ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΠΡΟΒΛΕΨΗ ΚΑΤΑΝΑΛΩΣΗΣ ΗΛΕΚΤΡΙΚΟΥ ΦΟΡΤΙΟΥ ΜΕ ΧΡΗΣΗ ΝΕΥΡΟ-ΑΣΑΦΩΝ ΣΥΣΤΗΜΑΤΩΝ

Αντώνιος Χ. Συντζανάκης

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:

Γεώργιος Ατσαλάκης

Χανιά, Οκτώβριος 2010

Ευχαριστίες

Για την εκπόνηση της παρούσας διπλωματικής εργασίας, θα ήθελα να ευχαριστήσω κατά κύριο λόγο τον επιβλέποντα καθηγητή μου κ. Ατσαλάκη Γεώργιο, ο οποίος με την συμπαράσταση και τις συμβουλές του υπήρξε ο κύριος καθοδηγητής μου στην μελέτη αυτή.

Πίνακας περιεχομένων

Πρόλογος

Από την αρχή της δημιουργίας των συστημάτων ηλεκτρικής ενέργειας, κρίθηκε απαραίτητη η προσέγγιση των μελλοντικών ζητούμενων φορτίων. Αυτή η ανάγκη πήγαζε κυρίως από το γεγονός ότι, όταν παράγεται η ηλεκτρική ενέργεια από κάποιον σταθμό παραγωγής, θα πρέπει να καταναλώνεται επί τόπου καθώς υπάρχει ελάχιστη δυνατότητα αποθήκευσης. Συνεπώς το κάθε σύστημα παραγωγής ηλεκτρικής ενέργειας θέλει να γνωρίζει με ικανοποιητική ακρίβεια τα μελλοντικά φορτία, ώστε η παραγόμενη ισχύς του να καλύπτει τα φορτία μαζί με τις απώλειες του δικτύου.

Μια πρώτη, αλλά σημαντική προσέγγιση αυτού του προβλήματος έγινε μέσω των κλασικών μεθόδων που περιλαμβάνουν τη χρήση πινάκων και την κατασκευή μαθηματικών συναρτήσεων, οι οποίες επεξεργάζονται ιστορικά δεδομένα και επιστρέφουν ως αποτέλεσμα την πρόβλεψη για τα μελλοντικά φορτία.

Την τελευταία δεκαετία κυρίως, έχει γίνει μια πολύ σοβαρή προσπάθεια για την αντιμετώπιση του προβλήματος με τη χρήση νέων εξελιγμένων μεθόδων, όπως η ασαφής λογική και τα τεχνητά νευρωνικά δίκτυα, μέθοδοι που βέβαια απαιτούν σημαντική υπολογιστική ισχύ.

Σκοπός της συγκεκριμένης διπλωματικής εργασίας λοιπόν, είναι κυρίως να παρουσιάσει την ανάπτυξη μεθοδολογίας για την πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας (φορτίου) με χρήση νεύρο-ασαφών συστημάτων. Στόχος της εργασίας είναι η δημιουργία ενός μοντέλου που θα παρέχει αξιόπιστες προβλέψεις αλλά και η εξερεύνηση των δυνατοτήτων που μας δίνονται στο τομέα της πρόβλεψης του ηλεκτρικού φορτίου χάρη στον συνδυασμό της ασαφούς λογικής με τα νευρωνικά δίκτυα. Στα πλαίσια της εργασίας αναπτύχθηκε κώδικας σε γλώσσα προγραμματισμού MATLAB όπου το νευρωνικό δίκτυο εκπαιδεύεται και παράγει προβλεπόμενες τιμές κατανάλωσης ηλεκτρικού φορτίου. Τα αποτελέσματα κρίνονται και εξάγονται συμπεράσματα.

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

1.1 Υπόβαθρο

Η ηλεκτρική ενέργεια είναι μία μορφή ενέργειας, η οποία συνδέεται με την παρουσία ή την κίνηση ηλεκτρικών φορτίων. Το βασικό χαρακτηριστικό της είναι ότι μπορεί εύκολα και αποδοτικά να μεταφερθεί και να μετατραπεί σε άλλες μορφές, όντας «καθαρή» στη χρήση της και διαθέσιμη όποτε της ζητηθεί. Εξαιτίας τούτων η ηλεκτρική ενέργεια παρουσιάζει συνεχώς αυξανόμενη διείσδυση, στην κάλυψη των αναγκών της ανθρώπινης κοινωνίας. Παράλληλα άρχισαν να γεννιούνται, να αναπτύσσονται και να εξελίσσονται τα συστήματα ηλεκτρικής ενέργειας, που έχουν ως βασικό τους στόχο να εξυπηρετήσουν τις ανάγκες των καταναλωτών, με όσον το δυνατόν λιγότερο κόστος και μεγαλύτερο επίπεδο αξιοπιστίας και ποιότητας υπηρεσιών. Αυτό έχει ως συνέπεια να δημιουργηθούν μια σειρά από προβλήματα, που σχετίζονται με το σχεδιασμό, τη κατασκευή και τη λειτουργία των συστημάτων. Ορισμένα από τα βασικότερα είναι τα ακόλουθα:

- ο βέλτιστος σχεδιασμός των συστημάτων παραγωγής, μεταφοράς και διανομής.
- η μελέτη και η κατασκευή των διαφόρων στοιχείων του συστήματος, όπως γεννητριών, αγωγών, μετασχηματιστών κτλ. ,
- η πρόβλεψη φορτίου σε βραχυχρόνιο και μακροχρόνιο χρονικό επίπεδο.
- ο προσδιορισμός κατανομής ισχύος στις διάφορες μονάδες για οικονομική λειτουργία του συστήματος.
- η ανάλυση μόνιμης και μεταβατικής ευστάθειας του συστήματος για ομαλή παροχή ισχύος.
- ο προσδιορισμός της ροής φορτίου στα στοιχεία του δικτύου για φόρτιση σε επιτρεπτά όρια.
- η αξιοπιστία στην παροχή και στη ποιότητα ηλεκτρικής ισχύος στον καταναλωτή.

Τα θέματα προς επίλυση στα πραγματικά ηλεκτρικά δίκτυα, αποτελούν τμήματα των προβλημάτων που θίχτηκαν, αλλά και άλλων που σχετίζονται με τα συστήματα ενέργειας, όπως περιβαλλοντικά προβλήματα κτλ. Δηλαδή δεν περιορίζονται μόνο στην εύρεση τάσεων και ρευμάτων, χρησιμοποιώντας τους νόμους Kirchhoff, αλλά εκτείνονται σε μεγάλο εύρος. Το θέμα με το οποίο θα ασχοληθούμε στην παρούσα διπλωματική εργασία είναι η πρόβλεψη φορτίου με χρήση νεύρο-ασαφών τεχνικών.

1.2 Ανάγκη Για Πρόβλεψη Φορτίου

Για να κατανοήσουμε καλύτερα τη χρησιμότητα της πρόβλεψης και γενικά της ηλεκτρικής ενέργειας, είναι σκόπιμο να γνωρίζουμε το βασικό της χαρακτηριστικό, που είναι η παροχή της, τη στιγμή της ζήτησης και η δυνατότητα της αποθήκευσης της. Η διαθεσιμότητα της ηλεκτρικής ενέργειας, ανά πάσα στιγμή στο σημείο της ζήτησης πρέπει επιπλέον να πληροί τις κατάλληλες προϋποθέσεις, ούτως ώστε να καλύπτονται οι απαιτήσεις του καταναλωτή, σε ποιότητα – επίπεδα τάσης, συχνότητας, συντελεστή παραμόρφωσης αρμονικών κτλ. και σε ποσότητα – επάρκεια ισχύος και ενέργειας. Δεδομένων των σημερινών κοινωνικών συνθηκών που επικρατούν στις αναπτυγμένες τεχνολογικά χώρες και της τεράστιας ανάγκης για ηλεκτρική ενέργεια 24 ώρες το 24ωρο σε ολόκληρο το πλανήτη, γίνεται αντιληπτό το γεγονός ότι με κάποιους τρόπους, πρέπει η ζήτηση της ηλεκτρικής ενέργειας να καλύπτεται στο παρόν και στο μέλλον. Αναπτύχθηκαν λοιπόν κατάλληλες μεθοδολογίες πρόβλεψης του φορτίου (της ηλεκτρικής ενέργειας που θα ζητάνε στο μέλλον οι καταναλωτές), ώστε το εκάστοτε δίκτυο παροχής ενέργειας να μπορεί να καλύπτει τις ανάγκες του και να είναι αξιόπιστο. Αυτές οι μέθοδοι διακρίνονται στις ακόλουθες τέσσερις κατηγορίες:

- την πολύ βραχυπρόθεσμη
- την βραχυπρόθεσμη
- την μεσοπρόθεσμη
- την μακροπρόθεσμη

Αναλύοντας την κάθε κατηγορία ξεχωριστά έχουμε:

Η *πολύ βραχυπρόθεσμη* πρόβλεψη φορτίου έχει ως στόχο, την κάλυψη των αναγκών των συστημάτων αυτομάτου ελέγχου των γεννητριών, τα οποία χρειάζονται όσο το δυνατό καλύτερη πληροφόρηση για τις μελλοντικές αλλαγές φόρτισης τους, ώστε να προχωρήσουν στις αναγκαίες μεταβολές της παροχής καυσίμου και ρύθμιση των επιπέδων τάσης εξόδου τους. Με αυτό τον τρόπο είναι δυνατός ο έλεγχος της σχέσης συχνότητας –φορτίου και της ασφάλειας του συστήματος.

Η *βραχυπρόθεσμη* πρόβλεψη έχει ως στόχο, την ρύθμιση των βασικών επιπέδων λειτουργίας των μονάδων, όπως επίσης παίζει σημαντικό ρόλο στη διαμόρφωση της σειράς ένταξης τους με βάση τα κριτήρια της οικονομικής κατανομής και λαμβάνοντας υπόψη τα αντίστοιχα προγράμματα συντήρησης. Βασικά είναι υπεύθυνη για την ενεργειακή διαχείριση του συστήματος.

Η *μεσοπρόθεσμη* πρόβλεψη αφορά κυρίως τη ρύθμιση των προγραμμάτων συντήρησης των μονάδων παραγωγής και αξιοποίησης των διαθέσιμων υδάτινων και λοιπών πόρων στις καταλλληλότερες χρονικές περιόδους του έτους. Παραδείγματος χάρη, στο ηπειρωτικό ελληνικό σύστημα, οι θερμικές μονάδες υφίστανται κατά κύριο λόγο τη μια ή τις δύο ετήσιες συντηρήσεις τους συνήθως σε μη καλοκαιρινούς μήνες, ώστε να καλυφθεί η ανάγκη της ετήσιας αιχμής που εμφανίζεται κατά το καλοκαίρι και οφείλεται κατά κύριο λόγο στη χρήση κλιματιστικών, όπως μας δίνεται από τις αντίστοιχες προβλέψεις.

Η *μακροπρόθεσμη* πρόβλεψη φορτίου αφορά κυρίως την ετήσια αιχμή φορτίου –ή και τις εποχιακές –και τη συνολική ετήσια ενέργεια για τα επόμενα 10 ή 20 χρόνια, ώστε να είναι δυνατός ο σχεδιασμός και η κατασκευή του συστήματος ηλεκτρικής ενέργειας, όπως η δημιουργία νέων μονάδων παραγωγής, η κατασκευή νέων γραμμών μεταφοράς ή η αναδιαμόρφωση του δικτύου διανομής μιας μεγάλης περιοχής, π.χ. η ανύψωση του δικτύου μέσης τάσης στα 20kV από τα 15kV για τη περιοχή της Θεσσαλονίκης ή η κατασκευή του υδροηλεκτρικού σταθμού στο Θησαυρό, δηλαδή έργων αρκετά χρονοβόρων και δαπανηρών.
Τα προαναφερθέντα χαρακτηριστικά της κάθε κατηγορίας καθώς και χρονικός ορίζοντας πρόβλεψης στην κάθε περίπτωση, συνοψίζονται στον παρακάτω πίνακα.

Είδος Πρόβλεψης	Χρονικός Ορίζοντας	Λειτουργίες για τις οποίες χρειάζεται
Πολύ βραχυπρόθεσμη	μερικά sec - μερικά min	Αυτόματος Έλεγχος Παραγωγής Οικονομική Κατανομή Φορτίου
Βραχυπρόθεσμη	μερικές h - 1 εβδομάδα	Ένταξη Μονάδων Υδροθερμική Συνεργασία Ανταλλαγές Ισχύος
Μεσοπρόθεσμη	μερικοί μήνες - 5 έτη	Προγραμματισμός Καυσίμων Χρήση Υδάτινων Πόρων Προγραμματισμός Συντήρησης
Μακροπρόθεσμη	10 - 25 έτη	Προγραμματισμός Ένταξης Νέων Μονάδων

Πίνακας 1.1: *Χαρακτηρισμός προβλέψεων φορτίου ανάλογα με το χρονικό ορίζοντα.*

Είναι σκόπιμο να πούμε εδώ ότι στη βιβλιογραφία για τις μεθόδους πρόβλεψης φορτίου υπάρχει μια σύγχυση, όσον αφορά τους όρους «μεσοπρόθεσμη» και «μακροπρόθεσμη». Αυτό συμβαίνει γιατί και οι δυο όροι αναφέρονται σε πρόβλεψη για μεγάλο χρονικό διάστημα, όποτε για άλλους οι δυο όροι ταυτίζονται και για άλλους αποτελούν ξεχωριστές μεθόδους.

1.3 Παράγοντες Μεταβολής Του Φορτίου

Γενικά, το φορτίο μιας ηλεκτρικής χρήσης απαρτίζεται από πολύ διαφορετικές καταναλωτικές μονάδες. Ένα μεγάλο μέρος του ηλεκτρισμού καταναλώνεται σε βιομηχανικές δραστηριότητες. Ένα άλλο μέρος φυσικά χρησιμοποιείται από ιδιώτες με σκοπό τη θέρμανση, το φωτισμό, το μαγείρεμα κ.ά. Επίσης, πολλές κοινωφελείς υπηρεσίες απαιτούν ηλεκτρισμό, π.χ. ο φωτισμός των δρόμων, ο ηλεκτρικός σιδηρόδρομος κ.α. Οι παράγοντες που επηρεάζουν το φορτίο εξαρτώνται από τις επιμέρους μονάδες κατανάλωσης. Το βιομηχανικό φορτίο συνήθως καθορίζεται περισσότερο από το επίπεδο της παραγωγής. Το φορτίο είναι συνήθως αρκετά σταθερό και είναι δυνατό να εκτιμηθεί η εξάρτηση του σε διαφορετικά επίπεδα παραγωγής. Παρ' όλα αυτά, από την σκοπιά της πώλησης ενέργειας, η βιομηχανική χρήση συχνά προσθέτει αβεβαιότητα στις προβλέψεις. Το πρόβλημα είναι η πιθανότητα απροσδόκητων συμβάντων, όπως μηχανικές βλάβες ή απεργίες, οι οποίες μπορούν να προκαλέσουν μεγάλες απρόβλεπτες διακυμάνσεις στο επίπεδο φορτίου. Στην περίπτωση των ιδιωτών, οι παράγοντες που επηρεάζουν το φορτίο είναι πολύ πιο δύσκολο να καθοριστούν. Κάθε άνθρωπος συμπεριφέρεται με τον δικό του ιδιαίτερο τρόπο και η ανθρώπινη ψυχολογία σχετίζεται με κάθε καταναλωτική απόφαση. Πολλοί κοινωνικοί παράγοντες αλλά και παράγοντες συμπεριφοράς μπορούν να παρατηρηθούν. Για παράδειγμα μεγάλα γεγονότα, αργίες ακόμη και τηλεοπτικά προγράμματα επηρεάζουν το φορτίο . Ο καιρός είναι ο πιο σημαντικός παράγοντας κι αυτό εξαιτίας της ηλεκτρικής θέρμανσης των σπιτιών, η οποία γίνεται πιο έντονη όσο μειώνεται η θερμοκρασία .

Γενικά, θα λέγαμε ότι οι παράγοντες οι οποίοι επηρεάζουν το ηλεκτρικό φορτίο μπορούν να χωριστούν σε τέσσερις κατηγορίες:

- Οικονομικοί
- Χρονικοί
- Καιρικοί
- Τυχαίοι

Για τη δημιουργία ενός μοντέλου που περιγράφει το φορτίο θα πρέπει να γίνει κατανοητή η επίδραση όλων των παραγόντων στη διαμόρφωση της ζήτησης. Θα κάνουμε τώρα μια σύντομη αναφορά στους προαναφερθείς παράγοντες που επηρεάζουν τα φορτία, δηλαδή τη ζήτηση της ηλεκτρικής ενέργειας από τους καταναλωτές.

1) ***Οικονομικοί παράγοντες***, δηλαδή η οικονομική κατάσταση, το οικονομικό περιβάλλον και οι ρυθμοί ανάπτυξης κάποιας περιοχής επηρεάζουν καθοριστικά τη ζήτηση της ηλεκτρικής ενέργειας. Αυτό συμβαίνει, διότι αλλάζει το βιοτικό επίπεδο των καταναλωτών, εξελίσσεται δημογραφικά ο πληθυσμός της χώρας, γίνεται αυτοματοποίηση των δραστηριοτήτων και βελτιστοποίηση της λειτουργίας των συσκευών και των μηχανών, επιδρώντας σε μεσοπρόθεσμη και μακροπρόθεσμη βάση. Κάτω από αυτούς τους παράγοντες μεταβάλλονται:

- η σύνθεση των καταναλωτών, αύξηση του αστικού πληθυσμού και μείωση του γεωργικού.
- η ποσότητα της καταναλισκόμενης ενέργειας καθώς από τη μια πλευρά βελτιώνεται το βιοτικό επίπεδο των καταναλωτών και από την άλλη υπάρχουν τα φαινόμενα κορεσμού και η χρησιμοποίηση λιγότερο ενεργοβόρων συσκευών.
- οι μέθοδοι καλλιέργειας, άρδευσης και αυτοματοποίησης της βιομηχανίας.

2) ***Εποχιακοί παράγοντες***, καθώς τα φόρτια παρουσιάζουν έντονες διακυμάνσεις κατά τη διάρκεια τόσο του 24ώρου όσο και του έτους. Οι αλλαγές αυτές οφείλονται κυρίως στις μεταβολές των καιρικών συνθηκών και τις ανθρώπινων δραστηριοτήτων. Ιδιαίτερα ο τελευταίος παράγοντας επηρεάζεται σε σημαντικό βαθμό από τη γεωγραφική περιοχή, καθώς παραδείγματος χάρη οι αστικές περιοχές κατά τη διάρκεια των καλοκαιρινών διακοπών παρουσιάζουν μείωση της ζήτησης, ενώ αντίστοιχα οι παραθαλάσσιες αύξηση.

3) ***Μετεωρολογικοί παράγοντες*** διότι η θερμοκρασία κατά κύριο λόγο – και σε αρκετά μικρότερο ποσοστό η υγρασία, η ανεμόπτωση και η ηλιοφάνεια – επιδρούν στη κατανάλωση ηλεκτρικής ενέργειας. Καθώς, η λειτουργία ενός μεγάλου ποσοστού ηλεκτρικών συσκευών (θέρμανση, κλιματισμός, αρδεύσεις) εξαρτάται από τις καιρικές συνθήκες. Σε πολλά συστήματα, η θερμοκρασία έχει τη σημαντικότερη επίδραση στη διαμόρφωση του φορτίου. Για συστήματα τα οποία καλύπτουν μεγάλη γεωγραφική έκταση χρειάζεται να ληφθούν υπόψη οι θερμοκρασίες σε διάφορες περιοχές για να υπολογιστεί η ακριβής επίδραση στο φορτίο.

4) ***Τυχαίοι παράγοντες***, σε αυτήν την κατηγορία ανήκουν όλοι οι παράγοντες που επηρεάζουν το φορτίο και δεν ανήκουν σε καμία από τις προηγούμενες κατηγορίες. Η λειτουργία του συστήματος δέχεται συνεχώς διαταραχές οι οποίες οφείλονται στο γεγονός ότι ζήτηση του φορτίου συνίσταται από ένα μεγάλο αριθμό επί μέρους ζητήσεων των καταναλωτών. Εκτός από τους χιλιάδες μικρούς καταναλωτές, υπάρχουν και μεγάλοι καταναλωτές (χαλυβουργεία κ.ά.) που προκαλούν σχετικά μεγάλες τυχαίες διαταραχές στη ζήτηση φορτίου. Υπάρχουν επίσης ορισμένα άλλα γεγονότα, όπως μεγάλες απεργίες, εκλογές, ειδικά προγράμματα στην τηλεόραση, τα οποία, αν και είναι εκ των προτέρων γνωστά, είναι αρκετά δύσκολο να εκτιμηθεί η επίδρασή τους στη διαμόρφωση της ζήτησης .

5) ***Απροσδιόριστοι παράγοντες***, δηλαδή παράγοντες που είναι πολύ δύσκολο να προβλεφθούν και επηρεάζουν μακροπρόθεσμα τη κατανάλωση, όπως οι προοπτικές ανάπτυξης μιας περιοχής, ο ρυθμός αύξησης του πληθυσμού και η πορεία της οικονομίας.

Ποιοι, όμως, από όλους αυτούς τους προαναφερθέντες παράγοντες λαμβάνονται υπόψη στη μελέτη μας; Ένα μεγάλο κομμάτι της κατανάλωσης οφείλεται σε ιδιώτες και άλλους μικρούς ηλεκτρικούς καταναλωτές. Μία συνήθης προσέγγιση της πρόβλεψης φορτίου είναι να επικεντρωθούμε στο συνολικό φορτίο. Η επίδραση των καιρικών φαινομένων φαίνεται ότι στη μακροπρόθεσμη πρόβλεψη φορτίου δεν διαδραματίζει σημαντικό ρόλο, σε αντίθεση με την βραχυπρόθεσμη που οι μετεωρολογικές συνθήκες προκαλούν μεγάλες αποκλίσεις στο συνολικό φορτίο. Ωστόσο η μετατόπιση της αιχμής του φορτίου το καλοκαίρι για το ελληνικό σύστημα λόγω της χρήσης των κλιματιστικών μπορεί να επιφέρει κάποια μεταβολή. Βέβαια σ’ αυτήν την περίπτωση πρέπει να τονιστεί ότι έγινε η διείσδυση μιας νέας συσκευής τελικής χρήσης, οπότε θα πρέπει ο μελετητής, πριν την εφαρμογή του οποιουδήποτε μοντέλου, να προδιαγράψει την αντίστοιχη επίδραση στο φορτίο, αλλιώς θα υπάρξει σε κάθε είδους πρόβλεψη σημαντική απόκλιση.

ΚΕΦΑΛΑΙΟ 2

Νευρωνικά δίκτυα-Ασαφής λογική και πρόβλεψη φορτιού

Τα τελευταία χρόνια έχει αναπτυχθεί ένα μεγάλο ενδιαφέρον σχετικά με την εφαρμογή του Ευφυούς Ελέγχου στις βιομηχανίες. Αυτό οφείλεται στην ικανοποιητική απόδοση των εφαρμογών του σε πολλούς κρίσιμους τομείς, όπως αυτός της πρόβλεψης ηλεκτρικού φορτίου. Στην αιχμή του δόρατος των μοντέλων πρόβλεψης, βρίσκονται τα νευροασαφή συστήματα αποτελούν μια υβριδική κατηγορία του Ευφυούς Ελέγχου, που έχει αναπτυχθεί σε μια προσπάθεια συνδυασμού των καλύτερων στοιχείων των δύο κατηγοριών(νευρωνίκα δίκτυα-ασαφής λογικη). Τα ασαφή και τα νευρωνικά συστήματα μοντελοποιούν τη συμπεριφορά ενός ειδικού κατά την αντιμετώπιση μιας πολύπλοκης διεργασίας. Αυτό σημαίνει ότι στόχος κατά το σχεδιασμό τέτοιων συστημάτων δεν είναι η μοντελοποίηση του ίδιου του προβλήματος κατασκευάζοντας ένα αναλυτικό μαθηματικό μοντέλο, αλλά η αξιοποίηση της γνώσης και της εμπειρίας ενός ειδικού, κατασκευάζοντας ένα σύστημα ελέγχου που θα ελέγχει τη διεργασία όπως ο ειδικός, συνδυάζοντας τα πλεονεκτήματα και των δύο μεθόδων και εξαλείφοντας κάποια από τα μειονεκτήματα τους.

Η ικανότητα εκπαίδευσης των νευρωνικών συστημάτων είναι ο κύριος λόγος συνδυασμού τους με τα ασαφή συστήματα. Αυτός ο συνδυασμός μπορεί να δημιουργήσει κανόνες για το ασαφές σύστημα ή να βελτιστοποιήσει τους ήδη υπάρχοντες, αυτοματοποιώντας τη ρύθμιση των παραμέτρων του ασαφούς συστήματος.

Όσον αφορά στα νευρωνικά συστήματα, ο συνδυασμός τους με τα ασαφή αναιρεί τη συμπεριφορά τους ως «μαύρα κουτιά» κι έτσι οποιαδήποτε εκ των προτέρων γνώση για το σύστημα μπορεί να χρησιμοποιηθεί κατά τη διαδικασία εκμάθησης περιορίζοντας το χρόνο που αυτή απαιτεί.

Παρακάτω θα δούμε νευρωνικά μοντέλα καθώς και νεύρο-ασαφή μοντέλα τα οποία έχουν εφαρμοστεί στην πρόβλεψη ηλεκτρικού φορτίου και τα οποία βρήκαμε στο διαδίκτυο στις ακόλουθες ιστοσελίδες.

www.sciencedirect.com
www.google.gr .

Ιεραρχικά νευρωνικά δίκτυα κοντινού χρονικού ορίζοντα

Το παρόν άρθρο, προτείνει ένα νέο νευρωνικό μοντέλο για τη λύση του προβλήματος της πρόβλεψης φορτίου κοντινού χρονικού ορίζοντα (STLF). Περιγραφικά, το εξεταζόμενο νευρωνικό μοντέλο αποτελείται από δυο SOM (self organizing map) δίκτυα, το ένα στην κορυφή του αλλού. Από το μοντέλο απαιτήθηκε να προβλέψει μια φορά την ώρα το ηλεκτρικό φορτίο κατά την διάρκεια του επόμενου εικοσιτετραώρου. Η ιεραρχική αυτή δομή που έχει του δίνει τη δύναμη να λειτουργεί αποδοτικά με τα δεδομένα που εγκλείονται στις εισόδους, το μοντέλο έχει καλή μνήμη από δεδομένα του παρελθόντος και αυτό το κάνει να παράγει καλά αποτελέσματα, δηλαδή προβλέψεις. Το HNM αποδίδει καλυτέρα αποτελέσματα σε ποιο σύντομο χρονικό διάστημα και η ανάπτυξη του κοστίζει λιγότερο από το μοντέλο παλινδρόμησης για αυτό και το αντικατέστησε. Το μοντέλο έχει επτά εισόδους στις οποίες εισάγονται τα φορτία του παρελθόντος. Στην πρώτη το τρέχον φορτίο, στην δεύτερη το φορτίο της αμέσως προηγουμένης ώρας και ακολούθως το φορτίο πριν 24 ώρες (3) , πριν μια εβδομάδα (4) και πριν μια εβδομάδα και 24 ώρες (5) ενώ στις εισόδους 6 και 7 έχουμε τριγωνομετρική κωδικοποίηση για την ώρα πρόβλεψης. Το μοντέλο έχει μια έξοδο την πρόβλεψη του φορτίου. Η εκπαίδευση του μοντέλου γίνεται με συνηθισμένες τιμές (συγκεκριμένου εύρους) σε κανονικές συνθήκες. Το HNM δίνει μικρότερο σφάλμα σε σχέση με το MLP συμφωνά με τα MAPE και MSE σφάλματα ενώ το αντίστροφο ισχύει για ME και MAX σφάλματα σε συνηθισμένες τιμές φορτιών. Στην περίπτωση ασυνήθιστων τιμών φορτίων το HNM δίνει χαμηλότερο σφάλμα συμφωνά και με τα 4 παραπάνω σφάλματα σε σχέση με το MLP. Στην συνέχεια παρατίθενται η σχηματική απεικόνιση του μοντέλου.

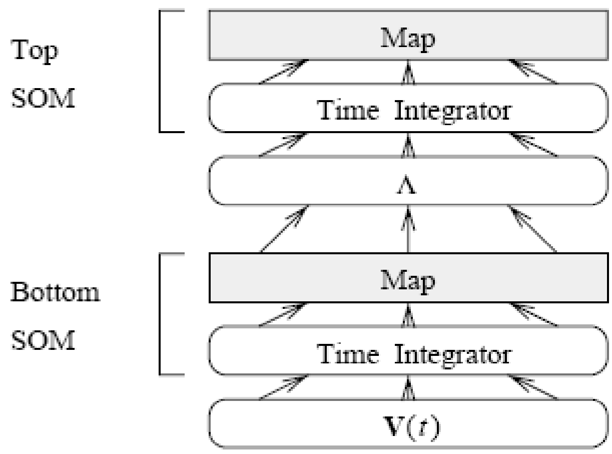


Fig. 1. HNM.

Σχήμα 2.1: *Σχηματική αναπαράσταση HNM δικτύου.*

RBF νευρωνικά δίκτυα και συνδυασμένη διόρθωση παραμένοντων λαθών για πρόβλεψη φορτίων κλιματισμού

Ο κύριος στόχος αυτού του άρθρου, είναι να παρουσιαστεί μια νέα προσέγγιση για την υψηλής ακρίβειας πρόβλεψη του ηλεκτρικού φορτίου που απαιτείται για την χρήση των κλιματιστικών. Η ακριβής πρόβλεψη κλιματισμού είναι η προϋπόθεση για τον βέλτιστο έλεγχο και την εξοικονόμηση ενέργειας των HVAC συστημάτων. Έχουν αναπτυχθεί πολλές μέθοδοι πρόβλεψης, όπως η πολλαπλή γραμμική οπισθοδρόμηση (MLR), ο αυτό-ανάδρομος κινούμενος μέσος όρος (ARIMA), το γκρίζο πρότυπο (GM) και τεχνητά νευρωνικά δίκτυα (ANN), στον τομέα της πρόβλεψης φορτίων κλιματισμού. Εντούτοις, καμία από αυτές δεν έχει αρκετή ακρίβεια για να ικανοποιήσει την πρακτική ζήτηση, έτσι βάσει αυτών των προτύπων αναπτύσσεται μια νέα μέθοδος πρόβλεψης που ονομάζεται RBF νευρωνικό δίκτυο (radial basis function neural network) με συνδυασμένη διόρθωση παραμενόντων λαθών. Το νέο πρότυπο υιοθετεί τον προηγμένο αλγόριθμο των νευρωνικών δικτύων που βασίζεται στις συναρτήσεις radial basis για την πρόβλεψη του φορτίου και χρησιμοποιεί το συνδυασμένο μοντέλο πρόβλεψης, που είναι ο συνδυασμός των MLR, του ARIMA και του GM, για να υπολογίσει τα παραμένοντα λάθη και να διορθώσει τα ολικά αποτελέσματα πρόβλεψης. Το συγκεκριμένο μοντέλο χρησιμοποιεί ως εισόδους:

- την χρονική στιγμή(ώρα)
- την εξωτερική θερμοκρασία
- την εξωτερική υγρασία
- την ηλιακή ακτινοβολία

Το μοντέλο πρόβλεψης RBFNN με την συνδυασμένη διόρθωση παραμενόντων λαθών αποδεικνύεται ότι δίνει ακριβέστερα αποτελέσματα από ότι το RBFNN μονό του ή από το RBFNN με την απλή διόρθωση . Η αρχιτεκτονική του μοντέλου απεικονίζεται στα ακόλουθα σχήματα.

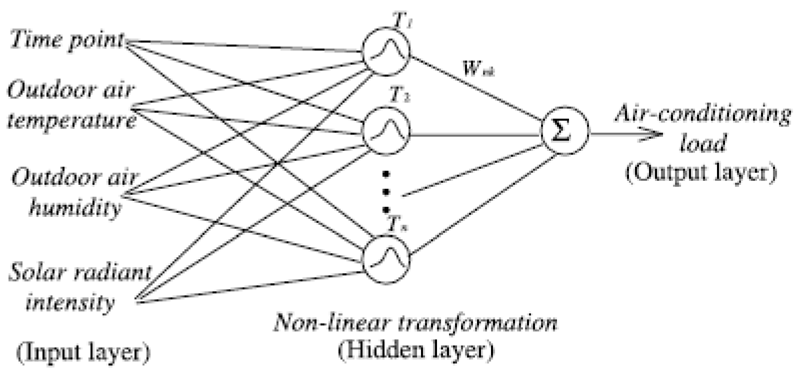


Fig. 2. Using RBF network to model air-conditioning load forecasting.

Σχήμα 2.2: *Χρησιμοποιούμενο RBF δίκτυο για την μοντελοποίηση της πρόβλεψης κλιματισμού.*

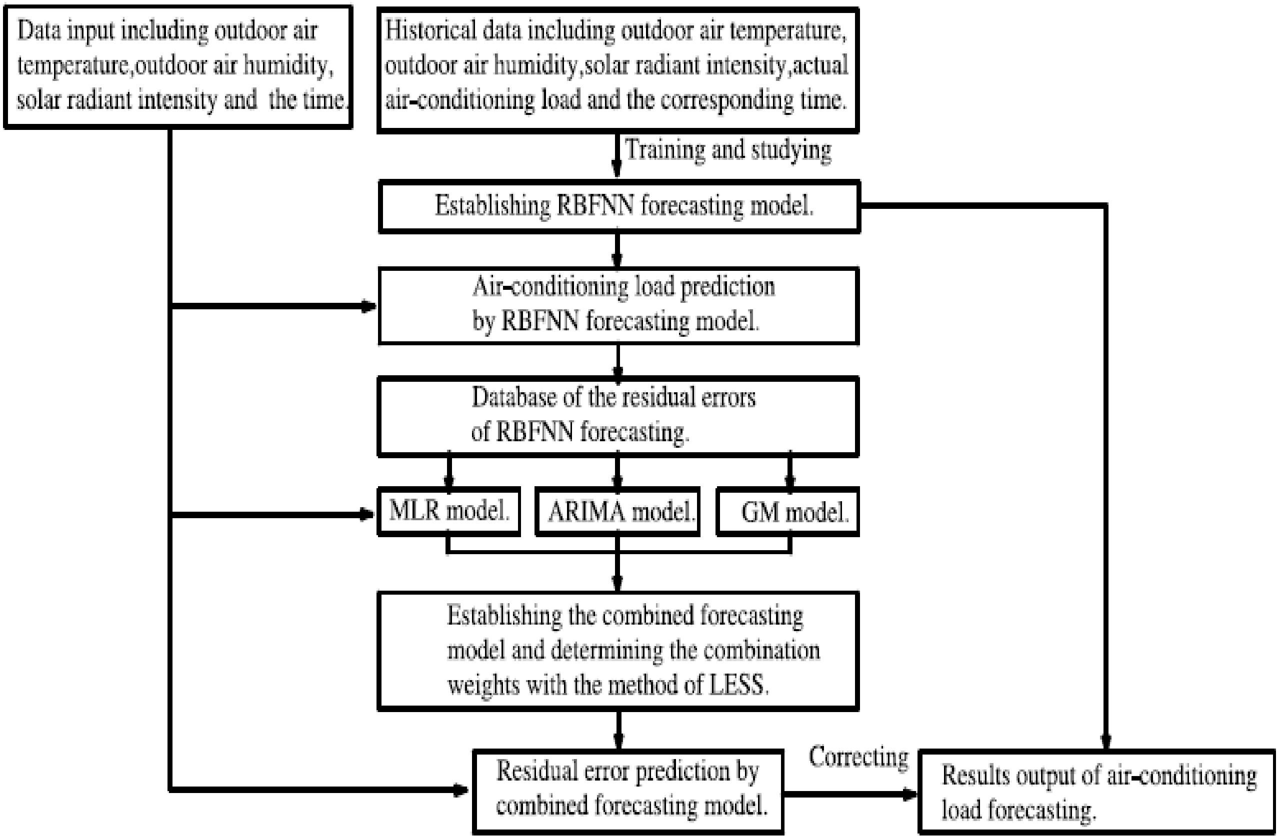


Fig. 3. Forecasting flowchart of RBFNN model with combined residual error.

<u>Σχήμα 2.3:</u> *Διάγραμμα ροής πρόβλεψης του RBFNN μοντέλου με συνδυασμένη διόρθωση παραμενόντων λαθών.*

Αλγόριθμοι προοδευτικής εκπαίδευσης για την βραχυπρόθεσμη πρόβλεψη φορτίου

Αυτό το άρθρο, παρουσιάζει την ανάπτυξη ενός βελτιωμένου νευρωνικού δικτυού, το οποίο βασίζεται σε μοντέλα βραχυπρόθεσμης πρόβλεψης του ηλεκτρικού φορτίου για το σύστημα κατανάλωσης ισχύος του νησιού της Κρήτης. Διάφορες προσεγγίσεις, όπως τα δίκτυα με συναρτήσεις radial basis (RBFN), τα δυναμικά νευρωνικά δίκτυα και τα δίκτυα ασαφούς τύπου, προτείνονται και αξιολογούνται. Τα αποτελέσματα δείχνουν ότι τα μοντέλα πρόβλεψης φορτίου που αναπτύσσονται κατ' αυτόν τον τρόπο παρέχουν ακριβέστερες προβλέψεις από τα συμβατικά μοντέλα αναστροφής διάδοσης (back-propagation). Αυτές οι μέθοδοι, αναπτύχθηκαν για πρόβλεψη του ωριαίου ηλεκτρικού φορτίου για μια ημέρα μπροστά χρησιμοποιώντας μια μορφοποιημένη αρχιτεκτονική η οποία απεικονίζεται παρακάτω.

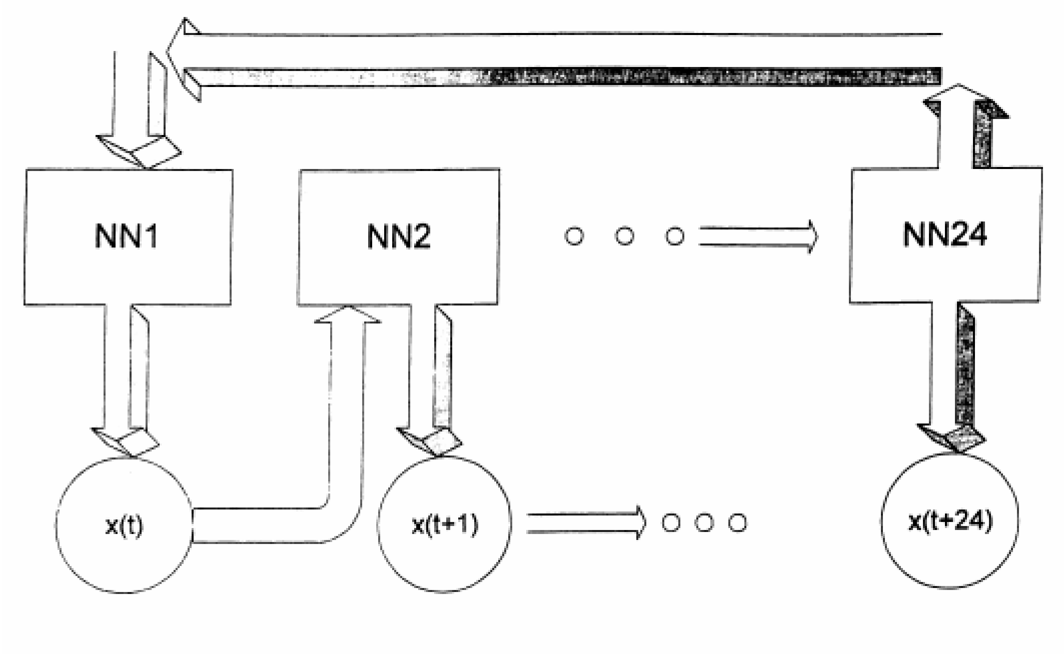


Fig. 1. Proposed modular architecture for the STLF peoblem.

Σχήμα 2.4: *Προτεινόμενη αρθρωτή αρχιτεκτονική για την επίλυση του STLF προβλήματος.*

Αξιολόγηση του νεύρο-ασαφούς ιεραρχικού BSP μοντέλου

Ο κύριος σκοπός αυτού του άρθρου, είναι να παρουσιάσει και να αξιολογήσει την απόδοση του ιεραρχικού νεύρο ασαφούς BSP-HNFB μοντέλου για ωριαία πρόβλεψη φορτίου. Για να επιτευχθεί αυτός ο στόχος, τρία συστήματα σχεδιάστηκαν και δοκιμάστηκαν: ένα σύστημα βασισμένο σε απλά νευρωνικά δίκτυα με σκοπό την σύγκριση με τις υβριδικές μεθόδους, ένα ιεραρχικό νεύρο ασαφές σύστημα και ένα υβριδικό νευρωνικό/νεύρο ασαφές σύστημα, ο σκοπός των οποίων είναι να αξιολογήσουν κατά πόσον είναι πιθανόν ένα νεύρο ασαφές σύστημα να βελτιώσει τα αποτελέσματα που προέρχονται από ένα νευρωνικό δίκτυο. Ως είσοδοι στα πειράματα χρησιμοποιούνται ιστορικές σειρές φορτίων και πρόσθετες μεταβλητές που επηρεάζουν την συμπεριφορά των φορτίων όπως η θερμοκρασία, το προφίλ της κατανάλωσης και ο δείκτης καλής συμπεριφοράς. Τα αποτελέσματα αποκαλύπτουν το δυναμικό των προτεινόμενων νεύρο-ασαφών και νευρωνικών νεύρο-ασαφών μοντέλων στην πρόβλεψη φορτίου όταν συγκρίνονται με νευρωνικά δίκτυα. Πιο συγκεκριμένα, ενώ τα νευρωνικά δίκτυα παρουσιάζουν καλή απόδοση στις προβλέψεις έχουν το μειονέκτημα να απαιτούν πολύ χρόνο για εκπαίδευση. Από την άλλη πλευρά, το προτεινόμενο ιεραρχικό νεύρο ασαφές μοντέλο βασιζόμενο στα αποτελέσματα που έχουν ήδη επιτευχθεί αποδεικνύει ότι με χρήση υβριδικών έξυπνων συστημάτων μπορούμε να επιτύχουμε πολύ καλύτερη απόδοση και με πολύ λιγότερο χρόνο εκπαίδευσης.

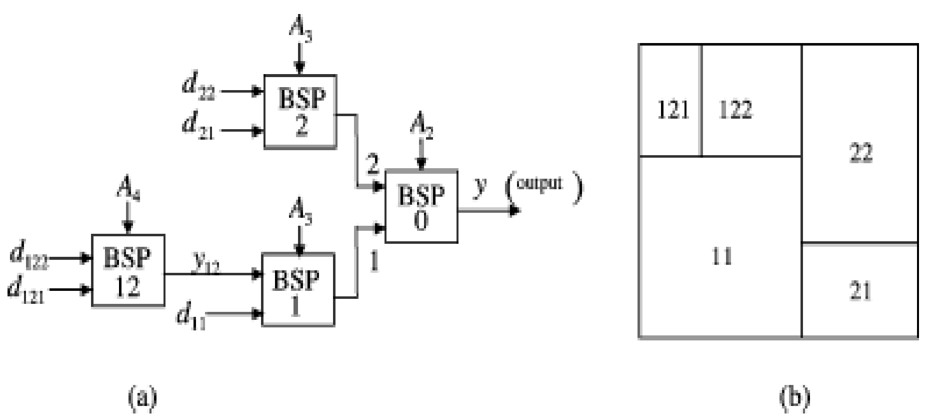


Fig. 3. (a) Example of an HNFB system. (b) input space partitioning of the HNFB system.

Σχήμα 2.5: *Παράδειγμα ενός HNFB συστήματος με τον αντίστοιχο διαχωρισμό του εισαγόμενου διαστήματος.*

Ασαφή και Νεύρο-Ασαφή Υπολογιστικά Μοντέλα για Πρόβλεψη Ηλεκτρικού Φορτίου.

Στο άρθρο αυτό, παρουσιάζονται δύο νέα υπολογιστικά μοντέλα, ένα ασαφές έμπειρο σύστημα και ένα υβριδικό νευρωνικό δίκτυο-ασαφές έμπειρο σύστημα για την πρόβλεψη χρόνο-σειρών ηλεκτρικών φορτίων. Το έμπειρο σύστημα που βασίζεται σε ασαφή λογική αξιοποιεί την ιστορική σχέση μεταξύ των φορτίων και των καιρικών συνθηκών για μια δοσμένη περίοδο και προβλέπει δίκαια και με ακρίβεια τα ηλεκτρικά φορτία σε βάθος χρόνου έως είκοσι τέσσερις ώρες μπροστά. Στην περίπτωση του υβριδικού νεύρο-ασαφούς συστήματος, το διάνυσμα εισόδου αποτελείται από τις σχέσεις μεταξύ των τιμών , των λεκτικών ιδιοτήτων που έχουν τα παρελθοντικά φορτία καθώς και των παραμέτρων του καιρού. Το διάνυσμα εξόδου καθορίζεται στις συνθήκες των σχέσεων μεταξύ των τιμών στις ασαφείς κλάσεις των φορτίων που έχουν προβλεφθεί. Τα γραμμικά και μη-γραμμικά προσαρμοζόμενα ασαφή σχέδια διορθώσεων χρησιμοποιούνται για να αυξηθεί η έξοδος από το νεύρο-ασαφές υπολογιστικό μοντέλο, έτσι ώστε να παραχθεί η τελική πρόβλεψη. Εκτεταμένες μελέτες έχουν διεξαχθεί για την πρόβλεψη φορτίων για όλες τις εποχές και τα αποτελέσματα για μια χαρακτηριστική χειμερινή μέρα δίνονται ως επιβεβαίωση της αποτελεσματικότητας αυτών των μοντέλων. Το συμπέρασμα λοιπόν που εξάγεται είναι, ότι το νεύρο ασαφές υπολογιστικό μοντέλο βρέθηκε να είναι αρκετά ακριβές σε σχέση με το ασαφές μοντέλο, καθώς οι ασαφείς διορθώσεις προέρχονται από μικρά σφάλματα στις μεταβλητές του καιρού. Επιπλέον, ο αριθμός των κανόνων που απαιτούνται στην περίπτωση ενός ασαφούς έμπειρου συστήματος είναι αρκετά μεγαλύτερος από αυτούς που απαιτούνται σε ένα υβριδικό νευρωνικό δίκτυο-ασαφές έμπειρο σύστημα, και γι’ αυτόν τον λόγο απαιτείται ένα ιεραρχικό σχέδιο έτσι ώστε να μειωθεί ο υπολογιστικός φόρτος. Αυτά τα αποτελέσματα είναι αρκετά σημαντικά σε σχέση με τις κανονικές τεχνικές πρόβλεψης ή τις ANN τεχνικές, που παρουσιάστηκαν νωρίτερα.

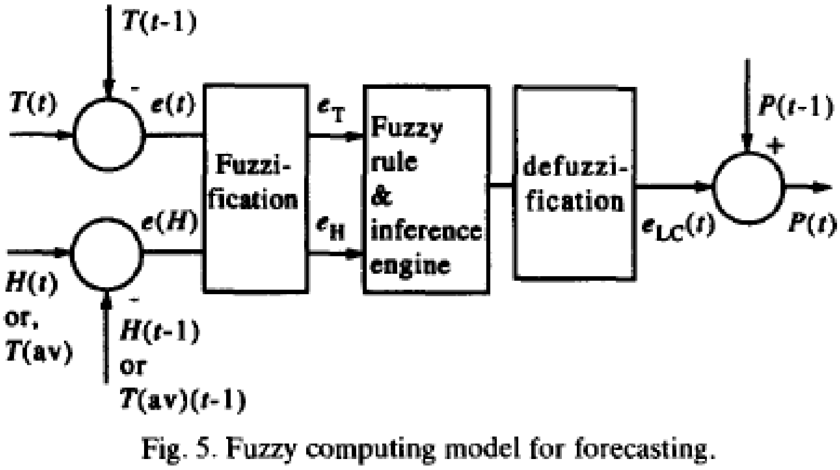
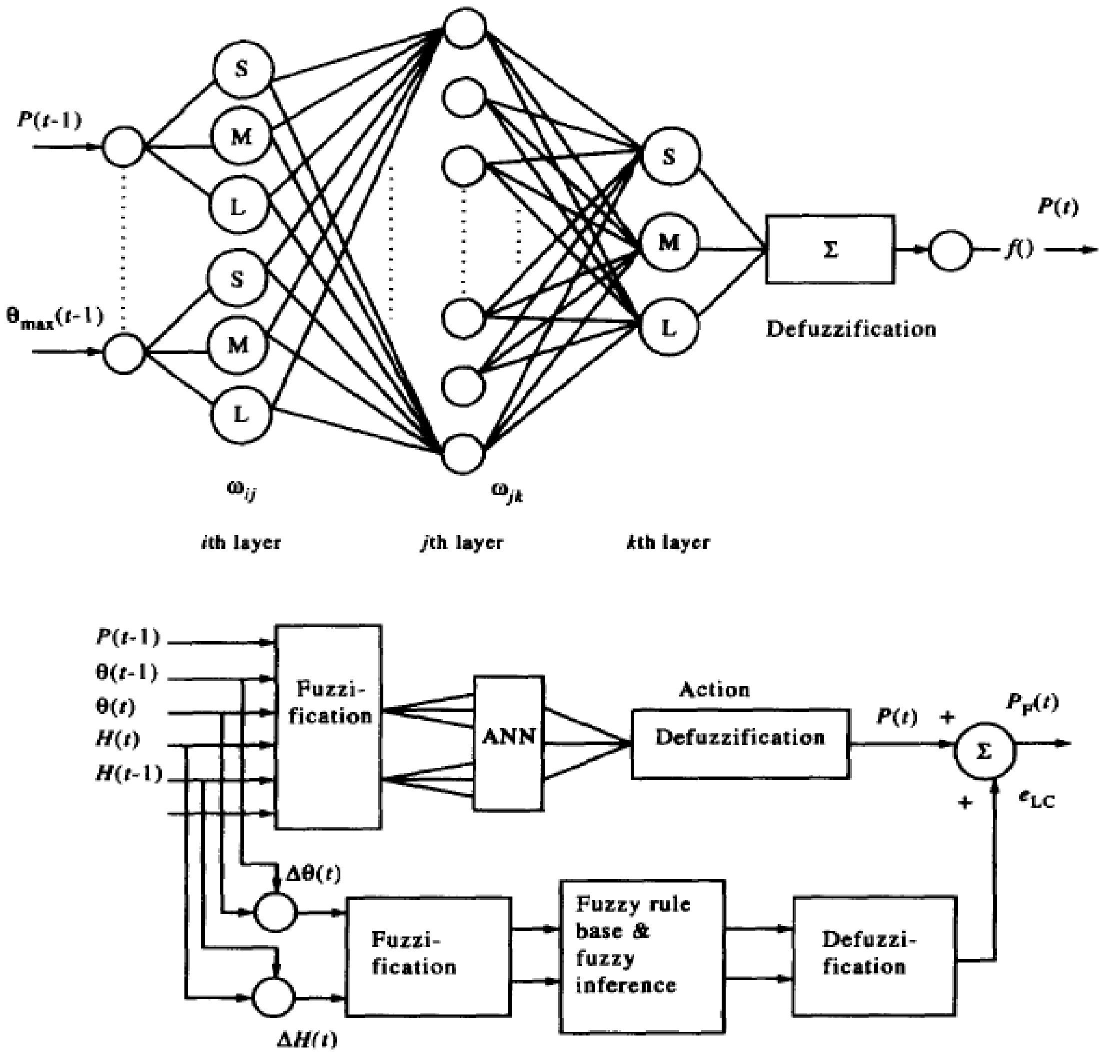


Fig. 5. Fuzzy computing model for forecasting.

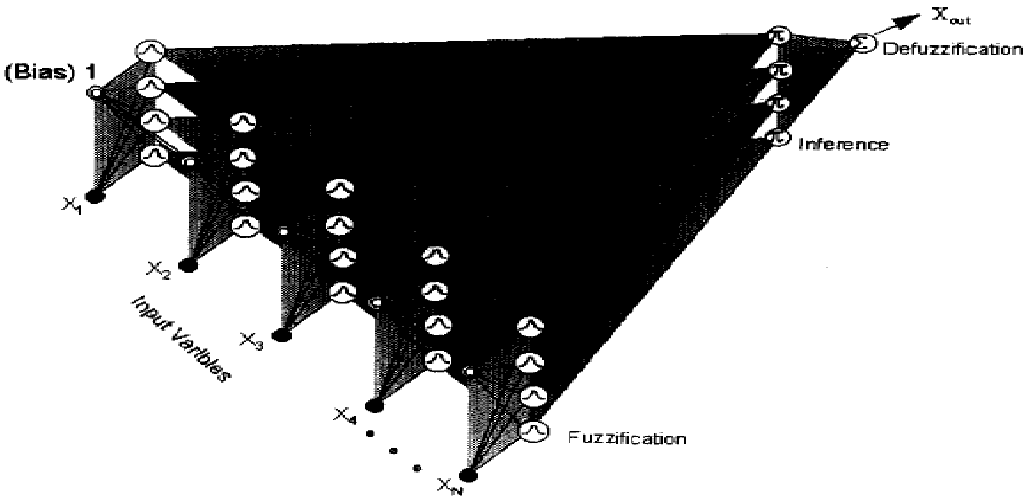
Σχήμα 2.6: *Ασαφές υπολογιστικό Μοντέλο.*



<u>Σχήμα 2.7</u>: *Νεύρο ασαφές υπολογιστικό μοντέλο για πρόβλεψη ηλεκτρικού φορτίου.*

Βραχυπρόθεσμη πρόβλεψη αιχμής και μέσου φορτίου πραγματικού χρόνου με χρήση αυτό-οργανωμένων νευρωνικών δικτύων

Αυτό το άρθρο παρουσιάζει ένα αυτό-οργανωμένο ασαφές νευρωνικό δίκτυο, που βασίζεται σε ένα βραχυπρόθεσμο σύστημα πρόβλεψης ηλεκτρικού φορτίου για μια υλοποίηση πραγματικού χρόνου. Το προαναφερθέν δίκτυο λοιπόν, αποτελείται από τα εξής επίπεδα: επίπεδο εισόδου, επίπεδο ασαφοποίησης, επίπεδο συμπερασμάτων και το επίπεδο απόασαφοποίησης όπως φαίνονται και στο Σχήμα 2.8. Το δίκτυο αποτελείται από N μεταβλητές εισόδου, με N νευρώνες στο διάνυσμα εισόδου και R αριθμό κανόνων, με R νευρώνες για εξαγωγή συμπερασμάτων. Έτσι, ο αριθμός των νευρώνων στο διάνυσμα ασαφοποίησης είναι N x R. Τα παρελθοντικά φορτία λαμβάνονται υπόψη για την βελτίωση της ικανότητας πρόβλεψης. Τόσο τα βάρη στα διαφορετικά επίπεδα του δικτύου, όσο και η δομή της συναρτήσεως συμμετοχής του δικτύου βελτιστοποιούνται με χρήση ενός καινοτομικού αλγόριθμου αναβάθμισης. Έτσι το δίκτυο συμπεριφέρεται ικανοποιητικά, ξεκινώντας από ένα αρχικό σετ τυχαίων βαρών. Ο προτεινόμενος αλγόριθμος εκμεταλλεύεται την έννοια του σφάλματος ανάστροφης διάδοσης (back propagation). Το προτεινόμενο δίκτυο δίνει λογαριασμό για εποχιακά και ημερήσια χαρακτηριστικά, όσο καλά και για περίεργες συνθήκες (π.χ. διακοπές). Είναι ικανό να προβλέψει στο ανεκτό χρονικό διάστημα από μία ημέρα έως μία εβδομάδα. Η εβδομαδιαία τιμή του σφάλματος (MAPE) είναι κυρίως από 0.5-1.5% για πρόβλεψη αιχμής φορτίου 24-ώρες μπροστά και από 0.05-0.15% για πρόβλεψη μέσου φορτίου, η οποία είναι πολύ καλή ως αποτέλεσμα σε σχέση με αυτές των BP νευρωνικών δικτύων.



Σχήμα 2.8: *Μοντέλο πρόβλεψης φορτίου με χρήση ενός SFNN (self-organising fuzzy neural network).*

Βραχυπρόθεσμη πρόβλεψη φορτίου βασιζόμενη στην παράλληλη υλοποίηση νευρωνικών δικτύων

Αυτό το άρθρο, παρουσιάζει την ανάπτυξη και την εφαρμογή των προηγμένων νευρωνικών δικτύων για την επιτυχή αντιμετώπιση του προβλήματος της βραχυπρόθεσμης πρόβλεψης του ηλεκτρικού φορτίου. Αρκετές προηγμένες νευρωνικές αρχιτεκτονικές δοκιμάζονται συμπεριλαμβανομένων, των MLP (multilayer perceptrons) με τον καθιερωμένο BP αλγόριθμο, των δικτύων τυχαίας ενεργοποίησης βαρών (RAWN) και των εκπαιδευόμενων (RAWN) με κινούμενο παράθυρο (MWRAWN), των δικτύων με συναρτήσεις radial basis (RBFN), των επαναληπτικών δικτύων πραγματικού χρόνου (RT-RN), των αναδρομικών επαναληπτικών νευρωνικών δικτύων (ARNN) αλλά και των καινοτομικών παραλλαγών που παρουσιάζονται και συγκρίνονται σε αυτό το έγγραφο. Η απόδοση της κάθε μιας από τις παρουσιαζόμενες δομές, αξιολογείται με τη βοήθεια μιας εκτενούς μελέτης προσομοίωσης, στην οποία χρησιμοποιούνται πραγματικά ωριαία δεδομένα φορτίων από το σύστημα ισχύος του νησιού της Κρήτης, στην Ελλάδα. Τα στατιστικά αποτελέσματα του σφάλματος πρόβλεψης, αντιστοιχώντας στο ελάχιστο και στο μέγιστο των χρόνο-σειρών φορτίου, υποδεικνύουν ότι τα προτεινόμενα μοντέλα πρόβλεψης φορτίου παρέχουν σημαντικά ακριβέστερες προβλέψεις, έναντι των συμβατικών προτύπων πρόβλεψης όπως της μεθόδου αναδρομής (autoregressive method) και του αλγορίθμου αναστροφής διάδοσης (Back Propagation). Τέλος, προτείνεται και εφαρμόζεται μια προσέγγιση παράλληλης επεξεργασίας για πρόβλεψη 24 ωρών μπροστά. Σύμφωνα με αυτήν την διαδικασία, το ζητούμενο φορτίο για κάθε συγκεκριμένη ώρα προβλέπεται, όχι μόνο χρησιμοποιώντας τις χρόνο-σειρές φορτίων για αυτήν την συγκεκριμένη ώρα από τις προηγούμενες ημέρες, αλλά επίσης και τα προβλεπόμενα δεδομένα φορτίων των πιο κοντινών προηγούμενων χρονικών βημάτων της ίδιας ημέρας. Κατά συνέπεια, οι ακριβείς αυτές προβλέψεις φορτίων λαμβάνονται χωρίς την ανάγκη των καιρικών στοιχείων που αυξάνουν την πολυπλοκότητα των συστημάτων, την απαίτηση σε αποθήκευση αλλά και το κόστος.

Βραχυπρόθεσμη πρόβλεψη φορτίου διά μέσου της νεύρο-ασαφούς συνεργασίας

Ένα σημαντικό στοιχείο ενός αποτελεσματικού συστήματος χειρισμού ενέργειας, είναι το καλά σχεδιασμένο βραχυπρόθεσμο χρονοδιάγραμμα των μονάδων παραγωγής ενέργειας. Οι διαχειριστές των συστημάτων ενέργειας χρησιμοποιούν τα ιστορικά δεδομένα φορτίων, για να προγραμματίσουν έτσι ώστε οι διαθέσιμες μονάδες παραγωγής να ικανοποιούν τα ωριαία φορτία του συστήματος σε ένα οικονομικό αλλά και αξιόπιστο συμβιβασμό. Αυτό το άρθρο λοιπόν περιγράφει, πως ένα ασαφούς λογικής (FL) έμπειρο σύστημα ολοκληρώνεται με τεχνητά νευρωνικά δίκτυα (ANN) για πιο ακριβείς βραχυπρόθεσμες προβλέψεις φορτίου. Τα 24 ώρες μπροστά, προβλεπόμενα φορτία επιτυγχάνονται με δύο βήματα. Πρώτα, μια FL υπό-μονάδα χαρτογραφεί την μη-γραμμική σχέση μεταξύ των παραμέτρων του καιρού και της επίδρασης τους πάνω στην ημερήσια αιχμή ηλεκτρικού φορτίου. Στην συνέχεια, 12 ANN υπό-μονάδες εκπαιδεύονται χρησιμοποιώντας παρελθοντικά ωριαία φορτία και δεδομένα του καιρού συνδυαζόμενα με τα δεδομένα εξόδου του FL, έτσι ώστε να εκτελεστεί η τελική πρόβλεψη. Συγκρίσεις που έγιναν μεταξύ αυτού του μοντέλου, ενός ANN μοντέλου και ενός ARMA (Autoregressive Moving Average) μοντέλου, δείχνουν την αποτελεσματικότητα και την ακρίβεια της νέας αυτής προσέγγισης η οποία παρουσιάζει σχετικά μικρό σφάλμα. Μεγαλύτερα σφάλματα κατ' απόλυτη τιμή παρουσιάστηκαν κατά την καλοκαιρινή περίοδο, αυτό οφείλεται στην υψηλότερη πολυπλοκότητα της σχέσης μεταξύ καιρού-φορτίου σε σχέση με τις άλλες περιόδους.

