



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

Εκτίμηση κατάστασης και τοπολογίας συστημάτων διανομής με χρήση τεχνητής νοημοσύνης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χαράλαμπος Δ. Σαπουντζάκης

Επιβλέπων : Άρης Ευάγγελος Δημέας
Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Μάρτιος 2024



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

Εκτίμηση κατάστασης και τοπολογίας συστημάτων διανομής με χρήση τεχνητής νοημοσύνης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χαράλαμπος Δ. Σαπουντζάκης

Επιβλέπων : Άρης Ευάγγελος Δημέας
Επίκουρος Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 20^η Μαρτίου 2024

.....
Δημέας Άρης-Ευάγγελος
Επίκ. Καθηγητής Ε.Μ.Π.

.....
Κορρές Γεώργιος
Καθηγητής Ε.Μ.Π.

.....
Γεωργιάκης Παύλος
Καθηγητής Ε.Μ.Π.

Αθήνα, Μάρτιος 2024

.....

Χαράλαμπος Δ. Σαπουντζάκης

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

Copyright © Χαράλαμπος Δ. Σαπουντζάκης, 2024.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Τα δίκτυα διανομής είναι συνήθως μεγάλα και πολύπλοκα συστήματα αποτελούμενα από πολυάριθμα διασυνδεδεμένα στοιχεία, γεγονός που δυσχαιραίνει την επίγνωση της κατάστασης του δικτύου σε μεμονωμένες στιγμές. Επίσης, τα σύγχρονα δίκτυα διανομής αντιμετωπίζουν συνεχώς νέες προκλήσεις, όπως είναι η ολοένα αυξανόμενη ενσωμάτωση των ανανεώσιμων πηγών ενέργειας, η αβεβαιότητα που αυτές συνεπάγονται, η εξάπλωση των καταναμημένων ενεργειακών πόρων και η αμφίδρομη ροή ισχύος που αυτό προκαλεί. Ο τρόπος αντιμετώπισης αυτών των προκλήσεων τα μεταμορφώνει από παθητικές σε ενεργητικές οντότητες και τα μετασχηματίζει σε έξυπνα δίκτυα που αξιοποιούν τις ψηφιακές τεχνολογίες, δίκτυα επικοινωνίας και την αυτοματοποίηση.

Η εκτίμηση κατάστασης στα συστήματα ηλεκτρικής ενέργειας είναι μία διαδικασία απαραίτητη για την διασφάλιση της αξιοπιστίας και αποδοτικής λειτουργίας τους. Αυτό επειδή παρέχει ακριβείς πληροφορίες σε πραγματικό χρόνο σχετικά με την κατάσταση και τη συμπεριφορά του συστήματος, οι οποίες είναι απαραίτητες για τη βελτιστοποίηση των λειτουργιών, την ενίσχυση της αξιοπιστίας και τη δημιουργία προηγμένων στρατηγικών ελέγχου.

Ένα σημαντικό εμπόδιο στην εκτίμηση της κατάστασης των δικτύων διανομής είναι η έλλειψη παρατηρησιμότητας τους, λόγω του περιορισμένου αριθμού αισθητήρων ή της μη βέλτιστης τοποθέτησης τους στο δίκτυο. Ως αποτέλεσμα, οι κλασικοί εκτιμητές κατάστασης (όπως ο WLS – σταθμισμένων ελαχίστων τετραγώνων) που υποθέτουν ντετερμινιστική κατάσταση του συστήματος δεν μπορούν να παρέχουν εγγυημένα ακριβείς μετρήσεις με αξιοπίστο τρόπο.

Η εκτίμηση της κατάστασης για μη παρατηρήσιμα συστήματα πρέπει να συμπεριλαμβάνει πρόσθετες ιδιότητες πέραν του μοντέλου μέτρησης που ορίζεται από τα ισοζύγια ροής ισχύος. Για το σκοπό αυτό, στην παρούσα εργασία, αναπτύσσεται μια Μπεϋζιανή προσέγγιση, όπου οι καταστάσεις του συστήματος θεωρούνται οι φασιθέτες της τάσης των κόμβων του δικτύου.

Το σημαντικότερο πλεονέκτημα αυτής της προσέγγισης είναι ότι η παρατηρησιμότητα δεν απαιτείται πλέον. Σε αντίθεση με τις παραδοσιακές τεχνικές εκτίμησης κατάστασης που ελαχιστοποιούν το σφάλμα μοντελοποίησης, ο εκτιμητής αυτός ελαχιστοποιεί άμεσα το σφάλμα εκτίμησης. Ένα βαθύ νευρωνικό δίκτυο κατασκευάζεται για την προσέγγιση αυτού του εκτιμητή κατάστασης, λόγω των εξαιρετικών δυνατοτήτων προσέγγισης που έχει αποδειχθεί πως αυτά διαθέτουν.

Ωστόσο, καθώς η τοπολογία ενός δικτύου διανομής αλλάζει με το χρόνο, είναι σημαντικό να εξεταστούν οι επιπτώσεις αυτής της αλλαγής στην εκτίμηση κατάστασης. Έτσι, κατασκευάστηκε κι ένα ακόμη βαθύ νευρωνικό δίκτυο, το οποίο αναγνωρίζει την τοπολογία του δικτύου και στη συνέχεια, αν αυτή αλλάξει, χρησιμοποιείται μια τεχνική που ονομάζεται transfer learning έτσι ώστε το προηγούμενο, νευρωνικό δίκτυο για εκτίμηση κατάστασης να προσαρμοστεί σε αυτή την αλλαγή. Τα δύο αυτά νευρωνικά δίκτυα εκπαιδεύονται με δεδομένα που παράγονται μέσω προσομοιώσεων Monte Carlo. Τέλος, αναπτύχθηκε ένας αλγόριθμος βέλτιστης τοποθέτησης μετρητών στο δίκτυο, έτσι ώστε να χρησιμοποιηθεί ο

ελάχιστος δυνατός αριθμός PMUs για την απόκτηση αξιόπιστων και ακριβών αποτελεσμάτων.

Λέξεις Κλειδιά

Εκτίμηση Κατάστασης, Εκτίμηση Τοπολογίας, Δίκτυα Διανομής, Τεχνητή Νοημοσύνη

Abstract

Distribution networks are usually large and complex systems consisting of numerous interconnected elements, which makes it difficult to know the state of the network at individual moments. Furthermore, modern distribution networks are constantly facing new challenges, such as the ever-increasing integration of renewable energy sources, the uncertainty they entail, the spread of distributed energy resources and the two-way power flow this causes. How to address these challenges transforms them from passive to active entities and into smart grids that leverage digital technologies, communication networks and automation.

State estimation in power systems is a process essential to ensure their reliable and efficient operation. This is because it provides accurate, real-time information about system status and behaviour, which is essential for optimising operations, enhancing reliability and creating advanced control strategies.

A major obstacle to assessing the status of distribution networks is their unobservability, due to the limited number of measurement devices or their suboptimal placement in the network. As a result, classical state estimators (such as WLS - weighted least squares) that assume deterministic system state cannot provide guaranteed accurate measurements in a reliable way.

State estimation for unobservable systems must include additional properties beyond the measurement model defined by power flow equilibria. For this purpose, in this paper, a Bayesian approach is developed, where the system states are considered as the voltage phasors of the network nodes.

The main advantage of this approach is that complete observability is no longer required. Unlike traditional state estimation techniques that minimize the modelling error, this estimator directly minimizes the estimation error. A deep neural network is constructed to approximate this state estimator because of the excellent approximation capabilities that they have been shown to possess.

However, as the topology of a distribution network changes over time, it is important to consider the impact of it on state estimation. Thus, another deep neural network was built, which identifies the network topology and then if it changes, a technique called transfer learning is used so that the previous, state estimation neural network can adapt to this change. These two neural networks are trained with data generated through Monte Carlo simulations. Finally, an algorithm for optimal placement of meters in the network was developed so that the minimum number of PMUs possible is used to obtain reliable and accurate results.

Keywords

State estimation, Topology Identification, Distribution Networks, Artificial Intelligence

Ευχαριστίες

Θα ήθελα αρχικά να ευχαριστήσω τον καθηγητή κ. Άρη Δημέα για την επίβλεψη της εργασίας κατά το περασμένο εξάμηνο και για την ευκαιρία που μου έδωσε να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα και που με ώθησε να ενδιαφερθώ για τον κλάδο της τεχνητής νοημοσύνης, με τον οποίο δεν είχα πρωτύτερα ασχοληθεί. Επίσης θα ήθελα να ευχαριστήσω θερμά τον υποψήφιο διδάκτορα Κυριάκο Ανδρεσάκη και τον Δρ. Θεμιστοκλή Ξύγκη για την πολύτιμη βοήθεια τους κατά την εκπόνηση αυτής της εργασίας. Τέλος θα ήθελα να ευχαριστήσω τον αδερφό μου για την βοήθεια και την συμπαράσταση που μου προσέφερε καθ'όλη την διάρκεια των σπουδών μου.

Περιεχόμενα

Θεωρητικό Μέρος

<u>Κεφάλαιο 1</u>	14
<u>1.1 Συστήματα Ηλεκτρικής Ενέργειας</u>	14
<u>1.2 Δίκτυα Διανομής</u>	15
<u>1.2.1 Εξοπλισμός Δικτύων Διανομής</u>	17
<u>1.2.2 Τοπολογία δικτυων διανομης</u>	20
<u>1.2.3 Τα Δίκτυα Διανομής Σήμερα</u>	21
<u>1.3 Ροή Φορτίου</u>	23
<u>1.4 Synchrophasor Measurement Devices – SMDs</u>	26
<u>1.4.1 μ-PMUs – D-PMUs</u>	28
<u>Κεφάλαιο 2</u>	31
<u>2.1 Εκτίμηση κατάστασης</u>	31
<u>2.2 Εκτίμηση κατάστασης σε ΣΗΕ</u>	34
<u>2.3 Μαθηματικός Ορισμός</u>	36
<u>2.4 Λόγοι Δυσκολίας Εκτέλεσης DSSE</u>	37
<u>2.5 Επεξεργαστής Τοπολογίας - Βιβλιογραφική Ανασκόπηση</u>	39
<u>2.6 Παρατηρησιμότητα</u>	40
<u>2.7 Τύποι Μετρήσεων - Βιβλιογραφική Ανασκόπηση</u>	42
<u>2.8 Βέλτιστη τοποθέτηση PMUs - Βιβλιογραφική Ανασκόπηση</u>	43
<u>2.9 Διατύπωση Προβλήματος – Βιβλιογραφική Ανασκόπηση DSSE</u>	44
<u>2.10 Προτεινόμενη Μεθοδολογία</u>	49
<u>Κεφάλαιο 3</u>	51
<u>3.1 Τεχνητή Νοημοσύνη</u>	51
<u>3.2 Μηχανική Μάθηση</u>	52
<u>3.2.1 Βαθιά Μάθηση</u>	53
<u>3.3 Νευρωνικά Δίκτυα</u>	55
<u>3.3.1 Συναρτήσεις Ενεργοποίησης</u>	59
<u>3.3.2 Hyperparameter tuning</u>	61
<u>3.4 Transfer Learning</u>	64
<u>3.5 Feature Selection</u>	66
<u>3.6 Sequential Forward Selection</u>	69

3.7 Hierarchical Clustering	70
3.8 Spearman Correlation Coefficient	74
3.9 Pearson Correlation Coefficient	75
<u>Πρακτικό Μέρος</u>	
<u>Κεφάλαιο 4</u>	78
4.1 Κατασκευή Dataset για τοποθέτηση SMDs	78
4.2 Αλγόριθμος Βέλτιστης Τοποθέτησης SMDs	84
4.2.1 Αναλυτική Περιγραφή	88
4.2.2 Υλοποίηση	94
4.3 Υλοποίηση Νευρωνικών Δικτύων	96
4.3.1 Ρύθμιση Νευρωνικών Δικτύων	101
4.3.2 Υλοποίηση Transfer Learning	103
<u>Κεφάλαιο 5</u>	105
5.1 Δίκτυα Προσομοιώσεων	105
5.1.1 IEEE34	105
5.1.2 IEEE33	109
5.2 Αποτελέσματα Αναγνώρισης Τοπολογίας	112
5.3 Αποτελέσματα Εκτίμησης Κατάστασης	119
5.4 Αποτελέσματα Transfer Learning	135
<u>Επέκταση: Random Forest</u>	136
<u>Συμπεράσματα</u>	147
<u>Μελλοντικές Επεκτάσεις</u>	148
<u>Βιβλιογραφία</u>	150

Πίνακας Ακρονομίων	
ΣΗΕ	Συστήματα Ηλεκτρικής Ενέργειας
ΔΔ	Δίκτυα Διανομής
ΑΠΕ	Ανανεώσιμες Πηγές Ενέργειας
SE	State Estimation
DSSE	Distribution System State Estimation
TI	Topology identification
SFS	Sequential Forward Selection
HC	Hierarchical Clustering
RF	Random Forest
NN	Neural Networks
SMD	Synchrophasor Measurement Unit
PMU	Phasor Measurement Unit

Θεωρητικό Μέρος

Κεφάλαιο I

Συστήματα ηλεκτρικής ενέργειας

Ένα σύστημα ηλεκτρικής ενέργειας ορίζεται ως ένα σύνολο υλικών, εγκαταστάσεων, εξοπλισμού και μέσων – δικτύων - μεταφοράς που επιτρέπουν την παραγωγή, τη μεταφορά και τη διανομή ηλεκτρικής ενέργειας από τους σταθμούς παραγωγής στους καταναλωτές. Το πιο απλό παράδειγμα ενός συστήματος ενέργειας είναι το ηλεκτρικό δίκτυο που παρέχει ενέργεια σε οικιακά συγκροτήματα και στις βιομηχανίες.

Τα συστήματα ηλεκτρικής ενέργειας περιλαμβάνουν πολλά συνδεδεμένα μέρη που συνεργάζονται για να παρέχουν σταθερή, ασφαλή και αποτελεσματική ροή ηλεκτρικής ενέργειας. Συνήθως τα συστήματα ηλεκτρικής ενέργειας είναι πολύπλοκα δίκτυα που απαιτούν συνεχή επίβλεψη, συντήρηση και ανάπτυξη τους για να εξασφαλιστεί η αξιοπιστία και η αποδοτικότητα τους.

Ένα από τα βασικά στοιχεία ενός τέτοιου συστήματος είναι οι πάροχοι, οι οποίοι παράγουν ενέργεια, δηλαδή αποτελούν την πηγή ενέργειας του συστήματος. Η ισχύς συνεχούς ρεύματος συνήθως παρέχεται από μπαταρίες, κυψέλες καυσίμου ή και φωτοβολταϊκά στοιχεία, ενώ η ισχύς εναλλασσόμενου ρεύματος παρέχεται συνήθως από γεννήτριες (όπως π.χ. Ατμοστόβιλοι, αεροστρόβιλοι κλπ). Ένα ακόμη βασικό στοιχείο είναι τα φορτία που λαμβάνουν ενέργεια για να εκτελέσουν μια λειτουργία, τα οποία μπορεί να είναι από οικιακές συσκευές έως βιομηχανικά μηχανήματα. Τα περισσότερα φορτία απαιτούν για να λειτουργήσουν μια συγκεκριμένη τάση και, για συσκευές εναλλασσόμενου ρεύματος, μια συγκεκριμένη συχνότητα και αριθμό φάσεων. Οι συσκευές που βρίσκονται σε οικιακές εγκαταστάσεις, για παράδειγμα, θα είναι τυπικά μονοφασικές που λειτουργούν στα 50 Hz με τάση 230 βολτ (με τα εθνικά πρότυπα της Ελλάδας). Όλες οι ηλεκτρικές συσκευές έχουν επίσης ονομαστική ισχύ, η οποία καθορίζει την ποσότητα ισχύος που καταναλώνει η συσκευή. Οποιαδήποτε στιγμή, η καθαρή ποσότητα ισχύος που καταναλώνεται από τα φορτία σε ένα σύστημα ισχύος πρέπει να ισούται με την καθαρή ποσότητα ισχύος που παράγεται από τις παροχές μείον την ισχύ που χάνεται κατά τη μετάδοση. Η διασφάλιση ότι η τάση, η συχνότητα, η ποσότητα και η ποιότητα της ισχύος που παρέχεται στα φορτία είναι σύμφωνη με τις προσδοκίες και παραμένει σταθερή, είναι η μεγαλύτερη ίσως πρόκληση που αντιμετωπίζει ένα ΣΗΕ.

Γενικά τα συστήματα ηλεκτρικής ενέργειας χωρίζονται συνήθως σε τρία κύρια μέρη: την παραγωγή, τη μεταφορά και τη διανομή ηλεκτρικής ενέργειας.

Η παραγωγή ηλεκτρικής ενέργειας συνήθως λαμβάνει χώρα σε σταθμούς παραγωγής όπως για παράδειγμα τα εργοστάσια καύσης λιθάνθρακα. Οι παραδοσιακές πηγές ηλεκτρικής ενέργειας είναι λοιπόν τα παραδοσιακά καύσιμα όπως ο άνθρακας, το πετρέλαιο και το φυσικό αέριο τα οποία μέσω της παραγωγής ατμού από την καύση τους περιστρέφουν τον δρομέα των γεννητριών κι έτσι παράγουν ηλεκτρική ενέργεια.

Παράλληλα, υπάρχουν και οι ανανεώσιμες πηγές ενέργειας (ΑΠΕ), η αιολική, η ηλιακή, η υδροηλεκτρική, η γεωθερμική κλπ μέσω των οποίων

στους κατάλληλους σταθμούς (π.χ. φωτοβολταϊκά πάρκα), ουσιαστικά συλλέγεται ενέργεια χωρίς τη χρήση καυσίμων.

Η μεταφορά της ηλεκτρικής ενέργειας περιλαμβάνει το σύνολο των καλωδίων, των γραμμών μεταφοράς και των υποσταθμών που μεταφέρουν την ηλεκτρική ενέργεια από τους σταθμούς παραγωγής έως τους υποσταθμούς που τροφοδοτούν τα μεγάλα κέντρα κατανάλωσης, όπου και συνεχίζουν τη μεταφορά ενέργειας τα δίκτυα διανομής. Τα συστήματα μεταφοράς έχουν κατά κανόνα βροχοειδή διάταξη έτσι ώστε να υπάρχουν όσο το δυνατόν περισσότερες διαδρομές ροής ενέργειας και να εξυπηρετούνται αποδοτικότερα οι σκοποί της μεταφοράς.

Η ηλεκτρική ενέργεια παράγεται συνήθως σε υψηλές τάσεις στα σημεία παραγωγής έτσι ώστε να ελαχιστοποιηθούν οι απώλειες κατά τη μεταφορά, έτσι ώστε να μειωθεί το ρεύμα μέσω του νόμου του Ωm έτσι ώστε οι απώλειες που είναι ανάλογες του τετραγώνου του ρεύματος (νόμος του Τζάουλ), να μειωθούν κι αυτές αντίστοιχα. Στη συνέχεια οι γραμμές μεταφοράς υψηλής τάσης, καταλήγουν σε υποσταθμούς, οι οποίοι αποτελούν τους κόμβους του συστήματος και όπου η υψηλή τάση μετατρέπεται σε μέση η οποία μπορεί να χρησιμοποιηθεί από τα δίκτυα διανομής. Με αυτό τον τρόπο, σε αυτό το σημείο, η ενέργεια έχει μειωθεί σε ένα επίπεδο που είναι κατάλληλο για τη διανομή σε σπίτια, επιχειρήσεις και άλλες καταναλωτικές εγκαταστάσεις.. Οι υποσταθμοί λειτουργούν επίσης ως σημεία έλεγχου πριν η ενέργεια φτάσει στους τελικούς καταναλωτές.

Κατά τη μεταφορά, συστήματα ασφαλείας, έλεγχου, και τεχνολογίες παρακολούθησης χρησιμοποιούνται για τη διασφάλιση της σταθερής λειτουργίας του συστήματος. Η αποδοτικότητα του συστήματος είναι ουσιώδης για τη μείωση των απωλειών, με τεχνολογίες όπως η HVDC (υψηλής τάσης γραμμές μεταφοράς DC) η οποία αποτελεί μια προηγμένη μέθοδο μεταφοράς ενέργειας, καθώς επιτρέπει τη μεταφορά σε μεγάλες αποστάσεις με ελάχιστες απώλειες και η οποία είναι ιδιαίτερα χρήσιμη όταν η ενέργεια πρέπει να μεταφερθεί από απομακρυσμένες πηγές παραγωγής προς τις περιοχές κατανάλωσης.

Τα επίπεδα τάσης των συστημάτων μεταφοράς κυμαίνονται από 66 kV μέχρι 750 kV και χωρίζονται σε τρεις βαθμίδες. Οι τάσεις μέχρι 220 kV αποτελούν την βαθμίδα υψηλών τάσεων, από 275 kV μέχρι 500 kV τη βαθμίδα των υπερυψηλών τάσεων και όσες ξεπερνούν τα 500 kV ανήκουν στη βαθμίδα των εξαιρετικά υψηλών τάσεων.

Δίκτυα Διανομής

Τα δίκτυα διανομής, το τελικό στάδιο στη μεταφορά της ηλεκτρικής ενέργειας, ορίζονται ως το σύστημα που μεταφέρει την ηλεκτρική ενέργεια από τους υποσταθμούς μετατροπής της υψηλής τάσης και τη διανείμει σε καταναλωτές, επιχειρήσεις και βιομηχανίες. Σε αντίθεση με τα δίκτυα μεταφοράς που συνήθως είναι όμοια σε διεθνές επίπεδο, τα δίκτυα διανομής διαφέρουν από χώρα σε χώρα [7].

Στα σύγχρονα ΣΗΕ η αξία των εγκαταστάσεων διανομής αποτελεί περίπου το 30% του συνόλου των εγκαταστάσεων, ενώ η ποσοστιαία

συμμετοχή στις συνολικές δαπάνες εκμετάλλευσης, αν αφαιρεθούν οι δαπάνες καυσίμων), είναι πάνω από 35%.

Τα δίκτυα διανομής συνήθως αποτελούνται από μεγάλο πλήθος στοιχείων, σε σύγκριση με τα δίκτυα μεταφοράς, όπως υποσταθμοί, γραμμές μεταφοράς χαμηλής και μέσης τάσης, μετασχηματιστές και συστήματα ελέγχου και προστασίας, ενώ οι απώλειες ενέργειας στο επίπεδο της διανομής, κυρίως λόγω της χαμηλότερης τάσης, είναι σχεδόν διπλάσιες από αυτές του συστήματος μεταφοράς.

Οι υποσταθμοί διανομής συνδέονται με το σύστημα μεταφοράς και μειώνουν την τάση μετάδοσης σε μεσαία τάση που κυμαίνεται μεταξύ 2 kV και 33 kV με τη χρήση μετασχηματιστών. Οι πρωτεύουσες γραμμές διανομής μεταφέρουν αυτήν τη μεσαία τάση ενέργειας σε μετασχηματιστές διανομής που βρίσκονται κοντά στις εγκαταστάσεις του τελικού καταναλωτή. Οι μετασχηματιστές διανομής μειώνουν ξανά την τάση στο επίπεδο χρήσης που χρησιμοποιείται για φωτισμό, βιομηχανικό εξοπλισμό και οικιακές συσκευές. Συχνά, πολλοί πελάτες τροφοδοτούνται από έναν μετασχηματιστή μέσω δευτερευόντων γραμμών διανομής. Πελάτες που απαιτούν μεγαλύτερη ποσότητα ενέργειας μπορεί να συνδεθούν απευθείας στο πρωτεύον επίπεδο διανομής ή στο δεύτερο επίπεδο διανομής.

Έτσι τα δίκτυα διανομής μπορούν να χωριστούν σε μέσης τάσης, (ή πρωτεύουσας διανομής) και χαμηλής τάσης, (ή δευτερεύουσας διανομής). Τα δίκτυα μέσης (MT) τάσης λειτουργούν σε επίπεδα τάσης που κυμαίνονται συνήθως από 2 kV έως 33 kV. Αυτό το επίπεδο τάσης χρησιμοποιείται για τη μεταφορά ενέργειας από τους υποσταθμούς μεταφοράς προς τους υποσταθμούς διανομής. Συνήθως, ξεκινούν από ένα υποσταθμό υποβιβασμού υψηλής τάσης σε μέση ο οποίος τροφοδοτείται από γραμμές του συστήματος μεταφοράς.

Οι γραμμές διανομής μέσης τάσης μεταφέρουν τη μέση τάση από τους υποσταθμούς μέχρι τις περιοχές κατανάλωσης. Συνήθως, αυτές οι γραμμές διανομής διακλαδίζονται για να εξυπηρετούν διάφορες περιοχές. Το δίκτυο μέσης τάσης τροφοδοτεί τους υποσταθμούς διανομής MT/XT, οι οποίοι βρίσκονται κοντά στους τελικούς καταναλωτές, καθώς και τους καταναλωτές μέσης τάσης όπως για παράδειγμα τα νοσοκομεία και τις ελαφρύτερες βιομηχανίες.

Τα δίκτυα διανομής χαμηλής τάσης (XT) είναι αυτά που λαμβάνουν την ηλεκτρική ενέργεια από τους υποσταθμούς διανομής μέσης τάσης και τους καθιστούν διαθέσιμους για τους τελικούς καταναλωτές (οικιακούς και εμπορικούς). Οι καταναλωτές συνδέονται στο δίκτυο μέσω των service drops, ονομασία των καλωδιακών γραμμών που συνδέουν τον υποσταθμό διανομής με τους πρώτους.

Τα δίκτυα διανομής χαμηλής τάσης λειτουργούν σε τάσεις που συνήθως κυμαίνονται στα 230 V για μονοφασικά συστήματα και 400 V για τριφασικά συστήματα [36]. Στις χώρες της Ευρωπαϊκής Ένωσης, ανάμεσα τους και η Ελλάδα, τα δίκτυα αυτά είναι τριφασικά και λειτουργούν σε τάση 400 V μεταξύ των των φάσεων και 230 V μεταξύ μιας φάσης και του ουδετέρου. Οι τάσεις είναι γενικά χαμηλές καθώς τα δίκτυα αυτά διέρχονται μέσα από κατοικημένες περιοχές.

Δύο ακόμη κατηγορίες στις οποίες μπορούν να διακριθούν τα δίκτυα διανομής είναι σε εναέρια και υπόγεια, δηλαδή αν οι εγκαταστάσεις που τα συνοδεύουν βρίσκονται πάνω ή κάτω από την επιφάνεια της γης. Τα εναέρια

δίκτυα συνήθως έχουν μικρότερο κόστος κατασκευής και πιο εύκολη συντήρηση καθώς τα σφάλματα που προκύπτουν μπορούν να αποκατασταθούν ταχύτερα, αλλά είναι πολύ πιο ευάλωτα στις καιρικές συνθήκες με αποτέλεσμα να εμφανίζουν συχνές βλάβες. Τα υπόγεια δίκτυα είναι ανθεκτικά στις καιρικές συνθήκες και παρέχουν προστασία από πιθανούς κινδύνους όπως οι κεραυνοί αλλά η κατασκευή και η συντήρηση τους είναι αρκετά πιο απαιτητική οικονομικά.

Η επιλογή γίνεται συνήθως ανάλογα με τα γεωγραφικά χαρακτηριστικά της εξυπηρετούμενης περιοχής. Έτσι, στα αστικά κέντρα τα δίκτυα διανομής είναι σε μεγάλο βαθμό υπόγεια αφενός καθώς δεν υπάρχει ο απαιτούμενος χώρος ώστε να τηρούνται οι αποστάσεις ασφαλείας από τα κτίρια ενώ αντίθετα, στην ύπαιθρο που δεν υπάρχουν οι αντίστοιχοι περιορισμοί, τα δίκτυα είναι κυρίως εναέρια με καλωδιακές γραμμές να τοποθετούνται στύλους ή πύργους. Επομένως, τα χαρακτηριστικά δόμησης των πόλεων και γενικά η χωροταξική διαμόρφωση των επιμέρους περιοχών μιας χώρας καθορίζουν σε μεγάλο βαθμό την διαμόρφωση των δικτύων διανομής.

Εξοπλισμός Δικτυών Διανομής

Το βασικό στοιχείο των υποσταθμών υποβιβασμού της τάσης (YT/MT και MT/XT), είναι οι μετασχηματιστές [37], οι οποίοι και επιτελούν το έργο μείωσης της τάσης στα επίπεδα προς κατανάλωση. Στην Ελλάδα, οι συχνότεροι μετασχηματιστές υποβιβασμού YT/MT είναι οι 150/20kV και 150/33kV, και οι συχνότεροι μετασχηματιστές υποβιβασμού MT/XT είναι οι 20kV/400V και 33kV/400V.

Πολλοί μετασχηματιστές διανομής είναι εξοπλισμένοι με συστήματα ασφαλείας που αποτρέπουν τυχόν προβλήματα, κυριότερο από τα οποία είναι το Σύστημα Αλλαγής Τάσης Υπό Φορτίο (ΣΑΤΥΦ), το οποίο παρέχει την δυνατότητα της αλλαγής του λόγου του μετασχηματιστή κατά τη διάρκεια της λειτουργίας του, μέσω ενός συστήματος πολλαπλών λήψεων στο πρωτεύον κύκλωμα του μετασχηματιστή και έναν μηχανισμό κινητών επαφών στο δευτερεύον κύκλωμα. Αυτό επιτρέπει τη βηματική αλλαγή του λόγου μετασχηματισμού κάτω από φορτίο. Επομένως, όταν η τάση σε μια γραμμή ξεπερνά τα επιθυμητά όρια, το σύστημα αυτό με μία ή περισσότερες αλλαγές λήψεων μπορεί να επαναφέρει την τάση του μετασχηματιστή εντός των επιθυμητών ορίων, καθώς η αλλαγή της λήψης στο πρωτεύον τύλιγμα του μετασχηματιστή αλλάζει την τάση στο δευτερεύον, αφού η επαγόμενη τάση ανά σπείρα παραμένει σταθερή. Αυτή η δυνατότητα είναι ιδιαίτερα χρήσιμη για τη διατήρηση της σταθερότητας του δικτύου και την προστασία από πιθανούς κινδύνους υπερτάσεων.

Επίσης απαραίτητος μηχανισμός για την ασφαλή λειτουργία κάθε υποσταθμού είναι η γείωση, η οποία ανάλογα με τον σκοπό, χωρίζεται σε τρεις κύριες κατηγορίες:

- Γείωση Προστασίας (Protective Grounding): Αυτή η γείωση σχετίζεται με την προστασία από βραχυκυκλώματα και άλλες επικίνδυνες συνθήκες συνδέοντας όλα τα αγωγίμα μέρη του δικτύου με την γη. Κατά την περίπτωση ενός βραχυκυκλώματος ή μιας βλάβης, η γείωση προσφέρει μια εύθυμη διέξοδο για το ρεύμα, βοηθώντας στην αποτροπή επικίνδυνων καταστάσεων.

- Γείωση Ασφαλείας (Safety Grounding): Αυτή η γείωση σχετίζεται με την προστασία των ανθρώπων από ηλεκτρικά πεδία. Σε περιπτώσεις όπου υπάρχει η πιθανότητα εμφάνισης υψηλής τάσης υπό μη ασφαλείς συνθήκες, η γείωση ασφαλείας προστατεύει τους ανθρώπους από ηλεκτροπληξία καθώς μεταφέρει τα στατικά φορτία στη γη.
- Γείωση Λειτουργίας (Operational Grounding): Αυτή η γείωση συνδέεται με τη λειτουργία του εξοπλισμού και των συστημάτων. Αποτελεί τμήμα του κυκλώματος και αφορά τη γείωση του ουδέτερου κόμβου των μετασχηματιστών υποβιβασμού και την κατά διαστήματα γείωση του ουδέτερου κόμβου ενός δικτύου ΧΤ.

Επιπροσθέτως, για την διατήρηση της τάσης σε επιθυμητό επίπεδο στα διάφορα σημεία του δικτύου χρησιμοποιούνται συχνά ρυθμιστές τάσης. Υπάρχουν δύο κύριες κατηγορίες ρυθμιστών τάσης. Η πρώτη είναι οι αυτόματοι ρυθμιστές τάσης (AVR - Automatic Voltage Regulators), οι οποίοι λειτουργούν αυτόματα για τη διατήρηση σταθερής τάσης σε συγκεκριμένα σημεία του δικτύου και λειτουργούν προσαρμόζοντας με βάση μια προκαθορισμένη αναλογία το ποσοστό της τάσης προκειμένου να την διατηρήσουν εντός των επιθυμητών ορίων.

Η δεύτερη είναι τα συστήματα αυτόματης αντιστάθμισης (Compensation Systems) τα οποία περιλαμβάνουν τη χρήση πυκνωτών και αντισταθμιστικών συσκευών για τη διατήρηση της τάσης σε επιθυμητά επίπεδα. Τα συστήματα αυτά μπορούν να αλλάξουν την ισχύ που αντισταθμίζεται ανάλογα με τις απαιτήσεις του δικτύου και των φορτίων.

Στις γραμμές μεγάλου μήκους που χαρακτηρίζονται από υψηλή φόρτιση τοποθετούνται περισσότεροι από ένας ρυθμιστές τάσης. Στις περιπτώσεις αυτές το ΣΑΤΥΦ δε μπορεί να κρατήσει τις τάσεις εντός ορίων λόγω του μήκους των γραμμών και της υψηλής τους φόρτισης και μόνο η τοποθέτηση ρυθμιστών τάσης σε σειρά διορθώνει το πρόβλημα, ανυψώνοντας σταδιακά την τάση μέχρι το επιθυμητό επίπεδο.

Για την βελτίωση της απόδοσης και της αξιοπιστίας του δικτύου χρησιμοποιούνται και σε αυτή την περίπτωση πυκνωτές. Με αυτό τον τρόπο επιτυγχάνεται έλεγχος του συντελεστή ισχύος και ρυθμίζεται η τάση καθώς αντισταθμίζουν την άεργο ισχύ που δημιουργείται, χάρη στην επαγωγική αντίδραση των γραμμών. Οι διάφοροι τύποι πυκνωτών που χρησιμοποιούνται κατηγοριοποιούνται σε συνδεδεμένους σε σειρά (series capacitors) και συνδεδεμένους εγκάρσια (shunt capacitors), αλλά και ανάλογα με τον τρόπο ελέγχου τους.

Τέλος, στα δίκτυα διανομής τα λεγόμενα μέσα ζεύξεως και προστασίας αποτελούν σημαντικό κομμάτι του εξοπλισμού για τη διασφάλιση της αξιοπιστίας και της ομαλής λειτουργίας του ηλεκτρικού συστήματος, καθώς στοχεύουν στον εντοπισμό του σφάλματος, στην απομόνωσή του και στην επαναφορά του συστήματος στην κανονική του λειτουργία, με την αποκοπή όσο το δυνατόν λιγότερων καταναλωτών στην κάθε περίπτωση.

Τα μέσα ζεύξεως χρησιμοποιούνται για την ζεύξη - απόζευξη τμημάτων του δικτύου και αποτελούνται από τρεις κύριες κατηγορίες:

- Αποζεύκτες (Isolators ή Disconnect Switches): αποτελούν συσκευές παρόμοιες με τους διακόπτες και χρησιμοποιούνται για την ασφαλή αποσύνδεση μέρους του δικτύου σε περίπτωση έκτακτης ανάγκης,

κυρίως όταν δεν υπάρχει φορτίο, δηλαδή όταν το ρεύμα είναι αμελητέο και για λόγους συντήρησης.

- Οι διακόπτες φορτίου (Switches): είναι συσκευές που έχουν την ικανότητα να διακόπτουν, ή να αποκαθιστούν τη ροή του ηλεκτρικού ρεύματος στις κανονικές συνθήκες λειτουργίας ενός κυκλώματος, καθώς και να μειώνουν, αλλά όχι και να κόβουν, τις υψηλές τιμές των ρευμάτων βραχυκύκλωσης. Χωρίζονται σε διακόπτες απομόνωσης, οι οποίοι συνδέουν δύο τμήματα γραμμών της ίδιας γραμμής τροφοδοσίας και διακόπτες διασύνδεσης, οι οποίοι ενώ σε συνθήκες κανονικής λειτουργίας είναι ανοικτοί, όταν κλείσουν μπορούν να συνδέσουν δύο διαφορετικές γραμμές τροφοδοσίας δημιουργώντας κάποιο βρόχο.
- Οι διακόπτες ισχύος (βραχυκυκλώσεως) (Circuit Breakers): αποτελούν συσκευές που χρησιμοποιούνται για τον έλεγχο και τη διακοπή της ροής ρεύματος σε περίπτωση βλάβης ή ανάγκης αποσύνδεσης μέρους του δικτύου. Διαθέτουν την ικανότητα να διακόπτουν ή να αποκαθιστούν τη ροή του ρεύματος υπό κανονικές συνθήκες λειτουργίας αλλά και υπό μη κανονικές συνθήκες, όπως για παράδειγμα βραχυκυκλώματα. Ονομάζονται έτσι γιατί έχουν μεγαλύτερη ισχύ από τους διακόπτες φορτίου.

Παράλληλα, τα μέσα προστασίας στα δίκτυα διανομής χρησιμοποιούνται για την αυτόματη απομόνωση τμημάτων του δικτύου σε περίπτωση σφάλματος και στην πρόληψη βλαβών με στόχο την προστασία της ανθρώπινης ζωής και των εγκαταστάσεων. Τα πιο συνήθη μέσα προστασίας είναι:

- Οι ασφάλειες, οι οποίες διακόπτουν το κύκλωμα αν η τιμή του ρεύματος ξεπεράσει μια ορισμένη τιμή.
- Οι ηλεκτρονόμοι (relays), οι οποίοι είναι ηλεκτρικοί διακόπτες που ανοίγουν και κλείνουν ένα ηλεκτρικό κύκλωμα υπό τον έλεγχο ενός άλλου ηλεκτρικού κυκλώματος. Υπάρχουν πολλές κατηγορίες ηλεκτρονόμων όπως οι ηλεκτρονόμοι προστασίας υπερφόρτωσης (Overcurrent Protection Relays) που αντιλαμβάνονται υπερβολικά υψηλές τιμές ρεύματος στο δίκτυο και διακόπτουν το κύκλωμα για να αποτρέψουν ζημιές λόγω υπερφόρτωσης, οι ηλεκτρονόμοι προστασίας βραχυκυκλώματος (Short Circuit Protection Relays) που αντιδρούν σε βραχυκυκλώματα, οι ηλεκτρονόμοι προστασίας διακοπής ρεύματος (Loss of Voltage Protection Relays) που αναγνωρίζουν την απώλεια τάσης και κατόπιν διακόπτουν το κύκλωμα για την προστασία των φορτίων.
- Οι διακόπτες ισχύος που είναι τροποποιημένοι ώστε να δέχονται εντολές μέσω ηλεκτρονόμων προστασίας και να αποτελούν ένα ηλεκτρικό κύκλωμα τηλεχειρισμού. Επιλέγονται μέσω κατάλληλων διαγραμμάτων που καθορίζουν το χρόνο απόκρισης του διακόπτη ανάλογα με το ρεύμα που τον διαπερνά.

Πρακτικά, χρησιμοποιούνται σχεδόν πάντα συνδυασμοί των παραπάνω μέσων ζεύξεως και προστασίας, με τους πιο συνηθισμένους να είναι ο συνδυασμός αποζεύκτη και διακόπτη ισχύος και ο συνδυασμός διακόπτη

φορτίου και ασφάλειας, ώστε να υπάρχει η δυνατότητα χειρισμών και να εξασφαλίζεται ταυτόχρονα η προστασία από βραχυκυκλώματα.

Τοπολογία Δικτύων Διανομής

Η σχεδίαση των δικτύων διανομής στηρίζεται στη μονόδρομη ροή ισχύος από το σύστημα μεταφοράς προς τους καταναλωτές, γεγονός όμως που τα τελευταία χρόνια τείνει να αλλάξει λόγω της υψηλής διείσδυσης των ανανεώσιμων πηγών ενέργειας. Έτσι, τα δίκτυα διανομής λειτουργούν συνήθως ακτινικά, δηλαδή υπάρχει μόνο μια διαδρομή μέσω της οποίας γίνεται η ροή ισχύος, από τους υποσταθμούς στους κάθε καταναλωτές, κυρίως λόγω της απλότητας της, της αποτελεσματικότητας ως προς το κόστος και την ευκολία συντήρησης που παρουσιάζει.

Γενικότερα, η τοπολογία των δικτύων διανομής εξαρτάται από ποικίλες παραμέτρους μερικές από τις οποίες είναι το μήκος του επιθυμητού δικτύου, η πυκνότητα του φορτίου και κυρίως τα γεωγραφικά χαρακτηριστικά της περιοχής που εξυπηρετεί καθώς και το αν το δίκτυο είναι εναέριο ή υπόγειο. Η μοναδική ουσιαστικά μορφή τοπολογίας που χρησιμοποιείται τόσο σε εναέρια όσο και σε υπόγεια δίκτυα και ιδίως στα δίκτυα μέσης τάσης είναι η ακτινική [3].

Ένα ακτινικό δίκτυο διανομής αποτελείται από έναν κορμό, την κεντρική γραμμή του δικτύου, και τις διακλαδώσεις, οι οποίες καταλήγουν στους υποσταθμούς MT/XT και τους καταναλωτές μέσης ή χαμηλής τάσης. Ο κορμός τροφοδοτείται στο ένα άκρο του από τον υποσταθμό YT/MT, ελεγχόμενος από ένα διακόπτη ισχύος που εκτελεί αυτόματες επαναφορές και είναι κατασκευασμένος από αγωγό μεγαλύτερης διατομής από τους αντίστοιχους των διακλαδώσεων. Οι διακλαδώσεις ελέγχονται από ασφάλειες και αποζεύκτες.

Ωστόσο, παρ'όλο που οι ακτινικές τοπολογίες αποτελούν με διαφορά πιο κοινά απαντούμενη μορφή τοπολογίας, η ανάγκη για επανατροφοδότηση των υποσταθμών MT/XT και των καταναλωτών MT μετά από μια βλάβη και η σύγχρονη μεταβολή των δικτύων διανομής προς τα έξυπνα δίκτυα και η αμφίδρομη παραγωγή ενέργειας που αναφέρονται παρακάτω, οδήγησε στην ενσωμάτωση πιο σύνθετων μορφών τοπολογιών, ιδίως σε αστικές περιοχές ή σε κρίσιμες υποδομές όπου ο πλεονασμός και η αξιοπιστία είναι ζωτικής σημασίας [38]. Αυτές οι διαμορφώσεις χρησιμοποιούνται συνήθως σε συγκεκριμένα τμήματα του δικτύου διανομής, όπου τα οφέλη της αυξημένης αξιοπιστίας υπερτερούν της πολυπλοκότητας και του κόστους που συνδέονται με πιο περίπλοκες τοπολογίες.

Μία τέτοια μορφή τοπολογίας είναι η βροχοειδής. Στην περίπτωση αυτή, το δίκτυο συνδέεται σε δύο σημεία του ίδιου υποσταθμού YT/MT ή με δύο διαφορετικούς υποσταθμούς, η τροφοδοσία όμως δε γίνεται ταυτόχρονα αφού ο ένας διακόπτης φορτίου παραμένει πάντα ανοικτός. Ουσιαστικά με αυτό τον τρόπο η μορφή αυτή τοπολογίας λειτουργεί ακτινικά. Αυτού του είδους η τοπολογία συναντάται κυρίως στα υπόγεια δίκτυα μέσης τάσης και βοηθάει κυρίως στην ευκολότερη προστασία του δικτύου και στην εξοικονόμηση γραμμών. Επίσης, είναι σχετικά απλή, αλλά ως κύριο μειονέκτημα έχει πως δεν αξιοποιούνται μέσω αυτής πλήρως τα καλώδια των γραμμών, καθώς η αναχώρηση κάθε βρόχου μπορεί να φορτίζεται μόνο μέχρι

το 50% της τιμής του καλωδίου. Αυτό συμβαίνει έτσι ώστε σε περίπτωση βλάβης του πρώτου τμήματος να μπορεί να τροφοδοτηθεί το υπόλοιπο.

Ένα αραχνοειδές δίκτυο, είναι παρόμοιο και στην ουσία είναι και αυτό βροχοειδές την εξής όμως διαφορά: οι γραμμές είναι πλέον τρεις και υπάρχουν και πλευρικές συνδέσεις. Έτσι, στην περίπτωση βλάβης σε μια από τις τρεις γραμμές, οι άλλες δύο καλύπτουν τις ανάγκες των φορτίων μέσω χειρισμών των διακοπών φορτίου.

Μερικές ακόμη κατηγορίες είναι τα ατρακτοειδή δίκτυα στα οποία οι γραμμές των υποσταθμών ΥΤ/ΜΤ καταλήγουν σε άλλους υποσταθμούς ΥΤ/ΜΤ ή ζεύξης, και τα δίκτυα διπλής τροφοδότησης τα οποία αποτελούνται από παράλληλες ακτινικές γραμμές έτσι ώστε σε κάθε υποσταθμό να καταλήγει και μια εφεδρική γραμμή τροφοδοσίας. Με αυτή τη μορφή τοπολογίας, γίνεται καλύτερη αξιοποίηση των καλωδίων αλλά δημιουργούνται υποσταθμοί ζεύξεως οι οποίοι αυξάνουν περισσότερο την πολυπλοκότητα και γενικά επιβαρύνουν το δίκτυο.

Επιπλέον ειδικότερες μορφές τοπολογιών που χρησιμοποιούνται κυρίως σε δίκτυα χαμηλής τάσης είναι η δακτυλιοειδής στο οποίο οι γραμμές που ξεκινούν από έναν υποσταθμό αποτελούν τμήματα δακτυλίων έτσι ώστε να αυξάνεται η εφέδρεια και να μειώνεται η πτώση τάσης, επιτρέποντας τη τροφοδότηση των καταναλωτών από διάφορες κατευθύνσεις. Σε περίπτωση αστοχίας ή βλάβης σε ένα σημείο του δακτυλίου, η ηλεκτρική ενέργεια μπορεί να εξακολουθήσει να ρέει προς την αντίθετη κατεύθυνση, διατηρώντας έτσι τη συνέχεια της παροχής στα συνδεδεμένα φορτία. Το πλεονεκτήματα αυτής της τοπολογίας είναι ο πλεονασμός, η υψηλή αξιοπιστία και η αποδοτικότητα των δεδομένων. Τα μειονεκτήματα είναι η πολυπλοκότητα και οι προκλήσεις συντήρησης. Αυτή η μορφή τοπολογίας εφαρμόζεται συνήθως σε βιομηχανικές εγκαταστάσεις, εταιρικά δίκτυα, συστήματα μεταφορών και τοπικές κρίσιμες υποδομές. Το κόστος μιας τοπολογίας δακτυλίου είναι γενικά μεσαίο σε σύγκριση με την ακτινική.

Τον μεγαλύτερο βαθμό εξυπηρέτησης παρουσιάζει το αραχνοειδές δίκτυο το οποίο όμως λόγω αυξημένης πολυπλοκότητας έχει μεγαλύτερο κόστος κατασκευής. Η δακτυλοειδής μορφή τοπολογίας ουσιαστικά έχει τα ίδια πλεονεκτήματα σε μικρότερο βαθμό αλλά και με μικρότερο κόστος.

Γενικά, η μορφή που έχει καθιερωθεί τα τελευταία χρόνια και χρησιμοποιείται ολοένα και περισσότερο και στην Ελλάδα, συνδυάζει τη βροχοειδή διάταξη, (μερικές φορές κι αυτή σε συνδυασμό με την ατρακτοειδή), με την ακτινική λειτουργία του δικτύου.

Μία τοπολογία χαρακτηρίζεται εφικτή όταν ικανοποιεί συγκεκριμένες λειτουργίες και απαιτήσεις όπως η αξιόπιστη παροχή ηλεκτρισμού στους καταναλωτές, η ισορροπία των φορτίων (δηλαδή ή ίση διανομή τους στα διάφορα τμήματα του δικτύου), η ανοχή ή/και η αποκατάσταση βλαβών, αποτελεσματική ρύθμιση τάσης, η ευελιξία στις αλλαγές στη ζήτηση της ενέργειας στις απαιτήσεις συντήρησης.

Τα Δίκτυα Διανομής Σήμερα

Τα δίκτυα διανομής σήμερα [35] λόγω των μεγάλων εξελίξεων γενικότερα στον τομέα των ΣΗΕ, γίνονται πιο προσαρμοστικά και πιο ικανά να ανταποκριθούν στα πολύπλοκα προβλήματα που εγείρουν οι σύγχρονες ενεργειακές απαιτήσεις. Η συνεχής έρευνα και οι εξελίξεις, ακόμη και σε

τομείς όπως η τεχνητή νοημοσύνη, επηρεάζουν και διαμορφώνουν το μέλλον των δικτύων διανομής με τρόπο απρόβλεπτο πλέον, αναδιαμορφώνοντας τα συνεχώς και προσθέτοντας τους νέες ιδιότητες.

Επιπρόσθετα, τα σύγχρονα δίκτυα εξελίσσονται σε έξυπνα δίκτυα, ενσωματώνοντας προηγμένες τεχνολογίες επικοινωνίας και ελέγχου. Η αμφίδρομη επικοινωνία διευκολύνει τον συντονισμό σε πραγματικό χρόνο μεταξύ των επιχειρήσεων κοινής ωφελείας και των καταναλωτών, ενώ η χρήση αισθητήρων και συσκευών παρακολούθησης παρέχει λεπτομερείς πληροφορίες για την απόδοση του δικτύου, επιτρέποντας βελτιστοποίηση της λειτουργίας και ταχεία απόκριση σε πάσης φύσεως προβλήματα.

Η ενσωμάτωση των ΑΠΕ, όπως η ηλιακή και η αιολική ενέργεια, με την αστάθεια και την τυχειότητα που τις χαρακτηρίζει απαιτεί χρήση εξελιγμένων μηχανισμών ελέγχου. Οι αναπτυσσόμενες τεχνικές αποθήκευσης ενέργειας και οι τεχνικές εξισορρόπησης των φορτίων του δικτύου είναι κομβικής σημασίας στη σταθεροποίηση του δικτύου παρά τις διακυμάνσεις της παραγωγής ηλεκτρικής ενέργειας με ανανεώσιμες πηγές.

Παράλληλα, οι κατανεμημένοι ενεργειακοί πόροι στους οποίους συμπεριλαμβάνονται π.χ. οι ηλιακοί συλλέκτες στις πολυκατοικίες και οι ανεμογεννήτριες μικρής κλίμακας συμβάλλουν σε ένα αποκεντρωμένο ενεργειακό τοπίο, το οποίο ενισχύει την ανθεκτικότητα του αφού μειώνεται η εξάρτηση από κεντρικούς σταθμούς ηλεκτροπαραγωγής.

Έτσι, δημιουργούνται μικροδίκτυα τα οποία λειτουργούν ως τοπικά, μικρότερης κλίμακας ενεργειακά συστήματα που μπορούν να είναι ανεξάρτητα ή να συνδέονται με το κύριο δίκτυο και τα οποία προσφέρουν στις κοινότητες μεγαλύτερο έλεγχο του ενεργειακού εφοδιασμού τους, ειδικότερα κατά τη διάρκεια διακοπών του δικτύου. Επίσης, οι νέες λύσεις αποθήκευσης ενέργειας, ειδικά λόγω της διακύμανσης των συνεχώς αυξανόμενων ανανεώσιμων τρόπων παραγωγής, βασισμένων κυρίως σε τεχνολογίες μπαταριών, αποθηκεύουν περίσσεια ενέργειας σε περιόδους χαμηλής ζήτησης. Αυτή η αποθηκευμένη ενέργεια μπορεί να απελευθερωθεί κατά τη διάρκεια της ζήτησης αιχμής, συμβάλλοντας στη σταθερότητα και την αξιοπιστία του δικτύου.

Επιπλέον ο αυτοματισμός δικτύου μέσω εκτεταμένων δικτύων αισθητήρων για παρακολούθηση σε πραγματικό χρόνο έτσι ώστε να ανιχνεύονται σφάλματα και να καταγράφονται οι συνθήκες φορτίου, ενισχύει τις ικανότητες απορρόφησης και αποκατάστασης βλαβών χωρίς ανθρώπινη επίβλεψη, ενώ η τάση προς την ηλεκτροδότηση των μεταφορών (όπως π.χ. ηλεκτρικά αυτοκίνητα και τρένα υψηλών ταχυτήτων) εισάγει νέες απαιτήσεις στο σύστημα διανομής, απαιτώντας αναβαθμίσεις της υποδομής για την υποστήριξη αυξημένων απαιτήσεων κοστολόγησης.

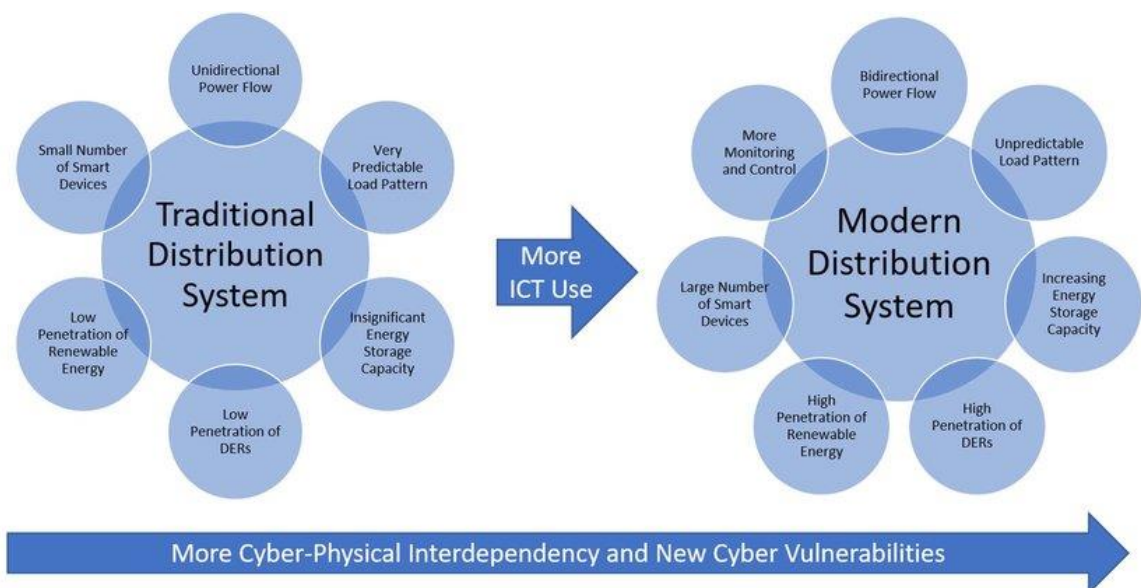
Τέλος, οι συνεχιζόμενες προσπάθειες επικεντρώνονται στην ελαχιστοποίηση των απωλειών ενέργειας κατά τη μεταφορά και τη διανομή, μέσω βελτιώσεων στον σχεδιασμό του δικτύου, στην αποδοτικότητα του εξοπλισμού και στις λειτουργικές πρακτικές έτσι ώστε να υπάρχει μέγιστη εξοικονόμηση ενέργειας.

Παραδοσιακά, τα δικτυακά συστήματα λειτουργούσαν ως απλές γραμμές διανομής, όπου η ηλεκτρική ενέργεια από τα δίκτυα μεταφοράς θα μοιραζόταν μεταξύ των καταναλωτών. Σήμερα, τα δίκτυα διανομής ενσωματώνουν στενά τις ΑΠΕ μέσω των πόρων κατανεμημένης παραγωγής, όπως η ηλιακή και η αιολική ενέργεια. Ως αποτέλεσμα, τα δίκτυα διανομής

γίνονται ολοένα και περισσότερο ανεξάρτητα από τα δίκτυα μεταφοράς. Η ισορροπία στη σχέση προσφοράς-ζήτησης σε αυτά τα μοντέρνα δίκτυα διανομής είναι εξαιρετικά προκλητική και απαιτεί τη χρήση διάφορων τεχνολογικών και λειτουργικών μέσων για τη λειτουργία τους. Επιπλέον, η ανάπτυξη των έξυπνων δικτύων επιτρέπει τη δημιουργία διακριτών περιοχών ενέργειας, όπου μπορούν να ενσωματωθούν τα τοπικά συστήματα αποθήκευσης και παραγωγής, προσφέροντας ταυτόχρονα ευκαιρίες για αυτοτελή λειτουργία και εκμετάλλευση των ανανεώσιμων πηγών ενέργειας. Οι προσπάθειες αυτές ανταποκρίνονται στην ανάγκη για μια πιο βιώσιμη και έξυπνη παραγωγή και κατανομή ενέργειας, προωθώντας τη μετάβαση προς ένα πιο αποδοτικό ενεργειακό μοντέλο.

Σταδιακά, τα δίκτυα διανομής μετασχηματίζονται σε έξυπνα δίκτυα, δηλαδή προηγμένα συστήματα που ενσωματώνουν ψηφιακές τεχνολογίες, αμφίδρομα συστήματα επικοινωνίας και τεχνικές αυτοματισμού για τη βελτιστοποίηση της παραγωγής, της διανομής και της κατανάλωσης ηλεκτρικής ενέργειας. Αυτή η μετάβαση, επιτρέπει την καλύτερη διαχείριση των ενεργειακών πόρων, ενισχύει την αξιοπιστία, προωθεί την αποδοτικότητα και υποστηρίζει την ενσωμάτωση ανανεώσιμων πηγών ενέργειας.

Στο παρακάτω σχήμα απεικονίζονται οι αλλαγές που επιτελούνται σήμερα στα δίκτυα διανομής [9] :



Όπου, με τον όρο ICT Use εννοείται η αυξανόμενη χρήση τεχνολογιών απόκτησης και μετάδοσης πληροφορίας, και DERs οι καταναλωμένοι πόροι παραγωγής (distributed energy resources).

Ροή φορτίου

Ροή φορτίου σε ένα ηλεκτρικό σύστημα ονομάζεται η μεταφορά της ηλεκτρικής ενέργειας από τις πηγές παραγωγής προς τους καταναλωτές και τα υπόλοιπα φορτία. Είναι δηλαδή η κίνηση της ηλεκτρικής ενέργειας μέσα από το δίκτυο μεταφοράς και διανομής, καθορίζοντας πώς η ενέργεια μεταφέρεται και διανέμεται σε όλο το σύστημα.

Η ροή φορτίου ελέγχεται και διαχειρίζεται προκειμένου να εξασφαλιστεί ότι η ηλεκτρική ενέργεια φτάνει στους καταναλωτές με τρόπο αποτελεσματικό, αξιόπιστο και ασφαλές. Σε μεγάλα ηλεκτρικά δίκτυα, όπως τα δίκτυα διανομής, η ροή φορτίου ελέγχεται με τη χρήση διάφορων μέσων, συμπεριλαμβανομένης της ρύθμισης τάσης, των μετασχηματιστών, και άλλων συστημάτων ελέγχου.

Η ανάλυση της ροής φορτίου είναι μια αριθμητική επίλυση της ροής ηλεκτρικής ενέργειας σε ένα συνδεδεμένο σύστημα. Στην πράξη, περιλαμβάνει την εκτίμηση των τάσεων, των ρευμάτων, της άεργου ισχύος που ανταλλάσσεται και άλλων παραμέτρων στο δίκτυο για διάφορες περιπτώσεις φορτίου. Συνήθως κατά την διαδικασία αυτή χρησιμοποιούνται τεχνικές απλοποίησης όπως τα διαγράμματα μονής γραμμής (που οι τρεις αναπαρίσταται με μία γραμμή για παράδειγμα), και το σύστημα ανά μονάδα. Τα ηλεκτρικά συστήματα αναλύονται κατά την κανονική και σταθερή λειτουργία τους και συνήθως η κύρια πληροφορία που αναζητείται είναι το πλάτος και η γωνία της τάσης σε κάθε σταθμό, καθώς και η πραγματική και άεργος ισχύς που ρέουν σε κάθε γραμμή.

Η ανάλυση της ροής φορτίου δείχνει την ικανότητα του συστήματος να τροφοδοτεί επαρκώς το συνδεδεμένο φορτίο. Με αυτό τον τρόπο, είναι βασική για το σχεδιασμό νέων συστημάτων ηλεκτρικής ενέργειας, καθώς βοηθάει στην εκτίμηση των απαιτήσεων για τις νέες γραμμές μεταφοράς, τους νέους υποσταθμούς και γενικά τον νέο εξοπλισμό. Επίσης η ανάλυση ροής φορτίου βοηθάει στη βελτιστοποίηση της λειτουργίας του συστήματος υπό κανονικές συνθήκες λειτουργίας, καθώς μέσω αυτής μπορούν να βρεθούν οι βέλτιστες ρυθμίσεις και να αποφευχθούν υπερφορτώσεις. Επιπλέον, για αντίστοιχους λόγους η κατανόηση της ροής φορτίου είναι σημαντική για τον σχεδιασμό συστημάτων προστασίας και η παρακολούθηση της επιτρέπει την άμεση ανίχνευση προβλημάτων στο σύστημα.

Τα σύγχρονα ενεργειακά συστήματα και ιδίως τα δίκτυα διανομής που αποτελούν την πιο σύνθετη υποκατηγορία τους, είναι συνήθως ιδιαίτερα πολύπλοκα και καθιστούν αδύνατη τη χειροκίνητη επίλυση της ροής ισχύος. Έτσι, συχνά χρησιμοποιούνται υπολογιστικά συστήματα για να παρέχουν αριθμητικές επιλύσεις του προβλήματος που βρίσκονται εντός αποδεκτών ορίων.

Ειδικότερα, το πρόβλημα έγκειται στο γεγονός πως πρέπει να επιλυθεί ένα μη γραμμικό σύστημα εξισώσεων που περιγράφει τη ροή ενέργειας μέσω κάθε γραμμής μεταφοράς. Ο στόχος είναι να βρεθούν πλήρως οι γωνίες και τα μέτρα των τάσεων για κάθε ζυγό του συστήματος υπό καθορισμένο φορτίο και συνθήκες πραγματικής ισχύος και τάσης της γεννήτριας. Κατόπιν, μέσω των τάσεων, μπορεί να προσδιοριστεί αναλυτικά η ροή πραγματικής και άεργου ισχύος σε κάθε κλάδο καθώς και η έξοδος άεργου ισχύος της γεννήτριας.

Το πρόβλημα είναι μη γραμμικό επειδή η ροή ισχύος προς τις αντιστάσεις φορτίου είναι συνάρτηση του τετραγώνου των τάσεων. Λόγω αυτής της μη γραμμικότητας, σε πολλές περιπτώσεις η ανάλυση μεγάλων δικτύων μέσω του μοντέλου ροής εναλλασσόμενου ρεύματος (AC) δεν είναι εφικτή, και χρησιμοποιείται αντ' αυτού ένα γραμμικό (αλλά λιγότερο ακριβές) μοντέλο ροής συνεχούς ρεύματος (DC). Μια περαιτέρω απλοποίηση είναι η υπόθεση σε ένα τριφασικό σύστημα εναλλασσόμενου ρεύματος πως έχει ισορροπημένο φορτίο σε όλες τις τρεις φάσεις και πως βρίσκεται σε

ημιτονοειδή λειτουργία σταθερής κατάστασης, χωρίς παροδικές αλλαγές στη ροή ισχύος ή στην τάση λόγω μεταβολών φορτίου ή παραγωγής. Δηλαδή, θεωρείται πως όλες οι κυματομορφές ρεύματος και τάσης είναι ημιτονικές συναρτήσεις του χρόνου χωρίς αντιστάθμιση συνεχούς ρεύματος και έχουν την ίδια σταθερή συχνότητα.

Η επίλυση πραγματοποιείται με τον εντοπισμό των γνωστών και των άγνωστων μεταβλητών στο σύστημα ανάλογα με τον τύπο των ζυγών που υπάρχουν. Οι συνηθισμένοι ζυγοί χωρίς γεννήτριες συνδεδεμένες σε αυτούς ονομάζεται ζυγοί φορτίου (Load Bus). Αν ένας ζυγός έχει τουλάχιστον μία γεννήτρια συνδεδεμένη σε αυτόν ονομάζεται ζυγός γεννήτριας, (Generator Bus). Τέλος, αυθαίρετα επιλέγεται ένας ζυγός που έχει γεννήτρια ως ζυγός αναφοράς (Slack Bus). Στους ζυγούς φορτίου υποθέτουμε πως είναι γνωστές η πραγματική και η άεργος ισχύς, στους ζυγούς γεννήτριας θεωρείται γνωστή η πραγματική ισχύς που παράγεται και το πλάτος της τάσης, ενώ στον ζυγό αναφοράς συνήθως θεωρείται πως το πλάτος της τάσης είναι 1 ανά μονάδα και τη γωνία της είναι 0 μοίρες. Έτσι ψάχνουμε για κάθε ζυγό φορτίου το πλάτος και τη γωνία της τάσης και για κάθε ζυγό γεννήτριας μόνο την γωνία.

Στη συνέχεια καταστρώνεται ένα σύστημα εξισώσεων με βάση τις εξισώσεις του ισοζυγίου πραγματικής και του ισοζυγίου αέργου ισχύος για κάθε ζυγό φορτίου και το ισοζύγιο πραγματικής ισχύος για κάθε ζυγό γεννήτριας, (η άεργος ισχύς που εισέρχεται στο δίκτυο από αυτούς τους ζυγούς θεωρείται άγνωστη και δεν την συμπεριλαμβάνουμε για να μην εισαχθούν περισσότερες άγνωστες μεταβλητές στο σύστημα).

Σε πολλά συστήματα μεταφοράς, η σύνθετη αντίσταση των γραμμών του δικτύου είναι κυρίως επαγωγική, δηλαδή οι γωνίες φάσης της σύνθετης αντίστασης των γραμμών είναι συνήθως σχετικά μεγάλες και αρκετά κοντά στις 90 μοίρες. Επομένως, καθώς υπάρχει έτσι μεγάλη συσχέτιση μεταξύ πραγματικής ισχύος και γωνίας τάσης, και μεταξύ άεργου ισχύος και μεγέθους τάσης, η πραγματική ισχύς τείνει να μεταδίδεται από τον ζυγό με την μεγαλύτερη γωνία τάσης στο δίαυλο με την μικρότερη γωνία τάσης και η άεργος ισχύς συνήθως μεταδίδεται από τον ζυγό με μεγαλύτερο πλάτος τάσης στον δίαυλο με το μικρότερο πλάτος τάσης. Ωστόσο, αυτή η προσέγγιση δεν ισχύει όταν η γωνία φάσης της σύνθετης αντίστασης των γραμμών είναι σχετικά μικρή.

Η πιο συχνά χρησιμοποιούμενη αριθμητική μέθοδος του προβλήματος ροής φορτίου, η οποία χρησιμοποιήθηκε και σε αυτή την εργασία, είναι μέθοδος Newton-Raphson. Πρόκειται για μία επαναληπτική μέθοδο η οποία μέσω διαδοχικών προσεγγίσεων βρίσκει προσεγγιστικά τις ρίζες μιας πραγματικής συνάρτησης. Ο επαναληπτικός τύπος της μεθόδου είναι:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Έτσι, η διαδικασία με βάση τον παραπάνω τύπο συνεχίζει επαναληπτικά προσαρμόζοντας την προσέγγιση με βάση τον λόγο της τιμής της συνάρτησης προς την κλίση της (παράγωγο) στο τρέχον σημείο, έως ότου η προσέγγιση συγκλίνει σε μια λύση, όπου η τιμή της συνάρτησης πέφτει κάτω από μηδέν ή κάτω από ένα προκαθορισμένο όριο.

Στο πλαίσιο της ροής φορτίου, αρχικά γίνεται μια εικασία για τα άγνωστα μεγέθη (τις τάσεις) αντιστοιχίζοντας τους τυχαίες τιμές. Συνήθως

πραγματοποιείται μια επίπεδη εκκίνηση, δηλαδή όλες οι γωνίες τάσης ρυθμίζονται στο μηδέν κι όλα τα πλάτη στο ένα ανά μονάδα. Στην συνέχεια καταστρώνονται κι επιλύονται οι εξισώσεις των ισοζυγίων ισχύος, χρησιμοποιώντας κάθε φορά τις πιο πρόσφατες τιμές. Κατόπιν, το σύστημα των εξισώσεων γραμμικοποιείται με βάση ξανά τις πιο πρόσφατες τιμές του πλάτους και της γωνίας των τάσεων και λύνεται για να βρεθούν οι νέες τιμές γωνίας και πλάτους. Τέλος, ενημερώνονται οι τιμές του πλάτους και της γωνίας των τάσεων και αν οι λύσεις είναι αποδεκτές σύμφωνα με τα προκαθορισμένα όρια, ο αλγόριθμος σταματά, αλλιώς συνεχίζει τις επαναλήψεις.

Υπάρχουν κι άλλες μέθοδοι αριθμητικών λύσεων της ροής φορτίου όπως η μέθοδος Gauss-Seidel, η Fast-decoupled-load-flow method και η Holomorphic embedding load flow method αλλά η Newton-Raphson είναι η πιο διαδεδομένη κυρίως λόγω της πολύ γρήγορης σύγκλισης της καθώς οι αρχικές τιμές που ανατέθηκαν τυχαία είναι συχνά στις πραγματικές. Επίσης η μέθοδος αυτή συγκλίνει τετραγωνικά, δηλαδή σε κάθε επανάληψη, ο αριθμός των σωστών ψηφίων στη λύση διπλασιάζεται περίπου, γεγονός που κάνει την σύγκλιση ιδιαίτερα αποτελεσματική. Η γρήγορη σύγκλιση είναι ουσιώδης για τα συστήματα ηλεκτρικής ενέργειας καθώς συνήθως είναι πολύ πολύπλοκα και χρειάζεται να βρεθούν γρήγορες λύσεις για την εποπτεία και τον έλεγχο τους σε πραγματικό χρόνο.

Επίσης, ένας ακόμα λόγος που προτιμάται η μέθοδος Newton-Raphson είναι η προσαρμοστικότητα της, καθώς μπορεί να χειριστεί αποτελεσματικά ενημερώσεις στα μοντέλα συστημάτων, ενσωματώνοντας νέα στοιχεία ή αλλαγές στις συνθήκες λειτουργίας.

Το κύριο μειονέκτημα της είναι πως ενδέχεται να μην συγκλίνει ή να συγκλίνει σε λάθος ρίζες εάν οι αρχικές εκτιμήσεις απέχουν πολύ από την πραγματική λύση ή σε περιπτώσεις όπου οι συνθήκες του συστήματος οδηγούν σε δυσκολίες σύγκλισης (π.χ. νησιδοποίηση). Επίσης, γενικά όπως γίνεται κατανοητό από αυτά που προαναφέρθηκαν, η επιτυχία της μεθόδου εξαρτάται πολύ από τις τυχαίες αρχικοποιημένες τιμές κι έτσι δεν υπάρχει εγγύηση πως η μέθοδος θα συγκλίνει στις πραγματικές λύσεις κι όχι σε κάποιο τοπικό ελάχιστο. Τέλος, η μέθοδος Newton-Raphson μπορεί να αποβεί υπολογιστικά δύσκολη καθώς γίνεται χρήση παραγώγων, ο υπολογισμός των οποίων σε πολύπλοκα συστήματα με πλήθος ζυγών μπορεί να αποβεί χρονοβόρος.

Synchrophasor Measurement Devices – SMDs

Μία Μονάδα Μέτρησης Φασιθέτη, καλούμενη με το αγγλικό ακρωνύμιο PMU ή synchrophasor, (Συγχρονισμένος φασιθέτης), είναι μια συσκευή που μετρά φασιθέτες. Ένας φασιθέτης δείχνει το μέγεθος και τη γωνία φάσης για την εναλλασσόμενη τάση ή ρεύμα σε μια συγκεκριμένη θέση σε μια γραμμή ηλεκτρικού ρεύματος. Αυτή η πληροφορία μπορεί επίσης να χρησιμοποιηθεί για τον προσδιορισμό της συχνότητας και για τον προσδιορισμό και την ανάλυση των συνθηκών που επικρατούν στο ΣΗΕ. Τα PMUs μπορούν να παράσχουν έως και 60 μετρήσεις ανά δευτερόλεπτο, πολύ περισσότερες από τις τυπικές μετρήσεις κάθε 2 έως 4 δευτερολέπτων που παρέχουν τα συμβατικά συστήματα SCADA, (Supervisory Control and Data

Acquisition – ένα σύστημα ελέγχου που συνδυάζει hardware με software παρέχοντας υψηλού επιπέδου επίβλεψη μηχανών και διαδικασιών). Ένα PMU μπορεί να λειτουργεί ανεξάρτητα ή να είναι ενσωματωμένο σε κάποιο ρελέ προστασίας ή κάποια άλλη συσκευή.

