

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ



Πολυτεχνείο Κρήτης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΘΕΜΑ : ΠΡΟΒΛΕΨΗ ΠΩΛΗΣΕΩΝ ΜΕ ΧΡΗΣΗ
ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΚΑΙ ΝΕΥΡΟΑΣΑΦΩΝ
ΣΥΣΤΗΜΑΤΩΝ

Επιβλέπων καθηγητής : Ατσαλάκης Γεώργιος
Επιμέλεια : Κιαγιάς Ιωάννης

ΧΑΝΙΑ 2009

ΠΡΟΛΟΓΟΣ

Η διπλωματική αυτή εργασία αυτή εκπονήθηκε στα πλαίσια του προπτυχιακού επιπέδου του τμήματος Μηχανικών Παραγωγής και Διοίκησης του Πολυτεχνείου Κρήτης, υπό την επίβλεψη του Καθηγητή κ. Ατσαλάκη Γεώργιου. Θα ήθελα να ευχαριστήσω βαθύτατα τον επιβλέποντα καθηγητή μου κ. Ατσαλάκη Γεώργιο που μου έδωσε την ευκαιρία να ασχοληθώ με το αντικείμενο της πρόβλεψης πωλήσεων. Καθ' όλη τη διάρκεια της εκπόνησης της διατριβής αυτής δημιουργήθηκε ένα άρτιο κλίμα συνεργασίας που με τη συνεχή στήριξη μέσω πολύτιμων συμβουλών και καθοδήγησης, συνέβαλε στην επιτυχημένη ολοκλήρωση της.

Θα ήθελα να ευχαριστήσω επίσης τα υπόλοιπα δύο μέλη της Τριμελούς Εξεταστικής Επιτροπής, τον καθηγητή κ. Μουστάκη Βασίλειο και τον Καθηγητή κ. Σκιαδά Χρήστο για τις υποδείξεις και τις παρατηρήσεις τους.

Την εργασία μου θα ήθελα να την αφιερώσω πρωταρχικά στη αγαπημένη μου μητέρα, η οποία βρισκόταν πάντα στο πλευρό μου και που δυστυχώς δεν πρόλαβε να με δει πτυχιούχο. Την αφιερώνω επίσης στην οικογένεια μου και στους φίλους μου που με στήριξαν όλα αυτά τα χρόνια της πορείας μου στο πολυτεχνείο.

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ.....	4
2. ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ.....	5
3. ΠΡΟΤΕΙΝΟΜΕΝΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ.....	18
3.1. Τεχνητά Νευρωνικά Δίκτυα.....	18
Ορισμός νευρωνικού δικτύου.....	18
Ιστορικά Στοιχεία.....	19
Από τα Βιολογικά Νευρωνικά Δίκτυα στα Τεχνητά.....	22
Χρήση Νευρωνικών Δικτύων και Εφαρμογές.....	24
Αρχιτεκτονική Νευρωνικών Δικτύων.....	26
Είδη Αρχιτεκτονικών Δομών.....	30
Λειτουργία Νευρωνικών Δικτύων.....	33
Εκπαίδευση Νευρωνικών Δικτύων.....	34
Ο Αλγόριθμος Μάθησης του Perceptron (Αισθητήρα).....	36
Ο Αλγόριθμος Ελάχιστου Μέσου Τετραγωνικού (EMT) Σφάλματος.....	39
Ο Αλγόριθμος Πίσω Διάδοσης (Π.Δ.) του Λάθους.....	40
3.2. Ασαφής Λογική (Fuzzy Logic)	43
Συναρτήσεις συμμετοχής.....	44
Ιδιότητες των ασαφών συνόλων.....	47
Λογικές πράξεις στα ασαφή σύνολα.....	49
Ασαφείς σχέσεις.....	53
Ασαφής συλλογιστική.....	54
3.3. Συστήματα Ασαφούς Συλλογιστικής-Νευροασαφή Συστήματα.....	56
Κανόνες αν-τότε (if-then rules)	56
Συστήματα ασαφούς συλλογιστικής (fuzzy inference systems)	59
Νευροασαφή συστήματα.....	67
3.4. ANFIS (Adaptive Neuro-Fuzzy Inference System)	72
Εισαγωγή - Ορισμός.....	72
Δομή Αρχιτεκτονικής του ANFIS.....	73
Περιγραφή του μοντέλου.....	76
Αριθμητικό Παράδειγμα Μοντέλου.....	79
4. ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ (case study).....	88
4.1. Εισαγωγή	88
4.2. Δεδομένα.....	88
4.3. Επεξεργασία.....	96
4.4. Αποτελέσματα.....	100
5. ΣΥΜΠΕΡΑΣΜΑΤΑ.....	107
6. ΒΙΒΛΙΟΓΡΑΦΙΑ.....	108

1. ΕΙΣΑΓΩΓΗ

Αντικειμενικός σκοπός μίας επιχείρησης είναι το κέρδος και πιο συγκεκριμένα η μεγιστοποίησή του. Για το σκοπό αυτό λοιπόν οι διευθυντές παραγωγής των εργοστασίων πρέπει να κανονίζουν την παραγωγή σύμφωνα με τη ζήτηση του εκάστοτε προϊόντος. Σε περίπτωση που δεν τη λάβουν υπόψη τους, αυτόματα αυτό θα σημαίνει είτε μείωση κερδών για την επιχείρηση είτε ακόμη και ζημία. Για τους λόγους αυτούς έχουν αναπτυχθεί διάφορες τεχνικές με στόχο την πρόβλεψη των πωλήσεων. Όσο πιο αποτελεσματική είναι η πρόβλεψη αυτή, τόσο πιο επιτυχημένη είναι η επιχειρηματική δραστηριότητα. Οι περισσότερες διαδικασίες στον επιχειρηματικό τομέα είναι δυναμικές και μη-γραμμικές, γεγονός που σημαίνει πως δεν μπορούν να προσεγγιστούν με συμβατικές τεχνικές.

Στην παρούσα διπλωματική εργασία επιχειρούμε να χρησιμοποιήσουμε μία εναλλακτική τεχνική η οποία βασίζεται σε προγραμματισμό με το λογισμικό της MATLAB και συγκεκριμένα στην πρόβλεψη των μηνιαίων πωλήσεων πέντε (5) διαφορετικών προϊόντων χρησιμοποιώντας τις πωλήσεις προηγούμενων μηνών και το ANFIS, το οποίο αποτελεί χαρακτηριστικό δείγμα των νευρωνικών δικτύων και της ασαφούς λογικής. Αρχικά γίνεται μία αναφορά στη βιβλιογραφία ώστε να γνωρίζουμε την πρόοδο που έχει επιτευχθεί μέχρι και σήμερα. Έπειτα καταγράφεται η θεωρητική βάση που είναι απαραίτητη για την κατανόηση του αντικειμένου και αναλύεται η περίπτωση μελέτης μας. Τέλος, εξάγονται συμπεράσματα και παρατίθεται η βιβλιογραφία και ο κώδικας που χρησιμοποιήθηκε.

1. ABSTRACT

The objective of business is to maximize the benefit. In order to achieve that goal, managers of production need to arrange the product quantity according to demand. If they product demand is not estimated, automatically that causes reducing profits or even loss. For these reasons, techniques have been developed to forecast sales. The more effective the prediction is, the more succeeded enterprising efficiency takes place. Most of the procedures are dynamic or non-linear, so they can't be approximated with conventional methods.

This research attempts to use an alternative technique which is based in programming using MATLAB software and specifically in order to forecast monthly sales with ANFIS. Firstly, bibliography reference is given so that progress which has been done until today in forecasting sales is presented. Secondly, the theoretical basis is recorded in order to comprehend prediction methods and our case study is analyzed. Finally, conclusions are been exported and matlab code is shown.

2.ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

Είναι σημαντικό να ερευνήσουμε τις προσπάθειες που έχουν γίνει μέχρι τώρα στην πρόβλεψη πωλήσεων, τα μοντέλα που έχουν χρησιμοποιηθεί στο πέρασμα του χρόνου, η αποτελεσματικότητα τους καθώς και τα συμπεράσματα που εξήχθησαν. Για τον λόγο αυτό στη συνέχεια παρατίθενται επιστημονικά άρθρα που διαπραγματεύονται τα παραπάνω.

Forecasting Sales Using Neural Networks

Στο άρθρο αυτό των Frank M. Thiesing και Oliver Vornberger (*Department of Mathematics and Computer Science, University of Osnabriick, Germany*), εκπαιδεύεται ένα νευρωνικό δίκτυο με τον αλγόριθμο οπισθοδιάδοσης και εφαρμόζεται για να προβλέψει τις μελλοντικές τιμές της χρονοσειράς που αποτελούν την εβδομαδιαία ζήτηση προϊόντων σε ένα supermarket. Οι επηρεασμένοι δείκτες τιμών, διαφημιστικές καμπάνιες και η περίοδος διακοπών έχουν ληφθεί υπόψη. Ο σχεδιασμός και η εκτέλεση ενός συστήματος πρόβλεψης με χρήση νευρωνικών δικτύων, έχει εγκατασταθεί ως πρότυπο σε Γερμανική εταιρεία supermarket ώστε να βοηθήσει στη διαδικασία του management και να καθορίσει τα αριθμητικά στοιχεία των αναμενόμενων πωλήσεων. Η επίδοση του δικτύου εκτιμάται συγκρίνοντας το με δύο τεχνικές που χρησιμοποιούνται τη δεδομένη στιγμή στο supermarket. Η σύγκριση δείχνει ότι τα νευρωνικά δίκτυα υπερτερούν από τις συμβατικές μεθόδους όσον αφορά την ποιότητα της πρόβλεψης.

Sales Forecasting for TFT-LCD Products with a Hybrid Neural Network Model

Στη βιομηχανία κατασκευής TFT-LCD οθονών, υπάρχουν συγκεκριμένα χαρακτηριστικά όπως απροσδόκητη διακύμανση της ζήτησης, εξατομικευμένα προϊόντα στα οποία κάθε καταναλωτής θα προσδιορίσει ένα εξάρτημα κλειδί και μικρός κύκλος ζωής προϊόντος. Η έρευνα που έγινε από τους Pei-Chann Chang (*Department of Information Management*), Chin-Yuan Fan (*Department of Industrial Engineering and Management*), Cyser Chen (*Department of Industrial Engineering and Management*), Baw-Jhiune Liu (*Department of Computer Science and Engineering*), παρουσιάζει το μοντέλο νευρωνικού δικτύου για μηνιαία πρόβλεψη πωλήσεων οθονών TFT-LCD στην Ταϊβάν. Οι πρακτική κατάσταση και οι απαιτήσεις της εξηγούνται στο παρόν άρθρο και διαπραγματεύονται δύο προσεγγίσεις: α) ένα νευρωνικό δίκτυο επίβλεψης μάθησης της λειτουργικής συμπεριφοράς των δεδομένων χρονοσειράς και μία χοντρική εκτίμηση της και β) ένας γενετικός αλγόριθμος για τη μοντελοποίηση της συμπεριφοράς του συστήματος με τη πιθανότητα άντλησης πληροφοριών από ειδικούς και συστηματική βελτιστοποίηση. Τα πλεονεκτήματα και τα μειονεκτήματα αυτών των προσεγγίσεων παρουσιάζονται διεξοδικά στο άρθρο αυτό. Και τα δύο αυτά συστήματα έχουν δοκιμαστεί με αποτελέσματα πολύ ικανοποιητικά, τα οποία παρουσιάζονται με πρακτικά δεδομένα.

Forecasting Financial Time Series using Neural Network and Fuzzy System-based Techniques

Η πρόβλεψη συναλλαγματικών ισοτιμιών είναι ένα πάρα πολύ σημαντικό πρόβλημα το οποίο επιδέχεται αυξανόμενης προσοχής, ιδιαιτέρως λόγω της ουσιαστικής δυσκολίας και της πρακτικής εφαρμογής. Κατά τη διάρκεια των τελευταίων ετών έχουν προταθεί πολλά μη γραμμικά μοντέλα για να εξασφαλίσουμε ακριβή αποτελέσματα πρόβλεψης, σε μια προσπάθεια να βελτιώσουμε τις επιδόσεις των κλασικών γραμμικών προσεγγίσεων. Ανάμεσα σε αυτές, έχουν χρησιμοποιηθεί νευρωνικά δίκτυα με πολύ ενθαρρυντικά αποτελέσματα. Στην εργασία αυτή από τους V. Kodogiannis (*Department of Computer Science, University of Westminster, London*) και A. Lolis (*Ford Motor Company, R&D Center, Dunton, Essex, UK*), παρουσιάζονται βελτιωμένα νευρωνικά δίκτυα και ασαφή μοντέλα που χρησιμοποιούνται στην πρόβλεψη συναλλαγματικών ισοτιμιών. Ποικίλες προσεγγίσεις συμπεριλαμβανομένων πολύ-στρωματικών αισθητήρων, συναρτήσεις ακτινικής βάσης, δυναμικά νευρωνικά δίκτυα και νευρο-ασαφή συστήματα, πραγματεύονται και προτείνονται στη μελέτη αυτή. Επίσης έχουν εκτιμηθεί οι επιδόσεις των προβλέψεων ενός βήματος και πολλών πρόσω βημάτων, χρησιμοποιώντας πραγματικά καθημερινά δεδομένα της ισοτιμίας του Αμερικάνικου δολαρίου έναντι της Βρετανικής λίρας.

Sales forecasts for existing consumer products and services: Do purchase intentions contribute to accuracy?

Είναι σύνηθες να χρησιμοποιείται η πρόθεση αγοράς στην πρόβλεψη πωλήσεων υφιστάμενων προϊόντων και υπηρεσιών. Ενώ παλαιότερες μελέτες έχουν δείξει ότι η πρόθεση είναι προφητική για τις πωλήσεις, αυτές έχουν μόνο εξετάσει την απόλυτη ακρίβεια των προθέσεων και όχι τη συσχέτιση της ακρίβειας σε σχέση με άλλες μεθόδους πρόβλεψης. Για παράδειγμα, καμία έρευνα δεν έχει επιδείξει ότι μία πρόβλεψη βασισμένη στην αγοραστική πρόβλεψη είναι αποτελεσματικότερη από μία απλή πρόβλεψη που στηρίζεται σε δεδομένα παλαιότερων πωλήσεων. Οι J. Scott Armstrong (*The Wharton School, University of Pennsylvania*), Vicki G. Morwitz (*Stern School of Business, New York University*) και V. Kumar (*Marketing Department, University of Houston*), εξέτασαν τη συγκριτική ακρίβεια τεσσάρων μεθόδων που προβλέπουν πωλήσεις βασισμένες σε αγοραστικές τάσεις. Δοκίμασαν τις μεθόδους χρησιμοποιώντας τέσσερα σετ δεδομένων συμπεριλαμβανομένων διαφορετικών προϊόντων και διαφορετικού χρονικού ορίζοντα. Για κάθε μια από τις τέσσερις μεθόδους βρέθηκε ότι είναι πιο ακριβής από μια απλή που χρησιμοποιεί μόνο δεδομένα προηγούμενων πωλήσεων. Σε ακόμα όμως καλύτερα αποτελέσματα οδηγούμαστε όταν χρησιμοποιούμε συνδυασμό αυτών των μεθόδων εξισώνοντας τα βάρη. Επετεύχθη μείωση του λάθους μέχρι και στο 1/3 του αρχικού. Επομένως γίνεται κατανοητό ότι η πρόθεση αγοράς είναι ένα πολύτιμο εργαλείο που βελτιώνει την αποτελεσματικότητα των προβλέψεων όταν χρησιμοποιείται στην είσοδο των δεδομένων μας.

A Hybrid Sales Forecasting Method Based on Stable Seasonal Pattern Models and BPNN

Από τη στιγμή που πολλές δραστηριότητες των εταιρειών εξαρτώνται από την πρόβλεψη πωλήσεων είναι πολύ σημαντικό για αυτές να προκαθορίζουν τα δεδομένα μελλοντικών πωλήσεων με ακρίβεια και αξιοπιστία. Οι περισσότερες επιχειρήσεις λιανικής πώλησης είναι εποχιακές και για αυτό το λόγο χρησιμοποιούν για την πρόβλεψη εποχιακά δεδομένα ώστε να επιτύχουν καλύτερα αποτελέσματα. Στο άρθρο αυτό, οι Xiao Jiwei, Wu Yaohua, Wang Qian, Liao Li and Hu Hongchun (*The Logistics Institute Shandong University*), προτείνουν ένα υβριδικό μοντέλο πρόβλεψης που συνδυάζει σταθερό εποχιακό πρότυπο μοντέλο και νευρωνικό δίκτυο οπισθο-διάδοσης BPNN (back propagation neural network) για να προβλέψουν πωλήσεις λιανικής. Τα πρακτικά αποτελέσματα αποδεικνύουν ότι το μοντέλο αυτό είναι πιο ακριβές από ένα απλό μοντέλο πρόβλεψης.

Developing and using forecasting models of consumer durables

Για μία επιχείρηση είναι μια από τις πιο κρίσιμες δραστηριότητες η πρόβλεψη των μελλοντικών πωλήσεων γιατί με το τρόπο αυτό μπορεί να σχεδιάζει και να ελέγχει πολλές δραστηριότητες. Η πρόβλεψη πωλήσεων διαρκών καταναλωτικών αγαθών είναι ένα ιδιαιτέρως δύσκολο πεδίο και πολύ ανταγωνιστικό αφού η αγορά με τα προϊόντα συνεχώς μεταβάλλεται. Οι Barry L. Bayus, Saman Hong και Russell P. Labe, παρουσιάζουν το σχέδιο ενός μοντέλου πρόβλεψης καθοδηγούμενο από την αγορά και διαπραγματεύονται την πρόβλεψη πωλήσεων έγχρωμων τηλεοράσεων RCA. Μία σημαντική σημείωση για την εφαρμογή αυτή είναι ότι η απλή προσέγγιση του μοντέλου ήταν ανεπαρκής. Οικονομετρικά μοντέλα και μοντέλα εξομοίωσης τα οποία αναπτύχθηκαν περιγράφονται, όπως και οι επιδόσεις πρόβλεψης και η χρήση τους.

A model of export sales forecasting behavior and performance: development and testing

Το άρθρο αυτό από τους Heidi Winklhofer (*University of Nottingham, UK*) και Αδαμάντιο Διαμαντόπουλο (*Marketing and Business Research , Loughborough University Business School, Leicestershire*), παρουσιάζει και δοκιμάζει ένα μοντέλο συμπεριφοράς και επιδόσεων πρόβλεψης πωλήσεων εξαγωγών, ενσωματώνοντας οργανωτικά χαρακτηριστικά και εξειδικευμένα στοιχεία εξαγωγών. Το μοντέλο δοκιμάζεται σε ένα δείγμα Βρετανών εξαγωγέων και δείχνει να έχει καλή αποτελεσματικότητα. Τα εξαγόμενα αποτελέσματα μας δείχνουν ότι το κλειδί στη πρόβλεψη εξαγωγών είναι η οργανωτική ανάληψη υποχρεώσεων μαζί με τους πόρους που διατίθενται στη συνάρτηση πρόβλεψης. Ακόμα, αναλύονται τα συμπεράσματα από τα ευρήματα του μοντέλου και αναγνωρίζονται μελλοντικές οδοί έρευνας.

Three complementary sales forecasting models for textile distributors

Με σκοπό να μειώσουν το απόθεμα τους, να αποφύγουν πιθανές ελλείψεις και να κατευθύνουν την πολιτική τους πάνω στο marketing, οι κλωστοϋφαντουργικές εταιρείες πρέπει να βελτιώσουν την εφοδιαστική τους αλυσίδα. Ο κλάδος αυτός απαιτεί συστήματα πρόβλεψης προσαρμοσμένα στο αβέβαιο περιβάλλον της κλωστοϋφαντουργίας. Συμπεριλαμβάνοντας και τους περιορισμούς παράδοσης, οι κλωστοϋφαντουργικές εταιρείες χρειάζονται πρόβλεψη βάθους μίας περιόδου για να ξεκινήσουν την παραγωγή και πρόβλεψη εβδομαδιαία για να καθορίσει το απόθεμα στις αποθήκες. Το σύστημα που προτείνεται στο παρόν άρθρο από τους Sebastien Thomassey, Michel Happiette και Marie Castelain, κάνει πρόβλεψη πωλήσεων μέσου και βραχυπρόθεσμου χρονικού ορίζοντα. Από τις αθροιστικές αυτές προβλέψεις το σύστημα εξαγεί μία εκτίμηση πωλήσεων. Τέλος, για να υπολογιστεί η ακρίβεια του μοντέλου, γίνεται προσομοίωση με αληθινά δεδομένα από μία σημαντική εταιρεία πώλησης έτοιμων ενδυμάτων.

A Comparison of extrapolation models on yearly sales forecasts

Η μελέτη αυτή του Steven P. Schnaars (*Baruch College of the City University of New York*) συγκρίνει τις επιδόσεις επτά επεκτεινόμενων μοντέλων και τριών συνδυασμών αυτών σε περίπου 1500 ετήσιες προβλέψεις πωλήσεων μονάδων και έχει επικεντρωθεί κυρίως σε δύο θέματα: α) στην εξακρίβωση των περιπτώσεων που τα μοντέλα αυτά έχουν μεγάλες πιθανότητες να επιτύχουν και β) ποια προσέγγιση προσφέρει καλύτερη δυναμική. Στο άρθρο αυτό τέλος επαναλαμβάνονται πολλά από τα ευρήματα του M-competition και προστίθενται μερικά ακόμα στοιχεία.

A fuzzy case-based reasoning model for sales forecasting in print circuit board industries

Είναι πλέον αποδεδειγμένο ότι η αξιόπιστη πρόβλεψη πωλήσεων, βελτιώνει την ποιότητα της στρατηγικής που ακολουθεί μια επιχείρηση. Η τεχνική CBR (case-based reasoning), μία από τις πλέον γνωστές τεχνικές τεχνητής νοημοσύνης (AI), έχει ήδη αποδείξει την αποτελεσματικότητά της σε πολυάριθμες μελέτες. Σε αυτή τη μελέτη από τους Pei-Chann Chang (*Department of Information Management, Yuan Ze University*), Chen-Hao Liu (*Department of Industrial Engineering and Management, Yuan Ze University*) και Robert K. Lai (*Department of Computer Science and Engineering, Yuan Ze University*), αναπτύσσεται ένα ασαφές CBR (FCBR-fuzzy case-based reasoning) και ερευνάται ενδεχόμενη χρήση στην πρόβλεψη μελλοντικών πωλήσεων σε εργοστάσια τυποποιημένων κυκλωμάτων. Αριθμητικά δεδομένα από διάφορους παράγοντες και πραγματική ζήτηση των τελευταίων 5 ετών από την πώληση των κυκλωμάτων συλλέχτηκαν και χρησιμοποιήθηκαν σαν είσοδος στο FCBR για να πάρουμε σαν έξοδο μηνιαίες μελλοντικές πωλήσεις. Τα πειραματικά αποτελέσματα δείχνουν την αποτελεσματικότητα του FCBR σε σχέση με άλλες προσεγγίσεις.

A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm

Η πρόβλεψη πωλήσεων είναι περίπλοκο ζήτημα και επηρεάζεται από εσωτερικούς και εξωτερικούς παράγοντες και παίζει πολύ σημαντικό ρόλο στη στρατηγική που ακολουθούν οι επιχειρήσεις. Πρόσφατα, τεχνητά νευρωνικά δίκτυα (ANNs) εφαρμόζονται στη πρόβλεψη λόγω των πολλά υποσχόμενων δυνατοτήτων τους στον τομέα ελέγχου και αναγνώρισης προτύπου. Περαιτέρω βελτίωση πάραυτα είναι αναγκαία από τη στιγμή που ιδιαίτερες περιστάσεις όπως η προώθηση προϊόντων προκαλούν αλλαγές στο μοτίβο των πωλήσεων. Εν προκειμένω λοιπόν, ο R.J. Kuo (*National Taipei University of Technology, Taipei*) χρησιμοποιεί ασαφή νευρωνικά δίκτυα με αρχικά βάρη τα οποία προκύπτουν από γενετικούς αλγορίθμους (GFNN) σε μια προσπάθεια εκπαίδευσης του κανόνα IF-THEN υπό την επίβλεψη ειδικών στην προώθηση προϊόντων. Τα αποτελέσματα του (GFNN) είναι πιο ολοκληρωμένα σε σχέση με τη πρόβλεψη μέσω ANN (Artificial Neural Networks). Αυτό ακριβώς υποδεικνύουν και όλα τα μοντέλα αξιολόγησης του προτεινόμενου συστήματος, ότι δηλαδή είναι πιο αποτελεσματικό από οποιοδήποτε απλό ANN.

Development of a fuzzy sales forecasting system for vending machines

Οι αυτόματες μηχανές πώλησης λειτουργούν 24 ώρες την ημέρα, επιτρέποντας στους καταναλωτές να αγοράζουν το προϊόν οποιαδήποτε ώρα της ημέρας ή νύχτας επιθυμούν για επιπλέον εξυπηρέτηση. Όμως καταναλώνουν μεγάλα ποσά ενέργειας, που ισούνται με τη συνολική παραγωγή δύο πυρηνικών αντιδραστήρων στην Ιαπωνία. Καθώς το περιβάλλον παγκοσμίως καταστρέφεται, η ανάγκη μείωσης κατανάλωσης ενέργειας αυξάνεται όλο και πιο πολύ. Επομένως είναι άμεση ανάγκη μείωσης της κατανάλωσης των μηχανημάτων αυτών. Στη παρούσα μελέτη γίνεται προσπάθεια από τους Hidetaka Sakai, Hideki Nakajima, Minoru Higashihara, Masashi Yasuda, Masato Oosumi (*SANYO Electric Co. Ltd, Mechatronics Research Center*), να αναπτύξουν ένα σύστημα το οποίο θα μειώσει την κατανάλωση που απαιτείται για την ψύξη. Γίνεται λοιπόν πρόβλεψη για την ημερήσια πώληση αναψυκτικών έτσι ώστε θα χρησιμοποιείται ηλεκτρισμός για να ψύξει μόνο αυτά τα κουτιά αναψυκτικών. Επίσης χρησιμοποιήθηκε ασαφής λογική και ένα πολλαπλό μοντέλο οπισθοδρόμησης για να διορθώσει προγενέστερες προβλέψεις, ώστε να μεγαλώσει η ακρίβεια της πρόβλεψης. Τέλος, έγινε προσομοίωση του μοντέλου και επιβεβαιώθηκε ότι η μείωση της κατανάλωσης ενέργειας, ανέρχεται περίπου στο 1/10.

Forecasting restaurant sales using self-selectivity models

Ένα ανεξάρτητο εστιατόριο συχνά ανοίγει και κλείνει διαφορετικές ώρες της βδομάδας ή του έτους. Αυτό επιφέρει προκατάληψη διότι οι πωλήσεις παρατηρούνται μόνο όταν το εστιατόριο είναι ανοιχτό και με τις μικρότερες προσδοκίες πωλήσεων. Επιπλέον ο χρήστης του μοντέλου θα έχει δυσκολία στο να εκτιμήσει ποια πολιτική είναι η πιο κατάλληλη. Οι Michael S Morgan (*Savitz Research Center, Dallas, USA*)

και Pradeep K Chintagunta (*Graduate School of Business, University of Chicago*) εξέτασαν τρία εναλλακτικά μοντέλα. Βρέθηκε ότι η καλύτερη πρόβλεψη επιτυγχάνεται με ένα περιορισμένο μοντέλο οπισθοδρόμησης. Αυτό επιβεβαίωσε τη χρησιμότητα του ως εργαλείο πρόβλεψης και ότι ο χρήστης ακολουθεί τη σωστή πολιτική ανοίγματος-κλεισίματος.

Product Demand Forecasting with a Novel Fuzzy CMAC

Η πρόβλεψη ζήτησης προϊόντων ήταν ανέκαθεν μεγάλη πρόκληση για τους managers αφού διαδραματίζει σημαντικό ρόλο στη λήψη αποφάσεων και στο σχεδιασμό των αποθεμάτων. Οι αποφάσεις αυτές είναι κρίσιμες για την επιβίωση της επιχείρησης και συντελούν στην εύρεση απαιτήσεων των πελατών. Στην εργασία αυτή, οι D. Shi, C. Quek, R. Tilani (*School of Computer Engineering, Nanyang Technological University, Singapore*) και J. Fu (*School of Electric and Information Engineering, Heilongjiang Institute of Science and Technology, Harbin, China*), παρουσιάζουν ένα νέο ασαφές μοντέλο με ελεγκτή παρεγκεφαλικής διάρθρωσης (FCMAC - Cerebellar-Model-Articulation-Controller) με περιορισμό πραγματικής τιμής (TVR - Truth Value Restriction). Γίνεται σύγκριση και ερευνάται η επίδοση του μοντέλου αυτού με ήδη υπάρχουσες τεχνικές όπως της Μονής Εκθετικής Εξομάλυνσης (Single Exponential Smoothing), η Γραμμική Ροπή του Holt (Holt's Linear Trend), Πρόσθετης Μεθοδολογίας των Holt και Winter (Holt-Winter's Additive methods), Μοντέλου ARIMA των Box και Jenkins (Box-Jenkin's ARIMA model), Δίκτυα με Συναρτήσεις Ακτινικής Βάσης (radial basis function networks) και Πολύ-στρωματικών Αισθητήρων (multi-layer perceptrons). Τα πειραματικά δεδομένα προήλθαν από το US Census Bureau και τον M3 Competition. Τα αποτελέσματα αποκαλύπτουν ότι το παραπάνω μοντέλο (FCMAC) αποφέρει μικρότερα λάθη για τα συγκεκριμένα σετ δεδομένων. Τέλος, παρατίθενται οι συνθήκες κάτω από τις οποίες το μοντέλο FCMAC επέδειξε ανωτερότητα σε σχέση με τα προϋπάρχοντα μοντέλα.

Sales Forecast Errors for New Product Projects

Είναι εξαιρετικά δύσκολο να αποφευχθούν λάθη στις προβλέψεις πωλήσεων σε καινούρια βιομηχανικά προϊόντα, ειδικά όταν οι προβλέψεις αυτές γίνονται στο αρχικό στάδιο κατασκευής του προϊόντος. Όμως εάν οι managers είναι ενήμεροι ότι η πρόβλεψη έχει μεγάλη πιθανότητα να περιέχει σημαντικά λάθη, μπορούν να ακολουθήσουν κάποια βήματα να αντιμετωπίσουν αυτή την αβεβαιότητα. Οι Blair Little and Roger A. More έκαναν μία μελέτη αναλύοντας 185 νέα προγράμματα ανάπτυξης προϊόντων για να καταλήξουν σε τέσσερις μείζονος σημασίας δείκτες λαθών πρόβλεψης. Οι δείκτες αυτοί μπορούν να εξαχθούν από τυπικές διαδικασίες για την προφύλαξη του σχεδίου ανάπτυξης και μπορούν να χρησιμοποιηθούν για να προειδοποιήσουν τους managers των προϊόντων για συγκεκριμένα προβλήματα που πιθανότατα λαμβάνουν χώρα.

A decision support system for sales forecasting through fuzzy neural networks with asymmetric fuzzy weights.

Η πρόβλεψη πωλήσεων παίζει πολύ σημαντικό ρόλο στη στρατηγική που ακολουθούν οι επιχειρήσεις. Πολυάριθμες προσεγγίσεις του προβλήματος έχουν γίνει χρησιμοποιώντας στατιστικές μεθόδους αυτής της οπισθοδρόμησης (regression) και του κινούμενου μέσου όρου (ARMA). Παρόλα αυτά η πρόβλεψη πωλήσεων είναι περίπλοκο ζήτημα και επηρεάζεται από εσωτερικούς και εξωτερικούς παράγοντες. Πρόσφατα, τεχνητά νευρωνικά δίκτυα (ANNs) εφαρμόζονται στη πρόβλεψη λόγω των πολλά υποσχόμενων δυνατοτήτων τους στον τομέα ελέγχου και αναγνώρισης προτύπου. Περαιτέρω βελτίωση είναι αναγκαία από τη στιγμή που ιδιαίτερες περιστάσεις όπως η προώθηση προϊόντων προκαλούν αλλαγές στο μοτίβο των πωλήσεων. Εν προκειμένω λοιπόν, οι R.J. Kuo και K.C. Xue του τμήματος μηχανικών βιομηχανίας (*National Taipei University of Technology, Taipei*) και επιστήμης διοίκησης αντίστοιχα (*I-Shou University, Kaohsiung County*) χρησιμοποιούν ασαφή λογική και προτείνουν το δίκτυο FNN (Fuzzy Neural Networks) που ουσιαστικά συνιστούν την ασαφοποίηση του συμβατικού μοντέλου νευρωνικού δικτύου σε μια προσπάθεια εκπαίδευσης του κανόνα IF-THEN υπό την επίβλεψη ειδικών στην προώθηση προϊόντων. Τα αποτελέσματα του FNN είναι πιο ολοκληρωμένα σε σχέση με τη πρόβλεψη μέσω ANN (Artificial Neural Networks). Αυτό ακριβώς υποδεικνύουν και όλα τα μοντέλα αξιολόγησης του προτεινόμενου συστήματος, ότι δηλαδή είναι πιο αποτελεσματικό από οποιοδήποτε απλό ANN.

Effectiveness of sales forecasting methods

Η αποτελεσματικότητα της πρόβλεψης πωλήσεων έχει γίνει η βάση για ένα επιτυχημένο management. Δυστυχώς όμως, παρά το γεγονός ότι έχει αναγνωριστεί αυτή η αναγκαιότητα, δεν έχουν επιτευχθεί καλύτερα αποτελέσματα. Φαίνεται ότι τα λάθη στην πρόβλεψη πωλήσεων αυξάνονται αντί να μειώνονται. Ο James T. Rothe, στο άρθρο του αυτό, εκθέτει τα ευρήματα του στη μελέτη συστημάτων πρόβλεψης σε κατασκευαστικές εταιρείες. Η σπουδαιότητα πρόβλεψης για τις εταιρείες, χρήστες και προπαρασκευαστές της πρόβλεψης, τεχνικές πρόβλεψης, χρήση δεδομένων μακροοικονομίας, λάθη και αποκλίσεις στην πρόβλεψη και διαχείριση συστημάτων πρόβλεψης, είναι οι τομείς που καλύπτει το παρόν άρθρο. Τα συμπεράσματα και οι προτάσεις για βελτίωση των συστημάτων πρόβλεψης προκύπτουν από τα ευρήματα της μελέτης.

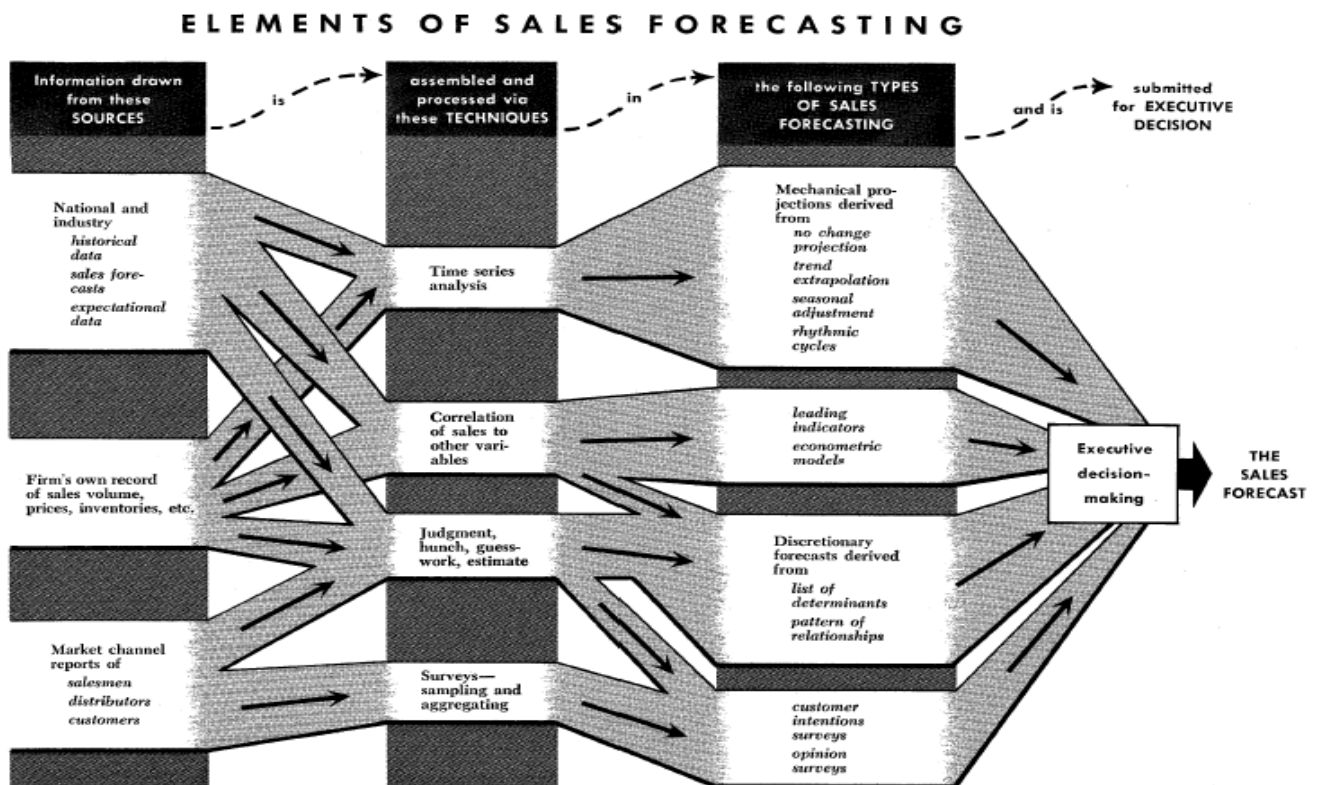
Sales Prediction for Franchise Stores Using Artificial Neural Networks

Η εργασία αυτή του Anubhav Tewari (Spectrum Analysis Australia) έχει σαν σκοπό να περιγράψει τη ωφέλεια των τεχνητών νευρωνικών δικτύων ως χρήσιμα εργαλεία πρόβλεψης πωλήσεων σε καταστήματα franchise. Η μοντελοποίηση πρόβλεψης πωλήσεων συνήθως εμπεριέχει τους λόγους που οδηγούν σε πωλήσεις, τη σημασία των επιρροών αυτών στις επιδόσεις των πωλήσεων και τέλος αναπτύσσεται το μοντέλο με τη μεγαλύτερη δυνατή ικανότητα πρόβλεψης, που ενσωματώνει αυτές τις πληροφορίες. Επίσης οι πληροφορίες αυτές βοηθούν τους αποφασίζοντες να

κατανοήσουν τους κύριους λόγους των πωλήσεων τους, να κάνουν ανασκόπηση της επίδοσης των καταστημάτων τους και να εδραιώσουν νέα προσοδοφόρα καταστήματα. Παραδοσιακά, στατιστικά μοντέλα που συνήθως χρησιμοποιούν φόρμα οπισθοδρόμησης, έχουν βοηθήσει σε τέτοιου είδους προβλέψεις. Τα τεχνητά νευρωνικά δίκτυα είναι μη-παραμετρικοί στατιστικοί εκτιμητές και χρησιμοποιούνται σε περιπτώσεις όπου τα δεδομένα είναι λίγα ή όταν δεν υπάρχει διαθέσιμη εκ των προτέρων γνώση. Η μελέτη αυτή χρησιμοποιεί έναν αλγόριθμο πολύ-στρωματικού αισθητήρα με οπισθοδιάδοση. Η επιλογή των μεταβλητών και η ανάλυση των βαρών των τελικών σετ, υπολογίζονται μέσω της μεθόδου της ανάλυσης ευαισθησίας. Το τελικό αποτέλεσμα είναι μία απλή μέθοδος με 10 οδηγούς και υψηλή δυνατότητα πρόβλεψης.

