης της κρίσιμης σημασίας που έχει σήμερα το φυσικό αέριο για την κάλυψη της ζήτησης ηλεκτρικής ενέργειας στην Ευρώπη, γίνεται εμφανές ότι η πρόβλεψη των ροών φυσικού αερίου παρέχει αφενός ένα ισχυρό πλεονέκτημα στον εκάστοτε παίκτη της αγοράς για την αξιοποίηση των αντίστοιχων δεδομένων κατά τη λήψη σχετικών αποφάσεων, και αφετέρου θέτει τα θεμέλια για την εγκαθίδρυση συνθηκών ομαλότητας, εποπτείας και κοινωνικής δικαιοσύνης στο διαμοιρασμό και την κοστολόγησή των πόρων ενέργειας. Στην παρούσα διπλωματική εργασία γίνεται μια απόπειρα μοντελοποίησης και πρόβλεψης αυτών των διακυμάνσεων των ροών φυσικού αερίου με χρήση Μηχανικής Μάθησης. Συγκεκριμένα αξιοποιώντας καταλλήλως δεδομένα από επίσημες και ανοιχτά προσβάσιμες ευρωπαϊκές και κρατικές πηγές καθώς και ιδιωτικά δεδομένα, τροφοδοτήθηκε κι εκπαιδεύθηκε ένα κατάλληλα παραμετροποιημένο εποπτευόμενο νευρωνικό δίκτυο χρησιμοποιώντας επί τω πλείστον μοντέρνα προγραμματιστικά εργαλεία ανοιχτού κώδικα. Γίνεται εκτενής ανάλυση της δομής τόσο του μοντέλου όσο και των δεδομένων εισόδου και εξόδου καθώς και οι μετρικές απόδοσης ως προς την προβλεπτική του ικανότητα σε συνάρτηση με την εκάστοτε παραμετροποίηση. Παράλληλα παρουσιάζονται τα συμπεράσματα που εξήχθησαν ως προς τις τρέχουσες υφιστάμενες δυσκολίες κατά τη συλλογή και αξιοποίηση των απαραίτητων δεδομένων ενώ παρατίθενται ορισμένες κατευθύνσεις επέκτασης και βελτίωσης των αποτελεσμάτων.

Λέξεις-Κλειδιά: Μηχανική Μάθηση, Νευρωνικά Δίκτυα, Αγορά Ενέργειας, Φυσικό Αέριο, Δίκτυο Διανομής Ενέργειας, Εποπτευόμενο Προβλεπτικό Μοντέλο, Χρονοσειρά

Abstract

One of the most critical issues in the modern energy markets is the prediction of the production, demand, and clearing price of the next day on the energy supply and distribution network. This is a multi-faceted problem influenced by a plethora of parameters at every stage of the energy load's journey from producers to consumers. Indicatively, environmental and technological conditions (among others) affect production nodes, global geopolitical conditions affect the network and distribution methods, while economic conditions affect the behavior of demand nodes on the energy network. Given the critical importance of natural gas in meeting Europe's electricity demand, it is evident that forecasting natural gas flows provides, on the one hand, a strong advantage for market players to leverage relevant data in decision-making, and on the other hand, lays the foundation for establishing conditions of fairness, oversight, and social justice in the distribution and pricing of energy resources. This thesis attempts to model and forecast these fluctuations in natural gas flows using Machine Learning. Specifically, by appropriately utilizing data from official and publicly accessible European and state sources as well as private data, a suitably parameterized supervised neural network was fed and trained using state-of-the-art programming tools at each stage of the procedure. An extensive analysis is made of the structure of both the model and the input-output data, as well as the performance metrics in terms of its predictive ability in relation to the respective parameterization. At the same time, the conclusions drawn regarding the current existing difficulties in collecting and utilizing the necessary data are presented, while some directions for the extension and improvement of the results are also provided.

Keywords: Machine Learning, Neural Networks, Energy Market, Natural Gas, Energy Distribution Network, Supervised Predictive Model, Time Series

Ευχαριστίες

Η εκπόνηση της παρούσας διπλωματικής εργασίας έγινε στα πλαίσια της φοίτησής μου στο τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου. Θα ήθελα να ευχαριστήσω τον καθηγητή μου κύριο Άρη-Ευάγγελο Δημέα καθώς και τους κ.κ. Γ.Κορρέ και Β. Νικολαΐδη για τη δυνατότητα που μου έδωσαν να ασχοληθώ με το συγκεκριμένο θέμα καθώς και την εμπιστοσύνη που μου έδειξαν κατά την εκπόνησή του. Παράλληλα θα ήθελα να ευχαριστήσω τον κ. Κυριάκο Ανδρεσάκη για τη βοήθεια και την καθοδήγηση στα διάφορα στάδια της διπλωματικής μου εργασίας. Ακόμη θα ήθελα να ευχαριστήσω φίλους και συναδέλφους για την ψυχική και έμπρακτη υποστήριξή τους. Τέλος θα ήθελα να ευχαριστήσω τους γονείς μου για την συμπαράσταση και την αμέριστη στήριξη τους όλα αυτά τα χρόνια.

Περιεχόμενα

Κατάλογος Σχημάτων.....	10
Κεφάλαιο 1. Εισαγωγή.....	12
1.1 Γενικό Πλαίσιο.....	12
1.2 Μοντέλο Πρόβλεψης.....	12
1.3 Δομή της Εργασίας.....	13
Κεφάλαιο 2. Φυσικό Αέριο Στην ΕΕ.....	15
2.1 Χρήσεις Φυσικού Αερίου στην ΕΕ.....	15
2.2 Τύποι Φυσικού αερίου στην ΕΕ.....	17
2.3 Δίκτυο Αγωγών.....	18
2.4 Ο ρόλος του ENTSOG.....	19
2.5 Παρεχόμενα Δεδομένα από τον ENTSOG.....	20
Κεφάλαιο 3. Μηχανισμός Συλλογής Δεδομένων Φυσικού Αερίου.....	23
3.1 DIEM.....	23
3.2 Συλλογή δεδομένων - DIEM.....	23
3.3 Επεξεργασία δεδομένων - Πρόβλεψη - DIEM.....	24
Κεφάλαιο 4. Νευρωνικά Δίκτυα και Μηχανική Μάθηση.....	25
4.1 Σύντομη εισαγωγή.....	25
4.2 Δομή Νευρώνα.....	26
4.3 Συναρτήσεις Ενεργοποίησης.....	27
4.4 Αρχιτεκτονικές νευρωνικών δικτύων.....	28
4.5 Εκμάθηση νευρωνικών δικτύων.....	31
4.6 Αλγόριθμος Οπισθοδιάδοσης (Back-propagation).....	34
4.7 Ρυθμός μάθησης και υπερ-παράμετροι (Hyper-parameters).....	38
4.8 Κανονικοποίηση των εισόδων.....	40
4.9 Γενίκευση, υπο-προσαρμογή, υπερ-προσαρμογή και cross-validation.....	40
4.10 Η κατάρα της Διαστατικότητας.....	43
Κεφάλαιο 5. Προβλέψεις Ροών Φυσικού Αερίου.....	43
5.1 Τεχνολογίες και εργαλεία.....	43
5.2 Συλλογή δεδομένων για εκπαίδευση.....	44
5.3 Οπτικοποίηση και ανάλυση δεδομένων.....	44
5.4 Επεξεργασία δεδομένων.....	49
5.5 Μηχανική Χαρακτηριστικών (Feature Engineering).....	52
5.6 Εκπαίδευση Μοντέλων.....	53
5.7 Αποτελέσματα.....	55
5.8 Σύγκριση αποτελεσμάτων.....	61
5.9 Συμπεράσματα.....	64
5.10 Επόμενα Βήματα.....	65
Παράρτημα.....	67
Βιβλιογραφία.....	69

Κατάλογος Σχημάτων

Σχήμα 1	Τελική κατανάλωση ενέργειας ανα τομέα	15
Σχήμα 2	Ακαθάριστη Εσωτερική Κατανάλωση Αερίου	17
Σχήμα 3.α	Ροές Φυσικού αερίου στην Ευρώπη	19
Σχήμα 3.β	Πληροφορίες για έναν Διαχειριστή Συστήματος Μεταφοράς Αερίου	22
Σχήμα 4	Μη-γραμμικό μοντέλο τεχνητού νευρώνα k	26
Σχήμα 5	Παραδείγματα συναρτήσεων ενεργοποίησης	27
Σχήμα 6	Δίκτυο πρόσθιας τροφοδότησης ενός επιπέδου	29
Σχήμα 7	Δίκτυο πρόσθιας τροφοδότησης δύο επιπέδων	30
Σχήμα 8	Δίκτυο με βρόχους ανάδρασης μοναδιαίας καθυστέρησης (z^{-1})	31
Σχήμα 9	Το βασικό μοντέλο Ενισχυτικής Μάθησης	32
Σχήμα 10	Κατηγορίες μηχανικής μάθησης	33
Σχήμα 11	Γράφημα ροής σήματος των νευρώνων j,k	35
Σχήμα 12	Επιπτώσεις ενός «μεγάλου» ρυθμού μάθησης	40
Σχήμα 13	Ποιοτικά παραδείγματα 3 διαφορετικών εκπαιδευμένων μοντέλων	41
Σχήμα 14	Ποιοτική αναπαράσταση του σημείου γενίκευσης ενός μοντέλου	42
Σχήμα 15	Οπτικοποίηση των δεδομένων εκπαίδευσης	46
Σχήμα 16	Αυτοσυσχέτιση χρονοσειρών δεδομένων	48
Σχήμα 17	Οπτικοποίηση γραμμικής παρεμβολής	50
Σχήμα 18	Ομοιότητες ελλিপών δεδομένων σε διαφορετικές εισόδους	51
Σχήμα 19	Απόσπασμα αρχείου csv για την περίπτωση 3 εισόδων	51
Σχήμα 20	Απόσπασμα αρχείου csv για την περίπτωση 3*24 εισόδων	52
Σχήμα 21	Αποτελέσματα Εκπαίδευσης Μοντέλων	55
Σχήμα 22	Γραφική παράσταση σφάλματος. Μοντέλο 1, περίπτωση 3 εισόδων	56
Σχήμα 23.α	Γραφική παράσταση σφάλματος. Μοντέλο 2, 3*72 εισόδων	57
Σχήμα 23.β	Γραφική παράσταση σφάλματος. Μοντέλο 3, 3*72 εισόδων	57
Σχήμα 23.γ	Γραφική παράσταση σφάλματος. Μοντέλο 4, 3*72 εισόδων	58
Σχήμα 24	Αποτελέσματα Εκπαίδευσης Επιλεγμένων Μοντέλων με τεχνικές	59

	άμβλυνσης υπερ-προσαρμογής	
Σχήμα 25	Γραφική παράσταση σφάλματος. Μοντέλο 2 με τεχνική Early Stop + Regularization	60
Σχήμα 26	Σύγκριση αποτελεσμάτων μεταξύ νευρωνικού και παλινδρόμησης	61
Σχήμα 27.α	Δεδομένα εκπαίδευσης (Train): Προβλέψεις μοντέλου 2 και γραμμικής παλινδρόμησης για εισροές και εκροές φυσικού αερίου	62
Σχήμα 27.β	Δεδομένα επαλήθευσης (Validation): Προβλέψεις μοντέλου 2 και γραμμικής παλινδρόμησης για εισροές και εκροές φυσικού αερίου	63
Σχήμα 27.γ	Δεδομένα δοκιμής (Test): Προβλέψεις μοντέλου 2 και γραμμικής παλινδρόμησης για εισροές και εκροές φυσικού αερίου	64

Κεφάλαιο 1. Εισαγωγή

1.1 Γενικό Πλαίσιο

Σκοπός της εργασίας είναι η πρόβλεψη της ροής του φυσικού αερίου στην Ελλάδα μέσω της ανάπτυξης ενός μοντέλου Μηχανικής Μάθησης. Για την εκπαίδευση και αξιολόγηση του μοντέλου χρησιμοποιήθηκαν ανοικτά προσβάσιμα δεδομένα διαθέσιμα στο διαδίκτυο από επίσημους κρατικούς και πανευρωπαϊκούς φορείς καθώς και ιδιωτικά δεδομένα, ενώ για την ανάπτυξη του μοντέλου χρησιμοποιήθηκαν ως επί το πλείστον εργαλεία ανοικτού κώδικα.

Αφορμή για την ανάπτυξη του συγκεκριμένου εργαλείου στα πλαίσια της διπλωματικής εργασίας στάθηκε ασφαλώς η καίρια σημασία αυτού του προβλήματος, η επίλυση του οποίου έχει πολύπλευρες δυνατότητες αξιοποίησης και ωφέλειας τόσο στη παρούσα εξειδικευμένη μορφή αλλά και ακόμη περισσότερο στις δυνατές επεκτάσεις του όσον αφορά τη γενικότερη προσφορά και ζήτηση του ενεργειακού φορτίου. Πράγματι σε πρώτο επίπεδο μέσω της πρόβλεψης των ροών φυσικού αερίου βρίσκονται σε πλεονεκτική θέση οι παίκτες της αγοράς (παραγωγοί και καταναλωτές - με διαφορετικό τρόπο ο καθένας) σε οικονομικό επίπεδο καθώς μπορούν να βελτιστοποιήσουν τις κινήσεις τους με σκοπό τόσο την ελαχιστοποίηση των ζημιών (ή και μεγιστοποίηση του κέρδους) όσο και τον καλύτερο προγραμματισμό της παραγωγής/κατανάλωσης τους. Ακόμη περισσότερο ένα τέτοιο εργαλείο έχει άμεσες επιρροές τόσο σε περιβαλλοντικό επίπεδο καθώς η πρόβλεψη των αναγκών σε ενέργεια οδηγεί σε καλύτερη διαχείριση και ελαχιστοποίηση των απωλειών, αλλά και σε κοινωνικό επίπεδο καθώς μπορεί να αποτελέσει ένα μοχλό στο να διευκολύνει τους αρμόδιους θεσμούς στην εποπτεία και ρύθμιση του διαμοιρασμού των ενεργειακών πόρων με διαφάνεια και δικαιοσύνη.

1.2 Μοντέλο Πρόβλεψης

Για το μοντέλο προβλέψεων της παρούσας εργασίας ακολουθήθηκαν σε γενικές γραμμές τα εξής βήματα που αποτελούν τη συνήθη πορεία ανάπτυξης μοντέλων μηχανικής μάθησης στην επιστημονική κοινότητα.

- Συλλογή Ιστορικών και Στατιστικών Δεδομένων:

Η συλλογή των δεδομένων έγινε από επίσημους ευρωπαϊκούς και κρατικούς φορείς μέσω των πλατφόρμων ανοικτού διαμοιρασμού που παρέχουν κάνοντας χρήση κατάλληλου κώδικα καθώς και από ιδιωτικά δεδομένα. Τα δεδομένα αφορούσαν εισροές και εκροές φυσικού αερίου στη χώρα καθώς και έμμεσα επιδραστικά στοιχεία όπως οι προβλέψεις παραγωγής ΑΠΕ και ζήτησης ενέργειας κατά τις σχετικές ημερομηνίες.

- Επεξεργασία-Φιλτράρισμα Δεδομένων:

Μεγάλο και κρίσιμο κομμάτι της διαδικασίας αποτελεί η επεξεργασία των δεδομένων εισόδου σε κατάλληλη μορφή για το μοντέλο. Αυτό αφορά αφενός τόσο την αναγκαία επεξεργασία ώστε να απαλειφθούν τυχόν δύσμορφες εγγραφές όσο και το να γίνει η κατάλληλη διαχείριση ελλιπών δεδομένων ώστε να μην δημιουργούνται απρόσμενα κενά και ανωμαλίες κατά τη τροφοδότηση του μοντέλου. Αφετέρου πρόκειται και για την απαραίτητη μορφοποίηση και κανονικοποίηση των δεδομένων εισόδου σε μορφή που να είναι τόσο πρακτικά (από πλευράς μεγέθους, διαστάσεων και χωρητικότητας) όσο και στατιστικά (από πλευράς επίδοσης και απόδοσης του μοντέλου) κατάλληλη.

- Αλγόριθμος και Αξιολόγηση Πρόβλεψης:

Τέλος στο τρίτο στάδιο με χρήση κάποιου κατάλληλου αλγορίθμου Μηχανικής Μάθησης, εν προκειμένω με χρήση νευρωνικών δικτύων, τροφοδοτείται το μοντέλο με τα δεδομένα εισόδου για την εκπαίδευση του ενώ ταυτόχρονα γίνεται η κατάλληλη ρύθμιση των υπερ-παραμέτρων του μοντέλου για τη βελτιστοποίηση των προβλεπτικών του ικανοτήτων. Η αξιολόγηση των τελευταίων γίνεται με χρήση κατάλληλων μετρικών που αναλύουμε στη συνέχεια και συχνά αποτελεί σημείο ανάδρασης για την ορθότερη ρύθμιση ή/και επεξεργασία της εισόδου μέχρις ότου το μοντέλο να φτάσει σε επιθυμητά επίπεδα ακρίβειας των προβλέψεών του.

1.3 Δομή της Εργασίας

Ακολουθεί μια συνοπτική παρουσίαση της δομής της παρούσας εργασίας μαζί με σύντομη περιγραφή των περιεχομένων κάθε κεφαλαίου:

- Κεφάλαιο 1: Εισαγωγή

Παρουσιάζεται το γενικό πλαίσιο του υπό μελέτη προβλήματος μαζί με τη γενική δομή της υπόλοιπης εργασίας.

- Κεφάλαιο 2: Φυσικό Αέριο στην ΕΕ

Παρατίθενται κάποια βασικά στατιστικά στοιχεία που αφορούν τη παραγωγή και κατανάλωση των διαφόρων μορφών του φυσικού αερίου καθώς και τον τρόπο διακίνησης του στο ευρωπαϊκό δίκτυο. Παράλληλα γίνεται μια εισαγωγή στον

ευρωπαϊκό φορέα ENTSG που εποπτεύει και ρυθμίζει τη διακίνηση του καθώς και στον στενά συνδεδεμένο ENTSG που είναι υπεύθυνος για το διαμοιρασμό της ηλεκτρικής ενέργειας.

- Κεφάλαιο 3: Μηχανισμός Συλλογής Δεδομένων Φυσικού Αερίου

Παρουσιάζεται ο φορέας DIEM που συλλέγει και οργανώνει μεταξύ άλλων και δεδομένα που σχετίζονται με το δίκτυο και τη χρήση του Φυσικού Αερίου, καθώς και τη μορφή και συχνότητα με την οποία τα παρέχει.

- Κεφάλαιο 4: Νευρωνικά Δίκτυα και Μηχανική Μάθηση

Γίνεται μια εισαγωγή στις βασικές αρχές της Μηχανικής Μάθησης όσον αφορά τα στάδια της εκπαίδευσης και ρύθμισης των παραμέτρων καθώς και τα συνήθη ζητήματα που καλείται να αντιμετωπίσει ο ερευνητής του σχετικού πεδίου (εσφαλμένης υπερ/υπο-προσαρμογής, διαστατικότητα, κανονικοποίηση κλπ). Παράλληλα παρουσιάζονται οι βασικές αρχές των Νευρωνικών Δικτύων που αποτελούν το βασικό μοντέλο που χρησιμοποιήθηκε στη παρούσα εργασία μαζί με κάποιες βασικές μαθηματικές έννοιες και αποδείξεις (back-propagation) που αφορούν την εκτέλεση του σχετικού αλγορίθμου.

- Κεφάλαιο 5: Προβλέψεις Κατανάλωσης Φυσικού Αερίου

Σε αυτό το κεφάλαιο παρουσιάζονται τα εργαλεία που χρησιμοποιήθηκαν, σκιαγραφείται η αρχική μορφή και η επεξεργασία που εφαρμόστηκε στα δεδομένα εισόδου καθώς και παρουσιάζονται και αξιολογούνται τα αποτελέσματα του τελικού μοντέλου που αναπτύχθηκε και παραμετροποιήθηκε καταλλήλως. Βασιζόμενοι σε αυτά, γίνεται μια παράθεση των αδυναμιών του μοντέλου, των εφικτών τρόπων βελτίωσης του καθώς και των δυνατών κατευθύνσεων που μπορεί να επεκταθεί και αξιοποιηθεί στο μέλλον.

Στο τέλος της εργασίας παρατίθενται η Βιβλιογραφία που χρησιμοποιήθηκε και το Παράρτημα με τους βασικούς κώδικες που χρειάστηκαν για την επεξεργασία και τον υπολογισμό των δεδομένων καθώς και της κεντρικής ροής εκτέλεσης του μοντέλου.

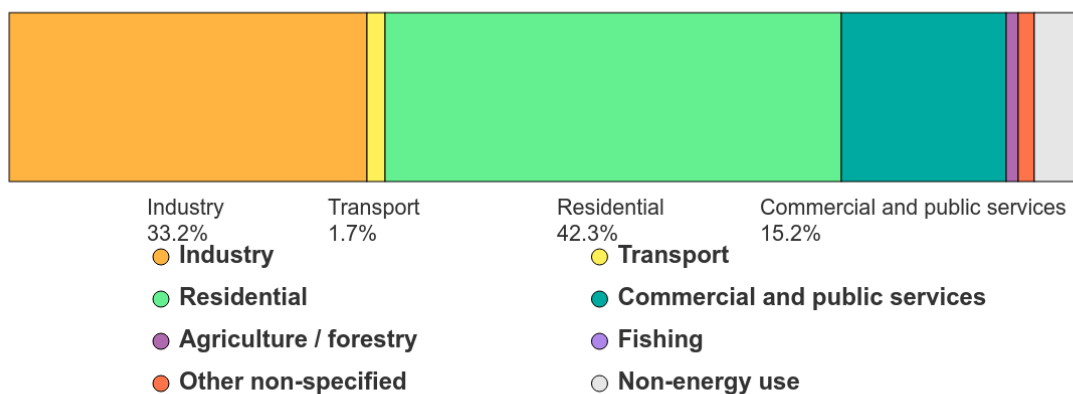
Κεφάλαιο 2. Φυσικό Αέριο Στην ΕΕ

2.1 Χρήσεις Φυσικού Αερίου στην ΕΕ

Το φυσικό αέριο είναι σημαντική πηγή ενέργειας για την ΕΕ, καθώς αποτελεί την μορφή ενέργειας με το 21,5% της συνολικής πρωτογενούς¹ κατανάλωσης ενέργειας. Πρόκειται για την μορφή ενέργειας με το μεγαλύτερο ποσοστό κατανάλωσης για τα νοικοκυριά, κατέχοντας περίπου το 32% της πρωτογενούς κατανάλωσης ενέργειας των νοικοκυριών καθώς και το 40% της ζήτησης αερίου του οικιστικού τομέα το 2021 (European Union Agency for the Cooperation of Energy Regulators (ACER), 2021). Η χρήση του στα νοικοκυριά της ΕΕ κατά συντριπτική πλειοψηφία αφορά θέρμανση (63,5% θέρμανση χώρου και 14,9% θέρμανση νερού) με τις ηλεκτρικές συσκευές και τον φωτισμό να αποτελούν το 13,9% για το 2022 (Eurostat, 2024).

Στην ευρύτερη Ευρώπη γενικότερα παρατηρούμε παρόμοιες τάσεις για τον οικιστικό τομέα, καθώς επίσης ότι ο βιομηχανικός τομέας κατέχει το 33,2%, χωρίς να συμπεριλαμβάνεται η παραγωγή ηλεκτρικής ενέργειας.

Final consumption of gas by sector, Europe, 2022



Source: International Energy Agency. Licence: CC BY 4.0

Σχήμα 1 Τελική κατανάλωση ενέργειας ανα τομέα

Το μερίδιο του φυσικού αερίου για την παραγωγή ηλεκτρικής ενέργειας στην Ευρώπη ανέρχεται στο 20,9% και είναι κρίσιμης σημασίας καθώς τα εργοστάσια παραγωγής ηλεκτρικής ενέργειας με φυσικό αέριο είναι σε θέση να τεθούν σε λειτουργία και να παυθούν σχετικά γρήγορα σε σχέση με άλλες πηγές ενέργειας που αυτονόητα παρουσιάζουν λιγότερη ευελιξία, όπως οι ανανεώσιμες που εξαρτώνται κυρίως από τις καιρικές συνθήκες, δίνοντας έτσι την δυνατότητα να

¹ Ενέργεια που περιέχεται σε ορυκτά καύσιμα και σε ανανεώσιμες πηγές ενέργειας και η οποία δεν έχει υποστεί μετατροπή ή μετασχηματισμό πχ σε ηλεκτρική.

καλυφθούν αιχμές στην ζήτηση ενέργειας και να ισορροπήσει το δίκτυο. (International Energy Agency, n.d.)

Ακόμα, είναι διαχρονικά σημαντική πηγή της ακαθάριστης εσωτερικής κατανάλωσης ενέργειας² της ΕΕ καθώς κατείχε την δεύτερη θέση του ενεργειακού μίγματος το 2000 με ποσοστό 20,6% και εξακολουθούσε να κατέχει την δεύτερη θέση το 2021 με ποσοστό 24,0%. (European Commission: Directorate-General for Energy, 2023, 22).

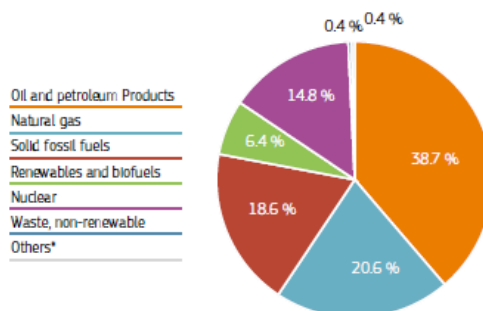
² Ακαθάριστη εσωτερική κατανάλωση ενέργειας = Πρωτογενής παραγωγή + Ανακτημένα & Ανακυκλωμένα προϊόντα + Εισαγωγές — Εξαγωγές + Μεταβολές αποθεμάτων — Διεθνείς ναυτιλιακές δεξαμενές

1.2.3 Gross Inland Consumption

ENERGY MIX (%) – PRIMARY PRODUCTS ONLY

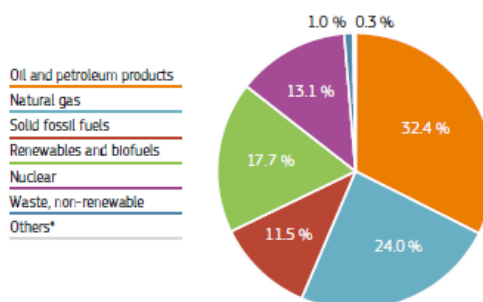
TOTAL PRIMARY PRODUCTS 2000: 1 497.4 Mtoe

(Total Primary and secondary products 2000: 1 498.2 Mtoe)



TOTAL PRIMARY PRODUCTS 2021: 1 423.4 Mtoe

(Total Primary and secondary products 2021: 1 421.6 Mtoe)



*Others = manufactured gases, peat and peat products, oil shale and oil sands

source: Eurostat April 2023

Methodology and Notes: see appendices

22

Σχήμα 2 Ακαθάριστη Εσωτερική Κατανάλωση Αερίου

2.2 Τύποι Φυσικού αερίου στην ΕΕ

Η εγχώρια παραγωγή φυσικού αερίου της ΕΕ για το 2023 αποτελούσε το 10% των αναγκών της, ως εκ τούτου είναι αναγκαία η εισαγωγή του. Η ΕΕ από τον Μάιο του 2022 προσπαθεί να μειώσει την εξάρτηση από το ρωσικό φυσικό αέριο αλλά και γενικότερα το φυσικό αέριο ως καύσιμο και να αυξήσει περαιτέρω τις ανανεώσιμες πηγές ενέργειας. (Directorate-General for Energy, n.d.). Το 2021 το 80%(41% Ρωσία, 39% άλλες χώρες) του αερίου που εισήγαγε η ΕΕ ήταν σε αέρια μορφή μέσω αγωγών υπό πίεση και το 2023 το ποσοστό πλέον είναι 58% (8% Ρωσία, 50% άλλες χώρες). Οι υπόλοιπες εισαγωγές αφορούν το υγροποιημένο φυσικό αέριο (LNG) το οποίο μεταφέρεται μέσω κατάλληλα διαμορφωμένων

πλοίων, και κατά το 2023 οι ΗΠΑ παρείχαν το 46% του υγροποιημένου αερίου που εισήγαγε η ΕΕ.

