Πολυτεχνειο Κρητης
Τμημα Μηχανικων Παραγωγης & Διοικησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΘΕΜΑ: ΠΡΟΒΛΕΨΗ ΚΑΤΑΝΑΛΩΣΗΣ ΗΛΕΚΤΡΙΚΟΥ ΦΟΡΤΙΟΥ ΜΕ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΚΑΙ ΓΕΝΕΤΙΚΟΥΣ ΑΛΓΟΡΙΘΜΟΥΣ

Η διπλωματική εργασία υπεβλήθη για την απόκτηση του διπλώματος του μηχανικού παραγωγής και διοίκησης από τον,
Σταματάκη Παναγιώτη

Η διπλωματική εργασία εγκρίνεται από τους :

κ. Ατσαλάκη
κ. Σκιαδά
κ. Μουστάκη

Χανιά 17-1-2008

ΠΕΡΙΕΧΟΜΕΝΑ

Πρόλογος……. 1
Περίληψη……. 2

Κεφάλαιο 1
Πρόβλεψη ηλεκτρικού φορτίου

Θεωρητικό υπόβαθρο……..... 3
Παράγοντες που επηρεάζουν το ηλεκτρικό φορτίο …… 7
α) οικονομικοί παράγοντες……..... 8
β) χρονικοί παράγοντες…….. 8
γ) καιρικοί παράγοντες……... 10
δ)τυχαίοι παράγοντες ……. 10

Κεφάλαιο 2
Νευρωνικά δίκτυα

Ιστορική αναδρομή …… 12
Από τα βιολογικά στα τεχνητά νευρωνικά δίκτυα …… 14
Τύποι συναρτήσεων ενεργοποίησης (μεταφοράς)…… 18
Αρχιτεκτονική νευρωνικών δικτύων…….. 21
1.Ενός-επιπέδου Εμπρός-Τροφοδότησης Δίκτυα…….. 23
2.Πολλαπλών-Επιπέδων Εμπρός-Τροφοδότησης Δίκτυα…….. 23
3.Αναδρομικά Δίκτυα……. 25
4.Δικτυωτές Δομές……. 27
Λειτουργία νευρωνικών δικτύων……. 29
Εκπαίδευση νευρωνικών δικτύων…… 30
Ο αλγόριθμος μάθησης του Perceptron (Αισθητήρα)…….. 32
Ο αλγόριθμος Ελάχιστου Μέσου Τετραγωνικού (ΕΜΤ) λάθους…… 36
Ο αλγόριθμος Πίσω Διάδοσης (Π.Δ.) του λάθους……... 37
Εφαρμογές των Ν.Δ……. 39
Πλεονεκτήματα των Νευρωνικών Δικτύων……. 40

Κεφάλαιο 3
Γενετικοί αλγόριθμοι

Εισαγωγή στους γενετικούς αλγορίθμους……... 42
Πλεονεκτήματα των Γενετικών Αλγορίθμων……...45
Τι Προκαλεί Δυσπιστία……47
Βασικά Χαρακτηριστικά Γενετικών Αλγορίθμων……49
Βασικά στοιχειά γενετικού αλγορίθμου…….. 51
Εφαρμογές……55

Κεφάλαιο 4
Νευρωνικά μοντέλα

Νευρωνικά δίκτυα και πρόβλεψη φορτιού……..56
1.Μακροπρόθεσμα τεχνητά νευρωνικά δίκτυα……...57
2.Υβριδικα νευρωνικά δίκτυα κοντινού χρονικού ορίζοντα…….. 58
3. Ιεραρχικά νευρωνικά δίκτυα κοντινού χρονικού ορίζοντα…… 59
4.Τεχνητά νευρωνικά δίκτυα που συνδυάζουν όμοιες ημερήσιες προσεγγίσεις …….. 60
5. MLP τεχνητά νευρωνικά δίκτυα που προβλέπουν κατανάλωση φορτίου και καυσίμου …….. 61
6.Τεχνητά νευρικά δίκτυα με ευρετικούς κανόνες…… 62
7.Τεχνητά νευρωνικά δίκτυα εξελισσόμενα από γενετικούς αλγόριθμους…….. 63
8.Καταλληλοτητα των MLP νευρωνικών δικτύων στην πρόβλεψη του ηλεκτρικού φορτίου……... 64
9.Ιεραρχικό υβριδικό νευρωνικό δίκτυο…… 65
10.Τεχνητα νευρωνικά δίκτυα εκπαιδευόμενα με την τεχνική βελτιστοποίησης σμηνών μορίων…….. 66
11. Σύγκριση τεχνικών πρόβλεψης ηλεκτρικού φορτίου……... 67
12. Πρόβλεψη κατανάλωσης ηλεκτρικού φορτίου ένα πρόβλημα πρωτίστως γραμμικό…… 68
13.Τεχνητά νευρωνικά δίκτυα μέσου χρονικού ορίζοντα……. 69
14. Συγκριση MLP και επαναληπτικών (recurrent) νευρωνικών δικτύων για μακροπρόθεσμη πρόβλεψη…… 70

15.Βελτιωμενα νευρωνικά δίκτυα βασισμένα σε μοντέλα πρόβλεψης….. 71
16.Συνδυασμός ελεγχόμενων και μη τεχνητών νευρωνικών δικτύων……. 72
17.Σύγκριση Multilayer Perceptron και Radial Basis Function νευρωνικών δικτύων……. 73
18.Τεχνητό νευρωνικό δίκτυο με δείκτες ημέρας και ώρας……. 74
19. RBFNN νευρωνικά δίκτυα που χρησιμοποιούν την συνδυασμένη διόρθωση υπολοίπων λαθών……75
20.Κυματοειδή νευρωνικά δίκτυα με την τεχνική bacterial foraging……...76
21.RSAN νευρωνικά δίκτυα……77
22.Τεχνητά νευρωνικά δίκτυα με ανατροφοδότηση……78
23.Τεχνητά νευρωνικά δίκτυα και ανάλυση ευαισθησίας των παραγόντων 79
24.Σύνθετα νευρωνικά δίκτυα : Adaptive Combiner ANN……..80
25.Δυναμικά τεχνητά νευρωνικά δίκτυα : το πρότυπο (DAN2)…….. 81
26.Συνδυασμος τεχνητών νευρωνικών δικτύων και εμπείρων εξειδικευμένων συστημάτων……82
27.Ασταθή νευρωνικά δίκτυα……...83

Τεχνικές πρόβλεψης ηλεκτρικού φορτίου

Η μέθοδος αναδρομής ……84
Προχωρημένοι αλγόριθμοι αναστροφής διάδοσης για MLP δίκτυα……...84
Δίκτυα τυχαίας ενεργοποίησης βαρών εκπαιδευόμενα με κινούμενο παράθυρο αναδρομής …….. 87
Δίκτυα με συναρτήσεις radial basis……90
Επαναληπτικά δίκτυα πραγματικού χρόνου……92
Αναδρομικά επαναληπτικά δίκτυα……..93
Άλλες δυνατές προσεγγίσεις…….. 96

Κεφάλαιο 5
Νευρωνικό μοντέλο

Συνοπτική περιγραφή νευρωνικού μοντέλου…….98

Κεφάλαιο 6

Αποτελέσματα- συμπεράσματα νευρωνικού μοντέλου ……102

Βιβλιογραφία ……107

Παραρτήματα

Παράρτημα α Παρουσίαση ολοκλήρου του νευρωνικού μοντέλου ….. 110
Παράρτημα β Δεδομένα…….……..131
Παράρτημα γ Περίληψη βιβλιογραφίας ……..146

ΠΡΟΛΟΓΟΣ

Η παρούσα εργασία αποτελεί τη Διπλωματική μου Εργασία στα πλαίσια των σπουδών μου στο Τμήμα Μηχανικών Παραγωγής και Διοίκησης του Πολυτεχνείου Κρήτης. Η εκπόνησή της ξεκίνησε το Σεπτέμβριο του 2007 και ολοκληρώθηκε τον Ιανουάριο του 2008, υπό την επίβλεψη του κ. Γεωργίου Ατσαλάκη.

Με την ευκαιρία της παρουσίασης της εργασίας μου αυτής, θεωρώ χρέος μου να εκφράσω τις ευχαριστίες μου σε όσους συνέβαλαν αποφασιστικά στην όλη μου προσπάθεια. Αρχικά, θα ήθελα να ευχαριστήσω τον επιβλέποντα Καθηγητή μου κ. Ατσαλάκη Γιώργο, ο οποίος μου έδωσε τη δυνατότητα να ασχοληθώ με το αντικείμενο της πρόβλεψης του ηλεκτρικού φορτίου με νευρωνικά δίκτυα και γενετικούς αλγορίθμους και με την βοήθεια του είτε με τη μορφή καθοδήγησης είτε με τις πολύτιμες συμβουλές που παρείχε, δημιούργησε ένα άρτιο κλίμα συνεργασίας σε όλη την πορεία της εργασίας, και συνέβαλε τα μέγιστα για την επιτυχή της ολοκλήρωση.

Τέλος θέλω να αφιερώσω αυτή την εργασία στην οικογένεια μου η οποία με στήριξε όλα αυτά τα χρονιά στο πολυτεχνείο, αλλά και στον εαυτό μου για τον όποιο κόπο μου όλα αυτά τα χρονιά.

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία παρουσιάζει την ανάπτυξη μεθοδολογίας για τη πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας (φορτίου) με χρήση νευρωνικών δικτύων και γενετικών αλγορίθμων. Στόχος της εργασίας είναι η δημιουργία ενός μοντέλου που θα παρέχει αξιόπιστες προβλέψεις αλλά και η εξερεύνηση των δυνατοτήτων που μας δίνονται ελεώ νευρωνικών δικτύων και γενετικών αλγορίθμων στο τομέα της πρόβλεψης του ηλεκτρικού φορτίου. 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.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Νευρωνικά Δίκτυα, Γενετικοί αλγόριθμοι, Πρόβλεψη κατανάλωσης ηλεκτρικού φορτίου

ΚΕΦΑΛΑΙΟ 1

ΠΡΟΒΛΕΨΗ ΗΛΕΚΤΡΙΚΟΥ ΦΟΡΤΙΟΥ

ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

Η συνεχής παρακολούθηση του φορτίου από την παραγωγή του συστήματος είναι μία βασική απαίτηση της λειτουργίας των συστημάτων ηλεκτρικής ενέργειας.
Για τον έλεγχο και την οικονομική λειτουργία του συστήματος, η παρακολούθηση του φορτίου από την παραγωγή, πρέπει να γίνεται συνεχώς και σε ένα μεγάλο φάσμα χρονικών διαστημάτων. Σε διάστημα δευτερολέπτων, όπου οι διακυμάνσεις του φορτίου είναι σχετικά μικρές και τυχαίες, **ο Αυτόματος Έλεγχος Παραγωγής (ΑΕΠ)*** φροντίζει ώστε η παραγωγή του συστήματος να αντισταθμίζει τη ζήτηση του φορτίου. Σε διαστήματα μερικών λεπτών, όπου οι διακυμάνσεις του φορτίου είναι μεγαλύτερες**, η Οικονομική Κατανομή Φορτίου*** φροντίζει για την κατανομή του συνολικού φορτίου του συστήματος στις ήδη ενταγμένες στο δίκτυο μονάδες κατά τον πλέον οικονομικό τρόπο. Σε διαστήματα μερικών ωρών ή και ημερών, οι διακυμάνσεις του φορτίου είναι ακόμη μεγαλύτερες. Η κάλυψη του φορτίου για τέτοια χρονικά διαστήματα απαιτεί την εκκίνηση ή την κράτηση μονάδων παραγωγής ή ανταλλαγή ισχύος με τα γειτονικά συστήματα ηλεκτρικής ενέργειας.

Λειτουργίες βραχυπρόθεσμου προγραμματισμού της παραγωγής, όπως η **Οικονομική Ένταξη Μονάδων Παραγωγής*, η Υδροθερμική Συνεργασία* και οι Οικονομικές Ανταλλαγές Ισχύος***, φροντίζουν για την κάλυψη του φορτίου για χρονικά διαστήματα που κυμαίνονται από τις λίγες επόμενες ώρες έως και την επόμενη εβδομάδα. Για χρονικές περιόδους μηνών έως και πέντε χρόνων γίνεται προγραμματισμός καυσίμων, χρήσης υδάτινων πόρων, συντήρησης των μονάδων κατά τρόπο που να εξασφαλίζεται η αξιόπιστη κάλυψη του φορτίου με τις ήδη εγκατεστημένες στο δίκτυο μονάδες. Τέλος, για μεγάλα χρονικά διαστήματα, μέχρι και τα 25 επόμενα χρόνια, γίνεται προγραμματισμός ένταξης (εγκατάστασης) νέων μονάδων στο σύστημα, ώστε να καλυφθεί ενδεχόμενη αύξηση του φορτίου του συστήματος .

* Οι λειτουργίες: Αυτόματος Έλεγχος Παραγωγής (ΑΕΠ), Οικονομική Κατανομή Φορτίου, Οικονομική Ένταξη Μονάδων Παραγωγής, Υδροθερμική Συνεργασία και Οικονομικές Ανταλλαγές Ισχύος, που αναφέρονται παραπάνω και σχετίζονται με τον προγραμματισμό και τον έλεγχο της παραγωγής ηλεκτρικής ενέργειας, είναι βασικές λειτουργίες του υποσυστήματος παραγωγής, το οποίο υπάγεται στη λειτουργία της Ενεργειακής Διαχείρισης του Λογισμικού Εφαρμογών των Συστημάτων Ελέγχου Ενέργειας στη χώρα μας.

Η πρόβλεψη φορτίου είναι, λοιπόν, μία από τις βασικές λειτουργίες στα συστήματα ηλεκτρικής ενέργειας. Το κίνητρο για ακριβείς προβλέψεις βρίσκεται στη φύση του ηλεκτρισμού ως ένα αγαθό και εμπορικό είδος. Ο ηλεκτρισμός δεν μπορεί να αποθηκευτεί, κάτι το οποίο σημαίνει ότι για μια ηλεκτρική χρήση η εκτίμηση της μελλοντικής απαίτησης είναι απαραίτητη για τη διαχείριση της παραγωγής και αγοράς με έναν οικονομικό τρόπο. Στην Ελλάδα η αγορά ενέργειας έχει πρόσφατα ανοίξει, γεγονός που αυξάνει τον ανταγωνισμό.

Βλέπουμε, επομένως, πόσο σημαντική είναι η πρόβλεψη του ηλεκτρικού φορτίου, δηλαδή της ζήτησης ηλεκτρικής ενέργειας, για διάφορους χρονικούς ορίζοντες, για τη σωστή και καλύτερη τεχνοοικονομικά διαχείριση και λειτουργία ενός συστήματος ηλεκτρικής ενέργειας. Υπάρχει, επομένως, η ανάγκη για μία κατά το δυνατόν αξιόπιστη και με μικρές αποκλίσεις από την πραγματικότητα πρόβλεψη του φορτίου, με στόχο τα αποτελέσματα αυτής να είναι τα καλύτερα δυνατά ούτως ώστε να μπορούν να χρησιμοποιηθούν για το σχεδιασμό των απαραίτητων ενεργειών που αφορούν τα συστήματα ηλεκτρικής ενέργειας.

Οι μέθοδοι πρόβλεψης φορτίου μπορούν να χωριστούν σε πολύ βραχυπρόθεσμα, βραχυπρόθεσμα, μεσοπρόθεσμα και μακροπρόθεσμα μοντέλα ανάλογα με το χρονικό διάστημα . Σε πολύ βραχυπρόθεσμες προβλέψεις φορτίου ο προβλεπόμενος χρόνος μπορεί να είναι ακόμη και μερικά λεπτά, ενώ σε μακροπρόθεσμες προβλέψεις μπορεί να είναι από μερικά χρόνια έως και μερικές δεκαετίες.

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

Πίνακας 1 : Χαρακτηρισμός προβλέψεων φορτίου ανάλογα με το χρονικό ορίζοντα.

Οι προβλέψεις του φορτίου ανάλογα με το χρονικό ορίζοντα πρόβλεψης χαρακτηρίζονται όπως φαίνεται στον πίνακα 1 .
Η πρόβλεψη φορτίου (load forecasting) αποτελεί ένα ευρέως διαδεδομένο θέμα στη βιβλιογραφία των συστημάτων ηλεκτρικής ενέργειας (ΣΗΕ). Ένας λόγος είναι ότι πρόσφατες επιστημονικές καινοτομίες μας έχουν φέρει πιο κοντά στη λύση του προβλήματος. Η ανάπτυξη της τεχνολογίας των ηλεκτρονικών υπολογιστών έχει αυξήσει τις δυνατότητες για αυτές και άλλες μεθόδους που λειτουργούν σε περιβάλλον πραγματικού χρόνου. Ένας ακόμη λόγος μπορεί να είναι ότι υπάρχει μία διεθνής τάση αύξησης του ανταγωνισμού στις ηλεκτρικές αγορές .
Ακόμη κι αν πολλές μέθοδοι πρόβλεψης έχουν ελεγχθεί και αποδειχθεί επιτυχείς, καμία δεν έχει κατορθώσει να καθιερωθεί σαν μια ευρέως εφαρμοζόμενη μέθοδος. Ο λόγος είναι ότι οι συνθήκες και οι απαιτήσεις της κάθε μεθόδου έχουν σημαντική επίδραση στην επιλογή της κατάλληλης μεθόδου. Τα αποτελέσματα που έχουν παρουσιασθεί σε σχετική βιβλιογραφία συνήθως δεν είναι άμεσα συγκρίσιμα μεταξύ τους.

Μια μειονότητα των πρόσφατων προσεγγίσεων βασίζεται στις τεχνικές νευρωνικών δικτύων. Πολλές έρευνες έχουν επιφέρει ικανοποιητικά αποτελέσματα. Η ελκυστικότητα των μεθόδων βρίσκεται στη δυνατότητα την οποία έχουν τα νευρωνικά δίκτυα να είναι σε θέση να μάθουν ιδιότητες του φορτίου, που διαφορετικά θα απαιτούσε προσεκτική ανάλυση για να εντοπιστούν. Εξίσου ενδιαφέρουσες αποτελούν και οι τεχνικές πρόβλεψης που στηρίζονται στη χρήση της ασαφούς λογικής για την ανάπτυξη αυτόνομων συστημάτων πρόβλεψης φορτίου (νευρωασαφή δίκτυα), αλλά και στην ενσωμάτωση ή το συνδυασμό της ασαφούς λογικής σε κλασσικά νευρωνικά δίκτυα.
Παρόλα αυτά, η ανάπτυξη των μεθόδων δεν έχει ολοκληρωθεί και η έλλειψη συγκρίσιμων αποτελεσμάτων σε παραλλαγές διαφορετικού είδους μοντέλων αποτελεί ένα πρόβλημα. Συνεπώς, για να κάνουμε χρήση των τεχνικών αυτών σε πραγματικές εφαρμογές μια συγκριτική ανάλυση των ιδιοτήτων των διαφόρων τύπων μοντέλων δείχνει να είναι απαραίτητη.

ΠΑΡΑΓΟΝΤΕΣ ΠΟΥ ΕΠΗΡΕΑΖΟΥΝ ΤΟ ΦΟΡΤΙΟ

Γενικά, το φορτίο μιάς 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.ά.

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:

(α)Οικονομικοί
(β)Χρονικοί
(γ) Καιρικοί
(δ) Τυχαίοι

Για τη δημιουργία ενός μοντέλου που περιγράφει το φορτίο θα πρέπει να γίνει κατανοητή η επίδραση όλων των παραγόντων στη διαμόρφωση της ζήτησης. Ακολουθεί μία σύντομη περιγραφή της επίδρασης των παραπάνω παραγόντων στη διαμόρφωση της ζήτησης της ηλεκτρικής ενέργειας.

(α) Οικονομικοί Παράγοντες

Το οικονομικό περιβάλλον εντός του οποίου λειτουργεί μία ηλεκτρική εταιρεία επιδρά αναμφίβολα στη διαμόρφωση της ηλεκτρικής ζήτησης. Δημογραφικοί παράγοντες, επίπεδα βιομηχανικής δραστηριότητας, ανάπτυξη ή και κορεσμός της αγοράς νέων ηλεκτρικών συσκευών, αλλαγές στον αγροτικό τομέα, κρατικές παρεμβάσεις και γενικά η οικονομική δραστηριότητα, έχουν μεγάλη επίδραση στο ρυθμό αύξησης ή μείωσης του φορτίου του συστήματος. Αλλαγές στην τιμολογιακή πολιτική της επιχείρησης, καθώς και προγράμματα εξοικονόμησης ενέργειας και διαχείρισης της ζήτησης επιδρούν επίσης στη διαμόρφωση της ζήτησης. Τυπικά, οι οικονομικοί παράγοντες έχουν χρονικές σταθερές σημαντικά μεγαλύτερες από μία εβδομάδα και δεν επηρεάζουν βραχυπρόθεσμα τη ζήτηση.

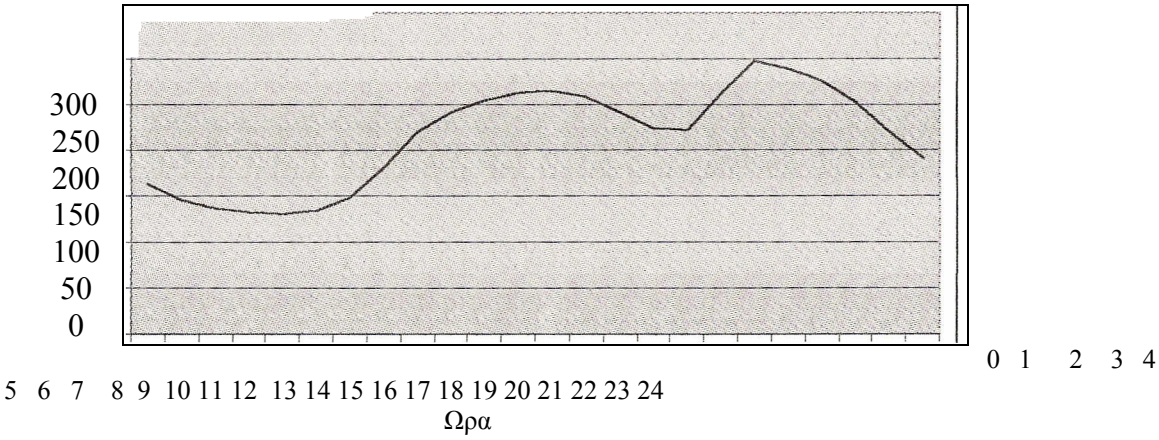
(β) Χρονικοί Παράγοντες

Τρεις βασικοί χρονικοί παράγοντες - εποχιακές μεταβολές, εβδομαδιαίος/ημερήσιος κύκλος και εθνικές και θρησκευτικές εορτές/αργίες επηρεάζουν σημαντικά το ηλεκτρικό φορτίο. Οι εποχιακές μεταβολές επιδρούν στη μορφή της καμπύλης φορτίου. Στο σχήμα 1 δίνεται μία τυπική ημερήσια καμπύλη φορτίου χειμώνα και μία θέρους του εθνικού διασυνδεδεμένου συστήματος. Με απλή εποπτεία, παρατηρούμε τις διαφορές στην ώρα εμφάνισης της βραδινής αιχμής, καθώς και την τρίτη αιχμή που εμφανίζεται το χειμώνα λόγω της θερμοσυσσώρευσης. Ορισμένες μεταβολές στην καμπύλη φόρτισης γίνονται σταδιακά με την αλλαγή διαφόρων εποχιακών μεταβλητών, όπως η θερμοκρασία και οι ώρες ηλιοφάνειας. Άλλες εποχιακές μεταβολές δημιουργούν απότομες αλλαγές στην καμπύλη φόρτισης. Τέτοιες αλλαγές είναι η αλλαγή ώρας (Χειμερινή – Θερινή) (σε αυτήν την περίπτωση παρατηρείται μετατόπιση της καμπύλης φορτίου), η έναρξη της σχολικής χρονιάς, η πιθανή εποχιακή μεταβολή στις χρονικές ζώνες φθηνού τιμολογίου κ.ά.

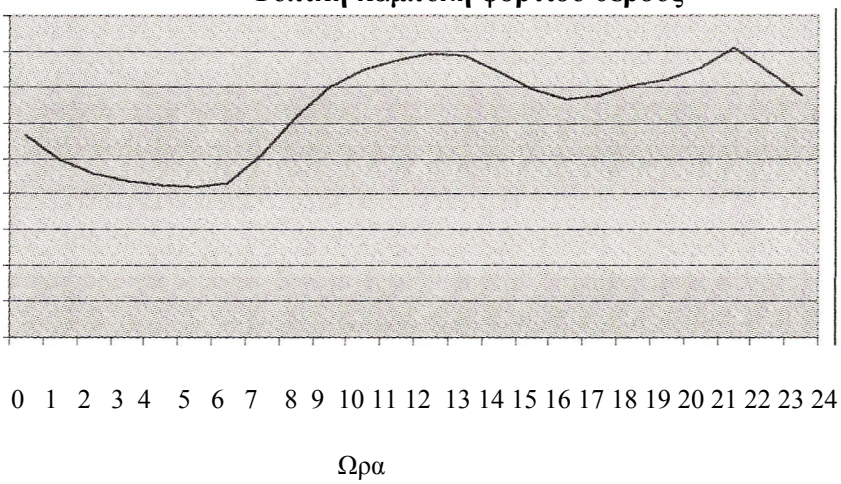
Η εβδομαδιαία/ημερήσια περιοδικότητα του φορτίου είναι αποτέλεσμα της περιοδικότητας του κύκλου εργασίας - ανάπαυσης του πληθυσμού. Υπάρχουν προκαθορισμένες μορφές καμπυλών φορτίου για μία «τυπική» εβδομάδα κάθε εποχής. Το Σαββατοκύριακο η ζήτηση είναι μειωμένη και το επίπεδο της ζήτησης, καθώς και η μορφή της καμπύλης φορτίου, της Δευτέρας και της Παρασκευής είναι κάπως διαφορετική από τις υπόλοιπες εργάσιμες μέρες λόγω γειτνίασης με το Σαββατοκύριακο.

Η Τρίτη, η Τετάρτη και η Πέμπτη έχουν συνήθως παρόμοιες καμπύλες φορτίου «εργάσιμης μέρας». Τέλος, η ζήτηση φορτίου κατά τη διάρκεια των εθνικών ή των θρησκευτικών εορτών είναι σημαντικά χαμηλότερη από τα φυσιολογικά επίπεδα. Σημαντική μείωση της ζήτησης παρατηρείται επιπλέον και κατά τη διάρκεια των εθνικών ή των θρησκευτικών εορτών όπου είναι σημαντικά χαμηλότερη από τα φυσιολογικά επίπεδα . Σημαντική μείωση της ζήτησης παρατηρείται επιπλέον και κατά τη διάρκεια εορταστικών τριημέρων.

Τυπική καμπύλη φορτίου χειμώνα



Τυπική καμπύλη φορτίου θέρους



Σχήμα 1: Τυπικές καμπύλες φορτίου για μία ημέρα του χειμώνα (Τετάρτη 17/11/1994) και μία ημέρα του καλοκαιριού (Τετάρτη 11/7/1994)

(γ) Καιρικοί Παράγοντες

Οι καιρικές συνθήκες προκαλούν σημαντικές αλλαγές στην καμπύλη φορτίου. Αυτό συμβαίνει διότι 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. Άλλοι παράγοντες, όπως υγρασία, βροχόπτωση, άνεμος, νεφώσεις, επηρεάζουν επίσης τη ζήτηση.

(δ) 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 .
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. Αυτή είναι και η προσέγγιση που θα κάνουμε στη μελέτη μας. Το γεγονός αυτό μειώνει τον αριθμό των παραγόντων που πρέπει να ληφθούν υπόψη. Οι πιο σημαντικοί είναι :
-Βραχυπρόθεσμα, οι μετεωρολογικές συνθήκες προκαλούν μεγάλες αποκλίσεις στο συνολικό φορτίο. Εκτός της θερμοκρασίας, και η ταχύτητα του ανέμου ,τα σύννεφα και η υγρασία ασκούν κάποια επιρροή .
- Μακροπρόθεσμα, οικονομικοί και δημογραφικοί παράγοντες παίζουν τον πιο σημαντικό ρόλο στον καθορισμό της εξέλιξης της ηλεκτρικής ζήτησης.
- Από τη σκοπιά της πρόβλεψης, οι χρονικοί παράγοντες είναι βασικοί. Διάφορες περιοδικές επιδράσεις και επαναλαμβανόμενες καταστάσεις καθώς και νόμιμες και θρησκευτικές αργίες λαμβάνονται υπόψη.

Οι άλλοι παράγοντες που προκαλούν διαταραχές και μπορούν να καταταχθούν σαν τυχαίοι παράγοντες είναι συνήθως μικροί στην περίπτωση των ατομικών καταναλωτών. Παρόλ' αυτά μεγάλα κοινωνικά γεγονότα και δημοφιλή τηλεοπτικά προγράμματα προσθέτουν αβεβαιότητα στις προβλέψεις. Από την άλλη βιομηχανικές μονάδες μπορούν να προκαλέσουν σχετικά μεγάλες διαταραχές.
Η απόφαση να συνδυάσουμε όλες τις καταναλωτικές μονάδες σε ένα συνολικό φορτίο σημαίνει ότι η πρόβλεψη θα στηριχθεί κυρίως στην προηγούμενη συμπεριφορά του φορτίου. Οι χρονικοί παράγοντες παίζουν τον καθοριστικό ρόλο στην ανάλυση αυτής της μελέτης.

Κεφαλαίο 2
ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Ιστορική αναδρομή

Η μοντέρνα περίοδος των νευρωνικών δικτύων λέγεται ότι άρχισε με την πρωτοποριακή δουλειά των McCulloch και Pitts (1943). Ο πρώτος ήταν ψυχίατρος και ο δεύτερος μεγαλοφυής μαθηματικός. Σύμφωνα με τον Rall (1990) η κλασσική εργασία των McCulloch και Pitts έγινε μέσα σε μια κοινωνία που ασχολούνταν με τους νευρώνες στο πανεπιστήμιο του Σικάγο, για πάνω από 5 χρόνια. Αυτή η εργασία περιγράφει το λογικό λογισμό των νευρωνικών δικτύων. Είναι αξιοσημείωτο το ότι ο von Neumann χρησιμοποίησε ιδεατά στοιχεία καθυστέρησης τα οποία είχαν υπολογιστεί από τα ιδεατά στοιχεία νευρώνων των McCulloch και Pitts, για την κατασκευή του EDVAC (Electronic Discrete Variable Automatic Computer), ο οποίος κατάληξε στον ENIAC, τον πρώτο γενικού σκοπού υπολογιστή.
Η επόμενη μεγάλη ανάπτυξη πάνω στα νευρωνικά δίκτυα, ήρθε το 1949 με την έκδοση του βιβλίου του Hebb με τίτλο "The Organization of Behavior", στο οποίο έγινε για πρώτη φορά μια ιδιαίτερη δήλωση ενός φυσιολογικού κανόνα μάθησης για συνοπτικές τροποποιήσεις. Πιο συγκεκριμένα ο Hebb πρότεινε ότι η συνδετικότητα του εγκεφάλου συνεχώς αλλάζει καθώς ο οργανισμός μαθαίνει διάφορες εργασίες, και ότι οι νευρωνικοί συγκέντρωτες δημιουργούνται από τέτοιες αλλαγές. Επίσης πρότεινε το διάσημο αίτημα μάθησης σύμφωνα με το οποίο η αποτελεσματικότητα μιας σύναψης μεταβλητής ανάμεσα σε δύο νευρώνες αυξάνεται από την επαναλαμβανόμενη ενεργοποίηση του ενός νευρώνα από τον άλλο κατά μήκος της σύναψης.
Η αναφορά των Rochester , Holland , Habit και Duda (1956) είναι ίσως η πρώτη προσπάθεια για χρήση υπολογιστή που χρησιμοποιεί την εξομοίωση , για να ελεγχθεί μια καλά σχηματισμένη νευρωνική θεωρία βασισμένη στο αίτημα μάθησης του Hebb. Η εξομοίωση έδειξε ότι χρειαζόταν να προστεθεί παρεμπόδιση ώστε η θεωρία να δουλέψει πραγματικά. Τον ίδιο χρόνο ο Uttley παρουσίασε την αποκαλούμενη "διαρρέουσα ολοκλήρωση" ή "νευρώνας φωτιάς" που αργότερα αναλύθηκε από τον Caianielo.
Το 1952 εκδόθηκε το βιβλίο του Ashby με τίτλο "Design for a brain: The Origin of Adaptive Behavior", το οποίο ασχολήθηκε με την βασική έννοια ότι η προσαρμοζόμενη συμπεριφορά δεν είναι έμφυτη αλλά μαθαίνεται. Το 1954 ο Minsky έγραψε τη διδακτορική του διατριβή με τίτλο "Theory of Neural-Analog Reinforcement Systems and Its Application to the Brain-Model Problem" και το 1961 ο ίδιος έγραψε μια εργασία με τίτλο "Steps Toward Artificial Intelligence". Επίσης το 1954 η ιδέα των μη γραμμικών προσαρμοζόμενων φίλτρων προτάθηκε από τον Gabor (πρωτοπόρος της θεωρίας επικοινωνιών και εφευρέτης του ολογραφήματος), ο οποίος υλοποίησε μια τέτοια μηχανή στην οποία η μάθηση επιτυγχανόταν με τροφοδότηση δειγμάτων στοχαστικών διαδικασιών στη μηχανή, μαζί με τη συνάρτηση-στόχο, την οποία ήταν αναμενόμενο να παράγει η μηχανή.
Ένα θέμα ιδιαίτερου ενδιαφέροντος για τα νευρωνικά δίκτυα είναι αυτό της σχεδίασης ενός αξιόπιστου δικτύου με νευρώνες που μπορούν να θεωρηθούν σαν μη αξιόπιστα στοιχεία. Αυτό το σημαντικό πρόβλημα λύθηκε από τον von Neumann (1956) χρησιμοποιώντας την ιδέα του πλεονασμού. 15 χρόνια μετά την έκδοση της εργασίας

προτύπων έγινε από τον Rosenblatt (1958) στην εργασία του πάνω στο αισθητήριο των McCulloch και Pitts, μια νέα προσέγγιση πάνω στο πρόβλημα της αναγνώρισης (perceptron). Το ιδιαίτερο επίτευγμα του ήταν το αποκαλούμενο θεώρημα σύγκλισης αισθητηρίου (perceptron convergence theorem). Το 1960 οι Widrow και Hoff πρότειναν τον αλγόριθμο ελάχιστου μέσου τετραγώνου (least mean-square-LMS) και τον χρησιμοποίησαν για να σχηματίσουν το Adaline (adaptive linear element). Η διαφορά ανάμεσα στο αισθητήριο και στο Adaline βρίσκεται στον τρόπο μάθησης. Ένα από τα από τα πρόσφατα εκπαιδεύσιμα νευρωνικά δίκτυα με πολλαπλά στοιχεία είναι η δομή Madaline (Widrow). Κατά την διάρκεια της κλασσικής περιόδου του perceptron ('60) πιστευόταν ότι τα νευρωνικά δίκτυα μπορούσαν να κάνουν τα πάντα. Αλλά τότε εκδόθηκε το βιβλίο των Minsky και Papert που με τη βοήθεια των μαθηματικών απέδειξε ότι υπάρχουν όρια πάνω στο τι μπορεί να υπολογιστεί από τα αισθητήρια. Ένα σημαντικό πρόβλημα πάνω στη σχεδίαση ενός πολυεπίπεδου αισθητηρίου είναι το πρόβλημα της ανάθεσης εμπιστοσύνης (credit assignment problem), το οποίο βρήκε την λύση του μόλις την δεκαετία του 1980.

Κατά την δεκαετία του '70 λόγω των προβλημάτων εγκαταλείφτηκε το ενδιαφέρον πάνω στα νευρωνικά δίκτυα. Μια σημαντική ενέργεια την δεκαετία αυτή ήταν οι χάρτες αυτοοργάνωσης με ανταγωνιστική μάθηση.

Το 1980 έγιναν πολλές εργασίες πάνω στην θεωρία αλλά και στον σχεδιασμό των νευρωνικών δικτύων. Ο Grossberg (1980) ανέπτυξε μια καινούργια αρχή αυτοοργάνωσης που συνδυάζει φιλτράρισμα από "κάτω προς τα πάνω" και αντίθετη αύξηση σε μικρή μνήμη με από "πάνω προς τα κάτω" ταίριασμα προτύπων και σταθεροποίηση του κώδικα μάθησης. Δεδομένης μιας τέτοιας ικανότητας, και αν το πρότυπο εισόδου ταιριάζει με την ανάδραση μάθησης, λαμβάνει χώρα μία δυναμική κατάσταση που καλείται adaptive resonance. Αυτό το φαινόμενο δίνει την βάση για μια νέα κατηγορία νευρωνικών δικτύων γνωστά σαν adaptive resonance theory (ART). Το 1982 ο Hopfield χρησιμοποίησε την ιδέα μια συνάρτησης ενέργειας για να φτιάξει ένα νέο τρόπο κατανόησης του υπολογισμού που γίνεται από τα δίκτυα με συμμετρικές συνοπτικές συνδέσεις. Επιπλέον καθιέρωσε τον *ισομορφισμό* ανάμεσα σε τέτοια περιοδικά δίκτυα και σε ένα Ising μοντέλο που χρησιμοποιείται στην στατιστική. Αυτή η αναλογία άνοιξε τον δρόμο για ένα κατακλυσμό θεωριών για τα νευρωνικά δίκτυα. Αυτή η συγκεκριμένη τάξη νευρωνικών δικτύων με ανατροφοδότηση έτυχε ιδιαίτερης προσοχής κατά τη δεκαετία του '80 και με το χρόνο έγιναν γνωστά σαν δίκτυα Hopfield.

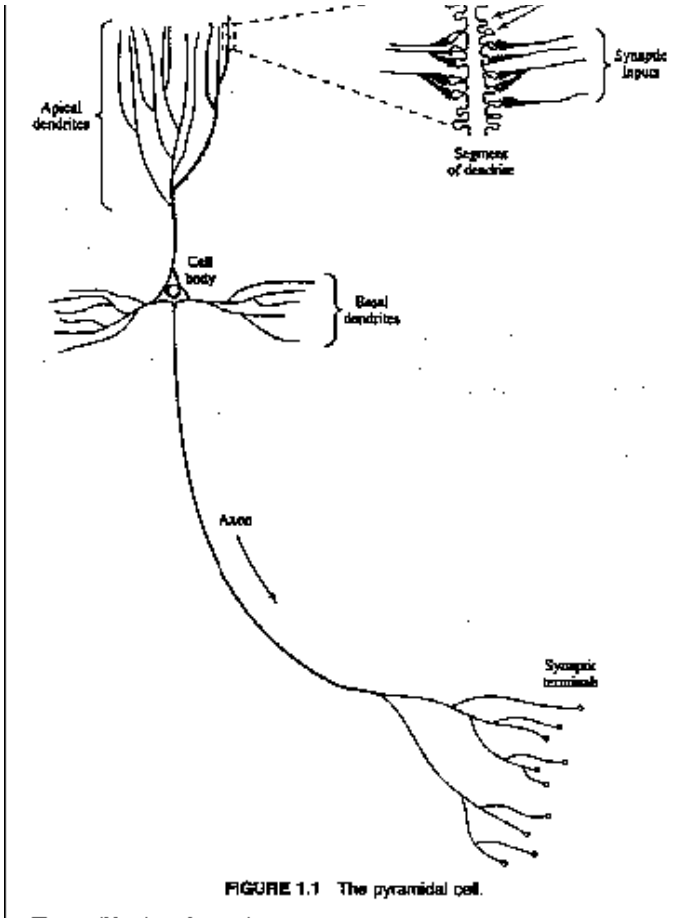
Το 1983 οι Cohen και Grossberg έδωσαν μια νέα αρχή για σχεδίαση μιας διευθυνσιοδοτούμενης μνήμης (content-addressable memory) που περιλαμβάνει την έκδοση συνεχούς χρόνου του δικτύου Hopfield σαν μια ιδιαίτερη περίπτωση. Μια ακόμα σημαντική ανάπτυξη το 1982 ήταν η έκδοση της εργασίας του Kohonen πάνω στους χάρτες αυτοοργάνωσης, χρησιμοποιώντας μιας ή δύο διαστάσεων δικτυωτές δομές. Το 1983 οι Kirkpatrick, Gallat και Vecchi περιέγραψαν μια νέα διαδικασία που λέγεται εξομοιωμένη ανόπτηση για λύση προβλημάτων συνδυαστικής βελτιστοποίησης. Η εξομοιωμένη ανόπτηση χρησιμοποιείται στην στατιστική θερμοδυναμική και βασίζεται σε μια απλή τεχνική. Την ίδια χρονιά εκδόθηκε μια εργασία από τους Burto, Sutton και Anderson πάνω στην ενισχυμένη μάθηση, η οποία δημιούργησε μεγάλο ενδιαφέρων πάνω στην ενισχυμένη μάθηση και την εφαρμογή της. Το 1984 ο Braitenberg εξέδωσε ένα βιβλίο με τίτλο "Vehicles: Experiments in Synthetic Psychology" το οποίο

περιγράφει διάφορες μηχανές με απλή εσωτερική αρχιτεκτονική, και το οποίο ενσωματώνει μερικές σημαντικές αρχές της αυτοοργανούμενης απόδοσης. Το 1986 η ανάπτυξη του αλγορίθμου για πίσω διάδοση (back-propagation algorithm) παρουσιάστηκε για πρώτη φορά από τον Rumelhart.
Αυτός ο αλγόριθμος έγινε πολύ δημοφιλής και έδωσε νέα ώθηση στις εφαρμογές των νευρωνικών δικτύων. Το 1988 ο Linsker περιέγραψε μια νέα αρχή για την αυτοοργάνωση σε ένα δίκτυο αισθητηρίων (perceptrons). Η αρχή αυτή σχεδιάστηκε ώστε να διατηρεί μέγιστη πληροφορία σχετικά με τα πρότυπα ενεργειών, που υπόκεινται σε περιορισμούς όπως συνοπτικές συνδέσεις και δυναμικές περιοχές σύναψης. Ο Linsker χρησιμοποίησε αφηρημένες έννοιες πάνω στη θεωρία πληροφοριών ώστε να σχηματίσει τη αρχή της διατήρησης μέγιστης πληροφορίας.
Επίσης το 1988 οι Broomhead και Lowe περιέγραψαν μία διαδικασία για το σχεδιασμό "προς τα εμπρός τροφοδότησης" (feedforward) δικτύων χρησιμοποιώντας συναρτήσεις ακτινικής βάσης (Radial Basis Functions - RBF), που είναι μια παραλλαγή των πολυεπίπεδων αισθητηρίων. Το 1989 εκδόθηκε το βιβλίο του Mead με τίτλο "VLSI and Neural Systems". Αυτό το βιβλίο δίνει μια ασυνήθιστη μίξη περιεχομένων από την νευροβιολογία και την τεχνολογία VLSI. Ίσως περισσότερο από κάθε άλλη έκδοση, η εργασία του Hopfield (1982) και το δίτομο βιβλίο των Rummelhart και McLelland (1986), να ήταν οι πιο σημαίνουσες εκδόσεις υπεύθυνες για την αναζωογόνηση του ενδιαφέροντος για τα νευρωνικά δίκτυα στην δεκαετία του '80.
Τα νευρωνικά δίκτυα έχουν σίγουρα διανύσει πολύ δρόμο από την εποχή των McCulloch και Pitts. Πραγματικά έχουν εγκαθιδρυθεί σαν ενδοπειθαρχικό αντικείμενο με βαθιές ρίζες στην επιστήμη των νευρώνων, στην ψυχολογία, στα μαθηματικά, στις φυσικές επιστήμες και στην μηχανική. Δεν είναι αναγκαίο να πούμε ότι είναι εδώ για να μείνουν και θα συνεχίσουν να αναπτύσσονται σε θεωρία, σχεδιασμό και εφαρμογές.

