

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ



Πολυτεχνείο Κρήτης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΘΕΜΑ : ΣΥΓΚΡΙΣΗ ΠΡΟΒΛΕΨΕΩΝ
ΠΑΡΑΔΟΣΙΑΚΩΝ ΚΑΙ ΝΕΩΝ ΤΕΧΝΟΛΟΓΙΩΝ
ΠΗΓΩΝ ΕΝΕΡΓΕΙΑΣ.

Επιβλέπων καθηγητής : Ατσαλάκης Γεώργιος
Επιμέλεια : Μωραϊτίδης Θεόδωρος

ΧΑΝΙΑ 2013

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ.....	3
2. ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ.....	4
3. ΠΡΟΤΕΙΝΟΜΕΝΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ.....	16
3.1. Τεχνητά Νευρωνικά Δίκτυα.....	16
Ορισμός νευρωνικού δικτύου.....	16
Ιστορικά Στοιχεία.....	17
Από τα Βιολογικά Νευρωνικά Δίκτυα στα Τεχνητά.....	20
Χρήση Νευρωνικών Δικτύων και Εφαρμογές.....	22
Αρχιτεκτονική Νευρωνικών Δικτύων.....	24
Είδη Αρχιτεκτονικών Δομών.....	28
Λειτουργία Νευρωνικών Δικτύων.....	31
Εκπαίδευση Νευρωνικών Δικτύων.....	32
Ο Αλγόριθμος Μάθησης του Perceptron (Αισθητήρα).....	34
Ο Αλγόριθμος Ελάχιστου Μέσου Τετραγωνικού (EMT) Σφάλματος.....	37
Ο Αλγόριθμος Πίσω Διάδοσης (Π.Δ.) του Λάθους.....	38
3.2. Ασαφής Λογική (Fuzzy Logic)	41
Συναρτήσεις συμμετοχής.....	43
Ιδιότητες των ασαφών συνόλων.....	45
Λογικές πράξεις στα ασαφή σύνολα.....	47
Ασαφείς σχέσεις.....	51
Ασαφής συλλογιστική.....	52
3.3. Συστήματα Ασαφούς Συλλογιστικής-Νευροασαφή Συστήματα.....	54
Κανόνες αν-τότε (if-then rules)	54
Συστήματα ασαφούς συλλογιστικής (fuzzy inference systems)	57
Νευροασαφή συστήματα.....	65
3.4. ANFIS (Adaptive Neuro-Fuzzy Inference System)	70
Εισαγωγή - Ορισμός.....	70
Δομή Αρχιτεκτονικής του ANFIS.....	71
Περιγραφή του μοντέλου.....	74
Αριθμητικό Παράδειγμα Μοντέλου.....	75
4. ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ (case study).....	83
4.1. Εισαγωγή	83
4.2. Επεξεργασία.....	85
4.3. Αποτελέσματα.....	86
5. ΣΥΜΠΕΡΑΣΜΑΤΑ.....	94
6.ΒΙΒΛΙΟΓΡΑΦΙΑ.....	95

1. ΕΙΣΑΓΩΓΗ

Ο άνθρωπος δεν είναι ούτε γρήγορος ούτε ακριβής στους υπολογισμούς του, όπως ένας σύγχρονος υπολογιστής. Σε πολλούς τομείς, όμως, η ανθρώπινη ικανότητα υπερτερεί κατά πολύ των ικανοτήτων ενός υπολογιστή. Για παράδειγμα, ο άνθρωπος αναγνωρίζει πολύ εύκολα αντικείμενα και κατανοεί τις σχέσεις μεταξύ τους μέσα στο φυσικό περιβάλλον έστω και αν αυτά είναι παραμορφωμένα ή δεν είναι εξολοκλήρου ορατά.

Η ικανότητα μάθησης μέσω της εμπειρίας συνεπώς είναι από τα κύρια χαρακτηριστικά της ανθρώπινης ευφυΐας. Η ανθρώπινη μνήμη έχει την ικανότητα να αποθηκεύει μεγάλη ποσότητα και ποικιλία γνώσης και να συσχετίζει πληροφορίες πολύ γρήγορα και χωρίς ιδιαίτερη προσπάθεια. Σε αντίθεση, ο υπολογιστής έχει την ικανότητα να απομνημονεύει τεράστιες ποσότητες πληροφοριών αλλά δυσκολεύεται να τις εκμεταλλευθεί.

Τα τελευταία χρόνια έχουν γίνει πολλές προσπάθειες, ανάπτυξης τεχνολογιών για την κατασκευή μηχανών που θα προσομοιώνουν την λειτουργία του ανθρώπινου εγκεφάλου, συνδυάζοντας έτσι την γρηγοράδα ενός υπολογιστή και τον τρόπο συσχετισμού πληροφοριών του ανθρώπου. Αρκετές από αυτές τις προσπάθειες είχαν ως κατάληξη τεχνολογίες όπως τα νευρωνικά δίκτυα, τα ασαφή συστήματα, τους εξελεγκτικούς αλγορίθμους και την τεχνητή νοημοσύνη.

Στην παρούσα διπλωματική εργασία επιχειρούμε να χρησιμοποιήσουμε μία τεχνική η οποία βασίζεται σε προγραμματισμό με το λογισμικό της MATLAB και συγκεκριμένα στην πρόβλεψη των καθημερινών μετρήσεων για την παραγωγή ενέργειας από οχτώ (8) διαφορετικές τεχνολογίες χωρίζοντας τις σε δύο ομάδες, νέες και παραδοσιακές και στη συνέχεια συγκρίνοντας τις, χρησιμοποιώντας τις μετρήσεις προηγούμενων ετών. Χρησιμοποιούμε 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.

2. ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

Είναι σημαντικό να ερευνήσουμε τις προσπάθειες που έχουν γίνει μέχρι τώρα, 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.

Energy Consumption Forecasting of Iran Using Recurrent Neural Networks.

Στη παρούσα μελέτη των Boroushaki και Avami προτείνεται ένα ευφύες μοντέλο που χρησιμοποιεί ένα επαναλαμβανόμενο νευρωνικό δίκτυο για την πρόβλεψη κατανάλωσης ενέργειας στο Ιράν έχοντας σα δεδομένα εισόδου τον πληθυσμό της χώρας και το ακαθάριστο εγχώριο προϊόν. Το συγκεκριμένο επαναλαμβανόμενο νευρωνικό δίκτυο έχει μη γραμμική συνάρτηση ενεργοποίησης.

Η κατασκευή του μοντέλου πρόβλεψης έγινε χρησιμοποιώντας δεδομένα της περιόδου 1976-2001 και η αποτελεσματικότητά του δοκιμάστηκε με δεδομένα της περιόδου 2002-2006.

Προσομοιώσεις που πραγματοποιήθηκαν έδειξαν ότι αυτό το μοντέλο μπορεί να προβλέψει την κατανάλωση ενέργειας στο Ιράν με ακρίβεια και με αποκλίσεις που είναι αποδεκτές. Τέλος αναφέρεται ότι το μοντέλο πραγματοποιεί ετήσιες προβλέψεις.

Influence of artificial neural network structure on quality of short-term electric energy consumption forecast.

Στις μέρες μας η πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας είναι πολύ σημαντική για τις εταιρείες διανομής ηλεκτρικής ενέργειας και τις βιομηχανικές επιχειρήσεις. Η ακρίβεια των προβλέψεων εξασφαλίζει την ελαχιστοποίηση του κόστους για αυτές.

Η εργασία των Parol και Baczynski σχολιάζει κατά πόσο επηρεάζεται η δομή ενός τεχνητού νευρωνικού δικτύου, σε σχέση με την ποιότητα των βραχυπρόθεσμων προβλέψεων κατανάλωσης ηλεκτρικής ενέργειας από μία μέθοδο που βασίζεται σε αυτό. Οι συγγραφείς αναφέρουν την ύπαρξη πολλών διαφορετικών μεθόδων για την πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας και τονίζουν ότι οι ακριβέστερες προβλέψεις γίνονται από μεθόδους που βασίζονται σε τεχνητά νευρωνικά δίκτυα. Γίνεται επίσης στην έρευνα μία εκτεταμένη προσπάθεια να διερευνηθεί η συσχέτιση της ποιότητας των μεθόδων προβλέψεων που βασίζονται σε τεχνητά νευρωνικά δίκτυα με τη δομή (αριθμός των νευρώνων εισόδου και εξόδου και των υπολογιστικών ή κρυμμένων νευρώνων) και την ικανότητα μάθησης των αντίστοιχων νευρωνικών δικτύων.

Τέλος αναφέρεται ότι έγιναν δοκιμές με δεδομένα που πάρθηκαν από εταιρείες διανομής ηλεκτρικής ενέργειας και τα αποτελέσματα έδειξαν ότι η δομή των τεχνητών νευρωνικών δικτύων έχει άμεση επίδραση στην ποιότητα των προβλέψεων κατανάλωσης ηλεκτρικής ενέργειας.

Transport energy demand forecast using multi-level genetic programming.

Στη μελέτη των Doustmohammadi, Hasanzadeh Samira, Shakouri G και Forouzanfar Mehdi παρουσιάζεται ένα μοντέλο πολυεπίπεδου γενετικού προγραμματισμού για την πρόβλεψη κατανάλωσης ενέργειας στις μεταφορές στο Ιράν. Η μελέτη έδειξε ότι η μέθοδος πραγματοποιεί προβλέψεις με μικρότερο σφάλμα, σε σχέση με μοντέλα προβλέψεων που χρησιμοποιούν νευρωνικά δίκτυα και μοντέλα (τεχνικών) ασαφούς (fuzzy) γραμμικής παλινδρόμησης. Το μοντέλο κατασκευάστηκε χρησιμοποιώντας δεδομένα κατανάλωσης ενέργειας των μεταφορών στο Ιράν και οι προβλέψεις βασίζονται σε τρεις παραμέτρους: το ακαθάριστο εγχώριο προϊόν, τον πληθυσμό και τον αριθμό των μεταφορικών μέσων. Ο όρος "πολυεπίπεδος γενετικός προγραμματισμός" αναφέρεται στην χρήση γενετικού προγραμματισμού από το μοντέλο σε δύο διαφορετικά επίπεδα για την πραγματοποίηση προβλέψεων. Στο πρώτο επίπεδο εισάγονται στο μοντέλο δεδομένα για τις τρεις παραμέτρους που έχουν συλλεχθεί και με βάση τα δεδομένα αυτά γίνονται από το μοντέλο προβλέψεις για τις παραμέτρους.

Στο δεύτερο επίπεδο το μοντέλο κάνει πρόβλεψη κατανάλωσης ενέργειας στις μεταφορές χρησιμοποιώντας καταγεγραμμένα δεδομένα της κατανάλωσης ενέργειας σε συνδυασμό με τις προβλέψεις που πραγματοποιήθηκαν στο πρώτο στάδιο. Η κατασκευή του μοντέλου έγινε χρησιμοποιώντας ιστορικά δεδομένα της περιόδου 1968-2002 και με δεδομένα της περιόδου 2003-2005 δοκιμάστηκε η αποτελεσματικότητά του.

Application of SOM neural networks to short-term load forecasting: The Spanish electricity market case study.

Η πρόβλεψη ζήτησης ηλεκτρικής ενέργειας με χρήση των νευρωνικών δικτύων είναι ένα θέμα για το οποίο έχουν γίνει πολλές έρευνες την τελευταία δεκαετία. Οι χάρτες αυτοοργάνωσης (Self-Organizing Maps, SOM) είναι ένας τύπος νευρωνικού δικτύου που έχει χρησιμοποιηθεί σχεδόν ελάχιστα για την πρόβλεψη της ζήτησης ηλεκτρικής ενέργειας. Η έρευνα που έγινε από τους Gabaldon, Aparicio, Senabre, Valero και López παρουσιάζει ένα μοντέλο πρόβλεψης με βάση το συγκεκριμένο είδος νευρωνικών δικτύων. Ο σκοπός του μοντέλου δεν είναι μόνο να αποδείξει ότι οι χάρτες αυτοοργάνωσης μπορούν να χρησιμοποιηθούν για την πρόβλεψη της ζήτησης ηλεκτρικής ενέργειας, αλλά και να προβάλει τη δυνατότητα της εφαρμογής τους να κάνουν ανάλυση των προβλέψεων, βασιζόμενη σε πραγματικά δεδομένα. Τα δεδομένα που χρησιμοποιούνται από το μοντέλο για να πραγματοποιηθεί η πρόβλεψη είναι η κατανάλωση ενέργειας και μετεωρολογικές μετρήσεις της Ισπανίας για την περίοδο 2001-2010. Στη μελέτη γίνεται εκτεταμένη αναφορά στη σπουδαιότητα των μετρολογικών μεταβλητών για τις προβλέψεις του μοντέλου αλλά και για το πως σχετίζονται μεταξύ τους οι μεταβλητές αυτές, όταν σαν δεδομένα εισόδου του μοντέλου χρησιμοποιούνται πρόσφατες τιμές της ζήτησης ενέργειας. Το μοντέλο κατασκευάστηκε για την πρόβλεψη της ζήτησης ενέργειας σε καθημερινή βάση. Τέλος αναφέρεται ότι οι προβλέψεις αυτές προορίζονται για δεδομένα εισόδου σε ένα σύστημα υποστήριξης αποφάσεων, που χρησιμοποιείται από διάφορες εταιρείες της ισπανικής αγοράς.

Nuclear mass systematics using neural networks.

Ραδιενέργεια είναι το φαινόμενο της εκπομπής σωματιδίων ή ηλεκτρομαγνητικής ακτινοβολίας από τους πυρήνες ορισμένων χημικών στοιχείων, που γι' αυτό το λόγο ονομάζονται ραδιενεργά. Τα τελευταία χρόνια παρατηρείται αύξηση των μελετών για την ανάπτυξη στατιστικών μοντέλων πρόβλεψης, που βασίζονται σε νευρωνικά δίκτυα πρόσθιας τροφοδότησης, με σκοπό την πρόβλεψη της ατομικής μάζας ραδιενεργών στοιχείων. Ο αριθμός αυτός δηλώνει την κατά μέσο όρο μάζα του ατόμου του στοιχείου μετρημένη σε μονάδες ατομικής μάζας (u) ή τη μάζα του στοιχείου μετρημένη σε γραμμάρια που περιέχεται σε ένα mol ατόμων του στοιχείου (g/mol).

Στόχος των μελετών αυτών είναι η κατασκευή μοντέλων πρόβλεψης, με γνωστά δεδομένα την μάζα του πυρήνα ενός ατόμου κάποιου στοιχείου και τον αριθμό των πρωτονίων και των νετρονίων του. Ένας άλλος στόχος των μελετών είναι η κατασκευή αξιόπιστων μοντέλων για την πρόβλεψη της ατομικής μάζας για στοιχεία που δεν ανήκουν στην κοιλάδα σταθερότητας. Στην παρούσα εργασία των Athanassopoulos, Mavrommatis, Gernoth και Clark προτείνεται ένα μοντέλο πρόβλεψης της ατομικής μάζας ραδιενεργών στοιχείων, με γνωστά δεδομένα την μάζα του πυρήνα ενός ατόμου κάποιου στοιχείου και τον αριθμό των πρωτονίων και των νετρονίων του που βασίζεται σε ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης. Αναφέρεται επίσης ότι το μοντέλο αυτό πραγματοποιεί σημαντικά ακριβέστερες προβλέψεις σε σχέση με παλιότερα μοντέλα προβλεψής της ατομικής μάζας ραδιενεργών στοιχείων, που βασίζονται σε διαφορετικά είδη νευρωνικών δικτύων.

Short-term load forecast using trend information and process reconstruction.

Η σταδιακή απελευθέρωση της αγοράς ηλεκτρικής ενέργειας έχει ως αποτέλεσμα την κατάργηση των μονοπώλιων της παραγωγής, μεταφοράς και διανομής της. Στα πλαίσια του ανταγωνισμού των εταιρειών προμήθειας ηλεκτρικής ενέργειας, σημαντικός παράγοντας καθίσταται η πρόβλεψη ζήτησης ηλεκτρικής ενέργειας. Μοντέλα βραχυπρόθεσμης πρόβλεψης έχουν κατασκευαστεί και χρησιμοποιούνται για τη βελτίωση της λειτουργίας και του σχεδιασμού κατασκευής συστημάτων ηλεκτρικής ενέργειας. Οι Santos, Martins, Pires, Martins και Mendes στην εργασία τους αναφέρουν τέτοια μοντέλα που είναι βασισμένα σε τεχνητά νευρωνικά δίκτυα και εστιάζουν στην σημαντικότητα της κατάλληλης επιλογής των μεταβλητών εισόδων τους. Οι συγγραφείς παρουσιάζουν τον σχεδιασμό ενός μοντέλου πρόβλεψης βασισμένο σε ένα νευρωνικό δίκτυο.

Στη πρώτη φάση της κατασκευής πραγματοποιήθηκε η εκπαίδευση του νευρωνικού δικτύου, μέσω μίας επαναληπτικής διαδικασίας σταδιακής προσαρμογής των παραμέτρων του δικτύου. Αναφέρεται ότι η μέθοδος εκπαίδευσης που χρησιμοποιήθηκε ήταν η στοχαστική μάθηση.

Στο δεύτερο στάδιο της κατασκευής του μοντέλου, πραγματοποιήθηκαν προσομοιώσεις για το μοντέλο, με δεδομένα εισόδου τις τιμές της κατανάλωσης προηγούμενων ημερών που είχαν αντίστοιχες καιρικές συνθήκες. Το αποτέλεσμα ήταν η δημιουργία ενός μοντέλου πρόβλεψης που δεν είχε άμεση συσχέτιση με τις καιρικές συνθήκες.

Using adaptive network based fuzzy inference system to forecast regional electricity loads.

Τα τελευταία χρόνια έχουν αναπτυχθεί διάφορες μέθοδοι για την πρόβλεψη της ζήτησης ηλεκτρικής ενέργειας, με σκοπό την βελτίωση της απόδοσης της διαχείρισης της ηλεκτρικής βιομηχανίας

Στην παρούσα μελέτη που πραγματοποιήθηκε από τους Pan και Ying, χρησιμοποιήθηκε ένα μοντέλο πρόβλεψης, βασισμένο στο νεύρο-ασαφές δίκτυο ANFIS, για την πρόβλεψη της ζήτησης ηλεκτρικής ενέργειας στην Ταϊβάν.

Αποδείχθηκε ότι οι προβλέψεις του μοντέλου ήταν πολύ ακριβείς.

Μεταξύ μοντέλων παλινδρόμησης, μηχανών διανυσματικής υποστήριξης γενετικών αλγορίθμων, τεχνητών νευρωνικών δικτύων, μοντέλων HEFST (hybrid ellipsoidal fuzzy systems for time series) και του μοντέλου ANFIS, πραγματοποιήθηκε σύγκριση για αποτελεσματικότητα των προβλέψεων τους, η οποία βασίστηκε σε στατιστικά αποτελέσματα και μέσα απόλυτα ποσοστιαία σφάλματα και καταδείχτηκε ότι οι προβλέψεις του τελευταίου είναι οι πιο ακριβείς.

Το αποτέλεσμα της παρούσας εργασίας ήταν να ορίσει το μοντέλο ANFIS ως μια πολλά υποσχόμενη εναλλακτική λύση για την πρόβλεψη της ζήτησης ηλεκτρικής ενέργειας.

The Prediction of Under saturated Crude Oil Viscosity: An Artificial Neural Network and Fuzzy Model Approach.

Το ιξώδες είναι η αντίσταση που παρουσιάζει κατά τη ροή της η ύλη και κυρίως τα υγρά και τα αέρια. Για τη μεταφορά του πετρελαίου μέσα από αγωγούς και πορώδη υλικά, το ιξώδες είναι από τις σημαντικότερες παραμέτρους και ο υπολογισμός του καθίσταται καθοριστικός. Αυτό συνεπάγεται ότι η χρήση μίας υψηλής ακρίβειας μεθόδου πρόβλεψης του ιξώδους του πετρελαίου να είναι σημαντική σε διάφορες συνθήκες λειτουργίας του. Μέχρι σήμερα, διάφορες εμπειρικές συσχετίσεις έχουν προταθεί για την πρόβλεψη του ιξώδους του πετρελαίου χωρίς όμως να θεωρούνται ικανοποιητικοί. Με στόχο την κατασκευή δύο πειραματικών μοντέλων πρόβλεψης για το ιξώδες του πετρελαίου, λήφθηκαν δείγματα πετρελαίου από δεξαμενές στο Ιράν και πραγματοποιήθηκαν πειραματικές μετρήσεις για να εξαρθούν δεδομένα για το ιξώδες σε κάθε περίπτωση.

Κατασκευάστηκε ένα μοντέλο βασισμένο σε νευρωνικά δίκτυα και ένα βασισμένο στις αρχές της ασαφούς λογικής.

Η εγκυρότητα και η ακρίβεια των προβλέψεων που λαμβάνονται από τα δύο μοντέλα κρίνεται ικανοποιητική, με συγκρίσεις που πραγματοποιήθηκαν μεταξύ των προβλέψεων και δειγμάτων που λήφθηκαν από δεξαμενές στο Ιράν.

A Neuro-fuzzy-stochastic frontier analysis approach for long-term natural gas consumption forecasting and behavior analysis: The cases of Bahrain, Saudi Arabia, Syria, and UAE.