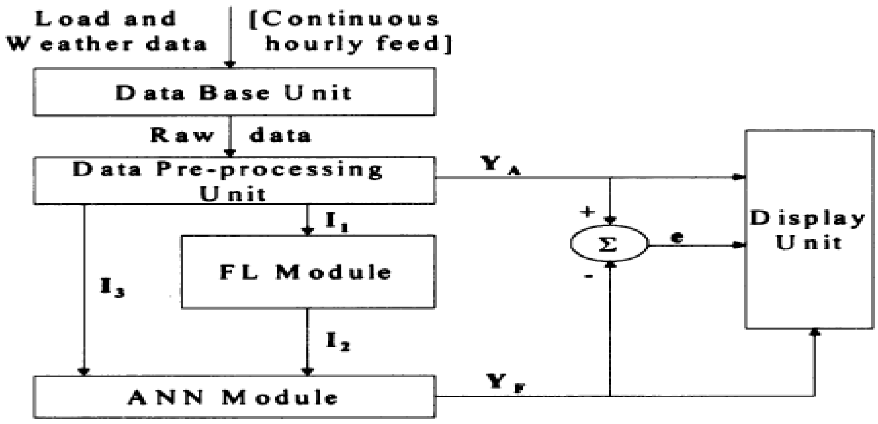


Fig. 1. The FL–ANN modular network.

Σχήμα 2.9: *Το FL-ANN δίκτυο και οι υπό-μονάδες του.*

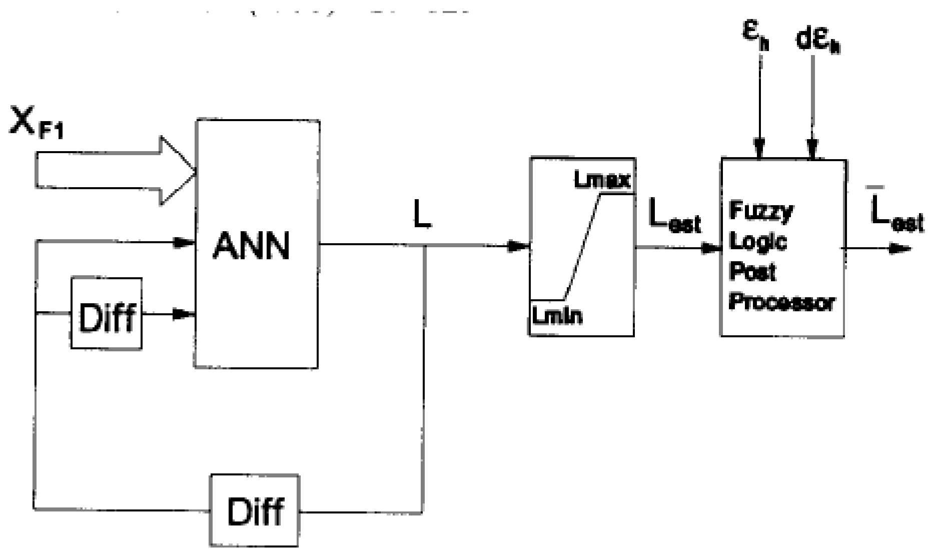
Όπου I_1 και I_2 η είσοδος και η έξοδος δεδομένων στην FL υπό-μονάδα, I_3 η είσοδος δεδομένων στην ANN υπό-μονάδα σε συνδυασμό με την I_2 και Y_F : προβλεπόμενο φορτίο, Y_A: το πραγματικό φορτίο.

Διαστάσεις της ακρίβειας και της οικονομικής αξίας στην βραχυπρόθεσμη πρόβλεψη ενέργειας με χρήση τεχνητών νευρωνικών δικτύων

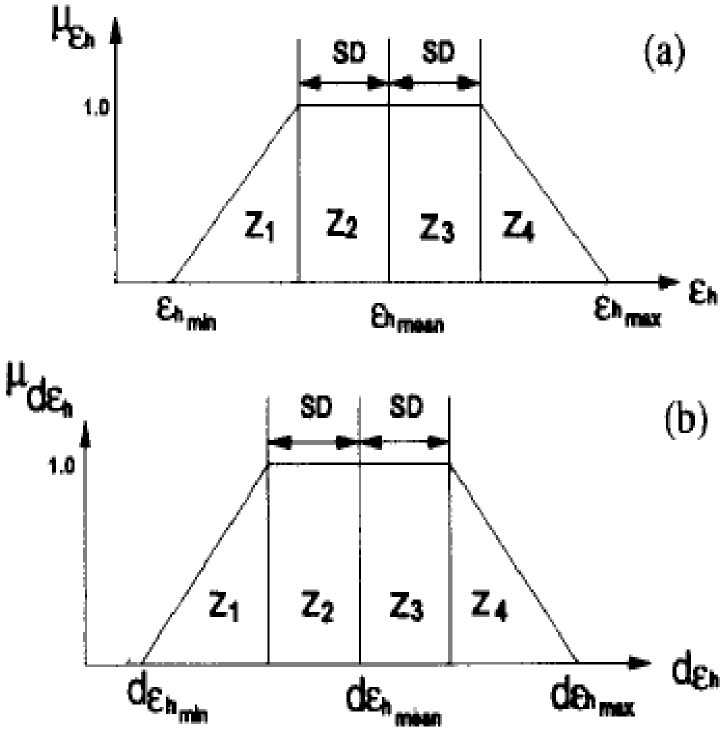
Το άρθρο αυτό, αναφέρεται στην πρόβλεψη του ηλεκτρικού φορτιού αλλά και του καυσίμου (φυσικό αέριο) για κοντινό χρονικό ορίζοντα με χρήση τεχνητών νευρωνικών δικτύων. Τα τεχνητά νευρωνικά δίκτυα, σε αυτήν την περίπτωση, παράγουν ωριαίες προβλέψεις ηλεκτρικών φορτίων μέχρι και 30 ημέρες μπροστά και έχουν τρία επίπεδα τροφοδότησης που περιλαμβάνουν το επίπεδο εισόδου, το κρυμμένο επίπεδο και το επίπεδο εξόδου. Σιγμοειδείς συναρτήσεις ενεργοποίησης χρησιμοποιούνται για τις συναρτήσεις μεταφοράς των κόμβων του κρυμμένου επιπέδου καθώς και του επιπέδου εξόδου. Η τελευταία έκδοση των προτεινόμενων ANN, αποτελείται από δυο διαφορετικά νευρωνικά δίκτυα, των οποίων οι προβλέψεις συνδυάζονται μεταξύ τους σε ένα μοντέλο παλινδρόμησης. Και τα δυο ANN έχουν 79 εισόδους οι οποίες χωρίζονται σε: 24 ωριαία φορτία της προηγουμένης μέρας, 24 παράμετροι του καιρού, 24 ωριαίες προβλέψεις φορτίου για την επόμενη μέρα και 7 δείκτες, ένας για κάθε μέρα της εβδομάδας. Τα δύο νευρωνικά δίκτυα διαφέρουν ως προς τις εξόδους τους, το ένα προβλέπει κατ' ευθείαν το ωριαίο φορτίο της επόμενης μέρας, ενώ το άλλο προβλέπει την αλλαγή στο ωριαίο φορτίο από την προηγουμένη μέρα στην επόμενη. Για την εκπαίδευση των δικτύων χρησιμοποιείται ο BP (back-propagation) αλγόριθμος. Το μοντέλο χρησιμοποιεί ξεχωριστά σετ στοιχείων για εκπαίδευση και αξιολόγηση. Το σφάλμα MAPE έχει ένα εύρος από 1% έως 4,5% με ελάχιστες περιπτώσεις να είναι πάνω από 3% δηλαδή είναι από αρκετά έως και πολύ πιο ακριβείς από τις προηγουμένας τις οποίες αντικατέστησε. Η οικονομική αξία της μείωσης του σφάλματος στη πρόβλεψη του ηλεκτρικού φορτίου υπολογίζεται ότι εξοικονόμησε περίπου 800.000 $ το χρόνο ανά περίπτωση. Το νευρωνικά δίκτυα, βελτιώνουν κατά πολύ την ακρίβεια της πρόβλεψης άρα και την διαχείριση των αποφάσεων .

Ένα νέο νεύρο-ασαφές βασιζόμενο στην αυτό-διόρθωση του διασυνδεμένου ηλεκτρικού φορτίου μοντέλο πρόβλεψης.

Αυτό το άρθρο παρουσιάζει ένα βραχυπρόθεσμο, νεύρο-ασαφές μοντέλο πρόβλεψης φόρτου (STLF). Η προτεινόμενη ANN προσεγγιστική λειτουργία μοντελοποιεί τις σχέσεις μεταξύ της αιχμής ηλεκτρικού φόρτου ανά ώρα και των μεταβλητών που την επηρεάζουν. Ονομαστικά είναι, η μεταβλητότητα του καιρού και της θερμοκρασίας, το είδος και η ώρα της ημέρας, οι ενυπάρχουσες παράμετροι του ιστορικού του φορτίου που υποδεικνύουν τις τάσεις, κυκλικές κυματώσεις και τέλος συνηθισμένα εποχικά και ασυνήθιστα ειδικά γεγονότα. Το εισερχόμενο διάνυσμα πρόβλεψης φόρτου επεκτάθηκε για να εκτιμήσει τις περισσότερες από τις εισερχόμενες επικρατούσες μεταβλητές, οι οποίες επηρεάζουν την βραχυπρόθεσμη πρόβλεψη φόρτου. Το μοντέλο λοιπόν, υλοποιεί μια προ-επεξεργασία για την δημιουργία του εισερχόμενου διανύσματος καθώς και την ταξινόμηση προτεραιοτήτων χρησιμοποιώντας το ιστορικό του φορτίου και τα δεδομένα του συστήματος. Ένα σύνολο μετά-επεξεργασίας ασαφούς λογικής παρέχει διόρθωση λαθών, φιλτράρισμα των δεδομένων, συντονισμό στο δίκτυο καθώς και ρύθμιση των δεδομένων της πρόβλεψης ηλεκτρικού φορτίου. Στην συνέχεια, παρατίθενται δύο σχήματα, το πρώτο απεικονίζει το προτεινόμενο STLF μοντέλο και το δεύτερο απεικονίζει τις λεπτομέρειες των ασαφών λειτουργιών των σχέσεων βάρους.



<u>Σχήμα 2.10:</u> *Μοντέλο βραχυπρόθεσμης πρόβλεψης φορτίου βασιζόμενο στα ANN.*



Σχήμα 2.11: *(α)Λανθασμένη λειτουργία σχέσης:* $\varepsilon_{h,max}$ *=0.289,* $\varepsilon_{h,min}$ *=-0.4639,* $\varepsilon_{h,mean}$ *=0.0016,* $\varepsilon_{h,SD}$ *=0.0822 (β) Λανθασμένη εκτίμηση λειτουργίας σχέσης:* $\varepsilon_{h,max}$ *0,5833,* $\varepsilon_{h,min}$ *=-0.6588,* $\varepsilon_{h,mean}$ *=-0,0019,* $\varepsilon_{h,SD}$ *=0.1003*

Επίδραση της θερμοκρασίας στην βραχυπρόθεσμη πρόβλεψη φορτίου με χρήση ολοκληρωμένων ANN

Μια ολοκληρωμένη προσέγγιση των τεχνητών νευρωνικών δικτύων (ANN) για τη λύση του προβλήματος της βραχυπρόθεσμης πρόβλεψης φορτίου (STLF), προτείνεται σε αυτό το άρθρο. Η προτεινόμενη προσέγγιση λοιπόν, αποτελείται από τέσσερις ενότητες, το βασικό τεχνητό νευρωνικό δίκτυο ANN, το μέγιστης και ελάχιστης τιμής ANN (Peak and Valley), το Averager and Forecaster ANN και το Adaptive Combiner ANN , οι οποίες ενσωματώνονται σε μια μέθοδο για την πρόβλεψη του φορτίου. Το βασικό ANN χρησιμοποιεί τα ιστορικά στοιχεία του φορτίου και της θερμοκρασίας για να προβλέψει το φορτίο για τις επόμενες 24 ώρες, ενώ το μέγιστης και ελάχιστης τιμής ANN (Peak and Valley), χρησιμοποιεί τα προηγούμενα στοιχεία μέγιστης και ελάχιστη τιμής του φορτίου και των θερμοκρασιών, αντίστοιχα. Το Averager εντοπίζει την απόκλιση του μέσου όρου των φορτίων σε σχέση με την προηγούμενη συμπεριφορά φορτίων, ενώ το Adaptive Combiner ANN χρησιμοποιεί το σταθμισμένο συνδυασμό των εξόδων από το βασικό ANN και τον Averager and Forecaster ANN για να προβλέψει το τελικό φορτίο. Στο παρακάτω σχήμα παρουσιάζεται η αρχιτεκτονική του σύνθετου αυτού νευρωνικού δικτυού.

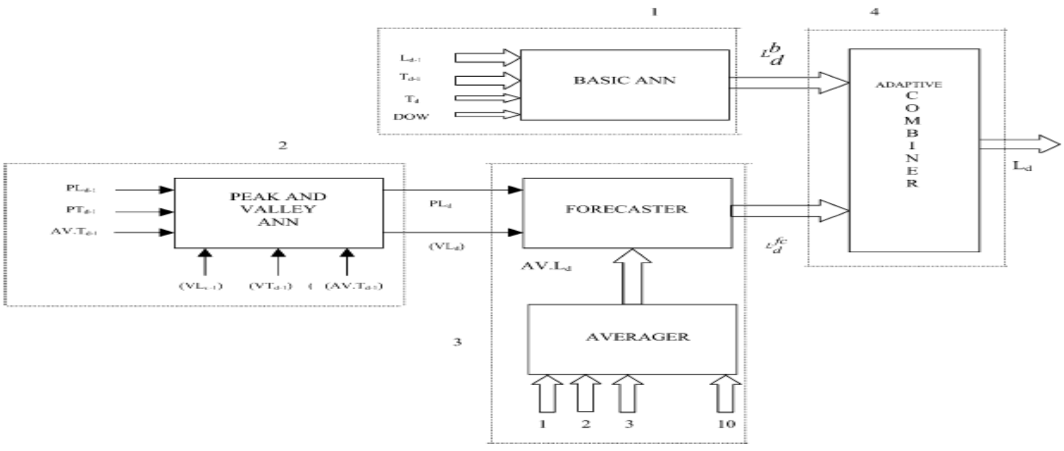


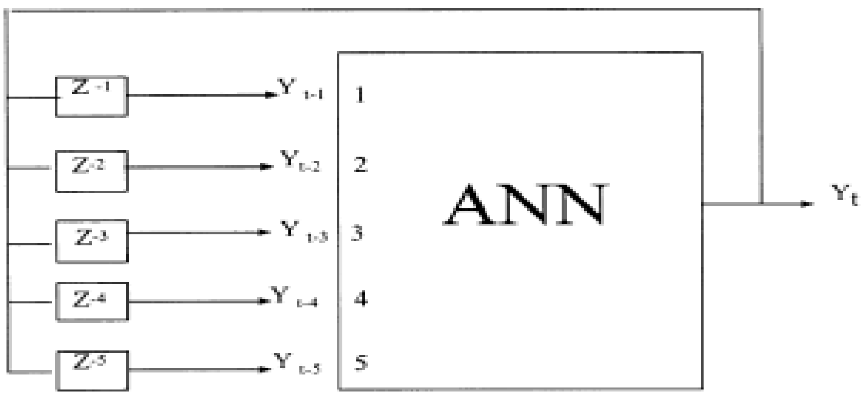
Fig. 1. Block diagram of proposed architecture.

<u>Σχήμα 2.12</u>: *Διάγραμμα της προτεινόμενης αρχιτεκτονικής Adaptive Combiner*

Τα αποτελέσματα που παίρνουμε από το Adaptive Combiner ANN, που είναι ο συνδυασμός των Basic ANN και του Averager and Forecaster ANN, είναι πιο ακριβή και έχουν χαμηλότερο σφάλμα από αυτά που παίρνουμε από το κάθε ένα ξεχωριστά. Ακόμα, ο αριθμός των σχεδίων εκπαίδευσης αλλά και η θερμοκρασία παίζουν σπουδαίο ρολό στην ακρίβεια των αποτελεσμάτων.

Εφαρμογή τεχνητών νευρωνικών δικτύων στην μακροπρόθεσμη πρόβλεψη ζήτησης.

Αυτό το άρθρο, αναφέρεται στην εφαρμογή των τεχνητών νευρωνικών δικτύων στην πρόβλεψη μακράς διαρκείας. Δηλαδή, πρόβλεψη σε μεγάλο χρονικό ορίζοντα ο οποίος μπορεί να είναι σε βάθος ενός, δέκα ή και τριάντα χρόνων. Λόγω του μεγάλου ορίζοντα πρόβλεψης υπάρχουν και μεγάλες αβεβαιότητες στα αποτελέσματα πράγμα που μετατρέπει την όλη υπόθεση σε αρκετά μεγάλο υπολογιστικό πρόβλημα. Κύριο χαρακτηριστικό του μοντέλου είναι η μικρή απαίτηση σε δεδομένα. Το μοντέλο των τεχνητών νευρωνικών δικτύων χρησιμοποιεί τον BP αλγόριθμο και έχει μερικά χαρακτηριστικά που κάνουν την όλη διαδικασία πιο αξιόπιστη. Το πρώτο χαρακτηριστικό είναι ότι η εκπαίδευση του BP σχεδιάστηκε, έτσι ώστε να ελαχιστοποιεί το σφάλμα ελαχίστων τετραγώνων, μεταξύ της εξόδου και των πραγματικών τιμών. Η υλοποίηση της διαδικασίας των τεχνητών νευρωνικών δικτύων συνοψίζεται στην συνέχεια. Τα δεδομένα τα οποία χρησιμοποιούνται ως είσοδοι, είναι τα ηλεκτρικά φορτία προηγούμενων χρονικών περιόδων. Η ακρίβεια των δεδομένων δεν είναι σημαντική καθώς τα τεχνητά νευρωνικά δίκτυα είναι εκπαιδευμένα να αγνοούν τον θόρυβο στα δεδομένα. Ο αριθμός των νευρωνικών εισόδων του μοντέλου ποικίλει ανάλογα τη περίπτωση ενώ η έξοδος, δηλαδή η προβλεπόμενη τιμή του φορτίου, είναι πάντα μια. Ανάλογα τον αριθμό των εισόδων στο σύστημα παίρνουμε και τα αντίστοιχα αποτελέσματα, δηλαδή όσο περισσότερες οι είσοδοι στο σύστημα, τόσο καλύτερη θα είναι η προβλεπόμενη τιμή (τόσο καλύτερη η προσέγγιση αυτής στην πραγματική τιμή, συνεπώς τόσο μικρότερο το σφάλμα. Συγκριτικά, με αλλά μοντέλα πρόβλεψης τα αποτελέσματα έδειξαν ότι τα τεχνητά νευρωνικά δίκτυα μπορούν να δώσουν ακριβή αποτελέσματα, χρησιμοποιώντας ένα ελάχιστο αριθμό δεδομένων σε σχέση με αλλά μοντέλα. Παρακάτω δίδεται σχήμα το οποίο περιγράφει το πώς λειτουργεί το νευρωνικό μοντέλο, Y_t είναι η έξοδος ενώ τα Y_{t-n} είναι οι είσοδοι (δεδομένα) για n χρονικές περιόδους. Η τιμή της πρόβλεψης Yt' εξαρτάται από τις τιμές των εισόδων δηλαδή των προηγουμένων χρονικών στιγμών.



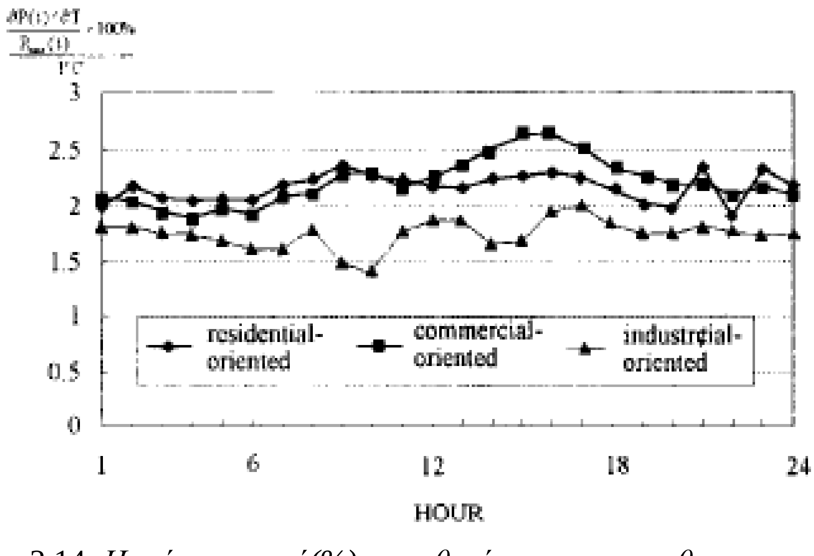
Σχήμα 2.13: *Προτεινόμενη ANN αρχιτεκτονική.*

Εφαρμογή τεχνητών νευρωνικών δικτύων στην πρόβλεψη φορτίων υποσταθμών.

Σε αυτό το άρθρο, μια εποπτεύουσα λειτουργική τεχνική τεχνητών νευρωνικών δικτύων (ANN) καλείται να επιλύσει το πρόβλημα της πρόβλεψης του φορτίου τριών Taipower υποσταθμών οι οποίοι εξυπηρετούν τους διαφορετικούς τύπους πελατών. Οι υποσταθμοί σε ένα σύστημα διανομής κατηγοριοποιούνται σε τρία είδη:

- Τους οικιακούς
- Τους εμπορικούς και
- Τους βιομηχανικούς

Κάθε υποσταθμός, έχει ένα διαφορετικό σχέδιο κατανάλωσης ισχύος ανάλογα με τον συνδυασμό των πελατών που εξυπηρετεί. Η ακρίβεια της πρόβλεψης φορτίων εμπλουτίζεται, λαμβάνοντας υπόψη την επίδραση της θερμοκρασίας στην ζήτηση φορτίων του υποσταθμού. Το μοντέλο χρησιμοποιεί ως εισόδους προηγούμενα και αντίστοιχα ωριαία στοιχειά καιρού όπως η θερμοκρασία και η σχετική υγρασία, αλλά και προηγούμενα και αντίστοιχα στοιχειά φορτίων, ενώ ξεχωρίζει τις ημέρες σε καθημερινές και Σαββατοκύριακα χρησιμοποιώντας ένα δείκτη. Το μοντέλο πρόβλεψης δίνει αποτελέσματα με σχετικά καλή ακρίβεια το σφάλμα πρόβλεψης του φορτίου είναι 1,93%, 2,02% και 2,87% για τα αντίστοιχα τρία είδη υποσταθμών (κατοικιών, εμπορικών και βιομηχανικών), η αντίστοιχη ακρίβεια πρόβλεψης των φορτίων ενισχύεται με την εξέταση της επίδρασης της θερμοκρασίας στη ζήτηση των φορτίων των υποσταθμών. Η ανάλυση της ευαισθησίας του συστήματος στην θερμοκρασία έδειξε ότι για κάθε ένα βαθμό κελσίου αλλαγής της θερμοκρασίας τα αποτελέσματα της πρόβλεψης αλλάζουν περισσότερο από 2% στην περίπτωση των κατοικιών, από 1.6 έως 2.1 % στην περίπτωση των εμπορικών καταντημάτων και έως 1.6% στην περίπτωση των βιομηχανιών για τις καθημερινές. Στο ακόλουθο σχήμα φαίνεται η επίδραση της θερμοκρασίας στην πρόβλεψη φορτίου για τα σαββατοκύριακα.



Σχήμα 2.14: *Η επί τοις εκατό(%) ευαισθησία ως προς την θερμοκρασία τριών διαφορετικών τύπων υποσταθμών για τα σαββατοκύριακα.*

Ημερήσια πρόβλεψη του προφίλ αστικών ηλεκτρικών φορτίων με χρήση των ANN.

Το άρθρο αυτό, επεξηγεί μια συνδυασμένη προσέγγιση που βασίζεται σε ελεγχόμενα και μη τεχνητά νευρωνικά δίκτυα, για τη βραχυπρόθεσμη πρόβλεψη ζήτησης του ηλεκτρικού φορτίου σε μια προαστιακή περιοχή με βάθος πρόβλεψης τις 24 ώρες. Μια προληπτική ταξινόμηση των ιστορικών δεδομένων του φορτίου διενεργείται κατά την διάρκεια του μη-επιβλεπόμενου σταδίου. Η πραγματική πρόβλεψη επιτυγχάνεται με χρήση ενός νευρωνικού μοντέλου το οποίο αποτελείται από δυο επίπεδα (το ένα εξ' αυτών κρυμμένο) και είναι πρόσθιας τροφοδότησης (Feed forward).Το παραπάνω μοντέλο εκπαιδεύεται με χρήση του αλγόριθμου Back propagation (όπισθεν διάδοσης). Με σκοπό να εξετάσουμε την επιρροή της μεταβλητότητας του κλίματος στην ηλεκτρική κατανάλωση, το νευρωνικό δίκτυο εκπαιδεύεται με δεδομένα καιρού, όπως η θερμοκρασία, η σχετική υγρασία και η παγκόσμια ηλιακή ακτινοβολία παράλληλα με παλαιοτέρα δεδομένα φορτίων. Τα στοιχεία αυτά αποτελούν τις εισόδους. Η επιβεβαίωση του μοντέλου διενεργείται συγκρίνοντας τις προβλέψεις του μοντέλου με τα δεδομένα του φορτίου που δεν χρησιμοποιήθηκαν στην εκπαίδευση του δικτύου.

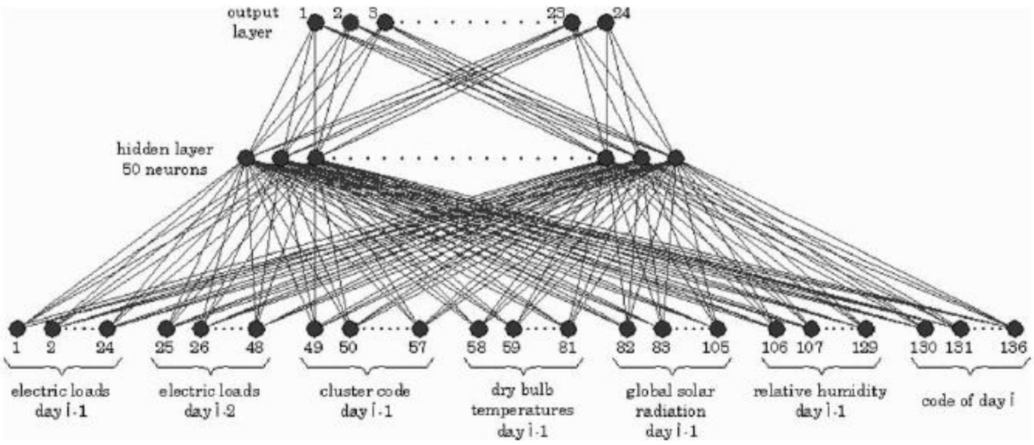


Fig. 14. The MLP architecture.

Σχήμα 2.15: *Αρχιτεκτονική του παραπάνω MLP δικτύου.*

Παρά τον μικρό αριθμό των διαθεσίμων σχεδίων εκπαίδευσης, το μοντέλο δίνει πολύ καλά αποτελέσματα, με μέσο σφάλμα μόλις 1,97%, επιβεβαιώνοντας την καταλληλότητα του μοντέλου για την λύση του προβλήματος του αστικού ενεργειακού προγραμματισμού.

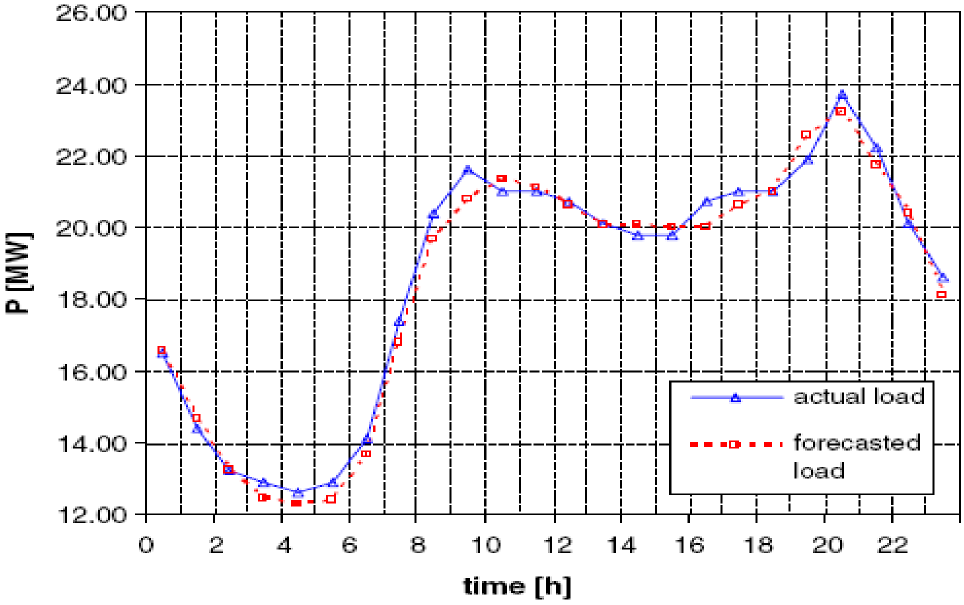


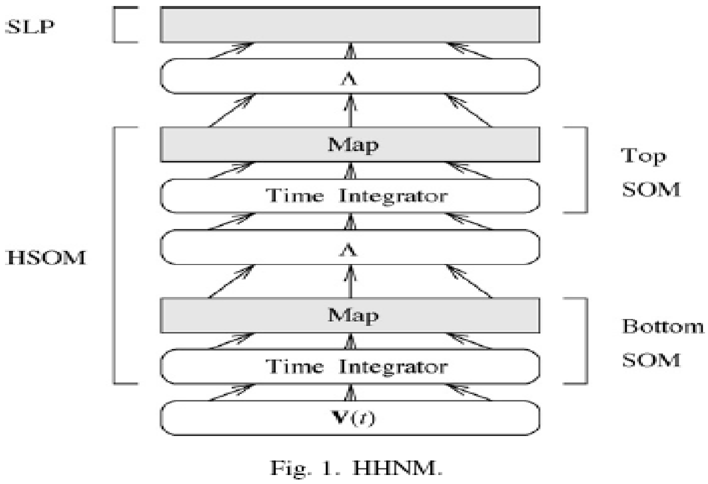
Fig. 16. Comparison between actual and forecasted loads for 30/06/01.

Σχήμα 2.16: Σφάλμα πρόβλεψης μοντέλου.

Ιεραρχικό νευρωνικό μοντέλο για μακροπρόθεσμη πρόβλεψη.

Σε αυτό το άρθρο, προτείνεται ένα νέο ιεραρχικό υβριδικό νευρωνικό μοντέλο, για την επίλυση του προβλήματος της μακροπρόθεσμης πρόβλεψης φορτίων. Το νευρωνικό αυτό μοντέλο αποτελείται από δύο αυτορρυθμιζόμενου χάρτη (SOM) δίκτυα-το ένα πάνω από το άλλο-καθώς και ενός απλού επιπέδου perceptron (SLP). Αυτό έχει εφαρμογές σε τομείς οι οποίοι απαιτούν ανάλυση σε χρονικό ορίζοντα. Το μοντέλο αυτό συγκρίνεται με ένα πολυεπίπεδο perceptron. Και τα δύο προαναφερόμενα μοντέλα εκπαιδεύονται και αξιολογούνται με δεδομένα φορτίων που εξάγονται από μια βορειοαμερικανική ηλεκτρική εταιρία και προβλέπουν είτε μια φορά την εβδομάδα είτε μια φορά τον μήνα, την μέγιστη τιμή του φορτίου και το μέσο φορτίο κατά την διάρκεια των επόμενων δύο ετών. Στην συνέχεια παρουσιάζεται σχηματικά το ιεραρχικό υβριδικό νευρωνικό δίκτυο, το οποίο έχει τρεις νευρωνικές εισόδους και οι οποίες παρουσιάζονται επιγραμματικά.

- Η <u>πρώτη είσοδος</u> αντιπροσωπεύει είτε την μέγιστη τιμή του φορτίου είτε το μέσο φορτίο της τρέχουσας εβδομάδας ή του τρέχοντα μήνα.
- Η <u>δεύτερη</u> και η <u>τρίτη είσοδος</u> αναπαριστούν μια τριγωνομετρική κωδικοποίηση για την εβδομάδα ή τον μήνα που θα γίνει η πρόβλεψη.



<u>Σχήμα 2.17</u>: *Ιεραρχικό υβριδικό νευρωνικό δίκτυο.*

Τα πειραματικά αποτελέσματα δείχνουν ότι η απόδοση του ιεραρχικού μοντέλου στην μακροπρόθεσμη πρόβλεψη του μέγιστου και του μέσου φορτίου είναι πολύ καλύτερη από αυτή του πολύ-επίπεδου perceptron. Αυτή η πολύ καλή απόδοση, δικαιολογείται από την ιεραρχική τοπολογία του. Παρόλα αυτά, τα αποτελέσματα που έχουν επιτευχθεί από το ιεραρχικό μοντέλο έχουν περιθώρια βελτίωσης.

Μια αποτελεσματική προσέγγιση για βραχυπρόθεσμη πρόβλεψη με χρήση τεχνητών νευρωνικών δικτύων(ANN).

Το μοντέλο πρόβλεψης φορτίου ενός συστήματος είναι μεγάλης σπουδαιότητας εργαλείο απόφασης έτσι ώστε ένα σύστημα ηλεκτρικής ενέργειας να λειτουργεί οικονομικά και με ασφάλεια. Χάρη στην ικανότητα χαρτογράφησης εισόδων-εξόδων τους, τα τεχνητά νευρωνικά δίκτυα ταιριάζουν πολύ καλά με αυτόν τον τύπο εφαρμογών. Σε αυτό το άρθρο, διεξάγεται μια έρευνα πάνω στην χρήση των τεχνητών νευρωνικών δικτύων για βραχυπρόθεσμη πρόβλεψη φορτίου στην οποία περιγράφονται οι δυνατότητες των ANN στην πρόβλεψη ηλεκτρικού φορτίου χωρίς την χρήση ιστορικών φορτίων στην είσοδο. Επιπλέον, από τις καιρικές μεταβλητές οι οποίες επηρεάζουν την κατανάλωση φορτίου μόνο η θερμοκρασία χρησιμοποιείται ως είσοδος στο μοντέλο αυτό, καθώς τα αποτελέσματα δείχνουν ότι άλλες μεταβλητές όπως η κατάσταση του ουρανού (κάλυψη από τα σύννεφα) αλλά και η ταχύτητα του ανέμου δεν έχουν κάποια σοβαρή επίδραση και δεν υπολογίζονται στην διαδικασία πρόβλεψης φορτίου. Πρόκειται λοιπόν, για ένα απλό πολύ-επίπεδο πρόσθιας τροφοδότησης (feed-forward) νευρωνικό δίκτυο το οποίο χρησιμοποιεί ως εισόδους:

- ένα δείκτη ώρας, ο οποίος δείχνει το μέγεθος του φορτίου ανάλογα την ώρα της ημέρας.
- ένα δείκτη ημέρας, ο οποίος δείχνει το μέγεθος του φορτίου ανάλογα την ημέρα της εβδομάδας που διανύουμε, αποδεικνύεται ότι το αναγκαίο φορτίο είναι αρκετά μικρότερο το σαββατοκύριακο σε σχέση με τις καθημερινές.
- εκτιμώμενη θερμοκρασία στην τρέχουσα χρονική στιγμή.
- εκτιμώμενη θερμοκρασία πριν μια ώρα .
- εκτιμώμενη θερμοκρασία πριν δυο ώρες.

Το μοντέλο εκπαιδεύεται διαιρώντας τα δεδομένα ενός χρόνου σε αρκετά κομμάτια, τα μισά από αυτά χρησιμοποιούνται στην εκπαίδευση και τα άλλα μισά χρησιμοποιούνται στον έλεγχο του ANN. Πιο συγκεκριμένα, σε κάθε μήνα η πρώτη και η τρίτη εβδομάδα χρησιμοποιούνται για εκπαίδευση, ενώ η δεύτερη και η τέταρτη για έλεγχο. Η εκπαίδευση για όλα τα κομμάτια δεδομένων γίνεται ταυτόχρονα και όλες οι είσοδοι-έξοδοι κανονικοποιούνται πριν την εκπαίδευση. Ο στόχος είναι να ρυθμιστούν τα βάρη του ANN, έτσι ώστε να ελαχιστοποιείται το σφάλμα μεταξύ των επιθυμητών και των πραγματικών εξόδων του μοντέλου μας. Ο αλγόριθμος που χρησιμοποιείται στην εκπαίδευση είναι ο Levenberg-Marquard και τα αποτελέσματα που δίνει είναι αρκετά καλά με αρκετά χαμηλό σφάλμα περίπου 0,981% σε σχέση με τις παραδοσιακές μεθόδους που δίνουν σφάλμα 1.083%.

Table 1
Correlation matrix for: load (*L*), temperature (*T*), wind velocity (*W*) and cloud cover (*C*)

	L	*T*	*W*	*C*
L	1.00	0.86	0.27	0.07
T	0.86	1.00	0.16	0.18
W	0.27	0.16	1.00	0.23
C	0.07	0.18	0.23	1.00

Πίνακάς 2.1: *Συσχετίσεις για το φορτίο(L), την θερμοκρασία(T), την ταχύτητα ανέμου(W), την κάλυψη από τα σύννεφα(C).*

Μια νεύρο- ασαφής προσέγγιση για την μοντελοποίηση της απαίτησης ηλεκτρισμού.

Αυτό το άρθρο, αξιολογεί την χρήση δύο δημοφιλών soft-computing τεχνικών και συμβατικών στατιστικών προσεγγίσεων που βασίζονται στο αυτό-ανάδρομο, ενσωματωμένο, κινούμενου μέσου μοντέλο (ARIMA) του Box-Jenkins, με σκοπό την πρόβλεψη της απαίτησης ηλεκτρισμού στην πολιτεία της Βικτώρια (Αυστραλία). Οι προς μελέτη τεχνικές λοιπόν είναι ένα εξελισσόμενο ασαφές νευρωνικό δίκτυο (EFuNN) και ένα τεχνητό νευρωνικό δίκτυο (ANN). Τα οποία εκπαιδεύονται με χρήση του βαθμονομημένου CGA (Conjugate gradient algorithm) αλγορίθμου και του αναστροφής διάδοσης BP (back-propagation) αλγόριθμου. Η ακρίβεια της πρόβλεψης συγκρίνεται με τις προβλέψεις που χρησιμοποιούνται από την VPX (Victorian Power Exchange) και με την αληθινή ενεργειακή απαίτηση. Για την αξιολόγηση, θεωρήσαμε σχέδιο απαίτησης φορτίων για 10 διαδοχικούς μήνες που λήφθηκαν κάθε 30 λεπτά για την εκπαίδευση διαφορετικών μοντέλων πρόβλεψης. Τα αποτελέσματα της δοκιμής έδειξαν ότι το νεύρο ασαφές δίκτυο λειτούργησε καλύτερα από τα άλλα νευρωνικά δίκτυα, το μοντέλο ARIMA και τις προβλέψεις VPX.

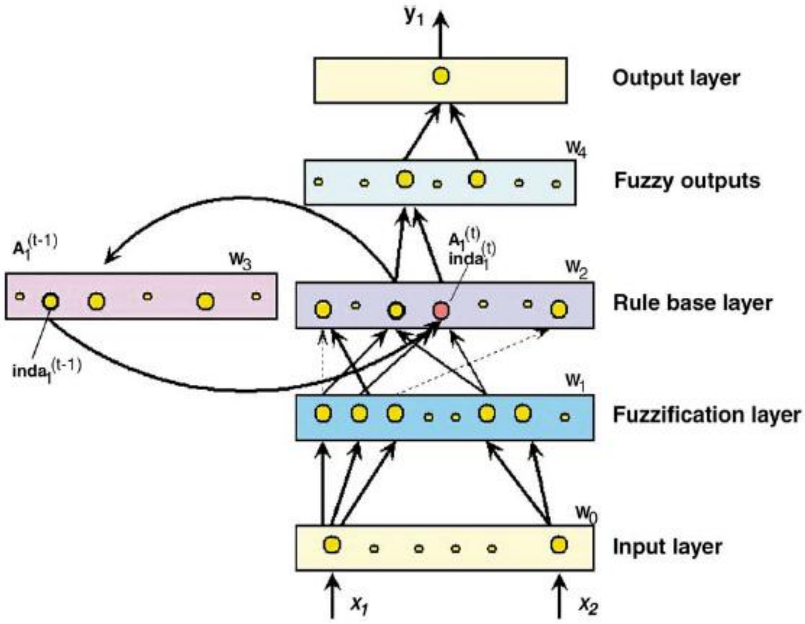


Fig. 1. Architecture of EFuNN.

Σχήμα 2.18: *Απεικόνιση Αρχιτεκτονικής EFuNN.*

Νεύρο Ασαφείς Προσεγγίσεις στην βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φόρτου.

Σε αυτό το άρθρο, διερευνάται η εφαρμογή των μοντέλων του Takagi-Sugeno στην βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φόρτου. Τα αποτελέσματα, συζητούνται και συγκρίνονται, για μοντέλα, με αρκετούς προσαρμοσμένης μάθησης αλγορίθμους και το κλασσικό πολύστρωματικό αντιληπτό δίκτυο. Για την ταυτοποίηση των μοντέλων με γραμμικούς προγόνους έχει χρησιμοποιηθεί ο συνδυασμός της εκτίμησης κλάσεων και της συνηθισμένης μεθόδου των ελαχίστων τετραγώνων. Για τον σκοπό της μοντελοποίησης μη γραμμικών προγόνων, έχει χρησιμοποιηθεί συνδυασμός εναλλασσόμενα ασαφών και ανατροφοδοτούμενων νευρωνικών δικτύων. Τα καλύτερα αποτελέσματα που έχουν επιτευχθεί για το σύστημα, με μη γραμμικές λειτουργίες στους προηγηθείς κανόνες, μοντελοποιούνται από τα ανατροφοδοτούμενα νευρωνικά δίκτυα. Η αποδοτικότητα των μοντέλων συγκρίνεται για δύο μέρες με την κορυφή της πρόβλεψης φόρτου στο δίκτυο διανομής. Σε προβλέψεις μικρού χρονικού ορίζοντα, ο μελλοντικός φόρτος επηρεάζεται κυρίως από τις παρελθοντικές τιμές και τον παράγοντα του καιρού (π.χ. θερμοκρασία, πίεση, υγρασία). Αυτή η σχέση είναι αρκετά σύνθετη και μη γραμμική. Η γλωσσική περιγραφή του συστήματος, πλαισιώνεται από μια βάση κανόνων, που αποτελείται από ένα ασαφές σετ ΕΑΝ-ΤΟΤΕ κανόνων, όπου το K είναι ο αριθμός των κανόνων, $x_1,\dots,x_n$: οι μεταβλητές εισόδου, y: η μεταβλητή εξόδου. Στην συνέχεια παρατίθεται γραφική απεικόνιση των παραπάνω.

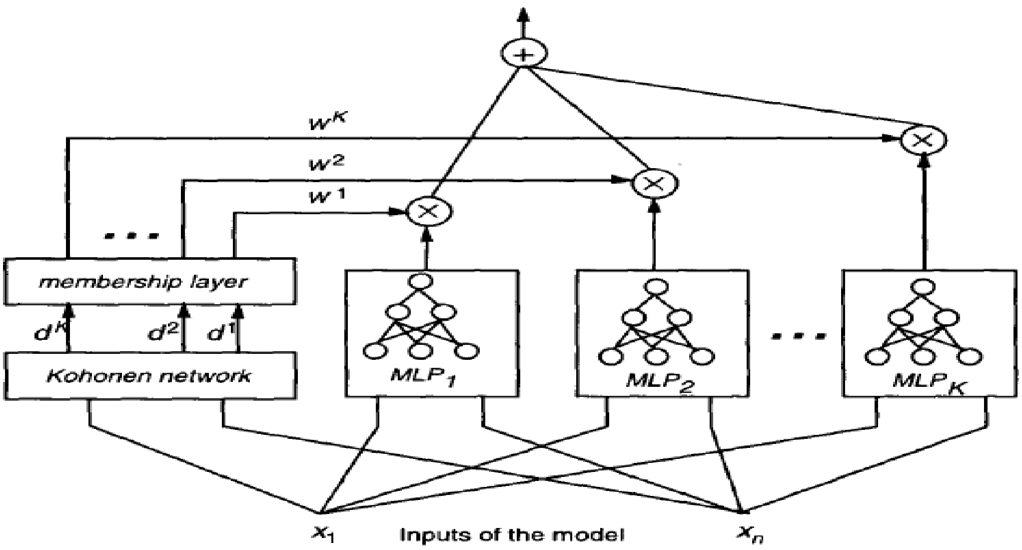
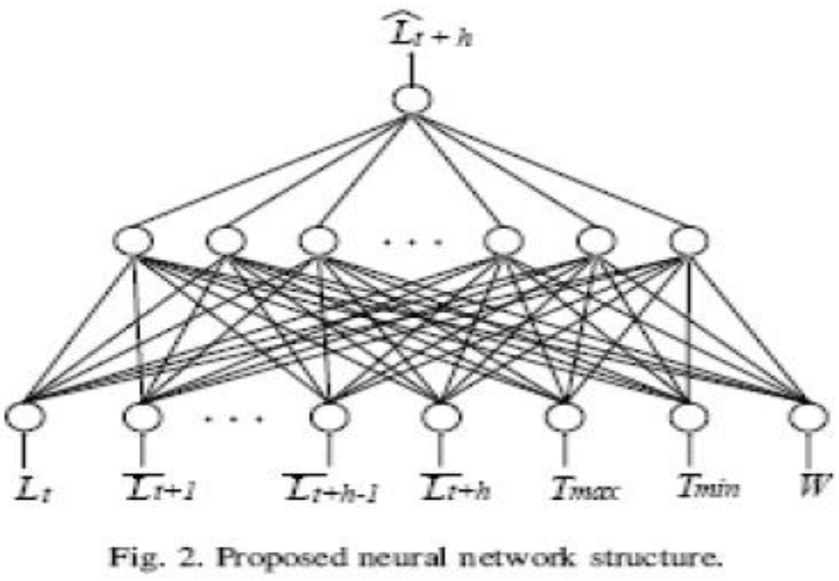


Fig. 1. Takagi Sugeno model with neural networks in the rule antecedents

Σχήμα 2.19: *Απεικόνιση Μοντέλου Tagagi Sugeno.*

Νευρωνικά δίκτυα βασισμένα στην πρόβλεψη φορτίων αρκετών ωρών μπροστά χρησιμοποιώντας την προσέγγιση παρόμοιων ημερών.

Στο άρθρο αυτό, παρουσιάζεται μια πρακτική μέθοδος λύσης για το πρόβλημα πρόβλεψης φορτίου σε κοντινό χρονικό ορίζοντα χρησιμοποιώντας τεχνητά νευρωνικά δίκτυα (ANN) συνδυάζοντας όμοιες ημερήσιες προσεγγίσεις. Τα νευρωνικά δίκτυα χρησιμοποιούνται σε παραδοσιακές μεθόδους πρόβλεψης χρησιμοποιώντας όμοια δεδομένα μεταξύ ημερών για να μάθουν την τάση της ομοιότητας . Παρ όλα αυτά, το να μαθαίνει όλα τα ημερήσια δεδομένα είναι μια σύνθετη εργασία και δεν ταιριάζει με την συνήθη εκπαίδευση των νευρωνικών δικτύων. Μια Ευκλείδεια νόρμα με σταθμισμένους παράγοντες (ύπαρξη βαρών) χρησιμοποιείται στην αξιολόγηση των πληροφοριών ημερών που είναι όμοιες σε καιρικές συνθήκες με την ημέρα πρόβλεψης. Η ακρίβεια της προτεινόμενης μεθόδου ενισχύεται με την προσθήκη του μείζονα κλιματολογικού παράγοντα, της θερμοκρασίας ενώ δίνεται μεγάλη προσοχή στις διαφορετικές εποχές (έτσι ώστε να είναι ακριβής η σχεδίαση) . Το σφάλμα της πρόβλεψης από μια έως και έξι ώρες μπροστά (MAPE) κυμαίνεται από 0,98% έως 2,43%. Τα μέγιστα και ελάχιστα ποσοστά λαθών καθώς και οι MAPE τιμές (που έχουν αποκτηθεί από αποτελέσματα της πρόβλεψης φορτίου), επιβεβαιώνουν ότι η μέθοδος μας δίνει αξιόπιστες προβλέψεις για αρκετές ώρες μπροστά. Το μέσο δείγμα του MAPE για πρόβλεψη μιας ώρας μπροστά είναι 0.98%, αυτός ο αριθμός αυξάνει σε μόλις 2.43% για προβλέψεις 6 ωρών μπροστά. Το μοντέλο αυτό δουλεύει καλά στην περίπτωση ξαφνικών κλιματολογικών αλλαγών. Στο ακόλουθο σχήμα αναπαρίσταται ένα τέτοιο δίκτυο.



Σχήμα 2.20: *Προτεινόμενη Δομή Νευρωνικού Δικτύου.*

Πρόβλεψη ηλεκτρικού φορτίου χρησιμοποιώντας ένα ασαφές ART και ARTMAP νευρωνικό δίκτυο

Σε αυτό το άρθρο, παρουσιάζεται ένα νευρωνικό δίκτυο που βασίζεται στην αρχιτεκτονική ART (adaptive resonance theory), ονομαζόμενο ασαφές ART και ARTMAP νευρωνικό δίκτυο και εφαρμόζεται στο πρόβλημα πρόβλεψης ηλεκτρικού φορτίου. Τα νευρωνικά δίκτυα που βασίζονται στην αρχιτεκτονική ART, έχουν δύο θεμελιώδη χαρακτηριστικά που είναι εξαιρετικά σημαντικά για την απόδοση των δικτύων (σταθερότητα και πλαστικότητα), τα οποία επιτρέπουν την εφαρμογή συνεχούς εκπαίδευσης. Το ασαφές ART&ARTMAP νευρωνικό δίκτυο στοχεύει να μειώσει την ανακρίβεια των αποτελεσμάτων πρόβλεψης με ένα μηχανισμό που διαχωρίζει τα αναλογικά και τα δυαδικά δεδομένα, με το να τα επεξεργάζεται χωριστά. Επομένως, αυτό αντιπροσωπεύει μια μείωση στον χρόνο επεξεργασίας και βελτιωμένη ποιότητα των αποτελεσμάτων, σε σύγκριση με το BP (back-propagation) νευρωνικό δίκτυο και καλύτερα από τις κλασσικές μεθόδους πρόβλεψης (ARIMA των Box και Jenkins μεθόδων). Τελειώνοντας την εκπαίδευση, το ασαφές ART&ARTMAP νευρωνικό δίκτυο είναι ικανό να προβλέψει ηλεκτρικά φορτία 24h εκ' των προτέρων. Για να επικυρωθεί αυτή η μεθοδολογία, χρησιμοποιούμε δεδομένα από μια βραζιλιάνικη ηλεκτρική εταιρία.

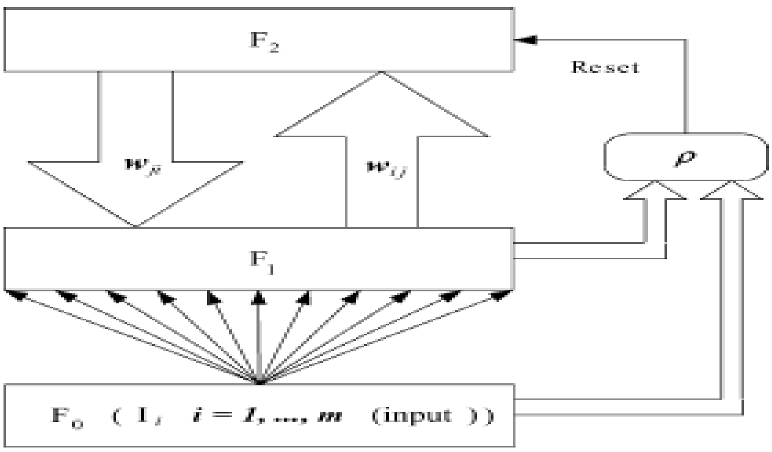


Fig. 1. ART network structure.

Σχήμα 2.21: *Διάταξη ART Δικτύου.*

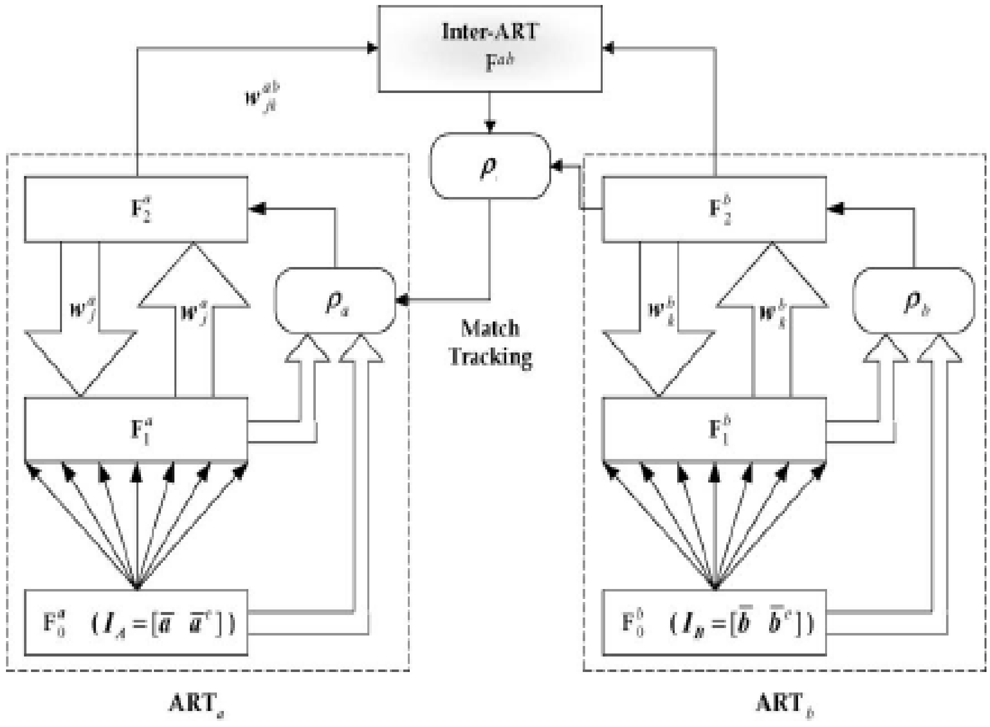


Fig. 2. Structure of the ARTMAP network.

<u>Σχήμα 2.22:</u> *Διάταξη ARTMAP Δικτύου.*

Πρόβλεψη της τάσης στην ζήτηση ηλεκτρικής ενέργειας και των μηνιαίων διακυμάνσεων με νευρωνικά δίκτυα.

Η πρόβλεψη της ζήτησης ηλεκτρικής ενέργειας, είναι βασικό εργαλείο για τις εταιρίες παραγωγής και διανομής επειδή παρέχει σε αυτές μια πρόβλεψη της ζήτησης στην αγορά σε ηλεκτρική ενέργεια. Δύο είναι τα είδη πρόβλεψης που μπορούν να διεξαχθούν: κοντινού και μέσου-μακρού χρονικού ορίζοντα. Αυτή η εργασία εστιάζει στη μηνιαία πρόβλεψη, η οποία είναι χρήσιμη για τον προγραμματισμό της συντήρησης αλλά και ως έρευνα αγοράς για τους παραγωγούς και τους μεταπωλητές της ηλεκτρικής ενέργειας. Η χρονομετρημένη σειρά των μηνιαίων ηλεκτρικών φορτιών παρουσιάζει μια αυξανόμενη τάση λόγω της επιρροής της οικονομικής και τεχνολογικής εξέλιξης στην ηλεκτρική αγορά. Σε αυτήν την γενική τάση είναι ενσωματωμένη μια διακύμανση που προκαλείται από τη διαφορά της ζήτησης από μήνα σε μήνα. Το έγγραφο προτείνει εξαγωγή εκείνης της τάσης έτσι ώστε να εκτελεστούν ξεχωριστά οι προβλέψεις της τάσης και της διακύμανσης με την χρήση νευρωνικών δικτύων, οι οποίες θα αθροιστούν έτσι ώστε να έχουμε πρόβλεψη σειρών. Μεταξύ των διαφορετικού τύπου νευρωνικών δικτύων που παρουσιάζει η βιβλιογραφία, το πολύ-επίπεδο perceptron (MLP) έχει επιλεχτεί για να εκτελέσει την πρόβλεψη, δεδομένου ότι έχει χρησιμοποιηθεί ευρέως για την πρόβλεψη χρονικών σειρών λόγω της δυνατότητάς του να προσδιορίσει το χρόνο εξέλιξης ενός δυναμικού συστήματος.

Το δίκτυο αποτελείται από:

- Το επίπεδο εισαγωγής, του οποίου τα στοιχεία είναι αυτά που υποβάλλονται σε επεξεργασία.
- Το επίπεδο εξόδου το όποιο παρέχει την πρόβλεψη του δικτύου.
- Ένα ή περισσότερα κρυμμένα επίπεδα, τα οποία επεξεργάζονται τις εισερχόμενες πληροφορίες.

Από τα αποτελέσματα που εξάγονται παρατηρούμε ότι η προτεινόμενη μέθοδος παρέχει μια πολύ αξιόπιστη πρόβλεψη, καθώς το σφάλμα πρόβλεψης κυμαίνεται μεταξύ 5% και 7%.

Πρόβλεψη ωριαίας κατανάλωσης ενέργειας σε κτήρια βασιζόμενη σε ένα ανατροφοδοτούμενο ANN.

Στο συγκεκριμένο άρθρο, παρουσιάζεται μια νέα προσέγγιση της βραχυπρόθεσμης πρόβλεψης ηλεκτρικού φορτίου στα κτίρια. Η μέθοδος βασίζεται σε ένα ειδικό είδος των τεχνητών νευρωνικών δικτύων (ANN), τα οποία ανατροφοδοτούν ένα μέρος των εξόδων τους. Το δίκτυο εκπαιδεύεται από ένα υβριδικό αλγόριθμο. Το νέο σύστημα χρησιμοποιεί τις τρέχουσες αλλά και προβλεπόμενες τιμές της θερμοκρασίας, το τρέχων φορτίο και την ημέρα και την ώρα ως εισόδους . Η απόδοση του προβλέπων αξιολογείται με χρήση πραγματικών δεδομένων και αποτελεσμάτων από διεθνείς διαγωνισμούς. Ο τύπος του νευρωνικού δικτύου στην περίπτωση μας, είναι perceptron και αποτελείται από ένα κρυμμένο επίπεδο και νευρώνες με συναρτήσεις μεταφοράς, υπερβολικής εφαπτομένης χάρη στα αναφερόμενα αποτελέσματα για την βελτίωση, σε σχέση με τον απαιτούμενο αριθμό επαναλήψεων κατά την διάρκεια της εκμάθησης ενός νευρωνικού δικτύου που χρησιμοποιεί τον BP (backpropagation) αλγόριθμο. Παρακάτω έχουμε τις σχηματικές απεικονίσεις της δομής ανατροφοδότησης καθώς και της γενικής δομής του προτεινόμενου τεχνητού νευρωνικού δικτύου.