Οι μετρήσεις φάσεων από τα PMU's κατασκευάζονται από συνημιτονοειδή κύματα, που ακολουθούν την δομή:

$$A \cos(\omega t + \theta)$$

Το A σε αυτή τη συνάρτηση είναι μια βαθμωτή τιμή η οποία αντιπροσωπεύει το πλάτος της τάσης ή του ρεύματος. Το θ είναι η μετατόπιση της γωνίας φάσης από κάποια καθορισμένη αρχική θέση και το ω είναι η γωνιακή συχνότητα της κυματομορφής (συνήθως $2\pi/50$ ακτίνια/δευτερόλεπτο ή $2\pi/60$ ακτίνια/δευτερόλεπτο). Συνήθως, τα PMU μετρούν μόνο το πλάτος της τάσης και τη γωνία φάσης και υποθέτουν ότι η γωνιακή συχνότητα είναι σταθερή, κι έτσι την αγνοούν στη μέτρηση της φάσης. Όταν η κυματομορφή είναι μη ημιτονοειδής, όπως συμβαίνει με την τάση π.χ. του δικτύου κατά τη διάρκεια μιας πτώσης τάσης ή σφάλματος η μονάδα PMU δεν μπορεί να την προσαρμόσει ακριβώς στη συνημιτονική καμπύλη με αποτέλεσμα όσο λιγότερο ημιτονοειδής είναι η κυματομορφή, τόσο λιγότερο ακριβής να γίνεται η αναπαράσταση της φάσης.

Το μεγάλο τους πλεονέκτημα έναντι των παραδοσιακών μεθόδων συλλογής δεδομένων είναι πως όλα τα δεδομένα των PMUs φέρουν χρονική σήμανση χρησιμοποιώντας δεδομένα του Παγκόσμιου Συστήματος Εντοπισμού (GPS). Μπορούν να λάβουν επίσης πολλαπλές πηγές χρόνου, συμπεριλαμβανομένων αναφορών εκτός GPS, εφόσον είναι όλες βαθμονομημένες και λειτουργούν συγχρονισμένα. Δηλαδή, τα δεδομένα που συλλέγονται σε ένα πλέγμα συγχρονίζονται χρησιμοποιώντας την ίδια ακριβώς μέθοδο συσχέτισης του χρόνου με τα δεδομένα. Εξού και τα PMUs ονομάζονται και synchrophasors. Στην πραγματικότητα οι όροι δεν είναι ταυτόσημοι καθώς synchrophasor καλείται η μετρούμενη τιμή, (μετρήσεις φασιθετών που πραγματοποιούνται ταυτόχρονα σε οποιαδήποτε απόσταση), ενώ το PMU είναι η συσκευή μέτρησης.

Τα PMUs παρέχουν μια λεπτομερή και ακριβή εικόνα της ποιότητας της ισχύος σε ένα ευρύ γεωγραφικό πλέγμα. Τα συλλεγόμενα δεδομένα δίνουν τις απαραίτητες πληροφορίες στον διαχειριστή του συστήματος έτσι ώστε να καταλαβαίνει εάν η τάση, το ρεύμα και η συχνότητα διατηρούνται εντός καθορισμένων ορίων. Η δυνατότητα αυτή μπορεί να χρησιμοποιηθεί για:

- Για τη βελτίωση της ακρίβειας του συστήματος μοντελοποίησης.
- Για την ανίχνευση και πρόβλεψη της πιθανής αστάθειας του συστήματος.
- Για την παροχή πληροφοριών για ανάλυση συμβάντος σε περίπτωση που παρατηρηθεί μια διαταραχή.
- Για τον εντοπισμό ανεπαρκειών του συστήματος.
- Για την πρόβλεψη και διαχείριση συμφόρησης των γραμμών.
- Για την εποπτεία μιας περιοχής μέσω εκτίμησης κατάστασης, σε υπερδίκτυα ευρείας περιοχής, περιφερειακά δίκτυα μεταφοράς και τοπικά δίκτυα διανομής.

Ιστορικά, μόνο ένας μικρός αριθμός PMUs έχει χρησιμοποιηθεί για την παρακολούθηση γραμμών μεταφοράς με αποδεκτά σφάλματα κι αυτά ήταν πιο “πρωτόγονες” συσκευές που εγκαταστάθηκαν με μόνο σκοπό την αποτροπή καταστροφικώς διακοπών ρεύματος. Τα τελευταία χρόνια όμως, δεκάδες χιλιάδες PMU έχουν εγκατασταθεί σε δίκτυα μεταφοράς σε όλο τον κόσμο και σε ορισμένες περιπτώσεις χρησιμοποιούνται και σε δίκτυα διανομής. Τώρα, με την εισαγωγή της τεχνολογίας micro-phasor, πολλά περισσότερα PMUs από αυτά είναι δυνατό να εγκατασταθούν σε δίκτυα διανομής όπου η ισχύς μπορεί να παρακολουθείται με πολύ υψηλό βαθμό ακρίβειας. Σε συνδυασμό με έξυπνους ελεγκτές, τα PMUs προσφέρουν τη δυνατότητα να αντικατασταθούν οι συνήθως χειροκίνητες προσαρμογές που απαιτούνται από τα συστήματα SCADA, με ένα σύστημα που λαμβάνει αποφάσεις και στέλνει σήματα ελέγχου αυτόνομα.

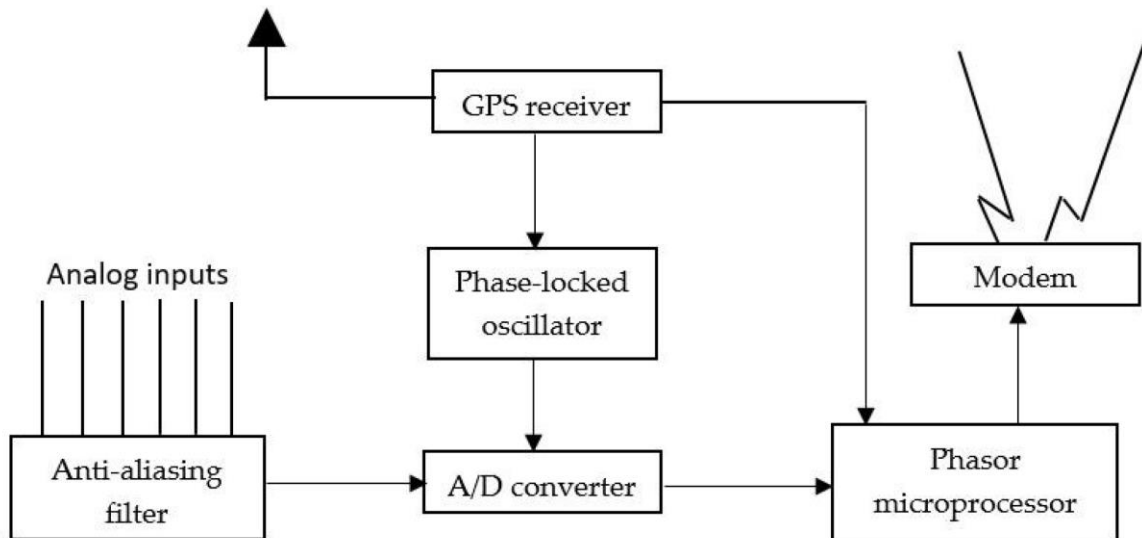
Μέχρι τώρα, τα PMUs έχουν αναπτυχθεί κυρίως βασισμένα σε δίκτυα μεταφοράς, καθώς είναι ακριβή η κατασκευή και η εγκατάστασή τους. Παράλληλα, τα δίκτυα διανομής είναι πολλά και έχουν εκτενές μέγεθος. Επιπλέον, τα δίκτυα διανομής είναι πολύ περίπλοκα, λόγω των πολλών κόμβων, των μικρών αποστάσεων, του μικρού πλάτους και των μικρών διαφορών γωνιών μεταξύ των κόμβων και των υπόλοιπων λόγων που εξηγούνται στην ενότητα για την εκτίμηση κατάστασης. Επίσης μία από τις μεγαλύτερες προκλήσεις στη λειτουργία του δικτύου διανομής είναι η περιορισμένη εποπτεία της κατάστασης τους σε πραγματικό χρόνο που καθιστά πιο περίπλοκο τον έλεγχο ενός δικτύου με υψηλή διεύθυνση αβέβαιων πόρων όπως είναι οι ανανεώσιμες πηγές ενέργειας.

Οι χαμηλότερες αποστάσεις στα δίκτυα διανομής οδηγούν σε χαμηλότερες διαφορές πλάτους και φάσης μεταξύ ηλεκτρικών μεγεθών σε διαφορετικούς κόμβους: επομένως, για να μετρηθούν σωστά αυτές οι διαφορές, απαιτείται μεγαλύτερη ακρίβεια. Επίσης, η μεγαλύτερη ευελιξία που αναμένεται στα έξυπνα δίκτυα λόγω πιθανής διακοπτόμενης συμπεριφοράς σε γεννήτριες, φορτία και συσκευές αποθήκευσης ενδέχεται να χρειάζεται χρονική απόκριση πάνω από τα πρότυπα όρια. Επιπλέον, η εντονότερη παραμόρφωση που συνήθως ανιχνεύεται στα δίκτυα διανομής καθιστά απαραίτητο τον εκ νέου καθορισμό των ορίων συμμόρφωσης με αρμονικές περισσότερο από το υπάρχον πρότυπο. Οι πληροφορίες που προέρχονται από τα PMU πρέπει να είναι άκρως αξιόπιστες, καθώς ο στόχος είναι η αυτοματοποίηση του δικτύου χωρίς ανθρώπινη επίβλεψη.

μ-PMUs – D-PMUs

Αυτές λοιπόν οι πολυπλοκότητες έχουν αυξήσει την ανάγκη ανάπτυξης νέων συστημάτων παρακολούθησης υψηλής ακρίβειας που υποστηρίζουν την επίγνωση της κατάστασης στα δίκτυα διανομής και επιτρέπουν στους διαχειριστές διανομής να λαμβάνουν καλύτερες αποφάσεις έλτιστη λειτουργία του δικτύου και για την αντιμετώπιση ποικίλων διαταραχών. Έτσι, η μικρο - μονάδα μέτρησης φασιθετών, (μ-PMU) ή μονάδα μέτρησης φασιθέτη σε επίπεδο διανομής, (D-PMU), έχει αναπτυχθεί σε δίκτυα διανομής. Αυτή η μονάδα είναι σε θέση να μετρήσει τις συγχρονισμένες φάσεις τάσης και ρεύματος (τόσο το μέγεθος όσο και η γωνία φάσης) σε πραγματικό χρόνο σε υψηλότερη ανάλυση και ακρίβεια, με σκοπό να προσφέρει ένα υψηλότερο επίπεδο εποπτείας σε ένα δίκτυο

διανομής. Το μPMU καταγράφει δύο θεμελιώδεις μετρήσεις σε τρεις φάσεις, επομένως έχει $2 \times 3 = 6$ κανάλια μέτρησης. Αυτές οι δύο μετρήσεις είναι το πλάτος τάσης, η γωνία φάσης τάσης, ως ο πρώτος φασιθέτης και το πλάτος ρεύματος και η γωνία φάσης ρεύματος, ως ο δεύτερος φασιθέτης, ανά φάση με μέγιστο ρυθμό δειγματοληψίας 120 Hz.



Διάγραμμα μπλοκ μίας μικρό μονάδας μέτρησης φασιθετών

Το σχήμα παραπάνω δείχνει το λειτουργικό μπλοκ διάγραμμα ενός μ-PMU. Έχει υψηλή ταχύτητα επεξεργασίας, εκτίμηση φάσης και ακριβή χρονική σήμανση των σημάτων εισόδου. Το φίλτρο anti-aliasing έχει το ρόλο της αφαίρεσης των στοιχείων του σήματος των οποίων η συχνότητα είναι ίση ή μεγαλύτερη από το μισό του ρυθμού Nyquist. Παράλληλα, οι αναλογικές κυματομορφές AC (φασιθέτες τάσης και ρεύματος) στην είσοδο ψηφιοποιούνται από έναν αναλογικό σε ψηφιακό μετατροπέα. Ένας ταλαντωτής φάσης μετατρέπει το σήμα GPS ενός παλμού ανά δευτερόλεπτο σε μια ακολουθία παλμών συγχρονισμού υψηλής ταχύτητας που χρησιμοποιούνται στη δειγματοληψία κυματομορφής. Κατόπιν, ο μικροεπεξεργαστής φάσης εκτελεί τους υπολογισμούς της φάσης και οι υπολογισμένες φάσεις συνδυάζονται για να σχηματίσουν μετρήσεις θετικής ακολουθίας. Σε κάθε φασιθέτη τελικά ορίζεται μία χρονική ταμπέλα με τις πληροφορίες συγχρονισμού που παρέχονται από το ρολόι και τη μέτρηση το δευτερόλεπτο που παρέχεται από τον δέκτη GPS. Η τελική τιμή της φάσης μεταδίδεται σε ένα κέντρο δεδομένων μέσω ενός μόντεμ.

Τα μ-PMUs επιτρέπουν στους χειριστές ενός κέντρου ελέγχου ενέργειας να επιτηρούν την τρέχουσα κατάσταση λειτουργίας ενός δικτύου διανομής και να εξετάζουν τις συνθήκες λειτουργίας του με την πάροδο του χρόνου μέσω εφαρμογών παρακολούθησης και διάγνωσης. Επιπλέον τους βοηθούν και να αντιμετωπίσουν τα εξιδεικευμένα προβλήματα των δικτύων διανομής που προαναφέρθηκαν καθώς και τα νέα που προκύπτουν από την αυξανόμενη διείσδυση ανανεώσιμων πηγών ενέργειας, τα ευέλικτα φορτία κ.α. Για παράδειγμα, ορισμένα συμβάντα στα δίκτυα διανομής είναι

ημιτονοειδή ή μη ημιτονοειδή μεταβατικά σε κυματομορφές τάσης και ρεύματος που μπορεί να προκληθούν από σφάλματα, αλλαγές τοπολογίας, τη συμπεριφορά του φορτίου και τη δυναμική συμπεριφορά της πηγής, συμβάντα που μπορούν να ανιχνευθούν σε πραγματικό χρόνο, αξιοποιώντας δεδομένα από τα εγκατεστημένα μ-PMUs στο δίκτυο. Επίσης, χρησιμοποιώντας μ-PMUs, η αντίστροφη ροή ισχύος στα δίκτυα διανομής μπορεί εύκολα να ανιχνευθεί όπως και αλλαγές της συχνότητας του συστήματος διανομής, ή η αναντιστοιχία ισχύος μεταξύ παραγωγής και ζήτησης φορτίου που μπορεί να προκληθεί από διαταραχές του συστήματος ισχύος. Άλλα δυσχερή προβλήματα που μπορούν να αντιμετωπιστούν με τη χρήση μ-PMUs είναι:

- Ανίχνευση σφάλματος υψηλής σύνθετης αντίστασης: πρόκειται για ένα δύσκολο πρόβλημα για τις συμβατικές μετρήσεις, λόγω της μικρής ροής ρεύματος του σφάλματος. Με τα δεδομένα των μ-PMUs, η αναγνώριση της θέσης σφάλματος γίνεται σχετικά εύκολα με τη χρήση συγχρονισμένων αρμονικών φάσεων και εστίαση στις τρίτες αρμονικές συνιστώσες των μετρήσεων πριν και μετά το σφάλμα.
- Προσδιορισμός θέσης σφάλματος: μπορεί να βοηθήσει στη μείωση της διάρκειας της διακοπής λειτουργίας και του κόστους λειτουργίας των δικτύων διανομής. Οι μετρήσεις μ-PMU συμβάλλουν στη βελτίωση της ακρίβειας της εύρεσης της θέσης του σφάλματος.
- Εκτίμηση κατάστασης: Στο δίκτυο διανομής, ο μονοφασικός εκτιμητής κατάστασης που εφαρμόζεται στο δίκτυο μεταφοράς δεν μπορεί να χρησιμοποιηθεί, επειδή το δίκτυο μεταφοράς θεωρείται ότι είναι λειτουργεί σε συμμετρική τριφασική κατάσταση φόρτισης, ενώ το δίκτυο διανομής όχι. Γενικά, η εκτίμηση κατάστασης για τα δίκτυα διανομής είναι δύσκολη, λόγω των πολυάριθμων κόμβων και της έλλειψης συγχρονισμένων μετρήσεων. Με δεδομένα μ-PMU, αυτές οι δυσκολίες μπορούν να αμβλυνθούν, επειδή οι φάσεις τάσης και ρεύματος μετρώνται άμεσα. Ο Πίνακας 3 δείχνει τη σύνοψη των διαφορετικών μεθόδων που εφαρμόζονται για την εκτίμηση της κατάστασης του συστήματος διανομής.

Τα PMUs, τα μ-PMUs και τα d-PMUs αναφέρονται μαζί ως συσκευές μέτρησης συγχρονισμένων φασιθετών (synchrophasor measurement devices), με το ακρονύμιο SMDs. Δηλαδή, τα SMDs και τα PMUs είναι μια γενική ορολογία που περιγράφει τις μονάδες που μετρούν συγχρονισμένους φασιθέτες. Υποκατηγορία αυτών είναι τα μPMUs και τα D-PMUs που αφορούν τη διανομή.

Κεφάλαιο II

Εκτίμηση κατάστασης

Η θεωρία της εκτίμησης (Estimation Theory), αποτελεί έναν κλάδο της στατιστικής ο οποίος ασχολείται με την ανάπτυξη μαθηματικών μοντέλων και μεθόδων για τον υπολογισμό των εκτιμήσεων των παραμέτρων ενός συστήματος, οι οποίες περιέχουν κάποιον παράγοντα τυχαιότητας, βασιζόμενος σε περιορισμένες πληροφορίες. Οι παράμετροι περιγράφουν ένα υποκείμενο φυσικό περιβάλλον με έναν τρόπο που η τιμή τους επηρεάζει την κατανομή των δεδομένων που μετρώνται. Αυτή η θεωρία εφαρμόζεται σε πληθώρα επιστημονικών πεδίων πέρα από την στατιστική, όπως στην μηχανική, στην ηλεκτρονική, στην ηλεκτρική ενέργεια, κλπ.

Η θεωρία αυτή συνήθως περιλαμβάνει την κατάστρωση εκτιμητών που προσπαθούν να προσεγγίσουν τις αγνωστες παραμέτρους με τον καλύτερο δυνατό τρόπο, λαμβάνοντας υπόψη τον βαθμό αβεβαιότητας των εμπειρικών μετρήσεων. Συνήθως, χρησιμοποιούνται στατιστικές μέθοδοι για την αξιολόγηση της αξιοπιστίας των εκτιμητών και την ποσότητα της αβεβαιότητας στις εκτιμήσεις.

Στη θεωρία της εκτίμησης, γενικά λαμβάνονται υπόψη δύο προσεγγίσεις, η πιθανοτική προσέγγιση, η οποία υποθέτει ότι τα μετρούμενα δεδομένα είναι τυχαία με πιθανοτική κατανομή που εξαρτάται από τις παραμέτρους που μας ενδιαφέρουν και η προσέγγιση μέλους του συνόλου (Set-Membership) που δεν υποθέτει μια συγκεκριμένη πιθανοτική διαδικασία για τα δεδομένα, αλλά αντιμετωπίζει τα δεδομένα ως μέλη ενός συνόλου. Η πρώτη προσέγγιση υποθέτει ότι τα μετρημένα δεδομένα προέρχονται από μια τυχαία διαδικασία κι έχει ως στόχο την εκτίμηση αυτών των παραμέτρων μέσω κατάλληλων στατιστικών μεθόδων. Στη δεύτερη προσέγγιση, το διάνυμα των μετρήσεων ανήκει σε ένα ορισμένο σύνολο, το οποίο εξαρτάται από το διάνυμα των παραμέτρων. Δεν κάνει σαφείς πιθανοτικές υποθέσεις, αλλά παρέχει ένα περιθώριο βεβαιότητας για την εκτίμηση των παραμέτρων.

Δημοφιλείς στατιστικές μεθόδους που χρησιμοποιούνται για τις εκτιμήσεις αποτελούν με λίγα λόγια οι:

- Μέθοδος των Ελαχίστων Τετραγώνων (Least Squares Method): Αναγάγει το πρόβλημα της εκτίμησης παραμέτρων σε πρόβλημα βελτιστοποίησης ελαχίστων τετραγώνων, όπως η προσαρμογή ευθειών ή παραβολών σε δεδομένα. Λειτουργεί ελαχιστοποιώντας το άθροισμα των τετραγώνων των αποκλίσεων μεταξύ των εκτιμώμενων και πραγματικών τιμών, είναι όμως αρκετά πιο δυσχερής υπολογιστικά και χρειάζεται μεγάλη ποικιλία και πλήθος μετρήσεων. Το κύριο πλεονέκτημα της είναι πως είναι απλή και γρήγορη στη χρήση, αλλά είναι ευαίσθητη σε λανθασμένα δεδομένα.
- Μέθοδος Μέγιστης Πιθανοφάνειας (Maximum Likelihood Estimation - MLE): Χρησιμοποιείται για να εκτιμήσει τις παραμέτρους που μεγιστοποιούν τη συνάρτηση πιθανοφάνειας που δίνει την πιθανότητα να παρατηρηθούν τα υπάρχοντα δεδομένα.

- Μέθοδος Σταθμισμένων Ελαχίστων Τετραγώνων (Weighted Least Squares): Είναι μια επέκταση της μεθόδου των ελαχίστων τετραγώνων με την διαφορά πως ορίζει διαφορετικά βάρη σε κάθε παρατήρηση, καθώς διαφορετικές παρατηρήσεις μπορεί να έχουν διαφορετικά επίπεδα ακρίβειας. Έτσι, αντικατοπτρίζεται η αξιοπιστία της κάθε μέτρησης, γεγονός που επιτρέπει μια πιο αποτελεσματική διαδικασία εκτίμησης, αφού θα δίνεται μεγαλύτερη σημασία σε πιο ακριβείς μετρήσεις και μικρότερο βάρος σε δυνητικά λιγότερο αξιόπιστες. Το κύριο πλεονέκτημα της είναι πως είναι απλή κι αποδοτική στην χρήση, αλλά ως μειονέκτημα έχει πως είναι ευαίσθητη σε λανθασμένα δεδομένα.
- Φίλτρα Κάλμαν: παρέχουν μια αποτελεσματική και αναδρομική προσέγγιση εκτίμησης, ιδίως σε σενάρια όπου το σύστημα υπόκειται σε αβεβαιότητες, υψηλά επίπεδα θορύβου και δυναμικές αλλαγές. Ο αλγόριθμος του φίλτρου Kalman αποτελείται από δύο βασικά βήματα: το βήμα πρόβλεψης και το βήμα ενημέρωσης. Στο βήμα πρόβλεψης, το φίλτρο προβλέπει την επόμενη κατάσταση με βάση τη δυναμική συμπεριφορά του συστήματος και την τρέχουσα εκτίμηση της κατάστασης. Αυτή η πρόβλεψη ενημερώνεται στη συνέχεια χρησιμοποιώντας τις μετρήσεις που λαμβάνονται από το σύστημα στο βήμα ενημέρωσης. Ο αλγόριθμος προσαρμόζει δυναμικά την εκτίμηση της κατάστασης με βάση τη διαφορά μεταξύ των εκτιμώμενων και των μετρούμενων τιμών, λαμβάνοντας υπόψη τις αβεβαιότητες που σχετίζονται με τις μετρήσεις και το μοντέλο του συστήματος. Τα φίλτρα Kalman είναι προσαρμοστικά και ικανά να χειρίζονται μη γραμμικότητες και χρονικά μεταβαλλόμενα χαρακτηριστικά των υπό μελέτη συστημάτων, ωστόσο η εξάρτηση και η ευαισθησία τους από το γραμμικό μοντέλο που χρησιμοποιούν για τις προσεγγίσεις, το οποίο μπορεί να μην είναι ακριβές και οι υπολογιστικές απαιτήσεις που χρειάζονται τα καθιστούν δύσκολα εφαρμόσιμα σε περίπλοκα δίκτυα όπως είναι συχνά τα δίκτυα διανομής.
- Εκτίμηση Bayes: Βασίζεται στο θεώρημα του Bayes, (Η πιθανότητα του A αν το B είναι αληθές αποτελείται από το γινόμενο της πιθανότητας του B όταν το A είναι αληθές επί την πιθανότητα του A ανεξαρτήτως του B και διά την πιθανότητα του B ανεξαρτήτως του A), και συνδυάζει την επεξεργασία μετρητικών δεδομένων που λαμβάνονται σε πραγματικό χρόνο και προϋπάρχουσων πληροφοριών.
- Παραγοντική Ανάλυση (Factor Analysis): Χρησιμοποιείται για να εκτιμήσει τις σχέσεις μεταξύ μεταβλητών και να ανακαλύψει τους παράγοντες που επηρεάζουν τα δεδομένα.
- Δέντρα Επίλυσης (Decision Trees): Χρησιμοποιούνται σε μη παραμετρικές μεθόδους εκτίμησης καθώς είναι ευέλικτα και εύκολα ερμηνεύσιμα. Ουσιαστικά πρόκειται για γραφικές αναπαραστάσεις που λαμβάνουν αποφάσεις βάσει σειράς ερωτήσεων που τίθενται σχετικά με τα χαρακτηριστικά των δεδομένων. Κάθε ερώτηση διακλαδώνει το δέντρο, και τα φύλλα του τελευταίου επιπέδου αντιστοιχούν σε διακριτικές κλάσεις ή εξόδους.

Κατά τα τελευταία έτη, παρατηρείται μια στροφή των μεθοδολογιών εκτίμησης προς τη μηχανική μάθηση και τα τεχνητά νευρωνικά δίκτυα [51], [55], [75] ως μια επιπλέον λύση συγκριτικά με τις ανωτέρω συμβατικές τεχνικές. Έτσι, στο πλαίσιο της παρούσας εργασίας, αξιοποιούνται ως εκτιμητές κατάστασης νευρωνικά δίκτυα καθώς αποτελούν ένα καθολικό εκτιμητή, ικανό να προσεγγίσει οποιαδήποτε συνάρτηση με οποιαδήποτε επιθυμητή ακρίβεια, υπό την προϋπόθεση ότι διαθέτει επαρκή αριθμό νευρώνων ή κρυφών επιπέδων, σύμφωνα με το Θεώρημα Καθολικής Προσέγγισης (Universal Approximation Theorem).

Όπως προαναφέρθηκε, τα τεχνητά νευρωνικά δίκτυα είναι συνδυασμοί πολλών απλών μαθηματικών συναρτήσεων που υλοποιούν πιο περίπλοκες συναρτήσεις από, συνήθως, πραγματικούς διανυσματικούς χώρους σε πραγματικούς διανυσματικούς χώρους. Οι χώροι των πολυμεταβλητών συναρτήσεων που μπορούν να υλοποιηθούν από ένα δίκτυο καθορίζονται από τη δομή του δικτύου, το σύνολο των απλών συναρτήσεων και τις παραμέτρους. Δηλαδή, τα τεχνητά νευρωνικά δίκτυα αποτελούνται από πολλαπλά στρώματα νευρώνων, και με τη χρήση συγκεκριμένων μαθηματικών συναρτήσεων και παραμέτρων, μπορούν να μοντελοποιήσουν πολύπλοκες συναρτήσεις που απεικονίζουν πραγματικά δεδομένα.

Στο μαθηματικό υπόβαθρο των τεχνητών νευρωνικών δικτύων, τα θεωρήματα καθολικής προσέγγισης συγκροτούν ένα σύνολο αποτελεσμάτων που θέτουν όρια σε αυτό που μπορούν να μάθουν θεωρητικά κατά το στάδιο της εκπαίδευσης τα νευρωνικά δίκτυα. Αναλυτικότερα, με δεδομένο έναν αλγόριθμο που παράγει δίκτυα εντός μιας κλάσης συναρτήσεων, τα θεωρήματα καθολικής προσέγγισης καθιερώνουν την πυκνότητα των παραγόμενων συναρτήσεων μέσα σε έναν συγκεκριμένο χώρο συναρτήσεων που μας ενδιαφέρει. Συνήθως, αυτά τα αποτελέσματα αφορούν τις ικανότητες προσέγγισης ενός νευρωνικού δικτύου του τύπου feedforward στο χώρο των συνεχών συναρτήσεων μεταξύ δύο Ευκλείδειων χώρων, με την προσέγγιση να γίνεται σε σχέση με την τοπολογία συμπαγούς σύγκλησης.

Ενώ κάποιες συναρτήσεις μπορούν να προσεγγίσουν εξαιρετικά καλά σε μια περιοχή, οι αποδείξεις δεν ισχύουν έξω από αυτήν την περιοχή, δηλαδή, οι προσεγγιζόμενες συναρτήσεις δεν εκτείνονται πέρα από την περιοχή αυτή. Αυτό ισχύει για όλες τις μη-περιοδικές συναρτήσεις ενεργοποίησης που χρησιμοποιούνται και οι περισσότερες αποδείξεις το υποθέτουν.

Έχουν βρεθεί επίσης παρόμοια αποτελέσματα και για μη-Ευκλείδειους χώρους και για άλλες συνήθως χρησιμοποιούμενες αρχιτεκτονικές, όπως τα συνελκτικά νευρωνικά δίκτυα (CNN), τις συναρτήσεις βασικής ακτινικής συνάρτησης κλπ με αποτέλεσμα τα θεωρήματα να εφαρμόζονται και σε αυτά.

Γενικά, τα περισσότερα θεωρήματα καθολικής προσέγγισης μπορούν να χωριστούν σε δύο κατηγορίες. Η πρώτη κατηγορία ποσοτικοποιεί τις ικανότητες προσέγγισης των νευρωνικών δικτύων με οποιοδήποτε αριθμό τεχνητών νευρώνων ("περίπτωση αυθαίρετου πλάτους"), ενώ η δεύτερη επικεντρώνεται στην περίπτωση με αυθαίρετο αριθμό κρυμμένων στρωμάτων, καθεμιά από τις οποίες περιέχει περιορισμένο αριθμό τεχνητών νευρώνων ("περίπτωση αυθαίρετου βάθους"). Εκτός από αυτές τις δύο κατηγορίες, υπάρχουν επίσης θεωρήματα καθολικής προσέγγισης για νευρωνικά δίκτυα με περιορισμένο αριθμό κρυμμένων στρωμάτων και

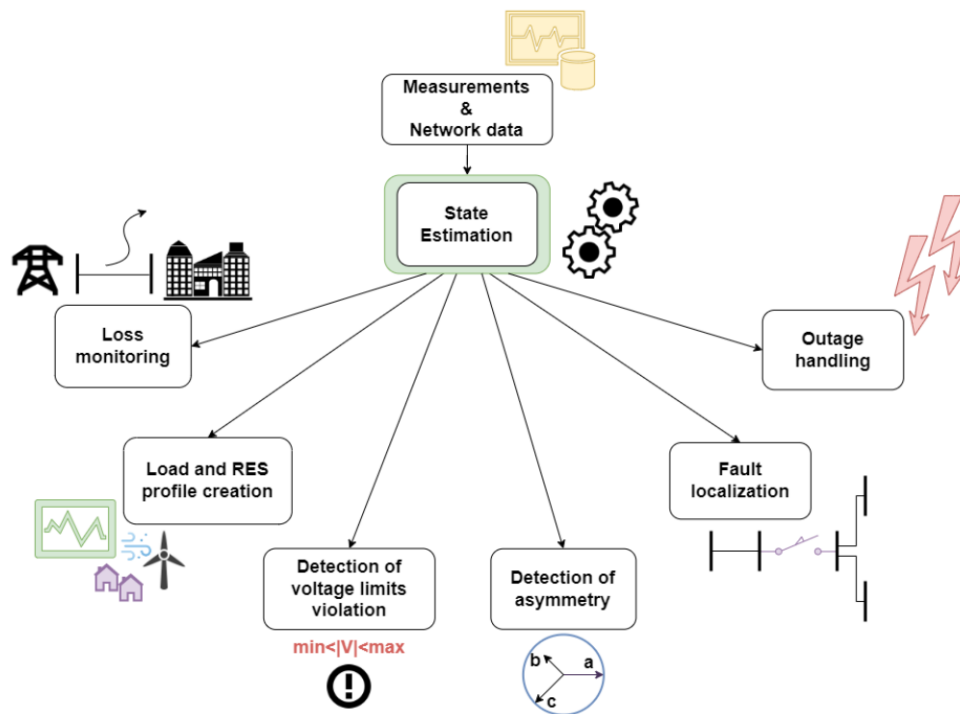
περιορισμένο αριθμό νευρώνων σε κάθε στρώμα ("περίπτωση περιορισμένου βάθους και περιορισμένου πλάτους").

Τα θεωρήματα παγκόσμιας προσέγγισης υποδηλώνουν ότι τα νευρωνικά δίκτυα μπορούν να αναπαραστήσουν μια ευρεία ποικιλία συναρτήσεων με κατάλληλα βάρη, αλλά δεν παρέχουν συγκεκριμένες οδηγίες για την εύρεση των τιμών των βαρών, απλώς διατυπώνουν ότι μια τέτοια κατασκευή είναι δυνατή. Για να κατασκευαστούν τα βάρη, τα νευρωνικά δίκτυα εκπαιδεύονται, και μπορεί να συγκλίνουν στα σωστά βάρη ή όχι (δηλαδή να κολλήσουν σε ένα τοπικό μέγιστο). Εάν το δίκτυο είναι πολύ μικρό (αναφορικά με τις διαστάσεις των εισόδων δεδομένων), τα θεωρήματα παγκόσμιας προσέγγισης δεν ισχύουν, καθώς τα δίκτυα δεν μαθαίνουν ποτέ.

Εν κατακλείδει, έχει αποδειχτεί πως τα νευρωνικά δίκτυα με μία μη περιορισμένη – δηλαδή μη πολυωνυμική - συνάρτηση ενεργοποίησης έχουν την ιδιότητα της καθολικής προσέγγισης.

Εκτίμηση Κατάστασης σε ΣΗΕ

Η εκτίμηση κατάστασης ενός ΣΗΕ είναι η διαδικασία υπολογισμού των τιμών των μεταβλητών κατάστασης του συστήματος χρησιμοποιώντας έναν περιορισμένο αριθμό μετρήσεων σε συγκεκριμένα σημεία του συστήματος, σε πραγματικό χρόνο. Έτσι, η εκτίμηση κατάστασης είναι βασικά μια αριθμητική διαδικασία η οποία στοχεύει στην αντιστοίχιση δεδομένων από μετρήσεις στο δίκτυο με μεταβλητές κατάστασης. Τα τελευταία χρόνια, με νέες υποδομές διεσπαρμένης παραγωγής, ΑΠΕ, συστημάτων αποθήκευσης ενέργειας, έξυπνων μετρητικών συστημάτων με δυνατότητες διαχείρισης της ζήτησης κ.α. να αναπτύσσονται διαρκώς προκειμένου να υποστηριχθεί η εξέλιξη προς την ενεργειακή μετάβαση στα ευφυή δίκτυα ηλεκτρικής ενέργειας, τα δίκτυα διανομής μετασχηματίζονται από παθητικά σε ενεργά συστήματα. Σε αυτό έντονα δυναμικό πλαίσιο λειτουργίας, η εκτίμηση κατάστασης δικτύων διανομής (Distribution system state estimation - DSSE), αναμένεται να γίνει γίνει ο ακρογωνιαίος λίθος στην εποπτεία και τη διαχείριση των έξυπνων δικτύων ηλεκτρικής ενέργειας μέσης και χαμηλής τάσης. Ένα γενικό σχήμα των λειτουργιών που επιτελούνται μέσω της εκτίμησης κατάστασης, παρουσιάζεται παρακάτω [10]:



Η εκτίμηση κατάστασης είναι αναγκαία καθώς επιτρέπει την παρακολούθηση του δικτύου διανομής σε πραγματικό χρόνο και παρέχει την αρχική κατάσταση/συνθήκη για πολλές εφαρμογές Διαχείρισης Δικτύου Διανομής (Distribution Management System - πλατφόρμες λογισμικού σχεδιασμένες για την αποτελεσματική εποπτεία, έλεγχο και βελτιστοποίηση της λειτουργίας του δικτύου διανομής), όπως είναι για παράδειγμα το SCADA. Οι εφαρμογές αυτές αναμένεται να πραγματοποιούν έναν αυξανόμενο αριθμό λειτουργιών, όπως ο έλεγχος τάσης / αέργου ισχύος, (volt/VAr optimization), η βελτιστοποίηση της χρήσης των μετασχηματιστών διανομής, η ελαχιστοποίηση των απωλειών ενέργειας – εντοπισμός σφαλμάτων, η αναδιαμόρφωση των ΔΔ μέσω τοπολογικών αλλαγών, η υποστήριξη λειτουργίας αγορών ηλεκτρικής ενέργειας. Γενικότερα, η ακρίβεια με την οποία γίνεται η εκτίμηση κατάστασης ΔΔ έχει υψηλές επιπτώσεις σε όλες τις λειτουργίες του δικτύου.

Η εκτίμηση κατάστασης βοηθάει σημαντικά στη βελτίωση της παρατηρησιμότητας καθώς εκτιμώντας τις τιμές των μη μετρημένων ή λανθασμένως μετρημένων μεταβλητών, παρέχεται μια πιο αξιόπιστη εικόνα του δικτύου διανομής. Παράλληλα, ένας από τους κύριους ρόλους της εκτίμησης κατάστασης ΔΔ είναι να καθιστά δυνατή την παρακολούθηση σε πραγματικό χρόνο του δικτύου διανομής. Εκτιμώντας συνεχώς τις μεταβλητές κατάστασης, οι οποίες συνήθως είναι οι τάσεις των κόμβων του ΔΔ, η εκτίμηση κατάστασης ΔΔ παρέχει συνεχώς ενημερωμένες πληροφορίες, οι οποίες είναι ουσιώδεις για τη διασφάλιση της αξιόπιστης και ασφαλούς λειτουργίας του δικτύου.

Επίσης, η ακριβής εκτίμηση της κατάστασης του συστήματος συνεισφέρει στην ασφάλεια του δικτύου καθώς βοηθά στον εντοπισμό και την αντιμετώπιση πιθανών προβλημάτων αμέσως, μειώνοντας την πιθανότητα κατάρρευσης. Ο πιο γρήγορος εντοπισμός προβλημάτων επιτρέπει και πιο ταχείες αντιδράσεις, μειώνοντας το χρόνο αδράνειας και ενισχύοντας τη συνολική αξιοπιστία του δικτύου. Αυτή η προληπτική

προσέγγιση στη διαχείριση βλαβών συντελεί στη βελτίωση της απόδοσης του συστήματος κι ενισχύει την συνολική ανθεκτικότητα του ηλεκτρικού δικτύου, εξασφαλίζοντας αξιόπιστη παροχή ενέργειας στους καταναλωτές. Επίσης, με την αυξανόμενη πολυπλοκότητα των δικτύων διανομής και την ενσωμάτωση των ψηφιακών τεχνολογιών, η DSSE βοηθά στο πεδίο της κυβερνοασφάλειας και της συνολικής ασφάλειας του δικτύου, καθώς μέσω του εντοπισμού ανωμαλιών και ασυνήθιστων καταστάσεων υποδεικνύει γρήγορα και τους κινδύνους αυτού του τύπου.

Παράλληλα, παρέχοντας μια σαφή απεικόνιση της κατάστασης του συστήματος, η εκτίμηση κατάστασης ΔΔ βοηθάει την προσαρμογή του δικτύου κατά την αυξανόμενη εισαγωγή ΑΠΕ και στη βελτιστοποίηση της λειτουργίας του δικτύου, την ελαχιστοποίηση των απωλειών και την βελτίωση της συνολικής αποδοτικότητας της ενέργειας.

Στο πλαίσιο των έξυπνων δικτύων, δηλαδή δικτύων που είναι προηγμένα και ψηφιακά βελτιωμένα και τα οποία αξιοποιούν τις σύγχρονες τεχνολογίες επικοινωνίας και ελέγχου αλλά και υπολογιστές για τη βελτιστοποίηση της παραγωγής και της διανομής ηλεκτρικής ενέργειας, η DSSE γίνεται ακόμα πιο κρίσιμη επειδή επιτρέπει την ενσωμάτωση αυτών των προηγμένων τεχνολογιών, όπως έξυπνα όργανα μέτρησης κι αισθητήρες παρέχοντας τα απαραίτητα δεδομένα για τη λήψη αποφάσεων και στρατηγικών ελέγχου.

Οι σχεδιαστές των δικτύων διανομής πολλές φορές χρησιμοποιούν DSSE για να εκτιμήσουν και να εξετάσουν το ιστορικό και την τρέχουσα φόρτιση των μετασχηματιστών και των γραμμών αλλά και για τη δοκιμή προτεινόμενων καταστάσεων φόρτισης. Τέλος, οι πληροφορίες που παράγονται από την DSSE, μπορούν να χρησιμοποιηθούν στις λειτουργίες της αγοράς καθώς εκτός από την συνεισφορά στον υπολογισμό των σημάτων τιμολόγησης που προαναφέρθηκε, βοηθούν σε αποτελεσματικότερες συναλλαγές στην αγορά και στη βέλτιστη κατανομή των πόρων.

Μαθηματικός Ορισμός

Μαθηματικά, ως εκτιμητής του x : $\hat{x}(y)$, με λάθος: $e = \hat{x} - x$, ορίζεται οποιαδήποτε συνάρτηση της μέτρησης y . Στην Μπεϋζιανή προσέγγιση που ακολουθείται, τα x και y θεωρούνται ως τυχαίες μεταβλητές. Για την εκτίμηση κατάστασης χρησιμοποιήθηκε ένας εκτιμητής ελάχιστου μέσου τετραγωνικού σφάλματος (MMSE - minimum mean squared error), ο οποίος δίνεται από τον τύπο:

$\hat{x}_{\text{MMSE}}(y) = \text{argmin}_x(\text{MSE})$, όπου το μέσο τετραγωνικό σφάλμα (MSE) είναι:

$$\min_{\hat{x}} \mathbf{E}(\|x - \hat{x}\|^2)$$

Ο όρος argmin αναφέρεται στην τιμή εισόδου (ή στις τιμές) μιας συνάρτησης που ελαχιστοποιεί την έξοδο της συνάρτησης, δηλαδή, βρίσκει το σύνολο όλων των τιμών εισόδου για τις οποίες η εσωτερική συνάρτηση παίρνει την ελάχιστη τιμή της. Ως $\mathbf{E}[\cdot]$ συμβολίζεται ο τελεστής μαθηματικής προσδοκίας, ο οποίος αντιπροσωπεύει τη μέση τιμή μιας τυχαίας μεταβλητής σε όλα τα πιθανά αποτελέσματα, σταθμισμένη με τις αντίστοιχες πιθανότητες.

Ο εκτιμητής αυτός ελαχιστοποιεί άμεσα το σφάλμα εκτίμησης (το μέσο τετραγωνικό σφάλμα (MSE), το οποίο είναι ο μέσος όρος των τετραγωνικών διαφορών μεταξύ των εκτιμώμενων τιμών και των πραγματικών τιμών), ενώ οι κλασικοί εκτιμητές, όπως η μέθοδος ελαχίστων τετραγώνων, ελαχιστοποιούν το σφάλμα μοντελοποίησης ενσωματωμένο μέσω μιας συνάρτησης μέτρησης που συσχετίζει τις μετρήσεις με τις καταστάσεις. Παρακάμπτοντας την ανάγκη της χρήσης μιας συνάρτησης μέτρησης, η αναγκαία προϋπόθεση για παρατηρησιμότητα σε πραγματικό χρόνο παρακάμπτεται επίσης σε αυτόν τον εκτιμητή κατάστασης. Επίσης, ο εκτιμητής MMSE είναι ιδιαίτερα πολύτιμος σε περιπτώσεις όπου υπάρχει αβεβαιότητα ή θόρυβος στα δεδομένα, καθώς με την ελαχιστοποίηση του MSE, μειώνει ταυτόχρονα τόσο την παράμετρο προτίμησης (bias - συστηματικό σφάλμα) όσο και τη διακύμανση (τυχαίο σφάλμα), αποδίδοντας εκτιμήσεις που είναι ακριβείς και αξιόπιστες. Καθώς όλες οι μετρήσεις στα δίκτυα διανομής μέσω των SMDs υπόκεινται σε σφάλματα λόγω του θορύβου των οργάνων μετρήσεων, αυτή η ιδιότητα είναι ιδιαίτερα σημαντική.

Στην παρούσα εργασία, ο εκτιμητής κατάστασης MMSE προσεγγίζεται με ένα βαθύ νευρωνικό δίκτυο. Τα νευρωνικά δίκτυα, όπως προαναφέρθηκε, αποδεδειγμένα μπορούν να προσεγγίσουν οποιαδήποτε συνάρτηση με επιτυχία, δεδομένου ενός αρκετά μεγάλου αριθμού παραμέτρων. Ενώ οι εκτιμητές MMSE περιλαμβάνουν συνήθως πολύπλοκες μαθηματικές διατυπώσεις και μπορεί να μην έχουν ρητές λύσεις κλειστής μορφής, τα DNN προσφέρουν μια ευέλικτη και καθοδηγούμενη από τα δεδομένα προσέγγιση της υποκείμενης σχέσης μεταξύ των μεταβλητών εισόδου και εξόδου.

Επίσης, τα νευρωνικά δίκτυα έχουν την ικανότητα να μοντελοποιούν εξαιρετικά μη γραμμικές σχέσεις και να συλλαμβάνουν περίπλοκα μοτίβα στα δεδομένα. Αυτή η ιεραρχική αναπαράσταση των δεδομένων που μαθαίνουν, τους επιτρέπει να αποτυπώνουν πολύπλοκες εξαρτήσεις μεταξύ των μεταβλητών εισόδου και εξόδου, καθιστώντας τα κατάλληλα για την προσέγγιση μη γραμμικών συναρτήσεων, όπως αυτές που συναντώνται σε προβλήματα εκτίμησης MMSE.

Ουσιαστικά, μέσω του νευρωνικού δικτύου, βρίσκεται μια απεικόνιση, $\mathcal{K}(\cdot)$, που συνδέει τα x και y . Το νευρωνικό δίκτυο μαθαίνει να προσεγγίζει αυτή την απεικόνιση επεξεργαζόμενο τις μετρήσεις εισόδου y και παράγοντας μια εκτίμηση της πραγματικής κατάστασης x .

Λόγοι Δυσκολίας Εκτέλεσης DSSE

Μετά από πάνω από τέσσερις δεκαετίες ανάπτυξης, στα συστήματα μεταφοράς η εκτίμηση κατάστασης (TSSE – transmission system state estimation) είναι μία αρκετά ανεπτυγμένη και ευρέως χρησιμοποιούμενη διαδικασία. Οι επαρκείς μετρήσεις έχουν επιτρέψει την πλήρη παρατηρησιμότητα του συστήματος μεταφοράς και την επεξεργασία λανθασμένων δεδομένων για την εκτίμηση κατάστασης. Υποθέτοντας ισορροπημένη λειτουργία του δικτύου, οι καταστάσεις των διακοπών

κυκλώματος, η θέση του on-line tap changer (συσκευής που αλλάζει την αναλογία ενός μετασχηματιστή χωρίς να διακόψει τη ροή ισχύος) και αναλογικές μετρήσεις, όπως η πραγματική και η άεργος ροή ισχύος, οι εγχύσεις ισχύος στους ζυγούς, οι τάσεις και οι φασικές γωνίες χρησιμοποιούνται για την εκτίμηση κατάστασης συστημάτων μεταφοράς (TSSE). Συνήθως, το πλάτος και η γωνία φάσης της τάσης χρησιμοποιούνται ως μεταβλητές κατάστασης καθώς εάν οι τιμές τους είναι γνωστές, μπορούμε να προσδιορίσουμε την τιμή της ισχύος καθώς και την κατεύθυνση της ροής ισχύος σε κάθε κλάδο. Η εκτίμηση κατάστασης συστημάτων μεταφοράς είναι ένα βασικό εργαλείο στο κέντρο ελέγχου ενέργειας και συνήθως εκτελείται σε συνδυασμό με λειτουργίες αξιολόγησης ασφαλείας κάθε 2 λεπτά ή και λιγότερο για να διασφαλιστεί η ασφαλής λειτουργία του συστήματος.

Η χρήση της εκτίμηση κατάστασης όμως στο επίπεδο των δικτύων διανομής εξακολουθεί να είναι αντικείμενο έρευνας. Ήδη από τη δεκαετία του 1960 χρησιμοποιούνταν υπολογιστές για αναλύσεις ροής φορτίου σε πραγματικό χρόνο των πρωτογενών συστημάτων διανομής με βάση εκτιμώμενα μοντέλα φορτίου. Οι αλγόριθμοι που έχουν αναπτυχθεί για εκτίμηση κατάστασης συστημάτων μεταφοράς πρέπει να προσαρμοστούν για να είναι κατάλληλοι για την εκτίμηση κατάστασης ΔΔ Αυτό αποτελεί πρόκληση λόγω πολλών παραγόντων που βασίζονται στις σημαντικές διαφορές μεταξύ των συστημάτων μεταφοράς και διανομής.

Αρχικά, σε αντίθεση με τα συστήματα μετάδοσης, τα δίκτυα διανομής είναι συνήθως μη παρατηρήσιμα, δηλαδή ο αριθμός των οργάνων που λαμβάνουν μετρήσεις είναι γενικά πολύ πιο μικρός σε σχέση με τον τεράστιο όγκο του συστήματος. Αυτή η έλλειψη επαρκών μετρήσεων πραγματικού χρόνου μπορεί να αντισταθμιστεί αν χρησιμοποιηθούν ψευδομετρήσεις. Αυτές, είναι μετρήσεις που υπολογίζονται χρησιμοποιώντας βραχυπρόθεσμες προβλέψεις ή ιστορικά δεδομένα, αλλά ως αποτέλεσμα, δεν είναι τόσο ακριβείς όσο οι πραγματικές. Ωστόσο στην χρήση τους δεν απαιτείται περαιτέρω ενέργεια ή κόστος.

Επιπροσθέτως στα δίκτυα διανομής συνήθως ο λόγος της επαγωγικής αντίστασης προς την ωμική είναι αρκετά χαμηλός λόγω των πολυάριθμων μη ισορροπημένων τριφασικών κλάδων, των ασύμμετρων φορτίων και της μικρής απόστασης ανάμεσα στους κόμβους. Αυτό καθιστά τις τεχνικές εκτίμησης της κατάστασης συνεχούς ρεύματος που χρησιμοποιούνται στα συστήματα μεταφοράς ακατάλληλες για εφαρμογή στο επίπεδο της διανομής. Παράλληλα, τα δίκτυα διανομής στην πράξη είναι μη ισορροπημένα, δηλαδή οι τρεις φάσεις του ηλεκτρικού συστήματος δεν μεταφέρουν την ίδια ποσότητα ρεύματος. Αυτό οδηγεί σε υψηλότερα επίπεδα πολυπλοκότητας κατά τη διατύπωση του προβλήματος της εκτίμησης κατάστασης (SE) – δηλαδή εισάγονται περισσότερες άγνωστες μεταβλητές που χαρακτηρίζονται από μη γραμμικές εξισώσεις.

Η πολυπλοκότητα του προβλήματος αυξάνεται επίσης λόγω της δομής του δικτύου. Τα δίκτυα διανομής όπως προαναφέρθηκε περιλαμβάνουν έναν τεράστιο αριθμό υποσταθμών, γεννήτριων, μετασχηματιστών και φορτίων, τα οποία είναι οργανωμένα συνήθως σε ακτινικές τοπολογίες. Αντίστοιχα, τα διαθέσιμα μετρητικά δεδομένα που διαθέτουμε είναι συγκριτικά πολύ λιγότερα με συνέπεια την ελλιπή πληροφόρηση για την τρέχουσα τοπολογία του δικτύου. Επιπλέον η συνεχώς αυξανόμενη διείδυση των ανανεώσιμων πηγών ενέργειας εισάγει ένα υψηλότερο επίπεδο αβεβαιότητας στη

λειτουργία του δικτύου λόγω της φυσικής έλλειψης προβλεψιμότητας της μορφής παραγωγής ενέργειας. Συμπερασματικά οι προκλήσεις για την εκτίμηση κατάστασης σε ΔΔ γίνονται εντονότερες σε αυτό το περιβάλλον ελλειπών και συχνά αβέβαιων δεδομένων.

Επιπλέον, τα δίκτυα διανομής χαρακτηρίζονται από αδυναμίες στην επικοινωνία και τη μετάδοση δεδομένων. Στα σύγχρονα ΣΗΕ, υπάρχει μια αυξανόμενη εξάρτηση από τα δίκτυα επικοινωνίας για τη διευκόλυνση της ανταλλαγής πληροφοριών μεταξύ διαφορετικών συσκευών, όπως αισθητήρες, ευφυείς ηλεκτρονικές συσκευές (όπως για παράδειγμα τα PMUs) και τα κέντρα ελέγχου. Αυτές οι συσκευές συλλέγουν δεδομένα σε πραγματικό χρόνο από το δίκτυο και τα κοινοποιούν σε κεντρικά συστήματα για παρακολούθηση, έλεγχο και ανάλυση. Όμως η ποσότητα των δεδομένων που μπορεί να μεταδοθεί μέσω ενός δικτύου επικοινωνίας περιορίζεται από το εύρος ζώνης του και την χωρητικότητά του. Στο πλαίσιο των δικτύων διανομής, ανταλλάσσεται μεγάλος όγκος δεδομένων, όπως μετρήσεις, πληροφορίες κατάστασης και εντολές ελέγχου, όμως εάν το εύρος ζώνης είναι ανεπαρκές ή το σύστημα λειτουργεί κοντά στη χωρητικότητά του με αποτέλεσμα τη δημιουργία καθυστερήσεων, την απώλεια δεδομένων και την μείωση του ρυθμού μετάδοσης.

Τέλος, το θέμα της κυβερνοασφάλειας εισάγει νέες προκλήσεις στη διαχείριση και τον έλεγχο των δικτύων διανομής. Τα σύγχρονα δίκτυα περιλαμβάνουν την ενσωμάτωση προηγμένων τεχνολογιών για την αποτελεσματική και ευέλικτη διαχείριση της ηλεκτρικής ενέργειας, όπως τα ψηφιακά συστήματα επικοινωνίας και αυτοματισμού στα έξυπνα δίκτυα. Όμως, επειδή αυτά τα συστήματα βασίζονται σε διασυνδεδεμένες συσκευές, δίκτυα επικοινωνίας και λογισμικό είναι δυνητικά, όπως άλλωστε και όλα τα υπολογιστικά συστήματα, ευάλωτα σε απειλές στον κυβερνοχώρο, καθώς η μη εξουσιοδοτημένη πρόσβαση ή οι επιθέσεις στον κυβερνοχώρο θα μπορούσαν να διαταράξουν την κανονική λειτουργία του συστήματος διανομής.

Επεξεργαστής Τοπολογίας

Η τοπολογία ενός δικτύου διανομής αλλάζει με την πάροδο του χρόνου, γεγονός το οποίο έχει αντίκτυπο και στην εκτίμηση κατάστασης. Έτσι έχουν προταθεί διάφορες μέθοδοι εκτίμησης της τοπολογίας του δικτύου προτού εφαρμοστεί η εκτίμηση κατάστασης ΔΔ.

Στο [12] έχει προταθεί η χρήση γραμμικού προγραμματισμού μεικτών ακεραίων για την εκτίμηση της τοπολογίας, όμως για την υλοποίηση αυτή της μεθόδου είναι απαραίτητες μετρήσεις ροής ισχύος γραμμών και έξυπνων μετρητών τα οποία όπως προαναφέρθηκε συνήθως δεν είναι διαθέσιμα στα δίκτυα διανομής. Στο [13] χρησιμοποιήθηκε μια μέθοδος βελτιστοποίησης βασισμένη στη θεωρία των γράφων η οποία εκτιμά σε πραγματικό χρόνο την ακτινική τοπολογία των δικτύων διανομής χρησιμοποιώντας έναν περιορισμένο αριθμό μετρήσεων πραγματικού χρόνου. Ωστόσο, και εδώ οι μετρήσεις που χρησιμοποιήθηκαν υπόκεινται στους ίδιους περιορισμούς και δεν συμπεριλήφθηκαν στη μέθοδο αυτή μη ακτινικά δίκτυα. Στο [14] ένα πιθανοτικό μοντέλο που βασίζεται σε δεδομένα έξυπνων μετρητών

χρησιμοποιήθηκε για την εκτίμηση τοπολογίας. Όμως, αυτού τους είδους τα δεδομένα είναι ακατάλληλα για αναγνώση τοπολογίας σε πραγματικό χρόνο.

Παράλληλα, στο [15] μια μέθοδος επαλήθευσης βασισμένη σε χρονοσειρές προτάθηκε η οποία καταγράφει τις αλλαγές στην τοπολογία των δικτύων με βάση σειρών μετρήσεων από micro-PMUs. Οι δυνατότητες της μεθόδου αυτή περιορίζονται ωστόσο από τις προϋποθέσεις του ενός ανοιγοκλείματος διακόπτη τη φορά και της πρότερης γνώσης της κατάστασης των διακοπών, οι οποίες απαιτήθηκαν. Στο [16] προτάθηκε ένα πλαίσιο βασισμένο στην μηχανική μάθηση για την αναγνώριση της τοπολογίας. Ωστόσο χρησιμοποιήθηκαν ρεύματα κόμβων (ζυγών), τάσεις και γωνίες της διαφοράς φάσης μεταξύ του ρεύματος και της τάσης, δεδομένα τα οποία λόγω της δυσκολίας απόκτησης τους για δίκτυα διανομής και του πλήθους καθιστούν αυτήν την μέθοδο δύσκολα εφαρμόσιμη στην πράξη για πραγματικό χρόνο. Επίσης, στο [17] προτάθηκε μια αριθμητική μέθοδος δύο βημάτων η οποία όμως είναι πολύ χρονοβόρα κατά τον έλεγχο του δικτύου σε πραγματικό χρόνο και λειτουργεί μόνο για συμμετρικά δίκτυα, ενώ στο [18] πραγματοποιήθηκε αναγνώριση της τοπολογίας σε μη ισορροπημένα δίκτυα με τη χρήση ψευδομετρήσεων οι οποίες παράχθηκαν με βάση προβλεπόμενα δεδομένα φορτίων, τα οποία όμως έχειδειχθεί πως μειώνουν την απόδοση. Επίσης καμία από αυτές τις μεθόδους δεν ασχολήθηκε με την εύρεση των κατάλληλων σημείων που μπορούν να τοποθετηθούν μετρητές για να αυξηθεί η παρατηρησιμότητα του δικτύου κι έτσι να γίνει πιο αποδοτική η εκτίμηση της τοπολογίας και της κατάστασης.

Παρατηρησιμότητα

Η παρατηρησιμότητα σε ένα σύστημα ελέγχου αναφέρεται στη δυνατότητα να παρακολουθηθεί ή να εκτιμηθεί η κατάσταση του συστήματος χρησιμοποιώντας μόνο τα μετρήσιμα σήματα εξόδου, για κάθε πιθανή εξέλιξη των διανυσμάτων κατάστασης και ελέγχου. Αυτό σημαίνει ότι, με βάση τα δεδομένα που είναι διαθέσιμα από τις μετρήσεις, μπορούμε να καταλάβουμε ή να προβλέψουμε την κατάσταση του συστήματος.

Ένα σύστημα ονομάζεται παρατηρήσιμο όταν οι εσωτερικές του καταστάσεις μπορούν να εξαχθούν ή να εκτιμηθούν με βάση τα μετρήσιμα σήματα εξόδου του, δηλαδή ότι μπορούμε να παρακολουθήσουμε ή να εκτιμήσουμε τις καταστάσεις του συστήματος με βάση τις διαθέσιμες παρατηρήσεις. Αυτό σημαίνει ότι δεν υπάρχουν ανεξάρτητες κρυφές (μη μετρήσιμες) καταστάσεις που δεν επηρεάζουν τα μετρήσιμα σήματα.

Η παρατηρησιμότητα είναι σημαντική για τον σχεδιασμό των συστημάτων ελέγχου, καθώς επιτρέπει την αξιοποίηση των πληροφοριών που προκύπτουν από τα μετρήσιμα σήματα για τον έλεγχο και την παρακολούθηση του συστήματος. Επίσης, η παρατηρησιμότητα είναι ουσιώδης για τον σχεδιασμό ελεγκτών, τη διάγνωση βλαβών, και άλλες εφαρμογές στην ελέγχου και την αυτοματοποίηση. Αν ένα σύστημα δεν είναι παρατηρήσιμο, μπορεί να είναι δύσκολη ή αδύνατη η εκτέλεση των διαδικασιών αυτών με αξιοπιστία.

Στο πλαίσιο των ΣΗΕ, η παρατηρησιμότητα ενός συστήματος αναφέρεται στη δυνατότητα παρακολούθησης και εκτίμησης των καταστάσεων (όπως οι τάσεις και τα ρεύματα ή ακόμα και οι διάφορες ροές ισχύος ανάμεσα στους

κλάδους) στο εσωτερικό του συστήματος με βάση τις διαθέσιμες μετρήσεις, σε διάφορα σημεία του δικτύου.

Σε πολύ μεγάλα ηλεκτρικά δίκτυα, όπου υπάρχει πληθώρα σημείων μέτρησης, η παρατηρησιμότητα μπορεί να είναι πιο εύκολα επιτεύξιμη. Ωστόσο, σε περιπτώσεις περιορισμένων σημείων μέτρησης, η πρόκληση είναι να εξασφαλίσετε ότι το σύστημα παραμένει παρατηρήσιμο, δηλαδή, αν μπρούν να εκτιμηθούν οι καταστάσεις που δεν μπορούν να μετρηθούν απευθείας.

Στα δίκτυα διανομής, η παρατηρησιμότητα μπορεί να είναι πρακτικά δύσκολη λόγω του γεγονότος ότι συνήθως δεν έχουμε επαρκή αριθμό μετρήσεων σε σημεία του δικτύου. Σε ένα μεγάλο δίκτυο, η εγκατάσταση μετρητικών συσκευών, (όπως τα PMUs), σε κάθε κόμβο μπορεί να είναι ανεφάρμοστη λόγω κόστους και τεχνικών δυσκολιών. Επιπλέον σε κάποια σημεία του δικτύου, ενδέχεται να μην υπάρχουν καν μετρήσεις, προκαλώντας προβλήματα παρατηρησιμότητας. Επίσης, τα δεδομένα που λαμβάνουμε μπορεί να είναι επιρρεπή σε μεγάλα σφάλματα και να περιέχουν μεγάλο βαθμό αβεβαιότητας, γεγονός που μπορεί να επηρεάσει την ικανότητά μας εξάγουμε ποιοτικές εκτιμήσεις της κατάστασης του δικτύου.

Η παρατηρησιμότητα αναφέρεται επίσης στη δυνατότητα του χρήστη ενός συστήματος να λύσει το πρόβλημα εκτίμησης κατάστασης. Αυτό εξαρτάται από τον αριθμό και την τοποθεσία των μετρητικών μονάδων στο υπό μελέτη ΣΗΕ. Επίσης, η διαθεσιμότητα και η ποιότητα των κρίσιμων δειγμάτων μέτρησης σε πραγματικό χρόνο έχει κρίσιμη επίδραση στην παρατηρησιμότητα του ΣΗΕ.

Σύμφωνα με την παραδοσιακή προσέγγιση, αριθμητικές και τοπολογικές μέθοδοι χρησιμοποιούνται για να αξιολογήσουν την παρατηρησιμότητα των συστημάτων μεταφοράς όσον αφορά τον αριθμό και την τοποθεσία των συσκευών μέτρησης. Όμως, σε αντίθεση με τα συστήματα μεταφοράς που έχουν υψηλό επίπεδο ακόμη και περίσσεια μετρητικών δεδομένων, τα δίκτυα διανομής είναι, όπως προαναφέρθηκε γενικά όχι επαρκώς επιτηρούμενα και με χαμηλό επίπεδο παρατηρησιμότητας. Επομένως, η ακρίβεια της εκτίμησης κατάστασης διανομής μπορεί να επηρεαστεί σημαντικά από την ποιότητα και τη διαθεσιμότητα των αισθητήρων.

Για να βελτιωθεί η παρατηρησιμότητα των συστημάτων διανομής, το σύνολο μετρήσεων πρέπει να επεκταθεί τεχνητός (για να αντισταθμιστούν τα δεδομένα που λείπουν) ή να διορθωθεί (για να αντισταθμιστούν τα λανθασμένα δεδομένα). Αυτό μπορεί να γίνει μέσω της χρήσης "ψευδομετρήσεων" που είναι τεχνητά παραγόμενα δεδομένα, όπως ενεργού ή αέργου ισχύος, τάσης κλπ, βασισμένα στο ιστορικά δεδομένα των δικτύων διανομής και σε ποικίλες άλλες πηγές όπως προβλεπόμενη κατανάλωση, προφίλ καιρικών συνθηκών κλπ.

Στο πλαίσιο αυτής της εργασίας, για να παρακαμφθεί το πρόβλημα της έλλειψης μετρητών η οποία οδηγεί σε μείωση του βαθμού παρατηρησιμότητας του δικτύου, χρησιμοποιήθηκε ο εκτιμητής κατάστασης MMSE που προαναφέρθηκε στην ενότητα παραπάνω για την εκτίμηση κατάστασης

Τύποι Μετρήσεων - Βιβλιογραφική Ανασκόπηση

Τρεις διαφορετικές κατηγορίες μετρήσεων γενικά χρησιμοποιούνται κατά την εκτίμηση κατάστασης σε δίκτυα διανομής.

Η πρώτη είναι οι μετρήσεις πραγματικού χρόνου. Συνήθως συλλέγονται μετρήσεις τάσεων στις γραμμές τροφοδοσίας, ρευματοφόρων κλάδων, ροών ισχύος και κατάστασης διακοπών σε λίγα σημεία της τροφοδοσίας. Πιο συγκεκριμένα δεδομένα SCADA, (Supervisory Control and Data Acquisition – αυτοματοποιημένο σύστημα συλλογής και ελέγχου δεδομένων), από τους υποσταθμούς διανομής είναι διαθέσιμα κάθε λίγα λεπτά [19]. Επίσης, συνεχώς εγκαθίστανται έξυπνων μετρητές και IEDs με δυνατότητα διπλής κατεύθυνσης επικοινωνίας κι έξυπνοι μετρητές πελατών οι οποίοι αναφέρουν δεδομένα ζήτησης κάθε δεκαπέντε λεπτά ή περισσότερο ανάλογα με τις απαιτήσεις του χειριστή. Πληροφορίες συγχρονισμού από IEDs είναι διαθέσιμες σε ρυθμιζόμενες ταχύτητες 1–60 μηνύματα ανά δευτερόλεπτο.

Τέλος, αρκετές μέθοδοι DSSE που έχουν ερευνηθεί τον τελευταίο καιρό επικεντρώνονται στη συμπερίληψη δεδομένων PMUs (phasor measurement units) στον αλγόριθμο. Για παράδειγμα, ορισμένοι αλγόριθμοι όπως αυτοί που έχουν αναπτυχθεί στα [20], [21], [22] εκμεταλλεύονται τις πληροφορίες των γωνιών φάσης με μετρήσεις από PMUs για να απλοποιήσουν τη διαδικασία υπολογισμού και να αυξήσουν την αποδοτικότητα της μεθόδου.

Ένα πρόβλημα που πρέπει να αντιμετωπιστεί όταν χρησιμοποιούνται διάφοροι τύποι μετρήσεων είναι η έλλειψη συγχρονισμού. Για να ενσωματωθούν μη συγχρονισμένες μετρήσεις, μπορεί να εισαχθεί ένας τελεστής συγχρονισμού που επιτρέπει τον επανασυγχρονισμό των μετρήσεων στο χρονικό αναφοράς [23]. Οι περιορισμένες μετρήσεις πραγματικού χρόνου προκαλούν πρόβλημα παρατηρησιμότητας, ενώ πολλές μετρήσεις απαιτούν υποδομή επικοινωνίας με μεγάλο εύρος ζώνης και υψηλή αξιοπιστία, οδηγώντας σε υπερφόρτωση δεδομένων και υψηλό κόστος [24]. Με επαρκή αριθμό μετρήσεων έτσι ώστε να καλύπτεται η παρατηρησιμότητα του συστήματος, μπορεί να χρησιμοποιηθεί συμπύεση δεδομένων, έτσι ώστε τα μεταδιδόμενα δεδομένα μέτρησης να μειωθούν [25].

Η δεύτερη κατηγορία δεδομένων που χρησιμοποιούνται για DSSE αποτελούν οι ψευδομετρήσεις, όπως προαναφέρθηκε. Οι ψευδομετρήσεις της ενέργειας που εισέρχεται σε σταθμούς μπορεί να θεωρηθεί ότι ανήκουν σε κανονικές κατανομές με μέση τιμή τα δεδομένα χρέωσης των πελατών ή τα τυπικά προφίλ φορτίου. Επίσης χρησιμοποιούνται καμπύλες φορτίου πελατών κατηγοριοποιημένες με στοχαστικό περιεχόμενο. Η συμπεριφορά των καταναλωτών όμως μπορεί να αλλάξει λόγω των νέων ανατιμολογήσεων και των ανανεώσιμων πηγών ενέργειας, που τους επιτρέπει από πελάτες να γίνονται παραγωγοί, και επομένως οι συμβατικές τεχνικές μοντελοποίησης φορτίου πρέπει να προσαρμοστούν για να ληφθούν υπόψη αυτές οι νέες αβεβαιότητες. Το γκαουσιάνο / κανονικών κατανομών μοντέλο μείγματος (GMM), μια συνδυαστική τεχνική πολλαπλών κανονικών κατανομών, μπορεί να χρησιμοποιηθεί για να αναπαραστήσει τη συνάρτηση πυκνότητας πιθανότητας φορτίου εάν τα φορτία δεν ακολουθούν καμία συνάρτηση κατανομής (βλέπε παρακάτω).

Η κατηγοριοποίηση των πελατών με τη στατιστική επεξεργασία ιστορικών δεδομένων μπορεί να βοηθήσει στην κατανομή δεδομένων φορτίων επιλεγμένων κόμβων μεταξύ μη μετρήσιμων κόμβων [26], [27].

Επίσης η ανάλυση συσχέτισης και η δυναμική ανάλυση των μετρήσεων πραγματικού χρόνου στους υποσταθμούς και των μη ελεγχόμενων μεταβλητών σε άλλους κόμβους με παράλληλη χρήση μιας συνάρτησης πυκνότητας πιθανότητας φορτίου έχει χρησιμοποιηθεί για τη βελτίωση της μοντελοποίησης των ψευδομετρήσεων (φορτίων), [26], [27].

Τα δεδομένα των έξυπνων μετρητών γενικά ενημερώνονται λιγότερο συχνά (π.χ. κάθε 15 λεπτά) και συνήθως περιλαμβάνουν σημαντικές καθυστερήσεις (έως και μια μέρα). Χωρίς αποτελεσματική προεπεξεργασία, αυτά τα δεδομένα δεν μπορούν να χρησιμοποιηθούν άμεσα ως είσοδοι σε μια διαδικασία εκτίμησης κατάστασης.

Τέλος, την τρίτη κατηγορία αποτελούν οι εικονικές μετρήσεις. Οι εικονικές μετρήσεις κατά κύριο λόγο αφορούν μηδενικής εγχύσεις ισχύος σε κόμβους του δικτύου που δεν τροφοδοτούν φορτία και δεν συνδέουν κάποια μονάδα διεσπαρμένης παραγωγής (κόμβοι μηδενικής ισχύος). Η ανάθεση υψηλών βαρών στις εικονικές μετρήσεις και χαμηλών βαρών στις ψευδομετρήσεις μπορεί να προκαλέσει να προβλήματα στη λειτουργία μεθόδων όπως αυτή των σταθμισμένων ελαχίστων τετραγώνων. Έτσι, έχει προταθεί η χρήση πολλαπλασιαστών Lagrange για τη διαχείριση των εικονικών μετρήσεων [29], [30].

Βέλτιστη τοποθέτηση PMUs - Βιβλιογραφική Ανασκόπηση

Η παρατηρησιμότητα ενός δικτύου διανομής είναι απαραίτητη για να λειτουργεί με ασφάλεια και αποτελεσματικότητα. Η θέση των μ -PMU μπορεί να συμβάλει καίρια στην παρατηρησιμότητα του δικτύου, ωστόσο τα μ -PMUs όπως προαναφέρθηκε είναι αρκετά δαπανηρή και είναι πρακτικά αδύνατο να τοποθετηθεί ένας μεγάλος αριθμός τους σ'ένα δίκτυο διανομής.

