



Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

**ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΚΡΗΤΗΣ**  
**Σχολή Τεχνολογικών Εφαρμογών**  
**Τμήμα Μηχανικών Πληροφορικής**



**Πτυχιακή Εργασία:**

Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

Σπουδαστής : Αθανασιάδης Στέφανος (ΑΜ 2493)

Επιβλέπων Καθηγητής : Δόκτωρ Τριανταφυλλίδης Γεώργιος



***Abstract***

*The purpose of this research is to find the best methods for forecasting electricity load for a power plant. To achieve the prediction modern methods of artificial intelligence algorithms and also a large variety of mathematical methods and ideas have been used for load forecasting. The development and improvements of appropriate mathematical tools will lead to the development of more accurate load forecasting techniques.*

*Load Forecasting is a necessary component of a modern power plant that is called smart grid. Load forecasts are extremely important for energy suppliers, financial institutions, and other participants in electric energy generation, transmission, distribution, and markets. Short-term load forecasting can help to estimate load flows and to make decisions that can prevent overloading. Timely implementations of such decisions lead to the improvement of network reliability and to the reduced occurrences of equipment failures and blackouts.*



## **Περίληψη**

*Ο σκοπός της έρευνας είναι η εύρεση των βέλτιστων μεθόδων για την πρόβλεψη του ηλεκτρικού φορτίου που πρέπει να παραγάγει μια μονάδα παραγωγής ηλεκτρικού φορτίου. Για να επιτευχθεί η πρόβλεψη χρησιμοποιούνται σύγχρονοι μέθοδοι και αλγόριθμοι τεχνητής νοημοσύνης, μια μεγάλη ποικιλία από μαθηματικές μεθόδους και ιδέες. Η ανάπτυξη και η βελτίωση της κατάλληλων μαθηματικών εργαλείων, θα οδηγήσουν στην ανάπτυξη πιο ακριβών τεχνικών πρόβλεψη φορτίου*

*Είναι ένα απαραίτητο τμήμα μιας σύγχρονης μονάδας παραγωγής ηλεκτρικού φορτίου που ονομάζεται έξυπνο δίκτυο. Οι προβλέψεις φορτίου είναι εξαιρετικά σημαντικές για τους προμηθευτές ενέργειας, τα χρηματοπιστωτικά ιδρύματα (και άλλους συμμετέχοντες), στην παραγωγή, στη μεταφορά, στη διανομή και στις αγορές ηλεκτρικής ενέργειας. Η Βραχυπρόθεσμη πρόβλεψη φορτίου, λαμβάνεται υπόψη για να βοηθήσει στην εκτίμηση των ροών φορτίου και στη λήψη αποφάσεων που μπορούν να αποτρέψουν την υπερφόρτωση. Έγκαιρες υλοποιήσεις τέτοιων αποφάσεων οδηγούν στη βελτίωση της αξιοπιστίας του δικτύου, στις μειωμένες εμφανίσεις βλαβών του εξοπλισμού και διακοπές ρεύματος.*



## Περιεχόμενα

1. Εισαγωγή – Ηλεκτρισμός και Παραγωγή Ηλεκτρικού Φορτίου .....	6
1.1 Ιστορία .....	6
1.2 Απορρύθμιση.....	7
1.3 Ο ορισμός του «Δικτύου» (Grid).....	8
1.3.1 Παροχή ηλεκτρικού ρεύματος .....	9
1.3.2 Χαρακτηριστικά και Δομή των δικτύων διανομής .....	10
1.3.3 Γεωγραφία των δικτύων μεταφοράς .....	11
1.3.4 Ανανεώσιμες πηγές ενέργειας .....	12
1.4 Γήρανση υποδομών .....	16
1.5 Σύγχρονες Τάσεις .....	17
1.6 Μελλοντικές Τάσεις .....	18
1.7 Αναδύομενο Έξυπνο Δίκτυο (Smart Grid) .....	19
1.7.1 Σύστημα διαχείρισης Διανομής .....	20
1.7.2 Γιατί Σύστημα διαχείρισης Διανομής (DMS) .....	21
1.7.3 Λειτουργίες Συστήματος διαχείρισης Διανομής .....	22
1.7.4 Ανάλυση Συνδεσιμότητας Δικτύου (EAA) .....	22
1.7.5 Πρόγραμμα Εναλλαγής & Διαχείρισης της Ασφάλειας .....	23
1.7.6 Εκτίμηση κράτους (SE) .....	23
1.7.7 Εφαρμογές ροής φορτίου (LFA) .....	24
1.7.8 Έλεγχος Volt-VAR (VVC) .....	25
1.7.9 Εφαρμογή Περικοπής Φορτίου (LSA) .....	25
1.7.10 Διαχείριση Βλάβης & Αποκατάσταση Συστήματος.....	26
1.7.11 Εξισορρόπηση φορτίου μέσω τροφοδότη Αναδιάρθρωσης.....	26
1.7.12 Πρότυπα με βάση την ένταξη.....	27
1.7.13 IEC 61968 πρότυπα με βάση την ολοκλήρωση.....	27
1.7.14 Διανομή Προβλεπόμενου Φορτίου (DLF).....	28
2. Αναγνώριση Προτύπων & Μηχανική μάθηση.....	30
2.1 Εισαγωγή.....	30
2.1.1 Βασική Ορολογία.....	30
2.1.2 Τυπική Διαδικασία.....	30
2.2 Αλγόριθμοι.....	31
2.2.1 Κανόνες Κατηγοριοποίησης & κατηγοριοποίηση στην Μηχανική μάθηση.....	32
2.2.2 Σχέση με άλλα προβλήματα.....	33
2.2.3 Μπεϋζιανή Κατηγοριοποίηση.....	33
2.2.3 Διανύσματα Χαρακτηριστικών.....	34
2.2.4 Δυαδική και πολλών κλάσεων Κατηγοριοποίηση.....	34
2.2.5 Γραμμικοί Ταξινομητές.....	34
2.3 Δέντρα Αποφάσεων (Decision Trees).....	35
2.3.1 Λειτουργία της μεθόδου Fit.....	38
2.3.2 Βελτίωση των classification & regression trees.....	38
2.4 Μάθηση Συνόλου (Ensemble Learning).....	42
2.4.1 Επισκόπηση.....	42
2.4.2 Θεωρία Συνόλου.....	42
2.4.3 Τύποι των συνόλων.....	43
2.5 Τυχαία Δάση (Random Forests).....	44
2.5.1 Λειτουργία Αλγορίθμου Random Forest.....	44
2.6 Μηχανή Υποστήριξης Διανυσμάτων (Support Vector Machine - SVM).....	47



Τμήμα Μηχανικών Πληροφορικής	
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου	
2.6.1 Διαχωρίσιμα δεδομένα.....	47
2.6.2 Μη διαχωρίσιμα δεδομένα.....	49
2.6.3 Μη γραμμικές μετατροπές με Πυρήνες (Kernel).....	51
3. Πρόβλεψη Ηλεκτρικού Φορτίου (Load Forecasting).....	52
3.1 Εισαγωγή.....	52
3.2 Σημαντικοί παράγοντες για προβλέψεις.....	53
3.3 Μέθοδοι Πρόβλεψης.....	54
3.4 Μέθοδοι Μεσαίας-και μακροπρόθεσμης πρόβλεψης φορτίου.....	55
3.5 Μέθοδοι Βραχυπρόθεσμης πρόβλεψης φορτίου.....	59
3.6 Μελλοντικές ερευνητικές κατευθύνσεις.....	63
3.7 Συμπέρασμα.....	64
4. Υλοποίηση.....	65
4.1 Χρησιμότητα του εργαλείου για την πρόβλεψη φορτίου.....	65
4.2 Στόχος.....	65
4.3 Απαιτήσεις.....	66
4.4 Προκλήσεις στην ενσωμάτωση ενός συστήματος πρόβλεψης φορτίου. (Από τρίτους).....	66
4.5 Λύσεις στην ενσωμάτωση ενός συστήματος πρόβλεψης φορτίου. (Από τρίτους).....	67
4.6 Γενική Ροή Μοντέλου Ανάπτυξης.....	68
4.7 Αρχιτεκτονική Μοντέλου.....	69
4.7.1 Δεδομένα.....	69
4.7.2 Σύνοψη Δεδομένων.....	72
4.7.3 Παραγωγή Προγνωστικών Παραγόντων & Επιλογή Βραχυπρόθεσμης ή Μακροπρόθεσμης Πρόβλεψης & Χωρισμός σε σετ δεδομένων εκπαίδευσης και επαλήθευσης.....	73
4.7.4 Επιλογή & Ρύθμιση Μοντέλου Πρόβλεψης.....	74
4.7.5 Σύνοψη Επιλογής & Ρύθμισης Μοντέλου Πρόβλεψης.....	78
4.7.6 Παρουσίαση αποτελεσμάτων.....	78
4.7.7 Σύνοψη Παρουσίαση Αποτελεσμάτων.....	85
4.8 Γενικό Συμπέρασμα.....	86
5. Βιβλιογραφία & Πηγές.....	87
6. Παράρτημα (Appendix).....	88
6.1 Παράρτημα α) Κώδικας για την δοκιμή διαφόρων τιμών για την μεταβλητή που καθορίζει τον ελάχιστο αριθμό δεδομένων ανά φύλλο.....	88
6.2 Παράρτημα 1: Κώδικας εισαγωγής Δεδομένων ( import.m).....	88
6.3 Παράρτημα 2: Κώδικας Χωρισμός σε σετ δεδομένων εκπαίδευσης και επαλήθευσης.....	89
6.4 Παράρτημα 3: κώδικας παραγωγής των νέων προγνωστικών παραμέτρων... ..	89
6.5 Παράρτημα 4: Κώδικας του μοντέλου του νευρωνικού δικτύου.....	90
6.6 Παράρτημα 5: Προσδιορισμός αριθμού των δένδρων και κατάλληλου μεγέθους Φύλλου.....	90
6.7 Παράρτημα 6: Καθορισμός σημαντικότητας των χαρακτηριστικών.....	91
6.8 Παράρτημα 7: Εύρεση βέλτιστου C και γ μέσα από το Search grid cross validation. & εκπαίδευση μοντέλου χρησιμοποιώντας αυτές τις τιμές.....	91
6.9 Παράρτημα 8: Υπολογισμός Σφάλματος, Ποσοστό Σφάλματος, και Μέγιστου ημερήσιου Μέσου απόλυτου Ποσοστιαίου Σφάλματος.....	92



## 1. Εισαγωγή – Ηλεκτρισμός και Παραγωγή Ηλεκτρικού Φορτίου

### 1.1 Ιστορία

Από την ίδρυση της παραγωγής ηλεκτρικού φορτίου στην βιομηχανική εποχή, το ηλεκτρικό δίκτυο είναι η εξέλιξη ενός νησιωτικού συστήματος, το οποίο εξυπηρετούσε μια συγκεκριμένη γεωγραφική περιοχή σε ένα ευρύτερο, επεκτατικό δίκτυο που ενσωμάτωσε πολλές περιοχές. Σε ένα χρονικό διάστημα, όλη η ενέργεια παραγόταν κοντά στη συσκευή. Στις αρχές του 19ου αιώνα, η ηλεκτρική ενέργεια ήταν μια νέα εφεύρεση, που αγωνίστηκε με ατμό, υδραυλικά, άμεση θέρμανση και ψύξη, φως, και κυρίως φυσικό αέριο. Κατά τη διάρκεια αυτής της περιόδου, η παραγωγή φυσικού αερίου και παράδοσή του είχε γίνει το πρώτο κεντρικό στοιχείο στη σύγχρονη ενεργειακή βιομηχανία. Ήταν το πρώτο που παράγεται στις εγκαταστάσεις του πελάτη, αλλά αργότερα εξελίχθηκε σε μεγάλο παραγωγό αερίου που απολαμβάνει οικονομίες κλίμακας. Σχεδόν κάθε πόλη στις ΗΠΑ και την Ευρώπη το φυσικό αέριο της πόλης, είχε διοχετευθεί μέσω των δήμων τους, δεδομένου ότι ήταν μια κυρίαρχη μορφή της χρήσης της ενέργειας των νοικοκυριών. Μέχρι τα μέσα του 19ου αιώνα, το ηλεκτρικό τόξο σε μικρό χρονικό διάστημα έγινε συμφέρουσα μέθοδος, σε σύγκριση με τους λαμπτήρες πηκτικών αερίων και τους λαμπτήρες αερίου που παράγουν κακό φως και τεράστια σπατάλη θερμότητας, οι οποίες έκαναν τα δωμάτια ζεστά και «καπνιστά», και δημιουργούσαν επιβλαβή στοιχεία, με τη μορφή υδρογόνου και μονοξειδίου του άνθρακα. Μοντελοποιώντας μετά τη βιομηχανία φωτισμού αερίου, ο Thomas Edison εφήυρε το πρώτο ηλεκτρικό σύστημα που παρέχεται ενέργεια μέσω εικονικών δικτύου σε διήθηση του φωτός σε αντίθεση με καυστήρες φυσικού αερίου. Με αυτό, τα συστήματα παραγωγής ηλεκτρικής ενέργειας επωφελήθηκαν από τις οικονομικές κλίμακες και μετακόμισε στην κεντρική παραγωγή ηλεκτρικής ενέργειας, τη διανομή και τη διαχείριση του συστήματος.

Κατά τη διάρκεια του 20ου αιώνα, η θεσμική ρύθμιση της παραγωγής ηλεκτρικής ενέργειας άλλαξε. Στην αρχή, ήταν απομονωμένα συστήματα χωρίς σύνδεση με άλλες επιχειρήσεις κοινής ωφέλειας και εξυπηρετούσαν μια συγκεκριμένη περιοχή. Στη δεκαετία του 1920, επιχειρήσεις κοινής ωφέλειας ενώνονται για την ίδρυση ενός ευρύτερου δικτύου. Ως κοινές επιχειρήσεις είδαν τα οφέλη από την ανταλλαγή φορτίου αιχμής κάλυψη και δύναμη δημιουργίας αντιγράφων ασφαλείας. Επίσης, οι εταιρείες παραγωγής ηλεκτρικής ενέργειας έπαιρναν χρηματοδοτήσεις από Wall Street ιδιώτες-επενδυτές που υποστηρίζονταν από πολλές επιχειρήσεις τους. Το 1934, με το πέρασμα της Κοινοφελούς Επιχείρησης Holding Company Act (ΗΠΑ), η παραγωγή ηλεκτρικής ενέργειας αναγνωρίστηκε ως δημόσιο αγαθό μαζί με το φυσικό αέριο, το νερό και το τηλέφωνο. Ως αποτέλεσμα έθεσαν στις εταιρείες περιορισμούς και ρυθμιστική εποπτεία των δραστηριοτήτων τους. Αυτό αποτέλεσε την απαρχή της Χρυσής Εποχής του κανονισμού για περισσότερα από 60 χρόνια.

Ωστόσο, με το πέρασμα των χρόνων και την ανάπτυξη των βιομηχανιών τηλεπικοινωνιών στα τέλη της δεκαετίας του 1970, ο Ενεργειακός Πολιτικός Νόμος



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου (EPAct) του 1992 υποστήριξε την απελευθέρωση της παραγωγής ηλεκτρικής ενέργειας με τη δημιουργία χονδρικής ηλεκτρικών αγορών. Απαιτούσε από τους ιδιοκτήτες γραμμής μεταφοράς να επιτρέψουν τις ηλεκτρικές εταιρείες παραγωγής την ανοικτή πρόσβαση στο δίκτυό τους.

## 1.2 Απορρύθμιση

Η απορρύθμιση έγινε σε αντίθεση με το παραδοσιακό καθετοποιημένο μονοπώλιο που επιβλέπει τις εργασίες του, σε όλο το πλέγμα του. Νεώτεροι συμμετέχοντες εισήλθαν στην αγορά, συμπεριλαμβανομένων των ανεξάρτητων φορέων παροχής ρεύματος (IPPs), που αποφάσισαν να κατασκευάσουν τη νέα εγκατάσταση. Οι Εταιρείες Μεταφοράς (TRANSCOs) οι οποίες κατασκεύασαν και παρείχαν τον εξοπλισμό μετάδοσης. Οι λιανοπωλητές που προμηθεύουν τους τελικούς πελάτες, κατασκεύασαν την ηλεκτρική υπηρεσία τους και τιμολόγησαν ολοκληρωμένες επιχειρήσεις ενέργειας. Επίσης, με το πέρασμα του χρόνου οι μακροπρόθεσμες λειτουργίες άλλαξαν μορφή. Οι προσθήκες στην υποδομή που είχαν μακροπρόθεσμο σχεδιασμό, μετάλλαξαν στην κατασκευή μιας νέας μονάδας παραγωγής ενέργειας, με οικονομικά ζητήματα (φορολογία, το κόστος εργασίας και υλικών) και απέκτησαν τη δυνατότητα να λάβουν χρηματοδότηση. Το τμήμα διαχείρισης της προμήθειας το οποίο ήταν στο πλαίσιο του μεσοπρόθεσμου σχεδιασμού, μετατράπηκε σε τμήμα διαχείρισης κινδύνων. Ως ιδιωτικές επιχειρήσεις κοινής ωφέλειας έπρεπε να διαχειριστούν τα περιουσιακά στοιχεία των τελικών πελατών, με την προτίμηση του κινδύνου της εταιρείας. Ο προγραμματισμός της επόμενης μέρας και η διαχείριση του δικτύου σε πραγματικό χρόνο στο βραχυπρόθεσμο σχεδιασμό που περιλαμβάνει την πρόβλεψη της ζήτησης και το χρονοδιάγραμμα αποστολής έγιναν στοιχεία διαχείρισης, όπως σταθμοί παραγωγής ηλεκτρικής ενέργειας και εξοπλισμοί δικτύου ήταν στοιχεία για να μπορούν να προγραμματιστούν. Εδώ, το ISO θέτει το χρονοδιάγραμμα αποστολής στην τιμή εκκαθάρισης της αγοράς, όπου οι προσφορές προμήθειας μονάδων παραγωγής εξισορροπούνται με τις προσφορές ζήτησης των λιανοπωλητών.

Πολλοί μηχανικοί υποστηρίζουν τα ατυχή μειονεκτήματα που απορρέουν από την απορρύθμιση. Στη περίπτωση των ρυθμιζόμενων μονοπωλίων, χρησιμοποιήθηκαν γραμμές ενέργειας μεγάλων αποστάσεων για καταστάσεις έκτακτης ανάγκης ως εφεδρικές σε περίπτωση διακοπών, τώρα, ιδιαίτερα στη Βόρεια Αμερική, η πλειοψηφία της εγχώριας παραγωγής πωλείται πάνω από συνεχώς αυξανόμενες αποστάσεις σχετικά με τη χονδρική αγορά πριν από την παράδοση στους πελάτες. Ως εκ τούτου, στο δίκτυο ηλεκτρικής ενέργειας υπάρχουν διακυμάνσεις ροών ηλεκτρικής ενέργειας που έχουν επιπτώσεις στη σταθερότητα και στην αξιοπιστία του συστήματος. Για να μειωθούν τα σφάλματα του συστήματος, η ροή ισχύος μιας γραμμής μεταφοράς πρέπει να λειτουργεί κάτω από την ικανότητα της γραμμής μεταφοράς του. Επιπλέον, όπως οι επιχειρήσεις κοινής ωφέλειας ανταλλάσσουν ισχύς μεταξύ τους, και η ενέργεια ρέει κατά μήκος όλων των



Τμήμα Μηχανικών Πληροφορικής

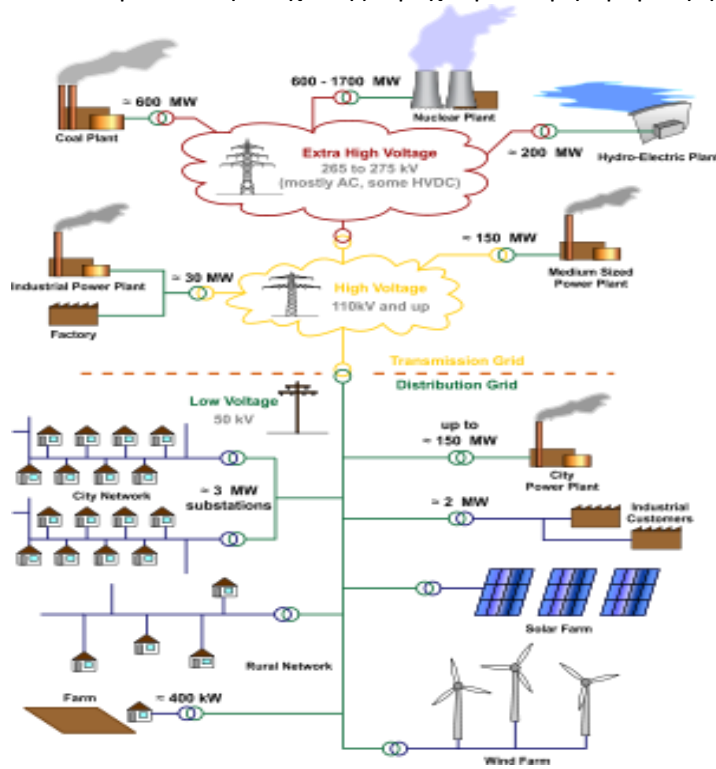
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

διαδρομών της σύνδεσης. Ως εκ τούτου, οποιαδήποτε αλλαγή σε ένα σημείο της παραγωγής και μεταφοράς επηρεάζει το φορτίο σε όλα τα άλλα σημεία. Πολλές φορές, αυτό είναι απρόβλεπτο και ανεξέλεγκτο. Συνήθως, η χωρητικότητα μιας μεγάλης γραμμής είναι μικρότερη από την χωρητικότητα μιας μικρότερης γραμμής. Αν όχι, εμφανίζεται αστάθεια τροφοδοσίας που έχει ως αποτέλεσμα το σπάσιμο ή κρέμασμα των γραμμών μεταφοράς. Τέτοια φάση και διακυμάνσεις τάσης προκαλούν διακοπές στο σύστημα όπως μαρτυρούν στα βορειοανατολικά Blackout του 1965 και το 2003 (η οποία αφορούσε μια χαλαρωμένη γραμμή σε ένα δέντρο). Επιπλέον, οι IPPs προσθέτουν νέες μονάδες παραγωγής σε συγκεκριμένες θέσεις μακριά από τις κεντρικές καταναλωτικές περιοχές επηρεάζοντας πολύ την παροχή ηλεκτρικού ρεύματος. Επίσης, οι επιχειρήσεις κοινής ωφέλειας, λόγω των ανταγωνιστικών αναγκών πληροφόρησης, δεν δημοσιεύουν απαιτητικά στοιχεία για την πρόβλεψη και την αντίδραση στο στρες συστήματος, όπως τις ροές ενέργειας και στατιστικά στοιχεία διακοπών ρεύματος. Συνολικά, τα οικονομικά του ηλεκτρικού δικτύου δεν ευθυγραμμίζονται επαρκώς με τη φυσική του δικτύου. Οι ειδικοί υποστηρίζουν ότι πρέπει να γίνουν θεμελιώδεις αλλαγές για να αποφευχθούν σοβαρές συνέπειες στο εγγύς μέλλον.

### 1.3 Ο ορισμός του "Δικτύου" (Grid)

Μια πόλη έχει επιτύχει σύνδεση με το δίκτυο όταν είναι συνδεδεμένο με πολλές περιττές πηγές και όταν περιλαμβάνει μεγάλες αποστάσεις μετάδοσης. Τα υφιστάμενα εθνικά ή περιφερειακά δίκτυα παρέχουν απλώς τη διασύνδεση των εγκαταστάσεων. Το ακριβές στάδιο της ανάπτυξης του δικτύου είναι η δημιουργία ενός σωστού πλέγματος για τον εφοδιασμό ηλεκτρικού ρεύματος στους καταναλωτές. Παρομοίως, ο όρος εθνικό πλέγμα είναι αναχρονιστικά σε πολλά μέρη του κόσμου, όπως καλώδια μεταφοράς τώρα διασχίζουν συχνά τα εθνικά σύνορα. Οι όροι δίκτυο διανομής για τις τοπικές συνδέσεις και δίκτυο μεταφοράς για μεγάλες αποστάσεις μεταδόσεις ως εκ τούτου προτιμώνται, αλλά εθνικό δίκτυο συχνά χρησιμοποιείται ακόμα για τη συνολική δομή. Ηλεκτρικό δίκτυο είναι ένα διασυνδεδεμένο δίκτυο για τη μεταφορά ηλεκτρικής ενέργειας από του παραγωγούς στους καταναλωτές. Συνίσταται από τρία μέρη:





Σχεδιάγραμμα ηλεκτρικού δικτύου

Παραγωγή Ηλεκτρικής Ενέργειας: Τα εργοστάσια παραγωγής ηλεκτρικής ενέργειας παράγουν ηλεκτρική ενέργεια χρησιμοποιώντας ορυκτά καύσιμα (άνθρακα, φυσικό αέριο, βιομάζα) ή τον αέρα, το νερό, τα πυρηνικά καύσιμα και τον ήλιο,

Μεταφορά Ηλεκτρικής Ενέργειας: Οι γραμμές μεταφοράς μεταφέρουν την ηλεκτρική ενέργεια από τα εργοστάσια στα κέντρα ζήτησης,

Διανομή Ηλεκτρικής Ενέργειας: Η ενέργεια φτάνει στους υποσταθμούς, γίνεται υποβιβασμός τάσης με τη βοήθεια μετασχηματιστών και μεταφέρεται μέσω γραμμών διανομής. με περαιτέρω υποβιβασμό τάσης, γίνεται κατάλληλη για χρήση από οικιακά δίκτυα.

### 1.3.1 Παροχή ηλεκτρικού ρεύματος

Η διανυσματική μορφή σε τριγωνομετρικούς κύκλους των φάσεων εμφανίζεται από αριστερά προς τα δεξιά: μονοφασική παροχή, τριφασική παροχή, εξαφασική παροχή.

Η παροχή του ηλεκτρικού ρεύματος γίνεται από το υφιστάμενο ηλεκτρικό δίκτυο ή δίκτυο ηλεκτροδότησης. Γενικά, το παρεχόμενο ηλεκτρικό ρεύμα είναι εναλλασσόμενο ημιτονοειδές ηλεκτρικό ρεύμα, ενεργής τάσης 230V και συχνότητας 50hz. Λιγότερο διαδεδομένο είναι το ηλεκτρικό ρεύμα με χαρακτηριστικά 110V και συχνότητα 60hz, που χρησιμοποιείται στις Ηνωμένες Πολιτείες Αμερικής.



## Τμήμα Μηχανικών Πληροφορικής

### Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

Σε κάθε καταναλωτή παρέχεται πάντα ένα τριπλό καλώδιο, που το αποτελούν το ουδέτερο (Ο) στον οποίο δεν διαπερνά το ρεύμα, και οι φάσεις, στα οποία παρέχεται το ρεύμα, ενώ ο καταναλωτής είναι υποχρεωμένος να παρέχει για ασφάλεια μία γείωση. Ανάλογα με τις φάσεις, η παροχή είναι μονοφασική ή τριφασική.

- Μονοφασική: Στον καταναλωτή παρέχεται μία φάση (R). Αν συνδεθεί μέσω ενός κυκλώματος η φάση με τον ουδέτερο τότε δημιουργείται κύκλωμα εναλλασσόμενου ρεύματος.
- Τριφασική: Στον καταναλωτή παρέχονται τρεις φάσεις (R,S,T). Η κάθε φάση διαφέρει από τις άλλες δύο κατά 120 μοίρες. Έτσι, αν συνδεθούν δύο φάσεις θα προκύψει ρεύμα, ενώ θα παρουσιαστεί και βραχυκύκλωμα εν μέρει. Αν συνδεθούν και οι τρεις φάσεις θα δημιουργηθεί πλήρως βραχυκύκλωμα και το τελικό καλώδιο δε θα φέρει καθόλου ρεύμα. Συνήθως η ηλεκτρολογική εγκατάσταση χωρίζεται σε τρία μέρη, όπου κάθε μέρος ηλεκτροδοτείται από μία φάση, όπως και στη μονοφασική παροχή.

Η διαφορά φάσης μεταξύ δύο διαδοχικών φάσεων είναι συγκεκριμένη και χαρακτηριστική της συγκεκριμένης παροχής. Γενικά, η διαφορά φάσης δύο

$$\Delta\phi = \frac{360^\circ}{\nu}$$

διαδοχικών φασών σε μια παροχή  $\nu$  φάσεων δίνεται από τον τύπο:

Στην Ελλάδα σήμερα το ηλεκτρικό ρεύμα παρέχεται από τη ΔΕΗ, νομικό πρόσωπο μεικτού δικαίου και υπηρεσία κοινής ωφελείας (ΔΕΚΟ).

Τα τελευταία 20 χρόνια γίνονται έντονες προσπάθειες αύξησης του ποσοστού ηλεκτρικής ενέργειας που παράγεται με τη χρήση Ανανεώσιμων Πηγών Ενέργειας (Α.Π.Ε.).

### 1.3.2 Χαρακτηριστικά και Δομή των δικτύων διανομής

Η δομή, ή "τοπολογία" ενός πλέγματος μπορεί να ποικίλλει σημαντικά. Η φυσική διάταξη είναι συχνά αναγκασμένη από τις διαθέσιμες περιοχές και τη γεωλογία τους. Η λογική τοπολογία μπορεί να ποικίλει ανάλογα με τους περιορισμούς του προϋπολογισμού, τις απαιτήσεις για την αξιοπιστία του συστήματος, καθώς και τα χαρακτηριστικά του φορτίου και την παραγωγή.

Η φθηνότερη και απλούστερη τοπολογία για ένα δίκτυο διανομής ή μετάδοσης είναι η ακτινική δομή. Αυτή είναι ένα σχήμα δέντρου όπου η διανομή, από μια μεγάλη προσφορά, ακτινοβολεί έξω σε προοδευτική μείωση της τάσης έως ότου τροφοδοτηθούν τα σπίτια και οι επιχειρήσεις.

Η δαπάνη των τοπολογιών τύπου ματιού περιορίζουν την εφαρμογή τους στα δίκτυα μεταφοράς και δικτύων διανομής μέσης τάσης. Ο πλεονασμός επιτρέπει αποτυχία της γραμμής να συμβεί και το δρομολόγιο της ισχύος αλλάζει, ενώ εργάτες επισκευάσουν την κατεστραμμένη και απενεργοποιημένη γραμμή.

Άλλες τοπολογίες που χρησιμοποιούνται είναι συστήματα τύπου βρόγχου που βρέθηκαν στην Ευρώπη και δεσμευμένα δίκτυα δακτυλίου. Σε πόλεις και κωμοπόλεις



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

της Βόρειας Αμερικής, το δίκτυο τείνει να ακολουθήσει την κλασική ακτινική σχεδίαση τροφοδοσίας. Ένας υποσταθμός λαμβάνει την ισχύ του από το δίκτυο μεταφοράς, η ισχύς κατεβαίνει με ένα μετασχηματιστή και αποστέλλεται σε ένα δίαυλο και από κει τροφοδοτές διασκορπίζουν προς όλες τις κατευθύνσεις σε όλη την ύπαιθρο. Αυτοί οι τροφοδοτές μεταφέρουν τριφασικό ρεύμα, και τείνουν να ακολουθούν τους κεντρικούς δρόμους κοντά στον υποσταθμό.. Αυτή η δομή τύπου δέντρου αναπτύσσεται προς τα έξω από τον υποσταθμό, αλλά για λόγους αξιοπιστίας, συνήθως περιέχει τουλάχιστον μια αχρησιμοποίητη εφεδρική σύνδεση σε ένα κοντινό υποσταθμό. Αυτή η σύνδεση μπορεί να ενεργοποιηθεί σε περίπτωση έκτακτης ανάγκης, έτσι ώστε ένα τμήμα του εδάφους υπηρεσίας ενός υποσταθμού να μπορεί εναλλακτικά να τροφοδοτείται από ένα άλλο υποσταθμό.

### 1.3.3 Γεωγραφία των δικτύων μεταφοράς

Τα δίκτυα μετάδοσης είναι πιο περίπλοκα με περιττά μονοπάτια. Ένα συγχρονισμένο πλέγμα μεγάλης περιοχής ή «διασύνδεσης» είναι μια ομάδα από περιοχές διανομής που όλα λειτουργούν με εναλλασσόμενο ρεύμα (AC), συχνότητες συγχρονισμένες (έτσι ώστε οι κορυφές να συμβαίνουν την ίδια στιγμή). Αυτό επιτρέπει την μετάδοση του εναλλασσόμενου ρεύματος σε όλη την περιοχή, που συνδέει ένα μεγάλο αριθμό των παραγωγών ηλεκτρικής ενέργειας και των καταναλωτών και, ενδεχομένως, επιτρέποντας πιο αποτελεσματικές αγορές ηλεκτρικής ενέργειας.

Σε ένα σύγχρονο δίκτυο όλες οι γεννήτριες λειτουργούν όχι μόνον με την ίδια ταχύτητα, αλλά επίσης στην ίδια φάση, κάθε γεννήτρια διατηρείται από ένα τοπικό κυβερνήτη που ρυθμίζει την ροπή οδήγησης κατά τον έλεγχο της τροφοδοσίας ατμού προς τον στρόβιλο οδήγησης. Η παραγωγή και κατανάλωση πρέπει να εξισορροπούνται σε ολόκληρο το πλέγμα, επειδή η ενέργεια καταναλώνεται σχεδόν ακαριαία καθώς παράγεται. Ενέργεια αποθηκεύεται στο άμεσο βραχυπρόθεσμα από την περιστροφική κινητική ενέργεια των γεννητριών.

Μια μεγάλη αποτυχία σε ένα τμήμα του δικτύου - εκτός αν αποτραπεί γρήγορα - μπορεί να προκαλέσει αλλαγή της διαδρομής του ρεύματος και να ρέει από τις υπόλοιπες γεννήτριες στους καταναλωτές μέσω των γραμμών μετάδοσης λόγω ανεπαρκούς χωρητικότητας, προκαλώντας περαιτέρω αποτυχίες. Ένα δίκτυο με κακή αρχιτεκτονική τοπολογίας έχει ως αποτέλεσμα τη διαδοχική αποτυχία και εκτεταμένη διακοπή ρεύματος. Μια κεντρική αρχή συνήθως ορίζεται για να διευκολύνει την επικοινωνία και την ανάπτυξη πρωτοκόλλων για να διατηρηθεί ένα σταθερό πλέγμα. Για παράδειγμα, η North American Electric Reliability Corporation απέκτησε δεσμευτικές εξουσίες στις Ηνωμένες Πολιτείες το 2006, και έχει συμβουλευτικές αρμοδιότητες στα σχετικά μέρη του Καναδά και το Μεξικό



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

Ορισμένες περιοχές, για παράδειγμα οι αγροτικές κοινότητες στην Αλάσκα, δεν λειτουργούν σε ένα μεγάλο πλέγμα, αλλά χρησιμοποιούν τοπικές γεννήτριες ντίζελ. Γραμμές συνεχούς ρεύματος υψηλής τάσης ή μετασηματιστές μεταβλητής συχνότητας μπορούν να χρησιμοποιηθούν για τη σύνδεση δύο δικτύων εναλλασσόμενου ρεύματος διασύνδεσης τα οποία δεν συγχρονίζονται μεταξύ τους. Αυτό παρέχει το πλεονέκτημα της διασύνδεσης χωρίς την ανάγκη να συγχρονίσει μια ακόμη ευρύτερη περιοχή.

### 1.3.4 Ανανεώσιμες πηγές ενέργειας

Οι ανανεώσιμες μορφές ενέργειας (ΑΠΕ) ή ήπιες μορφές ενέργειας, ή νέες πηγές ενέργειας, είναι μορφές εκμεταλλεύσιμης ενέργειας που προέρχονται από διάφορες φυσικές διαδικασίες, όπως ο άνεμος, η γεωθερμία, η κυκλοφορία του νερού και άλλες. Ο όρος «ήπιες» αναφέρεται σε δυο βασικά χαρακτηριστικά. Καταρχάς, για την εκμετάλλευσή τους δεν απαιτείται κάποια ενεργητική παρέμβαση, όπως εξόρυξη, άντληση ή καύση, όπως με τις μέχρι τώρα χρησιμοποιούμενες πηγές ενέργειας, αλλά η εκμετάλλευση της ήδη υπάρχουσας ροής ενέργειας στη φύση. Δεύτερον, πρόκειται για «καθαρές» μορφές ενέργειας, πολύ «φιλικές» στο περιβάλλον, που δεν αποδεδεσμεύουν υδρογονάνθρακες, διοξείδιο του άνθρακα ή τοξικά και ραδιενεργά απόβλητα, όπως οι υπόλοιπες πηγές ενέργειας που χρησιμοποιούνται σε μεγάλη κλίμακα. Έτσι οι ΑΠΕ θεωρούνται από πολλούς μία αφετηρία για την επίλυση των οικολογικών προβλημάτων που αντιμετωπίζει η Γη.

Ως «ανανεώσιμες πηγές» θεωρούνται γενικά οι εναλλακτικές των παραδοσιακών πηγών ενέργειας (π.χ. του πετρελαίου ή του άνθρακα), όπως η ηλιακή και η αιολική. Ο χαρακτηρισμός «ανανεώσιμες» είναι κάπως καταχρηστικός, μιας και ορισμένες από αυτές τις πηγές, όπως η γεωθερμική ενέργεια δεν ανανεώνονται σε κλίμακα χιλιετιών. Σε κάθε περίπτωση οι ΑΠΕ έχουν μελετηθεί ως λύση στο πρόβλημα της αναμενόμενης εξάντλησης των (μη ανανεώσιμων) αποθεμάτων ορυκτών καυσίμων. Τελευταία από την Ευρωπαϊκή Ένωση, αλλά και από πολλά μεμονωμένα κράτη, υιοθετούνται νέες πολιτικές για τη χρήση ανανεώσιμων πηγών ενέργειας, που προάγουν τέτοιες εσωτερικές πολιτικές και για τα κράτη μέλη. Οι ΑΠΕ αποτελούν τη βάση του μοντέλου οικονομικής ανάπτυξης της πράσινης οικονομίας και κεντρικό σημείο εστίασης της σχολής των οικολογικών οικονομικών, η οποία έχει κάποια επιρροή στο οικολογικό κίνημα.

Οι ήπιες μορφές ενέργειας βασίζονται κατ' ουσίαν στην ηλιακή ακτινοβολία, με εξαίρεση τη γεωθερμική ενέργεια, η οποία είναι ροή ενέργειας από το εσωτερικό του φλοιού της γης, και την ενέργεια απ' τις παλίρροιες που εκμεταλλεύεται τη βαρύτητα. Οι βασιζόμενες στην ηλιακή ακτινοβολία ήπιες πηγές ενέργειας είναι ανανεώσιμες, μιας και δεν πρόκειται να εξαντληθούν όσο υπάρχει ο ήλιος, δηλαδή για μερικά ακόμα δεκαετομύρια χρόνια. Ουσιαστικά είναι ηλιακή ενέργεια "συσκευασμένη" κατά τον ένα ή τον άλλο τρόπο, όπως η βιομάζα είναι ηλιακή



## Τμήμα Μηχανικών Πληροφορικής

### Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

ενέργεια δεσμευμένη στους ιστούς των φυτών μέσω της φωτοσύνθεσης. Η αιολική εκμεταλλεύεται τους ανέμους που προκαλούνται απ' τη θέρμανση του αέρα ενώ αυτές που βασίζονται στο νερό εκμεταλλεύονται τον κύκλο εξάτμισης-συμπύκνωσης του νερού και την κυκλοφορία του. Η γεωθερμική ενέργεια δεν είναι ανανεώσιμη, καθώς τα γεωθερμικά πεδία κάποια στιγμή εξαντλούνται.

Χρησιμοποιούνται είτε άμεσα (κυρίως για θέρμανση) είτε μετατρεπόμενες σε άλλες μορφές ενέργειας (κυρίως ηλεκτρισμό ή μηχανική ενέργεια). Υπολογίζεται ότι το τεχνικά εκμεταλλεύσιμο ενεργειακό δυναμικό απ' τις ήπιες μορφές ενέργειας είναι πολλαπλάσιο της παγκόσμιας συνολικής κατανάλωσης ενέργειας. Όμως η υψηλή μέχρι πρόσφατα τιμή των νέων ενεργειακών εφαρμογών, τα τεχνικά προβλήματα εφαρμογής καθώς και πολιτικές και οικονομικές σκοπιμότητες που έχουν να κάνουν με τη διατήρηση του παρόντος στάτους κβο στον ενεργειακό τομέα εμπόδισαν την εκμετάλλευση έστω και μέρους αυτού του δυναμικού.

