



Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Προηγμένα Συστήματα Πληροφορικής»

Μεταπτυχιακή Διατριβή

Τίτλος Διατριβής	Υλοποίηση Υπερ-ταξινομητών πολυμεσικών δεδομένων
Όνοματεπώνυμο Φοιτητή	Κάτσαρης Δημήτριος του Ανδρέα
Αριθμός Μητρώου	ΜΠΣΜ/06054
Κατεύθυνση	Ευφυείς Τεχνολογίες Επικοινωνίας Ανθρώπου-Υπολογιστή
Επιβλέπων	Γεώργιος Τσιχριντζής, Καθηγητής

Πανεπιστήμιο Πειραιώς-Τμήμα Πληροφορικής
Πρόγραμμα Μεταπτυχιακών Σπουδών στα
Προηγμένα Συστήματα Πληροφορικής

Ημερομηνία Παράδοσης **Απρίλιος 2010**

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΑΙΑ

Τριμελής Εξεταστική Επιτροπή

(υπογραφή)

Γεώργιος Τσιχριντζής
Καθηγητής

(υπογραφή)

Δημήτριος Αποστόλου
Λέκτορας

(υπογραφή)

Χαρ. Κωνσταντόπουλος
Λέκτορας

Copyright © Δημήτριος Α. Κάτσαρης 2010.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς το συγγραφέα.

Περίληψη

Η παρούσα μεταπτυχιακή διατριβή πραγματεύεται το πρόβλημα της ταξινόμησης προτύπων τα οποία ανήκουν σε περισσότερες από δύο κλάσεις. Για το σχεδιασμό του υπέρ-ταξινομητή μελετήθηκε η μέθοδος 'one against all'. Σύμφωνα με τη μέθοδο αυτή για κάθε κλάση δεδομένων δημιουργείται ένας ταξινομητής ο οποίος εκπαιδεύεται με θετικά και αρνητικά δεδομένα. Τα θετικά δεδομένα αποτελούνται από τα στιγμιότυπα της κλάσης στην οποία αντιστοιχίζεται κάθε φορά ο ταξινομητής ενώ τα αρνητικά δεδομένα αποτελούνται από τα στιγμιότυπα όλων των υπόλοιπων κλάσεων. Αν το πλήθος των θετικών και αρνητικών στιγμιότυπων είναι ίσο τότε μιλάμε για ισοζυγισμένο ταξινομητή ενώ σε αντίθετη περίπτωση για μη-ισοζυγισμένο ταξινομητή. Μετά την εκπαίδευση όλων των επιμέρους ταξινομητών ελέγχοντας ένα δοσμένο στιγμιότυπο με όλους τους παραπάνω ταξινομητές μπορούμε να βγάλουμε συμπέρασμα σε ποια κλάση αυτό το στιγμιότυπο μπορεί να ανήκει.

Στηριζόμενοι πάνω σε αυτή τη μέθοδο, δημιουργήσαμε δικά μας υβριδικά μοντέλα τα οποία μας βοήθησαν στην τελική σχεδίαση του υπέρ-ταξινομητή μας. Η προσέγγιση του προβλήματος στηρίχθηκε στην αρχή στη χρήση του ταξινομητή μηχανών διανυσμάτων στήριξης (Support Vector Machines). Προκειμένου όμως να βγάλουμε πιο ασφαλή συμπεράσματα για τη μέθοδο 'one against all' και των υβριδικών μοντέλων που δημιουργήσαμε χρησιμοποιήσαμε και νευρωνικά δίκτυα τα οποία όμως κατά την αξιολόγηση μας έδωσαν χειρότερα αποτελέσματα.

Η αξιολόγηση των παραπάνω μεθόδων σχεδίασης του υπέρ-ταξινομητή έγινε σε τρία προβλήματα ταξινόμησης. Τα δύο αναφέρονται σε προβλήματα ταξινόμησης μουσικών κομματιών για τα οποία έχουν υπολογιστεί κατάλληλα διανύσματα χαρακτηριστικών και το τρίτο αναφέρεται στην ταξινόμηση ανθρώπινων εκφράσεων προσώπου. Τα αποτελέσματα αναδεικνύουν ότι η τελική υβριδική σχεδιαστική μέθοδος του υπέρ-ταξινομητή φαίνεται να δίνει οριακά λίγο καλύτερα αποτελέσματα από αυτά της απλής μεθόδου 'one against all'. Επίσης φαίνεται ότι σημαντικό ρόλο στα τελικά αποτελέσματα παίζει η ποιότητα του συνόλου των χαρακτηριστικών που δίνονται σαν είσοδος στον ταξινομητή.

Abstract

This graduate dissertation is dealing with the case of classification of patterns which belong to more than two classes. The design of the super-classifier was based in a method called *“one against all”*. According to this method, for each data class, there is created a classifier that is trained by using positive and negative training data population. Positive data consist of snapshots of the class, which the classifier corresponds with each time, while negative data consist of snapshots of remaining classes. If the population of positive data and the population of negative data is equal, then we have balanced classifier, otherwise, we have an unbalanced classifier. After the integrated training of all the particular classifiers, we can conclude which class a give snapshot possibly belongs to by checking this snapshot with all the above particular classifiers.

Based on this method, we created our own models that assisted us to design the final model of our super-classifier. Initially, the approach of the problem was based on the use of the classifier SVM (Support Vector Machines). However, in order to have a safer conclusion for the method *“one against all”*, we also used neural networks that unfortunately returned worse results during the evaluation process.

The evaluation of the above design methods of the super-classifier was focused on three particular classification problems. The two of them refer to classification problems of musical pieces, for which suitable characteristic vectors have been computed, and the third one refers to the classification of human face expressions. The results show that the final hybrid design method of the super-classifier seems to have slightly better performance that of the pure (original) method *“one against all”*. Also, it seems that the quality of the set of characteristics, which are given as an input to the classifier, play a considerable role in the deduction of the final results.

Ευχαριστίες

Θα ήθελα να εκφράσω την ευγνωμοσύνη μου σε όλους όσους στάθηκαν αρωγοί και συνετέλεσαν, άμεσα ή έμμεσα, στην ολοκλήρωση της συγκεκριμένης εργασίας.

Ευχαριστώ την επιβλέποντα καθηγήτή μου, κ. Γεώργιο Τσιχριντζή, Καθηγητή του τμήματος Πληροφορικής του Πανεπιστημίου Πειραιώς, που μου έδωσε την ευκαιρία να ασχοληθώ με το συγκεκριμένο θέμα

Επίσης θα ήθελα να εκφράσω ένα μεγάλο ευχαριστώ στον υποψήφιο διδάκτορα Διονύσιο Σωτηρόπουλο που επί ένα περίπου ένα χρόνο με βοήθησε να ξεπεράσω όποια εμπόδια χρειάστηκε να αντιμετωπίσω. Οι συμβουλές και η καθοδήγησή του αποτέλεσαν καθοριστικό παράγοντα για την επιτυχή ολοκλήρωση της παρούσας εργασίας.

Τέλος, θα πρέπει να σταθώ στην ηθική και ψυχολογική συμπαράσταση που μου παρείχαν η οικογένεια και οι φίλοι μου, καθώς και τις όποιες σχετικές γνώσεις τους μέχρι την ολοκλήρωση της μεταπτυχιακής διατριβής.

1	Εισαγωγή	9
1.1	Αντικείμενο της διπλωματικής εργασίας	9
1.2	Διάρθρωση της εργασίας	9
2	Θεωρητικό υπόβαθρο	10
2.1	Εισαγωγή	10
2.2	Μηχανική Μάθηση	11
2.3	Ταξινόμηση	12
2.3.1	Διανυσματική αναπαράσταση	12
2.4	Μάθηση με πλήρη επίβλεψη – χωρίς επίβλεψη	12
2.5	Μέτρα αξιολόγησης	13
2.5.1	Πίνακας σύγχυσης-Confusion matrix	13
2.5.2	Ακρίβεια	14
2.5.3	Ορθότητα	14
2.5.4	Μέτρο DecisionValue	14
3	Ταξινομητής SVM	15
3.1	Support Vector Machines	15
3.1.1	Γραμμικές SVM	16
3.1.2	Μη γραμμικές SVM	18
4	Ταξινομητής Νευρωνικών Δικτύων	20
4.1	Βιολογικά νευρωνικά δίκτυα	21
4.2	Τεχνητά Νευρωνικά Δίκτυα	23
4.3	Μοντέλα νευρώνων	24
4.3.1	Απλός νευρώνας (χωρίς bias)	24
4.3.2	Απλός νευρώνας (με bias)	24
4.3.3	Νευρώνας με είσοδο διάνυσμα	25
4.4	Συναρτήσεις μεταφοράς	25
4.5	Αρχιτεκτονικές δικτύων	26
4.5.1	Επίπεδο νευρώνων	27
4.5.2	Πολλαπλά επίπεδα νευρώνων (Multiple Layers of Neurons)	27
4.6	Perceptron	29
4.7	Back-error Propagation (BP)	30
5	Σχεδιασμός Ταξινομητή	31
5.1	Μέθοδος ‘One against all’	31
5.2	Μέθοδος ‘One against all balanced’	36
5.3	Μέθοδος ‘One against all balanced with weight’	38
5.4	Μέθοδος ‘One against all balanced with double passing’	39
6	Δεδομένα	44
6.1	Μουσικά Δεδομένα	44
6.1.1	Χαρακτηριστικά ηχητικής χροιάς	45
6.1.2	Χαρακτηριστικά ρυθμού	46
6.1.3	Χαρακτηριστικά τόνου φωνής	46
6.2	Δεδομένα εκφράσεων προσώπου	47
7	Πειραματική αξιολόγηση	49
7.1	Μέθοδος ‘One against all’	50
7.2	Μέθοδος ‘One against all balanced’	53
7.2.1	Ταξινομητής SVM	53
7.2.2	Ταξινομητής νευρωνικών δικτύων	72
7.3	Μέθοδος ‘One against all balanced with weight’	80

7.3.1 Ταξινομητής SVM.....	80
7.3.2 Ταξινομητής νευρωνικών δικτύων	82
7.4 Μέθοδος 'One against all balanced with double passing'	83
7.4.1 Ταξινομητής SVM.....	83
7.4.2 Ταξινομητής νευρωνικών δικτύων	85
8 Επίλογος.....	86
8.1 Συμπεράσματα.....	86
8.2 Μελλοντικές κατευθύνσεις	86
9 Αναφορές.....	86

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΔΑΝ

1 Εισαγωγή

Η κατασκευή μηχανών ικανών να μαθαίνουν από την εμπειρία έχει αποτελέσει για πολύ καιρό αντικείμενο τόσο φιλοσοφικών όσο και τεχνικών συζητήσεων. Η τεχνική πτυχή των συζητήσεων έχει λάβει μια τεράστια ώθηση από την εμφάνιση των ηλεκτρονικών υπολογιστών που έχουν καταδείξει ότι οι μηχανές μπορούν να επιδείξουν ένα σημαντικό επίπεδο ικανότητας μάθησης, αν και τα όρια αυτής της ικανότητας δεν έχουν ακόμα σαφώς καθοριστεί.

Η ύπαρξη αξιόπιστων συστημάτων εκμάθησης είναι αναγκαία, δεδομένου ότι υπάρχουν πολλά προβλήματα για τα οποία δεν υπάρχει διαθέσιμο κάποιο μαθηματικό πρότυπο που να τα περιγράφει, οπότε και δεν μπορούν να λυθούν από τις κλασσικές τεχνικές προγραμματισμού. Παραδείγματος χάριν δεν είναι γνωστό πώς να γράψει κάποιος ένα πρόγραμμα υπολογιστή για την αναγνώριση χειρόγραφων χαρακτήρων ή την αναγνώριση κάποιου μουσικού μοτίβου, αν και υπάρχει αφθονία διαθέσιμων παραδειγμάτων. Είναι επομένως φυσικό να ρωτήσει κανείς αν ένας υπολογιστής μπορεί να εκπαιδευτεί να αναγνωρίζει το γράμμα 'Α' από παραδείγματα, ή να αναγνωρίζει ότι το μουσικό κομμάτι 'Dust in the Wind' είναι Rock μουσικό κομμάτι. Εξάλλου αυτός δεν είναι και ο τρόπος που ο άνθρωπος μαθαίνει να διαβάζει ή να ξεχωρίζει το είδος της μουσικής που ακούει;

Ο ίδιος συλλογισμός ισχύει και σε άλλα προβλήματα: εύρεση γονιδίων σε μια ακολουθία DNA, φιλτράρισμα ηλεκτρονικού ταχυδρομείου, εντοπισμός ή αναγνώριση αντικειμένων στην όραση υπολογιστών, κ.ά.. Η επίλυση κάθε ενός από αυτά τα προβλήματα έχει τη δυνατότητα να επιφέρει επανάσταση σε κάποια πτυχή της ζωής μας, και για κάθε ένα από αυτά οι αλγόριθμοι μηχανικής μάθησης μπορούν να προσφέρουν το κλειδί για την λύση του προβλήματος.

1.1 Αντικείμενο της διπλωματικής εργασίας

Στην παρούσα εργασία επιχειρείται μια μελέτη του ήδη υπάρχοντος και υλοποιημένου με το εργαλείο Matlab ταξινομητή SVM και η δημιουργία ενός υπέρ-ταξινομητή ο οποίος θα έχει τη δυνατότητα να ταξινομήσει περισσότερες από δύο κλάσεις. Η σχεδίαση του υπέρ-ταξινομητή θα στηριχθεί πάνω στη μέθοδο 'one against all' και υβριδικών μεθόδων που μπορούν να εξαχθούν από αυτή. Για να έχουμε μία καλύτερη εικόνα της λειτουργίας του υπέρ-ταξινομητή κατά την υλοποίησή του εκτός από τον SVM θα χρησιμοποιηθούν και ταξινομητές νευρωνικών δικτύων.

Η μεθοδολογία SVM είναι αρκετά πρόσφατη, αφού προτάθηκε την προηγούμενη δεκαετία. Έχει ήδη προσελκύσει αυξημένο ενδιαφέρον, μιας και φαίνεται να έχει εντυπωσιακά αποτελέσματα σε πολλές εφαρμογές. Σε αντίθεση με άλλους αλγόριθμους μηχανικής μάθησης που στηρίζονται στην ελαχιστοποίηση του εμπειρικού σφάλματος (π.χ αριθμός σφαλμάτων εκπαίδευσης), η μάθηση με SVM βασίζεται στην αρχή της ελαχιστοποίησης του δομικού σφάλματος. Ως εκ τούτου, οι SVM τείνουν να έχουν καλύτερα αποτελέσματα όταν εφαρμόζονται σε δεδομένα που δεν ανήκουν στο σύνολο εκπαίδευσης, έχουν μεγαλύτερη ικανότητα γενίκευσης και βρίσκουν ταχύτερα την βέλτιστη λύση.

Ειδικότερα λοιπόν, θα μελετηθεί ο υπέρ-ταξινομητής πώς συμπεριφέρεται και πώς ταξινομεί δεδομένα μετά από την εκπαίδευσή του. Τα δεδομένα για τα οποία γίνεται λόγος θα ανήκουν σε δύο διαφορετικές κλάσεις και θα είναι είτε ισοζυγισμένα είτε όχι (μέθοδος 'one against all'). Με τη βοήθεια αυτών των συμπερασμάτων θα είναι δυνατή η υλοποίηση ενός ταξινομητή ο οποίος θα έχει τη δυνατότητα να ταξινομεί δεδομένα τα οποία θα ανήκουν σε περισσότερες από δύο κλάσεις. Τα δεδομένα στα οποία θα γίνουν τα πειράματα και οι έλεγχοι προέρχονται από μουσικά κομμάτια διαφορετικών μοτίβων και από αρχεία εικόνων προσώπου διαφορετικών εκφράσεων. Μετά από κατάλληλη επεξεργασία για τα μεν μουσικά δεδομένα έχουν δημιουργηθεί 1000 στιγμιότυπα όπου το κάθε στιγμιότυπο αποτελείται από 30 χαρακτηριστικά και άλλη μία σειρά από 1000 τα οποία περιέχουν 280 χαρακτηριστικά ενώ για τα δεδομένα εκφράσεων προσώπου έχουν δημιουργηθεί 2500 στιγμιότυπα όπου το κάθε στιγμιότυπο αποτελείται από 7 χαρακτηριστικά αντίστοιχα.

1.2 Διάρθρωση της εργασίας

Το υπόλοιπο της εργασίας οργανώνεται ως ακολούθως: Στο κεφάλαιο 2 παρατίθεται το θεωρητικό υπόβαθρο πάνω στο οποίο στηρίχθηκε η εργασία. Το κεφάλαιο 3 περιγράφει τον αλγόριθμο SVM που αναπτύσσεται με τη βοήθεια του εργαλείου Matlab καθώς πάνω στον οποίο θα στηριχθούμε για να αναπτυχθεί ο υπέρ-ταξινομητής. Στο κεφάλαιο 4 περιγράφονται κάποιοι νευρωνικοί ταξινομητές που θα χρησιμοποιήσουμε στην εργασία μας για να έχουμε περισσότερα αποτελέσματα πάνω στη

μεθοδολογία 'one against all' που εφαρμόζεται για την ανάπτυξη του υπέρ ταξινομητής μας και των υβριδικών μοντέλων όπως αυτά παρουσιάζονται αναλυτικά στο κεφάλαιο 5. Στο κεφάλαιο 6 γίνεται αναφορά στο περιβάλλον διεξαγωγής των πειραμάτων στο οποίο διενεργήθηκε η αξιολόγηση των μεθόδων που συζητούνται, ενώ η περιγραφή των αποτελεσμάτων που προέκυψαν από την εκτέλεση τους παρουσιάζεται στο κεφάλαιο 7. Η εργασία ολοκληρώνεται με το κεφάλαιο 8, όπου ανακεφαλαιώνονται τα ζητήματα που θίχθηκαν και προτείνονται κατευθύνσεις που θα άξιζε να διερευνηθούν μελλοντικά.

2 Θεωρητικό υπόβαθρο

2.1 Εισαγωγή

Όταν οι υπολογιστές χρησιμοποιούνται για να λύσουν ένα πρακτικό πρόβλημα, η συνήθης περίπτωση είναι να μπορεί να περιγραφεί ρητά η μέθοδος παραγωγής της απαιτούμενης εξόδου από το σύνολο δεδομένων εισόδου. Οπότε ο στόχος του σχεδιαστή του συστήματος και τελικά του προγραμματιστή για την εφαρμογή των προδιαγραφών είναι να μεταφράσει αυτή την μέθοδο σε μια ακολουθία βημάτων που ο υπολογιστής θα ακολουθήσει για να επιτύχει το επιθυμητό αποτέλεσμα.

Όταν όμως οι υπολογιστές χρησιμοποιούνται για να λύσουν πιο σύνθετα προβλήματα, μπορούν να προκύψουν καταστάσεις στις οποίες δεν υπάρχει κάποια γνωστή μέθοδος για τον υπολογισμό της επιθυμητής εξόδου από το σύνολο των δεδομένων εισόδου ή ο υπολογισμός αυτός να είναι πολύ ακριβός. Παραδείγματα τέτοιων καταστάσεων μπορεί να είναι η μοντελοποίηση μιας σύνθετης χημικής αντίδρασης, όπου οι ακριβείς αλληλεπιδράσεις των διάφορων αντιδραστηρίων δεν είναι γνωστές, η ταξινόμηση πρωτεϊνικών τύπων βασισμένων στην ακολουθία DNA από την οποία παράγονται, ή η ταξινόμηση όσους έχουν κάνει αίτηση για δάνειο σε εκείνους που θα αθετήσουν ή σε εκείνους που θα το ξεπληρώσουν.

Αυτές οι εργασίες δεν μπορούν να λυθούν από μια παραδοσιακή προγραμματιστική προσέγγιση, αφού ο σχεδιαστής του συστήματος δεν μπορεί να καθορίσει ακριβώς την μέθοδο με την οποία η σωστή έξοδος υπολογίζεται από τα δεδομένα εισόδου. Μια εναλλακτική στρατηγική για την επίλυση τέτοιων προβλημάτων είναι να προσπαθήσει ο υπολογιστής να μάθει την λειτουργία εισόδου / εξόδου από παραδείγματα, με τον ίδιο τρόπο που τα παιδιά μαθαίνουν ποια είναι τα αγωνιστικά αυτοκίνητα απλά λέγοντας τους ποια από ένα μεγάλο αριθμό αυτοκινήτων είναι αγωνιστικά, και όχι δίνοντας τους μια ακριβή προδιαγραφή του αγωνιστικού αυτοκινήτου. Η προσέγγιση της χρησιμοποίησης παραδειγμάτων για την κατασκευή ενός προγράμματος είναι γνωστή ως μεθοδολογία μάθησης (learning methodology), και στην ειδική περίπτωση που τα παραδείγματα είναι ζευγάρια εισόδου / εξόδου καλείται μάθηση με επίβλεψη (supervised learning). Τα παραδείγματα αναφέρονται ως δεδομένα εκπαίδευσης (training data).

Τα ταιριάσματα εισόδου / εξόδου συνήθως φανερώνουν μια λειτουργική σχέση που απεικονίζει τα δεδομένα εισόδου σε έξοδο, αν και αυτό δεν ισχύει πάντα όταν η έξοδος αλλοιώνεται από θόρυβο. Όταν υπάρχει μια υποκείμενη συνάρτηση από τα δεδομένα εισόδου στην έξοδο, αυτή αναφέρεται ως συνάρτηση στόχου (target function). Η προσέγγιση της συνάρτησης στόχου που μαθαίνεται ή δημιουργείται από τον αλγόριθμο μάθησης είναι γνωστή ως λύση (solution) του προβλήματος μάθησης. Στην περίπτωση της ταξινόμησης, αυτή η συνάρτηση αναφέρεται μερικές φορές ως συνάρτηση απόφασης (decision function). Η λύση επιλέγεται από ένα σύνολο υποψήφιων συναρτήσεων που απεικονίζουν το χώρο εισόδου στην περιοχή εξόδου. Συνήθως θα επιλέξουμε ένα συγκεκριμένο σύνολο ή μια κατηγορία από υποψήφιες συναρτήσεις γνωστές ως υποθέσεις (hypotheses) πριν ξεκινήσουμε να μάθουμε τη σωστή συνάρτηση. Παραδείγματος χάριν, τα αποκαλούμενα δέντρα απόφασης (decision trees) είναι υποθέσεις που δημιουργούνται κατασκευάζοντας ένα δυαδικό δέντρο με απλές συναρτήσεις απόφασης στους εσωτερικούς κόμβους και τιμές εξόδου στα φύλλα. Ως εκ τούτου, μπορούμε να δούμε την επιλογή από τον χώρο υποθέσεων (hypotheses space) ως ένα από τα βασικά συστατικά της μάθησης. Ο αλγόριθμος που παίρνει τα δεδομένα εκπαίδευσης ως είσοδο και επιλέγει μια υπόθεση από τον χώρο υποθέσεων είναι το δεύτερο σημαντικό συστατικό και αναφέρεται ως αλγόριθμος μάθησης (learning algorithm).

Στην περίπτωση της μάθησης να διακρίνει κάποιος τα αγωνιστικά αυτοκίνητα, η έξοδος είναι ένα απλά ναι ή όχι, την οποία μπορούμε να θεωρήσουμε ως δυαδική τιμή εξόδου. Για το πρόβλημα της αναγνώρισης του μοτίβου ενός μουσικού κομματιού η τιμή της εξόδου θα είναι μία από ένα πεπερασμένο αριθμό κατηγοριών, ενώ οι τιμές εξόδου για την μοντελοποίηση μιας χημικής αντίδρασης μπορεί να είναι οι συγκεντρώσεις των αντιδραστηρίων ως πραγματικοί αριθμοί.

Ένα πρόβλημα μάθησης με δυαδική τιμή εξόδου αναφέρεται ως πρόβλημα δυαδικής ταξινόμησης (binary classification). Ένα πρόβλημα με ένα πεπερασμένο αριθμό από κατηγορίες αναφέρεται ως πρόβλημα ταξινόμησης πολλαπλών κατηγοριών (multi-class classification). Ένα πρόβλημα με πραγματικούς αριθμούς ως έξοδο αναφέρεται ως παλινδρόμηση (regression).

Η μάθηση χωρίς επίβλεψη (unsupervised learning) εξετάζει την περίπτωση όπου δεν υπάρχει κάποια τιμή εξόδου και ο στόχος της μάθησης είναι να κατανοηθεί η διαδικασία που παρήγαγε τα δεδομένα. Αυτός ο τύπος μάθησης περιλαμβάνει την εκτίμηση παραμέτρων, την ομαδοποίηση, κ.α.. Υπάρχουν επίσης μοντέλα μάθησης που εξετάζουν πιο σύνθετες αλληλεπιδράσεις μεταξύ του εκπαιδευόμενου και του περιβάλλοντος τους. Ίσως η απλούστερη περίπτωση είναι όταν ο εκπαιδευόμενος μπορεί να ρωτήσει το περιβάλλον για την έξοδο που συνδέεται με μια συγκεκριμένη είσοδο. Η μελέτη για το πώς αυτό έχει επιπτώσεις στη δυνατότητα του εκπαιδευόμενου να μάθει είναι γνωστή ως μάθηση ερωτήσεων (query learning). Περαιτέρω πολυπλοκότητες στην αλληλεπίδραση εξετάζονται στην μάθηση με ενίσχυση (reinforcement learning), όπου ο εκπαιδευόμενος έχει μια σειρά από ενέργειες στην διάθεση του που μπορεί να πάρει για να προσπαθήσει να κινηθεί προς καταστάσεις όπου αναμένει υψηλές ανταμοιβές. Η μεθοδολογία μάθησης μπορεί να παίξει ρόλο στην μάθηση με ενίσχυση αν μεταχειριστούμε τη βέλτιστη ενέργεια ως έξοδο μιας συνάρτησης της τρέχουσας κατάστασης του εκπαιδευόμενου. Υπάρχουν, εντούτοις, σημαντικές περιπλοκές, αφού η ποιότητα της εξόδου μπορεί να αξιολογηθεί μόνο έμμεσα όταν εμφανιστούν οι συνέπειες μιας ενέργειας.

Άλλοι τύποι μάθησης καθορίζονται από τον τρόπο με τον οποίο τα δεδομένα εκπαίδευσης παράγονται αλλά και από το πώς παρουσιάζονται στον εκπαιδευόμενο. Παραδείγματος χάριν, υπάρχει μια διάκριση που γίνεται μεταξύ της batch μάθησης στην οποία όλα τα δεδομένα δίνονται στον εκπαιδευόμενο στην έναρξη της μάθησης, και της online μάθησης στην οποία ο εκπαιδευόμενος λαμβάνει ένα παράδειγμα τη φορά, και δίνει την εκτίμηση της εξόδου, πριν να λάβει τη σωστή της τιμή. Στην on-line μάθηση ο εκπαιδευόμενος ενημερώνει την τρέχουσα υπόθεσή του βάση του νέου παραδείγματος και η ποιότητα της μάθησης αξιολογείται από το συνολικό αριθμό λαθών που γίνονται κατά τη διάρκεια της μάθησης.

2.2 Μηχανική Μάθηση

Η Μηχανική Μάθηση (Machine Learning) έχει ως αντικείμενο την κατασκευή προγραμμάτων, ικανών να βελτιώνονται αυτόματα με την εμπειρία που αποκτούν κατά την εκτέλεση τους. Χρησιμοποιεί έννοιες από διάφορα επιστημονικά πεδία και κυρίως από τη στατιστική, την τεχνητή νοημοσύνη και τη θεωρία πληροφορίας. Ο όρος μάθηση χρησιμοποιείται εδώ με την έννοια ότι ένα πρόγραμμα μαθαίνει από την υπάρχουσα εμπειρία E σε σχέση με κάποια εργασία T που πρέπει να επιτελέσει και με κάποιο μέτρο απόδοσης P , όταν η απόδοσή του στην εργασία T , όπως μετριέται από το P , βελτιώνεται από την εμπειρία E . Για παράδειγμα, η εργασία ενός προγράμματος μπορεί να είναι η οπτική αναγνώριση χειρόγραφων λέξεων. Σε αυτή την περίπτωση, ως μέτρο αξιολόγησης μπορεί να θεωρηθεί το ποσοστό των λέξεων που αναγνωρίστηκαν σωστά από ένα σύνολο δοκιμής και η εμπειρία θα είναι ένα σύνολο με χειρόγραφες λέξεις που είναι χαρακτηρισμένες εκ των προτέρων. Αυτό το σύνολο δεδομένων αξιοποιείται στην εκπαίδευση του προγράμματος.

Βασική αρχή λειτουργίας ενός αλγορίθμου μηχανική μάθησης είναι η προσπάθεια εξαγωγής γενικευμένων κανόνων, ικανών να αντιμετωπίσουν το προς επίλυση πρόβλημα από ένα περιορισμένο σύνολο δεδομένων εκπαίδευσης, προσέγγιση δηλαδή της συνάρτησης (έννοιας) στόχου από το σύνολο των πιθανών λύσεων, τις υποθέσεις. Όμως, η μόνη διαθέσιμη πληροφορία που υπάρχει είναι το πεπερασμένο δείγμα εκπαίδευσης. Συνεπώς, πρόκειται για μία διαδικασία επαγωγικού συλλογισμού (inductive reasoning), δηλαδή συμπερασμού μίας γενικότερης έννοιας από την ειδικότερη έννοια που αντιστοιχεί στο συγκεκριμένο δείγμα. Επομένως, το μόνο που είναι εγγυημένο είναι ότι η προσέγγιση της συνάρτησης στόχου είναι ακριβής για τα δεδομένα εκπαίδευσης, ενώ για όλα τα άλλα στιγμιότυπα μπορούμε μόνο να το υποθέσουμε. Σε αυτό το ζήτημα αναφέρεται η θεμελιώδης παραδοχή της επαγωγικής μάθησης:

Αξίωμα: Οποιαδήποτε υπόθεση έχει βρεθεί να προσεγγίζει καλά την συνάρτηση στόχου σε ένα επαρκώς μεγάλο σύνολο παραδειγμάτων εκπαίδευσης, θα την προσεγγίζει καλά και σε άγνωστα στιγμιότυπα.

Παρόλα αυτά, τα δεδομένα εκπαίδευσης δεν είναι επαρκή. Έχει δειχθεί ότι, για να μπορέσει ένας αλγόριθμος μάθησης να κατατάξει άγνωστα στιγμιότυπα με κάποια λογική βάση, πρέπει να έχουν γίνει εκ των προτέρων κάποιες παραδοχές ως προς τη φύση της συνάρτησης στόχου. Με αυτό τον τρόπο περιορίζεται ο χώρος των υποθέσεων στον οποίο αναζητείται η συνάρτηση στόχου και δίνεται μια λογική κατεύθυνση στη απαραίτητη γενίκευση πέρα από τα παραδείγματα εκπαίδευσης, έτσι ώστε

να συναχθεί η συνάρτηση στόχου. Οι παραδοχές αυτές είναι γενικά διαφορετικές για κάθε αλγόριθμο μάθησης και δεν είναι απαραίτητο να γνωρίζουμε αν όντως ισχύουν. Το σύνολο των παραδοχών-ισχυρισμών για κάποιον αλγόριθμο που αυτός ακολουθεί για να πλοηγηθεί στο χώρο υποθέσεων, καλείται επαγωγική προδιάθεση (inductive bias) του αλγορίθμου. Πιο τυπικά, η επαγωγική προδιάθεση ενός αλγορίθμου L είναι οποιοδήποτε ελάχιστο σύνολο παραδοχών B , τέτοιο ώστε για κάθε συνάρτηση στόχο C και αντίστοιχα παραδείγματα εκπαίδευσης D_c , η δυαδική τιμή κατάταξης που ανατίθεται σε κάθε στιγμιότυπο x_i , $L(x_i, D_c)$ να συνεπάγεται λογικά από τα B , D_c και x_i :

$$\forall x_i [(B \wedge D_c \wedge x_i) \rightarrow L(x_i, D_c)]$$

Με αυτό τον τρόπο επιτυγχάνεται η ισοδυναμία μιας διαδικασίας επαγωγικής μάθησης με μία παραγωγική (deductive) διαδικασία, δηλαδή με μια διαδικασία που με τυπικά ορθό τρόπο συμπεραίνει ειδικότερες έννοιες από γενικότερες.

2.3 Ταξινόμηση

Το πρόβλημα της ταξινόμησης (classification) συνιστάται στην εκμάθηση μιας συνάρτησης στόχου ικανής να αντιστοιχεί άγνωστα αντικείμενα σε προκαθορισμένο σύνολο κατηγοριών. Στόχος της διαδικασίας μάθησης είναι η δημιουργία ενός μοντέλου πρόβλεψης της μη παρατηρούμενης ιδιότητας που είναι το διακριτικό (όνομα / τιμή) της κατηγορίας ενός αγνώστου αντικειμένου, βάσει των τιμών των χαρακτηριστικών που το προσδιορίζουν. Τυπικά παραδείγματα εφαρμογών ταξινόμησης αποτελούν η αυτόματη ταξινόμηση μηνυμάτων ηλεκτρονικού ταχυδρομείου σε κατηγορίες που καθορίζονται από τον χρήστη, ο διαχωρισμός καλοσηθών από κακοήθεις όγκους μέσω της παρατήρησης των αποτελεσμάτων ιατρικών εξετάσεων, κ.ά.. Για την αντιμετώπιση του συγκεκριμένου προβλήματος προϋποτίθεται η συγκέντρωση ενός αριθμού από αντιπροσωπευτικά για κάθε κατηγορία παραδείγματα (στιγμιότυπα του προβλήματος), τα οποία έχουν ταξινομηθεί από τον άνθρωπο-εκπαιδευτή. Στην συνέχεια, τα εν λόγω στιγμιότυπα (σώμα δεδομένων εκπαίδευσης) δίνονται σε έναν αλγόριθμο μηχανικής μάθησης, ο οποίος, κατόπιν της παρατήρησης των χαρακτηριστικών τους, παράγει ως έξοδο ένα εκτιμώμενο μοντέλο, ικανό να αντιστοιχίζει άγνωστα στιγμιότυπα στις προδιαγεγραμμένες κατηγορίες. Το τελικό προϊόν της εκπαίδευσης καλείται ταξινομητής (classifier).

2.3.1 Διανυσματική αναπαράσταση

Για να είναι σε θέση ο αλγόριθμος μηχανικής μάθησης να χειρισθεί τα δεδομένα που δίνονται ως είσοδος, απαιτείται ο μετασχηματισμός τους σε μια μορφή που να τον διευκολύνει στο έργο του. Μια από τις δημοφιλέστερες μορφές αναπαράστασης είναι η διανυσματική. Κάθε παράδειγμα εκπαίδευσης αναπαριστάται ως ένα διάνυσμα l χαρακτηριστικών (attributes/features), μετρήσιμων δηλαδή ποσοτήτων που έχουν επιλεγεί με τέτοιο τρόπο ώστε ο αλγόριθμος να εκμαιεύσει από τις τιμές τους τη γνώση που χρειάζεται για την εκμάθηση της συνάρτησης στόχου. Ανάλογα με το είδος της πληροφορίας που κωδικοποιεί ένα χαρακτηριστικό, αυτό μπορεί να είναι συνεχές (continuous) – π.χ. ένας πραγματικός αριθμός ή ονομαστικό (nominal) – ένα σύνολο διακριτών τιμών. Με αυτόν τον τρόπο, τα δεδομένα του προβλήματος απεικονίζονται σε έναν χώρο l διαστάσεων, το χώρο των στιγμιότυπων (instance space), με κάθε στιγμιότυπο (ήτοι διάνυσμα l χαρακτηριστικών) να αποτελεί ένα σημείο του χώρου. Στην περίπτωση του προβλήματος της ταξινόμησης, η κλάση που συνοδεύει τα στιγμιότυπα εκπαίδευσης κωδικοποιείται κι αυτή ως ονομαστικό χαρακτηριστικό.

Στην περίπτωση που τα παραδείγματα είναι κείμενα, αυτά αναπαρίστανται ως σύνολα λέξεων (Bag of Words), δηλαδή διανύσματα αριθμητικών χαρακτηριστικών. Καθένα από αυτά αριθμεί το πλήθος των εμφανίσεων στο εκάστοτε κείμενο, της λέξης που αντιστοιχεί στη θέση του χαρακτηριστικού μέσα στο διάνυσμα. Το σύνολο όλων των λέξεων που εμφανίζονται στα κείμενα ονομάζεται λεξικό και συμβολίζεται με $V = \{w_1, \dots, w_{|V|}\}$.

2.4 Μάθηση με πλήρη επίβλεψη – χωρίς επίβλεψη

Η διαδικασία της μάθησης υπό πλήρη επίβλεψη στοχεύει στη δημιουργία ενός μοντέλου, ικανό να προβλέπει κάποιες μη παρατηρούμενες ιδιότητες σε άγνωστα αντικείμενα, χρησιμοποιώντας στο σώμα δεδομένων εκπαίδευσης παραδείγματα που έχουν αυτές τις ιδιότητες γνωστές. Ως παραδείγματα αυτής της κατηγορίας μάθησης θα μπορούσαμε να επικαλεστούμε την εκτίμηση του βάρους ενός ανθρώπου δεδομένου του ύψους και άλλων χαρακτηριστικών του, ή του θέματος ενός κειμένου βάσει των συμβολοσειρών (λέξεων, αριθμών και σημείων στίξης) που το απαρτίζουν.

Μια διαφορετική από την προηγούμενη προσέγγιση είναι η μάθηση χωρίς επίβλεψη, στην οποία τα παραδείγματα του σώματος δεδομένων εκπαίδευσης δεν έχουν γνωστή την τιμή της ιδιότητας που θέλουμε να προβλέψουμε, οπότε ο «μαθητής» αφήνεται ελεύθερος. Από τις σημαντικότερες υπο-περιοχές της μάθησης χωρίς επίβλεψη, ξεχωρίζουν η ομαδοποίηση (clustering), και η εκτίμηση παραμέτρων (parameter estimation). Η ομαδοποίηση συνιστάται στην ανακάλυψη ομάδων από αντικείμενα που παρουσιάζουν μεταξύ τους κοινά χαρακτηριστικά. Οι κατηγορίες στην περίπτωση αυτή δεν είναι γνωστές εκ των προτέρων (δεν δίνονται δηλαδή από τον χρήστη) αλλά προκύπτουν δυναμικά κατά την εκτέλεση του αλγορίθμου ομαδοποίησης. Στην εκτίμηση παραμέτρων κατασκευάζουμε ένα στατιστικό μοντέλο του κόσμου του προς αντιμετώπιση προβλήματος, που διαθέτει ένα σύνολο παραμέτρων, τις τιμές των οποίων επιθυμούμε να προσδιορίσουμε. Για το σκοπό αυτό χρησιμοποιούμε έναν αλγόριθμο μάθησης που εκπαιδεύεται σε αντιπροσωπευτικά δεδομένα του κόσμου μας, προκειμένου να προσεγγίσει τις ζητούμενες παραμέτρους.

Όπως είδαμε, τα παραδείγματα του σώματος δεδομένων εκπαίδευσης πρέπει όλα να έχουν γνωστή την τιμή της ιδιότητας. Για το πρόβλημα της ταξινόμησης αυτό σημαίνει να είναι ταξινομημένα, να γνωρίζουμε δηλαδή σε ποια κλάση ανήκουν. Σε αυτή την κατηγορία ανήκει και ο αλγόριθμος ταξινόμησης SVM (Support Vector Machines – SVM) τον οποίο θα αναλύσουμε στο αντίστοιχο κεφάλαιο.

Η κατασκευή ενός αντιπροσωπευτικού σώματος εκπαίδευσης, παρ' όλης ωστόσο της κρισιμότητας της, αποδεικνύεται τις περισσότερες φορές ιδιαίτερα επίπονη και χρονοβόρα διαδικασία. Συχνά, οι πηγές άντλησης ποιοτικών παραδειγμάτων από τον πραγματικό κόσμο είναι περιορισμένες, καθιστώντας επιτακτική την ανάγκη αξιοποίησης τους με τον καλύτερο δυνατό τρόπο. Εξίσου σημαντικός παράγοντας είναι το κόστος απόκτησης των δεδομένων. Σε ορισμένα πεδία, όπως αυτό της μοντελοποίησης βιομηχανικών διαδικασιών, ένα παράδειγμα εκπαίδευσης μπορεί να χρειασθεί αρκετές μέρες για να παραχθεί, και μάλιστα με διόλου ευκαταφρόνητο κόστος. Σε άλλα πεδία, όπου το κόστος απόκτησης των δεδομένων είναι ουσιαστικά μηδαμινό (π.χ. φιλτράρισμα ανεπιθύμητης αλληλογραφίας), απαιτείται από τον άνθρωπο να αφιερώσει αρκετές ώρες για την αρχική ταξινόμηση των δεδομένων εκπαίδευσης.

Στα ταξινομημένα παραδείγματα πολλές φορές υπάρχει θόρυβος με τη μορφή παραδειγμάτων που περιέχουν ελλιπείς ή εσφαλμένες τιμές στα χαρακτηριστικά τους. Το γεγονός αυτό αποτελεί ιδιαίτερα σοβαρό κίνδυνο για την ικανότητα γενίκευσης του επαγόμενου μοντέλου, καθιστώντας αναγκαία τη χρήση μεθόδων «καθαρισμού» του σώματος εκπαίδευσης, για τον περιορισμό των όποιων πιθανών επιπτώσεων. Επιπλέον, σημαντικό είναι και το πρόβλημα των άσχετων δεδομένων (outliers). Πρόκειται για δεδομένα που δεν πρόσκεινται ούτε των θετικών ούτε και των αρνητικών παραδειγμάτων του προβλήματος, με αποτέλεσμα να μην είναι δυνατόν από αυτά να εξαχθούν ουσιαστικά συμπεράσματα για το οπότε σύνολο των περιπτώσεων που αντιπροσωπεύουν, σε σχέση με τις περιπτώσεις που μας ενδιαφέρουν άμεσα. Αρκετοί ταξινομητές αποπροσανατολίζονται από την παρουσία τέτοιων παραδειγμάτων, καταλήγοντας στην εκμάθηση ασήμαντων λεπτομερειών των δεδομένων αυτών. Το φαινόμενο αυτό γνωστό και ως φαινόμενο του υπερταυριάσματος (overfitting effect), πλήττει την ικανότητα γενίκευσης των επαγόμενων μοντέλων, και κατά συνέπεια την αποτελεσματικά τους κατά την αντιμετώπιση αγνώστων παραδειγμάτων.

2.5 Μέτρα αξιολόγησης

2.5.1 Πίνακας σύγχυσης-Confusion matrix

	Positive Class	Negative Class
Classified as Positive	TP	FP(εII)
Classified as Negative	FN(εI)	TN

Πίνακας: Πίνακας σύγχυσης ενός προβλήματος ταξινόμησης δύο κλάσεων.

Στον παραπάνω πίνακα εμφανίζεται ο πίνακας σύγχυσης (confusion matrix) που αναπαριστά τις τέσσερις δυνατές εκβάσεις μιας προσπάθειας δυαδικής ταξινόμησης. Στην κύρια διαγώνιο του πίνακα οι συμβολισμοί TP (True Positive) και TN (True Negative) αντιστοιχούν στις περιπτώσεις ορθής

ταξινόμησης ενός θετικού παραδείγματος στην κλάση των θετικών και ενός αρνητικού στην κλάση των αρνητικών. Οι άλλες δύο περιπτώσεις αναφέρονται σε εσφαλμένες ταξινομήσεις, που συμμετέχουν στον υπολογισμό του συνολικού σφάλματος του ταξινομητή. Η περίπτωση FN (False Negative), που αντιστοιχεί στην εσφαλμένη ταξινόμηση ενός θετικού παραδείγματος ως αρνητικό, αναφέρεται στη βιβλιογραφία ως Σφάλμα Τύπου I (Error Type I ή eI). Η αντίθετη περίπτωση FP (False Positive) αποτελεί το Σφάλμα Τύπου II – eII. Ας σημειωθεί επίσης ότι το άθροισμα TP + FN ισούται με το σύνολο των θετικών παραδειγμάτων που χρησιμοποιήθηκαν κατά τον έλεγχο του ταξινομητή, και κατ' επέκταση το FP + TN με το πλήθος των αρνητικών που είτε ταξινομήθηκαν σωστά ή εσφαλμένα.

2.5.2 Ακρίβεια

Ως ακρίβεια (accuracy) ορίζουμε την αναλογία όλων των προβλέψεων που ήταν σωστές:

$$accuracy = \frac{\# \text{σωστών προβλέψεων}}{\# \text{προβλέψεων}}$$

Για ένα πρόβλημα ταξινόμησης δύο κλάσεων, η ποσότητα εκφράζεται με την βοήθεια του πλήθους των σωστών και εσφαλμένων ταξινομήσεων ενός συστήματος ως εξής:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2.5.3 Ορθότητα

Ως ορθότητα (precision – degree of soundness) ορίζουμε τη δεσμευμένη πιθανότητα αν ταυτίζεται η κλάση που προβλέπει ένας ταξινομητής για ένα στιγμιότυπο με την πραγματική του κλάση:

$$precision_c = \frac{\# \text{σωστών προβλέψεων κλάσης } c}{\# \text{προβλέψεων κλάσης } c}$$

Για ένα πρόβλημα ταξινόμησης δύο κλάσεων έχουμε:

$$precision_p = \frac{TP}{TP + FP}$$

$$precision_N = \frac{TN}{TN + FN}$$

2.5.4 Μέτρο DecisionValue

Στα προβλήματα μάθησης όπου υπάρχουν αρνητικά και θετικά παραδείγματα, μπορούμε να υπολογίσουμε την τιμή της συνάρτησης απόφασης. Αυτή παίρνει τιμές από -1 έως 1. Ανάλογα με το πόσο σίγουρος είναι ο ταξινομητής στην τοποθέτηση ενός στιγμιότυπου στην κλάση A ή στην συμπληρωματική της θα δώσει κ την κατάλληλη τιμή σε αυτή τη συνάρτηση. Έτσι λοιπόν για μεγάλη βεβαιότητα σωστής τοποθέτησης θα δώσει τιμές κοντά στο ένα ενώ αντίθετα για μικρή βεβαιότητα θα δώσει τιμές κοντά στο -1.

Είναι δυνατόν ο ταξινομητής να δώσει θετική τιμή στη συνάρτηση απόφασης αν έχει βεβαιότητα ότι ένα στιγμιότυπο ανήκει στην συμπληρωματική και όχι στην κλάση A. Τότε όμως το rating για το συγκεκριμένο στιγμιότυπο θα είναι μηδέν γιατί θα έχει προβλεφθεί σε λάθος μεριά το στιγμιότυπο ανεξάρτητα από τη βεβαιότητα που έχει ο ταξινομητής για το αντίστοιχο συγκεκριμένο στιγμιότυπο.

