



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΡΕΥΣΤΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΘΕΡΜΙΚΩΝ ΣΤΡΟΒΙΛΟΜΗΧΑΝΩΝ

ΔΙΑΓΝΩΣΤΙΚΗ ΑΕΡΙΟΣΤΡΟΒΙΛΩΝ ΜΕ ΧΡΗΣΗ
ΜΗΧΑΝΩΝ ΔΙΑΝΥΣΜΑΤΩΝ ΥΠΟΣΤΗΡΙΞΗΣ
(SUPPORT VECTOR MACHINES – SVM)

Διπλωματική εργασία

ΧΡΗΣΤΟΥ Α. ΚΑΛΛΑΘΑΚΗ

Επίβλεψη: Κ. Μαθιουδάκης

ΑΘΗΝΑ
Φεβρουάριος 2010



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΡΕΥΣΤΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΘΕΡΜΙΚΩΝ ΣΤΡΟΒΙΛΟΜΗΧΑΝΩΝ

Περίληψη της Διπλωματικής Εργασίας:

ΔΙΑΓΝΩΣΤΙΚΗ ΑΕΡΙΟΣΤΡΟΒΙΛΩΝ ΜΕ ΧΡΗΣΗ ΜΗΧΑΝΩΝ ΔΙΑΝΥΣΜΑΤΩΝ ΥΠΟΣΤΗΡΙΞΗΣ (SUPPORT VECTOR MACHINES – SVM)

Χρήστος Α. Καλαθάκης

Στη παρούσα διπλωματική εργασία εξετάζεται η δυνατότητα εφαρμογής των Μηχανών Διανυσμάτων Υποστήριξης (Support Vector Machines – SVM) για τη διάγνωση βλαβών συνιστωσών αεριοστροβίλων.

Κατά τη λειτουργία ενός αεριοστροβίλου μπορεί να παρουσιασθεί βλάβη σε κάποια συνιστώσα του. Η βλάβη αυτή μπορεί να επιφέρει μείωση της απόδοσης της μηχανής ή ακόμα και την καταστροφή της. Ένας αεριοστροβίλος μπορεί να παρουσιάζει διάφορες καταστάσεις λειτουργίας ανάλογα με το εάν είναι “υγιής” ή αν λειτουργεί παρουσία βλάβης, καθώς και με το μέγεθος της βλάβης. Συνεπώς είναι απαραίτητο να μπορούμε να διαπιστώσουμε τη βλάβη και το μέγεθός της έγκαιρα και αξιόπιστα. Αυτό επιτυγχάνεται με τη χρήση διαγνωστικών συστημάτων που βασίζονται σε διάφορες διαγνωστικές τεχνικές.

Οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines – SVM) εφευρέθηκαν από τον Vapnik και αποτελούν μια μέθοδο επεξεργασίας δεδομένων με μεγάλη επιτυχία σε πληθώρα προβλημάτων ταξινόμησης. Η χρήση των SVM στην περίπτωση της ταξινόμησης καταφέρνει τον διαχωρισμό ενός ζεύγους κλάσεων από μια διαχωριστική ευθεία (ή επίπεδο ή πιο γενικά υπερεπίπεδο).

Αν κάθε κατάσταση λειτουργίας ενός αεριοστροβίλου την αντιμετωπίσουμε ως μία κλάση, τότε, με τη χρήση των SVM, μπορούμε να επιτύχουμε τον διαχωρισμό τους και να κατατάσσουμε κάθε νέα εξεταζόμενη κατάσταση λειτουργίας σε μία από

τις κλάσεις – καταστάσεις λειτουργίας που διαχωρίσαμε. Τα παραπάνω αποτελούν την αρχή λειτουργίας του διαγνωστικού προγράμματος που δημιουργήθηκε και εφαρμόστηκε για διάφορες καταστάσεις λειτουργίας και βλάβες συγκεκριμένου αεριοστροβίλου.

Συγκεκριμένα, στη παρούσα διπλωματική αναπτύσσονται και περιγράφονται τα εξής:

- Βασικές αρχές της θεωρίας των SVM.
- Προτεινόμενη μεθοδολογία εφαρμογής των SVM για την διάγνωση βλαβών συνιστωσών αεριοστροβίλων.
- Εφαρμογή της μεθόδου των SVM για τη διάγνωση ρεαλιστικών αεροθερμοδυναμικών βλαβών αεριοστροβίλου διπλού ρεύματος πολιτικής αεροπορίας.
- Εφαρμογή της μεθόδου των SVM για τη διάγνωση πραγματικών αεροθερμοδυναμικών και μηχανικών βλαβών συμπιεστών αεριοστροβίλου.
- Αξιολόγηση των διαγνωστικών δυνατοτήτων/επιδόσεων της προτεινόμενης μεθόδου των SVM, με σύγκριση των διαγνωστικών συμπερασμάτων στα οποία αυτή οδήγησε με τα αντίστοιχα συμπεράσματα ήδη υπάρχοντων διαγνωστικών μεθόδων.



NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF MECHANICAL ENGINEERING
FLUIDS SECTION
LAB. OF THERMAL TURBOMACHINES

Abstract:

FAULT DIAGNOSIS OF THERMAL TURBOMACHINES USING SUPPORT VECTOR MACHINES (SVM)

Christos A. Kalathakis

This thesis examines the ability of Support Vector Machines (SVM) for fault diagnosis of gas turbine engine's components.

When operating a turbine failure may occur in any component. This damage can lead to a reduction of engine performance or even destruction. A gas turbine can operate in various conditions depending on whether the engine is "healthy" or is operating in the presence of damage. Therefore, an in time and reliable fault detection and isolation is required. This is achieved through the use of diagnostic systems, based on various diagnostic techniques.

The Support Vector Machines (SVM) were invented by Vapnik and provide a data processing method with great success in many classification problems. The use of SVM in the case of classification manages to separate a pair of classes by a dividing line (or plane, or more generally hyperplane).

If each mode of a gas turbine deals as a class, then, using the SVM, we can achieve separation and to classify each new test mode in one of the classes - operating modes as separate. These are the operating principles of the diagnostic program created and implemented for various modes of operation and damages for a specific gas turbine.

Specifically, this mission is developed and described as follows:

- Principles of SVM theory.

- Proposed methodology for the application of SVM for fault diagnosis in gas turbines components.
- Application of the SVM method in realistic aerothermodynamic faults of a civil aviation turbofan engine.
- Application of the SVM method in actual aerothermodynamic and mechanical faults of gas turbines compressors.
- Evaluation of diagnostic capabilities / performance of the proposed method of SVM, by comparison of the diagnostic results with the corresponding results of other pre-existing diagnostic methods.

Ευχαριστίες...

Στο σημείο αυτό θα ήθελα να ευχαριστήσω ιδιαίτερα τον επιβλέποντα Καθηγητή Κ. Μαθιουδάκη για την ευκαιρία που μου έδωσε να εκπονήσω ένα τέτοιο ενδιαφέρον θέμα, για την καθοδήγησή του και για τις χρήσιμες συμβουλές του.

Ιδιαίτερα, ευχαριστώ τον Δρ. Χ. Ρωμέση για την υπομονή, επιμονή και συνεχή βοήθειά του για την περάτωση της παρούσας διπλωματικής εργασίας. Ακόμα, ευχαριστώ όλη την ερευνητική ομάδα του Εργαστηρίου Θερμικών Στροβιλομηχανών για την βοήθεια που μου παρείχε.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου για τις θυσίες που έχει κάνει και για τη διαρκή στήριξή της όλα αυτά τα χρόνια, όπως επίσης και τους φίλους μου για τη συμπαράστασή τους.

Καλαθάκης Α. Χρήστος
Φεβρουάριος 2010

Πίνακας Περιεχομένων

Κεφάλαιο 1: Γενικά για τις Μηχανές Διανυσμάτων Υποστήριξης.....	1
1.1 Σύντομη περιγραφή.....	1
1.2 Βιβλιογραφική ανασκόπηση και υπάρχουσες εφαρμογές.....	2
1.3 Σκοπός της παρούσας διπλωματικής εργασίας.....	5
Κεφάλαιο 2: Μηχανές Διανυσμάτων Υποστήριξης (SVM).....	7
2.1 Εισαγωγικά.....	7
2.2 Θεωρητικό υπόβαθρο.....	10
2.2.1 Σφάλμα και Αβεβαιότητα.....	10
2.2.2 Χωρητικότητα οικογένειας συναρτήσεων (VC-Dimension).....	12
2.2.3 Ελαχιστοποίηση Κατασκευαστικού Ρίσκου (SRM).....	14
2.3 Γενικές περιπτώσεις προβλημάτων.....	14
2.3.1 Γραμμικά πλήρως διαχωρίσιμες κλάσεις.....	15
2.3.2 Μη-απόλυτα γραμμικά διαχωρίσιμες κλάσεις.....	23
2.3.3 Γραμμικά μη διαχωρίσιμες κλάσεις.....	27
2.4 Συνάρτηση κελύφους (kernel function).....	30
2.4.1 Θεωρητικό υπόβαθρο.....	30
2.5 Η περίπτωση των πολλών κλάσεων (multiclassification).....	31
2.5.1 One-against-all.....	32
2.5.2 All-together.....	33
2.5.3 One-against-one.....	33
2.5.4 DAGSVM.....	35
2.5.5 Αξιολόγηση των μεθόδων.....	36
Κεφάλαιο 3: Εφαρμογή των SVMs στη διαγνωστική αεριοστροβίλων.....	37
3.1 Γενικά.....	37
3.2 Διάγνωση βλαβών αεριοστροβίλων.....	37

3.3	Ανάπτυξη των SVMs για ταξινόμηση βλαβών αεριοστροβίλου.....	39
3.3.1	Εκπαίδευση.....	39
3.3.2	Συναρτήσεις κελύφους.....	41
3.3.3	Κατάταξη.....	42
Κεφάλαιο 4: Διαγνωστική αεριοστροβίλου με χρήση SVM.....		47
4.1	Εντοπισμός βλαβών αεριοστροβίλου διπλού ρεύματος με στοιχεία θερμοδυναμικών μεγεθών.....	47
4.1.1	Διάγνωση βλαβών συνιστωσών αεριοστροβίλου.....	50
4.1.1.1	Δεδομένα Εκπαίδευσης των SVMs.....	51
4.1.1.2	Βελτιστοποίηση των παραμέτρων των SVMs.....	53
4.1.1.3	Εξέταση Βλαβών.....	55
4.1.1.4	Αποτελέσματα set B (Running Average Points) με SVM-NS.....	56
4.1.1.5	Αποτελέσματα set A (Single Points) με SVM-NS.....	59
4.1.1.6	Αποτελέσματα set B (Running Average Points) με SVM-NF.....	61
4.1.1.7	Εντοπισμός Μεγέθους Βλάβης.....	62
4.1.1.8	Σύγκριση με άλλες μεθόδους.....	65
4.1.2	Διάγνωση βλαβών αισθητήρων.....	67
4.2	Εντοπισμός βλαβών συμπιεστών με στοιχεία θερμοδυναμικών – μηχανικών μεγεθών.....	70
4.2.1	Αξονικός συμπιεστής.....	71
4.2.2	Ακτινικός συμπιεστής.....	75
Κεφάλαιο 5: Ανακεφαλαίωση – Προτάσεις – Συμπεράσματα.....		81
5.1	Ανακεφαλαίωση.....	81
5.2	Συμπεράσματα.....	82
5.3	Προτάσεις.....	83
Παράρτημα I.....		87
Π1.1	Το διαγνωστικό πρόγραμμα που δημιουργήθηκε.....	87
Π1.2	Χρησιμοποιούμενα αρχεία.....	87

Π1.3 Χρησιμοποιούμενες Μεταβλητές Εισόδου/Εξόδου.....	88
Π2.1 Ανάλυση του προγράμματος.....	90
Π2.2 Χρησιμοποιούμενες μεταβλητές από το πρόγραμμα.....	90
Π2.3 Λογικό διάγραμμα προγράμματος.....	92
Π2.4 Περιγραφή υπορουτινών.....	94
Παράρτημα II.....	101
Π1 Το πρόγραμμα επίλυσης τετραγωνικών προβλημάτων QPROG (DQPROG)....	101
Βιβλιογραφικές Αναφορές.....	107

Κεφάλαιο 1: Γενικά για τις Μηχανές Διανυσμάτων Υποστήριξης

1.1 Σύντομη περιγραφή

Οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines) αποτελούν ένα εργαλείο για την επεξεργασία δεδομένων που χρησιμοποιείται σε πολλές εφαρμογές. Τα SVMs βασίζονται στην στατιστική θεωρία εκμάθησης και αναπτύχθηκαν από τον *Vapnik (1995)*.

Τα SVMs ανήκουν σε μια οικογένεια μεθόδων που ονομάζεται μηχανές εκμάθησης. Στην οικογένεια αυτή ανήκουν και τα Τεχνητά Νευρωνικά Δίκτυα, μια μέθοδος ευρέως γνωστή και χρησιμοποιούμενη. Οι μηχανές εκμάθησης χρησιμοποιούνται για να μάθουν την αντιστοιχία δοσμένων εισόδων-εξόδων ώστε να αντιστοιχίσουν μια καινούρια είσοδο σε μια έξοδο με βάση την εκπαίδευση που τους έγινε. Έτσι και τα SVMs, εκπαιδεύονται, μαθαίνουν δηλαδή, την αντιστοιχία εισόδου-εξόδου και συνδέουν μια καινούρια είσοδο με τη σωστή έξοδο. Κάθε Μηχανή Διανυσμάτων Υποστήριξης μπορεί να μάθει να αντιστοιχεί κάποια είσοδο σε μία από συνολικά δύο εξόδους. Έτσι, στην περίπτωση των πολλών εξόδων χρησιμοποιούνται πολλές Μηχανές Διανυσμάτων Υποστήριξης.

Τα SVMs μπορούν να χρησιμοποιηθούν στη διαγνωστική ενός αεριοστροβίλου. Μια μηχανή μπορεί να παρουσιάζει διάφορες καταστάσεις λειτουργίας ανάλογα με το αν είναι “υγιής” ή αν λειτουργεί υπό την παρουσία βλάβης καθώς επίσης και με το είδος και μέγεθος της βλάβης. Η κατάσταση λειτουργίας μιας μηχανής καθορίζει την τιμή που παίρνουν ορισμένα μεγέθη τα οποία και μπορούμε να μετρήσουμε. Έτσι, ανάλογα με την τιμή αυτών των μεγεθών μπορούμε να βγάλουμε συμπέρασμα για την κατάσταση λειτουργίας της μηχανής. Τα SVMs μπορούν να εκπαιδευτούν να μάθουν την αντιστοιχία κατάσταση λειτουργίας – τιμές μετρούμενων μεγεθών. Έτσι, για κάθε νέα είσοδο (τιμές μετρούμενων μεγεθών) θα μπορούν να προσδιορίσουν την έξοδο (κατάσταση λειτουργίας).

1.2 Βιβλιογραφική ανασκόπηση και υπάρχουσες εφαρμογές

Κατά τη λειτουργία ενός αεριοστροβίλου μπορούν να εμφανιστούν βλάβες σε διάφορες συνιστώσες του. Αυτές επιφέρουν μείωση της απόδοσης της μηχανής και σε κάποιες περιπτώσεις, την καταστροφή της. Για παράδειγμα, στο θάλαμο καύσης εμφανίζονται ρωγμές στο φλογομανδύα, απορύθμιση των ακροφυσίων έγχυσης κ.α. Στο συμπιεστή, εκτός των επικαθίσεων στα πτερύγια, παρατηρείται μηχανική διάβρωσή τους ή και καταστροφή τους από αναρρόφηση αντικειμένων. Οι παραμορφώσεις στην πρώτη σειρά σταθερών πτερυγίων, όπως και η δημιουργία επιφανειακών ανωμαλιών και ρωγμών αποτελούν τις βλάβες που συναντάμε στο στρόβιλο.

Για την διάγνωση τέτοιων βλαβών, τον προσδιορισμό δηλαδή της κατάστασης λειτουργίας της μηχανής, χρησιμοποιούμε μετρήσεις μεγεθών που λαμβάνουμε από τη μηχανή. Ανάλογα με το είδος αυτών των μετρήσεων, οι διαγνωστικές μέθοδοι ταξινομούνται σε δύο κατηγορίες: Α) διαγνωστικές μέθοδοι που κάνουν χρήση αεροθερμοδυναμικών μετρήσεων και Β) διαγνωστικές μέθοδοι που κάνουν χρήση μεθόδων επεξεργασίας μετρήσεων ταχείας απόκρισης.

Οι μέθοδοι της δεύτερης περίπτωσης είναι κατάλληλες για βλάβες που σχετίζονται είτε μόνο με μηχανικά μέρη είτε με μικρές βλάβες των επιμέρους συνιστωσών που δεν επηρεάζουν κατά ανάγκη τις συνολικές επιδόσεις της μηχανής. Αναφορά και μελέτη αυτών των μεθόδων γίνεται από τον *Αρετάκης (2000)*.

Οι μέθοδοι της πρώτης περίπτωσης είναι κατάλληλες για αποτίμηση της συνολικής κατάστασης της μηχανής και για την αναγνώριση βλαβών των συνιστωσών που πραγματοποιούν τις μεταβολές κατάστασης του εργαζόμενου μέσου και καθορίζουν τη λειτουργία της μηχανής. Χρησιμοποιούν τεχνικές που ποικίλουν, από απλούς υπολογισμούς μέχρι τεχνικές ανάλυσης διέλευσης αερίου (Gas Path Analysis) και έμπειρα συστήματα (expert systems). Ανασκόπηση αυτών των μεθόδων γίνεται από τους *Σταμάτης (1990)*, *Mathioudakis (2003)*. Οι μέθοδοι αυτές, χωρίζονται σε δύο βασικές κατηγορίες: στις μεθόδους εκτίμησης και στις μεθόδους ταξινόμησης.

Οι μέθοδοι εκτίμησης χρησιμοποιούν ένα υπολογιστικό μοντέλο λειτουργίας του αεριοστροβίλου και, αξιοποιώντας τις διαθέσιμες μετρήσεις, εκτιμούν την τιμή παραμέτρων που είναι ενδεικτικές της λειτουργικής κατάστασης του αεριοστροβίλου.

Κατά κύριο λόγο, οι μέθοδοι αυτές επιλύουν ένα πρόβλημα ελαχιστοποίησης, στηριζόμενες στη μέθοδο των ελαχίστων τετραγώνων, των φίλτρων Kalman και των παραγώγων τους. Πάνω στις μεθόδους εκτίμησης έχουν γίνει αρκετές εργασίες και συνοπτικά αναφέρουμε τους: *Merrington et al (1990)*, *Urban et al (1992)*, *Doel (1992)*, *Stamatis et al (1990)*, *Tsalavoutas et al (2000a)* και *(2000b)*, *Mathioudakis et al (2001)*.

Οι μέθοδοι ταξινόμησης, ακολουθούν την εξής βασική διαδικασία: λαμβάνονται μετρήσεις αεροθερμοδυναμικών μεγεθών της εν λειτουργία μηχανής και οι τιμές τους συγκρίνονται με τις αντίστοιχες σε υγιή λειτουργία, ώστε να δημιουργηθεί η αντίστοιχη υπογραφή. Η υπογραφή αυτή στη συνέχεια συγκρίνεται με υπογραφές γνωστών βλαβών και βγαίνει συμπέρασμα για την κατάσταση λειτουργίας της μηχανής. Η σύγκριση των υπογραφών έγκειται στην ομοιότητά τους κι έτσι οι μέθοδοι αυτής της κατηγορίας διαφέρουν στην τεχνική σύγκρισης της ομοιότητας.

Βασική τεχνική αποτελούν τα Τεχνητά Νευρωνικά Δίκτυα που έχουν βρει μεγάλη εφαρμογή λόγω των αξιόλογων αποτελεσμάτων τους. Ο πιο δημοφιλής αλγόριθμος ΤΝΔ για την ταξινόμηση βλαβών αεριοστροβίλων είναι ο Back-Propagation (BP), ωστόσο και άλλοι αλγόριθμοι έχουν χρησιμοποιηθεί με αξιόλογα αποτελέσματα. Ανάμεσα στις πιο αντιπροσωπευτικές διαγνωστικές μεθόδους με χρήση ΤΝΔ διακρίνουμε τις εργασίες των *Σταυρόπουλος (1996)*, *Kanelopoulos et al. (1997)*, *Roemer (1997)*, *Gluch (1997)*, *Volponi et al. (2000)*, *Alcolk et al. (2003)*. Σε αυτή τη κατηγορία έχουμε μια άλλη προσέγγιση, αυτή της αναγνώρισης προτύπων (παλαιότερη των ΤΝΔ) η οποία έχει υιοθετηθεί για παράδειγμα από τους: *Barschdorf et al. (1986)*, *Loukis (1994)*, *Αρετάκης (2000)* και άλλους.

Μια σχετικά καινούρια τεχνική που εντάσσεται στην προσέγγιση της αναγνώρισης προτύπων και που ακόμα αναπτύσσεται και έχει αρχίσει να βρίσκει εφαρμογή στη διαγνωστική αεριοστροβίλων είναι οι Μηχανές Διανυσμάτων Υποστήριξης (SVM). Η αρχή των SVM μπορούμε να πούμε ότι έγινε με τις εργασίες των *Vapnik and Chervonenkis (1974)* πάνω στην θεωρία στατιστικής εκμάθησης (statistical learning theory) και ειδικά με την περαιτέρω ανάπτυξη αυτής της θεωρίας από τον *Vapnik (1979)*. Η αρχή της μεθόδου έγκειται στον διαχωρισμό κλάσεων (σύνολο σημείων με ίδια ιδιότητα) με επίπεδα (υπερεπίπεδα) και με βάση αυτόν το διαχωρισμό γίνεται η κατάταξη ενός νέου σημείου σε κάποια από τις κλάσεις που διαχωρίστηκαν. Η τεχνική αυτή άρχισε πολύ γρήγορα να αναπτύσσεται και από

άλλους επιστήμονες όπως οι *Anlauf and Biehl (1989)* που πρότειναν τα υπερεπίπεδα διαχωρισμού με μέγιστο διάκενο, οι *Poggio and Girosi (1990)* μελετούν τη χρήση των συναρτήσεων κελύφους και τα SVM περίπου με τη μορφή που τα γνωρίζουμε προτάθηκαν στο συνέδριο COLT το 1992 (*Boser et al. (1992)*). Η ανάπτυξη της μεθόδου συνεχίστηκε και, από τους *Cortes and Vapnik (1995)* έχουμε την μελέτη που αναπτύσσει τους διαχωριστές “χαλαρού” διακένου (soft margin classifiers) όπως επίσης και την επέκταση των SVM στα προβλήματα προσέγγισης συνάρτησης (regression) (*Vapnik (1995)*).

Τα SVM έχουν χρησιμοποιηθεί σε διάφορα προβλήματα όπως στη βιολογία, στην ιατρική, στη γενετική, στη χημεία, στην αναγνώριση εικόνων, κειμένων και χαρακτήρων, στις οικονομικές επιστήμες καθώς επίσης και στη μηχανολογία στον τομέα της διαγνωστικής. Στη βιολογία, την ιατρική και τη γενετική βλέπουμε εφαρμογές όπως η αναγνώριση ομοιότητας πρωτεϊνών με τη χρήση των SVM (*Rangwala and Karypis (2005)*), η αναγνώριση λειτουργίας πρωτεΐνης σύμφωνα με τη δομή της (*Kim and Park (2003)*), η επιλογή χρήσης χημειοθεραπείας σε καρκίνο του στήθους (*Lee et al. (2001)*). Στα προβλήματα αναγνώρισης εικόνων και γενικά προτύπων έχουμε πληθώρα εργασιών και συνοπτικά αναφέρουμε την αναγνώριση έκφρασης προσώπου (*Ghent and McDonald (2005a), (2005b), (2005c)*), την κατάταξη κειμένων σύμφωνα με το περιεχόμενο (*Dumais et al. (1998)*) κ.α.

Χρήση των SVMs έχει γίνει και στη διαγνωστική. Χρησιμοποιήθηκαν για την διάγνωση βλάβης στην περίπτωση μορφοποίησης φύλλου μετάλλου με πρέσα (*Ge et al. 2003*) όπως και για τη διάγνωση βλάβης λόγω σπασίματος των στοιχείων του κλωβού επαγωγικού κινητήρα (*Kurek and Osowski (2009)*). Οι *Cho et al. (2006)* χρησιμοποίησαν τα SVMs σε συνδυασμό με τα ΤΝΔ για να διαγνώσουν βλάβες σε μετασχηματιστές ισχύος.

Τα SVMs στους αεριοστροβίλους χρησιμοποιούνται σχεδόν αποκλειστικά για τη διάγνωση βλαβών. Πάνω σε αυτό το αντικείμενο έχουν γίνει μελέτες και εργασίες που δείχνουν τα αποτελέσματα της χρήσης μόνο των SVM καθώς και την χρήση τους με άλλες διαγνωστικές τεχνικές. Οι *Matuck et al. (2009)* με την εργασία τους συνδύασαν την τεχνική των ΤΝΔ με άλλες τεχνικές, όπως τα SVM, για βελτιστοποίηση των αποτελεσμάτων για τον εντοπισμό και προσδιορισμό πολλαπλών βλαβών σε αεριοστροβίλους. Στην εργασία τους αναφέρουν ότι η χρήση των SVM επιτυγχάνει πάντα την εύρεση της βέλτιστης λύσης στην εκπαίδευση και δεν απαιτείται η γνώση των νευρωνικών δικτύων στο κρυφό πλέγμα αφού προσδιορίζεται

από τα SVM. Τέλος, διαπιστώνουν πως υπάρχουν πολύ λίγες εργασίες πάνω στη διαγνωστική αεριοστροβίλων με χρήση των SVM και για αυτό απαιτείται περαιτέρω μελέτη. Οι *Cook and Valentine (2008)* ανέπτυξαν μια διαδικασία που περιλαμβάνει διαγνωστικά συστήματα (ανάμεσά τους και τα SVM) και εξετάζει σε πραγματικό χρόνο το επίπεδο ετοιμότητας των αεριοστροβίλων που υπάρχουν σε στρατιωτικά πλοία. Και εδώ τα SVM χρησιμοποιούνται σαν μέρος ενός διαγνωστικού συστήματος της λογικής “Fusion”, δηλαδή τη χρήση πολλών τεχνικών για την εξαγωγή συμπεράσματος. Οι *Lee et al. (2008)* ανέπτυξαν τη μελέτη τους πάνω στη διαγνωστική αεριοστροβίλου που κινεί μη επανδρωμένο αεροσκάφος. Χρησιμοποιούνται οι τεχνικές των ΤΝΔ και των SVM για την υλοποίηση της διάγνωσης. Υπάρχουν ακόμα και άλλες εργασίες πάνω στη διαγνωστική των αεριοστροβίλων οι οποίες χρησιμοποιούν τα SVM σαν κομμάτι ενός διαγνωστικού συστήματος και όχι σαν διαγνωστικό σύστημα (*Mitchell et al. (2008)*, *Donat et al. (2007)*, *Rausch et al. (2005)*). Βλέπουμε δηλαδή τη συνεργασία πολλών διαγνωστικών μεθόδων σε ένα διαγνωστικό πρόγραμμα.

1.3 Σκοπός της παρούσας διπλωματικής εργασίας

Οι Μηχανές Διανυσμάτων Υποστήριξης, όπως είδαμε έχουν βρεί εφαρμογή σε μεγάλο εύρος προβλημάτων αφού αποτελούν μια πρόσφατη μέθοδο που διαρκώς αναπτύσσεται, βελτιώνεται και συνδυάζεται με άλλες μεθόδους.

Η χρήση τους στη διαγνωστική αεριοστροβίλων, αν και έχει πραγματοποιηθεί, είναι σε πρώιμο στάδιο. Η παρούσα διπλωματική ως σκοπό έχει να εφαρμόσει τη μέθοδο των Μηχανών Διανυσμάτων Υποστήριξης στη διαγνωστική αεριοστροβίλων.

Σκοπός είναι να “μάθουν” τα SVMs την αντιστοιχία μετρούμενων μεγεθών και κατάστασης λειτουργίας έτσι ώστε να μπορεί να καταταχθεί η εξεταζόμενη κατάσταση της μηχανής σε μια από τις υπάρχουσες καταστάσεις λειτουργίας. Αυτό υλοποιείται ως εξής:

- για συγκεκριμένες καταστάσεις λειτουργίας λαμβάνουμε τις τιμές των μετρούμενων μεγεθών

- εκπαιδεύουμε τα SVMs να μάθουν την αντιστοιχία τιμές μετρούμενων μεγεθών-κατάσταση λειτουργίας (ουσιαστικά βρίσκουμε τα όρια στα οποία βρίσκεται η κάθε κατάσταση λειτουργίας και έτσι τις διαχωρίζουμε).
- μια νέα, εξεταζόμενη κατάσταση της μηχανής αντιπροσωπεύεται από τις τιμές των μετρούμενων μεγεθών και τα SVMs την κατατάσσουν σε μία από τις καταστάσεις λειτουργίας με τις οποίες έχουν εκπαιδευτεί.

Με βάση την παραπάνω λογική δημιουργήθηκε κώδικας σε Fortran 77 ο οποίος υλοποιεί την λογική των SVMs και πραγματοποιεί διάγνωση σε μια μηχανή αεριοστροβίλου.

Συνοπτικά, η δομή της παρούσας διπλωματικής εργασίας είναι η εξής:

Στο 2^ο κεφάλαιο γίνεται μια εισαγωγή στα SVMs και αναπτύσσονται οι σχέσεις καθώς και το μαθηματικό υπόβαθρο που τα διέπουν.

Στη συνέχεια, στο 3^ο κεφάλαιο, περιγράφεται η μεθοδολογία της υλοποίησης του διαγνωστικού συστήματος αεριοστροβίλου με τη χρήση των Μηχανών Διανυσμάτων Υποστήριξης.

Στο κεφάλαιο 4 παρουσιάζονται τα δεδομένα που χρησιμοποιήθηκαν για την εφαρμογή του διαγνωστικού προγράμματος καθώς και τα αποτελέσματα της διάγνωσης και συγκρίνονται με αυτά άλλων μεθόδων.

Τέλος, στο 5^ο κεφάλαιο γίνεται μια σύντομη ανακεφαλαίωση της παρούσας εργασίας, περιγράφονται τα συμπεράσματα που εξήχθησαν και προτείνονται τρόποι για βελτίωση της μεθόδου.

Η εργασία συνοδεύεται από δύο παραρτήματα. Στο ένα περιγράφεται το διαγνωστικό πρόγραμμα που δημιουργήθηκε και στο άλλο παρουσιάζεται η λειτουργία της υπορουτίνας DQPROG η οποία επιλύει τετραγωνικά προβλήματα υπό περιορισμούς.

Κεφάλαιο 2: Μηχανές Διανυσμάτων Υποστήριξης (SVM)

2.1 Εισαγωγικά

Οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines - SVMs) είναι μια τεχνική η οποία ανήκει στην ομάδα των μηχανών εκμάθησης (learning machines) και ως στόχο έχει την επεξεργασία δεδομένων. Χρησιμοποιείται σε προβλήματα ταξινόμησης και στην προσέγγιση της μορφής της συνάρτησης σε προβλήματα παλινδρόμησης.

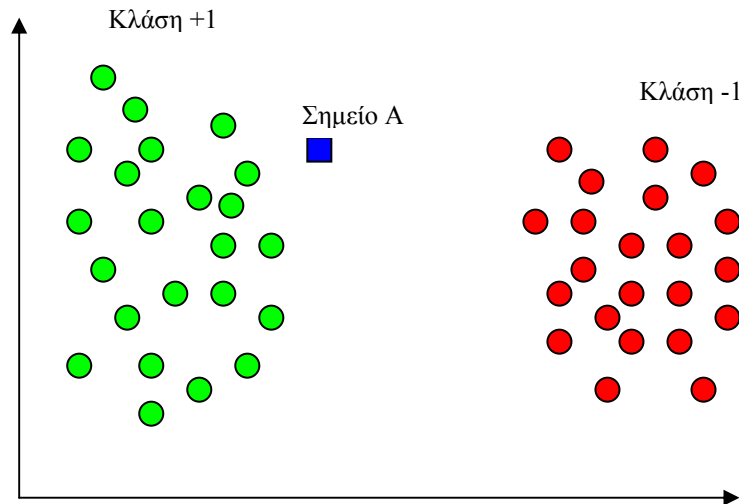
Η λογική μια μηχανής εκμάθησης είναι να δίνει την τιμή y_i μιας συνάρτησης (άγνωστη προς εμάς) που αντιστοιχεί σε δοσμένο σημείο \vec{x}_i . Αυτό γίνεται ως εξής:

Για δεδομένο σύνολο l σημείων $\vec{x}_i \in \mathbb{R}^N$ και έχοντας τις αντίστοιχες τιμές $y_i \in \mathbb{R}$ που παίρνει η άγνωστη συνάρτηση, εκπαιδεύουμε τη μηχανή εκμάθησης να μάθει τη σχέση που συνδέει τα \vec{x}_i με τα y_i . Δηλαδή, η μηχανή μαθαίνει την αντιστοίχιση $\vec{x}_i \rightarrow y_i$ και έτσι για ένα σημείο \vec{x}_m , διαφορετικό από αυτά του συνόλου l της εκμάθησης, θα μας δώσει την τιμή y_m που θα έπαιρνε η άγνωστη συνάρτηση.

Στην περίπτωση ταξινόμησης με τα SVM, το σύνολο των σημείων l αποτελείται από δύο υποσύνολα τα k και n . Έτσι, το αποτέλεσμα της συνάρτησης θα είναι $+1$ ή -1 ($y_i=+1$ ή $y_i=-1$) ανάλογα σε ποιο υποσύνολο ανήκει το δοθέν σημείο \vec{x}_i . Τα δύο αυτά υποσύνολα ονομάζονται κλάσεις και η τιμή $+1$ (-1) είναι η “ετικέτα” της κλάσης. Δηλαδή, σε αυτή τη περίπτωση τα SVM μαθαίνουν να κατατάσσουν σωστά τα σημεία \vec{x}_i στις δύο κλάσεις.

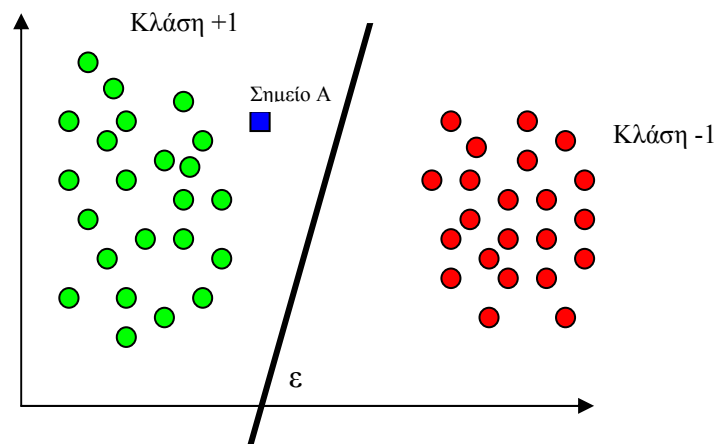
Τα σημεία \vec{x}_i και οι αντίστοιχες τιμές τους, y_i , αποτελούν την πληροφορία εκπαίδευσης (training set). Τα σημεία \vec{x}_i ονομάζονται πρότυπα εκπαίδευσης (training patterns) ενώ οι τιμές y_i που αντιστοιχούν σε αυτά, στόχοι εκπαίδευσης (training targets).

Στο παράδειγμα στο Σχήμα 2.1 βλέπουμε ένα σύνολο σημείων \bar{x}_i στο δισδιάστατο χώρο, καθώς και την κλάση στην οποία ανήκουν (+1 ή -1). Στόχος μας είναι να εκπαιδεύσουμε τα SVM ώστε να μάθουν την σχέση σημείου – κλάσης για τα δοσμένα σημεία, έτσι ώστε, για ένα άλλο σημείο (έστω A) να μας προβλέψουν σε ποία κλάση ανήκει.



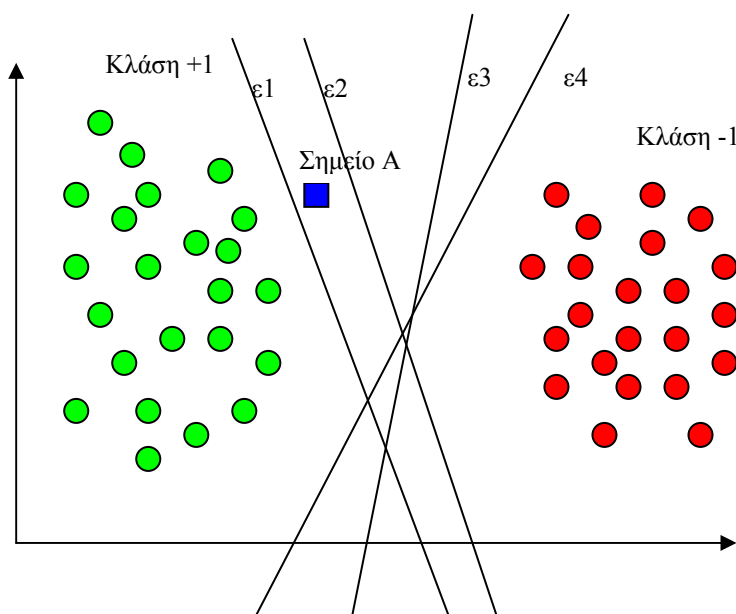
Σχήμα 2.1. Τυχαίο σύνολο σημείων δύο κλάσεων και τυχαίο σημείο προς κατάταξη.

Για να αποφασιστεί σε ποια κλάση ανήκει το νέο σημείο, θα πρέπει να βρεθεί που είναι το όριο της κάθε κλάσης, δηλαδή να βρεθεί μια γραμμή (δισδιάστατος χώρος) που να διαχωρίζει τις δύο κλάσεις. Τα SVMs, για να πετυχαίνουν το διαχωρισμό των κλάσεων χρησιμοποιούν ευθείες. Έτσι, από τη σχετική θέση του προς κατάταξη σημείου και της διαχωριστικής ευθείας θα μπορεί να βγει το συμπέρασμα σε ποία κλάση ανήκει το σημείο. Στο Σχήμα 2.2 βλέπουμε μια διαχωριστική ευθεία (ϵ). Επειδή το σημείο A βρίσκεται στα αριστερά της ευθείας, εκεί δηλαδή που υπάρχει η κλάση +1, συμπεραίνουμε πως το σημείο A ανήκει στην κλάση +1.



Σχήμα 2.2. Διαχωρισμός τυχαίων κλάσεων και κατάταξη τυχαίου σημείου.

Από τα παραπάνω, το βασικό ερώτημα που γεννάται είναι ποία θα ναι η διαχωριστική ευθεία. Στο Σχήμα 2.3 βλέπουμε μια ομάδα ευθειών που όλες διαχωρίζουν τις δύο κλάσεις, αλλά ποια είναι η καλύτερη; Στο απλό αυτό παράδειγμα βλέπουμε πως το αποτέλεσμα της κατάταξης του σημείου A εξαρτάται άμεσα από την ευθεία διαχωρισμού. Στην περίπτωση της ευθείας ϵ_1 το σημείο κατατάσσεται στην κλάση -1, ενώ στις άλλες περιπτώσεις στην κλάση +1.



Σχήμα 2.3. Πολλαπλός διαχωρισμός τυχαίων κλάσεων και κατάταξη σημείου.

Με βάση τα παραπάνω, στόχος μας είναι να προσδιοριστεί η βέλτιστη ευθεία διαχωρισμού των κλάσεων, για την οποία θα επιτυγχάνεται η ελαχιστοποίηση του σφάλματος κατάταξης. Δηλαδή, να καταταχθεί σωστά, στην κλάση που πραγματικά ανήκει, όσο το δυνατόν μεγαλύτερος αριθμός σημείων. Για να το πετύχουμε αυτό, μια λογική υπόθεση θα ήταν να απαιτήσουμε η διαχωριστική μας ευθεία (που καθορίζει το όριο των κλάσεων) να μην βρίσκεται κοντά στα δεδομένα σημεία των κλάσεων.