Από τα Βιολογικά Νευρωνικά Δίκτυα στα Τεχνητά

Η μελέτη των τεχνητών νευρωνικών δικτύων, κοινώς γνωστά σαν 'νευρωνικά δίκτυα' υποκινήθηκε άμεσα από την αναγνώριση ότι οι εγκεφαλικοί υπολογιστές είναι ένας εξ' ολοκλήρου διαφορετικός δρόμος από τους συμβατικούς ψηφιακούς υπολογιστές.
Τυπικά οι νευρώνες του εγκεφάλου είναι 5-6 τάξεις μεγέθη αργότεροι από τις λογικές πύλες σιλικόνης. Ο εγκέφαλος είναι ένας πολύ πολύπλοκος, μη-γραμμικός και παράλληλος υπολογιστής. Έχει τη δυνατότητα να οργανώνει τους νευρώνες έτσι ώστε να εκτελεί συγκεκριμένους υπολογισμούς πολύ πιο γρήγορα από τους πιο γρήγορους ψηφιακούς υπολογιστές που υπάρχουν.
Πώς τα καταφέρνει όμως ο εγκέφαλος; Κατά τη γέννησή του ο εγκέφαλος έχει την ικανότητα να κατασκευάζει τους δικούς του κανόνες, κοινώς "εμπειρία", η οποία μεγαλώνει με την πάροδο του χρόνου. Κατά τα 2 πρώτα χρόνια ζωής, έχουμε τη μέγιστη ανάπτυξη, όπου περίπου 1 εκατομμύριο συνάψεις (synapses) δημιουργούνται στο δευτερόλεπτο. Οι συνάψεις είναι οι βασικές δομικές και λειτουργικές μονάδες που μεσολαβούν στην ενδοεπικοινωνία των νευρώνων, όπως φαίνεται στο σχήμα 2 που δείχνει τη βασική δομή ενός νευρώνα . Άρα:

Ένα νευρωνικό δίκτυο είναι ένας συμπαγής παράλληλος κατανεμημένος επεξεργαστής, που έχει τη φυσική κλίση να αποθηκεύει εμπεριστατωμένη γνώση και να την κάνει διαθέσιμη για χρήση.



Σχήμα 2: Βασική δομή ενός νευρώνα

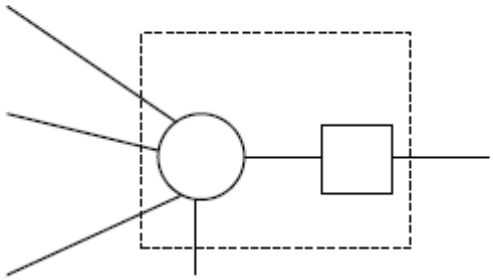
Ένα Τ.Ν.Δ. μοιάζει με τον εγκέφαλο στα εξής:

• Η γνώση αποκτάται από το δίκτυο μέσα από διαδικασία μάθησης.
• Οι δυνάμεις σύνδεσης των νευρώνων, γνωστές σαν συναπτικά (synaptic) βάρη, χρησιμοποιούνται για την αποθήκευση γνώσης.

Η διαδικασία για την εκμάθηση ονομάζεται “αλγόριθμος μάθησης”. Ο καθορισμός του τρόπου μεταβολής των συναπτικών (synaptic) βαρών, αποτελεί την παραδοσιακή μέθοδο για το σχεδιασμό των νευρωνικών δικτύων.
Τα τελευταία χρόνια έχει αναπτυχθεί ένα μεγάλο ενδιαφέρον σχετικά με την εφαρμογή της τεχνητής νοημοσύνης στις βιομηχανίες. Κύριο πλεονέκτημά τους αποτελεί το γεγονός ότι δεν απαιτείται περίπλοκη μαθηματική σχηματοποίηση ή ποσοτική συσχέτιση μεταξύ των εισόδων και εξόδων ενός συστήματος. Επίσης, δεν είναι απαραίτητα τα δεδομένα πολλών ετών. Έτσι, η ικανοποιητική απόδοση της τεχνητής νοημοσύνης έχει οδηγήσει σε επιτυχείς εφαρμογές σχετικά με την βραχυπρόθεσμη πρόβλεψη του ηλεκτρικού φορτίου. Πιο δημοφιλή όλων, είναι τα τεχνητά νευρωνικά δίκτυα (Artificial Neural Network, ANN). Τα νευρωνικά δίκτυα αναφέρονται σε μία κατηγορία μοντέλων τα οποία είναι εμπνευσμένα από τα βιολογικά νευρικά συστήματα. Τα μοντέλα αποτελούνται από πολλά προγραμματιζόμενα στοιχεία που συνήθως ονομάζονται νευρώνες και λειτουργούν παράλληλα. Τα στοιχεία συνδέονται με συνδετικά βάρη, κάτι το οποίο τους επιτρέπει να προσαρμόζονται μέσω μιας διαδικασίας εκμάθησης. Τα νευρωνικά δίκτυα μπορούν να ερμηνευθούν ως προσαρμόσιμες μηχανές που μπορούν να αποθηκεύουν γνώση μέσω της διαδικασίας εκμάθησης.

Η έρευνα πάνω στο πεδίο αυτό έχει μια ιστορία αρκετών δεκαετιών, αλλά, μετά από μία μείωση του ενδιαφέροντος στη δεκαετία του 1970, μία μαζική ανάπτυξη ξεκίνησε στις αρχές του 1980. Το κυρίαρχό τους πλεονέκτημα σε σχέση με τα στατιστικά μοντέλα έχει να κάνει με το γεγονός ότι μοντελοποιούν ένα πολυδιάστατο πρόβλημα χωρίς να απαιτούν περίπλοκες υποθέσεις ανάμεσα στις μεταβλητές εισόδου. Υπάρχουν πολλών τύπων μοντέλα νευρωνικών δικτύων. Το κοινό χαρακτηριστικό σε αυτά είναι η σχέση με τη γενικότερη ιδέα των βιολογικών συστημάτων. Δηλαδή, είναι ένα σύστημα επεξεργασίας δεδομένων που προσομοιώνουν τη λειτουργία του ανθρώπινου εγκεφάλου.
Πρακτικά, τα νευρωνικά δίκτυα αποτελούνται από ένα σύνολο στοιχειωδών μονάδων επεξεργασίας που ονομάζονται νευρώνες. Στο παρακάτω σχήμα (Σχήμα 3.1) φαίνεται το μαθηματικό μοντέλο του νευρώνα. Όπως φαίνεται στο Σχήμα 3.1 ένας νευρώνας μπορεί να δέχεται n εισόδους (x_1, x_2, …, x_n) και να παράγει μια έξοδο, y. Κάθε νευρώνας χαρακτηρίζεται από μια πόλωση b, η οποία μπορεί να είναι μη μηδενική ή μηδενική.
Σε κάθε μια από τις n εισόδους x_1, x_2, …, x_n αντιστοιχεί ένα βάρος w_1, w_2, …, w_n, αντίστοιχα. Η συνολική είσοδος z του νευρώνα υπολογίζεται ως εξής:

$$z = w_1 \cdot x_1 + w_2 \cdot x_2 + \cdots + w_n \cdot x_n + b$$



Σχήμα 3.1. Μαθηματικό μοντέλο νευρώνα

Ενώ, η έξοδος y του νευρώνα υπολογίζεται ως εξής:

$$y = f(z) = f(w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_n \cdot x_n + b)$$

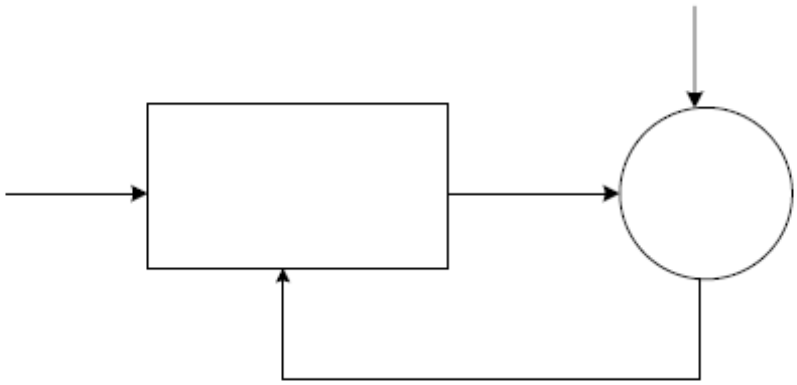
Η συνάρτηση f ονομάζεται συνάρτηση ενεργοποίησης ή συνάρτηση μεταφοράς. Δύο τυπικές συναρτήσεις ενεργοποίησης που χρησιμοποιούνται συχνά στα νευρωνικά δίκτυα είναι η σιγμοειδής συνάρτηση:

$$f(z) \equiv \log sig(z) = \frac{1}{1+e^{-z}}$$

και η συνάρτηση υπερβολικής εφαπτομένης:

$$f(z) \equiv \tan sig(z) \equiv \tanh(z) = \frac{2}{1+e^{-2z}} - 1$$

Η βασική φιλοσοφία των νευρωνικών δικτύων παρουσιάζεται στο Σχήμα 3.2.



Σχήμα 3.2. Φιλοσοφία λειτουργίας νευρωνικού δικτύου

Τύποι συναρτήσεων ενεργοποίησης

Η συνάρτηση ενεργοποίησης φ(.), ορίζει την έξοδο ενός νευρώνα συναρτήσει του επιπέδου ενεργοποίησης της εισόδου. Έχουμε 3 βασικούς τύπους συναρτήσεων ενεργοποίησης.

1.Συνάρτηση Κατωφλιού.

Για αυτόν τον τύπο συνάρτησης ενεργοποίησης, που περιγράφεται στο σχήμα 4a, έχουμε:

$$\phi(\upsilon) = \begin{cases} 1, \upsilon \geq 0 \\ 0, \upsilon < 0 \end{cases}$$

Επομένως η έξοδος ενός νευρώνα k, που έχει την παραπάνω συνάρτηση ενεργοποίησης, έχει τη μορφή:

$$y_k = \begin{cases} 1, \upsilon_k \geq 0 \\ 0, \upsilon_k < 0 \end{cases}$$

όπου υ_κ είναι το εσωτερικό επίπεδο ενεργοποίησης του νευρώνα και δίνεται από τη σχέση:

2.Τμηματικά Γραμμική Συνάρτηση.

Για την τμηματικά γραμμική συνάρτηση, που περιγράφεται στο σχήμα 4b, έχουμε:

$$\phi(\upsilon) = \begin{cases} 1, \upsilon \geq 1/2 \\ \upsilon, 1/2 > \upsilon > -1/2 \\ 0, \upsilon \leq 1/2 \end{cases}$$

Αυτή η μορφή συνάρτησης ενεργοποίησης μπορεί να θεωρηθεί σαν μια προσέγγιση ενός μη-γραμμικού ενισχυτή.

3.Σιγμοειδής.

Είναι η πιο συνηθισμένη μορφή συνάρτησης ενεργοποίησης που χρησιμοποιείται στην κατασκευή τεχνητών νευρωνικών δικτύων Ορίζεται σαν αυστηρά αύξουσα συνάρτηση, η οποία παρουσιάζει εξομάλυνση και ασυμπτωτικές ιδιότητες. Ένα παράδειγμα είναι η *λογιστική συνάρτηση,* που ορίζεται από τη σχέση:

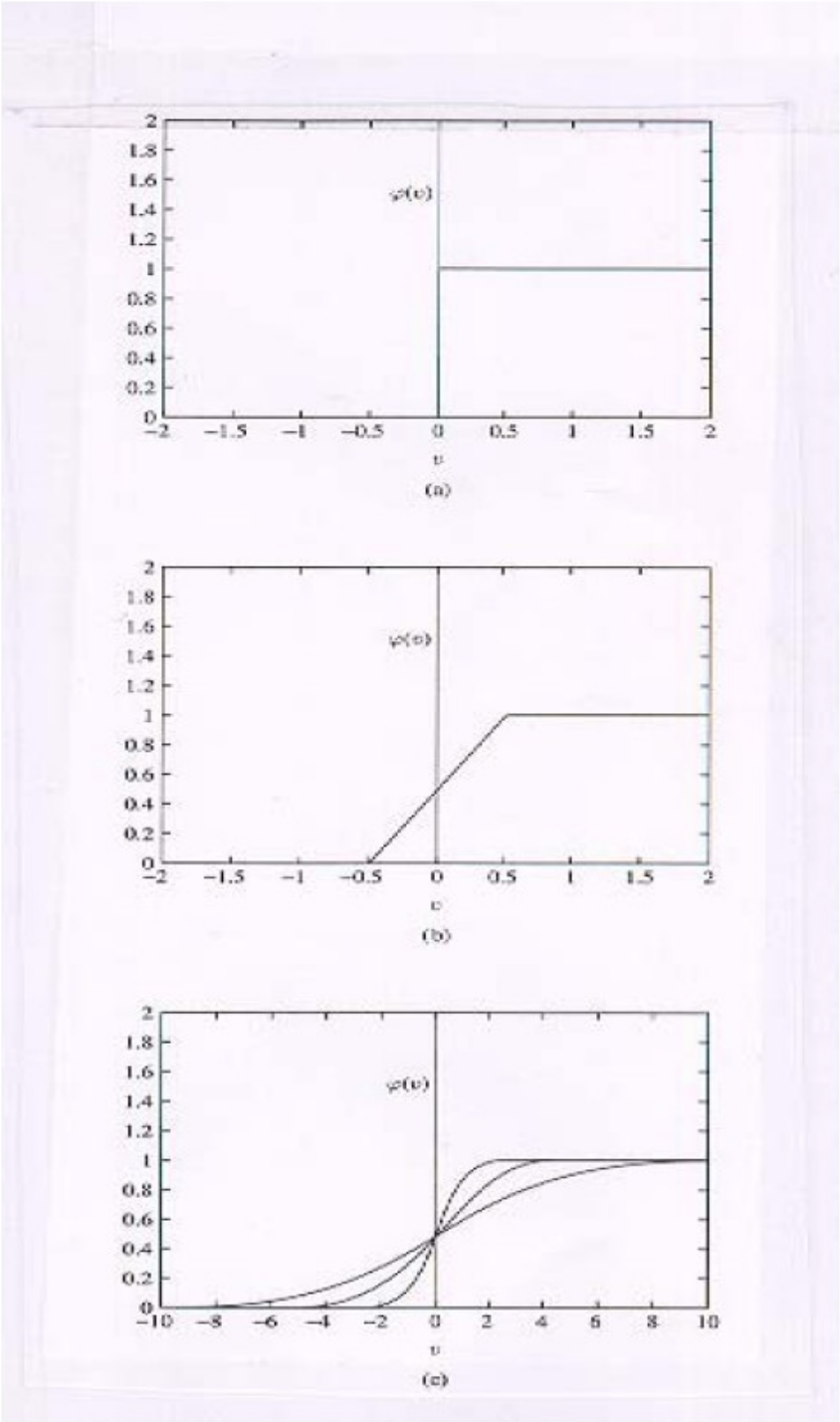
$$\phi(\upsilon) = \frac{1}{1+\exp(-\alpha\upsilon)}$$

Μεταβάλλοντας την παράμετρο *α* παίρνουμε διαφορετικές συναρτήσεις, όπως φαίνεται στο σχήμα 4c. Οι προηγούμενες συναρτήσεις ενεργοποίησης κυμαίνονται από 0 ως +1. Αν θέλουμε συνάρτηση που να κυμαίνεται από -1 ως +1, επαναπροσδιορίζουμε την συνάρτηση κατωφλιού ως:

$$\phi(\upsilon)\begin{cases} 1, \upsilon > 0 \\ 0, \upsilon = 0 \\ -1, \upsilon < 0 \end{cases}$$

που ονομάζεται συνάρτηση *προσήμου (signum) και* συμβολίζεται σαν sgn(.).

Στο σχήμα 4 στην επόμενη σελίδα παρουσιάζεται η γραφική παράσταση των παραπάνω τριών βασικών συναρτήσεων ενεργοποίησης.



Σχήμα 4: (α) Συνάρτηση καωφλίου, (b) Τμηματικά γραμμική συνάρτηση, (c)Σιγμοειδής συνάρτηση.

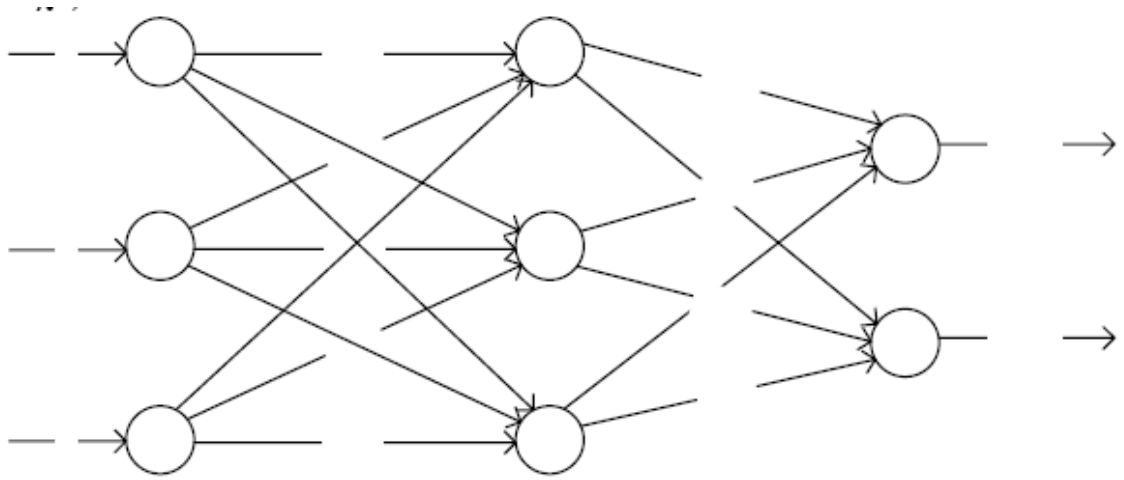
Αρχιτεκτονική νευρωνικών δικτύων

Το τεχνητό νευρωνικό δίκτυο μπορεί να θεωρηθεί ως μια συλλογή από νευρώνες αλληλοσυνδεδεμένους μεταξύ τους. Τα πλέον ευρέως χρησιμοποιούμενα νευρωνικά δίκτυα είναι τα νευρωνικά δίκτυα πολλών στρωμάτων, τα οποία είναι νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης (NNET). Τα NNET αποτελούνται από ένα στρώμα εισόδου, ένα ή περισσότερα κρυφά στρώματα και ένα στρώμα εξόδου. Κάθε στρώμα με διαφορετικά βάρη.

Οι νευρώνες στο στρώμα εισόδου είναι παθητικοί, δηλαδή ο καθένας απλά μεταδίδει μια τιμή μέσω των διασυνδετικών βαρών στους κρυφούς νευρώνες. Οι κρυφοί νευρώνες και οι νευρώνες εξόδου επεξεργάζονται τις εισόδους τους ως εξής:

Κάθε κρυφός νευρώνας πολλαπλασιάζει κάθε είσοδο με το βάρος της, προσθέτει τα γινόμενα, προσθέτει την πόλωση, και μετά περνάει το άθροισμα μέσω μιας συνάρτησης μεταφοράς (ενεργοποίησης) για να παράγει το αποτέλεσμα.

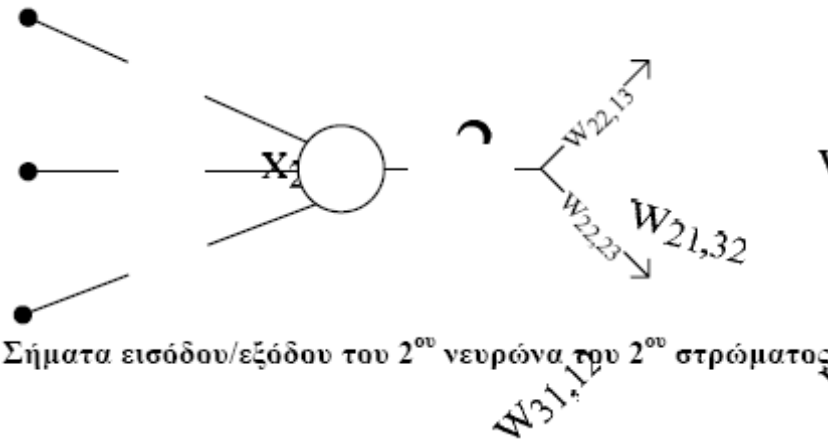
Στο Σχήμα 3.3 φαίνεται ένα νευρωνικό δίκτυο με 3 νευρώνες στο στρώμα εισόδου, 3 νευρώνες στο κρυφό στρώμα, και 2 νευρώνες στο στρώμα εξόδου. Αυτό το νευρωνικό δίκτυο συμβολίζεται ως δίκτυο αρχιτεκτονικής 3-3-2 (από τους αριθμούς των νευρώνων στο στρώμα εισόδου, το κρυφό στρώμα, και το στρώμα εξόδου αντίστοιχα).



Σχήμα 3.3. Νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης αρχιτεκτονικής 3-3-2

Το νευρωνικό αυτό δίκτυο έχει 3·3=9 συνδέσεις μεταξύ των 3 νευρώνων εισόδου και των 3 κρυφών νευρώνων, και 3·2=6 συνδέσεις μεταξύ των 3 κρυφών νευρώνων και των 2 νευρώνων εξόδου. Συνολικά το δίκτυο έχει 15 συνδέσεις (links). Σε κάθε σύνδεση αντιστοιχεί και μια τιμή βάρους $w_{pj,qk}$ που ενώνει το νευρώνα p του στρώματος j με το νευρώνα q του στρώματος k. Οι κρυφοί νευρώνες και οι νευρώνες εξόδου του νευρωνικού αυτού δικτύου έχουν μηδενικές πολώσεις. Στους κρυφούς νευρώνες και στους νευρώνες εξόδους χρησιμοποιείται κάποια συνάρτηση ενεργοποίησης (μεταφοράς), για παράδειγμα, η σιγμοειδής συνάρτηση.

Στο Σχήμα 3.4 φαίνονται αναλυτικά τα σήματα εισόδου και εξόδου του δεύτερου νευρώνα του δεύτερου στρώματος (κρυφό στρώμα) που αντιστοιχούν στο νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης στο Σχήμα 3.3. Με INP_{pj} και OUT_{pj} συμβολίζονται η συνολική είσοδος και η έξοδος, αντίστοιχα, του νευρώνα p του στρώματος j.



Σχήμα 3.4. Σήματα εισόδου/εξόδου του 2ου νευρώνα του 2ου στρώματος

Η συνολική είσοδος του δεύτερου νευρώνα του δεύτερου στρώματος είναι:

$$INP_{22} = w_{11,22} \cdot x_1 + w_{21,22} \cdot x_2 + w_{31,22} \cdot x_3$$

Επειδή στους κρυφούς νευρώνες και στους νευρώνες εξόδου χρησιμοποιείται η σιγμοειδής συνάρτησης ενεργοποίησης, η έξοδος του δεύτερου νευρώνα του δεύτερου στρώματος είναι:

$$OUT_{22} = \frac{1}{1 + e^{-INP_{22}}}$$

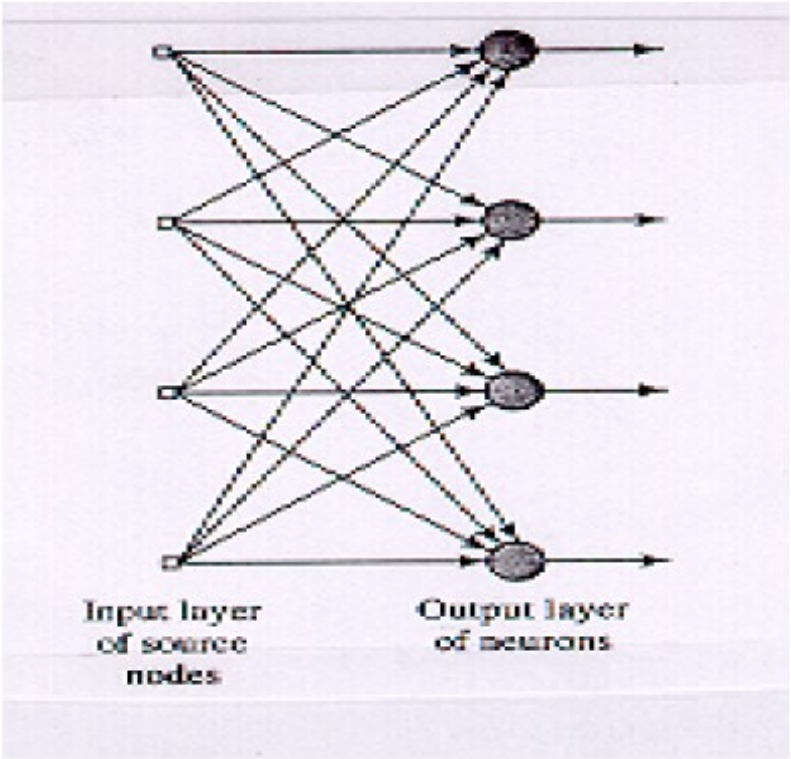
Με αντίστοιχο τρόπο παράγονται τα σήματα εξόδου και των υπόλοιπων νευρώνων του δικτύου εμπρόσθιας τροφοδότησης.

Ο τρόπος με τον οποίο οι νευρώνες ενός νευρωνικού δικτύου δομούνται, είναι στενά συνδεδεμένος με τον αλγόριθμο εκμάθησης που χρησιμοποιείται για την εκπαίδευση του δικτύου. Σαν εμπρός τροφοδότησης, αναφέρονται τα δίκτυα, στα οποία τα σήματα κατευθύνονται από την είσοδο στην έξοδο. Όταν οι έξοδοι κάποιων νευρώνων, γίνονται είσοδοι σε νευρώνες προηγούμενων επιπέδων (προς το μέρος της εισόδου του δικτύου), τότε έχουμε ανάδραση.

Μπορούμε να διακρίνουμε 4 διαφορετικές κλάσεις αρχιτεκτονικών δομών:

1.Ενός-επιπέδου Εμπρός-Τροφοδότησης Δίκτυα.

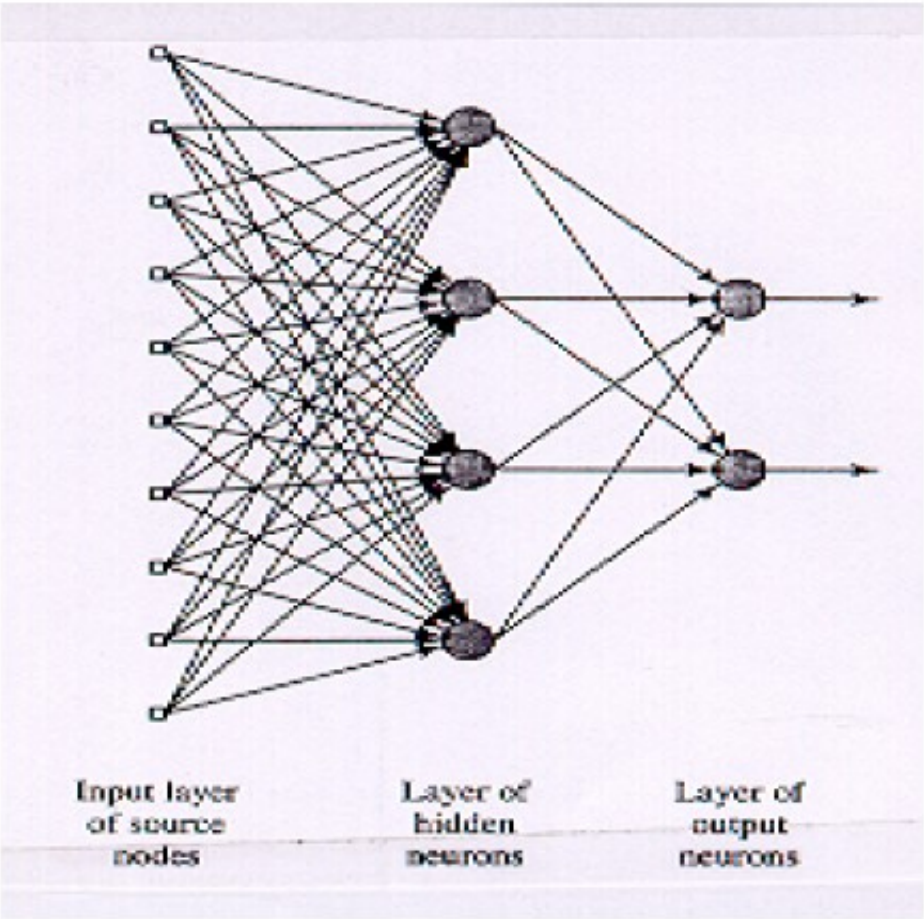
Σε ένα τέτοιο δίκτυο, οι νευρώνες είναι οργανωμένοι σε μορφή επιπέδων. Οι νευρώνες του πηγαίου επιπέδου δείχνουν στους νευρώνες του επόμενου επιπέδου αλλά όχι αντίστροφα, όπως στο Σχήμα 5.



Σχήμα 5: Εμπρός-Τροφοδότησης δίκτυο με ένα επίπεδο νευρώνων.

2.Πολλαπλών-Επιπέδων Εμπρός-Τροφοδότησης Δίκτυα.

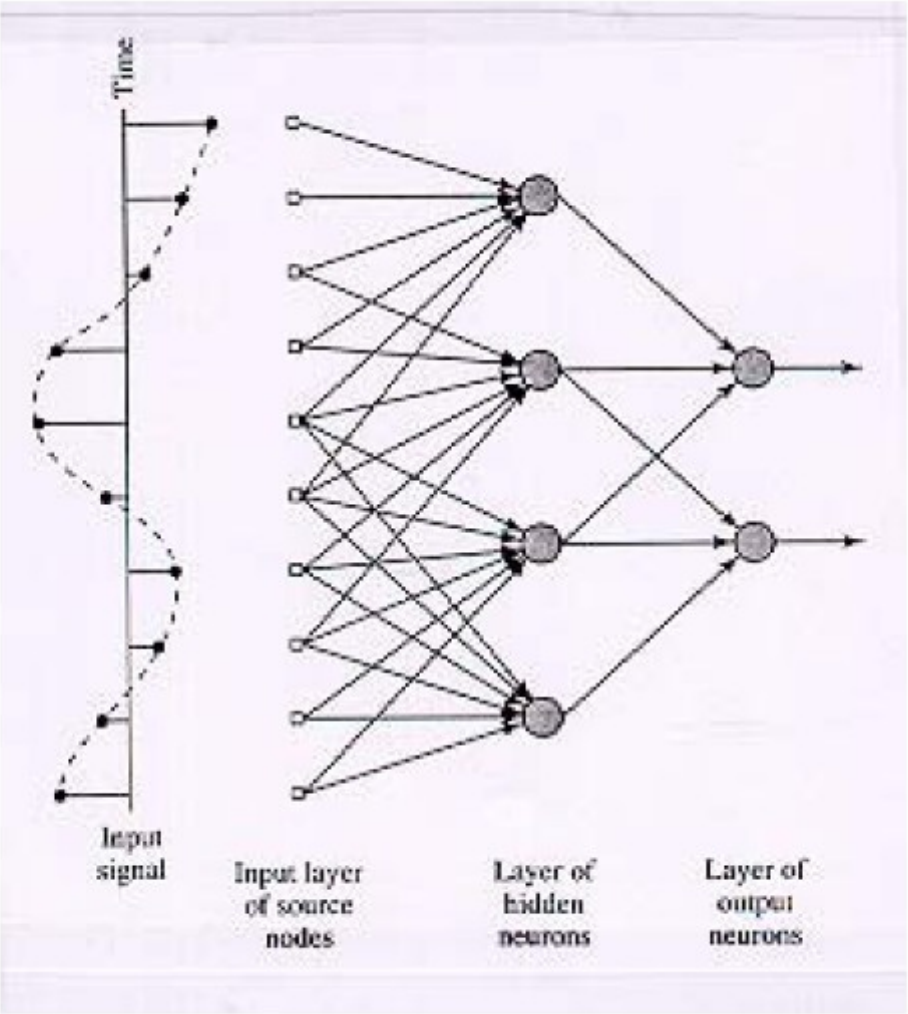
Εδώ έχουμε περισσότερα του ενός κρυφά επίπεδα, των οποίων οι κόμβοι υπολογισμού ονομάζονται ”κρυφοί νευρώνες”. Τυπικά, οι νευρώνες σε κάθε επίπεδο έχουν σαν εισόδους τα σήματα εξόδου του προηγούμενου μόνο επιπέδου.
Στο σχήμα 6 στην επόμενη σελίδα έχουμε ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο, με την έννοια ότι κάθε κόμβος συνδέεται με όλους τους κόμβους του αμέσως επόμενου επιπέδου.



Σχήμα 6: Πλήρως συνδεδεμένο δίκτυο εμπρός-τροφοδότησης με ένα κρυφό επίπεδο και ένα επίπεδο εξόδου.

Ένα τέτοιο δίκτυο περιγράφεται συνοπτικά με το συμβολισμό 10-4-2. Αυτός ο συμβολισμός σημαίνει ότι το Ν.Δ. έχει δέκα εισόδους, ένα κρυφό επίπεδο με τέσσερις νευρώνες (κόμβους) και το επίπεδο εξόδου έχει δύο νευρώνες ή κόμβους. Γενικά, ένα πολυεπίπεδο δίκτυο εμπρός τροφοδότησης με p εισόδους, m κρυφά επίπεδα με h_j, j=1, …,m κόμβους ανά επίπεδο και n κόμβους εξόδου, συμβολίζεται σαν: p-h_1,h_2,…h_m-n.

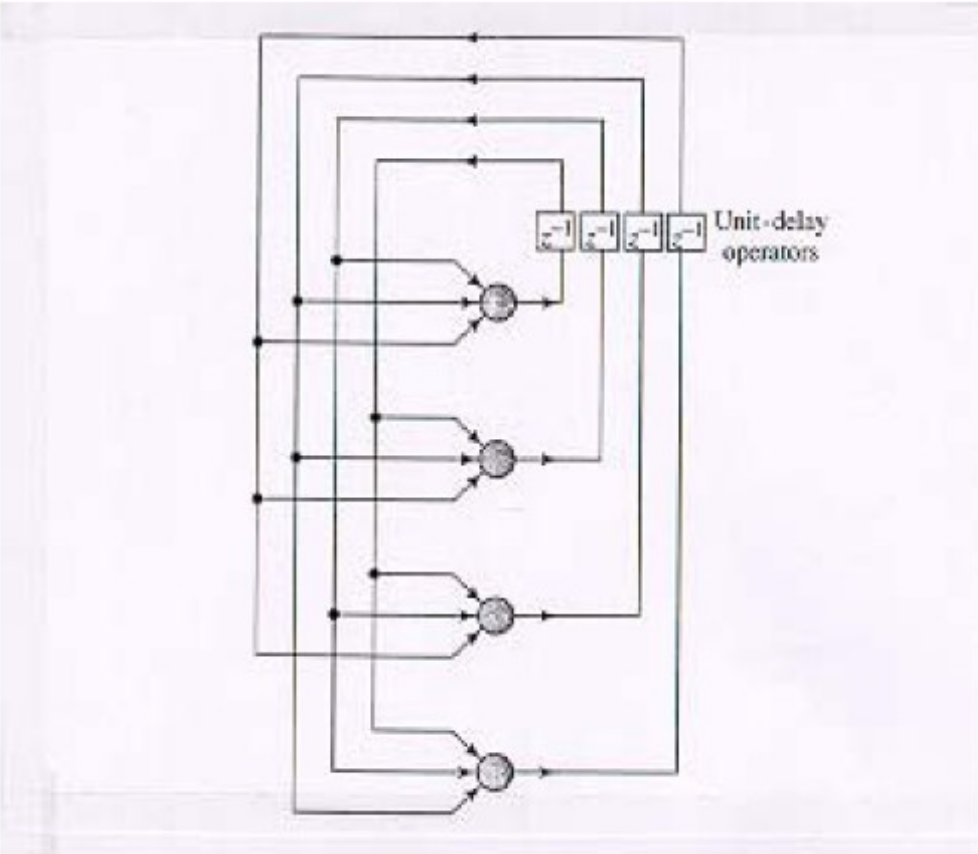
Αντίθετα στο σχήμα 7 έχουμε ένα μερικώς συνδεδεμένο νευρωνικό δίκτυο.



Σχήμα 7: Μερικώς συνδεδεμένο δίκτυο εμπρός-τροφοδότησης.

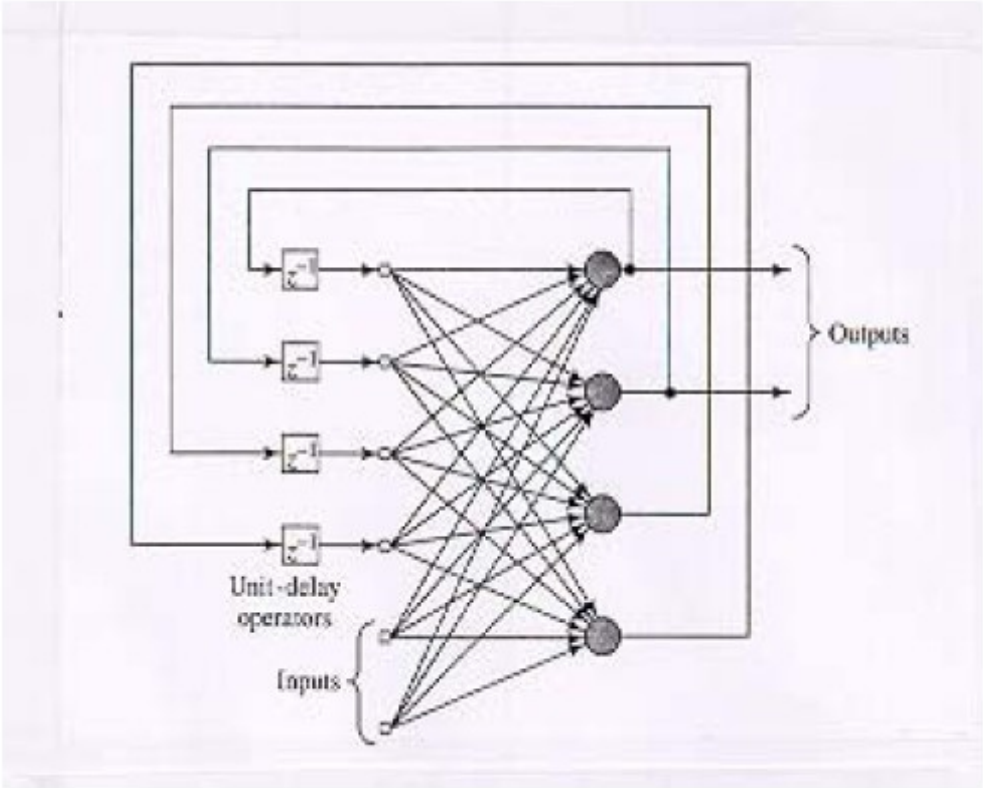
3.Αναδρομικά Δίκτυα.

Η διαφορά με τα Δίκτυα Ανατροφοδότησης είναι ότι έχει ένα τουλάχιστον βρόχο ανάδρασης. Στα σχήματα 8 και 9 που ακολουθούν στις επόμενες σελίδες φαίνονται δύο αναδρομικά δίκτυα το πρώτο χωρίς και το δεύτερο με κρυφούς νευρώνες.



Σχήμα 8: Αναδρομικό δίκτυο χωρίς ανατροφοδότηση και χωρίς κρυφούς νευρώνες.

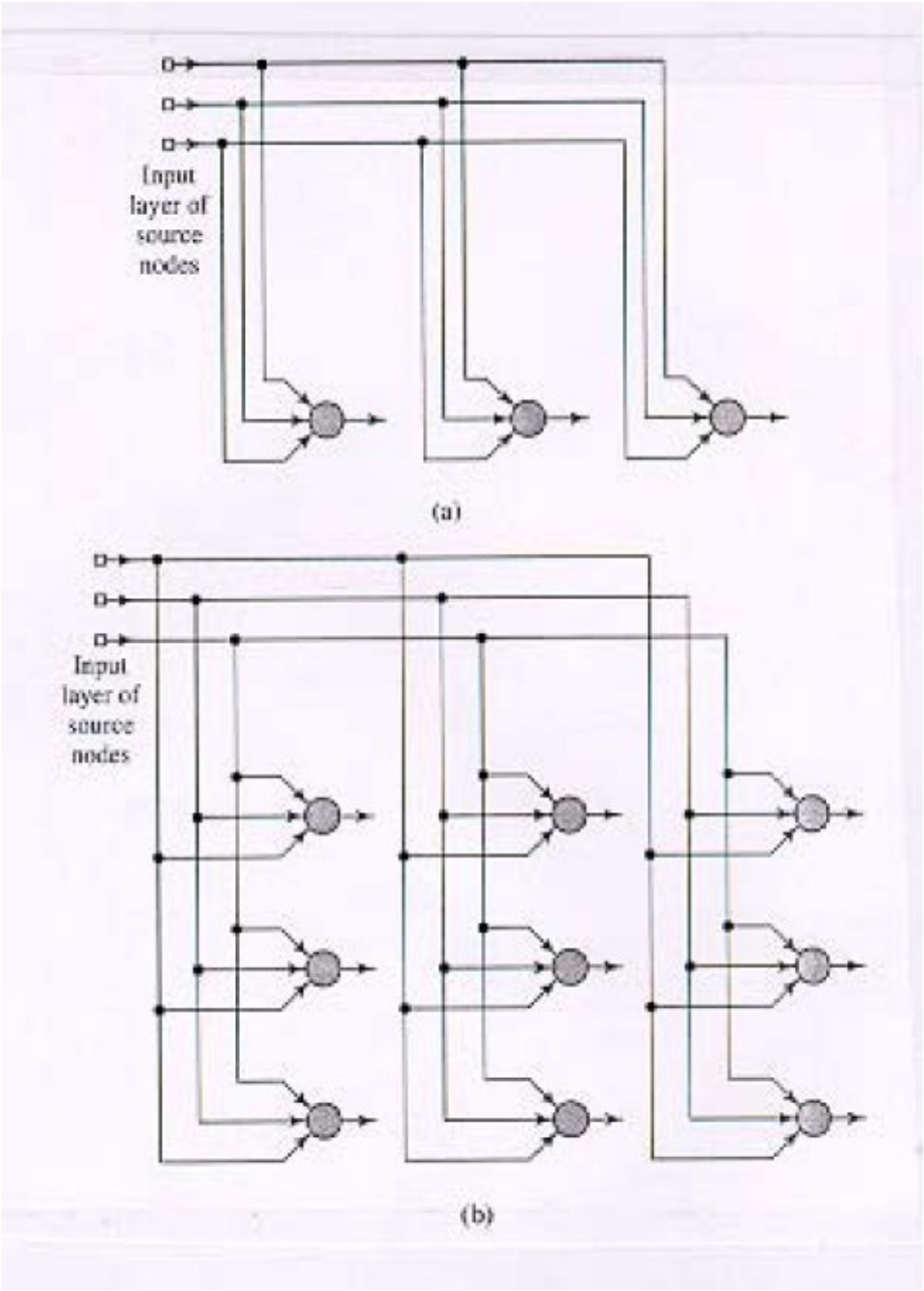
Το δίκτυο του σχήματος 8 λέγεται και δίκτυο Hopfield. Είναι μια μη-γραμμική συσχετιστική μνήμη ή μνήμη διευθυνσιοδοτούμενη από τα περιεχόμενα. Η κύρια λειτουργία μιας τέτοιας μνήμης είναι η ανάκτηση ενός προτύπου, που έχει αποθηκευτεί σε αυτήν, σαν απόκριση σε μια ελλιπή ή θορυβώδη έκδοση αυτού του προτύπου.



Σχήμα 9: Αναδρομικό δίκτυο με κρυφούς νευρώνες.

4.Δικτυωτές Δομές.

Ένα πλέγμα, αποτελείται από έναν πίνακα μιας, δύο ή μεγαλύτερης διάστασης από νευρώνες, με ένα αντίστοιχο σύνολο από πηγαίους κόμβους, που παρέχουν τα σήματα εισόδου στον πίνακα, όπως φαίνεται στο Σχήμα 10 που ακολουθεί στην επόμενη σελίδα.



Σχήμα 10: (a) Μονοδιάστατο πλέγμα με 3 νευρώνες. (b) Δυσδιάστατο πλέγμα με 3 x 3 νευρώνες.