2.3 Δίκτυο Αγωγών

Η ΕΕ εισάγει το φυσικό αέριο της από ένα εκτενές δίκτυο αγωγών το οποίο της επιτρέπει να διαφοροποιεί τον εφοδιασμό της, επιτρέποντας μια κάποια γεωπολιτική ανεξαρτησία. Η παροχή στα βόρεια αποτελείται από την παροχή φυσικού αερίου από την Βόρεια Θάλασσα, όπου ο κύριος πάροχος είναι η Νορβηγία. Στο παρελθόν υπήρχαν οι αγωγοί NordStream 1 και 2 που παρείχαν φυσικό αέριο από την Ρωσία στην Γερμανία, οι οποίοι όμως καταστράφηκαν από δολιοφθορές το 2022. Από τα ανατολικά, υπάρχουν αγωγοί όπως ο Turkstream, Yamal-Europe και άλλοι που μεταφέρουν αέριο κυρίως από Ρωσία μέσω Λευκορωσίας, Ουκρανίας και Τουρκίας στην ΕΕ, όπου επίσης παρατηρείται σημαντική μείωση μετά και την θέληση της ΕΕ να απεμπλακεί από το ρωσικό αέριο. Από την Κασπία Θάλασσα έρχεται ο αγωγός TANAP, μέρος του Νότιου Διαδρόμου Φυσικού Αερίου ο οποίος παρέχει αέριο κυρίως στις Νοτιοανατολικές Χώρες της Ευρώπης και αποτέλεσε το 2,4% και 3,4% των συνολικών εισαγωγών της ΕΕ για το 2022 και 2023 αντίστοιχα (Directorate-General for Energy, n.d.). Ακόμα, από την Βόρεια Αφρική η ΕΕ προμηθεύεται αέριο από αγωγούς σε Λιβύη, Αλγερία, Τυνησία και Μαρόκο. Τέλος, υπάρχουν διάφορες μονάδες αποθήκευσης και επαναεριοποίησης (FSRU) ανά την επικράτεια της ΕΕ που προμηθεύονται υγροποιημένο φυσικό αέριο από ναυτικές οδούς, και το μετατρέπουν σε φυσικό αέριο κατάλληλο να μεταφερθεί μέσω αγωγών. (ENTSOG, n.d.)

Flows Map

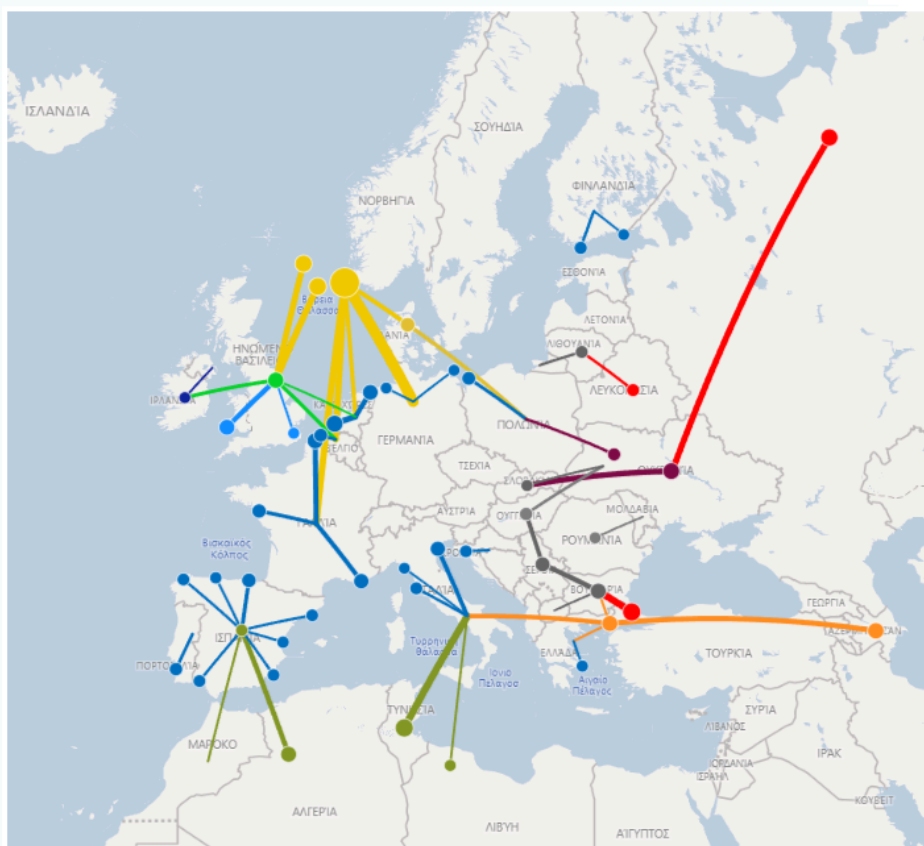
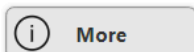
Corridors

LNG

Flows to Europe

Flows from EU

Last Updated: 10/13/2024



Σχήμα 3.α) Ροές Φυσικού αερίου στην Ευρώπη (ENTSOG Gas Dashboard, n.d.)

2.4 Ο ρόλος του ENTSOG

Ο οργανισμός ENTSOG (Ευρωπαϊκά δίκτυα διαχειριστών συστήματος μεταφοράς αερίου) ιδρύθηκε ως αποτέλεσμα του 3ου ενεργειακού πακέτου (Directorate for the Media, 2008) και έχει σκοπό την εύρυθμη λειτουργία των αγορών αερίου, την βελτίωση της ενεργειακής ασφάλειας της ΕΕ και την ενίσχυση της συνεργασίας των διαχειριστών αερίου. Η αποστολή του περιλαμβάνει την ανάπτυξη ενός πανευρωπαϊκού συστήματος μεταφοράς που να ικανοποιεί τις ανάγκες, τις νομικές πρωτοβουλίες και τις αποφάσεις της ΕΕ. Μέρος αυτών είναι η απανθρακοποίηση, αξιοποιώντας ανανεώσιμα και χαμηλής ρύπανσης αέρια καύσιμα. Μέρος των καθηκόντων του είναι η ανάπτυξη και η εφαρμογή κωδικών δικτύου που ορίζουν θέματα όπως η κατανομή χωρητικότητας, η συνδεσιμότητα και η ασφάλεια του δικτύου απέναντι σε κακόβουλους οργανισμούς. Ακόμη, κάνει προβλέψεις για την προσφορά και ζήτηση φυσικού αερίου καθώς επίσης παρέχει εργαλεία για την διαφάνεια και τον συντονισμό των λειτουργιών του δικτύου, (ENTSOG, n.d.) και αναλαμβάνει την εκπόνηση δεκαετών σχεδίων ανάπτυξης των δικτύων φυσικού αερίου. Τέλος, μέχρι και την ίδρυση του ENNOH μέσα στο 2025, ο ENTSOG θα

εκπονεί και τα δεκαετή σχέδια ανάπτυξης υδρογόνου (ENNOH Website, n.d.)(Hancher et al., 2022, 50).

2.5 Παρεχόμενα Δεδομένα από τον ENTSOG

Σύμφωνα με το Παράρτημα Ι Παράγραφος 3 του κανονισμού της Ευρωπαϊκής Ένωσης σχετικά με τις εσωτερικές αγορές ανανεώσιμου αερίου, φυσικού αερίου και υδρογόνου [...] (Regulation (EU) 2024/1789) ο ENTSOG σε συνεργασία με τους διαχειριστές συστήματος μεταφοράς, οφείλει να παρέχει δεδομένα «σε ισότοπο προσβάσιμο δωρεάν από το κοινό και χωρίς εγγραφή ή άλλη διαδικασία εξακρίβωσης από τον διαχειριστή συστήματος μεταφοράς [...] σε τακτική/κυλιόμενη βάση· η συχνότητα εξαρτάται από τις αλλαγές που επέρχονται και από τη διάρκεια της υπηρεσίας· [...] με φιλικό προς τον χρήστη τρόπο· [...] με ουσιώδη, ποσοτικώς σαφή και εύκολα προσβάσιμο τρόπο που δεν συνεπάγεται διακρίσεις [...] σε συμβατικές μονάδες, και συγκεκριμένα σε kWh (σε θερμοκρασία καύσης αναφοράς 298,15 K) ως τη μονάδα ενεργειακού περιεχομένου και σε m³ (σε 273,15 K και 1,01325 bar) ως τη μονάδα όγκου.[...]».

Τα παρεχόμενα δεδομένα αφορούν σχετικά σημεία, τα οποία ορίζονται ως όλα τα σημεία εισόδου και εξόδου προς και από το δίκτυο του εκάστοτε Διαχειριστή Συστήματος Μεταφοράς, εξαιρουμένων των σημείων εξόδου που συνδέονται με μεμονωμένους τελικούς καταναλωτές και των σημείων εισόδου που συνδέονται άμεσα με εγκαταστάσεις παραγωγής ενός μόνο παραγωγού εντός της ΕΕ. Τα παραπάνω εξαιρούμενα σημεία θα αντιπροσωπεύονται ως δύο σημεία, ένα σημείο με συγκεντρωτική μορφή για όλους τους μεμονωμένους τελικούς καταναλωτές, και ένα σημείο με συγκεντρωτική μορφή για όλους τους “μόνους” παραγωγούς. Ακόμα, σχετικά σημεία είναι και αυτά τα οποία συνδέουν ζώνες εξισορρόπησης διαχειριστών συστήματος, καθώς και κάθε σημείο σύνδεσης του δικτύου ενός διαχειριστή συστήματος μεταφοράς με τερματικούς σταθμούς LNG, κόμβους παράδοσης φυσικού αερίου, αποθηκευτικές εγκαταστάσεις και μονάδες παραγωγής που δεν εμπίπτουν στον παραπάνω κανόνα των μεμονωμένων παραγωγών/καταναλωτών.

Με βάση τα παραπάνω, ο ENTSOG διαθέτει πλατφόρμα διαφάνειας (*ENTSOG Transparency Platform*) όπου δημοσιεύει σχετικά δεδομένα σε ωριαία ή ημερήσια βάση και όταν τα υπάρχοντα δεδομένα αφορούν περιόδους μικρότερες της ημέρας, τότε δημοσιεύονται τα αντίστοιχα δεδομένα και με αναγωγή σε ημερήσια βάση. Αυτά τα δεδομένα αφορούν τα παραπάνω ορισμένα σχετικά σημεία, και μεταξύ άλλων είναι πληροφορίες δυναμικότητας, λειτουργικά δεδομένα, διακοπές προγραμματισμένες και μη, τιμολογήσεις υπηρεσιών, πληροφορίες ποιότητας φυσικού αερίου και επείγοντα μηνύματα. (Stannov, 2018)

Αναλυτικότερα, οι πληροφορίες δυναμικότητας (capacity) έχουν να κάνουν με πληροφορίες για την μέγιστη αμετάβλητη δυνατή ροή φυσικού αερίου που μπορεί να προσφερθεί στους χρήστες του δικτύου εντός λογικών πλαισίων για την ασφάλεια του δικτύου και των εγκαταστάσεων, την αμετάβλητη δυνατή ροή μεταφοράς και διανομής φυσικού αερίου και υδρογόνου για την οποία υπάρχει συμβατική εγγύηση του όποιου διαχειριστή συστήματος, την δυναμικότητα μεταφοράς που είναι διακοπτόμενη βάσει κάποιων προκαθορισμένων όρων σύμβασης μεταφοράς, στοιχεία χωρητικότητας ταμειυτήρων φυσικού αερίου και άλλες παρεμφερείς πληροφορίες.

Οι πληροφορίες λειτουργικότητας μεταξύ άλλων είναι οι πραγματικές ροές φυσικού αερίου μέσα σε δοσμένα χρονικά πλαίσια, οι πραγματικές ροές αερίου που προτίθενται να εισάγουν ή να εξάγουν από το σύστημα χρήστες του δικτύου, οι οποίες γνωστοποιούνται προς το διαχειριστή πριν πραγματοποιηθεί η έγχυση/απόσυρση, καθώς επίσης και όποιες διορθώσεις αυτών.

Οι πληροφορίες διακοπής αφορούν προγραμματισμένες και πραγματικές διακοπές των διακοπτόμενων ροών καθώς επίσης προγραμματισμένες ή/και απροσδόκητες διακοπές στις αμετάβλητες υπηρεσίες. Ακόμα, περιέχονται πληροφορίες για επερχόμενες διαδικασίες συντήρησης και η πιθανή χρονική διάρκεια αυτών.

Οι πληροφορίες τιμολόγησης είναι σχετικές παραδείγματος χάριν με την τιμολόγηση της ροής, όπου υπάρχουν διαφοροποιήσεις στην τιμολόγηση ανάλογα με το αν είναι αδιάλειπτη ή διακοπτόμενη ροή, αλλά και με το αν το συμβόλαιο είναι ετήσιο, τριμηνιαίο, μηνιαίο, ημερήσιο ή και ενδοημερήσιο. Επίσης υπάρχουν πληροφορίες τιμολόγησης των δημοπρασιών που πραγματοποιούνται καθώς και άλλων που μπορούν να αφορούν διαφορετικούς λόγους ανά διαχειριστή δικτύου όπως κατά περίπτωση χρεώσεις μετατροπής ποιότητας, χρεώσεις βιοαερίου κλπ. Ακόμα δίνονται μελλοντικές προβλέψεις τιμολογήσεων.

Points	Zones	Balancing Info	Tariff Info	Capacity Info	General Info
--------	-------	----------------	--------------------	---------------	--------------

General Remarks

On 1 January 2017 bayernets GmbH tariffs for transport capacity as well as for metering operation, metering and billing will be adjusted. The network tariffs are reported on a net basis and are subject to additional statutory VAT. They include all applicable network tariffs from the selected entry/exit point to the virtual point (VP) of the market area.

Specific Remarks

Tariff Type	Remark
Firm Capacity	Regulated tariff. Details see price sheet: https://www.bayernets.de/en/transparency/tariffs
Interruptible Capacity	Regulated tariff. Details see price sheet: https://www.bayernets.de/en/transparency/tariffs
Auction	Capacity shall be auctioned according to Regulation (EU) No 2017/459 via PRISMA booking platform. The reserve price for each auction shall be the regulated tariff. Auction premiums shall apply according to the auction algorithms if the total demand for available capacity exceeds the total offer.
Commodity	n/a
Others	In addition to the regulated tariffs there are network point specific additional costs levied such as cost for Metering Operation, Metering and Billing, Quality Conversion Charge and Biogas levy. Details see price sheet: https://www.bayernets.de/en/transparency/tariffs

Σχήμα 3.β) Πληροφορίες για έναν Διαχειριστή Συστήματος Μεταφοράς Αερίου «Bayernets» (*ENTSOG Transparency Platform*, n.d.)

Επίσης, παρέχονται πληροφορίες ενδείξεων ποιότητας φυσικού αερίου όπως η μεικτή θερμογόνος αξία, ο δείκτης Wobbe και η περιεκτικότητα σε οξυγόνο που είναι χρήσιμα για την σύγκριση του ενεργειακού περιεχομένου του φυσικού αερίου μεταξύ διαφορετικών σημείων/διαχειριστών.

Τέλος, οι χρήστες έχουν την δυνατότητα να δουν έκτακτες ενημερώσεις (UMM) που αφορούν το δίκτυο.

Κεφάλαιο 3. Μηχανισμός Συλλογής Δεδομένων Φυσικού Αερίου

3.1 DIEM

Το DIEM είναι μία διαδικτυακή πλατφόρμα που αναπτύχθηκε με σύμπραξη της SmartRUE και της Enargia A.G. Η SmartRUE (Smart Grids Research Unit ECE NTUA) ιδρύθηκε από τον καθηγητή Νίκο Χατζηαργυρίου και είναι μια από τις ερευνητικές ομάδες του Ινστιτούτου Επικοινωνιών και Υπολογιστικών Συστημάτων που συνδέεται με το Εργαστήριο Ηλεκτρικής Ενέργειας (EESL) της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου και εστιάζει στη βέλτιστη λειτουργία συστημάτων διανομής, τη συμμετοχή σε ενεργειακές αγορές, την ανθεκτικότητα του δικτύου ηλεκτρικής ενέργειας, τα προηγμένα συστήματα ελέγχου για έξυπνα δίκτυα, μικροδίκτυα, σε κανονιστικά πλαίσια ενεργειακών κοινοτήτων, έξυπνους μετρητές, ηλεκτρικά οχήματα, διαχείριση από πλευράς φορτίου και άλλες παρεμφερείς δραστηριότητες. (*Smart RUE of ECE-NTUA*, n.d.). Η Enargia A.G. είναι ελβετική εταιρία που εστιάζει στην προηγμένη ενεργειακή ανάλυση, παρέχοντας λύσεις βασισμένες σε δεδομένα για τη βελτιστοποίηση των ενεργειακών συστημάτων. Οι υπηρεσίες της περιλαμβάνουν λεπτομερή ανάλυση ενεργειακού συστήματος, μοντελοποίηση και συμβουλευτικές υπηρεσίες για την υποστήριξη στρατηγικών βιώσιμης ενέργειας για πελάτες (Enargia A.G., n.d.).

Σκοπός της πλατφόρμας DIEM είναι να παρέχει υπηρεσίες σχετικές με τις αγορές ενέργειας σε ενδιαφερόμενους πελάτες και απευθύνεται κυρίως στους ενεργά συμμετέχοντες στην ενεργειακή αγορά. Τέτοιες υπηρεσίες είναι η παροχή ενεργειακών δεδομένων σε μορφή φιλική προς τον χρήστη, μέσω διεπαφών API για την εύκολη ενσωμάτωση δεδομένων σε εργαλεία του χρήστη, ιστοσελίδων με συγκεντρωτικά και γεωχωρικά γραφήματα που συμπεριλαμβάνουν δεδομένα και προβλέψεις που ενημερώνονται σε πραγματικό χρόνο καθώς επίσης και εργαλεία ανάλυσης δεδομένων για διάφορους συνδυασμούς και χρονικές περιόδους. Επίσης, σε συγκεκριμένους χρήστες παρέχεται η δυνατότητα πρόσβασης σε ενεργειακά δεδομένα ιδιωτικών φορέων.

3.2 Συλλογή δεδομένων - DIEM

Η πλατφόρμα του DIEM συλλέγει σχετικά δεδομένα από διάφορες δημόσιες και ιδιωτικές πηγές. Όσον αφορά τις δημόσιες πηγές ηλεκτρικής ενέργειας και φυσικού αερίου, μερικές από αυτές είναι οι: Όμιλος Χρηματιστηρίου Ενέργειας, ο ΑΔΜΗΕ, η ΔΑΠΕΕΠ ΑΕ, η ΔΕΣΦΑ ΑΕ ο ENTSOG και ο ENTSO-e.

Ο Όμιλος Χρηματιστηρίου Ενέργειας διαχειρίζεται και εποπτεύει τις αγορές ηλεκτρικής ενέργειας, και φυσικού αερίου στην Ελλάδα. Αναλαμβάνει την εκκαθάριση και το διακανονισμό συναλλαγών στις παραπάνω αγορές (EnExGroup, n.d.).

Ο ΑΔΜΗΕ (Ανεξάρτητος Διαχειριστής Μεταφοράς Ηλεκτρικής Ενέργειας) είναι υπεύθυνος για τη διαχείριση, τη συντήρηση και την ανάπτυξη του Συστήματος Μεταφοράς Ηλεκτρικής Ενέργειας στην Ελλάδα, με στόχο τη διασφάλιση του επαρκούς, αποδοτικού και αξιόπιστου εφοδιασμού της χώρας με ηλεκτρική ενέργεια. Παράλληλα, έχει τη διαχείριση της Αγοράς Εξισορρόπησης και του διασυννοριακού εμπορίου, με γνώμονα τις αρχές της διαφάνειας, της ισότητας και του ελεύθερου ανταγωνισμού (ΑΔΜΗΕ, n.d.).

Η ΔΑΠΕΕΠ ΑΕ έχει σκοπό την διαχείριση και λειτουργία της ελληνικής αγοράς ηλεκτρικής ενέργειας που παράγεται από Ανανεώσιμες Πηγές Ενέργειας (ΑΠΕ) και Μονάδες Συμπαράγωγής Ηλεκτρισμού και Θερμότητας Υψηλής Απόδοσης (ΣΗΘΥΑ) στο πλαίσιο του Εθνικού Διασυνδεδεμένου Συστήματος και του Δικτύου Μεταφοράς Ηλεκτρικής Ενέργειας, καθώς και την έκδοση Εγγυήσεων Προέλευσης για ηλεκτρική ενέργεια παραγόμενη από ΑΠΕ και ΣΗΘΥΑ (ΔΑΠΕΕΠ, n.d.).

Ο σκοπός του ΔΕΣΦΑ είναι η λειτουργία, η διαχείριση, η εκμετάλλευση και η ανάπτυξη του Εθνικού Συστήματος Φυσικού Αερίου (ΕΣΦΑ) και των διασυνδέσεών του, με γνώμονα το οικονομικό συμφέρον και την τεχνική αρτιότητα καθώς και η βέλτιστη εξυπηρέτηση των χρηστών του με ασφάλεια, αξιοπιστία και επάρκεια, ενώ παράλληλα συμβάλλει στην ασφάλεια εφοδιασμού, στη διαφοροποίηση των πηγών τροφοδοσίας και στην ανάπτυξη ανταγωνισμού στην ελληνική αγορά ενέργειας, με στόχο τη μείωση των εκπομπών αερίων θερμοκηπίου (ΔΕΣΦΑ, n.d.).

Για τον σκοπό ύπαρξης του ENTSGO έχουμε προαναφερθεί στην παράγραφο 2.4, ενώ ο ENTSO-e είναι ο αντίστοιχος οργανισμός που ιδρύθηκε για τα δίκτυα ηλεκτρικής ενέργειας όλης της ΕΕ με σκοπό την εξασφάλιση της ασφάλειας και αξιοπιστίας του διασυνδεδεμένου ευρωπαϊκού συστήματος ηλεκτρικής ενέργειας. Συντονίζει και αναπτύσσει το δίκτυο μεταφοράς στην Ευρώπη, προάγει την ενσωμάτωση ανανεώσιμων πηγών ενέργειας και συμβάλλει στην ανάπτυξη της εσωτερικής ενεργειακής αγοράς. Επίσης, στηρίζει την απανθρακοποίηση του τομέα μέσω σχεδιασμού συστημάτων και ανάπτυξης τεχνικών προτύπων (ENTSO-e, n.d.).

3.3 Επεξεργασία δεδομένων - Πρόβλεψη - DIEM

Τα δεδομένα που συλλέγονται από τις παραπάνω πηγές επεξεργάζονται οι data analysts του DIEM έτσι ώστε η πλατφόρμα να παρέχει προβλέψεις για Day Ahead Market, 48 ωρών και παραπάνω για την παραγωγή ΑΠΕ καθώς και του ηλεκτρικού φορτίου. Επίσης, παρέχονται ενδοημερήσιες προβλέψεις παραγωγής και ζήτησης ανανεώσιμων πηγών για την υποστήριξη των συμμετεχόντων στην ελληνική αγορά ενέργειας όπως επίσης και προσαρμοσμένες λύσεις για πρόβλεψη παραγωγής ΑΠΕ έως και ένα μήνα μετά.

Κεφάλαιο 4. Νευρωνικά Δίκτυα και Μηχανική Μάθηση

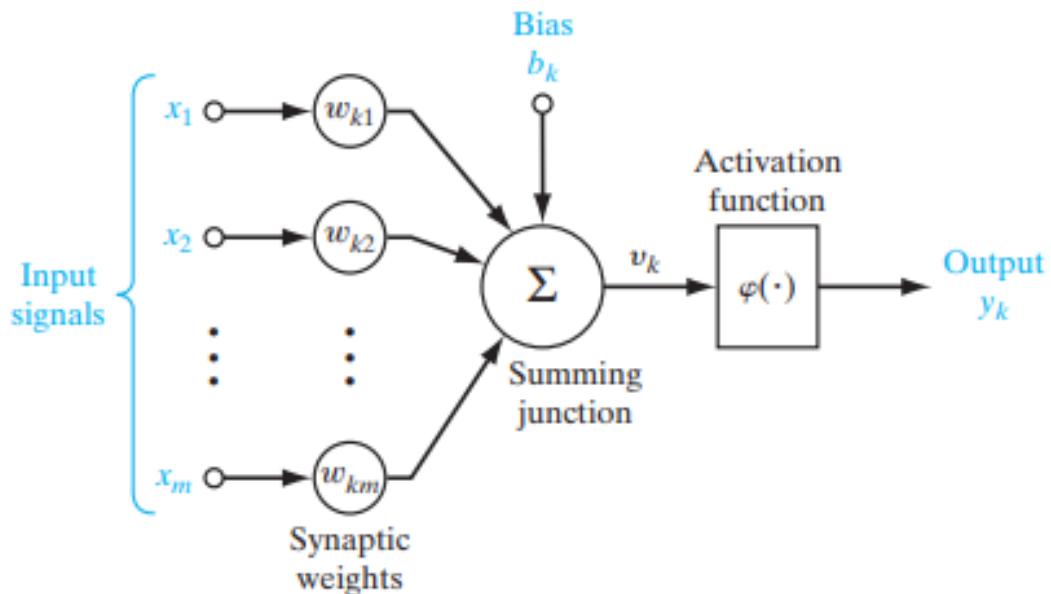
4.1 Σύντομη εισαγωγή

«Ένα νευρωνικό δίκτυο είναι ένας τεράστιος παράλληλος επεξεργαστής με κατανεμημένη αρχιτεκτονική ο οποίος αποτελείται από απλές μονάδες επεξεργασίας και έχει από τη φύση του τη δυνατότητα να αποθηκεύει εμπειρική γνώση και να την καθιστά διαθέσιμη για χρήση. Μοιάζει με τον ανθρώπινο εγκέφαλο σε δύο σημεία:

- 1) το δίκτυο προσλαμβάνει τη γνώση από το περιβάλλον του μέσω μιας διαδικασίας μάθησης,
 - 2) η ισχύς των συνδέσεων μεταξύ των νευρώνων που αποκαλείται συναπτικό βάρος, χρησιμοποιείται για την αποθήκευση της γνώσης που αποκτάται
- » (Aleksander, 1990)

Τα νευρωνικά δίκτυα σύμφωνα με τα παραπάνω είναι χρήσιμα για την εύρεση λύσεων που προσεγγίζουν την ιδανική λύση σε δυσεπίλυτα αλγοριθμικά προβλήματα, καθώς αντί για την επινόηση και υλοποίηση του σωστού και ικανοποιητικά γρήγορου αλγορίθμου, ένας επιστήμονας δεδομένων μπορεί να ορίσει ένα νευρωνικό δίκτυο το οποίο θα εκπαιδεύσει δίνοντάς του ένα σύνολο δεδομένων για τα οποία γνωρίζει την λύση, με την προσδοκία ότι μετά την εκπαίδευση του νευρωνικού δικτύου με κατάλληλες παραμέτρους και μεθόδους, αυτό θα είναι σε θέση εκτός από το να προβλέπει τις λύσεις στα δεδομένα εισόδου που έμαθε, να είναι σε θέση να γενικεύει και να προβλέπει λύσεις σε δεδομένα εισόδου που δεν έχει ακόμα συναντήσει.