Fitting the Sales Forecast to Your Firm

Αν και η προετοιμασία της πρόβλεψης πωλήσεων είναι αρκετά δύσκολη, είναι διαδικασία που αποτελεί μονόδρομο. Δυστυχώς, καμία πηγή πληροφόρησης, τεχνική, ή είδος πρόβλεψης δεν είναι κατάλληλο για όλες τις φίρμες. Οι επιχειρηματίες σύμφωνα με τον James B. Boulden, θα πρέπει να αναλύουν από μόνοι τους τα προϊόντα και την αγορά, ώστε να καθορίσουν ποια μεθοδολογία είναι πιο συμβατή σε αυτούς. Για να κινηθούν έξυπνα πρέπει να γνωρίζουν καλά τις διάφορες εναλλακτικές που είναι διαθέσιμες, αντικείμενο που συζητείται εκτενώς στο παρόν άρθρο. Στο σχεδιάγραμμα που ακολουθεί παρουσιάζεται η διαδικασία της πρόβλεψης πωλήσεων: εν ολίγοις, έχουν σχεδιαστεί πληροφορίες από διάφορες πηγές και έχουν επεξεργαστεί από μία ή παραπάνω τεχνικές πρόβλεψης. Κάθε μία από τις τεχνικές αυτές μπορεί να χρησιμοποιηθεί με ποικίλους τρόπους, κάτι που εξαρτάται από τον τύπο της πρόβλεψης. Οι τύποι αυτοί δεν είναι αμοιβαία αποκλειόμενοι αλλά μπορούν να συνδυαστούν οπότε η φίρμα μπορεί αντλήσει πληροφορίες από πολλές πηγές, να χρησιμοποιήσει παραπάνω από μία τεχνική και να τις εφαρμόσει ποικιλοτρόπως. Όλες οι προβλέψεις πωλήσεων σταματάνε στο στάδιο της εκτέλεσης της απόφασης. Το υπόλοιπο της εργασίας αφιερώνεται στον καθορισμό και αποσαφήνιση της πολύπλοκης αυτής διαδικασίας.



Neuro-fuzzy approach to forecast returns of scrapped products to recycling and remanufacturing

Η πρόβλεψη ελαττωματικών προϊόντων που καταλήγουν στην ανακύκλωση, λύνει πολλά προβλήματα στις εταιρείες ανακατασκευής και ανακύκλωσης εξαιτίας της αβεβαιότητας του χρόνου και της ποσότητας που θα επιστραφούν. Στο άρθρο αυτό των Marx-Gomez J., Rautenstrauch C., Nurnberger A. και Kruse R. Γίνεται μια εκτεταμένη έρευνα ως προς τη μεθοδολογία πρόβλεψης που επιστρέφει σαν αποτέλεσμα τιμές (ποσότητας και χρόνου) ελαττωματικών προϊόντων για ανακύκλωση. Το προτεινόμενο μοντέλο βασίζεται σε συναφείς παράγοντες που τα επηρεάζουν και στον κύκλο ζωής των προϊόντων και έχει εφαρμοστεί σε ένα case study για αξιολόγηση και εξέταση χρησιμότητας. Η προσέγγιση χρησιμοποιεί μία μελέτη προσομοίωσης, το σχεδιασμό ενός ασαφούς συμπερασματικού συστήματος για τη πρόβλεψη της επιστροφής σε μία προκαθορισμένη περίοδο και το σχεδιασμό ενός νευρο-ασαφούς συστήματος που επιστρέφει τιμές που αφορούν το χρόνο.

Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods

Όπως πολλές οικονομικές χρονοσειρές, έτσι και οι αθροιστικές πωλήσεις προϊόντων λιανικής των Ηνωμένων Πολιτειών έχουν ισχυρές τάσεις και εποχιακά πρότυπα. Η μοντελοποίηση και η πρόβλεψη αυτών των προτύπων είναι ένα μακροχρόνιο θέμα στην ανάλυση χρονοσειρών. Το άρθρο αυτό των Ilan Alon (*Department of Business and Economics, State University of New York*), Min Qi (*College of Business Administration, Kent State University*) και Robert J. Sadowski (*Market Analysis, Greater Cleveland Growth Association*), συγκρίνει τεχνητά νευρωνικά δίκτυα με παραδοσιακές μεθόδους συμπεριλαμβανομένων της Εκθετικής Εξομάλυνσης Winter (Winter's Exponential Smoothing), του μοντέλου ARIMA των Box και Jenkins (Box-Jenkin's ARIMA model) και της μεταβλητής οπισθοδρόμησης (multivariate regression). Τα αποτελέσματα φανερώουν ότι κατά μέσο όρο το μοντέλο που χρησιμοποιεί ANN αποδίδει καλύτερα, ακολουθούμενο από το μοντέλο ARIMA των Box και Jenkins. Παρά την απλότητα του, το μοντέλο του Winter δείχνει να είναι βιώσιμη λύση για προβλέψεις πολλαπλών βημάτων κάτω από σχετικά σταθερές οικονομικές συνθήκες. Η περεταίρω ανάλυση δείχνει τέλος, ότι το νευρωνικό μοντέλο έχει τη δυνατότητα να εντοπίσει τις μη γραμμικές τάσεις και τα εποχιακά πρότυπα, καθώς επίσης και τις αλληλεπιδράσεις τους.

Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing

Εξαιτίας του δυνατού ανταγωνισμού που υπάρχει σήμερα, οι περισσότεροι οργανισμοί βρίσκονται σε συνεχή προσπάθεια να αυξήσουν τα πλεονεκτήματά τους και να μειώσουν το κόστος. Η ακριβής πρόβλεψη πωλήσεων είναι σίγουρα ένας πολύ οικονομικός τρόπος για την επίτευξη των προαναφερθεισών στόχων, εφόσον αυτό οδηγεί σε βελτιωμένες υπηρεσίες προς τους πελάτες, μείωση χαμένων πωλήσεων και επιστροφές προϊόντων όπως επίσης και πιο αποτελεσματικό σχεδιασμό προϊόντων. Ειδικότερα για τη βιομηχανία τροφίμων, η επιτυχημένη πρόβλεψη πωλήσεων μπορεί να είναι πολύ ευεργετική λόγω της μικρής διάρκειας ζωής των προϊόντων αλλά και της σπουδαιότητας της ποιότητας που σχετίζεται άμεσα με την ανθρώπινη υγεία. Στη παρούσα μελέτη των Philip Doganis, Alex Alexandridis, Panagiotis Patrinos και Haralambos Sarimveis (*School of Chemical Engineering, National Technical University of Athens*), παρουσιάζεται ένας ολοκληρωμένος σκελετός ο οποίος μπορεί να χρησιμοποιηθεί στην ανάπτυξη μοντέλων πρόβλεψης μη γραμμικών χρονοσειρών. Η μέθοδος είναι ένας συνδυασμός δύο τεχνολογιών τεχνητής νοημοσύνης που ονομάζεται νευρωνικό δίκτυο συνάρτησης ακτινικής βάσης (radial basis function - RBF) μαζί με έναν ειδικά σχεδιασμένο γενετικό αλγόριθμο. Η παραπάνω μεθοδολογία εφαρμόζεται επιτυχώς σε δεδομένα πωλήσεων φρέσκου γάλακτος από μεγάλη εταιρεία παραγωγής γαλακτοκομικών προϊόντων.

Sales forecasting using extreme learning machine with applications in fashion retailing

Η πρόβλεψη πωλήσεων είναι μία μεγάλη πρόκληση που οφείλεται στην ελαστικότητα της ζήτησης η οποία εξαρτάται με τη σειρά της από πολλούς παράγοντες. Στο κόσμο της μόδας αυτό η ελαστικότητα αυτή γίνεται παραπάνω αισθητή και απαιτείται ένα πολύ καλό πολύπλευρο μοντέλο πρόβλεψης των πωλήσεων. Η εργασία αυτή των Zhan-Li Sun, Tsan-Ming Choi, Kin-Fan Au και Yong Yu (*Institute of Textiles and Clothing, The Hong Kong Polytechnic University*), εφαρμόζει μία νέα τεχνική που ονομάζεται μηχανή ακραίας εκμάθησης (extreme learning machine - ELM) που ερευνά τη σχέση ανάμεσα στο ποσό των πωλήσεων και κάποιους σημαντικούς παράγοντες που επιδρούν στη ζήτηση (όπως π.χ. σχεδιαστικούς παράγοντες). Η επιδόσεις του μοντέλου υπολογίζονται χρησιμοποιώντας πραγματικά δεδομένα από βιομηχανία ένδυσης στο Hong Kong. Τα πειραματικά αποτελέσματα αποδεικνύουν ότι η μέθοδος που παρουσιάστηκε, υπερτερεί από εκείνες που βασίζονται σε νευρωνικά δίκτυα οπισθο-διάδοσης.

A neural clustering and classification system for sales forecasting of new apparel items

Οι υπεύθυνοι διανομής του δικτύου ένδυσης απαιτούν μεγάλη ακρίβεια σε προβλέψεις πωλήσεων ώστε να ελαχιστοποιήσουν το κόστος και να ικανοποιήσουν τους πελάτες τους. Για μία τέτοια στρατηγική, οι διανομείς βασίζονται πολύ στα συστήματα πρόβλεψης ώστε να ανταποκριθούν στις απαιτήσεις της πολυδιάστατης αγοράς ένδυσης. Εν τούτοις, οι περιορισμοί στις πωλήσεις ρουχισμού (πολυάριθμα νέα προϊόντα, μικρή διάρκεια ζωής), περιπλέκουν ακόμα παραπάνω τη διαδικασία πρόβλεψης και οι διανομείς προτιμούν να χρησιμοποιούν περισσότερο το ένστικτο τους, παρά τα υπάρχοντα μοντέλα πρόβλεψης. Στο άρθρο αυτό από τους Sébastien Thomassey και Michel Harpette, προτείνεται ένα βοηθητικό σύστημα απόφασης που βασίζεται σε νευρωνικά δίκτυα, το οποίο εκτελεί αυτόματα πρόβλεψη πωλήσεων προϊόντων. Οι επιδόσεις του μοντέλου υπολογίζονται χρησιμοποιώντας δεδομένα μεγάλης Γαλλικής εταιρείας ένδυσης.

Forecasting sales of slow and fast moving inventories

Παραδοσιακά, το σύστημα καταγραφής αποθέματος, βασίζεται στην εκθετική εξομάλυνση για να προβλέψει τη ζήτηση σε ταχύτατα εξελισσόμενα αποθέματα. Οι μεθοδολογίες σε βραδυκίνητα αποθέματα ποικίλουν, όμως συνήθως χρησιμοποιείται η μέθοδος Croston. Στη μελέτη αυτή του Ralph Snyder (*Department of Econometrics and Business Statistics, Monash University, Australia*), παρουσιάζονται δύο χαρακτηριστικά της εκθετικής εξομάλυνσης: α) το ότι ενσωματώνει τη διαδικασία Bernoulli με σκοπό τη σύλληψη της σποραδικής φύσης της ζήτησης και β) ότι επιτρέπει τη μέση διακύμανση να μεταβάλλεται με το χρόνο. Κριτικάρεται η προσέγγιση του Croston, γίνονται διορθώσεις στη βαθύτερη θεωρία και μετατροπές προτείνονται ώστε να ξεπεραστούν διάφορες εκτελεστικές δυσκολίες. Τέλος σκιαγραφείται μία παραμετρική προσέγγιση παράκαμψης που ολοκληρώνει την πρόβλεψη ζήτησης με έλεγχο αποθεμάτων. Η προσομοίωση εκτελείται με πραγματικά δεδομένα για εξαρτήματα αυτοκινήτων.

A comparative study of linear and nonlinear models for aggregate retail sales forecasting

Σκοπός αυτού του άρθρου από τους Ching-Wu Chu και Guoqiang Peter Zhang (*Department of Shipping and Transportation Management, National Taiwan Ocean University και Department of Management, J. Mack Robinson College of Business, Georgia State University αντίστοιχα*), είναι να συγκρίνει την ακρίβεια διαφόρων γραμμικών και μη γραμμικών μοντέλων πρόβλεψης συνολικών πωλήσεων λιανικής. Εξαιτίας των δυνατών εποχιακών επιρροών που παρατηρούνται στις πωλήσεις αυτές, έχουν εφαρμοστεί πολλά παραδοσιακά μοντέλα πρόβλεψης όπως προσέγγιση χρονοσειρών και αυτό της οπισθοδρόμησης χρησιμοποιώντας εποχιακές μεταβλητές και τριγωνομετρικές συναρτήσεις. Οι μη γραμμικές εκδοχές αυτών των προσεγγίσεων εκτελούνται μέσω νευρωνικών δικτύων τα οποία είναι γενικευμένες μορφές μη γραμμικής λειτουργικής προσέγγισης. Ερευνώνται θέματα εποχιακής προσαρμοστικότητας των μοντέλων χρονικών σειρών. Μέσα από αυτή τη μελέτη προκύπτει ότι η αποδοτικότητα των μη γραμμικών μοντέλων είναι σαφώς καλύτερη απ την αντίστοιχη γραμμικών και βελτιώνεται ακόμα και πριν την εποχιακή προσαρμογή των δεδομένων μας. Συνολικά, το καλύτερο μοντέλο είναι το νευρωνικό δίκτυο που βασίζεται στην κανονικοποίηση της εποχικότητας των δεδομένων χρονικής σειράς.

A neuro-fuzzy based forecasting approach for rush order control applications

Στο άρθρο αυτό των Wen-Pai Wang και Ze Chen (*Department of Industrial Engineering and Management, National Chin-Yi University of Technology, Taiwan*), προτείνεται ένα προσαρμοστικό νευρο-ασαφές συμπερασματικό σύστημα (ANFIS) και ένα σύστημα Kernel για να λύσει το πρόβλημα πρόβλεψης παραγγελιών αιχμής και επιπλέον να μπορέσει να κανονικοποιήσει το μηχανισμό διευθέτησης της χωρητικότητας. Η προσέγγιση αυτή γενικεύει τους κανόνες σύνδεσης μεταξύ των παραγγελιών αυτών όπως επίσης να προβλέψει ποσότητες παραγόμενων προϊόντων και πιθανώς να επιτύχει παραγγελίες μέσω εκμάθησης του μοντέλου με πραγματικά δεδομένα εταιρείας πωλήσεων ηλεκτρονικών ειδών. Ειδικότερα γίνεται σύγκριση των αποτελεσμάτων σε σχέση με παραδοσιακή ανάλυση οπισθοδρόμησης για να εξασφαλίσουμε πιο επιθυμητά αποτελέσματα. Τελικά καταλήγουμε σε υπεροχή του μοντέλου ANFIS με ορθότητα που αγγίζει το 83% σε αντίθεση με το μοντέλο οπισθοδρόμησης που φτάνει το 63%.

Sales forecasting using longitudinal data models

Η μελέτη αυτή των Edward W. Frees και Thomas W. Miller (*University of Wisconsin—Madison, School of Business, USA*), δείχνει τον τρόπο πρόβλεψης χρησιμοποιώντας μία κατηγορία μεικτών πινακοποιημένων δεδομένων. Οι προβλέψεις προκύπτουν ως ειδικές περιπτώσεις των καλύτερων γραμμικών αμερόληπτων δεικτών, γνωστοί ως BLUPs. Η πρόβλεψη BLUP αναδύεται μέσω τριών στοιχείων: 1) μία μέθοδο πρόβλεψης βασιζόμενη σε υποθετικό τρόπο απόκρισης 2) μία σύνθεση μεταβλητών χρόνου 3) μία σειριακή συσχέτιση διόρθωσης. Οι τεχνικές πρόβλεψης είναι εφαρμόσιμες σε μία ευρεία γκάμα ρυθμίσεων. Τα δεδομένα που χρησιμοποιούνται προέρχονται από την Wisconsin State Lottery και είναι διαθέσιμες 40 εβδομάδες πωλήσεων για κάθε έναν από 50 ταχυδρομικούς κωδικούς. Δεδομένου επίσης οικονομικών και δημογραφικών χαρακτηριστικών για κάθε ταχυδρομικό κωδικό, προβλέπονται οι πωλήσεις για κάθε έναν ξεχωριστά.

3. ΠΡΟΤΕΙΝΟΜΕΝΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ

3.1 Τεχνητά Νευρωνικά Δίκτυα

Ορισμός νευρωνικού δικτύου

Τα τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks), ή απλώς νευρωνικά δίκτυα (Neural Networks) είναι ένα μαθηματικό μοντέλο για την επεξεργασία πληροφορίας που προσεγγίζει την υπολογιστική και αναπαραστατική δυνατότητα μέσω συνάψεων. Το μοντέλο είναι εμπνευσμένο από τα βιοηλεκτρικά δίκτυα που δημιουργούνται στον εγκέφαλο ανάμεσα στους νευρώνες (νευρικά κύτταρα) και στις συνάψεις (σημεία επαφής των νευρικών απολήξεων).

Οι νευρώνες είναι τα δομικά στοιχεία του δικτύου. Υπάρχουν δύο είδη νευρώνων, οι νευρώνες εισόδου και οι υπολογιστικοί νευρώνες: Οι νευρώνες εισόδου δεν υπολογίζουν τίποτα, μεσολαβούν ανάμεσα στις εισόδους του δικτύου και τους υπολογιστικούς νευρώνες. Οι υπολογιστικοί νευρώνες πολλαπλασιάζουν τις εισόδους τους με τα συναπτικά βάρη και υπολογίζουν το άθροισμα του γινομένου. Το άθροισμα που προκύπτει είναι το όρισμα της συνάρτησης μεταφοράς.

Εάν x_{ki} είναι η i -οστή είσοδος του k νευρώνα, w_{ki} : το i -οστό συναπτικό βάρος του k νευρώνα και $\phi(\cdot)$ η συνάρτηση μεταφοράς (ή "συνάρτηση ενεργοποίησης") του νευρωνικού δικτύου, τότε η έξοδος y_k του k νευρώνα δίνεται από την εξίσωση:

$$y_k = \phi\left(\sum_{i=0}^N x_{ki} w_{ki}\right)$$

Στον k -οστό νευρώνα υπάρχει ένα συναπτικό βάρος w_{k0} με ιδιαίτερη σημασία, το οποίο καλείται πόλωση ή κατώφλι (bias, threshold). Η τιμή της εισόδου του είναι πάντα η μονάδα, $x_{k0} = 1$. Εάν το συνολικό άθροισμα από τις υπόλοιπες εισόδους του νευρώνα είναι μεγαλύτερο από την τιμή αυτή, τότε ο νευρώνας ενεργοποιείται. Εάν είναι μικρότερο, τότε ο νευρώνας παραμένει ανενεργός. Η ιδέα προέκυψε από τα βιολογικά νευρικά κύτταρα.

Ιστορικά Στοιχεία

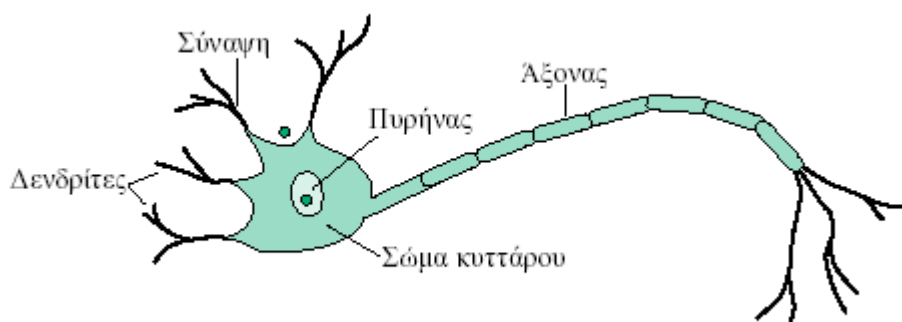
Τα νευρωνικά δίκτυα είναι συστήματα που επεξεργάζονται παράλληλα ένα μεγάλο όγκο δεδομένων και όπως προαναφέρθηκε, αναπαριστούν μια νέα υπολογιστική τεχνική που βασίζεται στον τρόπο διαχείρισης πληροφοριών από τον ανθρώπινο εγκέφαλο. Η εξέλιξη τεχνητών νευρωνικών δικτύων, από την αρχική ιδέα του νευροφυσιολόγου Heb (1949) για τη δομή και τη συμπεριφορά ενός βιολογικού νευρωνικού δικτύου έως το πιο πρόσφατο μοντέλο τεχνητών νευρωνικών δικτύων, πέρασε πολλά στάδια. Ο πρώτος ακρογωνιαίος λίθος ήρθε από τους νευρολόγους McCulloch και Pitts (1943), οι οποίοι χρησιμοποιώντας κοινή λογική μοντελοποίησαν νευρωνικά δίκτυα χρησιμοποιώντας τους νευρώνες ως δυαδικές συσκευές με προσδιορισμένα κατώφλια που αλληλοσυνδέονται μέσω συνάψεων. 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 συγκεντρωτές 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. 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. 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 Caianiello. Το πρώτο νευρωνικό δίκτυο αναπτύχθηκε το 1958 από τον Frank Rosenblatt και ονομαζόταν perceptron. Λόγω της μικρής όμως επεξεργαστικής ισχύος που ήταν διαθέσιμη την εποχή εκείνη, το ενδιαφέρον για αυτά έμεινε μειωμένο μέχρι τη στιγμή που παρουσιάστηκε ο αλγόριθμος ανάστροφου σφάλματος στα μέσα της δεκαετίας του 1980. Ωστόσο, οι Widrow και Hoff (1960) πρότειναν τον αλγόριθμο ελάχιστου μέσου τετραγώνου (least mean-square-LMS) και τον χρησιμοποίησαν για να δημιουργήσουν μια αναλογική ηλεκτρονική συσκευή γνωστή ως ADALINE (Adaptive Linear Element) και αργότερα MADALINE (Multiple ADALINE) που μπορεί να μιμηθεί τους νευρώνες. Κατά την διάρκεια της κλασσικής περιόδου του perceptron (‘60) υπήρχε η πεποίθηση 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. Το 1969 οι Minsky και Papert, έφτιαξαν το πορτρέτο του perceptron με μεγάλη επιτυχία, όμως η προσέγγιση τους ότι τα νευρωνικά δίκτυα πολλών στρωμάτων έχουν παρόμοια περιορισμένη ικανότητα μάθησης με ένα απλό νευρωνικό δίκτυο, καταρρίφθηκε αργότερα από τους Rumelhart και McClelland (1986). 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 απόπτηση για λύση προβλημάτων συνδυαστικής βελτιστοποίησης. Η εξομοιωμένη απόπτηση 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 ενδιαφέρον 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 McClelland (1986), να ήταν οι πιο σημαίνουσες εκδόσεις υπεύθυνες για την αναζωογόνηση του ενδιαφέροντος για τα νευρωνικά δίκτυα στην δεκαετία του ‘80. Οι Rumelhart και McClelland έδειξαν στην πραγματικότητα ότι τα νευρωνικά δίκτυα πολλών στρωμάτων έχουν απίστευτες δυνατότητες και είναι ικανά να εκπαιδεύονται σε πολύ πιο σύνθετα συστήματα μέσω του αλγορίθμου ανάστροφου σφάλματος. Μέχρι σήμερα, τα πιο ενδεδειγμένα μελετημένα και περισσότερο διαδεδομένα, είναι τα νευρωνικά δίκτυα πολλών στρωμάτων (Multilayer Perceptrons – MLPs). Τα MLPs είναι νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης που αποτελούνται από ένα στρώμα εισόδου (Input Layer), ένα ή περισσότερα κρυφά στρώματα (Hidden Layers) και ένα στρώμα εξόδου (Output Layer). Κάθε στρώμα αποτελείται από νευρώνες και καθένας από αυτούς συνδέεται με τους νευρώνες στο γειτονικό στρώμα με διαφορετικά βάρη (Haykin 1994).

Βιολογικά νευρωνικά δίκτυα

Λαμβάνοντας υπόψιν τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου και με δεδομένο ότι τα τεχνητά νευρωνικά δίκτυα προσπαθούν να μοιάσουν σε αυτό, τα ΝΔ διαθέτουν παρόμοιες ιδιότητες όπως την ικανότητα να μαθαίνει από εμπειρίες, να γενικεύει την υπάρχουσα γνώση και να εκτελεί λογικές αφαιρέσεις. Ο νευρώνας είναι το θεμελιακό δομικό στοιχείο του ανθρώπινου νευρικού συστήματος. Οι νευρώνες είναι ζώντα κύτταρα. Ο ανθρώπινος εγκεφάλος έχει περίπου $1.5 \cdot 10^{10}$ νευρώνες διαφόρων μορφών. Κάθε νευρώνας λαμβάνει σήματα μέσω περίπου 10^4 συνάψεων. Όλα οι νευρώνες ανεξάρτητα από το είδος τους, το σχήμα και το μέγεθος τους αποτελούνται από τα ίδια βασικά μέρη τα οποία είναι το κυτταρικό σώμα, τους δενδρίτες και τον άξονά τους. Το κυτταρικό σώμα είναι το κεντρικό μέρος του νευρώνα και το σχήμα τους μπορεί να είναι στρογγυλό, τριγωνικό, σταγονοειδές μυτερό στα δύο άκρα κλπ ανάλογα με το είδος το νευρώνα: οπτικό, αφής, μυϊκό κοκ. Τα παραπάνω φαίνονται στο παρακάτω σχήμα.



(σχήμα 3.1.1)

Η σημαντικότερη προσπάθεια μοντελοποίησης του βιολογικού κυττάρου έγινε από τους Mc Culloch και Pitts. Το μοντέλο των Mc Culloch και Pitts δεν είναι ένα πλήρες μοντέλο του νευρωνικού κυττάρου αλλά μία απλή παρουσίαση δύο καταστάσεων της λογικής διαδικασίας που λαμβάνει χώρα σε αυτό. Οι εισόδοι διαίρουνται σε δύο κατηγορίες: στις εισόδους που διεγείρουν το κύτταρο και στις απαγορευτικές εισόδους. Οι εισόδοι διέγερσης τείνουν να διεγείρουν το κύτταρο ενώ οι απαγορευτικές να απαγορεύσουν την διέγερση του κυττάρου

Από τα Βιολογικά Νευρωνικά Δίκτυα στα Τεχνητά

Ένα Τ.Ν.Δ. μοιάζει με το βιολογικό 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 ($x_1, x_2, \dots, x_n$) 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 $x_1, x_2, \dots, x_n$ αντιστοιχεί ένα βάρος $w_1, w_2, \dots, w_n$,

αντίστοιχα. Η συνολική είσοδος z του νευρώνα υπολογίζεται ως εξής:

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

Ενώ, η έξοδος y του νευρώνα υπολογίζεται ως εξής:

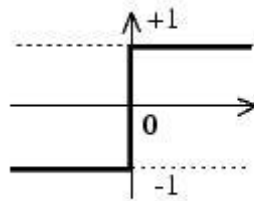
$$y = f(z) = f(w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n + b)$$

Συναρτήσεις μεταφοράς

Η συνάρτηση μεταφοράς μπορεί να είναι βηματική (step transfer function), γραμμική (linear transfer function), μη γραμμική (non-linear transfer function), στοχαστική (stochastic transfer function)

Η βηματική συνάρτηση μεταφοράς μπορεί να είναι:

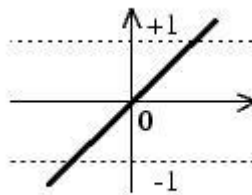
$$\phi(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



ή οποιαδήποτε άλλη βηματική συνάρτηση.

Η γραμμική συνάρτηση μεταφοράς μπορεί να είναι:

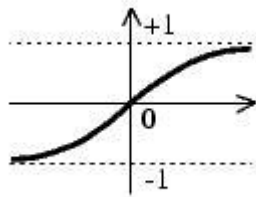
$$\phi(x) = x$$



ή οποιαδήποτε άλλη γραμμική συνάρτηση.

Η μη γραμμική συνάρτηση μεταφοράς που χρησιμοποιείται συνήθως στα νευρωνικά δίκτυα καλείται σιγμοειδής συνάρτηση:

$$\phi(x) = \frac{1}{1 + e^x}$$



Χρήση Νευρωνικών Δικτύων και Εφαρμογές.

Τα νευρωνικά δίκτυα, από το αρχικό στάδιο της πρακτικής τους χρήσης, έχουν αποδειχθεί πολύτιμο εργαλείο για την ανάλυση σημάτων, εξαγωγή χαρακτηριστικών, ταξινόμηση δεδομένων, αναγνώριση σχεδίων κτλ. Εξαιτίας της ικανότητας τους να εκπαιδεύονται και να μπορούν να κάνουν γενικεύσεις από δεδομένα που προκύπτουν μέσω παρατήρησης, τα δίκτυα αυτά έχουν εξαπλωθεί και βρίσκουν μεγάλη ανταπόκριση από μηχανικούς και ερευνητές σαν εργαλείο επεξεργασίας και ανάλυσης δεδομένων. Και αυτό κυρίως διότι τα νευρωνικά δίκτυα μειώνουν σε τεράστιο βαθμό τον υπολογιστικό φόρτο για την επίλυση προβλημάτων, και λόγω του χαρακτηριστικού τους να δουλεύουν παράλληλα, επιταχύνουν τη διαδικασία επίλυσης. Αυτός ο λόγος ήταν αρκετός για να ξεφύγει αυτή η τεχνολογία δικτύων από τα εργαστήρια ερευνών και να διεισδύσουν στη βιομηχανία, επιχειρήσεις κτλ. Για παράδειγμα η προσέγγιση μέσω μεθοδολογιών νευρωνικών δικτύων έχουν λύσει ουσιαστικά προβλήματα στην ανάλυση χρονοσειρών και πρόβλεψης χρησιμοποιώντας δεδομένα καταγεγραμμένα από παρατήρηση.

Γενικότερα, η πρακτική χρήση των νευρωνικών δικτύων, υφίσταται λόγω των κάτωθι χαρακτηριστικών:

- Μη γραμμική συσχέτιση ανάμεσα σε ένα υποσύνολο της παρελθούσης δεδομένης χρονοσειράς και της μελλοντικής προς πρόβλεψη χρονοσειράς.
- Η ικανότητα να αιχμαλωτίζει ουσιώδεις λειτουργικές σχέσεις ανάμεσα στα δεδομένα, οι οποίες είναι ιδιαίτερα χρήσιμες όταν αυτές οι σχέσεις δεν είναι γνωστές εκ των προτέρων ή είναι πολύ δύσκολο να περιγραφούν μαθηματικά ή όταν τα δεδομένα μας περιέχουν θόρυβο.
- Ικανότητα παγκόσμιας κανονικοποίησης συναρτήσεων που επιτρέπει τη μοντελοποίηση αυθαίρετων συνεχών μη γραμμικών συναρτήσεων σε οποιοδήποτε βαθμό ακρίβειας.
- Δυνατότητα εκπαίδευσης και γενίκευσης από παραδείγματα που χρησιμοποιούν προσέγγιση από αυτοπροσδιοριζόμενα δεδομένα.
- Ο 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.

- Η Προσαρμοστικότητα των Νευρωνικών Δικτύων που έχουν τη δυνατότητα να διαμορφώνουν 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.
- 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. Αυτό ονομάζεται Αποδεικτική Απόκριση.
- Συναφής Πληροφορία: Η γνώση αναπαριστάνεται από την πολύ δομημένη και ενεργή κατάσταση του Νευρωνικού Δικτύου.
- 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.
- Η συμπαγής παράλληλη φύση του Νευρωνικού Δικτύου, κάνει δυνατή την υλοποίηση του σε VLSI (Very-large-scale integration) τεχνολογία, έτσι ώστε τα νευρωνικά δίκτυα να μπορούν να χρησιμοποιηθούν σε εφαρμογές πραγματικού χρόνου.
- 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.
- 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.

Η ευρεία εφαρμογή των νευρωνικών δικτύων σε μία ποικιλία προβλημάτων, σε πολλά διαφορετικά πεδία, τα κάνει πολύ ελκυστικά. Επίσης, οι πιο 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 κωδικών, 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, δυναμική μοντελοποίηση 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.

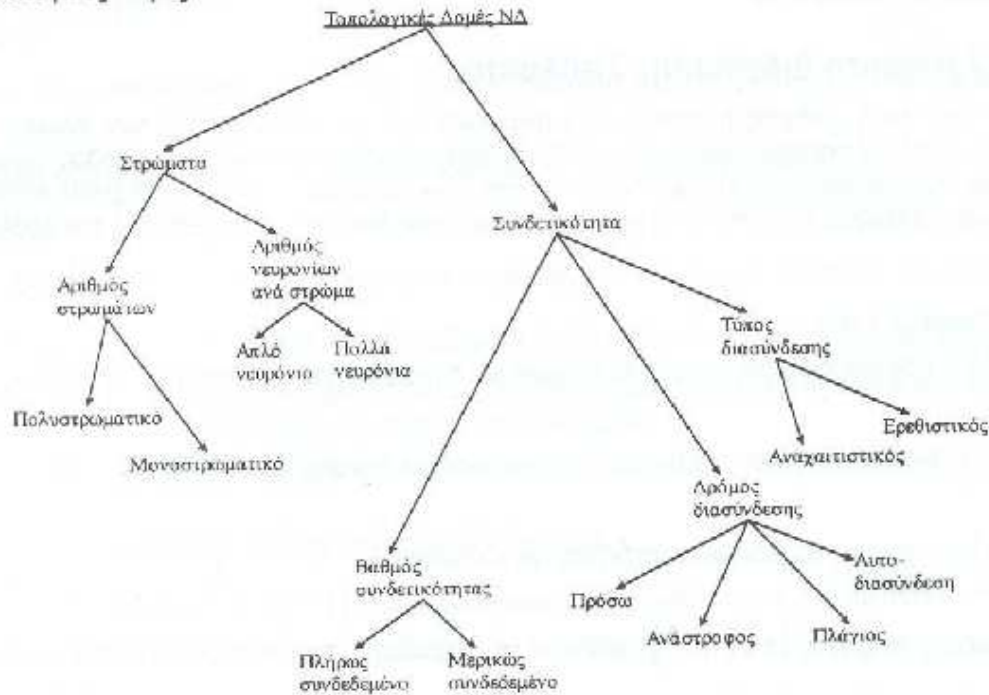
Μεταφορές: Συστήματα διάγνωσης βλαβών φρένων, χρονοπρογραμματισμός οχημάτων, συστήματα δρομολόγησης.

Από το παραπάνω πλήθος εφαρμογών, μπορούμε να συμπεράνουμε ότι τα νευρωνικά δίκτυα έχουν αναπτυχθεί ταχύτατα. Για αυτό το λόγο έχουν επενδυθεί πολλά χρήματα τόσο για την ανάπτυξη λογισμικού όσο και για την ανάπτυξη υλικού.

Αρχιτεκτονική Νευρωνικών Δικτύων

Η τοπολογική δομή είναι το κύριο χαρακτηριστικό των νευρωνικών δικτύων και αναφέρεται στην αρχιτεκτονική στην οποία διευθετούνται και διασυνδέονται πολλαπλοί νευρώνες. Οι δύο βασικές ιδιότητες που καθορίζουν την αρχιτεκτονική ενός νευρωνικού δικτύου είναι το πλήθος των στρωμάτων και οι συνδέσεις ανάμεσα στους νευρώνες. Πολύ σημαντική παράμετρος είναι ο αλγόριθμος μάθησης που χρησιμοποιείται για την εκπαίδευση του δικτύου. Παρακάτω δίνονται οι τοπολογικές δομές των νευρωνικών δικτύων (σχήμα 3.1.2).