Στο παρόν άρθρο οι Sheikalishahi, Tajvidi, Nadimi, Saberi και Asadzadeh παρουσιάζουν ένα μοντέλο πρόβλεψης για τη μακροπρόθεσμη κατανάλωση φυσικού αερίου σε μία χώρα, βασισμένο στο νευρο-ασαφές δίκτυο ANFIS. Παρουσιάζεται επίσης ότι χρησιμοποιείται η παραμετρική μέθοδος με στοχαστική εν δυνάμει ανάλυση (SFA) για να εξεταστεί η τάση της κατανάλωσης φυσικού αερίου στο παρελθόν αλλά και να αναλυθεί η πιθανή κλίμακα που θα κυμαίνεται στο μέλλον. Οι μεταβλητές εισόδου του μοντέλου είναι το ακαθάριστο εγχώριο προϊόν και ο πληθυσμός μίας χώρας. Με σκοπό την καλύτερη απόδοση του μοντέλου, γίνεται επεξεργασία των αριθμητικών δεδομένων εισόδου του και χρησιμοποιείται μια σταθερή κλίμακα για τις μονάδες των τιμών τους. Στη συνέχεια τα αποτελέσματα εξόδου (προβλέψεις) του, επεξεργάζονται έτσι ώστε να ανάγονται στις αρχικές μονάδες. Η εγκυρότητα και η αποτελεσματικότητα του μοντέλου ελέγχθηκαν με μετρήσεις της κατανάλωσης του φυσικού αερίου στο Μπαχρέιν, τη Σαουδική Αραβία, τη Συρία, και στα Ηνωμένα Αραβικά Εμιράτα για την χρονική περίοδο 1980-2007. Τα αποτελέσματα έδειξαν ότι το μοντέλο είναι σε θέση να κάνει ακριβείς προβλέψεις σε σχέση με την πολυπλοκότητα, την αβεβαιότητα και την τυχαιότητα του υπολογισμού της κατανάλωσης του φυσικού αερίου. Οι συγγραφείς κάνουν λόγο επίσης για τη χρήση ενός αυτοπαλίνδρομου μοντέλου ώστε να υπολογιστούν τιμές του ακαθάριστου εγχώριου προϊόντος και του πληθυσμού για την περίοδο 2008-2015 για τις προαναφερθείσες χώρες. Οι τιμές που υπολογίστηκαν χρησιμοποιήθηκαν σε δεδομένα εισόδου του μοντέλου πρόβλεψης και έγινε πρόβλεψη της κατανάλωσης φυσικού αερίου στις παραπάνω χώρες για αυτήν την περίοδο.

Artificial Neural Network based Hydro Electric Generation Modelling.

Ο υδροηλεκτρισμός είναι η ηλεκτρική ενέργεια που παράγεται από υδατοπτώσεις. Υδροηλεκτρικά εργοστάσια καλούνται οι εγκαταστάσεις παραγωγής ηλεκτρικής ενέργειας που εκμεταλλεύονται την δυναμική ενέργεια του νερού. Από υδροηλεκτρικά εργοστάσια είναι εφικτό να παραχθεί ηλεκτρική ενέργεια από νερό που συγκρατείται σε τεχνητές λίμνες από φράγματα που κατασκευάζονται. Τα υδροηλεκτρικά εργοστάσια αποτελούνται από τα εξής τμήματα: τουρμπίνα, γεννήτριες και γραμμές μεταφοράς. Κατά μήκος του ποταμού Τζαμούνα στην Ινδία με κατεύθυνση προς το Μπαγκλαντές, έχει κατασκευαστεί ένα υδροηλεκτρικό εργοστάσιο μεγάλης κλίμακας το οποίο παράγει έως και 120 MW ισχύος. Στο άρθρο των Sharma και Yadav παρουσιάζεται η κατασκευή ενός μοντέλου πρόβλεψης βασισμένο σε τεχνητό νευρωνικό δίκτυο με σκοπό να προβλέπει τη παραγωγή ενέργειας από το υδροηλεκτρικό εργοστάσιο στον ποταμό Τζαμούνα. Με χρήση αλγορίθμων οπισθοδρόμησης πραγματοποιήθηκε έλεγχος της ακρίβειας των προβλέψεων και του μέσου τετραγωνικού σφάλματος του μοντέλου.

Hybrid energy system evaluation in water supply systems: artificial neural network approach and methodology.

Ένα δίκτυο ύδρευσης αποτελείται από αγωγούς, αντλιοστάσια και δεξαμενές πίεσης. Για τη λειτουργία των αντλιοστασίων και των δεξαμενών πίεσης είναι απαραίτητη η κατανάλωση μεγάλων ποσών ενέργειας, γεγονός που καθιστά προτεραιότητα των συστημάτων παροχής νερού την βέλτιστη ενεργειακή απόδοσή τους. Η μελέτη των Gonçalves και Ramos παρουσιάζει μία μέθοδο βασισμένη σε ένα νευρωνικό δίκτυο με σκοπό την βελτιστοποίηση του ελέγχου πίεσης και της υδραυλικής συμπεριφοράς ενός υβριδικού συστήματος παροχής νερού. Για την εκπαίδευση του τεχνητού νευρωνικού δικτύου στην οποία στηρίχτηκε η μέθοδος, χρησιμοποιήθηκαν παράμετροι του δικτύου από προσομοιώσεις που πραγματοποιήθηκαν για τη κατανάλωση ενέργειας, τις οικονομικές και τις υδραυλικές απαιτήσεις του συστήματος. Τα αποτελέσματα του μοντέλου που παρουσιάστηκε, κρίθηκαν τόσο ικανοποιητικά, που το καθιστούν ένα σημαντικό εργαλείο στον τομέα της υποστήριξης αποφάσεων για υδροηλεκτρικά συστήματα.

Upgrade of an artificial neural network prediction method for electrical consumption forecasting using an hourly temperature curve model.

Η εργασία των Roldán-Blay, Escrivá-Escrivá, Álvarez-Bel, Roldán-Porta, Rodríguez-García (Energy & Buildings. May2013) παρουσιάζει μία μέθοδο (EUs method) που αναπτύχθηκε από τους παραπάνω για την πρόβλεψη κατανάλωσης ενέργειας ενός κτιρίου. Το μοντέλο χρησιμοποιεί στατιστικά δεδομένα θερμοκρασίας-χρόνου για να κάνει την πρόβλεψη. Από τη μέθοδο περιλαμβάνεται χρήση τεχνητών νευρωνικών δικτύων. Η κατανάλωση ενέργειας του κτιρίου γίνεται συνάρτηση των εξωτερικών μεταβλητών που επικρατούν στο περιβάλλον και η πρόβλεψη πραγματοποιείται από τη μέθοδο, λαμβάνοντας αυτές τις μεταβλητές ως δεδομένα. Οι παράμετροι που επηρεάζουν την κατανάλωση είναι οι καιρικές συνθήκες και η ώρα. Για τη χρήση της μεθόδου απαιτούνται ιστορικά στοιχεία της συνολικής κατανάλωσης και οι τιμές μέγιστης και ελάχιστης ημερήσιας θερμοκρασίας. Η μέθοδος έχει δοκιμαστεί με επιτυχία για την πρόβλεψη της κατανάλωσης του πολυτεχνείου της Βαλένθιας για έναν χρόνο.

Overview of U.S. Coal Production and Consumption.

Στο άρθρο από το Congressional Digest Jan2013, εξετάζεται η βιομηχανία άνθρακα των ΗΠΑ, από το 2012. Σημειώνεται ότι το μεγαλύτερο μέρος του άνθρακα που χρησιμοποιείται στις ΗΠΑ πηγαίνει στην παραγωγή ηλεκτρικής ενέργειας, και προβλέπεται η διακοπή λειτουργίας παλιότερων μονάδων ηλεκτροπαραγωγής με καύσεις ρυπογόνου άνθρακα. Αναφέρονται επίσης παράγοντες όπου θα διαμορφώσουν τη ζήτηση των ΗΠΑ για τον άνθρακα, συμπεριλαμβανομένων των περιβαλλοντικών κανονισμών, τη ζήτηση ηλεκτρικής ενέργειας, και τις τιμές των ανταγωνιστικών μορφών ενέργειας, κυρίως φυσικού αερίου. Τέλος δίνονται προβλέψεις για τη βιομηχανία του άνθρακα.

Forecasting of Coal Consumption Using an Artificial Neural Network and Comparison with Various Forecasting Techniques.

Η πρόβλεψη της κατανάλωσης ενέργειας είναι σημαντική για κάθε χώρα, για να υπολογίσει τη μελλοντική ενεργειακή της ζήτηση και η εκάστοτε κυβέρνηση να διαμορφώσει ανάλογα τη στρατηγική της πολιτικής της. Στο παρόν άρθρο των Goic, Iniyan και Jebaraj σχολιάζεται η χρήση άνθρακα σε διαφόρους τομείς στην Ινδία, όπως μεταφορές, οικιακή χρήση και παραγωγή ενέργειας. Γίνεται επίσης αναφορά στην διαμόρφωση διαφόρων μοντέλων πρόβλεψης για την κατανάλωση του άνθρακα. Οι συγγραφείς προτείνουν ένα μοντέλο πρόβλεψης για την κατανάλωση άνθρακα που βασίζεται σε ένα τεχνητό νευρωνικό δίκτυο, το οποίο έχει καλύτερα αποτελέσματα από κλασικά μοντέλα πρόβλεψης. Σκοπός της εργασίας είναι η παρουσίαση αυτού του μοντέλου και η σύγκριση των προβλέψεων για την πρόβλεψη κατανάλωσης άνθρακα σε σχέση με τις προβλέψεις από μοντέλα παλινδρόμησης. Για την σύγκριση χρησιμοποιήθηκαν δεδομένα μετρήσεων της συνολικής κατανάλωσης άνθρακα στην Ινδία και διαπιστώθηκε ότι το μοντέλο νευρωνικού δικτύου που προτάθηκε, στις περισσότερες περιπτώσεις παρείχε τις καλύτερες προβλέψεις. Τέλος αναφέρεται ότι έγινε με χρήση του μοντέλου πρόβλεψη της συνολικής κατανάλωσης άνθρακα στην Ινδία για τα έτη 2010, 2020, 2030 και προβλέφθηκε να είναι 695.518, 890.143 και 1.594.844 χιλιάδες τόνους, αντίστοιχα.

Forecasting of Wind Speed Using Artificial Neural Networks

Η αιολική ενέργεια κατέχει πολύ σημαντική θέση στις ανανεώσιμες πηγές ενέργειας όσον αφορά την παραγωγή ηλεκτρικής ενέργειας. Ολοένα και περισσότερα συστήματα μετατροπής αιολικής ενέργειας σε ηλεκτρική εγκαθίστανται ανά τον κόσμο. Μία πολύ σημαντική παράμετρος για την μετατροπή αιολικής ενέργειας από τα συστήματα μετατροπής είναι η ταχύτητα του ανέμου. Αυτό έχει ως αποτέλεσμα, ο σχεδιασμός ενός συστήματος μετατροπής αιολικής ενέργειας σε ηλεκτρική αλλά και η επιλογή της κατάλληλης περιοχής για την κατασκευή του, να επηρεάζονται άμεσα από τις προβλέψεις της ταχύτητας του ανέμου.

Στο άρθρο των Ramesh Babu και Arulmozhivarman εξετάζεται η βραχυπρόθεσμη πρόβλεψη της ταχύτητας του ανέμου με στατιστικές μεθόδους ανάλυσης δεδομένων και μεθόδους που χρησιμοποιούν τεχνητά νευρωνικά δίκτυα. Τα τεχνητά νευρωνικά δίκτυα στα οποία γίνεται αναφορά είναι: προώθηση με ανάδραση (back propagation neural network - BPN), γενικευμένη παλινδρόμηση (generalized regression neural network - GRNN) και συνάρτηση ακτινικής βάσης (Radial Basis Function Neural Networks – RBFNN).

Για την εκπαίδευση των νευρωνικών δικτύων χρησιμοποιήθηκαν αλγόριθμοι μάθησης με δεδομένα εισόδου την ταχύτητα του ανέμου που μετράται σε χρονικό διάστημα 15 λεπτών. Τέλος αναφέρεται ότι στο άρθρο γίνεται σύγκριση της αποτελεσματικότητας των προβλέψεων της ταχύτητας του ανέμου, μεταξύ των μεθόδων που χρησιμοποιούν τεχνητά νευρωνικά δίκτυα συνάρτησης ακτινικής βάσης (Radial Basis Function Neural Networks – RBFNN) με άλλες μεθόδους πρόβλεψης.

Neural Network Based Voltage and Frequency Controller for Isolated Wind Power Generation

Η εργασία των Singh και Sharma ασχολείται με ένα τεχνητό νευρωνικό δίκτυο για τον έλεγχο της τάσης και της συχνότητας μίας αυτόνομης γεννήτριας αιολικής ενέργειας.

Στο σύστημα λειτουργεί ένας γραμμικός νευρώνας ο οποίος μέσω ενός αλγορίθμου που βασίζεται στη μέθοδο ελαχίστων τετραγώνων, υπολογίζει την τάση και τη συχνότητα της γεννήτριας. Χρησιμοποιείται ένας μετατροπέας πηγής τάσης, ο οποίος μετατρέπει την αιολική ενέργεια σε ισχύς του συστήματος όπου αποθηκεύεται σε ένα σύστημα ενέργειας σε μπαταρία. Ο σχεδιασμός του συστήματος αποθήκευσης ενέργειας έγινε χρησιμοποιώντας το υπολογιστικό πρόγραμμα MATLAB και πιο συγκεκριμένα με τη βοήθεια του toolbox Simulink. Από τα αποτελέσματα της προσομοίωσης προκύπτουν οι δυνατότητες του συστήματος αποθήκευσης ενέργειας.

Hybrid energy system evaluation in water supply system energy production: neural network approach.

Τα συστήματα ύδρευσης έχουν υψηλή κατανάλωση ενέργειας. Γενικότερα προτείνεται η χρήση των υβριδικών συστημάτων για την κάλυψη της απαιτούμενης ενέργειας για τα συστήματα ύδρευσης. Στην εργασία των Gonçalves Fábio, Ramos Helena και Reis Luisa Fernanda παρουσιάζεται ένα μοντέλο πρόβλεψης που βασίζεται σε νευρωνικά δίκτυα για την διαμόρφωση ενός υβριδικού συστήματος με σκοπό την παραγωγή ενέργειας για συστήματα ύδρευσης. Στην μελέτη οι πηγές ενέργειας που προτείνονται για τα υβριδικά συστήματα είναι ο ηλεκτρισμός που προέρχεται από το εθνικό δίκτυο ηλεκτρικής ενέργειας, οι ανεμογεννήτριες και ο υδροηλεκτρισμός. Το τεχνητό νευρωνικό δίκτυο που αναπτύσσεται, αποτελείται από έξι στρώματα που χρησιμοποιούν δεδομένα, που προκύπτουν από προσομοίωση που πραγματοποιείται από το μοντέλο πρόβλεψης. Τα αποτελέσματα δείχνουν ότι το μοντέλο μπορεί να χρησιμοποιηθεί σε σύστημα υποστήριξης αποφάσεων για τον σχεδιασμό υβριδικών συστημάτων τα οποία βελτιώνουν την απόδοσή των συστημάτων ύδρευσης,

A method based on neural networks for generating solar radiation map.

Η πρόβλεψη μετρήσεων της ηλιακής ακτινοβολίας σε μία περιοχή είναι σημαντική για τις εφαρμογές που έχουν σαν πηγή ενέργειας την ηλιακή ενέργεια και ιδιαίτερα για τον σχεδιασμό τους.

Στο άρθρο των Ramedani, Keyhani και Omid παρουσιάζεται μία μέθοδος που χρησιμοποιεί τεχνητά νευρωνικά δίκτυα για την πρόβλεψη καθημερινών μετρήσεων για την ηλιακή ακτινοβολία. Βασιζόμενη σε υπάρχουσες μετρήσεις της ηλιακής ακτινοβολίας σε ορισμένες περιοχές, προβλέπει μετρήσεις για περιοχές που δεν υπάρχουν δεδομένα από το παρελθόν.

Προκειμένου να επιλεγεί το πιο αξιόπιστο τεχνητό νευρωνικό δίκτυο, χρησιμοποιήθηκαν ιστορικά δεδομένα για την ηλιακή ακτινοβολία της πόλης Karaj του Ιράν. Στα δεδομένα εισόδου για τα νευρωνικά δίκτυα χρησιμοποιήθηκαν η μέγιστη θερμοκρασία, η διάρκεια της ηλιοφάνειας και η ηλεκτρομαγνητική ακτινοβολία. Σχεδιάστηκαν αρκετά μοντέλα και επιλέχτηκε αυτό με τις πιο ακριβείς προβλέψεις. Το νευρωνικό δίκτυο που επιλέχτηκε χρησιμοποιήθηκε για την πρόβλεψη τιμών της ηλιακής ακτινοβολίας και σε άλλης πόλεις του Ιράν. Τα δεδομένα για τις προβλέψεις συλλέχθηκαν από 3 σταθμούς στις πόλεις Hashtgerd, Taleghan και Chitgar.

Οι προβλέψεις που προέκυψαν για την κάθε περιοχή στο Ιράν οδήγησαν στον σχεδιασμό ενός χάρτη με προβλέψεις για την ηλιακή ακτινοβολία.

Short-term prediction of photovoltaic energy generation by intelligent approach (2012).

Η αύξηση του παγκόσμιου πληθυσμού και η μείωση των αποθεμάτων ορυκτών καυσίμων δημιουργούν την ανάγκη ανάπτυξης και χρήσης των ανανεώσιμων πηγών ενέργειας. Για τους διευθυντές εγκαταστάσεων (facility managers) των βιομηχανιών, η πρόβλεψη των ενεργειακών απαιτήσεων έχει μετατραπεί σε προτεραιότητα και η πρόβλεψη της παραγωγής ενέργειας παίζει σημαντικό ρόλο στη διαχείριση και τον προγραμματισμό στη βιομηχανία.

Αξιόπιστες προβλέψεις για τη ζήτηση ενέργειας μπορούν να βοηθήσουν στην αποφυγή ελλείμματος της απαραίτητης ενέργειας στις βιομηχανίες και να παράξουν ζωτικής σημασίας πληροφορίες για αποφάσεις που λαμβάνονται.

Η ηλιακή ενέργεια στο σύνολό της είναι πρακτικά ανεξάντλητη και είναι μια ανανεώσιμη πηγή ενέργειας που μπορεί να παρέχει τεράστια ποσά ηλεκτρικής ενέργειας. Στη μελέτη αναφέρεται ότι τα τελευταία χρόνια έχει μειωθεί η κατασκευή μοντέλων πρόβλεψης παραγωγής ενέργειας από φωτοβολταϊκά συστήματα που χρησιμοποιούν ηλιακή ενέργεια.

Στην εργασία που έγινε από τους Chow Stanley, Lee Eric και Li Danny εξετάστηκε η αποτελεσματικότητα και οι παράγοντες που επηρεάζουν ένα μοντέλο βραχυπρόθεσμων προβλέψεων, που στηρίζεται σε τεχνητό νευρωνικό δίκτυο, το οποίο μιμείται τη μη γραμμική συσχέτιση μεταξύ των μετρολογικών παραμέτρων και της ενέργειας που παράγεται από ένα φωτοβολταϊκό σύστημα. Αναφέρεται επίσης ότι η ακρίβεια των προβλέψεων επηρεάζεται από την ηλιακή γωνία.

Short-Term Load Forecasting for Microgrids Based on Artificial Neural Networks.

Στις μέρες μας η ηλεκτρική ενέργεια είναι πολύ βασική για την οικονομία μίας χώρας. Προκειμένου οι επιχειρήσεις ηλεκτρισμού να προσφέρουν στους καταναλωτές υψηλής ποιότητας υπηρεσίες σε ανταγωνιστικές τιμές, στοχεύουν να παράγουν ποσά ηλεκτρικής ενέργειας ανάλογα με αυτά της ζήτησης. Για το σκοπό αυτό γίνονται προβλέψεις για τη ζήτηση από τις επιχειρήσεις ηλεκτρισμού προσπαθώντας αυτές να είναι όσο το δυνατόν πιο ακριβείς.

Η πρόβλεψη ζήτησης της ηλεκτρικής ενέργειας επηρεάζεται από πολλούς παράγοντες και είναι μία πολύπλοκη διαδικασία, η οποία δε μπορεί να υπολογιστεί με ακρίβεια από τις στατιστικές μεθόδους.

Στο άρθρο των Hernandez Luis, Baladrón Carlos, Aguiar Javier M., Carro Belén, Sanchez-Esguevillas, Antonio J. και Lloret Jaime παρουσιάζεται η κατασκευή ενός μοντέλου πρόβλεψης της ζήτησης ηλεκτρικής ενέργειας βασισμένο σε τεχνητά νευρωνικά δίκτυα, που εκτελεί βραχυπρόθεσμες προβλέψεις (Short-Term-Load-Forecasting (STLF)). Αναφέρεται ότι οι προβλέψεις του μοντέλου ήταν εξαιρετικά ακριβείς και η δομή του πολύ απλή. Οι προβλέψεις που πραγματοποιήθηκαν για την μελέτη αφορούσαν μικροδίκτυα.

Modeling and optimization of biogas production on saw dust and other co-substrates using Artificial Neural network and Genetic Algorithm

Η εξάντληση των ορυκτών καυσίμων και η ρύπανση του περιβάλλοντος είναι οι κύριοι λόγοι για την αναζήτηση εναλλακτικών ανανεώσιμων πηγών ενέργειας. Μια από αυτές είναι το βιοαέριο. Το βιοαέριο Αποτελείται τυπικά από 65% μεθάνιο και 35% διοξείδιο του άνθρακα

Το άρθρο των Gueguim Kana, Oloke, Lateef και Adesiyan παρουσιάζει την δημιουργία ενός μοντέλου με σκοπό την αύξηση παραγωγής βιοαερίου από έναν συνδυασμό από πριονίδι, ζωική κοπριά, βλαστούς μπανάνας και χαρτί, το οποίο θα χρησιμοποιεί νευρωνικά δίκτυα σε συνδυασμό με γενετικούς αλγόριθμους.