Το ενδιαφέρον για τις ήπιες μορφές ενέργειας ανακινήθηκε τη δεκαετία του 1970, ως αποτέλεσμα κυρίως των απανωτών πετρελαϊκών κρίσεων της εποχής, αλλά και της αλλοίωσης του περιβάλλοντος και της ποιότητας ζωής από τη χρήση κλασικών πηγών ενέργειας. Ιδιαίτερα ακριβές στην αρχή, ξεκίνησαν σαν πειραματικές εφαρμογές. Σήμερα όμως λαμβάνονται υπόψη στους επίσημους σχεδιασμούς των ανεπτυγμένων κρατών για την ενέργεια και, αν και αποτελούν πολύ μικρό ποσοστό της ενεργειακής παραγωγής, ετοιμάζονται βήματα για παραπέρα αξιοποίησή τους. Το κόστος δε των εφαρμογών ήπιων μορφών ενέργειας πέφτει συνέχεια τα τελευταία είκοσι χρόνια και ειδικά η αιολική και υδροηλεκτρική ενέργεια, αλλά και η βιομάζα, μπορούν πλέον να ανταγωνίζονται στα ίσα παραδοσιακές πηγές ενέργειας όπως ο άνθρακας και η πυρηνική ενέργεια. Ενδεικτικά, στις Η.Π.Α. ένα 6% της ενέργειας προέρχεται από ανανεώσιμες πηγές, ενώ στην Ευρωπαϊκή Ένωση το 2010 το 25% της ενέργειας θα προέρχεται από ανανεώσιμες πηγές (κυρίως υδροηλεκτρικά και βιομάζα).

### Πλεονεκτήματα

- Είναι πολύ φιλικές προς το περιβάλλον, έχοντας ουσιαστικά μηδενικά κατάλοιπα και απόβλητα.
- Δεν πρόκειται να εξαντληθούν ποτέ, σε αντίθεση με τα ορυκτά καύσιμα.
- Μπορούν να βοηθήσουν την ενεργειακή αυτάρκεια μικρών και αναπτυσσόμενων χωρών, καθώς και να αποτελέσουν την εναλλακτική πρόταση σε σχέση με την οικονομία του πετρελαίου.
- Είναι ευέλικτες εφαρμογές που μπορούν να παράγουν ενέργεια ανάλογη με τις ανάγκες του επί τόπου πληθυσμού, καταργώντας την ανάγκη για τεράστιες μονάδες παραγωγής ενέργειας (καταρχήν για την ύπαιθρο) αλλά και για μεταφορά της ενέργειας σε μεγάλες αποστάσεις.
- Ο εξοπλισμός είναι απλός στην κατασκευή και τη συντήρηση και έχει μεγάλο χρόνο ζωής.
- Επιδοτούνται από τις περισσότερες κυβερνήσεις.



### Μειονεκτήματα

- Έχουν αρκετά μικρό συντελεστή απόδοσης, της τάξης του 30% ή και χαμηλότερο. Συνεπώς απαιτείται αρκετά μεγάλο αρχικό κόστος εφαρμογής σε μεγάλη επιφάνεια γης. Γι' αυτό το λόγο μέχρι τώρα χρησιμοποιούνται σαν συμπληρωματικές πηγές ενέργειας.
- Για τον παραπάνω λόγο προς το παρόν δεν μπορούν να χρησιμοποιηθούν για την κάλυψη των αναγκών μεγάλων αστικών κέντρων.
- Η παροχή και απόδοση της αιολικής, υδροηλεκτρικής και ηλιακής ενέργειας εξαρτάται από την εποχή του έτους αλλά και από το γεωγραφικό πλάτος και το κλίμα της περιοχής στην οποία εγκαθίστανται.
- Για τις αιολικές μηχανές υπάρχει η άποψη ότι δεν είναι κομψές από αισθητική άποψη κι ότι προκαλούν θόρυβο και θανάτους πουλιών. Με την εξέλιξη όμως της τεχνολογίας τους και την προσεκτικότερη επιλογή χώρων εγκατάστασης (π.χ. σε πλατφόρμες στην ανοιχτή θάλασσα).
- Για τα υδροηλεκτρικά έργα λέγεται ότι προκαλούν έκλυση μεθανίου από την αποσύνθεση των φυτών που βρίσκονται κάτω απ' το νερό κι έτσι συντελούν στο φαινόμενο του θερμοκηπίου.

### Είδη ήπιων μορφών ενέργειας

- Αιολική ενέργεια. Χρησιμοποιήθηκε παλιότερα για την άντληση νερού από πηγάδια καθώς και για μηχανικές εφαρμογές (π.χ. την άλεση στους ανεμόμυλους). Έχει αρχίσει να χρησιμοποιείται ευρέως για ηλεκτροπαραγωγή.
- Ηλιακή ενέργεια. Χρησιμοποιείται περισσότερο για θερμικές εφαρμογές (ηλιακοί θερμοσίφωνες και φούρνοι) ενώ η χρήση της για την παραγωγή ηλεκτρισμού έχει αρχίσει να κερδίζει έδαφος, με την βοήθεια της πολιτικής προώθησης των Ανανεώσιμων Πηγών Ενέργειας από το ελληνικό κράτος και την Ευρωπαϊκή Ένωση.
- Υδραυλική ενέργεια. Είναι τα γνωστά υδροηλεκτρικά έργα, που στο πεδίο των ήπιων μορφών ενέργειας εξειδικεύονται περισσότερο στα μικρά υδροηλεκτρικά. Είναι η πιο διαδεδομένη μορφή ανανεώσιμης ενέργειας.
- Βιομάζα. Χρησιμοποιεί τους υδατάνθρακες των φυτών (κυρίως αποβλήτων της βιομηχανίας ξύλου, τροφίμων και ζωοτροφών και της βιομηχανίας ζάχαρης) με σκοπό την αποδέσμευση της ενέργειας που δεσμεύτηκε απ' το φυτό με τη φωτοσύνθεση. Ακόμα μπορούν να χρησιμοποιηθούν αστικά απόβλητα και απορρίμματα. Μπορεί να δώσει βίο-αιθανόλη και βιοαέριο, που είναι καύσιμα πιο φιλικά προς το περιβάλλον από τα παραδοσιακά. Είναι μια πηγή ενέργειας με πολλές δυνατότητες και εφαρμογές που θα χρησιμοποιηθεί πλατιά στο μέλλον.
- Γεωθερμική ενέργεια. Προέρχεται από τη θερμότητα που παράγεται απ' τη ραδιενεργό αποσύνθεση των πετρωμάτων της γης. Είναι εκμεταλλεύσιμη εκεί όπου η θερμότητα αυτή ανεβαίνει με φυσικό τρόπο στην επιφάνεια, π.χ. στους



#### Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

θερμοπίδακες ή στις πηγές ζεστού νερού. Μπορεί να χρησιμοποιηθεί είτε απευθείας για θερμικές εφαρμογές είτε για την παραγωγή ηλεκτρισμού. Η Ισλανδία καλύπτει το 80-90% των ενεργειακών της αναγκών, όσον αφορά τη θέρμανση, και το 20%, όσον αφορά τον ηλεκτρισμό, με γεωθερμική ενέργεια.

- Ενέργεια από τη θάλασσα
  1. Ενέργεια από παλίρροιες. Εκμεταλλεύεται τη βαρύτητα του Ήλιου και της Σελήνης, που προκαλεί ανύψωση της στάθμης του νερού. Το νερό αποθηκεύεται καθώς ανεβαίνει και για να ξανακατέβει αναγκάζεται να περάσει μέσα από μια τουρμπίνα, παράγοντας ηλεκτρισμό. Έχει εφαρμοστεί στην Αγγλία, τη Γαλλία, τη Ρωσία και αλλού.
  2. Ενέργεια από κύματα. Εκμεταλλεύεται την κινητική ενέργεια των κυμάτων της θάλασσας.
  3. Ενέργεια από τους ωκεανούς. Εκμεταλλεύεται τη διαφορά θερμοκρασίας ανάμεσα στα στρώματα του ωκεανού, κάνοντας χρήση θερμικών κύκλων. Βρίσκεται στο στάδιο της έρευνας.

### Αιολικό πάρκο

Το αιολικό πάρκο είναι μεγάλη περιοχή με εγκατεστημένες πολλές ανεμογεννήτριες που παράγουν ρεύμα περιστρεφόμενες από την ενέργεια του ανέμου, προκειμένου να τροφοδοτήσουν μία κατοικημένη περιοχή, είτε είναι μία πόλη, είτε ένα χωριό. Το αιολικό πάρκο δε μολύνει την ατμόσφαιρα με διοξείδιο του άνθρακα ή άλλα αέρια που συμβάλλουν στο φαινόμενο του θερμοκηπίου.



### Πλεονεκτήματα

- Τα αιολικά πάρκα δε μολύνουν την ατμόσφαιρα με επιβλαβή αέρια.
- Παράγουν ρεύμα από μία ανανεώσιμη πηγή ενέργειας.



### Μειονεκτήματα

- Κάνουν θόρυβο.
- Μπορεί τα πτερύγια των ανεμογεννητριών να σκοτώσουν πτηνά.
- Δεν παράγουν τόσο ρεύμα όσο ένα ατμοηλεκτρικό εργοστάσιο.
- Υπάρχει μεγάλο κόστος και χρειάζεται μεγάλη έκταση για να κατασκευαστεί ένα αιολικό πάρκο.
- Χρειάζεται άνεμο για να παράγουν ρεύμα και σε μία περιοχή δεν φυσάει συνέχεια όλο το χρόνο.

### **1.4 Γήρανση υποδομών**

Παρά τις νέες θεσμικές ρυθμίσεις και τα σχέδια του δικτύου του ηλεκτρικού πλέγματος, η υποδομή της παράδοσης ισχύος υποφέρει από «γήρανση» σε όλο τον ανεπτυγμένο κόσμο. Τέσσερις παράγοντες που συμβάλλουν στην τρέχουσα κατάσταση του ηλεκτρικού δικτύου και οι συνέπειές της είναι οι εξής:

- Γήρανση εξοπλισμού ισχύος – Ο παλαιότερος εξοπλισμός έχει υψηλότερα ποσοστά αποτυχίας, που οδηγεί σε ποσοστά διακοπής των πελατών επηρεάζοντας την οικονομία και την κοινωνία. Επίσης, τα μεγαλύτερα περιουσιακά στοιχεία και οι εγκαταστάσεις οδηγούν σε αύξηση του κόστους συντήρησης και περαιτέρω κόστους επισκευής / αποκατάστασης.
- Ξεπερασμένη διάταξη του συστήματος - Οι παλιότερες περιοχές απαιτούν σημαντικούς επιπλέον χώρους υποσταθμών με τρόπο που δεν μπορεί να επιτευχθεί με τη σημερινή περιοχή και είναι αναγκασμένες να χρησιμοποιούν τις υφιστάμενες, ανεπαρκείς εγκαταστάσεις.
- Παλιά μηχανική – Τα παραδοσιακά εργαλεία για το σχεδιασμό και την παράδοση μηχανικής είναι αναποτελεσματικά στην αντιμετώπιση των σημερινών προβλημάτων, των παλιών εξοπλισμών, απαρχαιωμένων διατάξεων του συστήματος, και των σύγχρονων απελευθερωμένων επίπεδων φόρτωσης.
- Παλιά πολιτιστική αξία – Ο σχεδιασμός, η κατασκευή, και η λειτουργία του συστήματος με τη χρήση εννοιών και διαδικασιών που εργάστηκαν σε καθετοποιημένη βιομηχανία επιδεινώνουν το πρόβλημα κάτω από μια απελευθερωμένη βιομηχανία.

### **1.5 Σύγχρονες Τάσεις**

Καθώς ο 21ος αιώνας προχωρά, η βιομηχανία παραγωγής ηλεκτρικού ρεύματος επιδιώκει να επωφεληθεί από νέες προσεγγίσεις για να ανταποκριθεί στην





## Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου αυξανόμενη ζήτηση ενέργειας. Βοηθητικά συστήματα βρίσκονται υπό πίεση για να εξελίσσουν τις κλασικές τοπολογίες σε φιλοξενούμενη αποκεντρωμένη παραγωγή. Ενώ η παραγωγή από ηλιακές και αιολικές γεννήτριες γίνεται πιο κοινή, οι διαφορές μεταξύ διανομής και δικτύων μεταφοράς θα συνεχίσουν να θολώνουν. Επίσης, η ανταπόκριση ζήτησης είναι μια τεχνική διαχείρισης δικτύου όπου πελάτες χονδρικής ή λιανικής ζήτησαν είτε ηλεκτρονικά είτε με το χέρι να μειώσουν το βάρος τους.

Με τα πάντα να διασυνδέονται, και με τον ελεύθερο ανταγωνισμό που συμβαίνουν σε μια ελεύθερη οικονομία της αγοράς, αρχίζει να έχει νόημα και να επιτρέπει ακόμη και την ενθάρρυνση της καταναεμημένης παραγωγής (DG). Οι μικρότερες γεννήτριες, συνήθως δεν ανήκουν στο βοηθητικό σύστημα, και μπορούν να ασκηθούν σε απευθείας σύνδεση για να βοηθήσουν τον εφοδιασμό της ανάγκης για ισχύ. Η μικρότερη μονάδα παραγωγής θα μπορούσε να είναι ένα σπίτι-ιδιοκτήτη με περίσσεια ενέργεια από ηλιακά πάνελ ή ανεμογεννητριών. Θα μπορούσε να είναι ένα μικρό γραφείο με μια γεννήτρια ντίζελ. Οι πόροι αυτοί μπορούν να εισαχθούν σε απευθείας σύνδεση είτε κατόπιν εντολής του βοηθητικού συστήματος, ή από τον ιδιοκτήτη της παραγωγής, σε μια προσπάθεια να πωλήσει ηλεκτρική ενέργεια. Πολλές μικρές γεννήτριες μπορούν να πωλούν ηλεκτρική ενέργεια πίσω στο δίκτυο για την ίδια τιμή που θα πληρώσει για να το αγοράσει. Επιπλέον, πολλές προσπάθειες βρίσκονται σε εξέλιξη για την ανάπτυξη ενός «έξυπνου δικτύου». Στις ΗΠΑ, η Energy Policy Act του 2005 και του τίτλου XIII του για την ενεργειακή ανεξαρτησία και ασφάλεια Νόμου του 2007 παρέχουν χρηματοδότηση για την ενθάρρυνση της ανάπτυξης έξυπνων δικτύων. Η ελπίδα είναι να μπορέσουν επιχειρήσεις κοινής ωφέλειας να προβλέψουν καλύτερα τις ανάγκες τους, και σε ορισμένες περιπτώσεις τη συμμετοχή των καταναλωτών. Τα κεφάλαια έχουν επίσης διατεθεί για την ανάπτυξη πιο αξιόπιστων τεχνολογιών ελέγχου της ενέργειας.

Δημιουργούνται διάφορα σχεδιαζόμενα και προτεινόμενα συστήματα για να αυξηθεί δραματικά η δυναμικότητα μεταφοράς που είναι γνωστά ως σούπερ, ή μέγα δίκτυα. Τα υποσχόμενα οφέλη περιλαμβάνουν την βιομηχανία ανανεώσιμης ενέργειας για την πώληση ηλεκτρικής ενέργειας σε μακρινές αγορές, η δυνατότητα να αυξηθεί η χρήση των πηγών διαλείπουσας ενέργειας με την εξισορρόπηση τους σε τεράστιες γεωλογικές περιοχές, καθώς και η άρση της συμφόρησης που εμποδίζει τις αγορές ηλεκτρικής ενέργειας από την άνθηση. Η τοπική αντιπολίτευση για χωροθέτηση νέων γραμμών και το σημαντικό κόστος των έργων αυτών είναι σημαντικά εμπόδια για τα σούπερ δίκτυα. Μια μελέτη για ένα ευρωπαϊκό σούπερ πλέγμα εκτιμά ότι 750 GW της επιπλέον δυναμικότητας μεταφοράς θα απαιτείται χωρητικότητα που θα φιλοξενηθούν γραμμές σε βήματα των 5 GW συνεχούς ρεύματος υψηλής τάσης. Με 750 GW νέας δυναμικότητα μεταφοράς συνεχούς ρεύματος υψηλής τάσης απαιτείται για ένα ευρωπαϊκό σούπερ πλέγμα, η γη και τα χρήματα που απαιτούνται για νέες γραμμές μεταφοράς θα είναι αξιοσημείωτα.



## 1.6 Μελλοντικές τάσεις

Όπως περαιτέρω απορρύθμιση συνεχίζεται, οι επιχειρήσεις κοινής ωφέλειας οδηγούνται να πουλήσουν περιουσιακά τους στοιχεία, όπως η αγορά ενέργειας ακολουθεί κατά σειρά, μαζί με την αγορά φυσικού αερίου στη χρήση των συμβολαίων μελλοντικής εκπλήρωσης και των αγορών και των λοιπών χρηματοπιστωτικών ρυθμίσεων. Ακόμη και η παγκοσμιοποίηση με τις ξένες αγορές λαμβάνουν χώρα. Πρόσφατα, U.K 's National Grid, η μεγαλύτερη ιδιωτική ηλεκτρική εταιρεία στον κόσμο, αγόρασε ηλεκτρικό σύστημα της Νέας Αγγλίας για 3,2 δισ. δολάρια. Σε εσωτερικό επίπεδο, οι τοπικές επιχειρήσεις ηλεκτρικών και φυσικού αερίου αρχίζουν να συγχωνεύουν λειτουργίες που θεωρούν πλεονέκτημα της κοινής υπαγωγής κυρίως με την μείωση του κόστους της ένωσης. Η τεχνολογική πρόοδος θα λάβει χώρα στις ανταγωνιστικές αγορές χονδρικής ηλεκτρικών, τέτοια παραδείγματα που έχουν ήδη χρησιμοποιηθεί περιλαμβάνουν κυψέλες καυσίμου που χρησιμοποιούνται σε διαστημικές πτήσεις, αεριωθήσεις αεριοστρόβιλων που χρησιμοποιούνται σε αεροσκάφη jet, ηλιακή τεχνολογία και τα φωτοβολταϊκά συστήματα, off-shore αιολικά πάρκα, και τις προόδους της επικοινωνίας που γεννήθηκαν από το ψηφιακό κόσμο.

Η ηλεκτρική ενέργεια αναμένεται να έχει αυξανόμενη ζήτηση στο μέλλον. Η επανάσταση των πληροφοριών είναι σε μεγάλο βαθμό στην ηλεκτρική ενέργεια. Άλλοι τομείς που περιλαμβάνουν την ανάπτυξη των αναδυόμενων νέων τεχνολογιών που βασίζονται καθαρά στην ηλεκτρική ενέργεια, οι εξελίξεις στο χώρο κλιματισμό, βιομηχανική διαδικασία, και τη μεταφορά (π.χ. υβριδικά οχήματα, μηχανές).

## 1.7 Αναδυόμενο έξυπνο δίκτυο

Όπως προαναφέρθηκε, το ηλεκτρικό δίκτυο αναμένεται να εξελιχθεί σε ένα νέο πρότυπο δίκτυο - έξυπνο δίκτυο. Τα παραδοσιακά ηλεκτρικά πλέγματα που χρησιμοποιούνται γενικά για τη μεταφορά ρεύματος διανέμουν από λίγες κεντρικές γεννήτριες σε μεγάλο αριθμό χρηστών ή πελατών. Σε αντίθεση, το νέο αναδυόμενο έξυπνο δίκτυο χρησιμοποιεί αμφίδρομη ροή του ηλεκτρικού ρεύματος και πληροφορίες για να δημιουργήσει ένα αυτοματοποιημένο και κατανεμημένο προηγμένο δίκτυο παροχής ενέργειας.

Πολλά ερευνητικά έργα έχουν πραγματοποιηθεί για τη διερεύνηση της έννοιας του έξυπνου δικτύου. Σύμφωνα με μια νεότερη έρευνα για έξυπνο δίκτυο, η έρευνα επικεντρώνεται κυρίως σε τρία συστήματα, το σύστημα των υποδομών, το σύστημα διαχείρισης και το σύστημα προστασίας.

Το σύστημα των υποδομών είναι η ενέργεια, η πληροφορία και η επικοινωνιακή υποδομή που στηρίζονται από το έξυπνο δίκτυο και που υποστηρίζει



1) προηγμένη παραγωγή ηλεκτρικής ενέργειας, παράδοση, και κατανάλωση, 2) προηγμένες πληροφορίες μέτρησης, παρακολούθησης, διαχείρισης και 3) προηγμένες τεχνολογίες επικοινωνίας. Κατά τη μετάβαση από τη συμβατικό δίκτυο ισχύος σε έξυπνο δίκτυο, θα αντικατασταθεί μια φυσική υποδομή με μια ψηφιακή. Οι ανάγκες και οι αλλαγές που παρουσιάζει ο κλάδος ενέργειας είναι μία από τις μεγαλύτερες προκλήσεις που έχει αντιμετωπίσει ποτέ.

Το σύστημα διαχείρισης είναι το υποσύστημα στο έξυπνο δίκτυο που παρέχει προηγμένες υπηρεσίες διαχείρισης και ελέγχου. Τα περισσότερα από τα υπάρχοντα έργα στοχεύουν στη βελτίωση της ενεργειακής απόδοσης, το προφίλ της ζήτησης, τη χρησιμότητα, το κόστος και τις εκπομπές, με βάση την υποδομή μέσω της βελτιστοποίησης, **μηχανικής μάθησης**, και τη θεωρία των παιγνίων. Στο πλαίσιο προηγμένης υποδομής του έξυπνου δικτύου, όλο και περισσότερες νέες υπηρεσίες και εφαρμογές διαχείρισης αναμένεται να προκύψουν και τελικά να φέρει επανάσταση στην καθημερινή ζωή των καταναλωτών.

Το σύστημα προστασίας είναι το υποσύστημα στο έξυπνο δίκτυο που παρέχει προηγμένη ανάλυση της αξιοπιστίας του δικτύου, την προστασία αποτυχιών, την ασφάλεια των υπηρεσιών και προστασίας της ιδιωτικής ζωής. Πρέπει να σημειωθεί ότι η προηγμένη υποδομή που χρησιμοποιείται στο έξυπνο δίκτυο αφενός δίνει τη δυνατότητα να αναπτυχθούν πιο ισχυροί μηχανισμοί για την υπεράσπιση ενάντια στις επιθέσεις και τις αποτυχίες, αλλά από την άλλη πλευρά, ανοίγει πολλές νέες ευπάθειες. Για παράδειγμα, το NIST επεσήμανε ότι το σημαντικό όφελος που παρέχεται από το έξυπνο δίκτυο είναι η δυνατότητα απόκτησης πλουσιότερων δεδομένων από και προς των έξυπνων μετρητών του πελάτη και άλλων ηλεκτρικών συσκευών. Το σύστημα προστασίας είναι επίσης η αχίλλειος πτέρνα του έξυπνου δικτύου από την άποψη της ιδιωτικής ζωής. Η προφανής ανησυχία της ιδιωτικής ζωής είναι ότι η χρήση πληροφοριών ενέργειας που αποθηκεύεται στο μέτρο, λειτουργεί ως κανάλι πλούσιο σε πληροφορίες. Αυτή η πληροφορία μπορεί να αντληθεί και να ανακτηθεί από τα ενδιαφερόμενα μέρη και να αποκαλύψουν προσωπικές πληροφορίες, όπως τις συνήθειες, συμπεριφορές του ατόμου, τις δραστηριότητες, ακόμα και τις πεποιθήσεις.

### 1.7.1 Σύστημα διαχείρισης Διανομής (Distribution Management System)

Τα τελευταία χρόνια, η χρήση της ηλεκτρικής ενέργειας αυξάνεται εκθετικά και απαίτηση των πελατών και οι ορισμοί της ποιότητας της ενέργειας έχουν αλλάξει πάρα πολύ. Δεδομένου ότι η ηλεκτρική ενέργεια έγινε ένα ουσιαστικό μέρος της καθημερινής ζωής, η βέλτιστη χρήση και η αξιοπιστία του έγινε σημαντική. Η προβολή δικτύων πραγματικού χρόνου και δυναμικές αποφάσεις έχουν παίξει ζωτικό ρόλο για τη βελτιστοποίηση των πόρων και τη διαχείριση των απαιτήσεων, καθιστώντας έτσι ένα σύστημα διαχείρισης διανομής το οποίο θα μπορούσε να χειριστεί την ορθή ροή εργασίας, κάτι που είναι πολύ κρίσιμο.



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

Ένα **Σύστημα Διαχείρισης Διανομής (Distribution Management System-DMS)** είναι μια συλλογή από εφαρμογές που έχουν σχεδιαστεί για την παρακολούθηση και τον έλεγχο ολόκληρου του δικτύου διανομής αποτελεσματικά και αξιόπιστα. Λειτουργεί ως ένα σύστημα υποστήριξης αποφάσεων για να βοηθήσει την αίθουσα ελέγχου και στον τομέα της λειτουργίας του προσωπικού με την παρακολούθηση και τον έλεγχο του ηλεκτρικού συστήματος διανομής. Η βελτίωση της αξιοπιστίας και της ποιότητας των παρεχόμενων υπηρεσιών όσον αφορά τη μείωση των διακοπών, την ελαχιστοποίηση του χρόνου διακοπής, διατηρώντας αποδεκτή συχνότητα και τα επίπεδα τάσης είναι τα βασικά παραδοτέα ενός DMS.

Οι περισσότερες επιχειρήσεις διανομής έχουν πλήρως τη χρήση Πληροφοριακών τεχνολογιών μέσω του **Συστήματος Διαχείρισης Διακοπής (Outage Management System)—OMS** που κάνει χρήση άλλων συστημάτων όπως το Σύστημα Πληροφοριών Πελατών (CIS), Γεωγραφικών Συστημάτων Πληροφοριών (GIS) και Συστήματος Interactive Voice Response (IVRS). Συνδυάζοντας τις θέσεις των κλήσεων διακοπών από τους πελάτες με τη γνώση των θέσεων των συσκευών προστασίας (όπως διακόπτες κυκλώματος) στο δίκτυο, μια μηχανή που λειτουργεί με κανόνες χρησιμοποιείται για να προβλέψει τις θέσεις των διακοπών. Με βάση αυτό, οι δραστηριότητες αποκατάστασης χαρτογραφούνται από το πλήρωμα και αποστέλλονται για την ίδια.

Παράλληλα με αυτό, οι επιχειρήσεις διανομής άρχισαν να αναπτύσσουν συστήματα Εποπτικού Ελέγχου και Συλλογής Δεδομένων (SCADA), αρχικά μόνο σε υποσταθμούς υψηλότερης τάσης. Με την πάροδο του χρόνου, η χρήση του SCADA έχει επεκταθεί σταδιακά προς τα κάτω σε δικτυακούς τόπους σε χαμηλότερα επίπεδα τάσης.

Το DMS έχει πρόσβαση σε δεδομένα πραγματικού χρόνου και παρέχει όλες τις πληροφορίες με μια ενιαία κονσόλα στο κέντρο ελέγχου με ολοκληρωμένο τρόπο. Η ανάπτυξή τους ποικίλλει στις διάφορες γεωγραφικές περιοχές. Στις ΗΠΑ, για παράδειγμα, με τα DMS αυξήθηκαν τα Συστήματα Διαχείρισης Διακοπών στο επόμενο επίπεδο αυτοματοποιώντας πλήρως τις ακολουθίες και παρέχοντας μια από άκρο σε άκρο, ολοκληρωμένη άποψη όλου το φάσματος της διανομής. Στο Ηνωμένο Βασίλειο, αντίθετα, το πολύ πυκνότερο και οι πιο πυκνες τοπολογίες, σε συνδυασμό με την «Υγεία & Ασφάλεια» ρύθμιση, είχε σαν αποτέλεσμα την πρόωρη συγκέντρωση λειτουργιών μεταγωγών υψηλής τάσης χρησιμοποιώντας αρχικά αρχεία σε χαρτί και σχηματικά διαγράμματα που εκτυπώνονται σε μεγάλες «πλακέτες» που είχαν μαγνητικά σύμβολα για να δείξουν τις τρέχουσες καταστάσεις λειτουργίας. Εκεί, τα DMS αυξήθηκαν αρχικά από τα συστήματα SCADA, όπως αυτά έχουν επεκταθεί για να επιτρέψουν σε αυτά και κεντρικό έλεγχο των διαδικασιών διαχείρισης της ασφάλειας που πρέπει να διεκπεραιωθεί με ηλεκτρονικά μέσα. Αυτά τα DMS απαιτούνται ακόμη πιο λεπτομερή συνδετικά μοντέλα και σχήματα από αυτά που απαιτούνται από τις αρχές του OMSs όπως κάθε πιθανή απομόνωση και έπρεπε να συμπεριληφθεί το σημείο γείωσης για τα δίκτυα. Σε περιοχές όπως το Ηνωμένο Βασίλειο, ως εκ τούτου, τα μοντέλα τύπου συνδετικού δικτύου αναπτύσσονται συνήθως στο DMS πρώτα.



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

Η τυπική ροή δεδομένων σε ένα DMS έχει το σύστημα SCADA, το σύστημα αποθήκευσης & Ανάκτηση πληροφοριών (ISR) , Servers επικοινωνίας (COM) , Front-End επεξεργαστές (FEPs) & Field Remote Terminal Units (FRTUs).

### 1.7.2 Γιατί Σύστημα διαχείρισης Διανομής (DMS);

- Μείωση τη διάρκειας των διακοπών
- Βελτίωση της ταχύτητας και της ακρίβειας των προβλέψεων διακοπής.
- Μείωση της περιπολίας πληρώματος και τους χρόνους κίνησης μέσω της βελτίωσης του εντοπισμού της διακοπής.
- Βελτίωση της λειτουργικής αποτελεσματικότητας
- Καθορισμός των πόρων του πληρώματος αναγκαία για την επίτευξη των στόχων αποκατάστασης.
- Αποτελεσματική αξιοποίηση των πόρων μεταξύ των περιοχών λειτουργίας.
- Καθορισμός της βέλτιστης περίπτωσης για τον προγραμματισμό της βοήθειας αμοιβαίων πληρωμάτων.
- Αύξηση της ικανοποίησης των πελατών
- Ένα DMS ενσωματώνει IVR και άλλες κινητές τεχνολογίες, μέσω της οποίας υπάρχει μια βελτίωση στην επικοινωνία των πελατών για την αντιμετώπιση διακοπών.
- Παροχή στους πελάτες με πιο ακριβή εκτιμήσεις για την ώρα της αποκατάστασης των διακοπών.
- Βελτίωση της αξιοπιστίας παροχής υπηρεσιών με την παρακολούθηση των πελατών που επηρεάζονται από την διακοπή, του καθορισμού ηλεκτρικών διαμορφώσεων της κάθε συσκευής σε κάθε τροφοδότη.

### 1.7.3 Λειτουργίες Συστήματος διαχείρισης Διανομής

Προκειμένου να ενισχυθεί η κατάλληλη διαδικασία λήψης αποφάσεων και Ο & Μ δραστηριοτήτων, η DMS λύση θα πρέπει να υποστηρίζει τις ακόλουθες λειτουργίες:

- Οπτικοποίηση Δικτύου & εργαλεία υποστήριξης
- Οι αιτήσεις για την αναλυτική και διορθωτική δράση
- Εργαλεία Προγραμματισμού Utility
- Σχέδια του συστήματος προστασίας



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

Οι διάφορες επιμέρους λειτουργίες, πραγματοποιούνται από το DMS που παρατίθενται κατωτέρω:

#### 1.7.4 Ανάλυση Συνδεσιμότητας Δικτύου (NCA)

Το δίκτυο διανομής καλύπτει συνήθως σε μεγάλη έκταση και διανέμει το ηλεκτρικό ρεύμα σε διαφορετικούς πελάτες σε διαφορετικά επίπεδα τάσης. Έτσι ο εντοπισμός απαιτούμενων πηγών και των φορτίων σε ένα μεγαλύτερο περιβάλλον GIS / Operator είναι συχνά πολύ δύσκολος. Η μετακίνηση και η μεγέθυνση παρέχονται με κανονικό σύστημα SCADA και το GUI δεν καλύπτει την ακριβή επιχειρησιακή απαίτηση. Το δίκτυο ανάλυσης συνδεσιμότητας είναι ένας φορέας ειδικής λειτουργίας που βοηθά το χειριστή να αναγνωρίσει ή να εντοπίσει το προτιμώμενο δίκτυο πολύ εύκολα. Το NCA κάνει τις απαιτούμενες αναλύσεις και παρέχει απεικόνιση του σημείου τροφοδοσίας των διαφόρων φορτίων του δικτύου. Με βάση την κατάσταση όλων των συσκευών μεταγωγής όπως διακόπτη (CB), Ring Main Unit (RMU) και / ή μονωτές που επηρεάζουν την τοπολογία του δικτύου μοντελοποιούνται, και η επικρατούσα τοπολογία του δικτύου ορίζεται. Το NCA βοηθά περαιτέρω το χειριστή να γνωρίζει λειτουργική κατάσταση του δικτύου διανομής δείχνοντας την ακτινική λειτουργία, βρόχους και παραλληλισμούς στο δίκτυο.

#### 1.7.5 Πρόγραμμα Εναλλαγής & Διαχείρισης της Ασφάλειας

Σε περιοχές όπως το Ηνωμένο Βασίλειο μια βασική λειτουργία του DMS ήταν πάντα να υποστηρίζει την ασφαλή μεταγωγή και λειτουργία στα δίκτυα. Μηχανικοί ελέγχου προετοιμάζουν τα χρονοδιαγράμματα των ασφαλών μεταγωγών για να απομονωθεί και να δημιουργηθεί ένα ασφαλές τμήμα του δικτύου πριν πραγματοποιηθεί καμία εργασία, και το DMS επικυρώνει τα προγράμματα χρησιμοποιώντας το μοντέλο του δικτύου της. Εναλλάσσοντας προγράμματα μπορούν να συνδυάσουν με τηλεχειρισμό και χειροκίνητες (on-site) λειτουργίες μεταγωγής. Όταν το απαιτούμενο τμήμα έχει γίνει ασφαλές, το DMS επιτρέπει να εκδοθεί έγγραφο που θα επιτρέπει την άδεια Εργασίας (Pemit to work PTW). Μετά την ακύρωσή του, όταν το έργο έχει ολοκληρωθεί, το πρόγραμμα μεταγωγών διευκολύνει την αποκατάσταση των κανονικών ρυθμίσεων λειτουργίας.

Το μοντέλο συνδετικού δικτύου, και τα συναφή διαγράμματα, πρέπει πάντα να είναι ενημερωμένα. Το χρονοδιάγραμμα εναλλαγής συνεπώς επιτρέπει επίσης «μπαλώματα» στο μοντέλο δικτύου για να εφαρμοστεί στην ζωντανή έκδοση στο κατάλληλο στάδιο(-α) των θέσεων εργασίας. Ο όρος «μπάλωμα» προέρχεται από τη



Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
μέθοδο που χρησιμοποιήθηκε προηγουμένως για τη διατήρηση των διαγραμμάτων  
πλακετών.

### 1.7.6 Εκτίμηση κράτους (SE)

Ο εκτιμητής κατάστασης αποτελεί αναπόσπαστο μέρος του συνόλου των συστημάτων παρακολούθησης και ελέγχου των δικτύων μεταφοράς. Ο εκτιμητής κατάστασης αποσκοπεί στην παροχή μιας αξιόπιστης εκτίμησης των τάσεων του συστήματος. Η πληροφορία από την εκτιμητή κατάσταση ρέει στο κέντρο ελέγχου και στους εξυπηρετητές (σέρβερς) βάσεων δεδομένων σε όλο το δίκτυο. Οι σημαντικές μεταβλητές είναι ενδεικτικές για τις παραμέτρους, όπως τα περιθώρια για τα όρια λειτουργίας, η ασφάλεια του εξοπλισμού και οι απαιτούμενες ενέργειες του χειριστή. Οι εκτιμητές επιτρέπουν τον υπολογισμό αυτών των σημαντικών μεταβλητών με υψηλή αυτοπεποίθηση, παρά τα γεγονότα ότι οι μετρήσεις μπορεί να είναι κατεστραμμένες από το θόρυβο, ή θα μπορούσε να λείπουν.

Ακόμα κι αν μπορεί να μην είναι σε θέση να παρατηρήσουν άμεσα την κατάσταση, μπορεί να συναχθεί από μια σάρωση των μετρήσεων που υποτίθεται ότι πρέπει να συγχρονιστούν. Οι αλγόριθμοι πρέπει να λάβουν υπόψη το γεγονός ότι η παρουσία του θορύβου ενδέχεται να τσαλακώσει τις μετρήσεις. Σε ένα τυπικό σύστημα παραγωγής ηλεκτρικής ενέργειας, η κατάσταση είναι ημι-στατική. Οι σταθερές χρόνου είναι επαρκώς γρήγορες, ώστε το σύστημα διάσπασης δυναμικής να φθίρειται γρήγορα (σε σχέση με τη συχνότητα μέτρησης). Το σύστημα φαίνεται να εξελίσσεται μέσα από μια σειρά στατικών καταστάσεων που οδηγούνται από διάφορες παραμέτρους, όπως οι αλλαγές στο προφίλ του φορτίου. Οι εισοδοί του εκτιμητή κατάστασης μπορούν να δοθούν σε διάφορες εφαρμογές όπως η ανάλυση της ροής φορτίου, ανάλυση έκτακτης ανάγκης, καθώς και άλλες εφαρμογές.

### 1.7.7 Εφαρμογές ροής φορτίου (LFA)

Η μελέτη της ροής φορτίου είναι ένα σημαντικό εργαλείο που περιλαμβάνει αριθμητική ανάλυση σε ένα σύστημα ισχύος. Η μελέτη ροής φορτίου χρησιμοποιεί συνήθως απλοποιημένους συμβολισμούς, όπως ένα διάγραμμα μιας γραμμής και επικεντρώνεται σε διάφορες μορφές του εναλλασσόμενου ρεύματος και όχι τάσης και ρεύματος. Αναλύει τα συστήματα ισχύος σε κανονική σταθερή κατάσταση λειτουργίας. Ο στόχος της μελέτης ροής ισχύος είναι να ληφθεί πλήρης γωνία τάσης και πληροφορίες μεγέθους για κάθε διάλο σε ένα σύστημα ισχύος για συγκεκριμένο φορτίο και γεννήτριες πραγματικής ισχύος και τις συνθήκες τάσης. Μόλις αυτή η πληροφορία είναι γνωστή, η πραγματική και η αντιδραστική ισχύς για κάθε



Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
διακλάδωση, καθώς και η γεννήτρια αντιδραστικής ισχύς μπορούν να προσδιοριστούν αναλυτικά.

Λόγω της μη γραμμικής φύσεως του προβλήματος αυτού, οι αριθμητικές μεθόδους που χρησιμοποιούνται για να ληφθεί μια λύση είναι μέσα σε ένα αποδεκτό εύρος τιμών. Το μοντέλο φορτίου πρέπει να υπολογίσει αυτόματα τα φορτία για να ταιριάζει με το Telemeter ή τα προβλεπόμενα ρεύματα τροφοδοσίας. Χρησιμοποιεί τύπο «πελάτη», των χαρακτηριστικών και άλλες πληροφορίες για τη κατάλληλη διανομή του φορτίου σε κάθε μετασχηματιστή διανομής. Μελέτες Ροής Φορτίου ή Ροής Ισχύος είναι σημαντικές για το σχεδιασμό μελλοντικής επέκτασης των συστημάτων ηλεκτρικής ενέργειας, καθώς και τον προσδιορισμό της καλύτερης λειτουργίας των υφιστάμενων συστημάτων.

### 1.7.8 Έλεγχος Volt-VAR (VVC)

Έλεγχος Volt-VAR ή VVC αναφέρεται στη διαδικασία της διαχείρισης των επιπέδων τάσης και αντιδραστικής ισχύος (VAR) σε όλα τα συστήματα διανομής ηλεκτρικής ενέργειας. Θα μπορούσαν να υπάρχουν φορτία που περιέχουν αντιδραστικά συστατικά, όπως πυκνωτές και πηνία (όπως ηλεκτρικοί κινητήρες) που βάζουν επιπλέον πίεση στο δίκτυο. Αυτό οφείλεται στο γεγονός ότι το αντιδραστικό μέρος αυτών των φορτίων τους αναγκάζει να επιστήσει περισσότερο ρεύμα. Τα αποτελέσματα ασταθούς ρεύματος και στις δύο over-voltage/under-voltage παραβιάσεις, καθώς και το ζέσταμα του εξοπλισμού όπως μετασχηματιστές, αγωγοί, κ.λπ., τα οποία μπορεί να χρειαστεί ακόμη και αλλαγή μεγέθους για να φέρει το συνολικό ρεύμα. Ένα σύστημα ισχύος πρέπει να ελέγχεται σχεδιάζοντας την παραγωγή, απορρόφηση και τη ροή της αντιδραστικής ισχύος σε όλα τα επίπεδα του συστήματος.

