

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ



Πολυτεχνείο Κρήτης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
ΠΡΟΒΛΕΨΗ ΕΙΣΑΓΩΓΗΣ ΦΥΣΙΚΟΥ ΑΕΡΙΟΥ ΜΕ
ΝΕΥΡΟΑΣΑΦΗ ΣΥΣΤΗΜΑΤΑ (ANFIS) ΚΑΙ ΝΕΥΡΩΝΙΚΑ
ΔΙΚΤΥΑ

Επιβλέπων καθηγητής : Ατσαλάκης Γεώργιος
Επιμέλεια : Ιωάννου Σοφία-Ευαγγελία

ΧΑΝΙΑ 2013

ΠΕΡΙΕΧΟΜΕΝΑ

1	ΕΙΣΑΓΩΓΗ.....	4
2	ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ.....	5
3	ΠΡΟΤΕΙΝΟΜΕΝΗ ΠΡΟΣΕΓΓΙΣΗ	11
3.1	ΑΣΑΦΗΣ ΛΟΓΙΚΗ.....	11
3.1.1	Συνάρτηση συμμετοχής.....	11
3.1.2	Ιδιότητες των ασαφών συνόλων.....	15
3.1.3	Λογικές πράξεις στα ασαφή σύνολα	17
3.1.4	Ασαφείς σχέσεις	20
3.1.5	Ασαφής συλλογιστική	22
3.2	ΣΥΣΤΗΜΑΤΑ ΑΣΑΦΟΥΣ ΣΥΛΛΟΓΙΣΤΙΚΗΣ-ΝΕΥΡΟΑΣΑΦΗ ΣΥΣΤΗΜΑΤΑ... ..	23
3.2.1	Κανόνες αν-τότε (If-then-rules).....	24
3.2.2	Συστήματα ασαφούς συλλογιστικής (fuzzy inference systems).....	26
3.2.3	Νευροασαφή συστήματα	33
3.3	ANFIS (Adaptive Neuro-Fuzzy Inference System)	37
3.3.1	Αρχιτεκτονική του ANFIS	37
3.3.2	Αριθμητικό Παράδειγμα Μοντέλου	39
3.4	ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.....	47
3.4.1	Ορισμός νευρωνικού δικτύου.....	47
3.4.2	Ιστορικά στοιχεία	47
3.4.3	Βιολογικά νευρωνικά δίκτυα.....	50
3.4.4	Από τα Βιολογικά Νευρωνικά Δίκτυα στα Τεχνητά	50
3.4.5	Χρήση νευρωνικών δικτύων και εφαρμογές	52
3.4.6	Αρχιτεκτονική νευρωνικών δικτύων	55
3.4.7	Είδη αρχιτεκτονικών δομών	58
3.4.8	Λειτουργία νευρωνικών δικτύων.....	60
3.4.9	Εκπαίδευση νευρωνικών δικτύων	62
3.4.10	Ο Αλγόριθμος μάθησης του Perceptron (Αισθητήρα).....	64
3.4.11	Ο Αλγόριθμος Ελάχιστου Μέσου Τετραγωνικού (EMT) Σφάλματος	67
3.4.12	Ο Αλγόριθμος Πίσω Διάδοσης (Π.Δ.) του Λάθους.....	67
4	ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ (case study).....	70
4.1	ΕΙΣΑΓΩΓΗ.....	70
4.2	ΔΕΔΟΜΕΝΑ.....	70

4.3	ΠΑΡΑΜΕΤΡΟΙ ANFIS.....	70
4.4	ΑΠΟΤΕΛΕΣΜΑΤΑ.....	71
4.4.1	Ελλάδα.....	72
4.4.2	Νέα Αγγλία.....	75
4.5	ΓΡΑΦΗΜΑΤΑ.....	77
4.5.1	Ελλάδα.....	77
4.5.2	Νέα Αγγλία.....	84
5	ΣΥΜΠΕΡΑΣΜΑΤΑ.....	90
6	ΒΙΒΛΙΟΓΡΑΦΙΑ.....	91

1 ΕΙΣΑΓΩΓΗ

Η εισαγωγή του φυσικού αερίου στο ενεργειακό ισοζύγιο της Ελλάδας έχει ήδη επηρεάσει σημαντικούς κλάδους της οικονομικής και κοινωνικής ζωής της χώρας. Αναμένεται μάλιστα ακόμα μεγαλύτερη διείσδυση καθώς αποτελεί την καθαρότερη πηγή πρωτογενούς ενέργειας, μετά τις ανανεώσιμες μορφές, και η τιμή του κυμαίνεται σε χαμηλότερα επίπεδα από αυτά του πετρελαίου.

Το φυσικό αέριο είναι ένα φυσικό προϊόν που βρίσκεται σε υπόγεια κοιτάσματα της γης και είτε συναντάται μόνο του είτε συνυπάρχει με κοιτάσματα πετρελαίου. Είναι μίγμα υδρογονανθράκων σε αέρια κατάσταση, αποτελούμενο κυρίως από μεθάνιο (σε ποσοστό άνω του 85%), που είναι ο ελαφρύτερος υδρογονάνθρακας, είναι πολύ καθαρό, χωρίς προσμίξεις και θειούχα συστατικά για το λόγο αυτό το φυσικό αέριο αποτελεί το φιλικότερο συμβατικό καύσιμο στο περιβάλλον και στον άνθρωπο. Το φυσικό αέριο υπάρχει σε μεγάλα αποθέματα που ήδη έχει διαπιστωθεί ότι επαρκούν τουλάχιστον 100 έτη σε χώρες όπως η πρώην Σοβιετική Ένωση, το Ιράν, το Κατάρ, το Ιράκ, η Νιγηρία, η Αλγερία, οι ΗΠΑ κ.ά. Η Ελλάδα σήμερα προμηθεύεται φυσικό αέριο από 3 διαφορετικές πηγές:

- από τη Ρωσία (μέσω Βουλγαρίας) μέσω αγωγών σε αέρια μορφή,
- από την Αλγερία με δεξαμενόπλοια σε υγροποιημένη μορφή (στις εγκαταστάσεις της νήσου Ρεβυθούσας, στον κόλπο των Μεγάρων) και
- από το 2007, από το Αζερμπαϊτζάν (μέσω Τουρκίας) μέσω αγωγών σε αέρια μορφή.

Στη παρούσα εργασία, πραγματοποιείται πρόβλεψη εισαγωγής φυσικού αερίου στην Ελλάδα και στη Νέα Αγγλία (ΗΠΑ). Η πρόβλεψη χρησιμοποιεί ανάπτυξη κώδικα που τρέχει στο περιβάλλον της Matlab και βασίζεται στο μοντέλο Adaptive Neural Fuzzy Inference System (ANFIS) και στο μοντέλο Neural Networks (NN). 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, παρατίθενται τα αποτελέσματα από την εφαρμογή του μοντέλου (γραφήματα και σφάλματα), ο αναλυτικός σχολιασμός τους και η σύγκριση του μοντέλου με τα αντίστοιχα AR και ARMA.



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.

Analysis about the seasonality of China's crude oil import based on X-12-ARIMA

Ο σκοπός των Zhou και Dong στη παρούσα μελέτη ήταν να εξετάσουν την πιθανή εποχικότητα των εισαγωγών αργού πετρελαίου της Κίνας, με την ελπίδα ότι θα βοηθήσουν τους ενδιαφερόμενους φορείς στον έλεγχο της απογραφής και τον προγραμματισμό της παραγωγής. Έτσι χρησιμοποίησαν τη X-12-ARIMA μέθοδο για την προσαρμογή της μηνιαίας και της τριμηνιαίας εισαγωγής καθαρού αργού πετρελαίου στη Κίνας τα τελευταία 16 χρόνια. Τα αποτελέσματα έδειξαν ότι η τριμηνιαία σειρά είναι καλύτερη απ' ότι η μηνιαία σειρά. Και όσο αφορά τη προσαρμογή της εποχικότητας, οι εποχιακοί παράγοντες τείνουν να είναι θετικοί την άνοιξη και το καλοκαίρι, ενώ αρνητικοί το φθινόπωρο και το χειμώνα. Επιπλέον συμπέραναν ότι οι τρεις εποχιακούς παράγοντες γίνονται ισχυρότεροι, ενώ ο συντελεστής του χειμώνα ασθενέστερος κατά τα τελευταία έτη.

Assessing New England's waterborne energy imports (1995–2004)

Εδώ ο Marti χρησιμοποιώντας ετήσια δεδομένα και με τη χρήση της μεθόδου αντιστοιχίας μελέτησε τη πλωτή εισαγωγή ενέργειας στη Νέα Αγγλία, στο λιμάνι της εισόδου, από το 1995 μέχρι και το 2004. Η καταγραφή των στοιχείων που αφορούν την εισαγωγή προϊόντων ενέργειας στη Νέα Αγγλία αποκάλυψε ότι υπάρχει μια ιεραρχική σύνθεση των μεγάλων, μεσαίων και μικρών λιμανιών. Ήταν σαφές ότι τα προϊόντα ενέργειας ήταν ο στυλοβάτης των ροών εισαγωγής στα περισσότερα λιμάνια , σε ετήσια βάση , όταν η κίνηση της ενέργειας υπήρχε, συχνά κέρδιζαν τιμές κοντά ή πάνω από το 90 % του συνόλου των πλωτών κυκλοφορίας.

Forecasting natural gas supply in China: Production peak and import trends

Οι Lin και Wang χρησιμοποιώντας ετήσια δεδομένα για τις χρονιές 1996-2006 και με τη χρήση της λογικής καμπύλης και της καμπύλης Gauss προέβλεψαν τη κατανάλωση φυσικού αερίου στη Κίνα και τη παραγωγή. Έτσι η καθαρή εισαγωγή αερίου ως εκ τούτου μπορεί να υπολογιστεί βάσει του σεναρίου 2. Όσο κι αν ο ρυθμός αύξησης της ζήτησης του φυσικού αερίου είναι χαμηλή, μεσαία ή υψηλή, ο όγκος των καθαρών εισαγωγών φυσικού αερίου θα αυξηθεί και η ξένη εξάρτηση θα είναι όλο και υψηλότερη. Κατέληξαν στο συμπέρασμα πως με το ρυθμό αύξησης που έχει η ζήτηση του φυσικού αερίου και με τη πρόβλεψη παραγωγής, το ποσό εισαγωγών θα αυξάνεται με γρήγορο ρυθμό μέχρι και το 2030.

Kalman filter estimation of monthly U.S oil imports

Οι Martz, Burris, Bruckner και Bivins το 1984 έχοντας για δεδομένα μετρήσεις από 20 μήνες (Γενάρης 1979- Αύγουστος 1980) χρησιμοποίησαν το φίλτρο Kalman για την πρόβλεψη μηνιαίων εισαγωγών πετρελαίου στις ΗΠΑ με δύο μήνες οδηγό. Η εξίσωση κατάστασης που χρησιμοποιήθηκε ήταν το μοντέλο ARIMA και η εξίσωση παρατήρησης ήταν η οπισθοδρόμηση της. Τα αποτελέσματα έδειξαν ότι οι εκτιμήσεις του φίλτρου Kalman είναι ανώτερες στην εταιρεία τιμών που παρέχονται. Οι συνολικές προβλέψεις είναι σε συμφωνία με πραγματικές αξίες και εντός των 95% ορίων εμπιστοσύνης.

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.

Forecasting natural gas consumption in Istanbul 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 παρόν 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.

A statistical model for the estimation of natural gas consumption

Σε αυτό το άρθρο οι Vondracek, Pelikan, Konar, Cermakova, Eben, Maly και Brabec παρουσιάζουν μια στατιστική προσέγγιση για την εκτίμηση της κατανάλωσης φυσικού αερίου μεμονωμένων κατοικιών και μικρών εμπορικών πελατών. Η προσέγγιση αυτή βασίστηκε στην αρχή της μη γραμμικής παλινδρόμησης. Οι παράμετροι υπολογίστηκαν χρησιμοποιώντας κυρίως δύο πραγματικά σύνολα δεδομένων, συνήθεις (περίπου ετήσια) ενδείξεις μέτρησης για σχεδόν όλους τους πελάτες και επιπλέον (περίπου μηνιαία) ενδείξεις μέτρησης σχεδιασμένες και να λειτουργούν στο πλαίσιο της συνεργασίας μεταξύ του Ινστιτούτου Πληροφορικής της Τσεχικής Ακαδημίας Επιστημών (ICS) και της Εταιρείας Διανομής Αερίου της Δυτικής Βοημίας, ένα μέρος του ομίλου RWE (WBG). Το μοντέλο δοκιμάστηκε σε διάφορα σύνολα δεδομένων. Έχει ευρεία εφαρμογή σε πολλούς τομείς της βιομηχανίας φυσικού αερίου. Ο κύριος στόχος αυτής της εργασίας ήταν αντί να εισάγει μια νέα στατιστική μέθοδο να δημιουργήσει μια όψη στο έργο της εφαρμογής προσανατολισμένης στον ερευνητή, ο οποίος οφείλει να εφαρμόζει και να τροποποιεί γνωστές μεθόδους το οποίο περιορίζεται από την ποιότητα και την ποσότητα των δεδομένων αλλά και από τις απαιτήσεις του τελικού χρήστη.

A mathematical modeling for incorporating energy price hikes into total natural gas consumption forecasting

Οι Dalfard, NazariAsli, Asadzadeh, Sajjadi και Nazari-Shirkouhi προτείνουν ένα ολοκληρωμένο προσαρμοζόμενο ασαφές σύστημα (FIS) για την πρόβλεψη μακροπρόθεσμης κατανάλωσης φυσικού αερίου, όταν οι τιμές βιώνουν μεγάλη αύξηση. Για να ενσωματώσουν τις επιπτώσεις της αύξησης των τιμών στην μοντελοποίηση, πρότειναν μια νέα διαδικασία για την κατασκευή και προσαρμογή του Takagi-Sugeno ασαφούς συστήματος (TS-FIS). Χρησιμοποιήθηκαν γραμμικές παλινδρομήσεις για την κατασκευή μιας πρώτης τάξης TS- FIS και για την πρόβλεψη κατανάλωσης φυσικού αερίου σε σταθμούς ηλεκτροπαραγωγής χρησιμοποιήθηκε προσαρμοζόμενο με βάση το δίκτυο FIS (ANFIS). Για να αντιμετωπίσει με τυχαία αβεβαιότητα σε μικρά σύνολα δεδομένων , χρησιμοποιείται η Monte Carlo προσομοίωση για την παραγωγή των δεδομένων κατάρτισης για ANFIS. Τα δεδομένα

που συλλέχθηκαν από τη στατιστική έκθεση της βιομηχανίας ηλεκτρικής ενέργειας του Ιράν είναι ετήσια για τα έτη 1997-2009. Τα αποτελέσματα της μελέτης έδειξαν ότι το 2010-2016, οι υψηλές τιμές της ενέργειας μειώνουν σημαντικά την κατανάλωση του φυσικού αερίου. Παρά το γεγονός ότι, αυτή η πρόβλεψη έχει τις ρίζες της στην γνώση των εμπειρογνομώνων του τομέα, όμως υπάρχουν και άλλα μέτρα βελτίωσης της κατανάλωσης που μπορούν να υποστηρίξουν η σταθερότητα και η ευρωστία της πρόβλεψης. Με την αύξηση των τιμών εμφανίζεται αλλαγή στη συμπεριφορά των τελικών χρηστών της κατανάλωσης και το πιο σημαντικό, η στροφή της παραγωγής ενέργειας προς τις ανανεώσιμες καθιστούν πιο πιθανό να γίνει πραγματικότητα η πρόβλεψη.

A nonlinear mixed effects model for the prediction of natural gas consumption by individual customers

Η μελέτη αυτή οι Brabec, Konár, Pelikán και Malý ασχολήθηκαν με την περιγραφή και την πρόβλεψη της ημερήσιας κατανάλωσης φυσικού αερίου σε επίπεδο μεμονωμένων πελατών. Αντίθετα από τις παραδοσιακές προσεγγίσεις μέσου όρου της ομάδας, βρισκόμαστε αντιμέτωποι με τις παρατυπίες των επιμέρους σειρών κατανάλωσης που θέτει η μεταξύ των ατόμων ετερογένεια, περιλαμβανομένων και των μηδενικών, δεδομένα που λείπουν, και οι αιφνίδιες αλλαγές στο καταναλωτικό πρότυπο. Το μοντέλο που χρησιμοποίησαν είναι η μη γραμμική τύπου παλινδρόμηση, με ειδικές παραμέτρους για κάθε πελάτη που, παρ'όλα αυτά, έχουν μια κοινή διανομή που αντιστοιχεί στο μη γραμμικό μοντέλο μικτού πλαισίου. Η απόδοση της πρόβλεψης του προτεινόμενου μοντέλου αποδεικνύεται χρησιμοποιώντας ένα πραγματικό σύνολο δεδομένων από 62 μεμονωμένους πελάτες, και σε σύγκριση με δύο από τις πιο παραδοσιακές προσεγγίσεις: ARIMAX και ARX.

An eclectic approach in energy forecasting: a case of Natural Resources Canada's (NRCan's) oil and gas outlook

Εδώ οι Persaud και Kumar παρουσίασαν χρήσιμες γνώσεις σχετικά με τις πολυπλοκότητες της προβολής εφοδιασμού σε πετρέλαιο και φυσικό αέριο, ανέπτυξαν ένα αναλυτικό πλαίσιο που εξηγεί την προσέγγιση που χρησιμοποιείται από τους φυσικούς πόρους του Καναδά (NRCan) στην προετοιμασία του πετρελαίου και του φυσικού αερίου για προβλέψεις εφοδιασμού, έλυσαν το μοντέλο εφοδιασμού (OGSM) και έδωσαν τις προβλέψεις προσφοράς και ζήτησης του πετρελαίου και του φυσικού αερίου για το έτος 2020. Τα δεδομένα που χρησιμοποιήθηκαν στο μοντέλο ήταν ετήσια.

Modeling and forecasting natural gas demand in Bangladesh

Στο παρόν άρθρο οι Wadud, S.Dey, Kabir και Khan υπάρχουν δύο δυνητικά ανταγωνιστικά μοντέλα για τη ζήτηση φυσικού αερίου: μία όπου οι ερμηνευτικές μεταβλητές περιλαμβάνουν το συνολικό ΑΕΠ (μοντέλο Α), και εκείνη που περιέχει το κατά κεφαλήν ΑΕΠ και τον πληθυσμό (μοντέλο Β) που προφανώς έχουν άλλες μεταβλητές, όπως η τιμή του φυσικού αερίου και η κατανάλωσή του. Τα δεδομένα

που εισήγαγαν στα δύο μοντέλα ήταν ετήσιας μορφής. Τα στατιστικά μοντέλα και των δύο είναι παρόμοια με το μοντέλο A οριακά καλύτερο από το μοντέλο B. Έτσι βρήκαν μια μακροχρόνια ανταπόκριση της ζήτησης του φυσικού αερίου σε σχέση με το ΑΕΠ / έσοδα. Χρησιμοποιώντας το οικονομετρικό τους μοντέλο, άντλησαν και τις προβλέψεις για την κατανάλωση φυσικού αερίου στο Μπαγκλαντές. Οι προβλέψεις τους ήταν σημαντικά μεγαλύτερες από τις προηγούμενες προβλέψεις της ζήτησης από πολυεθνικές υπηρεσίες, αλλά το πιο πιθανό θεωρώντας υψηλό ρυθμό αύξησης του ΑΕΠ του Μπαγκλαντές την επόμενη δεκαετία

Univariate modeling and forecasting of monthly energy demand time series using abductive and neural networks

Ο Abdel-Aal σε αυτό το άρθρο χρησιμοποίησε μονομεταβλητή μοντελοποίηση των μηνιαίων χρονοσειρών ζήτησης βασιζόμενο μόνο σε στοιχεία 6 ετών για την πρόβλεψη της ζήτησης για το έβδομο έτος. Και τα νευρωνικά και τα απαγωγικά δίκτυα χρησιμοποιήθηκαν για τη μοντελοποίηση και συγκρίθηκε η απόδοσή τους. Επίσης χρησιμοποίησε μια απλή τεχνική για την αφαίρεση της ανοδικής τάσης ανάπτυξης πριν τη διαμόρφωση της χρονοσειράς ζήτησης για την αποφυγή προβλημάτων που συνδέονται με την παρέκταση πέρα από τη χρησιμοποιούμενη περιοχή δεδομένων. Ερευνήθηκαν και συγκρίθηκαν δύο προσεγγίσεις μοντελοποίησης: στη πρώτη χρησιμοποιώντας επαναληπτικά ένα επόμενο προβλεπόμενο μήνα, και στη δεύτερη απασχολώντας 12 ειδικά μοντέλα για την πρόβλεψη των 12 επιμέρους μηνών. Τα αποτελέσματα δείχνουν καλύτερη απόδοση στην πρώτη προσέγγιση, με μέσο όρο ποσοστό σφάλματος (MAPE) της τάξης του 3% για απαγωγικά δίκτυα.

The degree-day method to estimate the residential heating natural gas consumption in Turkey: a case study

Θέλοντας να μελετήσουν τη κατανάλωση φυσικού αερίου, στη Τουρκία, για θέρμανση οι Sarak και Satman χρησιμοποίησαν μια μέθοδο για το βαθμό ημερήσιας θέρμανσης. Για την έρευνα επέλεξαν πόλεις που βρίσκονται κοντά στους ήδη υπάρχοντες, υπό κατασκευή ή ακόμα και σε προγραμματισμένους αγωγούς. Τα αποτελέσματα έδειξαν ότι το 2023, αν το 100% των καταναλωτών στη Τουρκία χρησιμοποιούν φυσικό αέριο για θέρμανση χώρου, η κατανάλωση θα αυξηθεί σημαντικά. Επιπλέον για ελέγξουν την αξιοπιστία της μεθόδου σύγκριναν τα αποτελέσματα της πρόβλεψης με τα πραγματικά στοιχεία για το 1998 και το 2000.

European energysecurity:Ananalys is of future Russian natural gas production and exports

Οι Soderbergh, Jakobsson και Aleklett προβλέπουν μια αύξηση της τάξης του 87% όσο αφορά την εισαγωγή φυσικού αερίου στην Ευρώπη μεταξύ των ετών 2006 και 2030. Ως εκ τούτου, θέλησαν να μελετήσουν τη παραγωγή φυσικού αερίου στη Ρωσία για τα επόμενα έτη. Για να τα καταφέρουν χρησιμοποίησαν το μοντέλο γιγάντιων κοιτασμάτων αερίου (GGF-model) στο οποίο για τη χρήση του

απαιτήθηκαν πολλά ποσά παραγωγής. Το αποτέλεσμα της πεδίο μελέτης των 83 γιγαντιαίων κοιτασμάτων φυσικού αερίου καταδεικνύει ότι οι μεγάλες παραγωγές ρωσικών κοιτασμάτων φυσικού αερίου βρίσκονται σε παρακμή, και από το 2013 πολύ μεγαλύτερες προμήθειες από τη χερσόνησο Γιαμάλ και το πεδίο Shtokman θα απαιτηθούν προκειμένου να αποφευχθεί η μείωση της παραγωγής. Φυσικό αέριο από κοιτάσματα στην Ανατολική Σιβηρία και την Άπω Ανατολή θα προμηθεύουν κυρίως χώρες της Ασίας και του Ειρηνικού. Ως αποτέλεσμα, η μέγιστη αύξηση των εξαγωγών προς τις ευρωπαϊκές και να ανέρχεται μόνο σε περίπου 45% για την περίοδο 2015-2030.

Forecasting the growth of Chinese natural gas consumption

Η χρήση του φυσικού αερίου στην Κίνα είναι ακόμη σχετικά ανώριμη , καθώς η παραγωγή φυσικού αερίου παρέχει μόνο ένα μικρό ποσοστό του εγχώριου ενεργειακού συστήματος. Αντίθετα , η κινεζική οικονομία βασίζεται κυρίως στον άνθρακα με μερίδιο 67 % της συνολικής πρωτογενούς ενέργειας . Οι περιβαλλοντικές επιπτώσεις από την υψηλή εξάρτηση από τον άνθρακα είναι σημαντική και θέλουν να καταφύγουν σε καθαρότερες πηγές ενέργειας . Το φυσικό αέριο είναι τόσο καθαρό και γενικά πιο αποτελεσματικό από τον άνθρακα και λόγω αυτών των γεγονότων η κατανάλωση φυσικού αερίου αυξάνεται γρήγορα. Για τον ποσοτικό προσδιορισμό της τάσης αυτής , οι Li, Dong, Shangguan και Høök χρησιμοποίησαν ένα πρότυπο σύστημα δυναμικής με λογισμικό STELLA και δημιούργησαν μια πιθανή προοπτική. Τα δεδομένα που εισήγαγαν στο σύστημα ήταν ετήσια από το 1997 μέχρι το 2007 Τα αποτελέσματα είχαν δείξει ότι η κατανάλωση φυσικού αερίου στην Κίνα θα συνεχίσει να αυξάνεται γρήγορα σε 89,5 δισεκατομμύρια κυβικά μέτρα το 2010 και το 2030 θα φτάσει τα 340,7 δισεκατομμύρια κυβικά μέτρα. Επίσης, διαπίστωσαν ότι η δομή της κατανάλωσης θα αλλάξει και χρησιμοποιήθηκε για την καλύτερη εκτίμηση ανάλυση σεναρίου. Τέλος , το παρόν έγγραφο δίνει κάποιες προτάσεις πολιτικής για το φυσικό αέριο και την ανάπτυξη, κατασκευή υποδομών και τεχνικές καινοτομίες για την προώθηση της βιώσιμης ανάπτυξης της βιομηχανίας φυσικού αερίου της Κίνας .

Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods

Θέλοντας να προβλέψουν την ημερήσια κατανάλωση φυσικού αερίου οι Taspinar, Celebi και Tutkun και χρησιμοποιώντας τρία είδη μοντέλων, το μοντέλο SARIMAX, το ANN-MLP μοντέλο και το ANN-RBF μοντέλο, τα οποία χρησιμοποιούν κάποια μετεωρολογικά δεδομένα (υγρασία, ατμοσφαιρική πίεση, ταχύτητα του ανέμου και θερμοκρασία του περιβάλλοντος). Για να χρησιμοποιήσουν τα μοντέλα χρειάστηκαν 1800 ημερήσια δεδομένα κατανάλωσης φυσικού αερίου που είναι στο διάστημα των ετών 2007-2011. Παρόλο που ο αριθμός των δεδομένων ήταν μικρός, ο προτεινόμενος αλγόριθμος λειτούργησε καλά για τη βραχυπρόθεσμη πρόβλεψη της κατανάλωσης φυσικού αερίου και είχαν ενθαρρυντικά αποτελέσματα για τη μελλοντική επενδυτική πολιτική της ενέργειας.

3 ΠΡΟΤΕΙΝΟΜΕΝΗ ΠΡΟΣΕΓΓΙΣΗ

3.1 ΑΣΑΦΗΣ ΛΟΓΙΚΗ

Ο θεωρητικός φορέας για την υλοποίηση μιας μεγάλης κατηγορίας ευφύων συστημάτων είναι η Ασαφής Λογική (FuzzyLogic), που εισήχθη το 1962 από τον LotfiA Zadeh του πανεπιστημίου της Καλιφόρνιας στο Berkley των ΗΠΑ. Η θεωρία των ασαφών συνόλων καθιερώθηκε σαν μια μέθοδος 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.

Τα λογικά παράδοξα και η αρχή της απροσδιοριστίας του Heisenberg συνέβαλαν στην ανάπτυξη της πλειότιμης λογικής (multi-valuedlogic) στη δεκαετία του 1920. Το επόμενο βήμα επέτρεψε βαθμούς αβεβαιότητας και θεωρούσε τους χαρακτηρισμούς αλήθειας και ψεύδους ως τις ακραίες τιμές ενός πεδίου μεταβολής της ασάφειας. Το 1937 ο μαθηματικός MaxBlack εφάρμοσε τη συνεχή λογική σε σύνολα στοιχείων και συμβόλων και ονόμασε τη δομή αυτή αβεβαιότητα. Το 1965, ο Zadeh στη θεμελιώδη δημοσίευσή του με τίτλο “FuzzySets”, παρουσίασε τη θεωρία των ασαφών συνόλων (fuzzysettheory) χρησιμοποιώντας τον όρο Ασαφή Λογική και πρωτοστάτησε σε ένα νέο και σημαντικό κύμα επιστημονικής δραστηριότητας, που συνεχίζεται μέχρι σήμερα. Δεν ήταν λίγοι, όμως, και ειδικά στις ΗΠΑ, που αμφισβήτησαν τη θεωρία του Zadeh. Αντίθετα, η Κομφουκιανή Ιαπωνία έχει ενστερνιστεί τη θεωρία απόλυτα και η χρήση της αποδίδει τεράστια οικονομικά οφέλη.

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

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

3.1.1 Συνάρτηση συμμετοχής

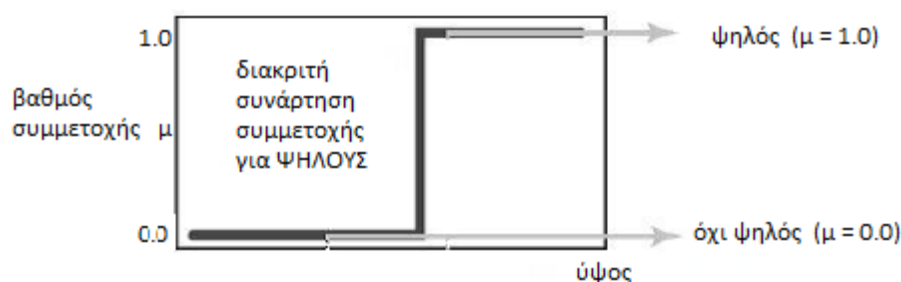
Ένα ασαφές σύνολο A ορίζεται ως ένα σύνολο διατεταγμένων ζευγών $(x, \mu_A(x))$ όπου $x \in X$ και $\mu_A(x) \in [0,1]$. Το σύνολο X αποτελεί ένα ευρύτερο σύνολο αναφοράς (universe of discourse) που περιλαμβάνει όλα τα αντικείμενα στα οποία μπορεί να γίνει αναφορά. Η τιμή $\mu_A(x)$ λέγεται βαθμός αλήθειας, συμβολίζει το βαθμό συγγένειας του x στο A και παίρνει τιμές στο διάστημα $[0,1]$. Τέλος η συνάρτηση μ_A ονομάζεται συνάρτηση συμμετοχής (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 [$\mu_A(x)=0$]. Άρα η ασαφής θεωρία συνόλων μεταπίπτει στην αντίστοιχη κλασσική, όταν οι δυνατές τιμές της συνάρτησης συμμετοχής είναι μόνο 0 ή 1. 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)

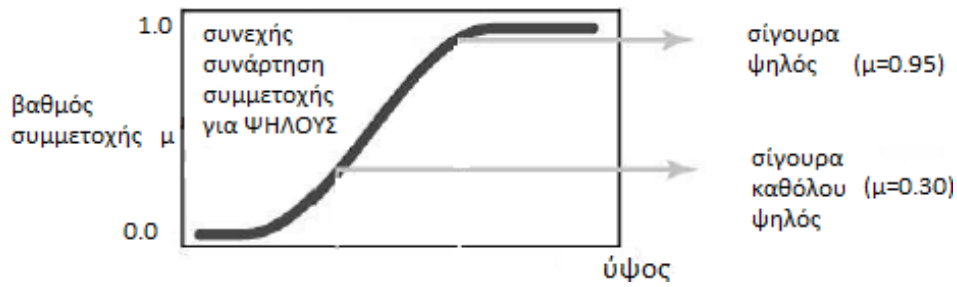
Παρακάτω αναπτύσσεται ένα παράδειγμα με το οποίο μπορεί να γίνει περισσότερο κατανοητή η έννοια της συνάρτησης συμμετοχής.

Ας υποθέσουμε ότι ο χώρος αναφοράς X είναι το σύνολο όλων των ανθρώπων. Ένα ασαφές υποσύνολο του χώρου αυτού είναι οι ψηλοί άνθρωποι. Τα πιθανά ύψη έστω ότι κυμαίνονται από 1.20 μέχρι 2.50. Η λέξη ψηλός μπορεί να συσχετισθεί με μια καμπύλη η οποία δείχνει κατά πόσο ένας άνθρωπος είναι ψηλός ή όχι. Αν χρησιμοποιήσουμε τις αρχές των κλασσικών συνόλων τότε για να ορίσουμε το σύνολο των ψηλών ανθρώπων θα πρέπει να ορίσουμε μια συγκεκριμένη τιμή ύψους, η οποία θα διαχωρίζει τους ανθρώπους σε ψηλούς και κοντούς. Π.χ. ας υποθέσουμε ότι η τιμή αυτού του ύψους είναι 1.75m τότε ένας άνθρωπος με ύψος 1.73 θα χαρακτηρίζεται κοντός ενώ ένας άνθρωπος με ύψος 1.77 θα χαρακτηρίζεται ψηλός. Ο παραπάνω διαχωρισμός φαίνεται παράλογος αφού έχουμε αντιστοιχήσει σε δύο ανθρώπους με αμελητέα διαφορά ύψους δύο αντίθετες μεταξύ τους έννοιες.



Σχήμα 3.1 Κλασσική θεωρία των συνόλων

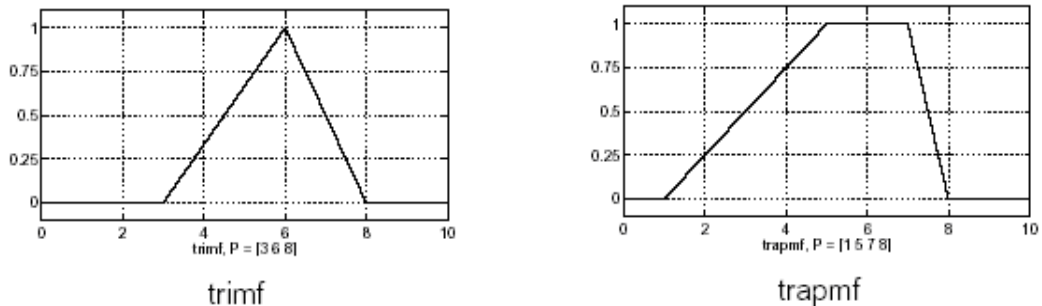
Ένας άλλος τρόπος να ορίσουμε την έννοια ψηλός είναι μέσω μιας καμπύλης που έχει ομαλή διακύμανση και μεταβαίνει από την έννοια ψηλός στην έννοια κοντός. Αυτή η καμπύλη είναι η συνάρτηση συμμετοχής του ασαφούς συνόλου των ψηλών ανθρώπων. Με άλλα λόγια δεχόμαστε ότι όλοι οι άνθρωποι είναι σε κάποιο βαθμό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



Σχήμα 3.2 Θεωρία των Ασαφών συνόλων

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

Οι απλούστερες συναρτήσεις συμμετοχής είναι αυτές που σχηματίζονται από ευθείες γραμμές. Η απλούστερη από αυτές είναι η τριγωνική συνάρτηση συμμετοχής, που δεν είναι τίποτα άλλο από ένα τρίγωνο. Στην ίδια κατηγορία ανήκει και η τραπεζοειδής συνάρτηση συμμετοχής (σχήμα 3.3)



Σχήμα 3.3 Συναρτήσεις συμμετοχής, τριγωνική και τραπεζοειδής

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.

Ενώ η μαθηματική έκφραση της τριγωνικής συνάρτησης συμμετοχής (trimf) είναι η εξής:

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

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

Ενώ η μαθηματική έκφραση της τραπεζοειδούς συνάρτησης συμμετοχής (trapmf) είναι:

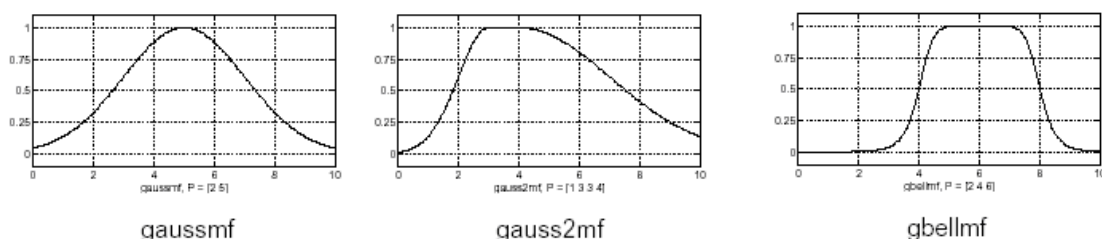
$$A = \begin{cases} 0, & x \leq a \\ (x-a)/(b-a), & x \in (a,b) \\ 1, & x \in (b,c) \\ (d-x)/(d-c), & x \in (c,d) \end{cases}$$

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

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

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

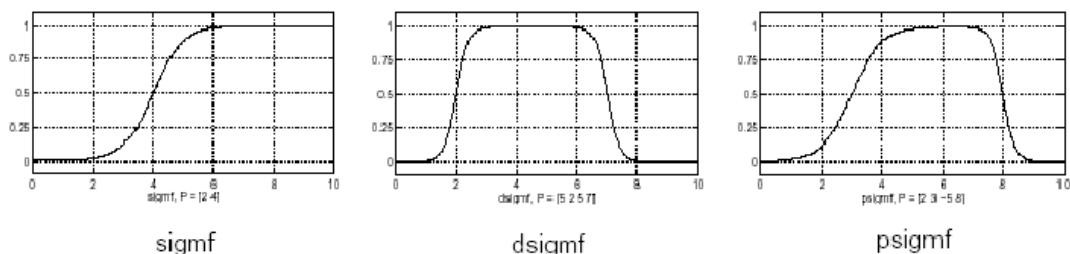
$$\text{gbellmf}(x, a, b, c) = \frac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}}$$



Σχήμα 3.4 Συναρτήσεις συμμετοχής, Γκαουσιανή, Διπλή Γκαουσιανή, Καμπανοειδής

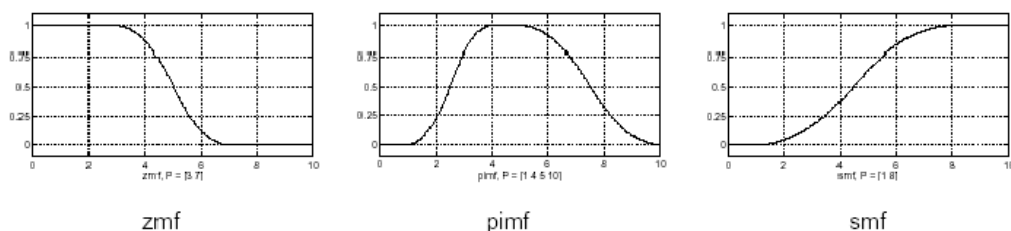
Παρά το γεγονός ότι η γκαουσιανή συνάρτηση συμμετοχής και η συνάρτηση καμπάνας επιτυγχάνουν ομαλή διακύμανση δε μπορούν ωστόσο να ορίσουν ασύμμετρες συναρτήσεις συμμετοχής που είναι χρήσιμες σε πολλά πρακτικά προβλήματα. Για τον λόγο αυτό χρησιμοποιείται η σιγμοειδής συνάρτηση συμμετοχής η οποία είναι ασύμμετρη και ανοικτή είτε προς τα δεξιά είτε προς τα αριστερά. Κλειστές συναρτήσεις συμμετοχής αυτού του τύπου μπορούν να

παραχθούν αν συνθέσουμε δύο σιγμοειδείς. Έτσι προκύπτει η διαφορά μεταξύ δύο σιγμοειδών και το άθροισμα τους. (σχήμα 3.5)



Σχήμα 3.5 Συναρτήσεις συμμετοχής σιγμοειδής, διαφορά σιγμοειδών, άθροισμα σιγμοειδών

Επίσης υπάρχουν πολλές πολωνυμικές καμπύλες (σχήμα 3.6) που τις χρησιμοποιούμε σε συναρτήσεις συμμετοχής. Τρεις από αυτές είναι η Z η S και η Π οι οποίες έχουν ονομασθεί έτσι εξαιτίας του σχήματός τους. Η Z είναι μια ασύμμετρη πολωνυμική καμπύλη που είναι ανοικτή στα αριστερά, η S είναι η κατοπτρική της Z και η Π είναι μια ασύμμετρη κλειστή καμπύλη σχήματος Π.



Σχήμα 3.6 Συναρτήσεις συμμετοχής Z, Π, S

3.1.2 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 .

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

Το σύνολο A μπορεί να γραφεί και ως:

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

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

Στο συνεχή χώρο αναφοράς:

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

Ασαφές δυναμοσύνολο:

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

Υποσύνολο:

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

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

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

Ασαφής διαμέριση:

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

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

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

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

Κενό ασαφές σύνολο:

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

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

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

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

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:

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

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:

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

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.

3.1.3 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 εφαρμοστούν. Για παράδειγμα ας θεωρήσουμε τον παρακάτω πίνακα αληθείας:

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

AND

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

OR

A	not A
0	1
1	0

NOT

Σχήμα 3.7 Λογικοί κανόνες

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		

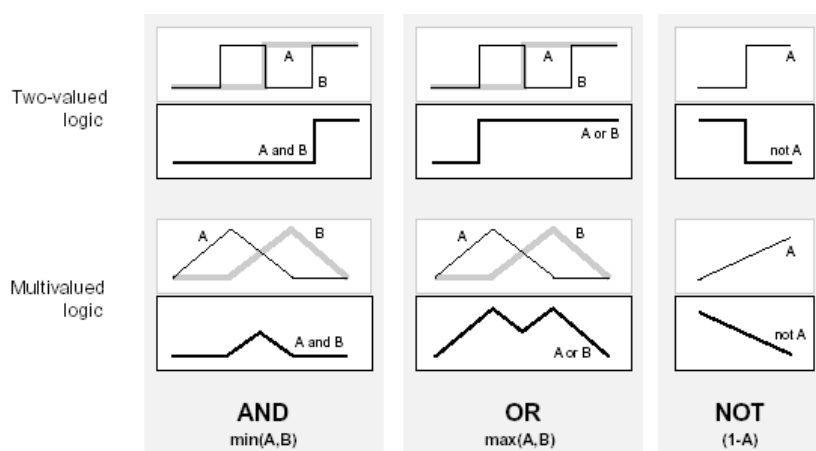
AND

OR

NOT

Σχήμα 3.8 Λογικοί κανόνες

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



Σχήμα 3.9: Κατασκευή νέων καμπυλών από ήδη υπάρχουσες, με τη βοήθεια των τελεστών

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

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

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

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

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.

Μερικοί από τους πιο γνωστούς τελεστές T-norm είναι:

- Συνήθης τομή $T(a,b)=\min(a,b)$
- Αλγεβρικό γινόμενο $T(a,b)=ab$
- Φραγμένη διαφορά $T(a,b)=\max(0,a+b-1)$
- Συνάρτηση Hamacher Η συνάρτηση Hamacher δίνεται από τη σχέση:

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

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

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

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

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

Παραδείγματα σ-νορμών είναι τα εξής:

Συνήθης ένωση $s(a,b)=\max(a,b)$

Αλγεβρικό άθροισμα $s(a,b)=a+b-ab$

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

Συμπλήρωμα

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

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

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:

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

3.1.4 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:

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

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:

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

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:

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

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

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

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

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

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

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

3.1.5 Ασαφής συλλογιστική

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

Οι συλλογιστικοί τρόποι που κυρίως χρησιμοποιούνται είναι τρεις:

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

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

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

όπου A και B συγκεκριμένα γεγονότα.

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

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

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

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

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

Πρέπει, λοιπόν να προσδιοριστεί ο βαθμός στον οποίο πληρείται το γεγονός B. Αυτός εξαρτάται από το βαθμό στον οποίο πληρείται το γεγονός A και από το είδος της

συνεπαγωγής που εφαρμόζουμε. Η πράξη της ασαφούς συνεπαγωγής υλοποιεί τη μαθηματική σχέση $A \Rightarrow B$, όταν τα A και B είναι ασαφή γεγονότα.

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

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

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

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

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

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

3.2 ΣΥΣΤΗΜΑΤΑ ΑΣΑΦΟΥΣ ΣΥΛΛΟΓΙΣΤΙΚΗΣ-ΝΕΥΡΟΑΣΑΦΗ

ΣΥΣΤΗΜΑΤΑ

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, 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)

3.2.1 Κανόνες αν-τότε (If-then-rules)

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

if x is A then y is B

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

if service is good then tip is average

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

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

if service == good then tip = average

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

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

Αν πάρουμε την απλή περίπτωση των δυαδικών τιμών οι κανόνες if-then είναι αρκετά εύκολο να εφαρμοσθούν. Αν η προϋπόθεση είναι αληθής τότε και η συνέπεια θα είναι αληθής. Πώς όμως επεκτείνεται η παραπάνω συλλογιστική για να

συμπεριλάβει και ασαφείς τιμές; Όταν η προϋπόθεση είναι αληθής σε κάποιο βαθμό τότε και η συνέπεια είναι αληθής στον ίδιο βαθμό. Με άλλα λόγια:

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

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

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

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

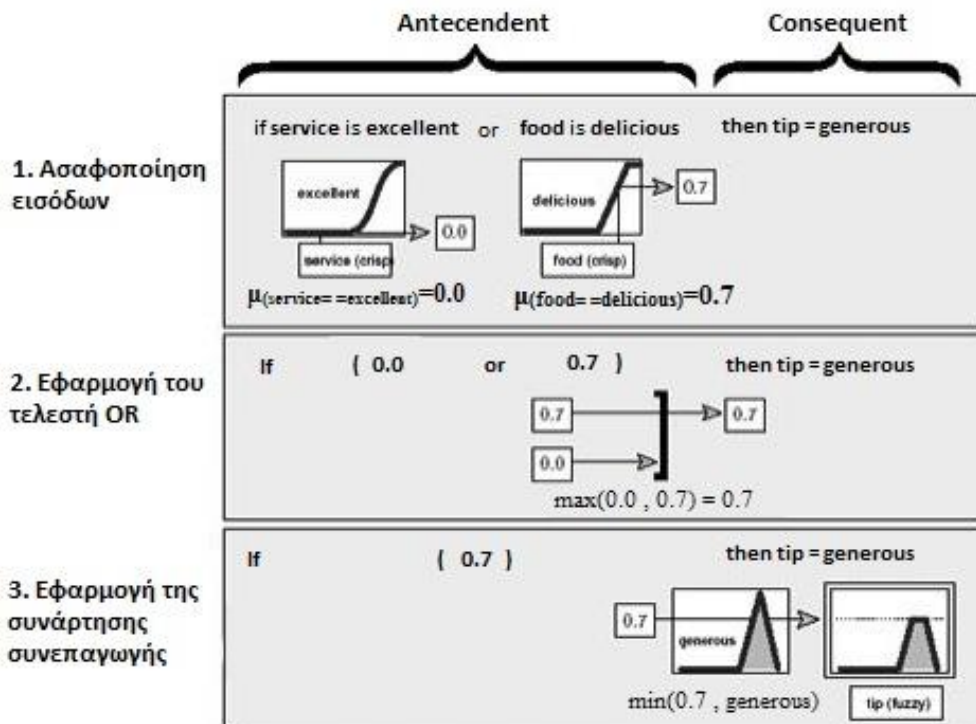
Σ' αυτή την περίπτωση όλα τα μέρη της προϋπόθεσης υπολογίζονται ταυτόχρονα και καταλήγουμε σε ένα μοναδικό νούμερο ανάλογα με τους λογικούς τελεστές που χρησιμοποιούμε.

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

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

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

Μια πολύ συνηθισμένη συνάρτηση για να γίνει αυτό είναι η αποκοπή με τη χρήση του τελεστή $\min$. Στο παρακάτω σχήμα (σχήμα 3.10) βλέπουμε τον τρόπο με τον οποίο εφαρμόζονται οι κανόνες if-then.



Σχήμα 3.10: Κανόνες if-then στην ασαφή λογική

3.2.2 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 σε 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

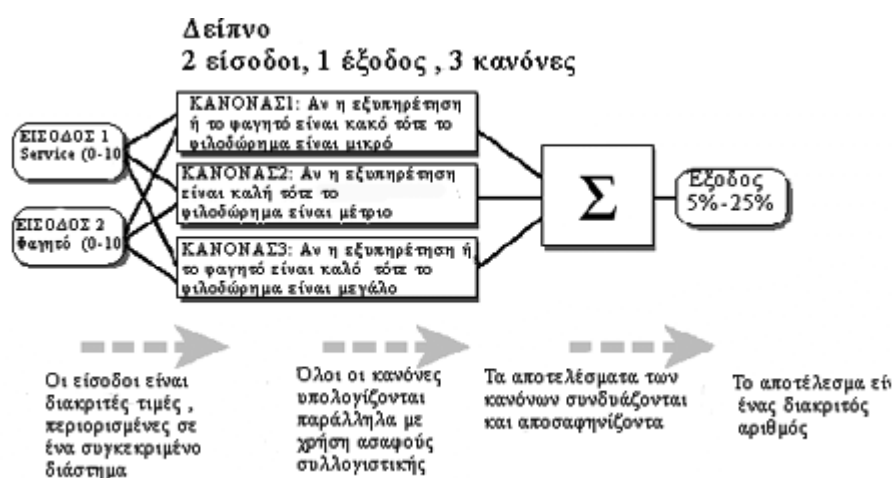
Η ασαφής συλλογιστική είναι η διαδικασία κατά την οποία σχηματίζεται η απεικόνιση από μια δεδομένη είσοδο σε μια έξοδο, χρησιμοποιώντας τις αρχές της ασαφούς λογικής. Η απεικόνιση αυτή θέτει τις βάσεις με τις οποίες μπορεί να ληφθεί μια απόφαση ή να διακριθούν πρότυπα (patterns). Η διαδικασία αυτή της 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 εισόδοι 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 παρακάτω:



Σχήμα 3.11: Σύστημα 2 εισόδων, 3 κανόνων, 1 εξόδου.