3 Ταξινόμητης SVM

Στόχος της συγκεκριμένης εργασίας ήταν η δημιουργία ενός υπέρ-ταξινομητή ο οποίος θα είναι σε θέση να μπορεί να ταξινομήσει στιγμιότυπα σε περισσότερες από δύο κλάσεις. Στην μελέτη μας, προκειμένου να πετύχουμε αυτό το στόχο χρησιμοποιήσαμε ένα σχετικά καινούριο αλγόριθμο μηχανικής μάθησης, τις support vector machines (SVM). Οι SVM προτάθηκαν τη δεκαετία του '90, ωστόσο η χρήση τους είναι ήδη ιδιαίτερα διαδεδομένη σε πολλές εφαρμογές, μιας και τα αποτελέσματα που παρέχουν είναι πολύ ικανοποιητικά. Η έρευνα γύρω από τα SVM ακόμη συνεχίζεται και νέα δεδομένα προκύπτουν, με αποτέλεσμα οι εφαρμογές τους συνεχώς να επεκτείνονται. Στο παρόν κεφάλαιο θα συζητηθούν οι βασικές αρχές μηχανικής μάθησης και θα περιγραφεί αναλυτικά ο αλγόριθμος των SVM.

Μπορούμε να διακρίνουμε γενικά τρεις τρόπους μάθησης: με επίβλεψη, χωρίς επίβλεψη και την ενισχυτική μάθηση. Η καθεμία λειτουργεί με διαφορετικό τρόπο και η επιλογή τους εξαρτάται σε μεγάλο βαθμό από τη φύση του προβλήματος που θέλουμε να αντιμετωπίσουμε. Οι αλγόριθμοι μηχανικής μάθησης βρίσκουν εφαρμογή σε πολλά προβλήματα, όπως είναι η προσέγγιση συναρτήσεων, η αναγνώριση συστημάτων, η πρόβλεψη σημάτων, η ταξινόμηση προτύπων (στην οποία εντάσσεται το θέμα που διαπραγματευόμαστε στη συγκεκριμένη εργασία), προβλήματα αυτόματου ελέγχου, προβλήματα ανάλυσης δεδομένων σε κύριες συνιστώσες κ.α.

Τα SVM προσφέρουν τη δυνατότητα μάθησης με επίβλεψη, δηλαδή η μάθηση και η εκπαίδευση γίνεται με αυτόνομο τρόπο. Το σύστημα τροφοδοτείται με εισόδους που η επιθυμητή έξοδος τους διακρίνεται σε δύο κατηγορίες (δυναμικό πρόβλημα) και καλείται να εκτιμήσει τη σχέση που υπάρχει ανάμεσα στις εισόδους και τις επιθυμητές εξόδους, ώστε να μπορεί να ταξινομήσει στη συνέχεια και άλλα πρότυπα, των οποίων η έξοδος είναι άγνωστη. Η συνάρτηση που δείχνει τη σχέση εισόδων-εξόδων καλείται συνάρτηση στόχος, ενώ όταν πρόκειται συγκεκριμένα για προβλήματα ταξινόμησης ονομάζεται *συνάρτηση απόφασης*. Ιδιαίτερα σημαντική είναι η αξιολόγηση της συνάρτησης απόφασης, ή με άλλα λόγια η ικανότητα του αλγόριθμου μάθησης στη σωστή ταξινόμηση των προτύπων.

Άλλοι αλγόριθμοι, όπως τα νευρωνικά δίκτυα που είναι ευρύτερα γνωστά και χρησιμοποιούνται σε ποικίλες εφαρμογές, στηρίζονται στην ελαχιστοποίηση των σφαλμάτων κατά τη διάρκεια της εκπαίδευσης, το οποίο είναι το λεγόμενο εμπειρικό σφάλμα (empirical risk). Όσο μικρότερο το εμπειρικό σφάλμα, τόσο καλύτερη θεωρείται η απόδοση του αλγόριθμου. Ωστόσο, συχνά παρατηρείται το φαινόμενο ο αλγόριθμος να έχει ακριβή αποτελέσματα στα δεδομένα του συνόλου εκπαίδευσης, αλλά οι προβλέψεις για άγνωστα πρότυπα να είναι ασυσχέτιστες. Απαιτείται συνεπώς από έναν αλγόριθμο όσο το δυνατόν καλύτερη *ικανότητα γενίκευσης*, που είναι η ικανότητα να φέρει ικανοποιητικά αποτελέσματα όχι μόνο στο σύνολο εκπαίδευσης, αλλά να λειτουργεί καλά και με άγνωστα πρότυπα.

Πράγματι, οι SVM παρουσιάζουν καλύτερη ικανότητα γενίκευσης, αφού δεν στηρίζονται στην ελαχιστοποίηση του εμπειρικού σφάλματος αλλά αντίθετα τείνουν να ελαχιστοποιήσουν το γενικευμένο σφάλμα, ή αλλιώς δομικό σφάλμα (structural risk) για το οποίο θα συζητήσουμε στη συνέχεια. Προς το παρόν, προχωράμε στην ανάλυση της μεθοδολογίας των SVM.

3.1 Support Vector Machines

Οι Support Vector Machines είναι μια τεχνική μηχανικής μάθησης που χρησιμοποιείται για την ταξινόμηση και αναγνώριση προτύπων. Συγκεκριμένα, πρόκειται για δυναμικό πρόβλημα, αφού η ταξινόμηση των διανυσμάτων (προτύπων) γίνεται σε δύο κατηγορίες. Η μεθοδολογία SVM στηρίζεται σε βασικές γεωμετρικές έννοιες, μιας και βασίζεται στην εύρεση ενός υπερεπιπέδου που θα διαχωρίζει τα διανύσματα εισόδου με όσο το δυνατόν λιγότερα λάθη.

Θα κατηγοριοποιήσουμε τα SVM σε δύο κατηγορίες: τις γραμμικές και τις μη γραμμικές SVM. Παρόλο που η δεύτερη κατηγορία είναι η πιο γενική, θα ξεκινήσουμε με την ανάλυση της πρώτης που είναι απλούστερη και θα βοηθήσει τον αναγνώστη να καταλάβει τις βασικές αρχές που διέπουν την μεθοδολογία. Επειδή η τεχνική είναι όμοια και για τις δύο κατηγορίες, τελικά καταλήγουμε σε κάθε περίπτωση στην ανάγκη επίλυσης ενός προβλήματος βελτιστοποίησης, το οποίο θα συζητηθεί στην τελευταία παράγραφο αυτού του κεφαλαίου.

3.1.1 Γραμμικές SVM

Ξεκινάμε με την απλούστερη περίπτωση, όπου τα δεδομένα του συνόλου εκπαίδευσης είναι γραμμικά διαχωρίσιμα. Ας θεωρήσουμε λοιπόν ότι το σύνολο εκπαίδευσης είναι το $S = \{\{x_i, y_i\}, i=1,2,\dots,m\}$, όπου $x \in R^n$ είναι το διάνυσμα εισόδου και η μεταβλητή y δείχνει την κλάση (τάξη) του αντίστοιχου διανύσματος με $y \in \{-1,1\}$. Σκοπός μας είναι να βρούμε μια συνάρτηση $f(x)$ η οποία θα ταξινομεί σωστά το διάνυσμα εισόδου x .

Αφού τα δεδομένα είναι γραμμικά διαχωρίσιμα στον χώρο R^n , θα υπάρχει ένα υπερεπίπεδο που θα τα διαχωρίζει αποτελεσματικά, δηλαδή διανύσματα ίδιας κλάσης θα βρίσκονται στην ίδια πλευρά του υπερεπιπέδου. Σύμφωνα με βασικές γνώσεις Γραμμικής Άλγεβρας, η εξίσωση του υπερεπιπέδου αυτού είναι η:

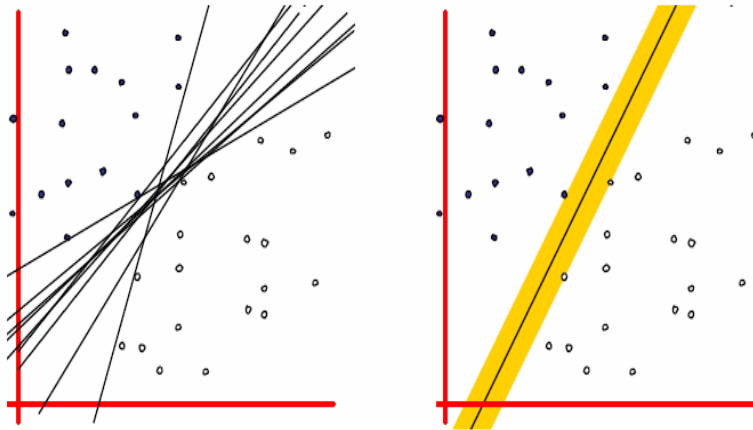
$$w \cdot x + b = 0 \quad (1)$$

όπου $w \in R^n$ είναι ένα διάνυσμα κάθετο στο υπερεπίπεδο και $b \in R$. Μπορούμε τότε να χρησιμοποιήσουμε την συνάρτηση:

$$f(x) = \text{sgn}(w \cdot x + b) \quad (2)$$

για να ταξινομήσουμε ένα πρότυπο x που δεν ανήκει απαραίτητα στο σύνολο εκπαίδευσης.

Είναι προφανές ότι υπάρχει απεριόριστος αριθμός τέτοιων επιπέδων που μπορούν να διαχωρίσουν αποτελεσματικά τα διανύσματα του συνόλου εκπαίδευσης. Η μέθοδος SVM προσπαθεί να υπολογίσει το επίπεδο που επιτυγχάνει μέγιστο εύρος κατά τον διαχωρισμό των διανυσμάτων εισόδου, όπως φαίνεται στο παρακάτω σχήμα. Ο λόγος για τον οποίο γίνεται αυτό είναι ότι αφ' ενός οπτικά φαίνεται ασφαλέστερο, αφ' ετέρου ελαχιστοποιείται με αυτόν τον τρόπο η πιθανότητα λάθους (κακής ταξινόμησης).



Σχήμα: (Αριστερά) Πιθανά επίπεδα διαχωρισμού, (Δεξιά) Βέλτιστο υπερεπίπεδο μέγιστου εύρους

Για τον υπολογισμό του επιπέδου μέγιστου εύρους αρκεί η ελαχιστοποίηση της συνάρτησης:

$$\frac{1}{2} \cdot \|w\|^2 \quad (3)$$

όπου $\|w\|$ είναι η Ευκλείδεια νόρμα του διανύσματος w , τηρώντας όμως τις συνθήκες:

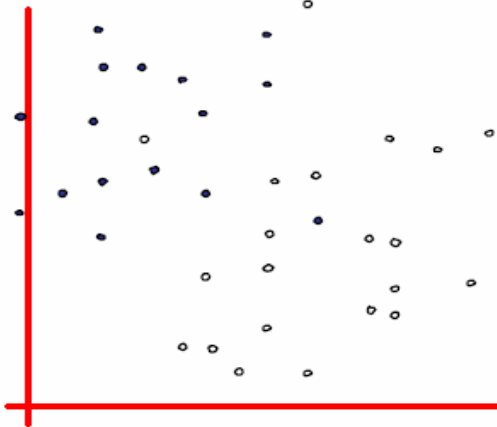
$$y_i (w \cdot x_i + b) - 1 \geq 0, i=1..l \quad (4)$$

Η εύρεση λοιπόν του επιπέδου με μέγιστο εύρος προϋποθέτει την επίλυση ενός προβλήματος δευτεροβάθμιου προγραμματισμού, με συνάρτηση κόστους την (3) και περιορισμούς τις σχέσεις (4). Το γεωμετρικό εύρος που επιτυγχάνεται μετά την επίλυση του παραπάνω προβλήματος ισούται με:

$$\gamma = 1 / \|w\| \quad (5)$$

Ωστόσο, σπάνια στην πράξη συναντάμε εφαρμογές όπου τα δεδομένα εισόδου είναι γραμμικώς διαχωρίσιμα. Στη συνηθέστερη περίπτωση δεν υπάρχει υπερεπίπεδο που να καταφέρνει τον

ακριβή διαχωρισμό των δεδομένων, όπως δείχνει το παρακάτω σχήμα. Το σκεπτικό που αναλύθηκε προηγουμένως στηρίζεται στην εύρεση ενός επιπέδου με «σκληρά» όρια, προϋποθέτει δηλαδή τον διαχωρισμό των δεδομένων χωρίς κανένα λάθος στην εκπαίδευση, με αποτέλεσμα η επίλυση του προηγούμενου προβλήματος να μην είναι δυνατή σε τέτοιου είδους σύνολα. Προκειμένου λοιπόν να αντιμετωπιστεί το συγκεκριμένο πρόβλημα πρέπει να επεκτείνουμε το προηγούμενο σκεπτικό με τέτοιο τρόπο ώστε να είναι πλέον επιτρεπτά τα λάθη στην εκπαίδευση, να γίνουν δηλαδή «διάτρητα» τα όρια του υπερεπιπέδου.



Σχήμα: Πρόβλημα γραμμικών μη διαχωρίσιμων δεδομένων

Προς την κατεύθυνση αυτή εισάγουμε στους περιορισμούς (4) μία μεταβλητή ξ_i η οποία δείχνει την απόσταση του σημείου από το υπερεπίπεδο σε περίπτωση λάθους. Προφανώς, $\xi_i \geq 0$ και μάλιστα σε περίπτωση σωστής ταξινόμησης $\xi_i = 0$. Με την εισαγωγή των συγκεκριμένων μεταβλητών πρέπει πλέον να αλλάξουμε και την προς ελαχιστοποίηση συνάρτηση. Το άθροισμα $\sum \xi_i$ αποτελεί ένα εύλογο όριο για τον αριθμό των σφαλμάτων εκπαίδευσης. Με τις αλλαγές που έγιναν, το πρόβλημα ελαχιστοποίησης διατυπώνεται πλέον ως εξής:

$$\text{ελαχιστοποίηση} \quad \frac{1}{2} \cdot \|\mathbf{w}\|^2 + C \sum \xi_i \quad (6)$$

$$\text{με περιορισμούς} \quad y_i (\mathbf{w} \cdot \mathbf{x}_i + b) - 1 + \xi_i \geq 0, i=1..l, \xi_i \geq 0 \quad (7)$$

Είναι σημαντικό εδώ να κάνουμε κάποιες παρατηρήσεις για τη συνάρτηση ελαχιστοποίησης. Πρόκειται για το *δομικό σφάλμα* (structural risk) που αναφέραμε ήδη στην προηγούμενη ενότητα αυτού του κεφαλαίου. Αποτελείται από δύο όρους: ο πρώτος όρος δείχνει την πολυπλοκότητα του μοντέλου, ενώ ο δεύτερος όρος είναι το εμπειρικό σφάλμα. Η παράμετρος C επιλέγεται από τον χρήστη για την φάση της εκπαίδευσης. Μία μικρή τιμή της παραμέτρου αναγκάζει τον αλγόριθμο για την εύρεση του υπερεπιπέδου με βέλτιστο εύρος να μεγιστοποιήσει την απόσταση $1/\|\mathbf{w}\|$, το εύρος δηλαδή, ‘αδιαφορώντας’ για τα σφάλματα εκπαίδευσης. Αντίθετα, μία μεγάλη τιμή της παραμέτρου προκαλεί την τάση για ελαχιστοποίηση του εμπειρικού σφάλματος, εις βάρος βέβαια της μεγιστοποίησης του εύρους.

Και για τις δύο περιπτώσεις που εξετάσαμε, γραμμικά διαχωρίσιμα και γραμμικά μη διαχωρίσιμα δεδομένα, στο τέλος έχουμε να αντιμετωπίσουμε ένα πρόβλημα ελαχιστοποίησης συνάρτησης, συνοδευόμενο από μια σειρά περιορισμών. Για την επίλυση αυτού του προβλήματος απαιτείται να εργαστούμε με διανύσματα μεγάλης διάστασης. Προκειμένου να αποφύγουμε αυτούς τους δύσκολους και χρονοβόρους υπολογισμούς, θα χρησιμοποιήσουμε την θεωρία Lagrange για να απλοποιήσουμε το πρόβλημα βελτιστοποίησης. Σύμφωνα με την θεωρία Lagrange, πολλαπλασιάζουμε την κάθε ανισότητα με έναν μη-αρνητικό πολλαπλασιαστή a_i (πολλαπλασιαστές Lagrange) και στην συνέχεια την αφαιρούμε από την συνάρτηση κόστους. Από την εφαρμογή λοιπόν του κανόνα για την περίπτωση των *γραμμικά* διαχωρίσιμων δεδομένων, οδηγούμαστε στην ελαχιστοποίηση της ακόλουθης συνάρτησης κόστους:

$$L = \frac{1}{2} \cdot \|\mathbf{w}\|^2 - \sum a_i \cdot y_i (\mathbf{w} \cdot \mathbf{x}_i + b) + \sum a_i \quad (8)$$

Οι περιοριστικές συνθήκες είναι $a_i \geq 0$. Θέλουμε τώρα την παραπάνω σχέση να την ελαχιστοποιήσουμε ως προς \mathbf{w} , b απαιτώντας συγχρόνως να απαλείφονται τα a_i με μηδενισμό των

μερικών πρώτων παραγώγων. Επειδή όμως για λόγους απλοποίησης θέλουμε να εργαστούμε με τους πολλαπλασιαστές a_i , θα λύσουμε το δυαδικό πρόβλημα του προηγούμενου. Δηλαδή, θα *μεγιστοποιήσουμε* την σχέση (8) ως προς a_i απαιτώντας την απαλοιφή των w, b με μηδενισμό των μερικών πρώτων παραγώγων, τηρώντας όμως τους περιορισμούς $a_i \geq 0$. Μηδενίζοντας λοιπόν τις παραγώγους έχουμε:

$$\partial L / \partial w = 0 \rightarrow w = \sum a_i \cdot y_i \cdot x_i \quad (9)$$

$$\partial L / \partial b = 0 \rightarrow \sum a_i \cdot y_i = 0 \quad (10)$$

Αντικαθιστώντας τις (9), (10) στην (8) και λαμβάνοντας υπ' όψιν και τις νέες περιοριστικές συνθήκες, καταλήγουμε στην διατύπωση του ακόλουθου προβλήματος:

$$\text{μεγιστοποίηση } -\frac{1}{2} \cdot \sum \sum a_i \cdot y_i \cdot a_j \cdot y_j \cdot x_i \cdot x_j + \sum a_i \quad (11)$$

$$\text{με περιορισμούς } a_i \geq 0, \sum a_i \cdot y_i = 0 \quad (12)$$

Οι σχέσεις (11), (12) περιγράφουν το προς βελτιστοποίηση πρόβλημα στην επίλυση του οποίου στοχεύει η μεθοδολογία SVM για την περίπτωση των γραμμικά διαχωρίσιμων δεδομένων.

Όπως όμως ήδη τονίσαμε μεγαλύτερο ενδιαφέρον υπάρχει για την περίπτωση των μη διαχωρίσιμων δεδομένων, αφού αυτή συναντάται συχνότερα στην πράξη και άλλωστε αποτελεί γενικότερη περίπτωση. Η πρώτη περίπτωση που μόλις αναλύσαμε αποτελεί στην ουσία υποπερίπτωση της κατηγορίας των μη διαχωρίσιμων δεδομένων. Εφαρμόζοντας πάλι την θεωρία Lagrange στις σχέσεις (6), (7) και μηδενίζοντας όπως πριν τις παραγώγους καταλήγουμε στην ακόλουθη διατύπωση του προβλήματος:

$$\text{μεγιστοποίηση } -\frac{1}{2} \cdot \sum \sum a_i \cdot y_i \cdot a_j \cdot y_j \cdot x_i \cdot x_j + \sum a_i \quad (13)$$

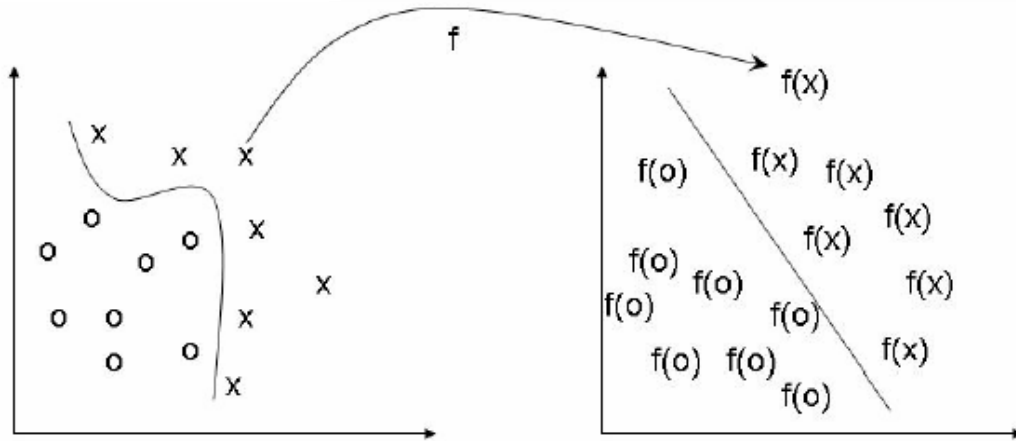
$$\text{με περιορισμούς } 0 \leq a_i \leq C, \sum a_i \cdot y_i = 0 \quad (14)$$

Είναι αξιοσημείωτο ότι η διατύπωση του προβλήματος είναι ίδια με την προηγούμενη, με την διαφορά ότι στους περιορισμούς για τους πολλαπλασιαστές Lagrange προβλέπεται ένα άνω όριο στην τιμή τους, που είναι μάλιστα η παράμετρος C για την οποία μιλήσαμε προηγουμένως. Να τονίσουμε το γεγονός ότι οι μεταβλητές ξ_i που εισήχθησαν νωρίτερα για την επίλυση, δεν εμφανίζονται πλέον πουθενά στην τελική μορφή του προβλήματος.

Η μέθοδος που χρησιμοποιήθηκε για την επίλυση του συγκεκριμένου προβλήματος θα συζητηθεί αργότερα. Προηγουμένως, θα αναφερθούμε στην δεύτερη περίπτωση που αναφέρεται σε μη γραμμικά δεδομένα.

3.1.2 Μη γραμμικές SVM

Μέχρι στιγμής, εξετάσαμε την γραμμική περίπτωση, όπου δηλαδή η συνάρτηση απόφασης για την τελική ταξινόμηση είναι γραμμική συνάρτηση των δεδομένων. Πρέπει να βρούμε έναν τρόπο ώστε να γενικεύσουμε την μεθοδολογία για περιπτώσεις που τα δεδομένα δεν είναι γραμμικά, η οποία άλλωστε είναι και η γενικότερη περίπτωση. Αυτό μπορεί να γίνει πολύ απλά, αρκεί πρώτα να εφαρμόσουμε μία μη γραμμική απεικόνιση $F: R^n \rightarrow H$, προκειμένου να απεικονίσουμε τα δεδομένα εισόδου σ' έναν χώρο μεγαλύτερης διάστασης. Με αυτόν τον τρόπο καταφέρνουμε να μεταφέρουμε τα μη γραμμικά δεδομένα εισόδου σε έναν άλλο χώρο, που ενδεχομένως να είναι πλέον γραμμικά διαχωρίσιμα, οπότε και μπορούμε να εφαρμόσουμε όσα παρουσιάστηκαν στην προηγούμενη υποενότητα. Αυτό απεικονίζεται γραφικά στο παρακάτω σχήμα.



Σχήμα: Απεικόνιση των δεδομένων εισόδου σε νέο χώρο με χρήση συνάρτησης f

Εργαζόμαστε ακριβώς όπως στην γραμμική περίπτωση, με την διαφορά ότι αντί για το διάνυσμα εισόδου χρησιμοποιούμε σε όλες τις παραπάνω σχέσεις το διάνυσμα $F(x_i)$. Ακολουθώντας λοιπόν την ίδια πορεία, οδηγούμαστε για την περίπτωση των μη γραμμικών SVM στην επίλυση του ακόλουθου προβλήματος:

$$\text{μεγιστοποίηση } -\frac{1}{2} \cdot \sum \sum a_i \cdot y_i \cdot a_j \cdot y_j \cdot F(x_i) \cdot F(x_j) + \sum a_i \quad (15)$$

$$\text{με περιορισμούς } 0 \leq a_i \leq C, \sum a_i \cdot y_i = 0 \quad (16)$$

Παρατηρούμε λοιπόν ότι η μορφή του προβλήματος βελτιστοποίησης είναι ίδια με πριν, με την διαφορά ότι αντί για το εσωτερικό γινόμενο $x_i \cdot x_j$, πρέπει να υπολογίσουμε το γινόμενο $F(x_i) \cdot F(x_j)$. Φυσικά, ο υπολογισμός του γινομένου αυτού είναι ιδιαίτερα χρονοβόρος, ειδικά αν πρόκειται για διανύσματα μεγάλης διάστασης, με αποτέλεσμα να αυξάνεται η υπολογιστική πολυπλοκότητα. Η συγκεκριμένη μορφή μας ωθεί στο να χρησιμοποιήσουμε αντί του παραπάνω εσωτερικού γινομένου μία συνάρτηση πυρήνα (kernel) $K(x_i, x_j)$ τέτοια ώστε:

$$K(x_i, x_j) = F(x_i) \cdot F(x_j) \quad (17)$$

με την προϋπόθεση ο πυρήνας αυτός να ικανοποιεί το θεώρημα του Mercer. Σύμφωνα με το συγκεκριμένο θεώρημα, κάθε *συμμετρική, θετικά ορισμένη* συνάρτηση $K(x_1, x_2)$ θεωρείται πυρήνας, υπάρχει δηλαδή μια απεικόνιση Φ τέτοια ώστε να είναι δυνατό:

$$K(x_1, x_2) = \langle \Phi(x_1) \cdot \Phi(x_2) \rangle \quad (18)$$

όπου $\langle \Phi(x_1) \cdot \Phi(x_2) \rangle$ είναι το εσωτερικό γινόμενο των $\Phi(x_1)$ και $\Phi(x_2)$.

Με την συγκεκριμένη αντικατάσταση δεν ενδιαφερόμαστε πλέον για την επιλογή της απεικόνισης F , μιας και αυτή δεν δηλώνεται πλέον πουθενά στην επίλυση του προβλήματος. Με την χρήση της συνάρτησης K μειώνουμε την υπολογιστική πολυπλοκότητα, αφού δεν απαιτείται πλέον η μεταφορά των δεδομένων σε χώρο μεγαλύτερης διάστασης και ο ακόλουθος υπολογισμός των εσωτερικών γινομένων. Η απεικόνιση των δεδομένων γίνεται πλέον εμμέσως κατά την διάρκεια της εκπαίδευσης. Για τον λόγο αυτό, μοναδικός μας στόχος είναι η κατάλληλη επιλογή του πυρήνα K . Δεν υπάρχει κάποια θεωρία σχετική με την επιλογή βέλτιστου πυρήνα, μιας και η σωστή επιλογή σχετίζεται άμεσα με τα δεδομένα του συνόλου εκπαίδευσης. Οι πιο γνωστοί πυρήνες που χρησιμοποιήθηκαν στη συγκεκριμένη εργασία είναι οι ακόλουθοι:

- 1) γκαουσιανός RBF: $K(x, y) = e^{-\gamma \|x-y\|^2}$ (19)

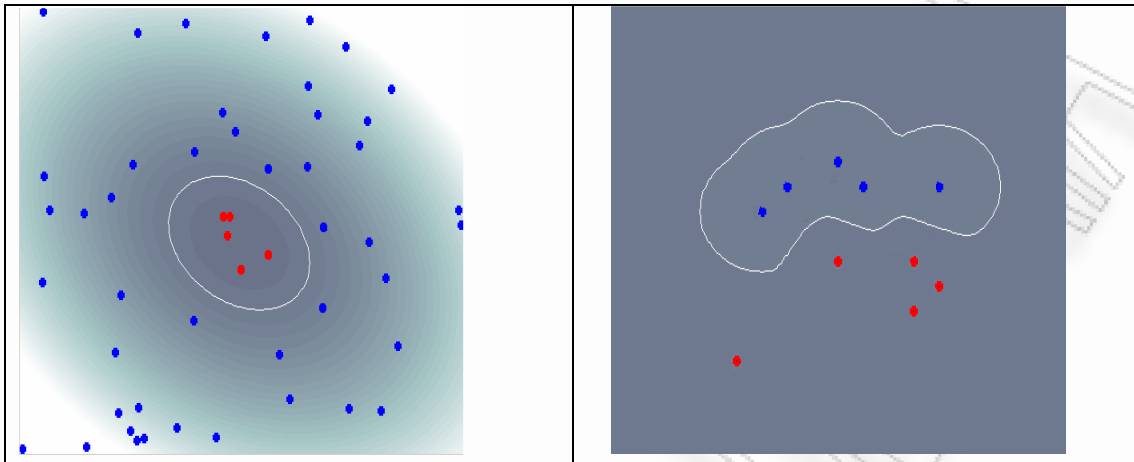
- 2) σιγμοειδής: $K(x, y) = \tanh(\gamma \cdot x \cdot y + 1)$ (20)

- 3) πολυωνυμικός: $K(x, y) = (x \cdot y + 1)^p$ (21)

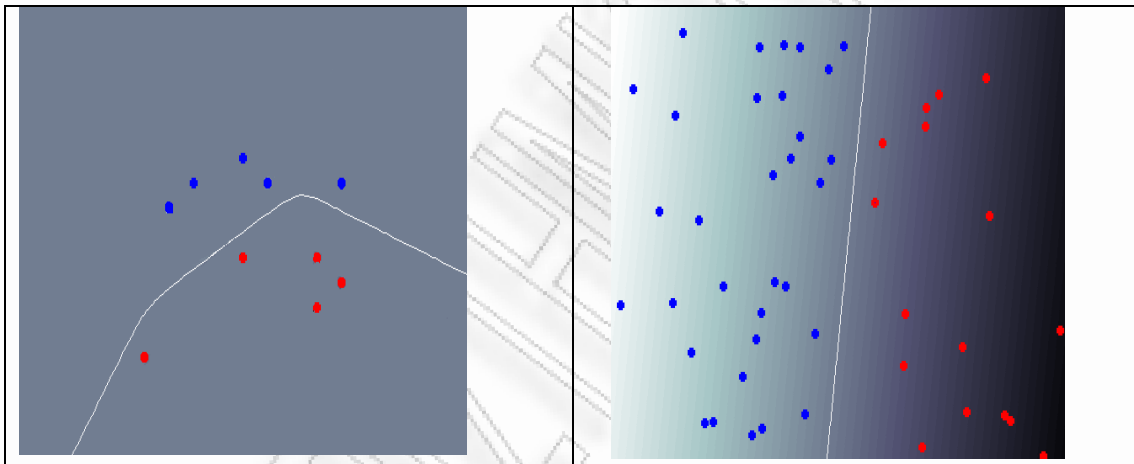
- 4) γραμμικός: $K(x, y) = x \cdot y$ (22)

όπου γ είναι μια σταθερά και p ο βαθμός του πολυωνύμου. Να τονίσουμε βέβαια εδώ ότι ο γραμμικός πυρήνας αναφέρεται στην ουσία στην περίπτωση των γραμμικών SVM. Οι συγκεκριμένες παράμετροι μαζί με την παράμετρο C που αναφέραμε προηγουμένως επιλέγονται από τον χρήστη για να χρησιμοποιηθούν στη φάση της εκπαίδευσης. Προφανώς, διαφορετικά ζευγάρια τιμών αυτών των παραμέτρων οδηγούν σε διαφορετικά αποτελέσματα εκπαίδευσης και κατ' επέκταση σε διαφορετικές επιδόσεις ταξινόμησης. Για τον λόγο αυτό, απαιτείται ιδιαίτερη προσοχή στην επιλογή των τιμών, ώστε να επιλεγθούν αυτές που παρέχουν την καλύτερη επίδοση

Στα επόμενα σχήματα, παρουσιάζονται παραδείγματα ταξινόμησης με χρήση διαφορετικών πυρήνων. Είναι φανερό, ότι η επιλογή του κατάλληλου πυρήνα δεν είναι απόλυτη, αλλά εξαρτάται από την μορφή του συνόλου εκπαίδευσης.



Σχήμα: (Αριστερά) Ταξινόμηση με πολυωνυμικό πυρήνα, (Δεξιά) Ταξινόμηση με RBF πυρήνα



Σχήμα: (Αριστερά) Ταξινόμηση με σιγμοειδή πυρήνα, (Δεξιά) Ταξινόμηση με γραμμικό πυρήνα

4 Ταξινομητής Νευρωνικών Δικτύων

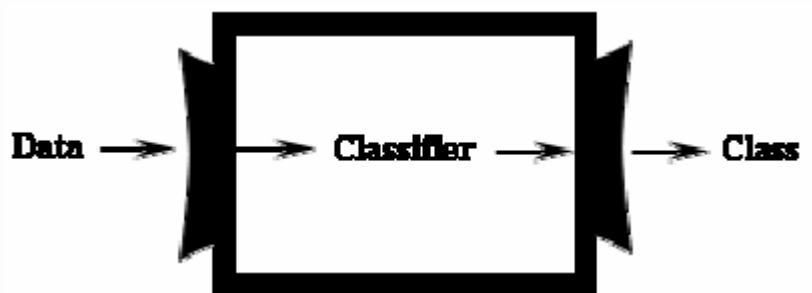
Το Τεχνητό Νευρωνικό Δίκτυο (ΤΝΔ, Artificial Neural Network (ANN)), το οποίο συχνά καλείται Προσομοιωμένο Νευρωνικό Δίκτυο ή απλά Νευρωνικό δίκτυο (ΝΔ), είναι ένα μαθηματικό ή υπολογιστικό μοντέλο βασισμένο στα βιολογικά νευρωνικά δίκτυα. Αποτελείται από διασυνδεδεμένες ομάδες τεχνητών νευρώνων και επεξεργάζεται τις πληροφορίες χρησιμοποιώντας την συνδετική προσέγγιση (connectionist approach) στους υπολογισμούς. Στις περισσότερες περιπτώσεις ένα ΝΔ αποτελεί ένα προσαρμόσιμο / δυναμικό σύστημα, το οποίο αλλάζει τη δομή του είτε λόγω εξωτερικών είτε λόγω εσωτερικών πληροφοριών που προκύπτουν κατά τη φάση της εκπαίδευσης.

Τα Τεχνητά Νευρωνικά Δίκτυα (Τ.Ν.Δ.) είναι ένα από τα σημαντικά εργαλεία στον τομέα της Τεχνητής Νοημοσύνης. Ουσιαστικά, τα Τεχνητά Νευρωνικά Δίκτυα δημιουργήθηκαν όταν, μετά από έρευνες πάνω στην οργάνωση του εγκεφάλου και τους βιολογικούς νευρώνες, έγινε δυνατή η κατασκευή του Τεχνητού Νευρώνα, δηλαδή ενός τεχνητού υπολογιστικού στοιχείου που θα προσομοιώνει την δυνατότητα του βιολογικού νευρώνα να "μαθαίνει" και να "εκπαιδεύεται", με απώτερο σκοπό την χρησιμοποίηση των Τεχνητών Νευρωνικών Δικτύων σε εφαρμογές αναγνώρισης φωνής και εικόνας, ανάκτησης πληροφοριών, λήψης αποφάσεων και γενικά την προσομοίωση της λειτουργίας του ανθρώπινου εγκεφάλου.

Εδώ υφίσταται και η έννοια της Τεχνητής Νοημοσύνης. Το Τ.Ν.Δ. θεωρείται "έξυπνο" γιατί είναι σε θέση να αναγνωρίζει κάποιο πρότυπο, το οποίο έχει κάποιες ομοιότητες με τα ήδη γνωστά πρότυπα, αλλά δεν ταυτίζεται ακριβώς. Το φαινόμενο αυτό ονομάζεται "ανοχή στον θόρυβο"

(noise tolerance) και είναι αυτό που κάνει τόσο σημαντική την θέση των Τ.Ν.Δ. στις εφαρμογές αναγνώρισης φωνής και εικόνας, όπου ο θόρυβος έχει την μορφή ήχων, διαφορετικής χροιάς, μουτζούρας, διαφορετικού χρώματος και φωτισμού και γενικά οτιδήποτε μπορεί να παραμορφώσει τα εισαγόμενα δεδομένα

Τα δεδομένα που δίδονται στην είσοδο του Τ.Ν.Δ. αντιπροσωπεύουν τα χαρακτηριστικά του αντικειμένου προς αναγνώριση, κατάλληλα διαμορφωμένα έτσι ώστε να γίνονται κατανοητά από το δίκτυο. Τα αποτελέσματα της λειτουργίας ενός Τ.Ν.Δ. έχουν την μορφή της ταξινόμησης, δηλαδή το Τ.Ν.Δ. για κάθε πρότυπο εισόδου εξάγει (επίσης κατάλληλα διαμορφωμένα) σε ποια κλάση αντιστοιχεί το αντικείμενο προς αναγνώριση. Για αυτό και τα Τ.Ν.Δ. ονομάζονται και *classifiers* (ταξινομητές).



Σχήμα: Λειτουργία ταξινομητή

Παραδείγματος χάριν: εκπαιδεύουμε ένα Τ.Ν.Δ. να αναγνωρίζει οχήματα. Τα πρότυπα εισόδου περιέχουν διάφορα χαρακτηριστικά όπως πλήθος τροχών, βάρος κ.λ.π. και να εξάγει ως αποτέλεσμα τον χαρακτηριστικό αριθμό για την κατηγορία στην οποία ανήκει το όχημα προς αναγνώριση: 0 για φορτηγά, 1 για αυτοκίνητα, 2 για μοτοσυκλέτες κ.ο.κ.

Αυτά είναι μερικά από τα κοινά χαρακτηριστικά των Τ.Ν.Δ. Μεταξύ τους, όμως, υπάρχουν σημαντικές διαφορές όπως η δομή και ο τρόπος εκπαίδευσης οι οποίες και είναι αρκετές ώστε να παρουσιάζουν διαφορές στον τρόπο λειτουργίας του κάθε Τ.Ν.Δ.

Οι πρώτες προσπάθειες για κατασκευή μοντέλων που θα εξομοίωνουν την λειτουργία του εγκεφάλου άρχισαν στα μέσα της δεκαετίας του 1940. Ο ενθουσιασμός της επιστημονικής κοινότητας τα πρώτα χρόνια ήταν μεγάλος και υπήρχε έντονη ερευνητική δραστηριότητα, καθώς πολλοί είχαν πειστεί ότι η κατασκευή ενός τεχνητού εγκεφάλου ήταν ζήτημα χρόνου. Όμως, οι περιορισμοί των πρώτων Τ.Ν.Δ., όπως και η μικρή, τότε, διαθέσιμη επεξεργαστική ισχύ επέτρεπε την χρήση των νευρωνικών μοντέλων για επίλυση πολύ απλών προβλημάτων.

Αυτοί ήταν μερικοί από τους λόγους για τους οποίους τα Τ.Ν.Δ. δέχτηκαν σφοδρές επιθέσεις που υποβάθμιζαν τις δυνατότητες εφαρμογής τους όπως και τον ιδιότητά τους ως "μονάδες τεχνητής νοημοσύνης". Το αποτέλεσμα ήταν η έρευνα πάνω σε αυτά να μειωθεί ιδιαίτερα στο τέλος της δεκαετίας του '60. Τα υπόλοιπα χρόνια υπάρχει κάποια σημαντική εργασία πάνω στο αντικείμενο αλλά από ένα μικρό μόνο πλήθος ερευνητών, αφού τα Τ.Ν.Δ. έχουν σχεδόν εγκαταλειφθεί.

Αυτή η κατάσταση παραμένει μέχρι τις αρχές του '80, καθώς τα επαρκή πια υπολογιστικά συστήματα και οι νέες τεχνικές εκπαίδευσης που εμφανίζονται, στρέφουν το ενδιαφέρον πάλι στα Τ.Ν.Δ. και απαλείφουν και τις τελευταίες αρνητικές αντιδράσεις απέναντί τους, αφού πλέον βρίσκουν εφαρμογή σε πραγματικά προβλήματα όπως η αναγνώριση προτύπων, φωνής και εικόνας.

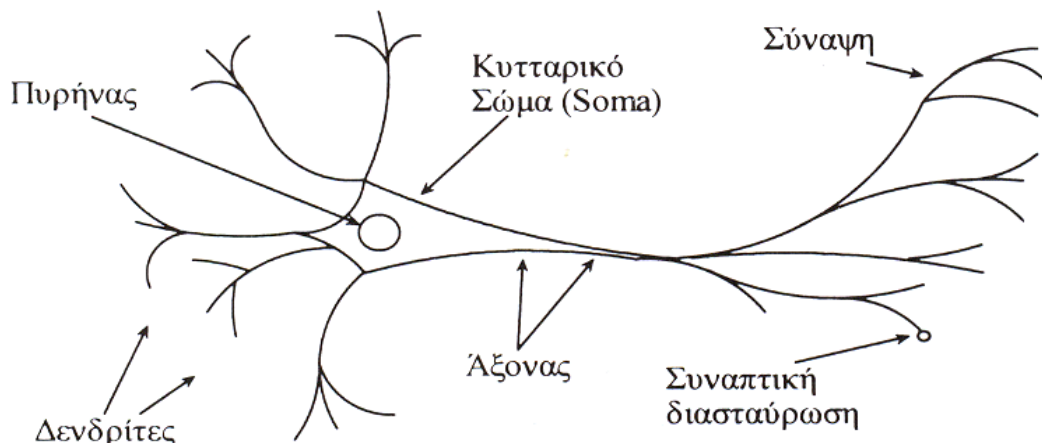
4.1 Βιολογικά νευρωνικά δίκτυα

Με τον όρο "Νευρωνικό Δίκτυο" (ΝΔ) εννοείται κάθε αρχιτεκτονική υπολογισμού η οποία περιλαμβάνει ένα μεγάλο αριθμό διασυνδεδεμένων απλών «νευρωνικών» επεξεργαστών. Με προσεκτική απομίμηση της λειτουργίας του ανθρώπινου εγκεφάλου, τα τεχνητά νευρωνικά δίκτυα έχουν εξοπλισθεί με χαρακτηριστικές ιδιότητες ανάλογες του εγκεφάλου, όπως είναι η ικανότητα να μαθαίνει από εμπειρίες, να γενικεύει την υπάρχουσα γνώση και να εκτελεί λογικές αφαιρέσεις. Έτσι, ισχύει για τα ΝΔ αυτό που είπε ο Σωκράτης για τον εαυτό του (για τον άνθρωπο), δηλαδή «γηράσκω αεί διδασκόμενος» ή ισοδύναμα «είμαι μόνιμος μαθητής».

Συνεπώς, για την καλύτερη κατανόηση των τεχνητών νευρωνικών δικτύων, κρίνεται σκόπιμο να γίνει εν συντομία μια περιγραφή της δομής των βιολογικών νευρωνικών δικτύων. Η αρχική εργασία των νευρο-φυσιολόγων επικεντρώθηκε στις θεμελιώδεις ηλεκτροφυσιολογικές λειτουργίες μεμονωμένων νευρώνων και στον τρόπο με τον οποίο οι νευρώνες αυτοί επικοινωνούν ανταλλάσσοντας μεταξύ τους ηλεκτρικά σήματα. Ο νευρώνας (ή αλλιώς νευρόνιο) είναι το θεμελιακό δομικό στοιχείο του ανθρώπινου νευρικού συστήματος. Οι νευρώνες είναι ζώντα κύτταρα. Ο

ανθρώπινος εγκέφαλος έχει κατά προσέγγιση 1.5×10^{10} νευρώνες διαφόρων μορφών. Κάθε νευρώνας λαμβάνει σήματα μέσω περίπου 10 συνάψεων. Ο εγκέφαλος αποτελείται ουσιαστικά από έναν αριθμό μικρότερων λειτουργικών περιοχών όπως είναι ο φλοιός κίνησης (motor cortex) , ο φλοιός όρασης (visual cortex) , ο φλοιός ακοής (auditory cortex) , η παρεγκεφαλίδα (cerebellum), κ.α.

Όλοι οι νευρώνες, ανεξάρτητα από το είδος, το σχήμα και το μέγεθος τους, αποτελούνται από τα ίδια βασικά μέρη, ήτοι: το *κυτταρικό σώμα*, τους *δενδρίτες* και τον *άξονα* όπως δείχνει το παρακάτω Σχήμα. Το κυτταρικό σώμα είναι το κεντρικό μέρος του νευρώνα και το σχήμα του μπορεί να είναι στρογγυλό, τριγωνικό, σταγονοειδές, μυτερό στα δύο άκρα κλπ, ανάλογα με το είδος του νευρώνα: οπτικός, ακουστικός, αφής, μυϊκός (κίνησης), κοκ.



Σχήμα: Δομή τυπικού βιολογικού νευρώνα (νευρώνας κίνησης).

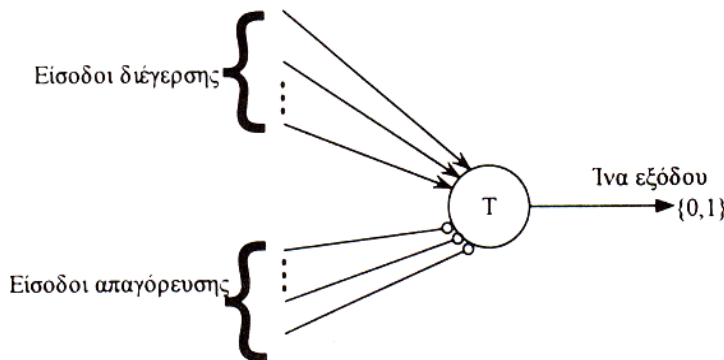
Αναλυτικότερα, το μέγεθος του «σώματος» είναι περίπου 10-80 μm . Το σώμα περιέχει τον «πυρήνα» και το αντίστοιχο συνδετικό πρωτόπλασμα. Ο ρυθμός μεταβολισμού του νευρώνα είναι πολύ υψηλός. Η σύνδεση του νευρώνα με τους πολυάριθμους γειτονικούς νευρώνες γίνεται μέσω του άξονα και των συνάψεων. Οι δενδρίτες συνθέτουν την εξωτερική επιφάνεια του νευρώνα και μεταφέρουν πληροφορίες στο σώμα του κυττάρου. Ο άξονας, που αποτελεί το τμήμα εξόδου του νευρικού κυττάρου, είναι μια λεπτή κυλινδρική ίνα η οποία μπορεί να μεταφέρει πληροφορίες ηλεκτροχημικά με βάση την κατάσταση του κυττάρου. Οι δενδρίτες και οι άξονες έχουν διάμετρο ολίγων μm . Η δομή αυτή μπορεί να υποστεί σημαντική δένδροποίηση καθώς τελειώνει στους καταληκτικούς βολβούς (endbulbs) κοντά στους δενδρίτες άλλων νευρώνων. Οι διασυνδέσεις (δηλαδή ο τύπος αλληλεπίδρασης) ενός καταληκτικού βολβού με το κύτταρο στο οποίο εισβάλλει είναι αυτό που ονομάστηκε από τον Sherrington «συνάψεις» (synapses) και αποτέλεσε το θεμέλιο της καλούμενης «συναπτολογίας». Ένας παλμός που ταξιδεύει κατά μήκος του άξονα τερματίζεται στους καταληκτικούς βολβούς και, εάν το επιτρέπουν οι συνθήκες, αναπαράγεται από τον επόμενο νευρώνα της αλυσίδας. Η ταχύτητα της παραγωγής του παλμού κατά μήκος της ίνας του άξονα ισούται προσεγγιστικά με: «ταχύτητα = συχνότητα \times διάμετρος», και μπορεί σε ορισμένους μυϊκούς νευρώνες να φθάσει μέχρι 10 m/sec.