2.2 Θεωρητικό υπόβαθρο

Στη συνέχεια, γίνεται μια ανάλυση για την προσέγγιση του σφάλματος και της αβεβαιότητας και τρόπος περιορισμού αυτών, έτσι ώστε να βγάλουμε συμπεράσματα και σχέσεις για την υλοποίηση της μεθόδου των Μηχανών Διανυσμάτων Υποστήριξης.

2.2.1 Σφάλμα και Αβεβαιότητα

Έστω ότι έχουμε ένα σύνολο εκπαίδευσης μεγέθους l σημείων και τις αντίστοιχες ετικέτες τους y_i . Θεωρούμε ότι υπάρχει μια (άγνωστη) κατανομή πιθανότητας $P(\bar{x}_i, y_i)$ από την οποία έχει προέλθει το σύνολο εκπαίδευσης. Αυτό σημαίνει ότι έχουμε μια κατανομή των y για δεδομένο x . Αυτή η υπόθεση είναι πιο γενική από το να θεωρήσουμε συγκεκριμένο y για κάθε x . Σε αυτή τη περίπτωση οι ετικέτες y αντιστοιχίζονται στα σημεία x μέσω μιας συγκεκριμένης κατανομής P . Η υπόθεση της ύπαρξης κατανομής πιθανότητας είναι γενική θεώρηση. Στις περιπτώσεις των προβλημάτων που αντιμετωπίζουμε πρακτικά υπάρχει συγκεκριμένη αντιστοιχία $\bar{x}_i \rightarrow y_i$.

Στόχος μας είναι τώρα, να εκπαιδεύσουμε τη μηχανή εκμάθησης ώστε να “μάθει” την αντιστοιχία $\bar{x}_i \rightarrow y_i$, δηλαδή να κατατάσσει σωστά τα \bar{x}_i τα οποία έχουν προέλθει από την ίδια κατανομή πιθανότητας $P(\bar{x}_i, y_i)$ με αυτά του συνόλου εκπαίδευσης. Ή πιο απλά να βρεθεί μια συνάρτηση f τέτοια ώστε: $\bar{x}_i \rightarrow f(\bar{x}_i, \vec{a}) = y_i$.

Η παράμετρος a ονομάζεται ρυθμιστική παράμετρος της συνάρτησης και για επιλεγμένο a και για δεδομένα \bar{x}_i η συνάρτηση f θα μας δίνει πάντα το ίδιο αποτέλεσμα (ντετερμινιστικά).

Τώρα θα οριστούν δύο ποσότητες, το αναμενόμενο/πραγματικό ρίσκο (expected/actual risk) και εμπειρικό ρίσκο (empirical risk), που συνδέονται με το σφάλμα, δηλαδή τις επιδόσεις της μηχανής εκμάθησης.

Το αναμενόμενο ή πραγματικό ρίσκο (*Burges (1998)*) ορίζεται ως:

$$R(a) = \int \frac{1}{2} |y - f(x, a)| dP(x, y) \quad (2.1)$$

και μας δείχνει την απόκλιση της εκτίμησης κατάταξης $f(\bar{x}_i, \bar{a})$ από την πραγματική y στα νέα σημεία (εκτός του συνόλου εκπαίδευσης).

Το εμπειρικό ρίσκο ορίζεται ως:

$$R_{emp}(a) = \frac{1}{2l} \sum_{i=1}^l |y_i - f(x_i, a)| \quad (2.2)$$

και ουσιαστικά αποτελεί το μέσο όρο του σφάλματος κατάταξης των l σημείων του συνόλου εκπαίδευσης.

Από τα παραπάνω βλέπουμε πως το πραγματικό ρίσκο αναφέρεται στο σφάλμα της μηχανής σε νέα σημεία \bar{x}_i , εκτός του συνόλου εκπαίδευσης, ενώ το εμπειρικό ρίσκο αναφέρεται μόνο στα σημεία του συνόλου εκπαίδευσης. Όμως, το πραγματικό ρίσκο δεν μπορεί να υπολογιστεί διότι στη σχέση ορισμού του βλέπουμε πως υπάρχει η άγνωστη κατανομή πιθανότητας $P(\bar{x}_i, y_i)$ ενώ αντίθετα στο εμπειρικό ρίσκο δεν έχουμε τέτοιο πρόβλημα.

Η θεώρηση ότι η ελαχιστοποίηση του εμπειρικού ρίσκου θα επιφέρει και μείωση του πραγματικού ρίσκου είναι η αρχή της θεωρίας της Ελαχιστοποίησης Εμπειρικού Ρίσκου (Empirical Risk Minimization, ERM) η οποία συνοψίζεται στην εξής σχέση: $\min R_{emp}(a) \rightarrow \min R(a)$ (*Burges (1998)*). Η θεωρία αυτή αποτελεί την αρχή λειτουργίας πολλών μηχανών εκμάθησης όπως των Τεχνητών Νευρωνικών Δικτύων.

Στην πραγματικότητα όμως, η σχέση μεταξύ του εμπειρικού και του πραγματικού ρίσκου είναι μια ανισότητα της μορφής (*Burges (1998)*):

$$R(a) \leq R_{emp}(a) + Z \quad (2.3)$$

Στην σχέση (2.2) του εμπειρικού ρίσκου υπάρχει η ποσότητα $\frac{1}{2} |y_i - f(x_i, a)|$ (απώλεια - loss) η οποία στην περίπτωση ταξινόμησης κλάσεων με τη μέθοδο SVM παίρνει τιμές 0 και 1. Κατά τον *Burges (1998)*, θεωρούμε μια παράμετρο n τέτοια ώστε: $0 \leq n \leq 1$. Η πιθανότητα για την οποία η ποσότητα $\frac{1}{2} |y_i - f(x_i, a)|$ παίρνει τις τιμές 0 και 1, είναι $1-n$. Έτσι, έχουμε την εξής μορφή της παραμέτρου Z της σχέσης (2.3) (*Vapnik (1995)*):

$$Z = \sqrt{\left(\frac{h(\log(2l/h) + 1 - \log(n/4))}{l}\right)} \quad (2.4)$$

Με βάση τα παραπάνω η ανισότητα που συνδέει το εμπειρικό με το πραγματικό ρίσκο (σχέση 2.3) θα είναι η:

$$R(a) \leq R_{emp}(a) + \sqrt{\left(\frac{h(\log(2l/h) + 1 - \log(n/4))}{l}\right)} \quad (2.5)$$

Το δεξί μέλος της ανισότητας ονομάζεται όριο ρίσκου (risk bound) και ο δεύτερος όρος αυτού του μέλους ονομάζεται VC-confidence εξαιτίας του όρου h (θετικός και ακέραιος αριθμός) που περιέχει και ο οποίος ονομάζεται VC-dimension (ή χωρητικότητα οικογένειας συναρτήσεων) από τους Vapnik – Chervonenkis που εισήγαγαν τον όρο αυτό (*Burges (1998)*).

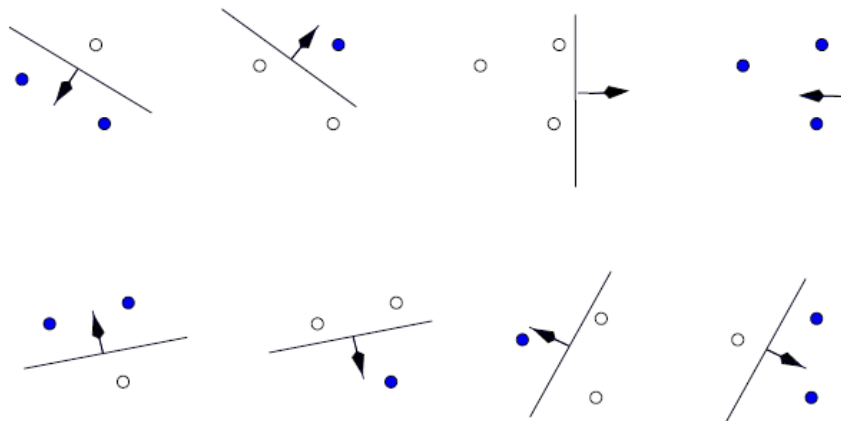
2.2.2 Χωρητικότητα οικογένειας συναρτήσεων (VC-Dimension)

Θεωρούμε πάλι το γνωστό σύνολο σημείων \vec{x}_i μεγέθους l και τις αντίστοιχες ετικέτες y_i (+1,-1) των σημείων, που υποδηλώνουν σε ποια κλάση εκ των δύο ανήκει

το κάθε σημείο. Για αυτό το σύνολο των l σημείων και των δύο κλάσεων έχουμε 2^l πιθανούς συνδυασμούς κατάταξης. Αν μια οικογένεια συναρτήσεων $f(\vec{x}, \vec{a})$ έχει ως μέλη συναρτήσεις οι οποίες (για κάποιο α) μπορούν να κατατάξουν σωστά όλα τα σημεία \vec{x}_i (για κάθε έναν από τους 2^l πιθανούς συνδυασμούς), τότε λέμε ότι η οικογένεια αυτή σαρώνει επιτυχώς το σύνολο l των σημείων και ότι έχει χωρητικότητα l .

Γενικά, χωρητικότητα μιας οικογένειας συναρτήσεων (VC-dimension) ονομάζουμε τον μέγιστο αριθμό σημείων ενός συνόλου που μπορούν να διαχωριστούν με όλους τους δυνατούς τρόπους από τα μέλη της οικογένειας αυτής.

Στο Σχήμα 2.4 βλέπουμε μια ομάδα τριών σημείων και τους δυνατούς τρόπους διαχωρισμού τους (*Burges (1998)*). Το χρώμα των σημείων υποδηλώνει σε κάθε περίπτωση την κλάση που ανήκουν τα σημεία. Η κατεύθυνση του βέλους της διαχωριστικής γραμμής των σημείων, μας δείχνει ότι τα σημεία που βρίσκονται από αυτή τη πλευρά της γραμμής θα καταταχθούν στην κλάση +1 ενώ τα άλλα στην κλάση -1. Η οικογένεια αυτών των συναρτήσεων (ευθείες διαχωρισμού) διαχωρίζει με όλους τους δυνατούς συνδυασμούς τα τρία σημεία και άρα λέμε ότι έχει χωρητικότητα 3.



Σχήμα 2.4. Δυνατοί διαχωρισμοί τριών σημείων στο επίπεδο (*Burges (1998)*)

Είναι φανερό ότι η οικογένεια αυτή των συναρτήσεων διαχωρίζει με κάθε δυνατό συνδυασμό τα τρία σημεία. Αν ακολουθήσουμε την ίδια διαδικασία για τέσσερα σημεία, δεν έχουμε την ίδια συμπεριφορά αφού οι δυνατοί θεωρητικά

συνδυασμοί είναι $2^4=16$ ενώ θα παρατηρήσουμε πως η οικογένεια αυτή καταφέρνει σωστό διαχωρισμό μόνο 14 συνδυασμών.

2.2.3 Ελαχιστοποίηση Κατασκευαστικού Ρίσκου (SRM)

Από την ανισότητα που συνδέει το πραγματικό με το εμπειρικό ρίσκο (σχέση 2.5) βλέπουμε πως για να ελαχιστοποιήσουμε το πραγματικό ρίσκο (σφάλμα) πρέπει να ελαχιστοποιήσουμε το εμπειρικό ρίσκο αλλά με τη χρήση της κατάλληλης οικογένειας συναρτήσεων της οποίας η χωρητικότητα VC είναι η αρμόζουσα για το δοσμένο σύνολο. Τα παραπάνω αποτελούν την αρχή της Ελαχιστοποίησης Κατασκευαστικού Ρίσκου (Structural Risk Minimization – SRM) (*Burges (1998)*).

Επειδή είναι δύσκολο να βρεθεί η κατάλληλη συνάρτηση από μια οικογένεια συναρτήσεων για να χρησιμοποιηθεί, χωρίζεται η οικογένεια αυτή σε “ομάδες” συναρτήσεων που μπορούμε να υπολογίσουμε την χωρητικότητά τους ή έστω ένα άνω όριο της τιμής της χωρητικότητάς τους. Οπότε τώρα, η εφαρμογή της μεθόδου SRM έγκειται στον να βρεθεί ποια “ομάδα” ελαχιστοποιεί το άνω όριο του πραγματικού ρίσκου.

Ο αλγόριθμος ταξινόμησης που χρησιμοποίησαν οι Vapnik – Chervonenkis, (δηλαδή οι Μηχανές Διανυσμάτων Υποστήριξης) βασίζεται στην αρχή της SRM και χρησιμοποιεί μια τέτοια οικογένεια συναρτήσεων ώστε να επιτυγχάνεται η μεγιστοποίηση της απόστασης της διαχωριστικής ευθείας από τις δύο κλάσεις ταυτόχρονα.

2.3 Γενικές περιπτώσεις προβλημάτων

Ζητούμενο από τα SVM είναι να μπορέσουν να βρουν την γραμμή αυτή (ή επίπεδο αν $\vec{x}_i \in \mathcal{R}^3$ ή υπερεπίπεδο (hyperplane) αν $\vec{x}_i \in \mathcal{R}^N$, $N \geq 4$) που διαχωρίζει βέλτιστα τις κλάσεις, ώστε να γίνεται εύκολα η κατάταξη νέων σημείων.

Τα προβλήματα που αντιμετωπίζονται με τις Μηχανές Διανυσμάτων Υποστήριξης μπορούν να ενταχθούν σε τρεις κατηγορίες:

- α) Γραμμικά πλήρως διαχωρίσιμες κλάσεις, όπου τα σημεία εκπαίδευσης ανήκουν σε κλάσεις που διαχωρίζονται πλήρως με μία ευθεία γραμμή (ή πιο γενικά υπερεπίπεδο)
- β) Μη-απόλυτα γραμμικά διαχωρίσιμες κλάσεις, όπου τα σημεία εκπαίδευσης ανήκουν σε κλάσεις που δεν διαχωρίζονται απόλυτα με μια ευθεία γραμμή (ή υπερεπίπεδο). Δηλαδή, οι κλάσεις έχουν μια περιοχή που αλληλεπικαλύπτονται.
- γ) Γραμμικά μη διαχωρίσιμες κλάσεις, όπου σε αυτή τη περίπτωση οι κλάσεις δεν μπορούν να διαχωριστούν με ευθεία (ή υπερεπίπεδο). Ένα τέτοιο παράδειγμα είναι όταν τα σημεία της μίας κλάσης βρίσκονται γύρω από τα σημεία της άλλης κλάσης.

Στη συνέχεια θα γίνει ανάλυση των τριών περιπτώσεων και θα βρεθούν οι μαθηματικές εκφράσεις που διέπουν την κάθε περίπτωση.

2.3.1 Γραμμικά πλήρως διαχωρίσιμες κλάσεις

Έστω η απλή δισδιάστατη περίπτωση ενός συνόλου εκπαίδευσης μεγέθους l (\vec{x}_i, y_i) . Όπως είδαμε και στο Σχήμα 2.3 υπάρχουν πολλές ευθείες που θα μπορούσαν να διαχωρίσουν τις δύο κλάσεις. Εμείς όμως ζητούμε τη βέλτιστη, δηλαδή αυτή που θα υπόκειται στην αρχή της μεθόδου SRM.

Όλες αυτές οι διαχωριστικές ευθείες είναι της γενικής μορφής $\vec{w} \cdot \vec{x} + b = 0$, όπου \vec{x} τα σημεία του συνόλου εκπαίδευσης (σημεία των κλάσεων προς διαχωρισμό), \vec{w} το διάνυσμα προσανατολισμού της διαχωριστικής ευθείας (κάθετο στην ευθεία) και b μια σταθερά. Η ποσότητα $|b| / \|\vec{w}\|$ είναι η απόσταση της ευθείας διαχωρισμού από την αρχή των αξόνων. Έστω d_+ μια μικρή απόσταση από την ευθεία διαχωρισμού προς την πλευρά της κλάσης +1 και d_- μια μικρή απόσταση προς την πλευρά της κλάσης -1. Θεωρούμε δυο ευθείες εκατέρωθεν της διαχωριστικής σε αποστάσεις d_+ και d_- αντίστοιχα και παράλληλες προς αυτήν. Η βέλτιστη διαχωριστική ευθεία θα είναι αυτή που θα απέχει περισσότερο ταυτόχρονα και από τις δύο κλάσεις (από τα ακραία σημεία τους). Δηλαδή θέλουμε να μεγιστοποιήσουμε την απόσταση $(d_+ + d_-)$ που ονομάζεται διάκενο ή διαχωριστική ζώνη (margin) (Burges (1998)). Όπως γίνεται εύκολα αντιληπτό στην περίπτωση του μέγιστου margin, οι δύο παράλληλες ευθείες εκατέρωθεν της διαχωριστικής θα κείτονται επί του ακραίου (ακραίων) σημείου της κάθε κλάσης. Έτσι, όσα σημεία βρίσκονται από

τη μια πλευρά της ευθείας προς την κλάση +1 θα ανήκουν σε αυτή τη κλάση. Αντίστοιχα για την κλάση -1. Οι εξισώσεις των δύο αυτών ευθειών θα είναι:

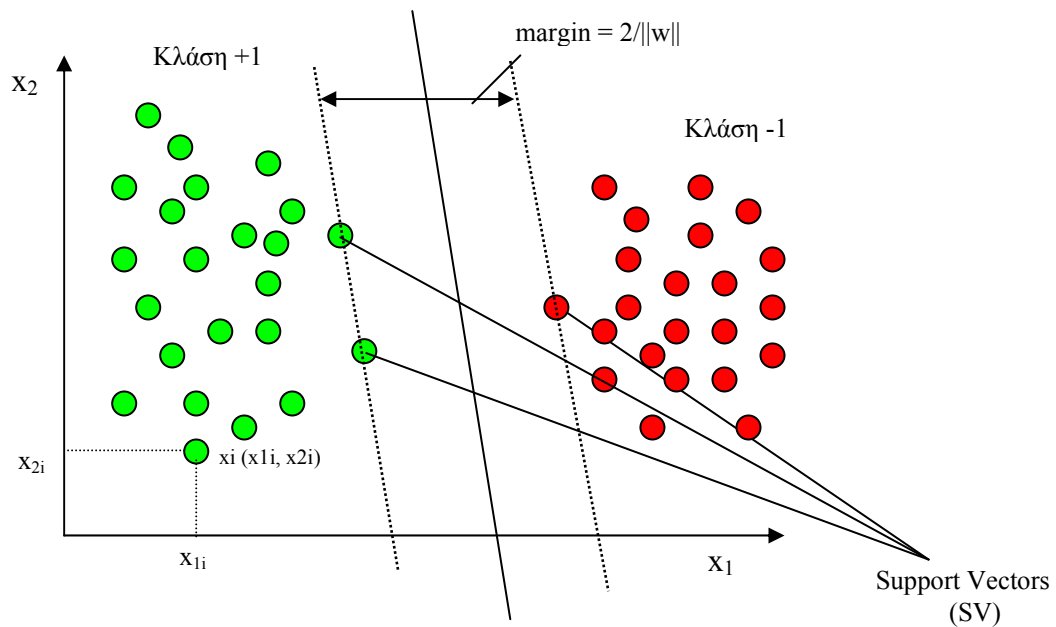
$$\bar{w} \cdot \bar{x} + b = +1, y = +1 \quad (2.6)$$

$$\bar{w} \cdot \bar{x} + b = -1, y = -1 \quad (2.7)$$

και η εξίσωση της διαχωριστικής ευθείας θα είναι:

$$\bar{w}^* \cdot \bar{x} + b^* = 0 \quad (2.8)$$

Στο Σχήμα 2.5 οπτικοποιούνται όλα τα παραπάνω, για διανύσματα διαστάσεως 2, για να γίνουν πιο κατανοητά.



Σχήμα 2.5. Οπτικοποίηση εννοιών της μεθόδου.

Με βάση τα παραπάνω, τα σημεία για τα οποία ισχύει $\bar{w} \cdot \bar{x} + b \geq +1$ θα ανήκουν στην κλάση που έχει ετικέτα +1, ενώ αυτά για τα οποία ισχύει $\bar{w} \cdot \bar{x} + b \leq -1$ θα ανήκουν στην κλάση με ετικέτα -1. Οι δύο αυτές περιπτώσεις μπορούν να παρουσιαστούν με μία ενιαία μορφή:

$$y_i(\bar{w} \cdot \bar{x}_i + b) - 1 \geq 0, \quad \forall i \quad (2.9)$$

Τα ακραία σημεία των κλάσεων επί των οποίων κείτονται οι δύο ευθείες ονομάζονται Διανύσματα Υποστήριξης (Support Vectors) διότι αυτά ουσιαστικά υπαγορεύουν την θέση των ευθειών.

Η απόσταση ($d_+ + d_-$) με βάση τα παραπάνω μπορεί να εκφραστεί με τη μορφή $\frac{2}{\|\bar{w}\|}$ όπου $\|\bar{w}\|$ η ευκλείδεια νόρμα του διανύσματος προσανατολισμού \bar{w} .

Δηλαδή, ζητούμενο είναι η μεγιστοποίηση της ποσότητας $\frac{2}{\|\bar{w}\|}$ ή η ελαχιστοποίηση του $\|\bar{w}\|$ ή η ελαχιστοποίηση του $1/2\|\bar{w}\|^2$ υπό τους περιορισμούς της σχέσης (2.9).

Στην περίπτωση αυτή των πλήρως διαχωρίσιμων κλάσεων έχουμε μηδενικό εμπειρικό ρίσκο αφού όλα τα σημεία του συνόλου εκπαίδευσης κατατάσσονται σωστά λόγω της πλήρους διαχωρισιμότητας των κλάσεων. Έτσι από την σχέση (2.5) αρκεί να ελαχιστοποιήσουμε τον δεύτερο όρο του δεξιού μέλους της, δηλαδή την VC-confidence ($\sqrt{\frac{h(\log(2l/h) + 1 - \log(n/4))}{l}}$). Αποδεικνύεται ότι όσο μεγαλύτερο είναι το διάκενο (margin), τόσο μικρότερη είναι η VC-dimension (h) και έτσι ικανοποιείται η αρχή της SRM (Vapnik (1998)).

Όπως προαναφέρθηκε, στόχος μας είναι να πετύχουμε μέγιστο διάκενο δηλαδή να ελαχιστοποιήσουμε την ποσότητα:

$$1/2\|\bar{w}\|^2 = \frac{1}{2} \sum_{i=1}^l w_i \cdot w_i \quad (2.10)$$

υπό τους περιορισμούς της σχέσης (2.9). Το πρόβλημα αυτό αποτελεί χαρακτηριστική περίπτωση προβλήματος τετραγωνικού προγραμματισμού (Quadratic Programming Problem – QP). Για την επίλυση αυτού του είδους των προβλημάτων μια μέθοδος (η οποία και θα χρησιμοποιηθεί) είναι ο μετασχηματισμός Lagrange (Borges (1998)) διότι στο κατά Lagrange μετασχηματισμένο πρόβλημα οι περιορισμοί είναι πιο εύκολα αντιμετωπίσιμοι και γιατί, όπως θα φανεί παρακάτω, θα μας είναι πιο εύκολα αντιμετωπίσιμο το πρόβλημα των μή γραμμικά διαχωρίσιμων κλάσεων. Η αρχή του μετασχηματισμού Lagrange για εξισώσεις με ανισοτικούς περιορισμούς είναι η εξής: Έστω μια συνάρτηση $F_0(x)$ με περιορισμούς $f_i(x) \geq 0$ την

οποία θέλουμε να μεγιστοποιήσουμε. Ο κατά Lagrange μετασχηματισμός της, δίνει:
 $L = F_0(x) - \sum a_i f_i(x)$.

Έτσι με βάση τη θεωρία θα εισάγουμε πολλαπλασιαστές Lagrange ($a_i \geq 0$) οι οποίοι θα είναι όσα και τα σημεία μας, δηλαδή l .

Με βάση τα παραπάνω, για τη συνάρτηση (2.10) υπό τους περιορισμούς (2.9) θα έχουμε:

$$\begin{aligned} L &= \frac{1}{2} \|\vec{w}\|^2 - \sum_{i=1}^l \alpha_i [y_i (w \cdot x + b) - 1] = \\ &= \frac{1}{2} \sum_{i=1}^l w_i w_i - \sum_{i=1}^l \alpha_i [y_i (\sum_{j=1}^l w_j \cdot x_j + b) - 1] \end{aligned} \quad (2.11)$$

Ακόμα, η βέλτιστη λύση θα πρέπει να ικανοποιεί τις συνθήκες Karush-Kuhn-Tucker (KKT) (*Burges (1998), Kecman (2001)*). Η ικανοποίηση των συνθηκών KKT είναι απαραίτητη για να θεωρηθεί βέλτιστη μια λύση σε μη γραμμικά προβλήματα. Οι συνθήκες KKT αποτελούν μια γενίκευση της μεθόδου του κατά Lagrange μετασχηματισμού σε ένα πρόβλημα με ανισοτικούς περιορισμούς. Οι συνθήκες KKT για το πρόβλημά μας είναι:

$$\frac{\partial L}{\partial w_j} = w_j - \sum_{i=1}^l a_i y_i x_j = 0 \Rightarrow w_j = \sum_{i=1}^l a_i y_i x_j \quad (2.12)$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^l a_i y_i = 0 \quad (2.13)$$

$$a_i [y_i (w \cdot x_i + b) - 1] = 0 \quad \forall i \quad (2.14)$$

$$a_i \geq 0 \quad \forall i \quad (2.15)$$

Αν αντικαταστήσουμε τις συνθήκες KKT (2.12, 2.13) στην (2.11) έχουμε:

$$Ld = L = \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l w_i a_j y_j x_i + \sum_{i=1}^l a_i - \sum_{i=1}^l \sum_{j=1}^l a_i y_i x_j w_j - b \sum_{i=1}^l a_i y_i \Rightarrow$$

$$Ld = \sum_{i=1}^l a_i + \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l w_i a_j y_j x_i - \sum_{i=1}^l \sum_{j=1}^l a_i y_i x_j w_j \Rightarrow$$

$$Ld = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l w_i a_j y_j x_i \Rightarrow$$

$$Ld = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j x_i x_j \quad (2.16)$$

Σύμφωνα με τη θεωρία του Wolfe (*Fletcher (1987)*), η σχέση (2.16) με τους περιορισμούς (2.13, 2.15) αποτελεί το δυαδικό (dual) του προβλήματος όπως αυτό εκφράζεται από τη σχέση (2.10) με τους περιορισμούς (2.9). Η ελαχιστοποίηση της (2.10) υπό τους περιορισμούς (2.9) είναι ισοδύναμη με την μεγιστοποίηση της (2.16) υπό τους περιορισμούς (2.13, 2.15)

Άρα τελικά καταλήγουμε στο εξής πρόβλημα:

$$\left[\begin{array}{l} \max : Ld = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j x_i x_j \\ \sum_{i=1}^l a_i y_i = 0 \\ a_i \geq 0 \end{array} \right] \quad (2.17)$$

Το παραπάνω πρόβλημα ανήκει και αυτό στην οικογένεια των προβλημάτων τετραγωνικού προγραμματισμού (Quadratic Programming Problem). Αν το σύστημα (2.17) εκφραστεί με τη μορφή πινάκων έχουμε:

$$\left[\begin{array}{l} \max : Ld = -0.5a^T H a + f^T a \\ y^T a = 0 \\ a \geq 0 \end{array} \right] \quad (2.17a)$$

όπου: Hessian matrix: $H_{ij} = y_i y_j x_i^T x_j$ και $f = [11 \dots 1]^T$

Από την παρατήρηση της σχέσης (2.12) βλέπουμε πως το w το οποίο μας δίνει την κλίση της διαχωριστικής ευθείας προέρχεται από σχέση που επίδραση έχουν μόνο τα σημεία x_i με μη-μηδενικούς πολλαπλασιαστές Lagrange ($a_i \geq 0$). Αυτοί οι πολλαπλασιαστές Lagrange αντιστοιχούν στα ακραία σημεία των κλάσεων (Support Vectors - SVs) αφού στη σχέση (2.14) αν $a_i > 0$ πρέπει $y_i (w \cdot x_i + b) - 1 = 0$ η οποία

ικανοποιείται μόνο για τα ακραία σημεία των κλάσεων. Έτσι αποδεικνύεται αυτό που υποθέσαμε πριν από το Σχήμα 2.5, ότι δηλαδή η μορφή της ευθείας εξαρτάται μόνο από τα ακραία σημεία των κλάσεων.

Η λύση του προβλήματος θα έλθει με την επίλυση του συστήματος (2.17). Με την μεγιστοποίηση δηλαδή της συνάρτησης υπό τους περιορισμούς της, ώστε να προκύψει η λύση, δηλαδή οι πολλαπλασιαστές Lagrange α_i . Έπειτα, από τη σχέση (2.12) θα προκύψει ο παράγοντας \vec{w}^* ο οποίος καθορίζει την κλίση-μορφή της βέλτιστης διαχωριστικής ευθείας. Πλέον, η σχέση (2.14) για κάποιο σημείο i για το οποίο ισχύει $\alpha_i \geq 0$ (ακραίο σημείο (SV)), γίνεται:

$$\vec{w}^* \cdot \vec{x}_i + b^* = \pm 1 \quad (2.18)$$

Από τη σχέση (2.18) μπορούμε να υπολογίσουμε το b^* της βέλτιστης διαχωριστικής ευθείας από οποιοδήποτε σημείο που λειτουργεί ως SV ($\alpha > 0$). Για πιο ακριβή αποτελέσματα, προτιμότερο είναι να λυθεί η (2.18) για όλα τα SVs (έστω ns το πλήθος τους) και έτσι να προκύψουν τα αντίστοιχα b_i^* . Προφανώς, τα b_i^* που θα προκύψουν θα είναι όσα και τα SVs, δηλαδή ns . Τέλος, το b^* που θα χρησιμοποιήσουμε για τη διαχωριστική ευθεία θα προκύψει ως ο μέσος όρος των b_i^* (*Burges (1998)*).

Γενικεύοντας, μπορούμε να πούμε ότι για τα σημεία που θα ισχύει $\alpha_i > 0$, αποτελούν τα διανύσματα υποστήριξης (SVs), δηλαδή τα ακραία σημεία των κλάσεων, ενώ για όλα τα άλλα ισχύει $\alpha_i = 0$.

Πλέον, έχουμε διαχωρίσει τις δύο κλάσεις αφού έχουν βρεθεί όλες οι παράμετροι της εξίσωσης της διαχωριστικής ευθείας. Αναγκαίο τώρα είναι να ειπωθεί και ο τρόπος κατάταξης ενός νέου σημείου.

Από τον *Vapnik (1998)* προτάθηκε μια συνάρτηση απόφασης (ή συνάρτηση κατάταξης) με τη μορφή:

$$f = (\vec{w}^* \cdot \vec{x} + b^*) \quad (2.19)$$

όπου \vec{x} το προς κατάταξη σημείο. Ελέγχουμε το πρόσημο της f , δηλαδή $sign(f) = sign(\vec{w}^* \cdot \vec{x} + b^*)$. Αν το πρόσημο της συνάρτησης απόφασης είναι θετικό,

το σημείο κατατάσσεται στην κλάση με ετικέτα +1 ενώ αν το πρόσημο είναι αρνητικό το σημείο κατατάσσεται στην κλάση με ετικέτα -1. Όπως βλέπουμε, η σχέση (2.19) αποτελεί τη συνάρτηση της διαχωριστικής ευθείας (σχέση (2.8)) υπολογισμένη για το προς κατάταξη σημείο.

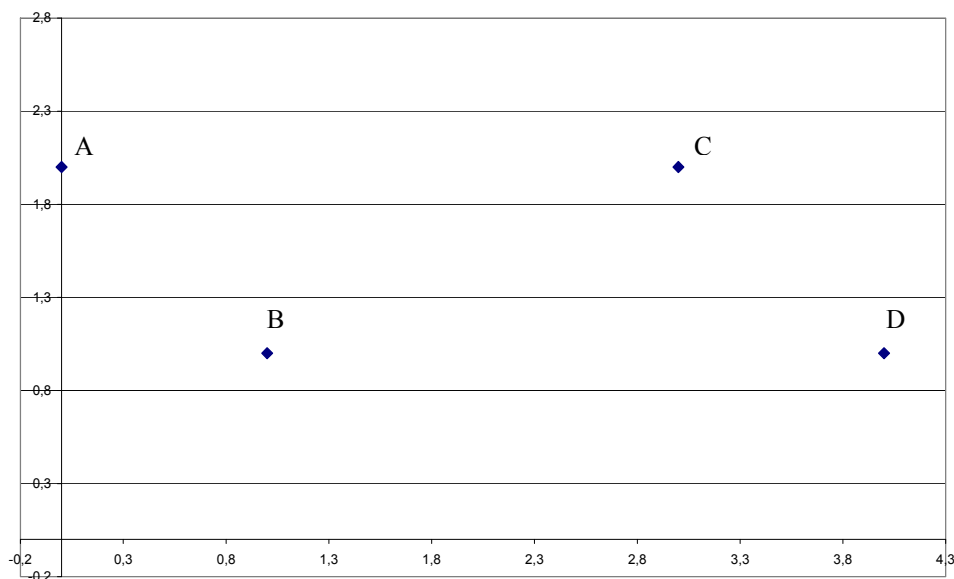
Εδώ πρέπει να αναφερθεί πως η παραπάνω ανάλυση δε σταματά στο δισδιάστατο πρόβλημα αλλά επεκτείνεται και στο \mathbb{R}^N , μόνο που πλέον μιλάμε για διαχωριστικό υπερεπίπεδο (hyperplane) (Vapnik (1995), (1998)) και τα διανύσματα w και x έχουν μεγαλύτερη διάσταση. Οι σχέσεις που προαναφέρθηκαν ισχύουν και επιλύονται με τη διαδικασία που προαναφέρθηκε.

Παράδειγμα

Για να γίνει πιο κατανοητή η μέθοδος θα επιλυθεί βήμα-βήμα ένα απλό παράδειγμα στο δισδιάστατο χώρο με 2 σημεία στην κλάση +1 και δύο σημεία στην κλάση -1.

Έστω τα σημεία και δίπλα η κλάση τους και η απεικόνισή τους στο Σχήμα 2.6.

- A: (0,2) +1
- B: (1,1) +1
- C: (3,2) -1
- D: (4,1) -1



Σχήμα 2.6. Απεικόνιση των 4 σημείων του παραδείγματος (2 κλάσεις)

Για αυτά τα δεδομένα από την ομάδα σχέσεων (2.17) έχουμε:

$$Ld = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - 2\alpha_1^2 - 2\alpha_1\alpha_2 + 4\alpha_1\alpha_3 + 2\alpha_1\alpha_4 - \alpha_2^2 + 5\alpha_2\alpha_3 + 5\alpha_2\alpha_4 - 6,5\alpha_3^2 - 14\alpha_3\alpha_4 - 8,5\alpha_4^2$$

Με ένα απλό πρόγραμμα σε Fortran 77 που σαρώνει τις τιμές των α_i και απορρίπτει αυτές που δεν πληρούν τους περιορισμούς της ομάδας σχέσεων (2.17) βρίσκουμε εύκολα ότι η Ld μεγιστοποιείται για $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (0, 0.4, 0.4, 0)$.

Από τη λύση βλέπουμε πως SVs είναι τα σημεία B: (1,1), C: (3,2).

Από την εφαρμογή της σχέσης (2.12) για το w^* έχουμε:

$$\vec{w}^* = \sum_{i=1}^{ns} a_i y_i x_j = 0.4 \cdot 1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0.4 \cdot (-1) \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -0.8 \\ -0.4 \end{bmatrix}$$

Από την εφαρμογή της σχέσης (2.18) για το b^* και για τα δύο σημεία B και C που λειτουργούν ως SVs έχουμε:

$$B: \vec{w}^* \cdot \vec{x} + b^* = +1 \Rightarrow \begin{bmatrix} -0.8 \\ -0.4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + b_1^* = 1 \Rightarrow b_1^* = 2.2$$

$$C: \vec{w}^* \cdot \vec{x} + b^* = -1 \Rightarrow \begin{bmatrix} -0.8 \\ -0.4 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix} + b_2^* = -1 \Rightarrow b_2^* = 2.2$$

και όπως προαναφέρθηκε το b^* λαμβάνεται ως ο μέσος όρος των b_1^* και b_2^*

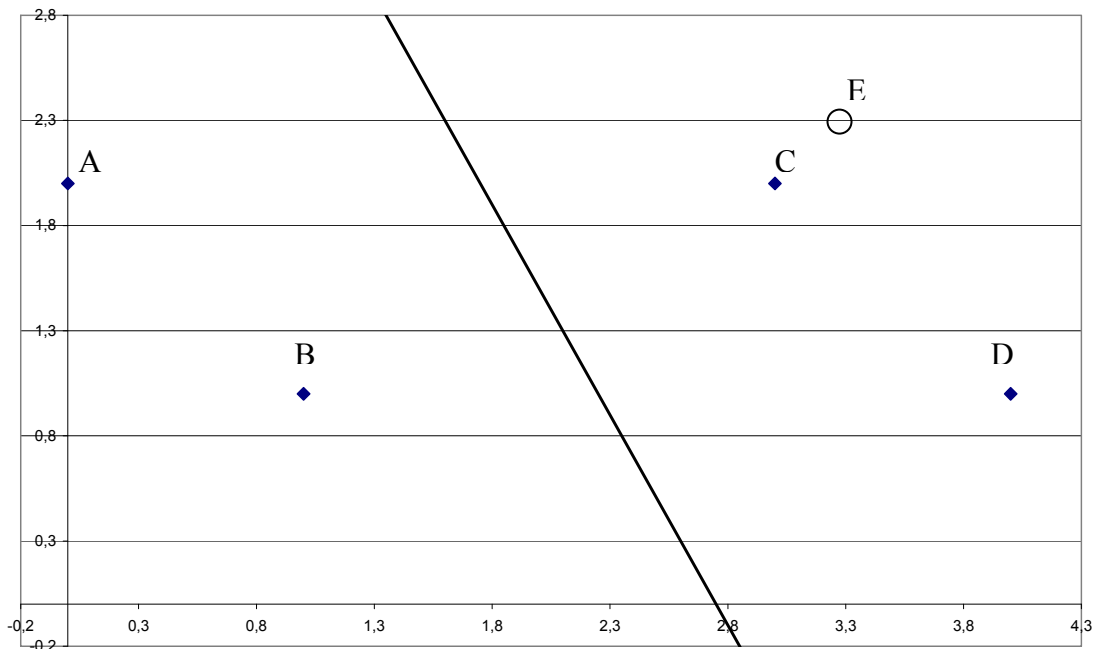
$$b^* = (b_1^* + b_2^*) / 2 = 2.2$$

Έτσι η εξίσωση της διαχωριστικής ευθείας είναι:

$$\vec{w} \cdot \vec{x} + b = 0 \Rightarrow y = -2x + 5.5$$

η οποία και φαίνεται στο Σχήμα 2.7.

Στο Σχήμα 2.7 επίσης απεικονίζεται και ένα καινούριο σημείο E για το οποίο θα εφαρμοστεί η διαδικασία της κατάταξης. Για το σημείο έχουμε E: (2.8, 2.3).



Σχήμα 2.7. Διαχωρισμός των κλάσεων του παραδείγματος και κατάταξη σημείου.