4.2 Δομή Νευρώνα



Σχήμα 4 Μη-γραμμικό μοντέλο τεχνητού νευρώνα k (Haykin, 2009, 11)

Ο τεχνητός νευρώνας, εφεξής νευρώνας, είναι το δομικό συστατικό ενός νευρωνικού δικτύου. Στο σχήμα 4 βλέπουμε το μοντέλο ενός τέτοιου νευρώνα που θα ονομάσουμε k . Μπορούμε να διαχωρίσουμε τρία βασικά συστατικά που ορίζουν τον νευρώνα:

Αρχικά, έχουμε στα αριστερά του σχήματος 4 ένα σύνολο συνάψεων (ή αλλιώς συνδέσεων), καθεμία από τις οποίες χαρακτηρίζεται από ένα βάρος w . Συγκεκριμένα, ένα σήμα x_j στην είσοδο της σύναψης j , όπου $j=0,1,\dots,m$ που συνδέεται με τον νευρώνα k , πολλαπλασιάζεται με το συναπτικό βάρος w_{kj} .

Επισημαίνεται ότι ο τρόπος γραφής των δεικτών στο συναπτικό βάρος έχει την εξής σημασία: Ο δείκτης k αναφέρεται στο όνομα του συγκεκριμένου νευρώνα, ενώ ο δεύτερος δείκτης j αναφέρεται στην είσοδο της σύναψης με την οποία σχετίζεται το βάρος w .

Έπειτα, ένας αθροιστής αναλαμβάνει το άθροισμα των σημάτων εισόδου, όπου σταθμίζονται από τα αντίστοιχα συναπτικά βάρη του νευρώνα. Μετά τον αθροιστή εφαρμόζεται η πόλωση, που είναι μία σταθερή τιμή. Οι λειτουργίες αυτές συνιστούν έναν γραμμικό συνδυαστή.

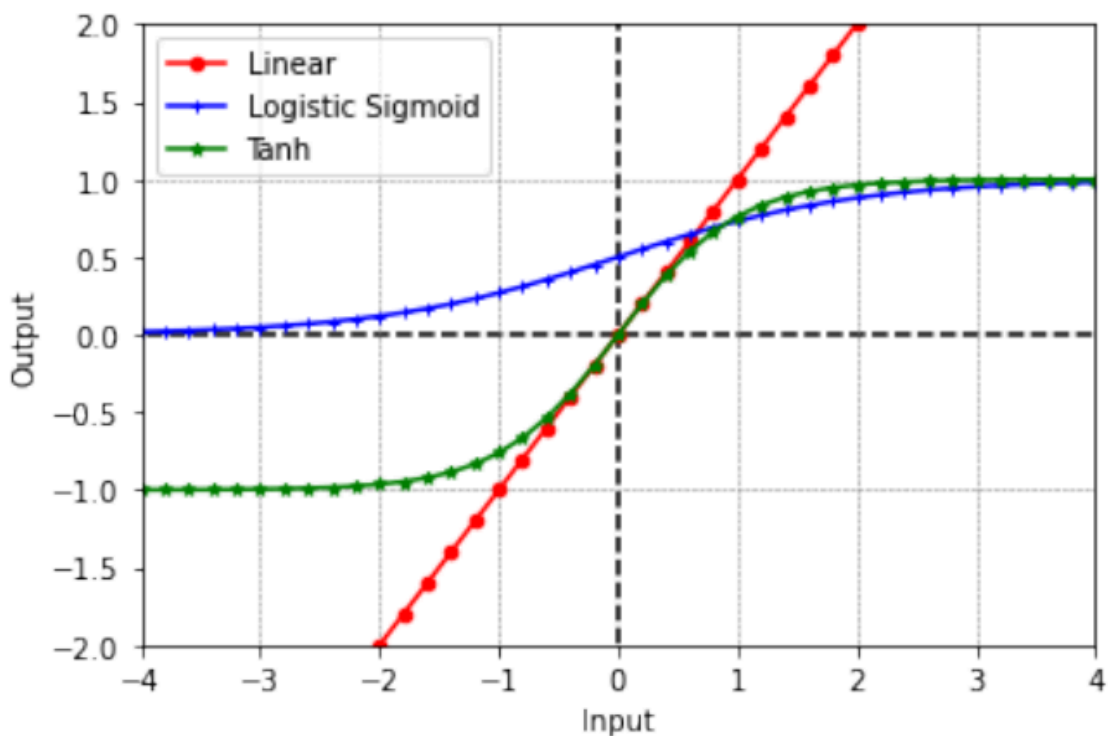
Τέλος, στα δεξιά του σχήματος 4 έχουμε την συνάρτηση ενεργοποίησης φ (activation function) η οποία φροντίζει για τον περιορισμό του εύρους της εξόδου ενός νευρώνα. Η συνάρτηση ενεργοποίησης ονομάζεται επίσης συνάρτηση

συμπίεσης (squashing function), καθώς περιορίζει (συμπιέζει) το επιτρεπτό εύρος πλάτους του σήματος εξόδου σε μια πεπερασμένη τιμή. Η λειτουργία της συνάρτησης ενεργοποίησης έχει σημαντικό ρόλο καθώς επιτρέπει την αναπαράσταση μη-γραμμικών συναρτήσεων στην έξοδο του νευρώνα, και αυτό με την σειρά του επιτρέπει την επίλυση γραμμικών αλλά κυρίως και μη-γραμμικών προβλημάτων.

Με μαθηματικούς όρους, τα παραπάνω συνοψίζονται ως εξής:

$$y_k = \phi\left(b_k + \sum_{j=1}^m w_{kj}x_j\right)$$

4.3 Συναρτήσεις Ενεργοποίησης



Σχήμα 5 Παραδείγματα συναρτήσεων ενεργοποίησης (Dubey et al., 2021)

Θα παρουσιαστούν 4 συναρτήσεις ενεργοποίησης, όμως πρέπει να σημειωθεί ότι υπάρχουν περισσότερες από αυτές.

1) Συνάρτηση κατωφλίου (threshold):

Πρόκειται για την μοναδιαία βηματική συνάρτηση ή αλλιώς συνάρτηση Χέβισαϊντ (Heaviside), και ορίζεται ως:

$$\varphi(x) = \begin{cases} 1 & \text{αν } x \geq 0 \\ 0 & \text{αν } x < 0 \end{cases}$$

Λόγω της μη-διαφορισιμότητας αυτής της συνάρτησης, δεν την συναντάμε συχνά σε νευρωνικά. Ο λόγος είναι ότι ο αλγόριθμος οπισθοδιάδοσης που θα αναλύσουμε αργότερα, χρειάζεται διαφορίσιμες συναρτήσεις για την λειτουργία του. Παρουσιάζεται ως αναφορά στο έργο των McCulloch-Pitts που μοντελοποίησαν μαθηματικά την συνάρτηση ενεργοποίησης των βιολογικών νευρώνων ως “όλα ή τίποτα”. (McCulloch & Pitts, 1990, 99)

2) Λογιστική συνάρτηση (logistic sigmoid):

Η αγγλική ονομασία αυτής της συνάρτησης προέρχεται από την τροχιά «S» που δίνει η γραφική της παράσταση (βλέπε Σχήμα 5). Η συνάρτηση αυτή που χρησιμοποιείται ευρέως στα νευρωνικά δίκτυα έχει την μορφή:

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

όπου α η παράμετρος κλίσης της συνάρτησης. Αν το $\alpha \rightarrow \infty$ τότε εκφυλίζεται σε μια συνάρτηση κατωφλίου. Σε αντίθεση όμως με την συνάρτηση κατωφλίου, η σιγμοειδής συνάρτηση είναι διαφορίσιμη. Επίσης, επειδή η έξοδός της είναι μεταξύ 0 και 1, είναι ιδιαίτερα χρήσιμη για προβλήματα κατηγοριοποίησης όπου η βεβαιότητα για την πρόβλεψη του μοντέλου μας δίνεται ως ποσοστό.

3) Γραμμική συνάρτηση (linear):

Η γνωστή συνάρτηση $\varphi(x) = \alpha x + \beta$. Έχει σχετικά μικρή χρησιμότητα για τα νευρωνικά, καθώς ο γραμμικός μετασχηματισμός που επιφέρει αυτή η συνάρτηση μπορεί να εφαρμοστεί ισοδύναμα στα συναπτικά βάρη και στην πόλωση. Άλλωστε η ισχύς των νευρωνικών δικτύων είναι η μη-γραμμικότητα. Συχνά χρησιμοποιείται στην έξοδο ενός νευρωνικού για να επισημανθεί ότι το πεδίο τιμών είναι μεταξύ $(-\infty, +\infty)$.

4) Συνάρτηση ReLU (rectified linear unit):

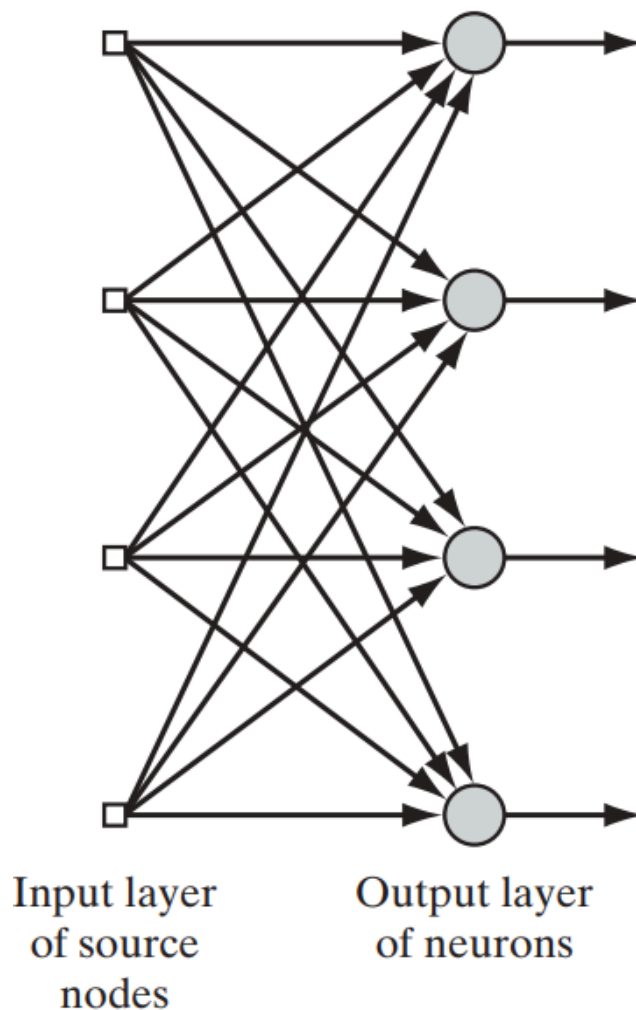
Αυτή η συνάρτηση είναι πολύ διαδεδομένη στα νευρωνικά λόγω της απλότητας και της εύκολης διαφορισιμότητάς της. Ο τύπος της είναι:

$$\varphi(x) = \begin{cases} x & \text{αν } x \geq 0 \\ 0 & \text{αν } x < 0 \end{cases}$$

Ο λόγος που η συνάρτηση ενεργοποίησης ReLU είναι χρησιμότερη της γραμμικής, είναι η μη-γραμμικότητά της που προέρχεται από την **τμηματικά** γραμμική της συμπεριφορά.

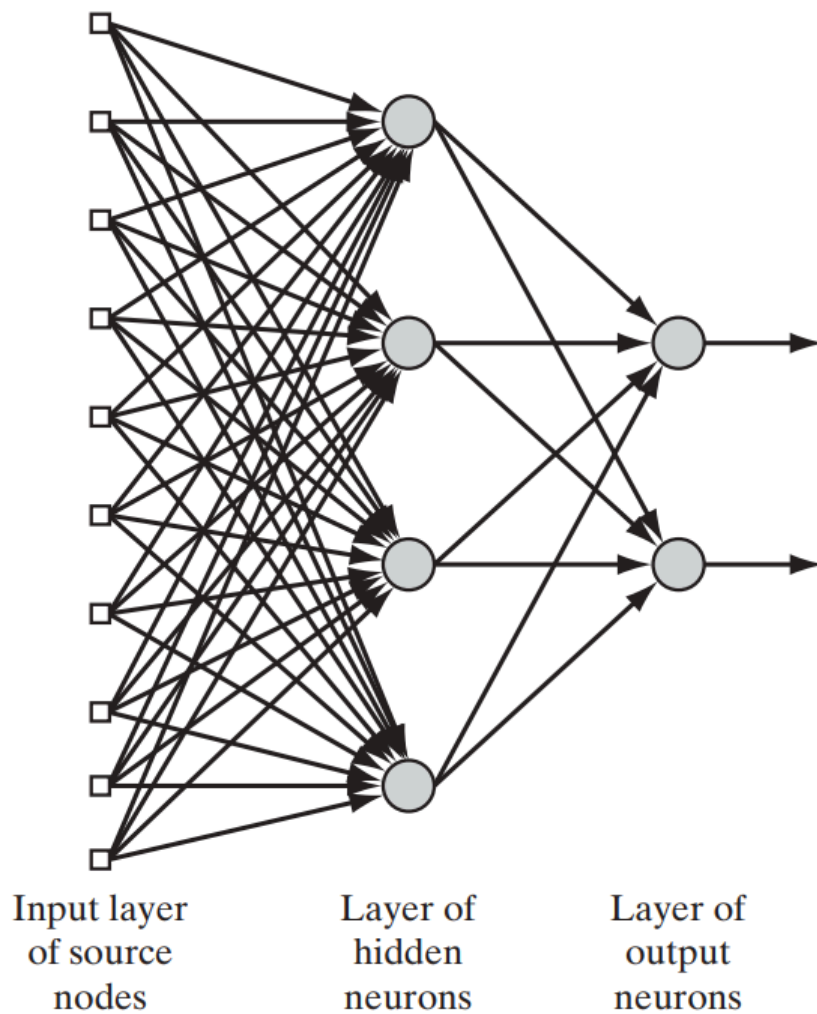
4.4 Αρχιτεκτονικές νευρωνικών δικτύων

Η ισχύς των νευρωνικών δικτύων προέρχεται από τον συνδυασμό πολλαπλών νευρώνων, και η διάταξη αυτών μπορεί να χωριστεί σε τρεις βασικές κατηγορίες. Πρώτη κατηγορία είναι τα δίκτυα πρόσθιας τροφοδότησης ενός επιπέδου. Με αυτόν τον όρο χαρακτηρίζουμε τα δίκτυα όπου οι νευρώνες είναι όλοι μέρος της εξόδου, και άρα σχηματίζουν ένα επίπεδο χωρίς κάποια σύνδεση των εξόδων τους με τις εισόδους κανενός. Γι αυτό τον λόγο ονομάζονται ενός επιπέδου.



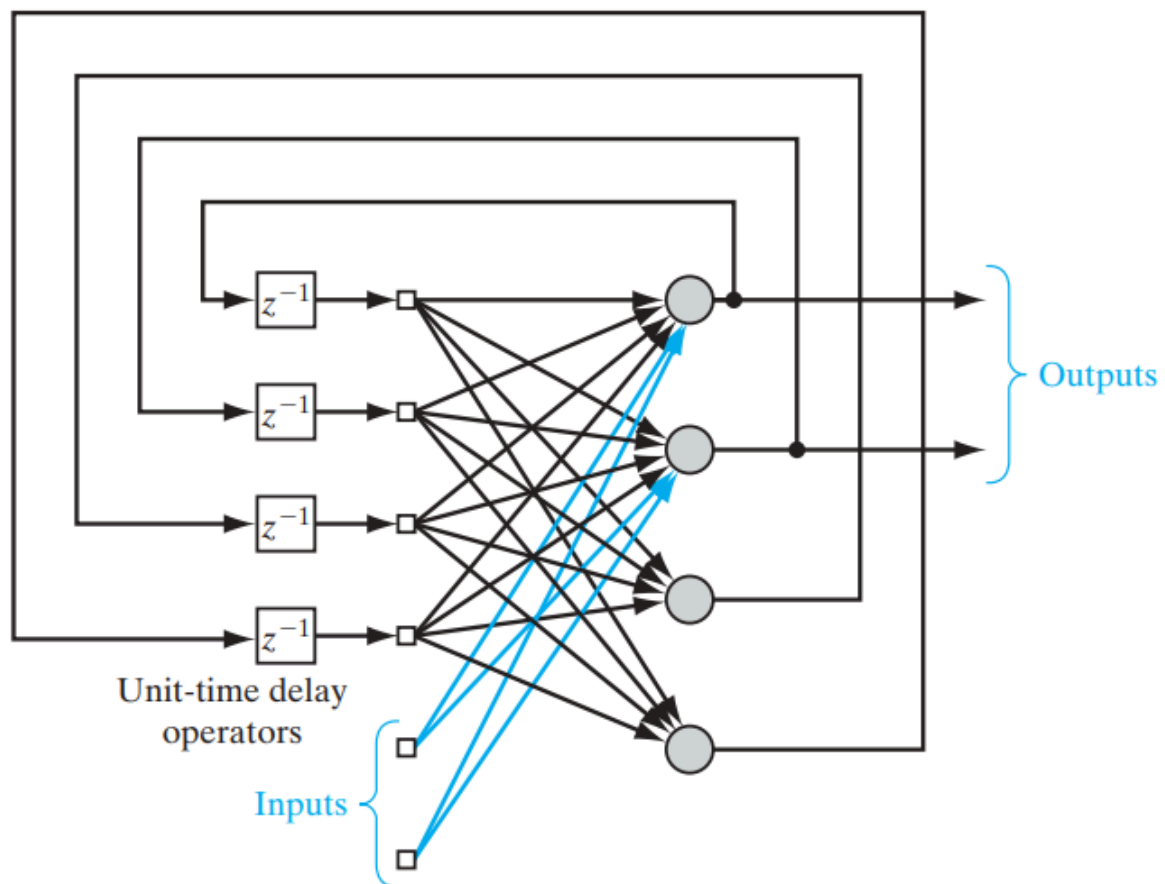
Σχήμα 6 Δίκτυο πρόσθιας τροφοδότησης ενός επιπέδου (Haykin, 2009, 21)

Η δεύτερη κατηγορία είναι τα πολυεπίπεδα δίκτυα πρόσθιας τροφοδότησης. Πρακτικά αυτά είναι η γενίκευση της προηγούμενης δομής και τα επίπεδα μεταξύ της εισόδου και της εξόδου ονομάζονται «κρυφά επίπεδα» (hidden layers). Η χρησιμότητά τους έγκειται στη δυνατότητα να παράξουν γνώση που συνδυάζει την γνώση που εξήχθη από τα προηγούμενα επίπεδα, και άρα να αποκτηθεί μια «γενικότερη» γνώση. Στο παρακάτω σχήμα 7, κάθε στοιχείο (είσοδος, νευρώνες) του δικτύου συνδέεται με όλα τα επόμενα και ως εκ τούτου ονομάζεται πλήρως συνδεδεμένο, ενώ όταν λείπει έστω μία ονομάζεται μερικώς συνδεδεμένο.



Σχήμα 7 Δίκτυο πρόσθιας τροφοδότησης δύο επιπέδων (Haykin, 2009, 22)

Τέλος, η τρίτη κατηγορία αφορά δίκτυα με ανάδραση, τα επονομαζόμενα αναδρομικά νευρωνικά δίκτυα, όπου έστω μία έξοδος νευρώνα πηγαίνει «προς τα πίσω» και συνδέεται με νευρώνα του ίδιου ή προηγούμενου επιπέδου. Η ανάδραση επιτρέπει σε αυτά τα δίκτυα να χρησιμοποιούν κάποιες ή όλες τις εξόδους τους ως εισόδους, επιτρέποντας έτσι δυναμικές προβλέψεις ιδιαίτερα όσον αφορά ακολουθίες και χωροχρονικά μοτίβα. Δύο μοντέλα που ανήκουν σε αυτήν την κατηγορία δικτύων και χρησιμοποιούνται ευρέως από την επιστημονική κοινότητα είναι τα LSTM και GRU (Fang et al., 2021, 98)



Σχήμα 8 Δίκτυο με βρόχους ανάδρασης μοναδιαίας καθυστέρησης (z^{-1})
(Haykin, 2009, 24)

4.5 Εκμάθηση νευρωνικών δικτύων

Μπορούμε να χωρίσουμε την εκμάθηση των νευρωνικών δικτύων σε δυο ευρείες κατηγορίες με βάση την διαδικασία με την οποία αποκτούν γνώση, την μάθηση με εκπαιδευτή και την μάθηση χωρίς εκπαιδευτή. Η μάθηση χωρίς εκπαιδευτή μπορεί να διαχωριστεί περαιτέρω σε μη-επιβλεπόμενη μάθηση και ενισχυτική μάθηση. Θα περιγράψουμε στις επόμενες παραγράφους κάθε μια από αυτές τις κατηγορίες.

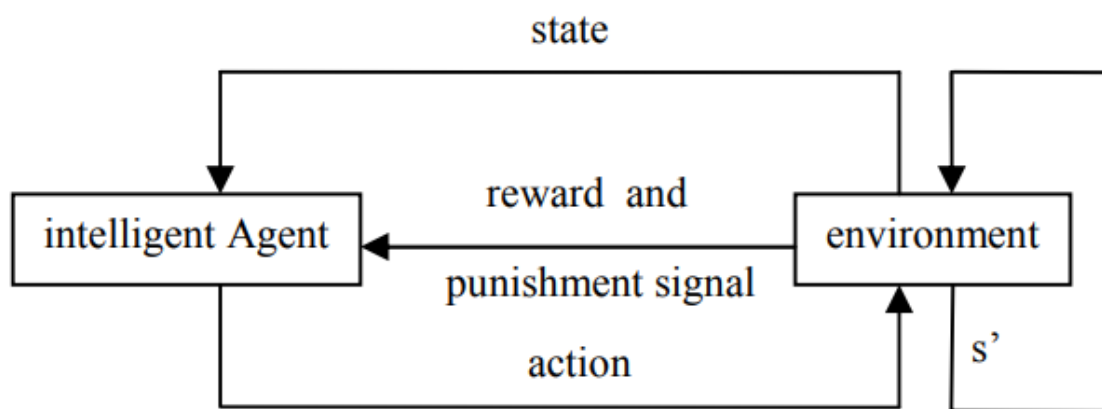
Μάθηση με εκπαιδευτή:

Γνωστή και ως επιβλεπόμενη μάθηση (Supervised Learning), περιγράφει την διαδικασία όπου ο εκπαιδευτής έχει ένα σύνολο δεδομένων για το οποίο γνωρίζει τις εισόδους, καθώς και τις επιθυμητές εξόδους. Το νευρωνικό αρχικά παράγει εξόδους μη γνωρίζοντας κάτι για το περιβάλλον ή τις επιθυμητές εξόδους, οπότε

αυτονόητα προκύπτουν σφάλματα μεταξύ της εξόδου του νευρωνικού και της γνώσης στον εκπαιδευτή επιθυμητής εξόδου. Η γνώση προσλαμβάνεται επαναληπτικά και σταδιακά με τα σφάλματα που προκύπτουν να είναι οδηγοί της εκμάθησης, καθώς το νευρωνικό δίκτυο αναπροσαρμόζει τα συναπτικά του βάρη με σκοπό να ελαχιστοποιήσει αυτό το σφάλμα. Η βέλτιστη εκπαίδευση δεν μπορεί να οριστεί αυστηρά, και κρίνεται με στατιστικές μεθόδους. Για παράδειγμα, ένα νευρωνικό δίκτυο απομαγνητοφώνησης μπορεί να μην αναγνωρίσει όλες τις ηχογραφήσεις τέλεια, αλλά αν είναι σε θέση να πετύχει μια αυθαίρετη μετρική, όπως θα ήταν ένα καλύτερο αποτέλεσμα από ανθρώπινες προσπάθειες, θα μπορούσε να χαρακτηριστεί επιτυχημένο. Όταν επιτευχθεί μια ικανοποιητική προσέγγιση της λύσης από το νευρωνικό, τότε μπορούμε να το θεωρήσουμε εκπαιδευμένο, με την γνώση να είναι αποθηκευμένη στα συναπτικά του βάρη. Πλέον, μπορούμε να απαλλαγούμε από τον εκπαιδευτή.

Ενισχυτική Μάθηση:

Η ενισχυτική μάθηση (Reinforcement Learning) αποτελεί ένα παράδειγμα μηχανικής μάθησης, όπου το νευρωνικό δίκτυο, που στην συγκεκριμένη περίπτωση ονομάζεται ως έξυπνος παράγοντας (intelligent agent), μαθαίνει να λαμβάνει αποφάσεις μέσω αλληλεπίδρασης με το περιβάλλον του, με σκοπό τη μεγιστοποίηση της συνολικής ανταμοιβής. Σε αντίθεση με την εποπτευόμενη μάθηση, που απαιτεί εκπαιδευτή (δηλαδή γνωστές εξόδους), ή τη μη εποπτευόμενη μάθηση, που όπως θα δούμε αποσκοπεί στην ανακάλυψη προτύπων, η ενισχυτική μάθηση χρησιμοποιεί μεθόδους δοκιμής και σφάλματος (trial-and-error) για να αναπτύξει βέλτιστες στρατηγικές λήψης αποφάσεων. Ένα παράδειγμα της ισχύς αυτής της μεθόδου είναι το πρόγραμμα AlphaGo Zero της Google Deepmind όπου εκπαιδεύτηκε με έναν αλγόριθμο ενισχυτικής μάθησης χωρίς πρότερες γνώσεις σε παιχνίδια όπως το σκάκι και το Go και μετά από 24 ώρες εκπαίδευσης σε ισχυρά υπολογιστικά συστήματα, είναι σε θέση να νικάει συστηματικά κορυφαίους παίκτες αλλά και άλλες μηχανές σκακιού. (Silver et al., 2017)

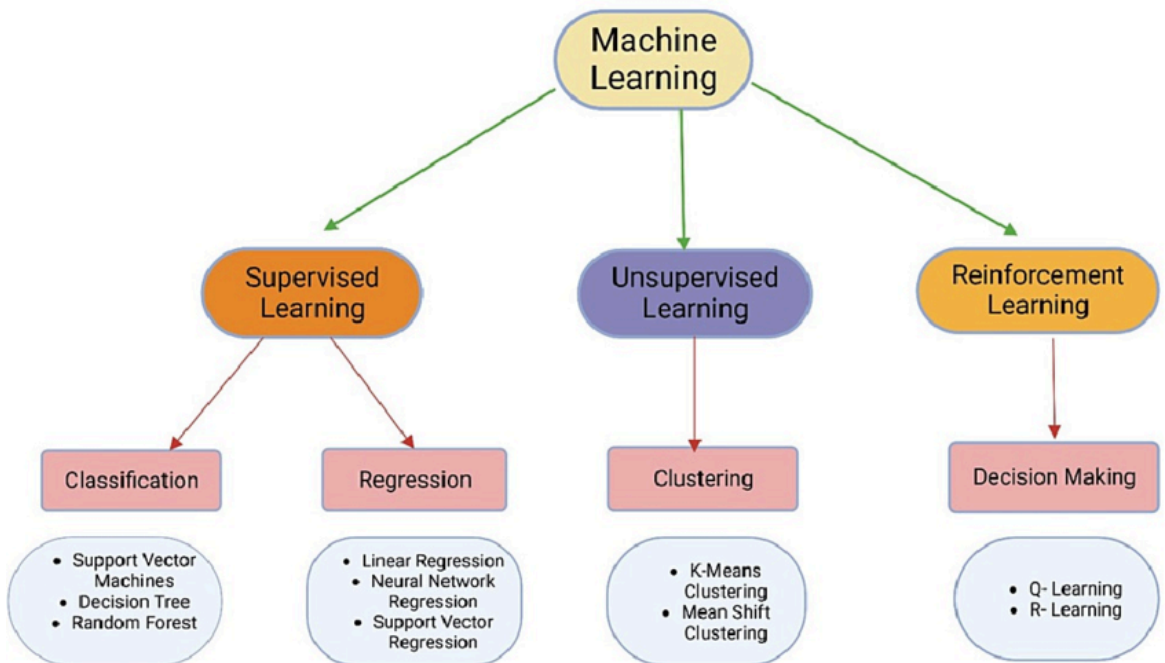


Σχήμα 9 Το βασικό μοντέλο Ενισχυτικής Μάθησης (Qiang & Zhongli, 2011, 1143)

Μη-επιβλεπόμενη Μάθηση:

Αυτού του είδους η εκμάθηση, γνωστή και ως αυτο-οργανούμενη (Unsupervised Learning / Self-organizing) λαμβάνει χώρα χωρίς εξωτερική εποπτεία ή γνωστές εξόδους. Σε αντίθεση με τις προηγούμενες κατηγορίες, δεν υπάρχει "δάσκαλος" ή "κριτής" για να καθοδηγήσει το σύστημα. Αντίθετα, το σύστημα βασίζεται σε έναν ανεξάρτητο από το έργο δείκτη ποιότητας για να βελτιστοποιήσει τις εσωτερικές αναπαραστάσεις των δεδομένων εισόδου. Αυτός ο δείκτης αφορά την ανίχνευση στατιστικών κανονικοτήτων, προτύπων ή δομών στα δεδομένα, επιτρέποντας στο σύστημα να ομαδοποιεί ή να κατηγοριοποιεί τα δεδομένα και να δημιουργεί αυτόνομα αναπαραστάσεις. (Haykin, 2009, 37)

Απλούστερα, η μάθηση χωρίς επίβλεψη μπορεί να θεωρηθεί ως εύρεση προτύπων σε δεδομένα πέρα από αυτό που θα θεωρείτο καθαρός μη-δομημένος θόρυβος. Δύο πολύ απλά κλασικά παραδείγματα μάθησης χωρίς επίβλεψη είναι η συσταδοποίηση, που αφορά το έργο της ομαδοποίησης ενός συνόλου αντικειμένων με τέτοιο τρόπο ώστε τα αντικείμενα που ανήκουν στην ίδια ομάδα να έχουν παρόμοιες ιδιότητες, καθώς και η μείωση διαστάσεων, δηλαδή η μέθοδος μείωσης των μεταβλητών σε ένα σύνολο δεδομένων εκπαίδευσης που χρησιμοποιείται για την ανάπτυξη μοντέλων μηχανικής μάθησης. (Bousquet et al., 2011, 72-112)



Σχήμα 10 Κατηγορίες μηχανικής μάθησης (Singh et al., 2023)

4.6 Αλγόριθμος Οπισθοδιάδοσης (Back-propagation)

Ο αλγόριθμος οπισθοδιάδοσης είναι ένας δημοφιλής και ευρέως χρησιμοποιούμενος αλγόριθμος στη μηχανική μάθηση, ο οποίος επιτρέπει τη βελτιστοποίηση των βαρών των τεχνητών νευρωνικών δικτύων, με στόχο να κάνει το μοντέλο να δίνει την ακριβέστερη δυνατή πρόβλεψη. Ο αλγόριθμος αυτός χρησιμοποιείται για τον αποδοτικό υπολογισμό των παραγώγων της συνάρτησης κόστους ως προς τα βάρη του δικτύου, διευκολύνοντας την προσαρμογή τους για τη βελτίωση της απόδοσης του μοντέλου. Η διαδικασία χρειάζεται δύο περάσματα, πρώτα προς τα «εμπρός» και μετά προς τα «πίσω».