Ένας νευρώνας παράγει έξοδο μόνο όταν εμφανιστούν αρκετοί παλμοί στην περιοχή των δενδριτών για ένα μικρό χρονικό διάστημα που ονομάζεται «λανθάνουσα περίοδος άθροισης». Στην πραγματικότητα, κάποιες εισοδοί εμποδίζουν την ενεργοποίηση του κυττάρου, με αποτέλεσμα ένας νευρώνας ενεργοποιείται εάν οι ενεργές εισοδοί διέγερσης ξεπερνούν τις ενεργές απαγορευτικές εισόδους κατά μία ποσότητα ίση με την «τιμή κατωφλίου» του κυττάρου. Ένα τμήμα του σώματος, ονομαζόμενο «εκκινήτης κορυφής» πραγματοποιεί την πρόσθεση αυτή και τη σύγκριση με το κατώφλι παράγοντας ένα παλμικό σήμα εξόδου για τις υπάρχουσες καταστάσεις εισόδου. Ο μέγιστος ρυθμός παραγωγής παλμών μέσα σε ένα νευρικό κύτταρο είναι περίπου 500 παλμοί/sec, ενώ το τυπικό πλάτος ενός νευρικού παλμού

είναι περίπου 100 μV με αποτέλεσμα ο παλμός να πλησιάζει τα +30mV. Ο χρόνος καθυστέρησης ανάμεσα στη διέγερση του νευρώνα και την παραγωγή του παλμού κυμαίνεται από 0.1 msec μέχρι 1 msec. Η τιμή του κατωφλίου ενός νευρώνα είναι συνάρτηση του χρόνου. Έτσι λ.χ. ύστερα από μια ενεργοποίηση, του νευρώνα εισέρχεται σε μια απαγορευτική περίοδο κατά τη διάρκεια της οποίας δεν μπορεί να ενεργοποιηθεί ξανά. Επαναλαμβανόμενες ενεργοποιήσεις αυξάνουν την τιμή του κατωφλίου με αποτέλεσμα το κύτταρο να οδηγείται σε «κόπωση».

Η σημαντικότερη προσπάθεια μοντελοποίησης του βιολογικού νευρώνα έγινε από τους McCulloch και Pitts. Το μοντέλο των McCulloch - Pitts δεν είναι ένα πλήρες μοντέλο του νευρικού

κυττάρου, αλλά μια απλή παρουσίαση δύο καταστάσεων της λογικής διαδικασίας που λαμβάνει χώρα σ' αυτό. Το βασικό μοντέλο νευρώνα των McCulloch και Pitts έχει τη μορφή του σχήματος .



Σχήμα: Μοντέλο νευρώνα McCulloch – Pitts.

Η ένα εξόδο του κυττάρου συσχετίζει τις πληροφορίες που αφορούν στην κατάσταση του νευρώνα (διέγερση ή όχι) και τελικά γίνεται είσοδος σε ένα άλλο κύτταρο (ή και στο ίδιο). Επιτρέπεται οποιοσδήποτε αριθμός εισόδων και οι εισοδοι διαιρούνται σε δύο κατηγορίες: στις εισόδους που διεγείρουν το κύτταρο και στις απαγορευτικές εισόδους. Οι εισοδοι διέγερσης που αναπαρίστανται με μια γραμμή, η οποία καταλήγει σε ένα βελάκι, τείνουν στο να διεγείρουν το κύτταρο κατά αυξανόμενο τρόπο μέχρι την τελική διέγερση. Οι απαγορευτικές εισοδοι αναπαρίστανται από μια γραμμή η οποία καταλήγει σε ένα μικρό κύκλο. Οι εισοδοι αυτές είναι ικανές να απαγορεύσουν εντελώς τη διέγερση του κυττάρου.

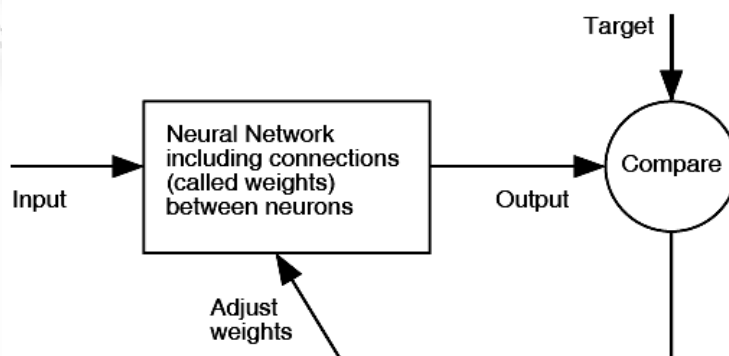
Έστω T η αριθμητική τιμή του κατωφλίου διέγερσης του κυττάρου. Το κύτταρο McCulloch - Pitts θα διεγερθεί όταν το άθροισμα των εισόδων διέγερσης γίνει μεγαλύτερο ή ίσο με την τιμή κατωφλίου και δεν υπάρχουν απαγορευτικές εισοδοι.

Έτσι για το απλό κύτταρο η έξοδος τη χρονική στιγμή $t+\Delta T$ δεν είναι συνάρτηση της εξόδου στο χρόνο t . Οι συνήθεις λογικές συναρτήσεις Boole μπορούν να υλοποιηθούν από αυτό το λογικό σχήμα. Προς τούτο υπάρχει δυνατότητα χρήσης λογικών πυλών AND και OR.

4.2 Τεχνητά Νευρωνικά Δίκτυα

Ο τρόπος με τον οποίο πραγματοποιούνται οι διάφοροι υπολογισμοί μέσω ΝΔ διαφέρει ιδιαίτερα από αυτόν που χρησιμοποιείται στον κλασικό προγραμματισμό. Η κύρια διαφορά τους είναι ότι το ΝΔ μπορεί να εκπαιδευτεί χρησιμοποιώντας δείγματα (στη διεθνή βιβλιογραφία αναφέρονται ως teaching signals) για να εκτελέσουν συγκεκριμένες εργασίες.

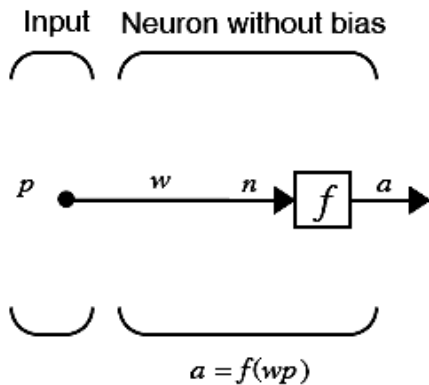
Τα ΝΔ αποτελούνται από απλά στοιχεία τα οποία λειτουργούν εν παραλλήλω. Όπως έχει αναφερθεί και παραπάνω, αυτά τα στοιχεία είναι εμπνευσμένα από το βιολογικό νευρικό σύστημα. Όπως στη φύση, η λειτουργία του δικτύου καθορίζεται σε μεγάλο βαθμό από τις συνδέσεις μεταξύ αυτών των στοιχείων. Συνήθως, τα ΝΔ προσαρμόζονται ή εκπαιδεύονται ώστε μια συγκεκριμένη είσοδος να οδηγεί σε μια συγκεκριμένη έξοδο. Παρακάτω παρουσιάζεται στο επόμενο σχήμα μια τέτοια κατάσταση σχηματικά. Συγκεκριμένα, το δίκτυο προσαρμόζεται, βάση της σύγκρισης που γίνεται μεταξύ των εξόδων του ΝΔ και των αντίστοιχων επιθυμητών στόχων. Αυτή η διαδικασία συνεχίζεται έως ότου αυτές οι ποσότητες ταυτιστούν, δηλαδή μηδενιστεί το σφάλμα (πρακτικά η διαδικασία σταματά όταν το σφάλμα λάβει μια αποδεκτή τιμή). Στο σημείο αυτό πρέπει να σημειωθεί ότι για την εκπαίδευση του ΝΔ απαιτούνται πολλά τέτοια ζευγάρια εισόδων / εξόδων.



4.3 Μοντέλα νευρώνων

4.3.1 Απλός νευρώνας (χωρίς bias)

Παρακάτω παρουσιάζεται σχηματικά ένας νευρώνας με βαθμωτή είσοδο και χωρίς bias.

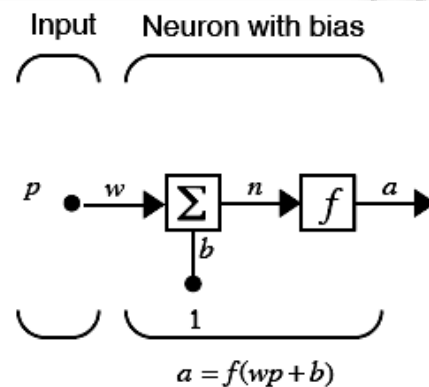


Σχήμα: Μοντέλο νευρώνα με βαθμωτή είσοδο, χωρίς bias.

Η βαθμωτή είσοδος p μεταδίδεται μέσω μιας σύνδεσης, η οποία πολλαπλασιάζει το μέγεθος της με το συντελεστή βαρύτητας w ώστε να σχηματιστεί το wp , το οποίο με τη σειρά του είναι και αυτό βαθμωτό μέγεθος. Η ποσότητα wp αποτελεί το όρισμα της συνάρτησης μεταφοράς f , η οποία τελικά παράγει τη βαθμωτή έξοδο a .

4.3.2 Απλός νευρώνας (με bias)

Ένας νευρώνας με βαθμωτή είσοδο και bias έχει την παρακάτω μορφή.



Σχήμα: Μοντέλο νευρώνα με βαθμωτή είσοδο και bias.

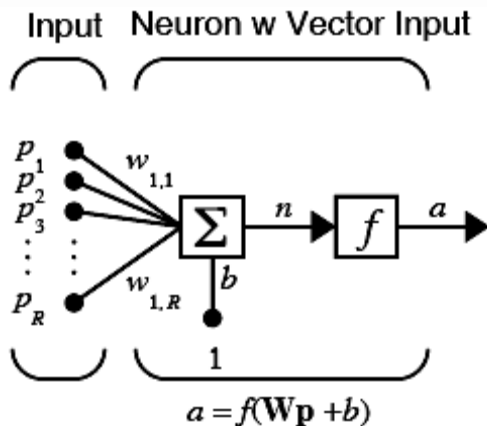
Σε αυτήν την περίπτωση το όρισμα της συνάρτησης μεταφοράς f θα είναι το καθαρό μέγεθος n , όπου $n=wp+b$. Το bias μπορεί να θεωρηθεί ένας συντελεστής βαρύτητας, με τη διαφορά ότι έχει σταθερή είσοδο τη μονάδα.

Στο σημείο αυτό πρέπει να επισημανθεί το γεγονός ότι τα μεγέθη w και b αποτελούν προσαρμόσιμες παραμέτρους του νευρώνα. Η κύρια ιδέα των νευρωνικών δικτύων είναι ότι τέτοιες παράμετροι μπορούν να μεταβληθούν ώστε τελικά το ΝΔ να έχει την επιθυμητή συμπεριφορά. Η εύρεση αυτών των τιμών, που θα δώσουν τα επιθυμητά αποτελέσματα είναι ο σκοπός της εκπαίδευσης του δικτύου.

Όπως αναφέρθηκε και προηγουμένως, το bias b είναι μια προσαρμόσιμη παράμετρος του νευρώνα. Δεν αποτελεί κάποια είσοδο. Παρόλα αυτά η σταθερά 1 η οποία κατευθύνει το bias είναι είσοδος και πρέπει να αντιμετωπίζεται ως είσοδος όταν αναφερόμαστε στη γραμμική εξάρτηση των διανυσμάτων εισόδου.

4.3.3 Νευρώνας με είσοδο διάνυσμα

Ένας νευρώνας με είσοδο ένα διάνυσμα p με R στοιχεία έχει την παρακάτω μορφή.



Σχήμα: Μοντέλο νευρώνα με είσοδο διάνυσμα.

Σε αυτήν την περίπτωση τα μεγέθη p_1, p_2, \dots, p_R πολλαπλασιάζονται με τους συντελεστές βαρύτητας $w_{1,1}, w_{1,2}, \dots, w_{1,R}$ και το σταθμισμένο αποτέλεσμα οδηγείται στην αθροιστική διασταύρωση. Το άθροισμα είναι απλά Wp , το δεκαδικό μέρος μιας γραμμής του πίνακα W και του διανύσματος p .

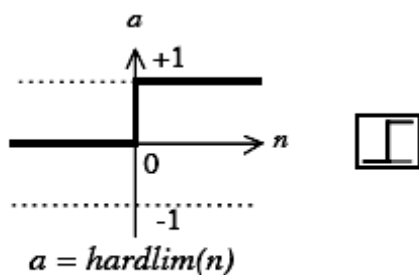
Ο νευρώνας έχει ένα bias b , το οποίο αθροίζεται με τις σταθμισμένες εισόδους για να σχηματίσει την καθαρή είσοδο n . Το άθροισμα n αποτελεί το όρισμα της συνάρτησης μεταφοράς f .

$$n = w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b$$

4.4 Συναρτήσεις μεταφοράς

Παρακάτω αναλύονται οι τέσσερις πιο συχνά χρησιμοποιούμενες συναρτήσεις μεταφοράς. Υπάρχουν βέβαια αρκετές ακόμα, οι οποίες χρησιμοποιούνται σε ιδιαίζουσες περιπτώσεις.

- Βηματική συνάρτηση μεταφοράς (Hard-limit)

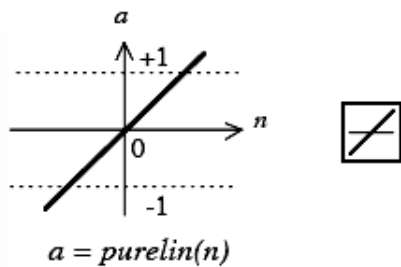


Hard-Limit Transfer Function

Σχήμα: Βηματική συνάρτηση μεταφοράς.

Η συνάρτηση θέτει ως έξοδο του νευρώνα είτε το 0, εάν το καθαρό (net) όρισμα εισόδου είναι μικρότερο από 0, είτε το 1, σε οποιαδήποτε άλλη περίπτωση. Συχνά, χρησιμοποιείται στα perceptrons ώστε να δημιουργηθούν νευρώνες κατάλληλοι για αποφάσεις κατηγοριοποίησης.

- Γραμμική συνάρτηση μεταφοράς (Linear Transfer Function)

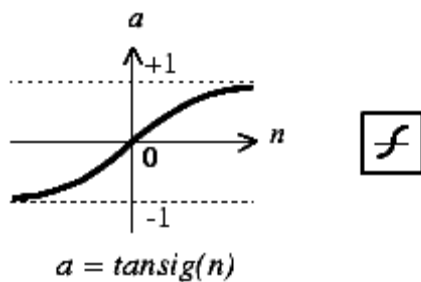


Linear Transfer Function

Σχήμα: Γραμμική συνάρτηση μεταφοράς.

Κυρίως, χρησιμοποιείται σε νευρώνες που προορίζονται για γραμμική προσέγγιση στα γραμμικά φίλτρα.

- Εφαπτομενική - σιγμοειδής συνάρτηση μεταφοράς (Tan-sigmoid Transfer Function)

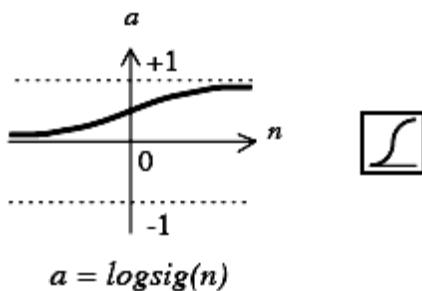


Tan-Sigmoid Transfer Function

Σχήμα: Εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς.

Η είσοδος της μπορεί να είναι οποιαδήποτε τιμή στο διάστημα $(-\infty, +\infty)$. Το πεδίο τιμών της όμως, περιορίζεται στο διάστημα $(-1, 1)$. Αποτελεί την κύρια επιλογή σε δίκτυα τύπου back-propagation λόγω της διαφορισμότητας της.

- Λογαριθμική – σιγμοειδής συνάρτηση μεταφοράς (Log-sigmoid Transfer Function)



Log-Sigmoid Transfer Function

Σχήμα: Λογαριθμική – σιγμοειδής συνάρτηση μεταφοράς.

Αυτή η συνάρτηση μεταφοράς παίρνει μια είσοδο, η οποία μπορεί να έχει τιμές από $-\infty$ έως $+\infty$ και περιορίζει την είσοδο στο πεδίο τιμών $(0, 1)$. Είναι κατάλληλη για δίκτυα τύπου back-propagation εφόσον είναι διαφορίσιμη.

4.5 Αρχιτεκτονικές δικτύων

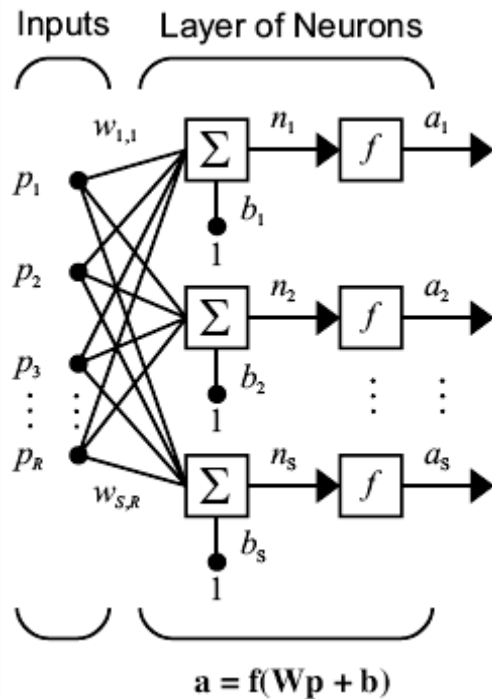
Οι αρχιτεκτονικές των Τ.Ν.Δ. χωρίζονται σε δύο μεγάλες κατηγορίες, τα feed-forward (μονής κατεύθυνσης) και τα recurrent (ανατροφοδότησης) δίκτυα. Η διαφορά τους έγκειται στο ότι τα recurrent δίκτυα ανατροφοδοτούν τους νευρώνες εισόδου με τα αποτελέσματα των νευρώνων εξόδου, ενώ τα feed-forward όχι.

Η τοπολογική δομή είναι το κύριο χαρακτηριστικό των Ν.Δ. και αναφέρεται στην αρχιτεκτονική στην οποία διευθετούνται και διασυνδέονται πολλαπλοί νευρώνες. Οι δύο βασικές ιδιότητες που καθορίζουν την αρχιτεκτονική ενός ΝΔ είναι το πλήθος των επιπέδων (layers) και οι

συνδέσεις μεταξύ των νευρώνων. Το τρίτο χαρακτηριστικό, το οποίο σχετίζεται φυσιολογικά με τον τρόπο κατά τον οποίο είναι δομημένοι οι νευρώνες, είναι ο *αλγόριθμος μάθησης* που χρησιμοποιείται για την εκπαίδευση του δικτύου.

4.5.1 Επίπεδο νευρώνων

Όταν δύο ή περισσότεροι νευρώνες συνδυαστούν τότε σχηματίζουν ένα επίπεδο νευρώνων. Ακολουθεί η γραφική αναπαράσταση ενός δικτύου με ένα επίπεδο νευρώνων με R στοιχεία εισόδου και S νευρώνες.



Σχήμα: Επίπεδο νευρώνων.

Σε αυτό το δίκτυο κάθε στοιχείο του διάνυσματος εισόδου συνδέεται με την είσοδο κάθε νευρώνα μέσω του πίνακα των συντελεστών βαρύτητας W . Ο i -οστός νευρώνας έχει έναν αθροιστή ο οποίος συλλέγει τις σταθμισμένες εισόδους και τα bias για να σχηματίσει τη δική του βαθμωτή είσοδο $n(i)$. Εάν συνδυαστούν τα διάφορα $n(i)$, δημιουργούν ένα διάνυσμα εισόδου n με πλήθος στοιχείων S . Τελικά, το νευρωνικό επίπεδο έχει ως έξοδο ένα διάνυσμα a .

Να σημειωθεί ότι είναι σύνηθες ο αριθμός των εισόδων σε ένα επίπεδο να είναι διαφορετικός από τον αριθμό των νευρώνων που το αποτελούν.

Ο πίνακας των συντελεστών βαρύτητας έχει την παρακάτω μορφή.

$$W = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,R} \\ w_{2,1} & w_{2,2} & \dots & w_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S,1} & w_{S,2} & \dots & w_{S,R} \end{bmatrix}$$

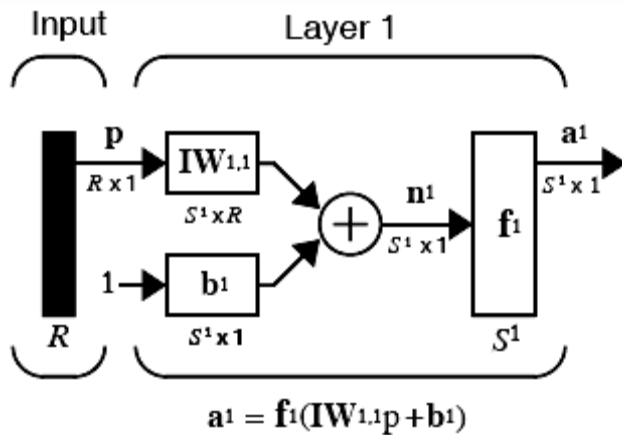
Η γραμμή του πίνακα υποδεικνύει τον νευρώνα στον οποίο θα εφαρμοστεί αυτό το βάρος, ενώ η στήλη υποδηλώνει την είσοδο.

4.5.2 Πολλαπλά επίπεδα νευρώνων (Multiple Layers of Neurons)

Για να περιγραφούν τα δίκτυα πολλαπλών επιπέδων, πρέπει να γίνουν κάποιες διευκρινήσεις. Ειδικότερα, πρέπει να γίνει διαχωρισμός μεταξύ των πινάκων των βαρών που συνδέονται με τις εισόδους και των πινάκων των συντελεστών βαρύτητας, οι οποίοι συνδέονται με τα διάφορα επίπεδα.

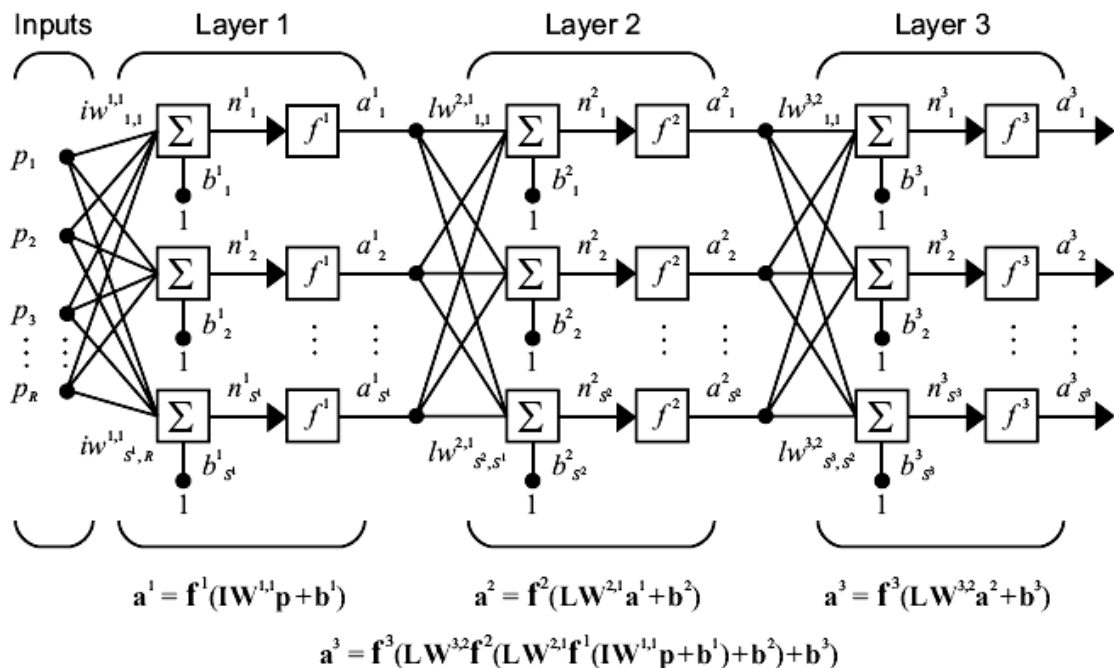
Θα καλούνται λοιπόν οι πίνακες που συνδέονται με τις εισόδους, *βάρη εισόδων*, ενώ οι πίνακες που συνδέονται με τις εξόδους των νευρωνικών επιπέδων, *βάρη επιπέδων*. Για να γίνεται αντιληπτό σε ποιο επίπεδο αναφέρεται η κάθε παράμετρος γίνεται χρήση ενός δείκτη. Από εδώ και στο εξής θα υιοθετηθεί η παραπάνω σύμβαση για την περαιτέρω περιγραφή των ΝΔ.

Το παρακάτω σχήμα αναπαριστά γραφικά ένα επίπεδο ενός πολυεπίπεδου ΝΔ.



Σχήμα: Μοντέλο ενός επιπέδου πολυεπίπεδου ΝΔ.

Όπως αναφέρθηκε και παραπάνω ένα ΝΔ μπορεί να αποτελείται από αρκετά επίπεδα. Κάθε επίπεδο έχει ένα πίνακα W , ένα bias b και ένα διάνυσμα εξόδου a . Ο πάνω συντελεστής που χρησιμοποιείται στα διάφορα μεγέθη σκοπό έχει να καθορίσει σε ποιο επίπεδο αναφέρονται οι πίνακες βαρών, τα διανύσματα εξόδου κτλ. (βλ. Σχήμα 3.16).



Σχήμα: Πολυεπίπεδο ΝΔ.

Το παραπάνω δίκτυο έχει R_1 εισόδους, S_1 νευρώνες στο πρώτο επίπεδο, S_2 νευρώνες στο δεύτερο επίπεδο κτλ. Είναι σύνηθες τα διαφορετικά επίπεδα να έχουν διαφορετικό αριθμό νευρώνων. Η είσοδος του bias είναι σταθερή ποσότητα για κάθε νευρώνα, ίση με 1.

Από το σχήμα φαίνεται επίσης ότι η έξοδος ενός ενδιάμεσου επιπέδου αποτελεί την είσοδο για το ακόλουθο επίπεδο. Επομένως, τα επίπεδα 2 μπορεί να αναλυθεί ως ένα επίπεδο με S_1 εισόδους, S_2 νευρώνες και έναν $S_2 \times S_1$ πίνακα βαρών W_2 . Η είσοδος του επιπέδου 2 είναι a^1 και η έξοδος του a^2 . Ύστερα, από τον καθορισμό των παραπάνω παραμέτρων μπορεί το συγκεκριμένο επίπεδο να αντιμετωπιστεί ως ένα ΝΔ ενός νευρωνικού επιπέδου. Αυτή η προσέγγιση μπορεί να γίνει για κάθε επίπεδο του δικτύου.

Τα διάφορα επίπεδα ενός πολυεπίπεδου ΝΔ διαδραματίζουν διαφορετικούς ρόλους το καθένα. Το επίπεδο που παράγει την έξοδο του δικτύου καλείται επίπεδο εξόδου (output layer). Όλα τα υπόλοιπα επίπεδα καλούνται κρυφά επίπεδα (hidden layers).

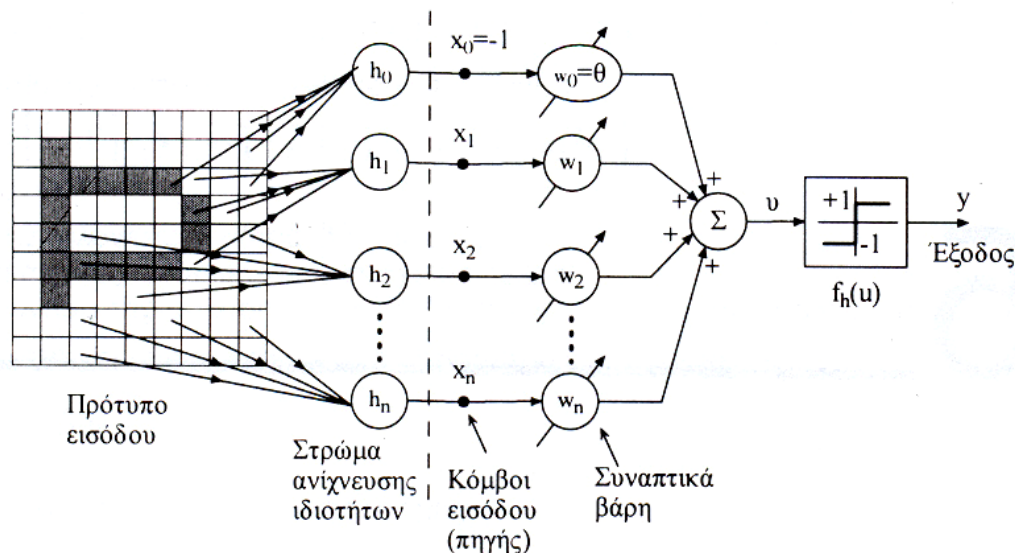
4.6 Perceptron

Το ΝΔ Perceptron αναπτύχθηκε από τον Rosenblatt (1958) και αποτελεί μια από τις πρώτες προσπάθειες σχεδίασης και κατασκευής ευφυών συστημάτων με δυνατότητες *αυτό-μάθησης*. Το perceptron είναι το απλούστερο ΝΔ που χρησιμοποιείται για την ταξινόμηση γραμμικά διαχωρίσιμων προτύπων δηλαδή προτύπων τα οποία διαχωρίζονται από ένα υπερεπίπεδο. Αποτελείται από έναν απλό νευρώνα με προσαρμόσιμα βάρη ο οποίος ακολουθείται από μια διπολική συνάρτηση ενεργοποίησης. Η δομή του perceptron (ενός νευρώνα) φαίνεται στο επόμενο σχήμα. Τούτο μπορεί να πραγματοποιήσει ταξινόμηση στην περίπτωση που έχουμε μόνο δύο κατηγορίες. Για να είναι δυνατή η ταξινόμηση περισσότερων γραμμικά διαχωρίσιμων κατηγοριών (κλάσεων) το επίπεδο εξόδου χρειάζεται να έχει περισσότερους από έναν νευρώνες.

Η έξοδος του perceptron δίνεται από τη σχέση:

$$y(t) = f_h(u(t))$$

$$u(t) = \sum_{i=0}^n w_i x_i = w^T(t)x(t)$$



Σχήμα: Perceptron ενός νευρώνα όπου το κατώφλι $w_0 = \theta$ έχει συμπεριληφθεί ως κανονική είσοδος ($x_0 = -1$). Οι εισοδοί του perceptron λαμβάνονται από ένα επίπεδο ανίχνευσης ιδιοτήτων (feature detection) με σταθερά βάρη h_1, h_2, \dots, h_n .

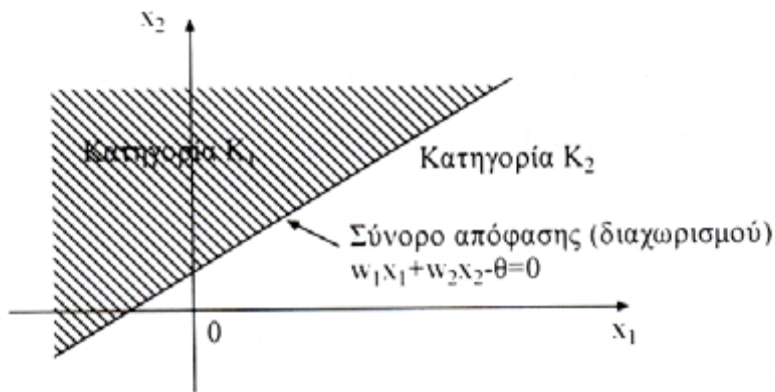
όπου $f_h(u)$ είναι η συνάρτηση του σκληρού περιοριστή (διπολικής συνάρτησης) και

$$x(t) = [-1, x_1(t), x_2(t), \dots, x_n(t)]^T$$

$$w(t) = [\theta(t), w_1(t), w_2(t), \dots, w_n(t)]^T$$

Ο σκοπός του perceptron είναι να ταξινομήσει τα εξωτερικά πρότυπα x_1, x_2, \dots, x_n σε μια από δύο κατηγορίες $K1$ ή $K2$. Τούτο γίνεται εκχωρώντας ένα σημείο $[x_1, x_2, \dots, x_n]^T$ στην κατηγορία $K1$ εάν η έξοδος y είναι $+1$ και στην κατηγορία $K2$ εάν $y = -1$. Η διδασκαλία του perceptron πρέπει να γίνει με καλά παραδείγματα κάθε μιας κατηγορίας. Εάν οι προς διαχωρισμό κατηγορίες μοιάζουν πολύ μεταξύ τους τότε είναι δύσκολο να εκπαιδευθεί το perceptron να ταξινομεί σωστά τα πρότυπα που του παρουσιάζονται στη μια ή την άλλη κατηγορία.

Για να γίνει κατανοητός ο τρόπος λειτουργίας ενός ταξινομητής προτύπων, θεωρείται η περίπτωση δύο μεταβλητών x_1 και x_2 όπου το σύνορο απόφασης (διαχωρισμού ανάμεσα στις κατηγορίες $K1$ και $K2$) είναι μια ευθεία γραμμή όπως δείχνει το επόμενο σχήμα.



Σχήμα: Γραμμικός διαχωρισμός δύο κλάσεων K1 και K2 στην περίπτωση δύο διαστάσεων (το κατώφλι θ απλά μεταθέτει το σύνορο διαχωρισμού).

Οποιοδήποτε σημείο βρίσκεται πάνω από τη διαχωριστική γραμμή ταξινομείται στην κατηγορία K1 διαφορετικά αποδίδεται στην κατηγορία K2. Ο αλγόριθμος επιλογής (ανανέωσης) των βαρών (αλγόριθμος perceptron) στηρίζεται στο γεγονός ότι εάν οι κλάσεις K1 και K2 είναι γραμμικά διαχωρίσιμες, τότε υπάρχει ένα διάνυσμα βαρών w τέτοιο ώστε

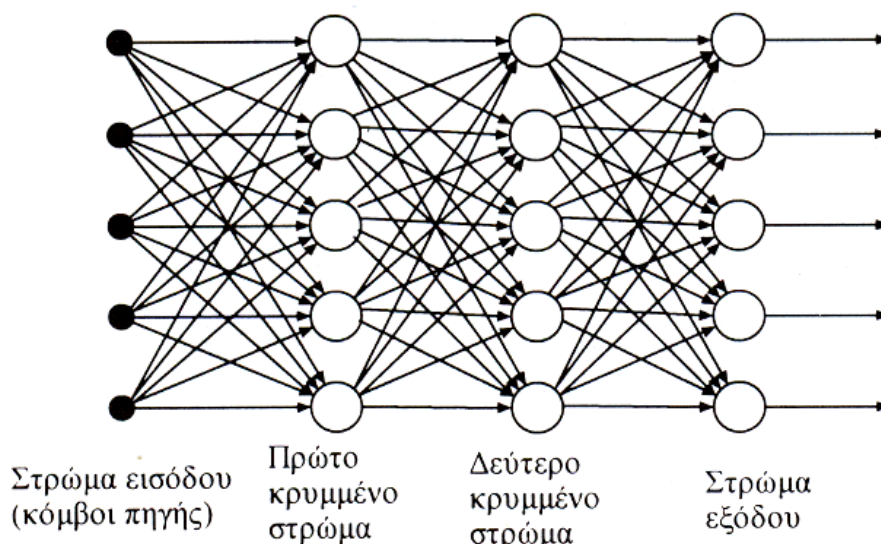
- $w^T x \geq 0$, όταν το X ανήκει στην K1
- $w^T x < 0$, όταν το ανήκει στην K2

Συνεπώς όταν το perceptron λαμβάνει ένα ζευγάρι εκπαίδευσης (x_1, x_2) πρέπει να προσδιορίσει ένα διάνυσμα βάρους w τέτοιο ώστε να ικανοποιούνται οι δύο παραπάνω ανισότητες.

4.7 Back-error Propagation (BP)

Το BP δίκτυο είναι ένα feed-forward T.N.Δ. πολλών επιπέδων στο οποίο οι έξοδοι των νευρώνων ενός επιπέδου αποτελούν τις εισόδους των νευρώνων του επόμενου επιπέδου και αποτελεί ένα από τα πιο απλά μέλη της οικογένειας αλγορίθμων γνωστών ως "gradient descent".

Το BP περιλαμβάνει, εκτός από το επίπεδο κόμβων εισόδου (πηγής) και το επίπεδο εξόδου (αποκρίσεων), ένα ή περισσότερα επίπεδα κρυμμένων κόμβων όπως δείχνει το παρακάτω σχήμα.



Σχήμα: BP με δύο κρυφά επίπεδα.

Το BP είναι κατάλληλο για την επίλυση πλήθους πολύπλοκων προβλημάτων και εκπαιδεύεται με επιβλεπόμενη μάθηση χρησιμοποιώντας τον αλγόριθμο ανάστροφης διάδοσης (BP: Back Propagation) ο οποίος στηρίζεται στον κανόνα μάθησης διόρθωσης σφάλματος.

Η λογική του είναι να ελαχιστοποιεί το συνολικό σφάλμα του δικτύου προσαρμόζοντας τα βάρη των συνάψεων και η εκπαίδευση σε γενικές γραμμές ακολουθεί τα παρακάτω βήματα:

- ▣ Αρχικοποίηση του δικτύου με τοποθέτηση τυχαίων, σχετικά μικρών, τιμών στα βάρη.
- ▣ Παρουσίαση διανύσματος εισόδου και του επιθυμητού αποτελέσματος.
- ▣ Υπολογισμός του διανύσματος εξόδου χρησιμοποιώντας, συνηθέστερα, την σιγμοειδή συνάρτηση ως συνάρτηση ενεργοποίησης.
- ▣ Προσαρμογή βαρών.

Επανάληψη της διαδικασίας από το βήμα 2.

Σε αντίθεση με τη διπολική μη γραμμική συνάρτηση ενεργοποίησης (σκληρό περιοριστή) που χρησιμοποιείται στο perceptron ενός επιπέδου του Rosenblatt, εδώ χρησιμοποιείται μια παραγωγίσιμη σιγμοειδής συνάρτηση ενεργοποίησης:

$$y_k = f(u_k) = 1/[1 + e^{-\lambda u_k}]$$

ή η υπερβολική εφαπτομένη:

$$y_k = f(u_k) = (1 - e^{-u_k}) / (1 + e^{-u_k})$$

όπου $u_k = u_k - \theta_k$ είναι η καθαρή εσωτερική δραστηριότητα (κατάσταση) του k νευρώνα και y_k είναι η έξοδος του. Συνήθως, τα BP έχουν υψηλή συνδεσιμότητα, δηλαδή έχουν μεγάλους αριθμούς συναπτικών βαρών (διασυνδέσεων). Ο αλγόριθμος BP περιλαμβάνει τη «φάση - προς - τα εμπρός» (ορθή φάση) και τη «φάση - προς - τα πίσω» (ανάστροφη φάση). Κατά τη «φάση - προς - τα εμπρός» (φάση περάσματος) ένα πρότυπο εισόδου ταξιδεύει προς την έξοδο «επίπεδο - με - επίπεδο» και όταν φθάσει σ' αυτή παράγει τις πραγματικές εξόδους. Κατά τη «φάση - προς - τα εμπρός» όλα τα συναπτικά βάρη κρατούν σταθερές τιμές. Κατά «φάση - προς - τα πίσω» το σφάλμα εξόδου διαδίδεται ανάστροφα από την έξοδο προς την είσοδο και τα συναπτικά βάρη όλων των νευρώνων ανανεώνονται μέσω του κανόνα διόρθωσης σφάλματος έτσι ώστε η πραγματική έξοδος να πλησιάσει όσο γίνεται πιο πολύ την επιθυμητή έξοδο του. Άρα ο όρος ανάστροφη διάδοση (Back Propagation) οφείλεται στον τρόπο με το οποίο διαδίδονται τα σφάλματα μέσω των επιπέδων του ΝΔ.

5 Σχεδιασμός Ταξινομητή

5.1 Μέθοδος 'One against all'

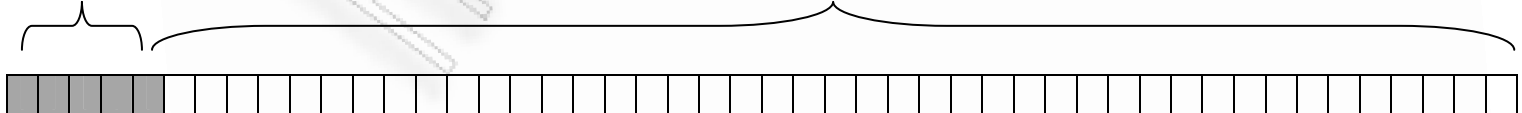
Στην πρώτη προσπάθειά μας αποφασίσαμε να χρησιμοποιήσουμε το 90% των δεδομένων για την εκπαίδευση του ταξινομητή και το άλλο 10% για τον έλεγχο του ταξινομητή. Αυτό έγινε ώστε τα δεδομένα τα οποία θα χρησιμοποιούσαμε στον έλεγχο του ταξινομητή να μην είναι τα ίδια με αυτά που θα χρησιμοποιούσαμε κατά την εκπαίδευση ώστε να πάρουμε αντικειμενικά αποτελέσματα.

Προκειμένου όμως να πάρουμε ακόμα πιο αντικειμενικά αποτελέσματα αποφασίσαμε τα δεδομένα ελέγχου να μην είναι από ένα συγκεκριμένο σημείο του συνόλου των δεδομένων αλλά από όλα τα σημεία του συνόλου των δεδομένων. Αφού λοιπόν θα χρησιμοποιούσαμε το 10% του συνόλου των δεδομένων για έλεγχο δημιουργήσαμε 10 διαφορετικές ομάδες δεδομένων προς εκπαίδευση τις οποίες από εδώ και στο εξής θα τις ονομάζουμε Folds (αν N είναι το πλήθος των δεδομένων τότε $N/10$ θα είναι το πλήθος των δεδομένων προς εκπαίδευση και $(9*N)/10$ θα είναι το πλήθος των δεδομένων εκπαίδευσης). Όπως φαίνεται από τα παραπάνω δημιουργούνται 10 διαφορετικά είδη δεδομένων για κάθε κλάση ανάλογα με το πιο σημείο επί του συνόλου θα γίνει η επιλογή των δεδομένων για έλεγχο.

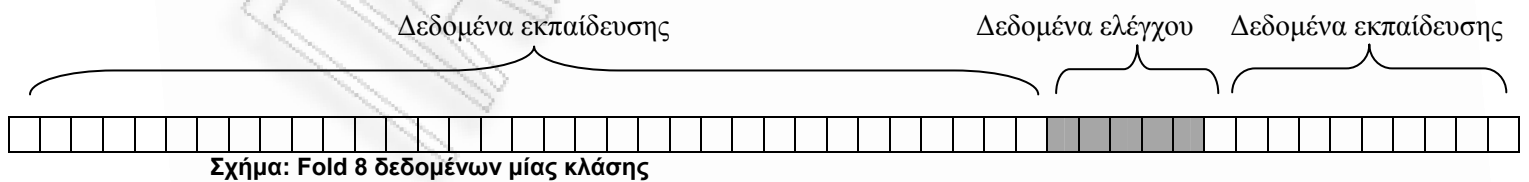
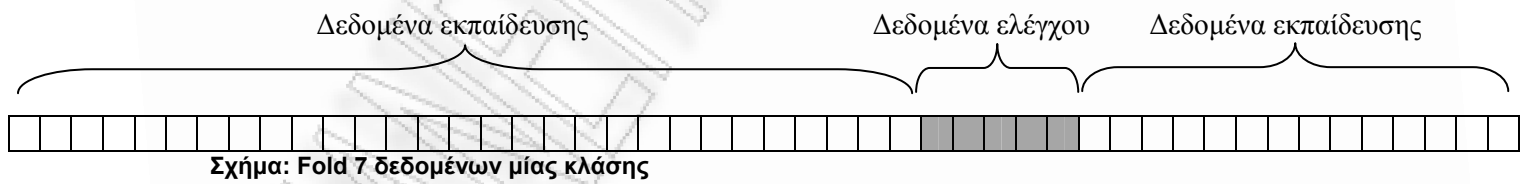
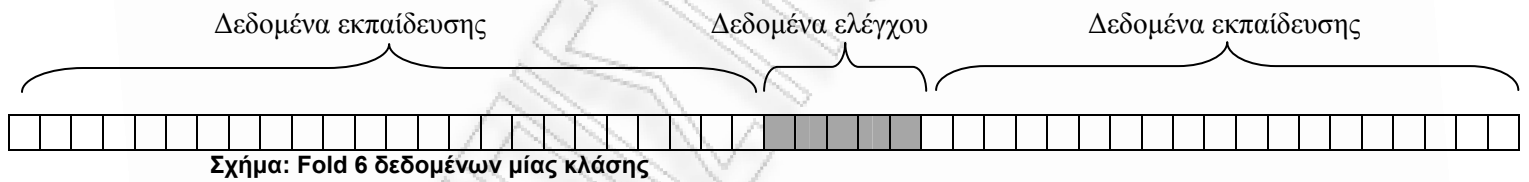
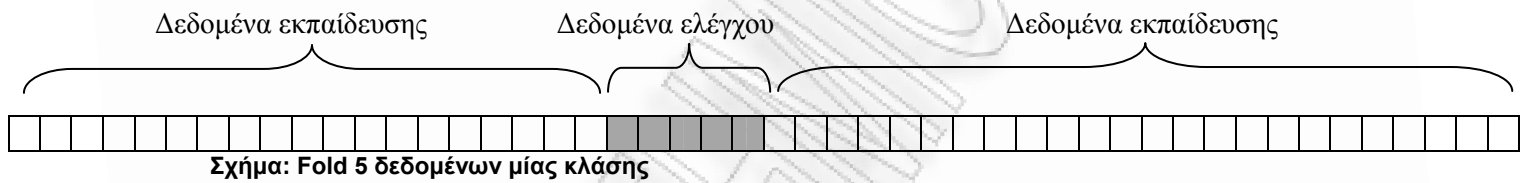
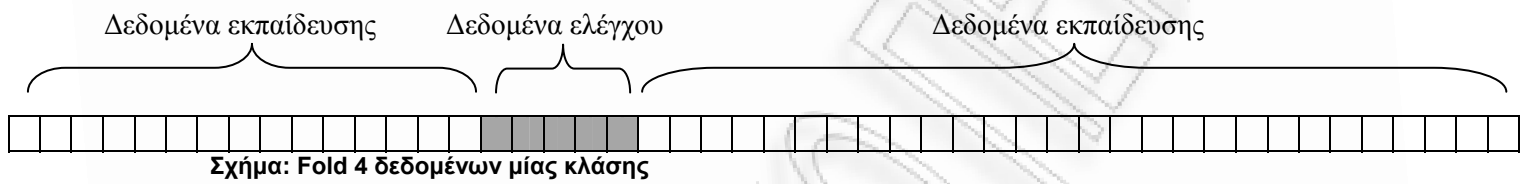
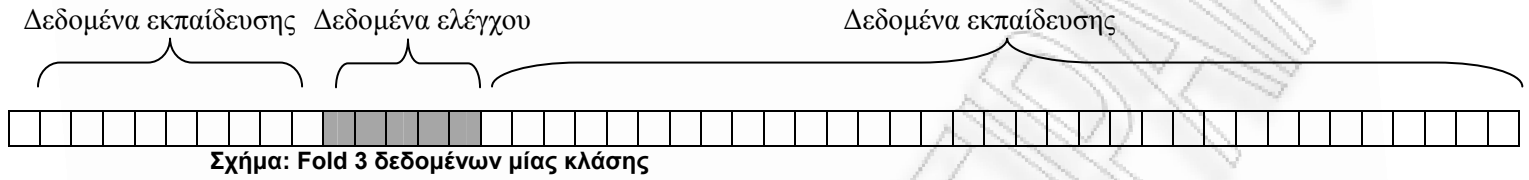
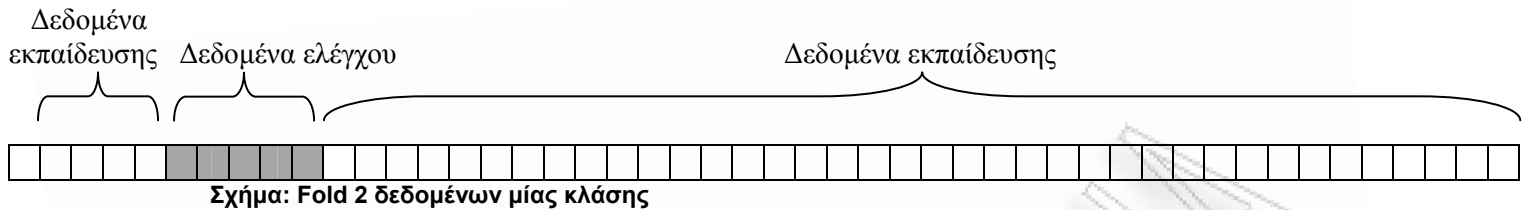
Όλα αυτά που περιγράψαμε παραπάνω γίνονται πιο φανερά στα παρακάτω διαγραμματικά σχήματα που ακολουθούν όπου παρουσιάζονται για κάθε κλάση ο τρόπος με τον οποίο λαμβάνονται τα δεδομένα για εκπαίδευση και για έλεγχο.