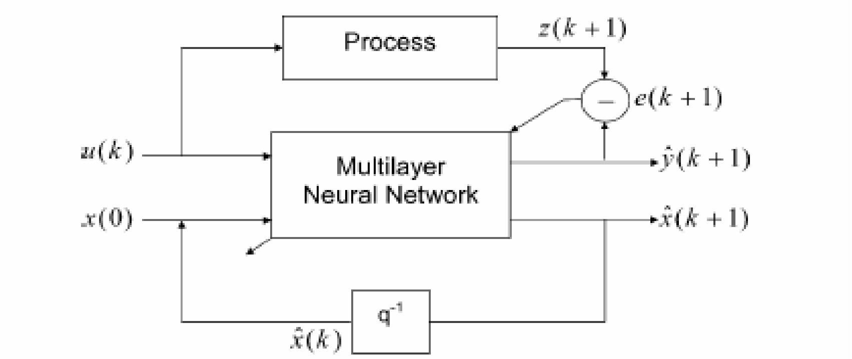


Fig. 1. General schema of the feedback neural network designed in [6].

Σχήμα 2.23: *Γενικό σχήμα της δομής ανατροφοδότησης του τεχνητού νευρωνικού δικτύου.*

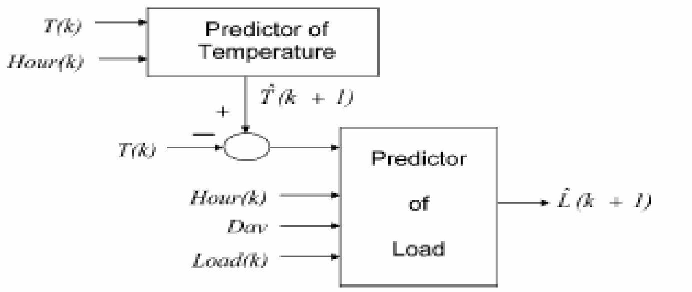


Fig. 2. General Structure of the load predictor.

Σχήμα 2.24: *Γενική δομή του προβλέποντος τεχνητού νευρωνικού δικτύου.*

Μακροπρόθεσμη πρόβλεψη φορτίου με χρήση νευρωνικών δικτύων.

Το συγκεκριμένο άρθρο, βασιζόμενο στα ήδη υπάρχοντα αποτελέσματα (από προηγούμενη έρευνα) τα οποία έχουν αξιολογηθεί ασχολείται με την επίλυση του προβλήματος της πρόβλεψης του μέγιστου ηλεκτρικού φορτίου μακροπρόθεσμα. Δυο τεχνητά νευρωνικά δίκτυα ένα τριών επιπέδων αναστροφής διάδοσης (three-layer back propagation)και ένα επαναληπτικό (recurrent) σχεδιαστήκαν για αυτό το σκοπό. Αντίθετα με την βραχυπρόθεσμη πρόβλεψη φορτίου η μακροπρόθεσμη εξαρτάται κυρίως από οικονομικούς παράγοντες όπως φαίνεται στο *σχήμα 2.25* παρά από τις καιρικές συνθήκες (παράγοντες). Τα οικονομικά αυτά δεδομένα είναι : το ακαθάριστο εθνικό προϊόν, το ακαθάριστο εγχώριο προϊόν, ο πληθυσμός, ο αριθμός των οικογενειών, ο αριθμός των κλιματιστικών, το μέγεθος της ρύπανσης από το CO2, ο δείκτης της βιομηχανικής παραγωγής, η τιμή του πετρελαίου, η κατανάλωση της ενεργείας και η τιμή του ηλεκτρισμού. Τα δεδομένα χρησιμοποιούνται με την μορφή του πραγματικού ποσοστού αύξησης και του επαυξημένου ποσοστού αύξησης σε σχέση με τον προηγούμενο χρόνο . Επίσης δυο ακόμη μελέτες χρησιμοποιούνται για την πρόβλεψη της πολύ μακροπρόθεσμης πρόβλεψης του ηλεκτρικού φορτίου (μέχρι και 2020) : *α) η συντήρηση της θέσης (preservation of the status) και β) η μεταρρύθμιση δομών (structure reform).* Και τα δυο μοντέλα πρόβλεψης δίνουν σχετικά ακριβείς προβλέψεις σε σχέση με τις πραγματικές τιμές και μας δείχνουν ότι η αλλαγές στο φορτίο είναι μια αντανάκλαση της οικονομίας στη περίπτωση μας. Γενικώς, στη μακροπρόθεσμη πρόβλεψη είναι αποδεκτό ένα σφάλμα του10% όμως τα μοντέλα αυτά έχουν την ικανότητα να προβλέψουν με σφάλμα μονό 3% που δηλώνει μεγάλη ακρίβεια.

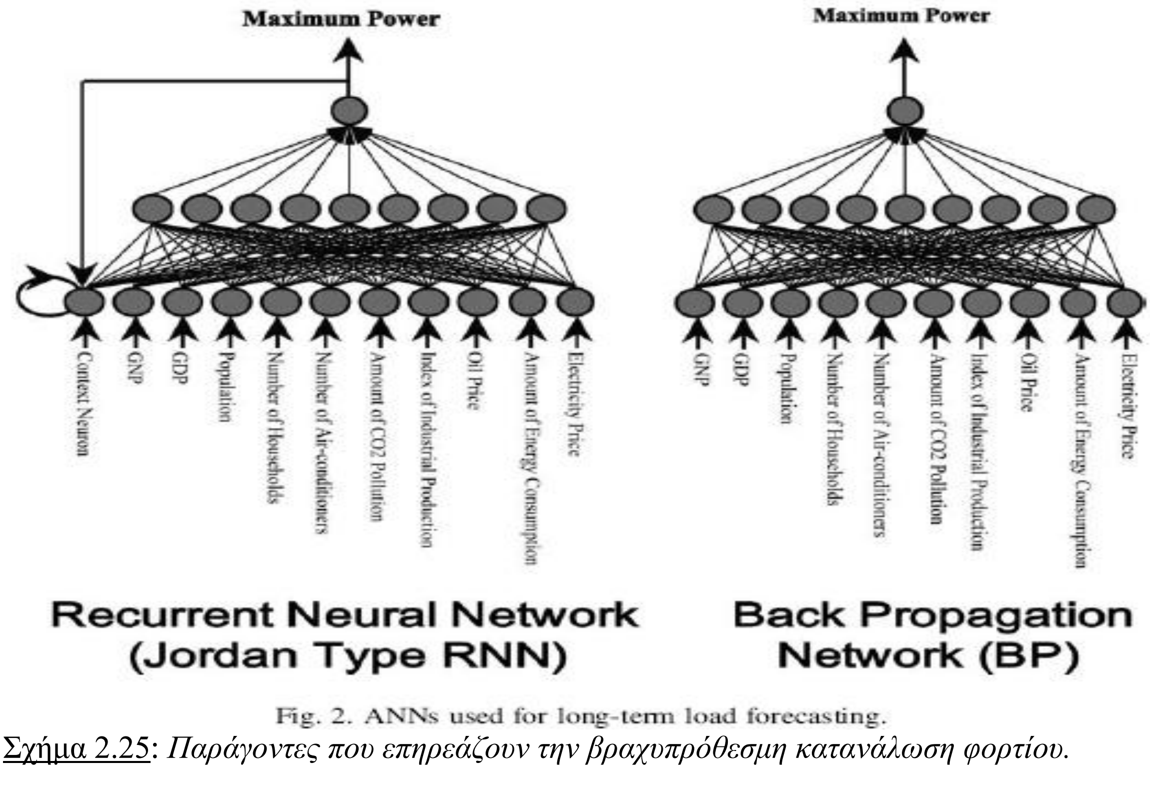


Fig. 2. ANNs used for long-term load forecasting.

Σχήμα 2.25: *Παράγοντες που επηρεάζουν την βραχυπρόθεσμη κατανάλωση φορτίου.*

Σύγκριση Multilayer Perceptron και Radial Basis Function νευρωνικών Δικτύων.

Έχει αποδειχθεί από πλήθος ερευνητών, ότι τα τεχνητά νευρωνικά δίκτυα(ANN) μπορούν να είναι περισσότερο αξιόπιστα και αποτελεσματικά στην πρόβλεψη ενέργειας κτηρίων σε σχέση με τα παραδοσιακά μοντέλα προσομοίωσης. Αυτό το άρθρο, παρουσιάζει τα αποτελέσματα από τις συγκρίσεις ως προς την ακρίβεια πρόβλεψης, δυο ευρέως χρησιμοποιούμενων νευρωνικών δικτύων που χρησιμοποιούνται για την πρόβλεψη κατανάλωσης ηλεκτρικού φορτίου σε καταστήματα λιανικής πώλησης τροφίμων. Τα νευρωνικά δίκτυα που χρησιμοποιούνται είναι το Multilayer Perceptron (MLP) και το Radial Basis Function (RBF). Τα δεδομένα που χρησιμοποιούν τα δυο νευρωνικά μοντέλα ως είσοδο είναι :

- ημέρα της εβδομάδας
- ώρα της ημέρας
- εσωτερική και εξωτερική θερμοκρασία
- εσωτερική και εξωτερική υγρασία

Τα δεδομένα εκπαίδευσης που λαμβάνονται είναι για μια περίοδο δυο μηνών και ακολούθως τα εκπαιδευμένα δίκτυα ελέγχονται για πρόβλεψη ακριβείας πάνω στις επόμενες έξι ημέρες. Το MLP δίκτυο που χρησιμοποιείται είναι τριών επιπέδων με έξι κόμβους στο επίπεδο εισόδου, τέσσερις κόμβους στο μοναδικό κρυμμένο επίπεδο και έναν κόμβο στο επίπεδο εξόδου. Η συνάρτηση μεταφοράς που χρησιμοποιείται είναι η υπερβολική εφαπτομένη και ο χρησιμοποιούμενος αλγόριθμος είναι ο ανάστροφης διάδοσης (back-propagation). Από την άλλη πλευρά το RBF δίκτυο, έχει και αυτό έξι κόμβους στο επίπεδο εισόδου. Ο αριθμός των κέντρων ποικίλει με τιμές μεταξύ του πέντε και του πενήντα με βήμα αύξησης κατά πέντε. Η μη-γραμμική βασική λειτουργία που χρησιμοποιείται είναι η spline. Από τα αποτελέσματα των δυο μοντέλων, τα MLP δίκτυα βρέθηκαν να αποδίδουν καλύτερες προβλέψεις σε σχέση με τα RBF (ιδιαιτέρως στην πρόβλεψη των διακυμάνσεων της ηλεκτρικής ενέργειας γύρω από την βάση αλλά και των μέγιστων φορτίων) καθώς οι τιμές τους είναι πιο κοντά στις πραγματικές και συνεπώς το σφάλμα τους είναι μικρότερο. Στον ακόλουθο πίνακα, διακρίνουμε την διαφορά των σφαλμάτων μεταξύ των δυο μοντέλων.

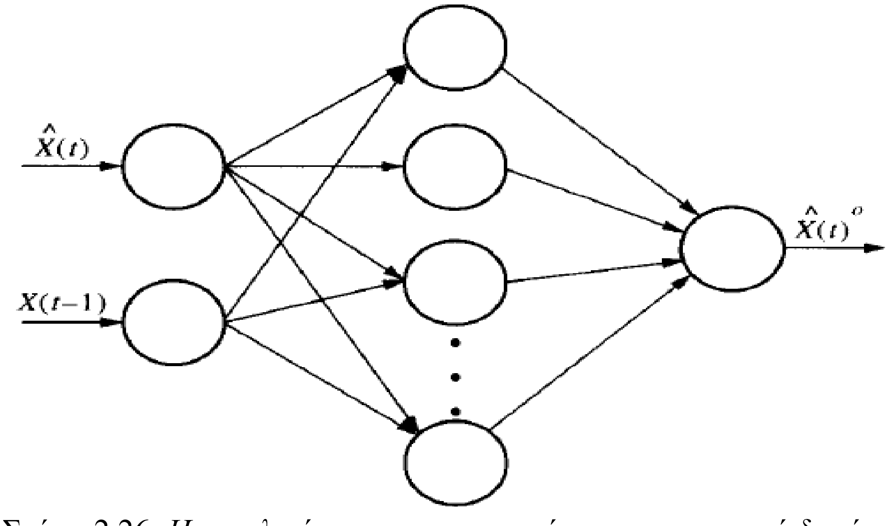
Table 1
Output error measures

Output	RMS error	Mean absolute	Mean absolute (%)
MLP network	24.19	17.38	4.59
RBF network	34.92	26.92	7.11

Πίνακας 2.2: *Σφάλμα πρόβλεψης για το Multilayer Perceptron (MLP) και το Radial Basis Function (RBF) νευρωνικό δίκτυο.*

Συνδυασμός ενός νευρωνικού δικτύου με ένα βασιζόμενο σε κανόνες έμπειρο σύστημα στην βραχυπρόθεσμη πρόβλεψη φορτίου.

Σε αυτό το άρθρο, ένα ανάστροφης διάδοσης (*back-propagation)* νευρωνικό δίκτυο με την έξοδο που παρέχεται από ένα βασιζόμενο σε κανόνες έμπειρο σύστημα σχεδιάζεται για βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φορτίου. Τα ωριαία φορτία για μια τυπική ημέρα αξιολογούνται και σε αυτήν την περίπτωση, ο συνυπολογισμός της πρόβλεψης του βασισμένου σε κανόνες έμπειρου συστήματος βελτιώνει σημαντικά την πρόβλεψη του νευρωνικού δικτύου για το ηλεκτρικό φορτίο. Το προτεινόμενο δίκτυο λοιπόν, αποτελείται από τρία επίπεδα μονάδων, το επίπεδο εισόδου έχει δυο στοιχεία (κόμβους) για να προβλέψει στον χρόνο t. Ο συνυπολογισμός της μεταβλητής εισόδου, X(t - 1) γίνεται εξαιτίας της υψηλής αυτοσυσχέτισης της και βασίζεται στα αποτελέσματα ταυτοποίησης που έγιναν από τον Ho et al. (1990). Σε αυτήν την δουλειά, πέντε ARIMA μοντέλα αναπτύσσονται με χρήση της μεθόδου του Box-Jenkins. Σε αυτήν την έρευνα λοιπόν, τα αποτελέσματα της πρόβλεψης από το βασιζόμενο σε κανόνες έμπειρο σύστημα του Ho et al.(1990), η X(t) χρησιμοποιείται όπως οι άλλες μεταβλητές εισόδου με σκοπό να παρέχει πρόσθετη ακριβή πληροφορία για την εκπαίδευση του δικτύου. Βελτιστοποιώντας αυτό το έμπειρο σύστημα, ο Ho επιστράτευσε την εμπειρία του χειριστή και έναν ικανό κανόνα, για να ελαχιστοποιήσει την επίδραση από τα ειδικά γεγονότα. Καθώς υπάρχουν μόνο δύο κόμβοι εισόδου και ένας κόμβος εξόδου, ο αρχικός αριθμός των κρυφών κόμβων για τον έλεγχο σε αυτό το παράδειγμα επιλέχθηκε να είναι 2, 3, 4, 5 και 6. Το δίκτυο μας απεικονίζεται στο παρακάτω σχήμα.



Σχήμα 2.26: *Η τοπολογία του χρησιμοποιούμενου νευρωνικού δικτύου.*

Όπου:

- *X^ (t)* = το προβλεπόμενο αποτέλεσμα στον χρόνο t παρεχόμενο από το βασισμένο σε κανόνες έμπειρο σύστημα.
- X(t-l) = η πραγματική ισχύς του φορτίου στον χρόνο t-l.
- $X(t)^o$ = το προβλεπόμενο αποτέλεσμα στον χρόνο προερχόμενο από τον συνδυασμό των δύο συστημάτων.

Η προτεινόμενη συνδυαστική προσέγγιση συγκλίνει πολύ γρηγορότερα σε σχέση με το συμβατικό νευρωνικό δίκτυο και την βασισμένη σε κανόνες μέθοδο έμπειρου συστήματος.

Τεχνητά νευρωνικά δίκτυα εξελισσόμενα για βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φορτίου.

Αυτό το άρθρο, παρουσιάζει τεχνικά νευρωνικά δίκτυα (ANN) τα οποία εξελίσσονται από έναν γενετικό αλγόριθμο με σκοπό την βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φόρτου. Αυτό επιτυγχάνεται, χρησιμοποιώντας τις πραγματικές τιμές φορτίων καθώς και τα προβλεπόμενα καιρικά δεδομένα. Τα ANN δίκτυα έχουν εκπαιδευτεί για την ακριβή πρόβλεψη του ωριαίου ηλεκτρικού φορτίου μια ημέρα μπροστά σε οποιαδήποτε ημέρα της εβδομάδας. Τα ANN δίκτυα που έχουν εκπαιδευθεί και εξελιχθεί από τον γενετικό αλγόριθμο δείχνουν να έχουν ισχυρό πλεονέκτημα έναντι μιας στατιστικής μεθόδου. Η απόδοση ενός ANN δικτυού δεν εξαρτάται μόνο από τις επιλογές των εισαγόμενων μεταβλητών, αλλά επίσης και από την διάσταση του. Το προτεινόμενο μοντέλο πρόβλεψης ηλεκτρικών φορτίων είναι τριών επιπέδων με ανατροφοδότηση το οποίο εκπαιδεύεται από το γενετικό αλγόριθμο. Στην αρχή της διαδικασίας εκπαίδευσης, η αρχιτεκτονική του μοντέλου διαμορφώνεται όπως στο παρακάτω σχήμα και τα βάρη παίρνονται σύμφωνα με τους δοσμένους περιορισμούς. Οι είσοδοι που απεικονίζονται στο *σχήμα 2.27* είναι:

- η ωριαία θερμοκρασία της ημέρας πρόβλεψης.
- οι μέσες ωριαίες τιμές φορτίων που προκύπτουν από τον υπολογισμό του μέσου όρου των φορτιών των τριών προηγούμενων εβδομάδων για τον ίδιο τύπο μέρας.
- η μέση τιμή των φορτίων για την προηγούμενη ώρα, που προκύπτει από τον μέσο όρο των δεδομένων των τριών προηγούμενων εβδομάδων.

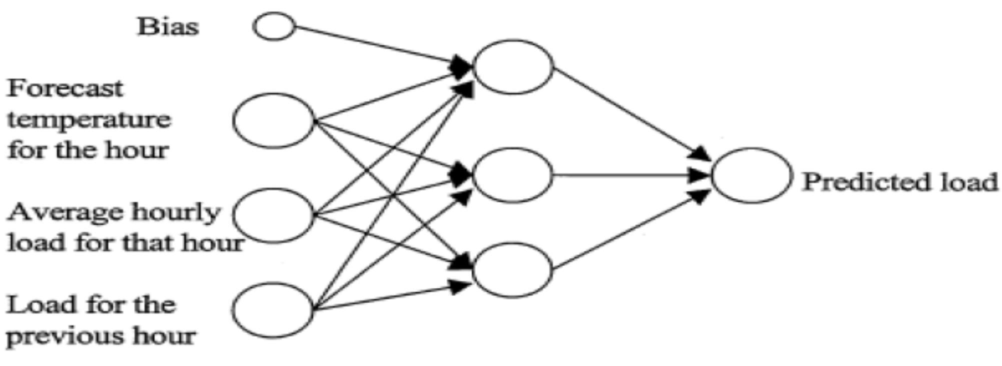


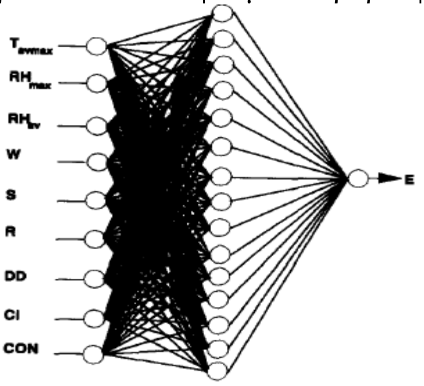
Fig. 1. The ANN model.

Σχήμα 2.27: *Απεικόνιση εισόδων νευρωνικού μοντέλου.*

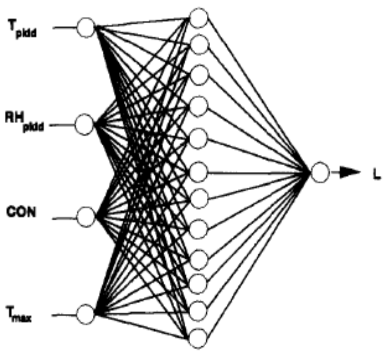
Επιπλέον, αυτή η προσέγγιση αποφεύγει την χρήση μεγάλου ιστορικού δεδομένων και την συχνή επανεκπαίδευση. Ένα σημαντικό μειονέκτημα αυτού του μοντέλου είναι ότι δεν είναι σε θέση να ανιχνεύσει ξαφνικές αλλαγές του φορτίου που μπορεί να συμβούν.

Τεχνητά νευρωνικά δίκτυα μέσου χρονικού ορίζοντα.

Σε αυτό το άρθρο, παρουσιάζονται τα νέα τεχνητά νευρωνικά δίκτυα (ANN) τα οποία βασιζόμενα στην σχέση καιρού-φορτίου και καιρού-ενέργειας εξελίχθηκαν να προβλέπουν το ηλεκτρικό φορτίο και την ενέργεια για 24 μήνες μπροστά, δηλαδή είναι μέσου χρονικού ορίζοντα (MTLF). Η κατανάλωση φορτίου και ενέργειας, εξαρτάται κυρίως από τον καιρό όπως και η αύξηση του φορτίου και της ενεργειακής απαίτησης εξαρτάται κατά ένα μεγάλο μέρος από τον αριθμό των καταναλωτών που συνδέονται στο σύστημα. Τα δεδομένα του μοντέλου και στις δυο περιπτώσεις (του φορτίου και της ενεργείας) αποτελούνται α) από *ένα σύνολο μεταβλητών του καιρού* όπως η απόλυτη μέγιστη θερμοκρασία, η μέση μέγιστη θερμοκρασία, η μέση μέγιστη σχετική υγρασία, η μέση σχετική υγρασία, η ταχύτητα του ανέμου, η διάρκεια της ηλιοφάνειας, η ατμοσφαιρική ακτινοβολία, οι βροχοπτώσεις και η ατμοσφαιρική πίεση από τις οποίες γίνεται επιλογή των μεταβλητών που θα χρησιμοποιηθούν (με κριτήρια τη συσχέτιση που έχουν με το φορτίο η την ενεργεία αλλά και την ανθρώπινη διαίσθηση) καθώς και β) *ιστορικά στοιχειά του φορτίου και της ενεργειακής απαίτησης* από προηγούμενα χρονιά (1986 έως 1990) από το σύστημα ισχύος μιας ηλεκτρικής εταιρίας. Με αυτά κάνει πρόβλεψη μέχρι και δυο χρόνια μπροστά (24 μήνες). Τα αποτελέσματα της πρόβλεψης σε σύγκριση με τα πραγματικά στοιχεία για το 1991-1992, δείχνουν ότι η μηνιαία απαιτούμενη ηλεκτρική ενέργεια και το φορτίο μπορούν να προβλεφθούν με ένα μέγιστο λάθος 6% και 10%, αντίστοιχα, για τις δυο περιπτώσεις ακόμη και με προβλεπόμενα στοιχειά του καιρού. Τα προτεινόμενα νευρωνικά μοντέλα (ANN), παρέχουν αποτελέσματα με μεγαλύτερη ακρίβεια από τα προηγουμένως αναπτυγμένα μοντέλα όπως φαίνεται και στον παρακάτω πίνακα όπου παρατίθενται τα σφάλματα πρόβλεψης των μεθόδων.



Σχήμα 2.28: *ANN μοντέλο βασισμένο στην ενέργεια.*



Σχήμα 2.29: *ANN μοντέλο βασισμένο στο φορτίο.*

Error (%) comparison in energy forecasting

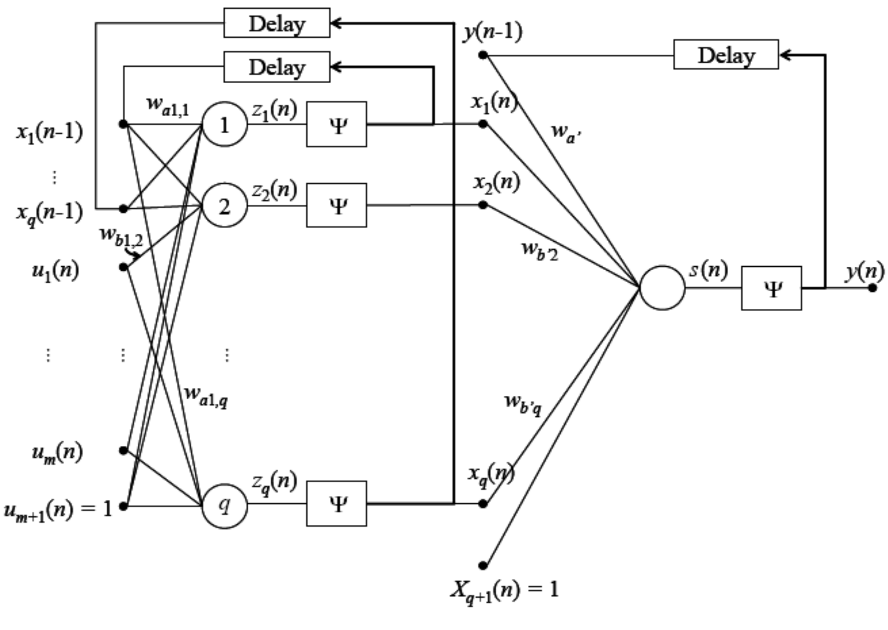
Month in 1991	Actual	ANN forecast	ANN error	SE model forecast	SE model error
January	134600	130314	(3.18)	138698	3.04
February	122000	124722	2.23	135768	11.29
March	161100	159103	(1.24)	171529	6.47
April	242800	247781	2.05	248880	2.50
May	296100	289318	(2.29)	276390	(6.66)
June	346500	355173	2.50	308607	(10.94)
July	372100	367742	(1.17)	310459	(16.57)
August	326900	320375	(2.00)	250919	(23.24)
September	307800	311567	1.22	259899	(15.56)
October	257200	249171	(3.12)	234642	(8.77)
November	179500	185340	3.25	187997	4.37
December	157600	157718	0.07	156754	(0.54)

R^2 (ANN) = 0.996; R^2 (SE model) = 0.917.

Πίνακας 2.2: *Σφάλμα πρόβλεψης των αναφερόμενων μεθόδων.*

Υβριδικά νευρωνικά δίκτυα κοντινού χρονικού ορίζοντα.

Ο σκοπός αυτού του άρθρου, είναι η πρόβλεψη της απαίτησης σε ηλεκτρικό φορτίο χρησιμοποιώντας υβριδικά νευρωνικά δίκτυα. Τα δίκτυα αυτά είναι κοντινής διάρκειας δηλαδή έχουν σύντομο χρονικό ορίζοντα πρόβλεψης. Η υλοποίηση γίνεται σε πραγματικά παρελθοντικά δεδομένα της τούρκικης εταιρίας ηλεκτρισμού. Η συνολική διαδικασία έχει ως εξής: Ξεκινώντας με τυχαία βάρη και παίρνοντας μη-ρεαλιστικές προβλέψεις, οι οποίες δεν θα μπορούσαν να γίνουν αποδεκτές σε μια πραγματικού χρόνου εφαρμογή, στην συνέχεια τα διαθέσιμα παρελθοντικά δεδομένα χρησιμοποιούνται off-line και το μοντέλο προετοιμάζεται για τις on-line προβλέψεις. Επιπλέον, μια μέθοδος που βασίζεται στην ωριαία κατανάλωση ενέργειας προτείνεται για να ομαδοποιήσει τα εισερχόμενα δεδομένα. Πρακτικά πρόκειται για επαναλαμβανόμενα νευρωνικά δίκτυα τα οποία περιέχουν ένα ή περισσότερους βρόχους ανατροφοδότησης, οι οποίοι συνεισφέρουν στην μείωση της απαιτούμενης μνήμης. Στο ακόλουθο σχήμα, έχουμε απεικόνιση του μοντέλου καθώς και των βρόχων ανατροφοδότησης .



Σχήμα 2.30: *Επαναλαμβανόμενη MLP κατασκευή.*

Στο εξεταζόμενο μοντέλο υπάρχει μονό μια μεταβλητή y(n) για το ηλεκτρικό φορτίο που προβλέπεται κάθε φορά αυτό και έχει μονό μια έξοδο. Το μέσο σφάλμα των δικτύων είναι 2,31%, εδώ φαίνεται η επιτυχία της off-line εκπαίδευσης του μοντέλου. Τα υβριδικά αυτά μοντέλα δίνουν καλύτερα αποτελέσματα από τα απλά μοντέλα διότι έχουν το πλεονέκτημα να χρησιμοποιούν την εμπειρία από τα δεδομένα του παρελθόντος ενώ μειώνουν και το χρόνο πρόβλεψης, αφού τα βάρη που χρησιμοποιούνται είναι ήδη κοντά στις βέλτιστες τιμές τους. Τα μοντέλα αυτά προσαρμόζονται πολύ γρήγορα σε πραγματικές καταστάσεις και δεν ταλαντώνονται πολύ σε απότομες αλλαγές τους περιβάλλοντος ενώ τα αποτελέσματα που μας δίνουν έχουν μικρότερο σφάλμα σε σύγκριση με τα απλά, οπού τα στοιχειά δεν ομαδοποιούνται. Γενικώς, τα υβριδικά μοντέλα χρησιμοποιούνται σε εφαρμογές οπού υπάρχουν διαθέσιμα ιστορικά στοιχεία και η ενημέρωση των βαρών γίνεται επαναλαμβάνοντας την off-line φάση σε συγκεκριμένες περιόδους προσθέτοντας τα πρόσφατα στοιχεία.

Ωριαία πρόβλεψη του Αιγυπτιακού ενωμένου δικτύου ηλεκτρικής ισχύος με χρήση νευρωνικών δικτύων

Αυτό το άρθρο, παρουσιάζει ένα τεχνητό νευρωνικό δίκτυο (ANN) βασιζόμενο στην εφαρμογή της ωριαίας πρόβλεψης φορτίου στο Αιγυπτιακό ενωμένο δίκτυο ηλεκτρικής ισχύος (EUG). Το ANN σχεδιάστηκε με χρήση της πολύ-επίπεδης ανάστροφης διάδοσης (multilayer back propagation), τεχνικής εκπαίδευσης. Στο επίπεδο εισόδου, λαμβάνονται όλες τις κατάλληλες πληροφορίες που μπορούν να συμβάλουν καθοριστικά στην διαδικασία πρόβλεψης, αποκλείοντας τις πληροφορίες του καιρού. Πιο συγκεκριμένα, το επίπεδο εισόδου λαμβάνει πληροφορίες σχετικά με: την κατηγορία του τύπου της ημέρας, την ώρα στον χρόνο της ημέρας, το φορτίο στην προηγούμενη ώρα, το φορτίο της ίδιας ώρας στην προηγούμενη ημέρα, το μέσο φορτίο της προηγούμενης ημέρας, το φορτίο αιχμής της προηγούμενης ημέρας, το ελάχιστο φορτίο της προηγούμενης ημέρας και τέλος τις τελευταίες τέσσερις μετρήσεις της προηγούμενης εβδομάδας. Το επίπεδο εξόδου παρέχει τα ωριαία φορτία που έχουν προβλεφθεί. Το ANN μοντέλο πρόβλεψης εκπαιδεύεται βασιζόμενο σε έναν ιστορικό τομέα γνώσης και οι δοκιμές αποδεικνύουν ότι μπορεί να προβλέψει αποτελεσματικά και με ακρίβεια ωριαία φορτία. Τα τελικά αποτελέσματα έδειξαν μέσα σφάλματα της τάξης 0.18% (εκπαίδευση) και 0.49% (αξιολόγηση) και τυπικές αποκλίσεις της τάξης 2.32 και 2.92% αντίστοιχα. Το ANN μοντέλο συγκρινόμενο με μια στατιστική μέθοδο (MLR- multilinear regression) όπως φαίνεται και στον παρακάτω πίνακα, δείχνει να έχει πολύ καλύτερη απόδοση καθώς το μέγιστο σφάλμα είναι αρκετά μικρότερο (το ANN 6.708% και το MLR 8.614%).

Table 2. Forecasting % error results for ANN and MLR methods

Hour	Friday 22 January			Sunday 31 January		
	Load	ANN	MLR	Load	ANN	MLR
1	5202	0.700	−0.82132	5014	1.213	−2.30024
2	4702	1.270	0.380072	4751	0.578	−2.00808
3	4200	6.708	7.204405	4548	0.165	−2.55031
4	4464	−2.194	2.722625	4475	0.460	−1.88342
5	4425	0.636	4.700949	4600	−2.503	−4.85493
6	4678	1.493	5.305857	4953	−4.026	−6.83489
7	4703	1.387	5.250542	5331	−5.617	−8.57229
8	4604	3.587	7.709427	5106	1.604	−2.64898
9	4800	1.762	5.908479	5056	2.427	−1.58113
10	5001	0.149	3.058468	5118	2.844	−0.78271
11	4988	2.849	5.00816	5103	3.811	0.005154
12	5152	0.156	2.465295	5032	5.359	1.552266
13	4856	6.041	8.614271	5087	4.098	0.561176
14	4835	5.260	8.359979	5112	2.968	−0.67181
15	4807	3.912	7.349449	5267	−0.552	−3.81124
16	4872	3.749	7.280911	5522	−2.859	−6.20681
17	5245	1.455	4.601449	5577	−0.319	−3.58614
18	6642	0.316	3.296582	6957	−1.977	−3.25436
19	6640	4.362	5.813946	7026	0.550	−1.3376
20	6457	4.562	6.41781	6925	0.278	−1.53479
21	6227	5.283	7.866436	6680	1.410	−0.49075
22	6030	1.677	4.987081	6246	1.491	−0.77464
23	5730	−1.476	1.999651	5627	4.040	2.198294
24	5257	0.893	5.850257	5328	0.007	−0.96368

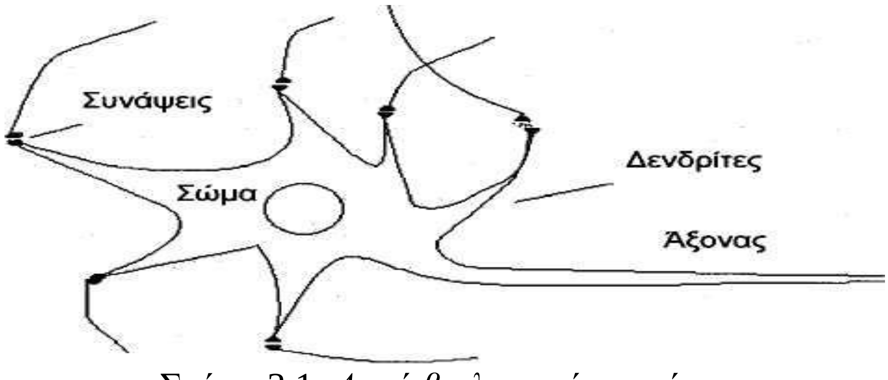
Πίνακας 2.3: *Σφάλμα πρόβλεψης της ANN και της MLR μεθόδου.*

ΚΕΦΑΛΑΙΟ 3

Εισαγωγή στη θεωρία των τεχνητών νευρωνικών δικτύων

3.1 Γενικά χαρακτηριστικά των νευρωνικών δικτύων

Η ικανότητα του ανθρώπου να σκέπτεται , να θυμάται και να επιλύει προβλήματα εντοπίζεται στον εγκέφαλό του. Όπως είναι γνωστό από τη βιολογία , η δομική μονάδα του εγκεφάλου είναι ο νευρώνας. Ένας τυπικός βιολογικός νευρώνας αποτελείται , όπως φαίνεται και από το παρακάτω σχήμα από το σώμα που περιλαμβάνει τον πυρήνα του, του δενδρίτες μέσω των οποίων λαμβάνει σήματα από γειτονικούς νευρώνες (σημεία εισόδου) και τον άξονα που είναι η έξοδος του νευρώνα και το μέσο σύνδεσής του με τους άλλους νευρώνες. Σε κάθε δενδρίτη υπάρχει ένα απειροελάχιστο κενό που ονομάζεται σύναψη. Οι συνάψεις μέσω χημικών διαδικασιών επιταχύνουν ή επιβραδύνουν τη ροή ηλεκτρονίων προς το σώμα του νευρώνα. Η ικανότητα μάθησης και μνήμης που παρουσιάζει ο εγκέφαλος οφείλεται στην ικανότητά των συνάψεων να μεταβάλλουν την αγωγιμότητά τους. Τα ηλεκτρικά σήματα που εισέρχονται στο σώμα του νευρώνα μέσω των δενδριτών συνδυάζονται και αν το αποτέλεσμα ξεπερνά κάποια τιμή κατωφλίου το σήμα διαδίδεται με τη βοήθεια του άξονα και στους άλλους νευρώνες.



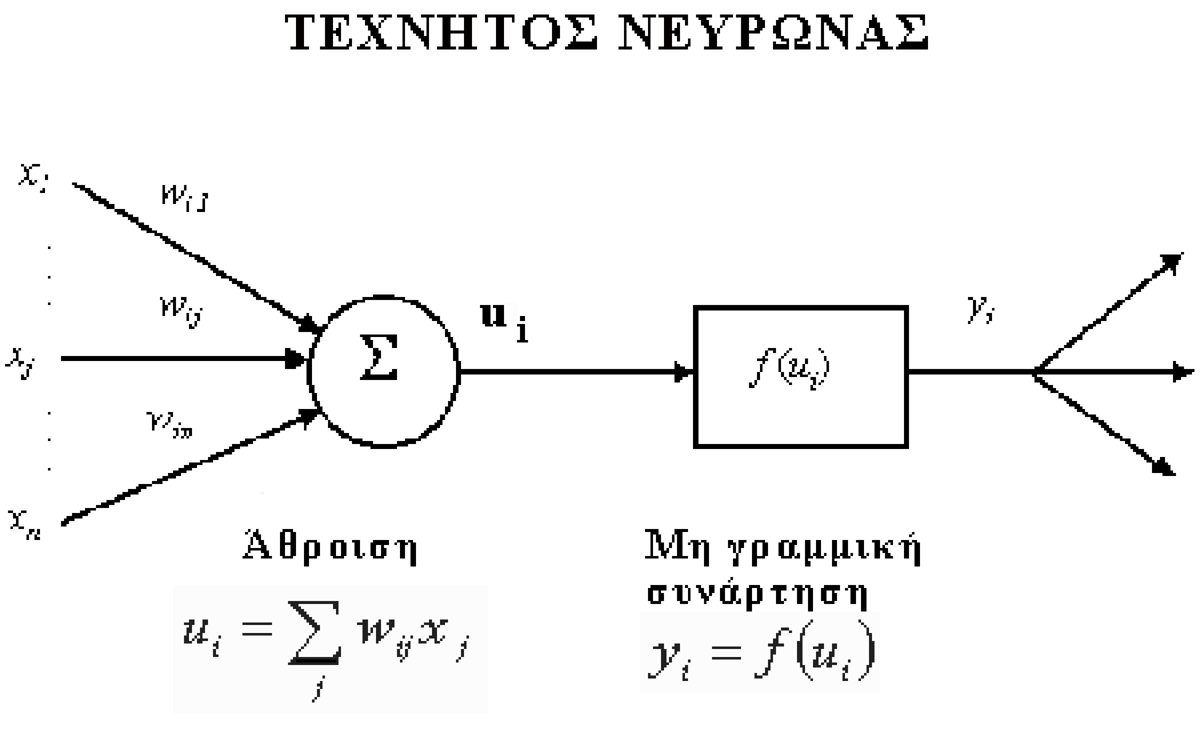
<u>Σχήμα 3.1:</u> *Δομή βιολογικού νευρώνα.*

Τα τεχνητά νευρωνικά δίκτυα είναι εμπνευσμένα από το βιολογικό νευρωνικό σύστημα. Τα στοιχεία που αποτελούν ένα τεχνητό νευρωνικό δίκτυο είναι οι νευρώνες οι οποίοι συνδέονται μεταξύ τους μέσω βαρών. Τα βάρη αυτά μπορούν να προσαρμόζονται κατά τη διαδικασία της μάθησης.

Κάθε νευρώνας αποτελείται από τα ακόλουθα βασικά στοιχεία:

- Ένα σύνολο από συνάψεις κάθε μία από τις οποίες συνδέει το νευρώνα με μία είσοδο και της αντιστοιχείται κάποιο βάρος.
- Μία συγκεκριμένη συνάρτηση μεταφοράς η οποία χρησιμοποιεί ως παραμέτρους τα βάρη των συνάψεων και εφαρμόζεται πάνω στις εισόδους των νευρώνων.
- Μία συνάρτηση ενεργοποίησης η οποία παίρνει ως είσοδο το αποτέλεσμα της συνάρτησης μεταφορά και παράγει την έξοδο του νευρώνα. Συνήθως σκοπός της συνάρτησης ενεργοποίησης είναι να περιορίζει το βαθμό ενεργοποίησης του νευρώνα.

Στο παρακάτω σχήμα απεικονίζεται η μορφή ενός τεχνητού νευρώνα:



<u>Σχήμα 3.2:</u> *Δομή τεχνητού νευρώνα.*

Τα τεχνητά νευρωνικά δίκτυα αναπτύχθηκαν ιδιαίτερα κατά τη δεκαετία του 80 και βρίσκουν σήμερα εφαρμογή σε πάρα πολλά προβλήματα. Μερικά από αυτά παρουσιάζονται παρακάτω.

- Κατηγοριοποίηση προτύπων

Το έργο της κατηγοριοποίησης έχει να κάνει με την κατάταξη σε κάποια κατηγορία ενός διανύσματος εισόδου που περιλαμβάνει δείγματα ενός πρωτογενούς ή επεξεργασμένου σήματος.

- Διαμόρφωση συναθροίσεων (clustering)

Η διαδικασία αυτή μοιάζει με την προηγούμενη με τη μόνη διαφορά ότι στη διαμόρφωση συναθροίσεων δεν είναι εκ των προτέρων γνωστός ο αριθμός των κατηγοριών.

- Προσέγγιση συναρτήσεων

Με δεδομένα ζεύγη εισόδων-εξόδων που είναι δείγματα από μια άγνωστη συνάρτηση y=f(x) ένα τεχνητό νευρωνικό δίκτυο μπορεί να εκπαιδευτεί ώστε να προσεγγίσει αυτή τη συνάρτηση.

- Πρόβλεψη Χρονοσειρών

Μέσω των τεχνητών νευρωνικών δικτύων μπορούν να προβλεφθούν οι μελλοντικές τιμές ενός σήματος ,δεδομένων των τιμών του σήματος σε χρονικές τιμές του παρελθόντος.

- Βελτιστοποίηση

Τα τεχνητά νευρωνικά δίκτυα εφαρμόζονται σε προβλήματα βελτιστοποίησης. Ένα χαρακτηριστικό παράδειγμα είναι το πρόβλημα του πωλητή που ταξιδεύει (Travelling salesman) , που είναι πρόβλημα βέλτιστης διαδρομής σε γράφο.

- Συσχετιστική Μνήμη

Μέσω των τεχνητών νευρωνικών δικτύων μπορούν υλοποιηθούν συστήματα συσχετιστικής μνήμης στα οποία η ανάκληση υλοποιείται μέσω παρουσίασης του περιεχόμενου της μνήμης.

- Αυτόματος έλεγχος

Τα τεχνητά νευρωνικά δίκτυα εφαρμόζονται σε ευρεία κλίμακα σε συστήματα αυτομάτου ελέγχου, όπως π.χ. έλεγχο κίνησης μηχανών , έλεγχο ρομποτικών μελών κτλ.

Τα τεχνητά νευρωνικά δίκτυα συνιστούν μια νέα τεχνική η οποία έχει εφαρμοσθεί επιτυχώς στην βραχυπρόθεσμη πρόβλεψη του φορτίου. Είναι γνωστό ότι υπάρχει μια αδυναμία εύρεσης αυστηρά καθορισμένων μαθηματικών σχέσεων που να προσδιορίζουν το φορτίο των επόμενων ημερών. Μέσω των νευρωνικών δικτύων όμως μπορούμε να κάνουμε πρόβλεψη χωρίς να γνωρίζουμε με ποιο ακριβώς τρόπο επηρεάζει κάθε παράμετρος το φορτίο.

Τα βασικά τους πλεονεκτήματα στην πρόβλεψη του φορτίου έναντι των κλασσικών στατιστικών τεχνικών που αναφέραμε παραπάνω είναι ότι μπορούν να αναγνωρίσουν μη γραμμικές διαχωριστικές επιφάνειες ανάμεσα στις εισόδους και μπορούν ακόμα να λειτουργήσουν ικανοποιητικά με μη πλήρη δεδομένα. Επιπλέον τα νευρωνικά δίκτυα δεν προγραμματίζονται αλλά μαθαίνουν. Έτσι δε χρειάζεται να έχουμε κατανοήσει πλήρως ένα πρόβλημα για να το επιλύσουμε μέσω των νευρωνικών δικτύων. Βέβαια πρέπει να σημειώσουμε ότι τα τεχνητά νευρωνικά δίκτυα δεν μπορούν να δώσουν απάντηση στο πως έφτασαν σε ένα αποτέλεσμα, δεν είναι δηλαδή κατάλληλα για εξαγωγή συμπερασμάτων (extrapolation) ,αλλά μόνο για παρεμβολή (interpolation).

Τα τεχνητά νευρωνικά δίκτυα μπορούν να διακριθούν σε δύο σημαντικές κατηγορίες ανάλογα με τον τρόπο με τον οποίο γίνεται η διαδικασία της μάθησης. Έτσι υπάρχουν τα δίκτυα στα οποία η μάθηση γίνεται *με επίβλεψη (supervised learning)* και τα δίκτυα στα οποία η *μάθηση γίνεται χωρίς επίβλεψη (unsupervised learning).* Παρακάτω εξηγούνται τα χαρακτηριστικά της κάθε κατηγορίας.

ΔΙΚΤΥΑ ΜΕ ΕΠΙΒΛΕΨΗ (SUPERVISED)

Όταν η μάθηση γίνεται με τη διαδικασία της επίβλεψης , το δίκτυο τροφοδοτείται με παραδείγματα εισόδων και τις αντίστοιχες επιθυμητές εξόδους. Κατά τη διαδικασία της μάθησης τα βάρη που συνδέουν τους νευρώνες προσαρμόζονται έτσι ώστε η διαφορά ανάμεσα στην έξοδο του δικτύου και την επιθυμητή έξοδό του να ελαχιστοποιείται.

ΔΙΚΤΥΑ ΧΩΡΙΣ ΕΠΙΒΛΕΨΗ (UNSUPERVISED)

Τα δίκτυα στα οποία η μάθηση γίνεται χωρίς επίβλεψη, τροφοδοτούνται μόνο με εισόδους και όχι με αντίστοιχες επιθυμητές εξόδους. Ο μηχανισμός της μάθησης στην περίπτωση αυτή συμπεριλαμβάνει την ομαδοποίηση ομοειδών δεδομένων.

Μελετούμε παρακάτω διάφορους σημαντικούς τύπους νευρωνικών δικτύων.

3.2 Δίκτυο εμπρόσθιας τροφοδότησης (feedforward network)

Στο δίκτυο αυτά τα δεδομένα εισόδου τροφοδοτούνται στο δίκτυο από τους νευρώνες εισόδου και μεταφέρονται προς τα επόμενα στρώματα του δικτύου. Στο τελευταίο στρώμα του δικτύου παράγεται η έξοδός του. Έτσι έχουμε μια απεικόνιση των εισόδων στις εξόδους. Η απεικόνιση αυτή γίνεται μέσω κάποιων παραμέτρων που δεν είναι άλλες από τα βάρη που συνδέουν τους νευρώνες μεταξύ τους. Στην κατηγορία των νευρωνικών δικτύων εμπρόσθιας τροφοδότησης ανήκει το πολυεπίπεδο perceptron (multi-layer perceprton , MLP) το οποίο χρησιμοποιούμε για να κάνουμε τις προβλέψεις μας. Για το λόγο αυτό θα αναλύσουμε διεξοδικότερα παρακάτω το MLP.

3.2.1 Multi-layer perceptron

Το πολυεπίπεδο perceptron συνιστά μια υποκατηγορία των feedforward networks και είναι ένα από τα δημοφιλέστερα νευρωνικά δίκτυα, αφού χρησιμοποιείται πολύ συχνά και αναφέρεται σε πάρα πολλές δημοσιεύσεις. Η βασική συστατική του μονάδα είναι το perceptron. Το perceptron είναι μια υπολογιστική μονάδα η οποία παίρνει το γραμμικό συνδυασμό των εισόδων της και μέσω της συνάρτησης ενεργοποίησης (activation function) παράγει την έξοδό της. Η έξοδος επομένως ενός perceptron θα δίδεται από την παρακάτω μαθηματική σχέση:

$$o = f\left(\sum_{i=1}^{n} w_i x_i - b\right) \qquad (3.1)$$

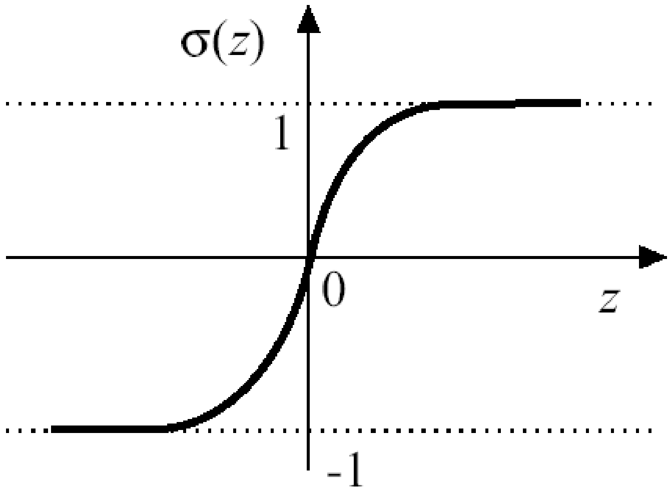
όπου:

- o η έξοδος
- x_i οι είσοδοι
- w_i τα βάρη των νευρώνων
- b ο συντελεστής πόλωσης
- f η συνάρτηση ενεργοποίησης.

Η μορφή της συνάρτησης ενεργοποίησης παίζει σημαντικό ρόλο και πρέπει να επιλεγεί κατάλληλα ανάλογα με το πρόβλημα που αντιμετωπίζουμε κάθε φορά. Υπάρχουν διάφορες μορφές που συνήθως παίρνει η συνάρτηση ενεργοποίησης όπως γραμμική, βηματική, λογαριθμική, υπερβολική εφαπτομένη.

Υπερβολική εφαπτομένη

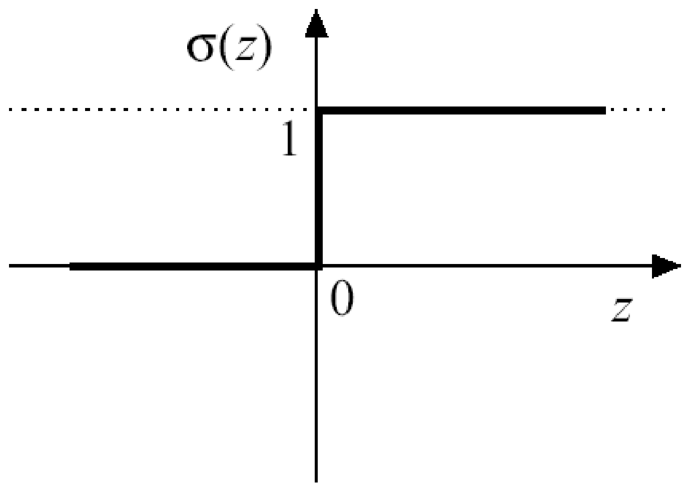
Μια συνάρτηση ενεργοποίησης που χρησιμοποιείται πολύ συχνά είναι η υπερβολική εφαπτομένη. Τη συνάρτηση αυτή θα χρησιμοποιήσουμε και στο τεχνητό νευρωνικό δίκτυο που θα κάνει βραχυπρόθεσμη πρόβλεψη φορτίου. Στο παρακάτω διάγραμμα φαίνεται η μορφή της:



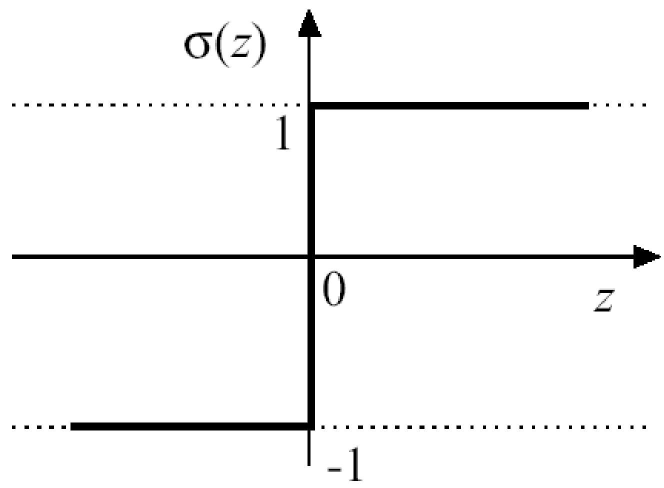
Σχήμα 3.3: *Συνάρτηση ενεργοποίησης τύπου Υπερβολική εφαπτομένη.*

Βηματική συνάρτηση

Σε περιπτώσεις που επιθυμούμε να κάνουμε διαχωρισμό των δεδομένων εισόδου ενδείκνυται να χρησιμοποιήσουμε τη βηματική συνάρτηση η οποία μπορεί να πάρει δύο μορφές. Η μία είναι η μη συμμετρική και η δίνει τιμές 0 ή 1 και η άλλη είναι συμμετρική και δίνει τιμές από το –1 στο +1. στο παρακάτω διάγραμμα φαίνονται και οι δύο αυτές συναρτήσεις:



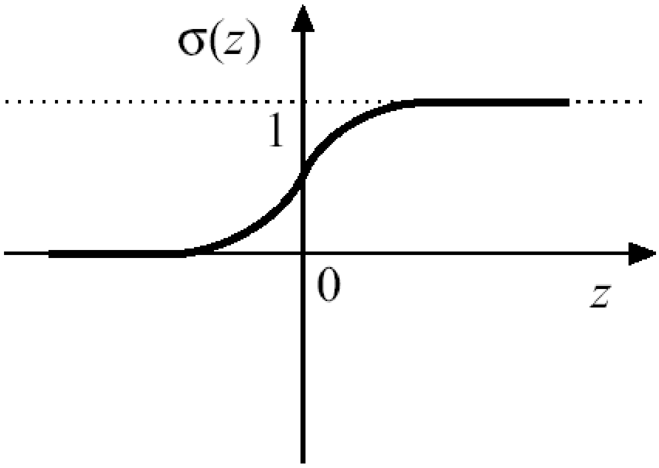
Σχήμα 3.4: *Μη συμμετρική βηματική συνάρτηση ενεργοποίησης.*



Σχήμα 3.5: *Συμμετρική βηματική συνάρτηση ενεργοποίησης.*

Λογαριθμική συνάρτηση

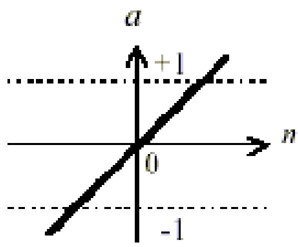
Η λογαριθμική συνάρτηση μοιάζει στη μορφή της με την συνάρτηση της υπερβολικής εφαπτομένης με τη διαφορά ότι η λογαριθμική συνάρτηση δίνει τιμές από το 0 έως το 1 , ενώ η υπερβολική εφαπτομένη από το –1 έως το 1. Στο παρακάτω διάγραμμα βλέπουμε τη μορφή της λογαριθμικής συνάρτησης



Σχήμα 3.6: *Λογαριθμική συνάρτηση ενεργοποίησης.*

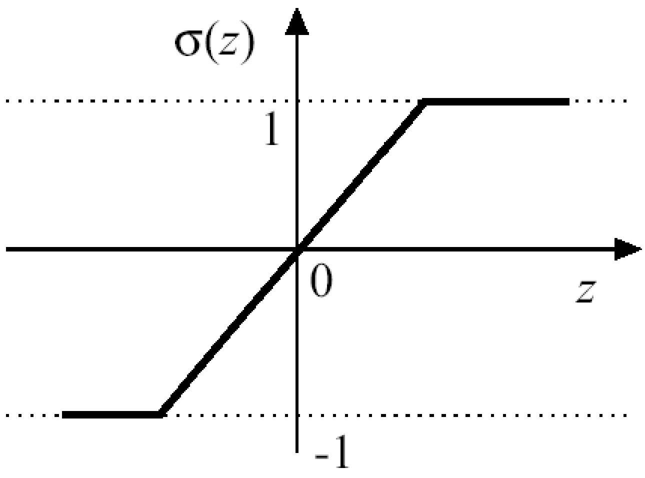
Γραμμική συνάρτηση

Μια απλή μορφή συνάρτησης ενεργοποίησης είναι η γραμμική. Αυτή έχει την παρακάτω μορφή:



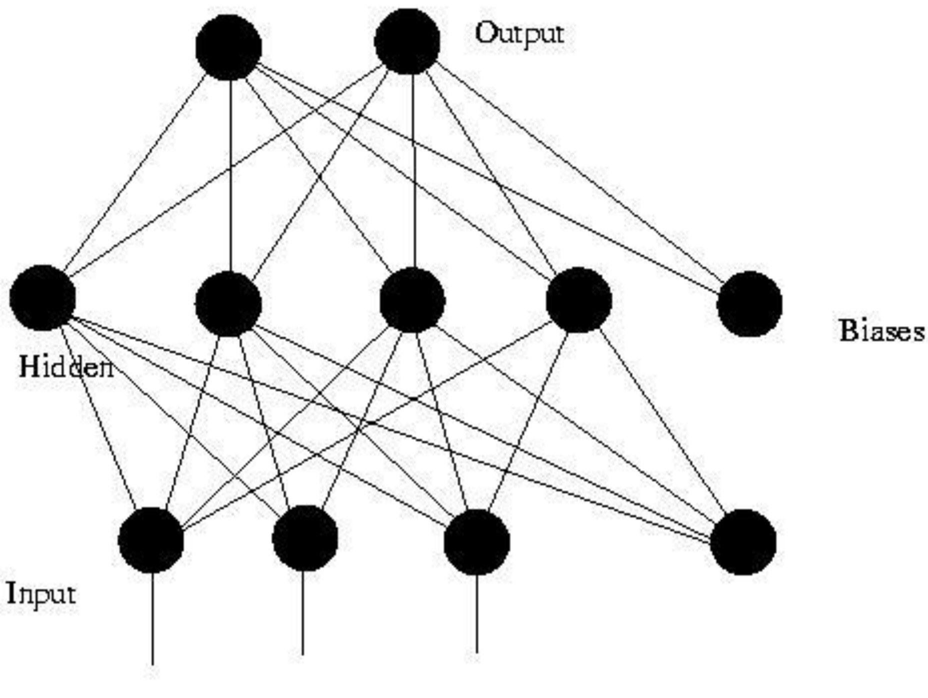
Σχήμα 3.7: *Γραμμική συνάρτηση ενεργοποίησης.*

Συνήθως αποφεύγουμε να χρησιμοποιήσουμε μια τέτοια συνάρτηση ως συνάρτηση ενεργοποίησης, και κυρίως στους κρυμμένους νευρώνες. Μια παραλλαγή της γραμμικής συνάρτησης είναι η αποκομμένη γραμμική συνάρτηση που φαίνεται στο παρακάτω διάγραμμα:



Σχήμα 3.8: *Αποκομμένη γραμμική συνάρτηση ενεργοποίησης.*

Ένα νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης έχει την παρακάτω μορφή:



Σχήμα 3.9: *Νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης.*

Παρατηρούμε ότι αποτελείται από στρώματα νευρώνων. Κάθε νευρώνας σε ένα συγκεκριμένο στρώμα ενώνεται μέσω συνάψεων με όλους τους νευρώνες του αμέσως επόμενου στρώματος. Δεν υπάρχουν αναδράσεις στο πολυεπίπεδο perceptron όπως γενικότερα δεν υπάρχουν αναδράσεις στα feedforward δίκτυα.

Η είσοδος του δικτύου τροφοδοτείται με ένα διάνυσμα N διατάσεων , ενώ η έξοδος του δίνει ένα διάνυσμα M διαστάσεων. Έτσι το δίκτυο μπορεί να θεωρηθεί ότι είναι μια συνάρτηση από τον χώρο των N-διαστάσεων στο χώρο των M-διαστάσεων. Αυτή η συνάρτηση μπορεί να γραφεί ως εξής:

$$y = f(x, W) = \sigma(W_n \sigma(W_{n-1} \sigma(\ldots \sigma(W_1 x)\ldots))) \qquad (3.2)$$

όπου:

- y είναι το διάνυσμα εξόδου.
- x είναι το διάνυσμα εισόδου.
- W_i είναι ένας πίνακας που περιέχει τα βάρη των νευρώνων του i-στου κρυμμένου στρώματος.