Στο πρώτο πέρασμα προς τα εμπρός, δεδομένα εισόδου διέρχονται από το νευρωνικό δίκτυο, παράγοντας μια έξοδο. Η έξοδος αυτή συγκρίνεται με την πραγματική τιμή-στόχο, και το σφάλμα μεταξύ αυτών υπολογίζεται με τη χρήση μιας συνάρτησης κόστους (cost function).

Το δεύτερο πέρασμα ξεκινά από την έξοδο του νευρωνικού με κατεύθυνση προς τα πίσω επίπεδα του νευρωνικού, βρίσκοντας ανά επίπεδο του δικτύου μια κλίση για κάθε νευρώνα κάνοντας χρήση του σφάλματος και της παραγώγου της συνάρτησης με στόχο αυτή η κλίση να οδηγεί το νευρωνικό δίκτυο σε μια ακριβέστερη πρόβλεψη.

Ύστερα, ένας αλγόριθμος κατάβασης κλίσης (gradient descent) αποφασίζει με βάσει τις υπολογισμένες κλίσεις του αλγορίθμου οπισθοδιάδοσης σε πόσο μεγάλο βαθμό θα ενημερώσει τα βάρη κάθε νευρώνα. Αυτό γίνεται για πολλαπλές επαναλήψεις, καθώς η βελτίωση των βαρών του νευρωνικού δικτύου είναι σταδιακή και επηρεάζεται από τον ρυθμό μάθησης. Όταν όλα τα δεδομένα εκπαίδευσης εκτελεστούν μία φορά, αυτό ορίζεται ως εποχή (epoch) και χρειάζονται πολλαπλές εποχές για την ανανέωση των βαρών του νευρωνικού. Υπάρχουν πολλοί αλγόριθμοι κατάβασης κλίσης και η εύρεση αποδοτικότερων αλγορίθμων είναι ένα ενεργό πεδίο έρευνας.

Προτού εμφανιστούν περισσότεροι πολύπλοκοι αλγόριθμοι κατάβασης κλίσης όπως ο Adagrad (Duchi et al., 2011), οι δύο επιλογές των ερευνητών μηχανικής μάθησης ήταν η μαζική μάθηση (batch) και η άμεση (on-line) μάθηση. Η διαφορά τους έγκειται ότι στην μαζική μάθηση γίνονται ανανεώσεις στα βάρη του δικτύου μετά τον υπολογισμό των κλίσεων για όλες τις εισόδους που διαθέτουμε, ενώ η άμεση μάθηση ενημερώνει τα βάρη μετά από κάθε είσοδο. Η μαζική μάθηση έχει πλεονεκτήματα όπως την σύγκλιση της μεθόδου σε τοπικό ελάχιστο και την παραλληλοποίησή της, όμως η χρήση ολόκληρης της εισόδου την κάνει πρακτικά απαγορευτική για μεγάλα μεγέθη εισόδου (Haykin, 2009, 128). Αντίθετα, η άμεση μάθηση έχει δείξει εμπειρικά καλύτερη ταχύτητα σύγκλισης με συγκρίσιμη ακρίβεια. (Wilson & Martinez, 2003, 1429)

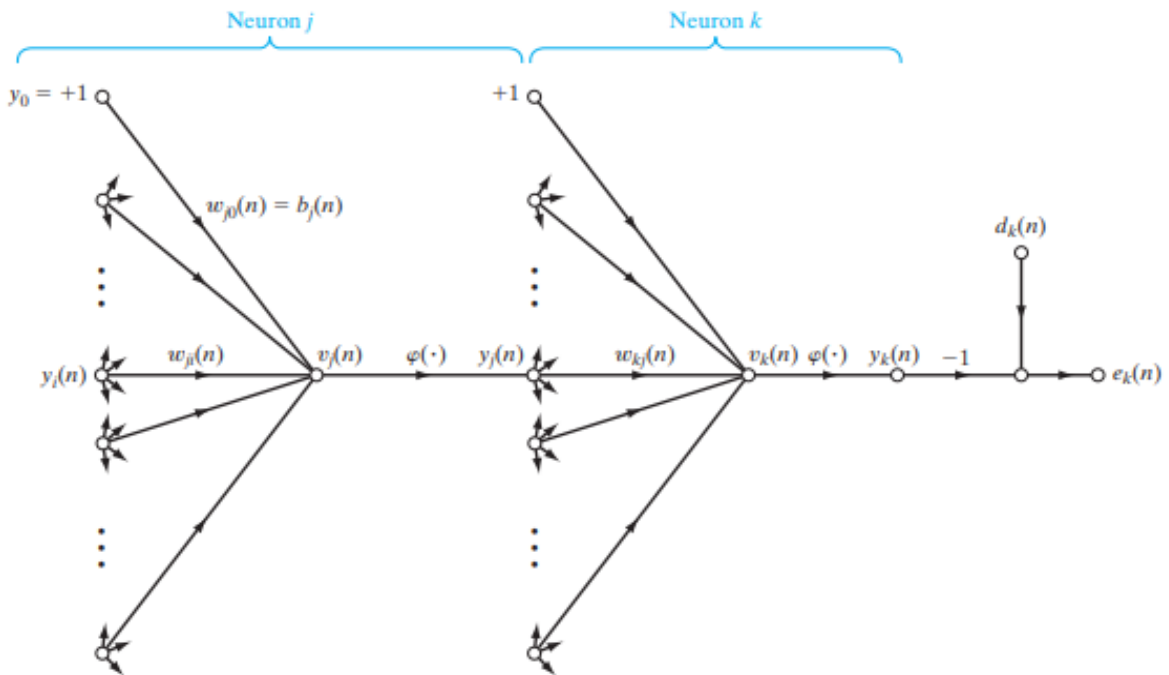
Θα εξετάσουμε τον αλγόριθμο οπισθοδιάδοσης αναλυτικότερα με βάση το παρακάτω σχήμα 11 για την περίπτωση της άμεσης κατάβασης. Έχουμε τα εξής:

Θέλουμε να ελαχιστοποιήσουμε την συνάρτηση κόστους E , όπου στην περίπτωση της άμεσης κατάβασης είναι η συνολική στιγμιαία ενέργεια σφάλματος

$$E(n) = \frac{1}{2} \sum_{\alpha \in C} e_{\alpha}^2(n) \text{ όπου το } \alpha \in C \text{ αναφέρεται σε κάποιον νευρώνα } \alpha \text{ του}$$

επιπέδου εξόδου, το σφάλμα $e_{\alpha}(n)$ ισούται με $d_{\alpha}(n) - y_{\alpha}(n)$ όπου d η επιθυμητή απόκριση, y η προβλεφθείσα απόκριση του νευρωνικού και το n αναφέρεται σε κάποιο συγκεκριμένο δείγμα εισόδου/εκπαίδευσης, καθώς στην περίπτωση άμεσης κατάβασης ενημερώνουμε τα βάρη με κάθε ξεχωριστό δείγμα.

Έστω νευρώνας εξόδου k ($k \in C$) και ένας νευρώνας j ($j \notin C$) του τελευταίου κρυφού επιπέδου.



Σχήμα 11 Γράφημα ροής σήματος των νευρώνων j,k (Haykin, 2009, 132)

Όπως είδαμε στην παράγραφο 4.2 για τον νευρώνα k ισχύει:

$$y_k = \varphi_k \left(b_k + \sum_{i=1}^m w_{ki} x_i \right) \text{ όπου } m \text{ ο συνολικός αριθμός εισόδων και } \varphi_k \text{ η συνάρτηση}$$

ενεργοποίησης του νευρώνα k . Ορίζουμε ότι οι είσοδοι x είναι οι έξοδοι y των νευρώνων προηγούμενου επιπέδου καθώς επίσης θα θεωρήσουμε την πόλωση b σαν είσοδο από έναν πλασματικό νευρώνα $z = 0$ προηγούμενου επιπέδου που δίνει πάντα έξοδο $x_0 = 1$ και το συναπτικό βάρος αυτού $w_{k0} = b_k$, τότε η συνάρτηση είναι ισοδύναμη και απλουστεύεται ως εξής:

$$y_k(n) = \varphi \left(\sum_{i=0}^m w_{ki}(n) x_i(n) \right). \text{ Θα ορίσουμε το τοπικό πεδίο } u \text{ του νευρώνα } k \text{ ως}$$

$$u_k(n) = \sum_{i=0}^m w_{ki}(n) x_i(n) \text{ και άρα έχουμε } y_k(n) = \varphi_k(u_k(n)).$$

Για να υπολογίσουμε την κλίση $\frac{\partial E(n)}{\partial w_{ki}(n)}$ για κάθε βάρος του νευρώνα k , χρησιμοποιούμε τον κανόνα αλυσίδας και έχουμε:

$$\frac{\partial E(n)}{\partial w_{ki}(n)} = \frac{\partial E(n)}{\partial e_k(n)} \frac{\partial e_k(n)}{\partial y_k(n)} \frac{\partial y_k(n)}{\partial u_k(n)} \frac{\partial u_k(n)}{\partial w_{ki}(n)}$$

Αναλύοντας τον κάθε όρο παίρνουμε:

$$\begin{aligned} \frac{\partial E(n)}{\partial e_k(n)} &= \frac{\partial (\frac{1}{2} \sum_{k' \in C} e_{k'}^2(n))}{\partial e_k(n)} = e_k(n) \\ \frac{\partial e_k(n)}{\partial y_k(n)} &= \frac{\partial (d_k(n) - y_k(n))}{\partial y_k(n)} = -1 \\ \frac{\partial y_k(n)}{\partial u_k(n)} &= \frac{\partial (\phi(u_k(n)))}{\partial u_k(n)} = \phi'(u_k(n)) \\ \frac{\partial u_k(n)}{\partial w_{ki}(n)} &= \frac{\partial (\sum_{i=0}^m w_{ki}(n) x_i(n))}{\partial w_{ki}(n)} = x_i(n) \end{aligned}$$

Από τα παραπάνω, η κλίση παίρνει την μορφή $\frac{\partial E(n)}{\partial w_{ki}(n)} = -e_k(n) \phi'(u_k(n)) x_i(n)$

Όλα τα παραπάνω είναι γνωστά για τον νευρώνα k που ανήκει στο επίπεδο εξόδου.

Ακόμα, εδώ παρατηρούμε την ανάγκη να έχουμε διαφορίσιμη συνάρτηση ενεργοποίησης ϕ ώστε να μπορούμε να χρησιμοποιήσουμε τον αλγόριθμο οπισθοδιάδοσης.

Δεδομένου ότι θέλουμε να ελαχιστοποιήσουμε την $E(n)$, θα χρησιμοποιήσουμε την κλίση αυτή για να αλλάξουμε τα βάρη $w_{ki}(n)$.

Η αλλαγή στα βάρη δίνεται από τον τύπο $\Delta w_{ki}(n) = -\mu \frac{\partial E(n)}{\partial w_{ki}(n)}$ όπου μ ο ρυθμός μάθησης του αλγορίθμου.

Στην περίπτωση του κρυφού νευρώνα j που βρίσκεται στο τελευταίο κρυφό επίπεδο, έχουμε το πρόβλημα ότι δεν γνωρίζουμε το $e_j(n)$ και αυτό διότι

$e_j(n) = d_j(n) - y_j(n)$ και δεν γνωρίζουμε την επιθυμητή απόκριση $d_j(n)$ παρά μόνο στην περίπτωση των νευρώνων εξόδου, οπότε θα εργαστούμε διαφορετικά για την εύρεση του $\frac{\partial E(n)}{\partial w_{ji}(n)}$. Με χρήση του κανόνα αλυσίδας έχουμε:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial u_j(n)} \frac{\partial u_j(n)}{\partial w_{ji}(n)}$$

Για τους δυο τελευταίους όρους, εργαζόμαστε όμοια με την προηγούμενη περίπτωση και παίρνουμε:

$$\frac{\partial y_j(n)}{\partial u_j(n)} = \phi'(u_j(n)) \text{ και } \frac{\partial u_j(n)}{\partial w_{ji}(n)} = x_i(n)$$

Για τον όρο $\frac{\partial E(n)}{\partial y_j(n)}$ παρατηρούμε ότι το y_j είναι η έξοδος του νευρώνα j αλλά και ταυτόχρονα μέρος της εισόδου του τελευταίου επιπέδου, και άρα από την εξίσωση $E(n) = \frac{1}{2} \sum_{\alpha \in C} e_a^2(n)$ για α νευρώνες του επιπέδου εξόδου, διαφορίζοντας ως προς

$$y_j \text{ έχουμε } \frac{\partial E(n)}{\partial y_j(n)} = \frac{\partial(\frac{1}{2} \sum_{\alpha \in C} e_a^2(n))}{\partial y_j(n)} = \sum_{\alpha \in C} \left[e_a(n) \frac{\partial(e_a(n))}{\partial y_j(n)} \right] \text{ και συνεχίζοντας με}$$

$$\text{κανόνα αλυσίδας } \frac{\partial E(n)}{\partial y_j(n)} = \sum_{\alpha \in C} \left[e_a(n) \frac{\partial(e_a(n))}{\partial u_\alpha(n)} \frac{\partial(u_\alpha(n))}{\partial y_j(n)} \right] \text{ όπου}$$

$$\frac{\partial(e_a(n))}{\partial u_\alpha(n)} = \frac{\partial(d_\alpha(n) - y_\alpha(n))}{\partial u_\alpha(n)} = \frac{\partial(-y_\alpha(n))}{\partial u_\alpha(n)} = \frac{\partial(-\varphi_\alpha(u_\alpha(n)))}{\partial u_\alpha(n)} = -\varphi'_\alpha(u_\alpha(n))$$

και

$$\frac{\partial(u_\alpha(n))}{\partial y_j(n)} = \frac{\partial(\sum_{i=0}^m w_{\alpha i}(n)x_i(n))}{\partial y_j(n)} \text{ όμως η κάθε είσοδος } x_i(n) \text{ των νευρώνων } \alpha \text{ του επιπέδου εξόδου, είναι ταυτόχρονα έξοδος } y_i(n) \text{ από νευρώνες του προτέρου επιπέδου, άρα}$$

$$\frac{\partial(u_\alpha(n))}{\partial y_j(n)} = \frac{\partial(\sum_{i=0}^m w_{\alpha i}(n)y_i(n))}{\partial y_j(n)} = w_{\alpha j}(n), \text{ το συναπτικό βάρος του νευρώνα } \alpha \text{ για την είσοδο } y_j \text{ που είναι ταυτόχρονα έξοδος του νευρώνα } j$$

Με τα παραπάνω, έχουμε:

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_{\alpha \in C} \left[e_\alpha(n) \frac{\partial(e_\alpha(n))}{\partial u_\alpha(n)} \frac{\partial(u_\alpha(n))}{\partial y_j(n)} \right] = - \sum_{\alpha \in C} \left[e_\alpha(n) \varphi'_\alpha(u_\alpha(n)) w_{\alpha j}(n) \right]$$

και συνολικά παίρνουμε:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial u_j(n)} \frac{\partial u_j(n)}{\partial w_{ji}(n)} = -\varphi'(u_j(n)) \sum_{\alpha \in C} \left[e_\alpha(n) \varphi'_\alpha(u_\alpha(n)) w_{\alpha j}(n) \right]$$

Με όμοιο τρόπο υπολογίζεται και το προτελευταίο κρυφό επίπεδο, έστω νευρώνας j' στο προτελευταίο επίπεδο:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial y_{j'}(n)} \frac{\partial y_{j'}(n)}{\partial u_{j'}(n)} \frac{\partial u_{j'}(n)}{\partial w_{ji}(n)}$$

Για τους δυο τελευταίους όρους, εργαζόμαστε όμοια με την προηγούμενη περίπτωση και παίρνουμε:

$$\frac{\partial y_{j'}(n)}{\partial u_{j'}(n)} = \varphi'(u_{j'}(n)) \text{ και } \frac{\partial u_{j'}(n)}{\partial w_{ji}(n)} = x_{j'}(n)$$

$$\begin{aligned} \frac{\partial E(n)}{\partial y_{j'}(n)} &= \sum_{\alpha \in \mathcal{C}} \left[e_{\alpha}(n) \frac{\partial(e_{\alpha}(n))}{\partial u_{\alpha}(n)} \frac{\partial(u_{\alpha}(n))}{\partial y_j(n)} \frac{\partial(y_j(n))}{\partial y_{j'}(n)} \right] = \\ &= - \sum_{\alpha \in \mathcal{C}} \left[e_{\alpha}(n) \varphi'_{\alpha}(u_{\alpha}(n)) w_{\alpha j}(n) \varphi'_j(u_j(n)) w_{jj'}(n) \right] = \\ &= - \varphi'_j(u_j(n)) w_{jj'}(n) \sum_{\alpha \in \mathcal{C}} \left[e_{\alpha}(n) \varphi'_{\alpha}(u_{\alpha}(n)) w_{\alpha j}(n) \right] \end{aligned}$$

$$\begin{aligned} \frac{\partial E(n)}{\partial w_{ji}(n)} &= \frac{\partial E(n)}{\partial y_{j'}(n)} \frac{\partial y_{j'}(n)}{\partial u_{j'}(n)} \frac{\partial u_{j'}(n)}{\partial w_{ji}(n)} = \\ &= - \varphi'_{j'}(u_{j'}(n)) \varphi'_j(u_j(n)) w_{jj'}(n) \sum_{\alpha \in \mathcal{C}} \left[e_{\alpha}(n) \varphi'_{\alpha}(u_{\alpha}(n)) w_{\alpha j}(n) \right] x_{j'}(n) \end{aligned}$$

Παρατηρούμε ότι μπορούμε να μειώσουμε την υπολογιστική πολυπλοκότητα αν ορίσουμε $\delta_{\lambda}(n) = \frac{\partial E(n)}{\partial u_{\lambda}(n)}$, όπου δ ορίζουμε την τοπική κλίση. Αν λ είναι

νευρώνας εξόδου, τότε $\delta_k(n) = - \frac{\partial E(n)}{\partial u_k(n)} = e_k(n) \varphi'_k(u_k(n))$, ενώ αν είναι

$$\begin{array}{ccccccc} \text{σε} & & \text{κρυφό} & & \text{επίπεδο} & & \text{τότε} \\ \delta_j(n) = - \frac{\partial E(n)}{\partial u_j(n)} & = - \frac{\partial E(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial u_j(n)} & = - \frac{\partial E(n)}{\partial y_j(n)} \varphi'_j(u_j(n)) & = - \frac{\partial E(n)}{\partial y_j(n)} \varphi'_j(u_j(n)) \end{array}$$

και άρα

$$\delta_j(n) = \sum_{\alpha \in \mathcal{C}} \left[e_{\alpha}(n) \varphi'_{\alpha}(u_{\alpha}(n)) w_{\alpha j}(n) \right] \varphi'_j(u_j(n)) = \varphi'_j(u_j(n)) \sum_{\alpha \in \mathcal{C}} \left[\delta_{\alpha}(n) w_{\alpha j}(n) \right]$$

Με τον παραπάνω τρόπο μπορούμε να υπολογίζουμε αναδρομικά το δ , επαναχρησιμοποιώντας βέλτιστα υπολογισμούς. (Haykin, 2009, 166)

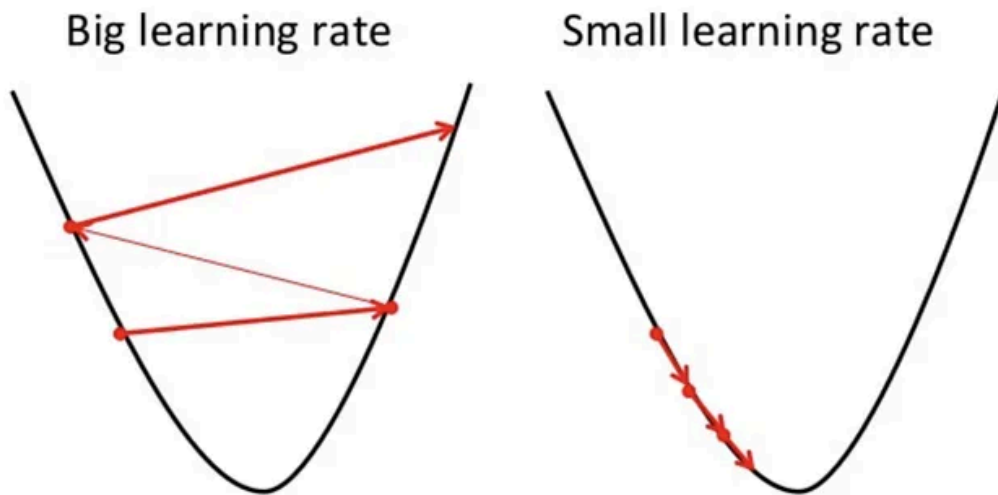
4.7 Ρυθμός μάθησης και υπερ-παράμετροι (Hyper-parameters)

Όπως είδαμε παραπάνω, η διόρθωση στα βάρη με τον αλγόριθμο οπισθοδιάδοσης δόθηκε από τον τύπο $\Delta w_{ki}(n) = - \mu \frac{\partial E(n)}{\partial w_{ki}(n)}$ όπου “ μ ” ο ρυθμός

μάθησης του αλγορίθμου. Το μ είναι μία υπερ-παράμετρος του νευρωνικού μοντέλου. Γενικά, τα μοντέλα μηχανικής μάθησης περιλαμβάνουν δύο διακριτές κατηγορίες παραμέτρων: τις παραμέτρους του μοντέλου, οι οποίες αρχικοποιούνται και ενημερώνονται κατά τη διαδικασία εκμάθησης (π.χ. τα βάρη των νευρώνων όπως είδαμε), και τις υπερ-παραμέτρους, οι οποίες δεν μπορούν να εκτιμηθούν απευθείας από τα δεδομένα και πρέπει να οριστούν εκ των προτέρων, καθώς καθορίζουν τη δομή του μοντέλου. Η ανάπτυξη ενός αποδοτικού μοντέλου μηχανικής μάθησης αποτελεί μία πολύπλοκη και χρονοβόρα διαδικασία που απαιτεί την επιλογή του κατάλληλου αλγορίθμου και τη βέλτιστη διαμόρφωση του μοντέλου μέσω της προσεκτικής ρύθμισης των υπερ-παραμέτρων (hyperparameter tuning). Οι υπερ-παράμετροι χρησιμοποιούνται είτε για τη ρύθμιση του μοντέλου ML (π.χ. η παράμετρος κανονικοποίησης σε έναν αλγόριθμο υποστηρικτικών διανυσμάτων ή ο ρυθμός μάθησης για την εκπαίδευση ενός νευρωνικού δικτύου) είτε για τον καθορισμό του αλγορίθμου που ελαχιστοποιεί τη συνάρτηση κόστους (π.χ. η επιλογή της συνάρτησης ενεργοποίησης και του βελτιστοποιητή σε ένα νευρωνικό δίκτυο ή ο τύπος του πυρήνα σε έναν αλγόριθμο υποστηρικτικών διανυσμάτων SVM). (Yang & Shami, 2020, 295)

Αναλυτικότερα, η επιλογή του ρυθμού μάθησης μ είναι μια άσκηση ισορροπίας. Ένας χαμηλός ρυθμός μάθησης πιθανόν να ανανεώνει απαγορευτικά αργά τα βάρη του δικτύου, ενώ ένας υψηλός ρυθμός εκμάθησης οδηγεί σε πιο επιθετική ενημέρωση των παραμέτρων αλλά μπορεί να οδηγήσει σε ταλάντωση όπως στο σχήμα 12, ή σε υπερπροσαρμογή, όπως θα εξηγηθεί παρακάτω.