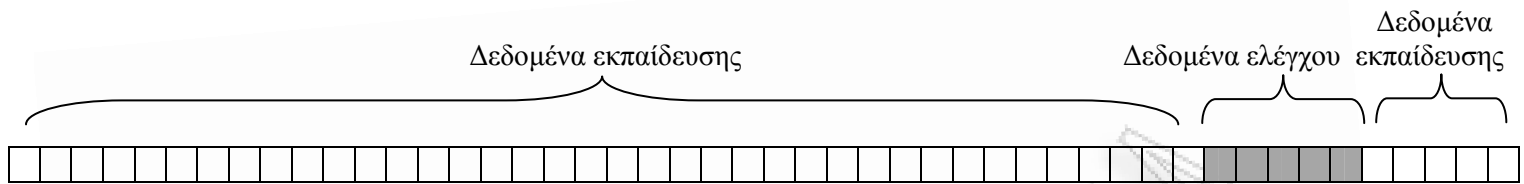
Δεδομένα ελέγχου

Δεδομένα εκπαίδευσης

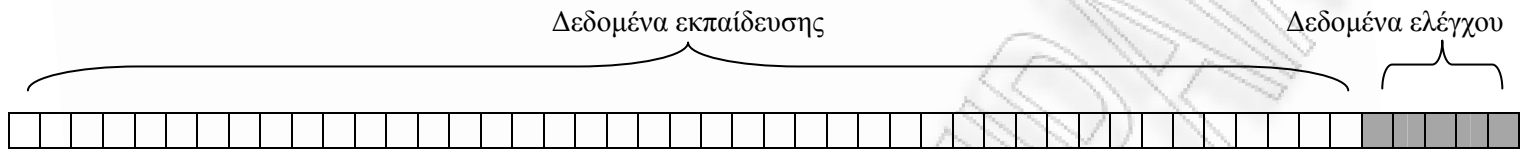


Σχήμα: Fold1 δεδομένων μίας κλάσης





Σχήμα: Fold 9 δεδομένων μίας κλάσης

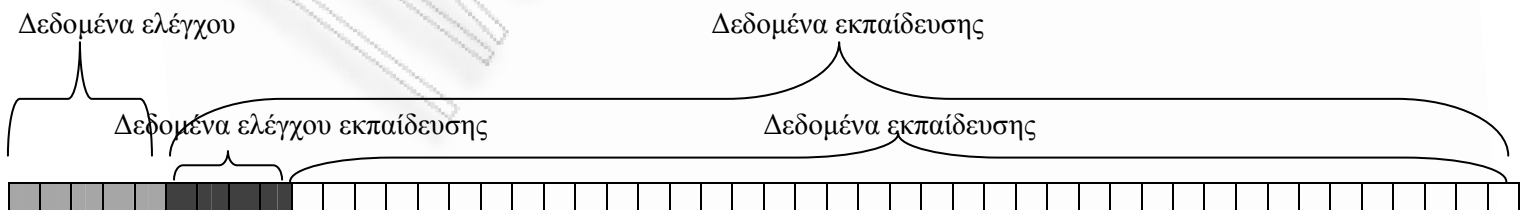


Σχήμα: Fold 10 δεδομένων μίας κλάσης

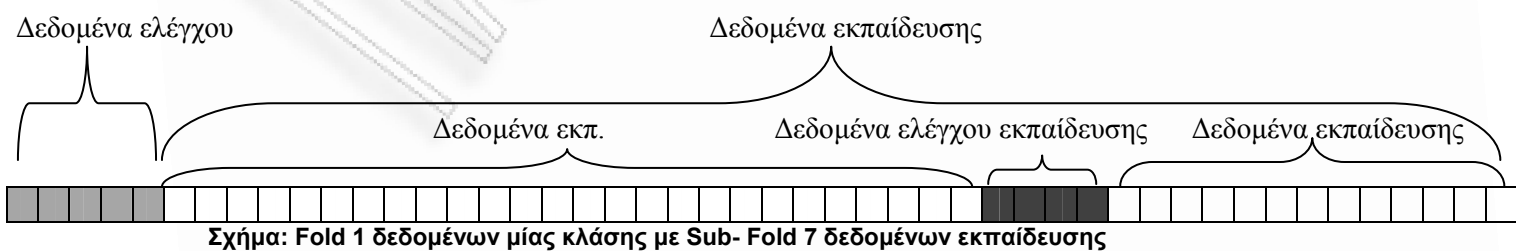
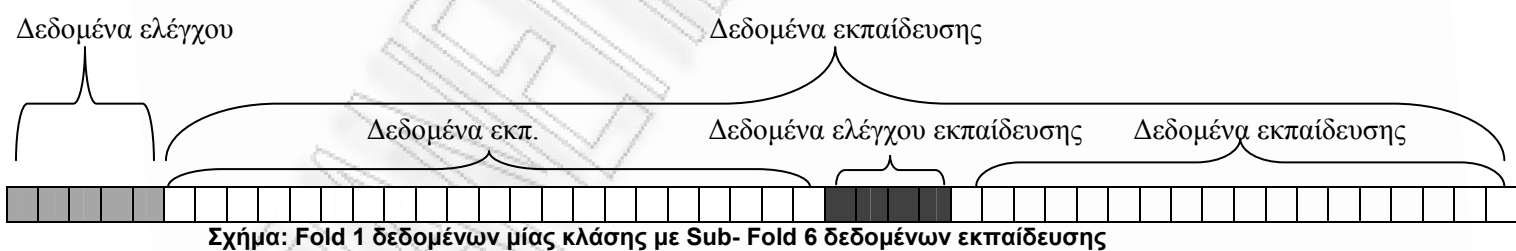
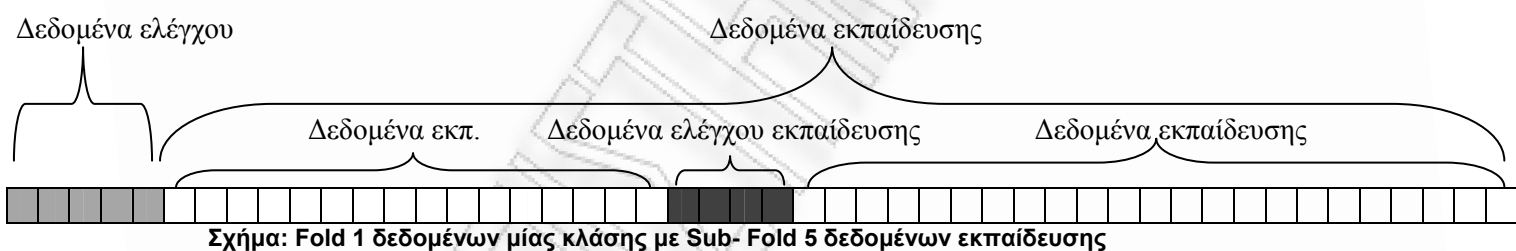
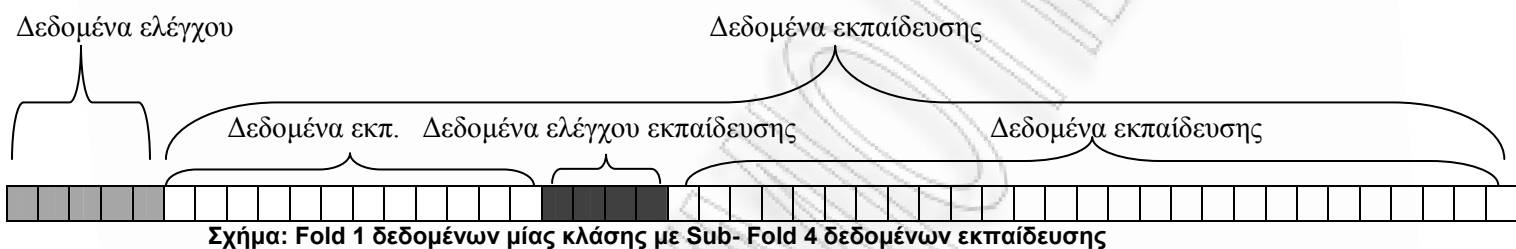
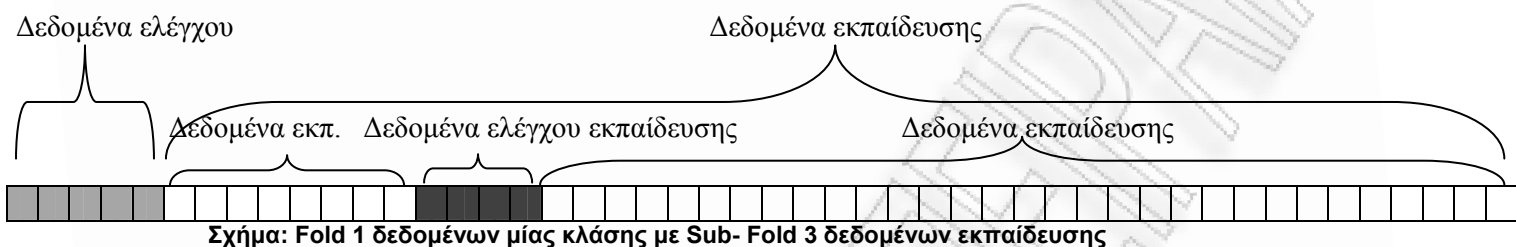
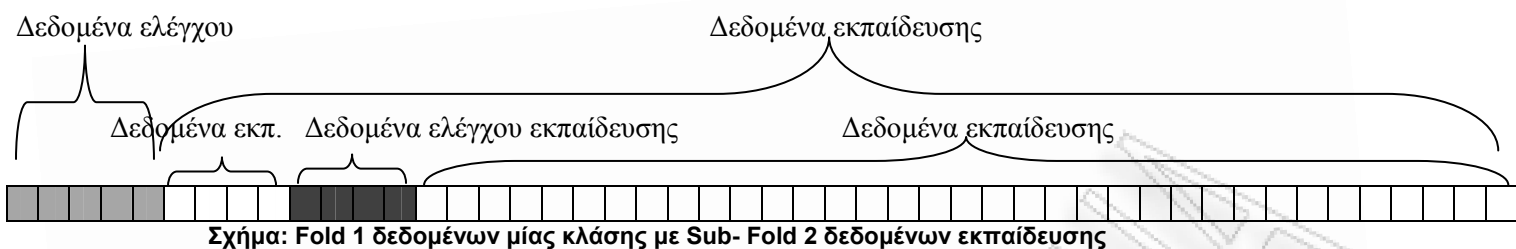
Αφού περιγράψαμε τον τρόπο με τον οποίο θα επιλεγθούν τα δεδομένα για έλεγχο και εκπαίδευση θα πρέπει τώρα να περιγράψουμε τον υπέρ-ταξινομητή. Η γενική ιδέα για τη δημιουργία του υπέρ-ταξινομητή είναι να δημιουργήσουμε τόσους ταξινομητές όσες είναι και οι κλάσεις στις οποίες πρέπει να ταξινομηθούν τα δεδομένα μας. Έτσι στη γενική περίπτωση που έχουμε M κλάσεις τότε θα πρέπει να δημιουργήσουμε M ταξινομητές. Τα θετικά δεδομένα εκπαίδευσης θα προέρχονται αντίστοιχα από την κάθε κλάση στην οποία αναφέρεται ο κάθε ταξινομητής ενώ τα αρνητικά δεδομένα από τις υπόλοιπες $M-1$ κλάσεις. Επειδή όπως εξηγήσαμε προηγουμένως θέλουμε να έχουμε μεγαλύτερη ακρίβεια στα αποτελέσματά μας, ανάλογα με το ποιο σημείο από το σύνολο των δεδομένων θα πάρουμε τα δεδομένα ελέγχου θα δημιουργηθεί και διαφορετικός ταξινομητής. Οπότε για την πρώτη κλάση θα δημιουργηθούν συνολικά 10 διαφορετικοί ταξινομητές όσες και οι διαφορετικές ομάδες δεδομένων. Συνολικά στη γενική περίπτωση M κλάσεων θα έχουμε $M*10$ ταξινομητές. Από τον έλεγχο των αποτελεσμάτων των $M*10$ ταξινομητών θα μπορούμε να επιλέξουμε σε ποια κλάση ανήκει το συγκεκριμένο στιγμιότυπο που δόθηκε στον υπέρ-ταξινομητή προς ταξινόμηση.

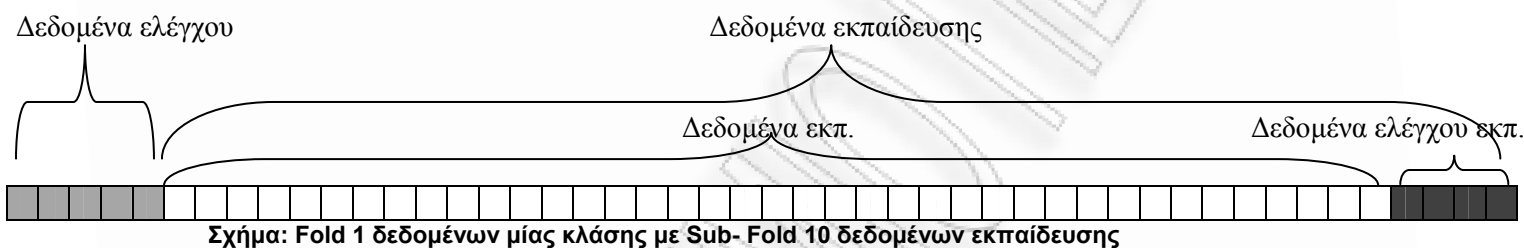
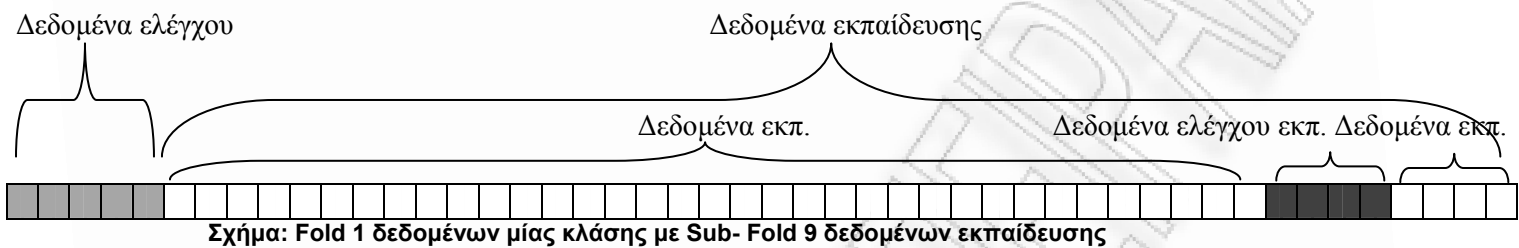
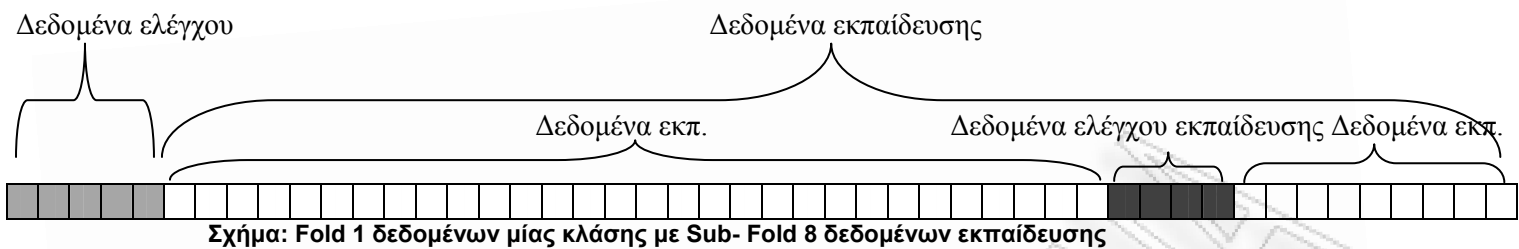
Προκειμένου να δημιουργήσουμε τον τελικό υπέρ-ταξινομητή ο οποίος θα ταξινομεί ένα στιγμιότυπο μεταξύ των κλάσεων θα χρησιμοποιήσουμε από κάθε ταξινομητή την τιμή της συνάρτησης απόφασης. Αυτή η τιμή για κάθε στιγμιότυπο προσδιορίζει κατά πόσο είναι σίγουρος ο κάθε ταξινομητής ότι έχει ταξινομήσει σωστά ένα στιγμιότυπο που του έχει δοθεί. Αν το στιγμιότυπο ανήκει στα θετικά δεδομένα τότε αυτή η τιμή θα έχει θετικό πρόσημο διαφορετικά θα έχει αρνητικό πρόσημο. Οπότε εμείς δεν έχουμε να κάνουμε τίποτε περισσότερο από το να επιλέξουμε τον ταξινομητή αυτόν, ο οποίος θα έχει δώσει τη μεγαλύτερη τιμή στην συνάρτηση απόφασης. Η σκέψη μας είναι ότι η τιμή της συνάρτησης απόφασης θα είναι μεγαλύτερη στον ταξινομητή που έχει εκπαιδευτεί να ταξινομεί στιγμιότυπα της συγκεκριμένης κλάσης.

Όμως για να πάρουμε σωστά αποτελέσματα κατά τον έλεγχο του υπέρ-ταξινομητή και να μην έχουμε επικαλύψεις μεταξύ των στιγμιότυπων κατά τον έλεγχο και την εκπαίδευση θα πρέπει να 'σπάσουμε' εκ νέου τα αρχικά στιγμιότυπα εκπαίδευσης που δημιουργήσαμε για ακόμη μία φορά σε δεδομένα εκπαίδευσης και ελέγχου. Αυτό γίνεται γιατί αν δε το κάνουμε, κατά τον έλεγχο του υπέρ-ταξινομητή ως δεδομένα ελέγχου θα παίρνουμε δεδομένα τα οποία έχουν ήδη χρησιμοποιηθεί για την εκπαίδευση του ταξινομητή της αντίστοιχης ομάδας. Οπότε θα λαμβάνουμε λανθασμένα σωστά αποτελέσματα αφού τα ίδια δεδομένα θα χρησιμοποιούνται στον ίδιο ταξινομητή τόσο για εκπαίδευση όσο και για έλεγχο. Σύμφωνα λοιπόν με αυτά που είπαμε για την κάθε κλάση τα δεδομένα που λαμβάνονται παίρνουν την παρακάτω μορφή όπως φαίνεται στα διαγράμματα:



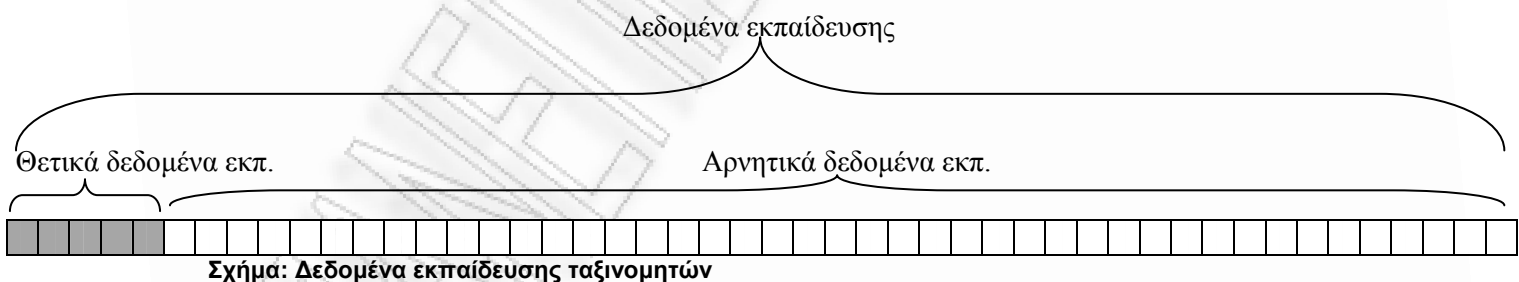
Σχήμα: Fold 1 δεδομένων μίας κλάσης με Sub-Fold 1 δεδομένων εκπαίδευσης





Από τα παραπάνω διαγράμματα γίνεται φανερό ότι δημιουργούνται για κάθε Fold της κάθε κλάσης 10 Sub-Folds δεδομένων εκπαίδευσης οπότε συνολικά θα έχουμε συνολικά 100 ταξινομητές για την κάθε κλάση δεδομένων. Στη γενική περίπτωση που έχουμε M κλάσεις δεδομένων και K Folds δεδομένων ανά κλάση θα δημιουργηθούν για κάθε κλάση K^2 διαφορετικοί ταξινομητές και συνολικά ο έλεγχος στον υπέρ-ταξινομητή θα γίνει σε $M \cdot K^2$ ταξινομητές για να έχουμε πιο αντικειμενικά αποτελέσματα.

Σύμφωνα λοιπόν με όσα περιγράψαμε τα δεδομένα εκπαίδευσης για κάθε ταξινομητή θα έχουν την παρακάτω μορφή όπως φαίνεται στο σχήμα:



Όπως έχουμε αναφέρει και προηγουμένως με τον όρο θετικά δεδομένα εννοούμε τα δεδομένα αυτά που προέρχονται από μία συγκεκριμένη κλάση ενώ αρνητικά τα δεδομένα που προέρχονται από όλες τις υπόλοιπες κλάσεις. Όλα αυτά τα δεδομένα αφού τους δοθούν οι κατάλληλες 'ετικέτες' ανάλογα αν είναι θετικά ή αρνητικά αντίστοιχα θα χρησιμοποιηθούν για να εκπαιδεύσουν τον ταξινομητή ο οποίος θα μπορεί στη συνέχεια να ελέγχει αν ένα στιγμιότυπο ανήκει ή δεν ανήκει στην κλάση από την οποία προέρχονται τα θετικά δεδομένα.

Στη γενική περίπτωση λοιπόν που έχουμε M κλάσεις και η κάθε κλάση έχει N στιγμιότυπα και επίσης έχουμε αποφασίσει ότι την κάθε κλάση θα τη χωρίσουμε σε K Folds τότε θα έχουμε N / K στιγμιότυπα για έλεγχο σε κάθε Fold και $(K - 1) \cdot N / K$ στιγμιότυπα για εκπαίδευση αντίστοιχα σε κάθε Fold. Πρέπει να σημειώσουμε εδώ ότι η διαίρεση N / K πρέπει να είναι τέλεια για να μπορέσουμε να συνεχίσουμε τη διαδικασία της εκπαίδευσης. Όμως αναφέραμε ότι τα στιγμιότυπα

εκπαίδευσης θα σπάσουν εκ νέου σε νέα Sub-Fold εκπαίδευσης δεδομένων και ελέγχου δεδομένων όπως εξηγήσαμε προηγουμένως για να μην έχουμε επικαλύψεις. Βέβαια δεν είναι απαραίτητο το πλήθος των Folds και των Sub-Folds να είναι το ίδιο. Αν θεωρήσουμε ότι αυτό το πλήθος είναι K' τότε θα πρέπει και πάλι η διαίρεση $(K-1) \cdot \frac{N}{K} / K'$ να είναι τέλεια. Επίσης το πλήθος των στιγμιότυπων που τελικά θα χρησιμοποιηθούν από κάθε κλάση για εκπαίδευση θα είναι:

$$((K-1) \cdot \frac{N}{K} / K') \cdot (K'-1)$$

Αυτό θα είναι και το πλήθος των θετικών στιγμιότυπων (δεδομένων) για κάθε ταξινομητή ενώ το πλήθος των αρνητικών στιγμιότυπων (δεδομένων) για κάθε ταξινομητή αν θεωρήσουμε ότι το συνολικό πλήθος των κλάσεων είναι M θα είναι:

$$\left[((K-1) \cdot \frac{N}{K} / K') \cdot (K'-1) \right] \cdot (M-1)$$

Συνολικά λοιπόν για την εκπαίδευση του κάθε ταξινομητή θετικά και αρνητικά δεδομένα θα χρησιμοποιηθούν: $\left[((K-1) \cdot \frac{N}{K} / K') \cdot (K'-1) \right] \cdot M$ στιγμιότυπα.

Αφού λοιπόν δημιουργήσαμε τα δεδομένα για εκπαίδευση των επιμέρους ταξινομητών δεν μένει τίποτε άλλο από το να τους εκπαιδεύσουμε με αυτά. Στη συνέχεια όλοι αυτοί οι ταξινομητές θα εκλεχθούν με τα δεδομένα ελέγχου. Κατά τον έλεγχο εμείς γνωρίζουμε σε ποια κλάση ανήκουν τα δεδομένα ελέγχου. Σκοπός μας είναι να διαπιστώσουμε πόσο καλά δουλεύει ο υπέρ-ταξινομητής και πόσο καλά ταξινομεί τα άγνωστα για αυτόν στιγμιότυπα κάθε κλάσης.

Όπως είπαμε και παραπάνω η λογική του υπέρ-ταξινομητή είναι να κατατάξει ένα στιγμιότυπο μίας κλάσης σύμφωνα με την τιμή της συνάρτησης απόφασης. Πιο συγκεκριμένα ένα στιγμιότυπο ελέγχεται από όλους τους ταξινομητές του υπέρ-ταξινομητή που έχουν δημιουργηθεί αντίστοιχα για κάθε κλάση δεδομένων. Κατά τον έλεγχο (testing) του συγκεκριμένου στιγμιότυπου δημιουργούνται κάποια αποτελέσματα που μας δείχνουν πως το κάθε στιγμιότυπο συμπεριφέρεται όταν περνάει από τον κάθε ένα ταξινομητή. Ένα από αυτά τα αποτελέσματα είναι και αυτό της συνάρτησης απόφασης. Αυτός ο συντελεστής που παίρνει τιμές από -1 έως 1 προσδιορίζει αν ένα στιγμιότυπο ανήκει ή δεν ανήκει στην κλάση για την οποία έχει δημιουργηθεί ο αντίστοιχος ταξινομητής που αυτό 'περνάει'. Αν η τιμή της συνάρτησης απόφασης είναι θετική τότε αυτό το στιγμιότυπο θεωρείται ότι ανήκει στην κλάση που εκφράζει ο ταξινομητής. Αν είναι αρνητική αυτή η τιμή τότε δεν ανήκει. Βέβαια ο ταξινομητής μπορεί να βγάλει και λανθασμένα αποτελέσματα αν τα δεδομένα με τα οποία έχει εκπαιδευτεί δεν είναι επαρκή ή λανθασμένα.

Επίσης η τιμή της συνάρτησης απόφασης εκτός από την κατάταξη του στιγμιότυπου σε μία κλάση ή όχι εκφράζει και τη βεβαιότητα με την οποία ένα στιγμιότυπο ανήκει ή δεν ανήκει σε αυτή την κλάση. Αν οι τιμές της συνάρτησης απόφασης βρίσκονται στα άκρα του διαστήματος $[-1, 1]$ τότε ο ταξινομητής δηλώνει τη βεβαιότητά του ότι το στιγμιότυπο ανήκει ή δεν ανήκει στην κλάση. Αντίθετα αν η τιμή βρίσκεται κοντά στο κέντρο του κλειστού αυτού διαστήματος, δηλαδή λίγο πάνω ή λίγο κάτω από το 0, τότε ο ταξινομητής δεν είναι και τόσο σίγουρος για αυτή του την απόφαση. Ένα ακόμα στοιχείο που προκύπτει αρκετά συχνά είναι το γεγονός ότι δύο διαφορετικοί ταξινομητές οι οποίοι έχουν εκπαιδευτεί με δεδομένα διαφορετικών κλάσεων και φυσικά με την ιδιότητα να κατατάσσουν τα αντίστοιχα στιγμιότυπά τους στις κλάσεις τους, μπορεί να δίνουν θετική τιμή στη συνάρτηση απόφασης για το ίδιο στιγμιότυπο. Φυσικά το στιγμιότυπο αυτό δεν μπορεί να ανήκει και στις δύο κλάσεις ταυτόχρονα. Η επιλογή σε αυτή την περίπτωση γίνεται με βάση την τιμή της συνάρτησης απόφασης. Η μεγαλύτερη τιμή, η οποία δίνει και τη μεγαλύτερη βεβαιότητα, επιλέγεται οπότε αντίστοιχα επιλέγεται και ο ταξινομητής που εκφράζει την αντίστοιχη κλάση δεδομένων. Έτσι λοιπόν λειτουργεί ο υπέρ-ταξινομητής και κατατάσσονται τα δεδομένα ελέγχου στις αντίστοιχες κλάσεις. Κατά πόσο καλά λειτουργεί αυτός ο μηχανισμός θα παρουσιαστεί στο κεφάλαιο με τα αποτελέσματα των πειραμάτων που έχουν γίνει από δύο είδη δεδομένων στους αντίστοιχους υπέρ-ταξινομητές που έχουμε δημιουργήσει.

5.2 Μέθοδος 'One against all balanced'

Κατά τη δεύτερη προσπάθειά μας η γενική ιδέα για τη δημιουργία του υπέρ-ταξινομητή είναι η ίδια με αυτήν που εφαρμόσαμε και προηγουμένως. Έτσι λοιπόν σαν σκοπό έχουμε πάλι να δημιουργήσουμε τόσους ταξινομητές όσες είναι και οι κλάσεις στις οποίες πρέπει να ταξινομηθούν τα δεδομένα μας. Στη γενική περίπτωση που έχουμε M κλάσεις τότε θα πρέπει να δημιουργήσουμε M ταξινομητές. Τα

θετικά δεδομένα εκπαίδευσης θα προέρχονται αντίστοιχα από την κάθε κλάση στην οποία αναφέρεται ο κάθε ταξινομητής ενώ τα αρνητικά δεδομένα από τις υπόλοιπες $M-1$ κλάσεις. Επειδή όμως θέλουμε να έχουμε μεγαλύτερη ακρίβεια στα αποτελέσματά μας, ανάλογα με το ποιο σημείο από το σύνολο των δεδομένων θα πάρουμε τα δεδομένα ελέγχου θα δημιουργηθεί και διαφορετικός ταξινομητής. Οπότε για την πρώτη κλάση θα δημιουργηθούν συνολικά 10 διαφορετικοί ταξινομητές όσες και τα διαφορετικά Folds δεδομένων. Συνολικά στη γενική περίπτωση M κλάσεων θα έχουμε $M \cdot K$ ταξινομητές αν θεωρήσουμε ότι το πλήθος των Folds δεδομένων για κάθε κλάση είναι K . Από τον έλεγχο των αποτελεσμάτων των $M \cdot K$ ταξινομητών θα μπορούμε να επιλέξουμε σε ποια κλάση ανήκει το συγκεκριμένο στιγμιότυπο που δόθηκε στον υπέρ-ταξινομητή προς ταξινόμηση.

Για τα μουσικά δεδομένα έχουμε 10 διαφορετικά είδη μουσικής συνεπώς και 10 διαφορετικές κλάσεις. Με άλλα λόγια το M είναι 10. Επίσης δημιουργούμε για κάθε κλάση 10 Folds συνεπώς και το K είναι 10. Οπότε θα δημιουργηθούν $10 \cdot 10 = 100$ διαφορετικοί ταξινομητές. 10 ταξινομητές για κάθε κλάση δεδομένων. Επίσης έχουμε πει ότι η κάθε κλάση αποτελείται από 100 στιγμιότυπα οπότε έχουμε ότι το N είναι ίσο με 100.

Η διαφορά της πρώτης προσπάθειας με την παρούσα είναι ότι τώρα κάθε ταξινομητής θα εκπαιδευτεί από ισοζυγισμένα δεδομένα και όχι από μη ισοζυγισμένα. Δηλαδή το πλήθος των θετικών όσο και των αρνητικών δεδομένων εκπαίδευσης θα είναι ίδιο. Αυτό το κάνουμε για να εκπαιδευτεί ο κάθε ταξινομητής της αντίστοιχης κλάσης πιο σωστά και να μπορεί να προβλέπει το ίδιο σωστά τόσο για θετικά όσο και για αρνητικά στιγμιότυπα. Εμείς κατά κύριο λόγο ενδιαφερόμαστε για τα θετικά στιγμιότυπα αφού αυτά καθορίζουν για ποια κλάση έχει δημιουργηθεί ο αντίστοιχος ταξινομητής, ενώ τα αρνητικά προέρχονται από όλες τις άλλες κλάσεις.

Πρέπει να σημειώσουμε επίσης ότι και εδώ όπως και στην προηγούμενη μέθοδο για να μη δημιουργηθεί το φαινόμενο της επικάλυψης και των λανθασμένα σωστών αποτελεσμάτων θα πρέπει τα δεδομένα μας να τα σπάσουμε πάλι σε Folds και Sub-Folds. Οπότε δεν υπάρχει κάτι να αλλάξει σε αυτό το κομμάτι. Έτσι το πλήθος των στιγμιότυπων που θα χρησιμοποιηθούν από κάθε κλάση για εκπαίδευση θα είναι:

$$((K-1) \cdot \frac{N}{K} / K') \cdot (K'-1)$$

Όπου K είναι το πλήθος των Folds, K' είναι το πλήθος των Sub-Folds και N είναι το πλήθος των στιγμιότυπων που περιέχει κάθε κλάση. Οπότε για την εκπαίδευση του κάθε ταξινομητή τελικά το πλήθος των στιγμιότυπων που θα χρησιμοποιηθούν είναι:

$$2 \cdot ((K-1) \cdot \frac{N}{K} / K') \cdot (K'-1)$$

Ενώ το πλήθος των στιγμιότυπων που θα χρησιμοποιηθούν κατά τον τελικό έλεγχο του υπέρ-ταξινομητή θα είναι:

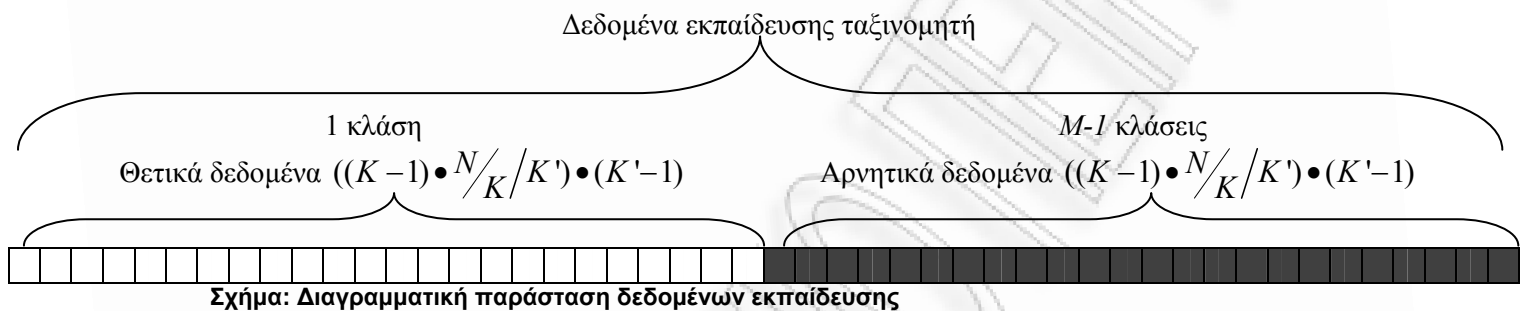
$$M \cdot \left(\frac{N}{K}\right)$$

Όπου M είναι το πλήθος των κλάσεων όπως περιγράψαμε στην προηγούμενη ενότητα

Για να γίνει αυτό πιο κατανοητό ο τρόπος με τον οποίο θα γίνει κυρίως η επιλογή των αρνητικών δεδομένων ας θεωρήσουμε το παρακάτω παράδειγμα. Έστω ότι θέλουμε να δημιουργήσουμε τον ταξινομητή 3 μιας ομάδας (η ομάδα δεν παίζει ρόλο στην παρούσα φάση αφού τα δεδομένα θα έχουν ήδη διαχωριστεί σε δεδομένα ελέγχου και σε δεδομένα εκπαίδευσης). Τότε δεν έχουμε τίποτε περισσότερο να κάνουμε από το να επιλέξουμε ως θετικά δεδομένα τα δεδομένα εκπαίδευσης που ανήκουν στην κλάση 3 με τον τρόπο που περιγράψαμε. Από την άλλη πλευρά ως αρνητικά δεδομένα θα επιλέξουμε τα δεδομένα εκπαίδευσης από όλες τις υπόλοιπες κλάσεις της αντίστοιχης ομάδας. Το λεπτό σημείο στην φάση αυτή είναι ότι η επιλογή των στιγμιότυπων θα είναι σε πλήθος τόσα όσα και τα θετικά δεδομένα εκπαίδευσης. Οπότε θα πρέπει να επιλέξουμε $((K-1) \cdot \frac{N}{K} / K') \cdot (K'-1)$ 'αρνητικά' στιγμιότυπα από τα συνολικά $\left[((K-1) \cdot \frac{N}{K} / K') \cdot (K'-1) \right] \cdot (M-1)$ όσα είναι και τα θετικά στιγμιότυπα εκπαίδευσης. Αφού όμως έχουμε $M-1$ 'αρνητικές' κλάσεις θα επιλέξουμε $\left[((K-1) \cdot \frac{N}{K} / K') \cdot (K'-1) \right] / (M-1)$ στιγμιότυπα από κάθε 'αρνητική' κλάση (ως αρνητικές κλάσεις μπορούμε να θεωρήσουμε όλες τις κλάσεις εκτός από αυτήν που ο ταξινομητής για την εκπαίδευσή του παίρνει θετικά δεδομένα εκπαίδευσης). Η επιλογή των 'αρνητικών' στιγμιότυπων από τα στιγμιότυπα προς εκπαίδευση που έχει η κάθε 'αρνητική' κλάση αποφασίσαμε να γίνεται με τυχαίο τρόπο και όχι με κάποιο συγκεκριμένη διάταξη (το πλήθος που τελικά χρειαζόμαστε από κάθε

αρνητική κλάση είναι πολύ μικρότερο από αυτό που διαθέτουμε). Αυτό είναι ίσως και το πιο σωστό για να μην υπάρξει κάποιο μεροληπτικό λάθος κατά την εκπαίδευση. Πρέπει να σημειώσουμε ότι αν η τελευταία διαίρεση δεν είναι τέλεια τότε κάποιες κλάσεις μπορεί να δώσουν ένα περισσότερο στιγμιότυπο από τις υπόλοιπες. Αυτό αν συμβεί θα είναι για τα αρνητικά δεδομένα τόσο εκπαίδευσης όσο και ελέγχου. Επίσης για να μην υπάρχει κάποιο πρόβλημα στη σωστή εκπαίδευση των ταξινομητών οι κλάσεις οι οποίες αν τύχει να δώσουν ένα περισσότερο στιγμιότυπο στην εκπαίδευση ή στον έλεγχο του ταξινομητή δεν θα είναι σταθερές αλλά με κάποιο κυκλικό αλγόριθμο θα εναλλάσσονται μεταξύ του συνόλου των κλάσεων. Όμως όπως αναφέραμε και στην προηγούμενη ενότητα οι διαιρέσεις N/K και $(K-1) \cdot N/K/K'$ πρέπει να είναι τέλειες.

Στο παρακάτω σχήμα φαίνεται διαγραμματικά η μορφή των δεδομένων που θα χρησιμοποιηθούν για την εκπαίδευση των αντίστοιχων ταξινομητών.



5.3 Μέθοδος ‘One against all balanced with weight’

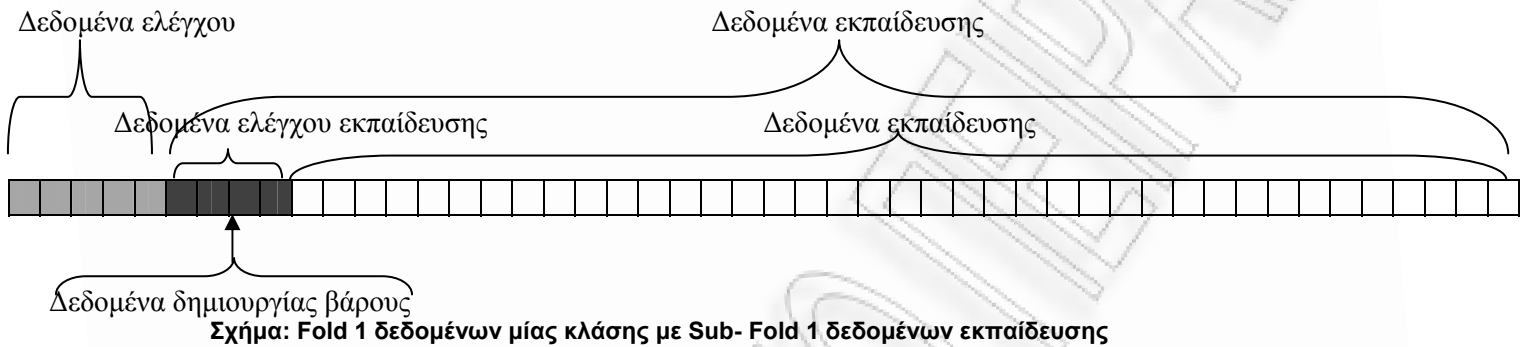
Στην προηγούμενη σχεδίαση του υπέρ-ταξινομητή για να έχουμε καλύτερα αποτελέσματα αποφασίσαμε κατά την εκπαίδευση του κάθε ταξινομητή που αντιστοιχεί σε μία συγκεκριμένη κλάση να χρησιμοποιήσουμε τον ίδιο αριθμό στιγμιότυπων τόσο για τα θετικά όσο και για τα αρνητικά δεδομένα. Όπως θα παρουσιάσουμε σε επόμενο κεφάλαιο με τα αποτελέσματα των πειραμάτων θα φανεί ότι αυτό είχε θετικό αποτέλεσμα.

Εδώ θα προχωρήσουμε ένα ακόμα βήμα παραπέρα. Είπαμε ότι η γενική ιδέα του υπέρ-ταξινομητή είναι κάθε στιγμιότυπο μίας κλάσης να ελέγχεται από όλους τους ταξινομητές που έχουν σχεδιαστεί για κάθε κλάση αντίστοιχα και ανάλογα με την τιμή της συνάρτησης απόφασης το συγκεκριμένο στιγμιότυπο να κατατάσσεται στον αντίστοιχο ταξινομητή άρα και στην αντίστοιχη κλάση που αυτός αντιπροσωπεύει. Το πρόβλημα που συναντήσαμε είναι όταν δύο διαφορετικοί ταξινομητές δημιουργημένοι να αναγνωρίζουν διαφορετικές κλάσεις δεδομένων δίνουν και οι δύο θετική τιμή στη συνάρτηση απόφασης, για το ίδιο στιγμιότυπο. Τότε είπαμε ότι θα επιλέγεται ο ταξινομητής που θα έχει δώσει τη μεγαλύτερη τιμή. Βέβαια και πάλι δημιουργούνται προβλήματα γιατί αρκετά συχνά ταξινομητές που είναι εκπαιδευμένοι να κατατάσσουν συγκεκριμένα δεδομένα σε συγκεκριμένες κλάσεις κατατάσσουν και άλλα δεδομένα στις κλάσεις τους δίνοντας καλύτερες τιμές στη συνάρτηση απόφασης απ’ ότι ο ταξινομητής που έχει εκπαιδευτεί για αυτήν την κλάση. Αυτό που θα μπορούσε να βοηθήσει τον τελικό υπέρ-ταξινομητή να δίνει ακόμα καλύτερα αποτελέσματα είναι να μπορούσε να δημιουργηθεί ένα ‘βάρος’ σε κάθε επιμέρους ταξινομητή αντίστοιχα με την αξιοπιστία του κάθε επιμέρους ταξινομητή. Αυτό το ‘βάρος’ θα μπορούσε να πολλαπλασιαστεί στη συνέχεια με την τιμή της συνάρτησης απόφασης που θα έδινε ο κάθε ταξινομητής για ένα δεδομένο στιγμιότυπο ώστε τελικά να έχουμε καλύτερα αποτελέσματα.