Η δομή του νευρωνικού δικτύου που χρησιμοποιήθηκε από το μοντέλο προήλθε από τα αποτελέσματα είκοσι πέντε πιλοτικών μελετών για την παραγωγή βιοαερίου. Το μοντέλο χρησιμοποίησε αντικειμενική συνάρτηση ενός γενετικού αλγόριθμου.

Το αποτέλεσμα του μοντέλου ήταν αύξηση παραγωγής βιοαερίου από συνδυασμό από πριονίδι, ζωική κοπριά, βλαστούς μπανάνας και χαρτί.

The Integration of Artificial Neural Networks and Particle Swarm Optimization to Forecast World Green Energy Consumption.

Η η μαθηματική μέθοδος βελτιστοποίησης σμήνους σωματιδίων (particle swarm optimization technique, PSO) είναι μια τεχνική που εντάσσεται στη λεγόμενη νοημοσύνη σμήνους. Πρόκειται για ένα στοχαστικό αλγόριθμο που βασίστηκε στην προσπάθεια προσομοίωσης της κίνησης ενός σμήνους πτηνών ή ενός κοπαδιού ψαριών ως μέρος μιας μελέτης πάνω στην ιδέα της συλλογικής νοημοσύνης των βιολογικών πληθυσμών. Στην έρευνα που έγινε από τους Ghanbarzadeh, Behrang και Assareh εξετάζεται η πρόβλεψη της κατανάλωσης ορυκτών καυσίμων, πρωτογενών μορφών ενέργειας και πράσινης ενέργειας με χρήση της μεθόδου βελτιστοποίησης σμήνους σωματιδίων καθώς και τεχνητών νευρωνικών δικτύων. Η έρευνα χωρίστηκε σε δύο στάδια:

Στο πρώτο στάδιο εφαρμόστηκε η μέθοδος βελτιστοποίησης σμήνους σωματιδίων για να προβλεφθεί η ζήτηση και η κατανάλωση ορυκτών καυσίμων και μορφών πρωτογενούς ενέργειας όπως ο άνθρακας, το αργό πετρέλαιο, ο άνεμος, το ηλιακό φως ή το φυσικό αέριο βασισμένες σε κοινωνικούς και οικονομικούς δείκτες. Οι δείκτες που χρησιμοποιήθηκαν είναι ο πληθυσμός, το ακαθόριστο εγχώριο προϊόν, η ζήτηση και η κατανάλωση του πετρελαίου και του φυσικού αερίου στο εμπόριο. Μελετήθηκαν δυο σενάρια για τον υπολογισμό των κοινωνικών και οικονομικών δεικτών. Στο πρώτο σενάριο δημιουργήθηκαν γραφήματα (polynomial trend line) βασισμένα σε δεδομένα που είχαν παρατηρηθεί για τον κάθε δείκτη και ορίστηκαν οι μέγιστες τιμές που μπορούν να λάβουν σε κάποιο συγκεκριμένο μελλοντικό χρονικό διάστημα. Στο δεύτερο σενάριο χρησιμοποιήθηκε νευρωνικό δίκτυο για να γίνει υπολογισμός του κάθε κοινωνικού και οικονομικού δείκτη για σε ένα συγκεκριμένο μελλοντικό χρονικό διάστημα. Στο δεύτερο στάδιο με τη χρήση μεθόδου βελτιστοποίησης σμήνους σωματιδίων (PSO), προβλέφθηκε η κατανάλωση πράσινης ενέργειας που βασίζεται στην κατανάλωση ορυκτών καυσίμων και πρωτογενών μορφών ενέργειας. Από τις βέλτιστες προβλέψεις του πρώτου σταδίου προβλέφθηκε στο δεύτερο στάδιο η κατανάλωση πράσινης ενέργειας μέχρι το 2040.

Promises of fuel cell in Great Britain

Σε αυτό το άρθρο του Williamson περιγράφονται ευκαιρίες που προσφέρουν οι κυψέλες καυσίμου για ενεργοβόρες βιομηχανίες στη Μεγάλη Βρετανία. Αναφέρεται ότι οι κυψέλες καυσίμου έχουν δυνητικές εφαρμογές στην παραγωγή ηλεκτρικής ενέργειας σε όλο τον ενεργειακό φάσμα, όπως ο άνθρακας, το φυσικό αέριο και ενέργεια από απόβλητα. Τέλος παρουσιάζεται επίσης μια κυψέλη καυσίμου στερεού οξειδίου που αναπτύχθηκε από την Power Ceres, η οποία μπορεί να χρησιμοποιήσει συμβατικά καύσιμα, όπως το φυσικό αέριο ή προπάνιο, για οικιακή χρήση.

Forecasting natural gas consumption in İstanbul using neural networks and multivariate time series methods.

Οι αλλαγές και οι εξελίξεις στην παγκόσμια οικονομία έχουν αυξήσει αισθητά της κατανάλωση ενέργειας, ως εκ τούτου ο σχεδιασμός της κατανάλωσης των ενεργειακών αποθεμάτων έχει γίνει πολύ σημαντικός. Από τα κύρια εργαλεία που χρησιμοποιούνται για τον ενεργειακό προγραμματισμό είναι μέθοδοι πρόβλεψης όπως γενετικοί αλγόριθμοι και νευρωνικά δίκτυα. Στο παρών άρθρο των Demirel Ömer, Zaim Selim, Fahrettin, Özuyar Pinar και Çalışkan Ahmet αναφέρεται ότι οι ακριβείς προβλέψεις της κατανάλωσης φυσικού αερίου είναι καθοριστικές καθώς η πλειοψηφία των συμφωνιών προμήθειας φυσικού αερίου βασίζονται σε αυτές. Γίνεται επίσης αναφορά ότι λανθασμένες προβλέψεις μπορούν να έχουν σαν αποτέλεσμα είτε αχρησιμοποίητες ποσότητες φυσικού αερίου ή έλλειψη διαθεσιμότητας φυσικού αερίου στην αγορά. Τέλος παρουσιάζεται ένα μοντέλο πρόβλεψης κατανάλωσης φυσικού αερίου με χρήση νευρωνικών δικτύων, το οποίο δοκιμάστηκε στη Κωνσταντινούπολη και προτείνεται από τους συγγραφείς δεδομένου ότι η πλειοψηφία των προβλέψεων συμφωνεί με τις πραγματικές τιμές της κατανάλωσης και οι αποκλίσεις είναι σχεδόν μηδαμινές.

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 συνοπτικά 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 συνοπτικό 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 συνοπτικό 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

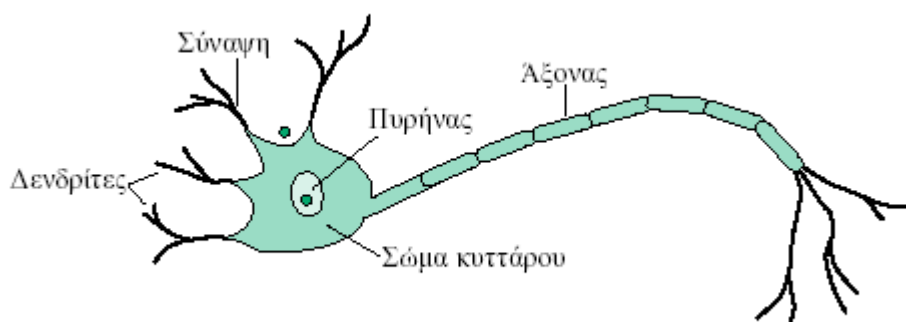
Τα νευρωνικά δίκτυα είναι συστήματα που επεξεργάζονται παράλληλα ένα μεγάλο όγκο δεδομένων και όπως προαναφέρθηκε, αναπαριστούν μια νέα υπολογιστική τεχνική που βασίζεται στον τρόπο διαχείρισης πληροφοριών από τον ανθρώπινο εγκέφαλο. Η εξέλιξη τεχνητών νευρωνικών δικτύων, από την αρχική ιδέα του νευροφυσιολόγου 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 λογογραφήματος), 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 εγκαταλείφθηκε το ενδιαφέρον πάνω στα νευρωνικά δίκτυα. Μια σημαντική ενέργεια την δεκαετία αυτή ήταν οι χάρτες αυτοοργάνωσης με ανταγωνιστική μάθηση. Το 1980 έγιναν πολλές εργασίες πάνω στην θεωρία αλλά και στον σχεδιασμό των νευρωνικών δικτύων. Ο Grossberg (1980) ανέπτυξε μια καινούργια αρχή αυτοοργάνωσης που συνδυάζει φιλτράρισμα από "κάτω προς τα πάνω" και αντίθετη αύξηση σε μικρή μνήμη με από "πάνω προς τα κάτω" ταίριασμα προτύπων και σταθεροποίηση του κώδικα μάθησης. Δεδομένης μιας τέτοιας ικανότητας, και αν το πρότυπο εισόδου ταιριάζει με την ανάδραση μάθησης, λαμβάνει χώρα μία δυναμική κατάσταση που καλείται adaptive resonance. Αυτό το φαινόμενο δίνει την βάση για μια νέα κατηγορία νευρωνικών δικτύων γνωστά σαν adaptive resonance theory (ART). Το 1982 ο Hopfield χρησιμοποίησε την ιδέα μια συνάρτησης ενέργειας για να φτιάξει ένα νέο τρόπο κατανόησης του υπολογισμού που γίνεται από τα δίκτυα με συμμετρικές συνοπτικές συνδέσεις. Επιπλέον καθιέρωσε τον ισομορφισμό ανάμεσα σε τέτοια περιοδικά δίκτυα και σε ένα Ising μοντέλο που χρησιμοποιείται στην στατιστική. Αυτή η αναλογία άνοιξε τον δρόμο για ένα κατακλυσμό θεωριών για τα νευρωνικά δίκτυα. Αυτή η συγκεκριμένη τάξη νευρωνικών δικτύων με ανατροφοδότηση έτυχε ιδιαίτερης προσοχής κατά τη δεκαετία του '80 και με το χρόνο έγιναν γνωστά σαν δίκτυα Hopfield. Το 1983 οι Cohen και Grossberg έδωσαν μια νέα αρχή για σχεδίαση μιας διευθυνσιοδοτούμενης μνήμης (content-addressable memory) που περιλαμβάνει την έκδοση συνεχούς χρόνου του δικτύου Hopfield σαν μια ιδιαίτερη περίπτωση. Μια ακόμα σημαντική ανάπτυξη το 1982 ήταν η έκδοση της εργασίας του Kohonen πάνω στους χάρτες αυτοοργάνωσης, χρησιμοποιώντας μιας ή δύο διαστάσεων δικτυωτές δομές. Το 1983 οι Kirkpatrick, Gallat και Vecchi περιέγραψαν μια νέα διαδικασία που λέγεται εξομοιωμένη ανόπτηση για λύση προβλημάτων συνδυαστικής βελτιστοποίησης. Η εξομοιωμένη ανόπτηση χρησιμοποιείται στην στατιστική θερμοδυναμική και βασίζεται σε μια απλή τεχνική. Την ίδια χρονιά εκδόθηκε μια εργασία από τους Burto, Sutton και Anderson πάνω στην ενισχυμένη μάθηση, η οποία δημιούργησε μεγάλο 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 Rumelhart 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 εγκέφαλος 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 διαιρούνται 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:

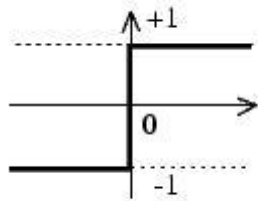
$$y = f(z) = f(w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n + b)$$

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:

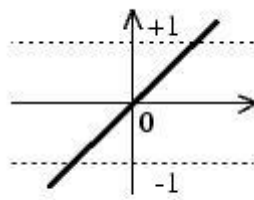
$$\phi(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



ή οποιαδήποτε άλλη βηματική συνάρτηση.

Η γραμμική συνάρτηση μεταφοράς μπορεί να είναι:

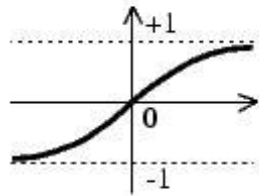
$$\phi(x) = x$$



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 κωδίκων, μορφοποίηση ολοκληρωμένων κυκλωμάτων, έλεγχος διεργασιών, διάγνωση βλαβών ολοκληρωμένων κυκλωμάτων, μηχανική όραση, σύνθεση φωνής.

Οικονομία: Οικονομική ανάλυση, πρόβλεψη τιμών συναλλάγματος.

Κοινωνική ασφάλιση: Αξιολόγηση εφαρμοζόμενης πολιτικής, βελτιστοποίηση παραγωγής.

Βιομηχανία: Βιομηχανικός έλεγχος διεργασιών, ανάλυση και σχεδίαση προϊόντων, συστήματα ποιοτικού ελέγχου, διάγνωση βλαβών διεργασιών και μηχανών, ανάλυση σχεδιασμού χημικών προϊόντων, 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.

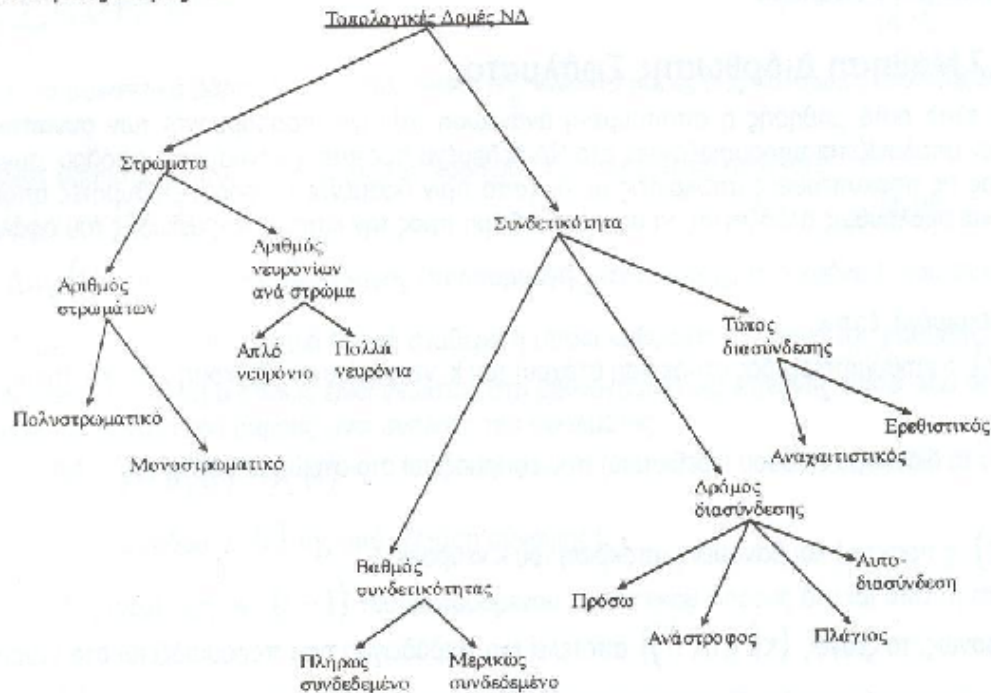
Μεταφορές: Συστήματα διάγνωσης βλαβών φρένων, χρονοπρογραμματισμός οχημάτων, συστήματα δρομολόγησης.

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)

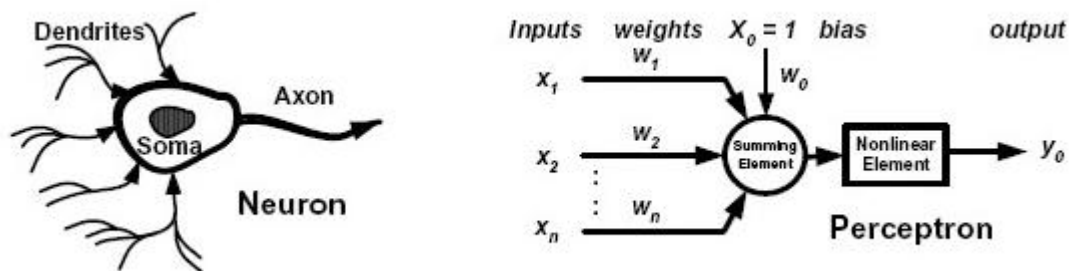
Το μοντέλο του βασικού στοιχείου ενός νευρωνικού δικτύου (βλ. νευρώνα), όπως χρησιμοποιείται ακόμα και σήμερα έχει πρωτοανακαλυφθεί από τους Widrow και Hoff (1960). Αυτοί προσδιόρισαν το perceptron ως το προσαρμοστικό στοιχείο που φέρει ομοιότητες με το νευρώνα. Ένας νευρώνας, που είναι το θεμέλιο ενός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

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

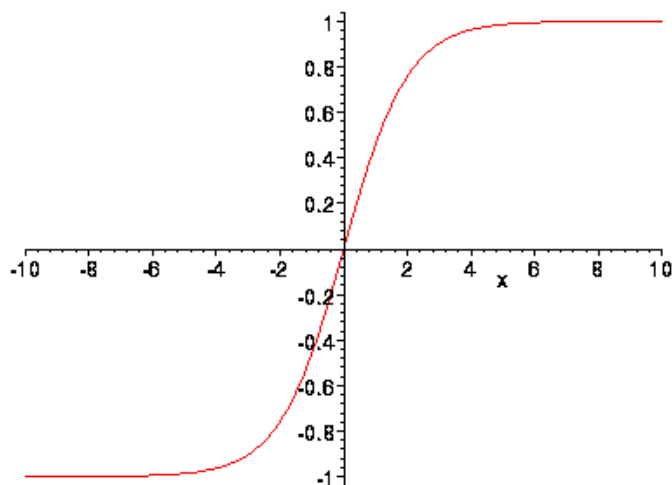
και το κατώφλι ακολουθεί την ανισότητα:

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

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

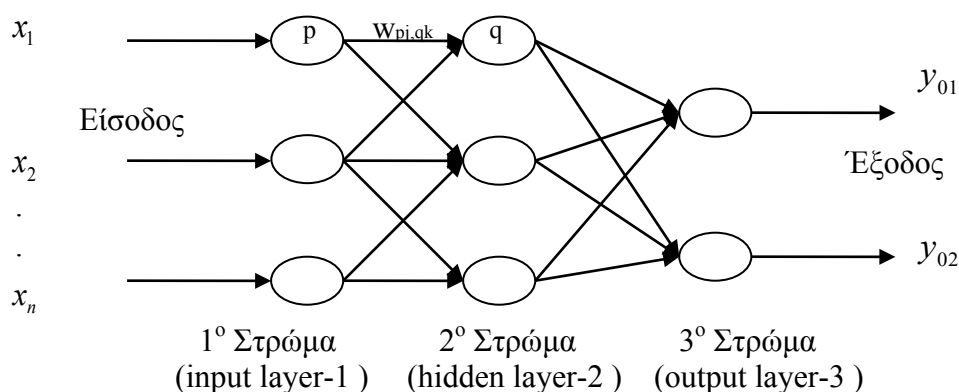
του κανόνα Δέλτα (Delta rule) κατά τον οποίο προστίθεται στην προσωρινή τιμή του βάρους $w(k)$ ο συμψηφιστικός όρος $\eta\epsilon(k)x(k)$, ώστε να δημιουργηθεί η νέα τιμή

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

όπου η είναι ένας αναλογικός όρος, $\epsilon(k)$ 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.

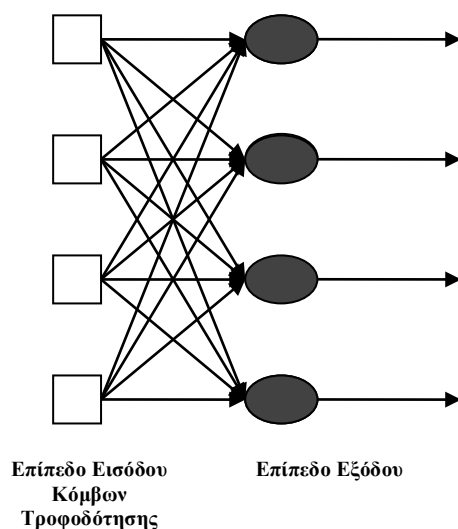


(σχήμα 3.1.5)

Είδη Αρχιτεκτονικών Δομών

1.Ενός Επιπέδου Πρόσθιας Τροφοδότησης Δίκτυα.

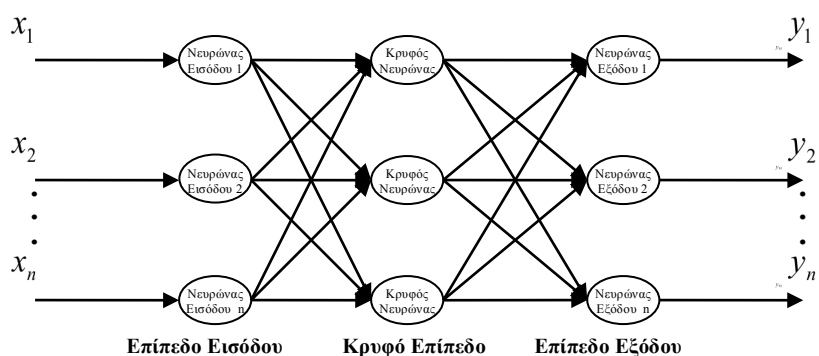
Σε ένα τέτοιο δίκτυο, οι νευρώνες είναι οργανωμένοι σε μορφή επιπέδων. Οι νευρώνες του πηγαίου επιπέδου δείχνουν στους νευρώνες του επόμενου επιπέδου αλλά όχι αντίστροφα, σύμφωνα με το σχήμα 3.1.6 που ακολουθεί.