Ο βέλτιστος ρυθμός μάθησης εξαρτάται από διάφορους παράγοντες, όπως το μέγεθος του νευρωνικού δικτύου, η πολυπλοκότητα του προβλήματος και ο όγκος των δεδομένων εκπαίδευσης. Συχνά απαιτείται πειραματισμός με διαφορετικές τιμές του ρυθμού μάθησης προκειμένου να προσδιοριστεί η πλέον κατάλληλη τιμή για το εκάστοτε μοντέλο. Ένας υπερβολικά υψηλός ρυθμός εκμάθησης μπορεί να οδηγήσει σε υπερπροσαρμογή του μοντέλου στα δεδομένα εκπαίδευσης, γεγονός που συνεπάγεται ότι το μοντέλο θα μάθει λεπτομέρειες των δεδομένων αυτών, αλλά δεν θα μπορέσει να γενικεύσει σε νέα δεδομένα. Από την άλλη πλευρά, ένας υπερβολικά χαμηλός ρυθμός εκμάθησης μπορεί να προκαλέσει υποπροσαρμογή, δηλαδή το μοντέλο δεν θα καταφέρει να μάθει τα υποκείμενα μοτίβα των δεδομένων εκπαίδευσης και δεν θα είναι σε θέση να πραγματοποιεί ακριβείς προβλέψεις. (Mishra, 2023) (Haykin, 2009, 137)



Σχήμα 12 Επιπτώσεις ενός «μεγάλου» ρυθμού μάθησης (Mishra, 2023)

4.8 Κανονικοποίηση των εισόδων

Ένας ακόμη τρόπος να επιταχύνουμε την εκμάθηση ενός νευρωνικού δικτύου είναι η κανονικοποίηση των εισόδων. Η σύγκλιση είναι συνήθως ταχύτερη όταν ο μέσος όρος κάθε εισόδου του συνόλου εκπαίδευσης είναι κοντά στο μηδέν. Για να γίνει κατανοητός ο λόγος, ας εξετάσουμε την ακραία περίπτωση όπου όλες οι εισοδοί είναι θετικές. Τα βάρη σε έναν συγκεκριμένο νευρώνα στο πρώτο επίπεδο ενημερώνονται κατά ένα ποσό ανάλογο με το σφάλμα στον νευρώνα αυτό και το διάνυσμα εισόδου. Όταν όλα τα συστατικά ενός διανύσματος εισόδου είναι θετικά, όλες οι ενημερώσεις των βαρών που τροφοδοτούν έναν κόμβο θα έχουν το ίδιο πρόσημο (δηλαδή, το πρόσημο του σφάλματος). Συνεπώς, αυτά τα βάρη θα μειώνονται ή θα αυξάνονται όλα ταυτόχρονα για κάθε δεδομένο μοτίβο εισόδου. Επομένως, αν ένα διάνυσμα βαρών πρέπει να αλλάξει κατεύθυνση, μπορεί να το κάνει μόνο με «ζιγκ-ζαγκ», κάτι που είναι αναποτελεσματικό και συνεπώς πολύ αργό. (LeCun et al., 2012, 16)

4.9 Γενίκευση, υπο-προσαρμογή, υπερ-προσαρμογή και cross-validation

Δεν υπάρχει ένας ντετερμινιστικός τρόπος για τον τερματισμό της οπισθοδιάδοσης όπως περιγράφηκε (Haykin, 2009, 139). Αντ' αυτού, η πιο συνήθης μέθοδος που χρησιμοποιείται για τον τερματισμό της διαδικασίας είναι το νευρωνικό να επιτυγχάνει επαρκή σύγκλιση και επαρκή γενίκευση. Επαρκής σύγκλιση εν προκειμένω σημαίνει ότι το νευρωνικό είναι σε θέση να δίνει τις αναμενόμενες τιμές στο σύνολο των δεδομένων εισόδου στο οποίο εκπαιδεύτηκε με επαρκώς

μικρό σφάλμα. Επαρκής γενίκευση σημαίνει ότι το νευρωνικό είναι σε θέση να δίνει τιμές σε δεδομένα εισόδου στα οποία δεν έχει εκπαιδευτεί με όσο το δυνατόν μικρότερο σφάλμα. Η ασάφεια του όρου επαρκής παραπάνω, υποδηλώνει ότι δεν έχει κάθε πρόβλημα τις ίδιες απαιτήσεις σύγκλισης και γενίκευσης, για παράδειγμα, αν ένα εκπαιδευμένο νευρωνικό δίκτυο αναγνώρισης ζώων σε εικόνες καταφέρει να αναγνωρίσει το 90% των εικόνων με ζώα, ενδεχομένως να θεωρήσουμε υπό μία έννοια ότι απέτυχε να φτάσει το 100%, αν όμως ένα τυχαίο δείγμα ανθρώπων αναγνωρίσει σωστά μόλις το 85%, τότε μπορούμε να πούμε ότι είναι επαρκώς εκπαιδευμένο ως προς την δυνατότητα να πετυχαίνει συγκρίσιμη ακρίβεια με έναν άνθρωπο.

Εδώ πρέπει να τονιστεί ότι παρότι η επαρκής σύγκλιση είναι προφανώς επιθυμητή, υπάρχει και η έννοια της υπερβολικής σύγκλισης. Πράγματι, ένα νευρωνικό με μόλις ένα κρυφό επίπεδο είναι σε θέση να προσεγγίσει οποιαδήποτε φραγμένη συνάρτηση (Hornik, 1991), αυτό όμως δεν είναι πάντοτε επιθυμητό γιατί μπορεί να οδηγήσει σε απώλεια της γενίκευσης όπως θα δούμε παρακάτω. Όταν έχουμε καλή σύγκλιση στα δεδομένα εισόδου αλλά χάνεται η δυνατότητα γενίκευσης, λέμε ότι το μοντέλο μας είναι υπερ-προσαρμοσμένο (overfit) στα δεδομένα της εισόδου. Όταν το μοντέλο δεν έχει την ικανότητα να μάθει από την είσοδο, τότε λέμε ότι είναι υπο-προσαρμοσμένο (underfit)- αυτό δύναται να συμβεί είτε όταν δεν αφήνουμε αρκετό χρόνο στον αλγόριθμο μάθησης του νευρωνικού ώστε για να μεταβάλλει τα βάρη των νευρώνων είτε όταν η πολυπλοκότητα του μοντέλου δεν είναι αντιπροσωπευτική του προβλήματος ή όταν τα δεδομένα εισόδου δεν είναι ικανά να προβλέπουν την έξοδο (Nelles, 2001, 161).



Σχήμα 13 Ποιοτικά παραδείγματα 3 διαφορετικών εκπαιδευμένων μοντέλων.

(Αριστερά) Υποπροσαρμογή, όπου το μοντέλο στερείται επαρκούς πολυπλοκότητας.

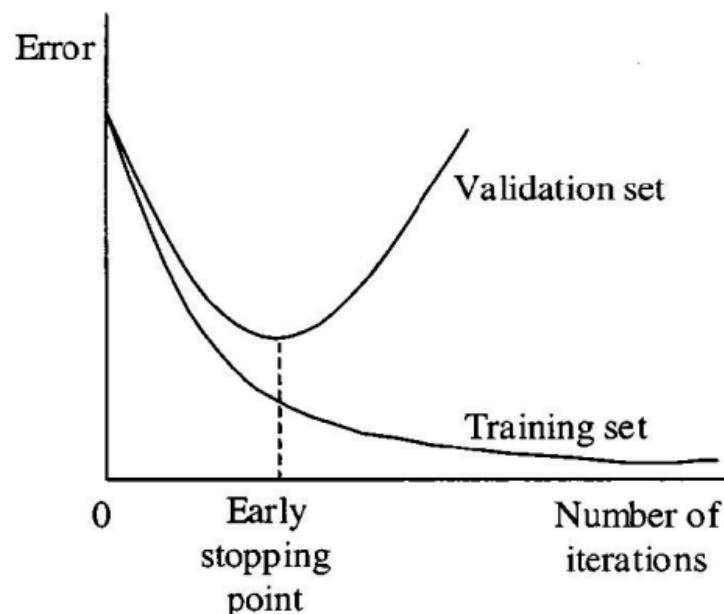
(Μέση) Καλή προσαρμογή, όπου το μοντέλο διαθέτει ισορροπημένη πολυπλοκότητα, επιτρέποντάς του να αποτυπώσει τα ουσιώδη χαρακτηριστικά των δεδομένων χωρίς υπερβολική γενίκευση.

(Δεξιά) Υπερπροσαρμογή, όπου το μοντέλο εμφανίζει υπερβολική πολυπλοκότητα, οδηγώντας στην αποτύπωση όχι μόνο των βασικών προτύπων των δεδομένων αλλά και του θορύβου και των ιδιομορφιών. (Oyedotun et al., 2017, 319)

Η υποπροσαρμογή είναι εύκολα ανιχνεύσιμη, καθώς ο ίδιος ο αλγόριθμος οπισθοδιάδοσης υπολογίζει το σφάλμα σε κάθε επανάληψη για την αλλαγή των βαρών όπως αναφέρθηκε, και αυτό το σφάλμα θα είναι μεγάλο. Η ανίχνευση υπερπροσαρμογής είναι λιγότερο προφανής και χρειάζεται μια μέθοδο ώστε να ελέγξουμε το σφάλμα του μοντέλου μας σε σχέση με δεδομένα που δεν έχει εκπαιδευτεί.

Μια τέτοια μέθοδος είναι η διασταυρωμένη επικύρωση (cross-validation). Η λογική είναι απλή, χωρίζουμε το σύνολο των δεδομένων εκπαίδευσης σε δύο υποσύνολα, όπου το πρώτο θα το ορίσουμε ως υποσύνολο εκτίμησης και το δεύτερο υποσύνολο επικύρωσης. Το σύνολο εκτίμησης είναι αυτό με το οποίο θα εκπαιδεύουμε το μοντέλο μας, ενώ το σύνολο επικύρωσης χρησιμοποιείται σε κάθε εποχή εκπαίδευσης για να δούμε την ικανότητα του μοντέλου να γενικεύει σε δεδομένα στα οποία δεν έχει ξαναδεί, δηλαδή να αποφεύγει την υπερπροσαρμογή.

Πέραν της ανίχνευσης της υπερπροσαρμογής, υπάρχουν διάφοροι τρόποι άμβλυνσής της. Ένας τρόπος είναι να τιμωρήσουμε τα μεγάλα βάρη, προσθέτοντας έναν όρο στην συνάρτηση σφάλματος που ελαχιστοποιεί ο αλγόριθμος οπισθοδιάδοσης, όρος που θα ορίζει το κόστος πολυπλοκότητας. Η τεχνική αυτή ονομάζεται ομαλοποίηση (regularization) (Moody, 1991, 849). Ένας δεύτερος τρόπος είναι ο τερματισμός του αλγορίθμου οπισθοδιάδοσης νωρίς (early stopping) (Sjöberg & Ljung, 1992, 73). Το νωρίς εν προκειμένω είναι το σημείο όπου το σφάλμα στο υποσύνολο επικύρωσης παύει να μειώνεται, όπως φαίνεται στο Σχήμα 14. Οι δύο παραπάνω μέθοδοι είναι αποτελεσματικές και ισοδύναμες ως προς την δυνατότητα γενίκευσης ενός νευρωνικού (Hagiwara, 2002, 937).



Σχήμα 14 Ποιοτική αναπαράσταση του σημείου γενίκευσης ενός μοντέλου (Oyedotun et al., 2017, 322)

4.10 Η κατάρα της Διαστατικότητας

Ο όρος αυτός αποδίδεται στον Richard Bellman (Bellman, 2013, ix) και περιγράφει πως όσο αυξάνονται οι μεταβλητές και άρα οι διαστάσεις ενός προβλήματος, η πολυπλοκότητά του αυξάνεται εκθετικά. Αυτό πρακτικά σημαίνει ότι η αύξηση των μεταβλητών εισόδου σε ένα νευρωνικό μοντέλο δεν το κάνει απαραίτητα ικανότερο να βρει κατάλληλη αντιστοιχία της εξόδου με την είσοδο. Αντιθέτως, αυξάνει κατά πολύ την ανάγκη περισσότερων δεδομένων εισόδου.

Ένας από τους βασικούς παράγοντες για την επιτυχή ανάπτυξη αλγορίθμων μάθησης είναι η ύπαρξη επαρκούς ποσότητας δεδομένων εκπαίδευσης, ώστε να καλύπτουν τον μαθηματικό χώρο ή το τμήμα του χώρου όπου το μοντέλο πρέπει να είναι έγκυρο. Είναι προφανές ότι, διατηρώντας σταθερούς όλους τους άλλους περιορισμούς, ο απαιτούμενος αριθμός δεδομένων εκπαίδευσης αυξάνεται εκθετικά με τη διάσταση του χώρου εισόδου. Για παράδειγμα, εάν 10 δεδομένα είναι επαρκή για να εκπαιδεύσουν ένα ομαλό μονοδιάστατο μοντέλο, απαιτούνται 100 δεδομένα για ένα διδιάστατο μοντέλο με την ίδια ομαλότητα, 1000 δεδομένα για ένα τριδιάστατο μοντέλο κ.ο.κ. Αυτή η εκθετική αύξηση αποτελεί την κύρια συνέπεια αυτού που ονομάζεται κατάρα της διαστατικότητας. (Verleysen & François, 2005, 761)

Κεφάλαιο 5. Προβλέψεις Ροών Φυσικού Αερίου

5.1 Τεχνολογίες και εργαλεία

Επιλέχθηκε η γλώσσα Python για τη συγγραφή του κώδικα, καθώς είναι ευρέως χρησιμοποιούμενη για εφαρμογές μηχανικής μάθησης λόγω της εκτεταμένης υποστήριξης σχετικών βιβλιοθηκών όπως Numpy, Pandas, Scikit-learn, TensorFlow και PyTorch καθώς και της ευκολίας χρήσης της. Χρησιμοποιήθηκαν Jupyter Notebooks τα οποία παρέχουν ένα διαδραστικό περιβάλλον που συνδυάζει την εκτέλεση κώδικα, την απεικόνιση και την τεκμηρίωση σε μια ενιαία πλατφόρμα. Το πρότζεκτ Jupyter είναι έργο ανοιχτού λογισμικού μη κερδοσκοπικού χαρακτήρα όπου καθένας μπορεί να συνεισφέρει στην ανάπτυξή του με την συναίνεση της κοινότητάς του (Jupyter, n.d.) Αυτό διευκολύνει την ερευνητική ανάλυση δεδομένων καθώς δίνει την δυνατότητα σε πραγματικό χρόνο να αλλάξουν κώδικας και δεδομένα και ταυτόχρονα επιτρέπει την παρουσίαση αποτελεσμάτων στο ίδιο αρχείο για την απρόσκοπτη παρουσίαση ευρημάτων. Επίσης έγινε χρήση του Google Colab, το οποίο είναι ένα εργαλείο που επεκτείνει τη λειτουργικότητα των Jupyter Notebooks, προσφέροντας υπολογιστικούς

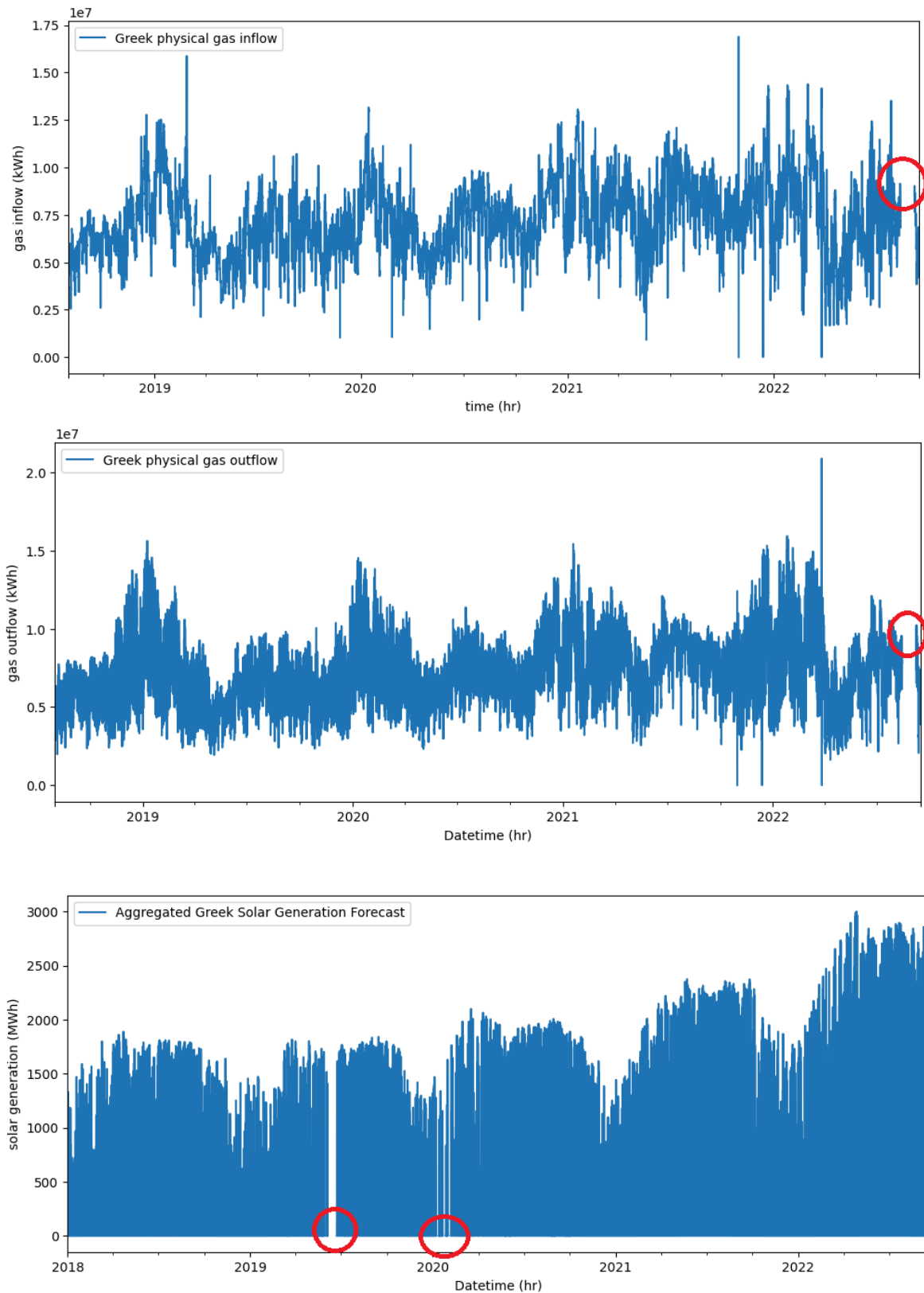
πόρους στο cloud, συμπεριλαμβανομένης της πρόσβασης σε GPUs και TPUs. Αυτό το καθιστά ιδιαίτερα χρήσιμο για την εκπαίδευση μεγάλων μοντέλων μηχανικής μάθησης χωρίς την ανάγκη επένδυσης σε υλικό. Το Google Colab είναι μέχρι και την στιγμή της συγγραφής αυτής της διπλωματικής εργασίας δωρεάν προς χρήση με κάποιους περιορισμούς σε πόρους, και δυνατότητα πληρωμής για περισσότερους πόρους (Google, n.d.). Το scikit-learn είναι μία ευρέως χρησιμοποιούμενη βιβλιοθήκη ανοιχτού κώδικα για μηχανική μάθηση που είναι γραμμένη σε Python. Παρέχει εργαλεία για ανάλυση δεδομένων, μοντελοποίηση και εξόρυξη γνώσης, τα οποία είναι κατάλληλα τόσο για ερευνητικούς όσο και για βιομηχανικούς σκοπούς. Η βιβλιοθήκη αναπτύχθηκε αρχικά ως τμήμα του έργου Google Summer of Code το 2007 και σταδιακά εξελίχθηκε από μια κοινότητα προγραμματιστών και ερευνητών. Βασίζεται σε βιβλιοθήκες όπως οι NumPy, Pandas, SciPy και matplotlib, καθιστώντας την ιδιαίτερα αποδοτική (scikit-learn, n.d.). Το TensorFlow που προαναφέρθηκε είναι ένα εργαλείο ανοιχτού κώδικα για την ανάπτυξη μοντέλων μηχανικής μάθησης που αναπτύχθηκε από την Google Brain, με έμφαση στις νευρωνικές αρχιτεκτονικές. Η ευελιξία του επιτρέπει στους ερευνητές να κατασκευάζουν, να εκπαιδεύουν και να βελτιστοποιούν μοντέλα, από απλά νευρωνικά δίκτυα (ANN) έως προηγμένα συστήματα όπως συνελκτικά (CNN) και επαναλαμβανόμενα δίκτυα (RNN) (TensorFlow, n.d.). Μαζί με το TensorFlow θα γίνει χρήση και του Keras, ενός API υψηλού επιπέδου ενσωματωμένο στο TensorFlow το οποίο απλοποιεί τη διαδικασία σχεδιασμού και εκπαίδευσης μοντέλων βαθιάς μάθησης. Παρέχει μια διεπαφή για την ταχεία πρωτοτυποποίηση μοντέλων, επιτρέποντας στους ερευνητές να επικεντρωθούν στον πειραματισμό και την καινοτομία (Keras, n.d.).

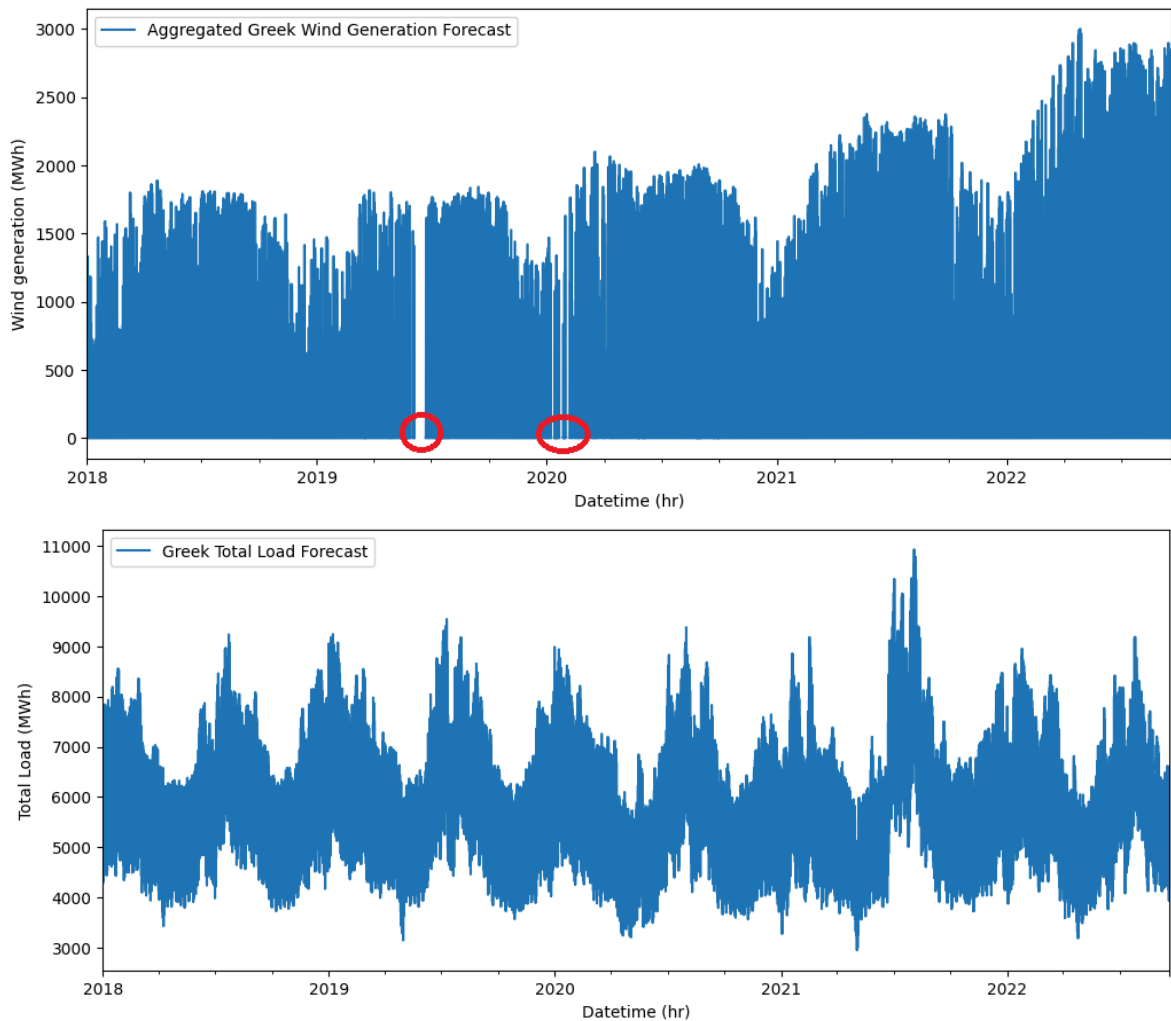
5.2 Συλλογή δεδομένων για εκπαίδευση

Όλα τα δεδομένα εκπαίδευσης στην ακατέργαστη μορφή τους είναι αρχεία csv (comma-separated values) που αφορούν την περίοδο 2018-2022 και οι πληροφορίες είναι δοσμένες σε ωριαία βάση. Τα δεδομένα αυτά αντλήθηκαν από την διεπαφή προγραμματισμού εφαρμογών (API) της πλατφόρμας DIEM και περιέχουν τις ροές φυσικού αερίου από και προς την Ελλάδα σε κιλοβατώρες, την συνολική πρόβλεψη παραγωγής επόμενης μέρας από Ανανεώσιμες Πηγές Ενέργειας (ηλιακή και αιολική) σε μεγαβατώρες, και την πρόβλεψη επόμενης μέρας της συνολικής ζήτησης ηλεκτρικής ενέργειας της Ελλάδος σε μεγαβατώρες.