Μια VVC εφαρμογή βοηθά τον χειριστή να μετριάσει τέτοιες συνθήκες προτείνοντας απαιτούμενα σχέδια δράσης. Το σχέδιο θα δώσει την απαιτούμενη θέση και πυκνωτή για να εξασφαλίσει την τάση στο όριο της και κατά συνέπεια βελτιστοποίηση Volt Vary λειτουργία ελέγχου για τη χρησιμότητα.

### 1.7.9 Εφαρμογή Περικοπής Φορτίου (LSA)

Τα Ηλεκτρικά Συστήματα Διανομής έχουν μεγάλα τμήματα γραμμής μεταφοράς, πολλαπλά ενωτικά σημεία και κυμαινόμενη ζήτηση των καταναλωτών. Αυτά τα χαρακτηριστικά είναι εγγενώς ευάλωτα σε ασταθείς ή απρόβλεπτες συνθήκες του συστήματος που μπορεί να οδηγήσουν σε κρίσιμη αποτυχία. Η αστάθεια προκύπτει συνήθως από ταλαντώσεις ισχύος του συστήματος λόγω βλαβών, μέγιστου ελλείμματος ή αποτυχημένες προστασίες. Διανομή περικοπής φορτίου και





Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
σχεδιαγράμματα αποκατάστασης παίζουν ζωτικό ρόλο στη λειτουργία έκτακτης  
ανάγκης και έλεγχο σε κάθε χρησιμότητα.

Μια αυτοματοποιημένη Εφαρμογή Περικοπής Φορτίου εντοπίζει προκαθορισμένες συνθήκες ενεργοποίησης του δικτύου διανομής και εκτελεί προκαθορισμένες σειρές δράσεων ελέγχου, όπως το άνοιγμα ή το κλείσιμο μη κρίσιμων τροφοδοτών, την αναμόρφωση διανομής ή τις πηγές των ενώσεων, ή την εκτέλεση ενός ελέγχου χείλους σε ένα μετασχηματιστή. Όταν ένα δίκτυο διανομής είναι πολύπλοκο και καλύπτει μια ευρύτερη περιοχή, ενέργειες έκτακτης ανάγκης που λαμβάνονται προς τα κάτω μπορούν να μειώσουν την επιβάρυνση των ανάντη τμημάτων του δικτύου. Σε ένα μη-αυτοματοποιημένο σύστημα, την ευαισθητοποίηση και χειροκίνητη παρέμβαση του χειριστή παίζουν καθοριστικό ρόλο στο μετριασμό προβλήματος. Αν δεν αντιμετωπιστούν τα προβλήματα αρκετά γρήγορα, μπορούν να διαλυθούν εκθετικά και να προκαλέσουν σημαντική καταστροφική αποτυχία.

Το DMS πρέπει να παρέχει μια σπονδυλωτή αυτόματη περικοπή φορτίου και αποκατάσταση εφαρμογής που αυτοματοποιεί τη λειτουργία έκτακτης ανάγκης και τις απαιτήσεις ελέγχου για οποιαδήποτε χρησιμότητα. Η εφαρμογή θα πρέπει να καλύπτει διάφορες δραστηριότητες, όπως Κάτω Περικοπή Φορτίου συχνότητας (UFLS), παραβίαση ορίου και βάση την ώρα της ημέρας αποκοπής φορτίου με τα συστήματα που συνήθως εκτελούνται από τον χρήστη.

### **1.7.10 Διαχείριση Βλάβης & Αποκατάσταση Συστήματος (FMSR)**

Η αξιοπιστία και η ποιότητα της παροχής ενέργειας αποτελούν τις βασικές παραμέτρους που πρέπει να εξασφαλίζεται από οποιαδήποτε χρησιμότητα. Η Μειωμένη διάρκεια του χρόνου διακοπής για τον πελάτη, πρέπει να βελτιωθεί σε όλους τους δείκτες αξιοπιστίας ως εκ τούτου FMSR και οι αυτοματοποιημένες εφαρμογές μεταγωγής παίζουν σημαντικό ρόλο. Τα δύο κύρια χαρακτηριστικά που απαιτούνται από ένα FMSR είναι: Ενεργοποίηση της διαχείρισης & Προτεινόμενη αλλαγή σχεδίου

Η εφαρμογή DMS λαμβάνει πληροφορίες σφαλμάτων από το σύστημα SCADA και επεξεργάζεται το ίδιο για τον εντοπισμό των σφαλμάτων και την λειτουργία εφαρμογής διαχείρισης μεταγωγών. Τα αποτελέσματα μετατρέπονται σε σχέδια δράσης από τις εφαρμογές. Το σχέδιο δράσης περιλαμβάνει τους διακόπτες ON / OFF αλλάζοντας την αυτόματη διακοπή φόρτισης / RMU / Sectionalizer. Το σχέδιο δράσης μπορεί να επαληθευτεί σε λειτουργία μελέτης που παρέχεται από την λειτουργικότητα. Η αλλαγή της διαχείρισης μπορεί να είναι χειροκίνητη / αυτόματη ανάλογα με τη διαμόρφωση.



### 1.7.11 Εξισορρόπηση φορτίου μέσω τροφοδότη Αναδιάρθρωσης (LBFR)

Η εξισορρόπηση φορτίου μέσω του τροφοδότη αναδιαμόρφωσης είναι απαραίτητη για την εφαρμογή κοινής ωφέλειας, όταν έχουν πολλούς τροφοδότες να τροφοδοτούν ένα φορτίο κορεσμένης περιοχής. Για την εξισορρόπηση των φορτίων σε ένα δίκτυο, ο χειριστής ορίζει ξανά τα φορτία σε άλλα τμήματα του δικτύου. Η διαχείριση φορτίου τροφοδοσίας (FLM) είναι απαραίτητη για να μπορεί να διαχειριστεί την παράδοση της ενέργειας στο ηλεκτρικό σύστημα διανομής και να εντοπίσει τις προβληματικές περιοχές. Η διαχείριση φορτίου τροφοδοσίας παρακολουθεί τα ζωτικά σημεία του συστήματος διανομής και προσδιορίζει τους τομείς που προκαλούν ανησυχία, έτσι ώστε ο χειριστής διανομής να έχει προειδοποιηθεί και να μπορεί να εστιάσει την προσοχή αποτελεσματικά όπου χρειάζεται περισσότερο. Παρέχει τη δυνατότητα για ταχύτερη διόρθωση των υφιστάμενων προβλημάτων και επιτρέπει δυνατότητες για την αποφυγή προβλημάτων, οδηγώντας στην βελτίωση της αξιοπιστίας και παράδοση ενεργειακής απόδοσης.

Σε μια παρόμοια σημείωση, ο τροφοδότης αναδιάρθρωσης χρησιμοποιείται επίσης για την ελαχιστοποίηση απώλειας. Λόγω πολλών χρηστών του δικτύου ο τροφοδότης μπορεί να λειτουργήσει στη μέγιστη ικανότητά του χωρίς να γνωρίζει τις συνέπειες των απωλειών. Οι συνολικές απώλειες ενέργειας και οι απώλειες εσόδων που οφείλονται σε αυτές τις πράξεις θα πρέπει να ελαχιστοποιηθούν για την αποτελεσματική λειτουργία. Το πρόβλημα ελαχιστοποίησης ζημιών λύνεται με το βέλτιστο αλγόριθμο ροής ισχύος και τα σχέδια αλλάζουν και στη συνέχεια δημιουργούνται παρόμοια με καλύτερη λειτουργία.

### 1.7.12 Πρότυπα με βάση την ένταξη

Σε κάθε μοντέλο ολοκληρωμένης παροχής ενέργειας λειτουργικής χρησιμότητας, υπάρχουν διαφορετικές λειτουργικές ενότητες όπως η λύση GIS, Billing & μέτρησης, ERP, σύστημα διαχείρισης στοιχείων που λειτουργούν παράλληλα και υποστηρίζουν τις συνήθεις διαδικασίες. Αρκετά συχνά, κάθε μία από αυτές τις λειτουργικές ενότητες πρέπει να ανταλλάσσουν περιοδικά ή σε πραγματικό χρόνο δεδομένα μεταξύ τους για την αξιολόγηση της παρούσας κατάστασης λειτουργίας του δικτύου, των ροών εργασίας, και τους πόρους (όπως το πλήρωμα, κλπ.). Σε αντίθεση με άλλα τμήματα του συστήματος ισχύος, το σύστημα διανομής αλλάζει ή μεγαλώνει κάθε ημέρα, και αυτό θα μπορούσε να οφείλεται στην προσθήκη ενός νέου καταναλωτή, μιας νέα γραμμής μετάδοσης ή αντικατάσταση του εξοπλισμού. Εάν οι διαφορετικές λειτουργικές μονάδες λειτουργούν σε ένα μη τυποποιημένο περιβάλλον και χρησιμοποιεί απλά APIs και διεπαφές βάσεων



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

δεδομένων, η προσπάθεια για τη διαχείριση των μηχανικών θα πρέπει να γίνει πολύ μεγάλη. Σύντομα θα γίνει δύσκολο να διαχειριστούν τις αυξανόμενες αλλαγές και προσθήκες που θα μπορούσε να οδηγήσει στη λήψη ολοκληρωμένων μη λειτουργικών συστημάτων. Συνεπώς βοηθητικά προγράμματα μπορούν να κάνουν την σωστή σύνδεση μεταξύ των λειτουργικών ενοτήτων και σε ορισμένες περιπτώσεις. Τα συστήματα μπορεί να χρειαστεί ακόμη να μεταφερθούν σε κατάλληλα περιβάλλοντα με πολύ υψηλό κόστος.

Δεδομένου ότι αυτά τα προβλήματα ήρθαν στο φως, διάφορες διαδικασίες τυποποίησης για την ανταλλαγή δεδομένων μεταξύ της εφαρμογής ξεκίνησαν. Ήταν κατανοητό ότι ένα βασισμένο πρότυπο ολοκλήρωσης διευκολύνει την ενσωμάτωση με άλλες λειτουργικές μονάδες και ότι βελτιώνει επίσης την επιχειρησιακή απόδοση. Εξασφαλίζει ότι η υπηρεσία μπορεί να είναι σε ένα ουδέτερο περιβάλλον για μελλοντικές επεκτάσεις, η οποία με τη σειρά της σημαίνει ότι η υπηρεσία εύκολα προσθέτει νέες λειτουργικές μονάδες πάνω από την υπάρχουσα λειτουργικότητα και εύκολα ωθεί ή να τραβά τα δεδομένα αποτελεσματικά χωρίς να χρειάζεται νέες προσαρμογές διεπαφών.

### 1.7.13 IEC 61968 πρότυπο με βάση την ολοκλήρωση

Το IEC 61968 είναι ένα πρότυπο που αναπτύχθηκε από την Ομάδα Εργασίας 14 της Τεχνικής Επιτροπής 57 του IEC και καθορίζει τα πρότυπα για την ανταλλαγή πληροφοριών μεταξύ των ηλεκτρικών εφαρμογών του συστήματος διανομής. Σκοπός του είναι να υποστηρίξει την εφαρμογή μεταξύ της ολοκλήρωσης μιας επιχείρησης κοινής ωφέλειας που πρέπει να συλλέγουν δεδομένα από διαφορετικές εφαρμογές, οι οποίες θα μπορούσαν να είναι νέα ή κληρονομιά.

Σύμφωνα με το IEC 61968, το DMS συμπυκνώνει διάφορες δυνατότητες, όπως την παρακολούθηση και τον έλεγχο του εξοπλισμού για την παροχή ενέργειας, τις διαδικασίες διαχείρισης ώστε να διασφαλίζεται η αξιοπιστία του συστήματος, η διαχείριση τάσης, διαχείριση της ζήτησης, η διαχείριση διακοπής, διαχείριση έργου, αυτοματοποιημένη χαρτογράφηση και διαχείριση εγκαταστάσεων. Η ουσία των προτύπων IEC 61968 είναι το Μοντέλο Αναφοράς Διασύνδεσης (IRM) που καθορίζει διάφορες τυποποιημένες διεπαφές για κάθε κατηγορία εφαρμογών. Τα απόλυτα (Λογικά) στοιχεία παρατίθενται να εκπροσωπούν συμπαγές (φυσικές) εφαρμογές. Για παράδειγμα, η λειτουργία των επιχειρήσεων, όπως Λειτουργίας και Διαχείρισης Δικτύου (NO) θα μπορούσε να εκπροσωπείται από διάφορες υπο-λειτουργίες όπως η παρακολούθηση Λειτουργίας και Διαχείρισης Δικτύου (NMON), το οποίο με τη σειρά του θα εκπροσωπηθεί από αφηρημένα στοιχεία, όπως κρατικό έλεγχο υποσταθμού, Δίκτυο κατάσταση επίβλεψη, την εποπτεία και συναγερμού.

Το IEC 61968 συστήνει ότι οι διεπαφές του συστήματος μιας συμβατής χρησιμότητας εσωτερικής εφαρμογής των υποδομών πρέπει να ορίζεται με βάση Ενοποιημένη Γλώσσα Μοντελοποίησης (UML). Το UML περιλαμβάνει μια σειρά από γραφικές τεχνικές σημειογραφίας που μπορούν να χρησιμοποιηθούν για να



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

δημιουργήσουν οπτικά μοντέλα αντικειμενοστραφούς λογισμικού εντατικών συστημάτων. Η IEC 61968 σειρά προτύπων επεκτείνει το Μοντέλο Κοινών Πληροφοριών (CIM), το οποίο διατηρείται ως ένα μοντέλο UML, για να καλύψει τις ανάγκες της διανομής ηλεκτρικού ρεύματος. Για δομημένη ανταλλαγή εγγράφων ιδίως μέσω του Διαδικτύου, η μορφή των δεδομένων που χρησιμοποιούνται μπορεί να είναι το Extensible Markup Language (XML). Μια από τις πρωταρχικές χρήσεις του είναι η ανταλλαγή πληροφοριών μεταξύ των διαφορετικών και ενδεχομένως ασύμβατων συστημάτων πληροφορικής. Το XML είναι έτσι καλά προσαρμοσμένη στο πεδίο των διεπαφών του συστήματος για τη διαχείριση της διανομής. Μετατρέπει το μήνυμα, έτσι ώστε να φορτώσει το ίδιο για διάφορες μεταφορές μνημάτων, όπως τα μηνύματα SOAP (Simple Object Access πρωτόκολλο), κλπ.

### 1.7.14 Διανομή Προβλεπόμενου Φορτίου (DLF)

Διανομή Προβλεπόμενου Φορτίου (DLF) παρέχει ένα δομημένο περιβάλλον για τη δημιουργία, τη διαχείριση και την ανάλυση των προβλέψεων φορτίου. Ακριβή μοντέλα πρόβλεψης για ηλεκτρικό φορτίο ισχύος είναι απαραίτητα για τη λειτουργία και τον προγραμματισμό μιας εταιρείας κοινής ωφελείας. Το DLF βοηθά ένα ηλεκτρικό βοηθητικό πρόγραμμα για να κάνουν σημαντικές αποφάσεις, συμπεριλαμβανομένων των αποφάσεων για την αγορά ηλεκτρικής ενέργειας, τη μεταγωγή του φορτίου, καθώς και την ανάπτυξη των υποδομών.

Η πρόβλεψη φορτίου έχει ταξινομηθεί από την άποψη της διαφορετικής διάρκειας σχεδιασμού: βραχυπρόθεσμη πρόβλεψη φορτίου ή STLF (μέχρι 1 ημέρα, μεσοπρόθεσμη πρόβλεψη φορτίου ή MTLF (1 ημέρα έως 1 έτος), καθώς και τη μακροπρόθεσμη πρόβλεψη φορτίου ή LTLF (1 - 10ετία). Για την πρόβλεψη του φορτίου με ακρίβεια όλο το χρόνο, διάφοροι εξωτερικοί παράγοντες, όπως καιρικές συνθήκες, η ηλιακή ακτινοβολία, ο πληθυσμός, το κατά κεφαλήν ακαθάριστο εγχώριο προϊόν εποχές και τις αργίες πρέπει να ληφθούν υπόψη. Για παράδειγμα, κατά τη χειμερινή περίοδο, ο μέσος συντελεστής ψυχρού αέρος θα μπορούσε να προστεθεί ως επεξηγηματική μεταβλητή επιπλέον εκείνων που χρησιμοποιούνται στο μοντέλο καλοκαιριού. Στις μεταβατικές εποχές, όπως την άνοιξη και το φθινόπωρο, η τεχνική μετασχηματισμού μπορεί να χρησιμοποιηθεί. Για τις διακοπές, ένα φορτίο επίδρασης διακοπών μπορεί να αφαιρεθεί από το κανονικό φορτίο για να εκτιμηθεί το πραγματικό φορτίο διακοπών καλύτερα.

Διάφορα μοντέλα πρόβλεψης έχουν αναπτυχθεί για την πρόβλεψη φορτίου με βάση διάφορες τεχνικές όπως η πολλαπλή παλινδρόμηση, εκθετική εξομάλυνση, επαναληπτική επαναστάθμιση ελαχίστων τετραγώνων, προσαρμοστική πρόβλεψη φορτίου, στοχαστική χρονοσειρά, ασαφή λογική, νευρωνικά δίκτυα και συστήματα βασισμένα στη γνώση των εμπειρογνομώνων. Μεταξύ αυτών, το πιο δημοφιλές STLF ήταν στοχαστικά μοντέλα χρονοσειρών, όπως αυτοπαλίνδρομο (AR) μοντέλο, κινούμενο αυτοπαλίνδρομο μοντέλο μέσου όρου (ARMA), αυτοπαλίνδρομο



Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
ολοκληρωμένο κινητό μέσο όρο (ARIMA) μοντέλο και άλλα μοντέλα με τη χρήση  
ασαφούς λογικής και νευρωνικών δικτύων.

Το DLF παρέχει δεδομένα συνάθροισης και δυνατότητες πρόβλεψης που έχει  
ρυθμιστεί για να ανταποκριθούν στις σημερινές απαιτήσεις και να προσαρμοστεί για  
να αντιμετωπίσει τις μελλοντικές απαιτήσεις και θα πρέπει να έχουν τη δυνατότητα  
να παράγουν επαναλαμβανόμενες και ακριβείς προβλέψεις.



## 2. Αναγνώριση Προτύπων & Μηχανική μάθηση

### 2.1 Εισαγωγή

Η μηχανική μάθηση (machine learning) είναι ένας τομέας της τεχνητής νοημοσύνης ο οποίος αφορά αλγόριθμους και μεθόδους που επιτρέπουν στους υπολογιστές να «μαθαίνουν». Με τη μηχανική μάθηση καθίσταται εφικτή η κατασκευή προσαρμόσιμων (adaptable) προγραμμάτων υπολογιστών τα οποία λειτουργούν με βάση την αυτοματοποιημένη ανάλυση συνόλων δεδομένων και όχι τη διαίσθηση των μηχανικών που τα προγραμμάτισαν. Η μηχανική μάθηση επικαλύπτεται σημαντικά με τη στατιστική, αφού και τα δύο πεδία μελετούν την ανάλυση δεδομένων.

Οι αλγόριθμοι μηχανικής μάθησης κατηγοριοποιούνται ανάλογα με το επιθυμητό αποτέλεσμα του αλγορίθμου. Οι συνηθέστερες κατηγορίες είναι οι εξής:

- Επιτηρούμενη μάθηση, επιβλεπόμενη μάθηση ή μάθηση με επίβλεψη (supervised learning), όπου ο αλγόριθμος κατασκευάζει μια συνάρτηση που απεικονίζει δεδομένες εισόδους σε γνωστές, επιθυμητές εξόδους (σύνολο εκπαίδευσης), με απώτερο στόχο τη γενίκευση της συνάρτησης αυτής και για εισόδους με άγνωστη έξοδο (σύνολο ελέγχου).
- Μη επιτηρούμενη μάθηση, μάθηση χωρίς επίβλεψη (unsupervised learning), όπου ο αλγόριθμος κατασκευάζει ένα μοντέλο για κάποιο σύνολο εισόδων χωρίς να γνωρίζει επιθυμητές εξόδους για το σύνολο εκπαίδευσης.
- Ενισχυτική μάθηση (reinforcement learning), όπου ο αλγόριθμος μαθαίνει μια στρατηγική ενεργειών για ένα δεδομένο στιγμιότυπο.

Η ανάλυση των αλγορίθμων μηχανικής μάθησης είναι ένας κλάδος της στατιστικής που ονομάζεται θεωρία μάθησης.

#### 2.1.1 Βασική Ορολογία

- *Σώμα ή Σύνολο ή σει Εκπαίδευσης* (training set): Σύνολο εισαγόμενων δεδομένων στο οποίο έχουν εκχωρηθεί εκ των προτέρων *ετικέτες* (labels).
- *Στιγμιότυπο* (instance): Εισαγόμενο αντικείμενο στο οποίο αναμένουμε να αποδοθεί μία τιμή από το σύστημα.
- *Χαρακτηριστικό ή Ιδιότητα* (attribute): Γνώρισμα του στιγμιότυπου βάσει του οποίου θα γίνει η ταξινόμηση των νέων αντικειμένων.

#### 2.1.2 Τυπική Διαδικασία

1. *Καθορισμός τύπου δεδομένων.*



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

2. *Δημιουργία σώματος εκπαίδευσης.* Το σώμα εκπαίδευσης πρέπει να είναι αντιπροσωπευτικό της πραγματικής φύσης του προβλήματος. Από μετρήσεις ή εξωτερική ανθρώπινη παρέμβαση αναθέτουμε τις ετικέτες στα δεδομένα του σώματος.
3. *Επιλογή κατάλληλων χαρακτηριστικών.* Μετατρέπουμε κάθε νέο αντικείμενο σε *διάνυσμα χαρακτηριστικών* (input vector) προκειμένου να ακολουθήσει η ταξινόμησή του. Είναι σημαντικό να επιλεγούν τα γνωρίσματα εκείνα που περιέχουν την απαραίτητη πληροφορία για την σωστή ανάθεση κλάσης ενώ παράλληλα ο αριθμός τους πρέπει να είναι υπολογιστικά διαχειρίσιμος. Επίσης πρέπει να γνωρίζουμε τον τύπο τιμής κάθε ιδιότητας (ονομαστικής (nominal), τακτικής (ordinal), αναλογικής (ratio)).
4. *Επιλογή αλγορίθμου εκπαίδευσης.* Υπάρχει πληθώρα αλγορίθμων και η επιλογή πρέπει να γίνει ανάλογα με το ποιος κρίνεται πιο αποδοτικός για τη συγκεκριμένη κατηγορία προβλημάτων.
5. *Αξιολόγηση της διαδικασίας.* Μετά την ολοκλήρωση της εκπαίδευσης η ακρίβεια της παραγόμενης σχέσης θα πρέπει να αξιολογηθεί με ένα σύνολο αξιολόγησης (evaluation set).

## 2.2 Αλγόριθμοι

Οι αλγόριθμοι για την αναγνώριση προτύπων εξαρτώνται από τον τύπο της εξόδου, σχετικά με το εάν η μάθηση είναι υπό επίβλεψη ή χωρίς και με το εάν ο αλγόριθμος είναι στατιστικός. Οι στατιστικοί αλγόριθμοι μπορούν να κατηγοριοποιηθούν περαιτέρω ως παραγωγικοί ή διακριτικοί.

- **Κατηγοριοποίησης (classification):** Αντιμετωπίζουν το πρόβλημα κατάταξης νέων αντικειμένων ή παρατηρήσεων (*labeling*) σε ήδη υπάρχουσες υποομάδες ενός δεδομένου συνόλου αντικειμένων (επιβλεπόμενη μάθηση). Ταξινομητής μέγιστης εντροπίας: Η λογιστική παλινδρόμηση είναι ένας αλγόριθμος για ταξινόμηση παρά την ονομασία του. (Το όνομα προέρχεται από το γεγονός ότι η λογιστική παλινδρόμηση χρησιμοποιεί μια επέκταση ενός γραμμικού μοντέλου παλινδρόμησης για να μοντελοποιήσει την πιθανότητα μιας εισόδου να ανήκει σε μια συγκεκριμένη κλάση). Άλλοι πασίγνωστοι αλγόριθμοι κατηγοριοποίησης είναι ο ταξινομητής Naive Bayes, Δέντρα αποφάσεων, λίστες αποφάσεων, Διάνυσμα υποστήριξης μηχανής (Support Vector Machines - SVM), Αλγόριθμοι εκτίμησης πυρήνα και K-κοντινότερου γείτονα, Perceptron, Νευρωνικά δίκτυα (perceptron πολλών επιπέδων).
- **Συσταδοποίησης (clustering):** Αντιμετωπίζουν το πρόβλημα διαχωρισμού ενός υπάρχοντος συνόλου αντικειμένου σε κατηγορίες (*clusters*) με βάση κριτήρια ομοιότητας (μη επιβλεπόμενη μάθηση). Οι πιο γνωστοί αλγόριθμοι είναι τα κατηγορηματικά ανάμικτα μοντέλα, η K-μέσων συσταδοποίηση, η ιεραρχική συσταδοποίηση και η ανάλυση κύριου συστατικού του πυρήνα (Kernel PCA).



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

- Παλινδρόμησης (Regression): Προβλέπουν την τιμή μίας μεταβλητής συνεχών τιμών συναρτήσει άλλων μεταβλητών.

## 2.2.1 Κανόνας Κατηγοριοποίησης & κατηγοριοποίηση στην Μηχανική μάθηση

Δεδομένου ένας πληθυσμός του οποίου τα μέλη μπορούν να διαχωριστούν σε έναν αριθμό διαφορετικών συνόλων ή τάξεις, ένας κανόνας κατηγοριοποίησης είναι μια διαδικασία στην οποία τα στοιχεία του συνόλου του πληθυσμού ανατίθενται το καθένα σε μια από τις κλάσεις. Ένα τέλειο τεστ είναι τέτοιο ώστε κάθε στοιχείο στον πληθυσμό να έχει εκχωρηθεί στην κατηγορία που ανήκει πραγματικά. Μια ατελής δοκιμή είναι τέτοια ώστε να εμφανίζονται μερικά λάθη και, στη συνέχεια η στατιστική ανάλυση θα πρέπει να εφαρμοστεί για την ανάλυση της ταξινόμησης.

Ένα ιδιαίτερο είδος του κανόνα ταξινόμησης είναι δυαδικές ταξινομήσεις. Στην μηχανική μάθηση και στην Στατιστική, η κατηγοριοποίηση είναι το ζήτημα της αναγνώρισης μιας νέας παρατήρησης (observation ή στιγμιότυπο) σε ποιο σύνολο κατηγοριών ανήκει, με βάση ένα σύνολο εκπαίδευσης που περιέχει τις παρατηρήσεις (ή περιπτώσεις) των οποίων η κατηγορία μελών είναι γνωστή. Οι επιμέρους παρατηρήσεις αναλύονται σε ένα σύνολο ποσοτικών ιδιοτήτων, γνωστές ως επεξηγηματικές μεταβλητές, χαρακτηριστικά, κλπ. Αυτές οι ιδιότητες μπορούν να είναι: *κατηγορηματικές* (π.χ. "Α", "Β", "ΑΒ" ή "Ο", για τον τύπο του αίματος) , *τακτικές* (π.χ. "μεγάλο", "μεσαίο" ή "μικρό"), *ακέραιες* (π.χ. τον αριθμό των εμφανίσεων μιας λέξης μέρος σε ένα e-mail) ή σε *πραγματικές* (π.χ. μέτρηση της πίεσης του αίματος). Ορισμένοι αλγόριθμοι λειτουργούν μόνο ως προς τα διακριτά στοιχεία και απαιτούν τα πραγματικά ή ακέραια δεδομένα να χωρίζονται σε ομάδες (π.χ. λιγότερο από 5, μεταξύ 5 και 10, ή μεγαλύτερη από 10). Ένα παράδειγμα θα ήταν η ανάθεση ενός δεδομένου ηλεκτρονικού ταχυδρομείου σε «spam» ή «μη-spam» τάξεις ή αναθέτοντας μια διάγνωση σε ένα δεδομένο ασθενή, όπως περιγράφεται από παρατηρούμενα χαρακτηριστικά του ασθενούς (φύλο, την πίεση του αίματος, παρουσία ή απουσία ορισμένων συμπτωμάτων, κλπ. ).

Ένας αλγόριθμος που υλοποιεί την ταξινόμηση, ειδικά σε μια συγκεκριμένη εφαρμογή, είναι γνωστό ως ένας *ταξινομητής*. Ο όρος «ταξινομητής» μερικές φορές αναφέρεται επίσης στην μαθηματική συνάρτηση, που εφαρμόζεται από έναν αλγόριθμο ταξινόμησης, που καθοδηγούν τα δεδομένα εισόδου σε μια κατηγορία.

## 2.2.2 Σχέση με άλλα προβλήματα

Η Ταξινόμηση και ομαδοποίηση είναι παραδείγματα της αναγνώρισης προτύπων, οι οποίες είναι οι εκχωρήσεις κάποιου είδους τιμή εξόδου σε μία δεδομένη τιμή





Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου εισόδου. Άλλα παραδείγματα είναι η παλινδρόμηση (regression), η οποία αποδίδει μια πραγματική τιμή εξόδου για κάθε είσοδο, εν συνεχεία η παλινδρόμηση εκχωρεί μια τάξη σε κάθε μέλος μια σειρά τιμών για παράδειγμα, μέρος της σήμανσης μιας ομιλίας, η οποία αποδίδει ένα μέρος της ομιλίας σε κάθε λέξη σε μια πρόταση εισόδου και τέλος η διαχώριση/ανάλυση parse, όπου η παλινδρόμηση αποδίδει ένα δέντρο parse σε μια πρόταση εισόδου, που περιγράφει την συντακτική δομή της πρότασης.

Μια κοινή υποκατηγορία της ταξινόμησης είναι η πιθανοτική ταξινόμηση. Αλγόριθμοι αυτού του είδους χρησιμοποιούν στατιστική επαγωγή για να βρεθεί η βέλτιστη κλάση για ένα δεδομένο στιγμιότυπο. Σε αντίθεση με άλλους αλγορίθμους που έχουν ως έξοδο μια «βέλτιστη» κλάση, οι πιθανοτικοί αλγόριθμοι έχουν έξοδο μια πιθανότητα του στιγμιότυπου που είναι μέλος σε καθεμία πιθανή κλάση. Η καλύτερη κλάση επιλέγεται με την μεγαλύτερη πιθανότητα. Ωστόσο, ένας τέτοιος αλγόριθμος έχει πολυάριθμα πλεονεκτήματα έναντι των μη-πιθανοτικών ταξινομητών:

- Μπορεί να εξάγει μια τιμή εμπιστοσύνης που σχετίζεται με την επιλογή της (σε γενικές γραμμές, ένας ταξινομητής που μπορεί να κάνει αυτό είναι γνωστό ως ένα ταξινομητής με σίγουρα βάρη).
- Αντίστοιχα, μπορεί να απόσχει όταν είναι σίγουρος για την επιλογή οποιαδήποτε συγκεκριμένης παραγωγής που είναι πολύ χαμηλή.
- Λόγω των πιθανοτήτων εξόδου, πιθανοτικοί ταξινομητές μπορούν να ενσωματωθούν πιο αποτελεσματικά σε μεγαλύτερα καθήκοντα μηχανικής μάθησης, με τρόπο ώστε μερικώς ή πλήρως να αποφεύγουν το πρόβλημα της διάδοσης των σφαλμάτων.

### 2.2.3 Μπεϋζιανή Κατηγοριοποίηση.

Οι διαδικασίες Bayesian ταξινόμησης παρέχουν ένα φυσικό τρόπο, λαμβάνοντας υπόψη όλες τις διαθέσιμες πληροφορίες σχετικά με τα αντίστοιχα μεγέθη των υπο-πληθυσμών που σχετίζονται με τις διάφορες ομάδες εντός του συνολικού πληθυσμού. Οι διαδικασίες Bayesian τείνουν να είναι υπολογιστικά δαπανηρές, όμως μετά την Markov - Monte Carlo αλυσίδα οι υπολογισμοί αναπτύχθηκαν, οι προσεγγίσεις για κανόνες Bayesian ομαδοποίησης επινοήθηκαν.

Μερικές Bayesian διαδικασίες περιλαμβάνουν τον υπολογισμό των ομαδικών συμμετοχικών πιθανοτήτων: αυτές μπορούν να θεωρηθούν ως παροχή μιας πληρέστερης αποτελεσμάτων μιας ανάλυσης δεδομένων από μια απλή απόδοση μιας ενιαίας ομάδας σε κάθε νέο στιγμιότυπο.



### 2.2.3 Διανύσματα Χαρακτηριστικών

Οι περισσότεροι αλγόριθμοι περιγράφουν ένα ξεχωριστό στιγμιότυπο του οποίου η κατηγορία προβλέπεται χρησιμοποιώντας ένα διάνυσμα χαρακτηριστικών και μετρήσιμων ιδιοτήτων του περιστατικού. Κάθε ιδιότητα έχει οριστεί ως χαρακτηριστικό, επίσης γνωστό στην Στατιστική ως ερμηνευτική μεταβλητή (ή ανεξάρτητη μεταβλητή, αν και σε γενικές γραμμές διαφορετικά χαρακτηριστικά μπορούν ή δεν μπορούν να είναι στατιστικά ανεξάρτητες). Εάν το στιγμιότυπο είναι μια εικόνα, τότε οι τιμές των χαρακτηριστικών ενδέχεται να αντιστοιχούν στα εικονοστοιχεία της μιας εικόνας. Εάν το στιγμιότυπο είναι ένα κομμάτι του κειμένου, τότε οι τιμές των χαρακτηριστικών ενδέχεται να είναι η συχνότητα εμφάνισης διαφορετικών λέξεων. Ορισμένοι αλγόριθμοι λειτουργούν μόνο ως προς τα διακριτά στοιχεία και απαιτούν τα πραγματικά ή ακέραια δεδομένα να χωρίζονται σε ομάδες (π.χ. λιγότερο από 5, μεταξύ 5 και 10, ή μεγαλύτερη από 10).

Το διάστημα διανυσμάτων που σχετίζεται με αυτά τα διανύσματα συχνά αποκαλείται ως το διάστημα χαρακτηριστικών. Προκειμένου να μειωθεί η διάσταση του διαστήματος χαρακτηριστικών, ένας αριθμός τεχνικών μείωσης της διάστασης μπορούν να χρησιμοποιηθούν.

### 2.2.4 Δυαδική και πολλών κλάσεων Κατηγοριοποίηση.

Η ταξινόμηση μπορεί να θεωρηθεί ως δύο ξεχωριστά προβλήματα - *δυαδική ταξινόμηση* και *ταξινόμηση πολλών κλάσεων (multiclass)*. Στην δυαδική ταξινόμηση, ένα καλύτερα κατανοητό έργο, μόνο δύο κλάσεις συμμετέχουν, ενώ στην κατάταξη multiclass συνεπάγεται την τοποθέτηση ενός αντικειμένου σε μία από τις πολλές κλάσεις. Επειδή πολλές μέθοδοι ταξινόμησης έχουν αναπτυχθεί ειδικά για δυαδικής ταξινόμησης, η ταξινόμηση multiclass συχνά απαιτεί τη συνδυασμένη χρήση πολλαπλών δυαδικών ταξινομητών.

### 2.2.5 Γραμμικοί Ταξινομητές

Ένας μεγάλος αριθμός των αλγορίθμων ταξινόμησης μπορεί να διατυπωθεί ως μια γραμμική συνάρτηση που αποδίδει ένα σκορ σε κάθε πιθανή κατηγορία  $k$  συνδυάζοντας το διάνυσμα χαρακτηριστικών ενός στιγμιότυπου με ένα διάνυσμα βαρών, χρησιμοποιώντας ένα εσωτερικό γινόμενο. Η προβλεπόμενη κατηγορία είναι αυτή με την υψηλότερη βαθμολογία. Αυτός ο τύπος λειτουργίας σκορ είναι γνωστή ως μία γραμμική συνάρτηση πρόβλεψης και έχει την ακόλουθη γενική μορφή:



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

$$\text{score}(\mathbf{X}_i, k) = \beta_k \cdot \mathbf{X}_i,$$

Όπου  $\mathbf{X}_i$  είναι το διάνυσμα χαρακτηριστικών, για παράδειγμα  $i$ ,  $\beta_k$  είναι το διάνυσμα των βαρών που αντιστοιχούν στην κατηγορία  $k$ , και το σκορ  $(\mathbf{X}_i, k)$  είναι το σκορ που σχετίζονται με την ανάθεση στιγμιότυπου  $i$  στην κατηγορία  $k$ . Στη διακριτική θεωρία επιλογής, όπου τα στιγμιότυπα αντιπροσωπεύουν ανθρώπους και οι κατηγορίες αντιπροσωπεύουν επιλογές, το σκορ θεωρείται το βοηθητικό εργαλείο που συνδέεται με άτομο  $i$  επιλέγοντας  $k$  κατηγορία.

Οι Αλγόριθμοι με αυτή τη βασική ρύθμιση είναι γνωστή ως γραμμικοί ταξινομητές. Αυτό που τους διακρίνει είναι η διαδικασία για τον προσδιορισμό (εκπαίδευση) των βέλτιστων βαρών / συντελεστών και τον τρόπο με τον οποίο το σκορ ερμηνεύεται.

### 2.3 Δέντρα Αποφάσεων (Decision Trees)

Τα Δέντρα αποφάσεων (Decision Trees) διακρίνονται σε δύο κατηγορίες, κατηγοριοποίησης και παλινδρόμησης (classification & regression trees). Η διαφορά μεταξύ αυτών των κατηγοριών είναι το είδος της προβλεπόμενης κατηγοριοποίησης (responses). Τα classification trees προβλέπουν και κατηγοριοποιούν τα δεδομένα σε ονομαστικές τάξεις (Nominal), δηλαδή 'True' & 'False', ενώ τα regression trees δίνουν αριθμητικές κατηγορίες (Numeric).

#### Θετικά των Δέντρων Αποφάσεων:

- Εύκολα στην κατανόηση και εξήγηση.
- Απαιτούν ελάχιστη προετοιμασία τα δεδομένα. Άλλες τεχνικές απαιτούν τα δεδομένα να μην έχουν κενές τιμές.
- Ικανά να αντιμετωπίσουν αριθμητικά και κατηγορηματικά δεδομένα.
- Χρησιμοποιούν το μοντέλο «άσπρο κουτί», δηλαδή σε μια κατάσταση που παρατηρείται από τρίτους, είναι εύκολο να κατανοηθεί χρησιμοποιώντας Boolean λογική.
- Εκτελούν ικανοποιητικά με μεγάλο όγκο δεδομένων σε σύντομο χρονικό διάστημα.

#### Αρνητικά των Δέντρων Αποφάσεων:

- Ορισμένες φορές δημιουργούνται πολύπλοκα δέντρα τα οποία δεν κάνουν καλή κατηγοριοποίηση. Το φαινόμενο αυτό λέγεται «over fitting», και μηχανισμοί όπως το «κλάδεμα» είναι απαραίτητοι για την αποφυγή του προβλήματος.



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

- Υπάρχουν ειδικές περιπτώσεις όπου το σύστημα δεν μπορεί να τις περιγράψει με δέντρα αποφάσεων, Όπως το πρόβλημα του λογικού XOR και πολυπλεξίας.
- Για δεδομένα που περιέχουν κατηγορικές μεταβλητές με διαφορετικό αριθμό επιπέδων, η πληροφορία που αντλείται από τα δέντρα αποφάσεων είναι υπέρ αυτών των ιδιοτήτων με περισσότερα επίπεδα.

Εφαρμογές των classification & regression trees με χρήση του statistical toolbox του προγράμματος Matlab2011a στα σετ δεδομένων «fisheriris» & «carsmall».