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

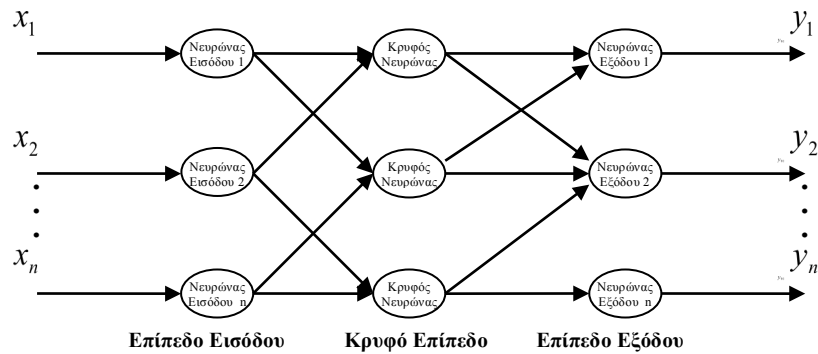
2. Πολλαπλών Επιπέδων Πρόσθιας Τροφοδότησης Δίκτυα.

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

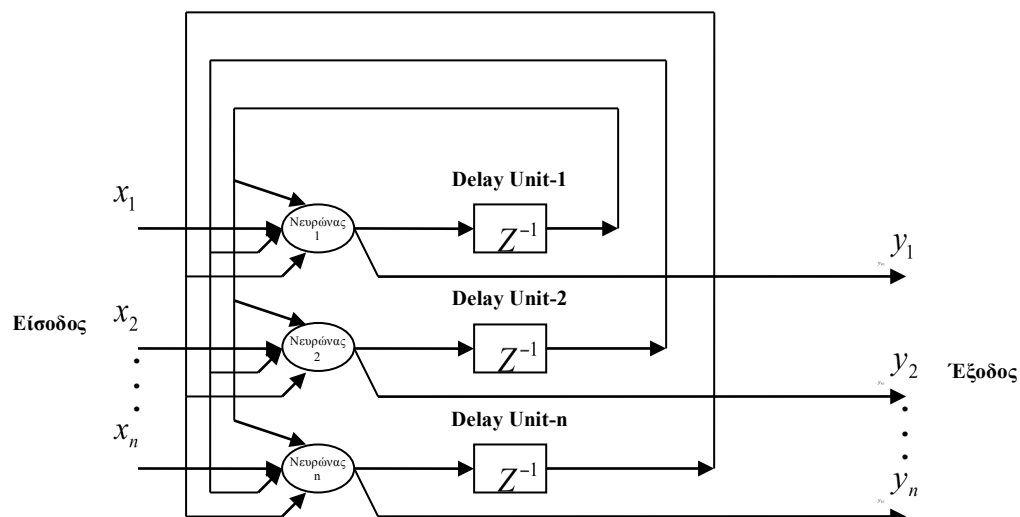
Στο σχήμα που ακολουθεί, παρουσιάζεται ένα μερικώς συνδεδεμένο δίκτυο πρόσθιας τροφοδότησης.



(Σχήμα 3.1.8): Μερικώς συνδεδεμένο δίκτυο πρόσθιας τροφοδότησης

3. Αναδρομικά Δίκτυα

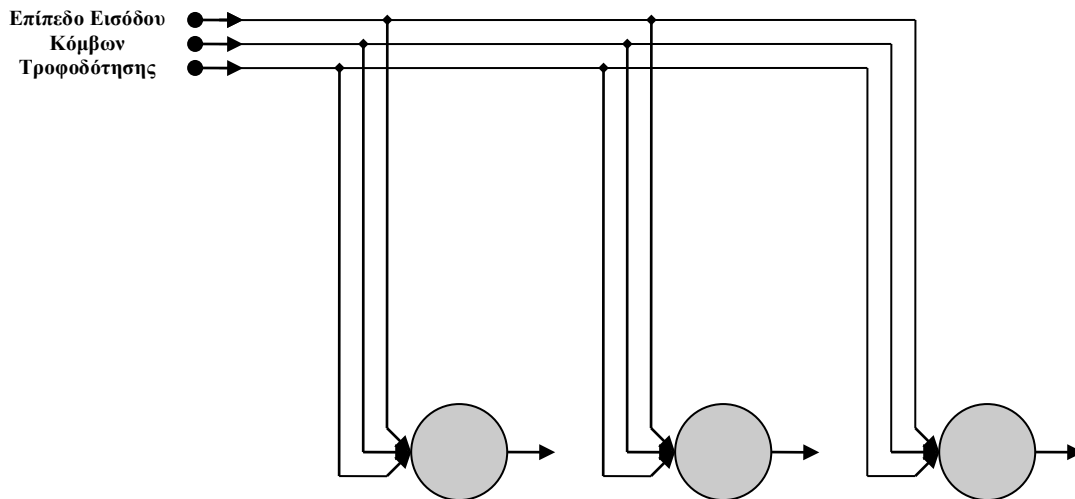
Ένα αναδρομικό νευρωνικό δίκτυο έχει εκτός από τις προς τα εμπρός και προς τα πίσω συνδέσεις. Έχει δηλαδή βρόγχους αναδρομής από τις εξόδους του προς τις εισόδους του. Η παρουσία τέτοιων βρόγχων έχει ισχυρή επίπτωση στην ικανότητα μάθησης του δικτύου. Χαρακτηριστικό παράδειγμα αναδρομικού δικτύου αποτελεί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



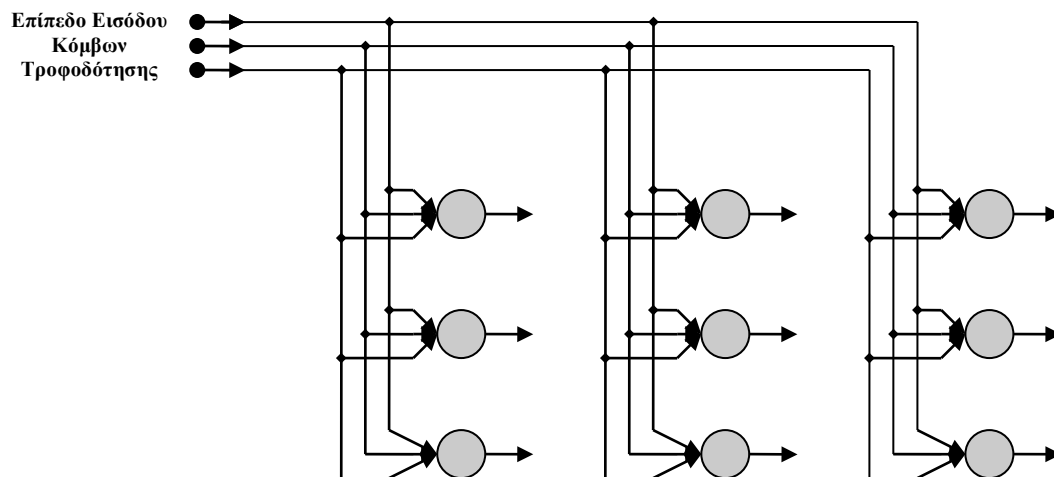
(Σχήμα 3.1.9): Σχηματισμός Δικτύου

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).

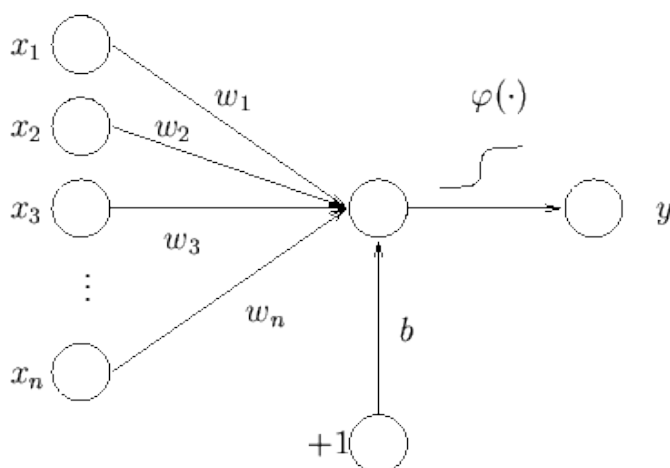
Η αναδρομική διαδικασία εκπαίδευσης (τίθεται $n=n+1$) εκτελείται μέχρι ένα συγκεκριμένο αριθμό επαναλήψεων είτε εναλλακτικά μέχρι να επιτευχθεί ικανοποιητική πρόβλεψη των τιμών εξόδου. Συνήθη κριτήρια τερματισμού του αλγόριθμου εκπαίδευσης περιλαμβάνουν επιτήρηση του αθροίσματος των τετραγώνων των σφαλμάτων,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

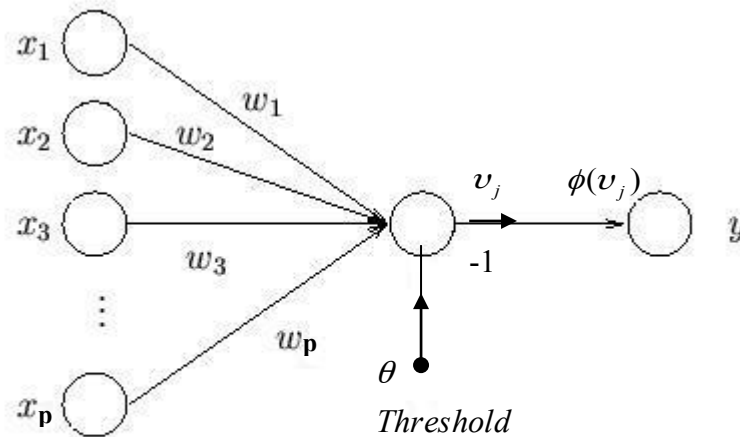
Ένα τέτοιο δίκτυο φαίνεται στο σχήμα που ακολουθεί με x_1 έως x_n να είναι η είσοδος, y η έξοδος και b το κατώφλι (threshold).



(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:

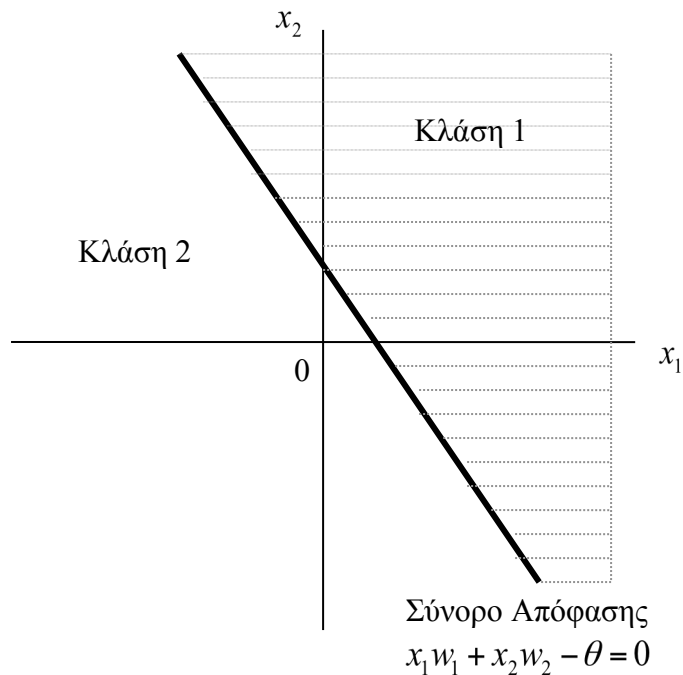
$$v = \sum_{i=1}^p w_i x_i - \theta$$

Σκοπός του Perceptron είναι να ταξινομήσει ένα σύνολο εισόδων (προτύπων) $x_1, x_2, \dots, x_n$ σε μία από τις κλάσεις l_1 και l_2 . Ο κανόνας απόφασης για την ταξινόμηση είναι ο εξής:

ανάθεσε το σημείο που αναπαριστούν $x_1, x_2, \dots, x_n$ στην κλάση l_1 , αν $y = +1$ και στην κλάση l_2 αν $y = -1$. Οι περιοχές απόφασης διαχωρίζονται από το υπερεπίπεδο που ορίζεται από τη σχέση:

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

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

Αλγόριθμος Σύγκλισης του Perceptron

Μεταβλητές και Παράμετροι

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

$\mathbf{y}(n)$ = actual response (πραγματική έξοδος)

$\mathbf{d}(n)$ = 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 εφαρμόζοντας 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:

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

όπου:

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

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 τυχειότητα τείνει να κάνει το ψάξιμο στο χώρο των βαρών σε κάθε κύκλο μάθησης σε στοχαστική διαδικασία, έτσι αποφεύγουμε τον κίνδυνο να γίνουν λιγότεροι κύκλοι από ότι πρέπει. Για ένα δεδομένο σύνολο διάδοσης μπορούμε να ακολουθήσουμε ένα από τους εξής δύο τρόπους:

1. Τρόπος Προτύπων (Pattern Mode). Στο pattern mode η τροποποίηση των βαρών γίνεται με την προβολή κάθε διανύσματος του συνόλου εκπαίδευσης.

2. Σωρηδόν Τρόπος (Batch Mode). Στο batch mode η τροποποίηση γίνεται μετά από την προβολή όλων των διανυσμάτων εκπαίδευσης που αποτελούν ένα epoch, στο δίκτυο. Ο αλγόριθμος πίσω-διάδοσης γενικά δεν συγκλίνει ούτε υπάρχουν σαφώς ορισμένα κριτήρια για να σταματούν την λειτουργία του, αλλά έχει λογικά κριτήρια που μπορούν να χρησιμοποιηθούν για να τερματίσουν της ρυθμίσεις τον βαρών. Για να διαμορφώσουμε ένα τέτοιο κριτήριο το λογικό είναι να σκεφτούμε σε σχέση με τις μοναδικές ιδιότητες του τοπικού ή ολικού ελαχίστου της επιφάνειας λάθους. Έστω ότι το διάνυσμα βαρών $\mathbf{w}^*$ 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:

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

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

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)

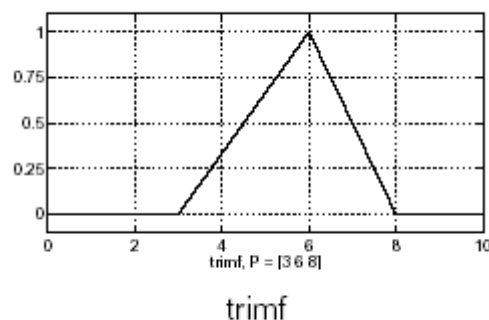
Μια τριγωνικής μορφής συνάρτηση συμμετοχής προσδιορίζεται από τρεις παραμέτρους $\{a,b,c\}$ ως ακολούθως:

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

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:**

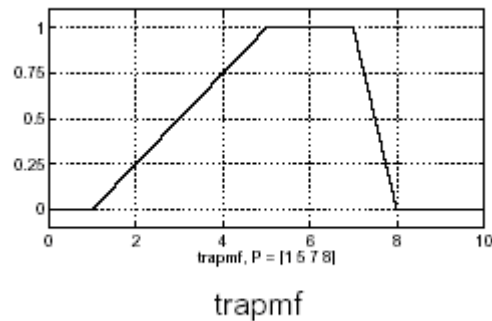
Η τραπεζοειδής συνάρτηση συμμετοχής προσδιορίζεται από τέσσερις παραμέτρους $\{a,b,c,d\}$ ως ακολούθως:

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

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:

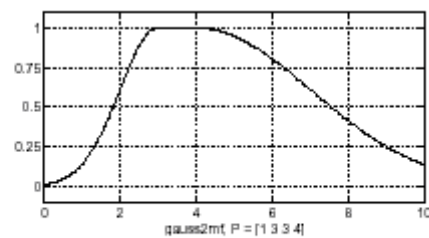
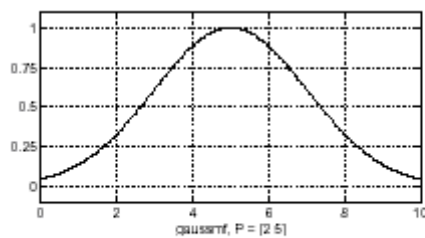
$$A = \begin{cases} 0, x \leq a \\ (x-a)/(b-a), x \in (a,b) \\ 1, x \in (b,c) \\ (d-x)/(d-c), x \in (c,d) \end{cases}$$



3. Gaussian MFs:

Μια Γκαουσιανή συνάρτηση συμμετοχής προσδιορίζεται από δύο παραμέτρους $\{c, \sigma\}$

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



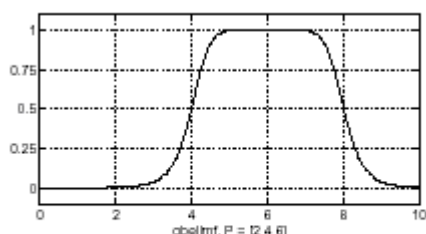
Η Γκαουσιανή συνάρτηση συμμετοχής προσδιορίζεται επακριβώς από δύο παραμέτρους $\{c, \sigma\}$. Το c αναπαριστά το κέντρο της, και το σ το πλάτος της.

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.

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

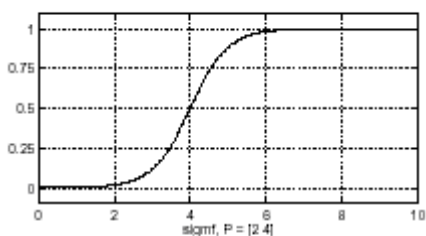


gbellmf

5. Sigmoidal MF

Η σιγμοειδής συνάρτηση συμμετοχής προσδιορίζεται ως ακολούθως:

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



sigmf

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:

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

$$0 < \sum_{i=1}^n A_i(x_i) < m, \forall i \in N_n$$

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.

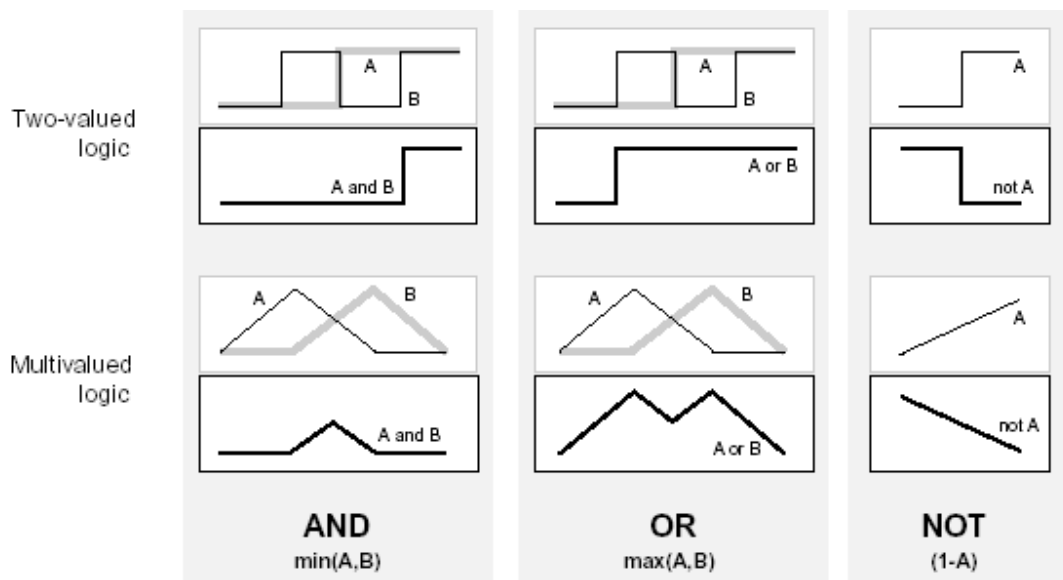
A	B	$\min(A,B)$	A	B	$\max(A,B)$	A	$1-A$
0	0	0	0	0	0	0	1
0	1	0	0	1	1	1	0
1	0	0	1	0	1		
1	1	1	1	1	1		

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: $\forall a,b \in [0,1],$ 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 $R_2: Y \times Z \rightarrow [0,1]$ ορίζεται από την εξίσωση:

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

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:

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

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:

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

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:

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

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 $\Rightarrow$). Αντίθετα όταν εμφανίζεται στο δεύτερο μέρος του κανόνα έχει την έννοια της καταχώρησης (δηλαδή την έννοια που δίνουμε στο σύμβολο $\Rightarrow$). 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.