Το πολυεπίπεδο perceptron αποτελείται από τρία στρώματα:

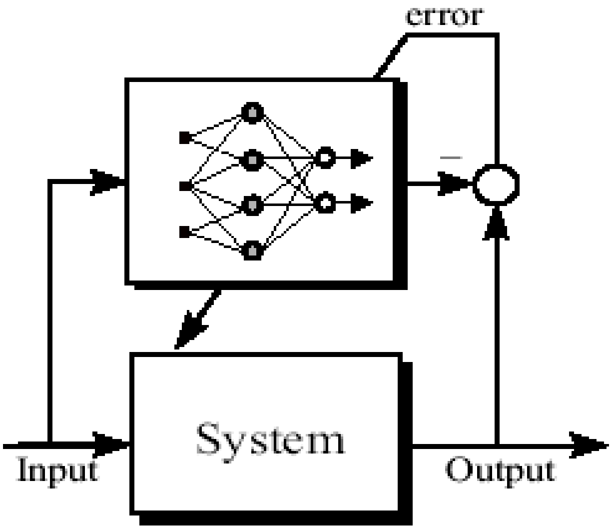
1. Το επίπεδο εισόδου
2. Το κρυφό στρώμα
3. Το επίπεδο εξόδου

Συνήθως η συνάρτηση ενεργοποίησης που χρησιμοποιείται στους κρυφούς νευρώνες είναι μη γραμμική (π.χ. σιγμοειδής). Στο εξωτερικό στρώμα μπορούν χρησιμοποιηθούν και γραμμικές συναρτήσεις ενεργοποίησης.

3.3 Μάθηση

Τα νευρωνικά δίκτυα μαθαίνουν μέσα από παραδείγματα , με τρόπο δηλαδή παρόμοιο μ' αυτόν που μαθαίνει ο άνθρωπος. Η ιδέα είναι να τροφοδοτηθεί το δίκτυο με μια σειρά από σήματα εισόδων και τις αντίστοιχες εξόδους που θέλουμε να προκύπτουν ως αποτέλεσμα αυτών των εισόδων. Όπως είδαμε παραπάνω τα βάρη που συνδέουν τους νευρώνες προσαρμόζονται κατά τη διαδικασία της μάθησης έτσι ώστε να ελαχιστοποιείται το σφάλμα ανάμεσα στην πραγματική έξοδο του δικτύου και την αντίστοιχη επιθυμητή.

Κατά τη μάθηση τροφοδοτούμε συνεχώς το δίκτυο με το πρότυπο εισόδων – εξόδων. Κάθε πλήρης προβολή του συνόλου εκπαίδευσης καλείται εποχή (epoch). Η εκπαίδευση του δικτύου γίνεται μέσα από διαδοχικές εποχές μέχρι να σταθεροποιηθούν τα βάρη και να συγκλίνει το άθροισμα των τετραγωνικών σφαλμάτων. Η διαδικασία της εκπαίδευσης φαίνεται σχηματικά στην παρακάτω εικόνα:



Σχήμα 3.10: *Σχηματική απεικόνιση της* διαδικασίας εκπαίδευσης.

Υπάρχουν διάφοροι αλγόριθμοι με τους οποίους επιτυγχάνουμε την εκπαίδευση ενός νευρωνικού δικτύου. Ο δημοφιλέστερος από αυτούς είναι ο αλγόριθμος της προς τα πίσω διάδοσης σφάλματος (backpropagation algorithm). Ο αλγόριθμος αυτός είναι μια γενίκευση της μεθόδου μάθησης των Widrow-Hoff για δίκτυα πολλαπλών επιπέδων και μη γραμμικές διαφορίσιμες συναρτήσεις μεταφοράς. Τα διανύσματα εξόδου χρησιμοποιούνται για να εκπαιδεύσουν το δίκτυο μέχρι να μπορέσει αυτό να προσεγγίσει μια συνάρτηση ή να ταξινομήσει τις εισόδους.

Ο τυπικός αλγόριθμος backpropagation είναι μια μέθοδος πτώσης κλίσης (gradient descent) κατά την οποία τα βάρη του δικτύου μετακινούνται προς την αρνητική μεριά της κλίσης. Υπάρχουν διάφορες τροποποιήσεις του αλγορίθμου, οι οποίες μπορούν να χρησιμοποιηθούν για να καλύψουν τις ανάγκες του εκάστοτε προβλήματος. Κατά την εκπαίδευση του νευρωνικού δικτύου που χρησιμοποιούμε για την πρόβλεψη φορτίου δεν έχουμε εφαρμόσει τον τυπικό αλγόριθμο backpropagation αλλά μια παραλλαγή αυτού (Levenberg-Marquardt algorithm) στην οποία αναφερόμαστε αναλυτικά παρακάτω.

Ας εξετάσουμε όμως πρώτα την τυπική μορφή που έχει ο αλγόριθμος της προς τα πίσω διάδοσης του σφάλματος. Η απλούστερη εφαρμογή του αλγορίθμου backpropagation ανανεώνει τις τιμές των βαρών και των πολώσεων του δικτύου προς την κατεύθυνση εκείνη που μειώνεται το μέσο τετραγωνικό σφάλμα ανάμεσα στις πραγματικές εξόδους του δικτύου και στις εξόδους -στόχους, πιο γρήγορα. Μια επανάληψη του αλγορίθμου αυτού έχει την παρακάτω μορφή:

$$\Delta w_{pl,qk}(n+1) = l_r \cdot \delta_{qk} \cdot OUT_{pj} + m\Delta w_{pl,qk}(n)$$

$$w_{pl,qk}(n+1) = w_{pl,qk+}(n) + \Delta w_{pl,qk}(n+1) \qquad (3.3)$$

όπου:

n:	ο αύξων αριθμός ανακύκλωσης της διαδικασίας εκπαίδευσης.
$w_{pl,qk}(n)$:	η τιμή του βάρους που συνδέει το νευρώνα p του στρώματος j με το νευρώνα q του στρώματος k κατά τη διάρκεια της ανακύκλωσης n.
$\Delta w_{pj,qk}(n)$:	η μεταβολή του βάρους $w_{pj,qk}$ κατά τη διάρκεια της ανακύκλωσης n.
δ_{qk} :	ο συντελεστής μεταβολής του νευρώνα q του στρώματος k.
OUT_{pj}:	Η έξοδος του νευρώνα p του στρώματος j.
l_r:	ο ρυθμός μάθησης.
m:	ο συντελεστής ορμής.

Ο τρόπος υπολογισμού του δ εξαρτάται από τη θέση του νευρώνα στο δίκτυο. Αν είναι νευρώνας εξόδου τότε το δ υπολογίζεται από την παρακάτω σχέση:

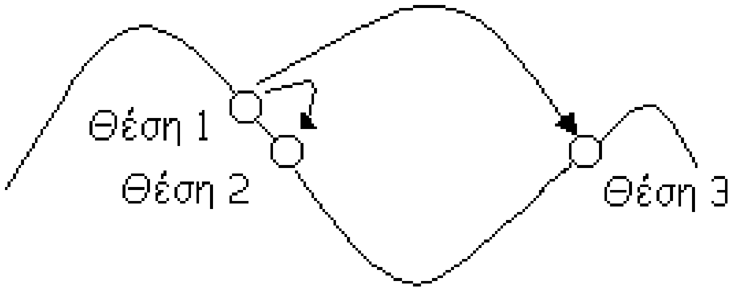
$$\delta_{qk} = OUT_{qk} \cdot (1 - OUT_{qk}) \cdot (OUT_{desired\_qk} - OUT_{qk}) \qquad (3.4)$$

Για τους υπόλοιπους νευρώνες χρησιμοποιείται η εξής σχέση:

$$\delta_{pj} = OUT_{pj} \cdot (1 - OUT_{pj}) \sum_q \delta_{qk} \cdot w_{pj,qk} \qquad (3.5)$$

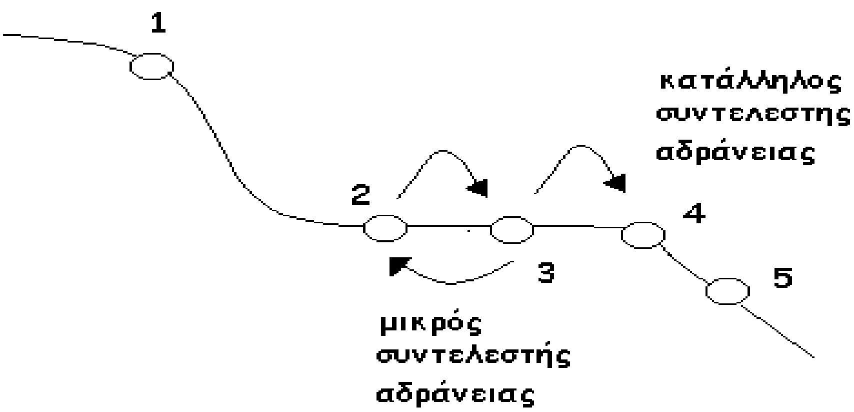
Ο ρυθμός μάθησης είναι μια παράμετρος που ελέγχει το ρυθμό με τον οποίο τα υπολογιζόμενα σφάλματα αλλάζουν τα βάρη του δικτύου. Η επιλογή της παραμέτρου αυτής είναι πολύ σημαντική. Αυτό γιατί ένας μεγάλος ρυθμός μάθησης εξασφαλίζει από τη μια ταχύτητα στη σύγκλιση, όμως υπάρχει από την άλλη ο κίνδυνος να υπερπηδηθεί η θέση ελαχίστου της αντικειμενικής συνάρτησης.

Για παράδειγμα στο παρακάτω σχήμα , με ένα μικρό συντελεστή μάθησης μεταβαίνουμε από τη θέση 1 στη θέση 2 , οδεύοντας προς τη θέση ελαχίστου. Αντίθετα με ένα μεγάλο ρυθμό μάθησης μεταβαίνουμε από τη θέση 1 στη θέση 3, υπερπηδώντας τη θέση ελαχίστου.



Σχήμα 3.11*: Παράδειγμα σχετικά με τον ρυθμό μάθησης*.

Ο συντελεστής ορμής παίρνει τιμές συνήθως από 0.01 έως 1.0. Ο ρόλος του είναι να επιτρέπει την αξιοποίηση πληροφορίας από προηγούμενες τιμές σφάλματος και να εμποδίζει τον εγκλωβισμό σε τοπικό ελάχιστο. Στο παρακάτω σχήμα βλέπουμε πως ένας μικρός συντελεστής ορμής μπορεί να οδηγήσει σε εγκλωβισμό του δικτύου σε θέση τοπικού ελαχίστου.



Σχήμα 3.12: *Παράδειγμα σχετικά με τον* συντελεστή ορμής.

Υπάρχουν δύο διαφορετικοί τρόποι με τους οποίους μπορούμε να εφαρμόσουμε αυτόν τον αλγόριθμο. Είτε σταδιακά (incremental mode) είτε συνολικά (batch mode).

Incremental mode: Σύμφωνα με αυτή τη μέθοδο η κλίση υπολογίζεται και τα βάρη ανανεώνονται με την εφαρμογή κάθε εισόδου.

Batch mode: Σύμφωνα με αυτή τη μέθοδο πρώτα εφαρμόζονται όλες οι είσοδοι στο δίκτυο και στη συνέχεια ανανεώνονται οι τιμές των βαρών και των πολώσεων

3.3.1 Levenberg-Marquardt αλγόριθμος

Όπως αναφέρθηκε και παραπάνω κατά την εκπαίδευση των νευρωνικών δικτύων που χρησιμοποιούμε για βραχυπρόθεσμη πρόβλεψη φορτίου δε χρησιμοποιούμε τον κλασσικό αλγόριθμο backpropagation αλλά μια παραλλαγή του. Αυτός ο αλγόριθμος επιτρέπει γρηγορότερη σύγκλιση, που είναι απαραίτητη. Η μαθηματική μορφή μιας επανάληψης του αλγορίθμου αυτού είναι η εξής:

$$x_{k+1} = x_k - \left(H + \mu \cdot \mathrm{I}\right) g \qquad (3.6)$$

Όπου:

- x_k είναι το άνυσμα των βαρών και των πολώσεων του δικτύου.
- H είναι η μήτρα των δευτέρων παραγώγων (Hessian matrix).
- I είναι ο μοναδιαίος πίνακας.
- g είναι η κλίση gradient.
- μ είναι ένας συντελεστής που καθορίζει το μέγεθος κάθε βήματος.

Η μήτρα H υπολογίζεται από την ακόλουθη σχέση:

$$H = J^T J \qquad (3.7)$$

Όπου η μήτρα J είναι η ιακωβιανή που περιέχει τις πρώτες παραγώγους της συνάρτησης σφάλματος ως προς τα βάρη και τις πολώσεις του δικτύου. Η κλίση μπορεί να υπολογισθεί από τη σχέση:

$$g = J^T e \tag{3.8}$$

Όπου το e, είναι ένα διάνυσμα των σφαλμάτων του δικτύου. Έτσι μια επανάληψη του αλγορίθμου θα έχει τη μορφή:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\mathbf{J}^T\mathbf{J} + \mu\mathbf{I}]^{-1}\mathbf{J}^T\mathbf{e} \tag{3.9}$$

Όταν ο συντελεστής *μ είναι 0* τότε ο αλγόριθμος εκφυλίζεται σε μορφή αλγορίθμου Newton. Όταν είναι *μεγάλο* τότε μειώνεται και το βήμα της κάθε επανάληψης. Έτσι το μ μειώνεται μετά από κάθε επιτυχές βήμα και αυξάνεται μετά από ένα ανεπιτυχές βήμα.

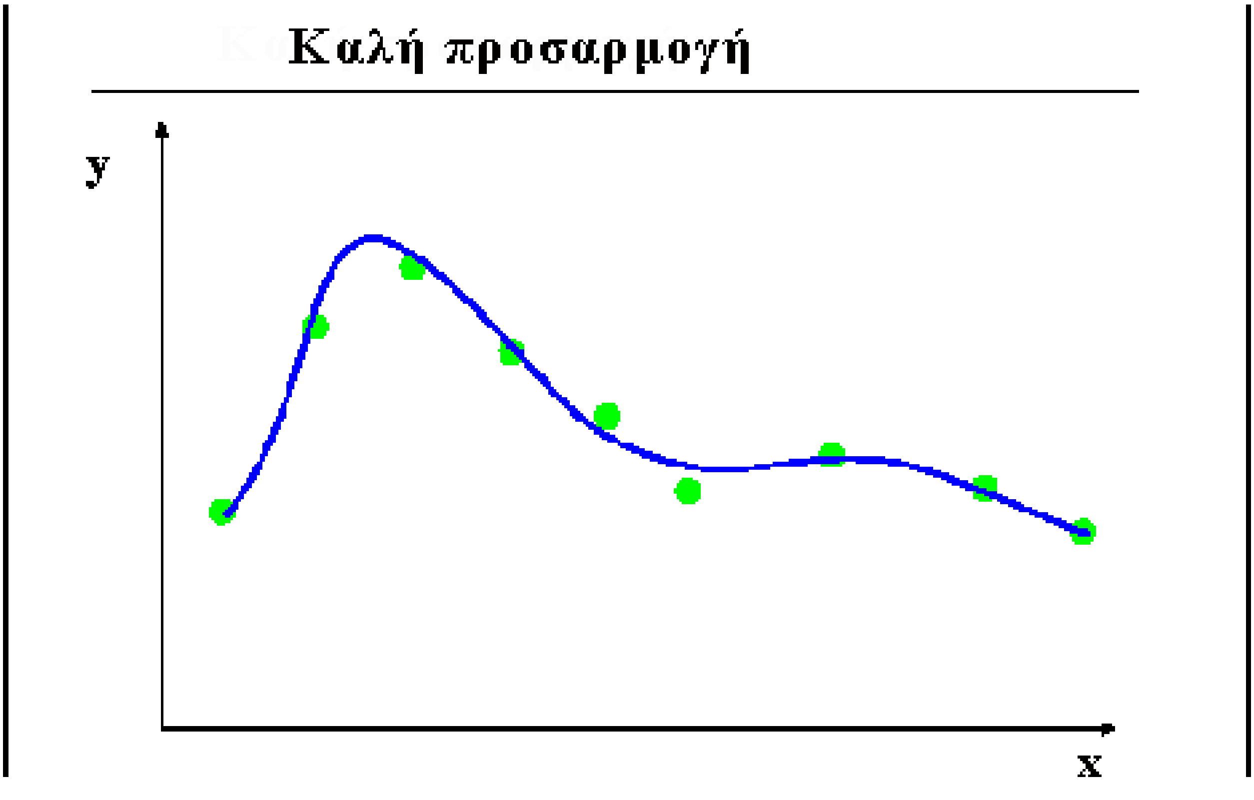
Ο αλγόριθμος της προς τα πίσω διάδοσης σφάλματος, παρά το γεγονός ότι μπορεί να εφαρμοσθεί με επιτυχία για την εκπαίδευση ενός νευρωνικού δικτύου, ωστόσο παρουσιάζει και κάποια **μειονεκτήματα**. Ένα από αυτά είναι *η μεγάλη διαδικασία εκπαίδευσης*, η οποία είναι χρονοβόρα, ιδιαίτερα όταν ο ρυθμός μάθησης δεν είναι κατάλληλα επιλεγμένος. Πολύ σημαντικό είναι επίσης και το πρόβλημα της *παράλυσης του δικτύου*. Καθώς το δίκτυο εκπαιδεύεται , τα βάρη του μπορούν να πάρουν πολύ υψηλές τιμές. Η είσοδος λοιπόν με την οποία τροφοδοτείται ένας νευρώνας θα είναι μεγάλη , και η έξοδος του εξαιτίας της σιγμοειδούς συνάρτησης ενεργοποίησης που συνήθως χρησιμοποιείται θα είναι ή 0 ή 1. Επομένως οι προσαρμογές που θα γίνονται στα βάρη στη συνέχεια , δε θα επιφέρουν καμιά ουσιαστική βελτίωση και η εκπαίδευση θα μένει στάσιμη. Σημειώνεται ακόμη ότι ο αλγόριθμος της προς τα πίσω διάδοσης του σφάλματος *δεν εγγυάται ότι θα συγκλίνει στο ελάχιστο της αντικειμενικής συνάρτησης*. Αντίθετα σε πολλές περιπτώσεις οδηγούμαστε σε σύγκλιση σε τοπικά ελάχιστα.

3.4 Γενίκευση (generalization)

Η διαδικασία της εκπαίδευσης αποσκοπεί στην ελαχιστοποίηση του σφάλματος ανάμεσα στην επιθυμητή έξοδο του δικτύου και στην αντίστοιχη πραγματική. Το να επιτύχουμε όμως ένα πολύ μικρό σφάλμα δε σημαίνει απαραίτητα ότι έχουμε καταφέρει να δημιουργήσουμε ένα λειτουργικό μοντέλο που να είναι ικανό να ανταποκριθεί στις απαιτήσεις μας με δεδομένα που δεν ανήκουν στο σύνολο μάθησης.

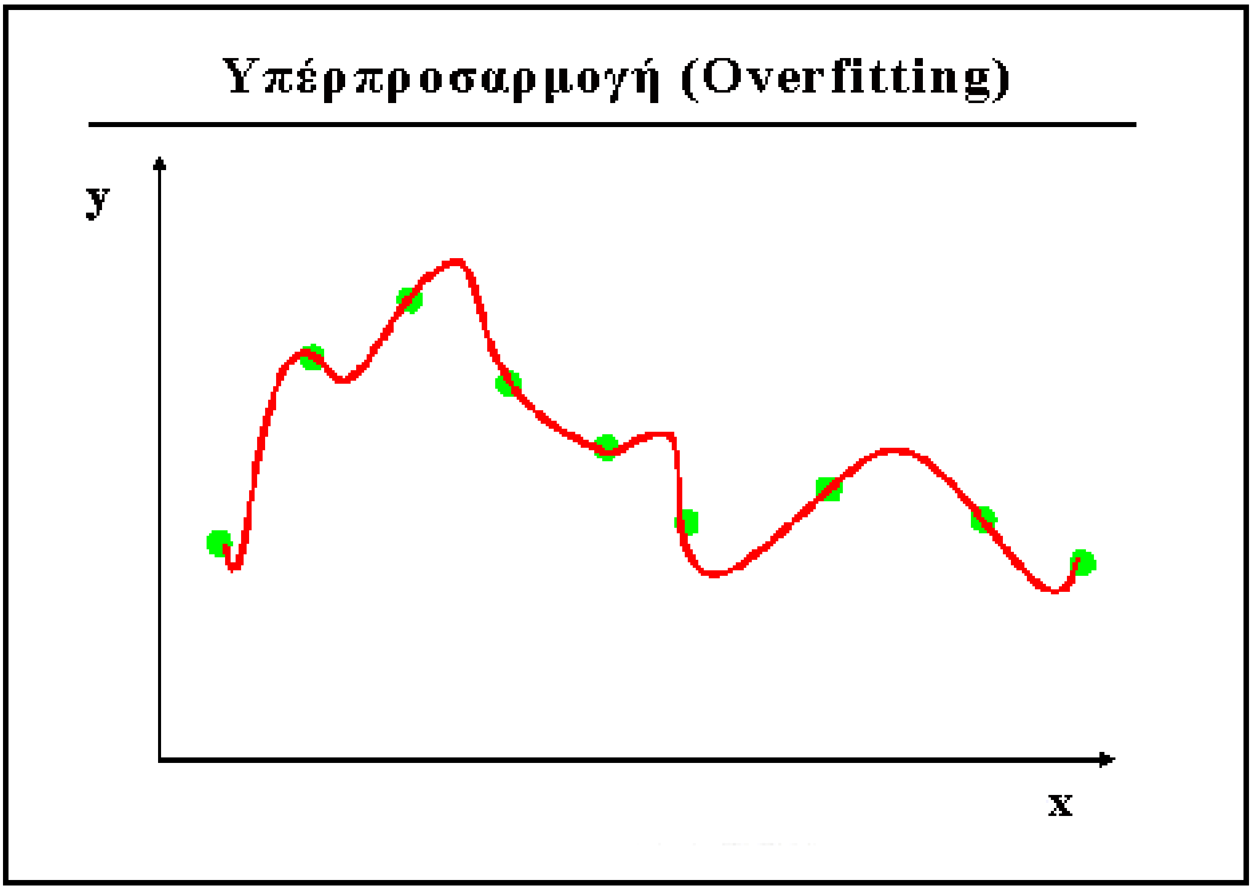
Για να είναι το νευρωνικό δίκτυο επιτυχημένο πρέπει να έχει την ικανότητα της γενίκευσης. Αυτό σημαίνει ότι πρέπει να δίνει ακριβείς εξόδους και όταν δέχεται δεδομένα εισόδων εκτός του συνόλου μάθησης.

Ένα πρόβλημα που συναντάται συχνά κατά τη σχεδίαση και εκπαίδευση ενός νευρωνικού δικτύου είναι η υπερπροσαρμογή (overfitting), η απομνημόνευση δηλαδή του συνόλου μάθησης. Αυτό που επιδιώκουμε με τα τεχνητά νευρωνικά δίκτυα δεν είναι η απεικόνιση των δεδομένων εισόδων στα αντίστοιχα δεδομένα εξόδου μέσα από μια διαδικασία απομνημόνευσης. Αντίθετα θέλουμε μέσα από τα βάρη που ενώνουν τους νευρώνων να ευρεθούν πρότυπα (patterns) και συσχετίσεις ανάμεσα σε εισόδους και εξόδους. Στο παρακάτω σχήμα βλέπουμε ένα παράδειγμα σωστής προσαρμογής:



Σχήμα 3.13: Σχηματική απεικόνιση σωστής προσαρμογής.

Τα ίδια δεδομένα θα μπορούσαν να προσεγγιστούν καλύτερα με μια άλλη καμπύλη με τη διαφορά όμως ότι τότε ενδεχομένως να είχαμε υπερπροσαρμογή. Στο σχήμα που ακολουθεί βλέπουμε ένα παράδειγμα υπέρπροσαρμογής.



Σχήμα 3.14: Σχηματική απεικόνιση υπέρπροσαρμογής (Overfitting).

Η ικανότητα ενός τεχνητού νευρωνικού δικτύου για γενίκευση επηρεάζεται κυρίως από τους ακόλουθους παράγοντες:

- ***Μέγεθος συνόλου εκπαίδευσης***

Η επιλογή του πλήθους των εισόδων είναι ουσιαστική για να έχει το δίκτυο ικανότητα γενίκευσης. Έτσι θα πρέπει να επιλέγουμε τόσες εισόδους όσες μας επιτρέπει το σύνολο δεδομένων εκπαίδευσης που έχουμε στη διάθεσή μας. Όσο πιο μεγάλη είναι η διάσταση του χώρου εισόδου τόσο πιο αραιό θα είναι το σύνολο των παρατηρήσεων , πράγμα που διευκολύνει την απομνημόνευση των δεδομένων. Ακόμα η επιλογή εισόδων που δεν συσχετίζονται με τις εξόδους, μπορεί να δημιουργήσει συσχετίσεις ανάμεσα σε εισόδους και εξόδους , οι οποίες όμως δεν υπάρχουν στην πραγματικότητα. Βέβαια από την άλλη όλοι οι παράγοντες που επιδρούν στις εξόδους πρέπει να συμπεριληφθούν στο διάνυσμα των εισόδων

- ***Επάρκεια του συνόλου εκπαίδευσης***

Το σύνολο εκπαίδευσης πρέπει να είναι επαρκές έτσι ώστε να συμπεριλαμβάνει όλες τις πιθανές περιπτώσεις. Άλλωστε ένα μεγάλο σύνολο εκπαίδευσης είναι δύσκολο να απομνημονευθεί.

- ***Δομή του νευρωνικού δικτύου***

Ένα μεγάλο νευρωνικό δίκτυο συμπεριλαμβάνει ένα μεγάλο αριθμό ελευθέρων παραμέτρων και αυτό μπορεί να οδηγήσει σε υπερπροσαρμογή (overfitting). Κάθε κρυμμένο στρώμα νευρώνων φέρει έναν αριθμό παραμέτρων. Έτσι για να αποφύγουμε την ύπαρξη πολλών παραμέτρων που θα επιβαρύνουν το μοντέλο πρέπει να χρησιμοποιήσουμε λελογισμένο πλήθος κρυμμένων στρωμάτων. Για MLPs, 3 στρωμάτων υπάρχει ένας κανόνας που μπορεί να χρησιμοποιηθεί για την επιλογή του πλήθους των κρυμμένων νευρώνων. Το πλήθος των ελευθέρων παραμέτρων σε ένα τέτοιο μοντέλο είναι:

$$W = H(N + M) \qquad (3.10)$$

όπου:

H: είναι το πλήθος των κρυμμένων νευρώνων.
N: η διάσταση του διανύσματος εισόδων.
M: η διάσταση του διανύσματος εξόδου.

Σύμφωνα με τον κανόνα αυτό θα πρέπει το πλήθος των ελευθέρων παραμέτρων να είναι μικρότερο από το ένα πέμπτο του συνόλου εκπαίδευσης. Δηλαδή αν υποθέσουμε ότι το πλήθος των παρατηρήσεων του συνόλου εκπαίδευσης είναι T θα πρέπει να ισχύει:

$$H \le \frac{T}{5(N + M)} \qquad (3.11)$$

Διαπιστώνουμε λοιπόν ότι είναι αναγκαίο όταν σχεδιάζουμε και εκπαιδεύουμε ένα δίκτυο να ελέγχουμε αν αυτό έχει την ικανότητα γενίκευσης. Για να το επιτύχουμε αυτό πρέπει να επαληθεύσουμε το μαθηματικό μας μοντέλο. Έτσι εκτός από το σύνολο μάθησης έχουμε και ένα σύνολο επαλήθευσης ,του οποίου τα δεδομένα εισόδου τροφοδοτούμε στο δίκτυο και μετράμε το σφάλμα ανάμεσα στις επιθυμητές εξόδους και τις πραγματικές εξόδους του δικτύου. Στις περισσότερες περιπτώσεις το σφάλμα αυτό είναι λίγο μεγαλύτερο από το σφάλμα που αντιστοιχεί στο σύνολο μάθησης. Αν όμως είναι κατά πολύ μεγαλύτερο τότε αυτό είναι μια ένδειξη υπερπροσαρμογής και επομένως το μοντέλο μας ενδέχεται να μην είναι κατάλληλο.

Είναι λοιπόν απαραίτητο να σχεδιάζουμε το δίκτυο μας με τέτοιο τρόπο που να επιτυγχάνεται η γενίκευση. Για να το κάνουμε αυτό υπάρχουν κάποιες μεθοδολογίες οι οποίες μπορούν να μας βοηθήσουν.

Ένας τρόπος είναι να σχεδιαστεί το δίκτυο έτσι ώστε να είναι τόσο μεγάλο (σε πλήθος κρυμμένων στρωμάτων , νευρώνων και εισόδων) , όσο ακριβώς χρειάζεται για να επιτευχθεί μια επαρκής προσαρμογή. Αν ένα δίκτυο δεν είναι πολύ μεγάλο τότε δε θα έχει και την ισχύ για να κάνει overfitting. Όμως δεν υπάρχει μια συστηματική μέθοδος για να το κάνουμε αυτό. Έτσι ο μόνος τρόπος για να το επιτύχουμε είναι να με δοκιμή και σφάλμα (trial and error). Όμως αυτός ο τρόπος είναι εξαιρετικά επίπονος και βέβαια είναι αμφίβολο αν θα οδηγήσει σε ικανοποιητικό αποτέλεσμα. Υπάρχουν ωστόσο και δύο άλλοι τρόποι για να βελτιώσουμε τη γενίκευση. Αυτοί είναι η κανονικοποίηση και το έγκαιρο σταμάτημα στην εκπαίδευση.

3.4.1 Κανονικοποίηση (Regularization)

Η μέθοδος αυτοί περιλαμβάνει την τροποποίηση της συνάρτησης επίδοσης (performance function), η οποία συνήθως είναι το άθροισμα των τετραγωνικών σφαλμάτων ανάμεσα στις εξόδους του δικτύου και τους στόχους που έχουμε ορίσει. Παρακάτω εξηγούμε πως μπορεί να τροποποιηθεί η συνάρτηση επίδοσης έτσι ώστε να βελτιωθεί η ικανότητα του δικτύου για γενίκευση.

Τροποποιημένη συνάρτηση επίδοσης

Η τυπική συνάρτηση επίδοσης που χρησιμοποιείται κατά την εκπαίδευση δικτύων εμπρόσθιας τροφοδότησης έχει τη μορφή:

$$F = mse = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2$$

(3.12)

Μπορούμε να τροποποιήσουμε την παραπάνω συνάρτηση και να συμπεριλάβουμε σ' αυτήν τη μέση τιμή των τετραγώνων των παραμέτρων του δικτύου (βάρη και πολώσεις). Τότε η συνάρτηση επίδοσης θα έχει την ακόλουθη μορφή.

$$msereg = \gamma mse + (1-\gamma) msw \tag{3.13}$$

Όπου ο συντελεστής γ, ονομάζεται **λόγος επίδοσης** και

$$msw = \frac{1}{n}\sum_{j=1}^{n} w_j^2$$

(3.14)

Χρησιμοποιώντας αυτή τη συνάρτηση οδηγούμε το δίκτυο στο να έχει μικρότερα βάρη και πολώσεις και ως εκ τούτου να αποκρίνεται ομαλότερα στις εισόδους οδηγώντας σε μικρότερη πιθανότητα για υπερπροσαρμογή.

Είναι πολλές φορές επιθυμητό να καθορίσουμε τις βέλτιστες παραμέτρους κανονικοποίησης με έναν αυτόματο τρόπο. Μια προσέγγιση σ' αυτό το πρόβλημα είναι η Bayesian. Σ' αυτή την προσέγγιση τα βάρη και οι πολώσεις θεωρούνται τυχαίες μεταβλητές με συγκεκριμένες κατανομές. Οι παράμετροι κανονικοποίησης σχετίζονται με τις άγνωστες μεταβλητότητες αυτών των κατανομών. Με στατιστικές τεχνικές μπορούμε να εκτιμήσουμε αυτές τις παραμέτρους.

Εγκαιρο σταματημα (Early stopping)

Μια άλλη μέθοδος για να βελτιώσουμε τη γενίκευση είναι το έγκαιρο σταμάτημα της εκπαίδευσης. Σ' αυτή την τεχνική τα δεδομένα χωρίζονται σε τρεις κατηγορίες. Η πρώτη ομάδα δεδομένων είναι τα δεδομένα που θα χρησιμοποιηθούν για την εκπαίδευση του δικτύου. Η δεύτερη ομάδα είναι τα δεδομένα που χρησιμοποιούνται για validation. Αρχικά των σφάλμα στη δεύτερη ομάδα θα μειώνεται ταυτόχρονα με το σφάλμα στην πρώτη ομάδα. Καθώς όμως προχωρεί η διαδικασία της εκπαίδευσης , το δίκτυο αρχίζει να κάνει υπερπροσαρμογή και έτσι παρατηρείται το μεν σφάλμα στα δεδομένα εκπαίδευσης να μειώνεται και το σφάλμα στα δεδομένα ελέγχου να αυξάνεται. Τότε είναι η κατάλληλη στιγμή για να σταματήσουμε τη διαδικασία εκπαίδευσης. Το τρίτο σετ δεδομένων δε χρησιμοποιείται στη διαδικασία της εκπαίδευσης αλλά μετά , με στόχο να συγκριθούν διαφορετικές δομές νευρωνικών δικτύων

Κλάδεμα του δικτύου

Όπως είπαμε και παραπάνω μια από τις αιτίες στις οποίες οφείλεται η απομνημόνευση του συνόλου εκπαίδευσης είναι ο υπερβολικός αριθμός των κρυμμένων στρωμάτων και των νευρώνων που υπάρχουν σ' αυτά. Δόθηκε ένας εμπειρικός κανόνας που για MLPs τριών στρωμάτων δίνει τον μέγιστο αριθμό των κρυμμένων νευρώνων που μπορεί να έχει το δίκτυο ώστε να είναι ικανό να κάνει γενίκευση. Ωστόσο ο κανόνας αυτός δεν μπορεί να εφαρμοστεί αυστηρά σε όλες τις περιπτώσεις.

Μια άλλη προσέγγιση που μπορούμε να κάνουμε για να βελτιώσουμε την ικανότητα γενίκευσης του δικτύου είναι το κλάδεμα του δικτύου. Η διαδικασία αυτή έχει ως στόχο να δημιουργήσει δίκτυα με όσο το δυνατό λιγότερα στρώματα και λιγότερους νευρώνες. Το κλάδεμα γίνεται μετά την εκπαίδευση του δικτύου. Αρχικά χρησιμοποιείται μια μεγαλύτερη τοπολογία δικτύου από αυτή που χρειάζεται. Στη συνέχεια με τη βοήθεια διαφόρων αλγορίθμων κλαδέματος αφαιρούνται νευρώνες και βάρη από το πλήρως συνδεδεμένο δίκτυο και έτσι βελτιώνεται η ικανότητα γενίκευσης του δικτύου.

3.5 Κλιμάκωση – κανονικοποίηση δεδομένων εισόδου-εξόδου

Η κλιμάκωση των μεταβλητών έχει ως στόχο την κανονικοποίηση τους και είναι απαραίτητη στα τεχνητά νευρωνικά δίκτυα για τους εξής λόγους:

- Αποφεύγεται η πιθανότητα το δίκτυο να συντονισθεί σε μια συγκεκριμένη περιοχή των δεδομένων εισόδου-εξόδου.

- Κλιμακώνονται οι τιμές εισόδου στους νευρώνες ώστε να βρίσκονται εντός της περιοχής λειτουργίας. Εάν θεωρήσουμε για παράδειγμα ότι η συνάρτηση των νευρώνων είναι μη συμμετρική (π.χ. σιγμοειδής) , στο διάστημα [0,1] τότε οι είσοδοι που βρίσκονται έξω από την περιοχή λειτουργίας θα οδηγήσουν το νευρώνα στην τιμή 1 ή στην τιμή 0. Η παράγωγος τότε της σιγμοειδούς θα είναι ίση με μηδέν και δεν θα μπορεί να συνεχιστεί η μάθηση. Θα έχουμε δηλαδή παράλυση του δικτύου.

- Η απ' ευθείας τροφοδότηση του δικτύου με τις πρωτογενείς τιμές των μεταβλητών εισόδου θα έχει ως αποτέλεσμα μεγάλες διαφορές στις περιοχές τιμών ανάμεσα στις αντίστοιχες εισόδους. Τις διαφορές αυτές το δίκτυο θα μπορούσε να τις εκλάβει ως σημαίνουσες τη βαρύτητα των αντίστοιχων εισόδων. Τότε το μοντέλο θα εξαρτάται σημαντικά από τις μεταβλητές με υψηλές τιμές και λιγότερο από τις μεταβλητές με χαμηλές τιμές. Αυτό όμως δεν είναι πληροφοριακά σωστό, γιατί οι υψηλές τιμές που μπορεί να πάρει μια μεταβλητή εισόδου δεν συνεπάγονται και μεγαλύτερη εξάρτηση της εξόδου από την μεταβλητή αυτή.

Επιδιώκουμε λοιπόν να μετασχηματίσουμε τις μεταβλητές εισόδου με τέτοιο τρόπο ώστε να έχουν μηδενική μέση τιμή και μοναδιαία τυπική απόκλιση. Έτσι εφαρμόζουμε την ακόλουθη μαθηματική σχέση:

$$\widetilde{x}_i(t) = \frac{x_i(t) - \bar{x}}{\sigma_x}$$

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{N} \left(x_i(t) - \bar{x}\right)^2}{N-1}} \qquad (3.15)$$

Ανάλογα με τη μορφή της συνάρτησης ενεργοποίησης , οι μεταβλητές εισόδου θα πρέπει να κανονικοποιηθούν σε διαφορετικά διαστήματα τιμών. Αν για παράδειγμα χρησιμοποιείται ως συνάρτηση ενεργοποίησης η σιγμοειδής , τότε οι τιμές των εξόδων θα πρέπει να είναι κανονικοποιημένες στο διάστημα [0.1,0.9] ή [0.15,0.85]. Ο μετασχηματισμός που χρησιμοποιείται είναι:

$$\tilde{y}(t) = SC \cdot y(t) + OFFSET$$

$$SC = \frac{Max - Min}{y_{\max} - y_{\min}} \qquad (3.16)$$

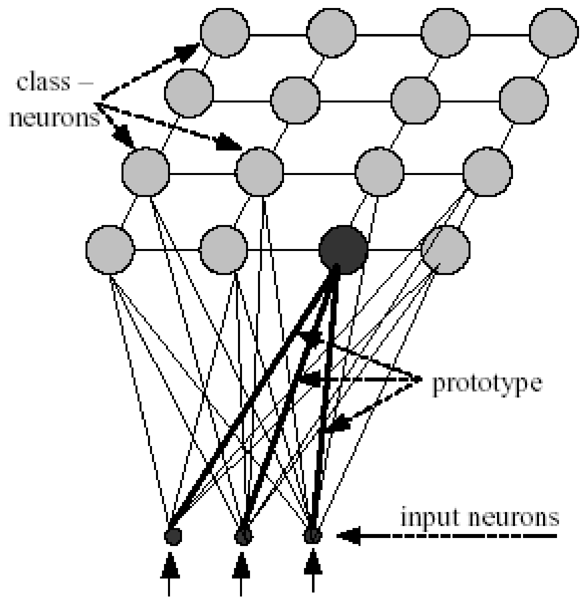
$$OFFSET = Max - SC \cdot y_{\max}$$

όπου:

y(t): Η χρονοσειρά του φορτίου
Max: Η μέγιστη τιμή που επιτρέπεται να έχουν οι στόχοι
Min: Η ελάχιστη τιμή που επιτρέπεται να έχουν οι στόχοι

3.6 Αυτόματα οργανωμένοι χάρτες (self organized maps)

Ένας τύπος δικτύων στον οποίο η μάθηση γίνεται χωρίς επίβλεψη, είναι τα δίκτυα Kohonen γνωστά και ως self organizing maps (SOMs). Ένα τέτοιο δίκτυο μαθαίνει μόνο του και παράγει εξόδους χωρίς να έχει πληροφορία για την σωστή τιμή. Η πληροφορία που χρειάζεται για την παραγωγή του αποτελέσματος αυτού δημιουργείται κατά τη διαδικασία της μάθησης και αποθηκεύεται στα βάρη των συνάψεων. Τα βάρη αυτά σχηματίζουν το σύνολο προτύπων.

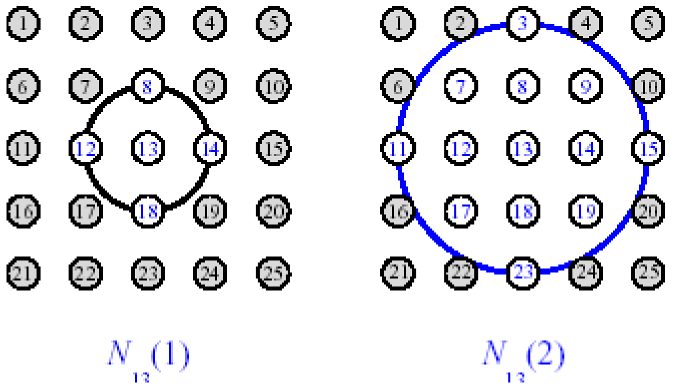


Σχήμα 3.15: *Δίκτυο τύπου self organizing map (SOM).*

Το δίκτυο Kohonen χρησιμοποιεί ένα πλέγμα νευρώνων. Κάθε νευρώνας αυτού του δικτύου συσχετίζεται με μια κατηγορία ή κλάση (class). Κάθε νευρώνας μπορεί να περιγραφεί από τη θέση του στο πλέγμα και από το πρότυπο. Το πρότυπο είναι το βάρος της σύναψης ανάμεσα στο νευρώνα αυτό και τους νευρώνες εισόδου. Η διαδικασία της αυτόματης οργάνωσης του χάρτη λαμβάνει χώρα κατά τη διάρκεια της εκπαίδευσης.
Σε πρώτο στάδιο όλα τα πρότυπα των νευρώνων αρχικοποιούνται σε πολύ μικρές , τυχαίες τιμές. Στη συνέχεια όλες οι παρατηρήσεις του συνόλου εκπαίδευσης εφαρμόζονται μια – μια στο στρώμα εισόδου. Κάθε παρατήρηση x_p συγκρίνεται με τα πρότυπα w_u των νευρώνων , και ο νευρώνας με τη μικρότερη απόσταση $||w_u-x_p||$,θεωρείται ότι είναι ο νικητής νευρώνας (winning neuron). Έπειτα το πρότυπο του νικητή νευρώνα τροποποιείται έτσι ώστε να πλησιάσει περισσότερο στην τρέχουσα παρατήρηση. Ακόμα τροποποιούνται τα πρότυπα όλων των νευρώνων όπου βρίσκονται στην γειτονιά του winning neuron. Με τον όρο γειτονιά εννοούμε το σύνολο των νευρώνων που βρίσκονται σε μια ακτίνα d από τον νευρώνα –νικητή.

$$N_i(d) = \{ j \mid d_{ij} \le d \} \qquad (3.17)$$

Στο παρακάτω σχήμα βλέπουμε τη γειτονιά του νευρώνα 13 για d=1 και για d=2.



Σχήμα 3.16: *Γειτονιά του νευρώνα 13 για d=1 και για d=2.*

ΚΕΦΑΛΑΙΟ 4

Ασαφή Σύνολα

4.1 Ασάφεια

Το 1965 ο Zadeh θεμελίωσε τη θεωρία των ασαφών συνόλων σα μια μέθοδο διαχείρισης της ανακρίβειας που αντιμετωπίζεται σε πολλά πρακτικά προβλήματα. Η ανακρίβεια , ή η ασάφεια είναι ο πυρήνας των ασαφών συνόλων και της ασαφούς λογικής. Τα ασαφή σύνολα ουσιαστικά αποτελούν μια γενίκευση των κλασσικών συνόλων. Από τις αρχές της δεκαετίας του 1980 τα ασαφή σύνολα βρήκαν πολλές εφαρμογές , ιδίως σε συστήματα ελέγχου.

Οι άνθρωποι σκέπτονται συνήθως όχι με όρους ακριβών συμβόλων και αριθμών αλλά με ασαφείς όρους. Αυτοί οι ασαφείς όροι προσδιορίζουν κατηγορίες όχι όμως απόλυτα διαχωρισμένα και σαφώς καθορισμένα σύνολα. Η μετάβαση από τη μια κατηγορία στην άλλη γίνεται σταδιακά , μεταβαίνοντας από καταστάσεις με περισσότερη ή λιγότερη συσχέτιση με την κατηγορία. Με βάση λοιπόν την πρακτική των ανθρώπων να χρησιμοποιούν ελαστικά σύνολα , ο Zadeh πρότεινε την ιδέα των ασαφών συνόλων.

Τα ασαφή σύνολα είναι συναρτήσεις που απεικονίζουν μια τιμή , η οποία μπορεί να είναι μέλος του συνόλου , σε έναν αριθμό από το 0 έως το 1. Ο αριθμός αυτός δηλώνει το βαθμό συσχέτισης της τιμής με το ασαφές σύνολο. Όταν ο βαθμός αυτός είναι 0 τότε αυτό υποδηλώνει ότι η τιμή δεν ανήκει στο σύνολο, ενώ όταν είναι 1 τότε αυτό σημαίνει ότι η τιμή αυτή αντιπροσωπεύει πλήρως το ασαφές σύνολο. Ο βαθμός αυτός καθορίζεται από τη συνάρτηση συμμετοχής του ασαφούς συνόλου. Στην συνέχεια γίνεται αναφορά για τις συναρτήσεις συμμετοχής.

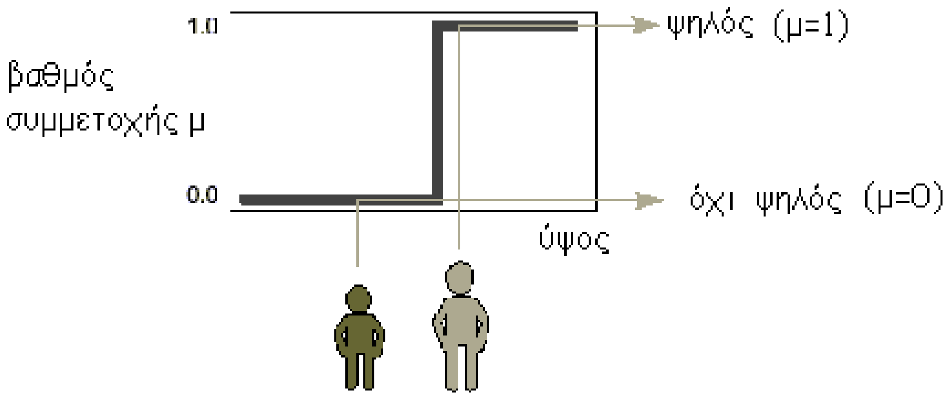
4.2 Συναρτησεις συμμετοχής.

Ένα ασαφές σύνολο Α ορίζεται ως ένα σύνολο διατεταγμένων ζευγών (x, μ_A(x) όπου x∈X και μ_A(x)∈[0,1]). Το σύνολο X αποτελεί ένα ευρύτερο σύνολο αναφοράς (universe of discourse) που περιλαμβάνει όλα τα αντικείμενα στα οποία μπορεί να γίνει αναφορά. Η τιμή μ_A(x) λέγεται βαθμός αλήθειας , συμβολίζει το βαθμό συγγένειας του x στο Α και παίρνει τιμές στο διάστημα [0,1]. Τέλος η συνάρτηση μ_A ονομάζεται συνάρτηση συμμετοχής (membership function). Στην πράξη η συνάρτηση συμμετοχής μπορεί να προέρχεται από:

- Υποκειμενικές εκτιμήσεις
- Προκαθορισμένες (ad hoc) και απλοποιημένες μορφές
- Συχνότητες εμφανίσεων και πιθανότητες
- Φυσικές μετρήσεις
- Διαδικασίες μάθησης και προσαρμογής (π.χ. με νευρωνικά δίκτυα)

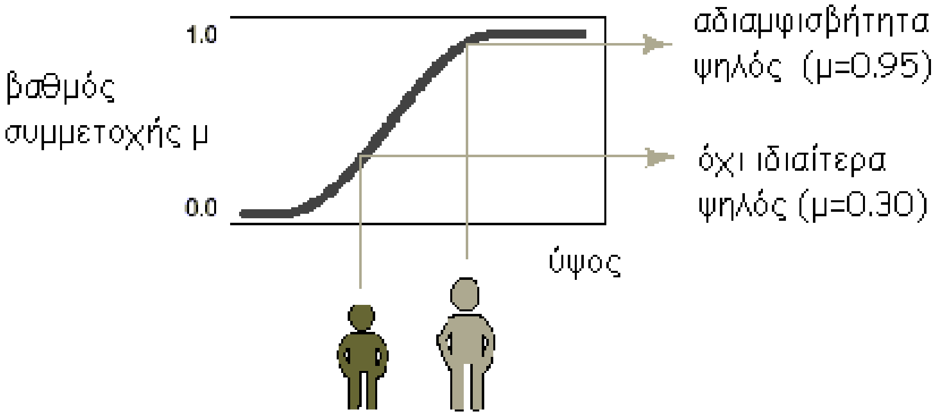
Η διαφορά των ασαφών συνόλων συγκριτικά με την κλασσική θεωρία συνόλων είναι ότι στην κλασσική θεωρία συνόλων ισχύει μ_A(x)∈{0,1}, δηλαδή το x είτε ανήκει στο Α (μ_A(x)=1) ή δεν ανήκει (μ_A(x)=0). Άρα η ασαφής θεωρία συνόλων μεταπίπτει στην αντίστοιχη κλασσική , όταν οι δυνατές τιμές της συνάρτησης συμμετοχής είναι μόνο 0 ή 1. Παρακάτω αναπτύσσεται ένα παράδειγμα με το οποίο μπορεί να γίνει περισσότερο κατανοητή η έννοια της συνάρτησης συμμετοχής.

Ας υποθέσουμε ότι ο χώρος αναφοράς X είναι το σύνολο όλων των ανθρώπων. Ένα ασαφές υποσύνολο του χώρου αυτού είναι οι ψηλοί άνθρωποι. Τα πιθανά ύψη έστω ότι κυμαίνονται από 1.20 μέχρι 2.50. Η λέξη ψηλός μπορεί να συσχετισθεί με μια καμπύλη η οποία δείχνει κατά πόσο ένας άνθρωπος είναι ψηλός ή όχι. Αν χρησιμοποιήσουμε τις αρχές των κλασσικών συνόλων τότε για να ορίσουμε το σύνολο των ψηλών ανθρώπων θα πρέπει να ορίσουμε μια συγκεκριμένη τιμή ύψους , η οποία θα διαχωρίζει τους ανθρώπους σε ψηλούς και κοντούς. Π.χ. ας υποθέσουμε ότι η τιμή αυτού του ύψους είναι 1.75m. τότε ένας άνθρωπος με ύψος 1.74 θα χαρακτηρίζεται κοντός ενώ ένας άνθρωπος με ύψος 1.76 θα χαρακτηρίζεται ψηλός. Ο παραπάνω διαχωρισμός φαίνεται παράλογος αφού έχουμε αντιστοιχήσει σε δύο ανθρώπους με αμελητέα διαφορά ύψους δύο αντίθετες μεταξύ τους έννοιες.



Σχήμα 4.1: *Κατηγοριοποίηση σύμφωνα με τα κλασσικά σύνολα.*

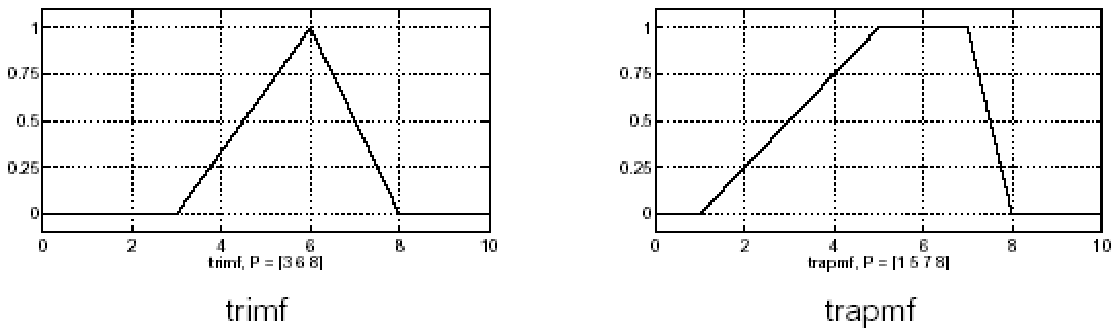
Ένας άλλος τρόπος να ορίσουμε την έννοια ψηλός είναι μέσω μιας καμπύλης που έχει ομαλή διακύμανση και μεταβαίνει από την έννοια ψηλός στην έννοια κοντός. Αυτή η καμπύλη είναι η συνάρτηση συμμετοχής του ασαφούς συνόλου των ψηλών ανθρώπων. Με άλλα λόγια δεχόμαστε ότι όλοι οι άνθρωποι είναι σε κάποιο βαθμό ψηλοί άλλα δεν είναι όλοι στον ίδιο βαθμό ψηλοί.



Σχήμα 4.2: *Κατηγοριοποίηση σύμφωνα με τα ασαφή σύνολα.*

Μπορούμε να παρατηρήσουμε ότι υποκειμενικοί παράγοντες ενυπάρχουν στα χαρακτηριστικά της δομής ενός ασαφούς συνόλου. Η μορφή δηλαδή της καμπύλης δεν μπορεί να είναι η ίδια όταν αναφερόμαστε στους ενήλικες και στους ανήλικους, σε γυναίκες και άντρες κλπ. Η μορφή επίσης της καμπύλης επιλέγεται αυθαίρετα σύμφωνα με την αντίληψη που έχει κάθε άνθρωπος για την έννοια ψηλός. Η μόνη προϋπόθεση που πρέπει να ικανοποιεί μια συνάρτηση συμμετοχής είναι να βρίσκεται στο διάστημα τιμών [0 1]. Το σχήμα της επιλέγεται μεν αυθαίρετα αλλά και με τρόπο που να διασφαλίζει όσο είναι δυνατό την απλότητα.

Οι απλούστερες συναρτήσεις συμμετοχής είναι αυτές που σχηματίζονται από ευθείες γραμμές. Η απλούστερη από αυτές είναι η τριγωνική συνάρτηση συμμετοχής, που δεν είναι τίποτα άλλο από ένα τρίγωνο. Στην ίδια κατηγορία ανήκει και η τραπεζοειδής συνάρτηση συμμετοχής. Αυτές οι δύο συναρτήσεις εξασφαλίζουν την απαίτηση για απλότητα.



Σχήμα 4.3: *Απεικόνιση τριγωνικής και τραπεζοειδούς συναρτήσεως συμμετοχής.*

Η μαθηματική έκφραση της τριγωνικής συνάρτησης συμμετοχής είναι η εξής:

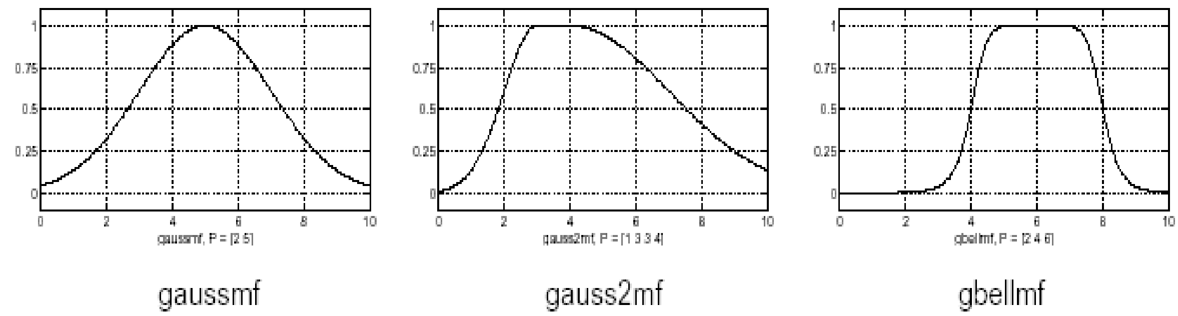
$$A = \begin{cases} 0, x \leq a \\ (x-a)/(b-a), x \in (a,b) \\ (c-x)/c-b), x \in (b,c) \\ 0, x \geq c \end{cases}$$

Παρακάτω δίδεται η μαθηματική έκφραση της τραπεζοειδούς συνάρτησης συμμετοχής:

$$A = \begin{cases} 0, x \leq a \\ (x-a)/(b-a), x \in (a,b) \\ 1, x \in (b,c) \\ (d-x)/d-c), x \in (c,d) \end{cases}$$

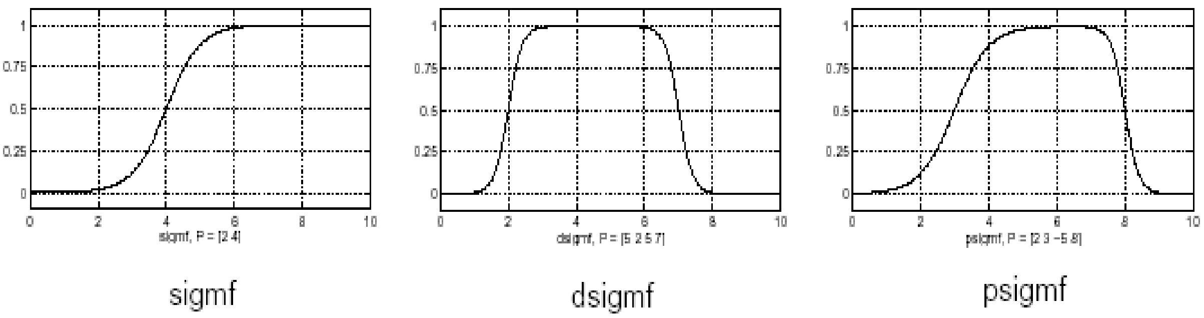
Δύο συναρτήσεις συμμετοχής που είναι δομημένες πάνω στη μορφή της κατανομής Gauss είναι μια απλή γκαουσιανή και μια σύνθεση δύο διαφορετικών γκαουσιανών.

Η γενικευμένη συνάρτηση συμμετοχής με μορφή καμπάνας έχει τρεις παραμέτρους, μια παραπάνω από την γκαουσιανή. Η γκαουσιανή και η καμπάνα μπορούν να χρησιμοποιούνται συχνά στα ασαφή σύνολα λόγο της ομαλότητάς τους. Έχουν δε το πλεονέκτημα να διατηρούν μη μηδενικές τιμές σε όλα τα σημεία.



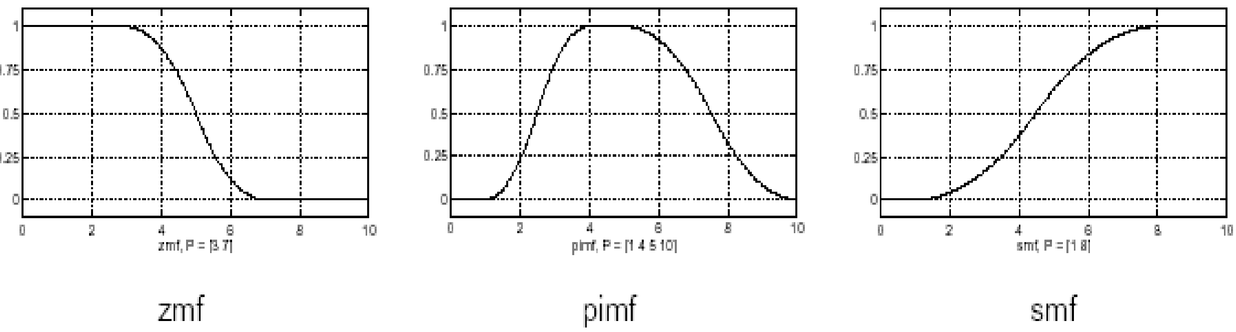
<u>Σχήμα 4.4</u>: *Απεικόνιση καμπανοειδούς και Gauss συναρτήσεως συμμετοχης.*

Παρά το γεγονός ότι η γκαουσιανή συνάρτηση συμμετοχής και η συνάρτηση καμπάνας επιτυγχάνουν ομαλή διακύμανση δε μπορούν ωστόσο να ορίσουν ασύμμετρες συναρτήσεις συμμετοχής που είναι χρήσιμες σε πολλά πρακτικά προβλήματα. Για τον λόγο αυτό χρησιμοποιείται η σιγμοειδής συνάρτηση συμμετοχής η οποία είναι ασύμμετρη και ανοικτή είτε προς τα δεξιά είτε προς τα αριστερά. Κλειστές συναρτήσεις συμμετοχής αυτού του τύπου μπορούν να παραχθούν αν συνθέσουμε δύο σιγμοειδείς. Έτσι προκύπτει η διαφορά μεταξύ δύο σιγμοειδών και το άθροισμα τους.