Θεωρούμε πως έχουμε πλήρη παρατηρησιμότητα όταν όλοι οι φασιθέτες των τάσεων των ζυγών και των ρευμάτων των γραμμών είναι μοναδικά προσδιορισμένα. Ένα μ -PMU μετρά τους φασιθέτες της τάσης του ζυγού στον οποίο έχει εγκατασταθεί και τους φασιθέτες του ρεύματος των γραμμών που βρίσκονται συνδεδεμένες σε αυτόν τον ζυγό. Οι φασιθέτες τάσης των παρακείμενων ζυγών σε αυτόν μπορούν να υπολογιστούν χρησιμοποιώντας τις παραμέτρους των γραμμών. Ως εκ τούτου, όλοι οι που συνδέονται άμεσα με αυτόν στον οποίο έχει εγκατασταθεί ένα μ -PMU θεωρούνται παρατηρήσιμοι.

Στην προτεινόμενη λύση του προβλήματος η οποία παρατίθεται σε αυτή την εργασία χρησιμοποιούνται δεδομένα από SMD σε πραγματικό χρόνο, ωστόσο όπως προαναφέρθηκε, δεν είναι οικονομικά εφικτή η τοποθέτηση μεγάλου αριθμού SMDs. Έτσι λοιπόν καθίσταται αναγκαία η δημιουργία ενός αλγορίθμου που θα βρίσκει συστηματικά τις καλύτερες τοποθεσίες SMDs σ'ένα δίκτυο έτσι ώστε μ'ένα μικρό αριθμό SMDs να παίρνουμε ακριβείς εκτιμήσεις τις μετρήσεις οι οποίες επηρεάζουν περισσότερο την ακρίβεια της εκτίμησης τοπολογίας και κατάστασης.

Ποικίλες μέθοδοι έχουν προταθεί για την ανίχνευση του ελάχιστου αριθμού μ -PMUs έτσι ώστε να υπάρχει πλήρης παρατηρησιμότητα του δικτύου. Για παράδειγμα, για να βρεθεί ο ελάχιστος αριθμός μ -PMU για πλήρη παρατηρησιμότητα, μια γραφική θεωρητική μέθοδος προτείνεται στο [4]. Η μέθοδος αυτή βασίζεται στη συλλογή του συνόλου των κορυφών και

των ακμών. Αρχικά, το δίκτυο διανομής μετατρέπεται σε γράφο θεωρώντας τις γραμμές ως ακμές και τους ζυγούς ως κορυφές. Ελάχιστο γεννητικό δέντρο, (ένα δέντρο με κορυφές αυτές του γράφου και ακμές ένα υποσύνολο των ακμών του γράφου, τέτοιο ώστε το άθροισμα των βαρών τους να είναι το ελάχιστο δυνατό), θα αποτελείται έτσι από δύο τύπους κορυφής: κρεμαστές κορυφές, δηλαδή αυτές που στην «γειτονιά» δεν υπάρχει άλλη κορυφή, και κομμένες κορυφές, δηλαδή κορυφές που αν κοπούν – αποσυνδεθούν από τον γράφο μαζί με τις προσπίπτουσες γραμμές τους αποσυνδέουν τον γράφο αυτό. Η προτεινόμενη μέθοδος τοποθέτησης διατυπώνεται σε δύο βήματα: (i) Επιλέξτε τις κρεμαστές κορυφές και (ii) επιλέξτε τις κομμένες κορυφές που συνδέονται με τις κρεμαστές ως θέσεις εγκατάστασης ενός μ -PMU. Οι μη παρατηρήσιμες κορυφές ομαδοποιούνται και το μ -PMU εγκαθίσταται στην κορυφή που καθιστά παρατηρήσιμη κάθε ομάδα.

Μία άλλη υβριδική μέθοδος που έχει προταθεί στο [31], βασίζεται σε αλγόριθμο παγκόσμιας αναζήτησης για να επιλεγθούν οι βέλτιστες τοποθεσίες των μ -PMUs. Αυτή η μέθοδος έχει τρία κύρια βήματα: (i) Αναγνώριση υποψηφίων τοποθεσιών, (ii) αναζήτηση ελάχιστου αριθμού μ PMUs, και (iii) σύγκριση περιττών PMUs. Η προτεινόμενη εδώ μέθοδος τοποθέτησης βασίζεται στο κριτήριο ελάχιστου σφάλματος εκτίμησης κατάστασης (MSEE). Μια άλλη πρόταση για τη βέλτιστη τοποθέτηση μ -PMUs για πλήρη παρατηρησιμότητα, [6], περιορίζει την ακτινική τοπολογία των δικτύων διανομής με βάση την κλίμακα των δικτύων – δηλαδή, μειώνει το μέγεθος του δικτύου φαινομενικά κατά την μελέτη του. Η μέθοδος αυτή βρίσκει μια λύση πιο γρήγορα από τη μέθοδο εξαντλητικής αναζήτησης με σχεδόν καθόλου μείωση της απόδοσης, ωστόσο, καταναλώνει προοδευτικά περισσότερο χρόνο όσο το μέγεθος του δικτύου αυξάνεται. Όμως χρησιμοποιώντας το μειωμένο κατά την κλίμακα σύστημα μπορεί η διαδικασία αυτή να επιταχυνθεί.

Τέλος, μία μέθοδος που έχει προταθεί εδώ [5] χρησιμοποιεί το συνδυαστικό σύστημα παρακολούθησης με τα μ -PMUs και μετρήσεις ροής ισχύος (PFMs) για την ελαχιστοποίηση του κόστους εγκατάστασης, εξασφαλίζοντας παράλληλα πλήρη παρατηρησιμότητα. Η μέθοδος αυτή βασίζεται σε έναν αλγόριθμο μη γραμμικού προγραμματισμού, με την βέλτιστη τοποθέτηση των μ -PMUs να επιτυγχάνεται μέσω της αξιολόγησης των πληροφοριών (ψευδομετρήσεων φορτίου κυρίως) και της χρήσης ενός αλγορίθμου εξαντλητικής αναζήτησης για να επιλεγούν οι βέλτιστοι κόμβοι τοποθέτησης.

Διατύπωση Προβλήματος – Βιβλιογραφική Ανασκόπηση DSSE

Ουσιαστικά, οι αλγόριθμοι εκτίμησης κατάστασης παρέχουν ένα μέσο για την εξάλειψη ή την ελαχιστοποίηση των θορύβων και σφαλμάτων μέτρησης και πιθανών σφαλμάτων τοπολογίας που διαφορετικά θα εμπόδιζαν τον ακριβή προσδιορισμό των τιμών της κατάστασης του συστήματος σε όλους τους ζυγούς.

Λόγω όλων αυτών των διαφορών που αναφέρθηκαν παραπάνω των δικτύων διανομής με τα συστήματα μετάδοσης, η διατύπωση του προβλήματος της DSSE έχει σημαντικές αποκλίσεις από την συμβατική

εκτίμηση κατάστασης. Η κύρια διαφορά είναι ο τρόπος μοντελοποίησης της συνάρτησης μέτρησης ($x(y)$ στον τύπο σελίδας 38) στην DSSE, καθώς αυτή η συνάρτηση αντικατοπτρίζει τα ισοζύγια ροής ισχύος στο ηλεκτρικό σύστημα. Έτσι, βασισμένη στην επιλογή των καταστάσεων και των μετρούμενων μεγεθών, της AC έναντι της DC ροής ισχύος και της αναπαράστασης των φάσεων στα ισοζύγια ροής ισχύος (για εφαρμογή σε μη ισορροπημένα συστήματα), η συνάρτηση μέτρησης μπορεί να έχει διάφορες μορφές. Με αυτό τον τρόπο, η DSSE μπορεί να διαχωριστεί σε δύο βασικές θεωρητικές υποκατηγορίες με βάση την διατύπωση [43].

Η πρώτη είναι η DSSE βασισμένη στις τάσεις. Παραδοσιακά, τα μέτρα και οι γωνίες φάσης των τάσεων στους ζυγούς σε πολικές ή ορθογώνιες μορφές ορίζονται ως μεταβλητές κατάστασης για την εκτίμηση κατάστασης στα συστήματα μετάδοσης. Αυτή η κλασική προσέγγιση μπορεί να χρησιμοποιηθεί επίσης στην DSSE.

Αντίθετα, η δεύτερη προσέγγιση εκτίμησης κατάστασης βασίζεται στα ρεύματα των κλάδων (BCSE – Branch Current based SE). Πολλές εργασίες σήμερα υιοθετούν τα ρεύματα των κλάδων ως μεταβλητές κατάστασης και ίσως, αυτός να είναι πιο αποτελεσματικός τρόπος διατύπωσης διατύπωσης της εκτίμησης κατάστασης για δίκτυα διανομής. Αυτό συμβαίνει επειδή τα ρεύματα σχετίζονται πιο άμεσα με τη ροή ηλεκτρικής ενέργειας, λόγω της ακτινικής λειτουργίας των δικτύων διανομής και της πολύπλοκης τοπολογίας τους με πολλές διακλαδώσεις. Ωστόσο, η προσέγγιση αυτή του προβλήματος θα πρέπει να διερευνηθεί παραπάνω προτού φανούν σαφή συμπεράσματα.

Για την εκπόνηση αυτής της εργασίας, εκτελέστηκε DSSE με βάση τις τάσεις καθώς η εφαρμογή της έτσι ήταν απλούστερη αφού ως μεταβλητές κατάστασης είχαν θεωρηθεί οι φασιθέτες τάσεως κάθε ζυγού. Επίσης, τα δεδομένα της ροής του ρεύματος είχαν χρησιμοποιηθεί ήδη κατά την διαδικασία της εκτίμησης τοπολογίας, οπότε κρίθηκε πιο αποδοτική η αξιοποίηση και των υπολοίπων δεδομένων που μπορούν να προσφέρουν τα SMDs.

Από την άλλη πλευρά, οι προσεγγίσεις της DSSE μπορούν να χωριστούν σε πέντε κατηγορίες ανάλογα με τον τρόπο προσέγγισης της επίλυσης του προβλήματος SE που επιλέγεται, με βάση την διάκριση που γίνεται στο [42]:

- Στατική DSSE σταθμισμένων ελαχίστων τετραγώνων: Οι εκτιμητές σταθμισμένων ελαχίστων τετραγώνων (WLS) είναι οι πιο δημοφιλείς και έχουν καταβληθεί σημαντικές προσπάθειες για τη μείωση των υπολογιστικών απαιτήσεων αυτής της μεθόδου. Οι κύριες διαφορές μεταξύ των προτάσεων είναι η επιλογή των μεταβλητών κατάστασης, οι απλουστεύσεις για την επιτάχυνση της εκτίμησης και οι τεχνικές για την ενσωμάτωση ετερογενών μετρήσεων. Δύο κύριες υποκατηγορίες με βάση την επιλογή των μεταβλητών κατάστασης, αποτελούν η τάση των κόμβων και τα ρεύματα διακλάδωσης. Και οι δύο μπορούν να διατυπωθούν σε πολικές και ορθογώνιες συντεταγμένες.
- DSSE με προσαρμογή φορτίου: τα φορτία προσαρμόζονται με βάση μιας τεχνικής μοντελοποίησης φορτίου η οποία βασίζεται στις καμπύλες προφίλ φορτίου του πελάτη. Οι μέθοδοι προσαρμογής φορτίου SE ρυθμίζουν γενικά τις τιμές φορτίου (έγχυση ρεύματος ζυγών ή έγχυση ισχύος) έτσι ώστε οι τιμές να

συμμορφώνονται με τις μετρήσεις. Τα δεδομένα των μετρήσεων ορίζονται ως περιορισμοί της λύσης στον αλγόριθμο.

- Ανθεκτικές μέθοδοι DSSE: Η ανίχνευση και ο εντοπισμός των λανθασμένων δεδομένων είναι ζωτικής σημασίας για την απόκτηση ακριβή αποτελεσμάτων μέσω της SE. Ένας εκτιμητής ανθεκτικότητας (robust estimator) γενικά καταστέλλει τις επιρροές των δεδομένων αυτών κατά τη διαδικασία επίλυσης μειώνοντας τα βάρη που αντιστοιχούν στα σημεία που υποπτεύεται η ύπαρξη λανθασμένων δεδομένων. Όταν οι εκτιμώμενες καταστάσεις παραμένουν ανεπηρέαστες από μεγάλες αποκλίσεις σε ένα περιορισμένο αριθμό μετρήσεων, ο εκτιμητής που χρησιμοποιείται σε αυτή την περίπτωση μπορεί να θεωρηθεί στατιστικά ανθεκτικός (robust).
- Δυναμικές μέθοδοι DSSE: Η δυναμική εκτίμηση κατάστασης, επίσης γνωστή ως εκτίμηση κατάστασης υποβοηθούμενη από προγνώσεις, (Forecasting-Aided SE – FASE), είναι μια αναδρομική μέθοδος εκτίμησης που βασίζεται σε αρκετά στιγμιότυπα μετρήσεων σε μια χρονική ακολουθία. Εάν η απαίτηση υπολογισμού ενός μεγάλου συστήματος αποτελεί πρόβλημα και οι μετρήσεις φθάνουν σε διαφορετικές συχνότητες, οι πιο πρόσφατα εισερχόμενες μετρήσεις μπορούν να υποβληθούν σε επεξεργασία μαζί με τις διαθέσιμες εκ των προτέρων εκτιμήσεις (προγνώσεις) και στη συνέχεια να χρησιμοποιηθούν για την πρόβλεψη της μεταβολής της κατάστασης μεταξύ πλήρων εκτελέσεων του αλγορίθμου DSSE.
- Κατανεμημένες μέθοδοι DSSE: Καθώς τα δίκτυα διανομής αποτελούνται από πολυάριθμους περιφερειακούς υποσταθμούς, τροφοδότες και κόμβους, απαιτείται υψηλός υπολογιστικός χρόνος για να εκτελεστεί SE ολόκληρο το δίκτυο. Η κατανεμημένη DSSE, γνωστή και ως εκτίμηση κατάστασης πολλαπλών περιοχών (MASE – multiple area SE) διαιρεί το δίκτυο διανομής σε διάφορες υποπεριοχές ανάλογα με κριτήρια γεωγραφικά και τοπολογικά. Στη συνέχεια, ανάλογα με τα διάφορα σημεία μέτρησης που έχουν επιλεγεί, επιλύεται το πρόβλημα με τοπικούς εκτιμητές. Αυτή η μέθοδος μπορεί να εκτελεί λοιπόν πολλές DSSE παράλληλα, για διάφορα τμήματα του δικτύου.

Οι αλγόριθμοι DSSE μπορούν επίσης να διαφοροποιηθούν ανάλογα με τους τύπους μετρήσεων που χρησιμοποιούν και τον τρόπο ενσωμάτωσής τους στο μοντέλο του εκτιμητή. Όπως προαναφέρθηκε στο τμήμα για την παρατηρησιμότητα, πολύ συχνά στα δίκτυα διανομής χρησιμοποιούνται ψευδομετρήσεις, δηλαδή τεχνητά παραγόμενα δεδομένα βασισμένα στο ιστορικά δεδομένα των δικτύων διανομής. Γενικότερα, κατά την DSSE απαιτούνται δεδομένα από πολλές πηγές από το σύστημα που διαχειρίζεται το δίκτυο διανομής.

Μια βασική προσέγγιση είναι η χρήση τυπικών προφίλ φορτίου για τη δημιουργία ψευδομετρήσεων, [44], όμως δεδομένου ότι αυτά τα σημεία δεδομένων δεν είναι πολύ ακριβή, εισάγουν υψηλά επίπεδα διακύμανσης

στον πίνακα βαρών (W), πράγμα που μπορεί να οδηγήσει σε λανθασμένες καταστάσεις το πρόβλημα εκτίμησης κατάστασης.

Η δημιουργία ψευδομετρήσεων γενικά μπορεί να διαχωριστεί σε δύο κατηγορίες διαδικασιών, αυτές που βασίζονται σε πιθανοτικές και στατιστικές προβλέψεις και αυτές που βασίζονται σε τεχνικές της μηχανικής μάθησης.

Οι αλγόριθμοι μηχανικής μάθησης έχουν επίσης τραβήξει την επιστημονική προσοχή για την επίλυση προβλημάτων DSSE, συμπεριλαμβανομένης της αντιμετώπισης του προβλήματος δημιουργίας ψευδο-μετρήσεων ενεργού/αέργου ισχύος και της αξιολόγησης της αβεβαιότητας τους. Στο [50] και χρησιμοποιούνται Πιθανοτικά Νευρωνικά Δίκτυα (PNNs) και συσταδοποίηση χρησιμοποιώντας ιστορικά δεδομένα από έξυπνους μετρητές (SM) και καμπύλες διάρκειας φορτίου. Αυτή η προσέγγιση μπορεί να χρησιμοποιηθεί και για τυχαίες κατανομές φορτίων κι εμφανίζει την συσχέτιση μεταξύ της κατανάλωσης και του ενεργειακού κόστους αλλά εξαρτάται υπερβολικά πολύ από τον αριθμό των συστάδων κατά την ομαδοποίηση.

Αντίστοιχα, στο [51] ακολουθείται μια μίξη τεχνητών νευρωνικών δικτύων ANNs με GMMs χρησιμοποιώντας ως βάση δεδομένα γενικής κατανάλωσης. Αυτή η μέθοδος έχει πλεονέκτημα πως μπορεί να εφαρμοστεί ανεξαρτήτως της κατανομής των σφαλμάτων εκτίμησης αλλά έχει υψηλό υπολογιστικό κόστος. Στο [52] χρησιμοποιήθηκαν τεχνικές συσταδοποίησης (clustering) σε συνδυασμό με GMMs και ιστορικά δεδομένα από έξυπνους μετρητές μαζί με καμπύλες διάρκειας φορτίου, αλλά ενώ ξανά δεν χρειάζεται τα σφάλματα να ακολουθούν κάποια κατανομή για να εφαρμοστεί κι ενώ μπορεί να εκφράσει τις χρονικές διαφορές από μήνα σε μήνα της συσχέτισης των καταναλώσεων, είναι ξανά δύσκολη υπολογιστικά και εξαρτάται και από των αριθμών των συστάδων που έχει επιλεγεί και από τις κατανομές που απαρτίζουν τα GMMs. Τέλος, στο [53] χρησιμοποιήθηκαν εξιδεικευμένα δίκτυα Narx (Nonlinear Auto-Regressive with eXogenous inputs – Μη γραμμικά αυτοπαλινδρομικά δίκτυα με εξωγενή δεδομένα – δηλαδή που δεν ανήκουν στην προς πρόβλεψη χρονική σειρά). Αυτά με βάση ιστορικά δεδομένα PMUs, μεταβλητές καιρικών συνθηκών και καμπύλες διάρκειας φορτίου παράγουν ψευδομετρήσεις οι οποίες δεν εξαρτώνται καθόλου από την κατανομή των φορτίων και μπορούν εύκολα να προσαρμοστούν αν αυτή αλλάξει, με κόστος τα σφάλματα να ακολουθούν υποχρεωτικά κανονική κατανομή ως προϋπόθεση.

Εν κατακλείδι, όπως έχει δειχθεί από τα [54] και [55] πολλές φορές έχουν χρησιμοποιηθεί προβλέψεις φορτίων για τη δημιουργία ψευδομετρήσεων, αλλά όπως έχει δειχθεί στο [56] αν χρησιμοποιηθούν γενικά ψευδομετρήσεις σε πραγματικό χρόνο, η ακρίβεια της εκτίμησης κατάστασης μειώνεται.

Έτσι, κρίνεται αναγκαία μια νέα προσέγγιση για να μπορέσει το σύστημα να θεωρηθεί παρατηρήσιμο με στόχο να εκτελείται μια ακριβής εκτίμηση κατάστασης όταν τροφοδοτούνται μετρήσεις σε πραγματικό χρόνο. Για τον σκοπό αυτό, αναπτύσσεται μία μεθοδολογία βασισμένη στην τεχνητή νοημοσύνη, η οποία αξιοποιεί δεδομένα προσομοιώσεων monte carlo, συμπεριλαμβανομένων μετρήσεων πληροφοριών τοπολογίας, για την εκμάθηση των σύνθετων μοτίβων και των σχέσεων μεταξύ των μεταβλητών κατάστασης. Αναλύοντας αυτά τα δεδομένα, οι αλγόριθμοι τεχνητής

νοημοσύνης μπορούν να εντοπίσουν αυτά τα κρυμμένα μοτίβα, τις συσχετίσεις και τις εξαρτήσεις που μπορεί να μην είναι εμφανείς με τη χρήση παραδοσιακών αναλυτικών μεθόδων, κι έτσι να εκτελέσουν SE πιο αποδοτικά. Τα ειδικά χαρακτηριστικά των δικτύων διανομής, όπως η πολυπλοκότητα του, οι μη γραμμικές αλληλεπιδράσεις μεταξύ διαφόρων στοιχείων, όπως μετασχηματιστές, διακόπτες και πόροι κατανεμημένης παραγωγής, οι εγγενής αβεβαιότητες και διαταραχές στις οποίες υπόκεινται και τέλος το αρκετά μεγάλο συνήθως μέγεθος τους, καθιστούν την τεχνητή νοημοσύνη πιο αξιόπιστη επιλογή για εκτέλεση SE, καθώς μπορεί να ανταπεξέλθει στα προβλήματα αυτά πολύ πιο εύκολα από τις παραδοσιακές μεθόδους.

Στο [57] προτείνεται μια Μπεϋζιανή προσέγγιση ανάλογη με αυτήν που πραγματοποιήθηκε σε αυτή την εργασία, στην οποία εκπαιδεύεται ένα βαθύ νευρωνικό δίκτυο για την παράκαμψη της έλειψης παρατηρησιμότητας του δικτύου σε πραγματικό χρόνο, το οποίο επίσης προσεγγίζει έναν εκτιμητή MMSE. Ωστόσο, η προσέγγιση αυτή δεν επικυρώθηκε για μη ισορροπημένα συστήματα διανομής. Στο [58] ένα τεχνητό νευρωνικό δίκτυο κατασκευάστηκε για την εκτέλεση τριφασικής DSSE σε μη εξισορροπημένα φορτία. Ωστόσο, οι μετρήσεις των έξυπνων μετρητών δεν λήφθηκαν υπόψη στην ανάλυση αυτή, (χρησιμοποιήθηκαν μόνο μετρήσεις από micro-PMUs). Στο [59] αναπτύχθηκε ένας εκτιμητής κατάστασης αραιάς παρακολούθησης για μη ισορροπημένα συστήματα διανομής που δεν είναι πλήρως παρατηρήσιμα από D-PMUs. Ωστόσο, απαιτούσε πρόσθετες πληροφορίες από διάφορα συμβάντα που μπορεί να μην είναι πάντα διαθέσιμες και περιορίστηκε σε ακτινικά δίκτυα. Στο [60] προτάθηκε μια διαδικασία με τη συγχώνευση κατά Bayes, έτσι ώστε να ληφθούν υπόψη οι διαφορετικές χρονικές πτυχές των καταστάσεων και των μετρήσεων. Όμως, εξαιτίας του μεγάλου υπολογιστικού φόρτου της διαδικασίας αυτής δεν είναι δυνατόν να χειριστεί μη γκαουσιανά φορτία και θόρυβο μετρήσεων. Επιπλέον, σε όλες αυτές τις μεθόδους, [57]-[60], η τοπολογία του συστήματος θεωρήθηκε σταθερή.

Ως προς την εκτίμηση της τοπολογίας, μία παρόμοια μέθοδος με αυτή της εργασίας έχει αναπτυχθεί στο [61], όπου χρησιμοποιήθηκε γραμμικός προγραμματισμός μικτού ακεραίου (mixed integer) για την εκτίμηση τοπολογίας των δικτύων διανομής, όμως, η μεθοδολογία αυτή απαιτούσε μετρήσεις σε πραγματικό χρόνο από μετρητές ροής ρεύματος στις γραμμές και έξυπνους μετρητές, οι οποίοι δεν είναι διαθέσιμοι στους περισσότερους σταθμούς διανομής. Στο [62], προτάθηκε ένα πλαίσιο βελτιστοποίησης βασισμένο σε γράφους για την ανάκτηση της τοπολογίας ακτινικών δικτύων διανομής χρησιμοποιώντας περιορισμένο αριθμό μετρητών πραγματικού χρόνου. Εντούτοις δίκτυα με τοπολογίες πλέγματος και ασύμμετρα πολυφασικά συστήματα διανομής δεν ελήφθησαν υπόψη. Στο [63] ένα πιθανοτικό μοντέλο δικτύου χρησιμοποιήθηκε για την εκτίμηση της τοπολογίας, όμως βασίζεται σε δεδομένα έξυπνων μετρητών, γεγονός που την καθιστά ακατάλληλη, όπως προαναφέρθηκε, για την εκτίμηση της τοπολογίας του δικτύου σε πραγματικό χρόνο. Στο [64] χρησιμοποιήθηκε μια μέθοδος επαλήθευσης με βάση χρονοσειρές για την παρακολούθηση των αλλαγών της τοπολογίας από μια ροή μετρήσεων μικρο-PMU. Ένα ανοιγόκλειμα όμως διακόπτη κάθε φορά καθώς και η εκ των προτέρων επίγνωση της κατάστασης των διακοπών είναι δύο υποθέσεις που τέθηκαν

οι οποίες περιορίζουν τη χρησιμότητα αυτής της μεθόδου. Στο [65] προτάθηκε ένα πλαίσιο εκτίμησης τοπολογίας βασισμένο στην μηχανική μάθηση. Ωστόσο, η ανάγκη για μετρήσεις των ρευμάτων στους κόμβους, των τάσεων και της γωνίας του συντελεστή ισχύος κάθε φάσης περιορίζει την εφαρμογή του σε πραγματικό χρόνο. Στο [66] μια αριθμητική μέθοδος δύο βημάτων αναπτύχθηκε για την εκτίμηση της τοπολογίας, η οποία όμως είναι πολύ αργή για παρακολούθηση σε πραγματικό χρόνο και περιορίζεται σε ισορροπημένα δίκτυα. Τέλος, στο [67] προτάθηκε μία μέθοδος που εκτελεί εκτίμηση τοπολογίας σε μη ισορροπημένα δίκτυα διανομής σε πραγματικό χρόνο. Ωστόσο, γίνεται χρήση δεδομένων προβλεπόμενου φορτίου ως ψευδομετρήσεις, γεγονός που μπορεί να επιδεινώσει την απόδοσή του, όπως προαναφέρθηκε. Καμία εκ των παραπάνω μεθόδων, [61]-[67], δεν αναπτύχθηκε μια συστηματική προσέγγιση για τον εντοπισμό των θέσεων τοποθέτησης μετρητών έτσι ώστε ενισχύεται η απόδοση της εκτίμησης κατάστασης.

Προτεινόμενη Μεθοδολογία

Μετά από τη μελέτη του προβλήματος της εκτίμησης κατάστασης και την εξέταση της σχετικής βιβλιογραφίας, στην παρούσα εργασία εξετάζεται μία μεθοδολογία εκτίμησης κατάστασης δικτύων διανομής κατά την οποία αρχικώς εκτιμάται σε πραγματικό χρόνο η τοπολογία του δικτύου και κατόπιν η κατάσταση λειτουργίας του, δηλαδή οι φασιθέτες τάσης όλων των κόμβων του. Ως είσοδο του ο προτεινόμενος εκτιμητής κατάστασης δέχεται δεδομένα από βέλτιστα τοποθετημένες μονάδες SMD.

Πλεονέκτημα της μεθοδολογίας η οποία βασίζεται σε τεχνητά νευρωνικά δίκτυα, είναι ότι δεν απαιτεί πλήρη παρατηρησιμότητα του δικτύου σε πραγματικό χρόνο από SMDs [1]. Έτσι, δεν χρησιμοποιεί μετρήσεις άλλου τύπου ούτε ψευδομετρήσεις, με αποτέλεσμα να μην υπάρχει πρόβλημα έλλειψης συγχρονισμού. Η εκπαίδευση των τεχνητών νευρωνικών δικτύων γίνονται με δεδομένα που παράγονται μέσω προσομοιώσεων Monte Carlo.

Αρχικά, προτείνεται μια μέθοδος εκτίμησης τοπολογίας βασισμένη σε βαθειά νευρωνικά δίκτυα η οποία ουσιαστικά εκτιμά την κατάσταση των διακοπών του συστήματος. Στην συνέχεια, όταν η τοπολογία βρεθεί πραγματοποιείται DSSE για να βρεθούν τα μέτρα και οι γωνίες φάσης των ζυγών του συστήματος με γρήγορο, χρονικά συγχρονισμένο τρόπο για τόσο ακτινικά όσο και βροχοειδή δίκτυα διανομής. Τέλος, χρησιμοποιείται μεταφορά μάθησης (transfer learning) για να ληφθούν υπόψη οι επιδράσεις των αλλαγών της τοπολογίας στην DSSE. Έτσι, όταν συμβαίνει μια αλλαγή στην τοπολογία, πρώτα ανιχνεύεται αυτή η αλλαγή χρησιμοποιώντας ένα νευρωνικό δίκτυο που έχει κατασκευαστεί για εντοπισμό τοπολογίας, και στη συνέχεια χρησιμοποιείται μεταφορά μάθησης μέσω της τεχνικής του λεπτομερούς συντονισμού για την ενημέρωση του νευρωνικού δικτύου που κάνει DSSE σε πραγματικό χρόνο για τη νέα αυτή τοπολογία. Με αυτόν τον τρόπο, παρακάμπτεται η ανάγκη για μεγάλες βάσεις δεδομένων και μεγάλων περιόδων χρόνου για την επανεκπαίδευση του νευρωνικού δικτύου που κάνει DSSE.

Επιπλέον, παρουσιάζεται μια προσέγγιση για την βέλτιστη τοποθέτηση μετρητών (SMDs) σε βέλτιστα σημεία ώστε το δίκτυο να είναι παρατηρήσιμο για τις διαφορετικές τοπολογίες λειτουργίας του και να ενισχύεται η ακρίβεια της εκτίμησης κατάστασης. Η μέθοδος αυτή ανήκει στις άπληστες μεθόδους αναζήτησης, κι έτσι η στρατηγική αυτή επιλογής μετρήσεων για την τοποθέτηση SMD δεν είναι εγγυημένα βέλτιστη. Ωστόσο, θεωρείται αποδεκτή επειδή το πρόβλημα που επιλύεται εδώ είναι άλυτο με την κλασική έννοια του όρου (δηλ. έχει μεγαλύτερο αριθμό άγνωστους από τους γνωστούς), επομένως κι έτσι είναι πιθανό να μην υπάρχει κάποιος βέλτιστος αλγόριθμος τοποθέτησης που να δίνει σταθερά τα καλύτερα αποτελέσματα υπό όλες τις συνθήκες λειτουργίας. Επιπλέον, η εστίαση του προβλήματος δεν είναι η βελτιστοποίηση της τοποθέτησης αισθητήρων, αλλά η επίτευξη λογικών αποτελεσμάτων εκτίμησης τοπολογίας και εκτίμησης κατάστασης, τα οποία όπως θα δειχθεί αυτή η στρατηγική τοποθέτησης είναι σε θέση να παρέχει.

Ωστόσο λόγω του χρονοβόρου της διαδικασίας, διερευνώνται και άλλες μέθοδοι τοποθέτησης έξυπνων μετρητών σε βέλτιστα σημεία, όπως ο αλγόριθμος Random Forest.

Κεφάλαιο III

Τεχνητή Νοημοσύνη

Η τεχνητή νοημοσύνη (TN) αφορά τη ‘νοημοσύνη’ μηχανών ή λογισμικού που έχουν κατασκευαστεί από ανθρώπους. Πρόκειται δηλαδή για ένα πεδίο μελέτης στην επιστήμη των υπολογιστών που έχει ως στόχο την μελέτη κι ανάπτυξη υπολογιστικών συστημάτων που μιμούνται στοιχεία της ανθρώπινης νόησης τα οποία υπονοούν έστω και στοιχειώδη ευφυΐα. Η τεχνητή νοημοσύνη χρησιμοποιεί στοιχεία από ποικίλες επιστήμες όπως η πληροφορική, η ψυχολογία, η φιλοσοφία, η νευρολογία κι η γλωσσολογία με στόχο τη σύνθεση ευφυούς συμπεριφοράς, με στοιχεία συλλογιστικής, μάθησης και προσαρμογής στο περιβάλλον.

Η τεχνητή νοημοσύνη ιδρύθηκε ως ακαδημαϊκή αντικείμενο το 1956, όμως επειδή τα αποτελέσματα δεν ήταν τα αναμενόμενα συνοδεύτηκε γρήγορα από απογοήτευση και απώλεια χρηματοδότησης. Αρχικά αναπτύχθηκαν απλοί αλγόριθμοι λήψης αποφάσεων που παρείχαν τη δυνατότητα επίλυσης σύνθετων μαθηματικών εκφράσεων και επεξεργασίας σειρών λέξεων. Στη συνέχεια αναπτύχθηκαν πολύπλοκα συστήματα χρησιμοποιώντας λογικούς κανόνες και συλλογιστικούς αλγόριθμους που μιμούνται μερικές από τις λειτουργίες του ανθρώπινου νου. Έτσι ξεκίνησε η άνοδος συστημάτων που έμαθαν τους «κανόνες» ενός συγκεκριμένου τομέα γνώσης, αλλά δεν ήταν ικανά για πολύπλοκους συλλογισμούς και δεν μπορούσαν να μάθουν νέους κανόνες για να εξελίσσονται και να επεκτείνουν τη δυνατότητα τους για λήψη αποφάσεων. Τελικά η χρηματοδότηση και το ενδιαφέρον αυξήθηκε σημαντικά μετά το 2012, όταν η τεχνική της βαθιάς μάθησης ξεπέρασε όλες τις προηγούμενες σε αποτελεσματικότητα και κόστος και αφού τα νευρωνικά δίκτυα είχαν αναπτυχθεί σε επαρκές επίπεδο έτσι ώστε να επιδεικνύουν την ανωτερότητα τους στην επίλυση προβλημάτων σε σχέση με πιο παραδοσιακούς αλγόριθμους.

Η τεχνολογία της τεχνητής νοημοσύνης χρησιμοποιείται ήδη ευρέως και συνεχώς ενσωματώνεται περισσότερο σχεδόν σε όλους τους τομείς της σύγχρονης ζωής. Ορισμένες προβεβλημένες εφαρμογές περιλαμβάνουν: τις μηχανές αναζήτησης στο διαδίκτυο, (π.χ., Google, Yahoo), τα συστήματα προτάσεων με βάση τα ενδιαφέροντα του χρήστη, (π.χ., χρησιμοποιούμενα από το YouTube, το Skroutz και το Netflix), την κατανόηση κι επεξεργασία της ανθρώπινης φωνής μέσω της οποίας έχουν αναπτυχθεί οι ψηφιακοί βοηθοί, (όπως το Google Assistant, ο Siri και ο Alexa). Ταυτόχρονα όπως επιδεικνύεται και με αυτή την εργασία οι τεχνικές και τα εργαλεία που παρέχει η τεχνητή νοημοσύνη μπορούν να εφαρμοστούν σχεδόν παντού και να βοηθήσουν σημαντικά στην επίλυση προβλημάτων που μέχρι πρότεινος θεωρούταν πολύ δύσκολη ή ανώφελη λόγω του μεγάλου κόστους η επίλυση τους.

Η τεχνητή νοημοσύνη ενσωματώνει τεχνικές μηχανικής μάθησης (machine learning) που χαρακτηρίζονται από ακριβείς μαθηματικούς αλγόριθμους και μεθόδους στατιστικής ανάλυσης.

Μηχανική Μάθηση

Η μηχανική μάθηση είναι ένα υποπεδίο της τεχνητής νοημοσύνης που ασχολείται με την ανάπτυξη και τη μελέτη στατιστικών αλγορίθμων που μπορούν να μαθαίνουν από δεδομένα και να γενικεύουν σε μέχρι στιγμής άγνωστα αλλά ίδιου πλαισίου δεδομένα, εκτελώντας έτσι διεργασίες χωρίς σαφείς οδηγίες. Δηλαδή η μηχανική μάθηση είναι η μελέτη αλγορίθμων που μπορούν να βελτιώσουν αυτόματα την απόδοσή τους σε μια συγκεκριμένη εργασία.

Οι αλγόριθμοι αυτοί λειτουργούν κατασκευάζοντας μοντέλα από πειραματικά δεδομένα, προκειμένου να κάνουν προβλέψεις βασιζόμενες σε αυτά ή να εξάγουν αποφάσεις που θα εκφράζονται ως το αποτέλεσμα. Η στατιστική, καθώς το αντικείμενο του τομέα αυτού είναι η ανάλυση και πρόβλεψη δεδομένων, είναι μια σημαντική πηγή μεθοδολογιών και εργαλείων για το πεδίο αυτό.

Η Μηχανική Μάθηση εφαρμόζεται κυρίως σε υπολογιστικές εργασίες, όπου τόσο ο σχεδιασμός όσο και ο ρητός προγραμματισμός των αλγορίθμων είναι ανέφικτος. Για παράδειγμα, οι προσεγγίσεις της μηχανικής μάθησης έχουν εφαρμοστεί σε μεγάλα μοντέλα επεξεργασίας φυσικής γλώσσας, όρασης υπολογιστών, αναγνώρισης φωνής, φιλτραρίσματος ηλεκτρονικών μηνυμάτων, γεωργίας και ιατρικής, όπου είναι πολύ οικονομικά απαιτητικό να αναπτυχθούν αλγόριθμοι για την εκτέλεση των απαιτούμενων εργασιών. Επίσης χρησιμοποιείται ευρέως σε προβλήματα επιχειρήσεων, που καθιστά χωριστό κλάδο μελέτης.

Οι μαθηματικές βάσεις της MM παρέχονται από μαθηματικές μεθόδους βελτιστοποίησης (δυναμικός προγραμματισμός). Η εξόρυξη δεδομένων είναι ένα παράλληλο πεδίο μελέτης που επικεντρώνεται στην εξερευνητική ανάλυση δεδομένων μέσω μη καθοδηγούμενης μάθησης. Από θεωρητικής άποψης, η "Πιθανώς Προσεγγιστικά Σωστή Μάθηση" ,(Probably approximately correct learning) παρέχει ένα πλαίσιο για την περιγραφή της μηχανικής μάθησης το οποίο στοχεύει και βοηθάει την κατανόηση των δυνατοτήτων και των περιορισμών των διάφορων μοντέλων που χρησιμοποιούνται στις προβλέψεις.

Η αναζωπύρωση του ενδιαφέροντος για τη μηχανική μάθηση οφείλεται κυρίως στον αυξανόμενο όγκο και τις ποικιλίες των διαθέσιμων δεδομένων, στην υπολογιστική επεξεργασία που έχει γίνει φθηνότερη και πιο ισχυρή και στην προσιτή αποθήκευση δεδομένων. Έτσι καθίσταται δυνατή η γρήγορη και αυτόματη παραγωγή μοντέλων που μπορούν να αναλύσουν μεγαλύτερο όγκο πιο σύνθετων δεδομένων και να προσφέρουν ταχύτερα, ακριβέστερα αποτελέσματα - ακόμη και σε πολύ μεγάλη κλίμακα. Με την κατασκευή ακριβών μοντέλων, ένας οργανισμός έχει περισσότερες πιθανότητες να εντοπίσει κερδοφόρες ευκαιρίες - ή να αποφύγει άγνωστους κινδύνους.

Υπάρχουν διάφορα είδη μηχανικής μάθησης. Η εκμάθηση χωρίς επίβλεψη/καθοδήγηση, (Unsupervised Learning), αναλύει μια ροή δεδομένων, εντοπίζει μοτίβα και κάνει προβλέψεις χωρίς κάποια άλλη καθοδήγηση. Οι αλγόριθμοι της εν λόγω μάθησης αναφέρονται ως αυτό-οργανώσιμοι και βασίζονται μόνο σε τοπική πληροφορία καθ' όλη τη διάρκεια της εκπαίδευσης του νευρωνικού δικτύου.

Η εποπτευόμενη εκμάθηση. (Supervised Learning), απαιτεί από έναν άνθρωπο να επισημάνει πρώτα τα δεδομένα εισόδου και έχει δύο βασικές παραλλαγές:

- Τα προβλήματα ταξινόμησης, όπου το πρόγραμμα μαθαίνει να κατηγοριοποιεί τα δεδομένα σε διακριτές κλάσεις ή ετικέτες, για παράδειγμα "spam" ή "not spam" στο φιλτράρισμα email
- και παλινδρόμησης, όπου το πρόγραμμα πρέπει να συμπεράνει μια αριθμητική συνάρτηση με βάση μια αριθμητική είσοδο, δηλαδή να προβλέπει συνεχείς, αριθμητικά αποτελέσματα όπως νούμερα πωλήσεων ή θερμοκρασιών.

Οι περισσότερες διαδικασίες εκπαίδευσης σε νευρωνικά δίκτυα διεξάγονται off line, δηλαδή χρησιμοποιείται όλο το δείγμα προτύπων για την τροποποίηση των τιμών των βαρών (βλέπε πιο κάτω στην ενότητα για τα νευρωνικά δίκτυα), πριν την τελική χρήση του δικτύου ως εφαρμογή. Στους αλγορίθμους εκπαίδευσης off line απαιτείται όλα τα παραδείγματα που χρησιμοποιούνται για την εκπαίδευση του δικτύου να είναι διαθέσιμα εκ των προτέρων. Το γεγονός αυτό αποκλείει την πιθανότητα εισαγωγής νέων πληροφοριών μέσω νέων προτύπων, κατά τη διάρκεια της εκπαίδευσης του μοντέλου. Κατά συνέπεια, αν κάποιος θέλει να εισάγει νέες πληροφορίες στο μοντέλο, θα πρέπει να ξανακάνει ολόκληρη τη διαδικασία εκπαίδευσης, συμπεριλαμβανομένων όλων των παραδειγμάτων. Το πλεονέκτημα των δικτύων που χρησιμοποιούν off line διαδικασίες εκπαίδευσης επικεντρώνεται κυρίως στη δυνατότητα που τους παρέχεται έτσι να δίνουν καλύτερες λύσεις σε δύσκολα προβλήματα.

Υπάρχουν και στρατηγικές εκπαίδευσης για νευρωνικά δίκτυα που εφαρμόζονται online ή με ενημέρωση. Σε αντίθεση με την offline εκπαίδευση όπου το δίκτυο εκπαιδεύεται σε ένα σύνολο δεδομένων που είναι διαθέσιμο εξ αρχής, η online εκπαίδευση επιτρέπει την ενημέρωση του μοντέλου με νέα δεδομένα καθώς αυτά γίνονται διαθέσιμα, καθώς όταν έρχεται ένα νέο παράδειγμα δεδομένων, το μοντέλο ενημερώνεται αμέσως χωρίς την ανάγκη επανεκπαίδευσης σε ολόκληρο το σύνολο των δεδομένων. Αυτό είναι χρήσιμο σε περιπτώσεις όπου τα δεδομένα λαμβάνονται σε πραγματικό χρόνο και το μοντέλο πρέπει να προσαρμόζεται συνεχώς σε νέα παραδείγματα. Αυτές οι μέθοδοι επιτρέπουν την ενημέρωση των βαρών του μοντέλου με κάθε νέα δεδομένα που έρχεται, χωρίς να απαιτείται πλήρης παρασκευή του συνόλου εκπαίδευσης.

Βαθιά Μάθηση

Η βαθιά μάθηση είναι μια κατηγορία αλγορίθμων μηχανικής μάθησης που χρησιμοποιεί πολλά ενδιάμεσα στρώματα για να εξάγει σταδιακά υψηλού επιπέδου χαρακτηριστικά από την ακατέργαστη είσοδο. Για παράδειγμα, στην επεξεργασία εικόνας, τα χαμηλότερα επίπεδα μπορεί να εντοπίζουν τις ακμές, ενώ τα υψηλότερα επίπεδα μπορεί να εντοπίζουν τις έννοιες που είναι πιο σημαντικές για έναν άνθρωπο, όπως τα ψηφία, τα γράμματα ή τα πρόσωπα. Οι μέθοδοι που χρησιμοποιούνται μπορούν να είναι με ή χωρίς επίβλεψη.

Από άλλη οπτική γωνία, η βαθιά μάθηση αναφέρεται στην "προσομοίωση" από υπολογιστές ή "αυτοματοποίηση" των ανθρώπινων διαδικασιών μάθησης από μια πηγή (π.χ., μια εικόνα σκύλων) σε ένα μαθημένο αντικείμενο (σκύλοι). Συνεπώς αναφέρεται στην πλήρως αυτοματοποιημένη μάθηση από μια πηγή σε ένα τελικό μαθημένο αντικείμενο. Η βαθύτερη μάθηση αναφέρεται έτσι σε ένα συνδυαστικό σύστημα μάθησης: μια διαδικασία μάθησης από τον άνθρωπο από μια πηγή σε ένα μαθημένο ημι-αντικείμενο, ακολουθούμενη από μια διαδικασία μάθησης του υπο-αντικειμένου που μάθαινε ο άνθρωπος για το τελικό μαθημένο αντικείμενο.

Τα περισσότερα σύγχρονα μοντέλα βαθιάς μάθησης βασίζονται σε πολυεπίεδα τεχνητά νευρωνικά δίκτυα, όπως βαθιά νευρωνικά δίκτυα, δίκτυα πεποίθησης, επαναλαμβανόμενα νευρωνικά δίκτυα, συνελκτικά νευρωνικά δίκτυα και transformers, αν και μπορούν επίσης να περιλαμβάνουν προτασιακούς τύπους ή λανθάνουσες μεταβλητές οργανωμένες κατά στρώματα σε μοντέλα βαθιάς παραγωγής.

Στη βαθιά μάθηση, κάθε επίπεδο συνήθως μαθαίνει να μετατρέπει τα δεδομένα εισόδου του σε μια ελαφρώς πιο αφηρημένη και σύνθετη αναπαράσταση. Σε μια εφαρμογή αναγνώρισης εικόνας, η ακατέργαστη είσοδος μπορεί να είναι μια μήτρα pixels. Το πρώτο αναπαραστατικό επίπεδο μπορεί να αφαιρέσει τα pixels και να κωδικοποιήσει τις άκρες. Το δεύτερο στρώμα μπορεί να συνθέτει και να κωδικοποιεί διατάξεις ακμών. Το τρίτο στρώμα μπορεί να αναγνωρίζει τη μύτη και τα μάτια. και το τέταρτο στρώμα μπορεί να αναγνωρίσει ότι η εικόνα περιέχει ένα πρόσωπο. Είναι σημαντικό ότι μια διαδικασία βαθιάς μάθησης μπορεί να μάθει από μόνη της ποια χαρακτηριστικά θα τοποθετήσει βέλτιστα σε ποιο επίπεδο.

Τα συστήματα βαθιάς μάθησης έχουν ένα σημαντικό credit assignment path (CAP), όπου ως CAP ορίζεται η αλυσίδα των μετασχηματισμών από την είσοδο στην έξοδο. Τα CAP περιγράφουν πιθανές αιτιώδεις συνδέσεις μεταξύ εισόδου και εξόδου. Για ένα νευρωνικό δίκτυο, το βάθος των CAP είναι αυτό του δικτύου και είναι ο αριθμός των κρυφών επιπέδων συν ένα, καθώς το επίπεδο εξόδου είναι επίσης παραμετροποιημένο. Για επαναλαμβανόμενα νευρωνικά δίκτυα, στα οποία ένα σήμα μπορεί να διαδοθεί μέσω ενός στρώματος περισσότερες από μία φορές, το βάθος CAP είναι δυνητικά απεριόριστο. Δεν υπάρχει καθολικά αποδεκτό όριο βάθους που χωρίζει τη ρηχή μάθηση από τη βαθιά μάθηση, αλλά οι περισσότεροι ερευνητές συμφωνούν ότι η βαθιά μάθηση περιλαμβάνει βάθος CAP μεγαλύτερο από 2. Το CAP του βάθους 2 έχει αποδειχθεί ότι είναι μια καθολική προσέγγιση με την έννοια ότι μπορεί να μιμηθεί οποιαδήποτε συνάρτηση. Ωστόσο περισσότερα επίπεδα δεν προσθέτουν απαραίτητα στην ικανότητα προσέγγισης συναρτήσεων του δικτύου.

Οι αρχιτεκτονικές βαθιάς μάθησης μπορούν να κατασκευαστούν προσθέτοντας διαδοχικά επίπεδα με άπληστο τρόπο. Η βαθιά μάθηση βοηθά στο να ξεμπλέξει αυτές τις αφαιρέσεις και στην επιλογή των χαρακτηριστικών που βελτιώνουν την απόδοση.

Οι αλγόριθμοι βαθιάς μάθησης μπορούν να εφαρμοστούν σε εργασίες μάθησης χωρίς επίβλεψη. Αυτό είναι ένα σημαντικό πλεονέκτημα επειδή τα δεδομένα χωρίς ετικέτες είναι πιο άφθονα από τα δεδομένα με ετικέτες. Παραδείγματα δομών σε βάθος που μπορούν να εκπαιδευτούν με τρόπο χωρίς επίβλεψη είναι τα δίκτυα βαθιάς πεποίθησης, (deep belief networks).

Για εποπτευόμενες εργασίες εκμάθησης, οι μέθοδοι βαθιάς μάθησης εξαλείφουν τη μηχανική χαρακτηριστικών, μεταφράζοντας τα δεδομένα σε συμπαγείς ενδιάμεσες αναπαραστάσεις που μοιάζουν με τα κύρια συστατικά και εξάγουν δομές σε επίπεδα που αφαιρούν τον πλεονασμό της πληροφορίας.

Αρχιτεκτονικές βαθιάς μάθησης έχουν εφαρμοστεί σε τομείς όπως η όραση υπολογιστών, η αναγνώριση φωνής, η επεξεργασία φυσικής γλώσσας, η μετάφραση μηχανής, η βιοπληροφορική, ο σχεδιασμός φαρμάκων, η ανάλυση ιατρικών εικόνων, η επιστήμη του κλίματος, ο έλεγχος υλικού και προγράμματα επιτραπέζιων παιχνιδιών, όπου έχουν παράγει αποτελέσματα που συγκρίνονται με την απόδοση ανθρώπινων ειδικών και σε ορισμένες περιπτώσεις υπερέχουν. Επίσης οι επενδυτές χρησιμοποιούν όλο και περισσότερο τεχνικές βαθιάς μάθησης για να προβλέψουν και να αναλύσουν τις τάσεις στις αγορές μετοχών και συναλλάγματος.

Τα τεχνητά νευρωνικά δίκτυα εμπνεύστηκαν από την επεξεργασία πληροφοριών και τους κόμβους κατανεμημένης επικοινωνίας στα βιολογικά συστήματα. Τα ΤΝΔ έχουν πολλαπλές διαφορές από τους βιολογικούς εγκεφάλους. Συγκεκριμένα, τα τεχνητά νευρωνικά δίκτυα τείνουν να είναι στατικά και συμβολικά, ενώ ο βιολογικός εγκέφαλος των περισσότερων ζώων είναι δυναμικός και αναλογικός. Τα ΤΝΔ θεωρούνται γενικά ως χαμηλής ποιότητας μοντέλα σε σχέση με τη λειτουργία του εγκεφάλου.

Νευρωνικά Δίκτυα

Τα Τεχνητά Νευρωνικά Δίκτυα, (κοινώς γνωστά ως νευρωνικά δίκτυα), αποτελούν έναν ιδιαίτερο τύπο μοντέλου μηχανικής μάθησης που δομείται με βάση τις αρχές της νευρωνικής οργάνωσης όπως αυτές συναντώνται στα βιολογικά νευρωνικά δίκτυα που απαρτίζουν τα εγκεφαλικά συστήματα των εγκεφάλων. Δηλαδή πρόκειται για ένα αφηρημένο αλγοριθμικό κατασκεύασμα το οποίο στοχεύει τελικά στην απομίμηση του κεντρικού νευρικού συστήματος και των διεργασιών που αυτό επιτελεί.

Ένα νευρωνικό δίκτυο αποτελείται από απλούς υπολογιστικούς κόμβους (νευρώνες ή νευρώνια), διασυνδεδεμένους μεταξύ τους οι προσομοιώνουν προσεγγιστικά τα νευρώνια σε έναν βιολογικό εγκέφαλο. Κάθε σύνδεση, όπως οι συνάψεις σε ένα βιολογικό εγκέφαλο, μπορεί να μεταδώσει ένα σήμα σε άλλα νευρώνια. Έτσι κάθε τέτοιος κόμβος δέχεται ένα σύνολο αριθμητικών εισόδων από διαφορετικές πηγές (είτε από άλλους νευρώνες, είτε από τον χρήστη), επιτελεί έναν υπολογισμό με βάση αυτές τις εισόδους και παράγει μία έξοδο. Το σήμα που προωθείται λοιπόν μέσω μιας τέτοιας σύνδεσης είναι ένα πραγματικός αριθμός, και το αποτέλεσμα – η έξοδος – κάθε νευρώνα υπολογίζεται από κάποια μη γραμμική συνάρτηση του άθροισματος των εισόδων του. Για να βρούμε την έξοδο του νευρώνα, παίρνουμε το πολλαπλασιασμένο με το εκάστοτε βάρος άθροισμα όλων των εισόδων, επισημασμένο με τα βάρη των συνδέσεων από τις εισόδους προς το νευρώνα. Στη συνέχεια προσθέτουμε έναν όρο πόλωσης σε αυτό το άθροισμα. Αυτό το συνολικό λοιπόν άθροισμα ονομάζεται ενεργοποίηση. Έπειτα, αυτό το άθροισμα περνά μέσα από μια - συνήθως μη γραμμική - συνάρτηση ενεργοποίησης που παίρνει τιμές 0 ή 1 για να παράγει την έξοδο.

Τελικά, η εν λόγω έξοδος είτε κατευθύνεται στο περιβάλλον, είτε τροφοδοτείται ως είσοδος σε άλλους νευρώνες του δικτύου.

Υπάρχουν τρεις τύποι νευρώνων: οι νευρώνες εισόδου, οι νευρώνες εξόδου και οι υπολογιστικοί νευρώνες ή κρυμμένοι νευρώνες. Οι νευρώνες εισόδου δεν επιτελούν κανέναν υπολογισμό, μεσολαβούν απλώς ανάμεσα στις περιβαλλοντικές εισόδους του δικτύου και στους υπολογιστικούς νευρώνες. Οι νευρώνες εξόδου διοχετεύουν στο περιβάλλον τις τελικές αριθμητικές εξόδους του δικτύου.

Οι συνδέσεις μεταξύ των νευρώνων ονομάζονται ακμές. Συνήθως, οι νευρώνες και οι ακμές έχουν ένα βάρος που προσαρμόζεται καθώς η μάθηση προχωρά. Το βάρος αυξάνεται ή μειώνεται για να ενισχύσει ή να αποδυναμώσει το σήμα σε μια σύνδεση. Επίσης, στα νευρώνια μπορεί να έχει οριστεί ένα όριο (όρος πόλωσης), έτσι ώστε ένα σήμα να σταλεί μόνο αν το συγκεντρωμένο σήμα ξεπεράσει σε τιμή αυτό το κατώφλι.

Οι υπολογιστικοί νευρώνες πολλαπλασιάζουν κάθε είσοδό τους με το αντίστοιχο συναπτικό βάρος και υπολογίζουν το ολικό άθροισμα των γινομένων. Το άθροισμα αυτό τροφοδοτείται ως όρισμα στη συνάρτηση ενεργοποίησης, την οποία υλοποιεί εσωτερικά κάθε κόμβος. Η τιμή που λαμβάνει η συνάρτηση για το εν λόγω όρισμα είναι και η έξοδος του νευρώνα για τις τρέχουσες εισόδους και τα βάρη.

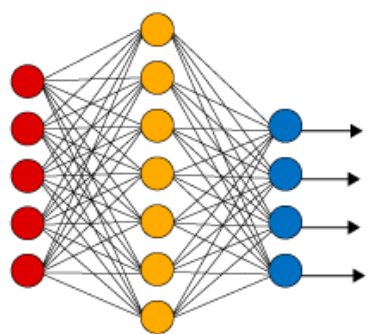
Συνήθως, τα νευρώνια συγκεντρώνονται σε πολλαπλά στρώματα, ιδιαίτερα στη βαθειά μάθηση. Τα νευρώνια ενός στρώματος συνδέονται μόνο με νευρώνια του αμέσως προηγούμενου και του αμέσως επόμενου στρώματος. Διάφορα στρώματα μπορεί να εκτελούν διάφορες μετασχηματίσεις στις εισόδους τους.

Το στρώμα που λαμβάνει εξωτερικά δεδομένα είναι το επίπεδο εισόδου. Το στρώμα που παράγει το τελικό αποτέλεσμα είναι το επίπεδο εξόδου. Ανάμεσά τους υπάρχουν μηδέν ή περισσότερα κρυφά στρώματα. Τα σήματα ταξιδεύουν από το πρώτο στρώμα (το επίπεδο εισόδου) στο τελευταίο στρώμα (το επίπεδο εξόδου), πιθανώς μετά από πολλαπλές διασχίσεις των στρωμάτων. Όπως προαναφέρθηκε, ένα δίκτυο ονομάζεται συνήθως βαθύ νευρωνικό δίκτυο αν έχει τουλάχιστον 2 κρυφά στρώματα. Χρησιμοποιούνται επίσης δίκτυα με μόνο ένα ή δίχως στρώματα, αλλά είναι πιο σπάνια.

Μεταξύ δύο στρωμάτων, είναι δυνατά πολλά μοτίβα σύνδεσης. Τα πιο δημοφιλή είναι τα "πλήρως διασυνδεδεμένα", (fully connected), με κάθε νευρώνα ενός στρώματος να συνδέεται με κάθε νευρώνα του επόμενου στρώματος. Νευρώνες με μόνο τέτοιου είδους συνδέσεις αποτελούν έναν κατευθυνόμενο ακυκλικό γράφο και είναι γνωστά σαν δίκτυα μετάδοσης. Ο όρος αυτός ουσιαστικά σημαίνει πως οι πληροφορίες σε τέτοια μοντέλα ρέουν μόνο σε μία κατεύθυνση - από τους κόμβους εισόδου, μέσω των κρυμμένων κόμβων, (αν υπάρχουν), και στους κόμβους εξόδου, χωρίς κύκλους ή βρόγχους.

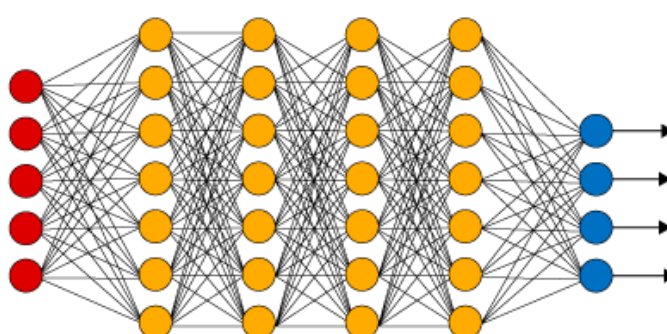
Σε αντίθεση με αυτά βρίσκονται τα δίκτυα που επιτρέπουν συνδέσεις μεταξύ νευρώνων στα ίδια ή προηγούμενα στρώματα με αποτέλεσμα η έξοδος κάποιων νευρώνων να επηρεάζει τις επακόλουθες εισόδους αυτών, τα οποία ονομάζονται αναδρομικά δίκτυα, (recurrent neural networks).

Simple Neural Network



● Input Layer ● Hidden Layer ● Output Layer

Deep Learning Neural Network



Στην παραπάνω εικόνα, [68] φαίνεται η βασική αρχιτεκτονική ενός απλού νευρωνικού δικτύου αριστερά κι ενός βαθιού νευρωνικού δικτύου δεξιά. Φαίνεται, όπως προαναφέρθηκε, πως βαθύ νευρωνικό δίκτυο είναι ένα με πάνω από ένα κρυφά στρώματα. Με κόκκινο αναπαρίστανται οι νευρώνες του στρώματος εισόδου, με κίτρινο οι νευρώνες των κρυφών στρωμάτων και με μπλε οι νευρώνες του στρώματος εξόδου.

Ιδιαίτερα κατά την δεκαετία των 2010s κι έπειτα έχει υπάρξει μία έκρηξη ενδιαφέροντος για τα νευρωνικά δίκτυα καθώς εφαρμόζονται με μεγάλη επιτυχία σε ένα ασυνήθιστα μεγάλο φάσμα τομέων της επιστήμης και της τεχνολογίας. Γενικά μιλώντας, μπορούν να εισαχθούν για την επίλυση οποιουδήποτε προβλήματος πρόβλεψης, ταξινόμησης ή ελέγχου, ενώ η τεράστια επιτυχία τους, μπορεί να αποδοθεί σε δύο βασικά στοιχεία τους: την αποτελεσματικότητά τους και την ευχρηστία τους.

Ως προς την αποτελεσματικότητά, τα νευρωνικά δίκτυα είναι πολύ εξελιγμένες τεχνικές μη γραμμικής μοντελοποίησης, ικανές να προσομοιώσουν εξαιρετικά πολύπλοκες λειτουργίες. Η γραμμική μοντελοποίηση υπήρξε ευρέως διαδεδομένη για πολύ καιρό, δεδομένου ότι στα γραμμικά μοντέλα εφαρμόζονται πολύ γνωστές στρατηγικές βελτιστοποίησης. Στις συνήθειες, όμως, περιπτώσεις όπου η γραμμική προσέγγιση δεν ήταν έγκυρη ή εφικτή, τα παραδοσιακά αυτά μοντέλα αποτύγχαναν αναλόγως. Τα νευρωνικά δίκτυα βέβαια επιτρέπουν τη μη γραμμικότητα μέσω της χρήσης μη γραμμικών συναρτήσεων ενεργοποίησης, όμως με τη σειρά τους δημιουργούν πρόβλημα στο ζήτημα της διάστασης (του πλήθους των διαφορετικών εισόδων και εξόδων), το οποίο δυσχεραίνει σημαντικά τις προσπάθειες μοντελοποίησης μη γραμμικών συναρτήσεων με μεγάλο αριθμό μεταβλητών.

Ως προς την ευχρηστία, τα νευρωνικά δίκτυα εκπαιδεύονται με παραδείγματα. Ο χρήστης συγκεντρώνει κι επεξεργάζεται τα κατάλληλα δεδομένα και στη συνέχεια, καθώς τα τροφοδοτεί συστηματικά στο δίκτυο μέσω των επιλεγμένων αλγορίθμων εκπαίδευσης, το δίκτυο αντιλαμβάνεται αυτόματα τη δομή των δεδομένων και εκφράζει αυτή τη γνώση ως τις κατάλληλες επιλογές συναπτικών βαρών. Επομένως το τελικό αποτέλεσμα της εκπαίδευσης με ένα συγκεκριμένο σύνολο παραδειγμάτων είναι η εύρεση βαρών εκείνων του δικτύου που δίνουν ποσοσισταία το ελάχιστο σφάλμα. Το επίπεδο των γνώσεων του χρήστη που απαιτούνται για μια επιτυχημένη

εφαρμογή των νευρωνικών δικτύων, μπορεί είναι πολύ χαμηλότερο συγκριτικά με κάποια περίπτωση που θα χρησιμοποιούνταν ορισμένες πιο παραδοσιακές, μη γραμμικές στατιστικές μέθοδοι, μέχρι κάποιο επίπεδο. Αυτό δεν σημαίνει πως ο χρήστης δεν χρειάζεται να έχει ουσιώδεις κι εξειδικευμένες γνώσεις σχετικά με τον τρόπο προετοιμασίας των δεδομένων, τον τρόπο εκλογής του κατάλληλου νευρωνικού δικτύου (hyperparameter tuning), καθώς και στον τρόπο ερμηνείας των αποτελεσμάτων.

Υπάρχουν ποικίλα είδη νευρωνικών δικτύων που χρησιμοποιούνται σήμερα, εκ των οποίων μερικά από τα πιο δημοφιλή είναι τα:

- multilayer perceptron (MLP – δίκτυα πολυεπίπεδης τροφοδοσίας: Είναι ο κλασικός τύπος νευρωνικών δικτύων που χρησιμοποιήθηκε και σε αυτή την εργασία, αποτελούμενο από ένα στρώμα εισόδου, ένα ή περισσότερα κρυφά στρώματα, των οποίων κάθε νευρώνας σε αυτά τα επίπεδα συνδέεται με όλους τους νευρώνες των προηγούμενων και επόμενων επιπέδων, και ένα στρώμα εξόδου που παρέχει την έξοδο του μοντέλου. Εκπαιδεύεται συνήθως με τη μέθοδο της καθολικής προσαρμογής (backpropagation), στην οποία τα δεδομένα εισόδου προωθούνται προς τα εμπρός μέσα στο δίκτυο, και τα αποτελέσματα συγκρίνονται με τα πραγματικά αποτελέσματα. Έπειτα, αρχίζει η αντίθετη κίνηση (backpropagation) από το τέλος του δικτύου προχωρώντας προς τα πίσω. Κατά την επιστροφή προς τα πίσω, υπολογίζεται η κλίση (η παράγωγος) της συνάρτησης κόστους ως προς τα βάρη του δικτύου και στη συνέχεια αυτά ενημερώνονται χρησιμοποιώντας μια μέθοδο βελτιστοποίησης, όπως ο επαναληπτικός αλγόριθμος της καθολικής προσαρμογής. Το είδος αυτό χρησιμοποιείται σχεδόν σε όλους τους τομείς που προτιμώνται νευρωνικά δίκτυα και ιδίως για την μοντελοποίηση πολύπλοκων συναρτήσεων, αν και ολοένα και αντικαθίσταται από νέα πιο ανεπτυγμένα μοντέλα.
- Συνελικτικά Νευρωνικά Δίκτυα (CNN – convolutional neural networks): είναι εξειδικευμένα νευρωνικά δίκτυα σχεδιασμένα για την αναγνώριση προτύπων σε εικόνες και γενικότερα δεδομένα που ακολουθούν μια δομική ιεραρχία – εξού και χρησιμοποιούνται κυρίως για την αναγνώριση κι επεξεργασία εικόνων και βίντεο. Τα cnn αποτελούνται από δύο είδη επιπέδων, τα συνελικτικά και τα πλήρως διασυνδεδεμένα. Κατά την διάρκεια της εκπαίδευσης αρχικά εφαρμόζονται φίλτρα τα οποία ανιχνεύουν χαρακτηριστικά σε διάφορες υποπεριοχές των δεδομένων, (τα οποία συνήθως είναι πάνω από μίας διάστασης), όπως ακμές, χρώματα και σχήματα. Στη συνέχεια στρώματα συνέλιξης (convolutional layers) και στρώματα υποδειγματοληψίας (pooling layers) χρησιμοποιούνται για τη μείωση των διαστάσεων και τη δημιουργία μιας ιεραρχίας στην αναγνώριση των χαρακτηριστικών, γεγονός το οποίο βελτιώνει σημαντικά την αποδοτικότητα του δικτύου. Τέλος, χρησιμοποιούνται τα πλήρη διασυνδεδεμένα επίπεδα (fully connected layers) τα οποία ταξινομούν τα χαρακτηριστικά που βρέθηκαν πρωτίστως. Η χρήση κοινόχρηστων βαρών σε αυτά τα

επίπεδα επιτρέπει τη γενίκευση των εκπαιδευμένων χαρακτηριστικών σε διάφορες θέσεις της εικόνας.

- Αναδρομικά Νευρωνικά Δίκτυα (RNN): αποτελούν μια κατηγορία εξειδικευμένων νευρωνικών δικτύων που είναι σχεδιασμένα για την επεξεργασία ακολουθιών δεδομένων, καθώς έχουν την ικανότητα να διατηρούν στη μνήμη τους την προηγούμενη είσοδο, επιτρέποντας την αντιμετώπιση δεδομένων σε χρονική σειρά. Αυτή η ικανότητα οφείλεται στο βασικό χαρακτηριστικό τους, τους αναδρομικούς κόμβους που επιτρέπουν την επαναχρησιμοποίηση των εσωτερικών καταστάσεών τους καθώς επεξεργάζονται νέες εισόδους. Έτσι, είναι ιδανικά για προβλήματα όπου η σειρά των δεδομένων έχει σημασία, όπως στην επεξεργασία της γλώσσας, τη μουσική ή τις χρονοσειρές.
- Δίκτυα Μακράς Βραχυπρόθεσμης Μνήμης (LSTMs - Long Short-term memory Networks): σχεδιάστηκαν για να αντιμετωπίσουν το πρόβλημα της δυσκολίας στην επίλυση προβλημάτων με συσχετίσεις με δεδομέν μακρυνά στο χρόνο που εμφανίζουν τα RNNs καθώς καθώς να ξεχνούν πληροφορίες από τις αρχικές αναδρομές. Ουσιαστικά βασίζονται σε ειδικούς μηχανισμούς πύλης οι οποίοι ελέγχουν πότε και πώς η πληροφορία κυκλοφορεί εντός του δικτύου. Οι τρεις κύριες πύλες που χρησιμοποιούνται είναι η πύλη εισόδου (input gate), η πύλη λήψης αποφάσεων (forget gate) και η πύλη εξόδου (output gate), οι οποίες ελέγχουν αντίστοιχα πότε και πώς η πληροφορία εισέρχεται, εξέρχεται και διατηρείται στη μνήμη του κάθε κόμβου του δικτύου. Έτσι, τα LSTMs μπορούν να διατηρούν μακροχρόνιες εξαρτήσεις και να αντιμετωπίζουν πιο αποτελεσματικά τέτοια προβλήματα όπως η αναγνώριση συντακτικής δομής σε φράσεις, η μετάφραση γλωσσών και η αναγνώριση φωνής.
- Transformers: είναι ίσως το ισχυρότερο μοντέλο αρχιτεκτονικής που εισήχθηκε αρχικά για προβλήματα μετάφρασης κειμένου αλλά έκτοτε χρησιμοποιούνται και σε άλλους τομείς της επεξεργασίας φυσικής γλώσσας, στην επεξεργασία εικόνων κλπ. Το κύριο χαρακτηριστικό τους είναι η ικανότητα να επεξεργάζονται ακολουθίες δεδομένων χωρίς την ανάγκη της πρωτύτερης γνώση της δομής τους η οποία οφείλεται στους μηχανισμούς προσοχής (attention mechanisms) που διαθέτουν, οι οποίοι επιτρέπουν στο μοντέλο να δίνει διαφορετική βαρύτητα σε διάφορα τμήματα της εισόδου κατά την επεξεργασία. Αυτό τους καθιστά ιδιαίτερα αποτελεσματικούς στην αντιμετώπιση ακολουθιών με μεταβλητό μήκος, όπως είναι οι προτάσεις κειμένου και η αναγνώριση μουσικών μελωδιών.

Συναρτήσεις Ενεργοποίησης

Σε κάθε στρώμα ενός νευρωνικού δικτύου χρειάζεται να έχει οριστεί μια συνάρτηση ενεργοποίησης, η οποία είναι μια συνάρτηση που υπολογίζει

την έξοδο του κόμβου, βασιζόμενη στις εισόδους του και τα βάρη. Τα περισσότερα προβλήματα μπορούν να λυθούν χρησιμοποιώντας μη γραμμικές συναρτήσεις ενεργοποίησης. Μερικές από τις σύγχρονες συναρτήσεις ενεργοποίησης που χρησιμοποιούνται ευρέως είναι:

- η ReLu (ανορθωτής), η οποία χρησιμοποιεί μία συνάρτηση ράμπας και είναι ανάλογη με την ανόρθωση μισού κύματος σε πιο ηλεκτρολογικούς όρους. Έχει διαπιστωθεί ότι επιτρέπει καλύτερη εκπαίδευση βαθύτερων δικτύων σε σχέση με τις υπόλοιπες κι έτσι μετά το 2011 είναι η πιο δημοφιλής συνάρτηση ενεργοποίησης για βαθιά νευρωνικά δίκτυα,
- η ομαλή έκδοση της ReLu, η οποία χρησιμοποιεί μια συνεχή συνάρτηση που προσεγγίζει τη μορφή της συνάρτησης ράμπας, αλλά είναι πιο ομαλή γύρω από το μηδέν, γεγονός που την καθιστά πιο χρήσιμη σε προβλήματα που απαιτούν ομαλή ενεργοποίηση και μπορεί να βελτιώσει τη σύγκλιση του μοντέλου κατά την εκπαίδευση καθώς έχει μία συνεχή και ομαλή παράγωγο παντού,
- την σιγμοειδή, (sigmoid), η οποία ορίζεται από τον μαθηματικό τύπο $\sigma(x) = 1/(1 + e^{-x})$. Οι τιμές της σιγμοειδούς βρίσκονται στο διάστημα από μηδέν έως ένα, κάτι που την καθιστά κατάλληλη για προβλήματα όπου η έξοδος πρέπει να ερμηνευτεί ως πιθανότητα ή βάρος. Επομένως συνήθως χρησιμοποιείται σε προβλήματα όπου η έξοδος αντιπροσωπεύει πιθανότητες και σε προβλήματα ταξινόμησης δυαδικών κατηγοριών (binary classification).
- Την GELU, (Gaussian Error Linear Unit), η οποία έχει τον τύπο:

$$GELU(x) = 0.5x \left(1 + \tanh \left(\sqrt{\frac{2}{\pi}} (x + 0.044715x^3) \right) \right)$$

και η οποία έχει εισαχθεί καθαρά για χρήση σε τεχνητά νευρωνικά δίκτυα. Οφείλει το όνομά της στην κανονική κατανομή που χρησιμοποιείται και προσπαθεί να συνδυάσει τις καλές ιδιότητες της σιγμοειδούς συνάρτησης (για τη μετατροπή σε περιορισμένο εύρος) με την απλότητα της συνάρτησης ενεργοποίησης ReLU. Όμως, είναι περισσότερο απαιτητική σε υπολογιστικούς πόρους συγκριτικά με κάποιες άλλες συναρτήσεις ενεργοποίησης, καθώς περιλαμβάνει την χρήση υπερβολικής εφαπτομένης.