### Δέντρο Κατηγοριοποίησης.

```
tc = ClassificationTree.fit(meas, species)
```

Η συνάρτηση ClassificationTree.fit δέχεται 2 κύρια ορίσματα, τον πίνακα με το σετ των δεδομένων και τον πίνακα των responses. Η «view» μας παρουσιάζει κάποια κύρια χαρακτηριστικά και ιδιότητες του δέντρου καθώς μπορούμε να δούμε την γραφική παράσταση του δέντρου .

```
tc =
```

ClassificationTree:

```
PredictorNames: {'x1' 'x2' 'x3' 'x4'}
```

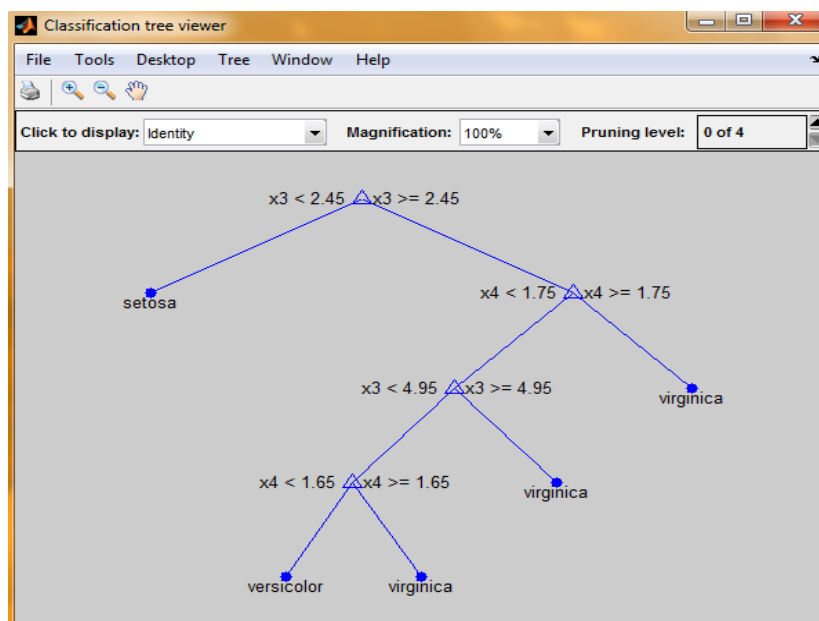
```
CategoricalPredictors: []
```

```
ResponseName: 'Y'
```

```
ClassNames: {'setosa' 'versicolor' 'virginica'}
```

```
ScoreTransform: 'none'
```

```
NObservations: 150
```

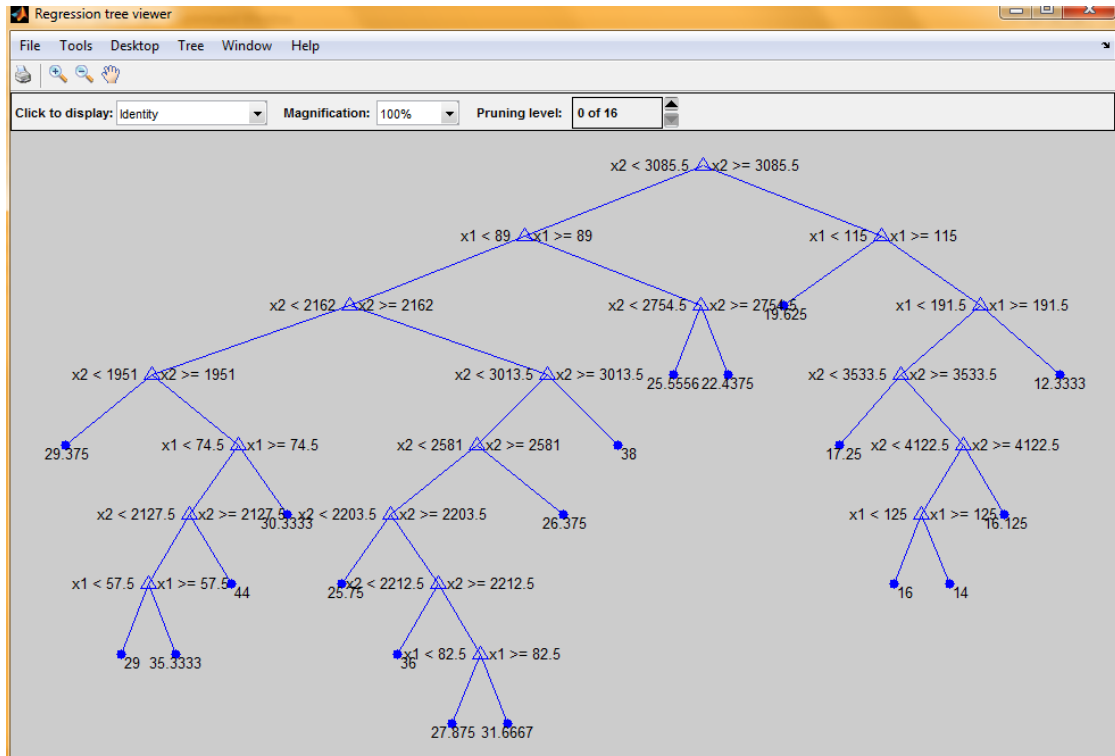




## Δέντρο Παλνδρόμησης

```
X = [Horsepower Weight];  
rtree = RegressionTree.fit(X,MPG);
```

Ομοίως εδώ χρησιμοποιούμε την συνάρτηση `RegressionTree.fit`, εκτελώντας το παραπάνω κομμάτι κώδικα αντλούμε τα εξής.



### 2.3.1 Λειτουργία της μεθόδου Fit.

Οι μέθοδοι `ClassificationTree.fit` & `RegressionTree.fit` ακολουθούν τα εξής βήματα για την δημιουργία του δέντρου:

- «Ξεκίνα» με όλα τα δεδομένα εισόδου, και «υπολόγισε» κάθε δυαδικό διαχωρισμό για κάθε μεταβλητή.
- «Διάλεξε» τον διαχωρισμό που έχει το καλύτερο κριτήριο βελτιστοποίησης: Εάν ο διαχωρισμός οδηγεί σε παιδί κόμβο που έχει λίγα observations (λιγότερους από την `Minleaf` (δλδ ο ελάχιστος αριθμός observations ανά φύλλο) παράμετρο), «επέλεξε» διαχωρισμό με καλύτερο κριτήριο βελτιστοποίησης ως προς τον περιορισμό του `minleaf`.
- «Επέβαλλε» τον διαχωρισμό.
- «Επανάλαβε» αναδρομικά και για τους δύο κόμβους παιδί.



### 2.3.2 Βελτίωση των classification & regression trees

Ρυθμίζοντας παραμέτρους στα δέντρα μπορούμε να πετύχουμε καλύτερη ταξινόμηση ή πρόβλεψη δεδομένων.

- Μελέτη του Resubstitution Error.
- Διασταυρωμένη επικύρωση (Cross Validation).
- Έλεγχος βάθους του δέντρου (leafiness).
- Κλάδεμα (prune).

Το Resubstitution σφάλμα είναι η διαφορά μεταξύ του σετ απάντησης δεδομένων εκπαίδευσης και των προβλέψεων που κάνει το δέντρο των απαντήσεων βάσει του σετ εισόδου δεδομένων εκπαίδευσης. Αν το σφάλμα είναι πολύ μεγάλο, τότε οι αναμενόμενες προβλέψεις δεν θα είναι τόσο ακριβείς.

*classification*

```
resuberror = resubLoss(ct)  
resuberror = 0.002
```

*regression*

```
resuberror = resubLoss(rtree)  
resuberror = 5.8158
```

Για να έχουμε καλύτερη αίσθηση της προγνωστικής ακρίβειας για τα νέα δεδομένα εφαρμόζουμε την μέθοδο cross Validate στο δέντρο. Η μέθοδος διαχωρίζει το σετ εκπαίδευσης σε 10 τυχαία κομμάτια. Εκπαιδεύει 10 νέα δέντρα, το καθένα από 9 τμήματα δεδομένων. Εν συνέχεια εξετάζει την προγνωστική ακρίβεια του κάθε νέου δέντρου χωρίς να περιλαμβάνει την εκπαίδευση αυτού του δέντρου. Αυτή η μέθοδος δίνει καλή εκτίμηση για την προγνωστική ακρίβεια του δέντρου.

*Regression*

```
cvrtree = crossval(rtree);  
cvloss = kfoldLoss(cvrtree)  
cvloss = 25.5703
```

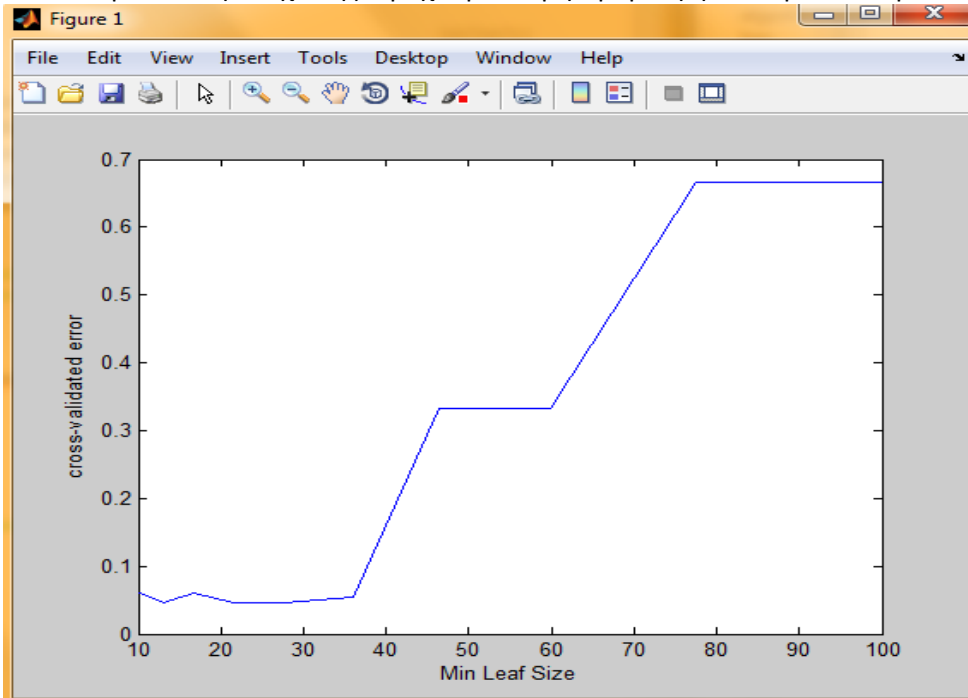
Το «cross validation» σφάλμα είναι σχεδόν 25, που σημαίνει ότι ένα τυπικό προγνωστικό σφάλμα του δέντρου πάνω στα νέα δεδομένα είναι περίπου 5. Αυτό αναδεικνύει ότι το «cross validation» σφάλμα είναι συνήθως μεγαλύτερο από το Resubstitution σφάλμα.

Όταν επεκτείνεται ένα μεγάλο δέντρο, εξετάζεται η απλότητα και η προγνωστική ικανότητά του. Ένα δέντρο με μεγάλο βάθος είναι πολύ ακριβές στα δεδομένα εκπαίδευσης. Όμως, το δέντρο δεν εγγυάται ότι θα είναι ακριβές για ανεξάρτητα δεδομένα ελέγχου. Έτσι ο σωστός συνδυασμός είναι εκείνος που θα κάνει το δέντρο να είναι ακριβές στις προβλέψεις και θα κάνει το δέντρο απλό.

Στο παράρτημα α) παρουσιάζεται ο κώδικας με τις διάφορες δοκιμές της μεταβλητής που καθορίζει τον ελάχιστο αριθμό δεδομένων ανά φύλλο.

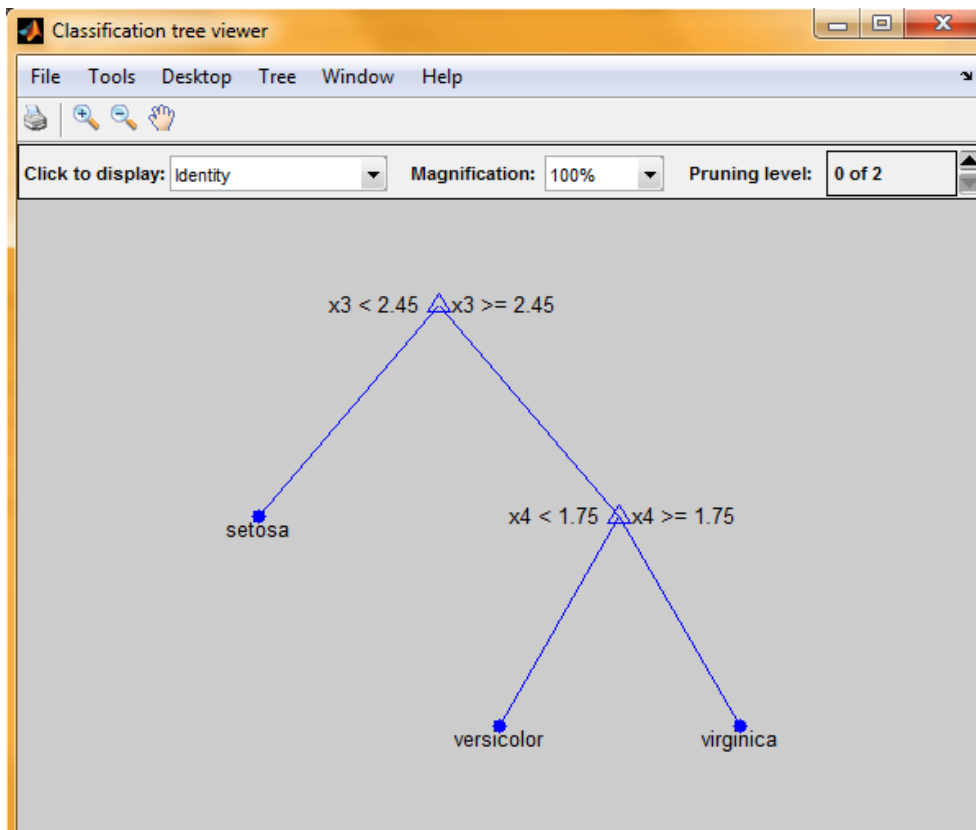


Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου



% Παρατηρούμε ότι το καλύτερο μέγεθος φύλλου είναι 20 με 35 observations ανά φύλλο.

```
OptimalTree = ClassificationTree.fit(X,Y,'minleaf',40);  
view(OptimalTree,'mode','graph')
```



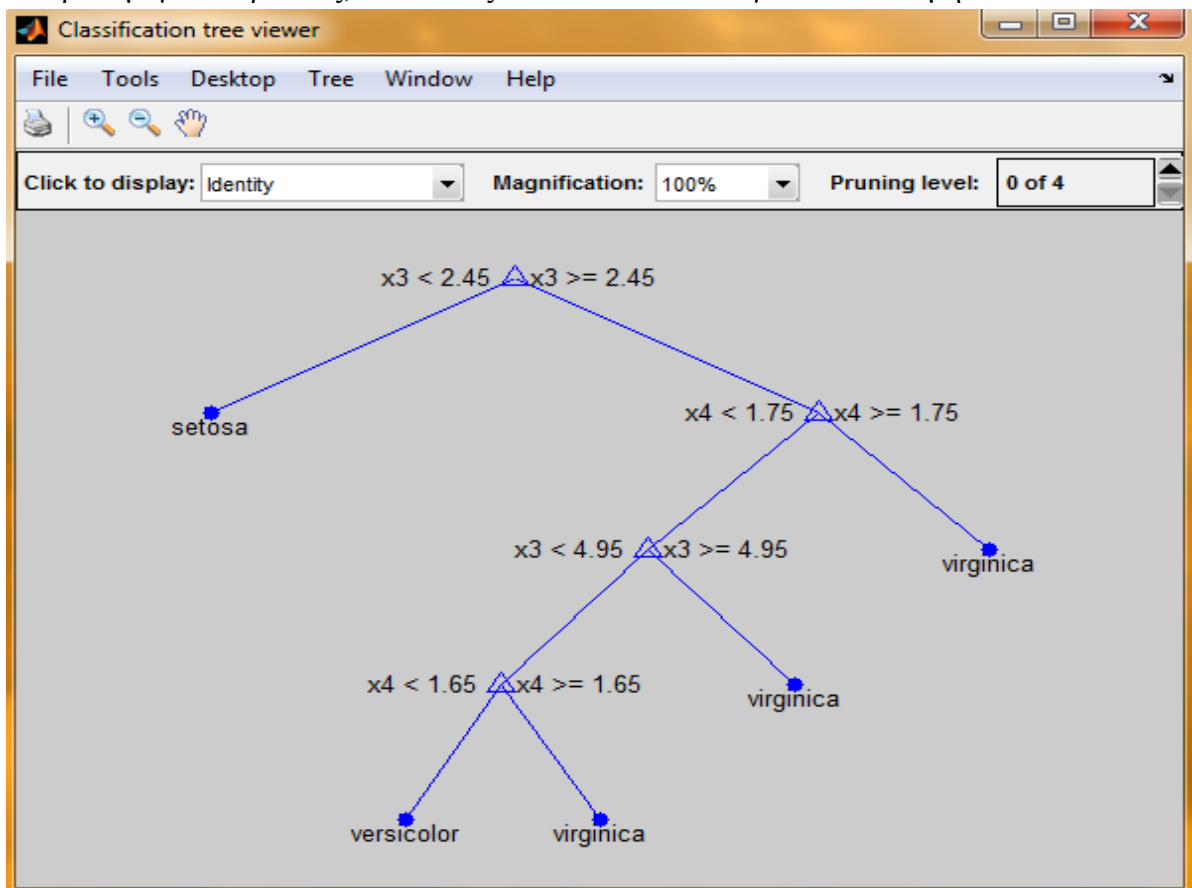


Αν υπολογίσουμε το σφάλμα του αρχικού δέντρου και του καινούριου έχουμε τα εξής:

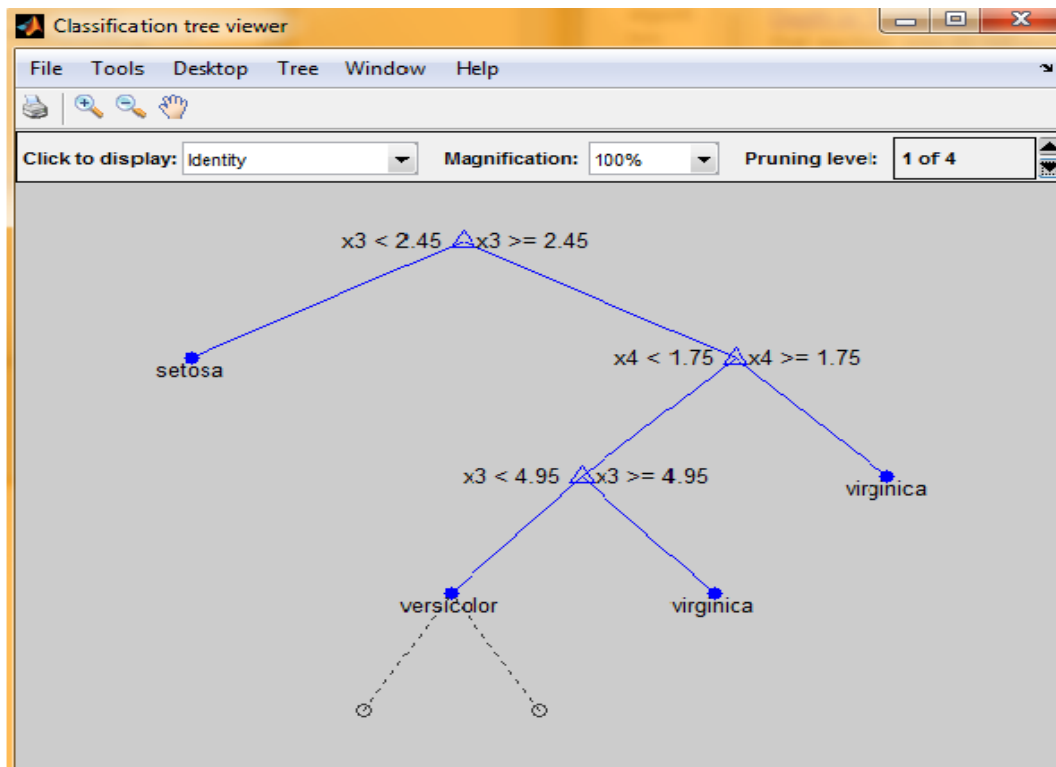
```
resubOpt = resubLoss(OptimalTree);  
lossOpt = kfoldLoss(crossval(OptimalTree));  
resubDefault = resubLoss(DefaultTree);  
lossDefault = kfoldLoss(crossval(DefaultTree));  
  
resubOpt = 0.0400  
resubDefault = 0.0200  
  
lossOpt = 0.0600  
lossDefault = 0.0400
```

Το καινούριο δέντρο είναι πολύ μικρότερο από το αρχικό, αλλά δίνει μεγαλύτερο resubstitution σφάλμα, ομοίως δίνει αντίστοιχη ακρίβεια για τα διασταυρωμένα (cross validated) δεδομένα.

Μια άλλη μέθοδος βελτιστοποίησης των δέντρων είναι το κλάδεμα (pruning). Η μέθοδος συγχωνεύει φύλλα στο ίδιο κλαδί. Έτσι δεν χρειάζεται να δημιουργηθεί καινούριο δέντρο για κάθε μέγεθος κόμβου, αντί αυτού προτιμάται η δημιουργία δέντρου μεγάλου βάθους, και τέλος κλαδεύεται ανάλογα στο επιθυμητό επίπεδο.







## 2.4 Μάθηση Συνόλου (Ensemble Learning)

Στις στατιστικές και μηχανικής μάθησης, οι μέθοδοι συνόλου χρησιμοποιούν πολλαπλά μοντέλα για να ληφθεί καλύτερη απόδοση πρόβλεψης από ό, τι θα μπορούσε να ληφθεί από κάποιο από τα συστατικά μοντέλα. Σε αντίθεση με ένα σύνολο στατιστικών σε στατιστική μηχανική, η οποία συνήθως είναι άπειρη, ένα σύνολο μηχανής μάθησης αναφέρεται μόνο σε ένα συγκεκριμένο πεπερασμένο σύνολο εναλλακτικών μοντέλων.

### 2.4.1 Επισκόπηση

Οι αλγόριθμοι μάθησης Με επίβλεψη συνήθως περιγράφονται ως την εκτέλεση της εργασίας της αναζήτησης μέσα από ένα διάστημα υποθέσεων για να βρεθεί μια κατάλληλη υπόθεση που θα κάνει καλές προβλέψεις με ένα συγκεκριμένο πρόβλημα. Ακόμα και αν ο χώρος υποθέσεων περιέχει υποθέσεις που είναι πολύ κατάλληλος για ένα συγκεκριμένο πρόβλημα, μπορεί να είναι πολύ δύσκολο να βρεθεί ένα καλό αποτέλεσμα. Σύνολα συνδυάζουν πολλαπλές υποθέσεις για να σχηματίσουν ένα (ευελπιστούμε) την καλύτερη υπόθεση. Με άλλα λόγια, ένα σύνολο



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

είναι μια τεχνική για το συνδυασμό πολλών αδύναμων μαθητών σε μια προσπάθεια να παράγουν ένα ισχυρό μαθητή. Ο όρος σύνολο συνήθως προορίζεται για μεθόδους που δημιουργούν πολλαπλές υποθέσεις χρησιμοποιώντας το ίδιο βασικό μαθητή. Ο ευρύτερος όρος του ταξινομητή πολλαπλών συστημάτων καλύπτει επίσης υβριδοποίηση των υποθέσεων που δεν προκαλείται από τον ίδιο βασικό μαθητή.

Η αξιολόγηση της πρόβλεψης ενός συνόλου απαιτεί συνήθως περισσότερους υπολογισμούς από την αξιολόγηση της πρόβλεψης ενός ενιαίου μοντέλου, έτσι σύνολα μπορούν να θεωρηθούν ως ένας τρόπος για την αντιστάθμιση των αλγορίθμων κακής μάθησης εκτελώντας πολλούς επιπλέον υπολογισμούς. Γρήγοροι αλγόριθμοι όπως δέντρα αποφάσεων χρησιμοποιούνται συνήθως με σύνολα (για παράδειγμα Random Forest - Τυχαίο Δάσος), αν και πιο αργοί αλγόριθμοι μπορούν να επωφεληθούν από τις τεχνικές συνόλου.

## 2.4.2 Θεωρία Συνόλου

Ένα σύνολο είναι από μόνο του ένας αλγόριθμος εκπαίδευσης με επίβλεψη, επειδή μπορεί να εκπαιδευτεί και στη συνέχεια να χρησιμοποιηθεί για να κάνει προβλέψεις. Το εκπαιδευμένο σύνολο, συνεπώς, αντιπροσωπεύει μία μόνο υπόθεση. Η υπόθεση αυτή, ωστόσο, δεν είναι αναγκαστικά περιεχόμενη εντός του χώρου υπόθεσης των μοντέλων από τα οποία έχει κατασκευαστεί. Έτσι, τα σύνολα αποδεδειγμένα έχουν μεγαλύτερη ευελιξία στις λειτουργίες που μπορούν να εκπροσωπούν. Η ευελιξία αυτή μπορεί, θεωρητικά, να τους δώσει τη δυνατότητα να ταιριάζει - Fit τα δεδομένα εκπαίδευσης περισσότερο από ένα ενιαίο μοντέλο, αλλά στην πράξη, ορισμένες τεχνικές συνόλου (ειδικά ενσακκίσεως - bagging) τείνουν να μειώσουν τα προβλήματα που σχετίζονται με υπερβολική στερέωση - overfitting των δεδομένων εκπαίδευσης.

Εμπειρικά, τα σύνολα τείνουν να αποδίδουν καλύτερα αποτελέσματα όταν υπάρχει μια σημαντική διαφοροποίηση μεταξύ των μοντέλων. Πολλές μέθοδοι συνόλου, ως εκ τούτου, επιδιώκουν να προωθήσουν την ποικιλομορφία μεταξύ των μοντέλων που συνδυάζουν. Οι τυχαίοι αλγόριθμοι (όπως τυχαία δέντρα αποφάσεων) μπορούν να χρησιμοποιηθούν για να παραχθεί ένα ισχυρότερο σύνολο από πιο συνειδητοποιημένους αλγόριθμους (όπως εντροπίας αναγωγικά δέντρα αποφάσεων). Χρησιμοποιώντας μια ποικιλία ισχυρών αλγορίθμων μάθησης, ωστόσο, έχει αποδειχθεί ότι είναι πιο αποτελεσματική από τη χρήση τεχνικών που επιχειρούν να αχρηστεύσουν τα μοντέλα για την προώθηση της ποικιλομορφίας.

## 2.4.3 Τύποι των συνόλων

### Bootstrap aggregating (bagging)

Bootstrap συγκέντρωση, συχνά συντομογραφία ενσάκκιση bagging, περιλαμβάνει κάθε μοντέλο στην ψηφοφορία συνόλου με την ίδια βαρύτητα. Προκειμένου να προωθηθεί η μοντέλο διακύμανση, η ενσάκκιση εκπαιδεύει κάθε μοντέλο στο σύνολο χρησιμοποιώντας ένα τυχαίο υποσύνολο του συνόλου



Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
εκπαίδευσης. Ως ένα παράδειγμα, ο αλγόριθμος τυχαίων δασών συνδυάζει τυχαία  
δέντρα αποφάσεων ενσακκίσεως να επιτύχουμε πολύ υψηλή ακρίβεια ταξινόμησης.

### Ενίσχυση (Boosting)

Η ενίσχυση – boosting περιλαμβάνει αυξητικά την δημιουργία ενός συνόλου με την εκπαίδευση κάθε μοντέλο νέου στιγμιότυπου για να τονίσει τα στιγμιότυπα εκπαίδευσης που τα προηγούμενα μοντέλα ταξινόμησαν εσφαλμένα. Σε ορισμένες περιπτώσεις, η ενίσχυση έχει αποδειχθεί να δώσει καλύτερη ακρίβεια από την ενσάκκιση, αλλά τείνει επίσης να υπερ-ταιριάζει - overfit τα δεδομένα εκπαίδευσης. Μέχρι στιγμής, η πιο κοινή εφαρμογή της Ενίσχυσης είναι η Adaboost, αν και ορισμένοι νεότεροι αλγόριθμοι έχουν αναφερθεί ότι πετυχαίνουν καλύτερα αποτελεσμάτα.

## **2.5 Τυχαία Δάση (Random Forests)**

Το Τυχαίο Δάσος (Random Forest) είναι ένας ταξινομητής ο οποίος αποτελείται από πολλά δέντρα απόφασης και έχει ως έξοδο την τάξη που παρουσιάζεται περισσότερο από τα επιμέρους δέντρα. Ο αλγόριθμος ο οποίος περιείχε τα τυχαία δέντρα αναπτύχθηκε από τους Leo Breiman και Adele Cutler, και τα “Random Forests” αποτελούν το σήμα κατατεθέν τους. Ο όρος προήλθε από τα τυχαία δάση αποφάσεων που προτάθηκε για πρώτη φορά από τον Tim Ho Kam των Bell Labs το 1995. Η μέθοδος αυτή συνδυάζει την ιδέα «ενσάκκιση – bagging» του Breiman και την τυχαία επιλογή των χαρακτηριστικών, που εισήχθη ανεξάρτητα από Ho και Amit και Geman, προκειμένου να κατασκευάσει μια συλλογή από δέντρα αποφάσεων με ελεγχόμενη διακύμανση. Η επιλογή ενός τυχαίου υποσυνόλου χαρακτηριστικών είναι ένα παράδειγμα της μεθόδου τυχαίας υποχώρος, η οποία, στη διαμόρφωση του Ho, είναι ένας τρόπος για την εφαρμογή στοχαστική διακρίσεων προτεινόμενη από τον Eugene Kleinberg.

Όταν το σετ εκπαίδευσης για το τρέχον δέντρο έχει σχεδιαστεί με δειγματοληψία και αντικατάσταση, περίπου το ένα τρίτο των περιπτώσεων παραβλέπονται. Αυτά τα «εκτός» δεδομένα (oob, out-of-bag) χρησιμοποιούνται για την τρεχούμενη αμερόληπτη εκτίμηση του σφάλματος ταξινόμησης, όπως τα δέντρα έχουν προστεθεί στο δάσος.



### 2.5.1 Λειτουργία Αλγορίθμου Random Forest

Κάθε δέντρο κατασκευάζεται χρησιμοποιώντας τον ακόλουθο αλγόριθμο:

- Έστω ο αριθμός των στιγμιότυπων κατάρτισης/εκπαίδευσης να είναι  $N$ , και ο αριθμός των μεταβλητών στο ταξινομητή να  $M$ .
- Μας λένε ότι ο αριθμός  $m$  των μεταβλητών εισόδου που θα χρησιμοποιηθεί για τον προσδιορισμό της απόφασης σε ένα κόμβο του δέντρου, το  $m$  θα πρέπει να είναι πολύ μικρότερο από  $M$ .
- Επιλέξτε ένα σύνολο εκπαίδευσης για αυτό το δέντρο επιλέγοντας  $n$  φορές με αντικατάσταση από όλες τις διαθέσιμες περιπτώσεις εκπαίδευσης  $N$  (δηλαδή, να λάβει ένα δείγμα bootstrap). Χρησιμοποιήστε τις υπόλοιπες περιπτώσεις να υπολογιστεί το σφάλμα του δέντρου, με την πρόβλεψη των τάξεων τους.
- Για κάθε κόμβο του δέντρου, επιλέξτε τυχαία  $m$  μεταβλητές στις οποίες θα βασιστεί η απόφαση σε αυτόν τον κόμβο. Υπολογίζεται η καλύτερη διάσπαση με βάση αυτές τις μεταβλητές  $m$  στο σύνολο εκπαίδευσης.
- Κάθε δέντρο έχει αναπτυχθεί πλήρως και δεν κλαδεύεται (όπως μπορεί να γίνει με την κατασκευή ενός κανονικού δέντρου ταξινόμησης).

Για την πρόβλεψη ένα νέο δείγμα πιέζεται προς τα κάτω το δέντρο. Ανατίθεται η ετικέτα του δείγματος εκπαίδευσης στον τερματικό κόμβο. Η διαδικασία αυτή επαναλαμβάνεται για όλα τα δέντρα στο σύνολο, και η λειτουργία ψηφοφορίας όλων των δένδρων αναφέρεται ως πρόβλεψη τυχαίου δάσους.

Τα πλεονεκτήματα των τυχαίων δασών είναι:

- Είναι ανυπέρβλητοι σε ακρίβεια μεταξύ των σημερινών αλγορίθμων.
- Λειτουργεί αποτελεσματικά σε μεγάλες βάσεις δεδομένων.
- Μπορεί να χειριστεί χιλιάδες μεταβλητές εισόδου χωρίς να διαγράψει καμία.
- Δίνει εκτιμήσεις για το τι μεταβλητές είναι σημαντικές στην κατάταξη.
- Έχει μια αποτελεσματική μέθοδο για την εκτίμηση των ελλειπόντων στοιχείων και διατηρεί την ακρίβεια, όταν ένα μεγάλο μέρος των δεδομένων λείπουν.
- Προσφέρει μια πειραματική μέθοδο για την ανίχνευση αλληλεπιδράσεων των μεταβλητών
- Παράγει μια εσωτερική αμερόληπτη εκτίμηση του γενικευμένου σφάλματος καθώς προχωρά η οικοδόμηση δασών.
- Έχει μεθόδους για την εξισορρόπηση σφάλματος σε ταξινομημένα σύνολα πληθυσμού ανισόρροπων σετ δεδομένων.
- Πρωτότυπα υπολογίζονται που δίνουν πληροφορίες για τη σχέση μεταξύ των μεταβλητών και την ταξινόμηση.



## Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

- Υπολογίζουν proximities μεταξύ ζευγών των περιπτώσεων που μπορούν να χρησιμοποιηθούν για την ομαδοποίηση, εντοπίζοντας ακραίες τιμές, ή (κατά την κλιμάκωση) δίνουν ενδιαφέρουσες απόψεις των δεδομένων.
- Οι δυνατότητες των παραπάνω μπορούν να επεκταθούν σε μη ταξινομημένα δεδομένα, που οδηγούν σε μη επιβλεπόμενη ομαδοποίηση προβολών δεδομένων και ανίχνευση ακραίων τιμών.

### Τα αρνητικά των τυχαίων δασών είναι:

- Έχει παρατηρηθεί ότι κατηγοριοποιεί «εσφαλμένα» το σετ δεδομένων σε θορυβώδη classification/regression εργασίες.
- Για τα δεδομένα, συμπεριλαμβανομένων των κατηγορικών μεταβλητών με διαφορετικό αριθμό επιπέδων, το Random Forest μεροληπτεί υπέρ αυτών των ιδιοτήτων με περισσότερα επίπεδα. Συνεπώς ο αλγόριθμος δεν είναι αξιόπιστος για αυτά τα δεδομένα.
- Εάν τα δεδομένα περιέχουν ομάδες συσχετισμένων χαρακτηριστικών παρόμοιας σημασίας για την έξοδο, τότε οι μικρότερες ομάδες ευνοούνται από τις μεγαλύτερες ομάδες.

Υλοποίηση Random Forest σε περιβάλλον Matlab2011a στο σετ δεδομένων «fisheriris». Το πρόγραμμα χρησιμοποιεί την συνάρτηση treebagger (ρυθμίζοντας κάποιες παραμέτρους η συνάρτηση καλεί τον αλγόριθμο Random Forest).

```
b = TreeBagger(50, meas, species, 'NVarToSample', 2, 'OOBPred', 'on')
```

Με την συνάρτηση δημιουργούνται 50 δέντρα στο σετ δεδομένων. Έχοντας δώσει μια τιμή στην παράμετρο NVarToSample καλείται ο αλγόριθμος των Random forests.

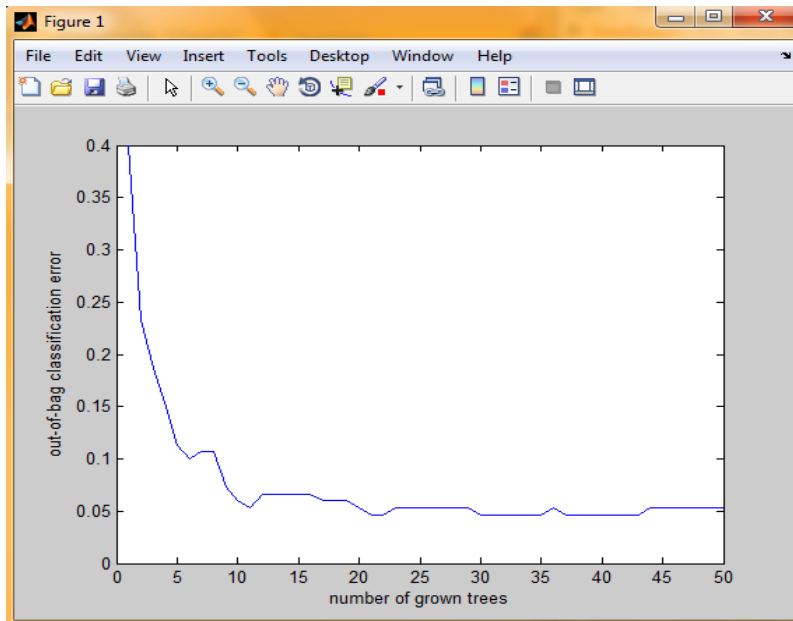
Ensemble with 50 bagged decision trees:

```
Training X: [150x4]
Training Y: [150x1]
Method: classification
Nvars: 4
NVarToSample: 2
MinLeaf: 1
FBoot: 1
SampleWithReplacement: 1
ComputeOOBPrediction: 1
ComputeOOBVarImp: 0
```



Proximity: []

ClassNames: 'setosa' 'versicolor' 'virginica'



Παρατηρούμε ότι όσα περισσότερα δέντρα έχουμε το σφάλμα oob μικραίνει όλο και περισσότερο.



## 2.6 Μηχανή Υποστήριξης Διανυσμάτων (Support Vector Machine - SVM)

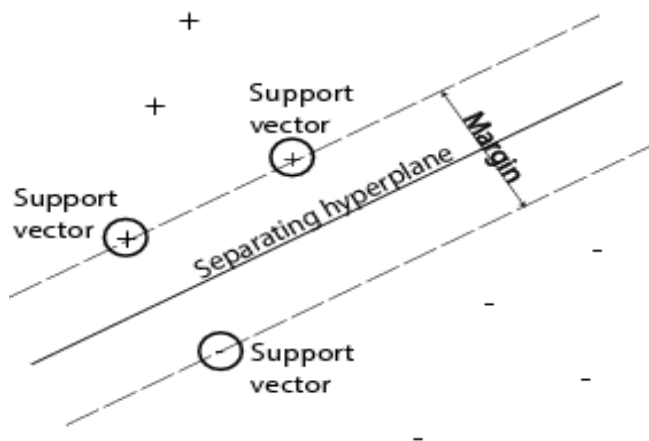
Κατηγορίες για την κατανόηση των SVM:

- Διαχωρίσιμα δεδομένα.
- Μη διαχωρίσιμα δεδομένα.
- Μη γραμμικές μετατροπές με Πυρήνες (Kernel).

### 2.6.1 Διαχωρίσιμα δεδομένα

Μπορεί να χρησιμοποιηθεί μια μηχανή υποστήριξης διανυσμάτων (SVM), όταν τα δεδομένα έχουν ακριβώς δύο κατηγορίες. Μια SVM ταξινομεί τα δεδομένα από την εύρεση του καλύτερου υπερεπίπεδου που διαχωρίζει όλα τα σημεία δεδομένων μιας κλάσης από εκείνες της άλλης κατηγορίας. Το καλύτερο υπερεπίπεδο για SVM σημαίνει εκείνο με το μεγαλύτερο περιθώριο μεταξύ των δύο κατηγοριών. Περιθώριο νοείται το μέγιστο πλάτος της πλάκας παράλληλα με το υπερεπίπεδο που δεν έχει εσωτερικά σημεία δεδομένων.

Τα διανύσματα υποστήριξης είναι τα σημεία δεδομένων που βρίσκονται πλησιέστερα στο διαχωριστικό υπερεπίπεδο. Αυτά τα σημεία είναι στο σύνορο της πλάκας. Η παρακάτω εικόνα απεικονίζει αυτούς τους ορισμούς, με + υποδεικνύοντας τα σημεία δεδομένων του τύπου 1, και - δείχνοντας σημεία δεδομένων του τύπου -1.



Μαθηματική διατύπωση: Primal. Τα δεδομένα για την εκπαίδευση είναι ένα σύνολο σημείων (διανύσματα)  $x_i$  μαζί με κατηγορίες  $y_i$ . Για κάποια διάσταση  $d$ , το  $x_i \in R^d$ , και το  $y_i = \pm 1$ . Η εξίσωση του υπερεπίπεδο είναι:

$$\langle w, x \rangle + b = 0,$$

Όπου το  $w \in R^d$ , και  $\langle w, x \rangle$  είναι το εσωτερικό γινόμενο του  $w$  και  $x$  και το  $b$  είναι ένας πραγματικός αριθμός.



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

Το παρακάτω πρόβλημα καθορίζει το καλύτερο διαχωρισμό υπερεπιπέδου. Εύρεση  $w$  και  $b$  που ελαχιστοποιούν  $\|w\|$  τέτοια ώστε για όλα τα σημεία δεδομένων  $(x_i, y_i)$ ,

$$y_i(\langle w, x_i \rangle + b) \geq 1.$$

Τα διανύσματα υποστήριξης είναι το  $x_i$  στο σύνορο, εκείνες για τις οποίες  $y_i(\langle w, x_i \rangle + b) = 1$ .