Από τη σχέση (2.19) έχουμε:

$$f = (\vec{w}^* \cdot \vec{x} + b^*) = \begin{bmatrix} -0.8 \\ -0.4 \end{bmatrix} \cdot \begin{bmatrix} 2.8 \\ 2.3 \end{bmatrix} + 2.2 = -3.16$$

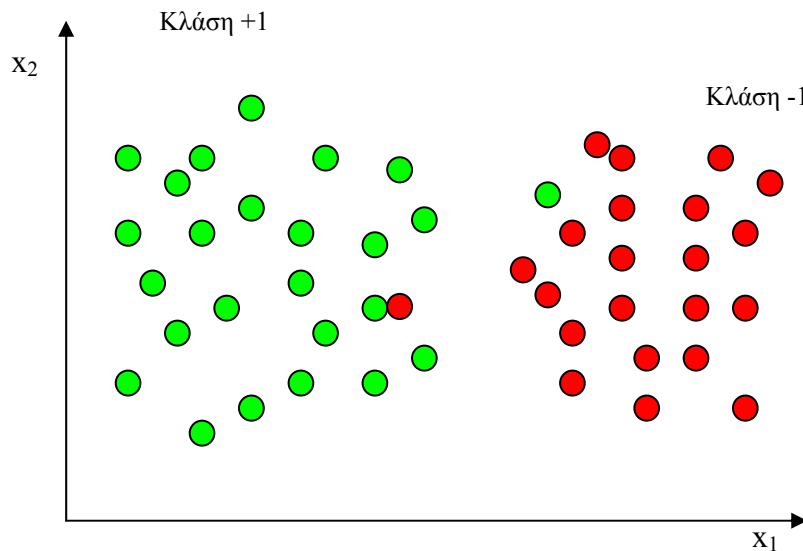
Όπως βλέπουμε το πρόσημο της συνάρτησης κατάταξης f είναι αρνητικό και άρα το σημείο E θα καταταχθεί στην κλάση με ετικέτα -1. Αυτό συμφωνεί και με την εικόνα που παίρνουμε από το Σχήμα 2.7.

2.3.2 Μη-απόλυτα γραμμικά διαχωρίσιμες κλάσεις

Στην παράγραφο 2.3.1 αναλύσαμε την περίπτωση των πλήρως διαχωρίσιμων κλάσεων. Όμως, αυτή η περίπτωση δεν αποτελεί την πλειοψηφία των προβλημάτων.

Τώρα θα αναλύσουμε την περίπτωση που οι δύο κλάσεις δεν είναι πλήρως γραμμικά διαχωρίσιμες με κάποια ευθεία (ή γενικότερα υπερεπίπεδο).

Έστω η δισδιάστατη περίπτωση ενός δισδιάστατου συνόλου εκπαίδευσης μεγέθους $l(\bar{x}_i, y_i)$ (Σχήμα 2.8).



Σχήμα 2.8. Δισδιάστατο σύνολο εκπαίδευσης μη πλήρως γραμμικά διαχωρίσιμων κλάσεων.

Όπως γίνεται εύκολα αντιληπτό από το Σχήμα 2.8, επειδή οι δύο κλάσεις έχουν μια περιοχή αλληλοεπικάλυψης, δεν μπορεί να βρεθεί μια ευθεία γραμμή που να τις διαχωρίσει απόλυτα. Οποιαδήποτε ευθεία γραμμή και να χαραχθεί θα αφήνει κάποιο (κάποια) σημείο της κλάσης +1 στα δεξιά της (αντίστοιχα για την -1) με αποτέλεσμα να μη διαχωρίζονται οι κλάσεις πλήρως. Άρα, η εφαρμογή της μεθόδου όπως αυτή αναπτύχθηκε για τις πλήρως διαχωρίσιμες κλάσεις, δεν θα μας έδινε κάποιο αποτέλεσμα.

Μια αντιμετώπιση αυτού του προβλήματος θα ήταν να χαραχθεί μια γραμμή (όχι ευθεία) που θα περνά ανάμεσα από τα “ακραία” σημεία των δύο κλάσεων και έτσι να τις διαχωρίσει απόλυτα. Αυτό όμως θα είχε σαν αποτέλεσμα τη μείωση του διακένου (margin) σε κάποιες περιοχές της γραμμής, όπως επίσης και την αύξηση της πολυπλοκότητας της μεθόδου.

Μια δεύτερη λύση (η οποία και είναι αυτή που εφαρμόζεται) είναι να “χαλαρώσουμε” τους περιορισμούς της σχέσης (2.9) όταν αυτό κρίνεται απαραίτητο.

Έτσι, οι δύο κλάσεις θα διαχωρίζονται με μια ευθεία (γενικά υπερεπίπεδο) και τα σημεία των δύο κλάσεων που ευθύνονται για την αλληλοεπικάλυψη, δεν θα παίζουν τόσο σημαντικό ρόλο, δηλαδή δεν θα παίζουν τον ρόλο των ακραίων σημείων που διαμορφώνουν τη μορφή της διαχωριστικής ευθείας.

Για να πετύχουμε αυτή τη “χαλάρωση” των περιορισμών εισάγουμε στους περιορισμούς τις μεταβλητές χαλάρωσης (slack variables) ξ_i , $i=1,l$ (Cortes and Vapnik (1995)). Έτσι οι περιορισμοί (2.9) γίνονται:

$$\left[\begin{array}{l} \bar{w} \cdot \bar{x}_i + b \geq +1 - \xi_i, y_i = +1 \\ \bar{w} \cdot \bar{x}_i + b \leq -1 + \xi_i, y_i = -1 \\ \xi_i \geq 0 \quad \forall i \end{array} \right] \quad (2.20)$$

Έτσι, το $\sum_i \xi_i$ αποτελεί ένα άνω όριο του σφάλματος κατάταξης των σημείων εκπαίδευσης (Burges (1998)).

Σε αυτή τη περίπτωση η προς ελαχιστοποίηση ποσότητα της (2.10) τροποποιείται ως εξής:

$$\frac{1}{2} \|\bar{w}\|^2 + C \left(\sum_i \xi_i \right)^k \quad (2.21)$$

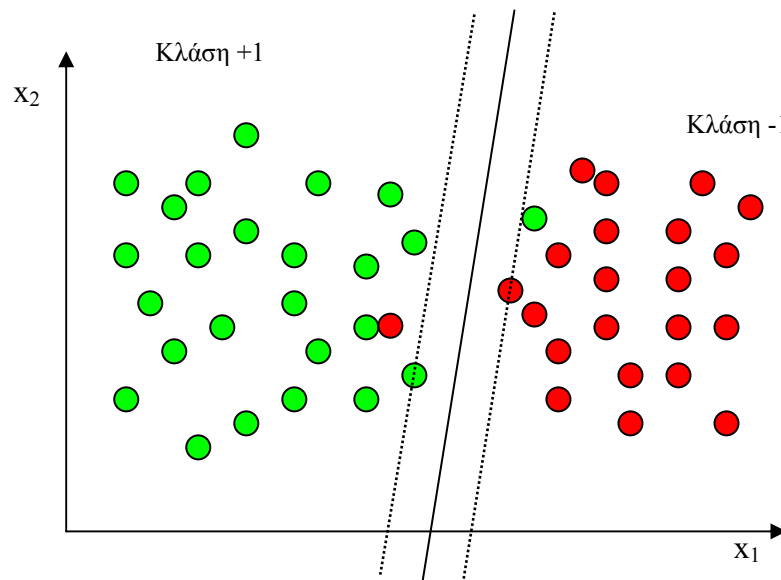
Η C είναι μια παράμετρος που καθορίζεται από το χρήστη και σχετίζεται με την “ποινή” που θα υπάρχει στην περίπτωση λάθους στην κατάταξη. Μεγαλύτερο C σημαίνει μεγαλύτερη ποινή. Ο εκθέτης k μπορεί να πάρει μια οποιαδήποτε ακέραια τιμή. Θα δεχθούμε $k=1$ διότι είναι βολικός αφού το πρόβλημα συνεχίζει να είναι πρόβλημα τετραγωνικού προγραμματισμού και στη δυαδική συνάρτηση μεγιστοποίησης δεν υπεισέρχονται τα ξ_i , ούτε και οι πολλαπλασιαστές Lagrange τους (Burges (1998)).

Σκοπός μας τώρα είναι να ελαχιστοποιήσουμε την ποσότητα της (2.21) υπό τους περιορισμούς (2.20). Ο πρώτος όρος της (2.21) σχετίζεται με την αύξηση του διακένου ενώ ο δεύτερος με τη μείωση του σφάλματος ταξινόμησης.

Όπως αναλύθηκε στην περίπτωση των πλήρως διαχωρίσιμων κλάσεων με τη βοήθεια του μετασχηματισμού Lagrange το πρόβλημά μας καταλήγει στην εξής μορφή:

$$\left[\begin{array}{l} \max : Ld = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j x_i x_j \\ \sum_{i=1}^l a_i y_i = 0 \\ 0 \leq a_i \leq C \end{array} \right] \quad (2.22)$$

Στο Σχήμα 2.9 βλέπουμε τον διαχωρισμό των κλάσεων στην περίπτωση που αυτές είναι μη πλήρως γραμμικά διαχωρίσιμες ύστερα από την επίδραση των μεταβλητών χαλάρωσης.



Σχήμα 2.9. Διαχωρισμός μη πλήρως γραμμικά διαχωρίσιμων κλάσεων.

Από τη σύγκριση των σχέσεων (2.17) και (2.22) βλέπουμε ότι αυτές διαφέρουν μόνο στην ύπαρξη ενός άνω ορίου στην τιμή που μπορούν να πάρουν οι πολλαπλασιαστές Lagrange στην περίπτωση των μη πλήρως διαχωρίσιμων κλάσεων (2.22). Αυτό μας επιτρέπει να πούμε πως η (2.22) είναι γενίκευση της (2.17) και μετασχηματίζεται σε αυτήν αν $C \rightarrow \infty$. Στην περίπτωση των μη πλήρως διαχωρίσιμων κλάσεων η τιμή του C είναι πεπερασμένη.

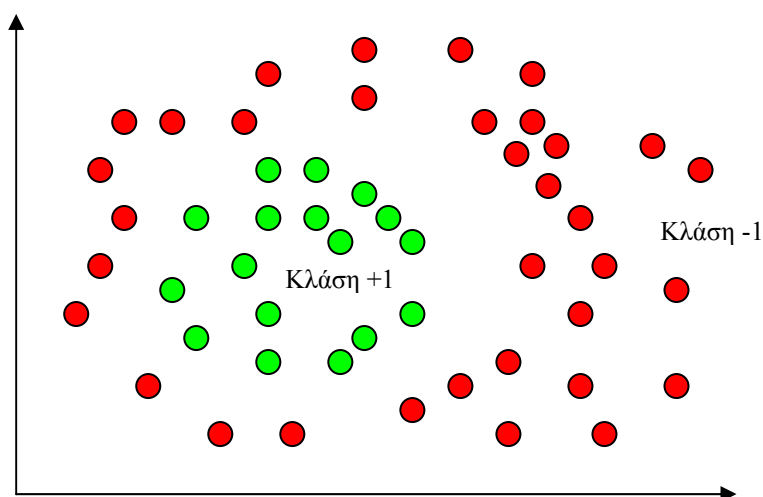
Στην εφαρμογή της ομάδας εξισώσεων (2.22) δε μπορούμε να γνωρίζουμε εκ των προτέρων ποια τιμή θα δώσουμε στην παράμετρο C . Για παράδειγμα, τα σημεία που ευθύνονται για τη μη πλήρη διαχωρισιμότητα μπορεί να είναι λίγα σε σχέση με

των ολικό αριθμό των σημείων ή και να περιέχουν θόρυβο που ευθύνεται για την αλληλοεπικάλυψη. Έτσι, η τιμή της παραμέτρου C δεν χρειάζεται να είναι μεγάλη αφού δεν μας ενδιαφέρει ο αυστηρός διαχωρισμός στα σημεία αλληλοεπικάλυψης. Άρα, η κατάλληλη τιμή της παραμέτρου C θα πρέπει να βρεθεί με βάση την εμπειρία, το είδος των δεδομένων καθώς και με τη βοήθεια δοκιμών.

Όπως και στην περίπτωση των πλήρως διαχωρίσιμων κλάσεων, έτσι κι εδώ, μπορούμε να πούμε ότι οι εξισώσεις του προβλήματος ισχύουν για το N -διάστατο χώρο χωρίς αλλαγές. Πλέον μιλάμε για υπερεπίπεδο και όχι για διαχωριστική ευθεία.

2.3.3 Γραμμικά μη διαχωρίσιμες κλάσεις

Οι δύο παραπάνω περιπτώσεις αφορούσαν κλάσεις που μπορούσαν (έστω και με κάποιο μικρό σφάλμα) να διαχωριστούν γραμμικά. Όμως, η περίπτωση αυτή αποτελεί τη μειονότητα των προβλημάτων. Η πιο γενική περίπτωση είναι αυτή των γραμμικά μη διαχωρίσιμων κλάσεων (μη διαχωρίσιμες από υπερεπίπεδο στο N -διάστατο χώρο). Παράδειγμα τέτοιας περίπτωσης φαίνεται στο Σχήμα 2.10 όπου τα σημεία της μίας κλάσης (δισδιάστατος χώρος) βρίσκονται γύρω από τα σημεία της άλλης κλάσης με αποτέλεσμα να μην μπορούμε να φέρουμε μια ευθεία που να τις διαχωρίζει.



Σχήμα 2.10. Γραμμικά μη διαχωρίσιμες κλάσεις.

Η περίπτωση διαχωρισμού των κλάσεων με κάποια άλλη γραμμή (καμπύλη) θεωρείται επιβεβλημένη, καθώς μόνο με μια τέτοια αντιμετώπιση θα μπορούσαμε να διαχωρίσουμε τις κλάσεις. Στη περίπτωση που απεικονίζεται στο Σχήμα 2.10, μια γραμμή που θα διαχώριζε τις κλάσεις θα μπορούσε να είναι κυκλικής μορφής.

Η αναζήτηση γραμμών (καμπυλών) μεγαλύτερου βαθμού από την περίπτωση της ευθείας γραμμής, όπως αναφέρθηκε στην προηγούμενη περίπτωση (παράγραφος 2.3.2) θα αύξανε την πολυπλοκότητα. Σκοπός μας είναι να μπορέσουμε να αντιμετωπίσουμε το πρόβλημα με τις υπάρχουσες απλές εξισώσεις οι οποίες όμως δεν θα αναφέρονται σε ευθεία.

Τα σημεία του παραδείγματος βλέπουμε πως στο δισδιάστατο χώρο είναι μη γραμμικά διαχωρίσιμα. Αν όμως απεικονιστούν σε χώρο μεγαλύτερων διαστάσεων θα μπορούν να διαχωριστούν από κάποιο υπερεπίπεδο, θα γίνονταν δηλαδή γραμμικά διαχωρίσιμα. Η απεικόνιση του διαχωριστικού υπερεπιπέδου στο δισδιάστατο χώρο θα μπορούσε να είναι η καμπύλη κυκλικής μορφής. Η απεικόνιση Φ του χώρου R^d στον χώρο H , είναι $\Phi: R^d \rightarrow H$.

Στις σχέσεις της μεθόδου, όπως είδαμε, τα σημεία εκπαίδευσης παρουσιάζονται μόνο με τη μορφή του εσωτερικού γινομένου αυτών $(x_i \cdot x_j)$ π.χ.

$$Ld = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j x_i x_j.$$

Έτσι, αν ανάγουμε το πρόβλημα σε κάποιο χώρο μεγαλύτερων διαστάσεων H , θα έχουμε κι εκεί μόνο την εμφάνιση του εσωτερικού γινομένου. Όμως, μας ενδιαφέρει το αποτέλεσμα του εσωτερικού γινομένου στο χώρο H και όχι ο χώρος αυτός καθαυτός. Έτσι, αν στις σχέσεις που αναπτύχθηκαν αντικαταστήσουμε το εσωτερικό γινόμενο των σημείων εκπαίδευσης του χώρου που βρίσκονται με το αποτέλεσμα του εσωτερικού γινομένου στο χώρο μεγαλύτερων διαστάσεων H , θα έχουμε αντιμετωπίσει το πρόβλημά μας.

Με απλά λόγια, εφαρμόζουμε πάλι τις σχέσεις του γραμμικού διαχωρισμού αλλά σε έναν άλλο χώρο H στον οποίο οι κλάσεις είναι διαχωρίσιμες από κάποιο υπερεπίπεδο.

Με αυτό το τρόπο μπορούμε να αντιμετωπίσουμε και την περίπτωση των μη πλήρως γραμμικά διαχωρίσιμων κλάσεων, αφού με την απεικόνιση, οι κλάσεις θα γίνουν γραμμικά διαχωρίσιμες.

Το πρόβλημά μας παίρνει τη μορφή:

$$\left[\begin{array}{l} \max : Ld = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j \langle x_i x_j \rangle \\ \sum_{i=1}^l a_i y_i = 0 \\ a_i \geq 0 \end{array} \right] \quad (2.23)$$

Και αν θεωρήσουμε και μεταβλητές χαλάρωσης, ώστε να αντιμετωπίσουμε την περίπτωση που το πρόβλημά μας είναι μη απόλυτα γραμμικά διαχωρίσιμο στον χώρο H , έχουμε:

$$\left[\begin{array}{l} \max : Ld = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j \langle x_i x_j \rangle \\ \sum_{i=1}^l a_i y_i = 0 \\ 0 \leq a_i \leq C \end{array} \right] \quad (2.24)$$

Ο όρος $\langle x_i x_j \rangle$ αποτελεί το αποτέλεσμα του εσωτερικού γινομένου των σημείων στο χώρο μεγαλύτερων διαστάσεων. Δηλαδή είναι $\langle x_i x_j \rangle = K(x_i, x_j)$, όπου K μια συνάρτηση που μας δίνει το αποτέλεσμα του εσωτερικού γινομένου στο χώρο H χωρίς να είναι απαραίτητο να γνωρίζουμε το χώρο αυτό. Η συνάρτηση K ονομάζεται συνάρτηση κελύφους (Kernel function) και θα αναλυθεί στη συνέχεια.

Όπως παρατηρούμε αν ισχύει $K(x_i, x_j) = x_i \cdot x_j$, τότε η (2.23) μεταπίπτει στην (2.17), δηλαδή στην περίπτωση των πλήρως γραμμικά διαχωρίσιμων κλάσεων.

Η κατάταξη των σημείων, όπως είναι λογικό, θα πρέπει να γίνει και αυτή στο χώρο H . Η συνάρτηση κατάταξης (2.19), αφού εξαρτάται και αυτή από το εσωτερικό γινόμενο, παίρνει τώρα τη μορφή:

$$f = \text{sign}\left(\sum_{i=1}^l y_i a_i K(x, x_i) + b^*\right) \quad (2.25)$$

Όπως και πριν, αν το πρόσημο της συνάρτησης απόφασης f είναι θετικό, το σημείο κατατάσσεται στην κλάση +1, ενώ αν είναι αρνητικό στην κλάση -1

2.4 Συνάρτηση κελύφους (kernel function)

Όπως αναφέρθηκε προηγουμένως, στην εφαρμογή των τύπων της μεθόδου τα σημεία εκπαίδευσης εμφανίζονται μόνο με τη μορφή του εσωτερικού γινομένου τους. Έτσι, δεν χρειάζεται να ξέρουμε το χώρο (και το μετασχηματισμό Φ) των μεγαλύτερων διαστάσεων H στον οποίο το πρόβλημά μας γίνεται γραμμικά διαχωρίσιμο. Μας αρκεί το αποτέλεσμα του εσωτερικού γινομένου των σημείων εκπαίδευσης σε αυτό το χώρο. Το αποτέλεσμα αυτό μας το δίνει η συνάρτηση κελύφους (kernel function).

2.4.1 Θεωρητικό υπόβαθρο

Το σημαντικό ερώτημα που τίθεται τώρα, είναι ποιες συναρτήσεις κελύφους μπορούν να χρησιμοποιηθούν με βάση την παραπάνω λογική. Δηλαδή, ποιες συνθήκες πρέπει να ικανοποιεί μια συνάρτηση ώστε να ονομάζεται συνάρτηση κελύφους. Η απάντηση είναι ότι πρέπει να ικανοποιούνται οι συνθήκες Mercer (*Vapnik (1995), Courant and Hilbert (1953)*). Το θεώρημα Mercer διατυπώνεται ως εξής:

Μια συνάρτηση K μας δίνει το εσωτερικό γινόμενο μιας απεικόνισης Φ , δηλαδή:

$$K(x, y) = \sum_i \Phi(x)_i \Phi(y)_i \quad (2.26)$$

αν και μόνον αν για κάθε συνάρτηση $g(x)$ τέτοια ώστε:

$$\int g(x)^2 dx \in \mathfrak{R} \quad \text{ισχύει} \quad \int K(x, y)g(x)g(y)dxdy \geq 0 \quad (2.27)$$

Οι πρώτες συναρτήσεις κελύφους που χρησιμοποιήθηκαν και αναλύθηκαν για την περίπτωση του διαχωρισμού κλάσεων ήταν (*Burges (1998)*):

$$K(x, y) = x \cdot y \quad \text{Γραμμικό kernel (όχι απεικόνιση σε χώρο H)}$$

$$K(x, y) = (k(x \cdot y) + n)^p \quad \text{Πολυωνυμικό kernel (Polynomial kernel)}$$

$$K(x, y) = e^{-\gamma \|x-y\|^2 / 2\sigma^2} \quad \text{RBF kernel}$$

$$K(x, y) = \tanh(k(x \cdot y) - \delta) \quad \text{Σιγμοειδές kernel (Sigmoid kernel)}$$

Όπως βλέπουμε οι συναρτήσεις κελύφους έχουν κάποιες μεταβλητές που καθορίζονται από το χρήστη. Οι μεταβλητές αυτές δεν έχουν καθορισμένη τιμή και σε κάθε περίπτωση πρέπει να επιλεγεί σωστά η τιμή τους. Ένας τρόπος είναι να χωρίσουμε τα δεδομένα εκπαίδευσης που έχουμε σε δύο ομάδες. Με τη μία ομάδα να κάνουμε την εκπαίδευση και με την άλλη την κατάταξη. Αυτή η διαδικασία πρέπει να γίνει για διάφορες τιμές των παραμέτρων των kernels. Οι βέλτιστες τιμές των παραμέτρων των kernels είναι αυτές για τις οποίες είχαμε τα λιγότερα λάθη στην διαδικασία της κατάταξης.

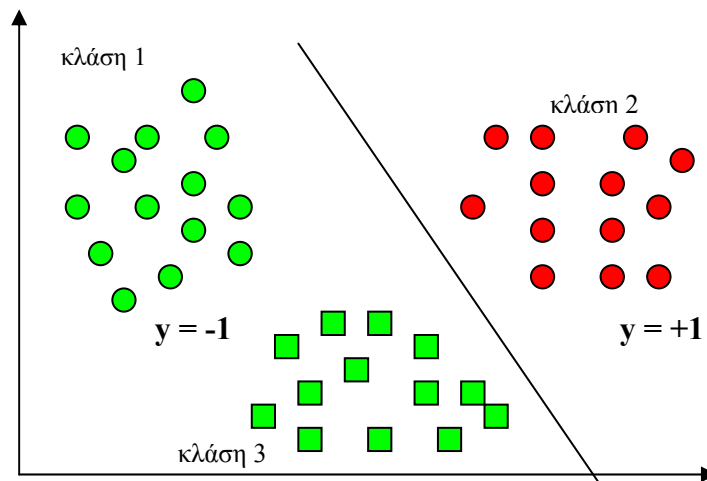
2.5 Η περίπτωση των πολλών κλάσεων (multiclassification)

Στις προηγούμενες παραγράφους περιγράφηκε ο τρόπος διαχωρισμού δυο κλάσεων σε όλες τις δυνατές περιπτώσεις. Στην πλειοψηφία των προβλημάτων όμως θα πρέπει να γίνει η κατάταξη σε περισσότερες των δύο κλάσεων. Η μέθοδος των SVM που αναπτύχθηκε, αφορά στο διαχωρισμό δύο κλάσεων με κάποιο υπερεπίπεδο και άρα δεν είναι δυνατόν να γίνει αυτή η διαδικασία για περισσότερες των δύο κλάσεων.

Επειδή η μέθοδός μας διαχωρίζει τις κλάσεις ανά δύο, θα πρέπει το σύνολο των περισσότερων κλάσεων να το φέρουμε σε αυτή τη μορφή ώστε να είναι εύκολα αντιμετωπίσιμο. Με βάση αυτή τη λογική έχουν προταθεί διάφορες μέθοδοι, οι σημαντικότερες των οποίων θα περιγραφούν στη συνέχεια εν συντομία. Αυτές οι μέθοδοι είναι οι (*Debnath et al. (2004)*): one-against-all, all-together, one-against-one και DAGSVM.

2.5.1 One-against-all

Η αρχή της μεθόδου αυτής είναι ότι εκπαιδεύουμε κάθε κλάση εναντίων όλων των άλλων. Σε ένα πρόβλημα που έχουμε k κλάσεις δημιουργούνται και k μοντέλα SVM. Για παράδειγμα, η κλάση i παίρνει ετικέτα $+1$ και όλες οι άλλες $k-1$ κλάσεις θεωρούνται ως μία που έχει ετικέτα -1 όπως φαίνεται στο Σχήμα 2.11.

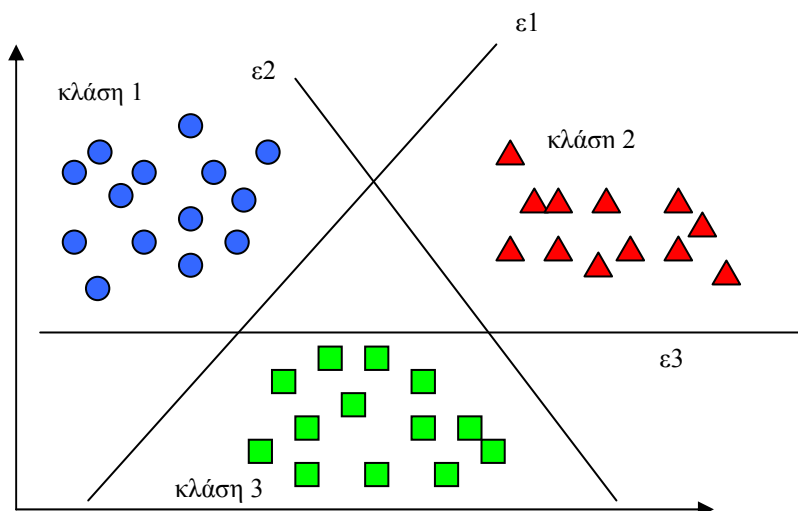


Σχήμα 2.11. Διαχωρισμός τριών κλάσεων με τη μέθοδο one-against-all.

Με αυτόν τον τρόπο προσπαθούμε να βρούμε τα όρια της κάθε κλάσης εναντίον όλων των άλλων μαζί και έτσι έχουμε k συναρτήσεις απόφασης αφού έχουμε k διαχωριστικά υπερεπίπεδα. Τα όρια της κάθε κλάσης προέρχονται ύστερα από συμψηφισμό των αποτελεσμάτων των k SVM.

Για την διαδικασία της κατάταξης τώρα, υπολογίζουμε την τιμή των k συναρτήσεων απόφασης και κατατάσσουμε το σημείο στην κλάση που αντιστοιχεί στην συνάρτηση απόφασης με τη μεγαλύτερη τιμή. Αυτή η λογική επιφέρει λάθη κατά τη διαδικασία της κατάταξης αφού εξετάζει μόνο τη τιμή της συνάρτησης απόφασης.

Στο Σχήμα 2.12 φαίνεται η αρχή της μεθόδου αυτής σε ένα διδιάστατο παράδειγμα με τρεις κλάσεις και τα διαχωριστικά επίπεδα που προήλθαν από αυτή τη μέθοδο. Η ευθεία ϵ_1 διαχωρίζει την κλάση 1 από όλες τις υπόλοιπες. Ομοίως, οι ευθείες ϵ_2 και ϵ_3 , διαχωρίζουν τις κλάσεις 2 και 3 από όλες τις υπόλοιπες, αντίστοιχα.



Σχήμα 2.12. Διαχωρισμός τριών κλάσεων με τη μέθοδο one-against-all.

2.5.2 All-together

Η μέθοδος αυτή προτάθηκε από τον εφευρέτη των SVM, Vapnik. Η μέθοδος αυτή είναι παραπλήσια της one-against-all. Δημιουργούνται k συναρτήσεις απόφασης για τις k κλάσεις, όπου, για παράδειγμα η n συνάρτηση απόφασης διαχωρίζει την n κλάση από όλες τις άλλες. Στην one-against-all επιλύονται k SVM για να βρεθούν τα όρια της μίας κλάσης, ενώ εδώ, συνοψίζονται στη λύση ενός.

Η συνάρτηση απόφασης αυτής της μεθόδου είναι η ίδια με την one-against-all και το σημείο κατατάσσεται στην κλάση που έχει τη μέγιστη τιμή η αντίστοιχη συνάρτηση απόφασης.

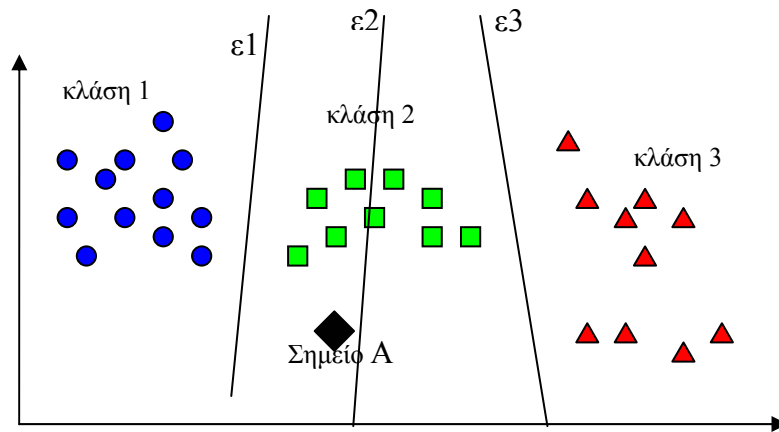
2.5.3 One-against-one

Σε αντίθεση με τη one-against-all, εδώ θεωρούνται όλα τα πιθανά ζεύγη των κλάσεων και λύνεται το πρόβλημα διαχωρισμού τους. Δηλαδή, για k κλάσεις έχουμε $k!/2(k-2)!$ δυνατούς συνδυασμούς των κλάσεων. Επίσης, τόσες είναι και οι συναρτήσεις απόφασης. Πιο απλά, εφαρμόζουμε τη μεθοδολογία που αναπτύχθηκε

στις προηγούμενες παραγράφους και βρίσκουμε το διαχωριστικό υπερεπίπεδο κάθε πιθανού ζεύγους κλάσεων.

Η πιο διαδεδομένη μέθοδος για την κατάταξη ενός σημείου αν έχουμε ακολουθήσει τη διαδικασία της one-against-one είναι η max-wins. Με τον αλγόριθμο αυτό, κάθε προς κατάταξη σημείο ελέγχεται με όλες τις συναρτήσεις απόφασης και κάθε φορά κατατάσσεται σε μια κλάση. Η κλάση στην οποία έχει καταταχθεί τις περισσότερες φορές είναι και η κλάση στην οποία τελικά τοποθετείται το σημείο.

Στο Σχήμα 2.13 βλέπουμε ένα διδιάστατο πρόβλημα τριών κλάσεων και τις διαχωριστικές ευθείες. Επίσης παρουσιάζεται και η διαδικασία της κατάταξης ενός σημείου.



Σχήμα 2.13. Δισδιάστατο πρόβλημα τριών κλάσεων που διαχωρίζονται με τη μέθοδο one-against-one, καθώς επίσης και η κατάταξη ενός τυχαίου σημείου.

Η ευθεία ϵ_1 διαχωρίζει τις κλάσεις 1-2, η ϵ_2 τις 1-3 και η ϵ_3 τις 2-3. Το σημείο A με βάση τη συνάρτηση απόφασης των 1-2 (ευθεία ϵ_1) κατατάσσεται στην κλάση 2, ενώ με βάση την ϵ_2 στην κλάση 1. Τέλος με βάση την ευθεία ϵ_3 κατατάσσεται στην κλάση 2. Άρα, συγκεντρωτικά έχουμε:

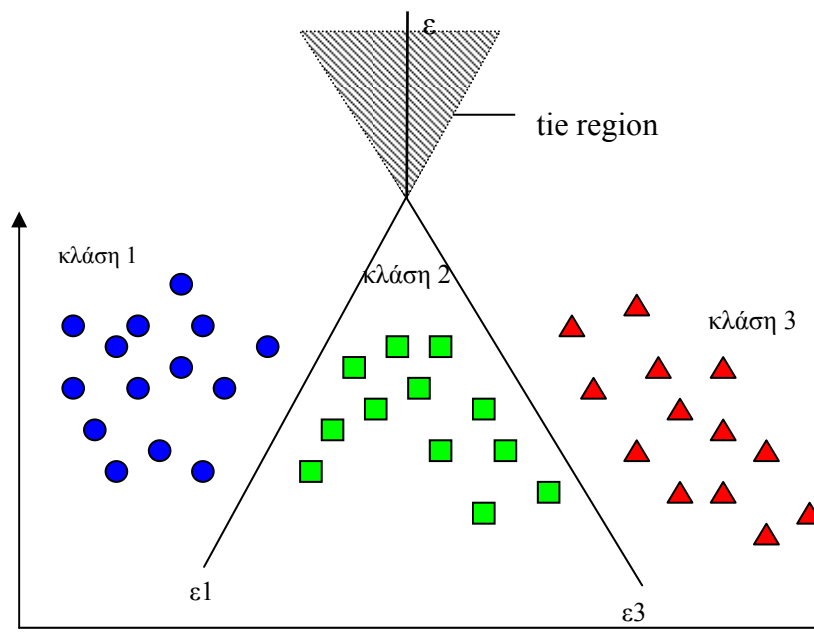
	κλάση 1	κλάση 2	κλάση 3
ευθεία ϵ_1	0	1	0
ευθεία ϵ_2	1	0	0
ευθεία ϵ_3	0	1	0
σύνολο	1	2	0

Πίνακας 2.1. Κατάταξη του σημείου A στις κλάσεις του παραδείγματος στο Σχήμα 2.13.

Από τον πίνακα βλέπουμε πως το σημείο A κατατάχθηκε 2 φορές στην κλάση 2 που είναι και ο μεγαλύτερος αριθμός κατατάξεων σε μία κλάση. Άρα, τελικά το σημείο A κατατάσσεται σωστά στην κλάση 2.

Σε πολλές περιπτώσεις με την εφαρμογή αυτής της διαδικασίας έχουμε περιοχές αλληλοεπικάλυψης (tie regions) όπως φαίνεται στο Σχήμα 2.14. Σε αυτές τις περιπτώσεις ακολουθείται η εξής διαδικασία (Debnath et al. (2004)):

Οι συναρτήσεις (ευθείες στο παράδειγμά μας) που ευθύνονται για την αλληλοεπικάλυψη, αντικαθίστανται με μία μόνο ευθεία στην περιοχή αλληλοεπικάλυψης. Στο Σχήμα 2.14 φαίνεται η περιοχή αλληλοεπικάλυψης (διαγραμμισμένη περιοχή), καθώς και η ευθεία αντικατάστασης ϵ .



Σχήμα 2.14. Περιοχή αλληλοεπικάλυψης (tie region) και η ευθεία αντικατάστασης.

Σε επόμενο κεφάλαιο θα προταθεί μια τεχνική που αντιμετωπίζει το πρόβλημα αυτό υπολογιστικά. Δηλαδή, στην περίπτωση που δύο (ή και παραπάνω κλάσεις) ισοψηφούν, πως θα αποφασιστεί η κλάση κατάταξης του σημείου.

2.5.4 DAGSVM

Η μέθοδος αυτή ακολουθεί τη λογική του δέντρου απόφασης. Τα φύλλα του “δέντρου” αντιπροσωπεύουν τις κλάσεις του προβλήματός μας και τα κλαδιά τη

κατεύθυνση που θα ακολουθήσουμε κατά την κατάταξη ενός σημείου ώστε να φτάσουμε στο φύλλο (κλάση). Τα σημεία που ξεκινούν τα κλαδιά ονομάζονται κόμβοι και αντιπροσωπεύουν τις συναρτήσεις απόφασης. Έτσι, κάθε κόμβος έχει 0-2 κλαδιά που αναχωρούν από αυτόν. Σε ένα πρόβλημα k κλάσεων υπάρχουν k φύλλα και $k(k-1)/2$ κόμβοι. Η διαδικασία της εκπαίδευσης είναι όμοια με αυτή της μεθόδου one-against-one. Για να κατατάξουμε ένα σημείο ξεκινάμε από τη “ρίζα” του δέντρου και υπολογίζουμε τη συνάρτηση απόφασης αυτού του κόμβου (της ρίζας). Ανάλογα με το αποτέλεσμα, κινούμαστε στο αντίστοιχο κλαδί από αυτά που αναχωρούν από τη ρίζα. Στη συνέχεια συναντάμε τον επόμενο κόμβο, υπολογίζουμε τη συνάρτηση απόφασης του και πάλι ανάλογα του αποτελέσματος κινούμαστε στο αντίστοιχο κλαδί που αναχωρεί από τον κόμβο αυτό. Προχωράμε έτσι ως ότου καταλήξουμε σε κάποιο φύλλο και κατατάσσουμε το σημείο στην κλάση που αντιστοιχεί στο φύλλο που καταλήξαμε.

Η διαδικασία της εκπαίδευσης καταναλώνει τον ίδιο χρόνο με της μεθόδου one-against-one αλλά η διαδικασία της κατάταξης χρειάζεται λιγότερο χρόνο. Όμως, η μέθοδος αυτή δε μπορεί να αντιμετωπίσει τις καταστάσεις αλληλοεπικάλυψης των ορίων των κλάσεων (tie regions).

2.5.5 Αξιολόγηση των μεθόδων

Έχει παρατηρηθεί με τη χρήση των παραπάνω μεθόδων σε διάφορα προβλήματα (*Hsu and Lin (2002)*) ότι η επιτυχία όλων των παραπάνω μεθόδων είναι παραπλήσια και έτσι δε μπορεί να βγει κάποιο στατιστικό συμπέρασμα για το αν υπερτερεί κάποια εξ’ αυτών. Ακόμα, η one-against-all και η all-together, χρησιμοποιούν όλα τα σημεία εκπαίδευσης κατά την διαδικασία της εκπαίδευσης και για αυτό η διαδικασία αυτή απαιτεί αρκετό χρόνο. Όσον αφορά στο χρόνο εκπαίδευσης, παρατηρήθηκε πως η one-against-one και η DAGSVM είναι οι καλύτερες. Με βάση την παρατήρηση της συμπεριφοράς των μεθόδων αυτών σε διάφορα προβλήματα προτείνεται η χρήση των one-against-one και DAGSVM.

Κεφάλαιο 3: Εφαρμογή των SVMs στη διαγνωστική αεριοστροβίλων

3.1 Γενικά

Τα όσα αναφέρθηκαν ως τώρα, αφορούν στον διαχωρισμό κλάσεων και στην κατάταξη ενός νέου σημείου σε κάποια κλάση. Η λογική των κλάσεων και η κατάταξη ενός σημείου μπορούν να εφαρμοστούν στη διαγνωστική ώστε να δημιουργηθεί ένας αλγόριθμος διάγνωσης βλαβών.

Μια μηχανή μπορεί να παρουσιάζει διάφορες καταστάσεις λειτουργίας ανάλογα με το αν υπάρχει βλάβη στη μηχανή ή όχι, όπως επίσης και με το μέγεθος της βλάβης. Αυτές τις δυνατές καταστάσεις λειτουργίας μπορούμε να τις θεωρήσουμε ως κλάσεις και έτσι μπορούμε να τις διαχωρίσουμε με τις Μηχανές Διανυσμάτων Υποστήριξης. Με αυτό το τρόπο, αφού έχουμε επιτύχει το διαχωρισμό των κλάσεων, κάθε εξεταζόμενη κατάσταση λειτουργίας θα κατατάσσεται σε μια από αυτές τις κλάσεις όπως γινόταν και στην περίπτωση της κατάταξης ενός σημείου σε κάποια κλάση.