- Την Softmax, η οποία ορίζεται ως:

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^N e^{x_j}}$$

και η οποία χρησιμοποιείται συχνά στο τελευταίο επίπεδο ενός νευρωνικού δικτύου, όταν προσπαθεί να λύσει ένα πρόβλημα ταξινόμησης.

Το κύριο χαρακτηριστικό των νευρωνικών δικτύων είναι η εγγενής ικανότητα μάθησης. Ως μάθηση μπορεί να οριστεί η σταδιακή βελτίωση της ικανότητας του δικτύου να επιλύει κάποιο πρόβλημα (π.χ. η σταδιακή προσέγγιση μίας συνάρτησης). Η μάθηση επιτυγχάνεται μέσω της εκπαίδευσης, μίας επαναληπτικής διαδικασίας σταδιακής προσαρμογής των

παραμέτρων του δικτύου (συνήθως των βαρών και της τιμής σταθερής πόλωσής του) σε τιμές κατάλληλες ώστε να επιλύεται με ικανοποιητική ακρίβεια το πρόβλημα. Αφού ένα δίκτυο εκπαιδευτεί, οι παράμετροί του συνήθως «παγώνουν» στις κατάλληλες τιμές και από εκεί κι έπειτα είναι σε λειτουργική κατάσταση. Το ζητούμενο είναι το λειτουργικό δίκτυο να έχει μία ικανότητα γενίκευσης, δηλαδή να δίνει σωστές εξόδους για καινούργιες εισόδους καινοφανείς – που δεν ανήκουν δηλαδή στα δεδομένα με τα οποία εκπαιδεύτηκε.

Για τον σκοπό της αποδοτικής αξιολόγησης έχει αναπτυχθεί η τεχνική «train test split» η οποία χωρίζει τα δεδομένα σας σε τρία σύνολα: ένα για την εκπαίδευση του μοντέλου (train set) κι ένα για τον έλεγχο της απόδοσής του (test set). Μερικές φορές ορίζεται κι ένα τρίτο σύνολο για τον σκοπό της επαλήθευσης των αποτελεσμάτων (validation set).

Με βάση ένα προκαθορισμένο ποσοστό, συνήθως 80% για εκπαίδευση, 20% για έλεγχο ή 80% για εκπαίδευση, 10% για έλεγχο και 10% για επαλήθευση, τα δεδομένα επιλέγονται τυχαία και διαχωρίζονται. Έτσι, το μοντέλο εκπαιδεύεται χρησιμοποιώντας μόνο τα δεδομένα εκπαίδευσης, ενώ αξιολογείται σε ξεχωριστά δεδομένα ελέγχου που δεν χρησιμοποιήθηκαν κατά την εκπαίδευση. Μέσω αυτής της διαδικασίας μπορεί να αξιολογηθεί η ικανότητα γενίκευσης του μοντέλου.

Hyperparameter Optimization

Υπερπαράμετρος ορίζεται ως μία παράμετρος που δεν διαμορφώνεται απευθείας από τα δεδομένα κατά τη διάρκεια της εκπαίδευσης ενός μοντέλου μηχανικής μάθησης. Οι υπερπαράμετροι καθορίζουν τη δομή του μοντέλου και την απόδοση της εκπαίδευσης. Μετά την δημιουργία ενός δικτύου, οι υπερπαράμετροι πρέπει να προσαρμόζονται και να επιλέγονται πριν από την εκπαίδευση του μοντέλου, μέσω πειραμάτων, με στόχο την βελτιστοποίηση του μοντέλου, μια διαδικασία που είναι γνωστή ως hyperparameter tuning.

Υπερπαράμετροι που πρέπει να οριστούν πριν ένα δίκτυο προχωρήσει στην εκπαίδευση αποτελούν για παράδειγμα:

- Η αρχιτεκτονική του δικτύου, δηλαδή τον αριθμό των επιπέδων, των νευρώνων, τη συνάρτηση ενεργοποίησης κ.λπ.
- Ο ρυθμός μάθησης, δηλαδή το μέγεθος των βημάτων που κάνει ο αλγόριθμος βελτιστοποίησης κατά την εκπαίδευση. Το πόσο γρήγορα αλλάζουν τα βάρη σε πιο πρακτικούς όρους. Ένας υψηλός ρυθμός μάθησης μπορεί να οδηγήσει σε ταχεία σύγκλιση, αλλά υπάρχει ο κίνδυνος λόγω υπερβολικών βημάτων να παρακάμψει το μοντέλο την ελάχιστη τιμή της συνάρτησης κόστους. Αντίθετα, ένας χαμηλός ρυθμός μάθησης μπορεί να αυξήσει τον χρόνο εκπαίδευσης και να κολλήσει σε τοπικά ελάχιστα.
- Το βάρος πολυπλοκότητας για τον έλεγχο της υπερεκπαίδευσης. Η υπερεκπαίδευση (overfitting) είναι ένα φαινόμενο που το δίκτυο μαθαίνει 'υπερβολικά' καλά τα ειδικά παραδείγματα

που του έχουν τροφοδοτηθεί και αδυνατεί να γενικεύει σωστά όταν αντιμετωπίζει το test set.

- Οι εποχές εκπαίδευσης, δηλαδή πόσες φορές το σύνολο των δεδομένων θα περάσει από το μοντέλο κατά την εκπαίδευση. Κατά τη διάρκεια μιας εποχής, όλα τα δεδομένα εισόδου περνούν από το μοντέλο, υπολογίζονται οι απώλειες (losses), ενημερώνονται οι παράμετροι του μοντέλου και επαναλαμβάνεται η διαδικασία και ούτω καθεξής.
- Το μέγεθος δείγματος (Batch Size), δηλαδή τον αριθμό των δειγμάτων που χρησιμοποιούνται σε κάθε εποχή εκπαίδευσης.
- Την συνάρτηση evaluation που χρησιμοποιείται, η οποία υποδεικνύει την απόσταση μεταξύ των προβλέψεων του μοντέλου και των πραγματικών τιμών του στόχου. Σε προβλήματα κατηγοριοποίησης, μπορεί να χρησιμοποιηθούν συναρτήσεις κόστους όπως η cross-entropy για πολυκατηγορικές τάξεις ή η δυαδική εντροπία, μία ειδική περίπτωση της, για δυαδικές τάξεις, οι οποίες έχουν ως στόχο να μειωθεί η εντροπία, προκειμένου το μοντέλο να παράγει προβλέψεις που πλησιάζουν τις πραγματικές τιμές. Σε προβλήματα παλινδρόμησης, μπορεί να χρησιμοποιηθούν συναρτήσεις evaluation όπως η μέση τετραγωνική απόκλιση (mean squared error), η οποία μετράει το μέσο τετραγωνικό σφάλμα μεταξύ των προβλέψεων του μοντέλου και των πραγματικών τιμών.
- Ο κατάλληλος βελτιστοποιητής, (optimizer). Οι βελτιστοποιητές είναι αλγόριθμοι που χρησιμοποιούνται για την ελαχιστοποίηση της συνάρτησης κόστους (ή απώλειας) κατά την εκπαίδευση του μοντέλου. Η συνάρτηση κόστους μετράει το σφάλμα μεταξύ των προβλέψεων του μοντέλου και των πραγματικών τιμών. Οι βελτιστοποιητές λειτουργούν επιδιώκοντας να ενημερώσουν τις παραμέτρους του μοντέλου έτσι ώστε η συνάρτηση κόστους να μειωθεί. Κάθε ενημέρωση παραμέτρων γίνεται με βάση το κλάσμα του αρνητικού κλίσης (ή παραγώγου) της συνάρτησης κόστους ως προς τις παραμέτρους, πολλαπλασιασμένο με τον ρυθμό μάθησης. Ορισμένοι δημοφιλείς βελτιστοποιητές είναι οι:
 - *Stochastic Gradient Descent (SGD)*: Ο βασικός αλγόριθμος βελτιστοποίησης που ενημερώνει τις παραμέτρους με βάση την αρνητική κλίση της συνάρτησης κόστους και ο οποίος χρησιμοποιεί τυχαία επιλογή δειγμάτων δεδομένων σε κάθε επανάληψη, εξού και ονομάζεται και στοχαστικός
 - *Adam (Adaptive Moment Estimation)*: Συνδυασμός του Stochastic Gradient Descent με την έννοια της έκτασης (momentum), η οποία βοηθάει στην επιτάχυνση της σύγκλισης, ειδικά σε κατευθύνσεις με μεγάλη κλίση, και με την προσαρμογή του ρυθμού μάθησης για κάθε παράμετρο ξεχωριστά.
 - *RMSprop (Root Mean Square Propagation)*: Αντίστοιχος με τον Adam, χωρίς την έννοια της έκτασης.

- ο *Adagrad (Adaptive Gradient Algorithm)*: Προσαρμόζει τον ρυθμό μάθησης για κάθε παράμετρο με βάση τη συχνότητα εμφάνισής της.

Όπως προαναφέρθηκε, η βελτιστοποίηση των υπερπαραμέτρων επιτελείται κυρίως «χειρονακτικά», ακολουθώντας τη μέθοδο δοκιμής και λάθους. Πέρα όμως από τους εμπειρικούς αυτούς κανόνες έχουν αναπτυχθεί και μερικές άλλες τεχνικές βελτιστοποίησης των υπερπαραμέτρων οι οποίες βασίζονται στην αυτόματη δοκιμή πολλαπλών τιμών των παραμέτρων και στην τελική επιλογή των καλύτερων.

Μία από αυτές είναι το *Grid search*, το οποίο διαιρεί τις διάφορες τιμές υπερπαραμέτρων που θέλει ο χρήστης να δοκιμάσει σε ένα "πλέγμα" (*grid* – εξού και το όνομα). Αρχικά, επιλέγονται οι υπερπαραμέτροι που χρειάζονται βελτιστοποίηση και στη συνέχεια ορίζεται ένα σύνολο τιμών που μπορούν να λάβουν προς εξέταση. Κατόπιν ο αλγόριθμος εκπαιδεύει και αξιολογεί το μοντέλο σε κάθε συνδυασμό υπερπαραμέτρων που προκύπτει από το πλέγμα και τελικά επιλέγεται αυτός ο συνδυασμός που παρέχει το ελάχιστο σφάλμα. Η τεχνική αυτή είναι χρήσιμη για τον ταχύ κι αυτοματοποιημένο εντοπισμό καλών υπερπαραμέτρων, ωστόσο, επειδή με την αύξηση του αριθμού των υπερπαραμέτρων και των τιμών, η διαδικασία μπορεί να αποβεί πολύ χρονοβόρα είναι καλό να χρησιμοποιείται σε συνδυασμό με την εμπειρική εποπτεία του χρήστη έτσι ώστε να δοκιμαστούν οι ελάχιστες απαραίτητες τιμές.

Το *Random Search*, (τυχαία αναζήτηση), είναι μια άλλη τεχνική βελτιστοποίησης υπερπαραμέτρων που σε αντίθεση με το *Grid Search* επιλέγει τυχαία συνδυασμούς για αξιολόγηση, μέσα από έναν καθορισμένο εύρος τιμών. Η τυχειότητα μπορεί να οδηγήσει σε γρηγορότερη εξερεύνηση του χώρου των υπερπαραμέτρων, καθώς δεν απαιτεί την εξέταση όλων των πιθανών συνδυασμών, αλλά καθιστά πιθανό το σενάριο ο αλγόριθμος μην εξερευνήσει όλο τον χώρο των υπερπαραμέτρων κι έτσι να αποτύχει να βρει τις βέλτιστες ρυθμίσεις. Έτσι, η τεχνική *Random Search* χρησιμοποιείται κυρίως όταν ο χώρος των υπερπαραμέτρων είναι μεγάλος και δύσκολος να εξερευνηθεί πλήρως.

Τέλος, η τεχνική *Bayesian search*, βασίζεται σε μοντέλα Μπέυζ που προσπαθούν να μοντελοποιήσουν τη σχέση μεταξύ των υπερπαραμέτρων και της απόδοσης του μοντέλου. Αναλυτικότερα η μέθοδος αυτή χρησιμοποιεί προσπαθώντας να κρατήσει μια ισορροπία την εξερεύνηση του χώρου των υπερπαραμέτρων και την εκμετάλλευση των γνωστών καλών συνδυασμών. Αυτό επιτυγχάνεται με τη χρήση μιας συνάρτησης κατανομής, συνήθως του κριτηρίου *Ενστάσεων-Εκτίμησης (Expected Improvement)*. Καθώς νέα δεδομένα γίνονται διαθέσιμα, το μοντέλο ανανεώνεται, και η διαδικασία εξερεύνησης-εκμετάλλευσης επαναπροσαρμόζεται. Αυτό επιτρέπει στη μέθοδο να προσαρμόζεται στη σχετική σημασία των υπερπαραμέτρων. Όπως και η *Random Search*, η βελτιστοποίηση Μπέυζ είναι χρήσιμη όταν ο χώρος των υπερπαραμέτρων είναι σύνθετος και δαπανηρός στην εξέταση όλων των δυνατών συνδυασμών, ενώ είναι πιο αποτελεσματική για μοντέλα μηχανικής μάθησης που απαιτούν πολύ χρόνο εκπαίδευσης και αξιολόγησης.

Στο πλαίσιο αυτής της διπλωματικής χρησιμοποιήθηκε η τεχνική *Grid Search*.

Transfer Learning

Καθώς τα δίκτυα που εκπαιδεύονται offline, όπως αυτά που χρησιμοποιήθηκαν για αυτή την εργασία, χρειάζονται εξ αρχής όλα τα δεδομένα διαθέσιμα, προκύπτει το πρόβλημα της ενσωμάτωσης νέων δεδομένων όταν είναι αναγκαίο να προστεθούν σε ένα τέτοιο δίκτυο; Το πρόβλημα αυτό μπορεί να περιλαμβάνει τη διαφορά στο χαρακτήρα των νέων δεδομένων, τη διαφορετική τους κλίμακα ή ακόμη και την εμφάνιση εντελώς νέων παραμέτρων. Σε μεγάλα δίκτυα η επανεκπαίδευση όλου του δικτύου είναι μια διαδικασία αρκετά χρονοβόρα και οικονομικά δύσκολη που πρακτικά είναι αδύνατο να γίνεται συνέχεια.

Για την αντιμετώπιση της πρόκλησης αυτής υπάρχει η μέθοδος transfer learning (μεταφορά μάθησης), που χρησιμοποιείται όταν ένα μοντέλο που έχει εκπαιδευτεί σε ένα συγκεκριμένο αρχικό πρόβλημα χρησιμοποιείται με ζητούμενη βέλτιστη απόδοση σε ένα διαφορετικό, αλλά σχετικό πρόβλημα. Η επαναχρησιμοποίηση ή η μεταφορά πληροφοριών από προηγούμενες εκπαιδευμένες διεργασίες σε νέες διεργασίες μπορεί να βελτιώσει σημαντικά την αποτελεσματικότητα της μάθησης.

Άλλοι λόγοι για την χρήση transfer learning είναι η έλλειψη επαρκών δεδομένων, καθώς πολλές περιπτώσεις, η ανάπτυξη ενός μεγάλου συνόλου δεδομένων για την εκπαίδευση από το μηδέν είναι χρονοβόρα και δαπανηρή ενώ το transfer learning επιτρέπει τη χρήση υπάρχοντος μοντέλου που έχει εκπαιδευτεί σε ένα παρελθόν πρόβλημα, ανεξαρτήτως της πολυπλοκότητάς του. Επίσης χρησιμοποιείται για τη βελτιστοποίηση συνεχούς εκπαίδευσης, σε περιπτώσεις που υπάρχουν συνεχώς νέα δεδομένα, καθώς το transfer learning επιτρέπει στο μοντέλο να επαναχρησιμοποιήσει τις γνώσεις του χωρίς να ξανακάνει πλήρη εκπαίδευση.

Το transfer learning προσφέρει διάφορα πλεονεκτήματα που καθιστούν αυτήν την τεχνική ιδιαίτερα χρήσιμη σε πολλές εφαρμογές της μηχανικής μάθησης. Τα κυριότερα από αυτά είναι:

- Χρήση Υπάρχοντος Κώδικα: Το transfer learning επιτρέπει τη χρήση υπάρχοντος κώδικα (pre-trained models) από άλλες εφαρμογές ή πηγές, περιορίζοντας την ανάγκη για πολύπλοκη και δαπανηρή ανάπτυξη από το μηδέν.
- Απόκτηση Γνώσης: Το μοντέλο που έχει εκπαιδευτεί σε ένα πρόβλημα μπορεί να μεταφέρει τις γνώσεις που έχει αποκτήσει σε ένα νέο, σχετικό πρόβλημα, βοηθώντας στην αύξηση της απόδοσης.
- Εξοικονόμηση Υπολογιστικού Χρόνου: Η χρήση προ-εκπαιδευμένων μοντέλων συνήθως εξοικονομεί χρόνο υπολογιστικής ισχύος, καθώς δεν απαιτείται πλήρης εκπαίδευση από το μηδέν.
- Βελτιστοποίηση σε Συγκεκριμένα Προβλήματα: Το μοντέλο μπορεί να προσαρμοστεί (fine-tuning) σε συγκεκριμένα προβλήματα, προσαρμόζοντας ορισμένα μέρη του μόνο στις ιδιαιτερότητες του καινούργιου προβλήματος.

Μέχρι το 2020, [69] είχαν προταθεί τέσσερις προσεγγίσεις για την υλοποίηση της επαγωγικής μεταφοράς μάθησης.

Η πρώτη είναι η μεταφορά χαρακτηριστικών αναπαράστασης (feature-representation transfer) η οποία επικεντρώνεται στις αναπαραστάσεις χαρακτηριστικών που μαθαίνει το νευρωνικό δίκτυο από το αρχικό πρόβλημα, με σκοπό την αξιοποίηση αυτών των αναπαραστάσεων για ένα συναφές αλλά διαφορετικό πρόβλημα - στόχο. Οι αναπαραστάσεις χαρακτηριστικών αποτυπώνουν τα υποκείμενα μοτίβα και τις εσωτερικές δομές στα δεδομένα εισόδου και η κεντρική ιδέα είναι ότι αυτές οι αναπαραστάσεις που μαθαίνονται μπορούν να έχουν κοινή συνάφεια ή χρησιμότητα για εργασίες στο πρόβλημα - στόχο. Έτσι, η διαδικασία περιλαμβάνει την εξαγωγή και την προσαρμογή των αναπαραστάσεων χαρακτηριστικών που αποκτήθηκαν κατά την εκπαίδευση σε έναν τομέα προέλευσης για την ενίσχυση της μαθησιακής απόδοσης κατά την επεξεργασία δεδομένων από έναν διαφορετικό τομέα.

Η δεύτερη είναι η μεταφορά περιπτώσεων (instance transfer). Παρόμοια με τη μεταφορά αναπαράστασης χαρακτηριστικών, εδώ η εστίαση είναι στη μεταφορά συγκεκριμένων στιγμιτύπων ή παραδειγμάτων, τα οποία μπορούν να είναι κατατοπιστικά και συναφή σε διαφορετικούς τομείς, από ένα αρχικό πρόβλημα σε ένα πρόβλημα - στόχο. Η διαδικασία αυτή περιλαμβάνει τη μεταφορά παραδειγμάτων μαζί με τις σχετικές ετικέτες τους, επιτρέποντας στο μοντέλο να προσαρμόζεται και να γενικεύει καλύτερα σε παρόμοιες εργασίες για το επικείμενο πρόβλημα.

Η Τρίτη προσέγγιση μεταφοράς μάθησης είναι γνωστή ως μεταφορά σχεσιακής γνώσης και αποσκοπεί στην αξιοποίηση της γνώσης σχετικά με τις σχέσεις που μαθαίνονται σε από ένα πρόβλημα προέλευσης για να βελτιωθεί η απόδοση του στο πρόβλημα - στόχο. Η έμφαση στη μεταφορά σχεσιακής γνώσης δίνεται στην καταγραφή και μεταφορά των εξαρτήσεων, των συνδέσεων και των αλληλεπιδράσεων μεταξύ διαφορετικών στοιχείων στα δεδομένα. Έτσι, ξεπερνά την άμεση μεταφορά των παραστάσεων των περιπτώσεων ή των χαρακτηριστικών και επικεντρώνεται στο ευρύτερο πλαίσιο του τρόπου με τον οποίο τα στοιχεία αλληλεπιδρούν. Πρακτικά, η μεταφορά σχεσιακής γνώσης περιλαμβάνει την κωδικοποίηση και τη μεταφορά πληροφοριών σχετικά με τις σχέσεις κατά το αρχικό πρόβλημα, επιτρέποντας στο μοντέλο να κατανοήσει και να αξιοποιήσει καλύτερα παρόμοιες σχέσεις στο τελικό πρόβλημα.

Τέλος, η τέταρτη προσέγγιση είναι η μεταφορά παραμέτρων η οποία εστιάζει συγκεκριμένα στη μεταφορά των παραμέτρων ενός προ-εκπαιδευμένου μοντέλου. Η βασική ιδέα πίσω από τη μεταφορά παραμέτρων είναι ότι η γνώση που κωδικοποιείται στις παραμέτρους ενός καλά εκπαιδευμένου μοντέλου σε έναν τομέα μπορεί να είναι πολύτιμη όταν εφαρμόζεται σε έναν διαφορετικό αλλά συναφή τομέα, καθώς οι παράμετροι αντιπροσωπεύουν τα μαθημένα βάρη και τις σταθερές που αποτυπώνουν τα υποκείμενα μοτίβα και χαρακτηριστικά των δεδομένων.

Υπάρχουν δύο κύριες μέθοδοι που χρησιμοποιούνται συνήθως στη μεταφορά παραμέτρων: ο διαμοιρασμός παραμέτρων (parameter-sharing) και ο λεπτομερής συντονισμός (fine tuning). Στον διαμοιρασμό παραμέτρων, η ιδέα είναι η άμεση αντιγραφή των παραμέτρων από το προ-εκπαιδευμένο μοντέλο στο μοντέλο που επιλύει το νέο πρόβλημα. Αυτές οι παράμετροι παγώνουν στη συνέχεια, δηλαδή δεν ενημερώνονται κατά τη διάρκεια της μεταγενέστερης εκπαίδευσης. Αυτό προϋποθέτει ότι οι παράμετροι είναι σε

μεγάλο βαθμό μεταβιβάσιμες και μπορούν να παρέχουν μια ευνοϊκή αρχικοποίηση για το μοντέλο στον τομέα-στόχο.

Από την άλλη πλευρά, κατά τον λεπτομερή συντονισμό θεωρείται πως ενώ οι παράμετροι από τον τομέα προέλευσης είναι χρήσιμες, ενδέχεται να χρειαστεί να προσαρμοστούν ή να επαναρυθμιστούν χρησιμοποιώντας περιορισμένα δεδομένα από το νέο πρόβλημα, επιτρέπει στο μοντέλο να προσαρμοστεί καλύτερα στα ειδικά χαρακτηριστικά του νέου αυτού προβλήματος. Πρακτικά, με αυτή τη μέθοδο ενημερώνονται τα βάρη ξανά κατά τη διάρκεια της νέας εκπαίδευσης, συνδυάζοντας τη γνώση από τον πηγαίο πρόβλημα. Ουσιαστικά, η μέθοδος αυτή παρέχει μια πιο αποτελεσματική αρχικοποίηση (σε σχέση με την τυχαία αρχικοποίηση), χρησιμοποιώντας τα βάρη από το μοντέλο που έχει προηγουμένως εκπαιδευτεί.

Ο διαμοιρασμός παραμέτρων υποθέτει ότι οι παράμετροι είναι ιδιαίτερα μεταβιβάσιμες, κι έτσι αυτές του μοντέλου για το αρχικό πρόβλημα μπορούν να αντιγραφούν άμεσα στον τομέα-στόχο, όπου διατηρούνται "παγωμένες". Ο λεπτομερής συντονισμός υποθέτει ότι οι παράμετροι του αρχικού μοντέλου είναι χρήσιμες, αλλά πρέπει να συνεχίσουν να εκπαιδεύονται με περιορισμένα δεδομένα του τελικού προβλήματος έτσι ώστε να προσαρμοστούν καλύτερα

Δεδομένου ότι το δεν υπάρχει καμία εγγύηση ότι οι παράμετροι των νευρωνικών δικτύων για DSSE θα είναι σε μεγάλο βαθμό μεταβιβάσιμες για διαφορετικές τοπολογίες, η τεχνική μεταφοράς μάθησης που χρησιμοποιείται στην παρούσα εργασία είναι η μεταφορά παραμέτρων μέσω λεπτομερούς συντονισμού, για την υλοποίηση DSSE όταν αλλάζει η τοπολογία.

Feature Selection

Το feature selection (επιλογή χαρακτηριστικών) αναφέρεται στη διαδικασία επιλογής ενός υποσυνόλου από τα χαρακτηριστικά (ή μεταβλητές) ενός συνόλου δεδομένων με σκοπό τη βελτιστοποίηση της απόδοσης του μοντέλου μηχανικής μάθησης. Οι χαρακτηριστικές μεταβλητές είναι οι διάφορες παράμετροι που χρησιμοποιούνται για την εκπαίδευση ενός μοντέλου. Οι κύριοι λόγοι για την εφαρμογή feature selection περιλαμβάνουν:

- Την βελτιστοποίηση της απόδοσης, καθώς η επιλογή των σημαντικότερων χαρακτηριστικών μπορεί να οδηγήσει σε βελτιωμένη απόδοση του μοντέλου αφού μειώνεται ο αριθμός των χαρακτηριστικών που πρέπει να εκπαιδευτούν.
- Την μείωση της πολυπλοκότητας, καθώς η εξάλειψη των λιγότερο σημαντικών χαρακτηριστικών μπορεί να μειώσει την πολυπλοκότητα του μοντέλου, καθιστώντας το πιο ευέλικτο και εύκολο στη συντήρηση.
- Την εξοικονόμηση χρόνου και πόρων, αφού αφαιρώντας τα μη σημαντικά χαρακτηριστικά, εξοικονομείται χρόνος και υπολογιστικοί πόροι κατά την εκπαίδευση του μοντέλου.
- Την κωδικοποίηση των εγγενών συμμετριών που υπάρχουν στον χώρο εισόδου, το οποίο σημαίνει ότι προσπαθούμε να αναπαραστήσουμε τις πληροφορίες για τη δομή ή τις ομοιότητες

που υπάρχουν μεταξύ διαφορετικών μερών των δεδομένων εισόδου. Αυτό μπορεί να συμβάλει στη βελτίωση της απόδοσης των μοντέλων μηχανικής μάθησης, επειδή επιτρέπει στο μοντέλο να αντιληφθεί και να αξιοποιήσει τις συμμετρίες αυτές κατά τη διάρκεια της εκπαίδευσης.

Η κεντρική ιδέα της επιλογής χαρακτηριστικών είναι ότι τα δεδομένα περιέχουν κάποια χαρακτηριστικά που είναι είτε περιττά είτε ασήμαντα, και μπορούν να αφαιρεθούν χωρίς να προκαλείται σημαντική απώλεια πληροφοριών. Ένα χαρακτηριστικό θεωρείται περιττό όταν παρέχει παρόμοια πληροφορία με ένα άλλο χαρακτηριστικό. Δηλαδή, μπορεί να υπάρχει συσχέτιση μεταξύ δύο χαρακτηριστικών, και το ένα από αυτά είναι περιττό αν το άλλο έχει ήδη παρέχει σημαντική πληροφορία. Επίσης, ένα χαρακτηριστικό θεωρείται άσχετο όταν δεν παρέχει σημαντική πληροφορία για το πρόβλημα που επιχειρούμε να λύσουμε. Αν είναι ανεξάρτητο από το υπόλοιπο σύνολο χαρακτηριστικών και δεν συνεισφέρει στην εξήγηση των δεδομένων, μπορεί να θεωρηθεί άσχετο. Πρακτικά κατά την εκπαίδευση ενός μοντέλου η επιθυμητή ακρίβεια επιτυγχάνεται με έναν μικρό σχετικά αριθμό χαρακτηριστικών ενώ τα υπόλοιπα ουσιαστικά απλώς προσθέτουν θόρυβο.

Οι τεχνικές επιλογής χαρακτηριστικών χρησιμοποιούνται συχνά σε περιβάλλοντα όπου υπάρχουν πολλά χαρακτηριστικά και μεγάλη συσχέτιση μεταξύ τους. Αν το σύνολο δεδομένων περιέχει πολλά χαρακτηριστικά (π.χ., χαρακτηριστικά μιας εικόνας, λέξεις σε ένα κείμενο), μπορεί να είναι δύσκολο να αντιμετωπιστεί η διαστατικότητα, ιδίως όταν ο αριθμός των χαρακτηριστικών είναι πολύ μεγάλος. Αντιθέτως, Όταν ο αριθμός των δειγμάτων (ή σημείων δεδομένων) είναι σχετικά λίγος σε σχέση με τον αριθμό των χαρακτηριστικών, υπάρχει ο κίνδυνος της υπερεκπαίδευσης (overfitting), καθώς το μοντέλο μπορεί να μαθαίνει τον θόρυβο στα δεδομένα.

Ένας αλγόριθμος επιλογής χαρακτηριστικών μπορεί να θεωρηθεί ως ο συνδυασμός μιας τεχνικής αναζήτησης για την πρόταση νέων υποσυνόλων χαρακτηριστικών με ένα μέτρο αξιολόγησης που βαθμολογεί τα διαφορετικά υποσύνολα χαρακτηριστικών. Η επιλογή του μέτρου αξιολόγησης επηρεάζει σε μεγάλο βαθμό τον αλγόριθμο και είναι το αντικείμενο που διαφοροποιεί μεταξύ των τριών βασικών κατηγοριών αλγορίθμων επιλογής χαρακτηριστικών: Wrappers (Περιτυλίγματα), Filters (Φίλτρα) και Ενσωματωμένοι Αλγόριθμοι (Embedded Methods).

Οι αλγόριθμοι που ανήκουν στην κατηγορία Wrappers χρησιμοποιούν ένα μοντέλο πρόβλεψης για να αξιολογήσουν την απόδοση διαφορετικών υποσυνόλων χαρακτηριστικών. Συνήθως, χρησιμοποιούνται με αλγορίθμους προβλέψεων, όπως τα δέντρα αποφάσεων και επαναλαμβάνουν τη διαδικασία εκπαίδευσης και αξιολόγησης για διάφορα υποσύνολα χαρακτηριστικών. Το κύριο πλεονέκτημα τους είναι ότι μπορούν να λάβουν υπόψη τους το συνολικό περιβάλλον προβλήματος, αλλάζοντας δυναμικά τα χαρακτηριστικά σε κάθε επανάληψη.

Οι αλγόριθμοι της κατηγορίας των φίλτρων, βασίζονται σε στατιστικές μετρικές για την επιλογή χαρακτηριστικών, ανεξάρτητα από το μοντέλο πρόβλεψης. Αξιολογούν τη σχέση κάθε χαρακτηριστικού με την εξάγωνη τιμή, τον μέσο όρο, την αυτοσυσχέτιση Pearson ή άλλες τέτοιες παρόμοιες μετρικές. Το πλεονέκτημα τους είναι πως είναι συνήθως πιο

γρήγοροι σε σχέση με τους Wrappers, αλλά λαμβάνουν υπόψη τους μόνο τις στατιστικές ιδιότητες των δεδομένων. Επιπροσθέτως μπορούν να χρησιμοποιηθούν και σαν βήμα προεπεξεργασίας για μεθόδους περιτυλίγματος, επιτρέποντας τη χρήση τέτοιων αλγορίθμων σε μεγαλύτερα και πολυπλοκότερα προβλήματα.

Τέλος, οι ενσωματωμένοι αλγόριθμοι ενσωματώνονται άμεσα στη διαδικασία εκπαίδευσης του μοντέλου και αναπτύσσουν τα χαρακτηριστικά κατά τη διάρκεια αυτής της διαδικασίας. Συχνά, η επιλογή χαρακτηριστικών ενσωματώνεται στην αξιολόγηση του σφάλματος ή της συνάρτησης κόστους. Το κύριο πλεονέκτημα τους είναι πως μπορεί να έχουν καλύτερη απόδοση από τους άλλους αλγορίθμους επιλογής χαρακτηριστικών, καθώς είναι ενσωματωμένοι στη διαδικασία εκπαίδευσης.

Σχετικά με τις τεχνικές αναζήτησης, η πιο δημοφιλής προσέγγιση είναι η "greedy hill climbing", η οποία αξιολογεί ένα υποψήφιο υποσύνολο χαρακτηριστικών με επαναληπτικό τρόπο και στη συνέχεια τροποποιεί το υποσύνολο και ελέγχει αν το νέο υποσύνολο είναι βελτίωση σε σχέση με το παλιό. Η αξιολόγηση των υποσυνόλων απαιτεί ένα μία μετρική που θα αξιολογεί το υποσύνολο χαρακτηριστικών. Η μέθοδος αυτή είναι γενικά ανέφικτη, κυρίως λόγω υπερβολικά μεγάλου χρόνου διεξαγωγής, οπότε σε κάποιο σημείο που καθορίζεται από τον χειριστή το υποσύνολο των χαρακτηριστικών με το υψηλότερο σκορ που έχει εντοπιστεί μέχρι εκείνη τη στιγμή επιλέγεται ως το ικανοποιητικότερο υποσύνολο χαρακτηριστικών. Το κριτήριο διακοπής μπορεί να είναι το σκορ ενός υποσυνόλου που υπερβαίνει ένα όριο, που έχει υπερβεί το μέγιστο χρόνο λειτουργίας του προγράμματος, κ.α.πλ.

Άλλες τεχνικές αναζήτησης αποτελούν οι:

- Exhaustive Search (Εξαντλητική Αναζήτηση):
 - Λειτουργία: Αξιολογεί όλα τα πιθανά υποσύνολα χαρακτηριστικών.
 - Πλεονεκτήματα: Εγγυάται την εύρεση του βέλτιστου υποσυνόλου.
 - Μειονεκτήματα: Υπολογιστικά απαιτητική, ιδιαίτερα όταν το πλήθος των χαρακτηριστικών είναι μεγάλο.
- Genetic Algorithms (Γενετικοί Αλγόριθμοι):
 - Λειτουργία: Μιμούνται τη βιολογική εξέλιξη για την αναζήτηση βέλτιστων υποσυνόλων.
 - Πλεονεκτήματα: Καλύτερη εκτίμηση σε μεγάλους χώρους χαρακτηριστικών, ανθεκτικότητα σε τοπικά ελάχιστα.
 - Μειονεκτήματα: Απαιτείται πολύς υπολογιστικός χρόνος.
- Recursive Feature Elimination (RFE - Αναδρομική Εξάλειψη Χαρακτηριστικών):
 - Λειτουργία: Ξεκινάει με όλα τα χαρακτηριστικά και σε κάθε επανάληψη εξαιρεί το λιγότερο σημαντικό.
 - Πλεονεκτήματα: Κατάλληλο για μοντέλα που δυσκολεύονται με μεγάλο αριθμό χαρακτηριστικών.
 - Μειονεκτήματα: Αυξημένη πολυπλοκότητα μεγάλων συνόλων δεδομένων.
- Principal Component Analysis (PCA - Ανάλυση Κύριων Συνιστωσών):
 - Λειτουργία: Μετασχηματίζει τα δεδομένα σε ένα χώρο μικρότερης διαστατικότητας, διατηρώντας την πληροφορία.

- Πλεονεκτήματα: Εξαγωγή χαρακτηριστικών μειώνοντας τη διαστατικότητα.
- Μειονεκτήματα: Δεν παρέχει ερμηνευσιμότητα στα χαρακτηριστικά.
- **Best First:**
 - Λειτουργία: Ξεκινάει από μια αρχική κατάσταση και επιλέγει το υποσύνολο χαρακτηριστικών που οδηγεί στη βέλτιστη επίδοση σύμφωνα με ένα κριτήριο.
 - Πλεονεκτήματα: Γρήγορη σύγκλιση σε καλές λύσεις.
 - Μειονεκτήματα: Μπορεί να παγιδευτεί σε τοπικά ελάχιστα.
- **Variable Neighborhood Search (Αλγόριθμος Μεταβλητής Γειτονιάς):**
 - Λειτουργία: Κάνει αλλαγές στη γειτονιά της τρέχουσας λύσης για την εύρεση βέλτιστων χαρακτηριστικών.
 - Πλεονεκτήματα: Αποφεύγει τα τοπικά ελάχιστα, καλή προσαρμογή σε διάφορες καταστάσεις.
 - Μειονεκτήματα: Υπολογιστική πολυπλοκότητα.
- **Simulated Annealing (Σιμουλίτεντ Άνιλινγκ):**
 - Λειτουργία: Μοντελοποιεί τη διαδικασία ψύξης μετάλλων για την αναζήτηση του χώρου λύσεων.
 - Πλεονεκτήματα: Βοηθά στον αποφυγισμό τοπικών ελαχίστων, καλή εφαρμογή σε προβλήματα με πολλαπλές λύσεις.
 - Μειονεκτήματα: Εξαρτάται από παραμέτρους που πρέπει να ρυθμιστούν.
- **Scatter Search (Αναζήτηση Διασκορπισμένων Λύσεων):**
 - Λειτουργία: Συνδυάζει και επεκτείνει υποψήφιες λύσεις για την αναζήτηση βέλτιστων χαρακτηριστικών.
 - Πλεονεκτήματα: Αποτελεσματική όταν ο χώρος λύσεων είναι μεγάλος, ανθεκτική σε τοπικά ελάχιστα.
 - Μειονεκτήματα: Υπολογιστικά απαιτητική.

Sequential Forward Selection

Sequential forward selection (σειριακή επιλογή χαρακτηριστικών) ονομάζεται, μια τεχνική feature selection η οποία ανήκει στους άπληστους αλγορίθμους, η οποία προσθέτει ή αφαιρεί χαρακτηριστικά από ένα σύνολο δεδομένων επαναληπτικά, με σκοπό τη βελτίωση της απόδοσης ενός προβλεπτικού μοντέλου.

Η μέθοδος αυτή, έχει δύο εκδοχές, μπορεί να είναι είτε προώθησης (forward selection) είτε υποχώρησης (backward selection). Η διαδικασία που ακολουθείται για την πρώτη περίπτωση είναι η εξής:

1. Ο επιλογέας αρχικοποιείται με ένα προβλεπτικό μοντέλο, τον αριθμό των χαρακτηριστικών που πρέπει να επιλεγούν, τη μετρική βαθμολογίας, και ένα κενό υποσύνολο επιλεγμένων χαρακτηριστικών
2. Ο επιλογέας εκπαιδεύει το προβλεπτικό μοντέλο χρησιμοποιώντας το τρέχων υποσύνολο επιλεγμένων χαρακτηριστικών, το οποίο σε κάθε επανάληψη αποτελείται από τα ήδη επιλεγμένα χαρακτηριστικά από προηγούμενα βήματα και ένα ακόμα από τα μη επιλεγμένα.
3. Το μοντέλο αξιολογείται στο σύνολο εκπαίδευσης χρησιμοποιώντας τη μετρική βαθμολογίας που έχει προεπιλεγθεί.

4. Το χαρακτηριστικό που βελτιώνει περισσότερο το σκορ cross-validation του μοντέλου προστίθεται στα επιλεγμένα χαρακτηριστικά.
5. Ο επιλογέας επαναλαμβάνει τα βήματα 2-4 μέχρι να επιλεγεί ο επιθυμητός αριθμός χαρακτηριστικών.

Η διαδικασία αντιστρέφεται στην περίπτωση του sequential backward selection. Ο επιλογέας ξεκινά με ολόκληρο το σύνολο των χαρακτηριστικών και επαναληπτικά αφαιρεί το χαρακτηριστικό που επηρεάζει λιγότερο την απόδοση του προβλεπτικού μοντέλου. Η διαδικασία επαναλαμβάνεται μέχρι να επιλεγεί ο απαιτούμενος αριθμός χαρακτηριστικών ή μέχρι να μην είναι δυνατή η εξάλειψη περισσότερων χαρακτηριστικών χωρίς να μειώνεται σημαντικά η απόδοση του μοντέλου.

Στο πλαίσιο της εργασίας αυτής χρησιμοποιήθηκε η τεχνική sequential forward selection, η οποία εν κατακλείδι εξαλείφει τα χαρακτηριστικά με μικρή σημασία, ενισχύοντας την αποτελεσματικότητα του μοντέλου – δηλαδή αφαιρείται ο θόρυβος των περιττών δεδομένων. Συνήθως, η διαδικασία συνεχίζεται μέχρι να επιτευχθεί ένα προκαθορισμένο κριτήριο, όπως ο αριθμός των επιλεγμένων χαρακτηριστικών ή η αύξηση της απόδοσης (και τα δύο χρησιμοποιήθηκαν κατά την υλοποίηση).

Hierarchical Clustering

Μία άλλη τεχνική επιλογής χαρακτηριστικών που αξιοποιήθηκε σε αυτή την εργασία βασίζεται στο hierarchical clustering (ιεραρχική συσταδοποίηση). Το hierarchical clustering είναι μια τεχνική συσταδοποίησης που χρησιμοποιείται στην ανάλυση δεδομένων για τη δημιουργία ενός ιεραρχικού δέντρου συσταδοποίησης. Συσταδοποίηση ονομάζεται η διαδικασία ομαδοποίησης παρόμοιων δειγμάτων σε ομάδες, γνωστές και ως συστάδες.

Οι στρατηγικές για τον ιεραρχικό συσταδοποιητή κατατάσσονται κατά κύριο λόγο σε δύο κατηγορίες:

- Συγκλητικός (Agglomerative): Αυτή είναι μια προσέγγιση "από τα κάτω προς τα πάνω". Κάθε δεδομένο ξεκινά σε μια δική του συστάδα, και ζεύγη συστάδων συγχωνεύονται καθώς προχωράμε προς τα πάνω στην ιεραρχία.
- Διαιρετικός (Divisive): Αυτή είναι μια προσέγγιση "από τα πάνω προς τα κάτω". Όλες οι παρατηρήσεις ξεκινούν σε μία συστάδα, και στη συνέχεια εκτελούνται διαιρέσεις αναδρομικά καθώς προχωράμε προς τα κάτω στην ιεραρχία.

Συνολικά, οι συγχωνεύσεις και οι διαιρέσεις καθορίζονται με άπληστο τρόπο. Σε αυτή την εργασία χρησιμοποιήθηκε agglomerative clustering, η λειτουργία του οποίου αναλυτικότερα είναι:

1. Αρχικοποίηση: Κάθε δεδομένο ξεκινάει σε ένα δικό του σύνολο-συστάδα.
2. Απόσταση Μεταξύ Συνόλων-Συστάδων: Υπολογίζεται η απόσταση μεταξύ όλων των διαφορετικών συνόλων-συστάδων.
3. Συγχώνευση: Συγχωνεύονται τα δύο σύνολα-συστάδες με τη μικρότερη απόσταση μεταξύ τους.

4. Επανάληψη: Επαναλαμβάνονται τα βήματα 2 και 3 μέχρι να δημιουργηθεί μια μεγάλη συστάδα που περιέχει όλες τις παρατηρήσει ή μέχρι ο αριθμός των συστάδων να φτάσει μία προκαθορισμένη τιμή.

Κατά τη διάρκεια της διαδικασίας, ο αλγόριθμος κατασκευάζει ένα δενδρόγραμμα (dendrogram), το οποίο απεικονίζει τις συνολικές συγχωνεύσεις και τις ιεραρχικές σχέσεις μεταξύ των συστάδων. Το δενδρόγραμμα αυτό μπορεί να χρησιμοποιηθεί για τον καθορισμό του αριθμού των συστάδων που θα επιλεγούν, ανάλογα με το επίπεδο κοπής (cutting level).

Το hierarchical clustering έχει το πλεονέκτημα ότι μπορεί να χρησιμοποιήσει οποιαδήποτε έγκυρη μετρική απόστασης. Στην πραγματικότητα, οι παρατηρήσεις δεν απαιτούνται: χρησιμοποιείται απλώς ένας πίνακας αποστάσεων, ο οποίος δείχνει πόσο “απέχουν” , δηλαδή σχετίζονται τα χαρακτηριστικά μεταξύ τους. Γ

Αναλυτικότερα, για να αποφασίσουμε ποιες συστάδες θα ενωθούν (στην περίπτωση του agglomerative clustering) ή πού μια συστάδα θα διασπαστεί (στην περίπτωση του divisive clustering), απαιτείται ένα μέτρο ανομοιογένειας μεταξύ του συνόλου των παρατηρήσεων - δεδομένων. Στις περισσότερες μεθόδους ιεραρχικής συσταδοποίησης, αυτό επιτυγχάνεται με τη χρήση μιας κατάλληλης απόστασης, όπως η Ευκλείδεια απόσταση, μεταξύ μεμονωμένων παρατηρήσεων του συνόλου δεδομένων και ενός κριτηρίου συνδεσμολογίας (linkage criterion), το οποίο καθορίζει την ανομοιότητα των συνόλων ως συνάρτηση των αντιστοιχιών των αποστάσεων μεταξύ των παρατηρήσεων στα σύνολα.

Η επιλογή της μετρικής καθώς και του συνδεσμολογικού κριτηρίου έχει σημαντική επίδραση στο αποτέλεσμα της συσταδοποίησης, καθώς η μετρική της απόστασης καθορίζει ποια αντικείμενα είναι πιο παρόμοια, ενώ το κριτήριο συνδεσμολογίας επηρεάζει το σχήμα των συστάδων. Για παράδειγμα, ο τύπος συνδέσμου complete-linkage τείνει να εξάγει πιο σφαιρικές συστάδες σε σχέση με τον single-linkage.

Οι πιο κοινά τύποι αποστάσεων που χρησιμοποιούνται περιλαμβάνουν:

- Ευκλείδεια Απόσταση (Euclidean Distance): Είναι η κλασική ευκλείδεια απόσταση μεταξύ δύο σημείων στον n -διάστατο χώρο.
- Απόσταση Manhattan (Manhattan Distance): Είναι το άθροισμα των απολύτων διαφορών μεταξύ των αντίστοιχων συντεταγμένων.
- Απόσταση Χειρονομίας (Chebyshev Distance): Είναι η μέγιστη απόλυτη διαφορά μεταξύ των αντίστοιχων συντεταγμένων.
- Απόσταση Πόλης (City Block Distance): Είναι η αθροιστική απόσταση, παρόμοια με την απόσταση Manhattan.

Τα πιο κοινά κριτήρια σύνδεσης (Linkage Tree) που χρησιμοποιούνται περιλαμβάνουν:

- Complete-Linkage (Σύνδεσμος Πλήρους Συνδέσμου): Συνδέει δύο συστάδες με βάση το μέγιστο μήκος απόστασης μεταξύ οποιωνδήποτε δύο σημείων, ένα από κάθε σύνολο. Δηλαδή, επιλέγει το ζεύγος σημείων που έχουν τη μεγαλύτερη απόσταση μεταξύ τους, και συνδέει τις δύο συστάδες με βάση αυτήν τη μέγιστη απόσταση.
- Single-Linkage (Σύνδεσμος Μοναδικού Συνδέσμου): Συνδέει δύο συστάδες με βάση το ελάχιστο μήκος απόστασης μεταξύ

οποιαδήποτε δύο σημείων, ένα από κάθε σύνολο. Δηλαδή, επιλέγει το ζεύγος σημείων που έχουν την ελάχιστη απόσταση μεταξύ τους και συνδέει τις δύο συστάδες με βάση αυτήν την ελάχιστη απόσταση.

- Average-Linkage (Σύνδεσμος Μέσου Συνδέσμου): Συνδέει δύο συστάδες με βάση το μέσο όρο των αποστάσεων μεταξύ όλων των δυνατών ζευγαριών σημείων, ένα από κάθε σύνολο. Υπολογίζει τον μέσο όρο των αποστάσεων μεταξύ όλων των σημείων που ανήκουν στις δύο συστάδες και συνδέει τις δύο συστάδες με βάση αυτόν τον μέσο όρο.
- Centroid Linkage (Σύνδεσμος Κέντρου Μάζας): Συνδέει τις συστάδες με βάση την απόσταση μεταξύ των κέντρων μάζας τους. Το κέντρο μάζας είναι το μέσο σημείο όλων των σημείων σε μια συστάδα.
- Median Linkage (Σύνδεσμος Μέσης Τιμής): Συνδέει τις συστάδες με βάση την απόσταση μεταξύ των μεσαίων τιμών τους.
- Κριτήριο Συνδεσμολογίας Πληθυσμού: Αυτό το κριτήριο λαμβάνει υπόψη τον αριθμό των σημείων σε κάθε σύνολο κατά τη σύνδεση. Για παράδειγμα, μπορεί να επιλέξει να συνδέσει συστάδες με μεγαλύτερο πληθυσμό προτιμώντας την αύξηση του πληθυσμού στις νέες συνδεδεμένες συστάδες.
- Ward's Method (Μέθοδος Ward): Αυτή η μέθοδος επιδιώκει να μειώσει τη διακύμανση εντός των συστάδων, προσπαθώντας να δημιουργήσει συστάδες που είναι ομοιόμορφες και συμπαγείς. Επιλέγει να συνδέσει τις συστάδες που, όταν ενωθούν, προκαλούν τη μικρότερη αύξηση στη συνολική διακύμανση. Η λειτουργία της μεθόδου Ward είναι η εξής:
 - Εκκίνηση: Αρχικά, κάθε παρατήρηση θεωρείται ως μια ξεχωριστή συστάδα.
 - Υπολογισμός των Κέντρων Μάζας: Για κάθε συστάδα υπολογίζεται το κέντρο μάζας, που είναι το μέσο σημείο όλων των παρατηρήσεων στη συστάδα.
 - Υπολογισμός της Διακύμανσης Εντός των Συστάδων: Υπολογίζεται η διακύμανση εντός των συστάδων, που αντιπροσωπεύει το πόσο "συμπαγείς" είναι οι παρατηρήσεις μέσα σε κάθε συστάδα. Αυτή η διακύμανση μπορεί να υπολογιστεί χρησιμοποιώντας διάφορες μεθόδους, αλλά η μέθοδος αυτή χρησιμοποιεί τη διαφορά των τετραγώνων των ευκλείδειων αποστάσεων μεταξύ του κέντρου μάζας και κάθε παρατήρησης, η οποία δίνεται από τον τύπο:

$$d(r, s) = \sqrt{\frac{2n_r n_s}{(n_r + n_s)}} \|\bar{x}_r - \bar{x}_s\|_2$$

Όπου $\|\cdot\|_2$ είναι η ευκλείδεια απόσταση, \bar{x} είναι τα κέντρα μάζας των συστάδων και n ο αριθμός των στοιχείων που καθεμία περιέχει.

- Επιλογή των Συστάδων προς Σύνδεση: Συνδέονται οι δύο συστάδες που η σύνδεσή τους θα οδηγήσει στη μικρότερη αύξηση της διακύμανσης εντός των συστάδων. Συγκεκριμένα, επιλέγονται αυτές οι δύο συστάδες που έχουν το ελάχιστο

κόστος σύνδεσης, με βάση τη διαφορά των διακυμάνσεων και του πληθυσμού τους.

Όπως προαναφέρθηκε, η μέθοδος Ward τείνει να δημιουργεί συστάδες που είναι ομοιόμορφες (με χαμηλή διακύμανση εντός της συστάδας) και συμπαγείς, καθιστώντας την κατάλληλη για περιπτώσεις όπου οι δημιουργούμενες συστάδες ζητείται να αντικατοπτρίζουν ομοιόμορφα κατανεμημένα δεδομένα. Επίσης, είναι πιο αποδοτική όταν οι συστάδες έχουν σχεδόν σφαιρικό σχήμα, δηλαδή όταν οι παρατηρήσεις μέσα σε μια συστάδα είναι ομοιόμορφα κατανεμημένες και έχουν περίπου ίδιο ακτίνα από το κέντρο της συστάδας.

Η μέθοδος Ward επιλέγεται συχνά σε περιπτώσεις συσταδοποίησης όταν ο χρήστης θέλει να εξετάσει τη δομή των συστάδων σε ένα σύνολο δεδομένων. Η Ward τείνει να δημιουργεί συστάδες που είναι ομοιόμορφες και συμπαγείς και σφαιρικές, κι έτσι θεωρούμε πως ο κοινός κόμβος κάθε τέτοιας συστάδας αν επιλεγθεί θα μπορεί να περιγράφει τη συστάδα του αρκετά καλύτερα σε σχέση με άλλες μεθόδους που δεν καταλήγουν σε ομοιόμορφες και συμπαγείς ομάδες, αφού το ζητούμενο είναι η επιλογή ενός μόνο κόμβου κάθε συστάδας από τον οποίο θα λαμβάνουμε δεδομένα. Επίσης καθώς οι περισσότερες συστάδες που δημιουργούνται είναι σφαιρικές, ο κοινός κόμβος βρισκόμενος κοντά στο κέντρο της κάθε σφαίρας θα αποτελεί άλλη μια συνθήκη που θα διασφαλίζει την επιλογή του με αυτή τη μέθοδο ως καλύτερη.

Παρά όλα αυτά, εκτός από την περίπτωση της απόστασης μοναδικού συνδέσμου (single linkage), κανένας από τους αλγόριθμους δεν μπορεί να εγγυηθεί την εύρεση της βέλτιστης λύσης.

Έτσι, επιπλέον, για λόγους σύγκρισης και επαλήθευσης των αποτελεσμάτων, διερευνήθηκε και η τεχνική k-means clustering για την επιλογή χαρακτηριστικών, η οποία αποτελεί έναν πιο παραδοσιακό αλγόριθμο clustering.

Η τεχνική αυτή, αρχικά προερχόμενη από την επεξεργασία σημάτων, χρησιμοποιείται για να ομαδοποιήσει ένα σύνολο δεδομένων σε k συστάδες. Ο στόχος της είναι να μειώσει τη διακύμανση εντός των συστάδων, ορίζοντας κέντρα συστάδας έτσι ώστε η απόσταση των παρατηρήσεων από αυτά να είναι ελάχιστη. Αναλυτικότερα, ο αλγόριθμος είναι ο εξής:

1. Επιλογή του Αριθμού των Συστάδων (k): καθορίζεται προκαταρκτικά ο αριθμός των συστάδων που θέλουμε να δημιουργηθούν.
2. Αρχικοποίηση των Κέντρων Συστάδας: Επιλέγονται τυχαία k παρατηρήσεις από το σύνολο δεδομένων ως αρχικά κέντρα των συστάδων.
3. Ανάθεση στα Σημεία στις Συστάδες: Κάθε παρατήρηση ανατίθεται στη συστάδα που έχει το κοντινότερο κέντρο.
4. Ενημέρωση των Κέντρων Συστάδας: Τα κέντρα των συστάδων ενημερώνονται ως ο μέσος όρος των παρατηρήσεων που τους έχουν ανατεθεί.
5. Επανάληψη των Βημάτων 3 και 4: Οι παρατηρήσεις ανατίθενται ξανά στις συστάδες, και τα κέντρα των συστάδων ενημερώνονται. Η διαδικασία επαναλαμβάνεται μέχρι τα κέντρα των συστάδων να σταθεροποιηθούν και η ανάθεση των παρατηρήσεων να μην αλλάζει σημαντικά.

Ο αλγόριθμος K-Means στοχεύει στο να ελαχιστοποιήσει τις διακυμάνσεις εντός των συστάδων, χρησιμοποιώντας τετραγωνικές Ευκλείδειες αποστάσεις μεταξύ των παρατηρήσεων.

Είναι μία αρκετά απλή μέθοδος στην υλοποίηση και αποδοτική σε πολύ μεγάλα σύνολα δεδομένων, ενώ επίσης μπορεί να κλιμακωθεί σε μεγάλες διαστάσεις και σε μεγάλα σύνολα δεδομένων. Επιπλέον τα αποτελέσματα είναι σχετικά εύκολα να οπτικοποιηθούν με τη χρήση διαγραμμάτων που δείχνουν τις συστάδες και τα κέντρα τους και δεν απαιτούνται στη χρήση της προκαθορισμένες ετικέτες για τη συσταδοποίηση, καθιστώντας την κατάλληλη για ανεπτυγμένα και μη επιβλεπόμενα σύνολα δεδομένων.

Το κύριο μειονέκτημα της τεχνικής αυτής είναι πως η απόδοσή της εξαρτάται σε μεγάλο βαθμό από την αρχικοποίηση των κέντρων των συστάδων στα πρώτα βήματα, καθώς μπορεί να επηρεάσει την τελική δομή των συστάδων και τα αποτελέσματα της συσταδοποίησης, εισάγοντας έναν παράγοντα τυχαιότητας που αναπόφευκτα μειώνει την πιθανότητα της εύρεσης των βέλτιστων συστάδων.

Spearman Correlation Coefficient

Το Spearman's rank correlation coefficient είναι ένας στατιστικός δείκτης που μετρά τον βαθμό στον οποίο οι αξίες δύο μεταβλητών συσχετίζονται, ανεξάρτητα από την ακριβή φύση της σχέσης τους, χρησιμοποιώντας μια μονοτονική συνάρτηση.

Ο υπολογισμός του Spearman's rank correlation coefficient απαιτεί τη μετατροπή των πρωτογενών δεδομένων σε αξιολογήσεις (ranks) και την έπειτα εφαρμογή μιας συσχέτισης Pearson στις αξιολογήσεις. Αναλυτικότερα, ο τρόπος υπολογισμού του Spearman's rank correlation coefficient περιλαμβάνει τα εξής βήματα:

1. Αξιολόγηση (Ranking): Για κάθε μεταβλητή, αναθέτουμε αξιολογήσεις (ranks) βάσει των τιμών τους. Δηλαδή, η μικρότερη τιμή παίρνει το ρυθμό 1, η δεύτερη μικρότερη παίρνει το ρυθμό 2, και ούτω καθεξής. Αν υπάρχουν επαναλαμβανόμενες τιμές, η αξιολόγηση τους είναι το μέσο όρο των θέσεών τους.
2. Υπολογισμός των διαφορών (Differences): Για κάθε ζεύγος δεδομένων, υπολογίζουμε τη διαφορά των αξιολογήσεών τους.
3. Υψωση στο τετράγωνο (Square Differences): Υψώνουμε κάθε διαφορά στο τετράγωνο.
4. Άθροισμα των τετραγώνων (Sum of Squares): Προσθέτουμε όλα τα τετράγωνα διαφορών.
5. Ο υπολογισμός γίνεται με τον τύπο:

$$\rho = 1 - \frac{6 \sum D^2}{n(n^2-1)}$$

όπου D είναι η διαφορά των αξιολογήσεων και n είναι ο αριθμός των δεδομένων.

Το Spearman's rank correlation coefficient δείχνει την κατεύθυνση της συσχέτισης μεταξύ της μεταβλητής X (ανεξάρτητη μεταβλητή) και της μεταβλητής Y (εξαρτημένη μεταβλητή). Εάν η Y τείνει να αυξάνεται όταν η X αυξάνεται, ο συντελεστής συσχέτισης Spearman είναι θετικός. Εάν η Y

τείνει να μειώνεται όταν η X αυξάνεται, ο συντελεστής συσχέτισης Spearman είναι αρνητικός. Ένας συντελεστής Spearman ίσος με μηδέν υποδηλώνει ότι δεν υπάρχει τάση για την αύξηση ή μείωση της Y όταν αυξάνεται η X . Ο συντελεστής Spearman αυξάνεται σε μέγεθος όσο η X και η Y γίνονται πιο κοντά σε μια τέλεια μονοτονική συνάρτηση μεταξύ τους. Έτσι, εάν δεν υπάρχουν επαναλαμβανόμενες τιμές δεδομένων, μια τέλεια Spearman συσχέτιση του $+1$ ή -1 συμβαίνει όταν κάθε μεταβλητή είναι μια τέλεια μονοτονική συνάρτηση της άλλης. Μια τέλεια μονοτονική αυξανόμενη σχέση σημαίνει ότι, για οποιοδήποτε ζεύγος τιμών δεδομένων X_i, Y_i και X_j, Y_j , οι διαφορές $X_i - X_j$ και $Y_i - Y_j$ έχουν πάντα το ίδιο πρόσημο. Μια τέλεια μονοτονική φθίνουσα σχέση σημαίνει ότι αυτές οι διαφορές έχουν πάντα αντίθετα πρόσημα.

Τα πλεονεκτήματα που προσφέρει ο συντελεστής Spearman's σε σύγκριση με άλλες μεθόδους είναι:

1. Μη Παραμετρικός Χαρακτήρας: Ο Spearman's ρ είναι μη παραμετρικός, που σημαίνει ότι δεν απαιτεί την υπόθεση κανονικής κατανομής των δεδομένων. Αυτό το καθιστά χρήσιμο όταν τα δεδομένα δεν ακολουθούν μια κανονική κατανομή. Επίσης, μια άλλη έννοια στην οποία ο συντελεστής συσχέτισης Spearman είναι μη παραμετρικός είναι ότι η ακριβής δειγματοληπτική του κατανομή μπορεί να προκύψει χωρίς να απαιτείται γνώση (δηλαδή, γνώση των παραμέτρων) της συνδυαστικής πιθανοκατανομής της X και της Y .
2. Αντισταθμίζει τυχαίες αποκλίσεις: Είναι λιγότερο επηρεάσιμος από ακραίες τιμές και τυχαίες αποκλίσεις σε σύγκριση με τον συντελεστή Pearson. Αυτό το καθιστά ανθεκτικό σε ανωμαλίες στα δεδομένα.
3. Κατάλληλο για Τακτοποιημένα Δεδομένα: Είναι κατάλληλο για τακτοποιημένα (ordinal) δεδομένα όπου η σειρά των τιμών είναι σημαντική, αλλά η ακριβής απόσταση μεταξύ τους δεν είναι γνωστή, εξού και προκύπτει η χρήση της σε συστήματα διανομής.

Ωστόσο, για λόγους υπολογιστικής ευκολίας, πολλές φορές γίνεται χρήση του συντελεστή συσχέτισης Pearson για την εκτίμηση του Spearman.

Pearson Correlation Coefficient

Ο συντελεστής συσχέτισης Pearson είναι ένας μέτρος που περιγράφει το βαθμό της γραμμικής σχέσης μεταξύ δύο μεταβλητών. Ορίζεται ως ο λόγος μεταξύ της συνδιακύμανσης δύο μεταβλητών και του γινομένου των τυπικών αποκλίσεων τους. Επομένως, είναι ουσιαστικά μια κανονικοποιημένη μέτρηση της συνδιακύμανσης, έτσι ώστε το αποτέλεσμα να έχει πάντα τιμή μεταξύ -1 και 1 , όπου:

- $+1$ υποδεικνύει την τέλεια θετική γραμμική σχέση.
- -1 υποδεικνύει την τέλεια αρνητική γραμμική σχέση.
- 0 υποδεικνύει απουσία γραμμικής σχέσης.

Ο υπολογισμός του συντελεστή συσχέτισης Pearson μεταξύ δύο μεταβλητών X και Y γίνεται ως εξής:

1. Υπολογισμός του μέσου (mean) των τιμών των X και Y.
2. Για κάθε ζεύγος τιμών (X_i, Y_i) υπολογισμός του όρου (X_i - \bar{X}) * (Y_i - \bar{Y}), όπου \bar{X} είναι το μέσο των τιμών του X και \bar{Y} είναι το μέσο των τιμών του Y.
3. Υπολογισμός του όρου (X_i - \bar{X})² για κάθε τιμή του X και (Y_i - \bar{Y})² για κάθε τιμή του Y.
4. Το σύνολο των όρων που υπολογίστηκαν διαιρείται με το γινόμενο των τυχαίων μεταβλητών.

Ο τύπος του συντελεστή συσχέτισης Pearson (r) είναι:

$$r = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}}$$

Όπου, X_i και Y_i είναι οι τιμές των μεταβλητών X και Y αντίστοιχα και \bar{X} και \bar{Y} είναι τα μέσα των τιμών των μεταβλητών X και Y αντίστοιχα.

Ο συντελεστής συσχέτισης Pearson είναι ιδιαίτερα χρήσιμος για τη μέτρηση της σχέσης μεταξύ δύο μεταβλητών, όταν αυτή η σχέση είναι γραμμική.

Η Spearman's συσχέτιση μεταξύ δύο μεταβλητών είναι ίση με τη συσχέτιση Pearson μεταξύ των βαθμολογημένων τιμών αυτών των δύο μεταβλητών. Ενώ η συσχέτιση του Pearson αξιολογεί γραμμικές σχέσεις, η συσχέτιση του Spearman αξιολογεί μονοτονικές σχέσεις (είτε γραμμικές είτε όχι).

Η προσέγγιση του συντελεστή spearman με τον pearson βασίζεται στο γεγονός ότι, αν οι κατανομές των δύο μεταβλητών είναι κανονικές, η σχέση μεταξύ των αξιολογήσεων της σειράς (ranks) και των ακριβών τιμών τους μπορεί να προσεγγιστεί από μια γραμμική συνάρτηση. Στην περίπτωση αυτή, ο συντελεστής συσχέτισης Pearson θα προσφέρει μια καλή προσέγγιση του Spearman.

Πρακτικό μέρος

Κεφάλαιο IV

Κατασκευή Dataset

Όπως προαναφέρθηκε, οι μετρήσεις των έξυπνων μετρητών (IEDs) γίνονται διαθέσιμες μετά από καθυστέρηση τουλάχιστον μερικών ωρών, γεγονός που σημαίνει ότι δεν μπορούν να χρησιμοποιηθούν άμεσα για DSSE σε πραγματικό χρόνο. Επομένως, αφού αυτού του είδους οι μετρήσεις, οι οποίες όπως είδαμε είναι πολύ συνηθισμένες, δεν μπορούν να χρησιμοποιηθούν, η προτεινόμενη μεθοδολογία χρησιμοποιεί δεδομένα προσομοιώσεων monte carlo.

στη διαδικασία εκπαίδευσης του DNN.

Για την δημιουργία του Dataset για την εκπαίδευση των μοντέλων πρόβλεψης χρησιμοποιήθηκε το πρόγραμμα Matlab, το οποίο ένα περιβάλλον προγραμματισμού πρωτίστως σχεδιασμένο για αριθμητικούς υπολογισμούς, ανάλυση δεδομένων και οπτικοποίηση αποτελεσμάτων, το οποίο παράλληλα παρέχει και μια υψηλού επιπέδου γλώσσα προγραμματισμού η οποία είναι γνωστή για την ευκολία με την οποία μέσω της αντιμετωπίζονται μαθηματικά και μηχανικά προβλήματα. Αυτό συμβαίνει επειδή σχεδιάστηκε για να είναι ευανάγνωστη και παρόμοια με τη μαθηματική γραφή κι έτσι είναι βασισμένη σε λειτουργίες πινάκων και διανυσμάτων.

Ένα από τα βασικά χαρακτηριστικά του Matlab είναι το εκτεταμένο σύνολο ενσωματωμένων συναρτήσεων και εργαλείων που καλύπτουν μια ευρεία γκάμα εφαρμογών, από επεξεργασία σημάτων και ανάλυση εικόνας έως συστήματα ελέγχου και μηχανική μάθηση. Αυτά τα εργαλεία επεκτείνουν την κεντρική λειτουργικότητα του Matlab, επιτρέποντας στους χρήστες να αξιοποιούν εξειδικευμένους αλγόριθμους και συναρτήσεις χωρίς να χρειάζεται να τις υλοποιήσουν από την αρχή.

Έτσι, για την παραγωγή των δεδομένων χρησιμοποιήθηκε το πακέτο MATPOWER το οποίο αποτελεί ένα πακέτο ανοικτού κώδικα για προσομοίωση και βελτιστοποίηση συστημάτων ηλεκτρικής ενέργειας που παρέχεται στα πλαίσια του Matlab. Ένα από τα κύρια χαρακτηριστικά του MATPOWER είναι η ικανότητά του να μοντελοποιεί και να προσομοιώνει διάφορες πτυχές των ηλεκτρικών συστημάτων, συμπεριλαμβανομένης της ανάλυσης της ροής ισχύος, της βέλτιστης ροής ισχύος και της ανάλυσης επικινδυνότητας. Σε αυτή την εργασία χρησιμοποιήθηκε μόνο για την ανάλυση της ροής ισχύος έτσι ώστε να προκύψουν τα δεδομένα που κατόπιν τροφοδοτήθηκαν για την εκπαίδευση των μοντέλων.