Η πληροφορία οδεύει από τα αριστερά προς τα δεξιά και καταλήγει σε μία μοναδική έξοδο. Ένα σημαντικό χαρακτηριστικό των συστημάτων ασαφούς λογικής είναι η παράλληλη εφαρμογή των κανόνων. Στα κλασικά 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

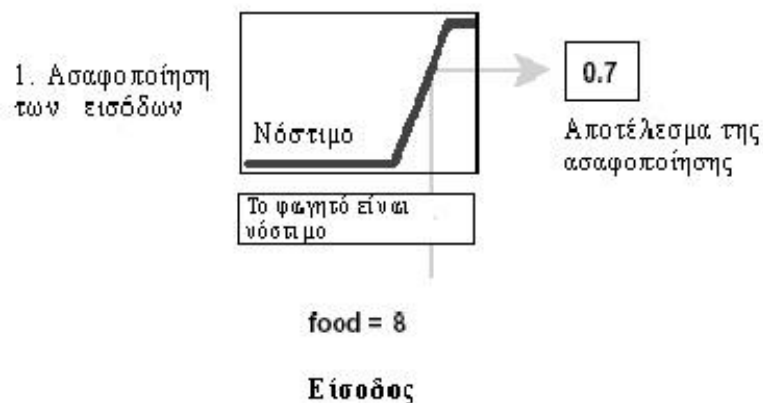
Θα εξετάσουμε κάθε βήμα από τα παραπάνω ξεχωριστά.

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

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

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

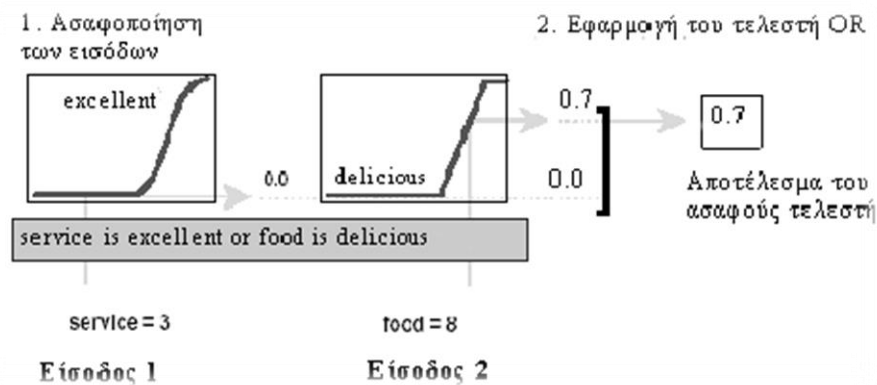


Σχήμα 3.12 Συνάρτηση συμμετοχής ασαφούς συνόλου νόστιμο

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 εισοδοί 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.

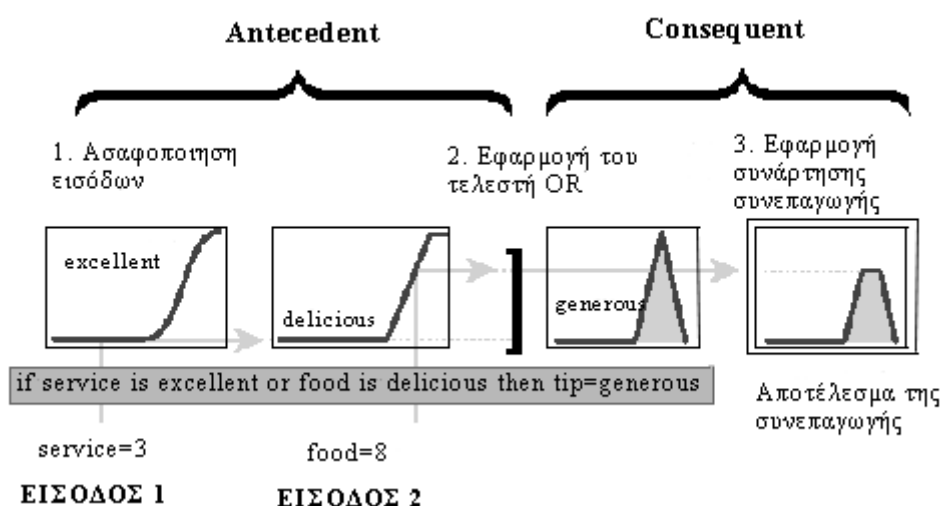


Σχήμα 3.13 Εφαρμογή του ασαφή τελεστή OR

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.

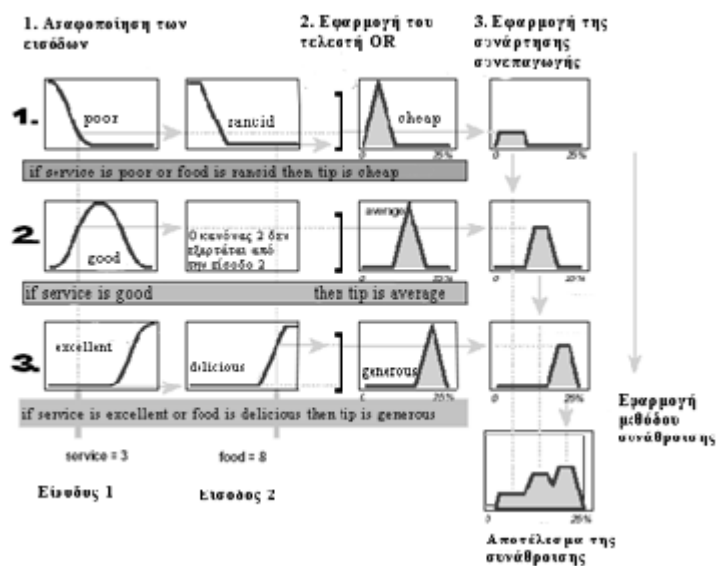


Σχήμα 3.14 Συνάρτηση συνεπαγωγής

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.



Σχήμα 3.15 FIS Mamdani

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 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

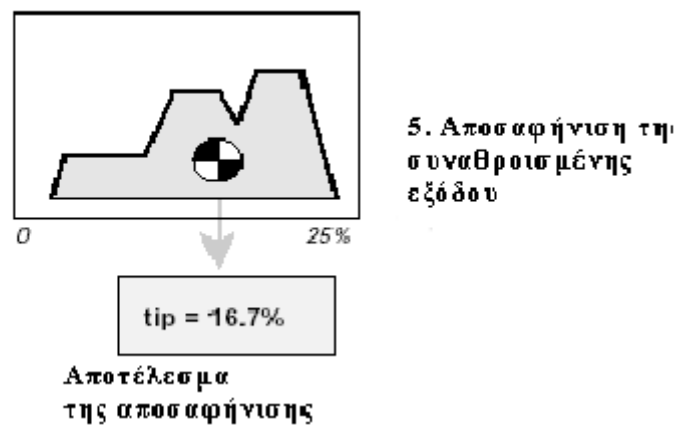
ο μέσος όρος τους (average-of-maximum) ή η μέγιστη τιμή τους (maximum-of-maximum) ή η ελάχιστη τιμή τους (minimum-of-maximum).

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.



Σχήμα 3.16 Αποσαφήνιση Cendroid

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.

Τα πλεονεκτήματα του συστήματος Mamdani είναι ότι έχει μεγαλύτερη αποδοχή από την ερευνητική κοινότητα και είναι πιο διαισθητικά. Επιπλέον ταιριάζουν καλύτερα στο τρόπο που περιγράφουν οι άνθρωποι τα προβλήματα.

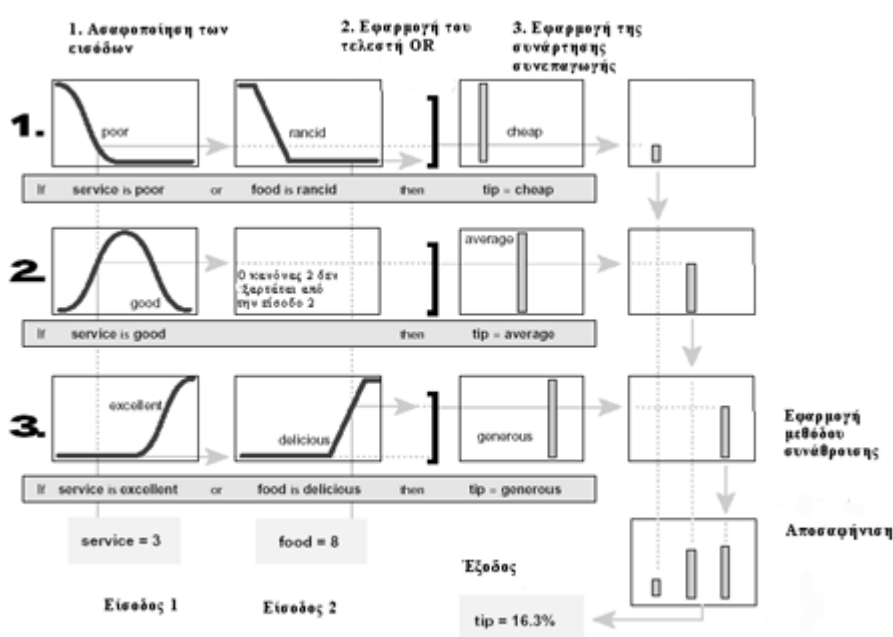
Συστήματα τύπου Sugeno

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.



Σχήμα 3.17 Ασαφές μοντέλο Sugeno

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.

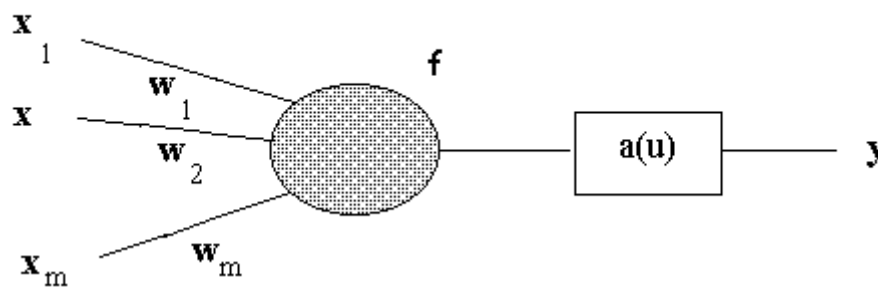
Συστήματα sugeno ανώτερης τάξης είναι εφικτά, όμως δεν προσφέρουν σημαντικές βελτιώσεις και ταυτόχρονα εισαγάγουν σημαντική πολυπλοκότητα.

3.2.3 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:



Σχήμα 3.18 Ασαφείς νευρώνες

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 ενεργοποίησης του νευρώνα
- y η έξοδος του νευρώνα

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

ΝΕΥΡΩΝΕΣ ΣΥΜΜΕΤΟΧΗΣ

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

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

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

ΛΕΙΤΟΥΡΓΙΚΟΙ ΝΕΥΡΩΝΕΣ

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

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

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

ΣΥΝΘΕΤΙΚΟΙ ΝΕΥΡΩΝΕΣ

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

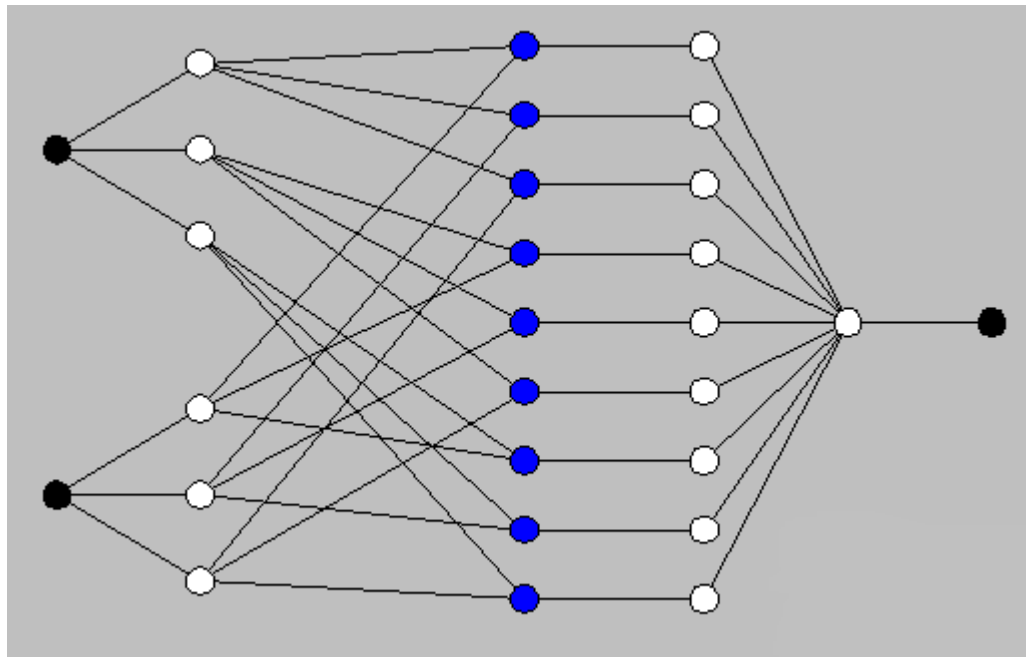
$$f = \text{union}_{t \in N_m} t(x_t, w_t)$$

$$y = a(u)$$

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

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 εισαγωγής φυσικού αερίου. Το δίκτυο αυτό στηρίζεται στην πολυυποθετική προσεγγιστική συλλογιστική και στην εξαγωγή συμπεράσματος που προτείνεται από τους Sugeno-Takagi. Η δομή του anfis φαίνεται στο παρακάτω διάγραμμα:



Σχήμα 3.19 Δομή του ANFIS

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 είσοδοι 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.

3.3 ANFIS (Adaptive Neuro-Fuzzy Inference System)

Ο αλγόριθμος ANFIS είναι από τους κυριότερους αλλά και από τους πρώτους που εφαρμόστηκαν για την επίλυση προβλημάτων στο πεδίο των νευροασαφών συστημάτων. Παρακάτω περιγράφεται το δίκτυο, υποθέτοντας ότι το πρόβλημα του οποίου προσπαθούμε να προσεγγίσουμε τη λύση, έχει δύο εισόδους (x, y) και μια έξοδο (z).

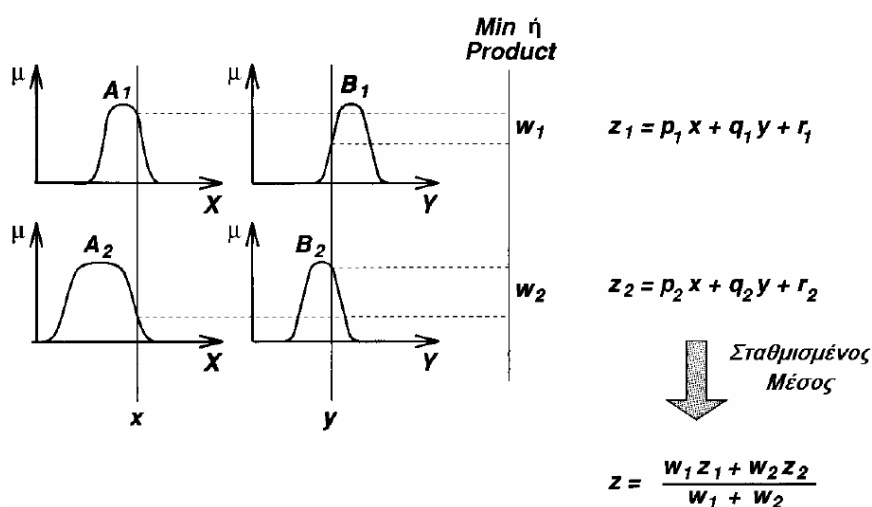
3.3.1 Αρχιτεκτονική του ANFIS

Υποθέτοντας ότι έχουμε ένα πρώτης τάξης μοντέλο sugeno, μία τυπική βάση κανόνων (rulebase) θα μπορούσε να είναι η εξής:

Rule 1: If x is A_1 and y is B_1 then $f_1 = p_1 \times x + q_1 \times y + r_1$

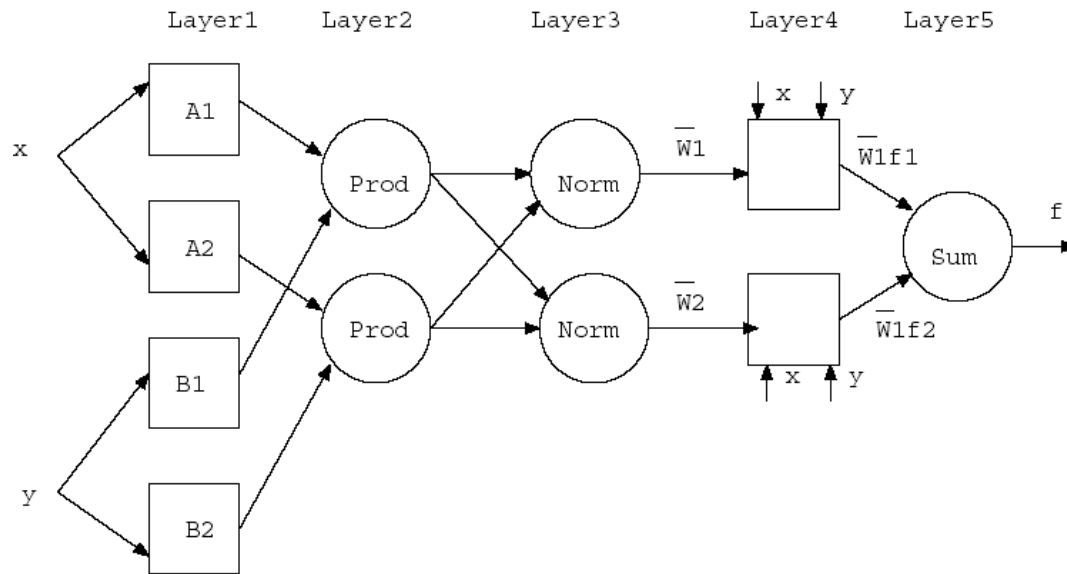
Rule 2: If x is A_2 and y is B_2 then $f_2 = p_2 \times x + q_2 \times y + r_2$

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 3.20 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 3.21 Αρχιτεκτονική του 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 χικανοποιεί 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:

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

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:

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

3.3.2 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:

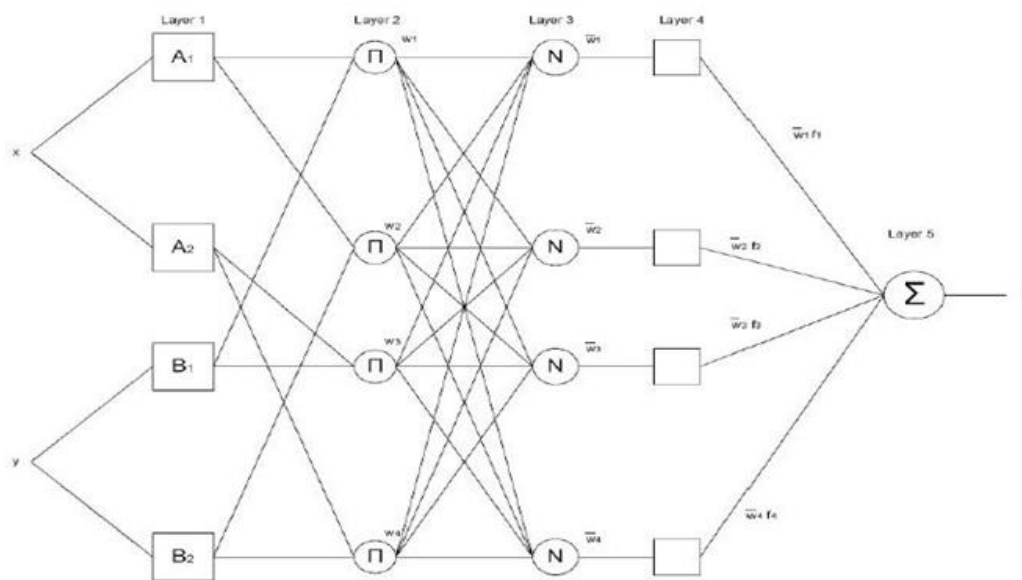
Αν (t) είναι small και (t-1) είναι small τότε (t+1) είναι verysmall

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 verylarge

Λαμβάνοντας υπόψη τα στοιχεία που αναφέρθηκαν παραπάνω φαίνεται στη συνέχεια (σχήμα 3.22) η αρχιτεκτονική του μοντέλου ANFIS. Ύστερα θα δούμε αναλυτικά τι συμβαίνει σε κάθε επίπεδο του μοντέλου και τέλος θα κάνουμε χρήση ενός παραδείγματος για την καλύτερη κατανόησή του. (Εδώ παρουσιάζουμε την έξοδο του i-οστού επιπέδου στο επίπεδο l σαν O_i^l .)



Σχήμα 3.22 Αρχιτεκτονική του μοντέλου ANFIS

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

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

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

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

Με άλλα λόγια, το O_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.

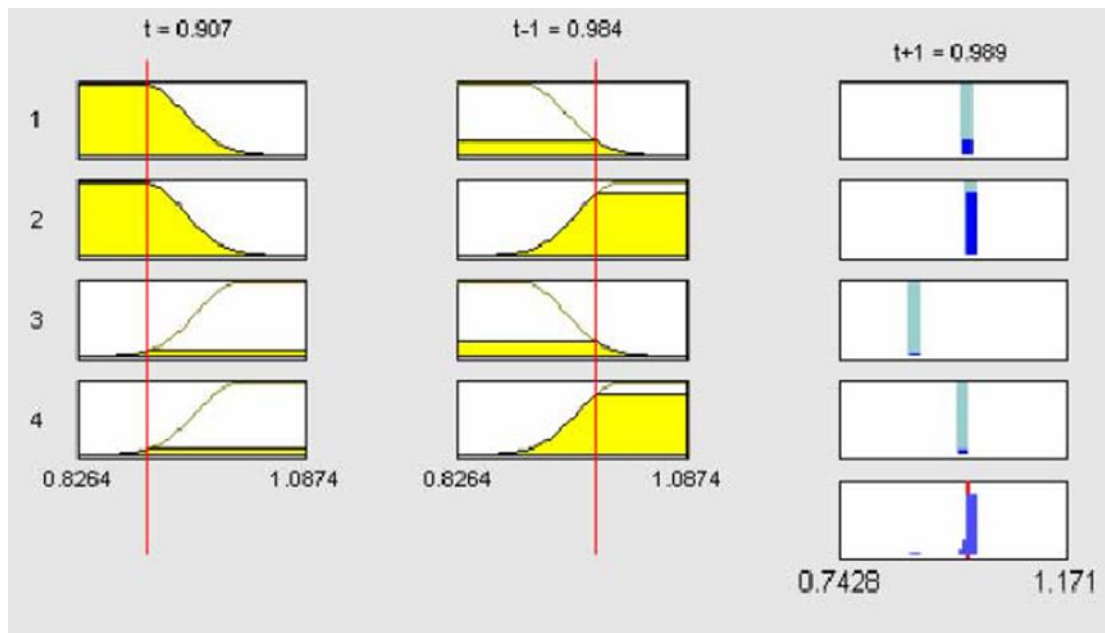
$$O_i^4 = \bar{w}_i * f_i = \bar{w}_i * (p_i * x + q_i * y + r_i)$$

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:

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

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



Σχήμα 3.23 Οι υπολογισμοί του ANFIS

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

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

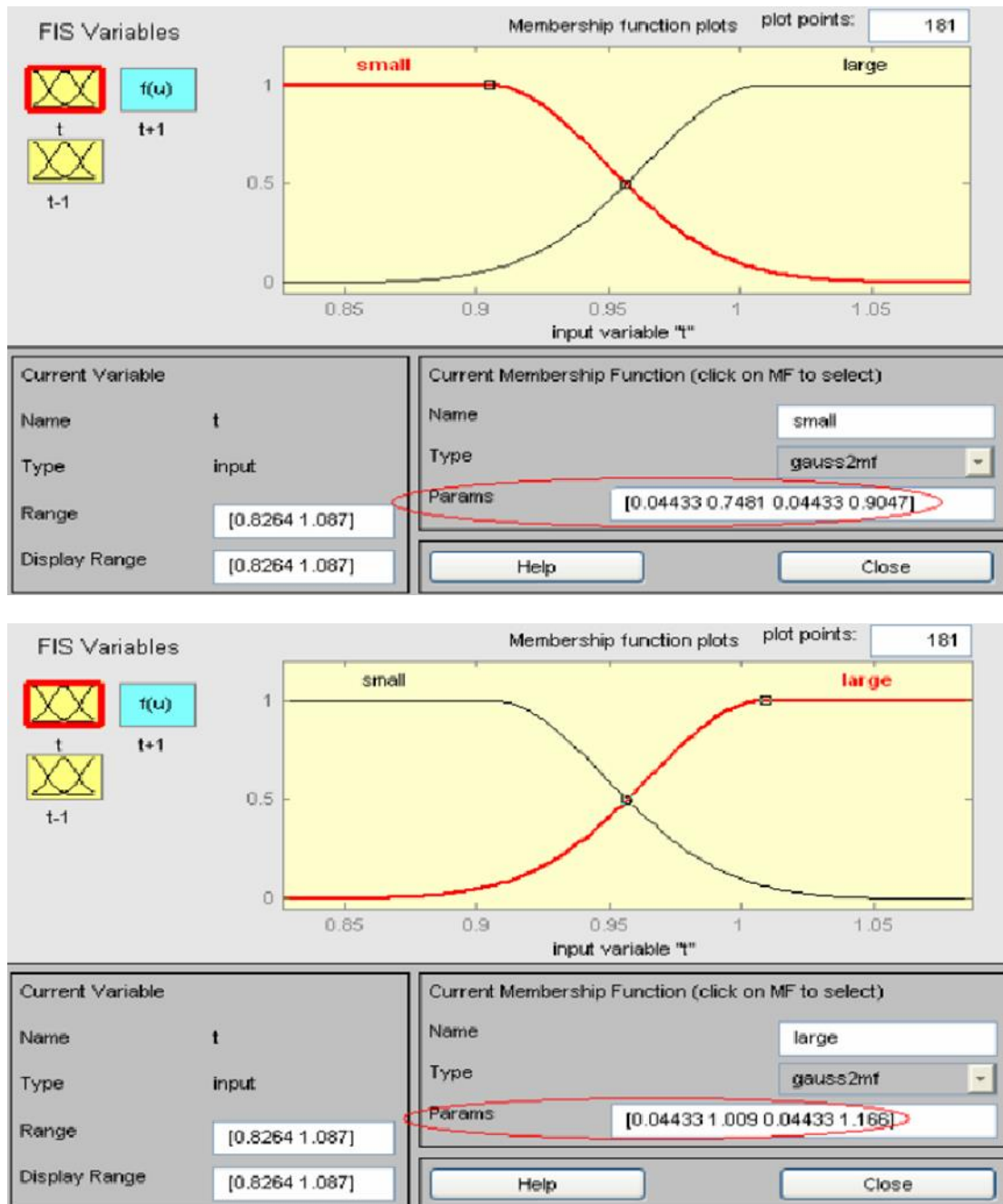
$$[a_1, c_1, a_2, c_2] = [0.04433, 0.7481, 0.04433, 0.9047]$$

αν η γλωσσική μεταβλητή είναι small και

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

αν η γλωσσική μεταβλητή είναι large

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



Σχήμα 3.24 Membership Function Editor

Από το σχήμα βλέπουμε ότι τα A_1 και B_1 είναι **small** και στη **δεξιά** καμπύλη, ενώ τα A_2 και B_2 είναι **large** και στην **αριστερή** καμπύλη. Αυτό σημαίνει ότι για τα A_1 και B_1 θα χρησιμοποιήσουμε το σύνολο παραμέτρων $[0.04433, 0.9047]$ και για τα A_2 και B_2 το σύνολο παραμέτρων $[0.04433, 1.009]$.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$O_4^3 = \bar{w}_3 = \frac{w_3}{w_1 + w_2 + w_3 + w_4} = \frac{0.0604}{0.2016 + 0.8518 + 0.0143 + 0.0604} = 0.0536$$

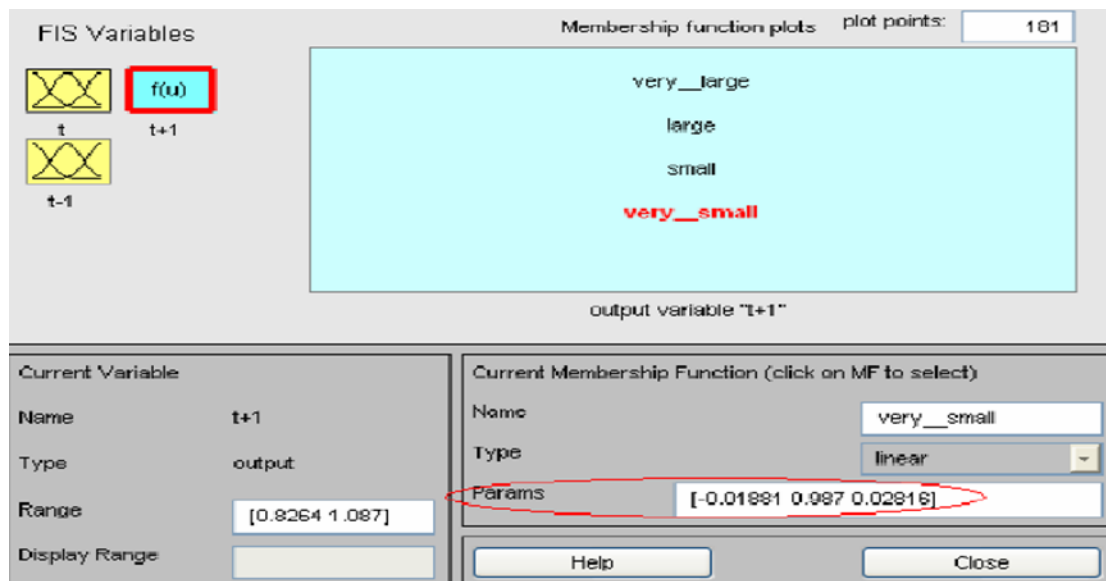
Επίπεδο 4: Για να μπορέσουμε να πραγματοποιήσουμε τους απαραίτητους υπολογισμούς σε αυτό το επίπεδο χρειαζόμαστε το σύνολο παραμέτρων για κάθε κόμβο. Αυτές οι παράμετροι φαίνονται στον membership function editor επιλέγοντας t+1 και είναι:

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

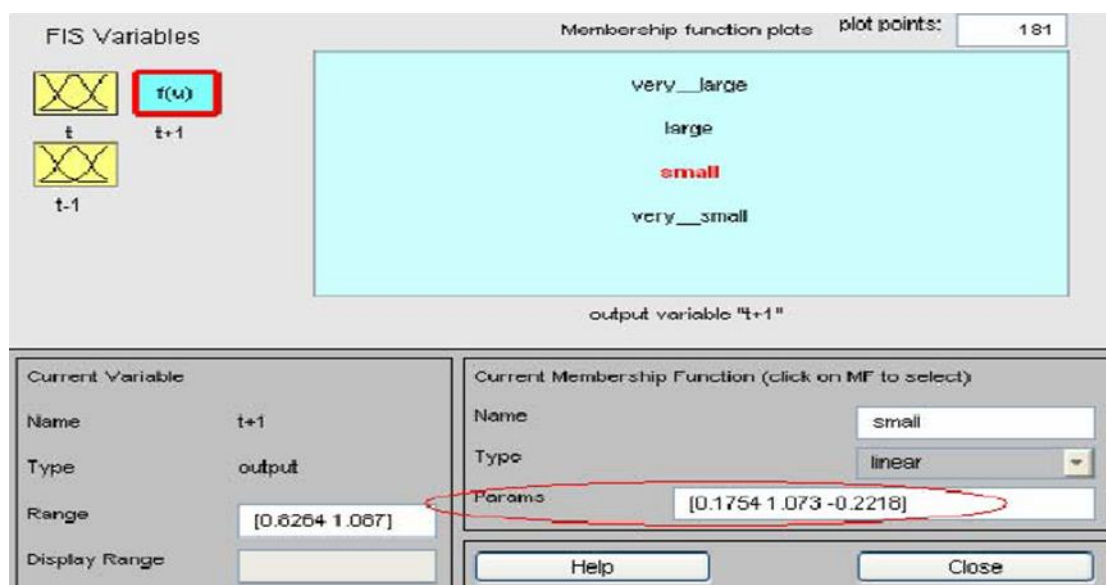
$$[p_2, q_2, r_2] = [0.1754, 1.073, -0.2218] \quad \text{αν η γλωσσική μεταβλητή είναι small}$$

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

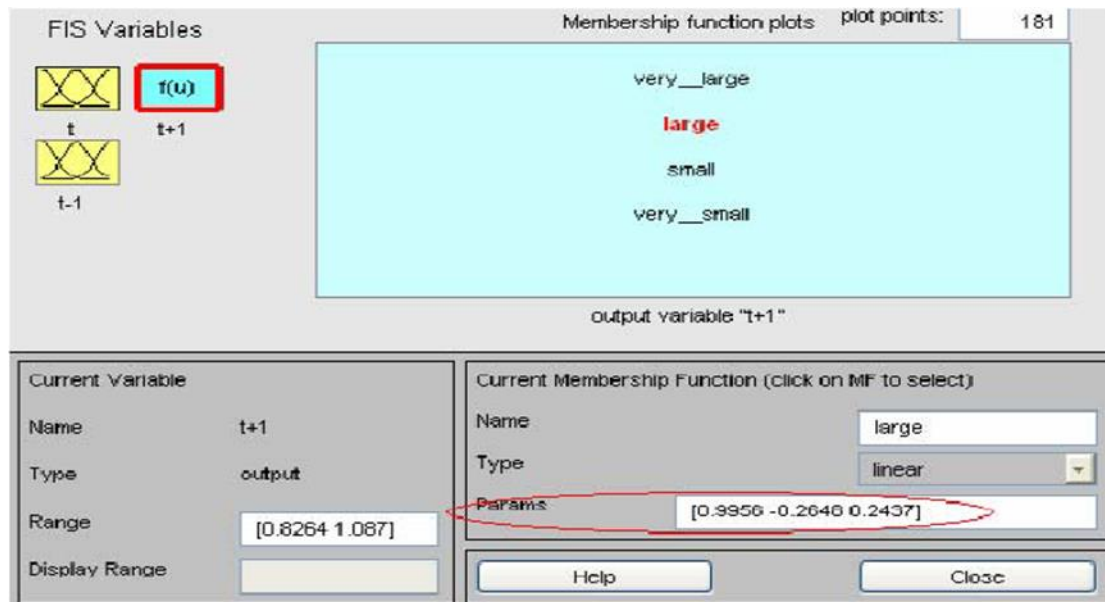
$$[p_4, q_4, r_4] = [0.09453, 0.9042, 0.0006383] \quad \text{αν η γλωσσική μεταβλητή είναι verylarge}$$



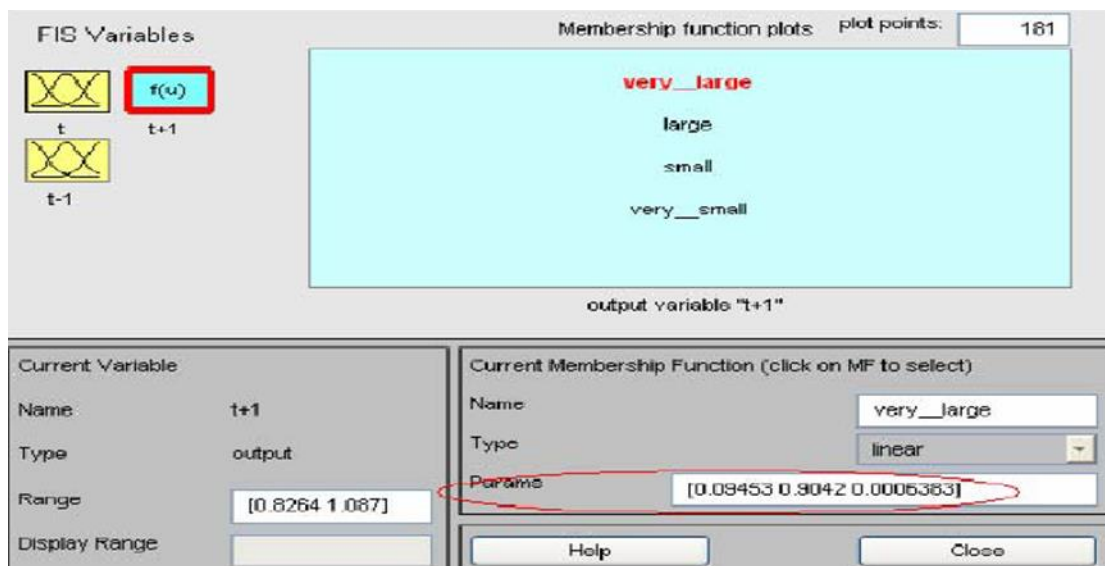
Σχήμα 3.25 Σύνολο παραμέτρων για very small



Σχήμα 3.26 Σύνολο παραμέτρων για small



Σχήμα 3.27 Σύνολο παραμέτρων για large



Σχήμα 3.28 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:

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

3.4 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

3.4.1 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 είναι η i -οστή είσοδος του k νευρώνα, το i -οστό συνοπτικό βάρος του k νευρώνα και η συνάρτηση μεταφοράς (ή "συνάρτηση ενεργοποίησης") του νευρωνικού δικτύου, τότε η έξοδος y_k του k νευρώνα δίνεται από την εξίσωση:

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

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.

3.4.2 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

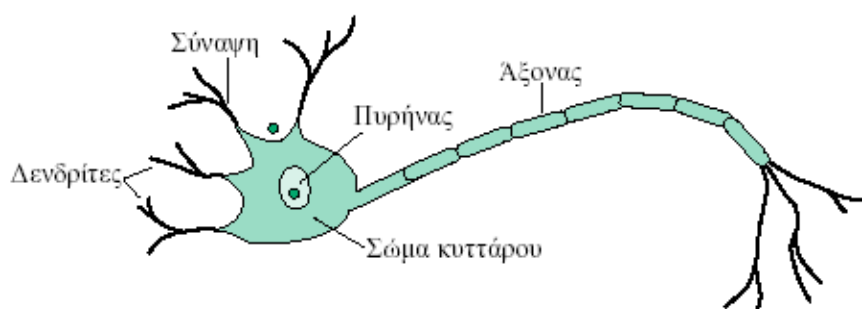
βιβλίου του Hebb με τίτλο “The Organization of Behavior”, στο οποίο έγινε για πρώτη φορά μια ιδιαίτερη δήλωση ενός φυσιολογικού κανόνα μάθησης για συνοπτικές τροποποιήσεις. Πιο συγκεκριμένα ο Hebb πρότεινε ότι η συνδετικότητα του εγκεφάλου συνεχώς αλλάζει καθώς ο οργανισμός μαθαίνει διάφορες εργασίες, 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 κλασικής 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 ενδελεχώς 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).

3.4.3 Βιολογικά νευρωνικά δίκτυα

Λαμβάνοντας υπ' όψιν 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 104 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 στρόγγυλο, τριγωνικό, σταγονοειδές μντερό στα δύο άκρα κλπ ανάλογα με το είδος το νευρώνα: οπτικό, αφής, μυϊκό κοκ.



Σχήμα 3.29 Δομή νευρώνα

Η σημαντικότερη προσπάθεια μοντελοποίησης του βιολογικού κυττάρου έγινε από τους McCulloch και Pitts. Το μοντέλο των McCulloch 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 είσοδοι διαιρούνται σε δύο κατηγορίες: στις εισόδους που διεγείρουν το κύτταρο και στις απαγορευτικές εισόδους. Οι είσοδοι διέγερσης τείνουν να διεγείρουν το κύτταρο ενώ οι απαγορευτικές να απαγορεύσουν την διέγερση του κυττάρου

3.4.4 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 φαίνεται το μαθηματικό μοντέλο του νευρώνα. Όπως φαίνεται στο 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:

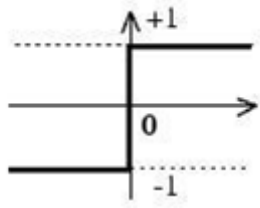
$$y = f(z) = f(w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n + b)$$

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:

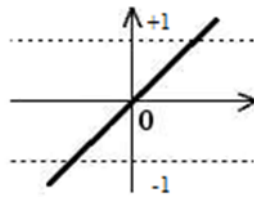
$$\varphi(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



ή οποιαδήποτε άλλη βηματική συνάρτηση.

Η γραμμική συνάρτηση μεταφοράς μπορεί να είναι:

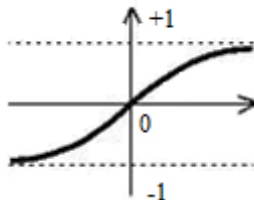
$$\varphi(x) = x$$



ή οποιαδήποτε άλλη γραμμική συνάρτηση.

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

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



3.4.5 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.

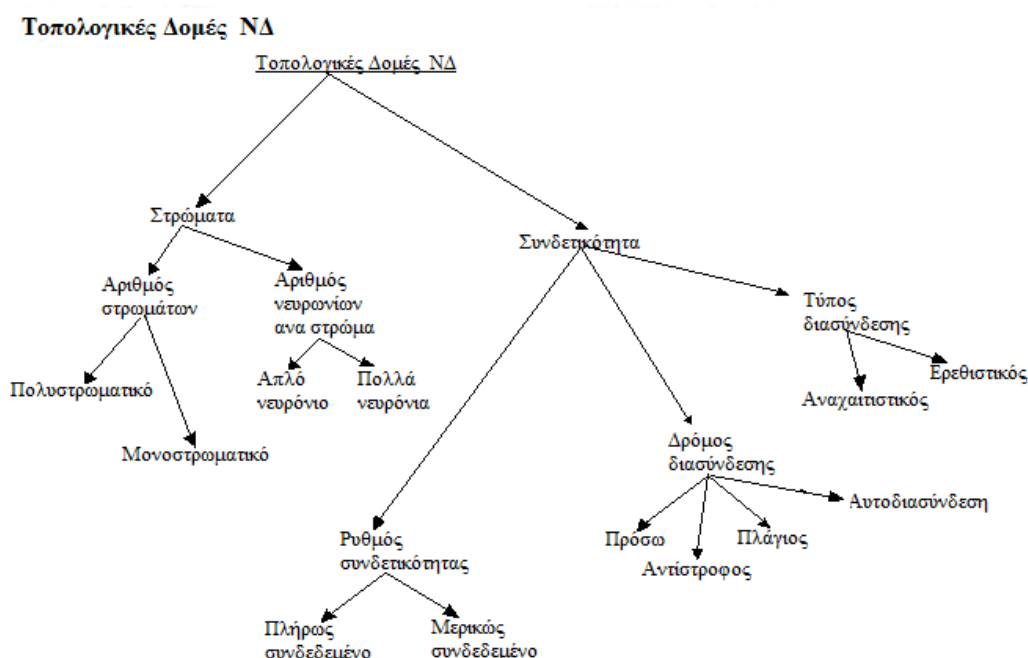
Τηλεπικοινωνίες: Συμπίεση εικόνας και δεδομένων, αυτοματοποιημένες υπηρεσίες πληροφοριών, μετάφραση πραγματικού χρόνου, συστήματα επεξεργασίας πληρωμών.

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.

3.4.6 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 3.30)



Σχήμα 3.30 Τοπολογικές δομές των νευρωνικών δικτύων

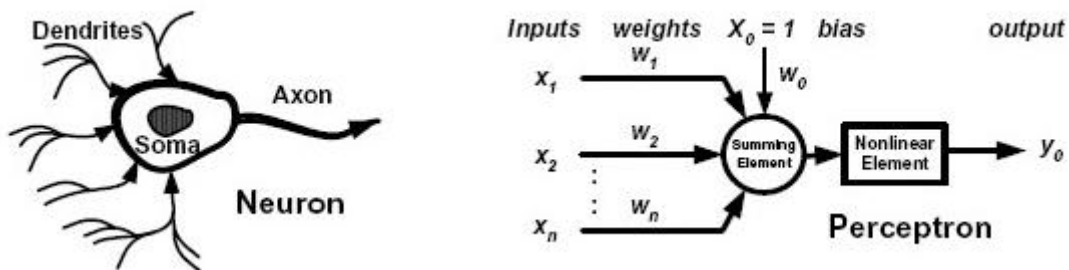
Το μοντέλο του βασικού στοιχείου ενός νευρωνικού δικτύου (βλ. νευρώνα), όπως χρησιμοποιείται ακόμα και σήμερα έχει πρωτοανακαλυφθεί από τους Widrow και Hoff (1960). Αυτοί προσδιόρισαν το perceptron ως το προσαρμοστικό στοιχείο που φέρει ομοιότητες με το νευρώνα. Ένας νευρώνας, που είναι το θεμέλιο ενός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

- το σώμα (cellbody) 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 3.31 Δομή perceptron

Το σήμα εξόδου δίνεται από τη συνάρτηση:

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

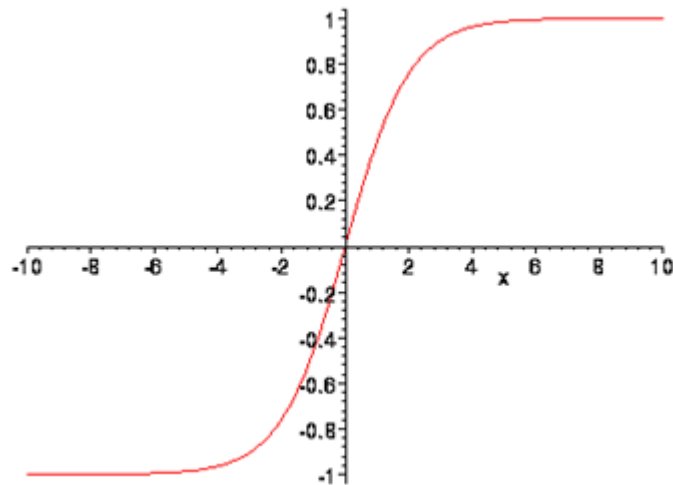
και το κατώφλι ακολουθεί την ανισότητα:

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

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.

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



Σχήμα 3.32 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 (Deltarule) κατά τον οποίο προστίθεται στην προσωρινή τιμή του βάρους $w(k)$ ο συμψηφιστικός όρος $\eta \varepsilon(k)x(k)$, ώστε να δημιουργηθεί η νέα τιμή

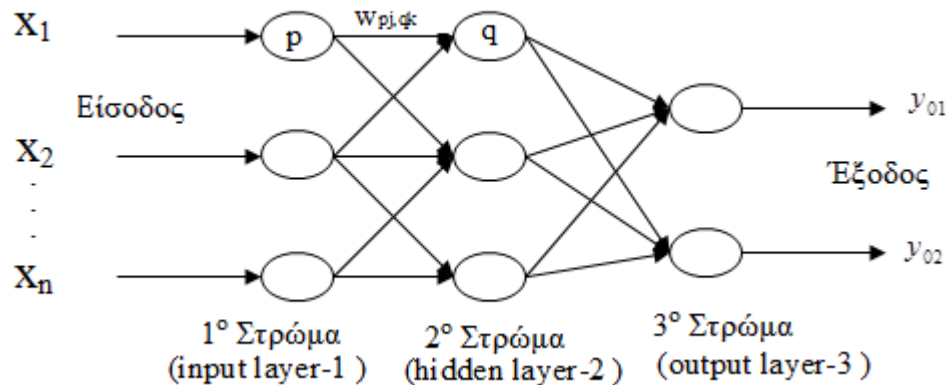
$$w(k+1) = w(k) + \eta \varepsilon(k)x(k) ,$$

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

Αν και αρκετά απλή, αυτή η μέθοδος εκπαίδευσης έχει στην πλειονότητα των περιπτώσεων επιδείξει μεγάλη ακρίβεια, αποτελεσματικότητα και υψηλή ταχύτητα σύγκλισης. Πάραυτα, ένα μοναχό perceptron δεν μπορεί να εκπαιδευτεί τόσο ώστε να είναι ικανό να επιλύσει σύνθετα προβλήματα διότι η ακτίνα υπολογιστικής της 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 (inputlayer), το στρώμα εξόδου (outputlayer) και τα κρυφά στρώματα (hiddenlayers), δημιουργώντας μια δομή με προχωρημένες ικανότητες διασύνδεσης. Στο σχήμα 3.33 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.

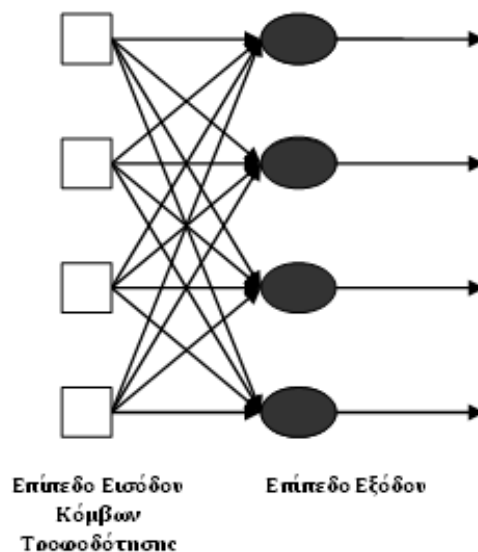


Σχήμα 3.33 Δομή νευρωνικού δικτύου

3.4.7 Είδη αρχιτεκτονικών δομών

1.Ενός Επιπέδου Πρόσθιας Τροφοδότησης Δίκτυα.