Για την μαθηματική ευκολία, το πρόβλημα συνήθως αναφέρεται ως ισοδύναμο πρόβλημα της ελαχιστοποίησης  $\langle w, w \rangle / 2$ . Αυτό είναι ένα πρόβλημα τετραγωνικού προγραμματισμού. Η βέλτιστη λύση  $w, b$  επιτρέπει την κατάταξη ενός διανύσματος  $Z$  ως εξής:

$$\text{class}(z) = \text{sign}(\langle w, z \rangle + b).$$

Μαθηματική διατύπωση: Dual. Είναι υπολογιστικά απλούστερο να λύσει το πρόβλημα διπλού τετραγωνικού προγραμματισμού. Για την απόκτηση του διπλού, λαμβάνει θετικούς πολλαπλασιαστές Lagrange  $\alpha_i$  πολλαπλασιαζόμενους με κάθε περιορισμό, και αφαιρεί από την αντικειμενική συνάρτηση:

$$L_P = \frac{1}{2} \langle w, w \rangle - \sum_i \alpha_i (y_i (\langle w, x_i \rangle + b) - 1),$$

όπου ψάχνεται για ένα σταθερό σημείο  $L_P$  πάνω  $w$  και  $b$  και με τη ρύθμιση της κλίσης του  $L_P$  σε 0, γίνεται:

$$\begin{aligned} w &= \sum_i \alpha_i y_i x_i \\ 0 &= \sum_i \alpha_i y_i \end{aligned} \quad (15-1)$$

Αντικαθιστώντας στην  $L_P$ , γίνεται το διπλό  $L_D$  σε:

$$L_D = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle,$$

όπου μεγιστοποιείται πάνω  $\alpha_i \geq 0$ . Σε γενικές γραμμές, πολλές  $\alpha_i$  είναι μηδέν. Τα μη μηδενικά  $\alpha_i$  στη λύση στο dual πρόβλημα ορίζουν το υπερεπίπεδο, όπως φαίνεται στην Εξίσωση 15-1, η οποία δίνει  $w$  ως το άθροισμα των  $\alpha_i y_i x_i$ . Τα σημεία δεδομένων  $x_i$  που αντιστοιχούν σε μη μηδενικά  $\alpha_i$  είναι τα διανύσματα υποστήριξης.

Η παράγωγος της  $L_D$  σε σχέση με ένα μη μηδενικό  $\alpha_i$  είναι μηδέν στο βέλτιστο. Αυτό δίνει:





$$y_i(\langle w, x_i \rangle + b) - 1 = 0.$$

Ειδικότερα, αυτό δίνει την τιμή του  $b$  στο λύση, με τη λήψη κάθε  $i$  με μη μηδενικό  $\alpha_i$

Το dual είναι ένα τυπικό τετραγωνικό πρόβλημα προγραμματισμού. Για παράδειγμα, το Optimization Toolbox™ quadprog solver λύνει αυτό το είδος του προβλήματος.

### 2.6.2 Μη διαχωρίσιμα δεδομένα

Τα δεδομένα είναι δυνατόν να μην επιτρέπουν για ένα διαχωριστικό υπερεπίπεδο. Σε αυτή την περίπτωση, το SVM μπορεί να χρησιμοποιήσει ένα μαλακό περιθώριο, που σημαίνει ένα υπερεπίπεδο να χωρίζει πολλά, αλλά όχι όλα τα σημεία δεδομένων.

Υπάρχουν δύο τυπικές συνθέσεις των μαλακών περιθωρίων. Και οι δύο περιλαμβάνουν την προσθήκη slack (χαλαρών) μεταβλητών  $s_i$  και μια «ποινή» παράμετρο  $C$ .

- Το  $L^1$ -norm πρόβλημα είναι:

$$\min_{w,b,s} \left( \frac{1}{2} \langle w, w \rangle + C \sum_i s_i \right)$$

Τέτοια ώστε:

$$y_i (\langle w, x_i \rangle + b) \geq 1 - s_i$$

$$s_i \geq 0.$$

Η  $L1$ -νόρμα αναφέρεται στη χρήση  $s_i$  ως slack (χαλαρών) μεταβλητών αντί των τετραγώνων τους. Η SMO svmtrain μέθοδος ελαχιστοποιεί το  $L1$  νόρμα πρόβλημα.

- Το  $L^2$ -norm πρόβλημα είναι:

$$\min_{w,b,s} \left( \frac{1}{2} \langle w, w \rangle + C \sum_i s_i^2 \right)$$

υπόκειται στους ίδιους περιορισμούς. Η QP svmtrain μέθοδος ελαχιστοποιεί το  $L2$ -νόρμα πρόβλημα.

Σε αυτές τις διαμορφώσεις, παρατηρείται ότι η αύξηση των  $C$  θέσεων βάζει περισσότερο βάρος στις slack μεταβλητές  $s_i$ , δηλαδή η βελτιστοποίηση προσπαθεί να κάνει ένα αυστηρότερο διαχωρισμό μεταξύ των τάξεων. Αντίστοιχα, η μείωση  $C$  προς το 0 κάνει εσφαλμένη ταξινόμηση λιγότερο σημαντική.

Μαθηματική διατύπωση: Dual. Για ευκολότερους υπολογισμούς, εξετάζεται το  $L1$  διπλό πρόβλημα σε αυτή τη διαμόρφωση χαλαρού περιθωρίου. Χρησιμοποιώντας



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
πολλαπλασιαστές Lagrange  $\mu_i$ , η συνάρτηση για την ελαχιστοποίηση του προβλήματος L1-κανόνας είναι:

$$L_P = \frac{1}{2} \langle w, w \rangle + C \sum_i s_i - \sum_i \alpha_i (y_i (\langle w, x_i \rangle + b) - (1 - s_i)) - \sum_i \mu_i s_i,$$

όπου αναζητείται για ένα σταθερό σημείο πάνω  $L_P$   $w$ ,  $b$ , και θετικό  $s_i$ . Η Ρύθμιση της κλίσης του LP σε 0, γίνεται:

$$\begin{aligned} b &= \sum_i \alpha_i y_i x_i \\ \sum_i \alpha_i y_i &= 0 \\ \alpha_i &= C - \mu_i \\ \alpha_i, \mu_i, s_i &\geq 0. \end{aligned}$$

Αυτές οι εξισώσεις οδηγούν απευθείας στην διπλή διατύπωση:

$$\max_{\alpha} \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle$$

υπόκεινται στους περιορισμούς:

$$\begin{aligned} \sum_i y_i \alpha_i &= 0 \\ 0 &\leq \alpha_i \leq C. \end{aligned}$$

Το τελικό σύνολο των ανισοτήτων,  $0 \leq \alpha_i \leq C$ , παρουσιάζει γιατί το  $C$  μερικές φορές ονομάζεται κουτί περιορισμού. Το  $C$  διατηρεί τις επιτρεπόμενες τιμές των πολλαπλασιαστών Lagrange  $\alpha_i$  σε ένα "κουτί", μια περιοχή που οριοθετείται.

Η εξίσωση για τη βαθμίδα  $b$  δίνει την λύση  $b$  από την άποψη του συνόλου των μη μηδενικών  $\alpha_i$ , που αντιστοιχούν στα διανύσματα στήριξης.

Μπορεί να γραφθεί και να λυθεί το διπλό του L2-νόρμα πρόβλημα με έναν ανάλογο τρόπο.

**svmtrain εφαρμογή.** Και τα δύο διπλά χαλαρά περιθώρια - προβλήματα είναι τετραγωνικά προβλήματα προγραμματισμού. Εσωτερικά, η svmtrain έχει πολλούς διαφορετικούς αλγόριθμους για την επίλυση των προβλημάτων. Ο προεπιλεγμένος Sequential Minimal Optimization (SMO) αλγόριθμος ελαχιστοποιεί το ένα-νόρμα πρόβλημα. Ο SMO είναι ένας σχετικά γρήγορος αλγόριθμος. Η εργαλειοθήκη Optimization Toolbox, μπορεί να επιλεγθεί για την χρήση του quadprog ως κατάλληλος αλγόριθμος. που ελαχιστοποιεί το L2-νόρμα πρόβλημα. επίσης χρησιμοποιεί αρκετή πολύ μνήμη, αλλά λύνει τετραγωνικά προγράμματα σε υψηλό βαθμό ακρίβειας.



### 2.6.3 Μη γραμμικές μετατροπές με Πυρήνες (Kernel).

Κάποια δυαδικά προβλήματα ταξινόμησης δεν έχουν ένα απλό υπερεπίπεδο ως ένα χρήσιμο κριτήριο διαχωρισμού. Για αυτά τα προβλήματα, υπάρχει μια παραλλαγή της μαθηματικής προσέγγισης που διατηρεί σχεδόν όλη την απλότητα ενός SVM διαχωρισμού υπερεπιπέδων.

Η προσέγγιση αυτή χρησιμοποιεί τα αποτελέσματα από τη θεωρία της αναπαραγωγής πυρήνων:

Υπάρχει μια κατηγορία λειτουργιών  $K(x, y)$  με την ακόλουθη ιδιότητα. Υπάρχει ένας γραμμικός χώρος  $S$  και μια συνάρτηση  $\varphi$  αντιστοίχισης του  $x$  στο  $S$  τέτοια ώστε:

$$K(x, y) = \langle \varphi(x), \varphi(y) \rangle.$$

Το εσωτερικό γινόμενο λαμβάνει χώρα στο χώρο  $S$ .

Αυτή η class των συναρτήσεων περιλαμβάνει:

- Πολυώνυμα (Polynomials): Για κάποιο θετικό ακέραιο  $d$ ,  
 $K(x, y) = (1 + \langle x, y \rangle)^d$ .
- Ακτινικών συναρτήσεων (Radial Basis): Για κάποιο θετικό αριθμό  $\sigma$ ,  
 $K(x, y) = \exp(-\langle (x-y), (x-y) \rangle / (2\sigma^2))$ .
- Πολυστρωματικό perceptron (νευρωνικό δίκτυο): Για ένα θετικό αριθμό  $p_1$  και ένα αρνητικό αριθμό  $p_2$ ,  
 $K(x, y) = \tanh(p_1 \langle x, y \rangle + p_2)$ .

Η μαθηματική προσέγγιση που χρησιμοποιεί πυρήνες εξαρτάται από την υπολογιστική μέθοδο των υπερεπιπέδων. Όλοι οι υπολογισμοί για την ταξινόμηση υπερεπιπέδων χρησιμοποιούν τίποτα περισσότερο από ό, τι τα εσωτερικά γινόμενα. Ως εκ τούτου, οι μη γραμμικές πυρήνες μπορούν να χρησιμοποιήσουν ίδιους υπολογισμούς και αλγόριθμους λύσεων, και να λάβουν ταξινομητές που είναι μη γραμμικές. Οι προκύπτουσες ταξινομητές είναι υπερεπιφάνειες σε κάποιο χώρο  $S$ , αλλά ο  $S$  χώρος δεν πρέπει να προσδιοριστεί ή να εξεταστεί.



### 3. Πρόβλεψη Ηλεκτρικού Φορτίου (Load Forecasting)

#### 3.1 Εισαγωγή

Ακριβή μοντέλα πρόβλεψης για την ισχύ ηλεκτρικού φορτίου είναι απαραίτητα για την λειτουργία και τον προγραμματισμό μιας εταιρείας κοινής ωφελείας. Η πρόβλεψη φορτίου βοηθά μια ηλεκτρική εταιρεία στη λήψη σημαντικών αποφάσεων, συμπεριλαμβανομένων των αποφάσεων για την αγορά και την παραγωγή ηλεκτρικής ενέργειας, μεταγωγή φορτίου, και τις υποδομές ανάπτυξης. Οι προβλέψεις φορτίου είναι εξαιρετικά σημαντικές για τους προμηθευτές ενέργειας, τα ISOs, τα χρηματοπιστωτικά ιδρύματα, και άλλους συμμετέχοντες στην παραγωγή, στη μεταφορά, στη διανομή, και στις αγορές ηλεκτρικής ενέργειας.

Προβλέψεις φορτίου μπορούν να χωριστούν σε τρεις κατηγορίες: τις βραχυπρόθεσμες προβλέψεις οι οποίες είναι συνήθως από μία ώρα έως μία εβδομάδα, μεσοπρόθεσμες προβλέψεις που είναι συνήθως από μία εβδομάδα έως ένα έτος, και μακροπρόθεσμες προβλέψεις που είναι περισσότερο από ένα έτος. Οι προβλέψεις για διαφορετικούς χρονικούς ορίζοντες είναι σημαντικές για διαφορετικές λειτουργίες μέσα σε μια εταιρεία κοινής ωφελείας. Οι φύσεις των προβλέψεων αυτών είναι διαφορετική επίσης. Για παράδειγμα, για μια συγκεκριμένη περιοχή, είναι δυνατόν να προβλέψει το επόμενο φορτίο ημέρας με ακρίβεια περίπου 1-3%. Ωστόσο, είναι αδύνατον να προβλεφθεί το επόμενο έτος αιχμής φορτίου με την ανάλογη ακρίβεια από ακριβείς μακροπρόθεσμες καιρικές προβλέψεις που δεν είναι διαθέσιμες. Για τις επόμενες προβλέψεις κορυφής στο χρόνο, είναι δυνατόν να δώσει την κατανομή πιθανότητας του φορτίου με βάση τα ιστορικά των καιρικών παρατηρήσεων. Είναι επίσης δυνατό, σύμφωνα με την πρακτική του κλάδου, να προβλέψει το λεγόμενο κανονικοποιημένο καιρικό φορτίο, το οποίο θα λάβει θέση για μέση ετήσια καιρικής αιχμής ή χειρότερη από το μέσο όρο αιχμής καιρικών συνθηκών για μια συγκεκριμένη περιοχή. Το κανονικοποιημένο φορτίο καιρού είναι το φορτίο που υπολογίστηκε για τις λεγόμενες κανονικές καιρικές συνθήκες οι οποίες είναι ο μέσος όρος των χαρακτηριστικών καιρού για ιστορικά φορτία αιχμής πάνω από ένα ορισμένο χρονικό διάστημα. Η διάρκεια της περιόδου αυτής ποικίλλει από ένα βοηθητικό πρόγραμμα στο άλλο. Οι περισσότερες εταιρείες λαμβάνουν δεδομένα τα τελευταία 25-30 χρόνια των καιρικών φαινομένων.

Η πρόβλεψη φορτίου ήταν πάντα σημαντική για την απόφαση σχεδιασμού και λειτουργίας που διεξάγεται από επιχειρήσεις κοινής ωφελείας. Ωστόσο, με την απορρύθμιση των ενεργειακών βιομηχανιών, η πρόβλεψη φορτίου είναι ακόμη πιο σημαντική. Με την προσφορά και τη ζήτηση διακυμάνσεων και οι αλλαγές των καιρικών συνθηκών και οι τιμές της ενέργειας αυξήθηκαν κατά ένα συντελεστή δέκα ή περισσότερο κατά τη διάρκεια των καταστάσεων αιχμής, η πρόβλεψη φορτίου είναι ζωτικής σημασίας για υπηρεσίες κοινής ωφέλειας. Η Βραχυπρόθεσμη πρόβλεψη φορτίου μπορεί να βοηθήσει στην εκτίμηση των ροών φορτίου και για τη λήψη αποφάσεων που μπορεί να αποτρέψει την υπερφόρτωση. Έγκαιρες υλοποιήσεις



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

τέτοιων αποφάσεων οδηγούν στη βελτίωση της αξιοπιστίας του δικτύου και στις μειωμένες εμφανίσεις στις βλάβες του εξοπλισμού και διακοπές ρεύματος. Πρόβλεψη φορτίου είναι επίσης σημαντική για τις αξιολογήσεις των συμβάσεων και των αξιολογήσεων των διαφόρων περίπλοκων χρηματοπιστωτικών προϊόντων για την τιμολόγηση της ενέργειας που προσφέρει η αγορά. Στην απελευθερωμένη οικονομία, οι αποφάσεις για κεφαλαιουχικές δαπάνες με βάση μακροπρόθεσμης πρόβλεψης είναι επίσης πιο σημαντικές από ό, τι στη μη απελευθερωμένη- οικονομία, όταν αυξάνεται το ποσοστό θα μπορούσε να δικαιολογηθεί από κεφαλαιουχικές δαπάνες έργων.

Οι περισσότεροι μέθοδοι πρόβλεψης χρησιμοποιούν στατιστικές τεχνικές ή αλγορίθμους τεχνητής νοημοσύνης, όπως η παλινδρόμηση, νευρωνικά δίκτυα, ασαφής λογική, και έμπειρα συστήματα. Δύο από τις μεθόδους, η λεγόμενη τελική χρήση και οικονομετρική προσέγγιση χρησιμοποιούνται ευρέως για τις μεσοπρόθεσμες και μακροπρόθεσμες πρόβλεψεις. Μία ποικιλία μεθόδων, οι οποίες περιλαμβάνουν την λεγόμενη παρόμοια προσέγγιση ημέρας, διάφορα μοντέλα παλινδρόμησης, χρονολογικές σειρές, νευρωνικά δίκτυα, στατιστικούς αλγόριθμους μάθησης, ασαφή λογική, και τα έμπειρα συστήματα, έχουν αναπτυχθεί για βραχυπρόθεσμη πρόβλεψη.

Μια μεγάλη ποικιλία από μαθηματικές μεθόδους και ιδέες έχουν χρησιμοποιηθεί για πρόβλεψη του φορτίου. Η ανάπτυξη και η βελτίωση των κατάλληλων μαθηματικών εργαλείων, θα οδηγήσουν στην ανάπτυξη πιο ακριβών τεχνικών πρόβλεψης φορτίου. Η ακρίβεια της πρόβλεψης φορτίου δεν εξαρτάται μόνο από τις τεχνικές, αλλά και για την ακρίβεια των προβλεπόμενων καιρικών σεναρίων. Πρόβλεψη καιρού είναι ένα σημαντικό θέμα το οποίο είναι έξω από το πεδίο εφαρμογής του παρόντος κεφαλαίου. Εδώ αναφέρεται η σημαντική πρόοδος όσον αφορά την ανάπτυξη των πληροφορικών συστημάτων καιρικής πρόβλεψης συμπεριλαμβανομένου του μοντέλου μέσης κλίμακας MM5 που αναπτυχθεί και υποστηρίζεται από μια κοινοπραξία των πανεπιστημίων.

### 3.2 Σημαντικοί παράγοντες για προβλέψεις

Για βραχυπρόθεσμη πρόβλεψη φορτίου διάφοροι παράγοντες πρέπει να εξεταστούν, όπως οι παράγοντες του χρόνου, πληροφορίες για τον καιρό, και τις πιθανές κατηγορίες πελατών. Οι μεσοπρόθεσμες και μακροπρόθεσμες προβλέψεις λαμβάνουν υπόψη το ιστορικό φορτίο και πληροφορίες για τον καιρό, το πλήθος των πελατών σε διάφορες κατηγορίες, τις συσκευές στην περιοχή και τα χαρακτηριστικά τους, όπως η ηλικία, τα οικονομικά και δημογραφικά δεδομένα και τις προβλέψεις τους, οι πωλήσεις συσκευών δεδομένων, και άλλους παράγοντες.

Οι παράγοντες του χρόνου περιλαμβάνουν την εποχή του χρόνου, την ημέρα της εβδομάδας, και η ώρα της ημέρας. Υπάρχουν σημαντικές διαφορές στο φορτίο μεταξύ των καθημερινών και Σαββατοκύριακων. Το φορτίο σε διαφορετικές ημέρες της εβδομάδας, επίσης, μπορεί να συμπεριφέρεται διαφορετικά. Για παράδειγμα Δευτέρες και Παρασκευές είναι δίπλα στα σαββατοκύριακα, μπορεί να έχουν δομικά



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

διαφορετικά φορτία από Τρίτη έως Πέμπτη. Αυτό ισχύει ιδιαίτερα κατά τη διάρκεια του καλοκαιριού. Οι Αργίες είναι πιο δύσκολο να προβλεφθούν από τις μη-διακοπές λόγω της σχετικής τους σπάνια εμφάνιση.

Οι καιρικές συνθήκες επηρεάζουν το φορτίο. Στην πραγματικότητα, προβλεπόμενοι καιρικοί παράμετροι είναι οι πιο σημαντικοί παράγοντες σε βραχυπρόθεσμες προβλέψεις φορτίου. Διάφορες καιρικές μεταβλητές θα μπορούσαν να θεωρηθούν για την πρόβλεψη στο φορτίο. Θερμοκρασία και η υγρασία είναι οι πλέον κοινώς χρησιμοποιούμενοι προγνωστικοί παράγοντες φορτίου. μια έρευνα πρόβλεψης ηλεκτρικού φορτίου που δημοσιεύθηκε έδειξε ότι από τις 22 εκθέσεις ερευνών που έχουν θεωρηθεί, 13 έκαναν χρήση μόνο της θερμοκρασίας, 3 έκαναν χρήση της θερμοκρασίας και της υγρασίας, 3 χρησιμοποίησαν πρόσθετες παραμέτρους καιρικών συνθηκών, και 3 χρησιμοποίησαν μόνο παραμέτρους φορτίου.

Μεταξύ των καιρικών μεταβλητών που αναφέρονται παραπάνω, δύο σύνθετες καιρικές λειτουργικές μεταβλητές, ο THI (δείκτης θερμοκρασίας-υγρασίας) και WCI (δείκτης ανέμου ψύξης), χρησιμοποιούνται ευρέως από εταιρείες κοινής ωφελείας. THI είναι ένα μέτρο δυσφορίας καλοκαιρινής θερμότητας και ομοίως WCI είναι ψυχρό στρες το χειμώνα.

Οι περισσότερες εταιρείες παραγωγής ηλεκτρικής ενέργειας εξυπηρετούν τους διαφορετικούς τύπους πελατών τους, όπως κατοικίες, εμπορικές και βιομηχανικές περιοχές. Το πρότυπο ηλεκτρικής χρήσης είναι διαφορετικό για τους πελάτες που ανήκουν σε διαφορετικές κατηγορίες, αλλά είναι κάπως όμοιο για τους πελάτες μέσα σε κάθε κατηγορία. Ως εκ τούτου, οι περισσότερες επιχειρήσεις κοινής ωφελείας διακρίνουν τη συμπεριφορά του φορτίου σε ταξική βάση ανά τάξη.

### 3.3 Μέθοδοι Πρόβλεψης

Κατά τη διάρκεια των τελευταίων δεκαετιών, μια σειρά από μεθόδους πρόβλεψης έχουν αναπτυχθεί. Δύο από τις μεθόδους, η τελική χρήση και η οικονομετρική προσέγγιση είναι ευρέως χρησιμοποιούμενες για την μεσοπρόθεσμη και μακροπρόθεσμη πρόβλεψη. Μία ποικιλία μεθόδων, οι οποίες περιλαμβάνουν την λεγόμενη παρόμοια προσέγγιση ημέρας, διάφορα μοντέλα παλινδρόμησης, χρονολογικές σειρές, νευρωνικά δίκτυα, έμπειρα συστήματα, ασαφούς λογικής και στατιστικών αλγορίθμων μάθησης, χρησιμοποιούνται για τη βραχυπρόθεσμη πρόβλεψη. Η ανάπτυξη, βελτίωση, και η έρευνα των κατάλληλων μαθηματικών εργαλείων, θα οδηγήσουν στην ανάπτυξη πιο ακριβών τεχνικών πρόβλεψης φορτίου.

Στατιστικές προσεγγίσεις συνήθως απαιτούν ένα μαθηματικό μοντέλο που αντιπροσωπεύει το φορτίο ως συνάρτηση διαφόρων παραγόντων όπως είναι ο χρόνος, ο καιρός και η κατηγορία πελατών. Οι δύο σημαντικές κατηγορίες τέτοιων μαθηματικών μοντέλων είναι: μοντέλα προσθετικών και μοντέλα πολλαπλασιαστικών. Διαφέρουν σε αν το φορτίο πρόβλεψης είναι το άθροισμα (προσθετικό) ενός αριθμού συστατικών ή το προϊόν (πολλαπλασιαστικός) από έναν αριθμό παραγόντων.



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

Για παράδειγμα, ο Chen et al. παρουσίασαν ένα προσθετικό μοντέλο που θα λαμβάνει τη μορφή πρόβλεψης φορτίου ως συνάρτηση των τεσσάρων συνιστωσών:

$$L = L_n + L_w + L_s + L_r$$

όπου  $L$  είναι το συνολικό φορτίο,  $L_n$  αντιπροσωπεύει το «κανονικό» μέρος του φορτίου, το οποίο είναι ένα σύνολο τυποποιημένων σχημάτων φορτίου για κάθε "τύπο" της ημέρας που έχει ταχτοποιηθεί,  $L_w$  αντιπροσωπεύει το καιρικό ευαίσθητο μέρος του φορτίου,  $L_s$  είναι ένα ειδικό εξάρτημα - γεγονός που δημιουργεί μια ουσιαστική απόκλιση από το σύνηθες πρότυπο φορτίο, και  $L_r$  που είναι εντελώς τυχαίος όρος, ο θόρυβος.

Chen et al. πρότεινε επίσης τη τιμολόγηση της ηλεκτρικής ενέργειας ως πρόσθετο όρο ο οποίος μπορεί να συμπεριληφθεί στο μοντέλο. Φυσικά, η τιμή μειώνεται / αυξάνεται και επηρεάζει την κατανάλωση ηλεκτρικής ενέργειας. Μεγάλο κόστος ευαίσθητων βιομηχανικών και θεσμικών φορτίων μπορούν να έχουν σημαντική επίδραση στα φορτία.. Οι συγγραφείς εκθέτουν ακριβείς εκτιμήσεις ότι επιτεύχθηκαν πιο γρήγορα με τη συμπερίληψη των δεδομένων τιμών. Ένα πολλαπλασιαστικό μοντέλο μπορεί να είναι της μορφής:

$$L = L_n \cdot F_w \cdot F_s \cdot F_r$$

όπου  $L_n$  είναι το κανονικό (βάση) φορτίο και οι συντελεστές διορθώσεως  $F_w$ ,  $F_s$ , και  $F_r$  είναι θετικοί αριθμοί που μπορούν να αυξήσουν ή να μειώσουν το συνολικό φορτίο. Αυτές οι διορθώσεις βασίζονται στον τρέχων καιρό ( $F_w$ ), στα ειδικά γεγονότα ( $F_s$ ), και στην τυχαία διακύμανση ( $F_r$ ). Τέτοιοι παράγοντες όπως η τιμολόγηση ηλεκτρικής ενέργειας ( $F_p$ ) και η αύξηση φορτίου ( $F_g$ ) μπορούν επίσης να συμπεριληφθούν. Ο Rahman παρουσίασε μια βασισμένη σε νόμους πρόβλεψη χρησιμοποιώντας ένα πολλαπλασιαστικό μοντέλο. Οι Μεταβλητές καιρικών συνθηκών και του βασικού φορτίου συνδέονται με τις καιρικές μετρήσεις που είχαν συμπεριληφθεί στο μοντέλο.

### 3.4 Μέθοδοι Μεσαίας-και μακροπρόθεσμης πρόβλεψης φορτίου

Η μοντελοποίηση τελικής χρήσης, μοντελοποίηση οικονομικής, και οι συνδυασμοί τους είναι οι πιο συχνά χρησιμοποιούμενες μέθοδοι για τη μεσοπρόθεσμη και μακροπρόθεσμη πρόβλεψη φορτίου. Περιγραφές των συσκευών που χρησιμοποιούνται από τους πελάτες, τα μεγέθη των σπιτιών, η ηλικία του εξοπλισμού, οι αλλαγές της τεχνολογίας, η συμπεριφορά των πελατών, και η δυναμική των πληθυσμών συνήθως περιλαμβάνονται στις στατιστικές και στα μοντέλα προσομοίωσης με βάση την λεγόμενη προσέγγιση τελικής χρήσης. Επιπλέον, οικονομικοί παράγοντες, όπως το κατά κεφαλήν εισόδημα, τα επίπεδα απασχόλησης, και οι τιμές της ηλεκτρικής ενέργειας περιλαμβάνονται στα οικονομικά μοντέλα. Αυτά τα μοντέλα συχνά χρησιμοποιούνται σε συνδυασμό με την προσέγγιση τελικής χρήσης. Μακροπρόθεσμες προβλέψεις περιλαμβάνουν τις προβλέψεις σχετικά με τις



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου αλλαγές του πληθυσμού, της οικονομικής ανάπτυξης, της βιομηχανικής κατασκευής, και της τεχνολογικής ανάπτυξης.

Μοντέλο Τελικής Χρήσης. Η προσέγγιση τελικής χρήσης υπολογίζει άμεσα την κατανάλωση ενέργειας με τη χρήση εκτενών πληροφοριών σχετικά με την τελική χρήση και των τελικών χρηστών, εν λόγω όπως συσκευές, η χρήση των πελατών, η ηλικία τους, τα μεγέθη των σπιτιών, και ούτω καθεξής. Οι στατιστικές πληροφορίες σχετικά με τους πελάτες, μαζί με τη δυναμική της αλλαγής είναι η βάση για την πρόβλεψη.

Μοντέλα τελικής χρήσης επικεντρώνονται στις διάφορες χρήσεις του ηλεκτρισμού στον οικιακό, εμπορικό, και βιομηχανικό τομέα. Αυτά τα μοντέλα βασίζονται στην αρχή ότι η ζήτηση ηλεκτρικής ενέργειας προέρχεται από τη ζήτηση των πελατών για φως, ψύξη, θέρμανση, ψύξη, κλπ. Έτσι μοντέλα τελικής χρήσης εξηγούν ζήτηση ενέργειας ως συνάρτηση του αριθμού των συσκευών στην αγορά.

Ιδανικά η προσέγγιση αυτή είναι πολύ ακριβής. Ωστόσο, είναι ευαίσθητη στην ποσότητα και στη ποιότητα της τελικής χρήσης των δεδομένων. Για παράδειγμα, σε αυτή τη μέθοδο η κατανομή της ηλικίας εξοπλισμού είναι σημαντική για ιδιαίτερους τύπους συσκευών. Η πρόβλεψη τελικής χρήσης απαιτεί λιγότερα ιστορικά δεδομένα, αλλά περισσότερες πληροφορίες σχετικά με τους πελάτες και τον εξοπλισμό τους.

Οικονομικό μοντέλο. Η οικονομική προσέγγιση συνδυάζει την οικονομική θεωρία και στατιστικές τεχνικές για την πρόβλεψη της ζήτησης ηλεκτρικής ενέργειας. Η προσέγγιση υπολογίζει τις σχέσεις μεταξύ της κατανάλωσης ενέργειας (εξαρτημένες μεταβλητές) και των παραγόντων που επηρεάζουν την κατανάλωση. Οι σχέσεις έχουν εκτιμηθεί από την ελαχίστων-τετραγώνων μέθοδο ή την μέθοδο χρονοσειρών.

Μία από τις επιλογές σε αυτό το πλαίσιο είναι να συγκεντρώνουν την οικονομική προσέγγιση, όταν η κατανάλωση σε διάφορους τομείς (οικιακό, εμπορικό, βιομηχανικό, κλπ) υπολογίζεται ως συνάρτηση των καιρικών συνθηκών, οικονομικών και άλλων μεταβλητών, και στη συνέχεια οι εκτιμήσεις συναρμολογούνται με τη χρήση πρόσφατων ιστορικών δεδομένων. Η ένταξη των οικονομικών προσεγγίσεων στην προσέγγιση τελικής χρήσης εισάγει στοιχεία συμπεριφοράς στις εξισώσεις τελικής χρήσης.

Μοντέλο Στατιστικής Εκμάθησης. Η τελική χρήση και η οικονομική μέθοδοι απαιτούν ένα μεγάλο ποσό των πληροφοριών σχετικά με τις συσκευές, πελάτες, οικονομία, κλπ. Η εφαρμογή τους είναι περίπλοκη και απαιτεί ανθρώπινη συμμετοχή. Επιπλέον, οι πληροφορίες αυτές δεν είναι συχνά διαθέσιμες σχετικά με συγκεκριμένους πελάτες και κρατούν ένα βοηθητικό πρόγραμμα και υποστηρίζουν το προφίλ ενός «μέσου» πελάτη ή το μέσο όρο πελατών για διαφορετικό τύπο πελάτη. Το πρόβλημα προκύπτει όταν το βοηθητικό πρόγραμμα θέλει να διεξάγει προβλέψεις για το επόμενο έτος για τις υπο-περιοχές, οι οποίες συχνά αποκαλούνται θύλακες φορτίου. Σε αυτή την περίπτωση, η ποσότητα της εργασίας που πρέπει να εκτελεστεί αυξάνεται αναλογικά με τον αριθμό των θυλάκων φορτίου. Επιπλέον, το προφίλ τελικής χρήσης και οικονομικά δεδομένα για διαφορετικές τσέπες φορτίου είναι συνήθως διαφορετικές. Τα χαρακτηριστικά για συγκεκριμένες περιοχές μπορούν να





Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

είναι διαφορετικά από το μέσο χαρακτηριστικών για την χρησιμότητα και μπορεί να μην είναι διαθέσιμα.

Για να απλουστευθούν οι μεσοπρόθεσμες προβλέψεις, να γίνουν πιο ακριβείς, και να αποφευχθεί η χρήση των μη διαθέσιμων πληροφοριών, ο Feinberg ανέπτυξε ένα στατιστικό μοντέλο που μαθαίνει τις παραμέτρους του μοντέλου φορτίου από τα ιστορικά δεδομένα. Ο Feinberg μελέτησε σύνολα δεδομένων φορτίου που παρέχονται από την επιχείρηση κοινής ωφέλειας στις βορειοανατολικές ΗΠΑ. Η εστίαση της μελέτης ήταν τα δεδομένα του καλοκαιριού. Έγινε σύγκριση με διάφορα μοντέλα φορτίου και κατέληξαν στο συμπέρασμα ότι το ακόλουθο πολλαπλασιαστικό μοντέλο είναι το πιο ακριβές.

$$L(t) = F(d(t), h(t)) \cdot f(w(t)) + R(t),$$

όπου  $L(t)$  είναι το πραγματικό φορτίο σε χρόνο  $t$ ,  $d(t)$  είναι η ημέρας της εβδομάδας,  $h(t)$  είναι η ώρα της ημέρας,  $F(d, h)$  είναι η ημερήσια και το ωριαίο συστατικό,  $w(t)$  είναι τα καιρικά δεδομένα που περιλαμβάνουν την θερμοκρασία και την υγρασία,  $f(w)$  είναι ο καιρικός παράγοντας και το  $R(t)$  είναι ένα τυχαίο σφάλμα.

Στην πραγματικότητα,  $w(t)$  είναι ένα διάνυσμα που αποτελείται από το ρεύμα και την καιρική χρονική υστέρηση μεταβλητών. Αυτό αντανακλά το γεγονός ότι το ηλεκτρικό φορτίο δεν εξαρτάται μόνο από τις τρέχουσες καιρικές συνθήκες, αλλά και με τον καιρό κατά τη διάρκεια των προηγούμενων ωρών και ημερών. Ειδικότερα, η γνωστή επίδραση του λεγόμενου θερμικού κύματος είναι ότι η χρήση των κλιματιστικών αυξάνεται όταν ο ζεστός καιρός συνεχίζεται για αρκετές ημέρες.

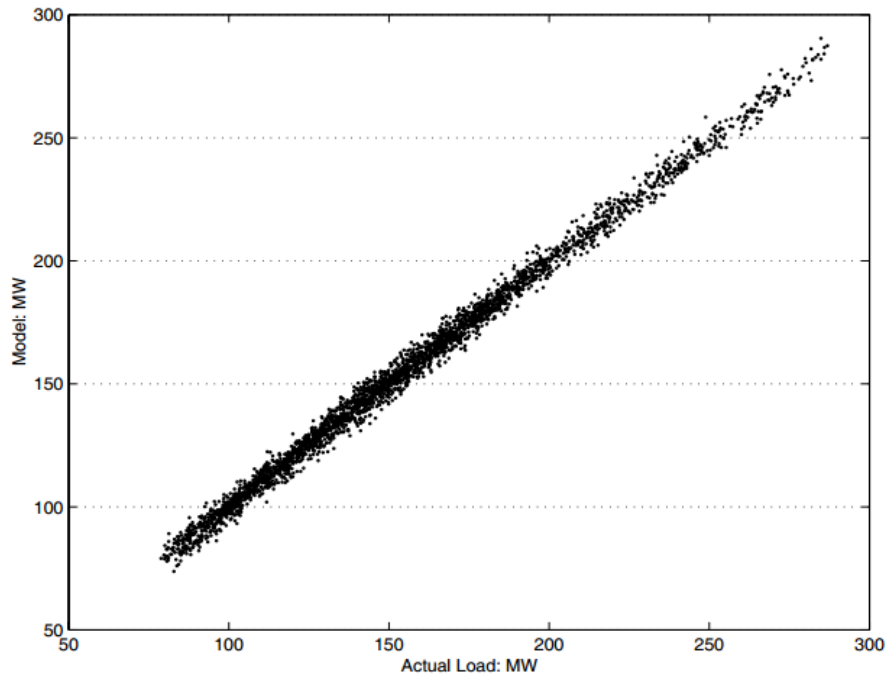
Για την εκτίμηση του καιρικού συντελεστή  $f(w)$ , χρησιμοποιείται το μοντέλο παλινδρόμησης:

$$f(w) = \beta_0 + \sum \beta_j X_j,$$

όπου  $X_j$  είναι οι επεξηγηματικές μεταβλητές οι οποίες είναι μη γραμμικές λειτουργίες των τρεχουσών και παλαιότερων καιρικών συνθηκών και παραμέτρους  $\beta_0, \beta_j$  είναι οι συντελεστές παλινδρόμησης.

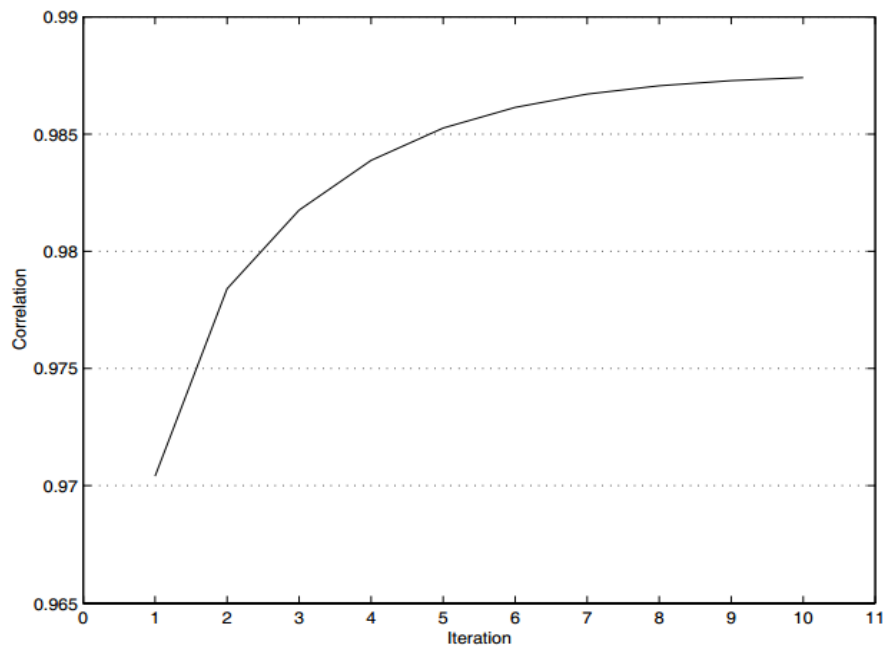
Οι παράμετροι του μοντέλου μπορούν να υπολογιστούν επαναληπτικά. Ξεκινάμε με  $F = 1$ . Στη συνέχεια, χρησιμοποιείται το παραπάνω μοντέλο παλινδρόμησης για την εκτίμηση  $f$ . Τότε εκτιμάται  $F$ , και ούτω καθεξής.

Ο αλγόριθμος που περιγράφηκε κατέδειξε την ταχεία σύγκλιση σε ιστορικό ωριαίο φορτίο και καιρικά δεδομένα. Έχει εφαρμοστεί σε πολλούς τομείς με πληθυσμό μεταξύ 50.000 και 250.000 πελατών. Σχήμα 12.1 παρουσιάζει ένα παράδειγμα ενός διαγράμματος διασποράς που συγκρίνει το μοντέλο και τις πραγματικές παραμέτρους. Σχήμα 12.2 καταδεικνύει την σύγκλιση της συσχέτισης μεταξύ του πραγματικού φορτίου και το υπόδειγμα για τη διαδικασία επανάληψης. Σχήμα 12.3 καταδεικνύει τη σύγκλιση των γραμμικών διαδικασιών παλινδρόμησης στον αλγόριθμο.