Αυτό έγινε μέσω της χρήση της εντολής:

```
results = runpf(casedata);
```

Αυτό σημαίνει πως η επίλυση της ροής ισχύος εκτελείται καλώντας το runpf με ως πρώτο όρισμα το όνομα του φορτωμένου αρχείου του συστήματος (casedata). Εκτός από την εκτύπωση της εξόδου στην οθόνη, που είναι η προεπιλεγμένη συμπεριφορά, η συνάρτηση runpf μπορεί προαιρετικά και να επιστρέψει τη λύση σε μια δομημένη μορφή που ονομάζεται "δομή αποτελεσμάτων" (results structure). Structures – δομές

ονομάζεται στην matlab ένας τύπος δεδομένων που επιτρέπει την ομαδοποίηση συναφή δεδομένων χρησιμοποιώντας πεδία με όνομα. Αυτή η δομή αποτελεσμάτων περιέχει λεπτομερείς πληροφορίες σχετικά με τη λύση της ροής ισχύος, συμπεριλαμβανομένων μεταβλητών όπως τάσεις διαύλων, ροές ισχύος κλάδων και άλλα σχετικά δεδομένα κι έτσι διευκολύνει τους χρήστες να έχουν πρόσβαση και να αναλύουν τα αποτελέσματα της ανάλυσης ροής ισχύος προγραμματιστικά. Η προεπιλεγμένη μέθοδος η οποία εφαρμόστηκε και σε αυτή την εργασία είναι η Newton – Raphson. Στο σχήμα παρακάτω παρατίθεται ένα παράδειγμα των αποτελεσμάτων της επίλυσης ροής φορτίου μέσω αυτής της εντολής:

```
MATPOWER Version 7.1, 08-Oct-2020 -- AC Power Flow (Newton)
```

```
Newton's method power flow (power balance, polar) converged in 4 iterations.
```

```
Converged in 1.60 seconds
```

```
=====
|      System Summary
|=====
```

How many?		How much?	P (MW)	Q (MVar)
Buses	33	Total Gen Capacity	10.0	-10.0 to 10.0
Generators	1	On-line Capacity	10.0	-10.0 to 10.0
Committed Gens	1	Generation (actual)	3.9	2.4
Loads	32	Load	3.7	2.3
Fixed	32	Fixed	3.7	2.3
Dispatchable	0	Dispatchable	-0.0 of -0.0	-0.0
Shunts	0	Shunt (inj)	-0.0	0.0
Branches	37	Losses (I ² * Z)	0.21	0.14
Transformers	0	Branch Charging (inj)	-	0.0
Inter-ties	0	Total Inter-tie Flow	0.0	0.0
Areas	1			

	Minimum	Maximum
Voltage Magnitude	0.904 p.u. @ bus 18	1.000 p.u. @ bus 1
Voltage Angle	-0.69 deg @ bus 18	0.50 deg @ bus 30
P Losses (I ² *R)	-	0.05 MW @ line 2-3
Q Losses (I ² *X)	-	0.03 MVar @ line 5-6

Bus Data						
Bus #	Voltage		Generation		Load	
	Mag(pu)	Ang(deg)	P (MW)	Q (MVAr)	P (MW)	Q (MVAr)
1	1.000	0.000*	3.93	2.44	-	-
2	0.997	0.015	-	-	0.10	0.06
3	0.983	0.097	-	-	0.09	0.04
4	0.975	0.163	-	-	0.12	0.08
5	0.968	0.230	-	-	0.06	0.03
6	0.949	0.136	-	-	0.06	0.02
7	0.946	-0.096	-	-	0.20	0.10
8	0.932	-0.249	-	-	0.20	0.10
9	0.926	-0.324	-	-	0.06	0.02
10	0.920	-0.388	-	-	0.06	0.02
11	0.919	-0.380	-	-	0.04	0.03
12	0.918	-0.368	-	-	0.06	0.04
13	0.912	-0.462	-	-	0.06	0.04
14	0.909	-0.542	-	-	0.12	0.08
15	0.908	-0.580	-	-	0.06	0.01
16	0.906	-0.604	-	-	0.06	0.02
17	0.904	-0.683	-	-	0.06	0.02
18	0.904	-0.693	-	-	0.09	0.04
19	0.996	0.004	-	-	0.09	0.04
20	0.993	-0.063	-	-	0.09	0.04
21	0.992	-0.083	-	-	0.09	0.04
22	0.992	-0.103	-	-	0.09	0.04
23	0.979	0.066	-	-	0.09	0.05
24	0.973	-0.023	-	-	0.42	0.20
25	0.969	-0.067	-	-	0.42	0.20
26	0.948	0.175	-	-	0.06	0.03
27	0.945	0.232	-	-	0.06	0.03
28	0.934	0.315	-	-	0.06	0.02
29	0.925	0.393	-	-	0.12	0.07
30	0.922	0.498	-	-	0.20	0.60
31	0.918	0.413	-	-	0.15	0.07
32	0.917	0.390	-	-	0.21	0.10
33	0.916	0.383	-	-	0.06	0.04
Total:			3.93	2.44	3.72	2.30

Branch Data										
Brnch #	From Bus	To Bus	From Bus P (MW)	Injection Q (MVar)	To Bus P (MW)	Injection Q (MVar)	Loss ($I^2 * Z$)			
							P (MW)	Q (MVar)		
1	1	2	3.93	2.44	-3.91	-2.44	0.012	0.01		
2	2	3	3.45	2.22	-3.40	-2.19	0.052	0.03		
3	3	4	2.37	1.69	-2.35	-1.68	0.020	0.01		
4	4	5	2.23	1.60	-2.21	-1.59	0.019	0.01		
5	5	6	2.15	1.56	-2.11	-1.53	0.039	0.03		
6	6	7	1.10	0.54	-1.10	-0.53	0.002	0.01		
7	7	8	0.90	0.43	-0.89	-0.42	0.012	0.01		
8	8	9	0.69	0.32	-0.68	-0.32	0.004	0.00		
9	9	10	0.62	0.30	-0.62	-0.29	0.004	0.00		
10	10	11	0.56	0.27	-0.56	-0.27	0.001	0.00		
11	11	12	0.52	0.24	-0.51	-0.24	0.001	0.00		
12	12	13	0.45	0.21	-0.45	-0.21	0.003	0.00		
13	13	14	0.39	0.17	-0.39	-0.17	0.001	0.00		
14	14	15	0.27	0.09	-0.27	-0.09	0.000	0.00		
15	15	16	0.21	0.08	-0.21	-0.08	0.000	0.00		
16	16	17	0.15	0.06	-0.15	-0.06	0.000	0.00		
17	17	18	0.09	0.04	-0.09	-0.04	0.000	0.00		
18	2	19	0.36	0.16	-0.36	-0.16	0.000	0.00		
19	19	20	0.27	0.12	-0.27	-0.12	0.001	0.00		
20	20	21	0.18	0.08	-0.18	-0.08	0.000	0.00		
21	21	22	0.09	0.04	-0.09	-0.04	0.000	0.00		
22	3	23	0.94	0.46	-0.94	-0.46	0.003	0.00		
23	23	24	0.85	0.41	-0.84	-0.40	0.005	0.00		
24	24	25	0.42	0.20	-0.42	-0.20	0.001	0.00		
25	6	26	0.95	0.97	-0.95	-0.97	0.003	0.00		
26	26	27	0.89	0.95	-0.88	-0.95	0.003	0.00		
27	27	28	0.82	0.92	-0.81	-0.91	0.011	0.01		
28	28	29	0.75	0.89	-0.75	-0.88	0.008	0.01		
29	29	30	0.63	0.81	-0.62	-0.81	0.004	0.00		
30	30	31	0.42	0.21	-0.42	-0.21	0.002	0.00		
31	31	32	0.27	0.14	-0.27	-0.14	0.000	0.00		
32	32	33	0.06	0.04	-0.06	-0.04	0.000	0.00		
33	21	8	0.00	0.00	0.00	0.00	0.000	0.00		
34	9	15	0.00	0.00	0.00	0.00	0.000	0.00		
35	12	22	0.00	0.00	0.00	0.00	0.000	0.00		
36	18	33	0.00	0.00	0.00	0.00	0.000	0.00		
37	25	29	0.00	0.00	0.00	0.00	0.000	0.00		
Total:								0.211	0.14	

Για να αναπαραχθούν οι πραγματικές μετρήσεις SMD, πρέπει να κατάλληλα σφάλμα μέτρησης πρέπει να προστεθούν στις τάσεις και στα ρεύματα που λαμβάνονται από τη λύση ροής ισχύος. Σύμφωνα με το πρότυπο της IEEE που περιγράφεται εδώ [70], τα συνολικά διανυσματικά σφάλματα (TVE) των SMDs πρέπει να πληρούν την απαίτηση να είναι έως 1%. Έχει επίσης ευρέως υποτεθεί ότι τα σφάλματα στα δεδομένα SMDs ακολουθούν μια γκαουσιανή κατανομή. Ωστόσο, τα SMD συνδέονται με το δίκτυο μέσω μιας πληθώρας οργάνων όπως μετασχηματιστές, καλώδια κλπ. Η αθροιστική επίδραση αυτών των απαραίτητων στοιχείων όχι μόνο προκαλεί το συνολικό σφάλμα μέτρησης να υπερβαίνει το 1% TVE όριο, αλλά αλλάζει επίσης τη μορφή της κατανομής των σφαλμάτων. Για παράδειγμα, μόνο το σφάλμα του μετασχηματιστή ενός SMD από μόνο του για μέτρα τάσης, γωνίες τάσης, μέτρα ρεύματος και γωνίες ρεύματος μπορεί να είναι παίρνει τιμές τόσο υψηλές όσο $\pm 1,2\%$, $\pm 1^\circ$, $\pm 2,4\%$ και $\pm 2^\circ$, αντίστοιχα.

Έχει δειχθεί στο [71] ότι ένα μοντέλο γκαουσιανής μίξης (GMM) (βιβλιογραφική αναφορά για SE, σελίδα) 3 συνιστωσών αντικατοπτρίζει καλύτερα τα σφάλματα των μετρήσεων από SMDs. Ωστόσο λόγω απλούστευσης των προσομοιώσεων και επειδή το θέμα της εργασίας δεν είναι η ακριβής παραγωγή δεδομένων, στην διαδικασία δημιουργίας του dataset υποτέθηκε πως τα σφάλματα ακολουθούν κανονική κατανομή. Όπως θα δειχθεί στην συνέχεια, τα αποτελέσματα, παρά αυτή την προϋπόθεση, αποδείχτηκαν ικανοποιητικά.

Έτσι για την δημιουργία των δεδομένων χρησιμοποιήθηκε ο εξής αλγόριθμος για κάθε πιθανή τοπολογία:

1. Ορίζουμε ως βάση του φορτίου ενεργής ισχύος κάθε ζυγού την αρχική τιμή του σύμφωνα με τα δεδομένα του κάθε δικτύου.
2. Αντιστοιχίζουμε την τιμή κάθε ζυγού με ένα τυχαίο αριθμό ανάμεσα σε 0 και 1
3. Αν αυτός ο αριθμός είναι πάνω από 0.5 αλλάζω την τιμή του φορτίου ενεργής ισχύος του εκάστοτε ζυγού μ'έναν τυχαίο αριθμό κανονικής κατανομής με μέση τιμή την βάση που ορίστηκε στο πρώτο βήμα και διασπορά το 20% αυτής της τιμής.
4. Επαναλαμβάνουμε τα παραπάνω βήματα και για το φορτίο αέργου ισχύος κάθε ζυγού.

Πρακτικά με αυτή την μέθοδο προστίθενται στις μετρήσεις τυχαία γκαουσιανά σφάλματα με την πιθανότητα εμφάνισης σφάλματος σε ένα ζυγό να ακολουθεί ομοιόμορφη κατανομή. Εκτελώντας τον παραπάνω αλγόριθμο χίλιες φορές για κάθε τοπολογία κατασκευάζεται ένα dataset με επαρκή αριθμό δεδομένων.

Για την εκπαίδευση του μοντέλου εκτίμησης τοπολογίας χρησιμοποιήθηκαν μόνο οι φασιθέτες του ρεύματος που προέκυψαν από την επίλυση της ροής ισχύος για τις χίλιες περιπτώσεις που αναγνωρίζουμε σε κάθε τοπολογία, επειδή το ανοιγοκλείσιμο των διακοπών θα έχει μεγαλύτερη επίδραση στα ρεύματα που διαρρέουν το δίκτυο. Αυτό γιατί μέσω των διακοπών ελέγχεται η ροή ενέργειας στο δίκτυο κι έτσι ο αντίκτυπος στα ρεύματα είναι συχνά πιο άμεσος από ότι στις τάσεις. Το άνοιγμα ή το κλείσιμο ενός διακόπτη μεταβάλλει την τοπολογία του δικτύου συνδέοντας ή αποσυνδέοντας κλάδου κι έτσι επηρεάζονται οι διαδρομές που είναι διαθέσιμες για τη ροή ρεύματος. Το ρεύμα ακολουθεί πάντα το

μονοπάτι της μικρότερης αντίστασης, επομένως η μεταβολή της δομής του δικτύου επηρεάζει άμεσα τον τρόπο και τα σημεία που κυκλοφορούν τα ρεύματα. Μόνο τα μέτρα των ρευμάτων επαρκούν καθώς αυτά επηρεάζονται σημαντικά από τις αλλαγές της τοπολογίας και όχι τόσο οι γωνίες τους, οι οποίες εξαρτώνται κυρίως από άλλα στοιχεία όπως η επαγωγή των γραμμών. Αυτό θα δειχθεί και παρακάτω από τα αποτελέσματα.

Η εντολή runpf του Matpower ωστόσο, δεν δίνει αποτελέσματα για τα ρεύματα που διέρχονται στις γραμμές του δικτύου κι έτσι έπρεπε να υπολογιστούν με βάση τα αποτελέσματα της επίλυσης της ροής ισχύος. Αρχικά, μέσω της εντολής makeYbus κατασκευάστηκαν οι μήτρες σύνθετων αγωγιμοτήτων των ζυγών και των κλάδων από και προς (Ybus, Yf, Yt αντίστοιχα). Η συνάρτηση αυτή δέχεται ως είσοδο τη βάση σε μεγαβολτοαμπέρ ανά μονάδα που χρησιμοποιείται στην εκάστοτε περίπτωση του συστήματος και τα δεδομένα των ζυγών και των κλάδων που υπολογίστηκαν μέσω της επίλυσης της ροής φορτίου. Στην συνέχεια οι πολικές μορφές των τάσεων των ζυγών μετατράπηκαν σε ορθογώνια μορφή για κάθε ζυγό. Τέλος, υπολογίστηκαν τα ρεύματα των ζυγών πολλαπλασιάζοντας τον πίνακα σύνθετης αγωγιμότητας των ζυγών (Ybus) με το διάνυσμα των τάσεων των ζυγών και αντίστοιχα υπολογίστηκαν τα ρεύματα κλάδων από (τον ζυγό που αναγράφεται) πολλαπλασιάζοντας τον πίνακα σύνθετων αγωγιμοτήτων από (Yf) με το διάνυσμα των τάσεων των κλάδων και τα ρεύματα κλάδων προς (τον ζυγό που αναγράφεται) πολλαπλασιάζοντας τον πίνακα σύνθετων αγωγιμοτήτων από (Yf) με το διάνυσμα των τάσεων των κλάδων.

Δοκιμάστηκε και ο υπολογισμός των ρευμάτων απευθείας μέσω του γνωστού τύπου:

$$I_{ij} = \frac{1}{Z_{ij}} \cdot (V_i - V_j)$$

Ωστόσο τα ρεύματα που υπολογίζονταν με αυτή την μέθοδο ήταν λανθασμένα καθώς δεν λαμβανόταν υπόψη οι ρυθμιστές τάσης που υπάρχουν στο δίκτυο (branches 7-8, 19-20).

Για κάθε μία από τις χίλιες περιπτώσεις κάθε τοπολογίας τα ρεύματα που υπολογίστηκαν αποθηκεύτηκαν σε χωριστά αρχεία excel, τα οποία συνδυάστηκαν μεταξύ τους μέσω διάφορων μικρών προγραμμάτων γραμμένα σε python. Κατόπιν ο πίνακας αντιμετατέθηκε και από την ορθογώνια μορφή των ρευμάτων που περιείχε υπολογίστηκαν τα μέτρα των ρευμάτων καθώς κρίθηκε πως μόνο αυτά είναι επαρκή για την εκπαίδευση ενός μοντέλου εκτίμησης τοπολογίας. Με αυτόν τον τρόπο για κάθε πιθανή τοπολογία προέκυψε ένας πίνακας με διαστάσεις 1000 x 35 π.χ. Για την περίπτωση του δικτύου IEEE 33 ζυγών (32 κλάδοι που συνδέουν τους 33 ζυγούς συν τρεις κλάδοι ανάμεσα στους διακόπτες). Οι στήλες αυτού του πίνακα αντιπροσωπεύουν τους ζυγούς από τους οποίους προέρχεται το ρεύμα στον αναγραφόμενο κλάδο και γραμμές τις διαφορετικές περιπτώσεις φορτίων. Τέλος, οι πίνακες αυτοί για κάθε τοπολογία προστέθηκαν μεταξύ τους έτσι ώστε προέκυψε ένας πίνακας 8000 x 35, ο οποίος και αποτελεί τα δεδομένα ρευμάτων.

Από την άλλη πλευρά, την εκτίμηση κατάστασης του δικτύου χρησιμοποιήθηκαν και φασιθέτες τάσης καθώς χρειαζόταν μια πιο συνοπτική και ακριβής αναπαράσταση των ηλεκτρικών μεγεθών του δικτύου. Οι γωνίες φάσης των φασιθετών τάσης χρησιμεύουν ως αναφορά για τη μέτρηση των

γωνιών όλων των μεγεθών του συστήματος ενώ τα μέτρα της τάσης αποτελούν και την πιο εύκολα προσβάσιμη πληροφορία ενός συστήματος μετά την επίλυση της ροής ισχύος.

Τα μέτρα και οι φάσεις των τάσεως υπολογίζονται άμεσα από την εντολή `runpf` του `Matpower`, οπότε σε κάθε περίπτωση από τις χίλιες και για κάθε τοπολογία εξήχθησαν σε χωριστά αρχεία τα οποία στη συνέχεια συνδυάστηκαν αντίστοιχα όπως έγινε και με τα δεδομένα των ρευμάτων. Ωστόσο σε αυτή την περίπτωση προέκυψαν για κάθε τοπολογία δύο πίνακες 1000 x 33, ένας για τα μέτρα της τάσης κι ένας για τις γωνίες φάσης, των οποίων οι στήλες αντιπροσωπεύουν ένα ζυγό και οι γραμμές κάθε μία περίπτωση. Οι τελικοί πίνακες δεν προστέθηκαν μεταξύ τους καθώς κάθε τοπολογία αντιμετωπίστηκε ξεχωριστά.

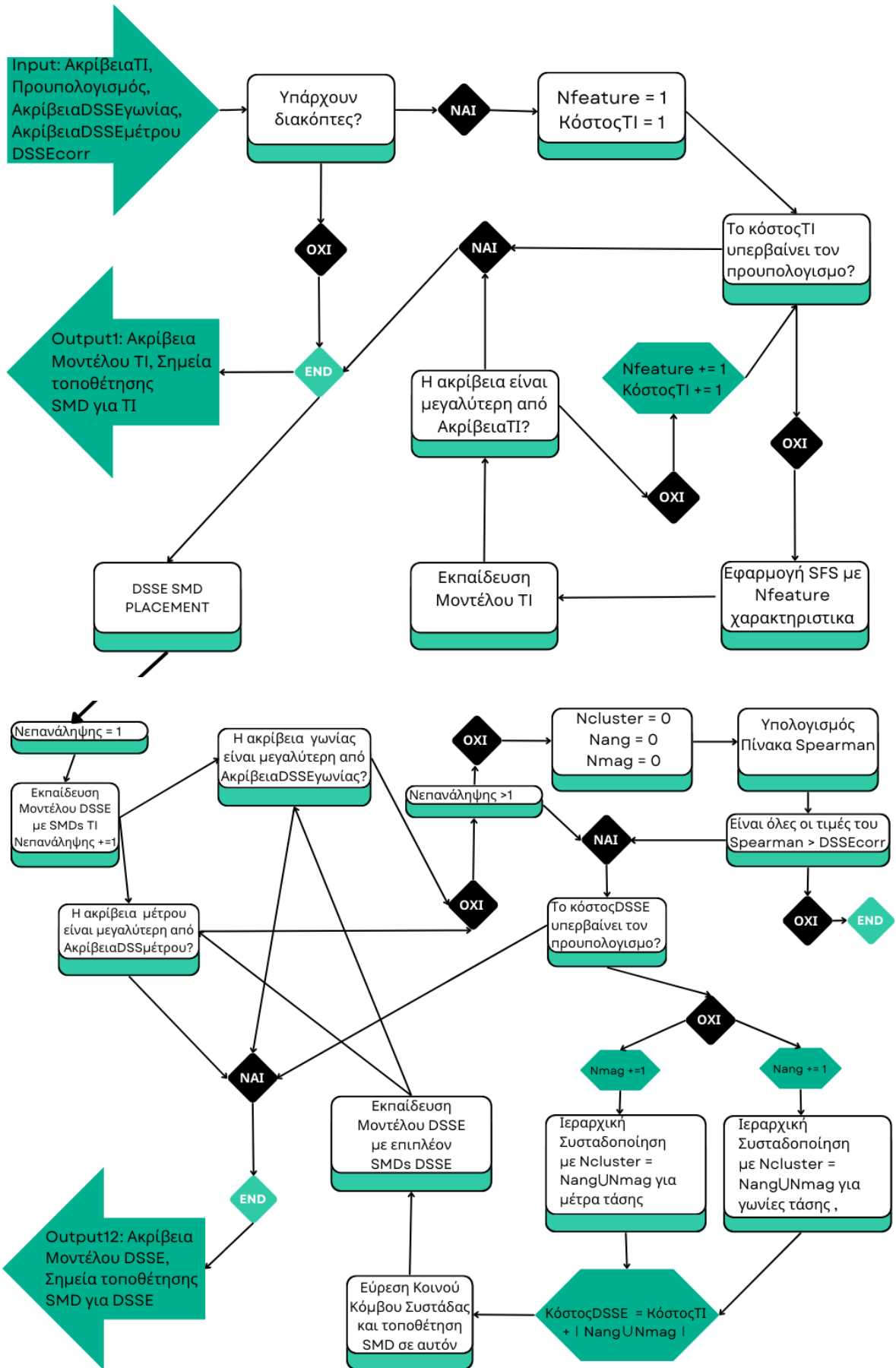
Το τελευταίο στάδιο της δημιουργίας του Dataset ήταν η προσθήκη θορύβου στις μετρήσεις. Έτσι προστέθηκε σφάλμα $\pm 0,1\%$ για τις μετρήσεις που αφορούν τα μέτρα της τάσης και του ρεύματος, και $\pm 0,5\%$ για τις μετρήσεις που αφορούν τις γωνίες τάσης και ρεύματος, με στόχο την αύξηση του ρεαλισμού των μετρήσεων έτσι ώστε τα δίκτυα που θα εκπαιδευτούν να αντεπεξέρχονται καλύτερα σε πραγματικές καταστάσεις και η εφαρμογή τους να αντικατοπτρίζει καλύτερα την πραγματικότητα.

Στην συνέχεια αναπτύχθηκε ο αλγόριθμος βέλτιστης τοποθέτησης SMDs.

Αλγόριθμος Βέλτιστης Τοποθέτησης SMDS

Η προσέγγιση για DSSE και TI που αναλύεται σε αυτή την εργασία χρησιμοποιεί δεδομένα από SMDs σε πραγματικό χρόνο για την εκτίμηση της κατάστασης και της τοπολογίας, αντίστοιχα. Όπως προαναφέρθηκε, κυρίως λόγω οικονομικών περιορισμών, δεν είναι βιώσιμη η τοποθέτηση πολλών SMDs σε ένα σύστημα διανομής. Έτσι αναπτύχθηκε ένας συστηματικός αλγόριθμος για την επιλογή θέσεων στο δίκτυο για την τοποθέτηση ενός μικρού αριθμού SMD για την απόκτηση αξιόπιστων και ακριβών αποτελεσμάτων τόσο για DSSE όσο και TI.

Δεδομένου ότι είναι ζωτικής σημασίας να γνωρίζουμε το μοντέλο μοντέλο του δικτύου πριν από την εκτέλεση της DSSE, πρέπει πρώτα να εκτελεστεί η TI έτσι ώστε εκτιμηθεί η τρέχουσα τοπολογία του δικτύου. Ως εκ τούτου, αρχικά ο αλγόριθμος βρίσκει τις τοποθεσίες που αν τοποθετηθούν μετρητές σε αυτές είναι ικανές να προσφέρουν δεδομένα με βάση τα οποία πραγματοποιείται η TI με ακρίβεια. Εάν αυτές οι τοποθεσίες δεν ικανοποιούν τα κριτήρια για την επιλογή χαρακτηριστικών για την DSSE, βρίσκουμε πρόσθετες τοποθεσίες όπου μας δίνουν δεδομένα που ικανοποιούν τα κριτήρια που έχουν τεθεί.



Για την δημιουργία του dataset υποτέθηκε πως το σύστημα είναι εξοπλισμένο με έξυπνους μετρητές σε όλους τους κόμβους και τα δεδομένα που παράγονται από αυτούς τους μετρητές έχουν αποθηκευτεί για κάποιο χρονικό διάστημα. Σε αυτό το σημείο, δεδομένου πως στην πραγματικότητα είναι αδύνατο όλα τα σημεία ενός δικτύου να έχουν πάνω τους εγκατεστημένο έναν μετρητή, το ζητούμενο ουσιαστικά είναι από ποια σημεία θα κρατήσουμε τις μετρήσεις κατά την εκπαίδευση των νευρωνικών δικτύων έτσι ώστε τα αποτελέσματα που θα ληφθούν να είναι ικανοποιητικά. Επομένως, το πρόβλημα ανήκει στην κατηγορία των προβλημάτων επιλογής χαρακτηριστικών, με στόχο να βρεθούν οι κατάλληλες θέσεις SMDs που απαιτούνται για την επίτευξη αποδεκτής απόδοσης TI και DSSE.

Όπως προαναφέρθηκε, υπάρχουν ποικίλες μέθοδοι επιλογής χαρακτηριστικών. Σε αυτό το σημείο επιλέχθηκαν δύο.

Αρχικά, για το πρόβλημα της επιλογής χαρακτηριστικών για την εκπαίδευση του μοντέλου εκτίμησης τοπολογίας επιλέχθηκε η μέθοδος Sequential Forward Selection (σειριακή προς τα εμπρός επιλογή χαρακτηριστικών). Η τεχνική αυτή περιλαμβάνει την επαναληπτική οικοδόμηση ενός προγνωστικού μοντέλου με τη σταδιακή προσθήκη ενός χαρακτηριστικού κάθε φορά από το αρχικό σύνολο χαρακτηριστικών. Μόλις προστεθεί ένα χαρακτηριστικό, η διαδικασία επαναλαμβάνεται, εξετάζοντας τα υπόλοιπα χαρακτηριστικά και επιλέγοντας εκείνο που παρέχει τη μεγαλύτερη βελτίωση στο επιλεγμένο κριτήριο. Αυτή η σταδιακή προσέγγιση συνεχίζεται έως ότου επιτευχθεί ένας προκαθορισμένος αριθμός χαρακτηριστικών ή δεν παρατηρείται περαιτέρω βελτίωση της απόδοσης του μοντέλου.

Η τεχνική αυτή ανήκει στην ευρύτερη οικογένεια των σειριακών μεθόδων επιλογής και είναι πολύ παρόμοια με την sequential backward selection, η οποία ουσιαστικά αφαιρεί ουσιαστικά χαρακτηριστικά μέχρι η επιθυμητή ακρίβεια να αρχίσει να μειώνεται σημαντικά. Ωστόσο αυτή η μέθοδος δεν επιλέχθηκε λόγω του προφανούς τεράστιου υπολογιστικού της κόστους.

Παρόλο που και η ίδια η sequential forward selection μπορεί να αποβεί υπολογιστικά εντατική, επιλέχθηκε για την επιλογή χαρακτηριστικών λόγω της αποτελεσματικότητάς της. Η SFS έχει την ικανότητά να χειρίζεται αποτελεσματικά σύνολα δεδομένων υψηλών διαστάσεων. Σε περιπτώσεις όπου ο αριθμός των χαρακτηριστικών είναι σημαντικά μεγαλύτερος από τον αριθμό των παρατηρήσεων, όπως στα δίκτυα διανομής, αξιολογεί συστηματικά τη συμβολή κάθε χαρακτηριστικού, τα πιο χρήσιμα από τα οποία προσθέτει ένα-ένα με βάση την επίδρασή τους στην ακρίβεια του μοντέλου. Έτσι, η SFS διευκολύνει τη δημιουργία πιο λιτών μοντέλων που επιτυγχάνουν ισορροπία μεταξύ προβλεπτικής ισχύος και απλότητας. Επίσης, ως εξαντλητική μέθοδος μπορεί να εξάγει με σιγουριά τα βέλτιστα χαρακτηριστικά ενός μοντέλου και είναι απλή κι εύκολη στην εφαρμογή. Η σταδιακή προσέγγισή του ευθυγραμμίζεται καλά με τον στόχο μου για συστηματική βελτίωση της απόδοσης του μοντέλου με την επαναληπτική ενσωμάτωση πληροφοριακών χαρακτηριστικών.

Ο αριθμός των χαρακτηριστικών που χρειάζονται είναι μία υπερπαράμετρος που πρέπει να ρυθμιστεί για να βρεθεί ο απαιτούμενος αριθμός SMDs για ένα προκαθορισμένο επίπεδο ακρίβειας πρόβλεψης του μοντέλου και με βάση τους περιορισμούς που θέτει ο τρέχων

προϋπολογισμός. Το πρόβλημα TI ανήκει στην κατηγορία των προβλημάτων κατηγοριοποίησης κι έτσι το μοντέλο θα πρέπει να προβλέπει σε ποια τοπολογία ανήκουν οι μετρήσεις από τα επιλεγμένα SMDs.

Από την άλλη πλευρά, το DSSE είναι ένα πρόβλημα παλινδρόμησης για το οποίο όλες οι τάσεις και τα ρεύματα του δικτύου διανομής μπορούν να είναι δυνητικά χαρακτηριστικά εισόδου. Μια ευρέως χρησιμοποιούμενη τεχνική για τον σκοπό αυτό είναι η αξιοποίηση των συντελεστών συσχέτισης. Οι συντελεστές αυτοί ποσοτικοποιούν την ισχύ και την κατεύθυνση της γραμμικής σχέσης μεταξύ δύο μεταβλητών.

Υπολογίζοντας τους συντελεστές συσχέτισης μεταξύ κάθε ανεξάρτητου χαρακτηριστικού και της μεταβλητής-στόχου, μπορεί κανείς να εκτιμήσει το βαθμό στον οποίο οι αλλαγές σε ένα χαρακτηριστικό αντιστοιχούν σε αλλαγές στη μεταβλητή-στόχο. Ένας υψηλός απόλυτος συντελεστής συσχέτισης, κοντά στο 1 ή στο -1, υποδηλώνει μια ισχυρή γραμμική σχέση. Οι θετικές συσχετίσεις υποδηλώνουν ότι καθώς αυξάνεται το χαρακτηριστικό, η μεταβλητή-στόχος τείνει να αυξάνεται, ενώ οι αρνητικές συσχετίσεις υποδηλώνουν μείωση της μεταβλητής-στόχου καθώς αυξάνεται το χαρακτηριστικό αυτό.

Η διαδικασία επιλογής χαρακτηριστικών με βάση τη συσχέτιση περιλαμβάνει συνήθως τον καθορισμό ενός κατωφλίου, με τα χαρακτηριστικά με συντελεστές συσχέτισης που υπερβαίνουν αυτό το κατώφλι να θεωρούνται σημαντικά και να επιλέγονται για συμπερίληψη στο μοντέλο παλινδρόμησης. Η προσέγγιση αυτή υποθέτει ότι τα χαρακτηριστικά με ισχυρές συσχετίσεις είναι πιθανότερο να συνεισφέρουν σημαντικές πληροφορίες στο έργο πρόβλεψης. Ωστόσο, η συσχέτιση δεν συνεπάγεται και αιτιώδη συνάφεια και ότι η μέθοδος καταγράφει κυρίως γραμμικές σχέσεις, χάνοντας ενδεχομένως πιο σύνθετες εξαρτήσεις.

Για αυτό τον λόγο, για τον υπολογισμό των συντελεστών συσχέτισης χρησιμοποιήθηκε η μέθοδος Spearman. Αυτή η μέθοδος χαρακτηρίζεται από την ανθεκτικότητα της σε μη γραμμικές σχέσεις, καθώς αξιολογεί μονοτονικές σχέσεις κι έτσι μπορεί να αξιολογεί τις σχέσεις μεταξύ δεδομένων όπου τα υποκείμενα πρότυπα είναι πιο σύνθετα και ποικίλα.

Επειδή ο συντελεστής συσχέτισης Spearman είναι ευαίσθητος στις μεταβολές των μονοτονικών τάσεων μεταξύ των μεταβλητών γίνεται ένα αποτελεσματικό εργαλείο για τον εντοπισμό χαρακτηριστικών των οποίων οι τάξεις μεταβάλλονται σταθερά με τις τάξεις της μεταβλητής-στόχου. Χαρακτηριστικά με υψηλούς συντελεστές συσχέτισης Spearman μπορούν να θεωρηθούν ότι έχουν ισχυρή, συνεπή επιρροή στη μεταβλητή-στόχο, καθιστώντας τα πολύτιμους υποψήφιους για συμπερίληψη σε μοντέλα παλινδρόμησης.

Επίσης, η συσχέτιση Spearman, είναι λιγότερο ευαίσθητη στις ακραίες τιμές σε σύγκριση με άλλες μεθόδους όπως η Pearson, καθώς βασίζεται στις τάξεις των δεδομένων και όχι στις πραγματικές τους τιμές. Οι ακραίες τιμές μπορεί να έχουν σημαντικό αντίκτυπο στους συντελεστές συσχέτισης κι έτσι η εξάρτηση της συσχέτισης Spearman από τις τάξεις βοηθά στην μετρίαση της επιρροής των ακραίων τιμών.

Τέλος, για τους παραπάνω λόγους, δηλαδή την μη παραμετρική φύση της, την μη εξάρτηση της από πραγματικές τιμές δεδομένων και την ανθεκτικότητα στις ακραίες τιμές, η συσχέτιση Spearman αποδίδει καλά

ακόμη και σε μικρά μεγέθη δείγματος, όπως είναι και το δικό μας dataset που περιέχει χίλιες περιπτώσεις για κάθε τοπολογία.

Στη συνέχεια για την επιλογή χαρακτηριστικών με βάση τον πίνακα των συντελεστών συσχέτισης που υπολογίστηκε με τη μέθοδο Spearman, χρησιμοποιήσαμε την τεχνική της ιεραρχικής συσταδοποίησης, η οποία διασπά τα χαρακτηριστικά σε συστάδες ανάλογα με την απόσταση που έχουν μεταξύ τους.

Η ιεραρχική συσταδοποίηση, δεν υποθέτει γραμμικότητα στις σχέσεις μεταξύ των χαρακτηριστικών. Μπορεί να συλλάβει μη γραμμικά μοτίβα και εξαρτήσεις μεταξύ των χαρακτηριστικών. Επίσης, η ιεραρχική συσταδοποίηση δεν απαιτεί προηγούμενη γνώση σχετικά με τον αριθμό των συστάδων στα δεδομένα. Οργανώνει αυτόματα τα χαρακτηριστικά σε μια ιεραρχική δομή, καθιστώντας την μια προσέγγιση κατάλληλη για καταστάσεις όπου η εγγενής δομή του χώρου των χαρακτηριστικών δεν είναι εκ των προτέρων γνωστή, όπως είναι γενικά τα δίκτυα διανομής. Τέλος, ένα ακόμη πλεονέκτημα της ιεραρχικής συσταδοποίησης είναι πως αντιμετωπίζει με επιτυχία την πολυσυγγραμμικότητα που μπορεί να υφίσταται μεταξύ των χαρακτηριστικών. Όταν τα χαρακτηριστικά συσχετίζονται σε μεγάλο βαθμό, το μοντέλο δυσκολεύεται να διαφοροποιήσει τις επιμέρους επιδράσεις κάθε χαρακτηριστικού στη μεταβλητή-στόχο. Αυτό μπορεί να οδηγήσει σε ασταθείς εκτιμήσεις συντελεστών, διογκωμένα τυπικά σφάλματα και δυσκολίες στην ερμηνεία της συμβολής κάθε χαρακτηριστικού στο αποτέλεσμα της παλινδρόμησης. Η ιεραρχική ομαδοποίηση μπορεί να βοηθήσει στον εντοπισμό συνόλων συσχετιζόμενων χαρακτηριστικών και στην επιλογή εκπροσώπων από κάθε ομάδα.

Ο τύπος απόστασης που χρησιμοποιήθηκε είναι η απόσταση Ward, η οποία τείνει να δημιουργεί συστάδες που είναι ομοιόμορφες και συμπαγείς και σφαιρικές, κι έτσι θεωρούμε πως ο κοινός κόμβος κάθε τέτοιας συστάδας αν επιλεγθεί θα μπορεί να περιγράφει τη συστάδα του αρκετά καλύτερα σε σχέση με άλλες μεθόδους που δεν καταλήγουν σε ομοιόμορφες και συμπαγείς ομάδες, αφού το ζητούμενο είναι η επιλογή ενός μόνο κόμβου κάθε συστάδας από τον οποίο θα λαμβάνουμε δεδομένα. Επίσης καθώς οι περισσότερες συστάδες που δημιουργούνται είναι σφαιρικές, ο κοινός κόμβος βρισκόμενος κοντά στο κέντρο της κάθε σφαίρας θα αποτελεί άλλη μια συνθήκη που θα διασφαλίζει την επιλογή του με αυτή τη μέθοδο ως καλύτερη.

Ωστόσο για λόγους σύγκρισης κι επιβεβαίωσης των αποτελεσμάτων εξετάστηκε και η ιεραρχική συσταδοποίηση με βάση την τεχνική k-means, η οποία αποτελεί μία από τις πιο δημοφιλείς εναλλακτικές.

Αναλυτική Περιγραφή Αλγορίθμου

Στο διάγραμμα που απεικονίζεται στην σελίδα 85 παρουσιάζεται το διάγραμμα block του αλγορίθμου που χρησιμοποιήθηκε για την εύρεση των βέλτιστων θέσεων τοποθέτησης SMDs στο δίκτυο διανομής που μελετάται. Ο αλγόριθμος χωρίζεται σε δύο μέρη. Το πρώτο αφορά την τοποθέτηση SMDs την εκτίμηση της τοπολογίας του δικτύου και στη συνέχεια, αν αυτά δεν επαρκούν, με το δεύτερο μέρος του αλγορίθμου προστίθενται επιπλέον έτσι ώστε η ακρίβεια της DSSSE να βρίσκεται πάνω από ένα όριο.

Τα δεδομένα εισαγωγής του αλγορίθμου για το πρώτο μέρος είναι η επιθυμητή ακρίβεια του μοντέλου, ο προυπολογισμός και οι πίνακες με τα δεδομένα εισαγωγής κι εξαγωγής του νευρωνικού δικτύου κατηγοριοποίησης, τα οποία αποτελούν έναν πίνακα με τα μέτρα από τους φασιθέτες ρεύματος και έναν ακόμη ίδιων διαστάσεων ο οποίος περιέχει τον αριθμό της τοπολογίας κωδικοποιημένο. Στους παρακάτω πίνακες απεικονίζεται η διαδικασία κωδικοποίησης της κάθε τοπολογίας, η οποία έγινε με την μέθοδο του one hot encoding.

Διακόπτης 25-> 29	Διακόπτης 12 -> 22	Διακόπτης 21 -> 8
ανοικτός	ανοικτός	ανοικτός
ανοικτός	ανοικτός	κλειστός
ανοικτός	κλειστός	ανοικτός
ανοικτός	κλειστός	κλειστός
κλειστός	ανοικτός	ανοικτός
κλειστός	ανοικτός	κλειστός
κλειστός	κλειστός	ανοικτός
κλειστός	κλειστός	κλειστός

Διακόπτης 28 -> 29	Διακόπτης 11 -> 12	Διακόπτης 7 -> 8
κλειστός	κλειστός	κλειστός
κλειστός	κλειστός	ανοικτός
κλειστός	ανοικτός	κλειστός
κλειστός	ανοικτός	ανοικτός
ανοικτός	κλειστός	κλειστός
ανοικτός	κλειστός	ανοικτός
ανοικτός	ανοικτός	κλειστός
ανοικτός	ανοικτός	ανοικτός

Όταν κλείνει κάποιος από τους διακόπτες του συστήματος, ο κλάδος που συνδέει τον ένα εκ των ζυγών που βρίσκονται στα άκρα του διακόπτη με τον προηγούμενο του ανοίγει, έτσι ώστε να αποφευχθούν οι τοπολογίες πλέγματος.

Αριθμός Τοπολογίας σε Δυαδικό	Αριθμός Τοπολογίας σε Δεκαδικό
000	1
001	2
010	3
011	4
100	5
101	6
110	7
111	8

Αριθμός Τοπολογίας σε Δυαδικό	Encoded Αριθμοί
000	00000001
001	00000010
010	00000100
011	00001000
100	00010000
101	00100000
110	01000000
111	10000000

Το one hot encoding είναι μια διαδικασία μετατροπής κατηγορικών μεταβλητών σε δυαδικό πίνακα, όπου κάθε κατηγορία αντιπροσωπεύεται από έναν μοναδικό δυαδικό κώδικα. Χρησιμοποιείται συνήθως για αλγόριθμους μηχανικής μάθησης που απαιτούν αριθμητική είσοδο, όπως τα νευρωνικά δίκτυα και ορισμένοι τύποι ταξινομητών και είναι ιδιαίτερα χρήσιμη για σενάρια όπου η κατηγορική μεταβλητή δεν έχει εγγενή τακτική σχέση, δηλαδή δεν υπάρχει ουσιαστική αριθμητική σειρά μεταξύ των κατηγοριών.

Με αυτή την τεχνική κωδικοποίησης κάθε μοναδική κατηγορία αντιστοιχίζεται σε ένα δυαδικό διάνυσμα που είναι τόσο μεγάλο όσο ο συνολικός αριθμός των διαφορετικών κατηγοριών στη μεταβλητή. Το δυαδικό διάνυσμα περιέχει όλα τα μηδενικά εκτός από ένα μόνο ένα στο δείκτη που αντιστοιχεί στην κατηγορία. Έτσι αφού στο συγκεκριμένο παράδειγμα του δικτύου IEEE33 που αναλύεται οι κατηγορίες είναι 8, όσες και οι πιθανές τοπολογίες, κάθε ετικέτα θα έχει μήκος 8 αριθμούς, με την πρώτη, για παράδειγμα, όπως φαίνεται από τους πίνακες, να είναι 00000001.

Η κωδικοποίηση είναι απαραίτητη καθώς επιτρέπει στο μοντέλο να διακρίνει μεταξύ των κατηγοριών χωρίς να εισάγει τακτικές σχέσεις που θα μπορούσαν να προκαλέσουν λανθασμένα αποτελέσματα. Επίσης, μέσω της κωδικοποίησης μπορούν να χρησιμοποιηθούν δεδομένα με μη αριθμητικές κατηγορικές μεταβλητές και να εξασφαλιστεί πως όλες οι κατηγορίες θα αντιμετωπίζονται ισότιμα ανεξαρτήτως του πλήθους των δεδομένων που αντιπροσωπεύουν. Επιπλέον, σε ένα νευρωνικό δίκτυο ταξινόμησης, το επίπεδο εξόδου χρησιμοποιεί συνήθως συναρτήσεις ενεργοποίησης που παράγουν συνεχείς τιμές μεταξύ 0 και 1, οι οποίες αντιπροσωπεύουν πιθανότητες. Η πιο συνηθισμένη συνάρτησης ενεργοποίησης για προβλήματα ταξινόμησης είναι η softmax η οποία μετασχηματίζει την ακατέργαστη έξοδο του δικτύου σε κατανομές πιθανοτήτων σε πολλαπλές κλάσεις. Για να λειτουργήσει αυτό, το δίκτυο αναμένει οι ετικέτες των κλάσεων να είναι κωδικοποιημένες με one hot encoding. Ακόμη, τα νευρωνικά δίκτυα χρησιμοποιούν μια συνάρτηση απωλειών για να μετρήσουν τη διαφορά μεταξύ της προβλεπόμενης εξόδου και των πραγματικών ετικετών κατά τη διάρκεια της εκπαίδευσης, η πιο συνηθισμένη των οποίων είναι η cross-entropy για προβλήματα ταξινόμησης. Για τον υπολογισμό αυτής της απώλειας, το δίκτυο χρειάζεται τις ετικέτες ξανά σε μορφή κωδικοποιημένη με one hot encoding.

Τα δεδομένα από τους φασιθέτες του ρεύματος κανονικοποιήθηκαν για να παίρνουν τιμές από 0 έως 1 μέσω της χρήσης της κλάσης StandardScaler από την βιβλιοθήκη scikit-learn της python. Για τον προυπολογισμό, καθώς

δεν υπάρχουν δεδομένα αναλυτικά για τα κόστη των SMD που χρησιμοποιούνται και καθώς το παράδειγμα που χρησιμοποιείται σε αυτή την εργασία είναι υποθετικό, θεωρήθηκε ένα σύστημα 'ανά μονάδα' για τα κόστη, δηλαδή κάθε SMD χρειάζεται μία μονάδα κόστους για την εγκατάσταση του. Ως προυπολογισμός τέθηκε μια μέση τιμή, δέκα μονάδες στην προκειμένη περίπτωση, οι οποίες όπως θα δειχθεί παρακάτω ήταν υπεραρκετές για το πρόβλημα.

Προχωρώντας με τον αλγόριθμο, αρχικά, ελέγχεται αν υπάρχουν διακόπτες. Έτσι, αρχικά ελέγχεται αν οι στήλες του πίνακα με τα δεδομένα του ρεύματος υπερβαίνουν τον αριθμό των ζυγών του δικτύου πλην ένα (33 ζυγοί στην περίπτωση μας). Αυτό γιατί αν υπάρχουν μόνο τόσοι κλάδοι έτσι ώστε κάθε ζυγός να συνδέεται μόνο με έναν άλλο, σημαίνει πως δεν υπάρχουν διακόπτες. Αν υπάρχουν παραπάνω στήλες, δηλαδή και παραπάνω κλάδοι, σημαίνει πως υπάρχουν διακόπτες.

Στην περίπτωση που δεν υπάρχουν, συνεπάγεται πως το δίκτυο έχει μόνο μία πιθανή τοπολογία οπότε το πρώτο μέρος τελειώνει εδώ. Στην περίπτωση που υπάρχουν όμως διακόπτες, όπως αυτή που εξετάζεται σε αυτή την εργασία, εισάγεται ένα SMD. Δηλαδή, με βάση την τεχνική sequential forward selection επιλέγεται μία στήλη από τον πίνακα με τα δεδομένα των φασιθετών του ρεύματος. Αυτό τελείται εκπαιδύοντας κάθε φορά το μοντέλο εκ νέου με είσοδο του μόνο το ένα χαρακτηριστικό προς εξέταση, δηλαδή την στήλη του πίνακα εισόδου, και συγκρίνοντας μετά την ακρίβεια της πρόβλεψης με αυτές των άλλων χαρακτηριστικών. Το χαρακτηριστικό με την καλύτερη ακρίβεια επιλέγεται και κρατείται σε μια λίστα επιλεγμένων χαρακτηριστικών.

Στην συνέχεια ο αλγόριθμος προσθέτει μία μονάδα κόστους και εξετάζει αν έχει ξεπεραστεί ο προυπολογισμός. Αν έχει ξεπεραστεί, σταματά τον αλγόριθμο και εξάγει τις τοποθεσίες των SMDs μέχρι στιγμής καθώς και την ακρίβεια του μοντέλου. Αν ο προυπολογισμός δεν έχει ξεπεραστεί, τα επιλεγμένα – μέχρι στιγμής - χαρακτηριστικά αφαιρούνται από την λίστα με τα χαρακτηριστικά που μπορούν να επιλεγθούν δυνητικά και τα οποία εξετάζει η sfs. Κατόπιν, μέσω της SFS το μοντέλο διενεργεί έναν έλεγχο στα εναπομείναντα χαρακτηριστικά και εκπαιδεύει ξανά το μοντέλο με είσοδο κάθε φορά το υπο εξέταση χαρακτηριστικό συν αυτό που επιλέχθηκε κατά την πρώτη επανάληψη του αλγορίθμου. Όταν τελειώσει αυτή η διαδικασία και συγκριθούν τα αποτελέσματα, το χαρακτηριστικό το οποίο σε συνδυασμό με το προεπιλεγμένο προσφέρει την μεγαλύτερη ακρίβεια, προστίθεται στη λίστα των επιλεγμένων. Αυτή η διαδικασία συνεχίζει επαναληπτικά μέχρι μία από τις συνθήκες διακοπής, οι οποίες είναι α) η ακρίβεια του μοντέλου να ξεπεράσει το προκαθορισμένο της όριο και β) το όριο που θέτει ο προυπολογισμός, να προκαλέσει την διακοπή του.

Ως όριο ακρίβειας του μοντέλου τέθηκε το 95%, η οποία υπολογίζεται με βάση τη μέση τιμή της ακρίβειας των προβλέψεων που κάνει το μοντέλο πάνω στο test set μετά την διαδικασία της εκπαίδευσης.

Το δεύτερο μέρος του αλγορίθμου που αναφέρεται στην τοποθέτηση SMDs για την εκτίμηση κατάστασης δέχεται ως δεδομένα εισαγωγής τον πίνακα με τα μέτρα από τους φασιθέτες ρεύματος που χρησιμοποιείται και κατά την τοποθέτηση SMDs για την εκτίμηση τοπολογίας, κρατώντας όμως μόνο τα δεδομένα που αναφέρονται σε μία τοπολογία. Η τοπολογία για την οποία εφαρμόστηκε πρώτα DSSE και για την οποία εκτελέστηκε ο

αλγόριθμος τοποθέτησης SMDs είναι όπως είναι φυσικό, η πρώτη, αυτή που είναι το δίκτυο στην αρχική του κατάσταση χωρίς διακόπτες. Επίσης, ως δεδομένα εισαγωγής προστέθηκαν και οι πίνακες με τα μέτρα και τις γωνίες φάσης της τάσης για αυτή την τοπολογία, οι τιμές της τάσης προς πρόβλεψη από το μοντέλο (η έξοδος του μοντέλου), η τιμή του ορίου συσχέτισης και η τιμή της ακρίβειας του μοντέλου παλινδρόμησης που θέλουμε να επιτύχουμε.

Αρχικά, αφού τα δεδομένα των γωνιών φάσεων και του ρεύματος κανονικοποιηθούν μέσω της κλάσης StandardScaler, το πρώτο βήμα του αλγορίθμου είναι η εκπαίδευση του μοντέλου χρησιμοποιώντας τα δεδομένα του ρεύματος αλλά μόνο από τα χαρακτηριστικά που έχουν επιλεγεί από το πρώτο μέρος του αλγορίθμου. Ουσιαστικά δηλαδή, εκπαιδεύεται το νευρωνικό δίκτυο της DSSE μόνο με δεδομένα από τις στήλες του πίνακα των ρευμάτων που αντιστοιχούν στα χαρακτηριστικά που επιλέχθηκαν με βάση το πρώτο βήμα του αλγορίθμου. Αν το μοντέλο DSSE επιτύχει τώρα την επιθυμητή ακρίβεια DSSE που τέθηκε στην αρχή, ο αλγόριθμος σταματάει χωρίς να προστεθεί κανέναν ακόμη SMD. Αν όμως δεν ξεπεράσει το όριο αυτό της ακρίβειας για DSSE και πρέπει να προστεθούν και άλλα SMDs, ο αλγόριθμος ακολουθεί της εξής διαδικασίας:

Στην αρχή, υπολογίζονται οι πίνακες των συντελεστών συσχέτισης Spearman από τους πίνακες των μέτρων της τάσης και των γωνιών φάσης της τάσης αντίστοιχα, ανάμεσα σε κάθε φασιθέτη τάσης που έχει μετρηθεί. Ουσιαστικά, κάθε κελί σε έναν πίνακα spearman αντιπροσωπεύει τον βαθμό συσχέτισης μεταξύ του χαρακτηριστικού που αντιστοιχεί στον αριθμό της γραμμής και αυτού του αριθμού της στήλης. Στη συνέχεια, όλα τα στοιχεία κάθε πίνακα ελέγχεται αν βρίσκονται πάνω από μια τιμή κατώφλι που ορίστηκε στην αρχή, και αν συμβαίνει αυτό, ο αλγόριθμος σταματάει χωρίς την τοποθέτηση SMDs. Ουσιαστικά, αυτό το βήμα προσφέρει ένα όριο – εγγύηση για το αν τα χαρακτηριστικά (οι στήλες) του πίνακα θα έχουν όλες υψηλή συσχέτιση και υψηλή μονοτονική σχέση μεταξύ τους, καθώς σε αυτή την περίπτωση δεν θα έχει νόημα η ιεραρχική συσταδοποίηση με βάση τον πίνακα των συντελεστών συσχέτισης.

Στη συνέχεια, πραγματοποιείται ιεραρχική συσταδοποίηση κάθε έναν από τους δύο πίνακες Spearman που έχουν παραχθεί, θέτοντας ως αριθμό συστάδων τον αριθμό ένα (στο Hierarchical Clustering ο αριθμός των συστάδων είναι μία παράμετρος που πρέπει να προσδιοριστεί εκ των προτέρων). Ουσιαστικά λοιπόν σε αυτό το ακόμη αρχικό βήμα δεν πραγματοποιείται ιεραρχική συσταδοποίηση καθώς θεωρείται πως όλοι οι κόμβοι (ζυγοί) ανήκουν στην ίδια συστάδα. Το ζητούμενο απλώς είναι η εύρεση του κοινού κόμβου της συστάδας – δηλαδή αυτού που βρίσκεται στο κέντρο της. Αυτό επιτυγχάνεται μέσω της σύγκρισης των τιμών της μέσης συσχέτισης που εμφανίζει κάθε χαρακτηριστικό. Το χαρακτηριστικό με την μεγαλύτερη μέση συσχέτιση επιλέγεται ως κοινός κόμβος και ως σημείο τοποθέτησης SMD. Από την διαδικασία αυτή εξαιρείται το χαρακτηριστικό της στήλης 0, καθώς αντιστοιχεί στον πρώτο ζυγό που είναι ζυγός αναφοράς και η επιλογή του δεν θα είχε νόημα. Αντί αυτού, αν τύχει και επιλεγεί, η πρώτη στήλη, δεν επιλέγεται κανένα σημείο για τοποθέτηση SMD. Δοκιμάστηκε και η επιλογή του σημείου με τη δεύτερη καλύτερη μέση συσχέτιση αλλά τα αποτελέσματα όπως θα δειχθεί αποδείχτηκαν χειρότερα.

Επίσης, αντί για την στήλη με τη μέση συσχέτιση δοκιμάστηκαν ως κριτήρια επιλογής του κοινού κόμβου και η κεντρικότητα (centrality) η οποία μετρά τον αριθμό των άμεσων συνδέσεων που έχει ένας κόμβος και το πόσο γρήγορα μπορεί να φτάσει στους άλλους κόμβους του δικτύου, και η ενδιάμεση κεντρικότητα (Betweenness Centrality) η οποία αξιολογεί τη σημασία ενός κόμβου αξιολογώντας τον αριθμό των συντομότερων διαδρομών που περνούν από αυτόν. Οι κόμβοι με υψηλή ενδιάμεση κεντρικότητα λειτουργούν ως γέφυρες σε ένα δίκτυο. Ουσιαστικά για την εύρεση της κεντρικότητας αθροίστηκαν οι τιμές των συσχετίσεων κάθε χαρακτηριστικού, ενώ για την εύρεση της ενδιάμεσης κεντρικότητας πρώτα κατασκευάζεται ένας γράφος για κάθε συστάδα κι ύστερα για κάθε κόμβο (που αντιπροσωπεύουν τα χαρακτηριστικά), υπολογίζεται ο αριθμός των συντομότερων διαδρομών που διέρχονται από αυτόν.

Αφού επιλεγεί ο κοινός κόμβος με κάποια από τις παραπάνω μεθόδους, προστίθεται μία μονάδα κόστους κι ελέγχεται αν ο προυπολογισμός έχει ξεπεραστεί για τα SMDs που έχουν τοποθετηθεί ήδη για TI και για τα νέα για DSSE. Αν όχι, εκπαιδεύεται το μοντέλο DSSE με δεδομένα εισαγωγής τις στήλες από τους φασιθέτες ρεύματος που επιλέχθηκαν κατά το μέρος του αλγορίθμου για TI και από τις στήλες από τους πίνακες μέτρων και γωνιών φάσης των τάσεων οι οποίες αντιστοιχούν στα χαρακτηριστικά που επιλέχθηκαν κατά την πρώτη επανάληψη της ιεραρχικής συσταδοποίησης.

Για την επιλογή χαρακτηριστικών μετά την διαδικασία της συσταδοποίησης θεωρείται πως αν τοποθετηθεί ένα SMD σε κάποιο σημείο, θα έχει την ικανότητα να μετράει και το μέτρο και την γωνία της τάσης. Επομένως τα χαρακτηριστικά που επιλέχθηκαν κατά την ιεραρχική συσταδοποίηση για τα μέτρα της τάσης παραδείγματος χάριν, θεωρήθηκε πως αντιστοιχούν και για τις γωνίες και το αντίστροφο.

Αν μετά την νέα εκπαίδευση του μοντέλου, η ακρίβεια του με βάση το test set είναι μεγαλύτερη από το όριο που έχει τεθεί εξαρχής, ο αλγόριθμος τερματίζει και εξάγονται η ακρίβεια αυτή και οι θέσεις των SMDs για DSSE. Αλλιώς, αν η ακρίβεια του μοντέλου είναι μικρότερη από την επιθυμητή για κάποιο εκ του μέτρο ή της γωνίας τάσης (καθώς μετράται χωριστά για καθένα από τα δύο), τότε πραγματοποιείται ξανά ιεραρχική συσταδοποίηση με δύο συστάδες για τον πίνακα Spearman της κατηγορίας αυτής. Στη συνέχεια, βρίσκονται εκ νέου οι κοινοί κόμβοι των δύο συστάδων και αφού θεωρηθεί πως τοποθετούνται εκεί SMDs κι έτσι μπορούν να κρατηθούν αυτές οι μετρήσεις, το μοντέλο εκπαιδεύεται εκ νέου με τα δεδομένα των χαρακτηριστικών που θα έχουν επιλεγεί μέχρι εδώ και ελέγχεται ξανά ο προυπολογισμός.

Ο αλγόριθμος ουσιαστικά συνεχίζει επαναληπτικά να διενεργεί ιεραρχική συσταδοποίηση για έναν από τους δύο ή και τους δύο πίνακες Spearman (για τα μέτρα και τις γωνίες φάση της τάσης) και να τοποθετεί έναν μετρητή στον κοινό κόμβο κάθε συστάδας, έως ότου το μοντέλο παλινδρόμησης δείξει την επιθυμητή ακρίβεια και για τις δύο αυτές κατηγορίες ή ξεπεραστεί το όριο του προυπολογισμού.

Ως όριο της ακρίβειας του μοντέλου παλινδρόμησης τέθηκε για τις γωνίες 0,15 μέσο απόλυτο σφάλμα και 0,30 μέσο απόλυτο ποσοστιαίο σφάλμα για τα μέτρα της τάσης. Ως κατώφλι συσχέτισης τέθηκε το

$DSS_{Ecorr} = 0.9$, το οποίο αποτελεί μια πολύ υψηλή τιμή του συντελεστή συσχέτισης μεταξύ των στηλών.

Υλοποίηση Αλγορίθμου Τοποθέτησης Μετρητών

Αρχικά, για την υλοποίηση του πρώτου μέρους του αλγορίθμου, η `sequential forward selection` υλοποιήθηκε αρχικά μέσω της βιβλιοθήκης `mlxtend` η οποία από την ενότητα `mlxtend.feature_selection` παρέχει τον αλγόριθμο αυτόν ως κλάση, η οποία ονομάζεται `SequentialFeatureSelector`. Χρησιμοποιώντας την, ο χρήστης μπορεί να διευκρινίσει:

- a) αν θέλει να πραγματοποιήσει `forward` ή `backward selection` οπότε κι επιλέχθηκε το `forward`
- b) τον αριθμό των χαρακτηριστικών που θα επιλεγθούν, ο οποίος σύμφωνα με όσα περιγράφηκαν παραπάνω αυξάνεται προοδευτικά,
- c) την μετρική βαθμολόγησης που χρησιμοποιείται για την αξιολόγηση της απόδοσης του μοντέλου κατά την επιλογή χαρακτηριστικών - σε αυτή την περίπτωση, ως μετρική βαθμολόγησης χρησιμοποιήθηκε η ακρίβεια

Η κλάση αυτή δουλεύει μέσω ενός μοντέλου ταξινόμησης από την βιβλιοθήκη `sklearn`, όπως ο `KNeighborsClassifier` ο οποίος επιλέχθηκε εδώ. Ο `KNeighborsClassifier` είναι ένας επιβλεπόμενος αλγόριθμος μηχανικής μάθησης, ο οποίος λειτουργεί με βάση την αρχή ότι παρόμοιες περιπτώσεις σε ένα χώρο χαρακτηριστικών τείνουν να ανήκουν στην ίδια κλάση. Ο `KNN` είναι ένας μη παραμετρικός και απλός αλγόριθμος που δεν κάνει υποθέσεις σχετικά με την υποκείμενη κατανομή των δεδομένων.

Στο πλαίσιο της ταξινόμησης, ο `KNeighborsClassifier` προβλέπει την ετικέτα κλάσης ενός σημείου δεδομένων λαμβάνοντας υπόψη την πλειοψηφική κλάση μεταξύ των k πλησιέστερων γειτόνων του στο χώρο χαρακτηριστικών. Η επιλογή του αριθμού των γειτόνων, που συμβολίζεται με " k ", είναι μια υπερπάρμετρος που επηρεάζει σημαντικά την απόδοση του αλγορίθμου και πρέπει να προεπιλεγθεί. Στο παράδειγμα αυτό χρησιμοποιήθηκαν $k=3$ και $k=4$ πλησιέστεροι γείτονες.

Για τη μέτρηση της εγγύτητας μεταξύ των σημείων δεδομένων χρησιμοποιούνται κοινές μετρικές αποστάσεων, όπως η ευκλείδεια απόσταση η οποία χρησιμοποιήθηκε κι εδώ.

Ωστόσο, αυτή η μεθοδολογία βασισμένη στην βιβλιοθήκη `mlxtend` δεν έχει την δυνατότητα να περιέχει ως μοντέλο ταξινόμησης ένα νευρωνικό δίκτυο, που είναι και το ζητούμενο σε αυτή την εργασία, οπότε και δεν αναπτύχθηκε περαιτέρω, παρά μόνο για σύγκριση των αποτελεσμάτων του νευρωνικού δικτύου με αυτά ενός 'συμβατικού' ταξινομητή.

Στη συνέχεια, εφαρμόστηκε μια πιο 'χειροκίνητη' προσέγγιση στην οποία η `sequential forward selection` υλοποιήθηκε με τη χρήση δύο βρόχους επανάληψης (`for loops`) μέσα σε ένα μεγαλύτερο βρόχο επανάληψης `while` (`while loop`). Σε αυτή την προσέγγιση, όσο τα χαρακτηριστικά που έχουν επιλεγθεί είναι λιγότερα από 10, ο αλγόριθμος εκπαιδεύει το μοντέλο ταξινόμησης για κάθε χαρακτηριστικό χωριστά και αποθηκεύει τις ακρίβειες που προκύπτουν σε μια νέα λίστα. Στην συνέχεια, με τον δεύτερο βρόχο επανάληψης βρίσκεται η μεγαλύτερη ακρίβεια οπότε κι επιλέγεται αυτό το

χαρακτηριστικό και ούτω καθεξής σύμφωνα με τα βήματα που περιγράφηκαν παραπάνω, έως ότου επιτευχθεί η επιθυμητή ακρίβεια ή εξαντληθεί το while loop.

Ως το δεύτερο μέρος του αλγορίθμου, αρχικά για τον υπολογισμό του πίνακα με τους συντελεστές συσχέτισης Spearman δοκιμάστηκαν τρεις διαφορετικές μέθοδοι. Η πρώτη προσέγγιση περιείχε τη χρήση της κλάσης spearmanr από την βιβλιοθήκη scipy, η οποία υπολογίζει αυτόματα τον συντελεστή spearman ανάμεσα σε δύο μεταβλητές, όπως θέλουμε να κάνουμε εδώ ανάμεσα σε όλους τους ζυγούς. Η κλάση αυτή δέχεται δύο μεταβλητές ως είσοδο κι επιστρέφει τον συντελεστή συσχέτισης τους spearman την τιμή p - δηλαδή την τιμή που αντιπροσωπεύει την πιθανότητα να παρατηρηθεί μια συσχέτιση τόσο ακραία όσο αυτή που υπολογίστηκε από τα δεδομένα του δείγματος, υποθέτοντας ότι δεν υπάρχει πραγματική συσχέτιση στα πραγματικά δεδομένα. Κρατώντας τους συντελεστές συσχέτισης κατασκευάστηκε ο πίνακας Spearman στον οποίο στη συνέχεια εφαρμόστηκε ιεραρχική συσταδοποίηση.

Ωστόσο, επειδή ο πρώτος ζυγός είναι ζυγός αναφοράς κι έτσι η πρώτη στήλη του πίνακα με τα δεδομένα των μέτρων τάσης έχει παντού την τιμή 1, ενώ αυτός με τις γωνίες φάσης της τάσης έχει πάντα την τιμή 0, η μέθοδος αυτή δεν μπορεί να εφαρμοστεί καθώς εμφανίζει σφάλμα λόγω των σταθερών ουσιαστικά χαρακτηριστικών. Δοκιμάστηκε η εισαγωγή ενός πολύ μικρού τυχαίου θορύβου, της τάξης των 10^{-6} , και η κανονικοποίηση των συντελεστών συσχέτισης έτσι ώστε να προσπελαστεί αυτό το σφάλμα, ωστόσο τα αποτελέσματα που προέκυψαν ήταν εσφαλμένα οπότε αυτή η μέθοδος απορρίφθηκε.

Η δεύτερη προσέγγιση αφορούσε την χρήση της συσχέτισης pearson ως μίας πιο εύκολης μεθόδου εκτίμησης των συντελεστών spearman. Έτσι, αρχικά υπολογίζονται οι βαθμοί των στοιχείων σε κάθε στήλη του πίνακα εισόδου. Το αποτέλεσμα είναι ένας πίνακας όπου κάθε στοιχείο του αντιπροσωπεύει τον βαθμό του αντίστοιχου στοιχείου στον πίνακα εισόδου. Στη συνέχεια, αφού μετατοπιστούν οι βαθμοί των στοιχείων κατά 1 μονάδα έτσι ώστε να μην ξεκινούν από το 0 αλλά από το 1, υπολογίζεται ο πίνακας συσχέτισης pearson με βάση τον πίνακα κατάταξης.

Όταν η σχέση μεταξύ των μεταβλητών είναι μονοτονική αλλά όχι αυστηρά γραμμική, η συσχέτιση pearson μπορεί να αποτυπώσει την μονοτονία με αρκετά καλό βαθμό ακρίβειας, υποθέτοντας πως η σχέση είναι γραμμική. Επίσης, η συσχέτιση pearson είναι υπολογιστικά λιγότερο δαπανηρή από τη συσχέτιση spearman, καθώς ο υπολογισμός των τάξεων στη δεύτερη περιλαμβάνει πρόσθετες λειτουργίες ταξινόμησης, γεγονός που την καθιστά πιο χρονοβόρα, ιδίως σε μεγάλα σύνολα δεδομένων. Επίσης, παρ'όλο που συσχέτιση pearson υποθέτει ότι οι μεταβλητές ακολουθούν μια διδιάστατη κανονική κατανομή κι ενώ η υπόθεση αυτή μπορεί να μην ισχύει για όλους τους τύπους δεδομένων, στην πράξη, για λογικά συμμετρικές και μονοτροπικές κατανομές η συνάρτηση pearson παρέχει αποτελέσματα αρκετά κοντά στα πραγματικά.

Όμως, για λόγους εγκυρότητας και για να μην βασιστεί όλη η DSSE σε μια ίσως λανθασμένη υπόθεση γραμμικότητας χρησιμοποιήθηκε και ένας αλγόριθμος παραγωγής των συντελεστών συσχέτισης spearman χωρίς προσεγγίσεις. Αρχικά, υπολογίζονται ξανά οι βαθμοί των στοιχείων κάθε στήλης, όπου ως βαθμοί λογίζονται οι δείκτες των ταξινομημένων στοιχείων

κι έτσι δημιουργείται ένας νέος πίνακας με τους βαθμούς. Στη συνέχεια, δημιουργείται ο πίνακας spearman μέσω του υπολογισμού για κάθε ζεύγος, της συνδιακύμανσης και της τυπικής απόκλισης μεταξύ τους κι έτσι και τον συντελεστή συσχέτισης spearman, ο οποίος αποδίδεται στην αντίστοιχη θέση στον νέο πίνακα spearman.

Προχωρώντας στον αλγόριθμο, τα δεδομένα των γωνιών φάσης και των ρευμάτων εισόδου κι εξόδου του μοντέλου παλινδρόμησης κλιμακώθηκαν έτσι ώστε όλα να βρίσκονται μεταξύ του 0 και του 1, χρησιμοποιώντας την κλάση StandardScaler από την βιβλιοθήκη sklearn. Τα δεδομένα των μέτρων της τάσης αφού είναι σε σύστημα ανά μονάδα και βρίσκονται ανάμεσα στο 0 και στο 1 δεν χρειάζεται να προεπεξεργαστούν με αυτή την διαδικασία.

Για την ιεραρχική συσταδοποίηση των πινάκων spearman αξιοποιήθηκε η κλάση AgglomerativeClustering ξανά από την βιβλιοθήκη sklearn, στην οποία καθορίστηκε ως μετρική απόστασης μεταξύ των συστάδων η απόσταση ward. Ο αλγόριθμος αυτός λειτουργεί δημιουργώντας μια ιεραρχία συστάδων με επαναληπτική συγχώνευση των πλησιέστερων συστάδων.

Για k-means clustering χρησιμοποιήθηκε ο αλγόριθμος Kmeans από την βιβλιοθήκη sklearn, θέτοντας ως παράμετρο τυχειότητας την default τιμή 40. Αυτή η τιμή καθορίζει τις αρχικές θέσεις των κέντρων των συστάδων, επηρεάζοντας το αποτέλεσμα της συσταδοποίησης κι εξασφαλίζει την αναπαραγωγιμότητα των αποτελεσμάτων.

Η μέση τιμή συσχέτισης κάθε στήλης για να βρεθεί ο κοινός κόμβος κάθε συστάδας υπολογίζεται μέσω της συνάρτησης mean από την βιβλιοθήκη numpy, η κεντρικότητα μέσω της εντολής np.sum που αθροίζει τα στοιχεία μιας στήλης πάλι από το numpy. Το κριτήριο της ενδιάμεσης κεντρικότητας υλοποιείται μέσω της βιβλιοθήκης networkx. Ένας γράφος G δημιουργείται με τη χρήση της networkx (nx.Graph()), με βάση ένα υποσύνολο του πίνακα spearman που περιέχει μόνο τις γραμμές και τις στήλες που αντιστοιχούν στα δεδομένα εντός της τρέχουσας συστάδας. Στη συνέχεια υπολογίζεται η ενδιάμεση κεντρικότητα για κάθε κόμβο του γραφήματος και ο κόμβος με την υψηλότερη προσδιορίζεται χρησιμοποιώντας την np.argmax().

Νευρωνικά Δίκτυα

Η κεντρική ιδέα της εργασίας αυτής είναι πως μπορεί να εκτελεστεί DSSE χρησιμοποιώντας για τις αναγκαίες προβλέψεις νευρωνικά δίκτυα. Έτσι, ο τρόπος κατασκευής των νευρωνικών αυτών μοντέλων πρόβλεψης διαδραματίζει πολύ σημαντικό ρόλο στα αποτελέσματα της εκτίμησης κατάστασης.

Αρχικά, για το μοντέλο ταξινόμησης που χρησιμοποιήθηκε για TI, ως βαθύ νευρωνικό δίκτυο, υποτέθηκε πως θα περιέχει τουλάχιστον ένα κρυφό στρώμα. Οι νευρώνες του στρώματος εισόδου δεν θα είναι σταθεροί, καθώς κατά την διαδικασία επιλογής χαρακτηριστικών μέσω sequential forward selection θα προστίθενται συνεχώς νέα χαρακτηριστικά. Οι νευρώνες του στρώματος εξόδου όμως θα είναι 8, καθώς ουσιαστικά παράγεται ένα διάλυμα του αριθμού της τοπολογίας το οποίο είναι one hot encoded κι

έτσι οι διαστάσεις του αντιστοιχούν στον αριθμό των πιθανών τοπολογιών. Ο αριθμός των ενδιάμεσων στρωμάτων και ο αριθμός των νευρώνων καθενός αποτελούν υπερπαραμέτρους οι οποίες απαιτούν την ρύθμιση του νευρωνικού δικτύου για την εύρεση τους, η οποία θα αναλυθεί σε επόμενο κεφάλαιο.

Στο στρώμα εισόδου και στα κρυφά στρώματα χρησιμοποιήθηκε η συνάρτηση ενεργοποίησης ReLU, η οποία αποτελεί ίσως την πιο δημοφιλή επιλογή για αυτά τα στρώματα, για την εισαγωγή μη γραμμικότητας και τη σύλληψη σύνθετων μοτίβων στα δεδομένα εισόδου. Ένα νευρωνικό δίκτυο χωρίς μη γραμμικές συναρτήσεις ενεργοποίησης θα ήταν ισοδύναμο με ένα γραμμικό μοντέλο, ανίκανο να συλλάβει περίπλοκα χαρακτηριστικά. Επίσης η ReLU είναι υπολογιστικά αποδοτική καθώς απλώς επιστρέφει την είσοδο για θετικές τιμές και μηδέν για αρνητικές τιμές. Αυτή η απλότητα καθιστά την εκπαίδευση βαθιών νευρωνικών δικτύων με τη ReLU ταχύτερη σε σύγκριση με ορισμένες άλλες συναρτήσεις ενεργοποίησης. Παράλληλα, η ReLU συμβάλλει στον μετριασμό του προβλήματος της εξαφανιζόμενης κλίσης, το οποίο είναι σύνηθες στα βαθιά δίκτυα. Οι παραδοσιακές συναρτήσεις ενεργοποίησης, όπως η σιγμοειδής (sigmoid) και η υπερβολική εφαπτομένη (hyperbolic tangent), μπορούν να οδηγήσουν σε κλίσεις που γίνονται εξαιρετικά μικρές κατά την διάδοση της πληροφορίας προς τα πίσω, εμποδίζοντας την εκπαίδευση βαθιών δικτύων, η ReLU όμως, καθώς δεν είναι κορεσμένη για θετικές τιμές, επιτρέπει στην κλίση να παίρνει πιο ομαλές τιμές, επιτρέποντας την αποτελεσματικότερη εκπαίδευση βαθιών αρχιτεκτονικών. Τέλος, η ReLU εισάγει ένα βαθμό μείωσης της πληροφορίας καθώς ενεργοποιείται μόνο για θετικές τιμές εισόδου, το οποίο μπορεί να βοηθήσει το δίκτυο να εστιάσει στα πιο σημαντικά χαρακτηριστικά και να μειωθεί η επίδραση του φαινομένου του overfitting.