Τοπολογικές Δομές ΝΔ



(σχήμα 3.1.2)

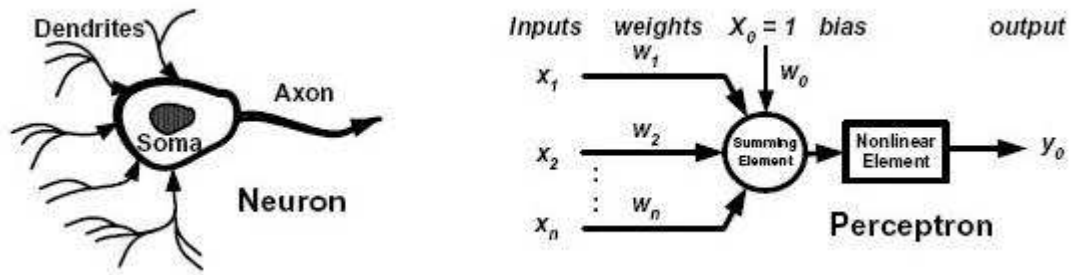
Το μοντέλο του βασικού στοιχείου ενός νευρωνικού δικτύου (βλ. νευρόνα), όπως χρησιμοποιείται ακόμα και σήμερα έχει πρωτοανακαλυφθεί από τους Widrow και Hoff (1960). Αυτοί προσδιόρισαν το perceptron ως το προσαρμοστικό στοιχείο που φέρει ομοιότητες με το νευρόνα. Ένας νευρώνας, που είναι το θεμέλιο ενός νευρωνικού συστήματος επεξεργασίας πληροφοριών, αποτελείται από (σχήμα 3.1.3):

- το σώμα (cell body) που περιέχει ένα νουκλεόνιο
- τους δενδρίτες που τροφοδοτούν με εξωτερικά σήματα το σώμα
- τους νευρίτες (axons) που μεταφέρουν το σήμα από το σώμα στα υπόλοιπα σώματα

Αυτή η διάταξη μπορεί να μεταφραστεί σε ανάλογους όρους υπολογιστικής τεχνολογίας, με τον εξής τρόπο:

- το κομμάτι του πυρήνα του στοιχείου, που ονομάζεται perceptron, περιέχει ένα αθροιστικό στοιχείο Σ και ένα μη γραμμικό στοιχείο NL
- τα πολλαπλά σήματα εισόδου x_i συνδέονται μέσω βαρών που μπορούν να ρυθμιστούν και να επαναπροσδιοριστούν, με τον πυρήνα του στοιχείου
- το σήμα εξόδου (που πολλές φορές μπορεί να είναι παραπάνω από ένα) y_d

Ένα επιπρόσθετο perceptron εισόδου w_0 , που ονομάζεται πόλωση (bias) και είναι ουσιαστικά το κατώφλι. Η τιμή που παίρνει πάντα είναι μονάδα. Εάν το συνολικό άθροισμα από τις υπόλοιπες εισόδους του νευρώνα είναι μεγαλύτερο από την τιμή αυτή, τότε ο νευρώνας ενεργοποιείται. Εάν είναι μικρότερο, τότε ο νευρώνας παραμένει ανενεργός.



(σχήμα 3.1.3)

Το σήμα εξόδου δίνεται από τη συνάρτηση:

$$y_0 = f\left(\sum_{i=1}^n w_i x_i + w_0\right)$$

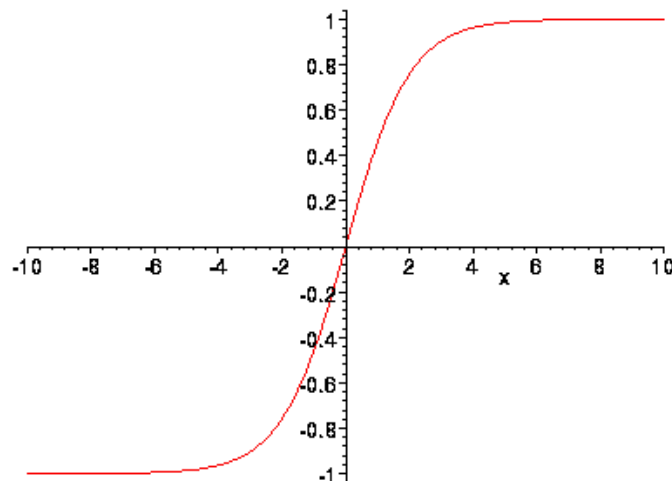
και το κατώφλι ακολουθεί την ανισότητα:

$$w^T x + w_0 \geq 0$$

που σημαίνει ότι το perceptron ενεργοποιείται και παράγει σήμα εξόδου όταν επαληθεύεται η παραπάνω σχέση.

Τώρα η προσοχή μας θα εστιαστεί στη μη γραμμική συνάρτηση που θα έπρεπε να χρησιμοποιηθεί στον πυρήνα του perceptron, εφόσον πρόκειται για τη συνάρτηση ενεργοποίησης. Η πρώτη απόπειρα του Block (1962) να επιλέξει τη δυαδική βηματική συνάρτηση για το σκοπό αυτό, αργότερα εγκαταλείφθηκε για να χρησιμοποιηθεί η σιγμοειδής συνάρτηση ενεργοποίησης (σχήμα 3.1.4).

$$f(x) = \frac{1}{1 + \exp(-x)}$$



(σχήμα 3.1.4)

Σιγμοειδής συνάρτηση

Το perceptron ουσιαστικά μαθαίνει μέσω μιας διαδικασίας εκπαίδευσης βασισμένη σε δεδομένα που έχουμε συλλέξει. Κατά τη διαδικασία της εκπαίδευσης, το perceptron αυξομειώνει τα βάρη διασύνδεσης ανάλογα με τα δεδομένα εισόδου του. Για τη ρύθμιση των βαρών αυτών, οι Widrow και Hoff (1960) πρότειναν τη χρήση

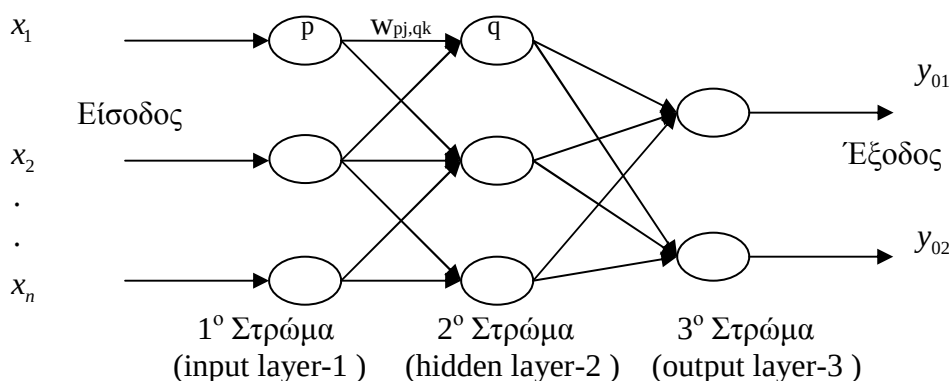
του κανόνα Δέλτα (Delta rule) κατά τον οποίο προστίθεται στην προσωρινή τιμή του βάρους $w(k)$ ο συμψηφιστικός όρος $\eta \varepsilon(k)x(k)$, ώστε να δημιουργηθεί η νέα τιμή

$$w(k+1) = w(k) + \eta \varepsilon(k)x(k),$$

όπου η είναι ένας αναλογικός όρος, $\varepsilon(k)$ είναι το σφάλμα στο ρυθμιστικό βήμα k και $x(k)$ η τιμή του σήματος εισόδου στο συγκεκριμένο βήμα k .

Αν και αρκετά απλή, αυτή η μέθοδος εκπαίδευσης έχει στην πλειονότητα των περιπτώσεων επιδείξει μεγάλη ακρίβεια, αποτελεσματικότητα και υψηλή ταχύτητα σύγκλισης. Πάραυτα, ένα μοναχό perceptron δεν μπορεί να εκπαιδευτεί τόσο ώστε να είναι ικανό να επιλύσει σύνθετα προβλήματα διότι η ακτίνα υπολογιστικής του δράσης είναι περιορισμένη από την απλότητα της δομής του. Στην πραγματικότητα μπορεί να επιλύσει μόνο γραμμικά προβλήματα, με αποτέλεσμα να το καθιστά αναποτελεσματικό μηχανισμό γενικής επίλυσης. Την άποψη του Minsky ότι ακόμα και perceptrons πολλών επιπέδων δεν μπορούσαν να λειτουργήσουν ικανοποιητικά, ήρθε να καταρρίψει η κατασκευή πολυστρωματικών, multilayer perceptrons (MLPs). Τα MLPs αποτελούνται από το στρώμα εισόδου (input layer), το στρώμα εξόδου (output layer) και τα κρυφά στρώματα (hidden layers), δημιουργώντας μια δομή με προχωρημένες ικανότητες διασύνδεσης. Στο σχήμα 3 που ακολουθεί, παρουσιάζεται 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).

Το νευρωνικό αυτό δίκτυο έχει $3 \cdot 3 = 9$ συνδέσεις μεταξύ των 3 νευρώνων εισόδου και των 3 κρυφών νευρώνων, και $3 \cdot 2 = 6$ 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 εξόδου χρησιμοποιείται κάποια συνάρτηση ενεργοποίησης (μεταφοράς), για παράδειγμα η σιγμοειδής συνάρτηση.

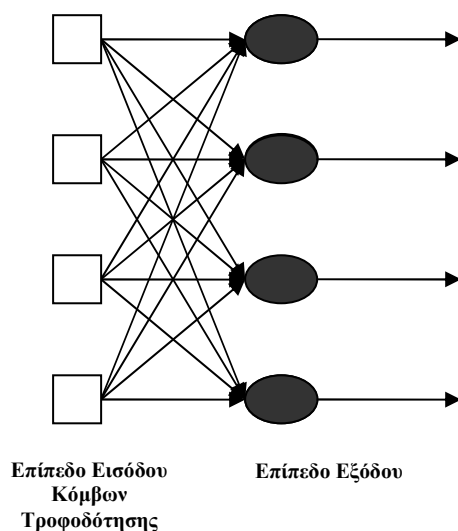


(σχήμα 3.1.5)

Είδη Αρχιτεκτονικών Δομών

1.Ενός Επιπέδου Πρόσθιας Τροφοδότησης Δίκτυα.

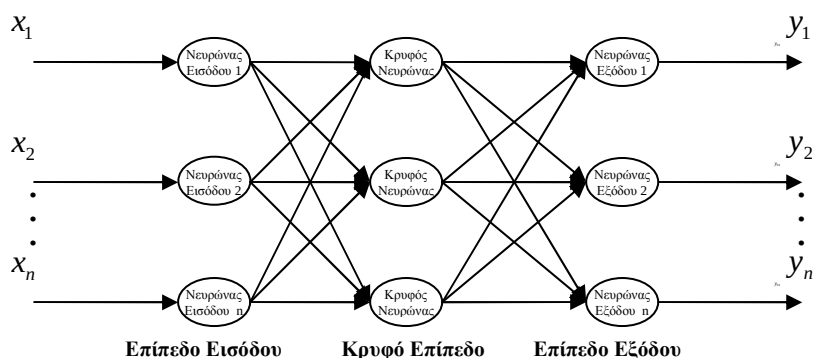
Σε ένα τέτοιο δίκτυο, οι νευρώνες είναι οργανωμένοι σε μορφή επιπέδων. Οι νευρώνες του πηγαίου επιπέδου δείχνουν στους νευρώνες του επόμενου επιπέδου αλλά όχι αντίστροφα, σύμφωνα με το σχήμα 3.1.6 που ακολουθεί.



(σχήμα 3.1.6): Πρόσθιας Τροφοδότησης δίκτυο με ένα επίπεδο νευρώνων.

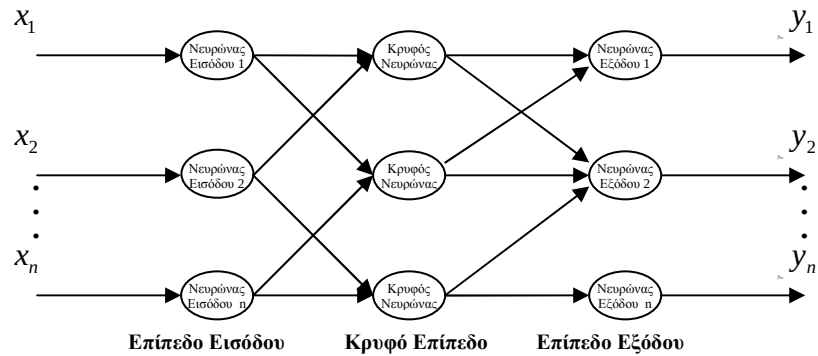
2. Πολλαπλών Επιπέδων Πρόσθιας Τροφοδότησης Δίκτυα.

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 3.1.7 παρακάτω έχουμε ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο, με την έννοια ότι κάθε κόμβος συνδέεται με όλους τους κόμβους του αμέσως επόμενου επιπέδου.



(σχήμα 3.1.7): Πλήρως συνδεδεμένο δίκτυο πρόσθιας τροφοδότησης

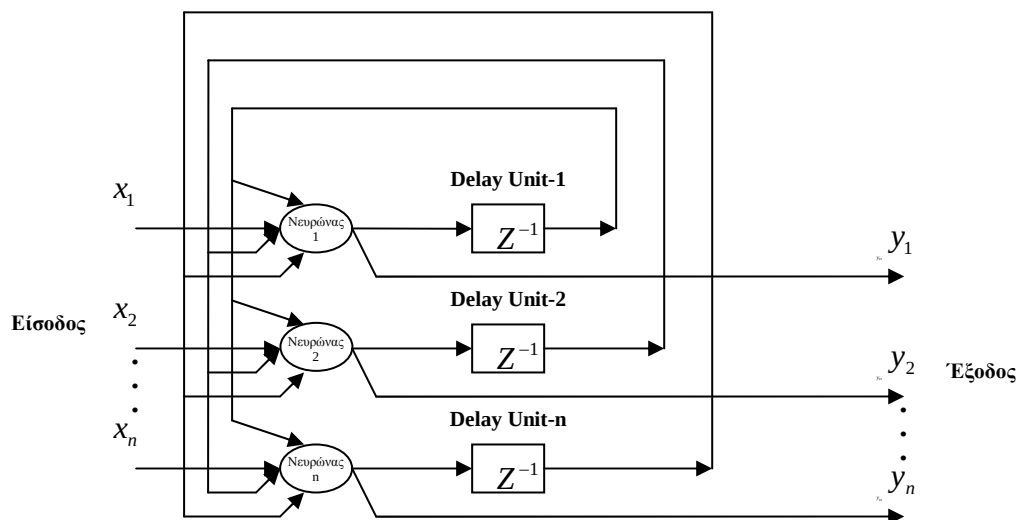
Στο σχήμα που ακολουθεί, παρουσιάζεται ένα μερικώς συνδεδεμένο δίκτυο πρόσθιας τροφοδότησης.



(Σχήμα 3.1.8): Μερικώς συνδεδεμένο δίκτυο πρόσθιας τροφοδότησης

3.Αναδρομικά Δίκτυα

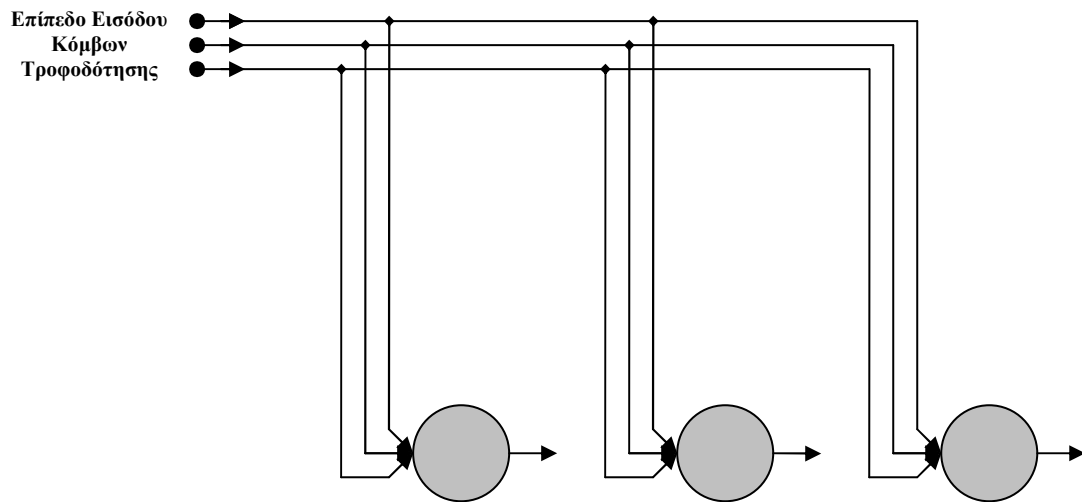
Ένα αναδρομικό νευρωνικό δίκτυο έχει εκτός από τις προς τα εμπρός και προς τα πίσω συνδέσεις. Έχει δηλαδή βρόγχους αναδρομής από τις εξόδους του προς τις εισόδους του. Η παρουσία τέτοιων βρόγχων έχει ισχυρή επίπτωση στην ικανότητα μάθησης του δικτύου. Χαρακτηριστικό παράδειγμα αναδρομικού δικτύου αποτελεί το δίκτυο Hopfield και παρουσιάζεται στο σχήμα 3.1.9 που ακολουθεί.



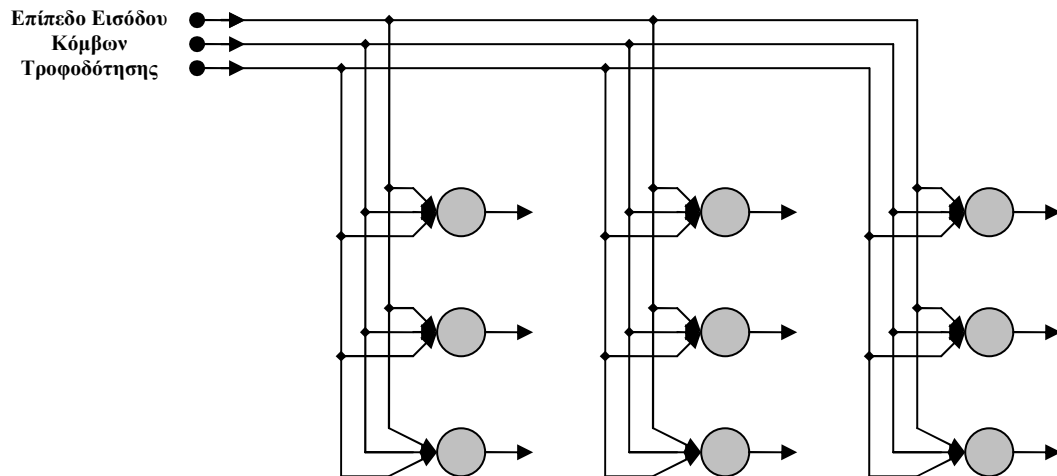
(Σχήμα 3.1.9): Σχηματισμός Δικτύου

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 στα Σχήματα 3.1.10 και 3.1.11 που ακολουθούν.



(Σχήμα 3.1.10): Μονοδιάστατο πλέγμα με 3 νευρώνες.



(Σχήμα 3.1.11): 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.

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

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 $\hat{y}$ 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 στρώματων 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 αλγόριθμος 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:

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

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

Όπου:

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).

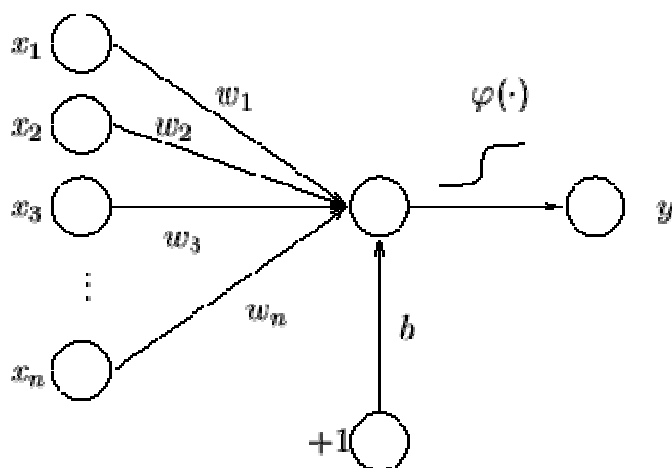
Η αναδρομική διαδικασία εκπαίδευσης (τίθεται $n=n+1$) εκτελείται μέχρι ένα συγκεκριμένο αριθμό επαναλήψεων είτε εναλλακτικά μέχρι να επιτευχθεί ικανοποιητική πρόβλεψη των τιμών εξόδου. Συνήθη κριτήρια τερματισμού του αλγόριθμου εκπαίδευσης περιλαμβάνουν επιτήρηση του αθροίσματος των τετραγώνων των σφαλμάτων, 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).

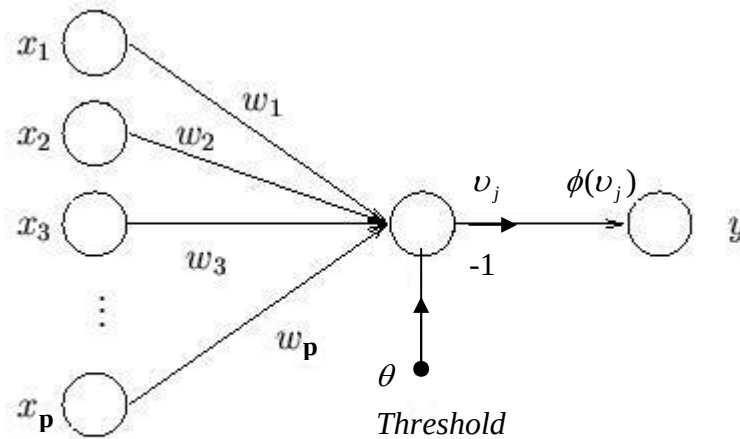
Ένα τέτοιο δίκτυο φαίνεται στο σχήμα που ακολουθεί με x_1 έως x_n να είναι η είσοδος, y η έξοδος και b το κατώφλι (threshold).



(Σχήμα 3.1.12): Perceptron ενός επιπέδου

Προκειμένου να εκπαιδευτεί ένα τέτοιο Ν.Δ., σαν αλγόριθμος εκπαίδευσης χρησιμοποιείται ο γνωστός κανόνας του Rosenblatt. 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 3.1.13.



(Σχήμα 3.1.13): Το διάγραμμα ροής σήματος του Perceptron.

Η έξοδος του γραμμικού συνδυαστή υπολογίζεται εύκολα από το διάγραμμα του σχήματος 11 και είναι:

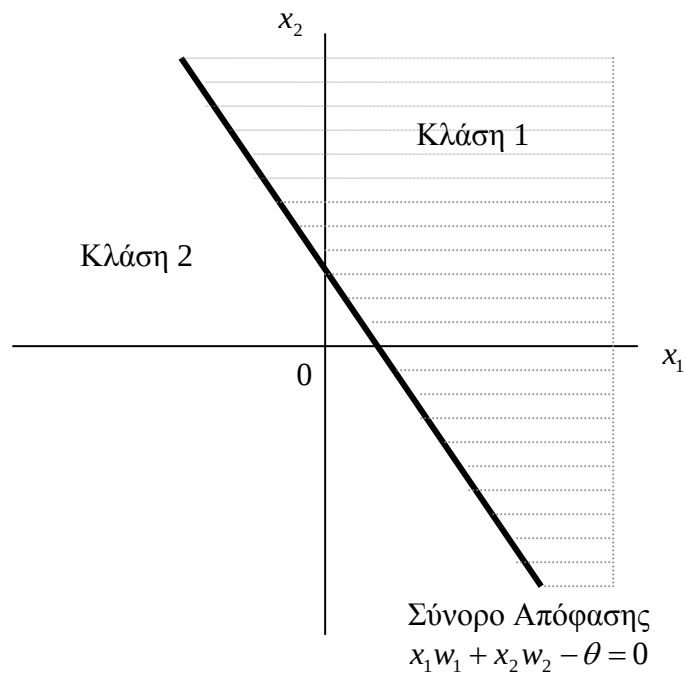
$$v = \sum_{i=1}^p w_i x_i - \theta$$

Σκοπός του Perceptron είναι να ταξινομήσει ένα σύνολο εισόδων (προτύπων) $x_1, x_2, \dots, x_n$ σε μία από τις κλάσεις I_1 και I_2 . Ο κανόνας απόφασης για την ταξινόμηση είναι ο εξής:

ανάθεσε το σημείο που αναπαριστούν $x_1, x_2, \dots, x_n$ στην κλάση I_1 , αν $y = +1$ και στην κλάση I_2 αν $y = -1$. Οι περιοχές απόφασης διαχωρίζονται από το υπερεπίπεδο που ορίζεται από τη σχέση:

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

Στο σχήμα 3.1.14 στην επόμενη σελίδα φαίνεται η γραμμική διαχωρισιμότητα για ένα δισδιάστατο πρόβλημα ταξινόμησης, με δύο κλάσεις.



Σχήμα 3.1.14: Το όριο και οι περιοχές απόφασης για ένα δισδιάστατο πρόβλημα ταξινόμησης δύο κλάσεων.

Από το παραπάνω σχήμα φαίνεται το αποτέλεσμα της εφαρμογής του κατώφλιου, 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

$\mathbf{x}(n)$ = $(p + 1) \times 1$ input vector

$[-1, x_1(n), x_2(n), \dots, x_p(n)]^T$

$\mathbf{w}(n)$ = $(p + 1) \times 1$ weight vector

$[\theta(n), w_1(n), w_2(n), \dots, w_p(n)]^T$

$\theta(n)$ = threshold (κατώφλι)

$y(n)$ = actual response (πραγματική έξοδος)

$d(n)$ = desired response (επιθυμητή έξοδος)

η = learning - rate parameter, θετική σταθερά < 1

Βήμα 1: Αρχικοποίηση

Θέσε $\mathbf{w}(0) = 0$. Κατόπιν κάνε τους υπολογισμούς για $\eta = 1, 2, \dots$

Βήμα 2: Ενεργοποίηση

Στο χρόνο n , ενεργοποίησε το Perceptron εφαρμόζονται το συνεχές διάνυσμα εισόδου $x(n)$ και το $d(n)$.

Βήμα 3: Υπολογισμός πραγματικής απόκρισης

Υπολόγισε την πραγματική απόκριση του Perceptron:

$$y(n) = \text{sgn}[w^T(n)x(n)]$$

Βήμα 4: Προσαρμογή διανύσματος βαρών

Προσάρμοσε τα βάρη του Perceptron:

$$w(n+1) = w(n) + n [d(n) - y(n)] x(n) \text{ (E.C.L. rule)}$$

όπου:

$$d(n) = \begin{cases} +1, & \text{εάν } x(n) \text{ ανήκει στην κλάση } I_1 \\ -1, & \text{εάν } x(n) \text{ ανήκει στην κλάση } I_2 \end{cases}$$

Βήμα 5: Αύξησε το χρόνο n κατά μια μονάδα και πήγαινε στο βήμα 2.

Ο Αλγόριθμος Ελάχιστου Μέσου Τετραγωνικού (EMT) Σφάλματος

Ο αλγόριθμος Ελάχιστου Μέσου Τετραγωνικού (EMT) σφάλματος εκπαιδεύει 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 βρει 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 $\mathbf{r}_x(\mathbf{j}, \mathbf{k})$ και της συνάρτησης ετεροσυσχέτισης $\mathbf{r}_{dk}(\mathbf{k})$. Μία σύνοψη του αλγορίθμου LMS παρουσιάζεται παρακάτω. Γίνεται αμέσως αντιληπτή η απλότητα του αλγορίθμου.

Αλγόριθμος LMS

1. Αρχικοποίηση. Θέσε

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

2. Υπολόγισε για $n=1,2,\dots$

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

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

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

Ο Αλγόριθμος Πίσω Διάδοσης (Π.Δ.) του Λάθους

Ο αλγόριθμος Πίσω Διάδοσης (Π.Δ.) του λάθους εκπαιδεύει 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 των 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 $\mathbf{w}^*$ δηλώνει ένα ελάχιστο, τοπικό ή ολικό. Για να είναι το $\mathbf{w}^*$ ένα ελάχιστο πρέπει το διάνυσμα κλίσης $\mathbf{g}(\mathbf{w})$ της επιφάνειας λάθους σε σχέση με το διάνυσμα βαρών $\mathbf{w}$ να είναι μηδέν όταν $\mathbf{w} = \mathbf{w}^*$. Αντίστοιχα, μπορούμε να διατυπώσουμε ένα λογικό κριτήριο σύγκλισης για την μάθηση πίσω-διάδοσης όπως παρακάτω (Kramer and Sangiovanni-Vincentelli, 1989):

• *Ο αλγόριθμος πίσω-διάδοσης συγκλίνει όταν η Ευκλείδεια νόρμα του διανύσματος κλίσης φτάσει σε 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 $\mathbf{g}(\mathbf{w})$. Άλλη μοναδική ιδιότητα ενός ελαχίστου είναι ότι η συνάρτηση κόστους ή μέτρο λάθους $E_{av}(\mathbf{w})$ είναι στάσιμο στο σημείο $\mathbf{w}=\mathbf{w}^*$ 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 αρκετά 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 $E_{av}(\mathbf{w})$ να είναι ίση ή μικρότερη από ένα αρκετά μικρό κατώφλι. Οι Kramer και Sangiovanni Vincentelli (1989) πρότειναν ένα υβριδικό κριτήριο σύγκλισης που συνίσταται από αυτό το τελευταίο κατώφλι και ένα κατώφλι κλίσης, όπως δηλώνεται παρακάτω:

- Ο αλγόριθμος πίσω-διάδοσης τερματίζεται στο διάνυσμα βαρών $\mathbf{w}_{final}$ όταν $\|\mathbf{g}(\mathbf{w}_{final})\| \leq \epsilon$, όπου ϵ είναι ένα αρκετά μικρό κατώφλι κλίσης, ή όταν $E_{av}(\mathbf{w}_{final}) \leq \tau$, όπου τ 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.

3.2 Ασαφής Λογική (Fuzzy Logic)

Τον όρο «ασαφής λογική» (fuzzy logic) εισήγαγε το 1962 με άρθρο του ο L.A. Zadeh, ο οποίος αναφέρθηκε στην αναγκαιότητα δημιουργίας μίας μαθηματικής θεωρίας που θα επεξεργάζεται ασαφείς-ανακριβείς έννοιες, οι οποίες δεν είναι δυνατό να μοντελοποιηθούν με τη θεωρία των πιθανοτήτων και θεμελίωσε τη θεωρία των ασαφών συνόλων σε μια μέθοδο διαχείρισης της ανακρίβειας που αντιμετωπίζεται σε πολλά πρακτικά προβλήματα. Η ανακρίβεια, ή η ασάφεια είναι ο πυρήνας των ασαφών συνόλων και της ασαφούς λογικής. Τα ασαφή σύνολα ουσιαστικά αποτελούν μια γενίκευση των κλασσικών συνόλων. Από τις αρχές της δεκαετίας του 1980 τα ασαφή σύνολα βρήκαν πολλές εφαρμογές, ιδίως σε συστήματα ελέγχου.

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

Ένα ασαφές σύνολο (fuzzy set) ορίζεται ως ένα σύνολο διατεταγμένων ζευγών $(x, \mu_A(x))$ όπου $x \in X$ και $\mu_A(x) \in [0,1]$. Το σύνολο X αποτελεί ένα ευρύτερο σύνολο αναφοράς (universe of discourse) που περιλαμβάνει όλα τα αντικείμενα στα οποία μπορεί να γίνει αναφορά. Η τιμή $\mu_A(x)$ λέγεται βαθμός αληθείας (degree of truth), συμβολίζει το βαθμό συγγένειας του x στο A και παίρνει τιμές στο διάστημα $[0,1]$. Τέλος η συνάρτηση μ_A ονομάζεται συνάρτηση συμμετοχής (συγγένειας) (membership function). Η διαφορά των ασαφών συνόλων συγκριτικά με την κλασική θεωρία συνόλων είναι ότι στη δεύτερη ισχύει $\mu_A(x) \in \{0,1\}$, δηλαδή το x είτε ανήκει στο A [$\mu_A(x) = 1$] ή δεν ανήκει [$\mu_A(x) = 0$].

Το ασαφές σύνολο ορίζεται ως:

$$A = \sum_{x_i \in X} \mu_A(x_i) / x_i \quad , \text{ όταν } X \text{ είναι ένα σύνολο διακριτών αντικειμένων και}$$

$$A = \int_X \mu_A(x) / x \quad , \text{ όταν το } X \text{ είναι ένα συνεχές διάστημα (συνήθως το } \mathbb{R}^+) \text{}$$

Στην πράξη η συνάρτηση συμμετοχής μπορεί να προέρχεται από:

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

Συναρτήσεις συμμετοχής

Παρακάτω θα παρουσιάσουμε ορισμένες μορφές μονοδιάστατων παραμετρικών συναρτήσεων συμμετοχής, δηλαδή συναρτήσεων συμμετοχής με μία μοναδική είσοδο (input).

1. **Triangular MFs** (συναρτήσεις συμμετοχής τριγωνικής μορφής)

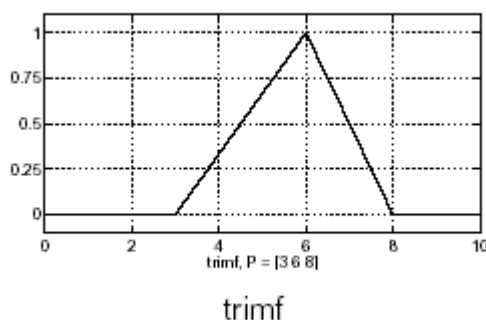
Μια τριγωνικής μορφής συνάρτηση συμμετοχής προσδιορίζεται από τρεις παραμέτρους $\{a, b, c\}$ ως ακολούθως:

$$\text{trimf}(x, a, b, c) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right)$$

Οι παράμετροι $\{a, b, c\}$ (όπου $a < b < c$) καθορίζουν τις συντεταγμένες x των τριών γωνιών της συγκεκριμένης τριγωνικής συνάρτησης συμμετοχής.

Η μαθηματική έκφραση της τριγωνικής συνάρτησης συμμετοχής είναι η εξής:

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



2. **Trapezoidal MFs:**

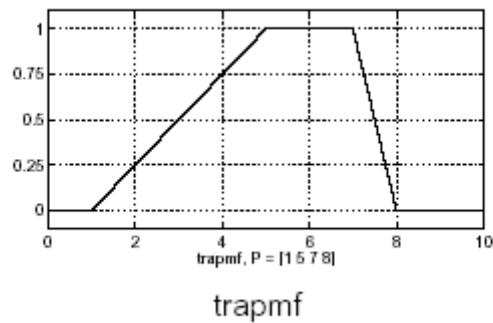
Η τραπεζοειδής συνάρτηση συμμετοχής προσδιορίζεται από τέσσερις παραμέτρους $\{a, b, c, d\}$ ως ακολούθως:

$$\text{trapmf}(x, a, b, c, d) = \max\left(\min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right)$$

Οι παράμετροι $\{a, b, c, d\}$ με $a < b < c < d$ καθορίζουν τις συντεταγμένες x των τεσσάρων γωνιών της συγκεκριμένης τραπεζοειδούς συνάρτησης συμμετοχής.

Παρακάτω δίδεται η μαθηματική έκφραση της τραπεζοειδούς συνάρτησης συμμετοχής:

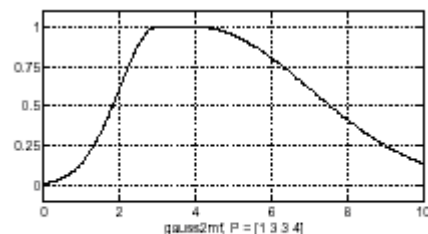
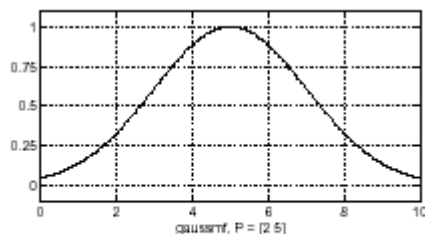
$$A = \begin{cases} 0, x \leq a \\ (x-a)/(b-a), x \in (a,b) \\ 1, x \in (b,c) \\ (d-x)/(d-c), x \in (c,d) \end{cases}$$



3. Gaussian MFs:

Μια Γκαουσιανή συνάρτηση συμμετοχής προσδιορίζεται από δύο παραμέτρους {c,σ}

$$\text{gaussmf}(x, c, \sigma) = e^{-\frac{1}{2} \left(\frac{x-c}{\sigma} \right)^2}$$



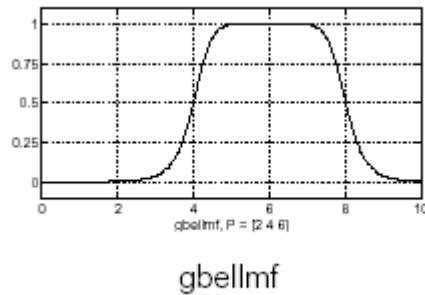
Η Γκαουσιανή συνάρτηση συμμετοχής προσδιορίζεται επακριβώς από δύο παραμέτρους {c,σ}. Το c αναπαριστά το κέντρο της, και το σ το πλάτος της.

Η Γκαουσιανή και η καμπάνα μπορούν να χρησιμοποιούνται συχνά στα ασαφή σύνολα λόγω της ομαλότητάς τους. Έχουν δε το πλεονέκτημα να διατηρούν μη μηδενικές τιμές σε όλα τα σημεία.

4. Generalized bell MFs:

Η γενικευμένη καμπανοειδής συνάρτηση συμμετοχής προσδιορίζεται από τρεις παραμέτρους {a,b,c}, όπου η παράμετρος b είναι συνήθως θετική.