Σχήμα 4.5: *Παραλλαγές σιγμοειδούς συναρτήσεως συμμετοχής.*

Επίσης υπάρχουν πολλές πολυωνυμικές καμπύλες που τις χρησιμοποιούμε σα συναρτήσεις συμμετοχής. Τρεις από αυτές είναι η Ζ η S και η Π οι οποίες έχουν ονομασθεί έτσι εξαιτίας του σχήματός τους. Η Ζ είναι μια ασύμμετρη πολυωνυμική καμπύλη που είναι ανοικτή στα αριστερά, η S είναι η κατοπτρική της Ζ και η Π είναι μια ασύμμετρη κλειστή καμπύλη σχήματος Π.



Σχήμα 4.6: *Απεικόνιση Z, S, Π συναρτήσεως συμμετοχης.*

4.3 Ιδιότητες των ασαφών συνόλων

Έστω X ένα σύνολο αντικειμένων , του οποίου τα στοιχεία συμβολίζονται με το γράμμα x. Η συμμετοχή σε ένα υποσύνολο A του συνόλου X είναι μια συνάρτηση συμμετοχής μ_A από το X στο διάστημα [0 1]. Το A είναι ένα ασαφές υποσύνολο του X , το οποίο όμως δεν έχει αυστηρά καθορισμένα σύνορα. μ_A είναι ο βαθμός συμμετοχής του στοιχείου x στο A. Όσο πιο κοντά στο 1 είναι το μ_A τόσο πιο πολύ ανήκει το x στο A. Το σύνολο A μπορεί να προσδιοριστεί επομένως από το σύνολο των παρακάτω ζευγών:

$$\mathbf{A=\{(x,\mu_A(x)),\ x\in X\}}$$

Κάθε ζευγάρι (x,μ_A(x)) ονομάζεται **μονο-σύνολο.**
Χώρος αναφοράς του ασαφούς συνόλου Α είναι το σύνολο των στοιχείων του Χ που έχουν μη μηδενικό βαθμό συμμετοχής στο Α.

$$\mathbf{Supp\ A=\{x\in X|\mu_A(x)>0\}.}$$

Το σύνολο Α μπορεί να γραφεί και ως:

$$\mathbf{A=\mu_1/x_1+\mu_2/x_2+\ldots}$$

$$\mathbf{A=\Sigma\mu_i/x_i}$$

Στο συνεχή χώρο αναφοράς :

$$A=\int_x \mu_A(x)/x$$

Ασαφές δυναμοσύνολο

Ασαφές δυναμοσύνολο, F(x) , του υπερσυνόλου αναφοράς X, ονομάζεται το σύνολο όλων των ασαφών υποσυνόλων του Χ.

Υποσύνολο
Το σύνολο Α είναι υποσύνολο του Β $A\subseteq B$ αν και μόνο αν:

$$\mu_A(x)\le\mu_B(x),\ \forall x\in X$$

Αν ταυτόχρονα τα Α και Β δεν είναι ίσα , τότε το Α θα ονομάζεται γνήσιο υποσύνολο του Β

Ασαφής διαμέριση

Μια οικογένεια ασαφών υποσυνόλων του Χ , θα λέγεται ασαφής διαμέριση $P^n(X)$ του Χ τάξης n (n$\in$N) και θα συμβολίζεται με $A^n=\{A_1,A_2,\ldots,A_n\}$, αν και μόνο αν:

$$A_j \neq A_i, \forall i, j \in N_n (i \neq j)$$

$$0 < \sum_{k=i}^{m} A_i(x_k) < m, \forall i \in N_n$$

Τα στοιχεία A_i $i \in N_n$ της A_n θα λέγονται κλάσεις της ασαφούς διαμέρισης.

Κενό ασαφές σύνολο

Ένα ασαφές σύνολο με χώρο αναφοράς τον X , λέγεται κενό αν για κάθε στοιχείο x που ανήκει στον X , η συνάρτηση συμμετοχής του A είναι μηδέν.

$$A \equiv 0 \quad \text{αν} \quad \mu_{\mathrm{A}}(x) = 0 \quad \forall x \in X$$

Κανονικό ασαφές σύνολο

Ένα ασαφές σύνολο A που ορίζεται στο χώρο αναφοράς X , λέγεται κανονικό αν υπάρχει τουλάχιστο ένα στοιχείο x του X για το οποίο η συνάρτηση συμμετοχής να παίρνει τιμή ίση με τη μονάδα.

A=κανονικό αν $\exists x_1 : \mu_{\mathrm{A}}(x_1) = 1$

Ισότητα ασαφών συνόλων

Δύο ασαφή σύνολα A και B που ορίζονται στο χώρο αναφοράς X , λέγονται ίσα αν για κάθε στοιχείο x του X οι συναρτήσεις συμμετοχής των A και B είναι ίσες. Δηλαδή:

A=B αν $\mu_{\mathrm{A}}(x) = \mu_{\mathrm{B}}(x) \quad \forall x \in X$

Συστολή ασαφών συνόλων

Έστω ασαφές σύνολο A που ορίζεται στο χώρο αναφοράς X. Η συστολή, CON(A) του συνόλου αυτού είναι ένα νέο ασαφές σύνολο με συνάρτηση συμμετοχής που ορίζεται ως εξής:

$$\mu_{CON(A)}(x) = (\mu_A(x))^2$$

Η συστολή του ασαφούς συνόλου αντιστοιχεί στην προσθήκη του όρου πολύ , μπροστά από τη λεκτική μεταβλητή που περιγράφει το ασαφές σύνολο. Δηλαδή η συστολή του ασαφούς συνόλου ψηλός είναι ένα νέο ασαφές σύνολο που αντιστοιχεί στην έννοια πολύ ψηλός.

Διαστολή ασαφών συνόλων

Αντιστοίχως η διαστολή ενός ασαφούς συνόλου A , είναι ένα νέο ασαφές σύνολο που συμβολίζεται με DIL(A) και έχει συνάρτηση συμμετοχής που περιγράφεται από την παρακάτω σχέση:

$$\mu_{DIL(A)}(x) = \sqrt{\mu_A(x)}$$

Η διαστολή του ασαφούς συνόλου αντιστοιχεί στην προσθήκη του όρου λίγο , μπροστά από τη λεκτική μεταβλητή που περιγράφει το ασαφές σύνολο. Δηλαδή η διαστολή του ασαφούς συνόλου ψηλός είναι ένα νέο ασαφές σύνολο που αντιστοιχεί στην έννοια λίγο ψηλός

4.4 Λογικες πραξεις στα ασαφή σύνολα

Μέχρι τώρα έχουμε αναφερθεί στα ασαφή σύνολα και καθόλου στην ασαφή λογική. Η ασαφής λογική δεν είναι παρά ένα υπερσύνολο της λογικής Boolean. Αν δηλαδή απομονώσουμε τους ακραίους βαθμούς συμμετοχής 0 (πλήρως ψευδές) και 1 (πλήρως αληθές) τότε οι κλασσικοί λογικοί τελεστές μπορούν να εφαρμοσθούν. Για παράδειγμα ας θεωρήσουμε τον παρακάτω πίνακα αληθείας:

A	B	A and B
0	0	0
0	1	0
1	0	0
1	1	1

AND

A	B	A or B
0	0	0
0	1	1
1	0	1
1	1	1

OR

A	not A
0	1
1	0

NOT

Σχήμα 4.7: *Απεικόνιση πίνακα αληθείας.*

Μεταβαίνοντας στον χώρο της ασαφούς λογικής πρέπει να έχομε κατά νου ότι οι έννοιες αληθές και ψευδές είναι θέμα βαθμού συμμετοχής. Επομένως ο πίνακας αυτός πρέπει να μετατραπεί με τρόπο που να συμπεριλαμβάνει αυτήν την αρχή. Οι τιμές των εισόδων A και B είναι τώρα πραγματικοί αριθμοί από το 0 μέχρι το 1. Πρέπει λοιπόν να ευρεθεί μια συνάρτηση που να διατηρεί τις ιδιότητες της συνάρτησης AND και ταυτόχρονα να μπορεί να επεκτείνεται για πραγματικούς αριθμούς. Μια πιθανή απάντηση μπορεί να είναι ο τελεστής min(A,B), η ελάχιστη δηλαδή τιμή των εισόδων A και B. Με βάση το ίδιο σκεπτικό μια συνάρτηση που μπορεί να αντικαταστήσει τον τελεστή OR της Boolean λογικής είναι ο τελεστής max(A,B). Τέλος ο τελεστής NOT A μπορεί να αντικατασταθεί με την πράξη 1-A. Παρακάτω βλέπουμε ότι ο πίνακας αληθείας παραμένει αμετάβλητος αν εφαρμόσουμε τις παραπάνω υποκαταστάσεις.

A	B	min(A,B)
0	0	0
0	1	0
1	0	0
1	1	1

AND

A	B	max(A,B)
0	0	0
0	1	1
1	0	1
1	1	1

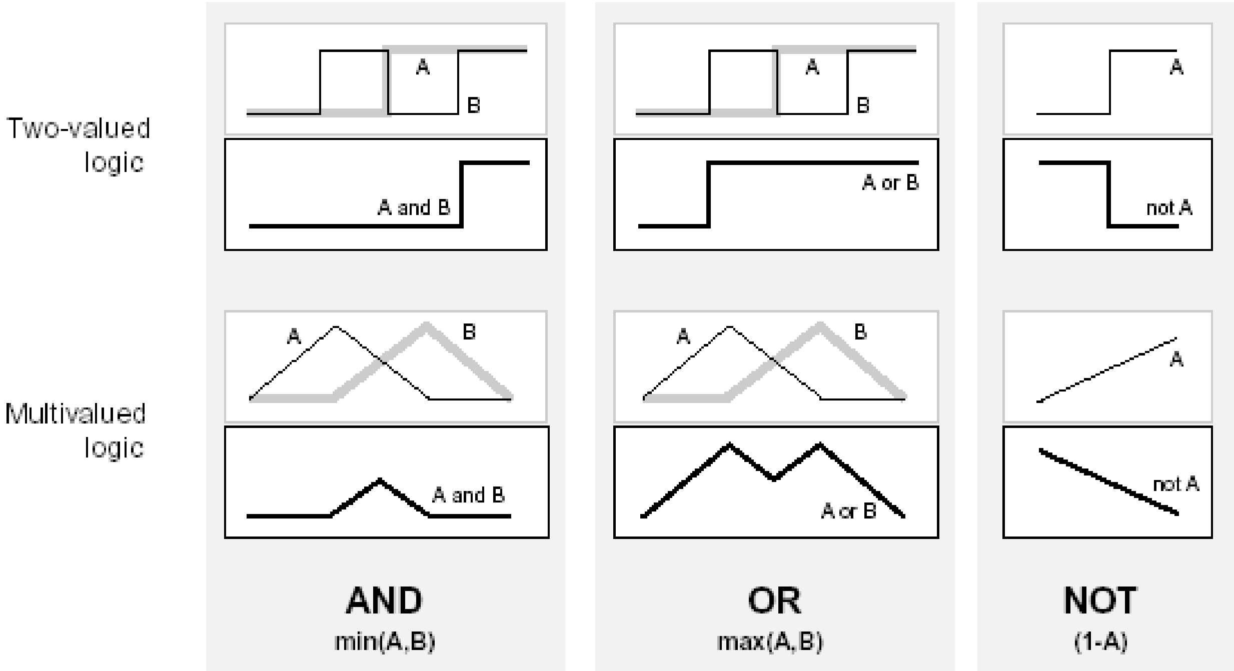
OR

A	1 - A
0	1
1	0

NOT

Σχήμα 4.8: *Απεικόνιση πίνακα αληθείας μετά την εφαρμογή των παραπάνω.*

Αφού έχουμε ορίσει συναρτήσεις που μπορούν ορίσουν τον πίνακα αληθείας, μπορούμε πλέον να επεκταθούμε και στην περίπτωση πραγματικών αριθμών. Στο παρακάτω σχήμα έχουμε αντικαταστήσει τον πίνακα αληθείας με ένα γράφημα δύο ασαφών συνόλων. Στο πάνω μέρος έχουμε την περίπτωση συνόλων με δύο τιμές ενώ στο κάτω φαίνεται ο τρόπος που φαίνεται πως λειτουργούν οι τελεστές στην περίπτωση που οι τιμές αληθείας A και B μεταβάλλονται συνεχώς από το 0 στο 1.



Σχήμα 4.9: *Σύνολα με δύο τιμές και τρόπος λειτουργίας οι τελεστών για συνεχώς μεταβαλλόμενες τιμές αλήθειας.*

Δεδομένων των παραπάνω συναρτήσεων μπορούμε να κατασκευάσουμε δομές με βάση ασαφή σύνολα και τους λογικούς κανόνες AND OR και NOT. Βέβαια το γεγονός ότι βρήκαμε συναρτήσεις που να επεκτείνουν τους τελεστές αυτούς από το χώρο της Boolean λογικής στο χώρο της ασαφούς λογικής δε σημαίνει σε καμία περίπτωση ότι οι συναρτήσεις αυτές είναι και οι μοναδικές. Έχουμε ορίσει την τομή ,την ένωση και το συμπλήρωμα ενός ασαφούς συνόλου με τους τελεστές min max 1-A οι όποιοι είναι οι πιο συχνά χρησιμοποιούμενοι. Ωστόσο μπορούμε να χρησιμοποιήσουμε και άλλες συναρτήσεις για τον ίδιο σκοπό.

Γενικά η τομή δύο ασαφών συνόλων μπορεί να παρασταθεί με μια δυαδική απεικόνιση T που αθροίζει (aggregates) τις δύο συναρτήσεις συμμετοχής ως ακολούθως:

$$\mu_{\mathrm{A}\cap\mathrm{B}} = T(\mu_{\mathrm{A}}(x), \mu_{\mathrm{B}}(x))$$

Για παράδειγμα ο δυαδικός τελεστής T μπορεί να αντιπροσωπεύει τον πολλαπλασιασμό των $\mu_A(x)$ και $\mu_B(x)$. Αυτού του είδους οι τελεστές τομής αναφέρονται συνήθως ως τ-νόρμες (τριγωνική νόρμα , triangular norm) και πρέπει να ικανοποιούν τις παρακάτω προϋποθέσεις:

- Οριακές συνθήκες: $T(0, 0) = 0$, $T(a, 1) = T(1, a) = a$

- Μονοτονία: $T(a, b) \leq T(c, d)$ *αν* $a \leq c$ *και* $b \leq d$

- Αντιμεταθετικότητα: $T(a, b) = T(b, a)$

- Προσεταιριστικότητα: $T(a, T(b, c)) = T(T(a, b), c)$

Η πρώτη απαίτηση επιβάλει τη γενίκευση σε σύνολα διακριτών τιμών (crisp sets).

Η δεύτερη υπονοεί ότι μια μείωση του βαθμού συμμετοχής στο Α ή στο Β δεν μπορεί να οδηγήσει σε αύξηση του βαθμού συμμετοχής στην τομή των Α και Β.

Η τρίτη απαίτηση δείχνει ότι ο τελεστής είναι αδιάφορος στον τρόπο διάταξης με τον οποίο τέμνονται τα δύο σύνολα Α και Β.

Τέλος η τέταρτη απαίτηση μας επιτρέπει να παίρνουμε την τομή οποιουδήποτε αριθμού συνόλων σε οποιαδήποτε διάταξη ζευγών.

Μερικοί από τους πιο γνωστούς τελεστές T-norm είναι:

Συνήθης τομή T(a,b)=min(a,b)

Αλγεβρικό γινόμενο T(a,b)=ab

Φραγμένη διαφορά T(a,b)=max(0,a+b-1)

Συνάρτηση Hamacher Η συνάρτηση Hamacher δίνεται από τη σχέση:

$$t(a,b) = \frac{ab}{r + (1-r)(a+b-ab)}$$

Όπως η ασαφής τομή έτσι και η ασαφής ένωση .προσδιορίζεται με μια δυαδική απεικόνιση S.

$$\mu_{A \cup B}(x) = S(\mu_A(x), \mu_B(x))$$

Για παράδειγμα ο τελεστής S μπορεί να είναι το άθροισμα των $\mu_A(x)$ και $\mu_B(x)$. Οι ασαφείς τελεστές ένωσης αναφέρονται συνήθως ως σ-νόρμες και πρέπει να πληρούν τις παρακάτω προϋποθέσεις:

- Οριακές συνθήκες: $S(1, 1) = 1, S(a, 0) = S(0, a) = a$
- Μονοτονία: $S(a, b) \leq S(c, d)$ αν $a \leq c$ και $b \leq d$
- Αντιμεταθετικότητα: $S(a, b) = S(b, a)$
- Προσεταιριστικότητα: $S(a, S(b, c)) = S(S(a, b), c)$

Παραδείγματα σ-νορμών είναι τα εξής:

Συνήθης ένωση s(a,b)=max(a,b)

Αλγεβρικό άθροισμα s(a,b)=a+b-ab

Φραγμένο άθροισμα s(a,b)=min(1,a+b)

Συμπλήρωμα

Το συμπλήρωμα A' ενός ασαφούς συνόλου A δίνεται από τη σχέση:

$$\mu_{\overline{A}}(x) = c(\mu_{\mathrm{A}}(x))$$

Όπου η συνάρτηση c πρέπει να ικανοποιεί τις παρακάτω συνθήκες:

- Οριακές συνθήκες: c(0)=1 και c(1)=0
- Μονοτονία: $\forall a, b \in [0,1]$, αν $a \leq b \Rightarrow c(a) \geq c(b)$
- Συνέχεια: c συνεχής στο [0,1].
- Εναγωγή: $\forall a \in [0,1]$ είναι c(c(a))=a

Το σύνηθες συμπλήρωμα δίνεται από τη σχέση:

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x)$$

4.5 Ασαφείς σχέσεις

Οι ασαφείς σχέσεις (fuzzy relations) είναι ασαφή σύνολα ορισμένα σε πεδία αναφοράς ανώτερης διάστασης (π.χ. X x X ,X x Y x Z κλπ). Ποιοτικά , μια ασαφής σχέση R θα μπορούσε να είναι μια έκφραση της μορφής «είναι βαρύτερο από» και η οποία θα συνδέει τα στοιχεία δύο άλλων συνόλων:

R = «x είναι βαρύτερο από y» x∈X, y∈Y και R∈ X x Y

Οι ασαφείς σχέσεις μπορεί να εκφραστούν με αναφορά όλων των ζευγών (τιμή, βαθμός συμμετοχής) , δηλαδή ζευγών της μορφής $((x,y),\mu_R(x,y))$. Ένας άλλος τρόπος αναπαράστασης , ιδιαίτερα χρήσιμος σε υπολογισμούς είναι σε μορφή πίνακα:

$$R = \begin{bmatrix} \mu_R(x_{1,}y_1) & \mu_R(x_1,y_2) & \cdots & \mu_R(x_1,y_n) \\ \mu_R(x_2,y_1) & \mu_R(x_2,y_2) & \cdots & \mu_R(x_2,y_n) \\ \cdot & & & \\ \cdot & & & \\ \mu_R(x_m,y_1) & \mu_R(x_m,y_2) & \cdots & \mu_R(x_m,y_n) \end{bmatrix}$$

Οι ασαφείς σχέσεις μπορούν να συνδυαστούν μεταξύ τους μέσω της διαδικασίας της σύνθεσης (composition). Αν για παράδειγμα συνδυαστεί η ασαφής σχέση $R_1(x,y)$ ορισμένη στο X x Y με την ασαφή σχέση $R_2(y,z)$ ορισμένη στο Y x Z τότε θα προκύψει μία ασαφής σχέση R(x,z) η οποία θα ορίζεται στο σύνολο X x Z και θα συσχετίζει άμεσα στοιχεία των συνόλων X και Z. Βέβαια είναι απαραίτητο να προσδιοριστεί επακριβώς η συνάρτηση συμμετοχής $\mu_R(x,z)$ της R με χρήση των συναρτήσεων συμμετοχής των R_1 και R_2.

Οι βασικές πράξεις που ορίζονται μεταξύ των ασαφών σχέσεων είναι η αντιστροφή και η σύνθεση.

- ***Αντιστροφή***

Αντίστροφη σχέση της R(X,Y) είναι η ασαφής σχέση $R^{-1}(Y,X)$ με τύπο: $R^{-1}(y,x) = R(x,y)$ για κάθε x που ανήκει στον X και κάθε y που ανήκει στον Y. Ο πίνακας συμμετοχής που παριστάνει την R^{-1} είναι ο ανάστροφος του R.

- ***Σύνθεση***

Η σύνθεση είναι πολύ σημαντική διαδικασία καθώς όπως θα δούμε παρακάτω οι κανόνες της μορφής if - then αντιστοιχούν σε ασαφείς σχέσεις και το πρόβλημα της ασαφούς συλλογιστικής είναι μαθηματικά ισοδύναμο με τη σύνθεση.

Αν $R_1(x,y)$ και $R_2(y,z)$ είναι δύο ασαφείς σχέσεις ορισμένες στα σύνολα X x Y και Y x Z αντίστοιχα , τότε η σύνθεση τους δίνει μια νέα σχέση $R_1 \circ R_2$

Η sup-t σύνθεση R: X x Y → [0,1] δύο ασαφών σχέσεων R1: X x Y→ [0,1] και R2: X x Y → [0,1] ορίζεται από την εξίσωση:

$$R(x,y) = \left[R_1 \circ^t R_2\right](x,y) = \sup_{y \in Y} t\left[R_1(x,y), R_2(y,z)\right]$$

Οι περισσότερο γνωστές μέθοδοι σύνθεσης ασαφών συνόλων είναι η σύνθεση max-min και η σύνθεση max-product. Η συνάρτηση συμμετοχής για την περίπτωση της max-min σύνθεση δίνεται από τη σχέση:

$$\mu_{R_1 \circ R_2}(x,z) = \vee_y \left[\mu_{R_1}(x,y) \wedge \mu_{R_2}(y,z)\right]$$

Στην περίπτωση της σύνθεσης max-product έχομε την παρακάτω σχέση:

$$\mu_{R_1 \circ R_2}(x,z) = \vee_y \left[\mu_{R_1}(x,y) \bullet \mu_{R_2}(y,z)\right]$$

Οι υπολογισμοί στο δεξιό μέρος των παραπάνω σχέσεων είναι παρόμοιοι με του πολλαπλασιασμού των πινάκων.

4.6 Ασαφής συλλογιστική.

Η βάση στην οποία στηρίζεται η λήψη αποφάσεων είναι η παραγωγή συλλογιστικής. Η ασαφής λογική ασχολείται με την παραγωγή συλλογιστικής σε περιβάλλον αβεβαιότητας. Για το σκοπό αυτό , θεμελιώνεται η δομή και η μαθηματική αναπαράσταση ενός ασαφούς γεγονότος με τον ορισμό των ασαφών συνόλων και καθορίζεται ο τρόπος με τον οποίο συνδυάζουμε τα γεγονότα για να παράγουμε λογικές προτάσεις ή σχέσεις και συνεπώς συμπεράσματα.

Οι συλλογιστικοί τρόποι που κυρίως χρησιμοποιούνται είναι τρεις:

- Ο modus ponens (MP)
- Ο modus tolens (MT)
- Ο υποθετικός συλλογισμός (HS)

Ο **modus ponens** παράγει συμπεράσματα από ένα σύνολο υποθέσεων σύμφωνα με το σχήμα:

$$(A \Rightarrow B) \wedge A \Rightarrow B$$

όπου Α και Β συγκεκριμένα γεγονότα.

Ο παραπάνω κανόνας ερμηνεύεται ως εξής: Αν το γεγονός Α συνεπάγεται το γεγονός Β και επιπλέον έχουμε ως υπόθεση ότι ισχύει το Α , τότε το συμπέρασμα που παίρνουμε είναι ότι ισχύει και το Β. Όμως σε περιβάλλον ασάφειας τα γεγονότα ισχύουν σε κάποιο βαθμό. Έτσι ο παραπάνω κανόνας πρέπει να τροποποιηθεί για να συμπεριλάβει και την ασάφεια. Καταλήγουμε μ ‘αυτό τον τρόπο στο γενικευμένο κανόνα modus ponens (generalized modus ponens) ο οποίος έχει την παρακάτω μορφή:

$$(A \Rightarrow B) \wedge A' \Rightarrow B'$$

Αντίστοιχα ο γενικευμένος κανόνας modus tolens (generalized modus tolens) διατυπώνεται ως εξής:

$$(A \Rightarrow B) \wedge B' \Rightarrow A'$$

Ας εξετάσουμε τώρα την ερμηνεία των παραπάνω κανόνων. Για παράδειγμα ο γενικευμένος κανόνας modus ponens ερμηνεύεται ως εξής: Αν το γεγονός Α συνεπάγεται το γεγονός Β και έχομε ως υπόθεση ότι ισχύει το Α σε κάποιο βαθμό , τότε θα ισχύει και το Β σε κάποιο βαθμό. Πρέπει, λοιπόν να προσδιοριστεί ο βαθμός στον οποίο πληρείται το γεγονός Β. Αυτός εξαρτάται από το βαθμό στον οποίο πληρείται το γεγονός Α και από το είδος της συνεπαγωγής που εφαρμόζουμε. Η πράξη της ασαφούς συνεπαγωγής υλοποιεί τη μαθηματική σχέση Α ==> Β, όταν τα Α και Β είναι ασαφή γεγονότα..

Το σχήμα που προτείνεται από τη θεωρία των ασαφών συνόλων για την εξαγωγή του γεγονότος Β’ από τα γεγονότα Α ,Α’,Β περιγράφεται από την παρακάτω σχέση που αποτελεί το συνθετικό κανόνα του Zadeh:

$$B'(y) = \sup_{x \in X} t[A'(x), \sigma(A(x), B(y))]$$

Όπου: σ είναι μια συνάρτηση που υλοποιεί την πράξη της ασαφούς συνεπαγωγής.

Επομένως η επιλογή της συνάρτησης που υλοποιεί την ασαφή συνεπαγωγή έχει ουσιαστικό ρόλο για την ασαφή συλλογιστική που παράγεται με βάση το παραπάνω σχήμα. Ένα από τα κριτήρια που χρησιμοποιείται γι’ αυτή την επιλογή είναι το κριτήριο της ανάκλησης (recall) , το οποίο διατυπώνεται ως εξής:

$$B(y) = \sup_{x \in X} t[A(x), \sigma(A(x), B(y))]$$

Το κριτήριο αυτό έχει την εξής ερμηνεία: η ασαφής συνεπαγωγή πρέπει να είναι τέτοια ώστε όταν η υπόθεση πληρείται ακριβώς , τότε να λαμβάνουμε το συμπέρασμα του κανόνα Α=>Β , δηλαδή το γεγονός Β. η απαίτηση αυτή είναι εύλογη αφού όταν δεν υπάρχει αβεβαιότητα η ασαφής συλλογιστική οφείλει να ταυτίζεται με την κλασσική συλλογιστική.

ΚΕΦΑΛΑΙΟ 5^{Ο}

Συστήματα ασαφούς συλλογιστικής-νευροασαφή συστήματα

5.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο, μελετήσαμε τα ασαφή σύνολα και τις πράξεις που γίνονται πάνω σ' αυτά. Στο παρόν κεφάλαιο εξετάζουμε τον τρόπο που μπορούν να εφαρμοσθούν τα παραπάνω σε ένα σύστημα λήψης αποφάσεων.

Θα εξετάσουμε ένα απλοϊκό παράδειγμα λήψης αποφάσεων και θα δούμε πως μπορεί να λυθεί με τη βοήθεια της ασαφούς λογικής. Εξετάζουμε το πρόβλημα του φιλοδωρήματος, πόσο δηλαδή είναι το φιλοδώρημα που πρέπει να δώσει ένας πελάτης σε ένα εστιατόριο. Οι παράμετροι του προβλήματος είναι η ποιότητα του φαγητού και η εξυπηρέτηση.

Είναι προφανές για παράδειγμα ότι αν το φαγητό είναι άριστης ποιότητας και η εξυπηρέτηση καλή τότε το φιλοδώρημα θα είναι υψηλό. Αντίστοιχα αν η ποιότητα της εξυπηρέτησης είναι χαμηλή και το φαγητό κακό τότε το φιλοδώρημα θα είναι χαμηλό. Τι γίνεται όμως αν έχουμε μια ενδιάμεση κατάσταση, π.χ. μέτριο φαγητό και καλή εξυπηρέτηση. Επιπλέον παραπάνω έχουμε αναφέρει προσδιορισμούς όπως καλό , καλό , μέτριο που εμπεριέχουν κάποια ασάφεια.

Αν προσπαθήσουμε να λύσουμε το πρόβλημα γραμμικά θα πρέπει να δώσουμε καθορισμένους κανόνες για κάθε μια περίπτωση. Μέσω όμως της ασαφούς λογικής το πρόβλημα αυτό μπορεί να λυθεί ευκολότερα. Πριν προχωρήσουμε στη λύση του προβλήματος θα δούμε πως μπορούμε να υλοποιήσουμε τους κανόνες αν-τότε (if-then rules) που προαναφέραμε (αν το φαγητό είναι κακό κτλ).

5.2 Κανόνες εαν-τότε (ιf-then rules)

Τα ασαφή σύνολα και οι ασαφείς τελεστές είναι τα υποκείμενα και τα ρήματα του ασαφούς λογισμού. Οι προτάσεις if-then χρησιμοποιούνται για να σχηματίσουν τις συνθήκες εκείνες που συνιστούν την ασαφή λογική. Ένας απλός κανόνας if-then έχει τη μορφή:

if x is A then y is B

Όπου Α και Β είναι οι γλωσσικές μεταβλητές που προσδιορίζονται από ασαφή σύνολα με χώρο αναφοράς X και Y αντίστοιχα. Το πρώτο κομμάτι του κανόνα 'if x is A' ονομάζεται υπόθεση (antecedent ή premise) ενώ το δεύτερο κομμάτι 'then y is B' ονομάζεται συνέπεια ή συμπέρασμα (consequent ή conclusion). Ένα παράδειγμα ενός τέτοιου κανόνα μπορεί να είναι το εξής;

if service is good then tip is average

Η λέξη ***good*** αντιπροσωπεύεται από έναν αριθμό που κυμαίνεται ανάμεσα στο 0 και το 1. Έτσι η υπόθεση είναι μια μετάφραση που επιστρέφει μια τιμή από το 0 έως το 1. Από την άλλη μεριά η λέξη ***average*** αποτελεί ένα ασαφές σύνολο και έτσι η συνέπεια του κανόνα είναι μια συσχέτιση της εξόδου y στο ασαφές σύνολο B. Στους κανόνες if-then η λέξη 'is' έχει διαφορετική έννοια ανάλογα με το αν εμφανίζεται στο πρώτο ή στο δεύτερο μέρος του κανόνα. Έτσι όταν εμφανίζεται στο πρώτο μέρος του κανόνα έχει την έννοια του ελέγχου ισότητας (δηλαδή την ίδια έννοια που δίνουμε στο σύμβολο = =). Αντίθετα όταν εμφανίζεται στο δεύτερο μέρος του κανόνα έχει την έννοια της καταχώρησης (δηλαδή την έννοια που δίνουμε στο σύμβολο =). Ο παραπάνω κανόνας λοιπόν θα μπορούσε να γραφεί συμβολικά και ως εξής:

if service == good then tip = average

Γενικά, η είσοδος σε έναν κανόνα if-then είναι η τρέχουσα τιμή της μεταβλητής εισόδου ενώ η έξοδος του κανόνα είναι ολόκληρο το ασαφές σύνολο. Από αυτό όμως το ασαφές σύνολο θα πρέπει να κρατήσουμε μια τιμή έτσι ώστε να μπορούμε να βγάλουμε κάποια συμπεράσματα. Γι' αυτό πρέπει να γίνει μια διαδικασία που ονομάζεται αποσαφήνιση (defuzzyfication). Για την αποσαφήνιση αναφερόμαστε εκτενέστερα παρακάτω. Για να διερμηνεύσουμε έναν κανόνα If-then πρέπει να ακολουθήσουμε κάποια στάδια:

1. Πρώτα πρέπει να υπολογίσουμε την είσοδο. Για να γίνει αυτό όμως πρέπει να γίνει ασαφοποίηση της εισόδου και να εφαρμοσθούν οι κατάλληλοι ασαφείς τελεστές.

2. Να εφαρμόσουμε το παραπάνω αποτέλεσμα στο δεύτερο μέρος του κανόνα.

Αν πάρουμε την απλή περίπτωση των δυαδικών τιμών οι κανόνες if-then είναι αρκετά εύκολο να εφαρμοσθούν. Αν η προϋπόθεση είναι αληθής τότε και η συνέπεια θα είναι αληθής. Πώς όμως επεκτείνεται η παραπάνω συλλογιστική για να συμπεριλάβει και ασαφείς τιμές. Όταν η προϋπόθεση είναι αληθής σε κάποιο βαθμό τότε και η συνέπεια είναι αληθής στον ίδιο βαθμό. Με άλλα λόγια:

Στη δυαδική λογική: $p \rightarrow q$ (Τα p και q είναι και τα δύο είτε πλήρως αληθή ή πλήρως ψευδή).

Στην ασαφή λογική: $0.5\ p \rightarrow 0.5\ q$ (Τα p και q είναι μερικώς αληθή και μερικώς ψευδή).

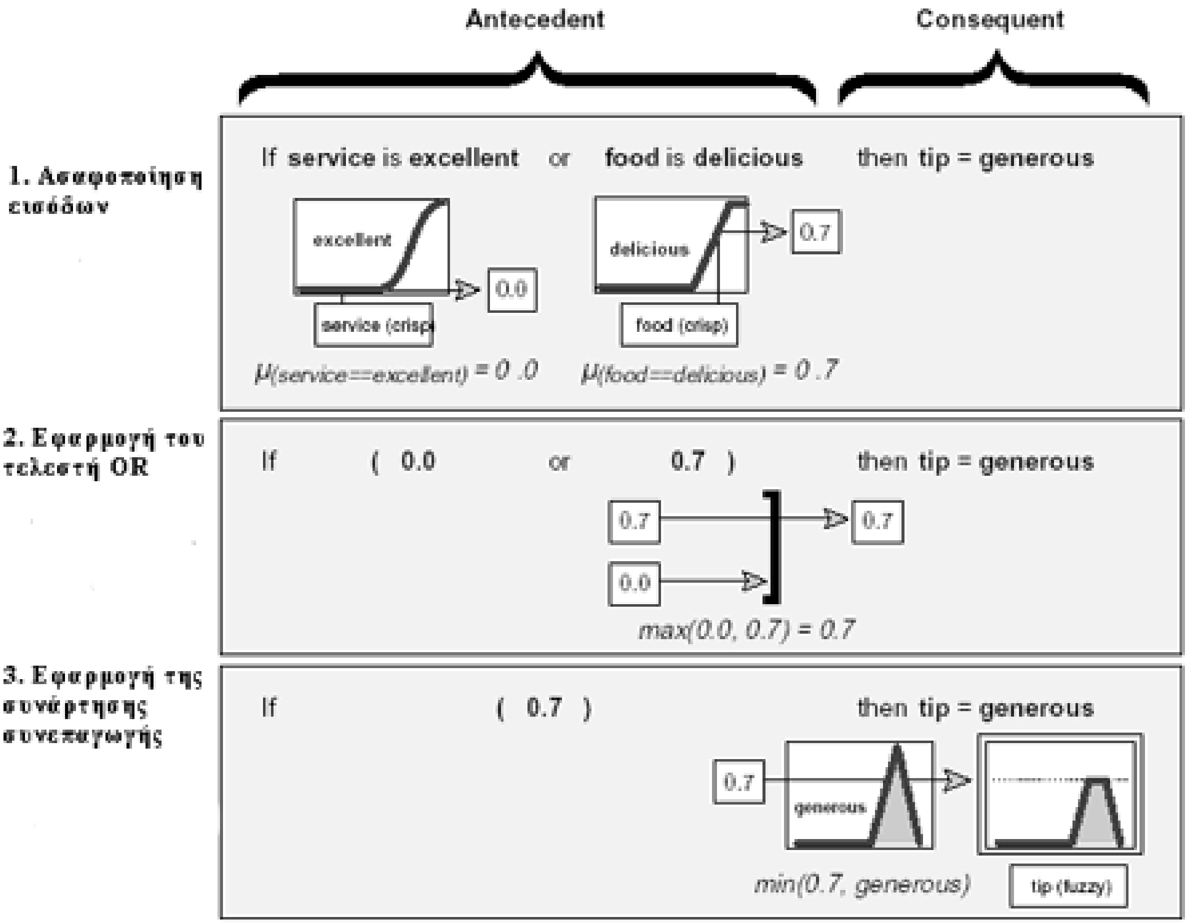
Η προϋπόθεση ενός κανόνα μπορεί να αποτελείται από περισσότερα του ενός μέρη. Όπως για παράδειγμα ο παρακάτω κανόνας:

if sky is gray and wind is strong and barometer is falling, then …

Σ' αυτή την περίπτωση όλα τα μέρη της προϋπόθεσης υπολογίζονται ταυτόχρονα και καταλήγουμε σε ένα μοναδικό νούμερο ανάλογα με τους λογικούς τελεστές που χρησιμοποιούμε. Αλλά και η συνέπεια ενός κανόνα μπορεί να αποτελείται από περισσότερα του ενός μέρη. Για παράδειγμα ο παρακάτω κανόνας:

if temperature is cold then hot water valve is open and cold water valve is shut.

Όλα τα μέρη της συνέπειας του κανόνα επηρεάζονται το ίδιο από το αποτέλεσμα της προϋπόθεσης. Η συνέπεια του κανόνα προσδιορίζει ένα ασαφές σύνολο το οποίο συσχετίζεται με την έξοδο. Η συνάρτηση συνεπαγωγής (*implication function*) στη συνέχεια τροποποιεί το ασαφές σύνολο στο βαθμό που προσδιορίζεται από την προϋπόθεση του κανόνα. Μια πολύ συνηθισμένη συνάρτηση για να γίνει αυτό είναι η αποκοπή με τη χρήση του τελεστή min. Στο παρακάτω σχήμα βλέπουμε τον τρόπο με τον οποίο εφαρμόζονται οι κανόνες if-then.



Σχήμα 5.1: *Κανόνες if-then στην ασαφή λογική.*

5.3 Συστήματα ασαφούς συλλογιστικής (fuzzy inference systems)

Γενικά ένας κανόνας από μόνος του δεν επαρκεί για τα περισσότερα πρακτικά προβλήματα. Αυτό που χρειαζόμαστε είναι περισσότεροι του ενός κανόνες οι οποίοι να αλληλεπιδρούν μεταξύ τους. Η έξοδος κάθε κανόνα θα είναι ένα ασαφές σύνολο. Τα ασαφή σύνολα που προκύπτουν σαν συνέπεια των κανόνων συγκεντρώνονται για να σχηματίσουν ένα ασαφές σύνολο εξόδου από το οποίο μέσω της διαδικασίας της αποσαφήνισης θα πάρουμε μια μοναδική τιμή. Παρακάτω αναπτύσσεται ο τρόπος με τον οποίο εξελίσσεται η παραπάνω διαδικασία μέσα από ένα σύστημα ασαφούς συλλογιστικής (fuzzy inference system, FIS) που ονομάζεται Mamdani.

5.3.1 Συστήματα Mamdani

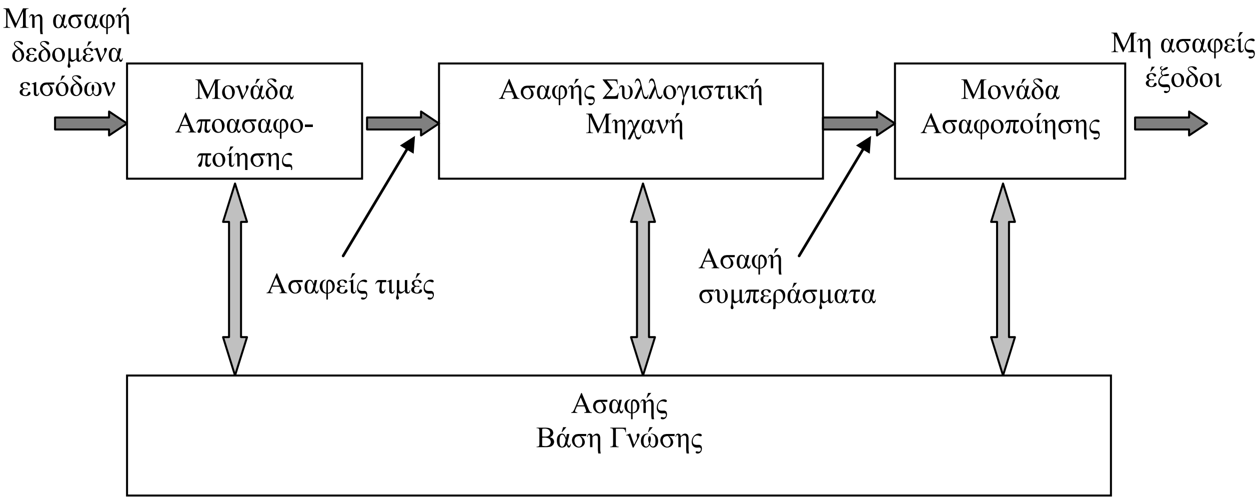
Η ασαφής συλλογιστική είναι η διαδικασία κατά την οποία σχηματίζεται η απεικόνιση από μια δεδομένη είσοδο σε μια έξοδο , χρησιμοποιώντας τις αρχές της ασαφούς λογικής. Η απεικόνιση αυτή θέτει τις βάσεις με τις οποίες μπορεί να ληφθεί μια απόφαση ή να διακριθούν πρότυπα (patterns). Η διαδικασία αυτή της ασαφούς συλλογιστικής περιλαμβάνει όλα εκείνα τα κομμάτια στα οποία αναφερθήκαμε προηγουμένως. Δηλαδή τις συναρτήσεις συμμετοχής, τους ασαφείς λογικούς τελεστές και τους κανόνες if-then. Υπάρχουν δύο είδη ασαφών συστημάτων παρεμβολής τα τύπου-Mamdani και τα τύπου Sugeno. Οι διαφορές ανάμεσα στα δύο αυτά συστήματα έγκεινται κυρίως στον τρόπο με τον οποίο σχηματίζεται η έξοδος του συστήματος. Τα συστήματα Mamdani είναι αυτά που συναντώνται συχνότερα. Η μεθοδολογία των συστημάτων Mamdani εισήχθηκε το 1975 από τον Ebrahim Mamdani.

Εφαρμογές των FIS έχουμε σε τομείς όπως ο αυτόματος έλεγχος, η ταξινόμηση δεδομένων, η ανάλυση αποφάσεων, τα έμπειρα συστήματα και η όραση υπολογιστών. Ένα Ασαφές Σύστημα (Fuzzy System) αποτελείται από τις ακόλουθες δομικές μονάδες:

- Μία βάση ασαφών κανόνων της μορφής ΕΑΝ - ΤΟΤΕ (Ασαφής Βάση Γνώσης).
- Μία ασαφή συλλογιστική μηχανή. Αυτή υλοποιεί την εξαγωγή συμπερασμάτων από τις εισόδους του συστήματος με βάση την ασαφή βάση γνώσης.
- Μία μονάδα ασαφοποίησης (ασαφοποιητική μονάδα διεπαφής), η οποία μετατρέπει τα δεδομένα εισόδου σε ασαφή σύνολα.
- Μία μονάδα αποασαφοποίησης (αποασαφοποιητική μονάδα διεπαφής), η οποία μετατρέπει τα ασαφή συμπεράσματα σε σαφώς καθορισμένη μορφή.

Ένα βασικό FIS μπορεί να πάρει είτε ασαφείς εισόδους, είτε σαφείς (συγκεκριμένες) με τη μορφή ασαφών singletons. Η έξοδος όμως που παράγει είναι πάντα ασαφή σύνολα. Μερικές φορές είναι αναγκαίο να έχουμε σαφή έξοδο, ειδικά όταν το FIS χρησιμοποιείται σαν ελεγκτής. Τότε χρειαζόμαστε μια μέθοδο αποσαφοποίησης, για να εξάγουμε ένα σαφές αποτέλεσμα που να αντιπροσωπεύει με τον καλύτερο τρόπο το ασαφές σύνολο.

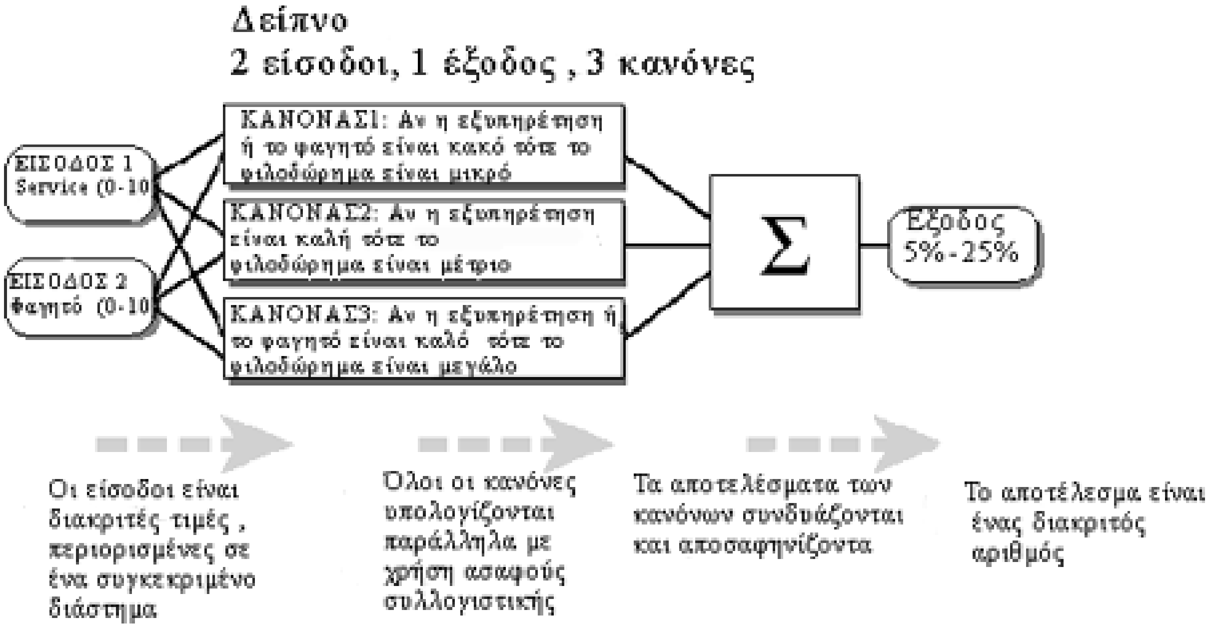
Με σαφείς εισόδους και εξόδους, ένα FIS υλοποιεί μια μη-γραμμική αντιστοίχιση, από τον χώρο εισόδου σε αυτόν της εξόδου, με έναν αριθμό κανόνων if-then, καθένας από των οποίων περιγράφει μια τοπική συμπεριφορά της αντιστοίχισης. Ειδικότερα, το μέρος των προϋποθέσεων ενός κανόνα προσδιορίζει μια ασαφή περιοχή στο χώρο της εισόδου, ενώ αυτός των συμπερασμάτων προσδιορίζει την έξοδο σε μια ασαφή περιοχή.



<u>Σχήμα 5.2:</u> *Δομή ασαφούς συστήματος εξαγωγής συμπερασμάτων.*

Θα εξετάσουμε τώρα πως λειτουργεί ένα FIS τύπου Mamdani μέσα από το πρόβλημα του φιλοδωρήματος που αναφέραμε παραπάνω. Το σύστημα έχει δύο εισόδους και τρεις κανόνες. Οι είσοδοι του συστήματος είναι η ποιότητα του φαγητού και η εξυπηρέτηση. Θεωρούμε ότι ο πελάτης βαθμολογεί την ποιότητα του φαγητού και την εξυπηρέτηση με μια κλίμακα από το 0 έως το 10. η έξοδος του συστήματος είναι το φιλοδώρημα που θα δώσει ο πελάτης και το οποίο κυμαίνεται από 5% έως 25% της τιμής του λογαριασμού.

Η βασική του δομή του συστήματος φαίνεται στο παρακάτω διάγραμμα:



<u>Σχήμα 5.3:</u> *Σύστημα 2 εισόδων, 3 κανόνων, 1 εξόδου.*

Η πληροφορία οδεύει από τα αριστερά προς τα δεξιά και καταλήγει σε μία μοναδική έξοδο. Ένα σημαντικό χαρακτηριστικό των συστημάτων ασαφούς λογικής είναι η παράλληλη εφαρμογή των κανόνων. Στα κλασσικά συστήματα δυαδικής λογικής έχουμε χρήση διακοπτών που ανοίγουν και κλείνουν ανάλογα με τις τιμές των εισόδων. Στα ασαφή συστήματα μεταβαίνουμε ομαλά από περιοχή σε περιοχή ανάλογα με τις συνέπειες των κανόνων. Η διαδικασία της απεικόνισης των εισόδων στις εξόδου μέσα από ένα σύστημα ασαφούς λογικής ακολουθεί πέντε βήματα:

1. Ασαφοποίηση των εισόδων.

2. Εφαρμογή των ασαφών λογικών τελεστών στις προϋπόθεσης των κανόνων.

3. Καθορισμός των συνεπαγωγών των κανόνων.

4. Συγκέντρωση όλων των συνεπαγωγών όλων των κανόνων.

5. Αποσαφήνιση .

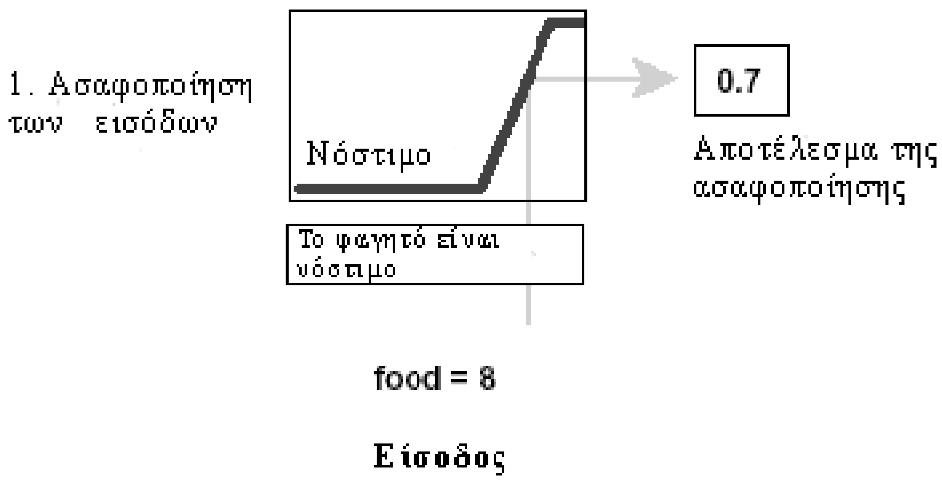
Θα εξετάσουμε κάθε βήμα από τα παραπάνω ξεχωριστά.

ΒΗΜΑ 1^{O} : Ασαφοποίηση εισόδων

Το πρώτο βήμα έχει να κάνει με την ασαφοποίηση των εισόδων δηλαδή αυτό που κάνουμε είναι να πάρουμε τις εισόδους και να καθορίσουμε σε ποιο βαθμό κάθε είσοδος ανήκει σε ένα ασαφές σύνολο μέσα από τις συναρτήσεις συμμετοχής. Οι είσοδοι είναι αριθμητικές τιμές μέσα στα όρια του χώρου αναφοράς (στην περίπτωση του προβλήματος που εξετάζουμε από 0 μέχρι 10) και οι έξοδοι είναι βαθμοί συμμετοχής στο προσδιορισμένο ασαφές σύνολο.

Η ασαφοποίηση των εισόδων μπορεί να γίνει είτε με χρήση πίνακα τιμών είτε μέσα από μια συνάρτηση. Το παράδειγμα μας είναι δομημένο σε τρεις κανόνες κάθε ένας από τους οποίους προϋποθέτει την κατάταξη των εισόδων σε διαφορετικές λεκτικές μεταβλητές service is poor, service is good, food is rancid, food is delicious.

Πριν να εφαρμοσθούν οι κανόνες θα πρέπει να ασαφοποιηθούν οι είσοδοι σύμφωνα με τις λεκτικές μεταβλητές. Για παράδειγμα σε ποιο βαθμό το φαγητό είναι νόστιμο. Αν υποθέσουμε ότι το φαγητό έχει αξιολογηθεί με μια συγκεκριμένη βαθμολογία από το 0 έως το 10 (π.χ. 7) τότε πως μπορούμε να συμπεράνουμε σε ποιο βαθμό είναι νόστιμο. Για το λόγο αυτό θα χρησιμοποιήσουμε μια γραφική απεικόνιση που ουσιαστικά αποτελεί τη συνάρτηση συμμετοχής του ασαφούς συνόλου νόστιμο.

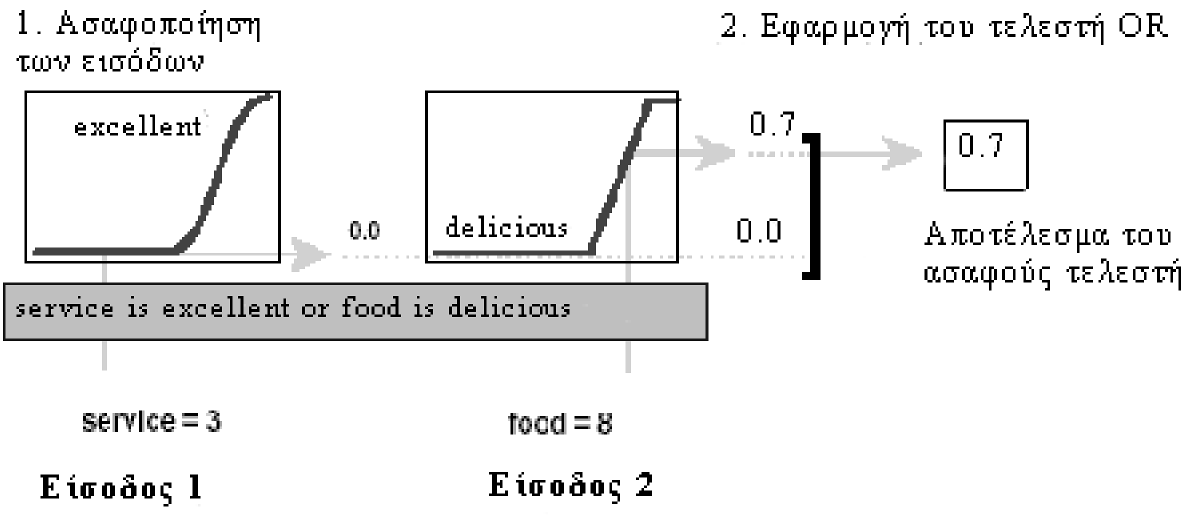


<u>Σχήμα 5.4:</u> *Συνάρτηση συμμετοχής του ασαφούς συνόλου νόστιμο.*

<u>ΒΗΜΑ 2ο: Εφαρμογή των ασαφών τελεστών</u>

Αφού έχουμε ασαφοποιήσει στις εισόδους μπορούμε να γνωρίζουμε σε ποιο βαθμό ικανοποιείται κάθε μέρος των προϋποθέσεων των if-then κανόνων. Αν ένας κανόνας if-then έχει στο σκέλος if περισσότερα του ενός μέρη τότε πρέπει να εφαρμοσθούν οι ασαφείς τελεστές ώστε να καθορισθεί ένα νούμερο που αντιπροσωπεύει το αποτέλεσμα της προϋπόθεσης του κανόνα. Το νούμερο αυτό θα εφαρμοσθεί έπειτα στη συνάρτηση εξόδου. Οι είσοδοι ενός λογικού τελεστή είναι δύο ή περισσότεροι βαθμοί συμμετοχής που έχουν προκύψει από τις ασαφοποιημένες μεταβλητές εισόδου. Η έξοδος του τελεστή είναι απλά ένας βαθμός αλήθειας.

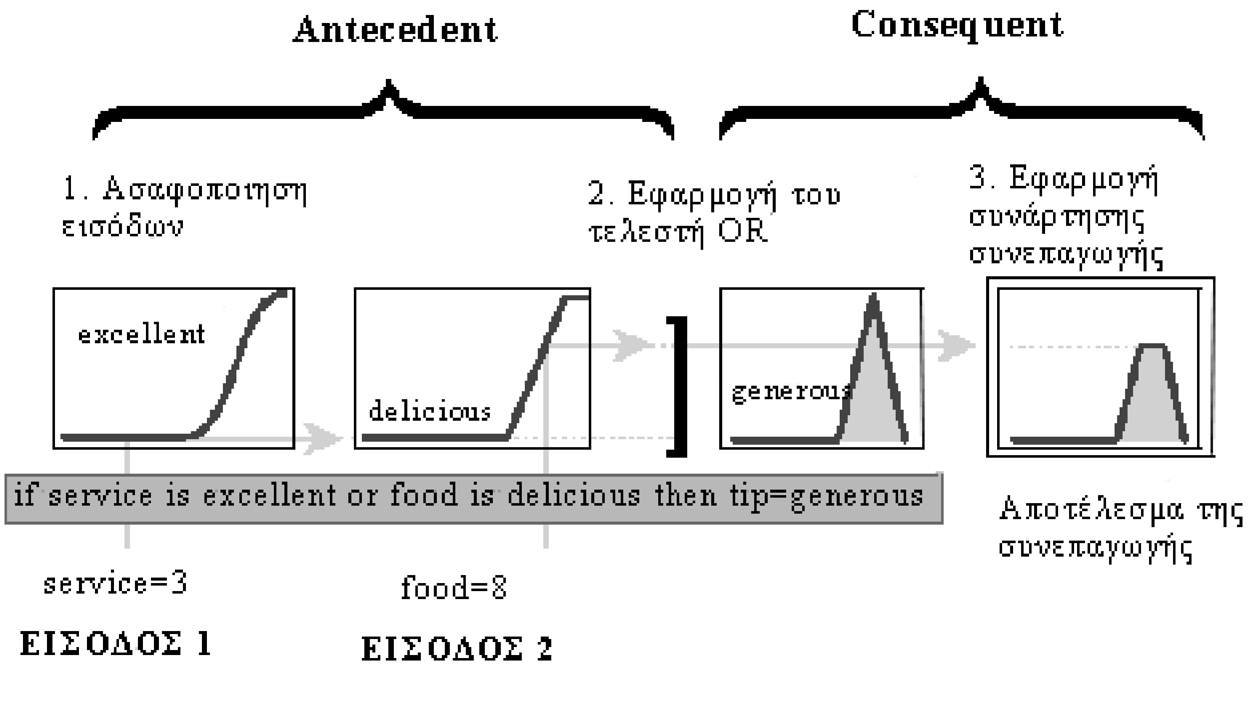
Όπως είδαμε κατά την αναφορά στους ασαφείς τελεστές υπάρχει ένας μεγάλος αριθμός συναρτήσεων που μπορούν να χρησιμοποιηθούν για να ορίσουν τους τελεστές AND και OR. Συνήθως χρησιμοποιούνται ο τελεστής max για το OR και min για το AND. Στο παρακάτω διάγραμμα βλέπουμε πως εφαρμόζονται οι τελεστές.