Η λογική που διέπει αυτή τη σκέψη είναι ότι με τη χρήση του βάρους ο υπέρ-ταξινομητής θα μπορεί να δώσει πιο ασφαλή αποτελέσματα. Αυτό θα συμβεί γιατί για ένα δοσμένο στιγμιότυπο το οποίο θέλουμε να το κατατάξουμε σε μία κλάση ένας ταξινομητής διαφορετικός από αυτόν της κλάσης του μπορεί να δίνει καλύτερα αποτελέσματα από αυτόν τον ταξινομητή που πραγματικά ανήκει, όμως αν υπάρχει η χρήση ‘βάρους’ τότε αυτός ο λανθασμένος ταξινομητής θα έχει μικρότερο βάρος από τον αληθινή ταξινομητή οπότε η τελική τιμή της συνάρτησης απόφασης γίνει μεγαλύτερη για τον σωστό ταξινομητή. Όλο αυτό βέβαια στηρίζεται στο γεγονός ότι οι ‘κακοί’ ταξινομητές έχουν μικρά βάρη ενώ οι ‘καλοί’ μεγάλα. Το τελευταίο ερώτημα είναι πως θα δημιουργηθούν αυτά τα βάρη.

Σε αυτόν τον σχεδιασμό του υπέρ-ταξινομητή δεν θα αλλάζουμε σχεδόν τίποτα από αυτά που μέχρι στιγμής έχουμε κάνει. Τα δεδομένα θα ‘σπάσουν’ έτσι όπως ακριβώς έχουμε περιγράψει και στις

δύο προηγούμενες περιπτώσεις. Επίσης κατά την εκπαίδευση του κάθε ταξινομητή θα χρησιμοποιηθούν 'ισοζυγισμένα' δεδομένα. Δηλαδή τόσο τα αρνητικά όσο και τα θετικά δεδομένα θα είναι του ίδιου πλήθους. Το μόνο που θα αλλάξει είναι ότι εκτός από τον τελικό έλεγχο σε κάθε ταξινομητή θα εκτελεστεί και ένας προ-έλεγχος ο οποίος θα καθορίσει και το βάρος για τον κάθε ταξινομητή. Αυτά τα δεδομένα θα είναι διαφορετικά για κάθε ταξινομητή και δεν θα έχουν καμία σχέση με τα δεδομένα για τον τελικό έλεγχο του υπέρ-ταξινομητή. Αυτό είναι αυτονόητο για να μην υπάρξουν επικαλύψεις και λανθασμένα αποτελέσματα. Επίσης κάθε ταξινομητής της ίδιας κλάσης δεδομένων διαφορετικού Sub-Fold πάλι θα έχει διαφορετικά δεδομένα προ-ελέγχου. Στο παρακάτω σχήμα φαίνονται τα δεδομένα προ-ελέγχου:



Το πλήθος των δεδομένων προ-ελέγχου του κάθε ταξινομητή ισούται με:

$$(K - 1) \cdot \frac{N}{K} / K'$$

όπου K είναι το πλήθος των Folds, K' είναι το πλήθος των Sub-Folds και N είναι το πλήθος των στιγμιότυπων της κάθε κλάσης.

Τα υπόλοιπα στοιχεία για κάθε κλάση και τον αντίστοιχο ταξινομητή είναι όπως τα έχουμε και νωρίτερα παρουσιάσει. Δηλαδή το πλήθος των δεδομένων εκπαίδευσης για κάθε ταξινομητή είναι:

$$2 \cdot ((K - 1) \cdot \frac{N}{K} / K') \cdot (K' - 1)$$

και το πλήθος των δεδομένων του τελικού ελέγχου είναι:

$$M \cdot \left(\frac{N}{K}\right)$$

όπου M είναι το πλήθος των διαφορετικών κλάσεων.

Πρέπει να σημειωθεί σε αυτό το σημείο ότι το βάρος στην ουσία αποτελεί το rate που δημιουργείται μετά τον έλεγχο του κάθε επιμέρους ταξινομητή με τα αντίστοιχα δεδομένα προ-ελέγχου. Δηλαδή πόσο σωστά ο κάθε επιμέρους ταξινομητής ταξινομεί τα δοσμένα στιγμιότυπα. Αυτός ο συντελεστής αποθηκεύεται μαζί με τα υπόλοιπα στοιχεία του κάθε ταξινομητή ώστε όταν φτάσουμε στον τελικό έλεγχο να είμαστε σε θέση να τον χρησιμοποιήσουμε. Στο κεφάλαιο με τα πειράματα θα παρουσιάσουμε και τα αποτελέσματα αυτού του υπέρ-ταξινομητή και πόσο καλά αυτός λειτουργεί σε σχέση με τους άλλους δύο που μέχρι στιγμής έχουμε περιγράψει.

5.4 Μέθοδος 'One against all balanced with double passing'

Στην τελευταία σχεδίαση του υπέρ-ταξινομητή προχωράμε ένα ακόμα βήμα παραπάνω προκειμένου να πετύχουμε ακόμα καλύτερα αποτελέσματα. Θα στηριχθούμε και πάλι στην τελευταία σχεδίασή του όπως περιγράψαμε λίγο νωρίτερα. Σε αυτό το στάδιο λοιπόν του σχεδιασμού θα προσπαθήσουμε να εκμεταλλευτούμε όχι απλά το 'rate' που δημιουργείται κατά τον προ-έλεγχο του κάθε ταξινομητή χωριστά αλλά τις ίδιες τις τιμές της συνάρτησης απόφασης που προκύπτουν από αυτόν τον προ-έλεγχο για τον κάθε ταξινομητή. Η σκέψη μας είναι να εκπαιδεύσουμε σε ένα δεύτερο στάδιο ένα νέο ταξινομητή με αυτές τις τιμές της συνάρτησης απόφασης ο οποίος τελικά να μας δώσει πολύ πιθανόν ακόμα καλύτερα αποτελέσματα. Θα πρέπει συνεπώς να δημιουργήσουμε πάλι με αυτά τα δεδομένα (τιμές συνάρτησης απόφασης) τόσες κλάσεις όσες είναι και οι αρχικές, να τα ξεχωρίσουμε σωστά ώστε να μη δημιουργηθεί κανένα λάθος, να τον εκπαιδεύσουμε και τέλος να τον ελέγξουμε με τα

δεδομένα ελέγχου που έχουμε δημιουργήσει όπως έχουμε αναφέρει. Στη συνέχεια θα περιγράψουμε αναλυτικά όλη αυτή τη διαδικασία.

Όπως έχουμε ήδη πει το πλήθος των στιγμιότυπων που χρησιμοποιούμε για τον προ-έλεγχο του κάθε ταξινομητή από κάθε κλάση είναι:

$$(K - 1) \cdot N / K / K'$$

όπου K είναι το πλήθος των Folds, K' είναι το πλήθος των Sub-Folds και N είναι το πλήθος των στιγμιότυπων της κάθε κλάσης.

Αυτό όμως στην πραγματικότητα είναι το πλήθος των θετικών στιγμιότυπων, δηλαδή των στιγμιότυπων που προέρχονται από την κλάση για την οποία έχει δημιουργηθεί ο κάθε ταξινομητής να αναγνωρίζει. Κατά τον έλεγχο όμως ενός ταξινομητή θέλουμε και στιγμιότυπα και από τις υπόλοιπες κλάσεις να δούμε πως ο ταξινομητής συμπεριφέρεται και με δεδομένα για τα οποία δεν είναι εκπαιδευμένος να τα κατατάσσει στην κλάση του. Πρέπει να σημειώσουμε εδώ, για να είμαστε πιο ακριβείς στην περιγραφή μας, ότι κατά την εκπαίδευση του κάθε ταξινομητή στα αρνητικά δεδομένα τα στιγμιότυπα προέρχονται από όλες τις υπόλοιπες κλάσεις, άρα στην πραγματικότητα ο ταξινομητής έχει εκπαιδευτεί να κατατάσσει και τα αρνητικά δεδομένα. Πιο αναλυτικά ο κάθε ταξινομητής εκπαιδεύεται με τέτοιο τρόπο έτσι ώστε να μπορεί να απαντά αν ένα δοσμένο στιγμιότυπο μετά τον έλεγχο του ανήκει στην κλάση για την οποία έχει εκπαιδευτεί ή ανήκει σε μία από τις υπόλοιπες $M-1$ κλάσεις χωρίς να ξέρουμε βέβαια σε ποια ακριβώς, αν θεωρήσουμε ότι M είναι το πλήθος των κλάσεων που έχουμε.

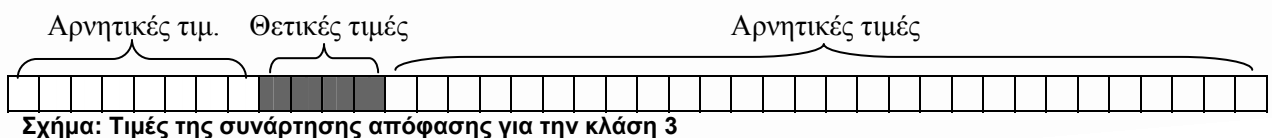
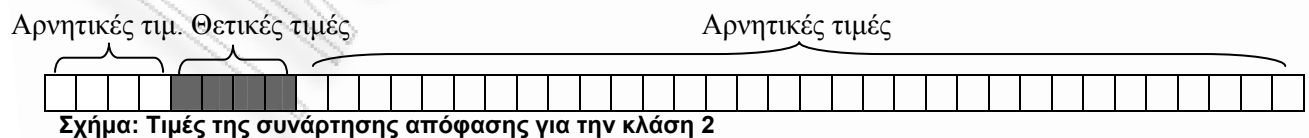
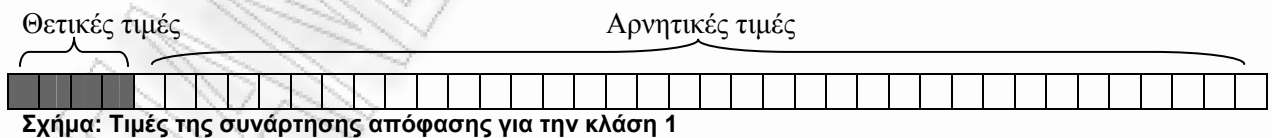
Οπότε σύμφωνα με τα παραπάνω το συνολικό πλήθος των στιγμιότυπων κατά τον προ-έλεγχο θα είναι:

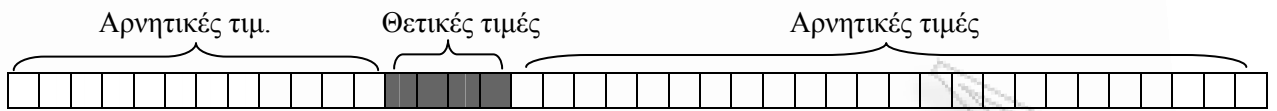
$$\left[(K - 1) \cdot N / K / K' \right] \cdot M$$

Τόσο βέβαια θα είναι και το πλήθος των τιμών της συνάρτησης απόφασης για ένα συγκεκριμένο ταξινομητή που έχει δημιουργηθεί για μία συγκεκριμένη κλάση και για ένα συγκεκριμένο Sub-Fold. Κάθε στιγμιότυπο δίνει και μία τιμή στη συνάρτηση απόφασης του ταξινομητή.

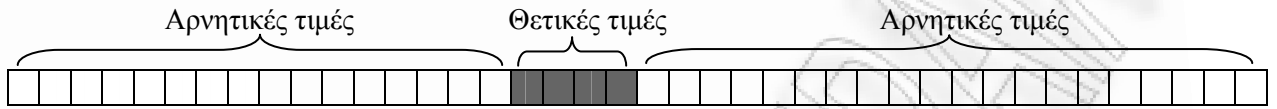
Στην ιδανική περίπτωση όπου ο κάθε ταξινομητής είναι τέλεια εκπαιδευμένος και αναγνωρίζει τέλεια τα δεδομένα ελέγχου που του δίνονται τότε θα δίνει: $(K - 1) \cdot N / K / K'$ θετικές τιμές και $\left[(K - 1) \cdot N / K / K' \right] \cdot (M - 1)$ αρνητικές τιμές. Επίσης αυτές οι τιμές θα πρέπει να αντιστοιχούν ακριβώς στα στιγμιότυπα δεδομένων με τη σειρά που δίνονται κατά τη διαδικασία του ελέγχου του κάθε ταξινομητή και θα μας εμφανίζονται με τη μορφή ενός μονοδιάστατου πίνακα $\left[1 \times \left(\left[(K - 1) \cdot N / K / K' \right] \cdot M \right) \right]$.

Στην πραγματικότητα βέβαια επειδή τα πράγματα δεν είναι τόσο ιδανικά όσο θα θέλαμε μεταξύ των θετικών και των αρνητικών τιμών συνάρτησης απόφασης θα παρεισφρήσουν και αρνητικές και θετικές τιμές αντίστοιχα. Στα παρακάτω σχήματα παρουσιάζεται η 'μορφή' του μονοδιάστατου πίνακα με τις τιμές της συνάρτησης απόφασης που θα πρέπει να έχει σύμφωνα με τις κλάσεις μας:

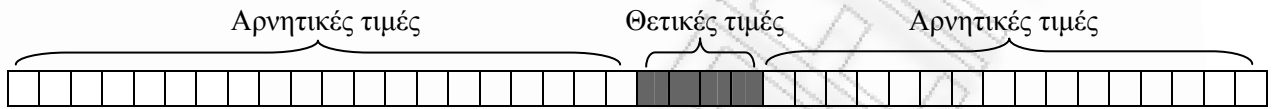




Σχήμα: Τιμές της συνάρτησης απόφασης για την κλάση 4



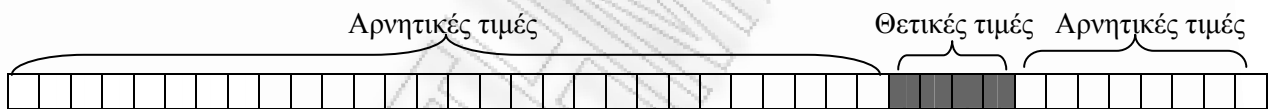
Σχήμα: Τιμές της συνάρτησης απόφασης για την κλάση 5



Σχήμα: Τιμές της συνάρτησης απόφασης για την κλάση 6



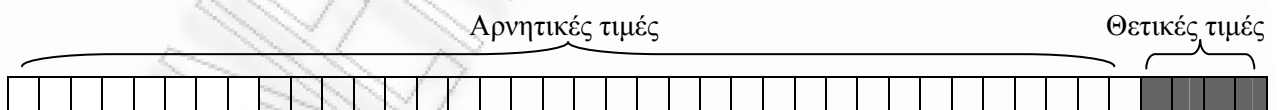
Σχήμα: Τιμές της συνάρτησης απόφασης για την κλάση 7



Σχήμα: Τιμές της συνάρτησης απόφασης για την κλάση 8



Σχήμα: Τιμές της συνάρτησης απόφασης για την κλάση 9



Σχήμα: Τιμές της συνάρτησης απόφασης για την κλάση 10

Αυτά τα δεδομένα όπως παρουσιάζονται στα παραπάνω σχήματα πρέπει στη συνέχεια να μετατραπούν σε στιγμιότυπα κλάσεων και χαρακτηριστικά για να μπορέσουμε σε δεύτερο στάδιο να εκπαιδεύσουμε τον υπέρ-ταξινομητή. Πριν συνεχίσουμε όμως την ανάλυσή μας πρέπει να επισημάνουμε ότι αυτές οι τιμές αναφέρονται σε ένα συγκεκριμένο Sub-Fold. Έχουμε όμως πει ότι τα δεδομένα εκπαίδευσης έχουν σπάσει σε K' Sub-Folds οπότε συνολικά τα παραπάνω δεδομένα πρέπει να τα πολλαπλασιάσουμε K' φορές για να πάρουμε όλο το σύνολο των τιμών της συνάρτησης απόφασης.

Έχουμε φτάσει τώρα στο σημείο να εξηγήσουμε πως οι παραπάνω τιμές της συνάρτησης απόφασης θα μετατραπούν σε χαρακτηριστικά και στιγμιότυπα κλάσεων. Θα εξηγήσουμε πως θα γίνει στο ένα Sub-Fold και στη συνέχεια θα επεκταθούμε στα K' Sub-Folds που διαθέτουμε.

Η ιδέα είναι πρώτα απ' όλα σε κάθε μονοδιάστατο πίνακα να γίνει αναστροφή οπότε από πίνακα γραμμής που έχουμε θα προκύψει πίνακας στήλης. Έτσι όταν μιλάμε για 10 κλάσεις που αποτελούν και τα δεδομένα μας από τα πειράματα με τα μουσικά αρχεία θα έχουμε και 10 στήλες. Αν

είχαμε M κλάσεις τότε θα μιλούσαμε για M στήλες. Αυτές οι M στήλες δίπλα δίπλα θα μας κάνουν ένα πίνακα διαστάσεων:

$$\left(\left[(K-1) \cdot \frac{N}{K} / K' \right] \cdot M \right) \times M$$

Έτσι λοιπόν σε αυτό τον πίνακα οι στήλες θα αποτελέσουν τα χαρακτηριστικά μας και οι γραμμές τα στιγμιότυπά μας. Οπότε με άλλα λόγια το πλήθος των κλάσεων μας γίνεται και πλήθος για τα χαρακτηριστικά του δεύτερου επιπέδου ταξινόμησης.

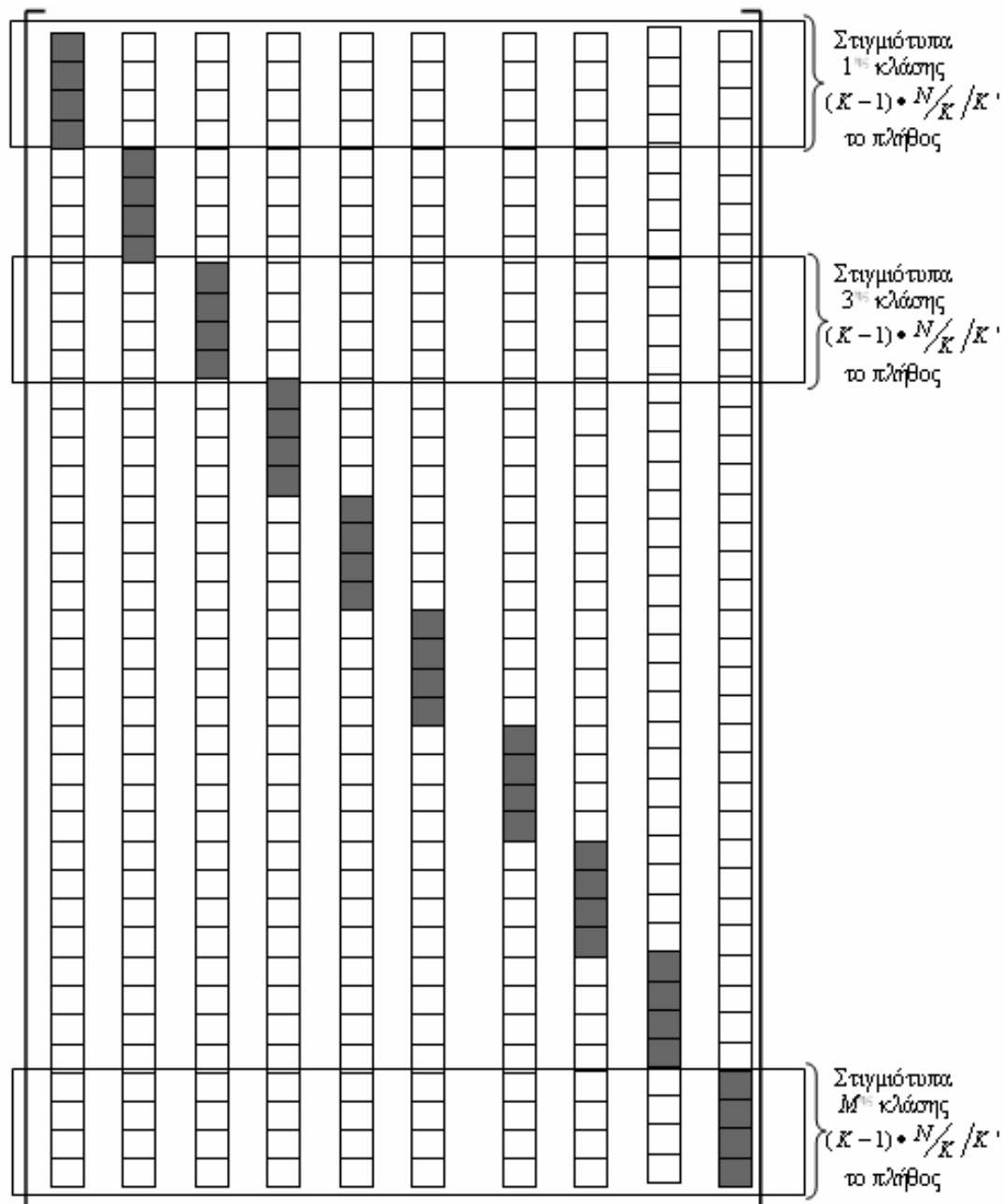
Το επόμενο σημείο που πρέπει να ξεδιαλύνουμε είναι πόσα και πια στιγμιότυπα θα ανήκουν σε κάθε κλάση ή ποιες γραμμές αφού μιλάμε για πίνακα πρέπει να αντιστοιχισθούν σε κάθε κλάση δεδομένων. Αυτό είναι απλό αν σκεφθεί κανείς από πού προέρχονται οι τιμές του πίνακα που οι τιμές αυτές αποτελούν τις τιμές της συνάρτησης απόφασης στο διάστημα του προ-ελέγχου. Στην ιδανική περίπτωση που θα είχαμε τέλειους ταξινομητές τέλεια εκπαιδευμένους με τέλεια αποτελέσματα κατά τον έλεγχο τους, σύμφωνα με τον τρόπο που έχουμε τακτοποιήσει τις τιμές της συνάρτησης απόφασης τότε σε κάθε γραμμή του πίνακα θα είχαμε μία θετική τιμή και όλες τις υπόλοιπες ($M-1$ τιμές) αρνητικές. Ανάλογα τώρα με τη στήλη που αυτή η τιμή θα κατείχε θα αντιστοιχούσε και στην ανάλογη κλάση. Για παράδειγμα η γραμμή που θα είχε τη θετική τιμή στην 5^η στήλη τότε αυτή η γραμμή (στιγμιότυπο) θα ανήκε στην 5^η κλάση δεδομένων. Τώρα όμως επειδή τα πράγματα δεν είναι ιδανικά δεν μπορούμε να στηριχθούμε σε αυτή την υπόθεση. Όμως ξέρουμε πόσες θα έπρεπε να είναι οι θετικές τιμές για την κάθε κλάση από τον αντίστοιχο ταξινομητή και σε ποιο σημείο αυτές θα έπρεπε να βρίσκονται στην ιδανική περίπτωση οπότε έτσι μπορούμε να αντιστοιχήσουμε στιγμιότυπα σε κλάσεις. Έτσι στις πρώτες $(K-1) \cdot \frac{N}{K} / K'$ γραμμές του πίνακα θα βρίσκονται τα στιγμιότυπα για την 1^η κλάση δεδομένων στις επόμενες $(K-1) \cdot \frac{N}{K} / K'$ γραμμές θα βρίσκονται τα στιγμιότυπα της 2^{ης} κλάσης μέχρι να φτάσουμε στην M κλάση. Βέβαια όλα αυτά για τα οποία έχουμε μιλήσει αφορούν μία υπό-ομάδα από τις K' που έχουμε συνολικά στα δεδομένα μας. Οπότε θα δημιουργηθούν άλλοι $K'-1$ αντίστοιχοι πίνακες με αυτόν που προαναφέραμε. Αν φανταστούμε τον πίνακα αυτόν ως μονοδιάστατο τότε ο επόμενος πίνακας θα δημιουργήσει την επόμενη γραμμή του τελικού πίνακα ο μεθεπόμενος της μεθεπόμενης γραμμής μέχρι που τελικά να εξαντλήσουμε όλους τους πίνακες. Τελικά θα δημιουργηθεί ένας τελικός πίνακας διαστάσεων:

$$\left(\left[(K-1) \cdot \frac{N}{K} / K' \right] \cdot M \cdot K' \right) \times M$$

Ο τελικός αριθμός των στιγμιότυπων ανά κλάση που θα δημιουργηθούν θα είναι:

$$\left[(K-1) \cdot \frac{N}{K} / K' \right] \cdot K'$$

Αυτά τα στιγμιότυπα μπορούν τώρα να χρησιμοποιηθούν για την εκπαίδευση σε δεύτερη φάση του υπέρ-ταξινομητή ώστε να είναι σε θέση να μπορεί στη συνέχεια να ταξινομήσει στιγμιότυπα που θα του δοθούν. Στο επόμενο σχήμα φαίνεται πιο καθαρά τα όσα περιγράψαμε για τη δημιουργία των στιγμιότυπων ενός Sub-Fold από τις τιμές της συνάρτησης απόφασης, όπως φαίνονται σε προηγούμενα σχήματα και την κατάταξη αυτών στις αντίστοιχες κλάσεις δεδομένων.



Σχήμα: Αναπαράσταση των τιμών συνάρτησης απόφασης

Πριν ολοκληρώσουμε την περιγραφή του τελευταίου υπέρ-ταξινομητή θα πρέπει να πούμε ότι στην προηγούμενη μορφή όπως την περιγράψαμε θα πρέπει να έρθουν και τα δεδομένα ελέγχου που έχουμε αφήσει στην άκρη κατά το πρώτο 'σπάσιμο' των αρχικών μας δεδομένων. Αυτό πρέπει να γίνει γιατί ο τελικός υπέρ-ταξινομητής έχει εκπαιδευτεί με αυτή τη μορφή δεδομένων οπότε αν του εισάγουμε κάποια άλλη μορφή τότε αυτός θα δώσει λανθασμένα αποτελέσματα. Για να έρθουν σε αυτή τη μορφή τα δεδομένα ελέγχου θα ακολουθήσουμε πάλι την ίδια φιλοσοφία. Θα 'περάσουν' τα στιγμιότυπα από τους αντίστοιχους ταξινομητές ανάλογα με την κλάση την οποία ανήκουν θα προκύψουν οι τιμές της συνάρτησης απόφασης και αυτές στη συνέχεια θα τακτοποιηθούν σε ένα πίνακα όπως ακριβώς περιγράψαμε λίγο νωρίτερα.

Τα αποτελέσματα της λειτουργίας αυτού του υπέρ-ταξινομητή θα παρουσιαστούν στο κεφάλαιο με την πειραματική αξιολόγηση.

6 Δεδομένα

6.1 Μουσικά Δεδομένα

Τα μουσικά δεδομένα προέρχονται από το χώρο της μουσικής που ακούμε στον υπολογιστή μας. Σκοπός είναι να ληφθεί ένα μουσικό κομμάτι σε 'audio' μορφή και να εξαχθεί το συμπέρασμα σε ποιο είδος μουσικής αυτό ανήκει. Οπότε ως είσοδο θα έχουμε ένα audio κομμάτι και ως έξοδο το είδος της μουσικής που αυτό ανήκει.

Ένα audio μουσικό κομμάτι προκειμένου να μπορεί να ταξινομηθεί μέσα από τον ταξινομητή θα πρέπει να επεξεργαστεί και να έρθει σε κατάλληλη μορφή ώστε να μπορέσει ο ταξινομητής να εργαστεί σε αυτό.

Μουσική ποιότητας CD περιέχει 44.100 δείγματα των 16 bit το δευτερόλεπτο. Οπότε για τα δείγματά μας τα οποία είναι της τάξης των 30 δευτερολέπτων, για κάθε στιγμιότυπο θα ισχύει:

$$X = \{0, \dots, 65535\}^{(30 \cdot 44.100)}$$

Όπως γίνεται φανερό αυτό το διάνυσμα είναι τεράστιο οπότε ο πρώτος στόχος μας είναι να μειωθεί η διάστασή του με κάποια ψηφιακή επεξεργασία. Αυτό θα γίνει με τη χρήση του λογισμικού MARSYAS. Με τη βοήθεια αυτού του λογισμικού θα εξαχθούν 30 πραγματικές τιμές για κάθε μουσικό κομμάτι τύπου audio. Οπότε θα έχουμε το παρακάτω διάνυσμα:

$$X = \{\text{Real}\}^{30}$$

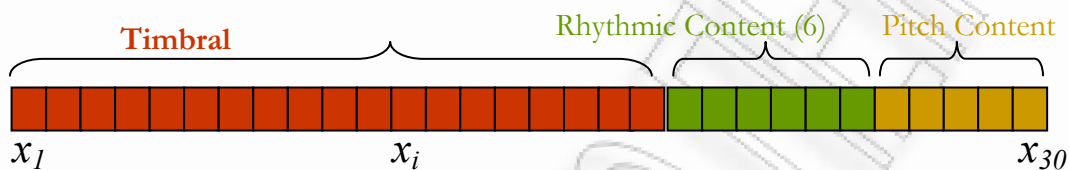
Έτσι λοιπόν κάθε μουσικό κομμάτι θα μετατραπεί σε ένα διάνυσμα 30 χαρακτηριστικών το οποίο αποτελεί μία 'μαθηματική αφαίρεση' της συνολικής πραγματικής πληροφορίας από την οποία αποτελείται κάθε μουσικό αρχείο. Σκοπός είναι να συλλέξουμε την κατάλληλη πληροφορία του ακουστικού σήματος με το παραπάνω λογισμικό που θα βοηθήσει τον ταξινομητή να κατατάξει το συγκεκριμένο μουσικό αρχείο.

Κάθε μουσικό κομμάτι που θα χρησιμοποιήσουμε είναι της τάξης των 30 δευτερολέπτων όπως έχουμε αναφέρει. Πιο συγκεκριμένα το κάθε αρχείο των 30 δευτερολέπτων έσπασε σε μικρότερα τμήματα της τάξης των 50msec που αναλύθηκαν ξεχωριστά για κάθε μουσικό κομμάτι. Αυτά τα τμήματα ονομάζονται 'παραθύρο ανάλυσης' ή πιο απλά 'πλαίσιο' (frame) και πρέπει να είναι αρκετά μικρό ώστε το φάσμα της έντασης του ήχου να είναι σχετικά σταθερό. Από την άλλη μεριά ο όρος 'περιεχόμενο παραθύρου' περιγράφει το μικρότερο παράθυρο (μικρότερη ποσότητα ήχου) που είναι απαραίτητο για την αναγνώριση ενός μουσικού περιεχομένου. Τα πραγματικά αντικειμενικά διανύσματα που χρησιμοποιούνται είναι το τρέχων μέσος ο διάμεσος και η σταθερή απόκλιση των χαρακτηριστικών των μουσικών σημάτων τα οποία υπολογίζονται μέσα από τα 'παραθύρα ανάλυσης'. Στον παρακάτω πίνακα φαίνονται αναλυτικά οι συνιστώσες για κάθε διάνυσμα χαρακτηριστικών.

Χαρακτηριστικό	Όνομα χαρακτηριστικού	Χαρακτηριστικό	Όνομα χαρακτηριστικού
1	Mean centroid	20	Beat A0
2	Mean rolloff	21	Beat A1
3	Mean flux	22	Beat RA
4	Mean zero-crossings	23	Beat P1
5	STD of centroid	24	Beat P2
6	STD of rolloff	25	Beat Sum
7	STD of flux	26	Pitch FA0
8	STD of zero-crossings	27	Pitch UP0
9	Low energy	28	Pitch FP0
10...19	MFCCs	29	Pitch IP0
		30	Pitch Sum

Πίνακας : συνιστώσες για κάθε διάνυσμα χαρακτηριστικών (όπου STD = Standard Deviation (κανονική απόκλιση))

Έτσι έχουμε λοιπόν το κάθε στιγμιότυπο μουσικού κομματιού να αποτελείται από 30 χαρακτηριστικά τα οποία να περιέχουν πληροφορίες για το αντίστοιχο στιγμιότυπο μουσικού κομματιού. Τα πρώτα 19 χαρακτηριστικά από τα 30 περιέχουν πληροφορίες για τη σύσταση της ηχητικής χροιάς του μουσικού κομματιού. Συγκεκριμένα περιέχει πληροφορίες για τη διάκριση μεταξύ φωνής και μουσικής καθώς και αναγνώριση φωνής. Αυτό γίνεται με τη βοήθεια του αλγορίθμου ‘Short Time Fourier Transform (STFT)’. Τα επόμενα 6 χαρακτηριστικά περιέχουν πληροφορία για τα περιεχόμενα του ρυθμού του μουσικού κομματιού. Πιο συγκεκριμένα καταγράφονται πληροφορίες για την ισχύ του ρυθμού του μουσικού κομματιού το πλάτος και την ανάλυση του ρυθμού. Αυτή η λήψη των διανυσμάτων γίνεται με τη βοήθεια του ‘Wavelet Transform’. Τα τελευταία 5 διανύσματα περιέχουν πληροφορίες για το βάθος του τόνου της φωνής. Καταγράφονται πληροφορίες σχετικά με το ποιος τόνος επικρατεί περισσότερο μέσα στο μουσικό κομμάτι καθώς και τα κενά μεταξύ των τόνων φωνής. Η συγκεκριμένη καταγραφή των διανυσμάτων γίνεται με τη βοήθεια του αλγορίθμου ‘Multipitch Detection Algorithm’. Όλα όσα περιγράψαμε φαίνονται διαγραμματικά στο παρακάτω σχήμα:



Σχήμα: Τα περιεχόμενα των διανυσμάτων για κάθε στιγμιότυπο μουσικού κομματιού

Στη συνέχεια θα περιγράψουμε πιο αναλυτικά τη δημιουργία των χαρακτηριστικών που αποτελούν το κάθε διάνυσμα για το κάθε στιγμιότυπο μουσικού κομματιού. Θα ξεκινήσουμε με τα πρώτα 19 χαρακτηριστικά που περιέχουν πληροφορίες για την ηχητική χροιά.

6.1.1 Χαρακτηριστικά ηχητικής χροιάς

Προκειμένου να περιγραφεί η ηχητική χροιά ενός μουσικού κομματιού η οποία θα βοηθήσει στην ταξινόμηση συλλέγονται στατιστικά δεδομένα του φάσματος του ηχητικού σήματος. Αυτά τα στατιστικά δεδομένα παρουσιάζονται παρακάτω:

-Spectral centroid: Αυτό το χαρακτηριστικό αποτυπώνει τη ‘λάμψη’ ηχητικού σήματος και υπολογίζεται ως το ισοζυγισμένο σημείο του φάσματος. Μπορεί να υπολογιστεί από τον παρακάτω τύπο:

$$C = \frac{\sum_{n=0}^{N-1} M_t[n] \cdot n}{\sum_{n=0}^{N-1} M_t[n]} \quad (1)$$

Όπου $M_t[n]$ είναι η ένταση του ήχου από το μετασχηματισμό Φουριέ (Fourier) ενός ‘παράθυρου’ (frame) t και συχνότητα n .

-Spectral rolloff: Αυτό το χαρακτηριστικό περιγράφει τη μορφή του φάσματος και ορίζεται ως η συχνότητα $R=R(r)$ που αντιστοιχεί στο ποσοστό $r\%$ της κατανομής της έντασης του ήχου. Μπορεί να θεωρηθεί ως γενίκευση του ‘spectral centroid’ καθώς αυτό αντιστοιχεί στο ποσοστό $r=50\%$ της κατανομής της έντασης του ήχου. Στα συγκεκριμένα δεδομένα έχουμε δώσει τιμή στο r ίση με 95% η οποία έχει καθοριστεί πειραματικά. Το ‘ N ’ είναι το μέγεθος του κάθε σήματος που αποθηκεύεται στο διάνυσμα x .

$$\sum_{n=0}^R M_t[n] = r \cdot \sum_{n=0}^{N-1} M_t[n] \quad (2)$$

-Spectral flux: Αυτό το χαρακτηριστικό περιγράφει την αλλαγή της συχνότητας με βάση το χρόνο και υπολογίζεται ως η διαφορά της έντασης του ήχου μικρού χρονικού διαστήματος στο μετασχηματισμό Φουριέ μεταξύ του τρέχοντος και του προηγούμενου ‘παράθυρου’ (frame). Γι’ αυτό το λόγο το ‘spectral flux’ είναι ένα μέτρο της τοπικής αλλαγής του φάσματος που δίνεται από την εξίσωση:

$$SF = \sum_{n=0}^{N-1} (N_t[n] - N_{t-1}[n])^2 \quad (3)$$

όπου $N_t[n]$ και $N_{t-1}[n]$ είναι η κονονικοποιημένη ένταση του ήχου του μικρού χρονικού διαστήματος μετασχηματισμού Φουριέ στο παράθυρο t και $t-1$ αντίστοιχα.

-Zero-crossings: Ένα zero-crossing συμβαίνει όταν διαδοχικά δείγματα σε ένα ψηφιακό σήμα έχουν διαφορετικές ενδείξεις σήματος. Αυτό το χαρακτηριστικό είναι χρήσιμο στην ανεύρεση του ποσού του θορύβου σε ένα σήμα και μπορεί να υπολογιστεί ως:

$$Z_n = \sum_m |\text{sgn}[x(m)] - \text{sgn}[x(m-1)]| \bullet w(n-m) \quad (4)$$

όπου:

$$\text{sgn}[x(n)] = \begin{cases} 1, & x(n) \geq 0 \\ 0, & x(n) < 0 \end{cases} \quad (5)$$

και

$$w(m) = \begin{cases} \frac{1}{2}, & 0 \leq m \leq N-1 \\ 0, & \text{διαφορετικά} \end{cases} \quad (6)$$

-Short-time energy function: Η short-time energy ενός ηχητικού σήματος $x(m)$ ορίζεται ως εξής:

$$E_n = \frac{1}{N} \sum_m [x(m) \bullet w(n-m)]^2 \quad (7)$$

όπου

$$w(m) = \begin{cases} 1, & 0 \leq m \leq N-1 \\ 0, & \text{διαφορετικά} \end{cases} \quad (8)$$

Στις εξισώσεις (7) και (8) $x(m)$ είναι το διακριτό ηχητικό σήμα, n είναι ο δείκτης του 'short time energy' και $w(m)$ είναι ένα ορθογώνιο παράθυρο. Αυτό το χαρακτηριστικό δίνει μία απλή απεικόνιση της αλλαγής στη διάρκεια του χρόνου του πλάτους του ηχητικού σήματος.

-Mel-frequency cepstral coefficients (MFCCs): Αυτοί οι συντελεστές είναι σχεδιασμένοι να συλλαμβάνουν μικρού χρονικού διαστήματος χαρακτηριστικά του φάσματος. Αφού πάρουμε το λογάριθμο του πλάτους του φάσματος το οποίο έχει ληφθεί από το μετασχηματισμό Φουριέ του κάθε παραθύρου (frame) οι συχνότητες ομαδοποιούνται και ομαλοποιούνται σύμφωνα με την 'Mel-frequency' κλίμακα η οποία έχει σχεδιαστεί σύμφωνα με την ανθρώπινη ακουστική αντίληψη. Οι συντελεστές MFCCs παράγονται εφαρμόζοντας στα 'Mel-spectral' διανύσματα έναν διακριτό μετασχηματισμό συνημίτονου.

6.1.2 Χαρακτηριστικά ρυθμού

Τα χαρακτηριστικά ρυθμού χαρακτηρίζουν την 'κίνηση' των μουσικών σημάτων κατά τη διάρκεια του χρόνου και περιέχουν πληροφορία όπως η ομαλότητα του ρυθμού. Το σύνολο των χαρακτηριστικών που απεικονίζουν το ρυθμό εξάγονται από ένα παλμικό ιστόγραμμα το οποίο είναι μία καμπύλη που περιγράφει τη δύναμη του κάθε παλμού (κτύπου) ως μία συνάρτηση από τιμές ρυθμού. Αυτές οι τιμές μπορεί να χρησιμοποιηθούν για να μας δώσουν πληροφορία για την πολυπλοκότητα του παλμού στο μουσικό αρχείο. Αυτό το σύνολο των χαρακτηριστικών βασίζεται στον εντοπισμό των πιο ήρεμων περιόδων του σήματος και συνήθως εξάγονται από το ιστόγραμμα που αναφέραμε λίγο νωρίτερα.

6.1.3 Χαρακτηριστικά τόνου φωνής

Τα χαρακτηριστικά του τόνου περιγράφουν την μελωδία και την αρμονία σε ένα μουσικό σήμα. Ένας αλγόριθμος ανίχνευσης τόνου αποσυνθέτει το σήμα σε δύο μπάντες συχνοτήτων και τα πλάτη που περικλείονται εξάγονται για κάθε μπάνα συχνοτήτων όπου η εξαγωγή εκτελείται μέσω half-way ανόρθωση και low-pass φιλτραρίσματος. Τα περιεχόμενα αθροίζονται και μια εμπλουτισμένη

συνάρτηση υπολογίζεται έτσι ώστε να ελαττωθεί η επίδραση των πολλαπλών ακέραιων κορυφών των συχνοτήτων να πολλαπλασιάσουν την ανίχνευση κορυφών. Οι κύριες κορυφές της συνάρτησης συγκεντρώνονται σε ιστογράμματα τόνου και τα χαρακτηριστικά του τόνου φωνής εξάγονται από αυτά τα ιστογράμματα. Τα χαρακτηριστικά του τόνου φωνής τυπικά περιλαμβάνουν πλάτη κ περιόδους από μέγιστα κορυφών, κενά τόνου φωνής μεταξύ των πιο σημαντικών κορυφών και γενικά αθροίσματα από τα ιστογράμματα.

Αφού περιγράψαμε τη διαδικασία δημιουργίας του κάθε στιγμιότυπου από 30 διαφορετικά διανύσματα για το κάθε μουσικό κομμάτι ακολούθησε η λήψη αυτών των δεδομένων από πραγματικά audio μουσικά αρχεία.

Επιλέχθηκαν 10 διαφορετικά είδη western μουσικής οπότε δημιουργήσαμε συνολικά 10 διαφορετικές κλάσεις όπως φαίνεται και στον παρακάτω πίνακα:

Class ID	Label
1	Blues
2	Classical
3	Country
4	Disco
5	Hip-Hop
6	Jazz
7	Metal
8	Pop
9	Reggae
10	Rock

Πίνακας: Οι 10 διαφορετικές κλάσεις μουσικών κομματιών

Για κάθε κλάση συλλέχθηκαν 100 δείγματα – στιγμιότυπα μουσικών κομματιών των 30 δευτερολέπτων το κάθε ένα. Οπότε συνολικά συλλέχθηκαν 1000 δείγματα με 30 διανύσματα για το κάθε στιγμιότυπο και δημιουργήθηκε ένας πίνακας 1000X30. Αυτός ο πίνακας περιείχε όλα τα δεδομένα πάνω στα οποία έγιναν όλα τα πειράματα και οι έλεγχοί μας. Από αυτά τα δεδομένα κάποια χρησιμοποιήσαμε για εκπαίδευση και κάποια άλλα για δοκιμή των ταξινομητών που δημιουργήσαμε. Κατά κύριο λόγο χρησιμοποιήθηκε το 90% των δεδομένων για εκπαίδευση και το υπόλοιπο 10% για έλεγχο. Πριν όμως χρησιμοποιήσουμε τα δεδομένα τα κανονικοποιήσαμε μεταξύ του διαστήματος [0,1] για να βγάλουμε καλύτερα αποτελέσματα.

Για τη συλλογή των μουσικών δεδομένων όπου το κάθε στιγμιότυπο αποτελείται από 280 χαρακτηριστικά ακολουθήθηκε η ίδια λογική με αυτή που περιγράφηκε λίγο πιο πάνω για τα στιγμιότυπα των 30 χαρακτηριστικών μόνο που τώρα λήφθηκαν ακόμα περισσότερα στοιχεία από κάθε μουσικό κομμάτι για να μπορέσουμε να πάρουμε μεγαλύτερη πληροφορία για το κάθε ένα.

6.2 Δεδομένα εκφράσεων προσώπου

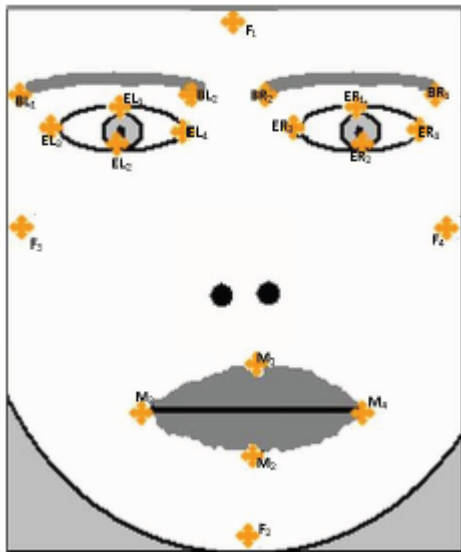
Τα δεδομένα εκφράσεων προσώπου προέρχονται από φωτογραφίες 250 ατόμων οι οποίες απεικονίζουν μόνο τα πρόσωπά τους. Για κάθε άτομο λαμβάνονται 6 φωτογραφίες οι οποίες αποτελούν 6 διαφορετικές εκφράσεις τους. Αυτές οι εκφράσεις είναι:

- Ουδέτερη
- Ευτυχία
- Λύπη
- Έκπληξη
- Θυμός
- Αηδία

Από αυτές τις 6 εκφράσεις που αποτελούν στην πραγματικότητα 6 διαφορετικές κλάσεις θα χρησιμοποιηθούν στον υπέρ-ταξινομητή μας μόνο οι 5. Η πρώτη έκφραση (ουδέτερη) θα χρησιμοποιηθεί μόνο στην βοήθεια να εξαχθούν τα χαρακτηριστικά από τα οποία θα αποτελείται κάθε στιγμιότυπο της εκάστοτε κλάσης. Όπως γίνεται αντιληπτό δουλειά του υπέρ-ταξινομητή με αυτά τα

δεδομένα θα είναι δοθέντος ενός στιγμιότυπου να βγάλει συμπέρασμα σε ποιο είδος έκφρασης προσώπου αυτό ανήκει.

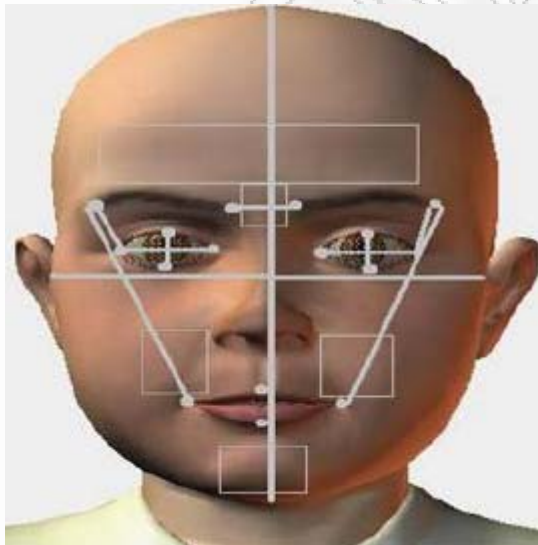
Τα πιο σημαντικά σημεία του προσώπου τα οποία θα βοηθήσουν στην εξαγωγή των χαρακτηριστικών του στιγμιότυπου των δεδομένων εκφράσεων προσώπου φαίνονται στο παρακάτω σχήμα:



Σχήμα: Τα πιο σημαντικά σημεία του προσώπου

Αυτά τα σημεία χρησιμοποιούνται ευρέως σε συστήματα επεξεργασίας προσώπου. Ο σκοπός της εξαγωγής αυτών των χαρακτηριστικών είναι να μετατραπούν τα απλά pixels της κάθε φωτογραφίας σε υψηλότερο επίπεδο πληροφορία όπως είναι το σχήμα, η κίνηση, το χρώμα, η υφή και η 'χωρική διαμόρφωση' του προσώπου και των συστατικών του (πως δηλαδή τα μάτια, το στόμα κ.α. αλλάζουν σχήμα μέσα στο πρόσωπο). Ιδιαίτερα δίνεται μεγάλη σημασία σε σημεία του προσώπου με γωνίες όπως είναι τα μάτια το στόμα τα φρύδια και υπολογίζονται αποκλίσεις στο μέγεθος ή στον προσανατολισμός που δημιουργείται σε σχέση με την αντίστοιχη ουδέτερη έκφραση. Επίσης εξάγονται συγκεκριμένες περιοχές του προσώπου όπως είναι το μέτωπο ή η περιοχή μεταξύ των φρυδιών με σκοπό να υπολογιστούν και πάλι αποκλίσεις στις τιμές σε σχέση με την ουδέτερη πάντα έκφραση.