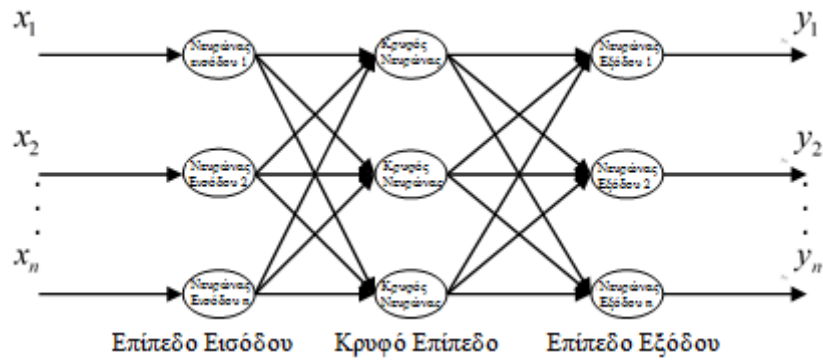
Σε ένα τέτοιο δίκτυο, οι νευρώνες είναι οργανωμένοι σε μορφή επιπέδων. Οι νευρώνες του πηγαίου επιπέδου δείχνουν στους νευρώνες του επόμενου επιπέδου αλλά όχι αντίστροφα, σύμφωνα με το σχήμα 3.34 που ακολουθεί.



Σχήμα 3.34 Πρόσθιας Τροφοδότησης δίκτυο με ένα επίπεδο νευρώνων.

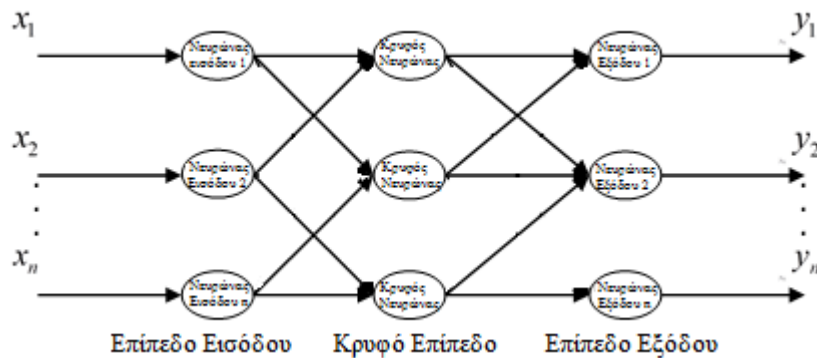
2. Πολλαπλών Επιπέδων Πρόσθιας Τροφοδότησης Δίκτυα.

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



Σχήμα 3.35 Πλήρως συνδεδεμένο δίκτυο πρόσθιας τροφοδότησης

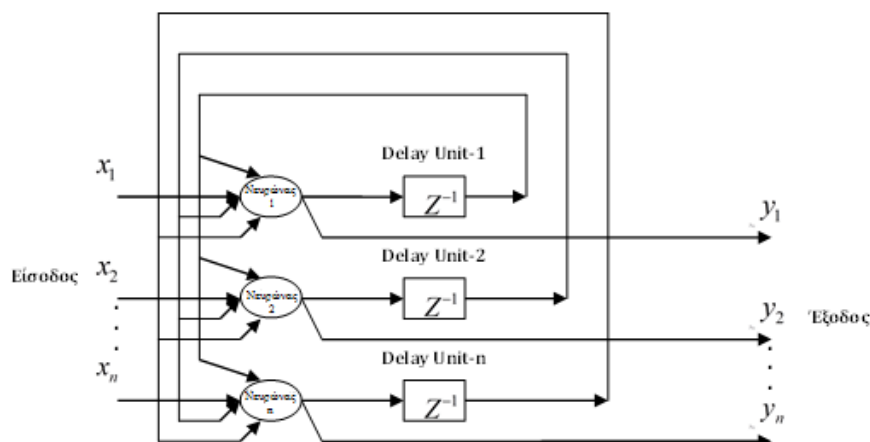
Στο σχήμα που ακολουθεί, παρουσιάζεται ένα μερικώς συνδεδεμένο δίκτυο πρόσθιας τροφοδότησης.



Σχήμα 3.36 Μερικώς συνδεδεμένο δίκτυο πρόσθιας τροφοδότησης

3. Αναδρομικά Δίκτυα

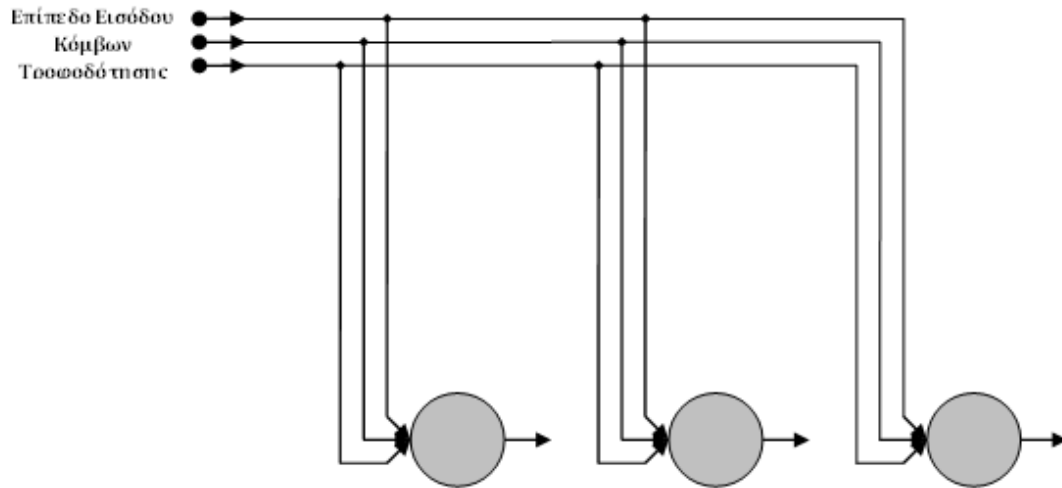
Ένα αναδρομικό νευρωνικό δίκτυο έχει εκτός από τις προς τα εμπρός και προς τα πίσω συνδέσεις. Έχει δηλαδή βρόγχους αναδρομής από τις εξόδους του προς τις εισόδους του. Η παρουσία τέτοιων βρόγχων έχει ισχυρή επίπτωση στην ικανότητα μάθησης του δικτύου. Χαρακτηριστικό παράδειγμα αναδρομικού δικτύου αποτελεί το δίκτυο Hopfield και παρουσιάζεται στο σχήμα 3.37 που ακολουθεί.



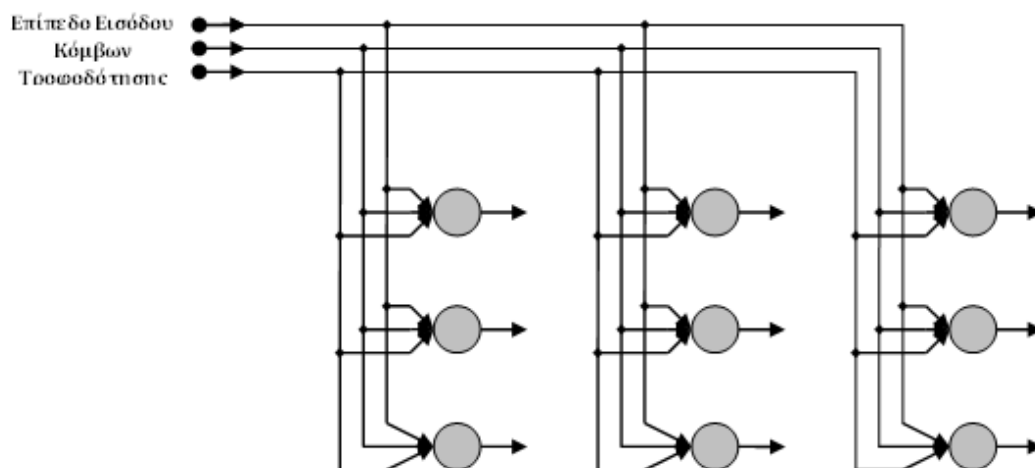
Σχήμα 3.37 Σχηματισμός Δικτύου

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 3.38 και 3.39 που ακολουθούν.



Σχήμα 3.38 Μονοδιάστατο πλέγμα με 3 νευρώνες



Σχήμα 3.39 Δυσδιάστατο πλέγμα με 3 x 3 νευρώνες

3.4.8 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 (trainingset), 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.

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

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 (testset), δηλαδή ένα σύνολο από πρότυπα ελέγχου. Στο στάδιο αυτό δίνονται μόνο τα διανύσματα εισόδου στο νευρωνικό δίκτυο και αυτό υπολογίζει τα προβλεπόμενα διανύσματα εξόδου. Ο υπολογισμός αυτός γίνεται, χρησιμοποιώντας τις τιμές των βαρών (και των πολώσεων, εφόσον υπάρχουν) που υπολογίστηκαν κατά το στάδιο της εκπαίδευσης. Το σφάλμα πρόβλεψης στο σύνολο ελέγχου προκύπτει από το σφάλμα των προβλεπόμενων εξόδων του νευρωνικού δικτύου ως προς τις επιθυμητές εξόδους για κάθε ένα από τα πρότυπα ελέγχου.

Για τη βέλτιστη λειτουργία του νευρωνικού δικτύου, θα πρέπει κατά τα δύο στάδια της λειτουργίας του να ληφθούν υπόψη ορισμένοι βασικοί παράγοντες που το χαρακτηρίζουν.

- Η δομή και ο καθορισμός της αρχιτεκτονικής του νευρωνικού δικτύου. Εξετάζεται ο αριθμός των κρυφών στρωμάτων, ο αριθμός των νευρώνων ανά στρώμα, η συνάρτηση ενεργοποίησης, ο αλγόριθμος εκπαίδευσης του νευρωνικού δικτύου, ο μέγιστος αριθμός επαναλήψεων του αλγορίθμου εκπαίδευσης, κτλ.
- Ο καθορισμός της κατάλληλης δομής των συνόλων εκπαίδευσης και ελέγχου του νευρωνικού δικτύου, για παράδειγμα ο βέλτιστος αριθμός των προτύπων εκπαίδευσης.
- Η ικανότητα γενίκευσης του νευρωνικού δικτύου, η οποία διασφαλίζεται όταν το νευρωνικό δίκτυο παρουσιάζει μικρό σφάλμα πρόβλεψης τόσο στο σύνολο εκπαίδευσης όσο και στο σύνολο ελέγχου.

Η εκπαίδευση του νευρωνικού δικτύου είναι μια σχετικά χρονοβόρα επαναληπτική διαδικασία, ιδίως όταν τα πρότυπα εκπαίδευσης είναι πολλά και όταν επίσης είναι πολλοί οι νευρώνες των στρωμάτων εισόδου και εξόδου.

Ο αριθμός των νευρώνων του στρώματος εισόδου είναι ίσος με τον αριθμό των μεταβλητών εισόδου του προβλήματος πρόβλεψης. Παρόμοια, ο αριθμός των νευρώνων του στρώματος εξόδου είναι ίσος με τον αριθμό των μεταβλητών εξόδου του προβλήματος πρόβλεψης.

Ο αριθμός των νευρώνων του κρυφού στρώματος, πρέπει να προσδιοριστεί με επαναληπτικές δοκιμές, καθώς δεν υπάρχει γενική μέθοδος προσδιορισμού του. Αν οι κρυφοί νευρώνες είναι υπερβολικά λίγοι, το νευρωνικό δίκτυο δε μπορεί να μάθει τις πολύπλοκες σχέσεις μεταξύ εισόδων και εξόδων και ίσως αντιμετωπίσει πρόβλημα σύγκλισης κατά τη διάρκεια της εκπαίδευσης του. Αν ο αριθμός των κρυφών νευρώνων είναι υπερβολικά μεγάλος, η διαδικασία εκπαίδευσης θα διαρκέσει περισσότερο και ίσως επηρεάσει αρνητικά την ικανότητα γενίκευσης του νευρωνικού δικτύου.

Ο αριθμός των νευρώνων του κρυφού στρώματος μεταβάλλεται για διαφορετικές εφαρμογές και συνήθως εξαρτάται από το μέγεθος του συνόλου εκπαίδευσης και τον αριθμό των νευρώνων του στρώματος εισόδου.

3.4.9 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).

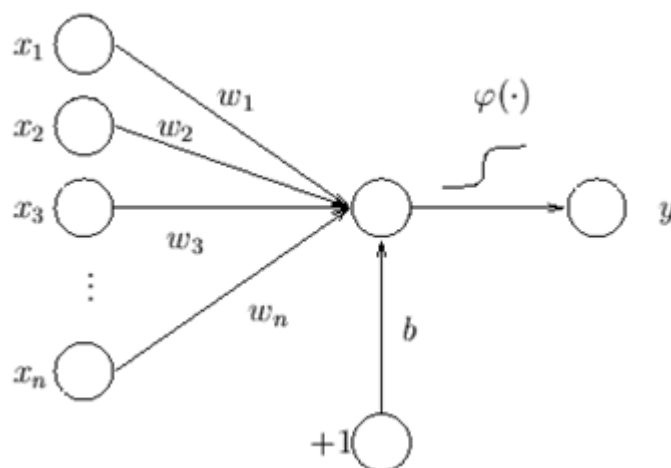
Η αναδρομική διαδικασία εκπαίδευσης (τίθεται $n=n+1$) εκτελείται μέχρι ένα συγκεκριμένο αριθμό επαναλήψεων είτε εναλλακτικά μέχρι να επιτευχθεί ικανοποιητική πρόβλεψη των τιμών εξόδου. Συνήθη κριτήρια τερματισμού του αλγόριθμου εκπαίδευσης περιλαμβάνουν επιτήρηση του αθροίσματος των τετραγώνων των σφαλμάτων,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

3.4.10 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).

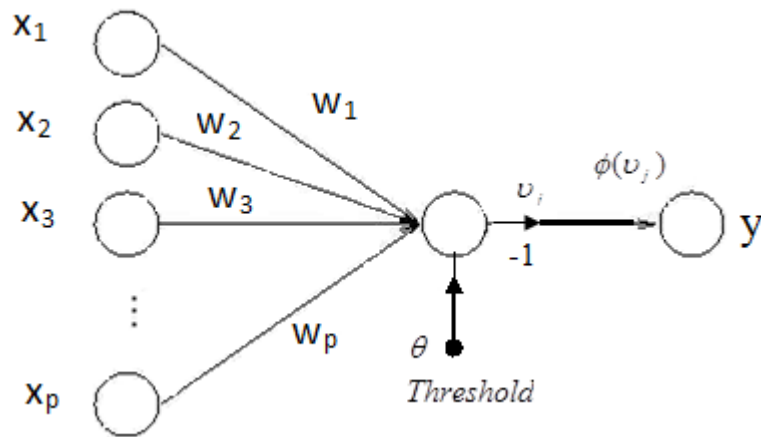
Ένα τέτοιο δίκτυο φαίνεται στο σχήμα που ακολουθεί με έως να είναι η είσοδος, η έξοδος και το κατώφλι (threshold).



Σχήμα 3.40 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 McCulloch – 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 3.41.



Σχήμα 3.41 Το διάγραμμα ροής σήματος του Perceptron

Η έξοδος του γραμμικού συνδυαστή υπολογίζεται εύκολα από το διάγραμμα του σχήματος 3.42 και είναι:

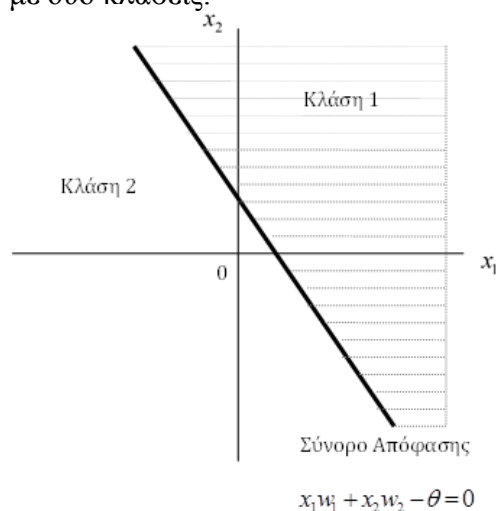
$$v = \sum_{i=1}^p w_i x_i - \theta$$

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 παρακάτω σχήμα φαίνεται η γραμμική διαχωρισιμότητα για ένα δισδιάστατο πρόβλημα ταξινόμησης, με δύο κλάσεις.



Σχήμα 3.42 Το όριο και οι περιοχές απόφασης για ένα δισδιάστατο πρόβλημα ταξινόμησης δύο κλάσεων.

Στο σχήμα 3.42 φαίνεται το αποτέλεσμα της εφαρμογής του κατωφλίου, το οποίο μετατοπίζει το όριο απόφασης από την αρχή των αξόνων. Τα συναπτικά βάρη του Perceptron, μπορούν να προσαρμοσθούν επαναληπτικά. Για την προσαρμογή του διανύσματος βαρών w , χρησιμοποιούμε έναν κανόνα διόρθωσης λάθους, που είναι γνωστός σαν κανόνας σύγκλισης του Perceptron.

Αλγόριθμος Σύγκλισης του Perceptron

Μεταβλητές και Παράμετροι

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

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

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

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

$\theta(n)$ = threshold (κατώφλι)

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

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

η = learning - rate parameter, θετική σταθερά < 1

Βήμα 1: Αρχικοποίηση

Θέσε $w(0) = 0$. 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 η κατά μια μονάδα και πήγαινε στο βήμα 2.

3.4.11 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 (LeastMeanSquare – 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 $n=1,2,\dots$

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

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

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

3.4.12 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 (hiddenlayers)

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 (MultiLayerPerceptrons- 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 BackProp. Στο αλγόριθμο πίσω διάδοσης, η μάθηση επιτυγχάνεται εφαρμόζοντας ένα σύνολο από διανύσματα εκπαίδευσης σαν είσοδο στο πολυεπίπεδο perceptron. Η προβολή όλων των διανυσμάτων εκπαίδευσης στο δίκτυο λέγεται 'κύκλος' (epoch). Η διαδικασία μάθησης προχωράει από epoch σε epoch, μέχρι να σταθεροποιηθούν τα βάρη και τα κατώφλια του δικτύου και το μέσο τετραγωνικό λάθος όλων των διανυσμάτων εκπαίδευσης τείνει σε κάποια ελάχιστη τιμή. Είναι καλή πρακτική να θέτουμε τα διανύσματα εκπαίδευσης σε μια τυχαία σειρά από ένα epoch στο άλλο. Αυτή η τυχαιότητα τείνει να κάνει το ψάξιμο στο χώρο των βαρών σε κάθε κύκλο μάθησης σε στοχαστική διαδικασία, έτσι αποφεύγουμε τον κίνδυνο να γίνουν λιγότεροι κύκλοι από ότι πρέπει. Για ένα δεδομένο σύνολο διάδοσης μπορούμε να ακολουθήσουμε ένα από τους εξής δύο τρόπους:

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

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

διατυπώσουμε ένα λογικό κριτήριο σύγκλισης για την μάθηση πίσω-διάδοσης όπως παρακάτω (Kramer and Sangiovanni-Vincentelli, 1989):

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

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

- *Ο αλγόριθμος πίσω-διάδοσης συγκλίνει όταν ο απόλυτος ρυθμός μεταβολής στο μέσο τετραγωνικό λάθος ανά κύκλο είναι αρκετά μικρός.*

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

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

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

4 ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ (case study)

4.1 ΕΙΣΑΓΩΓΗ

Τα τελευταία χρόνια έχει επιτευχθεί μεγάλη πρόοδος στη βελτιστοποίηση της μεθοδολογίας πρόβλεψης πωλήσεων και σε πολλούς τομείς για τη πρόβλεψη εφαρμόζονται τα νευρωνικά δίκτυα. Όμως όλο και περισσότερο εφαρμόζονται και νευρο-ασαφείς τεχνικές όπως στο μοντέλο αναγνώρισης και πρόγνωσης σε γραμμικά και μη γραμμικά συστήματα. Έτσι εδώ θα παρουσιάσουμε ένα νευρο-ασαφές μοντέλο και ένα νευρωνικό δίκτυο για την πρόβλεψη εισαγωγής φυσικού αερίου στην Ελλάδα και στη Νέα Αγγλία εφαρμόζοντας το μοντέλο Adaptive Neural Fuzzy Inference System (ANFIS) και το μοντέλο NeuralNetworks (NN). Αλλάζοντας τις παραμέτρους του μοντέλου μας, εξετάζεται η συμπεριφορά του και η αποτελεσματικότητά του, σχολιάζονται τα αποτελέσματα των προβλέψεων και τέλος καταγράφονται τα συμπεράσματα.

4.2 ΔΕΔΟΜΕΝΑ

Τα δεδομένα της μελέτης αντλήθηκαν από το site www.desfa.gr/, για την Ελλάδα και από το site <http://www.iso-ne.com> για τη Νέα Αγγλία (ΗΠΑ). Τα δεδομένα που καταγράφηκαν, αφορούν ημερήσια εισαγωγή φυσικού αερίου της χρονικής περιόδου 1/1/2008 έως και 31/3/2013 και η μονάδα μέτρησης είναι μεγαβατώρα (MWh). 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. Υπάρχει παράθεμα με τις μετρήσεις συγκεντρωτικά.

4.3 ΠΑΡΑΜΕΤΡΟΙ ANFIS

Κατά την επεξεργασία των δεδομένων μας ακολουθήσαμε την εξής μεθοδολογία: αρχικά χρησιμοποιήθηκαν τα δεδομένα που αφορούν τις ημερήσιες εισαγωγές στην Ελλάδα και στη Νέα Αγγλία. Οι αρχικές παράμετροι που τέθηκαν στο μοντέλο μας είναι $n=2$ ($mf_n=2$), μέγεθος βήματος 0.1 ($ss=0.1$), εποχές = 500 ($epochs=500$) και συνάρτηση συμμετοχής γενικευμένη καμπανοειδή ($gbellmf$). Αφού εκτελέσαμε το πρόγραμμα με τις παραπάνω ρυθμίσεις και αποθηκεύσαμε τα αποτελέσματα σε ένα φύλλο excel, επαναλάβουμε τη διαδικασία για 100 και 250 εποχές κρατώντας ως συνάρτηση συμμετοχής την $gbellmf$. Στη συνέχεια τρέξαμε τα δεδομένα μας πάλι για 500, 100, 250 εποχές χρησιμοποιώντας όμως όλες τις διαθέσιμες συναρτήσεις συμμετοχής $gaussmf$ και $gauss2mf$ (γκουσιανή συνάρτηση συμμετοχής), $trimf$ (συνάρτηση συμμετοχής τριγωνικής μορφής), smf (σιγμοειδής συνάρτηση συμμετοχής), $trapmf$ (τραπεζοειδής συνάρτηση συμμετοχής), $primf$ (συνάρτηση

συμμετοχής σχήματος Π) και zmf (συνάρτηση συμμετοχής σχήματος Z). Όλη η παραπάνω διαδικασία επαναλήφθηκε και για τα υπόλοιπα σετ δεδομένων και καταγράφηκαν αναλυτικά σε πίνακα που θα παρουσιαστεί στο κεφάλαιο των αποτελεσμάτων.

Στα πλαίσια της βαθύτερης κατανόησης του τρόπου λειτουργίας του ANFIS, θα παρουσιάσουμε τα κυριότερα γραφήματα που προκύπτουν από το πρόγραμμα της Matlab. Ακολουθεί πίνακας με χρήσιμες πληροφορίες σχετικά με το ANFIS:

Name (Όνομα)	ANFIS
Type (Τύπος ασαφούς μηχανισμού συμπερασμού)	Sugeno
NumInputs (Αριθμός εισόδων)	1
NumOutputs (Αριθμός εξόδων)	1
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)	384
Number of nodes (Αριθμός κόμβων)	12
Number of fuzzy rules (Αριθμός ασαφών κανόνων)	2

4.4 ΑΠΟΤΕΛΕΣΜΑΤΑ

Ακολουθεί η εξαγωγή τεσσάρων τύπων σφαλμάτων ώστε να γίνει η σύγκριση μεταξύ των μοντέλων AR, ARMA και ANFIS.

Τα σφάλματα αυτά είναι τα:

- Το μέσο τετραγωνικό σφάλμα (MSE)

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

- Η ρίζα του μέσου τετραγωνικού σφάλματος (RMSE)

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

- Το μέσο απόλυτο σφάλμα (MAE)

$$MAE = \frac{1}{N} \sum_{t=1}^N |e_t|$$

- Το μέσο απόλυτο ποσοστιαίο σφάλμα (MAPE)

$$MAPE = \frac{100}{N} \sum_{t=1}^N \left| \frac{F_t - A_t}{A_t} \right|$$

όπου A_t και F_t είναι πραγματικές τιμές.