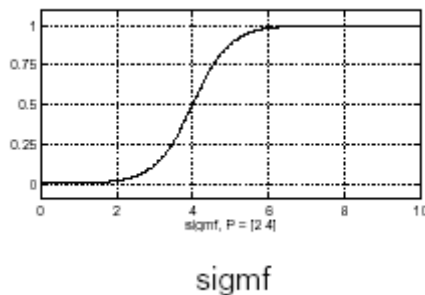
$$gbellmf(x, a, b, c) = \frac{1}{1 + \left| \frac{x - c}{a} \right|^{2b}}$$



5. Sigmoidal MF

Η σιγμοειδής συνάρτηση συμμετοχής προσδιορίζεται ως ακολούθως:

$$sigmf(x, a, c) = \frac{1}{1 + \exp[-a(x - c)]}$$



Το a ρυθμίζει την κλίση στο crossover point $x=c$. Ανάλογα με την τιμή που θα πάρει το a η σιγμοειδής MF είναι open left, open right και επομένως είναι κατάλληλη να χρησιμοποιηθεί στην αναπαράσταση εννοιών όπως «πολύ μεγάλο» ή «πολύ αρνητικό». Σιγμοειδείς συναρτήσεις αυτής της μορφής χρησιμοποιούνται ευρύτατα ως συναρτήσεις ενεργοποίησης (activation function) των τεχνητών νευρωνικών δικτύων.

Ιδιότητες των ασαφών συνόλων

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

Το σύνολο A μπορεί να προσδιοριστεί επομένως από το σύνολο των παρακάτω ζευγών:

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

Κάθε ζευγάρι $(x, \mu_A(x))$ ονομάζεται **μονοσύνολο**

Χώρος αναφοράς του ασαφούς συνόλου A είναι το σύνολο των στοιχείων του X που έχουν μη μηδενικό βαθμό συμμετοχής στο A .

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

Το σύνολο A μπορεί να γραφεί και ως:

$$A = \mu_1/x_1 + \mu_2/x_2 + \dots$$

$$A = \sum \mu_i/x_i$$

Στο συνεχή χώρο αναφοράς :

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

Ασαφές δυναμοσύνολο

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

Υποσύνολο

Το σύνολο A είναι υποσύνολο του B $A \subseteq B$ αν και μόνο αν

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

Αν ταυτόχρονα τα A και B δεν είναι ίσα, τότε το A θα ονομάζεται γνήσιο υποσύνολο του B

Ασαφής διαμέριση

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

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

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

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

Κενό ασαφές σύνολο

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

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

Κανονικό ασαφές σύνολο

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

$$A = \text{κανονικό} \quad \text{αν} \quad \exists x_1: \mu_A(x_1) = 1$$

Ισότητα ασαφών συνόλων

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

$$A=B \quad \text{αν} \quad \mu_A(x) = \mu_B(x) \quad \forall x \in X$$

Συστολή ασαφών συνόλων

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

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

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

Διαστολή ασαφών συνόλων

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

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

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

Λογικές πράξεις στα ασαφή σύνολα

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

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

AND

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

OR

A	not A
0	1
1	0

NOT

Μεταβαίνοντας στον χώρο της ασαφούς λογικής πρέπει να έχουμε κατά νου ότι οι έννοιες αληθές και ψευδές είναι θέμα βαθμού συμμετοχής. Επομένως ο πίνακας αυτός πρέπει να μετατραπεί με τρόπο που να συμπεριλαμβάνει αυτήν την αρχή. Οι τιμές των εισόδων A και B είναι τώρα πραγματικοί αριθμοί από το 0 μέχρι το 1. Πρέπει λοιπόν να ευρεθεί μια συνάρτηση που να διατηρεί τις ιδιότητες της συνάρτησης AND και ταυτόχρονα να μπορεί να επεκτείνεται για πραγματικούς αριθμούς.

Μια πιθανή απάντηση μπορεί να είναι ο τελεστής $\min(A,B)$, η ελάχιστη δηλαδή τιμή των εισόδων A και B . Με βάση το ίδιο σκεπτικό μια συνάρτηση που μπορεί να αντικαταστήσει τον τελεστή OR της Boolean λογικής είναι ο τελεστής $\max(A,B)$. Τέλος ο τελεστής NOT A μπορεί να αντικατασταθεί με την πράξη $1-A$.

Παρακάτω βλέπουμε ότι ο πίνακας αληθείας παραμένει αμετάβλητος αν εφαρμόσουμε τις παραπάνω υποκαταστάσεις.

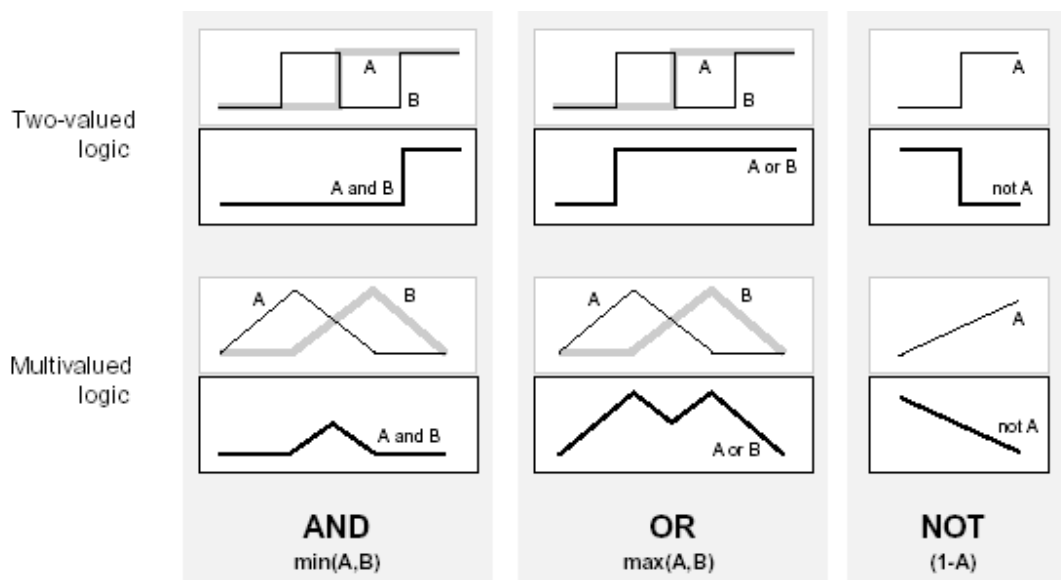
A	B	$\min(A,B)$	A	B	$\max(A,B)$	A	$1-A$
0	0	0	0	0	0	0	1
0	1	0	0	1	1	1	0
1	0	0	1	0	1		
1	1	1	1	1	1		

AND

OR

NOT

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



Δεδομένων των παραπάνω συναρτήσεων μπορούμε να κατασκευάσουμε δομές με βάση ασαφή σύνολα και τους λογικούς κανόνες AND OR και NOT. Βέβαια το γεγονός ότι βρήκαμε συναρτήσεις που να επεκτείνουν τους τελεστές αυτούς από το χώρο της Boolean λογικής στο χώρο της ασαφούς λογικής δε σημαίνει σε καμία περίπτωση ότι οι συναρτήσεις αυτές είναι και οι μοναδικές.

Έχουμε ορίσει την τομή, την ένωση και το συμπλήρωμα ενός ασαφούς συνόλου με τους τελεστές $\min$, $\max$, $1-A$ οι οποίοι είναι οι πιο συχνά χρησιμοποιούμενοι. Ωστόσο μπορούμε να χρησιμοποιήσουμε και άλλες συναρτήσεις για τον ίδιο σκοπό.

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

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

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

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

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

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

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

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

Παραδείγματα τ-νορμών είναι:

Συνήθης τομή $T(a,b)=\min(a,b)$

Αλγεβρικό γινόμενο $T(a,b)=ab$

Φραγμένη διαφορά $T(a,b)=\max(0,a+b-1)$

Συνάρτηση Hamacher Η συνάρτηση Hamacher δίνεται από τη σχέση:

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

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

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

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

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

Παραδείγματα σ-νορμών είναι τα εξής:

Συνήθης ένωση $s(a,b)=\max(a,b)$

Αλγεβρικό άθροισμα $s(a,b)=a+b-ab$

Φραγμένο άθροισμα $s(a,b)=\min(1,a+b)$

Συμπλήρωμα

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

$$\mu_{A'}(x) = c(\mu_A(x))$$

όπου η συνάρτηση c πρέπει να ικανοποιεί τις παρακάτω συνθήκες;

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

Το σύννηθες συμπλήρωμα δίνεται από τη σχέση:

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

Ασαφείς σχέσεις

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

$$R = \langle x \text{ είναι βαρύτερο από } y \rangle \quad x \in X, y \in Y \text{ και } R \in X \times Y$$

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

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

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

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

- **Αντιστροφή**

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

- **Σύνθεση**

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

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

Η sup-t σύνθεση $R: X \times Y \rightarrow [0,1]$ δύο ασαφών σχέσεων

$R_1: X \times Y \rightarrow [0,1]$ και $R_2: X \times Y \rightarrow [0,1]$ ορίζεται από την εξίσωση:

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

Οι περισσότεροι γνωστές μέθοδοι σύνθεσης ασαφών συνόλων είναι η σύνθεση max-min και η σύνθεση max-product.

Η συνάρτηση συμμετοχής για την περίπτωση της max-min σύνθεσης δίνεται από τη σχέση:

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

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

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

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

Ασαφής συλλογιστική

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

Οι συλλογιστικοί τρόποι που κυρίως χρησιμοποιούνται είναι τρεις:

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

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

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

όπου A και B συγκεκριμένα γεγονότα.

Ο παραπάνω κανόνας ερμηνεύεται ως εξής: Αν το γεγονός A συνεπάγεται το γεγονός B και επιπλέον έχουμε ως υπόθεση ότι ισχύει το A, τότε το συμπέρασμα που παίρνουμε είναι ότι ισχύει και το B. Όμως σε περιβάλλον ασάφειας τα γεγονότα

ισχύουν σε κάποιο βαθμό. Έτσι ο παραπάνω κανόνας πρέπει να τροποποιηθεί για να συμπεριλάβει και την ασάφεια. Καταλήγουμε μ' αυτό τον τρόπο στο γενικευμένο κανόνα *modus ponens* (*generalized modus ponens*) ο οποίος έχει την παρακάτω μορφή:

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

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

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

Ας εξετάσουμε τώρα την ερμηνεία των παραπάνω κανόνων. Για παράδειγμα ο γενικευμένος κανόνας *modus ponens* ερμηνεύεται ως εξής: Αν το γεγονός A συνεπάγεται το γεγονός B και έχουμε ως υπόθεση ότι ισχύει το A σε κάποιο βαθμό, τότε θα ισχύει και το B σε κάποιο βαθμό.

Πρέπει, λοιπόν να προσδιοριστεί ο βαθμός στον οποίο πληρείται το γεγονός B. Αυτός εξαρτάται από το βαθμό στον οποίο πληρείται το γεγονός A και από το είδος της συνεπαγωγής που εφαρμόζουμε. Η πράξη της ασαφούς συνεπαγωγής υλοποιεί τη μαθηματική σχέση $A \Rightarrow B$, όταν τα A και B είναι ασαφή γεγονότα..

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

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

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

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

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

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

3.3 Συστήματα ασαφούς συλλογιστικής-νευροασαφή συστήματα

Εισαγωγή

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

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

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

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

Κανόνες αν-τότε (if-then rules)

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

if x is A then y is B

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

if service is good then tip is average

Η λέξη **good** αντιπροσωπεύεται από έναν αριθμό που κυμαίνεται ανάμεσα στο 0 και το 1. Έτσι η υπόθεση είναι μια μετάφραση που επιστρέφει μια τιμή από το 0 έως το 1. Από την άλλη μεριά η λέξη **average** αποτελεί ένα ασαφές σύνολο και έτσι η συνέπεια του κανόνα είναι μια συσχέτιση της εξόδου y στο ασαφές σύνολο B.

Στους κανόνες if-then η λέξη 'is' έχει διαφορετική έννοια ανάλογα με το αν εμφανίζεται στο πρώτο ή στο δεύτερο μέρος του κανόνα. Έτσι όταν εμφανίζεται στο πρώτο μέρος του κανόνα έχει την έννοια του ελέγχου ισότητας (δηλαδή την ίδια έννοια που δίνουμε στο σύμβολο $=$). Αντίθετα όταν εμφανίζεται στο δεύτερο μέρος του κανόνα έχει την έννοια της καταχώρησης (δηλαδή την έννοια που δίνουμε στο σύμβολο $=$). Ο παραπάνω κανόνας λοιπόν θα μπορούσε να γραφεί συμβολικά και ως εξής:

if service == good then tip = average

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

1. Πρώτα πρέπει να υπολογίσουμε την είσοδο. Για να γίνει αυτό όμως πρέπει να γίνει ασαφοποίηση της εισόδου και να εφαρμοσθούν οι κατάλληλοι ασαφείς τελεστές
2. Να εφαρμόσουμε το παραπάνω αποτέλεσμα στο δεύτερο μέρος του κανόνα

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

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

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

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

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

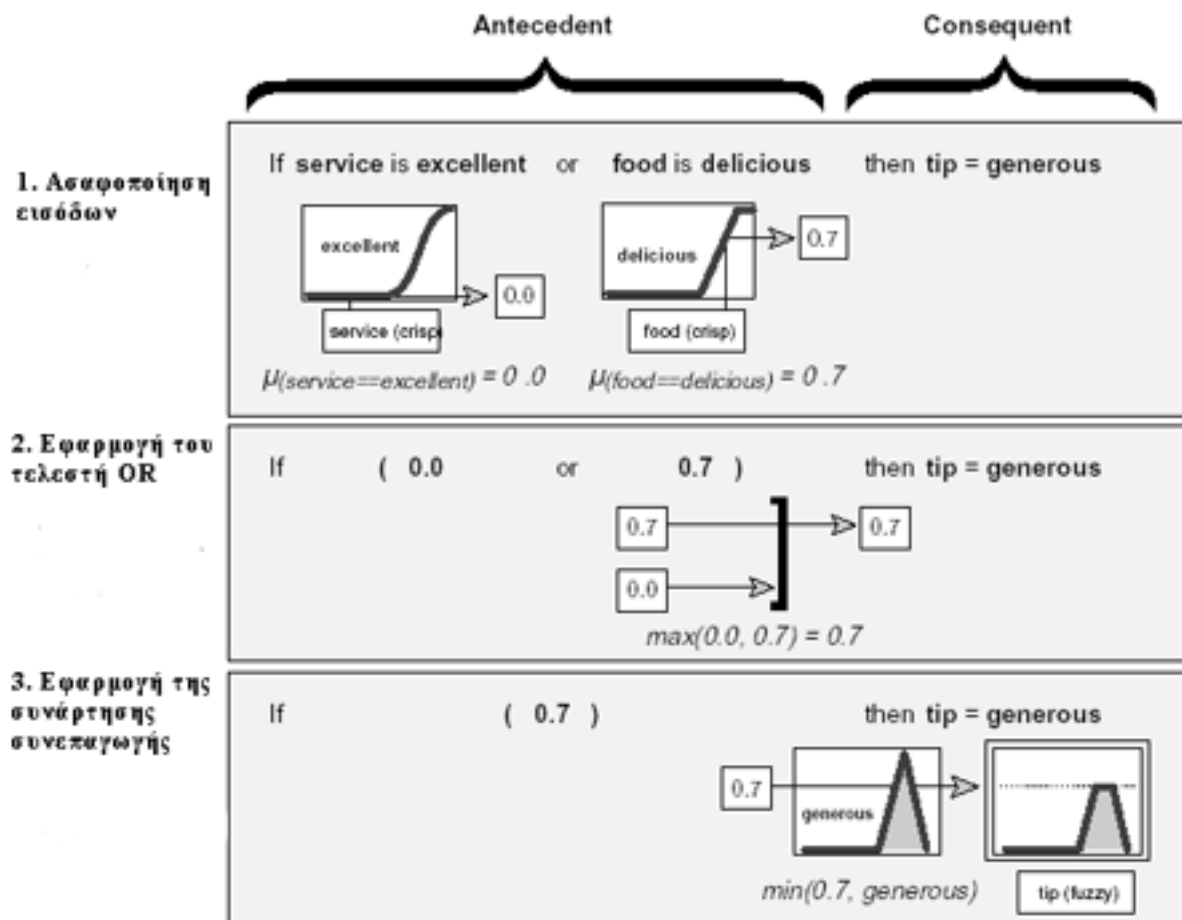
Σ' αυτή την περίπτωση όλα τα μέρη της προϋπόθεσης υπολογίζονται ταυτόχρονα και καταλήγουμε σε ένα μοναδικό νούμερο ανάλογα με τους λογικούς τελεστές που χρησιμοποιούμε.

Αλλά και η συνέπεια ενός κανόνα μπορεί να αποτελείται από περισσότερα του ενός μέρη. Για παράδειγμα ο παρακάτω κανόνας:

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

Όλα τα μέρη της συνέπειας του κανόνα επηρεάζονται το ίδιο από το αποτέλεσμα της προϋπόθεσης. Η συνέπεια του κανόνα προσδιορίζει ένα ασαφές σύνολο το οποίο συσχετίζεται με την έξοδο. Η συνάρτηση συνεπαγωγής (*implication function*) στη συνέχεια τροποποιεί το ασαφές σύνολο στο βαθμό που προσδιορίζεται από την προϋπόθεση του κανόνα.

Μια πολύ συνηθισμένη συνάρτηση για να γίνει αυτό είναι η αποκοπή με τη χρήση του τελεστή min. Στο παρακάτω σχήμα βλέπουμε τον τρόπο με τον οποίο εφαρμόζονται οι κανόνες if-then.



Συστήματα ασαφούς συλλογιστικής (fuzzy inference systems)

Γενικά ένας κανόνας από μόνος του δεν επαρκεί για τα περισσότερα πρακτικά προβλήματα. Αυτό που χρειαζόμαστε είναι περισσότεροι του ενός κανόνες οι οποίοι να αλληλεπιδρούν μεταξύ τους. Η έξοδος κάθε κανόνα θα είναι ένα ασαφές σύνολο. Τα ασαφή σύνολα που προκύπτουν σα συνέπεια των κανόνων συγκεντρώνονται για να σχηματίσουν ένα ασαφές σύνολο εξόδου από το οποίο μέσω της διαδικασίας της αποσαφήνισης θα πάρουμε μια μοναδική τιμή.

Παρακάτω αναπτύσσεται ο τρόπος με τον οποίο εξελίσσεται η παραπάνω διαδικασία μέσα από ένα σύστημα ασαφούς συλλογιστικής (fuzzy inference system ,FIS) που ονομάζεται Mamdani.

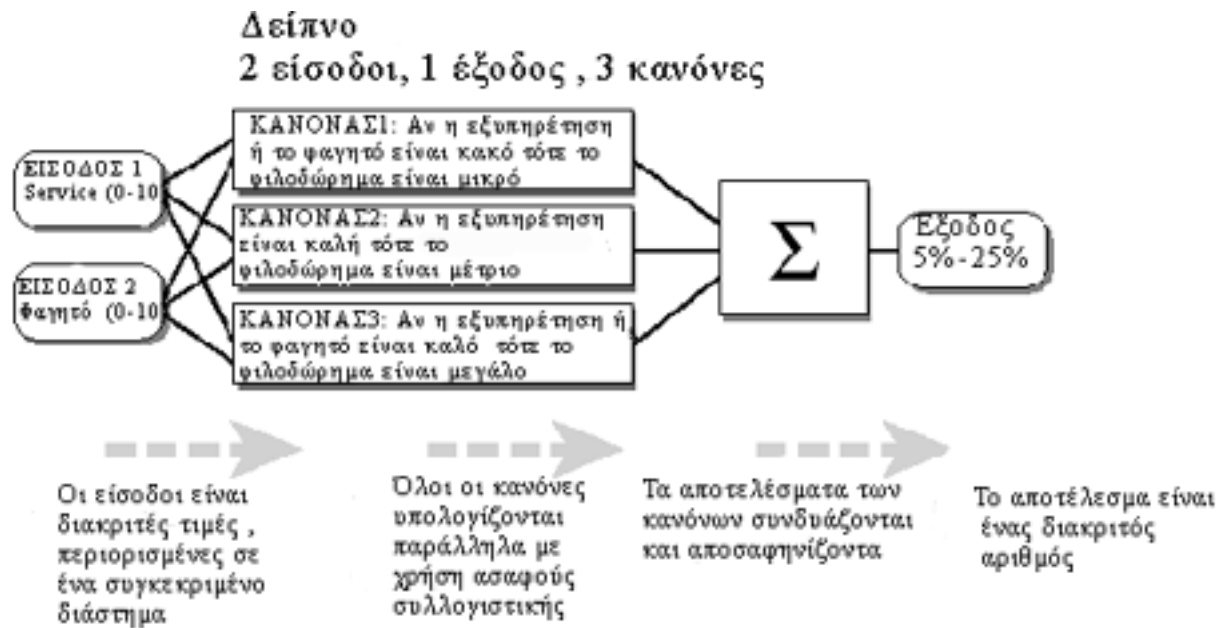
Συστήματα Mamdani

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

Θα εξετάσουμε τώρα πως λειτουργεί ένα FIS τύπου Mamdani μέσα από το πρόβλημα του φιλοδώρηματος που αναφέραμε παραπάνω .

Το σύστημα έχει δύο εισόδους και τρεις κανόνες. Οι είσοδοι του συστήματος είναι η ποιότητα του φαγητού και η εξυπηρέτηση. Θεωρούμε ότι ο πελάτης βαθμολογεί την ποιότητα του φαγητού και την εξυπηρέτηση με μια κλίμακα από το 0 έως το 10. η έξοδος του συστήματος είναι το φιλοδώρημα που θα δώσει ο πελάτης και το οποίο κυμαίνεται από 5% έως 25% της τιμής του λογαριασμού.

Η βασική του δομή του συστήματος φαίνεται στο παρακάτω διάγραμμα:



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

1. Ασαφοποίηση των εισόδων
2. Εφαρμογή των ασαφών λογικών τελεστών στις προϋπόθεσης των κανόνων
3. Καθορισμός των συνεπαγωγών των κανόνων
4. Συγκέντρωση όλων των συνεπαγωγών όλων των κανόνων
5. Αποσαφήνιση

Θα εξετάσουμε κάθε βήμα από τα παραπάνω ξεχωριστά.

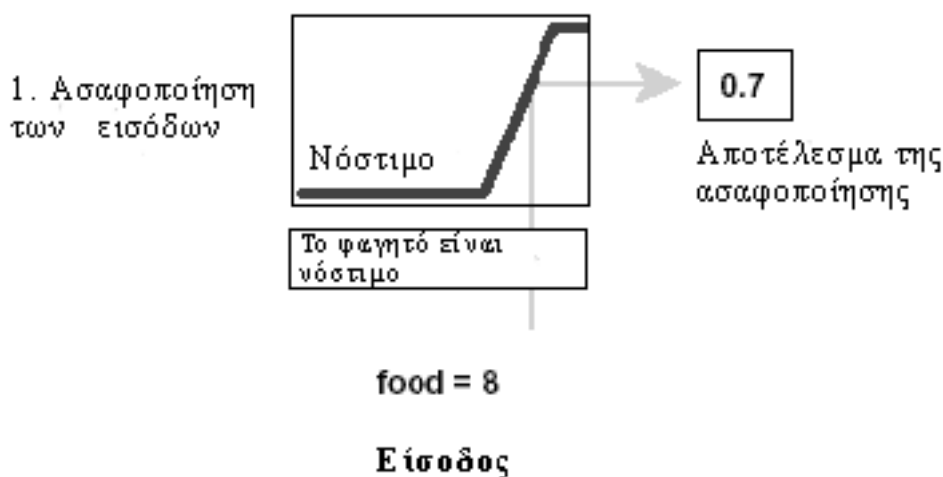
ΒΗΜΑ 1^ο : Ασαφοποίηση εισόδων

Το πρώτο βήμα έχει να κάνει με την ασαφοποίηση των εισόδων δηλαδή αυτό που κάνουμε είναι να πάρουμε τις εισόδους και να καθορίσουμε σε ποιο βαθμό κάθε είσοδος ανήκει σε ένα ασαφές σύνολο μέσα από τις συναρτήσεις συμμετοχής. Οι εισοδοι είναι αριθμητικές τιμές μέσα στα όρια του χώρου αναφοράς (στην περίπτωση

του προβλήματος που εξετάζουμε από 0 μέχρι 10) και οι έξοδοι είναι βαθμοί συμμετοχής στο προσδιορισμένο ασαφές σύνολο.

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

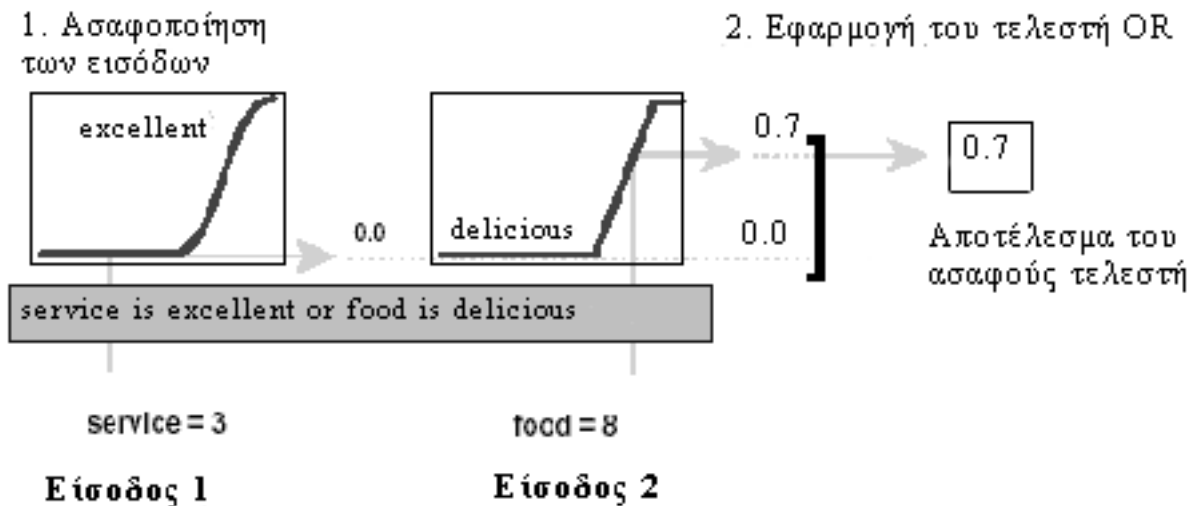
Πριν να εφαρμοσθούν οι κανόνες θα πρέπει να ασαφοποιηθούν οι είσοδοι σύμφωνα με τις λεκτικές μεταβλητές. Για παράδειγμα σε ποιο βαθμό το φαγητό είναι νόστιμο. Αν υποθέσουμε ότι το φαγητό έχει αξιολογηθεί με μια συγκεκριμένη βαθμολογία από το 0 έως το 10 (π.χ. 7) τότε πώς μπορούμε να συμπεράνουμε σε ποιο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



ΒΗΜΑ 2^ο : Εφαρμογή των ασαφών τελεστών

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

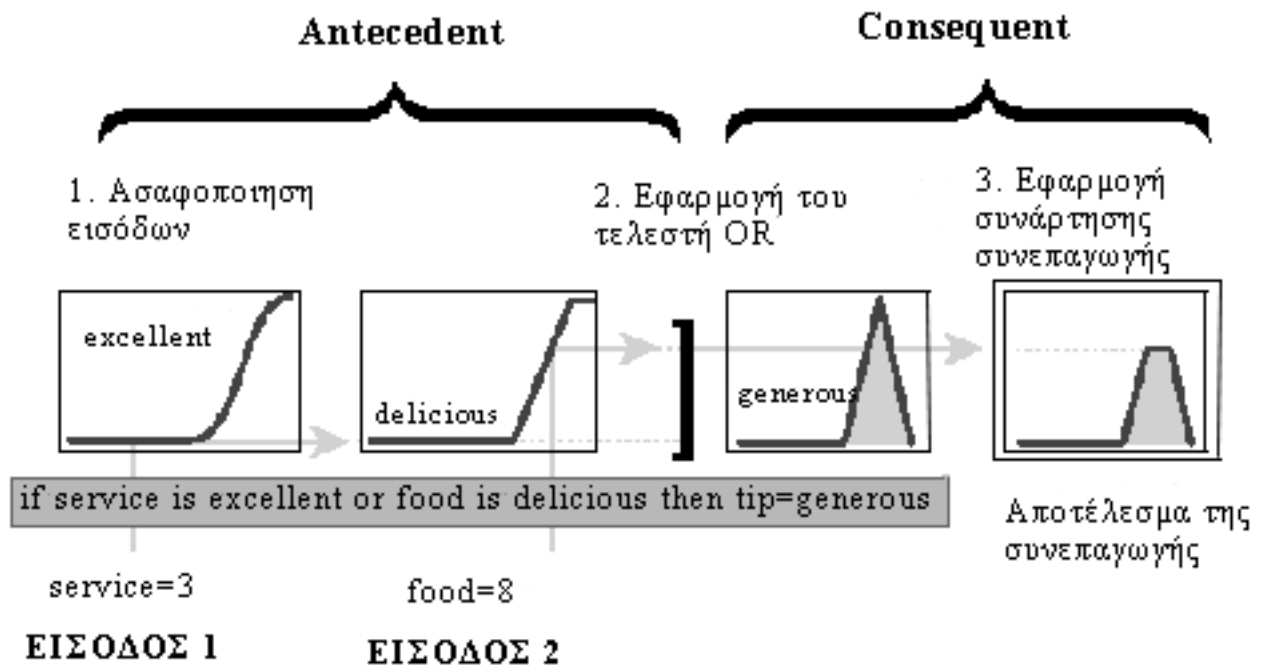
Όπως είδαμε κατά την αναφορά στους ασαφείς τελεστές υπάρχει ένας μεγάλος αριθμός συναρτήσεων που μπορούν να χρησιμοποιηθούν για να ορίσουν τους τελεστές AND και OR. Συνήθως χρησιμοποιούνται ο τελεστής max για το OR και min για το AND. Στο παρακάτω διάγραμμα βλέπουμε πως εφαρμόζονται οι τελεστές.



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

Για να εφαρμόσουμε την συνάρτηση συνεπαγωγής πρέπει να γνωρίζουμε το βάρος του κάθε κανόνα. Οι κανόνες if-then έχουν βάρη τα οποία εφαρμόζονται στο νούμερο που δίνει το κομμάτι της υπόθεσης και που μπορούν να κυμαίνονται από 0 έως 1. Συνήθως τα βάρη αυτά είναι ίσα με 1, ωστόσο μπορούμε να δώσουμε σ' αυτά μια οποιαδήποτε τιμή από 0 έως 1. Για παράδειγμα αν θέλουμε να δώσουμε μικρότερη έμφαση σε έναν κανόνα σε σχέση με κάποιον άλλο, αυτό μπορούμε να το κάνουμε μέσω των τιμών των βαρών τους.

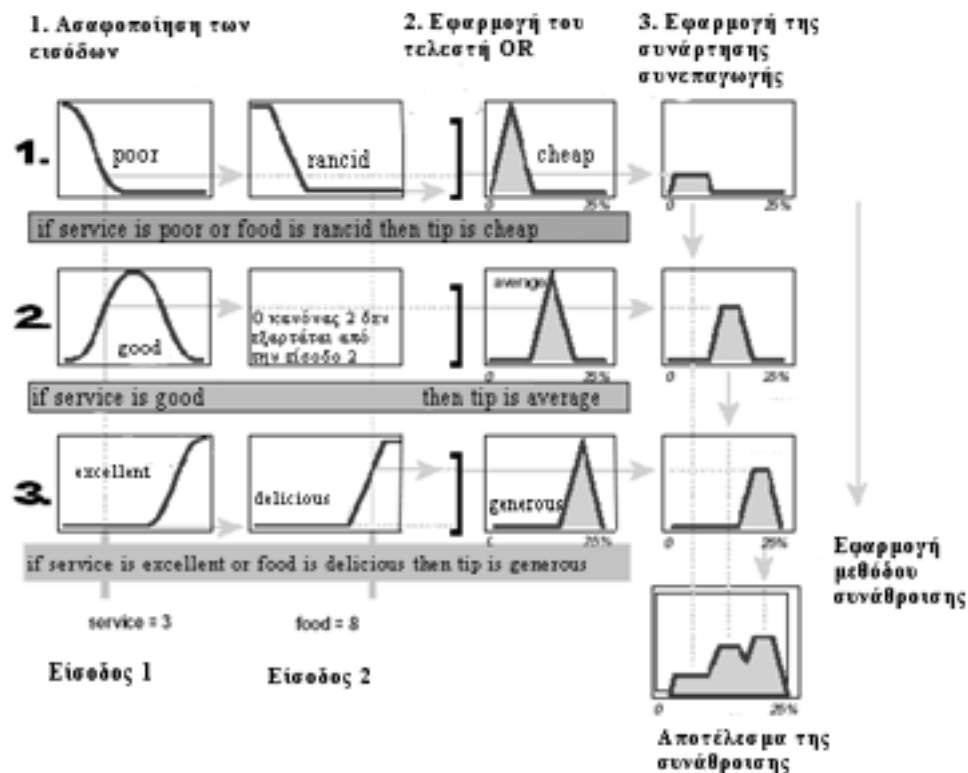
Αφού έχουμε αναθέσει στους κανόνες τα βάρη μπορούμε να εφαρμόσουμε την συνάρτηση συνεπαγωγής. Το αποτέλεσμα ενός κανόνα if-then είναι ένα ασαφές σύνολο που αντιπροσωπεύεται από μια συνάρτηση συμμετοχής που σταθμίζει κατάλληλα τα λεκτικά χαρακτηριστικά που έχουν ανατεθεί σ' αυτόν. Το αποτέλεσμα αυτό αναδιαμορφώνεται με τη χρήση μιας συνάρτησης που είναι συνδεδεμένη με το μέρος της υπόθεσης. Η είσοδος για τη διαδικασία συνεπαγωγής είναι ένας αριθμός που δίδεται από την υπόθεση και η έξοδος είναι ένα ασαφές σύνολο. Αυτή η διαδικασία εφαρμόζεται για κάθε κανόνα.



ΒΗΜΑ 4^ο : Συγκέντρωση όλων των εξόδων

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

Στο παρακάτω διάγραμμα βλέπουμε τους τρεις κανόνες του προβλήματος και τον τρόπο που συνδυάζονται τα αποτελέσματά τους σε ένα ασαφές σύνολο.



ΒΗΜΑ 5^ο : Αποσαφήνιση

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

Μέθοδος αποσαφήνισης MAXIMUM

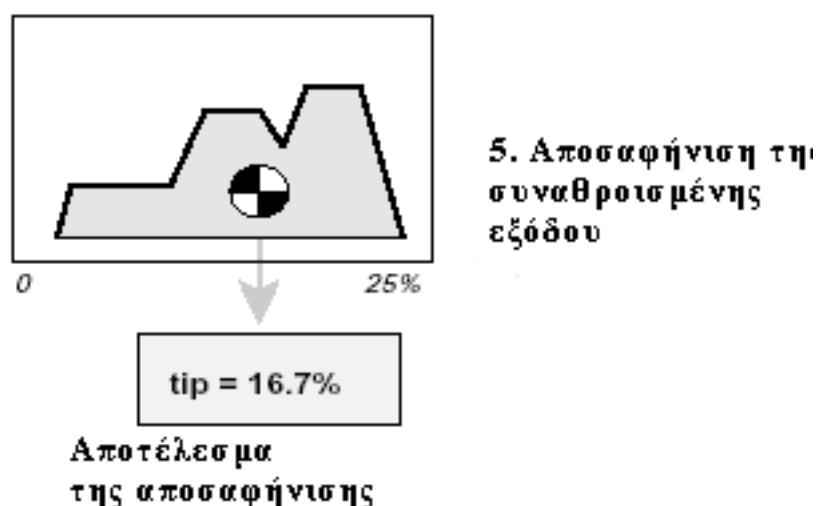
Σύμφωνα με τη μέθοδο αποσαφήνισης maximum, η διακριτή τιμή είναι αυτή που αντιστοιχεί στη μέγιστη τιμή συγγένειας του τελικού αποτελέσματος. Αν υπάρχουν περισσότερες από μία τέτοιες τιμές, τότε λαμβάνεται ανάλογα με την περίπτωση είτε ο μέσος όρος τους (average-of-maximum) ή η μέγιστη τιμή τους (maximum-of-maximum)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

Μέθοδος Cendroid

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

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

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



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

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

Διαγραμματική μέθοδος επίλυσης

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

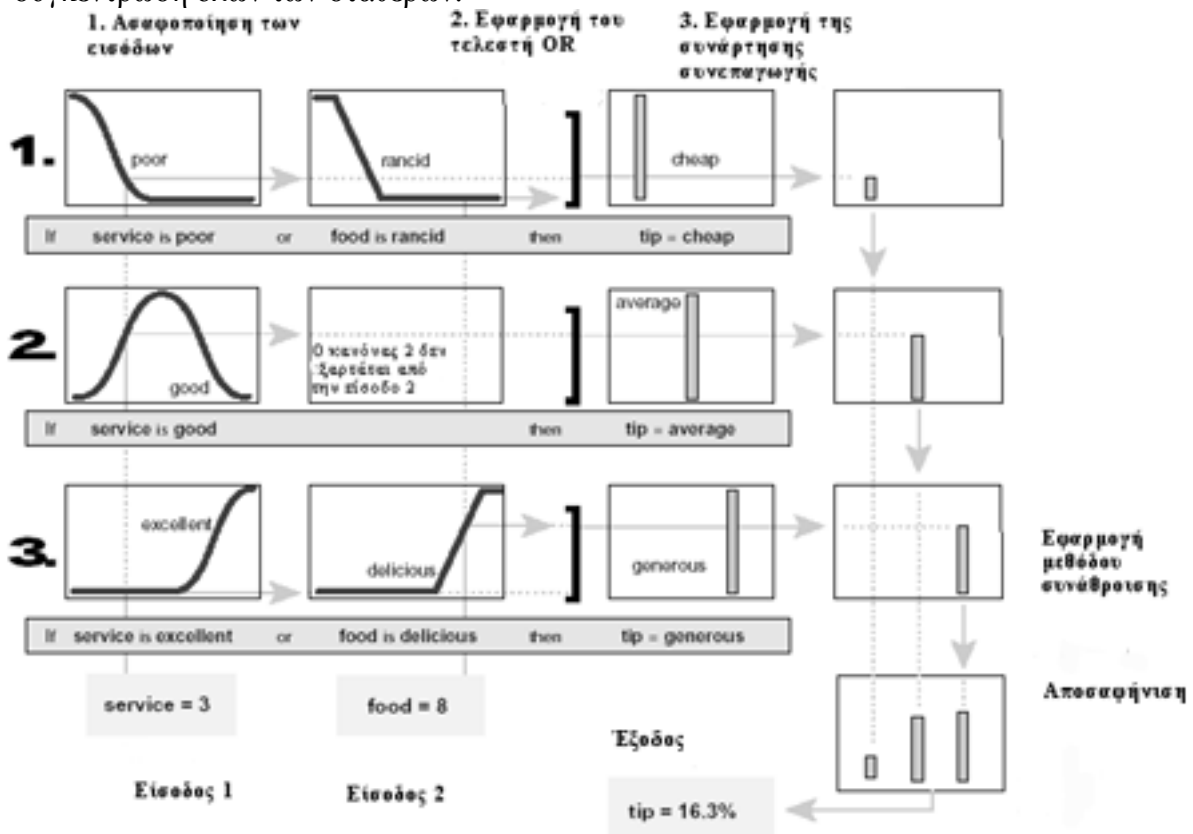
Συστήματα τύπου Sugeno

Παραπάνω ασχοληθήκαμε με τα συστήματα Mamdani που είναι και τα πιο ευρέως χρησιμοποιούμενα. Ωστόσο υπάρχει και η μέθοδος Sugeno που εισήχθηκε το 1985 η μέθοδος αυτή έχει ομοιότητες με τη μέθοδο των συστημάτων Mamdani. Για παράδειγμα τα 2 πρώτα βήματά της (ασαφοποίηση των εισόδων και εφαρμογή των τελεστών) είναι ακριβώς τα ίδια. Η κύρια διαφορά ανάμεσα στα δύο συστήματα έγκειται στο ότι οι συναρτήσεις συμμετοχής στην έξοδο των συστημάτων Sugeno είναι μόνο γραμμική ή σταθερά.