<u>Σχήμα 5.5:</u> *Παράδειγμα εφαρμογής τελεστών.*

ΒΗΜΑ 3ο : Μέθοδος εφαρμογής της συνεπαγωγής (Apply Implication Method)

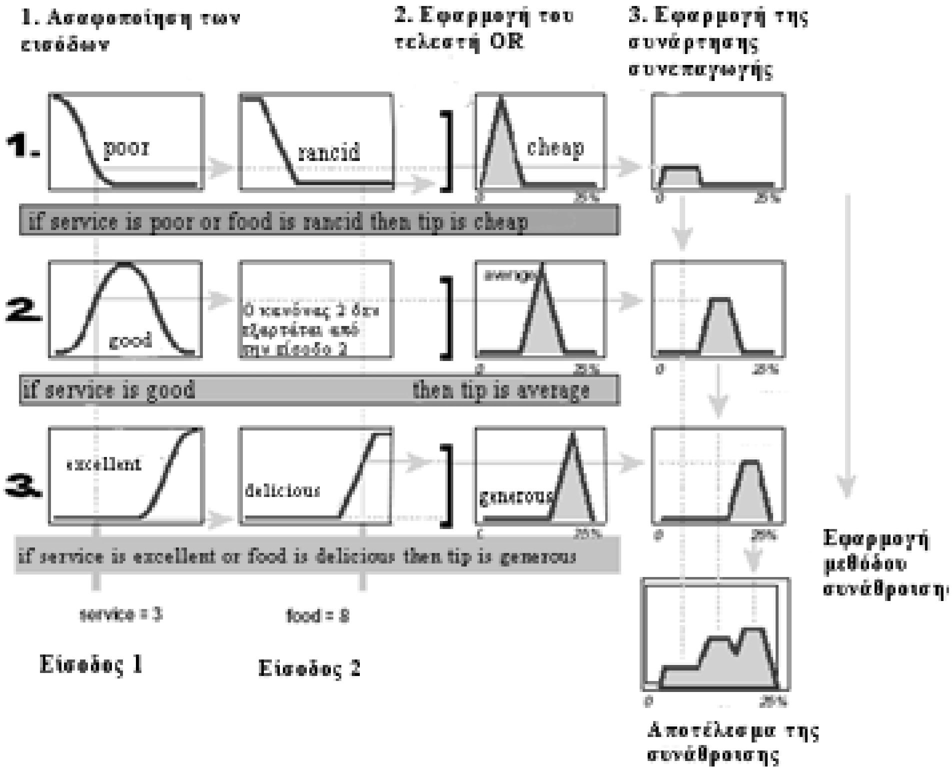
Για να εφαρμόσουμε την συνάρτηση συνεπαγωγής πρέπει να γνωρίζουμε το βάρος του κάθε κανόνα. Οι κανόνες if-then έχουν βάρη τα οποία εφαρμόζονται στο νούμερο που δίνει το κομμάτι της υπόθεσης και που μπορούν να κυμαίνονται από 0 έως 1. Συνήθως τα βάρη αυτά είναι ίσα με 1, ωστόσο μπορούμε να δώσουμε σε αυτά μια οποιαδήποτε τιμή από 0 έως 1. Για παράδειγμα αν θέλουμε να δώσουμε μικρότερη έμφαση σε έναν κανόνα σε σχέση με κάποιον άλλο , αυτό μπορούμε να το κάνουμε μέσω των τιμών των βαρών τους. Αφού έχομε αναθέσει στους κανόνες τα βάρη μπορούμε να εφαρμόσουμε την συνάρτηση συνεπαγωγής. Το αποτέλεσμα ενός κανόνα if-then είναι ένα ασαφές σύνολο που αντιπροσωπεύεται από μια συνάρτηση συμμετοχής που σταθμίζει κατάλληλα τα λεκτικά χαρακτηριστικά που έχουν ανατεθεί σ' αυτόν. Το αποτέλεσμα αυτό αναδιαμορφώνεται με τη χρήση μιας συνάρτησης που είναι συνδεδεμένη με το μέρος της υπόθεσης. Η είσοδος για τη διαδικασία συνεπαγωγής είναι ένας αριθμός που δίδεται από την υπόθεση και η έξοδος είναι ένα ασαφές σύνολο. Αυτή η διαδικασία εφαρμόζεται για κάθε κανόνα



Σχήμα 5.6: *Παράδειγμα εφαρμογής της συνεπαγωγής.*

<u>ΒΗΜΑ 4ο : Συγκέντρωση όλων των εξόδων</u>

Σε ένα FIS οι αποφάσεις βασίζονται στον έλεγχο όλων των κανόνων. Επομένως όλοι οι κανόνες πρέπει να συνδυαστούν προκειμένου να καταλήξουμε σε μια απόφαση. Η συγκέντρωση των κανόνων είναι μια διαδικασία κατά την οποία όλα τα ασαφή σύνολα που προκύπτουν ως έξοδοι των κανόνων συνδυάζονται ώστε να σχηματίσουν ένα μοναδικό ασαφές σύνολο. Η συγκέντρωση είναι η διαδικασία που προηγείται της αποσαφήνισης και δέχεται ως είσοδο τις αποκομμένες συναρτήσεις εξόδου του προηγούμενου βήματος. Η διαδικασία της συγκέντρωσης των εξόδων είναι αντιμεταθετική και έτσι η σειρά με την οποία εκτελούνται οι κανόνες δεν έχει σημασία στην έκβαση του αποτελέσματος. Στο παρακάτω διάγραμμα βλέπουμε τους τρεις κανόνες του προβλήματος και τον τρόπο που συνδυάζονται τα αποτελέσματά τους σε ένα ασαφές σύνολο.



<u>Σχήμα 5.7:</u> *Συνδυασμός αποτελεσμάτων των 3 κανόνων σε ένα ασαφές σύνολο.*

ΒΗΜΑ 5^{o} :Αποσαφήνιση

Η είσοδος για τη διαδικασία της αποσαφήνισης είναι το ασαφές σύνολο που προέκυψε από το προηγούμενο βήμα. Η έξοδος της αποσαφήνισης είναι ένας αριθμός. Η αποσαφήνιση είναι αναγκαία γιατί επιθυμούμε να έχομε ως αποτέλεσμα έναν αριθμό που θα μας επιτρέψει να πάρουμε συγκεκριμένες αποφάσεις. Υπάρχουν πολλοί τρόποι για να γίνει η αποσαφήνιση. Ένας ιδιαίτερα δημοφιλής είναι ο υπολογισμός του κέντρου βάρους της περιοχής που περικλείεται από την καμπύλη της συνάρτησης συμμετοχής του ασαφούς συνόλου. Άλλες επίσης δημοφιλείς μέθοδοι είναι η μέση τιμή , η μέγιστη η μέγιστη θέση του μεγίστου και η ελάχιστη θέση του μεγίστου.

Μέθοδος αποσαφήνισης MAXIMUM

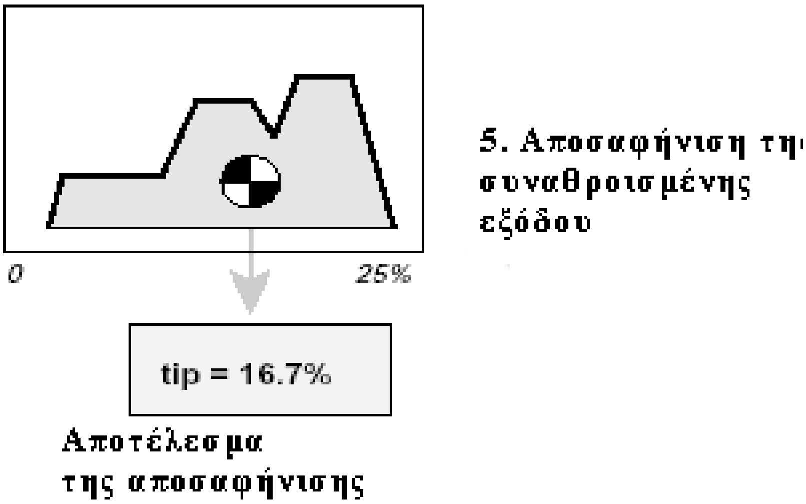
Σύμφωνα με τη μέθοδο αποσαφήνισης maximum , η διακριτή τιμή είναι αυτή που αντιστοιχεί στη μέγιστη τιμή συγγένειας του τελικού αποτελέσματος. Αν υπάρχουν περισσότερες από μία τέτοιες τιμές , τότε λαμβάνεται ανάλογα με την περίπτωση είτε ο μέσος όρος τους (average-of-maximum) ή η μέγιστη τιμή τους (maximum-of-maximum) ή η ελάχιστη τιμή τους (minimum-of-maximum).

Μέθοδος Cendroid

Σύμφωνα με τη μέθοδο αποσαφήνισης cendroid , η διακριτή τιμή είναι αυτή που προκύπτει από το κέντρο βάρους της τελικής συνάρτησης συμμετοχής για την ασαφή παράμετρο εξόδου. Το κέντρο βάρους μιας επιφάνειας που ορίζεται από μια συνάρτηση f(t) και τους καρτεσιανούς άξονες , βρίσκεται στη θέση t που ορίζεται από τη σχέση:

$$t_{\kappa\beta} = \frac{\int t \cdot f(t)dt}{\int f(t)dt}$$

Στο παρακάτω σχήμα βλέπουμε πως εφαρμόζεται η μέθοδος της αποσαφήνισης για το παράδειγμά μας.



Σχήμα 5.8: *Εφαρμογή της μεθόδου αποσαφήνισης.*

Στην περίπτωση διακριτού συνόλου αναφοράς , τα ολοκληρώματα στην παραπάνω σχέση αντικαθίστανται με διακριτό άθροισμα και γίνεται δειγματοληψία Ν σημείων στο σύνολο αναφοράς.

Ένα χαρακτηριστικό της μεθόδου αποσαφήνισης cendroid είναι ότι στην περίπτωση που έχει γίνει σύνθεση αποτελεσμάτων από επιμέρους κανόνες και υπάρχουν τυχόν αλληλεπικαλυπτόμενες περιοχές , αυτές λαμβάνονται υπόψη μία μόνο φορά. Επίσης στην περίπτωση που η συνάρτηση συμμετοχής είναι παντού 0 , το αποτέλεσμα της αποσαφήνισης ορίζεται κατά σύμβαση.

Διαγραμματική μέθοδος επίλυσης

Μία εναλλακτική προσέγγιση στην επίλυση προβλημάτων ασαφούς συλλογιστικής είναι η διαγραμματική επίλυση. Αυτή επιτρέπει τη γρήγορη, αλλά προσεγγιστική εκτίμηση της τελικής λύσης χωρίς να απαιτεί όλους τους αριθμητικούς υπολογισμούς που παρουσιάστηκαν παραπάνω. Βασική προϋπόθεση για την εφαρμογή της μεθόδου είναι οι συναρτήσεις συμμετοχής των παραμέτρων του προβλήματος να είναι συνεχείς καμπύλες και όχι σύνολο ζευγών.

5.3.2 Συστήματα ασαφούς συλλογιστικής τύπου Sugeno

Το ασαφές μοντέλο Sugeno δημιουργήθηκε σε μια προσπάθεια να αναπτυχθεί μια συστηματική προσέγγιση στη γέννηση ασαφών κανόνων από ένα δεδομένο σύνολο δεδομένων εισόδου-εξόδου. Σε πολλές περιπτώσεις είναι περισσότερο αποτελεσματική η χρήση σημείων-αιχμών σαν συναρτήσεις συμμετοχής εξόδου, από ένα συνεχές κατανεμημένο ασαφές σύνολο. Μια τέτοια αιχμή είναι γνωστή με τον αγγλικό όρο συνάρτηση συμμετοχής εξόδου singleton και θα μπορούσε να χαρακτηριστεί σαν προαποσαφοποιημένο ασαφές σύνολο. Η συνάρτηση συμμετοχής ενός singleton είναι μA(x)=1.

Με αυτόν τον τρόπο εμπλουτίζουμε την επάρκεια της διαδικασίας αποσαφοποίησης, διότι απλοποιούμε ιδιαιτέρως τους απαιτούμενους υπολογισμούς σε σχέση με την πιο γενική μέθοδο Mamdani, στην οποία βρίσκουμε το κέντρο μάζας μια δισδιάστατης συνάρτησης. Αντί λοιπόν να βρίσκουμε το κέντρο μάζας, χρησιμοποιούμε τον σταθμικό μέσο μερικών σημείων. Τα συστήματα τύπου Sugeno υποστηρίζουν αυτόν τον τύπο μοντέλου.

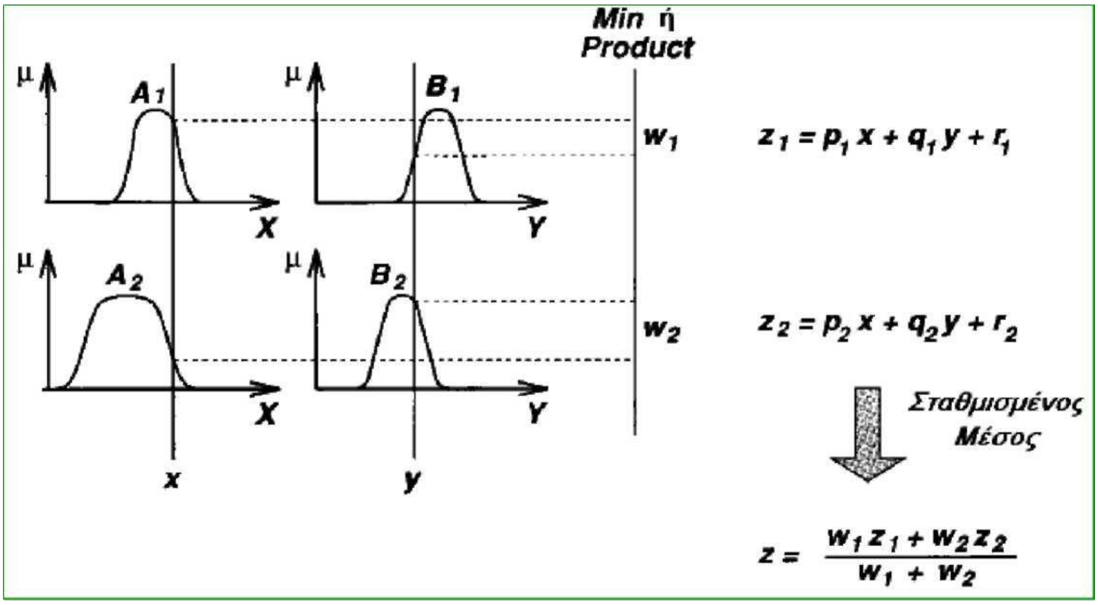
Η μέθοδος Sugeno εισήχθη το 1985 και έχει ομοιότητες με τη μέθοδο των συστημάτων Mamdani. Για παράδειγμα τα 2 πρώτα βήματά της (ασαφοποίηση των εισόδων και εφαρμογή των τελεστών) είναι ακριβώς τα ίδια. Η κύρια διαφορά ανάμεσα στα δύο συστήματα έγκειται στο ότι οι συναρτήσεις συμμετοχής στην έξοδο των συστημάτων Sugeno είναι μόνο γραμμικές ή σταθερές. Ένας κανόνας σε ασαφές μοντέλο Sugeno έχει τη μορφή:

***If x is A and y is B then z* = f(x, y).**

Όπου: A, B είναι τα ασαφή σύνολα των προϋποθέσεων και το z=f(x,y) είναι μια σαφής συνάρτηση στο χώρο των συμπερασμάτων. Συνήθως το f(x,y) είναι ένα πολυώνυμο των μεταβλητών εισόδου x, y, αλλά μπορεί να είναι μια οποιαδήποτε συνάρτηση, αρκεί να μπορεί να περιγράψει κατάλληλα την έξοδο του μοντέλου, μέσα στην ασαφή περιοχή που ορίστηκε από τις προϋποθέσεις των κανόνων. Όταν το f(x,y) είναι πολυώνυμο πρώτου βαθμού το FIS ονομάζεται ασαφές μοντέλο πρώτου βαθμού. Για παράδειγμα:

If input-1= x and input-2= y, then Output is z=ax+by+c.

Για ένα Sugeno μηδενικού βαθμού το επίπεδο εξόδου z είναι μια σταθερά (a=b=0) και μπορεί να θεωρηθεί και σαν μια ειδική περίπτωση Mamdani, στο οποίο το συμπέρασμα του κάθε κανόνα προσδιορίζεται από ένα ασαφές singleton (προαποσαφοποιημένο συμπέρασμα). Η έξοδος ενός Sugeno μηδενικού βαθμού είναι μια ομαλή συνάρτηση των μεταβλητών εισόδων του, εφόσον οι γειτονικές συναρτήσεις συμμετοχής στις προϋποθέσεις έχουν αρκετή επικάλυψη.



Σχήμα 5.9: *Ασαφές μοντέλο Sugeno πρώτου βαθμού.*

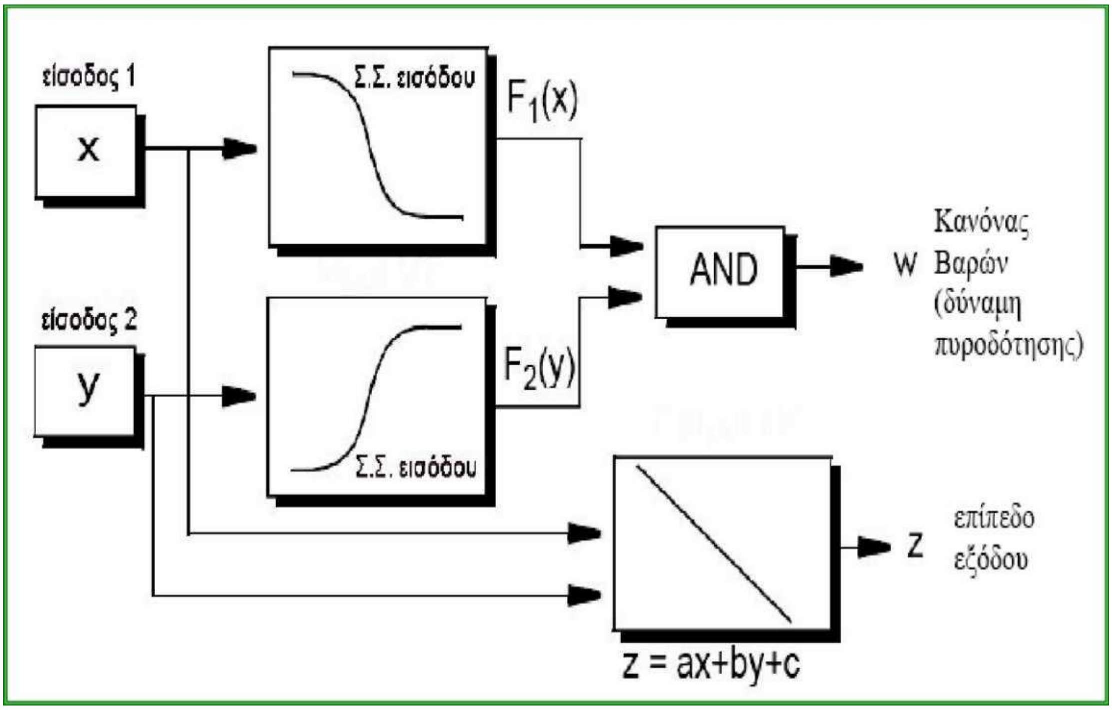
Στο παραπάνω σχήμα βλέπουμε ότι το επίπεδο εξόδου Zi για κάθε κανόνα είναι σταθμισμένο με βάρη η αλλιώς με μια **δύναμη πυροδότησης** (βαθμός ενεργοποίησης) Wi για κάθε κανόνα. Για παράδειγμα για τον τελεστή AND, με είσοδο-1= x και είσοδο-2= y, η δύναμη πυροδότησης είναι:

$$W_i = AND\,(F1(x), F2(x))$$

Όπου: F1 και F2 είναι οι συναρτήσεις συμμετοχής για τις αντίστοιχες εισόδους. Η τελική έξοδος του συστήματος είναι ο σταθμισμένος μέσος όλων των εξόδων και δίνεται από τον τύπο:

$$\text{Τελική έξοδος} = \frac{\sum_{i=1}^{N} w_i z_i}{\sum_{i=1}^{N} w_i}$$

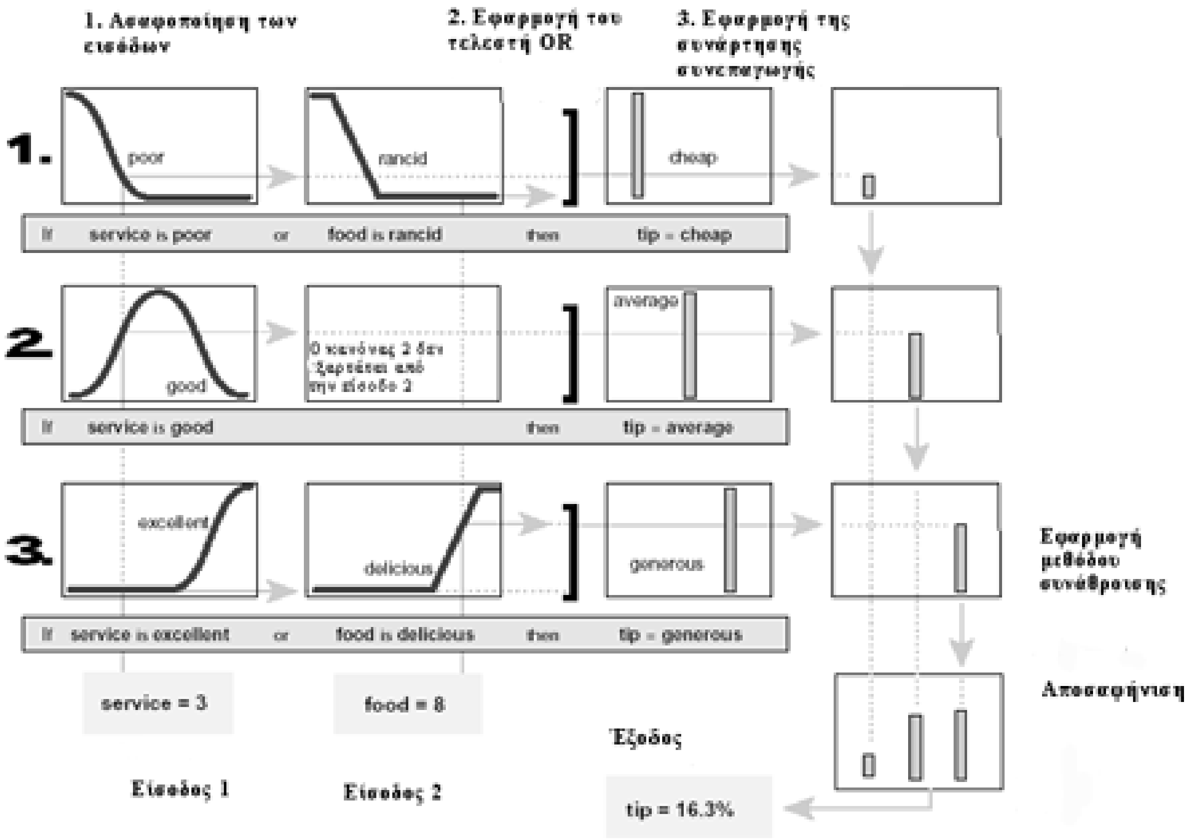
Στο παρακάτω σχήμα φαίνεται η λειτουργεί του Sugeno για το παράδειγμα που αναφέραμε.



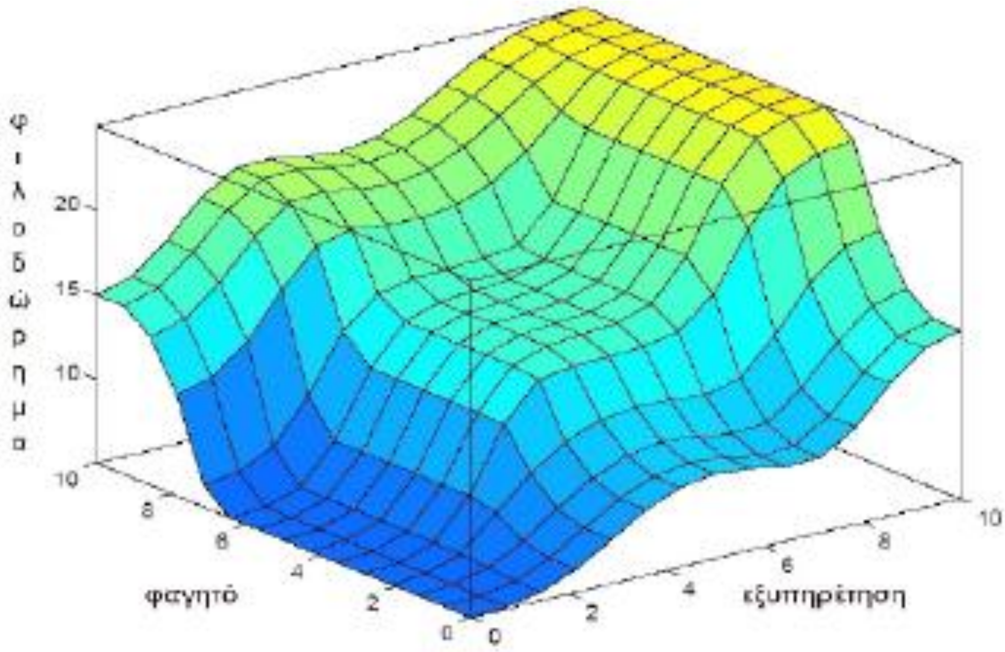
Σχήμα 5.10: *Ασαφές μοντέλο Sugeno, δύναμη πυροδότησης.*

Στα σχήματα που ακολουθούν παρακάτω φαίνεται σε πλήρη ανάπτυξη το FIS Sugeno για το πρόβλημα με το ποσοστό του φιλοδωρήματος, όπου η χρήση των συναρτήσεων singleton είναι απολύτως επαρκής.

Ο πιο εύκολος τρόπος να κατανοήσουμε τα ασαφή συστήματα Sugeno πρώτου βαθμού, είναι να σκεφτούμε πως ο κάθε κανόνας προσδιορίζει τη θέση ενός 'κινούμενου singleton'. Οι αιχμές δηλαδή των singleton της εξόδου μπορούν να κινηθούν με γραμμικό τρόπο στο χώρο εξόδου, ανάλογα με τις εισόδους.



Σχήμα 5.11: *Ασαφές μοντέλο Sugeno, ποσοστό φιλοδωρήματος.*



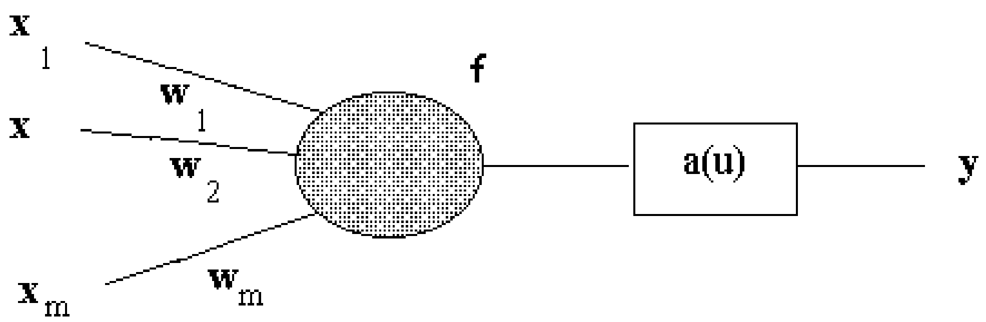
Σχήμα 5.12: *Επιφάνεια λύσης του προβλήματος για το φιλοδώρημα.*

5.4 Νεύρο-ασαφή συστήματα

Στην ενότητα αυτή εξετάζουμε τα νευροασαφή συστήματα. Η δομική μονάδα των συστημάτων αυτών είναι οι ασαφείς νευρώνες.

ΑΣΑΦΕΙΣ ΝΕΥΡΩΝΕΣ

Οι ασαφείς νευρώνες έχουν τη μορφή που φαίνεται στο παρακάτω σχήμα:



Σχήμα 5.13: *Ασαφείς Νευρώνες.*

Όπου :

$x_1, x_2, .., x_m$ οι είσοδοι του νευρώνα.
$w_1, w_2, .., w_m$ τα βάρη των συνάψεων.
f η συνάρτηση συμμετοχής του νευρώνα.
a η συνάρτηση ενεργοποίηση του νευρώνα.
y η έξοδος του νευρώνα.

Οι ασαφείς νευρώνες διακρίνονται σε νευρώνες συμμετοχής και σε λειτουργικούς νευρώνες. Παρακάτω γίνεται αναφορά στα χαρακτηριστικά κάθε κατηγορίας ασαφών νευρώνων.

ΝΕΥΡΩΝΕΣ ΣΥΜΜΕΤΟΧΗΣ

Οι συνηθέστερες μορφές συναρτήσεων συμμετοχής είναι η τριγωνική, η γκαουσιανή, η τραπεζοειδής , η κανονική κοκ. Οι παραπάνω συναρτήσεις συμμετοχής μπορούν να υλοποιηθούν με τη χρήση ενός νευρώνα. Η συνάρτηση συμμετοχής μπορεί να υλοποιηθεί θεωρώντας ότι ο νευρώνας δέχεται ως είσοδο το x και δίνει έξοδο A(x) απαιτώντας η συνάρτηση συμμετοχής του να έχει τη μορφή A. Αν δηλαδή υποθέσουμε ότι η συνάρτηση συμμετοχής είναι η κανονική , τότε για να υλοποιηθεί πρέπεί η συνάρτηση ενεργοποίησης να είναι:

$$a(u) = \exp\left(-\frac{(u-m)^2}{\sigma^2} \right)$$

Όπου m κέντρο της συνάρτησης συμμετοχής και σ το εύρος της.

ΛΕΙΤΟΥΡΓΙΚΟΙ ΝΕΥΡΩΝΕΣ

Με τον ίδιο τρόπο μπορούν να υλοποιηθούν οι βασικές πράξεις των ασαφών συνόλων. Οι πράξεις αυτές στηρίζονται σε λειτουργίες όπως η τ-νόρμα η σ-νόρμα , το ασαφές συμπλήρωμα κλπ. Οι παραπάνω λειτουργίες μπορούν να αντικαταστήσουν τη συνάρτηση ενεργοποίησης του νευρώνα αν θεωρήσουμε ότι οι είσοδοι δεν αθροίζονται πολλαπλασιασμένοι με βάρη όπως στους κλασσικούς νευρώνες. Για να μπορούμε όμως να εφαρμόσουμε τους αλγορίθμους μάθησης (πχ τον backpropagation αλγόριθμος αντίστροφης διάδοσης) θα πρέπει οι λειτουργίες να είναι παραγωγίσιμες. Για την υλοποίηση των ασαφών λειτουργιών μπορούμε να επεκτείνουμε την έννοια της συνάρτησης μεταφοράς του νευρώνα. Στην περίπτωση αυτή ο νευρώνας δεν υλοποιεί το αναλυτικό γινόμενο της εισόδου με το διάνυσμα των βαρών αλλά την πράξη της συγκεκριμένης ασαφούς λειτουργίας θεωρώντας ότι όλα τα βάρη είναι 1. ως συνάρτηση συμμετοχής λαμβάνεται η ακόλουθη:

$$a(u) = \begin{cases} 1, u > 1 \\ u, 0 \le u \le 1 \\ 0, u < 0 \end{cases}$$

ΣΥΝΘΕΤΙΚΟΙ ΝΕΥΡΩΝΕΣ

Παραπάνω αναφερθήκαμε στο στη σύνθεση των ασαφών σχέσεων. Η πράξη της σύνθεσης αποτελεί μια γενίκευση της πράξης του αναλυτικού γινομένου των διανυσμάτων και υλοποιείται από τη συνάρτηση μεταφοράς f των νευρώνων. Η μορφή του νευρώνα είναι η ίδια και ορίζεται από τις εξισώσεις:

$$f = \underset{i \in N_m}{union}\, t(x_t, w_t)$$

$$y = a(u)$$

Όπου: union είναι μια σ-νόρμα και t είναι μια τ-νόρμα. Ως συνάρτηση ενεργοποίησης θεωρείται η συνάρτηση αναρρίχησης. Παρατηρούμε ότι η ο συνθετικός νευρώνας είναι μια γενίκευση του κλασσικού νευρώνα, αφού η πράξη της φραγμένης άθροισης, η οποία εξασφαλίζεται από τη συνάρτηση ενεργοποίησης είναι μια ειδική περίπτωση της φραγμένης ένωσης, ενώ το γινόμενο είναι μια ειδική περίπτωση της ασαφούς τομής.

5.4.1 Περιγραφή του ANFIS (Adaptive Neuro-Fuzzy Inference System)

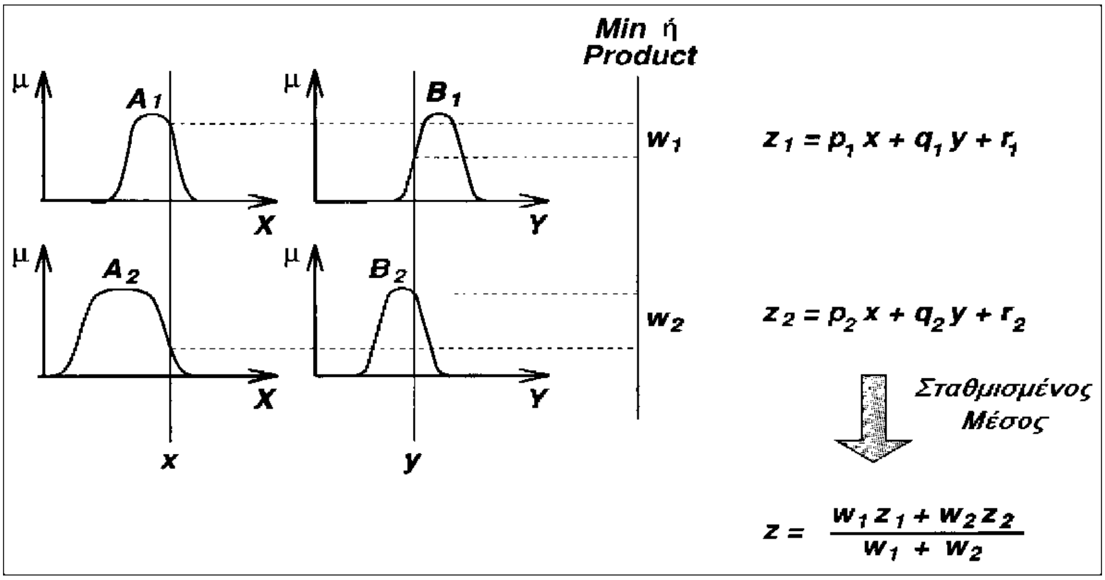
Το σύστημα ANFIS, Adaptive Neuro-Fuzzy Inference System είναι ένα από τα πρώτα που προτάθηκαν και εμπνευστής του είναι ο J.-S. Jang. Κύριο χαρακτηριστικό του είναι ότι χρησιμοποιεί ασαφή συλογιστική Takagi – Sugeno. Παραμένει ακόμα και σήμερα ένα από τα πλέον αποδοτικά και αξιόπιστα νευρο-ασαφή συστήματα και βρίσκει εφαρμογές σε πληθώρα περιοχών. Παρακάτω περιγράφεται το δίκτυο, υποθέτοντας ότι το πρόβλημα του οποίου προσπαθούμε να προσεγγίσουμε τη λύση, έχει δύο εισόδους (x,y) και μία έξοδο (z).

5.4.1.1 Αρχιτεκτονική του ANFIS

Υποθέτοντας ότι για ένα πρώτης τάξης μοντέλο Sugeno, μία τυπική βάση κανόνων (rule base) θα μπορούσε να είναι και η εξής:

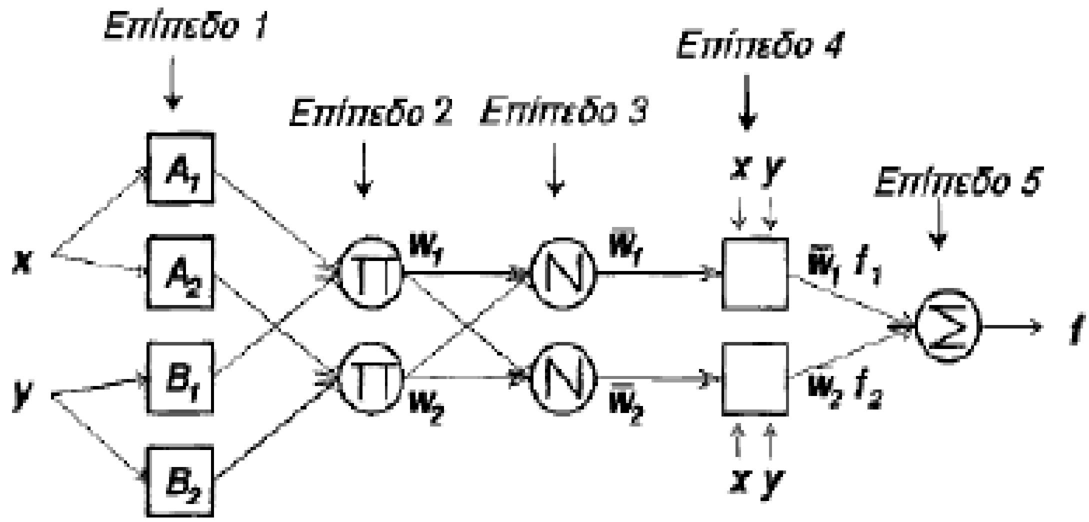
Κανόνας 1: **Εάν** x **είναι** A_1 **και** y **είναι** B_1 **τότε** $f_1 = p_1 \times x + q_1 \times y + r_1$
Κανόνας 2: **Εάν** x **είναι** A_2 **και** y **είναι** B_2 **τότε** $f_2 = p_2 \times x + q_2 \times y + r_2$

Στο παρακάτω σχήμα (5.14) δείχνει με απλό τρόπο τη διαδικασία συμπερασμού (inference procedure) του μοντέλου Sugeno, στην περίπτωση του t-operator έχει επιλεγεί η τομή των ασαφών συνόλων (A,B), οπότε $\mu_{A \cap B}(x) = \min [\mu_A(x), \mu_B(x)]$.



Σχήμα 5.14: *Ασαφής Συλλογιστική (fuzzy reasoning).*

Όπως γίνεται φανερό, η έξοδος z του πρωτοβάθμιου μοντέλου Sugeno είναι ένας σταθμικός μέσος όρος. Η αντίστοιχη αναπαράσταση του δικτύου ANFIS παρουσιάζεται στο παρακάτω σχήμα (5.15).



Σχήμα 5.15: *Αρχιτεκτονική του ANFIS.*

Στο παραπάνω σχήμα απεικονίζεται ο συλλογιστικός μηχανισμός (reasoning) για αυτό το μοντέλο Sugeno και η αντίστοιχη αρχιτεκτονική του ANFIS. Παρατηρούμε ότι οι κόμβοι, ίδιου επιπέδου, έχουν παρόμοιες συναρτήσεις. Παρακάτω παρουσιάζεται αναλυτικότερα η διεργασία που εκτελείται σε κάθε επίπεδο.

Επίπεδο 1: Κάθε κόμβος i σε αυτό το επίπεδο είναι ένας προσαρμόσιμος (adaptive) κόμβος με συνάρτηση κόμβου:

$$O_{1,i} = \mu_{A_i}(x) \quad \text{για i = 1,2 ή}$$

$$O_{1,i} = \mu_{B_{i-2}}(y) \quad \text{για i = 3,4 ή}$$

όπου x (ή y) η είσοδος στον κόμβο i, A_i (ή B_{i-2}) η γλωσσική μεταβλητή (small, large, κλπ.) που σχετίζεται με αυτή τη συνάρτηση του κόμβου. Με άλλα λόγια, το $O_{1,i}$ είναι ο βαθμός συμμετοχής του A (= A_1, A_2, B_1, B_2) και καθορίζει το βαθμό στον οποίο η είσοδος x (ή y) ικανοποιεί τον ποσοτικοποιητή A. Εδώ η συνάρτηση συμμετοχής για το Α μπορεί να είναι οποιαδήποτε κατάλληλη παραμετρική συνάρτηση συμμετοχής .

Καθώς οι τιμές των παραμέτρων αλλάζουν, οι συναρτήσεις διαμορφώνονται ανάλογα, παρουσιάζοντας έτσι διάφορες μορφές της αρχικής συνάρτησης συμμετοχής για το ασαφές σύνολο A. Οι παράμετροι σε αυτό το επίπεδο αναφέρονται ως αρχικοί παράμετροι (premise parameters).

Επίπεδο 2: Κάθε κόμβος σε αυτό το επίπεδο είναι ένας σταθερός (fixed) κόμβος Π, του οποίου η έξοδος είναι το γινόμενο όλων των εισερχόμενων σημάτων:

$$O_{2,i} = w_i = \mu_{A_1}(x) * \mu_{B_i}(y), \quad \text{για i = 1,2}$$

$$O_{2,i} = w_i = \mu_{A_1}(x) * \mu_{B_{i-2}}(y), \quad \text{για i = 3,4}$$

Επίπεδο 3: Κάθε κόμβος σε αυτό το επίπεδο είναι ένας σταθερός κόμβος Ν. Ο *i-ιστός* κόμβος υπολογίζει τη συμμετοχή του βαθμού ενεργοποίησης (firing strength) του *i-οστού* κανόνα στο άθροισμα των βαθμών ενεργοποίησης όλων των κανόνων:

$$O_{3,i} = \overline{w}_i = \frac{w_i}{w_1 + w_2 + w_3 + w_4}, \quad \text{i=1,2, …,4}$$

Οι έξοδοι αυτού του επιπέδου ονομάζονται κανονικοποιημένοι βαθμοί ενεργοποίησης (normalized firing strengths).

Επίπεδο 4: Κάθε κόμβος *i,* σε αυτό το επίπεδο είναι ένας προσαρμοστικός κόμβος με συνάρτηση κόμβου.

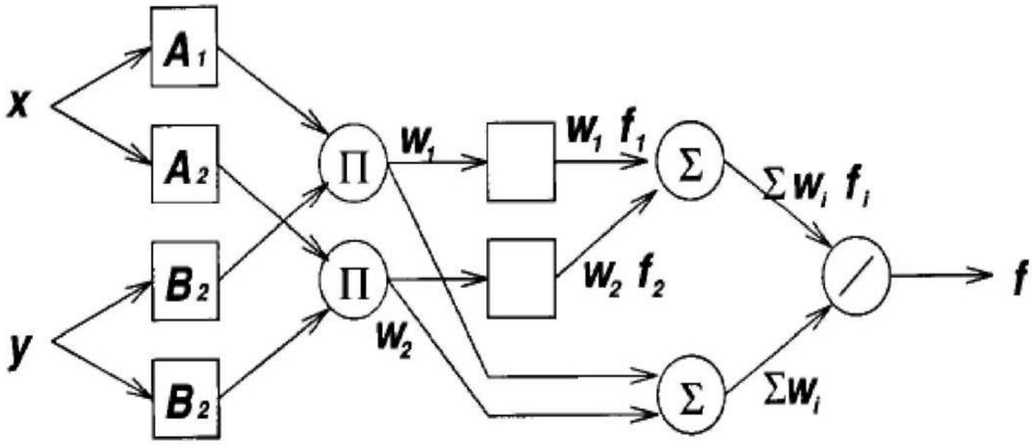
$$O_{4,i} = \overline{w}_i * f_i = \overline{w}_i * (p_i * x + q_i * y + s_i * z + r_i)$$

Όπου: w_i η έξοδος του επιπέδου 3, $\{p_i, q_i, s_i, r_i\}$ το σύνολο των παραμέτρων. Οι παράμετροι σε αυτό το επίπεδο αναφέρονται ως *επακόλουθοι* (consequent parameters).

Επίπεδο 5: Ο μοναδικός κόμβος σε αυτό το επίπεδο είναι ένας σταθερός κόμβος Σ που υπολογίζει τη συνολική έξοδο σαν το ολικό άθροισμα όλων των εισερχόμενων σημάτων:

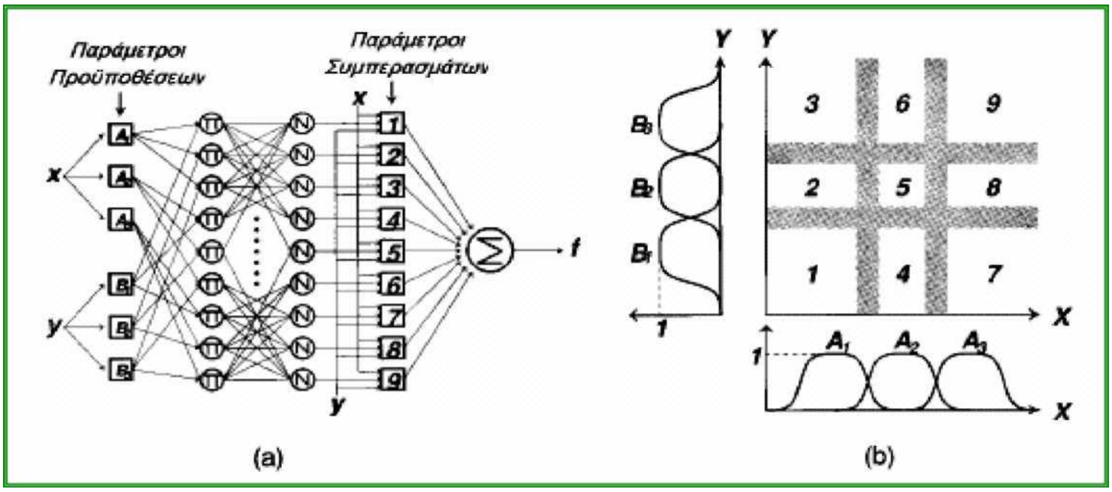
$$\text{overall output} = O_{5,i} = \sum_i \overline{w}_i * f_i = \frac{\sum_i w_i * f_i}{\sum_i w_i}$$

Αυτό το προσαρμοστικό δίκτυο είναι λειτουργικά ισοδύναμο με το ασαφές μοντέλο Sugeno. Μπορούμε να συνδυάσουμε τα επίπεδα 3 και 4 για να δημιουργήσουμε ένα ισοδύναμο δίκτυο με τέσσερα μόνο επίπεδα. Χρησιμοποιώντας το ίδιο δείγμα μπορούμε να πραγματοποιήσουμε την κανονικοποίηση των βαρών στο τελευταίο επίπεδο. Στο *σχήμα 5.16* απεικονίζεται ένα ANFIS τέτοιου τύπου. Σαν ακραία περίπτωση μπορούμε να συρρικνώσουμε όλο το δίκτυο, σε έναν μόνο προσαρμοστικό κόμβο με το ίδιο σύνολο παραμέτρων. Η ανάθεση συναρτήσεων στους κόμβους και η σύνθεση του δικτύου είναι αυθαίρετες. Κάθε κόμβος και κάθε επίπεδο πραγματοποιούν λειτουργίες που είναι σημαντικές και έχουν δυνατότητα να αποτελούνται από επιμέρους τμήματα.



Σχήμα 5.16: *Αρχιτεκτονική του ANFIS για ασαφές σύνολο Sugeno, όπου η κανονικοποίηση βαρών πραγματοποιείται στο τελευταίο επίπεδο.*

Στο ANFIS μπορεί επίσης να χρησιμοποιηθεί το μοντέλο Mamdami. Το Sugeno χρησιμοποιείται περισσότερο συχνά, μιας και διακρίνεται για την διαφάνεια και την αποτελεσματικότητα του.



<u>Σχήμα 5.17:</u> (α) Η αρχιτεκτονική του ANFIS για το ασαφές μοντέλο Sugeno με δύο εισόδους και εννέα κανόνες (β) ο χώρος εισόδου,

Στο σχήμα 5.17(α) βλέπουμε μια αρχιτεκτονική του ANFIS που είναι ισοδύναμη με ένα ασαφές μοντέλο Sugeno πρώτου βαθμού δυο εισόδων και εννέα κανόνων. Σε κάθε είσοδο θεωρούμε ότι αντιστοιχούν τρεις συναρτήσεις συμμετοχής. Στο σχήμα 5.17(β) φαίνεται πως είναι χωρισμένος ο χώρος εισόδου δύο διαστάσεων, σε εννέα υπέρθετες (overlapping) ασαφείς περιοχές, όπου κάθε μια ελέγχεται από ένα ασαφή κανόνα if-then. Αυτό σημαίνει ότι το μέρος των προϋποθέσεων ενός κανόνα προσδιορίζει μια ασαφή περιοχή, ενώ το μέρος των συμπερασμάτων προσδιορίζει την έξοδο μέσα στην περιοχή.

5.4.1.2 Εκτιμητής ελαχίστων τετραγώνων (Least-Squares Estimator).

Στο γενικό πρόβλημα ελαχίστων τετραγώνων, η έξοδος του γραμμικού μοντέλου ***y*** δίνεται από τη γραμμικά παραμετροποιημένη έκφραση:

$$y = \theta_1 f_1 (\mathbf{u}) + \theta_2 f_2 (\mathbf{u}) + \ldots + \theta_n f_n (\mathbf{u}) \qquad (5.1)$$

Όπου: $\mathbf{u} = [u_1, \ldots, u_p]^T$ είναι το διάνυσμα εισόδου του μοντέλου, $f_1, \ldots, f_n$ είναι γνωστές συναρτήσεις του **u** και $\theta_1, \ldots, \theta_n$ είναι άγνωστες παράμετροι που θα υπολογιστούν. Η παραπάνω εξίσωση (5.1) καλείται *συνάρτηση παλινδρόμησης,* και τα θ_i ονομάζονται *συντελεστές παλινδρόμησης.*

Για να προσδιοριστούν οι άγνωστες παράμετροι θ, συνήθως πρέπει να εκτελεστούν πειράματα για να βρεθεί ένα σύνολο δεδομένων εκπαίδευσης που αποτελείται από τα ζευγάρια δεδομένων $\{(u_i, y_i), i = 1,.., m\}$

τα οποία αντιπροσωπεύουν τα επιθυμητά ζευγάρια εισόδου-εξόδου του συστήματος στόχου που θα μοντελοποιηθεί. Η αντικατάσταση κάθε ζευγαριού στοιχείων στην (5.1)εξίσωση παράγει ένα σύνολο m γραμμικών εξισώσεων:

$$\begin{gathered} f_1(u_1)\theta_1 + f_2(u_1)\theta_2+\ldots+f_n(u_1)\theta_n=y_1 \\ f_1(u_2)\theta_1 + f_2(u_2)\theta_2+\ldots+f_n(u_2)\theta_n=y_2 \\ \ldots \\ f_1(u_m)\theta_1 + f_2(u_m)\theta_2+\ldots+f_n(u_m)\theta_n=y_m \end{gathered} \qquad (5.2)$$

Σε μορφή πινάκων, οι προηγούμενες εξισώσεις μπορούν να γραφτούν σε συνοπτική μορφή:

$$\mathbf{A\theta = y} \qquad (5.3)$$

Όπου A είναι ένας *m* χ *n* πίνακας *(πίνακας σχεδιασμού):*

$$\mathbf{A} = \begin{bmatrix} f_1(\mathbf{u}_1) & \cdots & f_n(\mathbf{u}_1) \\ \vdots & \vdots & \vdots \\ f_1(\mathbf{u}_m) & \cdots & f_n(\mathbf{u}_m) \end{bmatrix}$$

Το **θ** είναι ένα *n × 1* διάνυσμα άγνωστων παραμέτρων και y είναι ένα *m × 1* διάνυσμα εξόδου:

$$\mathbf{\theta} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Η *i-οστή* σειρά του ενωμένου πίνακα δεδομένων [A: y], που δηλώνεται με $[a_i^T : y_i]$, σχετίζεται με το *i-οστό* ζευγάρι δεδομένων $(u_i; y_i)$ μέσω της:

$$\mathbf{a}_i^T = [f_1(\mathbf{u}_i), \ldots, f_n(\mathbf{u}_i)]$$

Το μεγαλύτερο μέρος των υπολογισμών βασίζεται στους πίνακες A και y, μερικές φορές γίνεται αναφορά στο $(a_i^T; y_i)$ σαν το *i-οστό* ζευγάρι δεδομένων του συνόλου δεδομένων εκπαίδευσης. Για να προσδιοριστεί μεμονωμένα το άγνωστο διάνυσμα **θ**, είναι απαραίτητο να ισχύει $m \geq n$. Αν ο A είναι τετραγωνικός $(m = n)$ και αντιστρέψιμος, τότε η εξίσωση (5.3) μπορεί να λυθεί ως προς τον άγνωστο **x** και γίνεται:

$$\mathbf{\theta = A^{-1} y} \qquad (5.4)$$

Ο m είναι συνήθως μεγαλύτερος από τον n , που σημαίνει ότι υπάρχουν περισσότερα ζευγάρια στοιχείων από τις παραμέτρους. Σε αυτήν την περίπτωση, μια ακριβής λύση που να ικανοποιεί όλες τις m εξισώσεις δεν είναι πάντα δυνατή, δεδομένου ότι τα στοιχεία μπορεί να μολυνθούν από θόρυβο, ή το μοντέλο μπορεί να μην είναι κατάλληλο για την περιγραφή του συστήματος στόχου. Κατά συνέπεια η εξίσωση (5.3) πρέπει να τροποποιηθεί με την ενσωμάτωση ενός διανύσματος λάθους **e** για να αποτελέσει το τυχαίο λάθος θορύβου ή το τυχαίο λάθος διαμόρφωσης ως εξής:

$$\mathbf{A\theta} + \mathbf{e} = \mathbf{y} \tag{5.5}$$

Τώρα, αντί της εύρεσης της ακριβούς λύσης στην εξίσωση (5.9), πρέπει να βρεθεί το $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ που ελαχιστοποιεί το άθροισμα του τετραγωνικού σφάλματος και ορίζεται ως:

$$E(\theta) = \sum_{i=1}^{m} (y_i - a_i^r \theta)^2 = e^r e = (y - A\theta)^r (y - A\theta) \tag{5.6}$$

Όπου **e = y - Aθ** είναι το διάνυσμα λάθους που παράγεται από μια συγκεκριμένη επιλογή του **θ**. Πρέπει να σημειωθεί ότι το ***E(θ)*** είναι σε τετραγωνική μορφή και έχει ένα μοναδικό ελάχιστο $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. Το ακόλουθο θεώρημα δηλώνει έναν απαραίτητο όρο που ικανοποιείται από τον εκτιμητή ελαχίστων τετραγώνων $\hat{\boldsymbol{\theta}}$.

Εκτιμητής Ελαχίστων Τετραγώνων

Το τετραγωνικό σφάλμα στην εξίσωση (5.6) ελαχιστοποιείται όταν $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, ο οποίος καλείται εκτιμητής ελαχίστων τετραγώνων (LSE) και ο οποίος ικανοποιεί την κανονική εξίσωση:

$$A^r A \hat{\theta} = A^r y \tag{5.7}$$

Αν ο A^T A είναι αντιστρέψιμος, ο $\hat{\boldsymbol{\theta}}$ είναι μοναδικός και δίνεται από την:

$$\hat{\theta} = (A^T A)^{-1} A^{-T} y \tag{5.8}$$

5.4.1.3 Κανόνας Οπισθόδρομης Εκμάθησης (Backpropagation Rule).

Στην ενότητα αυτή παρουσιάζουμε ένα βασικό κανόνα εκμάθησης για προσαρμοστικά δίκτυα, που είναι στην ουσία η πιο απλή μέθοδος βαθμωτής ελαχιστοποίησης. Το κεντρικό μέρος αυτού του κανόνα εκμάθησης αφορά στο πως επιλέγεται ένα διάνυσμα κλίσης στο οποίο κάθε στοιχείο ορίζεται ως η παράγωγος του μέτρου σφάλματος ως προς μια παράμετρο. Αυτό γίνεται με τη βοήθεια του κανόνα αλυσίδας, ενός βασικού τύπου για τη διαφόριση σύνθετων συναρτήσεων. Η διαδικασία εύρεσης ενός διανύσματος κλίσης σε ένα δίκτυο αναφέρεται γενικά σαν *οπισθοδρόμηση (backpropagation)* επειδή το διάνυσμα κλίσης υπολογίζεται σε κατεύθυνση αντίθετη από τη ροή της εξόδου κάθε κόμβου. Μόλις επιλεγεί η κλίση, διάφορες τεχνικές βελτιστοποίησης και παλινδρόμησης βασισμένες στις παραγώγους είναι διαθέσιμες για την ενημέρωση των παραμέτρων. Ειδικότερα, εάν χρησιμοποιούμε το διάνυσμα κλίσης σε μια απλή μέθοδο βαθμωτής ελαχιστοποίησης, το προκύπτον παράδειγμα εκμάθησης αναφέρεται συχνά ως *κανόνας οπισθόδρομης εκμάθησης*.
Έστω ένα feed-forward προσαρμόσιμο δίκτυο. Στην αναπαράστασή του διακρίνουμε L επίπεδα και το επίπεδο l (όπου $l = 0,1,...,L; l = 0$ αντιπροσωπεύει το επίπεδο εισαγωγής) έχει $N(l)$ κόμβους. Τότε η έξοδος και η συνάρτηση του κόμβου i $[i = 1,...,N(l)]$ στο επίπεδο l μπορούν να αναπαρασταθούν σαν $x_{l,i}$ και $f_{l,i}$ αντίστοιχα. Ας υποτεθεί ότι δεν υπάρχει καμία σύνδεση μεταξύ μη συνεχόμενων στρωμάτων. Δεδομένου ότι η έξοδος ενός κόμβου εξαρτάται από τα εισερχόμενα σήματα και το σύνολο παραμέτρων του κόμβου, προκύπτει η ακόλουθη γενική έκφραση για τη συνάρτηση των κόμβων $f_{l,i}$:

$$x_{l,i} = f_{l,i}(x_{l-1,1},\dots x_{l-1,N(l-1)}, a, \beta, \gamma, \dots) \tag{5.9}$$

Όπου: α, β, γ κλπ. είναι οι παράμετροι αυτού του κόμβου.

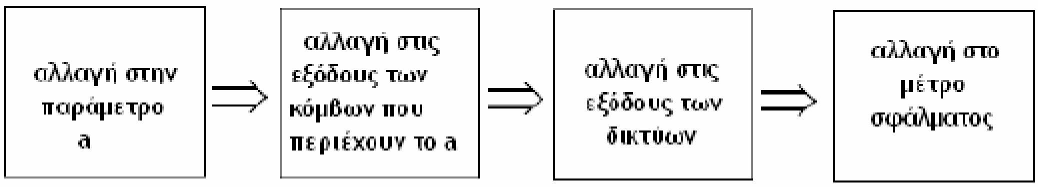
Υποθέτοντας ότι το δοσμένο σύνολο δεδομένων εκπαίδευσης έχει P καταχωρήσεις, μπορεί να οριστεί ένα μέτρο σφάλματος για την *p-οστή* $1 \le p \le P$ καταχώρηση των δεδομένων εκπαίδευσης σαν το άθροισμα των τετραγωνικών σφαλμάτων:

$$E_p = \sum_{k=1}^{N(l)} (d_k - x_{L,k})^2 \tag{5.10}$$

όπου d_k είναι το *k-οστό* συστατικό του *p-οστού* επιθυμητού διανύσματος εξόδου και $x_{L,k}$ είναι το *k-οστό* συστατικό του πραγματικού διανύσματος εξόδου που παράγεται με την παρουσίαση του *p-οστού* διανύσματος εισόδου στο δίκτυο. Προφανώς, όταν το E_p είναι ίσο με το μηδέν, το δίκτυο είναι ικανό να αναπαράγει ακριβώς το επιθυμητό διάνυσμα εξόδου στο *p-οστό* ζευγάρι δεδομένων εκπαίδευσης. Ο στόχος εδώ είναι να ελαχιστοποιηθεί ένα συνολικό μέτρο σφάλματος, που ορίζεται ως:

$$E = \sum_{p=1}^{P} E_P$$

Επιπλέον, υποθέτουμε ότι το E_p εξαρτάται μόνο από τους κόμβους εξόδου. Για να χρησιμοποιηθεί η βαθμωτή ελαχιστοποίηση ώστε να ελαχιστοποιηθεί το μέτρο σφάλματος, πρέπει πρώτα να βρεθεί το διάνυσμα κλίσης. Πριν υπολογιστεί το διάνυσμα κλίσης, πρέπει να παρατηρηθούν οι ακόλουθες αιτιώδεις σχέσεις:



όπου τα βέλη (→) δείχνουν τις αιτιώδεις σχέσεις. Δηλαδή, μια μικρή αλλαγή σε μια παράμετρο *a* θα επηρεάσει την έξοδο του κόμβου που περιέχει το α. Αυτό με τη σειρά του θα επηρεάσει την έξοδο του τελευταίου επιπέδου και συνεπώς το μέτρο σφάλματος. Η βασική αρχή στον υπολογισμό του διανύσματος κλίσης είναι να περαστούν μια σειρά από πληροφορίες παραγώγων ξεκινώντας από το επίπεδο εξόδου και πηγαίνοντας ανάποδα από επίπεδο σε επίπεδο έως ότου καταλήξει η διαδικασία στο επίπεδο εισόδου.

Το σήμα σφάλματος $e_{l,i}$ ορίζεται σαν τη παράγωγο του μέτρου σφάλματος E_p ως προς την έξοδο του κόμβου *i* στο επίπεδο *l*, λαμβάνοντας υπόψη και τις άμεσες και τις έμμεσες πορείες.

$$e_{l,i} = \frac{\partial^{+} E_P}{\partial x_{l,i}} \qquad (5.11)$$

Η έκφραση αυτή ονομάστηκε διατεταγμένη παράγωγος από τον Werbos.