Για το στρώμα εξόδου, χρησιμοποιήθηκε η softmax συνάρτηση ενεργοποίησης, η οποία αποτελεί μια γενίκευση ουσιαστικά της σιγμοειδούς που χρησιμοποιείται για δυαδικά προβλήματα ταξινόμησης. Η συνάρτηση ενεργοποίησης αυτή χρησιμοποιείται ευρέως στο στρώμα εξόδου σε προβλήματα ταξινόμησης πολλαπλών κλάσεων λόγω της ικανότητάς της να μετατρέπει την έξοδο ενός νευρωνικού δικτύου σε κατανομή πιθανότητας για πολλαπλές κλάσεις. Με την εφαρμογή της softmax, οι ακατέργαστες τιμές εξόδου του μοντέλου μετατρέπονται σε πιθανότητες, με κάθε κλάση να λαμβάνει μια βαθμολογία πιθανότητας μεταξύ 0 και 1 κι έτσι εξασφαλίζεται ότι το άθροισμα των πιθανοτήτων που προκύπτουν θα είναι 1, καθιστώντας την έξοδο κατάλληλη για την αναπαράσταση μιας έγκυρης κατανομής πιθανοτήτων. Οι πιθανότητες που προκύπτουν καθοδηγούν στη συνέχεια την απόφαση ταξινόμησης, με την υψηλότερη πιθανότητα να υποδεικνύει την προβλεπόμενη κλάση.

Κατά την μεταγλώττιση του μοντέλου, ως συνάρτηση απώλειας τέθηκε η categorical crossentropy καθώς διακρίνεται για την καταλληλότητα της για προβλήματα ταξινόμησης πολλαπλών κλάσεων, καθώς μετράει τη διαφορά μεταξύ της προβλεπόμενης κατανομής πιθανότητας και της πραγματικής κατανομής. Η έξοδος του νευρωνικού δικτύου συχνά ερμηνεύεται ως πιθανότητες κλάσεων και η κατηγορική διασταυρούμενη εντροπία ποσοτικοποιεί τη διαφορά μεταξύ αυτών των προβλεπόμενων πιθανοτήτων και της αλήθειας. Η categorical crossentropy υπολογίζεται με βάση τον τύπο:

$$\text{Categorical Crossentropy} = - \sum_{i=1}^C y_i \cdot \log(\hat{y}_i)$$

Ως βελτιστοποιητής τέθηκε ο adam, ο οποίος λόγω της φήμης του για την αποδοτικότητα και την αποτελεσματικότητά του, χρησιμοποιείται συνήθως ως προεπιλεγμένος βελτιστοποιητής. Το κύριο πλεονέκτημα του είναι πως διατηρεί έναν ξεχωριστό προσαρμοστικό ρυθμό μάθησης για κάθε παράμετρο, γεγονός που συχνά μειώνει την ανάγκη για εκτεταμένη χειροκίνητη ρύθμιση των υπερπαραμέτρων, καθιστώντας το μια βολική επιλογή για πειραματισμό. Τέλος, ως μετρικής για την διαδικασία της εκπαίδευσης ορίστηκε η ακρίβεια του μοντέλου.

Για την αξιολόγηση της απόδοσης του μοντέλου κατασκευάστηκε μια συνάρτηση η οποία δέχεται ένα μοντέλο και δεδομένα εκπαίδευσης. Μετά την εκπαίδευση, το μοντέλο χρησιμοποιείται για να κάνει προβλέψεις στα δεδομένα δοκιμής (test set) το οποίο έχει διαχωριστεί από τα δεδομένα εκπαίδευσης μέσω του `train_test_split` από την βιβλιοθήκη `sklearn` θέτοντας τον διαχωρισμό 80 – 20 %. Στη συνέχεια υπολογίζεται η κατηγορική ακρίβεια μεταξύ των πραγματικών ετικετών και των προβλεπόμενων. Κατόπιν, οι μεμονωμένες κατηγορικές ακρίβειες χρησιμοποιούνται για τον υπολογισμό της μέσης ακρίβειας σε όλες τις κλάσεις, η οποία αποτελεί και την έξοδο της συνάρτησης και την τελική ακρίβεια του μοντέλου την οποία ο αλγόριθμος επιλογής θέσεων SMDs έχει σκοπό να αυξήσει.

Η διαδικασία κατασκευής ενός δικτύου παλινδρόμησης για την DSSE αποδείχτηκε πιο περίπλοκη. Αρχικά, το δίκτυο δέχεται τρεις διαφορετικές εισόδους, τα μέτρα του ρεύματος, τα μέτρα της τάσης και οι γωνίες φάσης της τάσης, και σε δύο διαφορετικά στάδια – τα δεδομένα του ρεύματος προπορεύονται αυτά της τάσης. Επίσης, πρέπει να υπάρχουν δύο διαφορετικές έξοδοι, τα μέτρα και οι γωνίες φάσης ξανά που εκτιμώνται κατά την DSSE.

Το παραπάνω μοντέλο παλινδρόμησης θα μπορούσε να θεωρηθεί και ως ένας συνδυασμός δύο μοντέλων, το ένα αφιερωμένο αποκλειστικά στην πρόβλεψη του μέτρου της τάσης και το δεύτερο στην πρόβλεψη των γωνιών φάσης. Ωστόσο, αυτή η προσέγγιση αφενός θα ήταν πιο περίπλοκη καθώς θα χρειαζόταν να βρεθούν δύο νευρωνικά δίκτυα και να ρυθμιστούν κατάλληλα στην συνέχεια, και αφετέρου δεν θα λάμβανε υπόψη την αμοιβαία σύζευξη που υπάρχει μεταξύ του μέτρου και της γωνίας φάσης ενός μεγέθους. Έτσι, αναπτύχθηκε ένα μοντέλο, το οποίο σε κάποιο σημείο θα χωριζόταν σε δύο κλάδους, ο καθένας εκ των οποίων θα παράγει και την κατάλληλη έξοδο.

Σε αυτό το σημείο επίσης δοκιμάστηκαν δύο διαφορετικές προσεγγίσεις. Αρχικά κατασκευάστηκε ένα μοντέλο στο οποίο το στρώμα εισόδου δεχόταν τα δεδομένα του ρεύματος και κατόπιν τα δεδομένα της τάσης προστίθενται διαδοχικά σε επόμενα κρυφά κοινά στρώματα, δηλαδή προτού το νευρωνικό δίκτυο διασπαστεί σε δύο κλαδιά. Ωστόσο, στο μοντέλο αυτό παρουσιαζόταν το πρόβλημα της μηδενικής εισόδου. Στο πρώτο στάδιο του αλγορίθμου τοποθέτησης SMDs για DSSE το μοντέλο εκπαιδευόταν μόνο με τα δεδομένα του ρεύματος ως είσοδο, με αποτέλεσμα οι άλλες δύο εισοδοί να πρέπει να θεωρούνται μηδενικές. Στο μοντέλο αυτό, στο οποίο ουσιαστικά υπήρχαν τρία στρώματα εισόδου, στα δύο τελευταία από τα οποία που αντιστοιχούν στα δεδομένα της τάσης πρέπει να θεωρηθούν πως δέχονται μια είσοδο ίδιων διαστάσεων με τα επόμενα

στάδια, η οποία όμως ουσιαστικά θα είναι ένα μηδενικό διάνυσμα. Αυτό δυνητικά θα μπορούσε να προκαλέσει μεγάλη πτώση στην ακρίβεια του μοντέλου κι έτσι αυτή η ιδέα απορρίφθηκε.

Στη συνέχεια κατασκευάστηκε ένα νευρωνικό δίκτυο μ'ένα επίπεδο εισόδου, κρυφά στρώματα και δύο στρώματα εξόδου, ένα σε κάθε κλάδο του δικτύου. Τα στρώματα εξόδου αποτελούνται και τα δύο από 33 νευρώνες στο παράδειγμα του δικτύου IEEE33 που εξετάστηκε, δηλαδή ένας νευρώνας θα αντιστοιχεί σε κάθε ζυγό του δικτύου. Το στρώμα εισόδου θα έχει στην αρχή τόσους νευρώνες όσα τα SMDs που επιλέχθηκαν για TI και στη συνέχεια σε κάθε βήμα του αλγορίθμου τοποθέτησης SMDs για DSSE θα προστίθενται δύο ή τέσσερις νευρώνες, καθώς κάθε νέο SMD θα μετράει και ρεύμα και τάση και σε κάθε βήμα θα προστίθενται τουλάχιστον άλλο ένα SMD.

Όπως και στο δίκτυο ταξινόμησης χρησιμοποιήθηκε ως συνάρτηση ενεργοποίησης για τα ενδιάμεσα στρώματα εισόδου η ReLU, ενώ το στρώμα εισόδου περνάει απευθείας τα χαρακτηριστικά της εισόδου στο επόμενο επίπεδο χωρίς κάποιο μετασχηματισμό κι έτσι δεν έχει κάποια συνάρτηση ενεργοποίησης. Στο στρώμα εξόδου κάθε κλάδου, τέθηκε ως συνάρτηση ενεργοποίησης η linear, η οποία αποτελεί ίσως την πιο συνηθισμένη για το τελευταίο επίπεδο νευρωνικών μοντέλων παλινδρόμησης. Ο πρωταρχικός στόχος ενός μοντέλου παλινδρόμησης είναι να προβλέψει μια συνεχή αριθμητική τιμή, η οποία μπορεί να καλύπτει ένα ευρύ φάσμα πιθανών αριθμών. Η γραμμική συνάρτηση ενεργοποίησης επιτρέπει στο μοντέλο να εξάγει οποιονδήποτε πραγματικό αριθμό χωρίς να επιβάλλει ένα συγκεκριμένο εύρος ή οποιονδήποτε άλλο περιορισμό στον χώρο εξόδου. Επίσης, η χρήση της γραμμικής συνάρτησης ενεργοποίησης στο στρώμα εξόδου καθιστά τις προβλέψεις του μοντέλου άμεσα ερμηνεύσιμες καθώς δεν υπάρχουν μη γραμμικοί μετασχηματισμοί.

Επιπλέον, ξανά όπως στο μοντέλο ταξινόμησης ως βελτιστοποιητής θεωρήθηκε ο adam, με την διαφορά πως ρυθμός εκμάθησης (learning rate - το μέγεθος βήματος με το οποίο ενημερώνονται οι παράμετροι του μοντέλου κατά τη διάρκεια κάθε επανάληψης), θεωρήθηκε υπερπαράμετρος και η τιμή του εξετάστηκε στο επόμενο στάδιο, την ρύθμιση των υπερπαραμέτρων. Αντιθέτως, κατά την μεταγλώττιση του μοντέλου, ως συνάρτηση απώλειας τέθηκε η Mean Squared Error, η οποία δίνεται από τον τύπο:

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

Αυτή η συνάρτηση σφάλματος χρησιμοποιείται πολύ συχνά ως συνάρτηση απώλειας για μοντέλα παλινδρόμησης λόγω της συνεχής εξόδου του μοντέλου. Ο τετραγωνισμός των σφαλμάτων μεγεθύνει τον αντίκτυπο των μεγαλύτερων εξ αυτών, καθιστώντας το μοντέλο πιο ευαίσθητο σε σημαντικές αποκλίσεις από τις πραγματικές τιμές κι έτσι οδηγώντας σε πιο εύκολη βελτίωση της αποδοτικότητας. Επιπλέον, η Mean Squared Error οδηγεί σε κυρτά προβλήματα βελτιστοποίησης, γεγονός που εξασφαλίζει ότι η διαδικασία βελτιστοποίησης συγκλίνει σε ένα παγκόσμιο ελάχιστο αντί να κολλάει σε τοπικά ελάχιστα.

Τέλος, ως μετρική για την διαδικασία της εκπαίδευσης ορίστηκε, όπως είναι συνηθισμένο στα μοντέλα παλινδρόμησης, το μέσο απόλυτο σφάλμα του μοντέλου.

Ο αριθμός των κρυμμένων στρωμάτων κι εδώ θεωρήθηκε υπερπαράμετρος, όπως και ο αριθμός των νευρώνων κάθε στρώματος.

Επίσης, το δίκτυο θεωρήθηκε συμμετρικό, δηλαδή οι δύο κλάδοι θεωρήθηκε πως θα έχουν τον ίδιο αριθμό στρωμάτων και νευρώνων σε κάθε στρώμα έτσι ώστε να απλοποιηθεί η διαδικασία της ρύθμισης των υπερπαραμέτρων.

Για την αξιολόγηση της απόδοσης του μοντέλου κατασκευάστηκαν δύο συναρτήσεις. Η πρώτη μέσω των λειτουργιών `np.mean` και `np.abs` της βιβλιοθήκης `numpy` υπολογίζει την μέση απόλυτη διαφορά (Mean Absolute Error- Mae) μεταξύ των προβλεπόμενων και των πραγματικών τιμών των γωνιών φάσης. Η δεύτερη υπολογίζει το μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error – Mape) ανάμεσα των προβλεπόμενων τιμών των μέτρων των τάσεων, χρησιμοποιώντας ξανά τις λειτουργίες `np.mean` και `np.abs` της βιβλιοθήκης `numpy`, μέσω του τύπου:

$$\frac{|y_{\text{true}} - y_{\text{pred}}|}{y_{\text{true}}}$$

Για την κατασκευή των νευρωνικών δικτύων και στις δύο περιπτώσεις χρησιμοποιήθηκε η βιβλιοθήκη `keras`. Η βιβλιοθήκη `keras` αποτελεί ένα υψηλού επιπέδου API (Application Programming Interface) νευρωνικών δικτύων που απλοποιεί τη διαδικασία δημιουργίας, εκπαίδευσης και ανάπτυξης μοντέλων βαθιάς μάθησης, καθώς αφαιρεί τις λεπτομέρειες υλοποίησης χαμηλού επιπέδου, επιτρέποντας στους χρήστες να επικεντρωθούν στο σχεδιασμό μοντέλων και στον πειραματισμό. Αναπτύχθηκε με επίκεντρο τη δυνατότητα γρήγορου πειραματισμού με βαθιά νευρωνικά δίκτυα, καθιστώντας το προσιτό και εύκολο στη χρήση

Ωστόσο, ενώ το `keras` παρέχει το frontend API, βασίζεται σε ένα backend πλαίσιο για την αποτελεσματική εκτέλεση των υπολογισμών. Το `TensorFlow`, που αναπτύχθηκε από την `Google`, είναι ένα από τα πιο δημοφιλή πλαίσια βαθιάς μάθησης και χρησιμεύει ως το προεπιλεγμένο backend για το `Keras`. Όταν χρησιμοποιείτε το `Keras` μαζί με το `TensorFlow`, το δεύτερο χειρίζεται την κατασκευή και τη βελτιστοποίηση, αξιοποιώντας την υπολογιστική του απόδοση και την επεκτασιμότητά του.

Το `Keras` προσφέρει ένα ευρύ φάσμα ενσωματωμένων στρωμάτων, συναρτήσεων ενεργοποίησης, βελτιστοποιητών, συναρτήσεων απώλειας και βοηθητικών προγραμμάτων, καθιστώντας το ευέλικτο και προσαρμόσιμο σε διαφορετικές εργασίες και αρχιτεκτονικές. Με τη χρήση του γίνεται πιο εύκολη η κατασκευή σύνθετων νευρωνικών δικτύων στοιβάζοντας στρώματα, καθορίζοντας συναρτήσεις ενεργοποίησης, διαμορφώνοντας αλγόριθμους βελτιστοποίησης και ορίζοντας συναρτήσεις απωλειών - όλα αυτά με λίγες μόνο γραμμές κώδικα.

Παρακάτω απεικονίζονται ποιες κλάσεις του `keras` αξιοποιήθηκαν και από ποιο module του:

Submodule Keras	Κλάση
models	Sequential
layers	Dense
layers	Dropout
metrics	accuracy_score
metrics	categorical_accuracy
losses	MeanSquaredError
regularizers	l2
layers	Input
models	Model
layers	add
layers	BatchNormalization
layers	concatenate
optimizers	Adam
callbacks	EarlyStopping

Ρύθμιση Νευρωνικών Δικτύων

Όπως προαναφέρθηκε στο θεωρητικό μέρος της εργασίας, η ρύθμιση των υπερπαραμέτρων είναι η διαδικασία με την οποία μπορεί να βρεθεί το βέλτιστο σύνολο υπερπαραμέτρων για ένα δεδομένο μοντέλο. Αποτελεί ίσως το σημαντικότερο στάδιο κατά την δημιουργία ενός νευρωνικού δικτύου, κατά την διάρκεια του οποίου η απόδοση του δικτύου παίρνει την υψηλότερη τιμή της. Σχεδόν πάντα, οι προεπιλεγμένες υπερπαραμέτροι ενδέχεται να μην οδηγούν στο πιο αποτελεσματικό μοντέλο για ένα συγκεκριμένο σύνολο δεδομένων ή πρόβλημα κι έτσι είναι απαραίτητη η αναζήτηση στον χώρο των υπερπαραμέτρων για τον εντοπισμό του συνδυασμού που οδηγεί στην καλύτερη απόδοση του μοντέλου.

Οι υπερπαραμέτροι είναι εξωτερικές ρυθμίσεις διαμόρφωσης που δεν μαθαίνονται από τα δεδομένα αλλά είναι απαραίτητες για τη διαδικασία εκπαίδευσης. Παραδείγματα υπερπαραμέτρων περιλαμβάνουν τον ρυθμό μάθησης, τον αριθμό των κρυφών στρωμάτων, των αριθμών το νευρώνων ενός επιπέδου κλπ. Δυνητικά, μπορεί να επιλεγθεί ένας πολύ μεγάλος αριθμός υπερπαραμέτρων ο οποίος θα αφορά σχεδόν όλες τις πτυχές του δικτύου, από τις συναρτήσεις ενεργοποίησης κάθε στρώματος μέχρι τις κατάλληλες ρυθμίσεις του βελτιστοποιητή. Ωστόσο κάτι τέτοιο δεν είναι πρακτικό καθώς αύξηση του χώρου των υπερπαραμέτρων συνεπάγεται και πολλαπλάσια αύξηση του χρόνου και της υπολογιστικής δύναμης που απαιτούνται για την εύρεση του βέλτιστου συνδυασμού. Έτσι, ο χώρος των υπερπαραμέτρων προς εξέταση είναι προτιμότερο να περιέχει αυτές που η διαφοροποίηση τους επιδρά πιο καθοριστικά στην απόδοση του δικτύου.

Υπάρχουν διάφορες τεχνικές για την ρύθμιση των υπερπαραμέτρων, όπως η αναζήτηση πλέγματος (grid search), η τυχαία αναζήτηση (random search) και πιο προηγμένες μέθοδοι όπως η βελτιστοποίηση κατά Bayes. Σε αυτή την διπλωματική αξιοποιήθηκε η αναζήτηση πλέγματος καθώς είναι

εύκολη στην υλοποίηση και πιο αξιόπιστη στην εύρεση του καλύτερου συνδυασμού υπερπαραμέτρων. Δεδομένου ότι διερευνά συστηματικά όλους τους πιθανούς συνδυασμούς, υπάρχει βεβαιότητα ότι θα βρεθεί η βέλτιστη λύση εντός του καθορισμένου πλέγματος.

Το ζητούμενο σε αυτό το κεφάλαιο της εργασίας ήταν η εύρεση των βέλτιστων υπερπαραμέτρων για τα δύο νευρωνικά δίκτυα που αναπτύχθηκαν για TI και DSSE αντίστοιχα. Ωστόσο το πρόβλημα είναι πιο πολύπλοκο καθώς τα νευρωνικά δίκτυα κατασκευάζονταν παράλληλα με τον αλγόριθμο τοποθέτησης SMDs. Έτσι, ουσιαστικά το πρόβλημα που αντιμετωπίστηκε ήταν μία «δισδιάστατη» ρύθμιση παραμέτρων καθώς έπρεπε να βρεθούν οι υπερπαραμέτροι που οδηγούσαν παράλληλα αφενός στην μεγαλύτερη ακρίβεια του μοντέλου και αφετέρου στον μικρότερο αριθμό θέσεων τοποθέτησης SMDs.

Για το μοντέλο ταξινόμησης TI ως υπερπαραμέτροι λογίστηκαν: ο αριθμός των κρυφών στρωμάτων, ο αριθμός των νευρώνων των κρυφών στρωμάτων, ο αριθμός των εποχών (epochs) κατά την εκπαίδευση, το μέγεθος δέσμης (batch size) επίσης κατά την εκπαίδευση, ο αριθμός στρωμάτων dropout και τέλος το ποσοστό dropout κάθε τέτοιου στρώματος. Επιπλέον εξετάστηκε αν οι νευρώνες κάθε διαδοχικού κρυφού στρώματος θα αυξάνονται όσο το αντίστοιχο στρώμα θα πλησιάζει προς την έξοδο ή το αντίστροφο.

Αρχικά, δοκιμάζοντας ‘χειροκίνητα’ μια πληθώρα συνδυασμών βρέθηκαν μερικά δίκτυα τα οποία διέφεραν ελάχιστα στην ακρίβεια και στον αριθμό SMDs που χρειαζόνταν, οπότε και βέλτιστο θεωρήθηκε το απλούστερο από αυτά έτσι ώστε ο υπολογιστικός χρόνος να κρατηθεί όσο χαμηλότερα γίνεται (αφού κάθε αύξηση στον αριθμό των στρωμάτων ή των νευρώνων αυξάνει τον χρόνο της διαδικασίας της εκπαίδευσης). Εδώ, δεν κρίθηκε σκόπιμη η εκτέλεση μιας αναζήτησης πλέγματος καθώς είχε ελεγχθεί ήδη ‘χειροκίνητα’ το μεγαλύτερο μέρος των συνδυασμών των υπερπαραμέτρων. Η αύξηση του χώρου των υπερπαραμέτρων και κατόπιν η εκτέλεση μιας αναζήτησης πλέγματος απορρίφθηκε καθώς θα αυξανόταν σημαντικά ο χρόνος της εκπαίδευσης με τις νέες τιμές των υπερπαραμέτρων που θα εισάγονταν για δοκιμή.

Για απλοποίηση της διαδικασίας από ένα σημείο και μετά θεωρήθηκε πως κάθε επόμενο κρυφό στρώμα θα περιέχει τους μισούς νευρώνες από το προηγούμενο του. Στη συνέχεια, δοκιμάστηκε να αντιστραφεί αυτός ο κανόνας κι έτσι κάθε κρυφό στρώμα να περιέχει τους διπλάσιους νευρώνες από το προηγούμενο του, αλλά τα αποτελέσματα χειροτέρεψαν. Κατά την διάρκεια αυτής της διαδικασίας παρατηρήθηκε πως τα καλύτερα μοντέλα χρειαζόντουσαν 4 SMDs για να ξεπεράσουν το επιθυμητό όριο ακρίβειας, ωστόσο κανένας συνδυασμός υπερπαραμέτρων δεν μπορούσε να ελαττώσει τον αριθμό αυτό.

Για το μοντέλο παλινδρόμησης που χρησιμοποιήθηκε για DSSE, ως υπερπαραμέτροι θεωρήθηκαν: ο αριθμός των νευρώνων κάθε κοινού στρώματος και κάθε στρώματος κλάδου, ο αριθμός των κοινών κρυφών στρωμάτων, ο αριθμός των στρωμάτων καθενός από τους δύο κλάδους, ο αριθμός των εποχών, το μέγεθος της δέσμης και ο ρυθμός εκμάθησης του βελτιστοποιητή κατά την εκπαίδευση. Επίσης κι εδώ εξετάστηκε αν ο αριθμός των νευρώνων θα πρέπει να αυξάνεται ή να μειώνεται όσο το στρώμα στο οποίο αντιστοιχούν πλησιάζει την έξοδο του μοντέλου.

Αρχικά εκτελέστηκαν με το χέρι μια πληθώρα συνδυασμών κυρίως του αριθμού των κρυφών στρωμάτων και των στρωμάτων κάθε κλαδιού με στόχο την μείωση του χώρου των πιθανών συνδυασμών των υπερπαραμέτρων. Στην συνέχεια, εξετάστηκε η αντιστροφή του αριθμού των νευρώνων για διάφορα πλήθη στρωμάτων και αποδείχτηκε πως η απόδοση του μοντέλου ήταν καλύτερη όταν ο αριθμός των νευρώνων αυξανόταν όσο τα στρώματα πλησίαζαν την έξοδο. Επίσης, κατά την διάρκεια αυτών των ‘χειροκίνητων’ δοκιμών, ο αριθμός των SMDs για DSSE παρατηρήθηκε πως για μερικούς συνδυασμούς χρειάζεται να είναι μόνο ίσος με 2.

Τέλος, εκτελέστηκε μια σειρά από grid search θεωρώντας πως υπάρχουν μόνο δύο SMDs διαθέσιμα για DSSE (μαζί με αυτά που είχαν ήδη τοποθετηθεί για TI). Ο λόγος που δεν εκτελέστηκε ένα μεγάλο grid search εξαρχής ήταν για να μην διαρκέσει η διαδικασία πάρα πολύ, καθώς κατά τη διάρκεια της συνεχώς δημιουργούνται, εκπαιδεύονται και αποθηκεύονται νέα μοντέλα. Αυτό έχει ως αποτέλεσμα ο χρόνος να αυξάνεται εκθετικά για κάθε νέα υπερπαραμέτρο που προστίθεται.

Transfer Learning

Για το transfer learning δοκιμάστηκαν δύο τεχνικές – ουσιαστικά οι δύο μέθοδοι της μεταφοράς παραμέτρων (βλέπε θεωρητικό μέρος – transfer learning), ο διαμοιρασμός παραμέτρων (parameter-sharing) και ο λεπτομερής συντονισμός (fine tuning). Όπως προαναφέρθηκε σε εκείνο το σημείο, δεδομένου ότι το δεν υπάρχει καμία εγγύηση ότι οι παράμετροι των νευρωνικών δικτύων για DSSE θα είναι σε μεγάλο βαθμό μεταβιβάσιμες για διαφορετικές τοπολογίες, η τεχνική μεταφοράς μάθησης που χρησιμοποιείται κατά κύριο λόγο στην παρούσα εργασία είναι η μεταφορά παραμέτρων μέσω λεπτομερούς συντονισμού, για την υλοποίηση DSSE όταν αλλάζει η τοπολογία, καθώς ο διαμοιρασμός παραμέτρων απαιτεί οι παράμετροι είναι ιδιαίτερα μεταβιβάσιμες, κι έτσι αυτές του μοντέλου για το αρχικό πρόβλημα μπορούν να αντιγραφούν άμεσα στον τομέα-στόχο, όπου διατηρούνται "παγωμένες". Ωστόσο, για λόγους επαλήθευσης αυτής της συλλογιστικής εξετάστηκε σε περιορισμένο βαθμό και ο διαμοιρασμός παραμέτρων.

Η διαδικασία του fine tuning, που ουσιαστικά αποτελεί εκπαίδευση ξανά του μοντέλου παλινδρόμησης με τροφοδότηση νέων δεδομένων, υλοποιήθηκε μέσω των ίδιων συναρτήσεων που είχαν κατασκευαστεί για την αξιολόγηση του μοντέλου κατά την υλοποίηση του αλγορίθμου τοποθέτησης SMDs. Αρχικά το νευρωνικό δίκτυο παλινδρόμησης με τις υπερπαραμέτρους που βρέθηκαν ως βέλτιστες κατά το grid search εκπαιδεύτηκε με τα δεδομένα ρεύματος και τάσης για την πρώτη τοπολογία (χωρίς να κλείσει κανένας διακόπτης) κι ύστερα εκπαιδεύτηκε ξανά με τα δεδομένα της δεύτερης τοπολογίας (κλείσιμο πρώτου διακόπτη, 21 → 8) και αξιολογήθηκε η απόδοση του μοντέλου. Η διαδικασία αυτή εφαρμόστηκε διαδοχικά για όλες τις τοπολογίες και μπορεί να εφαρμοστεί και χωρίς αυτές να ακολουθούν απαραίτητα κάποια σειρά, π.χ. δηλαδή αν από την πρώτη περάσουμε στην έκτη τοπολογία με το κλείσιμο δύο διακοπών και το άνοιγμα αντίστοιχα των δύο κλάδων που τους αντιστοιχούν.

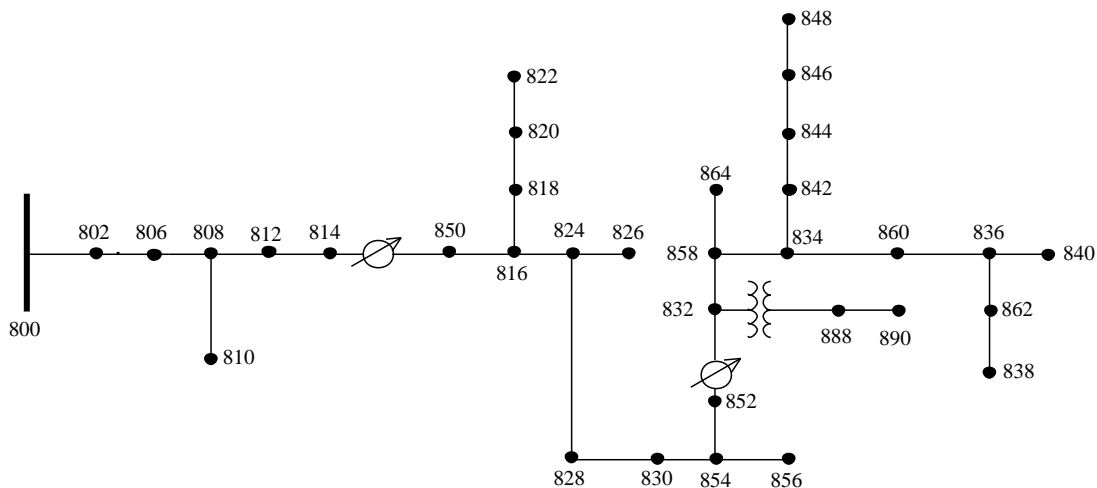
Ο διαμοιρασμός παραμέτρων υλοποιήθηκε παρόμοια. Αρχικά εκπαιδεύτηκε το βέλτιστο μοντέλο με τα δεδομένα της πρώτης τοπολογίας. Κατόπιν, τα κοινά στρώματα του μοντέλου παγώνουν, δηλαδή ορίζονται ως μη εκπαιδύσιμα και στην συνέχεια οι δύο κλάδοι του δικτύου αφαιρούνται και προστίθενται άλλοι δύο πανομοιότυποι στη θέση τους, οι οποίοι όμως δεν είναι εκπαιδευμένοι κι έτσι δεν περιέχουν τα σωστά βάρη. Στη συνέχεια, το δίκτυο εκπαιδεύεται (ουσιαστικά μόνο οι δύο κλάδοι κι όχι τα κοινά στρώματα), με τα δεδομένα της επόμενης τοπολογίας και η ακρίβεια του μοντέλου αξιολογείται.

Κεφάλαιο V

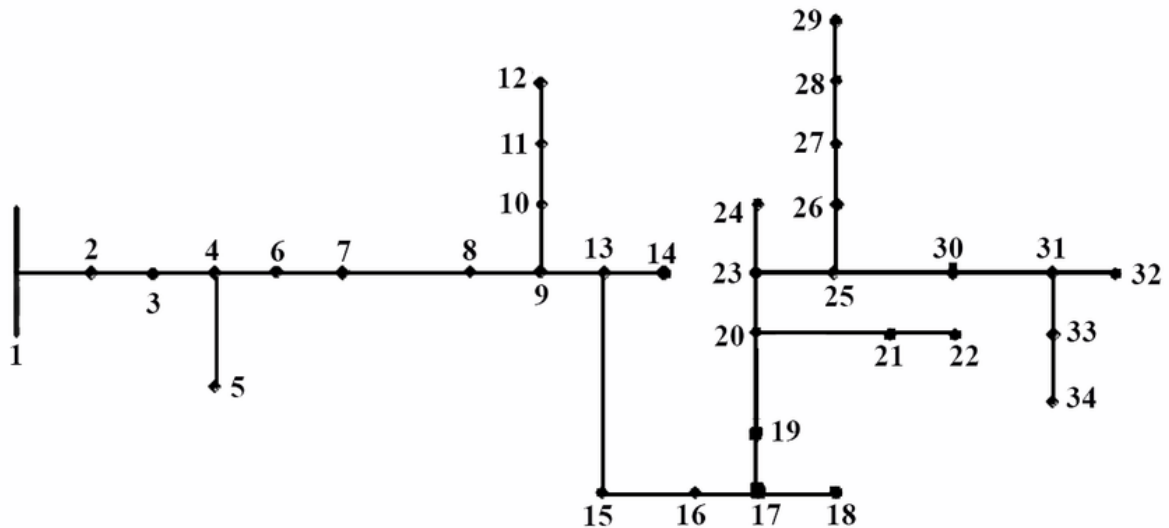
Δίκτυα Προσομοιώσεων

IEEE34

Αρχικά, για το δίκτυο προσομοίωσης του συστήματος που θα χρησιμοποιούνταν για την παραγωγή των μετρήσεων είχε επιλεγθεί το IEEE 34. Ουσιαστικά είναι ένα απλουστευμένο αλλά ρεαλιστικό μοντέλο ενός δικτύου διανομής 34 ζυγών που χρησιμοποιείται συχνά στην ανάλυση και την έρευνα συστημάτων ηλεκτρικής ενέργειας. Αποτελεί μέρος του IEEE Reliability Test System (RTS- μια συλλογή περιπτώσεων δοκιμών που αναπτύχθηκαν με σκοπό την παροχή τυποποιημένων συστημάτων δοκιμών για την αξιολόγηση της αξιοπιστίας και της απόδοσης των τεχνικών ανάλυσης του συστήματος ηλεκτρικής ενέργειας.



Στο παραπάνω σχήμα απεικονίζεται το δίκτυο IEEE34 όπως δίνεται επίσημα από την IEEE. Το σχήμα παρακάτω παρατίθεται για να φανεί η σωστή αρίθμηση των ζυγών που χρησιμοποιήθηκε:



Στον πίνακα παρακάτω παρατίθενται συνοπτικά τα δεδομένα των ζυγών του δικτύου:

3	1	0.05500	0.02900	0.00	0.00	1	1.00	0.00	*	BUS-806	*	24.9	1
5	1	0.01600	0.00800	0.00	0.00	1	1.00	0.00	*	BUS-810	*	24.9	1
11	1	0.03400	0.01700	0.00	0.00	1	1.00	0.00	*	BUS-820	*	24.9	1
12	1	0.13500	0.07000	0.00	0.00	1	1.00	0.00	*	BUS-822	*	24.9	1
13	1	0.00500	0.00200	0.00	0.00	1	1.00	0.00	*	BUS-824	*	24.9	1
14	1	0.04000	0.02000	0.00	0.00	1	1.00	0.00	*	BUS-826	*	24.9	1
15	1	0.00400	0.00200	0.00	0.00	1	1.00	0.00	*	BUS-828	*	24.9	1
16	1	0.05200	0.02300	0.00	0.00	1	1.00	0.00	*	BUS-830	*	24.9	1
18	1	0.00400	0.00200	0.00	0.00	1	1.00	0.00	*	BUS-856	*	24.9	1
22	1	0.45000	0.22500	0.00	0.00	1	1.00	0.00	*	BUS-890	*	4.0	1
23	1	0.01500	0.00700	0.00	0.00	1	1.00	0.00	*	BUS-858	*	24.9	1
24	1	0.00200	0.00100	0.00	0.00	1	1.00	0.00	*	BUS-864	*	24.9	1
25	1	0.03200	0.01700	0.00	0.00	1	1.00	0.00	*	BUS-834	*	24.9	1
27	1	0.41400	0.02000	0.00	0.00	1	1.00	0.00	*	BUS-844	*	24.9	1
28	1	0.04500	0.01700	0.00	0.00	1	1.00	0.00	*	BUS-846	*	24.9	1
29	1	0.08300	-0.39100	0.00	0.00	1	1.00	0.00	*	BUS-848	*	24.9	1
30	1	0.20600	0.12100	0.00	0.00	1	1.00	0.00	*	BUS-860	*	24.9	1
31	1	0.08200	0.04300	0.00	0.00	1	1.00	0.00	*	BUS-836	*	24.9	1
33	1	0.02800	0.01400	0.00	0.00	1	1.00	0.00	*	BUS-838	*	24.9	1
34	1	0.06700	0.04100	0.00	0.00	1	1.00	0.00	*	BUS-840	*	24.9	1
2	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-802	*	24.9	1
4	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-808	*	24.9	1
6	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-812	*	24.9	1
7	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-814	*	24.9	1
8	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-850	*	24.9	1
9	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-816	*	24.9	1
10	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-818	*	24.9	1
17	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-854	*	24.9	1
19	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-852	*	24.9	1
20	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-832	*	24.9	1
21	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-888	*	4.0	1
26	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-842	*	24.9	1
32	1	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-862	*	24.9	1
1	3	0.00000	0.00000	0.00	0.00	1	1.00	0.00	*	BUS-800	*	24.9	1

Οι στήλες αντιπροσωπεύουν αντίστοιχα:

Επίσης θεωρούμε πως υπάρχει μια σύγχρονη γεννήτρια συνδεδεμένη στον ζυγό 1 η οποία παράγει ενεργό ισχύ 5 MW και άεργο 2 MVAR

Για την δημιουργία διαφορετικών τοπολογιών εισήχθησαν διακόπτες ανάμεσα στους ζυγούς 7 → 8 και 19 → 20 (814 - 850 και 852 - 832 αντίστοιχα, φαίνονται και στο πρώτο σχήμα του δικτύου IEEE34), προσθέτοντας τις κατάλληλες γραμμές στα δεδομένα των ζυγών και των γραμμών. Ουσιαστικά κάθε διακόπτης μοντελοποιήθηκε σαν ένας καινούργιος ζυγός, με μηδενική αντίσταση όταν είναι κλειστός και με άπειρη (999 Ωμ) όταν είναι ανοιχτός. Ωστόσο, κατά την επίλυση της ροής ισχύος στο matpower όταν οι διακόπτες αυτοί άνοιγαν για να παραχθούν νέες τοπολογίες, παρατηρήθηκε ότι ο αλγόριθμος Netwon - Raphson δεν συνέκλινε. Αυτό συνέβη λόγω του φαινομένου της νησιδοποίησης (islanding).

Ως νησιδοποίηση στα συστήματα ηλεκτρικής ενέργειας αναφέρεται η κατάσταση κατά την οποία ένα τμήμα του ηλεκτρικού δικτύου απομονώνεται από το κύριο δίκτυο ηλεκτρικής ενέργειας, διατηρώντας παράλληλα την τοπική παραγωγή και το τοπικό φορτίο, με αποτέλεσμα λειτουργεί ανεξάρτητα από το κύριο δίκτυο και να μην είναι δυνατός ο έλεγχος του. Το φαινόμενο αυτό συνήθως εμφανίζεται κατά τη διάρκεια σφαλμάτων του συστήματος και βλαβών του εξοπλισμού. Σε τέτοιες περιπτώσεις οι ηλεκτρονόμοι προστασίας έχουν σχεδιαστεί για να ανιχνεύουν και να απομονώνουν την επηρεαζόμενη περιοχή. Ωστόσο, εάν τα συστήματα προστασίας δεν είναι σε θέση να εντοπίσουν και να διαχωρίσουν το ελαττωματικό τμήμα από το υπόλοιπο δίκτυο, μπορεί να σχηματιστεί μια ακούσια νησίδα, γεγονός μπορεί να δημιουργήσει σοβαρές προκλήσεις, καθώς η απομονωμένη νησίδα η οποία είναι εκτός ελέγχου μπορεί να μην μπορεί να διατηρήσει ισορροπία μεταξύ παραγωγής και ζήτησης ενέργειας κι έτσι να επέλθει αστάθεια της τάσης, αποκλίσεις στη συχνότητα και πιθανές περαιτέρω βλάβες του εξοπλισμού.

Αντιθέτως, μερικές φορές νησιδοποίηση προκαλείται σκόπιμα, ως μια προληπτική προσέγγιση για την ενίσχυση της ανθεκτικότητας του δικτύου και την μείωση της επίπτωσης των διαταραχών. Με την εσκεμμένη απομόνωση ενός τμήματος του δικτύου, μπορεί να ελαχιστοποιηθεί ο κίνδυνος αλυσιδωτών βλαβών και να επιτραπεί η δημιουργία αυτόνομων μικροδικτύων με τους δικούς τους καταναμημένους ενεργειακούς πόρους και συστήματα αποθήκευσης ενέργειας. Αυτή η αυτονομία ενισχύει την αξιοπιστία παρέχοντας μια τοπική και αυτόνομη παροχή ενέργειας. Επίσης, η σκόπιμη απομόνωση αποτελεί βασικό στοιχείο των στρατηγικών ανθεκτικότητας του δικτύου καθώς επιτρέπει τη δημιουργία ανθεκτικών ζωνών εντός του ευρύτερου δικτύου, διασφαλίζοντας ότι ορισμένες περιοχές μπορούν να διατηρήσουν τη λειτουργικότητά τους ακόμη και όταν αντιμετωπίζουν εξωτερικές διαταραχές. Γενικότερα, η αντιμετώπιση της νησιδοποίησης όταν αυτή δημιουργείται ακούσια, όπως στο παράδειγμα μας, αποτελεί μεγάλη πρόκληση για την διασφάλιση της σταθερότητας και της αξιοπιστίας των συστημάτων ηλεκτρικής ενέργειας.

Η εκτίμηση τοπολογίας γίνεται για τοπολογίες που επιτρέπουν την κανονική λειτουργία ενός ΣΗΕ χωρίς νησιδοποίηση, οι οποίες χαρακτηρίζονται και εφικτές (feasible topologies). Συνεπώς, για ένα ακτινικό δίκτυο όπως είναι το IEEE34, δεν είναι εφικτό να υπάρξει άνοιγμα

διακόπτη χωρίς νησιδοποίηση και λόγω αυτού του φαινομένου, το matpower δεν μπορεί να επιλύσει τη ροή ισχύος καθώς το απομονωμένο κομμάτι είναι πλήρως μη παρατηρήσιμο. Δηλαδή, οι πληροφορίες σχετικά με τις συνδέσεις και τις σχέσεις μεταξύ των διαφόρων τμημάτων του συστήματος διανομής χάνονται και οι μετρήσεις πλέον δεν αντιπροσωπεύουν όλο το δίκτυο αλλά μόνο το μη απομονωμένο κομμάτι.

IEEE33

Στη συνέχεια για την παραγωγή των τοπολογιών το δίκτυο IEEE33 το οποίο μοντελοποιήθηκε στο matlab μέσω του παραδείγματος «case33mg» που παρέχει το matpower, και το οποίο βασίζεται εδώ: [72]. Το δίκτυο αυτό αναπτύχθηκε το 1989 από τους Baran και Wu, (συγγραφείς του άρθρου), με σκοπό την ανάπτυξη της αναδιαμόρφωσης των δικτύων διανομής με στόχο την μείωση των απωλειών ισχύος και την εξισορρόπηση του φορτίου. Σήμερα, χρησιμοποιείται ευρέως ως σύστημα δοκιμών για τη μελέτη διαφόρων προβλημάτων στα δίκτυα διανομής.

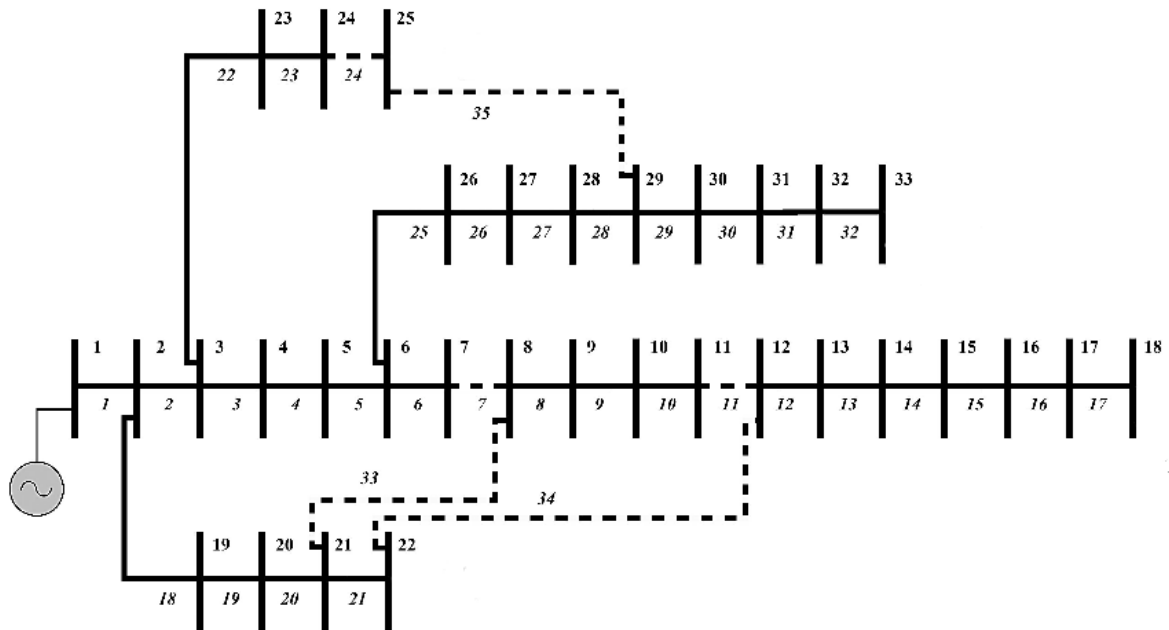
Το σύστημα αυτό αποτελείται από 33 ζυγούς, 32 γραμμές που συνδέουν τους ζυγούς, 5 διακόπτες και δεν περιέχει καμία μονάδα αντιστάθμισης αέργου ισχύος. Το δίκτυο τροφοδοτείται μόνο από μία γεννήτρια που συνδέεται με τον πρώτο ζυγό και παρέχει 4 MW ενεργής ισχύος και 2,5 MVAR αέργου ισχύος. Τα όρια τάσης των ζυγών ορίζονται από 0,9 έως 1,1 (p.u.). Στους παρακάτω πίνακες παρατίθενται τα δεδομένα των ζυγών και των κλάδων του δικτύου αντίστοιχα.

bus_i	type	Pd	Qd	Gs	Bs	area	Vm	Va	baseKV	zone	Vmax	Vmin
1	3	0	0	0	0	1	1	0	12.66	1	1	1;
2	1	100	60	0	0	1	1	0	12.66	1	1.1	0.9;
3	1	90	40	0	0	1	1	0	12.66	1	1.1	0.9;
4	1	120	80	0	0	1	1	0	12.66	1	1.1	0.9;
5	1	60	30	0	0	1	1	0	12.66	1	1.1	0.9;
6	1	60	20	0	0	1	1	0	12.66	1	1.1	0.9;
7	1	200	100	0	0	1	1	0	12.66	1	1.1	0.9;
8	1	200	100	0	0	1	1	0	12.66	1	1.1	0.9;
9	1	60	20	0	0	1	1	0	12.66	1	1.1	0.9;
10	1	60	20	0	0	1	1	0	12.66	1	1.1	0.9;
11	1	45	30	0	0	1	1	0	12.66	1	1.1	0.9;
12	1	60	35	0	0	1	1	0	12.66	1	1.1	0.9;
13	1	60	35	0	0	1	1	0	12.66	1	1.1	0.9;
14	1	120	80	0	0	1	1	0	12.66	1	1.1	0.9;
15	1	60	10	0	0	1	1	0	12.66	1	1.1	0.9;
16	1	60	20	0	0	1	1	0	12.66	1	1.1	0.9;
17	1	60	20	0	0	1	1	0	12.66	1	1.1	0.9;
18	1	90	40	0	0	1	1	0	12.66	1	1.1	0.9;
19	1	90	40	0	0	1	1	0	12.66	1	1.1	0.9;
20	1	90	40	0	0	1	1	0	12.66	1	1.1	0.9;
21	1	90	40	0	0	1	1	0	12.66	1	1.1	0.9;
22	1	90	40	0	0	1	1	0	12.66	1	1.1	0.9;
23	1	90	50	0	0	1	1	0	12.66	1	1.1	0.9;
24	1	420	200	0	0	1	1	0	12.66	1	1.1	0.9;
25	1	420	200	0	0	1	1	0	12.66	1	1.1	0.9;
26	1	60	25	0	0	1	1	0	12.66	1	1.1	0.9;
27	1	60	25	0	0	1	1	0	12.66	1	1.1	0.9;
28	1	60	20	0	0	1	1	0	12.66	1	1.1	0.9;
29	1	120	70	0	0	1	1	0	12.66	1	1.1	0.9;
30	1	200	600	0	0	1	1	0	12.66	1	1.1	0.9;
31	1	150	70	0	0	1	1	0	12.66	1	1.1	0.9;
32	1	210	100	0	0	1	1	0	12.66	1	1.1	0.9;
33	1	60	40	0	0	1	1	0	12.66	1	1.1	0.9;

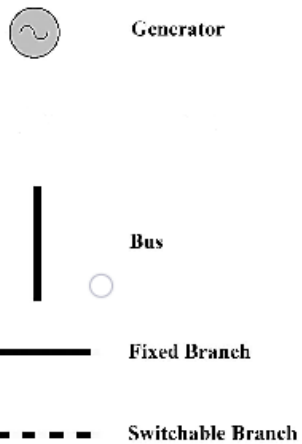
fbus	tbus	r($\Omega\mu$)	x($\Omega\mu$)	b	rateA	rateB	rateC	ratio	angle	status	angmin	angmax
1	2	0.0922	0.0470	0	0	0	0	0	0	1	-360	360
2	3	0.4930	0.2511	0	0	0	0	0	0	1	-360	360
3	4	0.3660	0.1864	0	0	0	0	0	0	1	-360	360
4	5	0.3811	0.1941	0	0	0	0	0	0	1	-360	360
5	6	0.8190	0.7070	0	0	0	0	0	0	1	-360	360
6	7	0.1872	0.6188	0	0	0	0	0	0	1	-360	360
7	8	0.7114	0.2351	0	0	0	0	0	0	1	-360	360
8	9	1.0300	0.7400	0	0	0	0	0	0	1	-360	360
9	10	1.0440	0.7400	0	0	0	0	0	0	1	-360	360
10	11	0.1966	0.0650	0	0	0	0	0	0	1	-360	360
11	12	0.3744	0.1238	0	0	0	0	0	0	1	-360	360
12	13	1.4680	1.1550	0	0	0	0	0	0	1	-360	360
13	14	0.5416	0.7129	0	0	0	0	0	0	1	-360	360
14	15	0.5910	0.5260	0	0	0	0	0	0	1	-360	360
15	16	0.7463	0.5450	0	0	0	0	0	0	1	-360	360
16	17	1.2890	1.7210	0	0	0	0	0	0	1	-360	360
17	18	0.7320	0.5740	0	0	0	0	0	0	1	-360	360
2	19	0.1640	0.1565	0	0	0	0	0	0	1	-360	360
19	20	1.5042	1.3554	0	0	0	0	0	0	1	-360	360
20	21	0.4095	0.4784	0	0	0	0	0	0	1	-360	360
21	22	0.7089	0.9373	0	0	0	0	0	0	1	-360	360
3	23	0.4512	0.3083	0	0	0	0	0	0	1	-360	360
23	24	0.8980	0.7091	0	0	0	0	0	0	1	-360	360
24	25	0.8960	0.7011	0	0	0	0	0	0	1	-360	360
6	26	0.2030	0.1034	0	0	0	0	0	0	1	-360	360
26	27	0.2842	0.1447	0	0	0	0	0	0	1	-360	360
27	28	1.0590	0.9337	0	0	0	0	0	0	1	-360	360
28	29	0.8042	0.7006	0	0	0	0	0	0	1	-360	360
29	30	0.5075	0.2585	0	0	0	0	0	0	1	-360	360
30	31	0.9744	0.9630	0	0	0	0	0	0	1	-360	360
31	32	0.3105	0.3619	0	0	0	0	0	0	1	-360	360
32	33	0.3410	0.5302	0	0	0	0	0	0	1	-360	360
21	8	2.0000	2.0000	0	0	0	0	0	0	0	-360	360
12	22	2.0000	2.0000	0	0	0	0	0	0	0	-360	360
25	29	0.5000	0.5000	0	0	0	0	0	0	0	-360	360

Από τους αρχικούς 5 διακόπτες του συστήματος αφαιρέθηκαν οι 9 \rightarrow 15 και 18 \rightarrow 33 προκειμένου να αποφευχθεί η περιττή πολυπλοκότητα. Η τροποποίηση αυτή απλώς απλοποιεί τη διαδικασία επαναδιαμόρφωσης με την εξάλειψη περιττών περιπλοκών. Ωστόσο, στη συνέχεια, προστέθηκαν τρεις νέοι διακόπτες, στους κλάδους 28 \rightarrow 29, 11 \rightarrow 12 και 7 \rightarrow 8 (κάθε φορά ανάμεσα στον ζυγό στον οποίο καταλήγει ένας διακόπτης και του προηγούμενου του), έτσι ώστε οι τοπολογίες που θα δημιουργηθούν να μην γίνουν πλέγματος αλλά να παραμείνουν ακτινικές. Με αυτό τον τρόπο, το IEEE33 θα αποτελεί ένα ακτινικό δίκτυο στο οποίο μέσω της χρήσης διακοπών τμηματοποίησης δημιουργούνται βρόγχοι κι έτσι μπορούν να

μελετηθούν και διαφορετικές τοπολογίες για εκτίμηση κατάστασης. Στο παρακάτω σχήμα αποτυπώνεται το δίκτυο που χρησιμοποιήθηκε:



με:



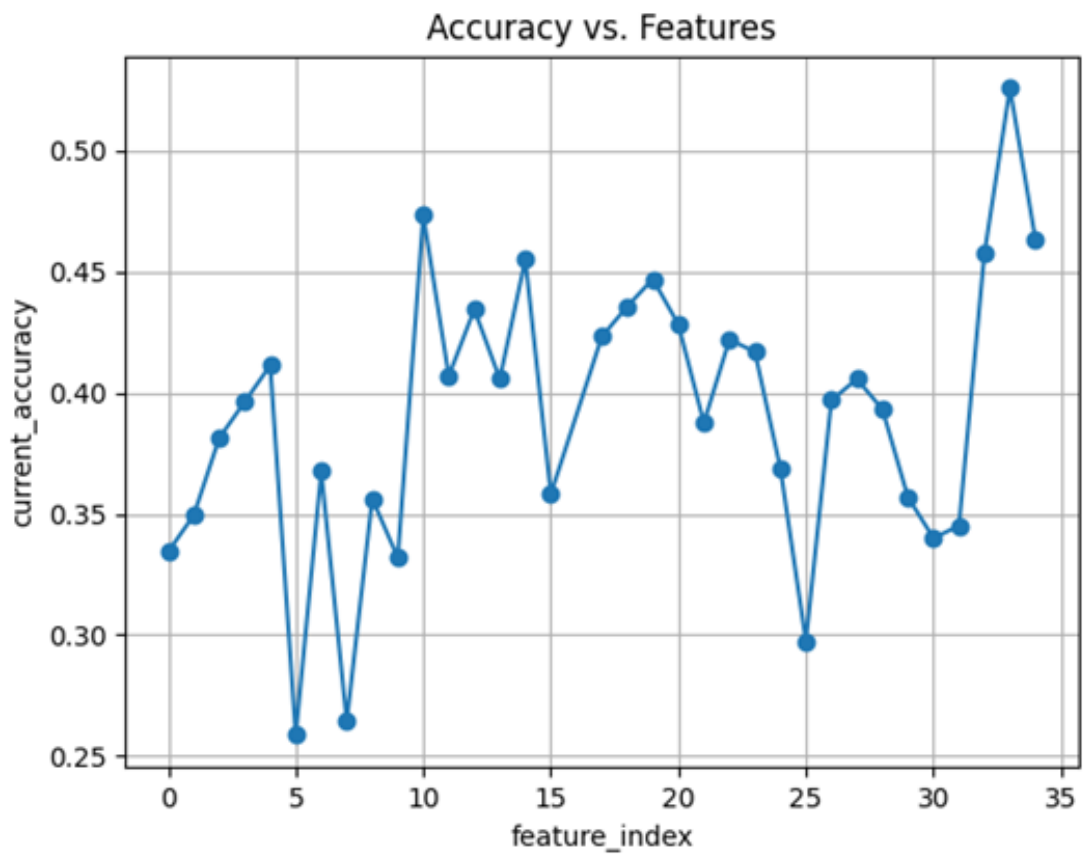
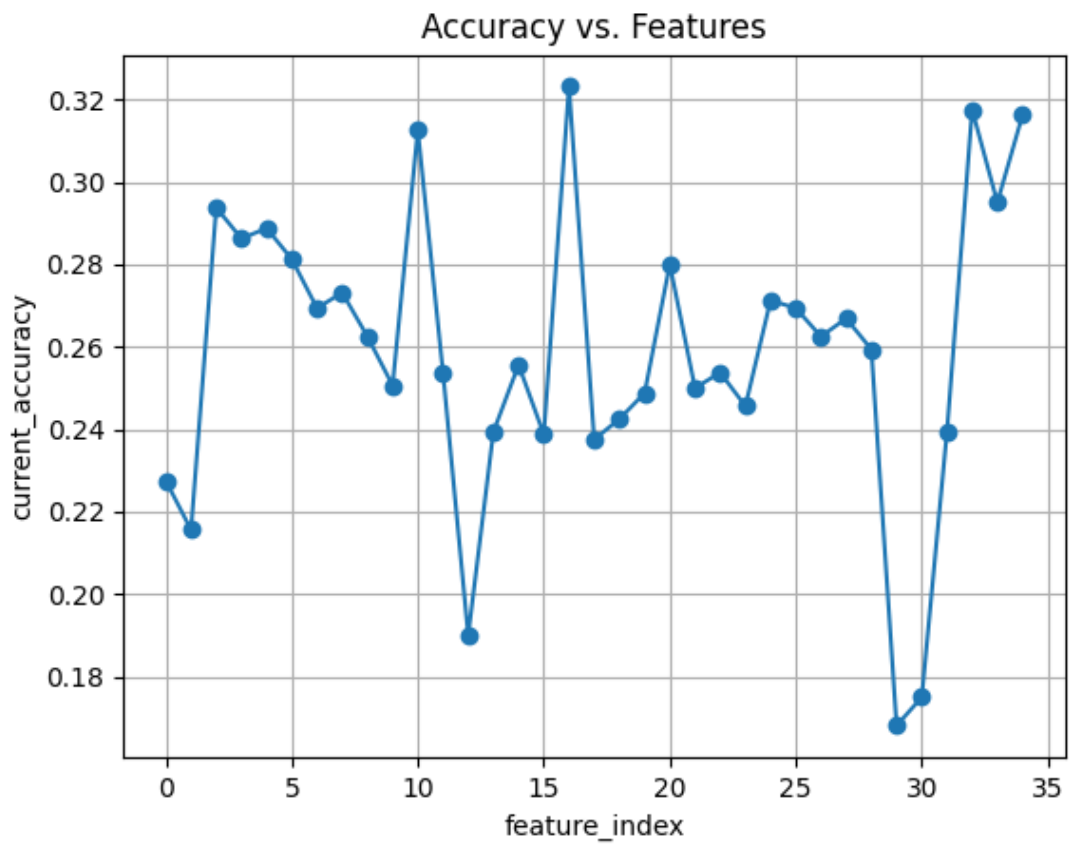
Στους πίνακες παρακάτω παρατίθενται ξανά οι πιθανοί συνδυασμοί των διακοπών που ορίζουν τις διαφορετικές τοπολογίες:

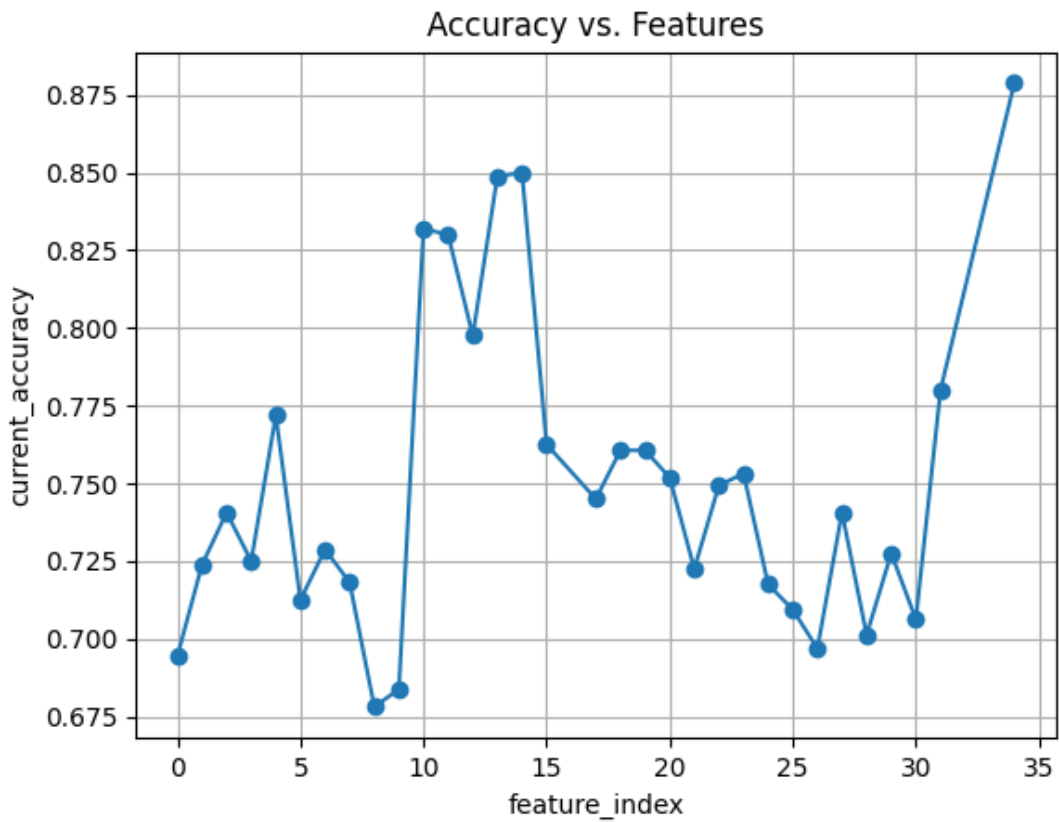
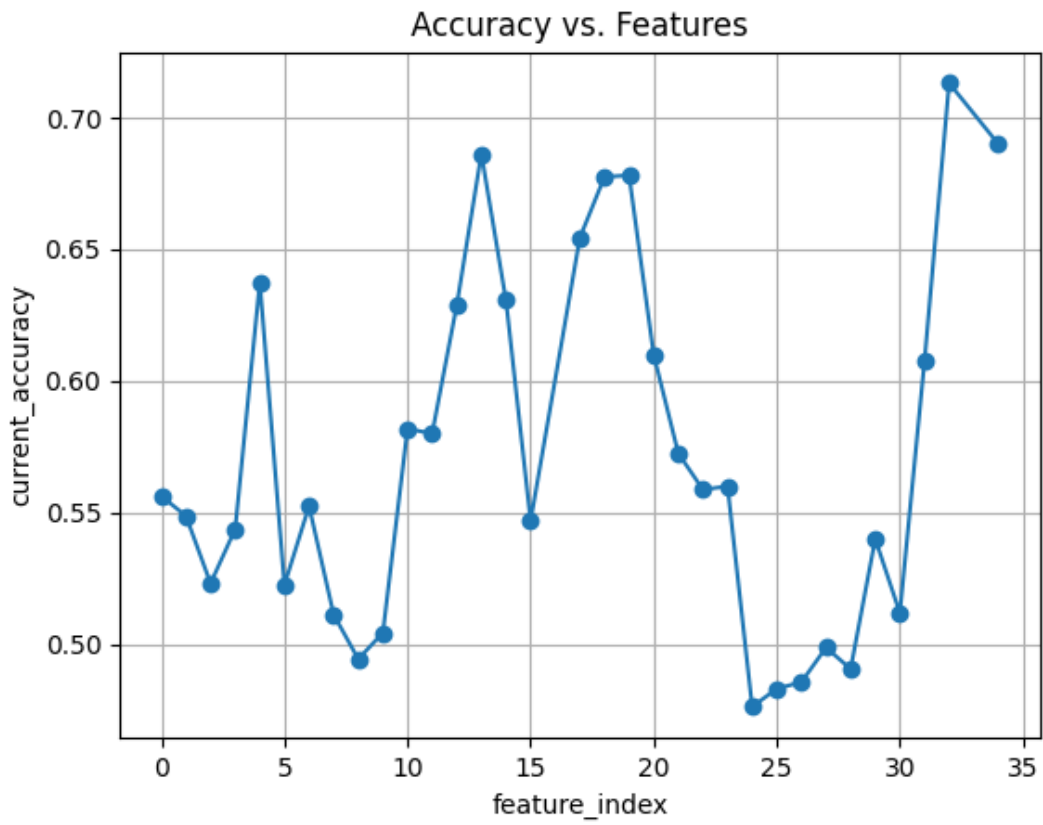
Διακόπτης 25-> 29	Διακόπτης 12 -> 22	Διακόπτης 21 -> 8
ανοικτός	ανοικτός	ανοικτός
ανοικτός	ανοικτός	κλειστός
ανοικτός	κλειστός	ανοικτός
ανοικτός	κλειστός	κλειστός
κλειστός	ανοικτός	ανοικτός
κλειστός	ανοικτός	κλειστός
κλειστός	κλειστός	ανοικτός
κλειστός	κλειστός	κλειστός

Διακόπτης 28 -> 29	Διακόπτης 11 -> 12	Διακόπτης 7 -> 8
κλειστός	κλειστός	κλειστός
κλειστός	κλειστός	ανοικτός
κλειστός	ανοικτός	κλειστός
κλειστός	ανοικτός	ανοικτός
ανοικτός	κλειστός	κλειστός
ανοικτός	κλειστός	ανοικτός
ανοικτός	ανοικτός	κλειστός
ανοικτός	ανοικτός	ανοικτός

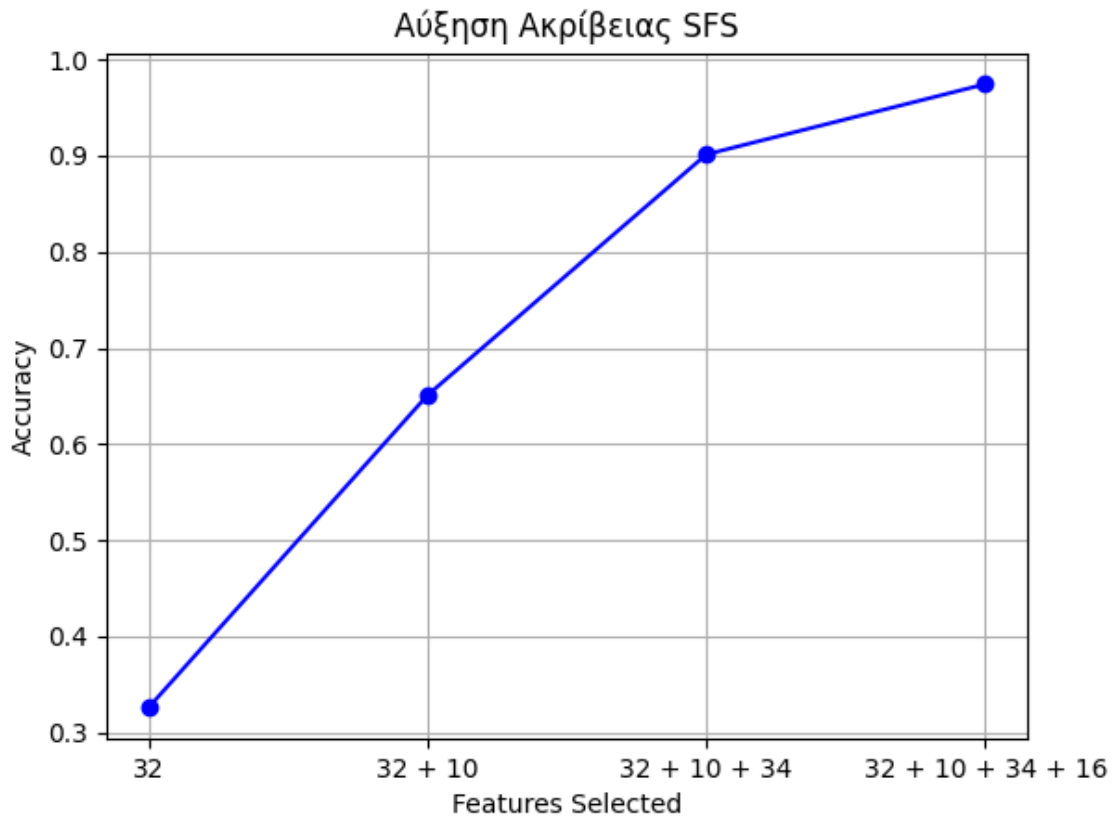
Αποτελέσματα Εκτίμησης Τοπολογίας

Στα παρακάτω διαγράμματα φαίνεται η διαδικασία επιλογής χαρακτηριστικών μέσω της χρήσης sfs. Το χαρακτηριστικό που δίνει την μεγαλύτερη ακρίβεια στο μοντέλο επιλέγεται κάθε φορά. Παρατηρείται πως η ακρίβεια του μοντέλου αυξάνεται σημαντικά στο επόμενο βήμα (επόμενο διάγραμμα), κάθε φορά που επιλέγεται κάποιο χαρακτηριστικό.



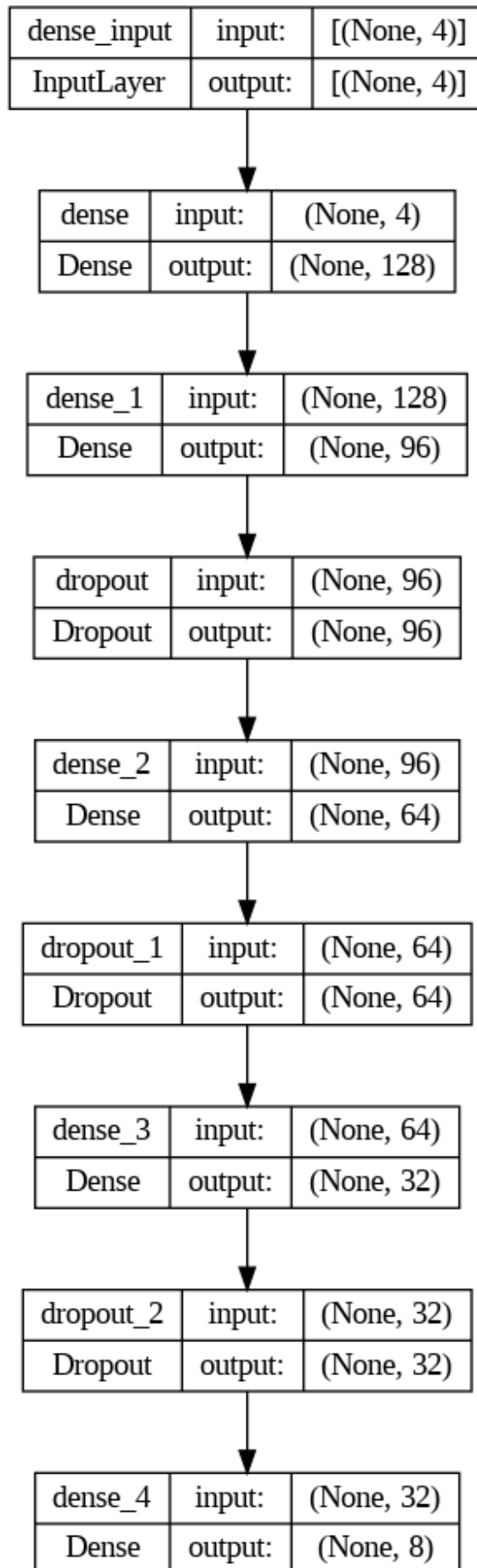


Στο παρακάτω διάγραμμα αποτυπώνεται η αύξηση της ακρίβειας του μοντέλου ταξινόμησης καθώς προστίθενται συνεχώς χαρακτηριστικά μέσω της sfs:



Έτσι, κατά την πρώτη επανάληψη της sfs επιλέγεται το χαρακτηριστικό 32, δηλαδή ο κλάδος $21 \rightarrow 8$ (διακόπτης 1), και μετά την εκπαίδευση, το μοντέλο έχει ακρίβεια 0.325625% στα δεδομένα πρόβλεψης. Κατά την δεύτερη επανάληψη, επιλέγεται η στήλη 10, δηλαδή ο κλάδος $11 \rightarrow 12$, και μετά την εκπαίδευση, το μοντέλο έχει ακρίβεια 0.650625%. Το τρίτο χαρακτηριστικό που επιλέγεται είναι το 34 που αντιστοιχεί στον κλάδο ανάμεσα στους ζυγούς 25 και 29 (διακόπτης 3) και όταν το μοντέλο εκπαιδεύεται μόνο με τα δεδομένα των στυλών 32, 10 και 34 αποδίδει ακρίβεια 0.90125%. Τέλος, όταν προστεθεί και η στήλη 16, δηλαδή τα δεδομένα του κλάδου από τον ζυγό 17 στον 18, το μοντέλο ταξινόμησης παρουσιάζει ακρίβεια 0.974375% οπότε και ο αλγόριθμος σταματά σύμφωνα με τα κριτήρια που τέθηκαν νωρίτερα.