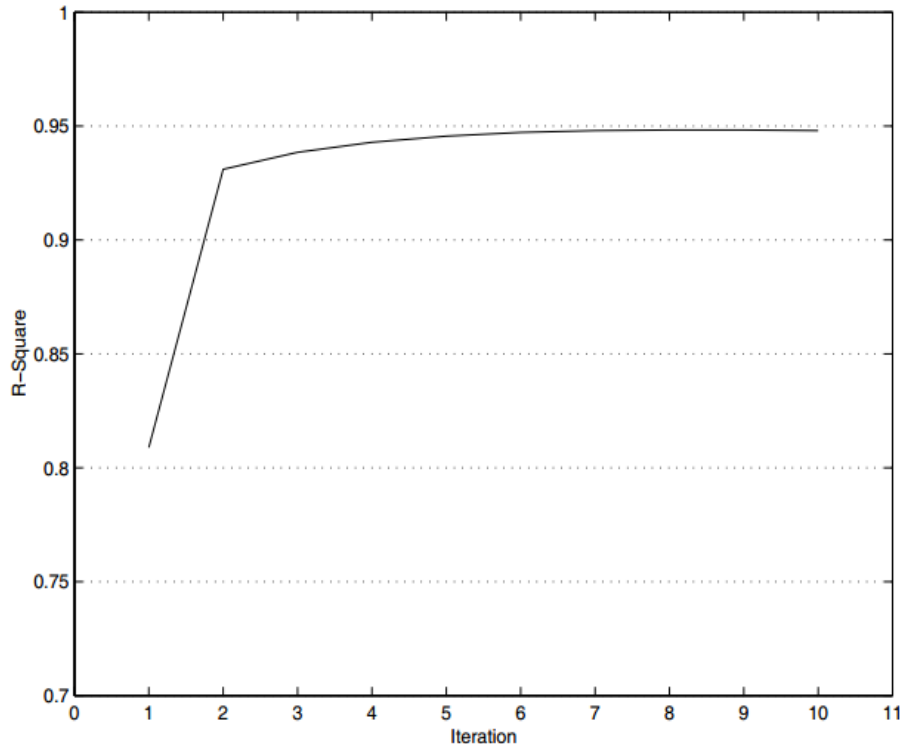


Σχήμα 12.1. Διάγραμμα διασποράς του πραγματικού φορτίου εναντίον του μοντέλου.

Το λογισμικό, που χρησιμοποιεί τη μέθοδο που περιγράφηκε, μαθαίνει στο μοντέλο παραμέτρους και κάνει προβλέψεις της επόμενης χρονιάς με βάση το μοντέλο φορτίων για τα τελευταία 25-30 χρόνια των δεδομένων. Αν και τα ιστορικά φορτία μπορούν να μην είναι διαθέσιμα, το λογισμικό εφαρμόζει τα τελευταία χρονικά Μοντέλα για τα ιστορικά καιρικά δεδομένα για την εκτίμηση διανομής κορυφής του επόμενου έτους.



Σχήμα 12.2. Συσχέτιση μεταξύ του πραγματικού φορτίου και του μοντέλου.



Σχήμα 12.3. Σύγκλιση του  $R^2$  για το πραγματικό φορτίο έναντι του μοντέλου.

Το λογισμικό παράγει πολλά σημαντικά χαρακτηριστικά. Για παράδειγμα, για κάθε θύλακα φορτίου και για το σύστημα, υπολογίζει ένα παράγοντα κανονικοποίησης καιρού που είναι ο λόγος του φορτίου αιχμής στο φορτίο που θα παρατηρούνταν υπό συνθήκες μέσης κορυφής. Παράγει επίσης κατανομές πιθανοτήτων κορυφών για τα επόμενα χρόνια. Οι περιγραφόμενες μέθοδοι μπορούν να εφαρμοστούν και σε μεσοπρόθεσμες και σε μακροπρόθεσμες προβλέψεις. Ωστόσο, οι μακροπρόθεσμες προβλέψεις θα πρέπει να ενσωματώνουν τις οικονομικές και πληθυσμιακές δυναμικές προβλέψεις ως παραμέτρους εισόδου.

### 3.5 Μέθοδοι Βραχυπρόθεσμης πρόβλεψης φορτίου

Μια μεγάλη ποικιλία των στατιστικών και τεχνικών τεχνητής νοημοσύνης έχουν αναπτυχθεί για τη βραχυπρόθεσμη πρόβλεψη φορτίου.

Προσέγγιση Παρόμοιας ημέρας. Η προσέγγιση αυτή βασίζεται στην αναζήτηση ιστορικών δεδομένων για ημέρες μέσα σε ένα, δύο ή τρία έτη, με παρόμοια χαρακτηριστικά στην πρόβλεψη της ημέρας. Παρόμοια χαρακτηριστικά περιλαμβάνουν τον καιρό, την ημέρα της εβδομάδα, και την ημερομηνία. Το φορτίο μιας παρόμοιας ημέρας θεωρείται ως πρόβλεψη. Αντί για ένα ενιαίο παρόμοιο φορτίο ημέρας, η πρόβλεψη μπορεί να είναι ένας γραμμικός συνδυασμός ή μια διαδικασία παλινδρόμησης που μπορεί να περιλαμβάνει πολλές παρόμοιες ημέρες. Οι



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
συντελεστές τάσης μπορεί να χρησιμοποιηθούν για παρόμοιες ημέρες στα προηγούμενα χρόνια.

Μέθοδοι παλινδρόμησης. Η Παλινδρόμηση είναι μία από τις πιο ευρέως χρησιμοποιούμενες στατιστικές τεχνικές. Για την πρόβλεψη του ηλεκτρικού φορτίου οι μέθοδοι παλινδρόμησης χρησιμοποιούνται για την αναπαράσταση της σχέσης της κατανάλωσης φορτίου και άλλους παράγοντες, όπως καιρός, τον τύπο ημέρας, και την κατηγορία του πελάτη. Ο Engle παρουσίασε διάφορα μοντέλα παλινδρόμησης για την πρόβλεψη επόμενης κορυφής της ημέρας. Τα μοντέλα τους ενσωματώνουν ντετερμινιστικές επιρροές, όπως οι διακοπές, στοχαστικές επιρροές, όπως τα μέσα φορτία, και εξωγενείς επιρροές, όπως ο καιρός.

Χρονοσειρές. Οι μέθοδοι χρονοσειρών βασίζονται στην υπόθεση ότι τα δεδομένα έχουν μια εσωτερική δομή, όπως αυτοσυσχέτιση, τάση, ή εποχική διακύμανση. Μέθοδοι πρόβλεψης χρονοσειράς ανιχνεύουν και διερευνούν μια τέτοια δομή. Οι χρονοσειρές έχουν χρησιμοποιηθεί εδώ και δεκαετίες σε τομείς όπως η οικονομία, η ψηφιακή επεξεργασία σήματος, καθώς και πρόβλεψη φορτίου ηλεκτρική. Ειδικότερα, ARMA (αυτοπαλίνδρομο κινητού μέσου όρου), ARIMA (αυτοπαλίνδρομο ολοκληρωμένου κινητού μέσου όρου), ARMAX (αυτοπαλίνδρομο κινητού μέσου όρου με εξωγενείς μεταβλητές), και ARIMAX (αυτοπαλίνδρομο ολοκληρωμένου μέσου όρου που διακινούνται με εξωγενείς μεταβλητές) είναι οι πιο συχνά χρησιμοποιούμενες κλασικές μεθόδους χρονοσειρών. Τα μοντέλα ARMA συνήθως χρησιμοποιούνται για σταθερές διαδικασίες, ενώ τα ARIMA είναι μια επέκταση της ARMA στις μη στάσιμες διαδικασίες. Τα ARMA και ARIMA χρησιμοποιούν το χρόνο και το φορτίο ως παραμέτρους εισόδου. Το φορτίο εξαρτάται γενικά από τον καιρό και την ώρα της ημέρας, Το ARIMAX είναι το πιο φυσικό εργαλείο για την πρόβλεψη φορτίου μεταξύ των κλασικών μοντέλων χρονοσειράς.

Οι Fan, McDonald και Cho περιγράφουν υλοποιήσεις των ARIMAX μοντέλων για την πρόβλεψη φορτίου. Ο Yang χρησιμοποίησε εξελικτική προγραμματιστική (ΕΚ) Προσέγγιση για την αναγνώριση των παραμέτρων του μοντέλου ARMAX για μία ημέρα έως μία εβδομάδα μπροστά στην ωριαία πρόβλεψη της ζήτησης φορτίου. Ο Εξελικτικός προγραμματισμός είναι μια μέθοδος για την προσομοίωση της εξέλιξης και αποτελεί ένα στοχαστικό αλγόριθμο βελτιστοποίησης. Οι Yang και Huang πρότειναν μια ασαφή αυτοπαλίνδρομο κινητό μέσο όρο μέθοδο με εξωγενείς μεταβλητές εισόδου (FARMAX) για μία ημέρα μπροστά ωριαίες προβλέψεις φορτίου.

Νευρωνικά δίκτυα. Η χρήση των τεχνητών νευρωνικών δικτύων (ΤΝΔ ή NN) έχει μια ευρέως μελετημένη τεχνική πρόβλεψης ηλεκτρικού φορτίου από το 1990. Τα νευρωνικά δίκτυα είναι ουσιαστικά μη-γραμμικά κυκλώματα που έχουν την ικανότητα να αποδείξουν ότι μπορούν να κάνουν καμπύλη μη γραμμικής προσαρμογής.

Οι έξοδοι ενός τεχνητού νευρωνικού δικτύου είναι η γραμμική ή μη γραμμική μαθηματική συνάρτηση των εισόδων του. Οι εισοδοί μπορούν να είναι οι έξοδοι



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

άλλων στοιχείων του δικτύου, καθώς και πραγματικές εισροές του δικτύου. Στην πράξη, τα στοιχεία του δικτύου διατάσσονται σε ένα σχετικά μικρό αριθμό συνδεδεμένων στρώματων των στοιχείων του δικτύου μεταξύ των εισροών και εκροών. Τα μονοπάτια Ανατροφοδότησης μερικές φορές χρησιμοποιούνται.

Κατά την εφαρμογή ενός νευρικού δικτύου σε πρόβλεψη ηλεκτρικού φορτίου, πρέπει κανείς να επιλέξει μια συγκεκριμένη αρχιτεκτονική (π.χ. Hopfield, πίσω πολλαπλασιασμού – back propagation, Boltzmann μηχανής), ο αριθμός και η συνδεσιμότητα των στρωμάτων και των στοιχείων, η χρήση διπλής κατεύθυνσης ή μονής κατεύθυνσης δεσμούς, και η μορφή αριθμού (π.χ. δυαδικό ή συνεχής) πρέπει να χρησιμοποιούνται από εισόδους και εξόδους, καθώς και εσωτερικά.

Η πιο δημοφιλής αρχιτεκτονική τεχνητών νευρωνικών δικτύων για την πρόβλεψη φορτίου ηλεκτρικού είναι η back propagation. Η Back propagation νευρωνικών δικτύων χρησιμοποιούν συνεχώς συναρτήσεις και επιβλεπόμενη μάθηση. Αυτά είναι, σύμφωνα με επιβλεπόμενη μάθηση, τα πραγματικά αριθμητικά βάρη στις εισόδους στοιχείων που καθορίζονται με το ταίριασμα ιστορικών δεδομένων (όπως είναι ο χρόνος και οι καιρικές συνθήκες) σε επιθυμητά αποτελέσματα (όπως ιστορικά ηλεκτρικά φορτία) σε μια προ-επιχειρησιακή εκπαίδευση. Τεχνητά νευρωνικά δίκτυα με μη επιβλεπόμενη μάθηση δεν απαιτούν προ-επιχειρησιακή εκπαίδευση.

Ο Μπακιρτζής ανέπτυξε ένα TNA με βάση το βραχυπρόθεσμο μοντέλο πρόβλεψης φορτίου για το Κέντρο Ελέγχου Ενέργειας της ελληνικής Δημόσιας Επιχείρησης Ηλεκτρισμού. Κατά την ανάπτυξη χρησιμοποίησαν ένα πλήρως συνδεδεμένο τριών στρωμάτων feedforward ANN και ο Back propagation αλγόριθμος χρησιμοποιήθηκε για την εκπαίδευση. Μεταβλητές εισόδου περιλαμβάνουν ιστορικά ωριαία δεδομένα φορτίου, θερμοκρασίας, και ημέρας της εβδομάδας. Το μοντέλο μπορεί να προβλέψει προφίλ φορτίου από μία έως επτά ημέρες. Επίσης ο Παπαλεξόπουλος ανέπτυξε και εφάρμοσε ένα πολυεπίπεδο feedforward ANN για το βραχυπρόθεσμο σύστημα πρόβλεψης φορτίου. Στο μοντέλο είναι τρεις οι τύποι μεταβλητών που χρησιμοποιούνται ως εισοδοί στο νευρωνικό δίκτυο: μεταβλητές εισόδου που έχουν σχέση με την εποχή, μεταβλητές εισόδου που σχετίζονται με τον καιρό και τα ιστορικά φορτία. Ο Khotanzad περιέγραψε ένα σύστημα πρόβλεψης φορτίου είναι γνωστό ως ANNSTLF. ANNSTLF βασίζεται σε στρατηγικές πολλαπλών ANN που αιχμαλωτίζουν διάφορες «τάσεις» στα δεδομένα. Κατά την ανάπτυξη χρησιμοποίησαν ένα πολυστρωματικό perceptron εκπαιδευμένο με τον Back propagation αλγόριθμο για την εύρεση σφάλματος. Το ANNSTLF μπορεί να εξετάσει την επίδραση της θερμοκρασίας και σχετική υγρασία με το φορτίο. Περιέχει, επίσης, προβλέψεις που μπορούν να δημιουργήσουν προβλέψεις των ωριαίων θερμοκρασιών και της σχετικής υγρασίας που απαιτούνται από το σύστημα. Στη νέα γενιά, το ANNSTLF περιλαμβάνει δύο ANN, το ένα προβλέπει το βασικό φορτίο και το άλλο προβλέπει την αλλαγή του φορτίου.

Η τελική πρόβλεψη υπολογίζεται από ένα προσαρμοστικό συνδυασμό αυτών των προβλέψεων. Οι επιδράσεις της υγρασίας και της ταχύτητας του ανέμου θεωρούνται μέσω ενός γραμμικού μετασχηματισμού της θερμοκρασίας. Το ANNSTLF χρησιμοποιούνταν κατά 35 επιχειρήσεις κοινής ωφέλειας σε όλη την ΗΠΑ και τον Καναδά. Ο Chen ανέπτυξε ένα τριών επιπέδων πλήρως συνδεδεμένο feedforward



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

ANN και ο Back propagation αλγόριθμος χρησιμοποιήθηκε ως μέθοδος εκπαίδευσης. Τα ANN τους κρίνει αν η τιμή της ηλεκτρικής ενέργειας ως ένα από τα κύρια χαρακτηριστικά του φορτίου του συστήματος. Πολλές δημοσιευμένες μελέτες χρησιμοποιούν τεχνητά νευρωνικά δίκτυα σε συνδυασμό με άλλες τεχνικές πρόβλεψης (όπως με δέντρα παλινδρόμησης, ο χρόνος σειρά, ή ασαφούς λογικής).

Έμπειρα συστήματα. Οι βασισμένες σε Κανόνα προβλέψεις κάνουν χρήση των κανόνων, που συχνά είναι ευρετικοί στη φύση, να κάνουν ακριβείς προβλέψεις. Τα έμπειρα συστήματα ενσωματώνουν τους κανόνες και τις διαδικασίες που χρησιμοποιούνται από τους εμπειρογνώμονες στον τομέα του ενδιαφέροντος σε λογισμικό που είναι σε θέση στη συνέχεια να κάνει αυτόματα προβλέψεις τω χωρίς ανθρώπινη βοήθεια.

Η χρήση των έμπειρων συστημάτων ξεκίνησε στη δεκαετία του 1960 για εφαρμογές όπως γεωλογική αναζήτηση και το σχεδιασμό ηλεκτρονικών υπολογιστών. Τα έμπειρα συστήματα λειτουργούν καλύτερα όταν ένας εμπειρογνώμονας είναι διαθέσιμος να συνεργαστεί με τους προγραμματιστές λογισμικού για ένα σημαντικό χρονικό διάστημα στην μετάδοση της γνώσης των εμπειρογνώμωνων για το λογισμικό του συστήματος εμπειρογνώμονα. Επίσης, η γνώση ενός ειδικού πρέπει να είναι κατάλληλη για την κωδικοποίηση των κανόνων σε λογισμικό (δηλαδή ο εμπειρογνώμονας πρέπει να είναι σε θέση να εξηγήσει / διαδικασία του την απόφαση στους προγραμματιστές). Ένα έμπειρο σύστημα μπορεί να κωδικοποιήσει έως και εκατοντάδες ή χιλιάδες κανόνες παραγωγής.

Ο Ho πρότεινε ένα έμπειρο σύστημα για την βραχυπρόθεσμη πρόβλεψη φορτίου του συστήματος ηλεκτρικής ενέργειας της Ταϊβάν. Η γνώση του διαχειριστή και οι ωριαίες παρατηρήσεις του φορτίου του συστήματος κατά τη διάρκεια των τελευταίων πέντε ετών είχαν χρησιμοποιηθεί για τη δημιουργία ένδεκα τύπους της ημέρας. Καιρικοί παράμετροι εξετάστηκαν επίσης. Ο αναπτυσσόμενος αλγόριθμος εκτελείται καλύτερα σε σύγκριση με τη συμβατική Box-Jenkins μέθοδο. Ο Rahman και Χαζίμ ανέπτυξαν μια τεχνική ιστοσελίδα-ανεξάρτητη για τη βραχυπρόθεσμη πρόβλεψη φορτίου. Οι γνώσεις σχετικά με το φορτίο και τους παράγοντες που την επηρεάζουν εξάγονται και εκπροσωπούνται σε ένα παραμετροποιημένο βάση τον κανόνα. Αυτός ο κανόνας συνοδεύεται από μια βάση δεδομένων παραμέτρων που διαφέρει από τόπο σε τόπο. Η τεχνική δοκιμάστηκε σε διάφορες τοποθεσίες στις Ηνωμένες Πολιτείες με χαμηλό ποσοστό σφαλμάτων πρόβλεψης.

Το μοντέλο φορτίου, οι κανόνες και οι παράμετροι που παρουσιάζονται στο έγγραφο έχουν σχεδιαστεί χωρίς την χρήση ειδικών γνώσεων σχετικά με κάποιο συγκεκριμένο χώρο. Τα αποτελέσματα μπορούν να βελτιωθούν αν είναι φορείς σε μια συγκεκριμένη τοποθεσία.

Ασαφής λογική - Fuzzy Logic. Ασαφής λογική είναι μια γενίκευση της συνήθους Boolean λογικής που χρησιμοποιείται για την ψηφιακή σχεδίαση του κυκλώματος. Μια είσοδος με Boolean λογική παίρνει μία τιμή αλήθεια της "0" ή "1". Η ασαφής λογική έχει μια είσοδο να σχετίζεται με μια ποιοτική ορισμένη περιοχή. Για παράδειγμα, ένα φορτίο μετασχηματιστή μπορεί να είναι "χαμηλό", "μέσο" και "υψηλό". Ασαφής λογική επιτρέπει σε κάποιον να συμπεράνει (λογικά) εξόδους από



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου ασαφείς εισόδους. Με αυτή την έννοια η ασαφής λογική είναι μια από τις τεχνικές χαρτογράφησης για εισόδους προς τις εξόδους (δηλαδή προσαρμογή καμπύλης).

Μεταξύ των πλεονεκτημάτων της ασαφούς λογικής είναι η απουσία ανάγκης για ένα μαθηματικό μοντέλο χαρτογράφησης εισροών στις εξόδους και η απουσία ανάγκης για ακριβή (ή ακόμη και χωρίς θόρυβο) εισροών. Με τέτοιους γενικούς κανόνες, σωστά σχεδιασμένα συστήματα ασαφούς λογικής μπορούν να είναι πολύ ισχυρά όταν χρησιμοποιούνται για την πρόβλεψη. Βέβαια, σε πολλές περιπτώσεις μια ακριβής έξοδος (π.χ. το ακριβές φορτίο 24:00) απαιτείται. Μετά την λογική επεξεργασία της ασαφούς εισόδου, μία «defuzzification» διαδικασία μπορεί να χρησιμοποιηθεί για την παραγωγή τέτοιων ακριβή εξόδων.

Support vector machines. Μηχανή Υποστήριξης Διανυσμάτων. Μηχανή Υποστήριξης Διανυσμάτων (SVMs) είναι μια πιο πρόσφατη ισχυρή τεχνική για την επίλυση προβλημάτων ταξινόμησης και παλινδρόμησης. Η προσέγγιση αυτή προήλθε από [35] στατιστική θεωρία μάθησης του Vapnik. Σε αντίθεση με νευρωνικά δίκτυα, τα οποία προσπαθούν να ορίσουν σύνθετες λειτουργίες του χώρου χαρακτηριστικών εισόδου, οι SVMs εκτελούν μια μη γραμμική χαρτογράφηση (με τη χρήση του πυρήνα – Kernel functions) των δεδομένων σε ένα υψηλό χώρο διαστάσεων (χαρακτηριστικό). Στη συνέχεια, μηχανές υποστήριξης διανυσμάτων χρησιμοποιούν απλές γραμμικές λειτουργίες για τη δημιουργία γραμμικών ορίων απόφασης στο νέο χώρο. Το πρόβλημα της επιλογής μιας αρχιτεκτονικής για ένα νευρωνικό δίκτυο αντικαθίσταται εδώ από το πρόβλημα της επιλογής ενός κατάλληλου Kernel για την μηχανή υποστήριξης διανυσμάτων.

Ο Mohandes εφάρμοσε τη μέθοδο των SVMs για τη βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φορτίου. Ο συγγραφέας συγκρίνει τις επιδόσεις του με τη μέθοδο αυτοπαλινδρόμησης. Τα αποτελέσματα δείχνουν ότι οι SVMs συγκρίνονται ευνοϊκά έναντι της μεθόδου αυτοπαλινδρόμησης. Chen πρότεινε ένα μοντέλο SVM να προβλέψει καθημερινή ζήτηση φορτίου ενός μηνός. Το πρόγραμμά τους ήταν η νικήτρια συμμετοχή του στο διαγωνισμό που διοργανώθηκε από το δίκτυο EUNITE. Οι Li και Fang χρησιμοποίησαν επίσης ένα μοντέλο SVM για τη βραχυπρόθεσμη πρόβλεψη φορτίου.

### 3.6 Μελλοντικές ερευνητικές κατευθύνσεις

Σε αυτό το κεφάλαιο έχουν συζητηθεί πολλές στατιστικές και τεχνικές τεχνητής νοημοσύνης που έχουν αναπτυχθεί για τη βραχυπρόθεσμη, μεσοπρόθεσμη και μακροπρόθεσμη πρόβλεψη φορτίου ηλεκτρικού. Πολλά στατιστικά μοντέλα και αλγόριθμους που έχουν αναπτυχθεί, λειτουργούν ad hoc. Η ακρίβεια των προβλέψεων θα μπορούσε να βελτιωθεί, αν κάποιος θα μελετούσε αυτά τα στατιστικά μοντέλα και να αναπτύξουν μια μαθηματική θεωρία που εξηγεί τη σύγκλιση αυτών των αλγορίθμων.

Οι ερευνητές, επίσης, θα πρέπει να διερευνήσουν τα όρια της εφαρμογής των προτεινόμενων μοντέλων και αλγορίθμων. Μέχρι στιγμής, δεν υπάρχει ένα ενιαίο μοντέλο ή αλγόριθμος που είναι ανώτερος για όλες τις επιχειρήσεις κοινής ωφέλειας. Ο λόγος είναι ότι οι τομείς παροχής υπηρεσιών κοινής ωφέλειας ποικίλλουν σε



διαφορετικά μείγματα βιομηχανικών, εμπορικών και οικιακών πελατών. Μπορούν επίσης να ποικίλουν σε γεωγραφικές, κλιματολογικές, οικονομικές, και κοινωνικά χαρακτηριστικά. Επιλέγοντας το πιο κατάλληλο αλγόριθμο με ένα βοηθητικό πρόγραμμα μπορεί να γίνει με εξέταση των αλγορίθμων σε πραγματικά δεδομένα. Στην πραγματικότητα, ορισμένες εταιρείες κοινής ωφέλειας χρησιμοποιούν διάφορες μεθόδους πρόβλεψης φορτίου παράλληλα. Σε ό, τι γνωρίζουμε, τίποτα δεν είναι γνωστό για ένα εκ των προτέρων αλγόριθμο που θα μπορούσε να εντοπίσει ποια μέθοδος πρόβλεψης είναι πιο κατάλληλη για μια δεδομένη περιοχή φορτίου. Ένα σημαντικό ερώτημα είναι να διερευνηθεί η ευαισθησία των αλγορίθμων και των μοντέλων πρόβλεψης φορτίου με τον αριθμό των πελατών, τα χαρακτηριστικά της περιοχής, τις τιμές ενέργειας, και άλλους παράγοντες.

Όπως αναφέρθηκε παραπάνω, ο καιρός είναι ένας σημαντικός παράγοντας που επηρεάζει το φορτίο. Η συνήθης προσέγγιση για την βραχυπρόθεσμη πρόβλεψη φορτίου χρησιμοποιεί το προβλεπόμενο σενάριο καιρού ως δεδομένο. Ωστόσο, μία από τις πιο σημαντικές πρόσφατες εξελίξεις στην πρόγνωση του καιρού είναι η λεγόμενη προσέγγιση συνόλου που αποτελείται από τον υπολογισμό πολλαπλών προβλέψεων. Τότε βάρη πιθανοτήτων μπορούν να αποδοθούν σε αυτά τα σύνολα.

Αντί να χρησιμοποιεί τη ενιαία πρόγνωση του καιρού, καιρικές προβλέψεις συνόλου μπορούν να χρησιμοποιηθούν ως πολλαπλές εισόδους για τις προβλέψεις φορτίου. Αυτές οι εισοδοί δημιουργούν πολλαπλές προβλέψεις φορτίου. Στα τελευταία έγγραφα, οι συγγραφείς περιγράφουν προβλέψεις φορτίου συνόλου βασισμένες σε 51 καιρικά σύνολα και διάφορες στατιστικές μεθόδους πρόβλεψης. Υπάρχουν δύο πλεονεκτήματα της πρόβλεψης φορτίου στη πιθανολογική μορφή: (i) μπορούν να οδηγήσουν σε μια πιο ακριβή ωριαία πρόβλεψη που λαμβάνεται με τη χρήση πολλαπλών συνόλων, για παράδειγμα, με τον μέσο όρο τους, (ii) η πιθανοτική περιγραφή του μελλοντικού φορτίου μπορεί να είναι χρησιμοποιηθεί ως πρώτη ύλη για τα συστήματα υποστήριξης αποφάσεων για να κάνουν σημαντικές παραγωγές, αγορά, και αλλαγή αποφάσεων.

Σε γενικές γραμμές, είναι γνωστό από τα κατάλληλα μαθηματικά μοντέλα ότι η γνώση της κατανομής της ζήτησης οδηγεί σε περισσότερες οικονομικά αποδοτικές αποφάσεις από τη γνώση της αναμενόμενης ζήτησης. Σε ευρύτερη κλίμακα, πιστεύουμε ότι η σημαντική έρευνα και ανάπτυξη κατευθύνσεων είναι οι εξής: (i) το συνδυασμό καιρού και πρόβλεψη φορτίου και (ii) ενσωμάτωση της πρόβλεψης φορτίου σε διάφορα συστήματα υποστήριξης αποφάσεων.

### 3.7 Συμπέρασμα

Η ακριβής πρόβλεψη φορτίου είναι πολύ σημαντική για εταιρείες παραγωγής ηλεκτρικής ενέργειας σε ανταγωνιστικό περιβάλλον που δημιουργείται από την απελευθέρωση της ηλεκτρικής βιομηχανίας. Στη παρούσα εργασία εξετάστηκαν κάποιες στατιστικές και τεχνικές τεχνητής νοημοσύνης που χρησιμοποιούνται για την πρόβλεψη ηλεκτρικού φορτίου. Αναφέρθηκαν, επίσης, παράγοντες που επηρεάζουν την ακρίβεια των προβλέψεων, όπως τα καιρικά δεδομένα, παράγοντες του χρόνου,





Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

κατηγορίες πελατών, καθώς και της οικονομικής και τελικής-χρήσης παράγοντες. Μέθοδοι πρόβλεψης φορτίου χρησιμοποιούν μοντελοποίηση προηγμένων μαθηματικών. Επιπλέον πρόοδος στην πρόβλεψη φορτίου και τη χρήση του σε βιομηχανικές εφαρμογές μπορεί να επιτευχθεί με την παροχή βραχυπρόθεσμων προβλέψεων φορτίου υπό τη μορφή των κατανομών πιθανοτήτων παρά των προβλεπόμενων αριθμών, για παράδειγμα η λεγόμενη προσέγγιση συνόλου μπορεί να χρησιμοποιηθεί. Θεωρείται ότι η πρόοδος για την πρόβλεψη του φορτίου θα επιτευχθεί σε δύο κατευθύνσεις: (i) τη βασική έρευνα στον τομέα της στατιστικής και της τεχνητής νοημοσύνης και (ii) την καλύτερη κατανόηση της δυναμικής φορτίου και στατιστικές ιδιότητες της να εφαρμόσει τα κατάλληλα μοντέλων.

## 4. Υλοποίηση

---

### 4.1 Χρησιμότητα του εργαλείου για την πρόβλεψη φορτίου

Το παρόν εργαλείο που κατασκευάστηκε για την πρόβλεψη ηλεκτρικού φορτίου αποτελεί μεγάλη σημαντικότητα για τους ηλεκτρολόγους μηχανικούς και για τους στατιστικολόγους. Είναι μια εφαρμογή που πρέπει να παράγει ακριβή αποτελέσματα με το μικρότερο δυνατό σφάλμα.

Από σκοπιά των ηλεκτρολόγων μηχανικών, οποιοσδήποτε που είναι υπεύθυνος για την ομαλή λειτουργία του ηλεκτρικού δικτύου και την σωστή λειτουργία των βοηθητικών προγραμμάτων, διαχειριστών του συστήματος και διαφόρων πολυάριθμων καταρροή τμημάτων, όπως τον προγραμματισμό των πόρων και το σύστημα διανομής, απαιτεί ακριβή πρόβλεψη για το ηλεκτρικό φορτίο.

Από την πλευρά των στατιστικολόγων, οποιοσδήποτε που ασχολείται με την αγορά ηλεκτρικής ενέργειας χρειάζεται την ακριβή πρόβλεψη φορτίου και τιμών με σκοπό να γνωρίζουν το ποσό που πρέπει να παράγει μια μονάδα παραγωγής για να μη έχει ζημιά η υπηρεσία κοινής ωφέλειας. Οι έμποροι ενέργειας έχουν ανάγκη για ακριβή μοντέλα πρόβλεψης καθώς αντιπροσωπεύουν ένα κλειδί που είναι χρήσιμο για την διατίμηση συμβολαίων, για την διαχείριση κινδύνου και άλλων πολυάριθμων στατιστικών εφαρμογών.

### 4.2 Στόχος

Οι στόχοι της έρευνας επιγραμματικά είναι οι εξής:

- Εύρεση Βέλτιστου Αλγορίθμου μηχανικής μάθησης. Ο αλγόριθμος πρέπει να:
  - Είναι γρήγορος στους υπολογισμούς.
  - Χρησιμοποιεί λίγους πόρους για τους υπολογισμούς.
  - Φέρει ακριβή αποτελέσματα πρόβλεψης ηλεκτρικού φορτίου.
- Βελτίωση της αξιοπιστίας του δικτύου και στις μειωμένες εμφανίσεις στις βλάβες του εξοπλισμού και διακοπές ρεύματος.
- Επίτευξη μεγαλύτερου κέρδους για την επιχείρηση κοινής ωφέλειας.



### 4.3 Απαιτήσεις

Οι απαιτήσεις για την υλοποίηση τέτοιου μοντέλου είναι η επιλογή ακριβούς μοντέλου πρόβλεψης, ώστε τα αποτελέσματα της πρόβλεψης είναι αρκετά ακριβή και να ικανοποιούν τους εκάστοτε περιορισμούς, για παράδειγμα, λόγω της σημασίας της ακρίβειας, απαιτείται ποσοστό επιτυχίας που είναι 97-99% ακριβές.

### 4.4 Προκλήσεις στην ενσωμάτωση ενός συστήματος πρόβλεψης φορτίου. (Από τρίτους)

Οι επιχειρήσεις κοινής ωφέλειας όταν επιθυμούν να αποκτήσουν ένα τέτοιο εργαλείο έχουν τρεις παραδοσιακούς τρόπους:

- Αγορά ενός (όχι γνωστού) έτοιμου προγράμματος.
- Πρόσληψη εξωτερικών φορέων για την ανάπτυξη τέτοιου συστήματος.
- Ανάπτυξη του συστήματος με παραδοσιακές προγραμματιστικές γλώσσες, όπως C , Java.

Με τους τρεις παραπάνω τρόπους οι επιχειρήσεις αντιμετωπίζουν αρκετά προβλήματα τα οποία κοστίζουν αρκετούς πόρους.

Με την πρώτη προσέγγιση, αγοράζοντας ένα πρόγραμμα από την αγορά δεν υπάρχει δυνατότητα για να προσαρμοστεί πλήρως στις ανάγκες της εταιρίας. Με αποτέλεσμα να μη φτάσουν την μέγιστη απαιτούμενη ακρίβεια.

Στο δεύτερο τρόπο προσλαμβάνοντας εξωτερικούς σύμβουλους το πρόγραμμα και σύστημα το οποίο παραδίδουν είναι μοντέλο τύπου «μαύρο κουτί» και δεν έχει διαφάνεια στα μοντέλα πρόβλεψης που χρησιμοποιεί. Στην περίπτωση που η εταιρία επιθυμεί να αναλύσει συγκεκριμένα το πρόγραμμα, δεν έχει την επίγνωση για το πώς αυτά τα μοντέλα παράγουν το προβλεπόμενο φορτίο. Στην περίπτωση που επιθυμείται να γίνει αναβάθμιση του προγράμματος αναγκαστικά η εταιρία είναι εξαρτώμενη από τους εξωτερικούς φορείς και αυτό μπορεί να κοστίζει εναλλακτικά.

Τέλος στην τρίτη περίπτωση όπου η εταιρία επιθυμεί να αναπτύξει τέτοιο σύστημα με παραδοσιακές προγραμματιστικές γλώσσες, θα δημιουργηθεί πάνω στις ανάγκες της, αλλά ο χρόνος ανάπτυξης του είναι αρκετά μεγάλος. Είναι ένα πρόγραμμα το οποίο χρειάζεται μήνες ίσως και χρόνο για να ολοκληρωθεί, κι από την οπτική γωνία των μηχανικών υπάρχει πιθανότητα να μην υπάρχουν πόροι για να δημιουργηθεί τέτοιο πρόγραμμα.

**4.5 Λύσεις στην ενσωμάτωση ενός συστήματος πρόβλεψης φορτίου. ( Από τρίτους).**

Χρησιμοποιώντας το Matlab για την ανάπτυξη τέτοιου συστήματος οι παραπάνω προκλήσεις εξαλείφονται.

Στην πρώτη πρόκληση στην ανικανότητα προσαρμογής, το matlab αντίθετα έχει εύκολη μοντελοποίηση. Αυτό γίνεται γιατί το εργαλείο έχει ολοκληρωμένο εργαλείο ανάπτυξης και βιβλιοθήκες που χρησιμοποιούν συναρτήσεις που μπορούν να προσαρμοστούν στις ανάγκες της εταιρίας. Επίσης λόγω της εύκολης μοντελοποίησης και παραμετροποίησης των συναρτήσεων είναι δυνατή και εύκολη η απεικόνιση των αποτελεσμάτων και ο χρόνος ανάπτυξης τους μπορεί να γίνει σε μικρό χρονικό διάστημα.

Στο δεύτερο πρόβλημα που παρουσιάζεται, λόγω των αλλαγών που μπορούν να γίνουν στις βιβλιοθήκες και στις εκάστοτε συναρτήσεις, οι μοντελοποιήσεις που γίνονται με το matlab είναι τύπου «άσπρου κουτιού». Έτσι η εταιρία που δημιουργεί το πρόγραμμα μπορεί να γνωρίζει και να επικυρώνει λεπτομερώς τα βήματα και τους υπολογισμούς που κάνει το σύστημα για την πρόβλεψη του ηλεκτρικού φορτίου, χρησιμοποιώντας το διαδραστικό debugging του matlab.

Στην τελευταία πρόκληση που αντιμετωπίζεται από τις εταιρίες, το matlab δεν επικεντρώνεται στον προγραμματισμό αλλά στην μοντελοποίηση για την δημιουργία του βέλτιστου συστήματος. Φυσικά απαιτείται από την εταιρία να γνωρίζει την προγραμματιστική λογική, αλλά επειδή το matlab προσφέρει τις έτοιμες βιβλιοθήκες και συναρτήσεις. Η ανάπτυξη του συστήματος είναι τύπου point-and-click μόνο με δύο – τρεις κινήσεις μπορούν να δημιουργηθούν γρήγορα πολυάριθμες απεικονίσεις των δεδομένων χωρίς κόπο. Η προγραμματιστική γλώσσα του προγράμματος είναι υψηλού επιπέδου και αρκετά κομμάτια του κώδικα παράγονται αυτόματα, έτσι ο χρόνος για την ολοκλήρωση του συστήματος είναι αρκετά μικρός, και μπορεί οποιοσδήποτε να χρησιμοποιήσει το matlab εφόσον δεν απαιτούνται μεγάλες προγραμματιστικές γνώσεις.

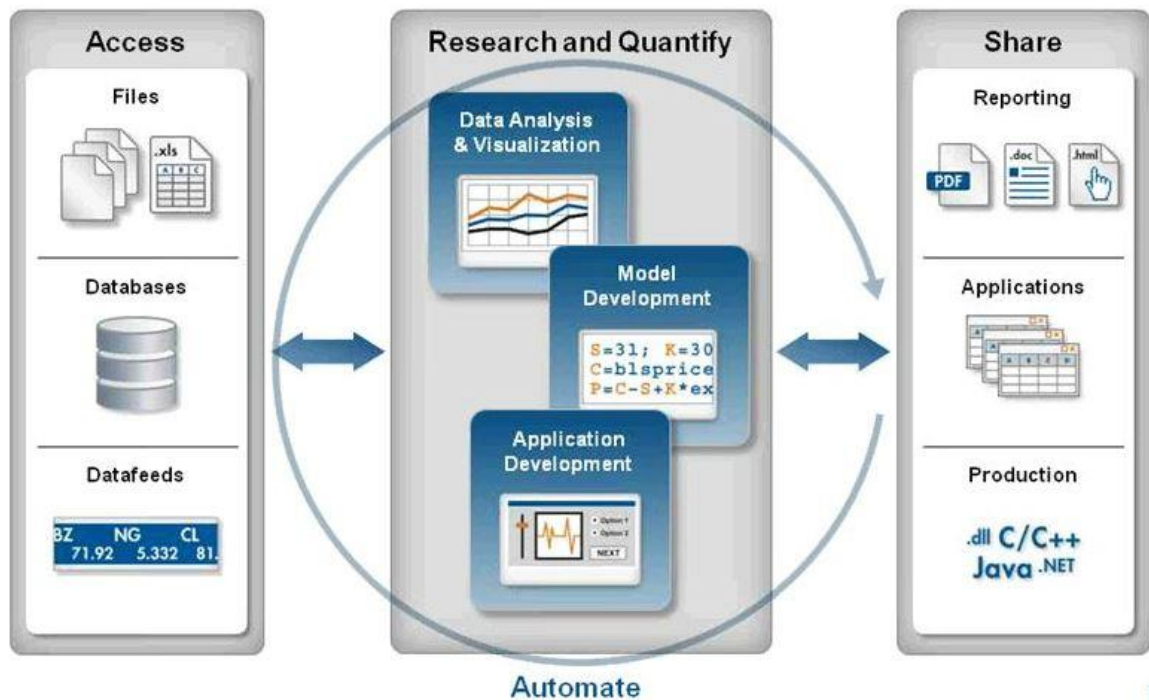
<b>Παραδοσιακές Προσεγγίσεις</b>	<b>Προκλήσεις</b>	<b>Λύσεις (Matlab)</b>
Πρόγραμμα Αγοράς	Ανικανότητα Προσαρμογής	Ευέλικτη μοντελοποίηση
Εξωτερικοί φορείς	Έλλειψη διαφάνειας	Τύπου «άσπρο κουτί» μοντελοποίηση
Ανάπτυξη με παραδοσιακές γλώσσες	Μεγάλο χρονικό διάστημα ανάπτυξης	Γρήγορη προτυποποίηση

*Πίνακας: Παρουσίαση των προσεγγίσεων απόκτησης ενός συστήματος προκλήσεων και λύσεων που παρουσιάζονται.*



## 4.6 Γενική Ροή Μοντέλου Ανάπτυξης.

# Model Development Workflow



Εικόνα: Αναπαράσταση γενικού μοντέλου ανάπτυξης

Για την δημιουργία ενός παρομοίου συστήματος χρησιμοποιείται το παραπάνω μοντέλο ανάπτυξης.