5.3 Οπτικοποίηση και ανάλυση δεδομένων

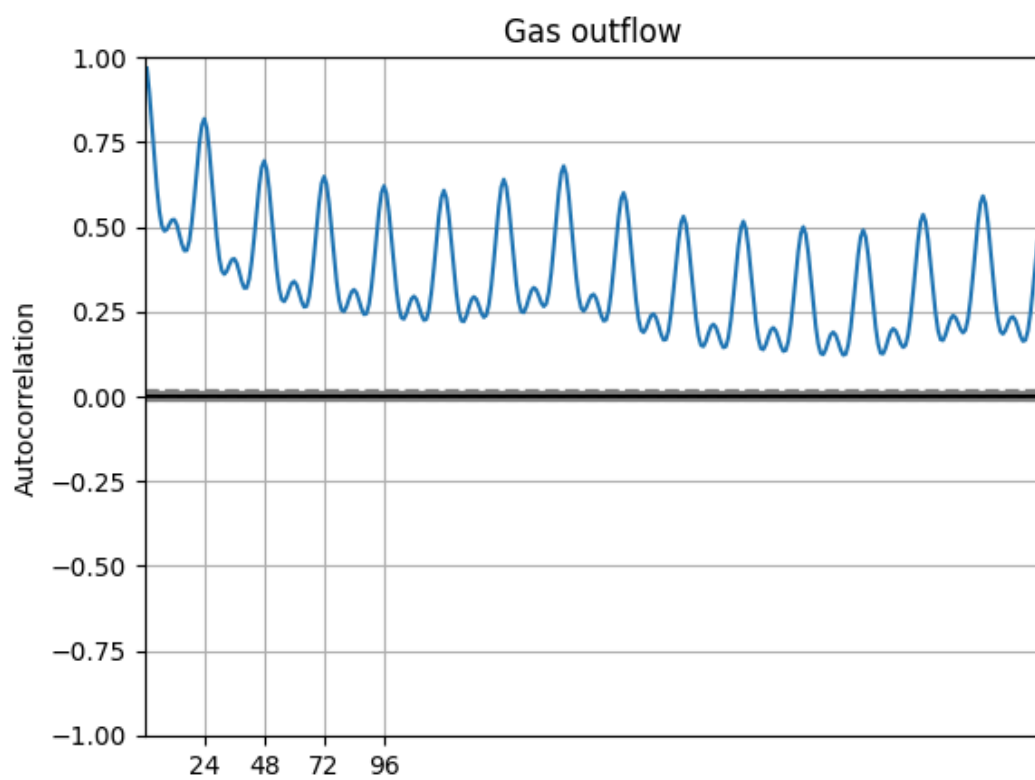
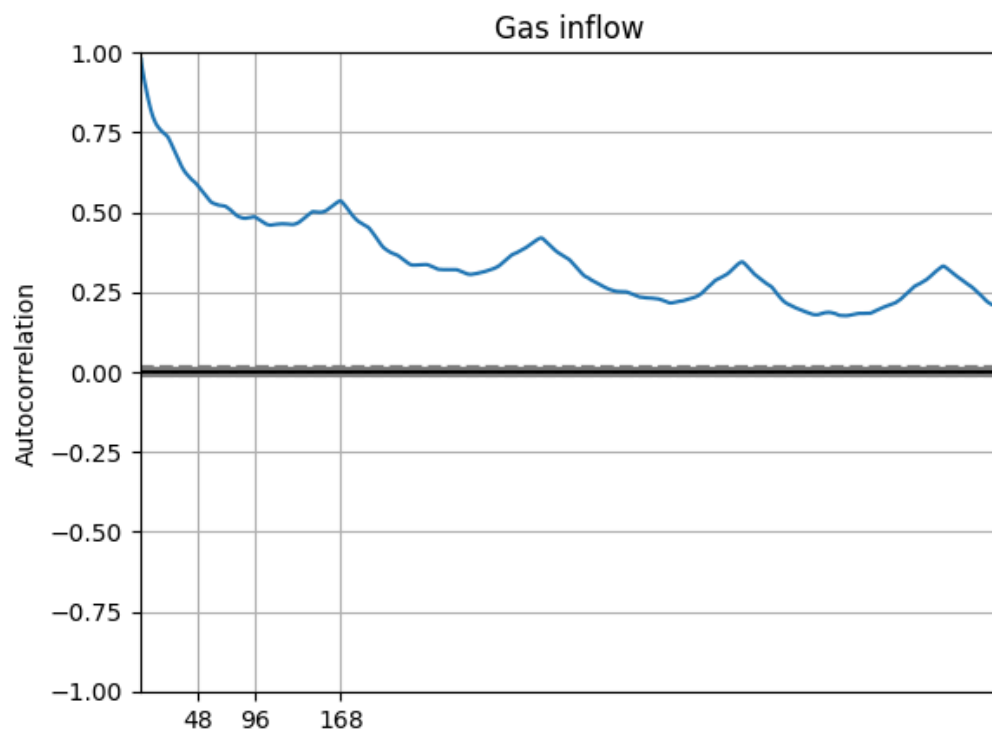
Με τη χρήση των παραπάνω εργαλείων, εισαγάγαμε και οπτικοποιήσαμε τα δεδομένα μας. Με μία γρήγορη ματιά παρατηρούμε τις ελλείψεις στα δεδομένα μας, επισημασμένα με κόκκινους κύκλους στο παρακάτω Σχήμα 15:

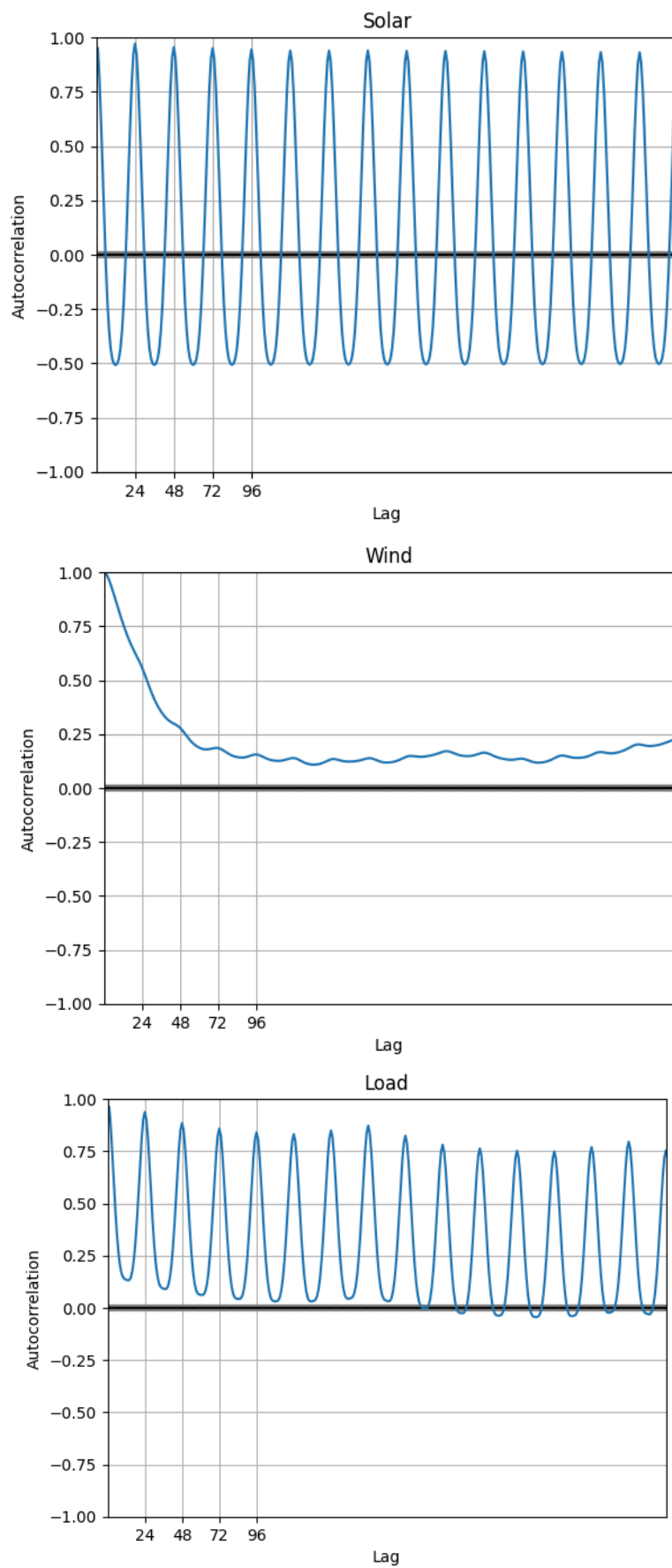




Σχήμα 15 Οπτικοποίηση των δεδομένων εκπαίδευσης
 Από πάνω προς τα κάτω: εισροές ΦΑ, εκροές ΦΑ, πρόβλεψη παραγωγής ηλιακής
 ενέργειας, πρόβλεψη παραγωγής ηλιακής ενέργειας, πρόβλεψη παραγωγής
 αιολικής ενέργειας, πρόβλεψη ζήτησης ηλεκτρικής ενέργειας

Στην ενότητα 5.4 θα επεξεργαστούμε τα κενά ώστε να μην συμπεριλάβουμε κενές τιμές στην εκπαίδευση του μοντέλου μας. Ύστερα θα υπολογίσουμε τις συναρτήσεις αυτοσυσχέτισης των παραπάνω δεδομένων-χρονοσειρών:





Σχήμα 16 Αυτοσυσχέτιση χρονοσειρών δεδομένων

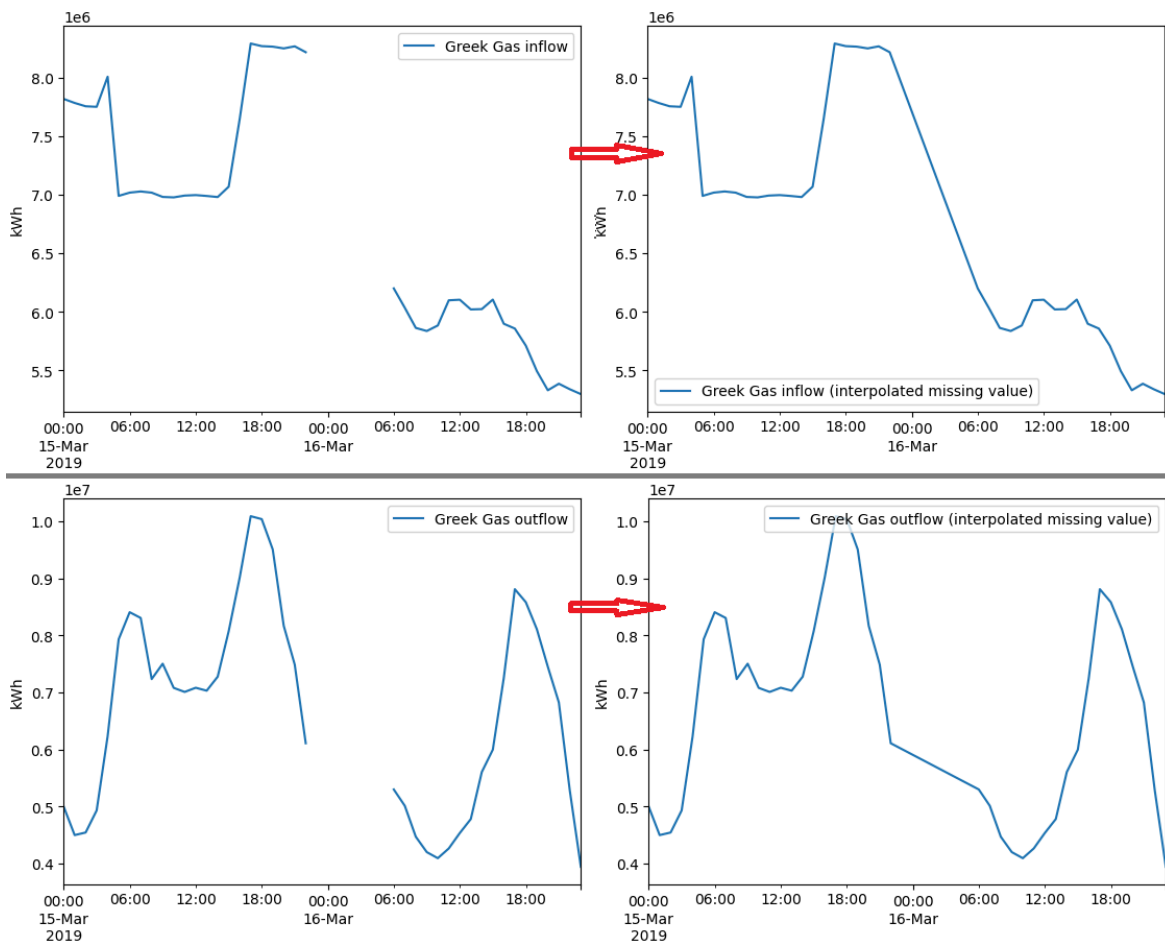
Από το σχήμα 16 βγάζουμε τα εξής συμπεράσματα, οι εισροές αερίου έχουν συσχέτιση αναμεταξύ τους σε κοντινές χρονικές αποστάσεις, δηλαδή αναμένουμε για ώρες x , $x\pm 1$, $x\pm 2$ παρόμοιες τιμές εισροών. Οι εκροές αερίου έχουν ισχυρή αυτοσυσχέτιση ανά 24ωρο και σχετικά ισχυρή ανά 48ωρο, υποδηλώνοντας μια περιοδικότητα στις εκροές αερίου. Αναμενόμενα η παραγωγή φωτοβολταϊκών έχει πολύ ισχυρή αυτοσυσχέτιση ανα 24, 48, 96 κ.ο.κ ώρες, καθώς είναι λογικό τα φωτοβολταϊκα να παράγουν (και αντίστοιχα να μην παράγουν) κατα μέσο όρο το ίδιο για ώρα x , $x+24$, $x+48$, $x+96$ ανά ημέρα αφού εξαρτάται από την ηλιοφάνεια εκείνων των ωρών/ημερών. Αντίθετα, η παραγωγή αιολικών έχει παραπλήσιες τιμές μόνο για βραχυπρόθεσμα παράθυρα των κοντινών ωρών, και χαμηλή αυτοσυσχέτιση γενικότερα. Η ζήτηση ηλεκτρικής ενέργειας έχει παρόμοια συμπεριφορά με την παραγωγή ηλιακής ενέργειας, με την υπόθεση ότι η ανθρώπινη δραστηριότητα και ανάγκη για ενέργεια έχει και αυτή συσχέτιση ανα 24ωρο.

5.4 Επεξεργασία δεδομένων

Με την χρήση κατάλληλης συνάρτησης, ο κώδικας της οποίας είναι αναρτημένος στο παράρτημα, βρίσκουμε ότι λείπουν τα εξής διαστήματα και από τις εισροές και από τις εκροές αερίου:

1. από τις 2019-03-15 23:00 η ώρα έως 2019-03-16 05:00 η ώρα,
2. το χρονικό σημείο 2021-10-31 00:00 και
3. από τις 2022-08-13 22:00 η ώρα έως 2022-09-07 03:00 η ώρα

Για το διάστημα τρία πάρθηκε η απόφαση να απαλειφθούν καθότι αφορά ελλιπή δεδομένα 24 ημερών και δεν κρίθηκε δόκιμη η εισαγωγή αυθαίρετων τιμών. Για το χρονικό σημείο δύο πάρθηκε η απόφαση να γίνει γραμμική παρεμβολή, καθότι αφορά μόνο ένα δείγμα του συνόλου. Όμοια έγινε και για το διάστημα ένα καθότι αφορά έξι χρονικές στιγμές, βλέπε Σχήμα 17.



Σχήμα 17 Οπτικοποίηση γραμμικής παρεμβολής για το διάστημα 1

Εργαζόμενοι με τον ίδιο τρόπο για τη χρονοσειρά της πρόβλεψης ηλιακής παραγωγής ενέργειας, βρίσκουμε 22 διαστήματα με ελλείψεις τιμών, όλα μεγαλύτερα ή ίσα των 23 ωρών. Παρατηρούμε ότι τα ίδια διαστήματα λείπουν και από την χρονοσειρά της πρόβλεψης αιολικής παραγωγής ενέργειας, όπου στην αιολική λείπει ένα ακόμα χρονικό σημείο (Σχήμα 18). Ως εκ τούτου, πάρθηκε η απόφαση να απαλειφθούν όλα τα παραπάνω ελλιπή δεδομένα. Όσον αφορά την πρόβλεψη ολικής ζήτησης ηλεκτρικής ενέργειας της επόμενης μέρας, δεν βρέθηκε καμία έλλειψη δεδομένων. Τέλος, περιορίσαμε το χρονικό εύρος μέχρι την 31 Ιανουαρίου 2022 για την αποφυγή της περιόδου έναρξης της ρωσικής εισβολής στην Ουκρανία όπου την ακολούθησαν αναταράξεις στην παροχή Φυσικού Αερίου.(Directorate-General for Energy, n.d.)

Ελλιπή δεδομένα πρόβλεψης παραγωγής ηλιακής ενέργειας:					Ελλιπή δεδομένα πρόβλεψης παραγωγής αιολικής ενέργειας:				
From		To			From		To		
0	2019-03-17 22:00:00	2019-03-18	21:00:00		0	2018-08-27 21:00:00	2018-08-28	20:00:00	
1	2019-04-30 21:00:00	2019-05-03	20:00:00		1	2019-03-17 22:00:00	2019-03-18	21:00:00	
2	2019-05-19 21:00:00	2019-05-21	20:00:00		2	2019-04-30 21:00:00	2019-05-03	20:00:00	
3	2019-05-24 21:00:00	2019-05-25	20:00:00		3	2019-05-19 21:00:00	2019-05-21	20:00:00	
4	2019-05-28 21:00:00	2019-06-01	20:00:00		4	2019-05-24 21:00:00	2019-05-25	20:00:00	
5	2019-06-04 21:00:00	2019-06-22	20:00:00		5	2019-05-28 21:00:00	2019-06-01	20:00:00	
6	2019-07-09 21:00:00	2019-07-11	20:00:00		6	2019-06-04 21:00:00	2019-06-22	20:00:00	
7	2019-07-15 21:00:00	2019-07-17	20:00:00		7	2019-07-09 21:00:00	2019-07-11	20:00:00	
8	2019-07-23 21:00:00	2019-07-25	20:00:00		8	2019-07-15 21:00:00	2019-07-17	20:00:00	
9	2019-07-29 21:00:00	2019-08-01	20:00:00		9	2019-07-23 21:00:00	2019-07-25	20:00:00	
10	2020-01-07 22:00:00	2020-01-12	21:00:00		10	2019-07-29 21:00:00	2019-08-01	20:00:00	
11	2020-01-13 22:00:00	2020-01-14	21:00:00		11	2020-01-07 22:00:00	2020-01-12	21:00:00	
12	2020-01-16 22:00:00	2020-01-18	21:00:00		12	2020-01-13 22:00:00	2020-01-14	21:00:00	
13	2020-01-20 22:00:00	2020-01-26	21:00:00		13	2020-01-16 22:00:00	2020-01-18	21:00:00	
14	2020-01-27 22:00:00	2020-01-28	21:00:00		14	2020-01-20 22:00:00	2020-01-26	21:00:00	
15	2020-01-30 22:00:00	2020-02-05	21:00:00		15	2020-01-27 22:00:00	2020-01-28	21:00:00	
16	2020-02-09 22:00:00	2020-02-13	21:00:00		16	2020-01-30 22:00:00	2020-02-05	21:00:00	
17	2020-05-25 21:00:00	2020-05-26	20:00:00		17	2020-02-09 22:00:00	2020-02-13	21:00:00	
18	2020-06-02 21:00:00	2020-06-03	20:00:00		18	2020-05-25 21:00:00	2020-05-26	20:00:00	
19	2020-08-26 21:00:00	2020-08-27	20:00:00		19	2020-06-02 21:00:00	2020-06-03	20:00:00	
20	2020-08-31 21:00:00	2020-09-01	20:00:00		20	2020-08-26 21:00:00	2020-08-27	20:00:00	
21	2020-09-23 21:00:00	2020-09-24	20:00:00		21	2020-08-31 21:00:00	2020-09-01	20:00:00	
22	2020-09-29 21:00:00	2020-09-30	20:00:00		22	2020-09-23 21:00:00	2020-09-24	20:00:00	
					23	2020-09-29 21:00:00	2020-09-30	20:00:00	

Σχήμα 18 Ομοιότητες ελλειπών δεδομένων σε διαφορετικές εισόδους

Ύστερα, θα προσθέσουμε στήλες με τις προηγούμενες 24 και 72 ώρες των δεδομένων πρόβλεψης, για να δοκιμάσουμε μοντέλα με περισσότερες εισόδους (features) στην εκπαίδευση του νευρωνικού μοντέλου μας. Ακόμα, μεταφέρουμε τις ροές φυσικού αερίου 24 ώρες στο παρελθόν, καθώς τα υπόλοιπα δεδομένα μας αφορούν πρόβλεψη επόμενης μέρας, και άρα οι ροές στην έξοδο του νευρωνικού θα αφορούν πρόβλεψη της επόμενης μέρας, πράγμα επιθυμητό. Τέλος, θα ενώσουμε τα δεδομένα σε ένα ενιαίο dataframe για κάθε μια από τις περιπτώσεις εισόδων που αναφέρθηκαν (3, 3*24, 3*72) και θα τα εξάγουμε σε αρχεία csv για περαιτέρω επεξεργασία (παραδείγματα csv στα Σχήματα 19,20).

	WindGenFo recast	SolarGenFo recast	TotalLoadF orecast	sum(physicalFlo w)0 inflow	sum(physical Flow)1 outflow
2018-08-02 04:00	470	158	5745	5332486	4084033
2018-08-02 05:00	520	478	6465	5425152	4611564
2018-08-02 06:00	550	871	7108	5411656	5204580
2018-08-02 07:00	570	1227	7494	5792693	5379787
2018-08-02 08:00	580	1473	7779	5735920	5706807
...	...	...	...	...	...

Σχήμα 19 Απόσπασμα αρχείου csv για την περίπτωση 3 εισόδων.

	WindG enFor ecast	WindG enFor ecast t-1	WindG enFor ecast t-2	...	TotalL oadFo recast t-22	TotalL oadFo recast t-23	sum(physi calFlow)0 inflow	sum(physic alFlow)1 outflow
2018-08-02 04:00	470	430	390	...	7170	6573	5332486	4084033
2018-08-02 05:00	520	470	430	...	7548	7170	5425152	4611564
2018-08-02 06:00	550	520	470	...	7798	7548	5411656	5204580
...	...	...	...	...	...	...	...	...

Σχήμα 20 Απόσπασμα αρχείου csv για την περίπτωση 3*24 εισόδων.

5.5 Μηχανική Χαρακτηριστικών (Feature Engineering)

Με τον όρο χαρακτηριστικά (features) εδώ εννοούμε τις εισόδους μας που δεν αφορούν τις επιθυμητές τιμές, συγκεκριμένα λοιπόν εννοούμε τις προβλέψεις (ηλιακή, αιολική, φορτίο) και όχι τις ροές φυσικού αερίου. Η μηχανική χαρακτηριστικών είναι χρήσιμη διαδικασία στην επιστήμη των δεδομένων, καθώς επηρεάζει σημαντικά την απόδοση των αλγορίθμων μηχανικής μάθησης. Περιλαμβάνει τη δημιουργία νέων χαρακτηριστικών ή τον μετασχηματισμό των υπαρχόντων με στόχο την βελτίωση της ικανότητας του μοντέλου να μαθαίνει πρότυπα και να κάνει ακριβείς προβλέψεις. Στα νευρωνικά δίκτυα συμβαίνει στα κρυφά επίπεδα να μαθαίνονται χαρακτηριστικά, μειώνοντας έτσι την ανάγκη ενός επιστήμονα δεδομένων να χρειαστεί να τα δημιουργήσει με το χέρι (Goodfellow et al., 2016, 5-6).

Επιλέξαμε να δημιουργήσουμε παρόλα αυτά το χαρακτηριστικό της ώρας που αντιπροσωπεύει κάθε είσοδο, δεδομένο που αλλιώς θα χανόταν καθώς ήταν μέρος του κλειδιού των δεδομένων μας. Κατόπιν, χωρίσαμε τα δεδομένα μας σε 3 κατηγορίες, κατηγορία εκπαίδευσης (training), κατηγορία επαλήθευσης (validation) και κατηγορία δοκιμής (test). Η πρώτη κατηγορία (train) είναι αυτή στην οποία θα εκπαιδευτεί το μοντέλο μας, η δεύτερη κατηγορία (validation) είναι αυτή η οποία θα ελέγχουμε κατά πόσο το μοντέλο μας είναι ικανό να γενικεύει και με βάση αυτό θα τροποποιούμε υπερ-παραμέτρους για καλύτερη γενίκευση/σύγκλιση και η τρίτη κατηγορία (test) μας επιτρέπει να συγκρίνουμε διαφορετικά μοντέλα, καθώς αυτή την κατηγορία δεν θα την χρησιμοποιήσουμε ούτε για εκπαίδευση, ούτε για ρύθμιση υπερ-παραμέτρων και άρα δεν θα έχει επηρεάσει την διαδικασία σχεδιασμού των μοντέλων. Δώσαμε 75% (~858 ημέρες), 15% (~172 ημέρες) και

10% (~114 ημέρες) της εισόδου σε κάθε κατηγορία 1,2 και 3 αντίστοιχα. Τέλος, κανονικοποιήσαμε όλα τα χαρακτηριστικά μας για να έχουμε ταχύτερη δυνατή σύγκλιση του αλγορίθμου μας (LeCun et al., 2012, 30) προσέχοντας να χρησιμοποιήσουμε μόνο τα δεδομένα της πρώτης κατηγορίας για υπολογισμό μέσου όρου/διακύμανσης, διαφορετικά θα εισαγάγαμε πληροφορίες από τις άλλες δύο κατηγορίες κατά την εκπαίδευση (data leakage). Η διαδικασία αναφέρεται συχνά στα Αγγλικά ως feature scaling.

5.6 Εκπαίδευση Μοντέλων

Για τα τρία αρχεία csv που δημιουργήσαμε με 3, 3*24 και 3*72 εισόδους, εφεξής all.csv, all_24.csv και all_72.csv αντίστοιχα, χρησιμοποιήθηκαν 3 Jupyter Notebooks (εφεξής notebooks) με πανομοιότυπο κώδικα ώστε να συγκρίνουμε τις ικανότητες πρόβλεψης ανάλογα με το πλήθος των χαρακτηριστικών στην διάθεση του μοντέλου, επιτρέποντάς μας να τα τρέξουμε παράλληλα με πόρους χωρίς χρέωση στο Google Colab. Για κάθε notebook επιλέξαμε 5 αρχιτεκτονικές πλήρως συνδεδεμένων νευρωνικών δικτύων:

1. 2 κρυφών επιπέδων με 16 νευρώνες σε κάθε επίπεδο (16-16)
2. 2 κρυφών επιπέδων με 32 νευρώνες στο πρώτο και 16 νευρώνες στο δεύτερο επίπεδο (32-16)
3. 2 κρυφών επιπέδων με 32 νευρώνες σε κάθε επίπεδο (32-32)
4. 3 κρυφών επιπέδων με 32 νευρώνες στο πρώτο, 32 νευρώνες στο δεύτερο και 16 νευρώνες στο τρίτο επίπεδο (32-32-16)
5. 3 κρυφών επιπέδων με 64 νευρώνες στο πρώτο, 32 νευρώνες στο δεύτερο και 16 νευρώνες στο τρίτο επίπεδο 64-32-16)

Όλα είχαν δύο νευρώνες στο επίπεδο εξόδου (λόγω πρόβλεψης εισροής και εκροής) με γραμμική συνάρτηση ενεργοποίησης. Για τους νευρώνες κρυφού επιπέδου χρησιμοποιήθηκαν συναρτήσεις ενεργοποίησης ReLU και καθώς δεν υπάρχει ομοφωνία για το ποιός είναι ο βέλτιστος αλγόριθμος διόρθωσης των βαρών στην οπισθοδιάδοση (Goodfellow et al., 2016, 306) επιλέχθηκε ο αλγόριθμος Adam (Kingma & Ba, 2014, 1). Τα μοντέλα έτρεξαν για 3000 εποχές και η εκπαίδευση διήρκεσε μεταξύ μιας και δύο ωρών για κάθε μοντέλο. Επίσης έγινε χρήση της τεχνικής early stop για να λάβουμε τα καλύτερα δυνατά βάρη.

Η διαδικασία που ακολουθήθηκε στα Notebooks μετά την εκπαίδευση κάθε μοντέλου έχει ως εξής:

- Επισήμανση του μέσου τετραγωνικού σφάλματος των σετ (κατηγορία) εκπαίδευσης και επαλήθευσης στο τέλος της εκπαίδευσης.
- Γραφική παράσταση των παραπάνω για κάθε εποχή.
- Ποιοτική αναπαράσταση των προβλέψεων για ένα μικρό εύρος των σετ εκπαίδευσης και επαλήθευσης.

Τέλος, θα διαλέξουμε το αποδοτικότερο μοντέλο και θα το συγκρίνουμε ποιοτικά στις προβλέψεις του με την γραμμική παλινδρόμηση.