Το νευρωνικό δίκτυο που τελικά κατασκευάστηκε για ΤΙ με βάση την απόδοση και τον αριθμό SMDs που χρειάζονται φαίνεται στο παρακάτω διάγραμμα:



Στους παρακάτω πίνακες απεικονίζονται ενδεικτικά μερικά παραδείγματα δοκιμών που έγιναν με διάφορες υπερπαραμέτρους για την εύρεση του ιδανικού μοντέλου:

Επιλεγμένα Χαρακτηριστικά	Ακρίβεια	Κρυφά Στρώματα και Νευρώνες
[4, 33, 32, 34, 15, 10, 27, 3, 11, 31]	0.945625	64 32
[4, 20, 32, 10, 34, 11, 15, 31, 17, 16]	0.935625	64
[10, 32, 34, 4, 0]	0.960625	128 64 32
[10, 32, 34, 19, 26, 31]	0.9575	128 64 32
[10, 32, 34, 29]	0.96625	128 96 64 32
[32, 10, 34, 4]	0.973125	128 96 64 32
[32, 10, 34, 29]	0.98125	256 128 96 64 32
[34, 32, 10, 16]	0.98875	256 128 96 64 32
[34, 32, 10, 3]	0.975	512 256 128 96 64 32
[34, 10, 32, 28]	0.973125	512 256 128 96 64 32
[32, 10, 34, 5]	0.96375	32 64 96 128
[10, 32, 34, 6]	0.955	32 64 96 128

Επιλεγμένα Χαρακτηριστικά	Ακρίβεια	Epochs	Batch Size	Dropout
[32, 10, 34, 4]	0.973125	10	64	0.2
[10, 32, 34, 29]	0.96625	10	64	0.2
[34, 10, 32, 1]	0.956875	10	64	0.3
[32, 10, 34, 31]	0.963125	10	64	0.3
[16, 33, 6, 31, 34, 32]	0.995	20	32	0.2
[16, 33, 34, 32, 31, 2]	0.9875	10	128	0.2

Ουσιαστικά, όπως φαίνεται, αρχικά έγιναν δοκιμές μόνο με τον αριθμό των στρωμάτων και των νευρώνων. Διαπιστώθηκε πως τα δίκτυα (128 96 64 32), (256 128 96 64 32) και (512 256 128 96 64 32) χρειάζονται σταθερά μόνο 4 χαρακτηριστικά έτσι ώστε με την εκπαίδευση πάνω σε αυτά το μοντέλο ταξινόμησης να ξεπεράσει την επιθυμητή ακρίβεια. Επιλέχθηκε το (128 96 64 32) έτσι ώστε να αποφευχθεί το overfitting όσο το δυνατόν γίνεται. Το αντίστροφο μοντέλο (32 64 96 128) δίνει κι αυτό σταθερά αποδεκτά αποτελέσματα χρησιμοποιώντας μόνο 4 χαρακτηριστικά, όμως αυτά όπως φαίνεται δεν είναι σταθερά αλλά μεταβάλλονται από εκτέλεση σε εκτέλεση, οπότε και αυτή η περίπτωση απορρίφθηκε.

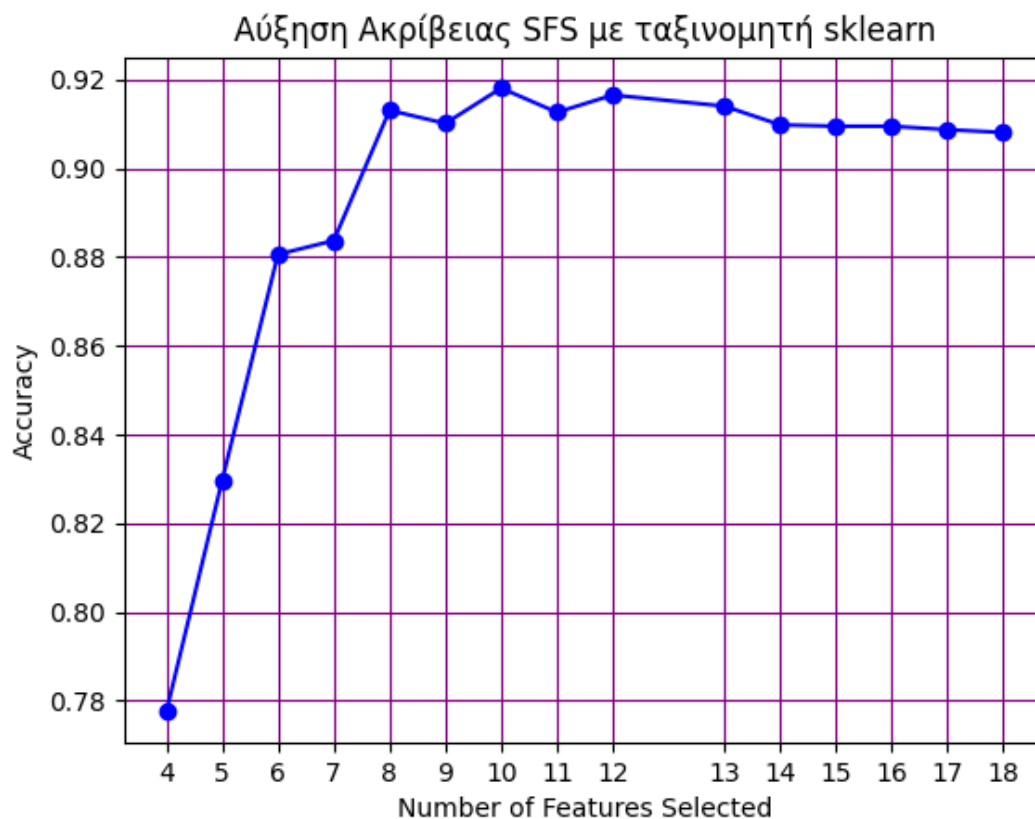
Στη συνέχεια, δοκιμάστηκε ένας συνδυασμός διαφορετικών τιμών epochs, batch size και τιμών στα στρώματα dropout που προστέθηκαν ανάμεσα από κάθε κρυφό στρώμα. Έτσι βρέθηκε πως ο συνδυασμός 10 epochs, batch size 64 και τιμή dropout κάθε τέτοιου στρώματος δίνει σταθερά ακρίβεια πάνω από 95% για 4 επιλεγμένα χαρακτηριστικά. Οι τιμές epochs ενδεικτικά θεωρούνται πολλαπλάσια του 10 και του batch size 32, 64

ή 128 επειδή έτσι συνηθίζεται κατά το hyperparameter tuning. Παρατηρείται πως τα πρώτα τρία επιλεγμένα χαρακτηριστικά είναι με διαφορά τα πιο σημαντικά, καθώς μόνο με αυτά η ακρίβεια του νευρωνικού δικτύου μετά την εκπαίδευση κυμαίνεται στο 91-94%. Το 4^ο χαρακτηριστικό που επιλέγεται έτσι ώστε η ακρίβεια του μοντέλου να αυξηθεί σε πάνω από 95% δεν είναι σταθερό αλλά μεταβάλλεται σχεδόν σε κάθε νέα εκτέλεση του αλγορίθμου. Δηλαδή η επιλογή του είναι σχεδόν τυχαία και δεν συμβάλλει όσο τα πρώτα τρία (32,10,34) τα οποία είναι πάντα σταθερά. Αυτό δείχνει ότι το τέταρτο χαρακτηριστικό είναι πολύ λιγότερο σημαντικό από τα πρώτα τρία και πιθανώς και περιττό. Ωστόσο δεν βρέθηκε κάποιο νευρωνικό δίκτυο που να ξεπερνά την επιθυμητή ακρίβεια χρησιμοποιώντας μόνο τα 3 πρώτα χαρακτηριστικά. Επίσης, εξετάστηκε και το learning rate, το οποίο από την default τιμή του adam optimizer 0.001 τέθηκε ως 0.0001 αλλά παρατηρήθηκε πως η ακρίβεια μειωνόταν ελαφρώς οπότε και η υπερπαράμετρος αυτή δεν εξετάστηκε περαιτέρω.

Από τις τεχνικές που εξετάστηκαν για να αποφευχθεί το overfitting, κρατήθηκαν μόνο τα στρώματα dropout ανάμεσα στα 3 τελευταία κρυφά στρώματα καθώς από τις δοκιμές φάνηκε πως αυξάνουν την ακρίβεια.

Για το επόμενο μέρος του αλγορίθμου θεωρήθηκε πως το τέταρτο χαρακτηριστικό θα είναι η στήλη 29 που αντιστοιχεί στον κλάδο ανάμεσα στους ζυγούς 30 και 31, καθώς βρέθηκε πως εμφανίζεται συγκριτικά λίγο περισσότερες φορές από τα υπόλοιπα, κατά την διάρκεια επανειλημμένων εκτελέσεων του αλγορίθμου.

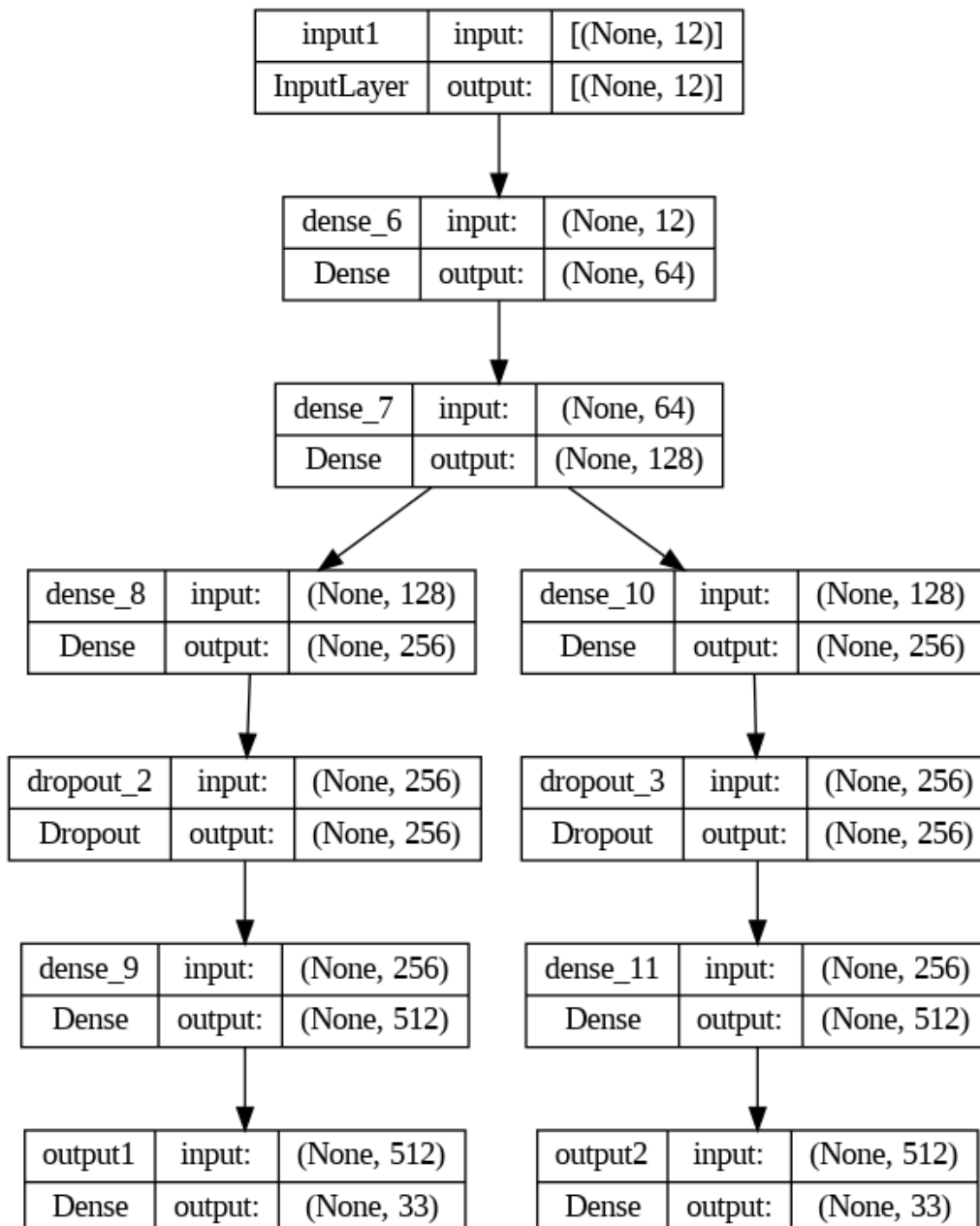
Στη συνέχεια, στο παρακάτω διάγραμμα παρατίθεται η ακρίβεια ενός τυπικού ταξινομητή από την βιβλιοθήκη sklearn (K-Nearest Neighbors), για σύγκριση με την απόδοση του νευρωνικού δικτύου.



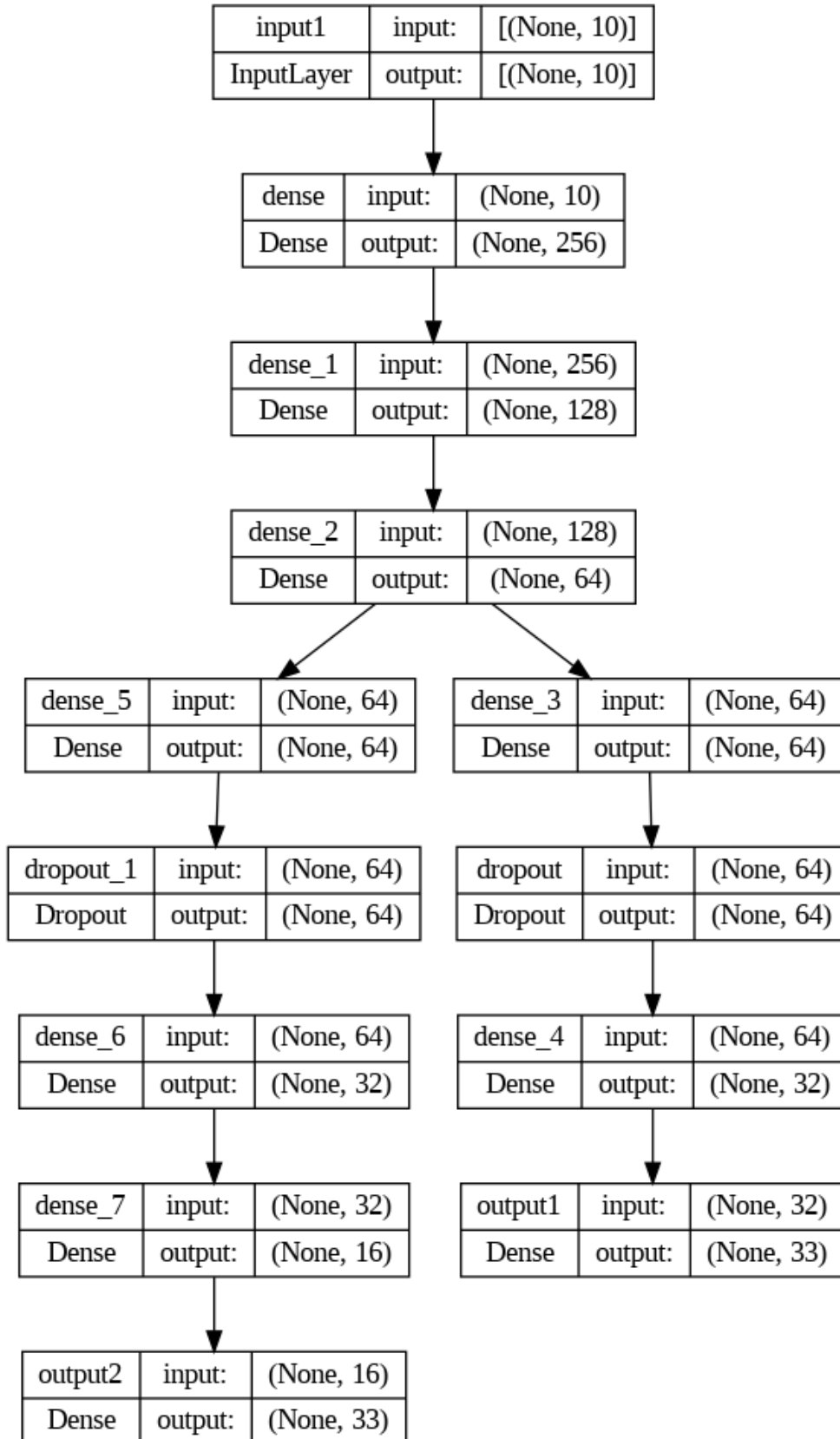
Παρατηρείται πως η ακρίβεια του ταξινομητή αγγίζει το 92% αλλά στη συνέχεια, παρόλο που προστίθενται συνεχώς και άλλα χαρακτηριστικά, αυτή δεν βελτιώνεται αλλά αντιθέτως χειροτερεύει ελαφρώς. Αυτό σημαίνει, πως από το 10^ο SMD και μετά ο ταξινομητής αντιλαμβάνεται ως θόρυβο τα επιπλέον χαρακτηριστικά.

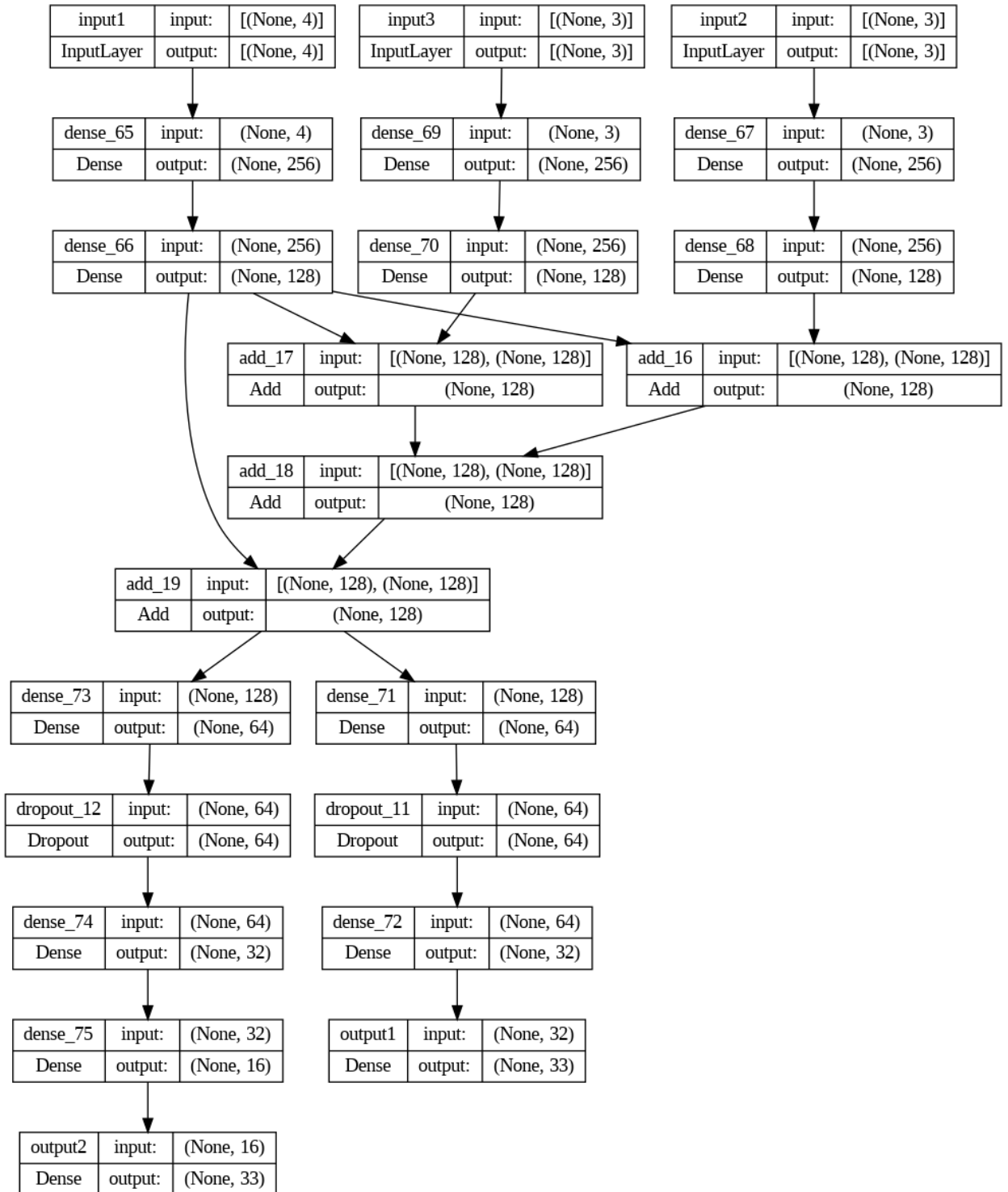
Αποτελέσματα Εκτίμησης Κατάστασης

Στο διάγραμμα παρακάτω απεικονίζεται το νευρωνικό δίκτυο παλινδρόμησης που τελικά κατασκευάστηκε για DSSE με βάση την απόδοση και τον αριθμό SMDs που χρειάζονται:

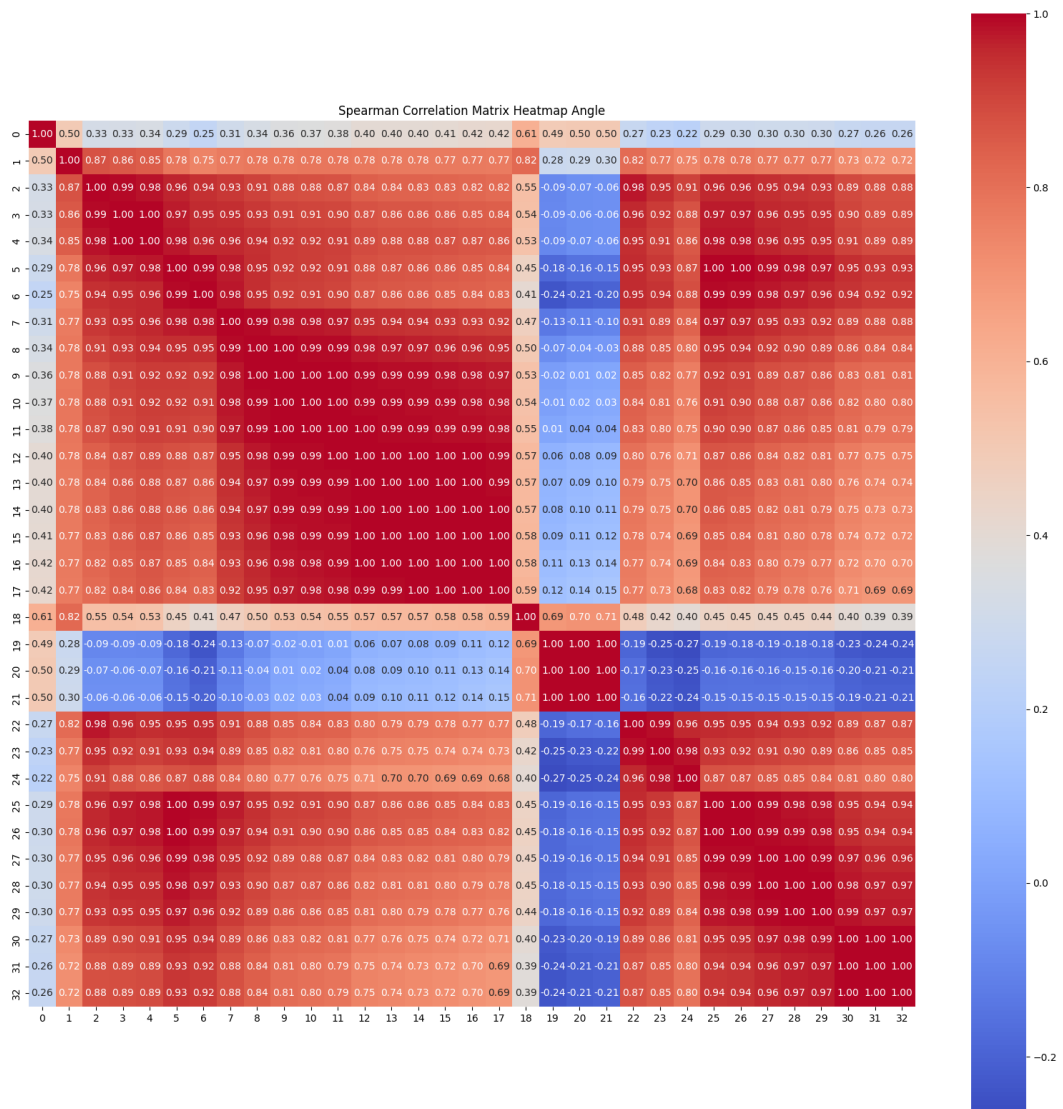


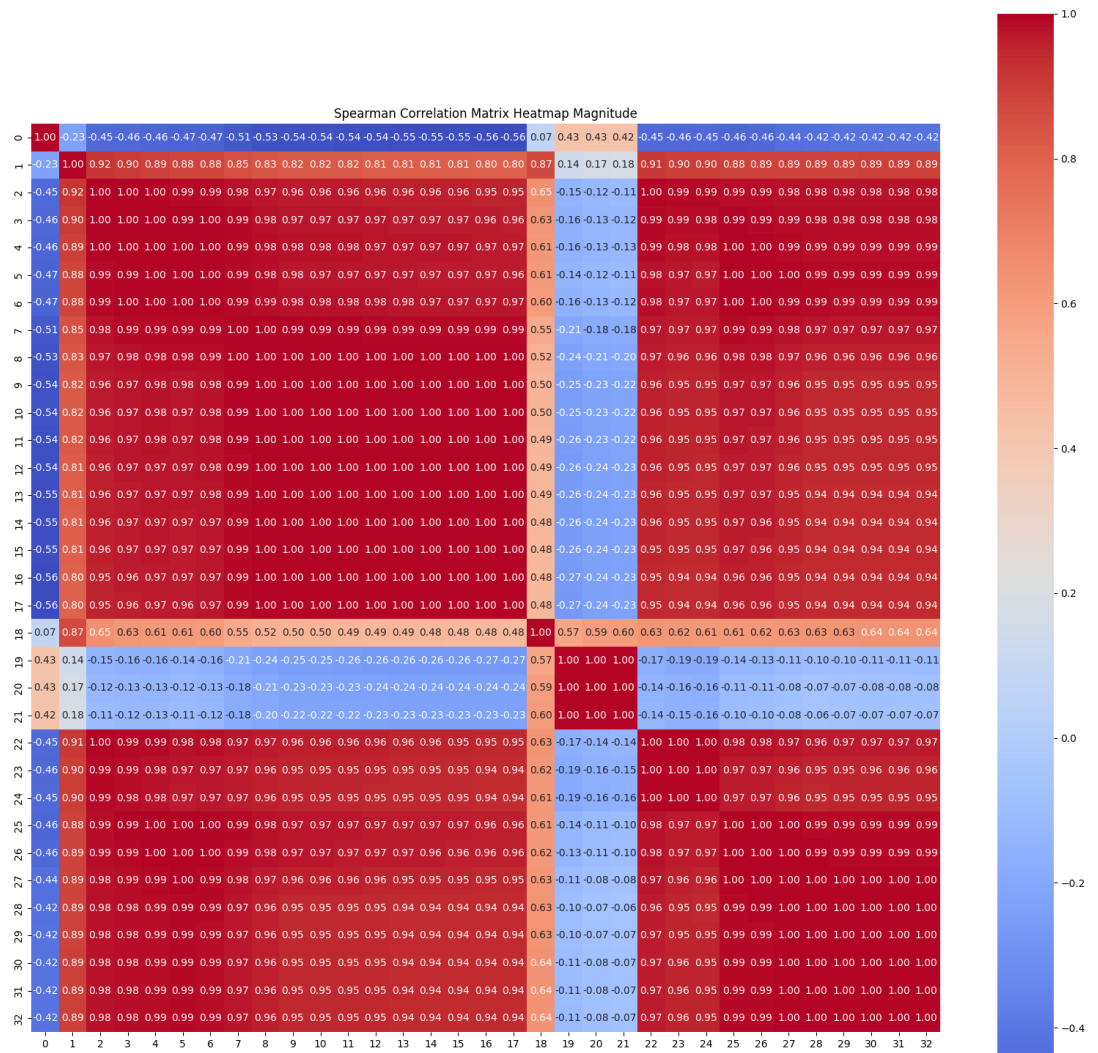
Παραδείγματα νευρωνικών που απορρίφθηκαν:



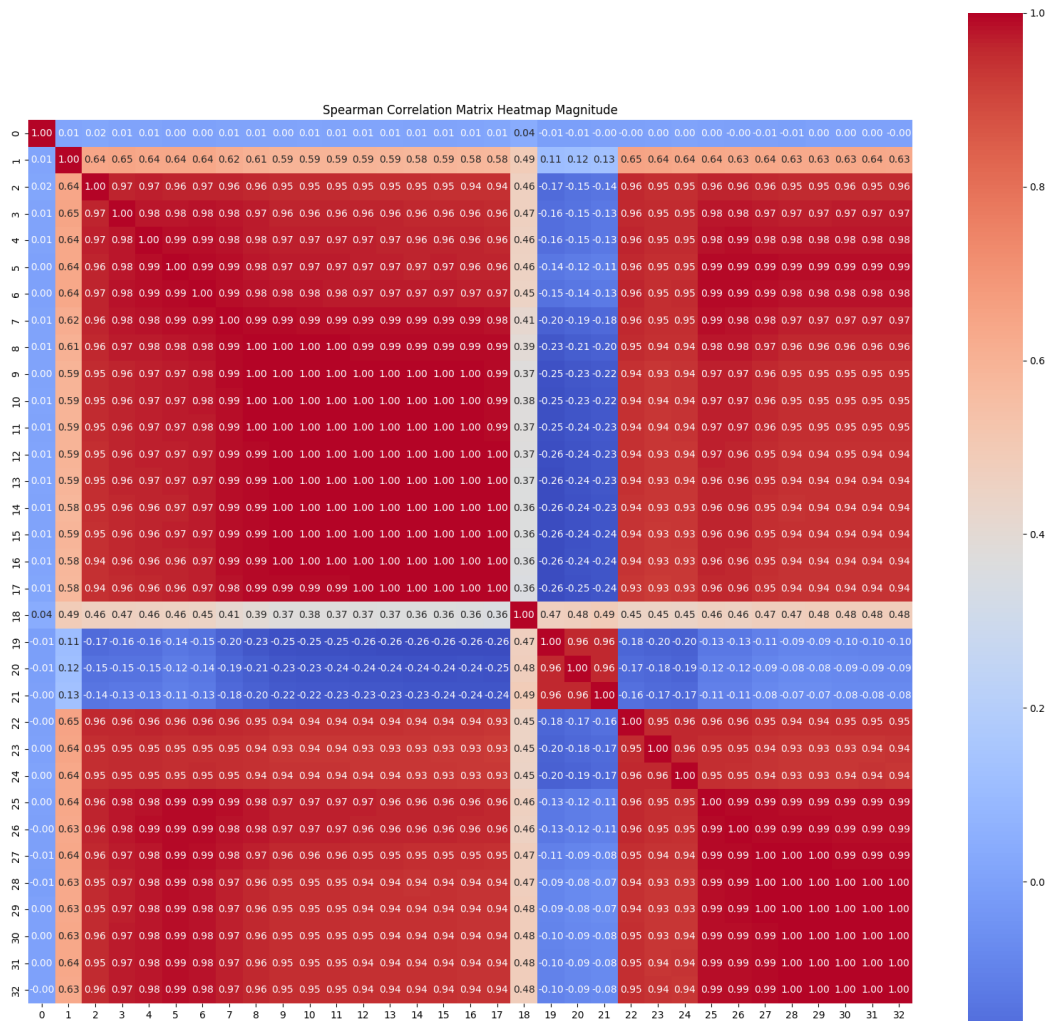


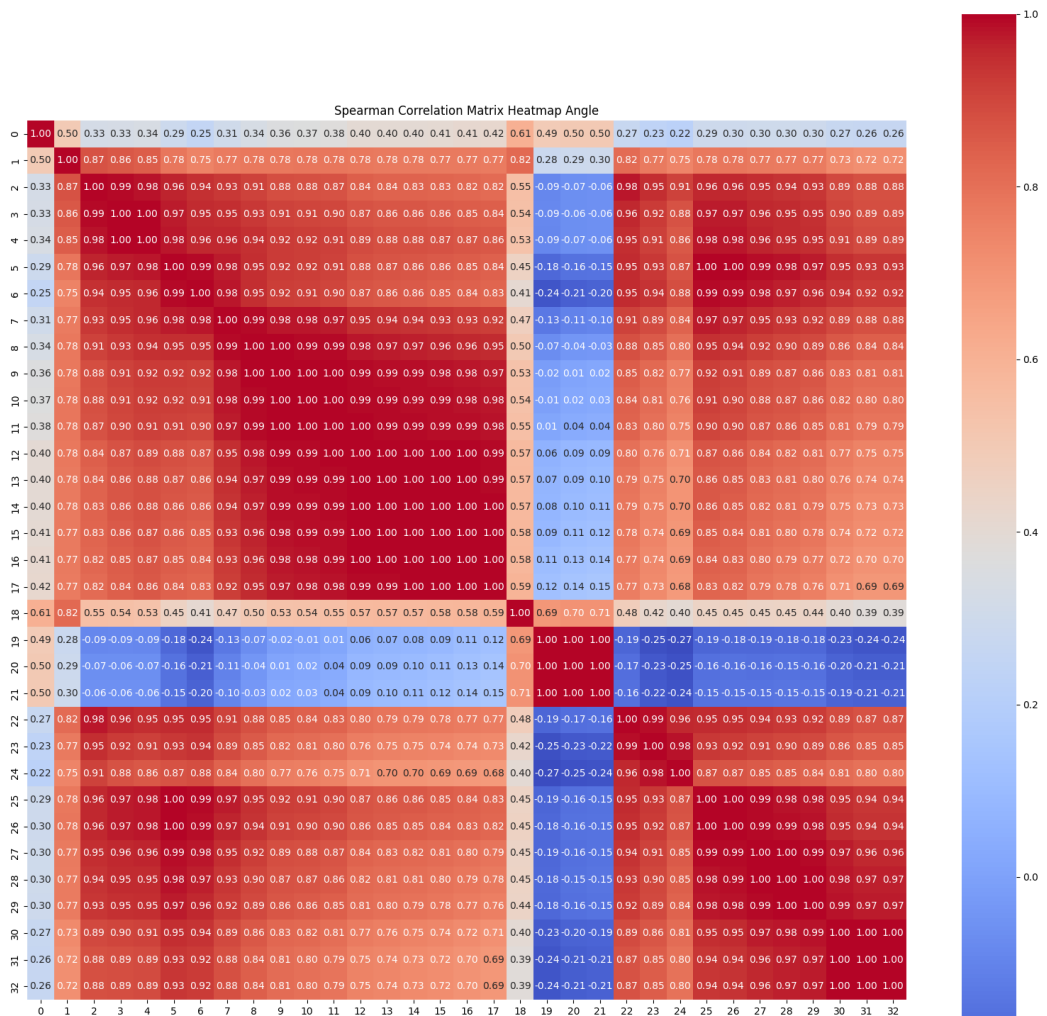
Παρακάτω παρουσιάζονται οι δύο πίνακες spearman που υπολογίστηκαν με βάση την προσέγγιση pearson και αξιολογήθηκαν, με τη μορφή heatmap, δηλαδή, το χρώμα κάθε κελιού ανάλογα με την κλίμακα υποδεικνύει την ισχύ και την κατεύθυνση της συσχέτισης μεταξύ των μεταβλητών που αντιπροσωπεύει





Κατόπιν, παρατίθενται τα heatmaps των πινάκων spearman όπως υπολογίστηκαν απευθείας, χωρίς την χρήση του pearson correlation coefficient ως προσέγγιση, για τα πλάτη και τις γωνίες αντίστοιχα:





Παρατηρείται πως υπάρχει πολύ μικρή διαφορά ανάμεσα στα δύο σχήματα και για τα δεδομένα των γωνιών και για τα δεδομένα των μέτρων της τάσης, κι έτσι και ανάμεσα στις δύο αυτές μεθόδους υπολογισμού του πίνακα spearman. Αυτό υποδηλώνει ότι οι μεταβλητές στο σύνολο των δεδομένων τόσο για τις γωνίες φάσης, όσο και για τα μέτρα τάσης έχουν μονοτονική σχέση, καθώς μέσω της χρήσης του pearson correlation coefficient μπορεί να προσεγγιστεί το spearman correlation coefficient με επιτυχία. Η διαγώνιος και στους τέσσερις πίνακες είναι παντού κόκκινη καθώς κάθε μεταβλητή έχει τέλεια μονοτονική σχέση με τον εαυτό της, ενώ οι γραμμές και στήλες που αποτελούνται από τις συσχετίσεις των μεταβλητών 19, 20 και 21 - (φαίνονται σαν σχηματίζουν μπλε σταυρούς στο

heatmap) - και στους τέσσερις πίνακες είναι παντού μπλε καθώς οι ζυγοί που αντιστοιχούν σε αυτές αποτελούν δικό τους χωριστό κλάδο στο δίκτυο για την τοπολογία που εξετάζεται τώρα (χωρίς κανέναν από τους επιπλέον διακόπτες κλειστό).

Στη συνέχεια παρατίθενται μερικά αποτελέσματα, αρχικά από την δοκιμή υπερπαραμέτρων και την υλοποίηση του αλγορίθμου:

Selected Features: ([7, 27, 17, 3, 13, 23, 21, 1, 20, 18, 6, 24, 19, 0, 22, 12], Accuracy: 0.5591516088813323) Layers: 128 96 64 32 8 epochs:10 batch size: 32

Μετρώντας λανθασμένα ως accuracy μόνο το mean absolute error. Σε αυτό το στάδιο ή έξοδος του μοντέλου ήταν μόνο μία. Στην συνέχεια, δοκιμάστηκαν μοντέλα με δύο εξόδους, μία για κάθε μέγεθος που υπολογίζεται στην SE.

Selected Features: ([6, 18, 20], [23, 20, 7], 0.4086719523929153, 8.138451198481215) Layers: 128 96 - branches (64 32 8) epochs:10 batch size: 32

Selected Features: ([6, 18, 20], [23, 20, 7], 0.4084653510411638, 6.275189845731085) extra shared layer

Selected Features: ([6, 18, 20], [23, 20, 7], 0.3726369268235612, 6.190544660199996) 2 extra shared layers

Selected Features: ([6, 18, 20], [23, 20, 7], 0.37452016149014755, 6.1582334844546) extra layer 64 neurons in every branch

Selected Features: ([5, 13, 20, 18], [7, 20, 24, 23], 0.21459017873217762, 0.5585600236971484) epochs 150 batch size 32

Selected Features: ([5, 13, 20, 18], [7, 20, 24, 23], 0.19107387021087308, 0.40595435583443373) epochs 500 batch size 16

Selected Features: ([6, 18, 20, 3], [23, 20, 7, 24], 0.20917877894594236, 0.5306709296937591) epochs 550 batch size 16

Όλες όμως οι παραπάνω προσεγγίσεις έγιναν χωρίς να θεωρηθούν ως δεδομένα και αυτά από τα SMDs που επιλέγονται στο πρώτο μέρος του αλγορίθμου. Έτσι, θεωρούμε αρχικά πως έχουν επιλεγθεί από τους πίνακες του ρεύματος τα χαρακτηριστικά που αντιστοιχούν στις στήλες: 32, 10, 34, 29. Επίσης, παρατηρήθηκε πως αν τα δεδομένα των γωνιών κλιμακωθούν με βάση τον Standard Scaler ξανά, η ακρίβεια βελτιώνεται. Για τα δεδομένα των μέτρων τάσης δεν χρειάζεται να γίνει αυτή η διαδικασία καθώς είναι σε ανά μονάδα ήδη. Έτσι, αφού προστέθηκαν, στους παρακάτω πίνακες απεικονίζονται τα πρώτα αποτελέσματα μετά τις δοκιμές υπεραπαραμέτρων:

Επιλεγμένα Χαρακτηριστικά	MAE Γωνιών	MAPE Μέτρων	Κρυφά Στρώματα	Κλαδιά	Epochs
[21, 7], [3, 20, 12, 1, 18, 29]	0.11579362111531	0.32243579535962	128 96 64	64 - 32	1000
[26, 20, 11, 1], [7, 20, 18, 1]	0.59433379767284	2.20946349614568	200 200 200 200	64 - 32	1000
[8], [26]	0.13528407020900	0.23079927895398	64 128 256	256 512	1000
[8], [4, 20]	0.11465436536579	0.29344252407654	64 128 256	256 512	1000
[21, 7], [4, 20]	0.13886619056762	0.27426037621439	32 64 128	128 256	1000
[21, 7], [4, 20, 12, 1, 18]	0.10798617507406	0.29381450908084	32 64 128	128 256	1000

Όπως φαίνεται, τα αποτελέσματα ποικίλουν ακόμα και για μοντέλα με ίδιες υπερπαραμέτρους, λόγω του ότι κάθε φορά στον αλγόριθμο το train test split γινόταν κατά την διάρκεια του hierarchical clustering ξανά κι έτσι γινόταν εισαγωγή ενός βαθμού τυχειότητας που δεν θα έπρεπε να υπάρχει. Έτσι, όταν ξεπεράστηκε αυτό το πρόβλημα:

Επιλεγμένα Χαρακτηριστικά	MAE Γωνιών	MAPE Μέτρων	Κρυφά Στρώματα	Κλαδιά	Epochs	
[8], [3, 20, 12, 20, 1, 18, 29]	0.12137955947922	0.34725300981877	32 64 128 256	256 512	1000	
[21, 7], [4, 20, 18, 20]	0.13293564290690	0.26901661723897	12 32 64 128	128 256	1000	
[7, 18, 20], [20, 4, 18]	0.16943446961767	0.50162032189008	16 32 64	64	500	
[7, 18, 20], [20, 4, 18]	0.14602205492570	0.73419807891308	16 32 64	64	1000	
[21, 7], [4, 20, 18, 20]	0.14303997900889	0.41552683546184	32 64 128	128	500	
[21, 7], [7, 20, 18, 1]	0.14536761528582	0.33514282209801	32 64 128	128	500	Μη επιλογή δεύτερο κοινού κόβου σε περίπτωση 0
[21, 7], [7, 20, 18, 1]	0.14340663242017	0.34613267944637	32 64 128	128	500	0.3 dropout layer between shared layers
[11, 20, 26], [4, 20, 18]	0.15405038878593	0.37677613587433	32 64 128	128	1000	0.3 dropout layer between shared layers
[11, 20, 26], [4, 20, 18]	0.16220303670287	0.32134476764504	32 64 128	128	1000	0.3 dropout layer between first 2 shared layers
[8], [4, 20, 12, 1, 18]	0.12741218501815	0.30693431395540	32 64 128	128	256	0.3 dropout layer between branch layers
[8], [20, 4, 18]	0.11469559785540	0.28291144599732	32 64 128 256	128 256	1000	
[11, 20, 26], [4, 20, 18]	0.18867902445659	0.43731143656214	32 64 128 256	128 256	1000	0.2 dropout between first 2 shared layers
[8], [25]	0.13067773610003	0.28325100259746	64 128 256	256 512	1000	
[8], [7, 20, 18, 1]	0.12877714108246	0.28270979455338	64 128 256	256 512	1000	

Είναι φανερό πως στην πλειοψηφία των περιπτώσεων, ο αλγόριθμος τερματίστηκε καθώς ξεπεράστηκε το κριτήριο του κόστους κι όχι επειδή επιτεύχθηκε η επιθυμητή ακρίβεια του μοντέλου παλινδρόμησης. Επίσης, το καλύτερο μοντέλο μέχρι στιγμή είναι το 64 128 256 shared layers και 256 512 branch layers με τα δύο Branches συμμετρικά. Δοκιμάστηκαν και μη συμμετρικά branches αλλά παρατηρήθηκε πως η ακρίβεια μειωνόταν και ο αριθμός των απαραίτητων SMDs γινόταν μεγαλύτερος.

Ωστόσο, μέχρι στιγμής είχε θεωρηθεί πως κάθε SMD που τοποθετείται με βάση το αποτέλεσμα της ιεραρχικής συσταδοποίησης για τον πίνακα με τα δεδομένα των γωνιών τάσης μετρά μόνο γωνία τάσης. Αντίστοιχα και για τα SMDs μέτρων τάσης. Όταν θεωρηθεί πως κάθε SMD μετρά και μέτρο και γωνία τάσης:

Επιλεγμένα Χαρακτηριστικά	MAE Γωνιών	MAPE Μέτρων	Κρυφά Στρώματα	Κλαδιά	Epochs	
[8], [25]	0.10953973811285	0.25848436644545	64 128 256	256 512	1000	no dropout
[8], [25]	0.11673848594012	0.26723498604113	64 128 256	256 512	1000	
[8], [4, 20]	0.10145280439400	0.24346608101974	64 128 256	256 512	1000	
[8], [25]	0.12001781279147	0.28666174692563	64 128 256	256 512	1000	dropout ανάμεσα στα branch layers
[8], [25]	0.11793921280987	0.26689036709879	64 128 256	256 512	1000	dropout ανάμεσα στα layers 64 128
[8], [7, 20, 18, 1]	0.10306982826415	0.28787789755801	64 128	128 256	1000	
[8] [4, 20]	0.10261828600145	0.31826099241417	64 128 256	256 512	1000	
[8] [20, 4, 18]	0.11385023966058	0.30216305102241	64 128 256	512 1024	1000	
[8], [25]	0.11957635153980	0.28777587721507	64 128 256	256	1000	

Στις παραπάνω δοκιμές παρατηρήθηκε πως τα dropout layers παίζουν πολύ μικρό ρόλο στην βελτίωση της ακρίβειας του μοντέλου ακόμη και όταν

τοποθετηθούν στα αρχικά στρώματα του νευρωνικού (οπότε και είναι φυσικό να το επηρεάζουν πιο πολύ). Αυτό συμβαίνει επειδή όπως θα δειχθεί παρακάτω παρατηρείται πολύ μικρό ποσοστό overfitting.

Στη συνέχεια εκτελέστηκαν δύο grid searches (αναζητήσεις πλέγματος) έτσι ώστε να βρεθούν οι καλύτερες υπερπαραμέτροι. Οι υπερπαραμέτροι που εξετάστηκαν φαίνονται στον παρακάτω πίνακα:

```
shared_layers_values = [1, 2, 3]
neurons_values = [32, 50, 64, 100, 128, 256, 512, 1024]
epochs_values = [100, 400, 500, 600, 750, 1000]
batch_size_values = [16, 32, 64, 128]
learning_rates = [0.01, 0.001, 0.0001]
```

Θεωρήθηκε πως τα δύο branches του δικτύου θα είναι συμμετρικά και πως θα αποτελούνται από δύο στρώματα, το καθένα με διπλάσιο αριθμό νευρώνων από το προηγούμενο, καθώς αυτό φάνηκε μέσα από τις δοκιμές με το 'χέρι' να αυξάνει την ακρίβεια και καθιστά την διαδικασία της ρύθμισης των υπερπαραμέτρων αρκετά πιο εύκολη. Ως shared layer values ορίστηκε ο αριθμός των κοινών στρωμάτων, ως neuron value ο αριθμός των νευρώνων του πρώτου κοινού στρώματος, ο οποίος θα είναι διπλάσιος στο επόμενο στρώμα και ούτω καθεξής.

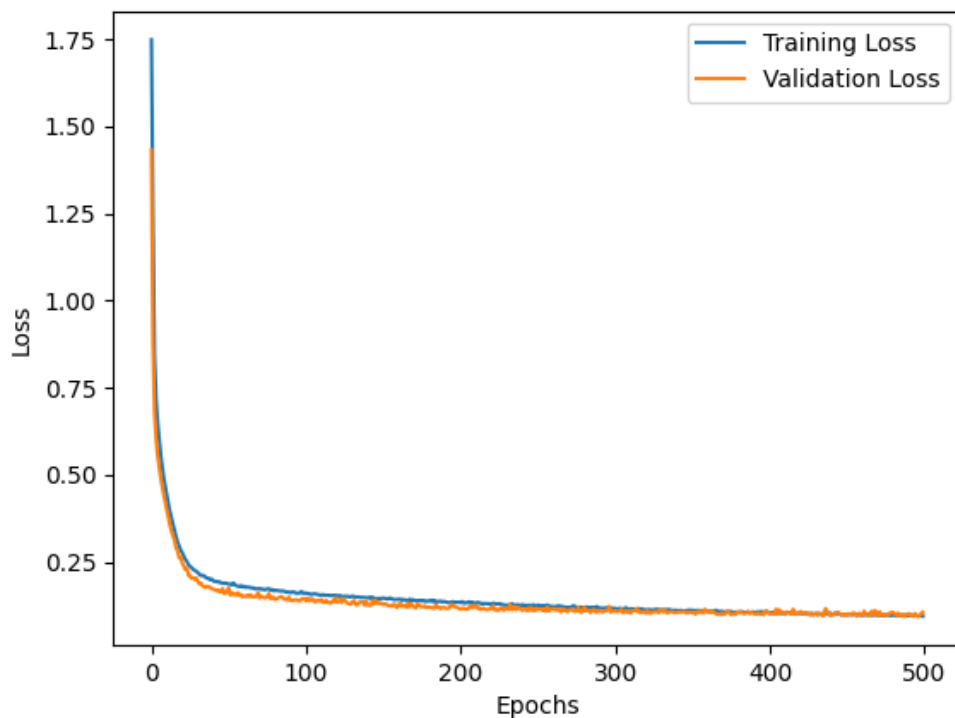
Για τον σκοπό των grid search θεωρήθηκε πως χρειάζονται μόνο δύο επιπλέον SMDs για DSSE, καθώς αυτά αποτελούν τον ελάχιστο αριθμό που απαιτείται, αφού το μοντέλο δεν εμφανίζει την επαρκή ακρίβεια με τα υπάρχοντα SMDs που τοποθετήθηκαν για TI στο πρώτο μέρος του αλγορίθμου. Παρακάτω φαίνονται μερικά από τα αποτελέσματα που παρουσίαζε το grid search όσο εκτελούνταν:

```
Shared Layers: 1, Neurons: 1024, Epochs: 200, Batch_Size: 64, Learning
Rate: 0.0001, MAE: 0.17989196878105274, MAPE: 1.6248742811989814
4/4 [=====] - 0s 9ms/step
Shared Layers: 1, Neurons: 1024, Epochs: 400, Batch_Size: 16, Learning
Rate: 0.01, MAE: 0.22509619633686553, MAPE: 0.6422537070901976
4/4 [=====] - 0s 15ms/step
Shared Layers: 1, Neurons: 1024, Epochs: 400, Batch_Size: 16, Learning
Rate: 0.001, MAE: 0.1668811859209273, MAPE: 0.6197004770403823
4/4 [=====] - 0s 13ms/step
Shared Layers: 1, Neurons: 1024, Epochs: 400, Batch_Size: 16, Learning
Rate: 0.0001, MAE: 0.1622994325405043, MAPE: 0.5012238920678094
4/4 [=====] - 0s 11ms/step
Shared Layers: 1, Neurons: 1024, Epochs: 400, Batch_Size: 32, Learning
Rate: 0.01, MAE: 0.22459532688656672, MAPE: 1.4570377752206078
4/4 [=====] - 0s 7ms/step
Shared Layers: 1, Neurons: 1024, Epochs: 400, Batch_Size: 32, Learning
Rate: 0.001, MAE: 0.156824039541243, MAPE: 0.6411913534444812
4/4 [=====] - 0s 8ms/step
Shared Layers: 1, Neurons: 1024, Epochs: 400, Batch_Size: 32, Learning
Rate: 0.0001, MAE: 0.16328032944177184, MAPE: 0.4931663522479102
4/4 [=====] - 0s 7ms/step
```


Shared Layers: 1, Neurons: 1024, Epochs: 400, Batch_Size: 64, Learning Rate: 0.01, MAE: 0.19550831308959715, MAPE: 0.9100210798686252
 4/4 [=====] - 0s 11ms/step
 Shared Layers: 1, Neurons: 1024, Epochs: 400, Batch_Size: 64, Learning Rate: 0.001, MAE: 0.1556311108529801, MAPE: 0.37517500258519265
 4/4 [=====] - 0s 8ms/step
 Shared Layers: 1, Neurons: 1024, Epochs: 400, Batch_Size: 64, Learning Rate: 0.0001, MAE: 0.17386460270847318, MAPE: 1.6362836007163442
 4/4 [=====] - 0s 9ms/step
 Shared Layers: 1, Neurons: 1024, Epochs: 500, Batch_Size: 16, Learning Rate: 0.01, MAE: 0.21729500084359057, MAPE: 0.7288496809502314

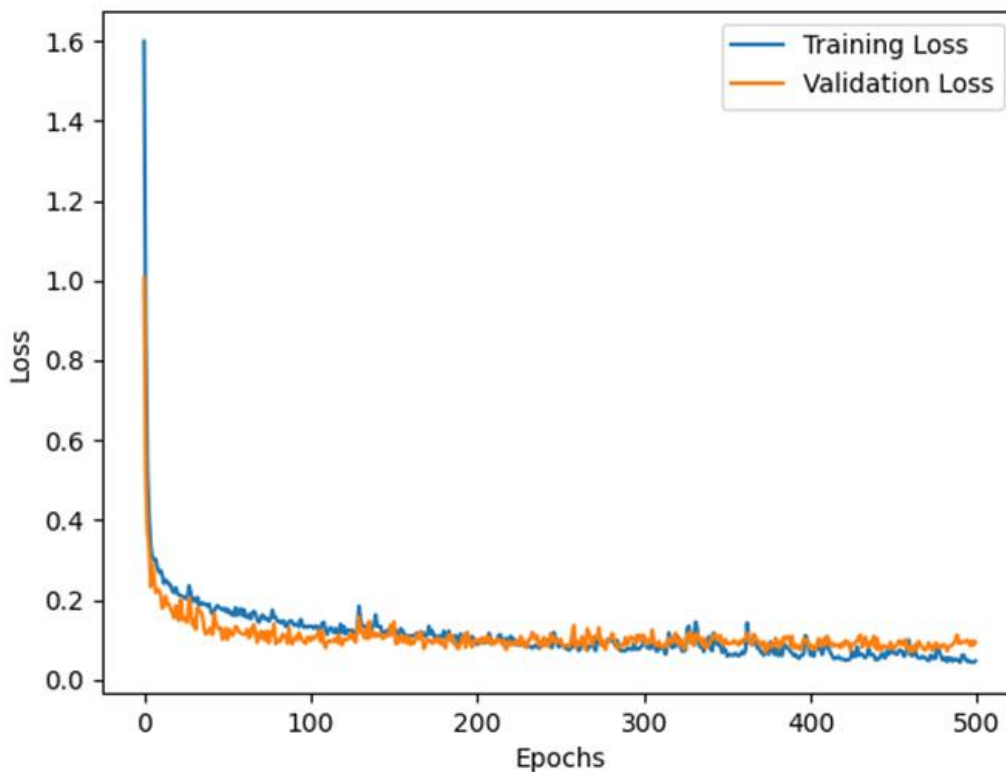
Ωστόσο, σε αυτό το πρώτο grid search ο χώρος αυτός των υπεραπαραμέτρων αποδείχτηκε πολύ μεγάλος με αποτέλεσμα η διαδικασία να αποβεί τόσο χρονοβόρα που αναγκαστικά διακόπη μετά από τέσσερις μέρες όταν είχε ολοκληρωθεί μόλις το 1/3. Στη συνέχεια εξετάστηκε ο αριθμός των κοινών στρωμάτων που βελτιώνει την ακρίβεια, ξανά με δοκιμές με το ‘χέρι’.

Οι δοκιμές με το χέρι εστίασαν κυρίως στον αριθμό των κοινών κρυφών στρωμάτων καθώς αφού οι αρχικές πιθανές τιμές ήταν μόνο τρεις, με την επιλογή της μίας η διάρκεια του grid search θα μειωνόταν κατά 67%. Αρχικά, για ένα θεωρήθηκε ένα δίκτυο με μόνο ένα κρυφό κοινό στρώμα, για το οποίο τέθηκαν 500 epochs για την εκπαίδευση έτσι ώστε να μειωθεί ο χρόνος της, καθώς σε αυτό το στάδιο δεν ψάχνουμε βέλτιστα αποτελέσματα. Στο παρακάτω διάγραμμα φαίνεται το mae του μοντέλου:



Παρατηρείται πως οι γραμμές του training loss και του validation loss συμβαδίζουν, οπότε δεν υπάρχει ούτε overfitting (υπερεκπαίδευση) ούτε underfitting (υποεκπαίδευση). Ωστόσο, σε μερικές άλλες πιο πρώιμες δοκιμές παρατηρήθηκε λίγο underfitting σε μοντέλα με μόνο ένα κρυφό στρώμα, δηλαδή το μοντέλο ήταν πολύ απλό για να αποτυπώσει την υποκείμενη δομή των δεδομένων και αδυνατούσε να μάθει αποτελεσματικά από τα δεδομένα εκπαίδευσης. Επίσης, χρειάζεται 5 SMDs ([11, 20, 26], [4, 20, 18]) έτσι ώστε να φτάσει στην επιθυμητή ακρίβεια.

Στη συνέχεια εξετάστηκαν τα μοντέλα με δύο και τρία κρυφά κοινά στρώματα, τα οποία χρειάζονται μόνο δύο SMDs για να επιτύχουν ακριβή DSSE. Στο παρακάτω διάγραμμα φαίνεται πως παρουσιάζεται ένας μικρός βαθμός overfitting και για τις δύο περιπτώσεις:



Για την μείωση του overfitting δοκιμάστηκε η εισαγωγή περισσότερων στρωμάτων dropout σε διάφορες θέσεις του δικτύου, όπως για παράδειγμα ανάμεσα στα στρώματα των κλάδων ή ανάμεσα σε κάθε κοινό στρώμα αλλά όπως φάνηκε από τους πίνακες και πριν αυτό είχε ως αποτέλεσμα να μειωθεί η ακρίβεια. Επίσης, δοκιμάστηκε η εισαγωγή L2 regularization, μια τεχνική σύμφωνα με την οποία ένας επιπλέον όρος προστίθεται στη συνάρτηση απώλειας, 'τιμωρώντας' τα μεγάλα βάρη στο δίκτυο. Αυτός ο όρος είναι ανάλογος του τετραγώνου των βαρών. Ως αποτέλεσμα, οι μαθημένες παράμετροι του μοντέλου τείνουν να είναι πιο κατανομημένες και λιγότερο επιρρεπείς σε ακραίες τιμές, μειώνοντας τον κίνδυνο υπερπροσαρμογής. Ωστόσο η ακρίβεια μειωνόταν σημαντικά με την προσθήκη regularization

στα κρυφά στρώματα, ανεξαρτήτως της τιμής του regularization strength (λ). Έτσι κι αυτή η προσθήκη απορρίφθηκε. Τέλος, δοκιμάστηκε και το early stopping, δηλαδή η διακοπή της εκπαίδευσης μόλις η απόδοση αρχίσει να υποβαθμίζεται. Κατά τη διάρκεια της εκπαίδευσης ενός μοντέλου μηχανικής μάθησης, η απόδοση σε ένα validation set (δεδομένα επικύρωσης) παρακολουθείται συνήθως σε τακτά χρονικά διαστήματα, όπως μετά από κάθε εποχή. Εάν η μετρική απόδοσης (π.χ. mae ή ακρίβεια) σταματήσει να βελτιώνεται ή αρχίσει να επιδεινώνεται για έναν ορισμένο αριθμό διαδοχικών εποχών, η διαδικασία εκπαίδευσης διακόπτεται και διατηρούνται οι παράμετροι του μοντέλου στο σημείο της καλύτερης απόδοσης. Ο αριθμός των διαδοχικών εποχών για τις οποίες η στασιμότητα ή η μείωση της ακρίβειας που μπορεί να ανεχθεί θεωρείται χωριστή υπερπαραμέτρος, η οποία ονομάζεται παράμετρος patience. Ωστόσο, όταν εφαρμόστηκε αυτή η τεχνική μέσω της προσθήκης της συνθήκης early stopping κατά την εκπαίδευση του μοντέλου, η ακρίβεια μειωνόταν, ακόμα και όταν ο αριθμός των εποχών γινόταν 1000 και το patience 100 ή και 200. Εν κατακλείδι, το overfitting ήταν πολύ μικρού βαθμού, οπότε η μείωση του ουσιαστικά δεν μπορούσε να γίνει σε τόσο μικρό επίπεδο που να μην προκαλέσει underfitting κατά την εκπαίδευση του μοντέλου. Στο τέλος, εφαρμόστηκε early stopping με παράμετρο patience ίση με 200 η οποία ουσιαστικά δεν επηρεάζει το μοντέλο και τοποθετήθηκε ένα στρώμα dropout με τιμή 0.3 ακριβώς μετά το τελευταίο κοινό στρώμα, έτσι ώστε να βρίσκεται προς την αρχή του μοντέλου ώστε να επηρεάζει περισσότερο την εκπαίδευση, αλλά όχι τόσο όσο να μειώνει την ακρίβεια.

Έτσι, μετά από αυτές τις δοκιμές εκτελέστηκε ένα δεύτερο grid search με τις εξής παραμέτρους ο συνδυασμός των οποίων οδηγεί σε πολύ λιγότερες περιπτώσεις:

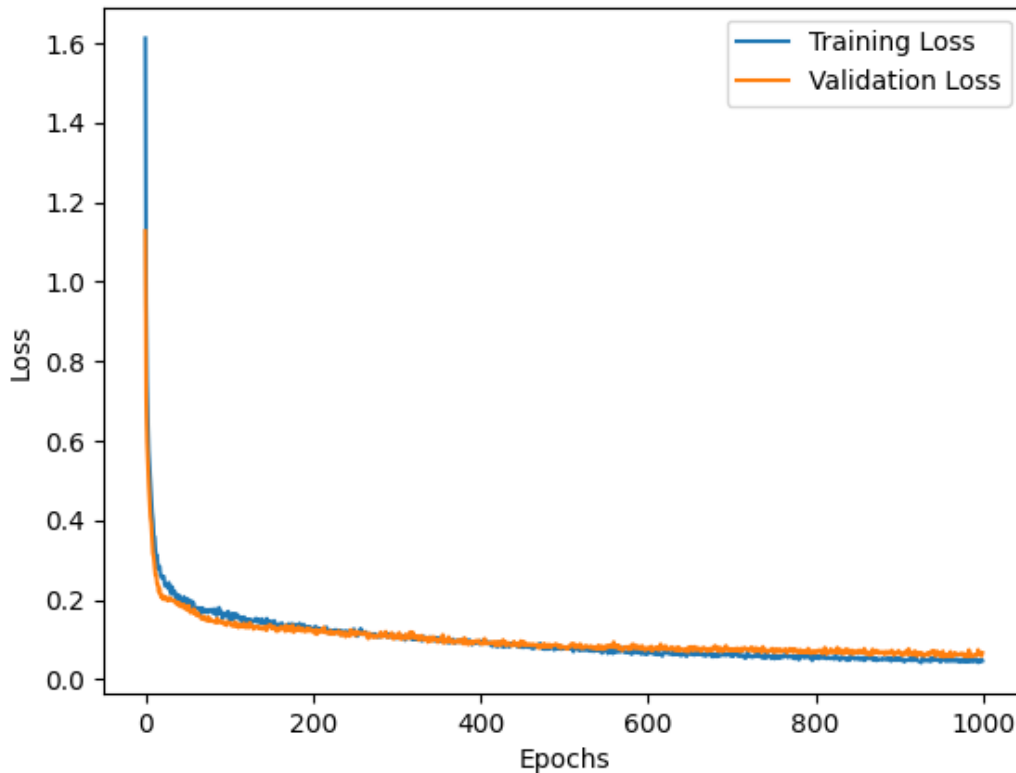
```
shared_layers_values = [2]
neurons_values = [32, 50, 64, 100, 128]
epochs_values = [500, 600, 750, 1000]
batch_size_values = [16, 32, 64]
learning_rates = [0.001, 0.0001]
```

Ξανά, θεωρήθηκε πως τα δύο branches του δικτύου θα είναι συμμετρικά και πως θα αποτελούνται από δύο στρώματα, το καθένα με διπλάσιο αριθμό νευρώνων από το προηγούμενο, καθώς αυτό φάνηκε μέσα από τις δοκιμές με το 'χέρι' να αυξάνει την ακρίβεια και καθιστά την διαδικασία της ρύθμισης των υπερπαραμέτρων αρκετά πιο εύκολη. Ως shared layer values ορίστηκε ο αριθμός των κοινών στρωμάτων, ως neuron value ο αριθμός των νευρώνων του πρώτου κοινού στρώματος, ο οποίος θα είναι διπλάσιος στο επόμενο στρώμα και ούτω καθεξής. Επίσης ξανά θεωρήθηκε πως για DSSE χρειάζονται μόνο δύο επιπλέον SMDs, στις θέσεις 8 και 26.

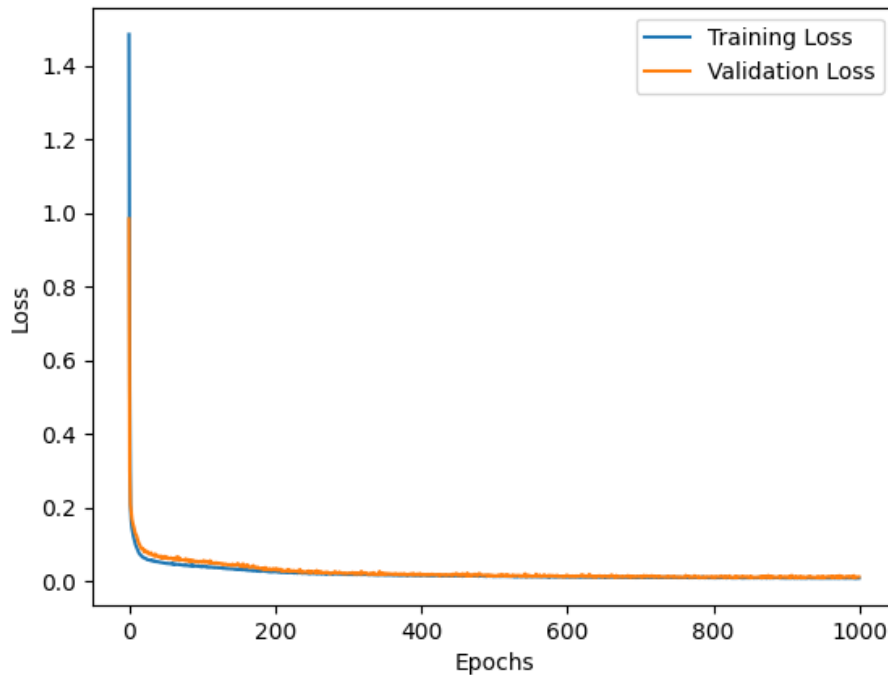
Αυτή η αναζήτηση πλέγματος έδωσε 33 διαφορετικούς συνδυασμούς υπερπαραμέτρων, αν με τις οποίες κατασκευαστούν νευρωνικά δίκτυα, μετά την εκπαίδευση τους η ακρίβεια ξεπερνά τα όρια που έχουν τεθεί ($MAE < 0,15$ και $MAPE < 0.3$ %). Ο καλύτερος από αυτούς, δηλαδή αυτός που το άθροισμα MAE και MAPE που δίνει το μοντέλο είναι μικρότερος είναι ο:

Best Shared Layers: 2
Best Neurons: 64
Best Epochs: 1000
Best Batch Size: 16
Best Learning Rate: 0.0001
Best combined: 0.5124444728295003
Best MAE: 0.128487067429024
Best MAPE: 0.1919787027002382

Παρακάτω φαίνεται το training και validation loss κατά την διάρκεια της εκπαίδευσης:



Παρατηρείται πως το overfitting είναι σχεδόν ανύπαρκτο, παρά τον αρκετά μεγάλο αριθμό των εποχών που έχουν επιλεχθεί. Ωστόσο, αν αφαιρεθεί το dropout στρώμα που είναι ουσιαστικά και το τελευταίο κοινό, το μοντέλο αυτό με τις υπεραπαραμέτρους που βρέθηκαν ως βέλτιστες θα εμφανίσει ακρίβεια: MAE = 0.05436857992950348 για τις γωνίες φάσης τάσης και MAPE = 0.2029360826588961 (*100%) για τα μέτρα τάσης. Το διάγραμμα τώρα που απεικονίζει τα losses κατά το training και το validation:



Παρατηρείται πως τώρα το overfitting είναι πλήρως ανύπαρκτο οπότε η προηγούμενη προσέγγιση με το στρώμα dropout ήταν λανθασμένη.

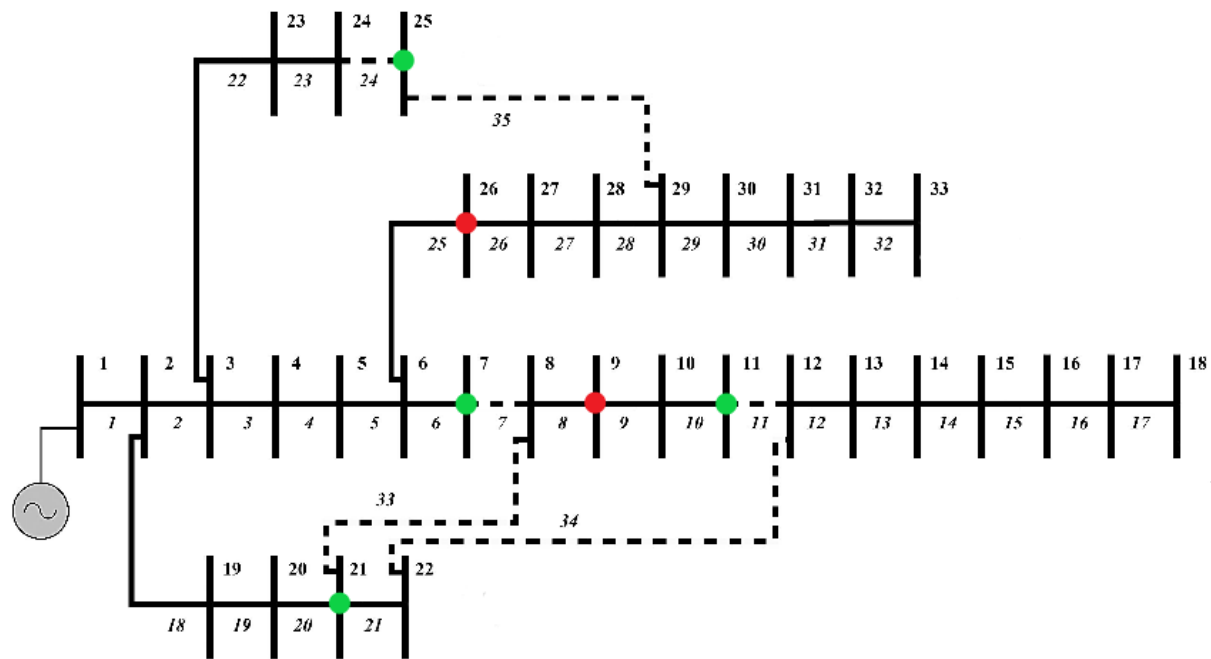
Τα αποτελέσματα αυτού του μοντέλου αν αντί ο κοινός κόμβος να επιλέγεται με βάση την μέση τιμή συσχέτισης που περιέχει κάθε στήλη (χαρακτηριστικό), να επιλέγεται με το άθροισμα των συσχετίσεων σύμφωνα με το κριτήριο centrality, θα είναι ίδια, MAE = 0.05792129258031085 και MAPE = 0.287299734201535, καθώς ξανά επιλέγονται οι κόμβοι 8 και 25.

Οι διάφορες εναλλακτικές που εφαρμόστηκαν παραπάνω σε περίπτωση που επιλεγεί το 0 ως κοινός κόμβος μετά το Clustering δεν έχουν σημασία, καθώς δεν επιλέγεται τώρα το 0 ως κοινός κόμβος στο παράδειγμα μας.

Τα αποτελέσματα αυτού του μοντέλου αν αντί για Agglomerative Clustering με απόσταση ward μεταξύ των συστάδων χρησιμοποιηθεί KMeans Clustering με τον αλγόριθμο κατά τα άλλα να παραμένει ίδιο θα είναι: MAE = 0.058821333139785444, MAPE = 0.23997109343651765

Ουσιαστικά επιλέγονται ξανά κι εδώ τα χαρακτηριστικά 8 και 25 (ζυγοί 9 και 26), οπότε αυτή η μέθοδος για το παράδειγμα μας δεν εμφανίζει καμία διαφορά. Η διαφορά της απόδοσης του μοντέλου σε αυτό το παράδειγμα οφείλεται στον βαθμό τυχειότητας που περιέχει το train test split εγγενώς σαν μέθοδος.