Λειτουργία νευρωνικών δικτύων

Ένα νευρωνικό δίκτυο, έχει δύο βασικές λειτουργίες:
• Εκπαίδευση
• Πρόβλεψη

Το πρώτο στάδιο είναι η εκπαίδευση του νευρωνικού δικτύου. Στο στάδιο αυτό δημιουργείται ένα σύνολο μάθησης (training set), δηλαδή ένα σύνολο από διανύσματα εισόδων και επιθυμητών εξόδων – αποτελεσμάτων. Τα διανύσματα αυτά ονομάζονται πρότυπα εκπαίδευσης. Χρησιμοποιώντας το σύνολο μάθησης και κατάλληλο αλγόριθμο, το νευρωνικό δίκτυο εκπαιδεύεται, δηλαδή υπολογίζει τα βάρη του (και τις πολώσεις του, εφόσον υπάρχουν). Τελικός σκοπός της εκπαίδευσης του νευρωνικού δικτύου είναι η ελαχιστοποίηση του σφάλματος πρόβλεψης στο σύνολο μάθησης.

$$e = \left| \hat{y} - y \right|$$

Ένας απλοϊκός τρόπος υπολογισμού του σφάλματος φαίνεται στην παραπάνω εξίσωση. Όπου, e το σφάλμα εξόδου του νευρωνικού. Με y συμβολίζεται η προβλεπόμενη τιμή – έξοδος, από το νευρωνικό δίκτυο και y είναι η πραγματική– επιθυμητή τιμή, σύμφωνα με το σύνολο μάθησης.
Μετά την εκπαίδευση του νευρωνικού δικτύου ακολουθεί το στάδιο της πρόβλεψης. Στο στάδιο αυτό δημιουργείται ένα σύνολο ελέγχου (test set), δηλαδή ένα σύνολο από πρότυπα ελέγχου. Στο στάδιο αυτό δίνονται μόνο τα διανύσματα εισόδου στο νευρωνικό δίκτυο και αυτό υπολογίζει τα προβλεπόμενα διανύσματα εξόδου. Ο υπολογισμός αυτός γίνεται, χρησιμοποιώντας τις τιμές των βαρών (και των πολώσεων, εφόσον υπάρχουν) που υπολογίστηκαν κατά το στάδιο της εκπαίδευσης. Το σφάλμα πρόβλεψης στο σύνολο ελέγχου προκύπτει από το σφάλμα των προβλεπόμενων εξόδων του νευρωνικού δικτύου ως προς τις επιθυμητές εξόδους για κάθε ένα από τα πρότυπα ελέγχου.

Για τη βέλτιστη λειτουργία του νευρωνικού δικτύου, θα πρέπει κατά τα δύο στάδια της λειτουργίας του να ληφθούν υπόψη ορισμένοι βασικοί παράγοντες που το χαρακτηρίζουν.
• Η δομή και ο καθορισμός της αρχιτεκτονικής του νευρωνικού δικτύου. Εξετάζεται ο αριθμός των κρυφών στρωμάτων, ο αριθμός των νευρώνων ανά στρώμα, η συνάρτηση ενεργοποίησης, ο αλγόριθμος εκπαίδευσης του νευρωνικού δικτύου, ο μέγιστος αριθμός επαναλήψεων του αλγορίθμου εκπαίδευσης, κτλ.
• Ο καθορισμός της κατάλληλης δομής των συνόλων εκπαίδευσης και ελέγχου του νευρωνικού δικτύου, για παράδειγμα ο βέλτιστος αριθμός των προτύπων εκπαίδευσης.
• Η ικανότητα γενίκευσης του νευρωνικού δικτύου, η οποία διασφαλίζεται όταν το νευρωνικό δίκτυο παρουσιάζει μικρό σφάλμα πρόβλεψης τόσο στο σύνολο εκπαίδευσης όσο και στο σύνολο ελέγχου.

Η εκπαίδευση του νευρωνικού δικτύου είναι μια σχετικά χρονοβόρα επαναληπτική διαδικασία, ιδίως όταν τα πρότυπα εκπαίδευσης είναι πολλά και όταν επίσης είναι πολλοί οι νευρώνες των στρωμάτων εισόδου και εξόδου.
Ο αριθμός των νευρώνων του στρώματος εισόδου είναι ίσος με τον αριθμό των μεταβλητών εισόδου του προβλήματος πρόβλεψης. Παρόμοια, ο αριθμός των νευρώνων του στρώματος εξόδου είναι ίσος με τον αριθμό των μεταβλητών εξόδου του προβλήματος πρόβλεψης.
Ο αριθμός των νευρώνων του κρυφού στρώματος, πρέπει να προσδιοριστεί με επαναληπτικές δοκιμές, καθώς δεν υπάρχει γενική μέθοδος προσδιορισμού του. Αν οι κρυφοί νευρώνες είναι υπερβολικά λίγοι, το νευρωνικό δίκτυο δε μπορεί να μάθει τις πολύπλοκες σχέσεις μεταξύ εισόδων και εξόδων και ίσως αντιμετωπίσει πρόβλημα σύγκλισης κατά τη διάρκεια της εκπαίδευσης του. Αν ο αριθμός των κρυφών νευρώνων είναι υπερβολικά μεγάλος, η διαδικασία εκπαίδευσης θα διαρκέσει περισσότερο και ίσως επηρεάσει αρνητικά την ικανότητα γενίκευσης του νευρωνικού δικτύου.
Ο αριθμός των νευρώνων του κρυφού στρώματος μεταβάλλεται για διαφορετικές εφαρμογές και συνήθως εξαρτάται από το μέγεθος του συνόλου εκπαίδευσης και τον αριθμό των νευρώνων του στρώματος εισόδου.

Εκπαίδευση νευρωνικών δικτύων

Στόχος της εκπαίδευσης ενός νευρωνικού δικτύου είναι να υπολογιστούν τα βάρη του (και οι πολώσεις του, εφόσον υπάρχουν) έτσι ώστε να ελαχιστοποιείται το σφάλμα πρόβλεψης στο σύνολο μάθησης. Υπάρχουν πολλές μέθοδοι που μπορούν να χρησιμοποιηθούν για την εκπαίδευση ενός νευρωνικού δικτύου. Όμως δεν υπάρχει τρόπος για να προσδιοριστεί εκ των προτέρων ποια μέθοδος εκπαίδευσης λειτουργεί καλύτερα σε μια συγκεκριμένη εφαρμογή. Στη συνέχεια αναπτύσσεται ο αλγόριθμος εκπαίδευσης με προς τα πίσω διάδοση του σφάλματος ο οποίος δίνει ικανοποιητικά αποτελέσματα για ένα μεγάλο πλήθος διαφορετικών προβλημάτων. Κατά την παρουσίαση του αλγόριθμου, γίνεται η θεώρηση ότι οι κρυφοί νευρώνες και οι νευρώνες εξόδου δεν έχουν πολώσεις.

Ο αλγόριθμους της προς τα πίσω διάδοσης του σφάλματος είναι μια μέθοδος επαναληπτικής ρύθμισης των βαρών μέχρι να επιτευχθεί η επιθυμητή ακρίβεια πρόβλεψης στο σύνολο μάθησης. Στηρίζεται σε μια μέθοδο βελτιστοποίησης που εφαρμόζεται σε μια συνάρτηση σφάλματος. Μια τυπική συνάρτηση σφάλματος είναι το άθροισμα τετραγώνων των σφαλμάτων, για N συνολικό πλήθος προτύπων εκπαίδευσης.

$$e = \sum_{i=1}^{N} (OUT_{forecast,i} - OUT_{actual,i})$$

όπου $OUT_{forecast,i}$ και $OUT_{actual,i}$ είναι η προβλεπόμενη έξοδος από το νευρωνικό δίκτυο και η επιθυμητή (πραγματική) έξοδος, αντίστοιχα, του i-οστού προτύπου εκπαίδευσης.

Το σύνολο των N προτύπων εκπαίδευσης αποτελεί το σύνολο εκπαίδευσης. Μπορεί να αναπτυχθεί ένας αναδρομικός αλγόριθμος ρύθμισης των βαρών, έτσι ώστε να ελαχιστοποιεί το σφάλμα, όπως ορίζεται από την προηγούμενη εξίσωση . Οι παρακάτω σχέσεις είναι οι αναδρομικές εξισώσεις εκπαίδευσης, και αποτελούν επίλυση του προβλήματος ελαχιστοποίησης με χρήση του γενικευμένου κανόνα δέλτα:

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

$$w_{pj,qk}(n+1) = w_{pj,qk}(n) + \Delta w_{pj,qk}(n+1)$$

με,

n	ο αύξων αριθμός ανακύκλωσης της διαδικασίας εκπαίδευσης
$w_{pj,qk}(n)$	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 $w_{pj,qk}(n)$ κατά τη διάρκεια της ανακύκλωσης n
δ_{qk}	ο συντελεστής μεταβολής δ του νευρώνα q του στρώματος k
OUT_{pj}	η έξοδος του νευρώνα p του στρώματος j
lr	ο ρυθμός μάθησης
m	ο συντελεστής ορμής

Ο τρόπος υπολογισμού της τιμής του δ εξαρτάται από τη θέση του συγκεκριμένου βάρους στο δίκτυο, δηλαδή αν συνδέει νευρώνα του στρώματος εισόδου με κρυφό νευρώνα, ή κρυφό νευρώνα με νευρώνα εξόδου.

$$\delta_{qk} = OUT_{qk} \cdot (1 - OUT_{qk}) \cdot (OUT_{actual,qk} - OUT_{qk})$$

Στη προηγούμενη σχέση το στρώμα k είναι το στρώμα εξόδου και το $OUT_{actual,qk}$ είναι η επιθυμητή (πραγματική) τιμή της εξόδου οποιουδήποτε νευρώνα q του στρώματος k. Οι τιμές $OUT_{actual,qk}$ είναι γνωστές από το σύνολο εκπαίδευσης. Η υπολογιζόμενη έξοδος του δικτύου συγκρίνεται με την επιθυμητή και παράγεται ένα σήμα σφάλματος. Το σήμα του σφάλματος διαδίδεται προς τα πίσω μέσω του δικτύου για να ρυθμίσει τα βάρη, όπως φαίνεται από τις προηγούμενες εξισώσεις . Για τους νευρώνες που ανήκουν σε οποιοδήποτε άλλο στρώμα, εκτός από το στρώμα εξόδου, το σφάλμα δεν μπορεί να υπολογιστεί απευθείας επειδή δε δίνεται η επιθυμητή έξοδος για αυτούς τους 'εσωτερικούς' νευρώνες ως τμήμα του συνόλου εκπαίδευσης. Οι τιμές σφάλματος για όλους τους νευρώνες εκτός των νευρώνων εξόδου υπολογίζονται ως τα βεβαρυμμένα αθροίσματα των σφαλμάτων του στρώματος εξόδου:

$$\delta_{pj} = OUT_{pj} \cdot (1 - OUT_{pj}) \cdot \sum_{q} \delta_{qk} w_{pj,qk}$$

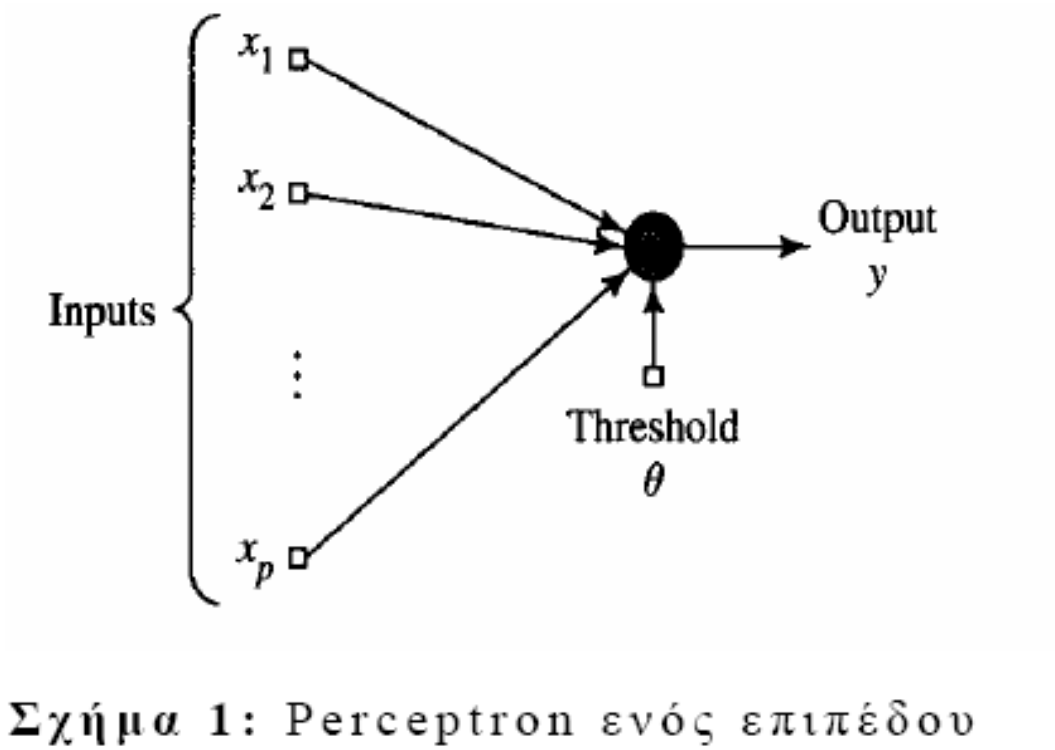
Ο συντελεστής lr ονομάζεται ρυθμός μάθησης και ελέγχει το ρυθμό με τον οποίο τα υπολογιζόμενα σφάλματα αλλάζουν τα βάρη του δικτύου. Ο ρυθμός μάθησης συνήθως επιλέγεται μεταξύ 0.01 και 1.0. Ο συντελεστής m ονομάζεται συντελεστής ορμής και επιτρέπει την αξιοποίηση της πληροφορίας από προηγούμενες τιμές σφάλματος, κατά τη διαδικασία ενημέρωσης των βαρών στη διάρκεια μιας ανακύκλωσης. Ο συντελεστής ορμής βοηθάει στην αποφυγή εγκλωβισμού σε τοπικό ελάχιστο και επιλέγεται μεταξύ 0.01 και 1.0 (αλλά σχεδόν πάντοτε μικρότερος από το ρυθμό μάθησης).
Η αναδρομική διαδικασία εκπαίδευσης (τίθεται n=n+1) εκτελείται μέχρι ένα συγκεκριμένο αριθμό επαναλήψεων είτε εναλλακτικά μέχρι να επιτευχθεί ικανοποιητική πρόβλεψη των τιμών εξόδου. Συνήθη κριτήρια τερματισμού του αλγόριθμου εκπαίδευσης περιλαμβάνουν επιτήρηση του αθροίσματος των τετραγώνων των σφαλμάτων, ή του μέγιστου σφάλματος, ή και των δύο, και η διαδικασία εκπαίδευσης σταματάει όταν η τιμή του σφάλματος είναι μικρότερη από μια ορισμένη τιμή (ανοχή). Η επιλογή της ανοχής εξαρτάται από το πρόβλημα και είναι δυνατόν να επιτευχθεί ή να μην επιτευχθεί. Δεν υπάρχει μαθηματική απόδειξη ότι ο αλγόριθμος εκπαίδευσης με προς τα πίσω διάδοση του σφάλματος θα συγκλίνει για μια δοσμένη ανοχή. Η μόνη εγγύηση είναι ότι οι αλλαγές των βαρών του ρυθμού ορμής επιτρέπει στο σφάλμα, όπως ορίστηκε προηγουμένως , να αυξηθεί προσωρινά αν η διαδικασία βελτιστοποίησης απομακρύνεται από ένα τοπικό ελάχιστο.

Παρακάτω θα παρουσιάσουμε τους δύο αλγορίθμους εκπαίδευσης απλών Ν.Δ., ενός επιπέδου. Αυτά τα δίκτυα είναι κατάλληλα για την ταξινόμηση προτύπων, που είναι γραμμικά διαχωριζόμενα. Στη συνέχεια θα ασχοληθούμε με δίκτυα πολλών επιπέδων που είναι γνωστά και σαν Perceptrons πολλών επιπέδων. Αυτά τα δίκτυα εκπαιδεύονται με τον αλγόριθμο Πίσω Διάδοσης του λάθους και είναι κατάλληλα για την ταξινόμηση προτύπων που δεν είναι γραμμικά διαχωριζόμενα. Αυτός είναι ο λόγος που αυτή η κατηγορία Τ.Ν.Δ. έχει χρησιμοποιηθεί για την επίλυση μιας μεγάλης ποικιλίας πρακτικών προβλημάτων.

Ο αλγόριθμος μάθησης του Perceptron (Αισθητήρα)

Το Perceptron είναι η απλούστερη μορφή Νευρωνικού δικτύου, το οποίο χρησιμοποιείται για την ταξινόμηση ενός ειδικού τύπου προτύπων, που είναι γραμμικά διαχωριζόμενα (δηλαδή πρότυπα που βρίσκονται στις αντίθετες πλευρές ενός υπερεπιπέδου, το οποίο ορίζει τις περιοχές απόφασης).

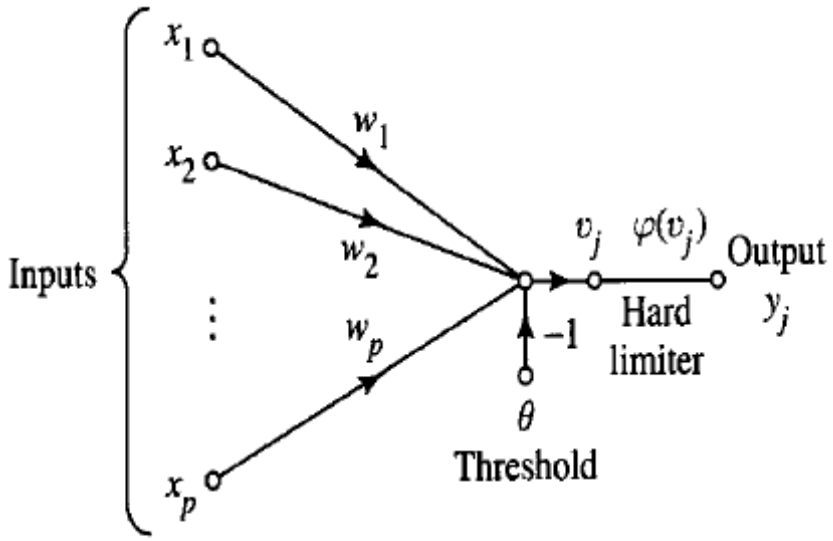
Ένα τέτοιο δίκτυο φαίνεται στο παρακάτω σχήμα 11:



Σχήμα 11: Perceptron ενός επιπέδου

Προκειμένου να εκπαιδευτεί ένα τέτοιο Ν.Δ., σαν αλγόριθμος εκπαίδευσης χρησιμοποιείται ο γνωστός κανόνας του Rosenblatt[1]. Αυτός ο κανόνας εφαρμόζεται στο γνωστό μοντέλο Mc Culloch – Pitts, για το νευρώνα και αποτελείται από ένα γραμμικό συνδυαστή ακολουθούμενο από ένα στοιχείο κατωφλίου και η παραγόμενη έξοδος παίρνει με τιμές ± 1.

Θεωρούμε το διάγραμμα ροής σήματος του Perceptron, που φαίνεται στο σχήμα 12.



Σχήμα 2: Το διάγραμμα ροής σήματος του Perceptron.

Σχήμα 12: Το διάγραμμα ροής σήματος του Perceptron

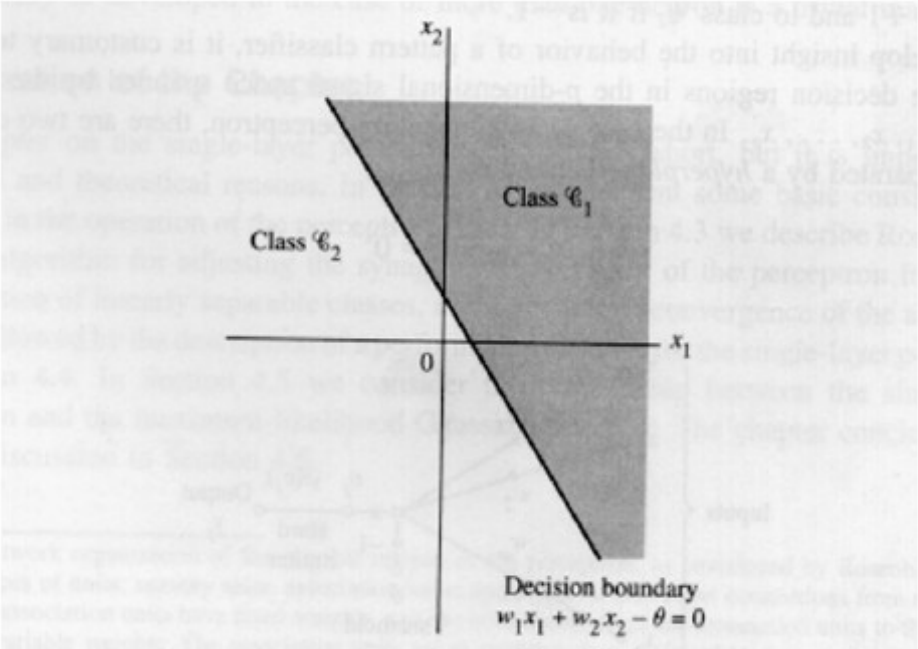
Η έξοδος του γραμμικού συνδυαστή υπολογίζεται εύκολα από το διάγραμμα του σχήματος 12 καί είναι:

$$\upsilon = \sum_{i=1}^{p} w_i x_i - \theta$$

Σκοπός του Perceptron είναι να ταξινομήσει ένα σύνολο εισόδων (προτύπων) x_1, x_2, … x_p σε μία από τις κλάσεις l_1 και l_2. Ο κανόνας απόφασης για την ταξινόμηση είναι:
ανάθεσε το σημείο που αναπαριστούν x_1, x_2, … x_p τα στην κλάση l_1, αν y = +1 και στην κλάση l_2 αν y = -1. Οι περιοχές απόφασης διαχωρίζονται από το υπερεπίπεδο που ορίζεται από τη σχέση:

$$\upsilon = \sum_{i=1}^{p} w_i x_i - \theta = 0 \quad \Leftrightarrow \quad w_1 x_1 + w_2 x_2 - \theta = 0$$

Στο σχήμα 13 στην επόμενη σελίδα φαίνεται η γραμμική διαχωρισιμότητα για ένα δισδιάστατο πρόβλημα ταξινόμησης, με δύο κλάσεις.



Σχήμα 13: Το όριο και οι περιοχές απόφασης για ένα δισδιάστατο πρόβλημα ταξινόμησης δύο κλάσεων.

Από το παραπάνω σχήμα φαίνεται το αποτέλεσμα της εφαρμογής του κατωφλίου, το οποίο μετατοπίζει το όριο απόφασης από την αρχή των αξόνων. Τα συναπτικά βάρη του Perceptron, μπορούν να προσαρμοσθούν επαναληπτικά. Για την προσαρμογή του διανύσματος βαρών w, χρησιμοποιούμε έναν κανόνα διόρθωσης λάθους, που είναι γνωστός σαν κανόνας σύγκλισης του Perceptron

Αλγόριθμος Σύγκλισης του Perceptron

Μεταβλητές και Παράμετροι

x(n) = (p + 1) x 1 input vector
[-1, x_1(n), x_2 (n), … x_ρ (n)]^T
w(n) = (p + 1) x 1 weight vector

= [θ(n), $\mathbf{w}_1$(n), $\mathbf{w}_2$ (n), … , $\mathbf{w}_p$(n)]^T
θ(n) = threshold (κατώφλι)
y(n) = actual response (πραγματική έξοδος)
d(n) = desired response (επιθυμητή έξοδος)
η = learning - rate parameter, θετική σταθερά < 1

Step 1: **Αρχικοποίηση**
Θέσε **w**(0) = 0. Κατόπιν κάνε τους υπολογισμούς για η = 1, 2, …

Step 2: Ενεργοποίηση
Στο χρόνο n, ενεργοποίησε το Perceptron εφαρμόζονται το συνεχές διάνυσμα εισόδου x(n) και το **d**(n).
Step 3: **Υπολογισμός πραγματικής απόκρισης**
Υπολόγισε την πραγματική απόκριση του Perceptron:
$\mathbf{y}(n) = \mathrm{sgn}\,[\mathbf{w}^T(n)\,\mathbf{x}(n)]$
Step 4: **Προσαρμογή διανύσματος βαρών**
Προσάρμοσε τα βάρη του Perceptron:
$\mathbf{w}(n+1) = \mathbf{w}(n) + \boldsymbol{n}\,[\mathbf{d}(n) - \mathbf{y}(n)]\,\mathbf{x}(n)$ (E.C.L. rule)
όπου:

$$\mathbf{d}(n) = \begin{cases} +1, & \text{αν } \mathbf{x}(n) \text{ ανήκει στην κλάση } l_1 \\ -1, & \text{αν } \mathbf{x}(n) \text{ ανήκει στην κλάση } l_2 \end{cases}$$

Step 5: **Αύξησε το χρόνο η κατά μια μονάδα και πήγαινε στο βήμα 2.**

Ο αλγόριθμος Ελάχιστου Μέσου Τετραγωνικού (ΕΜΤ) λάθους

Ο αλγόριθμος Ελάχιστου Μέσου Τετραγωνικού (ΕΜΤ) λάθους εκπαιδευει μία ″πρωτόγονη″ κατηγορία νευρωνικών δικτύων που αποτελούνται από ένα απλό νευρώνα και λειτουργούν κάτω από την υπόθεση της γραμμικότητας. Αυτή η κατηγορία νευρωνικών δικτύων είναι σπουδαία για τρεις λόγους:
α. Αναπτύσσεται καλά η θεωρία των γραμμικών προσαρμοζόμενων φίλτρων που χρησιμοποιούν το μοντέλο ενός απλού γραμμικού νευρώνα, με πάρα πολλές εφαρμογές, όπως ο αυτόματος έλεγχος, τα ραντάρ, τα σόναρ, κ.λ.π..
β. Είναι ένα προϊόν της πρωτοποριακής δουλειάς που έγινε στα νευρωνικά δίκτυα τη δεκαετία του 1960.
γ. Μια μελέτη των γραμμικών προσαρμοζόμενων φίλτρων ανοίγει το δρόμο για τη θεωρητική ανάπτυξη της πιο γενικής περίπτωσης των perceptrons πολλών-επιπέδων, που περιλαμβάνει τη χρήση μη-γραμμικών στοιχείων.
Στη συνέχεια διατυπώνεται ο αλγόριθμος Ελαχίστων Μέσων Τετραγώνων (Least Mean Square – LMS), που είναι επίσης γνωστός σαν Delta-rule ή σαν ο κανόνας των Widrow και Hoff (1960).
Ο αλγόριθμος LMS λειτουργεί με το μοντέλο ενός απλού γραμμικού νευρώνα, και έχει βρεί πολλές εφαρμογές. Πράγματι, ο LMS αλγόριθμος καθιερώθηκε σαν ένα σπουδαίο λειτουργικό κομμάτι στην συνεχώς επεκτεινόμενη περιοχή της προσαρμοζόμενης επεξεργασίας σημάτων.
Ο αλγόριθμος LMS βασίζεται στη χρήση στιγμιαίων εκτιμήσεων της συνάρτησης αυτοσυσχέτισης $\mathbf{r_x(j,k)}$ και της συνάρτησης ετεροσυσχέτισης $\mathbf{r_{dk}(k)}$.
Μία σύνοψη του αλγορίθμου LMS , φαίνεται στον παρακάτω πίνακα 2, από τον οποίο φαίνεται και η απλότητα του αλγορίθμου.

ΠΙΝΑΚΑΣ 2: Ο αλγόριθμος LMS.

1. Initialization. Set

$$\hat{w}_k(1) = 0 \quad \text{for } k = 1, 2, \ldots, p$$

2. Filtering. For time n= 1, 2, ... , compute:

$$y(n) = \sum_{j=1}^{p} \hat{w}_j(n) x_j(n)$$

$$e(n) = d(n) - y(n)$$

$$\hat{w}_k(n+1) = w_k(n) + \eta e(n) x_k(n) \quad \text{for } k = 1, 2, \ldots, p$$

Ο αλγόριθμος Πίσω Διάδοσης (Π.Δ.) του λάθους

Ο αλγόριθμος Πίσω Διάδοσης (Π.Δ.) του λάθους εκπαιδευει μία σπουδαία κλάση νευρωνικών δικτύων, τα δίκτυα εμπρός τροφοδότησης πολλών επιπέδων. Τυπικά ένα τέτοιο δίκτυο αποτελείται από ένα σύνολο αισθητήρων (πηγαίοι κόμβοι), που αποτελούν το επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα (hidden layers) υπολογιστικών κόμβων και ένα επίπεδο υπολογιστικών κόμβων εξόδου. Το σήμα εισόδου διαδίδεται μέσα στο δίκτυο σε μία προς τα εμπρός κατεύθυνση, από επίπεδο σε επίπεδο. Αυτά τα νευρωνικά δίκτυα αναφέρονται σαν Perceptrons πολλών επιπέδων (Multi Layer Perceptrons- MLPs) τα οποία είναι μια γενίκευση του απλού Perceptron.
Αυτός ο αλγόριθμος βασίζεται στον κανόνα μάθησης διόρθωσης του λάθους (error correction learning rule).
Βασικά η διαδικασία της πίσω διάδοσης του λάθους αποτελείται από δυο περάσματα διαμέσου των διαφορετικών επιπέδων του δικτύου ένα προς τα εμπρός πέρασμα (forward pass) και ένα προς τα πίσω πέρασμα (backward pass).
• Στο εμπρός πέρασμα ένα διάνυσμα εισόδου (input vector) εφαρμόζεται στους νευρώνες εισόδου του δικτύου, και η επίδραση του διαδίδεται μέσα στο δίκτυο από επίπεδο σε επίπεδο (layer by layer). Τελικά ένα σύνολο από εξόδους παράγεται ως η πραγματική απόκριση του δικτύου. Κατά τη διάρκεια του εμπρός περάσματος τα βάρη του δικτύου είναι σταθερά.
• Από την άλλη μεριά κατά τη διάρκεια της πίσω διάδοσης τα βάρη προσαρμόζονται σε συμφωνία με τον κανόνα διόρθωσης λάθους.
Πιο συγκεκριμένα, η πραγματική απόκριση του δικτύου αφαιρείται από την επιθυμητή απόκριση για την παραγωγή ενός σήματος λάθους, που διαδίδεται προς τα πίσω στο δίκτυο, αντίθετα από την κατεύθυνση των συνδέσεων, από το οποίο προκύπτει και

το όνομα πίσω διάδοσης του λάθους. Τα συναπτικά βάρη προσαρμόζονται έτσι ώστε να κάνουν την πραγματική απόκριση του δικτύου να πλησιάσει την επιθυμητή απόκριση.
Στην βιβλιογραφία ο αλγόριθμος πίσω διάδοσης του λάθους συχνά αναφέρεται και σαν αλγόριθμος πίσω διάδοσης (Back Propagation Algorithm) ή πιο απλά σαν Back Prop. Στο αλγόριθμο πίσω διάδοσης, η μάθηση επιτυγχάνεται εφαρμόζοντας ένα σύνολο από διανύσματα εκπαίδευσης σαν είσοδο στο πολυεπίπεδο perceptron. **Η προβολή όλων των διανυσμάτων εκπαίδευσης στο δίκτυο λέγεται ΄κύκλος΄ (epoch)**. Η διαδικασία μάθησης προχωράει από epoch σε epoch, μέχρι να σταθεροποιηθούν τα βάρη και τα κατώφλια του δικτύου και το μέσο τετραγωνικό λάθος όλων των διανυσμάτων εκπαίδευσης τείνει σε κάποια ελάχιστη τιμή . Είναι καλή πρακτική να θέτουμε τα διανύσματα εκπαίδευσης σε μια τυχαία σειρά από ένα epoch στο άλλο .
Αυτή η τυχαιότητα τείνει να κάνει το ψάξιμο στο χώρο των βαρών σε κάθε κύκλο μάθησης σε στοχαστική διαδικασία, έτσι αποφεύγουμε τον κίνδυνο να γίνουν λιγότεροι κύκλοι από ότι πρέπει. Για ένα δεδομένο σύνολο διάδοσης μπορούμε να ακολουθήσουμε ένα από τους εξής δύο τρόπους:

1. **Τρόπος Προτύπων (Pattern Mode)**. Στο pattern mode η τροποποίηση των βαρών γίνεται με την προβολή κάθε διανύσματος του συνόλου εκπαίδευσης.
2. **Σωρηδόν Τρόπος (Batch Mode)**. Στο batch mode η τροποποίηση γίνεται μετά από την προβολή όλων των διανυσμάτων εκπαίδευσης που αποτελούν ένα epoch, στο δίκτυο.
Ο αλγόριθμος πίσω-διάδοσης γενικά δεν συγκλίνει ούτε υπάρχουν σαφώς ορισμένα κριτήρια για να σταματούν την λειτουργία του, αλλά έχει λογικά κριτήρια που μπορούν να χρησιμοποιηθούν για να τερματίσουν της ρυθμίσεις τον βαρών. Για να διαμορφώσουμε ένα τέτοιο κριτήριο το λογικό είναι να σκεφτούμε σε σχέση με τις μοναδικές ιδιότητες του τοπικού ή ολικού ελαχίστου της επιφάνειας λάθους. Έστω ότι το διάνυσμα βαρών **w*** δηλώνει ένα ελάχιστο, τοπικό ή ολικό. Για να είναι το **w*** ένα ελάχιστο πρέπει το διάνυσμα κλίσης **g(w)** της επιφάνειας λάθους σε σχέση με το διάνυσμα βαρών **w** να είναι μηδέν όταν **w = w*.** Αντίστοιχα, μπορούμε να διατυπώσουμε ένα λογικό κριτήριο σύγκλισης για την μάθηση πίσω-διάδοσης όπως παρακάτω (Kramer and Sangiovanni-Vincentelli, 1989 [1]):

• *Ο αλγόριθμος πίσω-διάδοσης συγκλίνει όταν η Ευκλείδεια νόρμα του διανύσματος κλίσης φτάσει σ΄ένα αρκετά μικρό κατώφλι κλίσης.*

Το μειονέκτημα από αυτό το κριτήριο σύγκλισης είναι ότι ο χρόνος μάθησης μπορεί να είναι μεγάλος και χρειάζεται το υπολογισμό του διανύσματος κλίσης **g(w)**.
Άλλη μοναδική ιδιότητα ενός ελαχίστου είναι ότι η συνάρτηση κόστους ή μέτρο λάθους $\mathbf{E_{av}(w)}$ είναι στάσιμο στο σημείο **w=w*** και μπορούμε ως εκ τούτου να προτείνουμε ένα διαφορετικό κριτήριο σύγκλισης:

• *Ο αλγόριθμος πίσω-διάδοσης συγκλίνει όταν ο απόλυτος ρυθμός μεταβολής στο μέσο τετραγωνικό λάθος ανά κύκλο είναι αρκετά μικρός.*
Τυπικά, ο ρυθμός της μεταβολής στο μέσο τετραγωνικό λάθος θεωρείται αρκετά μικρός εάν βρίσκεται στο διάστημα 0.1 έως 1 % ανά κύκλο εκπαίδευσης , ενώ μερικές φορές χρησιμοποιείται μια τιμή που είναι αρκετή μικρή έως 0.01 % ανά κύκλο.
Ένα άλλο κριτήριο σύγκλισης του αλγορίθμου, παραλλαγή του προηγούμενου ,είναι να απαιτούμε η μέγιστη τιμή του μέσου τετραγωνικού λάθους $\mathbf{E_{av}(w)}$ να είναι ίση ή μικρότερη από ένα αρκετά μικρό κατώφλι . Οι Kramer και Sangiovanni-Vincentelli (1989) πρότειναν ένα υβριδικό κριτήριο σύγκλισης που συνίσταται απ' αυτό το τελευταίο κατώφλι και ένα κατώφλι κλίσης, όπως δηλώνεται παρακάτω:

• *Ο αλγόριθμος πίσω-διάδοσης τερματίζεται στο διάνυσμα βαρών* $\boldsymbol{w_{final}}$ *όταν* $\|\boldsymbol{g(w_{final})}\| <= \boldsymbol{\varepsilon}$*, όπου* $\boldsymbol{\varepsilon}$ *είναι ένα αρκετά μικρό κατώφλι κλίσης, ή όταν* $\boldsymbol{E_{av}(w_{final}) <= \tau}$*, όπου* $\boldsymbol{\tau}$ *είναι ένα αρκετά μικρό κατώφλι ενέργειας λάθους.*

Άλλο χρήσιμο κριτήριο σύγκλισης είναι το παρακάτω. Μετά από κάθε επανάληψη μάθησης, το δίκτυο δοκιμάζεται για την γενική του απόδοση, και αν η γενική απόδοση είναι αρκετή ή έχει κορυφωθεί τότε σταματάμε την διαδικασία μάθησης.

Εφαρμογές των Ν.Δ.

Η ευρεία εφαρμογή των νευρωνικών δικτύων σε μία ποικιλία προβλημάτων, σε πολλά διαφορετικά πεδία, τα κάνει πολύ ελκυστικά. Επίσης, οι ποιο γρήγοροι υπολογιστές και αλγόριθμοι έχουν κάνει δυνατή την χρήση νευρωνικών δικτύων στην επίλυση πολύπλοκων βιομηχανικών προβλημάτων, τα οποία απαιτούσαν πάρα πολλούς υπολογισμούς. Από την πρώτη τους εφαρμογή σε προσαρμοζόμενους εξισωτές καναλιών το 1984 μέχρι σήμερα έχουν εφαρμοστεί σε μεγάλη ποικιλία εφαρμογών όπως φαίνεται παρακάτω :
Αεροπορία: Υψηλής απόδοσης αυτόματοι πιλότοι αεροπλάνων, προσομοιωτές πτήσης, συστήματα αυτομάτου ελέγχου αεροπλάνων, συστήματα ανίχνευσης βλαβών.
Αυτοκίνηση: Αυτοκινούμενα συστήματα αυτόματης πλοήγησης.
Τραπεζικές εφαρμογές: Αναγνώστες επιταγών και άλλων παραστατικών, συστήματα αξιολόγησης αιτήσεων δανειοδότησης .
Άμυνα: Πλοήγηση όπλων, ανίχνευση στόχων, νέα είδη αισθητήρων, σόναρ, ραντάρ, ψηφιακή επεξεργασία σημάτων, συμπίεση δεδομένων, εξαγωγή χαρακτηριστικών, αναγνώριση σήματος / εικόνας.
Ηλεκτρονική: Πρόβλεψη ακολουθίας κωδίκων, μορφοποίηση ολοκληρωμένων κυκλωμάτων, έλεγχος διεργασιών, διάγνωση βλαβών ολοκληρωμένων κυκλωμάτων, μηχανική όραση, σύνθεση φωνής.
Οικονομία: Οικονομική ανάλυση, πρόβλεψη τιμών συναλλάγματος.
Κοινωνική ασφάλιση: Αξιολόγηση εφαρμοζόμενης πολιτικής, βελτιστοποίηση παραγωγής.

Βιομηχανία: Βιομηχανικός έλεγχος διεργασιών, ανάλυση και σχεδίαση προϊόντων, συστήματα ποιοτικού ελέγχου, διάγνωση βλαβών διεργασιών και μηχανών, ανάλυση σχεδιασμού χημικών προϊόντων, δυναμικό μοντελάρισμα συστημάτων χημικών διεργασιών, σχεδιασμός και διοίκηση.

Ιατρική: Ανάλυση καρκινικών κυττάρων, ανάλυση Ηλεκτροεγκεφαλογραφήματος και Ηλεκτροκαρδιογραφήματος, βελτιστοποίηση χρόνου νοσηλείας, μείωση νοσοκομειακού κόστους, βελτίωση ποιότητας νοσοκομείων.
Γεωλογικές έρευνες: Εντοπισμός πετρελαίου και φυσικού αερίου.
Ρομποτική: Έλεγχος τροχιάς και σύστημα όρασης ρομπότ.
Επεξεργασία φωνής: Αναγνώριση φωνής, συμπίεση φωνής, σύνθεση φωνής από κείμενο.
Χρηματιστηριακές εφαρμογές: Ανάλυση αγοράς, πρόβλεψη τιμών μετοχών.
Τηλεπικοινωνίες: Συμπίεση εικόνας και δεδομένων, αυτοματοποιημένες υπηρεσίες πληροφοριών, μετάφραση πραγματικού χρόνου, συστήματα επεξεργασίας πληρωμών.
Μεταφορές: Συστήματα διάγνωσης βλαβών φρένων, χρονοπρογραμματισμός οχημάτων, συστήματα δρομολόγησης.
Από το παραπάνω πλήθος εφαρμογών, μπορούμε να συμπεράνουμε ότι τα νευρωνικά δίκτυα έχουν αναπτυχθεί ταχύτατα. Γι΄αυτό το λόγο έχουν επενδυθεί πολλά χρήματα τόσο για την ανάπτυξη λογισμικού όσο και για την ανάπτυξη υλικού.

Πλεονεκτήματα των Νευρωνικών Δικτύων

Η συμπαγής παράλληλα κατανεμημένη δομή και η ικανότητα μάθησης των νευρωνικών δικτύων, κάνουν δυνατή την επίλυση πολύπλοκων προβλημάτων, τα οποία σπάνε σε μικρότερα έργα, που αναλαμβάνουν τα νευρωνικά δίκτυα ανάλογα με τις δυνατότητές τους.
Η χρήση των Νευρωνικών Δικτύων παρέχει τις ακόλουθες χρήσιμες ιδιότητες :
1. **Μη-γραμμικότητα**. Αυτή οφείλεται στο γεγονός ότι ένα Νευρωνικό Δίκτυο δομείται από τη σύνδεση νευρώνων, οι οποίοι είναι μη-γραμμικές συσκευές. Η μη- γραμμικότητα είναι πολύ σημαντική ιδιότητα, ιδιαίτερα αν ο φυσικός μηχανισμός για την παραγωγή των σημάτων εισόδου είναι μη-γραμμικός.
2. **Σχεδιασμός Εισόδου-Εξόδου.** Ένα συνηθισμένο παράδειγμα μάθησης που καλείται επιβλεπόμενη μάθηση, εμπλέκει μεταβολή των synaptic βαρών του Νευρωνικού Δικτύου, εφαρμόζοντας ένα σύνολο δειγμάτων εξάσκησης ή παραδείγματα έργων. Κάθε παράδειγμα αποτελείται από ένα μοναδικό σήμα εισόδου και την επιθυμητή απόκριση. Η εξάσκηση του δικτύου επαναλαμβάνεται για πολλά παραδείγματα, μέχρι το δίκτυο να φτάσει σε μια σταθερή κατάσταση, όπου πλέον δεν γίνονται αλλαγές στα βάρη. Έτσι το δίκτυο μαθαίνει από τα παραδείγματα, κατασκευάζοντας ένα σχεδιασμό εισόδου-εξόδου για το πρόβλημα στο χέρι.
3. **Προσαρμοστικότητα.** Τα Νευρωνικά Δίκτυα έχουν τη δυνατότητα να προσαρμόζουν τα βάρη τους στις αλλαγές του περιβάλλοντός τους. Μερικές φορές οι προσαρμογές οδηγούν σε μείωση της απόδοσης του συστήματος, γι’αυτό θα πρέπει να είναι επαρκείς δίλημμα σταθερότητας-πλαστικότητας.

4. **Αποδεικτική Απόκριση**. Ένα Νευρωνικό Δίκτυο σχεδιάζεται για να παρέχει πληροφορίες όχι μόνο για το συγκεκριμένο υπόδειγμα που επιλέγεται αλλά και για την εμπιστοσύνη στην απόφαση που παίρνεται. Αυτό βελτιώνει την απόδοση του συστήματος.
5. **Συναφής Πληροφορία.** Η γνώση αναπαριστάνεται από την πολύ δομημένη και ενεργή κατάσταση του Νευρωνικού Δικτύου.
6. **Αντοχή σε σφάλματα.** Ένα νευρωνικό δίκτυο, υλοποιημένο σε υλικό (hardware) έχει τη σημαντική ιδιότητα να είναι "ανεκτικό σε σφάλματα", με την έννοια ότι η απόδοσή του μειώνεται κάτω από αντίξοες λειτουργικές συνθήκες.
7. **Υλοποιησιμότητα σε VLSI.** Η συμπαγής παράλληλη φύση του Νευρωνικού Δικτύου, κάνει δυνατή την υλοποίηση του σε VLSI τεχνολογία, έτσι ώστε τα νευρωνικά δίκτυα να μπορούν να χρησιμοποιηθούν σε εφαρμογές πραγματικού χρόνου.
8. **Ομοιομορφία Ανάλυσης και Σχεδιασμού.** Η έννοια είναι ότι ο ίδιος συμβολισμός χρησιμοποιείται σε όλα τα πεδία που περιέχουν εφαρμογή των νευρωνικών δικτύων. Αυτό το χαρακτηριστικό υποδηλώνεται με διαφορετικούς τρόπους :
Οι νευρώνες παριστάνουν ένα συστατικό συνηθισμένο σε όλα τα νευρωνικά δίκτυα.
Αυτή η ιδιότητα κάνει δυνατό το διαμοιρασμό θεωριών και αλγορίθμων εκμάθησης σε διαφορετικές εφαρμογές των νευρωνικών δικτύων. Ρυθμιστικά (αναδρομικά) δίκτυα μπορούν να κατασκευαστούν μέσω μιας αμιγούς ολοκλήρωσης από modules.
9.**Αναλογία με Νευροβιολογία.** Ο σχεδιασμός νευρωνικών δικτύων γίνεται σε αναλογία με τον εγκέφαλο. Οι νευροβιολόγοι βλέπουν τα νευρωνικά δίκτυα σαν αντικείμενο έρευνας για την εξήγηση νευροβιολογικών φαινομένων. Ομοίως οι μηχανικοί βλέπουν στη νευροβιολογία για νέες ιδέες για την επίλυση πολύπλοκων προβλημάτων.