4.4.1 Ελλάδα

epochs=500		MSE	RMSE	MAE	MAPE
gbell	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7455e+08	2.3970e+04	1.8557e+04	8.3737
tri	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7452e+08	2.3963e+04	1.8610e+04	8.3877
Gauss ₂	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7357e+08	2.3949e+04	1.8551e+04	8.3632
gauss	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7268e+08	2.3931e+04	1.8532e+04	8.3467
trap	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7417e+08	2.3962e+04	1.8554e+04	8.3621
Pi	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419

	ANFIS	5.7418e+08	2.3962e+04	1.8557e+04	8.3786
Nn	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	NN	5.8080e+08	2.4100e+04	1.8721e+04	8.4057
smf	Unsupported				
zmf	Unsupported				

Πίνακας 1 Σφάλματα στα 3 μοντέλα για όλες τις συναρτήσεις για 500 εποχές

Ο πίνακας 1 παρουσιάζει τις τιμές και των τεσσάρων σφαλμάτων (MSE, RMSE, MAE, MAPE) για κάθε μοντέλο (AR, ARMA, ANFIS) σε κάθε συνάρτηση συμμετοχής (gbell, tri, gauss, gauss2, trap, pi, smf, zmf) για 500 εποχές. Συγκρίνοντας το RMSE (Root Mean Square Error) αφού είναι ο σημαντικότερος δείκτης για να μας δείξει πόσο αποτελεσματικό είναι το μοντέλο πρόβλεψης βλέπουμε πως τα καλύτερα αποτελέσματα τα παρέχει η συνάρτηση GAUSS σε συνδυασμό με το μοντέλο πρόβλεψης ANFIS.

epochs=100		MSE	RMSE	MAE	MAPE
gbell	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7317e+08	2.3941e+04	1.8552e+04	8.3627
tri	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7452e+08	2.3969e+04	1.8610e+04	8.3877
Gauss ₂	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7358e+08	2.3949e+04	1.8550e+04	8.3617
gauss	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7268e+08	2.3931e+04	1.8532e+04	8.3467
trap	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7465e+08	2.3972e+04	1.8581e+04	8.3795
pi	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7363e+08	2.3951e+04	1.8551e+04	8.3626
smf	Unsupported				
zmf	Unsupported				

Πίνακας 2 Σφάλματα στα 3 μοντέλα για όλες τις συναρτήσεις για 100 εποχές

Στο πίνακα 2 φαίνονται οι τιμές των τεσσάρων σφαλμάτων (MSE, RMSE, MAE, MAPE) για κάθε μοντέλο (AR, ARMA, ANFIS) σε κάθε συνάρτηση συμμετοχής (gbell, tri, gauss, gauss2, trap, pi, smf, zmf) για τις 100 εποχές. Καλύτερα αποτέλεσμα, για να χαρακτηριστεί το μοντέλο πρόβλεψης αξιόπιστο, είναι εκείνο που έχει χαμηλότερη τιμή του σφάλματος RMSE. Αυτό το παρέχει η συνάρτηση GAUSS σε συνδυασμό με το μοντέλο πρόβλεψης ANFIS.

epochs=250		MSE	RMSE	MAE	MAPE
gbell	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7435e+08	2.3966e+04	1.8553e+04	8.3699
tri	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7452e+08	2.3969e+04	1.8610e+04	8.3877
Gauss 2	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7363e+08	2.3951e+04	1.8553e+04	8.3641
gauss	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7268e+08	2.3931e+04	1.8532e+04	8.3467
trap	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7417e+08	2.3962e+04	1.8554e+04	8.3621
pi	AR	1.7076e+09	4.1324e+04	2.9495e+04	12.7988
	ARMA	1.3318e+09	3.6494e+04	2.8535e+04	12.5419
	ANFIS	5.7418e+08	2.3962e+04	1.8577e+04	8.3786
smf	Unsupported				
zmf	Unsupported				

Πίνακας 3 Σφάλματα στα 3 μοντέλα για όλες τις συναρτήσεις για 250 εποχές

Στο πίνακα 3 παρουσιάζονται τις τιμές και των τεσσάρων σφαλμάτων (MSE, RMSE, MAE, MAPE) για κάθε μοντέλο (AR, ARMA, ANFIS) σε κάθε συνάρτηση συμμετοχής (gbell, tri, gauss, gauss2, trap, pi, smf, zmf) για 250 εποχές. Τα καλύτερα αποτελέσματα, δηλαδή τη μικρότερη τιμή RMSE σφάλματος τα παρέχει η συνάρτηση GAUSS σε συνδυασμό με το μοντέλο πρόβλεψης ANFIS.

4.4.2 Νέα Αγγλία

epochs=500		MSE	RMSE	MAE	MAPE
gbell	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.5355e+08	1.5923e+04	1.2381e+04	9.8903
tri	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4680e+08	1.5710e+04	1.2159e+04	9.7757
Gauss ₂	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4702e+08	1.5717e+04	1.2241e+04	9.8055
gauss	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4767e+08	1.5738e+04	1.2256e+04	9.8151
trap	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4935e+08	1.5791e+04	1.2324e+04	9.8445
pi	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4771e+08	1.5739e+04	1.2255e+04	9.8101
nn	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	NN	2.6784e+08	1.6366e+04	1.2752e+04	10.0358
smf	Unsupported				
zmf	Unsupported				

Πίνακας 4 Σφάλματα στα 3 μοντέλα για όλες τις συναρτήσεις για 500 εποχές

Ο πίνακας 4 παρουσιάζει τις τιμές και των τεσσάρων σφαλμάτων (MSE, RMSE, MAE, MAPE) για κάθε μοντέλο (AR, ARMA, ANFIS) σε κάθε συνάρτηση συμμετοχής (gbell, tri, gauss, gauss2, trap, pi, smf, zmf) για 500 εποχές. Για τις συναρτήσεις συμμετοχής smf και zmf το πρόγραμμα δεν υποστηρίζεται ενώ συγκρίνοντας τα υπόλοιπα αποτελέσματα, τα καλύτερα τα παρέχει η συνάρτηση TRI σε συνδυασμό με το μοντέλο πρόβλεψης ANFIS.

epochs=100		MSE	RMSE	MAE	MAPE
gbell	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.5352e+08	1.5922e+04	1.2389e+04	9.8933

tri	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4689e+08	1.5713e+04	1.2159e+04	9.7762
Gauss ₂	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4701e+08	1.5717e+04	1.2241e+04	9.8054
gauss	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4767e+08	1.5737e+04	1.2256e+04	9.8151
trap	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4936e+08	1.5791e+04	1.2324e+04	9.8444
pi	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4771e+08	1.5739e+04	1.2255e+04	9.8101
smf	Unsupported				
zmf	Unsupported				

Πίνακας 5 Σφάλματα στα 3 μοντέλα για όλες τις συναρτήσεις για 100 εποχές

Στο πίνακα 5 εμφανίζονται τα σφάλματα MSE, RMSE, MAE και MAPE για κάθε μοντέλο (AR, ARMA, ANFIS) σε κάθε συνάρτηση συμμετοχής (gbell, tri, gauss, gauss2, trap, pi, smf, zmf) για 100 εποχές. Τα καλύτερα αποτελέσματα τα παρέχει η συνάρτηση TRI σε συνδυασμό με το μοντέλο πρόβλεψης ANFIS, στο οποίο εμφανίζονται τα μικρότερα σφάλματα.

epochs=250		MSE	RMSE	MAE	MAPE
gbell	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.5369e+08	1.5928e+04	1.2386e+04	9.8925
tri	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4684e+08	1.5711e+04	1.2159e+04	9.7758
Gauss ₂	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4701e+08	1.5717e+04	1.2241e+04	9.8054
gauss	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4767e+08	1.5738e+04	1.2256e+04	9.8152

trap	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4936e+08	1.5791e+04	1.2324e+04	9.8444
pi	AR	5.8281e+08	2.4141e+04	1.8942e+04	15.4453
	ARMA	5.6467e+08	2.3763e+04	1.9016e+04	15.4782
	ANFIS	2.4771e+08	1.5739e+04	1.2255e+04	9.8101
smf	Unsupported				
zmf	Unsupported				

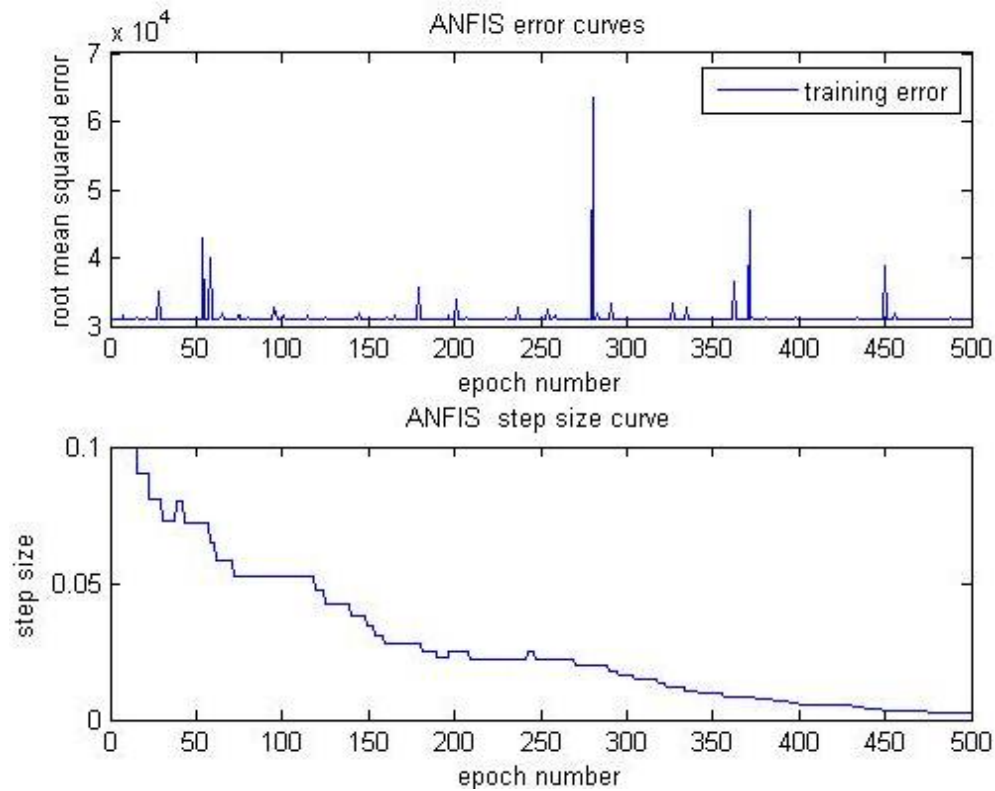
Πίνακας 6 Σφάλματα στα 3 μοντέλα για όλες τις συναρτήσεις για 250 εποχές

Ο πίνακας 6 παρουσιάζει τις τιμές και τεσσάρων σφαλμάτων (MSE, RMSE, MAE, MAPE) που έχουμε για κάθε μοντέλο (AR, ARMA, ANFIS) σε κάθε συνάρτηση συμμετοχής (gbell, tri, gauss, gauss2, trap, pi, smf, zmf) για 250 εποχές. Τα καλύτερα αποτελέσματα τα παρέχει η συνάρτηση TRI σε συνδυασμό με το μοντέλο πρόβλεψης ANFIS μιας και εκεί εμφανίζεται η μικρότερη τιμή του σφάλματος RMSE, που είναι ο σημαντικότερος δείκτης για να χαρακτηριστεί αξιόπιστη η πρόβλεψη.

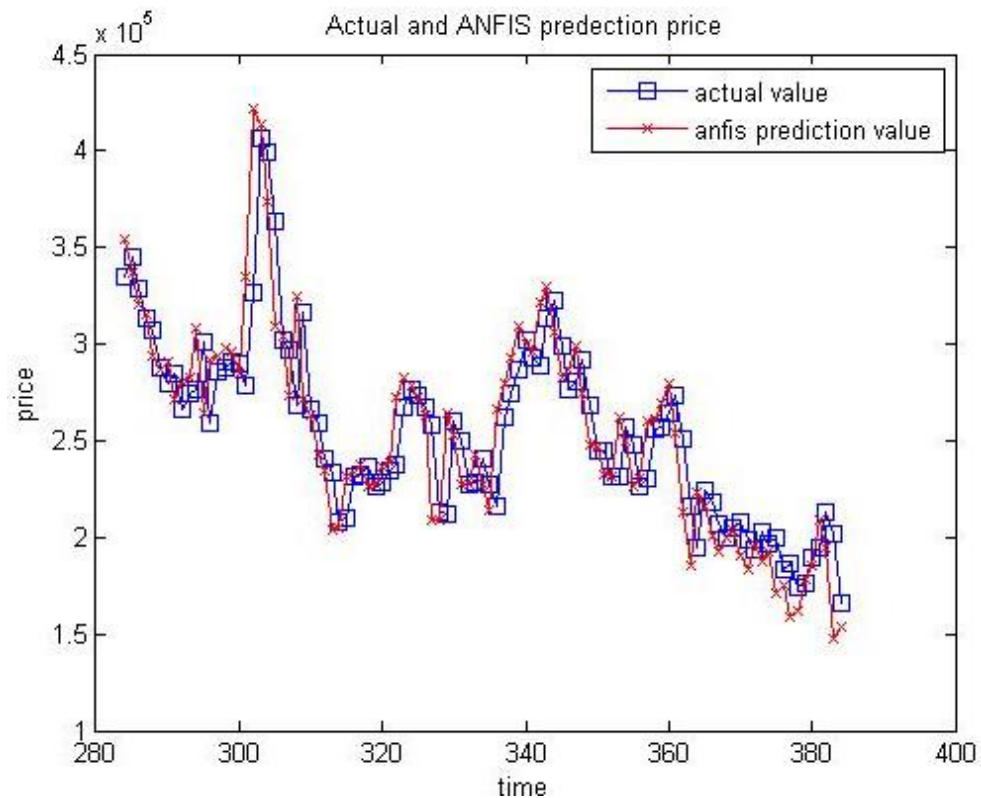
4.5 ΓΡΑΦΗΜΑΤΑ

4.5.1 Ελλάδα

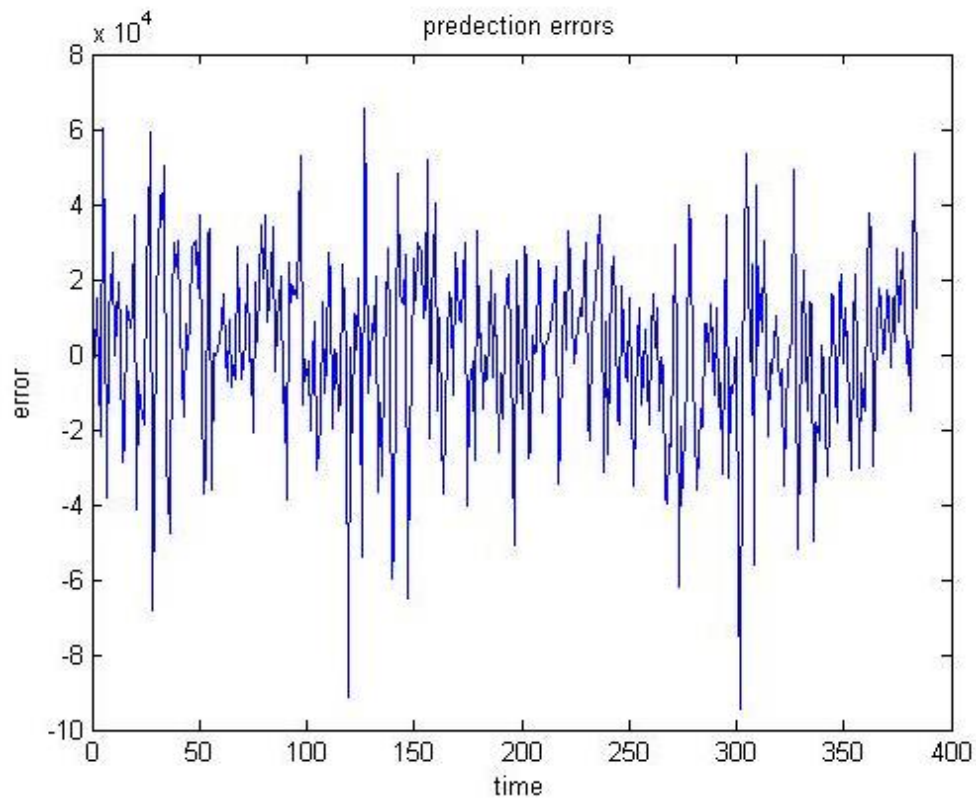
Παρακάτω απεικονίζονται ενδεικτικά κάποια γραφήματα από τα τελικά αποτελέσματα για τις διάφορες δοκιμές που πραγματοποιήσαμε. Συγκεκριμένα, αφορούν την ανάλυση με τη gbell συνάρτηση και αριθμό εποχών 500.



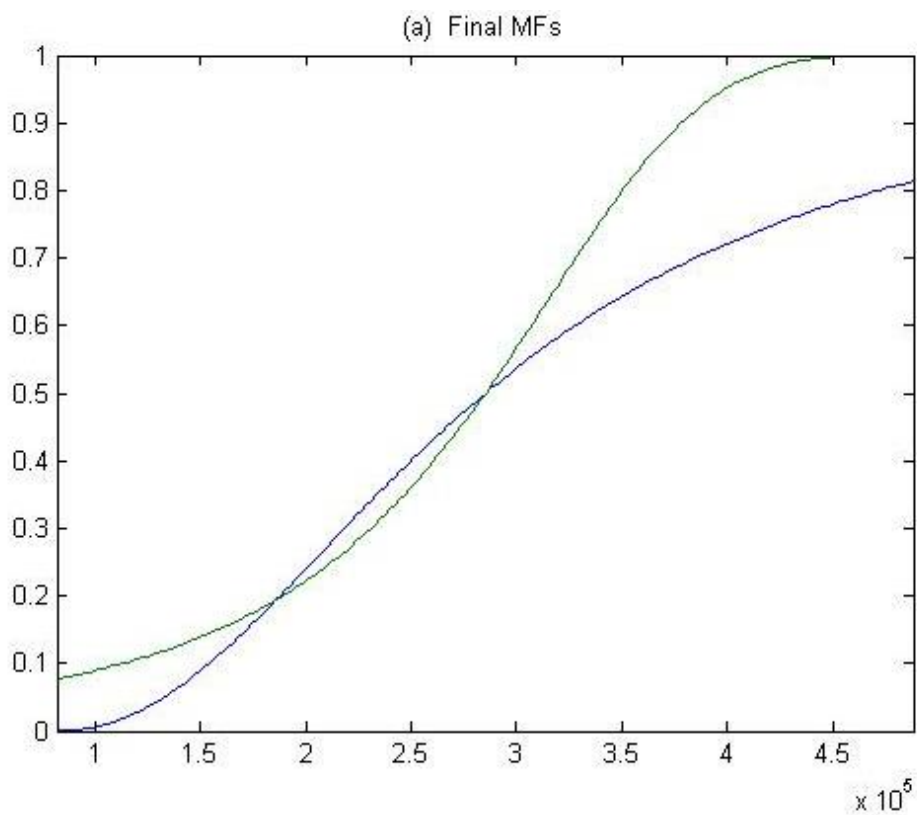
Γράφημα 1 Καμπύλες σφαλμάτων και καμπύλες μεγέθους βήματος της δοκιμής



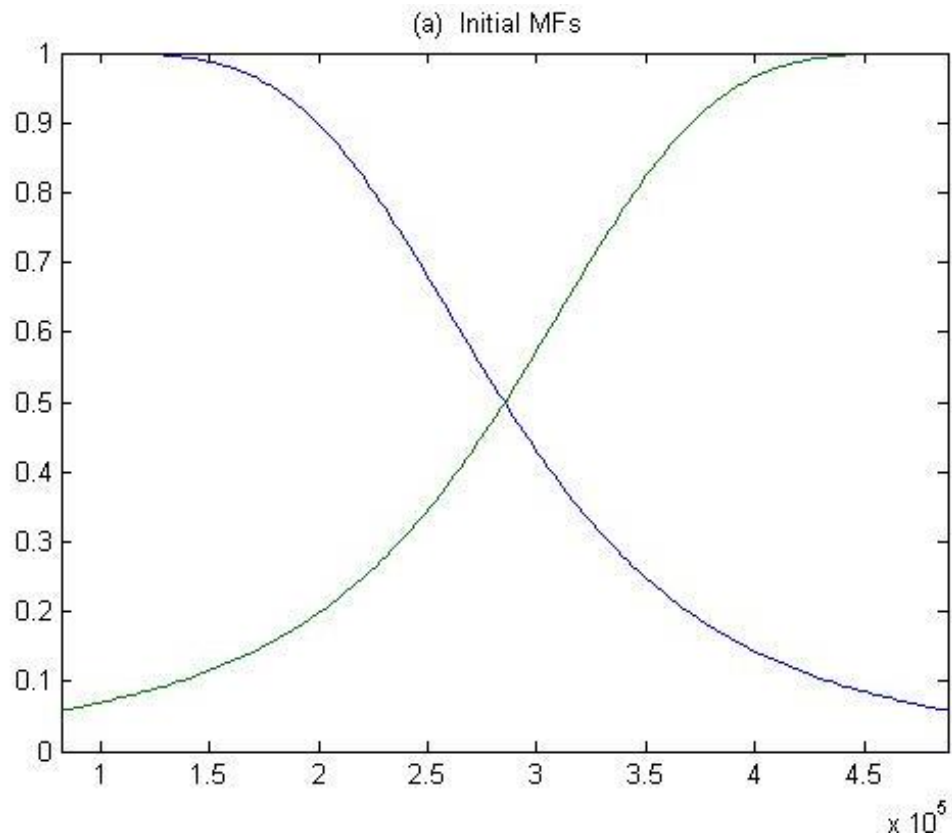
Γράφημα 2 Σύγκριση προβλέψεων του ANFIS έναντι των επιθυμητών τιμών της δοκιμής



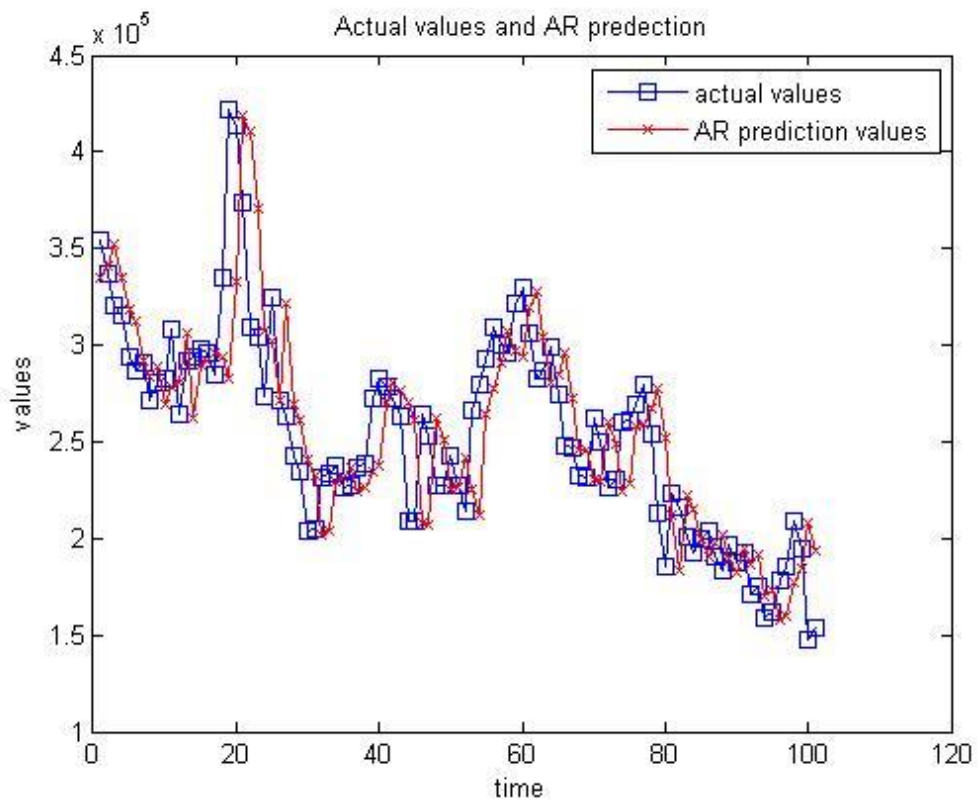
Γράφημα 3 Σφάλματα πρόβλεψης του ANFIS της δοκιμής



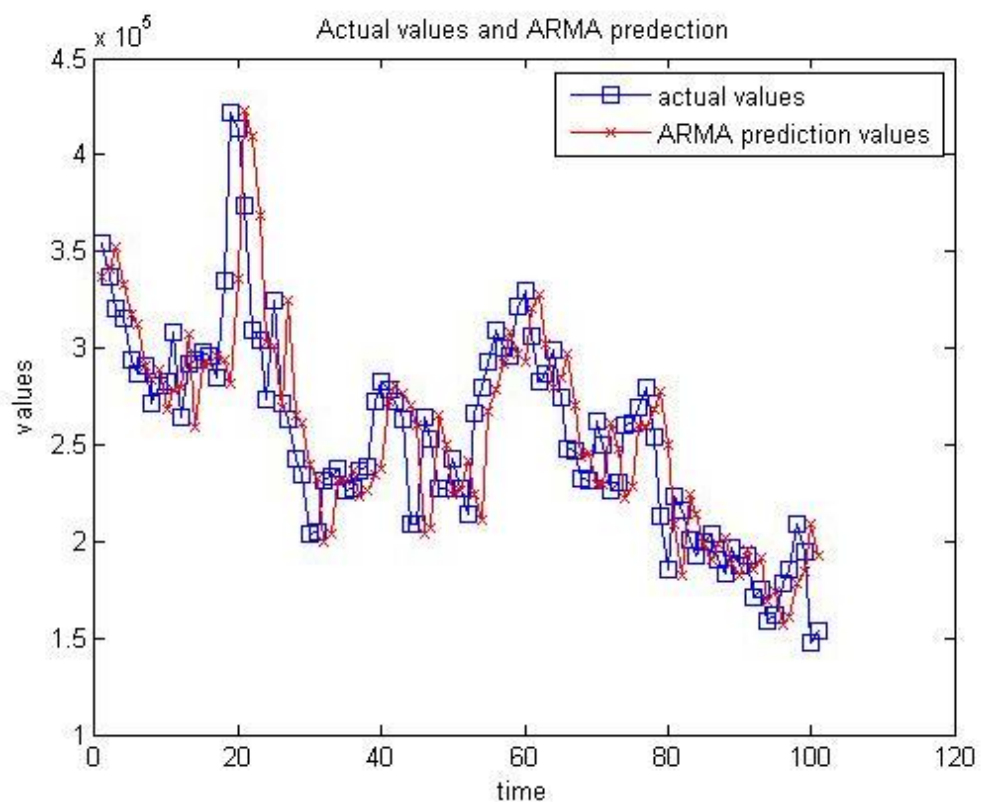
Γράφημα 4 Τελικές συναρτήσεις συμμετοχής της δοκιμής



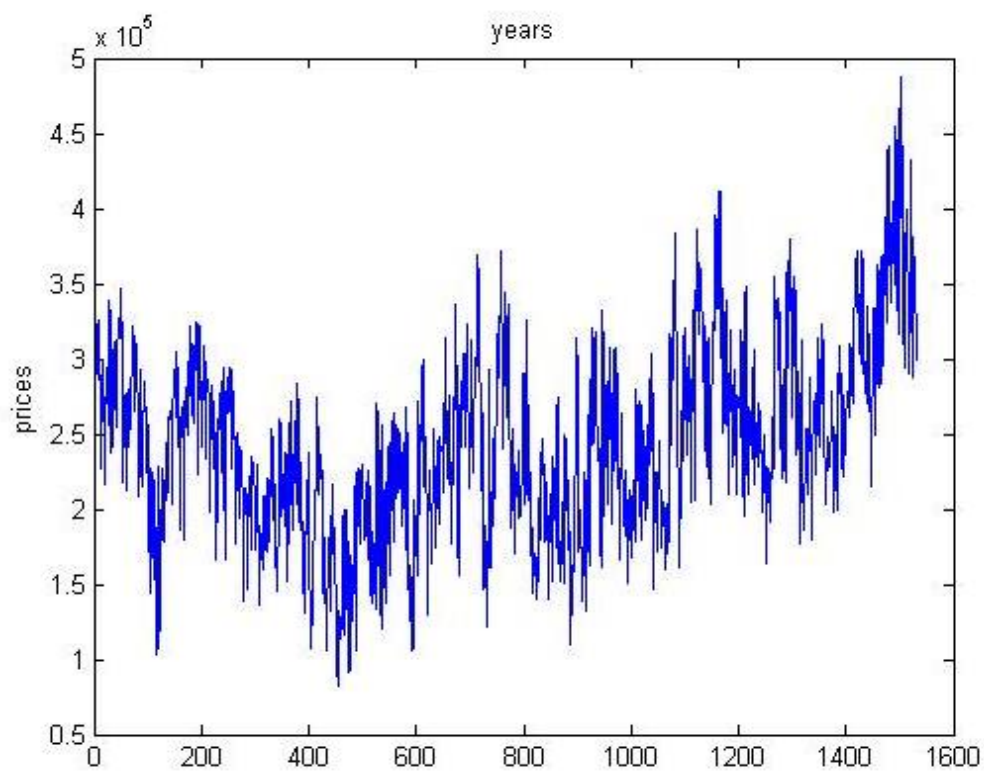
Γράφημα 5 Αρχικές συναρτήσεις συμμετοχής της δοκιμής



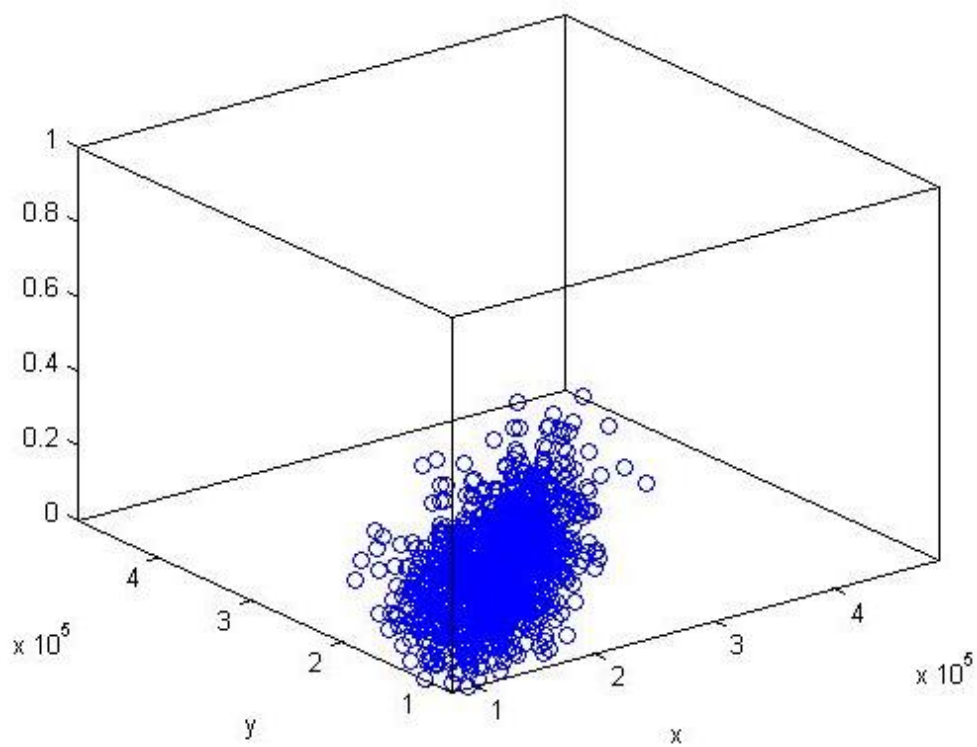
Γράφημα 6 Σύγκριση προβλέψεων του μοντέλου AR έναντι των επιθυμητών τιμών.



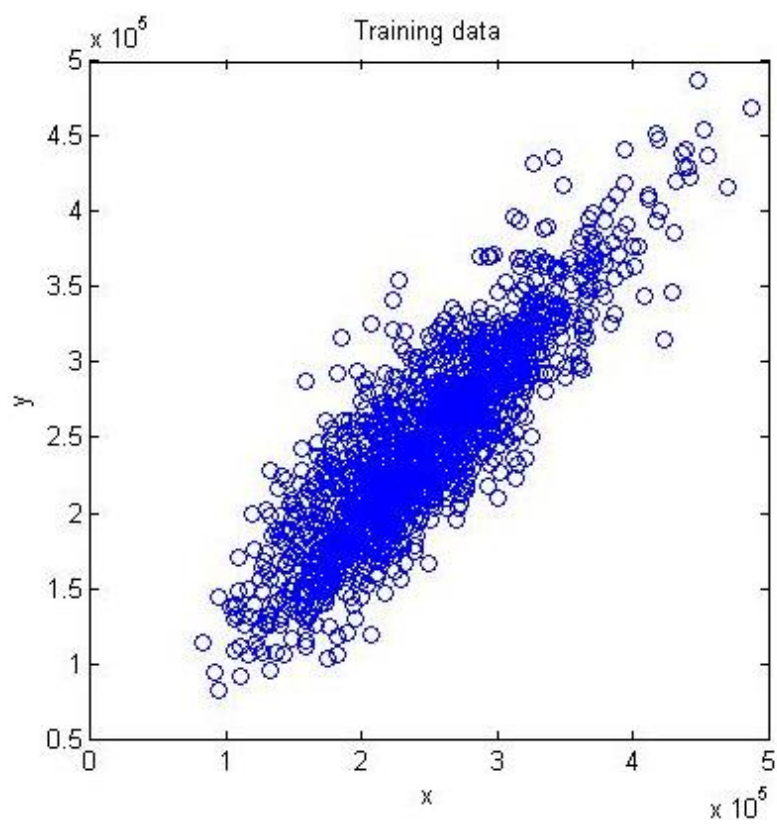
Γράφημα 7 Σύγκριση προβλέψεων του μοντέλου ARMA έναντι των επιθυμητών τιμών.



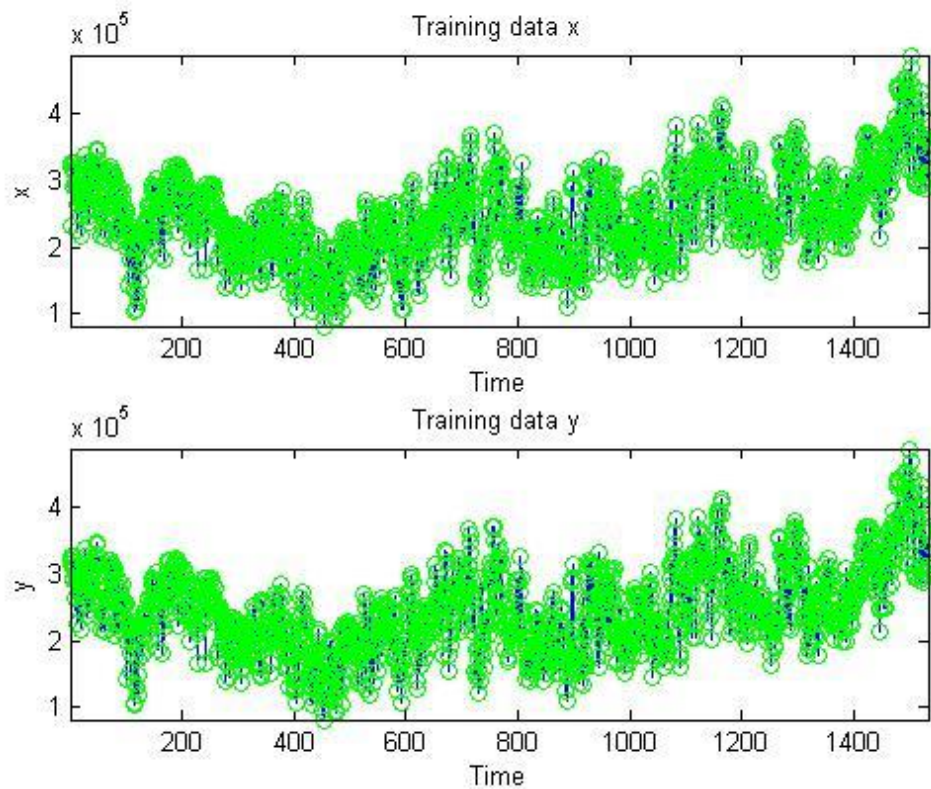
Γράφημα 8 Δεδομένα εκπαίδευσης



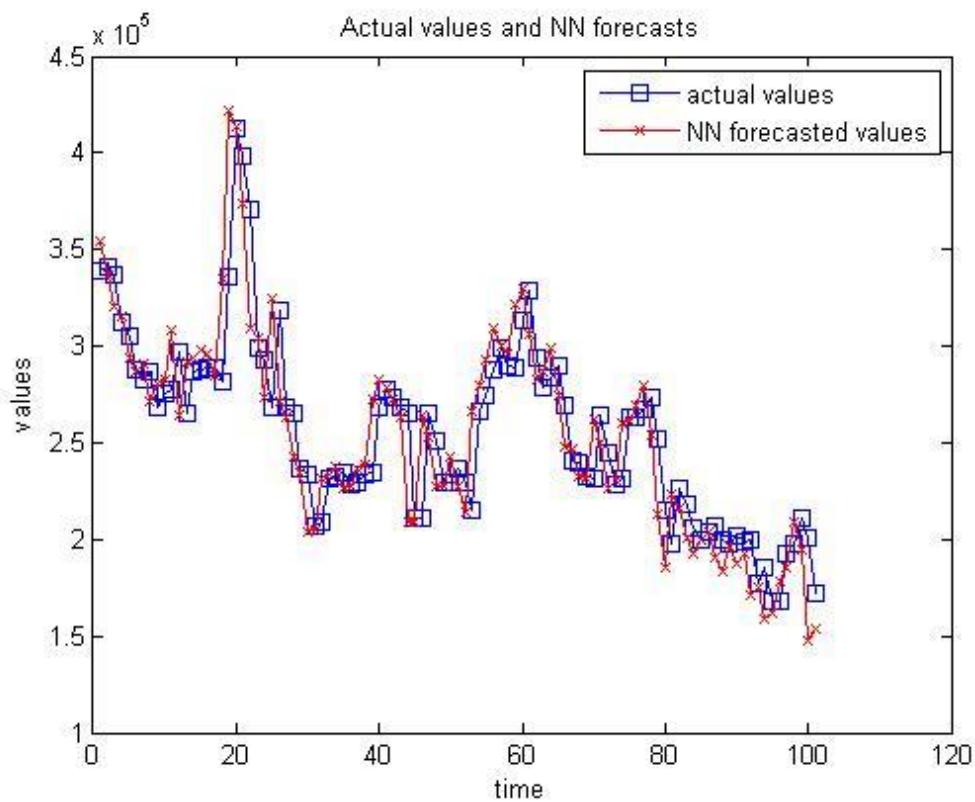
Γράφημα 9 Δεδομένα εκπαίδευσης (διασπορά) σε τρισδιάστατη μορφή



Γράφημα 10 Δεδομένα εκπαίδευσης (διασπορά) σε διοδιάστατη μορφή



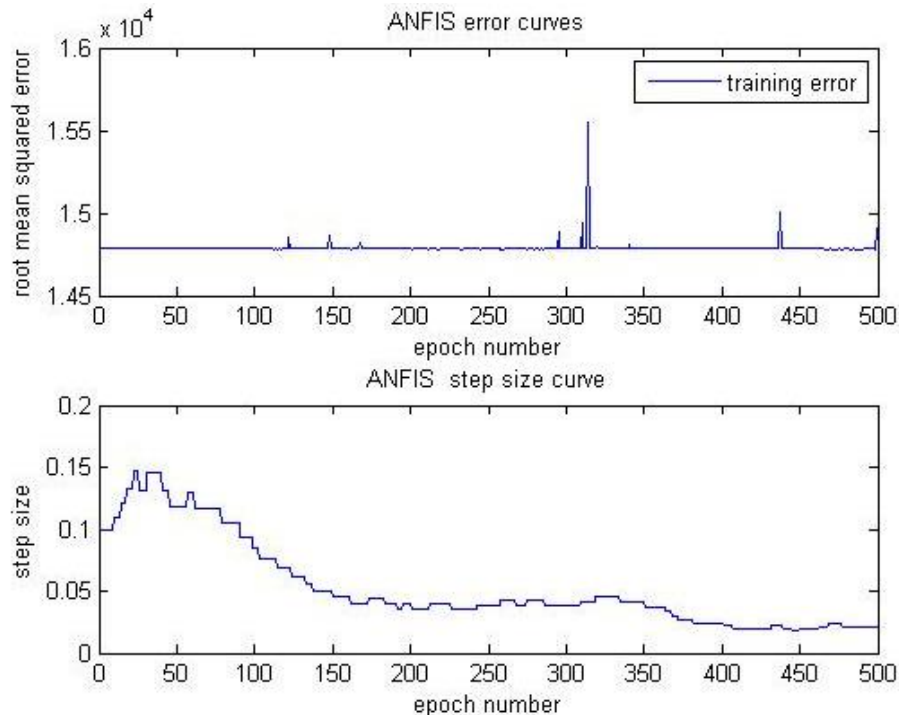
Γράφημα 11 Δεδομένα εκπαίδευσης (διασπορά)



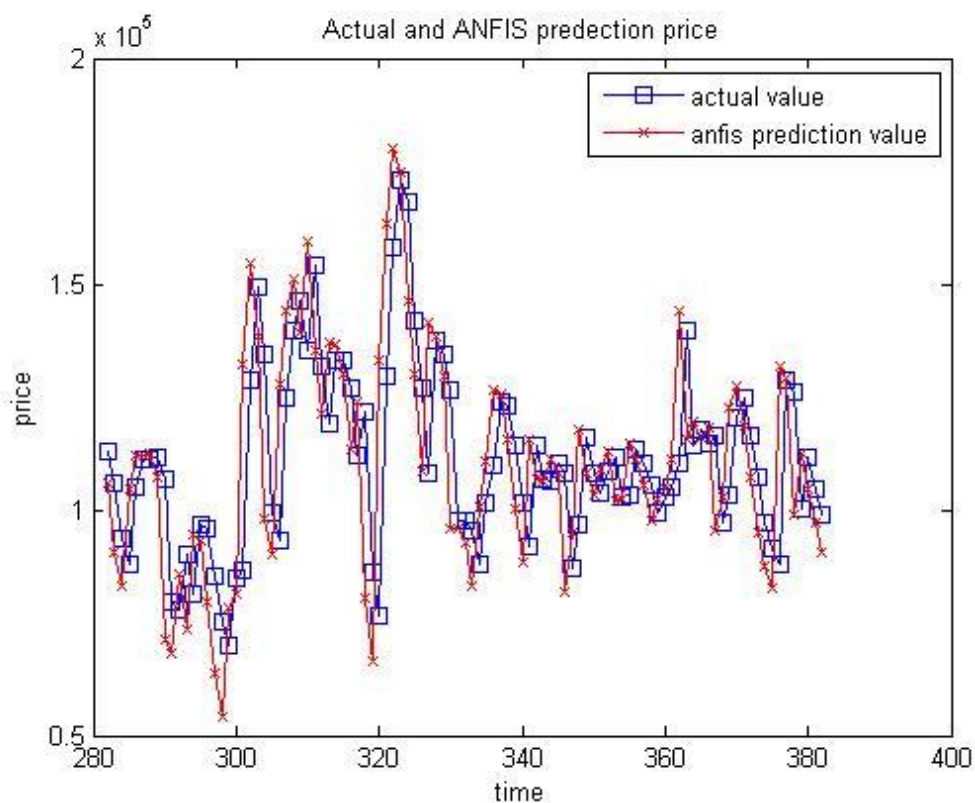
Γράφημα 12 Σύγκριση προβλέψεων του μοντέλου AR έναντι των επιθυμητών τιμών

4.5.2 Νέα Αγγλία

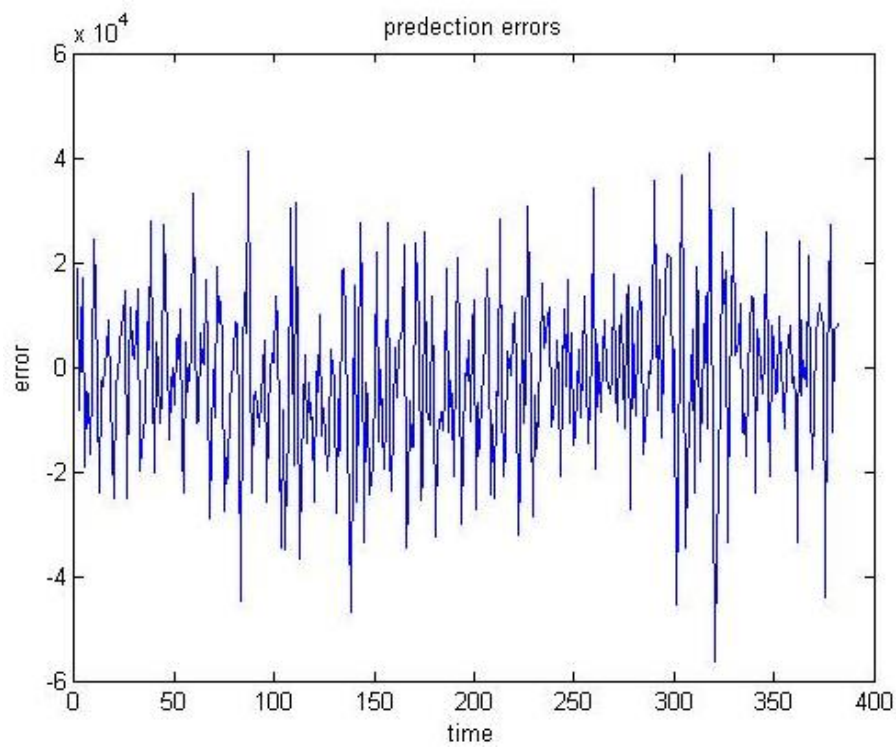
Παρακάτω απεικονίζονται ενδεικτικά κάποια γραφήματα από τα τελικά αποτελέσματα για τις διάφορες δοκιμές που πραγματοποιήσαμε. Συγκεκριμένα, αφορούν την ανάλυση με τη gbell συνάρτηση και αριθμό εποχών 500.



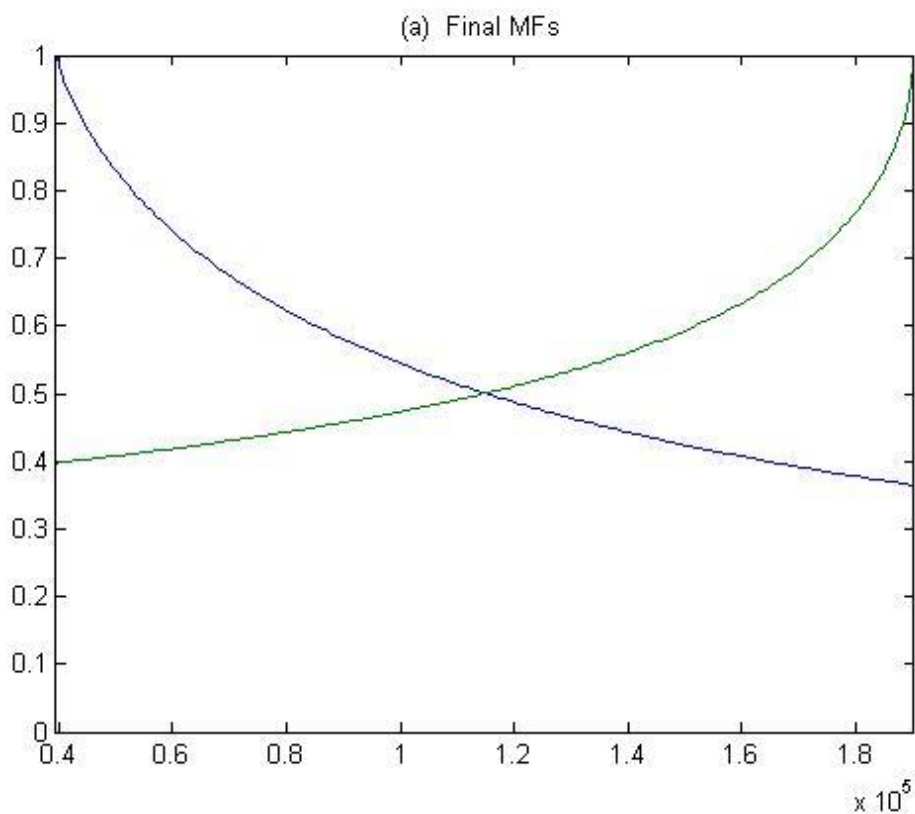
Γράφημα 13 Καμπύλες σφαλμάτων και καμπύλες μεγέθους βήματος της δοκιμής



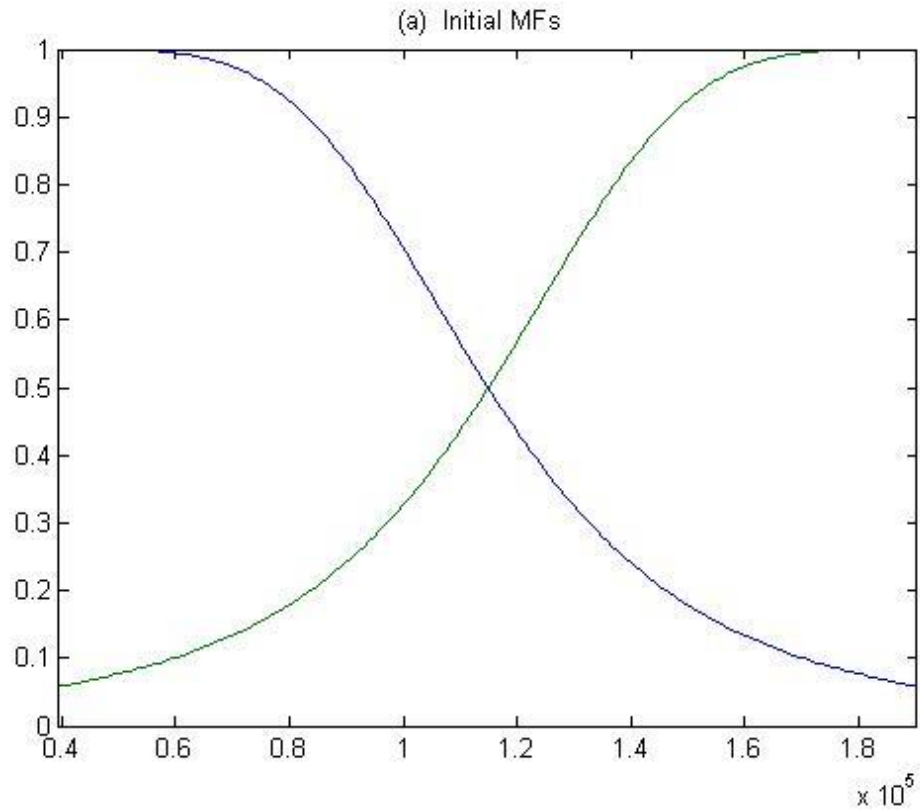
Γράφημα 14 Σύγκριση προβλέψεων του ANFIS έναντι των επιθυμητών τιμών της δοκιμής



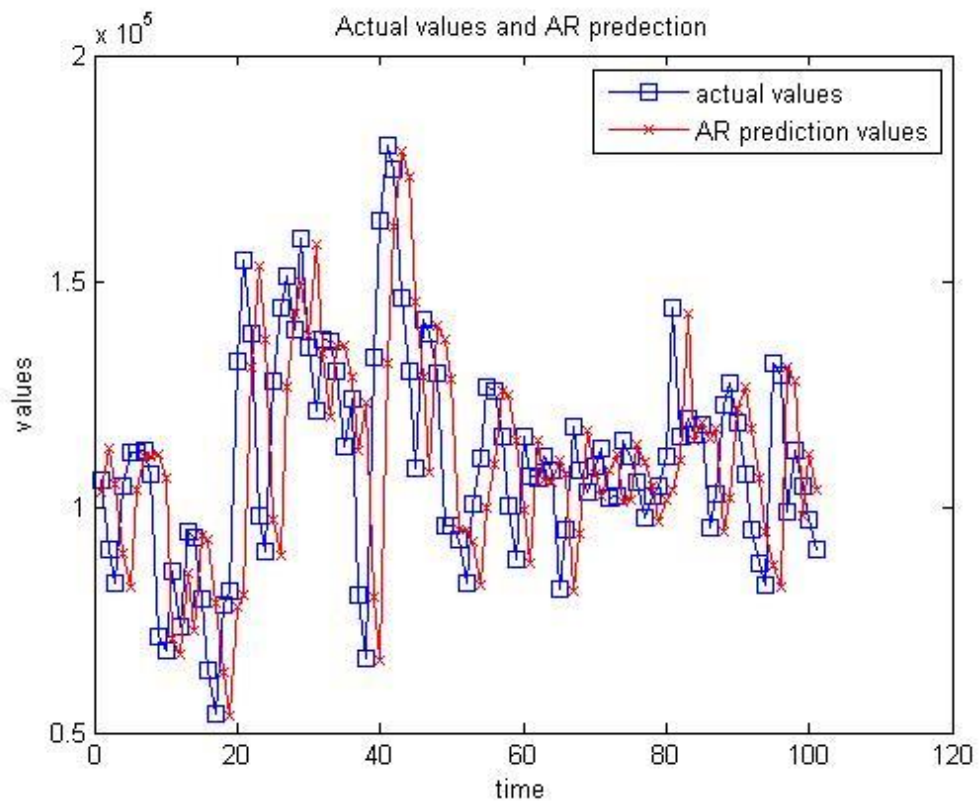
Γράφημα 15 Σφάλματα πρόβλεψης του ANFIS της δοκιμής



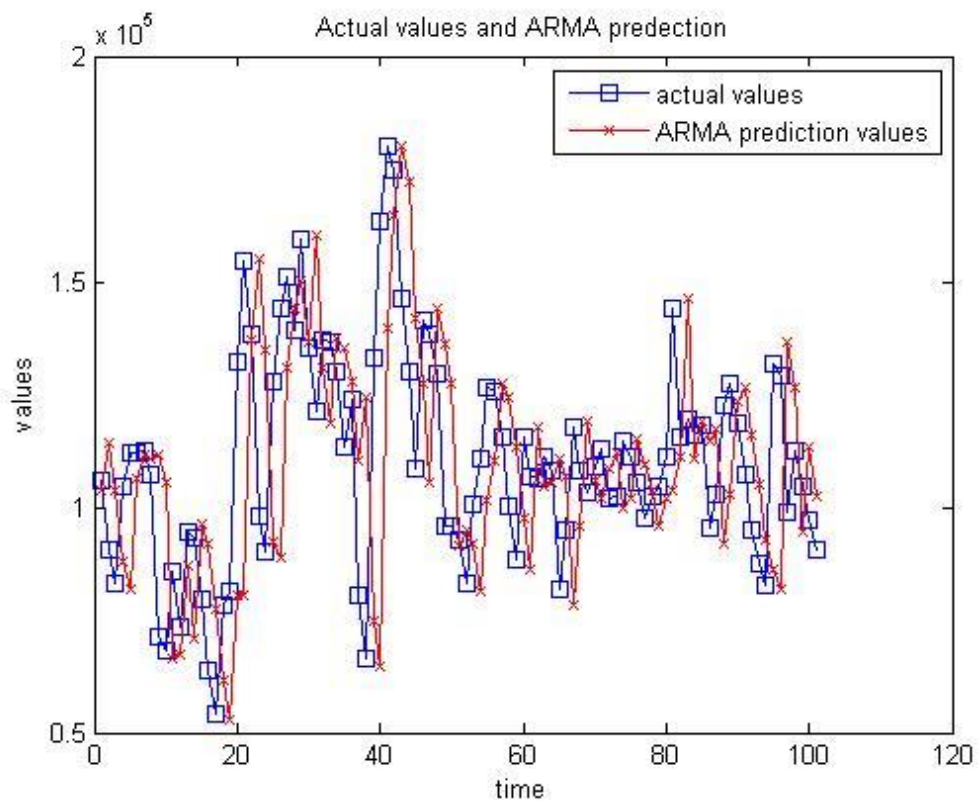
Γράφημα 16 Τελικές συναρτήσεις συμμετοχής της δοκιμής



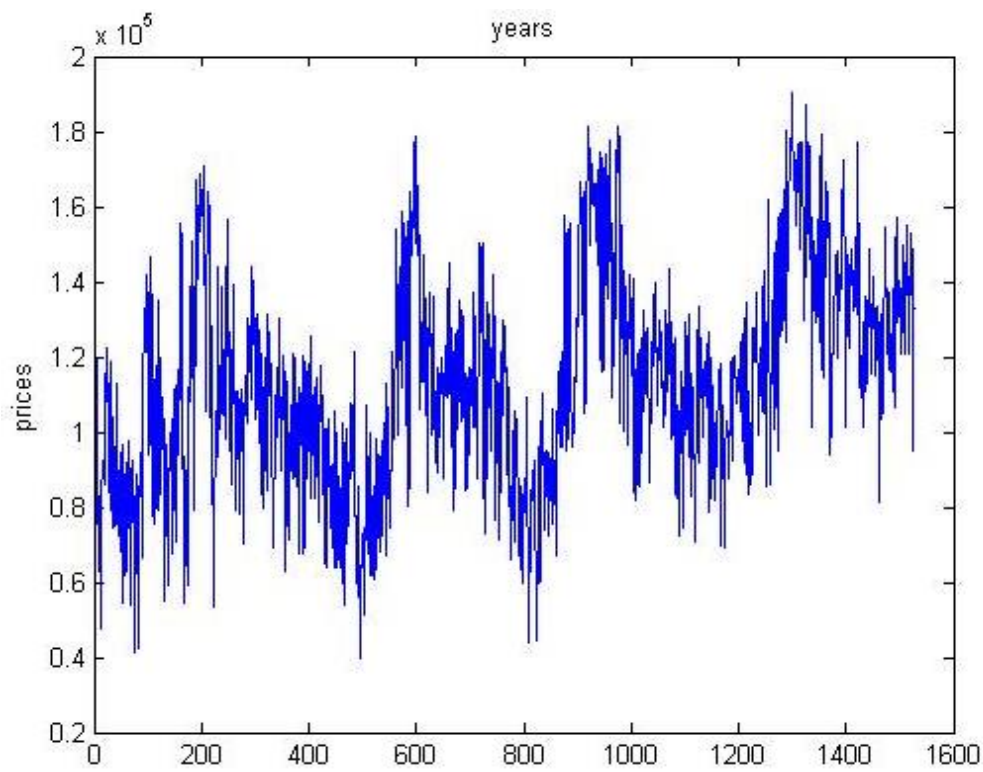
Γράφημα 17 Αρχικές συναρτήσεις συμμετοχής της δοκιμής



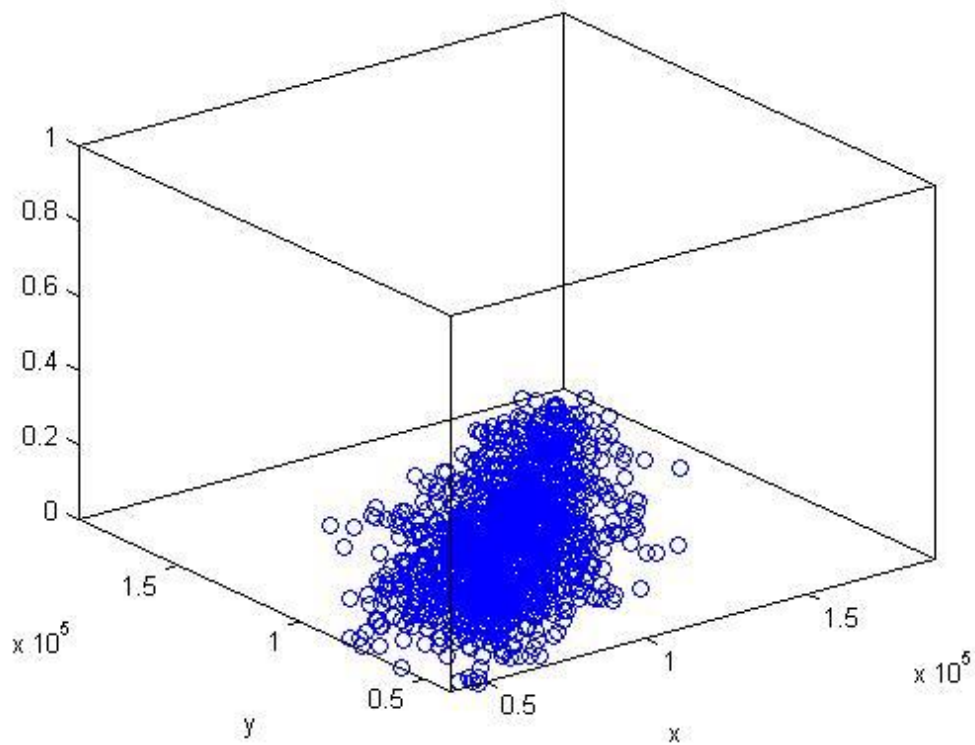
Γράφημα 18 Σύγκριση προβλέψεων του μοντέλου AR έναντι των επιθυμητών τιμών



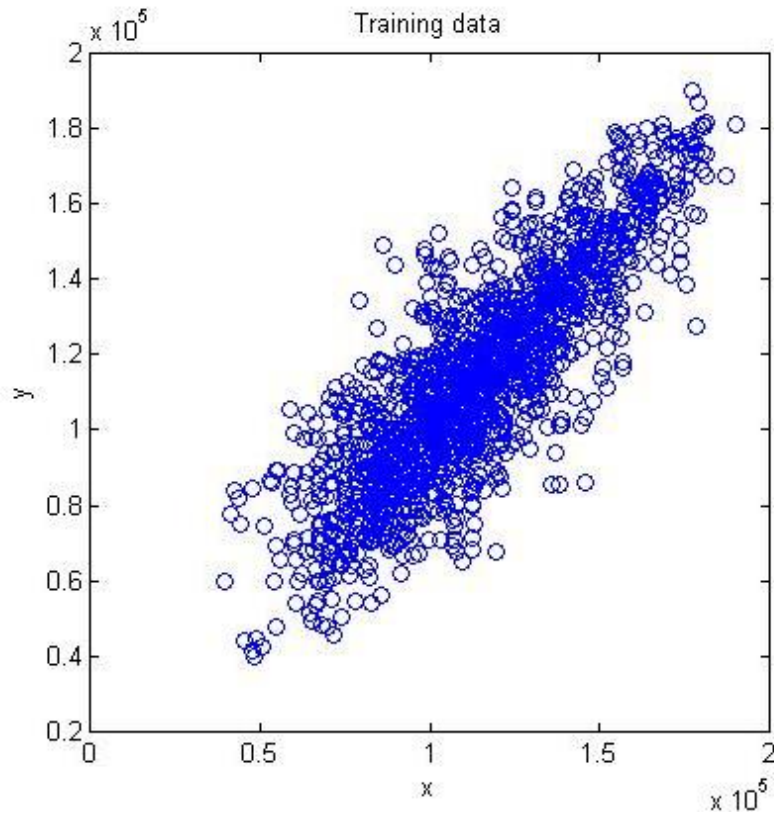
Γράφημα 19 Σύγκριση προβλέψεων του μοντέλου ARMA έναντι των επιθυμητών τιμών



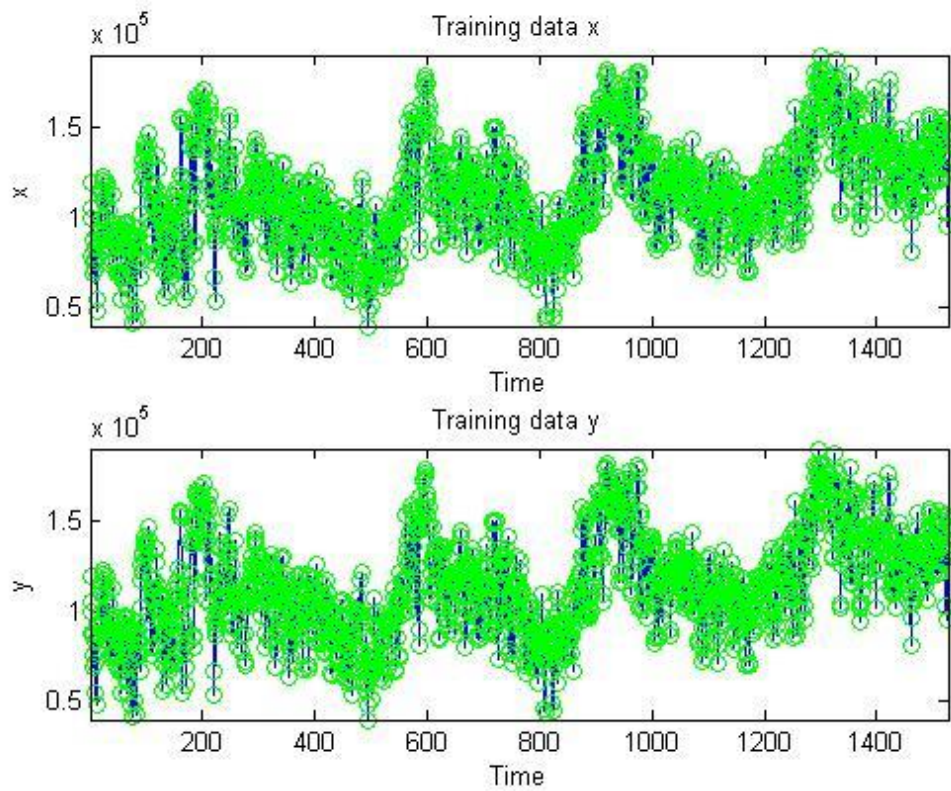
Γράφημα 20 Δεδομένα εκπαίδευσης



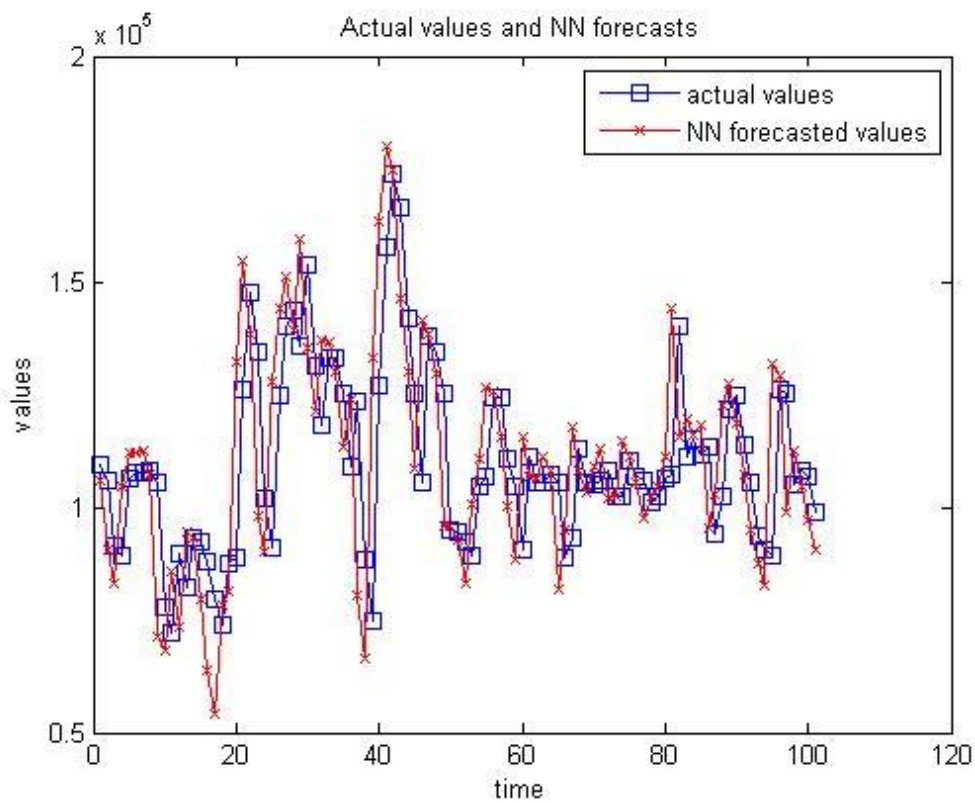
Γράφημα 21 Δεδομένα εκπαίδευσης (διασπορά) σε τρισδιάστατη μορφή



Γράφημα 22 Δεδομένα εκπαίδευσης (διασπορά) σε διδιάστατη μορφή



Γράφημα 23 Δεδομένα εκπαίδευσης (διασπορά)



Γράφημα 24 Σύγκριση προβλέψεων του μοντέλου NN έναντι των επιθυμητών τιμών

5 ΣΥΜΠΕΡΑΣΜΑΤΑ

Το μοντέλο μας χρησιμοποιεί ένα τμήμα των δεδομένων (insample) για την εκπαίδευση του και ένα τμήμα (out of sample) για τον έλεγχο των τιμών πρόβλεψης. 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 τους πίνακες των αποτελεσμάτων του ANFIS και για τις δύο περιοχές, γίνεται αντιληπτό ότι τα αποτελέσματα της πρόβλεψης ήταν ακριβή αφού οι τιμές των σφαλμάτων ήταν χαμηλές. Συγκεκριμένα, για την Ελλάδα τα καλύτερα αποτελέσματα εμφανίζονται στη συνάρτηση gauss με ίδιο αριθμό RMSE και για τις 3 εποχές. Ενώ για τη Νέα Αγγλία βάση το RMSE τα καλύτερα αποτελέσματα τα έχω για τη τριγωνική συνάρτηση συμμετοχής και για 500 εποχές, με μικρές όμως διαφορές με τις υπόλοιπες εποχές. Πρέπει να σημειωθεί ότι οι τιμές αρκετών αποτελεσμάτων ήταν κοινές για τις τρεις διαφορετικές εποχές που χρησιμοποιήθηκαν και στις δύο περιοχές. Ενώ σε σχέση με τα μοντέλα πρόβλεψης, τα καλύτερα αποτελέσματα τα παρέχει το ANFIS έναντι των AR, ARMA και NN.

Γενικότερα προβλέψεις που πραγματοποιούνται με νευρωνικά δίκτυα είναι πιο αξιόπιστες από αυτές των υπόλοιπων κλασικών μεθόδων. Με την απαίτηση όμως τα στατιστικά στοιχεία προς επεξεργασία να είναι μεγάλα σε πλήθος χρονικής περιόδου παρελθόντος. Σε περίπτωση που τα δεδομένα εκπαίδευσης δεν είναι πολλά αυξάνεται ο κίνδυνος μη ασφαλούς πρόβλεψης με υψηλά στατιστικά σφάλματα. Ως τελικό συμπέρασμα από την σύγκριση που πραγματοποιήθηκε στην εργασία αποδείχτηκε πως το μοντέλο που χρησιμοποιήθηκε μπορεί να πραγματοποιήσει πολύ ικανοποιητικές προβλέψεις για την εισαγωγή φυσικού αερίου.

Μέσα από αυτή την εργασία δείξαμε ότι η εισαγωγή φυσικού αερίου σε καθημερινή βάση, μπορεί να συλληφθεί από ένα σύστημα ANFIS το οποίο μπορεί να μας βοηθήσει για να προβλέψουμε τις βραχυπρόθεσμες ποσότητες εισαγωγής. Τα αποτελέσματα είναι πολύ ενθαρρυντικά και έχουν δείξει ότι το μοντέλο ANFIS αποδίδει καλύτερα από τις άλλες συμβατικές μεθόδους πρόβλεψης μιας και οι κανόνες του είναι απλούστεροι και ο χρόνος εκπαίδευσης του ήταν μικρότερος. Αυτό, με τη σειρά του, σημαίνει ότι το μοντέλο ANFIS μπορεί επιτυχώς να χρησιμοποιηθεί στον οικονομικό στρατηγικό προγραμματισμό.

6 ΒΙΒΛΙΟΓΡΑΦΙΑ

- Zhong-bing Zhou, Xiu-cheng Dong (2012). Analysis about the seasonality of China's crude oil import based on X-12-ARIMA. *Energy* 42 (2012), 281-288
- 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.
- V. Majazi Dalfard, M. Nazari Asli, S.M. Asadzadeh, S.M. Sajjadi, A. Nazari-Shirkouhi (2013). A mathematical modeling for incorporating energy price hikes into total natural gas consumption forecasting. *Applied Mathematical Modelling* 37 (2013), 5664–5679
- 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.
- Marek Brabec, Ondřej Konár, Emil Pelikán, Marek Malý (2008). A nonlinear mixed effects model for the prediction of natural gas consumption by individual customers. *International Journal of Forecasting* 24 (2008), 659–678
- Jiri Vondracek, Jana Cermakova, Krystof Eben, Marek Brabec, Ondřej Konár, Emil Pelikán, Marek Malý (2008). A statistical model for the estimation of natural gas consumption. *Applied Energy* 85 (2008), 362–370
- A. Jai Persaud, Uma Kumar (2001). An eclectic approach in energy forecasting: a case of Natural Resources Canada's (NRCan's) oil and gas outlook. *Energy Policy* 29 (2001) 303-313
- Bruce E. Marti (2008). Assessing New England's waterborne energy imports (1995–2004). *Marine Policy* 32 (2008), 740–748
- Bengt Soderbergh, Kristofer Jakobsson, Kjell Aleklett (2010). European energy security: An analysis of future Russian natural gas production and exports. *Energy Policy* 38(2010), 7827–7843
- Fatih Taspınara, Numan Celebi, Nedim Tutkun (2013). Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods. *Energy and Buildings* 56 (2013), 23–31
- Boqiang Lin, Ting Wang (2012). Forecasting natural gas supply in China: Production peak and import trends. *Energy Policy* 49 (2012), 225–233

Harry Martz, Ani-Hony Burris, Lawrence Bruckner, Robert Bivins (1986). Kalman filter estimation of monthly U.S oil imports. *Energy Vol. I I, No. 3. pp. 27 I-280.*

Junchen Li, Xiucheng Dong, Jianxin Shangguan, Mikael Höök (2011). Forecasting the growth of Chinese natural gas consumption. *Energy 36, 1380-1385*

Zia Wadud, Himadri S.Dey, Md.Ashfanor Kabir, Shahidul I.Khan (2011). Modeling and forecasting natural gas demand in Bangladesh. *Energy Policy 39(2011), 7372–7380.*

H. Sarak, A. Satman (2003). The degree-day method to estimate the residential heating natural gas consumption in Turkey: a case study. *Energy 28 (2003), 929–939*

R.E. Abdel-Aal (2008). Univariate modeling and forecasting of monthly energy demand time series using abductive and neural networks. *Computers & Industrial Engineering 54 (2008), 903–917*