5.7 Αποτελέσματα

all.csv (3 είσοδοι)		all_24.csv (3*24 είσοδοι)		all_72.csv (3*72 είσοδοι)	
MSE	CV MSE	MSE	CV MSE	MSE	CV MSE
Μοντέλο 1: 16-16					
$\sim 25 * 10^{11}$	$\sim 30 * 10^{11}$	$\sim 15 * 10^{11}$	$\sim 29 * 10^{11}$	$\sim 13 * 10^{11}$	$\sim 21 * 10^{11}$
Μοντέλο 2: 32-16					
$\sim 25 * 10^{11}$	$\sim 30 * 10^{11}$	$\sim 14 * 10^{11}$	$\sim 26 * 10^{11}$	$\sim 12 * 10^{11}$	$\sim 18 * 10^{11}$
Μοντέλο 3: 32-32					
$\sim 26 * 10^{11}$	$\sim 30 * 10^{11}$	$\sim 24 * 10^{11}$	$\sim 32 * 10^{11}$	$\sim 11 * 10^{11}$	$\sim 20 * 10^{11}$
Μοντέλο 4: 32-32-16					
$\sim 26 * 10^{11}$	$\sim 30 * 10^{11}$	$\sim 15 * 10^{11}$	$\sim 26 * 10^{11}$	$\sim 12 * 10^{11}$	$\sim 18 * 10^{11}$
Μοντέλο 5: 64-32-16					
$\sim 27 * 10^{11}$	$\sim 30 * 10^{11}$	$\sim 14 * 10^{11}$	$\sim 26 * 10^{11}$	$\sim 12 * 10^{11}$	$\sim 18 * 10^{11}$

Σχήμα 21 Αποτελέσματα Εκπαίδευσης Μοντέλων

Από το παραπάνω σχήμα 21 έχουμε για κάθε αρχείο csv και κάθε μοντέλο τα αποτελέσματα της εκπαίδευσής του. Τονίζεται ότι το MSE (mean squared error) είναι το μέσο τετραγωνικό σφάλμα εκπαίδευσης, και το CV (cross-validation) MSE είναι το μέσο τετραγωνικό σφάλμα επαλήθευσης. Παρατηρούμε τα εξής:

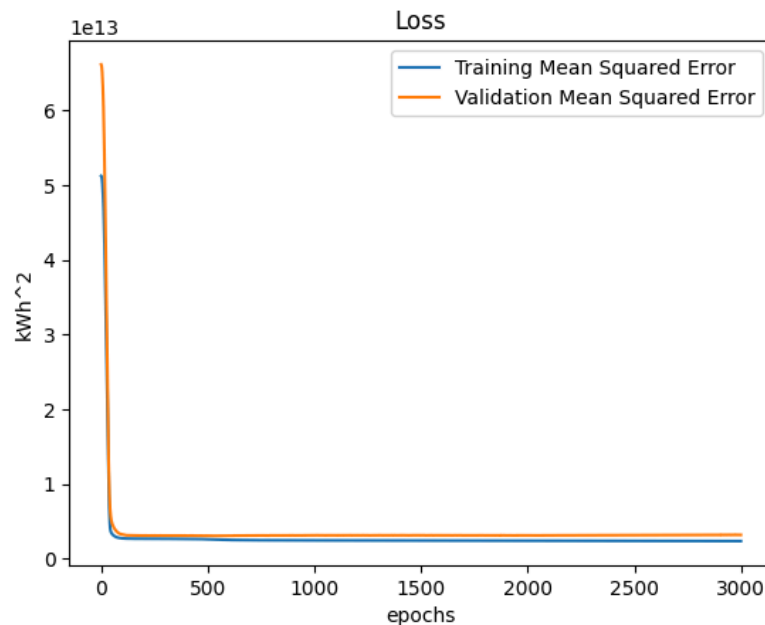
- Τα μοντέλα έχουν συναφή αρχιτεκτονική και άρα κοντινά αποτελέσματα μεταξύ τους.
- Η περίπτωση 3 εισόδων (all.csv) έχει τις χειρότερες επιδόσεις σε όλα τα μοντέλα και για όλες τις μετρικές σφάλματος, εκτός του σφάλματος CV MSE του μοντέλου 3 για την περίπτωση 3*24 εισόδων (all_24.csv)
- Όσο αυξάνονται οι είσοδοι, τόσο χαμηλότερο είναι το MSE για κάθε μοντέλο. Για τις 3*72 εισόδους (all_72.csv) παρατηρούμε ότι τόσο το σφάλμα MSE όσο και το CV MSE είναι χαμηλότερα σε σχέση με τις 3*24 εισόδους (all_24.csv).

Με βάσει τις παραπάνω παρατηρήσεις, θα απορρίψουμε την περίπτωση 1 εισόδου (all.csv) λόγω κακών επιδόσεων σύγκλισης ανεξαρτήτως μοντέλου. Δεν

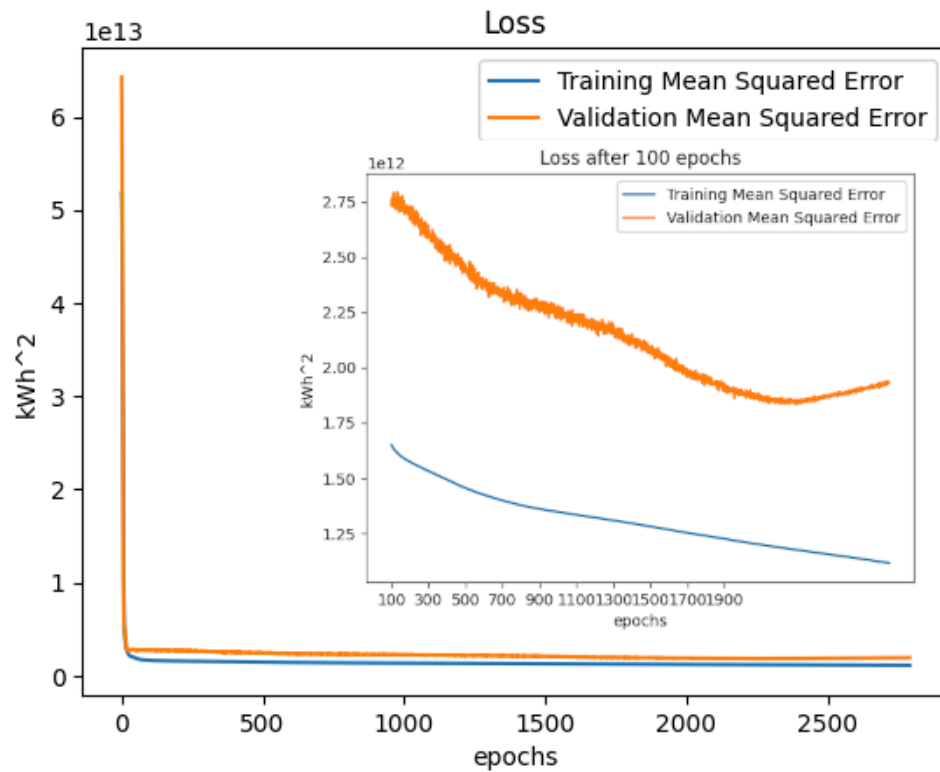
μπορεί να αποκλειστεί η περίπτωση που τα μοντέλα θα συνέκλιναν μετά από πολλές εποχές, πράγμα επιθυμητό, αλλά κάτι τέτοιο φαίνεται να είναι υπολογιστικά κοστοβόρο και δεν θα εξεταστεί στην παρούσα διπλωματική. Θα απορρίψουμε και την περίπτωση 3*24 εισόδων διότι παρότι η εικόνα είναι καλύτερη, η περίπτωση των 3*72 εισόδων είναι καλύτερη σε κάθε μετρική.

Ακόμα θα απορρίψουμε το μοντέλο 1 λόγω της αδυναμίας του να συγκριθεί με τα υπόλοιπα μοντέλα, πιθανώς λόγω απλότητας. Ακόμα θα απορρίψουμε το μοντέλο 5 διότι παρότι έχει περισσότερους νευρώνες, οι επιδόσεις είναι παρεμφερείς με το μοντέλο 4.

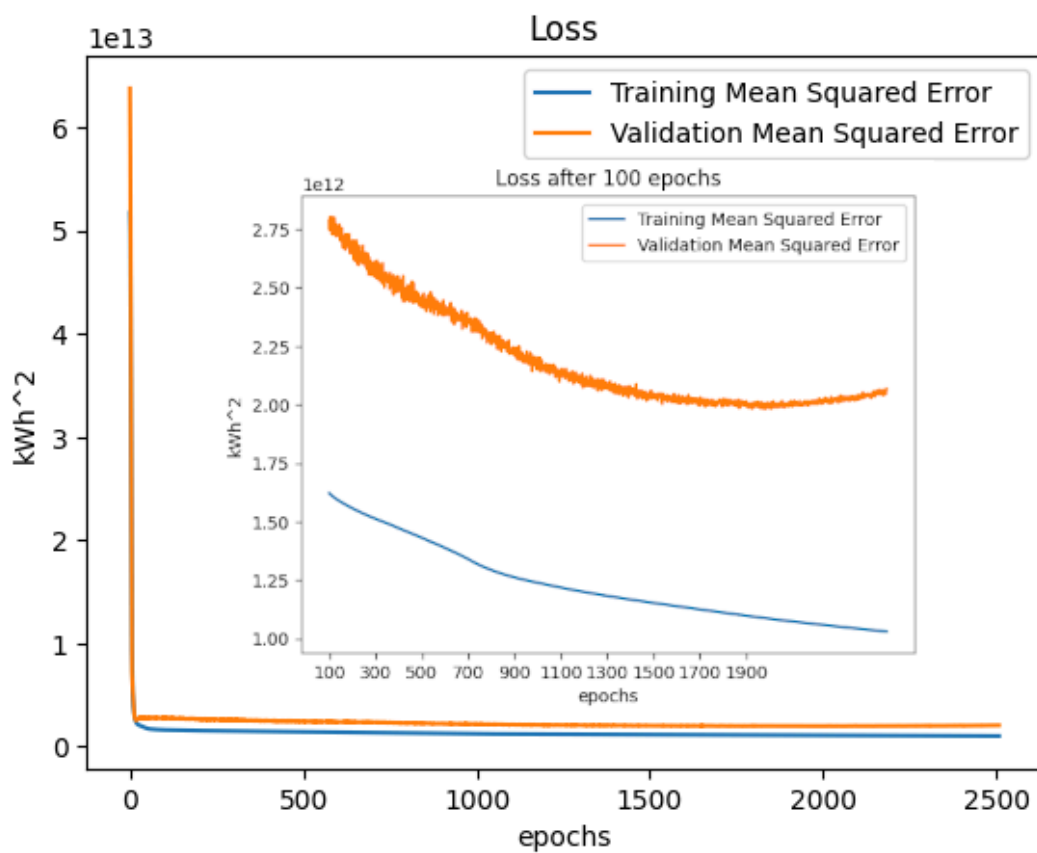
Στα σχήματα 22, 23 και 24 παρουσιάζονται οι γραφικές παραστάσεις εκπαίδευσης των μοντέλων και αρχείων που απέμειναν, καθώς και του μοντέλου 1 για την περίπτωση 3 εισόδων για την παρουσίαση ενός «κακού» παραδείγματος. Στους υπερσυνδέσμους του παραρτήματος είναι αναρτημένα τα Notebooks που περιέχουν γραφικές παραστάσεις για όλες τις περιπτώσεις που παρουσιάστηκαν.



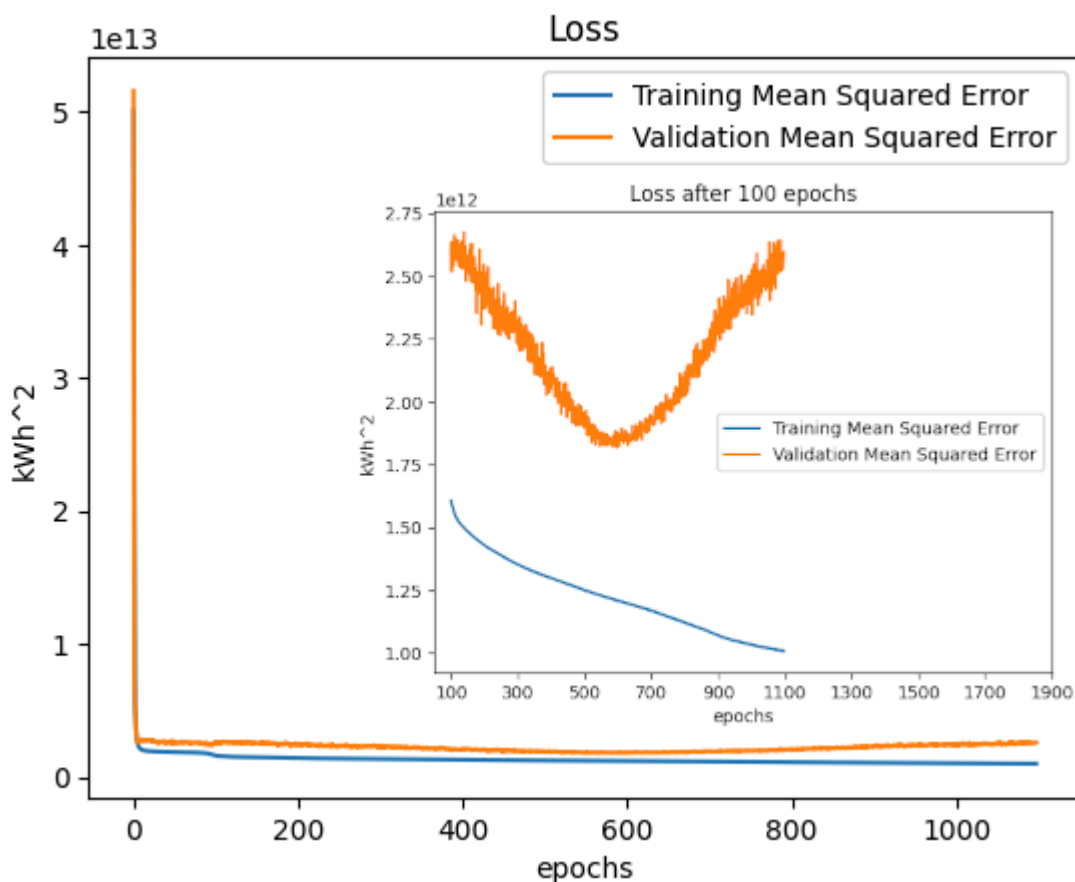
Σχήμα 22 Γραφική παράσταση σφάλματος. Μοντέλο 1, περίπτωση 3 εισόδων



Σχήμα 23.α) Γραφική παράσταση σφάλματος. Μοντέλο 2, 3*72 εισόδων.



Σχήμα 23.β) Γραφική παράσταση σφάλματος. Μοντέλο 3, 3*72 εισόδων.



Σχήμα 23.γ) Γραφική παράσταση σφάλματος. Μοντέλο 4, 3*72 εισόδων.

Στο σχήμα 22 παρατηρούμε την προαναφερθείσα δυσκολία του αρχείου των 3 εισόδων να εκπαιδευτεί περαιτέρω, στην συγκεκριμένη περίπτωση με το μοντέλο 1, αλλά στα δεδομένα μας παρατηρήθηκε το ίδιο για κάθε μοντέλο.

Στο σχήμα 23 παρατηρούμε ότι το μοντέλο 2 κάνει υπερ-προσαρμογή για 3*72 εισόδους, ομοίως και τα μοντέλα 3 και 4 οπότε και χρειάζεται early stop.

Για να πάρουμε το καλύτερα δυνατά μοντέλα, θα επανεκπαιδεύσουμε τα μοντέλα 2 και 4 για την περίπτωση 3*72 εισόδων και θα χρησιμοποιήσουμε την τεχνική ομαλοποίησης η οποία πρακτικά δίνει ποινή στα μεγάλα βάρη, με σκοπό να μειωθεί η υπερπροσαρμογή (κεφ 4.9). Το μοντέλο 3 θα το απορρίψουμε καθότι είχε το χειρότερο CV MSE μεταξύ των 2 και 4.

Στο σχήμα 24 παρουσιάζονται τα αποτελέσματα:

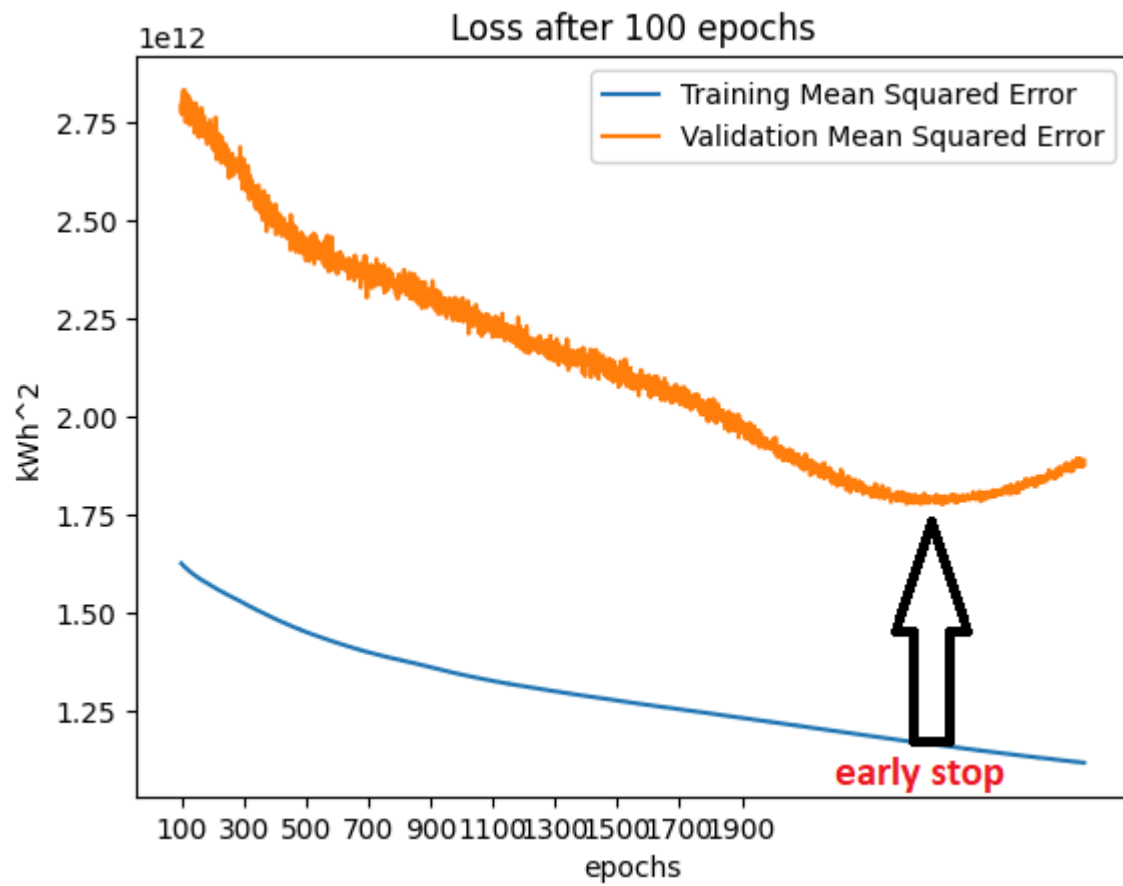
all_72.csv (3*72 είσοδοι)	
MSE	CV MSE
Μοντέλο 2 (Early stop + regularized) : 32-16	
Βέλτιστη Εποχή: 2547/3000	
$\sim 12 * 10^{11}$	$\sim 18 * 10^{11}$
Μοντέλο 4 (Early stop + regularized): 32-32-16	
Βέλτιστη Εποχή: 520/3000	
$\sim 12 * 10^{11}$	$\sim 18 * 10^{11}$

Σχήμα 24 Αποτελέσματα Εκπαίδευσης Επιλεγμένων Μοντέλων με τεχνικές άμβλυνσης υπερ-προσαρμογής

Η ομαλοποίηση που δοκιμάσαμε να εισάγουμε στα μοντέλα δεν άλλαξε σημαντικά τα αποτελέσματα.

Από το σχήμα 24 βλέπουμε ότι το μοντέλο 2 έχει όμοιες επιδόσεις με το μοντέλο 4, αλλά με λιγότερους νευρώνες και ένα κρυφό επίπεδο λιγότερο. Με αυτά τα δεδομένα θα επιλέξουμε ως τελικό μοντέλο το 2 με αρχείο εισόδου all_72.csv (3*72 εισόδους).

Τα ιστορικό της τελικής εκπαίδευσης φαίνεται στο παρακάτω σχήμα:



Σχήμα 25 Γραφική παράσταση σφάλματος. Μοντέλο 2 με τεχνική Early Stop + Regularization

5.8 Σύγκριση αποτελεσμάτων

Για την αξιολόγηση του τελικού μοντέλου, χρησιμοποιήσαμε το μέσο τετραγωνικό σφάλμα (MSE) καθώς και το μέσο απόλυτο ποσοστό σφάλματος (mean absolute percentage error - MAPE), του οποίου ο τύπος είναι:

$$MAPE = 100 * \frac{1}{n} \sum_{x=1}^n \left| \frac{actual_x - forecasted_x}{actual_x} \right|$$

Παρατηρούμε άμεσα ότι ο τύπος έχει περιορισμούς για τιμές όπου οι ροές μας είναι κοντά στο 0 καθώς είναι στον παρονομαστή του αθροίσματος. Αυτό δεν αποτελεί πρόβλημα στην προκειμένη καθώς οι ροές είναι αρκετά πάνω του μηδενός. Επίσης, δίνει μεγαλύτερο σφάλμα όταν η πραγματική τιμή είναι μικρότερη της εκτιμώμενης παρά το αντίθετο (Makridakis, 1993, 528), κάτι που την κάνει να προτιμά μοντέλα με προβλέψεις που υποτιμούν την πραγματική τιμή. Παρόλα αυτά, έχοντας υπόψιν τους περιορισμούς, είναι εξεφρασμένη επί τοις εκατό και άρα εύκολη στην κατανόηση.

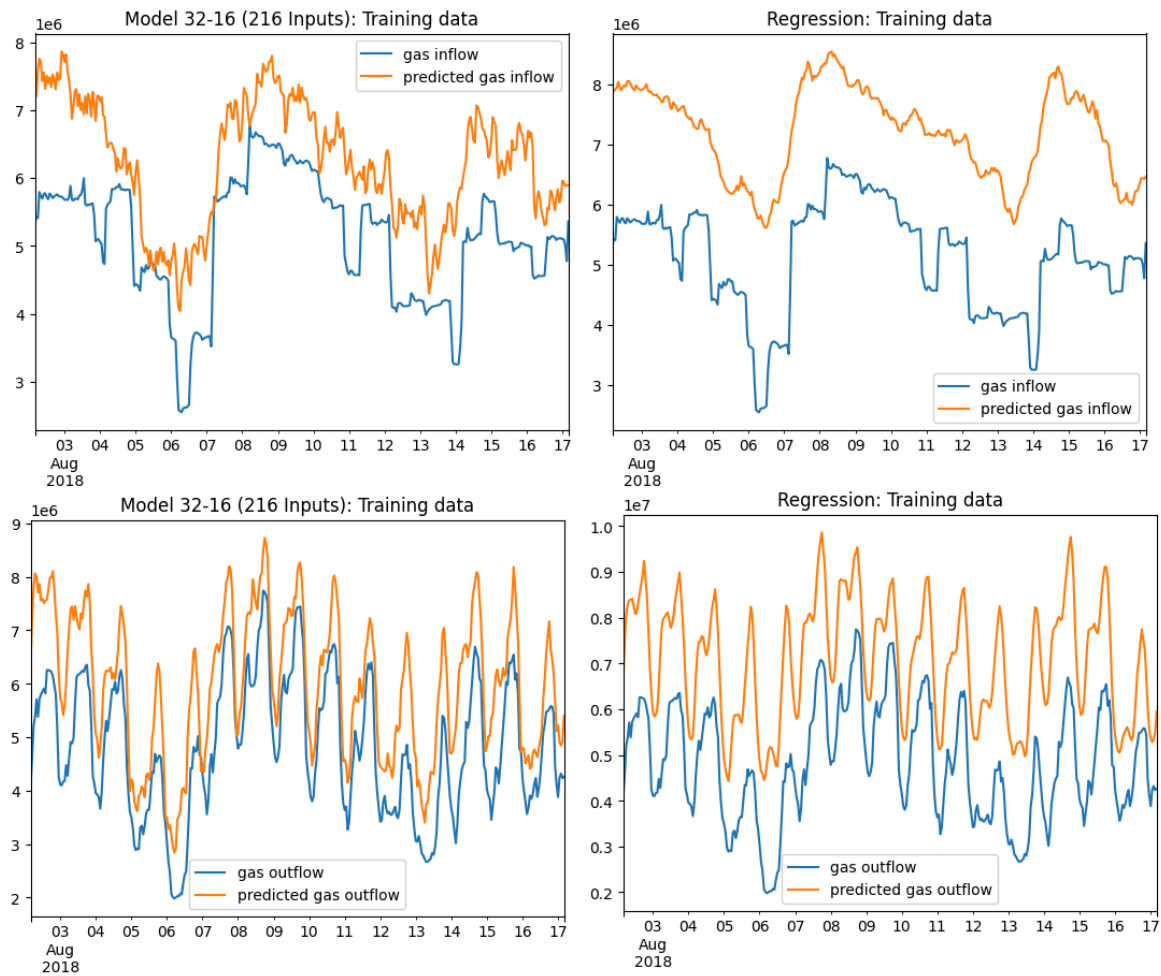
Θα εφαρμόσουμε τα δυο παραπάνω σφάλματα στο μοντέλο μας για τα σετ εκπαίδευσης, επαλήθευσης, καθώς και αυτό της δοκιμής (test) για να δοκιμάσουμε τις δυνατότητες πρόβλεψης σε άγνωστα δεδομένα. Ακόμα θα κάνουμε τα παραπάνω για ένα μοντέλο πολλαπλής γραμμικής παλινδρόμησης (Multiple Linear Regression - MLR) για να τα συγκρίνουμε.

Τα αποτελέσματα των παραπάνω δίνονται στον πίνακα του σχήματος 26:

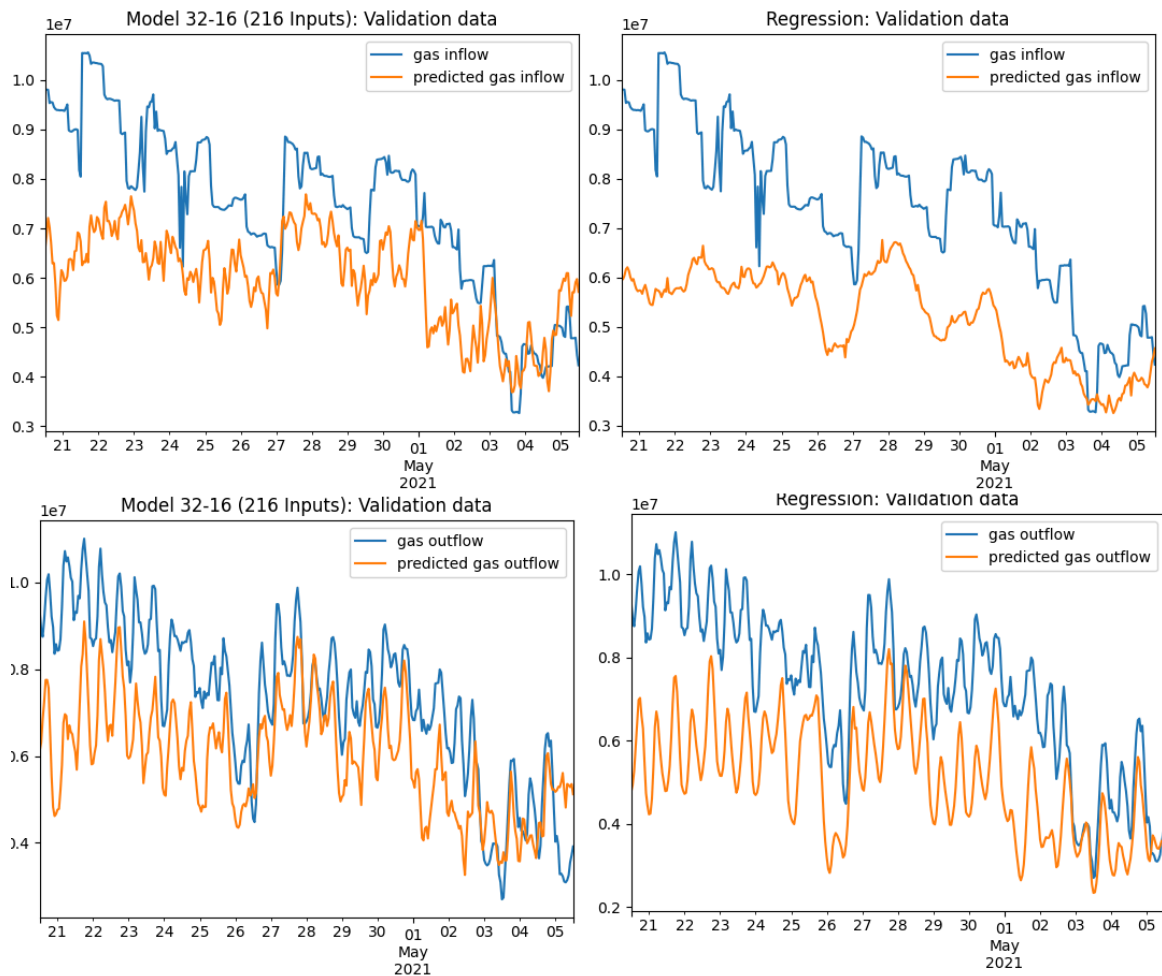
Μετρική - Σετ	Νευρωνικό Μοντέλο 2	Πολλαπλή Γρ. Παλινδρόμηση
MSE - Εκπαίδευση	$\sim 12 * 10^{11}$	$\sim 18 * 10^{11}$
MAPE - Εκπαίδευση	$\sim 13\%$	$\sim 17\%$
MSE - Επαλήθευση	$\sim 18 * 10^{11}$	$\sim 33 * 10^{11}$
MAPE - Επαλήθευση	$\sim 15\%$	$\sim 19\%$
MSE - Δοκιμή	$\sim 21 * 10^{11}$	$\sim 26 * 10^{11}$
MAPE - Δοκιμή	$\sim 15\%$	$\sim 15\%$

Σχήμα 26 Σύγκριση αποτελεσμάτων μεταξύ νευρωνικού και παλινδρόμησης

Τέλος, στο σχήμα 27 παρατίθενται ποιοτικές συγκρίσεις των προβλέψεων του μοντέλου 2 καθώς και της πολλαπλής γραμμικής παλινδρόμησης για αυθαίρετα επιλεγμένα χρονικά διαστήματα.