Στο παρακάτω σχήμα που ακολουθεί φαίνονται οι περιοχές του προσώπου για τις οποίες μιλήσαμε λίγο πριν και τα σημεία από όπου λαμβάνονται οι τιμές για να δημιουργηθούν τα χαρακτηριστικά του κάθε στιγμιότυπου.



Σχήμα: Σημεία υπολογισμού χαρακτηριστικών

Στη συνέχεια θα παρουσιάσουμε τα 7 χαρακτηριστικά που χρησιμοποιούμε για να περιγράψουμε το κάθε στιγμιότυπο που αποτελεί μία έκφραση προσώπου. Τα χαρακτηριστικά αυτά είναι:

- Αναλογία στόματος

$$\frac{\|M_1 M_2\| \|M_3 M_4\|}{\|M_{1Neu} M_{2Neu}\| \|M_{3Neu} M_{4Neu}\|}$$

όπου M είναι το αντίστοιχο σημείο του στόματος και M_{Neu} το αντίστοιχο σημείο του στόματος για την ουδέτερη έκφραση προσώπου.

- Αναλογία αριστερού ματιού

$$\frac{\|EL_1 EL_2\| \|EL_3 EL_4\|}{\|EL_{1Neu} EL_{2Neu}\| \|EL_{3Neu} EL_{4Neu}\|}$$

όπου EL είναι το αντίστοιχο σημείο του αριστερού ματιού και EL_{Neu} το αντίστοιχο σημείο του αριστερού ματιού για την ουδέτερη έκφραση προσώπου.

- Αναλογία δεξιού ματιού

$$\frac{\|ER_1 ER_2\| \|ER_3 ER_4\|}{\|ER_{1Neu} ER_{2Neu}\| \|ER_{3Neu} ER_{4Neu}\|}$$

όπου ER είναι το αντίστοιχο σημείο του αριστερού ματιού και ER_{Neu} το αντίστοιχο σημείο του αριστερού ματιού για την ουδέτερη έκφραση προσώπου.

- Αναλογία κεφαλιού

$$\frac{\|F_1 F_2\| \|F_3 F_4\|}{\|F_{1Neu} F_{2Neu}\| \|F_{3Neu} F_{4Neu}\|}$$

όπου F είναι το αντίστοιχο σημείο του αριστερού ματιού και F_{Neu} το αντίστοιχο σημείο του αριστερού ματιού για την ουδέτερη έκφραση προσώπου.

- Υφή μετώπου: μέτρηση της αλλαγής της υψής του μετώπου σε σχέση με την υφή του μετώπου στην ουδέτερη έκφραση.
- Υφή πιγουνιού: μέτρηση της αλλαγής της υψής στο πιγούνι σε σχέση με την υφή στο πιγούνι στην ουδέτερη έκφραση.
- Υφή της περιοχής μεταξύ των φρυδιών: μέτρηση της αλλαγής της υψής μεταξύ των φρυδιών σε σχέση με την υφή των φρυδιών στην ουδέτερη έκφραση.

Τελικά για την κάθε κλάση δεδομένων στα δεδομένα εκφράσεων προσώπου δημιουργούνται 250 στιγμιότυπα των 7 χαρακτηριστικών το κάθε στιγμιότυπο. Συνολικά αφού έχουμε 5 κλάσεις δημιουργείται ένας πίνακας διαστάσεων 1250X7. Πάνω σε αυτά τα δεδομένα έγιναν τα πειράματά μας όσον αφορά τα δεδομένα εκφράσεων προσώπου για να δούμε πως συμπεριφέρεται ο υπέρ-ταξινομητής μας μετά το σχεδιασμό του. Τα αποτελέσματα αυτά θα παρουσιάσουμε στο επόμενο κεφάλαιο που ακολουθεί.

7 Πειραματική αξιολόγηση

Έχοντας αναπτύξει στα κεφάλαια που προηγήθηκαν το θεωρητικό πλαίσιο στο οποίο τοποθετείται η εργασία, και έχοντας περιγράψει τη σχεδίαση του υπέρ-ταξινομητή που υλοποιήσαμε, σε αυτό το κεφάλαιο θα παρουσιάσουμε τα αποτελέσματα μιας σειράς πειραμάτων που αξιολογούν τον υπέρ-ταξινομητή μας με την εφαρμογή ενός αλγορίθμου μάθησης με επίβλεψη από θετικά και αρνητικά ταξινομημένα παραδείγματα. Στον υπέρ-ταξινομητή μας εφαρμόζονται τόσο ισοζυγισμένα όσο και σε μη ισοζυγισμένα σώματα δεδομένων που περιγράφονται από αριθμητικά χαρακτηριστικά.

Για την υλοποίηση των πειραμάτων στηριχθήκαμε στην προγραμματιστική υποδομή OSU SVM Classifier Matlab Toolbox (έκδοση 3.00) που υλοποιήθηκε από τους J. Ma, Y. Zhao και S. Ahalt στο πανεπιστήμιο του Οχάιο. Πρόκειται για μια ολοκληρωμένη συλλογή αλγορίθμων μηχανικής μάθησης που περιλαμβάνει τη μέθοδο ταξινόμησης SVM η οποία έχει υλοποιηθεί με το προγραμματιστικό εργαλείο Matlab.

Για την υλοποίηση των πειραμάτων που αφορούν την ταξινόμηση με βάση τεχνητά νευρωνικά δίκτυα χρησιμοποιήσαμε το σύνολο εργαλείων PRTools4 (έκδοση 4.1) η οποία έχει υλοποιηθεί επίσης με το προγραμματιστικό εργαλείο Matlab από τους R.P.W. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. de Ridder, D.M.J. Tax και S. Verzakon στο πανεπιστήμιο του Ντελφτ στην Ολλανδία.

7.1 Μέθοδος 'One against all'

Στη συνέχεια θα παρουσιάσουμε τα αποτελέσματα των πειραμάτων σε αντιστοιχία με τους υπέρ-ταξινομητές που δημιουργήσαμε. Θα ξεκινήσουμε από την πρώτη μέθοδο την 'one against all' και τα μουσικά δεδομένα των 30 χαρακτηριστικών. Στη συγκεκριμένη μέθοδο θα παρουσιάσουμε μόνο τον SVM ταξινομητή και θα εξηγήσουμε το λόγο γιατί δεν γίνεται αναφορά σε όλους τους ταξινομητές και τα δεδομένα που έχουμε.

Στα μουσικά δεδομένα λοιπόν έχουμε 10 διαφορετικά είδη μουσικής συνεπώς και 10 διαφορετικές κλάσεις. Σύμφωνα με όσα έχουμε περιγράψει κατά το σχεδιασμό του υπέρ-ταξινομητή θα δημιουργηθούν $10*10=100$ διαφορετικοί ταξινομητές. 10 ταξινομητές για κάθε κλάση δεδομένων. Επίσης έχουμε πει ότι η κάθε κλάση για τα συγκεκριμένα δεδομένα αποτελείται από 100 στιγμιότυπα οπότε θα έχουμε 90 στιγμιότυπα για εκπαίδευση από κάθε κλάση και 10 στιγμιότυπα για έλεγχο αντίστοιχα. Συνολικά θα έχουμε από όλες τις κλάσεις 900 στιγμιότυπα για εκπαίδευση και 100 στιγμιότυπα για έλεγχο των ταξινομητών.

Σύμφωνα με αυτά που είπαμε για κάθε ταξινομητή θα έχουμε 90 θετικά δεδομένα εκπαίδευσης και άλλα 810 αρνητικά δεδομένα εκπαίδευσης. Τα θετικά δεδομένα εκπαίδευσης θα προέρχονται από την κλάση που αναφέρεται ο ταξινομητής (ο ταξινομητής αυτός θα ταξινομεί ένα στιγμιότυπο σε σχέση με το αν αυτό ανήκει ή δεν ανήκει στην κλάση για την οποία έχει δημιουργηθεί ο ταξινομητής) και τα υπόλοιπα 810 θα προέρχονται από όλους τους άλλους ταξινομητές. Αντίστοιχα για τα δεδομένα ελέγχου τα 10 θετικά δεδομένα ελέγχου θα προέρχονται από την αντίστοιχη κλάση για την οποία έχει δημιουργηθεί ο ταξινομητής και τα υπόλοιπα 90 αρνητικά δεδομένα ελέγχου από όλες τις υπόλοιπες κλάσεις.

Μετά την εκπαίδευση των ταξινομητών που υλοποιήσαμε, τους ελέγχουμε σύμφωνα με τα δεδομένα ελέγχου που δημιουργήσαμε από το σύνολο των δεδομένων όπως περιγράψαμε παραπάνω. Τα αποτελέσματα για τον κάθε ταξινομητή και για το αντίστοιχη Fold δεδομένων φαίνεται στους δύο παρακάτω πίνακες:

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Ταξινομητής1	0,9	0,92	0,9	0,89	0,91
Ταξινομητής2	0,97	0,96	0,97	0,94	0,95
Ταξινομητής3	0,88	0,9	0,9	0,91	0,9
Ταξινομητής4	0,91	0,9	0,91	0,9	0,89
Ταξινομητής5	0,89	0,91	0,9	0,91	0,9
Ταξινομητής6	0,92	0,91	0,94	0,94	0,93
Ταξινομητής7	0,91	0,96	0,92	0,94	0,97
Ταξινομητής8	0,9	0,92	0,91	0,88	0,93
Ταξινομητής9	0,91	0,92	0,93	0,89	0,92
Ταξινομητής10	0,89	0,9	0,9	0,9	0,9

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου

	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
Ταξινομητής1	0,9	0,9	0,9	0,9	0,9
Ταξινομητής2	0,95	0,97	0,92	0,97	0,93
Ταξινομητής3	0,9	0,91	0,9	0,91	0,89
Ταξινομητής4	0,89	0,89	0,9	0,88	0,91
Ταξινομητής5	0,88	0,9	0,91	0,9	0,89
Ταξινομητής6	0,94	0,92	0,93	0,93	0,91
Ταξινομητής7	0,93	0,96	0,89	0,93	0,92
Ταξινομητής8	0,93	0,94	0,97	0,9	0,92
Ταξινομητής9	0,92	0,9	0,89	0,9	0,87
Ταξινομητής10	0,89	0,9	0,89	0,9	0,89

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου

Με μία μικρή παρατήρηση φαίνεται ότι ο κάθε ταξινομητής ξεχωριστά έχει εκπαιδευτεί σωστά και ταξινομεί τα άγνωστα για αυτόν στιγμιότυπα που του δίνονται με πολύ μεγάλο ποσοστό επιτυχίας.

Στη συνέχεια παρουσιάζεται ένα παράδειγμα ταξινόμησης ενός στιγμιότυπου το οποίο ξέρουμε ότι ανήκει στην πρώτη κλάση. Σύμφωνα με όσα έχουμε περιγράψει σε προηγούμενο κεφάλαιο ο ταξινομητής 1 ο οποίος έχει εκπαιδευτεί να ταξινομεί στιγμιότυπα της κλάσης 1 σε σχέση με τις υπόλοιπες κλάσεις θα δώσει μεγαλύτερη τιμή στη συνάρτηση απόφασης (Decision Value) απ' ότι οι υπόλοιποι ταξινομητές. Οπότε έχουμε βρει και τη ζητούμενη κλάση.

Ενδεικτικά στους επόμενους δύο πίνακες παρουσιάζουμε όλες τις τιμές της συνάρτησης απόφασης των ταξινομητών που ανήκουν στις αντίστοιχες κλάσεις και Fold δεδομένων που προκύπτουν μετά τον έλεγχο του πρώτου στιγμιότυπου με όλους τους επιμέρους ταξινομητές:

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Ταξινομητής1	-0,57588	0,047434	0,044981	-0,18125	0,078152
Ταξινομητής2	1,0735	1,0371	1,1041	1,3607	1,1447
Ταξινομητής3	1,647	1,6301	1,3915	1,6059	1,5126
Ταξινομητής4	1,1033	1,1572	1,1205	1,0825	1,0812
Ταξινομητής5	1,2494	1,0979	1,0894	1,2431	1,2635
Ταξινομητής6	1,4712	1,6779	1,6351	1,5814	1,4367
Ταξινομητής7	1,0906	1,0805	1,1921	1,2461	1,117
Ταξινομητής8	1,6321	1,6632	1,5295	1,3909	1,4508
Ταξινομητής9	0,50463	1,0002	1,0004	1,0005	0,99966
Ταξινομητής10	0,76324	0,99998	1,0001	1,0002	1,0002

Πίνακας: Τιμές της συνάρτησης απόφασης για το πρώτο στιγμιότυπο

	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
Ταξινομητής1	0,11073	0,15118	0,20706	0,20528	0,070881
Ταξινομητής2	1,056	1,0571	1,0007	1,1139	1,0253
Ταξινομητής3	1,5093	1,2542	1,6426	1,4575	1,5609
Ταξινομητής4	1,1382	1,4407	1,2488	1,0338	1,1623
Ταξινομητής5	1,2486	1,3852	1,1725	1,1504	1,2707
Ταξινομητής6	1,6468	1,6236	1,5735	1,7884	1,4898
Ταξινομητής7	1,0002	1,1469	1,1162	1,0195	1,3389
Ταξινομητής8	1,5623	1,7727	1,5384	1,5963	1,5444
Ταξινομητής9	0,95125	0,99231	0,99994	1,0002	1
Ταξινομητής10	0,99959	1,0002	1,0001	1,0003	0,99955

Πίνακας: Τιμές της συνάρτησης απόφασης για το πρώτο στιγμιότυπο

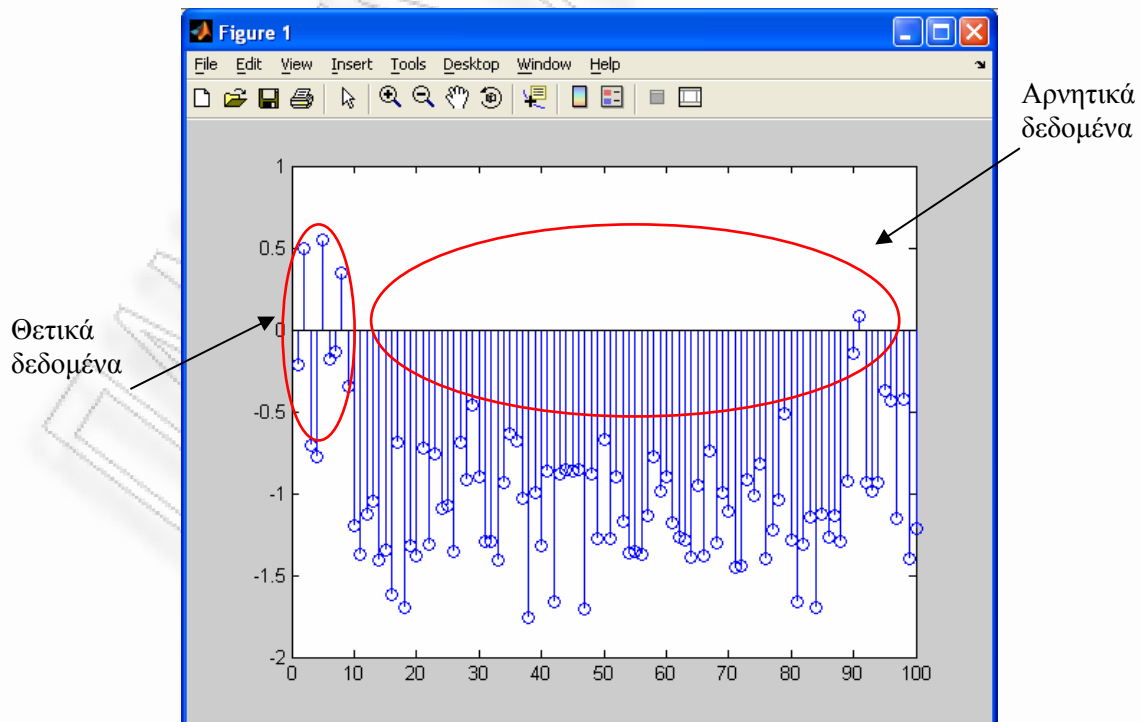
Μετά τον έλεγχο όλων των στιγμιότυπων παρατηρούμε με έκπληξη ότι ο υπέρ-ταξινομητής μας ταξινομεί τα δείγματά μας στις αντίστοιχες κλάσεις τους με ποσοστό επιτυχίας 10%. Ένα ελάχιστο μικρό ποσοστό σε σχέση με το ποσοστό επιτυχίας που έχει ο κάθε ταξινομητής ξεχωριστά. Παρατηρώντας όμως πιο προσεκτικά τα αποτελέσματα που παίρνουμε από τους ταξινομητές μετά την εκπαίδευσή τους και κυρίως τον πίνακα με τις τιμές της συνάρτησης απόφασης βλέπουμε ότι δεν ταξινομούν καθόλου σωστά ή σχεδόν καθόλου σωστά τα θετικά δεδομένα ενώ τα αρνητικά δεδομένα ταξινομούνται στην πλειοψηφία τους άριστα.

Στον επόμενο πίνακα παρουσιάζουμε ενδεικτικά τις τιμές που παίρνει η συνάρτηση απόφασης από τον έλεγχο που γίνεται στον ταξινομητή 1 του δεύτερου Fold δεδομένων με τα αντίστοιχα δεδομένα ελέγχου που έχουν δημιουργηθεί για αυτόν τον ταξινομητή.

	Fold 2				
Κλάση1	-0,21637	0,49862	-0,70506	-0,77981	0,54892
	-0,18372	-0,13225	0,34603	-0,34664	-1,2013
Κλάση2	-1,376	-1,1285	-1,0442	-1,4102	-1,3427
	-1,6147	-0,68633	-1,6938	-1,3234	-1,3835
Κλάση3	-0,72688	-1,3071	-0,75534	-1,0931	-1,0769
	-1,3572	-0,6846	-0,91493	-0,45954	-0,89604
Κλάση4	-1,2958	-1,298	-1,4055	-0,93797	-0,63513
	-0,68007	-1,0344	-1,7628	-0,99298	-1,3172
Κλάση5	-0,867	-1,6588	-0,88555	-0,85285	-0,86472
	-0,85894	-1,7022	-0,88253	-1,2785	-0,6732
Κλάση6	-1,2763	-0,90223	-1,1744	-1,3662	-1,3535
	-1,3705	-1,1382	-0,77605	-0,98916	-0,89677
Κλάση7	-1,1766	-1,2656	-1,2842	-1,387	-0,95405
	-1,3808	-0,74194	-1,3059	-0,99413	-1,1126
Κλάση8	-1,4487	-1,4428	-0,91594	-1,0161	-0,81934
	-1,4019	-1,2278	-1,0399	-0,5124	-1,289
Κλάση9	-1,6632	-1,314	-1,1491	-1,7007	-1,126
	-1,2652	-1,1327	-1,2973	-0,92361	-0,14897
Κλάση10	0,079399	-0,93658	-0,98749	-0,93168	-0,37648
	-0,43734	-1,1555	-0,42888	-1,4031	-1,2163

Πίνακας: Τιμές της συνάρτησης απόφασης για τον ταξινομητή 1 του Fold 2 δεδομένων μετά τον έλεγχο της από τα δεδομένα ελέγχου που της αντιστοιχούν

Τα παραπάνω αποτελέσματα φαίνονται ακόμα πιο καλά στο παρακάτω ραβδωτό διάγραμμα όπου φαίνονται συνολικά οι τιμές της συνάρτησης απόφασης της κλάσης 1 του Fold 2 σε σχέση με όλες τις υπόλοιπες κλάσεις. Είναι ευδιάκριτα τα σωστά και τα λανθασμένα αποτελέσματα που βγάζει ο ταξινομητής για τα αντίστοιχα δεδομένα που του δίνονται.



Σχήμα: Διάγραμμα με τις τιμές της συνάρτησης απόφασης του ταξινομητή 1 με δεδομένα από τον Fold 2

Σύμφωνα με τον τρόπο που έχει εκπαιδευτεί ο ταξινομητής 1 του Fold 2 τα θετικά δεδομένα για αυτόν ανήκουν στην κλάση 1 και τα αρνητικά δεδομένα στις υπόλοιπες 9 κλάσεις. Όπως φαίνεται ο ταξινομητής αδυνατεί να ταξινομήσει σωστά τα θετικά δεδομένα αυτά δηλαδή της κλάσης 1 ενώ τα αντίστοιχα αρνητικά ταξινομούνται σωστά. Αυτό γίνεται εμφανές όπως παρουσιάζει ο παραπάνω πίνακας από το γεγονός ότι η συνάρτηση απόφασης στην κλάση 1 δίνει θετικές τιμές μόλις σε τρία από τα δέκα στιγμιότυπα. Από την άλλη πλευρά η ίδια συνάρτηση απόφασης δίνει αρνητικές τιμές σε 89 από τα 90 στιγμιότυπα τα οποία ανήκουν στα αρνητικά δεδομένα. Αυτό είναι λογικό γιατί τα δεδομένα στα οποία εκπαιδεύτηκε ο ταξινομητής δεν ήταν ισοζυγισμένα και αυτό δεν βοήθησε στη σωστή εκπαίδευσή του. Αν σκεφτούμε τον τρόπο που έχει εκπαιδευτεί ο κάθε ταξινομητής όπως περιγράψαμε σε προηγούμενο κεφάλαιο χρησιμοποιήσαμε 90% των δεδομένων για αρνητικά και μόλις το 10% για θετικά αποτελέσματα.

Επίσης ήταν λογικό το ποσοστό επιτυχίας σωστής ταξινόμησης για τον κάθε επιμέρους ταξινομητή να είναι περίπου 90% αφού πραγματικά τα αρνητικά δεδομένα ταξινομούνταν σωστά και αποτελούσαν τον 90% του συνόλου των δεδομένων προς έλεγχο. Οπότε ενώ φαινόταν να δουλεύει σωστά ο κάθε επιμέρους ταξινομητής στην πραγματικότητα αυτό δεν ίσχυε οπότε και ο τελικός υπέρ-ταξινομητής δε μας έδωσε ικανοποιητικά αποτελέσματα. Γι' αυτό το λόγο δε θα παραθέσουμε τα υπόλοιπα πειραματικά αποτελέσματα για τα υπόλοιπα δεδομένα που διαθέτουμε και για τους υπόλοιπους ταξινομητές υλοποιημένους με νευρωνικά δίκτυα. Στην πραγματικότητα αυτή η υλοποίηση του υπέρ-ταξινομητή έχει σχεδιαστικό λάθος το οποίο δεν επιτρέπει να πάρουμε καλά αποτελέσματα για τον υπέρ-ταξινομητή μας οπότε δε χρήζει και περισσότερης έρευνας.

Στην δεύτερη σχεδιαστική προσπάθειά μας επιχειρούμε να εκπαιδεύσουμε τους ταξινομητές μας με ισοζυγισμένα δεδομένα όπως έχουμε περιγράψει. Το όλο εγχείρημά μας θα το περιγράψουμε στην επόμενη παράγραφο.

7.2 Μέθοδος 'One against all balanced'

7.2.1 Ταξινομητής SVM

Μουσικά δεδομένα 30 χαρακτηριστικών

Μετά την εκπαίδευση των ταξινομητών που υλοποιήσαμε όπως περιγράψαμε σε προηγούμενο κεφάλαιο, τους ελέγχουμε σύμφωνα με τα δεδομένα ελέγχου που δημιουργήσαμε από το σύνολο των δεδομένων. Τα ποσοστά σωστής ταξινόμησης για τον κάθε επιμέρους ταξινομητή και το αντίστοιχο Sub-Fold του Fold 1 παρουσιάζονται στον παρακάτω πίνακα:

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,8	0,82	0,79	0,82	0,85
Ταξινομητής2	0,88	0,84	0,9	0,89	0,9
Ταξινομητής3	0,79	0,84	0,84	0,83	0,79
Ταξινομητής4	0,72	0,79	0,74	0,77	0,71
Ταξινομητής5	0,78	0,78	0,77	0,78	0,77
Ταξινομητής6	0,8	0,77	0,79	0,73	0,77
Ταξινομητής7	0,8	0,81	0,81	0,82	0,77
Ταξινομητής8	0,83	0,83	0,82	0,83	0,83
Ταξινομητής9	0,83	0,8	0,83	0,85	0,81
Ταξινομητής10	0,67	0,65	0,6	0,55	0,67
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,79	0,87	0,81	0,86	0,72
Ταξινομητής2	0,89	0,89	0,91	0,85	0,87
Ταξινομητής3	0,85	0,83	0,83	0,76	0,76
Ταξινομητής4	0,76	0,78	0,74	0,78	0,78
Ταξινομητής5	0,78	0,79	0,79	0,78	0,77
Ταξινομητής6	0,76	0,83	0,84	0,76	0,76
Ταξινομητής7	0,84	0,8	0,84	0,82	0,82
Ταξινομητής8	0,82	0,81	0,77	0,82	0,82

Ταξινομητής9	0,79	0,76	0,83	0,82	0,8
Ταξινομητής10	0,72	0,69	0,63	0,6	0,63

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 1

Παρατηρώντας τα αποτελέσματα επί τοις εκατό σωστής ταξινόμησης φαίνεται ότι ο κάθε ταξινομητής έχει εκπαιδευτεί ικανοποιητικά και ταξινομεί τα άγνωστα για αυτόν στιγμιότυπα που του δίνονται με αρκετά μεγάλο ποσοστό επιτυχίας. Στους επόμενους πίνακες θα παρουσιάσουμε τα αποτελέσματα σωστής ταξινόμησης και των υπόλοιπων Folds που αντιστοιχούν στους ταξινομητές που έχουμε δημιουργήσει:

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,69	0,7	0,61	0,69	0,72
Ταξινομητής2	0,87	0,83	0,82	0,85	0,83
Ταξινομητής3	0,76	0,76	0,81	0,77	0,76
Ταξινομητής4	0,8	0,78	0,78	0,78	0,8
Ταξινομητής5	0,75	0,76	0,74	0,77	0,79
Ταξινομητής6	0,75	0,76	0,78	0,81	0,77
Ταξινομητής7	0,89	0,88	0,86	0,85	0,85
Ταξινομητής8	0,87	0,85	0,88	0,87	0,84
Ταξινομητής9	0,72	0,73	0,77	0,75	0,76
Ταξινομητής10	0,73	0,71	0,65	0,71	0,73
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,73	0,72	0,68	0,75	0,75
Ταξινομητής2	0,91	0,85	0,82	0,85	0,87
Ταξινομητής3	0,73	0,72	0,77	0,73	0,74
Ταξινομητής4	0,83	0,81	0,77	0,79	0,8
Ταξινομητής5	0,78	0,72	0,73	0,76	0,74
Ταξινομητής6	0,76	0,78	0,78	0,78	0,78
Ταξινομητής7	0,83	0,87	0,89	0,88	0,86
Ταξινομητής8	0,86	0,86	0,87	0,85	0,88
Ταξινομητής9	0,78	0,82	0,75	0,77	0,73
Ταξινομητής10	0,7	0,73	0,68	0,68	0,72

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 2

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,71	0,69	0,68	0,74	0,73
Ταξινομητής2	0,85	0,9	0,89	0,91	0,91
Ταξινομητής3	0,81	0,79	0,77	0,86	0,75
Ταξινομητής4	0,74	0,77	0,73	0,76	0,75
Ταξινομητής5	0,67	0,66	0,7	0,7	0,67
Ταξινομητής6	0,87	0,78	0,81	0,79	0,8
Ταξινομητής7	0,86	0,84	0,83	0,84	0,86
Ταξινομητής8	0,86	0,87	0,84	0,84	0,88
Ταξινομητής9	0,79	0,78	0,74	0,76	0,8
Ταξινομητής10	0,79	0,77	0,7	0,68	0,7
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,72	0,73	0,75	0,73	0,68
Ταξινομητής2	0,91	0,89	0,89	0,87	0,88
Ταξινομητής3	0,8	0,83	0,81	0,81	0,82
Ταξινομητής4	0,77	0,75	0,77	0,74	0,75
Ταξινομητής5	0,69	0,69	0,72	0,69	0,65

Ταξινομητής6	0,82	0,83	0,81	0,84	0,78
Ταξινομητής7	0,85	0,86	0,89	0,83	0,86
Ταξινομητής8	0,86	0,86	0,85	0,87	0,89
Ταξινομητής9	0,83	0,81	0,77	0,75	0,83
Ταξινομητής10	0,76	0,69	0,71	0,69	0,73

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 3

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,64	0,73	0,71	0,69	0,71
Ταξινομητής2	0,82	0,81	0,85	0,87	0,89
Ταξινομητής3	0,71	0,79	0,67	0,77	0,76
Ταξινομητής4	0,7	0,68	0,74	0,68	0,69
Ταξινομητής5	0,74	0,78	0,77	0,75	0,75
Ταξινομητής6	0,76	0,76	0,83	0,82	0,78
Ταξινομητής7	0,8	0,85	0,8	0,8	0,79
Ταξινομητής8	0,71	0,78	0,74	0,76	0,78
Ταξινομητής9	0,87	0,87	0,85	0,91	0,9
Ταξινομητής10	0,7	0,67	0,7	0,71	0,73
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,76	0,72	0,75	0,69	0,64
Ταξινομητής2	0,86	0,87	0,85	0,86	0,87
Ταξινομητής3	0,76	0,73	0,64	0,68	0,76
Ταξινομητής4	0,7	0,73	0,71	0,69	0,63
Ταξινομητής5	0,75	0,8	0,75	0,82	0,73
Ταξινομητής6	0,86	0,81	0,82	0,82	0,78
Ταξινομητής7	0,86	0,81	0,8	0,8	0,82
Ταξινομητής8	0,72	0,83	0,77	0,79	0,81
Ταξινομητής9	0,88	0,9	0,88	0,84	0,89
Ταξινομητής10	0,65	0,64	0,69	0,63	0,65

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 4

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,68	0,66	0,76	0,71	0,68
Ταξινομητής2	0,87	0,84	0,86	0,86	0,87
Ταξινομητής3	0,75	0,8	0,76	0,84	0,8
Ταξινομητής4	0,77	0,75	0,8	0,76	0,8
Ταξινομητής5	0,74	0,74	0,74	0,72	0,72
Ταξινομητής6	0,77	0,77	0,81	0,76	0,76
Ταξινομητής7	0,91	0,89	0,82	0,85	0,87
Ταξινομητής8	0,86	0,86	0,85	0,84	0,88
Ταξινομητής9	0,77	0,88	0,78	0,74	0,75
Ταξινομητής10	0,57	0,68	0,69	0,59	0,72
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,69	0,71	0,7	0,65	0,69
Ταξινομητής2	0,82	0,85	0,85	0,87	0,86
Ταξινομητής3	0,74	0,8	0,72	0,75	0,77
Ταξινομητής4	0,75	0,8	0,77	0,77	0,79
Ταξινομητής5	0,75	0,73	0,65	0,79	0,74
Ταξινομητής6	0,79	0,82	0,83	0,83	0,78

Ταξινομητής7	0,85	0,89	0,82	0,85	0,88
Ταξινομητής8	0,87	0,88	0,82	0,86	0,87
Ταξινομητής9	0,79	0,8	0,79	0,76	0,75
Ταξινομητής10	0,59	0,7	0,66	0,69	0,66

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 5

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,78	0,76	0,75	0,8	0,77
Ταξινομητής2	0,86	0,92	0,89	0,89	0,85
Ταξινομητής3	0,79	0,77	0,77	0,8	0,81
Ταξινομητής4	0,72	0,73	0,78	0,79	0,76
Ταξινομητής5	0,78	0,74	0,75	0,71	0,68
Ταξινομητής6	0,77	0,86	0,82	0,81	0,82
Ταξινομητής7	0,81	0,77	0,82	0,8	0,81
Ταξινομητής8	0,89	0,88	0,88	0,88	0,92
Ταξινομητής9	0,74	0,76	0,8	0,81	0,78
Ταξινομητής10	0,63	0,61	0,64	0,61	0,57
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,72	0,76	0,78	0,78	0,73
Ταξινομητής2	0,85	0,87	0,9	0,91	0,92
Ταξινομητής3	0,82	0,78	0,78	0,79	0,83
Ταξινομητής4	0,8	0,79	0,78	0,76	0,75
Ταξινομητής5	0,71	0,7	0,72	0,76	0,72
Ταξινομητής6	0,82	0,8	0,82	0,81	0,83
Ταξινομητής7	0,79	0,79	0,77	0,78	0,78
Ταξινομητής8	0,9	0,9	0,85	0,9	0,9
Ταξινομητής9	0,76	0,76	0,77	0,79	0,83
Ταξινομητής10	0,64	0,64	0,64	0,62	0,63

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 6

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,73	0,61	0,65	0,65	0,64
Ταξινομητής2	0,9	0,9	0,93	0,93	0,94
Ταξινομητής3	0,79	0,73	0,78	0,78	0,81
Ταξινομητής4	0,7	0,72	0,75	0,7	0,75
Ταξινομητής5	0,73	0,75	0,73	0,73	0,7
Ταξινομητής6	0,87	0,83	0,82	0,83	0,84
Ταξινομητής7	0,83	0,84	0,82	0,85	0,8
Ταξινομητής8	0,89	0,86	0,84	0,86	0,84
Ταξινομητής9	0,78	0,76	0,81	0,82	0,78
Ταξινομητής10	0,67	0,61	0,58	0,66	0,6
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,67	0,66	0,7	0,58	0,66
Ταξινομητής2	0,92	0,93	0,95	0,92	0,93
Ταξινομητής3	0,78	0,82	0,72	0,78	0,72
Ταξινομητής4	0,78	0,7	0,68	0,71	0,77
Ταξινομητής5	0,68	0,73	0,75	0,72	0,7
Ταξινομητής6	0,81	0,81	0,82	0,75	0,84
Ταξινομητής7	0,78	0,82	0,83	0,86	0,8

Ταξινομητής8	0,86	0,8	0,86	0,83	0,85
Ταξινομητής9	0,81	0,79	0,79	0,78	0,83
Ταξινομητής10	0,65	0,71	0,68	0,6	0,63

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 7

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,65	0,62	0,64	0,61	0,61
Ταξινομητής2	0,9	0,92	0,9	0,94	0,91
Ταξινομητής3	0,85	0,79	0,72	0,78	0,74
Ταξινομητής4	0,8	0,73	0,78	0,74	0,79
Ταξινομητής5	0,71	0,7	0,79	0,79	0,72
Ταξινομητής6	0,89	0,81	0,82	0,83	0,79
Ταξινομητής7	0,85	0,81	0,83	0,84	0,81
Ταξινομητής8	0,78	0,81	0,78	0,81	0,81
Ταξινομητής9	0,83	0,81	0,84	0,75	0,78
Ταξινομητής10	0,69	0,65	0,7	0,6	0,71
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,61	0,67	0,63	0,63	0,69
Ταξινομητής2	0,93	0,93	0,87	0,92	0,93
Ταξινομητής3	0,7	0,75	0,71	0,7	0,8
Ταξινομητής4	0,77	0,8	0,77	0,76	0,75
Ταξινομητής5	0,75	0,78	0,7	0,82	0,8
Ταξινομητής6	0,85	0,81	0,83	0,84	0,81
Ταξινομητής7	0,84	0,83	0,85	0,78	0,83
Ταξινομητής8	0,78	0,77	0,8	0,77	0,77
Ταξινομητής9	0,75	0,79	0,85	0,83	0,83
Ταξινομητής10	0,73	0,75	0,74	0,67	0,63

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 8

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,68	0,64	0,65	0,71	0,74
Ταξινομητής2	0,88	0,91	0,9	0,88	0,9
Ταξινομητής3	0,75	0,74	0,71	0,78	0,78
Ταξινομητής4	0,64	0,71	0,71	0,64	0,64
Ταξινομητής5	0,74	0,69	0,76	0,73	0,75
Ταξινομητής6	0,84	0,82	0,82	0,83	0,81
Ταξινομητής7	0,89	0,86	0,87	0,86	0,9
Ταξινομητής8	0,8	0,86	0,81	0,79	0,83
Ταξινομητής9	0,75	0,8	0,79	0,83	0,8
Ταξινομητής10	0,72	0,63	0,72	0,61	0,64
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,69	0,82	0,69	0,73	0,68
Ταξινομητής2	0,91	0,89	0,88	0,87	0,93
Ταξινομητής3	0,68	0,71	0,76	0,69	0,7
Ταξινομητής4	0,7	0,7	0,65	0,67	0,71
Ταξινομητής5	0,69	0,79	0,67	0,75	0,77
Ταξινομητής6	0,82	0,81	0,83	0,85	0,8
Ταξινομητής7	0,84	0,9	0,87	0,87	0,9
Ταξινομητής8	0,8	0,82	0,82	0,81	0,81

Ταξινομητής9	0,8	0,79	0,81	0,73	0,83
Ταξινομητής10	0,67	0,7	0,66	0,68	0,68

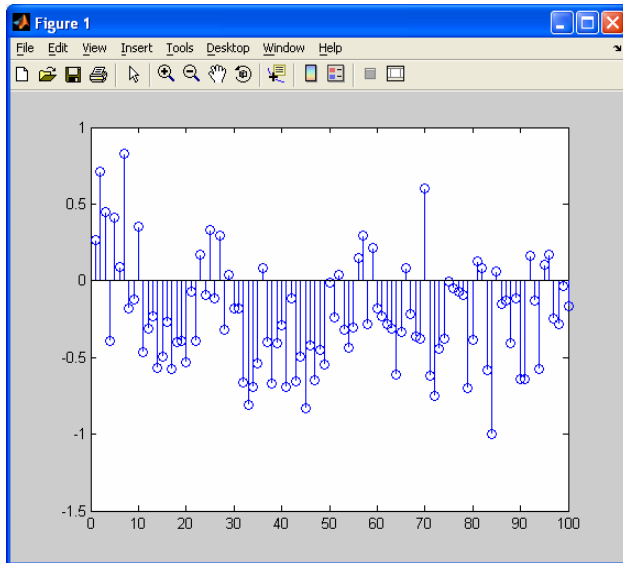
Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 9

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,72	0,72	0,7	0,75	0,79
Ταξινομητής2	0,86	0,86	0,88	0,89	0,89
Ταξινομητής3	0,71	0,73	0,76	0,75	0,76
Ταξινομητής4	0,82	0,77	0,83	0,78	0,75
Ταξινομητής5	0,76	0,83	0,81	0,79	0,82
Ταξινομητής6	0,76	0,75	0,9	0,83	0,85
Ταξινομητής7	0,79	0,75	0,8	0,79	0,74
Ταξινομητής8	0,9	0,88	0,88	0,92	0,9
Ταξινομητής9	0,79	0,83	0,8	0,8	0,79
Ταξινομητής10	0,74	0,64	0,64	0,56	0,63
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,76	0,73	0,81	0,81	0,72
Ταξινομητής2	0,84	0,89	0,88	0,9	0,9
Ταξινομητής3	0,78	0,76	0,74	0,72	0,72
Ταξινομητής4	0,79	0,82	0,75	0,83	0,8
Ταξινομητής5	0,8	0,83	0,77	0,82	0,83
Ταξινομητής6	0,8	0,87	0,78	0,83	0,78
Ταξινομητής7	0,76	0,78	0,71	0,66	0,74
Ταξινομητής8	0,85	0,87	0,89	0,91	0,86
Ταξινομητής9	0,78	0,81	0,78	0,83	0,81
Ταξινομητής10	0,66	0,63	0,61	0,59	0,54

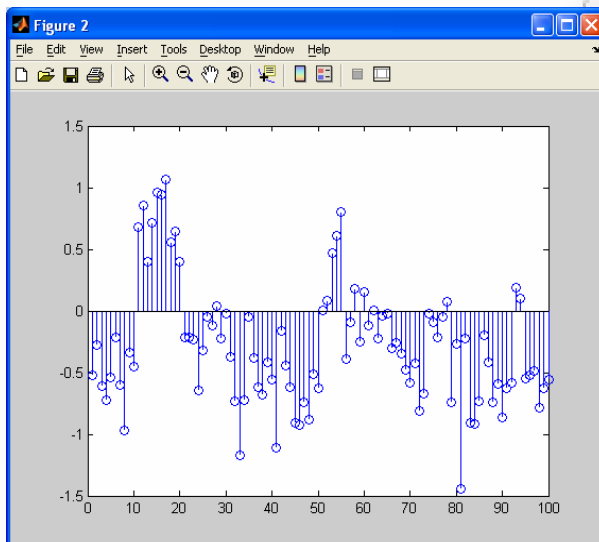
Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 10

Όπως ενεργήσαμε και στην πρώτη σχεδιάσή μας προκειμένου να δημιουργήσουμε τον τελικό υπέρ-ταξινομητή ο οποίος θα ταξινομεί ένα στιγμιότυπο μεταξύ των κλάσεων θα εκμεταλλευτούμε από κάθε ταξινομητή την τιμή της συνάρτησης απόφασης. Αυτή η τιμή για κάθε στιγμιότυπο όπως έχουμε εξηγήσει προσδιορίζει κατά πόσο είναι σίγουρος ο κάθε ταξινομητής ότι έχει ταξινομήσει σωστά ένα στιγμιότυπο που του έχει δοθεί. Αν το στιγμιότυπο ανήκει στα θετικά δεδομένα τότε αυτή η τιμή θα έχει θετικό πρόσημο διαφορετικά θα έχει αρνητικό πρόσημο. Οπότε εμείς δεν έχουμε να κάνουμε τίποτε περισσότερο από το να επιλέξουμε τον ταξινομητή αυτό ο οποίος θα έχει δώσει τη μεγαλύτερη τιμή στην συνάρτηση απόφασης. Η σκέψη μας είναι ότι η τιμή της συνάρτησης απόφασης θα είναι μεγαλύτερη στον ταξινομητή που έχει εκπαιδευτεί να ταξινομεί στιγμιότυπα της συγκεκριμένης κλάσης.

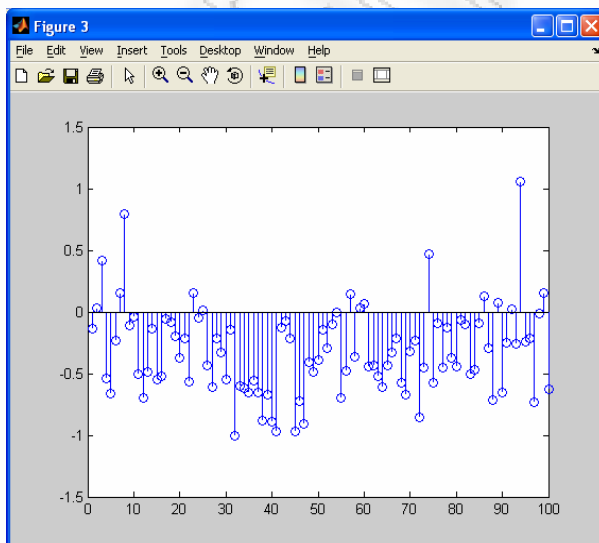
Στη συνέχεια θα παρουσιάσουμε ενδεικτικά τις τιμές της συνάρτησης απόφασης του Sub-Fold 1 του Fold 1 για τον κάθε ταξινομητή με τα αντίστοιχα δεδομένα ελέγχου. Θα πρέπει να σημειώσουμε ότι είναι αδύνατο να παρουσιάσουμε όλα τα αποτελέσματα με τις τιμές της συνάρτησης απόφασης γιατί είναι υπερβολικά πολλά. Όμως περίπου με τον ίδιο τρόπο συμπεριφέρονται οι ταξινομητές σε όλα τα Folds και τα Sub-Folds που έχουμε δημιουργήσει όπως έχουμε περιγράψει σε προηγούμενο κεφάλαιο. Επίσης πρέπει να πούμε ότι τα στιγμιότυπα ελέγχου για την κάθε κλάση δεδομένων για τα συγκεκριμένα μουσικά δεδομένα είναι 10 σε πλήθος και δίνονται με τη σειρά στον κάθε ταξινομητή. Οπότε οι πρώτες 10 τιμές της συνάρτησης απόφασης αντιστοιχούν στην πρώτη κλάση οι δεύτερες 10 τιμές της συνάρτησης απόφασης ανήκουν στη δεύτερη κλάση μας μέχρι τις τελευταίες 10 τιμές που ανήκουν στην 10^η κλάση μας.