Το πρώτο στάδιο είναι η πρόσβαση στα δεδομένα, τα οποία είναι σε επεξεργάσιμη μορφή από το σύστημα. Για παράδειγμα τέτοια αρχεία είναι .xls ή βάσεων δεδομένων ή ροές δεδομένων. Τα δεδομένα αυτά εισάγονται και χρησιμοποιούνται από το σύστημα.

Το Δεύτερο στάδιο είναι η επεξεργασία των δεδομένων. Στο παρόν στάδιο επιλέγεται και αναπτύσσεται το καταλληλότερο μοντέλο για την επεξεργασία των δεδομένων και για την απεικόνιση αυτών και των αποτελεσμάτων, δηλαδή η έξοδος του συστήματος. Είναι το σημαντικότερο στάδιο διότι εδώ γίνονται οι πιο σημαντικοί υπολογισμοί.

Το τρίτο και τελευταίο στάδιο είναι το Μοίρασμα του συστήματος. Το κομμάτι του συστήματος αυτό εστιάζεται στην παρουσίαση και μοίρασμα του προγράμματος σε τρίτους και σε άλλους φορείς. Για παράδειγμα η παραγωγή μιας αναφοράς σε μορφή .pdf, .doc, html η οποία είναι ευανάγνωστη και ευπαρουσίαστη. Άλλο παράδειγμα είναι η υλοποίηση μιας διαδραστικής εφαρμογής που μπορεί να χρησιμοποιηθεί από οποιονδήποτε είτε η παραγωγή του σε μια παραδοσιακή προγραμματιστική γλώσσα όπως Java και C/C++.



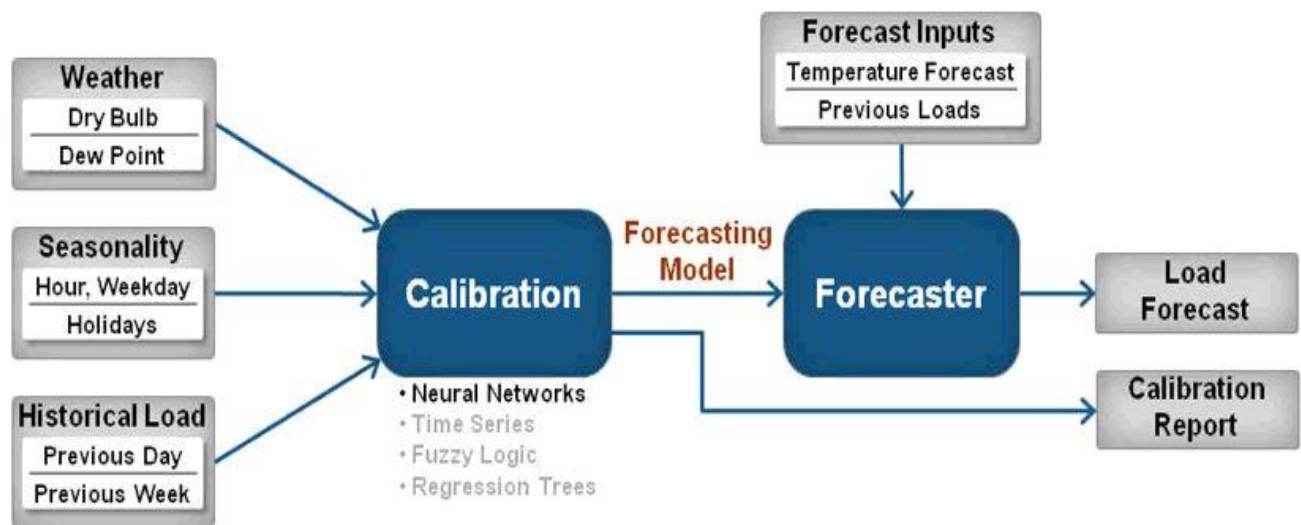
Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

Ένα σημαντικό χαρακτηριστικό του συστήματος πρέπει να είναι ο αυτοματισμός, δηλαδή όλες οι ενέργειες και τα στάδια υλοποίησης του συστήματος να γίνονται αυτόματα με το πάτημα ενός κουμπιού.

## 4.7 Αρχιτεκτονική Μοντέλου

Στην παρακάτω γραφική παράσταση παρουσιάζεται αναλυτικά τα στάδια για την πρόβλεψη του ηλεκτρικού φορτίου. Το μοντέλο είναι βασισμένο στο προαναφερόμενο μοντέλο ανάπτυξης.



Εικόνα:

Αρχιτεκτονική

συστήματος

### 4.7.1 Δεδομένα

Το πρώτο στάδιο, όπως αναφέρθηκε και στο προηγούμενο γενικό μοντέλο, είναι η είσοδος του συστήματος, δηλαδή η εισαγωγή των δεδομένων και η μετατροπή τους σε επεξεργάσιμη μορφή.

Τα δεδομένα που χρησιμοποιήθηκαν αντλήθηκαν από τις μονάδες παραγωγής ενέργειας και μετεωρολογικούς σταθμούς στην περιοχή της Νέας Αγγλίας. (Σημείωση: Τα πρωτότυπα δεδομένα ήταν των σταθμών παραγωγής ηλεκτρικής ενέργειας της Κρήτης, όμως λόγω της μη επεξεργάσιμης μορφής τους, ήταν δύσκολο να εισαχθούν στο σύστημα).

Ο συνολικός αριθμός των παρατηρήσεων (observations) είναι **43848** εκ των οποίων οι **35064** χρησιμοποιήθηκαν για εκπαίδευση (training set) και οι υπόλοιπες **8784** για επικύρωση / τεστ (validation / test set). Τα 35064 δεδομένα εκπροσωπούν τις ωριαίες θερμοκρασίες και φορτίο συστήματος παραγωγής ενέργειας από το 2004 έως



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου 2007 και τα υπόλοιπα 8784 δεδομένα του χρονολογικού έτους 2008. Η κάθε παρατήρηση είχε αρχικά 16 χαρακτηριστικά ( features) και στην συνέχεια προστέθηκαν επιπλέον άλλες δύο. Τα αρχικά χαρακτηριστικά είναι τα εξής:

<b>Date:</b>	Ημερομηνία σε μορφή MM/DD/YYYY.
<b>Hour</b>	Η τελική «τιμή» της ώρα με την ώρα 1 ίση με την ώρα που λήγει στις 1:00 το πρωί.
<b>DA_DEMD</b>	Η επαύριον ζήτηση που αποτελείται από σταθερές και ευαίσθητες προσφορές ζήτησης και προσφορές προσαύξησης και η προσφορές μεταβλητής μείωσης.
<b>DEMAND</b>	Η ζήτηση είναι η μη-PTF ζήτηση για ISO-NE CA( άθροισμα όλων των ζωνών και του σταθμού) και τις ζώνες φορτίου, όπως καθορίζεται από τη μέτρηση. Μη-PTF ζήτηση είναι το φορτίο που χρησιμοποιείται στη διαδικασία διευθέτησης και υπολογίζεται ως εξής: <i>PTF ζήτηση = [μη παραδοτέες + απεριόριστη + σταθμός]</i>
<b>DA_LMP</b>	Η γεωγραφική οριακή τιμή της επόμενης μέρας.
<b>DA_EC</b>	Η τιμή της ενεργειακής συνιστώσας της επόμενης μέρας.
<b>DA_CC</b>	Η τιμή της συμφόρησης συνιστώσας της επόμενης μέρας.
<b>DA_MLC</b>	Η τιμή της οριακής απώλειας συνιστώσας της επόμενης μέρας.
<b>RT_LMP</b>	Η γεωγραφική οριακή τιμή σε πραγματικό χρόνο.
<b>RT_EC</b>	Η τιμή της ενεργειακής συνιστώσας σε πραγματικό χρόνο.
<b>RT_CC</b>	Η τιμή της συμφόρησης συνιστώσας σε πραγματικό χρόνο.
<b>RT_MLC</b>	Η τιμή της οριακής απώλειας συνιστώσας σε πραγματικό χρόνο.
<b>DryBulb</b>	Η θερμοκρασία ξηρού βολβού σε βαθμούς Fahrenheit για τον αντίστοιχο μετεωρολογικό σταθμό.
<b>DewPnt *</b>	Η θερμοκρασία σημείου δρόσου σε βαθμούς Fahrenheit για τον αντίστοιχο μετεωρολογικό σταθμό.
<b>SYSLoad</b>	(για το ISO-NE CA μόνο) Το πραγματικό φορτίο του συστήματος σε MW, όπως καθορίζεται από τη μέτρηση. Το φορτίο του συστήματος που χρησιμοποιείται για την επαύριον και μακροπρόθεσμη πρόβλεψη και αναφορά. Υπολογίζεται ως εξής: Load System = [παραγωγή - άντληση + δίκτυο ανταλλαγής].
<b>RegCP</b>	(για το ISO-NE CA worksheet μόνο) Η τιμή εκκαθάρισης σύμφωνα με τον κανονισμό \$/MWh

\* **Σημείωση:** Ο μετεωρολογικός σταθμός που αντιστοιχεί σε κάθε ζώνη φορτίου παρουσιάζεται ως η "καλύτερη επιλογή", με βάση τη γεωγραφική εγγύτητα με τη ζώνη φορτίου και δεν αντιπροσωπεύει "πραγματικό καιρός" για τη ζώνη φορτίου, αλλά μόνο για το καταχωρημένο μετεωρολογικό σταθμό. Δεν διεξήχθη ανάλυση αυτών των μετεωρολογικών σταθμών για τον προσδιορισμό της καταλληλότητας της θέσης



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου σταθμού για κάθε ζώνη φορτίου κατά ISO New England ούτε οι μετεωρολογικοί σταθμοί που χρησιμοποιούνται σε αναλύσεις ή φορτίου προβλέψεις ζώνη, εκτός από το σταθμό οκτώ σταθμισμένο μέσο όρο της Νέας Αγγλίας.

Η θερμοκρασία της Νέας Αγγλίας είναι μια τιμή με βάρη των οκτώ μετεωρολογικών σταθμών της πολιτείας, και υπολογίζεται με βάση τους ακόλουθους παράγοντες που προέρχονται από τα ιστορικά στοιχεία των πωλήσεων ηλεκτρικής ενέργειας.

Μετεωρολογικός Σταθμός	Πολιτεία	Κωδικός Πολιτείας	Πλησιέστερη Ζώνη Φορτίου	Καλοκαιρινό Βάρος	Χειμερινό Βάρος
Boston	MA	BOS	NEMASSBost	0.201	0.214
Bridgeport	CT	BDR	--	0.070	0.075
Burlington	VT	BTV	VT	0.046	0.040
Concord	NH	CON	NH	0.058	0.055
Portland	ME	PWM	ME	0.085	0.082
Providence	RI	PVD	RI & SEMASS	0.049	0.048
WindsorLks	CT	BDL	CT	0.277	0.277
Worcester	MA	ORH	WCMASS	0.214	0.209

*Σημείωση: Για τους μήνες Απρίλιο και Οκτώβριο, το ωριαίο φορτίο του συστήματος και τα δεδομένα καιρού είναι κατά μέσες τιμές για τη διόρθωση της θερινής περιόδου. Τα LMP δεδομένα και της μη-PTF ζήτησης αναφέρονται ως μηδενικές τιμές για την ελλείπουσα ώρα. Τον Οκτώβριο, η μέση τιμή χρησιμοποιείται για την διπλή μετρημένη ώρα για το φορτίο του συστήματος και τις καιρικές συνθήκες για να διορθώσει την αλλαγή πίσω στην Ανατολική χειμερινή ώρα.*

Οι επιπλέον δύο μεταβλητές / χαρακτηριστικά που προστέθηκαν είναι η year (χρονιά) και το Numdate το οποίο είναι η ημερομηνία και ώρα σε μορφή ακέραιου αριθμού. Η μετατροπή αυτή είναι απαραίτητη για τον επιτυχή διαχωρισμό των δεδομένων στο σετ εκπαίδευσης και τεστ. Ο διαχωρισμός υλοποιείται στο εισαγωγικό στάδιο του συστήματος δηλαδή πριν την εισαγωγή στο σύστημα πρόβλεψης.



Στο αρχείο *import.m* του matlab εισάγουμε τα 6 αρχεία excel *YYYY\_smd\_hourly.xls* και δημιουργείται ένα dataset – NEData (σετ δεδομένων) που είναι όλες μαζί οι παρατηρήσεις σε έναν «πίνακα». Στην συνέχεια προστίθεται η μεταβλητή *year* και υπολογίζεται το *Numdate*. Τέλος λαμβάνονται υπόψη και σημειώνονται οι σημαντικές εορταστικές μέρες για τις οποίες το φορτίο που καταναλώνεται εκείνες τις ημερομηνίες είναι ασταθές και αρκετά αυξημένο από τις σύνηθες τιμές του. Οι ημερομηνίες που σημειώθηκαν είναι οι εξής:

Ημερομηνία	Εορταστική Μέρα
01/01	New Year's Day
19/01	Martin Luther King Day
16/02	President's Day
31/05	Memorial Day
05/07	Independence Day
06/09	Labor Day
25/11	Thanksgiving
26/11	Day After Thanksgiving
24/12	Christmas

*Πίνακας εορταστικών ημερομηνιών, σημείωση, επειδή τα δεδομένα είναι των Ηνωμένων Πολιτειών Αμερικής, λαμβάνονται υπόψη οι Εορτές της χώρας και όχι της Ελλάδας φυσικά*

#### 4.7.2 Σύνοψη Δεδομένων

Συνοψίζοντας στο συγκεκριμένο κομμάτι. Το πρώτο στάδιο της υλοποίησης του συστήματος είναι οι εισαγωγή των δεδομένων. Αυτά τα δεδομένα αποτελούνται από κάποια χαρακτηριστικά τα οποία χωρίζονται σε τρεις μικρές κατηγορίες: **Καιρικά** χαρακτηριστικά όπως η θερμοκρασία σημείου Δρόσου, και η θερμοκρασία ξηρού Βολβού, **Εποχιακά** χαρακτηριστικά όπως ημέρα, σαββατοκύριακο, εορταστικές μέρες, και τέλος χαρακτηριστικά **Ιστορικό Φορτίου** δηλαδή προηγούμενης μέρας και βδομάδας.

Στο [παράρτημα 1](#) παρουσιάζεται ο κώδικας για την εισαγωγή δεδομένων.





### 4.7.3 Παραγωγή Προγνωστικών Παραγόντων & Επιλογή Βραχυπρόθεσμης ή Μακροπρόθεσμης Πρόβλεψης & Χωρισμός σε σετ δεδομένων εκπαίδευσης και επαλήθευσης. (Generate Predictors & Selection short or long term forecasting & Split Data in Training and Test Set)

Ανάμεσα του σταδίου της εισαγωγής δεδομένων και της διαμέτρησης του μοντέλου πρόβλεψης, προηγούνται 2 μικρά στάδια επεξεργασίας των δεδομένων. Το πρώτο στάδιο είναι η δημιουργία των προγνωστικών χαρακτηριστικών, δηλαδή η επιλογή των σωστών χαρακτηριστικών από το αρχικό σύνολο δεδομένων με τα 18 χαρακτηριστικά. Ο πλεονασμός των χαρακτηριστικών των δεδομένων περιγράφουν πληροφορίες που είναι απαιτητικές για την πρόβλεψη των τιμών του φορτίου, έτσι πρέπει να γίνει σωστή επιλογή των παραμέτρων. Επίσης πρέπει να προσδιοριστεί το είδος της πρόβλεψης, δηλαδή αν είναι βραχυπρόθεσμη ή μακροπρόθεσμη πρόβλεψη, για παράδειγμα στην βραχυπρόθεσμη τα εποχιακά χαρακτηριστικά θα είναι η ώρα όταν είχε δημιουργηθεί το φορτίο, ενώ μακροπρόθεσμα θα ήταν το φορτίο του μήνα ή και χρόνου.

Χρησιμοποιώντας την συνάρτηση:

```
[X, dates, labels] = genPredictors(NEDData, term, holidays);
```

Τα νέα χαρακτηριστικά των δεδομένων που παράγονται είναι τα εξής:

- Θερμοκρασία ξηρού βολβού.
- Θερμοκρασία σημείου Δρόσου.
- Ωρα της ημέρας.
- Μέρα της Εβδομάδας.
- Μια Λογική τιμή που δηλώνει αν είναι εορταστική ημέρα ή Σαββατοκύριακο.
- Μέση Τιμή φορτίου προηγούμενης ημέρας.
- Φορτίο της ίδιας ώρας της προηγούμενης ημέρας.
- Φορτίο της ίδιας ώρας και ημέρας και προηγούμενης βδομάδας.

Στο [παράρτημα 2](#) παρουσιάζεται ο κώδικας για τον χωρισμό των δεδομένων στο σετ εκπαίδευσης και επαλήθευσης.

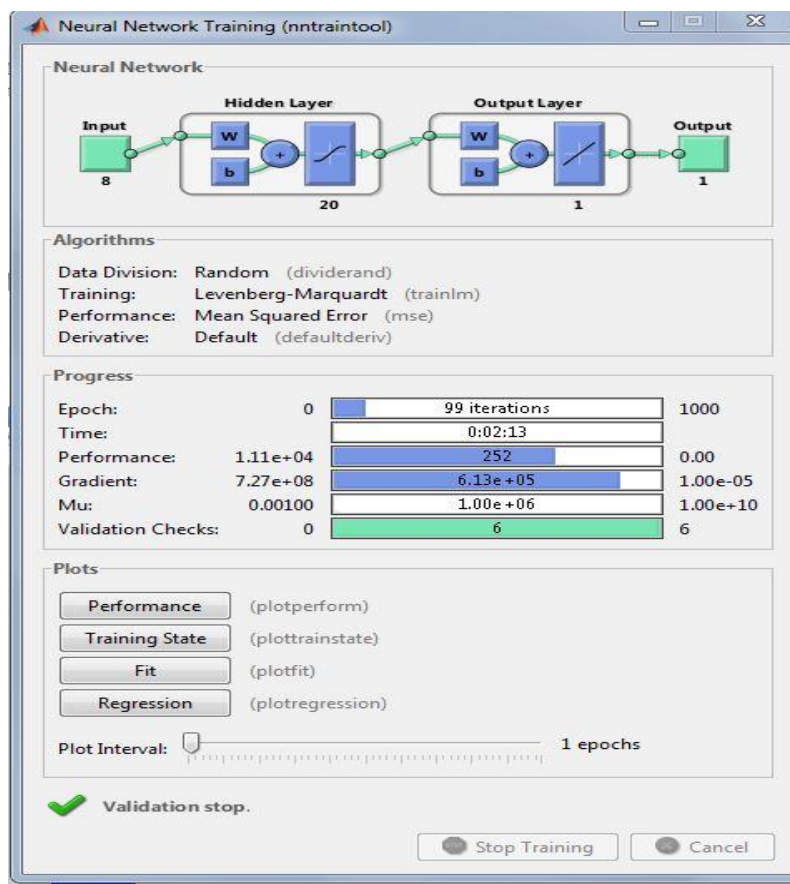
Στο [παράρτημα 3](#) παρουσιάζεται ο κώδικας παραγωγής των νέων προγνωστικών παραμέτρων.



#### 4.7.4 Επιλογή & Ρύθμιση Μοντέλου Πρόβλεψης.

Στο επόμενο στάδιο του συστήματος επιλέγεται το καταλληλότερο μοντέλο που θα πραγματοποιήσει την πρόβλεψη. Τα δεδομένα εκπαίδευσης χρησιμοποιούνται για την εκπαίδευση του συστήματος και η επικύρωση του συστήματος γίνεται με το τεστ σύνολο δεδομένων. Τρία είναι τα μοντέλα που αναπτύχθηκαν για εκπαίδευση του συστήματος τα οποία είναι Νευρωνικό Δίκτυο, Τυχαίο Δάσος, και SVM. Ο σκοπός είναι η επιλογή του μοντέλου που αποδίδει βέλτιστα με το μικρότερο απόλυτο σφάλμα. Αναλυτικότερα περιγράφονται τα μοντέλα παρακάτω.

##### Νευρωνικά Δίκτυα. (Neural Networks)



Το νευρωνικό δίκτυο που κατασκευάστηκε είναι της απλής μορφής regression. Η είσοδος του συστήματος είναι οι **35064** παρατηρήσεις με 8 χαρακτηριστικές ιδιότητες και οι υπόλοιπες **8784** χρησιμοποιήθηκαν για την επικύρωση της ακρίβειάς του. Το νευρωνικό δίκτυο αποτελείται από ένα μόνο κρυφό επίπεδο των 20 νευρώνων και χρησιμοποιεί το απόλυτο μέσο σφάλμα για την μέτρηση της απόδοσης. Στη συνέχεια, εκπαιδεύεται το δίκτυο με το προεπιλεγμένο αλγόριθμο Levenburg-Marquardt. Για λόγους αποδοτικότητας, ένα προ-εκπαιδευμένο δίκτυο είναι φορτωμένο εκτός εάν επανεκπαίδευση έχει εκτελεστεί ειδικά.



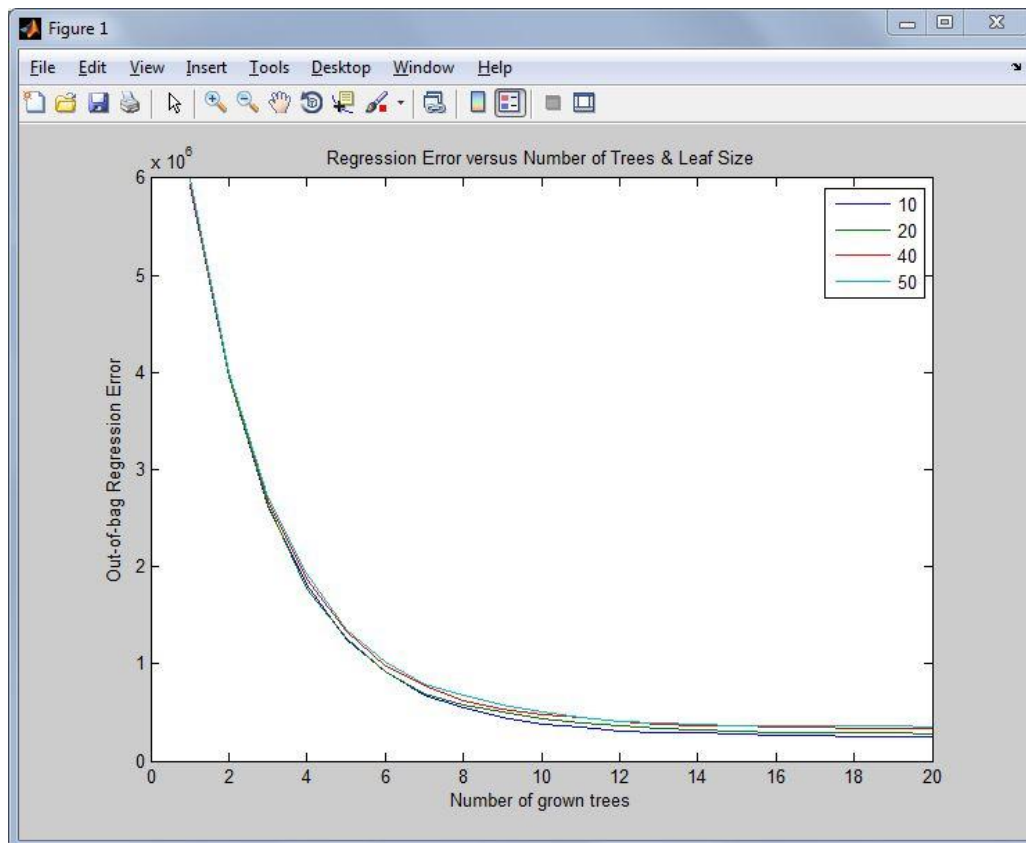
Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
Στο [παράρτημα 4](#) αναδειχεται ο κώδικα για την δημιουργία του Νευρωνικού Δικτύου

### Τυχαία Δάση.(Random Forests)

Το τυχαίο δάσος αποτελείται από πολλαπλά δέντρα αποφάσεων. Οι κύριοι παράμετροι που καθορίζουν αρκετά την ακρίβεια επιτυχίας του συστήματος είναι ο αριθμός των δέντρων αποφάσεων, ο ελάχιστος αριθμός που χωράει στα φύλλα και η επιλογή καταλληλότερων χαρακτηριστικών των δεδομένων. Ο σκοπός είναι η κατασκευή του μοντέλου που να έχει βέλτιστη απόδοση, δηλαδή η επίτευξη μέγιστης ακρίβειας χωρίς να γίνουν χρονοβόροι υπέρ-υπολογισμοί, έτσι με τον σωστό συνδυασμό των δέντρων, ελαχίστου μέγεθος φύλλου και καταλληλότερων χαρακτηριστικών μπορεί να επιτευχθεί το βέλτιστο σύστημα.

### Προσδιορισμός αριθμού των δένδρων και κατάλληλου μεγέθους Φύλλου

Σε αυτό το κομμάτι της έρευνας συγκρίνεται το εκτός-δείγματος σφάλμα πρόβλεψης για τα μεγέθη των φύλλων 10, 20, 40 και 50. Ο έλεγχος γίνεται με την ενεργοποίηση του `_oobpred_` παραμέτρου που επιστρέφει out-of-bag σφάλματα πρόβλεψης κατά τη διάρκεια της εκπαιδευτικής διαδικασίας.



Εικόνα: Σύγκριση σφάλματος, ανάλογα με το μέγεθος του φύλλου.

Από την γραφική παράσταση το συμπέρασμα που αντλείται είναι ότι όσο πιο μικρό το μέγεθος φύλλου τόσο μικρότερο και το σφάλμα. Αυτό οφείλεται στο



Τμήμα Μηχανικών Πληροφορικής

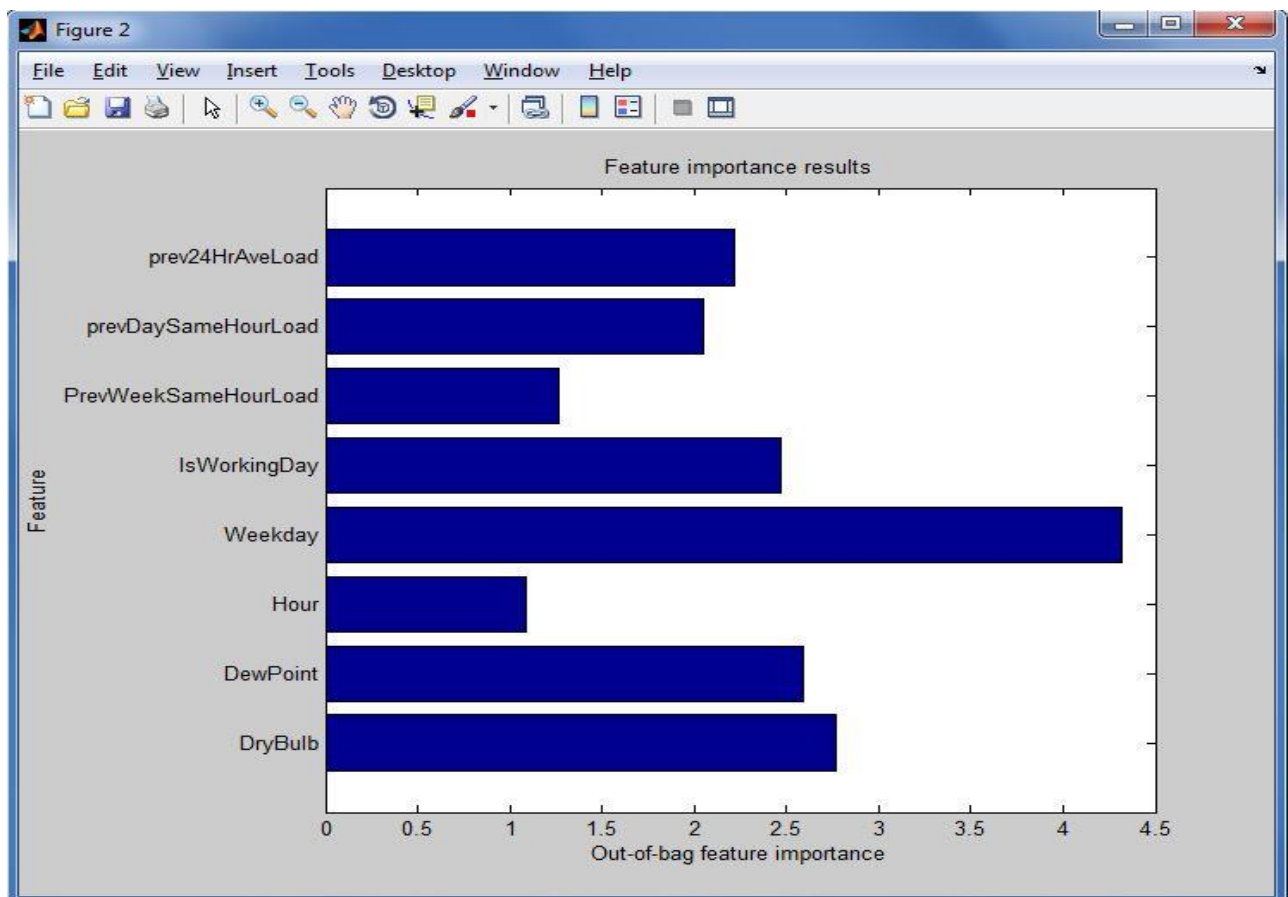
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
γεγονός ότι τα δεδομένα που πάνε να κατηγοριοποιηθούν στα φύλλα είναι λιγότερα έτσι γίνεται και πιο σωστός διαχωρισμός. Επίσης έγιναν και δοκιμές τριπλασιάζοντας τον αριθμό των δέντρων από 20 σε 60 και τα συμπεράσματα είναι ότι ο χρόνος εκτέλεσης της εκπαίδευσης είναι πολύ τεράστιος και δεν παίζει τόσο μεγάλο ρόλο γιατί τα δεδομένα έχουν χωριστεί από το 20<sup>ο</sup> δέντρο.

Στο [παράρτημα 5](#) παρουσιάζεται ο κώδικας που εκτελείται με τις διάφορες τιμές των φύλλων και τον αριθμό των δέντρων.

#### Καθορισμός σημαντικότητας των χαρακτηριστικών

Με την ενεργοποίηση της παραμέτρου `_oobVarImp_` παρουσιάζονται οι out of bag εκτιμήσεις της σημαντικότητας των χαρακτηριστικών.

Στο [παράρτημα 6](#) παρουσιάζεται ο κώδικας που καθορίζει τα πιο σημαντικά χαρακτηριστικά.



Εικόνα: Σύγκριση σημαντικότητας των χαρακτηριστικών.

Από την παραπάνω γραφική παράσταση το συμπέρασμα είναι ότι το πιο σημαντικό χαρακτηριστικό είναι η μέρα της βδομάδας. Για κάθε χαρακτηριστικό, το μέτρο είναι η αύξηση του σφάλματος πρόβλεψης, εάν οι τιμές αυτού του χαρακτηριστικού μετατεθούν σε όλες τις out-of-bag παρατηρήσεις. Το μέτρο υπολογίζεται για κάθε δέντρο, τότε κατά μέσο όρο πάνω σε ολόκληρο το σύνολο και διαιρείται με την τυπική απόκλιση πάνω σε ολόκληρο το σύνολο.



## Support Vector Machine SVM

Για την επίτευξη βέλτιστου ποσοστού επιτυχίας ακολούθησαν μια σειρά βημάτων επεξεργασίας των δεδομένων και εύρεση βέλτιστων τιμών κάποιων παραμέτρων. Τα βήματα είναι τα εξής:

- Μετατροπή των δεδομένων σε μορφή ενός πακέτου SVM.
- Διεξαγωγή απλής κλιμάκωσης στα δεδομένα.
- Χρήση της RBF radial basis function.  $K(x,y) = e^{-\gamma|x-y|^2}$
- Χρήση της μεθόδου cross validation για τον εντοπισμό της παραμέτρου C,  $\gamma$ .
- Εκπαίδευση του συστήματος με τις καλύτερες παραμέτρους που βρέθηκαν στο προηγούμενο βήμα
- Τεστ.

Το πρώτο βήμα δεν ήταν απαραίτητο, γιατί εκτελείται σε περιπτώσεις όπου έχουμε αφαριθμητικές τιμές, ενώ εδώ όλες είναι πραγματικές. Στο δεύτερο βήμα μετατρέπουμε το σύνολο εκπαίδευσης και τεστ σε ένα διάστημα τιμών [0 1]. Αυτή η απλή μετατροπή είναι απαραίτητη ώστε οι υπολογισμοί που θα πραγματοποιηθούν από την RBF kernel θα είναι πιο εύκολοι και σε ένα μικρό εύρος τιμών. Η RBF δεν μπορεί να αναλάβει μεγάλες τιμές.

```
[train, test, jn2] = scaleSVM(train, test, train, 0, 1);
```

Για το τρίτο βήμα υπάρχουν 4 γνωστές συναρτήσεις πυρηνών, αναλόγως με το είδος των δεδομένων γίνεται η επιλογή της κατάλληλης συνάρτησης. Στην προκειμένη περίπτωση τα δεδομένα δεν είναι γραμμικά και ο αριθμός των παρατηρήσεων είναι μεγαλύτερος από τον αριθμό των χαρακτηριστικών για αυτό τους λόγους χρησιμοποιείται η RBF, εν αντίθεση με την γραμμική που χρησιμοποιείται σε αντίθετα είδους δεδομένα. Με την RBF πρέπει να υπολογιστεί η παράμετρος  $\gamma$  και C. Η  $C > 0$  είναι η παράμετρος ποινής του όρου σφάλματος, και η  $\gamma$  είναι παράμετρος της RBF.

Στο τέταρτο στάδιο εφαρμόζεται η μέθοδος cross validation για την εύρεση των 2 κλειδιών παραμέτρων. Δυστυχώς οι τιμές τους δεν είναι γνωστές γι' αυτό μέσα από επαναλήψεις υπολογίζονται οι βέλτιστες τιμές και χρησιμοποιούνται για την εκπαίδευση του συστήματος. Η κλασσική μέθοδος που έχει περιγραφεί για την εκπαίδευση του συστήματος είναι ο χωρισμός των δεδομένων σε 2 σύνολα, σε αυτό της εκπαίδευσης και σε αυτό της επικύρωσης. Στην k-fold cross validation χωρίζεται μόνο το σύνολο εκπαίδευσης σε k τμήματα, συγκεκριμένα 5. Διαδοχικά ένα υποσύνολο εξετάζεται χρησιμοποιώντας τα υπόλοιπα k-1 υποσύνολα. Έτσι κάθε στιγμιότυπο/παρατήρηση ολόκληρου του συνόλου εκπαίδευσης προβλέπεται από μια φορά ώστε η cross validation ακρίβεια είναι το ποσοστό των δεδομένων που έχουν κατηγοριοποιηθεί.



Για την εύρεση των  $C$  και  $\gamma$  και cross validation συνιστάται η μέθοδος «πλέγμα εύρεσης». Η διαδικασία λειτουργεί ως εξής, γίνονται αρκετές επαναλήψεις για την εκπαίδευση του μοντέλου χρησιμοποιώντας συνδυασμούς διαφόρων δοκιμαστικών τιμών των  $C$  και  $\gamma$  (κατά προτίμηση δύναμη του 2). Η διαδικασία είναι απλή με έγκυρα και ασφαλή αποτελέσματα, αλλά είναι χρονοβόρα, γιατί απαιτούνται αρκετές επαναλήψεις μέχρι να βρεθούν οι βέλτιστες τιμές των παραμέτρων. Μόλις βρεθούν οι βέλτιστες τιμές για τις παραμέτρους τότε τις χρησιμοποιούμε για την εκπαίδευση του συστήματος και στο τέλος συγκρίνουμε τα αποτελέσματα με βάση του συνόλου τεστ.

Στο [παράρτημα 7](#) παρουσιάζεται ο κώδικας για την εύρεση βέλτιστου  $C$  και  $\gamma$  μέσα από το Search grid cross validation. & εκπαίδευση μοντέλου χρησιμοποιώντας αυτές τις τιμές.

\* Σημείωση: Λόγω των χρονοβόρων υπολογισμών η μέθοδος search grid ολοκλήρωσε 14 μόνο επαναλήψεις και οι τιμές των παραμέτρων παρουσιάζονται στον παρακάτω πίνακα.

C	$\epsilon=0$	1	2	3
$\gamma=2^{-1}$	270265,44	270257,10	270242,88	270229,53
$2^0$	216564,67	216573,32	216573,88	216559,45
$2^1$	157974,20	157970,66	157979,08	157986,34
$2^2$	114944,72	114935,93	114930,80	<b>114924,12</b>

Από τον παραπάνω πίνακα, επιλέγουμε την παράμετρο  $\gamma=4$  και  $C=114924$ . Όμως λόγω της μεγάλης τιμής της  $C$  οι υπολογισμοί είναι τόσο μεγάλοι όπου παρουσιάζεται overfitting με αποτέλεσμα το μοντέλο να μην αποδίδει σωστά.

#### 4.7.5 Σύνοψη Επιλογής & Ρύθμισης Μοντέλου Πρόβλεψης.

Από τα παραπάνω τρία μοντέλα μηχανικής μάθησης το πιο αποδοτικό και γρηγορότερο μοντέλο ήταν το **τυχαίο δάσος**, ενώ το πιο χρονοβόρο μοντέλο ήταν το **SVM** και τέλος το μοντέλο με τα ελάχιστο σφάλμα ήταν το **νευρωνικό δίκτυο**. Αναλόγως με την περίπτωση και δεδομένα πρέπει να επιλεγεί το πιο κατάλληλο μοντέλο για να υπάρξουν και τα πιο βέλτιστα αποτελέσματα.

#### 4.7.6 Παρουσίαση αποτελεσμάτων.

Αρχικά θα αναφερθούν τρία διαφορετικά σφάλματα που υπολογίστηκαν για τα 3 διαφορετικά μοντέλα, καθώς θα παρουσιαστούν και οι παλαιότερες τιμές των σφαλμάτων, δηλαδή εκείνες προτού υποστούν κατάλληλη επεξεργασία του μοντέλου. Τα είδη των σφαλμάτων είναι τα εξής: **MAE** (mean absolute error – μέσο απόλυτο σφάλμα), **MAPE** (mean absolute percentage error – μέσο απόλυτο ποσοστιαίο



Τμήμα Μηχανικών Πληροφορικής

Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου (σφάλμα) & **daily peak MAPE** (μέσο απόλυτο ποσοστιαίο σφάλμα ημερήσιας κορυφής).

Στο [παράρτημα 8](#) φαίνεται ο κώδικας για την υλοποίηση της παρακάτω γραφικής παράστασης.

Το MAE υπολογίζεται αφαιρώντας τις προβλεπόμενες τιμές από τις πραγματικές της εκάστοτε παρατήρησης (δηλαδή σφάλμα) και ύστερα υπολογίζεται ο μέσος όρος αυτών των διαφορών τιμών.

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| = \frac{1}{n} \sum_{i=1}^n |e_i|.$$

MAE = mean(abs(err));

όπου το  $y_i$  είναι η πραγματική τιμή, και  $f_i$  είναι η προβλεπόμενη. Επίσης ισχύει:  $e_i = |f_i - y_i|$

Το MAPE υπολογίζεται με τον ίδιο τρόπο με το MAE μόνο που υπολογίζεται ποσοστιαία, ο υπολογισμός αυτός είναι χρήσιμος διότι παρατηρείται επίσημα το ποσοστό επιτυχίας του συστήματος.

$$M = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|,$$

MAPE = mean(errpct(~isinf(errpct)));

Όπου  $A_t$  είναι η πραγματική τιμή, και  $F_t$  η προβλεπόμενη τιμή

Τέλος το, είναι το μέγιστο σφάλμα του συστήματος και υπολογίζεται παίρνοντας την απόλυτη διαφορά των μεγίστων του προβλεπόμενου και πραγματικού φορτίου και στην συνέχεια υπολογίζεται το ποσοστό σε τοις εκατό.

$$\text{daily peak MAPE} = \text{mean}(\text{peakerrpct})$$

Μοντέλο	Παραμετροποιήσεις κάθε μοντέλου	Μέσο απόλυτο Ποσοστιαίο Σφάλμα (MAPE):	Μέσο απόλυτο σφάλμα (MAE):	Μέγιστο ημερήσιο Μέσο απόλυτο Ποσοστιαίο Σφάλμα:
Νευρωνικό δίκτυο	Τυπικές	1.81%	272.28 MWh	1.85%
Τυχαίο δάσος	Δέντρα αποφάσεων: <b>20</b> Ελ. αρ. φύλλων: <b>20</b>	2.32%	348.56 MWh	2.36%
	Δέντρα αποφάσεων: <b>70</b> Ελ. αρ. φύλλων: <b>5</b>	2.02%	305.93 MWh	2.08%
	Δέντρα αποφάσεων: <b>100</b> Ελ. αρ. φύλλων: <b>1</b>	1.94%	293.24 MWh	2.01%*
SVM	Τυπικές	13.69%	2286.54 MWh	16.04%
	Μετατροπή δεδομένων στο διαστήμα [0 1]	3.52%	531.89 MWh	3.54%
	Έυρεση βέλτιστων παραμέτρων C, γ	2.11%	317.26 MWh	2.16%



Η Απόδοση των νευρωνικών δικτύων ήταν η καλύτερη και δεν χρειάστηκε καμιά τροποποίηση για την βελτίωση απόδοσης του συστήματος.