3.2 Διάγνωση βλαβών αεριοστροβίλου

Κατά τη λειτουργία μιας μηχανής με βλάβη σε κάποια συνιστώσα της, επηρεάζονται κάποια μεγέθη της συνιστώσας αυτής και συνεπώς της μηχανής. Το μέγεθος της επίδρασης αυτής εξαρτάται από το μέγεθος της βλάβης. Για παράδειγμα, σε έναν αεριοστροβίλο, μια βλάβη μπορεί να επηρεάσει το βαθμό απόδοσης, την παροχή του εργαζόμενου μέσου καθώς και την διατομή του ακροφυσίου. Οι τιμές αυτών των “διαγνωστικών” μεγεθών σε λειτουργία της μηχανής παρουσία βλάβης διαφέρουν από τις τιμές που θα είχαν αυτά τα μεγέθη κατά την υγιή λειτουργία. Το μέγεθος της διαφοράς εξαρτάται από το μέγεθος της βλάβης. Ο λόγος της τιμής ενός τέτοιου μεγέθους προς την τιμή που έχει σε υγιή λειτουργία ονομάζεται παράμετρος

λειτουργίας (ή παράμετρος υγείας). Οι λόγοι όλων αυτών των μεγεθών σχηματίζουν το διάνυσμα \vec{f} .

Για τον προσδιορισμό της τιμής των ‘διαγνωστικών’ μεγεθών και συνεπώς των παραμέτρων λειτουργίας, χρησιμοποιούμε τις μετρήσεις κάποιων μεγεθών που λαμβάνουμε από τη μηχανή και αποτελούν το διάνυσμα \bar{Y} . Ενδεικτικά μετρούμενα μεγέθη σε έναν αεριοστρόβιλο είναι: πιέσεις και θερμοκρασίες σε διάφορες θέσεις της μηχανής, στροφές του άξονα (ή των αξόνων) της μηχανής, κραδασμοί κτλ. Για αδιαστατοποίηση, από το διάνυσμα \bar{Y} που παρήγαμε αφαιρούμε το διάνυσμα \bar{Y}_0 του οποίου οι συνιστώσες είναι οι τιμές των μετρούμενων μεγεθών στην υγιή λειτουργία της μηχανής και διαιρούμε με αυτό. Έτσι, δημιουργούμε ένα διάνυσμα $\Delta\bar{Y} = \frac{\bar{Y} - \bar{Y}_0}{\bar{Y}_0}$.

Με την αδιαστατοποίηση αυτή εκφράζουμε τα μετρούμενα μεγέθη στη μορφή της ποσοστιαίας απόκλισής τους από την ονομαστική τους τιμή. Έτσι η τιμή τους επηρεάζεται μόνο από την κατάσταση υγείας της μηχανής και όχι από το σημείο λειτουργίας της (Μαθιουδάκης (2007)).

Όπως αναφέρθηκε, μια μηχανή μπορεί να παρουσιάζει διάφορες καταστάσεις λειτουργίας ανάλογα με το αν είναι υγιής ή αν υπάρχει βλάβη, όπως επίσης και με το μέγεθος της βλάβης. Σε κάθε κατάσταση λειτουργίας έχουμε απόκλιση λόγω βλάβης σε μία ή και περισσότερες παραμέτρους υγείας, οπότε μπορούμε να πούμε ότι οι καταστάσεις λειτουργίας ορίζονται από τις παραμέτρους υγείας που αποκλίνουν κάθε φορά. Αυτές οι καταστάσεις λειτουργίας αποτελούν τις κλάσεις με τις οποίες θα εκπαιδευτεί το διαγνωστικό σύστημα. Τα σημεία της κάθε κλάσης (κατάστασης λειτουργίας) είναι τα διανύσματα $\Delta\bar{Y}$, δηλαδή οι αδιαστατοποιημένες μετρήσεις των μεγεθών της μηχανής που λαμβάνονται όταν αυτή λειτουργεί με τη συγκεκριμένη βλάβη.

Έτσι, το πρόβλημα της διαγνωστικής ανάγεται στον διαχωρισμό των πιθανών καταστάσεων λειτουργίας (κλάσεις) και την κατάταξη της εξεταζόμενης κατάστασης λειτουργίας σε μία από αυτές τις κλάσεις. Δηλαδή, με βάση τα μετρούμενα μεγέθη της μηχανής (διάνυσμα $\Delta\bar{Y}$) μπορούμε να δούμε σε ποια κατάσταση λειτουργίας βρίσκεται αυτή. Με βάση τα παραπάνω, έχουμε ένα πρόβλημα πολλών κλάσεων όπως αυτό περιγράφηκε στο κεφάλαιο 2 και μπορεί να αντιμετωπιστεί με τη λογική της μεθόδου των SVMs.

3.3 Ανάπτυξη των SVM για ταξινόμηση βλαβών αεριοστροβίλου

Κάθε πιθανή κατάσταση λειτουργίας ενός αεριοστροβίλου αποτελεί και μία κλάση, στην οποία αποκλίνουν μέσα σε ένα εύρος κάθε φορά, μία ή περισσότερες παράμετροι υγείας. Τα σημεία κάθε τέτοιας κλάσης είναι τα διανύσματα $\Delta\bar{Y}$ (μετρούμενα μεγέθη) που λαμβάνονται από τη μηχανή κατά τη λειτουργία της με τη συγκεκριμένη βλάβη.

Αυτά τα διανύσματα $\Delta\bar{Y}$ της κάθε κλάσης μπορούν να παραχθούν είτε από λήψη μετρήσεων από πραγματική μηχανή που λειτουργεί με την παρουσία συγκεκριμένης βλάβης (απόκλιση παραμέτρων υγείας), είτε από υπολογιστικό μοντέλο που προσομοιώνει τη λειτουργία της μηχανής. Η εφαρμογή της δεύτερης περίπτωσης θα αναλυθεί στη συνέχεια.

3.3.1 Εκπαίδευση

Για έναν αεριοστροβίλο, θεωρούμε τις καταστάσεις λειτουργίας (κλάσεις) έτσι ώστε σε καθεμία να αποκλίνουν, μέσα σε ένα συγκεκριμένο εύρος κάποιες παράμετροι υγείας.

Στην περίπτωση ύπαρξης υπολογιστικού μοντέλου προσομοίωσης, για κάθε κατάσταση λειτουργίας που θεωρήθηκε, “σαρώνουμε” το εύρος της απόκλισης των παραμέτρων υγείας με ένα συγκεκριμένο βήμα και με τη βοήθεια του μοντέλου προσομοίωσης λαμβάνουμε κάθε φορά το διάνυσμα $\Delta\bar{Y}$. Δηλαδή, για κάθε “βήμα” στη διαδικασία της σάρωσης λαμβάνουμε ένα διάνυσμα $\Delta\bar{Y}$. Αφού η κάθε κλάση ορίζεται από την απόκλιση συγκεκριμένων παραμέτρων υγείας μέσα σε ένα συγκεκριμένο εύρος, με βάση τα παραπάνω, για κάθε κλάση θα έχουμε ένα σύνολο διανυσμάτων $\Delta\bar{Y}$.

Συνιστώσες του διανύσματος $\Delta\bar{Y}$ είναι οι ποσοστιαίες αποκλίσεις των τιμών των μετρούμενων μεγεθών στην συγκεκριμένη κατάσταση λειτουργίας από τις τιμές των μεγεθών αυτών στην υγιή κατάσταση λειτουργίας. Αν το κάθε τέτοιο διάνυσμα $\Delta\bar{Y}$ το δούμε ως σημείο με συντεταγμένες τις συνιστώσες του διανύσματος, μπορούμε να πούμε πως έχουμε τα σημεία της κάθε κλάσης, άρα και την πληροφορία

εκπαίδευσης. Έτσι, κάθε κλάση αποτελείται από σημεία \vec{x}_i που το καθένα είναι ένα διάνυσμα με συνιστώσες τις αποκλίσεις των μετρούμενων μεγεθών από την υγιή λειτουργία (διάνυσμα $\Delta\bar{Y}$). Με αυτό το τρόπο γνωρίζουμε ένα σύνολο σημείων εκπαίδευσης (training patterns), καθώς και την κλάση που ανήκουν. Συνεπώς είναι γνωστή η πληροφορία της εκπαίδευσης (training set) η οποία πραγματοποιείται όπως περιγράφηκε στο κεφάλαιο 2 για τον διαχωρισμό των κλάσεων.

Για το πρόγραμμα διάγνωσης που δημιουργήθηκε για την παρούσα διπλωματική, χρησιμοποιήθηκε η μέθοδος one-against-one. Έτσι οι κλάσεις (καταστάσεις λειτουργίας) χωρίζονται σε όλα τα πιθανά ζεύγη και βρίσκεται το υπερεπίπεδο που διαχωρίζει κάθε ζεύγος. Για την αντιμετώπιση της περίπτωσης αλληλοεπικάλυψης των κλάσεων ή αυτή των μη γραμμικά διαχωρίσιμων κλάσεων υπάρχει η επιλογή χρήσης συναρτήσεων κελύφους.

Κατά τη διαδικασία της εκπαίδευσης, για τον διαχωρισμό του κάθε ζεύγους κλάσεων, απαιτείται να προσδιοριστούν τα σημεία της κάθε κλάσης που λειτουργούν ως διανύσματα υποστήριξης (SVs) και προσδιορίζουν το υπερεπίπεδο που διαχωρίζει το ζεύγος. Όπως αναπτύχθηκε στο κεφάλαιο 2, αυτό πραγματοποιείται με την λύση ενός προβλήματος τετραγωνικού προγραμματισμού (Quadratic Programming Problem). Στην παρούσα εργασία, αυτό επιτυγχάνεται με τη χρήση έτοιμου αλγορίθμου (DQPROG), ο οποίος παρέχεται μέσα από τις συναρτήσεις βιβλιοθήκης της IMSL. Ο κώδικας αυτός βασίζεται στην υλοποίηση του αλγορίθμου των Goldfarb and Idnani για λύση δυαδικών τετραγωνικών προβλημάτων (dual quadratic programming) από τον M.J.D. Powell (*Visual Numerics*). Ο αλγόριθμος αυτός παρουσιάζεται στο Παράρτημα II του παρόντος, μέσα από το εγχειρίδιο χρήσης του από την IMSL.

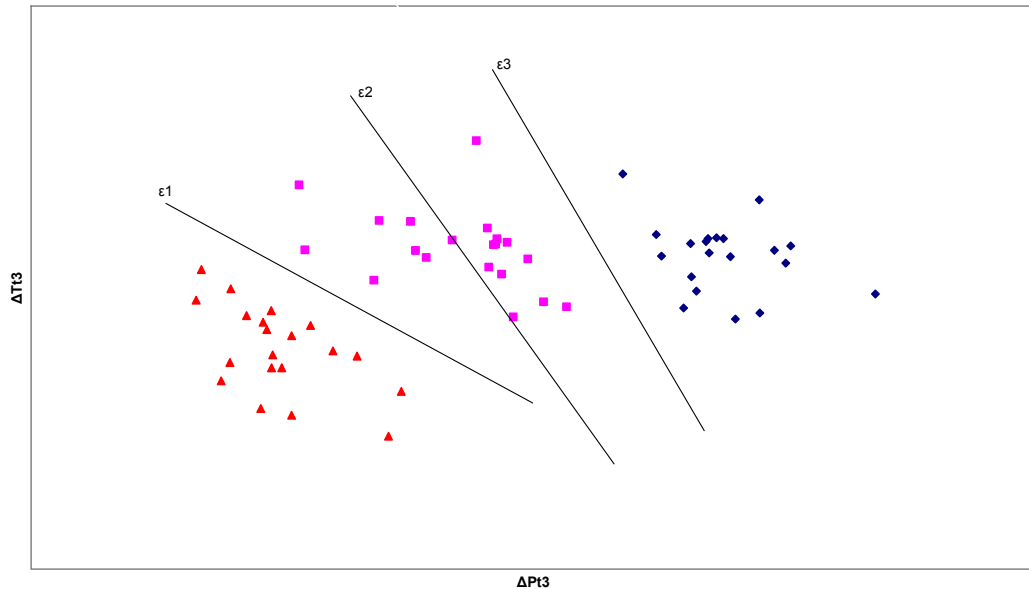
Στο Σχήμα 3.1 απεικονίζονται τα σημεία (διανύσματα $\Delta\bar{Y}$) όπως αυτά προέκυψαν για τρεις καταστάσεις λειτουργίας μιας μηχανής. Κάθε κατάσταση λειτουργίας αποτελεί και μία κλάση, έτσι έχουμε:

Κλάση I : απόκλιση λόγω βλάβης της παραμέτρου παροχής μάζας του στροβίλου υψηλής.

Κλάση II: απόκλιση λόγω βλάβης της παραμέτρου παροχής μάζας και της παραμέτρου βαθμού απόδοσης του συμπιεστή υψηλής.

Κλάση III : απόκλιση λόγω βλάβης της παραμέτρου επιφανείας του ακροφυσίου.

Στο συγκεκριμένο παράδειγμα, για ευκολία στην απεικόνιση, το διάνυσμα $\Delta\bar{Y}$ αποτελείται από δύο μόνο μεγέθη, την απόκλιση από την υγιή λειτουργία της ολικής θερμοκρασίας (ΔTt) και της ολικής πίεσης (ΔPt) στη θέση 3. Οι ευθείες ϵ_1 , ϵ_2 και ϵ_3 διαχωρίζουν τις τρεις κλάσεις (ϵ_1 : I-II, ϵ_2 : I-III, ϵ_3 : II-III).



Σχήμα 3.1. Διαχωρισμός τριών καταστάσεων λειτουργίας προσδιοριζόμενων από δύο μεγέθη (ΔTt_3 , ΔPt_3).

3.3.2 Συναρτήσεις κελύφους

Στο πρόγραμμα διαγνωστικής με τη μέθοδο SVM που αναπτύχθηκε, δίνεται η δυνατότητα στο χρήστη να επιλέξει μεταξύ κάποιων συναρτήσεων κελύφους. Δηλαδή, να επιλέξει την εξίσωση που δίνει το αποτέλεσμα του εσωτερικού γινομένου των σημείων στο χώρο μεγαλύτερων διαστάσεων (βλ. παράγραφο 2.3.3).

Επειδή η μέθοδος των SVM έχει χρησιμοποιηθεί ελάχιστα στη διαγνωστική αεριοστροβίλων, δεν είναι γνωστό ποιες συναρτήσεις κελύφους έχουν τη καλύτερη συμπεριφορά για τη φύση των δεδομένων του συγκεκριμένου προβλήματος. Για αυτό το λόγο, στο πρόγραμμα υπάρχει η δυνατότητα επιλογής τεσσάρων συναρτήσεων κελύφους, οι οποίες είναι και οι βασικές που χρησιμοποιήθηκαν γενικά για τα SVMs

και έχουν εν γένει την καλύτερη συμπεριφορά σε πληθώρα προβλημάτων, όπως αναφέρεται σε διάφορα σημεία της βιβλιογραφίας.

Τα kernels και οι παράμετροι που καθορίζονται από το χρήστη είναι τα εξής:

$$K(x, y) = x \cdot y \quad \text{Γραμμικό kernel (όχι απεικόνιση σε χώρο } H)$$

$$K(x, y) = (gama \cdot (x \cdot y) + coef)^{degree} \quad \text{Πολυωνυμικό kernel (Polynomial kernel)}$$

$$K(x, y) = e^{-gama \cdot \|x-y\|^2} \quad \text{RBF kernel}$$

$$K(x, y) = \tanh(gama \cdot (x \cdot y) - coef) \quad \text{Σιγμοειδές kernel (Sigmoid kernel)}$$

Οι συντελεστές $gama$, $coef$ και $degree$ καθορίζονται από το χρήστη και πρακτικά διαμορφώνουν το υπερεπίπεδο στο οποίο γίνεται η απεικόνιση χωρίς να χρειάζεται αυτό να είναι γνωστό.

Η επιλογή της βέλτιστης τιμής των μεταβλητών αυτών γίνεται πειραματικά, όπως θα φανεί καλύτερα και στις εφαρμογές που παρουσιάζονται στο κεφάλαιο 4.

3.3.3 Κατάταξη

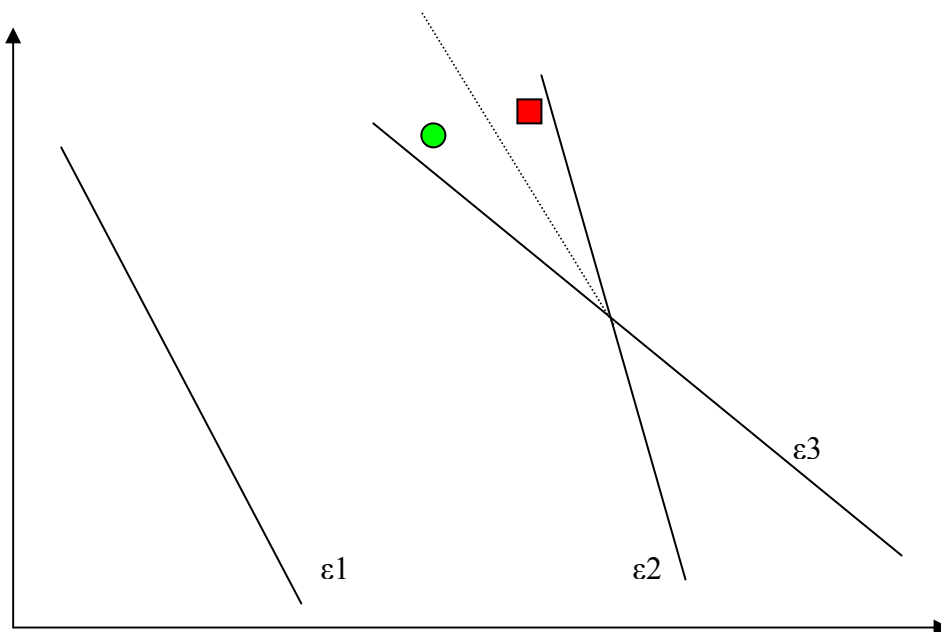
Κατά την λειτουργία της μηχανής μπορούμε να μετρήσουμε κάποια μεγέθη (διάνυσμα \bar{Y}), τα οποία καθορίζουν την κατάσταση λειτουργίας της μηχανής και που από αυτά προσδιορίζονται οι παράμετροι λειτουργίας. Αυτές οι μετρήσεις χρησιμοποιήθηκαν και για τον καθορισμό των σημείων εκπαίδευσης, όπως περιγράφηκε προηγουμένως.

Σκοπός μας τώρα είναι να δούμε σε ποία κατάσταση λειτουργίας (κλάση) αντιστοιχεί η εξεταζόμενη κατάσταση λειτουργίας της μηχανής που αντιπροσωπεύεται από το διάνυσμα \bar{Y} . Αρχικά και με χρήση του μοντέλου προσομοίωσης της μηχανής, δημιουργούμε το διάνυσμα $\Delta\bar{Y}$ το οποίο και

αντιπροσωπεύει την εξεταζόμενη κατάσταση λειτουργίας της μηχανής. Άρα, το πρόβλημα ανάγεται στην κατάταξη του σημείου (διάνυσμα $\Delta\bar{Y}$) σε μία από τις κλάσεις που έχουν ήδη διαχωριστεί. Η διαδικασία αυτή υλοποιείται όπως περιγράφηκε στο κεφάλαιο 2, ελέγχοντας τη σχετική θέση του προς κατάταξη σημείου με τα διαχωριστικά υπερεπίπεδα των κλάσεων με χρήση της συνάρτησης απόφασης.

Για την κατάταξη των σημείων, στον κώδικα, χρησιμοποιείται η μέθοδος max-wins. Έτσι, το κάθε σημείο ελέγχεται με τη συνάρτηση απόφασης (διαχωριστικό υπερεπίπεδο του ζεύγους) κάθε ζεύγους κλάσεων και τοποθετείται τελικά στην κλάση που κατατάχθηκε τις περισσότερες φορές.

Με τη μέθοδο max-wins μπορεί να εμφανιστούν κάποια προβλήματα, όπως η περίπτωση που το σημείο κατατάχθηκε ίσες φορές σε παραπάνω από μία κλάσεις. Παράδειγμα τέτοιας περίπτωσης στο δισδιάστατο χώρο για εύκολη απεικόνιση, παρουσιάζεται στο Σχήμα 3.2. Στο σχήμα αυτό παρουσιάζονται μόνο οι διαχωριστικές ευθείες τριών κλάσεων καθώς και δύο προς κατάταξη σημεία. Επίσης με διακεκομμένη γραμμή παρουσιάζεται και η ευθεία που αντικαθιστά τις δύο διαχωριστικές ευθείες στην περιοχή αλληλοεπικάλυψης (tie region), τεχνική που περιγράφηκε στην παράγραφο 2.5.3.



Σχήμα 3.2. Περίπτωση ισοβάθμισης κλάσεων κατά τη μέθοδο max-wins.

Οι ευθεία ε1 διαχωρίζει τις κλάσεις 1-2, η ε2 τις 1-3 και η ε3 τις 2-3. Έτσι για το σημείο που παριστάνεται με τον κύκλο και για το σημείο που παριστάνεται με το τετράγωνο θα έχουμε:

	κλάση 1	κλάση 2	κλάση 3
ευθεία ε1	0	1	0
ευθεία ε2	1	0	0
ευθεία ε3	0	0	1
σύνολο	1	1	1

Πίνακας 3.1. Κατάταξη του σημείου ‘κύκλος’.

	κλάση 1	κλάση 2	κλάση 3
ευθεία ε1	0	1	0
ευθεία ε2	1	0	0
ευθεία ε3	0	0	1
σύνολο	1	1	1

Πίνακας 3.2. Κατάταξη του σημείου ‘τετράγωνο’.

Όπως βλέπουμε από τους παραπάνω πίνακες κατάταξης, δε μπορεί να βγει συμπέρασμα σε ποια κλάση ανήκουν τα δύο αυτά σημεία.

Αν τώρα αντικαταστήσουμε τις ε2 και ε3 στην περιοχή της αλληλοεπικάλυψης με τη διακεκομμένη οι πίνακες κατάταξης διαμορφώνονται ως εξής:

	κλάση 1	κλάση 2	κλάση 3
ευθεία ε1	0	1	0
ευθεία ε2	1	0	0
ευθεία ε3	0	1	0
σύνολο	1	2	0

Πίνακας 3.3. Κατάταξη του σημείου ‘κύκλος’ με αντικατάσταση των ευθειών στην περιοχή αλληλοεπικάλυψης.

	κλάση 1	κλάση 2	κλάση 3
ευθεία ε1	0	1	0
ευθεία ε2	0	0	1
ευθεία ε3	0	0	1
σύνολο	0	1	2

Πίνακας 3.4. Κατάταξη του σημείου ‘τετράγωνο’ με αντικατάσταση των ευθειών στην περιοχή αλληλοεπικάλυψης.

Από τα παραπάνω, ο κύκλος κατατάσσεται στην κλάση 2 ενώ το τετράγωνο στην κλάση 3.

Η παραπάνω διαδικασία, αν εφαρμοζόταν στον κώδικα διάγνωσης, θα αύξανε την πολυπλοκότητα και το χρόνο που απαιτείται κατά την κατάταξη αφού θα χρειαζόταν να βρίσκεται η περιοχή αλληλοεπικάλυψης των δύο υπερεπιπέδων και στην περιοχή αυτή τα δύο διαχωριστικά υπερεπίπεδα να αντικαθίστανται από ένα. Για το λόγο αυτό, στα πλαίσια της παρούσας εργασίας, προτείνουμε μια καινούρια τεχνική, ισοδύναμη με την προαναφερθείσα, που έχει τα ίδια αποτελέσματα κατά την κατάταξη αλλά δεν απαιτεί την εύρεση της τομής των δύο υπερεπιπέδων και την αντικατάστασή τους από ένα (όπως χρειάζεται η διαδικασία που περιγράφηκε) και συνεπώς μπορούμε να υποθέσουμε πως είναι ταχύτερη και λιγότερο πολύπλοκη.

Η τεχνική αυτή είναι η εξής: Έστω ένα σημείο P που κατά τη διαδικασία της κατάταξης με τη μέθοδο 'max-wins', τοποθετείται ίσες φορές σε περισσότερες από μία κλάσεις. Ελέγχουμε την ευκλείδεια απόσταση αυτού του σημείου P από τα διαχωριστικά υπερεπίπεδα μόνο των ισοβαθμούντων κλάσεων. Έστω ότι το υπερεπίπεδο που διαχωρίζει τις κλάσεις i, j είναι αυτό με τη μικρότερη απόσταση από το σημείο P . Για το σημείο P υπολογίζουμε την συνάρτηση απόφασης F του υπερεπιπέδου αυτού ($F(P)$). Αν $F(P) > 0$ τότε το σημείο P κατατάσσεται στην κλάση j . Αν $F(P) < 0$ τότε το σημείο P κατατάσσεται στην κλάση i .

Στο παράδειγμα που απεικονίζεται στο Σχήμα 3.2, ο κύκλος βρίσκεται πιο κοντά στην ευθεία ϵ_3 , δηλαδή τη διαχωριστική των κλάσεων 2-3. Επειδή το σημείο αυτό βρίσκεται δεξιά της διαχωριστικής ευθείας (αρνητική η συνάρτηση απόφασης) θα καταταχθεί στην κλάση 2. Το τετράγωνο βρίσκεται πιο κοντά στην ευθεία ϵ_2 , δηλαδή σε αυτή που χωρίζει τις 1-3. Επειδή βρίσκεται στα αριστερά της ευθείας (θετική συνάρτηση απόφασης) θα καταταχθεί στην κλάση 3. Τα αποτελέσματα αυτά βλέπουμε ότι συμφωνούν με εκείνα της διαδικασίας αντικατάστασης των διαχωριστικών υπερεπιπέδων από ένα στη περιοχή της αλληλοεπικάλυψης.

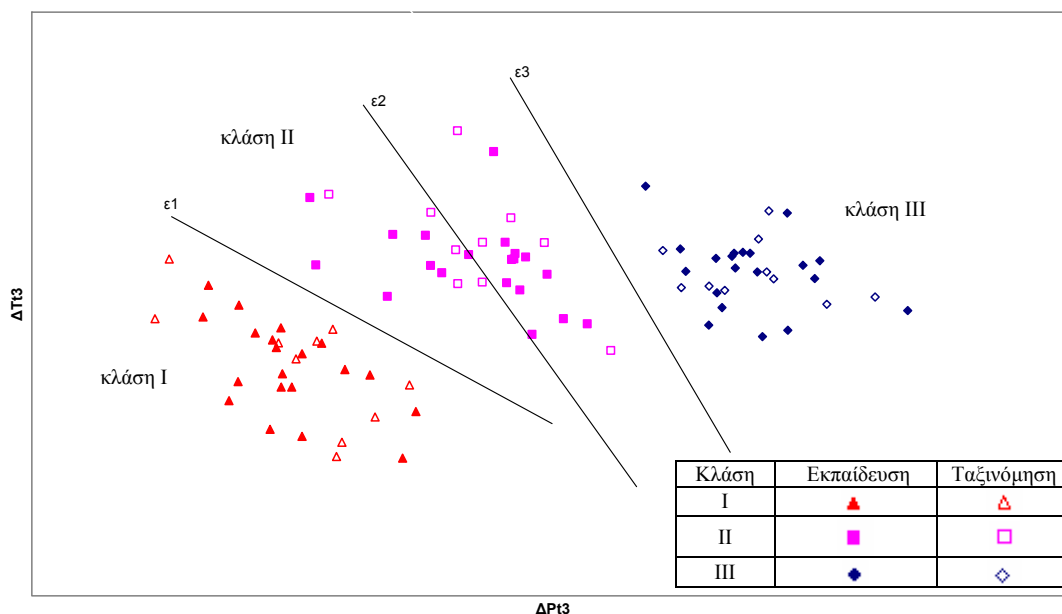
Ο παραπάνω αλγόριθμος χρησιμοποιήθηκε στον κώδικα διαγνωστικής που δημιουργήθηκε. Έτσι, αν ένα σημείο κατατάσσεται ίσες φορές σε πάνω από μια κλάσεις, τότε υπολογίζεται η απόστασή του από τα διαχωριστικά υπερεπίπεδα των ισοβαθμούντων κλάσεων και η απόφαση λαμβάνεται για το υπερεπίπεδο με τη μικρότερη απόσταση.

Αν ένα προς κατάταξη σημείο βρίσκεται πολύ κοντά στο διαχωριστικό υπερεπίπεδο δύο κλάσεων, ο κώδικας δίνει τη δυνατότητα το σημείο αυτό να καταταχθεί και στις δύο κλάσεις ταυτόχρονα. Η επιλογή αυτή δίνεται με τον καθορισμό ενός εύρους της συνάρτησης απόφασης. Αν η συνάρτηση απόφασης για

το σημείο αυτό βρίσκεται εντός αυτού του εύρους, τότε το σημείο κατατάσσεται και στις δύο κλάσεις. Το εύρος αυτό είναι της μορφής $[-\delta, +\delta]$. Η τελική τοποθέτηση του σημείου γίνεται σύμφωνα με την αρχή της max-wins, όπως σε οποιοδήποτε προς κατάταξη σημείο.

Όσο πιο κοντά στο διαχωριστικό υπερεπίπεδο βρίσκεται ένα προς κατάταξη σημείο, τόσο μεγαλύτερη αβεβαιότητα υπάρχει όσον αφορά στην κλάση που πραγματικά ανήκει το σημείο. Με την παραπάνω διαδικασία έχουμε τη δυνατότητα να “χαλαρώσουμε” τα όρια κατάταξης αν το σημείο βρίσκεται κοντά στο διαχωριστικό υπερεπίπεδο. Η τεχνική αυτή δεν επιφέρει προβλήματα στη διαδικασία της κατάταξης σε πραγματικά προβλήματα με μεγάλο αριθμό κλάσεων και σημείων, αφού δεν παίζει σημαντικό ρόλο η διπλή κατάταξη του σημείου σε μερικές από τις κλάσεις αυτές λόγω του μεγάλου αριθμού τους. Ακόμη όμως, και στην περίπτωση που λόγω αυτής της τεχνικής έχουμε ισοβάθμιση δύο κλάσεων το πρόβλημα λύνεται με τον αλγόριθμο της απόστασης του σημείου από το διαχωριστικό υπερεπίπεδο που παρουσιάστηκε προηγουμένως.

Στο Σχήμα 3.3 παρουσιάζεται η κατάταξη καινούριων σημείων (καταστάσεων λειτουργίας) στις κλάσεις του παραδείγματος που παρουσιάζει το Σχήμα 3.1. Από το σχήμα και τις διαχωριστικές ευθείες βλέπουμε πως όλες οι εξεταζόμενες καταστάσεις λειτουργίας έχουν καταταχθεί στις σωστές κλάσεις.



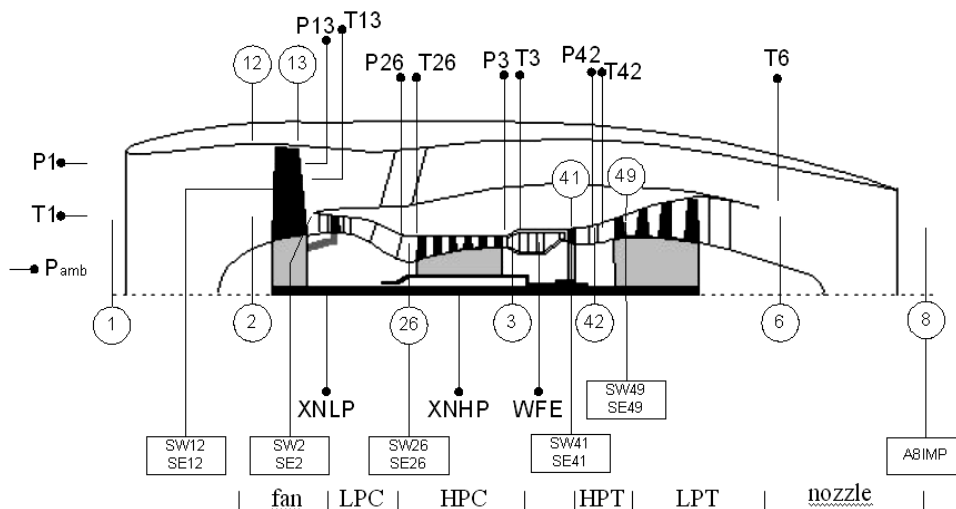
Σχήμα 3.3. Κατάταξη καινούριων σημείων (καταστάσεων λειτουργίας) στις κλάσεις του παραδείγματος που απεικονίζεται στο Σχήμα 3.1.

Κεφάλαιο 4: Διαγνωστική αεριοστροβίλου με χρήση SVM

4.1 Εντοπισμός βλαβών αεριοστροβίλου διπλού ρεύματος με στοιχεία θερμοδυναμικών μεγεθών

Η διαδικασία της διάγνωσης, όπως αυτή περιγράφηκε στο κεφάλαιο 3, εφαρμόστηκε για βλάβες θερμοδυναμικής φύσεως σε συγκεκριμένο αεριοστρόβιλο. Η περίπτωση αυτή του αεριοστροβίλου που εξετάστηκε δεν αντιστοιχεί σε υπαρκτό λειτουργούντα αεριοστρόβιλο, αλλά τα χαρακτηριστικά του αντικατοπτρίζουν έναν σύγχρονο αεροπορικό κινητήρα διπλού ρεύματος πολιτικής αεροπορίας. (*Stamatis et al. (2001)*).

Στο Σχήμα 4.1 απεικονίζεται ο κινητήρας που θεωρήθηκε, καθώς και οι μεταβλητές που χρησιμοποιούνται για να περιγράψουν τη λειτουργία του.



Σχήμα 4.1. Σχηματική απεικόνιση του κινητήρα με τις χρησιμοποιούμενες μεταβλητές.

Οι συνιστώσες της μηχανής, κατά τη φορά κίνησης του εργαζόμενου μέσου είναι: ο ανεμιστήρας (fan), ο συμπιεστής χαμηλής πίεσης (LPC), ο συμπιεστής

υψηλής πίεσης (HPC), ο στρόβιλος υψηλής πίεσης (HPT), ο στρόβιλος χαμηλής πίεσης (LPT), και το ακροφύσιο (nozzle). Από αυτές τις συνιστώσες, οι τρεις πρώτες αποτελούν το “ψυχρό τμήμα” της μηχανής, ενώ οι υπόλοιπες το “θερμό τμήμα”.

Το σημείο λειτουργίας της μηχανής καθορίζεται από 4 μεγέθη, που είναι τα εξής: πίεση περιβάλλοντος (P_{amb}), πίεση εισόδου του κινητήρα ($P1$), θερμοκρασία εισόδου του κινητήρα ($T1$) και παροχή καυσίμου (WFE). Τα μεγέθη αυτά φαίνονται στο Σχήμα 4.1 και αποτελούν το διάνυσμα \bar{u} .

Η κατάσταση υγείας κάθε συνιστώσας, εκτός του ακροφυσίου, αντιπροσωπεύεται από δύο λειτουργικούς δείκτες, την ‘παράμετρο παροχής μάζας’ και την ‘παράμετρο βαθμού απόδοσης’ που ορίζονται ως εξής:

$$\text{Παράμετρος παροχής μάζας: } SW_i = \frac{W_i \cdot \sqrt{T_i}}{p_i} \left/ \left(\frac{W_i \cdot \sqrt{T_i}}{p_i} \right)_{ref} \right. \quad (4.1)$$

$$\text{Παράμετρος βαθμού απόδοσης: } SE_i = \frac{\eta_i}{(\eta_i)_{ref}} \quad (4.2)$$

όπου i είναι η αρίθμηση της θέσης στο σημείο εισόδου της συνιστώσας, W_i η παροχή μάζας στη θέση i , T_i και p_i είναι αντίστοιχα η ολική θερμοκρασία και πίεση στη θέση i και η_i , ο ισεντροπικός βαθμός απόδοσης της συνιστώσας. Ο δείκτης ref αναφέρεται στις ονομαστικές τιμές των αντίστοιχων μεγεθών για την περίπτωση υγιούς λειτουργίας του κινητήρα.

Η κατάσταση λειτουργίας του ακροφυσίου αντιπροσωπεύεται από έναν λειτουργικό δείκτη, την ‘παράμετρο επιφάνειας ακροφυσίου’ (*Stamatis et al. (1989)*), που ορίζεται ως εξής:

$$A8IMP = \frac{A8}{A8_{ref}} \quad (4.3)$$

όπου $A8$ το εμβαδόν της επιφάνειας του ακροφυσίου στη θέση 8.

Οι λειτουργικοί αυτοί δείκτες (παράμετροι υγείας) εκφρασμένοι για κάθε συνιστώσα της μηχανής, όπως φαίνεται και στο Σχήμα 4.1, αποτελούν το διάνυσμα

\bar{f} . Έτσι, το διάνυσμα \bar{f} απεικονίζει την κατάσταση λειτουργίας όλων των συνιστωσών της μηχανής (κατά συνέπεια και της ίδιας της μηχανής) και αποτελείται από 11 παραμέτρους υγείας, όπως τις απεικονίζει ο Πίνακας 4.1.

Συμβολισμός	Περιγραφή
SW12	Παράμετρος παροχής μάζας ανεμιστήρα
SE12	Παράμετρος βαθμού απόδοσης ανεμιστήρα
SW2	Παράμετρος παροχής μάζας συμπιεστή χαμηλής
SE2	Παράμετρος βαθμού απόδοσης συμπιεστή χαμηλής
SW26	Παράμετρος παροχής μάζας συμπιεστή υψηλής
SE26	Παράμετρος βαθμού απόδοσης συμπιεστή υψηλής
SW41	Παράμετρος παροχής μάζας στροβίλου υψηλής
SE41	Παράμετρος βαθμού απόδοσης στροβίλου υψηλής
SW49	Παράμετρος παροχής μάζας στροβίλου χαμηλής
SE49	Παράμετρος βαθμού απόδοσης στροβίλου χαμηλής
A8IMP	Παράμετρος επιφάνειας ακροφυσίου

Πίνακας 4.1. Οι παράμετροι λειτουργίας που περιγράφουν την υγεία των συνιστωσών του κινητήρα.

Για την παρακολούθηση της λειτουργίας της μηχανής και τον προσδιορισμό των στοιχείων του διανύσματος \bar{f} χρειαζόμαστε ένα σύνολο μετρούμενων μεγεθών (διάνυσμα \bar{Y}). Ένα τέτοιο κατάλληλο σύνολο μετρούμενων μεγεθών (Mathioudakis et al. (2004)), απεικονίζει ο Πίνακας 4.2.

Συμβολισμός	Περιγραφή
XNLP	Στροφές άξονα χαμηλής
XNHP	Στροφές άξονα υψηλής
P13	Ολική πίεση εξόδου ανεμιστήρα
P3	Ολική πίεση εξόδου συμπιεστή υψηλής
T3	Θερμοκρασία εξόδου συμπιεστή υψηλής
T6	Θερμοκρασία εξόδου στροβίλου χαμηλής
T13	Θερμοκρασία εξόδου ανεμιστήρα
P26	Ολική πίεση εξόδου συμπιεστή χαμηλής
T26	Θερμοκρασία εξόδου συμπιεστή χαμηλής
P42	Ολική πίεση εισόδου στροβίλου υψηλής
T42	Θερμοκρασία εισόδου στροβίλου υψηλής

Πίνακας 4.2. Επιθυμητό σύνολο μετρούμενων μεγεθών για την παρακολούθηση της λειτουργίας του κινητήρα.

Το σύνολο που απεικονίζει ο Πίνακας 4.2 αποτελεί το επιθυμητό σύνολο των μετρούμενων μεγεθών για τον προσδιορισμό της κατάστασης λειτουργίας της κάθε συνιστώσας της μηχανής. Όμως, η μέτρηση των μεγεθών $P26$, $T26$, $P42$, $T42$ συνήθως δεν είναι εφικτή λόγω του μεγάλου κόστους του μετρητικού εξοπλισμού. Έτσι, η παρακολούθηση της λειτουργίας του κινητήρα γίνεται με ένα σετ επτά μετρήσεων που αποτελεί το διάνυσμα $\bar{Y} = (XNLP, XNHP, P13, P3, T3, T6, T13)$.