Το σήμα σφάλματος για τον *i-οστό* κόμβο εξόδου (στο επίπεδο *L)* μπορεί να υπολογιστεί απευθείας:

$$e_{1,i} = \frac{\partial^+ E_p}{\partial x_{1,i}} = \frac{\partial E_p}{\partial x_{1,i}} \tag{5.12}$$

Αυτό ισούται με $e_{L,i} = -2(d_i - x_{L,i})$ αν το E_p ορίζεται όπως στην εξίσωση (5.10). Για τον εσωτερικό κόμβο στην *i-οστή* θέση του επιπέδου *l,* το σήμα σφάλματος μπορεί να βρεθεί από τον κανόνα αλυσίδας:

$$e_{l,i} = \underbrace{\frac{\partial^+ E_p}{\partial x_{l,i}}}_{\substack{\text{error signal}\\ \text{at layer } l}} = \sum_{m=1}^{N(l+1)} \underbrace{\frac{\partial^+ E_p}{\partial x_{l+1,m}}}_{\substack{\text{error signal}\\ \text{at layer } l+1}} \frac{\partial f_{l+1,m}}{\partial x_{l,i}} = \sum_{m=1}^{N(l+1)} e_{l+1,m} \frac{\partial f_{l+1,m}}{\partial x_{l,i}} \tag{5.13}$$

Όπου: $0 \le l \le L - 1$**.** Δηλαδή το σήμα σφάλματος ενός εσωτερικού κόμβου στο στρώμα ***l*** μπορεί να εκφραστεί ως γραμμικός συνδυασμός των κόμβων στο στρώμα ***l+1.*** Επομένως, για οποιαδήποτε ***l*** και ***i*** **[** και $1 \le i \le N(l)$], μπορούν να βρεθούν τα:

$$\epsilon_{l,i} = \frac{\partial^+ E_p}{\partial x_{l,i}}$$

Εφαρμόζοντας πρώτα την εξίσωση (5.12) μία φορά για να υπολογιστούν τα σήματα σφάλματος στο επίπεδο εξόδου, και έπειτα εφαρμόζοντας την εξίσωση (5.13) επαναληπτικά έως ότου καταλήξει η διαδικασία στο επιθυμητό επίπεδο *l*. Η διαδικασία αυτή καλείται οπισθοδρόμηση δεδομένου ότι τα σήματα σφάλματος λαμβάνονται διαδοχικά από το επίπεδο εξόδου προς το επίπεδο εισόδου.

Το διάνυσμα κλίσης ορίζεται ως την παράγωγο του μέτρου σφάλματος ως προς κάθε παράμετρο, έτσι πρέπει να εφαρμοστεί ο κανόνας αλυσίδας ξανά για να βρεθεί το διάνυσμα κλίσης. Εάν *α* είναι μια παράμετρος του *i-οστού* κόμβου στο επίπεδο *l,* τότε ισχύει:

$$\frac{\partial^+ E_p}{\partial a} = \frac{\partial^+ E_p}{\partial x_{l,i}} \frac{\partial f_{l,i}}{\partial a} = \epsilon_{l,i} \frac{\partial f_{l,i}}{\partial a} \tag{5.14}$$

Ας σημειωθεί ότι αν επιτραπεί στην παράμετρο α να μοιράζεται μεταξύ διαφορετικών κόμβων, τότε η εξίσωση (5.14) πρέπει να αλλαχθεί σε μια πιο γενική μορφή:

$$\frac{\partial^+ E_p}{\partial a} = \sum_{x^* \in S} \frac{\partial^+ E_p}{\partial x^*} \frac{\partial f^*}{\partial a} \tag{5.15}$$

όπου S είναι το σύνολο των κόμβων που περιέχουν το ***a*** σαν παράμετρο ενώ x* και ***f*** * είναι η έξοδος και η συνάρτηση, αντίστοιχα, ενός γενικού κόμβου στο S.

Η παράγωγος του γενικού μέτρου σφάλματος ***E*** ως προς το ***a*** είναι:

$$\frac{\partial^+ E_P}{\partial a} = \sum_{p=1}^{P} \frac{\partial^+ E_p}{\partial a} \tag{5.16}$$

Συνεπώς, για την απλούστερη βαθμωτή ελαχιστοποίηση χωρίς ελαχιστοποίηση γραμμών, ο τύπος για τη γενική παράμετρο ***a*** είναι:

$$\Delta a = -\eta \frac{\partial^+ E_p}{\partial a} \tag{5.17}$$

Όπου: το n είναι ο ρυθμός εκμάθησης, το οποίο μπορεί να εκφραστεί περαιτέρω ως:

$$\eta = \frac{\kappa}{\sqrt{\sum_{\alpha} \left(\frac{\partial E}{\partial \alpha} \right)^2}} \tag{5.18}$$

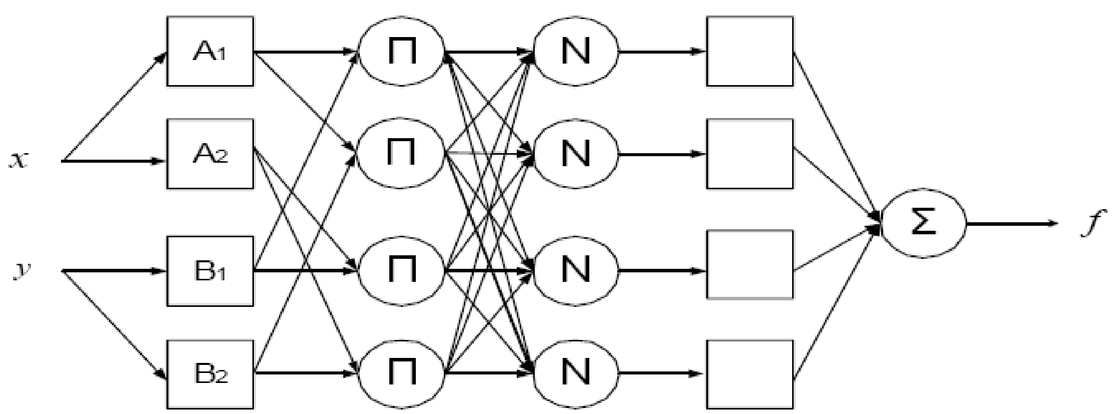
όπου το k είναι το μέγεθος βήματος, το μήκος δηλαδή κάθε μετάβασης κατά μήκος της κατεύθυνσης κλίσης στο διάστημα παραμέτρου. Συνήθως το μέγεθος βήματος μπορεί να αλλαχθεί για να μεταβληθεί η ταχύτητα της σύγκλισης. Όταν ένα feed-forward δίκτυο n-κόμβων αναπαριστάται στην τοπολογική του διάταξη, μπορεί να υπολογιστεί το μέτρο σφάλματος E_p σαν την έξοδο ενός επιπλέον κόμβου με ένδειξη n+1, του οποίου η συνάρτηση κόμβου f_{n+1} μπορεί να οριστεί από τις εξόδους κάθε κόμβου με μικρότερη ένδειξη. (Επομένως, το E_p μπορεί να εξαρτάται άμεσα από οποιουσδήποτε κόμβους.) Εφαρμόζοντας πάλι τον κανόνα αλυσίδας, ισχύει ο ακόλουθος συνοπτικός τύπος για τον υπολογισμό του σήματος σφάλματος.

$$e_i = {\partial E_p}/{\partial x_i}$$

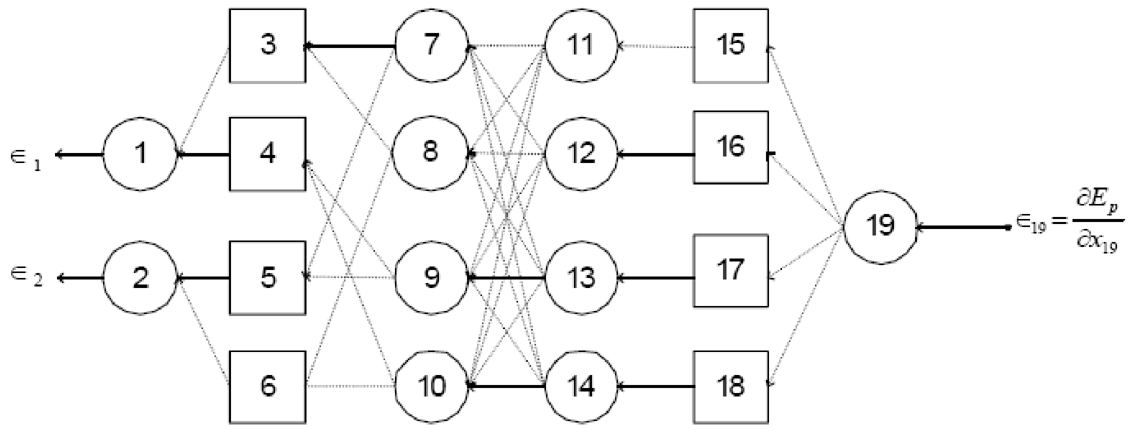
$$e_i = \frac{\partial f_{n+1}}{\partial x_i} + \sum_{i<j\leq n} e_j \frac{\partial f_j}{\partial x_i} \tag{5.19}$$

Όπου: ο πρώτος όρος δείχνει μια άμεση επίδραση του x_i στο E_p μέσω της άμεσης διαδρομής από τον κόμβο *i* στον κόμβο *n+1* και κάθε όρος παραγώγου στο άθροισμα δείχνει την έμμεση επίδραση του x_i στο E_p. Μόλις βρεθεί το σήμα σφάλματος για κάθε κόμβο, τότε το διάνυσμα κλίσης για τις παραμέτρους παράγεται όπως πριν.

Ένας άλλος συστηματικός τρόπος να υπολογιστούν τα σήματα σφάλματος είναι μέσω της αναπαράστασης του δικτύου διάδοσης σφάλματος (ή του μοντέλου ευαισθησίας), το οποίο λαμβάνεται από το αρχικό προσαρμόσιμο δίκτυο με την αντιστροφή των συνδέσεων και την παροχή των σημάτων σφάλματος στο επίπεδο εξόδου ως είσοδοι στο νέο δίκτυο. Το ακόλουθο σχήμα δείχνει το μοντέλο ANFIS και το δίκτυο διάδοσης σφάλματος του:



Σχήμα 5.18: Προσαρμοστικό δίκτυο.



Σχήμα 5.19: Δίκτυο διάδοσης σφάλματος.

Τώρα θα υπολογιστούν τα σήματα σφάλματος στους εσωτερικούς κόμβους. Χρησιμοποιούνται τα f_i και x_i για το συμβολισμό της συνάρτησης και της εξόδου του κόμβου *i.* Η έξοδος του κόμβου *i* είναι το σήμα σφάλματος αυτού του κόμβου στο πραγματικό προσαρμόσιμο δίκτυο. Σε σύμβολα, αν επιλεχθεί το τετραγωνικό μέτρο σφάλματος για E_P, τότε ισχύει $e_{19} = -2(d_{19} - x_{19})$.

Αυτό είναι επειδή ο κόμβος 19 είναι μόνο ένας κόμβος προσωρινής αποθήκευσης στο δίκτυο διάδοσης σφάλματος. Για τους κόμβους 15, 16, 17 και 18 ισχύουν:

$$e_{18} = \frac{\partial^+ E_p}{\partial x_{18}} = \frac{\partial^+ E_p}{\partial x_{19}} \frac{\partial f_{19}}{\partial x_{18}} = e_{19} \frac{\partial f_{19}}{\partial x_{18}}$$

$$e_{17} = \frac{\partial^+ E_p}{\partial x_{17}} = \frac{\partial^+ E_p}{\partial x_{19}} \frac{\partial f_{19}}{\partial x_{17}} = e_{19} \frac{\partial f_{19}}{\partial x_{17}}$$

$$e_{16} = \frac{\partial^+ E_p}{\partial x_{16}} = \frac{\partial^+ E_p}{\partial x_{19}} \frac{\partial f_{19}}{\partial x_{16}} = e_{19} \frac{\partial f_{19}}{\partial x_{16}}$$

$$e_{15} = \frac{\partial^+ E_p}{\partial x_{15}} = \frac{\partial^+ E_p}{\partial x_{19}} \frac{\partial f_{19}}{\partial x_{15}} = e_{19} \frac{\partial f_{19}}{\partial x_{15}}$$

Αυτό είναι επειδή όλοι αυτοί οι κόμβοι εξαρτώνται από τον κόμβο 19. Αυτό ισχύει και για τους κόμβους 11, 12, 13 και 14 με τη διαφορά ότι κάθε ένας από αυτούς εξαρτάται από διαφορετικό κόμβο:

$$e_{14} = \frac{\partial^+ E_p}{\partial x_{14}} = \frac{\partial^+ E_p}{\partial x_{18}} \frac{\partial f_{18}}{\partial x_{14}} = e_{18} \frac{\partial f_{18}}{\partial x_{14}}$$

$$e_{13} = \frac{\partial^+ E_p}{\partial x_{13}} = \frac{\partial^+ E_p}{\partial x_{17}} \frac{\partial f_{17}}{\partial x_{13}} = e_{17} \frac{\partial f_{17}}{\partial x_{13}}$$

$$e_{12} = \frac{\partial^+ E_p}{\partial x_{12}} = \frac{\partial^+ E_p}{\partial x_{16}} \frac{\partial f_{16}}{\partial x_{12}} = e_{16} \frac{\partial f_{16}}{\partial x_{12}}$$

$$e_{11} = \frac{\partial^+ E_p}{\partial x_{11}} = \frac{\partial^+ E_p}{\partial x_{15}} \frac{\partial f_{15}}{\partial x_{11}} = e_{15} \frac{\partial f_{15}}{\partial x_{11}}$$

Αντιθέτως, οι κόμβοι 7, 8, 9 και 10 εξαρτώνται από τέσσερις διαφορετικούς κόμβους όπως φαίνεται στο σχήμα 5.19. Έτσι, ισχύει:

$$e_{10} = \frac{\partial^+ E_p}{\partial x_{10}} = e_{14} \frac{\partial f_{14}}{\partial x_{10}} + e_{13} \frac{\partial f_{13}}{\partial x_{10}} + e_{12} \frac{\partial f_{12}}{\partial x_{10}} + e_{11} \frac{\partial f_{11}}{\partial x_{10}}$$

$$e_{9} = \frac{\partial^+ E_p}{\partial x_{9}} = e_{14} \frac{\partial f_{14}}{\partial x_{9}} + e_{13} \frac{\partial f_{13}}{\partial x_{9}} + e_{12} \frac{\partial f_{12}}{\partial x_{9}} + e_{11} \frac{\partial f_{11}}{\partial x_{9}}$$

$$e_{8} = \frac{\partial^+ E_p}{\partial x_{8}} = e_{14} \frac{\partial f_{14}}{\partial x_{8}} + e_{13} \frac{\partial f_{13}}{\partial x_{8}} + e_{12} \frac{\partial f_{12}}{\partial x_{8}} + e_{11} \frac{\partial f_{11}}{\partial x_{8}}$$

$$e_{7} = \frac{\partial^+ E_p}{\partial x_{7}} = e_{14} \frac{\partial f_{14}}{\partial x_{7}} + e_{13} \frac{\partial f_{13}}{\partial x_{7}} + e_{12} \frac{\partial f_{12}}{\partial x_{7}} + e_{11} \frac{\partial f_{11}}{\partial x_{7}}$$

Οι κόμβοι 3, 4, 5 και 6 εξαρτώνται από δύο διαφορετικούς κόμβους και σαν αποτέλεσμα προκύπτουν οι ακόλουθες σχέσεις:

$$e_6 = \frac{\partial^+ E_p}{\partial x_6} = \frac{\partial^+ E_p}{\partial x_{10}} \frac{\partial f_{10}}{\partial x_6} + \frac{\partial^+ E_p}{\partial x_8} \frac{\partial f_8}{\partial x_6} = e_{10} \frac{\partial f_{10}}{\partial x_6} + e_8 \frac{\partial f_8}{\partial x_6}$$

$$e_5 = \frac{\partial^+ E_p}{\partial x_5} = \frac{\partial^+ E_p}{\partial x_9} \frac{\partial f_9}{\partial x_5} + \frac{\partial^+ E_p}{\partial x_7} \frac{\partial f_7}{\partial x_5} = e_9 \frac{\partial f_9}{\partial x_5} + e_7 \frac{\partial f_7}{\partial x_5}$$

$$e_4 = \frac{\partial^+ E_p}{\partial x_4} = \frac{\partial^+ E_p}{\partial x_{10}} \frac{\partial f_{10}}{\partial x_4} + \frac{\partial^+ E_p}{\partial x_9} \frac{\partial f_9}{\partial x_4} = e_{10} \frac{\partial f_{10}}{\partial x_4} + e_9 \frac{\partial f_9}{\partial x_4}$$

$$e_3 = \frac{\partial^+ E_p}{\partial x_3} = \frac{\partial^+ E_p}{\partial x_8} \frac{\partial f_8}{\partial x_3} + \frac{\partial^+ E_p}{\partial x_7} \frac{\partial f_7}{\partial x_3} = e_8 \frac{\partial f_8}{\partial x_3} + e_7 \frac{\partial f_7}{\partial x_3}$$

Τέλος, οι κόμβοι 1 και 2 εξαρτώνται από δύο κόμβους και έτσι ισχύει:

$$e_2 = \frac{\partial^+ E_p}{\partial x_2} = \frac{\partial^+ E_p}{\partial x_6} \frac{\partial f_6}{\partial x_2} + \frac{\partial^+ E_p}{\partial x_5} \frac{\partial f_5}{\partial x_2} = e_6 \frac{\partial f_2}{\partial x_2} + e_5 \frac{\partial f_5}{\partial x_2}$$

$$e_1 = \frac{\partial^+ E_p}{\partial x_1} = \frac{\partial^+ E_p}{\partial x_4} \frac{\partial f_4}{\partial x_1} + \frac{\partial^+ E_p}{\partial x_3} \frac{\partial f_3}{\partial x_1} = e_4 \frac{\partial f_4}{\partial x_1} + e_3 \frac{\partial f_3}{\partial x_1}$$

5.4.1.4 Υβριδικός Αλγόριθμος Εκπαίδευσης.

Το ANFIS χρησιμοποιεί έναν υβριδικό αλγόριθμο εκμάθησης για να προσδιορίσει τις παραμέτρους του ασαφούς συστήματος τύπου Sugeno. Εφαρμόζει έναν συνδυασμό της μεθόδου ελαχίστων τετραγώνων (least-squares) και της μεθόδου οπισθόδρομης βαθμωτής ελαχιστοποίησης (backpropagation gradient descent) για την εκπαίδευση παραμέτρων των συναρτήσεων συμμετοχής του FIS ώστε να μιμηθεί ένα δοσμένο σύνολο δεδομένων εκπαίδευσης.

Η έξοδος ενός προσαρμοστικού δικτύου είναι γραμμική σε μερικές από τις παραμέτρους του δικτύου. Έτσι μπορούν να προσδιοριστούν αυτές οι γραμμικές παράμετροι με τη γραμμική μέθοδο ελαχίστων τετραγώνων που περιγράφτηκε παραπάνω. Από την αρχιτεκτονική δομή του ANFIS παρατηρούμε ότι οι τιμές των παραμέτρων από τις προϋποθέσεις είναι προκαθορισμένες (αμετάβλητες), η ολική έξοδος μπορεί να εκφραστεί σαν ο γραμμικός συνδυασμός των παραμέτρων από τα συμπεράσματα. Σε σύμβολα, η έξοδος f στο σχήμα 5.17(β) ξαναγράφεται:

$$f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2$$

$$= \overline{w}_1 (p_1 x + q_1 y + r_1) + \overline{w}_2 (p_2 x + q_2 y + r_2)$$

$$= (\overline{w}_1 x) p_1 + (\overline{w}_1 y) q_1 + (\overline{w}_1) r_1 + (\overline{w}_2 x) p_2 + (\overline{w_2 y}) q_2 + (\overline{w}_2) r_2$$

Η οποία είναι γραμμική στις παραμέτρους των συμπερασμάτων p_1, q_1, r_1, p_2, q_2, και r_2. Έστω ότι:

S=σύνολο όλων των παραμέτρων.
S_1= σύνολο των (μη γραμμικών) παραμέτρων από τις προϋποθέσεις .
S_2 =σύνολο των (γραμμικών) παραμέτρων από τα συμπεράσματα.

Το προσαρμοστικό δίκτυο έχει μία έξοδο που αναπαριστάται ως:

O = F (i , S) (5.20)

όπου i είναι το διάνυσμα των μεταβλητών εισόδου, S είναι το σύνολο των παραμέτρων, και *F* είναι η συνολική συνάρτηση που εφαρμόζεται από το προσαρμοστικό δίκτυο.

$$S = S_1 \oplus S_2 \qquad (5.21)$$

Οι H(.) και F(.) είναι η ταυτοτικές συναρτήσεις και η συνάρτηση του FIS αντιστοίχως στην εξίσωση, η $H \circ F$ είναι γραμμική στα στοιχεία του S_2, τότε εφαρμόζοντας την *H* στην εξίσωση (5.20), προκύπτει ότι $H(o) = H \circ F(B_i, S)$. Το $\oplus$ αντιπροσωπεύει την άμεση άθροιση (direct sum).

Η H(.) είναι η ταυτοτική συνάρτηση και η F(.,.) είναι η συνάρτηση του συστήματος ασαφούς συμπερασμού αντίστοιχα. Κατά συνέπεια, ο υβριδικός αλγόριθμος μάθησης που αναπτύχθηκε παραπάνω μπορεί να εφαρμοστεί άμεσα. Αναλυτικότερα, όπως συμβαίνει και στα νευρωνικά δίκτυα, θα ξεχωρίσουν δύο φάσεις στη διαδικασία εκπαίδευσης:

- Πέρασμα προς τα εμπρός (forward pass): Στη φάση αυτή το σήμα εισόδου διαδίδεται από το επίπεδο 1 μέχρι το επίπεδο 4 και οι παράμετροι (p_i, q_i, r_i), i = 1,2 εκτιμούνται με τη μέθοδο των ελαχίστων τετραγώνων.
- Πέρασμα προς τα πίσω (backward pass): Στη φάση αυτή έχοντας μία ένδειξη του λάθους πραγματοποιείται μία διόρθωση κατά μία ποσότητα που ορίζεται από τη μέθοδο της βαθμωτής κατάβασης (Gradient Descent), των μεταβλητών α_i, b_i, c_i.

Το ANFIS διασπά το σύνολο των παραμέτρων του σε δύο υποσύνολα, εκ των οποίων το ένα αποτελεί γραμμικό σύνολο παραμέτρων και, επομένως, είναι δυνατόν να εκπαιδευτεί με γραμμικούς αλγόριθμους, όπως η μέθοδος των ελαχίστων τετραγώνων (least squared). Οι γραμμικοί αλγόριθμοι μάθησης επιτυγχάνουν συνολικά ελάχιστα (global minimums) της συνάρτησης κόστους στο χώρο των παραμέτρων τους και είναι αποδοτικοί από πλευράς απαιτούμενου υπολογιστικού χρόνου. Το δεύτερο υποσύνολο παραμέτρων εκπαιδεύεται με αλγόριθμους που μπορούν να δημιουργήσουν μη-γραμμικές απεικονίσεις, όπως η βαθμωτή κατάβαση (gradient descent). Οι αλγόριθμοι αυτοί είναι απαιτητικοί από πλευράς απαιτούμενου χρόνου εκπαίδευσης και δεν υπάρχει εγγύηση για την πραγματοποίηση του συνολικού ελαχίστου της συνάρτησης κόστους στον χώρο των παραμέτρων τους. Ο αλγόριθμος μάθησης του ANFIS συνδυάζει περάσματα προς τα εμπρός και προς τα πίσω. Στο πέρασμα εμπρός γίνεται η μάθηση του συνόλου των γραμμικών παραμέτρων και στο πέρασμα προς τα πίσω γίνεται η προσαρμογή των μη-γραμμικών αντίστοιχα. Στον παρακάτω **πίνακα** συνοψίζονται οι δραστηριότητες του κάθε περάσματος.

	Εμπροσθόδρομο πέρασμα	**Οπισθόδρομο πέρασμα**
Παράμετροι των προϋποθέσεων	Προκαθορισμένες (αμετάβλητες)	Βαθμωτή φθίνουσα μέθοδος
Παράμετροι των συμπερασμάτων	Εκτιμήτρια ελαχίστων τετραγώνων	Προκαθορισμένες (αμετάβλητες)
Σήματα	Έξοδοι των κόμβων	Σήματα σφάλματος

Πίνακας 5.1: *Δραστηριότητες κατά την εκπαίδευση.*

Το πλεονέκτημα της μεθόδου έγκειται στο γεγονός της ύπαρξης τόσο γραμμικών όσο και μη γραμμικών μεθόδων, που την καθιστά ταχύτερη από τα κλασικά νευρωνικά δίκτυα. Οι παράμετροι των συμπερασμάτων που ευρίσκονται με αυτόν τον τρόπο, είναι βέλτιστες υπό την προϋπόθεση ότι οι παράμετροι των προϋποθέσεων είναι προκαθορισμένες (αμετάβλητες). Αναλόγως, η υβριδική προσέγγιση συγκλίνει πολύ γρηγορότερα μιας και μειώνει τις διαστάσεις του διαστήματος αναζήτησης της αρχικής καθαρής μεθόδου οπισθόδρομης διάδοσης.

Μέχρι τώρα χρησιμοποιούσαμε προκαθορισμένες και αυθαίρετα επιλεγμένες συναρτήσεις συμμετοχής. Πολλές φορές έχουμε μια συλλογή από δεδομένα εισόδου / εξόδου και δεν μπορούμε να χρησιμοποιήσουμε ένα προκαθορισμένο μοντέλο βασισμένο στα στοιχεία αυτά ώστε να ξέρουμε ποιες συναρτήσεις συμμετοχής και με ποιες παραμέτρους να χρησιμοποιήσουμε. Τότε μπορούμε να βοηθηθούμε από το ANFIS, ώστε να βρούμε τις καταλληλότερες συναρτήσεις συμμετοχής.

Οι νεύρο-προσαρμοστικές τεχνικές είναι αρκετά απλές. Έτσι παρέχεται μια μέθοδος ώστε το ασαφές μοντέλο να εκπαιδευτεί με τις πληροφορίες που του δίνει το σύνολο των δεδομένων, ώστε να υπολογιστούν οι παράμετροι των συναρτήσεων συμμετοχής με σκοπό την εύρεση του καλύτερου τρόπο ώστε να επιτρέπει στο σύστημα εξαγωγής συμπερασμάτων να ανιχνεύει τα δεδομένα εισόδου/ εξόδου. Η μέθοδος αυτή είναι παρόμοια με αυτή των νευρωτικών δικτύων.

Η προσαρμογή των παραμέτρων των συναρτήσεων συμμετοχής από ένα ANFIS γίνεται με αλγόριθμους εκμάθησης είτε μόνο back propagation είτε με έναν υβριδικό αλγόριθμο, συνδυασμό back propagation και μεθόδου ελαχίστων τετραγώνων.

Ο υπολογισμός των παραμέτρων αυτών, που αλλάζουν κατά τη διάρκεια της διαδικασίας εκμάθησης, διευκολύνεται από ένα βαθμωτό διάνυσμα που μας δείχνει πόσο καλά το FIS μοντελοποιεί τα δεδομένα εισόδου/ εξόδου για το εν λόγο σύνολο παραμέτρων. Μόλις αποκτήσουμε το βαθμωτό διάνυσμα μπορούμε να εφαρμόσουμε κάποιον από τους πολλούς αλγόριθμους βελτιστοποίησης, για να βελτιώσουμε αυτές τις παραμέτρους, ούτως ώστε να μειώσουμε το μέγεθος του σφάλματος (συνήθως το άθροισμα των τετραγώνων των διαφορών μεταξύ πραγματικών και επιθυμητών τιμών).

5.4.1.5 Περιορισμοί Του ANFIS

Οι κυριότεροι περιορισμοί του ANFIS που αναφέρονται στην διεθνή βιβλιογραφία είναι:

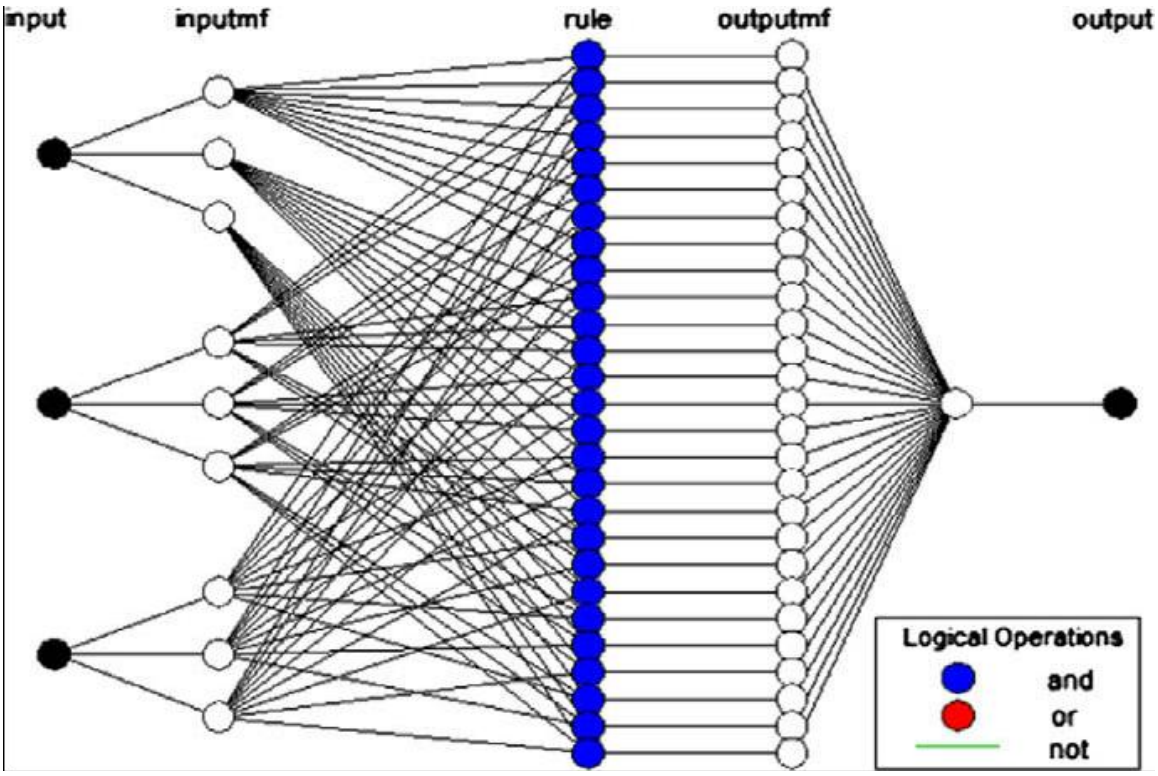
- Βασίζεται σε ένα FIS τύπου Sugeno.
- Έχει μια έξοδο, που λαμβάνεται με τη μέθοδο αποσαφοποίησης σταθμισμένου μέσου. Όλες οι συναρτήσεις συμμετοχής εξόδου πρέπει να είναι ίδιου τύπου, είτε γραμμικές είτε σταθερές.
- Δεν μπορεί να γίνεται κοινή χρήση κανόνων. Διαφορετικοί κανόνες δεν μπορούν να έχουν την ίδια συνάρτηση συμμετοχής εξόδου.
- Πρέπει να υπάρχουν βάρη σε κάθε κανόνα.
- Δεν μπορεί να χρησιμοποιηθεί οποιαδήποτε συνάρτηση συμμετοχής, παρά μόνο προκαθορισμένες που επιβάλλουν οι ορισμοί του ANFIS.

ΚΕΦΑΛΑΙΟ 6^{Ο}

Περιγραφή Νεύρο-Ασαφούς Μοντέλου

Αποτελέσματα-Συμπεράσματα

Η εφαρμογή του νευρο-ασαφούς μοντέλου έγινε στο Matlab, κάνοντας χρήση του fuzzy logic toolbox, το οποίο περιέχει όλα τα απαραίτητα εργαλεία για την ολοκληρωμένη ανάπτυξη και αναλυτική χρήση ενός Προσαρμοστικού Νεύρο-Ασαφούς Συστήματος Συμπερασμού (Adaptive Neural Fuzzy Inference, ANFIS). Γραφική απεικόνιση του οποίου αποτελεί το ακόλουθο σχήμα.



6.1 Επιλογή εισόδων

Αρχίζοντας το σχεδιασμό του νεύρο-ασαφούς συστήματος, πρέπει να ορίσουμε τον αριθμό και το είδος των εισόδων. Στην πιο απλή περίπτωση, μπορούμε να χρησιμοποιήσουμε τις δύο τελευταίες τιμές, για πρόβλεψη του επόμενου βήματος, δηλαδή:

είσοδοι: (τ-1), (τ-2)→ έξοδος: (τ)

Σε αυτή τη βασική δομή μπορούμε να προσθέσουμε επιπλέον εισόδους, προσέχοντας πάντα το κέρδος, που εκφράζεται σε μείωση του σφάλματος, να είναι σημαντικό σε σχέση με την αύξηση της πολυπλοκότητας του συστήματος.

είσοδοι: (τ-1), (τ-2), ..., (τ-n)→ έξοδος: (τ)

Σαν μια διαφορετική προσέγγιση για την επιλογή των εισόδων θα μπορούσαμε να χρησιμοποιήσουμε τιμές που θεωρούμε ότι σχετίζονται, στατιστικά, με κάποιο τρόπο με την έξοδο του συστήματος. Μια τέτοια πληροφορία μπορούμε να αντλήσουμε από τα διαγράμματα των συσχετίσεων και των μερικών αυτοσυσχετίσεων. Αν παρατηρήσουμε κάποια στατιστικά σημαντική τιμή, έστω **h,** τότε οι είσοδοι μπορεί να είναι:

είσοδοι: (τ-1), (τ-2), **...,** (τ-n), (τ- h) → έξοδος: (τ)

Σε κάθε περίπτωση θα πρέπει να εξασφαλίζουμε πως η αύξηση του αριθμού των εισόδων, που μεταφράζεται σε αύξηση της πολυπλοκότητας του συστήματος, επιφέρει ανάλογη μείωση του σφάλματος.

Το συγκεκριμένο μοντέλο χρησιμοποιεί την πιο απλή περίπτωση των δύο εισόδων καθώς τα αποτελέσματα του κρίνονται ικανοποιητικά, επομένως δεν υπάρχει λόγος αύξησης της πολυπλοκότητας του. Το μοντέλο χρησιμοποιεί ως εισόδους μηνιαία φορτία του παρελθόντος την περίοδο από **1/1973 έως 4/2006** τα οποία καταναλωθήκαν στις Η.Π.Α συμφωνά με την στατιστική υπηρεσία και τα οποία βρήκαμε στην ιστοσελίδα **www.economagic.com**, ενώ η έξοδος είναι η πρόβλεψη του επομένου μηνιαίου φορτίου.

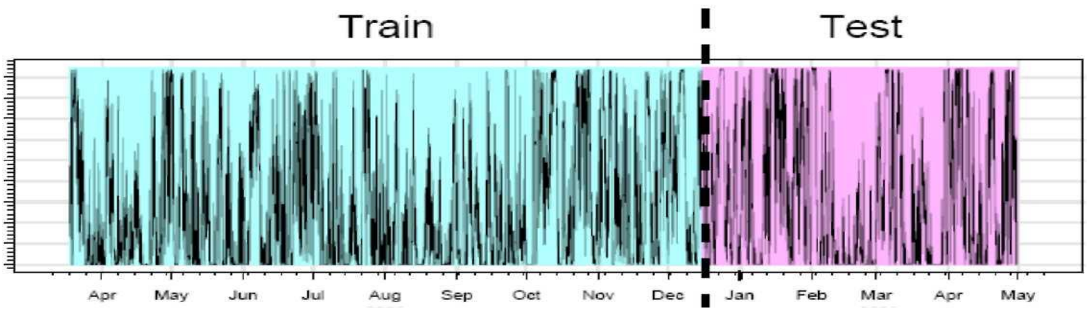
6.2 Εκπαίδευση και Αποτελέσματα πρόβλεψης

6.2.1 Διαδικασία Εκπαίδευσης

Έχοντας ορίσει τον επιθυμητό αριθμό και το είδος των εισόδων και των εξόδων, δημιουργούμε τα κατάλληλα υποδείγματα εισόδων και επιθυμητών εξόδων, που θα χρησιμοποιηθούν για την εκπαίδευση και την αξιολόγηση του μοντέλου.

Για την αξιολόγηση της απόδοσης ενός μοντέλου πρόβλεψης λοιπόν, η κάθε χρονοσειρά χωρίζεται σε δύο τμήματα.

- Το πρώτο σύνολο σημείων για την εκμάθηση του μοντέλου (training set). Για τα διάφορα μοντέλα αυτό το σύνολο σημείων δίνει τις απαραίτητες παραμέτρους που θα χρησιμοποιηθούν στην εφαρμογή.
- Το δεύτερο σύνολο σημείων χρησιμοποιείται στην εκτίμηση του λάθους πρόβλεψης (test set). Αυτά είναι τα σημεία αναφοράς για τα οποία γίνεται πρόβλεψη και συγκρίνεται με τις πραγματικές παρατηρήσεις.



Είναι σύνηθες στη στατιστική ανάλυση των χρονοσειρών να χρησιμοποιούμε το 80% των δεδομένων για την προσαρμογή ή την εκμάθηση του μοντέλου (training set) και το 20% για την αξιολόγηση ης προβλεπτικής του ικανότητας (test set). Γεγονός που ισχύει και στην δική μας περίπτωση, έτσι έχουμε λάβει ως δεδομένα εκπαίδευσης τις πρώτες 320 τιμές του δείγματος μας (Δηλαδή από 1/1973-8/1999) και δεδομένα ελέγχου από 321–400 (9/1999-4/2006).

Στη συνέχεια δημιουργούμε ένα ANFIS, με τη βοήθεια του fuzzy logic toolbox, το οποίο υπολογίζει τις καταλληλότερες παραμέτρους των συναρτήσεων συμμετοχής (MF parameters). Ορίζουμε τις παραμέτρους εκπαίδευσης και αφού ολοκληρωθεί η εκπαίδευση του συστήματος χρησιμοποιούμε τα δεδομένα ελέγχου για να αξιολογήσουμε το μοντέλο.

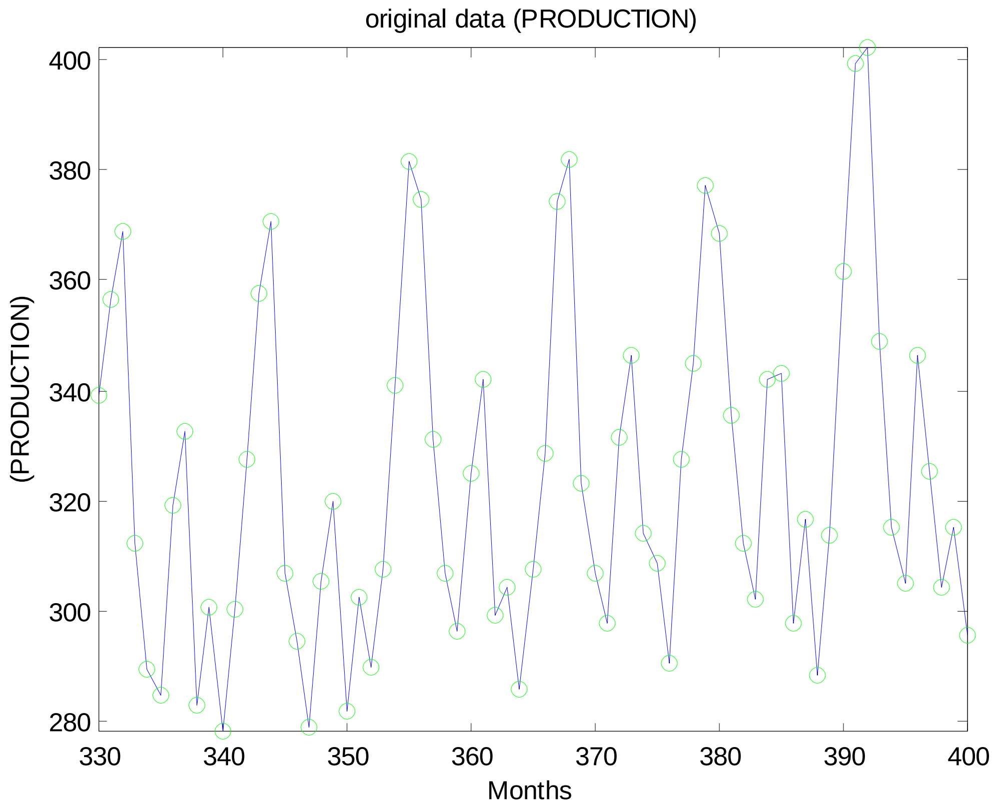
Η εκπαίδευση των νευροασαφών δικτύων γίνεται με μια υβριδική μέθοδο που είναι συνδυασμός του αλγορίθμου backpropagation και της μεθόδου των ελαχίστων τετραγώνων. Σε κάθε περίπτωση, όταν εκπαιδεύουμε ένα μοντέλο, δοκιμάζουμε αρκετούς συνδυασμούς παραμέτρων (MFs, epochs, step size), με σκοπό την επίτευξη μικρότερου σφάλματος.

Έχοντας εκπαιδεύσει το μοντέλο, προχωρούμε στη διενέργεια προβλέψεων εκτός δείγματος (test set). Δεδομένου ότι το μοντέλο μας έχει την ικανότητα να κάνει μια πρόβλεψη, χρησιμοποιώντας πραγματικές παρελθοντικές τιμές.

6.2.2 Αποτελέσματα Πρόβλεψης

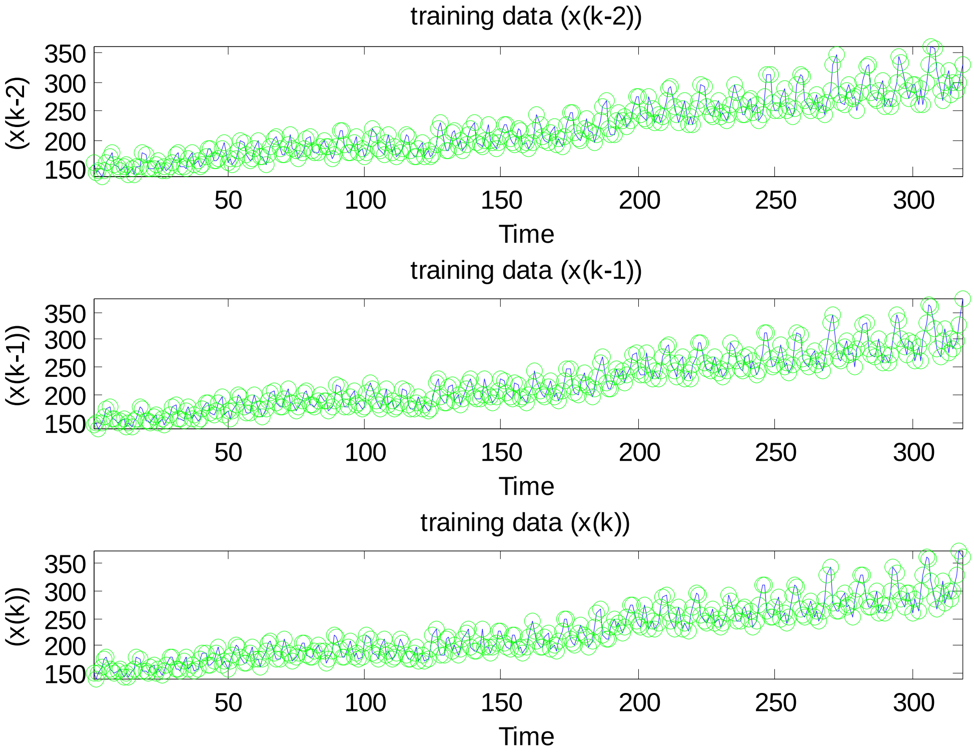
Τα αποτελέσματα από την εφαρμογή του Προσαρμοστικού Νεύρο-Ασαφούς Συστήματος Συμπερασμού (Adaptive Neural Fuzzy Inference, ANFIS) που έγινε στο λογισμικό της Matlab, παρατίθενται παρακάτω.

Στο παρακάτω διάγραμμα έχουμε την σχηματική αναπαράσταση των δεδομένων (φορτίων) ως προς τον χρόνο. Συγκεκριμένα, από τις 400 τιμές του δείγματος χρησιμοποιούμε τις τελευταίες 70.

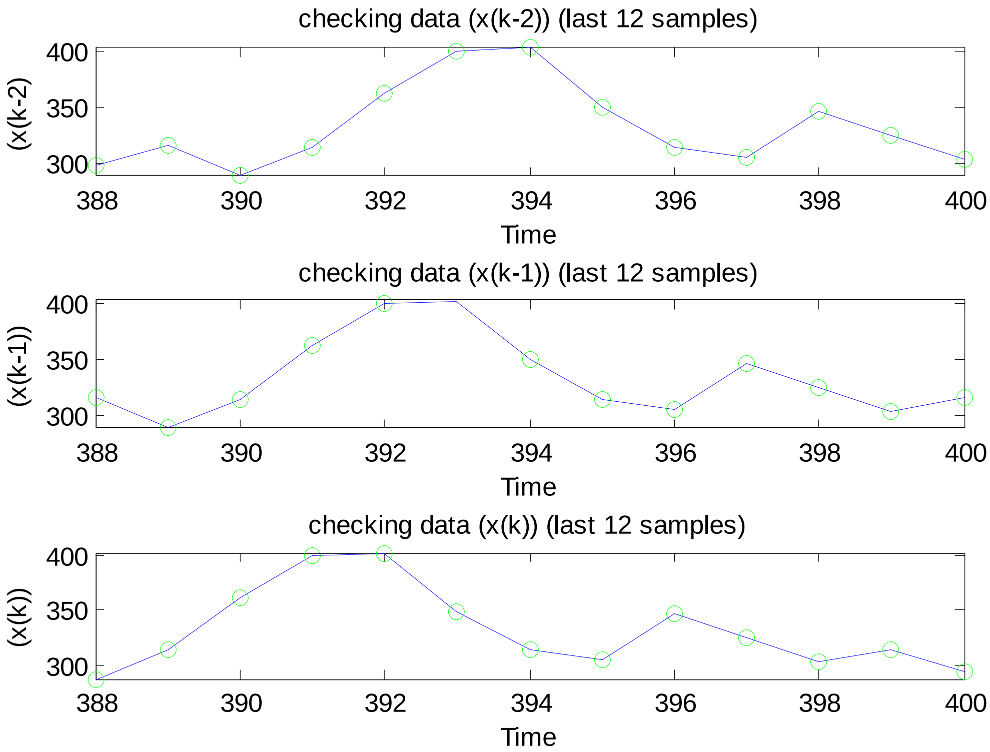


<u>Σχημα 6.1</u>: *Σχηματική απεικόνιση των τελευταίων 70 τιμών της χρονοσειρας φορτίου.*

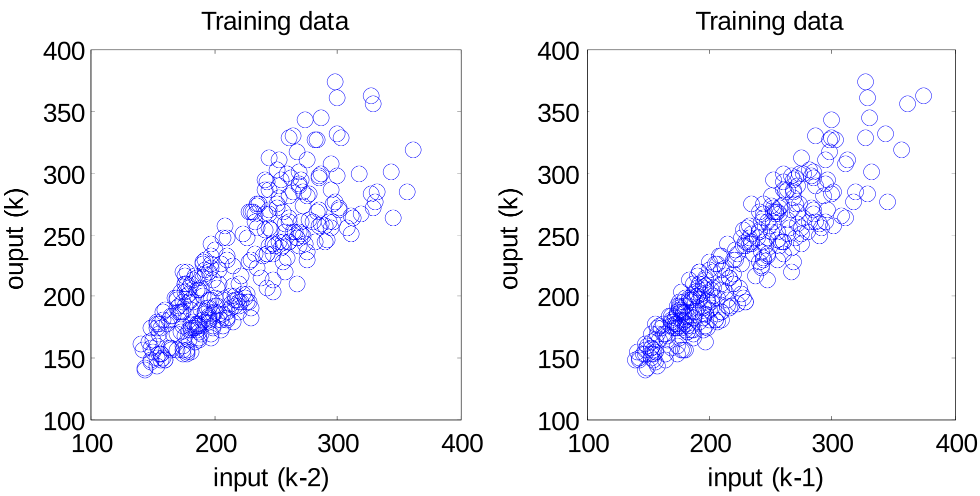
Είναι σύνηθες στη στατιστική ανάλυση των χρονοσειρών να χρησιμοποιούμε το 80% των δεδομένων για την προσαρμογή ή την εκμάθηση του μοντέλου (training set) και το 20% για την αξιολόγηση ης προβλεπτικής του ικανότητας (test set). Γεγονός που ισχύει και στην δική μας περίπτωση, έτσι έχουμε λάβει ως δεδομένα εκπαίδευσης τις πρώτες 320 τιμές του δείγματος μας (Δηλαδή από 1/1973-8/1999) και δεδομένα ελέγχου από 321–400 (9/1999-4/2006). Στην συνέχεια παρατίθενται οι γραφικές απεικονίσεις των δεδομένων εκπαίδευσης *Σχήμα 6.2*, όπως και των δεδομένων ελέγχου *Σχήμα 6.3*.



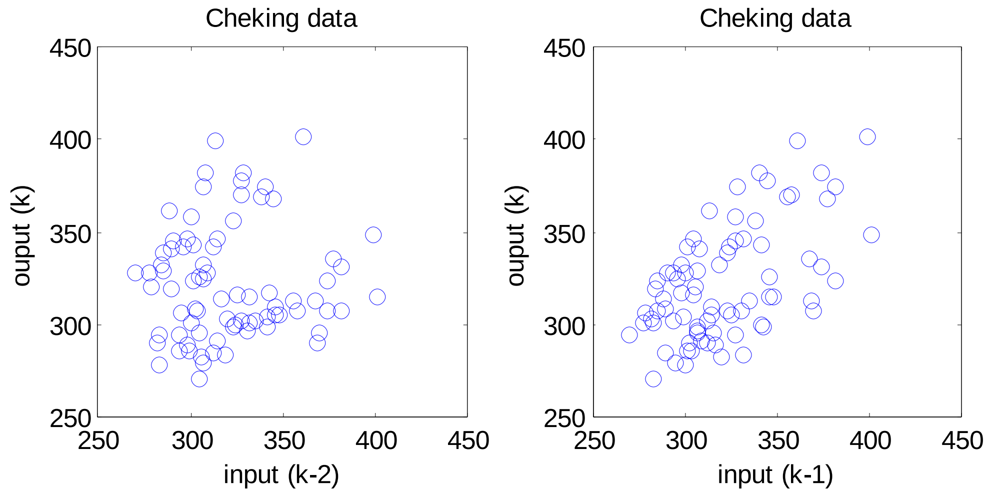
Σχήμα 6.2: *Απεικόνιση δεδομένων εκπαίδευσης του ελεγκτή ANFIS.*



Σχήμα 6.3: *Απεικόνιση δεδομένων ελέγχου για τις τελευταίες 12 τιμές.*

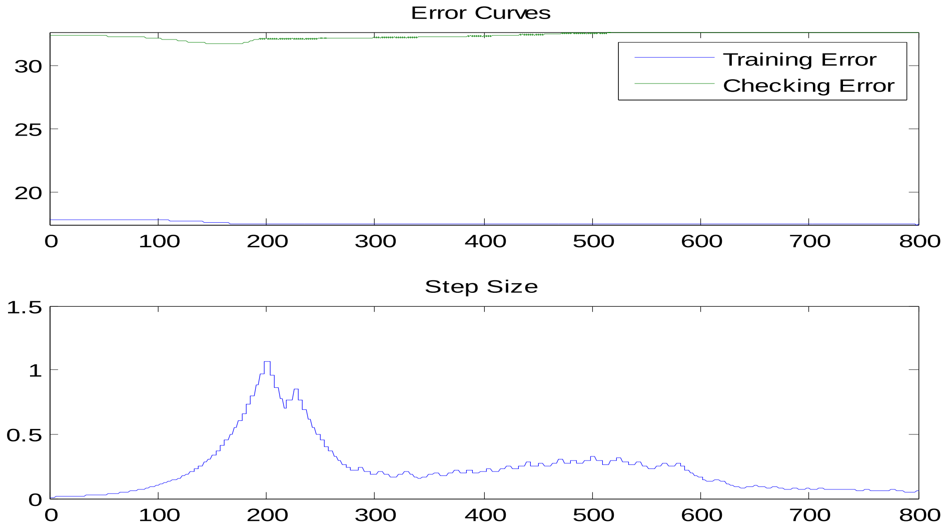


Σχήμα 6.4: *Δισδιάστατη απεικόνιση των δεδομένων εκπαίδευσης.*



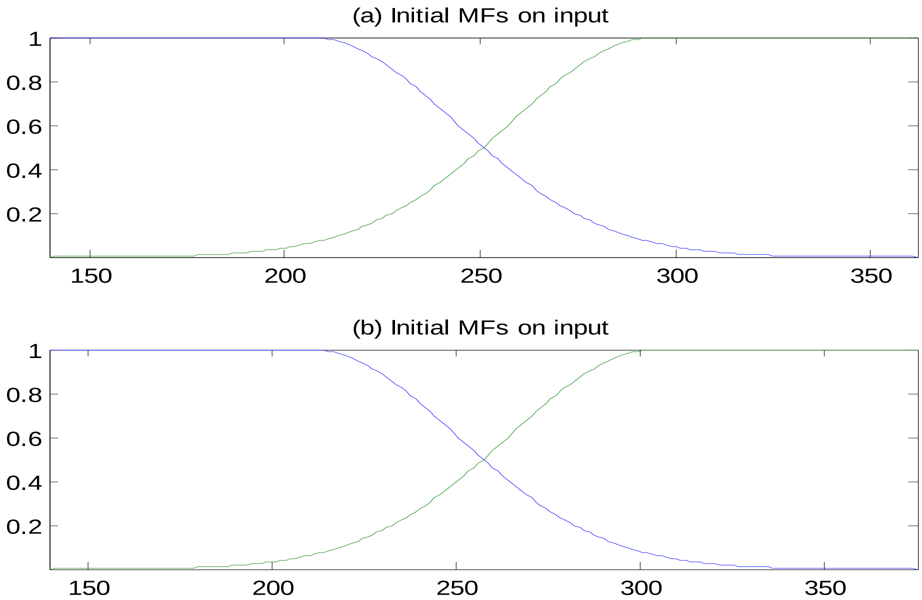
Σχήμα 6.5: *Δισδιάστατη απεικόνιση των δεδομένων ελέγχου.*

Στο *Σχήμα 6.6* παρουσιάζονται δύο ξεχωριστές γραφικές παραστάσεις, η πρώτη με το όνομα "error curves" και η δεύτερη με το όνομα "step size". Στην πρώτη γραφική παράσταση απεικονίζονται τα σφάλματα εκπαίδευσης και ελέγχου (ως τη ρίζα του μέσου τετραγωνικού σφάλματος-RMSE) συναρτήσει του αριθμού των επαναλήψεων (epochs). Στη δεύτερη γραφική παράσταση σχεδιάζουμε την καμπύλη του μεγέθους βήματος συναρτήσει των εποχών. Αυτό μας βοηθάει να προσαρμόσουμε το αρχικό μέγεθος βήματος στο βέλτιστο, βλέποντας τους αντίστοιχους ρυθμούς μείωσης και αύξησης, πάνω στο σχήμα. Μετά από πολλές επαναλήψεις δοκιμής-σφάλματος καθορίστηκε ότι για epochs=800, το σφάλμα RMSE ελαχιστοποιούνταν (περισσότερες επαναλήψεις (epochs) δεν θα το μείωναν περαιτέρω, έτσι το μέγεθος βήματος πρόσαρμόστηκε από *0.01 που ήταν η αρχική του τιμή*, στην *βέλτιστη 0.060173*.

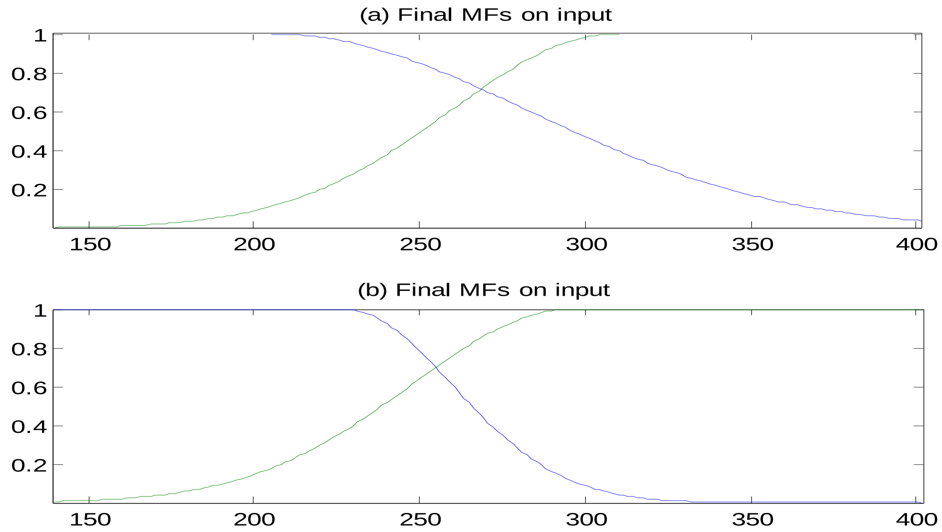


<u>Σχήμα 6.6:</u> *Καμπύλη σφάλματος και μέγεθος βήματος κατά την εκπαίδευση*.

Στην συνέχεια, παρουσιάζονται οι αρχικές συναρτήσεις συμμετοχής πριν την εκπαίδευση για τις 2 εισόδους, που προέκυψαν από την προσαρμογή των παραμέτρων τους με τη βοήθεια του ANFIS (Σχήμα 6.7 με τίτλο "initial MFs") και τις τελικές συναρτήσεις συμμετοχής που προέκυψαν μετά την εκπαίδευση του μοντέλου ANFIS (Σχήμα 6.8 με τίτλο "final MFs").

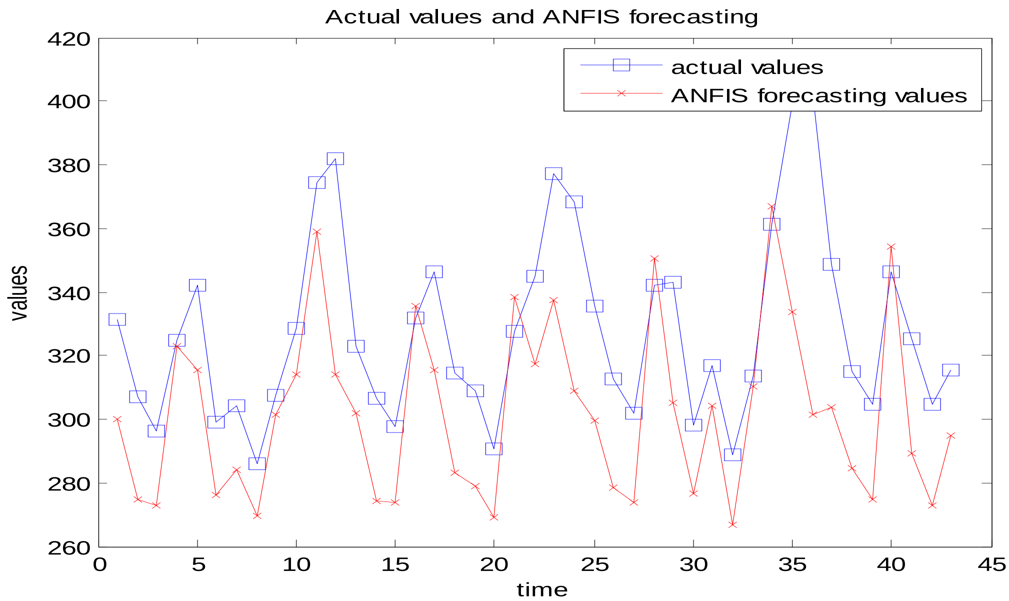


<u>Σχήμα 6.7:</u> *Αρχικές συναρτήσεις συμμετοχής*.

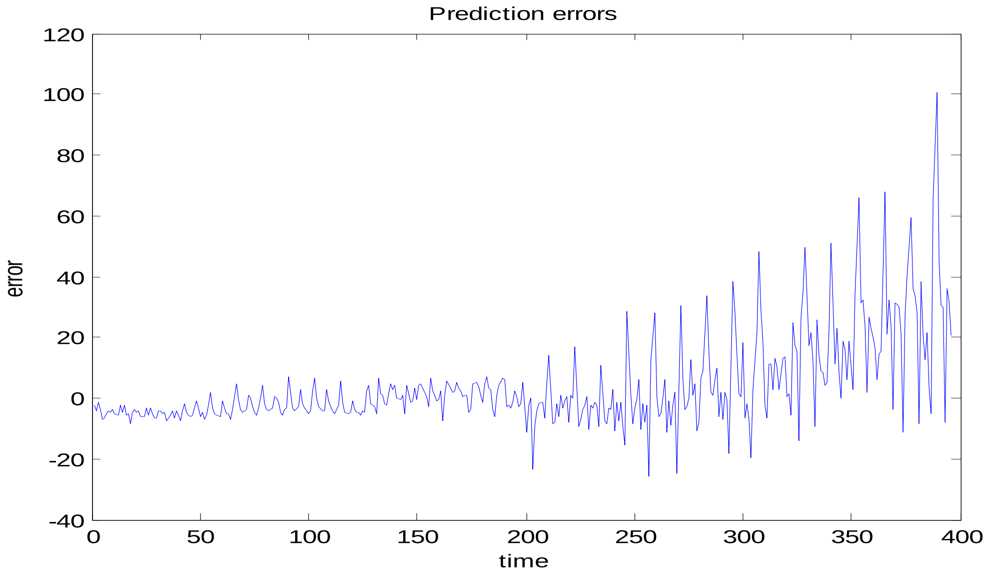


Σχήμα 6.8: *Τελικές συναρτήσεις συμμετοχής.*

Το διάγραμμα που ακολουθεί κατέχει σημαντικό ρόλο στην αξιολόγηση του ANFIS, καθώς μπορούμε να δούμε τη διαφορά μεταξύ ενός μέρους των πραγματικών και των προβλεπόμενων τιμών. Με μπλε χρώμα απεικονίζονται οι πραγματικές τιμές και με κόκκινο οι τιμές που προβλέπει το ANFIS.