Συνολικά, για ΤΙ τοποθετούνται SMDs στις θέσεις 32, 10, 34 και 6 και για DSSE στις θέσεις 8 και 25. Η ακρίβεια του νευρωνικού δικτύου που εκτελεί ΤΙ είναι 96,625 % ενώ του DSSE είναι 0.05436857992950348 μέσω σφάλμα για τις γωνίες φάσης τάσης και 0.2029360826588961 για τα μέτρα φάσης τάσης. Στο παρακάτω σχήμα φαίνονται οι θέσεις που επιλέχθηκαν για τοποθέτηση SMDs, με πράσινο για ΤΙ και με κόκκινο για DSSE:



με:



Generator



Bus



Fixed Branch



Switchable Branch

Αποτελέσματα Transfer Learning

Ως προς το transfer learning, στους παρακάτω πίνακες παρατίθενται τα δεδομένα των δύο μεθόδων και πόσο χρόνο παίρνει η εκπαίδευση της δεύτερης τοπολογίας για την καθεμία.

Αρχικά, μέσω της μεθόδου fine tuning, κατά την οποία η εκπαίδευση όταν αλλάξει η τοπολογία πραγματοποιείται ξανά με τα δεδομένα που προκύπτουν από την νέα τοπολογία, όμως με τα βάρη του νευρωνικού δικτύου αρχικοποιημένα από την προηγούμενη εκπαίδευση, έδωσε τα εξής αποτελέσματα:

	MAE για γωνία φάσης	MAPE για πλάτος	Χρόνος Εκπαίδευσης (s)
Αρχική τοπολογία	0.074380348	0.130255123	33.88485909
Δεύτερη τοπολογία	0.19186354	0.148508874	30.94050598
Τρίτη τοπολογία	0.065961739	0.080196201	30.67676902
Τέταρτη τοπολογία	0.053123306	0.150040422	31.80428624
Πέμπτη τοπολογία	0.204055239	0.161878914	31.75328541
Έκτη τοπολογία	0.082896134	0.14474521	30.72022319
Έβδομη τοπολογία	0.061126164	0.099223597	31.33136415
Όγδοη τοπολογία	0.054714343	0.102464397	32.39788055

Το νευρωνικό δίκτυο που χρησιμοποιήθηκε είναι αυτό που βρέθηκε ως βέλτιστο μέσω της διαδικασίας που αναλύθηκε στο κεφάλαιο που αναφέρεται στη ρύθμιση των νευρωνικών δικτύων.

Μέσω της τεχνικής parameter-sharing, κατά την οποία ουσιαστικά εκπαιδεύεται ξανά το δίκτυο από την αρχή, κρατώντας όμως παγωμένα τα κοινά κρυφά στρώματα, τα αποτελέσματα είναι τα εξής:

	MAE για γωνία φάσης	MAPE για μέτρο	Χρόνος Εκπαίδευσης (s)
Αρχική τοπολογία	0.069603437	0.186088875	32.91384673
Τελική τοπολογία	0.218097966	0.676178248	26.38749218

Όπως ήταν αναμενόμενο, ο χρόνος εκπαίδευσης είναι μικρότερος καθώς ουσιαστικά εκπαιδεύονται ξανά μόνο δύο στρώματα και η ακρίβεια του μοντέλου είναι αισθητά μικρότερη (τα σφάλματα αυξάνονται σημαντικά), για την δεύτερη τοπολογία, καθώς παράμετροι των νευρωνικών δικτύων για DSSE δεν είναι σε μεγάλο βαθμό μεταβιβάσιμες για τις διαφορετικές τοπολογίες.

Ο χρόνος της επόμενης εκπαίδευσης παρατηρείται πως μειώθηκε λίγο (της τάξης του 10%) το οποίο σε μεγάλα και πολύπλοκα δίκτυα διανομής μπορεί να αποβεί σημαντικό πλεονέκτημα. Η ακρίβεια της εκτίμησης κατάστασης ουσιαστικά παραμένει συνεχώς μέσα στα όρια που τέθηκαν εκτός μερικών περιπτώσεων, ωστόσο παρατηρείται πως όταν οι τοπολογίες διαφέρουν μεταξύ τους, όταν π.χ. ένα διαφορετικό ζευγάρι διακοπών κλείνει ή ανοίγει απ'ότι στην προηγούμενη τοπολογία (όπως συμβαίνει για παράδειγμα μεταξύ των τοπολογιών 1 και 2) η ακρίβεια μειώνεται λίγο. Αυτό είναι λογικό, καθώς το transfer learning παρουσιάζει καλύτερα αποτελέσματα όσο πιο πολύ δύο τοπολογίες μοιάζουν μεταξύ τους.

Επέκταση: Random Forest

Το σημαντικότερο μειονέκτημα της μεθόδου που περιγράφηκε παραπάνω που συνδυάζει sequential forward selection και hierarchical clustering για την επιλογή χαρακτηριστικών είναι πως αποτελεί μια ιδιαίτερα χρονοβόρα διαδικασία, ιδιαίτερα για μεγάλα και πολύπλοκα δίκτυα και την sfs. Ενδεικτικά στο τέλος αυτής της ενότητας παρατίθενται οι χρόνοι των δύο τμημάτων του αλγορίθμου έτσι όπως υπολογίστηκαν χρησιμοποιώντας την κλάση Current Time από την ενότητα time της pythοn, η οποία μετράει τον χρόνο του προβλήματος από μία προκαθορισμένη αρχή έως το τέλος.

Έτσι, διερευνήθηκε και η τεχνική random forest, για την επιλογή χαρακτηριστικών. Το Random Forest είναι ένας αλγόριθμος μηχανικής μάθησης που χρησιμοποιείται τόσο για προβλήματα ταξινόμησης όσο και για προβλήματα παλινδρόμησης. Για τα πρώτα, η έξοδος είναι η κλάση που επιλέγεται από τα περισσότερα δέντρα, ενώ για τα δεύτερα επιστρέφεται ο μέσος όρος ή η μέση πρόβλεψη των μεμονωμένων δέντρων.

Ανήκει στις μεθόδους μάθησης συνόλου, οι οποίες περιλαμβάνουν το συνδυασμό πολλαπλών μεμονωμένων μοντέλων για την παραγωγή μιας πιο ακριβούς και αξιόπιστης πρόβλεψης. Η βασική ιδέα πίσω από το Random Forest είναι η δημιουργία ενός μεγάλου αριθμού δέντρων απόφασης κατά τη διάρκεια της εκπαίδευσης και, στη συνέχεια, ο συνδυασμός των αποτελεσμάτων τους για την πραγματοποίηση προβλέψεων. Τα δέντρα αποφάσεων αποτελούν έναν αλγόριθμο μάθησης με επίβλεψη που μοντελοποιεί τις αποφάσεις με βάση μια δομή που μοιάζει με δέντρο, όπου κάθε εσωτερικός κόμβος αντιπροσωπεύει μια "απόφαση" με βάση ένα χαρακτηριστικό και κάθε κόμβος φύλλου αντιπροσωπεύει το αποτέλεσμα ή την πρόβλεψη. Το δέντρο αποφάσεων ξεκινά από τον κόμβο-ρίζα και χωρίζει αναδρομικά τα δεδομένα σε υποσύνολα με βάση τις τιμές των χαρακτηριστικών. Σε κάθε κόμβο, ο αλγόριθμος επιλέγει το χαρακτηριστικό που χωρίζει καλύτερα τα δεδομένα σε ομοιογενή υποσύνολα, με την διαδικασία αυτή να συνεχίζεται έως ότου ικανοποιηθεί ένα από τα κριτήρια διακοπής.

Αφού κατασκευαστεί το δέντρο αποφάσεων, μπορεί να χρησιμοποιηθεί για την πραγματοποίηση προβλέψεων σε νέα δεδομένα, διασχίζοντας το από τον κόμβο ρίζας μέχρι έναν κόμβο φύλλου κι ακολουθώντας τις αποφάσεις που λαμβάνονται σε κάθε διάσπαση. Στα προβλήματα ταξινόμησης, η πρόβλεψη είναι συνήθως η πλειοψηφούσα κλάση στον κόμβο-φύλλο, ενώ για τα προβλήματα παλινδρόμησης είναι η μέση τιμή της μεταβλητής-στόχου στον κόμβο-φύλλο.

Κάθε δέντρο απόφασης σε ένα Random Forest κατασκευάζεται χρησιμοποιώντας ένα υποσύνολο των δεδομένων εκπαίδευσης και ένα υποσύνολο των χαρακτηριστικών. Αυτή η τυχαιότητα που εισάγεται στο μοντέλο βοηθά στη μείωση του overfitting και βελτιώνει την απόδοση γενίκευσης του αλγορίθμου. Επιπλέον, τα τυχαία αυτά δέντρα μπορούν να χειριστούν μεγάλο αριθμό χαρακτηριστικών εισόδου χωρίς κλιμάκωση (scaling) των δεδομένων και είναι ανθεκτικά σε ακραίες τιμές και θορυβώδη δεδομένα.

Ο αλγόριθμος ουσιαστικά αποτελείται από 4 κύρια στάδια:

1. Δειγματοληψία - Bootstrapped Sampling (Bagging):

Το Random Forest ξεκινά με τη δημιουργία πολλαπλών υποσυνόλων του αρχικού συνόλου δεδομένων με τυχαία επιλογή δειγμάτων με αντικατάσταση. Αυτό σημαίνει ότι κάθε δείγμα στο αρχικό σύνολο δεδομένων έχει ίσες πιθανότητες να επιλεγεί για συμπερίληψη σε κάθε υποσύνολο, και ορισμένα δείγματα μπορεί να εμφανίζονται περισσότερες από μία φορές σε ένα υποσύνολο, ενώ άλλα μπορεί να μην εμφανίζονται καθόλου. Επιτρέποντας διπλά δείγματα και τυχαία επιλογή, η δειγματοληψία αυτού του είδους (bootstrap) δημιουργεί ποικίλα υποσύνολα δεδομένων. Κάθε υποσύνολο αντιπροσωπεύει μια διαφορετική προοπτική του αρχικού συνόλου δεδομένων, καταγράφοντας διαφορετικά μοτίβα και θόρυβο που υπάρχουν στα δεδομένα. Τέλος, κάθε δείγμα έχει το ίδιο μέγεθος με το αρχικό σύνολο δεδομένων, αλλά λόγω της δειγματοληψίας με αντικατάσταση, μπορεί να περιέχει ελαφρώς διαφορετικές κατανομές των σημείων δεδομένων.

2. Τυχαία επιλογή χαρακτηριστικών:

Για κάθε δέντρο απόφασης στο random forest, σε κάθε κόμβο, ένα τυχαίο υποσύνολο χαρακτηριστικών εξετάζεται για το διαχωρισμό των δεδομένων. Το μέγεθος αυτού του υποσυνόλου είναι συνήθως μικρότερο από τον συνολικό αριθμό χαρακτηριστικών στο σύνολο δεδομένων. Με αυτό τον τρόπο, τα μεμονωμένα δέντρα απόφασης στο δάσος είναι διαφορετικά και δεν έχουν υψηλή συσχέτιση μεταξύ τους. Με την εξέταση αυτή τυχαίων υποσυνόλων χαρακτηριστικών σε κάθε κόμβο, αξιολογείται έμμεσα η σημασία των χαρακτηριστικών παρατηρώντας πόσο συμβάλλουν στη μείωση του σφάλματος στις διαχωριστικές ομάδες.

3. Κατασκευή δέντρων απόφασης:

Με τα υποσύνολα των δεδομένων μετά την δειγματοληψία από το πρώτο βήμα και τα τυχαία υποσύνολα χαρακτηριστικών, το Random Forest κατασκευάζει ανεξάρτητα πολλαπλά δέντρα απόφασης. Κάθε δέντρο απόφασης εκπαιδεύεται σε ένα από αυτά τα υποσύνολα δεδομένων, εξασφαλίζοντας ότι βλέπει μια ελαφρώς διαφορετική προοπτική των δεδομένων. Επίσης, κάθε δέντρο αναπτύσσεται στο μέγιστο δυνατό βαθμό χωρίς κλάδεμα, δηλαδή, χωρίς κανέναν περιορισμό. Με αυτόν τον τρόπο εξασφαλίζεται ότι κάθε δέντρο μπορεί να συλλάβει όσο το δυνατόν περισσότερες πληροφορίες από τα δεδομένα, μεγιστοποιώντας την ποικιλομορφία και την προβλεπτική δύναμη του συνόλου.

4. Ψηφοφορία (για ταξινόμηση) / Μέση τιμή (για παλινδρόμηση):

Για προβλήματα ταξινόμησης, η τελική πρόβλεψη γίνεται με τη βάση την πλειοψηφία των ψήφων μεταξύ όλων των δέντρων. Κάθε δέντρο απόφασης προβλέπει ανεξάρτητα την ετικέτα κλάσης για το δείγμα εισόδου με βάση τα παρεχόμενα χαρακτηριστικά. Αφού όλα τα δέντρα απόφασης έχουν κάνει τις προβλέψεις τους, η ετικέτα κλάσης που λαμβάνει τις περισσότερες ψήφους επιλέγεται ως η τελική πρόβλεψη για το σύνολο Random Forest. Για προβλήματα παλινδρόμησης, η τελική πρόβλεψη είναι ο μέσος όρος των

προβλέψεων που γίνονται από όλα τα δέντρα. Κάθε δέντρο απόφασης προβλέπει ανεξάρτητα μια συνεχή τιμή για το δείγμα εισόδου με βάση τα παρεχόμενα χαρακτηριστικά. Αυτές οι προβλέψεις μπορεί να διαφέρουν, καθώς κάθε δέντρο απόφασης έχει εκπαιδευτεί σε διαφορετικό υποσύνολο των δεδομένων και καταγράφει διαφορετικές πτυχές των υποκείμενων σχέσεων μεταξύ των χαρακτηριστικών και της μεταβλητής-στόχου. Αυτή η διαδικασία οδηγεί σε μια ενιαία προβλεπόμενη τιμή που αντιπροσωπεύει την γενική συναίνεση μεταξύ των μεμονωμένων δέντρων απόφασης στο σύνολο

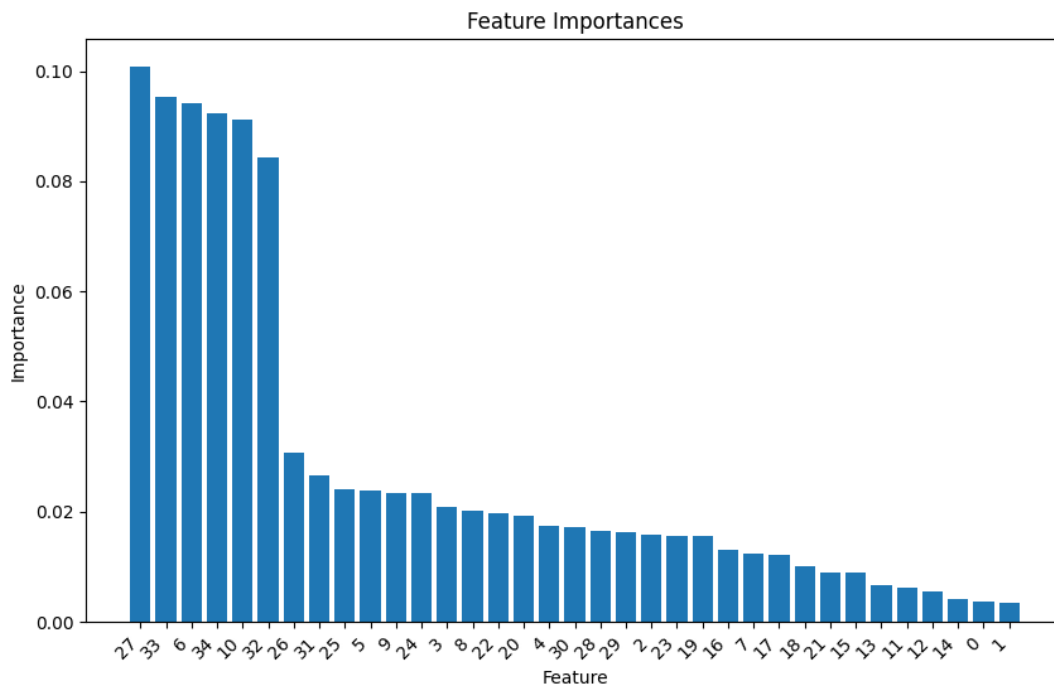
Τα κύρια πλεονεκτήματα της τεχνικής Random Forest αποτελούν η υψηλή ακρίβεια που παρέχεται τόσο σε προβλήματα ταξινόμησης όσο και σε παλινδρόμησης κι η ανθεκτικότητα στην υπερπροσαρμογή (overfitting). Με τη μέση τιμή των προβλέψεων πολλαπλών δέντρων απόφασης, το Random Forest μειώνει την υπερπροσαρμογή σε σύγκριση με τα μεμονωμένα δέντρα απόφασης. Επίσης, το Random Forest μπορεί να χειριστεί αποτελεσματικά μεγάλα σύνολα δεδομένων με μεγάλο αριθμό χαρακτηριστικών και να αντιμετωπίσει τις ακραίες τιμές και τις τιμές που λείπουν χωρίς την ανάγκη προεπεξεργασίας. Επιπλέον, καθώς το Random Forest παρέχει ένα μέτρο της σημασίας των χαρακτηριστικών, γίνεται πολύ πιο εύκολος ο προσδιορισμός των χαρακτηριστικών με τη μεγαλύτερη επιρροή κι έτσι και η διαδικασία της επιλογής χαρακτηριστικών. Τέλος, η εκπαίδευση κάθε δέντρου απόφασης σε ένα Random Forest μπορεί να γίνει ανεξάρτητα. Αυτό επιτρέπει ταχύτερη εκπαίδευση και σύγκλιση του αλγορίθμου.

Ο αλγόριθμος Random Forest υλοποιήθηκε μέσω της κλάσης RandomForestClassifier από την βιβλιοθήκη scikit-learn. Αυτή ουσιαστικά αποτελεί έναν έτοιμο προς χρήση ταξινομητή ο οποίος εκπαιδεύεται στα δεδομένα, στη συνέχεια προβλέπει με βάση το test set και εξάγει έτσι το πόσο σημαντικό είναι κάθε χαρακτηριστικό στην πρόβλεψη, με βάση την ιδιότητα `feature_importances_`. Για DSSE η μόνη διαφορά είναι πως καλείται η κλάση RandomForestRegressor, πάλι από την βιβλιοθήκη scikit-learn.

Κατόπιν, αναπτύχθηκαν δύο προσεγγίσεις. Η πρώτη, έτσι ώστε να παραμείνουμε σύμφωνοι με το πνεύμα του αλγορίθμου επιλογής SMDs που αναπτύχθηκε στα προηγούμενα κεφάλαια, βασίζεται στην ιδέα ξανά της σταδιακής αύξησης των χαρακτηριστικών μέχρις ότου το μοντέλο πρόβλεψης να φτάσει στην κατάλληλη ακρίβεια. Έτσι, αρχικά χρησιμοποιείται το χαρακτηριστικό που έχει κριθεί ως πιο σημαντικό (έχει την υψηλότερη `feature_importances_`) για την εκπαίδευση του μοντέλου ταξινόμησης. Αν η ακρίβεια του ξεπεράσει το επιθυμητό όριο, τότε ο αλγόριθμος σταματάει, αλλιώς προστίθεται το χαρακτηριστικό που θεωρείται από το Random Forest ως το δεύτερο πιο σημαντικό και ούτω καθεξής.

Ωστόσο και αυτή η προσέγγιση σε μεγάλα και πολύπλοκα δίκτυα διανομής μπορεί να αποβεί χρονοβόρα, έστω και λιγότερο από την `sfs`. Έτσι, η δεύτερη προσέγγιση αφορά την χρήση ενός διαγράμματος που δείχνει την 'χρησιμότητα' κάθε χαρακτηριστικού όπως την παρουσιάζει ο αλγόριθμος random forest. Με αυτόν τον τρόπο, μπορούν να επιλεγθούν αμέσως τα πιο σημαντικά χαρακτηριστικά, καθώς θεωρούνται ως αυτά τα οποία βρίσκονται

πριν από το σημείο ‘γονάτου’ του διαγράμματος. Για το παράδειγμα μας, το διάγραμμα που κατασκευάζεται είναι το:



Με αυτό τον τρόπο, φαίνεται πως τα πιο σημαντικά χαρακτηριστικά, τώρα που κατατάσσονται σύμφωνα με τη σημασία του καθενός, είναι οι στήλες 27, 33, 6, 34, 10 και 32. Παρατηρείται πως οι 3 από αυτές (33, 10, 34), έχουν επιλεγθεί επίσης και από την μέθοδο sequential forward selection. Επίσης, όλα τα χαρακτηριστικά αυτά αντιστοιχούν σε ζυγούς που βρίσκονται στο ένα άκρο ενός διακόπτη. Ωστόσο σύμφωνα με αυτή την τεχνική, έχουν βρεθεί 6 θέσεις SMDs, ενώ όπως θα δειχθεί παρακάτω, χρειάζονται μόνο 4.

Ακολουθώντας την πρώτη προσέγγιση, αναλόγως και με την μέθοδο που αναπτύχθηκε για TI με sfs, πραγματοποιήθηκε στην συνέχεια hyperparameter tuning, ελέγχοντας παράλληλα τον αριθμό των SMDs που χρειάζονται.

Έτσι, δοκιμάζοντας αρχικά το βέλτιστο μοντέλου που είχε βρεθεί με την μέθοδο sfs, βρέθηκε πως χρειάζονταν επτά με οχτώ SMDs τα οποία έδιναν ακρίβεια του νευρωνικού: 0.971875 και 0.9825 (x100%). Έτσι, με δοκιμές ‘με το χέρι’ προσαρμόστηκαν οι υπερπαραμέτροι έτσι ώστε τα αποτελέσματα να βελτιωθούν. Παρακάτω δίνεται ένας ενδεικτικός πίνακας με τις δοκιμές που πραγματοποιήθηκαν:

Επιλεγμένα Χαρακτηριστικά	Ακρίβεια	Κρυφά Στρώματα + Νευρώνες	Epochs	Batch Size
[27, 33, 6, 34, 10, 32, 9, 31, 25, 5]	0.975625	128 96 64 32	10	128
[27, 33, 6, 34, 10, 32, 9, 31, 25]	0.951250	32 64 96 128	10	128
[27, 33, 6, 34, 10, 32, 9, 31]	0.951875	32 64 128	10	128
[27, 33, 6, 34, 10, 32, 9, 31, 25, 5]	0.888750	32 64	10	128
[27, 33, 6, 34, 10, 32, 9, 31]	0.959375	32 64 128 256	10	128
[27, 33, 6, 34, 10, 32, 9, 31]	0.970000	2 64 128 256 512	10	128
[27, 33, 6, 34, 10, 32, 9, 31]	0.969375	32 64 128 256	20	128
[27, 33, 6, 34, 10, 32, 9]	0.955000	32 64 128 256	50	128
[27, 33, 6, 34, 10]	0.961250	32 64 128 256	100	128
[27, 33, 6, 34, 10]	0.976250	32 64 128 256	200	128
[27, 33, 6, 34, 10]	0.961875	32 64 128 256	400	128
[27, 33, 6, 34, 10]	0.991250	32 64 128 256	500	128
[27, 33, 6, 34, 10]	0.965625	32 64 128 256	750	128
[27, 33, 6, 34, 10]	0.965000	32 64 128 256	1000	128
[27, 33, 6, 34]	0.955625	32 64 128 256	500	64
[27, 33, 6, 34]	0.958125	32 64 128 256	500	32
[27, 33, 6, 34]	0.953750	32 64 128 256	500	16

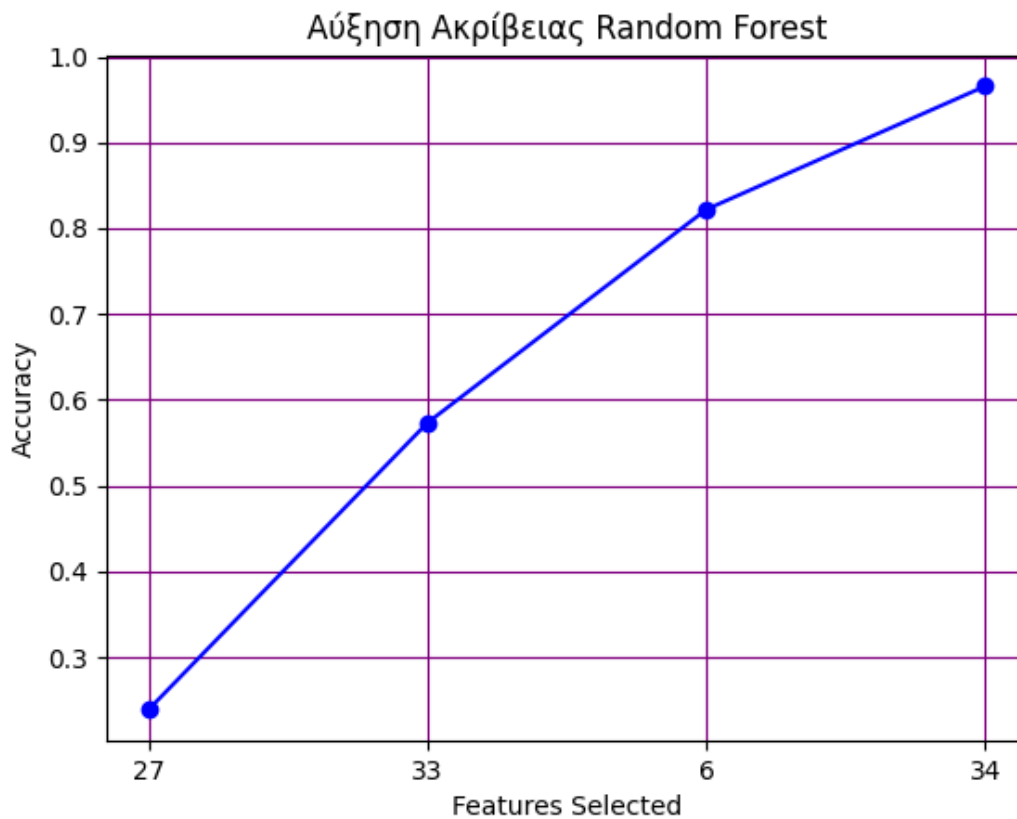
Στη συνέχεια εκτελέστηκε ένα grid search έτσι ώστε να βρεθούν οι υπερπαραμέτροι που δίνουν στο νευρωνικό δίκτυο την καλύτερη απόδοση, όταν αυτό εκπαιδευτεί με δεδομένα από τα χαρακτηριστικά 27, 33, 6 και 34. Οι παράμετροι που εξερευνήθηκαν ήταν οι εξής:

```
layers = [4, 5, 6]
neurons_values = [16, 32, 64, 128]
epochs_values = [10, 20, 50, 100, 200, 400, 500, 750, 1000]
batch_size_values = [16, 32, 64]
```

Έτσι, ως βέλτιστο μοντέλο επιλέχθηκε το αυτό με 4 στρώματα, 32 64 128 256 αντίστοιχα σε κάθε στρώμα, με 500 εποχές και batch size 64 καθώς παρουσιάζει επανειλημμένα την μεγαλύτερη ακρίβεια και χρειάζεται μόλις 4 SMDs. Αυτά τοποθετούνται στους κλάδους:

- 28: ζυγός 28 → ζυγός 29 (στήλη 27 πίνακα δεδομένων)
- 34: ζυγός 12 → ζυγός 22 (στήλη 33 πίνακα δεδομένων – διακόπτης 2)
- 7: ζυγός 7 → ζυγός 8 (στήλη 6 πίνακα δεδομένων)
- 35: ζυγός 25 → ζυγός 29 (στήλη 34 πίνακα δεδομένων – διακόπτης 3)

Στο παρακάτω διάγραμμα παρουσιάζεται η ακρίβεια του μοντέλου, όπως αυτή αυξάνεται διαδοχικά με τα χαρακτηριστικά που προστίθενται:



Μετά την εκπαίδευση, με επιλογή του χαρακτηριστικού 27 η ακρίβεια του μοντέλου είναι 0.23875, μετά την επιλογή του χαρακτηριστικού 33 0.573125, μετά την επιλογή του χαρακτηριστικού 6 0.821875 μετά την επιλογή του χαρακτηριστικού 34 0.965625.

Στη συνέχεια, για την ανάπτυξη του νευρωνικού για DSSE και για την τοποθέτηση επιπλέον SMDs για DSSE πραγματοποιήθηκε μία ανάλογη διαδικασία όπως και με την τοποθέτηση SMDs με hierarchical clustering. Η βασική αρχιτεκτονική του νευρωνικού δικτύου θεωρήθηκε παρόμοια και στις δύο περιπτώσεις, με αυτό για DSSE να θεωρούμε πως αποτελείται πάλι από μερικά κρυφά κοινά στρώματα και δύο κλάδους που βγάζουν δύο διαφορετικές εξόδους.

Έτσι, αρχικά ελέγχθηκε ο αριθμός των κοινών κρυφών στρωμάτων που χρειάζονται προσεγγιστικά για να επιτευχθεί ακρίβεια μοντέλου παραπάνω από το όριο. Παρακάτω, παρουσιάζονται τα αποτελέσματα μερικών εκ των προσεγγιστικών αυτών δοκιμών που πραγματοποιήθηκαν:

Επιλεγμένα Χαρακτηριστικά	MAE Γωνιών	MAPE Μέτρων	ρυφά Στρώματα + Νευρώνες	Epochs	Batch Size
[8, 3, 6], [7, 8, 9]	0.12933721164203	0.45928652735344	64 128 256 512	50	16
[8, 3, 6], [7, 8, 9]	0.14554137160979	0.36120151121757	32 64 128 256 512	50	16
[8, 3, 6], [7, 8, 9]	0.14312902006639	0.35129700019979	32 64 128 256 512 1024	50	16
[8, 3, 6], [7, 8, 9]	0.13948045827843	0.28596739495277	32 64 128 256 512 1024 2048	50	16

Με αυτό τον τρόπο φάνηκε πως ένα δίκτυο με 3 ή 4 κοινά κρυφά στρώματα είναι αρκετά για να επιτευχθεί η επιθυμητή ακρίβεια. Έτσι, στην συνέχεια εκτελέστηκε ξανά ένα grid search θεωρώντας ως επιπλέον SMDs

για DSSE τα χαρακτηριστικά 8 και 7 καθώς αυτά αποτελούν τον ελάχιστο αριθμό επιπλέον SMDs. Οι παράμετροι του grid search ήταν οι εξής:

```
shared_layers_values = [3, 4]
neurons_values = [32, 50, 64, 100, 128]
epochs_values = [10, 50, 100, 200, 400, 500, 750, 1000]
batch_size_values = [16, 32, 64]
learning_rates = [0.001, 0.0001]
```

Όπου, ξανά θεωρήθηκε πως τα δύο branches του δικτύου θα είναι συμμετρικά και πως θα αποτελούνται από δύο στρώματα, το καθένα με διπλάσιο αριθμό νευρώνων από το προηγούμενο, καθώς αυτό φάνηκε μέσα από τις δοκιμές με το 'χέρι' να αυξάνει την ακρίβεια και καθιστά την διαδικασία της ρύθμισης των υπερπαραμέτρων αρκετά πιο εύκολη. Ως shared layer values ορίστηκε ο αριθμός των κοινών στρωμάτων, ως neuron value ο αριθμός των νευρώνων του πρώτου κοινού στρώματος, ο οποίος θα είναι διπλάσιος στο επόμενο στρώμα και ούτω καθεξής.

Αυτή η αναζήτηση πλέγματος έδωσε 23 πιθανούς συνδυασμούς υπερπαραμέτρων έτσι ώστε με αυτούς το νευρωνικό δίκτυο μετά την εκπαίδευση να ξεπερνά τα όρια που έχουν τεθεί για τα σφάλματα. Το καλύτερο εξ αυτών, δηλαδή αυτό που παρουσίαζε το μικρότερο άθροισμα σφαλμάτων είναι το:

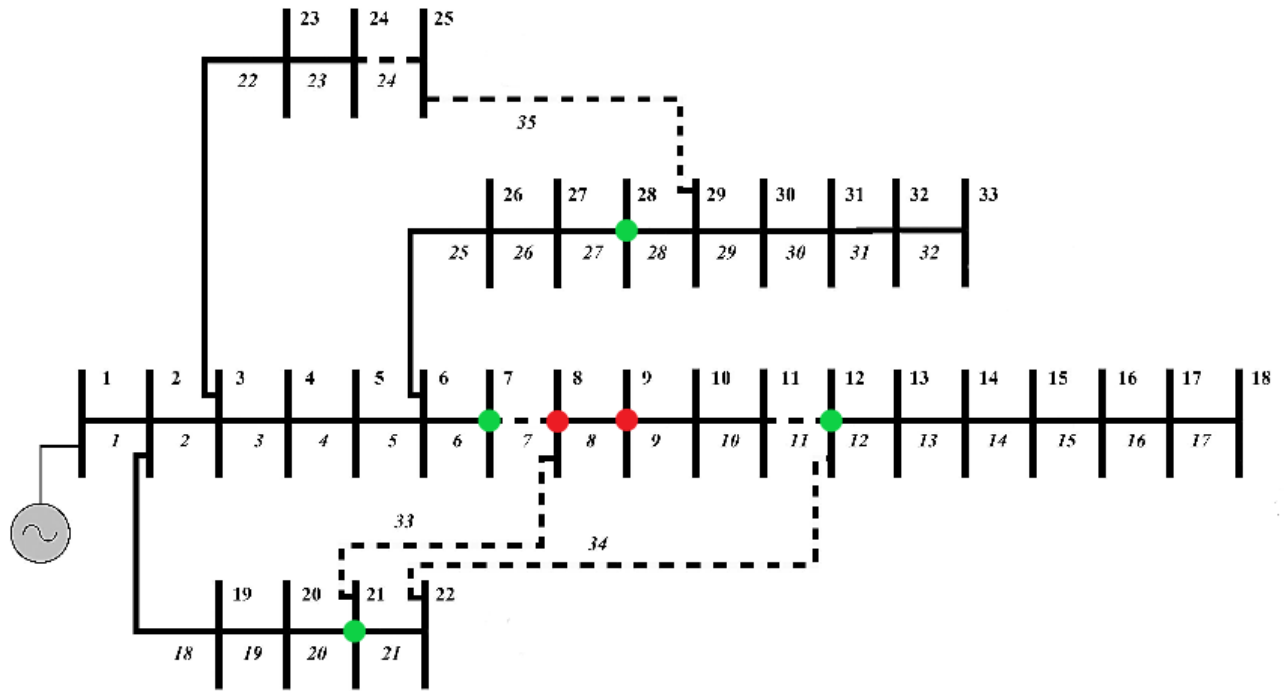
```
Shared Layers: 3
Neurons: 64
Epochs: 750
Batch_Size: 16
Learning Rate: 0.0001
MAE: 0.11709231035396636
MAPE: 0.2087770195902655
Combined Metric: 0.3258693299442319
```

Δηλαδή, οι βέλτιστες υπερπαραμέτροι είναι 3 κοινά κρυφά στρώματα, τα οποία μαζί με το στρώμα εισόδου, τα δύο στρώματα εξόδου και τα δύο κρυφά στρώματα σε κάθε κλαδί αποτελούν τα στρώματα του μοντέλου. Οι νευρώνες του πρώτου κοινού κρυφού στρώματος θα είναι 64 και στη συνέχεια κάθε επόμενο στρώμα θα έχει τον διπλάσιο αριθμό νευρώνων από το προηγούμενο του, με εξαίρεση τα στρώματα εξόδου τα οποία έχουν σταθερά 33 νευρώνες. Ο βέλτιστος αριθμός εποχών θα είναι 750, ενώ batch size 16 και ρυθμού μάθησης 0.0001. Η ακρίβεια που παρουσιάζει αυτό το μοντέλο μετά την εκπαίδευση με μόλις 2 επιπλέον SMDs θα είναι 0.11709231035396636 μέσο απόλυτο σφάλμα για τις γωνίες τάσης φάσης και 0.2087770195902655% μέσο απόλυτο ποσοστιαίο σφάλμα για τα μέτρα τάσης.

Συγκριτικά με την μέθοδο που αναπτύχθηκε πρωτίτερα βασισμένη σε sequential forward selection και hierarchical clustering, η Random Forest έχοντας το πλεονέκτημα της ταχύτητας της επιλογής χαρακτηριστικών, ιδίως έναντι της sfs, έχει την ίδια απόδοση καθώς χρειάζονται και στις δύο περιπτώσεις 4 SMDs για TI και 2 επιπλέον για DSSE.

Επίσης, παρατηρείται πως τα πιθανά μοντέλα που μπορούν να χρησιμοποιηθούν μετά το grid search είναι 23 έναντι 35, παρόλο που ο χώρος των υπερπαραμέτρων στο τελευταίο grid search που έγινε για την προσέγγιση με Random Forest ήταν σχεδόν ο διπλάσιος. Αυτό υποδηλώνει πως τα χαρακτηριστικά που επιλέγονται με sfs περιέχουν μεγαλύτερο εύρος πληροφοριών σε σύγκριση με τα χαρακτηριστικά με Random Forest, και συσχετίζονται περισσότερο μεταξύ τους.

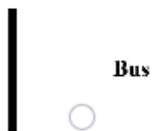
Στο παρακάτω σχήμα του δικτύου απεικονίζονται οι θέσεις τοποθέτησης SMDs για TI (με πράσινο) και για DSSE (με κόκκινο).



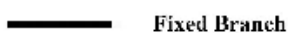
Όπου:



Generator



Bus



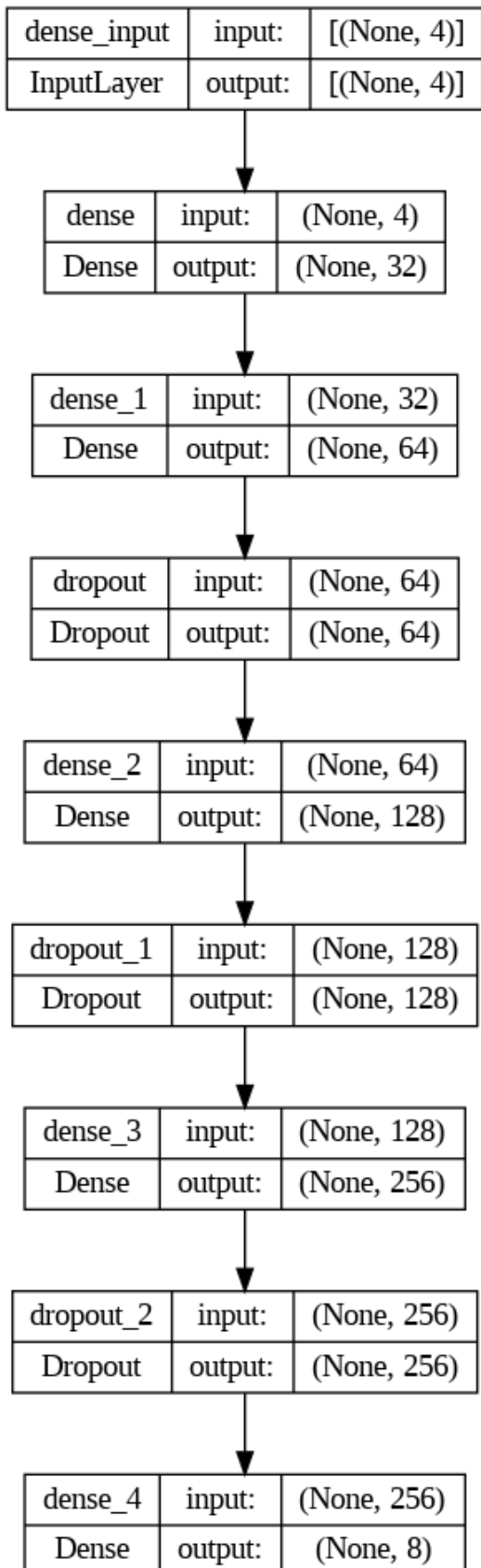
Fixed Branch



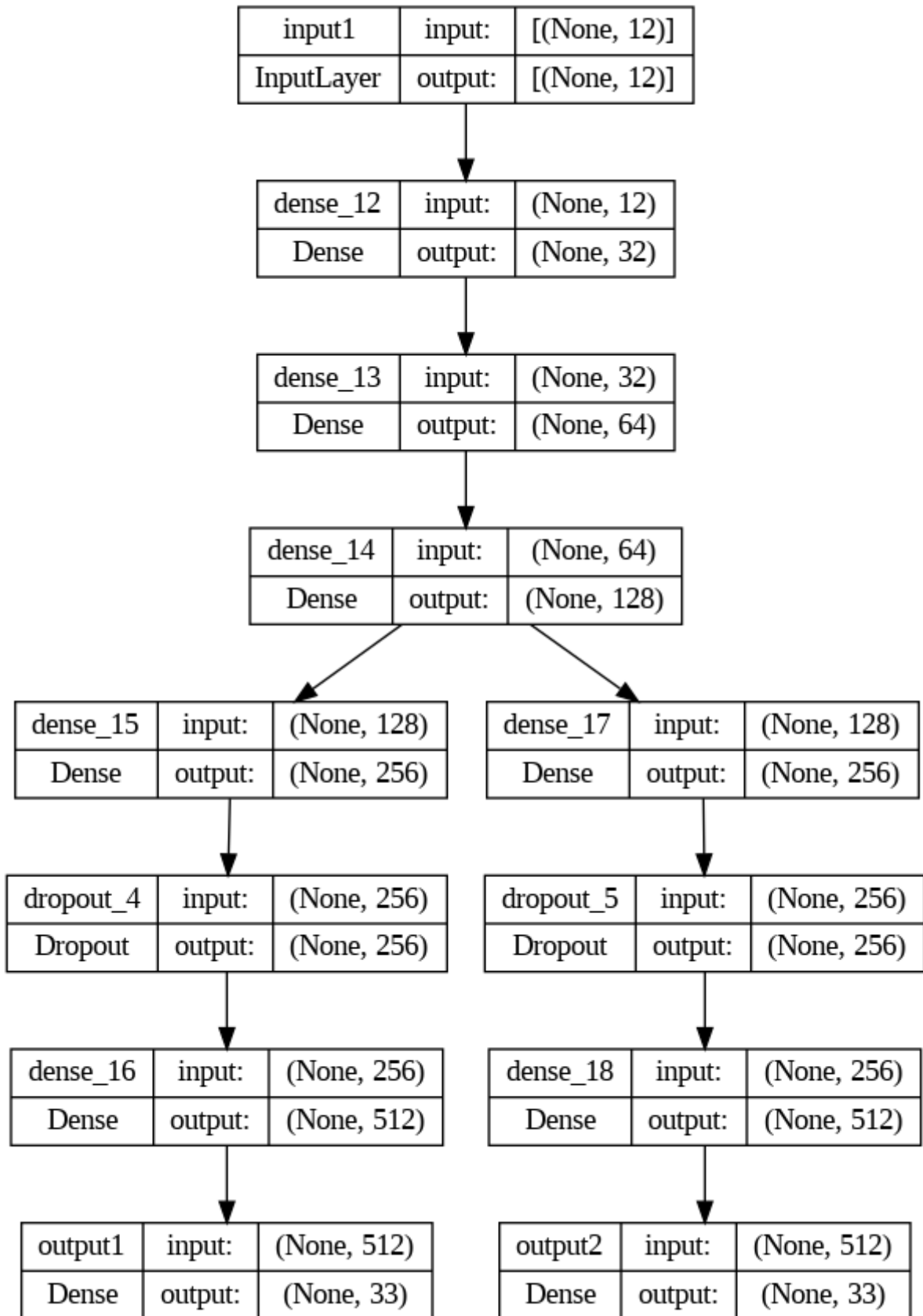
Switchable Branch

Τα τελικά νευρωνικά δίκτυα που χρησιμοποιήθηκαν απεικονίζονται στα παρακάτω διαγράμματα:

Για ΤΙ:



Γ1α DSSE:



Ο χρόνος ο οποίος χρειάζεται το νευρωνικό δίκτυο που αναπτύχθηκε για την επιλογή χαρακτηριστικών με Random Forest και στην συνέχεια ρυθμίστηκαν οι υπερπαραμέτροι του εκτελεί TI σε χρόνο: 350.26286458969116 seconds και δίνει τελική ακρίβεια: 0.965625 (*100%). Από την άλλη πλευρά, νευρωνικό δίκτυο που αναπτύχθηκε για την επιλογή χαρακτηριστικών με SFS και στην συνέχεια ρυθμίστηκαν οι υπερπαραμέτροι του εκτελεί TI σε χρόνο 329.7047920227051 seconds με τελική ακρίβεια: 0.975625 (*100%). Παρ'όλο που η Random Forest χρειάζεται πολύ πιο πολλές εποχές για να εκτελέσει TI (50πλάσιες ουσιαστικά), χρησιμοποιεί μόνο 4 εκπαιδεύσεις συνολικά κατά την εκτέλεση του αλγορίθμου, ενώ η SFS πραγματοποιεί αντίστοιχα 122 εκπαιδεύσεις. Γενικότερα, η Random Forest εκτελεί αριθμό εκπαιδεύσεων ίσο με τα χαρακτηριστικά που τελικά επιλέγονται, ενώ η SFS χρειάζεται εκπαιδεύσεις ίσες με:

$$\sum_{i=0}^{k-1} (N - i)$$

όπου k είναι ο αριθμός των χαρακτηριστικών που τελικά επιλέγονται και N ο συνολικός αριθμός των χαρακτηριστικών. Έτσι, ενώ ο χρόνος εκτέλεσης αυτών των δύο αλγορίθμων είναι παρόμοιος, αν χρησιμοποιηθούν σε μεγαλύτερα και πιο πολύπλοκα δίκτυα, η διαφορά θα γίνει πιο ορατή. Η ακρίβεια των δύο αυτών μεθόδων τελικά είναι σχεδόν ίδια, όπως ήταν αναμενόμενο καθώς και στα δύο χρησιμοποιούνται βαθιά νευρωνικά δίκτυα με παρόμοιες παραμέτρους.

Ως προς το DSSE επίσης η ακρίβεια είναι παρόμοια, όπως και ο χρόνος εκτέλεσης του Random Forest με το Hierarchical Clustering, οπότε η Random Forest θεωρείται πως αποτελεί απλώς εναλλακτική μεθοδολογία εκτέλεσης DSSE.

Συμπεράσματα

Στην παρούσα εργασία αναλύεται μια προσέγγιση βαθιάς μάθησης στην εκτίμησης κατάστασης δικτύων διανομής, για δίκτυα με διαφορετικές παραμέτρους, που δεν απαιτεί πλήρη παρατηρησιμότητα από τα SMD σε πραγματικό χρόνο. Η εκτίμηση κατάστασης εκτελείται με ένα βαθύ νευρωνικό δίκτυο το οποίο προσομοιώνει έναν εκτιμητή MMSE, ενώ μεταβλητές κατάστασης θεωρούνται οι φασιθέτες τάσεις των ζυγών του συστήματος. Όταν συμβαίνει μια αλλαγή στην τοπολογία, δηλαδή μια αλλαγή στην κατάσταση των διακοπών, στο προτεινόμενο πλαίσιο ανιχνεύεται πρώτα η αλλαγή αυτή χρησιμοποιώντας ένα ακόμη βαθύ νευρωνικό δίκτυο που έχει κατασκευαστεί για την εκτίμηση της τοπολογίας, και στη συνέχεια χρησιμοποιείται η μέθοδος transfer learning (μεταφορά μάθησης) μέσω της τεχνικής fine tuning για την ενημέρωση του αρχικού νευρωνικού δικτύου υπεύθυνου για την εκτέλεση της εκτίμησης κατάστασης, έτσι ώστε να διατηρηθούν όσο είναι δυνατόν τα μοτίβα που έχουν μαθευτεί από το νευρωνικό αυτό δίκτυο κατά την διάρκεια της εκπαίδευσης, κατά την μετάβαση από την παλιά στην νέα τοπολογία, κι έτσι να μειωθεί ο χρόνος της εκπαίδευσης, ο οποίος μπορεί να αποβεί σημαντικός για μεγάλα δίκτυα και μπορεί να μην είναι εφαρμόσιμος σε πραγματικό χρόνο.

Με αυτόν τον τρόπο, το πρόβλημα της εκτίμησης κατάστασης ανάγεται σε ένα πρόβλημα παλινδρόμησης, με στόχο την εκτίμηση των τάσεων των ζυγών του συστήματος, και αυτό της εκτίμησης τοπολογίας, ανάγεται σ'ένα πρόβλημα ταξινόμησης, με κλάσεις όσες και οι πιθανές εφαρμόσιμες τοπολογίες του δικτύου. Τα βαθιά νευρωνικά δίκτυα είναι ιδιαίτερα αποτελεσματικά για την επίλυση τέτοιων προβλημάτων και διαδραματίζουν καθοριστικό ρόλο στην υπέρβαση της υπολογιστικής πολυπλοκότητας υπολογισμού στην Μπεϋζιανή εκτίμηση κατάστασης, καθιστώντας αυτή την προσέγγιση σημαντικά καλύτερη από παραδοσιακές μεθόδους εκτίμησης κατάστασης, όπως φάνηκε από τα αποτελέσματα.

Δεδομένα που παράγονται μέσω προσομοιώσεων Monte Carlo χρησιμοποιήθηκαν για την εκπαίδευση των δύο βαθιών νευρωνικών δικτύων, παρμένες από προσομοιώσεις των δικτύων στο Matlab. Τα αποτελέσματα των προσομοιώσεων καταδεικνύουν τις δυνατότητες της Μπεϋζιανής εκτίμησης κατάστασης και της εφαρμογής τεχνικών της μηχανικής μάθησης για την ανάλυση συστημάτων ηλεκτρικής ενέργειας.

Επίσης, αναπτύχθηκε μια λεπτομερής μεθοδολογία για την τοποθέτηση SMDs (έξυπνων μετρητών) για την ενίσχυση των επιδόσεων της εκτίμησης κατάστασης και της εκτίμησης τοπολογίας, η οποία βασίζεται στις τεχνικές επιλογής χαρακτηριστικών Sequential Forward Selection και Hierarchical Clustering, και η οποία αποσκοπεί στην παράκαμψη της μη παρατηρησιμότητας που εμφανίζεται πολύ συχνά στα δίκτυα διανομής για λόγους κόστους.

Λόγω όμως του μεγάλου υπολογιστικού κόστους αυτών των δύο τεχνικών εξετάστηκε και η εφαρμογή του αλγορίθμου Random Forest για την επιλογή χαρακτηριστικών κι έτσι και βέλτιστων θέσεων τοποθέτησης SMDs. Η μέθοδος αυτή αποδείχθηκε πολύ πιο συμφέρουσα χρονικά, καθώς ο αριθμός των εκπαιδεύσεων του μοντέλου εκτίμησης τοπολογίας μειώνεται δραστικά.

Επιπλέον, παρατηρήθηκε πως για DSSE, στα νευρωνικά δίκτυα που είναι πιο αποτελεσματικά ο αριθμός των νευρώνων κάθε στρώματος αυξάνεται όσο πιο κοντά είναι αυτό το στρώμα στο στρώμα εξόδου, ενώ στα νευρωνικά δίκτυα για TI συμβαίνει το αντίθετο. Επίσης, για TI χρειάζονται περίπου 10 – 20 epochs ενώ για DSSE 500+ epochs, το οποίο οφείλεται κυρίως στο γεγονός πως το DSSE είναι ένα πολύ πιο περίπλοκο πρόβλημα που αφορά την εκτίμηση 66 μεγεθών, ενώ το TI μόνος μίας ετικέτας (της εκάστοτε κλάσης).

Συνοπτικά, στην παρούσα εργασία δείχτηκε πως τα βαθιά νευρωνικά δίκτυα είναι ικανά να εκτελέσουν υψηλής ακρίβειας εκτίμηση κατάστασης χρησιμοποιώντας μόνο έναν πολύ μικρό αριθμό μετρητών. Επίσης, μπορούν να αναγνωρίσουν αποτελεσματικά ποικίλες τοπολογίες δικτύου για αναδιαμορφώσιμα δίκτυα διανομής, και μαζί με αυτό να εκτελούν εκτίμηση κατάστασης αξιόπιστα για διαφορετικές τοπολογίες. Αυτές οι ικανότητες σε συνδυασμό με τους αλγορίθμους που αναπτύχθηκαν καθιστούν αυτήν την προσέγγιση εξαιρετικά αποτελεσματική σε μεγάλα και πολύπλοκα δίκτυα διανομής, για διαφορετικές τοπολογίες για ενισχυμένη παρακολούθηση, προστασία, ελέγχου και όλες τις άλλες λειτουργίες που επιτελούνται με την βοήθεια της εκτίμησης κατάστασης στα πραγματικά, πολύπλοκα συστήματα διανομής.

Επεκτάσεις

Η Μπεϋζιανή προσέγγιση που παρουσιάστηκε σε αυτή την εργασία έχει, όπως είναι φυσικό, τους περιορισμούς της. Περαιτέρω έρευνα και αξιολογήσεις εκτός του πεδίου εφαρμογής της παρούσας εργασίας είναι απαραίτητες.

Αρχικά, επειδή η εκ των προτέρων κατανομή των φορτίων παίζει σημαντικό ρόλο στις Μπεϋζιανές τεχνικές, ο εκτιμητής που χρησιμοποιήθηκε είναι αρκετά ευαίσθητος σε αλλαγές στο δίκτυο, όπως π.χ. διακοπές γραμμών και παραγωγής ενέργειας. Επίσης, η εκπαίδευση των βαθιών νευρωνικών δικτύων είναι ένας ερευνητικός τομέας που υφίσταται συνεχώς εντατική έρευνα. Καμία από τις διαθέσιμες τεχνικές δεν μπορεί να εγγυηθεί πως παράγει τα καλύτερα αποτελέσματα. Για το λόγο αυτό, ενώ η προσέγγιση που αναπτύχθηκε είναι αξιόπιστη και πολλά υποσχόμενη, όπως καθεμία στο πλαίσιο της τεχνητής νοημοσύνης, οι αποδόσεις που αναφέρθηκαν είναι αποτελέσματα προσεκτικής ρύθμισης ορισμένων από τις υπερπαραμέτρους των δικτύων για τα συγκεκριμένα δίκτυα που μελετήθηκαν.

Έτσι, αρκετές επεκτάσεις είναι δυνατές για την περαιτέρω βελτίωση αυτής της τεχνικής εκτίμησης κατάστασης με νευρωνικά δίκτυα. Για παράδειγμα, ενδιαφέρον παρουσιάζει η πιθανότητα χρήσης άλλων, πιο περίπλοκων κι εξειδικευμένων νευρωνικών δικτύων όπως τα CNNs, τα LSTMs και τα πρόσφατα δημιουργημένα και πολλά υποσχόμενα transformers, τα οποία δύναται να αυξήσουν σημαντικά την απόδοση χρησιμοποιώντας λιγότερους έξυπνους μετρητές, αλλά ενδέχεται να είναι υπολογιστικά πολύ πιο δύσκολη.

Πρακτικά, καθώς αυτή η εργασία συνδυάζει μερικούς πολύ ευρείς τομείς, σχεδόν όλες οι τεχνικές που χρησιμοποιήθηκαν χρήζουν περαιτέρω

διερεύνησης όπως για παράδειγμα η δημιουργία κατάλληλων και πιο κοντά στην πραγματικότητα δεδομένων, καθώς όπως έχει δειχθεί εδώ [71], τα φορτία δεν αναπαρίστανται με ακρίβεια να θεωρηθεί πως ακολουθούν κανονική κατανομή όπως έγινε εδώ και καθώς οι ψευδομετρήσεις όπως προαναφέρθηκε έχουν κι αυτές τους περιορισμούς και τα μειονεκτήματά τους, παρά την μεγάλη πρακτική του εφαρμογή. Είναι επίσης δυνατό να αξιοποιηθούν οι χρονικές εξαρτήσεις των δεδομένων.

Επιπλέον, για την τοποθέτηση μετρητών σε βέλτιστες θέσεις για την δημιουργία δεδομένων έχει ενδιαφέρον να διερευνηθούν κι άλλες τεχνικές επιλογής χαρακτηριστικών που έχουν αναπτυχθεί στο πλαίσιο της βαθιάς μάθησης, προκειμένου να αυξηθεί ακόμα περισσότερο η απόδοση των μοντέλων ταξινόμησης και παλινδρόμησης που χρησιμοποιούνται, όπως για παράδειγμα η ανάλυση κύριων παραγόντων (Principal Component Analysis - PCA) και η παλινδρόμηση Lasso.

Τέλος, θα ήταν χρήσιμο να βρεθούν πιο αποδοτικοί τρόποι μεταφοράς των μοτίβων που έχουν διδαχτεί τα νευρωνικά δίκτυα όταν αλλάζει η τοπολογία ενός δικτύου, καθώς το fine tuning, κατά το οποίο το νευρωνικό δίκτυο εκπαιδεύεται ξανά, μπορεί να αποβεί εξαιρετικά χρονοβόρο για μεγάλα δίκτυα.

Βιβλιογραφία:

- [1] B. Azimian, R. S. Biswas, S. Moshtagh, A. Pal, L. Tong and G. Dasarathy, "State and Topology Estimation for Unobservable Distribution Systems Using Deep Neural Networks," in IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-14, 2022, Art no. 9003514.
- [2] Zanjani, M.G.M.; Mazlumi, K.; Kamwa, I. Application of μ PMUs for Adaptive Protection of Overcurrent Relays in Microgrids. IET Gener. Transm. Distrib. 2018, 12, 4061–4068
- [3] Μ. Π. Παπαδόπουλος, Δίκτυα Διανομής Ηλεκτρικής Ενέργειας. Αθήνα: Εθνικό Μετσόβιο Πολυτεχνείο, 1994.
- [4] Tahabilder, A.; Ghosh, P.K.; Chatterjee, S.; Rahman, N. Distribution System Monitoring by Using Micro-PMU in Graph-Theoretic Way. In Proceedings of the 4th International Conference on Advances in Electrical Engineering (ICAEE), Dhaka, Bangladesh, 28–30 September 2017; pp. 159–163.
- [5] Dusabimana, E.; Nishimwe, H.L.F.; Yoon, S. Optimal Placement of Micro-Phasor Measurement Units and Power Flow Measurements to Monitor Distribution Network. In Proceedings of the International Council on Electrical Engineering (ICEE), Hong Kong, China, 2–6 July 2019.
- [6] Chen, X.; Chen, T.; Tseng, K.J.; Sun, Y.; Amaratunga, G. Customized Optimal μ PMU Placement Method for Distribution Networks. In Proceedings of the IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Xi'an, China, 25–28 October 2016; pp. 135–140
- [7] <https://utilitiesone.com/the-evolution-of-power-distribution-systems>
- [8] https://www.enerdynamics.com/Energy-Currents_Blog/What-Is-a-Phasor-Measurement-Unit-and-How-Does-it-Make-the-Grid-More-Reliable.aspx
- [9] Saleem, Danish & Siqueira de Carvalho, Ricardo. (2020). Recommended Functionalities for Improving Cybersecurity of Distributed Energy Resources.
- [10] Fotopoulou, M.; Petridis, S.; Karachalios, I.; Rakopoulos, D. A Review on Distribution System State Estimation Algorithms. Appl. Sci. 2022, 12, 11073.

- [11] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "Bayesian state estimation for unobservable distribution systems via deep learning," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4910-4920, Nov. 2019.
- [12] A. Gandluru, S. Poudel, and A. Dubey, "Joint estimation of operational topology and outages for unbalanced power distribution systems," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 605-617, Jan. 2020.
- [13] G. Cavraro, A. Bernstein, V. Kekatos, and Y. Zhang, "Real-time identifiability of power distribution network topologies with limited monitoring," *IEEE Control Syst. Letters*, vol. 4, no. 2, pp. 325-330, Apr. 2020.
- [14] W. Jiang et al., "A physical probabilistic network model for distribution network topology recognition using smart meter data," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6965-6973, Nov. 2019.
- [15] G. Cavraro and R. Arghandeh, "Power distribution network topology detection with time-series signature verification method," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3500-3509, Jul. 2018.
- [16] L. Ma, L. Wang and Z. Liu, "Topology identification of distribution networks using a split-EM based data-driven approach," *IEEE Trans. Power Syst.* Early Access 2021.
- [17] L. Ma, L. Wang and Z. Liu, "Topology identification of distribution networks using a split-EM based data-driven approach," *IEEE Trans. Power Syst.* Early Access 2021.
- [18] J. Zhang, Y. Wang, Y. Weng and N. Zhang, "Topology identification and line parameter estimation for non-PMU distribution network: a numerical method," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4440-4453, Sep. 2020.
- [19] D. Atanackovic and V. Dabic, "Deployment of real-time state estimator and load flow in BC Hydro DMS - Challenges and opportunities," in *Proc. 2013 IEEE Power & Energy Soc. General Meeting*, Vancouver, BC, Canada, 2013, pp. 1-5.
- [20] F. Aminifar, M. Fotuhi-Firuzabad, A. Safdarian, A. Davoudi, and M. Shahidehpour, "Synchrophasor measurement technology in power systems: panorama and state-of-the-art," *IEEE Access*, vol. 2, pp. 1607-1628, 2015.
- [21] D. A. Haughton and G. T. Heydt, "A linear state estimation formulation for smart distribution systems," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1187-1195, May 2013.]

- [22] D. Macii, G. Barchi, and L. Schenato, "On the role of phasor measurement units for distribution system state estimation," in Proc. IEEE Workshop Environ. Energy Struct. Monitoring Syst., Naples, FL, USA, 2014.
- [23] P. Janssen, T. Sezi, and J. C. Maun, "Distribution system state estimation using unsynchronized phasor measurements," in Proc. 3rd IEEE PES Int. Conf. Exhib. Innovative Smart Grid Technol. Eur., Berlin, Germany, 2012, pp. 1–6.
- [24] G. Celli, P. A. Pegoraro, F. Pilo, G. Pisano, and S. Sulis, "DMS cyberphysical simulation for assessing the impact of state estimation and communication media in smart grid operation," IEEE Trans. Power Syst., vol. 29, no. 5, pp. 2436–2446, Sep. 2014.
- [25] S. M. S. Alam, B. Natarajan, and A. Pahwa, "Distribution grid state estimation from compressed measurements," IEEE Trans. Smart Grid, vol. 5, no. 4, pp. 1631–1642, Jul. 2014.
- [26] A. K. Ghosh, D. L. Lubkeman, and R. H. Jones, "Load modeling for distribution circuit state estimation," IEEE Trans. Power Del., vol. 12, no. 2, pp. 999–1005, Apr. 1997.
- [27] E. Manitsas, R. Singh, B. Pal, and G. Strbac, "Modelling of pseudomeasurements for distribution system state estimation," in Proc. Int. Conf. Exhib. Elect. Distrib. Seminar: SmartGrids Distrib., Frankfurt, Germany, 2008, pp. 1–4.
- [28] R. Arritt and R. Dugan, "Comparing load estimation methods for distribution system analysis," in Proc. 22nd Int. Conf. Exhib. Elect. Distrib., Stockholm, Sweden, 2013, pp. 1–4.
- [29] W. M. Lin and J. H. Teng, "State estimation for distribution systems with zero-injection constraints," IEEE Trans. Power Syst., vol. 11, no. 1, pp. 518–524, Feb. 1996.
- [30] C. Muscas, M. Pau, P. A. Pegoraro, and S. Sulis, "An efficient method to include equality constraints in branch current distribution system state estimation," EURASIP J. Adv. Signal Process., vol. 2015, pp. 2–11, 2015.
- [31] Chen, X.; Chen, T.; Tseng, K.J.; Sun, Y.; Amaratunga, G. Hybrid Approach Based on Global Search Algorithm for Optimal Placement of μ PMU in Distribution Networks. In Proceedings of the IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia), Melbourne, VIC, Australia, 28 November–1 December 2016; pp. 559–563.
- [32] M. E. Baran and A. W. Kelley, "State estimation for real-time monitoring of distribution systems," IEEE Trans. Power Syst., vol. 9, no. 3, pp. 1601–1609, Aug. 1994.

- [33] D. A. Haughton and G. T. Heydt, "A linear state estimation formulation for smart distribution systems," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1187–1195, May 2013.
- [34] P. M. D. O. D. Jesus and A. A. R. Quintana, "Distribution system state estimation model using a reduced quasi-symmetric impedance matrix," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 2856–2866, Nov. 2015.
- [35] https://en.wikipedia.org/wiki/Electric_power_distribution.
- [36] <https://www.tisoft.com/el/support/help/electricaldesign/project/parameters/network/lowvoltage>.
- [37] <https://electrical-engineering-portal.com/distribution-substation>
- [38] Distribution network topology planning and optimization: A brief review, Ayoub El Berkaoui, Saida Bahsine, Aziz Oukennou, Fatima Ait Nouh and Bouchra Rzine, *E3S Web of Conf.*, 469 (2023) 00054 a.
- [39] M. E. Baran, J. Jung, and T. E. McDermott, "Including voltage measurements in branch current state estimation for distribution systems," In *IEEE Power & Energy Society General Meeting*, pp. 1–5, Jul. 2009.
- [40] H. Wang and N. N. Schulz, "A revised branch current-based distribution system state estimation algorithm and meter placement impact," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 207–213, Feb. 2004.
- [41] M. Pau, P. A. Pegoraro, and S. Sulis, "Efficient branch-current-based distribution system state estimation including synchronized measurements," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 9, pp. 2419–2429, Sep. 2013.
- [42] A. Primadianto and C. -N. Lu, "A Review on Distribution System State Estimation," in *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3875-3883, Sept. 2017.
- [43] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan and F. Bu, "A Survey on State Estimation Techniques and Challenges in Smart Distribution Systems," in *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2312-2322, March 2019.
- [44] A. Angioni, T. Schlosser, F. Ponci, and A. Monti, "Impact of pseudomeasurements from new power profiles on state estimation in lowvoltage grids," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 70–77, Jan. 2016.

- [45] A. K. Ghosh, D. L. Lubkeman, M. J. Downey, and R. H. Jones, "Distribution circuit state estimation using a probabilistic approach," *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 45–51, Feb. 1997.
- [46] D. T. Nguyen, "Modeling load uncertainty in distribution network monitoring," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2321–2328, Sep. 2015.
- [47] A. Arefi, G. Ledwich, and B. Behi, "An efficient DSE using conditional multivariate complex Gaussian distribution," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2147–2156, Jul. 2015.
- [48] A. Alimardani, F. Therrien, D. Atanackovic, J. Jatskevich, and E. Vaahedi, "Distribution system state estimation based on nonsynchronized smart meters," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2919–2928, Nov. 2015.
- [49] R. Singh, B. C. Pal, and R. A. Jabr, "Statistical representation of distribution system loads using Gaussian mixture model," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 29–37, Feb. 2010.
- [50] I. Dzafic and R. A. Jabr, "Real time multiphase state estimation in weakly meshed distribution networks with distributed generation," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4560–4569, Nov. 2017.
- [51] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp.1888–1896, Nov. 2012.
- [52] B. P. Hayes, J. K. Gruber, and M. Prodanovic, "A closed-loop state estimation tool for MV network monitoring and operation," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2116–2125, Jul. 2015.
- [53] Hayes, B.P.; Gruber, J.K.; Prodanovic, M. A Closed-Loop State Estimation Tool for MV Network Monitoring and Operation. *IEEE Trans. Smart Grid* 2015, 6, 2116–2125.
- [54] V. Zamani and M. E. Baran, "Meter placement for conservation voltage reduction in distribution systems," *IEEE Trans. Power Syst.*, vol. 33, no.2, pp. 2109-2116, Mar. 2018
- [55] A. S. Zamzam, and N. D. Sidiropoulos, "Physics-aware neural networks for distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4347-4356, Nov. 2020.
- [56] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "Bayesian state estimation for unobservable distribution systems via deep learning," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4910-4920, Nov. 2019.

- [57] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "Bayesian state estimation for unobservable distribution systems via deep learning," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4910-4920, Nov. 2019.
- [58] B. Zargar, A. Angioni, F. Ponci, and A. Monti, "Multiarea parallel data driven three-phase distribution system state estimation using synchrophasor measurements," *IEEE Trans. Instrum. Meas.*, vol. 69, no.9, pp. 6186-6202, Sep. 2020.
- [59] A. Akrami, S. Asif and H. Mohsenian-Rad, "Sparse tracking state estimation for low-observable power distribution systems using D-PMUs," *IEEE Trans. Power Syst.*
- [60] J. A. D. Massignan, J. B. A. London, M. Bessani, C. D. Maciel, R. Z. Fannucchi and V. Miranda, "Bayesian inference approach for information fusion in distribution system state estimation," *IEEE Trans. Smart Grid.*
- [61] A. Gandluru, S. Poudel, and A. Dubey, "Joint estimation of operational topology and outages for unbalanced power distribution systems," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 605-617, Jan. 2020.
- [62] G. Cavraro, A. Bernstein, V. Kekatos, and Y. Zhang, "Real-time identifiability of power distribution network topologies with limited monitoring," *IEEE Control Syst. Letters*, vol. 4, no. 2, pp. 325-330, Apr.2020.
- [63] W. Jiang et al., "A physical probabilistic network model for distribution network topology recognition using smart meter data," *IEEE Trans.Smart Grid*, vol. 10, no. 6, pp. 6965-6973, Nov. 2019.
- [64] G. Cavraro and R. Arghandeh, "Power distribution network topology detection with time-series signature verification method," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3500-3509, Jul. 2018.
- [65] L. Ma, L. Wang and Z. Liu, "Topology identification of distribution networks using a split-EM based data-driven approach," *IEEE Trans. Power Syst.*
- [66] J. Zhang, Y. Wang, Y. Weng and N. Zhang, "Topology identification and line parameter estimation for non-PMU distribution network: a numerical method," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4440-4453, Sep. 2020.
- [67] Xu, Yijun, et al. "An adaptive-importance-sampling-enhanced Bayesian approach for topology estimation in an unbalanced power distribution system," *IEEE Trans. Power Syst.*

- [68] https://cdn-images-1.medium.com/max/800/1*5egrX--WuyrLA7gBEXdg5A.png
- [69] S. Niu, Y. Liu, J. Wang and H. Song, "A decade survey of transfer learning (2010–2020)," *IEEE Trans. Artificial Intelligence*, vol. 1, no. 2, pp. 151-166, Oct. 2020.
- [70] "IEEE/IEC International Standard - Measuring relays and protection equipment - Part 118 -1: Synchrophasor for power systems - Measurements," *IEC/IEEE 60255 -118 -1:2018*, pp.1 -78, Dec. 2018.
- [71] T. Ahmad, and N. Senroy, "Statistical characterization of PMU error for robust WAMS based analytics," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 920-928, Mar. 2020.
- [72] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," in *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401-1407, April 1989.
- [73] Dolatabadi, Sarineh & Ghorbanian, Maedeh & Siano, Pierluigi & Hatziargyriou, Nikos. (2020). An Enhanced IEEE 33 Bus Benchmark Test System for Distribution System Studies. *IEEE Transactions on Power Systems*. PP. 1-1. 10.1109/TPWRS.2020.3038030.
- [74] Dusabimana, E.; Yoon, S.-G. A Survey on the Micro-Phasor Measurement Unit in Distribution Networks. *Electronics* 2020, 9, 305
- [75] Carmichael, J. P., & Liao, Y. (2022). Application of Deep Neural Networks to Distribution System State Estimation and Forecasting. *Frontiers in Sustainable Cities*, 3, Article 814037.
- [76] S. Kumar Kotha, B. Rajpathak, B. Ramesh and M. K. Khedkar, "Optimal Placement Of Micro- PMUs for Real-time Monitoring of Inter-Connected Smart Distribution Networks," 2022 2nd Asian Conference on Innovation in Technology (ASIANCON), Ravet, India, 2022, pp. 1-5.
- [77] X. Chen, K. J. Tseng and G. Amaratunga, "State estimation for distribution systems using micro-synchrophasors," 2015 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Brisbane, QLD, Australia, 2015, pp. 1-5.
- [78] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan and F. Bu, "A Survey on State Estimation Techniques and Challenges in Smart Distribution Systems," in *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2312-2322, March 2019.
- [79] Vijaychandra, J.; Prasad, B.R.V.; Darapureddi, V.K.; Rao, B.V.; Knypiński, Ł. A Review of Distribution System State Estimation Methods and Their Applications in Power Systems. *Electronics* 2023, 12, 603.

- [80] Vijaychandra, J.; Prasad, B.R.V.; Darapureddi, V.K.; Rao, B.V.; Knypiński, Ł. A Review of Distribution System State Estimation Methods and Their Applications in Power Systems. *Electronics* 2023, 12, 603.
- [81] Introduction to Probability, Statistics, and Random Processes, Hossein Pishro-Nik, Kappa Research, LLC, 2014.