Οι μετρούμενες τιμές των μεγεθών του διανύσματος \bar{Y} παρουσιάζουν μια απόκλιση από τις πραγματικές τιμές λόγω παρουσίας θορύβου. Ο Πίνακας 4.3 παρουσιάζει τη στάθμη του θορύβου για κάθε μετρούμενο μέγεθος.

Μέγεθος	Θόρυβος
Pamb	33,333 Pa
T1	0,667 K
P1	33,333 Pa
WFE	0,00067 gr/sec
XNLP	2 rpm
XNHP	4 rpm
P13	100 Pa
P26	150 Pa
T26	0,667 K
P3	1666,667 Pa
T3	0,667 K
T6	0,667 K
T13	0,667 K
P42	365 Pa
T42	0,667 K

Πίνακας 4.3. Ο θόρυβος στα μετρούμενα μεγέθη.

4.1.1 Διάγνωση βλαβών συνιστωσών αεριοστροβίλου

Αρχικά, εξετάζουμε βλάβες των συνιστωσών του προαναφερθέντος κινητήρα. Οι βλάβες αυτές επηρεάζουν την επίδοση και τη συμπεριφορά των αντίστοιχων συνιστωσών με αποτέλεσμα την απόκλιση των παραμέτρων υγείας τους.

4.1.1.1 Δεδομένα Εκπαίδευσης των SVMs

Αρχικά, θεωρήθηκαν 18 περιπτώσεις λειτουργίας της μηχανής που αντιπροσωπεύει τις πιθανές καταστάσεις λειτουργίας ενός αεριοστροβίλου διπλού ρεύματος όπως αυτός απεικονίζεται στο Σχήμα 4.1. Από τις 18 περιπτώσεις, οι 17 αντιστοιχούν σε λειτουργία της μηχανής με ύπαρξη βλάβης σε κάποια συνιστώσα της, ενώ η μία αντιστοιχεί στην λειτουργία χωρίς ύπαρξη βλάβης (υγιής κατάσταση λειτουργίας). Αυτές οι 17 περιπτώσεις προκύπτουν από βλάβες που γενικά έχει διαπιστωθεί ότι μπορούν να εμφανιστούν σε μια μηχανή.

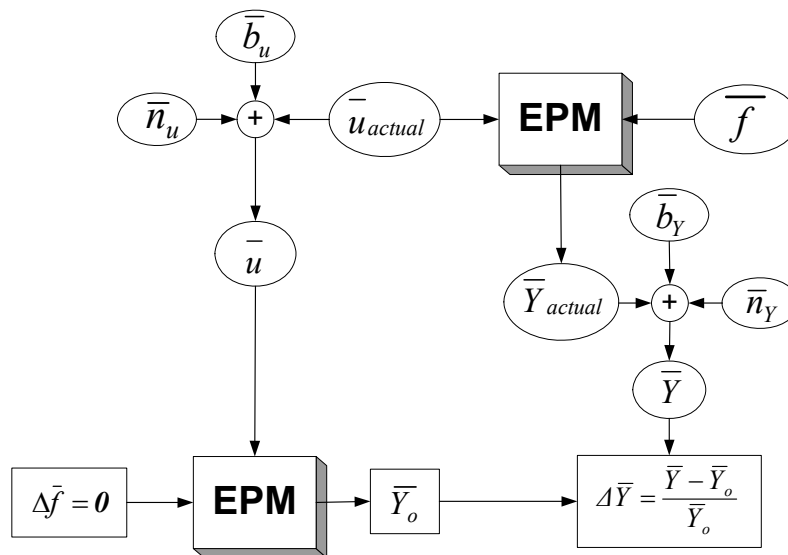
Ο καθορισμός των περιπτώσεων αυτών γίνεται γνωρίζοντας ότι σε περίπτωση λειτουργίας με βλάβη αποκλίνει, μέσα σε ένα συγκεκριμένο εύρος, μια τουλάχιστον παράμετρος υγείας της μηχανής.

Με βάση τα παραπάνω, ο Πίνακας 4.4 παρουσιάζει τις 18 περιπτώσεις λειτουργίας που θεωρήθηκαν, τη παράμετρο (ή τις παραμέτρους) υγείας που αποκλίνει, καθώς και το εύρος αυτής της απόκλισης. Οι 18 αυτές περιπτώσεις είναι οι 18 κλάσεις με τις οποίες θα εκπαιδευτεί το διαγνωστικό πρόγραμμα. Να σημειωθεί, πως θεωρούμε λειτουργία της μηχανής με ύπαρξη βλάβης αν μία τουλάχιστον παράμετρος υγείας παρουσιάζει απόκλιση $+0.5$ και άνω (ή -0.5 και κάτω).

α/α κλάσης	Παράμετροι που αποκλίνουν σημαντικά
1	$\Delta SW12=\alpha$
2	$\Delta SE12=\alpha$
3	$(\Delta SW12=\alpha, \Delta SE12=\alpha)$ ή $(\Delta SW12=\alpha, \Delta SE12=\alpha/2)$
4	$\Delta SW2=\alpha$
5	$\Delta SE2=\alpha$
6	$(\Delta SW2=\alpha, \Delta SE2=\alpha)$ ή $(\Delta SW2=\alpha, \Delta SE2=\alpha/2)$
7	$(\Delta SW12=\alpha, \Delta SE12=\alpha, \Delta SW2=\alpha, \Delta SE2=\alpha)$ ή $(\Delta SW12=\alpha, \Delta SE12=\alpha/2, \Delta SW2=\alpha, \Delta SE2=\alpha/2)$
8	$\Delta SW26=\alpha, \Delta SW26=-\alpha$
9	$\Delta SE26=\alpha$
10	$(\Delta SW26=\alpha, \Delta SE26=\alpha)$ ή $(\Delta SW26=\alpha, \Delta SE26=\alpha/2)$
11	$\Delta SW41=\alpha, \Delta SW41=-\alpha$
12	$\Delta SE41=\alpha$
13	$(\Delta SW41=\alpha, \Delta SE41=\alpha)$ ή $(\Delta SW41=\alpha, \Delta SE41=\alpha/2)$
14	$\Delta SW49=\alpha, \Delta SW49=-\alpha$
15	$\Delta SE49=\alpha$
16	$(\Delta SW49=\alpha, \Delta SE49=\alpha)$ ή $(\Delta SW49=\alpha, \Delta SE49=\alpha/2)$
17	$A8IMP=\alpha$
18	Οποιαδήποτε κατά β

Πίνακας 4.4. Οι περιπτώσεις λειτουργίας (κλάσεις) που θεωρήθηκαν ($\alpha \in [-2.5, -0.5]$, $\beta \in (-0.5, +0.5)$, βήμα: 0.1).

Για την παραγωγή της πληροφορίας εκπαίδευσης (training points) του διαγνωστικού προγράμματος της μηχανής διανυσμάτων υποστήριξης, έγινε προσομοίωση της λειτουργίας της μηχανής (Σχήμα 4.2) με τις αποκλίσεις των παραμέτρων υγείας που θεωρήθηκαν (Πίνακας 4.4). Έτσι, για κάθε σημείο λειτουργίας της μηχανής, παίρνουμε προσομοιωμένα τα μεγέθη που θα μετρούσαμε σε λειτουργούσα μηχανή και αποτελούν το διάνυσμα $\bar{Y} = (XNLP, XNHP, P13, P3, T3, T6, T13)$. Στη συνέχεια παράγεται το διάνυσμα $\Delta\bar{Y}$ που ως συνιστώσες έχει τις ποσοστιαίες αποκλίσεις από την υγιή λειτουργία των μετρούμενων μεγεθών (διάνυσμα \bar{Y}), όπως περιγράφηκε στο κεφάλαιο 3. Κάθε τέτοιο διάνυσμα $\Delta\bar{Y}$ είναι και ένα σημείο εκπαίδευσης για το διαγνωστικό πρόγραμμα.



Σχήμα 4.2. Διαδικασία προσομοίωσης περιπτώσεων λειτουργίας της μηχανής με χρήση αεροθερμοδυναμικού μοντέλου προσομοίωσης λειτουργίας της μηχανής Engine Performance Model – EPM)

Με αυτόν τον τρόπο παρήχθησαν δύο ομάδες σημείων (SVM-NS και SVM-NF). Η κάθε μία έχει 782 σημεία τα οποία αντιστοιχούν στις 18 προαναφερθείσες κλάσεις. Η πρώτη ομάδα εκ των δύο (SVM-NS), παράχθηκε προσθέτοντας θόρυβο στα θερμοδυναμικά μεγέθη κατά τη διαδικασία της προσομοίωσης, ενώ η δεύτερη

(SVM-NF) παράχθηκε χωρίς θόρυβο. Τα θεωρηθέντα επίπεδα θορύβου αντιστοιχούν στο θόρυβο της μέσης τιμής 20 μετρήσεων από τον κινητήρα. Έτσι, το επίπεδο θορύβου δίνεται από τη σχέση: $\sigma_\alpha = \frac{\sigma}{\sqrt{20}}$, όπου σ το επίπεδο θορύβου λόγω των χρησιμοποιούμενων μετρητικών οργάνων όπως το δίνει ο Πίνακας 4.3.

Ο Πίνακας 4.5 απεικονίζει το σημείο λειτουργίας της μηχανής που επιλέχθηκε γι' αυτή τη διαδικασία.

Μέγεθος	Pamb	T1	P1	WFE
Τιμή	23351,3 Pa	247,2073 K	36306,5 Pa	0,4060 gr/sec

Πίνακας 4.5. Το σημείο λειτουργίας που χρησιμοποιήθηκε για την προσομοίωση.

4.1.1.2 Βελτιστοποίηση των παραμέτρων των SVMs

Για την εκτίμηση της βέλτιστης τιμής των μεταβλητών των συναρτήσεων κελύφους ακολουθούμε τη διαδικασία που περιγράφηκε στο κεφάλαιο 3. Συγκεκριμένα, για τον προσδιορισμό των βέλτιστων τιμών, αρχικά παράγουμε δύο σεντ τιμών (ένα με θόρυβο και ένα χωρίς) των 782 σημείων (optimization points) το καθένα με προσομοίωση της λειτουργίας της μηχανής όπως έγινε και για την πληροφορία εκπαίδευσης. Αυτή τη φορά όμως, έχουμε διαφορά στα εύρη μεταβολής των παραμέτρων υγείας. Η διαφορά αυτή είναι μικρή ώστε τα σημεία που θα προκύψουν να ανήκουν πάλι στις ίδιες κλάσεις αλλά να μην είναι τα ίδια με τα σημεία της πληροφορίας εκπαίδευσης. Ο Πίνακας 4.6 παρουσιάζει τις 18 κλάσεις, τις παραμέτρους που αποκλίνουν καθώς και το εύρος τους, όπως θεωρήθηκαν για την παραγωγή των σημείων βελτιστοποίησης (optimization points).

α/α κλάσης	Παράμετροι που αποκλίνουν σημαντικά
1	$\Delta SW12=\alpha$
2	$\Delta SE12=\alpha$
3	$(\Delta SW12=\alpha, \Delta SE12=\alpha)$ ή $(\Delta SW12=\alpha, \Delta SE12=\alpha/2)$
4	$\Delta SW2=\alpha$
5	$\Delta SE2=\alpha$
6	$(\Delta SW2=\alpha, \Delta SE2=\alpha)$ ή $(\Delta SW2=\alpha, \Delta SE2=\alpha/2)$
7	$(\Delta SW12=\alpha, \Delta SE12=\alpha, \Delta SW2=\alpha, \Delta SE2=\alpha)$ ή $(\Delta SW12=\alpha, \Delta SE12=\alpha/2, \Delta SW2=\alpha, \Delta SE2=\alpha/2)$
8	$\Delta SW26=\alpha, \Delta SW26=-\alpha$
9	$\Delta SE26=\alpha$
10	$(\Delta SW26=\alpha, \Delta SE26=\alpha)$ ή $(\Delta SW26=\alpha, \Delta SE26=\alpha/2)$
11	$\Delta SW41=\alpha, \Delta SW41=-\alpha$
12	$\Delta SE41=\alpha$
13	$(\Delta SW41=\alpha, \Delta SE41=\alpha)$ ή $(\Delta SW41=\alpha, \Delta SE41=\alpha/2)$
14	$\Delta SW49=\alpha, \Delta SW49=-\alpha$
15	$\Delta SE49=\alpha$
16	$(\Delta SW49=\alpha, \Delta SE49=\alpha)$ ή $(\Delta SW49=\alpha, \Delta SE49=\alpha/2)$
17	$\Delta IMP=\alpha$
18	Οποιαδήποτε κατά β

Πίνακας 4.6. Οι περιπτώσεις λειτουργίας (κλάσεις) που θεωρήθηκαν για την παραγωγή σημείων βελτιστοποίησης. ($\alpha \in [-2.45, -0.45]$, $\beta \in (-0.45, +0.45)$ βήμα: 0.1%).

Για τον προσδιορισμό των βέλτιστων τιμών των παραμέτρων των kernels δημιουργήθηκε ένα πρόγραμμα βελτιστοποίησης που θα εκπαιδεύει τη μηχανή διανυσμάτων υποστήριξης με τα training points και θα κατατάσσει στις 18 κλάσεις τα optimization points για διάφορες τιμές των μεταβλητών των kernels. Επειδή γνωρίζουμε σε ποια κλάση ανήκει το κάθε σημείο, μπορούμε να δούμε πόσες λάθος κατατάξεις έχουν γίνει ελέγχοντας για κάθε σημείο την ομοιότητα της κλάσεως κατάταξης με την πραγματική. Τέλος, για κάθε kernel κρατάμε τις τιμές των μεταβλητών του για τις οποίες είχαμε τις λιγότερες λάθος κατατάξεις. Αυτή η διαδικασία γίνεται τόσο για το σετ των SVM-NS όσο και για το σετ των SVM-NF με, αντίστοιχα, το σετ βελτιστοποίησης που παράχθηκε με θόρυβο και με αυτό που παράχθηκε χωρίς θόρυβο. Ο Πίνακας 4.7 απεικονίζει τις βέλτιστες τιμές που προέκυψαν καθώς και το πλήθος των λανθασμένων κατατάξεων, για αυτές τις βέλτιστες τιμές, σε σύνολο 782 σημείων.

Μεταβλητές των kernels	set SVM							
	SVM-NS				SVM-NF			
	SIGMOID	POLYNOMIAL	LINEAR	RBF	SIGMOID	POLYNOMIAL	LINEAR	RBF
gamma	1.0	1.0	-	9.0	1.0	1.0	-	0.2
coef	5.0	1.0	-	-	3.0	1.0	-	-
degree	-	3.0	-	-	-	3.0	-	-
αριθμός λαθών	625	60	279	60	668	63	249	25

Πίνακας 4.7. Βέλτιστες τιμές των παραμέτρων των kernels.

4.1.1.3 Εξέταση Βλαβών

Για την εφαρμογή του διαγνωστικού προγράμματος των SVMs εξετάζονται 15 περιπτώσεις οι οποίες αποτελούν αντιπροσωπευτικές περιπτώσεις βλαβών ενός αεριοστροβίλου (Curnock (2001)). Τις 15 εξεταζόμενες περιπτώσεις βλαβών (a-o), τους λειτουργικούς δείκτες που αποκλίνουν κάθε φορά καθώς και τη μεταβολή τους παρουσιάζει ο Πίνακας 4.8.

Περίπτωση Βλάβης	Αποκλίσεις λειτουργικών δεικτών
a	$\Delta SW2=-0.7\%$, $\Delta SE2=-0.4\%$ $\Delta SW12=-1\%$, $\Delta SE12=-0.5\%$
b	$\Delta SE12=-1\%$
c	$\Delta SW26=-1\%$, $\Delta SE26=-0.7\%$
d	$\Delta SE26=-1\%$
e	$\Delta SW26=-1\%$
f	$\Delta SW41=+1\%$
g	$\Delta SW41=-1\%$, $\Delta SE41=-1\%$
h	$\Delta SE41=-1\%$
i	$\Delta SE49=-1\%$
j	$\Delta SW49=+1\%$, $\Delta SE49=-0.4\%$
k	$\Delta SW49=-1\%$
l	$\Delta SW49=+1\%$, $\Delta SE49=-0.6\%$
m	$A8IMP=+1\%$
n	$A8IMP=-1\%$
o	$A8IMP=+2\%$

Πίνακας 4.8. Εξεταζόμενες περιπτώσεις βλαβών.

Διαθέσιμα για κάθε περίπτωση εξεταζόμενης βλάβης είναι 50 σημεία (διανύσματα $\Delta \bar{Y}$) και έτσι και για τις 15 βλάβες έχουμε 750 σημεία. Τα σημεία αυτά αντικατοπτρίζουν τη λειτουργία της μηχανής με απόκλιση των παραμέτρων όπως

αυτές παρουσιάστηκαν πριν (Πίνακας 4.8) και περιέχουν θόρυβο. Το σύνολο αυτών των 750 σημείων, Single Points, αποτελεί το σετ κατάταξης A (classification set A).

Επίσης, δημιουργούμε και ένα δεύτερο σετ κατάταξης (classification set B) το οποίο περιέχει πάλι 750 σημεία, 50 για κάθε βλάβη. Το κάθε σημείο της κάθε βλάβης αυτού του σετ έχει προέλθει ως ο μέσος όρος των προηγούμενων του σημείων όπως αυτά υπάρχουν στο σετ A. Δηλαδή, το 3^ο σημείο της βλάβης c του σετ B είναι ο μέσος όρος των 3^{ων} πρώτων σημείων του σετ A της βλάβης c. Η διαδικασία αυτή γίνεται ώστε να παραχθεί ένα σετ κατάταξης (Running Average Points) στα σημεία του οποίου ο θόρυβος θα έχει μειωμένη επίδραση (θα έχει 'φιλτραριστεί').

4.1.1.4 Αποτελέσματα set B (Running Average Points) με SVM-NS

Στο πρόγραμμα των SVMs, η εκπαίδευση γίνεται με το set SVM-NS και η κατάταξη με το set B. Δηλαδή, η εκπαίδευση γίνεται με σημεία που περιέχουν θόρυβο και η κατάταξη γίνεται με σημεία στα οποία ο θόρυβος έχει φιλτραριστεί. Η διαδικασία αυτή γίνεται και για τα 4 kernels με τις τιμές των παραμέτρων τους όπως αυτές προέκυψαν κατά τη διαδικασία βελτιστοποίησής τους (Πίνακας 4.7).

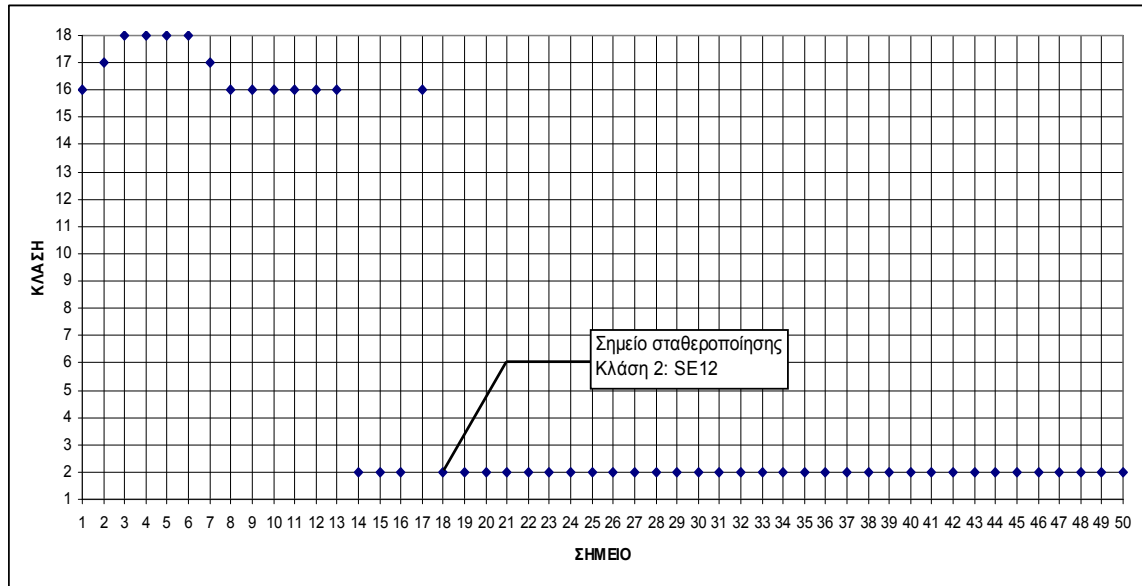
Τα αποτελέσματα παρουσιάζει ο Πίνακας 4.9 στον οποίο έχουμε τις περιπτώσεις βλάβης (a-o), τις παραμέτρους βλάβης που είναι οι παράμετροι που αποκλίνουν σημαντικά για την κάθε βλάβη, τις παραμέτρους κατάταξης που είναι οι παράμετροι που αποκλίνουν σημαντικά για κάθε μία κλάση από τις 18 και στην οποία το πρόγραμμα SVM κατέταξε τα σημεία και τέλος το σημείο από το οποίο και μετά η κατάταξη έχει σταθεροποιηθεί και κάθε μετέπειτα σημείο κατατάσσεται στην ίδια κλάση. Οι σκιασμένες παράμετροι υποδηλώνουν τις λάθος κατατάξεις.

Περίπτωση βλάβης	SVM								
	Παράμετροι βλάβης	SIGMOID		POLYNOMIAL		LINEAR		RBF	
		Παράμετροι κατάταξης	Σημείο	Παράμετροι κατάταξης	Σημείο	Παράμετροι κατάταξης	Σημείο	Παράμετροι κατάταξης	Σημείο
a	SW12, SE12 SW2, SE2	A8IMP	6	SW12, SE12 SW2, SE2	1	SW12, SE12 SW2, SE2	1	SW12, SE12 SW2, SE2	1
b	SE12	no fault	1	SE12	18	SE12	12	SE12	21
c	SW26, SE26	SW49, SE49	13	SE26	21	SW26, SE26	47	SW2, SE2	25
d	SE26	SW49, SE49	1	SE26	12	SW26, SE26	50	SE26	37
e	SW26	no fault	2	SW26	3	SE2	48	SW26	4
f	SW41	SW49, SE49	13	SW41	1	SW2	1	SW41	1
g	SW41, SE41	SW49, SE49	1	SW41, SE41	23	SW41, SE41	2	SW41, SE41	1
h	SE41	SW49, SE49	1	SE41	1	SW41, SE41	10	SE41	21
i	SE49	SW49, SE49	1	SE49	1	SW49, SE49	35	SE49	3
j	SW49, SE49	SW49, SE49	8	SW49, SE49	20	SW41, SE41	10	SE41	40
k	SW49	SW49, SE49	3	SW49	1	SW41, SE41	42	SW49	1
l	SW49, SE49	no fault	4	SW49, SE49	2	SW49, SE49	3	SW49, SE49	2
m	A8IMP	A8IMP	1	A8IMP	2	SW12, SE12	3	A8IMP	7
n	A8IMP	SW49, SE49	3	A8IMP	2	A8IMP	50	A8IMP	3
o	A8IMP	A8IMP	1	A8IMP	1	SW12, SE12	41	A8IMP	3

Πίνακας 4.9. Εντοπιζόμενη βλάβη για το σετ των Running Average Points με εκπαίδευση από το σετ SVM-NS.

Από τα αποτελέσματα συμπεραίνουμε ότι το sigmoid και το linear kernel απέτυχαν να κάνουν σωστή κατάταξη σχεδόν σε όλα τα σημεία, ενώ τα polynomial και RBF είχαν σημαντική επιτυχία. Αυτό συμφωνεί με την απόδοση που είχαν τα τέσσερα kernels στη διαδικασία της εύρεσης των βέλτιστων τιμών των παραμέτρων τους, όπως φαίνεται από τον αριθμό των λανθασμένων κατατάξεων που είχαμε για κάθε kernel (Πίνακας 4.7).

Για να γίνει πιο κατανοητή η έννοια του σημείου εντοπισμού, στο Σχήμα 4.3 φαίνονται τα 50 σημεία που αντιστοιχούν στην βλάβη b και οι κλάσεις στις οποίες κατατάσσονται για την περίπτωση του polynomial kernel. Να σημειωθεί πως η κλάση που αντιστοιχεί στις μεταβολές των παραμέτρων υγείας της βλάβης b είναι η 2.



Σχήμα 4.3. Κατάταξη των 50 σημείων που αντιστοιχούν στη βλάβη b για polynomial kernel.

Ο Πίνακας 4.10 για το RBF και ο Πίνακας 4.11 για το polynomial kernel απεικονίζει πόσα σημεία που πραγματικά ανήκουν στην εκάστοτε βλάβη κατατάχτηκαν λάθος καθώς επίσης και την κλάση κατάταξής τους. Η έντονη σκίαση υποδηλώνει τη σωστή κλάση, που αντιστοιχεί σε σημαντική απόκλιση όλων των παραμέτρων που πραγματικά αποκλίνουν, ενώ η ασθενέστερη αντιστοιχεί σε βλάβη που εντοπίζεται στην ίδια συνιστώσα της μηχανής στην οποία πραγματικά υπάρχει βλάβη.

ΚΛΑΣΗ		ERRORS	ΔSW12	ΔSE12	ΔSW12 ΔSE12	ΔSW2	ΔSE2	ΔSW2 ΔSE2	ΔSW12 ΔSE12 ΔSW2 ΔSE2	ΔSW26	ΔSE26	ΔSW26 ΔSE26	SW41	SE41	ΔSW41 ΔSE41	SW49	SE49	ΔSW49 ΔSE49	A8IMP	no fault	
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
a	SW12, SE12 SW2, SE2	0																			
b	SE12	20													1						19
c	SW26, SE26	49					10	26							13						
d	SE26	14										14									
e	SW26	3													1						2
f	SW41	0																			
g	SW41, SE41	0																			
h	SE41	9													1				8		
i	SE49	1													1						
j	SW49, SE49	46												42	4						
k	SW49	0																			
l	SW49, SE49	1													1						
m	A8IMP	6			4										2						
n	A8IMP	2													2						
o	A8IMP	2			2																

Πίνακας 4.10. Κλάσεις κατάταξης των λάθος κατατασσόμενων σημείων για το set B στην περίπτωση του RBF kernel.

ΚΛΑΣΗ		ERRORS	ΔSW12	ΔSE12	ΔSW12 ΔSE12	ΔSW2	ΔSE2	ΔSW2 ΔSE2	ΔSW12 ΔSE12 ΔSW2 ΔSE2	ΔSW26	ΔSE26	ΔSW26 ΔSE26	SW41	SE41	ΔSW41 ΔSE41	SW49	SE49	ΔSW49 ΔSE49	ABIMP	no fault	
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
a	SW12, SE12 SW2, SE2	0																			
b	SE12	14																	8	2	4
c	SW26, SE26	50					9				30		8				3				
d	SE26	8										8									
e	SW26	2																		2	
f	SW41	0																			
g	SW41, SE41	22									21			1							
h	SE41	0																			
i	SE49	0																			
j	SW49, SE49	16													16						
k	SW49	0																			
l	SW49, SE49	1													1						
m	ABIMP	1			1																
n	ABIMP	1																	1		
o	ABIMP	0																			

Πίνακας 4.11. Κλάσεις κατάταξης των λάθος κατατασόμενων σημείων για το set B στην περίπτωση του Polynomial kernel

Και τα δύο kernels είχαν καλή απόδοση αφού κατέταξαν σωστά σημαντικό αριθμό σημείων. Στην περίπτωση του RBF (Πίνακας 4.10) σημαντικό αριθμό λαθών έχουμε για τα σημεία που αντιστοιχούν στις βλάβες c και j, ενώ το polynomial (Πίνακας 4.11) παρουσίασε εν γένει καλύτερη συμπεριφορά από το RBF με σημαντικό αριθμό λαθών για τα σημεία που αντιστοιχούν στην κλάση c.

4.1.1.5 Αποτελέσματα set A (Single Points) με SVM-NS

Στο πρόγραμμα των SVMs, η εκπαίδευση γίνεται με το set SVM-NS και η κατάταξη με το set A. Δηλαδή, τα σημεία που χρησιμοποιούνται για την εκπαίδευση περιέχουν θόρυβο και στα σημεία για την κατάταξη ο θόρυβος δεν έχει ‘φιλτραριστεί’. Η διαδικασία αυτή γίνεται και για τα 2 kernels (RBF και Polynomial) που μέχρι στιγμής έχουν τη μεγαλύτερη επιτυχία, με τις τιμές των παραμέτρων τους όπως αυτές προέκυψαν κατά τη διαδικασία βελτιστοποίησής τους (Πίνακας 4.7).

Ο Πίνακας 4.12 για το RBF και ο Πίνακας 4.13 για το polynomial kernel απεικονίζουν πόσα σημεία που πραγματικά ανήκουν στην εκάστοτε βλάβη, κατατάχτηκαν λάθος καθώς επίσης και την κλάση κατάταξής τους. Η έντονη σκίαση υποδηλώνει τη σωστή κλάση, που αντιστοιχεί σε σημαντική απόκλιση όλων των

παραμέτρων που πραγματικά αποκλίνουν, ενώ η ασθενέστερη αντιστοιχεί σε βλάβη που εντοπίζεται στην ίδια συνιστώσα της μηχανής στην οποία πραγματικά υπάρχει βλάβη.

		ERRORS	Δ SW12	Δ SE12	Δ SW12 Δ SE12	Δ SW2	Δ SE2	Δ SW2 Δ SE2	Δ SW12 Δ SE12 Δ SW2 Δ SE2	Δ SW26	Δ SE26	Δ SW26 Δ SE26	SW41	SE41	Δ SW41 Δ SE41	SW49	SE49	Δ SW49 Δ SE49	A8IMP	no fault
ΚΛΑΣΗ			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
a	SW12, SE12 SW2, SE2	32	4		1										26					1
b	SE12	35						1				2			28			1		3
c	SW26, SE26	36					12	4							17			1		
d	SE26	49										30			19					
e	SW26	43		2		2						4			29	4		1		1
f	SW41	25						1							24					
g	SW41, SE41	0																		
h	SE41	45													30				15	
i	SE49	34										3			30				1	
j	SW49, SE49	37						2						6	26	3				
k	SW49	35				1				1			7		24				2	
l	SW49, SE49	30													30					
m	A8IMP	44	2		12				3						27					
n	A8IMP	33										1			31				1	
o	A8IMP	32			8										24					

Πίνακας 4.12. Κλάσεις κατάταξης των λάθος κατατασσόμενων σημείων για το set A στην περίπτωση του RBF kernel.

		ERRORS	Δ SW12	Δ SE12	Δ SW12 Δ SE12	Δ SW2	Δ SE2	Δ SW2 Δ SE2	Δ SW12 Δ SE12 Δ SW2 Δ SE2	Δ SW26	Δ SE26	Δ SW26 Δ SE26	SW41	SE41	Δ SW41 Δ SE41	SW49	SE49	Δ SW49 Δ SE49	A8IMP	no fault
ΚΛΑΣΗ			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
a	SW12, SE12 SW2, SE2	22	9	4	9															
b	SE12	15			1		2		1									6	1	4
c	SW26, SE26	45					22	4				8	1	2			6	1		1
d	SE26	10										4		6						
e	SW26	36		5		6	2	1	1			2				5		3	1	10
f	SW41	5				2		1							2					
g	SW41, SE41	16									11			3		2				
h	SE41	20												9					11	
i	SE49	17									4	1			3				9	
j	SW49, SE49	34						1						23		7	1			
k	SW49	18				3				1	5		6		1				2	
l	SW49, SE49	25		7			1					1	8	2				5		1
m	A8IMP	39	6	1	26				5				1							
n	A8IMP	27					1	3				6			2			1	12	
o	A8IMP	22	9		13															

Πίνακας 4.13. Κλάσεις κατάταξης των λάθος κατατασσόμενων σημείων για το set A στην περίπτωση του Polynomial kernel.

Από τη σύγκριση των παραπάνω αποτελεσμάτων (Πίνακας 4.10 - Πίνακας 4.13) βλέπουμε πως στην περίπτωση που τα σημεία κατάταξης περιέχουν θόρυβο (set A), τότε αυξάνονται και τα λάθη κατά την κατάταξη, ενώ όταν έχουμε ‘φιλτραρισμένο’ το θόρυβο στα σημεία κατάταξης (set B), τότε τα λάθη είναι λίγα. Έτσι, συμπεραίνουμε ότι η ύπαρξη θορύβου στα δεδομένα κατάταξης επηρεάζει

αρνητικά τα αποτελέσματα της διαδικασίας κατάταξης μειώνοντας την απόδοση του διαγνωστικού προγράμματος.

4.1.1.6 Αποτελέσματα set B (Running Average Points) με SVM-NF

Στην περίπτωση αυτή, η εκπαίδευση γίνεται με το set SVM-NF και η κατάταξη με το set B, δηλαδή χωρίς θόρυβο στα δεδομένα εκπαίδευσης και με φιλτραρισμένα τα δεδομένα κατάταξης. Η διαδικασία αυτή γίνεται για τα 2 kernels (RBF και Polynomial) με τις τιμές των παραμέτρων τους όπως αυτές προέκυψαν κατά τη διαδικασία βελτιστοποίησής τους (Πίνακας 4.7).

Τα αποτελέσματα απεικονίζει ο Πίνακας 4.14 στον οποίο έχουμε τις περιπτώσεις βλάβης (a-o), τις παραμέτρους βλάβης που είναι οι παράμετροι που αποκλίνουν σημαντικά για την κάθε βλάβη, τις παραμέτρους κατάταξης που είναι οι παράμετροι που αποκλίνουν σημαντικά για κάθε μία κλάση από τις 18 και στην οποία το πρόγραμμα SVM κατέταξε τα σημεία και τέλος το σημείο από το οποίο και μετά η κατάταξη έχει σταθεροποιηθεί και κάθε μετέπειτα σημείο κατατάσσεται στην ίδια κλάση. Οι σκιασμένες παράμετροι υποδηλώνουν τις λάθος κατατάξεις. Η λέξη “final” στην κλάση *h* στο polynomial kernel υποδηλώνει πως τα προηγούμενα σημεία κατατάσσονταν σωστά, αλλά τελικά η κατάταξη σταθεροποιήθηκε σε λανθασμένη κλάση.

Περίπτωση βλάβης	SVM				
	Παράμετροι βλάβης	POLYNOMIAL		RBF	
		Παράμετροι κατάταξης	Σημείο	Παράμετροι κατάταξης	Σημείο
a	SW12, SE12 SW2, SE2	SW12, SE12 SW2, SE2	1	SW12, SE12 SW2, SE2	1
b	SE12	SE12	11	SE12	6
c	SW26, SE26	SW2, SE2	16	SW2, SE2	15
d	SE26	SE26	16	SE26	2
e	SW26	SW26	3	SW26	4
f	SW41	SW41	1	SW41	1
g	SW41, SE41	SW41, SE41	12	SW41, SE41	18
h	SE41	SW49, SE49	42 (final)	SE41	1
i	SE49	SE49	1	SE49	1
j	SW49, SE49	SE41	44	SE41	44
k	SW49	SW49	1	SW49	1
l	SW49, SE49	SW49, SE49	4	SW49, SE49	4
m	A8IMP	A8IMP	7	A8IMP	7
n	A8IMP	A8IMP	8	A8IMP	31
o	A8IMP	A8IMP	24	A8IMP	17

Πίνακας 4.14. Εντοπιζόμενη βλάβη για το σετ των Running Average Points με εκπαίδευση από το σετ SVM-NF

Συγκρίνοντας τα αποτελέσματα της κατάταξης των ‘φιλτραρισμένων’ σημείων στις περιπτώσεις που η εκπαίδευση έγινε με σημεία με θόρυβο και χωρίς θόρυβο, (Πίνακας 4.9 και Πίνακας 4.14) βλέπουμε πως το kernel RBF είχε παραπλήσια συμπεριφορά ενώ έχουμε μικρή χειροτέρευση στα αποτελέσματα του polynomial. Αυτό δείχνει πως είναι δυνατό να γίνει η εκπαίδευση με δεδομένα χωρίς θόρυβο και τα αποτελέσματα που έχουμε κατά την διαδικασία της κατάταξης να παραμένουν σωστά. Το συμπέρασμα αυτό μπορεί να φανεί πολύ χρήσιμο σε περιπτώσεις που δεν γνωρίζουμε τη στάθμη του θορύβου των μεγεθών, αφού μπορούμε να κάνουμε την εκπαίδευση με δεδομένα χωρίς θόρυβο και να έχουμε σωστά αποτελέσματα κατάταξης.

4.1.1.7 Εντοπισμός Μεγέθους Βλάβης

Μέχρι στιγμής το διαγνωστικό πρόγραμμα χρησιμοποιήθηκε για να εντοπίσει τις παραμέτρους υγείας που αποκλίνουν σημαντικά σε κάθε περίπτωση. Σημαντικό

όμως είναι να γνωρίζουμε και το μέγεθος αυτής της απόκλισης των παραμέτρων υγείας. Για να το πετύχουμε αυτό, θεωρούμε 133 κλάσεις (αντί των 18 που είχαμε μέχρι τώρα) που η κάθε μία αντιπροσωπεύει συγκεκριμένο εύρος αποκλίσεων των λειτουργικών δεικτών όπως απεικονίζει ο Πίνακας 4.15.