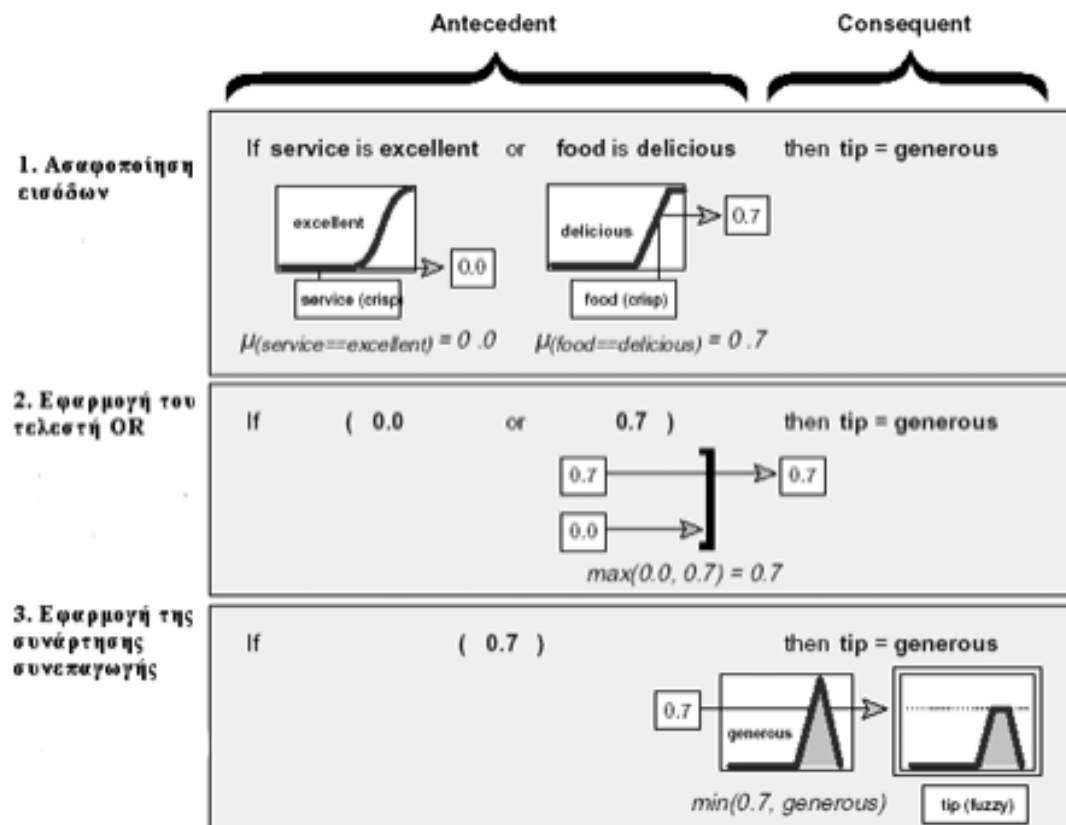
Αλλά και η συνέπεια ενός κανόνα μπορεί να αποτελείται από περισσότερα του ενός μέρη. Για παράδειγμα ο παρακάτω κανόνας:

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.

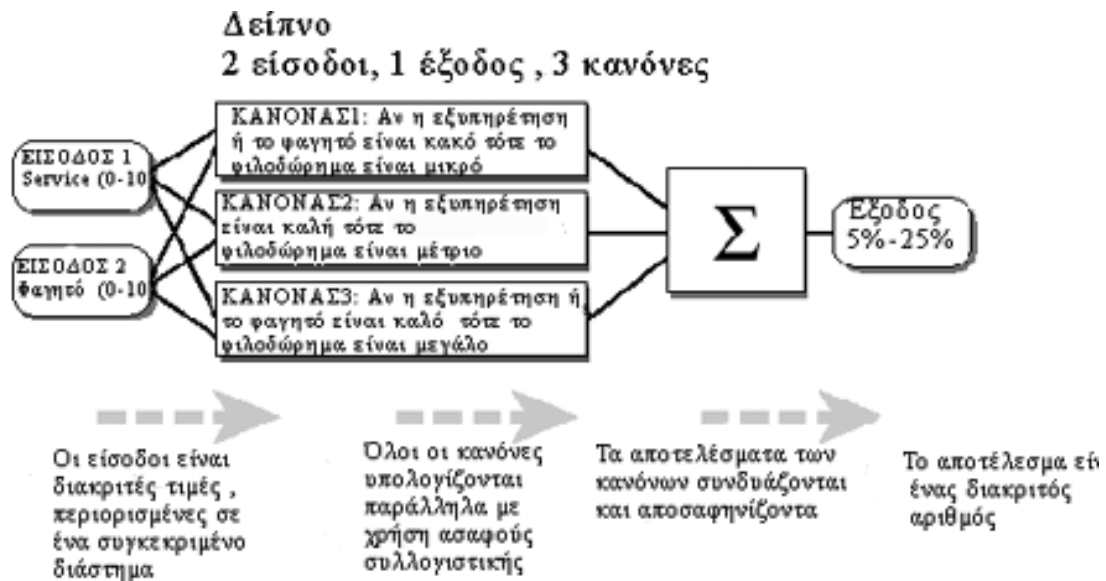
Συστήματα Mamdani

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

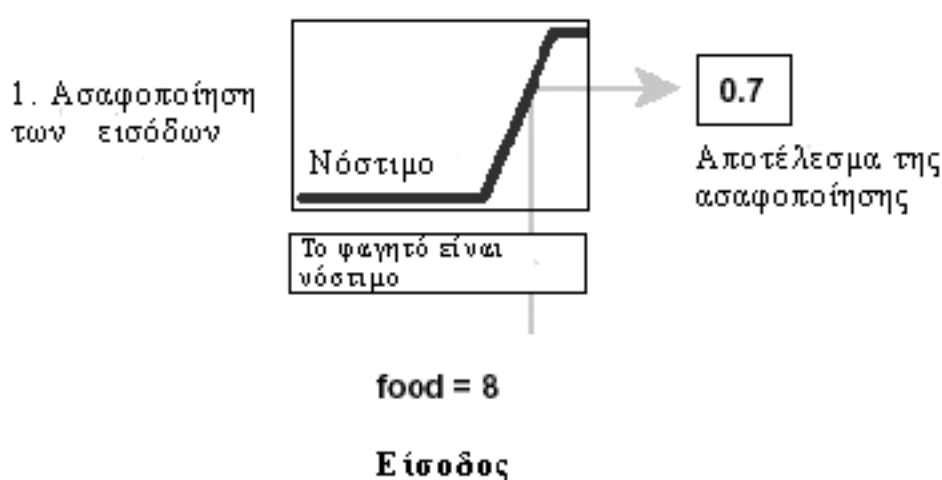
Θα εξετάσουμε κάθε βήμα από τα παραπάνω ξεχωριστά.

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 εισόδοι είναι αριθμητικές τιμές μέσα στα όρια του χώρου αναφοράς (στην περίπτωση του προβλήματος που εξετάζουμε από 0 μέχρι 10) και οι εξόδους είναι βαθμοί συμμετοχής στο προσδιορισμένο ασαφές σύνολο.

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

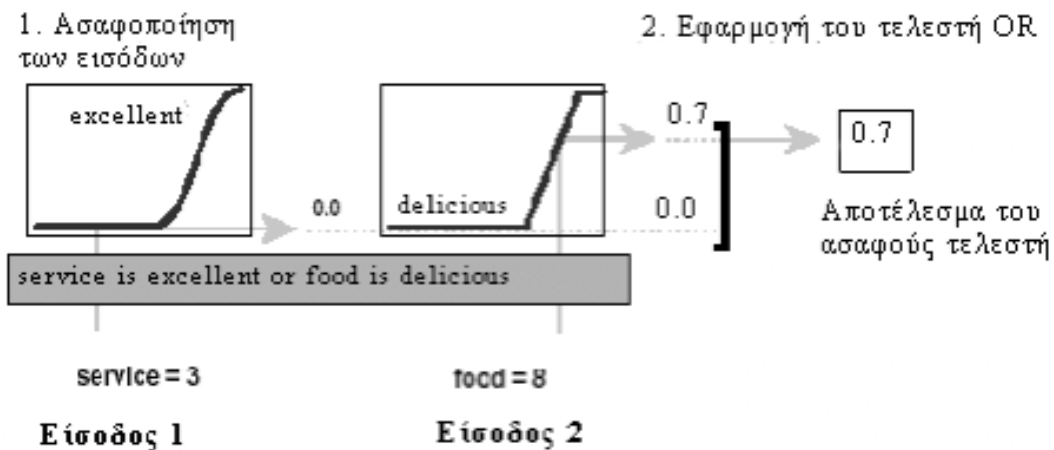
Πριν να εφαρμοσθούν οι κανόνες θα πρέπει να ασαφοποιηθούν οι είσοδοι σύμφωνα με τις λεκτικές μεταβλητές. Για παράδειγμα σε ποιο βαθμό το φαγητό είναι νόστιμο. Αν υποθέσουμε ότι το φαγητό έχει αξιολογηθεί με μια συγκεκριμένη βαθμολογία από το 0 έως το 10 (π.χ. 7) τότε πως μπορούμε να συμπεράνουμε σε ποιο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



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.

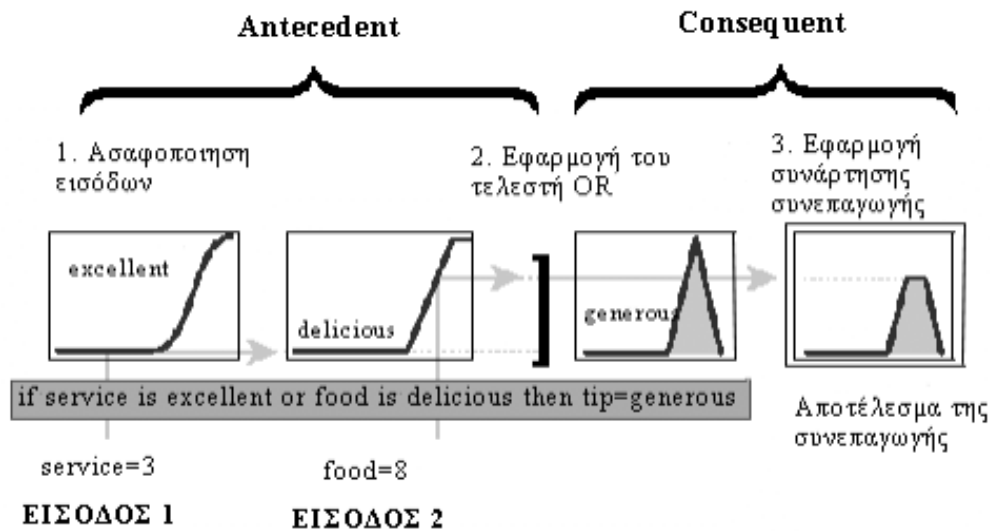
Όπως είδαμε κατά την αναφορά στους ασαφείς τελεστές υπάρχει ένας μεγάλος αριθμός συναρτήσεων που μπορούν να χρησιμοποιηθούν για να ορίσουν τους τελεστές AND και OR. Συνήθως χρησιμοποιούνται ο τελεστής max για το OR και min για το AND. Στο παρακάτω διάγραμμα βλέπουμε πως εφαρμόζονται οι τελεστές.



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.

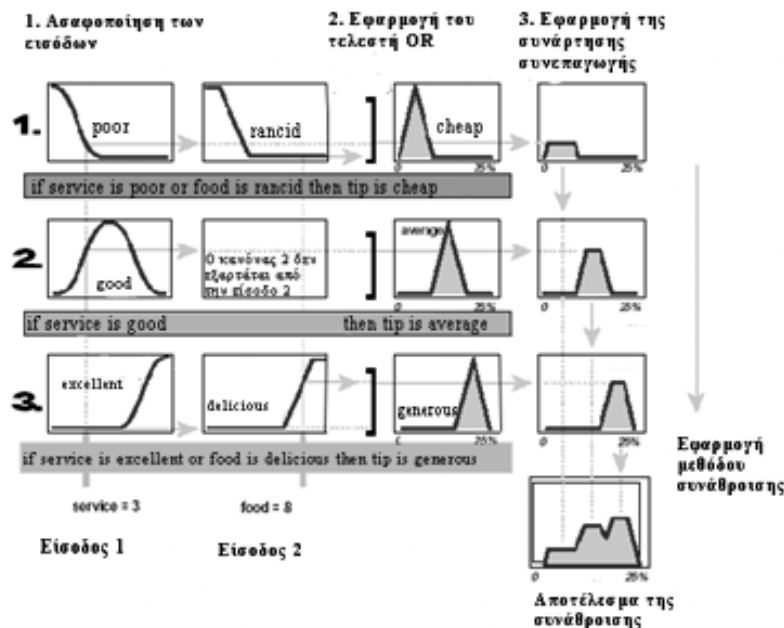
Αφού έχουμε αναθέσει στους κανόνες τα βάρη μπορούμε να εφαρμόσουμε την συνάρτηση συνεπαγωγής. Το αποτέλεσμα ενός κανόνα if-then είναι ένα ασαφές σύνολο που αντιπροσωπεύεται από μια συνάρτηση συμμετοχής που σταθμίζει κατάλληλα τα λεκτικά χαρακτηριστικά που έχουν ανατεθεί σ' αυτόν. Το αποτέλεσμα αυτό αναδιαμορφώνεται 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).

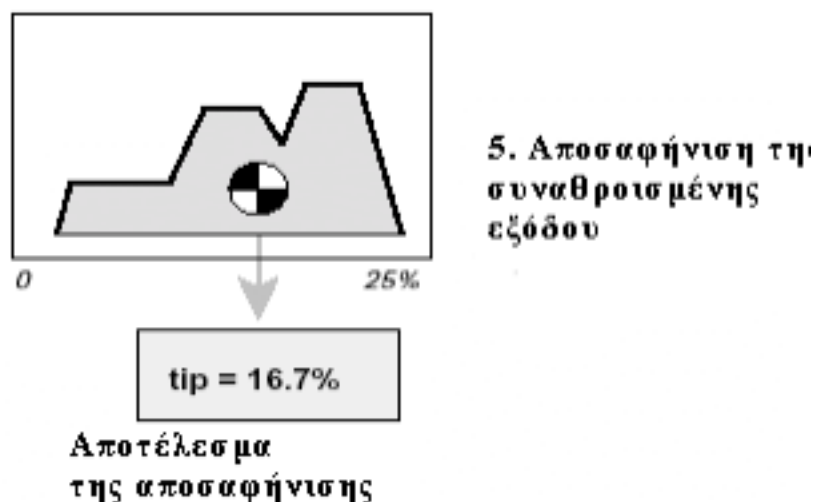
Μέθοδος Cendroid

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:

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

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.

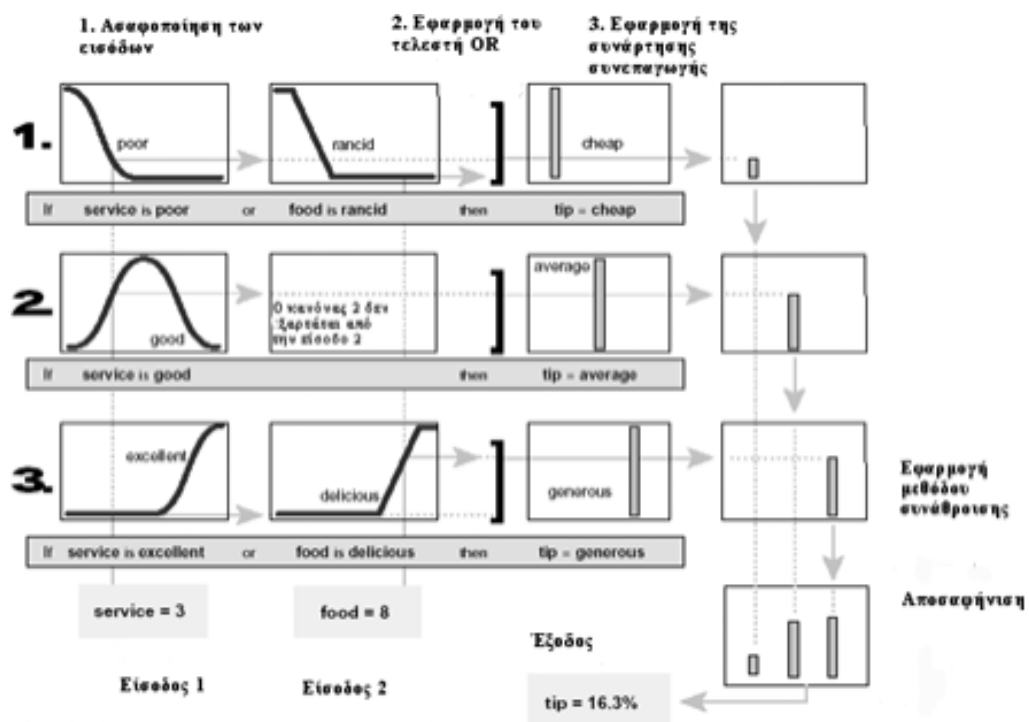
Συστήματα τύπου Sugeno

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.

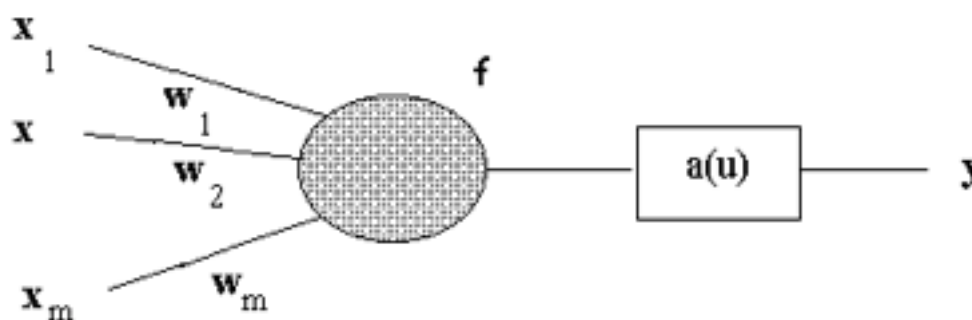
Συστήματα sugeno ανώτερης τάξης είναι εφικτά, όμως δεν προσφέρουν σημαντικές βελτιώσεις και ταυτόχρονα εισαγάγουν σημαντική πολυπλοκότητα.

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
a	η συνάρτηση ενεργοποίηση του νευρώνα
y	η έξοδος του νευρώνα

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

ΝΕΥΡΩΝΕΣ ΣΥΜΜΕΤΟΧΗΣ

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

αν δηλαδή υποθέσουμε ότι η συνάρτηση συμμετοχής είναι η κανονική , τότε για να υλοποιηθεί 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 αλγόριθμους 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:

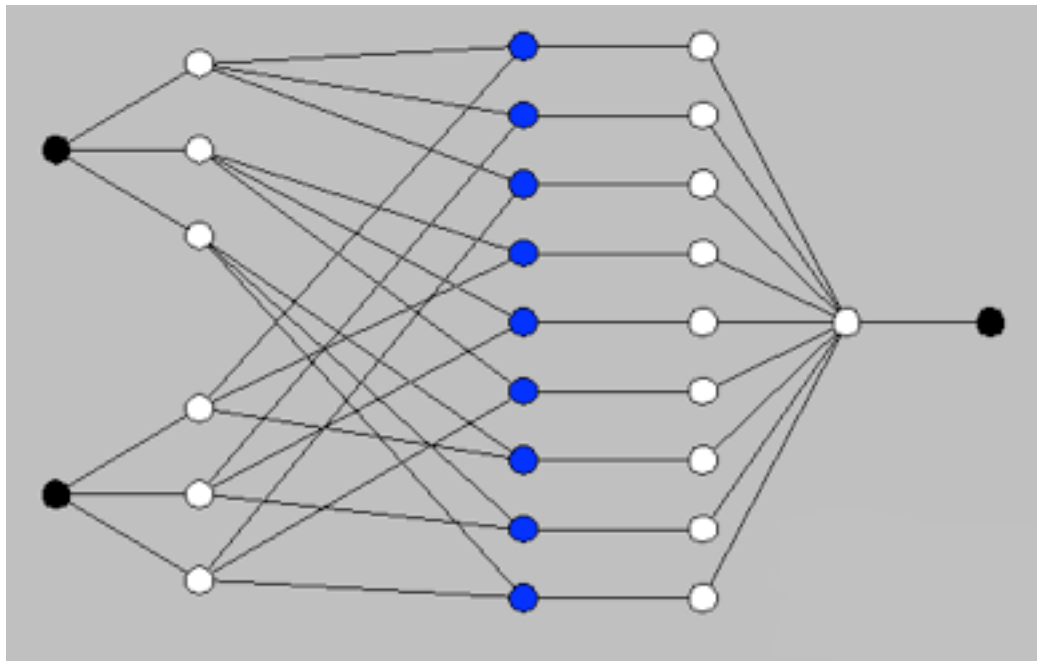
$$f = \text{union } t(x_i, w_i)_{i \in N_m}$$

$$y = a(u)$$

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 παραγωγής ενέργειας. 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:

ΚΑΝΟΝΑΣ 1

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

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

ΚΑΝΟΝΑΣ 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)