Ένας τυπικός ασαφής κανόνας σε συστήματα sugeno μηδενικής τάξης έχει την μορφή:

$$\text{if } x \text{ is } A \text{ and } y \text{ is } B \text{ then } z = k$$

όπου A και B είναι τα ασαφή σύνολα της προϋπόθεσης ενώ k είναι μια αριθμητική τιμή. Αφού η συνέπεια του κανόνα είναι μια σταθερά τότε αυτό σημαίνει ότι το βήμα 3 εκφυλίζεται σε ένα απλό πολλαπλασιασμό ενώ το βήμα 4 καταλήγει να είναι η συγκέντρωση όλων των σταθερών.



Ένα σύστημα sugeno πρώτης τάξης θα έχει κανόνες με τη γενική μορφή

$$\text{if } x \text{ is } A \text{ and } y \text{ is } B \text{ then } z = p*x + q*y + r$$

όπου A και B είναι τα ασαφή σύνολα της προϋπόθεσης ενώ τα p,q,r είναι σταθερές.

Ένας τρόπος για να δούμε τα συστήματα πρώτης τάξης είναι να θεωρήσουμε ότι κάθε κανόνας προσδιορίζει τη θέση ενός κινούμενου singleton. Το singleton αυτό μπορεί να κινείται στο χώρο της εξόδου, με γραμμικό τρόπο και η θέση του εξαρτάται από τις τιμές των εισόδων.

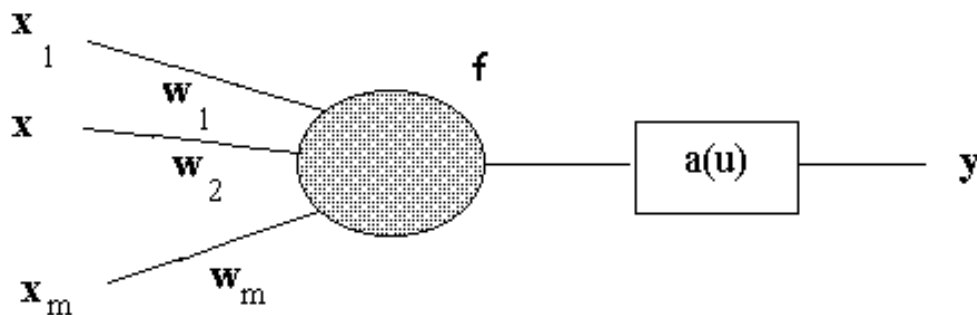
Συστήματα sugeno ανώτερης τάξης είναι εφικτά, όμως δεν προσφέρουν σημαντικές βελτιώσεις και ταυτόχρονα εισαγάγουν σημαντική πολυπλοκότητα.

Νευροασαφή συστήματα

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

ΑΣΑΦΕΙΣ ΝΕΥΡΩΝΕΣ

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



Όπου :

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

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

ΝΕΥΡΩΝΕΣ ΣΥΜΜΕΤΟΧΗΣ

Οι συνηθέστερες μορφές συναρτήσεων συμμετοχής είναι η τριγωνική, η γκαουσιανή, η τραπεζοειδής, η κανονική κοκ. Οι παραπάνω συναρτήσεις συμμετοχής μπορούν να υλοποιηθούν με τη χρήση ενός νευρώνα. Η συνάρτηση συμμετοχής μπορεί να υλοποιηθεί θεωρώντας ότι ο νευρώνας δέχεται ως είσοδο το x

και δίνει έξοδο $A(x)$ απαιτώντας η συνάρτηση συμμετοχής του να έχει τη μορφή A . αν δηλαδή υποθέσουμε ότι η συνάρτηση συμμετοχής είναι η κανονική , τότε για να υλοποιηθεί πρέπει η συνάρτηση ενεργοποίησης να είναι

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

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

ΛΕΙΤΟΥΡΓΙΚΟΙ ΝΕΥΡΩΝΕΣ

Με τον ίδιο τρόπο μπορούν να υλοποιηθούν οι βασικές πράξεις των ασαφών συνόλων. Οι πράξεις αυτές στηρίζονται σε λειτουργίες όπως η τ - νόρμα η σ -νόρμα , το ασαφές συμπλήρωμα κλπ. Οι παραπάνω λειτουργίες μπορούν να αντικαταστήσουν τη συνάρτηση ενεργοποίησης του νευρώνα αν θεωρήσουμε ότι οι είσοδοι δεν αθροίζονται πολλαπλασιασμένοι με βάρη όπως στους κλασσικούς νευρώνες. Για να μπορούμε όμως να εφαρμόσουμε τους αλγορίθμους μάθησης (πχ τον backpropagation αλγόριθμος αντίστροφης διάδοσης) θα πρέπει οι λειτουργίες να είναι παραγωγίσιμες.

Για την υλοποίηση των ασαφών λειτουργιών μπορούμε να επεκτείνουμε την έννοια της συνάρτησης μεταφοράς του νευρώνα. Στην περίπτωση αυτή ο νευρώνας δεν υλοποιεί το αναλυτικό γινόμενο της εισόδου με το διάνυσμα των βαρών αλλά την πράξη της συγκεκριμένης ασαφούς λειτουργίας θεωρώντας ότι όλα τα βάρη είναι 1. ως συνάρτηση συμμετοχής λαμβάνεται η :

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

ΣΥΝΘΕΤΙΚΟΙ ΝΕΥΡΩΝΕΣ

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

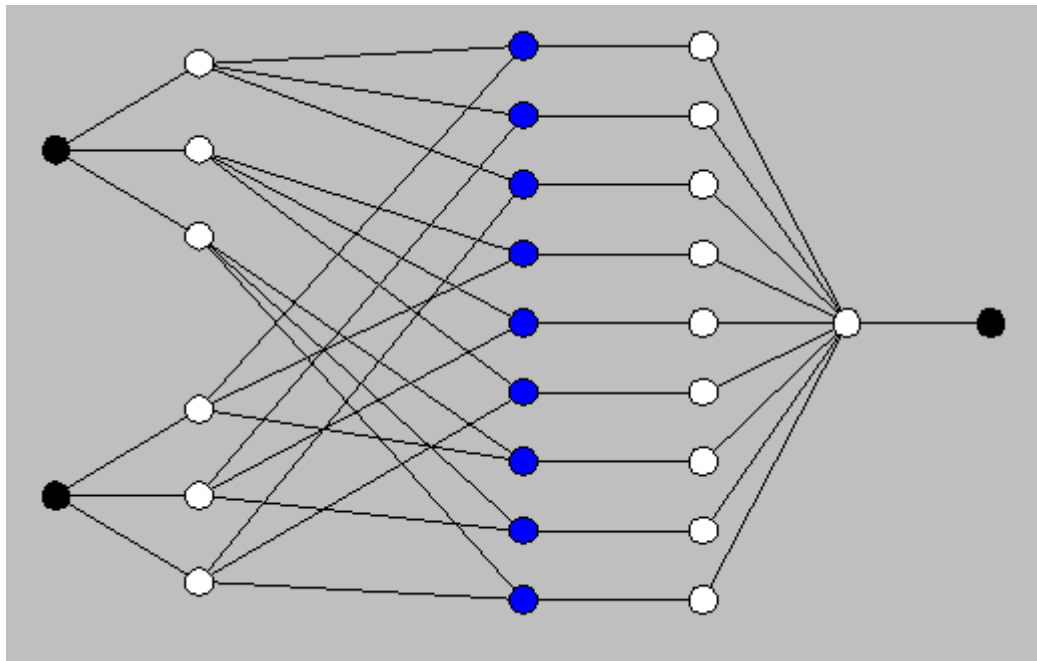
$$f = \underset{i \in N_m}{union} t(x_i, w_i)$$

$$y = a(u)$$

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

ANFIS (Adaptive Neuro-Fuzzy Inference System)

Στην παράγραφο αυτή θα ασχοληθούμε με τον τρόπο με τον οποίο μπορούμε να κατασκευάσουμε ένα νευροασαφές σύστημα, το οποίο προσαρμόζει τα βάρη του έτσι ώστε να αναπαριστά ένα άγνωστο σύστημα από ένα σύνολο δεδομένων. Συγκεκριμένα θα εστιάσουμε στο anfis, το οποίο είναι ένα νευροασαφές δίκτυο που προτάθηκε από τον Jang το 1992, και το οποίο θα χρησιμοποιήσουμε για την βραχυπρόθεσμη πρόβλεψη πωλήσεων. Το δίκτυο αυτό στηρίζεται στην πολυποθετική προσεγγιστική συλλογιστική και στην εξαγωγή συμπεράσματος που προτείνεται από τους Sugeno-Takagi. Η δομή του anfis φαίνεται στο παρακάτω διάγραμμα:



Η μορφή του κανόνα j που αναπαριστά το δίκτυο είναι η εξής:

$$AN \ x_1 \ EINAI \ A_1^{(j)} \ KAI \ x_2 \ EINAI \ A_2^{(j)} \ KAI \dots KAI \ x_d \ EINAI \ A_d^{(j)} \ TOTE:$$

$$y = f_j(x_1, x_2, \dots, x_d) = c_0^j + c_1^j \cdot x_1 + \dots + c_d^j \cdot x_d$$

Ας θεωρήσουμε ένα παράδειγμα συστήματος που αποτελείται από δύο εισόδους x_1 και x_2 και 2 κανόνες:

KANONΑΣ 1

$$AN \ x_1 \ EINAI \ A_1^{(1)} \ KAI \ x_2 \ EINAI \ A_2^{(2)} \ TOTE:$$

$$y = f_1(x_1, x_2) = c_0^1 + c_1^1 \cdot x_1 + c_2^1 \cdot x_2$$

KANONΑΣ 2

$$AN \ x_1 \ EINAI \ A_1^{(2)} \ KAI \ x_2 \ EINAI \ A_2^{(2)} \ TOTE:$$

$$y = f_2(x_1, x_2) = c_0^2 + c_1^2 \cdot x_1 + c_2^2 \cdot x_2$$

Όταν η είσοδος του συστήματος είναι $[x_1, x_2]$ τότε η έξοδος του θα δίδεται από τη σχέση:

$$y = \frac{r_1 \cdot f_1 + r_2 \cdot f_2}{r_1 + r_2} = r_1' \cdot f_1 + r_2' \cdot f_2$$

Όπου r_1, r_2 είναι οι βαθμοί συσχέτισης των εισόδων με τις υποθέσεις των κανόνων. Το δίκτυο αποτελείται από 6 επίπεδα. Παρακάτω αναλύονται οι λειτουργίες και τα χαρακτηριστικά κάθε επιπέδου.

Επίπεδο 1

Κάθε κόμβος του επιπέδου αυτού περνάει τις εισόδους στο επόμενο επίπεδο.

Επίπεδο 2

Το δεύτερο επίπεδο αποτελείται από νευρώνες συμμετοχής. Σε κάθε κόμβο αναπαριστάται μια συνάρτηση συμμετοχής ενός ασαφούς συνόλου που ορίζεται στο πεδίο ορισμού κάθε εισόδου. Οι νευρώνες αυτού του επιπέδου διαχωρίζονται σε υποκατηγορίες κάθε μία από τις οποίες απεικονίζει την ασαφή διαμέριση του πεδίου ορισμού της αντίστοιχης εισόδου. Στην περίπτωση του προηγούμενου παραδείγματος έχουμε δύο υποκατηγορίες νευρώνων συμμετοχής. Που αναπαριστούν τα ασαφή υποσύνολα $A_1^{(1)}, A_2^{(2)}, A_1^{(2)}, A_2^{(2)}$. Αν οι συναρτήσεις συμμετοχής έχουν τη μορφή καμπάνας τότε :

$$A_i^j(x_i) = \exp \left\{ - \left[\left(\frac{x_i - m_i^j}{\sigma_i^j} \right)^2 \right]^{b_i^j} \right\}$$

όπου m_i^j, σ_i^j, b_i^j οι παράμετροι των συναρτήσεων συμμετοχής που πρέπει να καθοριστούν κατά τη διάρκεια της μάθησης. Αυτές οι παράμετροι ονομάζονται παράμετροι υποθέσεων. Για να μπορούν να χρησιμοποιηθούν κλασσικοί αλγόριθμοι μάθησης θα πρέπει οι συναρτήσεις συμμετοχής που χρησιμοποιούνται να είναι συνεχείς και κατά διαστήματα παραγωγίσιμες.

Επίπεδο 3

Οι νευρώνες που συνιστούν αυτό το επίπεδο είναι λειτουργικοί νευρώνες που υλοποιούν την πράξη της τ-νόρμας. Κάθε νευρώνας αντιστοιχεί σε έναν κανόνα. Οι εισοδοί των νευρώνων αυτών είναι οι τιμές των συναρτήσεων συμμετοχής των υποθέσεων του κανόνα, που προέρχονται από τις εξόδους του επιπέδου 2. η έξοδος κάθε νευρώνα του επιπέδου 3 είναι ο βαθμός ενεργοποίησης του κανόνα.

Επίπεδο 4

Το επίπεδο αυτό περιέχει τον ίδιο αριθμό νευρώνων με το προηγούμενο και υπολογίζει τους κανονικοποιημένους βαθμούς ενεργοποίησης των κανόνων.

Επίπεδο 5

Κάθε νευρώνας j του επιπέδου αυτού υπολογίζει το συμπέρασμα του j κανόνα . Το επίπεδο αυτό έχει 2 κόμβους και 3 παραμέτρους για κάθε κόμβο. Οι παράμετροι αυτοί ονομάζονται παράμετροι συμπερασμάτων.

Επίπεδο 6

Στο επίπεδο αυτό υπάρχει ένας μόνο νευρώνας ο οποίος αθροίζει χωρίς βάρη τις εισόδους και δίδει το τελικό αποτέλεσμα στην έξοδο.

3.4. ANFIS (Adaptive Neuro-Fuzzy Inference System)

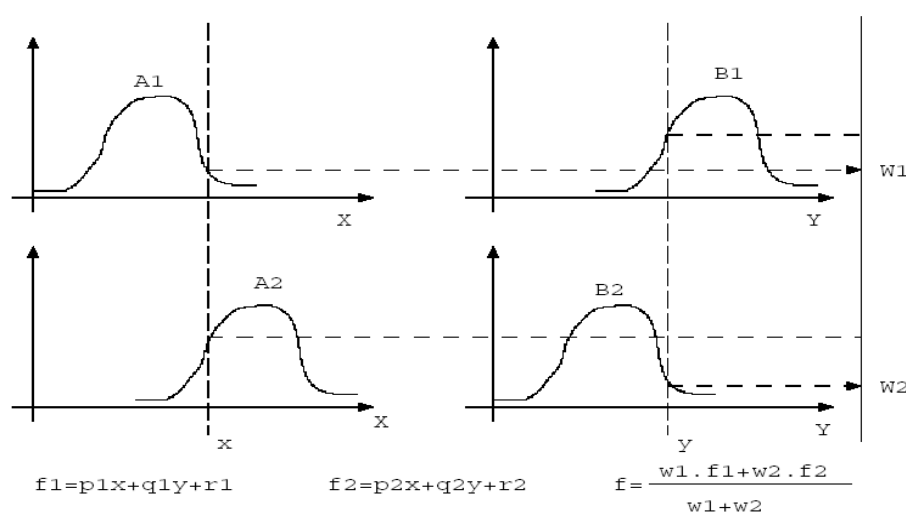
Εισαγωγή

Στο κεφάλαιο αυτό θα ασχοληθούμε με την ανάλυση του μοντέλου ANFIS που χρησιμοποιήσαμε στην έρευνά μας. Αρχικά θα δώσουμε έναν ορισμό του μοντέλου και θα παρουσιάσουμε τη δομή του. Στη συνέχεια θα εξετάσουμε κάθε επίπεδο του μοντέλου ξεχωριστά και θα παρουσιάσουμε τι είδους υπολογισμοί γίνονται σε κάθε επίπεδο. Τέλος, θα γίνει μια λεπτομερής περιγραφή του ANFIS και θα παρουσιάσουμε ένα αριθμητικό παράδειγμα όπου θα φαίνονται αναλυτικά όλοι οι υπολογισμοί που πραγματοποιεί το σε κάθε επίπεδο, έτσι ώστε να γίνει πλήρως κατανοητή η λειτουργία του μοντέλου.

Ορισμός

Ένα νευρο-ασαφές σύστημα ορίζεται ως ένας συνδυασμός τεχνητών νευρωνικών δικτύων (ANN) και Ασαφούς Συστήματος Συμπερασμού (Fuzzy Inference System, FIS) κατά τέτοιο τρόπο ώστε ο αλγόριθμος εκμάθησης νευρωνικών δικτύων χρησιμοποιείται για να καθορίσει τις παραμέτρους του FIS. Το Προσαρμοστικό Νευροασαφές Σύστημα Συμπερασμού (Adaptive Neural Fuzzy Inference System, ANFIS) είναι ένα σύστημα που ανήκει στην νευρο-ασαφή κατηγορία. Λειτουργικά, δεν υπάρχει σχεδόν κανένας περιορισμός στις συναρτήσεις κόμβων ενός προσαρμόσιμου (adaptive) δικτύου εκτός από τμήματα που μπορούν να διαφοριστούν (piecewise differentiability). Δομικά, ο μόνος περιορισμός της διαμόρφωσης δικτύων είναι ότι πρέπει να είναι τύπου πρόσθιας τροφοδότησης. Λόγω αυτού του μικρού περιορισμού, οι εφαρμογές του προσαρμόσιμου δικτύου είναι άμεσες και απέραντες σε διάφορους τομείς. Σε αυτό το τμήμα, προτείνουμε μια κατηγορία προσαρμόσιμων δικτύων, τα οποία είναι λειτουργικά ισοδύναμα με τα ασαφή συστήματα συμπερασμού.

Ο ασαφής μηχανισμός συμπερασμού παρουσιάζεται στο επόμενο σχήμα:



(σχήμα 3.4.1): Ασαφής Μηχανισμός Συμπερασμού

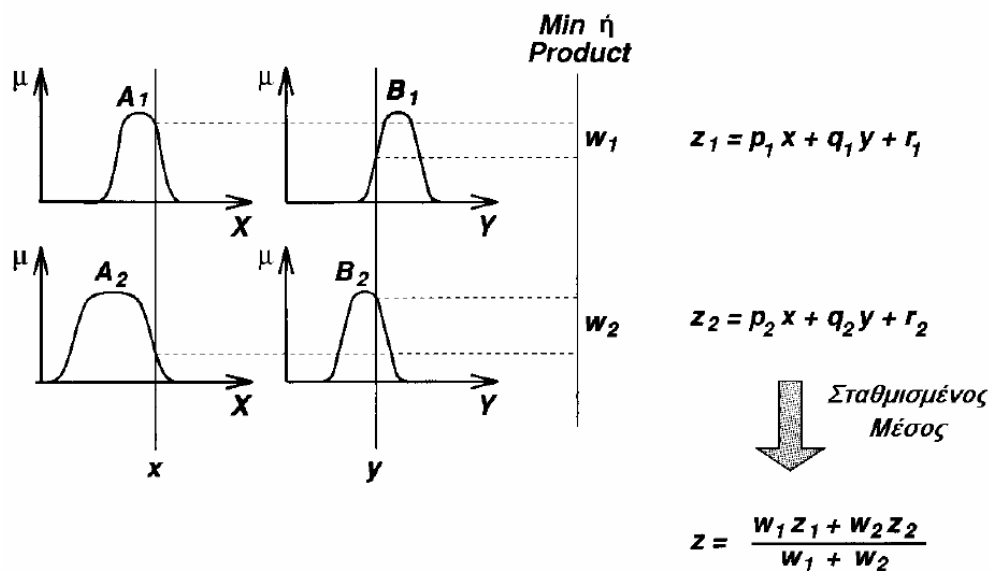
Δομή Αρχιτεκτονικής του ANFIS

Ο αλγόριθμος ANFIS είναι από τους κυριότερους αλλά και ταυτόχρονα από τους πρώτους που εφαρμόστηκαν στο πεδίο της νευροασαφούς προσέγγισης προβλημάτων. Θα παρουσιάσουμε το δίκτυο υποθέτοντας ότι το πρόβλημα το οποίο έχουμε να αναλύσουμε έχει δύο εισόδους x και y και μία έξοδο z . Υποθέτοντας ότι έχουμε ένα πρώτης τάξης μοντέλο sugeno, μία τυπική βάση κανόνων (rule base) θα μπορούσε να είναι η εξής:

Rule 1: If x is A_1 and y is B_1 then $f_1 = p_1x + q_1y + r_1$

Rule 2: If x is A_2 and y is B_2 then $f_2 = p_2x + q_2y + r_2$

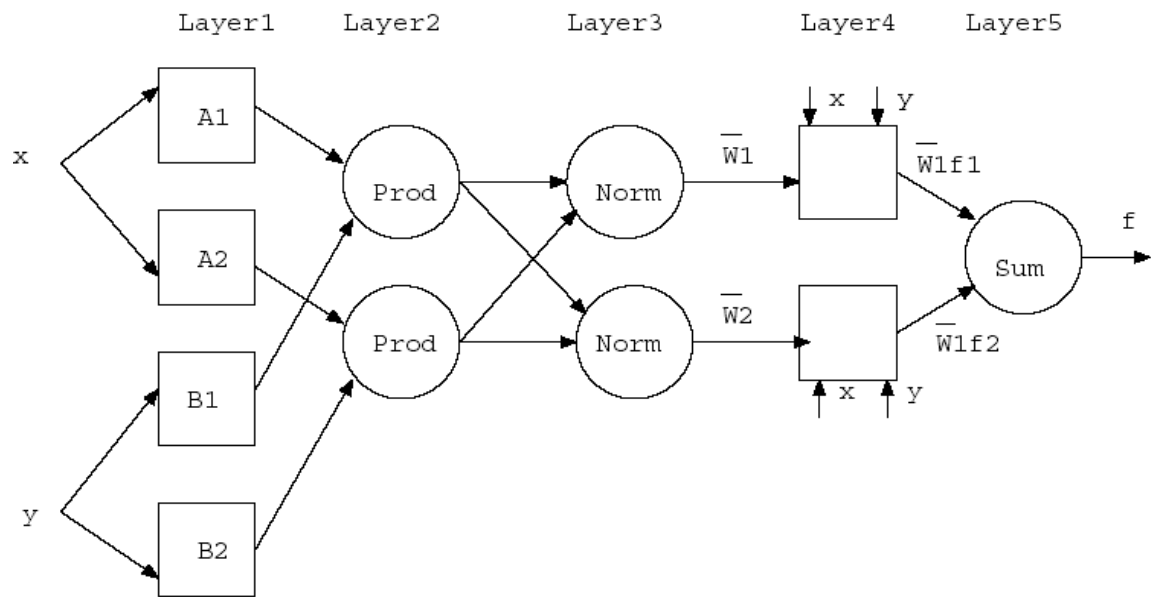
Το παρακάτω σχήμα δείχνει με απλό τρόπο τη διαδικασία συμπερασμού (inference procedure) του μοντέλου sugeno, στην περίπτωση όπου για t-operator έχουμε επιλέξει την τομή των δύο ασαφών συνόλων (A, B), οπότε $\mu_{A \cap B}(x) = \min[\mu_A(x), \mu_B(x)]$



(σχήμα 3.4.2): Ασαφής Συλλογιστική (fuzzy reasoning)

Όπως γίνεται φανερό, η έξοδος z του πρωτοβάθμιου μοντέλου Sugeno είναι ένας σταθμικός μέσος όρος.

Η αντίστοιχη αναπαράσταση του δικτύου ANFIS παρουσιάζεται στο παρακάτω σχήμα:



(σχήμα 3.4.3): Αρχιτεκτονική δομή ANFIS

Ο κάθε κόμβος και το κάθε επίπεδο του ANFIS έχει μία ιδιαίτερη χρήση, την οποία εξηγούμε παρακάτω.

Επίπεδο 1: Οι κόμβοι του πρώτου επιπέδου εμφανίζουν τη συμπεριφορά που χαρακτηρίζει τους νευρώνες των νευρωνικών δικτύων. Σε αυτό το επίπεδο είναι ένας προσαρμόσιμος κόμβος με μία συνάρτηση κόμβου:

$$O_i^1(x) = \mu_{A_i}(x)$$

όπου x είναι η είσοδος στον κόμβο i και A_i είναι η γλωσσική ταμπέλα (small, large, κλπ.) που σχετίζεται με αυτή τη συνάρτηση του κόμβου.

Με άλλα λόγια, το O_i είναι ο βαθμός συμμετοχής του A_i και καθορίζει το βαθμό στον οποίο η είσοδος x ικανοποιεί τον ποσοτικοποιητή A_i . Συνήθως επιλέγεται το $\mu_{A_i}(x)$ σε σχήμα καμπάνας με το μέγιστο ίσο με 1 και το ελάχιστο ίσο με 0, όπως η γενικευμένη καμπανοειδής συνάρτηση:

$$\mu_{A_i}(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}}$$

όπου a_i, b_i, c_i είναι το σύνολο παραμέτρων.

Καθώς οι τιμές αυτών των παραμέτρων αλλάζουν, οι συναρτήσεις ποικίλλουν ανάλογα, παρουσιάζοντας έτσι διάφορες μορφές της συνάρτησης συμμετοχής για τη γλωσσική ταμπέλα. Οι παράμετροι σε αυτό το επίπεδο αναφέρονται ως *αρχικοί παράμετροι* (premise parameters) A_i .

Επίπεδο 2: Κάθε κόμβος στο επίπεδο αυτό είναι ένας πολλαπλασιαστής (για αυτό και συμβολίζεται με το γράμμα Π) και έχει ως ρόλο του την εκτέλεση του πολλαπλασιασμού των σημάτων εξόδου του πρώτου επιπέδου:

$$O_i^2 = w_i = \mu_{A_i}(x) * \mu_{B_i}(y), i=1,2$$

Η κάθε έξοδος αυτού του επιπέδου αντιστοιχεί στην ισχύ ενεργοποίησης (firing strength) του κάθε κανόνα. Πρέπει να αναφέρουμε ότι εκτός του γινομένου μπορούμε να χρησιμοποιήσουμε οποιαδήποτε πράξη από τις αποδεκτές της ασαφούς λογικής που υλοποιούν την πράξη AND.

Επίπεδο 3: Κάθε κόμβος στο τρίτο επίπεδο επιτελεί με τη σειρά του τη λειτουργία που δεν επιδέχεται ρυθμίσεις. Ο i -κόμβος του επιπέδου υπολογίζει απλώς το πηλίκο της ισχύος του i -κανόνα ως προς το άθροισμα όλων των άλλων firing strengths που αντιστοιχούν στους:

$$\bar{w}_i = \frac{w_i}{w_1 + w_2}, i=1,2$$

Για ευκολία, οι έξοδοι αυτού του επιπέδου ονομάζονται κανονικοποιημένοι βαθμοί ενεργοποίησης (normalized firing strengths).

Επίπεδο 4: Κάθε κόμβος i στο τέταρτο επίπεδο είναι ένας κόμβος ο οποίος υπόκειται στη διαδικασία μάθησης. Η διαδικασία που εκτελεί είναι η εξής:

$$O_i^4 = \bar{w}_i * f_i = \bar{w}_i * (p_i * x + q_i * y + r_i)$$

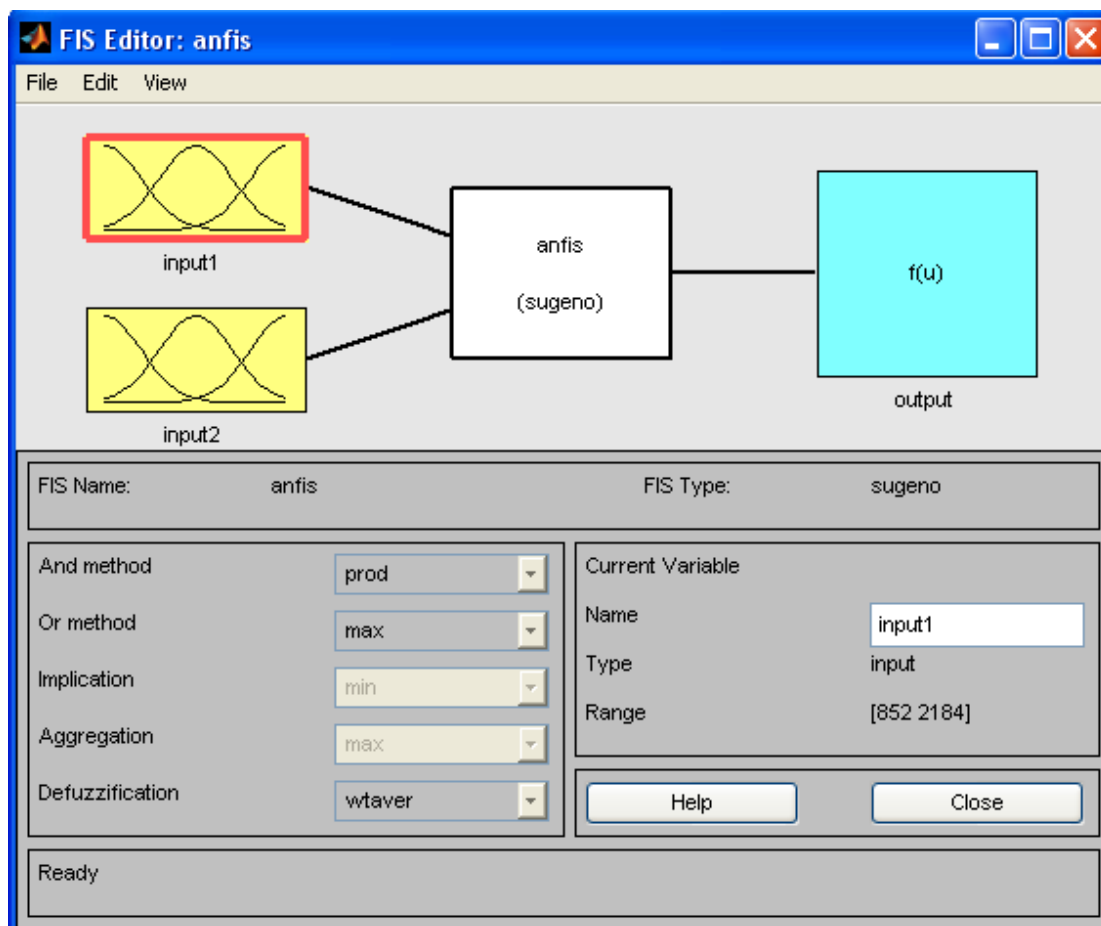
Όπου w_i η έξοδος του επιπέδου 3 και $\{p_i, q_i, r_i\}$ το σύνολο παραμέτρων. Οι παράμετροι σε αυτό το επίπεδο αναφέρονται ως *επακόλουθοι* (consequent parameters).