Αρ. Κλάσεων	Αποκλίσεις παραμέτρων
11	$\Delta f_i = \alpha, \alpha \in [-2.5\%, -2.0\%], f_i = A8IMP, SW_i, SE_i, i = 12, 2, 26, 41, 49$
11	$\Delta f_i = \alpha, \alpha \in (-2.0\%, -1.5\%], f_i = A8IMP, SW_i, SE_i, i = 12, 2, 26, 41, 49$
11	$\Delta f_i = \alpha, \alpha \in (-1.5\%, -1.0\%], f_i = A8IMP, SW_i, SE_i, i = 12, 2, 26, 41, 49$
11	$\Delta f_i = \alpha, \alpha \in (-1.0\%, -0.5\%], f_i = A8IMP, SW_i, SE_i, i = 12, 2, 26, 41, 49$
4	$\Delta f_i = \alpha, \alpha \in [0.5\%, 1.0\%], f_i = A8IMP, SW_i, i = 26, 41, 49$
4	$\Delta f_i = \alpha, \alpha \in (1.0\%, 1.5\%], f_i = A8IMP, SW_i, i = 26, 41, 49$
4	$\Delta f_i = \alpha, \alpha \in (1.5\%, 2.0\%], f_i = A8IMP, SW_i, i = 26, 41, 49$
4	$\Delta f_i = \alpha, \alpha \in (2.0\%, 2.5\%], f_i = A8IMP, SW_i, i = 26, 41, 49$
5	$(\Delta SW_i = \alpha, \Delta SE_i = \alpha), \alpha \in [-2.5\%, -2.0\%], i = 12, 2, 26, 41, 49$
5	$(\Delta SW_i = \alpha, \Delta SE_i = \alpha), \alpha \in (-2.0\%, -1.5\%], i = 12, 2, 26, 41, 49$
5	$(\Delta SW_i = \alpha, \Delta SE_i = \alpha), \alpha \in (-1.5\%, -1.0\%], i = 12, 2, 26, 41, 49$
5	$(\Delta SW_i = \alpha, \Delta SE_i = \alpha), \alpha \in (-1.0\%, -0.5\%], i = 12, 2, 26, 41, 49$
5	$(\Delta SW_i = \alpha, \Delta SE_i = \alpha/2), \alpha \in [-2.5\%, -2.0\%], i = 12, 2, 26, 41, 49$
5	$(\Delta SW_i = \alpha, \Delta SE_i = \alpha/2), \alpha \in (-2.0\%, -1.5\%], i = 12, 2, 26, 41, 49$
5	$(\Delta SW_i = \alpha, \Delta SE_i = \alpha/2), \alpha \in (-1.5\%, -1.0\%], i = 12, 2, 26, 41, 49$
5	$(\Delta SW_i = \alpha, \Delta SE_i = \alpha/2), \alpha \in (-1.0\%, -0.5\%], i = 12, 2, 26, 41, 49$
3	$(\Delta SW_i = \alpha, \Delta SE_i = -\alpha), \alpha \in [0.5\%, 1.0\%], i = 26, 41, 49$
3	$(\Delta SW_i = \alpha, \Delta SE_i = -\alpha), \alpha \in (1.0\%, 1.5\%], i = 26, 41, 49$
3	$(\Delta SW_i = \alpha, \Delta SE_i = -\alpha), \alpha \in (1.5\%, 2.0\%], i = 26, 41, 49$
3	$(\Delta SW_i = \alpha, \Delta SE_i = -\alpha), \alpha \in (2.0\%, 2.5\%], i = 26, 41, 49$
3	$(\Delta SW_i = \alpha, \Delta SE_i = -\alpha/2), \alpha \in [0.5\%, 1.0\%], i = 26, 41, 49$
3	$(\Delta SW_i = \alpha, \Delta SE_i = -\alpha/2), \alpha \in (1.0\%, 1.5\%], i = 26, 41, 49$
3	$(\Delta SW_i = \alpha, \Delta SE_i = -\alpha/2), \alpha \in (1.5\%, 2.0\%], i = 26, 41, 49$
3	$(\Delta SW_i = \alpha, \Delta SE_i = -\alpha/2), \alpha \in (2.0\%, 2.5\%], i = 26, 41, 49$
1	$(\Delta SW_{12} = \alpha, \Delta SE_{12} = \alpha, \Delta SW_2 = \alpha, \Delta SE_2 = \alpha), \alpha \in [-2.5\%, -2.0\%]$
1	$(\Delta SW_{12} = \alpha, \Delta SE_{12} = \alpha, \Delta SW_2 = \alpha, \Delta SE_2 = \alpha), \alpha \in (-2.0\%, -1.5\%]$
1	$(\Delta SW_{12} = \alpha, \Delta SE_{12} = \alpha, \Delta SW_2 = \alpha, \Delta SE_2 = \alpha), \alpha \in (-1.5\%, -1.0\%]$
1	$(\Delta SW_{12} = \alpha, \Delta SE_{12} = \alpha, \Delta SW_2 = \alpha, \Delta SE_2 = \alpha), \alpha \in (-1.0\%, -0.5\%]$
1	$(\Delta SW_{12} = \alpha, \Delta SE_{12} = \alpha/2, \Delta SW_2 = \alpha, \Delta SE_2 = \alpha/2), \alpha \in [-2.5\%, -2.0\%]$
1	$(\Delta SW_{12} = \alpha, \Delta SE_{12} = \alpha/2, \Delta SW_2 = \alpha, \Delta SE_2 = \alpha/2), \alpha \in (-2.0\%, -1.5\%]$
1	$(\Delta SW_{12} = \alpha, \Delta SE_{12} = \alpha/2, \Delta SW_2 = \alpha, \Delta SE_2 = \alpha/2), \alpha \in (-1.5\%, -1.0\%]$
1	$(\Delta SW_{12} = \alpha, \Delta SE_{12} = \alpha/2, \Delta SW_2 = \alpha, \Delta SE_2 = \alpha/2), \alpha \in (-1.0\%, -0.5\%]$
1	Οποιαδήποτε κατά β ($\beta \in (-0.5\%, +0.5\%)$)
133	Σύνολο κλάσεων

Πίνακας 4.15. Περιοχές αποκλίσεων των λειτουργικών δεικτών των κλάσεων που θεωρήθηκαν για τον εντοπισμό του μεγέθους της βλάβης (βήμα 0.1%).

Η εκπαίδευση έγινε πάλι με 782 σημεία με ύπαρξη θορύβου (set SVM-NS), όπως περιγράφηκε και στην περίπτωση των 18 κλάσεων (παράγραφος 4.1.1.4), ενώ η βελτιστοποίηση με το σετ των optimization points με θόρυβο. Η διαδικασία της κατάταξης έγινε με το set B (“φιλτραρισμένα” σημεία κατάταξης) και

χρησιμοποιήθηκαν τα δύο kernels (RBF και Polynomial) με την καλύτερη συμπεριφορά ως τώρα, αλλά με τη χρήση του polynomial δεν κατέστη εφικτή η εκπαίδευση. Ο Πίνακας 4.16 παρουσιάζει τις βέλτιστες τιμές των παραμέτρων των kernels. Στην περίπτωση του πολυωνυμικού, αναγράφεται η φράση ‘No SVs’ επειδή δεν κατέστη δυνατή η εκπαίδευση διότι δεν μπόρεσαν να προσδιοριστούν τα διανύσματα υποστήριξης (Support Vectors).

Μεταβλητές των kernels	set SVM			
	SVM-NS			
	SIGMOID	POLYNOMIAL	LINEAR	RBF
gamma				1.0
coef		No SVs		-
degree				-

Πίνακας 4.16. Βέλτιστες τιμές των παραμέτρων των kernels για τις 133 κλάσεις.

Ο Πίνακας 4.17 απεικονίζει τα πραγματικά και τα εκτιμώμενα μεγέθη απόκλισης των παραμέτρων καθώς και τα σημεία που επιτυγχάνεται αυτή η εκτίμηση. Οι εκτιμώμενες περιοχές αποκλίσεων των παραμέτρων απεικονίζονται στον πίνακα με τη μέση τιμή τους. Οι σκιασμένες παράμετροι υποδηλώνουν τις λάθος κατατάξεις. Στις βλάβες e , l , m υπάρχει η επισήμανση ‘other points...’. Αυτό σημαίνει ότι η κατάταξη, να μεν σταθεροποιήθηκε μετά το 33^ο σημείο (e και l), αλλά τα προηγούμενα σημεία κατατάσσονταν όπως υποδηλώνει η έκφραση ‘other points...’. Οι εκτιμώμενες αποκλίσεις των παραμέτρων κατάταξης αυτών των σημείων δεν διαφέρουν σχεδόν καθόλου από τις πραγματικές αποκλίσεις και έτσι μπορούμε να πούμε πως και για τα σημεία πριν από τη σταθεροποίηση έχουμε σωστή κατάταξη.

Περίπτωση βλάβης	SVM (RBF)		
	Πραγματικές Αποκλίσεις Παραμέτρων	Εκτιμώμενες Αποκλίσεις Παραμέτρων	Σημείο
a	SW2=-0.7%, SE2=-0.4% SW12=-1%, SE12=-0.5%	SW2=-0.7%, SE2=-0.4% SW12=-0.7%, SE12=-0.4%	6
b	SE12=-1%	SE12=-1.2%	11
c	SW26=-1%, SE26=-0.7%	SE2=-2.25%	47
d	SE26=-1%	SE26=-1.2%	11
e	SW26=-1%	SW26=-1.2% (other points SW26=-0.7%)	33
f	SW41=+1%	SW41=+0.75%	9
g	SW41=-1%, SE41=-1%	SW41=-1.2%, SE41=-1.2%	2
h	SE41=-1%	SE41=-1.2%	8
i	SE49=-1%	SE49=-1.2%	1
j	SW49=-1%, SE49=-0.4%	SW49=-1.2%, SE49=-0.6%	41
k	SW49=-1%	SW49=-1.2%	1
l	SW49=+1%, SE49=-0.6%	SW49=+0.75%, SE49=-0.4% (other points SW49=+1.3%, SE49=-0.7%)	33
m	A8IMP=+1%	A8IMP=+1.3% (other points A8IMP=0.75%)	46
n	A8IMP=-1%	A8IMP=-1.2%	5
o	A8IMP=+2%	A8IMP=+1.8%	11

Πίνακας 4.17. Εντοπιζόμενο μέγεθος βλάβης για το σετ των Running Average Points με εκπαίδευση από το σετ SVM-NS.

Από τα παραπάνω βλέπουμε πως είχαμε όχι μόνο σωστό εντοπισμό των παραμέτρων υγείας που αποκλίνουν αλλά και εντοπισμό του εύρους τις απόκλισης.

4.1.1.8 Σύγκριση με άλλες μεθόδους

Σε αυτό το κεφάλαιο είδαμε τα αποτελέσματα της διάγνωσης με τη χρήση του προγράμματος των SVMs. Κρίνεται σκόπιμο αυτά τα αποτελέσματα να συγκριθούν με αυτά άλλων μεθόδων ώστε να βγει κάποιο συμπέρασμα για την αποτελεσματικότητα των SVMs.

Για τη σύγκριση χρησιμοποιούμε τα αποτελέσματα δύο διαγνωστικών μεθόδων, των Πιθανολογικών Νευρωνικών Δικτύων (Probabilistic Neural Networks, PNN) και των Δικτύων Πιθανοτήτων (Bayesian Belief Networks, BBN), στην

περίπτωση διάγνωσης βλαβών που περιγράφηκε στην παράγραφο 4.1.3 (Ρωμέσης (2005)).

Τα δεδομένα αυτά είναι τα αποτελέσματα που παρήχθησαν κατά τον εντοπισμό τόσο των παραμέτρων που αποκλίνουν σημαντικά κάθε φορά, όσο και του μεγέθους αυτής της μεταβολής.

Περίπτωση βλάβης	SVM (RBF)		
	Πραγματικές Αποκλίσεις Παραμέτρων	Εκτιμώμενες Αποκλίσεις Παραμέτρων	Σημείο
a	SW2=-0.7%, SE2=-0.4% SW12=-1%, SE12=-0.5%	SW2=-0.7%, SE2=-0.4% SW12=-0.7%, SE12=-0.4%	6
b	SE12=-1%	SE12=-1.2%	11
c	SW26=-1%, SE26=-0.7%	SE2=-2.25%	47
d	SE26=-1%	SE26=-1.2%	11
e	SW26=-1%	SW26=-1.2% (other points SW26=-0.7%)	33
f	SW41=+1%	SW41=+0.75%	9
g	SW41=-1%, SE41=-1%	SW41=-1.2%, SE41=-1.2%	2
h	SE41=-1%	SE41=-1.2%	8
i	SE49=-1%	SE49=-1.2%	1
j	SW49=-1%, SE49=-0.4%	SW49=-1.2%, SE49=-0.6%	41
k	SW49=-1%	SW49=-1.2%	1
l	SW49=+1%, SE49=-0.6%	SW49=+0.75%, SE49=-0.4% (other points SW49=+1.3%, SE49=-0.7%)	33
m	A8IMP=+1%	A8IMP=+1.3% (other points A8IMP=0.75%)	46
n	A8IMP=-1%	A8IMP=-1.2%	5
o	A8IMP=+2%	A8IMP=+1.8%	11

Πίνακας 4.18. Εντοπιζόμενο μέγεθος απόκλισης των παραμέτρων στην περίπτωση των SVMs.

Περίπτωση Βλάβης	Πραγματικές Αποκλίσεις Παραμέτρων	Κλάση Διάγνωσης	Εκτιμώμενες Αποκλίσεις Παραμέτρων $\pm 0.167\%$	Σημείο Εντοπισμού μεγέθους Απόκλισης
a	SW2=-0.7%, SE2=-0.4%, SW12=-1%, SE12=-0.5%	sD	SW12=-0.33%, SW2=-1.33%	48
b	SE12=-1%	fD	SE12=-0.67%	43
c	SW26=-1%, SE26=-0.7%	sD	SE2=-1.667%	41
d	SE26=-1%	fD	SE26=-1%	7
e	SW26=-1%	fD	SW26=-1%	28
f	SW41=+1%	fD	SW41=+1%	12
g	SW41=-1%, SE41=-1%	fD	SW41=-1.33%, SE41=-1%	44
h	SE41=-1%	fD	SE41=-1%	1
l	SE49=-1%	fD	SE49=1.33%	30
j	SW49=-1%, SE49=-0.4%	sD	SE41=-0.667%	42
k	SW49=-1%	fD	SW49=-1%	35
l	SW49=+1%, SE49=-0.6%	cD	SW49=+0.66%, SE49=0%	46
m	A8IMP=+1%	fD	A8IMP=+1%	48
n	A8IMP=-1%	fD	A8IMP=-1%	30
o	A8IMP=+2%	fD	A8IMP=+2%	37

Πίνακας 4.19. Εντοπιζόμενο μέγεθος βλάβης με χρήση BBN.

Περίπτωση Βλάβης	Πραγματικές Αποκλίσεις Παραμέτρων	Κλάση Διάγνωσης	Εκτιμώμενες Αποκλίσεις Παραμέτρων $\pm 0.25\%$	Σημείο Εντοπισμού μεγέθους Απόκλισης
a	SW2=-0.7%, SE2=-0.4%, SW12=-1%, SE12=-0.5%	cD	$\Delta SW12=-0.75\%$, $\Delta SE12=-0.375\%$, $\Delta SW2=-0.75\%$, $\Delta SE2=-0.375\%$	6
b	SE12=-1%	fD	$\Delta SE12=-1.25\%$	40
c	SW26=-1%, SE26=-0.7%	fD	$\Delta SW26=-1.25\%$, $\Delta SE26=-1.25\%$	49
d	SE26=-1%	fD	$\Delta SE26=-1.25\%$	2
e	SW26=-1%	fD	$\Delta SW26=-1.25\%$	11
f	SW41=+1%	fD	$\Delta SW41=+0.75\%$	8
g	SW41=-1%, SE41=-1%	fD	$\Delta SW41=-1.25\%$, $\Delta SE41=-1.25\%$	1
h	SE41=-1%	fD	$\Delta SE41=-1.25\%$	15
l	SE49=-1%	fD	$\Delta SE49=-1.25\%$	1
j	SW49=-1%, SE49=-0.4%	cD	$\Delta SW49=-1.25\%$, $\Delta SE49=-0.625\%$	39
k	SW49=-1%	fD	$\Delta SW49=-1.25\%$	1
l	SW49=+1%, SE49=-0.6%	fD	$\Delta SW49=+0.75\%$, $\Delta SE49=-0.375\%$	28
m	A8IMP=+1%	fD	$\Delta A8IMP=+0.75\%$	29
n	A8IMP=-1%	fD	$\Delta A8IMP=-1.25\%$	25
o	A8IMP=+2%	fD	$\Delta A8IMP=+1.75\%$	3

Πίνακας 4.20. Εντοπιζόμενο μέγεθος βλάβης με χρήση PNN.

Από τη σύγκριση των παραπάνω αποτελεσμάτων βλέπουμε πως τα SVMs (Πίνακας 4.18) έχουν μεγαλύτερη επιτυχία από τα BBN (Πίνακας 4.19) τόσο στον εντοπισμό των παραμέτρων που πραγματικά αποκλίνουν κάθε φορά, όσο και στον εντοπισμό του μεγέθους αυτής της απόκλισης. Τα PNN (Πίνακας 4.20) παρουσιάζουν, σε κάποιες περιπτώσεις, καλύτερη συμπεριφορά από τα SVMs τόσο στον εντοπισμό των παραμέτρων που αποκλίνουν, όσο και στον προσδιορισμό του μεγέθους της απόκλισης αυτής ενώ σε άλλες υστερούν. Εδώ πρέπει να σημειωθεί πως τα αποτελέσματα των PNN εξήχθησαν με χρήση 7820 σημείων εκπαίδευσης σε αντίθεση με αυτά των SVMs που χρησιμοποιήθηκαν 782 σημεία για την εκπαίδευση (διαχωρισμός κλάσεων).

4.1.2 Διάγνωση βλαβών αισθητήρων

Στις προηγούμενες παραγράφους εξετάσαμε τη διάγνωση βλαβών των συνιστωσών του κινητήρα. Κρίνεται σκόπιμο τώρα να εξεταστεί και η διάγνωση βλαβών των αισθητήρων που χρησιμοποιούνται ώστε να ληφθεί η απαιτούμενη πληροφορία (διανύσματα \bar{Y}) για την διάγνωση βλαβών της μηχανής. Πριν προχωρήσουμε στην εκτίμηση της κατάστασης των συνιστωσών του αεριοστροβίλου, θα πρέπει να γνωρίζουμε κατά πόσο η πληροφορία που λαμβάνουμε από τη μηχανή με τη χρήση των αισθητήρων είναι αληθής.

Οι περιπτώσεις που εξετάστηκαν είναι 11. Οι 10 από αυτές αντιστοιχούν σε βλάβη κάποιου αισθητήρα, ενώ η 11^η αντιστοιχεί στην υγιή λειτουργία όλων των αισθητήρων. Τις περιπτώσεις αυτές απεικονίζει ο Πίνακας 4.21.

Κλάση	Αισθητήρας με βλάβη
1	T1
2	P1
3	WFE
4	XNLP
5	XNHP
6	P13
7	P3
8	T3
9	T6
10	T13
11	No Fault

Πίνακας 4.21. Οι περιπτώσεις που εξετάστηκαν για τον εντοπισμό βλάβης σε αισθητήρα.

Σε κάθε εξεταζόμενη κλάση που αντιστοιχεί σε βλάβη αισθητήρα, θεωρούμε αποκλίσεις του συγκεκριμένου αισθητήρα κατά $\{-2\%, -1\%, +1\%, +2\%\}$. Επιπλέον, σε όλες τις εξεταζόμενες κλάσεις θεωρούμε και ταυτόχρονη απόκλιση των παραμέτρων υγείας, μιας παραμέτρου κάθε φορά. Η απόκλιση αυτή παίρνει τιμές -2.5, -1.5 και 0. Δηλαδή, πραγματοποιούμε διάγνωση βλαβών αισθητήρων με ταυτόχρονη παρουσία βλάβης στις συνιστώσες της μηχανής.

Με τη χρήση του μοντέλου προσομοίωσης της μηχανής, για τα παραπάνω στοιχεία, λαμβάνουμε τα διανύσματα $\Delta\bar{Y}$ της κάθε κλάσης που αποτελούν και τα σημεία εκπαίδευσης. Δηλαδή, ακολουθούμε την διαδικασία που περιγράφηκε στο κεφάλαιο 3 και εφαρμόστηκε ήδη στην περίπτωση διάγνωσης βλαβών στις συνιστώσες της μηχανής.

Για την δημιουργία των δεδομένων για την εύρεση των βέλτιστων τιμών των παραμέτρων των kernels ακολουθούμε την ίδια διαδικασία μόνο που σε αυτή τη περίπτωση η απόκλιση του κάθε αισθητήρα παίρνει τις τιμές $\{-1.75\%, -1.25\%, +1.25\%, +1.75\%\}$. Ο Πίνακας 4.22 παρουσιάζει τις βέλτιστες τιμές των παραμέτρων καθώς και τις λάθος κατατάξεις για κάθε kernel στην περίπτωση της διαδικασίας βελτιστοποίησης.

Όμοια, δημιουργούμε τα δεδομένα της κατάταξης με απόκλιση των αισθητήρων κατά $\pm 1.5\%$.

Τόσο για την διαδικασία της εκπαίδευσης όσο και για την διαδικασία της βελτιστοποίησης, χρησιμοποιήθηκαν 943 σημεία που αντιστοιχούν στις 11 προαναφερθείσες κλάσεις. Για την διαδικασία της κατάταξης χρησιμοποιήθηκαν 460 σημεία που αντιστοιχούν στις 10 πρώτες κλάσεις.

Μεταβλητές των kernels	SVM			
	LINEAR	POLYNOMIAL	SIGMOID	RBF
gamma	-	1.0	10.0	1.0
coef	-	1.0	1.0	-
degree	-	3.0	-	-
Αριθμός λαθών	900	40	903	0

Πίνακας 4.22. Βέλτιστες τιμές των παραμέτρων των kernels για την περίπτωση διάγνωσης βλαβών αισθητήρων.

Ο Πίνακας 4.23 παρουσιάζει τον αριθμό των λαθών και το ποσοστό επιτυχίας για κάθε kernel κατά τη διαδικασία της κατάταξης.

SVM	Αριθμός λαθών	Ποσοστό επιτυχίας
LINEAR	460	0.0%
POLYNOMIAL	43	93.3%
SIGMOID	450	1.6%
RBF	1	99.8%

Πίνακας 4.23. Αριθμός λαθών και ποσοστό επιτυχίας για τα 4 kernel κατά τη διαδικασία της κατάταξης.

Όπως ήταν αναμενόμενο και από τα αποτελέσματα της βελτιστοποίησης (Πίνακας 4.22) το RBF πέτυχε την κατάταξη όλων των εξεταζόμενων περιπτώσεων στις σωστές κλάσεις. Το πολυωνυμικό kernel είχε και αυτό αξιόλογη επιτυχία. Ο Πίνακας 4.24 απεικονίζει τις κλάσεις που το kernel αυτό κατέταξε τις εξεταζόμενες περιπτώσεις κατά τη διαδικασία της κατάταξης. Η σκίαση υποδηλώνει τη σωστή κατάταξη.

		T1	P1	WFE	XNLP	XNHP	P13	P3	T3	T6	T13	No Fault
ΚΛΑΣΗ		1	2	3	4	5	6	7	8	9	10	11
1	T1	43	2						1			
2	P1		46									
3	WFE	1		36	1	1				6		1
4	XNLP		1	2	42			1				
5	XNHP		1	1		43	1					
6	P13		2				44					
7	P3			3				42			1	
8	T3								46			
9	T6		2	10	1					33		
10	T13	3	1								42	

Πίνακας 4.24. Κλάσεις κατάταξης των εξεταζόμενων περιπτώσεων για την περίπτωση του polynomial kernel.

4.2 Εντοπισμός βλαβών συμπίεστών με στοιχεία θερμοδυναμικών – μηχανικών μεγεθών

Η διαδικασία της διάγνωσης, όπως αυτή περιγράφηκε στο κεφάλαιο 3, εφαρμόστηκε για βλάβες λειτουργούντων αεριοστροβίλων με στοιχεία θερμοδυναμικών και μηχανικών μεγεθών. Σε αυτή τη περίπτωση δεν υπάρχει κάποιο μοντέλο προσομοίωσης της μηχανής που να μας δίνει δεδομένα για τις αποκλίσεις των μηχανικών μεγεθών γι αυτό τα δεδομένα που χρησιμοποιήθηκαν είναι από πειραματικές μετρήσεις για μηχανή με αξονικό συμπίεστή και για μηχανή με ακτινικό συμπίεστή. Για την παραγωγή των δεδομένων χρησιμοποιήθηκαν επιταχυνσιόμετρα σε διάφορες θέσεις της μηχανής καθώς και ένας μορφοτροπέας πίεσης, συσκευές δηλαδή που επιτρέπουν τη λήψη δεδομένων ταχείας απόκρισης. Τα σήματα που ελήφθησαν με αυτή τη διαδικασία αναλύθηκαν με μετασχηματισμό Fourier στις συνιστώσες τους, και αφαιρέθηκαν από τις κατά Fourier μετασχηματισμένες συνιστώσες των σημάτων της υγιούς λειτουργίας (Λουκής (1993), Αρετάκης (2000)).

4.2.1 Αξονικός συμπίεστής

Στην περίπτωση του αξονικού συμπίεστή, μελετήθηκαν 4 βλάβες (F2, F3, F4, F53) με δεδομένα 7 θερμοδυναμικά μεγέθη και μετρήσεις από 3 επιταχυνσιόμετρα και 1 μορφοτροπέα πίεσης (pressure transducer).

Για την κάθε βλάβη είχαμε 4 σημεία με μετρήσεις θερμοδυναμικών μεγεθών, 4 σημεία από κάθε επιταχυνσιόμετρο και 4 από το μορφοτροπέα πίεσης. Τα θερμοδυναμικά μεγέθη που χρησιμοποιήθηκαν είναι: στατική πίεση στην έξοδο του διαχύτη του συμπίεστή ($Ps4$), θερμοκρασία στην έξοδο του διαχύτη του συμπίεστή ($T4$), παροχή εισόδου (WI), στατική πίεση στην είσοδο του στροβίλου ισχύος ($Ps9$), θερμοκρασία στην είσοδο του στροβίλου ισχύος ($T9$), θερμοκρασία στην έξοδο του στροβίλου ισχύος ($T12$) και παροχή καυσίμου (Wf).

Στη μελέτη μας, θεωρούμε τον αριθμό των κλάσεων ίδιο με των αριθμό των εξεταζόμενων βλαβών. Έτσι έχουμε 4 κλάσεις.

Το διαγνωστικό πρόγραμμα εφαρμόζεται ξεχωριστά για τα θερμοδυναμικά δεδομένα, για τα δεδομένα του κάθε επιταχυνσιόμετρου καθώς και γι' αυτά του μορφοτροπέα πίεσης.

Από τα 4 σημεία που έχουμε για κάθε βλάβη-κλάση, τα 3 χρησιμοποιήθηκαν για την διαδικασία της εκπαίδευσης. Επίσης, με τα ίδια αυτά 3 σημεία έγινε η διαδικασία της βελτιστοποίησης κατά την οποία βρέθηκαν οι βέλτιστες τιμές των παραμέτρων των kernels (Πίνακας 4.25). Κατά την διαδικασία της εκπαίδευσης δεν επετεύχθη διαχωρισμός των κλάσεων με τη χρήση του polynomial kernel στα δεδομένα των επιταχυνσιομέτρων και του μορφοτροπέα πίεσης.

Για την διαδικασία της διάγνωσης χρησιμοποιήθηκε το τέταρτο σημείο από κάθε βλάβη-κλάση, αυτό δηλαδή που δεν χρησιμοποιήθηκε για τις διαδικασίες της εκπαίδευσης και βελτιστοποίησης.

(α)

Μεταβλητές των kernels	set SVM			
	SIGMOID	POLYNOMIAL	LINEAR	RBF
gamma	1.0	1.0	-	3.0
coef	1.0	1.0	-	-
degree	-	3.0	-	-

(β)

Μεταβλητές των kernels	set SVM							
	Επιταχυνσιόμετρο 1				Επιταχυνσιόμετρο 2			
	SIGMOID	POLYNOMIAL	LINEAR	RBF	SIGMOID	POLYNOMIAL	LINEAR	RBF
gamma	1.0		-	0.02	1.0		-	0.02
coef	1.0	No SVs	-	-	1.0	No SVs	-	-
degree	-		-	-	-		-	-
	Επιταχυνσιόμετρο 3				Μορφοτροπέας πίεσης			
	SIGMOID	POLYNOMIAL	LINEAR	RBF	SIGMOID	POLYNOMIAL	LINEAR	RBF
	gamma	1.0		-	0.2	1.0		-
coef	1.0	No SVs	-	-	1.0	No SVs	-	-
degree	-		-	-	-		-	-

Πίνακας 4.25. Βέλτιστες τιμές των παραμέτρων των kernels στην περίπτωση θερμοδυναμικών μεγεθών (α) και επιταχυνσιομέτρων (β).

Ο Πίνακας 4.26 παρουσιάζει τα αποτελέσματα της διάγνωσης με χρήση θερμοδυναμικών μεγεθών, ενώ ο Πίνακας 4.27 τα αποτελέσματα της διάγνωσης με χρήση των δεδομένων των επιταχυνσιομέτρων και του μορφοτροπέα πίεσης.

Στα αποτελέσματα (Πίνακας 4.27) βλέπουμε πως η διάγνωση έγινε με τη χρήση των τριών kernels αφού με τη χρήση του polynomial δεν μπορούσε να επιτευχθεί ο διαχωρισμός των κλάσεων λόγω αδυναμίας εύρεσης των διανυσμάτων υποστήριξης (Support Vectors).

Πραγματική κλάση	SVM			
	LINEAR	POLYNOMIAL	RBF	SIGMOID
F2	F2	F2	F2	F2
F3	F3	F3	F3	F3
F4	F3	F3	F3	F3
F53	F53	F53	F53	F53
απόδοση	3/4	3/4	3/4	3/4

Πίνακας 4.26 Αποτελέσματα με χρήση θερμοδυναμικών μεγεθών.

Πραγματική κλάση	SVM – επιταχυνσιόμετρο 1			
	LINEAR	POLYNOMIAL	RBF	SIGMOID
F2	F2		F2	F4
F3	F4	NO SVs	F4	F4
F4	F53		F53	F53
F53	F4		F53	F4
απόδοση	1/4	-	2/4	0/4

Πραγματική κλάση	SVM – επιταχυνσιόμετρο 2			
	LINEAR	POLYNOMIAL	RBF	SIGMOID
F2	F2		F2	F4
F3	F4	NO SVs	F4	F4
F4	F53		F53	F53
F53	F3		F3	F53
απόδοση	1/4	-	1/4	1/4

Πραγματική κλάση	SVM – επιταχυνσιόμετρο 3			
	LINEAR	POLYNOMIAL	RBF	SIGMOID
F2	F2		F2	F4
F3	F4	NO SVs	F4	F4
F4	F53		F53	F53
F53	F3		F3	F3
απόδοση	1/4	-	1/4	0/4

Πραγματική κλάση	SVM – μορφοτροπέας πίεσης (PT-2)			
	LINEAR	POLYNOMIAL	RBF	SIGMOID
F2	F2		F2	F3
F3	F4	NO SVs	F4	F4
F4	F53		F53	F53
F53	F3		F53	F4
απόδοση	1/4	-	2/4	0/4

Πίνακας 4.27. Αποτελέσματα με χρήση δεδομένων από επιταχυνσιόμετρα και του μορφοτροπέα πίεσης.

Στα αποτελέσματα που προήλθαν από τη χρήση των δεδομένων από τα επιταχυνσιόμετρα και του μορφοτροπέα πίεσης (Πίνακας 4.27), βλέπουμε πως δε γίνεται δυνατή η σωστή κατάταξη των σημείων. Αυτό οφείλεται στο μικρό αριθμό σημείων (μόλις 3 για κάθε βλάβη) με τα οποία εκπαιδεύτηκε το διαγνωστικό πρόγραμμα. Η μη εύρεση διανυσμάτων υποστήριξης (SVs) στην περίπτωση του polynomial kernel οφείλεται σε αυτόν το λόγο. Ο μικρός αριθμός σημείων εκπαίδευσης δεν μπορεί να δώσει σαφή πληροφορία για τα όρια των κλάσεων. Αντίθετα, τα αποτελέσματα με χρήση θερμοδυναμικών μεγεθών (Πίνακας 4.26) είναι πολύ ικανοποιητικά, παρόλο που και σε αυτή τη περίπτωση έχουμε μικρό αριθμό σημείων για κάθε κλάση. Αυτή η καλύτερη συμπεριφορά οφείλεται στη φύση των δεδομένων αυτών που καθιστούν καθορίσιμα τα όρια των κλάσεων με μικρό αριθμό σημείων.

Στην περίπτωση της χρήσης μηχανικών μεγεθών, για να καθοριστούν πιο αποτελεσματικά τα όρια πρέπει να χρησιμοποιήσουμε μεγαλύτερο αριθμό σημείων για κάθε κλάση. Επειδή δεν διαθέτουμε επιπλέον στοιχεία, δημιουργούμε νέα σημεία από τα υπάρχοντα. Για να το πετύχουμε αυτό, υπολογίζουμε τους μέσους όρους όλων των δυνατών συνδυασμών των σημείων κάθε κλάσης ανά δύο, τρία και τέσσερα σημεία. Με αυτή τη διαδικασία, από τα 4 σημεία που είχαμε για κάθε βλάβη-κλάση, τώρα έχουμε 11. Τα 11 αυτά σημεία που δημιουργήθηκαν για κάθε βλάβη (συνολικά $4 \times 11 = 44$ για όλες τις βλάβες) χρησιμοποιούνται ως δεδομένα για την εκπαίδευση του

διαγνωστικού προγράμματος. Τα υπάρχοντα 4 σημεία για κάθε βλάβη (συνολικά $4 \times 4 = 16$ για όλες τις βλάβες) χρησιμοποιούνται για την διαδικασία της βελτιστοποίησης των τιμών των παραμέτρων των kernels (Πίνακας 4.28), καθώς επίσης και για τη διαδικασία της κατάταξης. Η διαδικασία αυτή πραγματοποιήθηκε για τα δεδομένα των τριών επιταχυνσιομέτρων, καθώς και για αυτά του μορφοτροπέα πίεσης.

Μεταβλητές των kernels	set SVM							
	Επιταχυνσιόμετρο 1				Επιταχυνσιόμετρο 2			
	SIGMOID	POLYNOMIAL	LINEAR	RBF	SIGMOID	POLYNOMIAL	LINEAR	RBF
gamma	1.0		-	0.01	1.0		-	0.01
coef	1.0	No SVs	-	-	1.0	No SVs	-	-
degree	-		-	-	-		-	-
	Επιταχυνσιόμετρο 3				Μορφοτροπέας πίεσης			
	SIGMOID	POLYNOMIAL	LINEAR	RBF	SIGMOID	POLYNOMIAL	LINEAR	RBF
	gamma	1.0	-	0.01	1.0		-	0.01
coef	1.0	No SVs	-	-	1.0	No SVs	-	-
degree	-		-	-	-		-	-

Πίνακας 4.28. Βέλτιστες τιμές των παραμέτρων των kernels στην περίπτωση χρήσης παράγωγων δεδομένων.

Τα αποτελέσματα που προέκυψαν με την παραπάνω διαδικασία και για τα τέσσερα kernels απεικονίζει ο Πίνακας 4.29. Η φράση 'NO SVs' στην περίπτωση του πολυωνυμικού υποδηλώνει ότι δεν κατέστη δυνατή η εκπαίδευση για αυτό το kernel. Τέλος, η σκίαση υποδηλώνει τις λάθος κατατάξεις.

Πραγματική κλάση	SVM - επιταχυνσιόμετρο 1			
	LINEAR	POLYNOMIAL	RBF	SIGMOID
F2	F2		F2	F4
F2	F2		F2	F4
F2	F2		F2	F4
F2	F2		F2	F4
F3	F3		F3	F4
F3	F3		F3	F4
F3	F3	NO SVs	F3	F4
F3	F3		F3	F4
F4	F4		F4	F4
F4	F4		F4	F4
F4	F4		F4	F4
F4	F4		F4	F2
F53	F53		F53	F2
F53	F53		F53	F53
F53	F53		F53	F53
F53	F53		F53	F53
απόδοση	16/16	-	16/16	6/16

Πραγματική κλάση	SVM - επιταχυνσιόμετρο 2			
	LINEAR	POLYNOMIAL	RBF	SIGMOID
F2	F2		F2	F3
F2	F2		F2	F3
F2	F2		F2	F4
F2	F2		F2	F4
F3	F3		F3	F4
F3	F3		F3	F4
F3	F3	NO SVs	F3	F4
F3	F3		F3	F4
F4	F4		F4	F53
F4	F4		F4	F53
F4	F4		F4	F4
F4	F53		F53	F53
F53	F53		F53	F53
F53	F53		F53	F53
F53	F53		F53	F53
F53	F53		F53	F53
απόδοση	15/16	-	15/16	5/16

Πραγματική κλάση	SVM - επιταχυνσιόμετρο 3				Πραγματική κλάση	SVM - μορφοτροπέας πίεσης (PT-2)			
	LINEAR	POLYNOMIAL	RBF	SIGMOID		LINEAR	POLYNOMIAL	RBF	SIGMOID
F2	F2		F2	F3	F2		F2	F3	
F2	F2		F2	F3	F2		F2	F3	
F2	F2		F2	F53	F2		F2	F2	
F2	F2		F2	F2	F2		F2	F3	
F3	F3		F3	F4	F3		F3	F4	
F3	F3		F3	F4	F3		F3	F4	
F3	F3	NO SVs	F3	F4	F3	NO SVs	F3	F3	
F3	F3		F3	F4	F3	F4	F4	F4	
F4	F4		F4	F53	F4		F4	F3	
F4	F4		F4	F53	F4		F4	F4	
F4	F4		F4	F4	F4		F4	F3	
F4	F53		F53	F53	F4		F4	F53	
F53	F53		F53	F53	F53		F53	F53	
F53	F53		F53	F53	F53		F53	F53	
F53	F53		F53	F4	F53		F53	F53	
F53	F53		F53	F3	F53		F53	F3	
απόδοση	15/16	-	15/16	4/16	απόδοση	15/16	-	15/16	6/16

Πίνακας 4.29 Αποτελέσματα με χρήση των παράγωγων δεδομένων για τα επιταχυνσιόμετρα και το μορφοτροπέα πίεσης.

Η δημιουργία περισσότερων δεδομένων για κάθε βλάβη, όπως βλέπουμε από τον παραπάνω πίνακα, επέφερε την σχεδόν απόλυτη επιτυχία στη διάγνωση. Με τη χρήση περισσότερων σημείων για κάθε βλάβη, κατέστη δυνατός ο επιτυχής καθορισμός των ορίων των κλάσεων με αποτέλεσμα τελείως διαφορετική συμπεριφορά του διαγνωστικού προγράμματος σε σχέση με τη χρήση λίγων σημείων (Πίνακας 4.27).

4.2.2 Ακτινικός συμπίεστής

Στην περίπτωση του ακτινικού συμπίεστή, μελετήθηκαν 3 βλάβες ($M1$: διαταραχή διαχύτη, $M2$: επικαθίσεις περωτής και $M3$: διαταραχή εισόδου) με δεδομένα 7 θερμοδυναμικά μεγέθη (ανηγμένη παροχή μάζας, λόγος πίεσης περωτής, ισεντροπικός βαθμός απόδοσης περωτής, λόγος πίεσης συμπίεστή, ισεντροπικός βαθμός απόδοσης συμπίεστή, παράμετρος παροχής (Φ) και παράμετρος φόρτισης (Kis) και μετρήσεις από επιταχυνσιόμετρο.

Για την κάθε βλάβη είχαμε 4 σημεία με μετρήσεις θερμοδυναμικών μεγεθών και 2 σετ (διαφορετικές ημερομηνίες μετρήσεων) με 4 σημεία για κάθε βλάβη από το επιταχυνσιόμετρο (για κάθε σετ). Όπως και στην περίπτωση του αξονικού, έτσι και εδώ θεωρούμε τον αριθμό των κλάσεων ίδιο με τον αριθμό των εξεταζόμενων βλαβών. Έτσι έχουμε 3 κλάσεις.

Το διαγνωστικό πρόγραμμα εφαρμόζεται ξεχωριστά για τα θερμοδυναμικά δεδομένα και για τα δεδομένα του επιταχυνσιόμετρου.

Στην περίπτωση των θερμοδυναμικών δεδομένων, από τα 4 σημεία που έχουμε για κάθε βλάβη-κλάση, τα 3 χρησιμοποιήθηκαν για την διαδικασία της εκπαίδευσης. Επίσης, με τα ίδια αυτά 3 σημεία έγινε η διαδικασία της βελτιστοποίησης κατά την οποία βρέθηκαν οι βέλτιστες τιμές των παραμέτρων των kernels (Πίνακας 4.30). Το τέταρτο σημείο από κάθε βλάβη-κλάση χρησιμοποιήθηκε για τη διαδικασία της κατάταξης.

Στην περίπτωση των μετρήσεων του επιταχυνσιόμετρου, η εκπαίδευση έγινε με 3 σημεία για κάθε βλάβη από το πρώτο σετ, η βελτιστοποίηση με 3 σημεία για κάθε βλάβη από το δεύτερο (Πίνακας 4.30) ενώ η διαδικασία της κατάταξης με 2 σημεία για κάθε βλάβη (ένα από κάθε σετ).

(α) Μεταβλητές των kernels	set SVM			
	SIGMOID	POLYNOMIAL	LINEAR	RBF
gamma	1.0	1.0	-	3.0
coef	1.0	1.0	-	-
degree	-	3.0	-	-

(β) Μεταβλητές των kernels	set SVM			
	SIGMOID	POLYNOMIAL	LINEAR	RBF
gamma	1.0	1.0	-	3.0
coef	1.0	1.0	-	-
degree	-	3.0	-	-

Πίνακας 4.30. Βέλτιστες τιμές των παραμέτρων των kernels στην περίπτωση θερμοδυναμικών μεγεθών (α) και επιταχυνσιόμετρου (β).