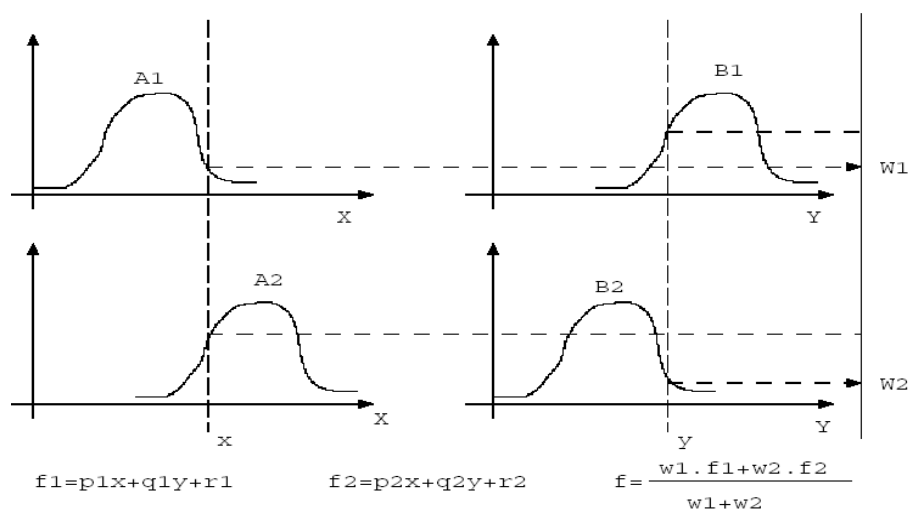
Εισαγωγή

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

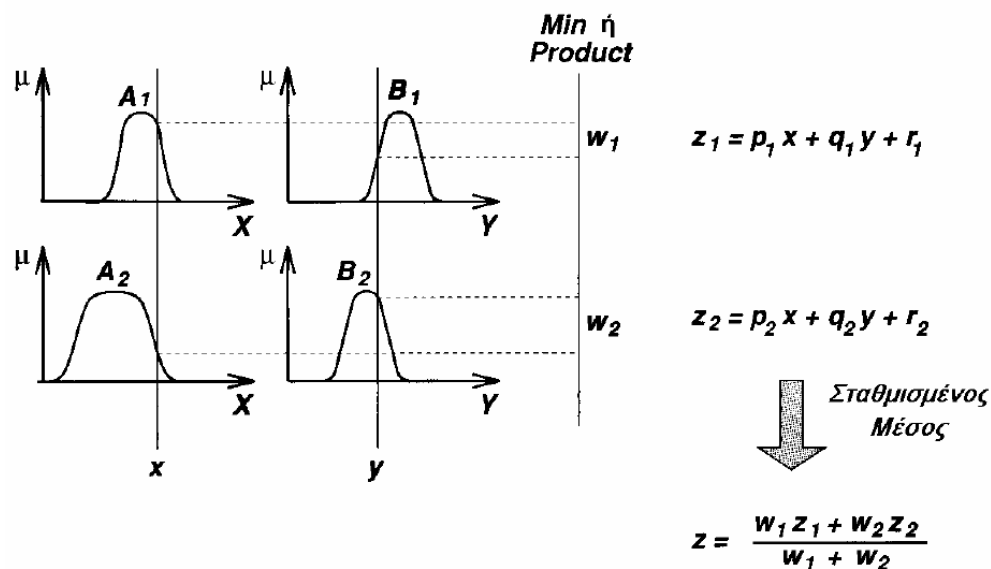
Δομή Αρχιτεκτονικής του ANFIS

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 $f_1 = p_1 \times x + q_1 \times y + r_1$

Rule 2: If x is A_2 and y is B_2 then $f_2 = p_2 \times x + q_2 \times y + r_2$

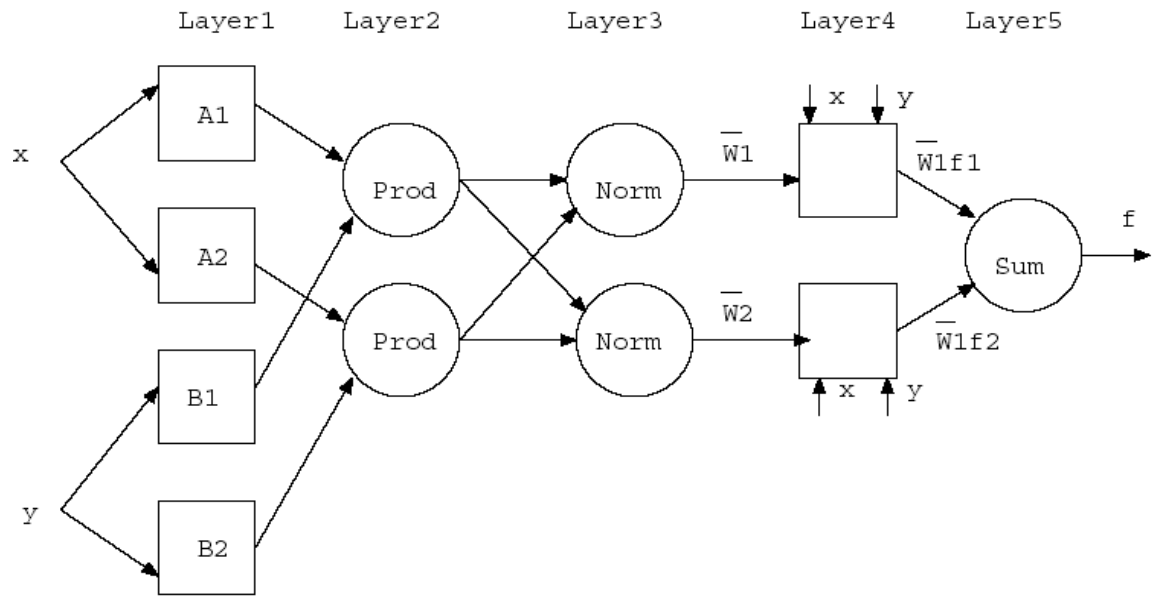
Το παρακάτω σχήμα δείχνει με απλό τρόπο τη διαδικασία συμπερασμού (inference procedure) του μοντέλου sugeno, στην περίπτωση όπου για t-operator έχουμε επιλέξει την τομή των δύο ασαφών συνόλων (A, B), οπότε $\mu_{A \cap B}(x) = \min[\mu_A(x), \mu_B(x)]$



(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:

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

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 μία είσοδο και μια έξοδο. Σαν είσοδο παίρνουμε την τιμή της παραγωγής ενέργειας από μία τεχνολογία μιας τυχαίας ημέρας (t), ενώ σαν έξοδο παίρνουμε την πρόβλεψη του ANFIS για την τιμή της παραγωγής ενέργειας από αυτήν την τεχνολογία της επόμενης από την τυχαία ημέρα ($t+1$).

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 της εισόδου (t) 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.

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:

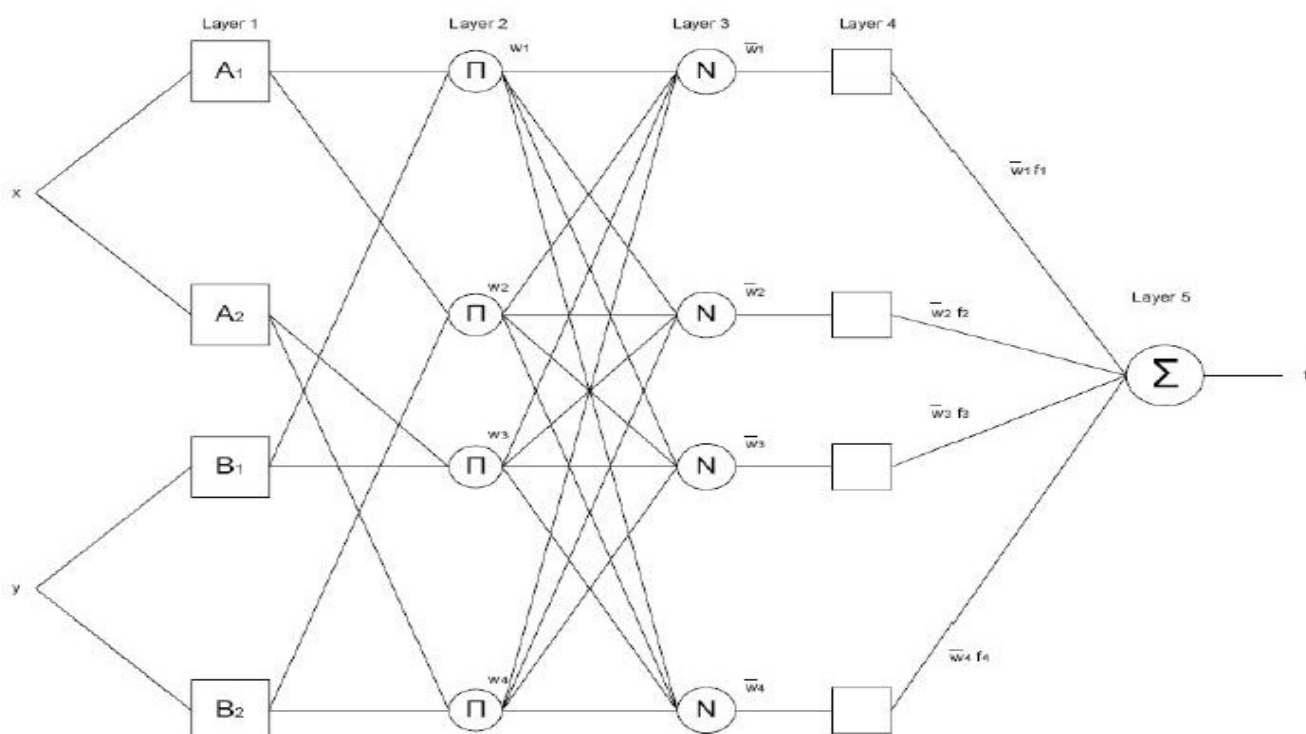
Αν (t) είναι small και (t-1) είναι small τότε (t+1) είναι very small

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_l^i .)



(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.

$$\begin{aligned} O_i^1 &= \mu_{A_i}(x) \quad \text{για } i=1,2 \text{ ή} \\ O_i^1 &= \mu_{B_{i-2}}(y) \quad \text{για } i=3,4 \end{aligned}$$

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 $O_{1,i}$ 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.

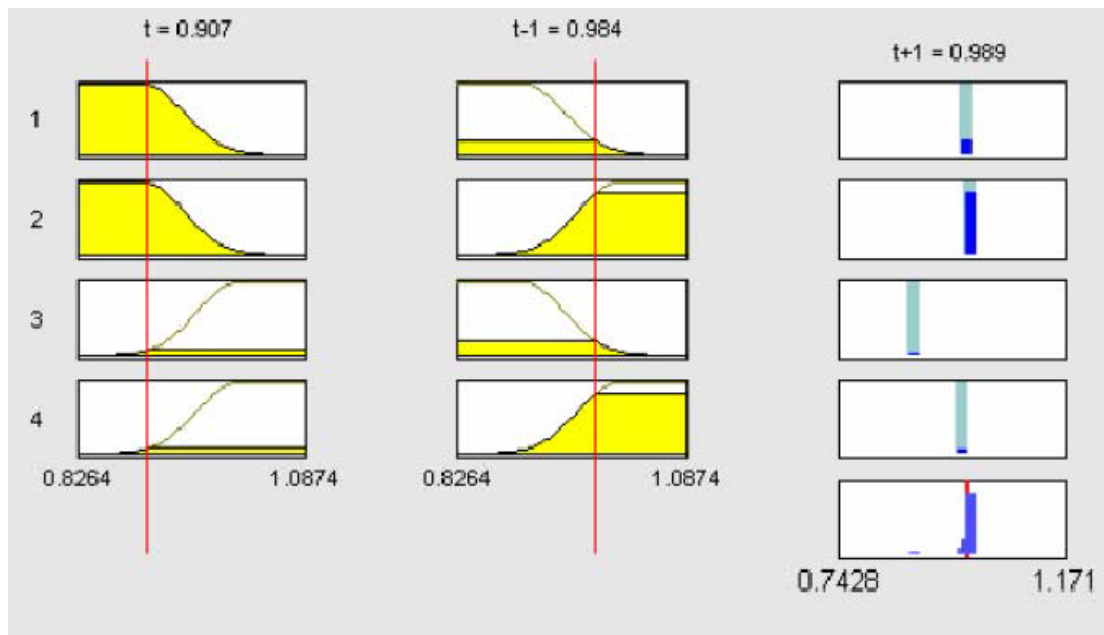
$$O_i^4 = \bar{w}_i * f_i = \bar{w}_i * (p_i * x + q_i * y + r_i)$$

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^1 για κάθε κόμβο. Το σύνολο παραμέτρων φαίνεται στον membership function editor και υπολογίζεται αυτόματα. Έτσι έχουμε:

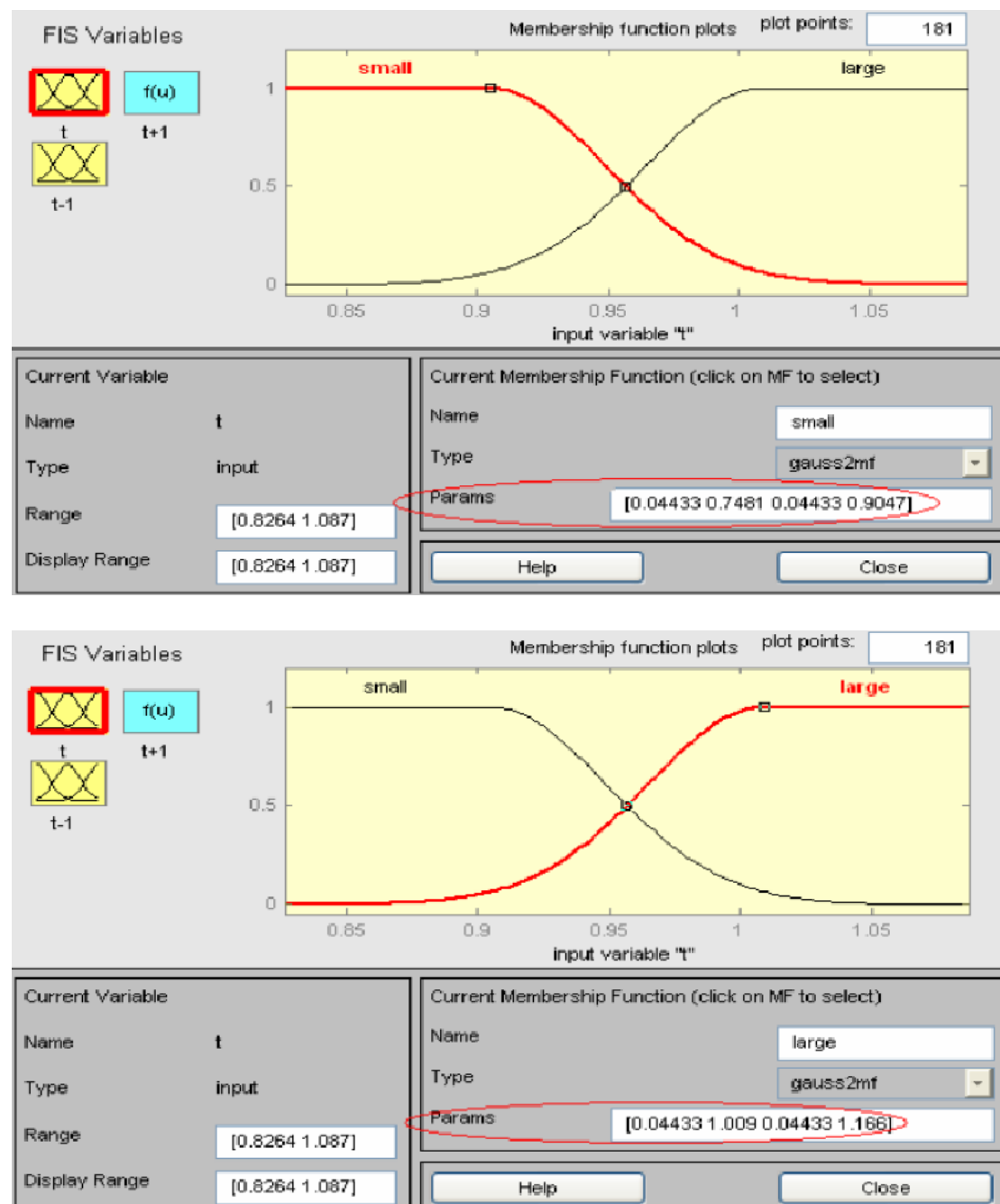
$$[a_1, c_1, a_2, c_2] = [0.04433, 0.7481, 0.04433, 0.9047]$$

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 εισοδοί 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:

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

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

$$O_3^1(y) = \mu_{B_2}(y) = 0.202$$

$$O_4^1(y) = \mu_{B_1}(y) = 0.853$$

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:

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

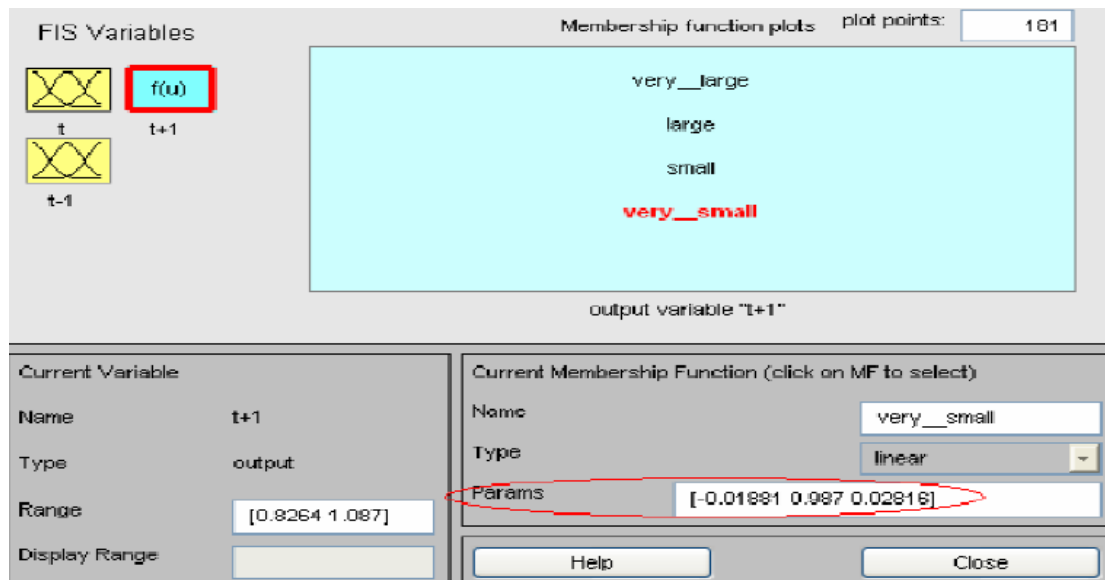
$$O_2^3 = \bar{w}_2 = \frac{w_2}{w_1 + w_2 + w_3 + w_4} = \frac{0.8518}{0.2016 + 0.8518 + 0.0143 + 0.0604} = 0.755$$

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

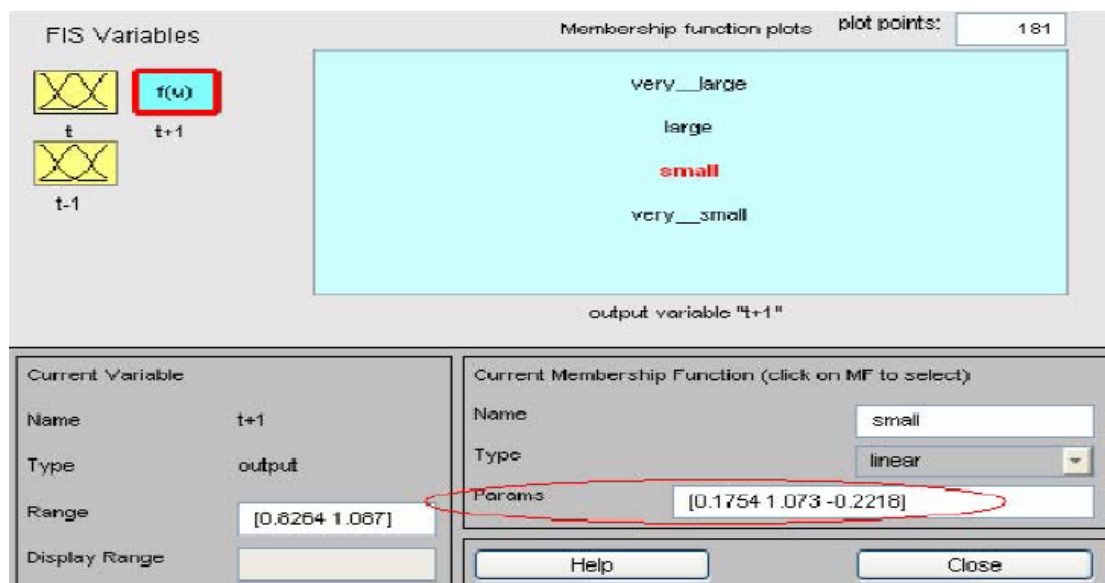
$$O_4^3 = \bar{w}_4 = \frac{w_4}{w_1 + w_2 + w_3 + w_4} = \frac{0.0604}{0.2016 + 0.8518 + 0.0143 + 0.0604} = 0.0536$$

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:

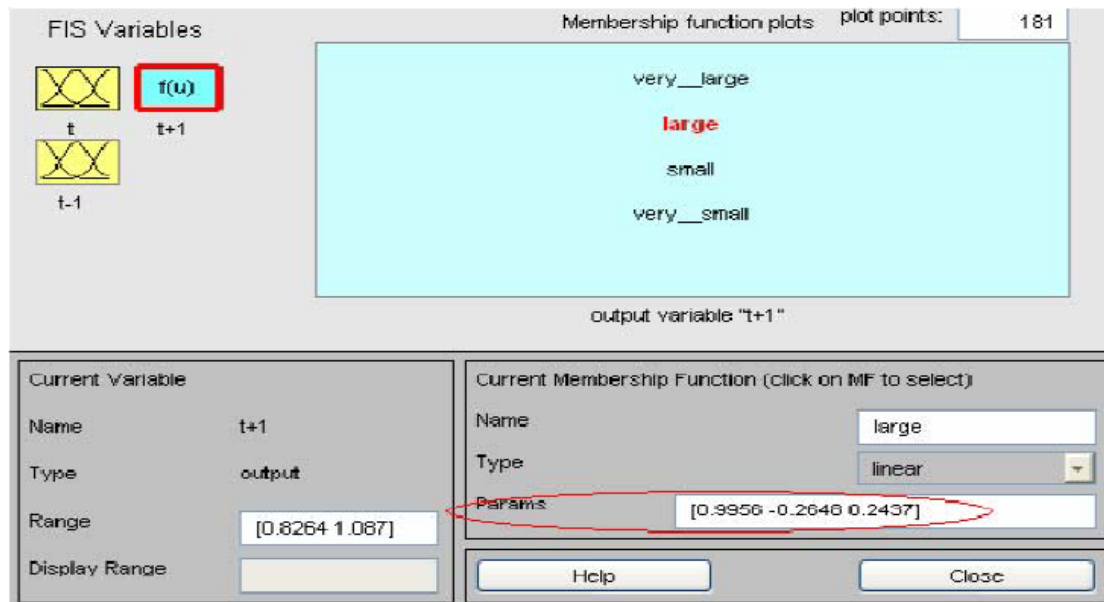
$[p_1, q_1, r_1] = [-0.01881, 0.987, 0.02816]$ αν η γλωσσική μεταβλητή είναι very small
 $[p_2, q_2, r_2] = [0.1754, 1.073, -0.2218]$ αν η γλωσσική μεταβλητή είναι small
 $[p_3, q_3, r_3] = [0.9956, -0.2648, 0.2437]$ αν η γλωσσική μεταβλητή είναι large
 $[p_4, q_4, r_4] = [0.09453, 0.9042, 0.0006383]$ αν η γλωσσική μεταβλητή είναι very large