Επίπεδο 5: Ο μοναδικός κόμβος που υπάρχει στο τελευταίο επίπεδο, είναι ένας κυκλικός κόμβος που συμβολίζεται με το γράμμα της πρόσθεσης Sum, και αποτελεί το επίπεδο εξόδου, αναλαμβάνει το καθήκον του υπολογισμού της εξόδου της μηχανής ασαφούς συμπερασμού μέσω της συνάθροισης όλων των εισερχόμενων σημάτων. Συγκεκριμένα έχουμε:

$$Overall\ output = O_i^5 = \sum_i \bar{w}_i * f_i = \frac{\sum_i w_i * f_i}{\sum_i w_i}$$

Περιγραφή του μοντέλου

Το μοντέλο μας αποτελείται από δύο εισόδους και μια έξοδο. Σαν εισόδους παίρνουμε την πώληση του προϊόντος μιας τυχαίας ημέρας (t) και την πώληση της ακριβώς προηγούμενης ημέρας ($t-1$), ενώ σαν έξοδο παίρνουμε την πρόβλεψη του ANFIS για την πώληση του προϊόντος της επόμενης από την τυχαία ημέρα ($t+1$). Η δομή του μοντέλου παρουσιάζεται στο ακόλουθο σχήμα:



(σχήμα 3.4.4): Δομή Μοντέλου

Από το σχήμα βλέπουμε ότι το ANFIS είναι τύπου Sugeno. Το ασαφές μοντέλο τύπου Sugeno προτάθηκε από τους Tagaki, Sugeno και Kang σε μια προσπάθεια να αναπτύξουν μια συστηματική προσέγγιση που θα τους επιτρέψει την παραγωγή ασαφών κανόνων από ένα συγκεκριμένο σύνολο δεδομένων εισόδου-εξόδου. Ένας τυπικός ασαφής κανόνας για ένα ασαφές μοντέλο τύπου Sugeno έχει την ακόλουθη μορφή:

$$\text{αν } x \text{ είναι } A \text{ και } y \text{ είναι } B \text{ τότε } z = f(x,y)$$

όπου τα A και B είναι ασαφή σύνολα στα αίτια (antecedent), ενώ το $z = f(x,y)$ είναι μια σαφής (crisp) συνάρτηση στο συμπέρασμα (consequent). Συνήθως το $f(x,y)$ είναι ένα πολυώνυμο των μεταβλητών εισόδου x και y , αλλά μπορεί να είναι οποιαδήποτε συνάρτηση αρκεί να μπορεί να περιγράψει κατάλληλα την έξοδο του μοντέλου εντός της ασαφούς περιοχής που καθορίζεται από τα αίτια του κανόνα. Εμείς έχουμε ένα

πρώτου βαθμού ασαφές μοντέλο τύπου Sugeno το οποίο σημαίνει ότι το $f(x,y)$ είναι ένα πρώτου βαθμού πολυώνυμο. Στο μοντέλο μας ισχύουν οι παρακάτω κανόνες:

Αν (t) είναι small και (t-1) είναι small τότε (t+1) είναι verysmall

Αν (t) είναι small και (t-1) είναι large τότε (t+1) είναι small

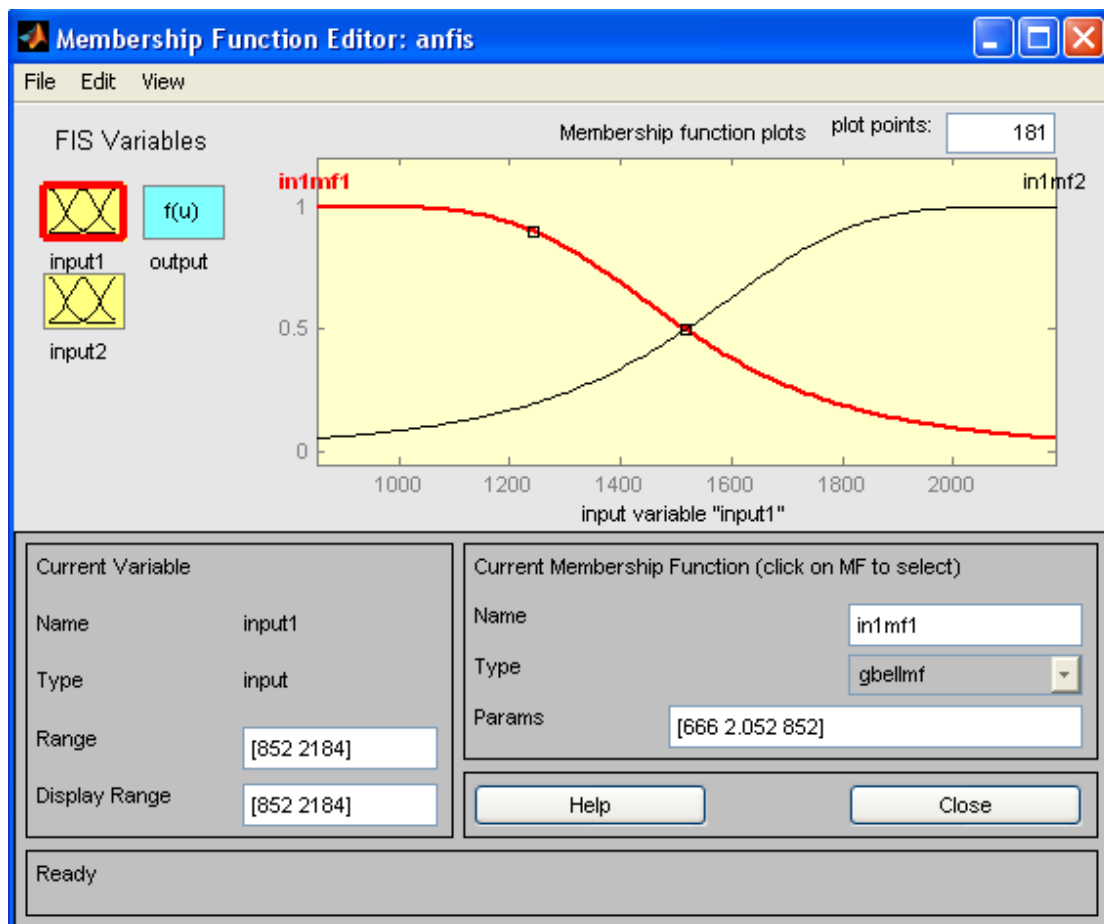
Αν (t) είναι large και (t-1) είναι small τότε (t+1) είναι large

Αν (t) είναι large και (t-1) είναι large τότε (t+1) είναι very large

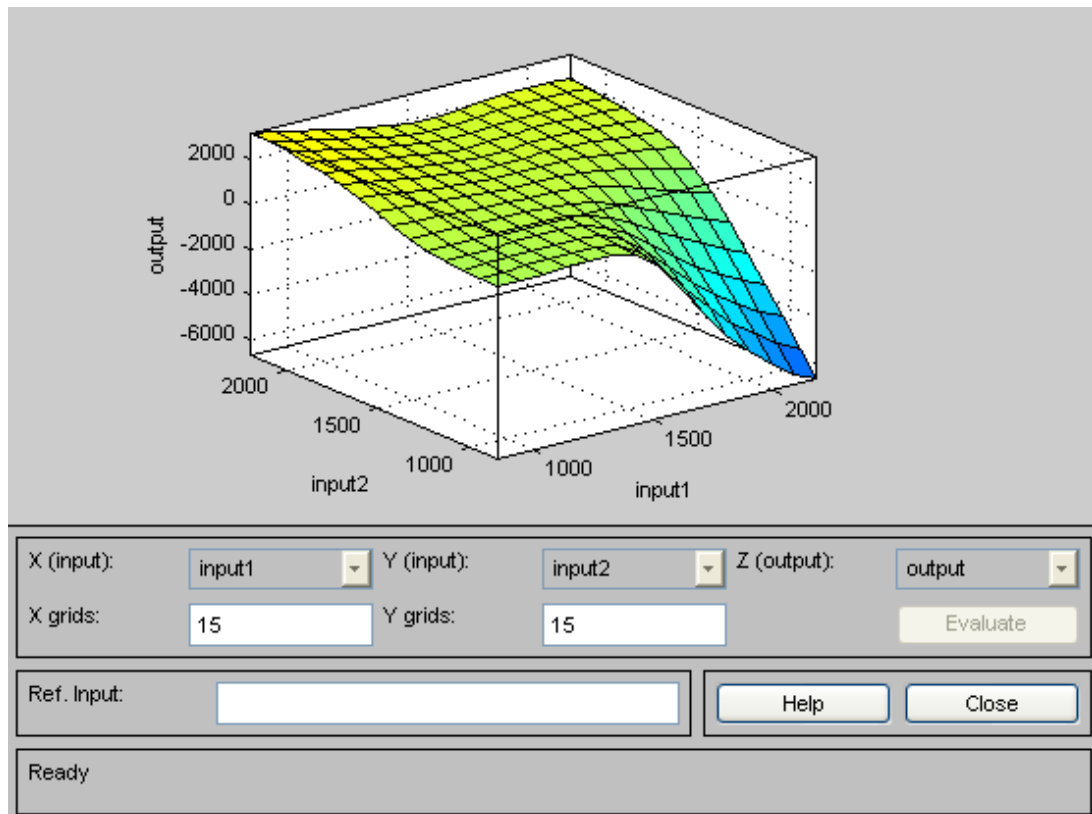
Οι συναρτήσεις συμμετοχής των εισόδων (t) και (t-1) είναι οι ίδιες, αφού και οι δύο συναρτήσεις είναι τύπου gbellmf. Η εξίσωση των συναρτήσεων τύπου gbellmf είναι:

$$\mu_A(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}}$$

όπου τα a_i, b_i, c_i είναι το σύνολο των παραμέτρων. Στο σχήμα 3.4.5 μπορούμε να δούμε τη συνάρτηση συμμετοχής των εισόδων (t) και (t-1), ενώ στο σχήμα 3.4.6 είναι η προκύπτουσα επιφάνεια εισόδου-εξόδου:

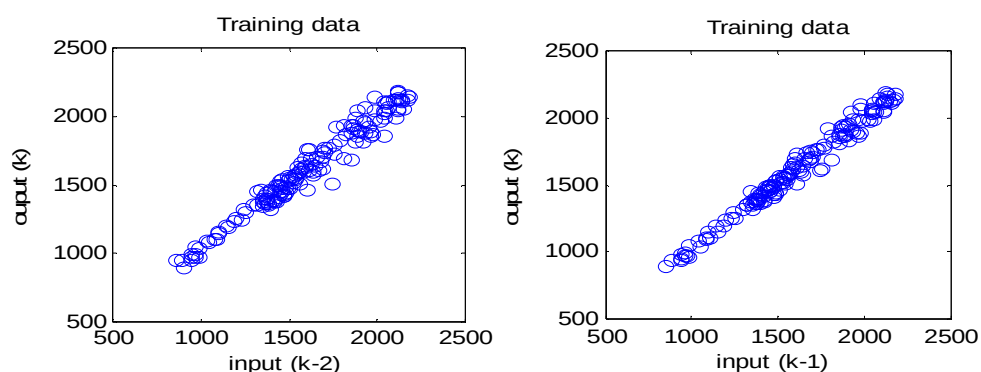


(σχήμα 3.4.5): Συνάρτηση Συμμετοχής



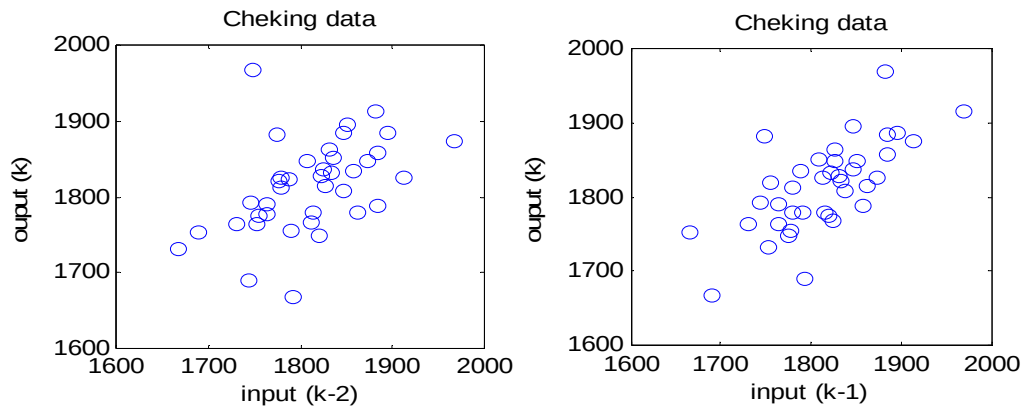
(σχήμα 3.4.6): Συνολική επιφάνεια εισόδου-εξόδου

Στο σχήμα 3.4.6 έχουμε μια τρισδιάστατη επιφάνεια του ασαφούς μοντέλου. Όπως μπορούμε να δούμε η επιφάνεια είναι αρκετά ομαλή, πράγμα που σημαίνει ότι έχουμε πάρει ένα ικανοποιητικό αριθμό δεδομένων και έτσι τα αποτελέσματα που θα πάρουμε από τη χρήση του μοντέλου που κατασκευάσαμε θα είναι αρκετά καλά. Εντούτοις, αυτό δεν φαίνεται να συμβαίνει στη κάτω δεξιά γωνία. Αυτή η ανωμαλία μπορεί να εξηγηθεί με την εξέταση των γραφημάτων διασποράς στα επόμενα σχήματα:



(σχήμα 3.4.7): Γράφημα διασποράς δεδομένων εκπαίδευσης

Τα δεδομένα εκπαίδευσης εκτείνονται για $k-2$ από 750 έως 2200.
Τα δεδομένα εκπαίδευσης εκτείνονται για $k-1$ από 750 έως 2200.



(σχήμα 3.4.8): Γράφημα διασποράς δεδομένων ελέγχου

Τα δεδομένα ελέγχου εκτείνονται για $k-2$ από 1670 έως 1980.
Τα δεδομένα ελέγχου εκτείνονται για $k-1$ από 1670 έως 1980.

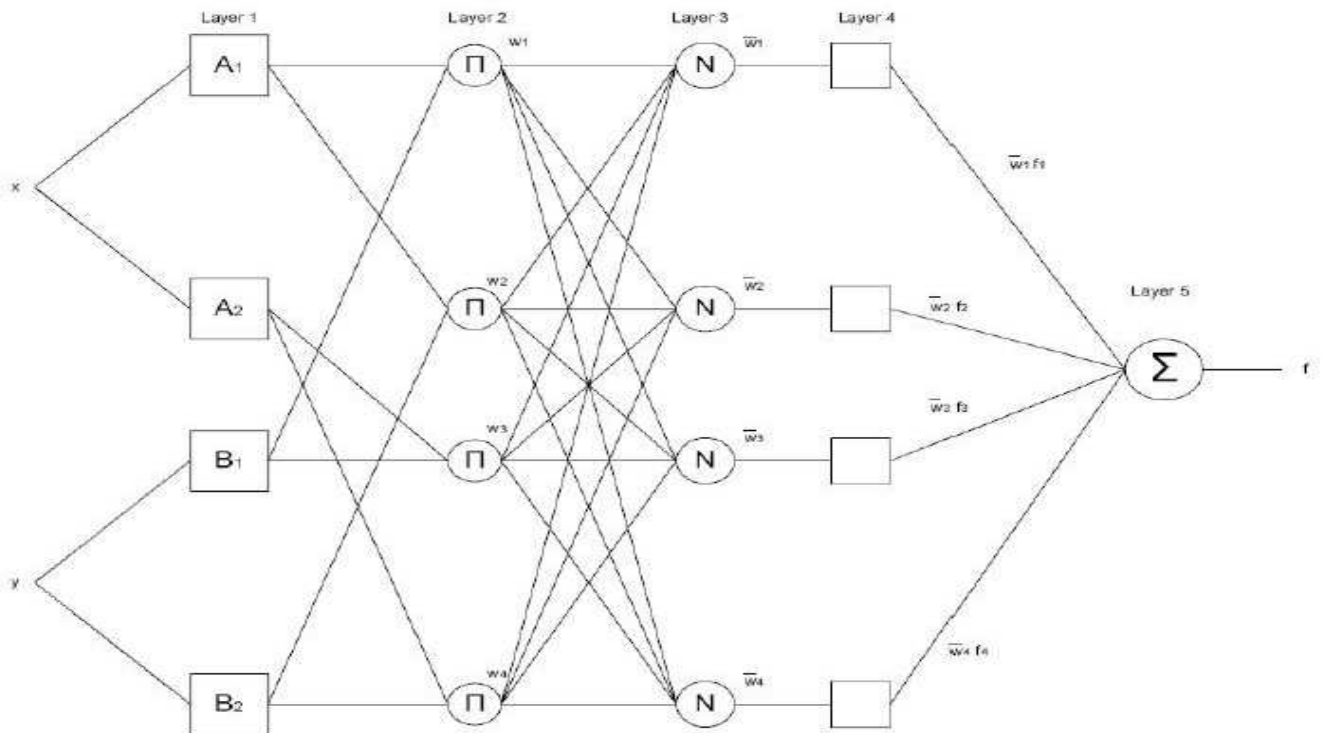
Αριθμητικό Παράδειγμα Μοντέλου

Για την καλύτερη κατανόηση του μοντέλου, παρατίθεται ένα αριθμητικό παράδειγμα που αποσαφηνίζει τις λειτουργίες που γίνονται σε κάθε επίπεδο αλλά και στους κόμβους κάθε επιπέδου, ώστε να προκύψει το τελικό αποτέλεσμα.

Όπως αναφέρθηκε στην προηγούμενη ενότητα, το μοντέλο μας είναι ένα ασαφές μοντέλο τύπου Sugeno πρώτου βαθμού το οποίο έχει δύο εισόδους και μία έξοδο. Επίσης, στο μοντέλο ισχύουν τέσσερις ασαφείς κανόνες οι οποίοι είναι οι εξής:

- Αν (t) είναι small και $(t-1)$ είναι small τότε $(t+1)$ είναι very small
- Αν (t) είναι small και $(t-1)$ είναι large τότε $(t+1)$ είναι small
- Αν (t) είναι large και $(t-1)$ είναι small τότε $(t+1)$ είναι large
- Αν (t) είναι large και $(t-1)$ είναι large τότε $(t+1)$ είναι very large

Λαμβάνοντας υπόψη τα στοιχεία που αναφέρθηκαν παραπάνω φαίνεται στη συνέχεια (σχήμα 3.4.9) η αρχιτεκτονική του μοντέλου ANFIS. Ύστερα θα δούμε αναλυτικά τι συμβαίνει σε κάθε επίπεδο του μοντέλου και τέλος θα κάνουμε χρήση ενός παραδείγματος για την καλύτερη κατανόησή του. (Εδώ παρουσιάζουμε την έξοδο του i -οστού επιπέδου στο επίπεδο l σαν O_i^l .)



(σχήμα 3.4.9): Αρχιτεκτονική του μοντέλου ANFIS

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

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

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

όπου x (ή y) – η είσοδος στον κόμβο i και A_i (ή B_{i-2}) - η γλωσσική μεταβλητή (small, large, κλπ.) που σχετίζεται με αυτή τη συνάρτηση του κόμβου.

Με άλλα λόγια, το $O_{1,1}$ είναι ο βαθμός συμμετοχής του $A(=A_1, A_2, B_1$ or $B_2)$ και καθορίζει το βαθμό στον οποίο η είσοδος x (ή y) ικανοποιεί τον ποσοτικοποιητή A . Εδώ η συνάρτηση συμμετοχής του A είναι τύπου Gaussian2:

$$\mu_{A_i}(x) = e^{-\frac{1}{2} \left(\frac{x-c_i}{\sigma_i} \right)^2}$$

όπου a_i, c_i είναι το σύνολο παραμέτρων.

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

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

$$O_i^2 = w_i = \mu_{A_i}(x) * \mu_{B_i}(y), i=1,2 \text{ και}$$

$$O_i^2 = w_i = \mu_{A_i}(x) * \mu_{B_{i-2}}(y), i=3,4$$

Επίπεδο 3: Κάθε κόμβος σε αυτό το επίπεδο είναι ένας σταθερός κόμβος N . Ο i -οστός κόμβος υπολογίζει το λόγο της βαθμού ενεργοποίησης (firing strength) του i -οστού κανόνα στο άθροισμα των βαθμών ενεργοποίησης όλων των κανόνων:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2 + w_3 + w_4} \quad i=1,2,3,4$$

Για ευκολία, οι έξοδοι αυτού του επιπέδου ονομάζονται κανονικοποιημένοι βαθμοί ενεργοποίησης (normalized firing strengths).

Επίπεδο 4: Κάθε κόμβος σε αυτό το επίπεδο είναι ένας προσαρμόσιμος κόμβος με μία συνάρτηση κόμβου.

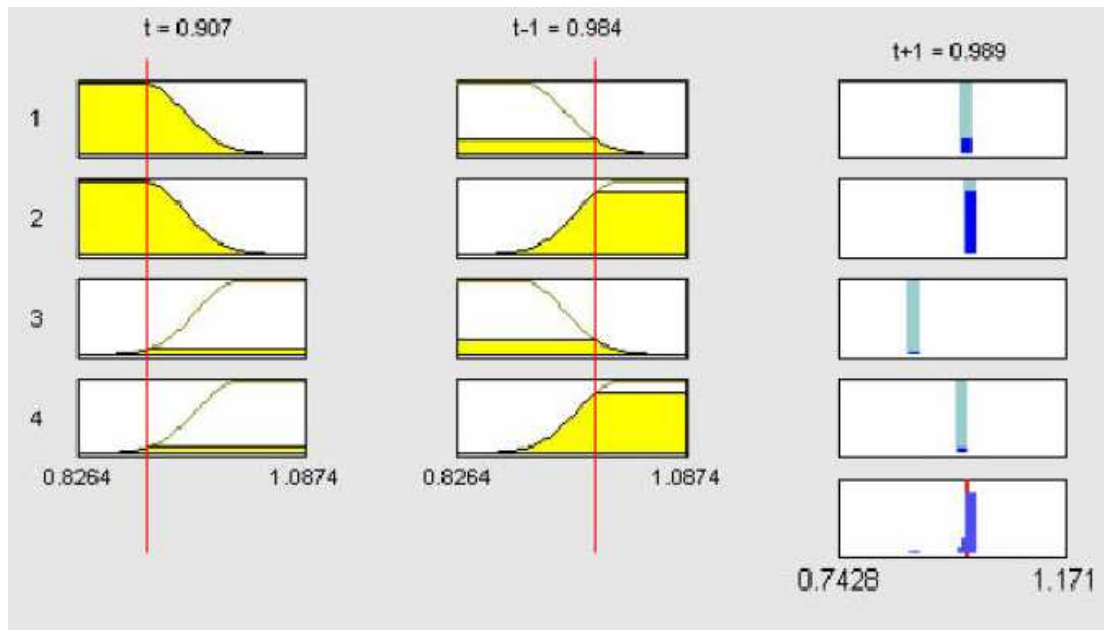
$$O_i^4 = \bar{w}_i * f_i = \bar{w}_i * (p_i * x + q_i * y + r_i)$$

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

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

$$\text{Overall output} = O_i^5 = \sum_i \bar{w}_i * f_i = \frac{\sum_i w_i * f_i}{\sum_i w_i}$$

Στο επόμενο σχήμα βλέπουμε πως το μοντέλο ANFIS βρίσκει το αποτέλεσμα που προκύπτει για δύο τυχαίες τιμές t και $t-1$:



(σχήμα 3.4.10): Οι υπολογισμοί του ANFIS

Όπως βλέπουμε στο σχήμα αν $t = 0.907$ και $t-1 = 0.984$ τότε $t+1 = 0.989$. Ας δούμε τώρα πως βγαίνει αυτό το αποτέλεσμα.

Επίπεδο 1: Εδώ χρησιμοποιούμε τη συνάρτηση συμμετοχής Gaussian2 για να υπολογίσουμε το O_i' για κάθε κόμβο. Το σύνολο παραμέτρων φαίνεται στον membership function editor και υπολογίζεται αυτόματα. Έτσι έχουμε:

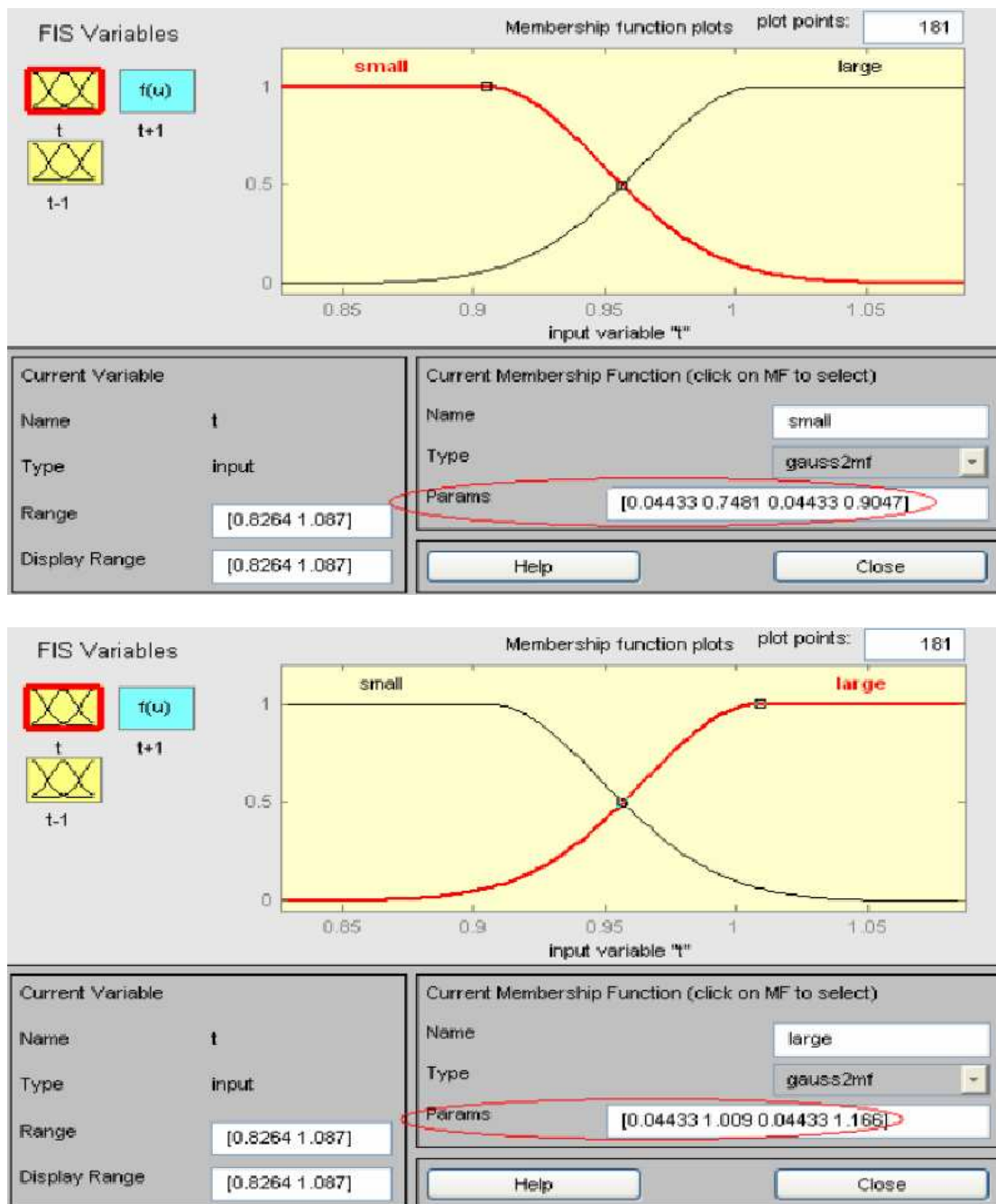
$$[a_1, c_1, a_2, c_2] = [0.04433, 0.7481, 0.04433, 0.9047]$$

αν η γλωσσική μεταβλητή είναι small και

$$[a_1, c_1, a_2, c_2] = [0.04433, 0.1009, 0.04433, 1.166]$$

αν η γλωσσική μεταβλητή είναι large

όπου a_1, c_1 είναι για την αριστερή καμπύλη της συνάρτησης Gaussian2 και a_2, c_2 είναι για την δεξιά καμπύλη, a είναι το πλάτος της καμπύλης και c το κέντρο της καμπύλης. Επίσης, αφού και οι δύο εισόδους είναι ίδιου τύπου, το σύνολο παραμέτρων είναι το ίδιο και για τις δύο εισόδους.



(σχήμα 3.4.11): Membership Function Editor

Από το σχήμα 3.4.11 βλέπουμε ότι τα $A1$ και $B1$ είναι **small** και στη **δεξιά** καμπύλη, ενώ τα $A2$ και $B2$ είναι **large** και στην **αριστερή** καμπύλη. Αυτό σημαίνει ότι για τα $A1$ και $B1$ θα χρησιμοποιήσουμε το σύνολο παραμέτρων [0.04433, 0.9047 και για τα $A2$ και $B2$ το σύνολο παραμέτρων [0.04433, 1.009].

Τώρα χρησιμοποιώντας όλα αυτά και για $x = 0.907$ και $y = 0.984$ παίρνουμε:

$$O_1^1(x) = \mu_{A_1}(x) = 0.999$$

$$O_2^1(x) = \mu_{A_2}(x) = 0.701$$

$$O_3^1(y) = \mu_{B_1}(y) = 0.202$$

$$O_4^1(y) = \mu_{B_2}(y) = 0.853$$

Επίπεδο 2: Εδώ κάνουμε τους υπολογισμούς χρησιμοποιώντας τις εξισώσεις που αναφέρθηκαν παραπάνω και έχουμε:

$$O_1^2 = w_1 = \mu_{A_1}(x) * \mu_{B_1}(y) = 0.999 * 0.202$$

$$O_2^2 = w_2 = \mu_{A_1}(x) * \mu_{B_2}(y) = 0.999 * 0.853$$

$$O_3^2 = w_3 = \mu_{A_2}(x) * \mu_{B_1}(y) = 0.701 * 0.202$$

$$O_4^2 = w_4 = \mu_{A_2}(x) * \mu_{B_2}(y) = 0.701 * 0.853$$

Επίπεδο 3: Όπως και στο επίπεδο 2 εδώ κάνουμε τους υπολογισμούς χρησιμοποιώντας την εξίσωση που αναφέρθηκε παραπάνω:

$$O_1^3 = \bar{w}_1 = \frac{w_1}{w_1 + w_2 + w_3 + w_4} = \frac{0.2016}{0.2016 + 0.8518 + 0.0143 + 0.0604} = 0.1787$$

$$O_2^3 = \bar{w}_2 = \frac{w_2}{w_1 + w_2 + w_3 + w_4} = \frac{0.8518}{0.2016 + 0.8518 + 0.0143 + 0.0604} = 0.755$$

$$O_3^3 = \bar{w}_3 = \frac{w_3}{w_1 + w_2 + w_3 + w_4} = \frac{0.0143}{0.2016 + 0.8518 + 0.0143 + 0.0604} = 0.0127$$

$$O_4^3 = \bar{w}_4 = \frac{w_4}{w_1 + w_2 + w_3 + w_4} = \frac{0.0604}{0.2016 + 0.8518 + 0.0143 + 0.0604} = 0.0536$$

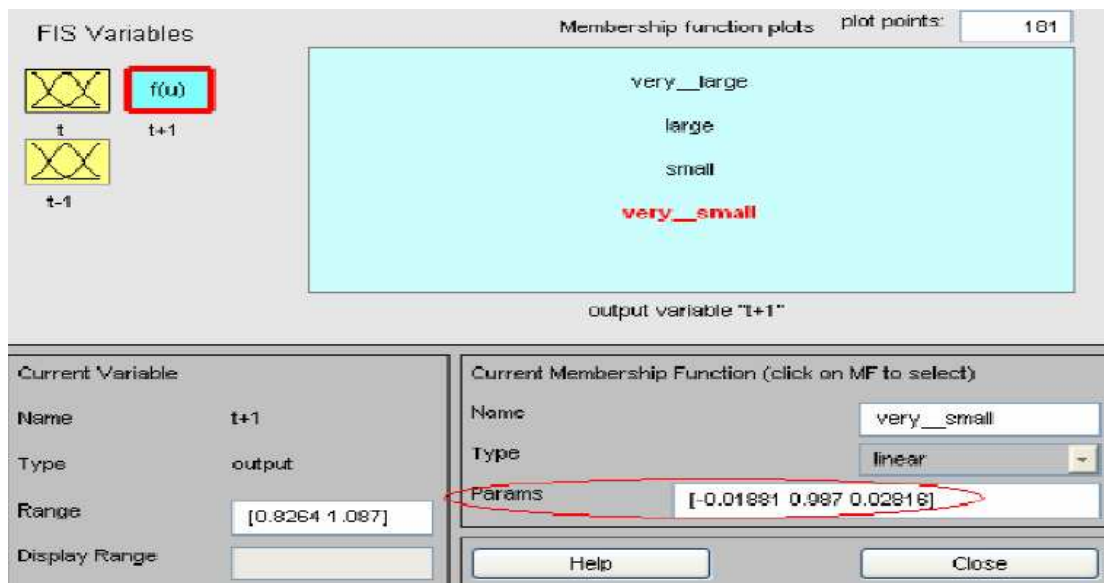
Επίπεδο 4: Για να μπορέσουμε να πραγματοποιήσουμε τους απαραίτητους υπολογισμούς σε αυτό το επίπεδο χρειαζόμαστε το σύνολο παραμέτρων για κάθε κόμβο. Αυτές οι παράμετροι φαίνονται στον membership function editor επιλέγοντας t+1 και είναι:

$$[p_1, q_1, r_1] = [-0.01881, 0.987, 0.02816] \quad \text{αν η γλωσσική μεταβλητή είναι very small}$$

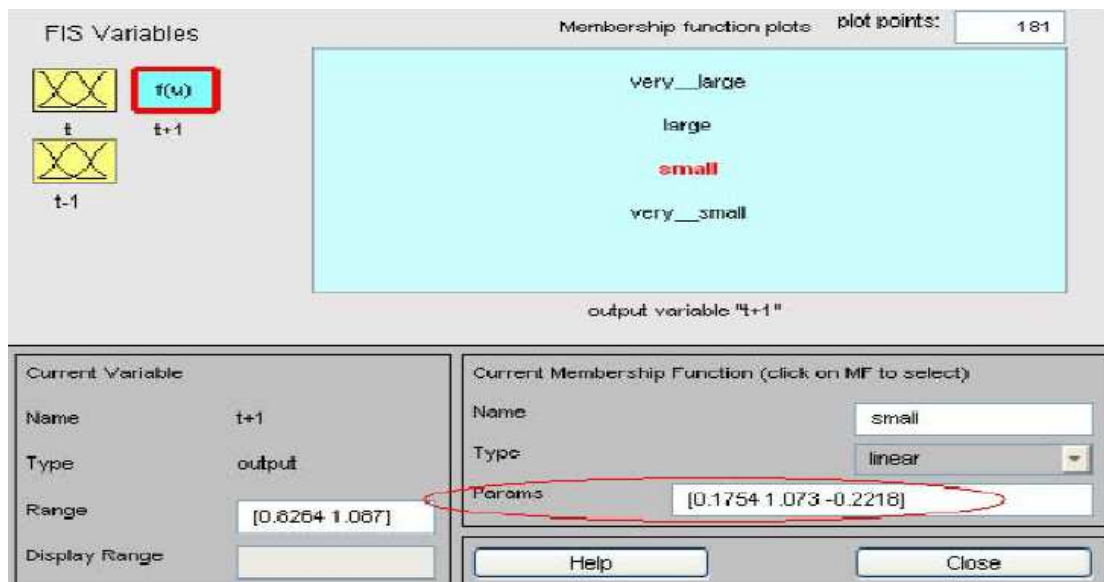
$$[p_2, q_2, r_2] = [0.1754, 1.073, -0.2218] \quad \text{αν η γλωσσική μεταβλητή είναι small}$$

$$[p_3, q_3, r_3] = [0.9956, -0.2648, 0.2437] \quad \text{αν η γλωσσική μεταβλητή είναι large}$$

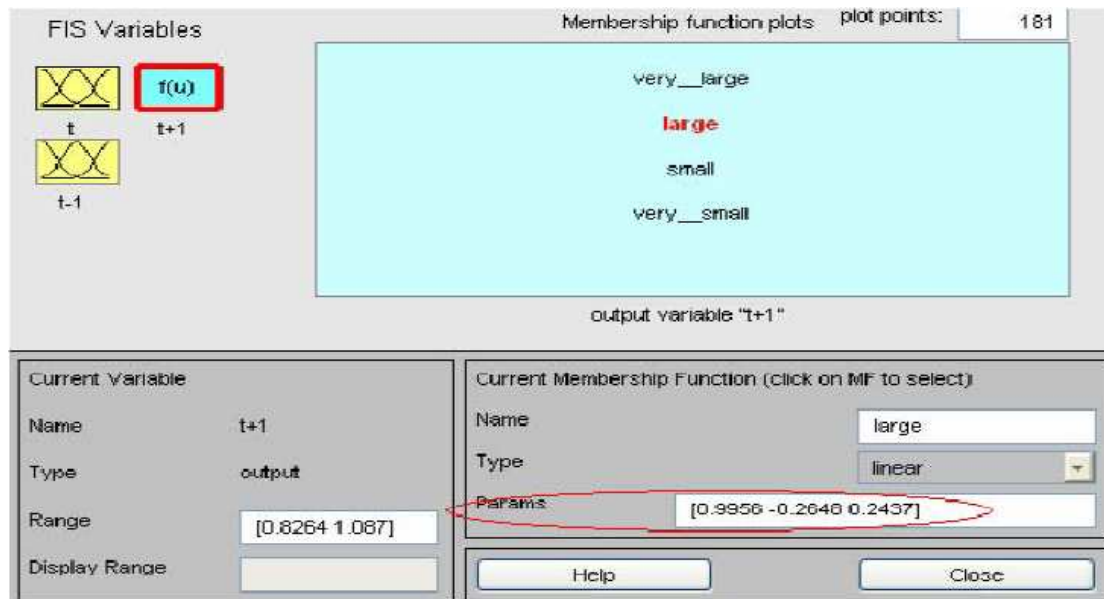
$$[p_4, q_4, r_4] = [0.09453, 0.9042, 0.0006383] \quad \text{αν η γλωσσική μεταβλητή είναι very large}$$



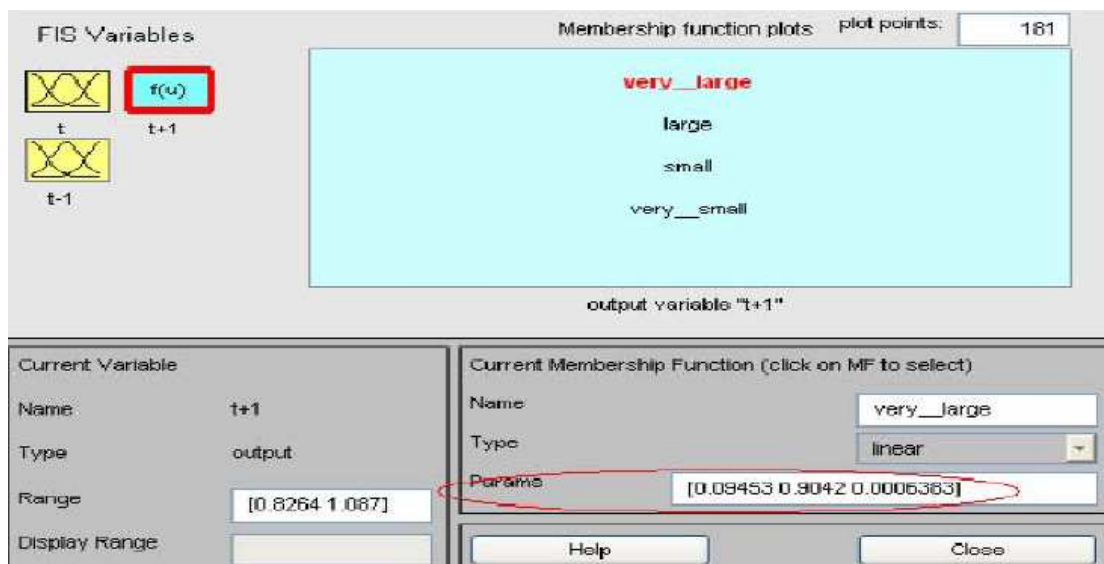
(σχήμα 3.4.12): Σύνολο παραμέτρων για *very small*



(σχήμα 3.4.13): Σύνολο παραμέτρων για *small*



(σχήμα 3.4.14): Σύνολο παραμέτρων για large



(σχήμα 3.4.15): Σύνολο παραμέτρων για very large

Χρησιμοποιώντας τα σύνολα παραμέτρων έχουμε :

$$O_1^4 = \bar{w}_1 * (p_1 * x + q_1 * y + r_1) = 0.1787 * (-0.01881 * 0.907 + 0.987 * 0.984 - 0.02816) = 0.176$$

$$O_2^4 = \bar{w}_2 * (p_2 * x + q_2 * y + r_2) = 0.755 * (0.1754 * 0.907 + 1.073 * 0.984 - 0.2218) = 0.75$$

$$O_3^4 = \bar{w}_3 * (p_3 * x + q_3 * y + r_3) = 0.0127 * (0.9956 * 0.907 - 0.2648 * 0.984 + 0.2437) = 0.011$$

$$O_4^4 = \bar{w}_4 * (p_4 * x + q_4 * y + r_4) = 0.05363 * (0.09453 * 0.907 - 0.9042 * 0.984 + 0.0006363) = 0.052$$

Επίπεδο 5: Στο τελευταίο αυτό επίπεδο αυτό βρίσκουμε την τιμή για t+1 χρησιμοποιώντας την εξίσωση που αναφέρθηκε στη θεωρία:

$$O_i^5 = \sum_i \bar{w}_i * f_i = 0.176 + 0.75 + 0.011 + 0.052 = 0.989$$

4. ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ (case study)

4.1 ΕΙΣΑΓΩΓΗ

Τα τελευταία χρόνια έχει επιτευχθεί μεγάλη πρόοδος στη βελτιστοποίηση της μεθοδολογίας πρόβλεψης πωλήσεων κι έχουν προταθεί πολλοί τρόποι προσέγγισης προβλημάτων όπως και διάφορες εναλλακτικές για την επίλυση τους. Γίνεται φανερό ανατρέχοντας στη βιβλιογραφία, ότι ολοένα και περισσότεροι κλάδοι και επιχειρήσεις, θεωρούν αναγκαίο εργαλείο την πρόβλεψη για τη σωστή διοίκηση και την ορθή χάραξη πολιτικής πωλήσεων. Στα πλαίσια λοιπόν της σπουδαιότητας της τεχνικής αυτής, θα επιχειρήσουμε να εφαρμόσουμε το μοντέλο ANFIS (*Adaptive Neuro-Fuzzy Inference System*) σε πέντε διαφορετικούς τύπους προϊόντων που πωλούνται στην αγορά των Ηνωμένων Πολιτειών. Αλλάζοντας τις παραμέτρους του μοντέλου μας, εξετάζεται η συμπεριφορά του και η αποτελεσματικότητά του, σχολιάζονται τα αποτελέσματα των προβλέψεων και τέλος καταγράφονται τα συμπεράσματα.