Ο Πίνακας 4.31 απεικονίζει τα αποτελέσματα της διάγνωσης στην περίπτωση που χρησιμοποιήθηκαν τα θερμοδυναμικά μεγέθη, ενώ ο Πίνακας 4.32 τα αποτελέσματα που προήλθαν από τη χρήση των δεδομένων του επιταχυνσιόμετρου.

Πραγματική κλάση	SVM			
	LINEAR	POLYNOMIAL	RBF	SIGMOID
M1	M1	M1	M1	M1
M2	M2	M2	M2	M2
M3	M3	M3	M3	M3
απόδοση	3/3	3/3	3/3	3/3

Πίνακας 4.31. Αποτελέσματα με χρήση θερμοδυναμικών μεγεθών.

Πραγματική κλάση	SVM			
	LINEAR	POLYNOMIAL	RBF	SIGMOID
M1	M1	M1	M1	M3
M1	M1	M1	M1	M3
M2	M3	M3	M3	M3
M2	M3	M3	M3	M3
M3	M3	M3	M1	M3
M3	M1	M1	M1	M3
απόδοση	3/6	3/6	2/6	2/6

Πίνακας 4.32. Αποτελέσματα με χρήση δεδομένων από επιταχυνσιόμετρο.

Όπως στην περίπτωση του αξονικού, έτσι κι εδώ βλέπουμε πως η επίδραση του μικρού αριθμού σημείων δρα αρνητικά στην ικανότητα της διάγνωσης στην περίπτωση της χρήσης των δεδομένων του επιταχυνσιόμετρο.

Για να φανεί το μέγεθος αυτής της επίδρασης, επαναλήφθηκε η προηγούμενη διαδικασία χρησιμοποιώντας 3 διαφορετικά σημεία από ότι πριν για την εκπαίδευση και την βελτιστοποίηση. Το τέταρτο σημείο για κάθε βλάβη (από κάθε σετ) χρησιμοποιήθηκε για την διαδικασία της κατάταξης. Η διαδικασία έγινε για τα δύο kernels που έδειξαν την καλύτερη συμπεριφορά ως τώρα. Τα αποτελέσματα δείχνει ο Πίνακας 4.33.

Πραγματική κλάση	SVM	
	POLYNOMIAL	RBF(3)
M1		M2
M1		M3
M2	NO SVs	M2
M2		M2
M3		M2
M3		M2
απόδοση	-	2/6

Πίνακας 4.33. Αποτελέσματα με χρήση δεδομένων από επιταχυνσιόμετρο.

Από τα αποτελέσματα του πίνακα αυτού βλέπουμε πως έχουμε τελείως διαφορετική συμπεριφορά κατάταξης και επί πλέον, ότι με τη χρήση του Polynomial kernel, αντίθετα με πριν, δεν κατέστη δυνατός ο διαχωρισμός των κλάσεων. Δηλαδή, τα αποτελέσματα εξαρτώνται άμεσα από το ποια σημεία θα χρησιμοποιηθούν παρουσιάζοντας έτσι στοχαστική συμπεριφορά. Αυτό συμβαίνει, όπως αναφέρθηκε και προηγουμένως, γιατί ο μικρός αριθμός σημείων και η φύση τους δεν επιτρέπουν τον σαφή προσδιορισμό των ορίων των κλάσεων.

Όπως στην περίπτωση του αξονικού συμπιεστή, έτσι κι εδώ, δημιουργούμε επιπλέον δεδομένα για κάθε βλάβη από τα υπάρχοντα ακολουθώντας τη διαδικασία που περιγράφηκε. Χρησιμοποιούμε το πρώτο σετ δεδομένων και έτσι, για κάθε βλάβη δημιουργούνται 11 σημεία (συνολικά 33 σημεία για τις τρεις βλάβες). Με αυτά τα 33 σημεία πραγματοποιούμε την εκπαίδευση. Οι διαδικασίες της βελτιστοποίησης και της κατάταξης πραγματοποιούνται με τη χρήση των δεδομένων του πρώτου σετ. Ο Πίνακας 4.34 παρουσιάζει τις βέλτιστες τιμές των παραμέτρων των kernels.

Μεταβλητές των kernels	set SVM			
	SIGMOID	POLYNOMIAL	LINEAR	RBF
gamma	1.0	0.01	-	0.01
coef	1.0	1.0	-	-
degree	-	3.0	-	-

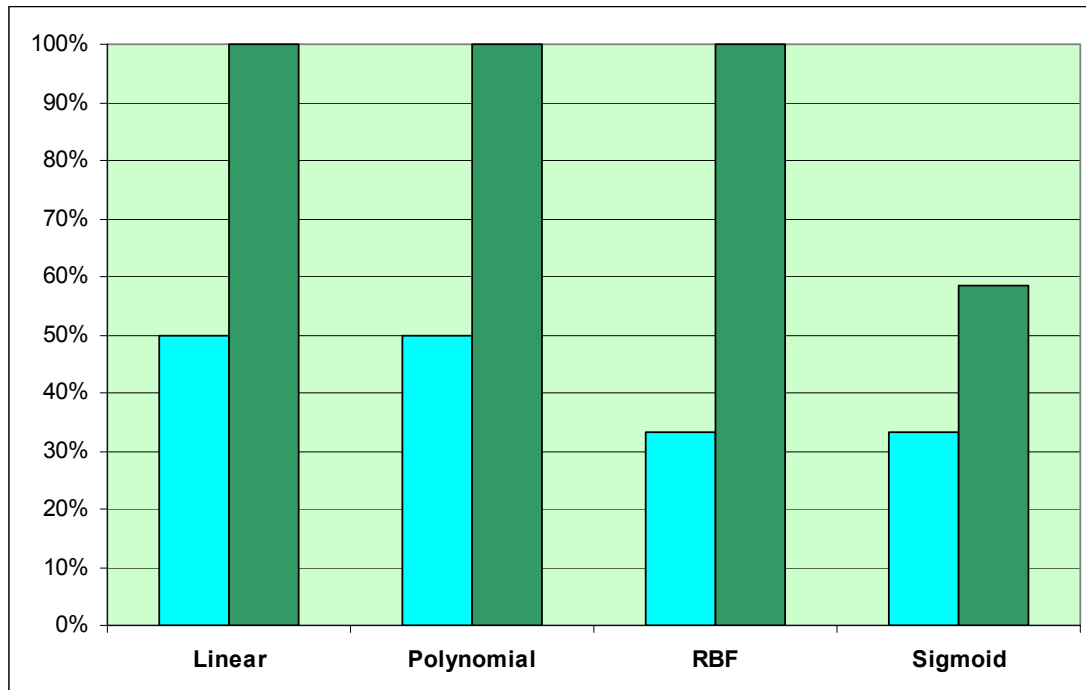
Πίνακας 4.34. Βέλτιστες τιμές των παραμέτρων των kernels στην περίπτωση χρήσης παράγωγων δεδομένων.

Τα αποτελέσματα που προέκυψαν κατά τη διαδικασία της κατάταξης και για τα τέσσερα kernels απεικονίζει ο Πίνακας 4.35. Η φράση ‘NO SVs’ στην περίπτωση του πολυωνυμικού υποδηλώνει ότι δεν κατέστη δυνατή η εκπαίδευση για αυτό το kernel. Τέλος, η σκίαση υποδηλώνει τις λάθος κατατάξεις.

Πραγματική κλάση	SVM - επιταχυνσιόμετρο			
	LINEAR	POLYNOMIAL	RBF	SIGMOID
M1	M1	M1	M1	M1
M1	M1	M1	M1	M3
M1	M1	M1	M1	M3
M1	M1	M1	M1	M3
M2	M2	M2	M2	M2
M2	M2	M2	M2	M3
M2	M2	M2	M2	M2
M2	M2	M2	M2	M3
M3	M3	M3	M3	M3
M3	M3	M3	M3	M3
M3	M3	M3	M3	M3
M3	M3	M3	M3	M3
απόδοση	12/12	12/12	12/12	7/12

Πίνακας 4.35. Αποτελέσματα με χρήση των παράγωγων δεδομένων για το επιταχυνσιόμετρο.

Από τον παραπάνω πίνακα και από το Σχήμα 4.4 γίνεται εμφανές ότι η χρήση περισσότερων δεδομένων για κάθε βλάβη καθιστά επιτυχώς καθορίσιμα τα όρια των κλάσεων με αποτέλεσμα την επιτυχή διάγνωση.



Σχήμα 4.4. Σύγκριση απόδοσης της διάγνωσης με χρήση πολλών (πράσινο) και λίγων σημείων (μπλέ) κατά την εκπαίδευση.

Κεφάλαιο 5: Ανακεφαλαίωση – Προτάσεις – Συμπεράσματα

5.1 Ανακεφαλαίωση

Τα κύρια σημεία της εργασίας που πραγματοποιήθηκε για την διαγνωστική αεριοστροβίλων με τη μέθοδο SVM μπορούν να συνοψιστούν στα εξής:

- Η τεχνική των SVM ανήκει στην οικογένεια των μεθόδων αναγνώρισης προτύπων. Αποτελούν μια σχετικά καινούρια τεχνική που χρειάζεται περαιτέρω μελέτη.
- Τα SVM βρίσκουν το υπερεπίπεδο που διαχωρίζει δύο κλάσεις με μεγιστοποίηση της απόστασης του υπερεπιπέδου από τις δύο αυτές κλάσεις ταυτόχρονα. Η σχετική θέση ενός σημείου με το διαχωριστικό υπερεπίπεδο καθορίζει σε ποια κλάση εκ των δύο θα καταταχθεί το σημείο.
- Τα προβλήματα που αντιμετωπίζονται με τη χρήση SVM χωρίζονται σε τρεις κατηγορίες: γραμμικά πλήρως διαχωρίσιμες κλάσεις, μη απόλυτα γραμμικά διαχωρίσιμες κλάσεις και μη γραμμικά διαχωρίσιμες κλάσεις. Στη πρώτη περίπτωση απλώς απαιτείται η εύρεση του βέλτιστου διαχωριστικού υπερεπιπέδου. Στη δεύτερη περίπτωση εισάγονται μεταβλητές χαλάρωσης ώστε να μην είναι αυστηρά τα όρια, ενώ στην τρίτη περίπτωση απαιτείται απεικόνιση των δεδομένων μας σε χώρο μεγαλύτερων διαστάσεων. Με τη τεχνική αυτή τα προβλήματα της δεύτερης περίπτωσης ανάγονται σε αυτά της πρώτης.
- Ουσιαστικό ρόλο για την απεικόνιση σε χώρο μεγαλύτερων διαστάσεων παίζουν οι συναρτήσεις κελύφους οι οποίες μας δίνουν το αποτέλεσμα του εσωτερικού γινομένου των σημείων σε αυτό το χώρο χωρίς να απαιτείται ο

προσδιορισμός αυτού του χώρου. Αυτό είναι χρήσιμο γιατί στις εξισώσεις των SVMs τα σημεία εμφανίζονται με την μορφή του εσωτερικού γινομένου τους.

- Για τον ταυτόχρονο διαχωρισμό πολλών κλάσεων απαιτείται η χρήση πολλών SVMs. Οι κλάσεις χωρίζονται σε ζεύγη και υπολογίζεται το διαχωριστικό υπερεπίπεδο κάθε ζεύγους. Η κατάταξη χρησιμοποιεί την τεχνική ‘max-wins’.
- Στα πλαίσια της παρούσας εργασίας αναπτύχθηκε πρόγραμμα που διαχωρίζει τις κλάσεις (κάθε κλάση μια κατάσταση λειτουργίας της μηχανής) και με βάση τις μετρήσεις της μηχανής κατατάσσει την εξεταζόμενη κατάσταση λειτουργίας σε κάποια από αυτές που έχει εκπαιδευτεί.
- Το πρόγραμμα αυτό εφαρμόστηκε στη διαγνωστική αεριοστροβίλου τόσο για τη διάγνωση βλαβών συνιστωσών, όσο και βλαβών αισθητήρων και οι κλάσεις ορίζονται από την μεταβολή συγκεκριμένων παραμέτρων υγείας ή αποκλίσεων των αισθητήρων και με συγκεκριμένο εύρος μεταβολής κάθε φορά. Οι καταστάσεις λειτουργίας που κλήθηκε να κατατάξει αποτελούσαν καταστάσεις λειτουργίας με βλάβη προσδιοριζόμενες από μετρήσεις πραγματικής μηχανής. Τα αποτελέσματα της διάγνωσης ήταν πάρα πολύ ικανοποιητικά.
- Με χρήση του προγράμματος, επιχειρήθηκε και ο προσδιορισμός του μεγέθους της βλάβης, πράγμα το οποίο επιτεύχθηκε επιτυχώς.
- Το πρόγραμμα χρησιμοποιήθηκε για τον προσδιορισμό τόσο θερμοδυναμικών βλαβών όσο και μηχανικών.

5.2 Συμπεράσματα

Από όσα προηγήθηκαν μπορούμε να εξάγουμε τα εξής συμπεράσματα:

- Η διάγνωση βλαβών σε αεριοστροβίλους με τη μέθοδο των SVM βρίσκεται σε αρχικό στάδιο με σημάδια έντονης ανάπτυξης. Ο ρόλος που μπορούν να

παίξουν σε αυτό το τομέα είναι σημαντικός και για αυτό αναμένουμε πληθώρα μελλοντικών εργασιών.

- Όπως διαπιστώνεται από τα αποτελέσματα της διάγνωσης αεριοστροβίλου που έγινε, έχουμε πολύ ικανοποιητικά συμπεράσματα και βλέπουμε πως η απόδοση της μεθόδου αυτής είναι υψηλή.
- Σημαντικές παράμετροι για την απόδοση της μεθόδου είναι η χρησιμοποιούμενη συνάρτηση κελύφους καθώς και το μέγεθος των δεδομένων εκπαίδευσης. Ανάλογα με τη συνάρτηση κελύφους μπορούμε να έχουμε έντονη διαφορά στην απόδοση της μεθόδου. Επίσης, όπως διαπιστώθηκε, όσο μεγαλύτερο όγκο δεδομένων εκπαίδευσης έχουμε, τόσο πιο σωστά υπολογίζονται τα όρια των κλάσεων και έτσι, έχουμε μεγαλύτερη πιθανότητα σωστής διάγνωσης.
- Από τη σύγκριση των αποτελεσμάτων των SVM με αυτά άλλων μεθόδων βλέπουμε ότι τα SVM υπερτερούν των BBN αφού εντοπίζουν τη βλάβη γρηγορότερα (μετά από λιγότερα σημεία) και ότι είναι παραπλήσια των PNN.

5.3 Προτάσεις

Από όσα είδαμε, μπορούμε να προτείνουμε τα εξής για την βελτίωση και την περαιτέρω ανάπτυξη του αλγόριθμου, ώστε να αποτελέσει ένα ισχυρό διαγνωστικό εργαλείο ειδικά στην περιοχή των αεριοστροβίλων.

- Το πιο χρονοβόρο κομμάτι του αλγορίθμου είναι η διαδικασία της κατάταξης καινούριων σημείων. Γι' αυτό, κρίνεται απαραίτητη η βελτιστοποίηση του αλγορίθμου ώστε να μειωθεί ο χρόνος που απαιτείται για την κατάταξη.
- Ο αλγόριθμος που δημιουργήθηκε δεν κάνει χρήση των εξισώσεων που περιέχουν τις μεταβλητές χαλάρωσης γιατί, όπως περιγράφηκε, θα αυξανόταν πολύ η πολυπλοκότητα και ο χρόνος με όχι βέβαιη βελτίωση των αποτελεσμάτων. Για πληρότητα και πιο αντικειμενική αντιμετώπιση, μπορούν

να χρησιμοποιηθούν οι εξισώσεις με τις μεταβλητές χαλάρωσης οι οποίες πρέπει να παίρνουν μια βέλτιστη τιμή για κάθε ζεύγος κλάσεων και όχι μια καθολική βέλτιστη τιμή για όλα τα ζεύγη.

- Ο αλγόριθμος, έτσι όπως σχεδιάστηκε, κάνει τον διαχωρισμό των κλάσεων σε ένα βέλτιστο χώρο. Θεωρούμε δηλαδή ότι απεικονίζουμε τα δεδομένα μας σε ένα χώρο καινούριο ώστε να επιτύχουμε καλύτερο διαχωρισμό των κλάσεων και έτσι λιγότερα λάθη κατάταξης. Μια πρόταση είναι, κάθε ζεύγος να απεικονίζεται σε διαφορετικό χώρο, τον βέλτιστο κάθε φορά. Δηλαδή το κάθε ζεύγος να απεικονίζεται στο χώρο που οι δύο κλάσεις διαχωρίζονται βέλτιστα. Η διαδικασία της κατάταξης θα γίνεται για κάθε ζεύγος ξεχωριστά, στο χώρο που το ζεύγος θα έχει διαχωριστεί βέλτιστα. Μπορεί να συνεχίσει να ακολουθείται η λογική του 'max-wins' και για τις κλάσεις που ισοψηφούν, να βρίσκεται (ίσως) η ευκλείδεια απόσταση του σημείου από το κέντρο των κλάσεων, στον χώρο στον οποίο ανήκουν πραγματικά τα δεδομένα μας και έτσι να λαμβάνεται η απόφαση.
- Στον αλγόριθμο που υλοποιήθηκε υπάρχει η επιλογή χρήσης ενός εκ τεσσάρων συναρτήσεων κελύφους οι οποίες βιβλιογραφικά είχαν την καλύτερη συμπεριφορά σε πληθώρα προβλημάτων. Κρίνεται αναγκαίο να γίνει περαιτέρω διερεύνηση των συναρτήσεων κελύφους ώστε να χρησιμοποιηθεί κάποια (καινούρια ίσως) που να έχει καλύτερη συμπεριφορά στο είδος των δεδομένων μας.
- Τέλος, μια ταχύτατα αναπτυσσόμενη τεχνική είναι αυτή της χρήσης πολλών διαγνωστικών τεχνικών ταυτόχρονα ώστε να εξαχθεί ένα συμπέρασμα (fusion). Προτείνεται η μελέτη της χρήσης των SVM με άλλες τεχνικές για περαιτέρω βελτίωση των αποτελεσμάτων.

ΠΑΡΑΡΤΗΜΑΤΑ

Παράρτημα I

Π1.1 Το διαγνωστικό πρόγραμμα που δημιουργήθηκε

Ο κώδικας που δημιουργήθηκε υλοποιεί τον διαχωρισμό των κλάσεων (εκπαίδευση) και κατατάσσει νέα σημεία στις κλάσεις με τις οποίες εκπαιδεύτηκε. Βρίσκει δηλαδή τα υπερεπίπεδα που διαχωρίζουν τις κλάσεις και με βάση αυτό το διαχωρισμό λαμβάνεται η απόφαση για την κατάταξη καινούριου σημείου.

Π1.2 Χρησιμοποιούμενα αρχεία

Τα χρησιμοποιούμενα από τον κώδικα αρχεία είναι:

Όνομα αρχείου	Μορφή ονόματος	Format	Πληροφορία
inputall.dat	Σταθερή	.dat file	Εισόδου
testing.dat	Σταθερή	.dat file	Εισόδου
LogFile.dat	Σταθερή	.dat file	Εξόδου
output.dat	Σταθερή	.dat file	Εξόδου

Περιγραφή χρησιμοποιούμενων αρχείων:

Αρχείο: inputall.dat

Γραμμή	Περιεχόμενο
1	datakernel(1) datakernel(2) datakernel(3) datakernel(4)
2	acura acurc
3	itest1y0n
4	iclasses
5	icolumns
6	idataclass(1) idataclass(2) ... idataclass(j)
7	xxall(1,1) xxall(1,2) ... xxall(1,j)
....	
6+i	xxall(i,1) xxall(i,2) ... xxall(i,j)

Αρχείο: testing.dat

Γραμμή	Περιεχόμενο
1	inubtest
2	xtest(1) xtest(2) ... xtest(j)
...	
1+inubtest	xtest(1) xtest(2) ... xtest(j)

Αρχείο: LogFile.dat

Γραμμή	Περιεχόμενο
1	TRAINING: "OK!" or "NO SVs OR ACCURACY PROBLEM"
2	CLASSIFICATION: "OK!" or "IMPOSSIBLE OR NOT SELECTED"
3	CLASSIFICATION: "ALL CLASSIFIED!" or "SECONDLY CLASSIFIED: i2classifier "

Αρχείο: output.dat

Γραμμή	Περιεχόμενο
1	"inubclass" or "inubclass,'S.C.'"
...	
inubtest	"inubclass" or "inubclass,'S.C.'"

Π1.3 Χρησιμοποιούμενες Μεταβλητές Εισόδου/Εξόδου

Όνομα	Format	Περιγραφή
datakernel(1)	real*8	* Επιλογή τύπου kernel (1:linear 2:polynomial 3:sigmoid 4:RBF)
datakernel(2)	real*8	* kernel variable: gama
datakernel(3)	real*8	* kernel variable: coef
datakernel(4)	real*8	* kernel variable: degree
acura	real*8	** Ακρίβεια για την εκπαίδευση
acurc	real*8	** Ακρίβεια για την κατάταξη
itestly0n	integer*4	0: μόνο εκπαίδευση 1: και κατάταξη
iclassess	integer*4	Αριθμός κλάσεων
icolumns	integer*4	Μέγεθος σημείου εκπαίδευσης (διάνυσμα μεγέθους j, jmax=12)
idataclass(j)	integer*4	Αριθμός σημείων σε κάθε κλάση
xxall(i,j)	real*8	Πίνακας με τα σημεία εκπαίδευσης (imax=1000, jmax=12)
inubtest	integer*4	Αριθμός σημείων για κατάταξη
xtest(j)	real*8	Προς κατάταξη σημείο (διάνυσμα μεγέθους j, jmax=12)
inubclass (S.C.)	integer*4	Κλάση κατάταξης (μέσω δεύτερης κατάταξης)
i2classifier	integer*4	*** Αριθμός σημείων που κατατάχθηκαν τελικώς μέσω δεύτερης κατάταξης

Επεξήγηση μεταβλητών Input/Output

* Ο χρήστης μπορεί να επιλέξει ένα εκ των τεσσάρων kernels που υπάρχουν στον κώδικα. Αυτά είναι:

$$K(x, y) = x \cdot y \quad \text{Γραμμικό kernel (Linear kernel)}$$

$$K(x, y) = (gama \cdot (x \cdot y) + coef)^{degree} \quad \text{Πολυωνυμικό kernel (Polynomial kernel)}$$

$$K(x, y) = e^{-gama \cdot \|x-y\|^2} \quad \text{RBF kernel}$$

$$K(x, y) = \tanh(gama \cdot (x \cdot y) - coef) \quad \text{Σιγμοειδές kernel (Sigmoid kernel)}$$

Οι μεταβλητές των kernels gama, coef, degree επιλέγονται από το χρήστη.

** Κατά τη διαδικασία της εκπαίδευσης, όσα σημεία έχουν πολλαπλασιαστική Lagrange θετικό, αυτά είναι και Support Vectors (SVs). Ο κώδικας επιτρέπει στο χρήστη να θέσει την τιμή 0 στους πολλαπλασιαστές Lagrange που έχουν τιμή μικρότερη ή ίση με την τιμή της μεταβλητής acura.

Κατά τη διαδικασία της κατάταξης κάποιο σημείο μπορεί να βρίσκεται πάρα πολύ κοντά σε κάποιο διαχωριστικό υπερεπίπεδο οπότε ο χρήστης να επιθυμεί αυτό το σημείο να καταταχθεί και στις δύο κλάσεις που διαχωρίζει αυτό το υπερεπίπεδο. Έτσι, αν η συνάρτηση απόφασης του σημείου για αυτό το υπερεπίπεδο έχει τιμή μικρότερη ή ίση (σε απόλυτη τιμή) με την τιμή της μεταβλητής acure, τότε το σημείο κατατάσσεται και στις δύο κλάσεις. Λόγω των πολλών κλάσεων δε δημιουργείται πρόβλημα με την κατάταξη ενός σημείου σε δυο κλάσεις ταυτόχρονα, ή και αν ακόμα υπάρξει ισοβαθμία δύο κλάσεων λόγω αυτής της διαδικασίας, το πρόβλημα λύνεται αυτόματα μέσω της διαδικασίας της δεύτερης κατάταξης.

*** Στους παραπάνω πίνακες αναφέρεται ότι κάποια σημεία τοποθετούνται στην τελική τους κλάση μέσω δεύτερης κατάταξης. Κάποιο σημείο μπορεί να καταταχθεί ίσες φορές σε πάνω από μια κλάσεις (με τη μέθοδο max-wins) λόγω περιοχών αλληλοεπικάλυψης των διαχωριστικών υπερεπιπέδων, οπότε για να βρεθεί η κλάση στην οποία θα τοποθετηθεί τελικά υπολογίζονται οι αποστάσεις του σημείου

από τα διαχωριστικά επίπεδα των ισοβαθμούντων κλάσεων και για το υπερεπίπεδο i,j με τη μικρότερη απόσταση ισχύει: Αν η συνάρτηση απόφασης του σημείου για αυτό το υπερεπίπεδο είναι >0 τότε το σημείο κατατάσσεται στην κλάση j , αλλιώς κατατάσσεται στην κλάση i .

Π2.1 Ανάλυση του προγράμματος

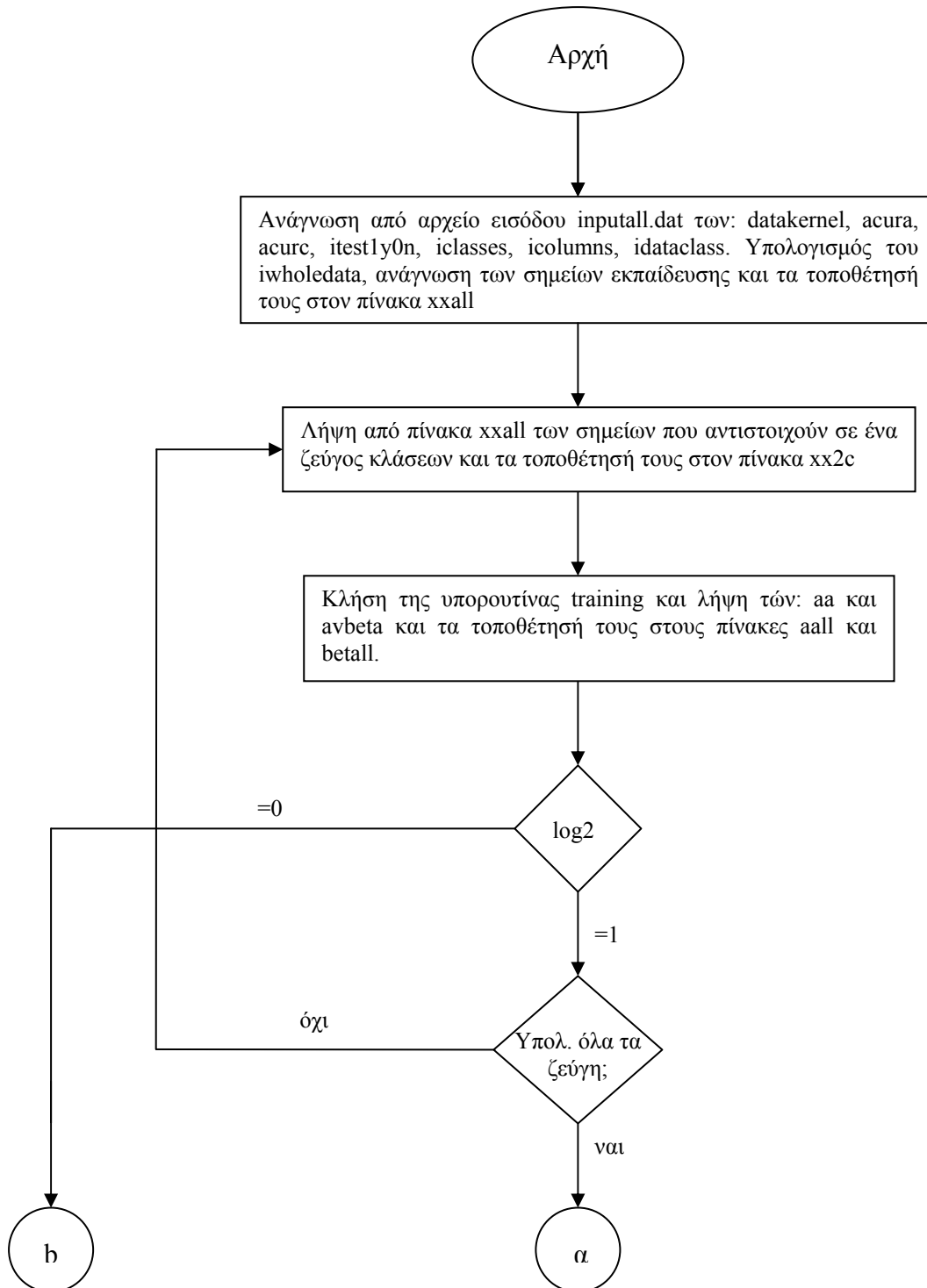
Η μέχρι τώρα επεξήγηση αφορά το τρόπο χρήσης του προγράμματος χωρίς να είναι απαραίτητη μια πιο βαθιά γνώση του τρόπου υλοποίησης των διαδικασιών που αυτό εκτελεί. Σε αυτή τη παράγραφο θα αναλυθεί ο τρόπος που το πρόγραμμα υλοποιεί τις απαιτούμενες διαδικασίες.

Π2.2 Χρησιμοποιούμενες μεταβλητές από το πρόγραμμα

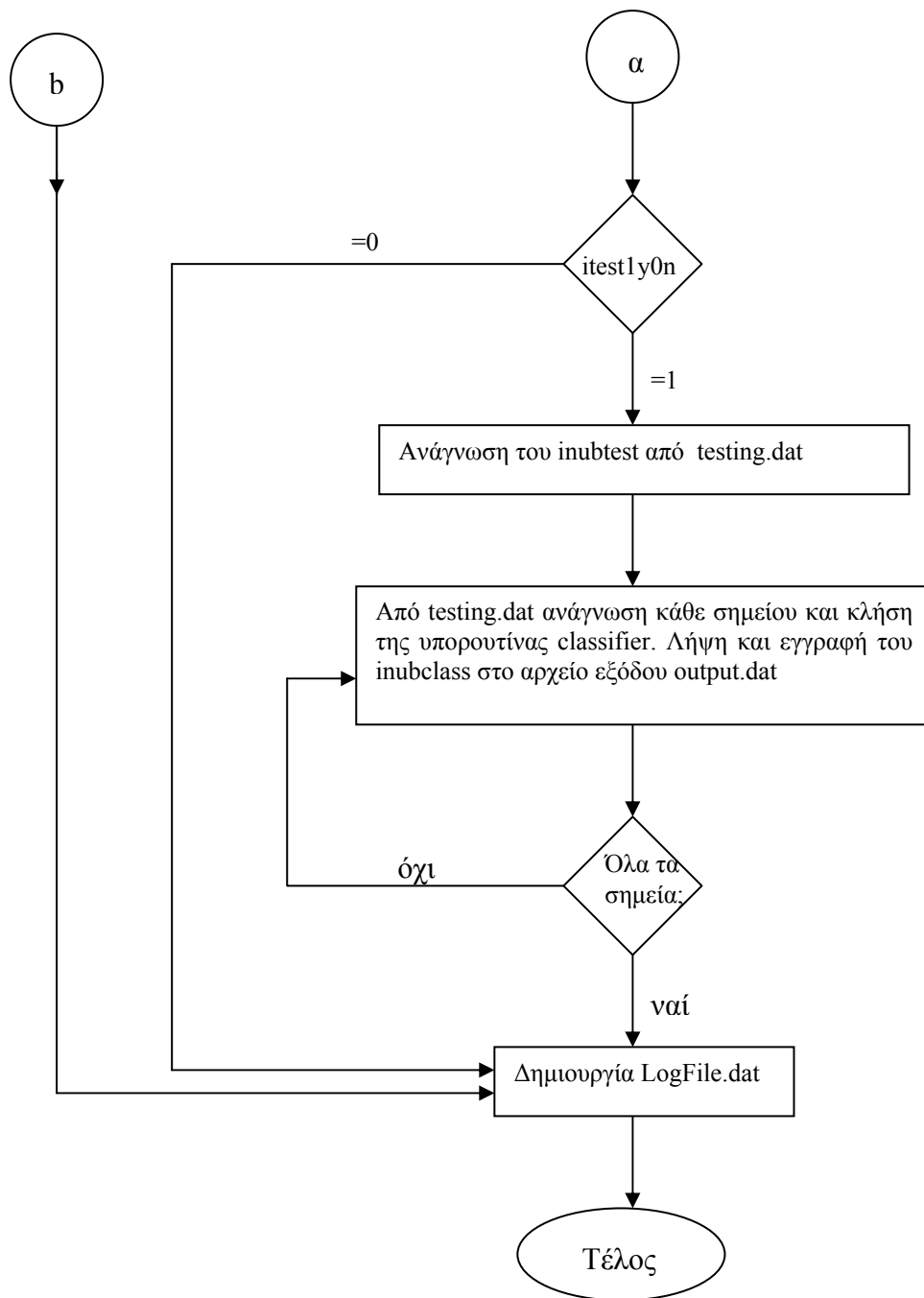
Όνομα	Format (διάσταση)	Περιγραφή
A_imsl(i,j)	real*8 (ncon,nvar)	πίνακας για την υπορουτίνα DQPROG της IMSL
aa(i)	real*8 (ik)	πολλαπλασιαστές Lagrange των σημείων του ζεύγους κλάσεων υπολογιζόμενοι κατά την εκπαίδευση
aa2c(i)	real*8 (lk)	πολλαπλασιαστές Lagrange των σημείων του ζεύγους κλάσεων για την διαδικασία της κατάταξης
aall(i)	real*8 (14*ik)	πολλαπλασιαστές Lagrange όλων των σημείων εκπαίδευσης
acura	real*8	ακρίβεια για την εκπαίδευση (βλ. προηγούμενα)
acurc	real*8	ακρίβεια για την κατάταξη (βλ. προηγούμενα)
aKxx	real*8	αποτέλεσμα εσωτερικού γινομένου δύο σημείων μέσω kernel
avbeta	real*8	μέσος όρος των υπολογιζόμενων beta στην εκπαίδευση
B_imsl(i)	real*8 (ncon)	πίνακας για την υπορουτίνα DQPROG της IMSL
beta(i)	real*8 (kbet)	οι υπολογιζόμενοι συντελεστές από τα SVs του ζεύγους κλάσεων
betall(i)	real*8 (14*ik)	οι συντελεστές beta των διαχωριστικών υπερεπιπέδων όλων των ζευγών κλάσεων
coef	real*8	μεταβλητή των kernels (βλ. προηγούμενα)
datakernel(i)	real*8 (4)	πληροφορίες για τα kernels (βλ. προηγούμενα)
degree	real*8	μεταβλητή των kernels (βλ. προηγούμενα)
distFx(i)	real*8 (ik)	απόσταση του κάθε προς κατάταξη σημείου από το κάθε διαχωριστικό υπερεπίπεδο
Fx	real*8	το πρόσημο της συνάρτησης απόφασης
G_imsl(i)	real*8 (nvar)	πίνακας για την υπορουτίνα DQPROG της IMSL
gama	real*8	μεταβλητή των kernels (βλ. προηγούμενα)
H_imsl	real*8	πίνακας για την υπορουτίνα DQPROG της IMSL
i2classifier	integer*4	αριθμός προς κατάταξη σημείων που χρειάστηκαν διαδικασία δεύτερης κατάταξης

Όνομα	Format (διάσταση)	Περιγραφή
iall dat	integer*4	αριθμός σημείων στο ζεύγος για την εκπαίδευση
icheck	integer*4	δείχνει αν το σημείο κατατάχθηκε με την πρώτη διαδικασία
iclassess	integer*4	αριθμός κλάσεων
icolumns	integer*4	μέγεθος διανύσματος σημείων εκπαίδευσης (και κατάταξης)
idat_class(i)	integer*4 (2)	αριθμός σημείων σε κάθε κλάση του ζεύγους
idataclass(i)	integer*4 (20)	αριθμός σημείων σε κάθε κλάση
idatanow	integer*4	δείχνει τη σειρά του πίνακα xxall που έχει χρησιμοποιηθεί ως τώρα
iflag	integer*4	μετρητής στην υπορουτίνα minimization που αναγκάζει να γίνουν 3 επαναλήψεις για σίγουρη σύγκλιση
iflagclass(i)	integer*4 (iclassess)	μετρά πόσες φορές κατατάχθηκε ένα σημείο στην κάθε κλάση
ik	integer*4	μέγιστος αριθμός σημείων εκπαίδευσης
inubclass	integer*4	κλάση τελικής κατάταξης του σημείου
inubtest	integer*4	αριθμός προς κατάταξη σημείων
isignFx(i)	integer*4 (ik)	1 if $F_x < 0$ 2 if $F_x > 0$
itest1y0n	integer*4	0: μόνο εκπαίδευση 1: και κατάταξη
iwholedata	integer*4	ο αριθμός όλων των σημείων για την εκπαίδευση
jk	integer*4	μέγιστο μέγεθος διανύσματος σημείου εκπαίδευσης (κατάταξη)
kbet	integer*4	μετρητής των SVs κάθε ζεύγους
lk	integer*4	μέγιστος αριθμός σημείων κάθε ζεύγους κλάσεων για την εκπαίδευση
log2	integer*4	=0 πρόβλημα στην εκπαίδευση
log3	integer*4	=0 αδύνατη η κατάταξη ή μη επιλεγμένη
log4	integer*4	πληροφορία για την διαδικασία κατάταξης
ncon	integer*4	μεταβλητή για την υπορουτίνα DQPROG
neq	integer*4	μεταβλητή για την υπορουτίνα DQPROG
nvar	integer*4	μεταβλητή για την υπορουτίνα DQPROG
nzero	integer*4	μετρά τους μηδενισμούς στην υπορουτίνα minimization
sol(i)	real*8 (nvar)	μεταβλητή υπορουτίνας DQPROG (το αποτέλεσμα της ελαχιστοποίησης) ($\Xi(\alpha(i))$)
sumbeta	real*8	άθροισμα των beta του κάθε ζεύγους
x1(i)	real*8 (icolumns)	σημείο για τον υπολογισμό εσωτερικού γινομένου από kernel
x2(i)	real*8 (icolumns)	σημείο για τον υπολογισμό εσωτερικού γινομένου από kernel
xtest(i)	real*8 (icolumns)	προς κατάταξη σημείο
xx2c(i,j)	real*8 (lk,jk)	όλα τα σημεία του ζεύγους των δύο κλάσεων
xxall(i,j)	real*8 (ik,jk)	όλα τα σημεία της εκπαίδευσης
yy(i)	real*8 (iall dat)	ετικέτα της κάθε κλάσης για το ζεύγος κλάσεων (1 ή -1)

Π2.3 Λογικό διάγραμμα προγράμματος



Συνεχίζεται στην επόμενη σελίδα



Π2.4 Περιγραφή υπορουτινών

Υπορουτίνα: training

Σκοπός αυτής της υπορουτίνας είναι να επιτύχει το διαχωρισμό ενός ζεύγους κλάσεων, δηλαδή να βρεί την εξίσωση του διαχωριστικού υπερεπιπέδου. Χρησιμοποιείται η υπορουτίνα minimization καθώς και μια συνάρτηση function που μας δίνει το αποτέλεσμα του εσωτερικού γινομένου δύο σημείων μέσω των kernels.