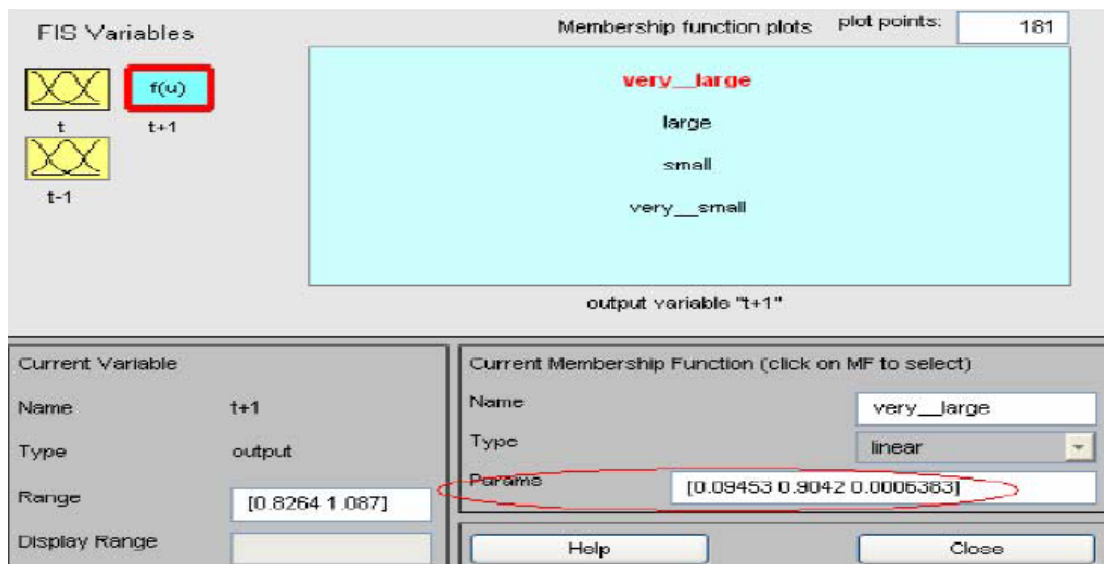
(σχήμα 3.4.12): Σύνολο παραμέτρων για very small



(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 :

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

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

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

$$O_4^4 = \bar{w}_4 * (p_4 * x + q_4 * y + r_4) = 0.05363 * (0.09453 * 0.907 - 0.9042 * 0.984 + 0.0006383) = 0.052$$

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

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

4. ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ (case study)

4.1 ΔΕΔΟΜΕΝΑ

Τα δεδομένα της μελέτης αντλήθηκαν από το site <http://www.iso-ne.com>, το οποίο περιέχει τεράστιο όγκο στατιστικών δεδομένων για τη χώρα της Αγγλίας.

Τα οχτώ είδη τεχνολογιών για τα οποία συλλέξαμε δεδομένα είναι τα εξής:

- Coal (άνθρακας)
- Gas (αέριο)
- Hydro (νερό)
- Nuclear (πυρηνική ενέργεια)
- Oil (πετρέλαιο)
- Refuse (απορρίματα)
- Solar (ηλιακή ενέργεια)
- Wind (αιολική ενέργεια)

Τα δεδομένα των μετρήσεων που καταγράφηκαν, αφορούν καθημερινές μετρήσεις της χρονικής περιόδου 1/1/2008 έως και 31/3/2013, τα ποσά των οποίων αναφέρονται σε μεγαβατώρες (MWh), το ποσό της ηλεκτρικής ενέργειας που παράγεται από κάθε τύπο καυσίμου κατά τη διάρκεια της εβδομάδας.

Χρησιμοποιήθηκαν για κάθε μία από τις περιπτώσεις που μελετάμε, 1917 μετρήσεις. 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 1 με μέρος των μετρήσεων συγκεντρωτικά.

Πίνακας 1: μέρος των δεδομένων

date	COAL	GAS	HYDRO	NUCLEAR	OIL	REFUSE	SOLAR	WIND
1/1/08	56247	78427	30414	110148	26643	22381	0	44
1/2/08	51842	99853	29064	110118	50865	22244	0	50
1/3/08	52524	104750	31225	110066	72884	21775	2	59
1/4/08	52317	119469	22700	110070	44827	20918	1	25
1/5/08	52049	86928	21813	110093	34710	19362	1	23
1/6/08	51690	75635	22970	110122	27222	18904	1	25
1/7/08	52783	87902	27100	110111	30195	18414	0	40
1/8/08	55008	68214	30586	110146	30255	19513	1	56
1/9/08	54637	75057	30764	110107	28054	20556	0	81
1/10/08	49644	83065	34384	110095	31052	22321	1	33
1/11/08	55777	70922	40411	109894	26317	21452	0	42
1/12/08	48843	54897	38156	110127	31336	20248	2	37
1/13/08	47771	47587	33936	110123	33418	19733	2	29
1/14/08	53211	84569	35789	110084	39492	20549	0	69
1/15/08	51738	90991	32959	110099	32820	21006	0	15
1/16/08	56444	92551	34810	110009	33754	21469	2	41
1/17/08	57656	87965	32592	109974	41109	20721	1	19
1/18/08	56583	92138	29212	109478	42961	20670	1	42
1/19/08	52547	85855	29716	108902	20438	20582	1	23
1/20/08	54230	99877	30235	80184	30582	20650	2	61
1/21/08	57158	113417	30103	80179	64862	21706	2	45
1/22/08	55780	115763	28353	80195	49896	21592	1	39
1/23/08	57468	109394	25217	80227	37877	20908	2	31
1/24/08	56693	120331	29341	80205	44640	21326	1	28
1/25/08	57521	122515	27833	80199	42739	22221	2	39
1/26/08	57729	94198	25857	80196	43813	21233	1	13
1/27/08	57622	83794	24714	80195	33925	20893	0	54
1/28/08	55469	99761	26851	80185	44093	21994	1	68
1/29/08	56230	109521	21512	80188	44741	21268	2	23
1/30/08	47278	116996	25766	80032	48905	20960	0	65
1/31/08	50043	118791	27539	81271	47680	21502	0	20

Στον πίνακα 1 παρουσιάζεται ένα δείγμα από τις μετρήσεις (από 1 μέχρι 31 Ιανουαρίου για το έτος 2008) τις οποίες χρησιμοποίησε το ANFIS. Στην κάθε γραμμή καταγράφονται οι μετρήσεις που πραγματοποιήθηκαν σε καθημερινή βάση για όλες τις πηγές ενέργειας που μελετήθηκαν. Στην πρώτη στήλη καταγράφονται οι ημερομηνίες στις οποίες αναφέρονται οι μετρήσεις και στις υπόλοιπες στήλες οι πηγές ενέργειας που χρησιμοποιήθηκαν από το μοντέλο. Στην δεύτερη στήλη καταγράφονται οι μετρήσεις για τον άνθρακα (coal), στην τρίτη στήλη οι μετρήσεις για το αέριο(gas) και αντίστοιχα μέχρι την ένατη στήλη οι μετρήσεις για το νερό (hydro), την πυρηνική ενέργεια (nuclear), το πετρέλαιο(oil) , τα απορρίμματα (refuse), την ηλιακή ενέργεια(solar) και την αιολική ενέργεια (wind).

4.2 Επεξεργασία

Κατά την επεξεργασία των δεδομένων μας ακολουθήσαμε την εξής μεθοδολογία: αρχικά χρησιμοποιήθηκαν τα δεδομένα της παραγωγής ηλεκτρικής ενέργειας που από κάθε τύπο καυσίμου. Οι αρχικές παράμετροι που τέθηκαν στο μοντέλο μας είναι $n=2$ ($mf_n=2$), μέγεθος βήματος 0.1 ($ss=0.1$), εποχές = 500 ($epochs=500$). Στη συνέχεια τρέξαμε τα δεδομένα μας χρησιμοποιώντας όλες τις διαθέσιμες συναρτήσεις ενεργοποίησης $gaussmf$ και $gauss2mf$ (γκαουσιανή 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 Z) και συνάρτηση ενεργοποίησης γενικευμένη καμπανοειδή ($gbellmf$). Όλη η παραπάνω διαδικασία επαναλήφθηκε και για τα υπόλοιπα τα δεδομένα της παραγωγής ηλεκτρικής ενέργειας που από κάθε τύπο καυσίμου και καταγράφηκαν αναλυτικά σε πίνακα που θα παρουσιαστεί στο κεφάλαιο των αποτελεσμάτων.

Ακολουθεί ο πίνακας 2 ο οποίος περιέχει παραμέτρους του ANFIS:

Πίνακας 2 : οι παράμετροι του ANFIS

Name (Όνομα)	ANFIS
Type (Τύπος ασαφούς μηχανισμού συμπερασμού)	Sugeno
Number of nodes (Αριθμός κόμβων)	12
Number of linear parameters (Αριθμός γραμμικών παραμέτρων)	4
Number of nonlinear parameters (Αριθμός μη-γραμμικών παραμέτρων)	6
Total number of parameters (Συνολικός αριθμός παραμέτρων)	10
Number of training data pairs (Αριθμός εκπαιδευόμενων ζευγαριών δεδομένων)	1532
Number of checking data pairs (Αριθμός ελεγχόμενων ζευγαριών δεδομένων)	385
Number of fuzzy rules (Αριθμός ασαφών κανόνων)	2
NumInputs (Αριθμός εισόδων)	1
NumOutputs (Αριθμός εξόδων)	1

4.3 ΑΠΟΤΕΛΕΣΜΑΤΑ

Έπειτα από τις επαναλαμβανόμενες δοκιμές με εφαρμογή διαφορετικών παραμέτρων, το ANFIS δίνει διάφορα αποτελέσματα για την κάθε πηγή ενέργειας τα οποία συγκεντρώθηκαν όλα σε ένα φύλλο excel.

Οι οχτώ πηγές ενέργειας που μελετήσαμε χωρίστηκαν σε δύο ομάδες, ανάλογα με τη χρονολογία που άρχισε να γίνεται η χρήση τους για την παραγωγή ηλεκτρικής ενέργειας, σε παραδοσιακές και νέες τεχνολογίες.

Οι παραδοσιακές τεχνολογίες αποτελούνται από τις εξής:

- Coal (άνθρακας)
- Hydro (νερό)
- Nuclear (πυρηνική ενέργεια)
- Oil (πετρέλαιο)

Οι νέες τεχνολογίες αποτελούνται από τις εξής:

- Gas (αέριο)
- Refuse (απορρίματα)
- Solar (ηλιακή ενέργεια)
- Wind (αιολική ενέργεια)

Στη συνέχεια τα αποτελέσματα που συγκεντρώθηκαν για την κάθε πηγή ενέργειας, προστέθηκαν ανάλογα τις ομάδες που ανήκουν και δημιουργήθηκαν τιμές για τις δύο ομάδες (πίνακας 4) . Ανά συνάρτηση συγκρίθηκαν οι τιμές αυτές και επιλέχθηκε η καλύτερη σε κάθε περίπτωση με γνώμονα την χαμηλότερη τιμή του RMSE.

Ο RMSE (Root Mean Square Error) είναι σημαντικότερος δείκτης που πρέπει να εξετασθεί,. Ο δείκτης αυτός ουσιαστικά καταγράφει πόσο αποτελεσματικό είναι το μοντέλο μας, δηλαδή πόσο καλή πρόβλεψη μπορεί να επιτύχει και δίνεται από τον τύπο:

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

Πίνακας 3: Στατιστικά σφάλματα για διάφορες συναρτήσεις συμμετοχής του ANFIS για όλες τις πηγές ενέργειας που χρησιμοποιήθηκαν.

ANFIS	MSE	RMSE	MAE	MAPE
Gbell				
Coal	20142541.15	4488.04	2954.28	61.06183
Hydro	8558495.29	2925.49	2292.36	11.7866
Nuclear	12830029.85	3581.9	1860.29	2.00431
Oil	61774956.1	7859.7	5854.77	130.3729
Gas	255023797.2	15969.5	12412	9.91231
Refuse	481046.6314	693.5753	543.433	2.96911
Solar	12457.32887	111.6124	73.813	120.43256
Wind	5039984.508	2244.991	1669.691	59.3722
Trimf				
Coal	68727070.74	8290.179174	6000.896544	75.39349194
Hydro	8544978.815	2923.179573	2292.159815	11.83667278
Nuclear	12774984.94	3574.211094	1826.256654	1.98644539
Oil	369865220.3	7231.880311	12089.7268	55.19672113
Gas	248120531.9	15751.84218	12191.17497	9.7986332
Refuse	479796.3847	692.6733608	542.9257576	2.968440944
Solar	12216.04442	110.526216	76.49341511	80.20447903
Wind	5259464.151	2293.352165	1696.846893	59.9499073
gauss2mf				
Coal	20171143.67	4491.229639	2976.087068	68.96738362
Hydro	8562880.89	2926.24006	2295.346957	11.80065933
Nuclear	12816307.41	3579.987068	1844.112313	1.98701512
Oil	58793306.72	7667.679357	5675.572449	119.815899
Gas	248205362.5	15754.53466	12268.07245	9.8255779
Refuse	479395.4137	692.3838629	542.7180939	2.965573907
Solar	6841.25825	82.71189908	57.7811984	99.0655518
Wind	6008401.098	2451.20401	1732.932829	60.82155537
Gaussmf				
Coal	20172229.19	4491.350486	2968.670855	65.87537633
Hydro	8561946.964	2926.080478	2293.160333	11.78884529
Nuclear	12835164.04	3582.619718	1858.582386	2.00279175
Oil	61851401.05	7864.56617	5856.483112	131.2165849
Gas	248880070.8	15775.93328	12281.91061	9.8341965
Refuse	481065.224	693.5886562	543.682005	2.970137832
Solar	10257.08931	101.2772892	72.75281193	133.4115432
Wind	5146113.344	2268.504649	1681.930557	59.48498988

TRAMPF				
Coal	20254384.57	4500.487148	3037.927934	90.91499129
Hydro	8573172.452	2927.998028	2294.356001	11.78941882
Neuclear	12839813.27	3583.268518	1849.402473	1.99435703
Oil	67747579.25	8230.891765	6178.38309	14.92495879
Gas	250525073.1	15827.98386	12349.54078	9.86314443
Refuse	482953.43	694.9485089	544.4600936	2.974333636
Solar	12871.61565	113.453143	80.80402106	81.85984956
Wind	5342872.039	2311.465344	1715.122224	60.61345293
PIMF				
Coal	563361514.9	23735.23783	20858.75932	191.0421254
Hydro	8571270.604	2927.673241	2297.176432	11.81428264
Neuclear	12846975.54	3584.267783	1825.904166	1.97035583
Oil	57313372.96	7570.55962	5584.935319	11.45198948
Gas	248902216.7	15776.63515	12281.95687	9.8303828
Refuse	482304.4593	694.4814319	544.1879263	2.974864903
Solar	12861.30297	113.4076848	80.71703342	82.10917194
Wind	5183660.988	2276.765466	1690.093916	60.2509356

Ο πίνακας 3 περιέχει στατιστικά σφάλματα για διάφορες συναρτήσεις ενεργοποίησης του ANFIS. Στον πίνακα περιέχονται 6 υποπίνακες, όπου ο κάθε ένας αντιστοιχεί σε κάθε συνάρτηση ενεργοποίησης που χρησιμοποιήθηκε. Οι συναρτήσεις ενεργοποίησης έχουν μπει σε κίτρινο πλαίσιο για να γίνονται εύκολα αναγνωρίσιμες. Οι συναρτήσεις που χρησιμοποιήθηκαν είναι οι gbell, trimf, gaussmf, gauss2mf, pimgf και trampf. Ο κάθε υποπίνακας περιέχει οχτώ γραμμές με τις πηγές ενέργειας που μελετήθηκαν σε κάθε περίπτωση. Τέλος ο πίνακας περιέχει τέσσερις στήλες οι οποίες αντιστοιχούν στα στατιστικά σφάλματα που εξήγαγε το AFNIS. Τα στατιστικά σφάλματα είναι τα MSE, RMSE, MAE και MAPE και για να γίνονται εύκολα αντιληπτά έχουν μπει σε κόκκινο πλαίσιο.

Πίνακας 4: Στατιστικά σφάλματα για διάφορες συναρτήσεις συμμετοχής του ANFIS αθροιστικά για δυο ομάδες πηγών ενέργειας (παραδοσιακές-νέας τεχνολογίας πηγές ενέργειας).

	MSE	RMSE	MAE	MAPE
Gbell				
traditional	103306022.4	18855.13	12961.7	205.22564
new	260557285.7	19019.6787	14698.937	192.68618
Trimf				
traditional	459912254.8	22019.45015	22209.03981	144.4133312
new	253872008.5	18848.39392	14507.44104	152.9214605
gauss2mf				
traditional	100343638.7	18665.13612	12791.11879	202.5709571
new	254700000.2	18980.83444	14601.50457	172.678259
Gaussmf				
traditional	103420741.2	18864.61685	12976.89669	210.8835982
new	254517506.4	18839.30387	14580.27598	205.7008674
TRAMPF				
traditional	109414949.5	19242.64546	13360.0695	119.6237259
new	256363770.2	18947.85086	14689.92712	155.3107804
PIMF				
traditional	642093134	37817.73848	30566.77523	216.2787533
new	254581043.4	18861.28973	14596.95574	155.1653552

Ο πίνακας 4 περιέχει τα στατιστικά σφάλματα για τις διάφορες συναρτήσεις συμμετοχής του ANFIS που παρουσιάστηκαν στον πίνακα 3. Οι συναρτήσεις συμμετοχής έχουν τονιστεί όπως και στον πίνακα 3 με κίτρινο πλαίσιο. Η διαφορά μεταξύ των δύο πινάκων είναι ότι ο κάθε υποπίνακας του πίνακα 4 έχει δύο γραμμές αντί για οχτώ που είχε ο πίνακας 3, οι οποίες αντιστοιχούν στα αθροίσματα των παραδοσιακών και των νέων τεχνολογιών, που έχουν χωριστεί οι πηγές ενέργειας. Στον πίνακα όπως και στον πίνακα 3 υπάρχουν τέσσερις στήλες όπου αντιστοιχούν στα αθροίσματα των στατιστικών σφαλμάτων των δύο ομάδων και έχουν τονιστεί με κόκκινο πλαίσιο.

Για μεγαλύτερη ευκολία έχουν τονιστεί με bold οι χαμηλότερες τιμές του RMSE ανά κατηγορία συνάρτησης συμμετοχής.

Πίνακας 5: Στατιστικά σφάλματα με βάση την συνάρτηση συμμετοχής του ANFIS που δίνει το μικρότερο RMSE αθροιστικά, για τις δυο ομάδες πηγών ενέργειας (παραδοσιακές-νέας τεχνολογίας πηγές ενέργειας).

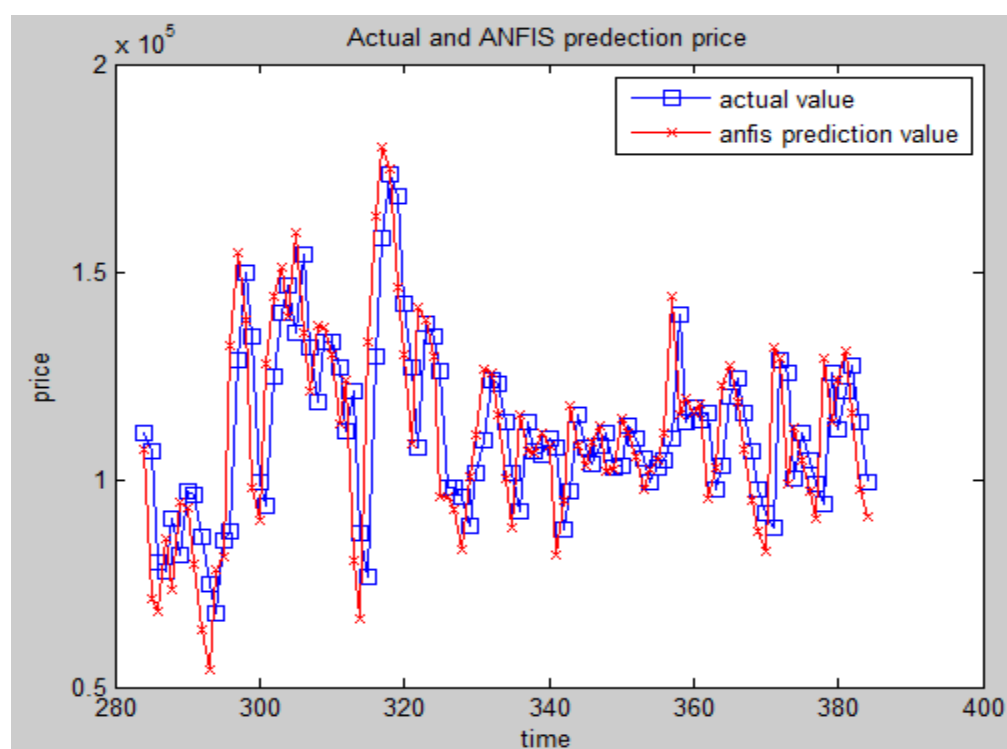
gauss2mf				
traditional	100343638.7	18665.13612	12791.11879	202.5709571
new	254700000.2	18980.83444	14601.50457	172.678259

Ο πίνακας 5 αντιστοιχεί στη συνάρτηση ενεργοποίησης gauss2mf στην οποία παρατηρήθηκε η μικρότερη τιμή του RMSE για τις δυο ομάδες που διαχωρίστηκαν οι πηγές ενέργειας (παραδοσιακές-νέας τεχνολογίας πηγές ενέργειας) από όλες τις συναρτήσεις ενεργοποίησης που χρησιμοποιήθηκαν.