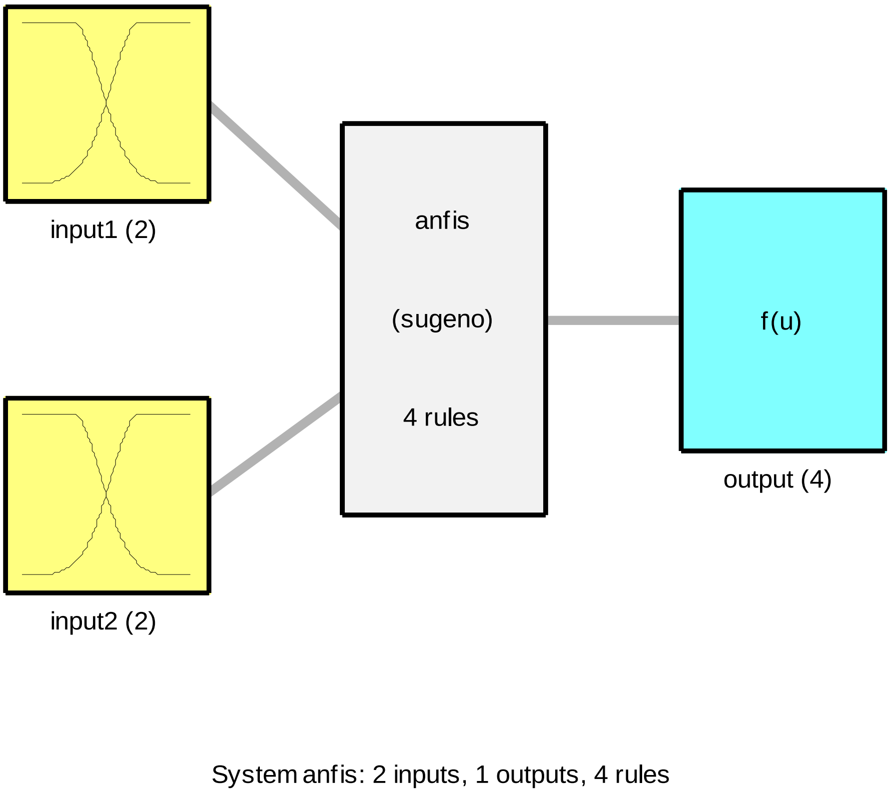


Σχήμα 6.9: *Π*ροβλέψεις του ANFIS έναντι των πραγματικών τιμών (42 τελευταίες τιμές).

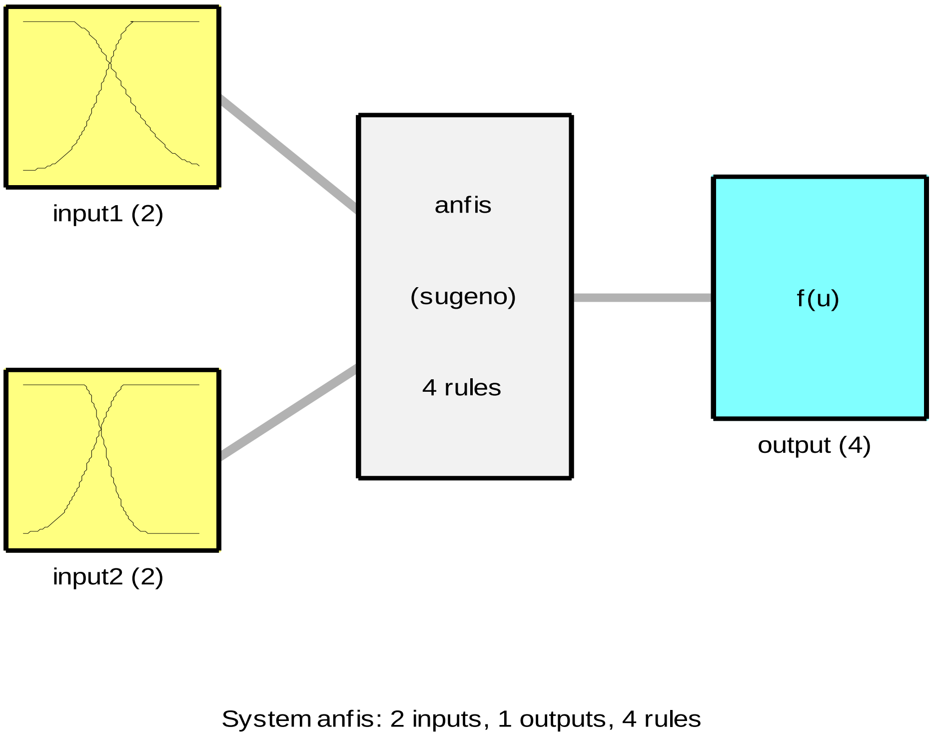


Σχήμα 6.10: *Διάγραμμα Σφάλματος Πρόβλεψης.*

Στο Σχήμα 6.11 παρουσιάζεται η αρχική δομή του ANFIS που χρησιμοποιήθηκε. Διακρίνονται οι δυο είσοδοι με τις αρχικές συναρτήσεις συμμετοχής, ο μηχανισμός συμπερασμού με τους τέσσερις κανόνες και η έξοδος του συστήματος.



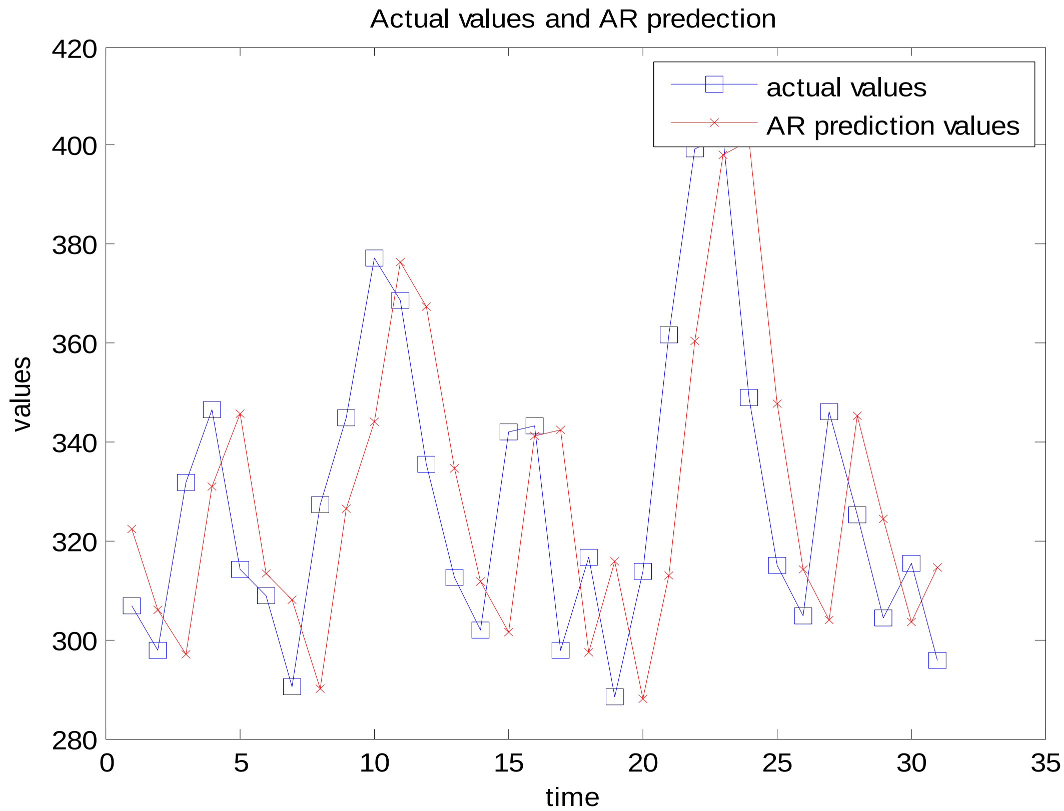
Σχήμα 6.11: *Παρουσίαση της αρχικής δομής του ANFIS.*



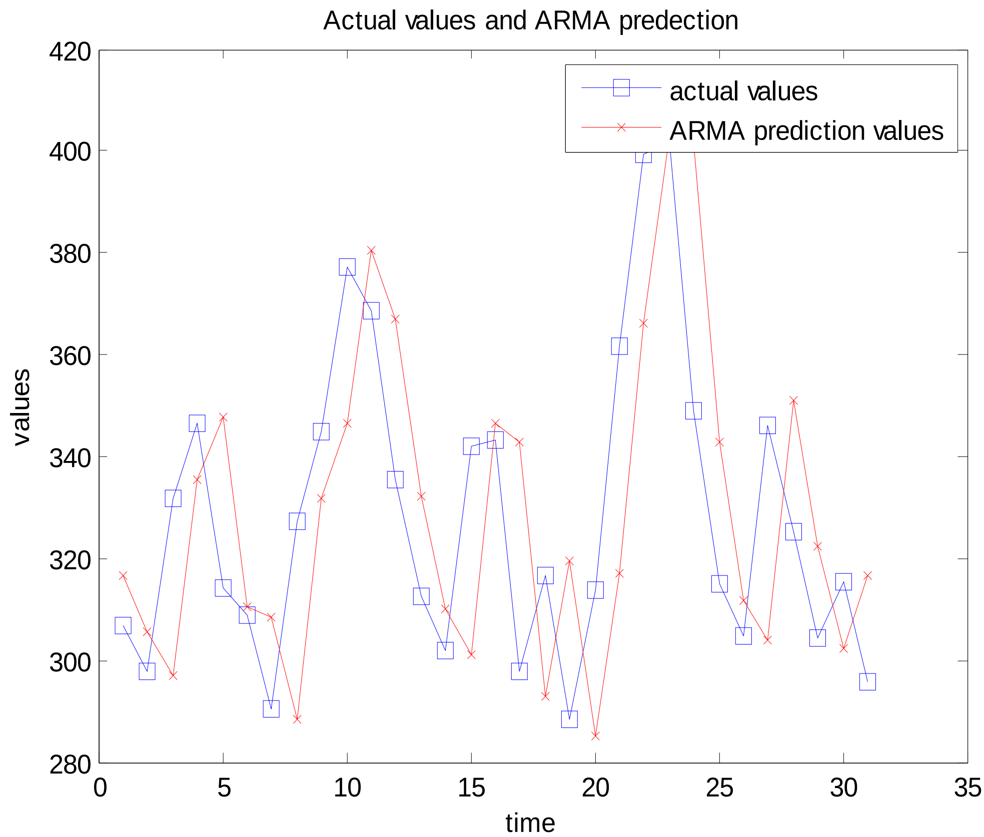
<u>Σχήμα 6.12:</u> *Παρουσίαση της τελικής δομής του ANFIS.*

Στο *Σχήμα 6.12*, έχουμε την τελική δομή του νεύρο-ασαφούς μας μοντέλου (ANFIS) με τις τελικές συναρτήσεις συμμετοχής που προέκυψαν από την διαδικασία της εκπαίδευσης.

Τα επόμενα δυο διαγράμματα κατέχουν σημαντικό ρόλο στην αξιολόγηση των AR και ARMA μοντέλων πρόβλεψης αφού σε καθένα απ' αυτά μπορούμε να δούμε τη διαφορά μεταξύ ενός μέρους των πραγματικών και των προβλεπόμενων τιμών. Καθώς και στην σύγκριση της απόδοσης των υπό μελέτη μοντέλων μας. Με μπλε χρώμα απεικονίζονται οι πραγματικές τιμές και με κόκκινο οι τιμές που προβλέπει το εκάστοτε μοντέλο.



Σχήμα 6.13: Π*ροβλέψεις του AR-μοντέλου έναντι των πραγματικών τιμών (30 τελευταίες τιμές).*



Σχήμα 6.14: *Π*ροβλέψεις του ARMA-μοντέλου έναντι των πραγματικών τιμών (30 τελευταίες τιμές).

Τέλος, στον επόμενο συγκεντρωτικό πίνακα έχουμε την παρουσίαση των σφαλμάτων για το κάθε μοντέλο πρόβλεψης.

		ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ		
		ANFIS	**AR**	**ARMA**
ΤΥΠΟΣ ΣΦΑΛΜΑΤΟΣ	**MSE**	221,1242	460,9707	456,8153
	RMSE	14,8702	21,4702	21,3732
	MAE	8,9296	17,5168	17,3351
	MAPE	3,2439	7,1159	7,0816

Πίνακας 6.1: *Τύποι σφαλμάτων για τα χρησιμοποιούμενα μοντέλα πρόβλεψης (ANFIS, AR, ARMA).*

Σημασία σφαλμάτων:

Το **MSE (Mean Square Error)** είναι το μέσο τετραγωνικό σφάλμα δηλαδή παίρνουμε μια μέση τιμή του τετραγώνου του σφάλματος δηλαδή της διαφοράς της πραγματικής από τη προβλεπόμενη τιμή κατά μήκος των δυο δειγμάτων είτε είναι το δείγμα εκπαίδευσης είτε το δείγμα πρόβλεψης οπότε και παίρνουμε και τα αντίστοιχα μέσα τετραγωνικά σφάλματα.
Το **RMSE (Root Mean Square Error)** είναι η ρίζα του μέσου τετραγωνικού σφάλματος το οποίο αναλύσαμε προηγουμένως.
Το **MAE (Mean Absolute Error)** είναι το μέσο απόλυτο σφάλμα δηλαδή παίρνουμε τη μέση τιμή του απολύτου σφάλματος δηλαδή της διαφοράς της πραγματικής από τη προβλεπόμενη τιμή κατά μήκος των δυο δειγμάτων εκπαίδευσης και πρόβλεψης και παίρνουμε τα αντίστοιχα μέσα απόλυτα σφάλματα εκπαίδευσης και πρόβλεψης.
Το **MAPE (Mean Absolute percentage Error)** μέσο ποσοστιαίο απόλυτο σφάλμα δηλαδή το μέσο απόλυτο σφάλμα εκφρασμένο σε ποσοστό επί τις εκατό.

Τέλος όσον αφορά το χρόνο που χρειάστηκε για να βγάλει αποτελέσματα το πρόγραμμα χρειαστήκαν 31.385834 seconds.

6.2.3 Συμπεράσματα Από Την Πρόβλεψη

Όπως αναφέραμε και στην εισαγωγή αυτού του συγγράμματος, η παροχή εγγυημένης ισχύος είναι ένα ζήτημα μείζονος σημασίας, σε κάθε Σύστημα Παραγωγής Ενέργειας. Η αποφυγή της αστάθειας, σε περιόδους αιχμής, κρίνεται επιτακτική. Σε τέτοιες περιπτώσεις απαιτούνται ειδικά μέτρα και κατάλληλη διαχείριση για την αποφυγή της κατάρρευσής του. Κρίνεται λοιπόν αναγκαία, η δημιουργία αποδοτικών μοντέλων πρόβλεψης που προσεγγίζουν τα μελλοντικά φορτία με ικανοποιητική ακρίβεια, έτσι ώστε η παραγόμενη ισχύς να καλύπτει τα φορτία μαζί με τις απώλειες του δικτύου.

Στην παρούσα εργασία ξεκίνησε με σκοπό την αξιολόγηση και την σύγκριση του Νεύρο-ασαφούς ANFIS, έναντι των συμβατικών μεθόδων πρόβλεψης (αυτοπαλινδρομικά μοντέλα (AR), αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου (autoregressive moving average ή ARMA models)).
Τα μοντέλα αυτά (AR, ARMA) μπορούν να χρησιμοποιηθούν μόνο για στάσιμες χρονοσειρές, πράγμα που τα κάνει να μειονεκτούν απέναντι σε ένα νεύρο-ασαφές μοντέλο. Επιπλέον, βάση του συγκεντρωτικού *πίνακα 6.1* , παρατηρούμε ότι το ANFIS παρουσίασε μικρότερη τιμή σφάλματος και στα τέσσερα είδη σφαλμάτων, έναντι των άλλων δύο μοντέλων πρόβλεψης. Πράγμα που το καθιστά αποτελεσματικότερο για την πρόβλεψη μας. Ακλουθεί το ARMA μοντέλο, όπου έχουμε και στις τέσσερις κατηγορίες σφάλματος σχεδόν διπλάσιες τιμές. Τελευταίο έρχεται το AR μοντέλο, με μικρές διαφορές σε όλους τους τομείς έναντι του προαναφερόμενου ARMA.

Τέλος, βάση των γραφημάτων προέκυψαν από την εφαρμογή του Προσαρμοστικού Νεύρο-Ασαφούς Συστήματος Συμπερασμού (Adaptive Neural Fuzzy Inference, ANFIS) στο λογισμικό της Matlab, βλέπουμε ότι τα αποτελέσματα είναι αρκετά ικανοποιητικά. Πιο συγκεκριμένα, βλέπουμε στην Δισδιάστατη απεικόνιση των δεδομένων εκπαίδευσης (*Σχήμα 6.4*) (sctter plot), ότι τα δεδομένα μας τοποθετούνται με μικρή απόκλιση από την ιδεατή γραμμή και σε συνδυασμό με το *Σχήμα 6.6* όπου βλέπουμε ότι το σφάλμα εκπαίδευσης διατηρείται σε χαμηλά επίπεδα.

Συνολικά λοιπόν, το ANFIS κρίνεται περισσότερο ακριβές σε σχέση με τα συγκρινόμενα μοντέλα, καθώς παρουσιάζει μικρότερο ποσοστό σφάλματος σε όλους τους τύπους σφαλμάτων.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] Otavio A.S. Carpinteiro [a,*], Agnaldo J.R. Reis [a], Alexandre P.A. da Silva [b] ´´ **A hierarchical neural model in short-term load forecasting**´´
[2] Ye Yao, Zhiwei Lian*, Zhijian Hou, Weiwei Liu ´´ **An innovative air-conditioning load forecasting model based on RBF neural network and combined residual error correction** ´´
[3] V.S. Kodogiannis *, E.M. Anagnostakis ´´ **A study of advanced learning algorithms for short-term load forecasting** ´´
[4] Marley Maria B.R. Vellasco, Marco Aurelio C. Pacheco, Luiz Sabino Ribeiro Neto, Flavio J. de Souza´´ **Electric load forecasting: evaluating the novel hierarchical neuro-fuzzy BSP model**´´
[5] P. K. DASH, A. C. LIEW, S. RAHMAN, S. DASH´´ **Fuzzy and Neuro-fuzzy Computing Models for Electric Load Forecasting**´´
[6] P. K. Dash[a], H. P. Satpathy[b] and A. C. Liew[a] ´´ **A real-time short-term peak and average load forecasting system using a self-organising fuzzy neural network**´´
[7] K. Kalaitzakis *, G.S. Stavrakakis, E.M. Anagnostakis ´´ **Short-term load forecasting based on artificial neural networks parallel implementation** ´´
[8] Mohammad Tamimi* , Robert Egbert´´**Short term electric load forecasting via fuzzy neural collaboration**´´
[9] Benjamin F. Hobbs!,*, Udi Helman!, Suradet Jitprapaikulsarn",
Sreenivas Konda", Dominic Maratukulam ´´**Artificial neural networks for short-term energy forecasting:**
Accuracy and economic value´´
[10] A.M. Sharaf [a], Tjing T. Lie [b]´´ **A novel neuro-fuzzy based self-correcting online electric load forecasting model**´´
[11] B. Satish [a,*], K.S. Swarup [b], S. Srinivas [c], A. Hanumantha Rao[a]
´´ **Effect of temperature on short term load forecasting using an integrated ANN** ´´
[12] Tawfiq Al-Saba, Ibrahim El-Amin* ´´**Artificial neural networks as applied to long-term demand forecasting**´´
[13] C.S. Chen [a,*], Y.M. Tzeng [a], J.C. Hwang [b] ´´ **The application of artificial neural networks to substation load forecasting** ´´
[14] M. Beccali, M. Cellura *, V. Lo Brano, A. Marvuglia ´´**Forecasting daily urban electric load profiles using artificial neural networks**´´
[15] Ot´avio A.S. Carpinteiro [a,*] , Rafael C. Leme[b,a], Antonio C. Zambroni de Souza [b,a],Carlos A.M. Pinheiro [a], Edmilson M. Moreira
´´ **Long-term load forecasting via a hierarchical neural model with time integrators** ´´

[16] Nahi Kandil [a], Rene´ Wamkeue [a,*], Maarouf Saad [b], Semaan Georges ´´ **An efficient approach for short term load forecasting using artificial neural networks** ´´
[17] Ajith Abraham*, Baikunth Nath´´**A neuro-fuzzy approach for modelling electricity demand in Victoria**´´
[18] Witold Bartkiewicz´´**NEURO-FUZZY APPROACHES TO SHORT-TERM ELECTRICAL LOAD FORECASTING**´´
[19] Paras Mandal [a],1, Tomonobu Senjyu [a,1], Naomitsu Urasaki [a,1], Toshihisa Funabashi [b] ´´**A neural network based several-hour-ahead electric load forecasting using similar days approach**´´
[20] Mara Lucia M. Lopes*, Carlos R. Minussi, Anna Diva P. Lotufo´´ **Electric load forecasting using a fuzzy ART&ARTMAP neural network**´´
[21] Eva Gonza´lez-Romera, Miguel A ´ ngel Jaramillo-Mora´n *, Diego Carmona-Ferna´ndez ´´ **Forecasting of the electric energy demand trend and monthly fluctuation with neural networks** ´´
[22] Pedro A. Gonza´lez, Jesu´s M. Zamarren˜o* ´´ **Prediction of hourly energy consumption in buildings based on a feedback artificial neural network**´´
[23] Bahman Kermanshasi, Kiroshi Iwamiya ´´ **Up to year 2020 load forecasting using neural nets** ´´
[24] D.Datta,S.A Tassou ´´ **Artificial neural networks based electrical load prediction for food retail stores** ´´
[25] CHIH-CHOU CHIU, LING-JING KAO´´ **Combining a Neural Network with a Rule-Based Expert System Approach for Short-Term Power Load Forecasting in Taiwan**´´
[26] Dipti Srinivasan* ´´**Evolving artificial neural networks for short term load Forecasting**´´
[27] Syed M. Islam, Saleh M. A1-Alawi, Khaled A. Ellithy ´´ **Forecasting monthly electric load and energy for a fast growing utility using an artificial neural network**
[28] Ayca Kumluca Topallia , Ismet Erkmenb ´´ **A hybrid learning for neural networks applied to short term load forecasting**´´
[29] E A Mohamed[a], M M Mansour[a], S EI-Debeiky[a] and K G Mohamed ´´ **Results of Egyptian unified grid hourly load forecasting using an artificial neural network with expert system interface** ´´
[30] George S. Atsalakis [a,*], Kimon P. Valavanis [b] ´´**Forecasting stock market short-term trends using a neuro-fuzzy based methodology**´´
[31] Atsalakis G., Ucenic C. (2005c) ´´**Forecasting of Electricity Demand Using a Neuro-fuzzy model**´´, International Conference: New Horizons In Industry, Business and Education, 25-26 August 2005, Corfu, Greece

[32] George Atsalakis, Dimitris Nezis, and Constantinos Zopounidis **´´Neuro-Fuzzy Versus Traditional Models for Forecasting Wind Energy Production´´**
[33] George S. Atsalakis[1], Christos H. Skiadas[1], Ilias Braimis[1] **´´Probability of trend prediction of exchange rate by ANFIS´´**
[34] Σταματάκης Παναγιώτης: **ΠΡΟΒΛΕΨΗ ΚΑΤΑΝΑΛΩΣΗΣ ΗΛΕΚΤΡΙΚΟΥ ΦΟΡΤΙΟΥ ΜΕ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΚΑΙ ΓΕΝΕΤΙΚΟΥΣ ΑΛΓΟΡΙΘΜΟΥΣ**, διπλωματική εργασία(2008).
[35] Εμμανουήλ Μ. Βουμβουλάκης: **ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ ΦΟΡΤΙΟΥ ΜΕ ΧΡΗΣΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΚΑΙ ΑΣΑΦΟΥΣ ΛΟΓΙΚΗΣ**, διπλωματική εργασία(2003).

[36] Ευθύμιος Ι. Νικολάου: « **ΣΥΓΚΡΙΤΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΕΦΑΡΜΟΓΗ ΓΡΑΜΜΙΚΩΝ, ΜΗ-ΓΡΑΜΜΙΚΩΝ ΚΑΙ ΝΕΥΡΟ-ΑΣΑΦΩΝ ΜΕΘΟΔΩΝ, ΓΙΑ ΤΗ ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ ΠΑΡΑΓΩΓΗΣ ΕΝΕΡΓΕΙΑΣ ΑΠΟ ΑΙΟΛΙΚΑ ΠΑΡΚΑ** », διπλωματική εργασία.
[37] Ατσαλάκης Γιώργος: **Συνοπτικές σημειώσεις για την λύση ασκήσεων Τεχνολογικής Πρόβλεψης**, πανεπιστημιακές σημειώσεις(2006)
[38] Neural network Toolbox, Users Guide, Mathworks INC
[39] Fuzzy Logic Toolbox, Users Guide, Mathworks INC

ΠΑΡΑΡΤΗΜΑ 1

Δεδομένα

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1973	1	160.217.989
1973	2	143.538.699
1973	3	148.158.370
1973	4	139.589.440
1973	5	147.395.124
1973	6	161.243.674
1973	7	173.733.007
1973	8	177.365.364
1973	9	156.875.032
1973	10	154.197.466
1973	11	148.137.609
1973	12	153.604.858
1974	1	157.554.691
1974	2	142.748.458
1974	3	150.342.206
1974	4	142.311.633
1974	5	153.812.873
1974	6	156.440.089
1974	7	178.247.304
1974	8	174.119.001
1974	9	152.467.442
1974	10	152.195.832
1974	11	150.070.855
1974	12	160.009.019
1975	1	164.623.080
1975	2	147.349.000
1975	3	155.760.211
1975	4	146.494.879
1975	5	153.531.080
1975	6	162.716.960

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1975	7	177.056.505
1975	8	179.930.892
1975	9	155.440.826
1975	10	155.187.742
1975	11	153.035.097
1975	12	169.628.296
1976	1	178.608.612
1976	2	156.965.799
1976	3	164.466.676
1976	4	153.467.392
1976	5	157.664.467
1976	6	173.673.719
1976	7	186.691.040
1976	8	186.638.696
1976	9	165.237.437
1976	10	164.009.445
1976	11	169.346.291
1976	12	184.144.107
1977	1	196.665.182
1977	2	162.949.088
1977	3	169.437.171
1977	4	157.117.355
1977	5	169.596.403
1977	6	181.031.212
1977	7	199.167.726
1977	8	196.363.166
1977	9	176.497.560
1977	10	166.645.282
1977	11	167.388.072
1977	12	184.589.270

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1978	1	198.107.635
1978	2	173.745.731
1978	3	173.460.887
1978	4	160.012.982
1978	5	175.548.628
1978	6	188.585.386
1978	7	202.947.293
1978	8	206.658.749
1978	9	185.802.058
1978	10	176.013.116
1978	11	176.389.093
1978	12	192.105.353
1979	1	209.986.749
1979	2	186.587.309
1979	3	183.154.004
1979	4	170.260.472
1979	5	178.409.356
1979	6	186.976.379
1979	7	202.522.477
1979	8	205.100.592
1979	9	180.974.646
1979	10	179.953.363
1979	11	177.769.508
1979	12	188.970.171
1980	1	200.296.330
1980	2	188.961.379
1980	3	187.744.800
1980	4	169.016.585
1980	5	176.066.486
1980	6	189.748.486

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1980	7	217.057.853
1980	8	215.629.382
1980	9	191.697.740
1980	10	178.761.034
1980	11	178.771.361
1980	12	195.848.927
1981	1	206.758.205
1981	2	179.859.910
1981	3	185.833.882
1981	4	172.841.342
1981	5	178.138.553
1981	6	203.020.751
1981	7	220.654.747
1981	8	210.639.415
1981	9	187.051.489
1981	10	181.557.590
1981	11	175.791.753
1981	12	195.825.699
1982	1	209.694.373
1982	2	180.545.573
1982	3	187.967.965
1982	4	172.876.502
1982	5	177.479.981
1982	6	186.447.119
1982	7	210.865.079
1982	8	205.891.676
1982	9	180.875.268
1982	10	173.172.248
1982	11	173.598.759
1982	12	184.957.943

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1983	1	195.870.844
1983	2	172.725.190
1983	3	182.768.722
1983	4	170.668.876
1983	5	174.724.813
1983	6	191.367.206
1983	7	220.446.836
1983	8	230.193.460
1983	9	195.817.014
1983	10	183.136.711
1983	11	183.170.681
1983	12	212.555.333
1984	1	216.923.729
1984	2	189.810.043
1984	3	200.387.475
1984	4	181.381.139
1984	5	192.550.198
1984	6	209.967.314
1984	7	221.526.162
1984	8	229.531.666
1984	9	195.411.388
1984	10	191.142.430
1984	11	190.601.865
1984	12	200.231.958
1985	1	228.147.888
1985	2	198.488.497
1985	3	195.250.388
1985	4	185.173.312
1985	5	197.122.843
1985	6	205.681.971

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1985	7	227.003.762
1985	8	226.286.447
1985	9	202.712.422
1985	10	194.995.216
1985	11	192.648.952
1985	12	219.490.423
1986	1	217.761.136
1986	2	192.582.016
1986	3	197.114.555
1986	4	186.370.454
1986	5	197.647.498
1986	6	215.334.156
1986	7	242.953.552
1986	8	225.402.481
1986	9	206.905.101
1986	10	197.959.776
1986	11	196.653.686
1986	12	213.786.542
1987	1	223.040.975
1987	2	194.280.625
1987	3	202.129.944
1987	4	189.792.417
1987	5	206.406.878
1987	6	225.907.915
1987	7	248.196.221
1987	8	247.880.760
1987	9	213.221.249
1987	10	203.214.924
1987	11	200.479.386
1987	12	220.736.372

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1988	1	238.188.434
1988	2	217.183.084
1988	3	214.293.615
1988	4	196.297.111
1988	5	208.703.604
1988	6	233.065.853
1988	7	257.742.125
1988	8	267.928.593
1988	9	220.392.242
1988	10	210.813.885
1988	11	209.814.992
1988	12	232.987.639
1989	1	246.781.617
1989	2	233.800.651
1989	3	241.963.196
1989	4	222.880.229
1989	5	234.834.779
1989	6	250.931.790
1989	7	273.469.901
1989	8	275.430.580
1989	9	242.803.941
1989	10	235.677.339
1989	11	234.261.233
1989	12	274.470.268
1990	1	255.202.858
1990	2	229.512.112
1990	3	244.775.474
1990	4	229.783.675
1990	5	241.786.514
1990	6	269.010.617

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1990	7	287.465.399
1990	8	290.670.233
1990	9	258.367.608
1990	10	244.382.150
1990	11	231.250.933
1990	12	255.780.702
1991	1	269.214.313
1991	2	228.290.131
1991	3	240.561.216
1991	4	227.798.971
1991	5	254.873.018
1991	6	269.099.254
1991	7	294.647.834
1991	8	292.158.580
1991	9	256.027.328
1991	10	245.378.242
1991	11	241.661.365
1991	12	254.088.633
1992	1	267.772.670
1992	2	239.514.031
1992	3	247.733.163
1992	4	233.405.563
1992	5	242.411.928
1992	6	261.076.926
1992	7	293.616.745
1992	8	281.927.430
1992	9	259.925.035
1992	10	244.993.925
1992	11	244.161.999
1992	12	267.342.791

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1993	1	271.020.902
1993	2	248.015.328
1993	3	261.247.551
1993	4	234.695.127
1993	5	244.325.583
1993	6	275.359.600
1993	7	312.224.835
1993	8	311.450.190
1993	9	264.032.481
1993	10	250.552.792
1993	11	252.102.138
1993	12	272.164.570
1994	1	289.768.406
1994	2	249.172.290
1994	3	257.998.052
1994	4	240.636.581
1994	5	252.744.936
1994	6	294.161.782
1994	7	311.256.823
1994	8	307.604.584
1994	9	266.262.327
1994	10	256.528.490
1994	11	251.868.348
1994	12	269.519.769
1995	1	279.773.101
1995	2	252.307.149
1995	3	261.342.562
1995	4	244.735.923
1995	5	264.288.242
1995	6	286.258.135

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1995	7	330.416.095
1995	8	345.780.079
1995	9	277.575.357
1995	10	263.978.440
1995	11	261.478.399
1995	12	285.553.882
1996	1	296.922.971
1996	2	270.685.247
1996	3	275.019.018
1996	4	251.612.875
1996	5	282.266.414
1996	6	302.716.899
1996	7	327.708.462
1996	8	329.286.330
1996	9	283.151.263
1996	10	271.446.494
1996	11	269.122.035
1996	12	284.249.613
1997	1	300.573.901
1997	2	258.131.029
1997	3	272.258.227
1997	4	258.284.482
1997	5	272.913.783
1997	6	299.091.561
1997	7	344.516.455
1997	8	332.898.630
1997	9	301.056.672
1997	10	284.970.629
1997	11	271.120.816
1997	12	296.356.097

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
1998	1	295.260.404
1998	2	260.589.918
1998	3	286.877.989
1998	4	261.229.604
1998	5	299.639.908
1998	6	328.903.002
1998	7	361.936.460
1998	8	357.365.610
1998	9	318.924.466
1998	10	284.445.547
1998	11	267.136.408
1998	12	297.986.181
1999	1	318.107.610
1999	2	276.826.953
1999	3	300.353.589
1999	4	283.411.496
1999	5	298.907.872
1999	6	327.821.999
1999	7	375.426.401
1999	8	362.896.031
1999	9	304.374.744
1999	10	282.751.814
1999	11	269.973.431
1999	12	293.957.868
2000	1	327.993.944
2000	2	294.169.282
2000	3	301.579.821
2000	4	285.577.749
2000	5	322.953.681
2000	6	339.053.660

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
2000	7	356.527.917
2000	8	368.669.365
2000	9	312.446.647
2000	10	289.451.903
2000	11	284.670.275
2000	12	319.010.798
2001	1	332.493.160
2001	2	282.940.198
2001	3	300.706.544
2001	4	278.078.871
2001	5	300.491.621
2001	6	327.693.978
2001	7	357.613.700
2001	8	370.532.828
2001	9	306.928.866
2001	10	294.733.613
2001	11	278.933.942
2001	12	305.496.328
2002	1	319.941.472
2002	2	281.825.708
2002	3	302.549.007
2002	4	289.848.252
2002	5	307.674.569
2002	6	341.023.145
2002	7	381.542.109
2002	8	374.585.823
2002	9	331.279.434
2002	10	307.059.462
2002	11	296.289.561
2002	12	324.833.704

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
2003	1	341.988.899
2003	2	299.248.775
2003	3	304.317.263
2003	4	285.755.700
2003	5	307.544.887
2003	6	328.693.699
2003	7	374.396.049
2003	8	381.816.318
2003	9	323.135.536
2003	10	306.740.934
2003	11	297.866.713
2003	12	331.680.431
2004	1	346.545.630
2004	2	314.279.911
2004	3	308.812.097
2004	4	290.559.918
2004	5	327.380.267
2004	6	345.085.080
2004	7	377.331.700
2004	8	368.439.042
2004	9	335.622.199
2004	10	312.450.226
2004	11	302.101.254
2004	12	341.947.965
2005	1	343.228.971
2005	2	297.939.518
2005	3	316.779.985
2005	4	288.565.631
2005	5	313.773.104
2005	6	361.471.996

ΕΤΟΣ	ΜΗΝΑΣ	ΦΟΡΤΙΟ εκ Kwh
2005	7	399.252.240
2005	8	401.977.976
2005	9	348.812.107
2005	10	315.034.251
2005	11	304.899.013
2005	12	346.253.918
2006	1	325.245.767
2006	2	304.455.783
2006	3	316.239.337
2006	4	295.569.818

ΠΑΡΑΡΤΗΜΑ 2

Περίληψη Βιβλιογραφίας

A hierarchical neural model in short-term load forecasting

Otavio A.S. Carpinteiro [a], Agnaldo J.R. Reis [a], Alexandre P.A. da Silva [b]

This paper proposes a novel neural model to the problem of short-term load forecasting (STLF). It was required to predict once every hour the electric load during the next 24 h. The model has a topology made up of two SOM networks, one on top of the other. It encodes and manipulates context information effectively. Some conclusions may be drawn from the experiments. First, the knowledge representation proposed for the HNM inputs seems to be adequate. It supplied the model with the necessary information to make it produce correct predictions. Second, the HNM performance on the forecasts was much better than that of the MLP. The results obtained have shown that the HNM was able to perform efficiently the prediction of the electric load in short forecasting horizons. Third, it is worth mentioning that MLP has been widely employed to tackle the problem of STLF so far. The results obtained thus suggest that HNM may offer a better alternative to approach such problem. A research and development project for a Brazilian electric utility is under course. The research will focus on the effects of the HNM time integrators on the predictions in order to produce a better adaptability. Besides, it will focus on the study of its performance on larger load databases. The forecasts should also span a larger number of days in order to be more significant statistically.

An innovative air-conditioning load forecasting model based on RBF neural network and combined residual error correction

Ye Yao, Zhiwei Lian, Zhijian Hou, Weiwei Liu

Accurate air-conditioning load forecasting is the precondition for the optimal control and energy saving operation of HVAC systems. They have developed many forecasting methods, such as multiple linear regression (MLR), autoregressive integrated moving average (ARIMA), grey model (GM) and artificial neural network (ANN), in the field of air-conditioning load prediction. However, none of them has enough accuracy to satisfy the practical demand. On the basis of these models existed, a novel forecasting method, called 'RBF neural network (RBFNN) with combined residual error correction', is developed in this paper. The new model adopts the advanced algorithm of neural network based on radial basis functions for the air-conditioning load forecasting, and uses the combined forecasting model, which is the combination of MLR, ARIMA and GM, to estimate the residual errors and correct the ultimate foresting results. A study case indicates that RBFNN with combined residual error correction has a much better forecasting accuracy than RBFNN itself and RBFNN with single-model correction.

A study of advanced learning algorithms for short-term load forecasting

V.S. Kodogiannis *, E.M. Anagnostakis

Neural networks are currently finding practical applications, ranging from `soft' regulatory control in consumer products, to the accurate modelling of nonlinear systems. This paper presents the development of improved neural-network-based short-term electric load forecasting models for the power system of the Greek island of Crete. Several approaches, including radial basis function networks, dynamic neural networks and fuzzy-neural-type networks, have been proposed, and are discussed in this paper. Their performances are evaluated through a simulation study, using metered data provided by the Greek Public Power Corporation. The results indicate that the load-forecasting models developed in this way provide more accurate forecasts, compared with conventional back-propagation network forecasting models. Finally, the embedding of the new model capability in a modular forecasting system is presented.

Electric load forecasting: evaluating the novel hierarchical neuro-fuzzy BSP model

Marley Maria B.R. Vellasco, Marco Aurelio C. Pacheco, Luiz Sabino Ribeiro Neto, Flavio J. de Souza

This paper introduces a new hybrid neuro-fuzzy model, called HNFB, and evaluates its performance in short-term load forecasting. To this end, two Brazilian electric power companies were used as case studies. A total of three intelligent forecasting systems were tested and compared: neural networks, neuro-fuzzy, and neural/neuro-fuzzy systems. As input variables, the experiments made use of historical load series and of additional variables that influence the load behavior, such as the temperature, the comfort index and the consumption profile. The results reveal the potential of the proposed neuro-fuzzy and neural/neuro-fuzzy models for load forecasting, when compared with neural networks; the mean absolute percentage errors varied between 0.44 and 1.95%, depending on the case study at hand.

Fuzzy and Neuro-fuzzy Computing Models for Electric Load Forecasting

P. K. DASH, A. C. LIEW, S. RAHMAN, S. DASH

Two new computing models, namely a fuzzy expert system and a hybrid neural network-fuzzy expert system for time series forecasting of electric load, are presented in this paper. The fuzzy-logic-based expert system utilizes the historical relationship between load and dry-bulb temperature, and predicts electric loads fairly accurately, 1-24 h ahead. In the case of the hybrid neural network-fuzzy expert system, the input vector consists of the membership values of the linguistic properties of the past load and weather parameters, and the output vector is defined in terms of the fuzzy class membership values of the forecasted load. Linear and nonlinear adaptive fuzzy correction schemes are then used to augment the output from the neuro-fuzzy computing model to yield the final forecast. Extensive studies have been performed for load prediction for all seasons, and the results for a typical winter day are given to confirm the effectiveness of these models.

A real-time short-term peak and average load forecasting system using a self-organising fuzzy neural network

P. K. Dash[a], H. P. Satpathy[b] and A. C. Liew[a]

This paper presents a self organising fuzzy-neural-network-based short-term electric load forecasting system for real-time implementation. A learning algorithm is devised for updating the connecting weights as well as the structure of the membership function of the network. The number of rules in the inferencing layer is optimised; this in turn optimises the network structure. The proposed algorithm exploits the notion of error back-propagation. The network is initialised with random weights. Experimental results of the system are discussed from a practical standpoint. The system accounts for seasonal and daily characteristics, as well as abnormal conditions, holidays and other conditions. It is capable of forecasting load with a lead time of one day to one week. The adaptive mechanism is used to train the network on-line. The results indicate that the proposed load forecasting system is robust and yields more accurate forecasts. Furthermore, it allows greater adaptability to sudden changes, compared with simple neural-network or statistical approaches. Extensive studies have been performed for all seasons, and some of them are presented in this paper. The new algorithm is tested with a typical load date of the Virginia Utility, and produces a very robust and accurate forecast, with a Mean Absolute of Percentage of Error (MAPE) mostly less than 1.8% for 24-hours-ahead peak load forecast, and 1.6% for a 168-hours-ahead forecast.

Short-term load forecasting based on artificial neural networks parallel implementation

K. Kalaitzakis *, G.S. Stavrakakis, E.M. Anagnostakis

This paper presents the development and application of advanced neural networks to face successfully the problem of the short-term electric load forecasting. Several approaches including Gaussian encoding backpropagation (BP), window random activation, radial basis function networks, real-time recurrent neural networks and their innovative variations are proposed, compared and discussed in this paper. The performance of each presented structure is evaluated by means of an extensive simulation study, using actual hourly load data from the power system of the island of Crete, in Greece. The forecasting error statistical results, corresponding to the minimum and maximum load time-series, indicate that the load forecasting models proposed here provide significantly more accurate forecasts, compared to conventional autoregressive and BP forecasting models. Finally, a parallel processing approach for 24 h ahead forecasting is proposed and applied. According to this procedure, the requested load for each specific hour is forecasted, not only using the load time-series for this specific hour from the previous days, but also using the forecasted load data of the closer previous time steps for the same day. Thus, acceptable accuracy load predictions are obtained without the need of weather data that increase the system complexity, storage requirement and cost.

Short term electric load forecasting via fuzzy neural collaboration

Mohammad Tamimi , Robert Egbert

An important element of effective power system operation is the well-planned short term scheduling of power generating units. Power system operators use historical load data to schedule available generating units to meet hourly system loads in an economical and reliable manner. This paper describes how a Fuzzy Logic (FL) expert system is integrated with Artificial Neural Networks (ANN) for a more accurate short-term load forecast. The 24 h ahead forecasted load is obtained through two steps. First, a FL module maps the highly nonlinear relationship between the weather parameters and their impact on the daily electric load peak. Second, 12 ANN modules are trained using historical hourly load and weather data combined with the FL output
data, to perform the final forecast. Comparisons made between this model, an ANN model, and an Autoregressive Moving Average (ARMA) model show the efficiency and accuracy of this new approach.

Artificial neural networks for short-term energy forecasting: Accuracy and economic value

Benjamin F. Hobbs, Udi Helman, Suradet Jitprapaikulsarn, Sreenivas Konda, Dominic Maratukulam

ANN technology for short-term energy forecasting holds the promise of improved accuracy and better operating decisions. We have documented significant reductions in forecast error relative to other systems presently used by 18 electric utilities and five gas utilities. Based on responses to a questionnaire, we estimate that the accuracy and convenience of an ANN-based forecasting system results in an average savings of about $800 000 yr electric utility for those utilities that reported that savings occurred.
The reasonableness of this estimate is confirmed by simulations of generation scheduling decisions under varying levels of accuracy. Future research should consider benefits of better forecasts that flow from improving other types of short-term commitments that are increasingly important in restructured power markets.

A novel neuro-fuzzy based self-correcting online electric load forecasting model

A.M. Sharaf a, Tjing T. Lie b

The paper presents a novel online self-adjusting short-term electric load forecast model with preprocessor data scaling/normalization, a fixed-structure feedforward two-layer stationary neural network, and a novel fuzzy logic postprocessor correction/filtering block. The fuzzy logic error-driven, error-adjusted postprocessor serves as a final online tuning mechanism, using actual accumulated forecasted/accurate data to tune the next forecast level. The concept is similar to a smoothing possibilistic FIR filter structure. The combined neural-fuzzy structure can be viewed as an elastic intelligent neurofuzzy, self-correcting, self-adaptive accurate forecast regression model. The postprocessor can generate possibilistic error limits that allow the operator to predict the forecast possibilistic confidence limits.

Effect of temperature on short term load forecasting using an integrated ANN

B. Satish, K.S. Swarup, S. Srinivas, A. Hanumantha Ra

An integrated Artificial Neural Network (ANN) approach to Short-Term Load Forecasting (STLF) is proposed in this paper. Four modules consisting of the Basic ANN, Peak and Valley ANN, Averager and Forecaster and Adaptive Combiner form the integrated method for load forecasting. The Basic ANN uses the historical data of load and temperature to predict the next 24 h load, while the Peak and Valley ANN uses the past peak and valley data of load and temperatures, respectively. The Averager captures the average variation of the load from the previous load behaviour, while the adaptive combiner uses the weighted combination of outputs from the Basic ANN and the Forecaster, to forecast the final load. The regression based and time series methods are conceptually incorporated into the ANN to obtain an integrated load forecasting approach.

Artificial neural networks as applied to long-term demand forecasting

Tawfiq Al-Saba, Ibrahim El-Amin*

This paper reports on the application of Artificial Neural Networks (ANN) to long-term load forecasting. The ANN model is used to forecast the energy requirements of an electric utility. It is then compared to time series models. The comparison reveals that the ANN produces results that are close to the actual data. The ANN model is then used to forecast the annual peak demand of a Middle Eastern utility up to the year 2006. The results compare favorably with the utility's forecast.

The application of artificial neural networks to substation load forecasting

C.S. Chen, Y.M. Tzeng, J.C. Hwang

The substation loading is highly correlated with the customers served. The substations in a distribution system can be categorized as residential, commercial and industrial. Each type has a different power consumption pattern. The substation loading will be varied according to the combination of the above three types of customers. In this paper, a supervisory functional artificial neural network (ANN) technique is applied to solve the load forecasting of three Taipower substations which serve the different customer types. The load forecasting accuracy is enhanced by considering the temperature effect on the substation load demand. With the converged ANN models derived by a training procedure, the temperature sensitivity of the substation load demand is easily obtained by the recall process. It is suggested that the substation load forecasting can be performed efficiently by the proposed method to support distribution operation effectively.

Forecasting daily urban electric load profiles using artificial neural networks

M. Beccali, M. Cellura *, V. Lo Brano, A. Marvuglia

The paper illustrates a combined approach based on unsupervised and supervised neural networks for the electric energy demand forecasting of a suburban area with a prediction time of 24 h. A preventive classification of the historical load data is performed during the unsupervised stage by means of a Kohonens self organizing map (SOM). The actual forecast is obtained using a two layered feed forward neural network, trained with the back propagation with momentum learning algorithm. In

order to investigate the influence of climate variability on the electricity consumption, the neural network is trained using weather data (temperature, relative humidity, global solar radiation) along with historical load data available for a part of the electric grid of the town of Palermo (Italy) from 2001 to 2003. The model validation is performed by comparing model predictions with load data that were not used for the network_s training. The results obtained bear out the suitability of the adopted methodology for the short term load forecasting (STLF) problem also at so small a spatial scale as the suburban one.

Long-term load forecasting via a hierarchical neural model with time integrators

Otavio A.S. Carpinteiro [a,∗], Rafael C. Leme[b,a], Antonio C. Zambroni de Souza [b,a], Carlos A.M. Pinheiro [a], Edmilson M. Moreira

The paper presents a novel artificial neural model to the problem of long-term load forecasting. The model has a topology made up of two self-organizing map networks – one on top of the other –, and a single-layer perceptron. It encodes and manipulates context information effectively. The novel hierarchical model is compared to a multilayer perceptron. Both are trained and assessed on load data extracted from a North-American electric utility. No pre-processing is made on data apart from ordinary normalization. The experiments show that the performance of the hierarchical model on long-term peak-load forecasts is much better than that of the multilayer perceptron. The superior performance displayed by the model seems to be justified by its hierarchical topology. The experiments also show that the performance of the hierarchical model on long-term mean-load forecasts is relatively better than that of the multilayer perceptron. The reasonable performance achieved by the multilayer perceptron seems to be justified by the smoothness of the mean-load curves. The results achieved by the hierarchical model still have space for improvements. Fine adjustments on its parameters – map sizes, radius of the Gaussian, and radius of the neighbourhood – as well as the usage of pre-processing techniques on load data will certainly lead to improvements. Finally, a comparison of both models endowed with time windows is under consideration. The hierarchical model has never been tested with time windows. It is worth verifying then its behaviour with such mechanism. In its turn, time-windowed multilayer perceptrons have been widely employed to tackle the problem of short-term load forecasting so far. Thus, it is worth analysing its behaviour on long-term horizons as well.

An efficient approach for short term load forecasting using artificial neural networks

Nahi Kandil, Rene Wamkeue, Maarouf Saad, Semaan Georges

In previous work, we applied artificial neural networks (ANN) for short term load forecasting using real load and weather data from the Hydro-Quebec databases where three types of variables were used as inputs to the neural network: (a) hour and day indicators, (b) weather related inputs and (c) historical loads. In general, for forecasting with a lead time of up to a few days ahead, load history (for the last few days) is not available, and therefore, estimated values of this load are used instead. However, a small error in these estimated values may grow up dramatically and lead to a serious problem in load forecasting since this error is fed back as an input to the forecasting procedure. In this paper, we demonstrate ANN capabilities in load forecasting without the use of load history as an input. In addition, only temperature (from weather variables) is used, in this application, where results show that other variables like sky condition (cloud cover) and wind velocity have no serious effect and may not be considered in the load forecasting procedure.

A neuro-fuzzy approach for modelling electricity demand in Victoria

Ajith Abraham*, Baikunth Nath

This paper evaluates the use of two popular soft computing techniques and conventional statistical approach based on Box–Jenkins autoregressive integrated moving average (ARIMA) model to predict electricity demand in the State of Victoria, Australia. The soft computing methods considered are an evolving fuzzy neural network (EFuNN) and an artificial neural network (ANN) trained using scaled conjugate gradient algorithm (CGA) and backpropagation (BP) algorithm. The forecast accuracy is compared with the forecasts used by Victorian Power Exchange (VPX) and the actual energy demand. To evaluate, we considered load demand patterns for 10 consecutive months taken every 30 min for training the different prediction models. Test results show that the neuro-fuzzy system performed better than neural networks, ARIMA model and the VPX forecasts.

NEURO-FUZZY APPROACHES TO SHORT-TERM ELECTRICAL LOAD FORECASTING

Witold Bartkiewicz

In this paper we investigate the application of the Takagi-Sugeno fuzzy models to short-term electrical load forecasting problem. Several learning algorithms for these type fuzzy systems are discussed. For identification of the models with linear antecedents has been applied the combination of the cluster estimation and ordinary least squares method. For nonlinear antecedents modelling purposes has been used the fuzzy switched ensemble of feedforward neural networks. The performance of the models is compared for two-day ahead peak load prediction in the distribution network.

A neural network based several-hour-ahead electric load forecasting using similar days approach

Paras Mandal a,1, Tomonobu Senjyu a,1, Naomitsu Urasaki a,1, Toshihisa Funabashi b,*

This paper presents a practical method for short-term load forecast problem using artificial neural network (ANN) combined similar days approach. Neural networks applied in traditional prediction methods all use similar days data to learn the trend of similarity. However, learning all similar days data is a complex task, and does not suit the training of neural network. A Euclidean norm with weighted factors is used to evaluate the similarity between the forecast day and searched previous days. According to similar days approach, load curve is forecasted by using information of the days that are similar to weather condition of the forecast day. An accuracy of the proposed method is enhanced by the addition of temperature as a major climate factor, and special attention was paid to model accurately in different seasons, i.e. Summer, Winter, Spring, and Autumn. The one-to-six hour-ahead forecast errors (MAPE) range from 0.98 to 2.43%. Maximum and minimum percentage errors, and MAPE values obtained from the load forecasting results confirm that ANN-based proposed method provides reliable forecasts for several-hour-ahead load forecasting.

Electric load forecasting using a fuzzy ART&ARTMAP neural network

Mara Lucia M. Lopes*, Carlos R. Minussi, Anna Diva P. Lotufo

This work presents a neural network based on the ART architecture (adaptive resonance theory), named fuzzy ART&ARTMAP neural network, applied to the electric load-forecasting problem. The neural networks based on the ART architecture have two fundamental characteristics that are extremely important for the network performance (stability and plasticity), which allow the implementation of continuous training. The fuzzy ART&ARTMAP neural network aims to reduce the imprecision of the forecasting results by a mechanism that separate the analog and binary data, processing them separately. Therefore, this represents a reduction on the processing time and improved quality of the results, when compared to the Back-Propagation neural network, and better to the classical forecasting techniques (ARIMA of Box and Jenkins methods). Finished the training, the fuzzy ART&ARTMAP neural network is capable to forecast electrical loads 24 h in advance. To validate the methodology, data from a Brazilian electric company is used.

Forecasting of the electric energy demand trend and monthly fluctuation with neural networks

Eva Gonzalez-Romera, Miguel Angel Jaramillo-Moran ,
Diego Carmona-Fernandez

Electric energy demand forecasting is a fundamental tool for production and distribution companies because it provides them with a prediction of the market demand of electric energy. Two kinds of forecasting may be performed: short term and medium to long term. This work is focused on monthly prediction, which is useful for the maintenance planning of grids and as market research for electricity producers and resellers. The timed series of monthly electric energy demands presents a rising tendency due to the influence of economic and technological evolution on the electric market. Embedded in this general trend is a fluctuation caused by the difference in demand from month to month. This paper proposes the extraction of that trend to perform separate predictions of both tendency and fluctuation with neural networks, which will be summed up to obtain the series forecasting. A Mean Absolute Percentage Error (MAPE) of about 2% has been obtained.

Prediction of hourly energy consumption in buildings based on a feedback artificial neural network

Pedro A. Gonzalez, Jesu´s M. Zamarreno

In this paper a new approach for short-term load prediction in buildings is shown. The method is based on a special kind of artificial neural network (ANN), which feeds back a part of its outputs. This ANN is trained by means of a hybrid algorithm. The new system uses current and forecasted values of temperature, the current load and the hour and the day as inputs. The performance of this predictor was evaluated using real data and results from international contests. The achieved results demonstrate the high precision reached with this system.

Up to year 2020 load forecasting using neural nets

Bahman Kermanshasi, Kiroshi Iwamiya

In this study, total system load forecast reflecting current and future trends is carried out for nine power companies in Japan. Two ANNs, a three-layered back-propagation and a recurrent neural network, were designed and tested for the purpose. Predictions were done for target years 1999, 2000, 2005, 2010, 2015, and 2020, respectively. Two case studies, *preservation of the status* and *structure reform*, were also tested for predicting the loads of years 2010 and 2020. Unlike short-term load forecasting, long-term load forecasting is mainly affected by economical factors rather than weather conditions. This study focuses on economical data that seem to influence long-term electric load demands. Here, 10 factors are selected as inputs for the proposed ANNs:

(1) gross national product, (2) gross domestic product, (3) population, (4) number of households, (5) number of air-conditioners, (6) amount of CO_2 pollution, (7) index of industrial production, (8) oil price, (9) energy consumption, and (10) electricity price. The data used are: actual yearly, incremental growth rate from the previous year, and both together (actual and incremental growth rate from the previous year). As a result, the demands for 2010 and 2020 are predicted to be 225.779 and 249.617 GW, respectively (preservation of the status). With structure reform, the demands for 2010 and 2020 are predicted to be 219.259 and 244.508 GW.

Artificial neural network based electrical load prediction for food retail stores

D. Datta*, S.A. Tassou

This paper presents the results from comparisons of the predictive accuracy of two commonly used neural networks employed for the prediction of the electrical load of a retail food store. The networks used were the multi-layered perceptron (MLP) and radial basis function (RBF). The MLP network was found to perform better than the RBF network particularly in the prediction of fluctuations of the electrical energy around the base and maximum loads. Further work will be carried out to optimize the structure and prediction accuracy of the two networks.

Combining a Neural Network with a Rule-Based Expert System Approach for Short-Term Power Load Forecasting in Taiwan

CHIH-CHOU CHIU AND LING-JING KAO, DEBORAH E COOK

A back-propagation neural network with the output provided by a rule-based expert system is designed for short-term power load forecasting. To demonstrate that the inclusion of the prediction from a rule-based expert system of a power system would improve the predictive capability of the network, load forecasting is performed on the Taiwan power system. The hourly load for one typical day was evaluated and, in that case, the inclusion of the rule-based expert system prediction significantly improved the neural network's prediction of power load. Moreover, the proposed combined approach converges much faster than the conventional neural network and the rule-based expert system method. Extensive studies were performed on the robustness of the built network model by using different specified censoring time. The prediction intervals of future power load series are also provided, to evaluate the prediction efficiency of the neural network model.

Evolving artificial neural networks for short term load Forecasting

Dipti Srinivasan

This paper presents artificial neural networks (ANN) evolved by a genetic algorithm for short-term load forecasting. Using real load values and forecast weather data these ANN have been tested for electric load forecasting on weekdays and weekends. For each day type, the best-evolved artificial neural network was found capable of accurately forecasting one-day ahead hourly loads. The forecasting results obtained using these best-evolved networks were observed to be consistently superior compared to a commonly used statistical method.

Forecasting monthly electric load and energy for a fast growing utility using an artificial neural network

Syed M. Islam, Saleh M. A1-Alawi, Khaled A. Ellithy

In this paper, novel artificial neural network (ANN) based weather-load and weather-energy models have been developed to forecast electric load and energy for 24 months ahead. A set of weather and other variables which have been identified for both models together with their correlations and contribution to the forecasted variable is reported. The proposed ANN models have been applied to historical energy, load, and weather data available for the Muscat power system from 1986 to 1990. Forecast results, when compared with the actual data for 1991-1992, show that monthly electric energy and load can be predicted within a maximum error of 6% and 10%, respectively, even with forecasted weather. The proposed ANN models provide better accuracy than previously developed models.

A hybrid learning for neural networks applied to short term load forecasting

Ayca Kumluca Topallia , Ismet Erkmenb

The aim of this work is to forecast total electric demand of Turkey one day in advance using neural networks. Starting withrandom weights and getting unrealistic forecasts could not be acceptable for a real-time operation; therefore, available past data, which is the actual load data obtained from Turkish Electricity Authority, is used o.-line and the model is prepared for on-line forecasts. Moreover, a method based on hourly electric consumption is proposed to cluster input data. In order to have an idea about the success of the model, several alternate models are formed. The proposed model shows considerably better results.

Egyptian Unified Grid hourly load forecasting using artificial neural network

E A Mohamed, M M Mansour, S EI-Debeiky and K G Mohamed

This paper presents an artificial neural network (ANN) based hourly load forecasting application to the Egyptian Unified Grid (EUG). The' ANN involved is designed using the multilayer back propagation learning technique. The ANN input layer receives all relevant information that can contribute extensively to ,the prediction process, excluding weather information. The input layer receives information on: the class of day type, the hour in day time, the load in hour-before, the load in day-before at same hour, the average load in day-before, the peak load in day-before, the minimum load in day-before, and also the latter four measurements but in the week before. Also, the ANN output layer provides the predicted hourly load. The ANN load forecasting model is trained based on a historical domain of knowledge. The required knowledge patterns were obtained for the EUG during the winter of 1993. When testing the trained ANN, it is proved that it can be applied to the prediction of hourly load very efficiently and accurately. Final results indicated average errors of 0.18% (training) and 0.49% (evaluation), and standard deviations of 2.32 and 2.92%, respectively.