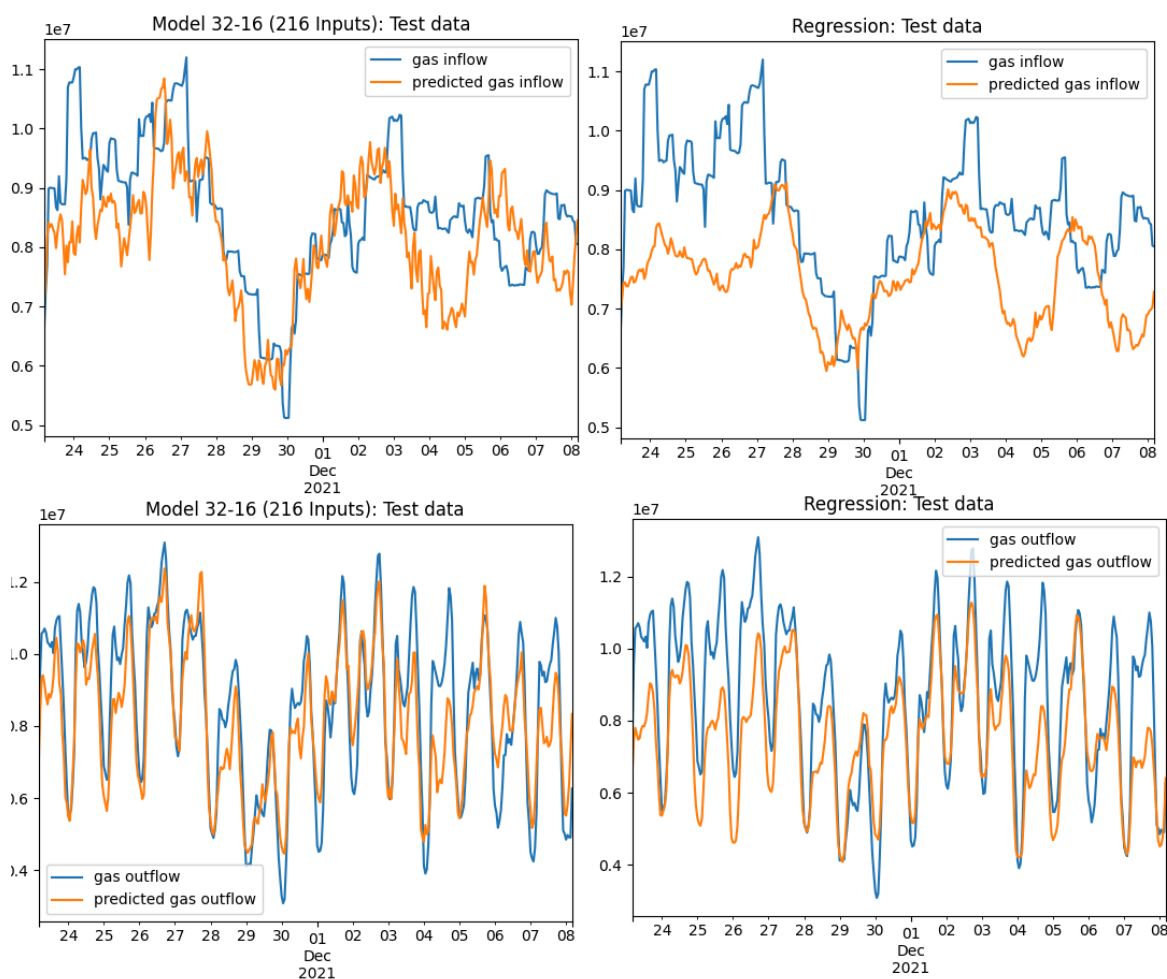


Σχήμα 27.α) Δεδομένα εκπαίδευσης (Train):
 Προβλέψεις μοντέλου 2 (32-16, αριστερά) και γραμμικής παλινδρόμησης (δεξιά)
 για εισροές (πάνω) και εκροές (κάτω) φυσικού αερίου. Με πορτοκαλί οι
 προβλέψεις και μπλε οι πραγματικές τιμές. Τιμές σε kWh.



Σχήμα 27.β) Δεδομένα επαλήθευσης (Validation):

Προβλέψεις μοντέλου 2 (αριστερά) και γραμμικής παλινδρόμησης (δεξιά) για εισροές (πάνω) και εκροές (κάτω) φυσικού αερίου. Με πορτοκαλί οι προβλέψεις και μπλε οι πραγματικές τιμές. Τιμές σε kWh.



Σχήμα 27.γ) Δεδομένα δοκιμής (Test):

Προβλέψεις μοντέλου 2 (αριστερά) και γραμμικής παλινδρόμησης (δεξιά) για εισροές (πάνω) και εκροές (κάτω) φυσικού αερίου. Με πορτοκαλί οι προβλέψεις και μπλε οι πραγματικές τιμές. Τιμές σε kWh.

5.9 Συμπεράσματα

Παρατηρούμε ότι το νευρωνικό μοντέλο 2 στον πίνακα του σχήματος 27 δείχνει καλύτερες επιδόσεις για το σετ εκπαίδευσης και επαλήθευσης σε σχέση με την πολλαπλή γραμμική παλινδρόμηση, και ικανοποιητικά καλές προβλέψεις στο σύνολο δοκιμής, κάτι το οποίο παρατηρείται και ποιοτικά στο σχήμα 27.γ. Μία πιθανή βελτίωση ενδεχομένως να ήταν ένας άλλος συνδυασμός νευρώνων, αλλαγή ρυθμού μάθησης, συναρτήσεων ενεργοποίησης ή/και σφάλματος, χαρακτηριστικών (features) κλπ που θα κατάφερνε να βελτιώσει περαιτέρω την ακρίβεια πρόβλεψης στο σετ δοκιμής.

Σε γενικές γραμμές πάντως παρατηρούμε και ποιοτικά από τα γραφήματα πως, παρότι η γραμμική παλινδρόμηση καταφέρνει να ακολουθήσει αρκετά πιστά τη κεντρική τάση των αντίστοιχων χρονοσειρών, το νευρωνικό μοντέλο πέρα από την

τάση υπερτερεί και στην αφομοίωση των επί μέρους μη περιοδικών αυξομειώσεων. Αυτό συμβαδίζει με τις φύσεις των δύο μοντέλων καθώς το πρώτο εξειδικεύεται στον εντοπισμό σχετικά απλών γραμμικών εξαρτήσεων, ενώ το δεύτερο στον εντοπισμό πιο σύνθετων μη γραμμικών σχέσεων. Ενδεχομένως αυτό θα μπορούσε να προτείνει την κατεύθυνση της κατασκευής ενός υβριδικού μοντέλου όπου κάθε κομμάτι θα επικεντρώνεται στην αντίστοιχη συνιστώσα της χρονοσειράς (δηλαδή γραμμική παλινδρόμηση για την κεντρική τάση και νευρωνικό δίκτυο για τις βραχυχρόνιες αυξομειώσεις).

Σε κάθε περίπτωση γίνεται εμφανές τόσο από τα νούμερα όσο και από τα γραφήματα ότι το μοντέλο, ακόμη και με τους τρέχοντες περιορισμούς του, είναι ικανό να προσεγγίσει σε αξιολογικά μη τετριμμένο βαθμό την τάση των πραγματικών εισροών και εκροών του φυσικού αερίου. Πράγματι είναι εμφανές ότι οι αποκλίσεις είναι μορφολογικά και ποσοτικά αρκετά εύλογες ώστε να μην υπάρχει κίνδυνος να οφείλονται σε κάποια θεμελιώδη ανεξαρτησία του ζητούμενου από τα διαθέσιμα δεδομένα, αλλά αντιθέτως η πιθανή αιτία τους να είναι ο θόρυβος που εκρέει εν μέρει από τις αναπόφευκτες ατέλειες της εκπαιδευτικής διαδικασίας του μοντέλου κι εν μέρει από την ίδια τη φύση των δεδομένων.

5.10 Επόμενα Βήματα

Για να επιτύχουμε καλύτερες προβλέψεις, μπορούμε να δοκιμάσουμε χειροκίνητη ή αυτοματοποιημένη ρύθμιση υπερ-παραμέτρων (hyperparameter tuning). Η χειροκίνητη ρύθμιση αφορά δοκιμές «με το χέρι» όπως εκτελέσαμε με τα 5 μοντέλα, είτε μπορούμε αυτόματα μέσω επιλεκτικής ή εξαντλητικής αναζήτησης των βέλτιστων υπερπαραμέτρων, ή αποδοτικότερα με πιο εκλεπτυσμένες μεθόδους όπως χρήση γενετικών αλγορίθμων ή Μπεϋζιανή βελτιστοποίηση (Loussaief & Abdelkrim, 2020, 258-259).

Εναλλακτικά, μπορούμε να δοκιμάσουμε άλλου είδους αρχιτεκτονικές νευρωνικών δικτύων, όπως τα LSTM που ανήκουν στην κατηγορία RNN και που παρέχουν μεγαλύτερη πολυπλοκότητα από τα απλά νευρωνικά δίκτυα καθώς μπορούν να επαναρυθμίζουν τη ροή της πληροφορίας εισόδου όπως αρμόζει. Τέλος, θα μπορούσαμε για πληρότητα να εξετάσουμε διαφορετικής φύσης μοντέλα όπως το XGBoost ή μοντέλα αυτοπαλινδρόμησης συναφή με προβλήματα χρονοσειρών όπως αυτοπαλινδρομική αθροιστική διαδικασία κινητού μέσου (ARIMA). (Paliari et al., 2021, 1)

Σε κάθε περίπτωση έχει νόημα η εξερεύνηση δρόμων τόσο για την αποδοτικότερη ρύθμιση και αναζήτηση των κατάλληλων (υπερ-)παραμέτρων όσο και για τη παράλληλη αποδοτική διαχείριση μεγαλύτερου όγκου δεδομένων - είτε αναφερόμαστε σε επέκταση του τρέχοντος μοντέλου είτε σε δοκιμή κάποιου εναλλακτικού όπως τα προαναφερθέντα. Αυτό είναι κάτι που ασφαλώς θα έχει

αυξημένες απαιτήσεις σε χρονικούς, υπολογιστικούς και χρηματικούς πόρους, αλλά που παράλληλα, όπως είδαμε από τα αποτελέσματα της παρούσας εργασίας, φαίνεται να αποτελεί μια επένδυση ικανή να καρποφορήσει ακόμη πιο χρήσιμα και προσοδοφόρα αποτελέσματα στο εγγύς μέλλον.

Παράρτημα

Ο πλήρης κώδικας Python που χρησιμοποιήθηκε για την συγγραφή της παρούσας διπλωματικής εργασίας βρίσκεται στον υπερσύνδεσμο

https://github.com/stratos1991/diplo24_25

Εδώ θα παρουσιαστούν μερικά από τα σημαντικότερα κομμάτια κώδικα:

- Υπολογισμός της χρονοσειράς με κενά δεδομένα από ένα πλαίσιο δεδομένων (εφεξής DataFrame):

```
def missing_hours_series(df: pd.DataFrame, start='2018-08-01 04:00:00',end='2022-09-15 09:00:00'):
    """index must be a DateTimeIndex for this function to make sense"""
    missingDatetimes = pd.date_range(start=start, end=end, freq='h').difference(df.index)
    return missingDatetimes.to_series()
```

- Εύρεση συνεχών τμημάτων μιας χρονοσειράς:

```
def contiguous_hours(s: pd.Series):
    """series s must be of Timestamp type (hourly data) for this function to make sense"""
    threshold = pd.Timedelta(hours=1)

    # Calculate the differences between consecutive datetime values
    diffs = s.diff()

    # Identify groups of contiguous datetime values
    groups = (diffs > threshold).cumsum()

    # Aggregate contiguous groups
    result = s.groupby(groups).agg(['min', 'max']).reset_index(drop=True)
    result['time']=result['max']-result['min']
    result.rename(columns={"min": "From", "max": "To"},inplace=True)
    return result
```

- Επέκταση ενός DataFrame με στήλες προηγούμενων ωρών:

```
def fill_last_X_hours(df:pd.DataFrame,X):
    """Index must be (hourly) timeseries"""
    temp=pd.DataFrame([])
    for i in range(X):
        temp=temp.join(df.shift(i,freq='h'),how='outer',suffix=' t-'+str(i))
    return temp
```

- Εύρεση «μικρών» (για τα δεδομένα μας) και μηδενικών τιμών προς απόρριψη:

```
print(data[data['sum(physicalFlow)0 inflow'] < 100000])
print(data[data['sum(physicalFlow)1 outflow'] < 100000])
```

- Δημιουργία γραφικής παράστασης αυτοσυσχέτισης με προσαρμοσμένο άξονα x:

```
x=pd.plotting.autocorrelation_plot(df_solar['AggregatedGenerationForecast']).set_xlim([0, 24*15])
```

```
plt.title('Solar')
plt.xticks(ticks=[24,48,72,96])
plt.show()
```

- Διαχωρισμός δεδομένων σε κατηγορίες εκπαίδευσης, επαλήθευσης και δοκιμής:

```
# https://datascience.stackexchange.com/questions/15135/train-test-validation-set-splitting-in-sklearn
from sklearn.model_selection import train_test_split

train_ratio = 0.75 # training
validation_ratio = 0.15 # hyperparameter tuning
test_ratio = 0.10 # final evaluation

X_train, X_test, y_train, y_test=train_test_split(X,y,test_size=test_ratio,shuffle=False,
random_state=1234)

X_train, X_val, y_train, y_val = train_test_split(X_train, y_train,
test_size=validation_ratio/(1-test_ratio),shuffle=False, random_state=1)

print('train percent',y_train.size/X.index.size, '\ttrain days',y_train.index.size/24 )
print('validation percent',y_val.index.size/X.index.size,'\tvalid days',y_val.index.size/24)
print('test percent',y_test.index.size/X.index.size,'\ttest days',y_test.index.size/24)
```

- Κανονικοποίηση χαρακτηριστικών (Feature Scaling) με την βοήθεια του scikit-learn:

```
from sklearn.preprocessing import StandardScaler
sc=StandardScaler()

X_train=sc.fit_transform(X_train)
X_val=sc.transform(X_val)
X_test=sc.transform(X_test)
```

- Δημιουργία και εκπαίδευση ενός πλήρως συνδεδεμένου νευρωνικού μοντέλου με 6+6 κρυφούς νευρώνες και 2 γραμμικές εξόδους:

```
def new_nn():
    ann=tf.keras.models.Sequential()
    ann.add(tf.keras.layers.Dense(units=6,activation='relu'))
    ann.add(tf.keras.layers.Dense(units=6,activation='relu'))
    ann.add(tf.keras.layers.Dense(units=2,activation='linear'))
    ann.compile(optimizer='adam',loss=tf.keras.losses.MeanSquaredError())
    return ann

new=new_nn()
history=new.fit(X_train,y_train,epochs=1800,validation_data=(X_val,y_val),verbose=2)
```

- Υπολογισμός MAPE και MSE με την βοήθεια του scikit-learn:

```
print('mape_train: ',mean_absolute_percentage_error(y_train,y_train_prediction))
print('mse_test: ',mean_squared_error(y_test,y_TEST_prediction))
```

Βιβλιογραφία

- ΑΔΜΗΕ. (n.d.). ΑΔΜΗΕ: Ανεξάρτητος Διαχειριστής Μεταφοράς Ηλεκτρικής Ενέργειας. Retrieved December 5, 2024, from <https://www.admie.gr/>
- ΔΑΠΕΕΠ. (n.d.). Retrieved 12 5, 2024, from <https://www.dapeep.gr/o-skopos-tis-etaireias/>
- ΔΕΣΦΑ. (n.d.). Retrieved 12 5, 2024, from <https://www.desfa.gr/company>
- Aleksander, I. (1990). *Introduction to Neural Computing*. Chapman & Hall.
- Bellman, R. (2013). *Dynamic Programming*. Dover Publications.
- Bousquet, O., Luxburg, U. v., & Rätsch, G. (Eds.). (2011). *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures*. Physica-Verlag.
https://link.springer.com/chapter/10.1007/978-3-540-28650-9_5
- Directorate for the Media. (2008). 3rd Energy Package gets final approval from MEPs.
[https://www.europarl.europa.eu/RegData/presse/pr_focus/2008/EN/03A-DV-PRESSE_FCS\(2008\)06-16\(31737\)_EN.pdf](https://www.europarl.europa.eu/RegData/presse/pr_focus/2008/EN/03A-DV-PRESSE_FCS(2008)06-16(31737)_EN.pdf)
- Directorate-General for Energy. (n.d.). *Diversification of gas supply sources and routes*. Energy - European Commission. Retrieved October 15, 2024, from https://energy.ec.europa.eu/topics/energy-security/diversification-gas-supply-sources-and-routes_en
- Directorate-General for Energy. (n.d.). *Liquefied natural gas*. Energy - European Commission. Retrieved October 3, 2024, from https://energy.ec.europa.eu/topics/carbon-management-and-fossil-fuels/liquefied-natural-gas_en
- Directorate-General for Energy. (n.d.). *Security of gas supply*. Energy - European Commission. Retrieved October 3, 2024, from https://energy.ec.europa.eu/topics/energy-security/security-gas-supply_en
- Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2021, 9 29). *Activation Functions in Deep Learning: A comprehensive survey and benchmark*. arXiv.org.
<https://arxiv.org/abs/2109.14545>
- Duchi, J., Hazan, E., & Singer, Y. (2011). *Adaptive Subgradient Methods for Online Learning and Stochastic Optimization**. Journal of Machine Learning Research. Retrieved January 9, 2025, from <https://www.jmlr.org/papers/volume12/duchi11a/duchi11a.pdf>
- Enargia A.G. (n.d.). *Enargia website*. Enargia | Advanced Energy Analytics. Retrieved 12 5, 2024, from <https://www.enargia.ch/>
- EnExGroup. (n.d.). Retrieved 12 5, 2024, from <https://www.enexgroup.gr/el/web/guest/enexgroup>
- ENNOH Website. (n.d.). ENNOH: Welcome. Retrieved October 16, 2024, from <https://ennoh.eu/>

- ENTSO-e. (n.d.). *About ENTSO-E*. ENTSO-e. Retrieved December 5, 2024, from <https://www.entsoe.eu/about/>
- ENTSOG. (n.d.). *About ENTSOG*. ENTSOG. Retrieved October 15, 2024, from <https://www.entsog.eu/about-entsog>
- ENTSOG Gas Dashboard*. (n.d.). European Gas Flow dashboard by ENTSOG. Retrieved October 15, 2024, from <https://gasdashboard.entsog.eu/>
- ENTSOG Transparency Platform*. (n.d.). TP. Retrieved November 7, 2024, from <https://transparency.entsog.eu/>
- European Commission: Directorate-General for Energy. (2023). *EU Energy in Figures: Statistical Pocketbook 2023*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2833/502436>
- European Union Agency for the Cooperation of Energy Regulators (ACER). (2021). *Gas factsheet* | www.acer.europa.eu. [acer.europa.eu](http://www.acer.europa.eu). Retrieved October 2, 2024, from <https://www.acer.europa.eu/gas-factsheet>
- Eurostat. (2024, June 5). *Energy consumption in households - Statistics Explained*. European Commission. Retrieved October 3, 2024, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households
- Fang, W., Chen, Y., & Xue, Q. (2021). Survey on research of RNN-Based Spatio-Temporal Sequence Prediction Algorithms. *Journal on big data*, 3(3). 10.32604/jbd.2021.016993
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. <https://www.deeplearningbook.org/>
- Google. (n.d.). *Google Colab*. Google Research. Retrieved January 17, 2025, from <https://research.google.com/colaboratory/faq.html>
- Hagiwara, K. (2002). Regularization learning, early stopping and biased estimator. *Neurocomputing*, 48(1-4), 937-955. 10.1016/s0925-2312(01)00681-6
- Hancher, L., Belmans, R., Conti, I., Meeus, L., Ferrari, A., Nouicer, A., Galdi, G., Reif, V., Kneebone, J., Olczak, M., Piebalgs, A., Pototschnig, A., Stampatori, D., & Schittekatte, T. (2022). *The EU Green Deal (2022 Ed.)* (L. Hancher, L. Meeus, A. Nouicer, & V. Reif, Eds.). European University Institute (EUI). 10.2870/00714
- Haykin, S. S. (2009). *Neural networks and learning machines*. Prentice Hall.
- Hornik, K. (1991). *Approximation capabilities of multilayer feedforward networks*. Approximation Capabilities of Multilayer Feedforward Networks. Retrieved January 14, 2025, from https://web.njit.edu/%7Eusman/courses/cs677_spring21/hornik-nn-1991.pdf
- International Energy Agency. (n.d.). *Europe – Countries & Regions - IEA*. Europe – Countries & Regions - IEA. Retrieved October 3, 2024, from <https://www.iea.org/regions/europe/natural-gas>
- Jupyter. (n.d.). Project Jupyter | Home. Retrieved January 17, 2025, from <https://jupyter.org/>

- Keras. (n.d.). Keras: Deep Learning for humans. Retrieved January 17, 2025, from <https://keras.io/>
- Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *Computer Science*. 10.48550/arXiv.1412.6980
- LeCun, Y., Bottou, L., Orr, G., & Müller, K.-R. (2012). Efficient BackProp. *Lecture Notes in Computer Science*, 7700, 9-48. https://doi.org/10.1007/978-3-642-35289-8_3
- Loussaief, S., & Abdelkrim, A. (2020, 12). Convolutional Neural Network Hyper-Parameters Optimization based on Genetic Algorithms. *International Journal of Advanced Computer Science and Applications*, 9(10), 252-266. 10.14569/ijacsa.2018.091031
- Makridakis, S. (1993). Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting*, 9(4), 527-529. 10.1016/0169-2070(93)90079-3
- McCulloch, W. S., & Pitts, W. (1990, 1 1). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 52(1-2), 99-115. 10.1007/bf02459570
- Mishra, M. (2023, May 27). *The Learning Rate: A Hyperparameter That Matters* | by Mohit Mishra | Medium. Mohit Mishra. Retrieved January 14, 2025, from <https://mohitmishra786687.medium.com/the-learning-rate-a-hyperparameter-that-matters-b2f3b68324ab>
- Moody, J. E. (1991, 12 2). The Effective Number of Parameters: An Analysis of Generalization and Regularization in Nonlinear Learning Systems. *Neural Information Processing Systems*, 4, 847-854. https://proceedings.neurips.cc/paper_files/paper/1991/file/d64a340bcb633f536d56e51874281454-Paper.pdf
- Nelles, O. (2001). *Nonlinear System Identification*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-04323-3_7
- Oyedotun, O., Olaniyi, E., & Khashman, A. (2017). A simple and practical review of over-fitting in neural network learning. *International Journal of Applied Pattern Recognition*, 4(4), 307-328. 10.1504/ijapr.2017.089384
- Paliari, I., Karanikola, A., & Kotsiantis, S. (2021, 7 12). A comparison of the optimized LSTM, XGBOOST and ARIMA in Time Series forecasting. *2021 12th International Conference on Information, Intelligence, Systems & Applications (IISA)*, 1-7. 10.1109/iisa52424.2021.9555520
- Qiang, W., & Zhongli, Z. (2011, 9). Reinforcement learning model, algorithms and its Application. *2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)*, 1143-1146. 10.1109/mec.2011.6025669
- Regulation (EU) 2024/1789. (2024, 6 13). *Κανονισμός (ΕΕ) 2024/1789 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου, της 13ης Ιουνίου 2024, σχετικά με τις εσωτερικές αγορές ανανεώσιμου αερίου, φυσικού αερίου και υδρογόνου[...]*. <https://eur-lex.europa.eu/legal-content/EL/TXT/?uri=CELEX:32024R1789>

- scikit-learn. (n.d.). scikit-learn: machine learning in Python — scikit-learn 1.6.1 documentation. Retrieved January 17, 2025, from <https://scikit-learn.org/>
- Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., & Hassabis, D. (2017, December 5). Mastering chess and shogi by self-play with a general reinforcement learning algorithm. <https://arxiv.org/abs/1712.01815v1>
- Singh, S., Kumar, R., Payra, S., & Singh, S. K. (2023, August 30). Artificial Intelligence and machine learning in pharmacological research: Bridging the gap between Data and Drug Discovery. *Cureus*. 10.7759/cureus.44359
- Sjöberg, J., & Ljung, L. (1992). Overtraining, Regularization, and Searching for Minimum in Neural Networks. *IFAC Proceedings Volumes*, 25(14), 73-78. 10.1016/s1474-6670(17)50715-6
- Smart RUE of ECE-NTUA. (n.d.). HMMY - Smart RUE of ECE-NTUA. Retrieved 12 5, 2024, from <https://www.ece.ntua.gr/gr/article/492>
- Stannov, K. (2018). *ENTSOG Transparency Platform Overview of publication options and obligations*. https://www.energy-community.org/dam/jcr:8160fcef-0588-43aa-b926-98cd46c802ba/REGSCHOOL_102018_ENTSOG_Platform.pdf
- TensorFlow. (n.d.). TensorFlow.org. Retrieved January 17, 2025, from <https://www.tensorflow.org/>
- Verleysen, M., & François, D. (2005). The Curse of Dimensionality in Data Mining and Time Series Prediction. *Computational Intelligence and Bioinspired Systems*, 758-770. 10.1007/11494669_93
- Where does the EU's gas come from? (n.d.). Consilium.europa.eu. Retrieved October 3, 2024, from <https://www.consilium.europa.eu/en/infographics/eu-gas-supply/>
- Wilson, D.R., & Martinez, T. R. (2003). The general inefficiency of batch training for gradient descent learning. *Neural Networks*, 16(10), 1429-1451. 10.1016/s0893-6080(03)00138-2
- Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415, 295-316. 10.1016/j.neucom.2020.07.061