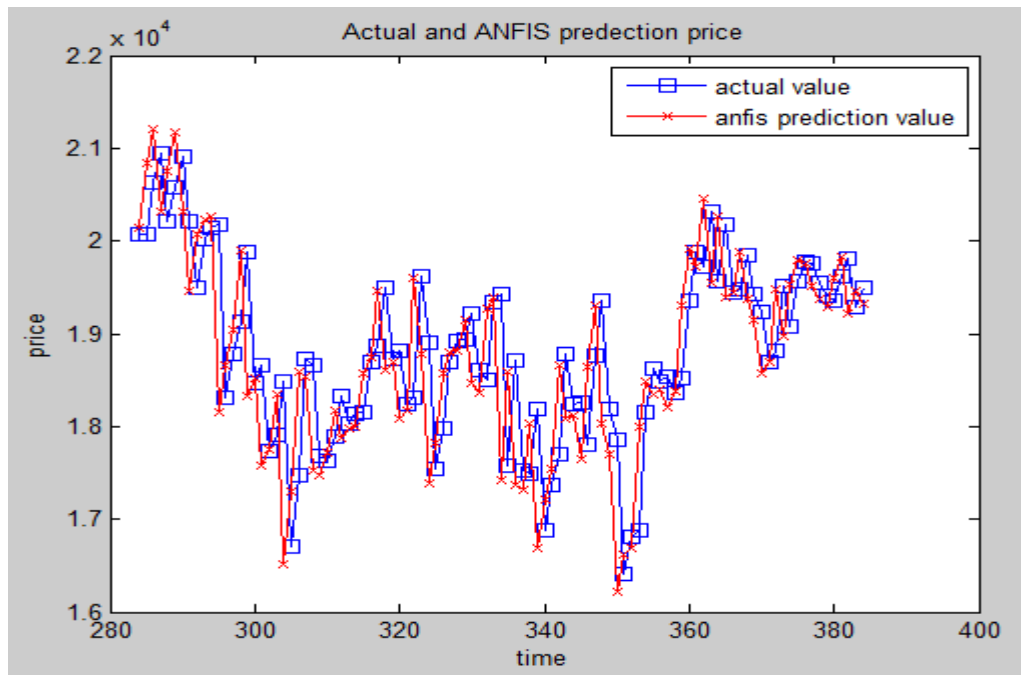
Η τιμή αυτή έχει τονιστεί με bold και βρίσκεται σε πράσινο πλαίσιο και αντιστοιχεί στις παραδοσιακής τεχνολογίας πηγές ενέργειας.

Στα σχήματα που ακολουθούν, παρουσιάζονται γραφήματα αντιστοιχίας των πραγματικών τιμών σε σχέση με αυτών της πρόβλεψης. Η επιλογή των γραφημάτων έχει γίνει με κριτήριο τη χαμηλότερη τιμή RMSE του πίνακα 4, που αντιστοιχεί στις νέες τεχνολογίες για τη συνάρτηση gauss2mf.

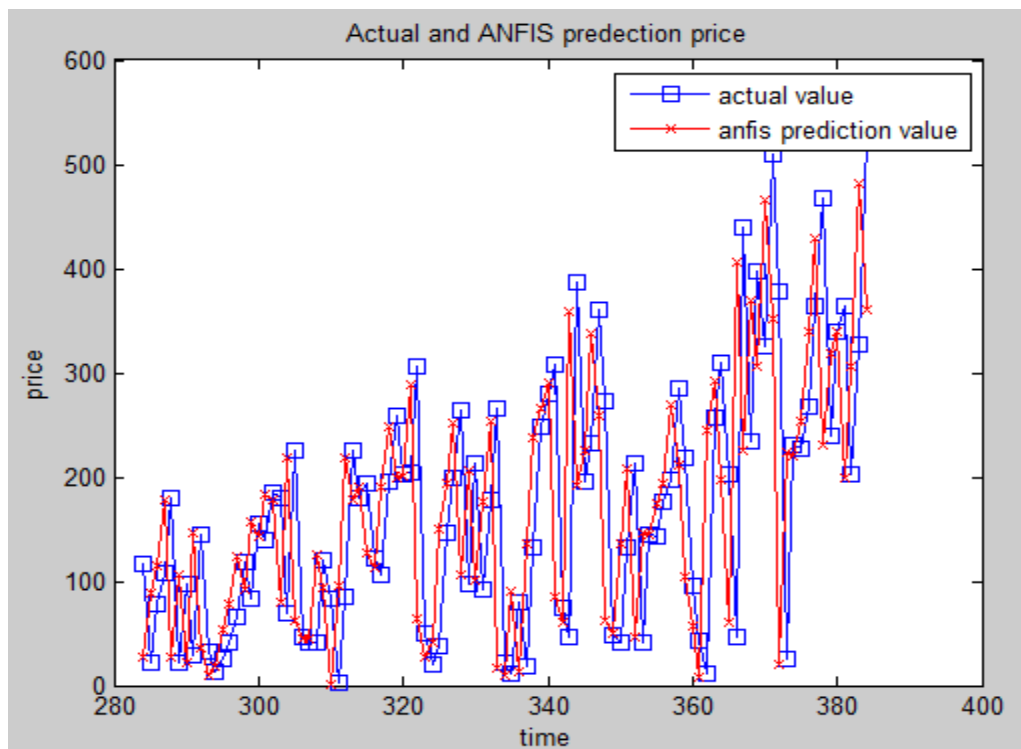
Με την παρακάτω απεικόνιση θα αποκτήσουμε πιο ολοκληρωμένη αντίληψη για την αποδοτικότητα του μοντέλου μας.



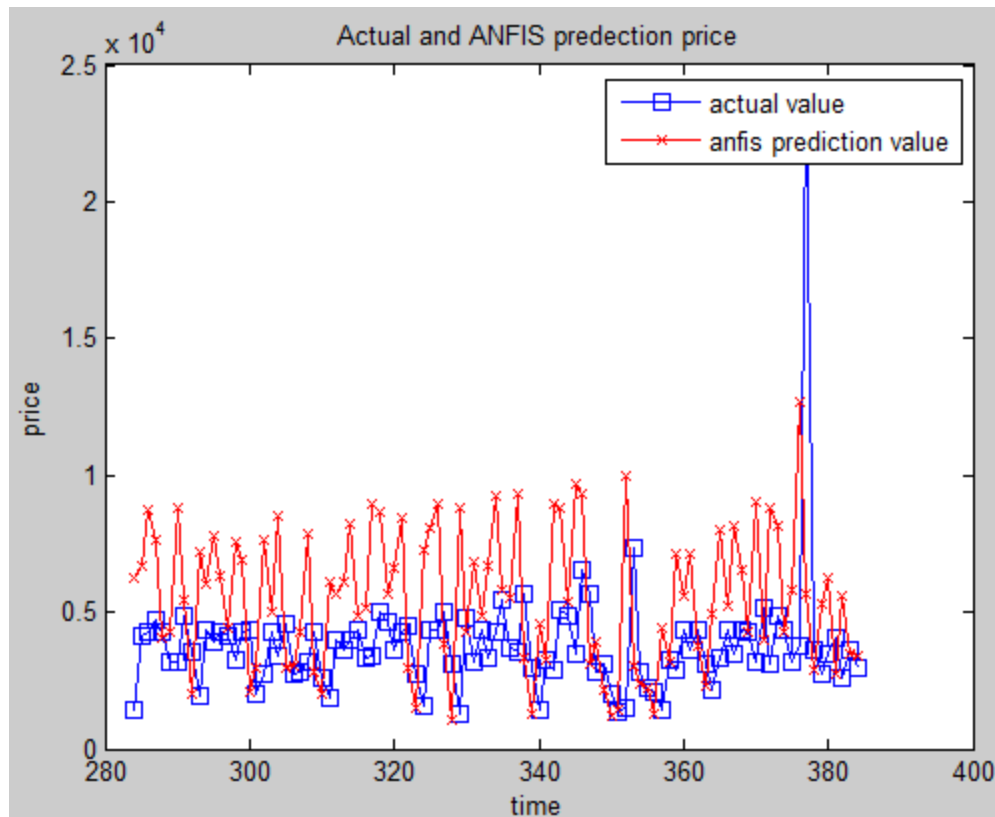
Γράφημα 1: σύγκρισης πραγματικών τιμών και τιμών πρόβλεψης που αφορά την παραγωγή ηλεκτρικής ενέργειας με πηγή ενέργειας το φυσικό αέριο με χρήση της συνάρτησης gauss2mf.



Γράφημα 2: σύγκρισης πραγματικών τιμών και τιμών πρόβλεψης που αφορά την παραγωγή ηλεκτρικής ενέργειας με πηγή ενέργειας απορρίμματα με χρήσης της συνάρτησης gauss2mf.



Γράφημα 3 σύγκρισης πραγματικών τιμών και τιμών πρόβλεψης που αφορά την παραγωγή ηλεκτρικής ενέργειας με πηγή ενέργειας την ηλιακή ενέργεια με χρήσης της συνάρτησης gauss2mf.



Γράφημα 4: σύγκρισης πραγματικών τιμών και τιμών πρόβλεψης που αφορά την παραγωγή ηλεκτρικής ενέργειας με πηγή ενέργειας την αιολική ενέργεια με χρήσης της συνάρτησης gauss2mf.

Στα γραφήματα 1, 2, 3 και 4 γίνεται σύγκριση των πραγματικών τιμών της ποσότητας παραγωγής ενέργειας με τις τιμές πρόβλεψης που πραγματοποίησε το μοντέλο για τις νέες τεχνολογίες στην συνάρτηση gauss2mf στην οποία παρατηρήθηκε η μικρότερη τιμή του σφάλματος στον πίνακα 4 όπου πραγματοποιήθηκε σύγκριση μεταξύ των παραδοσιακών και νέων τεχνολογιών για όλες τις συναρτήσεις ενεργοποίησης που χρησιμοποιήθηκαν. Οι πηγές ενέργειας των νέων τεχνολογιών στις οποίες αντιστοιχούν τα παραπάνω γραφήματα είναι:

Γράφημα 1: παραγωγή ηλεκτρικής ενέργειας με πηγή ενέργειας το φυσικό αέριο

Γράφημα 2: παραγωγή ηλεκτρικής ενέργειας με πηγή ενέργειας απορρίμματα

Γράφημα 3: παραγωγή ηλεκτρικής ενέργειας με πηγή ενέργειας την ηλιακή ενέργεια

Γράφημα 4: παραγωγή ηλεκτρικής ενέργειας με πηγή ενέργειας την αιολική ενέργεια

Με την παρατήρηση των παραπάνω γραφημάτων γίνεται αντιληπτό, ότι το μοντέλο στην περίπτωση που είχε τη βέλτιστη απόδοση του πραγματοποίησε προβλέψεις οι οποίες ήταν πολύ κοντά στις πραγματικές τιμές της παραγωγής ενέργειας και οι αποκλίσεις ήταν πολύ μικρές και επίσης η ανωμαλία στις καμπύλες σε κάθε γράφημα σχετίζεται άμεσα με τη διασπορά των δεδομένων που χρησιμοποιήθηκαν.

Από την σύγκριση που πραγματοποιήθηκε αποδείχτηκε πως το μοντέλο που χρησιμοποιήθηκε μπορεί να πραγματοποιήσει πολύ ικανοποιητικές προβλέψεις για την παραγωγή ενέργειας για όλες τις πηγές που μελετήθηκαν αλλά συνολικά για τις νέες τεχνολογίες σε σχέση με τις παραδοσιακές πραγματοποιεί ακριβέστερη πρόβλεψη με βάση το μέσο τετραγωνικό σφάλμα (RMSE).

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 1532 πρώτες μετρήσεις για να εκπαιδευτεί και τις υπόλοιπες 385 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 που παραθέτει τα στατιστικά σφάλματα για διάφορες συναρτήσεις συμμετοχής του ANFIS αθροιστικά για τις δυο ομάδες πηγών ενέργειας που δημιουργήθηκαν, γίνεται αντιληπτό ότι στην πλειοψηφία (τέσσερις από τις έξι) των συναντήσεων ενεργοποίησης οι προβλέψεις για τις νέες τεχνολογίες με βάση το RMSE δίνουν πιο ακριβές αποτέλεσμα, κάτι που μας κάνει να συμπεράνουμε πως η πρόβλεψη μπορεί συνολικά να είναι πιο ακριβής για τις νέες τεχνολογίες. Παρόλα αυτά είναι σαφές ότι στην περίπτωση των παραδοσιακών τεχνολογιών στη συνάντηση gauss2mf το μοντέλο μας απέδωσε τα μέγιστα με άθροισμα των RMSE 18665.13612 (πίνακας 5). Έτσι μπορούμε να παρατηρήσουμε ότι ενώ στην πλειοψηφία των συναντήσεων, οι προβλέψεις για τις νέες τεχνολογίες με βάση το RMSE δίνουν πιο ακριβές αποτέλεσμα, τη μεγαλύτερη ακρίβεια την είχαμε σε μία περίπτωση παραδοσιακών τεχνολογιών. Ωστόσο βλέπουμε πως σε κάποιες συναρτήσεις, οι τιμές των προβλέψεων για τις τιμές παραγωγής ενέργειας νέων και παραδοσιακών τεχνολογιών οι αποκλίσεις είναι πολύ μικρές. Μπορούμε επίσης να συμπεράνουμε από τα σχήματα σύγκρισης πραγματικών τιμών (γραφήματα 1, 2, 3 και 4) πως οι τιμές των προβλέψεων είναι πολύ κοντά στις πραγματικές τιμές και ότι η ανωμαλία των καμπυλών σχετίζεται άμεσα με τη διασπορά των δεδομένων μας.

Ως τελικό συμπέρασμα από την σύγκριση που πραγματοποιήθηκε στην εργασία αποδείχτηκε πως το μοντέλο που χρησιμοποιήθηκε μπορεί να πραγματοποιήσει πολύ ικανοποιητικές προβλέψεις για την παραγωγή ενέργειας για όλες τις πηγές που μελετήθηκαν αλλά συνολικά για τις νέες τεχνολογίες σε σχέση με τις παραδοσιακές μπορεί να πραγματοποιηθεί ακριβέστερη πρόβλεψη.

Γενικότερα μέσα από αυτή την εργασία δείξαμε ότι η παραγωγή ηλεκτρικής ενέργειας σε καθημερινή βάση από διάφορες πηγές ενέργειας, μπορεί να συλληφθεί από ένα σύστημα ANFIS το οποίο μπορεί να μας βοηθήσει για να προβλέψουμε τις βραχυπρόθεσμες ποσότητες παραγωγής. 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 βιομηχανίες παραγωγής ενέργειας έχει γίνει ένα βασικό μέρος στον οικονομικό στρατηγικό προγραμματισμό κατά τη διάρκεια των τελευταίων ετών.

6. ΒΙΒΛΙΟΓΡΑΦΙΑ

- Abedini, R. R., Esfandyari, M. M., Nezhadmoghadam, A. A., & Rahmanian, B. B. (2012). The Prediction of Undersaturated Crude Oil Viscosity: An Artificial Neural Network and Fuzzy Model Approach. *Petroleum Science & Technology*, 30(19), 2008-2021.
- Assareh, E. E., Behrang, M. A., & Ghanbarzadeh, A. A. (2012). The Integration of Artificial Neural Networks and Particle Swarm Optimization to Forecast World Green Energy Consumption. *Energy Sources Part B: Economics, Planning & Policy*, 7(4), 398-410.
- Athanassopoulos, S. S., Mavrommatis, E. E., Gernoth, K. A., & Clark, J. W. (2004). Nuclear mass systematics using neural networks. *Nuclear Physics A*, 743(4), 222-235.
- Avami, A. A., & Boroushaki, M. M. (2011). Energy Consumption Forecasting of Iran Using Recurrent Neural Networks. *Energy Sources Part B: Economics, Planning & Policy*, 6(4), 339-347.
- Azadeh, A. A., Asadzadeh, S. M., Saberi, M. M., Nadimi, V. V., Tajvidi, A. A., & Sheikalishahi, M. M. (2011). A Neuro-fuzzy-stochastic frontier analysis approach for long-term natural gas consumption forecasting and behavior analysis: The cases of Bahrain, Saudi Arabia, Syria, and UAE. *Applied Energy*, 88(11), 3850-3859.
- Baczynski, D. D., & Parol, M. M. (2004). Influence of artificial neural network structure on quality of short-term electric energy consumption forecast. *IEEE Proceedings -- Generation, Transmission & Distribution*, 151(2), 241-245.
- Chow, S. H., Lee, E. M., & Li, D. W. (2012). Short-term prediction of photovoltaic energy generation by intelligent approach. *Energy & Buildings*, 55, 660-667.
- Demirel, Ö., Zaim, S., Çalışkan, A., & Özuyar, P. (2012). Forecasting natural gas consumption in İstanbul using neural networks and multivariate time series methods. *Turkish Journal Of Electrical Engineering & Computer Sciences*, 20(5), 695-711.
- Forouzanfar, M., Doustmohammadi, A. A., Hasanzadeh, S., & Shakouri G, H. H. (2012). Transport energy demand forecast using multi-level genetic programming. *Applied Energy*, 91(1), 496-503.
- Gonçalves, F. V., Ramos, H. M., & Reis, L. R. (2010). Hybrid energy system evaluation in water supply system energy production: neural network approach. *International Journal Of Energy & Environment*, (1), 21-30.
- Gonçalves, F., & Ramos, H. (2012). Hybrid energy system evaluation in water supply systems: artificial neural network approach and methodology. *Journal Of Water Supply: Research & Technology-AQUA*, 61(2), 59-72.

- Gueguim Kana, E. B., Oloke, J. K., Lateef, A. A., & Adesiyan, M. O. (2012). Modeling and optimization of biogas production on saw dust and other co-substrates using Artificial Neural network and Genetic Algorithm. *Renewable Energy: An International Journal*, 46276-281.
- Hernandez, L., Baladrón, C., Aguiar, J. M., Carro, B., Sanchez-Esguevillas, A. J., & Lloret, J. (2013). Short-Term Load Forecasting for Microgrids Based on Artificial Neural Networks. *Energies (19961073)*, 6(3), 1385-1408
- Jebaraj, S. S., Iniyan, S. S., & Goic, R. R. (2011). Forecasting of Coal Consumption Using an Artificial Neural Network and Comparison with Various Forecasting Techniques. *Energy Sources Part A: Recovery, Utilization & Environmental Effects*, 33(14), 1305-1316.
- López, M. M., Valero, S. S., Senabre, C. C., Aparicio, J. J., & Gabaldon, A. A. (2012). Application of SOM neural networks to short-term load forecasting: The Spanish electricity market case study. *Electric Power Systems Research*, 9118-27.
- Overview of U.S. Coal Production and Consumption. (2013). *Congressional Digest*, 92(1), 2-32.
- Ramedani, Z. Z., Omid, M. M., & Keyhani, A. A. (2012). A method based on neural networks for generating solar radiation map. *International Journal Of Energy & Environment*, 3(5), 775-786.
- Ramesh Babu, N. N., & Arulmozhivarman, P. P. (2012). Forecasting of Wind Speed Using Artificial Neural Networks. *International Review On Modelling & Simulations*, 5(5), 2276-2280
- Roldán-Blay, C., Escrivá-Escrivá, G., Álvarez-Bel, C., Roldán-Porta, C., & Rodríguez-García, J. (2013). Upgrade of an artificial neural network prediction method for electrical consumption forecasting using an hourly temperature curve model. *Energy & Buildings*, 6038-46. doi:10.1016/j.enbuild.2012.12.009
- Santos, P. J., Martins, A. G., Pires, A. J., Martins, J. F., & Mendes, R. V. (2006). Short-term load forecast using trend information and process reconstruction. *International Journal Of Energy Research*, 30(10), 811-822
- Singh, B., & Sharma, S. (2011). Neural Network Based Voltage and Frequency Controller for Isolated Wind Power Generation. *IETE Journal Of Research*, 57(5), 467-477.
- Williamson, I. (2012). Promises of Fuel Cell in Great Britain. *TCE: The Chemical Engineer*, (856), 34-37.
- Yadav, D., & Sharma, V. (2010). Artificial Neural Network based Hydro Electric Generation Modelling. *International Journal Of Applied Engineering Research (0976-4259)*, 1(3), 343-359.

Ying, L., & Pan, M. (2008). Using adaptive network based fuzzy inference system to forecast regional electricity loads. *Energy Conversion & Management*, 49(2), 205-211. doi:10.1016/j.enconman.2007.06.015