### Τυχαίο Δάσος

Το τυχαίο δάσος είχε ικανοποιητικό ποσοστό επιτυχίας, κάνοντας αλλαγές στον αριθμό δέντρων και ελάχιστο αριθμό φύλλων. Παρακάτω παρουσιάζονται τα αποτελέσματα των σφαλμάτων χρησιμοποιώντας τις βέλτιστες παραμετροποιήσεις.

Ensemble with **70** bagged decision trees:

Training X: [35064x8]

Training Y: [35064x1]

Method: regression

Nvars: 8

NVarToSample: 3

MinLeaf: **5**

FBoot: 1

SampleWithReplacement: 1

ComputeOOBPrediction: 0

ComputeOOBVarImp: 0

*\*Σημείωση: Το γεγονός ότι το ποσοστιαίο σφάλμα είναι μικρότερο δεν είναι απαραίτητο ότι είναι σωστό, όταν ο ελάχιστος αριθμός φύλλων είναι τόσο μικρός η εκπαίδευση δεν ολοκληρώνεται σωστά και παρουσιάζεται το πρόβλημα του overfitting.*

### Support Vectors machine

Στην περιγραφή του μοντέλου αναλύθηκαν τα βήματα τα οποία ακολούθησαν ώστε το SVM μοντέλο να δημιουργήσει τα βέλτιστα αποτελέσματα. Το πρώτο βήμα ήταν η μετατροπή των δεδομένων στο διάστημα [0 1]. Αυτή η διαδικασία γίνεται με σκοπό τα δεδομένα να είναι σε ένα εύρος τιμών που μπορεί να γίνει σωστός διαχωρισμός. Το δεύτερο βήμα είναι η εύρεση των βέλτιστων παραμέτρων C και γ που καθορίζουν το διαχωριστικό υπερεπίπεδο.

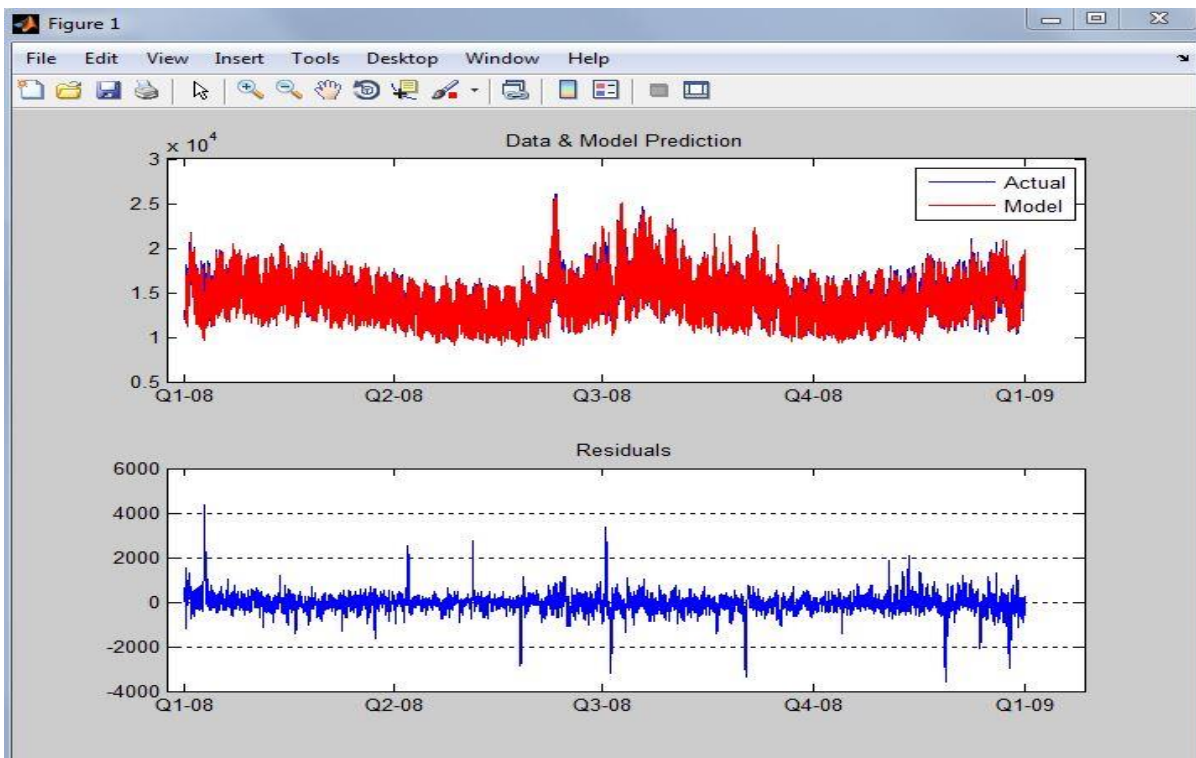
### Γραφικές παραστάσεις.

Στις παρακάτω γραφικές παραστάσεις απεικονίζονται διάφορα στατιστικά τα οποία βοηθούν στην καλύτερη κατανόηση της απόδοσης του συστήματος.





Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
Σύγκριση προβλεπόμενου και πραγματικού φορτίου

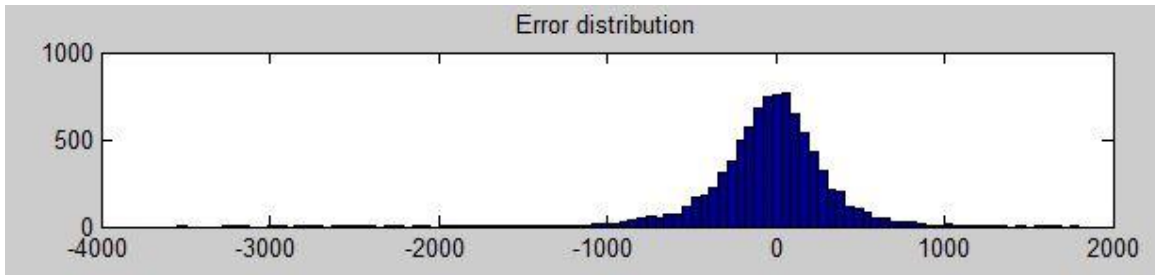


Στην παραπάνω πρώτη γραφική παράσταση απεικονίζεται το προβλεπόμενο & πραγματικό φορτίο σε συνάρτηση με τον χρόνο. Το σύνολο τεστ ήταν το φορτίο της χρονιάς 2008. Η κόκκινη παράσταση είναι το προβλεπόμενο φορτίο και η μπλε το πραγματικό, και ο σκοπός είναι η σύμπτυξη των δυο γραφικών παραστάσεων. Έτσι αντιλαμβάνεται πόσο ακριβές ήταν το μοντέλο πρόβλεψης.

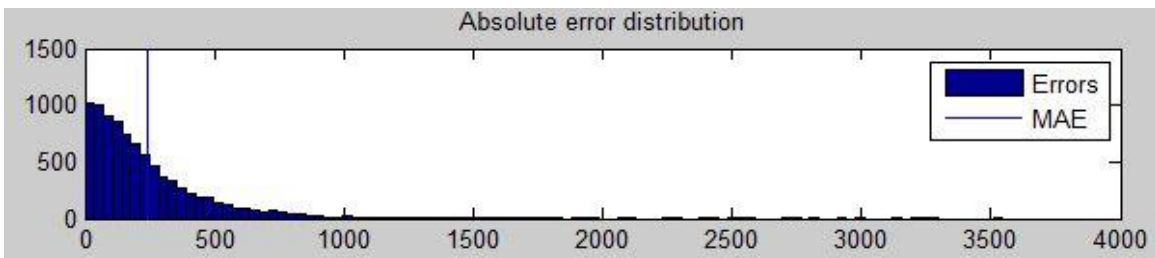
Στην δεύτερη γραφική παράσταση παρουσιάζεται η διαφορά του πραγματικού με το προβλεπόμενο φορτίο ( $\text{testY} - \text{forecast\_Load}$ ) κάθε παρατήρησης για όλο τον χρόνο. Στην γραφική παράσταση παρατηρούνται κάποιες τιμές/κορυφές που είναι αρκετά μεγάλες στους μήνες Ιανουάριο, Μάιο, Ιούλιο, Σεπτέμβριο και Δεκέμβριο, αυτές οι τιμές είναι γνωστές ως και outliers όπου η διαφορά (σφάλμα) είναι αρκετά μεγάλη. Το σύστημα απέτυχε να προβλέψει σωστά στους προαναφερθέντες μήνες γιατί αυτές οι ημερομηνίες είναι πολύ σημαντικές εορταστικές μέρες στις ΗΠΑ όπως η 4<sup>η</sup> Ιουλίου (μέρα ανεξαρτησίας) και υπάρχει απρόβλεπτη κατανάλωσης του φορτίου από τις συνοικίες της Νέας Αγγλίας. Έτσι η οποιαδήποτε υπηρεσία κοινής ωφέλειας πρέπει να είναι προετοιμασμένη για εκείνες τις συγκεκριμένες ημερομηνίες. Δεύτερη σημαντική πληροφορία που μπορεί να αντληθεί από την γραφική παράσταση είναι αν οι τιμές της διαφοράς είναι θετικές ή αρνητικές. Στην περίπτωση που είναι αρνητική η διαφορά σημαίνει ότι το σύστημα πρόβλεψε ότι το φορτίο που πρέπει να παράγει η μονάδα είναι περισσότερο από το πραγματικό που θα καταναλωθεί, έτσι θα περισσεύει φορτίο, αντίθετα στην περίπτωση που είναι θετικό τότε η μονάδα δεν θα παραγάγει αρκετό φορτίο για να καλύψει όλη την περιοχή και θα δημιουργηθεί μεγάλο πρόβλημα και ζημιά στην υπηρεσία.



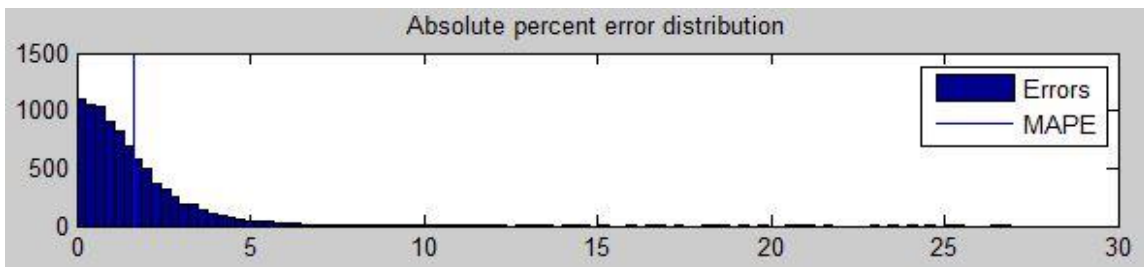
### Εξέταση της κατανομής των σφαλμάτων.



Στο πρώτο ιστόγραμμα ο άξονας  $\Psi$ , δηλαδή το ύψος της μπάρας δηλώνει τον αριθμό των παρατηρήσεων που έχουν το ίδιο σφάλμα  $X$ . Προτιμάται στο ιστόγραμμα οι παρατηρήσεις να τείνουν στο 0, ή αλλιώς οι μπάρες που βρίσκονται στο 0 να είναι οι υψηλότερες, γιατί αυτές είναι οι παρατηρήσεις με το μικρότερο σφάλμα.



Στη δεύτερη γραφική παράσταση απεικονίζεται το απόλυτο σφάλμα γι' αυτό το λόγο δεν υπάρχουν αρνητικές τιμές. Ο σκοπός αυτής της γραφικής παράστασης είναι να απεικονίσει την γραμμή  $x = \text{MAE}$  (μέσο απόλυτο σφάλμα) σε σύγκριση με τις υπόλοιπες τιμές. Όταν τείνει η γραμμή  $x$  στο 0 τότε το σφάλμα είναι μικρότερο.



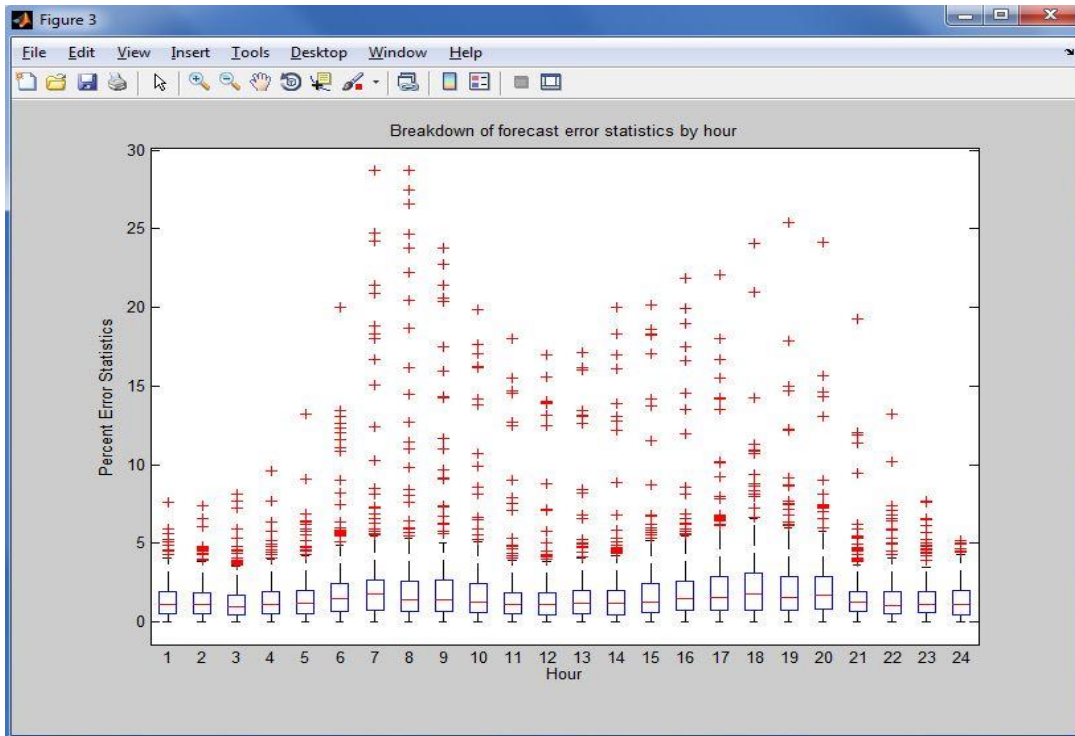
Το τρίτο ιστόγραμμα είναι παρόμοιο με το δεύτερο, μόνο που ο άξονας  $X$  δηλώνει το μέσο σφάλμα σε ποσοστό τοις εκατό. Με την ίδια λογική όσο η γραμμή  $x = \text{MAPE}$  τείνει στο 0 τόσο καλύτερα αποδίδει το σύστημα.

### Ομαδική ανάλυση των δεδομένων ( Group analysis error)

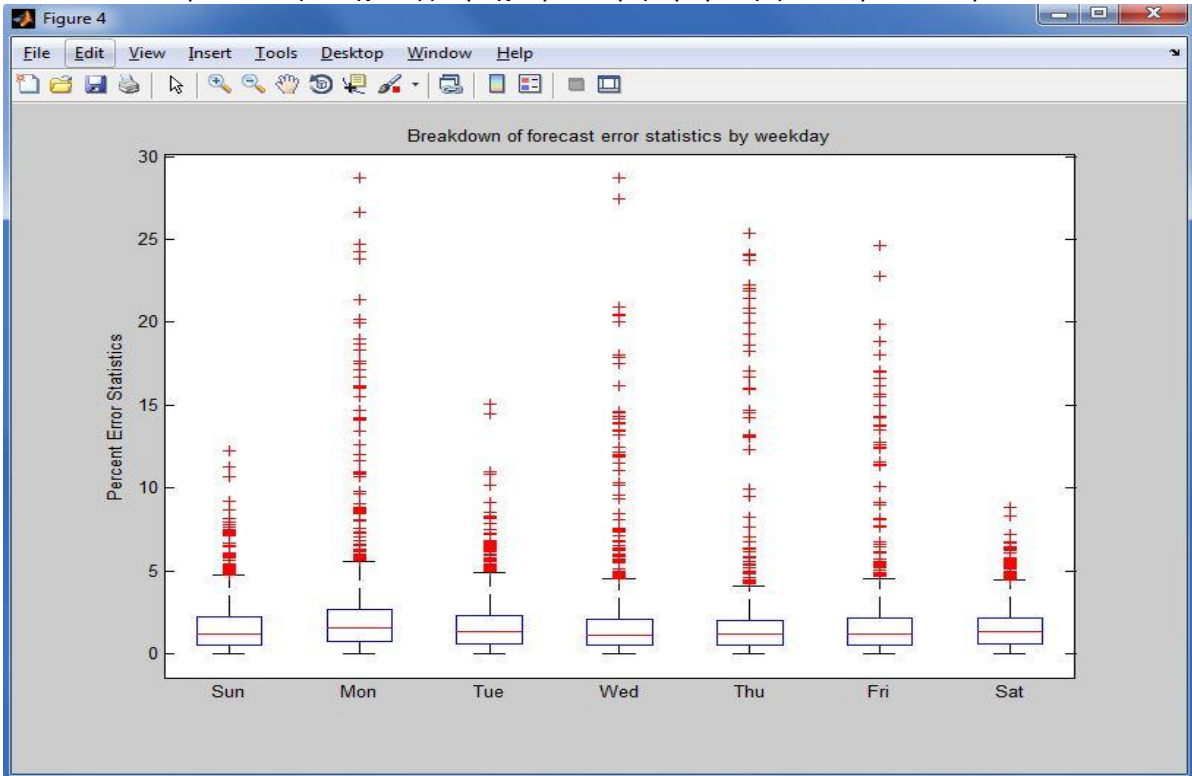
Οι παρακάτω γραφικές παραστάσεις γίνονται για περαιτέρω γνώσεις σχετικά με την απόδοση του συστήματος στις προβλέψεις, είναι δυνατόν να απεικονιστούν τα ποσοστιαία σφάλματα με την ώρα της ημέρας, ημέρα της εβδομάδας και μήνα του έτους. Αυτά τα διαγράμματα ονομάζονται box plots γιατί κάθε κουτί εκπροσωπεί μια ομάδα παρατηρήσεων που έχουν κάποιο κοινό χαρακτηριστικό. Πολύ σημαντική



Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
πληροφορία αποτελούν οι outliers τιμές. Με αυτές τις τιμές βγαίνουν κάποια σημαντικά πορίσματα, όπου η υπηρεσία κοινής ωφελείας μπορεί να δράσει για να τα αντιμετωπίσει τα σφάλματα του συστήματος.

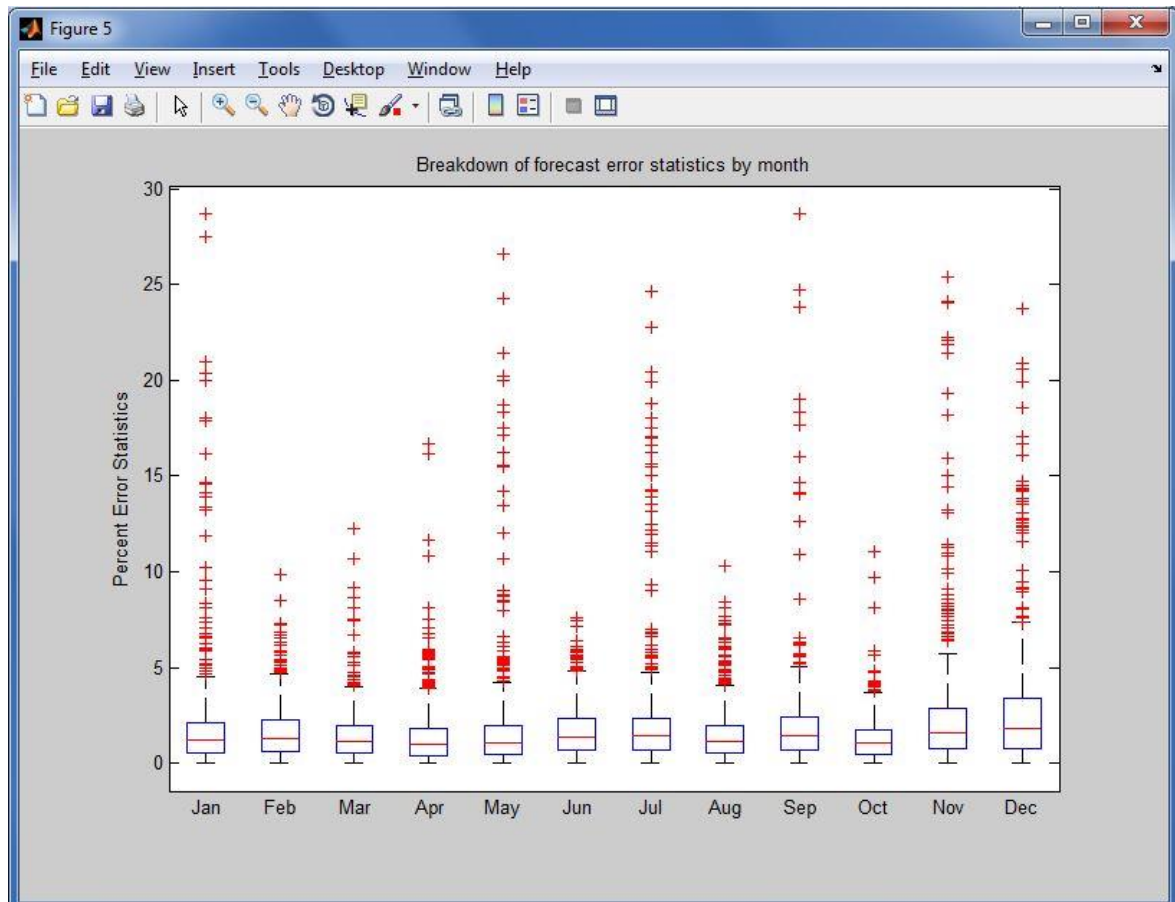


Μια σημαντική πληροφορία που παρατηρείται εδώ είναι ότι στις 7-8 το πρωί γίνονται τα πιο πολλά λάθη στην πρόβλεψη του φορτίου, γιατί εκείνες τις ώρες ο κόσμος προετοιμάζεται για να πάει στην δουλειά για να εργαστεί. Επίσης τα νοικοκυριά είναι σε λειτουργία ώστε να καλύψουν τις ανάγκες των ανθρώπων, όμως τα σαββατοκύριακα & διακοπές ( δηλαδή τις μη εργάσιμες μέρες), ο κόσμος δεν είναι υποχρεωμένος να ετοιμαστεί εκείνες τις ώρες. Επειδή δεν γίνεται σωστός διαχωρισμός των εργάσιμων και μη ημερών, παρουσιάζονται οι outliers τιμές.



Η κύρια παρατήρηση στην ομαδοποίηση ανά βδομάδα είναι ότι τις μη εργάσιμες μέρες είναι πιο εύκολο να προληφθεί το φορτίο διότι η κατανάλωση του φορτίου είναι πιο εύκολη υπολογίσιμη εκείνες τις μέρες.

01/01	New Year's Day
19/01	Martin Luther King Day
31/05	Memorial Day
05/07	Independence Day
06/09	Labor Day
25/11	Thanksgiving
26/11	Day After Thanksgiving
24/12	Christmas



Στην τελευταία παρακάτω γραφική παράσταση οι outliers τιμές οφείλονται στις απρόβλεπτες εορταστικές ημέρες. Για παράδειγμα όπως έχει προαναφερθεί, ο παρακάτω πίνακας δείχνει τις πιο σημαντικές εθνικές εορτές των ΗΠΑ. Έτσι η υπηρεσία προετοιμάζεται καταλλήλως ώστε να μην υπάρξουν προβλήματα εκείνες τις μέρες.

#### 4.7.7 Σύνοψη Παρουσίαση Αποτελεσμάτων.

Το τελευταίο κομμάτι του συστήματος, είναι η παρουσίαση των αποτελεσμάτων της εκπαίδευσης και απόδοσής του. Στην ουσία το τμήμα είναι η πρόσοψη του συστήματος, γιατί είναι ευπαρουσίαστα και ενόητα τα αποτελέσματα που παρουσιάζονται στους αγοραστές ηλεκτρικής ενέργειας και υπηρεσιών κοινής ωφέλειας. Μέσα από διαγράμματα και γραφικές παραστάσεις μπορούν να παρθούν σημαντικές αποφάσεις για την πορεία που θα ακολουθήσουν οι υπηρεσίες.



## 4.8 Γενικό Συμπέρασμα

Ακριβή μοντέλα πρόβλεψης για την ισχύς ηλεκτρικού φορτίου είναι απαραίτητα για την λειτουργία και τον προγραμματισμό μιας εταιρείας κοινής ωφελείας. Η πρόβλεψη φορτίου βοηθά μια ηλεκτρική εταιρεία στη λήψη σημαντικών αποφάσεων, συμπεριλαμβανομένων των αποφάσεων για την αγορά και την παραγωγή ηλεκτρικής ενέργειας, μεταγωγή φορτίου, και τις υποδομές ανάπτυξης.

Η Βραχυπρόθεσμη πρόβλεψη φορτίου μπορεί να βοηθήσει στην εκτίμηση των ροών φορτίου και για τη λήψη αποφάσεων που μπορεί να αποτρέψει την υπερφόρτωση. Έγκαιρες υλοποιήσεις τέτοιων αποφάσεων οδηγούν στη βελτίωση της αξιοπιστίας του δικτύου και στις μειωμένες εμφανίσεις στις βλάβες του εξοπλισμού και διακοπές ρεύματος.

Οι περισσότεροι μέθοδοι πρόβλεψης χρησιμοποιούν στατιστικές τεχνικές ή αλγορίθμους τεχνητής νοημοσύνης, όπως τα δέντρα αποφάσεων, νευρωνικά δίκτυα, ασαφής λογική, και έμπειρα συστήματα. Η ανάπτυξη και η βελτίωση της κατάλληλων μαθηματικών εργαλείων, θα οδηγήσουν στην ανάπτυξη πιο ακριβών τεχνικών πρόβλεψη φορτίου.

Η ακριβής πρόβλεψη φορτίου είναι πολύ σημαντική για εταιρείες παραγωγής ηλεκτρικής ενέργειας σε ανταγωνιστικό περιβάλλον που δημιουργείται από την απελευθέρωση της ηλεκτρική βιομηχανία. Στη παρούσα εργασία εξετάστηκαν κάποιες στατιστικές και τεχνικές τεχνητής νοημοσύνης που χρησιμοποιούνται για την πρόβλεψη ηλεκτρικού φορτίου. Αναφέρθηκαν, επίσης, παράγοντες που επηρεάζουν την ακρίβεια των προβλέψεων, όπως τα καιρικά δεδομένα, παράγοντες του χρόνου, κατηγορίες πελατών, και παράγοντες της οικονομικής και τελικής-χρήσης. Μέθοδοι πρόβλεψης φορτίου χρησιμοποιούν μοντελοποίηση προηγμένων μαθηματικών. Επιπλέον πρόοδο στην πρόβλεψη φορτίου και τη χρήση του σε βιομηχανικές εφαρμογές μπορεί να επιτευχθεί με την παροχή βραχυπρόθεσμων προβλέψεων φορτίου υπό τη μορφή των κατανομών πιθανοτήτων παρά των προβλεπόμενων αριθμών, για παράδειγμα η λεγόμενη προσέγγιση συνόλου μπορεί να χρησιμοποιηθεί. Θεωρείτο ότι η πρόοδος για την πρόβλεψη του φορτίου θα επιτευχθεί σε δύο κατευθύνσεις: (i) τη βασική έρευνα στον τομέα της στατιστικής και της τεχνητής νοημοσύνης και (ii) την καλύτερη κατανόηση της δυναμικής φορτίου και στατιστικές ιδιότητες της να εφαρμόσει τα κατάλληλα μοντέλων.



## 5. Βιβλιογραφία & Πηγές

---

### Ηλεκτρικό δίκτυο & Σύστημα διαχείρισης Διανομής

- [IEEE Power & Energy Society](#)
- [IEC-International Standards and Conformity Assessment for all electrical, electronic and related technologies](#)
- [Smart Grid - The New and Improved Power Grid: A Survey](#)

### Πρόβλεψη Φορτίου

- Eugene A. Feinberg, Dora Genethliou State University of New York, Stony Brook *Load Forecasting* <http://www.almozg.narod.ru/bible/lf.pdf>

### Δεδομένα

- Zonal hourly data (Excel Spreadsheets)  
[http://iso-ne.com/markets/hstdata/znl\\_info/hourly/index.html](http://iso-ne.com/markets/hstdata/znl_info/hourly/index.html)

### Νευρωνικά Δίκτυα

- Neural Networks Toolbox, Matlab 2012a

### Τυχαίο Δάσος

- Breiman, L. *Bagging Predictors*. Machine Learning 26, pp. 123–140, 1996
- Breiman, L. *Random Forests*. Machine Learning 45, pp. 5–32, 2001.
- Breiman, L. <http://www.stat.berkeley.edu/~breiman/RandomForests/>
- Breiman, L., et al. *Classification and Regression Trees*. Chapman & Hall, Boca Raton, 1993.
- Statistics Toolbox, Matlab 2012a

### Support Vector Machine (SVM)

- LIBSVM -- A Library for Support Vector Machines  
Chih-Chung Chang and [Chih-Jen Lin](#)
- Hsu, Chih-Wei, Chih-Chung Chang, and Chih-Jen Lin. *A Practical Guide to Support Vector Classification*. Available at <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>.
- Bottou, L., and Chih-Jen Lin. *Support Vector Machine Solvers*. Available at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.64.4209&rep=rep1&type=pdf>.
- Double Resampling  
[http://www.ece.umn.edu/users/cherkass/ee4389/double\\_resampling.pdf](http://www.ece.umn.edu/users/cherkass/ee4389/double_resampling.pdf)



- Mitchell, T. (1997). *Machine Learning*, McGraw Hill.

## 6. Παράρτημα (Appendix)

---

**Παράρτημα α):** Κώδικας για την δοκιμή διαφόρων τιμών για την μεταβλητή που καθορίζει τον ελάχιστο αριθμό δεδομένων ανά φύλλο.

```
Load fisheriris
leafs = logspace(1,2,10); % δημιουργεί τις ελάχιστες θέσεις
που μπορούν να συντελέσουν ένα φύλλο μεταξύ των τιμών 10 κ
100.
N = numel(leafs);
err = zeros(N,1);
for n=1:N
    t = ClassificationTree.fit(meas, species, 'crossval',
'on', 'minleaf',leafs(n));
    err(n) = kfoldLoss(t);
end
plot(leafs,err);
xlabel('Min Leaf Size');
ylabel('cross-validated error');
%δημιουργείται cross validated classification δέντρο.
```

**Παράρτημα 1:** Κώδικας εισαγωγής Δεδομένων ( import.m)

```
folder =
'C:\MATLAB\R2012a\bin\LoadPriceForecastingFileExchange\Electri
cityLoad&PriceForecasting\Data';

sheetname = 'ISONE CA';

% Import data for 2004 | SVHSTHKE H METAVLHTH yr META TO
FOLDER
if strcmp(sheetname, 'ISONE CA')
    NEData = dataset('XLSFile',
sprintf('%s\2004_smd_hourly.xls',folder), 'Sheet', 'NEPOOL');
else
    NEData = dataset('XLSFile',
sprintf('%s\2004_smd_hourly.xls',folder), 'Sheet',
sheetname);
end
% Add a column 'Year'
NEData.Year = 2004 * ones(length(NEData),1);

% Import data for other years
for yr = 2005:2008

    % Read in data into a dataset array
```





Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

```
x = dataset('XLSFile',  
sprintf('%s\\%d_smd_hourly.xls', folder, yr), 'Sheet',  
sheetname);
```

```
% Add a column 'Year'  
x.Year = yr*ones(length(x),1);  
  
% Concatenate the datasets together  
NEData = [NEData; x];  
end  
  
% Calculate numeric date  
NEData.NumDate = datenum(NEData.Date, 'dd/mm/yyyy') +  
(NEData.Hour-1)/24;  
  
save([folder '\' genvarname(sheetname) '_Data.mat'],  
'NEData');  
  
%Import Holidays Dates  
[num, text] =  
xlsread('C:\MATLAB\R2012a\bin\LoadPriceForecastingFileExchange  
\ElectricityLoad&PriceForecasting\Data\Holidays.xls');  
holidays = text(2:end,1);
```

**Παράρτημα 2:** Κώδικας Χωρισμού σε σετ δεδομένων εκπαίδευσης και επαλήθευσης.

```
% Create training set  
trainInd = NEData.NumDate < datenum('01/01/2008');  
trainX = X(trainInd,:);  
trainY = NEData.SYSLoad(trainInd);  
  
% %Create test set and save for later  
testInd = NEData.NumDate >= datenum('01/01/2008');  
testX = X(testInd,:);  
testY = NEData.SYSLoad(testInd);  
testDates = dates(testInd);
```

**Παράρτημα 3:** Κώδικας παραγωγής των νέων προγνωστικών παραμέτρων.

```
% Create Predictors  
  
% Short term forecasting inputs  
% Lagged load inputs  
prevDaySameHourLoad = [NaN(24,1); data.SYSLoad(1:end-24)];  
prevWeekSameHourLoad = [NaN(168,1); data.SYSLoad(1:end-168)];  
prev24HrAveLoad = filter(ones(1,24)/24, 1, data.SYSLoad);  
%prev24HrAveLoad = filter(ones(1,24)/24, 1, [NaN(24,1);  
data.SYSLoad(1:end-24)]);  
  
% Date predictors  
dayOfWeek = weekday(dates);
```



## Τμήμα Μηχανικών Πληροφορικής

### Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

```
% Non-business days
isWorkingDay = ~ismember(floor(dates), holidays) &
~ismember(dayOfWeek, [1 7]);
%[,~,isWorkingDay] = createHolidayDates(data.NumDate);

if strcmpi(term, 'long', 4);
    % Long Term Forecast Predictors
    X = [data.DryBulb data.DewPnt daily5dayHighAve
daily5dayLowAve data.Hour dayOfWeek isWorkingDay];
    labels = {'DryBulb', 'DewPoint', 'Prev5DayHighAve',
'Prev5DayLowAve', 'Hour', 'Weekday', 'IsWorkingDay'};
else
    % Short Term
    X = [data.DryBulb data.DewPnt data.Hour dayOfWeek
isWorkingDay prevWeekSameHourLoad prevDaySameHourLoad
prev24HrAveLoad];
    labels = {'DryBulb', 'DewPoint', 'Hour', 'Weekday',
'IsWorkingDay', 'PrevWeekSameHourLoad', 'prevDaySameHourLoad',
'prev24HrAveLoad'};
end
```

#### Παράρτημα 4: Κώδικας του μοντέλου του νευρωνικού δικτύου.

```
reTrain = false;
if reTrain || ~exist('Models\NNModel.mat', 'file')
    net = newfit(trainX', trainY', 20);
    net.performFcn = 'mae';
    net = train(net, trainX', trainY');
    save Models\NNModel.mat net
else
    load Models\NNModel.mat
end
```

#### Παράρτημα 5: Προσδιορισμός αριθμού των δένδρων και κατάλληλου μεγέθους Φύλλου

```
oobError = [];
leafSizes = [10 20 40 50];
for i = 1:length(leafSizes)
    model = TreeBagger(20, trainX, trainY, 'method',
'regression', ...
'oobpred', 'on', 'minleaf',
leafSizes(i));
    oobError = [oobError model.oobError];

    figure(1), plot(oobError);
    xlabel('Number of grown trees'), ylabel('Out-of-bag
Regression Error');
    title(sprintf('Regression Error versus Number of Trees &
Leaf Size'));
end
```



Τμήμα Μηχανικών Πληροφορικής  
Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου  
legend(num2str(leafSizes(1:i)'), drawnow;  
end

## Παράρτημα 6: Καθορισμός σημαντικότητας των χαρακτηριστικών

```
model = TreeBagger(20, trainX, trainY, 'method', 'regression',  
...  
                'oobvarimp', 'on', 'minleaf', 30);  
  
figure(2);  
barh(model.OOBPermutedVarDeltaError);  
ylabel('Feature');  
xlabel('Out-of-bag feature importance');  
title('Feature importance results');  
set(gca, 'YTickLabel', labels)
```

**Παράρτημα 7:** παρουσιάζεται ο κώδικας για την εύρεση βέλτιστου C και γ μέσα από το Search grid cross validation. & εκπαίδευση μοντέλου χρησιμοποιώντας αυτές τις τιμές.

```
param.s = 3; % epsilon SVR  
param.C = max(train.Y) - min(train.Y); % FIX C  
param.t = 2; % RBF kernel  
param.gset = 2.^[-1:2]; % range of the gamma parameter  
param.eset = [0:3]; % range of the epsilon parameter  
param.nfold = 5; % 5-fold CV  
  
Rval = zeros(length(param.gset), length(param.eset));  
  
for i = 1:param.nfold  
    % partition the training data into the learning/validation  
    % in this example, the 5-fold data partitioning is done by the  
    following strategy,  
    % for partition 1: Use samples 1, 6, 11, ... as validation  
    samples and  
    % the remaining as learning samples  
    % for partition 2: Use samples 2, 7, 12, ... as validation  
    samples and  
    % the remaining as learning samples  
    % :  
    % for partition 5: Use samples 5, 10, 15, ... as validation  
    samples and  
    % the remaining as learning samples  
  
    data = [train.Y, train.X];  
    [learn, val] = k_FoldCV_SPLIT(data, param.nfold, i);  
    lrndata.X = learn(:, 2:end);  
    lrndata.y = learn(:, 1);  
    valdata.X = val(:, 2:end);  
    valdata.y = val(:, 1);  
  
    for j = 1:length(param.gset)  
        param.g = param.gset(j);
```



## Τμήμα Μηχανικών Πληροφορικής

### Παρουσίαση Πτυχιακής: Βραχυπρόθεσμη Πρόβλεψη Ηλεκτρικού Φορτίου

```
for k = 1:length(param.eset)
    param.e = param.eset(k);
    param.libsvm = ['-s ', num2str(param.s), ' -t ',
num2str(param.t), ...
    ' -c ', num2str(param.C), ' -g ',
num2str(param.g), ...
    ' -p ', num2str(param.e)];

    % build model on Learning data
    model = svmtrain(lrndata.y, lrndata.X, param.libsvm);

    % predict on the validation data
    [y_hat, Acc, projection] = svmpredict(valdata.y,
valdata.X, model);

    Rval(j,k) = Rval(j,k) + mean((y_hat-valdata.y).^2);
end
end

end

Rval = Rval ./ (param.nfold);

[v1, i1] = min(Rval);
[v2, i2] = min(v1);
optparam = param;
optparam.g = param.gset( i1(i2) );
optparam.e = param.eset(i2);

optparam.libsvm = ['-s ', num2str(optparam.s), ' -t ',
num2str(optparam.t), ...
    ' -c', num2str(optparam.C), ' -g ', num2str(optparam.g), ...
    ' -p ', num2str(optparam.e)];

model = svmtrain(train.Y, train.X, optparam.libsvm);
```

### Παράρτημα 8: Υπολογισμός Σφάλματος, Ποσοστού Σφάλματος, και Μέγιστου ημερήσιου Μέσου απόλυτου Ποσοστιαίου Σφάλματος

*forecastLoad* = προβλεπόμενο φορτίο & *testY* = πραγματικό φορτίο.

```
err = test.Y-forecastLoad;
errpct = abs(err)./test.Y*100;
fL = reshape(forecastLoad, 24, length(forecastLoad)/24)';
tY = reshape(test.Y, 24, length(test.Y)/24)';

peakerrpct = abs(max(tY, [], 2) - max(fL, [], 2))./max(tY, [], 2) *
100;
```