4.2 ΔΕΔΟΜΕΝΑ

Τα δεδομένα της μελέτης αντλήθηκαν από το site www.economagic.com/, το οποίο περιέχει τεράστιο όγκο στατιστικών δεδομένων για τη χώρα των Ηνωμένων Πολιτειών Αμερικής. Επίσης, χρησιμοποιήθηκε το site www.econstats.com/ κυρίως για επαλήθευση της αξιοπιστίας των δεδομένων μας.

Τα πέντε είδη προϊόντων για τα οποία συλλέξαμε δεδομένα είναι τα εξής:

- Computer and software (υπολογιστές και λογισμικό)
- Furniture and home furnishings (έπιπλα)
- Electronic shopping and mail order (ηλεκτρονικό εμπόριο)
- Gasoline (βενζίνη)
- Shoes (υποδήματα)

Τα δεδομένα των πωλήσεων που καταγράφηκαν, αφορούν μηνιαίες πωλήσεις της χρονικής περιόδου Ιανουαρίου του 1992 έως και Φεβρουάριο του 2009, τα ποσά των οποίων αναφέρονται σε εκατομμύρια δολάρια. Κατά μέσο όρο λοιπόν, χρησιμοποιήθηκαν για κάθε μία από τις περιπτώσεις που μελετάμε, 206 μετρήσεις. Από αυτές, οι 161 χρησιμοποιήθηκαν στην εκπαίδευση του μοντέλου μας, ενώ οι υπόλοιπες 39 για την επαλήθευσή του. Η συλλογή και αποθήκευση των παραπάνω δεδομένων έγινε σε φύλλα excel, ώστε να γίνει εύκολα η επεξεργασία τους από το μοντέλο μας (ANFIS) μέσω του προγράμματος της Matlab. Ακολουθεί ο πίνακας 4.2.1. με τις μετρήσεις συγκεντρωτικά.

DATE	Computer and software stores (million \$)	Electronic shopping and mail order houses (million \$)	Furniture and home furnishings stores (million \$)	Gasoline stations (million \$)	Shoe stores (million \$)
Ιαν-92	906	2673	4276	12755	1627
Φεβ-92	852	2632	4344	12660	1588
Μαρ-92	891	2479	4328	12697	1567
Απρ-92	940	2760	4371	12767	1578
Μαϊ-92	943	2881	4336	12864	1515
Ιουν-92	940	3065	4408	12918	1520
Ιουλ-92	985	3250	4386	12939	1498
Αυγ-92	963	3135	4365	13099	1522
Σεπ-92	972	3074	4350	13292	1560
Οκτ-92	965	3064	4331	13283	1569
Νοε-92	988	3033	4382	13350	1528
Δεκ-92	1048	3038	4448	13575	1556
Ιαν-93	1038	3190	4627	13417	1593
Φεβ-93	1082	3235	4528	13599	1527
Μαρ-93	1091	3235	4454	13491	1524
Απρ-93	1104	3220	4584	13517	1560
Μαϊ-93	1101	3226	4642	13578	1575
Ιουν-93	1151	3171	4616	13491	1588
Ιουλ-93	1144	3236	4657	13543	1567
Αυγ-93	1189	3282	4651	13377	1602
Σεπ-93	1198	3441	4656	13460	1624
Οκτ-93	1232	3570	4674	13705	1597
Νοε-93	1254	3704	4715	13786	1614
Δεκ-93	1245	3714	4707	13562	1644
Ιαν-94	1298	3730	4561	13731	1637
Φεβ-94	1319	3749	4827	13892	1617
Μαρ-94	1350	3830	4911	14028	1679
Απρ-94	1347	4026	4918	13826	1607
Μαϊ-94	1375	3867	4990	13814	1623
Ιουν-94	1392	3907	5034	14102	1619

Ιουλ-94	1442	3896	5053	14324	1667
Αυγ-94	1459	3980	5167	14635	1660
Σεπ-94	1492	3938	5194	14621	1681
Οκτ-94	1533	4013	5248	14725	1696
Νοε-94	1564	4000	5224	14894	1710
Δεκ-94	1570	4081	5256	14916	1694
Ιαν-95	1588	4099	5254	15010	1663
Φεβ-95	1612	4014	5156	14959	1531
Μαρ-95	1627	4072	5141	14905	1707
Απρ-95	1604	4175	5187	15082	1707
Μαϊ-95	1693	4286	5222	15214	1715
Ιουν-95	1753	4274	5243	15348	1735
Ιουλ-95	1613	4404	5318	15308	1692
Αυγ-95	1726	4478	5326	15228	1695
Σεπ-95	1763	4592	5391	15104	1721
Οκτ-95	1769	4579	5429	15083	1698
Νοε-95	1794	4741	5551	15027	1770
Δεκ-95	1918	4816	5442	15248	1703
Ιαν-96	1829	4735	5290	15466	1745
Φεβ-96	1817	4848	5433	15527	1728
Μαρ-96	1867	4859	5560	15953	1776
Απρ-96	1930	4919	5599	16287	1807
Μαϊ-96	1912	4938	5608	16491	1800
Ιουν-96	1953	5028	5666	16385	1758
Ιουλ-96	1890	5099	5697	16059	1784
Αυγ-96	1925	5101	5672	16006	1791
Σεπ-96	2040	5143	5766	16198	1743
Οκτ-96	1970	5232	5791	16387	1785
Νοε-96	1860	5390	5784	16490	1765
Δεκ-96	1856	5479	5769	16832	1753
Ιαν-97	1879	5319	5826	16938	1753
Φεβ-97	1937	5428	5909	16962	1790
Μαρ-97	1984	5755	5919	17000	1830

Απρ-97	2065	5949	5972	16594	1702
Μαϊ-97	2042	5764	6018	16350	1769
Ιουν-97	2061	5846	6036	16443	1793
Ιουλ-97	2110	5852	6046	16465	1801
Αυγ-97	2114	5875	6154	16648	1789
Σεπ-97	2116	6172	6185	16795	1791
Οκτ-97	2124	6084	6185	16707	1799
Νοε-97	2184	6024	6265	16714	1811
Δεκ-97	2179	6087	6306	16505	1849
Ιαν-98	2140	6215	6359	16250	1824
Φεβ-98	2155	6371	6369	16070	1882
Μαρ-98	2111	6476	6352	15877	1859
Απρ-98	2056	6581	6383	15837	1831
Μαϊ-98	1985	6774	6365	16061	1832
Ιουν-98	2097	6365	6432	15932	1842
Ιουλ-98	2146	6838	6455	16067	1874
Αυγ-98	2119	6810	6445	15841	1845
Σεπ-98	2125	6861	6464	15797	1811
Οκτ-98	2133	6974	6525	15900	1898
Νοε-98	2174	7031	6680	16076	1878
Δεκ-98	2126	6995	6733	16376	1901
Ιαν-99	2119	7107	6591	16403	1916
Φεβ-99	2047	7377	6781	16318	1894
Μαρ-99	2055	7358	6837	16555	1883
Απρ-99	2043	7541	6813	17125	1871
Μαϊ-99	2048	7632	6890	17181	1918
Ιουν-99	2026	7706	6997	17037	1943
Ιουλ-99	1990	7848	6932	17661	1905
Αυγ-99	1972	8077	7170	18218	1892
Σεπ-99	1982	8238	7350	18417	1893
Οκτ-99	1888	8343	7308	18630	1869
Νοε-99	1864	8224	7298	19022	1867
Δεκ-99	1903	8585	7306	20039	1887

Ιαν-00	1939	8734	7447	19505	1885
Φεβ-00	1894	8805	7553	20334	1885
Μαρ-00	1875	9151	7620	21027	1925
Απρ-00	1858	9333	7711	20064	1891
Μαϊ-00	1808	9082	7687	20214	1900
Ιουν-00	1681	9381	7637	20979	1888
Ιουλ-00	1689	9499	7705	20952	1865
Αυγ-00	1671	9650	7708	20588	1921
Σεπ-00	1744	9852	7702	21171	1949
Οκτ-00	1609	9932	7734	21418	1923
Νοε-00	1509	10073	7613	21742	1922
Δεκ-00	1466	10320	7185	21760	1894
Ιαν-01	1473	9465	7759	21747	1908
Φεβ-01	1478	9466	7615	21483	1855
Μαρ-01	1426	9592	7571	20581	1858
Απρ-01	1418	9535	7565	21639	1941
Μαϊ-01	1377	9576	7574	22773	1938
Ιουν-01	1385	9674	7629	21853	1901
Ιουλ-01	1397	9681	7672	20848	1964
Αυγ-01	1356	9844	7650	20909	1963
Σεπ-01	1320	9096	7452	21444	1838
Οκτ-01	1346	9672	7551	20059	1877
Νοε-01	1447	9697	7720	19326	1927
Δεκ-01	1457	9709	7864	18901	1911
Ιαν-02	1447	9952	7854	19406	1965
Φεβ-02	1456	10039	7942	19202	1982
Μαρ-02	1484	9934	7936	19889	1937
Απρ-02	1431	10070	7891	21003	1969
Μαϊ-02	1479	10138	7966	20977	1940
Ιουν-02	1490	10087	7832	21764	1965
Ιουλ-02	1442	10257	7753	21380	1924
Αυγ-02	1478	10243	7849	21097	1934
Σεπ-02	1404	10211	7839	21233	1873

Οκτ-02	1423	10280	7816	21608	1924
Νοε-02	1384	10473	7985	21792	1912
Δεκ-02	1378	10495	7968	22300	1931
Ιαν-03	1361	10444	7854	23163	1865
Φεβ-03	1376	10580	7535	23917	1889
Μαρ-03	1362	10546	7905	23931	1920
Απρ-03	1398	10664	7946	22572	1867
Μαϊ-03	1419	10655	8118	21772	1895
Ιουν-03	1453	10868	8149	21779	1926
Ιουλ-03	1455	11002	8192	22190	1969
Αυγ-03	1492	11033	8220	22793	1994
Σεπ-03	1531	11230	8290	22957	1975
Οκτ-03	1522	11334	8336	22432	1945
Νοε-03	1515	11298	8449	23142	1977
Δεκ-03	1556	11415	8421	23397	1981
Ιαν-04	1542	11740	8394	24706	1957
Φεβ-04	1509	11866	8476	25157	1993
Μαρ-04	1534	11954	8710	25427	2009
Απρ-04	1547	11948	8649	25368	1981
Μαϊ-04	1551	12168	8498	26854	1970
Ιουν-04	1578	12096	8690	26841	1951
Ιουλ-04	1629	12160	9000	26500	1957
Αυγ-04	1594	12286	8799	26263	1929
Σεπ-04	1646	12323	8756	26635	1981
Οκτ-04	1641	12476	8930	28282	1986
Νοε-04	1641	12522	8865	29105	1942
Δεκ-04	1576	12725	9079	28252	2014
Ιαν-05	1598	12992	8962	28224	1994
Φεβ-05	1659	12846	9104	28738	2144
Μαρ-05	1695	12982	9058	29397	1959
Απρ-05	1701	13241	9188	29487	2091
Μαϊ-05	1706	13315	9149	29255	2096
Ιουν-05	1737	13585	9321	29794	2112

Ιουλ-05	1765	13655	9296	30755	2103
Αυγ-05	1746	14015	9390	32631	2077
Σεπ-05	1745	13784	9528	35039	2112
Οκτ-05	1793	14237	9541	35116	2201
Νοε-05	1690	14520	9580	32837	2206
Δεκ-05	1666	14836	9561	32139	2166
Ιαν-06	1753	15146	9797	34077	2255
Φεβ-06	1731	15350	9815	33726	2237
Μαρ-06	1764	15835	9820	33586	2127
Απρ-06	1764	15568	9778	35318	2218
Μαϊ-06	1790	15617	9763	36120	2186
Ιουν-06	1778	15762	9913	36152	2188
Ιουλ-06	1755	15593	9851	36929	2180
Αυγ-06	1820	16137	9849	37129	2214
Σεπ-06	1775	16247	9991	33678	2339
Οκτ-06	1748	16406	9785	32241	2290
Νοε-06	1882	16399	9771	33151	2243
Δεκ-06	1969	16725	9777	33896	2241
Ιαν-07	1914	16808	10024	33690	2242
Φεβ-07	1874	17264	10012	34427	2166
Μαρ-07	1826	17306	10037	35218	2274
Απρ-07	1848	17149	10030	36037	2159
Μαϊ-07	1837	17468	9987	38312	2185
Ιουν-07	1808	17539	9901	37013	2180
Ιουλ-07	1851	18129	10009	36716	2227
Αυγ-07	1848	17653	9967	36214	2236
Σεπ-07	1895	17727	9887	37388	2218
Οκτ-07	1885	17847	9866	38072	2187
Νοε-07	1884	18373	9797	41830	2238
Δεκ-07	1858	17908	9690	40633	2236
Ιαν-08	1788	17810	9670	41894	2210
Φεβ-08	1834	17673	9466	41523	2224
Μαρ-08	1822	17893	9477	42296	2210

Απρ-08	1832	18258	9455	42552	2257
Μαϊ-08	1827	18397	9502	43946	2275
Ιουν-08	1863	18504	9315	45334	2286
Ιουλ-08	1815	18312	9253	45352	2282
Αυγ-08	1779	18321	8983	43988	2294
Σεπ-08	1779	18125	8804	43934	2171
Οκτ-08	1813	17871	8658	38163	2142
Νοε-08	1825	17761	8525	31201	2156
Δεκ-08	1767	18101	8443	26422	2113
Ιαν-09			8392	27158	
Φεβ-09			8449	28092	

(πίνακας 4.2.1): Στατιστικά Δεδομένα

4.3 Επεξεργασία

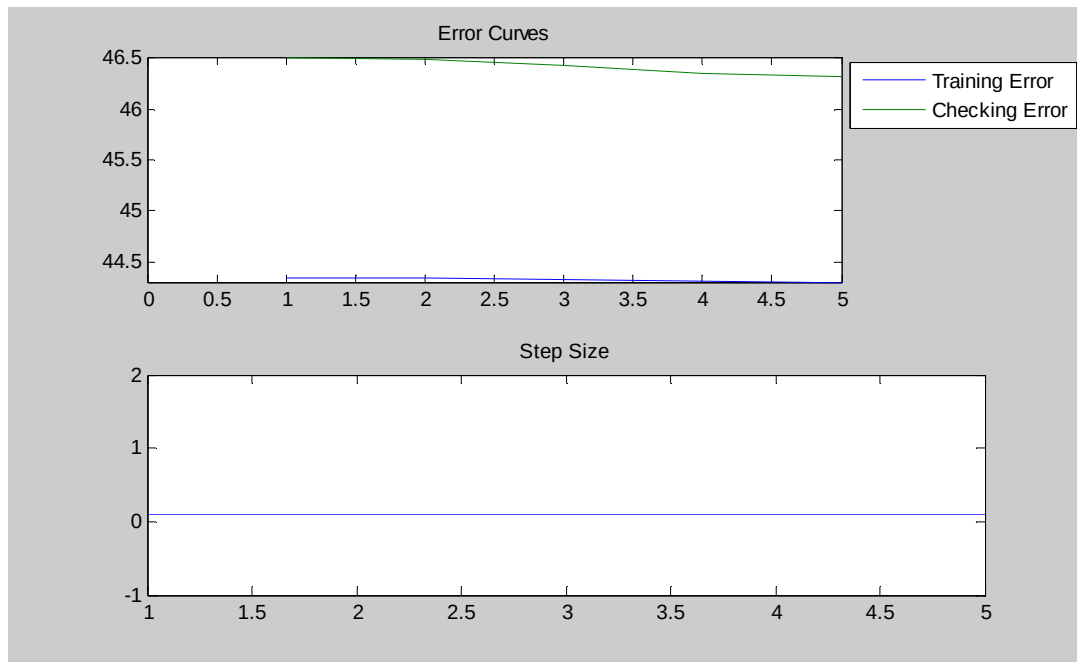
Κατά την επεξεργασία των δεδομένων μας ακολουθήσαμε την εξής μεθοδολογία: αρχικά χρησιμοποιήθηκαν τα δεδομένα που αφορούν πωλήσεις υπολογιστών και λογισμικού. Οι αρχικές παράμετροι που τέθηκαν στο μοντέλο μας είναι $n=2$ ($mf_n=2$), μέγεθος βήματος 0.1 ($ss=0.1$), εποχές = 5 ($epochs=5$) και συνάρτηση συμμετοχής γενικευμένη καμπανοειδή ($gbellmf$). Αφού εκτελέσαμε το πρόγραμμα με τις παραπάνω ρυθμίσεις και αποθηκεύσαμε τα αποτελέσματα σε ένα φύλλο excel, επαναλάβουμε τη διαδικασία για 10, 100, και 1000 εποχές κρατώντας ως συνάρτηση συμμετοχής την $gbellmf$. Στη συνέχεια τρέξαμε τα δεδομένα μας πάλι για 5, 10, 100, 1000 εποχές χρησιμοποιώντας όμως όλες τις διαθέσιμες συναρτήσεις συμμετοχής $gaussmf$ και $gauss2mf$ (γκουσιανή συνάρτηση συμμετοχής), $trimf$ (συνάρτηση συμμετοχής τριγωνικής μορφής), smf (σιγμοειδής συνάρτηση συμμετοχής), $trapmf$ (τραπεζοειδής συνάρτηση συμμετοχής), $rimf$ (συνάρτηση συμμετοχής σχήματος Π) και zmf (συνάρτηση συμμετοχής σχήματος Ζ). Όλη η παραπάνω διαδικασία επαναλήφθηκε και για τα υπόλοιπα σετ δεδομένων και καταγράφηκαν αναλυτικά σε πίνακα που θα παρουσιαστεί στο κεφάλαιο των αποτελεσμάτων.

Στα πλαίσια της βαθύτερης κατανόησης του τρόπου λειτουργίας του ANFIS, θα παρουσιάσουμε τα κυριότερα γραφήματα που προκύπτουν από το πρόγραμμα της Matlab, κατά την επεξεργασία των δεδομένων από τις πωλήσεις υπολογιστών.

Ακολουθεί πίνακας με χρήσιμες πληροφορίες σχετικά με το ANFIS:

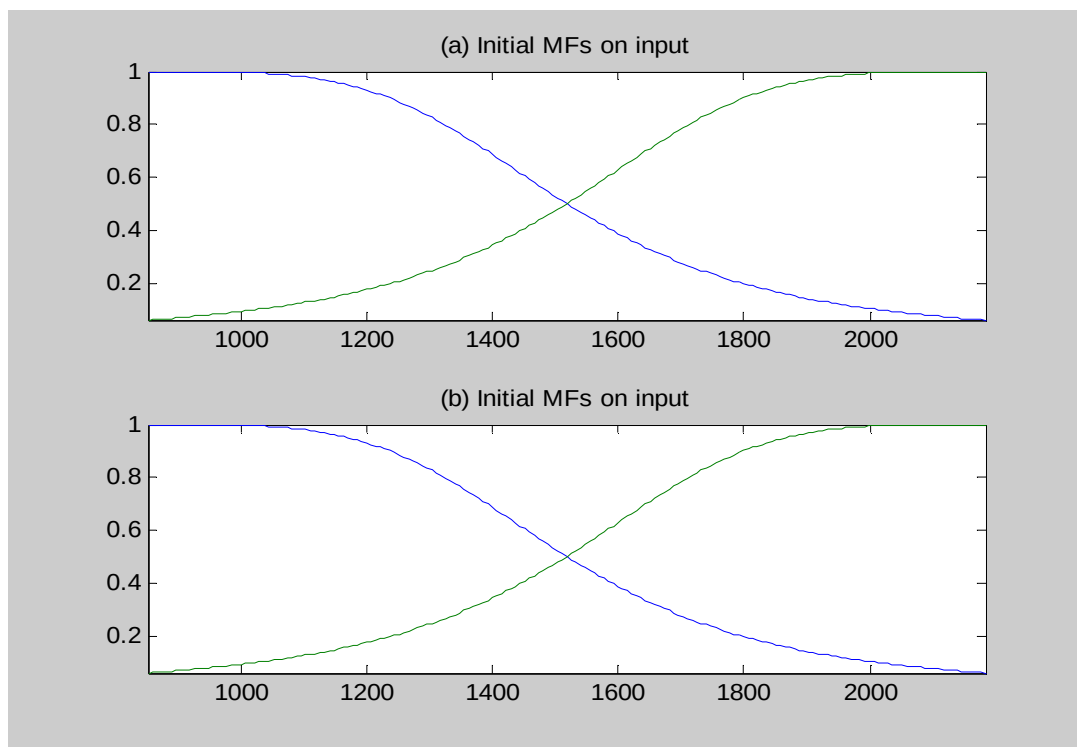
Name (Όνομα)	ANFIS
Type (Τύπος ασαφούς μηχανισμού συμπερασμού)	Sugeno
Number of nodes (Αριθμός κόμβων)	21
Number of linear parameters (Αριθμός γραμμικών παραμέτρων)	12
Number of nonlinear parameters (Αριθμός μη-γραμμικών παραμέτρων)	12
Total number of parameters (Συνολικός αριθμός παραμέτρων)	24
Number of training data pairs (Αριθμός εκπαιδευόμενων ζευγαριών δεδομένων)	161
Number of checking data pairs (Αριθμός ελεγχόμενων ζευγαριών δεδομένων)	39
Number of fuzzy rules (Αριθμός ασαφών κανόνων)	4
NumInputs (Αριθμός εισόδων)	2
NumOutputs (Αριθμός εξόδων)	1

Στο σχήμα 4.3.1 που ακολουθεί, παρουσιάζεται το σφάλμα εκπαίδευσης ελέγχου και μεγέθους βήματος:



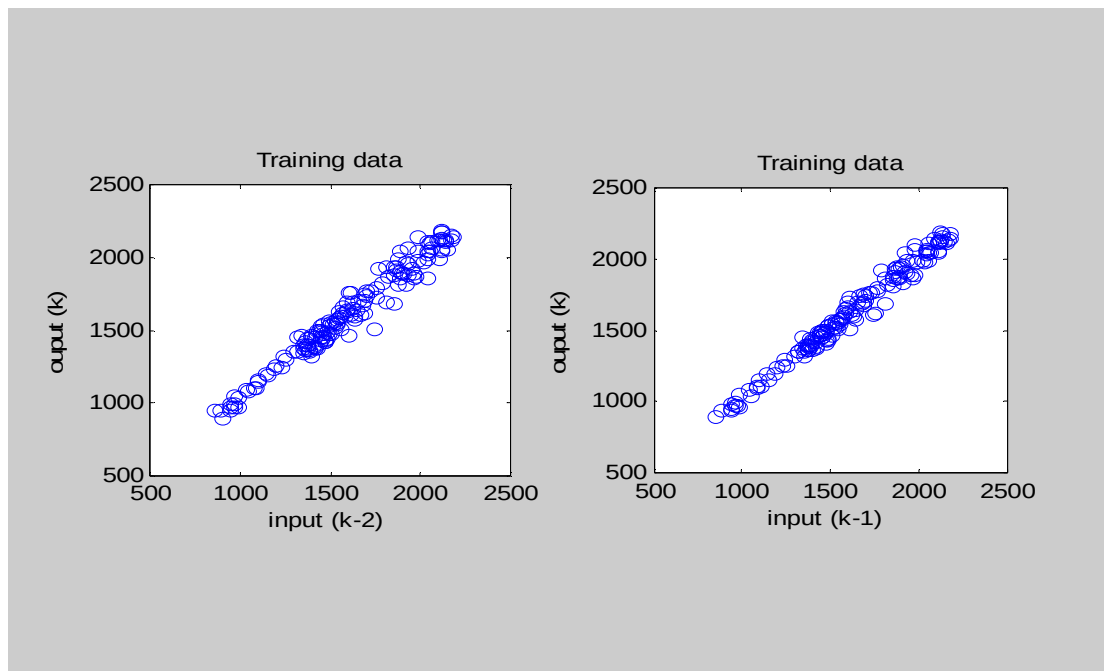
(σχήμα 4.3.1): Σφάλμα εκπαίδευσης ελέγχου και μεγέθους βήματος

Στο παρακάτω σχήμα παρατίθενται οι αρχικές συναρτήσεις συμμετοχής οι οποίες είναι καμπανοειδείς:



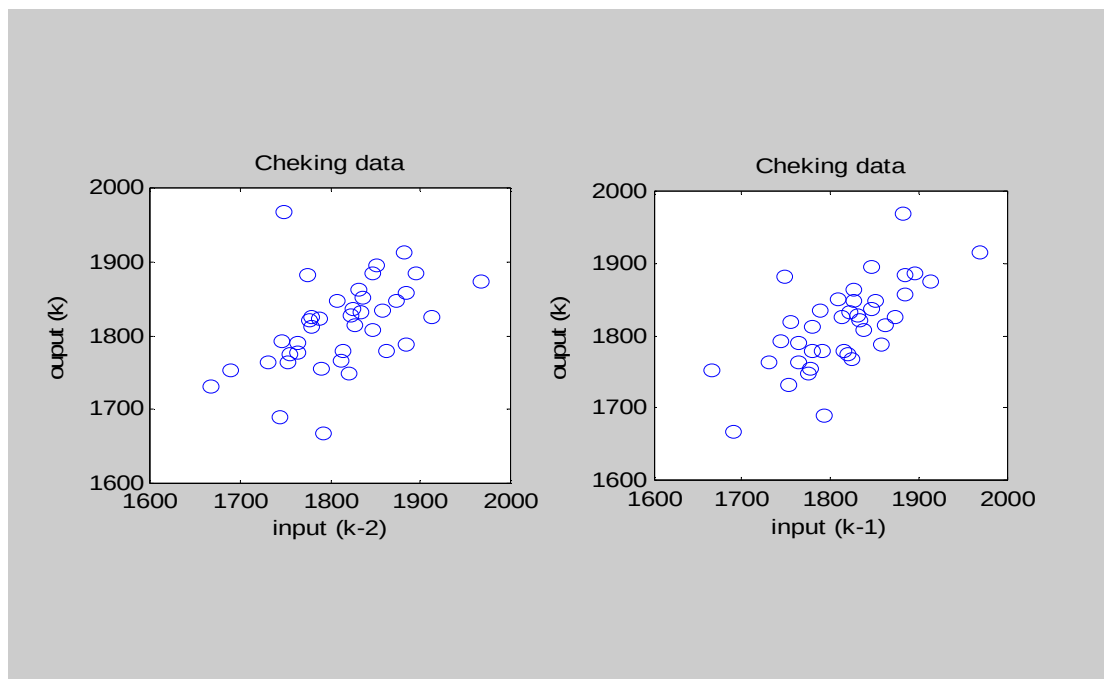
(σχήμα 4.3.2): Συναρτήσεις συμμετοχής για την gbellmf

Στο σχήμα 4.3.3. που ακολουθεί παρουσιάζεται το γράφημα διασποράς των δεδομένων εκπαίδευσης. Τα δεδομένα εκπαίδευσης εκτείνονται για $k-2$ από 750 έως 2200 και για $k-1$ επίσης από 750 έως 2200.



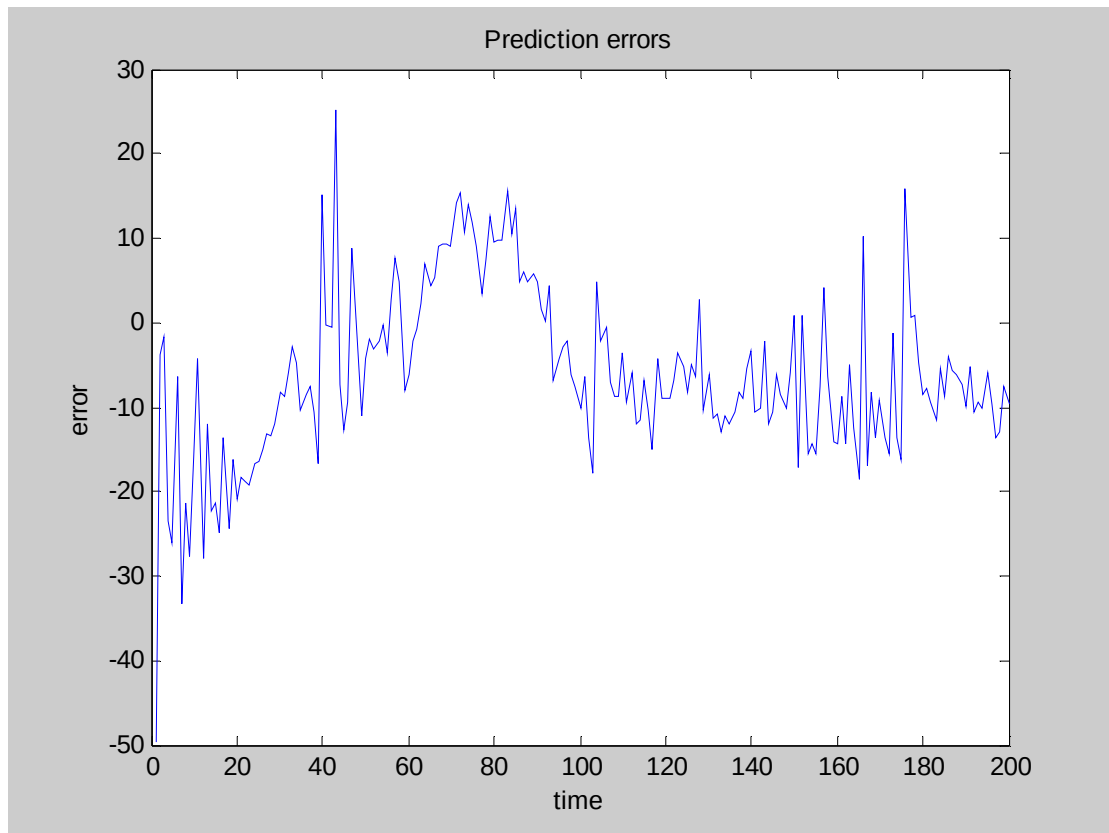
(σχήμα 4.3.3): Γράφημα διασποράς των δεδομένων εκπαίδευσης.

Στο σχήμα 4.3.4. που ακολουθεί παρουσιάζεται το γράφημα διασποράς των δεδομένων ελέγχου. Τα δεδομένα ελέγχου εκτείνονται για $k-2$ από 1670 έως 1980 και για $k-1$ επίσης από 1670 έως 1980.



(σχήμα 4.3.4): Γράφημα διασποράς των δεδομένων ελέγχου.

Το επόμενο και τελευταίο γράφημα εμφανίζει συνολικά τη διακύμανση του λάθους πρόβλεψης του μοντέλου μας.



(σχήμα 4.3.5): Γράφημα λάθους πρόβλεψης.

4.4 ΑΠΟΤΕΛΕΣΜΑΤΑ

Έπειτα από τις επαναλαμβανόμενες δοκιμές με εφαρμογή διαφορετικών παραμέτρων, το ANFIS δίνει διάφορα αποτελέσματα τα οποία συγκεντρώθηκαν όλα σε ένα φύλλο excel. Ο σημαντικότερος δείκτης που πρέπει να εξετασθεί, είναι ο RMSE (Root Mean Square Error). Ο δείκτης αυτός ουσιαστικά καταγράφει πόσο αποτελεσματικό είναι το μοντέλο μας, δηλαδή πόσο καλή πρόβλεψη μπορεί να επιτύχει και δίνεται από τον τύπο:

$$RMSE = \sqrt{\frac{\sum_{t=1}^N e_t^2}{N}}$$

Στη συνέχεια παρατίθενται τα εξαγόμενα αποτελέσματα για κάθε μία από τις περιπτώσεις που εξετάσαμε.

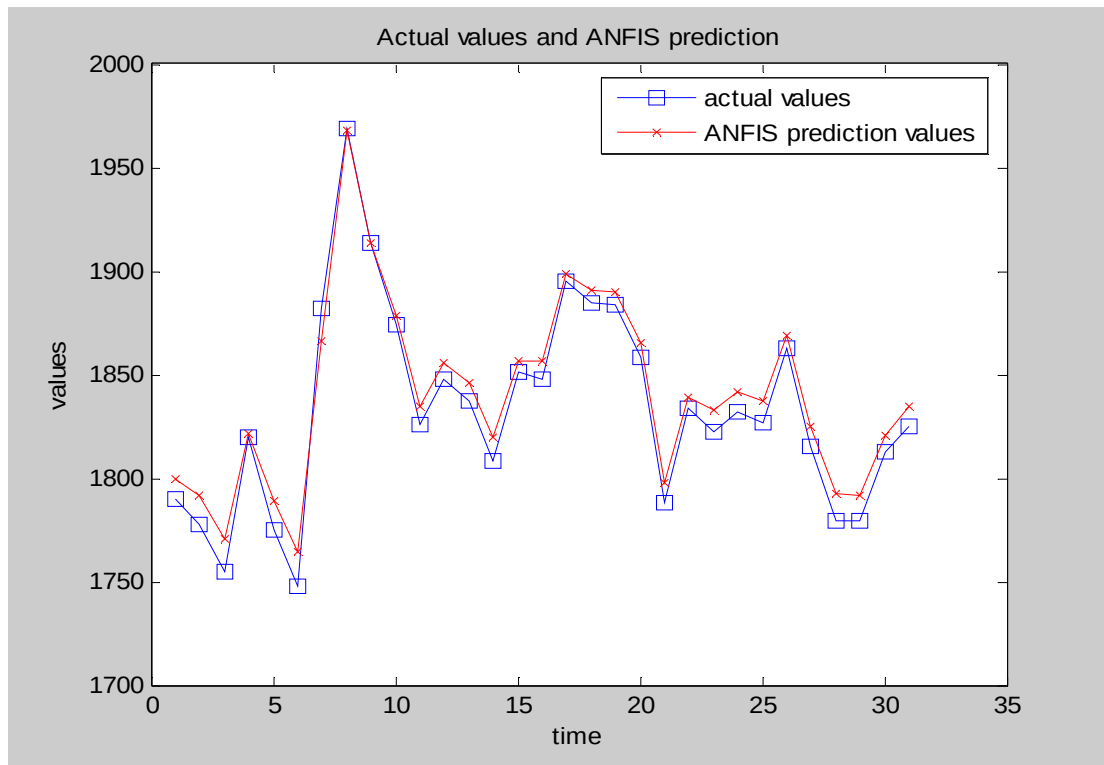
Συνάρτηση συμμετοχής gbellmf	mf_n=2 και ss=0.1			
	epochs=5 RMSE	epochs=10 RMSE	epochs=100 RMSE	epochs=1000 RMSE
Computer and software	11.77	12.24	12.75	12.79
Electronic shopping	164.50	320.44	697.01	697.01
Furniture	76.28	68.67	84.44	78.63
Gasoline stations	923.05	1193.5	22896	23197
Shoe stores	45.05	47.01	50.20	81.18
Συνάρτηση συμμετοχής trimf	mf_n=2 και ss=0.1			
	epochs=5 RMSE	epochs=10 RMSE	epochs=100 RMSE	epochs=1000 RMSE
Computer and software	11.11	11.22	11.23	11.23
Electronic shopping	1.7*10 ⁶			
Furniture	1.8*10 ⁴			
Gasoline stations	2.7*10 ⁷			
Shoe stores	2.9*10 ⁴			
Συνάρτηση συμμετοχής gauss2mf	mf_n=2 και ss=0.1			
	epochs=5 RMSE	epochs=10 RMSE	epochs=100 RMSE	epochs=1000 RMSE
Computer and software	11.73	11.77	11.94	31.33
Electronic shopping	254.46	252.70	218.45	507.57
Furniture	79.74	79.34	72.32	114.69
Gasoline stations	358.78	362.09	363.64	371.10
Shoe stores	28.16	28.21	33.71	119.09
Συνάρτηση συμμετοχής gaussmf	mf_n=2 και ss=0.1			
	epochs=5 RMSE	epochs=10 RMSE	epochs=100 RMSE	epochs=1000 RMSE
Computer and software	11.68	11.70	12.49	21.27
Electronic shopping	284.29	283.37	282.90	234.86
Furniture	84.83	84.93	85.05	85.15
Gasoline stations	864.72	864.22	864.22	864.52
Shoe stores	53.11	53.27	62.75	68.91

Συνάρτηση συμμετοχής smf	mf_n=2 και ss=0.1			
	epochs=5 RMSE	epochs=10 RMSE	epochs=100 RMSE	epochs=1000 RMSE
Computer and software	error			
Electronic shopping	error			
Furniture	error			
Gasoline stations	error			
Shoe stores	error			
Συνάρτηση συμμετοχής trapmf	mf_n=2 και ss=0.1			
	epochs=5 RMSE	epochs=10 RMSE	epochs=100 RMSE	epochs=1000 RMSE
Computer and software	13.18	13.21	14.31	14.44
Electronic shopping	210.77	210.80	211.08	210.90
Furniture	72.79	72.82	73.49	71.94
Gasoline stations	354.15	354.08	354.60	390.27
Shoe stores	28.29	28.30	30.73	30.73
Συνάρτηση συμμετοχής zmf	mf_n=2 και ss=0.1			
	epochs=5 RMSE	epochs=10 RMSE	epochs=100 RMSE	epochs=1000 RMSE
Computer and software	error			
Electronic shopping	error			
Furniture	error			
Gasoline stations	error			
Shoe stores	error			
Συνάρτηση συμμετοχής pimf	mf_n=2 και ss=0.1			
	epochs=5 RMSE	epochs=10 RMSE	epochs=100 RMSE	epochs=1000 RMSE
Computer and software	13.32	13.34	14.14	26.52
Electronic shopping	188.21	188.22	189.64	179.72
Furniture	72.98	73.00	72.86	72.01
Gasoline stations	398.20	398.20	401.98	497.32
Shoe stores	25.63	25.68	29.96	35.62

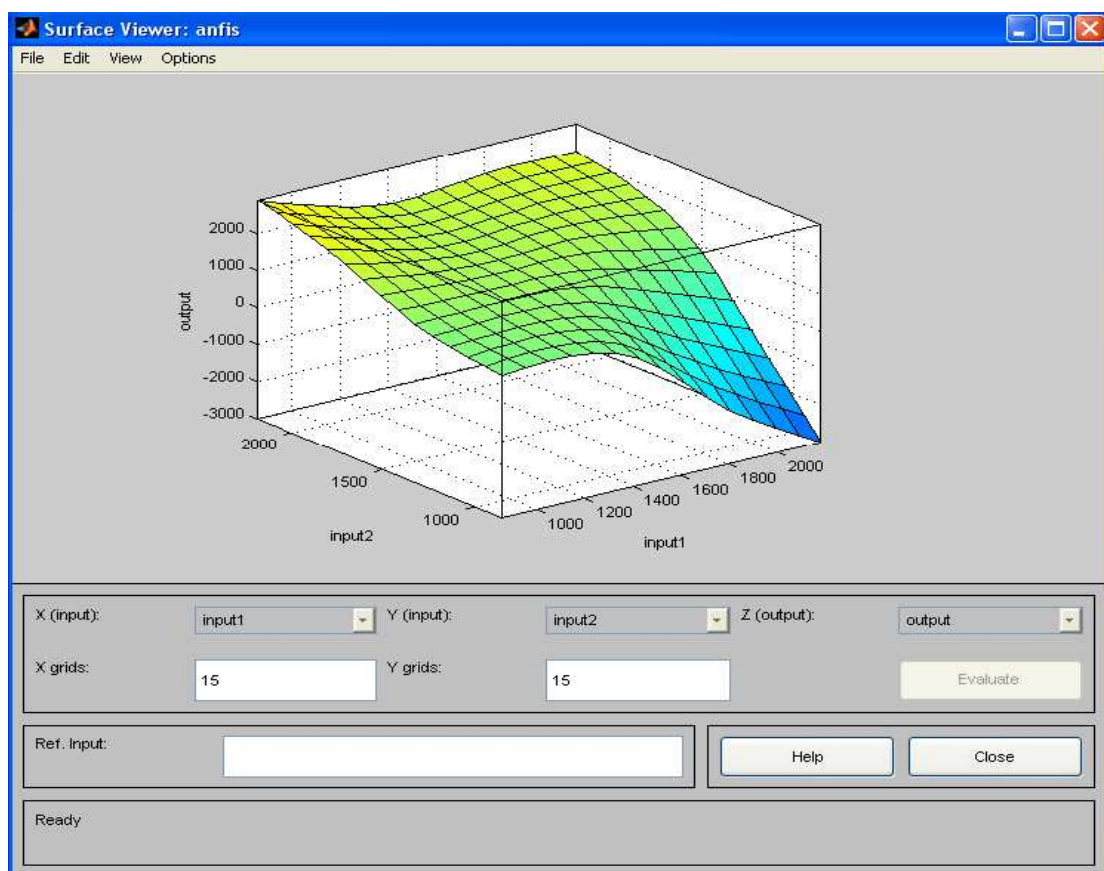
(πίνακας 4.4.1): Αποτελέσματα ANFIS

Για μεγαλύτερη ευκολία έχουν τονιστεί με bold οι χαμηλότερες τιμές του RMSE ανά κατηγορία συνάρτησης συμμετοχής, ανάλογα τις εποχές που έχουν οριστεί. Με έντονη κόκκινη πλάγια γραφή, έχει ορισθεί η συνολικά χαμηλότερη τιμή RMSE που επετεύχθη για κάθε κλάδο πωλήσεων που εξετάσαμε.