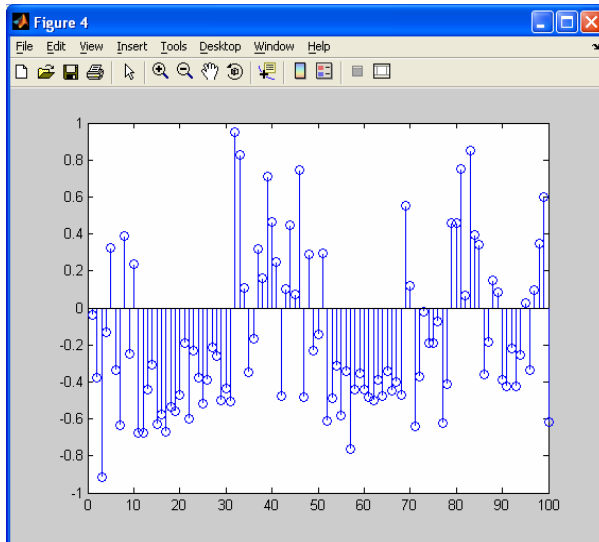
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 1



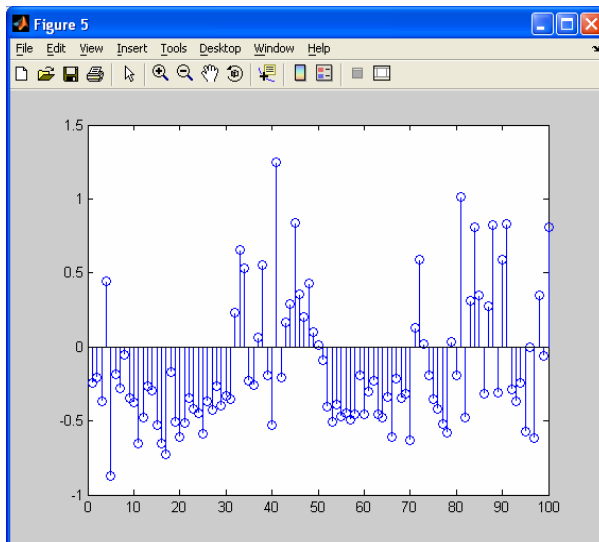
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 2



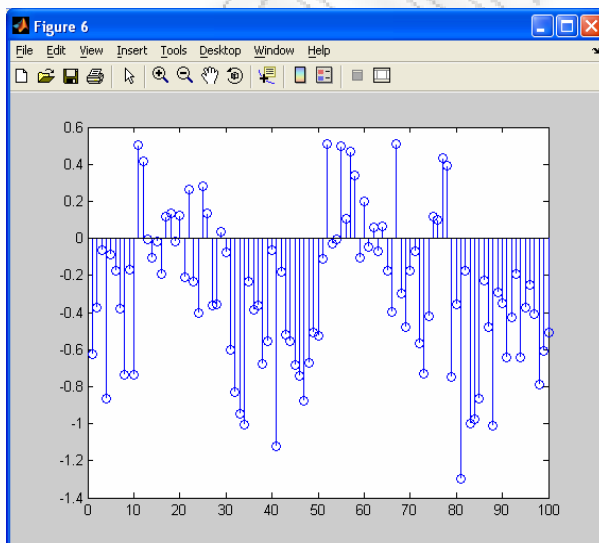
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 3



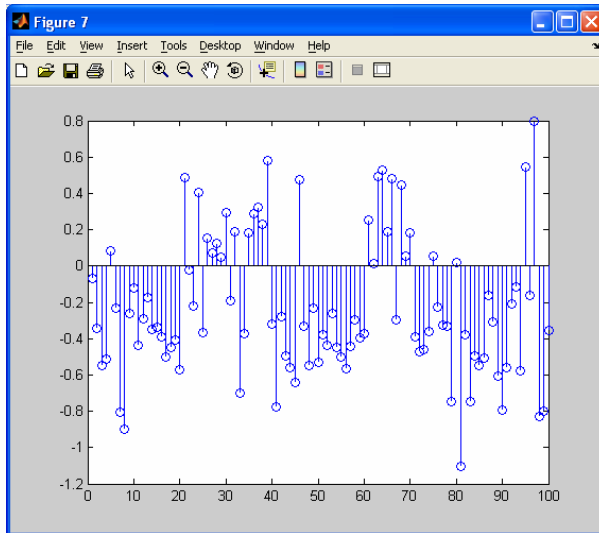
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 4



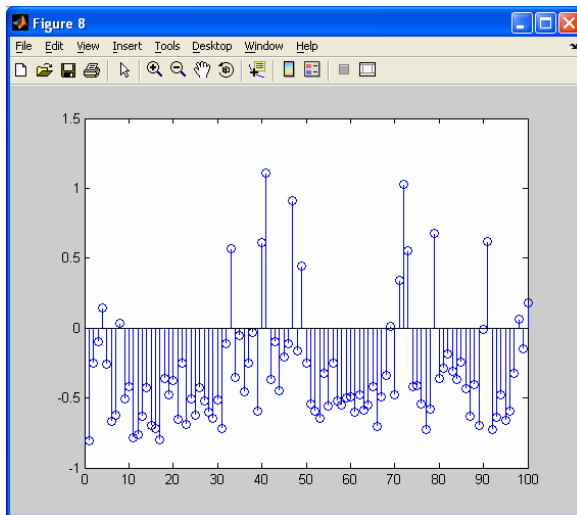
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 5



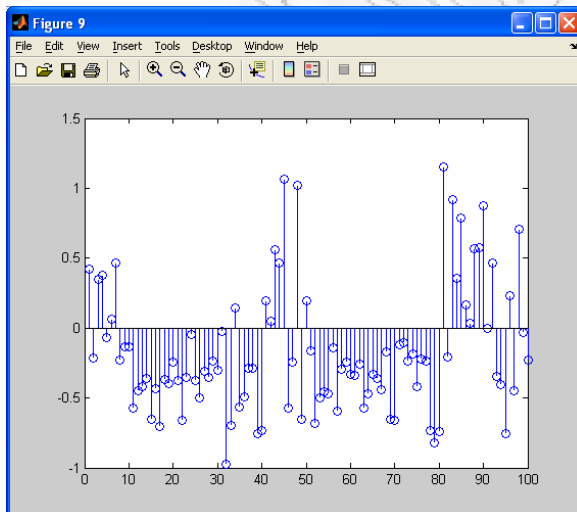
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 6



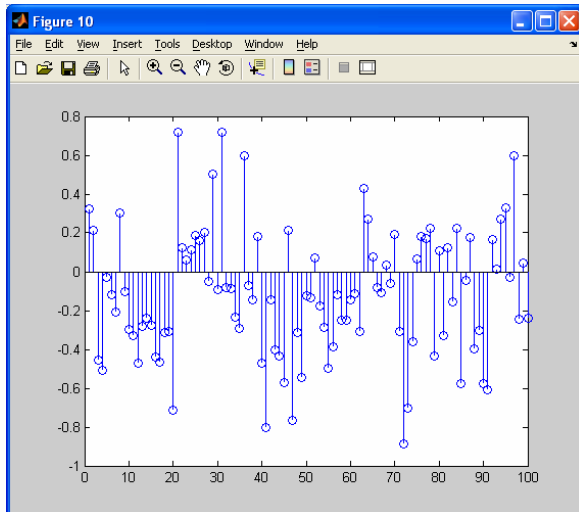
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 7



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 8



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 9



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 10

Όπως φαίνεται από τα προηγούμενα σχήματα ο ταξινομητής ο οποίος λειτουργεί καλύτερα είναι αυτός της δεύτερης κλάσης δεδομένων. Αυτός που δείχνει να κάνει τα περισσότερα λάθη για τα συγκεκριμένα δεδομένα είναι ο ταξινομητής 10 ο οποίος πιθανότατα να μπερδέψει περισσότερο από όλους τον υπέρ-ταξινομητή μας στην τελική του απόφαση.

Στους παρακάτω πίνακες που ακολουθούν παρουσιάζονται τα confusion matrixes για τους αντίστοιχους ταξινομητές του Sub-Fold 1 του Fold 1 όπως παρουσιάσαμε λίγο παραπάνω.

Ταξινομητής 1	Positive Class	Negative Class
Classified as Positive	0,7	0,3
Classified as Negative	0,18889	0,81111

Ταξινομητής 2	Positive Class	Negative Class
Classified as Positive	1	0
Classified as Negative	0,13333	0,86667

Ταξινομητής 3	Positive Class	Negative Class
Classified as Positive	0,2	0,8
Classified as Negative	0,14444	0,85556

Ταξινομητής 4	Positive Class	Negative Class
Classified as Positive	0,7	0,3
Classified as Negative	0,27778	0,72222

Ταξινομητής 5	Positive Class	Negative Class
Classified as Positive	0,9	0,1
Classified as Negative	0,23333	0,76667

Ταξινομητής 6	Positive Class	Negative Class
Classified as Positive	0,6	0,4
Classified as Negative	0,17778	0,82222

Ταξινομητής 7	Positive Class	Negative Class
Classified as Positive	0,9	0,1
Classified as Negative	0,21111	0,78889

Ταξινομητής 8	Positive Class	Negative Class
Classified as Positive	0,4	0,6
Classified as Negative	0,12222	0,87778

Ταξινομητής 9	Positive Class	Negative Class
Classified as Positive	0,9	0,1
Classified as Negative	0,17778	0,82222

Ταξινομητής 10	Positive Class	Negative Class
Classified as Positive	0,6	0,4
Classified as Negative	0,32222	0,67778

Πίνακες: Confusion matrixes των αντίστοιχων ταξινομητών

Όπως φαίνεται και από τα confusion matrixes γίνεται φανερό ότι ο ταξινομητής που δουλεύει καλύτερα είναι ο ταξινομητής 2 ενώ ο χειρότερος είναι ο ταξινομητής 10. Στον επόμενο πίνακα που ακολουθεί παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή για τα 10 διαφορετικά Folds που έχουμε δημιουργήσει. Τα αποτελέσματα αυτά είναι σε μορφή επί τοις εκατό.

Fold 1	44%
Fold 2	47%
Fold 3	49%
Fold 4	46%
Fold 5	51%
Fold 6	40%
Fold 7	44%
Fold 8	46%
Fold 9	45%
Fold 10	43%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή

Όπως φαίνεται από τον παραπάνω πίνακα ο υπέρ-ταξινομητής του Fold 5 δουλεύει καλύτερα από όλους. Τα αποτελέσματα αυτά είναι σίγουρα καλύτερα από αυτά του σχεδιασμού του πρώτου ταξινομητή αλλά σίγουρα όχι τόσο ικανοποιητικά όσο θα θέλαμε. Σε αυτό το γεγονός βέβαια συμβάλει και η φύση των δεδομένων που χρησιμοποιήσαμε. Στη συνέχεια θα παρουσιάσουμε τα αποτελέσματα του ίδιου ταξινομητή για τα ίδια μουσικά αρχεία όμως με στιγμιότυπα που περιέχουν πολύ περισσότερα χαρακτηριστικά και συνεπώς καλύτερη ποιότητα πληροφορίας.

Μουσικά δεδομένα 280 χαρακτηριστικών

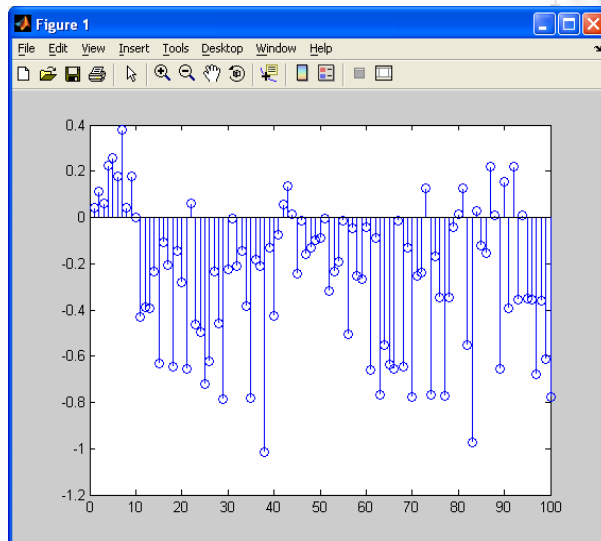
Για λόγους καθαρά πρακτικούς δε θα παρουσιάσουμε όπως πράξαμε προηγουμένως όλα τα ποσοστά σωστής ταξινόμησης για τον κάθε επιμέρους ταξινομητή και το αντίστοιχο Sub-Fold του κάθε Fold γιατί τα αποτελέσματα είναι υπερβολικά πολλά. Ενδεικτικά θα παρουσιάσουμε μόνο τα αποτελέσματα που αφορούν το Fold 2. Με αυτή τη λογική θα συνεχίσουμε να παρουσιάζουμε όλα τα επόμενα αποτελέσματά μας. Έτσι λοιπόν στον παρακάτω πίνακα φαίνονται οι επιμέρους ταξινομητές και πως τα έχουν πάει στο Fold 2:

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,87	0,84	0,89	0,69	0,78
Ταξινομητής2	0,98	0,98	0,99	0,97	0,99
Ταξινομητής3	0,92	0,91	0,94	0,93	0,92
Ταξινομητής4	0,83	0,84	0,81	0,88	0,81
Ταξινομητής5	0,86	0,89	0,87	0,86	0,81
Ταξινομητής6	0,95	0,82	0,83	0,81	0,82
Ταξινομητής7	0,92	0,92	0,89	0,91	0,91
Ταξινομητής8	0,89	0,88	0,87	0,89	0,83

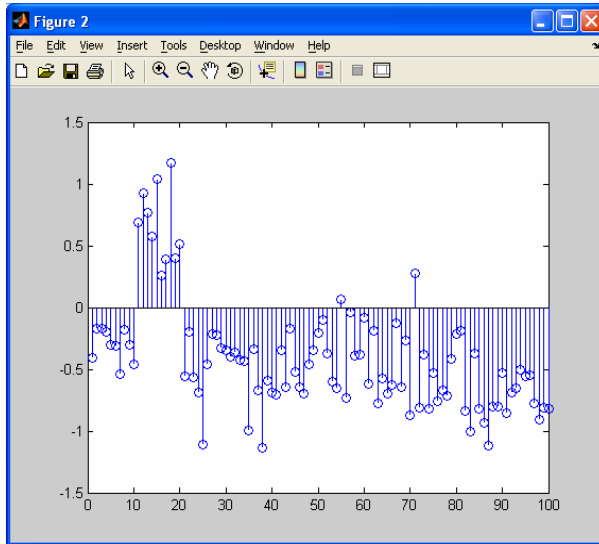
Ταξινομητής9	0,91	0,9	0,91	0,89	0,9
Ταξινομητής10	0,79	0,75	0,79	0,78	0,74
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,8	0,86	0,88	0,69	0,75
Ταξινομητής2	0,99	0,97	0,99	0,99	0,98
Ταξινομητής3	0,93	0,93	0,93	0,94	0,91
Ταξινομητής4	0,86	0,82	0,78	0,86	0,85
Ταξινομητής5	0,86	0,88	0,88	0,87	0,9
Ταξινομητής6	0,89	0,78	0,85	0,85	0,88
Ταξινομητής7	0,89	0,89	0,9	0,9	0,92
Ταξινομητής8	0,88	0,84	0,88	0,85	0,87
Ταξινομητής9	0,92	0,91	0,92	0,92	0,92
Ταξινομητής10	0,77	0,72	0,81	0,74	0,76

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 2

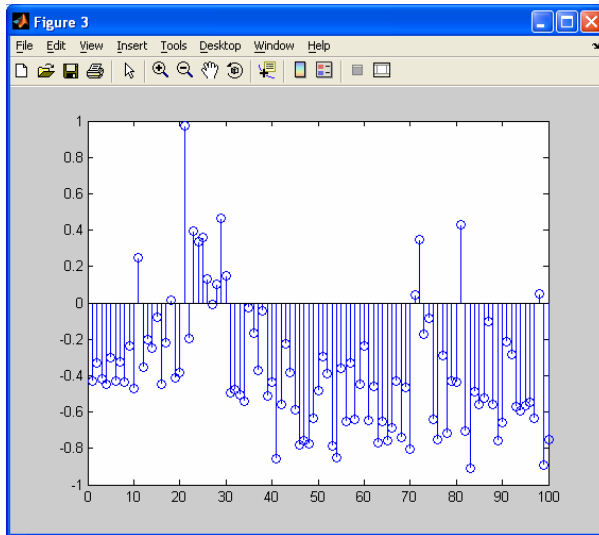
Ήταν λογικό και αναμενόμενο να έχουμε καλύτερα αποτελέσματα από αυτά που είχαμε με τα μουσικά δεδομένα των 30 χαρακτηριστικών. Στη συνέχεια θα παρουσιάσω και τις τιμές της συνάρτησης απόφασης για τα αντίστοιχα δεδομένα ελέγχου του κάθε ταξινομητή που αφορούν το Fold 2 και τον Sub-Fold 1.



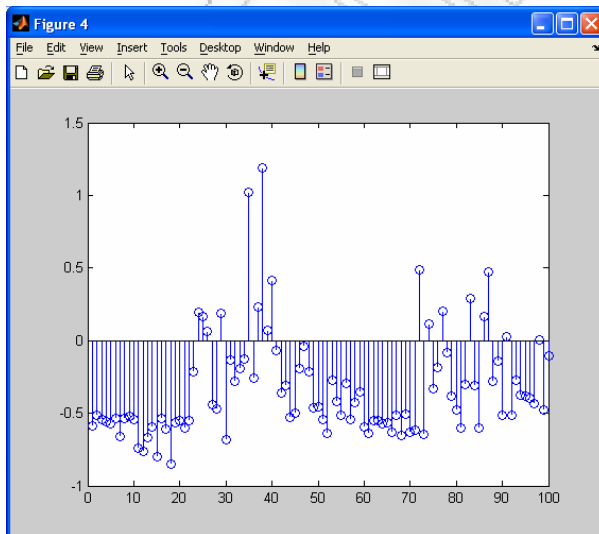
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 1



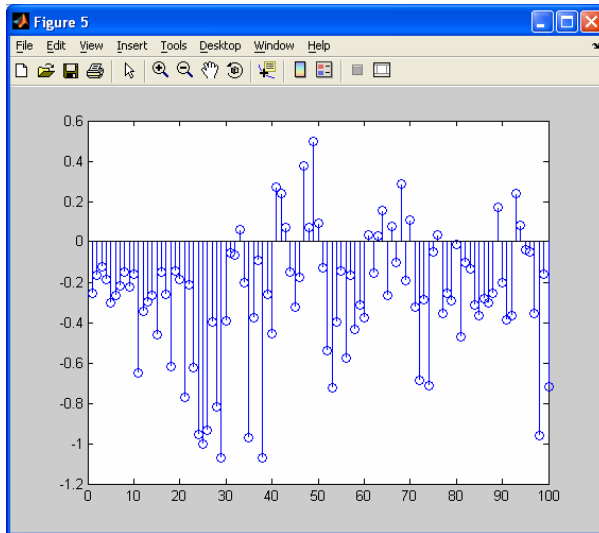
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 2



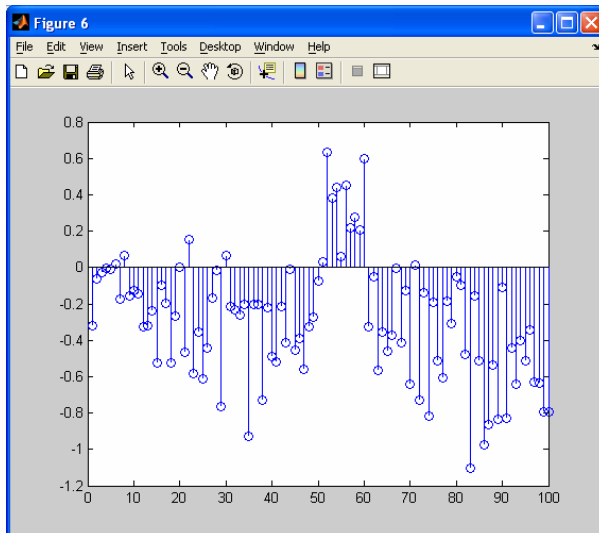
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 3



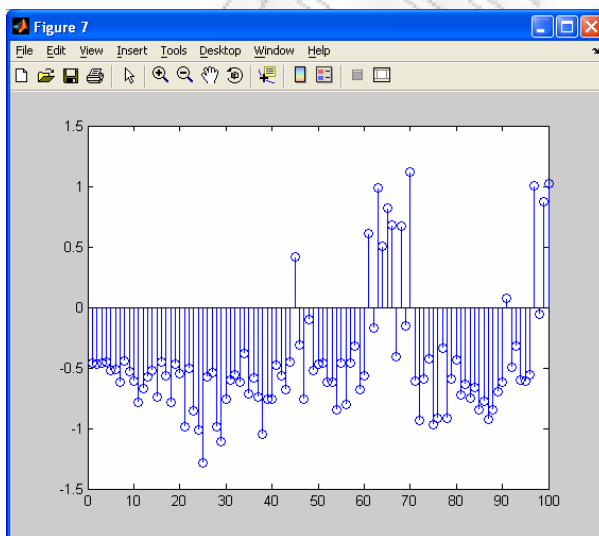
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 4



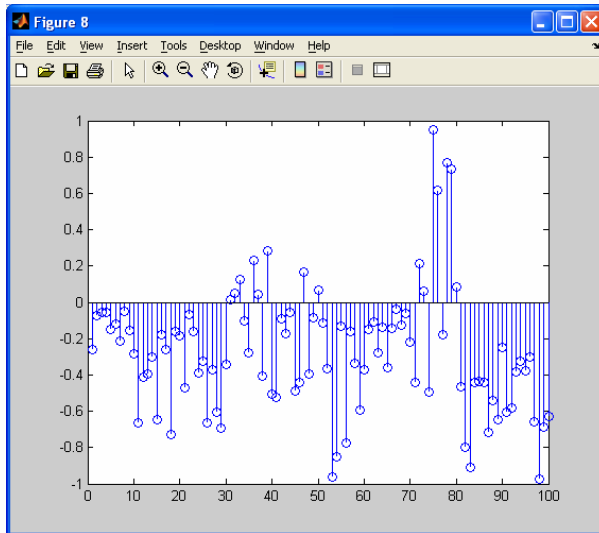
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 5



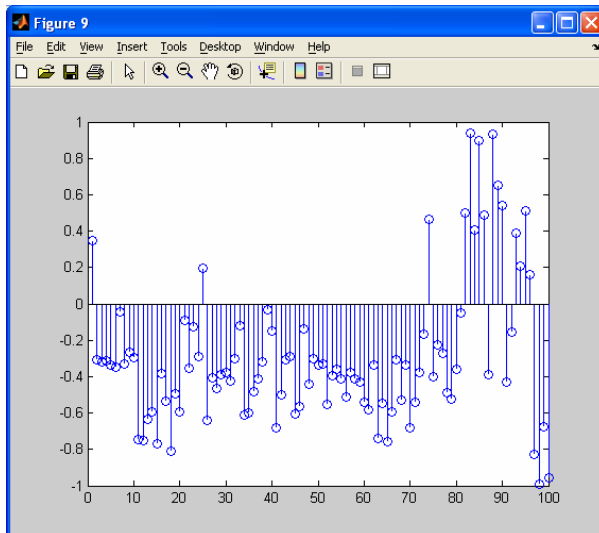
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 6



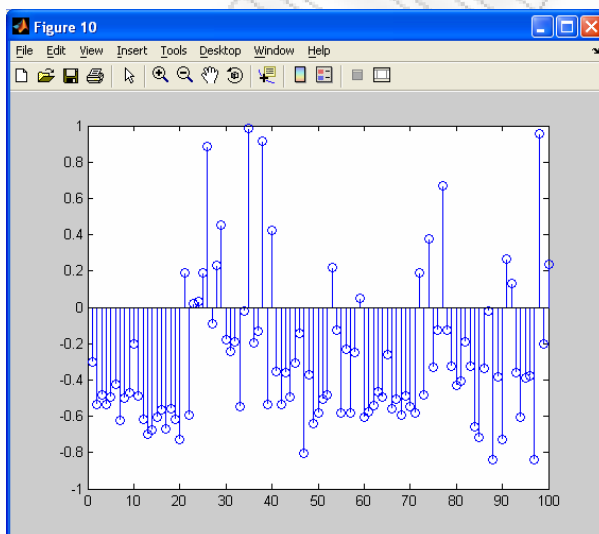
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 7



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 8



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 9



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 10

Από τα παραπάνω σχήματα γίνεται φανερό ότι έχουμε πολύ καλύτερα αποτελέσματα στους επιμέρους ταξινομητές απ' ό,τι είχαμε με τα προηγούμενα δεδομένα μας. Όμως και πάλι η τάση των

ταξινομητών είναι ίδια. Ο ταξινομητής 10 εξακολουθεί να έχει τα χειρότερα αποτελέσματα και ο ταξινομητής 2 τα καλύτερα. Αυτό πρέπει να οφείλεται στα δεδομένα μας στα οποία δείχνει η κλάση 10 να περιέχει τη μεγαλύτερη ασάφεια. Στη συνέχεια θα παρουσιάσουμε τα confusion matrixes για τους παραπάνω ταξινομητές που ελέγχονται στα αντίστοιχα δεδομένα.

Ταξινομητής 1	Positive Class	Negative Class
Classified as Positive	1	0
Classified as Negative	0,14444	0,85556

Ταξινομητής 2	Positive Class	Negative Class
Classified as Positive	1	0
Classified as Negative	0,022222	0,97778

Ταξινομητής 3	Positive Class	Negative Class
Classified as Positive	0,8	0,2
Classified as Negative	0,066667	0,93333

Ταξινομητής 4	Positive Class	Negative Class
Classified as Positive	0,5	0,5
Classified as Negative	0,13333	0,86667

Ταξινομητής 5	Positive Class	Negative Class
Classified as Positive	0,7	0,3
Classified as Negative	0,12222	0,87778

Ταξινομητής 6	Positive Class	Negative Class
Classified as Positive	1	0
Classified as Negative	0,055556	0,94444

Ταξινομητής 7	Positive Class	Negative Class
Classified as Positive	0,7	0,3
Classified as Negative	0,055556	0,94444

Ταξινομητής 8	Positive Class	Negative Class
Classified as Positive	0,7	0,3
Classified as Negative	0,088889	0,91111

Ταξινομητής 9	Positive Class	Negative Class
Classified as Positive	0,8	0,2
Classified as Negative	0,077778	0,92222

Ταξινομητής 10	Positive Class	Negative Class
Classified as Positive	0,4	0,6
Classified as Negative	0,16667	0,83333

Πίνακες: Confusion matrixes των αντίστοιχων ταξινομητών

Όπως φαίνεται και πάλι από τα confusion matrixes ο ταξινομητής που δουλεύει καλύτερα είναι ο ταξινομητής 2 ενώ ο χειρότερος είναι ο ταξινομητής 10. Στον επόμενο πίνακα που ακολουθεί παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή για τα 10 διαφορετικά Folds που έχουμε δημιουργήσει. Τα αποτελέσματα αυτά είναι σε μορφή επί τοις εκατό.

Fold 1	57%
Fold 2	67%
Fold 3	73%

Fold 4	62%
Fold 5	58%
Fold 6	59%
Fold 7	55%
Fold 8	62%
Fold 9	59%
Fold 10	57%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή

Παρατηρώντας τον πίνακα ο υπέρ-ταξινομητής για το Fold 3 δουλεύει καλύτερα από όλους. Στα προηγούμενα δεδομένα μας με τα λιγότερα χαρακτηριστικά ανά στιγμιότυπο είχαμε διαπιστώσει ότι καλύτερα δούλεψε ο υπέρ-ταξινομητής για το Fold 5. Αν και δεν είναι τελειώς ικανοποιητικά αυτά τα αποτελέσματα για έναν ταξινομητή σίγουρα είναι αρκετά καλά αν αναλογιστεί κανείς τη φύση των δεδομένων που έχουμε χρησιμοποιήσει στα πειράματα μας. Γνωρίζουμε από την αρχή ότι δεν περιέχουν τόσο καλή πληροφορία όση θα θέλαμε για να εκπαιδύσουμε τον υπέρ-ταξινομητή μας. Στη συνέχεια θα παρουσιάσουμε τα αποτελέσματα του ίδιου ταξινομητή για δεδομένα που αφορούν εκφράσεις προσώπου.

Δεδομένα εκφράσεων προσώπου

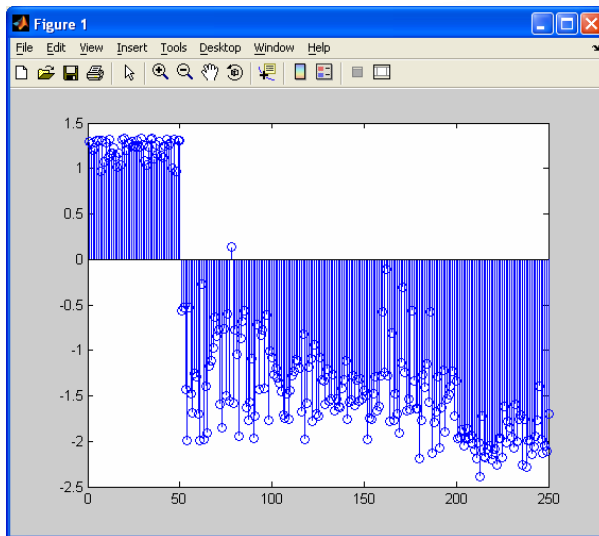
Πριν ξεκινήσουμε να παρουσιάσουμε τα αποτελέσματα του υπέρ-ταξινομητή μας μετά από την εκπαίδευση του με δεδομένα εκφράσεων προσώπου θα πρέπει να κάνουμε δύο σημαντικές παρατηρήσεις. Πρώτα απ' όλα τα αποτελέσματα όπως θα διαπιστώσουμε είναι πολύ καλύτερα από αυτά που παρουσιάσαμε στα μουσικά δεδομένα. Θα μπορούσαμε να πούμε ότι αγγίζουν το τέλειο. Αυτό οφείλεται στο γεγονός ότι έχουν συλλεχθεί με πολύ μεγάλη προσοχή και ίσως τα αποτελέσματα να ξεφεύγουν από πραγματικότητα. Επίσης τα δεδομένα αυτά αφορούν 5 κλάσεις και όχι 10 όπως τα μουσικά και τέλος λόγω του πλήθους των στιγμιότυπων ανά κλάση (έχουμε 50 στιγμιότυπα ελέγχου ανά κλάση) δημιουργήσαμε 5 Folds και όχι 10. Τα Sub-Folds είναι ακριβώς στο ίδιο πλήθος.

Αφού κάναμε μερικές παρατηρήσεις στα δεδομένα μας είμαστε τώρα σε θέση να παρουσιάσουμε τα αποτελέσματά μας πάλι για τον ίδιο ταξινομητή. Όπως και προηγουμένως έτσι και εδώ δεν θα δώσουμε τα αποτελέσματα από όλα τα Folds αλλά από ένα. Αποφασίσαμε να παρουσιάσουμε το Fold 3. Έτσι λοιπόν στον παρακάτω πίνακα φαίνονται οι επιμέρους ταξινομητές και τα ποσοστά σωστής ταξινόμησης στο Fold 3:

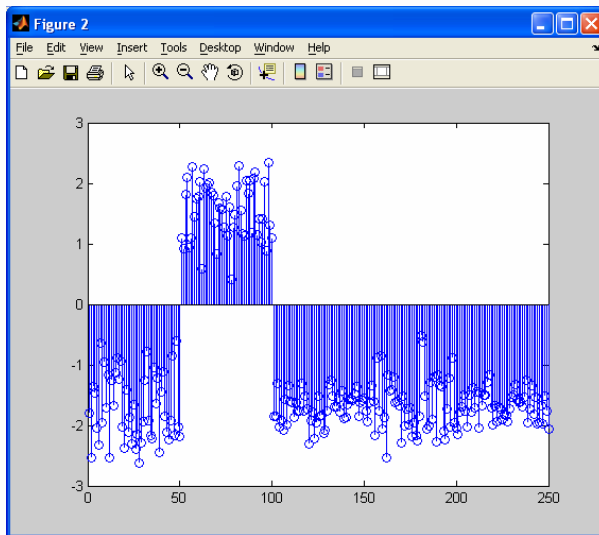
	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,996	0,996	0,992	0,996	0,996
Ταξινομητής2	1	1	1	1	1
Ταξινομητής3	1	1	1	1	1
Ταξινομητής4	0,932	0,928	0,944	0,936	0,936
Ταξινομητής5	0,932	0,932	0,944	0,928	0,928
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,996	0,996	0,996	0,996	0,996
Ταξινομητής2	1	1	1	1	1
Ταξινομητής3	1	1	1	1	1
Ταξινομητής4	0,932	0,936	0,932	0,928	0,944
Ταξινομητής5	0,932	0,928	0,928	0,936	0,92

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 3

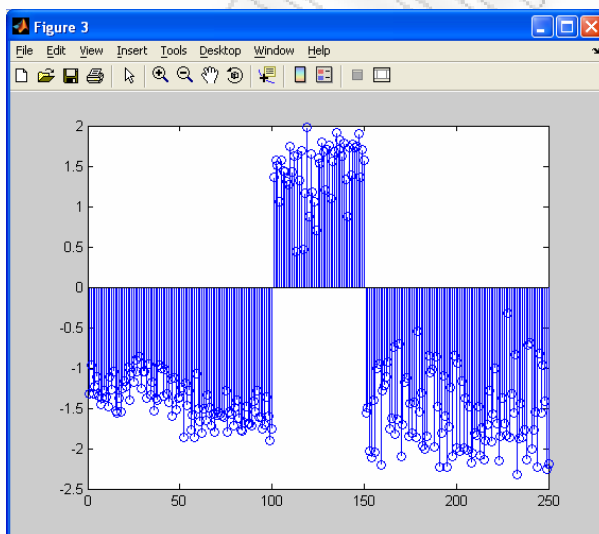
Είπαμε στην αρχή της ενότητας αυτής ότι έχουμε 5 κλάσεις οπότε για τον υπέρ-ταξινομητή μας δημιουργήσαμε και 5 επιμέρους ταξινομητές. Στη συνέχεια θα παρουσιάσω και τις τιμές της συνάρτησης απόφασης για τα αντίστοιχα δεδομένα ελέγχου του κάθε ταξινομητή που αφορούν το Fold 3 και τον Sub-Fold 2.



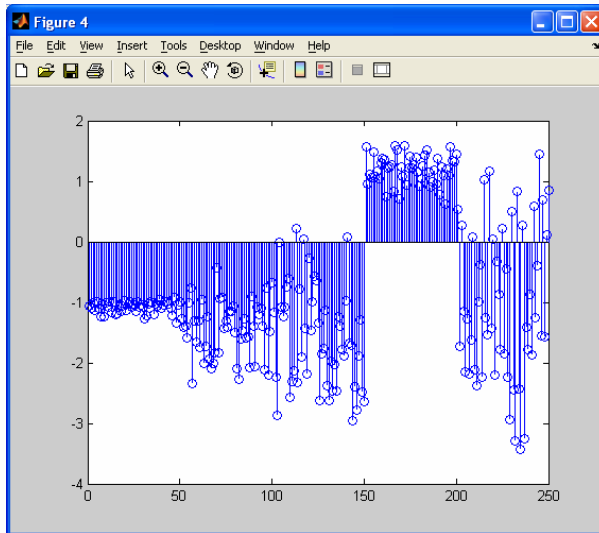
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 1



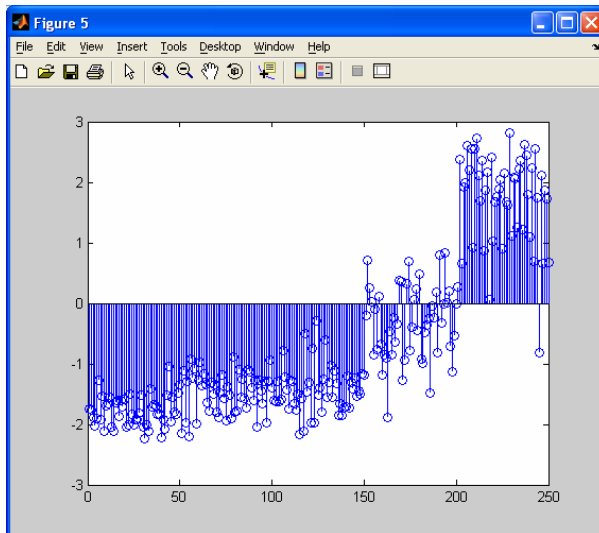
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 2



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 3



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 4



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 5

Παρατηρώντας τώρα τα αποτελέσματα τόσο στον πίνακα όσο και στα σχήματα που παρουσιάζουν τις τιμές της συνάρτησης απόφασης διαπιστώνουμε αυτό που λίγο πριν είπαμε. Οι ταξινομητές τα έχουν πάει πολύ καλά. Ειδικά οι ταξινομητές 2 και 3 έχουν 100% επιτυχία στην ταξινόμησή τους. Οι ταξινομητές 4 και 5 φαίνεται να εμφανίζουν κάποια μικρή αστάθεια κατά τον έλεγχο τους και κυρίως μπερδεύονται λίγο μεταξύ τους. Ενώ ο ταξινομητής 1 και αυτός φαίνεται να είναι σχεδόν άριστος. Όπως είπαμε κάθε κλάση έχει 50 στιγμιότυπα κατά τον έλεγχο των ταξινομητών γι' αυτό και έχουμε συνολικά 250 τιμές της συνάρτησης απόφασης.

Στη συνέχεια θα παρουσιάσουμε τα confusion matrixes για τους παραπάνω ταξινομητές που ελέγχονται στα αντίστοιχα δεδομένα.

Ταξινομητής	Positive Class	Negative Class
1	1	0
Classified as Positive	1	0
Classified as Negative	0,005	0,995

Ταξινομητής	Positive Class	Negative Class
2	1	0
Classified as Positive	1	0
Classified as Negative	0	1

Ταξινομητής 3	Positive Class	Negative Class
Classified as Positive	1	0
Classified as Negative	0	1

Ταξινομητής 4	Positive Class	Negative Class
Classified as Positive	1	0
Classified as Negative	0,09	0,91

Ταξινομητής 5	Positive Class	Negative Class
Classified as Positive	0,98	0,02
Classified as Negative	0,08	0,92

Πίνακες: Confusion matrixes των αντίστοιχων ταξινομητών

Όπως φαίνεται και πάλι από τα confusion matrixes όλοι οι ταξινομητές δουλεύουν πολύ καλά. Αυτός που υστερεί λίγο είναι ο ταξινομητής 5. Στον επόμενο πίνακα που ακολουθεί παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή για τα 5 διαφορετικά Folds που έχουμε δημιουργήσει. Τα αποτελέσματα αυτά είναι σε μορφή επί τοις εκατό.

Fold 1	95,2%
Fold 2	96,4%
Fold 3	97,6%
Fold 4	96,8%
Fold 5	94,4%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή

Δεν μπορούμε να ξεχωρίσουμε κανένα από τα 5 Folds αφού όλα έχουν πολύ καλά αποτελέσματα. Απλά μπορούμε να παρατηρήσουμε ότι στο Fold 3 ο υπέρ-ταξινομητής μας δείχνει να δουλεύει λίγο καλύτερα. Στη συνέχεια θα παρουσιάσουμε τα αποτελέσματα του υπέρ-ταξινομητή μας όταν αυτός υλοποιείται με νευρωνικό δίκτυο.

7.2.2 Ταξινομητής νευρωνικών δικτύων

Μουσικά δεδομένα 280 χαρακτηριστικών

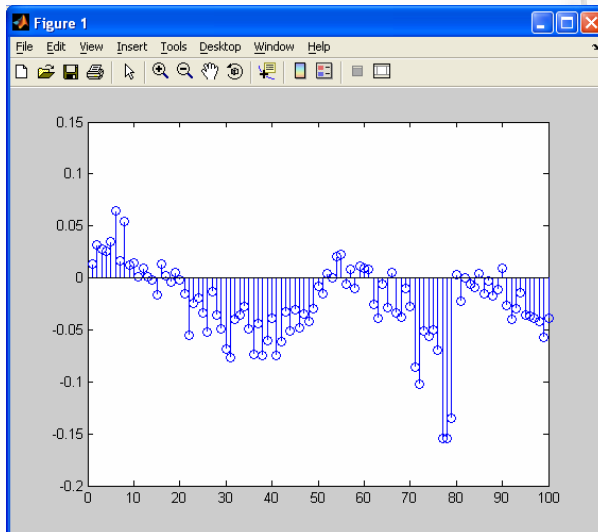
Στον παρακάτω πίνακα παρουσιάζονται τα αποτελέσματα σωστής ταξινόμησης των επιμέρους ταξινομητών του υπέρ-ταξινομητή μας που έχει δημιουργηθεί από τον perceptron για το Fold 4. Όπως φαίνεται από τα αποτελέσματα δείχνει να δουλεύει καλύτερα ο ταξινομητής με τα μουσικά δεδομένα των 280 χαρακτηριστικών.

	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	0,89	0,76	0,84	0,82	0,82
Ταξινομητής2	0,88	0,9	0,93	0,89	0,95
Ταξινομητής3	0,76	0,1	0,79	0,77	0,76
Ταξινομητής4	0,1	0,1	0,1	0,1	0,76
Ταξινομητής5	0,88	0,89	0,91	0,87	0,89
Ταξινομητής6	0,92	0,9	0,88	0,1	0,91
Ταξινομητής7	0,1	0,89	0,92	0,1	0,88
Ταξινομητής8	0,77	0,83	0,79	0,82	0,75
Ταξινομητής9	0,69	0,73	0,84	0,76	0,76
Ταξινομητής10	0,67	0,74	0,74	0,1	0,69
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10

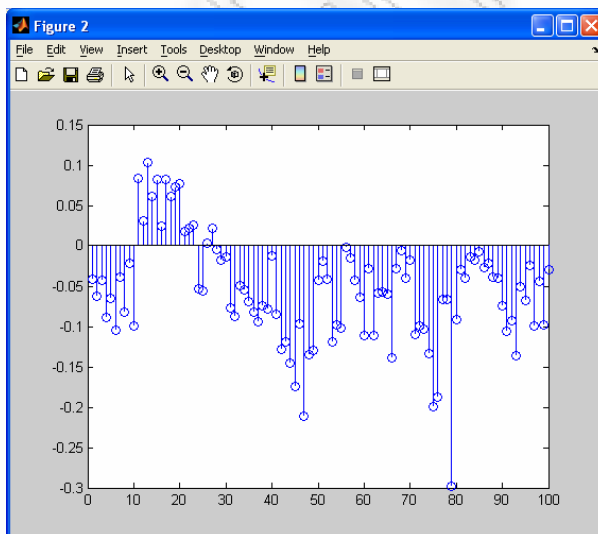
Ταξινομητής1	0,81	0,93	0,83	0,1	0,86
Ταξινομητής2	0,88	0,89	0,92	0,89	0,1
Ταξινομητής3	0,1	0,75	0,77	0,1	0,75
Ταξινομητής4	0,1	0,1	0,1	0,1	0,1
Ταξινομητής5	0,9	0,86	0,1	0,88	0,9
Ταξινομητής6	0,91	0,89	0,91	0,91	0,86
Ταξινομητής7	0,93	0,95	0,93	0,88	0,95
Ταξινομητής8	0,79	0,1	0,87	0,77	0,83
Ταξινομητής9	0,7	0,81	0,87	0,82	0,73
Ταξινομητής10	0,75	0,74	0,66	0,66	0,63

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 4

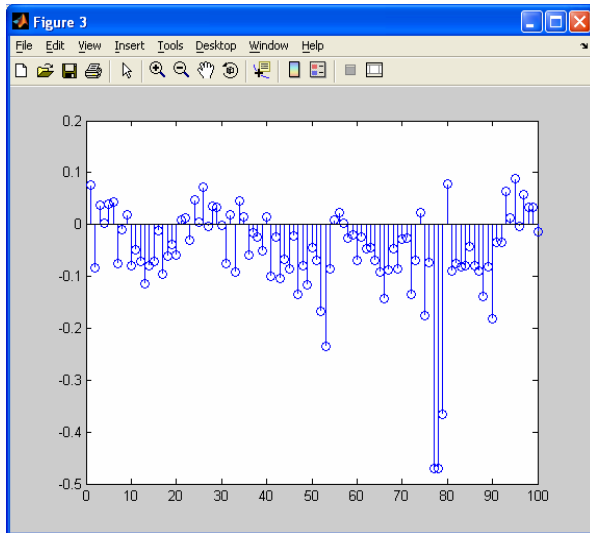
Παρατηρώντας λίγο καλύτερα τα αποτελέσματα διαπιστώνουμε ότι ο ταξινομητής 4 δίνει σχεδόν σε όλα τα Folds τιμές της τάξης του 10%. Αυτό συμβαίνει γιατί απλά λόγω της υψής των δεδομένων ο ταξινομητής δεν μπορεί να δώσει τιμές στη συνάρτηση απόφασης οπότε δίνει τιμές του τύπου 'NaN'. Αυτό όμως δεν αποτελεί σημαντικό πρόβλημα στην τελική απόφαση του υπέρ-ταξινομητή γιατί υπάρχουν Sub-Folds που ο ταξινομητής παίρνει τιμές για τη συνάρτηση απόφασης. Στα επόμενα σχήματα φαίνονται αυτές οι τιμές για το Sub-Fold 5 του Fold 4.



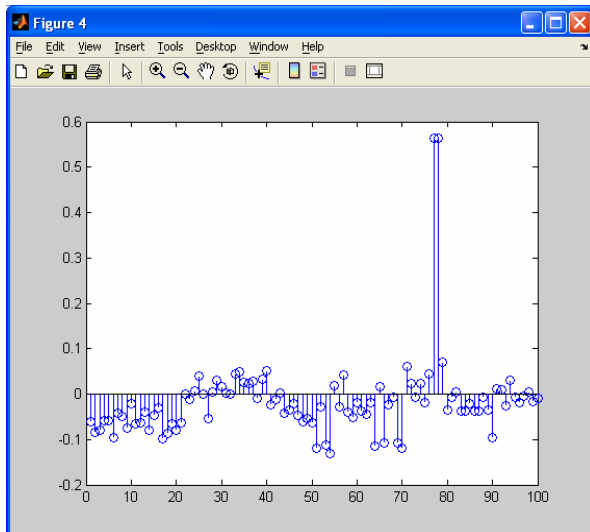
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 1



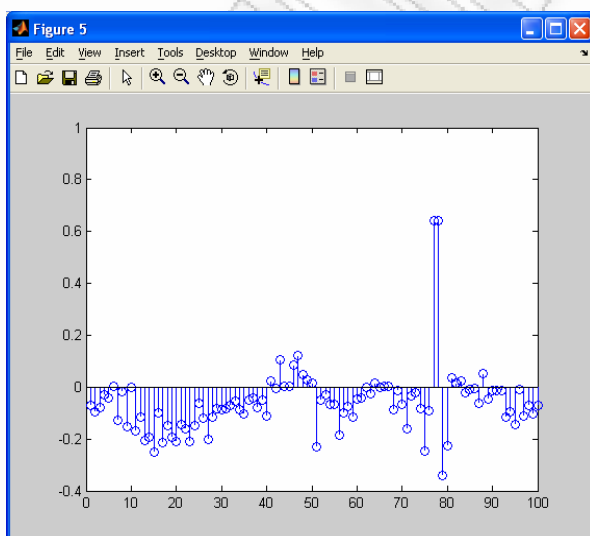
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 2



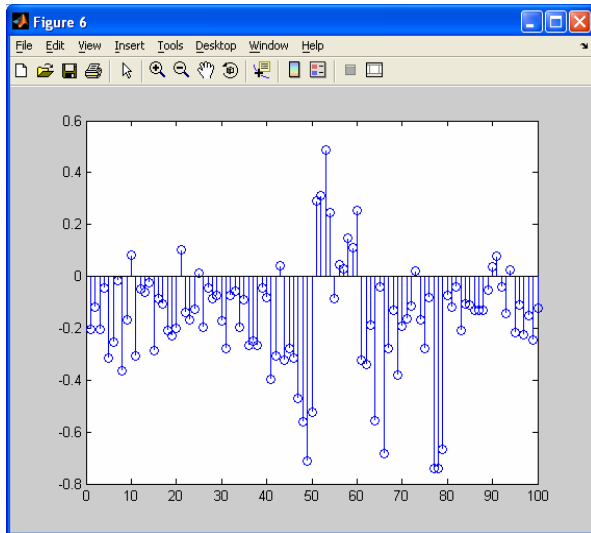
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 3



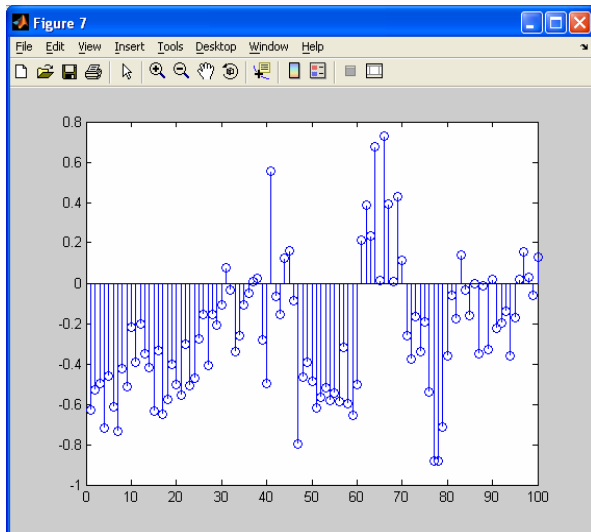
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 4



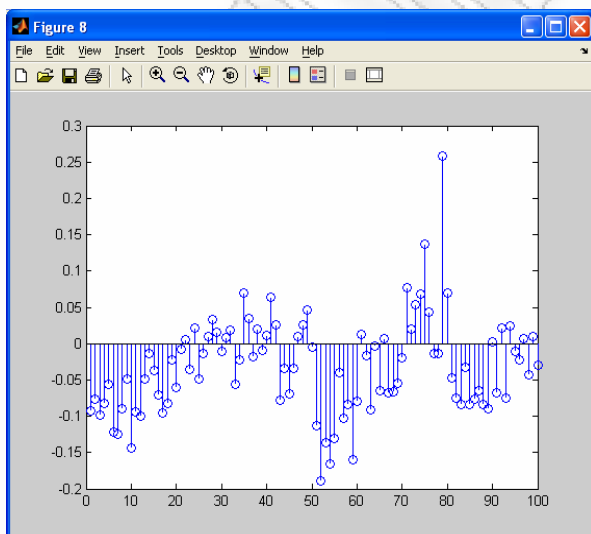
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 5



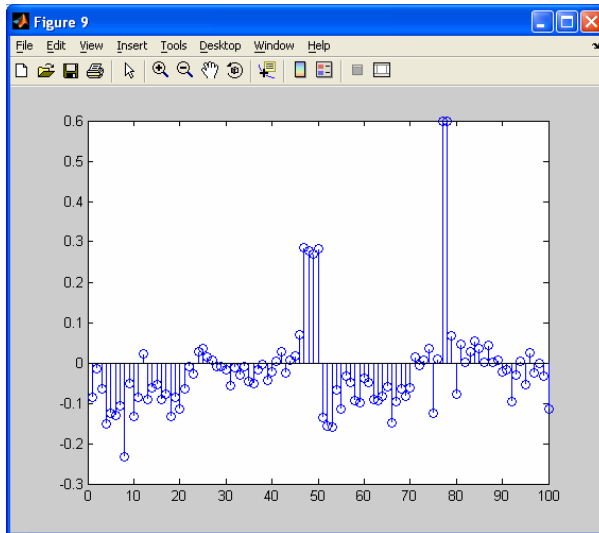
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 6



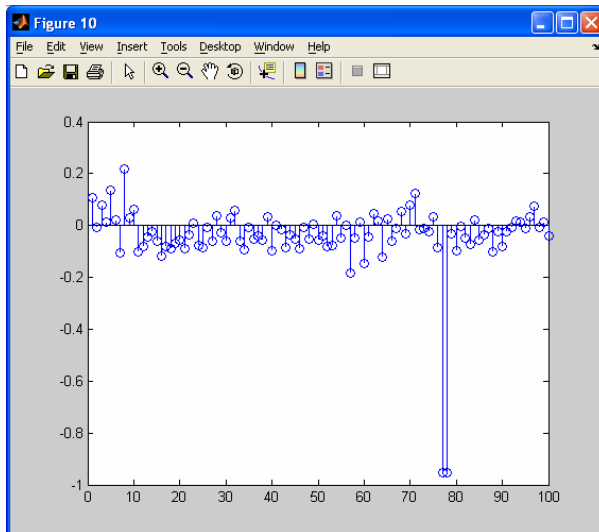
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 7



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 8



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 9



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 10

Σύμφωνα με τα παραπάνω σχήματα οι ταξινομητές 6 και 7 φαίνεται να λειτουργούν αρκετά καλά. Στον παρακάτω πίνακα παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή ο οποίος κατατάσσει τα μουσικά δεδομένα των 280 χαρακτηριστικών στις 10 κλάσεις.

Fold 1	51%
Fold 2	48%
Fold 3	61%
Fold 4	52%
Fold 5	48%
Fold 6	49%
Fold 7	48%
Fold 8	54%
Fold 9	56%
Fold 10	58%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή perceptron

Πρέπει να σημειωθεί ότι αυτά τα αποτελέσματα είναι χειρότερα από τα αντίστοιχα όταν χρησιμοποιήσαμε για την δημιουργία του υπέρ-ταξινομητή τον SVM ταξινομητή. Στη συνέχεια στον

επόμενο πίνακα παρουσιάζουμε τα τελικά αποτελέσματα του ταξινομητή νευρωνικού δικτύου back-propagation.

Fold 1	54%
Fold 2	64%
Fold 3	63%
Fold 4	57%
Fold 5	46%
Fold 6	54%
Fold 7	57%
Fold 8	54%
Fold 9	57%
Fold 10	60%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή back-propagation

Αυτό ο υπέρ-ταξινομητής λειτουργεί καλύτερα από τον ταξινομητή perceptron όπως είναι εμφανές από τα αποτελέσματα.

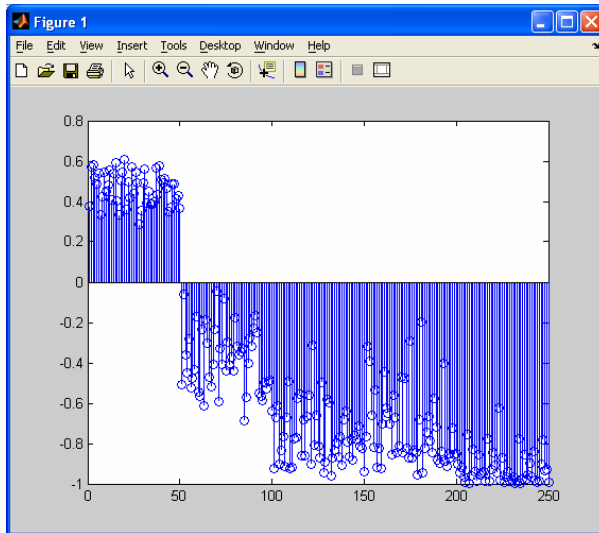
Δεδομένα εκφράσεων προσώπου

Στη συνέχεια θα παρουσιάσουμε τα αποτελέσματα των νευρωνικών ταξινομητών για τα δεδομένα εκφράσεων προσώπου. Έτσι λοιπόν στον παρακάτω πίνακα φαίνονται οι επιμέρους ταξινομητές και τα ποσοστά σωστής ταξινόμησης στο Fold 1:

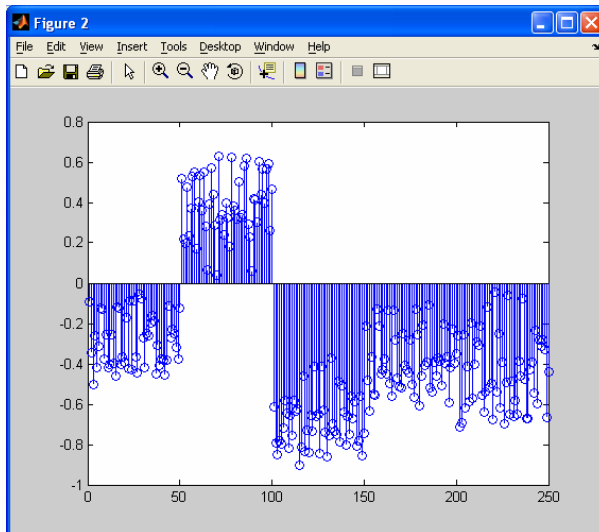
	Sub-Fold1	Sub-Fold2	Sub-Fold3	Sub-Fold4	Sub-Fold5
Ταξινομητής1	1	0,992	1	1	1
Ταξινομητής2	1	0,988	0,996	1	0,988
Ταξινομητής3	0,996	0,996	0,996	0,996	0,996
Ταξινομητής4	0,64	0,772	0,544	0,748	0,42
Ταξινομητής5	0,936	0,948	0,916	0,9	0,936
	Sub-Fold6	Sub-Fold7	Sub-Fold8	Sub-Fold9	Sub-Fold10
Ταξινομητής1	0,984	0,996	0,996	0,996	0,976
Ταξινομητής2	1	1	0,984	1	0,988
Ταξινομητής3	0,996	0,996	0,996	0,996	0,996
Ταξινομητής4	0,712	0,448	0,444	0,628	0,556
Ταξινομητής5	0,808	0,948	0,844	0,948	0,944

Πίνακας: Αποτελέσματα ταξινόμησης δεδομένων ελέγχου Fold 1

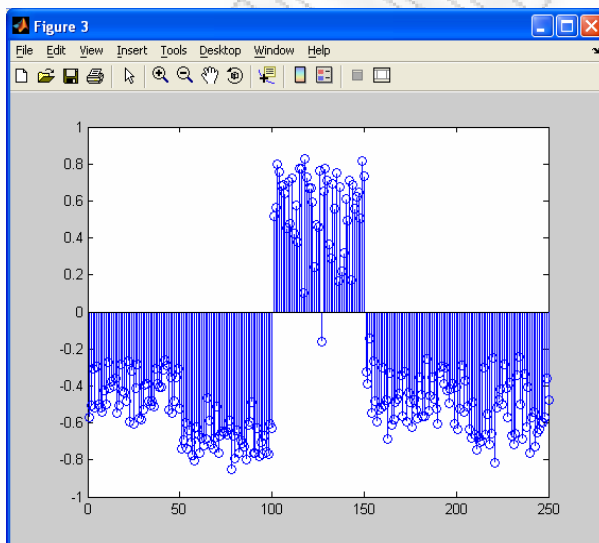
Και πάλι γίνεται εμφανές αυτό που είχαμε και με τον SVM υπέρ-ταξινομητή. Τα δεδομένα εκφράσεων προσώπου είναι πολύ καλά δομημένα με αποτέλεσμα να έχουμε πολύ καλά αποτελέσματα. Όμως δεν είναι καλύτερα από τα αντίστοιχα του SVM ταξινομητή. Επίσης έχουμε να παρατηρήσουμε ότι ο ταξινομητής 4 δεν μας δίνει τόσο καλά αποτελέσματα όσο οι υπόλοιποι. Αυτό θα γίνει πιο εμφανές όταν παρουσιάσουμε τις τιμές της συνάρτησης απόφασης του Sub-Fold 1 του Fold 1. Στα επόμενα σχήματα φαίνεται αυτό που μόλις περιγράψαμε.



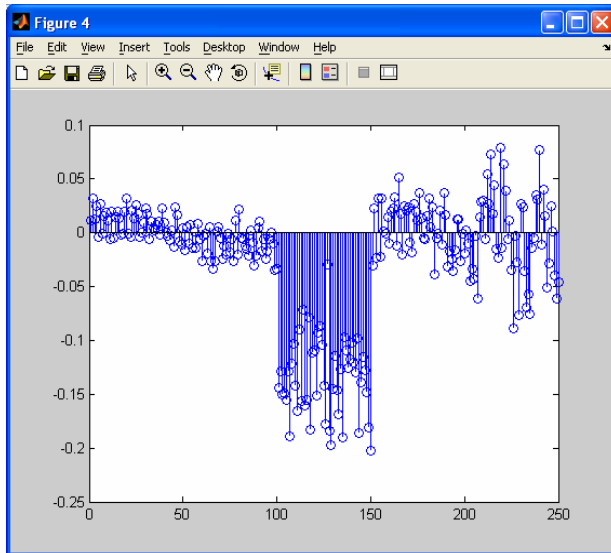
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 1



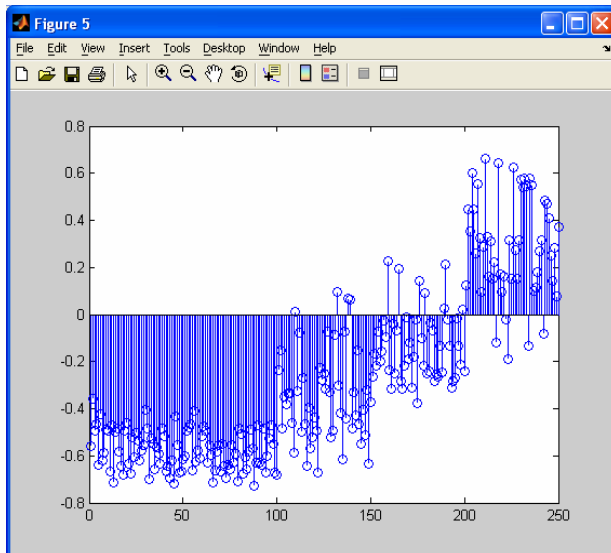
Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 2



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 3



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 4



Σχήμα: Τιμές της συνάρτησης απόφασης του ταξινομητή 5

Οι ταξινομητές 1 και 2 δουλεύουν τέλεια οι ταξινομητές 3 και 5 παρουσιάζουν κάποια μικρά προβλήματα στην ταξινόμησή τους και τέλος ο ταξινομητής 4 φαίνεται να μπερδεύεται αρκετά στα δεδομένα που του δίνονται. Όμως τα δεδομένα της κλάσης 3 τα κατατάσσει στα 'αρνητικά' με μεγάλη βεβαιότητα.

Στον επόμενο πίνακα παρουσιάζονται τα τελικά μας αποτελέσματα της σωστής ταξινόμησης του υπέρ-ταξινομητή μας ο οποίος έχει υλοποιηθεί με τη βοήθεια του νευρωνικού ταξινομητή perceptron.

Fold 1	82,8%
Fold 2	91,6%
Fold 3	91,2%
Fold 4	91,2%
Fold 5	85,6%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή perceptron

Όπως φαίνεται από τον παραπάνω πίνακα ο υπέρ-ταξινομητής με τα δεδομένα εκφράσεων προσώπου για ακόμη μία φορά τα πάει πολύ καλά. Πρέπει να παρατηρήσουμε όμως ότι δεν τα πάει καλύτερα από τον υπέρ-ταξινομητή βασισμένο στον SVM. Επίσης δεν τα πάει καλύτερα ούτε και από

τους υπέρ ταξινομητές που είναι βασισμένοι στον back-propagation και στον automatic νευρωνικό ταξινομητή του πακέτου PRTools4. Τα αποτελέσματα αυτά φαίνονται στους παρακάτω δύο πίνακες:

Fold 1	94,8%
Fold 2	95,6%
Fold 3	96%
Fold 4	96%
Fold 5	94%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή back-propagation

Fold 1	94,8%
Fold 2	95,6%
Fold 3	94,8%
Fold 4	96,8%
Fold 5	94,4%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή automatic neural network

Μία μικρή παρατήρηση που μπορούμε να κάνουμε στους δύο τελευταίους πίνακες είναι ότι ο automatic νευρωνικός ταξινομητής φαίνεται να δουλεύει λίγο καλύτερα από τον ταξινομητή back-propagation στα Folds 4 και 5. Στο Fold 3 δουλεύει λίγο χειρότερα ενώ στα Folds 1 και 2 μας δίνει ακριβώς τα ίδια αποτελέσματα. Συμπερασματικά λοιπόν δεν μπορούμε να πούμε ότι κάποιος δουλεύει καλύτερα από τον άλλον.

Στην τρίτη σχεδιαστική προσπάθειά εκπαιδεύουμε τους ταξινομητές μας και πάλι με ισοζυγισμένα δεδομένα όπως έχουμε περιγράψει αλλά προσπαθούμε να χρησιμοποιήσουμε βάρος για καλύτερα αποτελέσματα. Το όλο εγχείρημά μας μαζί με τα αποτελέσματα θα το περιγράψουμε στην επόμενη ενότητα.

7.3 Μέθοδος ‘One against all balanced with weight’

Μετά την εκπαίδευση των ταξινομητών που υλοποιήσαμε όπως περιγράψαμε σε προηγούμενο κεφάλαιο, τους ελέγχουμε σύμφωνα με τα δεδομένα ελέγχου που δημιουργήσαμε από το σύνολο των δεδομένων. Προκειμένου όμως να έχουμε καλύτερα αποτελέσματα στον τελικό μας υπέρ-ταξινομητή θα χρησιμοποιήσουμε ‘βάρος’ στον κάθε επιμέρους ταξινομητή πριν αυτός πάρει την τελική του απόφαση για την κατάταξη ενός στιγμιότυπου σε μία κλάση. Αυτό το βάρος στην πραγματικότητα δημιουργείται κατά τον προ-έλεγχο των επιμέρους ταξινομητών στα Sub-Folds.

Όπως έχουμε περιγράψει στο κεφάλαιο του σχεδιασμού του υπέρ-ταξινομητή τα δεδομένα μας τα ‘σπάμε’ σε Folds και σε Sub-Folds. Μέσα στα Sub-Folds έχουμε τόσο δεδομένα ελέγχου όσο και δεδομένα εκπαίδευσης. Αφού εκπαιδευτεί ο κάθε επιμέρους ταξινομητής στη συνέχεια ελέγχεται σε ένα πρώιμο στάδιο. Το ‘Rating’ που δημιουργείται για κάθε επιμέρους ταξινομητή χρησιμοποιείται στον τελικό υπέρ-ταξινομητή για την σωστή κατάταξη των στιγμιότυπων σε κλάσεις σύμφωνα με τα δεδομένα μας. Στη συνέχεια θα παρουσιάσουμε τα τελικά αποτελέσματα για κάθε υπέρ-ταξινομητή. Οι τιμές της συνάρτησης απόφασης δεν αλλάζουν απλά θα πολλαπλασιαστούν στον τελικό υπέρ-ταξινομητή με το βάρος για την τελική απόφαση κατάταξης. Οπότε δεν χρειάζεται να δώσουμε άλλα στοιχεία παρά μόνο τα τελικά αποτελέσματα κατάταξης των στιγμιότυπων στα διάφορα Folds που έχουμε κάθε φορά δημιουργήσει.

7.3.1 Ταξινομητής SVM

Μουσικά δεδομένα 30 χαρακτηριστικών

Στον παρακάτω πίνακα παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή βασισμένο στον SVM ταξινομητή για τα διάφορα Folds:

Fold 1	44%
Fold 2	49%
Fold 3	47%

Fold 4	50%
Fold 5	48%
Fold 6	42%
Fold 7	44%
Fold 8	45%
Fold 9	43%
Fold 10	39%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή SVM

Μελετώντας λίγο περισσότερο τα αποτελέσματα σε σχέση με αυτά του υπέρ-ταξινομητή χωρίς βάρος μπορούμε να παρατηρήσουμε ότι σε μερικά Folds έχουμε καλύτερα αποτελέσματα και σε μερικά άλλα χειρότερα. Η γενική τάση όμως είναι να βελτιώνονται τα αποτελέσματα.

Μουσικά δεδομένα 280 χαρακτηριστικών

Στον παρακάτω πίνακα παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή βασισμένο στον SVM ταξινομητή για τα διάφορα Folds που αφορούν τα μουσικά δεδομένα 280 χαρακτηριστικών:

Fold 1	59%
Fold 2	68%
Fold 3	74%
Fold 4	62%
Fold 5	61%
Fold 6	61%
Fold 7	56%
Fold 8	61%
Fold 9	62%
Fold 10	58%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή SVM

Το συμπέρασμα που βγάλαμε στα μουσικά δεδομένα των 30 χαρακτηριστικών γίνεται πιο εμφανές εδώ στα μουσικά δεδομένα των 280 χαρακτηριστικών. Σε όλα τα Folds έχουμε καλύτερα αποτελέσματα σε σχέση με τα αντίστοιχα του υπέρ-ταξινομητή χωρίς τη χρήση βάρους. Επίσης πρέπει να επισημάνουμε ότι παρατηρούνται και μεγάλες αποκλίσεις μεταξύ μερικών Folds. Για παράδειγμα στο Fold 3 ο υπέρ ταξινομητής μας βγάζει ένα ποσοστό περίπου στο 74% ενώ στο Fold 7 μόλις 56%. Η διαφορά είναι περίπου στις 20 ποσοστιαίες μονάδες αρκετά μεγάλη διαφορά. Αυτό μπορεί να δικαιολογηθεί μόνο στη φύση των δεδομένων. Στο Fold 3 να υπάρχουν στιγμιότυπα που να εκπαιδεύουν πολύ καλύτερα τον υπέρ-ταξινομητή μας απ' ότι στο Fold 7.

Δεδομένα εκφράσεων προσώπου

Στον παρακάτω πίνακα παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή βασισμένο στον SVM ταξινομητή για τα διάφορα Folds που αφορούν τα δεδομένα εκφράσεων προσώπου:

Fold 1	94,8%
Fold 2	96,4%
Fold 3	97,2%
Fold 4	96,8%
Fold 5	94,4%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή SVM

Όπως μπορούμε να δούμε τα αποτελέσματα αυτά είναι πολύ καλά όχι όμως καλύτερα από αυτά που εισπράτουμε από τον υπέρ-ταξινομητή χωρίς τη χρήση βάρους. Είναι σχεδόν παρεμφερή. Αυτό μπορεί να εξηγηθεί από το γεγονός ότι τα δεδομένα των εκφράσεων προσώπου είναι τόσο καλά δομημένα που δεν μπορεί η χρήση βάρους να βελτιώσει τα αποτελέσματα του υπέρ-ταξινομητή ακόμα περισσότερα αφού αυτά είναι ήδη πολύ καλά.

7.3.2 Ταξινομητής νευρωνικών δικτύων

Μουσικά δεδομένα 280 χαρακτηριστικών

Στον παρακάτω πίνακα παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή βασισμένο στον perceptron ταξινομητή για τα διάφορα Folds που αφορούν τα μουσικά δεδομένα των 280 χαρακτηριστικών:

Fold 1	50%
Fold 2	49%
Fold 3	59%
Fold 4	52%
Fold 5	49%
Fold 6	48%
Fold 7	51%
Fold 8	49%
Fold 9	59%
Fold 10	59%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή perceptron

Όπως φαίνεται από τον παραπάνω πίνακα τα αποτελέσματα είναι καλύτερα στα περισσότερα Folds σε σχέση με αυτά που μας δίνει ο υπέρ-ταξινομητής χωρίς τη χρήση βάρους αλλά σίγουρα όχι καλύτερα από αυτά που μας δίνει ο αντίστοιχος SVM υπέρ-ταξινομητής.

Στον επόμενο πίνακα παρουσιάζονται τα αποτελέσματα του υπέρ-ταξινομητή ο οποίος είναι στηριγμένος νευρωνικό ταξινομητή back-propagation:

Fold 1	59%
Fold 2	68%
Fold 3	70%
Fold 4	60%
Fold 5	48%
Fold 6	57%
Fold 7	59%
Fold 8	65%
Fold 9	64%
Fold 10	50%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή back-propagation

Τα αποτελέσματα εδώ είναι πολύ καλύτερα από αυτά του υπέρ-ταξινομητή υλοποιημένου με βάση τον υπέρ-ταξινομητή perceptron. Επίσης είναι αρκετά καλύτερα σε σχέση με τον υπέρ-ταξινομητή χωρίς τη χρήση βάρους ενώ εξακολουθούν να είναι λίγο χειρότερα από αυτά του υπέρ-ταξινομητή υλοποιημένου με βάση τον ταξινομητή SVM.

Δεδομένα εκφράσεων προσώπου

Στον παρακάτω πίνακα παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή βασισμένο στον ταξινομητή perceptron για τα διάφορα Folds που αφορούν τα δεδομένα εκφράσεων προσώπου:

Fold 1	82,4%
Fold 2	90,8%
Fold 3	90,8%
Fold 4	90,8%
Fold 5	84,4%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή perceptron

Σε όλα τα Folds ο ταξινομητής perceptron με τη χρήση βάρους παρουσιάζεται να μας δίνει χειρότερα αποτελέσματα σε σχέση με αυτά χωρίς τη χρήση βάρους. Επίσης ο συγκεκριμένος υπέρ-ταξινομητής φαίνεται να δίνει τα χειρότερα αποτελέσματα για τα αντίστοιχα δεδομένα σε σχέση με όλους τους υπόλοιπους υπέρ-ταξινομητές.

Στους επόμενους δύο πίνακες παρουσιάζονται τα αποτελέσματα των υπέρ-ταξινομητών οι οποίοι είναι στηριγμένοι στο νευρωνικό ταξινομητή back-propagation και automatic neural network ταξινομητή του πακέτου PRTools4 αντίστοιχα για τα δεδομένα εκφράσεων προσώπου.

Fold 1	95,2%
Fold 2	95,6%
Fold 3	95,6%
Fold 4	96%
Fold 5	94,4%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή back-propagation

Fold 1	95,2%
Fold 2	95,6%
Fold 3	94,4%
Fold 4	95,6%
Fold 5	97,2%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή automatic neural network

Όπως μπορούμε να δούμε τα αποτελέσματα αυτά είναι λίγο καλύτερα από αυτά που εισπράττουμε από τον υπέρ-ταξινομητή χωρίς τη χρήση βάρους. Είναι όμως αρκετά καλύτερα από αυτά που μας δίνει το υπέρ-ταξινομητής ο οποίος είναι υλοποιημένος με βάση το νευρωνικό ταξινομητή perceptron.

7.4 Μέθοδος ‘One against all balanced with double passing’

Στην τελευταία σχεδιαστική προσπάθεια του υπέρ-ταξινομητή εκπαιδεύουμε τους ταξινομητές μας και πάλι με ισοζυγισμένα δεδομένα όπως έχουμε περιγράψει και στις δύο προηγούμενες μεθόδους αλλά τώρα προσπαθούμε να εκμεταλλευτούμε τις τιμές της συνάρτησης απόφασης. Σκοπός μας είναι όπως έχουμε περιγράψει και στο κεφάλαιο του σχεδιασμού να διαπιστώσουμε μήπως οι τιμές της συνάρτησης απόφασης κρύβουν επιπλέον πληροφορία την οποία θα μπορούσαμε εμείς να εκμεταλλευτούμε για να βγάλουμε ακόμα καλύτερα αποτελέσματα στον υπέρ-ταξινομητή μας.

Έτσι λοιπόν μετά την εκπαίδευση των ταξινομητών που υλοποιήσαμε όπως περιγράψαμε και στη μέθοδο ‘one against all balanced’ όπως και στη μέθοδο ‘one against all balanced with weight’ ελέγχουμε τους επιμέρους ταξινομητές σε ένα πρώιμο στάδιο με δεδομένα ελέγχου τα οποία βρίσκονται στα Sub-Folds. Οπότε με τον προ-έλεγχο δημιουργούνται τιμές συνάρτησης απόφασης για κάθε ταξινομητή της αντίστοιχης κλάσης τις οποίες θα χρησιμοποιήσουμε σε δεύτερο στάδιο για την εκπαίδευση του τελικού υπέρ-ταξινομητή. Δηλαδή ο υπέρ-ταξινομητής μας εκπαιδεύεται δύο φορές, μία με τα αρχικά στιγμιότυπα των κλάσεων και μετά σε δεύτερη φάση με τις τιμές της συνάρτησης απόφασης που έχουν μετατραπεί σε κλάσεις και χαρακτηριστικά όπως περιγράψαμε σε προηγούμενο κεφάλαιο. Στη συνέχεια θα παρουσιάσουμε τα τελικά αποτελέσματα αυτής της υλοποίησης.

7.4.1 Ταξινομητής SVM

Μουσικά δεδομένα 30 χαρακτηριστικών

Στον παρακάτω πίνακα παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή βασισμένο στον SVM ταξινομητή για τα διάφορα Folds:

Fold 1	40,3%
Fold 2	44,6%
Fold 3	42%
Fold 4	46,4%
Fold 5	43%

Fold 6	41,4%
Fold 7	42,8%
Fold 8	45,2%
Fold 9	38,1%
Fold 10	42%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινόμητη SVM

Παρατηρώντας λίγο περισσότερο τα αποτελέσματα σε σχέση με αυτά του υπέρ-ταξινόμητη χωρίς βάρος και με αυτά με βάρος μπορούμε να παρατηρήσουμε ότι σε μερικά Folds έχουμε καλύτερα αποτελέσματα και σε μερικά άλλα χειρότερα.

Μουσικά δεδομένα 280 χαρακτηριστικών

Στον παρακάτω πίνακα παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινόμητη βασισμένο στον SVM ταξινόμητη για τα διάφορα Folds που αφορούν τα μουσικά δεδομένα 280 χαρακτηριστικών:

Fold 1	50,7%
Fold 2	68,2%
Fold 3	74,2%
Fold 4	63,7%
Fold 5	61,9%
Fold 6	58,5%
Fold 7	56,6%
Fold 8	60,6%
Fold 9	62,4%
Fold 10	60,8%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινόμητη SVM

Μελετώντας τα αποτελέσματα λίγο πιο προσεκτικά είναι ότι το συμπέρασμα που βγάλαμε στα μουσικά δεδομένα των 30 χαρακτηριστικών γίνεται πιο εμφανές εδώ στα μουσικά δεδομένα των 280 χαρακτηριστικών. Στα περισσότερα Folds έχουμε καλύτερα αποτελέσματα σε σχέση με τα αντίστοιχα του υπέρ-ταξινόμητη με τη χρήση βάρους και σίγουρα του υπέρ-ταξινόμητη χωρίς της χρήση βάρους. Πρέπει να επισημάνουμε όμως ότι οι διαφορές αυτές δεν είναι τόσο μεγάλες όσες αυτές μεταξύ του υπέρ-ταξινόμητη 'one against all balanced' και του υπέρ-ταξινόμητη 'one against all balanced with weight'. Η γενική τάση όμως είναι να έχουμε καλύτερα αποτελέσματα.

Δεδομένα εκφράσεων προσώπου

Στον παρακάτω πίνακα παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινόμητη βασισμένο στον SVM ταξινόμητη για τα διάφορα Folds που αφορούν τα δεδομένα εκφράσεων προσώπου:

Fold 1	95,52%
Fold 2	96,44%
Fold 3	96,68%
Fold 4	97,04%
Fold 5	95,12%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινόμητη SVM

Όπως μπορούμε να δούμε τα αποτελέσματα αυτά είναι πολύ καλά και λίγο καλύτερα από αυτά που εισπράτουμε από τον αντίστοιχο υπέρ-ταξινόμητη τόσο με τη χρήση βάρους όσο και χωρίς τη χρήση βάρους. Η γενική τάση είναι η ίδια με αυτήν που διαπιστώσαμε και με τα μουσικά δεδομένα των 280 χαρακτηριστικών. Η σχεδίαση αυτού του υπέρ-ταξινόμητη δουλεύει λίγο καλύτερα από τις υπόλοιπες. Αν και είναι λίγο πρόιμο να βγάλουμε συμπεράσματα αφού δεν έχουμε παραθέσει ακόμα όλα τα αποτελέσματα μπορούμε να πούμε ότι είναι μέχρι στιγμής ο καλύτερος.

7.4.2 Ταξινομητής νευρωνικών δικτύων

Πριν ξεκινήσουμε την παρουσίαση των αποτελεσμάτων που αφορούν τους υπέρ-ταξινομητές οι οποίοι έχουν υλοποιηθεί βασισμένοι σε ταξινομητές νευρωνικών δικτύων θα πρέπει να πούμε ότι λόγω της φύσης των μουσικών δεδομένων και κυρίως αυτών των μουσικών δεδομένων που αποτελούνται τόσο από 30 όσο και από 280 χαρακτηριστικά ανά στιγμιότυπο δεν κατέστη δυνατό κατά την εκπαίδευση κάποιων επιμέρους ταξινομητών να δημιουργηθούν τιμές στη συνάρτηση απόφασης. Το προγραμματιστικό εργαλείο του Matlab μας έδινε τιμές του τύπου 'NaN' οπότε δεν ήμασταν σε θέση να εκπαιδεύσουμε σε δεύτερη φάση τον υπέρ-ταξινομητή μας οπότε ούτε και να τον υλοποιήσουμε. Για όποια δεδομένα δεν καταφέραμε να υλοποιήσουμε τους αντίστοιχους υπέρ ταξινομητές τα αποτελέσματα παραλείπονται.

Δεδομένα εκφράσεων προσώπου

Τα τελευταία αποτελέσματα που θα παρουσιάσουμε αφορούν τα δεδομένα εκφράσεων προσώπου. Στον παρακάτω πίνακα παρουσιάζονται τα τελικά αποτελέσματα του υπέρ-ταξινομητή βασισμένο στον ταξινομητή perceptron για τα διάφορα Folds που αφορούν τα δεδομένα εκφράσεων προσώπου:

Fold 1	90,56%
Fold 2	91,16%
Fold 3	85,96%
Fold 4	90,8%
Fold 5	87,4%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή perceptron

Στα περισσότερα Folds ο ταξινομητής perceptron σχεδιασμένος με τη μέθοδο 'one against all balanced with double passing' παρουσιάζεται να μας δίνει καλύτερα αποτελέσματα σε σχέση με αυτά του ταξινομητή με τη χρήση βάρους, ενώ με τη μη χρήση βάρους εμφανίζεται να έχει περίπου τα ίδια αποτελέσματα. Η διαπίστωση που μπορεί να γίνει είναι ότι επειδή είναι τόσο καλά τα δεδομένα των εκφράσεων προσώπου οι τιμές της συνάρτησης απόφασης τελικά δεν κρύβουν κάποια επιπλέον πληροφορία που να μπορούμε να εκμεταλλευτούμε ώστε να μπορέσουμε να αποκομίσουμε καλύτερα αποτελέσματα κατά τον έλεγχο του τελικού μας υπέρ-ταξινομητή.

Στους επόμενους δύο πίνακες παρουσιάζονται τα αποτελέσματα των υπέρ-ταξινομητών οι οποίοι είναι στηριγμένοι στο νευρωνικό ταξινομητή back-propagation και στον automatic neural network ταξινομητή του πακέτου PRTools4 αντίστοιχα για τα δεδομένα εκφράσεων προσώπου.

Fold 1	94,2%
Fold 2	95,12%
Fold 3	95,52%
Fold 4	93,68%
Fold 5	93,88%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή back-propagation

Fold 1	93,6%
Fold 2	93,08%
Fold 3	93,8%
Fold 4	79,28%
Fold 5	79,44%

Πίνακας: Αποτελέσματα τελικού υπέρ-ταξινομητή automatic neural network

Όπως μπορούμε να δούμε τα αποτελέσματα αυτά είναι λίγο χειρότερα από αυτά που εισπράττουμε από τον υπέρ-ταξινομητή τόσο με τη χρήση όσο και χωρίς τη χρήση βάρους. Αυτό πιθανότατα γίνεται γιατί ο τελικός υπέρ-ταξινομητής σχεδιασμένος με τη μέθοδο 'one against all balanced with double passing' δεν εκμεταλλεύεται τις τιμές της συνάρτησης απόφασης και οι δύο ταξινομητές νευρωνικών δικτύων στους οποίους είναι στηριγμένος μπερδεύονται και δεν εκμεταλλεύονται πληροφορία που τυχόν μπορεί να κρύβεται πίσω από αυτές τις τιμές.

8 Επίλογος

8.1 Συμπεράσματα

Κύριο αντικείμενο της εργασίας αποτέλεσε ο σχεδιασμός ενός υπέρ-ταξινομητή ο οποίος θα είναι σε θέση να μπορεί να ταξινομήσει περισσότερες από δύο κλάσεις. Ο σχεδιασμός αυτός στηρίχθηκε στη μεθοδολογία ‘one against all’ όπου εξάχθηκαν και υβριδικά μοντέλα αυτής της μεθοδολογίας. Η υλοποίηση του ταξινομητή έγινε με τη βοήθεια του αλγορίθμου μηχανικής μάθησης SVM καθώς και με ταξινομητές νευρωνικών δικτύων όπως το perceptron και το back-propagation.

Στο κεφάλαιο 2 παρουσιάστηκαν οι βασικότερες έννοιες των υποπεριοχών της μηχανικής μάθησης από τις οποίες αντλήθηκε υλικό για την εκπόνηση της εργασίας, και τα κεφάλαια 3 και 4 περιγράφηκαν οι αλγόριθμοι που χρησιμοποιήθηκαν τόσο για τον SVM όσο και τους νευρωνικούς ταξινομητές. Αυτοί οι δυαδικοί ταξινομητές ήταν η βάση για τη δημιουργία του υπέρ ταξινομητή μας.

Στο επόμενο κεφάλαιο περιγράφηκε αναλυτικά η μεθοδολογία που χρησιμοποιήθηκε για το σχεδιασμό του υπέρ-ταξινομητή. Μέσα από τη μέθοδο ‘one against all’ δημιουργήσαμε ένα καινούριο υβριδικό μοντέλο ταξινόμησης το οποίο ονομάσαμε ‘one against all balanced with double passing’.

Η αποτελεσματικότητα του παραπάνω σχεδιασμού αξιολογήθηκε πειραματικά, πάνω σε τρία προβλήματα ταξινόμησης δεδομένων. Το ένα πρόβλημα αποτελούταν από δεδομένα διαφορετικών εκφράσεων προσώπου και το άλλο από μουσικά αρχεία διαφορετικών ειδών μουσικής. Για το δεύτερο πρόβλημα είχαμε δύο είδη δεδομένων. Στα μεν πρώτα το κάθε στιγμιότυπο αποτελούταν από 30 χαρακτηριστικά ενώ στα δε δεύτερα είχαμε 280 χαρακτηριστικά. Τα αποτελέσματα της λεπτομερούς ανάλυσης των πειραματικών μετρήσεων που διεξήχθησαν παρουσιάστηκαν στο αντίστοιχο κεφάλαιο της πειραματικής αξιολόγησης. Τελικά διαπιστώθηκε ότι το υβριδικό μοντέλο που δημιουργήσαμε έχει την τάση να βελτιώνει τα τελικά αποτελέσματα του υπέρ-ταξινομητή σε σχέση με αυτά του υπέρ-ταξινομητή σχεδιασμένου με βάση τη μέθοδο ‘one against all’. Επίσης διαπιστώθηκε ότι δεν αντιδρούν όλα τα δεδομένα με τον ίδιο τρόπο όταν τους εφαρμόζεται ο υπέρ-ταξινομητής ενώ δεδομένα πολύς καλής ποιότητας φαίνεται να μην βελτιώνουν τα αποτελέσματά τους περισσότερο με τη χρήση του υβριδικού μας μοντέλου.

8.2 Μελλοντικές κατευθύνσεις

Από τα αποτελέσματα της μελέτης αυτής πάνω στο πρόβλημα της ταξινόμησης δεδομένων που ανήκουν σε περισσότερες από δύο κλάσεις προκύπτει ότι για να επιτευχθεί αξιόλογη βελτίωση της απόδοσης των υπέρ-ταξινομητών θα πρέπει στο ανώτερο επίπεδο ταξινόμησης να ενσωματωθεί η γνώση που έχει προκύψει από την ταξινομητική συμπεριφορά των επιμέρους αλγορίθμων μηχανικής μάθησης. Στην εργασία μας κατά την εκπαίδευση του κάθε επιμέρους ταξινομητή επιλέγουμε τόσο τα θετικά όσο και τα αρνητικά δεδομένα με μηχανιστική μέθοδο χωρίς να ενδιαφερόμαστε καθόλου αν αυτά τα δεδομένα δημιουργούν τον καλύτερο δυνατό ταξινομητή. Δηλαδή η επιλογή των προτύπων από την αρνητική κλάση πραγματοποιείται με βάση την μέθοδο της τυχαίας δειγματοληψίας.

Η μελλοντική έρευνα θα μπορούσε να επικεντρωθεί στην αναζήτηση μεθόδων που θα καταφέρνουν να επιλέγουν από τα συνολικά δεδομένα μίας κλάσης εκείνα τα στιγμιότυπα τα οποία θα εκπαιδεύσουν καλύτερα τον ταξινομητή. Με άλλα λόγια, η επιλογή των δεδομένων από την αρνητική κλάση να υλοποιείται με την βοήθεια μεθόδων οι οποίες λαμβάνουν υπόψη την κατανομή των δεδομένων μέσα στην συγκεκριμένη κλάση. Επομένως, θα ήταν μελλοντικά εφικτό να πραγματοποιηθεί μια συγκριτική μελέτη μεταξύ αυτών των προσεγγίσεων. Θα είχε επίσης ενδιαφέρον, η χρήση μεθόδων για την σύνθεση των αποτελεσμάτων των επιμέρους ταξινομητών που χρησιμοποιούν στοιχεία της θεωρίας παιγνίων.

9 Αναφορές

1 Jonathan Milgram Mohamed Cheriet Robert Sabourin: “One Against One” or “One Against All”: Which One is Better for Handwriting Recognition with SVMs? École de Technologie Supérieure, Montréal, Canada

2 Zhonghui Hu, Yuangu Li, Yunze Cai, and Xiaoming Xu: Method of Combining Multi-class SVMs Using Dempster-Shafer Theory and its Application. In: Proceedings of American Control Conference Portland, OR, USA (June 8-10, 2005)

- 3 Yi Liu, Yuan F. Zheng: One-Against-All Multi-Class SVM Classification Using Reliability Measures. In: Proceedings of International Joint Conference on Neural Networks, Montreal, Canada (July 31 - August 4, 2005)
- 4 Hai Zhao and Bao-Liang Lu: A General Procedure for Combining Binary Classifiers and Its Performance Analysis. L. Wang, K. Chen, and Y.S. Ong (Eds.): ICNC 2005, LNCS 3610, pp. 303–312, Springer-Verlag Berlin Heidelberg 2005
- 5 Arturo Montejó R´aez, Lu´ıs Alfonso Ure˜na L´opez, Ralf Steinberger: Adaptive Selection of Base Classifiers in One-Against-All Learning for Large Multi-labeled Collections. J. L. Vicedo et al. (Eds.): EsTAL 2004, LNAI 3230, pp. 1–12 Springer-Verlag Berlin Heidelberg 2004
- 6 Kai-Bo Duan, S. Sathiya Keerthi: Which Is the Best Multiclass SVM Method? An Empirical Study. N.C. Oza et al. (Eds.): MCS 2005, LNCS 3541, pp. 278–285, Springer-Verlag Berlin Heidelberg 2005
- 7 Alina Beygelzimer, John Langford, Bianca Zadrozny: Weighted One-Against-All. Copyright c 2005, American Association for Artificial Intelligence
- 8 Yunqiang Chen, Xiang Zhou, and Thomas S. Huang: ONE-CLASS SVM FOR LEARNING IN IMAGE RETRIEVAL. In Proc. IEEE Int'l Conf. on Image Processing Thessaloniki, Greece (2001)
- 9 Larry M. Manevitz, Malik Yousef: One-Class SVMs for Document Classification. Journal of Machine Learning Research 2 139-154 (2001)
- 10 Dionysios N. Sotiropoulos, Aristomenis S. Lampropoulos, George A. Tsihrintzis: MUSIPER: a system for modeling music similarity perception based on objective feature subset selection. Published online: 25 September 2007. Springer Science+Business Media B.V. 2007
- 11 Foote, J.: Content-based retrieval of music and audio. In: Proc. Storage and Retrieval for Image and Video Databases (SPIE), vol. 3229, pp. 138–147 (1997)
- 12 Tzanetakis, G.: Manipulation, analysis and retrieval systems for audio signals. Ph.D. thesis, Princeton University (2002)
- 13 Tzanetakis, G., Cook, P.: MARSYAS: a framework for audio analysis. Organised Sound 4(3), 169–175 (2000)
- 14 Tzanetakis, G., Cook, P.: Musical genre classification of audio signals. IEEE Transact. Speech Audio Process. 10(5), 293–302 (2002)
- 15 Aristomenis S. Lampropoulos, Ioanna-Ourania Stathopoulou, George A. Tsihrintzis: Comparative performance evaluation of classifiers for Facial Expression Recognition. E. Damiani et al. (Eds.): New Direct. in Intel. Interactive Multimedia Sys., SCI 226, pp. 253–263. Springer-Verlag Berlin Heidelberg 2009
- 16 Stathopoulou, I.O., Tsihrintzis, G.: Facial Expression Classification: Specifying Requirements for an Automated System. In: Proceedings of the 10th International Conference on Knowledge-Based & Intelligent Information & Engineering Systems, Bournemouth, United Kingdom (October 2006)
- 17 Stathopoulou, I.O., Tsihrintzis, G.: A neural network-based facial analysis system. In: Proceedings of the 5th International Workshop on Image Analysis for Multimedia Interactive Services, Lisboa, Portugal (April 2004)

- 18 Stathopoulou, I.O., Tsihrintzis, G.: An Improved Neural Network-Based Face Detection and Facial Expression Classification System. In: IEEE International Conference on Systems, Man, and Cybernetics, The Hague, Netherlands (October 2004)
- 19 Stathopoulou, I.O., Tsihrintzis, G.: Pre-processing and expression classification in low quality face images. In: Proceedings of 5th EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services (July 2005)
- 20 Stathopoulou, I.O., Tsihrintzis, G.: Evaluation of the Discrimination Power of Features Extracted from 2-D and 3-D Facial Images for Facial Expression Analysis. In: Proceedings of the 13th European Signal Processing Conference, Antalya, Turkey (September 2005)
- 21 Stathopoulou, I.O., Tsihrintzis, G.: Detection and Expression Classification Systems for Face Images (FADECS). In: Proceedings of the IEEE Workshop on Signal Processing Systems (SiPS 2005), Athens, Greece (November 2005)
- 22 Stathopoulou, I.O., Tsihrintzis, G.: An Accurate Method for eye detection and feature extraction in face color images. In: Proceedings of the 13th International Conference on Signals, Systems, and Image Processing, Budapest, Hungary (September 2006)
- 23 Stathopoulou, I.O., Tsihrintzis, G.: NEU-FACES: A Neural Network-based Face Image Analysis System. In: Proceedings of the 8th International Conference on Adaptive and Natural Computing Systems, Warsaw, Poland (April 2007)
- 24 J. Ma, Y. Zhao and S. Ahalt: OSU SVM Classifier Matlab Toolbox (version 3.00)
http://www.ece.osu.edu/~maj/osu_svm/
- 25 R.P.W. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. de Ridder, D.M.J. Tax, S. Verzakov: PRTools4 A Matlab Toolbox for Pattern Recognition Version 4.1. August 2007
- 26 Ludmila I. Kuncheva: Combining Pattern Classifiers Methods and Algorithms. Book: A JOHN WILEY & SONS, INC., PUBLICATION (2004)
- 27 Steve R. Gunn: Support Vector Machines for Classification and Regression. Technical Report: Faculty of Engineering, Science and Mathematics School of Electronics and Computer Science (1998)
- 28 Brian R. Hunt, Ronald L. Lipsman, Jonathan M. Rosenberg with Kevin R. Coombes, John E. Osborn, and Garrett J. Stuck: A Guide to MATLAB for Beginners and Experienced Users. Cambridge University Press (2001)