Κεφάλαιο 3
ΕΙΣΑΓΩΓΗ ΣΤΟΥΣ ΓΕΝΕΤΙΚΟΥΣ ΑΛΓΟΡΙΘΜΟΥΣ

Τα τελευταία τριάντα χρόνια, παρατηρείται ένα συνεχώς αυξανόμενο ενδιαφέρον για την ανάπτυξη συστημάτων επίλυσης προβλημάτων βασισμένων στις αρχές της Φυσικής Εξέλιξης. Τα συστήματα αυτού του είδους λειτουργούν διατηρώντας έναν πληθυσμό κωδικοποιημένων πιθανών λύσεων του προβλήματος που προσπαθούμε να επιλύσουμε, και εφαρμόζοντας πάνω σε αυτόν διάφορες διαδικασίες εμπνευσμένες από τη βιολογική εξέλιξη. Έτσι, περνώντας από γενιά σε γενιά, τα συστήματα αυτά δημιουργούν συνεχώς νέους πληθυσμούς πιθανών λύσεων εξελίσσοντας τους προηγούμενους πληθυσμούς.

Οι Γενετικοί Αλγόριθμοι (Genetic Algorithms) είναι ένα παράδειγμα τέτοιου συστήματος που μαζί με τον Εξελικτικό Προγραμματισμό (Evolutionary Programming.), τις Στρατηγικές Εξέλιξης (Evolution Strategies), τα Συστήματα Ταξινόμησης (Classifier Systems) και το Γενετικό Προγραμματισμό (Genetic Programming) αποτελούν μια κατηγορία συστημάτων επίλυσης προβλημάτων που είναι ευρύτερα γνωστή με τον όρο Εξελικτικοί Αλγόριθμοι (Evolutionary Algorithms).

Η πρώτη εμφάνιση των Γενετικών Αλγόριθμων (Γ.Α.) χρονολογείται στις αρχές του 1950, όταν διάφοροι επιστήμονες από το χώρο της βιολογίας αποφάσισαν να χρησιμοποιήσουν υπολογιστές στην προσπάθειά τους να προσομοιώσουν πολύπλοκα βιολογικά συστήματα. Η συστηματική τους ανάπτυξη όμως, που οδήγησε στην μορφή με την οποία είναι γνωστοί σήμερα, πραγματοποιήθηκε στις αρχές του 1970 από τον John Holland [Holland 1975] και τους συνεργάτες του στο Πανεπιστήμιο του Michigan.

Η βασική ιδέα που κρύβεται πίσω από τους Γ.Α. είναι η μίμηση των μηχανισμών της βιολογικής εξέλιξης που απαντώνται στη φύση. Ας πάρουμε, για παράδειγμα, τους λαγούς και τον τρόπο που αναπαράγονται και εξελίσσονται από γενιά σε γενιά [Michalewicz, 1996]. Έστω ότι αρχίζουμε να παρατηρούμε ένα συγκεκριμένο πληθυσμό από λαγούς σε ένα οικοσύστημα. Όπως είναι φυσικό, κάποιοι από αυτούς θα είναι πιο γρήγοροι και πιο εύστροφοι από άλλους. Αυτοί οι λαγοί έχουν περισσότερες πιθανότητες να επιβιώσουν 5 στο φυσικό τους περιβάλλον (δηλαδή να εξασφαλίζουν τροφή και να ξεφεύγουν από τα διάφορα αρπακτικά που τους καταδιώκουν, όπως για παράδειγμα τις αλεπούδες) από ότι κάποιοι πιο αργοί ή λιγότερο έξυπνοι λαγοί. Φυσικά δεν είναι λίγοι οι αργοί ή λιγότερο έξυπνοι λαγοί που καταφέρνουν να

επιβιώνουν εξαιτίας της τύχης ή άλλων παραγόντων. Όλοι αυτοί οι λαγοί, που καταφέρνουν να επιβιώσουν, θα αρχίσουν την παραγωγή της επόμενης γενιάς τους, μιας γενιάς που θα συνδυάζει με διάφορους τρόπους όλα τα χαρακτηριστικά των μελών της προηγούμενης. Έτσι, μερικοί αργοί λαγοί θα αναμειχθούν με κάποιους γρήγορους, κάποιοι γρήγοροι με άλλους γρήγορους, κάποιοι εύστροφοι λαγοί με κάποιους μη εύστροφους και ούτω καθεξής, δημιουργώντας έτσι σταδιακά έναν πληθυσμό που απαρτίζεται από λαγούς που κατά μέσο όρο είναι εξυπνότεροι και ταχύτεροι από τους προγόνους τους. Ευτυχώς όμως, για τη διατήρηση της φυσικής ισορροπίας, και τα αρπακτικά υφίστανται την ίδια διαδικασία εξέλιξης από γενιά σε γενιά, διαφορετικά οι λαγοί θα γίνονταν υπερβολικά γρήγοροι και έξυπνοι για να μπορούν να τους πιάσουν.

Οι Γ.Α. χρησιμοποιούν ορολογία δανεισμένη από το χώρο της Φυσικής Γενετικής. Αναφέρονται σε άτομα (individuals) ή γενότυπους (genotypes) μέσα σε ένα πληθυσμό. Κάθε άτομο ή γενότυπος αποτελείται από χρωμοσώματα (chromosomes). Στους Γ.Α. αναφερόμαστε συνήθως σε άτομα με ένα μόνο χρωμόσωμα. Τα χρωμοσώματα αποτελούνται από γονίδια (genes) που είναι διατεταγμένα σε γραμμική ακολουθία. Κάθε γονίδιο επηρεάζει την κληρονομικότητα ενός ή περισσότερων χαρακτηριστικών. Τα γονίδια που επηρεάζουν συγκεκριμένα χαρακτηριστικά γνωρίσματα του ατόμου βρίσκονται και σε συγκεκριμένες θέσεις του χρωματοσώματος που καλούνται loci. Κάθε χαρακτηριστικό γνώρισμα του ατόμου (όπως για παράδειγμα το χρώμα μαλλιών) έχει τη δυνατότητα να εμφανιστεί με διάφορες μορφές, ανάλογα με την κατάσταση στην οποία βρίσκεται το αντίστοιχο γονίδιο που το επηρεάζει. Οι διαφορετικές αυτές καταστάσεις που μπορεί να πάρει το γονίδιο καλούνται alleles (τιμές χαρακτηριστικού γνωρίσματος).

Κάθε γενότυπος αναπαριστά μια πιθανή λύση σε ένα πρόβλημα. Το «αποκωδικοποιημένο» περιεχόμενο ενός συγκεκριμένου χρωμοσώματος καλείται φαινότυπος (phenotype) (π.χ. ένας ζωντανός οργανισμός είναι ο φαινότυπος των 6 χρωμοσωμάτων του). Μια διαδικασία εξέλιξης που εφαρμόζεται πάνω σε ένα πληθυσμό αντιστοιχεί σε ένα εκτενές ψάξιμο στο χώρο των πιθανών λύσεων. Απαραίτητη προϋπόθεση για την επιτυχημένη έκβαση ενός τέτοιου ψαξίματος αποτελεί η εξισορρόπηση δύο διαδικασιών που είναι προφανώς αντικρουόμενες, της εκμετάλλευσης και διατήρησης των καλύτερων λύσεων, και της όσο το δυνατόν καλύτερης εξερεύνησης όλου του διαστήματος.

Οι Γ.Α. διατηρούν έναν πληθυσμό πιθανών λύσεων, του προβλήματος που μας ενδιαφέρει, πάνω στον οποίο δουλεύουν, σε αντίθεση με άλλες μεθόδους αναζήτησης που επεξεργάζονται ένα μόνο σημείο του διαστήματος αναζήτησης. Ετσι ένας Γ.Α. πραγματοποιεί αναζήτηση σε πολλές κατευθύνσεις και υποστηρίζει καταγραφή και ανταλλαγή πληροφοριών μεταξύ αυτών των κατευθύνσεων. Ο πληθυσμός υφίσταται μια προσομοιωμένη γενετική εξέλιξη. Σε κάθε γενιά, οι σχετικά "καλές" λύσεις αναπαράγονται, ενώ οι σχετικά "κακές" απομακρύνονται. Ο διαχωρισμός και η αποτίμηση των διαφόρων λύσεων γίνεται με την βοήθεια μιας αντικειμενικής συνάρτησης(objective ή fitness function), η οποία παίζει το ρόλο του περιβάλλοντος μέσα στο οποίο εξελίσσεται ο πληθυσμός.

Η δομή ενός απλού Γ.Α. έχει σε γενικές γραμμές ως εξής [Michalewicz, 1996]: Κατά την διάρκεια της γενιάς t, ο Γ.Α. διατηρεί ένα πληθυσμό P(t) από n πιθανές λύσεις (individuals): P(t)={x1 t,….,xt n}. Κάθε individual xi t αποτιμάται και δίνει ένα μέτρο της καταλληλότητας και ορθότητάς του. Αφού ολοκληρωθεί η αποτίμηση όλων των μελών του πληθυσμού, δημιουργείται ένας νέος πληθυσμός (γενιά t+1) που προκύπτει από την επιλογή των πιο κατάλληλων στοιχείων του πληθυσμού της προηγούμενης γενιάς.

Μερικά μέλη από τον καινούργιο αυτό πληθυσμό υφίστανται αλλαγές με την βοήθεια των γενετικών διαδικασιών της διασταύρωσης και της μετάλλαξης σχηματίζοντας νέες πιθανές λύσεις. Η διασταύρωση συνδυάζει τα στοιχεία των χρωμοσωμάτων δύο γονέων για να δημιουργήσει δύο νέους απογόνους ανταλλάσσοντας κομμάτια από τους γονείς. Για παράδειγμα, έστω ότι δύο γονείς αναπαριστώνται με χρωματοσώματα πέντε γονιδίων (a1,b1,c1,d1,e1) και (a2,b2,c2,d2,e2) αντίστοιχα, τότε οι απόγονοι που θα προκύψουν από διασταύρωση με σημείο διασταύρωσης (crossover point) το σημείο 2 είναι οι 7 (a1,b1,c2,d2,e2) και (a2,b2,c1,d1,e1). Διαισθητικά μπορούμε να πούμε ότι η διασταύρωση εξυπηρετεί την ανταλλαγή πληροφοριών μεταξύ διαφορετικών πιθανών λύσεων. Η διαδικασία της μετάλλαξης αλλάζει αυθαίρετα ένα ή περισσότερα γονίδια ενός συγκεκριμένου χρωμοσώματος. Πραγματοποιείται με τυχαία αλλαγή γονιδίων με πιθανότητα ίση με το ρυθμό μετάλλαξης (mutation rate). Για παράδειγμα, έστω ότι ένας individual αναπαρίσταται με το διάνυσμα πέντε διαστάσεων (a1,b1,c1,d1,e1), τότε ο individual που θα προκύψει με μετάλλαξη στη δεύτερη και στην τέταρτη διάσταση είναι ο(a1,b1 *,c1,d1 *,e1). Διαισθητικά μπορούμε να πούμε ότι η μετάλλαξη εξυπηρετεί την εισαγωγή νέων πιθανών λύσεων, διαφορετικών από τις υπάρχουσες, στον ήδη υπάρχοντα πληθυσμό.

Συνοψίζοντας μπορούμε να πούμε ότι ένας Γ.Α. για ένα συγκεκριμένο πρόβλημα πρέπει να αποτελείται από τα παρακάτω πέντε συστατικά:

1. Μια γενετική αναπαράσταση των πιθανών λύσεων του προβλήματος.
2. Έναν τρόπο δημιουργίας ενός αρχικού πληθυσμού από πιθανές λύσεις (αρχικοποίηση).
3. Μια αντικειμενική συνάρτηση αξιολόγησης των μελών του πληθυσμού, που παίζει το ρόλο του περιβάλλοντος.
4. Γενετικούς τελεστές για τη δημιουργία νέων μελών (λύσεων).
5. Τιμές για τις διάφορες παραμέτρους που χρησιμοποιεί ο Γ.Α. (μέγεθος πληθυσμού, πιθανότητες εφαρμογής των γενετικών τελεστών, κ.τ.λ.).

Πλεονεκτήματα των Γενετικών Αλγορίθμων

Μερικά από τα σημαντικότερα πλεονεκτήματα που έχει η χρήση Γ.Α. για την επίλυση προβλημάτων είναι τα εξής:

1) Μπορούν να επιλύουν δύσκολα προβλήματα γρήγορα και αξιόπιστα. Ένας από τους σημαντικούς λόγους χρήσης των Γ.Α. είναι η μεγάλη τους αποδοτικότητα. Τόσο η θεωρία, όσο και η πράξη έχουν δείξει ότι προβλήματα που έχουν πολλές, δύσκολα προσδιορισμένες λύσεις μπορούν να αντιμετωπιστούν καλύτερα από Γ.Α. Είναι δε αξιοσημείωτο ότι συναρτήσεις που παρουσιάζουν μεγάλες διακυμάνσεις και καθιστούν ανεπαρκείς άλλες μεθόδους στην εύρεση των ακρότατών τους, για τους Γ.Α. δεν αποτελούν σημεία δυσχέρειας.

2) Μπορούν εύκολα να συνεργαστούν με τα υπάρχοντα μοντέλα και συστήματα. Οι Γ.Α. προσφέρουν το σημαντικό πλεονέκτημα της χρήσης τους με προσθετικό τρόπο στα μοντέλα που χρησιμοποιούνται σήμερα, μη απαιτώντας την επανασχεδίασή τους. Μπορούν εύκολα να συνεργαστούν με τον υπάρχοντα κώδικα, χωρίς μεγάλο κόπο. Αυτό συμβαίνει, διότι χρησιμοποιούν μόνο πληροφορίες της διαδικασίας ή συνάρτησης που πρόκειται να βελτιστοποιήσουν, δίχως να ενδιαφέρει άμεσα ο ρόλος της μέσα στο σύστημα ή η όλη δομή του συστήματος.

3) Είναι εύκολα επεκτάσιμοι και εξελίξιμοι. Όπως θα γίνει σαφές στα επόμενα κεφάλαια, οι Γ.Α. δεν αντιστέκονται σε αλλαγές, επεκτάσεις και μετεξελίξεις, ανάλογα με την κρίση του σχεδιαστή. Σε πολλές εφαρμογές, έχουν αναφερθεί λειτουργίες των Γ.Α. που δεν είναι δανεισμένες από τη φύση ή που έχουν υποστεί σημαντικές αλλαγές, πάντα προς όφελος της απόδοσης. Παραλλαγές στο βασικό σχήμα δεν είναι απλά αναγκαίες, αλλά σε ορισμένες περιπτώσεις επιβάλλονται.

4) Μπορούν να συμμετέχουν σε υβριδικές μορφές με άλλες μεθόδους. Αν και η ισχύς των Γ.Α. είναι μεγάλη, σε μερικές ειδικές περιπτώσεις προβλημάτων, όπου άλλες μέθοδοι συμβαίνει να έχουν πολύ υψηλή αποδοτικότητα, λόγω εξειδίκευσης, υπάρχει η δυνατότητα χρησιμοποίησης ενός υβριδικού σχήματος Γ.Α. με άλλη μέθοδο. Αυτό είναι αποτέλεσμα της μεγάλης ευελιξίας των Γ.Α.

5) Εφαρμόζονται σε πολύ περισσότερα πεδία από κάθε άλλη μέθοδο. Το χαρακτηριστικό που τους εξασφαλίζει αυτό το πλεονέκτημα είναι η ελευθερία επιλογής των κριτηρίων που καθορίζουν την επιλογή μέσα στο τεχνικό περιβάλλον.Έτσι, Γ.Α. μπορούν να χρησιμοποιηθούν στην οικονομία, στο σχεδιασμό μηχανών, στην επίλυση μαθηματικών εξισώσεων, στην εκπαίδευση Νευρωνικών Δικτύων και σε πολλούς άλλους τομείς.

6) Δεν απαιτούν περιορισμούς στις συναρτήσεις που επεξεργάζονται. Ο κύριος λόγος που καθιστά τις παραδοσιακές μεθόδους δύσκαμπτες και ακατάλληλες για πολλά προβλήματα είναι η απαίτησή τους για ύπαρξη περιορισμών, όπως ύπαρξη παραγώγων, συνέχεια, όχι "θορυβώδεις" συναρτήσεις κ.τ.λ. Τέτοιου είδους ιδιότητες είναι αδιάφορες για τους Γ.Α. πράγμα που τους κάνει κατάλληλους για μεγάλο φάσμα προβλημάτων.

7) Δεν ενδιαφέρει η σημασία της υπό εξέταση πληροφορίας. Η μόνη "επικοινωνία" του Γ.Α. με το περιβάλλον του είναι η αντικειμενική συνάρτηση. Αυτό εγγυάται την επιτυχία του ανεξάρτητα από την σημασία του προβλήματος. Βέβαια, δεν σημαίνει ότι δεν υπάρχουν άλυτα προβλήματα για τους Γ.Α. Όπου όμως δεν τα καταφέρνουν, η αιτία είναι η φύση του χώρου που ερευνούν και όχι το πληροφοριακό περιεχόμενο του προβλήματος.

8) Έχουν από τη φύση τους τα στοιχεία του παραλληλισμού. Οι Γ.Α. σε κάθε τους βήμα επεξεργάζονται μεγάλες ποσότητες πληροφορίας, αφού κάθε άτομο θεωρείται αντιπρόσωπος πολλών άλλων. Έχει υπολογιστεί ότι η αναλογία αυτή είναι της τάξεως *O(n3)*, δηλαδή 10 άτομα αντιπροσωπεύουν περίπου 1000. Είναι, λοιπόν, προφανές ότι μπορούν να καλύψουν με αποδοτικό ψάξιμο μεγάλους χώρους σε μικρούς χρόνους.

9) Είναι μία μέθοδος που κάνει ταυτόχρονα εξερεύνηση του χώρου αναζήτησης και εκμετάλλευση της ήδη επεξεργασμένης πληροφορίας. Ο συνδυασμός αυτός σπάνια συναντάται σε οποιαδήποτε άλλη μέθοδο. Με το τυχαίο ψάξιμο γίνεται καλή εξερεύνηση του χώρου, αλλά δεν γίνεται εκμετάλλευση της πληροφορίας. Αντίθετα, με το hill-climbing γίνεται καλή εκμετάλλευση της πληροφορίας, αλλά όχι καλή εξερεύνηση. Συνήθως τα δύο αυτά χαρακτηριστικά είναι ανταγωνιστικά και το επιθυμητό είναι να συνυπάρχουν και τα δύο προς όφελος της όλης διαδικασίας. Οι Γ.Α. επιτυγχάνουν το βέλτιστο συνδυασμό εξερεύνησης και εκμετάλλευσης, πράγμα που τους κάνει ιδιαίτερα αποδοτικούς και ελκυστικούς.

10) Επιδέχονται παράλληλη υλοποίηση. Οι Γ.Α. μπορούν να εκμεταλλευτούν τα πλεονεκτήματα των παράλληλων μηχανών, αφού λόγω της φύσης τους, εύκολα μπορούν να δεχτούν παράλληλη υλοποίηση. Το χαρακτηριστικό αυτό αυξάνει ακόμη περισσότερο την απόδοσή τους, ενώ σπάνια συναντάται σε ανταγωνιστικές μεθόδους.

Τι Προκαλεί Δυσπιστία

Η τεχνολογία των Γ.Α., αν και δεν αποτελεί πρόσφατη ανακάλυψη, άρχισε ουσιαστικά να εφαρμόζεται τα τελευταία χρόνια. Η δυσπιστία με την οποία αντιμετώπιζαν οι επιστήμονες το όλο θέμα έχει αρχίσει πλέον να υποχωρεί.. Ποιοι είναι όμως οι κυριότεροι λόγοι που ίσως θα μπορούσαν να σταθούν εμπόδιο στην εξάπλωση αυτής της τεχνολογίας;

1) Προβλήματα εξοικείωσης με τη Γενετική. Για τους περισσότερους, που ασχολούνται με την Επιστήμη των Υπολογιστών, οι έννοιες της Εξέλιξης και της Φυσικής Επιλογής μπορεί να μην ηχούν παράξενα, αλλά δεν είναι και από τις πιο οικείες. Η Βιολογία δεν έχει άμεση σχέση με τους υπολογιστές, γι' αυτό και οι γνώσεις σχεδόν όλων είναι σε πολύ γενικό επίπεδο. Παρ' όλ' αυτά, δεν απαιτούνται γνώσεις Γενετικής και Βιολογίας. Εκείνο που συμβαίνει με τους Γ.Α. είναι ότι μιμούνται με αφαιρετικό τρόπο κάποιες διαδικασίες που παρατηρούνται στη φύση, χωρίς να ενδιαφέρει σε μεγάλο βαθμό λεπτομέρειας η λειτουργία τους και χωρίς να είναι απαραίτητο το γνωστικό υπόβαθρο που έχουν οι βιολόγοι για να μελετήσουν αυτά τα φαινόμενα. Οι όροι είναι δανεισμένοι από τη βιολογία με σκοπό την καλύτερη εισαγωγή και κατανόηση του θέματος και όχι την παραπομπή του μελετητή στα άγνωστα πεδία μιας ξένης επιστήμης και, τελικά, τη σύγχυσή του.

Θα μπορούσε ίσως, να παραληφθεί η αναφορά στη Γενετική και να γίνει μια παρουσίαση των Γ.Α. ως "προσωπικές διαδικασίες για αναζήτηση και βελτιστοποίηση", αυτό όμως μάλλον θα έκανε τα πράγματα δυσκολότερα. Εξάλλου, είναι συνηθισμένο το φαινόμενο θεωρίες που είναι δανεισμένες από άλλες επιστήμες να διατηρούν την αυθεντική τους ορολογία (π.χ. στα Νευρωνικά Δίκτυα: νευρώνες, συνάψεις, κ.τ.λ.). Επιπλέον, το μέλλον και η εξέλιξη των Γ.Α. δεν εξαρτώνται σε καμία περίπτωση από τις αντίστοιχες θεωρίες της Βιολογίας. Το αρχικό μοντέλο είναι δανεισμένο από εκεί, όμως η εφαρμογή του στα Τεχνητά Συστήματα έγινε με πλήθος διαφοροποιήσεων, προσαρμόσεων και "παρεκτροπών" με στόχο πάντα τη βελτίωση της απόδοσης.
Πλέον, μπορούμε να μιλάμε για εξέλιξη και απογόνους των πρώτων Γ.Α. και για μια πορεία τους στο χρόνο που είναι πλήρως ανεξάρτητη και αυτοδύναμη.

2) Το πρόβλημα του χρόνου. Στη φύση ως γνωστό, η εξέλιξη λειτουργεί με ρυθμούς πολύ αργούς. Χρειάζονται να περάσουν χιλιάδες γενιές, άρα και αρκετός χρόνος, για να αλλάξουν τα χαρακτηριστικά των ειδών και να διαφοροποιηθούν οι ικανότητες και η συμπεριφορά τους. Θέτουν έτσι ορισμένοι το ερώτημα: πώς είναι δυνατό ένα μοντέλο αναζήτησης λύσεων να έχει καλές επιδόσεις χρόνου, όταν είναι εμπνευσμένο από μια φυσική διαδικασία που εξελίσσεται με ρυθμούς απίστευτα αργούς; Η απάντηση εδώ είναι απλή. Κατ' αρχήν, ακόμη και στη φύση, η εξέλιξη δεν είναι από μόνη της μια αργή διαδικασία. Εξέλιξη των ειδών συμβαίνει όταν αλλάζει τα περιβάλλον τους και πρέπει να προσαρμοστούν στα καινούργια δεδομένα, ώστε να επιβιώσουν. Αλλαγές όμως του περιβάλλοντος γίνονται με πολύ αργούς ρυθμούς και κατά συνέπεια και η εξέλιξη ακολουθεί αυτούς τους ρυθμούς. Αν οι αλλαγές του περιβάλλοντος γίνονται με γρηγορότερο τρόπο, τότε επιταχύνεται και η εξέλιξη. Αυτό άλλωστε παρατηρείται και στα βιολογικά εργαστήρια, όπου μικροοργανισμοί αλλάζουν την συμπεριφορά τους αμέσως, όταν τοποθετούνται σε νέες συνθήκες. Επιπλέον, στο πεδίο των υπολογιστών τα άτομα κωδικοποιούνται συνήθως ως συμβολοσειρές και οι συνθήκες του περιβάλλοντος μοντελοποιούνται με απλές μαθηματικές σχέσεις. Έτσι, το μοντέλο με το οποίο δουλεύει ο υπολογιστής δεν παρουσιάζει ιδιαίτερο υπολογιστικό φόρτο, συγκρινόμενο πάντα με αντίστοιχες μεθόδους.

Το πλήθος των ατόμων που κάθε φορά εξετάζεται είναι από λίγες δεκάδες έως μερικές χιλιάδες, δηλαδή αρκετές τάξεις μεγέθους κάτω από το πλήθος των γονιδίων των χρωμοσωμάτων μιας έμβιας οντότητας. Ο ρυθμός που μπορούν να ζευγαρώνουν τα άτομα στους πιο γρήγορους υπολογιστές μπορεί να φτάσει το ένα εκατομμύριο ανά δευτερόλεπτο. Όσο μεγάλος και αν είναι ο χώρος που καλείται ο αλγόριθμος να ψάξει, η επεξεργασία μερικών μόνο ατόμων αρκεί, γιατί, όπως θα αναπτυχθεί και παρακάτω, τα άτομα αυτά θεωρούνται αντιπρόσωποι ολόκληρων κλάσεων. Έτσι λοιπόν, οι ταχύτητες που μπορούν να επιτύχουν οι Γ.Α. είναι πολύ υψηλές.

Βασικά Χαρακτηριστικά Γενετικών Αλγορίθμων

Όπως αναφέρθηκε, οι Γ.Α. πλεονεκτούν αισθητά στη λύση προβλημάτων αναζήτησης και βελτιστοποίησης από τις παραδοσιακές μεθόδους. Αυτό συμβαίνει, διότι διαφέρουν θεμελιωδώς από αυτές. Τα κυριότερα νέα χαρακτηριστικά που τους διαφοροποιούν, αλλά και τους δίνουν υπεροχή σύμφωνα με τον D. Goldberg [Goldberg, 1989], είναι τα εξής:

1) Οι Γ.Α. δουλεύουν με μια κωδικοποίηση του συνόλου τιμών που μπορούν να λάβουν οι μεταβλητές και όχι με τις ίδιες τις μεταβλητές του προβλήματος: Για παράδειγμα, αναφέρεται το εξής πρόβλημα βελτιστοποίησης: Έστω ένα μαύρο κουτί με πέντε δυαδικούς διακόπτες (on-off). Για κάθε συνδυασμό s των διακοπτών παράγεται μία έξοδος f(s). Ζητείται ο συνδυασμός των διακοπτών που μεγιστοποιεί την έξοδο. Με τις παραδοσιακές μεθόδους, το μέγιστο θα εντοπιζόταν με "παίξιμο" των διακοπτών πηγαίνοντας από συνδυασμό σε συνδυασμό με ψάξιμο στα τυφλά, καθ' ότι δεν είναι γνωστός ο τύπος της συνάρτησης. Σε ένα Γ.Α. όμως, η πρώτη ενέργεια είναι η κωδικοποίηση των διακοπτών ως συμβολοσειρές πεπερασμένου μήκους. Μια απλή κωδικοποίηση θα μπορούσε να γίνει θεωρώντας μια δυαδική συμβολοσειρά μήκους πέντε, όπου η κάθε θέση αναπαριστά ένα διακόπτη. Το 0 αντιστοιχεί στη θέση off και το 1 στη θέση on. Δηλαδή, η συμβολοσειρά 11110 κωδικοποιεί το συνδυασμό κατά τον οποίο οι πρώτοι τέσσερις διακόπτες είναι on και ο τελευταίος off. Η κωδικοποίηση δεν είναι απαραίτητο να είναι πάντα δυαδική. Όπως θα φανεί και αργότερα, μπορεί να γίνει με πολλούς τρόπους, αρκετοί από τους οποίους ίσως και να μην είναι προφανείς. Το στοιχείο της κωδικοποίησης, όπως εξηγείται παρακάτω, είναι εκείνο που επιτρέπει στους Γ.Α. να κάνουν παράλληλη επεξεργασία δεδομένων.

2) Οι Γ.Α. κάνουν αναζήτηση σε πολλά σημεία ταυτόχρονα και όχι μόνο σε ένα: Σε πολλές μεθόδους βελτιστοποίησης, η επεξεργασία γίνεται βήμα προς βήμα, πηγαίνοντας προσεκτικά από σημείο σε σημείο του πεδίου ορισμού του προβλήματος. Αυτό το βήμα προς βήμα ενέχει αρκετούς κινδύνους, ο κυριότερος από τους οποίους είναι να περιοριστεί η αναζήτηση σε μια περιοχή τοπικού ακρότατου, που δεν είναι ολικό. Οι Γ.Α. εξαλείφουν αυτόν τον κίνδυνο ενεργώντας ταυτόχρονα πάνω σε ένα ευρύ σύνολο σημείων (σύνολο από συμβολοσειρές). Έτσι μπορούν να "ανεβαίνουν" πολλούς λόφους (hill-climbing) την ίδια στιγμή, ελαχιστοποιώντας την πιθανότητα να βρουν μια λάθος κορυφή. Γυρίζοντας στο παράδειγμα με το μαύρο κουτί, οι κλασσικές μέθοδοι θα ξεκινούσαν το ψάξιμο από ένα συνδυασμό των διακοπτών και στη συνέχεια, εφαρμόζοντας κάποιο κανόνα μετάβασης, θα δοκίμαζαν τον επόμενο (ψάξιμο δηλαδή σημείο προς σημείο). Αντιθέτως, ένας Γ.Α. αρχίζει το ψάξιμό του από ένα πληθυσμό συνδυασμών συμβολοσειρών και κατόπιν παράγει διαδοχικά καινούριους. Ένας αρχικός πληθυσμός θα μπορούσε να είναι, π.χ. 01101, 11000, 01000 και 10011.

Έπειτα, "τρέχοντας" ο αλγόριθμος δημιουργεί νέους πληθυσμούς που συγκλίνουν προς την επιθυμητή λύση. Διαλέγοντας ένα πληθυσμό που να καλύπτει αντιπροσωπευτικά ένα μεγάλο εύρος τιμών μπορούν να προκύψουν ικανοποιητικά αποτελέσματα.

3) Οι Γ.Α. χρησιμοποιούν μόνο την αντικειμενική συνάρτηση και καμία επιπρόσθετη πληροφορία: Πολλές μέθοδοι αναζήτησης απαιτούν αρκετές βοηθητικές πληροφορίες για τη συνάρτηση που επεξεργάζονται. Τέτοιου είδους πληροφορίες δεν προαπαιτούνται από τους Γ.Α. Το ψάξιμό τους είναι κατά κάποιο τρόπο "τυφλό". Αξιοποιούν μόνο όση πληροφορία περιέχεται στην αντικειμενική συνάρτηση πράγμα που τους προσδίδει μεγάλη ευελιξία. Έτσι όμως, προκύπτει το ερώτημα αν συμφέρεινα αγνοούνται οι βοηθητικές πληροφορίες. Γι' αυτό ακριβώς το λόγο έχουν αναπτυχθεί μορφές Γ.Α. που αξιοποιούν και τέτοιου είδους πληροφορίες (Knowledge-Based Genetic Algorithms).

4) Οι Γ.Α. χρησιμοποιούν πιθανοθεωρητικούς κανόνες μετάβασης και όχι ντετερμινιστικούς: Η χρήση πιθανοτικών κανόνων μετάβασης είναι κυρίαρχο γνώρισμα των Γ.Α., χωρίς αυτό να σημαίνει ότι η όλη διαδικασία βαδίζει στην τύχη. Δε λαμβάνονται αποφάσεις με το "στρίψιμο ενός νομίσματος". Το στοιχείο της τύχης χρησιμοποιείται ως οδηγός για αναζήτηση σε περιοχές που αναμένεται να δώσουν καλά αποτελέσματα. Τα τέσσερα προαναφερθέντα χαρακτηριστικά συμβάλουν αποφασιστικά ώστε να έχουν οι Γ.Α. την πολυπόθητη ιδιότητα της ευρωστίας.

ΒΑΣΙΚΑ ΣΤΟΙΧΕΙΑ ΓΕΝΕΤΙΚΟΥ ΑΛΓΟΡΙΘΜΟΥ

Σε αυτή την ενότητα θα παρουσιάσουμε τα βασικά χαρακτηριστικά ενός απλού Γενετικού Αλγορίθμου μέσα από την επίλυση ενός προβλήματος βελτιστοποίησης. Έστω ότι το πρόβλημα βελτιστοποίησης που θέλουμε να επιλύσουμε είναι ένα πρόβλημα μεγιστοποίησης μιας συνάρτησης *f* - στην περίπτωση τώρα που αντιμετωπίζουμε ένα πρόβλημα ελαχιστοποίησης μιας συνάρτησης *f*, αυτό μπορούμε να το μετασχηματίσουμε σε πρόβλημα μεγιστοποίησης της συνάρτησης *g = -f*. Επιπλέον, υποθέτουμε ότι η αντικειμενική συνάρτηση *f* παίρνει μόνο θετικές τιμές - διαφορετικά μπορούμε να εισάγουμε μια θετική σταθερά C, έτσι ώστε

$$max\ g(x) = max\ \{f(x) + C\}$$

Έστω, λοιπόν, ότι θέλουμε να μεγιστοποιήσουμε μια συνάρτηση k μεταβλητών, f(x1, …,xk): Rk→R. Κάθε μεταβλητή xi παίρνει τιμές στο διάστημα Di=[ai,bi] C R και f(xi, …,xk)>0, Vxi E Di, i=1,…,k.. Επιθυμούμε να βελτιστοποιήσουμε την f με κάποια απαιτούμενη ακρίβεια q δεκαδικών ψηφίων για κάθε μεταβλητή.

Ενα από τα βασικά χαρακτηριστικά του γενετικού αλγορίθμου είναι η γενετική αναπαράσταση των υποψηφίων λύσεων, η οποία στο συγκεκριμένο πρόβλημα θα είναι η δυαδική. Για να επιτευχθεί η ζητούμενη ακρίβεια, θα πρέπει κάθε διάστημα τιμών Di=[ai,bi] να διαχωριστεί σε (bi-ai) ×10q ίσα υποδιαστήματα. Έστω mi ο μικρότερος ακέραιος για τον οποίο ισχύει (bi-ai) ×10q ≤ 2mi −1 . Τότε, η αναπαράσταση των μεταβλητών σαν δυαδικές συμβολοσειρές μήκους mi ικανοποιεί την απαίτηση για ακρίβεια q δεκαδικών ψηφίων. Η ακόλουθη σχέση μετατρέπει κάθε τέτοια δυαδική συμβολοσειρά bin_str στον αντίστοιχο πραγματικό αριθμό :

$$x_i = a_i + decimal(bin_str) \cdot \frac{b_i - a_i}{2^{m_i} - 1},$$

όπου decimal(bin_str) επιστρέφει την αντίστοιχη δεκαδική τιμή για το δυαδικό αριθμό που περιέχει η bin_str.

Κατ' αυτόν τον τρόπο, κάθε χρωμόσωμα αναπαρίσταται από μια δυαδική συμβολοσειρά μήκους

$$m = \sum_{i=1}^{k} m_i$$

Τα πρώτα m_1 δυαδικά ψηφία κωδικοποιούν τη μεταβλητή x_1, δηλαδή
το διάστημα $[a_1, b_1]$, τα επόμενα m_2 κωδικοποιούν τη x_2 στο διάστημα $[a_2, b_2]$, κ.ο.κ.
Τα βασικά βήματα του απλού Γενετικού Αλγορίθμου, που επιλύει το παραπάνω πρόβλημα μεγιστοποίησης, είναι τα εξής:

1. Δημιουργία, με τυχαίο τρόπο, ενός αρχικού πληθυσμού δυνατών λύσεων *(Αρχικοποίηση - Initialization)*.

2. Αξιολόγηση κάθε λύσης χρησιμοποιώντας τη συνάρτηση *f* σαν αντικειμενική συνάρτηση.

3. Επιλογή ενός νέου πληθυσμού με βάση την απόδοση κάθε μέλους (δυνατής λύσης) του προηγούμενου πληθυσμού *(Επιλογή - Selection)*.

4. Εφαρμογή στον πληθυσμό που προκύπτει μετά τη διαδικασία της επιλογής των γενετικών τελεστών της διασταύρωσης (*Crossover*) και της μετάλλαξης (*Mutation*).

5. Με την ολοκλήρωση του προηγούμενου βήματος, έχει δημιουργηθεί η επόμενη γενιά, οπότε επιστρέφουμε στο βήμα 2.

6. Μετά από κάποιον αριθμό γενιών, και αφού καμιά βελτίωση δεν παρατηρείται πλέον, ο Γενετικός Αλγόριθμος τερματίζεται. Το καλύτερο χρωμόσωμα αντιστοιχεί σε μια βέλτιστη λύση (πιθανώς καθολικά βέλτιστη).
Στη συνέχεια θα περιγράψουμε πιο αναλυτικά τα βήματα του αλγορίθμου.

Αρχικοποίηση

Στη φάση της αρχικοποίησης δημιουργούμε έναν αρχικό πληθυσμό από δυνατές λύσεις. Αυτό γίνεται παράγοντας τυχαία *(pop_size × m)* δυαδικά ψηφία., όπου pop_size είναι το μέγεθος του πληθυσμού που θα επεξεργαστεί ο Γ.Α. Το μέγεθος του πληθυσμού παραμένει σταθερό καθ' όλη τη διάρκεια λειτουργίας του Γ.Α.

Επιλογή

Για τη διαδικασία επιλογής ενός νέου πληθυσμού χρησιμοποιείται μια *ρουλέτα με σχισμές (slotted roulette wheel)*. Η επιλογή γίνεται με βάση την απόδοση κάθε μέλους (ατόμου - δυνατής λύσης) του πληθυσμού, έτσι όσο καλύτερο είναι κάποιο μέλος τόσο μεγαλύτερη πιθανότητα έχει να επιλεγεί και να περάσει στην επόμενη γενεά. Τα διάφορα μέλη του πληθυσμού τοποθετούνται στη ρουλέτα ανάλογα με την απόδοσή τους (μέλη με μεγάλη απόδοση κατέχουν μεγαλύτερες σχισμές της ρουλέτας).

Η κατασκευή μιας τέτοιας ρουλέτας γίνεται ως εξής:

- Υπολογίζουμε την απόδοση eval(v_i) κάθε μέλους v_i, i=1,…,pop_size του πληθυσμού.
- Υπολογίζουμε τη συνολική απόδοση του πληθυσμού $F = \sum_{i=1}^{pop_size} eval(v_i)$.
- Υπολογίζουμε την πιθανότητα επιλογής p_i κάθε μέλους v_i, i=1,…,pop_size: $p_i = eval(v_i) / F$.
- Τέλος, υπολογίζουμε την αθροιστική (cumulative) πιθανότητα q_i κάθε μέλους v_i, i=1,…,pop_size: $q_i = \sum_{j=1}^{i} p_j$.

Για την επιλογή των μελών του νέου πληθυσμού εκτελούμε pop_size περιστροφές της ρουλέτας. Αυτό γίνεται ως εξής:

1. Επιλέγουμε τυχαία έναν πραγματικό αριθμό r μεταξύ 0 και 1.
2. Αν $r<q_1$, τότε επιλέγουμε το πρώτο χρωμόσωμα v_1, διαφορετικά επιλέγουμε το v_i ($2 \le i \le pop_size$), έτσι ώστε ($q_{i-1} < r \le q_i$).

Προφανώς, με αυτή τη μέθοδο επιλογής είναι δυνατόν κάποια μέλη του πληθυσμού να επιλεχθούν περισσότερες από μία φορές, με αυτά που είχαν την καλύτερη απόδοση στην προηγούμενη γενιά να έχουν τις περισσότερες πιθανότητες γι' αυτό.

Διασταύρωση

Στη συνέχεια, εφαρμόζεται ο τελεστής της διασταύρωσης - ο οποίος στο συγκεκριμένο Γ.Α. είναι διασταύρωση ενός σημείου - στο νέο πληθυσμό. Θεωρούμε ότι η πιθανότητα κάθε μέλους του πληθυσμού να επιλεγεί για διασταύρωση είναι pc. Για κάθε μέλος του πληθυσμού κάνουμε τα εξής:

• Επιλέγουμε τυχαία έναν πραγματικό αριθμό r μεταξύ 0 και 1.
• Αν r<pc, επιλέγουμε το τρέχον μέλος του πληθυσμού για διασταύρωση.

Μετά την επιλογή μελών του πληθυσμού για διασταύρωση (ο αναμενόμενος αριθμός αυτών των μελών είναι pc × pop_size), σχηματίζουμε ζευγάρια από μέλη και για κάθε ζευγάρι επιλέγεται τυχαία ένας ακέραιος αριθμός pos στο διάστημα [1, m-1], όπου m είναι το μήκος σε δυαδικά ψηφία του (μοναδικού στην περίπτωσή μας) χρωμοσώματος κάθε μέλους. Ο αριθμός pos προσδιορίζει το σημείο διασταύρωσης. Τα επιλεγμένα ζευγάρια διασταυρώνονται και την θέση τους στον πληθυσμό την παίρνουν οι απόγονοί τους. Έτσι, η διασταύρωση των δυο παρακάτω ατόμων:

$$(b_1 b_2 \ldots b_{pos} b_{pos+1} \ldots b_m)$$

$$(c_1 c_2 \ldots c_{pos} c_{pos+1} \ldots c_m)$$

θα δώσει το ακόλουθο ζευγάρι απογόνων:

$$(b_1 b_2 \ldots b_{pos} c_{pos+1} \ldots c_m)$$

$$(c_1 c_2 \ldots c_{pos} b_{pos+1} \ldots b_m).$$

Απόγονοι οι οποίοι θα αντικαταστήσουν τους «γονείς» τους στον πληθυσμό.

Μετάλλαξη

Η μετάλλαξη επιλέγει με τυχαίο τρόπο γονίδια από τα χρωμοσώματα των μελών του πληθυσμού και μεταβάλει την τιμή τους. Επειδή στο συγκεκριμένο Γ.Α. τα γονίδια παίρνουν μόνο δυαδικές τιμές ο τελεστής της μετάλλαξης απλώς τα αντιστρέφει. Έτσι στο Γ.Α. που παρουσιάζουμε η μετάλλαξη αντιμετωπίζει τον πληθυσμό των ατόμων σαν μια ακολουθία από δυαδικά ψηφία. Κάθε δυαδικό ψηφίο έχει την ίδια πιθανότητα να επιλεγεί προκειμένου να μεταλλαχθεί. Η πιθανότητα αυτή ισούται με την πιθανότητα μετάλλαξης pm. Ο αναμενόμενος αριθμός των ανεστραμμένων ψηφίων μετά τη διαδικασία της μετάλλαξης θα είναι (pm × m × pop_size). Η διαδικασία έχει ως εξής:
Για κάθε χρωμόσωμα μέλους και κάθε γονίδιο μέσα στο χρωμόσωμα:
• Επιλέγουμε τυχαία έναν πραγματικό αριθμό r μεταξύ 0 και 1.
• Αν r<pm, τότε μεταλλάζουμε το γονίδιο (αντιστρέφουμε το δυαδικό ψηφίο).

Εφαρμογές

Οι γενετικοί αλγόριθμοι έχουν ένα ευρύ φάσμα εφαρμογών όπως φαίνεται παρακάτω:

Σχεδίαση μηχανών αεροπλάνων
Αναγνώριση προσώπων
Προγραμματισμός υπολογιστών
Συστήματα αναγνώρισης εισβολών
Σχεδίαση τηλεπικοινωνιακών δικτύων
Βέλτιστη τοποθέτηση ενεργών στοιχείων για μείωση θορύβου

Κεφάλαιο 4

Νευρωνικά δίκτυα και πρόβλεψη φορτιού

Τα τελευταία χρόνια έχει αναπτυχθεί ένα μεγάλο ενδιαφέρον σχετικά με την εφαρμογή της τεχνητής νοημοσύνης στις βιομηχανίες. Κύριο πλεονέκτημά τους αποτελεί το γεγονός ότι δεν απαιτείται περίπλοκη μαθηματική σχηματοποίηση ή ποσοτική συσχέτιση μεταξύ των εισόδων και εξόδων ενός συστήματος. Επίσης, δεν είναι απαραίτητα τα δεδομένα πολλών ετών. Έτσι, η ικανοποιητική απόδοση της τεχνητής νοημοσύνης έχει οδηγήσει σε επιτυχείς εφαρμογές σχετικά με την βραχυπρόθεσμη πρόβλεψη του ηλεκτρικού φορτίου. Πιο δημοφιλή όλων, είναι τα τεχνητά νευρωνικά δίκτυα (Artificial Neural Network, ANN). Τα νευρωνικά δίκτυα αναφέρονται σε μία κατηγορία μοντέλων τα οποία είναι εμπνευσμένα από τα βιολογικά νευρικά συστήματα. Τα μοντέλα αποτελούνται από πολλά προγραμματιζόμενα στοιχεία που συνήθως ονομάζονται νευρώνες και λειτουργούν παράλληλα. Τα στοιχεία συνδέονται με συνδετικά βάρη, κάτι το οποίο τους επιτρέπει να προσαρμόζονται μέσω μιας διαδικασίας εκμάθησης. Τα νευρωνικά δίκτυα μπορούν να ερμηνευθούν ως προσαρμόσιμες μηχανές που μπορούν να αποθηκεύουν γνώση μέσω της διαδικασίας εκμάθησης.

Η έρευνα πάνω στο πεδίο αυτό έχει μια ιστορία αρκετών δεκαετιών, αλλά, μετά από μία μείωση του ενδιαφέροντος στη δεκαετία του 1970, μία μαζική ανάπτυξη ξεκίνησε στις αρχές του 1980. Το κυρίαρχό τους πλεονέκτημα σε σχέση με τα στατιστικά μοντέλα έχει να κάνει με το γεγονός ότι μοντελοποιούν ένα πολυδιάστατο πρόβλημα χωρίς να απαιτούν περίπλοκες υποθέσεις ανάμεσα στις μεταβλητές εισόδου. Υπάρχουν πολλών τύπων μοντέλα νευρωνικών δικτύων. Το κοινό χαρακτηριστικό σε αυτά είναι η σχέση με τη γενικότερη ιδέα των βιολογικών συστημάτων. Δηλαδή, είναι ένα σύστημα επεξεργασίας δεδομένων που προσομοιώνουν τη λειτουργία του ανθρώπινου εγκεφάλου.

Παρακάτω θα δούμε νευρωνικά μοντέλα τα οποία κάνουν πρόβλεψη ηλεκτρικού φορτίου και τα οποία βρήκαμε στο διαδίκτυο στην ιστοσελίδα www.sciencedirct.com.

1.Μακροπρόθεσμα τεχνητά νευρωνικά δίκτυα

Το άρθρο[1] αναφέρεται στις δυνατότητες που έχουν κάποια τεχνητά νευρωνικά δίκτυα τα οποία κάνουν πρόβλεψη μακράς διάρκειας δηλαδή σε μεγάλο χρονικό ορίζοντα ο οποίος μπορεί να είναι από ένα έως και τριάντα χρόνια . Λόγω του μεγάλου ορίζοντα πρόβλεψης υπάρχει και μεγάλη αβεβαιότητα στα αποτελέσματα πράγμα που το μετατρέπει σε μεγάλο υπολογιστικό πρόβλημα. Τα δεδομένα τα οποία χρησιμοποιούνται είναι τα ηλεκτρικά φορτία που απαιτούνταν σε προηγούμενες χρονικές περιόδους . Η ακρίβεια των δεδομένων δεν είναι σημαντική διότι τα τεχνητά αυτά νευρωνικά δίκτυα είναι εκπαιδευμένα να αγνοούν τον θόρυβο δηλαδή το σφάλμα των δεδομένων. Ο αριθμός των νευρωνικών εισόδων του μοντέλου(δηλαδή των φορτίων των προηγουμένων χρονικών περιόδων) ποικίλουν ανάλογα τη περίπτωση ενώ η έξοδος(δηλαδή η προβλεπόμενη τιμή του φορτιού) είναι πάντα μια.. Παρακάτω δίδεται σχήμα το οποίο περιγράφει το πώς λειτουργεί το νευρωνικό μοντέλο, Yt είναι η έξοδος ενώ τα Yt-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.

Αναλόγως τον αριθμό των εισόδων στο σύστημα παίρνουμε και τα κατάλληλα αποτελέσματα ,όσο περισσότερες είσοδοι τόσο καλύτερη είναι η προβλεπόμενη τιμή δηλαδή τόσο περισσότερο πλησιάζει την πραγματική τιμή και άρα τόσο μικρότερο το σφάλμα. Σε σύγκριση με αλλά μοντέλα πρόβλεψης όπως τα AR ,ARMA ,ARIMA τα αποτελέσματα έδειξαν ότι τα τεχνητά νευρωνικά δίκτυα μπορούν να βγάλουν αποτελέσματα με ακρίβεια όταν χρησιμοποιούν ένα ελάχιστο αριθμό δεδομένων που είναι μικρότερος από τα αλλά μοντέλα.. Δηλαδή μπορούν να μας δώσουν αποτελέσματα με μικρό σφάλμα

Σχήμα 14: Προτεινομένη αρχιτεκτονική μακροπρόθεσμων τεχνητών νευρωνικών δικτύων_

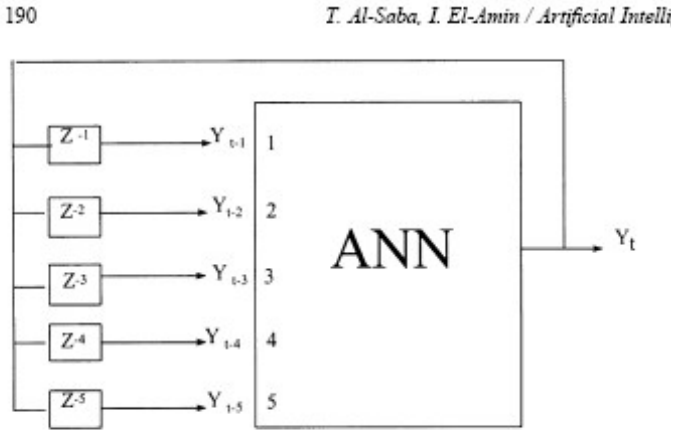


Fig. 1. Proposed ANN architecure.

2.Υβριδικά νευρωνικά δίκτυα κοντινού χρονικού ορίζοντα

Το συγκεκριμένο άρθρο [2] ασχολείται με την πρόβλεψη του απαιτούμενου ηλεκτρικού φορτιού χρησιμοποιώντας υβριδικά νευρωνικά δίκτυα δηλαδή δίκτυα τα οποία εκπαιδεύονται και δουλεύουν on-line και off-line.Τα δίκτυα αυτά είναι κοντινής διαρκείας δηλαδή έχουν σύντομο χρονικό ορίζοντα πρόβλεψης. Τα στοιχεία (δεδομένα) που χρησιμοποιούνται είναι πραγματικά στοιχεία (ηλεκτρικά φορτία) από το παρελθόν τα οποία βρέθηκαν από την τουρκική εταιρία ηλεκτρισμού. Πρόκειται για πολυεπίπεδα νευρωνικά επίπεδα τα οποία περιέχουν ένα ή περισσότερους βρόχους ανατροφοδότησης οι οποίοι μειώνουν την απαιτούμενη μνήμη σημαντικά .Στο παρακάτω σχήμα 15 φαίνεται το μοντέλο καθώς και οι βρόχοι ανατροφοδότησης .

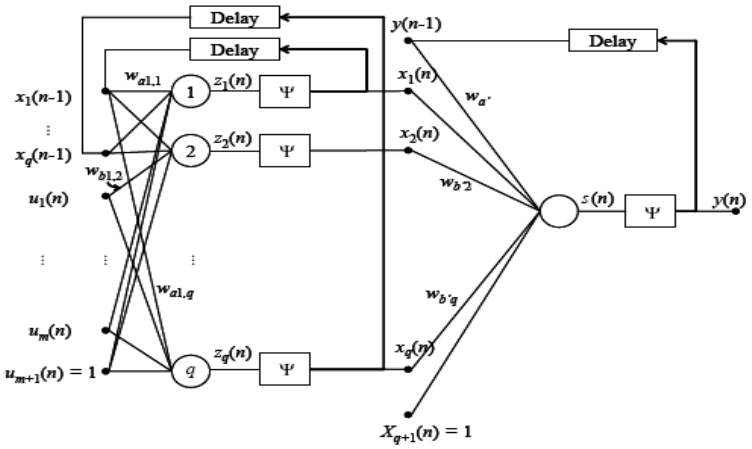


Fig. 1. Recurrent MLP structure.

Σχήμα: 15 Υβριδικά νευρωνικά δίκτυα κοντινού χρονικού ορίζοντα

Στο συγκεκριμένο 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 μειώνει 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..

3. Ιεραρχικά νευρωνικά δίκτυα κοντινού χρονικού ορίζοντα

Το άρθρο [3] παρουσιάζει ένα ιεραρχικό νευρωνικό δίκτυο(HNM) για τη λύση του προβλήματος της πρόβλεψης φορτίου κοντινού χρονικού ορίζοντα (STLF) το οποίο αποτελείται από δυο SOM (self organizing map) δίκτυα ,το ένα στην κορυφή του αλλού, όπως φαίνεται στο σχήμα. Από το μοντέλο απαιτήθηκε να προβλέψει μια φορά την ώρα το ηλεκτρικό φορτίο κατά την διάρκεια του εικοσιτετραώρου. Το μοντέλο αυτό είναι η επέκταση (συνεχεία) του SOM (self organizing map) μοντέλου του Kohonen. 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 γι΄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.

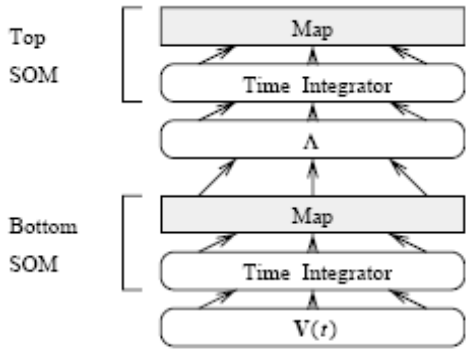


Fig. 1. HNM.

Σχήμα 16: Ιεραρχικά νευρωνικά δίκτυα κοντινού χρονικού ορίζοντα

4.Τεχνητά νευρωνικά δίκτυα που συνδυάζουν όμοιες ημερήσιες προσεγγίσεις

Το άρθρο [4]παρουσιάζει μια πρακτική μεθοδολογία λύσης του προβλήματος πρόβλεψης φορτίου σε κοντινό χρονικό ορίζοντα (STLF) χρησιμοποιώντας τεχνητά νευρωνικά δίκτυα συνδυάζοντας όμοιες ημερήσιες προσεγγίσεις. Τα νευρωνικά δίκτυα εφαρμόζονται σε παραδοσιακές μεθόδους πρόβλεψης χρησιμοποιώντας όμοια στοιχειά μεταξύ ημερών για να μάθουν την τάση της ομοιότητας . Αυτό είναι μια περιπλοκή εργασία και δεν ταιριάζει με την συνήθη εκπαίδευση των νευρωνικών δικτύων. Μια Ευκλείδεια νόρμα με σταθμισμένους (ύπαρξη βάρους) παράγοντες αξιολογεί την ομοιότητα μεταξύ της ημέρας πρόβλεψης και των προηγουμένων ημερών που έχουν ερευνηθεί. Η καμπύλη της πρόβλεψης (του φορτίου) χρησιμοποιεί τα στοιχεία από ημέρες που έχουν ίδιες συνθήκες καιρού με την ημέρα που γίνεται η πρόβλεψη. Η ακρίβεια ενισχύεται προσθέτοντας τον μείζων κλιματολογικό παράγοντα την θερμοκρασία ενώ δίνεται μεγάλη προσοχή στις διαφορετικές εποχές . Το σφάλμα της πρόβλεψης από μια έως και έξι ώρες μπροστά κυμαίνεται από 0,98% έως 2,43% . Αυτό μας δείχνει ότι η μέθοδος μας δίνει αξιόπιστες προβλέψεις για αρκετές ώρες μπροστά. Το μοντέλο αυτό δουλεύει καλά στην περίπτωση ξαφνικών κλιματολογικών αλλαγών. Στο σχήμα 17 που ακολουθεί αναπαρίσταται ένα τέτοιο δίκτυο

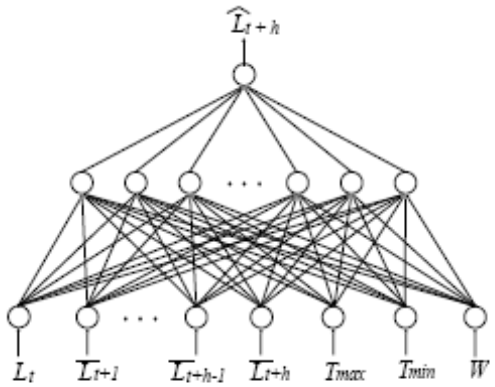


Fig. 2. Proposed neural network structure.

Σχήμα 17: Τεχνητά νευρωνικά δίκτυα που συνδυάζουν όμοιες ημερήσιες προσεγγίσεις

5. MLP τεχνητά νευρωνικά δίκτυα που προβλέπουν κατανάλωση φορτίου και καυσίμου

Το άρθρο[5] αναφέρεται στην πρόβλεψη του ηλεκτρικού φορτιού αλλά και του καυσίμου με τεχνητά νευρωνικά δίκτυα για κοντινό χρονικό ορίζοντα. Τα νευρωνικά αυτά δίκτυα έχουν τρία επίπεδα τροφοδότησης: επίπεδο εισόδου, κρυμμένο επίπεδο και επίπεδο εξόδου. Η τελευταία έκδοση τους αποτελείται από δυο διαφορετικά νευρωνικά δίκτυα, των οποίων οι προβλέψεις συνδυάζονται μεταξύ τους. Και τα δυο έχουν 79 εισόδους οι οποίες είναι: 24 ωριαία φορτία της προηγουμένης μέρας , 24 παράμετροι του καιρού, 24 ωριαίες προβλέψεις φορτίου για την επόμενη μέρα, και 7 δείκτες ένα για κάθε μέρα της εβδομάδας. Τα αποτελέσματα (έξοδοι) όμως διαφέρουν , το ένα προβλέπει το ωριαίο φορτίο της επόμενης μέρας ενώ το άλλο προβλέπει την αλλαγή στο ωριαίο φορτίο από την προηγουμένη μέρα στην επόμενη. Το μοντέλο χρησιμοποιεί ξεχωριστά σέτ 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 .

6.Τεχνητά νευρικά δίκτυα με ευρετικούς κανόνες

Σε αυτό το άρθρο παρουσιάζεται ένα τεχνητό νευρικό δίκτυο (ANN) το οποίο συνδυάζεται με ευρετικούς κανόνες για την παράγωγη ενός ισχυρού υβριδικού ευφυούς συστήματος (PREVER) για την πρόβλεψη του ηλεκτρικού φορτίου κοντινού και μέσου χρονικού ορίζοντα Οι ευρετικοί κανόνες χρησιμοποιούνται για να ρυθμίσουν τα αποτελέσματα του νευρωνικού δικτυού και να βελτιώσουν την απόδοση του συστήματος. Τα στοιχεία που χρησιμοποιήθηκαν σε αυτήν την εργασία παρθήκαν από την εταιρία CELPE και αντιστοιχούν στα ωριαία στοιχεία απαίτησης φορτίων την περίοδο από τον Ιανουάριο του 2000 μέχρι το Δεκέμβριο του 2004.Όλα τα στοιχεία ενοποιήθηκαν σε ένα ενιαίο αρχείο, όπου και ταξινομηθήκαν συμφωνά με πληροφορίες όπως το έτος, η ημέρα, ο μήνας , το φορτίο της ημέρας για κάθε ώρα (24 ώρες) και η ημέρα της εβδομάδας που προβλέπεται (Κυριακή, Δευτέρα.... Σάββατο). Η πρόβλεψη του ωριαίου φορτίου γίνεται για 3, 7, 15, 30 και 45 ημέρες μπροστά . Τα νευρωνικά αυτά δίκτυα που αναπτύσσονται έχουν 55 κόμβους στο επίπεδο εισόδου το οποία διανέμονται με τον ακόλουθο τρόπο: 24 κόμβοι αντιστοιχούν στην ωριαία κατανάλωση (ν + 1) ημέρων πριν από την ημέρα που προβλέπεται, 24 κόμβοι αντιστοιχούν στην ωριαία κατανάλωση ν ημερών πριν από την ημέρα που προβλέπεται (ν = 3 ..7 ..15 ..30, ή 45 ημέρες μπροστά), και τέλος, οι επόμενες 7 τιμές καθορίζουν την ημέρα της εβδομάδας που θα προβλεφθεί. Το επίπεδο εξόδου (πρόβλεψη) χαρακτηρίζεται από 24 τιμές, μια για κάθε ώρα της ημέρας, η οποία δείχνει την ωριαία κατανάλωση φορτίου της ημέρας που προβλέπεται.

Τα αποτελέσματα έδειξαν τη δυνατότητα και την καταλληλότητα του υβριδικού συστήματος (PREVER) έναντι του (CELPE) συστήματος της εταιρίας και των νευρωνικών δικτύων ANN που ενεργούν μόνα τους. Το υβριδικό σύστημα PREVER ήταν ακριβέστερο από τα αλλά συστήματα πρόβλεψης σε όλες τις περιπτώσεις προβλέψεων όπως φαίνεται και στο σχήμα για την πρόβλεψη 45 ημερών μπροστά.

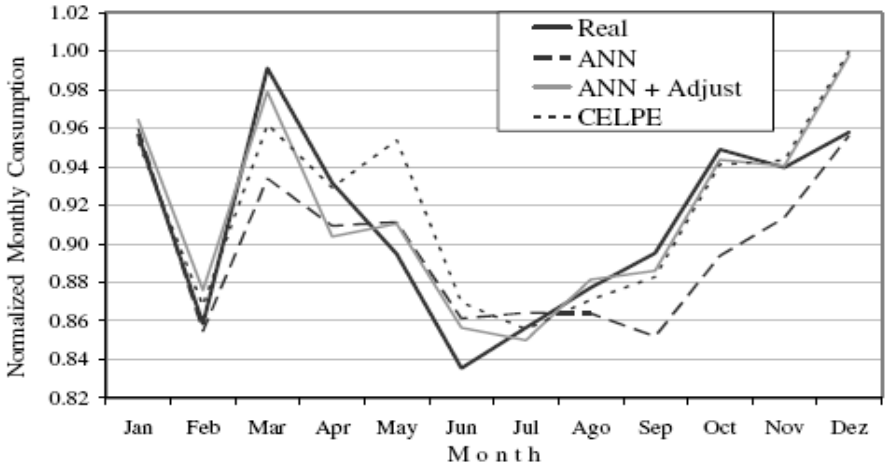


Fig. 2. Normalized Monthly Consumption in 45 days ahead

Σχήμα 18: Σφάλμα πρόβλεψης μεθόδων

7.Τεχνητά νευρωνικά δίκτυα εξελισσόμενα από γενετικούς αλγόριθμους

Αυτό το άρθρο παρουσιάζει τεχνικά νευρωνικά δίκτυα (ANN) τα οποία εξελίσσονται από έναν γενετικό αλγόριθμο για την λύση του προβλήματος της βραχυπρόθεσμης (δηλαδή κοντινού χρονικού ορίζοντα) πρόβλεψης φορτίων. Χρησιμοποιώντας τις πραγματικές τιμές φορτίων και τα προβλεπόμενα καιρικά στοιχεία τα ANN δίκτυα έχουν εκπαιδευτεί για την πρόβλεψη του ωριαίου ηλεκτρικού φορτίου μια ημέρα μπροστά με ακρίβεια οποιαδήποτε ημέρα της εβδομάδας.

Η απόδοση ενός ANN δικτυού δεν εξαρτάται μόνο από τις επιλογές των μεταβλητών εισαγωγής αλλά επίσης και από την διάσταση του.

Το μοντέλο που προτείνεται για το πρόβλημα πρόβλεψης φορτίων στο άρθρο είναι τριών επιπέδων το οποίο εκπαιδεύεται από το γενετικό αλγόριθμο.

Οι είσοδοι όπως φαίνονται στο σχήμα 19 είναι:

1. η ωριαία θερμοκρασία της ημέρας πρόβλεψης
2. μέσες ωριαίες τιμές φορτίων που λαμβάνονται με τον υπολογισμό του μέσου όρου των φορτιών των τριών προηγούμενων εβδομάδων για τον ίδιο τύπο μέρας.
3. μέση τιμή των φορτίων για την προηγούμενη ώρα, που λαμβάνεται με τον υπολογισμό του μέσου όρου των στοιχείων για τις τρεις προηγούμενες εβδομάδες.

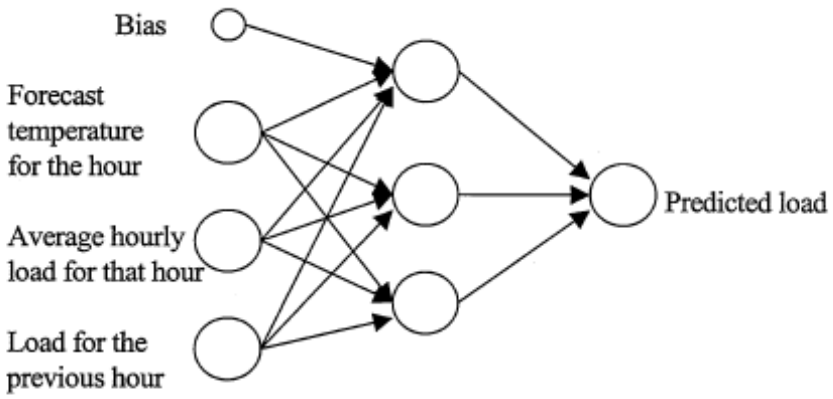


Fig. 1. The ANN model.

Σχήμα 19: Αναπαράσταση εισόδων νευρωνικού μοντέλου

Ένα σημαντικό μειονέκτημα αυτού του μοντέλου είναι ότι δεν είναι σε θέση να ανιχνεύσει ξαφνικές αλλαγές του φορτίου που μπορεί να συμβούν όπως ένα ποδοσφαιρικό παιχνιδίζου θα προκαλέσει αύξηση της κατανάλωσης φορτίου.

Το λάθος μεταξύ της πραγματικής αξίας και των τιμών πρόβλεψης υπολογίστηκε σύμφωνα με το μέσο απόλυτο σφάλμα ποσοστού (MAPE). Γενικά το μοντέλο δίνει το χαμηλότερο σφάλμα όλες τις ημέρες σε σχέση με τις στατιστικές μεθόδους που είχαν χρησιμοποιηθεί μέχρι εκείνη την περίοδο.

8.Καταλληλοτητα των MLP νευρωνικών δικτύων στην πρόβλεψη του ηλεκτρικού φορτίου

Στο συγκεκριμένο άρθρο παρουσιάζεται η λύση του προβλήματος πρόβλεψης της ηλεκτρικής ενεργειακής απαίτησης (φορτίο) σε κοντινό και μέσο-μάκρο χρονικό ορίζοντα με τη χρήση νευρωνικών δικτύων. Η πρόβλεψη της ηλεκτρικής ενεργειακής απαίτησης είναι ένα θεμελιώδες εργαλείο για τις επιχειρήσεις παραγωγής και διανομής επειδή τους παρέχει μια πρόβλεψη της ζήτησης της ηλεκτρικής ενέργειας στην αγορά.
Αυτή η εργασία στρέφεται 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 για να εκτελέσει χωριστά τις προβλέψεις της τάσης και της διακύμανσης με την χρήση νευρωνικών δικτύων. Μεταξύ των διαφορετικού τύπου νευρωνικών δικτύων που παρουσιάζει η βιβλιογραφία, το πολυεπίπεδο 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:
α) ένα επίπεδο εισαγωγής, του οποίου τα στοιχεία είναι αυτά που υποβάλλονται σε επεξεργασία
β) ένα επίπεδο εξόδου το όποιο παρέχει την πρόβλεψη του δικτύου και
γ)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 υπερβαίνει μετά βίας το 5% και είναι πάντα χαμηλότερο από 7%.

9.Ιεραρχικό υβριδικό νευρωνικό δίκτυο

Στο άρθρο 9 προτείνεται ένα νέο ιεραρχικό υβριδικό νευρωνικό πρότυπο στη λύση του προβλήματος της μακροπρόθεσμης πρόβλεψης φορτίων. Το νευρωνικό αυτό πρότυπο αποτελείται από δυο SOM (self-organizing map) δίκτυα το ένα πάνω από το άλλο και ενός απλού επιπέδου perceptron (single-layer perceptron). Το πρότυπο αυτό συγκρίνεται με ένα πολυεπίπεδο perceptron. Και τα δυο πρότυπα εκπαιδεύονται και αξιολογούνται με στοιχεία φορτίων που εξάγονται από μια βορειοαμερικανική ηλεκτρική εταιρία, και προβλέπουν είτε μία φορά κάθε εβδομάδα είτε μία φορά κάθε μήνα τη μέγιστη τιμή του φορτίου και το μέσο φορτίο κατά τη διάρκεια των επόμενων δύο ετών.
Τα δεδομένα είναι είτε εβδομαδιαία φορτία είτε μηνιαία φορτία.
Οι θερμοκρασίες δεν περιλήφθηκαν στα δεδομένα εισόδου.
Τρεις νευρωνικές είσοδοι χρησιμοποιούνται :
α) η πρώτη είσοδος αντιπροσωπεύει είτε την μέγιστη τιμή του φορτίου είτε το μέσο φορτίο τη τρέχουσας εβδομάδας ή του τρέχοντα μήνα.
β)Οι δεύτερη και τρίτη είσοδος αντιπροσωπεύουν μια τριγωνομετρική κωδικοποίηση για την εβδομάδα ή το μήνα που θα γίνει η πρόβλεψη.

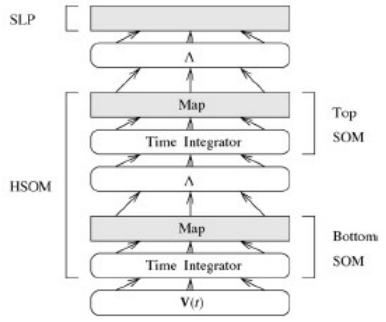


Fig. 1. HHNM.

Σχήμα 20: Ιεραρχικό υβριδικό νευρωνικό δίκτυο

Τα πειράματα δείχνουν ότι η απόδοση του ιεραρχικού υβριδικού αυτού μοντέλου (HHNN) στη μακροπρόθεσμη πρόβλεψη του μέγιστου αλλά και μέσου φορτιού είναι πολύ καλύτερη από αυτή του πολυεπίπεδου perceptron (MLP) . Η καλύτερη απόδοση που επιδεικνύεται από το υβριδικό μοντέλο (HHNN) φαίνεται να δικαιολογείται από την ιεραρχική δομή του.
Τα αποτελέσματα που επιτυγχάνονται από το υβριδικό ιεραρχικό μοντέλο (HHNN) έχουν ακόμα περιθώρια για βελτίωση η οποία μπορεί να έρθει κάνοντας λεπτές ρυθμίσεις στις παραμέτρους αλλά και κάνοντας χρήση της προεπεξεργασίας των στοιχείων .

10.Τεχνητά νευρωνικά δίκτυα εκπαιδευόμενα με την τεχνική βελτιστοποίησης σμηνών μορίων

Το άρθρο [10] συζητά και παρουσιάζει τα αποτελέσματα της χρησιμοποίησης νευρικών δικτύων για την λύση του προβλήματος της πρόβλεψης της ιορδανικής απαίτησης ηλεκτρικής ενέργειας σε κοντινό χρονικό ορίζοντα (STLF), τα οποία εκπαιδεύονται από την τεχνική βελτιστοποίησης σμηνών μορίων όπως στο σχήμα 21(particle swarm optimization technique) που είναι ένας νέος προσαρμοστικός αλγόριθμος βασιζόμενος σε μια κοινωνικό-ψυχολογική μεταφορά. Τα αποτελέσματα της χρησιμοποίησης αυτής της τεχνικής συγκρίνονται με τα αποτελέσματα της χρησιμοποίησης του αλγορίθμου της ανάστροφης διάδοσης (back-propagation) και της μεθόδου αυτοαναδρομής κινούμενου μέσου όρου (autoregressive moving average method).Το σύνολο των στοιχείων (δεδομένων) διαιρείται σε δύο τμήματα α) κατάρτιση και β) δοκιμή. Τα πρώτα 671 στοιχείων για την κατάρτιση, τα επόμενα 30 για την επικύρωση και τη δοκιμή. Κάθε δείγμα κατάρτισης περιλαμβάνει την ηλεκτρική απαίτηση(φορτίο) των προηγούμενων 2 ημερών ,την ελάχιστη καθημερινή θερμοκρασία, αλλά και δείκτες της ημέρας του μήνα και της πρόβλεψης του μελλοντικού φορτίου που απαιτείται. Ο αλγόριθμος PSO (particle swarm optimization) αποδίδει καλύτερα από τον αλγόριθμο BP(back-propagation) στα εκπαιδευτικά και εξεταστικά σύνολα, αλλά και από την τεχνική ARMA όπως φαίνεται και στο πίνακα 3 Αν και αλγόριθμος PSO απαιτεί γενικά πολύ περισσότερες συναρτήσεις αξιολόγησης για να βρει μια βέλτιστη λύση, σε σύγκριση με τον αλγόριθμο BP είναι μια ελπιδοφόρα προσέγγιση λόγω της αποτελεσματικότητάς του στην έρευνα των πολύ μεγάλων διαστημάτων και στη δυνατότητα του για σφαιρική αναζήτηση του καλύτερου μοντέλου πρόβλεψης.

Πίνακας 3

Table 1
The comparison between forecasted and actual measured load demand for different training algorithms

	Training set			Test set		
	PSO	BP	ARMA	PSO	BP	ARMA
η	0.67	0.49	0.44	0.38	0.32	0.018
MAPE (%)	2.42	3.2	3.98	2.52	2.82	3.93
MAXIMAL (MW)	57.5	131.36	135.31	61.43	100.05	176.6

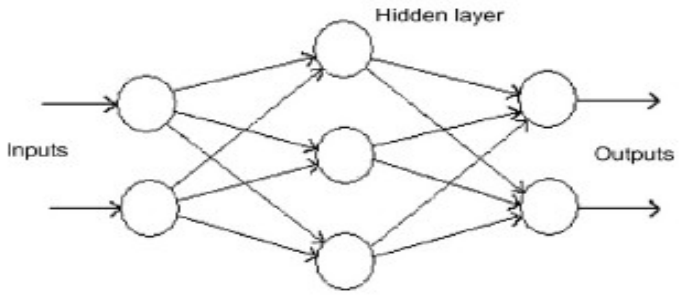


Fig. 1. Network configuration.

Σχήμα 21: Τεχνητά νευρωνικά δίκτυα εκπαιδευόμενα με την τεχνική βελτιστοποίησης σμηνών μορίων

11.Σύγκριση τεχνικών πρόβλεψης

Αυτό το άρθρο παρουσιάζει την ανάπτυξη και την εφαρμογή των προηγμένων νευρωνικών δικτύων για την αντιμετώπιση του προβλήματος της πρόβλεψης του ηλεκτρικού φορτίου για κοντινό χρονικό ορίζοντα(STLF). Διάφορες προσεγγίσεις συμπεριλαμβανομένου του Γκαουσσιανου αλγορίθμου ανάστροφης διάδοσης (Back Propagation, BP) ,τα δίκτυα τυχαίας ενεργοποίησης βαρών (RAWN) και εκπαιδευόμενα (RAWN) με κινούμενο παράθυρο (MWRAWN),τα δίκτυα με συναρτήσεις radial basis (RBFN) ,τα επαναληπτικά δίκτυα πραγματικού χρόνου(RT-RN),τα αναδρομικά επαναληπτικά νευρωνικά δίκτυα (ARNN) αλλά και οι παραλλαγές τους προτείνονται και συγκρίνονται σε αυτό το έγγραφο. Η απόδοση αυτών των δομών αξιολογείται με τη βοήθεια μιας εκτενούς μελέτης προσομοίωσης, στην οποία χρησιμοποιούνται πραγματικά ωριαία φορτία από το σύστημα ισχύος του νησιού της Κρήτης, στην Ελλάδα. Τα προτεινόμενα πρότυπα 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 μια προσέγγιση παράλληλης επεξεργασίας για πρόβλεψη 24 ωρών και εφαρμόζεται. Σύμφωνα με αυτήν την διαδικασία, το ζητούμενο φορτίο για κάθε συγκεκριμένη ώρα προβλέπεται, όχι μόνο χρησιμοποιώντας τα φορτία για αυτήν την συγκεκριμένη ώρα από τις προηγούμενες ημέρες, αλλά και τα προβλεπόμενα φορτία των πιο κοντινών προηγούμενων χρονικών βημάτων για την ίδια μέρα. 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.
Τα αποτελέσματα αποδεικνύουν ότι τα zero-order regularization RBF δίκτυα (ZORRBF), τα ARNN δίνουν τα ακριβέστερα αποτελέσματα στην πρόβλεψη του φορτίου κοντινού χρονικού ορίζοντα (STLF) όπως φαίνεται και στον πίνακα 4

Table 1
Statistical indices for the 2h00 hour load forecasting (min load case)

Models	AR	BP	ALRBP
Relative error (%)	3.86297	4.22628	2.73460
Standard deviation error	5.90894	5.16192	3.91741
RMSE (%)	5.80199	5.18811	3.80634
Models	**GE**	**RAWN**	**MWRAWN**
Relative error (%)	1.51737	6.82535	2.75417
Standard deviation error	2.13450	9.24337	4.19438
RMSE (%)	2.03138	8.98179	4.06330
Models	**ZORRBF**	**RTRL**	**ARNN**
Relative error (%)	1.35119	2.98850	1.22330
Standard deviation error	1.96127	4.35816	1.83768
RMSE (%)	1.86722	4.08409	1.71771

Πίνακας 4 :Σφάλματα πρόβλεψης μεθόδων

12.Πρόβλεψη κατανάλωσης ηλεκτρικού φορτίου ένα πρόβλημα πρωτίστως γραμμικό

Σε αυτό το άρθρο εξετάζουμε το πρόβλημα που αντιμετωπίζει κάθε προμηθευτής ηλεκτρικής ενέργειας, δηλ. την πρόβλεψη της βραχυπρόθεσμης (κοντινού χρονικού ορίζοντα) κατανάλωσης της ηλεκτρικής ενέργειας. Η εστίασή μας κατευθύνεται για το εάν το πρόβλημα είναι γραμμικό ή μη γραμμικό. Κατ' αρχάς, θεσπίζουμε ένα μη γραμμικό μέτρο της στατιστικής εξάρτησης. Δεύτερον, αναλύουμε τις γραμμικές και μη γραμμικές συναρτήσεις αυτοσυσχέτισης της ηλεκτρικής κατανάλωσης. Τρίτον, συγκρίνουμε τις προβλέψεις των μη γραμμικών προτύπων (τεχνητά νευρικά δίκτυα) με τα γραμμικά πρότυπα (ARMA).
Η ανάλυση της συσχέτισης προτείνει ότι η πρόβλεψη της βραχυπρόθεσμης εξέλιξης του τσεχικού ηλεκτρικού φορτίου είναι πρώτιστα ένα γραμμικό πρόβλημα. Αυτό επιβεβαιώνεται από τη σύγκριση των προβλέψεων των μοντέλων όπως φαίνεται παρακάτω.
Στην περίπτωση μοντέλου μιας μεταβλητής τα αποτελέσματα των δυο μοντέλων, γραμμικού και μη γραμμικού ,είναι λίγο διαφορετικά με καλύτερες προβλέψεις από το γραμμικό μοντέλο ARMA και είναι αναμενόμενο εφόσον το πρόβλημα είναι πρωτίστως γραμμικό. Ομοίως στην περίπτωση που έχουμε μοντέλο πολλών μεταβλητών εισάγοντας και την θερμοκρασία. Σε περιπτώσεις όμως όπου τα δεδομένα δεν μπορούν να προεπεξεργαστούν σε μια κατάλληλη μορφή γραμμικού μοντέλου εξαιτίας της διαφοροποίησης τους για παράδειγμα ωριαία φορτία με ημερήσιες θερμοκρασίες αντί ωριαίες τότε τα νευρωνικά δίκτυα μπορούν να ενσωματώσουν περισσότερη πληροφορία και να παράγουν καλύτερα αποτελέσματα όπως φαίνεται και στον πίνακα 5.

Πίνακας 5: Σφάλμα πρόβλεψης νευρωνικού μοντέλου

Table 4
Daily forecasting results for the ARMAX model and the feedforward neural network model, both with temperature input, working days only

Horizon (days)	NMSE (1994)	NMSE (1995)	MAPE (1994)	MAPE (1995)	maxAPE (1994)	maxAPE (1995)
$k = 1$ (ARMAX)	1.1%	1.2%	1.2%	1.4%	4%	7%
$k = 2$ (ARMAX)	2.0%	1.8%	1.7%	1.9%	6%	6%
$k = 1$ (neur. net.)	1.4%	1.6%	1.3%	1.5%	4%	6%
$k = 2$ (neur. net.)	2.4%	2.5%	1.9%	1.9%	5%	8%

13.Τεχνητά νευρωνικά δίκτυα μέσου χρονικού ορίζοντα

Σε αυτό το αρθρο, αναφέρονται τα νέα τεχνητά νευρωνικά δίκτυα (ANN) τα οποία βασιζόμενα στην σχέση καιρού-φορτίου και καιρού-ενέργειας επιλύουν το πρόβλημα της πρόβλεψης του ηλεκτρικού φορτίου και της ενέργειας 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 σφαιρική 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 πίνακα 6.

Πίνακας 6: **Σφάλματα πρόβλεψης μεθόδων**

Table 4
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.

14.Συγκριση MLP και επαναληπτικών (recurrent) νευρωνικών δικτύων για μακροπρόθεσμη πρόβλεψη

Το συγκεκριμένο άρθρο ασχολείται με την επίλυση του προβλήματος της πρόβλεψης του μέγιστου ηλεκτρικού φορτίου μακροπρόθεσμα (σε μακρύ χρονικό ορίζοντα ,LTLF) .
Δυο 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 22 παρά από τις καιρικές συνθήκες (παράγοντες). Τα οικονομικά αυτά δεδομένα είναι : το ακαθάριστο εθνικό προϊόν , το ακαθάριστο εγχώριο προϊόν, ο πλυθησμός, ο αριθμός των οικογενειών, ο αριθμός των κλιματιστικών, το μέγεθος της ρύπανσης από το CO_2, 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.

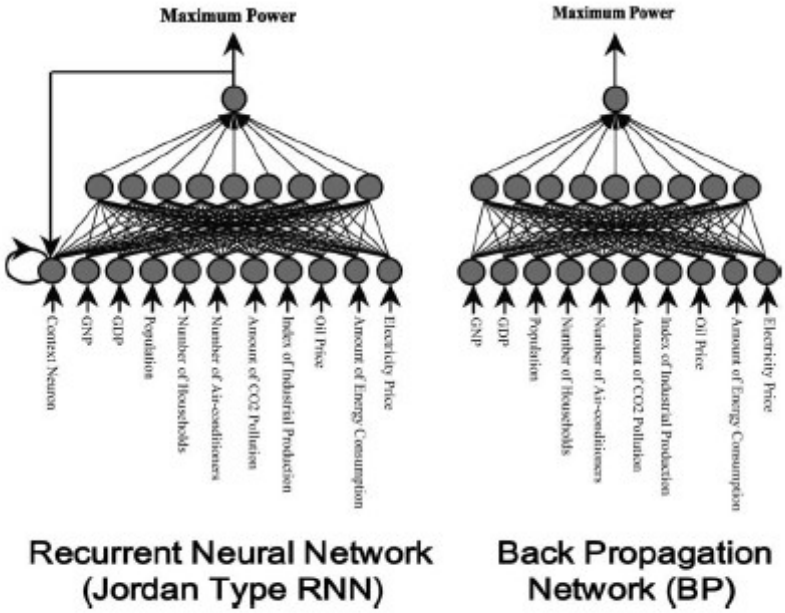


Fig. 2. ANNs used for long-term load forecasting.

Σχήμα 22: Παράγοντες που επηρεάζουν την βραχυπρόθεσμη κατανάλωση φορτίου

15.Βελτιωμένα νευρωνικά δίκτυα βασισμένα σε μοντέλα πρόβλεψης

Αυτό το άρθρο παρουσιάζει την ανάπτυξη του βελτιωμένου νευρωνικού δικτυού το οποίο βασίζεται σε μοντέλα πρόβλεψης του ηλεκτρικού φορτίου βραχυπρόθεσμου χρονικού ορίζοντα (STLF) 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 σε αυτό το έγγραφο. Τα αποτελέσματα δείχνουν ότι τα μοντέλα πρόβλεψης φορτίου που αναπτύσσονται κατ' αυτό τον τρόπο παρέχουν ακριβέστερες προβλέψεις από τα συμβατικά μοντέλα αναστροφής διάδοσης (back propagation). Αυτές οι μέθοδοι αναπτύχθηκαν για πρόβλεψη του ωριαίου ηλεκτρικού φορτίου μια ημέρα μπροστά χρησιμοποιώντας μια μορφοποιημένη αρχιτεκτονική όπως φαίνεται στο σχήμα 23.

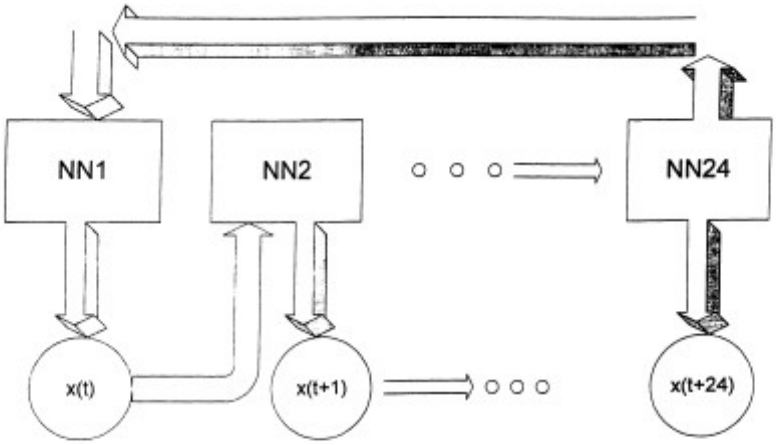


Fig. 1. Proposed modular architecture for the STLF peoblem.

Σχήμα 23: **Βελτιωμένα νευρωνικά δίκτυα βασισμένα σε μοντέλα πρόβλεψης**
Στον παρακάτω πίνακα 7φαίνεται το σφάλμα πρόβλεψης της κάθε μεθόδου.

Table 1
Results for the 2:00 h load

	Models						
Statistical properties	AR	WRAWN	ABP	SE	RBF	N-fuzzy	M. Elman
Relative error (%)	3.86297	2.75417	2.7346	1.5174	1.35119	1.202	1.21426
Standard deviation error	5.90894	4.19438	3.91741	2.1345	1.96127	1.962	1.95812
RMSE (%)	5.80199	4.0633	3.80634	2.0314	1.86722	1.815	1.80738

Πίνακας 7: Σφάλμα πρόβλεψης

16.Συνδυασμός ελεγχόμενων και μη τεχνητών νευρωνικών δικτύων

Το άρθρο αυτό περιγράφει ένα συνδυασμό ελεγχόμενων και μη τεχνητών νευρωνικών δικτύων για τη βραχυπρόθεσμη-ημερήσια πρόβλεψη του ηλεκτρικού φορτίου ενός προαστίου. Πρόκειται για ένα πρόσθιας τροφοδότησης (Feed forward) νευρωνικό μοντέλο το οποίο αποτελείται από δυο επίπεδα (το ένα εξ αυτών κρυμμένο) το οποίο εκπαιδεύεται με τον αλγόριθμο Back propagation (πίσω διάδοσης) .Το μοντέλο εκπαιδεύεται με στοιχεία καιρού, όπως η θερμοκρασία , η σχετική υγρασία και η παγκόσμια ηλιακή ακτινοβολία ,ούτως ώστε να ανακαλύψουμε την επιρροή των κλιματολογικών μεταβλητών στην κατανάλωση ηλεκτρικού φορτίου, παράλληλα με στοιχεία παλαιοτέρων φορτίων. Τα στοιχεία αυτά αποτελούν τις εισόδους .
Στο σχήμα 24 που ακολουθεί φαίνεται η αρχιτεκτονική του .

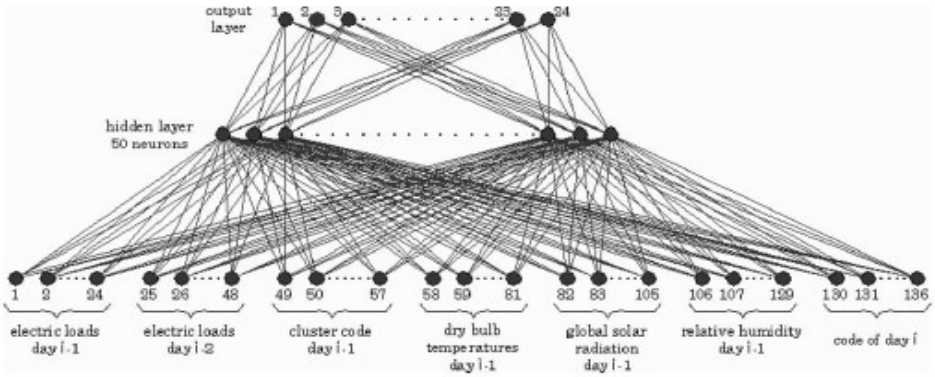
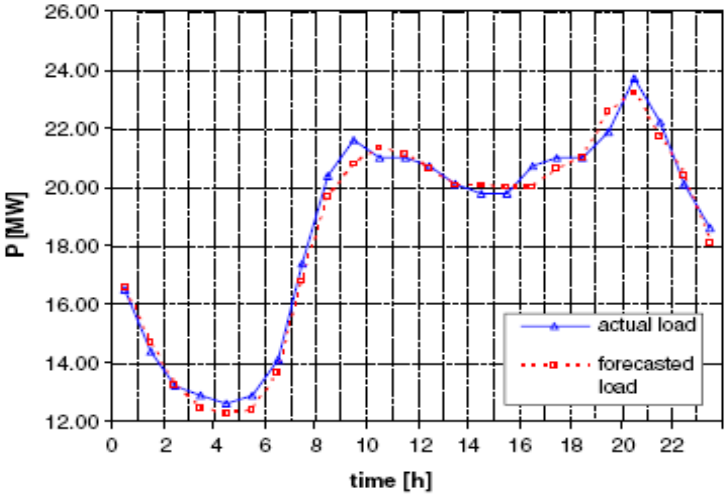


Fig. 14. The MLP architecture.

Σχήμα 24: Αρχιτεκτονική του MLP

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. Στο σχήμα 25 φαίνεται το σφάλμα πρόβλεψης.



Σχήμα 25: Σφάλμα πρόβλεψης μοντέλου

17.Σύγκριση Multilayer Perceptron και Radial Basis Function νευρωνικών δικτύων

Στο συγκεκριμένο άρθρο παρουσιάζεται η σύγκριση των αποτελεσμάτων δυο πολύ συνηθισμένων νευρωνικών δικτύων που χρησιμοποιούνται για την πρόβλεψη κατανάλωσης ηλεκτρικού φορτίου σε λιανικής πώλησης καταστήματα 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 είναι :
α) ημέρα της εβδομάδας ,β) ώρα , γ) εξωτερική και εσωτερική θερμοκρασία και δ) εξωτερική και εσωτερική υγρασία . Τα δίκτυα εκπαιδεύονται με στοιχεία δυο μηνών και ακολούθως κάνουν πρόβλεψη ακριβείας για τις επόμενες 6 ημέρες. Τα δυο δίκτυα είναι τριών επιπέδων με συνάρτηση μεταφοράς την υπερβολική εφαπτομένη για τα MLP δίκτυα και την spline για τα RBF δίκτυα.
Από τα αποτελέσματα των δυο μοντέλων παρατηρούμε ότι τα MLP δίκτυα έχουν καλύτερες προβλέψεις από τα RBF καθώς οι τιμές τους είναι πιο κοντά στις πραγματικές και άρα και το σφάλμα τους είναι μικρότερο.

Στον πίνακα 8 παρατίθενται τα σφάλματα για τα δυο μοντέλα

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 8: Σφάλμα πρόβλεψης

18.Τεχνητό νευρωνικό δίκτυο με δείκτες ημέρας και ώρας

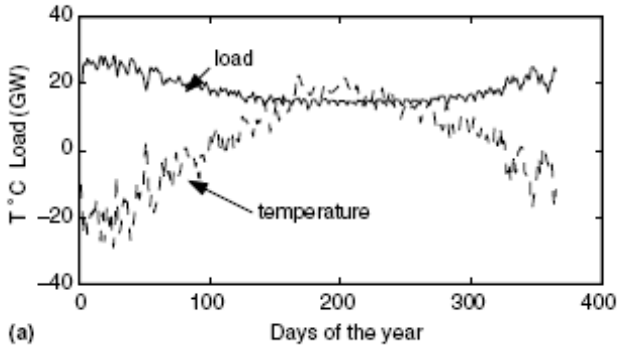
Στο συγκεκριμένο άρθρο παρουσιάζεται ένα τεχνητό νευρωνικό δίκτυο το οποίο ασχολείται με το πρόβλημα της βραχυπρόθεσμης πρόβλεψης ηλεκτρικού φορτίου . πρόκειται για ένα πολύ-επίπεδο πρόσθιας τροφοδότησης (feed-forward) νευρωνικό μοντέλο το οποίο χρησιμοποιεί για εισόδους :

α. ένα δείκτη ημέρας ο οποίος δείχνει το μέγεθος του φορτίου ανάλογα την ημέρα της εβδομάδας που διανύουμε , αποδυκνείεται ότι το αναγκαίο φορτίο είναι αρκετά μικρότερο το σαββατοκύριακο σε σχέση με τις καθημερινές.

β. ένα δείκτη ώρας ο οποίος δείχνει το μέγεθος του φορτίου ανάλογα την ώρα της ημέρας

γ. καιρικές μεταβλητές . Από τις καιρικές μεταβλητές οι οποίους επηρεάζουν την κατανάλωση φορτίου μόνο η θερμοκρασία χρησιμοποιείται ως είσοδος στο μοντέλο αυτό και όχι η ταχύτητα του ανέμου αλλά και η κάλυψη από τα σύννεφα. Αυτό συμβαίνει γιατί αποδυκνείεται ότι μόνο η θερμοκρασία έχει ισχυρή συσχέτιση με το φορτίο όπως φαίνεται και στο παρακάτω σχήμα 26.

δ. σε αυτή την εφαρμογή δεν χρησιμοποιούνται ως είσοδοι ιστορικά φορτία και αυτό γιατί στη βραχυπρόθεσμη πρόβλεψη όπως εδώ για μερικές μέρες μετά τα στοιχειά του φορτίου μερικές ώρες πριν δεν είναι διαθέσιμα με αποτέλεσμα οι εκτιμώμενες τιμές που πρέπει να τις αντικαταστήσουν να περιέχουν έστω και μικρό σφάλμα το οποίο μπορεί να μεγαλώσει δραματικά και να οδηγήσει σε μεγάλο σφάλμα την διαδικασία πρόβλεψης.



Σχήμα 26:Συσχέτιση θερμοκρασίας φορτίου

Το μοντέλο εκπαιδεύεται με τον αλγόριθμο Levenberg-Marquard και τα αποτελέσματα που δίνει είναι αρκετά καλά με χαμηλό σφάλμα περίπου 0,981%

19.RBFNN νευρωνικά δίκτυα που χρησιμοποιούν την συνδυασμένη διόρθωση υπολοίπων λαθών

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. Έχουν αναπτυχθεί πολλές μέθοδοι πρόβλεψης, όπως η πολλαπλάσια γραμμική οπισθοδρόμηση (MLR), ο αυτοανάδρομος κινούμενος μέσος όρος (ARIMA), το γκρίζο πρότυπο (GM) και τεχνητά νευρωνικά δίκτυα (ANN), στον τομέα της πρόβλεψης φορτίων κλιματισμού. Εντούτοις, κανένας από τους δεν έχει αρκετή ακρίβεια για να ικανοποιήσει την πρακτική ζήτηση έτσι βάσει αυτών των προτύπων αναπτύσσεται μια νέα μέθοδος που είναι το μοντέλο RBFNN (radial basis function neural network) με την συνδυασμένη διόρθωση υπολοίπων λαθών. Το νέο πρότυπο υιοθετεί τον προηγμένο αλγόριθμο των νευρωνικών δικτύων βασισμένο στις συναρτήσεις radial basis για την πρόβλεψη του φορτίου και χρησιμοποιεί το συνδυασμένο πρότυπο πρόβλεψης, που είναι ο συνδυασμός των MLR, του ARIMA και του GM, για να υπολογίσει τα υπόλοιπα λάθη και να διορθώσει τα τελευταία αποτελέσματα πρόβλεψης .Το συγκεκριμένο μοντέλο χρησιμοποιεί ως εισόδους την χρονική στιγμή(ώρα) την εξωτερική θερμοκρασία ,την εξωτερική υγρασία και την ηλιακή ακτινοβολία. Το μοντέλο πρόβλεψης RBFNN με την συνδυασμένη διόρθωση υπολοίπων λαθών αποδυκνείεται ότι δίνει ακριβέστερα αποτελέσματα (μικρότερο σφάλμα) από ότι το RBFNN μονό του η από το RBFNN με την απλή διόρθωση .Η αρχιτεκτονική του μοντέλου ακολουθεί στο σχήμα 27.

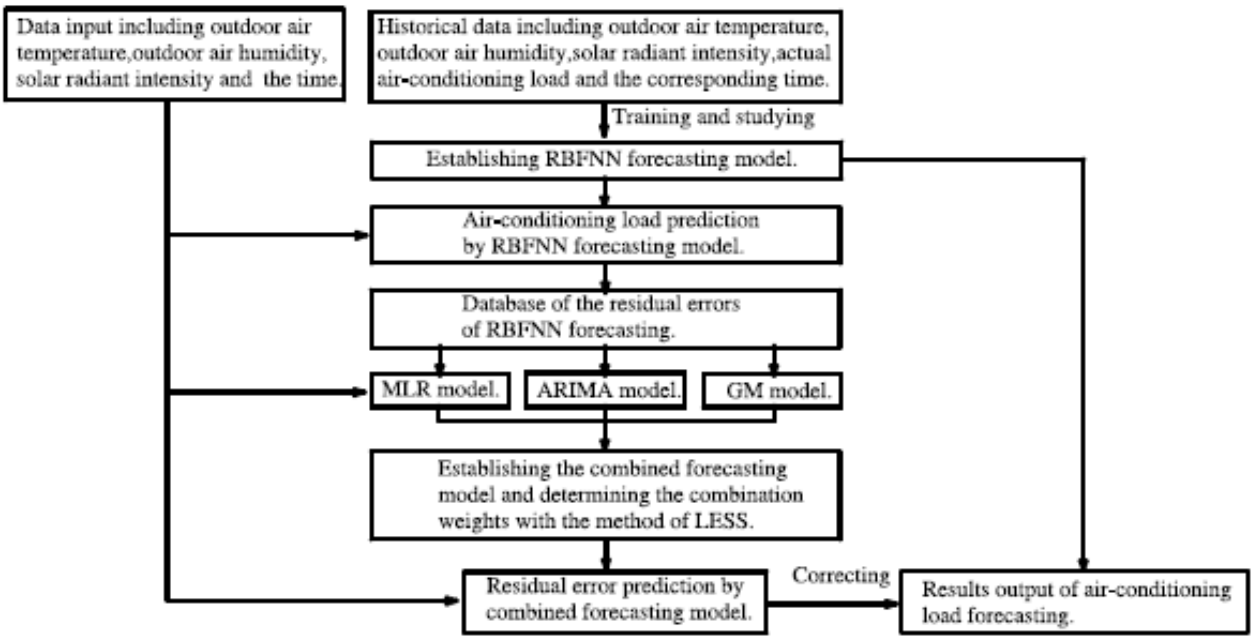


Fig. 3. Forecasting flowchart of RBFNN model with combined residual error.

Σχήμα 27: Αρχιτεκτονική του RBFNN μοντέλου

20.Κυματοειδή νευρωνικά δίκτυα με την τεχνική bacterial foraging

Σε αυτό το άρθρο[20] παρουσιάζεται μια νέα προσέγγιση βραχυπρόθεσμης πρόβλεψης του ηλεκτρικού φορτίου η οποία χρησιμοποιεί την τεχνική bacterial foraging (BFT) για να εκπαιδεύσει τα κυματοειδή νευρωνικά δίκτυα (Wavelet Neural Network) δηλαδή τα τεχνητά νευρωνικά δίκτυα (ANN) τα οποία έχουν συνάρτηση μεταφοράς την παραπάνω.
Το μοντέλο χρησιμοποιεί ως εισόδους φορτία του παρελθόντος συμφωνά με ημερολογιακές μεταβλητές. Για την πρόβλεψη του φορτιού μιας συγκεκριμένης ημέρας χρησιμοποιεί τα φορτία της ιδίας ημέρας της προηγουμένης εβδομάδας και της ιδίας ημέρας του προηγουμένου έτους τον ιδίου μηνά και της ιδίας εβδομάδας ,όπως φαίνεται και στο σχήμα 28.

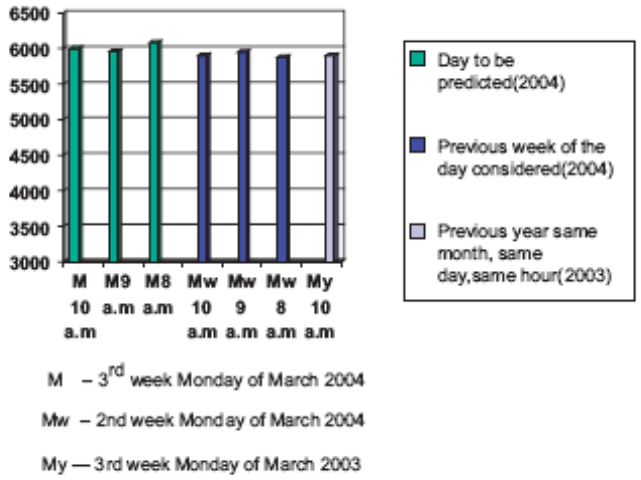


Fig. 1. Input variable selection.

Σχήμα 28: Είσοδοι μοντέλου

Από τα αποτελέσματα βλέπουμε ότι τα WNN και τα ANN που χρησιμοποιούν τη bacterial foraging τεχνική έχουν μεγαλύτερη ακρίβεια από αυτά που δεν την χρησιμοποιούν . Επίσης τα WNN δίκτυα σε σχέση με τα ANN δίκτυα έχουν μεγαλύτερη ακρίβεια δηλαδή μικρότερο σφάλμα . Τα αποτελέσματα των προβλέψεων παρατίθενται στο πίνακα 9.

Comparison of MAPE with and without BFT in ANN and WNN for different LF periods—TNEB 69 bus system

LF period	ANN with delta rule	ANN with BFT	WNN with delta rule	WNN with BFT
One day (November 11,004)	1.388625	0.778292	0.845929	0.528421
One week (November 7–14,004)	1.87	1.2	1.46	0.914
One month (November 1–30,004)	4.45	3.58	3.81	3.2

Πίνακας 9: Σφάλματα μεθόδων

21.RSAN νευρωνικά δίκτυα

Στο άρθρο [21] περιγράφεται το RSAN νευρωνικό δίκτυο δηλαδή ένας συνδυασμός ενός τεχνητού νευρωνικού δικτυού (ANN) με την θεωρία των σκληρών σέτ δεδομένων (Rough set theory RS) για τη λύση του προβλήματος της πρόβλεψης κατανάλωσης του φορτίου από τα κλιματιστικά. Η θεωρία RS χρησιμοποιείται για την επιλογή των καταλλήλων παραγόντων οι οποίοι επηρεάζουν την κατανάλωση του φορτίου ούτως ώστε να χρησιμοποιηθούν ως είσοδοι στο τεχνητά νευρωνικά δίκτυα (ANN). Κατόπιν τα νευρωνικά δίκτυα κάνουν πρόβλεψη του φορτιού. Για την αύξηση της ακριβείας των αποτελεσμάτων χρησιμοποιείται μια σύνθεση από RSAN δίκτυα έτσι το αποτέλεσμα τους είναι να δημιουργηθεί ένα πολλαπλό RSAN δίκτυο multiple-RSAN (MRAN) .
Η μορφή (αρχιτεκτονική) του μοντέλου φαίνεται στο σχήμα 29.

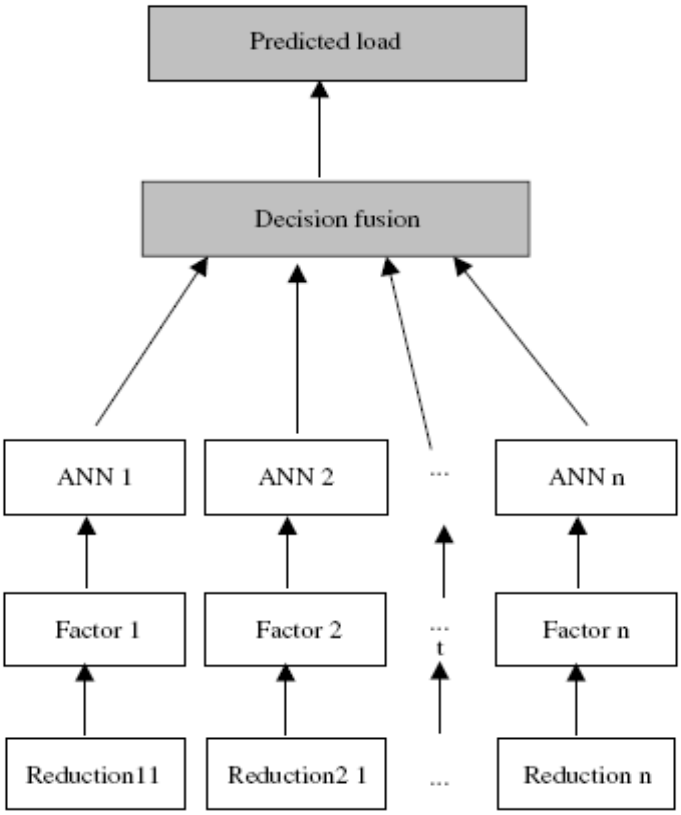


Fig. 2. Load prediction based on information data-fusion technique.

Σχήμα 29: Μορφή (αρχιτεκτονική) του RSAN μοντέλου

Το μοντέλο χρησιμοποιεί ως εισόδους την θερμοκρασία, την σχετική υγρασία, την πίεση και το ποσοστό ροής και προβλέπει το φορτίο. Το MRAN μοντέλο δίνει καλυτέρα αποτελέσματα δηλαδή μικρότερο σχετικό σφάλμα από ότι τα RSAN μοντέλα μονά τους αλλά και από ότι το μοντέλο ARIMA.

22.Τεχνητά νευρωνικά δίκτυα με ανατροφοδότηση

Στο συγκεκριμένο άρθρο παρουσιάζεται μια νέα προσέγγιση της βραχυπρόθεσμης πρόβλεψης ηλεκτρικού φορτίου στα κτίρια. Η μέθοδος βασίζεται σε ένα είδος τεχνητών νευρωνικών δικτύων τα οποία ανατροφοδοτούν ένα μέρος των εξόδων (προβλέψεων) τους. Το δίκτυο εκπαιδεύεται από ένα υβριδικό αλγόριθμο. Το μοντέλο χρησιμοποιεί τις τρέχων αλλά και προβλέψιμες τιμές της θερμοκρασίας ,το τρέχων φορτίο ,την ημέρα και την ώρα ως εισόδους .Τα αποτελέσματα που παίρνουμε από το μοντέλο αυτό είναι πολύ ακριβή με χαμηλό σφάλμα. Η μορφή του δικτυού ακολουθεί στα σχήματα 30 και 31.

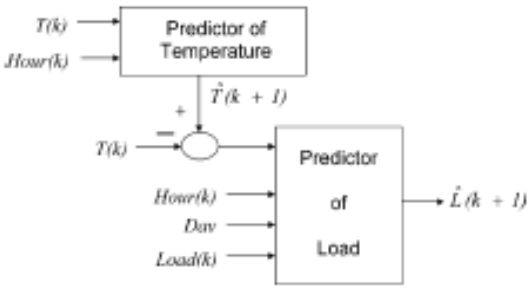


Fig. 2. General Structure of the load predictor.

Σχήμα 30: Είσοδοι και έξοδος (μέλλον φορτίο) του δικτυού

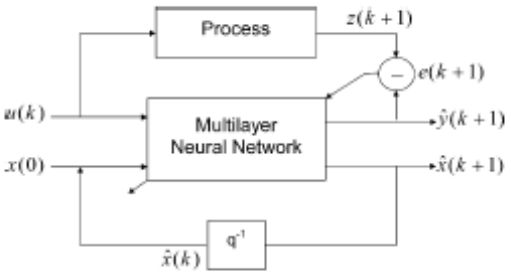


Fig. 1. General schema of the feedback neural network designed in [6].

Σχήμα 31: Το σύστημα ανατροφοδότησης του νευρωνικού δικτυού.

23.Τεχνητά νευρωνικά δίκτυα και ανάλυση ευαισθησίας των παραγόντων

Σε αυτό το άρθρο, παρουσιάζεται η εφαρμογή ενός τεχνητού νευρωνικού δικτύου (ANN) για την λύση της πρόβλεψης του φορτίου των υποσταθμών. Οι πελάτες σε ένα σύστημα διανομής είναι τριών ειδών:
α) κατοικίες, β) εμπορικά καταστήματα και γ) βιομηχανίες.
Κάθε υποσταθμός έχει ένα διαφορετικό σχέδιο κατανάλωσης ισχύος ανάλογα με τον συνδυασμό των παραπάνω πελατών που εξυπηρετείτο 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 τρία είδη υποσταθμών κατοικιών ,εμπορικών και βιομηχανικών 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 σχήμα 32 φαίνεται η επίδραση της θερμοκρασίας στην πρόβλεψη για τα σαββατοκύριακα.

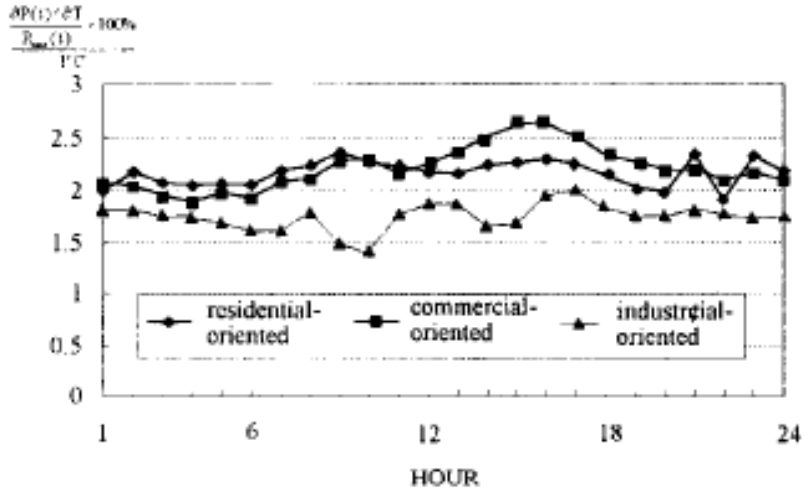


Fig. 12. Percentage temperature sensitivity of the three different types of substation on a weekend in summer.

Σχήμα 32: Επίδραση θερμοκρασίας στην πρόβλεψη για τα σαββατοκύριακα

24.Σύνθετα νευρωνικά δίκτυα : Adaptive Combiner ANN

Μια ολοκληρωμένη προσέγγιση των νευρωνικών δικτύων (ANN) για τη λύση του προβλήματος της πρόβλεψης του βραχυπρόθεσμου φορτίου (STLF) προτείνεται σε αυτό το άρθρο. Η προσέγγιση 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 επόμενο φορτίο 24 ώρες μπροστά, 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 χρησιμοποιεί μέσες τιμές του φορτίου από προηγούμενη συμπεριφορά φορτίων, ενώ το Adaptive Combiner ANN χρησιμοποιεί το σταθμισμένο συνδυασμό των αποτελεσμάτων (προβλέψεων) από το βασικό ANN και τον Averager and Forecaster ANN για να προβλέψει το τελικό φορτίο.
Στο σχήμα 33 παρουσιάζεται η αρχιτεκτονική του σύνθετου αυτού νευρωνικού δικτυού.

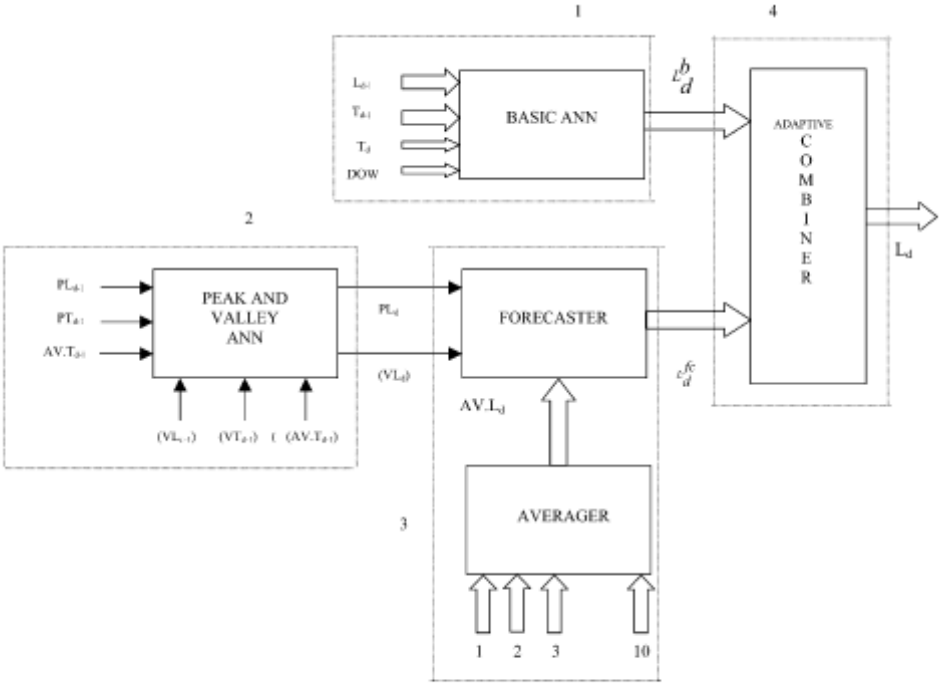


Fig. 1. Block diagram of proposed architecture.

Σχήμα 33: Αρχιτεκτονική του Adaptive Combiner τεχνητού νευρωνικού δικτυού

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 έχουν χαμηλότερο σφάλμα από αυτά που παίρνουμε από το κάθε ένα ξεχωριστά. Ο αριθμός των σχεδίων εκπαίδευσης αλλά και η θερμοκρασία παίζουν σπουδαίο ρολό στην ακρίβεια των αποτελεσμάτων.

25.Δυναμικά τεχνητά νευρωνικά δίκτυα : το πρότυπο (DAN2)

Αυτό το άρθρο παρουσιάζει την ανάπτυξη ενός δυναμικού τεχνητού νευρωνικού προτύπου (DAN2) για τη μεσοπρόθεσμη πρόβλεψη των ηλεκτρικών φορτίων (MTLF).Η ακριβής μεσοπρόθεσμη πρόβλεψη των ηλεκτρικών φορτίων MTLF παρέχει χρήσιμες πληροφορίες για το καλύτερο σχέδιο επέκτασης της ηλεκτρικής παραγωγής, για δραστηριότητες συντήρησης προγράμματος, για βελτιώσεις των συστημάτων, για την καλύτερη διαπραγμάτευση μελλοντικών συμβάσεων και για την ανάπτυξη αποδοτικότερων στρατηγικών στην αγορά καυσίμων . Παρουσιάζουμε ένα ετήσιο πρότυπο που χρησιμοποιεί μηνιαία φορτία τους παρελθόντος για να προβλέψει τις μελλοντικές ηλεκτρικές απαιτήσεις(φορτία). Επίσης δείχνουμε ότι ο συνυπολογισμός των καιρικών πληροφοριών (παράγοντες) βελτιώνει την ακρίβεια πρόβλεψης φορτίου. Τέτοια πρότυπα, εντούτοις, απαιτούν τις ακριβείς καιρικές προβλέψεις, οι οποίες είναι συχνά δύσκολο να ληφθούν. Επομένως, έχουμε αναπτύξει μια εναλλακτική λύση, δηλαδή εποχιακά πρότυπα που παρέχουν άριστες προβλέψεις χωρίς να απαιτείται μεγάλη εμπιστοσύνη επάνω στις καιρικές μεταβλητές. Και τα ετήσια όπως και τα εποχιακά πρότυπα παράγουν σφάλμα (MAPE) μικρότερο του 1%, καταδεικνύοντας την αποτελεσματικότητα του προτύπου DAN2 στην πρόβλεψη των μεσοπρόθεσμων φορτίων. Τέλος, συγκρίνουμε τα αποτελέσματά του προτύπου DAN2 με το πρότυπο MLR , το πρότυπο ARIMA και ένα παραδοσιακό νευρωνικό δίκτυο. Τα αποτελέσματα δείχνουν ότι το πρότυπο DAN2 δίνει χαμηλότερο σφάλμα από τα MLR, ARIMA και το παραδοσιακό νευρωνικό δίκτυο. Και τα ετήσια και τα εποχιακά πρότυπα DAN2 παρείχαν πρόβλεψη ακρίδας πάνω από 99% (MAPE τιμές κατω1%). Τέλος , τα εποχιακά πρότυπα ξεπέρασαν όλα τα ετήσια πρότυπα.

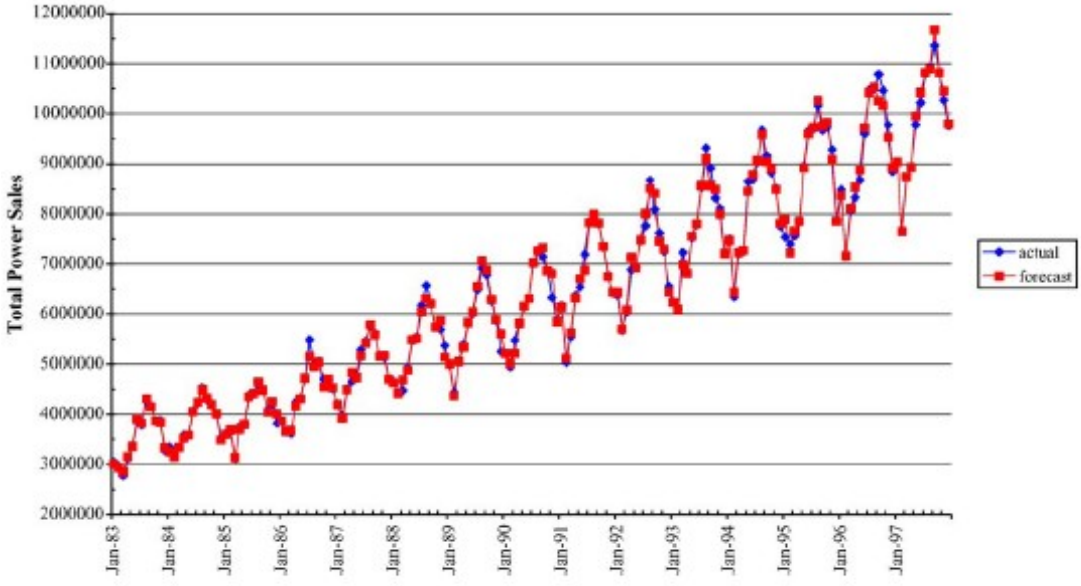


Fig. 3. Actual vs. forecasted loads for the yearly model.

Σχήμα 34: Σφάλμα πρόβλεψης ετησίου μοντέλου

26.Συνδυασμός τεχνητών νευρωνικών δικτύων και εμπείρων εξειδικευμένων συστημάτων

Το άρθρο [26] παρουσιάζει μια τεχνική βραχυπρόθεσμης πρόβλεψης ηλεκτρικού φορτίου που χρησιμοποιείται στην Αίγυπτο. Η τεχνική είναι βασισμένη σε ένα γενικευμένο πρότυπο που συνδυάζει τα χαρακτηριστικά γνωρίσματα ενός τεχνητού νευρωνικού δικτυού ANN και ενός έμπειρου εξειδικευμένου συστήματος(ES). Η παραπάνω μεθοδολογία καθιστά την τεχνική ικανή, ενημερωμένη ενώ επιτρέπει την επέμβαση χειριστών όταν και οπού χρειάζεται πράγμα που την καθιστά ιδιαίτερα κατάλληλη για την πρόβλεψη των φορτίων από την EUG στην περιοχή καθώς η κατανάλωση φορτίου επηρεάζεται περισσότερο λόγω κοινωνικών δραστηριοτήτων, ενώ ο καιρός συμβάλλει πολύ ελάχιστα στην κατανάλωση και κατά επέκταση στην πρόβλεψη των φορτίων. Για παράδειγμα, πολλές κοινωνικές συναθροίσεις εξαρτώνται από τις θρησκευτικές προτιμήσεις οι οποίους δεν μπορούν να αποφασιστούν καλά εκ των προτέρων. Το πρότυπο χρησιμοποιεί ως εισόδους στοιχειά όπως η ημέρα ,η ώρα, το προβλεπόμενο φορτίο μια ώρα πριν , το φορτίο της ιδίας ώρας της προηγουμένης ημέρας , το μέσο ,το μέγιστο και το ελάχιστο φορτίο της προηγουμένης ημέρας αλλά και το φορτίο της ιδίας ώρας μια εβδομάδα πριν όπως και το μέσο, το μέγιστο και το ελάχιστο φορτίο της προηγουμένης εβδομάδας. Στο σχήμα 35 φαίνεται η αρχιτεκτονική του δικτυού.

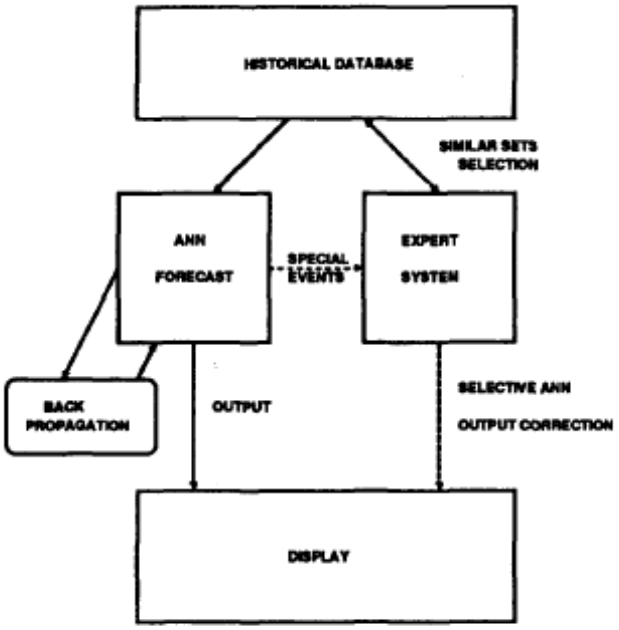


Fig. 1. The coupled ANN–ES scheme.

Σχήμα 35: Αρχιτεκτονική του δικτυού

Τα μέσα απόλυτα σφάλματα πρόβλεψης για την προτεινόμενη μεθοδολογία είναι 2,63% με μια σταθερή απόκλιση 2,62% καταδεικνύοντας σαφώς το πλεονέκτημα της μεθοδολογίας σε σχέση με το πρότυπο MLR αλλά και το ANN μονό του.

27.Ασταθή νευρωνικά δίκτυα

Σε αυτό το άρθρο αναπτύσσεται ένα νέο ευφυές σύστημα βραχυπρόθεσμης πρόβλεψης της μέγιστης τιμής του φορτίου (STLF) το οποίο κατασκευάζεται από ένα ασταθές νευρωνικό δίκτυο. Το σύστημα χρησιμοποιεί τον αλγόριθμο Back Propagation (BP) για την εκπαίδευση του . Οι είσοδοι του συστήματος είναι δεδομένα από παλαιοτέρα φορτία και καιρικούς παράγοντες όπως η θερμοκρασία. Η επιλογή των καταλλήλων παραγόντων οι οποίοι επηρεάζουν και έχουν μεγαλύτερη σχετικότητα με το φορτίο καθορίζονται από την Gray analysis (γκρίζα ανάλυση) ανάλυση της αμοιβαίας σχετικότητας μεταξύ φορτίου και θερμοκρασίας. Με αλλά λόγια οι ακατάλληλοι παράγοντες αποφεύγονται να χρησιμοποιηθούν στην εκπαίδευση και την εφαρμογή του νευρωνικού δικτυού γιατί μπορεί να μην εκπαιδεύσουν το πρότυπο ολοκληρωμένα με αποτέλεσμα να μην απόδοση σωστά το δίκτυο και να μεγαλώσει το σφάλμα. Στη συγκεκριμένη μελέτη το σύστημα μας αποτελείται από πέντε νευρωνικά δίκτυα όπως φαίνεται στο σχήμα 36 ανάλογα με τις ημέρες πρόβλεψης (1εως 5) που θέλουμε να κάνουμε.

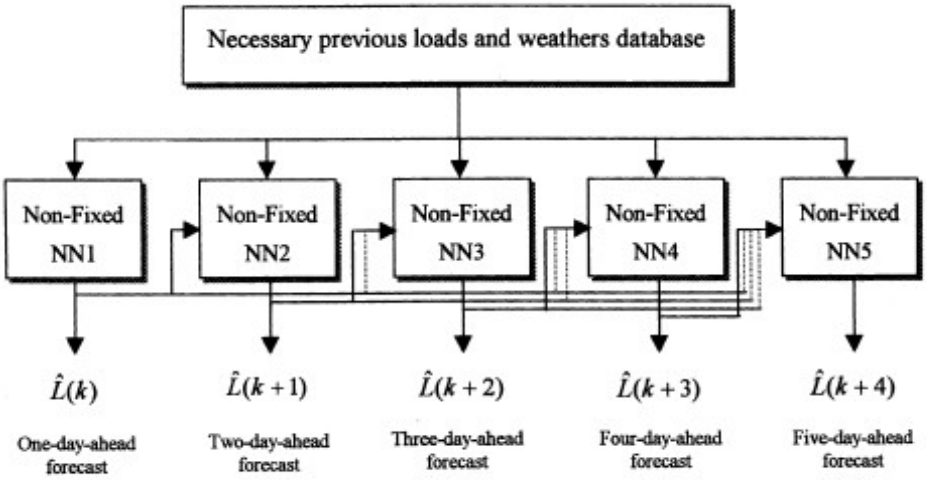


Fig. 1. The architecture of one-to-five-day-ahead peak load forecastor.

Σχήμα 36: Αρχιτεκτονική του ασταθούς δικτυού

Από τα αποτελέσματα του πίνακα 10 φαίνεται ότι το σύστημα έχει καλύτερα αποτελέσματα από τις άλλες τεχνικές πρόβλεψης.

Πίνακας 10: Σφάλμα πρόβλεψης

Table 4
The MAPEs and the days number of forecasted error > 600 MW of one-to-five-day-ahead peak load forecasts for 1998 data

	Time series model		Fixed NN module		ANNSTLF module		Non-fixed NN module	
	MAPE (%)	No. (>600 MW)	MAPE (%)	No. (>600 MW)	MAPE (%)	No. (>600MW)	MAPE (%)	No. (>600 MW)
One-day-ahead	3.35	65	2.22	11	1.81	5	1.75	4
Two-day-ahead	3.97	141	2.74	27	2.35	15	2.24	12
Three-day-ahead	4.36	170	3.01	48	2.63	20	2.56	19
Four-day-ahead	4.69	181	3.37	75	2.71	25	2.67	22
Five-day-ahead	5.01	188	3.58	113	2.80	33	2.74	28

ΤΕΧΝΙΚΕΣ ΠΡΟΒΛΕΨΕΙΣ ΗΛΕΚΤΡΙΚΟΥ ΦΟΡΤΙΟΥ

Στις ενότητες που ακολουθούν παρουσιάζονται τα διάφορα μοντέλα πρόβλεψης ηλεκτρικού φορτίου.

Η μέθοδος αναδρομής (autoregressive method)

Η συγκεκριμένη στατιστική μέθοδος αν και δεν χρησιμοποιείται πλέον ευρέως, ωστόσο παρουσιάζεται για λόγους σύγκρισης με τις υπόλοιπες μεθόδους. Έτσι, το μοντέλο του οποίου τις μελλοντικές τιμές θέλουμε να προβλέψουμε παριστάνεται από την εξίσωση:

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_p y_{t-p} + b_t \qquad (1)$$

όπου *Yt,Yt-1, ... ,Yt-p* είναι η παρούσα και οι παρελθόντες τιμές του φορτίου αντίστοιχα, α1, …αp οι άγνωστοι συντελεστές βάρους αυτών των τιμών, a_0 ένας σταθερός όρος και b_t τυχαίος θόρυβος. Ο υπολογισμός των συντελεστών αυτών επιτυγχάνεται με τη χρήση της αναδρομικής μεθόδου ελαχίστων τετραγώνων, σε συνδυασμό με τον αλγόριθμο αποσύνθεσης U-D για αριθμητικούς υπολογισμούς και ακριβής βελτιστοποίηση. Η τάξη p του μοντέλου υπολογίζεται χρησιμοποιώντας το κριτήριο Akaike .

Προχωρημένοι αλγόριθμοι ανάστροφης διάδοσης (backpropagation) για MLP δίκτυα.

Τα MLP νευρωνικά δίκτυα που εκπαιδεύονται με βάση τον κλασικό αλγόριθμο ανάστροφης διάδοσης (BP), θεωρούνται ως η πιο συνηθισμένη προσέγγιση για περίπλοκες αντιστοιχήσεις εισόδου εξόδου ενός δικτύου, αποτελώντας ουσιαστικά μια μη γραμμική συνάρτηση αντιστοίχησης. Ο γενικευμένος κανόνας δέλτα χρησιμοποιείται για τη ρύθμιση των βαρών των τροφοδοτούμενων προς τα εμπρός δικτύων (feed-forward networks) ελαχιστοποιώντας έτσι μια προκαθορισμένη συνάρτηση κόστους σφάλματος. Οι τιμές των βαρών αυτών ρυθμίζονται με βάση τον κανόνα:

$$w_{ik}^p(t+1) = w_{ik}^p(t) + m\delta_k^p y_i^p + \alpha \Delta w_{ik}^p(t) \qquad (2)$$

Οπού m η παράμετρος του ρυθμού εκμάθησης, a ο όρος momentnm (κυμαινόμενος μεταξύ 0 και 1) και d η αρνητική κλίση του συνολικού τετραγωνικού σφάλματος σε σχέση με την έξοδο του νευρώνα.
Ωστόσο, οι κλασικοί αλγόριθμοι BP αντιμετωπίζουν το σημαντικό πρόβλημα της αργής σύγκλισης. Για το σκοπό αυτό χρησιμοποιείται μια τεχνική αυτοδιδαχής που ονομάζεται νευρωνικό δίκτυο ανάστροφης διάδοσης με προσαρμοζόμενο ρυθμό εκμάθησης (adaptive learning rate backpropagation neural network ή ALRBP). Το ALRBP συσχετίζει το ρυθμό εκμάθησης με το συνολικό σφάλμα E, επιτυγχάνοντας έτσι την επιτάχυνση της διαδικασίας εκμάθησης. Ο κανόνας ανανέωσης των βαρών σε αυτήν την περίπτωση θα δίνεται από τη σχέση :

$$w_{ij}^{p}(t+1) = w_{ij}^{p}(t) - \frac{q(E)\dfrac{\partial E}{\partial w_{ij}^{p}}}{\left\| \dfrac{\partial E}{\partial w_{ij}^{p}} \right\|^{2}}$$

Όπου q(E) είναι μια συνάρτηση του σφάλματος E που δίνεται από την σχέση:

$$q(E) = \begin{cases} m \\ mE \\ m \tan \mathrm{h}\left(\dfrac{E}{E_o}\right) \end{cases}$$

Όπου οι m και E_0 αποτελούν σταθερούς μη αρνητικούς αριθμούς που αντιπροσωπεύουν το ρυθμό εκμάθησης και τον παράγοντα κανονικοποίησης του σφάλματος αντίστοιχα. Κύριο πλεονέκτημα του συγκεκριμένου αλγορίθμου αποτελεί το γεγονός της εξάρτισης του ρυθμού εκμάθησης από την ακαριαία τιμή του συνολικού τετραγωνικού σφάλματος E.
Αρκετά σημαντικό ρόλο στην ακρίβεια πρόβλεψης ενός νευρωνικού μοντέλου διαδραματίζει και η μέθοδος κωδικοποίησης των δεδομένων. Με τη συμβατική μέθοδο τα δεδομένα αντιπροσωπεύονται χρησιμοποιώντας ένα μονάχα κόμβο στα στρώματα εισόδου ή εξόδου. Μια εναλλακτική απεικόνιση των δεδομένων ονομάζεται Gaussian κωδικοποίηση. Στην περίπτωση αυτή κάθε δεδομένο αντιπροσωπεύεται ως η μέση τιμή ενός κυλιόμενου Gaussian προτύπου διέγερσης, σε αρκετούς κόμβους στην είσοδο και έξοδο του δικτύου. Αναλυτικότερα, κάθε μεταβλητή του δικτύου αντιστοιχίζεται σε ένα κυλιόμενο Gaussian πρότυπο N κόμβων, που περιλαμβάνει επιπρόσθετους κόμβους σε κάθε πλευρά του εύρους κάθε μεταβλητής.

Το επίπεδο ενεργοποίησης κάθε κόμβου βρίσκεται στο διάστημα [0.1,0.9]. Σε κάθε κόμβο αντιστοιχίζεται μια τιμή α_i, γραμμικά διευθετημένη κατά απόσταση *δ* έτσι ώστε να περιλάβει ολόκληρο το διάστημα του x, με το κέντρο του Gaussian προτύπου διέγερσης να αντιστοιχεί στην κωδικοποιημένη τιμή. Για μια συγκεκριμένη τιμή του x η διέγερση κάθε κόμβου καθορίζεται από μια συνάρτηση $\psi_i(x)$ που ικανοποιεί την απαίτηση :

$$\sum_{i=1}^{N} \alpha_i \psi_i(x) = \int \alpha \phi(\alpha - x)\, \mathrm{d}\alpha = \bar{\alpha} = x \qquad (5)$$

όπου Φ(α) η κανονικοποιημένη πυκνότητα πιθανότητας μιας Gaussian κατανομής. Η τεχνική αυτή προσομοιάζει με εκείνη της ασαφοποίησης δεδομένων (data-fuzzification). Η τελική (αποκωδικοποιημένη πλέον) έξοδος του δικτύου θα δίνεται από τη σχέση:

$$x = \frac{\sum_{i=1}^{N} \alpha_i \psi_i(x)}{\sum_{i=1}^{N} \psi_i(x)}$$

Μέσα και από τις ανάλογες δοκιμές τελικά, αποδεικνύεται ότι η ακρίβεια στην πρόβλεψη είναι καλύτερη από ότι στην περίπτωση των συμβατικών δικτύων.

Δίκτυα τυχαίας ενεργοποίησης βαρών (Random activation weight networks ή RAWN) και εκπαιδευόμενα RAWN με κινούμενο παράθυρο αναδρομής (moving window regression trained RAWN ή MWRAWN.)

Μια γενική συνάρτηση προσέγγισης μπορεί να επιτευχθεί με τα feedforward νευρωνικά δίκτυα που αποτελούνται από ένα μονάχα κρυφό στρώμα μη γραμμικών νευρώνων. Η πρωταρχική ιδέα πίσω από τα δίκτυα RAWN είναι ότι δεν απαιτείται η εκπαίδευση των βαρών μεταξύ του στρώματος εισόδου και του κρυφού στρώματος. Θεωρώντας ως τυχαίες τις αρχικές τιμές των βαρών ενεργοποίησης, η διαδικασία εκτίμησης των παραμέτρων μπορεί να θεωρηθεί ως γραμμικού τύπου, και έτσι είναι δυνατό να χρησιμοποιηθεί ένας γραμμικός εκτιμητής ελάχιστων τετραγώνων. Οι εξισώσεις του συγκεκριμένου δικτύου μπορούν να γραφτούν με την μορφή πινάκων:

$$
\begin{aligned}
Z &= XW^{h} && [N_e \times N_h] \\
U &= f(Z) = \alpha \tan \mathrm{h}(bZ) && [N_e \times N_h] \\
Y_n &= UW^{o} && [N_e \times N_o]
\end{aligned}
$$

όπου W^h ο πίνακας των βαρών μεταξύ εισόδου και κρυφού στρώματος, U ο πίνακας εξόδου του κρυφού στρώματος, W ο πίνακας βαρών μεταξύ κρυφού στρώματος και στρώματος εξόδου, και α και b το πλάτος και η φάση της συνάρτησης ενεργοποίησης αντίστοιχα. Θεωρώντας επίσης ως Y_d τα δεδομένα εξόδου στο σύνολο εκπαίδευσης, ο πίνακας Y_d θα δίνεται από τη σχέση

$$Y_d = UW^{o} + e$$

όπου e ο παράγοντας του σφάλματος. Αφού προσδιοριστούν αρχικά τα βάρη του πίνακα W^h „ ακολουθεί ο υπολογισμός των βαρών του πίνακα W^0, ελαχιστοποιώντας τη διαφορά ανάμεσα στην έξοδο του νευρωνικού δικτύου Y(=U W^h) και τις επιθυμητές τιμές εξόδου Y_d. Αυτό βέβαια αποτελεί ένα κανονικό πρόβλημα τύπου ελάχιστων τετραγώνων. Δεδομένου ότι ο πίνακας U είναι full rank (δηλαδή έχει ένα βαθμό του N_h), η λύση των ελάχιστων τετραγώνων προκύπτει:

$$\hat{W}^{o} = (U^{T}U)^{-1}U^{T}Y_d$$

Με άλλα λόγια, ο πίνακας U πρέπει να έχει N_h γραμμικές ανεξάρτητες στήλες. Σε αντίθετη περίπτωση δεν θα υπάρχει λύση για την προηγούμενη εξίσωση. Εάν ο πίνακας U μπορούσε να δοθεί από τον Z ως ένας μετασχηματισμός γραμμικής συνάρτησης, τότε ο U θα είχε το βαθμό του Z. Όμως, ο Z (ως το εξωτερικό γινόμενο X W^h) αποτελείται από γραμμικά εξαρτημένες στήλες και άρα ο U δεν θα μπορούσε να είναι full rank. Στην περίπτωση ωστόσο μιας μη γραμμικής εξίσωσης, ο πίνακας U είναι full rank, μιας και οι εξαρτημένες στήλες του Z μετασχηματίζονται με ένα μη γραμμικό τρόπο, εξασφαλίζοντας ότι δεν υπάρχουν δύο στήλες του πίνακα που να είναι ακριβώς οι ίδιες. Το τελευταίο επιτυγχάνεται δίνοντας τυχαίες τιμές στα βάρη ενεργοποίησης W^h . Με τη χρήση επαναληπτικών μεθόδων υπολογίζονται οι τελικές τιμές των βαρών.

Για ακριβείς και γρήγορους υπολογισμούς εφαρμόζεται στα προηγούμενα μια μέθοδος αναδρομής κινούμενου παραθύρου. Έτσι θεωρώντας:

$$U = \begin{bmatrix} u_1(1) & \cdots & u_m(1) \\ \vdots & & \vdots \\ u_1(k) & \cdots & u_m(k) \end{bmatrix} = \begin{bmatrix} u^T(1) \\ \vdots \\ u^T(k) \end{bmatrix} \quad \text{and}$$

$$Y_d = [y_d(1) \ldots y_d(k)]^T$$

στην περίπτωση ενός κινούμενου παραθύρου μήκους χρησιμοποιούνται οι ακόλουθες σχέσεις :

$$U_k = \begin{bmatrix} u^T(k - n_w + 1) \\ \vdots \\ u^T(k) \end{bmatrix} = \begin{bmatrix} u^T(k - n_w + 1) \\ \vdots \\ U(k, k - n_w + 2) \end{bmatrix}$$

$$U_{k+1} = \begin{bmatrix} u^T(k - n_w + 2) \\ \vdots \\ u^T(k + 1) \end{bmatrix} = \begin{bmatrix} U(k, k - n_w + 2) \\ \vdots \\ u^T(k + 1) \end{bmatrix}$$

ένα κινούμενο παράθυρο ελάχιστων τετραγώνων μπορεί να εξαχθεί από τις αναδρομικές εξισώσεις:

$$\bar{W}^{o}(k+1) = \bar{W}^{o}(k) - P(k+1) \times [\Gamma(k+1)\bar{W}^{o}(k) - \delta(k+1)]$$

Όπου

$$P^{-1}(k+1) = (U_{k+1}^{T} U_{k+1})^{-1} = P^{-1}(k) + \Gamma(k+1)$$

Ο πίνακας συμμεταβλητότητας της εκτίμησης W^0 (k+1), ενώ οι ποσότητες δ και Γ καθορίζονται ως :

$$\delta(k+1) = u(k+1)y(k+1) - u(k-n_w+1)y(k-n_w+1)$$
$$\Gamma(k+1) = u(k+1)u^{T}(k+1) - u(k-n_w+1)u^{T}(k-n_w+1)$$

Βασικό πλεονέκτημα της συγκεκριμένης μεθόδου αποτελεί το γεγονός ότι η τιμή *W (t)* εκτιμάται χρησιμοποιώντας την πληροφορία που ανακύπτει από τα τελευταία n_w δείγματα. Κατ' αυτόν τον τρόπο επιτυγχάνεται μια άυξηση στην ταχύτητα εκπαίδευσης του δικτύου, ανάλογη του μήκους του παραθύρου. Τέλος, αρκετά σημαντικό ρόλο παίζει η επιλογή των βαρών ενεργοποίησης W^h . Για το σκοπό αυτό επιλέγεται ένας πίνακας W^h που προκύπτει μέσα από μια τυχαία παραγόμενη κανονική κατανομή με μέση τιμή 0 και τυπική απόκλιση ίση με τη μονάδα. Έτσι, η κανονικοποιημένη τιμή για τον W^h θα δίνεται τελικά από την εξίσωση:

$$W^{h} = \sqrt{\frac{a}{\max_{k} \sum_{\ell=1}^{N_\ell+1} x_\ell^2(k)}} N(0,1)$$

Δίκτυα με συναρτήσεις radial basis (radial basis function networks ή RBFN's)

Ένα δίκτυο RBFN σχεδιάζεται ώστε να πραγματοποιεί μια αντιστοίχηση εισόδου - εξόδου, εκπαιδευόμενο με ένα σύνολο από p δείγματα *(xk,dk), k=1,2,…,ρ.* Το συγκεκριμένο υβριδικό δίκτυο ακολουθεί μια συνδυασμένη τεχνική επιτηρούμενης και μη επιτηρούμενης εκπαίδευσης. Οι κρυφοί κόμβοι του δικτύου ακολουθούν μια κανονικοποιημένη Gaussian συνάρτηση ενεργοποίησης:

$$z_q = g_q(\mathbf{x}) \triangleq \frac{R_q(\mathbf{x})}{\sum_{k=1}^{\ell} R_k(\mathbf{x})} = \frac{\exp[-|\mathbf{x} - m_q|^2/2\sigma_q^2]}{\sum_{k=1}^{\ell} \exp[-|\mathbf{x} - m_k|^2/2\sigma_k^2]}$$

Με αυτόν τον τρόπο ο κρυφός κόμβος *q* δίνει μια μέγιστη απόκριση στα διανύσματα εισόδου x τα οποία είναι κοντύτερα προς την τιμή m_q κάθε κρυφός κόμβος q έχει το δικό του δεκτικό πεδίο Rq(x) στο χώρο εισόδου, που είναι μια περιοχή με κέντρο την τιμή m_q με μέγεθος ανάλογο προς το σ_q^2 όπου τα m_q και σ_q^2 είναι η μέση τιμή και η διασπορά της q Gaussian συνάρτησης αντίστοιχα. Οι Gaussian συναρτήσεις αποτελούν βέβαια ένα συγκεκριμένο παράδειγμα των συναρτήσεων radial basis. Η τελική έξοδος του δικτύου θα είναι απλά το ζυγισμένο άθροισμα των εξόδων των κρυφών κόμβων:

$$y_i = \alpha_i\left(\sum_{q=1}^{\ell} w_{iq} z_q + \theta_i\right)$$

όπου αί( ) είναι η έξοδος της συνάρτησης ενεργοποίησης και η θ_i η τιμή κατωφλίου.

Στην μελέτη των Κ.Καλαϊτζάκή, Γ.Σταυρακάκη και Ε.Αναγνωστάκη επιλέγεται όχι ένα σταθερό και εκ των προτέρων ορισμένο κέντρο για κάθε μια από τις συναρτήσεις radial basis, αλλά με τη χρήση της μεθόδου των ορθογωνίων ελάχιστων τετραγώνων και τον αντίστοιχο αλγόριθμο εκπαίδευσης επιλέγονται διαδοχικά κατάλληλα RBF κέντρα ένα προς ένα από τα σημεία του σέτ εκπαίδευσης μέχρις ότου επιτευχθεί το βέλτιστο δίκτυο. Με τη χρήση της συγκεκριμένης μεθόδου παράγεται τελικά ένα δίκτυο RBF αρκετά μικρότερο από ότι ένα τυχαία επιλεγόμενο RBF δίκτυο. Ο χρόνος εκπαίδευσης ενός RBF δικτύου προκύπτει αρκετά μικρότερος από εκείνον ενός συμβατικό νευρωνικού δικτύου ανάστροφης διάδοσης.

Ένα σημαντικό μειονέκτημα της παραπάνω μεθόδου σχετίζεται με το μέγεθος διάστασης του χώρου εισόδου. Για μεγάλο αριθμό από μονάδες εισόδου, ο αριθμός των απαιτούμενων συναρτήσεων radial basis μπορεί να γίνει δραματικά μεγάλος. Κατά συνέπεια, η συνεπακόλουθη ύπαρξη πολλών κέντρων (των αντίστοιχων RBF συναρτήσεων) οδηγεί σε ένα επίσης μεγάλο αριθμό παραμέτρων προς υπολογισμό, με αποτέλεσμα το δίκτυο να είναι τελικά υπερευαίσθητο στις λεπτομέρειες του συγκεκριμένου συνόλου εκπαίδευσης οδηγώντας έτσι σε φτωχά αποτελέσματα όσον αφορά την ικανότητα γενίκευσης του δικτύου (υπερεκπαίδευση). Για την αποφυγή του προβλήματος αυτού χρησιμοποιείται η τεχνική της ομαλοποίησης μηδενικής τάξης (zero-order regularization RBF ZORRBF). Τα καλύτερα αποτελέσματα επιτυγχάνονται χρησιμοποιώντας εισόδους από 0 εώς1. Η δε τιμή της παραμέτρου ομαλοποίησης λ τέθηκε μέσα από τη διαδικασία της δοκιμής και του λάθους ίση με 0.00002 και 0.00008 για την περίπτωση του μέγιστου και ελάχιστου φορτίου αντίστοιχα.

Επαναληπτικά δίκτυα πραγματικού χρόνου (recurrent real-time networks)

Η περιγραφή της δομής του συγκεκριμένου δικτύου έχει ως εξής. Σε ένα δίκτυο με N νευρώνες και N_i εξωτερικές εισόδους, x(t) είναι η εξωτερική είσοδος που εφαρμόζεται στο δίκτυο τη στιγμή t και y(t+ 1) το διάνυσμα εξόδου τη στιγμή t+ 1. Το διάνυσμα εισόδου x(t) και το διάνυσμα εξόδου y(t+ 1) αποτελούν από κοινού το διάνυσμα u(t). Το δίκτυο έχει δύο στρώματα, ένα συγχωνευμένης εισόδου-εξόδου και ένα στρώμα επεξεργασίας. Το δίκτυο είναι πλήρως διασυνδεδεμένο, έτσι υπάρχουν Ni*N συνδέσεις προς τα εμπρός και N^2 αναδρομικές συνδέσεις. Θεωρώντας ότι W είναι ο N (Ni+N) πίνακας των βαρών του δικτύου, η είσοδος κάθε νευρώνα θα δίνεται από τη σχέση :

$$v_j(t) = \sum_{i=1}^{N_i+N} w_{ji}(t)u_i(t)$$

Την επόμενη χρονική στιγμή *t*+ 1, η έξοδος κάθε νευρώνα του στρώματος επεξεργασίας θα δίνεται από τη σχέση

$$y_j(t+1) = F(v_j(t))$$

παρατηρώντας ότι η εξωτερική είσοδος x(t) δεν επηρεάζει την έξοδο κανενός νευρώνα μέχρι τη στιγμή t+1. Η παρακάτω σχέση μας δίνει τη διαφορά της επιθυμητής από τη υπολογισμένη έξοδο κάθε νευρωνικού.

$$e_j(t) = d_j(t) - y_j(t)$$

Στόχος η ελαχιστοποίηση της συνάρτησης :

$$E_{\text{total}} = \sum_{t=1}^{t_{\text{final}}} E(t)$$

Οπού tfinal είναι το τέλος του τρεξίματος του νευρωνικού δικτυού (NN) .

Τελικά με βάση την παρακάτω απαίτηση

$$\frac{\partial E_{\text{total}}}{\partial W} = 0 \Rightarrow \sum_{t=1}^{t_{\text{final}}} \frac{\partial E_{\text{total}}}{\partial W} = 0 \Rightarrow \sum_{t=1}^{t_{\text{final}}} \nabla_{\mathbf{w}} E(t) = 0$$

προκύπτουν οι ακόλουθοι κανόνες ανανέωσης για τα βάρη ενός εξωτερικού νευρώνα (α) και ενός νευρώνα εξόδου (β).

Σχέση α

$$w_{k\ell}(t+1) = w_{k\ell}(t) + \eta \sum_{j=1}^{N} e_j(t)\, F'(v_j(t)) \times \sum_{i=1}^{N_l+N} [\delta_{kj} u_\ell(t)]$$

Σχέση β

$$w_{k\ell}(t+1) = w_{k\ell}(t) + \eta \sum_{j=1}^{N} e_j(t)\, F'(v_j(t)) \sum_{i=1}^{N_l+N} \left[w_{ji}(t) \frac{\partial y_j(t)}{\partial w_{k\ell}(t)} + \delta_{kj} u_\ell(t) \right]$$

Στην περίπτωση αυτή χρησιμοποιούνται 24 διαφορετικά νευρωνικά δίκτυα για την πρόβλεψη του φορτιού της επόμενης ημέρας.

Αναδρομικά επαναληπτικά νευρωνικά δίκτυα (autoregressive recurrent neural networks ARNN)

Η τελευταία κατηγορία νευρωνικών δικτύων που παρουσιάζεται αποτελεί ένα υβριδικό τύπο τροφοδοτούμενων προς τα εμπρός και προς τα πίσω νευρωνικών δικτύων. Υπάρχουν δύο κρυφά στρώματα με σιγμοειδείς συναρτήσεις μεταφοράς και με έναν μονάχα γραμμικό κόμβο εξόδου. Η τοπολογία του ARNN επιτρέπει την επαναληπτικότητα μονάχα στο πρώτο κρυφό στρώμα. Για το σκοπό αυτό το μοντέλο BP επεκτείνεται έτσι ώστε να περιλάβει μια αναδρομική (autoregressive) μνήμη, μια μορφή αυτόανάδρασης όπου η έξοδος εξαρτάται επίσης από το ζυγισμένο άθροισμα των προηγούμενων εξόδων. Για την εκπαίδευση του δικτύου χρησιμοποιείται ένας τροποποιημένος BP αλγόριθμος, ο οποίος περιλαμβάνει δυναμικές επαναληπτικές συναρτήσεις του χρόνου. Ο μαθηματικός προσδιορισμός του μοντέλου αυτού θα έχει ως εξής:

$$y(t) = O(t) = \sum_{\ell} W_{\ell}^{O} Q_{\ell}(t), \quad Q_{\ell} = f(S_{\ell}), \quad S_{\ell}$$
$$= \sum_{j} W_{j\ell}^{H} Z_{j}(t)$$

και

$$Z_j(t) = f(H_j(t)),$$

$$H_j(t) = \sum_{k=1}^{k=n} W_{jk}^{D} Z_j(t-k) + \sum_{i} W_{ij}^{I} I_i(t)$$

όπου $I_i(t)$ η i-οστή είσοδος του ARNN, $H_j(t)$ το άθροισμα των εισόδων του j-οστού επαναληπτικού νευρώνα στο πρώτο κρυφό στρώμα, $Z_j(t)$ η έξοδος του j-οστού επαναληπτικού νευρώνα, $S_l(t)$ το άθροισμα των εισόδων του j-οστού νευρώνα στο δεύτερο κρυφό στρώμα, $Q_l(t)$ η έξοδος του j-οστού νευρώνα και O(t) η τελική έξοδος του δικτύου. Οι πίνακες W με τους διαφορετικούς δείκτες αποτελούν τους πίνακες βαρών των αντίστοιχων στρωμάτων του δικτύου.
Συμβολίζοντας με d(t) και y(t) τις επιθυμητές και πραγματικές αποκρίσεις του ARNN αντίστοιχα, η συνάρτηση σφάλματος θα δίνεται από τη σχέση:

$$E = \frac{1}{2}[d(t) - y(t)]^2$$

Ο κανόνας ανανέωσης των βαρών θα δίνεται από τη σχέση:

$$W(n+1) = W(n) + n\left(-\frac{\partial E}{\partial W}\right) + \alpha \Delta W(n)$$

Όπου:

$$\frac{\partial E}{\partial W} = -e(t)\frac{\partial y(t)}{\partial W} = -e(t)\frac{\partial O(t)}{\partial W}$$

$$\frac{\partial O(t)}{\partial W_\ell^O} = Q_\ell(t)$$

$$\frac{\partial O(t)}{\partial W_{j\ell}^H} = f'(S_\ell)W_\ell^O Z_j(t)$$

$$\frac{\partial O(t)}{\partial W_{jk}^D} = \delta_j(t)\frac{\partial Z_j(t)}{\partial W_{jk}^D}$$

$$\frac{\partial O(t)}{\partial W_{ij}^I} = \delta_j(t)\frac{\partial Z_j(t)}{\partial W_{ij}^I} \quad \text{with}$$

$$\delta_j(t) = \sum_\ell f'(S_\ell)W_\ell^O W_{j\ell}^H$$

και

$$\frac{\partial Z_j(t)}{\partial W_{jk}^D} = f'(H_j)\left(Z_j(t-k) + \sum_{n=1}^{n=5} W_{jn}^D \frac{\partial Z_j(t-n)}{\partial W_{jk}^D}\right),$$

$$\frac{\partial Z_j(0)}{\partial W_{jk}^D} = 0$$

$$\frac{\partial Z_j(t)}{\partial W_{ij}^I} = f'(H_j)\left(I_i(t) + \sum_{n=1}^{n=5} W_{jn}^D \frac{\partial Z_j(t-n)}{\partial W_{ij}^I}\right),$$

$$\frac{\partial Z_j(0)}{\partial W_{ij}^I} = 0$$

ΔΥΝΑΤΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ

Έχουν αναπτυχθεί διαφορετικά μοντέλα για την επίτευξη ακρίβειας στην πρόβλεψη του φορτίου. Μεταξύ των μοντέλων αυτών είναι η παλινδρόμηση, οι στατιστικές μέθοδοι, οι μέθοδοι του χώρου κατάστασης, τα μοντέλα ARIMA, τα μοντέλα προσομοίωσης, τα τεχνητά νευρωνικά δίκτυα και άλλα.
Στη συνέχεια αναφέρονται χαρακτηριστικά ορισμένων βασικών μοντέλων.
Τα μοντέλα **απόκλισης εξαρτημένης στατιστικής μεταβλητής**, συνήθως θεωρούν ότι η ΟΤΣ μπορεί να διαιρεθεί σε ένα πρότυπο στοιχείο ΟΤΣ και σε ένα στοιχείο γραμμικά εξαρτημένο από κάποιες επεξηγηματικές μεταβλητές. Το μοντέλο αυτό γράφεται:

$$z(t) = b(t) + \sum_{i=1}^{n} a_i \cdot y_i(t) + \varepsilon(t)$$

όπου $b(t)$ είναι η πρότυπη ΟΤΣ, $\varepsilon(t)$ ένα στοιχείο του λευκού θορύβου και $yi(t)$ οι ανεξάρτητες επεξηγηματικές μεταβλητές. Οι πιο τυπικές επεξηγηματικές μεταβλητές είναι οι καιρικοί παράγοντες.

Από τις πιο διαδεδομένες κατηγορίες μοντέλων δυναμικών προβλέψεων είναι τα μοντέλα **στοχαστικών χρονοσειρών**. Στην κατηγορία αυτή εντάσσονται τα μοντέλα ARMA (Auto-Regressive Moving Average), τα μοντέλα ARIMA (Integrated Auto-Regressive Moving Average), οι μέθοδοι Box-Jenkins, μοντέλα γραμμικών χρονοσειρών κ.τ.λ.. Το βασικό μοντέλο ARIMA, γράφεται ως εξής:

$$\phi(B)\nabla^d z(t) = \theta(B)a(t)$$

με,

$z(t),\ t=1,\ 2,\ \ldots,\ N$	η μοντελοποιημένη χρονοσειρά
$\alpha(t),\ t=1,\ 2,\ \ldots,\ N$	η ακολουθία λευκού θορύβου
$\varphi(B)=1-\varphi_1 B-\ldots-\varphi_p B^p$	η AR πολυωνυμική παράμετρος
$\theta(B)=1-\theta_1 B-\ldots-\theta_q B^q$	η MA πολυωνυμική παράμετρος
B	είναι ο *backward shift operator* ($B^n(z(t))=z(t-n)$)
$\nabla = 1 - B$	είναι ο *backward difference operator'*

Τέλος, τα μοντέλα **χώρων κατάστασης**, η ΟΤΣ στο χρόνο t μπορεί να γραφτεί:

$$z(t) = c^T \cdot x(t)$$

όπου,

$$x(t+1) = A \cdot x(t) + B \cdot u(t) + w(t)$$

Το διάνυσμα κατάστασης στο χρόνο t είναι το $x(t)$. Η $u(t)$ είναι ένα διάνυσμα εισόδου βασισμένο στη μεταβλητή καιρού, και η $w(t)$ ένα διάνυσμα τυχαίων εισόδων λευκού θορύβου. Οι πίνακες, A, B και το διάνυσμα c είναι σταθερές που λαμβάνονται.
Στην πραγματικότητα το βασικό μοντέλου χώρου κατάστασης μπορεί να μετατραπεί σε μοντέλο ARIMA και αντιστρόφως. Έτσι δεν υπάρχει θεμελιώδης διαφορά ανάμεσα στις ιδιότητες των δύο τύπων μοντέλων.

Κεφάλαιο 5

ΠΕΡΙΓΡΑΦΗ ΝΕΥΡΩΝΙΚΟΥ ΜΟΝΤΕΛΟΥ

Το νευρωνικό μας μοντέλο είναι ένα πρόσθιας τροφοδότησης (feedforward) νευρωνικό δίκτυο το οποίο εκπαιδεύεται από ένα γενετικό αλγόριθμο. Το μοντέλο αποτελείται από τρία στρωματά:
Α) ένα στρώμα εισόδου το οποίο έχει δυο εισόδους ,
Β) ένα κρυμμένο στρώμα που αποτελείται από τρεις νευρώνες και
Γ) ένα στρώμα εξόδου το οποίο έχει μια έξοδο .
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.
Ο ρόλος του γενετικού αλγορίθμου είναι η εκπαίδευση του μοντέλου δηλαδή η εξεύρεση των καταλλήλων βαρών του μοντέλου ούτως ώστε να ελαχιστοποιηθεί το σφάλμα πρόβλεψης . Ο αλγόριθμος έχει μέγιστο αριθμό γενεών 20 που σημαίνει ότι θα αλλάξει 20 φορές τα βάρη που συνδέονται οι νευρώνες για να μειώσει το σφάλμα πρόβλεψης. Άρα κάθε φορά που εκπαιδεύει το μοντέλο , κάνει πρόβλεψη, υπολογίζει τα σφάλματα εκπαίδευσης και πρόβλεψης και επιστρέφει πάλι πίσω δίνοντας νέες τιμές στα βάρη (η επόμενη γενιά) από τα οποία θα πάρει νέα πρόβλεψη και νέα σφάλματα (εκπαίδευσης και πρόβλεψης) τα οποία πρέπει να είναι μικρότερα από τα προηγούμενα, άρα θα έχουμε και καλύτερη πρόβλεψη λόγω καλύτερης εκπαίδευσης.
Το μοντέλο χρησιμοποιεί το 80 % των δεδομένων για το στάδιο της εκπαίδευσης και το υπόλοιπο 20 % για το στάδιο της πρόβλεψης.
Συνοπτικά το πρόγραμμα διαβάζει τα στοιχειά χρησιμοποιώντας δυο εισόδους προχωρώντας ένα βήμα (ένας μηνάς) μπροστά για κάθε είσοδο .
Οι είσοδοι του μοντέλου αλλά και η έξοδος φαίνονται παρακάτω.

```
input11=xlsread ('Tablec','NT', 'd4:d401');
input12=xlsread ('Tablec','NT', 'd5:d402');
output1=xlsread ('Tablec','NT', 'd3:d400');
```

Ένας βρόχος τρέχει όλο το δείγμα των στοιχειών και διαχωρίζει τις μεταβλητές x (μηνάς, έτος του φορτίου) και y (τιμή φορτιού).

```
for j=1:length(output1)
x=mydata(:,1:2);
y=mydata(:,3);
```

Κατόπιν περνάει στο στάδιο της βελτιστοποίησης των βαρών. Η βελτιστοποίηση των βαρών γίνεται από το γενετικό αλγόριθμο.
Το μοντέλο ρυθμίζει τα στοιχειά συμφωνά με τον αριθμό των καθυστερήσεων των μεταβλητών:

```
if nlag == 0,
   y = assetx(1+delay:end,target);
   x = assetx(:,target(end)+1:end,:);
   else [y x] = mylagvv(assetx, nlag, target, delay);
end;
```

Τα φέρνει σε μορφή πίνακα:

```
yxmat = [y x];
  [nrow ncol] = size(x);
  [nrowy ncoly] = size(y);
```

Έπειτα διαχωρίζει τα στοιχειά σε δυο δείγματα: α) το δείγμα εκπαίδευσης δηλαδή τα στοιχειά που χρησιμοποιούνται για την εκπαίδευση του μοντέλου και β) το δείγμα πρόβλεψης δηλαδή τα στοιχειά που χρησιμοποιούνται για να κάνει πρόβλεψη το μοντέλο συμφωνά πάντα με το ποσοστό που έχει οριστεί και είναι 80% για την εκπαίδευση και 20% για τη πρόβλεψη:

```
nrow1 = round(percent * nrow);
nrow11 = nrow1 + 1;
[nrow12 nrow13] = size(x(1:nrow1,:));
yy = y(1:nrow1,:);
xx = [ x(1:nrow1,:)];
[nrow ncol] = size(x);
```

Ακολούθως μεταφέρει την τάση όλων των στοιχειών x και y και των δυο δειγμάτων με τη παρακάτω εντολή:

```
yz = detrend(y,0) ./ kron(ones(rx,1), sigy);
xz = detrend(x,0) ./ kron(ones(rx,1),sigx);;
```

Στο επόμενο του βήμα μετατρέπει τις τιμές των στοιχειών στο διάστημα 0 έως 1 με τη βοήθεια της μεθόδου Helga Peterson :

```
for i = 1:cy,
   ys(:,i) = hsquasher(y(:,i), smax, smin);
   yss(:,i) = 1 ./ (1+ exp(-(yz(:,i))));
end

for i = 1:cx,
   xs(:,i) = hsquasher(x(:,i), smax, smin);
   xss(:,i) = 1 ./(1+ exp(-(xz(:,i))));
end
```

και λύνει ένα γραμμικό μοντέλο για την εξεύρεση της πρόβλεψης το οποίο χρησιμοποιεί αντιστροφή και πολλαπλασιασμούς πινάκων:

```
betaols = inv(Praw1' * Praw1) * Praw1' * Traw;
```

όπου ο πίνακας Praw1 είναι ο πίνακας του πραγματικού x και ο πίνακας Traw είναι ο πίνακας του πραγματικού y.

Μόλις βρει την έξοδο η οποία είναι :

```
A1 = Praw1 * betaols ;
```

υπολογίζει το σφάλμα συγκρίνοντας πραγματικές και επιθυμητές τιμές:

```
err1 = Traw - A1;
```

Έπειτα καθορίζει τον αριθμό των παραμέτρων σε κάθε στρώμα.

```
if nlayer == 1, nparm = nneuron1*cp + nneuron1 + nneuron1 * ncoly +
ncoly;
     elseif nlayer == 2, nparm = nneuron1 * cp + nneuron1 + nneuron1 *
nneuron2 +  nneuron2 +  nneuron2 * ncoly + ncoly;
      else nparm = nneuron1 * cp + nneuron1 + nneuron1 * nneuron2 +
nneuron2 + nneuron2 * nneuron3 +  nneuron3 + nneuron3 * ncoly + ncoly;
end
```

Τέλος θα αντιστρέψει την μετατρεπόμενη έξοδο και τα στοιχειά πάλι στις κανονικές τιμές (από 0 έως 1που είχαν μετατραπεί προηγουμένως) με τη παρακάτω εντολή:

```
for i = 1:cy,
      A3x(:,i) = helgeyx(A3(:,i), maxy(i), miny(i),smax,smin);
      End
```

και θα υπολογίσει το σφάλμα με τις κανονικές τιμές αυτή τη φορά:

```
err3 = ydep - yhatnet;
```

Έπειτα θα κινήσει όλη αυτή τη διαδικασία και για το δήγμα πρόβλεψης .
Έτσι ολοκληρώνεται ένας κύκλος εκπαίδευσης –πρόβλεψης και θα αρχίσει ένα καινούργιος με σκοπό τη βελτιστοποίηση των βαρών και τη μείωση των σφαλμάτων . Όταν το μοντέλο φτάσει σε αυτό το σημείο της βέλτιστης λύσης τότε εκτυπώνει τα σφάλματα RMSE(root mean square error), MSE (mean square error) ,MAPE (Mean Absolute percentage Error), MAE (Mean Absolute Error), για την εκπαίδευση και τη πρόβλεψη. Επίσης εκτυπώνει το χρόνο που χρειάζεται το μοντέλο για να φτάσει στη λύση. Τα σφάλματα παρατίθενται από κάτω:

Η ρίζα του μέσου τετραγωνικού σφάλματος:

```
RMSE_nn_outof_SAMPLE=sqrt(norm(er_nn_ev)^2/length(er_nn_ev))
```

Μέσο απόλυτο σφάλμα:

```
MAE_nn_outof_SAMPLE=(1/length(er_nn_ev))*sum(abs(er_nn_ev))
```

Μέσο απόλυτο ποσοστιαίο σφάλμα:

```
MAPE_nn_outof_SAMPLE=(100/length(er_nn_ev))*sum(abs(er_nn_ev)./a
bs(yyyout(:,1)))
```

Το μέσο τετραγωνικό σφάλμα:

```
MSE_nn_outof_SAMPLE=(1/length(yyyout(:,1)))*norm(er_nn_ev)^2
```

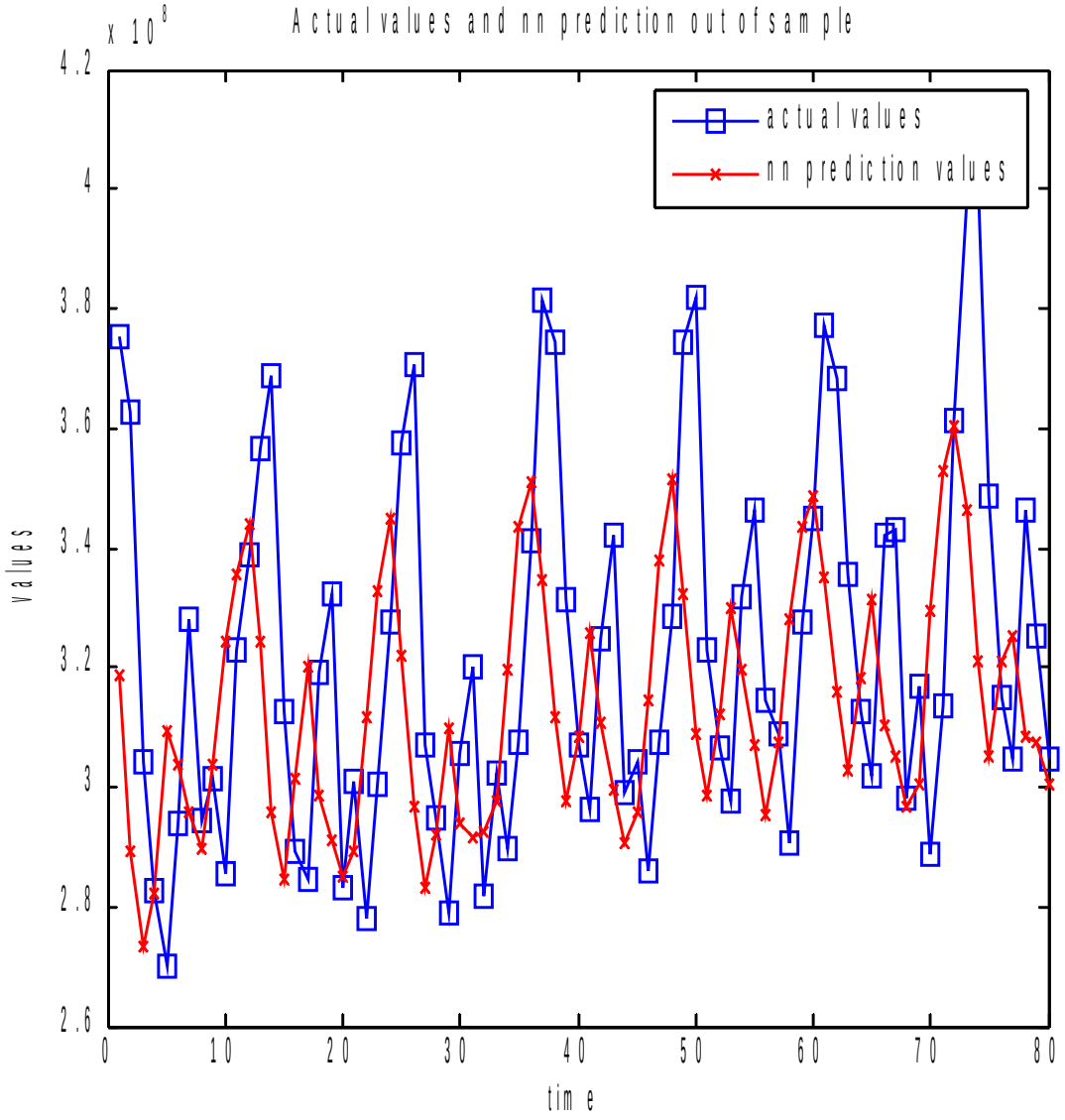
Ενώ ο χρόνος υπολογίζεται με την εντολή

```
toc
minutes=toc/60
```

Κεφαλαίο 6

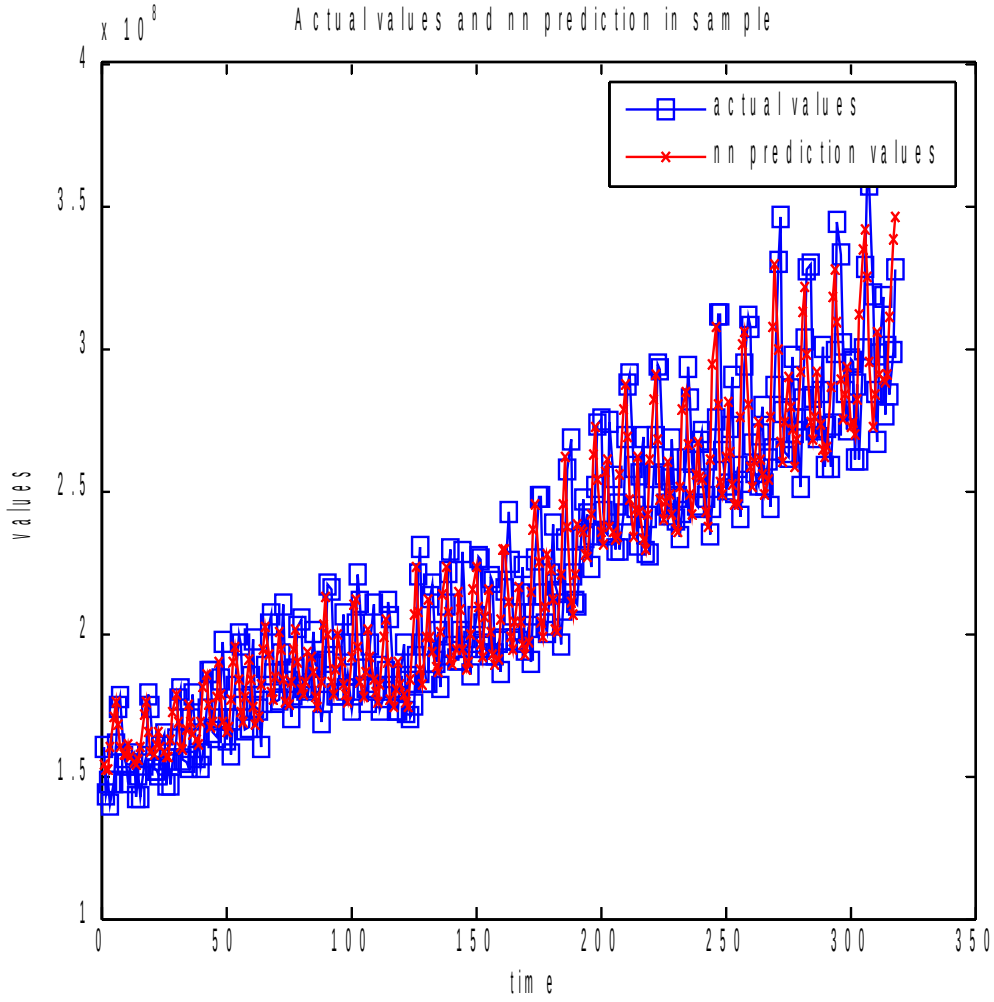
ΑΠΟΤΕΛΕΣΜΑΤΑ- ΣΥΜΠΕΡΑΣΜΑΤΑ

Από το παραπάνω πρόγραμμα παίρνουμε τα παρακάτω αποτελέσματα:
Στο σχήμα 37 μπορούμε να δούμε τη διαφορά μεταξύ ενός μέρους των πραγματικών και προβλέψιμων τιμών συμφωνά με το νευρωνικό μοντέλο πρόβλεψης με την χρήση δεδομένων εκτός δείγματος εκπαίδευσης (out of sample). Με μπλε χρώμα απεικονίζονται οι πραγματικές τιμές και με κόκκινο χρώμα οι τιμές που υπολογίζει το νευρωνικό δίκτυο.



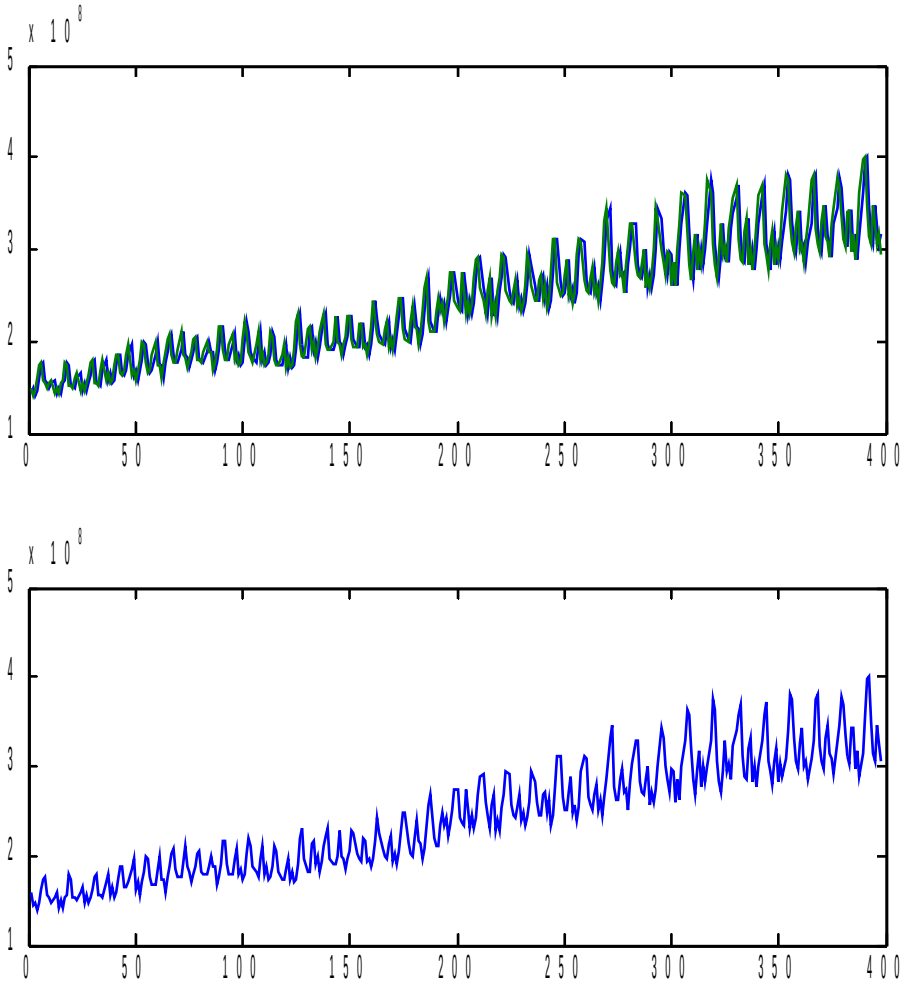
Σχήμα 37: Σφάλμα πρόβλεψης

Στο σχήμα 38 μπορούμε να παρατηρήσουμε τις πραγματικές και προβλέψιμες τιμές του νευρωνικού μοντέλου για το δείγμα εκπαίδευσης (in sample). Άρα μπορούμε να υπολογίσουμε και να δούμε την αυξομείωση του σφάλματος εκπαίδευσης.



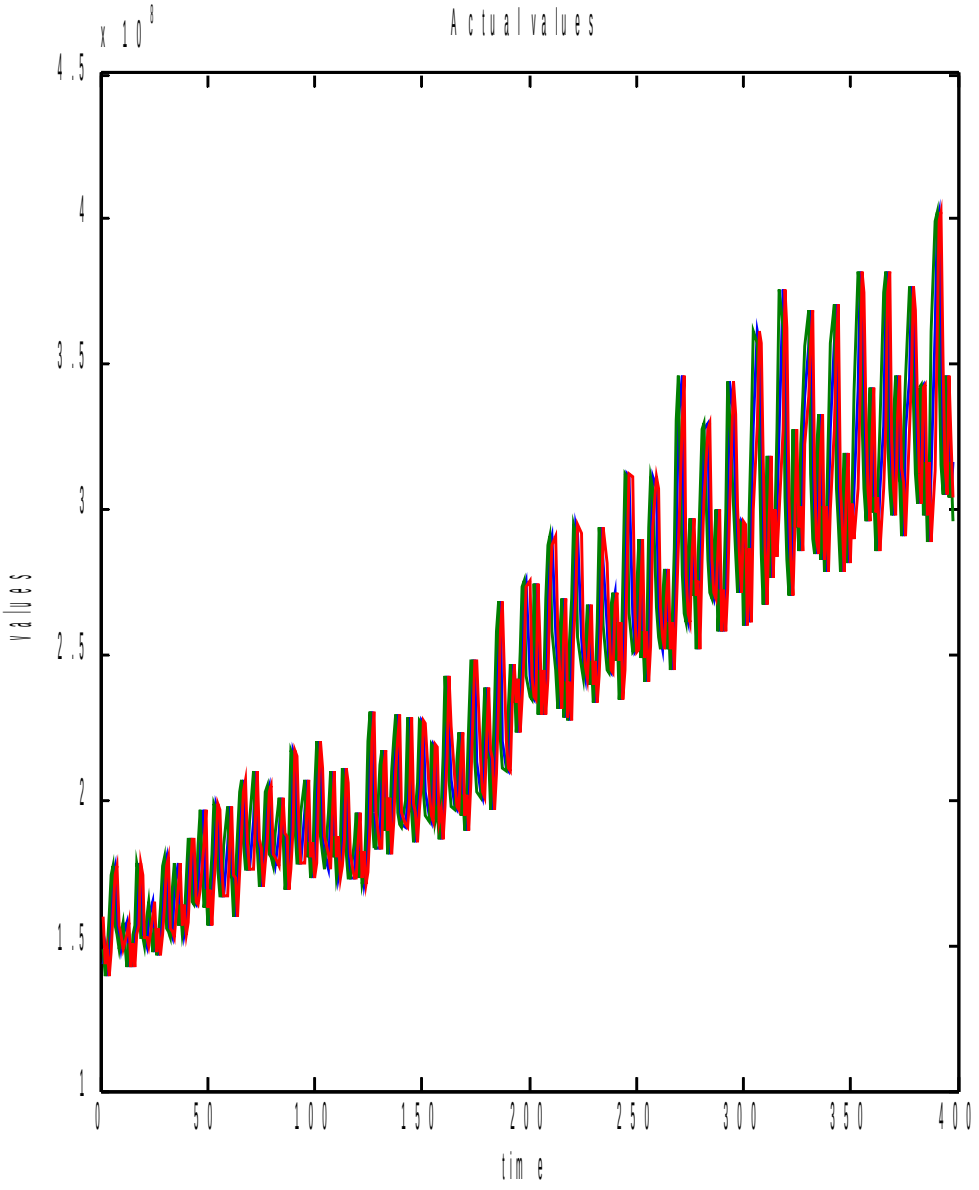
Σχήμα 38: Σφάλμα εκπαίδευσης

Στο σχήμα 39 παρατηρούμε τις δυο εισόδους του νευρωνικού μοντέλου



Σχήμα 39: Είσοδοι νευρωνικού μοντέλου

Τέλος στο σχήμα 40 μπορούμε να παρατηρήσουμε όλες τις πραγματικές τιμές δηλαδή όλο το δείγμα.



Σχήμα 40: Πραγματικές τιμές δείγματος

Τα σφάλματα για τα εντός δείγματος εκπαίδευσης (in sample) αλλά και τα εκτός δείγματος πρόβλεψης (out of sample) παρατίθενται στον πίνακα 11.

ΣΦΑΛΜΑΤΑ	ΕΝΤΟΣ ΔΕΙΓΜΑΤΟΣ (IN SAMPLE)	ΕΚΤΟΣ ΔΕΙΓΜΑΤΟΣ (OUT OF SAMPLE)
RMSE	2.2358e+007	3.3486e+007
MAE	1.8395e+007	2.7132e+007
MAPE	8.1315	8.1701
MSE	4.9988e+014	1.1213e+015

Πίνακας 11: σφάλματα πρόβλεψης και εκπαίδευσης νευρωνικού μοντέλου

Όπως φαίνεται παραπάνω τα αποτελέσματα που παίρνουμε έχουν μικρά σφάλματα άρα είναι αρκετά αξιόπιστα.

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 11.4513 min. ή ισοδύναμα 687.047000 seconds

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] Tawfiq Al-Saba, Ibrahim El-Amin* ΄΄Artificial neural networks as applied to long-term demand forecasting΄΄
[2] Ayca Kumluca Topallia , Ismet Erkmenb ΄΄ A hybrid learning for neural networks applied to short term load forecasting΄΄
[3] 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΄΄
[4] 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΄΄
[5] Benjamin F. Hobbs!,*, Udi Helman!, Suradet Jitprapaikulsarn",
Sreenivas Konda", Dominic Maratukulam ΄΄Artificial neural networks for short-term energy forecasting:
Accuracy and economic value΄΄
[6] Ronaldo R.B. de Aquino1, Aida A. Ferreira2, Manoel A. Carvalho Jr1,
Milde M.S. Lira1, Geane B. Silva1, and Otoni Nóbrega Neto1
΄΄ Combining Artificial Neural Networks and Heuristic Rules in a Hybrid Intelligent Load Forecast System΄΄
[7] Dipti Srinivasan* ΄΄Evolving artificial neural networks for short term load Forecasting΄΄
[8] 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 ΄΄
[9] Ot΄avio A.S. Carpinteiro a,*, Rafael C. Lemeb,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΄΄
[10] Mohammed El-Telbany a, Fawwaz El-Karmi b
΄΄ Short-term forecasting of Jordanian electricity demand using particle swarm optimization ΄΄

[11] K. Kalaitzakis *, G.S. Stavrakakis, E.M. Anagnostakis ´´ Short-term load forecasting based on artificial neural networks parallel implementation ´´
[12] Georges A. Darbellay*, Marek Slama ´´ Forecasting the short-term demand for electricity .Do neural networks stand a better chance? ´´
[13] 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´´
[14] Bahman Kermanshasi, Kiroshi Iwamiya ´´ Up to year 2020 load forecasting using neural nets ´´
[15] V.S. Kodogiannis *, E.M. Anagnostakis ´´ A study of advanced learning algorithms for short-term load forecasting ´´
[16] M. Beccali, M. Cellura *, V. Lo Brano, A. Marvuglia
´´ Forecasting daily urban electric load profiles using artificial neural networks´´
[17] D.Datta,S.A Tassou ´´ Artificial neural networks based electrical load prediction for food retail stores ´´
[18] Nahi Kandil a, Rene´ Wamkeue a,*, Maarouf Saad b, Semaan Georges ´´ An efficient approach for short term load forecasting
using artificial neural networks ´´
[19] 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 ´´
[20] M. Ulagammaia,_, P. Venkatesha, P.S. Kannana, Narayana Prasad Padhy ´´ Application of bacterial foraging technique trained artificial and wavelet neural networks in load forecasting´´
[21] Zhijian Hou a, Zhiwei Lian a,*, Ye Yao a, Xinjian Yuan
´´ Cooling-load prediction by the combination of rough set theory and an artificial neural-network based on data-fusion technique ´´
[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] C.S. Chen a,*, Y.M. Tzeng a, J.C. Hwang b ´´ The application of artificial neural networks to substation load forecasting ´´

[24] B. Satish a,*, K.S. Swarup b, S. Srinivas c, A. Hanumantha Raoa
´´ Effect of temperature on short term load forecasting using an integrated ANN ´´
[25] M. Ghiassi a,*, David K. Zimbra b, H. Saidane ´´ Medium term system load forecasting with a dynamic artificial neural network model´´
[26] E.A. Mohamad a M.M. Mansour a, S. E1-Debeiky a, K.G. Mohamad b, N.D. Rao ~'*,G. Ramakrishna ´´ Results of Egyptian unified grid hourly load forecasting using an artificial neural network with expert system interface ´´
[27] Huang-Chu Huang, Rey-Chue Hwang , Jer_guang Hsieh ´´ A new artificial intelligent peak power load forecaster based on non-fixed neural networks ´´
[28] Paras Mandal a, Tomonobu Senjyu a,*, Toshihisa Funabashi b
´´Neural networks approach to forecast several hour´´
[29] M. Beccali, M. Cellura, V. Lo Brano, A. Marvuglia ´´ Short-term prediction of household electricity´´
[30] Che-Chiang Hsu a,*, Chia-Yon Chen b´´ Regional load forecasting in Taiwan–applications´´
[31] Ping-Feng Pai a,., Wei-Chiang Hongb ´´Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms´´
[32] S.J. Yao, Y.H. Song*, L.Z. Zhang, X.Y. Cheng ´´ Wavelet transform and neural networks for short-term electrical load forecasting´´
[33] Atsalakis G., (2007) – Wind Energy production Forecasting by Neural Networks and Genetic Algorithms, European Computer conference. Verlag Spinger Athens, Greece
[34] Atsalakis G., Ucenic C., (2006) – *Forecasting the Electricity Demand Using a Neuro- fuzzy Approach versus Traditional Methods*, JOURNAL of WSEAS TRANSACTIONS on BUSINESS and ECONOMICS Issue 1, Volume 3, January 2006, pg.9-17, ISSN 1109-9526.
[35] 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

ΠΑΡΑΡΤΗΜΑ Α

Παρουσίαση ολοκλήρου του νευρωνικού μοντέλου

```
% Stam_ENERG CONSUME  FORECASTING BY NN
%for training is used genetic algorithms
% ATSALAKIS GEORGE 23/10/2007


close all  %clean the workspace
clear
clc

tic
%   preperation of data. The output coloumn concerns one step ahead
input11=xlsread ('Tablec','NT', 'd4:d401'); % retrieves the C column data of
range 1 to 365 from the data file
input12=xlsread ('Tablec','NT', 'd5:d402');

output1=xlsread ('Tablec','NT', 'd3:d400'); % the output is  from 3:366 (with
this way i do the one step ahead prediction
mydata1=[input11 input12  output1]; % all the data in columns



% Here the data for each of the plant is activated
mydata=mydata1;  % data of plant 1


figure(1) % a view of data
plot (mydata)
xlabel('time'); ylabel('values')
title('Actual values')
```

```
%loop that run thrue the number of samples

for j=1:length(output1)

x=mydata(:,1:2);
y=mydata(:,3);


figure(3)
subplot(2,1,1),plot(x)
subplot(2,1,2),plot(y)

data=[y x];

position= 1   %column number of dependent variable
architecture=[1 3 0 0]  %feddforward network with one hidden layer, with
three neurons
geneticdummy=1  %use genetic algoritm
maxgen=20
maxgen1=20 %number of generations for genetic algorithm
percent =0.8    % use 80 persent of data for all in sample estimation
nlags=0        %number of lags for variables
ndelay=0       %number of leads for the variables
niter=300       %number of interations for quasi-Newton method
helge=1 % EGO
delta=0.0001 %EGO
derdum=0 %EGO
nlag=0
delay=0
```

```
%  optimization of the weights of NN
%genetic algorith with gradient discentd (=1)
%genetic algorithm off, only gradient discent on (=0)
%only genetic algorithm,  gradient discent off (=2)

gendum=1    %%use genetic algoritm if gendum =1, use BFGS Quasi-
Newton optimization method if gendum= )


% from here starts the program  ffnet9.m
global nlayer nneuron1 nneuron2 nneuron3;
fun = 'ffnet9fun';
warning off;
nntwarn off;

info=[1 3 0 0];  % feedforward network with one hidden layer and 3 neurons
nlayer = info(1); % number of layers
nneuron1 = info(2);  % number of neurons in the 1st layer
nneuron2 = info(3);  % number of neurons in the 2nd layer
nneuron3 = info(4);  % number of neurons in the 3rd layer
target=1  %column number of dependent variable


assetx=data;
popsize = 50; pc = .9; pdes = 0; toler = .001; elite = 1;
[rr cc] = size(assetx); % rr=number of rows(100),  cc=number of columns(4)
% nlag = 1;   % number of lags and arguments.

yrhat = [assetx]; %an array 4 coloumns x 100 rows
```

```
% ajust the data according the lags
if nlag == 0,
   y = assetx(1+delay:end,target); % takes the first column as y - target=1 (is
defined as position= 1 in the base program)
   x = assetx(:,target(end)+1:end,:); % takes the columnes 2-end as x
   x = x(1:end-delay,:);
   else [y x] = mylagvv(assetx, nlag, target, delay);
end;


yxmat = [y x]; % an array 4 coloumn x 100 rows
   [nrow ncol] = size(x); % nrow=number of rows of x(100),  ncol=number
of columns of x (3)
      [nrowy ncoly] = size(y); % nrowy=number of rows of y (100),
ncoly=number of columns of y (1)

% separate the data acording the 'percent' value
nrow1 = round(percent * nrow); %number of rows -Round towards nearest
integer according the value of percent example (100*0.8=80)
nrow11 = nrow1 + 1; % =81
[nrow12 nrow13] = size(x(1:nrow1,:)); % example for sample data 100,
nrow12=80, nrow13=3(gives the number of coloumn)
yy = y(1:nrow1,:);  %takes the first 80 rows of y
xx = [ x(1:nrow1,:)]; %takes the first 80 rows of x
[nrow ncol] = size(x); % nrow=number of rows of x(100),  ncol=number of
columns of x(3)
nrow1 = round(percent * nrow); %number of rows -Round towards nearest
integer according the value of percent example (100*0.8=80)
nrow11 = nrow1 + 1; % =81
[nrow12 nrow13] = size(x(1:nrow1,:)); % example for sample data 100,
nrow12=80, nrow13=3(gives the number of coloumn)
yy = y(1:nrow1,:); %takes the first 80 rows of y
xx = x(1:nrow1,:); %takes the first 80 rows of x
smin = .1;
smax = .9;
```

```
% removes the trend
[rx, cx] = size(x);  %rx=number of rows(100),  cx=number of columns(3)
[ry, cy] = size(y);    %ry=number of rows(100),  cy=number of columns(1)
maxy = max(y);  %returns the indices of the maximum values in vector y
(0,0463)
miny = min(y);  %returns the indices of the minimum values in vector y
(-0.0471)
maxx = max(x);   %returns the indices of the maximum values in vector x
( 0.0670    0.0603    0.0726)
minx = min(x);  %returns the indices of the minimum values in vector x
( -0.0547   -0.0470   -0.0552)
meany = mean(y);  % calculates the mean of y (5.8602e-004)
sigy = std(y);    % calculates the standard daviation of y (0.0179)
yz = detrend(y,0) ./ kron(ones(rx,1), sigy);  % removes the trend of the all
data samples (100) of y
meanx = mean(x); % calculates the mean of x( 0.0025   -0.0003    0.0046)
sigx = std(x); % calculates the standard daviation of x (0.0192      0.0196
0.0214)
xz = detrend(x,0) ./ kron(ones(rx,1),sigx);;  % removes the trend of the all
data samples (100) of x

%squasing of data between 0-1 by Helge Petrson method McNelis book
page 84 & 64 & 24)
for i = 1:cy,  %cy=number of columns of y
   ys(:,i) = hsquasher(y(:,i), smax, smin); %squasing of data (100samples)
   yss(:,i) = 1 ./ (1+ exp(-(yz(:,i))));
end

for i = 1:cx,  %cx=number of columns of x
   xs(:,i) = hsquasher(x(:,i), smax, smin);
   xss(:,i) = 1 ./(1+ exp(-(xz(:,i))));
end
```

```
if helge == 0, PN = x; TN = y;
   elseif helge == 1, PN = xs; TN = ys; %PN=squashing x,  TN =squasing y
   elseif helge == 2, PN = xss; TN = yss;
   else PP = x; TT = y;
   [rp, cp] = size(PP);
      for i = 1:cp, PN(:,i) = (PP(:,i)- minx(i)) ./ (maxx(i)-minx(i));
      end
      for i = 1:cy,
      TN(:,i) = (TT(:,i)-miny(i))/(maxy(i)-miny(i));
      end
end

global P T;
P = PN(1:nrow1,:); %takes the first nrow1 (80) rows of squasing x
T = TN(1:nrow1,:);  %takes the first nrow1 (80) rows of squasing y
Praw = xx; %takes the original x (80 rows)
Traw = yy; %takes the original y (80 rows)
Praw1 = [Praw ones(length(Praw),1)]; % adds a column of ones next to
'Praw' column

%solving a linear system by inverting the matrix with inv(A)*b
%linear model
betaols = inv(Praw1' * Praw1) * Praw1' * Traw;  %is the inverse of the
square matrix (Praw(x) and Traw(y))
A1 = Praw1 * betaols; % output of 80 rows
err1 = Traw - A1; % calculate the error by comparison of actual y (80)and
estimated values
sse1 =  err1' * err1; % it gives 0,0121
Als = A1; % change the name of output of 80 rows
yhatls = Als;  %change the name of output A1 of 80 rows of the linear
model
[yyr yyc] = size(yhatls);  % yyr=80,  yyc=1
err1 = yy - yhatls; % calculate the error by comparison of actual y (80
samples) and estimated values
```

```
ssrsq1 = ones(1,yyc) - var(err1) ./ var(yy);  %it is 0.5300
hqols = nrow1 * log(sse1) + (ncol+1) * log(log(nrow1));     % it gives
(-353.3646+ 5.9100 =-347.4545)
[rp,cp] = size(P);

%nn
%defines the number of parameter in each layer (2*3)+2+(2*1)+1=11
if nlayer == 1, nparm = nneuron1*cp + nneuron1 + nneuron1 * ncoly +
ncoly;
     elseif nlayer == 2, nparm = nneuron1 * cp + nneuron1 + nneuron1 *
nneuron2 +  nneuron2 +  nneuron2 * ncoly + ncoly;
      else nparm = nneuron1 * cp + nneuron1 + nneuron1 * nneuron2 +
nneuron2 + nneuron2 * nneuron3 +  nneuron3 + nneuron3 * ncoly + ncoly;
end

nepoch = maxgen1; scale = 1; beta0 = randn(1,nparm); % initial values of
net parameters 1x11
tp = [25, nepoch, .02, .01, 1.07, .7, .9, 1.04];
pm = .33; elite = 1; pdes = 0;

% optimization of function  fun = 'ffnet9fun'
%  optimization of the weights of NN
%genetic algorith with gradient discentd (=1)
%genetic algorithm off, only gradient discent on (=0)
%only genetic algorithm,  gradient discent off (=2)

if gendum >= 1, beta = ...
     genetic5(fun,nparm,popsize,maxgen,pc,pm,elite,pdes,maxgen + 10,toler,
scale, beta0); %fun = 'ffnet9fun';

   else beta = .01 * ones(1, nparm);
end
[criterion,sse3,g,A3,W3,b3,W4,b4, W5, b5, W6, b6] = feval(fun,beta);
```

```
%'Display','iter' = displays output at each iteration
%'MaxIter' =Maximum number of iterations allowed.
%'maxfuneval'=Maximum number of function evaluations allowed

if gendum <= 1,
   options = optimset('MaxFunEvals', nepoch, 'MaxIter', nepoch,...
   'TolFun', delta);
   [beta,foutput,exitflag] = fminunc(fun, beta, options);

   [criterion,sse3,g,A3, W3,b3,W4,b4, W5, b5, W6, b6] = feval(fun,beta);
   else W3=W3; b3 = b3; W4 = W4; b4 = b4; W5 = W5, b5 = b5, W6 = W6,
b6 = b6;
end


%reverse transformation
if helge == 0, A3n = A3;
    elseif helge == 1, A3 = A3; %reverse transformation of A3(net output by
'helgeyx' function developed by Helge
                          %Petersohn if helge =1  (helgeyx)
      for i = 1:cy,
      A3x(:,i) = helgeyx(A3(:,i), maxy(i), miny(i),smax,smin);
      end
   A3 = A3;
   A3n = A3x; %Change the name to the A3x net output to A3n (table 80x1)
```

```
   elseif helge == 2, A3 = A3;  %reverse transformation (2) for DeLeo
   A3x = -log(1./A3- ones(size(A3)));
   A3x = real(A3x); [junkr, junkc] = size(A3x);
   A3x = kron(ones(junkr,1),meany) + A3x .* kron(ones(junkr,1),sigy);
   A3 = A3;
   A3n = A3x; A3n = real(A3n);
   else
      for i = 1:cy,  %reverse transformation
      A3n(:,i) = A3(:,i) * (maxy(:,i)-miny(:,i)) + miny(:,i);
      end
end

yhatnet = A3n; % %Change the name to the A3n net output to yhatnet (table
80x1)
ydep = yy;  % change the name of yy (original y data -60 samples
err3 = ydep - yhatnet; %calculates the error by comparison the actual y with
the output of the linear model
sse3 = sum(err3 .^2);  % it gives 0.0141
hqnet = nrow1 .* log(sse3) + (nparm) * log(log(nrow1)); % it gives
-340.7595+16.2526=-324.5069
% ssrsq3 = var(yhatnet) ./ var(ydep);
ssrsq3 = ones(1,yyc) - var(err3) ./ var(yy); % it gives 0.4498
sse = [sse1; sse3]; %sse =    0.0121
%                          0.0141
hqif = [hqols; hqnet]; %hqif = -347.4545 linear model
%                           -324.5069  net
ssrsq = [ssrsq1; ssrsq3]; %ssrsq = 0.5300 linear model
%                              0.4498 net
xxmean = mean(xx); %0.0050    0.0017    0.0064
xxend = xx(end,:); %takes the last row of original x (the 80th row)-0.0222
-0.0453   -0.0079
```

```
% deravative estimations??????
if derdum == 0,
   xstar = xxmean; %0.0050    0.0017    0.0064
   else xstar = xxend;
end

if helge == 0, pstar = xstar;
   elseif helge == 1,
      for i = 1:cx,
       pstar(1,i) = hsquasher(xstar(1,i), smax, smin, maxx(1,i), minx(1,i)); %it
gives 0.4894    0.4500    0.4806
      end
   elseif helge == 2, pstar = (1 ./ 1+ exp(-((xstar - meanx)))./ sigx);
   pstar = real(pstar);
   else
      for i=1:cp,
         pstar(:,i) = (xstar(:,i) - minx(i)) ./ (maxx(i) - minx(i));
      end
end
pstar = pstar;
hdelta = delta; %xhange the name from delta to hdelta (0)
[rp, cp] = size(P); %rp=80, cp=3
hdeltav = eye(cp) * hdelta;

%hdeltav =
%    1.0e-004 *
%    1.0000         0         0
%         0    1.0000         0
%         0         0    1.0000
```

```
if nlayer == 1,
   A3star =  pstar * W3+ b3; %A3star =   -0.2572   -1.2055
   A3star = 1./(1+exp(-A3star)); %A3star =    0.4361    0.2305
   A3star = A3star * W4 + b4; %A3star =    0.5128
elseif nlayer == 2,
   A3star = pstar * W3+ b3;
   A3star = 1./(1+exp(-A3star));
   A3star = A3star * W4 + b4;
   A3star = 1./(1+exp(-A3star));
   A3star = A3star * W5 + b5;
else
  A3star = pstar * W3+ b3;
  A3star = 1./(1+exp(-A3star));
  A3star = A3star * W4 + b4;
  A3star = 1./(1+exp(-A3star));
  A3star = A3star * W5 + b5;
  A3star = 1./(1+exp(-A3star));
  A3star = A3star * W6 + b6;
end


if helge == 0, A3star = A3star;
   elseif helge == 1,
   A3star = A3star;
      for ii = 1:cy,
        A3star(:,ii) = helgeyx(A3star(:,ii),maxy(1,ii), miny(1,ii),smax, smin);
%unscale the A3star
      end
   A3star = A3star; %A3star =  6.7903e-004
   elseif helge == 2,
   A3star = A3star;
   A3star = -log(1./A3star - ones(size(A3star)));
   A3star = meany + A3star .* sigy;
   A3star = A3star;
   else
```

```
for i = 1:cy,
        A3star(:,i) = A3star(:,i) .* (maxy(i)-miny(i)) + miny(i);
      end
end

for i = 1:cp, %cp=number of coloumn of x (1:3)
  xdel = xstar + hdeltav(i,:);
  if helge == 0,
      pstardel = xdel;
      elseif helge == 1,
        for j = 1:cx,
                pstardel(1,j)  =  hsquasher(xdel(1,j),  smax,  smin,  maxx(1,j),
minx(1,j));
        end  %pstardel =    0.4894    0.4500    0.4806
   elseif helge == 2,
   pstardel = logsig((xdel - meanx) ./ sigx); pstardel = real(pstardel);
   else
        for jj = 1:cx,
        pstardel(1,jj) = (xdel(1,jj) - minx(1,jj)) ./ (maxx(jj) - minx(jj));
        end

  end

%-------
   if nlayer == 1,
      A3d =  pstardel * W3+ b3;
      A3d = 1./(1+exp(-A3d));
      A3d = A3d * W4 + b4;
      elseif nlayer == 2,
      A3d = pstardel * W3+ b3;
      A3d = 1./(1+exp(-A3d));
      A3d = A3d * W4 + b4;
      A3d = 1./(1+exp(-A3d));
      A3d = A3d * W5 + b5;
```

```
 else A3d = pstardel * W3+ b3;
      A3d = 1./(1+exp(-A3d));
      A3d = A3d * W4 + b4;
      A3d = 1./(1+exp(-A3d));
      A3d = A3d * W5 + b5;
      A3d = 1./(1+exp(-A3d));
      A3d = A3d * W6 + b6;
  end
  %unscaling data
  if helge == 0, A3d = A3d;
     elseif helge == 1,
     A3d = A3d;
        for kk = 1:cy,
        A3d(1,kk) = helgeyx(A3d(1,kk), maxy(kk), miny(kk),smax,smin);
        end
     A3d = A3d;
     A3d = real(A3d);
     elseif helge == 2,
     A3d = A3d;
     A3d = -log(1./A3d - ones(size(A3d)));
     A3d = meany + A3d .* sigy;
     A3d = real(A3d);
     else
        for j=1:cy,
        A3d(:,j) =  A3d(:,j) .* (maxy(j)-miny(j)) + miny(j);
        end;
  end
   pdernum(i,:) = (A3d-A3star) ./ hdelta;  %pdernum = -0.0131
 %                                          -0.2881
 %                                          -0.2228
  clear A3d xdel; %A3d=6.5675e-004  xdel =    0.0050    0.0017    0.0065

end
```

```
%%%%%%%%%   out of sample evaluation    %%%%%%%%%

% out of sample linear model
% the data here must be squashed data
xout  = [x(nrow11:nrow,:)]; % original(without squashing) x (last 20 rows)-
evaluationa sample
yout1 = y(nrow11:nrow,:); % original (without squashing) y (last 20 rows)-
evaluationa sample
[n1 c1] = size(yout1); % n1=number of rows(20), c1=number of columns(1)
Poutraw = xout; %change the name of xout to Poutraw
Poutraw1 = [Poutraw ones(n1,1)]; % adds a 4th coloumn by ones
Toutraw = yout1; %change the name of yout1 to Toutraw
T1 = TN(nrow1+1:end,:); %it takes the last nrow+1:end (20) rows from the
squashed y data TN
P1 = PN(nrow1+1:end,:); %it takes the last nrow+1:end (20) rows from the
squashed x data TN

% calculation of linear model by out of sample data
A11 = Poutraw1 * betaols; %output of linear model by multiplication of
Poutraw with
%the betaols (the inverse of the square matrix (Praw(x) and Traw(y))
yhatls1 = A11; % change the name of output A11 to yhatls1
err11 = yout1 - yhatls1; %calculates the error copmared the actual y with the
estimated y
err11sq = err11 .^2; %calculates the ^2
rmsqe1 = sqrt(mean(err11sq)); %rmsqe1 =    0.0133
[rp1, cp1] = size(P1); % rp1=number of rows (20), cp1=number of coloumns
(3), for x data
```

```
%calculation of network output A31 by out of sample data
if nlayer == 1,
   A31 = P1 * W3+ kron(ones(rp1,1),b3);% The output of 1st neuron
   A31 = 1./(1+exp(-A31));
    A31 = A31 * W4 + kron(ones(length(T1),1),b4); %output of 2nd neuron
(squashed)
elseif nlayer == 2,
   A31 = P1 * W3+ kron(ones(rp1,1),b3);
   A31 = 1./(1+exp(-A31));
   A31 = A31 * W4 + b4;
   A31 = 1./(1+exp(-A31));
   A31 = A31 * W5 + b5;

else
  A31 = P1 * W3+ kron(ones(rp1,1),b3);
  A31 = 1./(1+exp(-A31));
  A31 = A31 * W4 + b4;
  A31 = 1./(1+exp(-A31));
   A31 = A31 * W5 + b5;
  A31 = 1./(1+exp(-A31));
  A31 = A31 * W6 + b6;
End


%reverse transformation (unsquashed of net output) by 'hsquasher' function
developed by Helge Petersohn

if helge == 0, yhatnet1 = A31;
   elseif helge == 1, A31z = A31; %change the name
      for i = 1:cy,
       A31x(:,i) = helgeyx(A31z(:,i),maxy(i), miny(i), smax, smin); %reverse
trasformation of net output
      end
  A31net = A31x;
  yhatnet1 = A31net; %unsquashed  net output
```

```
elseif helge == 2,
   A31z = A31;
     A31x = -log(1./A31z - ones(size(A31z)));
   A31x = real(A31x);
   A31x = kron(ones(n1,1),meany) + A31x .* kron(ones(n1,1), sigy);
   A31net = A31x;
   yhatnet1 = A31net;
else
   for i = 1:cy,
      A31net(:,i) = A31(:,i) .* (maxy(i)-miny(i)) + miny(i);
   end
        yhatnet1 = A31net;
end

err31 = yout1-yhatnet1; %error calculation by copmarison of actual y with
net output (yhatnet1)
err31sq = err31 .^2; % %calculates the ^2
rmsqe3 = (mean(err31sq)) .^.5; %0,0131
rmsqe = [rmsqe1; rmsqe3];  % rmsqe =  0.0133

%                              0.0131
% hintonwb(W3,b3); pause
% subplot(211); barerr(err1); grid; subplot(212); barerr(err3); grid; pause
% subplot(211); barerr(err11); grid; subplot(212); barerr(err31); grid;


hqnet = hqif(ncoly+1:end,:); %hqnet = -324.5069
pdernetc = pdernum; %pdernetc =  -0.0131
%                            -0.2881
%                            -0.2228
w1net = W3; %gives the weights that calculated during the training
w2net = W4;
w3net = W5;
w4net = W6;
```

```
ssebest = sse;
rsqbest = [ssrsq];
rmsqbest = real([rmsqe]);
hqifbest = [hqif];
W1 = betaols(1:end-1,:); %W1 = -0.1267
%                        -0.4174
%                        -0.2071
pderbest = [W1 pdernetc];  %pderbest =   -0.1267   -0.0131
%                                -0.4174   -0.2881
%                                -0.2071   -0.2228
w1best = w1net;
w2best = w2net;
w3best = w3net;
w4best = w4net;
b1best = b3;
b2best = b4;
b3best = b5;
b4best = b6;
yyy = [yy yhatls yhatnet; yout1 yhatls1 yhatnet1];
yyy = real(yyy); %the same as yyy
yyyhat = yyy(1:nrow1,:); %it takes the firsr (1:nrow1,:)rows - in sample (yy
yhatls yhatnet)
yyyout = yyy(nrow1+1:end,:); %t takes the last (nrow1+1:end,:)rows - out of
sample (yout1 yhatls1 yhatnet1)
yout = yyyout; %change the name
ndim = ncoly; %ndim =     1
for i = 1:ndim,

  erroroutls(:,i) = yout(:,ndim+i) - yout(:,i);
   erroroutnet(:,i) = yout(:,2*ndim+i) - yout(:,i);
   rrmsq(1,i) = sqrt(mean(erroroutls(:,i) .^2)); %rrmsq =  0.0133

  %                                        0.0131
 rrmsq(2,i) = sqrt(mean(erroroutnet(:,i) .^2));
end
rmsqbest = rrmsq;
```

```
%clear global P T nlayer nneuron1 nneuron2 nneuron3;
 %  here ends the program  ffnet9.m

                    %[rsqnet01,yxmat,yrhat,      yhatls,yhatnet,yout1,      yhatls1,
yhatnet1,yyyhat,yyyout,yout,ydep,beta]...
    %            =ffnet9_my(data,    position,    percent,    nlags,    ndelay,
architecture,geneticdummy, maxgen, niter)

% it can give until 13 output variables
rsqnet01=rmsqbest
rsqnet0(j)=rsqnet01(2)
 RSQ(j,:)=[rsqnet0(j)]
end %here ends the fist loop (for j=1:in4)

% results in sample
yout2=yyyhat(:,3);

figure (100)
plot(yyyhat(end-60:end,1), 'b-s'), hold, plot(yout2(end-60:end), 'r-x');
legend('actual values','nn prediction values')
xlabel('time')
ylabel('values')
title('Actual values and nn prediction in sample')
er_nn=[yyyhat(:,1)- yyyhat(:,3)];

%Root Mean Square Error (RMSE)
RMSE_nn_IN_SAMPLE=sqrt(norm(er_nn)^2/length(er_nn))

%Mean Absolute Error (MAE)
MAE_nn_IN_SAMPLE=(1/length(er_nn))*sum(abs(er_nn))

%Mean Absolute percentage Error (MAPE)
MAPE_nn_IN_SAMPLE=(100/length(er_nn))*sum(abs(er_nn)./abs(yyyhat(
:,1)))

MSE_nn_IN_SAMPLE=(1/length(yyyhat(:,1)))*norm(er_nn)^2
```

```
%%%%%%%% out of sample evaluation %%%%%%%%



yout2=yyyout(:,3);


figure (200)
plot(yyyout(end-60:end,1), 'b-s'), hold, plot(yout2(end-60:end,1), 'r-x');
legend('actual values','nn prediction values')
xlabel('time')
ylabel('values')
title('Actual values and nn prediction out of sample')

er_nn_ev=[yyyout(:,1)- yyyout(:,3) ];


%Root Mean Square Error (RMSE)
RMSE_nn_outof_SAMPLE=sqrt(norm(er_nn_ev)^2/length(er_nn_ev))

%Mean Absolute Error (MAE)
MAE_nn_outof_SAMPLE=(1/length(er_nn_ev))*sum(abs(er_nn_ev))


%Mean Absolute percentage Error (MAPE)
MAPE_nn_outof_SAMPLE=(100/length(er_nn_ev))*sum(abs(er_nn_ev)./a
bs(yyyout(:,1)))

MSE_nn_outof_SAMPLE=(1/length(yyyout(:,1)))*norm(er_nn_ev)^2
```

```
insample_error=[RMSE_nn_IN_SAMPLE;          MAE_nn_IN_SAMPLE;
MAPE_nn_IN_SAMPLE; MSE_nn_IN_SAMPLE]


outofsample_error=[RMSE_nn_outof_SAMPLE;
MAE_nn_outof_SAMPLE;                        MAPE_nn_outof_SAMPLE;
MSE_nn_outof_SAMPLE]

allerrors=[insample_error outofsample_error]

toc
minutes=toc/60
```

Το παραπάνω πρόγραμμα περιέχει σχόλια πάνω στο κώδικα τα οποία είναι σε πράσινο χρώμα για να γίνονται ευκολότερα αντιληπτά και πάντα προηγείται ο χαρακτήρας % .

ΠΑΡΑΡΤΗΜΑ Β

Δεδομένα

ΔΕΔΟΜΕΝΑ

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 Γ

Περίληψη βιβλιογραφίας

1. 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.

2. 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. Available past data, which is the actual load data obtained from Turkish Electricity Authority, is used off-line and the model is prepared for on-line forecasts. This hybrid learning approach gives better results than the standard on-line learning since it takes the advantage of using experience gained by past data. This model reduces convergence time of on-line learning considerably since the weights are already brought near to optimal values.

3. 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). The neural model is made up of two self-organizing map (SOM) nets, one on top of the other. It has been successfully applied to domains in which the context information given by former events plays a primary role. The model was trained on load data extracted from a Brazilian electric utility, and compared to a multilayer perceptron (MLP) load forecaster.

4. 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. A Euclidean norm with weighted factors is used to evaluate the similarity between the forecast day and searched previous days.

5. 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 according to this paper 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.

6. Combining Artificial Neural Networks and Heuristic Rules in a Hybrid Intelligent Load Forecast System

Ronaldo R.B. de Aquino1, Aida A. Ferreira2, Manoel A. Carvalho Jr1, Milde M.S. Lira1, Geane B. Silva1, and Otoni Nóbrega Neto1

In this work, an Artificial Neural Network (ANN) is combined to Heuristic Rules producing a powerful hybrid intelligent system for short and mid-term electric load forecasting. The Heuristic Rules are used to adjust the ANN output to improve the system performance. The results confirmed the potential and suitability of the hybrid intelligent system implemented in PREVER compared to CELPE´s load forecasting system and ANN acting alone.

7. 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 artficial neural network was found capable of accurately forecasting one-day ahead hourly loads.

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

Eva Gonza´lez-Romera, Miguel A ´ ngel Jaramillo-Mora´n *,
Diego Carmona-Ferna´ndez

This work is focused on electric energy demand short term and medium to long term forecasting that is a fundamental tool for production and distribution companies because it provides them with a prediction of the market demand of electric energy. 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.

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

Ot´avio A.S. Carpinteiro a,*, Rafael C. Lemeb,a, Antonio C. Zambroni de Souza b,a,
Carlos A.M. Pinheiro a, Edmilson M. Moreira

A novel hierarchical hybrid neural model to the problem of long-term load forecasting is proposed in this paper. The neural model is made up of two self-organizing map nets – one on top of the other –, and a single-layer perceptron. The model is compared to a multilayer perceptron. Both the hierarchical and the multilayer perceptron models are

trained and assessed on load data extracted from a North-American electric utility. They are required to predict either once every week or once every month the electric peak-load and mean-load during the next two years.

10. Short-term forecasting of Jordanian electricity demand using particle swarm optimization

Mohammed El-Telbany a, Fawwaz El-Karmi b

This paper discusses and presents the results of utilizing neural network for forecasting the Jordanian electricity demand that is trained by particle swarm optimization technique, which is a new adaptive algorithm based on a social-psychological metaphor. The results of using this technique are compared with the results of using back-propagation algorithm and autoregressive moving average method.

11. 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.

12. Forecasting the short-term demand for electricity Do neural networks stand a better chance?

Georges A. Darbellay*, Marek Slama

In this study we address the problem which is faced by every supplier of electricity, i.e. forecasting the short-term electricity consumption. The introduction of new techniques has often been justified by invoking the nonlinearity of the problem. Our focus is directed to the question of deciding whether the problem is indeed nonlinear. The correlational analysis suggests that forecasting the short-term evolution of the Czech electric load is primarily a linear problem. This is confirmed by the comparison of the predictions.

13. 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. 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.

14. Up to year 2020 load forecasting using neural nets

Bahman Kermanshasi, Kiroshi Iwamiya

In this study two Ann's ,a three-layered back propagation and a recurrent neural network , are designed and tested to solve the long term load forecasting problem. Also two case studies , preservation of the status and structure reform , were also tested for predicting loads of longer term until 2020. The used data are economical data and the results are very accurate.

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

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

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. The results indicate that the load-forecasting models developed in this way provide more accurate forecasts, compared with conventional backpropagation network forecasting models.

16.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.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. 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.

17.Artificial neural networks 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 and radial basis function. The MLP netwok 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.

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

Nahi Kandil a, Rene´ Wamkeue a,*, Maarouf Saad b, Semaan Georges

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.

19.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

In this paper is developed a novel forecasting method, called 'RBF neural network (RBFNN) with combined residual error correction',. 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.

20.Application of bacterial foraging technique trained artificial and wavelet neural networks in load forecasting

M. Ulagammaia,_, P. Venkatesha, P.S. Kannana, Narayana Prasad Padhy

A new load forecasting (LF) approach using bacterial foraging technique (BFT) trained wavelet neural network (WNN) is proposed in this paper. Artificial neural network (ANN) is combined with wavelet transform called wavelet neural network is applied for LF. The parameters of translation and dilation in the wavelet nodes and the weighting factors in the weighting nodes are tuned using BFT optimization. The performance of the BFT-based WNN is better than existing delta rule-based methods.

21.Cooling-load prediction by the combination of rough set theory and an artificial neural-network based on data-fusion technique

Zhijian Hou a, Zhiwei Lian a,*, Ye Yao a, Xinjian Yuan

A novel method integrating rough sets (RS) theory and an artificial neural network (ANN) based on data-fusion technique is presented to forecast an air-conditioning load. In this paper, RS theory is applied to find relevant factors to the load, which are used as inputs of an artificial neural-network to predict the cooling load. To improve the accuracy of load forecasting results, a general load-prediction model, by synthesizing multi-RSAN (MRAN), is presented so as to make full use of redundant information. Actual prediction results show that, the MRAN forecasting model is better than the individual RSAN and moving average (ARIMA) ones, whose relative error is within 4%.

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

Pedro A. Gonza´lez, Jesu´s M. Zamarren˜o*

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 achieved results demonstrate the high precision reached with this system.

23.The application of artificial neural networks to substation load forecasting

C.S. Chen a,*, Y.M. Tzeng a, J.C. Hwang b

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.

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

B. Satish a,*, K.S. Swarup b, S. Srinivas c, A. Hanumantha Raoa

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 adaptive combiner is a weighted combination of outputs of the Basic ANN and the Forecaster. Forecasted load obtained from the proposed architecture is better than obtained from a single ANN.

25.Medium term system load forecasting with a dynamic artificial neural network model

M. Ghiassi a,*, David K. Zimbra b, H. Saidane

This paper presents the development of a dynamic artificial neural network model (DAN2) for medium term electrical load forecasting (MTLF). We present a yearly model that uses past monthly system loads to forecast future electrical demands. We also show that the inclusion of weather information improves load forecasting accuracy. Such models, however, require accurate weather forecasts, which are often difficult to obtain. Therefore, we have developed an alternative seasonal models that provide excellent fit and forecasts without reliance upon weather variables. Both the yearly and seasonal models produce mean absolute percent error (MAPE) values below 1%, demonstrating the effectiveness of DAN2 in forecasting medium term loads.

26. Results of Egyptian unified grid hourly load forecasting using an artificial neural network with expert system interface

E.A. Mohamad a M.M. Mansour a, S. E1-Debeiky a, K.G. Mohamad b, N.D. Rao ~'*,G. Ramakrishna

This paper presents the hourly load forecasting results of the Egyptian unified grid (EUG). The technique is based on a generalized model combining the features of ANN and an expert system. The above methodology makes the technique robust, updatable and provides for operator intervention when necessary. This property makes it especially suitable for the EUG where the load patterns are influenced mostly because of social activities, and weather contributes very little to load forecast. The results clearly demonstrate the advantage of the above methodology over statistical based techniques. The average absolute forecast errors for the proposed methodology is 2.63% with a standard deviation of 2.62% whereas, the conventional multiple regression method scores an average absolute error of 4.69% with a standard deviation of 4.03%.

27.A new artificial intelligent peak power load forecaster based on non-fixed neural networks

Huang-Chu Huang, Rey-Chue Hwang , Jer_guang Hsieh

In this paper a new artificial intelligent peak load forecaster constructed by non-fixed neural networks is developed. Several techniques, including gray analysis and stochastic back propagation learning rule with dynamic learning rate and momentum, are used in this forecaster in order to attain more accurate prediction in forecasting operation.