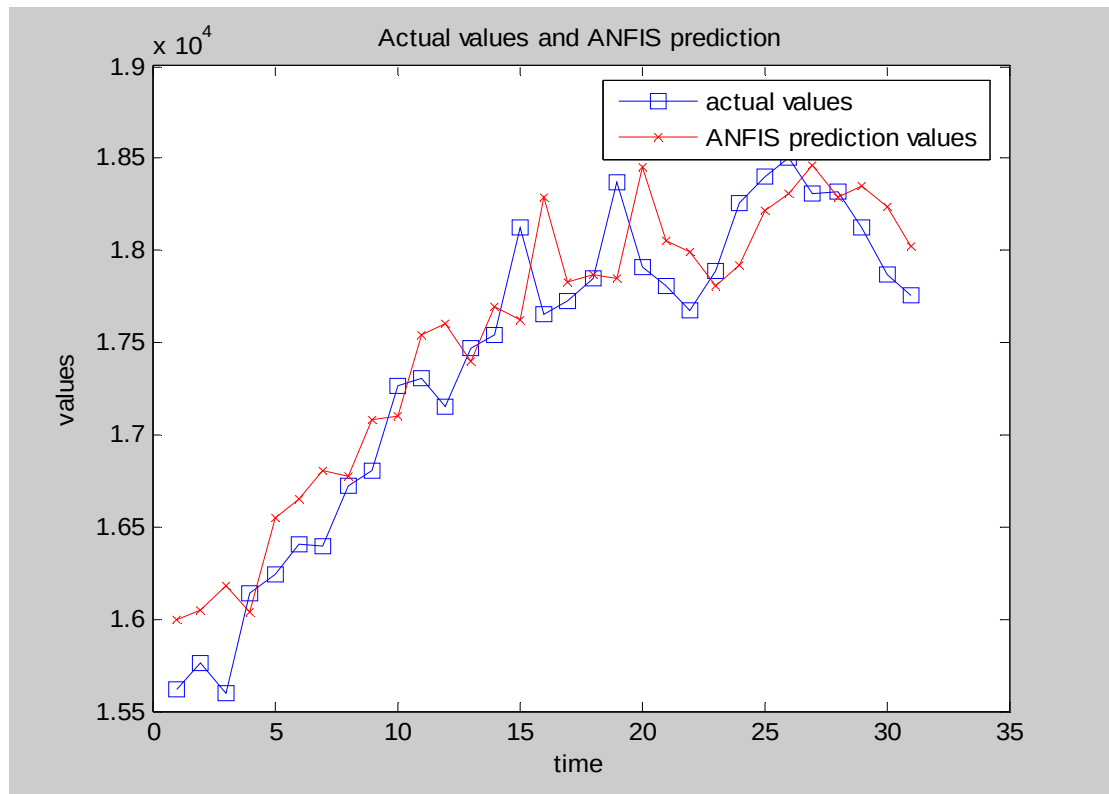
Στα σχήματα που ακολουθούν, παρουσιάζονται τα γραφήματα αντιστοιχίας των πραγματικών τιμών σε σχέση με αυτών της πρόβλεψης και η τρισδιάστατη επιφάνεια εισόδου-εξόδου, για κάθε ένα τύπο προϊόντος. Η επιλογή των γραφημάτων έχει γίνει με κριτήριο τη χαμηλότερη τιμή RMSE που προκύπτει από τον πίνακα 4.4.1. Με την παρακάτω απεικόνιση θα αποκτήσουμε πιο ολοκληρωμένη αντίληψη για την αποδοτικότητα του μοντέλου μας.



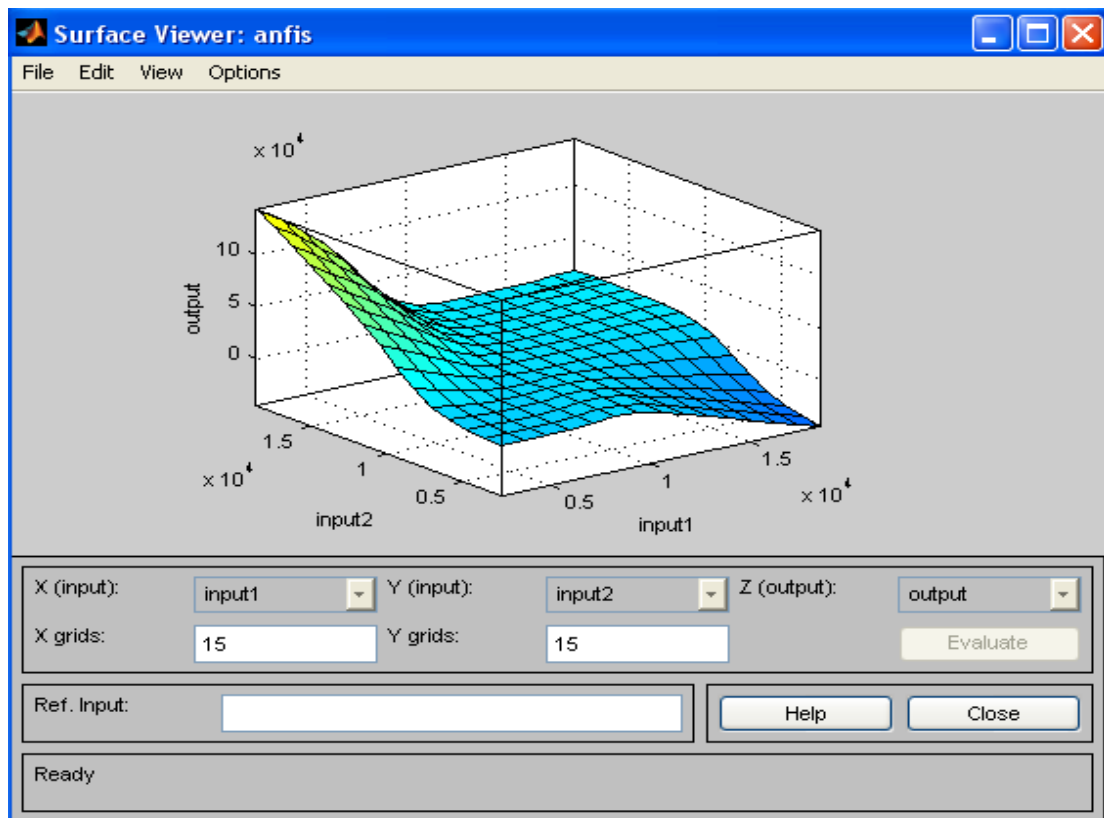
(σχήμα 4.4.2): Σύγκριση πραγματικών τιμών και τιμών πρόβλεψης που αφορά την πώληση υπολογιστών και λογισμικού



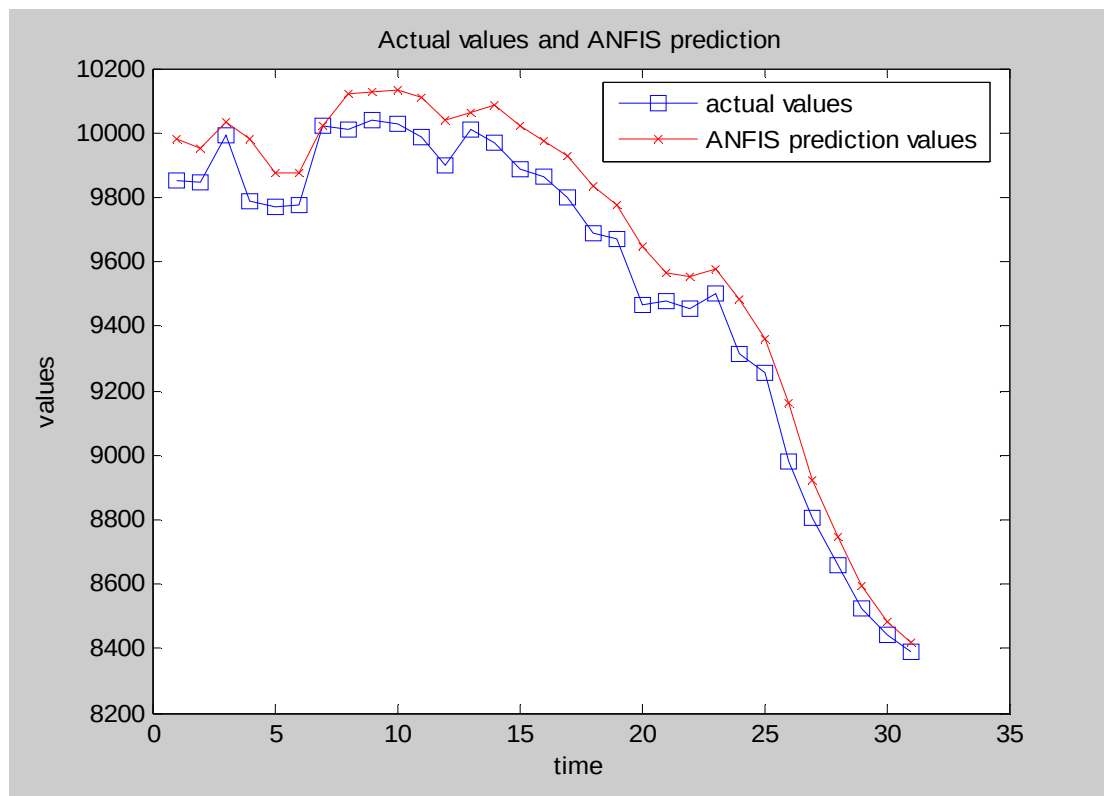
(σχήμα 4.4.3): Επιφάνεια εισόδου-εξόδου πώλησης υπολογιστών και λογισμικού



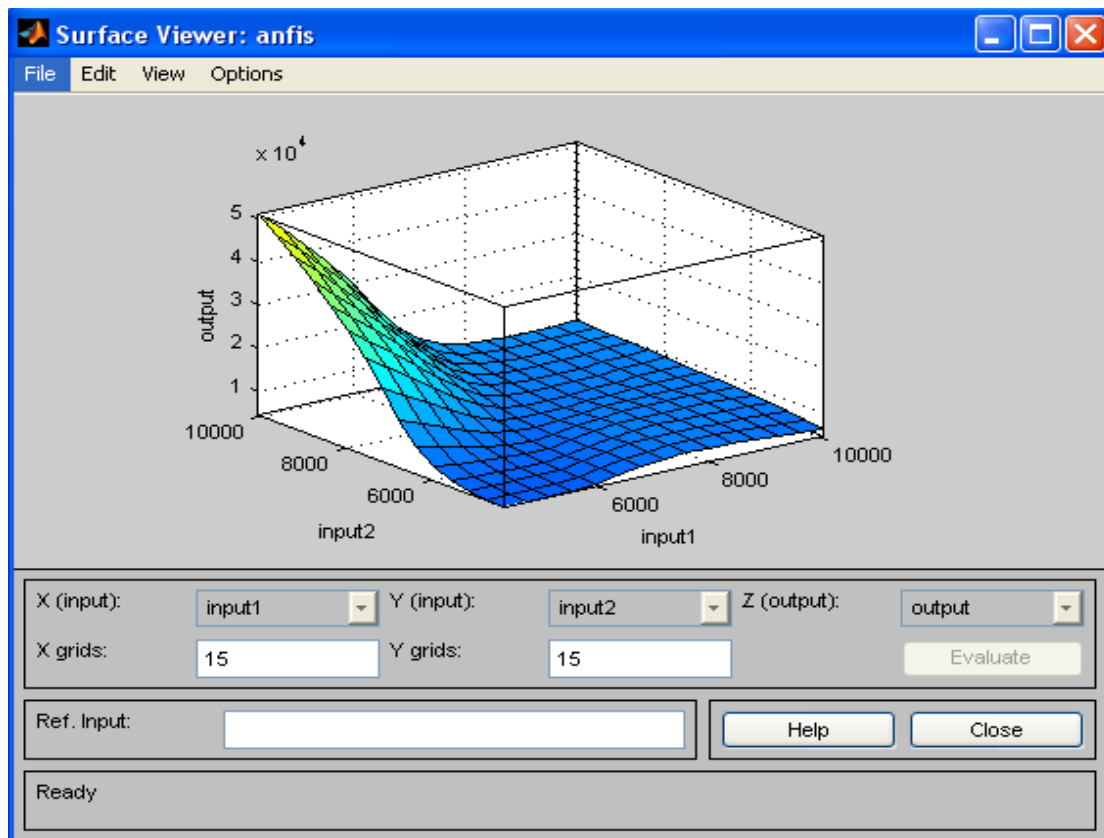
(σχήμα 4.4.4): Σύγκριση πραγματικών τιμών και τιμών πρόβλεψης που αφορά πωλήσεις ηλεκτρονικού εμπορίου



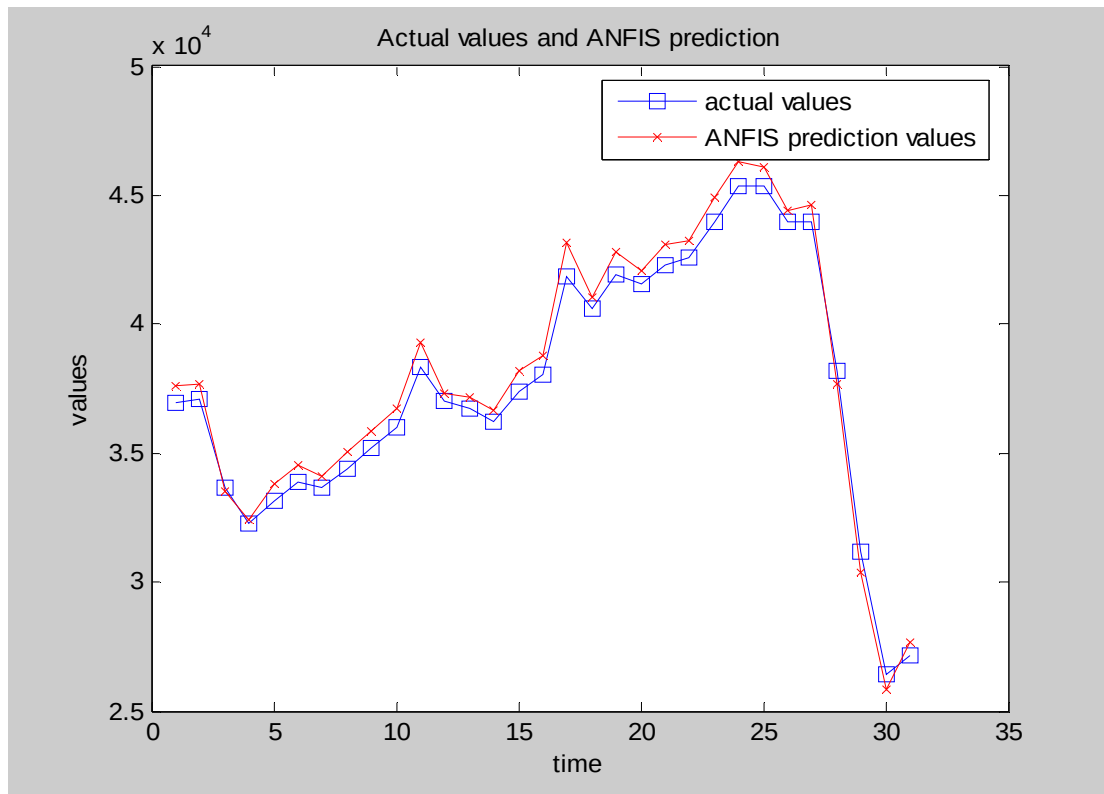
(σχήμα 4.4.5): Επιφάνεια εισόδου-εξόδου πώλησης ηλεκτρονικού εμπορίου



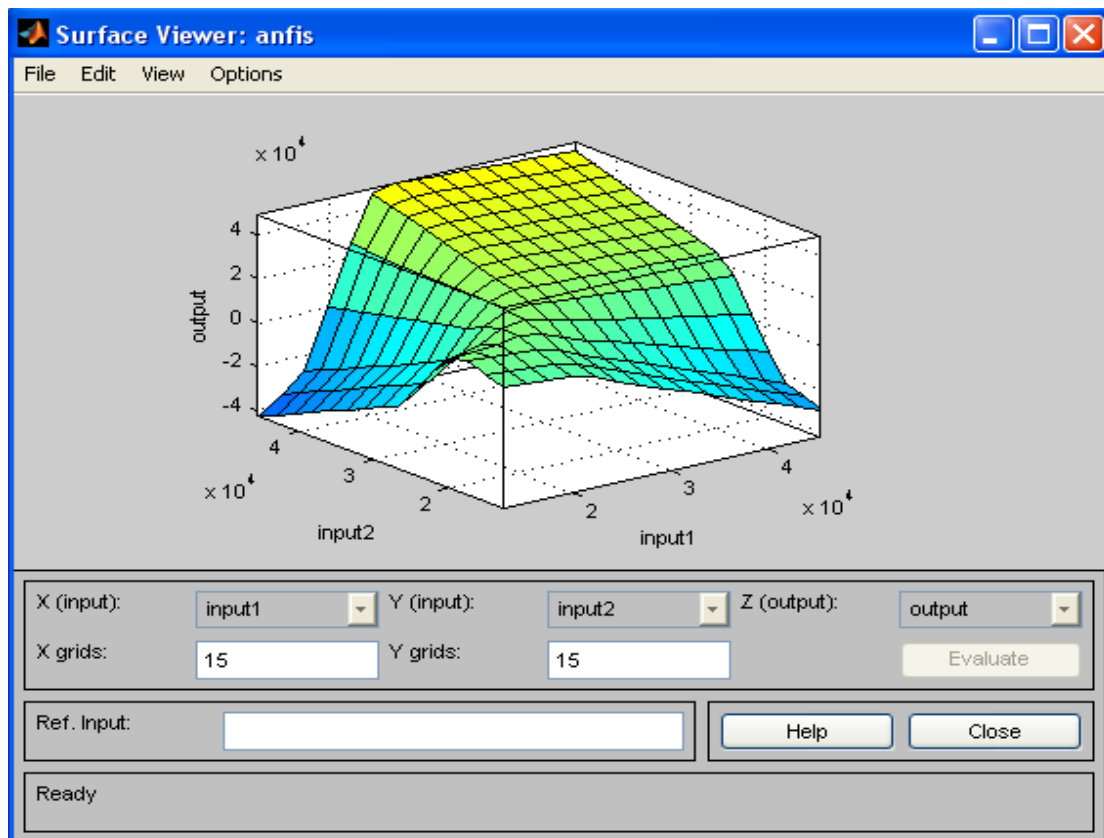
(σχήμα 4.4.6): Σύγκριση πραγματικών τιμών και τιμών πρόβλεψης που αφορά πωλήσεις επίπλων



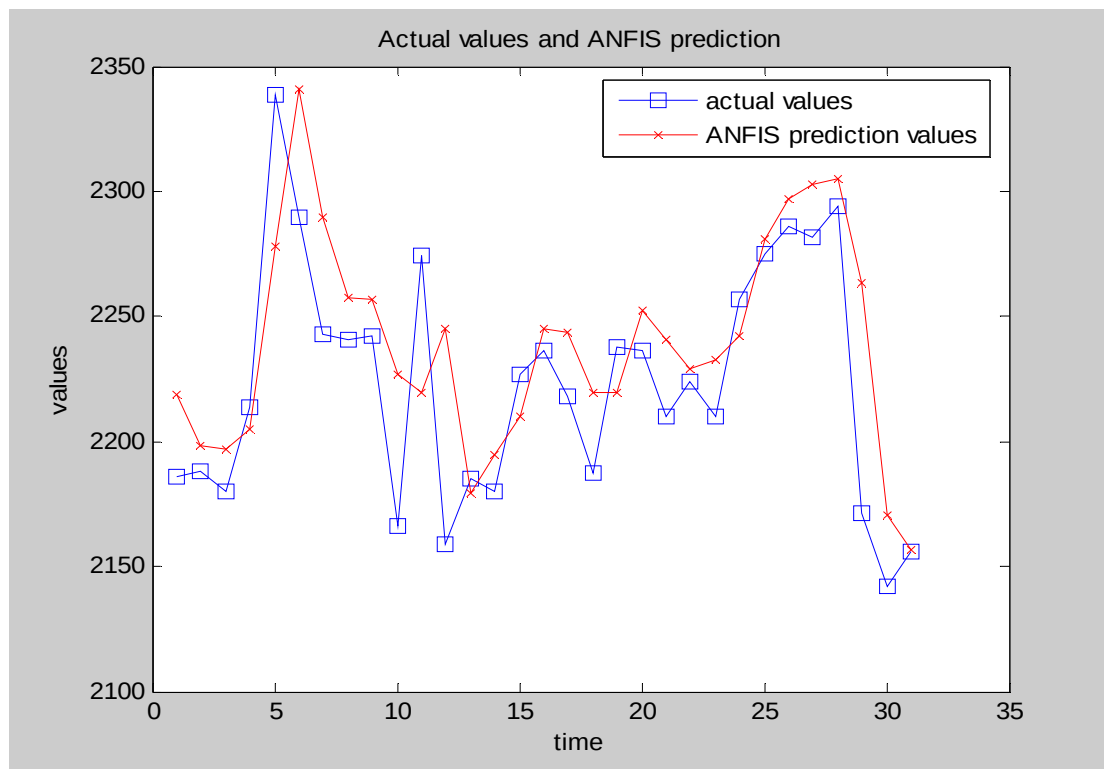
(σχήμα 4.4.7): Επιφάνεια εισόδου-εξόδου πώλησης επίπλων



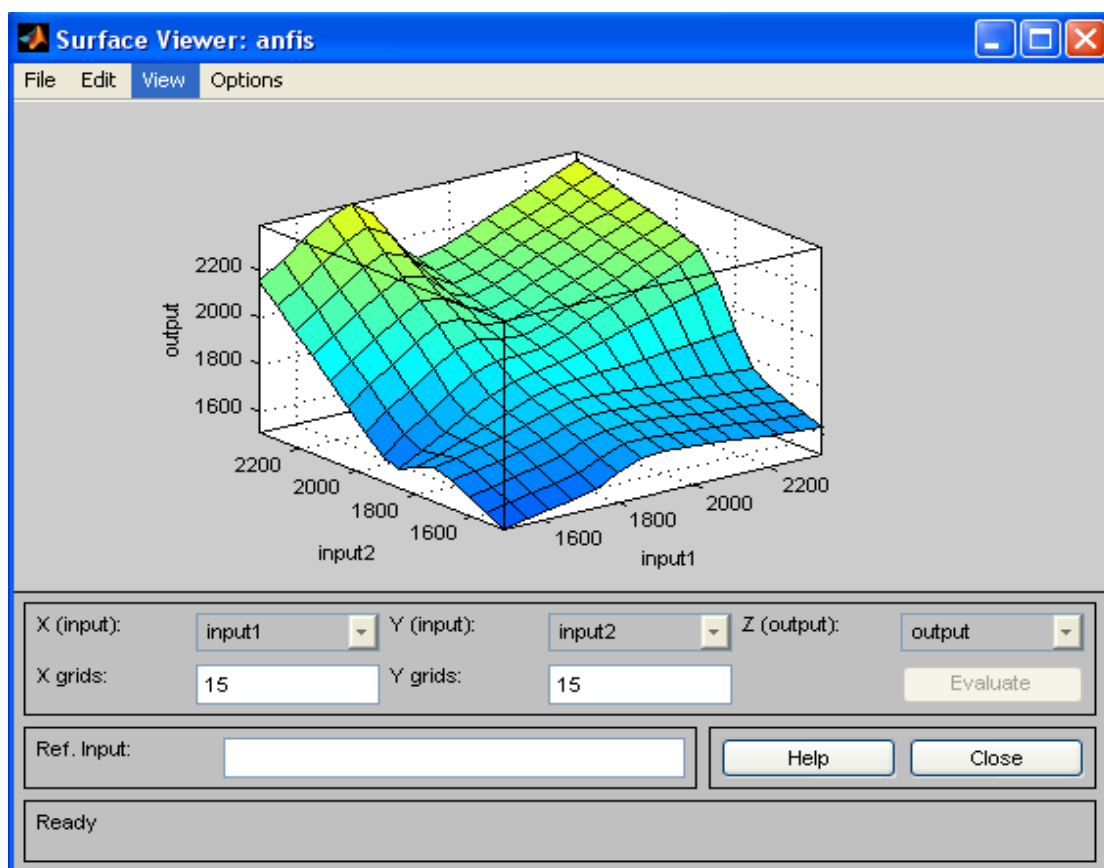
(σχήμα 4.4.8): Σύγκριση πραγματικών τιμών και τιμών πρόβλεψης που αφορά πωλήσεις βενζίνης



(σχήμα 4.4.9): Επιφάνεια εισόδου-εξόδου πώλησης βενζίνης



(σχήμα 4.4.10): Σύγκριση πραγματικών τιμών και τιμών πρόβλεψης που αφορά πωλήσεις υποδημάτων



(σχήμα 4.4.11): Επιφάνεια εισόδου-εξόδου πώλησης υποδημάτων

5. ΣΥΜΠΕΡΑΣΜΑΤΑ

Όπως αναφέρθηκε και σε προηγούμενη παράγραφο, το μοντέλο μας χρησιμοποιεί ένα τμήμα των δεδομένων μας (in sample) για την εκπαίδευση του και ένα τμήμα (out of sample) για τον έλεγχο των τιμών πρόβλεψης. Στην περίπτωση της μελέτης μας, το ANFIS χρησιμοποιεί τις 160 πρώτες μετρήσεις για να εκπαιδευτεί και τις υπόλοιπες 40 για τον έλεγχο. Για να γίνει αξιολόγηση της αποτελεσματικότητας του μοντέλου, θα εστιάσουμε την προσοχή μας στα εξαγόμενα διαγράμματα της Matlab όπως επίσης και στις τιμές σφάλματος που προέκυψαν μετά το πέρας της διαδικασίας. Ο σημαντικότερος δείκτης είναι το RMSE (Root Mean Square Error), η ρίζα δηλαδή του μέσου τετραγωνικού σφάλματος. Το μέσο τετραγωνικό σφάλμα MSE (Mean Square Error) προκύπτει παίρνοντας 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.

Από τον πίνακα 4.4.1 των αποτελεσμάτων του ANFIS, γίνεται αντιληπτό ότι σε κάποιους κλάδους τα αποτελέσματα της πρόβλεψης ήταν πολύ ακριβή και σε κάποιους άλλους λιγότερο. Είναι σαφές, ότι στην περίπτωση πρόβλεψης πωλήσεων υπολογιστών και λογισμικού, το μοντέλο μας απέδωσε τα μέγιστα με RMSE μόλις 11.11. Η επιτυχία του φαίνεται και στο διάγραμμα σύγκρισης πραγματικών τιμών με τις τιμές πρόβλεψης (σχήμα 4.4.2). Βλέπουμε τις τιμές να είναι πάρα πολύ κοντά κατά μήκος του διαγράμματος και μάλιστα κάποιες εξ' αυτών είναι σχεδόν ταυτόσημες. Επίσης, από το σχήμα 4.4.3, προκύπτει μια πολύ ομαλή καμπύλη δεδομένων εισόδου-εξόδου. Αυτό σημαίνει ότι τα δεδομένα εισόδου είναι ικανοποιητικά ώστε να δώσουν μία καλή πρόβλεψη. Η ανωμαλία της καμπύλης στο ένα της άκρο, σχετίζεται με τη διασπορά των δεδομένων μας.

Γενικότερα μέσα από αυτή την εργασία δείξαμε ότι η δυναμική των καθημερινών πωλήσεων ενός προϊόντος μπορεί να συλληφθεί από ένα σύστημα ANFIS και μπορεί να μας βοηθήσει για να προβλέψουμε τη βραχυπρόθεσμη τάση των πωλήσεων. Τα αποτελέσματα είναι πολύ ενθαρρυντικά και έχουν δείξει ότι το μοντέλο ANFIS αποδίδει καλύτερα από τις άλλες συμβατικές μεθόδους πρόβλεψης. Αυτό, με τη σειρά του, σημαίνει ότι το μοντέλο ANFIS μπορεί επιτυχώς να χρησιμοποιηθεί στον οικονομικό στρατηγικό προγραμματισμό, δεδομένου ότι η εφαρμογή των τεχνικών πρόβλεψης στις βιομηχανικές και εμπορικές αγορές έχει γίνει ένα βασικό μέρος στον οικονομικό στρατηγικό προγραμματισμό κατά τη διάρκεια των τελευταίων ετών.

6. ΒΙΒΛΙΟΓΡΑΦΙΑ

- Steven P. Schnaars “A comparison of extrapolation models on yearly sales forecasts” *International Journal of Forecasting* 2 (1986) 71-85, North-Holland
- R.J. Kuo, K.C. Xue “A decision support system for sales forecasting through fuzzy neural networks with asymmetric fuzzy weights” *Decision Support Systems* 24 1998 105–126
- Pei-Chann Chang, Chen-Hao Liu, Robert K. Lai “A fuzzy case-based reasoning model for sales forecasting in print circuit board industries”. *Expert Systems with Applications* (2007)
- Xiao Jiwei, Wu Yaohua, Wang Qian, Liao Li and Hu Hongchun “A Hybrid Sales Forecasting Method Based on Stable Seasonal Pattern Models and BPNN”. *International Conference on Automation and Logistics* August 18 - 21, 2007, Jinan, China
- Heidi Winklhofer , Adamantios Diamantopoulos “A model of export sales forecasting behavior and performance: development and testing”. *International Journal of Forecasting* 19 (2003) 271–285
- R.J. Kuo “A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm”. *European Journal of Operational Research* 129 (2001) 496-517
- Hidetaka Sakai, Hideki Nakajima, Minoru Higashihara, Masashi Yasuda, Masato Oosumi “Development of a fuzzy sales forecasting system for vending machines”. *Computers & Industrial Engineering* 36 (1999) 427- 449
- James B. Boulden “Fitting the Sales Forecast to Your Firm”
- V. Kodogiannis and A. Lolis “Forecasting Financial Time Series using Neural Network and Fuzzy System-based Techniques”. *Neural Comput & Applic* (2002) 11:90–102
- Michael S Morgan, Pradeep K Chintagunta “Forecasting restaurant sales using self-selectivity models”. *Journal of Retailing and Consumer Services*, Vol. 4, N o. 2, pp. 117-128, 1997
- Ralph Snyder “Forecasting sales of slow and fast moving inventories”. *European Journal of Operational Research* 140 (2002) 684–699
- Frank M. Thiesing and Oliver Vornberger “Forecasting Sales Using Neural Networks”. *Department of Mathematics and Computer Science University of Osnabrück*
- D. SHI, C. QUEK , R. TILANI and J. FU “Product Demand Forecasting with a Novel Fuzzy CMAC”. *Neural Processing Letters* (2007) 25:63–78
- Blair Little and Roger A. More “Sales Forecast Errors for New Product Projects”. *Industrial Marketing Management* 7, 49-53 (1978) 49
- Pei-Chann Chang, Chin-Yuan Fan, Baw-Jhiune Liu, and Cycer Chen “Sales Forecasting for TFT-LCD Products with a Hybrid Neural Network Model”. *ICIC 2007, CCIS 2*, pp. 704–713, 2007.
- Edward W. Frees, Thomas W. Miller “Sales forecasting using longitudinal data models”. *International Journal of Forecasting* 20 (2004) 99–114
- J. Scott Armstrong , Vicki G. Morwitz , V. Kumar “Sales forecasts for existing consumer products and services: Do purchase intentions contribute to accuracy? *International Journal of Forecasting* 16 (2000) 383–397

- Barry L. Bayus, Saman Hong and Russell P. Labe, Jr. "Developing and Using Forecasting Models of Consumer Durables: The case of color television". J PROD INNOV MAGAZINE 1989:6 5- 19
- Anubhav Tewari "Sales Prediction for Franchise Stores Using Artificial Neural Networks".
- James T.Rothe "Effectiveness of sales forecasting methods". Industrial Marketing Management 7, 114-118 (1978)
- Wen-Pai, Wang, Ze Chen "A neuro-fuzzy based forecasting approach for rush order control applications". Expert Systems with Applications: An International Journal Volume 35 , Issue 1-2 (July 2008) Pages 223-234
- Marx-Gomez J., Rautenstrauch C., Nurnberger A., Kruse R. "Neuro-fuzzy approach to forecast returns of scrapped products to recycling and remanufacturing". Knowledge-Based Systems, Volume 15, Number 1, January 2002 , pp. 119-128(10)
- Ilan Alon, Min Qi, Robert J. Sadowski "Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods". Journal of Retailing and Consumer Services, Volume 8, Issue 3, May 2001, Pages 147-156
- Roselina Sallehuddin, Siti Mariyam Hj. Shamsuddin, Siti Zaiton Mohd. Hashim, Ajith Abraham "Forecasting Time Series Data Using Hybrid Grey Relational Artificial Neural Network and Auto Regressive Integrated Moving Average Model". Neural Network World 6/07, 573-605
- Sebastien Tomassey, Michel Happiette "A neural clustering and classification system for sales forecasting of new apparel items". Applied Soft Computing, Volume 7 , Issue 4 (August 2007), Pages 1177-1187
- Zhan-Li Sun, Tsan-Ming Choi, Kin-Fan Au, Yong Yu "Sales forecasting using extreme learning machine with applications in fashion retailing". Decision Support Systems, Volume 46, Issue 1, December 2008, Pages 411-419
- Ching-Wu Chu, Guoqiang Peter Zhang "A comparative study of linear and nonlinear models for aggregate retail sales forecasting". International Journal of Production Economics, Volume 86, Issue 3, 11 December 2003, Pages 217-231
- James T. Luxhslj, Jens O. Riisb, Brian Stensballe "A hybrid econometric-neural network modeling approach for sales forecasting". Int. J. Production Economics 43 (1996) 175-192
- Sameh Ghwanmeh, Riyadh Al-Shalabi, Ghassan Kana, Luai Alnemi "Enhanced Neural Networks Model Based on a Single Layer Linear Couterpropagation for Prediction and Function Approximation". Information Technology Journal 5(1): 45-50, 2006 ISSN 1812-5638
- Ebrahim Abbasi, and Amir Abouec "Stock Price Forecast by Using Neuro-Fuzzy Inference System". World Academy of Science, Engineering and Technology 46 2008
- F. Owen Irvine "Models of Sales Expectations". Engineering Costs and Production Economics, 19 (1990) 37-46
- Nikola K. Kasabov "Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering".