



ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ

Διπλωματική εργασία:

ΜΕΘΟΔΟΛΟΓΙΕΣ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΝΟΗΜΟΣΥΝΗΣ ΣΤΗΝ
ΠΡΟΒΛΕΨΗ ΧΡΗΜΑΤΙΣΤΗΡΙΑΚΩΝ ΤΑΣΕΩΝ: ΕΦΑΡΜΟΓΗ ΣΤΟ
ΧΡΗΜΑΤΙΣΤΗΡΙΟ ΑΞΙΩΝ ΑΘΗΝΩΝ

Τσιώρης Ιωάννης

Χανιά 2008

Ευχαριστίες:

Με την εκπόνηση της διπλωματικής μου, ολοκληρώνεται η φοίτηση μου στο Πολυτεχνείο Κρήτης. Θα ήθελα να ευχαριστήσω θερμά τον επίκουρο καθηγητή Μιχάλη Δούμπο, για την εντατική επίβλεψη, το ενδιαφέρον και τις καθοριστικά σημαντικές παρεμβάσεις. Ακόμα θα ήθελα να ευχαριστήσω τον Ατσαλάκη Γιώργο για την συνεργασία του, στα θέματα που αφορούν το σύστημα ANFIS.

Έχοντας την ευκαιρία, θα ήθελα να ευχαριστήσω τους γονείς μου Νταϊάνα και Κώστα, που με υποστήριξαν αυτά τα χρόνια -συχνά με αυταπάρνηση-. Τέλος, ευχαριστώ ξεχωριστά καθέναν από τους φίλους μου: Αγγελιδάκη Ιωάννα, Γιαννάκη Γιώργο, και Κουρκουρίκη Γιάννη, για την ηθική συμπαράσταση.

Περιεχόμενα:

Κεφάλαιο 1:	Εισαγωγή	5
Κεφάλαιο 2:	Υπολογιστικές διαδικασίες στη χρημα-τιστηριακή πρόβλεψη	10
2.1	Μεθοδολογικές προσεγγίσεις	10
2.1.1	Τεχνητά νευρωνικά δίκτυα	11
2.1.2	Ασαφή λογική.....	14
2.1.3	Έμπειρα συστήματα	17
2.1.4	Γενετικοί αλγόριθμοι.....	20
2.2	Εφαρμογές στη χρηματιστηριακή πρόβλεψη	24
2.2.1	Χρηματιστηριακή πρόβλεψη με τεχνητά νευρωνικά δίκτυα.....	24
2.2.2	Χρηματιστηριακή πρόβλεψη με ασαφή λογική	25
2.2.3	Χρηματιστηριακή πρόβλεψη με έμπειρα συστήματα	25
2.2.4	Χρηματιστηριακή πρόβλεψη με γενετικούς αλγόριθμους	26
Κεφάλαιο 3:	Μεθοδολογικές προσεγγίσεις.....	27
3.1	Μηχανές διανυσμάτων υποστήριξης (SVM).....	27
3.1.1	Η γραμμικά διαχωρίσιμη περίπτωση.....	27
3.1.2	Η Γραμμικά Μη Διαχωρίσιμη Περίπτωση	31
3.1.3	Μη Γραμμικές Μηχανές Διανυσμάτων Υποστήριξης.....	33
3.1.4	Πλεονεκτήματα Μηχανών Διανυσμάτων Υποστήριξης	35
3.1.5	Εφαρμογές των SVM	36
3.2	Προσαρμοστικά νευροασαφή συστήματα εξαγωγής συμπερασμάτων (ANFIS) 37	
3.2.1	Γενικά	37
3.2.2	Ασαφή συστήματα εξαγωγής συμπερασμάτων (FIS)	37
	Ένα FIS αντιστοιχεί κατά σειρά:.....	37

3.2.3	Εκπαίδευση του ANFIS.....	38
3.2.4	Αρχιτεκτονική ANFIS	41
3.2.5	Πλεονεκτήματα του συστήματος ANFIS.....	43
3.2.6	Εφαρμογές του ANFIS	43
Κεφάλαιο 4:	Δεδομένα, μεθοδολογία ανάλυσης και αποτελέσματα.....	45
4.1	Δεδομένα	45
4.2	Δομή δεδομένων.....	47
4.3	Πειραματικός σχεδιασμός	49
4.4	Αποτελέσματα	52
4.4.1	SVM	52
4.4.2	ANFIS.....	57
4.5	Σύγκριση αποτελεσμάτων	61
Κεφάλαιο 5:	Συμπεράσματα.....	62

Κεφάλαιο 1: Εισαγωγή

Η αναζήτηση κερδοφόρων επενδύσεων αποτελεί μια πραγματικότητα διαχρονική. Διαχρονική είναι επίσης η ανάγκη των επενδυτών για όσο το δυνατόν υψηλότερες αλλά και ασφαλέστερες αποδόσεις. Η πλειοψηφία των επενδύσεων ωστόσο, συνοδεύεται από κάποια αβεβαιότητα, που θέτει σε κίνδυνο το επενδεδυμένο κεφάλαιο

Συχνή είναι η στροφή των επενδυτών στην αγορά μετοχών στο χρηματιστήριο για να αξιοποιήσουν τα κεφάλαια τους. Σε αυτή τη μορφή επένδυσης, ο κίνδυνος απώλειας κεφαλαίων είναι ακόμα μεγαλύτερος, εξαιτίας των συχνών διακυμάνσεων που εμφανίζει η αξία μιας μετοχής. Η πρόβλεψη μίας τέτοιας χρονοσειράς είναι από τις δυσκολότερες, ειδικά όταν επιζητείται να είναι βραχυπρόθεσμη (ημερήσια, ωριαία-λεπτού κ.ο.κ.).

Το θέμα που διαπραγματεύεται η παρούσα διπλωματική είναι η προσπάθεια επιτυχούς κατηγοριοποίησης-ταξινόμησης (**classification**), της κίνησης ανόδου ή καθόδου του γενικού δείκτη, του Χρηματιστηρίου Αξιών Αθηνών. Σκοπός είναι η πρόβλεψη, με ικανοποιητική ακρίβεια, της τάσης του γενικού δείκτη κατά την επόμενη συνεδρίαση, με τη χρήση μηχανών διανυσμάτων υποστήριξης (Support Vector Machines,SVM) από τη μια και του συστήματος ANFIS (Adaptive Neuro Fuzzy Inference System) από την άλλη. Τα δεδομένα που αξιοποιήθηκαν αφορούν στις ημερήσιες τιμές κλεισίματος του γενικού δείκτη, κατά τη χρονική περίοδο 1997-2007.

Οι SVM, αποτελούν μια σχετικά νέα και καινοτόμα προσέγγιση, στο πρόβλημα της ταξινόμησης (Vapnik, 1998). Είναι ένας αλγόριθμος αναγνώρισης προτύπων, που διαφοροποιείται σημαντικά από τους προγενέστερους του. Βασικό χαρακτηριστικό των SVM, είναι πως ελαχιστοποιούν ένα άνω όριο του αναμενόμενου σφάλματος γενίκευσης, ενώ οι άλλες μεθοδολογίες, ελαχιστοποιούν το μέσο σφάλμα πάνω στα δεδομένα εκμάθησης.

Το σύστημα ANFIS, δημιουργήθηκε από τον Jang (1993), ανήκει στα προσαρμοστικά δίκτυα και παρουσιάζει πολλές ομοιότητες με τα νευρωνικά δίκτυα. Αποτελεί έναν συνδυασμό νευρωνικών δικτύων (Artificial Neural Networks-ANN) και ασαφούς λογικής (Fuzzy logic-FL). Διαχειρίζεται τόσο αριθμητικά όσο και ποιοτικά

χαρακτηριστικά της διαδικασίας. Η εξαγωγή γνώσης από το σύστημα είναι άμεσα εφικτή, εν αντίθεση με τα νευρωνικά δίκτυα που αποτελούν μαύρο κουτί (black box). Κάτι τέτοιο είναι εφικτό καθώς τα νευροασαφή συστήματα παράγουν κανόνες οι οποίοι γίνονται άμεσα κατανοητοί. Το ANFIS, συνδυάζει τα πλεονεκτήματα των συστατικών του (ANN & FL). Έτσι περιεχεί ισχυρό υπόβαθρο για την αναπαράσταση της γνώσης, ενώ τα νευρωνικά δίκτυα παρέχουν δυνατότητες μάθησης και δημιουργίας αποδοτικών υπολογιστικών εφαρμογών με χρήση H/Y.

Βασική προϋπόθεση, για να υπάρξει δυνατότητα επιστημονικής προσέγγισης, είναι η ύπαρξη κάποιας συσχέτισης των δεδομένων. Πολλές είναι οι ακαδημαϊκές μελέτες που συμπεραίνουν πως η χρηματιστηριακή πρόβλεψη, δεν είναι δυνατή. Η θεωρία αυτή είναι γνωστή ως «**Υπόθεση της Αποτελεσματικής Αγοράς**» (Efficient Market Hypothesis-EMH). Σύμφωνα με την EMH οι τιμές των μετοχών αντικατοπτρίζουν πλήρως και σωστά όλες τις διαθέσιμες κάθε χρονική στιγμή πληροφορίες που τις αφορούν. Κάθε ανισορροπία θα αποκαθίσταται καθώς νέες πληροφορίες γνωστοποιούνται. Η θεωρία EMH αντιλαμβάνεται την αγορά ως ένα μηχανισμό με ταχύτατη απόκριση και προσαρμογή στις εξελίξεις, με αποτέλεσμα κανείς να μην μπορεί να επωφεληθεί από αυτές. Οι τρεις βασικές παραδοχές της θεωρίας EMH είναι:

- Η χρήση εμπιστευτικών πληροφοριών δεν επιτρέπεται
- Όλοι οι επενδυτές έχουν πανομοιότυπη πληροφόρηση και ισότιμη μεταχείριση.
- Δεν υπάρχουν επενδυτές με μονοπωλιακή ισχύ. Επενδυτές δηλαδή, που λόγω του όγκου μετοχών που κατέχουν μπορούν να επηρεάζουν την κατεύθυνση κίνησης των μετοχών.

Από τη άλλη δεν είναι λίγοι οι ερευνητές, που υποστηρίζουν το αντίθετο. Ότι δηλαδή οι ανισορροπίες της αγοράς απαιτούν κάποιο χρονικό διάστημα για να αποκατασταθούν, αφήνοντας έτσι περιθώριο αποκομιδής κερδών. Η ύπαρξη αυτού του χρονικού διαστήματος πηγάζει από το γεγονός πως η πληροφόρηση δεν φτάνει την ίδια χρονική στιγμή σε όλους. Οι ερευνητές αυτής της περισσότερο αισιόδοξης πεποίθησης, έχουν να αντιπαραθέσουν πλήθος επιστημονικών μεθόδων και αποτελεσμάτων στο θέμα της χρηματιστηριακής πρόβλεψης.

Στην προσπάθεια πρόβλεψης τιμών-τάσεων μετοχών, υπάρχουν δύο προσεγγίσεις: η θεμελιώδης ανάλυση και η τεχνική ανάλυση:

Η **θεμελιώδης ανάλυση** λαμβάνει υπόψη της μακροοικονομικούς δείκτες και δείκτες της επιχείρησης, που πηγάζουν από τις χρηματοοικονομικές της καταστάσεις. Μπορεί ακόμη να λαμβάνει υπόψη της, αλλά χαρακτηριστικά της επιχείρησης ή του εξωτερικού περιβάλλοντος (κυβερνητική πολιτική, οικονομικό-επιχειρηματικό περιβάλλον κ.α.), που ο αναλυτής θεωρεί πως μπορεί να βελτιώσουν την προβλεπτική ικανότητα του μοντέλου. Χρησιμοποιείται για μεσοπρόθεσμες ή μακροπρόθεσμες προβλέψεις. Ακρογωνιαίος λίθος της θεμελιώδους ανάλυσης είναι η υπόθεση πως η μετοχή αντανακλά την οικονομική κατάσταση στην οποία βρίσκεται η αντίστοιχη εταιρία. Βασικά μειονεκτήματα της προσέγγισης είναι:

- Απαιτείται μεγάλη ποσότητα πληροφοριών, πολλές από τις οποίες είναι δύσκολο να συλλεχθούν.
- Δεν μπορεί να συνυπολογίσει εξωγενείς παράγοντες, που μπορούν να επηρεάσουν την τιμή μιας μετοχής ή και ολόκληρη την αγορά. Ως αποτέλεσμα, η θεμελιώδης ανάλυση δεν μπορεί να προστατέψει τον επενδύτη από μια αποσταθεροποίηση της αγοράς.
- Συχνή δυσκολία ερμηνείας των απότομων ανόδων ή πτώσεων των τιμών. Αδυναμία που συναντάνε και οι άλλες μέθοδοι που θα διαπραγματευτούν στην παρούσα μελέτη, με τη διαφορά πως αυτές διαχειρίζονται τάσεις.

Η **τεχνική ανάλυση**, αξιοποιεί την συσχέτιση τιμών και όγκου συναλλαγών για να εκτιμήσει την πορεία μιας μετοχής. Η τεχνική ανάλυση αναζητά στα δεδομένα κάποιο είδος προτύπου-μορφοδιάταξης, η οποία μπορεί να υποδεικνύει κάποια επαναληπτικότητα. Θεωρεί πως τα στοιχεία που λαμβάνει υπόψη της η θεμελιώδης ανάλυση είναι ενσωματωμένα στις τιμές των μετοχών και επομένως η πρωτογενής αξιοποίηση τους δεν απαιτείται. Οι μέθοδοι που θα διαπραγματευτούν, εντάσσονται στην τεχνική ανάλυση και βασίζονται σε μια πεποίθηση: οι αριθμοί που εκφράζουν την πορεία ενός φαινομένου στο χρόνο, εμπεριέχουν κάποια κρυμμένη γνώση. Η γνώση αυτή είναι δυνατόν να ευρεθεί και να χρησιμοποιηθεί για μελλοντικές εκτιμήσεις. Η θεωρία της τεχνικής ανάλυσης αναπτύχθηκε αρχικά από τον Charles Dow το 1884.

Μια νεότερη εκδοχή δίδεται στο βιβλίο των Edwards και Magee (1997) Βασικά μειονεκτήματα της προσέγγισης αυτής είναι:

- Τα σημαντικά κέρδη που χάνονται κατά την φάση εντοπισμού της τάσης καθώς απαιτείται χρόνος μέχρι να εντοπιστεί. Δεν εντοπίζεται ή τάση κατά τη δημιουργία της.
- Η μεγάλη συχνότητα συναλλαγών που πραγματοποιούνται, η οποία οδηγεί σε αυξημένα έξοδα συναλλαγών.
- Οι μεθοδολογίες αυτές δυσκολεύονται να υποστηρίξουν με ισχυρό θεωρητικό υπόβαθρο τον τρόπο λειτουργίας τους.
- Απευθύνεται σε μικρούς επενδυτές και όχι σε θεσμικούς.

Βασικά πλεονεκτήματα της προσέγγισης είναι:

- Απαιτεί μικρή ποσότητα πληροφοριών που είναι εύκολα προσβάσιμη.
- Η αυτοματοποίηση των τεχνικών ώστε να είναι εύκολες στη χρήση συγκριτικά με τις τεχνικές θεμελιώδους ανάλυσης.
- Μεγάλη προσαρμοστικότητα. Η τεχνική ανάλυση εύκολα επεκτείνεται σε άλλους κλάδους της ίδιας αγοράς ή ακόμη σε άλλες αγορές.
- Μπορεί να αντιμετωπίζει καλύτερα καταστάσεις, όπου κάποιοι επενδυτές κατέχουν εσωτερική πληροφόρηση, καθώς προβλέπει τάσεις.

Στο σημείο αυτό θα πρέπει να σημειωθεί πως η θεμελιώδης ανάλυση και η τεχνική ανάλυση δεν αλληλοαναιρούνται. Αντιθέτως δύναται να αλληλοσυμπληρωθούν και να οδηγήσουν σε μια ασφαλέστερη και αποδοτικότερη απόφαση.

Η υπόλοιπη εργασία διαμορφώνεται ως εξής: Στο δεύτερο κεφάλαιο παρουσιάζονται οι σημαντικότερες μεθοδολογίες προσέγγισης του προβλήματος της χρηματιστηριακής πρόβλεψης, καθώς και κάποια αποτελέσματα, που έχουν επιτευχθεί σε κάθε μεθοδολογία. Στο τρίτο κεφάλαιο αναπτύσσεται η θεωρία των μηχανών διανυσμάτων υποστήριξης (SVM) και η θεωρία στην οποία βασίζεται ο νευροασαφής αλγόριθμος (ANFIS). Στο τέταρτο κεφάλαιο επεξηγούνται οι συνθήκες εκτέλεσης των πειραμάτων πρόβλεψης και παρατίθενται και συγκρίνονται τα αποτελέσματα. ως προς την ακρίβεια

των προβλέψεων τους. Τέλος στο πέμπτο κεφάλαιο δίνονται τα συμπεράσματα της εργασίας και προτείνονται μελλοντικές ερευνητικές προοπτικές και επεκτάσεις.

Κεφάλαιο 2: Υπολογιστικές διαδικασίες στη χρηματιστηριακή πρόβλεψη

Η διαδικασία εξαγωγής εκτίμησης, για την έκβαση ενός μελλοντικού γεγονότος, καλείται πρόβλεψη, όταν δε, η εκτίμηση αφορά κάποιο χρηματιστηριακό προϊόν, καλείται χρηματιστηριακή πρόβλεψη. Η χρηματιστηριακή πρόβλεψή είναι ένα δύσκολο έργο, κυρίως διότι οι μελετούμενες χρονοσειρές εμφανίζουν συμπεριφορά που προσεγγίζει αυτή του τυχαίου περιπάτου (random walk). Η ενασχόληση ενός επενδυτή με το χρηματιστήριο αφορά συνήθως δυο φάσεις. Η πρώτη φάση είναι εκείνη της επιλογής και η δεύτερη φάση είναι εκείνη της διαχείρισης. Στην πρώτη φάση ο επενδυτής καλείται να επιλέξει και να επενδύσει σε κάποια επενδυτικά προϊόντα. Η δεύτερη φάση, αφορά τους κατάλληλους χειρισμούς (πώληση – αγορά - διακράτηση κλπ), ώστε να υποστηρίξει και να βελτιώσει την απόδοση των επενδύσεων του. Το αντικείμενο της παρούσας έρευνας, εμπίπτει στη δεύτερη φάση. Η πραγματοποίηση προβλέψεων, είναι σαφώς πιο εύκολη όταν αφορά ανεπτυγμένη αγορά, διότι οι ανεπτυγμένες αγορές εμφανίζουν πολύ μικρότερη αστάθεια. Το ΧΑΑ εντάσσεται πλέον στις ανεπτυγμένες αγορές. Στις εφαρμογές των αλγορίθμων SVM και ANFIS που θα ακολουθήσουν επιχειρείται πρόβλεψη της **τάσης** του γενικού δείκτη. Άλλες πάλι μελέτες επιχειρούν πρόβλεψη της αριθμητικής **τιμής** του δείκτη. Οι συχνότητες των δεδομένων που χρησιμοποιούνται συνήθως για βραχυπρόθεσμη πρόβλεψη είναι η ημερήσια και εβδομαδιαία, ενώ για πιο μακροπρόθεσμη λαμβάνονται μηνιαίες, τριμηνιαίες και ετήσιες συχνότητες, χωρίς φυσικά να αποκλείονται και άλλες.

2.1 Μεθοδολογικές προσεγγίσεις

Πολλές είναι οι μεθοδολογικές προσεγγίσεις, στο θέμα της χρηματιστηριακής πρόβλεψης. Κάποιες από τις σημαντικότερες, που θα αναφερθούν στη συνέχεια περιλαμβάνουν:

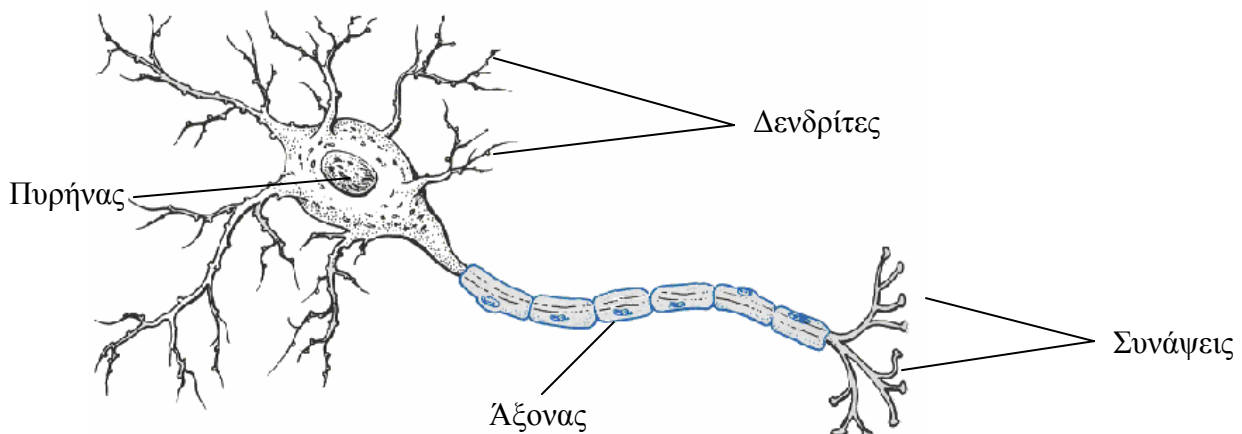
- Τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks-ANN)
- Ασαφής λογική (Fuzzy Logic-FL)

- Έμπειρα συστήματα (Expert Systems-ES)
- Γενετικούς αλγόριθμους (Genetic Algorithms-GAs)

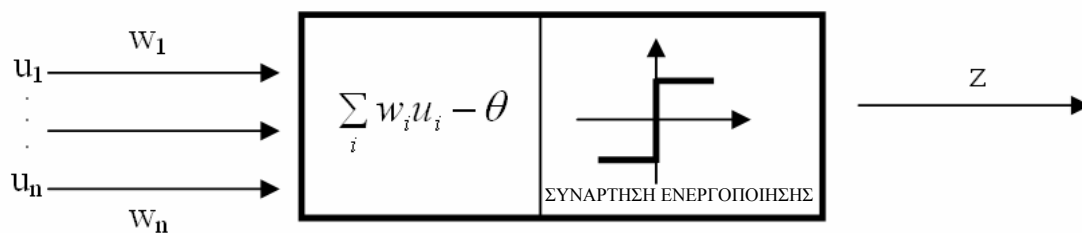
2.1.1 Τεχνητά νευρωνικά δίκτυα

Τα τεχνητά νευρωνικά δίκτυα είναι εμπνευσμένα από τα βιολογικά νευρωνικά δίκτυα. Ο ανθρώπινος εγκέφαλος υπολογίζεται πως αποτελείται από 10 δισεκατομμύρια νευρώνες και από 6000 φορές περισσότερες μεταξύ τους συνάψεις (Haykin ,1994). Ένας νευρώνας είναι συγκριτικά πιο αργός, από μια λογική δίοδο σιλικόνης, το εκπληκτικό ωστόσο πλήθος νευρώνων και συνάψεων καθιστά τον ανθρώπινο εγκέφαλο, ταχύτερο από οποιαδήποτε υπολογιστική μηχανή έχει κατασκευαστεί έως τώρα.

Ένας τεχνητός νευρώνας αποτελεί τη θεμελιώδη μονάδα επεξεργασίας δεδομένων ενός νευρωνικού δικτύου. Η καλύτερη κατανόηση του, γίνεται σε αναλογία με έναν βιολογικό νευρώνα (Σχήματα 2.1 , 2.2).



Σχήμα 2.1 Βιολογικός νευρώνας



Σχήμα 2.2 Τεχνητός νευρώνας

Από το σχήμα 2.2 φαίνεται πως οι συνδέσεις (συνάψεις) w_i μεταφέρουν τα σήματα (ερεθισμούς) u_i μέσα στο νευρώνα. Τα βάρη w_i μπορεί να γίνουν κατανοητά ως η «σημαντικότητα» της αντίστοιχης εισόδου. Μέσα στον νευρώνα υπολογίζεται το σταθμισμένο άθροισμα $w_i \cdot u_i$. Εφόσον το άθροισμα που υπολογίστηκε είναι μεγαλύτερο από μια σταθερά θ , ο νευρώνας θα παράγει ένα σήμα εξόδου z . Οι τιμές της εξόδου, είτε είναι συνεχής είτε είναι δυαδικές. Η μορφή τους εξαρτάται από τη συνάρτηση ενεργοποίησης (ή συνάρτηση μεταφοράς). Στις περισσότερες περιπτώσεις η συνάρτηση ενεργοποίησης είναι τέτοια ώστε η έξοδος του νευρώνα να ανήκει στα διαστήματα $[0,1]$ ή $[-1,1]$.

Σε μαθηματικούς όρους ο νευρώνας μπορεί να περιγραφεί ως εξής:

$$y = \sum_i w_i u_i - \theta \text{ και } z = \psi(y) \quad (2.1)$$

όπου $\psi(\cdot)$ είναι η συνάρτηση ενεργοποίησης.

Τρεις είναι οι βασικές συναρτήσεις ενεργοποίησης που χρησιμοποιούνται στα νευρωνικά δίκτυα:

- Η βηματική συνάρτηση (threshold function)

$$\psi(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2.2)$$

- Η κατά τμήματα γραμμική συνάρτηση (piecewise linear function)

$$\psi(x) = \begin{cases} 1 & x \geq \frac{1}{2} \\ x & -\frac{1}{2} < x < \frac{1}{2} \\ 0 & x < 0 \text{ ή } x \leq -\frac{1}{2} \end{cases} \quad (2.3)$$

- Η σιγμοειδής συνάρτηση (sigmoid function)

$$\psi(x) = \frac{1}{1 + e^{-a(x)}} \quad (2.4)$$

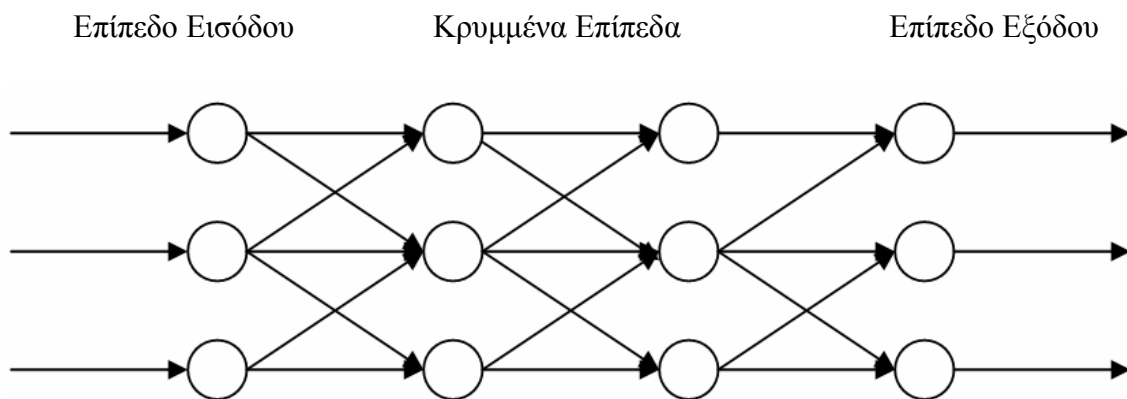
Άλλες συναρτήσεις που χρησιμοποιούνται σπανιότερα είναι η γραμμική και η γκαουσιανή. Η βηματική συνάρτηση (Εξ. 2.2) περιγράφει την ιδιότητα «αληθές/ψευδές». Η Εξ. 2.2 είναι παρόμοια με τη βηματική με μια επιπλέον γραμμική περιοχή. Η σιγμοειδής συνάρτηση (Εξ. 2.4) αποτελεί τη δημοφιλέστερη συνάρτηση ενεργοποίησης. Έχει τη δυνατότητα γραμμικής και μη γραμμικής συμπεριφοράς ενώ είναι συνεχής και παραγωγίσιμη.

Οι νευρώνες πληρούν δύο βασικές ιδιότητες:

- Η λειτουργία τους απαιτεί μόνο τοπική πληροφορία, δηλαδή η συνολική πληροφορία για την εξαγωγή του αποτελέσματος αποκτάται από τις εισόδους του νευρώνα.
- Παράγουν μια και μοναδική έξοδο.

Οι ιδιότητες αυτές, κάνουν εφικτή την παράλληλη επεξεργασία των νευρωνικών δικτύων.

Ένα νευρωνικό δίκτυο αποτελείται από ένα σύνολο κόμβων συνδεδεμένων μεταξύ τους (στο εξής οι νευρώνες θα αναφέρονται ως κόμβοι).



Σχήμα 2.3 Νευρωνικό δίκτυο

Το νευρωνικό δίκτυο του σχήματος 2.3 αποτελείται από το *επίπεδο εισόδου* (input layer) το *επίπεδο εξόδου* (output layer) και δύο *κρυμμένα επίπεδα* (hidden layers). Κάθε επίπεδο περιλαμβάνει τρεις κόμβους, καθένας από τους οποίους είναι συνδεδεμένος με όλους τους κόμβους του προηγούμενου και του επόμενου επιπέδου και μόνο με αυτούς.

Το δίκτυο αυτό ονομάζεται *πλήρως συνδεδεμένο* στην περίπτωση που λείπει κάποια σύνδεση ονομάζεται *μερικώς συνδεδεμένο*. Το πλήθος των κρυμμένων επιπέδων μεταβάλλεται ανάλογα με την εφαρμογή. Η σωστή σχεδίαση του δικτύου (διευθέτηση κόμβων και συνδέσεων) είναι ιδιαίτερα σημαντική. Υπάρχει μια έντονη συσχέτιση μεταξύ του αλγορίθμου εκμάθησης και της τοπολογίας του δικτύου, γεγονός που κάνει τη σχεδίαση πρωταρχικής σημασίας (Haykin, 1994).

Η ικανότητα των νευρωνικών δικτύων να «εκπαιδεύονται», αποτελεί το σημαντικότερο χαρακτηριστικό τους. Με τον όρο εκπαίδευση εννοείται η προσαρμογή των βαρών w_i των διασυνδέσεων, έτσι ώστε το δίκτυο να είναι σε θέση να λύσει κάποιο συγκεκριμένο πρόβλημα. Η εμβάθυνση στην διαδικασία εκπαίδευσης των νευρωνικών δικτύων είναι πέρα από το σκοπό της παρούσας εργασίας. Αναφέρεται επιλεκτικά ο αλγόριθμος «προς τα πίσω διάδοσης» (Back propagation) εκμάθησης ο οποίος αποτέλεσε σταθμό στην εξέλιξη των νευρωνικών δικτύων (Rumelhart, 1986).

Αναλυτική περιγραφή της μεθοδολογίας των τεχνητών νευρωνικών δικτύων δίνεται στα βιβλία των Bishop (1995), Ripley και Hjort (1996), Omidvar και Dayhoff (1997).

2.1.2 Ασαφή λογική

«Δεν υπάρχει τίποτα το ασαφές γύρω από την ασαφή λογική. Στηρίζεται σε ένα ισχυρό θεωρητικό υπόβαθρο και δίνει συγκεκριμένα αποτελέσματα. Επίσης δεν είναι άλλη μια θεωρία πιθανοτήτων. Με τις πιθανότητες μοντελοποιείται η έννοια της τυχαιότητας, ενώ με την ασαφή λογική, η έννοια της αβεβαιότητας» (Βασιλείου, 2006). Όσο η πολυπλοκότητα ενός συστήματος αυξάνει, γίνεται όλο και πιο δύσκολη και από ένα σημείο και μετά ανέφικτη, η επεξήγηση και μοντελοποίηση της συμπεριφοράς του. Σε πολλές περιπτώσεις η πολυπλοκότητα ενός συστήματος είναι τέτοια, που μόνο η ασαφής λογική μπορεί να προσεγγίσει το πρόβλημα. Καθημερινά, ο μέσος άνθρωπος επιλύει πολλά από τα προβλήματά του, ικανοποιητικά κυρίως με εμπειρικό τρόπο, δηλαδή χωρίς να γνωρίζει τα ακριβή δεδομένα που θα χρειαζόταν ένα συμβατικό σύστημα για να λύσει το ίδιο πρόβλημα, τα επιλύει λοιπόν με έναν ασαφή τρόπο. Η ασαφή λογική, σκοπό έχει να μιμηθεί την ανθρώπινη διαδικασία επίλυσης προβλημάτων. Ο όρος «ασαφής» χρησιμοποιήθηκε πρώτα από τον Lofti Zadeh το 1962, όποτε και πρότεινε την ύπαρξη ασαφών συνόλων. Για πολλούς τα ασαφή σύνολα

και η ασαφής λογική είναι σχεδόν συνώνυμα. Τα ασαφή σύνολα σκοπό έχουν την κατηγοριοποίηση εννοιών, που δεν διαχωρίζονται με ξεκάθαρο τρόπο. Τέτοιες έννοιες είναι: «ψηλός - μέτριος - κοντός», «λίγο - αρκετά - πολύ» κ.α.

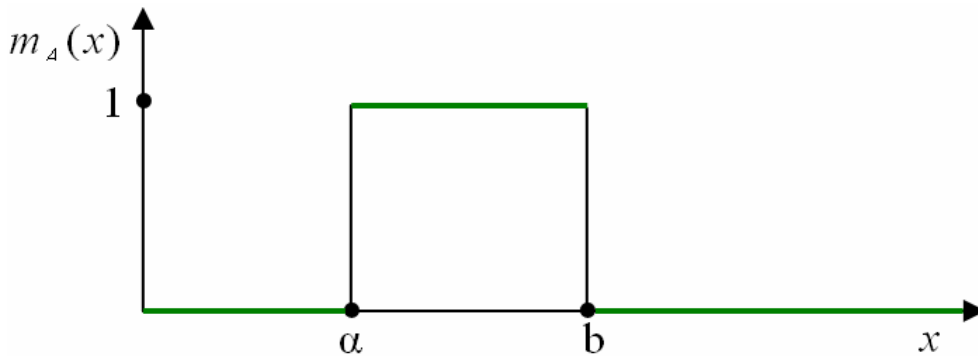
Ένα κλασσικό σύνολο A μπορεί να παρουσιαστεί ως εξής:

$$A = \{x \mid a \leq x \leq b\} \quad (2.5)$$

$$m_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \quad m_A(x) \in \{0,1\} \quad (2.6)$$

όπου $m_A(\cdot)$ ονομάζεται χαρακτηριστική συνάρτηση (ΧΣ)

Η ΧΣ λαμβάνει τιμές ένα ή μηδέν αναλόγως με το αν το x ανήκει ή όχι στο σύνολο A αντίστοιχα. Στο σχήμα 2.4, περιγράφεται η χαρακτηριστική συνάρτηση του κλασσικού συνόλου, που περιγράφεται από τις σχέσεις (2.5), (2.6) .



Σχήμα 2.4 Χαρακτηριστική συνάρτηση κλασσικού συνόλου

Η συνάρτηση $m_A(x)$ παρουσιάζεται με πράσινο χρώμα και καθορίζει τη βεβαιότητα με οποία ένα στοιχείο x ανήκει στο σύνολο A ή ακόμη τυπικότερα ορίζει το βαθμό συμμετοχής του x στο σύνολο A . Η συνάρτηση $m_A(x)$ για ασαφή σύνολα ονομάζεται συνάρτηση συμμετοχής και συμβολίζεται με $\mu_A(x)$. Όπως φαίνεται κάθε στοιχείο x , στο διάστημα $[a,b]$ ανήκει στο σύνολο A με βεβαιότητα ένα, ενώ τα x εκτός του διαστήματος αυτού με απόλυτη βεβαιότητα δεν ανήκουν στο σύνολο. Με αυτό το τρόπο δύναται να αναπαρασταθεί ένα κλασσικό σύνολο με ασαφή τρόπο.

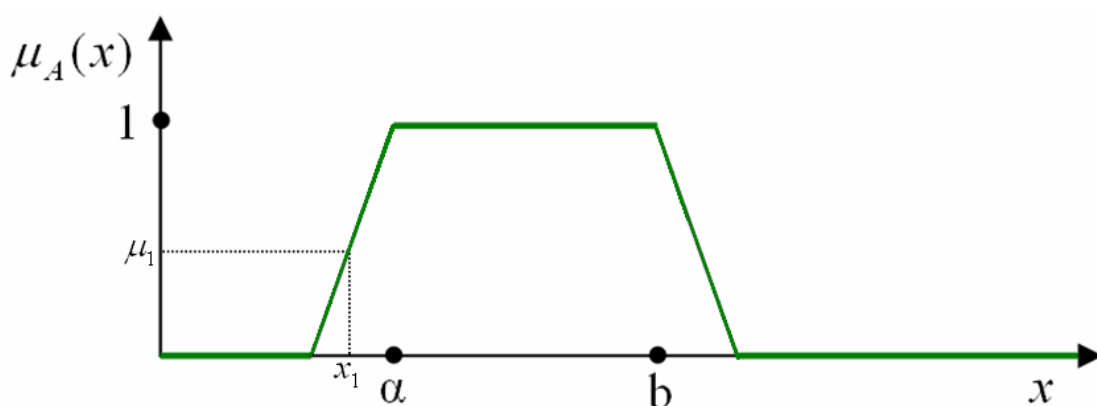
Ένα ασαφές σύνολο B μπορεί να παρουσιαστεί ως εξής:

$$x \text{ «σχεδόν» ανήκει στο } [a,b] \quad (2.7)$$

$$\mu_B(x) \in [0,1] \quad (2.8)$$

Οι σχέσεις (2.7), (2.8) είναι αντίστοιχες των (2.5), (2.6) για ασαφή σύνολα.

Σε ένα ασαφές σύνολο δεν υπάρχει βεβαιότητα για το ποία στοιχεία περιέχει το σύνολο, αλλά η συνάρτηση συμμετοχής, δίνει το βαθμό συμμετοχής στο σύνολο. Στο σχήμα 2.5 φαίνεται η συνάρτηση συμμετοχής ενός ασαφούς συνόλου. Υπάρχουν κάποια στοιχεία, σε τιμές λίγο μικρότερες από το a και σε τιμές λίγο μεγαλύτερες από το b τα οποία δε μπορούμε να αποκλείσουμε από το σύνολο. Έτσι για παράδειγμα το στοιχείο x_1 ανήκει στο σύνολο σε βαθμό $\mu_1 \in [0,1]$.



Σχήμα 2.5 Συνάρτηση συμμετοχής ασαφούς συνόλου

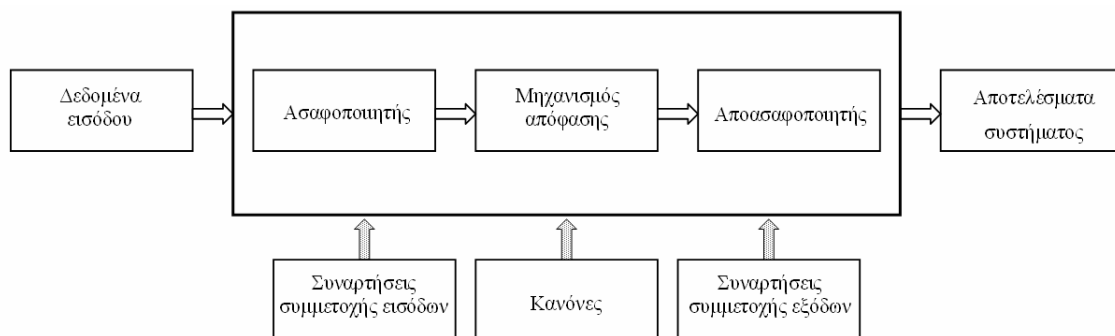
Οι πιο συνηθισμένες συναρτήσεις συμμετοχής είναι οι:

- Τραπεζοειδής
- Τριγωνική
- Gaussian
- Singleton

Τα βασικά τμήματα από τα οποία αποτελείται ένα ασαφές σύστημα είναι:

1. Ασαφοποιητής
2. Μηχανισμός απόφασης
3. Αποασαφοποιητής

Η γενική δομή ενός συστήματος ασαφούς λογικής παρουσιάζεται στο σχήμα 2.6.



Σχήμα 2.6 Γενική δομή ενός ασαφούς συστήματος

Ο ασαφοποιητής (fuzzifier) εκτελεί τη διαδικασία της ασαφοποίησης, δηλαδή της μετατροπής των αρχικών δεδομένων σε ασαφείς αριθμούς, μέσω των συναρτήσεων συμμετοχής. Τα δεδομένα που δέχεται σαν είσοδο είναι μια αριθμητική τιμή που ανήκει στο υπερσύνολο αναφοράς της εισαγόμενης μεταβλητής, ενώ η έξοδος είναι ένας ασαφής βαθμός συμμετοχής.

Ο μηχανισμός απόφασης (inference mechanism) ή αλλιώς μηχανισμός εξαγωγής συμπεράσματος είναι άρρηκτα συνδεδεμένος με μια βάση κανόνων. Η επεξεργασία για την παραγωγή εξόδου (output processing) περιλαμβάνει πράξεις ασαφών συνόλων.

Ο αποασαφοποιητής (defuzzifier) εκτελεί την αντίστροφη διαδικασία από αυτήν που εκτελεί ο ασαφοποιητής, δηλαδή μέσω των συναρτήσεων συμμετοχής μετατρέπει τις ασαφείς εισόδους στους αντιστοίχους αρχικούς τύπους δεδομένων κάθε μεταβλητής.

Αναλυτική περιγραφή της μεθοδολογίας της ασαφούς λογικής δίνεται στα βιβλία των Αναλυτική περιγραφή της FL δίνεται στα βιβλία των Bezdek και Sankan (1992), Altrock (1995).

2.1.3 Έμπειρα συστήματα

Ως έμπειρο σύστημα (ΕΣ), ορίζεται το λογισμικό που συνθέτει τη γνώση των ειδικών πάνω σε ένα συγκεκριμένο τομέα, με σκοπό την επίλυση ή υποβοήθηση της επίλυσης των προβλημάτων που μπορεί να προκύψουν. Τα έμπειρα συστήματα αποσκοπούν είτε στην αντικατάσταση του ειδικού είτε στην υποβοήθηση του. Ως ειδικός-

εμπειρογνώμονας ορίζεται το άτομο που κατέχει εμπειρία και εξειδικευμένη γνώση πάνω σε ένα γνωστικό πεδίο.

Τα ΕΣ επιχειρούν να αναπαραγάγουν ισάξια αποτελεσματικές λύσεις με τους ειδικούς αλλά πολύ ταχύτερα. Η σχεδίαση έμπειρων συστημάτων, απαιτεί έναν «μηχανικό γνώσης». Ως μηχανικός γνώσης (knowledge engineer) ορίζεται ένα πρόσωπο εξειδικευμένο σε θέματα προγραμματισμού και διαδικασιών τεχνητής νοημοσύνης. Η ενεργή συμμετοχή των ειδικών ενός τομέα, στη διαδικασία σχεδίασης είναι απόλυτα επιτακτική. Η διαρκής αλληλεπίδραση εμπειρογνομόνων και μηχανικού γνώσης θα επιχειρήσει να ξεκαθαρίσει στον δεύτερο, τα πιθανά προβλήματα που μπορεί να προκύψουν, τα αίτια τους, τους πιθανούς τρόπους επίλυσης και το σημαντικότερο και δυσκολότερο, τη διαδικασία επιλογής της πλέον ενδεδειγμένης λύσης σε κάθε πρόβλημα. Στο σημείο αυτό φαίνεται μια βασική διαφορά των έμπειρων συστημάτων από την ασαφή λογική. Η ασαφής λογική παράγει τους κανόνες επιλογής της λύσης, ενώ στα έμπειρα συστήματα οι κανόνες και η διαδικασία επιλογής της λύσης υπάρχει εκ κατασκευής.

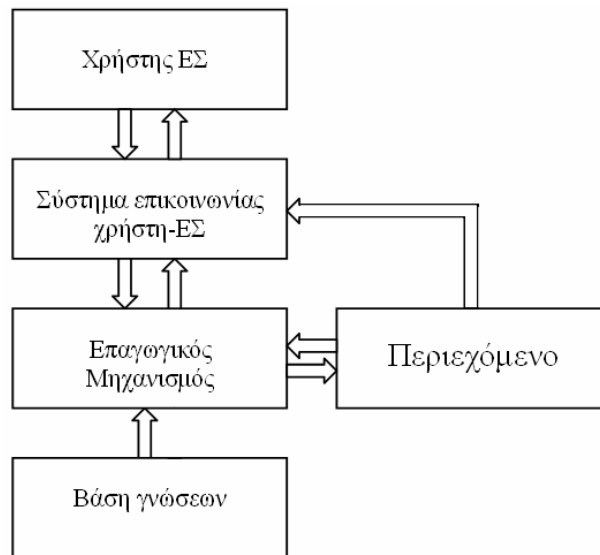
Μερικά από τα σημαντικότερα πλεονεκτήματα των ΕΣ είναι:

- Είναι πάντα διαθέσιμα ενώ ένας ειδικός όχι.
- Συνδυάζουν την γνώση πολλών ειδικών σε μία απόφαση.
- Υπάρχει η δυνατότητα παρακολούθησης (βήμα προς βήμα) της διαδικασίας με την οποία το σύστημα εξήγαγε μια απόφαση, υπάρχει δηλαδή σαφής τεκμηρίωση.
- Εμφανίζουν μειωμένο κόστος, σε περιπτώσεις όπου η ανάγκη για υποστήριξη του ειδικού παρουσιάζει υψηλή συχνότητα.
- Πάντα αποδίδουν το μέγιστο των δυνατοτήτων τους καθώς δεν υπόκεινται σε υποκειμενικούς παράγοντες.

Όσον αφορά τα μειονεκτήματα, είναι δυνατό να αριθμηθούν πολλά: Όλα ωστόσο πηγάζουν από την αδυναμία μοντελοποίησης των εννοιών: «έμπνευση-δημιουργικότητα-κοινή λογική». Σε αυτές τις έννοιες βασίζονται ανακαλύψεις, που θεμελιώνουν τις σύγχρονες επιστήμες. Σε ένα κακώς δομημένο πρόβλημα, η εποπτική, δημιουργική και

ευέλικτη προσέγγισή ενός ειδικού, είναι δύσκολο-αδύνατο να μοντελοποιηθεί για κάθε περίπτωση.

Η γενική δομή ενός έμπειρου συστήματος είναι:



Σχήμα 2.7 Έμπειρο σύστημα

Η **βάση γνώσης** (knowledge base) περιέχει το σύνολο των στοιχείων και ευρετικών κανόνων, που δύναται να χρησιμοποιηθούν προς επίλυση του προβλήματος. Περιέχει γεγονότα (facts) και κανόνες (rules). Τα γεγονότα καθορίζουν το τι είναι γνωστό στο συγκεκριμένο τομέα, σύμφωνα με κάποια θεωρία ή εμπειρία. Οι κανόνες παρέχουν την πληροφορία για τη διαδικασία αξιοποίησης της υπάρχουσας γνώσης και τη διαμόρφωση λύσεων.

Ο **επαγωγικός μηχανισμός** (inference engine) περιλαμβάνει τις τεχνικές εφαρμογής της βάσης γνώσης, πάνω στις πληροφορίες που παρέχει ο χρήστης, σχετικά με το πρόβλημα που αντιμετωπίζει. Ακόμα περιλαμβάνει τη διαδικασία σύγκρισης των εναλλακτικών λύσεων και επιλογής της καλύτερης.

Το **σύστημα επικοινωνίας χρήστη-ΕΣ** (interface) περιέχει τα κατάλληλα εργαλεία για τη διευκόλυνση της αλληλεπίδρασης του χρήστη με το ΕΣ. Μέσω αυτού μεταφέρονται τα στοιχεία του προς επίλυση προβλήματος στο ΕΣ, και τελικώς γνωστοποιείται η επίλυση ή η διαδικασία και τεκμηρίωση επίλυσης στο χρήστη. Επιθυμητά

χαρακτηριστικά για αυτό το σύστημα είναι η ευχρηστία-απλότητα και επάρκεια διαθέσιμων επιλογών αναζήτησης πληροφοριών (querying options).

Το **περιεχόμενο** (context) ή αλλιώς μνήμη εργασίας, χρησιμοποιείται από τον επαγωγικό μηχανισμό για να αποθηκεύει τις απαραίτητες πληροφορίες για την επίλυση του εκάστοτε προβλήματος ή για την αποθήκευση των ενδιάμεσων αποτελεσμάτων στην πορεία προς την τελική λύση. Πρόκειται για μια προσωρινή μνήμη.

Αναλυτική περιγραφή των έμπειρων συστημάτων δίνεται στα βιβλία των Hart (1986) και Hussain (1995)

2.1.4 Γενετικοί αλγόριθμοι

Η γενετικοί αλγόριθμοι (GAs) είναι μια τεχνική παράλληλης αναζήτησης λύσεων, εμπνευσμένη από την εξέλιξη των ειδών και τη φυσική επιλογή. Σκοπός είναι η βελτιστοποίηση μίας η περισσότερων αντικειμενικών συναρτήσεων. Η κωδικοποίηση κάθε πιθανής λύσης γίνεται σε μια δομή που ορίζεται ως χρωμόσωμα (chromosome). Κάθε χρωμόσωμα αποτελείται από γονίδια (genes), ενώ κάθε γονίδιο αποτελείται από μια συγκεκριμένη αλληλουχία στοιχείων, είναι δηλαδή ένα αλληλόμορφο (allelomorph). Ο τύπος των στοιχείων είναι αποτέλεσμα επιλογής του ερευνητή. Συνήθως όμως χρησιμοποιείται μια δυαδική αναπαράσταση. Σε κάθε επανάληψη ο αλγόριθμος μετασχηματίζει τις υπάρχουσες λύσεις, εφαρμόζοντας πάνω τους κάποιους τελεστές. Αποτέλεσμα είναι η παραγωγή νέων λύσεων. Το σύνολο των λύσεων κάθε επανάληψης καλείται γενεά (generation). Κατά κανόνα, τα χρωμοσώματα λόγω της μορφής τους δεν μπορούν να εισαχθούν στην αντικειμενική συνάρτηση. Με μια κατάλληλη αντιστοιχία με την αντικειμενική συνάρτηση, προκύπτει η συνάρτηση προσαρμογής (fitness function-ΣΠ). Η ΣΠ ποσοτικοποιεί την καταλληλότητα κάθε λύσης (χρωμοσώματος), ώστε να είναι δυνατή η σύγκριση των λύσεων.

Οι βασικοί τελεστές ενός απλού γενετικού αλγορίθμου (simple genetic algorithm) είναι οι ακόλουθοι:

- Επιλογή
- Διασταύρωση
- Μετάλλαξη

Η **επιλογή** (selection) καθορίζει ποια χρωμοσώματα θα αξιοποιηθούν για τη δημιουργία της επόμενης γενεάς. Τα χρωμοσώματα που θα επιλεγούν, ονομάζονται γονείς, ενώ τα χρωμοσώματα που θα προκύψουν αναφέρονται ως απόγονοι. Κυρίαρχο ρόλο έχει η συνάρτηση προσαρμογής στην επιλογή των χρωμοσωμάτων προς «επιβίωση». Έχουν προταθεί διάφορες τεχνικές υλοποίησης της φάσης αυτής, όπως επιλογή τουρνουά (tournament selection), μέθοδοι κατάταξης (ranking methods), επιλογή ρουλέτας (roulette wheel selection) κ.α. Ως παράδειγμα της διαδικασίας, θα επεξηγηθεί ενδεικτικά μέσω ενός παραδείγματος η επιλογή ρουλέτας:

Έστω ότι για πέντε χρωμοσώματα προκύπτουν οι ακόλουθες τιμές της συνάρτησης προσαρμογής:

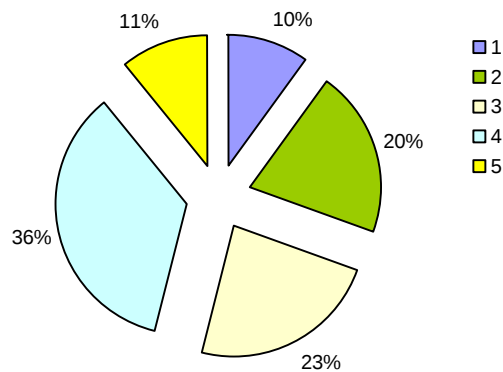
Χρωμόσωμα	Τιμή προσαρμογής f_i
1	101
2	205
3	235
4	357
5	108

Οι πιθανότητες επιλογής υπολογίζονται ως εξής:

$$p_i = \frac{f_i}{\sum_{j=1}^5 f_j} \quad \forall i \quad (2.9)$$

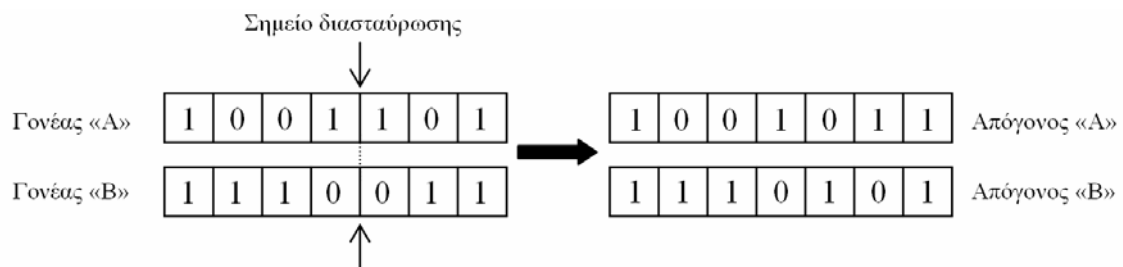
Από τη σχέση (2.9) γίνεται φανερό πως η πιθανότητα να επιλεγεί ένα χρωμόσωμα είναι ευθέως ανάλογη της τιμής προσαρμογής του. Μια γεννήτρια τυχαίων αριθμών προσαρμοσμένη στις πιθανότητες αυτές, θα εκτελέσει τόσες επιλογές όσες το μέγεθος του πληθυσμού.

Για το συγκεκριμένο παράδειγμα, ο τροχός της τύχης διαμορφώνεται ως εξής:

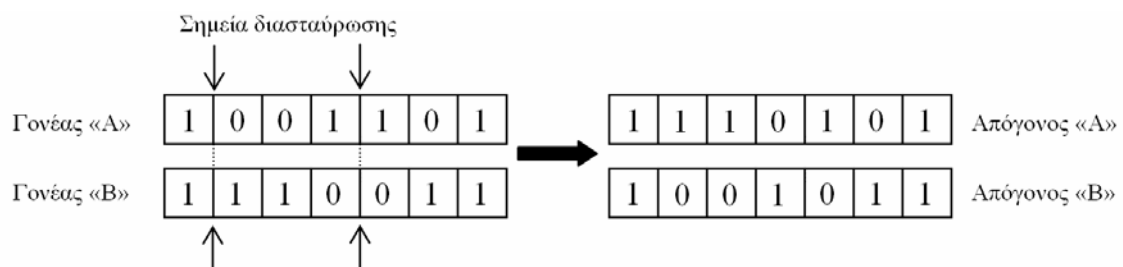


Σχήμα 2.8 Επιλογή «ρουλέτας»

Η **διασταύρωση** (crossover-reproduction) επιδρά πάνω σε δύο τυχαία επιλεγμένα χρωμοσώματα. Η συνηθέστερες τεχνικές που χρησιμοποιείται είναι η «διασταύρωση ενός σημείου» και η «διασταύρωση δύο σημείων». Σύμφωνα με την πρώτη μεθοδολογία, επιλέγεται τυχαία ένα σημείο και για τα δύο χρωμοσώματα (σημείο διασταύρωσης-crossover point). Από το σημείο αυτό και πέρα γίνεται ανταλλαγή των αλληλόμορφων. Στη δεύτερη μεθοδολογία, γίνεται επιλογή δύο σημείων και ανταλλάσσονται τα αλληλόμορφα που βρίσκονται μεταξύ αυτών. Στα σχήματα 2.9 και 2.10 εμφανίζονται παραστατικά οι δυο διαδικασίες:

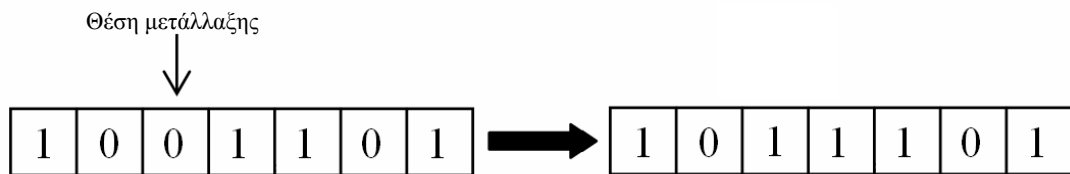


Σχήμα 2.9 Διασταύρωση ενός σημείου



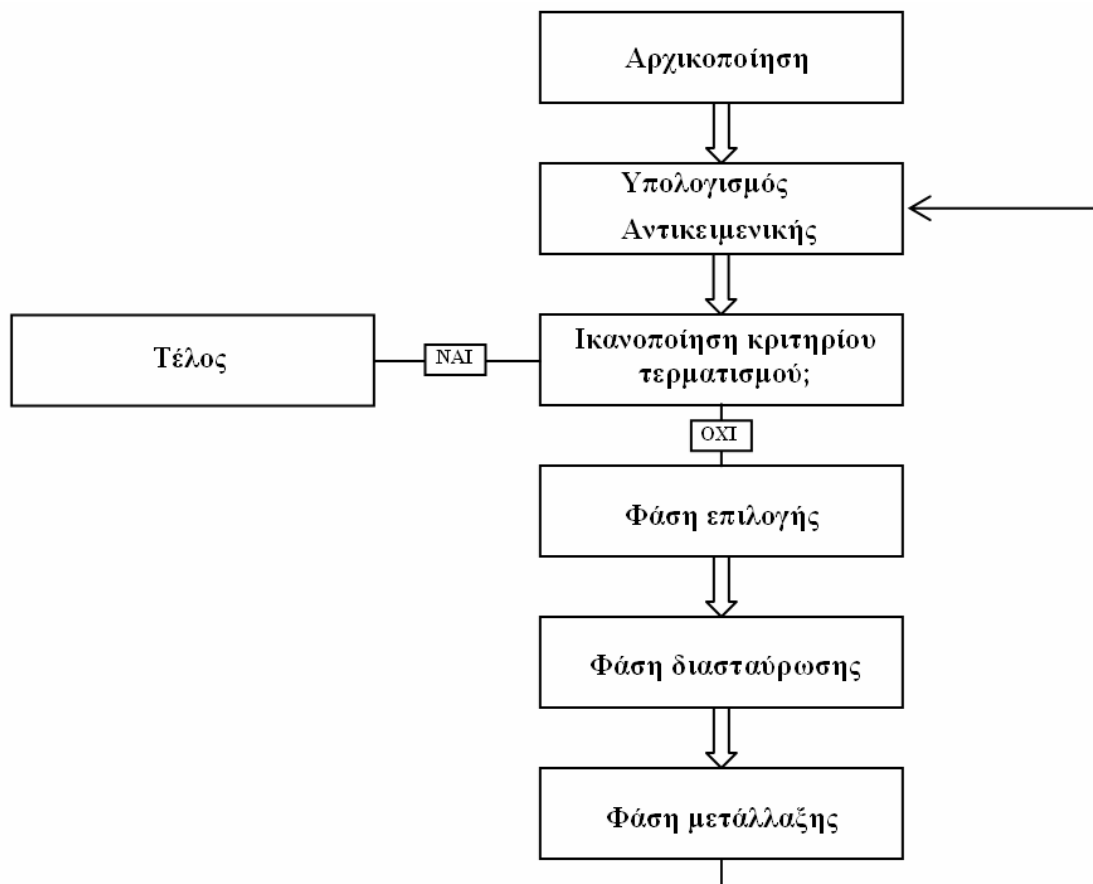
Σχήμα 2.10 Διασταύρωση δύο σημείων

Η **μετάλλαξη** (mutation) αποσκοπεί στο να προσφέρει λύσεις, που δε θα προέκυπταν αλλιώς. Βασική παράμετρος εδώ είναι η πιθανότητα μετάλλαξης ενός χρωμοσώματος, η οποία κατά κανόνα επιλέγεται να είναι μικρή. Η απλούστερη τεχνική είναι η επιλογή (τυχαία) μίας θέσης στο γονίδιο ενός χρωμοσώματος και η αντικατάστασή του με ένα άλλο στοιχείο, όπως φαίνεται στο σχήμα 2.11:



Σχήμα 2.11 Μετάλλαξη

Είναι στην ευχέρεια του ερευνητή η μετάλλαξη σε περισσότερες θέσεις ή η χρήση άλλων στρατηγικών μετάλλαξης. Ολοκληρώνοντας την ενότητα αυτή, παρουσιάζεται ο απλός γενετικός αλγόριθμος (σχήμα 2.12):



Σχήμα 2.12 Απλός γενετικός αλγόριθμος

Σε πρώτη φάση ο αλγόριθμος παράγει τυχαίες αρχικές λύσεις (χρωμοσώματα) ή τροφοδοτείται με «καλές» λύσεις από προηγούμενες επιλύσεις ή πραγματοποιείται ένας συνδυασμός των παραπάνω. Στη συνέχεια γίνεται υπολογισμός της τιμής της συνάρτησης προσαρμογής. Ακολουθούν οι φάσεις της επιλογή, διασταύρωσης και μετάλλαξης. Στη συνέχεια υπολογίζεται και πάλι η τιμή της αντικειμενικής. Εάν το κριτήριο σύγκλισης ικανοποιείται ο αλγόριθμος τερματίζει, ειδάλλως επαναλαμβάνεται η όλη διαδικασία.

Εκτενής περιγραφή των γενετικών αλγορίθμων δίνεται στα βιβλία των Winter (1995), Herrera και Verdegay (1996), Μανιαδάκης και Ζερβάκης (1999).

2.2 Εφαρμογές στη χρηματιστηριακή πρόβλεψη

Στην ενότητα αυτή θα παρουσιαστούν ορισμένες εφαρμογές των μεθοδολογιών που αναλύθηκαν προηγουμένως, στο πεδίο της χρηματιστηριακής πρόβλεψης. Για κάθε μεθοδολογία θα αναφερθεί ένας μικρός αριθμός ερευνητικών προσπαθειών και κατά πόσο αυτές πέτυχαν το σκοπό τους. Θα πρέπει να σημειωθεί πως στην πρόβλεψη τάσεων η επιτυχία ενός μοντέλου αρχίζει να διαφαίνεται σε ποσοστά άνω του 50% ποσοστά 50%. Χαμηλότερα ποσοστά επιτυχίας αποτελούν ένδειξη πως η χρονοσειρά είναι τυχαία.

2.2.1 Χρηματιστηριακή πρόβλεψη με τεχνητά νευρωνικά δίκτυα

Ο Walezak (1999) έκανε χρήση νευρωνικών δικτύων για τη πρόβλεψη του δείκτη Dow Jones αλλά και του γενικού δείκτη του χρηματιστηρίου της Σιγκαπούρης. Η επιλογή δύο τόσο διαφορετικών γενικών δεικτών, αποσκοπεί στο να δείξει πως το μοντέλο λειτουργεί καλά, σε ώριμες και αναπτυσσόμενες αγορές. Το ποσοστό σωστής πρόβλεψης του προτεινόμενου μοντέλου βρισκόταν στο 63%.

Ο Nygren (2004) πραγματοποίησε ημερήσια και εβδομαδιαία πρόβλεψη του δείκτη του χρηματιστηρίου της Σουηδίας με χρήση νευρωνικών δικτύων. Τα δεδομένα που χρησιμοποίησε είναι οι τιμές: κλεισίματος, υψηλότερη, χαμηλότερη και ο όγκος συναλλαγών. Ακόμη χρησιμοποίησε και δέκα δείκτες, ανάμεσα στους οποίους ο Nikkei 225, ο DAX η τιμή του χρυσού και άλλοι. Το μέσο ποσοστό επιτυχούς ημερήσιας πρόβλεψης κυμαινόταν στο 57%, ενώ στην εβδομαδιαία μειώθηκε στο 53%.

Οι Kimoto et al. (1998) πρότειναν ένα σύστημα πρόβλεψης του κατάλληλου χρόνου αγοράς και πώλησης στο χρηματιστήριο του Τόκιο. Το σύστημα χρησιμοποίησε ένα εβδομαδιαίο κινητό μέσο όρο για να μειώσει τις διακυμάνσεις. Για να επιβεβαιωθεί η αποτελεσματικότητα του συστήματος πραγματοποιήθηκε προσομοίωση αγοράς. Τα αποτελέσματα προέκυψαν ικανοποιητικά.

2.2.2 Χρηματιστηριακή πρόβλεψη με ασαφή λογική

Η ασαφής λογική βρίσκει ευρεία αποδοχή και πλήθος εφαρμογών στο τομέα της χρηματιστηριακής πρόβλεψης. Τις περισσότερες φορές χρησιμοποιείται σε συνδυασμό με κάποια άλλη τεχνική όπως νευρωνικά δίκτυα ή γενετικούς αλγορίθμους.

Οι Cheung και Kaymak (2006) ανέπτυξαν ένα σύστημα συναλλαγών βασιζόμενο στην ασαφή λογική, για την πρόβλεψη τάσεων και την παροχή επενδυτικών προτάσεων. Το σύστημα απευθύνεται σε κάποιο επενδυτή ο οποίος ενδιαφέρεται να επενδύσει σε ένα συνδυασμό από ευρωπαϊκά, αμερικάνικα και ιαπωνικά ομόλογα ή συνάλλαγμα. Χρησιμοποιήθηκαν οκτώ οικονομικοί δείκτες και το σύστημα μιμείται το τρόπο κατανόησης των δεικτών από τους ανθρώπους. Η επιλογή των παραμέτρων του ασαφούς συστήματος (συνάρτηση συμμετοχής) βελτιστοποιήθηκε με τη χρήση ενός γενετικού αλγορίθμου. Τα αποτελέσματα ήταν ικανοποιητικά και έδωσαν στις περισσότερες περιπτώσεις αρκετά υψηλότερη απόδοση από την τεχνική αγοράς και διακράτησης.

Οι Nishina et al. (1997) πρότειναν το νευροασαφές σύστημα FINN για την πρόβλεψη της τιμής μετοχών, το οποίο έχει τη δυνατότητα να εξάγει αυτόματα ασαφείς κανόνες της μορφής «Εάν-Τότε» από τα δεδομένα. Το μοντέλο δέχεται δέκα εισόδους. Οι τιμές πρόβλεψης του FINN προσεγγίζουν καλύτερα τις πραγματικές από ότι ένα τεχνητο νευρωνικό δίκτυο (BP).

2.2.3 Χρηματιστηριακή πρόβλεψη με έμπειρα συστήματα

Οι Lee και Jo (1999) ανέπτυξαν ένα έμπειρο σύστημα που επιχειρεί την ερμηνεία γραφημάτων χρονοσειρών. Στόχος είναι η πρόβλεψη των κατάλληλων σημείων αγοράς και πώλησης. Αυτό επιτυγχάνεται με πρόβλεψη τάσεων. Το σύστημα ταξινομεί την

πιθανή κίνηση μιας μετοχής σε πέντε κατηγορίες (πτωτική, ανοδική, στάσιμη, συνέχιση τάσης αναστροφή τάσης) Τα αποτελέσματα ήταν ιδιαίτερα ικανοποιητικά με μέση τιμή επιτυχούς πρόβλεψης 72%.

2.2.4 Χρηματιστηριακή πρόβλεψη με γενετικούς αλγόριθμους

Οι Grosan και Abraham (2005) εφάρμοσαν δυο τύπους γενετικών αλγορίθμων στην πρόβλεψη του δείκτη Nasdaq-100 και του S&P CNX NIFTY. Στη συνέχεια συνέκριναν τους γενετικούς αλγορίθμους με ένα νευρωνικό δίκτυο και ένα νευροασαφές σύστημα. Μέτρα σύγκρισης αποτέλεσαν: η ρίζα του μέσου τετραγωνικού σφάλματος (RMES), το μέγιστο απόλυτο ποσοστό σφάλματος (MAP) και το μέσο απόλυτο ποσοστό σφάλματος (MAPE). Τα αποτελέσματα των γενετικών αλγορίθμων στα παραπάνω μέτρα σύγκρισης βρέθηκαν ότι υπερτερούν σε πολλές περιπτώσεις από αυτά του νευρωνικού δικτύου και του νευροασαφούς συστήματος. Επίσης εξετάστηκε ο συνδυασμός δύο γενετικών αλγορίθμων μέσω ενός αλγορίθμου πολυκριτήριας βελτιστοποίησης. Ο νέος αλγόριθμος απέδιδε καλύτερα αποτελέσματα από κάθε έναν από τους παλιούς χωριστά.

Κεφάλαιο 3: Μεθοδολογικές προσεγγίσεις.

3.1 Μηχανές διανυσμάτων υποστήριξης (SVM)

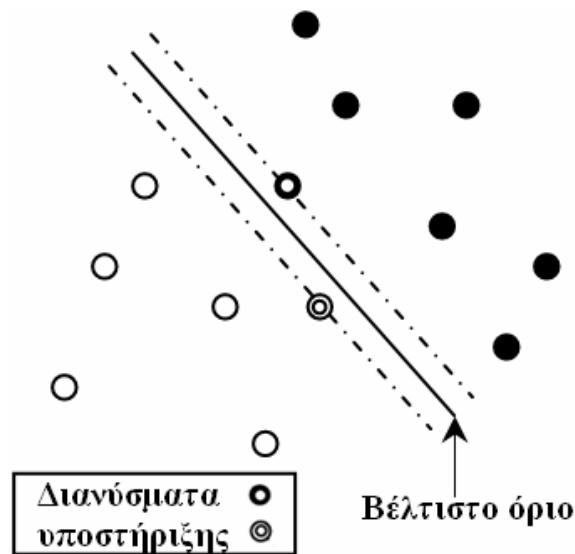
Οι SVM αποτελούν μια σχετικά νέα και καινοτόμα προσέγγιση στο πρόβλημα της ταξινόμησης (Vapnik, 1998). Είναι ένας αλγόριθμος αναγνώρισης προτύπων, που διαφοροποιείται σημαντικά από τους προγενέστερους του. Οι SVM χρησιμοποιούνται κυρίως για δυαδική κατηγοριοποίηση, χωρίς βέβαια να αποκλείεται η χρήση τους για ένταξη δεδομένων σε περισσότερες κατηγορίες, αφού προηγηθεί μια σειρά τροποποιήσεων τους. Οι SVM προσαρμόζονται ανάλογα με την περίπτωση και τη φύση των δεδομένων που είναι επιθυμητό να ταξινομηθούν. Στη συνέχεια θα περιγραφεί η μεθοδολογία ξεκινώντας από τη γραμμικά διαχωρίσιμη περίπτωση και συνεχίζοντας στις γενικότερες περιπτώσεις.

3.1.1 Η γραμμικά διαχωρίσιμη περίπτωση

Έστω ένα σύνολο l παρατηρήσεων εκπαίδευσης $\{x_i, y_i\}$ $i = 1 \dots l$, όπου το x_i είναι διάνυσμα διάστασης d , δηλαδή το x_i ανήκει στον χώρο R^d , ενώ η δυαδική μεταβλητή y_i παίρνει τιμές -1 ή 1 ανάλογα με την κατηγορία στην οποία ανήκει το x_i . Στο σχήμα 3.1 αναζητείται το όριο μεταξύ των δύο κατηγοριών.

Είναι οπτικά ευδιάκριτη η δυνατότητα να διαχωριστούν οι δυο κατηγορίες με μία γραμμή (η γραμμή θα είναι στη διάσταση των δεδομένων δηλαδή στον χώρο R^d). Οι κατηγορίες, στην περίπτωση αυτή είναι οι *λευκοί* και *μαύροι κύκλοι*.

Στην προσέγγιση των SVM, στόχος είναι να βρεθεί το «βέλτιστο» όριο που διαχωρίζει τις δύο ομάδες. Αυτό ορίζεται ως εκείνο το υπερεπίπεδο που μεγιστοποιεί την μεταξύ τους απόσταση. Το όριο αυτό απομονώνει μέγιστα τις δύο ομάδες, και έτσι αναμένεται να αποτελεί το «βέλτιστο ταξινομητή».



Σχήμα 3.1 Γραμμικά διαχωρίσιμη περίπτωση

Τα διανύσματα-παρατηρήσεις που ορίζουν το όριο απόφασης, ονομάζονται *διανύσματα υποστήριξης*. Τα υπόλοιπα στοιχεία του δείγματος εκπαίδευσης δεν επιδρούν στην μορφή του τελικού μοντέλου ταξινόμησης.

Κάθε υπερεπίπεδο στον χώρο $\mathbb{R}^d$ έχει τη μορφή:

$$w^T \cdot x + b = 0 \quad (3.1)$$

Δηλαδή παραμετροποιείται από το διάνυσμα (w) και την σταθερά (b). Όπου το w είναι το κάθετο στο υπερεπίπεδο διάνυσμα. Ένα υπερεπίπεδο που χαρακτηρίζεται από το ζεύγος παραμέτρων (w, b), μπορεί ισοδύναμα να εκφραστεί και από όλα τα ζευγάρια ($\lambda w, \lambda b$), όπου το λ ανήκει $\mathbb{R}^+$. Έτσι ως *κανονικό υπερεπίπεδο* (canonical hyperplane) ορίζεται εκείνο που θέτει σε «συναρτησιακή απόσταση» μεγαλύτερη ή ίση της μονάδας, τα διανύσματα των δύο ομάδων από το υπερεπίπεδο.

Δηλαδή λαμβάνουμε υπόψη εκείνα τα υπερεπίπεδα για τα οποία ισχύει:

$$w^T \cdot x_i + b \geq 1 \quad \text{όταν} \quad y_i = 1 \quad (3.2)$$

$$w^T \cdot x_i + b \leq -1 \quad \text{όταν} \quad y_i = -1 \quad (3.3)$$

ή απλούστερα :

$$y_i(w^T \cdot x_i + b) \geq 1 \quad (3.4)$$

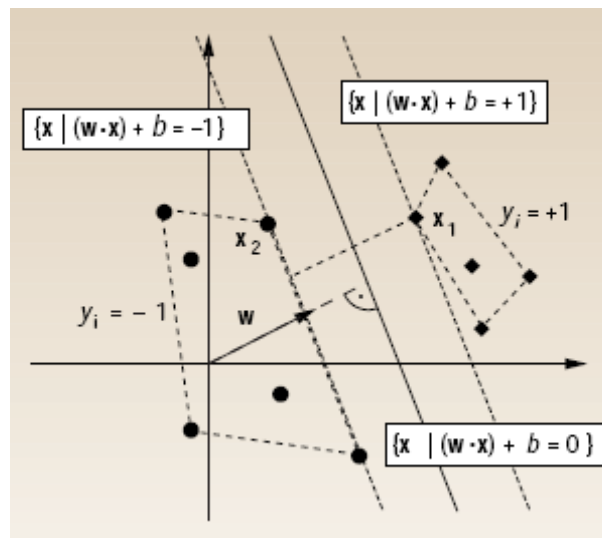
Όλα τα επίπεδα που ικανοποιούν αυτή τη σχέση έχουν «συναρτησιακή απόσταση» μεγαλύτερη ή ίση της μονάδας. Η συναρτησιακή απόσταση δε θα πρέπει να συγχέεται με τη γεωμετρική απόσταση. Για ένα δεδομένο επίπεδο (w, b) όλα τα ζεύγη $(\lambda w, \lambda b)$ με το λ να ανήκει στο $\mathbb{R}^+$, εκφράζουν ακριβώς το ίδιο υπερεπίπεδο, αλλά το καθένα έχει διαφορετική «συναρτησιακή απόσταση» από ένα γνωστό διάνυσμα δεδομένων.

Όπως έχει αναφερθεί οι SVM επιχειρούν να μεγιστοποιήσουν την απόσταση του υπερεπιπέδου ταξινόμησης από τα κοντύτερα σημεία των δύο κλάσεων. Στη συνέχεια υπολογίζεται αυτή η απόσταση.

Έστω λοιπόν ένα διάνυσμα x_1 που ανήκει στο υπερεπίπεδο P_1 : $w^T \cdot x + b = 1$ οπότε $w^T \cdot x_1 + b = 1$. Η γεωμετρική απόσταση του x_1 από το υπερεπίπεδο P_0 : $w^T \cdot x + b = 0$ προκύπτει διαιρώντας με $\|w\|$.

Οπότε:

$$d_1 = \text{distance}(P_0, x_1) = \frac{|w^T \cdot x_1 + b|}{\|w\|_2} = \frac{1}{\|w\|_2} \quad (3.6)$$



Σχήμα 3.2 Διαχωρίσιμη περίπτωση. Πηγή: Hearst, 1998

Όμοια έστω ένα δεύτερο διάνυσμα x_2 που ανήκει στο υπερεπίπεδο P_{-1} :

$w^T \cdot x + b = -1$, οπότε $w^T \cdot x_2 + b = -1$. Η την γεωμετρική απόσταση του x_2 από το υπερεπίπεδο $w^T \cdot x + b = 0$, είναι:

$$d_2 = \text{distance}(P_{-1}, x_2) = \frac{|w \cdot x_2 + b|}{\|w\|_2} = \frac{|-1|}{\|w\|_2} = \frac{1}{\|w\|_2}$$

Τα δύο επίπεδα είναι παράλληλα, επομένως οπότε η απόσταση μεταξύ τους είναι:

οπότε:

$$d_1 + d_2 = \frac{2}{\|w\|_2} = \frac{1}{\frac{\|w\|_2}{2}} \quad (3.7)$$

Η απόσταση αυτή θα πρέπει να μεγιστοποιηθεί. Για να γίνει αυτό αρκεί να ελαχιστοποιηθεί το $\frac{\|w\|_2}{2} = \frac{w^T w}{2}$. Τελικά, το πρόβλημα για τη γραμμικά διαχωρίσιμη περίπτωση, διαμορφώνεται ως εξής:

min	$\frac{1}{2} w^T w$	(3.8)
υ.π.	$y_i(w^T \cdot x_i + b) \geq 1$	(3.9)

Για το πρόβλημα αυτό, διαμορφώνεται η Λαγκραντσιανή L :

$$L = \frac{1}{2} w^T w - \sum_i \alpha_i [y_i(w^T \cdot x_i + b) - 1] \quad (3.10)$$

Όπου $\alpha_i \geq 0$ οι συντελεστές Lagrange.

Στη συνέχεια θα υπολογιστούν και θα μηδενιστούν οι μερικές παράγωγοι:

$$\frac{\partial L}{\partial w} = w - \sum_i \alpha_i y_i x_i = 0 \quad \Rightarrow \quad w = \sum_i \alpha_i y_i x_i \quad (3.11)$$

$$\frac{\partial L}{\partial b} = -\sum_i \alpha_i y_i = 0 \Rightarrow \sum_i \alpha_i y_i = 0 \quad \Rightarrow \quad \sum_i \alpha_i y_i = 0 \quad (3.12)$$

Αν γίνουν οι επιμεριστικές στην σχέση (10) προκύπτει μια άλλη μορφή της:

$$L = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_i \alpha_i y_i \mathbf{w}^T \mathbf{x}_i - b \sum_i \alpha_i y_i + \sum_i \alpha_i \quad (3.13)$$

Αν αντικαταστήσουμε τις εξισώσεις (11) και (12) στην (13), θα προκύψει:

$$L = \frac{1}{2} (\sum_i \alpha_i y_i \mathbf{x}_i)^T \sum_j \alpha_j y_j \mathbf{x}_j - \sum_i \alpha_i y_i (\sum_j \alpha_j y_j \mathbf{x}_j)^T \mathbf{x}_i + \sum_i \alpha_i \quad (3.14)$$

μετά από πράξεις η (14) παίρνει τη μορφή:

$$L = -\frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j + \sum_i \alpha_i$$

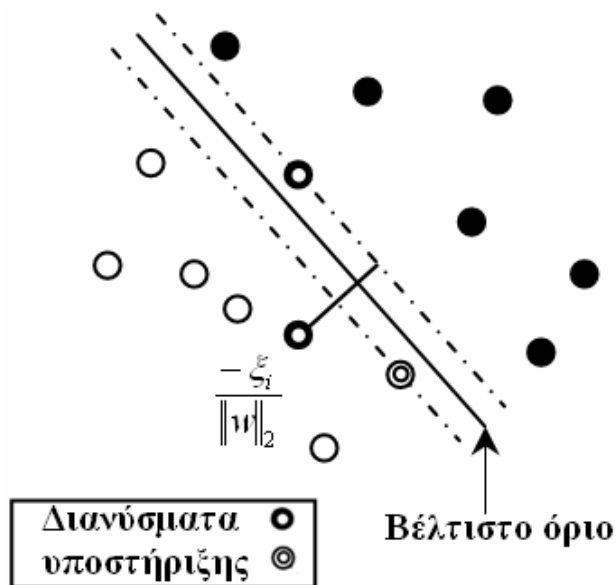
Το τελικό πρόβλημα τετραγωνικού προγραμματισμού διαμορφώνεται ως εξής:

$\begin{aligned} \max \quad & L = -\frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j + \sum_i \alpha_i \\ \text{υ.π.} \quad & \sum_i \alpha_i y_i = 0 \quad \quad \alpha_i \geq 0 \quad \forall i \end{aligned}$	(3.15)
--	--------

3.1.2 Η Γραμμικά Μη Διαχωρίσιμη Περίπτωση

Όσα προαναφέρθηκαν, μπορούν να εφαρμοστούν μόνο στις περιπτώσεις όπου οι ομάδες δεδομένων είναι γραμμικά διαχωρίσιμες.

Στο σχήμα 3.3, φαίνονται δυο κλάσεις, που δεν δύναται να διαχωριστούν με γραμμικό τρόπο. Δεν μπορεί λοιπόν με αυτά που παρουσιάστηκαν έως τώρα, να βρεθεί κάποιο υπερεπίπεδο που να προσδιορίζει το διαχωριστικό όριο των δύο ομάδων.



Σχήμα 3.3 Γραμμικά μη διαχωρίσιμη περίπτωση.

Αν όμως χαλαρωθεί ο περιορισμός (9) ώστε να επιτραπούν περιπτώσεις λανθασμένης ταξινόμησης, τότε μπορεί να βρεθεί το βέλτιστο υπερεπίπεδο ταξινόμησης, ακολουθώντας μια μεθοδολογία παρόμοια με αυτή που αναλύθηκε προηγούμενα. Ο περιορισμός (9) μπορεί να χαλαρωθεί με την εισαγωγή μιας μεταβλητής απόκλισης (slack variable) ξ_i :

$$y_i(w^T \cdot x_i + b) \geq 1 - \xi_i \quad (3.16)$$

Η μεταβλητή ξ_i είναι θετική και υποδηλώνει πως μια παρατήρηση, μπορεί πλέον να παραβεί τον περιορισμό.

Έτσι, το **πρωτεύον** πρόβλημα μπορεί να διαμορφωθεί πλέον ως εξής:

min	$\frac{1}{2} w^T w + C \sum_i \xi_i$	(3.17)
υ.π.	$y_i(w^T \cdot x_i + b) \geq 1 - \xi_i \quad \forall i$	
	$\xi_i \geq 0 \quad \forall i$	

Η παράμετρος C υποδηλώνει την βαρύτητα που δίνεται στην αποφυγή λανθασμένων κατατάξεων. Αποτελεί λοιπόν, έναν συντελεστή τιμωρίας για αυτό τον τύπο λάθους.

Η λαγκραντσιανή του πρωτεύοντος είναι:

$$L_P = \frac{1}{2} w w^T + C \sum_i \xi_i - \sum_i a_i [y_i (w^T x_i + b) - 1 + \xi_i] - \sum_i \mu_i \xi_i \quad (3.18)$$

$$\frac{\partial L_P}{\partial w} = w - \sum_i \alpha_i y_i x_i = 0 \quad \Rightarrow \quad w = \sum_i \alpha_i y_i x_i \quad (3.19)$$

$$\frac{\partial L}{\partial b} = -\sum_i \alpha_i y_i = 0 \quad \Rightarrow \quad \sum_i \alpha_i y_i = 0 \quad (3.20)$$

$$\frac{\partial L_P}{\partial \xi_i} = C - a_i - \mu_i = 0 \quad \Rightarrow \quad C = a_i + \mu_i \quad (3.21)$$

$$a_i \geq 0 \quad \forall i \quad \xi_i \geq 0 \quad \forall i \quad \mu_i \geq 0 \quad \forall i \quad (3.22)$$

Αντικαθιστώντας τις σχέσεις (19),(20),(21) στην 18, το πρόβλημα τετραγωνικού προγραμματισμού γίνεται:

max	$L = -\frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j + \sum_i \alpha_i$	(3.23)
υ.π.	$\sum_i \alpha_i y_i = 0$	
	$0 \leq \alpha_i \leq C \quad \forall i$	

Μπορούν να παρατηρηθούν τα εξής:

- (α) Δεν εμφανίζεται η μεταβλητή ξ_i ούτε ο λαγκρανσιανός συντελεστής της μ_i .
- (β) Αν συγκριθούν οι εξισώσεις της γραμμικά διαχωρίσιμης περίπτωσης (15) με τις εξισώσεις της μη γραμμικά διαχωρίσιμης περίπτωσης (18), είναι προφανές πως διαφέρουν μόνο ως προς το άνω όριο των δυϊκών συντελεστών α_i .

3.1.3 Μη Γραμμικές Μηχανές Διανυσμάτων Υποστήριξης

Μέχρι το σημείο αυτό, έχει παρουσιαστεί, ο τρόπος με τον οποίο οι SVM, μπορούν να διαχωρίσουν δυο κλάσεις, τα όρια των οποίων είναι ή τείνουν να είναι γραμμικά. Στις

περιπτώσεις όπου τα όρια είναι έντονα μη γραμμικά οι μη γραμμικές SVM αποτελούν λύση.

Στην προσέγγιση αυτή, οι SVM επιλύουν το πρόβλημα, αναπαριστώντας τα δεδομένα εισόδου σε ένα χώρο υψηλότερων διαστάσεων, όπου και είναι γραμμικά διαχωρίσιμα. Το γραμμικό όριο στο νέο χώρο $\mathcal{H}$, αντιστοιχεί σε ένα μη γραμμικό όριο, στον αρχικό χώρο διάστασης $\mathcal{L}$, όπου βρίσκονται τα δεδομένα ($\mathcal{L} \equiv \mathcal{R}^d$). Έτσι επιτυγχάνεται η κατασκευή ενός μη γραμμικού συνόρου.

Από υπολογιστική πλευρά, αυτή η αύξηση των διαστάσεων δεν δημιουργεί πρόβλημα στις SVM. Το 1992 οι Boser, Guyon και Vapnik, έδειξαν πως ένα τέχνασμα που βασίζεται στην μεθοδολογία του Aizerman et al. (1964), μπορεί να δώσει τη λύση. Το τέχνασμα βασίζεται στη διαπίστωση πως στις εξισώσεις (23) τα δεδομένα συσχετίζονται μόνο μέσω εσωτερικών γινομένων.

Έστω Φ η συνάρτηση που μεταφέρει τα δεδομένα από το χώρο $\mathcal{L}$ στον $\mathcal{H}$:

$$\Phi : \mathcal{L} \mapsto \mathcal{H}.$$

Η βελτιστοποίηση θα γίνει λοιπόν στο χώρο $\mathcal{H}$. Και εκεί τα δεδομένα θα συσχετίζονται μόνο με πράξη εσωτερικού γινομένου της ακόλουθης μορφής :

$$\Phi(x_i) \cdot \Phi(x_j)$$

Αν γνωρίζαμε μια συνάρτηση πυρήνα (Kernel function) τέτοια ώστε:

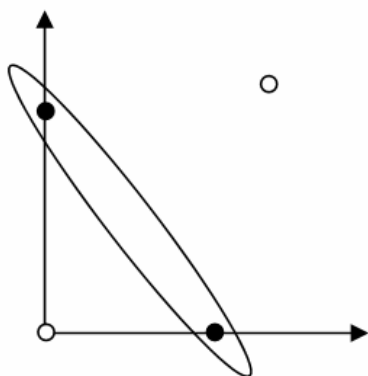
$$K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$$

δεν θα χρειαζόταν η επακριβής γνώση της Φ . Με τον τρόπο αυτό, αποφεύγονται οι πράξεις στον υψηλής διάστασης χώρο, οι οποίες θα αύξαναν υπερβολικά τον υπολογιστικό φόρτο.

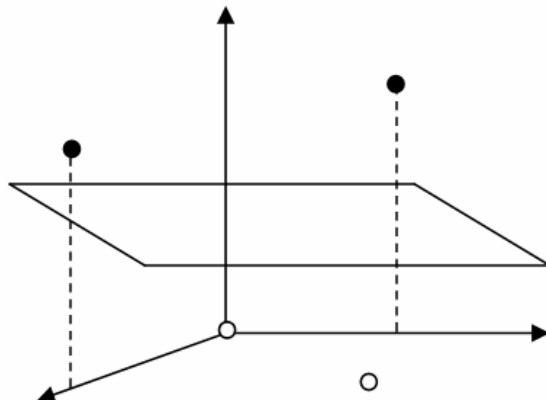
Το σύνορο στον χώρο $\mathcal{H}$ έχει τη μορφή:

$$w^T \Phi(x) + b = 0$$

Για την καλύτερη κατανόηση της λογικής της συνάρτησης πυρήνα, παρουσιάζεται ένα γραφικό παράδειγμα.



Σχήμα (α)



Σχήμα (β)

Σχήμα 3.4 Παράδειγμα αναπαράστασης δεδομένων σε υψηλότερη διάσταση με χρήση συνάρτησης πυρήνα

Στο Σχήμα 3.4α, φαίνεται η τοποθέτηση τεσσάρων σημείων στο δυσδιάστατο χώρο. Με την διάταξη που έχουν, ο γραμμικός διαχωρισμός των άσπρων και μαύρων κουκκίδων δε φαίνεται εφικτός. Αν όμως τα δεδομένα αναχθούν μέσω μιας συνάρτησης πυρήνα στην τρίτη διάσταση, η νέα τους διάταξη θα έχει όπως φαίνεται στο Σχήμα 3.4β. Τα δεδομένα πλέον είναι γραμμικά διαχωρίσιμα.

Γεννιέται το ερώτημα του γιατί χρησιμοποιείται η φράση «γραμμικά διαχωρίσιμα», αφού στην τρίτη διάσταση ένα επίπεδο, είναι που τελεί των διαχωρισμό των κλάσεων. Όπως η γραμμή χωρίζει ένα επίπεδο σε δύο ημιεπίπεδα (δηλαδή δύο κλάσεις), έτσι και ο τρισδιάστατος χώρος διχοτομείται με ένα επίπεδο. Γενικά ένας χώρος R^n διάστασης, διχοτομείται από ένα υπερεπίπεδο στον χώρο R^{n-1} .

3.1.4 Πλεονεκτήματα Μηχανών Διανυσμάτων Υποστήριξης

Ακολουθεί η παράθεση των κυριότερων πλεονεκτημάτων των SVM:

- Βασίζονται σε πολύ απλές και ξεκάθαρες ιδέες από τη θεωρία στατιστικής μάθησης,
- Έχουν πολύ καλή ικανότητα γενίκευσης, καθώς βασίζονται στην αρχή ελαχιστοποίησης του δομικού σφάλματος και όχι της μέσης τετραγωνικής απόκλισης από τα δεδομένα εκμάθησης. Με τον τρόπο αυτό οι μηχανές

αποκτούν καλή ικανότητα γενίκευσης, ενώ ξεπερνούν σε σημαντικό βαθμό το πρόβλημα υπερπροσαρμογής στα δεδομένα (overfitting).

- Η επίλυση ανάγεται στην επίλυση ενός προβλήματος τετραγωνικού κυρτού προγραμματισμού, το οποίο έχει μία και μοναδική βέλτιστη λύση.
- Η μέθοδος δεν απαιτεί τη γνώση της στατιστικής κατανομής των δεδομένων .
- Μπορούν να παράγουν περίπλοκα μη γραμμικά μοντέλα, που έχουν συγκεκριμένη συναρτησιακή διατύπωση.
- Έχουν ισχυρό θεωρητικό υπόβαθρο και έτσι μπορούμε να ακολουθήσουμε την διαδικασία βήμα-βήμα.
- Σε πληθώρα πραγματικών εφαρμογών, έχουν επιδείξει ισάξια η καλύτερη επίδοση συγκριτικά με άλλες ανταγωνιστικές μεθόδους.
- Υπάρχει η δυνατότητα υπολογισμού της διάστασης VC που αποτελεί ένα μέτρο της ικανότητας γενίκευσης (τα νευρωνικά δίκτυα για παράδειγμα , δεν διαθέτουν ένα τέτοιο μέτρο).
- Το πλήθος των παραμέτρων που απαιτούν ρύθμιση στις SVM είναι σημαντικά μικρότερο από .το αντίστοιχο ανταγωνιστικών μεθοδολογιών.

3.1.5 Εφαρμογές των SVM

Οι SVM έχουν βρει εφαρμογή σε πλήθος επιστημονικών πεδίων και προβλημάτων. Στη συνέχεια θα αναφερθούν ενδεικτικά κάποιες εφαρμογές. Στο οικονομικό πεδίο υπάρχουν εφαρμογές, στη χρηματιστηριακή πρόβλεψη (Kim, 2003), Στη βιολογία έχουν εφαρμοστεί στην κατηγοριοποίηση γονιδίων και ασθενών με βάση χαρακτηριστικά των γονιδίων τους (Jianmin et al, 2007). Στην ιατρική έχουν χρησιμοποιηθεί, στη διάγνωση καρκίνου (Furey et al, 2000; Statnikov et al, 2003) Ακόμη υπάρχουν εφαρμογές, στο πρόβλημα αναγνώρισης χειρόγραφου κειμένου (Hidetoshi et al, 2007), στο πρόβλημα παλινδρόμησης (Drucker et al, 1997 και Vapnik et al, 1997), στο πρόβλημα εντοπισμού προσώπων (Osuna et al, 1997), στην αναγνώριση προσώπων (Casiraghi et al, 2005), στην κατηγοριοποίηση κειμένου. (Roobaert και Hulle, 1999), στην αναγνώριση φωνής (Smith και Gales et al, 2002) ή

αντικειμένων (Heisele et al, 2002). Η συνεχής έρευνα, οδηγεί διαρκώς στην εύρεση περισσότερων εφαρμογών των SVM σε πραγματικά προβλήματα.

3.2 Προσαρμοστικά νευροασαφή συστήματα εξαγωγής συμπερασμάτων (ANFIS)

3.2.1 Γενικά

Το σύστημα ANFIS, δημιουργήθηκε από τον Jang (1993) και ανήκει στα προσαρμοστικά δίκτυα και παρουσιάζει πολλές ομοιότητες με τα νευρωνικά δίκτυα. Προσαρμοστικό ονομάζεται ένα δίκτυο που εκτός από σταθερούς κόμβους περιέχει και κόμβους, οι παράμετροι των οποίων προσαρμόζονται κατά τη φάση της εκπαίδευσης του δικτύου. Αποτελεί έναν συνδυασμό νευρωνικών δικτύων (Artificial Neural Networks-ANN) και ασαφούς λογικής (Fuzzy logic-FL). Διαχειρίζεται τόσο αριθμητικά όσο και ποιοτικά χαρακτηριστικά. Η εξαγωγή γνώσης από το σύστημα είναι άμεσα εφικτή εν αντίθεση με τα νευρωνικά δίκτυα που αποτελούν ένα μαύρο κουτί (black box). Κάτι τέτοιο είναι εφικτό καθώς τα νευροασαφή συστήματα παράγουν κανόνες, οι οποίοι γίνονται άμεσα κατανοητοί. Το ANFIS, συνδυάζει τα πλεονεκτήματα των συστατικών του (ANN & FL). Έτσι παρέχει ισχυρό υπόβαθρο για την αναπαράσταση της γνώσης, ενώ τα νευρωνικά δίκτυα, παρέχουν δυνατότητες μάθησης και δημιουργίας αποδοτικών υπολογιστικών εφαρμογών με χρήση H/Y.

Το ANFIS δέχεται ως είσοδο, ένα σύνολο δεδομένων εισόδου/εξόδου και κατασκευάζει ένα ασαφές σύστημα εξαγωγής συμπερασμάτων (Fuzzy Inference System-FIS). Οι παράμετροι των συναρτήσεων συμμετοχής του FIS, ρυθμίζονται είτε με τη χρήση ενός ανεξάρτητου αλγορίθμου οπισθόδρομης μάθησης (Back propagation-BP) είτε με έναν συνδυασμό του BP με τη μέθοδο ελαχίστων τετραγώνων. Το γεγονός αυτό επιτρέπει στο ασαφές σύστημα να μαθαίνει από τα δεδομένα που μοντελοποιεί.

3.2.2 Ασαφή συστήματα εξαγωγής συμπερασμάτων (FIS)

Ένα FIS αντιστοιχεί κατά σειρά:

-Τα δεδομένα εισόδου στις συναρτήσεις συμμετοχής εισόδου.

- Τις συναρτήσεις συμμετοχής εισόδου σε κανόνες.
- Τους κανόνες σε ένα σετ από χαρακτηριστικά εξόδου.
- Τα χαρακτηριστικά εξόδου στις συναρτήσεις συμμετοχής εξόδου.
- Τις συναρτήσεις συμμετοχής εξόδου σε μια ενιαία τιμή εξόδου ή σε μία απόφαση που σχετίζεται με την τιμή εξόδου.

Η διάταξη ενός FIS φαίνεται στην εικόνα (2.5).

Έστω ένα FIS με δύο εισόδους (x,y) και μία έξοδο z . Τα συστήματα Sugeno μηδενικού βαθμού παράγουν ασαφείς κανόνες της μορφής:

$$\text{Αν } x \text{ ανήκει στο } A \text{ και } y \text{ ανήκει στο } B, \text{ τότε } z = k$$

όπου A, B ασαφή σύνολα και k μια σταθερά

Τα συστήματα Sugeno πρώτου βαθμού παράγουν ασαφείς κανόνες της μορφής:

$$\text{Αν } x \text{ ανήκει στο } C \text{ και } y \text{ ανήκει στο } D, \text{ τότε } z = px + qy + r$$

όπου C, D ασαφή σύνολα και p, q, r σταθερές.

3.2.3 Εκπαίδευση του ANFIS

Σε πρώτη φάση οι παράμετροι των συναρτήσεων συμμετοχής αρχικοποιούνται σε κάποια επιθυμητή ή τυχαία τιμή. Στη συνέχεια το σύστημα τροφοδοτείται με τα σετ δεδομένα εισόδου/εξόδου και εκπαιδεύεται πάνω σε αυτά, τροποποιώντας κατάλληλα τις παραμέτρους του δικτύου μέχρι να ικανοποιηθεί κάποιο κριτήριο τερματισμού. Θα πρέπει να σημειωθεί πως το σύστημα ANFIS χρησιμοποιεί τη μέθοδο ασαφούς συμπεράσματος Takagi-Sugeno-Kang που πρωτοπαρουσιάστηκε το 1985. Στο εξής η μέθοδος για συντομία θα λέγεται Sugeno. Τα μοντέλα που είναι διαθέσιμα στο εμπορικό λογισμικό Matlab® που χρησιμοποιείται, είναι μοντέλα Sugeno είτε μηδενικού είτε πρώτου βαθμού.

Το σύνολο των δεδομένων μπορεί να χωριστεί σε τρία τμήματα ανάλογα με την χρήση τους από το ANFIS:

- Δεδομένα εκπαίδευσης μοντέλου (training data)
- Δεδομένα επαλήθευσης μοντέλου (validation data)

- Δεδομένα ελέγχου εκπαίδευσης (check data)

Τα δεδομένα εκπαίδευσης έχουν εξηγηθεί στη παράγραφο (4.2). Τα δεδομένα επαλήθευσης έχουν το ίδιο ακριβώς ρόλο που έχουν τα «δεδομένα ελέγχου» στα SVM, απλά διαφοροποιείται η ορολογία, δηλαδή είναι σετ δεδομένων, στα οποία δεν έχει εκπαιδευτεί το μοντέλο και ζητείται να τα εκτιμήσει. Κατά τον τρόπο αυτό επαληθεύεται η αποτελεσματική ή όχι ικανότητα πρόβλεψης. Τα δεδομένα ελέγχου εκπαίδευσης είναι αποτέλεσμα ειδικής προσέγγισης των ANFIS και θα επεξηγηθούν παρακάτω.

Εν γένει η εκπαίδευση και η προβλεπτική ικανότητα του μοντέλου θα είναι ικανοποιητική, εφόσον τα δεδομένα είναι αντιπροσωπευτικά της διαδικασίας που αναπαριστούν. Κάτι τέτοιο δεν ισχύει σε κάθε περίπτωση, ειδικά όταν πρόκειται για μια περίπλοκη διαδικασία, οπότε τα δεδομένα εμπεριέχουν ανακρίβεια και θόρυβο. Τέτοιες περιπτώσεις επιχειρεί να εντοπίσει μια ενσωματωμένη διαδικασία του ANFIS που στο εξής θα αποκαλείται *έλεγχος εκπαίδευσης*. Σκοπός του ελέγχου εκπαίδευσης είναι η αποφυγή της υπερπροσαρμογής του μοντέλου στα δεδομένα (overfitting). Η ιδέα στην οποία βασίζεται είναι η εξής: όσο το μοντέλο εκπαιδεύεται χωρίς να έχει υποστεί υπερπροσαρμογή, το σφάλμα επί των δεδομένων ελέγχου εκπαίδευσης μειώνεται, από το σημείο που το μοντέλο ξεκινά να υπερπροσαρμόζεται, παρατηρείται μια απότομη αύξηση του σφάλματος αυτού. Ένα μεγάλο σφάλμα ελέγχου εκπαίδευσης υποδεικνύει την ανάγκη λήψης δράσης, όπως μείωση του μεγέθους του δείγματος εκπαίδευσης, αλλαγή του τύπου των συναρτήσεων συμμετοχής κ.α. Στο σημείο αυτό ως δεδομένα ελέγχου εκπαίδευσης, ορίζεται το τμήμα εκείνο των συνολικών δεδομένων, που αξιοποιούνται για την τέλεση της διαδικασίας ελέγχου εκπαίδευσης.

Μια σημαντική παράμετρος στον έλεγχο εκπαίδευσης, είναι η σωστή επιλογή των δεδομένων ελέγχου εκπαίδευσης. Η επιλογή καθίσταται δύσκολη λόγω των δύο βασικών χαρακτηριστικών που πρέπει να έχουν τα δεδομένα αυτά: (1) θα πρέπει να είναι αντιπροσωπευτικά των δεδομένων που είναι επιθυμητό το μοντέλο να αναπαράγει και (2) θα πρέπει να διαφοροποιούνται σε κάποιο βαθμό από τα δεδομένα που εκπαιδεύτηκε το μοντέλο. Διαφορετικά η διαδικασία ελέγχου εκπαίδευσης μπορεί να καταστεί αναποτελεσματική.

Το σύστημα ANFIS υποστηρίζει μόνο συστήματα τύπου Sugeno (μηδενικού ή πρώτου βαθμού στο Matlab FL toolbox), γεγονός που δημιουργεί περιορισμούς για τα μοντελοποιούμενα συστήματα, όπως:

- Έχουν μία έξοδο, που λαμβάνεται με τη μέθοδο αποασαφοποίησης σταθμισμένου μέσου (weighted average defuzzification). Όλες οι συναρτήσεις συμμετοχής εξόδου πρέπει να είναι ίδιου τύπου γραμμικές ή σταθερές.
- Δεν επιτρέπεται η από κοινού χρήση κανόνων (διαφορετικοί κανόνες δεν μοιράζονται την ίδια συνάρτηση συμμετοχής εξόδου). Δηλαδή, το πλήθος των συναρτήσεων συμμετοχής εξόδου είναι ίσο με το πλήθος των κανόνων.
- Οι συναρτήσεις συμμετοχής (όπως και άλλες ρυθμίσεις) που είναι διαθέσιμες προς χρήση, είναι προκαθορισμένες, και δεν είναι δυνατή η κατασκευή τους από το χρήστη όπως γίνεται σε άλλα FIS.

Στη συνέχεια θα οριστούν το σφάλμα εκπαίδευσης (training error), επαλήθευσης (validation error), το σφάλμα ελέγχου (checking error), και μια παράμετρος των ANFIS που καλείται βήμα (step size)

Το σφάλμα εκπαίδευσης είναι η διαφορά μεταξύ της δεδομένης τιμής εξόδου ενός σετ εκπαίδευσης και της τιμής που παράγει το σύστημα όταν σε αυτό εισαχθούν τα δεδομένα εισόδου της ίδιας εξόδου. Το σύστημα ελαχιστοποιεί τη ρίζα του μέσου τετραγωνικού σφάλματος στα δεδομένα εκπαίδευσης.

Το σφάλμα επαλήθευσης αποτελεί μέτρο της ικανότητας πρόβλεψης του μοντέλου και ορίζεται ως η διαφορά μεταξύ της πραγματικής τιμής εξόδου ενός σετ επαλήθευσης και των αντίστοιχων εκτιμήσεων που εξάγει το μοντέλο. Το σύστημα ελαχιστοποιεί τη ρίζα του μέσου τετραγωνικού σφάλματος των σετ επαλήθευσης.

Σφάλμα ελέγχου είναι η διαφορά μεταξύ της γνωστής τιμής εξόδου του σετ ελέγχου και της τιμής που παράγει το μοντέλο όταν σε αυτό τροφοδοτηθούν τα δεδομένα εισόδου του ίδιου σετ.

Κατά τη φάση της εκπαίδευσης η ταχύτητα σύγκλισης του αλγορίθμου ρυθμίζεται από μία παράμετρο η οποία αναφέρεται ως **βήμα** και η οποία μεταβάλλεται δυναμικά καθώς εξελίσσεται η εκπαίδευση. Ο κανόνας ρύθμισης της, έχει ως εξής:

- Εκκίνηση με μια ορισμένη τιμή είτε από το χρήστη είτε από το σύστημα (π.χ. 0.01).
- Αν το σφάλμα παρουσιάσει πέντε συνεχόμενες μειώσεις, το βήμα θα αυξηθεί με πολλαπλασιασμό με μια σταθερά μεγαλύτερη της μονάδας.
- Αν το σφάλμα παρουσιάσει δυο διαδοχικές εναλλαγές αύξησης-μείωσης το βήμα θα μειωθεί με πολλαπλασιασμό με μια σταθερά μικρότερη της μονάδας.

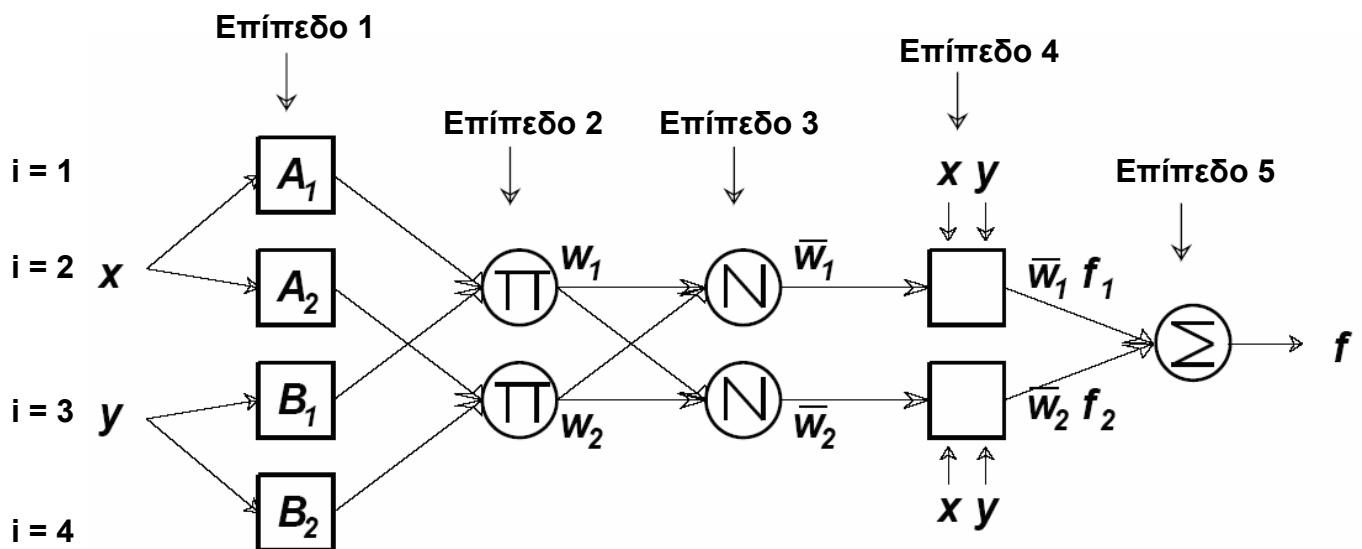
3.2.4 Αρχιτεκτονική ANFIS

Θα παρουσιαστεί η αρχιτεκτονική ενός ANFIS μέσω ενός παραδείγματος. Το σύστημα που θα περιγραφεί, έχει δύο εισόδους και υποθέτουμε πως η βάση κανόνων περιέχει δύο ασαφείς κανόνες της μορφής Εάν-Τότε τύπου Takagi-Sugeno :

1^{ος} κανόνας : Εάν $x \in A_1$ και $y \in B_1$ Τότε $f_1 = p_1x + q_1y + r_1$

2^{ος} κανόνας : Εάν $x \in A_2$ και $y \in B_2$ Τότε $f_2 = p_2x + q_2y + r_2$

όπου A_k, B_k ασαφή σύνολα και p_k, q_k, r_{ki} σταθερές με $k = 1, 2$



Σχήμα 3.5 Παράδειγμα αρχιτεκτονικής ANFIS με δύο εισόδους και δύο κανόνες
(Jang et al., 1997)

Στη συνέχεια θα επεξηγηθεί ο μηχανισμός λειτουργίας του ANFIS από επίπεδο σε επίπεδο. Η έξοδος κάθε κόμβου i που ανήκει στο επίπεδο j συμβολίζεται με $O_{j,i}$.

Επίπεδο 1: κάθε κόμβος i του επιπέδου ($i = 1..4$) έχει συνάρτηση εξόδου:

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

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

όπου A_i, B_{i-2} είναι γλωσσικές μεταβλητές (linguistic variables) που μπορεί να χαρακτηρίζουν τις εισόδους (ψήλος-κοντός κ.α.). Η έξοδος $O_{1,1}$ ορίζει το βαθμό συμμετοχής της εισόδου x στο ασαφές σύνολο A_1 , η έξοδος $O_{1,2}$ ορίζει το βαθμό συμμετοχής της εισόδου x στο ασαφές σύνολο A_2 . Κατά αναλογία επεξηγούνται και οι υπόλοιπες έξοδοι αυτού του επιπέδου.

Επίπεδο 2: κάθε κόμβος i του επιπέδου ($i = 2,3$) συμβολίζεται με Π και έχει συνάρτηση εξόδου:

$$O_{i,2} = O_{1,1} \times O_{2,1} = w_1 \quad \text{για } i = 2$$

$$O_{i,2} = O_{3,1} \times O_{4,1} = w_2 \quad \text{για } i = 3$$

Δηλαδή κάθε κόμβος του επιπέδου παράγει το γινόμενο των εισόδων του. Η έξοδος κάθε κόμβου ονομάζεται δύναμη πυροδότησης του κανόνα στον οποίο αντιστοιχεί.

Επίπεδο 3: κάθε κόμβος i του επιπέδου ($i = 2,3$) συμβολίζεται με N και έχει συνάρτηση εξόδου:

$$O_{3,i} = \frac{w_1}{w_1 + w_2} = \overline{w_1} \quad \text{για } i = 2$$

$$O_{3,i} = \frac{w_2}{w_1 + w_2} = \overline{w_2} \quad \text{για } i = 3$$

Κάθε κόμβος αυτού του επιπέδου υπολογίζει το λόγο της δύναμης πυροδότησης ενός κανόνα ως προς το άθροισμα των δυνάμεων πυροδότησης όλων των κανόνων. Κάθε έξοδος ονομάζεται κανονικοποιημένη δύναμη πυροδότησης.

Επίπεδο 4: κάθε κόμβος i του επιπέδου ($i = 2,3$) έχει συνάρτηση εξόδου:

$$O_{4,i} = \overline{w_1} f_1 = \overline{w_1} (p_1 x + q_1 y + r_1) \quad i = 2$$

$$O_{4,i} = \overline{w_2} f_2 = \overline{w_2} (p_2 x + q_2 y + r_2) \quad i = 3$$

Επίπεδο 5: Ο μοναδικός κόμβος αυτού του επιπέδου (έστω $i = 2$) υπολογίζει την ολική έξοδο ως άθροισμα όλων των σημάτων εισόδου:

$$O_5 = O_{4,2} + O_{4,3} = \overline{w_1} f_1 + \overline{w_2} f_2$$

3.2.5 Πλεονεκτήματα του συστήματος ANFIS

- Μπορεί να παράγει μια έντονα μη γραμμική αντιστοιχία των δεδομένων εισόδου με τα δεδομένα εξόδου.
- Δημιουργεί λιγότερες παραμέτρους που απαιτούν ρύθμιση σε σύγκριση με ένα απλό νευρωνικό δίκτυο και συγκλίνει ταχύτερα.
- Δεν απαιτεί γνώση των στατιστικών κατανομών των δεδομένων εισόδου ή εξόδου.
- Παράγει τους κανόνες που διέπουν ένα σύστημα χωρίς καμία γνώση ειδικού.
- Η εξαγωγή γνώσης από το σύστημα είναι άμεσα εφικτή εν αντίθεση με τα νευρωνικά δίκτυα που αποτελούν μαύρο κουτί (black box).
- Ενσωματώνει μηχανισμό αποφυγής της υπερπροσαρμογής των δεδομένων.
- Παρέχει μια μέθοδο αυτόματης ρύθμισης των παραμέτρων των συναρτήσεων συμμετοχής, ώστε να ελαχιστοποιείται το μέτρο του σφάλματος.
- Είναι ευέλικτο και ευπροσάρμοστο.

3.2.6 Εφαρμογές του ANFIS

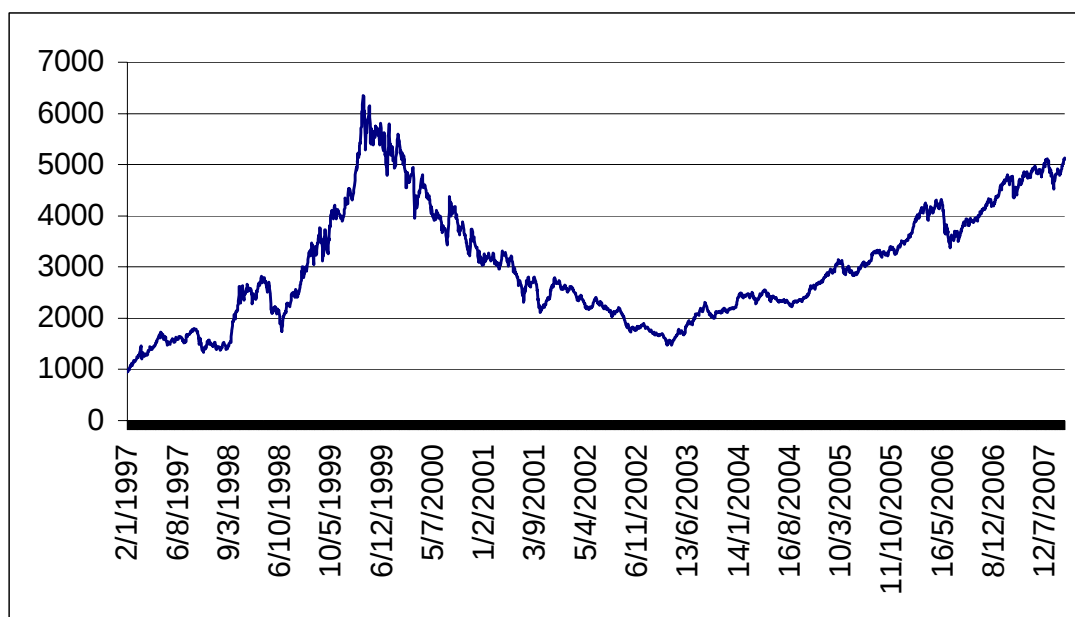
Πολλές είναι οι εφαρμογές του ANFIS σε πολλούς τομείς και προβλήματα. Στη χρηματιστηριακή πρόβλεψη (Ατσαλάκης, 2006). Στη ρομποτική, στον συντονισμό και τη συνεργασία ρομποτικών ποδοσφαιριστών (Yang et al, 2002). Στην ιατρική, για τη

διάγνωση του διαβήτη (Polat et al, 2006). Στη μηχανολογία, για τη μελέτη υποβάθμισης των παθητικών εξαρτημάτων ενός πυρηνικού εργοστασίου (Guimaraes et al, 2006) και στον έλεγχο αποθεμάτων και παραγγελιών (Wang et al, 2008). Στον αυτόματο έλεγχο, μέσω της κατασκευής εικονικών αισθητήρων για την πρόβλεψη της ασφάλειας χρήσης του υδρογόνου ως υποκατάστατο των σημερινών καυσίμων (Karria et al, 2008).

Κεφάλαιο 4: Δεδομένα, μεθοδολογία ανάλυσης και αποτελέσματα.

4.1 Δεδομένα

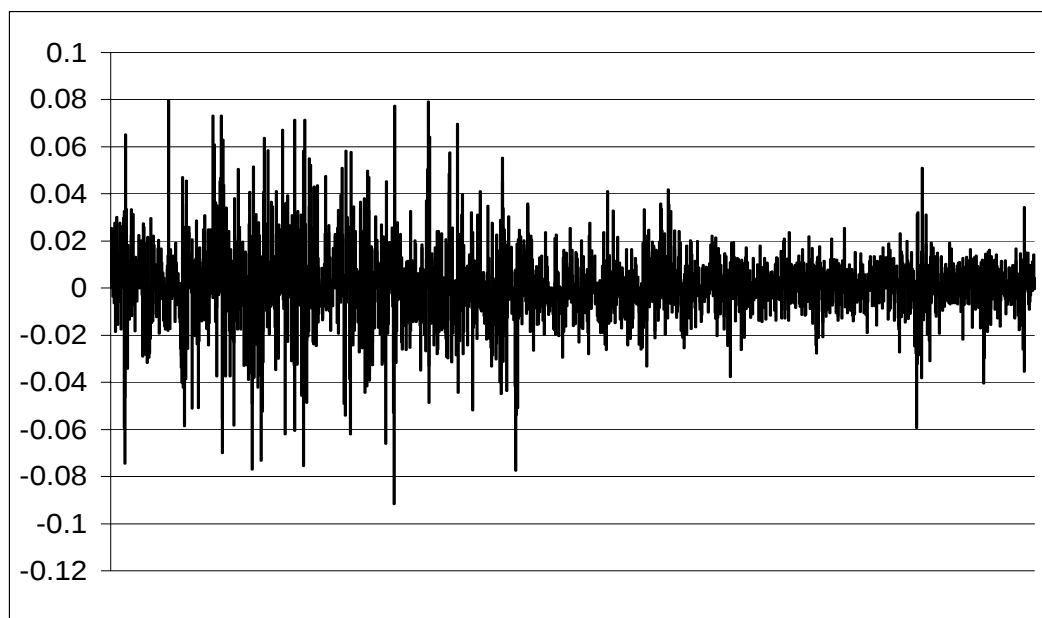
Τα δεδομένα αποτελούνται από τις ημερήσιες τιμές κλεισίματος, του γενικού δείκτη του Χρηματιστηρίου Αθηνών. Αφορούν την χρονική περίοδο μεταξύ 2/1/1997 και 28/9/2007. Το πλήθος των παρατηρήσεων είναι 2684. Η επιλογή της συγκεκριμένης περιόδου έγινε διότι από το 1997, το ενδιαφέρον των επενδυτών για το θεσμό του χρηματιστηρίου εντεινόταν, υπό το πρίσμα της προοπτικής ένταξης της Ελλάδας στη ζώνη του ευρώ.



Σχήμα 4.1 Πορεία γενικού δείκτη 1997-2007

Το σχήμα 4.1 παρουσιάζει την πορεία του γενικού δείκτη. Όπως φαίνεται η εξεταζόμενη χρονική περίοδος περιλαμβάνει ανοδικά και καθοδικά διαστήματα. Συγκεκριμένα από τον Ιανουάριο του 1997 έως το Μάιο του 1999, παρατηρείται σημαντική άνοδος, αλλά στη συνέχεια ακολουθεί μια τριετής καθοδική πορεία, που αντιστρέφεται στα τέλη του 2002. Από το σημείο αυτό έως το Σεπτέμβρη του 2007, η πορεία του γενικού δείκτη είναι ανοδική.

Στο σχήμα 4.2 παρουσιάζονται οι αποδόσεις του γενικού δείκτη κατά την επιλεγμένη περίοδο. Συνολικά, στην περίοδο 1997-2007 υπήρξαν 1392 ημέρες που ο δείκτης κινήθηκε ανοδικά, ενώ σε 1292 σημείωσε πτώση. Όπως και αναμενόταν το πλήθος των ανοδικών και καθοδικών δε διαφέρει σημαντικά.



Σχήμα 4.2 Ημερήσιες αποδόσεις-τάσεις γενικού δείκτη 1997-2007

Έτος	Ετήσια απόδοση	Αθροιστική	Τυπική απόκλιση
1997	0.539	0.539	0.0190
1998	0.847	1.842	0.0242
1999	0.895	4.386	0.0239
2000	-0.415	2.151	0.0204
2001	-0.229	1.429	0.0181
2002	-0.334	0.618	0.0109
2003	0.274	1.06	0.0121
2004	0.204	1.481	0.0093
2005	0.295	2.214	0.0084
2006	0.199	2.853	0.0119
2007*	0.147	3.420	0.0100

Πίνακας 4.1 Ετήσια στατιστικά μεγέθη γενικού δείκτη

Στο πίνακα 4.1 εμφανίζεται η ετήσια απόδοση, η αθροιστική απόδοση από την αρχή της εξεταζόμενης περιόδου και η τυπική απόκλιση των ημερήσιων αποδόσεων για κάθε έτος. Ο γενικός δείκτης εμφανίζει θετική απόδοση σε όλα τα έτη εκτός από την τριετία 2000-2002. Ιδιαίτερα υψηλή είναι η απόδοση για το 1999, φτάνοντας το 89,5 %. Ένας επενδυτής που διατηρούσε ένα χαρτοφυλάκιο αντιπροσωπευτικό του γενικού δείκτη για την συγκεκριμένη δεκαετία θα είχε απόδοση 342%. Η τυπική απόκλιση εμφανίζει σημαντική αύξηση (περίπου διπλασιάζεται) στα δύο χρονιά πριν και μετά την χρηματιστηριακή κρίση του 1999. Στα υπόλοιπα χρόνια η τυπική απόκλιση κυμαίνεται στο ίδιο επίπεδο. Το 2007 συνοδεύεται από αστερίσκο (*) διότι τα δεδομένα δεν αφορούν όλο το έτος, αλλά τους πρώτους εννέα μήνες του.

Στα πειράματα, μελετήθηκαν ξεχωριστά τρεις ορίζοντες πρόβλεψης, ημερήσιος , τριών ημερών και εβδομαδιαίος. Με τον τρόπο αυτό δίνεται χρήσιμη πληροφόρηση, όχι μόνο στον επενδυτή που πραγματοποιεί ημερήσιες συναλλαγές, αλλά και σε αυτούς που επιθυμούν να συναλλάσσονται ανά τρεις μέρες ή εβδομαδιαίος. Δίνεται κατά τον τρόπο αυτόν και μια δυνατότητα μείωσης του κόστους συναλλαγών.

4.2 Δομή δεδομένων

Έστω πως έχουμε ένα πλήθος από n ιστορικά στοιχεία για κάποιο μέγεθος, στην προκειμένη περίπτωση για τον γενικό δείκτη του Χρηματιστηρίου Αξιών Αθηνών. Το ζητούμενο είναι να εξεταστεί η δυνατότητα αξιοποίησης των ιστορικών αυτών στοιχείων για την εξαγωγή μιας εκτίμησης για τη μελλοντική εξέλιξη του δείκτη. Η εκτίμηση περιορίζεται κάθε φορά σε έναν προκαθορισμένο χρονικό ορίζοντα. Στους αλγορίθμους που μελετώνται, τα δεδομένα κατ' ουσία χωρίζονται σε δύο τμήματα:

1. Δεδομένα εκπαίδευσης (training data)
2. Δεδομένα ελέγχου (test data)

Τα δεδομένα εκπαίδευσης αποτελούν εκείνο το τμήμα των δεδομένων, τα οποία θα ληφθούν υπόψη από τον αλγόριθμο πρόβλεψης ,ώστε να επιτευχθεί η ρύθμιση των παραμέτρων του. Μια τυπική μορφή ενός μόνο set εκπαίδευσης είναι:

ΦΑΣΗ ΕΚΠΑΙΔΕΥΣΗΣ
Είσοδοι
P_{t-2}
P_{t-1}
P_t
P_{t+1}

Το συγκεκριμένο σύνολο εκπαίδευσης, εμπεριέχει τις παρατηρήσεις των τριών προηγούμενων ημερών (P_{t-2} , P_{t-1} , P_t) και το πρόσημο της έκβασης της τέταρτης μέρας (P_{t+1}), χωρίς αυτό να αποτελεί κανόνα. Εναλλακτικά τα δεδομένα εκπαίδευσης θα μπορούσε να συμπεριλαμβάνουν τις 5 προηγούμενες παρατηρήσεις (P_{t-4} , P_{t-3} , P_{t-2} , P_{t-1} , P_t) και το πρόσημο της έκβασης της έκτης μέρας (P_{t+1}) κ.ο.κ. Το πόσες προηγούμενες μέρες θα ληφθούν σε κάθε set απαιτεί διερεύνηση και θα αναλυθεί σε επόμενο κεφάλαιο.

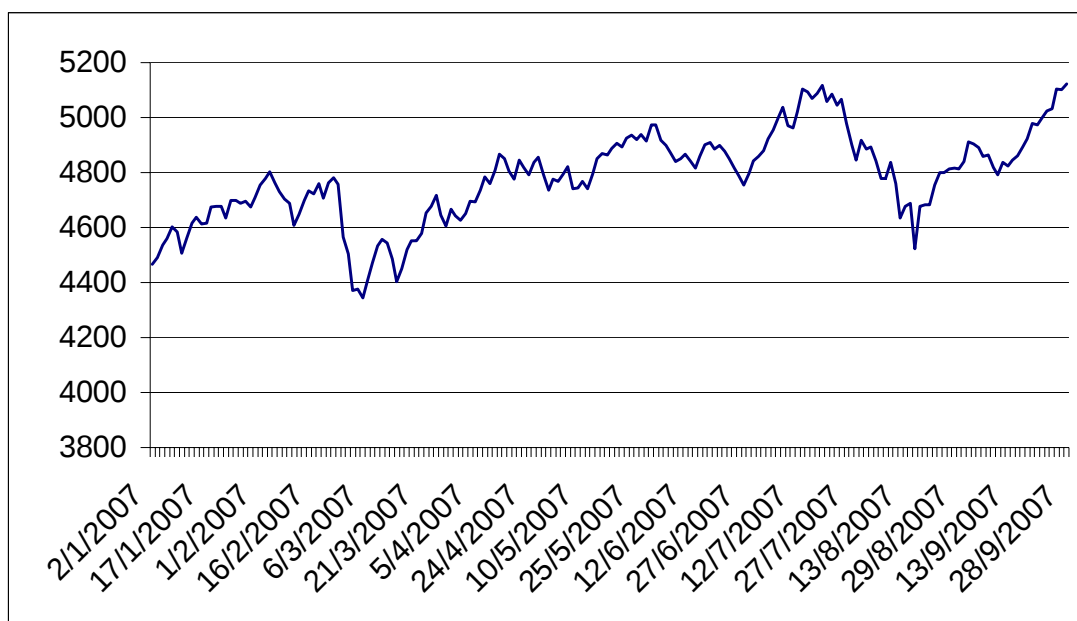
Τα δεδομένα ελέγχου αποτελούν ένα εντελώς ξεχωριστό τμήμα, από εκείνο των δεδομένων εκπαίδευσης. Το μοντέλο δεν έχει εκπαιδευτεί πάνω σε αυτά, ωστόσο του ζητείται να τα προβλέψει. Μέσω των δεδομένων ελέγχου εξετάζεται η ποιότητα του μοντέλου και η ικανότητα γενίκευσης του. Αν λοιπόν, το μοντέλο προβλέψει, σε ικανοποιητικό βαθμό τα δεδομένα ελέγχου, τότε θεωρείται ότι η ικανότητα γενίκευσης του είναι καλή. Θα πρέπει φυσικά να σημειωθεί, πως αν το δείγμα ελέγχου αλλάξει και διατηρηθούν τα δεδομένα εκπαίδευσης, η ικανότητα γενίκευσης του μοντέλου θα μεταβληθεί (προς το χειρότερο ή το καλύτερο). Όσον αφορά τη δομή, τα δεδομένα ελέγχου δε διαφοροποιούνται καθόλου από τα δεδομένα εκπαίδευσης. Η διαφορά τους έγκειται στο τρόπο χρήσης και μόνο.

Στη φάση λοιπόν του ελέγχου, εξετάζονται οι αποκλίσεις των προβλέψεων-εκτιμήσεων του μοντέλου σε σχέση με τα πραγματικά δεδομένα. Όσες περισσότερες φορές υπάρχει συμφωνία μεταξύ (P_{t+1} πραγματικό) και (P_{t+1} πρόβλεψη), τόσο καλύτερο το μοντέλο πρόβλεψης. Η παρούσα εργασία επιχειρεί πρόβλεψη τάσεων, έτσι λοιπόν τα (P_{t+1} πραγματικό) και (P_{t+1} πρόβλεψη) είναι τάσεις, ενώ οι (P_{t-2} , P_{t-1} , P_t), επιλέγεται να είναι οι αποδόσεις των αντίστοιχων ημερών όπως θα φανεί στη συνέχεια. Ως τάση ορίζεται

το πρόσημο της απόδοσης του γενικού δείκτη για την παρατήρηση P_{t+1} . Αν το πρόσημο είναι θετικό (αρνητικό) τότε και η τάση είναι θετική(αρνητική).

4.3 Πειραματικός σχεδιασμός

Στο υποκεφάλαιο αυτό θα παρουσιαστούν οι πειραματικές εφαρμογές που πραγματοποιήθηκαν στα πλαίσια της εργασίας αυτής. Τα αποτελέσματα των πειραμάτων βρίσκονται στην ενότητα 4.4. Η επιλογή των πρώτων εννέα μηνών του 2007 για περίοδο πρόβλεψης, έγινε διότι κατά το συγκεκριμένο διάστημα ο γενικός δείκτης δεν κινήθηκε έντονα μόνο προς μία κατεύθυνση αλλά προς όλες τις κατευθύνσεις, όπως φαίνεται στο σχήμα 4.3 που ακολουθεί .



Σχήμα 4.3 Πορεία γενικού δείκτη για το έτος πρόβλεψης.

Ακολουθώς περιγράφεται η δομή των δεδομένων εκπαίδευσης-εισόδου για την κάθε πειραματική περίπτωση

1^η πειραματική περίπτωση:

Έστω $r_1, \dots, r_n$ οι αποδόσεις n διαδοχικών ημερών για το γενικό δείκτη. Τα δεδομένα εκπαίδευσης διαμορφώνονται ως εξής:

Γ_1	Γ_2	Γ_3	Γ_4
Γ_2	Γ_3	Γ_4	Γ_5
.....			
Γ_{n-3}	Γ_{n-2}	Γ_{n-1}	Γ_n

Γίνεται φανερό πως τα σετ εκπαίδευσης για την **1^η Περίπτωση**, διαμορφώνονται κατά «κυλιόμενο» τρόπο. Ανάλογη είναι η μορφή των δεδομένων εκπαίδευσης στη **2^η Περίπτωση** με τη διαφορά πως λαμβάνονται σαν είσοδοι οι πέντε προηγούμενες ημέρες.

Χρονικό διάστημα εκπαίδευσης	2/1/1997 έως 29/12/2006
Χρονικό διάστημα πρόβλεψης	2/1/2007 έως 28/9/2007.
Είσοδοι στο μοντέλο	Οι αποδόσεις των τριών προηγούμενων ημερών
Έξοδος μοντέλου	Η πρόβλεψη της τάσης της επόμενης ημέρας
Πλήθος παρατηρήσεων εκπαίδευσης	2496
Πλήθος παρατηρήσεων ελέγχου	185

2^η πειραματική περίπτωση:

Χρονικό διάστημα εκπαίδευσης	2/1/1997 έως 29/12/2006
Χρονικό διάστημα πρόβλεψης	2/1/2007 έως 28/9/2007.
Είσοδοι στο μοντέλο	Οι αποδόσεις των πέντε προηγούμενων ημερών
Έξοδος μοντέλου	Η πρόβλεψη της τάσης της επόμενης ημέρας
Πλήθος παρατηρήσεων εκπαίδευσης	2496
Πλήθος παρατηρήσεων ελέγχου	185

Υπό το πρίσμα της μεταβολής του ορίζοντα πρόβλεψης σε τρεις ημέρες, η μορφή των δεδομένων εισόδου-εκπαίδευσης διαφοροποιείται. Οι αλγόριθμοι εκπαιδεύονται και πάλι σε ένα σύνολο τεσσάρων στοιχείων εισόδου, όπου τα τρία πρώτα στοιχεία είναι οι αποδόσεις του γενικού δείκτη για κάθε μία από τις προηγούμενες τρεις ημέρες. Το τέταρτο στοιχείο, πλέον δεν παίρνει την απόδοση της επόμενης ημέρας, αλλά την απόδοση των επόμενων τριών ημερών. Η δημιουργία ενός συνόλου εκπαίδευσης για την **3^η περίπτωση**, περιγράφεται ακολούθως:

Έστω $r_1, \dots, r_n$ οι αποδόσεις n διαδοχικών ημερών για το γενικό δείκτη, τα σετ εκπαίδευσης διαμορφώνονται ως εξής:

r_1	r_2	r_3	$(1+r_4)(1+r_5)(1+r_6)-1$
r_4	r_5	r_6	$(1+r_7)(1+r_8)(1+r_9)-1$
.....			
r_{n-5}	r_{n-4}	r_{n-3}	$(1+r_{n-2})(1+r_{n-1})(1+r_n)-1$

Κατά αντίστοιχο τρόπο διαμορφώνονται και τα δεδομένα εισόδου για την **4^η περίπτωση** (ορίζοντας πρόβλεψης πέντε ημερών).

3^η πειραματική Περίπτωση:

Χρονικό διάστημα εκπαίδευσης	2/1/1997 έως 29/12/2006
Χρονικό διάστημα πρόβλεψης	2/1/2007 έως 28/9/2007.
Είσοδοι στο μοντέλο	Οι αποδόσεις των τριών προηγούμενων ημερών
Έξοδος μοντέλου	Η πρόβλεψη της τάσης του επόμενου τριημέρου
Πλήθος παρατηρήσεων εκπαίδευσης	830
Πλήθος παρατηρήσεων ελέγχου	61

Ανάλογη είναι η μορφή των δεδομένων εκπαίδευσης στη **4^η Περίπτωση** με τη διαφορά πως λαμβάνονται σαν είσοδοι οι πέντε προηγούμενες ημέρες και ως έξοδος η τάση του επόμενου πενθημέρου.

4^η πειραματική περίπτωση:

Χρονικό διάστημα εκπαίδευσης	2/1/1997 έως 29/12/2006
Χρονικό διάστημα πρόβλεψης	4/7/2006 έως 28/9/2007.
Είσοδοι στο μοντέλο	Οι αποδόσεις των πέντε προηγούμενων ημερών
Έξοδος μοντέλου	Η πρόβλεψη της τάσης του επόμενου πενθήμερου
Πλήθος παρατηρήσεων εκπαίδευσης	473
Πλήθος παρατηρήσεων ελέγχου	61

4.4 Αποτελέσματα

Στην ενότητα αυτή κεφάλαιο θα παρουσιαστούν τα αποτελέσματα των SVM και ANFIS.

4.4.1 SVM

Για την περίπτωση των SVM, οι πυρήνες που θα χρησιμοποιηθούν είναι ο γραμμικός, ο πολυωνυμικός και ο RBF. Ο πολυωνυμικός πυρήνας είναι δευτέρου, τρίτου ή τέταρτου βαθμού. Η παράμετρος σ για τον RBF πυρήνα και η σταθερά παραχώρησης θα μεταβάλλονται ανάλογα με την περίπτωση.

4.4.1.1 Ημερήσια πρόβλεψη

Αρχικά εξετάζεται η **1^η πειραματική περίπτωση** της ενότητας 4.3. Σαν είσοδοι λαμβάνονται οι αποδόσεις των τριών προηγούμενων ημερών. Το δείγμα εκπαίδευσης περιλαμβάνει 2496 παρατηρήσεις, ενώ ο έλεγχος γίνεται σε 185 παρατηρήσεις. Θα δοκιμαστούν τιμές του σ 0.1 και 0.05. Η σταθερά παραχώρησης C , θα πάρει την τιμή 10 για τον γραμμικό και τους πολυωνυμικούς πυρήνες ενώ στον RBF θα λάβει τιμές 10,100,1000. Η πράξη δείχνει, πως η σταθερά παραχώρησης δεν επηρεάζει σημαντικά τα αποτελέσματα πρόβλεψης, στην περίπτωση γραμμικού ή πολυωνυμικού πυρήνα. Αυτός είναι ο λόγος που δεν εξετάζονται περισσότερες περιπτώσεις τιμών για τη

σταθερά C στους πυρήνες αυτούς. Αντίθετα η σταθερά C μεταβάλλει σημαντικά τα αποτελέσματα, στην περίπτωση του πυρήνα RBF.

Στον πίνακα 4.2 συνοψίζονται τα επιμέρους σενάρια που εξετάζονται για τις παραμέτρους των SVM, ενώ ο πίνακας 4.3 παρουσιάζει τα σχετικά αποτελέσματα..

Πυρήνας (παραμέτροι)	Τιμές παραμέτρων
Γραμμικός (Συντελεστής παραχώρησης C)	10
Πολυωνυμικός $2^{\text{ου}}$ $3^{\text{ου}}$ ή $4^{\text{ου}}$ βαθμού (Συντελεστής παραχώρησης C)	10
RBF (Συντελεστής παραχώρησης C , Παράμετρος σ)	10, {0.01,0.05,0.1}

Πίνακας 4. 2 Επιλογή παραμέτρων SVM για τη 1η πειραματική περίπτωση

Γραμμικός πυρήνας				
Παράμετροι		Επιτυχής πρόβλεψη ανόδων	Επιτυχής πρόβλεψη πτώσεων	Συνολικό ποσοστό επιτυχίας
C=10		0.55769	0.50617	0.5351
Πολυωνυμικός πυρήνας				
C=10	k=2	0.5	0.6049	0.54594
	k=3	0.6154	0.5309	0.5784
	k=4	0.5385	0.5803	0.5568
RBF πυρήνας				
C=10	$\sigma = 0.1$	0.6154	0.5062	0.5676
	$\sigma = 0.05$	0.5962	0.5432	0.5730
	$\sigma = 0.01$	0.5288	0.5802	0.5513
C=100	$\sigma = 0.1$	0.6154	0.5062	0.5676
	$\sigma = 0.05$	0.5769	0.5185	0.5513
	$\sigma = 0.01$	0.5865	0.5555	0.5730
C=1000	$\sigma = 0.1$	0.61538	0.4938	0.5622
	$\sigma = 0.05$	0.5961	0.5062	0.5567
	$\sigma = 0.01$	0.5961	0.54320	0.5730

Πίνακας 4. 3 Αποτελέσματα 1^{ης} πειραματικής περίπτωσης

Ο γραμμικός πυρήνας αδυνατεί να μοντελοποιήσει ικανοποιητικά τα συγκεκριμένα δεδομένα, δίνοντας επιτυχή πρόβλεψη για το 53,51% των παρατηρήσεων ελέγχου. Ο πολωνυμικός πυρήνας τρίτου βαθμού ανταποκρίνεται καλύτερα, δίδοντας επιτυχή πρόβλεψη 57,84%. Τέλος ο RBF πυρήνας εμφανίζει παρόμοια αποτελέσματα με τον πολωνυμικό με μέγιστη επιτυχή πρόβλεψη 57,30%.

Ο πίνακας 4.4 που ακολουθεί, αφορά την 2^η πειραματική περίπτωση της ενότητας 4.3. Σαν είσοδοι λαμβάνονται οι αποδόσεις των πέντε προηγούμενων ημερών. Η εκπαίδευση γίνεται πάνω σε 2491 παρατηρήσεις, ενώ ο έλεγχος σε 188. Η επιλογή των παραμέτρων είναι ίδια με αυτήν της 1^{ης} περίπτωσης, δηλαδή $C \in \{10,100,100\}$. και $\sigma \in \{0.1,0.05,0.01\}$.

Γραμμικός πυρήνας				
Παράμετροι		Επιτυχής πρόβλεψη ανόδων	Επιτυχής πρόβλεψη πτώσεων	Μέσο ποσοστό επιτυχίας
C=10		0.5514	0.5802	0.5638
Πολωνυμικός πυρήνας				
C=10	k=2	0.5234	0.5926	0.5532
	k=3	0.5701	0.5679	0.5691
	k=4	0.5421	0.5555	0.5479
RBF πυρήνας				
C=10	$\sigma = 0.1$	0.5888	0.5432	0.5691
	$\sigma = 0.05$	0.5981	0.5432	0.5745
	$\sigma = 0.01$	0.5607	0.58025	0.5691
C=100	$\sigma = 0.1$	0.6168	0.53086	0.5798
	$\sigma = 0.05$	0.5794	0.5432	0.5638
	$\sigma = 0.01$	0.5981	0.5062	0.5585
C=1000	$\sigma = 0.1$	0.6262	0.59259	0.6117
	$\sigma = 0.05$	0.5701	0.5185	0.5479
	0.01	0.6075	0.5309	0.5745

Πίνακας 4. 4 Αποτελέσματα 2^{ης} πειραματικής περίπτωσης

Ο γραμμικός πυρήνας ταξινομεί σωστά το 56,38% των περιπτώσεων ελέγχου. Ο πολωνυμικός πυρήνας τρίτου βαθμού επιτυγχάνει έως και 56,91% επιτυχή πρόβλεψη,

και επομένως δεν διαφοροποιείται σημαντικά από το γραμμικό. Τέλος ο πυρήνας RBF προβλέπει σωστά έως και το 61.17% των παρατηρήσεων ελέγχου και φαίνεται να είναι ο καταλληλότερος.

4.4.1.2 Τριήμερη πρόβλεψη

Τα αποτελέσματα του πίνακα 4.5 αφορούν την 3^η πειραματική περίπτωση της ενότητας 4.3. Σαν είσοδοι λαμβάνονται οι αποδόσεις των τριών προηγούμενων ημερών. Το μοντέλο εκπαιδεύεται σε 830 παρατηρήσεις και ο έλεγχος γίνεται σε 61 παρατηρήσεις. Η παράμετρος C θα λάβει τιμές {100,1000,10000} Οι τιμές του σ συνεχίζουν να είναι οι ίδιες.

Γραμμικός πυρήνας				
Παράμετροι		Επιτυχής πρόβλεψη ανόδων	Επιτυχής πρόβλεψη πτώσεων	Μέσο ποσοστό επιτυχίας
C=10		0.5294	0.5555	0.5409
Πολυωνυμικός πυρήνας				
C=10	k=2	0.5294	0.5185	0.5246
	k=3	0.5294	0.5555	0.5410
	k=4	0.6765	0.4444	0.5738
RBF πυρήνας				
C=10	$\sigma = 0.1$	0.6764	0.4814	0.5902
	$\sigma = 0.05$	0.6764	0.4814	0.5902
	$\sigma = 0.01$	0.5588	0.5926	0.5738
C=100	$\sigma = 0.1$	0.6765	0.4444	0.5738
	$\sigma = 0.05$	0.6470	0.5185	0.5902
	$\sigma = 0.01$	0.5588	0.4815	0.5246
C=1000	$\sigma = 0.1$	0.6470	0.4444	0.5574
	$\sigma = 0.05$	0.6176	0.4444	0.5310
	$\sigma = 0.01$	0.67647	0.4444	0.5604

Πίνακας 4.5 Αποτελέσματα 3^{ης} πειραματικής περίπτωσης

Ο γραμμικός πυρήνας ταξινομεί σωστά το 54,09% των περιπτώσεων ελέγχου. Ο πολυωνυμικός τέταρτου βαθμού φαίνεται να γενικεύει αποτελεσματικότερα,

πετυχαίνοντας επιτυχή πρόβλεψη στο 57,38% των περιπτώσεων. Ακόμη υψηλότερη ικανότητα γενίκευσης παρουσιάζει ο πυρήνας RBF ταξινομώντας έως και το 59,02% των παρατηρήσεων στο δείγμα ελέγχου σωστά.

4.4.1.3 Πενθήμερη πρόβλεψη

Τα αποτελέσματα του πίνακα 4.6 αφορούν την 4^η πειραματική περίπτωση της ενότητας 4.3. Η εκπαίδευση πραγματοποιείται σε 474 παρατηρήσεις και ο έλεγχος σε 61 παρατηρήσεις

Γραμμικός πυρήνας				
Παράμετροι		Επιτυχής πρόβλεψη ανόδων	Επιτυχής πρόβλεψη πτώσεων	Μέσο ποσοστό επιτυχίας
C=1000		0.625	0.5	0.5833
Πολυωνυμικός πυρήνας				
C=1000	k=2	0.6	0.5	0.5333
	k=3	0.575	0.35	0.5
	k=4	0.625	0.5	0.5833
RBF πυρήνας				
C=1000	$\sigma = 1$	0.6	0.5	0.5667
	$\sigma = 3$	0.6	0.55	0.5833
	$\sigma = 5$	0.625	0.55	0.6
C=10000	$\sigma = 1$	0.55	0.525	0.5333
	$\sigma = 3$	0.575	0.55	0.5667
	$\sigma = 5$	0.625	0.55	0.6
C=100000	$\sigma = 1$	0.525	0.5375	0.5333
	$\sigma = 3$	0.575	0.55	0.5667
	$\sigma = 5$	0.625	0.55	0.6

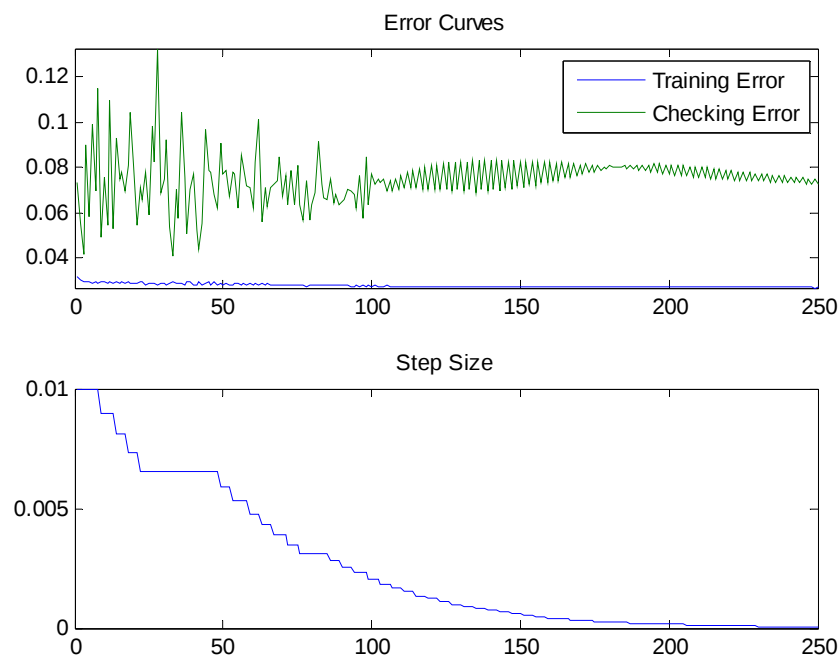
Πίνακας 4.6 Αποτελέσματα 4^{ης} πειραματικής περίπτωσης

Ο γραμμικός πυρήνας επιτυγχάνει επιτυχή πρόβλεψη στο 58,33% των περιπτώσεων ελέγχου. Ο πολυωνυμικός πυρήνας 4^{ου} βαθμού εμφανίζει την ίδια ικανότητα γενίκευσης με τον γραμμικό. Ο RBF πυρήνας οδηγεί σε επιτυχή πρόβλεψη, στο 60% των

περιπτώσεων πρόβλεψης, κατέχοντας την καλύτερη ικανότητα γενίκευσης για τη συγκεκριμένη περίπτωση.

4.4.2 ANFIS

Στη συνέχεια θα παρουσιαστούν τα αποτελέσματα του μοντέλου ANFIS, σε δομή αντίστοιχη με τα αποτελέσματα των SVM. Το πλήθος των συναρτήσεων συμμετοχής ανά είσοδο επιλέγεται να είναι είτε δύο είτε τρία. Οι συναρτήσεις συμμετοχής που εξετάστηκαν ήταν ο Gauss2 και ο Gbell. Το πλήθος των παρατηρήσεων στο δείγμα έλεγχου εκπαίδευσης για όλα τα πειράματα επιλέχθηκε να είναι 50. Για όλα τα πειράματα, το πλήθος των epochs εκκινούσε από τις 250 και προσαρμοζόταν ανάλογα με τα διαγράμματα: epochs-σφάλμα, epochs-βήμα. Ο τρόπος προσαρμογής των εποχών βασίζεται στην παρατήρηση, πως όταν τα epochs ρυθμίζονται ώστε να είναι αρκετές για τον περιορισμό του σφάλματος και της ταλάντωσης του βήματος αλλά όχι τόσες ώστε να ταλαντώνεται για πολύ το σύστημα, τα αποτελέσματα είναι συχνά καλύτερα. Ακολουθεί ένα παράδειγμα προσαρμογής των epochs:



Σχήμα 4.4 Πορεία σφάλματος και βήματος

Στο σχήμα 4.4 παρατηρούμε πως ήδη από τις 180 περίπου epochs το σφάλμα ελέγχου εκπαίδευσης έχει αρχίσει να ταλαντώνεται σε πολύ μικρό εύρος. Αν λοιπόν αντί για 250

epochs επιλέξουμε 180 epochs, είναι πιθανή μια βελτίωση της ικανότητας γενίκευσης του μοντέλου.

Ο χρόνος ολοκλήρωσης των πειραμάτων για ένα συγκεκριμένο υπολογιστικό σύστημα, εξαρτάται από τέσσερις παραμέτρους: (1) το πλήθος των εισόδων, (2) το πλήθος των συναρτήσεων συμμετοχής ανά είσοδο, (3) το πλήθος των δεδομένων εκπαίδευσης και (4) το πλήθος των epochs εκπαίδευσης του συστήματος. Όσο αυξάνεται η τιμή των παραμέτρων αυτών, αυξάνεται και ο χρόνος εκτέλεσης των πειραμάτων. Το πλήθος των κανόνων που παράγεται σε κάθε περίπτωση είναι ίσο με u^x όπου u το πλήθος των συναρτήσεων συμμετοχής και x το πλήθος των εισόδων στο μοντέλο.

4.4.2.1 Ημερήσια πρόβλεψη

Αρχικά εξετάζεται η **1^η πειραματική περίπτωση** της ενότητας 4.3. Σαν είσοδοι λαμβάνονται οι αποδόσεις των τριών προηγούμενων ημερών. Το δείγμα εκπαίδευσης περιλαμβάνει 2496 παρατηρήσεις, ενώ ο έλεγχος γίνεται σε 185 παρατηρήσεις. Στο πείραμα 1, εξετάζεται η συνάρτηση συμμετοχής Gauss2 και στο πείραμα 2 η συνάρτηση Gbell, για την περίπτωση τριών συναρτήσεων συμμετοχής ανά είσοδο. Ο χρόνος εκτέλεσης των πειραμάτων ήταν της τάξης των δέκα λεπτών.

Gauss2					
Αριθμός πειράματος	Epochs	Πλήθος συναρτήσεων συμμετοχής	Επιτυχής πρόβλεψη ανόδων	Επιτυχής πρόβλεψη πτώσεων	Μέσο ποσοστό επιτυχίας
Πείραμα 1	200	3	0.6538	0.4691	0.5730
Gbell					
Πείραμα 2	200	3	0.6346	0.4568	0.5568

Πίνακας 4.7 Αποτελέσματα 1^{ης} πειραματικής περίπτωσης

Για το πείραμα 1 επιτυχία πρόβλεψης βρίσκεται στο 58,33%. Στο πείραμα 2 η αλλαγή του τύπου συνάρτησης συμμετοχής, σε Gbell, δίνει χειρότερη πρόβλεψη (55.68%). Τα αποτελέσματα είναι σχετικά ικανοποιητικά.

Στη συνέχεια εξετάζεται η **2^η πειραματική περίπτωση** της ενότητας 4.3. Σαν είσοδοι λαμβάνονται οι αποδόσεις των πέντε προηγούμενων ημερών. Η εκπαίδευση γίνεται

πάνω σε 2491 παρατηρήσεις, ενώ ο έλεγχος σε 188. Και για τους δύο τύπους συναρτήσεων συμμετοχής (Gauss2, Gbell), λαμβάνονται οι περιπτώσεις πλήθους δύο και τριών συναρτήσεων συμμετοχής. Ο χρόνος εκτέλεσης κάθε πειράματος με πλήθος συναρτήσεων συμμετοχής δύο, είναι της τάξης των 10 λεπτών, ενώ για τρεις συναρτήσεις συμμετοχής, ο χρόνος είναι της τάξης των 10 ωρών.

Gauss2					
Αριθμός πειράματος	Epochs	Πλήθος συναρτήσεων συμμετοχής	Επιτυχής πρόβλεψη ανόδων	Επιτυχής πρόβλεψη πτώσεων	Μέσο ποσοστό επιτυχίας
Πείραμα 2α	150	2	0.6538	0.4474	0.574%
Πείραμα 2β	250	3	0.5769	0.4321	0.5135
Gbell					
Πείραμα 3α	150	2	0.6731	0.3704	0.5405
Πείραμα 3β	250	3	0.625	0.4197	0.5351

Πίνακας 4.8 Αποτελέσματα 2^{ης} πειραματικής περίπτωσης

Για τη συνάρτηση συμμετοχής τύπου Gauss2, η χρήση δυο συναρτήσεων συμμετοχής αποδίδει καλύτερα δίνοντας ποσοστό επιτυχούς πρόβλεψης 57,4%. Για τη συνάρτηση συμμετοχής Gbell, πλήθος δύο συναρτήσεων συμμετοχής, επίσης δίνει τα καλύτερα αποτελέσματα (54.05%. Η συνάρτηση τύπου Gauss2, φαίνεται να υπερέχει στην περίπτωση αυτή.

4.4.2.2 Τριήμερη πρόβλεψη

Εξετάζεται η **3^η πειραματική περίπτωση** της ενότητας 4.3. Σαν είσοδοι λαμβάνονται οι αποδόσεις των τριών προηγούμενων ημερών. Το μοντέλο εκπαιδεύεται σε 830 παρατηρήσεις και ο έλεγχος γίνεται σε 61 παρατηρήσεις. Ο χρόνος ολοκλήρωσης για κάθε πείραμα είναι της τάξης των 10 λεπτών.

Gauss2					
Αριθμός πειράματος	Epochs	Πλήθος συναρτήσεων συμμετοχής	Επιτυχής πρόβλεψη ανόδων	Επιτυχής πρόβλεψη πτώσεων	Μέσο ποσοστό επιτυχίας
Πείραμα 4α	250	2	0.7353	0.2963	0.5410
Πείραμα 4β	250	3	0.6765	0.4764	0.6065
Gbell2					
Πείραμα 5α	250	2	0.764706	0.333333	57.378
Πείραμα 5β	250	3	0.735294	0.37037	0.5738

Πίνακας 4.9 Αποτελέσματα 3^{ης} πειραματικής περίπτωσης

Για τη συνάρτηση συμμετοχής τύπου Gauss2, η χρήση τριών συναρτήσεων συμμετοχής αποδίδει καλύτερα δίνοντας ποσοστό επιτυχούς πρόβλεψης 60,65%. Για τη συνάρτηση συμμετοχής Gbell, η χρήση διαφορετικού αριθμού συναρτήσεων συμμετοχής, δεν διαφοροποιεί το μέσο ποσοστό επιτυχούς πρόβλεψης, ωστόσο στην περίπτωση 3 συναρτήσεων συμμετοχής η κατανομή της επιτυχούς πρόβλεψης μεταξύ ανόδων και καθόδων είναι πιο ομοιόμορφη και έτσι η περίπτωση αυτή μπορεί να θεωρηθεί καλύτερη.

4.4.2.3 Πενθήμερη πρόβλεψη

Τέλος εξετάζεται η **4^η πειραματική περίπτωση** της ενότητας 4.3. Η εκπαίδευση πραγματοποιείται σε 474 παρατηρήσεις και ο έλεγχος σε 61 παρατηρήσεις. Σαν είσοδοι λαμβάνονται οι πέντε προηγούμενες ημέρες. Ο χρόνος εκτέλεσης των πειραμάτων χρησιμοποιώντας δύο συναρτήσεις συμμετοχής, είναι της τάξης των 10 λεπτών και στην περίπτωση πλήθους τριών, της τάξης των 5 ωρών.

Gauss2					
Αριθμός πειράματος	Epochs	Πλήθος συναρτήσεων συμμετοχής	Επιτυχής πρόβλεψη ανόδων	Επιτυχής πρόβλεψη πτώσεων	Μέσο ποσοστό επιτυχίας
Πείραμα 6α	250	2	0.5609	0.5	0.5410
Πείραμα 6β	250	3	0.5583	0.4978	0.5349

Πίνακας 4.10 Αποτελέσματα 4^{ης} πειραματικής περίπτωσης

Στην περίπτωση αυτή η συνάρτηση τύπου Gbell δεν κατόρθωσε να πετύχει ποσοστό επιτυχούς πρόβλεψης μεγαλύτερο του 50% και τα αποτελέσματα δεν παρουσιάζονται. Η συνάρτηση συμμετοχής τύπου Gauss2, έδωσε αποτελέσματα άνω του 50% (54,1%), αλλά δεν είναι ικανοποιητικά.

4.5 Σύγκριση αποτελεσμάτων

Στον πίνακα 4.11 συνοψίζονται τα αποτελέσματα της εφαρμογής. Όπως φαίνεται στην ημερήσια πρόβλεψη οι SVM με 5 εισόδους πετυχαίνουν καλύτερη πρόβλεψη (61,17%). Στην τριήμερη πρόβλεψη, λίγο καλύτερο αποτέλεσμα πετυχαίνει το ANFIS. Τέλος, στην πενθήμερη πρόβλεψη οι SVM πετυχαίνουν σημαντικά καλύτερο αποτέλεσμα (60%). Τα αποτελέσματα των SVM γενικά κινούνται σε ισάξια ή υψηλότερα επίπεδα, ενώ η μέθοδος εξάγει αποτελέσματα εντός μερικών δευτερολέπτων εν αντίθεση με το ANFIS που απαιτεί υπερ-πολλαπλάσιο χρόνο.

Πλήθος εισόδων	Ορίζοντας Πρόβλεψης (ημέρες)	Μέγιστη επιτυχής πρόβλεψη SVM	Μέγιστη επιτυχής πρόβλεψη ANFIS	Συστημα που υπερέχει
3	1	0.5730	0.5730	ισοδύναμα
5	1	0.6117	0.5740	SVM
3	3	0.5902	0.6065	ANFIS
5	5	0.6	0.5410	SVM

Πίνακας 4.11 Σύγκριση αποτελεσμάτων

Κεφάλαιο 5: Συμπεράσματα

Ο στόχος της διπλωματικής ήταν η πρόβλεψη χρηματιστηριακών τάσεων στο Χρηματιστήριο Αθηνών. Ως ένα βαθμό ο στόχος επιτεύχθηκε. Τόσο η μεθοδολογία των SVM όσο και η μεθοδολογία των ANFIS έδωσαν θετικά αποτελέσματα, τα οποία βρίσκονται κοντά στα αποτελέσματα άλλων σχετικών ερευνών. Η εργασία λοιπόν, φαίνεται να αντίκειται στην υπόθεση της αποτελεσματικής αγοράς. Οι SVM φάνηκαν ιδιαίτερα εύχρηστες κατά την διαδικασία εκτέλεσης των πειραμάτων, ενώ ο χρόνος εξαγωγής του κάθε αποτελέσματος ήταν της τάξης των δέκα δευτερολέπτων. Η διερεύνηση των παραμέτρων στις SVM θεωρείται ικανοποιητική, το ίδιο δυστυχώς δεν ισχύει για το ANFIS. Η χρήση μεγαλύτερης ποικιλίας τιμών για τις παραμέτρους του ANFIS δεν κατέστη δυνατή στην συγκεκριμένη έρευνα, εξαιτίας της αυξημένης υπολογιστικής ισχύος που απαιτεί το σύστημα ANFIS και που έχει ως αποτέλεσμα την ιδιαίτερα χρονοβόρα εξαγωγή αποτελεσμάτων.

Στα πείραμα που εκτελέστηκαν, οι μηχανές διανυσμάτων υποστήριξης απέδωσαν καλύτερα από το προσαρμοστικό νευροασαφές σύστημα συμπερασμού. Η συγκριτική υστέρηση του ANFIS μπορεί να οφείλεται στο πλήθος των ιστορικών στοιχείων που χρησιμοποιήθηκε (η χρήση στοιχείων σε μεγαλύτερο βάθος στο παρελθόν, ίσως να απέδιδε καλύτερα αποτελέσματα για το ANFIS). Ακόμη μια εκτενέστερη διερεύνηση των παραμέτρων του συστήματος έχει νόημα προς το σκοπό αυτό.

Ακόμη μπορούν να διερευνηθούν κάποιες επεκτάσεις, όπως:

1. Χρήση δεικτών τεχνικής ανάλυσης
2. Εφαρμογή και άλλων τεχνικών
3. Διαφοροποίηση του χρονικού ορίζοντα και εφαρμογή σε άλλα χρηματιστήρια
4. Συνδυασμός με θεμελιώδη ανάλυση
5. Συνδυασμός με διαδικασίες βελτιστοποίησης χαρτοφυλακίων (π.χ. Markowitz)

Βιβλιογραφία:

1. Aizerman M., Braveman E., Rozonoer L., (1964) “Theoretical foundations of the potential function method in pattern recognition learning. Automation and remote control, 25: 821-837.
2. Altrock, C. (1995.), "*Fuzzy Logic and Neuro-Fuzzy Applications Explained*", New Jersey, Prentice Hall PTR,
3. Bezdek James. και Pal Sankar K. (1992) “Fuzzy models for pattern recognition: methods that search for structures in data” . IEEE CS Press
4. Bishop Christopher M (1995) “Neural Networks for Pattern Recognition” Oxford University Press
5. Boser Bernhard E., Isabelle M. Guyon, Vladimir N. Vapnik (1992), “ A training algorithm for optimal margin classifiers”. Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory ISBN:0-89791-497-X
6. Casiraghi Elena, Lanzarotti Raffaella , Lipori Giuseppe (2005), “A face detection system based on color and support vector machines”
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.105.4406&rep=rep1&type=pdf>
7. Cheung Wee Mien και Uzay Kaymak (2006) “A Fuzzy Logic Based Trading System”, Econometric Institute, Erasmus School of Economics, Erasmus University, Rotterdam, the Netherlands. [www.nisis.risk-technologies.com/\(S\(iiripdvjr5t2puqvka0qp45\)\)/events/symp2007/papers/BB25_p_kaymak.pdf](http://www.nisis.risk-technologies.com/(S(iiripdvjr5t2puqvka0qp45))/events/symp2007/papers/BB25_p_kaymak.pdf)
8. Doumpos Michael, Zopounidis Constantin, and Golfinopoulou Vassiliki “Additive Support Vector Machines for Pattern Classification” IEEE transactions on systems, man, and cybernetics-part b: cybernetics, vol. 37, no. 3, june 2007
9. Drucker Harris, Chris J.C. Burges, Linda Kaufman, Alex Smola, Vladimir Vapnik (1997) “Bell Labs AT&T Labs”
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.4845&rep=rep1&type=pdf>

10. Furey Terrence S., Cristianini Nello, Duffy Nigel, Bednarski David W., Schummer Michel, Haussler David. (2000), "Support vector classification and validation of cancer tissue samples using micro array expression data" *Bioinformatics* Vol. 16 no. 10 2000 Pages 906-914
11. Grosan Crina και Abraham Ajith (2005), "Stock Market Modeling Using Genetic Programming ensembles" <http://www.cs.ubbcluj.ro/~cgrosan/stock-chapter.pdf>
12. Guimaraes Antonio Cesar Ferreira, Denise Cunha Cabral, Celso Marcelo Franklin Lapa "Adaptive fuzzy system for degradation study in nuclear power plants' passive components. [Volume 48, Issue 7](#), September 2006, Pages 655-663
13. Hart A.(1986) "Knowledge Acquisition for Expert Systems" Kogan Page LTD ,London ISBN 1 85091 091 X
14. Haykin Simon, Macmillan (1994) "Neural networks: a comprehensive foundation" Cambridge University Press New York
15. Hearst Marti A. (1998) "Support vector machines" University of California, Berkeley www.cs.toronto.edu/~zemel/Courses/CS411/Papers/svmOverview.pdf
16. Heisele Bernd (2002), "Visual Object Recognition with Supervised Learning", *MIT Center for Biological and Computational Learning* 1094-7167/03 *Published by the IEEE Computer Society*
17. Herrera Francisio και Verdegay Jose Luis (1996), "Genetic algorithms and soft computing", Physica-Verlag ISBN:379080956X
18. Hidetoshi Miyao, Minoru Maruyama (2007) "An online handwritten music symbol recognition system", Shinsu University, Japan. *International Journal on Document Analysis and Recognition* Volume 9 , Issue 1 (February 2007) Pages: 49 - 58 ISSN:1433-2833
19. Jang Jyh-Shing Roger (1993) "Adaptive-Network-Based Fuzzy Inference System" Department of Electrical Engineering and Computer Science

- University of California, Berkeley. IEEE Trans. on Systems, Man and Cybernetics, vol. 23, no. 3, pp. 665-685
20. Karri Vishy, Tien Ho, Ole Madsen (2008), “Artificial neural networks and neuro-fuzzy inference systems as virtual sensors for hydrogen safety prediction”. International journal of hydrogen energy, Volume 33, Issue 11, June 2008, Pages 2857-2867
 21. Kim Kyoung-jae (2003) “Financial time series forecasting using support vector machines”, College of Business Administration, Dongguk University, Seoul, South Korea Neurocomputing, 55 (2003) 307-319
 22. Kimoto, T., Asakawa, K., Yoda, M., & Takeoka, M. (1990). “Stock market prediction system with modular neural network”, *Proceedings of the International Joint Conference on Neural Networks* (pp. 1–6). San Diego, CA.
 23. Lee K.H., G.S. Jo (1999) “Expert system for predicting stock market timing using a candlestick chart”, Inha University, ISSN 0957-4174 [Elsevier](#)
 24. Ma Jianmin, Nguyen M. N., Rajapakse J. C, (2005)“gene classification using codon usage analysis and support vector machines”. Proceedings of the 2005 IEEE Symposium.
<http://ieeexplore.ieee.org/iel5/8857/4359833/04359889.pdf?arnumber=4359889>
 25. Nishina, T. and M. Hagiwara (1997), ‘Fuzzy Interface Neural Network’, Neurocomputing, Vol. 14, 223-239.
 26. Nygren Karl (2004) “Stock Prediction – A Neural Network Approach”, Master Thesis, Royal Institute of Technology, KTH. <http://www.f.kth.se/~f98-kny/>
 27. Omidvar Omid and Judith Dayhoff (1997). “Neural networks and pattern recognition” ISBN: 978-0-12-526420-4
 28. Osuna Edgar, Fleund Robert, Girosi Federico (1997) Training support vector machines: an application to face detection. ISBN:0-8186-7822-4 IEEE Computer Society Washington, DC, USA

29. Partridge, D. and Hussain D.K.M (1995), "Knowledge based information systems" ISBN 0077076249 McGraw-Hill New York USA
30. Polat Kemal , Salih Günes, "An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease" Volume 17 , Issue 4 (July 2007) ISSN:1051-2004 Pages 702-710
31. Ripley B.D. and Hjort N.L. (1996) "Pattern Recognition and Neural Networks" Cambridge University Press.
32. Roobaert D., Van Hulle M. M., "View based 3D object recognition with support vector machines" ISBN: 0-7803-5673-x University Leuven, USA
33. Rumelhart D.E., Hinton G.E., Williams R.J (1986)"Learning internal representations by error propagation" in Rumelhart D.E and McClelland J. A. The MIT press
34. Smith N and Gales M. (2002.) "Speech recognition using SVMs". In T.G. Dietterich, S. Becker, and Z. Ghahramani, editors, Advances in Neural Information Processing Systems 14. MIT Press
35. Statnikov Alexander , Aliferis Constantin Fotios, Tsamandinos Ioannis (2004) "Methods for multi-category cancer diagnosis from gene expression data: a comprehensive evaluation to inform decision support system development" International Journal of Medical Informatics (2005) 74, 491-503
36. Vapnik V. N. (1998) "*Statistical Learning Theory*", John Wiley & Sons,. New York USA
37. Walezak, S. (1999), "Gaining Competitive Advantage For Trading In Emerging Capital Markets With Neural Networks", *Journal Of Management Information Systems*, Vol. 16, Issue 2, 178-194.
38. Wang Wen-Pai, Ze Chen (2008), "A neuro-fuzzy based forecasting approach for rush order control applications Volume 35 , Issue 1-2 (July 2008) Pages 223-234 Taiwan

39. Winter G. (1995) “Genetic algorithms in engineering and computer science”
John Wiley & Sons, Inc. New York, USA
40. Yang Gu, Junfeng Liu, Feng Pan, Lihui Cui, Qinglin Wang (2002) “ANFIS
Based Method to Implement Flexible Positioning of Secondary Attackers in
RoboCup Simulation League”,
<http://www.cs.cmu.edu/~guyang/Everest/doc/FPSA.PDF>
41. Ατσαλάκης Γ. (2006) “Σύστημα πρόβλεψης της βραχυχρόνιας τάσης της τιμής
των μετοχών με χρήση νευρο-ασαφών μεθόδων” Διδακτορική διατριβή.
Πολυτεχνείο Κρήτης, Χανιά.
42. Βασιλείου Α.Χ (2006) ,μεταπτυχιακή διατριβή: ”Ανάπτυξη αλγορίθμου
βασισμένου σε προσαρμοζόμενα νευροασαφή συστήματα ANFIS για τη
βελτίωση της λειτουργίας εν σειρά συστημάτων ελέγχου ποιότητας λιγνιτών”
Πολυτεχνείο Κρήτης, Χανιά.
43. Μανιαδάκης Μ. (1999), “Genetic algorithms for classification and function
identification applications” μεταπτυχιακή διατριβή Πολυτεχνείο Κρήτης, Χανιά