Κλήση:

call training(apura, idat_class, iall_dat, xx, datakernel, icolumns, avbeta, log2, aa)

Input: apura, idat_class, iall_dat, xx, datakernel, icolumns

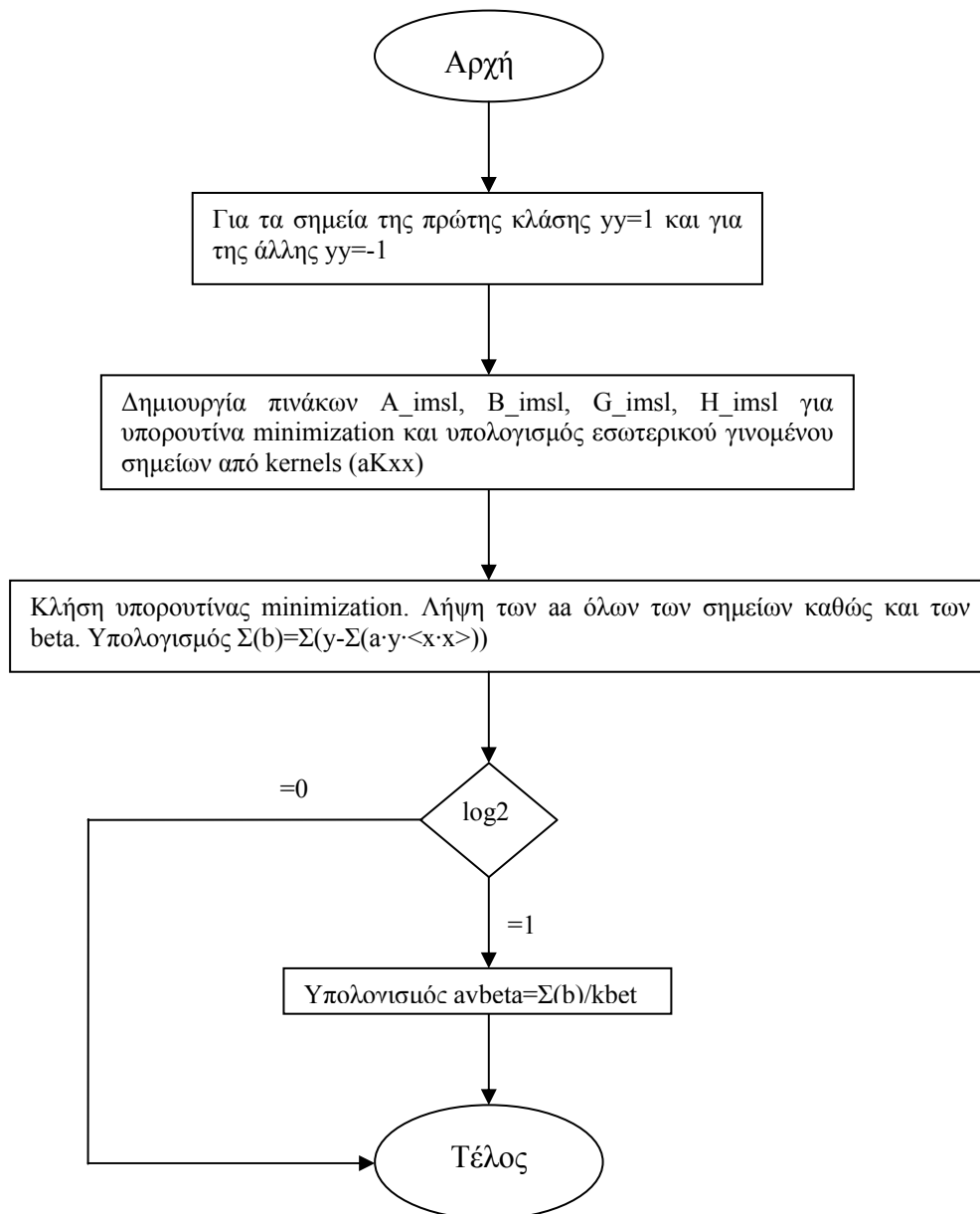
Output: avbeta, log2, aa

Μεταβλητές:

Όνομα	Format (διάσταση)	Περιγραφή
A_imsl(i,j)	real*8 (ncon,nvar)	πίνακας για την υπορουτίνα DQPROG της IMSL
aa(i)	real*8 (ik)	πολλαπλασιαστές Lagrange των σημείων του ζεύγους κλάσεων υπολογιζόμενοι κατά την εκπαίδευση
apura	real*8	ακρίβεια για την εκπαίδευση (βλ. προηγούμενα)
aKxx	real*8	αποτέλεσμα εσωτερικού γινομένου δύο σημείων μέσω kernel
avbeta	real*8	μέσος όρος των υπολογιζόμενων beta στην εκπαίδευση
B_imsl(i)	real*8 (ncon)	πίνακας για την υπορουτίνα DQPROG της IMSL
beta(i)	real*8 (kbet)	οι υπολογιζόμενοι συντελεστές από τα SVs του ζεύγους κλάσεων
coef	real*8	μεταβλητή των kernels (βλ. προηγούμενα)
datakernel(i)	real*8 (4)	πληροφορίες για τα kernels (βλ. προηγούμενα)
degree	real*8	μεταβλητή των kernels (βλ. προηγούμενα)
G_imsl(i)	real*8 (nvar)	πίνακας για την υπορουτίνα DQPROG της IMSL
gama	real*8	μεταβλητή των kernels (βλ. προηγούμενα)
H_imsl	real*8	πίνακας για την υπορουτίνα DQPROG της IMSL
iall_dat	integer*4	αριθμός σημείων στο ζεύγος για την εκπαίδευση
icolumns	integer*4	μέγεθος διανύσματος σημείων εκπαίδευσης (και κατάταξης)
idat_class(i)	integer*4 (2)	αριθμός σημείων σε κάθε κλάση του ζεύγους
idatclass(i)	integer*4 (20)	αριθμός σημείων σε κάθε κλάση
ik	integer*4	μέγιστος αριθμός σημείων εκπαίδευσης
jk	integer*4	μέγιστο μέγεθος διανύσματος σημείου εκπαίδευσης (κατάταξη)
kbet	integer*4	μετρητής των SVs κάθε ζεύγους
lk	integer*4	μέγιστος αριθμός σημείων κάθε ζεύγους κλάσεων για την εκπαίδευση
log2	integer*4	=0 πρόβλημα στην εκπαίδευση
ncon	integer*4	μεταβλητή για την υπορουτίνα DQPROG
neq	integer*4	μεταβλητή για την υπορουτίνα DQPROG
nvar	integer*4	μεταβλητή για την υπορουτίνα DQPROG

Όνομα	Format (διάσταση)	Περιγραφή
sol(i)	real*8 (nvar)	μεταβλητή υπορουτίνας DQPROG (το αποτέλεσμα της ελαχιστοποίησης) ($\equiv a(i)$)
sumbeta	real*8	άθροισμα των beta του κάθε ζεύγους
x1(i)	real*8 (icolumns)	σημείο για τον υπολογισμό εσωτερικού γινομένου από kernel
x2(i)	real*8 (icolumns)	σημείο για τον υπολογισμό εσωτερικού γινομένου από kernel
yy(i)	real*8 (iall_dat)	ετικέτα της κάθε κλάσης για το ζεύγος κλάσεων (1 ή -1)

Στη συνέχεια παρουσιάζεται το λογικό διάγραμμα της υπορουτίνας training.



Υπορουτίνα: minimization

Σκοπός αυτής της υπορουτίνας είναι να ελαχιστοποιήσει την $-Ld = -[\sum_{i=1}^l a_i - \frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j < x_i x_j >]$ με τους περιορισμούς: $\sum_{i=1}^l a_i y_i = 0$, $a_i \geq 0$.

Γίνεται χρήση της υπορουτίνας DQPROG της IMSL. Η DQPROG καλείται τόσες φορές έτσι ώστε να υπάρχει σύγκλιση, ενώ πριν από κάθε κάλεσμα μειώνεται η διάσταση του προβλήματος (μηδενισμός σειρών και στηλών πινάκων) ανάλογα με το πόσοι πολλαπλασιαστές Lagrange (α) πήραν την τιμή 0 (ή τιμή κοντά στο 0 και τέθηκε από την minimization 0). Εφόσον, τελικά, δεν υπάρχει περαιτέρω μείωση, η υπορουτίνα DQPROG καλείται να τρέξει τρεις επιπλέον φορές για βεβαίωση της σύγκλισης.

Κλήση:

call minimization(nvar, ncon, neq, a_imsl, b_imsl, g_imsl, h_imsl, aa)

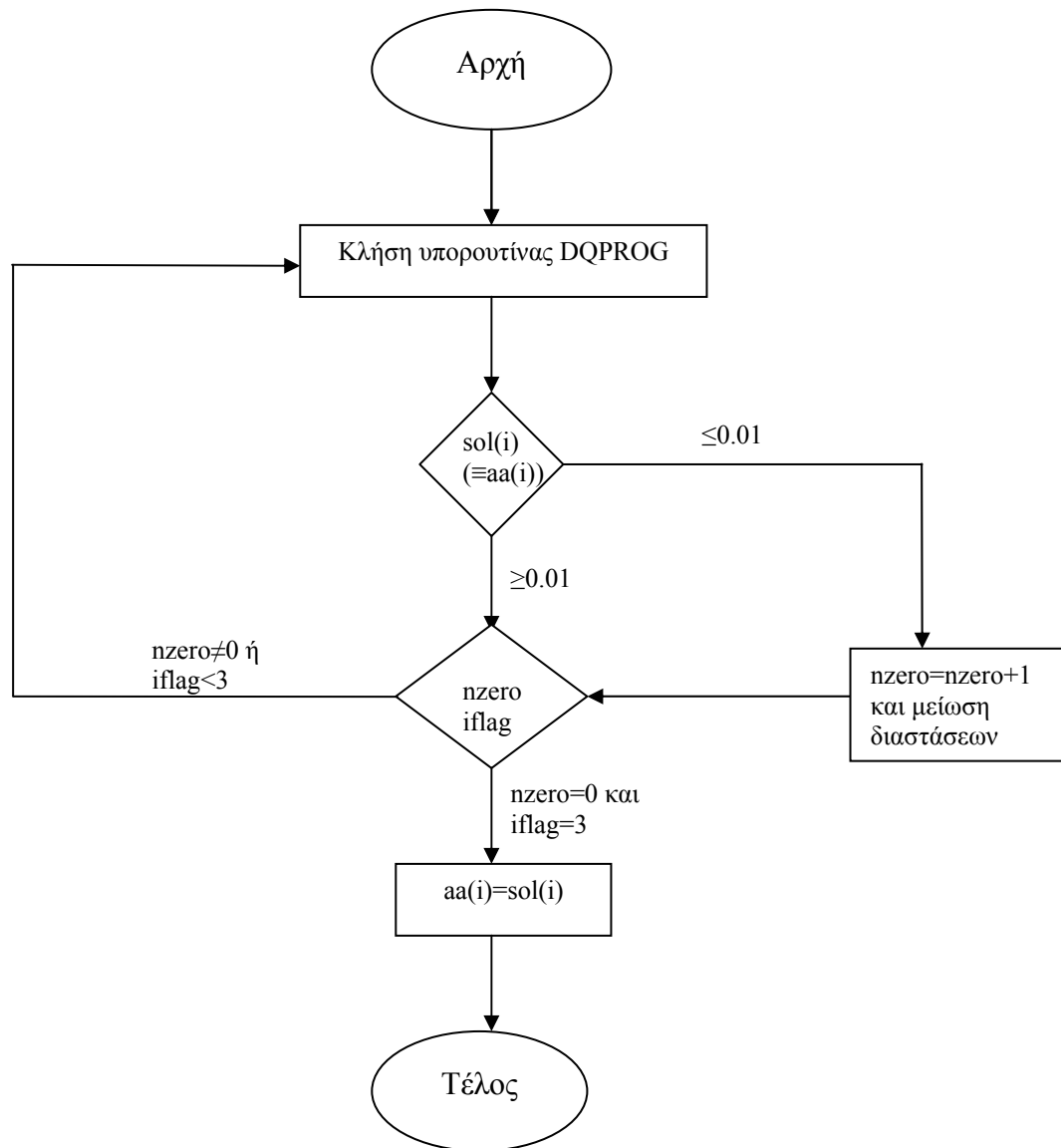
Input: nvar, ncon, neq, a_imsl, b_imsl, g_imsl, h_imsl

Output: aa

Μεταβλητές:

Όνομα	Format (διάσταση)	Περιγραφή
A_imsl(i,j)	real*8 (ncon,nvar)	πίνακας για την υπορουτίνα DQPROG της IMSL
aa(i)	real*8 (ik)	πολλαπλασιαστές Lagrange των σημείων του ζεύγους κλάσεων υπολογιζόμενοι κατά την εκπαίδευση
B_imsl(i)	real*8 (ncon)	πίνακας για την υπορουτίνα DQPROG της IMSL
G_imsl(i)	real*8 (nvar)	πίνακας για την υπορουτίνα DQPROG της IMSL
gama	real*8	μεταβλητή των kernels (βλ. προηγούμενα)
H_imsl	real*8	πίνακας για την υπορουτίνα DQPROG της IMSL
iflag	integer*4	μετρητής στην υπορουτίνα minimization που αναγκάζει να γίνουν 3 επαναλήψεις για σίγουρη σύγκλιση
ncon	integer*4	μεταβλητή για την υπορουτίνα DQPROG
neq	integer*4	μεταβλητή για την υπορουτίνα DQPROG
nvar	integer*4	μεταβλητή για την υπορουτίνα DQPROG
nzero	integer*4	μετρά τους μηδενισμούς στην υπορουτίνα minimization
sol(i)	real*8 (nvar)	μεταβλητή υπορουτίνας DQPROG (το αποτέλεσμα της ελαχιστοποίησης) ($\equiv \alpha(i)$)

Στη συνέχεια παρουσιάζεται το λογικό διάγραμμα της υπορουτίνας minimization.



Υπορουτίνα: classifier

Σκοπός της υπορουτίνας αυτής είναι να κατατάξει ένα σημείο στην κλάση που ανήκει.

Κλήση:

call classifier (iclasses, xxall, betall, aall, idataclass, acurc, xtest, inubclass, icolumns, icheck, datakernel)

Input: iclasses, xxall, betall, aall, idataclass, acurc, xtest, icolumns, datakernel

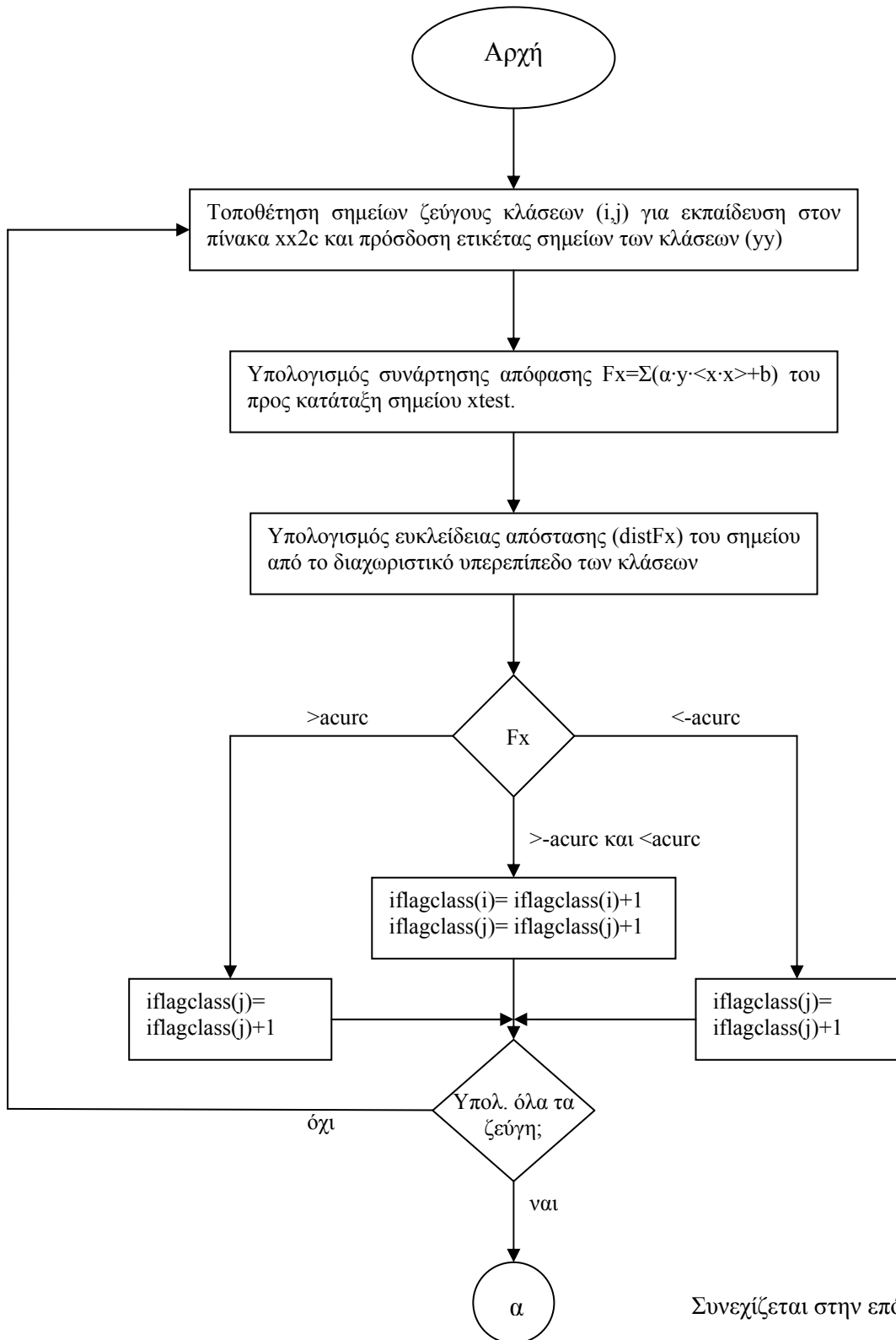
Output: inubclass, icheck

Μεταβλητές:

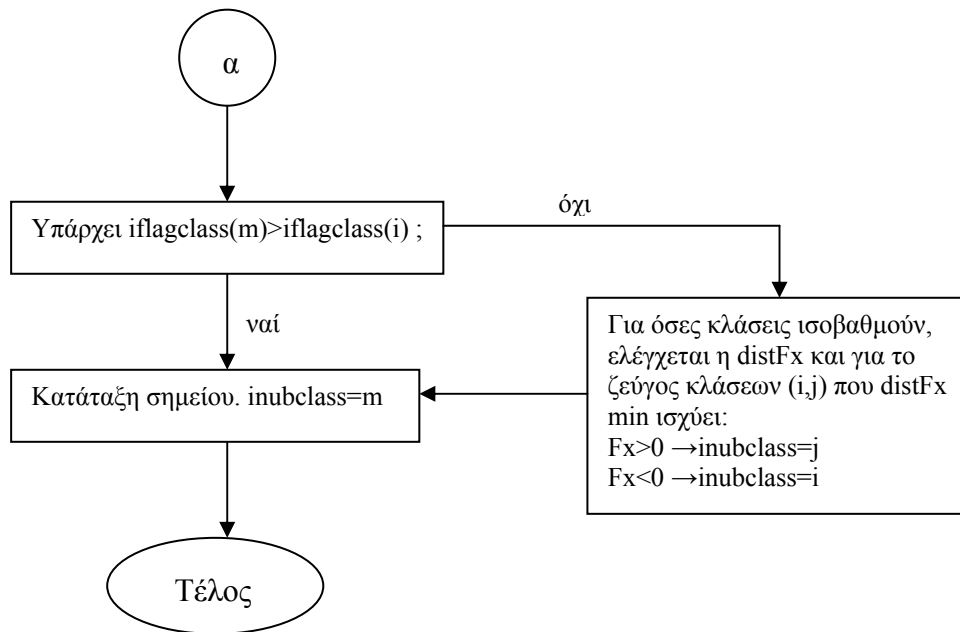
Όνομα	Format (διάσταση)	Περιγραφή
aa(i)	real*8 (ik)	πολλαπλασιαστές Lagrange των σημείων του ζεύγους κλάσεων υπολογιζόμενοι κατά την εκπαίδευση
aa2c(i)	real*8 (lk)	πολλαπλασιαστές Lagrange των σημείων του ζεύγους κλάσεων για την διαδικασία της κατάταξης
aall(i)	real*8 (14*ik)	πολλαπλασιαστές Lagrange όλων των σημείων εκπαίδευσης
aKxx	real*8	αποτέλεσμα εσωτερικού γινομένου δύο σημείων μέσω kernel
betall(i)	real*8 (14*ik)	οι συντελεστές beta των διαχωριστικών υπερεπιπέδων όλων των ζευγών κλάσεων
coef	real*8	μεταβλητή των kernels (βλ. προηγούμενα)
datakernel(i)	real*8 (4)	πληροφορίες για τα kernels (βλ. προηγούμενα)
degree	real*8	μεταβλητή των kernels (βλ. προηγούμενα)
distFx(i)	real*8 (ik)	απόσταση του κάθε προς κατάταξη σημείου από το κάθε διαχωριστικό υπερεπίπεδο
Fx	real*8	το πρόσημο της συνάρτησης απόφασης
gama	real*8	μεταβλητή των kernels (βλ. προηγούμενα)
icheck	integer*4	δείχνει αν το σημείο κατατάχθηκε με την πρώτη διαδικασία
iclasses	integer*4	αριθμός κλάσεων
icolumns	integer*4	μέγεθος διανύσματος σημείων εκπαίδευσης (και κατάταξης)
idataclass(i)	integer*4 (20)	αριθμός σημείων σε κάθε κλάση
iflagclass(i)	integer*4 (iclasses)	μετρά πόσες φορές κατατάχθηκε ένα σημείο στην κάθε κλάση
ik	integer*4	μέγιστος αριθμός σημείων εκπαίδευσης
inubclass	integer*4	κλάση τελικής κατάταξης του σημείου
isignFx(i)	integer*4 (ik)	1 if Fx<0 2 if Fx>0
jk	integer*4	μέγιστο μέγεθος διανύσματος σημείου εκπαίδευσης (κατάταξη)
lk	integer*4	μέγιστος αριθμός σημείων κάθε ζεύγους κλάσεων για την εκπαίδευση
x1(i)	real*8 (icolumns)	σημείο για των υπολογισμό εσωτερικού γινομένου από kernel
x2(i)	real*8 (icolumns)	σημείο για των υπολογισμό εσωτερικού γινομένου από kernel
xtest(i)	real*8 (icolumns)	προς κατάταξη σημείο
xx2c(i,j)	real*8 (lk,jk)	όλα τα σημεία του ζεύγους των δύο κλάσεων

Όνομα	Format (διάσταση)	Περιγραφή
xxall(i,j)	real*8 (ik,jk)	όλα τα σημεία της εκπαίδευσης
yy(i)	real*8 (iall dat)	ετικέτα της κάθε κλάσης για το ζεύγος κλάσεων (1 ή -1)

Στη συνέχεια παρουσιάζεται το λογικό διάγραμμα της υπορουτίνας classifier.



Συνεχίζεται στην επόμενη σελίδα



Παράρτημα II

Π1 Το πρόγραμμα επίλυσης τετραγωνικών προβλημάτων QPROG (DQPROG)

QPROG

Solves a quadratic programming problem subject to linear equality/inequality constraints.

Required Arguments

NEQ — The number of linear equality constraints. (Input)

A — $NCON$ by $NVAR$ matrix. (Input)

The matrix contains the equality constraints in the first NEQ rows followed by the inequality constraints.

B — Vector of length $NCON$ containing right-hand sides of the linear constraints. (Input)

G — Vector of length $NVAR$ containing the coefficients of the linear term of the objective function. (Input)

H — $NVAR$ by $NVAR$ matrix containing the Hessian matrix of the objective function. (Input)

H should be symmetric positive definite; if H is not positive definite, the algorithm attempts to solve the QP problem with H replaced by a $H +$

$DIAGNL * I$ such that

$H + DIAGNL * I$ is positive definite. See Comment 3.

SOL — Vector of length $NVAR$ containing solution. (Output)

Optional Arguments

NVAR — The number of variables. (Input)

Default: $NVAR = SIZE(A,2)$.

NCON — The number of linear constraints. (Input)

Default: $NCON = SIZE(A,1)$.

LDA — Leading dimension of A exactly as specified in the dimension statement of the calling program. (Input)

Default: $LDA = SIZE(A,1)$.

LDH — Leading dimension of H exactly as specified in the dimension statement of the calling program. (Input)
Default: $LDH = SIZE(H,1)$.

DIAGNL — Scalar equal to the multiple of the identity matrix added to H to give a positive definite matrix. (Output)

NACT — Final number of active constraints. (Output)

IACT — Vector of length $NVAR$ containing the indices of the final active constraints in the first $NACT$ positions. (Output)

ALAMDA — Vector of length $NVAR$ containing the Lagrange multiplier estimates of the final active constraints in the first $NACT$ positions. (Output)

FORTRAN 90 Interface

Generic: `CALL QPROG (NEQ, A, B, G, H, SOL [, ...])`

Specific: The specific interface names are `S_QPROG` and `D_QPROG`.

FORTRAN 77 Interface

Single: `CALL QPROG (NVAR, NCON, NEQ, A, LDA, B, G, H, LDH, DIAGNL, SOL, NACT, IACT, ALAMDA)`

Double: The double precision name is `DQPROG`.

Description

The routine `QPROG` is based on M.J.D. Powell's implementation of the Goldfarb and Idnani (1983) dual quadratic programming (QP) algorithm for convex QP problems subject to general linear equality/inequality constraints, i.e., problems of the form:

$$\min_{x \in \mathbb{R}^N} g^T x + \frac{1}{2} x^T H x$$

$$\text{subject to} \quad A_1 x = b_1$$

$$A_2 x \geq b_2$$

given the vectors b_1 , b_2 , and g and the matrices H , A_1 , and A_2 . H is required to be positive definite. In this case, a unique x solves the problem or the constraints are inconsistent. If H is not positive definite, a positive definite perturbation of H is used in place of H . For more details, see Powell (1983, 1985).

Comments

1. Workspace may be explicitly provided, if desired, by use of Q2ROG/DQ2ROG. The reference is:

CALL Q2ROG (NVAR, NCON, NEQ, A, LDA, B, G, H, LDH, DIAGNL, SOL, NACT, IACT, ALAMDA, WK)

The additional argument is:

WK — Work vector of length $(3 * NVAR**2 + 11 * NVAR)/2 + NCON$.

2. Informational errors

Type Code

3 1 Due to the effect of computer rounding error, a change in the variables fail to improve the objective function value; usually the solution is close to optimum.

4 2 The system of equations is inconsistent. There is no solution.

3. If a perturbation of H , $H + \text{DIAGNL} * I$, was used in the QP problem, then

$H + \text{DIAGNL} * I$ should also be used in the definition of the Lagrange multipliers.

Example

The quadratic programming problem

$$\begin{aligned} \min f(x) &= x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - 2x_2x_3 - 2x_4x_5 - 2x_1 \\ \text{subject to } &x_1 + x_2 + x_3 + x_4 + x_5 = 5 \\ &x_3 - 2x_4 - 2x_5 = -3 \end{aligned}$$

is solved.

USE QPROG_INT

USE UMACH_INT

IMPLICIT NONE

! Declare variables

```

INTEGER  LDA, LDH, NCON, NEQ, NVAR

PARAMETER (NCON=2, NEQ=2, NVAR=5, LDA=NCON, LDH=NVAR)

!

INTEGER  K, NACT, NOUT

REAL    A(LDA,NVAR), ALAMDA(NVAR), B(NCON), G(NVAR), &
        H(LDH,LDH), SOL(NVAR)

!

!           Set values of A, B, G and H.

!           A = ( 1.0  1.0  1.0  1.0  1.0)
!
!           ( 0.0  0.0  1.0 -2.0 -2.0)
!
!
!           B = ( 5.0 -3.0)
!
!
!           G = (-2.0  0.0  0.0  0.0  0.0)
!
!
!           H = ( 2.0  0.0  0.0  0.0  0.0)
!
!           ( 0.0  2.0 -2.0  0.0  0.0)
!
!           ( 0.0 -2.0  2.0  0.0  0.0)
!
!           ( 0.0  0.0  0.0  2.0 -2.0)
!
!           ( 0.0  0.0  0.0 -2.0  2.0)
!

DATA A/1.0, 0.0, 1.0, 0.0, 1.0, 1.0, 1.0, -2.0, 1.0, -2.0/

DATA B/5.0, -3.0/

DATA G/-2.0, 4*0.0/

DATA H/2.0, 5*0.0, 2.0, -2.0, 3*0.0, -2.0, 2.0, 5*0.0, 2.0, &

```

```
-2.0, 3*0.0, -2.0, 2.0/  
!  
CALL QPROG (NEQ, A, B, G, H, SOL)  
!  
CALL UMACH (2, NOUT)  
WRITE (NOUT,99999) (SOL(K),K=1,NVAR)  
99999 FORMAT (' The solution vector is', /, ' SOL = (', 5F6.1, &  
    ' )')  
!  
END
```

Output

```
The solution vector is  
SOL = ( 1.0 1.0 1.0 1.0 1.0 )
```


Βιβλιογραφία

- Alcock A., Singh R., Ogaji S.O.T., 2003, '*Application of artificial neural networks for fault diagnostics of military turbofan engines*', ISABE paper, ISABE-2003-1209
- Anlauf J. K., Biehl M., 1989, '*The adatron: An adaptive perceptron algorithm*', Europhysics Letters, p.687–692.
- Barschdorff D., Korthauer R., 1986, '*Aspects of Failure Diagnosis on Rotating Parts of Turbomachines Using Computer Simulation and Pattern Recognition Methods*', Int. Conf. On Condition Monitoring, Brighton, England, 21-23 May 1986.
- Boser, B.E., Guyon I.M., Vapnik V., 1992, '*A training algorithm for optimal margin classifiers*', Proceedings of the Fifth Annual Workshop on Computational Learning Theory. New York, NY, USA: ACM Press, pp. 144–152.
- Burges C.J.C., 1998, '*A Tutorial on Support Vector Machines for Pattern Recognition*', Data Mining and Knowledge Discovery, Volume 2 , Issue 2 (June 1998) Pages: 121 - 167
- Cho M-Y, Lee T-F, Kau S-W., Shih C-N, Chou C-J., 2006, '*Fault Diagnosis of Power Transformers Using SVM/ANN with Clonal Selection Algorithm for Features and Kernel Parameters Selection*' Proceedings of the First International Conference on Innovative Computing, Information and Control Vol. 1 p.26 - 30
- Cook T.C, Valentine S.G., 2008, '*Shipboard mission readiness assessment*', ASME paper, GT2008-51373
- Cortes C., Vapnik V., 1995, '*Support vector networks*', Machine Learning, 20:273–297, 1995.
- Courant R., Hilbert D., 1953, '*Methods of Mathematical Physics*', Interscience
- Curnock B., 2001, '*Steady- State Test Cases for Engine Deterioration*', OBIDICOTE Programme Work Package 4, Rolls Royce report, Document Number DNS78608, May 2001

- Debnath R., Takahide N., Takahashi H., 2004, '*A decision based one-against-one method for multi-class support vector machine*', Pattern Anal Applic (2004) 7: 164–175
- Doel D.L., 1990, '*The role of expert systems in commercial gas turbine engine monitoring*', ASME paper, 90-GT-374
- Doel D.L., 1992, '*TEMPER- A gas path analysis tool for commercial jet engines*', ASME paper, 92-GT-315
- Donat W., Choi K., An W., Singh S., Pattipati K., 2007, '*Data visualization, data reduction and classifier fusion for intelligent fault detection and diagnosis in gas turbine engines*', ASME paper, GT2007-28343
- Dumais S., 1998, '*Using SVMs for text categorization*', IEEE Intelligent Systems 13 (1998), p. 21–23
- Dumais S., Platt J., Heckerman D., Sahami, M., 1998, '*Inductive learning algorithms and representations for text categorization*', Proceedings of the seventh international conference on information and knowledge management (CIKM) p. 148–155
- Fletcher R., 1987, '*Practical Methods of Optimization*', John Wiley & Sons Inc., 2nd edition, 1987
- Ge M., Du R., Zhang G., Xu Y., 2003, '*Fault diagnosis using support vector machine with an application in sheet metal stamping operations*', Mechanical Systems and Signal Processing 18 (2004) 143–159
- Ghent J., McDonald J., 2005a, '*Facial Expression Classification using a One-Against-All Support Vector Machine*', proceedings of the Irish Machine Vision and Image Processing Conference, Aug 2005
- Ghent J., McDonald J., 2005b, '*Holistic Facial Expression Classification*', SPIE Opto-Ireland, pp 5823-18, April 2005
- Ghent J., McDonald J., 2005c, '*Photo-Realistic Facial Expression Synthesis*', Image and Vision Computing Journal, Vol 23 (12), p1029-1144
- Gluch J., 1997, '*Flow and heat diagnostic analysis for a group of steam turbine stages*', 2nd Eur. Conference on Turbomachines, March 5-7, 1997
- Hsu C-W., Lin C-J., 2002, '*A comparison of methods for multiclass support vector machines*', IEEE Trans Neural Networks, 13:415–425
- Kanelopoulos K., Stamatis A., Mathioudakis K., 1997, '*Incorporating neural networks into gas turbine performance diagnostics*', ASME paper, 97-GT-35
- Kecman V., 2001, '*Learning and Soft computing, Support Vector Machines, Neural Networks and Fuzzy Logic Models*', The MIT Press. 2001

- Kim H., Park H., 2003, '*Protein secondary structure prediction based on an improved support vector machines approach*', Protein Eng. 2003 Aug;16(8):553-60
- Kurek J., Osowski S., 2009, '*Support vector machine for fault diagnosis of the broken rotor bars of squirrel-cage induction motor*', Neural Computing & Applications, Springer London
- Lee S-M, Roh T-S., Choi D-W., 2008, '*Defect diagnostics of SUAV gas turbine engine using hybrid SVM-artificial neural network method*', Journal of Mechanical Science and Technology , The Korean Society of Mechanical Engineers
- Lee Y-J., Mangasarian O.L., Wolberg W.H., 2001, '*Survival-Time Classification of Breast Cancer Patients*', Data Mining Institute Technical Report 01-03, March 2001.
- Loukis E., Mathioudakis K., Papailiou K., 1994, '*Optimizing automated gas turbine fault detection using statistical pattern recognition*', Journal of Engineering for gas turbines and power, Jan.1994, vol. 116, pp.165171
- Mathioudakis K., 2003, '*Non linear methods for gas turbine fault diagnosis*', von Karman Institute for Fluid Dynamics, Lecture Series 2003-01 'Gas Turbine Condition Monitoring & Fault Diagnosis'
- Mathioudakis K., Stamatis A., Tsalavoutas A., 2001, '*Performance analysis of industrial gas turbines for engine condition monitoring*', Proceedings Of The Institution of Mechanical Engineers, PART A, Journal of Power and Energy, vol.215, No.A2, March 2001, pp173-184
- Mathioudakis K., Kamboukos P., 2004, '*Assessment of the Effectiveness of Gas Path Diagnostic Schemes*', ASME paper, GT2004-53862
- Matuck R. G., Barbosa R. J, Bringhenti C., Lima I., 2009, '*multiple faults detection of gas turbine by mlp neural network*', ASME paper, GT2009-59964
- Merrington G., Kwon O.-K., Goodwin G., Carlsson B., 1990, '*Fault detection and diagnosis in gas turbines*', ASME paper, 90-GT-339
- Mitchell D., Kulkarni A., Roesch E., Subramanian R., Burns A., Brogan J., Greenlaw R., Lostetter A., Schupbach M., Fraley J., Waits R., 2008, '*Development and f-class industrial gas turbine engine testing of smart components with direct write embedded sensors and high temperature wireless telemetry*' ASME paper, GT2008-51533
- Poggio T., Girosi F., 1990, '*Networks for approximation and learning*', Proceedings of the IEEE, 78(9), 1481–1497
- Rangwala H., Karypis G., 2005, '*Profile based direct kernels for remote homology detection and fold recognition*', Oxford University Press
- Rausch T. R., Goebel F. K., Eklund H. N., Brunell J. B., 2005, '*Integrated in-flight fault detection and accommodation: a model-based study*', ASME paper, GT2005-68300

- Roemer M.J., 1997, '*Real-Time Health Monitoring and Diagnostics For Gas Turbine Engines*', ASME Paper 97-GT-30
- Stamatis A., Mathioudakis K., Berios K., Papailiou K.D., 1989, '*Jet Engine Fault Detection with Differential Gas Path Analysis at Discrete Operating Points*', Journal of Propulsion and Power, AIAA, Vol 7, No 6, Nov-Dec 1991, pp1043-1048
- Stamatis A., Mathioudakis K., Papailiou K. D., 1990, '*Adaptive simulation of gas turbine performance*', Journal of engineering for gas turbines and power, April 1990. vol. 112, pg. 168-175
- Stamatis A., Mathioudakis K., Ruiz J., Curnock B., 2001, '*Real Time Engine Model Implementation For Adaptive Control & Performance Monitoring Of Large Civil Turbofans*', ASME paper 2001-GT-0362.
- Tsalavoutas A., Aretakis N., Mathioudakis K., Stamatis A., 2000a, '*Combining advanced data analysis methods for the constitution of an integrated gas turbine condition monitoring and diagnostic system*', ASME paper, 2000-GT-34
- Tsalavoutas A., Mathioudakis K., Stamatis A., Smith M., 2000b, '*Identifying faults in the variable geometry system of a gas turbine compressor*', ASME paper, 2000-GT-33
- Urban L.A., Volponi A.J., 1992, '*Mathematical methods of relative engine performance diagnostics*', SAE 1992 Transactions, vol. 101, Journal of aerospace, technical paper 922048
- Vapnik V., 1979, '*Estimation of Dependences Based on Empirical Data*', Moscow: Nauka.
- Vapnik V., 1998, '*Statistical Learning Theory*', John Wiley&Sons,Inc.,1998
- Vapnik V., Chervonenkis A., 1974, '*Teoriya raspoznavaniya obrazov: Statisticheskie problemy obucheniya. (Russian) [Theory of pattern recognition: Statistical problems of learning]*', Moscow: Nauka.
- Vapnik Vladimir N., 1995, '*The Nature of Statistical Learning Theory*' Springer-Verlag New York, Inc.
- Vapnik Vladimir, 1982, '*Estimation of Dependences Based on Empirical Data*' Springer Verlag.
- Visual Numerics ,
<http://www.vni.com/products/imsl/documentation/fort06/math/NetHelp/default.htm?url=qprog.htm>

- Volponi A.J., DePold H. Ganguli R., Daguang C., 2000, *'The use of kalman filter and neural network methodologies in gas turbine performance diagnostics: a comparative study'*, proceedings of ASME TURBOEXPO 2000,2000-GT-547
- Αρετάκης Ν., 2000, *'Συμβολή στην ανάπτυξη διαγνωστικών συστημάτων σε εγκαταστάσεις αεριοστροβίλων με χρήση μεθόδων ανάλυσης μετρήσεων ταχείας απόκρισης'*, Διδακτορική Διατριβή, ΕΜΠ, Αθήνα 2000.
- Λουκής Ε., 1993, *'Συμβολή στη διαγνωστική βλαβών αεριοστροβίλων με χρήση μεθόδων ανάλυσης μετρήσεων ταχείας απόκρισης'*, Διδακτορική διατριβή, ΕΜΠ, Αθήνα 1993.
- Μαθιουδάκης Κ., 2007, *'Διαγνωστική Αεριοστροβίλων'*, ΕΜΠ, Αθήνα 2007.
- Ρωμέσης Χ., 2005, *'Χρήση στοχαστικών μεθόδων στην ανάπτυξη συστημάτων παρακολούθησης λειτουργίας και διάγνωση βλαβών αεριοστροβίλων'*, Διδακτορική Διατριβή, ΕΜΠ, Αθήνα 2005.
- Σταμάτης Α., 1990, *'Η Διαγνωστική στις Εγκαταστάσεις Αεριοστροβίλων και η Εφαρμογή της στην Προληπτική Συντήρηση'*, Διδακτορική Διατριβή, ΕΜΠ, Αθήνα 1990.
- Σταυρόπουλος Μ., 1996, *'Διάγνωση Βλαβών Στροβιλομηχανών από Μετρήσεις Ταχείας Απόκρισης με Τεχνητά Νευρωνικά Δίκτυα'*, Διπλωματική Εργασία, ΕΜΠ Αθήνα 1996.