

Ατομική Διπλωματική Εργασία

**ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ
ΣΤΗΝ ΑΝΑΛΥΣΗ ΤΗΣ ΜΟΥΣΙΚΗΣ**

Ραφαέλλα Ιωάννου

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ



ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Μάιος 2013

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Υπολογιστική Νοημοσύνη στην ανάλυση της Μουσικής

Ραφαέλλα Ιωάννου

Επιβλέπων Καθηγητής

Χρίστος Ν. Σχίζας

Η Ατομική Διπλωματική Εργασία υποβλήθηκε προς μερική εκπλήρωση των
απαιτήσεων απόκτησης του πτυχίου Πληροφορικής του Τμήματος Πληροφορικής του
Πανεπιστημίου Κύπρου

Μάιος 2013

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέποντα Καθηγητή Χρίστο Ν. Σχίζα που με εμπιστεύθηκε για την εκπόνηση αυτής της διπλωματικής εργασίας, καθώς επίσης και για την υποστήριξη, την υπομονή, την κατανόηση και καθοδήγηση που μου παρείχε όλο αυτό το διάστημα.

Επίσης, θα ήθελα να ευχαριστήσω τον ερευνητή Αντρέα Νεοκλέους για τις ιδέες, τις συμβουλές και την καθοδήγηση που μου πρόσφερε με σκοπό να βελτιωθεί στο μέγιστο η εργασία αυτή.

Τέλος, ευχαριστώ την οικογένεια μου, τους φίλους μου και όσους ήταν στο πλάι μου όλο αυτό το διάστημα της εκπόνησης της εργασίας αυτής, τόσο για την υποστήριξη όσο και για τη βοήθεια που μου παρείχαν.

Περίληψη

Στόχος αυτής της διπλωματικής εργασίας είναι η μελέτη της υπολογιστικής νοημοσύνης και οι εφαρμογές της σε διάφορους τομείς και κυρίως στον τομέα της μουσικής. Συγκεκριμένα, στόχος είναι η υλοποίηση ενός συστήματος μάθησης που θα ασχολείται με την ανάλυση της Κυπριακής παραδοσιακής μουσικής και την μελέτη πιθανών ομοιοτήτων ή διαφορών με διάφορες μουσικές της Μεσογείου.

Για την υλοποίηση αυτή, χρησιμοποιήθηκε μια βάση δεδομένων με 127 μονοφωνικά τραγούδια, από τα οποία τα 37 ήταν Κυπριακά, τα 43 ήταν Τούρκικα και τα υπόλοιπα 47 ήταν Δυτικά. Σε κάθε τραγούδι εφαρμόστηκε μια συγκεκριμένη διαδικασία για την εξαγωγή κάποιων τονικών χαρακτηριστικών και χαρακτηριστικών χροιάς, τα οποία χρησιμοποιήθηκαν για τη δημιουργία διάφορων μοντέλων με διάφορες αρχιτεκτονικές που να μπορούν να «χαρακτηρίζουν» τη μουσική και να την τοποθετούν σε κατηγορίες. Δημιουργήθηκαν μοντέλα με επιβλεπόμενη μάθηση με τη χρήση νευρωνικών δικτύων Multilayer Perceptron ενός κρυφού επιπέδου, K κοντινότερου γείτονα και μηχανές διανυσμάτων (support vector machines) με πυρήνες 1,2,3.

Τελικά το σύστημα έδωσε πολύ καλά ποσοστά αποτελεσμάτων. Τα χαρακτηριστικά που χρησιμοποιήθηκαν, μπόρεσαν να διακρίνουν τη Δυτική μουσική από την Τούρκικη και να κατηγοριοποιήσουν σωστά 92% της Δυτικής μουσικής και 100% της Τούρκικης μουσικής.

Από τα αποτελέσματα αυτά, διάφοροι κλάδοι της επιστήμης όπως η μουσικολογία, εθνομουσικολογία, επιστήμη ηλεκτρονικών υπολογιστών θα επωφεληθούν και θα κατανοηθούν καλύτερα τα χαρακτηριστικά της Κυπριακής παραδοσιακής μουσικής όπου ίσως να προσδιορίσουν και τις πηγές που την επηρεάζουν.

Περιεχόμενα

Κεφάλαιο 1	Εισαγωγή.....	1
	1.1 Νοημοσύνη	1
	1.2 Υπολογιστική Νοημοσύνη	2
Κεφάλαιο 2	Τεχνικές Υπολογιστικής Νοημοσύνης.....	5
	2.1 K-Κοντινότερος γείτονας (K-nn)	5
	2.2 Νευρωνικά Δίκτυα	6
	2.2.1 Βιολογικοί Νευρώνες	7
	2.2.2 Τεχνητοί Νευρώνες	8
	2.2.3 Τεχνητά Νευρωνικά Δίκτυα	8
	2.2.4 Ιδιότητες Νευρωνικών Δικτύων	13
	2.3 Μηχανές Υποστήριξης διανυσμάτων (SVM)	13
	2.4 Εξελικτικά Συστήματα	15
	2.4.1 Θεωρία της Εξέλιξης	15
	2.4.2 Γενετικοί Αλγόριθμοι	16
	2.4.3 Γενετικός προγραμματισμός	18
	2.4.4 Στρατηγικές εξέλιξης	20
	2.4.5 Εξελικτικός προγραμματισμός	20
	2.5 Ασαφής Λογική	21
Κεφάλαιο 3	Προηγούμενες Μελέτες.....	24
	3.1 Εργασίες στο Πανεπιστήμιο Κύπρου	24
	3.2 Εργασίες σε Διεθνές επίπεδο	26
Κεφάλαιο 4	Ανάλυση θέματος και εξήγηση.....	30
	4.1 Ορισμός θέματος / προβλήματος	30
	4.2 Ερευνητική ερώτηση	31
	4.3 Ανάλυση	31
	4.4 Παραδοσιακές μέθοδοι αντιμετώπισης	39
	4.5 Σύγκριση και συζήτηση μεθόδων	39

Κεφάλαιο 5	Πρόταση εργασίας.....	41
	5.1 Προεπεξεργασία – Εξαγωγή Χαρακτηριστικών	42
	5.2 Υπολογιστική Νοημοσύνη	44
	5.3 Αποτελέσματα	45
Κεφάλαιο 6	Συμπεράσματα	54
	6.1 Γενικά συμπεράσματα	54
Κεφάλαιο 7	Θέματα μελλοντικής μελέτης	57
	7.1 Μελλοντική εργασία	57
Α ν α φ ο ρ έ ς		58
Π α ρ ά ρ τ η μ α Α.....		A-1

Κεφάλαιο 1

Εισαγωγή

1.1 Νοημοσύνη	1
1.2 Υπολογιστική Νοημοσύνη	2

1.1 Νοημοσύνη

Κατά τη διάρκεια του αιώνα αυτού, ο άνθρωπος ήρθε αντιμέτωπος με διάφορα δύσκολα προβλήματα σε διάφορους τομείς όπως στην ιατρική, οικολογία, οικονομία και άλλα. Ένα από τα πιο κύρια προβλήματα στην έρευνα ήταν η γρήγορη και σωστή επεξεργασία των δεδομένων, που πολλές φορές ήταν δύσκολο ακόμα και να καταγραφούν. Ο τεράστιος όγκος δεδομένων σε όλους τους τομείς, η χρήση και επαναχρησιμοποίηση τους σε εργασίες οδήγησαν τον άνθρωπο στην ανακάλυψη νέων μεθόδων με σκοπό την διευκόλυνση του. Λύση ήρθε να δώσει η επιστήμη της πληροφορικής, όπου πλέον όλα τα δεδομένα μπορούν να καταγραφούν και να αποθηκευτούν γρήγορα και με υψηλή ανάλυση σε σκληρούς δίσκους. Πλέον οι υπολογιστές μπορούν να καταγράφουν και να επεξεργάζονται τεράστιες βάσεις δεδομένων με πολύ γρήγορο ρυθμό, κάτι που ο άνθρωπος αδυνατεί να το κάνει.

Νοημοσύνη είναι η γενική ικανότητα που εκφράζεται μέσα από τις διαδικασίες υπολογισμών, συλλογισμών, λογικής, μάθησης, εξοικείωσης σε νέες καταστάσεις και νέο περιβάλλον για την επίτευξη κάποιου συγκεκριμένου στόχου. Υπάρχουν πολλά είδη νοημοσύνης όπως για παράδειγμα μουσική νοημοσύνη, γλωσσική, μαθηματική. Τεχνητή Νοημοσύνη είναι ένας γενικός όρος που υποδηλώνει κάθε είδος νοημοσύνης που δεν είναι βιολογική. Κυριαρχούν δύο προσεγγίσεις για την τεχνητή νοημοσύνη. Η ασθενής προσέγγιση όπου η ανθρώπινη γνώση μπορεί να προσομοιαστεί σε

υπολογιστικά συστήματα και η ισχυρή προσέγγιση όπου η ανθρώπινη γνώση μπορεί να εκδηλωθεί σε υπολογιστικά συστήματα αυτόματα.

1.2 Υπολογιστική Νοημοσύνη

Υπολογιστική Νοημοσύνη, που αποτελεί κομμάτι της Τεχνητής Νοημοσύνης, είναι ο τομέας της πληροφορικής που ασχολείται με την ανάλυση, σχεδίαση και υλοποίηση υπολογιστικών συστημάτων τα οποία έχουν την ικανότητα να μιμούνται διάφορα χαρακτηριστικά και συμπεριφορές του ανθρώπου. Τα χαρακτηριστικά αυτά μπορεί να είναι η εξαγωγή συμπερασμάτων για πρόβλεψη διάφορων καταστάσεων, ασθενειών ή καιρικών φαινομένων, η μάθηση, η εξέλιξη, η προσαρμοστικότητα σε νέες καταστάσεις, η ικανότητα επίλυσης προβλημάτων. Αρχικά η χρήση της υπολογιστικής νοημοσύνης ήταν σε εφαρμογές αναγνώρισης εικόνας, στην συνέχεια όμως εφαρμόστηκε σε πολλούς άλλους τομείς.

Οι τρεις κύριες υποστηρικτικές μέθοδοι που χρησιμοποιούνται στην Υπολογιστική Νοημοσύνη είναι οι εξής:

- Τεχνητά Νευρωνικά Δίκτυα
- Ασαφής λογική
- Εξελικτικός υπολογισμός

Υπάρχουν πολλές εφαρμογές της Υπολογιστικής Νοημοσύνης σε διάφορους τομείς. Ένα παράδειγμα στην ιατρική, είναι η συλλογή και επεξεργασία δεδομένων από αποτελέσματα κάποιων εξετάσεων που γίνονται σε ανθρώπους για την πρόβλεψη καρκίνου του πνεύμονα, για πρόβλεψη διαβήτη, πρόβλεψη χρωμοσωμικών ανωμαλιών. Στην μελέτη [5], παρουσιάζεται μια μέθοδος όπου οι γιατροί ελέγχουν ένα οζίδιο στον πνεύμονα βγάζοντας ακτίνες και από τις ακτίνες αυτές ανακτάται η γεωμετρία του πνεύμονα και το σχήμα του οζιδίου σε διαφορετικούς χρόνους. Επιλέγεται κατάλληλο μοντέλο με βάση την πραγματική εξέλιξη του όγκου στις εικόνες και εκτιμούνται οι παράμετροι του. Από την αρχική μορφή του οζιδίου, τρέχουν πολλές προσομοιώσεις του μοντέλου, χρησιμοποιώντας ένα σύνολο παραμέτρων. Βγαίνει μια συλλογή αποτελεσμάτων και γίνεται η πρόβλεψη. Η πρόβλεψη βασίζεται στις παραμέτρους υπολογισμού χρησιμοποιώντας τις εικόνες που πάρθηκαν. Το μοντέλο λαμβάνει υπόψη

τη θεραπευτική συγκέντρωση στον πνεύμονα του ασθενή και την εξέλιξη του πληθυσμού των κυττάρων του.

Επίσης στην ιατρική, άλλο ένα παράδειγμα είναι η συλλογή δεδομένων από εξετάσεις ασθενών για να γίνει πρόβλεψη καρκίνου του προστάτη. Στο άρθρο [2], αναφέρεται ότι δημιουργήθηκε ένα νευρωνικό δίκτυο με δεδομένα που πάρθηκαν από ιατρικές εξετάσεις 151 ανδρών. Τα δεδομένα αυτά ήταν η ηλικία, ο αριθμός κρεατίνης, ο συνολικός αριθμός ειδικού αντιγόνου προστάτη (PSA, Prostate Specific Antigen), η πυκνότητα του ειδικού αντιγόνου προστάτη, στοιχεία διατροφής, και γενικά στοιχεία υγείας του ασθενή. Το νευρωνικό δίκτυο έβγαλε ως πρόβλεψη ότι 37 στους 151 άνδρες έπασχαν από καρκίνο του προστάτη. Στοιχεία έδειξαν ότι περίπου 30% των ανδρών υποβάλλονται σε ριζική τομή αφαίρεσης του καρκίνου και το 25% των ανδρών που έχουν καρκίνο του προστάτη έχουν τιμή ειδικού αντιγόνου προστάτη λιγότερη από 4 ng/ml .

Ένα άλλο παράδειγμα στην τεχνολογία είναι στην εφαρμογή αναγνώρισης προσώπου. Για παράδειγμα στο άρθρο [11], γίνεται προσπάθεια να ανιχνευτούν πρόσωπα σε εικόνες. Εφαρμόζονται φίλτρα νευρωνικών δικτύων σε μια εικόνα σε διάφορες κλίμακες και μετά συνδυάζονται οι έξοδοι. Τα φίλτρα αυτά δέχονται ως είσοδο 20 X 20 pixels εικόνα και για να γίνει η ανίχνευση το φίλτρο εφαρμόζεται σε κάθε μέρος της εικόνας. Σαν πρώτο βήμα, ο αλγόριθμος φιλτραρίσματος προσαρμόζεται σε ένα παράθυρο-κομμάτι της εικόνας το οποίο περνά μέσω από ένα νευρωνικό δίκτυο για να αποφανθεί κατά πόσο το παράθυρο περιέχει πρόσωπο ή όχι. Το βήμα αυτό προσπαθεί να εξισώσει τις τιμές έντασης κατά μήκος του παραθύρου. Ταιριάζετε μια συνάρτηση η οποία μεταβάλλεται γραμμικά κατά μήκος του παραθύρου με τις τιμές έντασης σε μια οβάλ περιοχή στο παράθυρο. Τα pixels έξω από την οβάλ περιοχή αγνοούνται γιατί πιθανών να είναι το φόντο της εικόνας. Στη συνέχεια υπολογίζεται το ιστόγραμμα των pixels που βρίσκονται μέσα στην οβάλ περιοχή. Το νευρωνικό δίκτυο στο οποίο περνά το παράθυρο της εικόνας, έχει 3 διαφορετικούς τύπους κρυφών επιπέδων. Οι 3 αυτοί τύποι βοηθούν στην ανίχνευση τοπικών χαρακτηριστικών ενός προσώπου. Το νευρωνικό δίκτυο δίνει στην έξοδο τιμές από 1 έως -1 σηματοδοτώντας την παρουσία ή την απουσία προσώπου. Το σύστημα μπόρεσε να ανιχνεύσει 90,5% των προσώπων από ένα σύνολο δοκιμών 130 σύνθετων εικόνων.

Τέλος παράδειγμα στην ρομποτική είναι η κατασκευή ρομπότ με ανθρώπινη συμπεριφορά έτσι ώστε να βοηθούν τον άνθρωπο κάνοντας εργασίες πιο γρήγορα και πιο αποδοτικά. Συγκεκριμένο παράδειγμα στο βιβλίο [7], όπου ρομποτική και ιατρική συνεργάζονται χρησιμοποιώντας τα ρομπότ σαν βοηθούς στην χειρουργική. Κάνουν χρήση ενός συστήματος υπολογιστικής όρασης για να δημιουργήσουν ένα τρισδιάστατο μοντέλο της εσωτερικής ανατομίας του ασθενή και τέλος χρησιμοποιούν ρομποτικό έλεγχο για να καθοδηγήσουν την εισαγωγή ενός προσθετικού γοφού.

Στο κεφάλαιο 2, αναλύονται σε θεωρητικό υπόβαθρο οι τρεις μέθοδοι που αναφέρονται πιο πάνω. Στο κεφάλαιο 3, γίνεται αναφορά και ανάλυση σε εργασίες που έγιναν τόσο στο Πανεπιστήμιο Κύπρου όσο και σε διεθνές επίπεδο. Στο κεφάλαιο 4, ορίζεται το θέμα και η ανάλυση του θέματος της εργασίας αυτής, και γίνεται σύγκριση και συζήτηση των μεθόδων αντιμετώπισης του που αναφέρθηκαν σε προηγούμενο κεφάλαιο. Στο κεφάλαιο 5, αναλύονται κάποιες τεχνικές που χρησιμοποιήθηκαν, και αναλύονται τα αποτελέσματα. Στο κεφάλαιο 6, γίνεται μια κριτική ανάλυση των τεχνικών και αποτελεσμάτων και προτείνονται κάποιες εισηγήσεις προς τον ειδικό. Και τέλος στο κεφάλαιο 7, γίνεται αναφορά για θέματα μελλοντικής μελέτης.

Κεφάλαιο 2

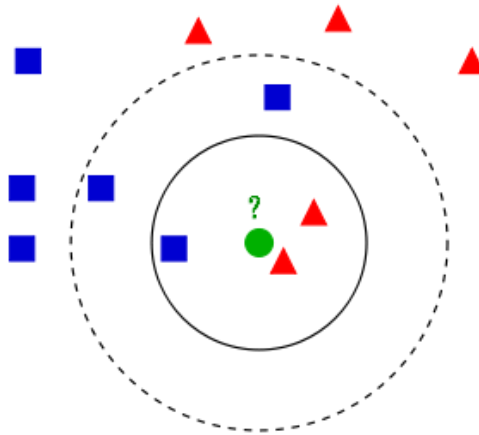
Τεχνικές Υπολογιστικής Νοημοσύνης

2.1 K-Κοντινότερος Γείτονας (K-nn)	5
2.2 Νευρωνικά Δίκτυα	6
2.2.1 Βιολογικοί Νευρώνες	7
2.2.2 Τεχνητοί Νευρώνες	8
2.2.3 Τεχνητά Νευρωνικά Δίκτυα	8
2.2.4 Ιδιότητες Νευρωνικών Δικτύων	13
2.3 Μηχανές Υποστήριξης Διανυσμάτων (SVM)	13
2.4 Εξελικτικά Συστήματα	15
2.4.1 Θεωρία της Εξέλιξης	15
2.4.2 Γενετικοί Αλγόριθμοι	16
2.4.3 Γενετικός προγραμματισμός	18
2.4.4 Στρατηγικές εξέλιξης	20
2.4.5 Εξελικτικός προγραμματισμός	20
2.5 Ασαφής Λογική	21

2.1 K-Κοντινότερος Γείτονας (K-nn)

Ο αλγόριθμος K κοντινότερου γείτονα [19] είναι μια μη παραμετρική μέθοδος για κατηγοριοποίηση αντικειμένων που βασίζονται στο πιο κοντινό παράδειγμα στο χώρο των χαρακτηριστικών. Η κατηγοριοποίηση του αντικειμένου γίνεται με τον ακόλουθο τρόπο (Σχήμα 2.1): το αντικείμενο κατατάσσεται στην κλάση που είναι η πιο κοινή μεταξύ των k κοντινότερων γειτόνων του. Οι γείτονες λαμβάνονται από ένα σύνολο αντικειμένων των οποίων η σωστή κατηγοριοποίηση είναι γνωστή. Ο αλγόριθμος λειτουργεί ως εξής: τα δεδομένα εκπαίδευσης είναι διανύσματα σε ένα πολυδιάστατο χώρο χαρακτηριστικών, όπου κάθε δεδομένο έχει μια επιγραφή (label) που αναγράφει

την κλάση του. Η φάση εκπαίδευσης του αλγορίθμου αποτελείται μόνο από την αποθήκευση των διανυσμάτων των χαρακτηριστικών και των επιγραφών (label) των κλάσεων των δεδομένων εκπαίδευσης. Στη φάση κατηγοριοποίησης, το k είναι μια σταθερά που ορίζεται από το χρήστη. Ένα διάνυσμα κατηγοριοποιείται αναθέτοντας την επιγραφή (label) που είναι η πιο συχνή μεταξύ των k δεδομένων εκπαίδευσης.



Σχήμα 2.1: Εδώ φαίνεται η κατηγοριοποίηση με τη χρήση KNN αλγόριθμου. Η πράσινη κουκίδα μπορεί να κατηγοριοποιηθεί είτε στη μια κατηγορία (τα μπλε τετράγωνα) είτε στην άλλη (κόκκινα τρίγωνα). Αν η τιμή του k είναι 3 (ο πρώτος μαύρος κύκλος) τότε κατηγοριοποιείται στη δεύτερη κατηγορία (κόκκινα τρίγωνα), γιατί υπάρχουν 2 τρίγωνα και μόνο 1 τετράγωνο στον κύκλο αυτό. Αν η τιμή του k είναι 5 (ο διακεκομμένος μαύρος κύκλος) τότε κατηγοριοποιείται στη πρώτη κατηγορία (μπλε τετράγωνα), γιατί υπάρχουν 2 κόκκινα τρίγωνα και 3 μπλε τετράγωνα στον κύκλο αυτό.

2.2 Νευρωνικά Δίκτυα

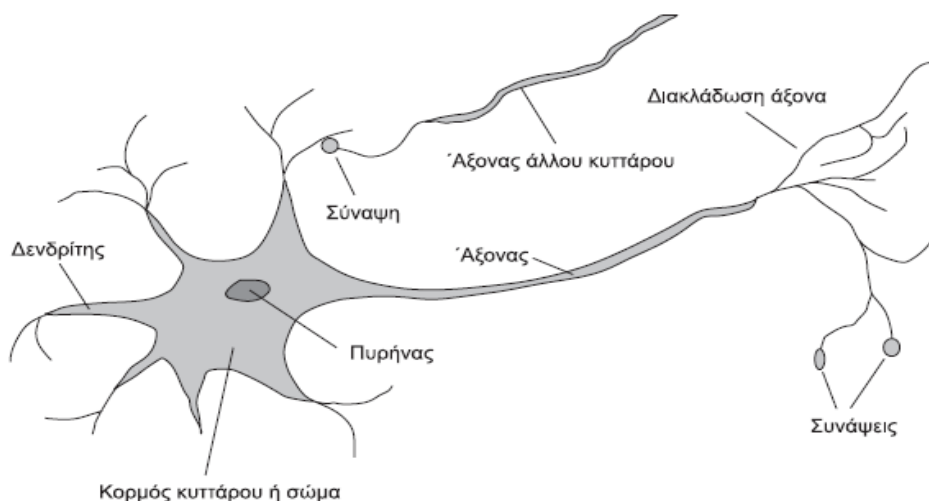
Τα νευρωνικά δίκτυα είναι υπολογιστικά συστήματα επεξεργασίας πληροφοριών τα οποία δέχονται εισόδους και παράγουν εξόδους για την επίλυση διαφόρων προβλημάτων. Χρησιμοποιούνται κυρίως στην αναγνώριση προτύπων. Τα συστήματα αυτά έχουν κατασκευαστεί με πρότυπο τους νευρώνες των βιολογικών νευρώνων του εγκεφάλου και αποτελούν μια προσπάθεια προσέγγισης της λειτουργίας του ανθρώπινου εγκεφάλου από μια μηχανή. Στόχος των δικτύων αυτών είναι η

προσομοίωση του τρόπου με τον οποίο ο ανθρώπινος εγκέφαλος επεξεργάζεται τις πληροφορίες, τις μαθαίνει και τις θυμάται.

2.2.1 Βιολογικοί Νευρώνες

Οι βιολογικοί νευρώνες είναι τα κύρια δομικά στοιχεία (κύτταρα) του εγκεφάλου και του νευρικού συστήματος. Ένας τέτοιος νευρώνας αποτελείται από το σώμα, τους δενδρίτες και τον άξονα (σχήμα 2.2). Ο νευρώνας δέχεται σήματα εισόδου από τους γειτονικούς νευρώνες μέσω των δενδρίτων και στέλλει σήματα εξόδου σε άλλους νευρώνες μέσω του άξονα του. Κάθε άξονας καταλήγει σε μια σύναψη όπου αντιπροσωπεύει τη σύνδεση μεταξύ του άξονα και του δενδρίτη. Ένας νευρώνας μπορεί να έχει πολλούς δενδρίτες άρα είναι πιθανόν να δέχεται πολλές διαφορετικές εισόδους την ίδια στιγμή. Ακολουθώς ο νευρώνας μετατρέπει τα σήματα εισόδου σε ένα σήμα εξόδου για να σταλεί σε γειτονικούς νευρώνες. Οι πληροφορίες σε ένα νευρώνα κωδικοποιούνται με ηλεκτροχημικές διαδικασίες και υπολογίζεται ότι το συνολικό μήκος των συνάψεων στον εγκέφαλο είναι περίπου 10^9 μέτρα.

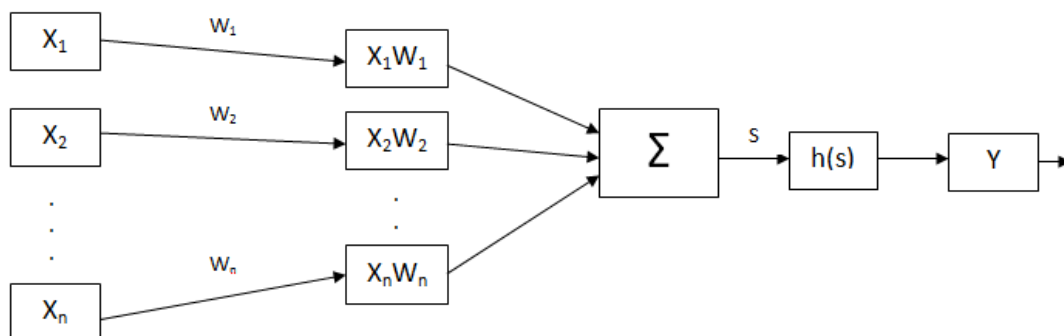
Τα βιολογικά νευρωνικά δίκτυα έχουν σαν πλεονέκτημα τη δυνατότητα να μαθαίνουν, να γενικεύουν και να χειρίζονται ανακριβείς πληροφορίες λύνοντας διάφορα προβλήματα.



Σχήμα 2.2: Δομή ενός βιολογικού νευρώνα

2.2.2 Τεχνητοί Νευρώνες

Ένας τεχνητός νευρώνας είναι ένας «επεξεργαστής» που δέχεται είσοδο από το περιβάλλον ή από άλλο νευρώνα και με βάση μια συνάρτηση ενεργοποίησης παράγει ένα σήμα. Το σήμα αυτό, το οποίο μεταβιβάζεται στον επόμενο νευρώνα, μπορεί να είναι μια απόφαση ή η απάντηση στο πρόβλημα το οποίο προσπαθεί να λύσει ο νευρώνας. Ένας τέτοιος νευρώνας μπορεί να έχει πολλές εισόδους αλλά μόνο μια έξοδο και μπορεί να προσαρμόζεται εύκολα σε διάφορα προβλήματα. Όπως φαίνεται και πιο κάτω (σχήμα 2.3), στις εισόδους του νευρώνα εφαρμόζεται ένα σύνολο δεδομένων με πραγματικές τιμές, όπου το καθένα είναι αποτέλεσμα της εξόδου ενός άλλου νευρώνα. Το κάθε στοιχείο X_i πολλαπλασιάζεται με το αντίστοιχο βάρος W_i και στη συνέχεια όλα τα γινόμενα αθροίζονται. Το άθροισμα αυτό, όταν περαστεί σαν είσοδος μέσα από μια συνάρτηση ενεργοποίησης $h(s)$, δίνει σαν αποτέλεσμα την έξοδο του νευρώνα, που αυτή θα είναι η είσοδος για τον επόμενο γειτονικό νευρώνα.



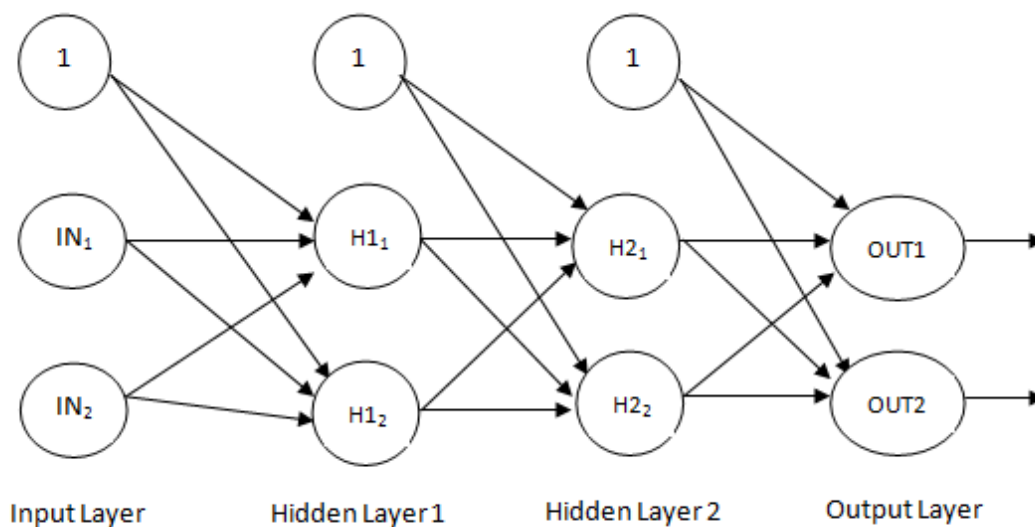
Σχήμα 2.3: Τεχνητός νευρώνας με πολλούς εισόδους αλλά μόνο μια έξοδο, όπου X_i είναι η είσοδος του νευρώνα, W_i το βάρος της κάθε εισόδου, Σ το άθροισμα των γινομένων όλων των νευρώνων, $h(s)$ μια συνάρτηση ενεργοποίησης και Y η έξοδος του νευρώνα.

2.2.3 Τεχνητά Νευρωνικά Δίκτυα

Τα τεχνητά νευρωνικά δίκτυα είναι η μελέτη δικτύων με κόμβους, οι οποίοι μέσω μιας διαδικασίας μάθησης από παραδείγματα που είναι αποθηκευμένα σε πειραματική γνώση, μπορούν να χρησιμοποιηθούν για λύση κάποιων προβλημάτων. Είναι μη γραμμικά συστήματα πολλών επεξεργαστών που παρουσιάζουν έξυπνη συμπεριφορά

και βοηθούν στη μελέτη βιολογικών συμπεριφορών, στη μελέτη συστημάτων ελέγχου, στην επεξεργασία δεδομένων και σημάτων, στην πρόγνωση φαινομένων και ασθενειών καθώς επίσης και στη διάγνωση ασθενειών και ανωμαλιών. Η ανάγκη επίλυσης προβλημάτων που δεν διέπονται από κανόνες (κάτι που ένας υπολογιστής δεν μπορεί να επιλύσει) ήταν η αιτία για τη χρήση των νευρωνικών δικτύων.

Αποτελούνται από πολλούς τεχνητούς νευρώνες (σχήμα 2.4), που είναι συνδεδεμένοι μεταξύ τους και τα περισσότερα νευρωνικά δίκτυα είναι χωρισμένα σε επίπεδα. Κάθε επίπεδο περιέχει ένα σύνολο τεχνητών νευρώνων που ονομάζονται κόμβοι. Στο πρώτο επίπεδο, επίπεδο εισόδου, κάθε κόμβος αντιστοιχεί και σε ένα χαρακτηριστικό. Στο τελευταίο επίπεδο, επίπεδο εξόδου, κάθε κόμβος αντιστοιχεί στην απάντηση που δίνει το δίκτυο για το κάθε πρόβλημα. Τα ενδιάμεσα επίπεδα είναι τα κρυφά επίπεδα. Οι συνδέσεις μεταξύ των κόμβων χαρακτηρίζονται από μια τιμή που ονομάζεται βάρος και παίζει σημαντικό ρόλο στην φάση εκπαίδευσης του δικτύου.



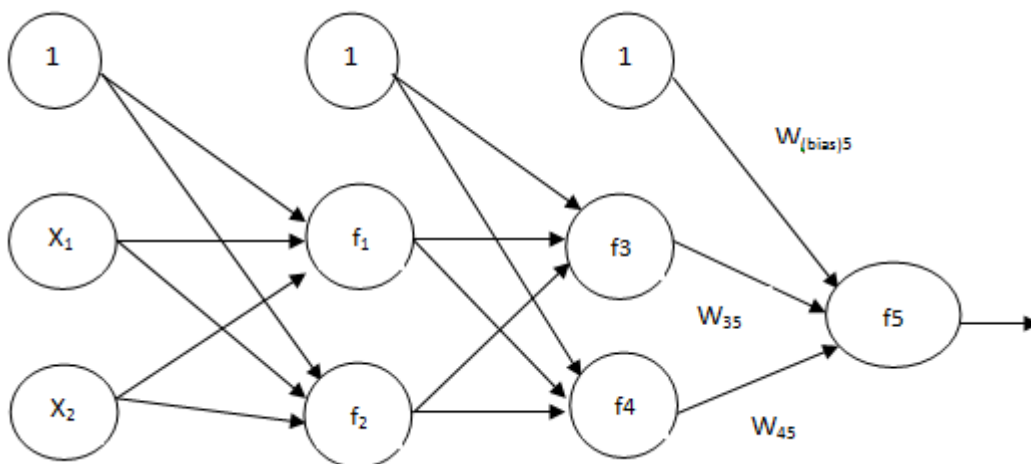
Σχήμα 2.4: Τεχνητό Νευρωνικό Δίκτυο (feed forward) με το επίπεδο εισόδου, τα κρυφά επίπεδα και το επίπεδο εξόδου.

Για να μπορέσει το δίκτυο να κάνει πρόβλεψη και να δώσει ένα αποτέλεσμα, πρέπει να περάσει από κατάλληλη φάση μάθησης. Υπάρχουν 3 μέθοδοι μάθησης: η επιβλεπόμενη μάθηση, η μη επιβλεπόμενη μάθηση και η ενισχυτική μάθηση. Στην επιβλεπόμενη μάθηση, είναι γνωστή η επιθυμητή έξοδος για συγκεκριμένες εισόδους και έτσι γίνεται

αλλαγή των βαρών των νευρώνων έτσι ώστε να ελαχιστοποιηθεί το σφάλμα μεταξύ του πραγματικού και του επιθυμητού αποτελέσματος. Στην μη επιβλεπόμενη μάθηση, είναι γνωστές μόνο οι εισοδοι στο δίκτυο και γίνεται η υπόθεση ότι υπάρχουν δεδομένα με κοινά χαρακτηριστικά έτσι ώστε να ομαδοποιηθούν μαζί. Και τέλος στην ενισχυτική μάθηση, υπάρχει ένας «κριτής» που αξιολογεί την πράξη του δικτύου και δίνει αμοιβή ή τιμωρία.

Back Propagation:

Ο αλγόριθμος αυτός χωρίζεται σε 2 φάσεις: στο Forward pass, και στο Backward pass. Στην πρώτη φάση (σχήμα 2.5), αφού αρχικοποιηθούν τα βάρη με μικρές τυχαίες τιμές, υπολογίζεται το άθροισμα του γινομένου των βαρών επί την είσοδο για τον κάθε νευρώνα για όλα τα επίπεδα μέχρι να βρεθεί το αποτέλεσμα του κόμβου του επιπέδου εξόδου. Εισάγεται μια επιπρόσθετη είσοδος στο κάθε επίπεδο νευρώνων, που είναι πάντα ενεργή (η τιμή της είναι πάντα 1) και το βάρος της συμβολίζεται με W_{bias} . Στη δεύτερη φάση (σχήμα 2.6, σχήμα 2.7), ξεκινώντας από το επίπεδο εξόδου, με βάση συγκεκριμένους τύπους βρίσκει τη διαφορά μεταξύ του πραγματικού και του επιθυμητού αποτελέσματος και τη μεταδίδει προς τα πίσω. Αυτό γίνεται για κάθε νευρώνα σε κάθε επίπεδο. Στη συνέχεια, ξεκινώντας από το επίπεδο εισόδου, χρησιμοποιώντας τους κατάλληλους τύπους γίνεται η αλλαγή των βαρών έτσι ώστε να ελαχιστοποιηθεί το σφάλμα. Η όλη διαδικασία συνεχίζεται μέχρι να ελαχιστοποιηθεί όσο το δυνατό καλύτερα το σφάλμα.



Σχήμα 2.5: Forward pass όπου υπολογίζεται το άθροισμα του γινομένου των βαρών επί την είσοδο για τον κάθε νευρώνα για όλα τα επίπεδα με τον ακόλουθο τύπο:

$$y_5 = f_5(w_{(\text{bias})5} + w_{35} \times y_3 + w_{45} \times y_4)$$

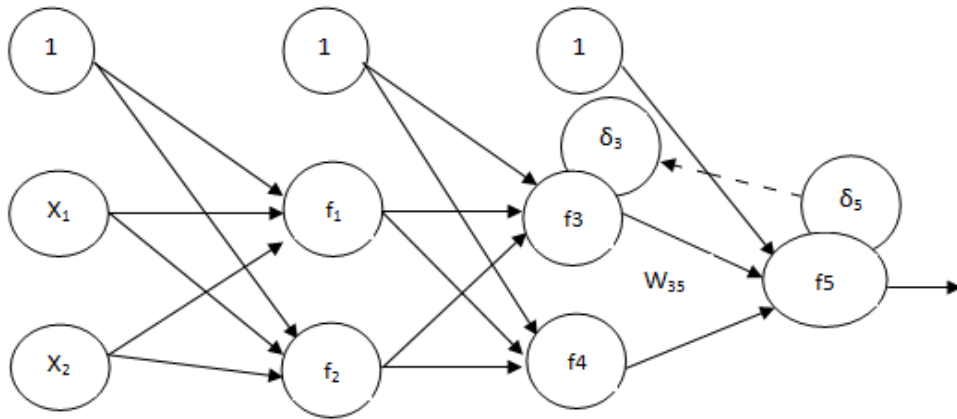
Όπου:

$w_{(\text{bias})i}$: το βάρος της επιπρόσθετης εισόδου i που είναι πάντα ενεργή

f : η συνάρτηση ενεργοποίησης

w_{ij} : το βάρος που ενώνει τον νευρώνα i με τον νευρώνα j

y_i : η έξοδος του νευρώνα i



Σχήμα 2.6: Backward pass όπου βρίσκουμε τη διαφορά μεταξύ του πραγματικού και του επιθυμητού αποτελέσματος και τη μεταδίδουμε προς τα πίσω. Η διαφορά υπολογίζεται με τον ακόλουθο τύπο:

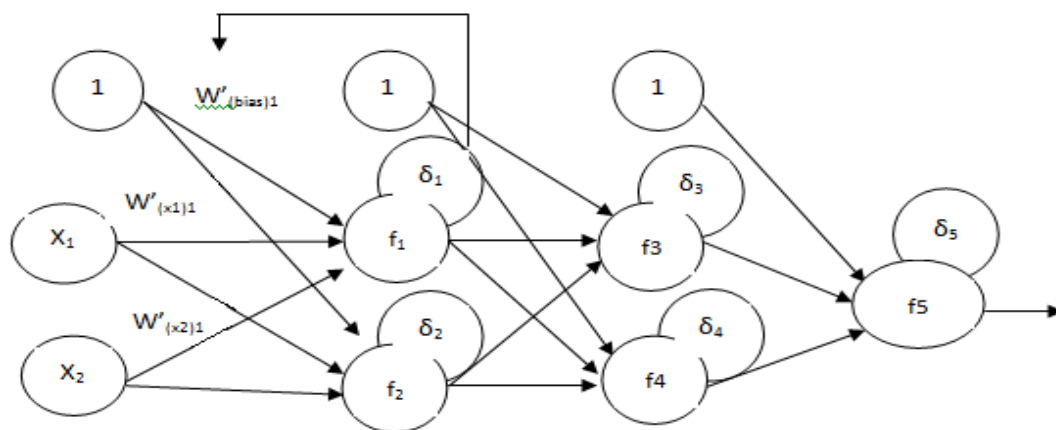
$$\delta_3 = y_3 \times (1 - y_3) \times (w_{35} \times \delta_5)$$

Όπου:

w_{ij} : το βάρος που ενώνει τον νευρώνα i με τον νευρώνα j

y_i : η έξοδος του νευρώνα i

δ_i : διαφορά πραγματικού και επιθυμητού αποτελέσματος



Σχήμα 2.7: Backward pass ξεκινώντας από το επίπεδο εισόδου, γίνεται η αλλαγή των βαρών έτσι ώστε να ελαχιστοποιηθεί το σφάλμα. Η αλλαγή των βαρών γίνεται με τους ακόλουθους τύπους:

$$w'_{(bias)1} = w_{(bias)1} - n \times \delta_1$$

$$w'_{(x1)1} = w_{(x1)1} - n \times \delta_1 \times x_1$$

$$w'_{(x2)1} = w_{(x2)1} - n \times \delta_1 \times x_2$$

Όπου:

$w_{(bias)i}$: το βάρος της επιπρόσθετης εισόδου i που είναι πάντα ενεργή

$w'_{(bias)i}$: το νέο βάρος της επιπρόσθετης εισόδου i που είναι πάντα ενεργή

w_{ij} : το βάρος που ενώνει τον νευρώνα i με τον νευρώνα j

w'_{ij} : το νέο βάρος που ενώνει τον νευρώνα i με τον νευρώνα j

x_i : τιμή νευρώνα i

δ_i : διαφορά πραγματικού και επιθυμητού αποτελέσματος

n : ρυθμός μάθησης

2.2.4 Ιδιότητες Νευρωνικών Διχτύων

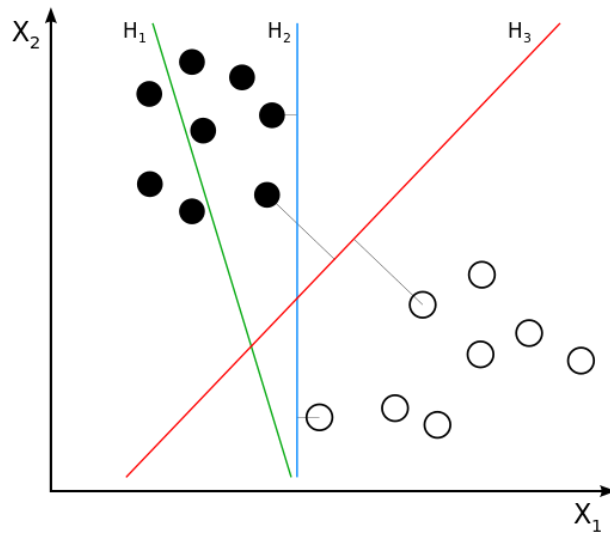
Τα νευρωνικά δίκτυα έχουν την ιδιότητα της ανθεκτικότητας και της σταθερότητας. Ακόμα και αν αφαιρεθούν νευρώνες από το δίκτυο, αυτό συνεχίζει να συμπεριφέρεται κανονικά. Επίσης είναι ανθεκτικά στα λάθη. Ακόμα και αν τα δεδομένα εισόδου είναι ασαφή, το δίκτυο συνεχίζει να συμπεριφέρεται ικανοποιητικά. Και τέλος έχουν την ικανότητα της γενίκευσης. Μπορούν δηλαδή να προτείνουν λύσεις, να δώσουν αποτελέσματα σε συγκεκριμένα προβλήματα και να χειριστούν ασαφείς και ανακριβείς πληροφορίες.

2.3 Μηχανές Υποστήριξης Διανυσμάτων (SVM)

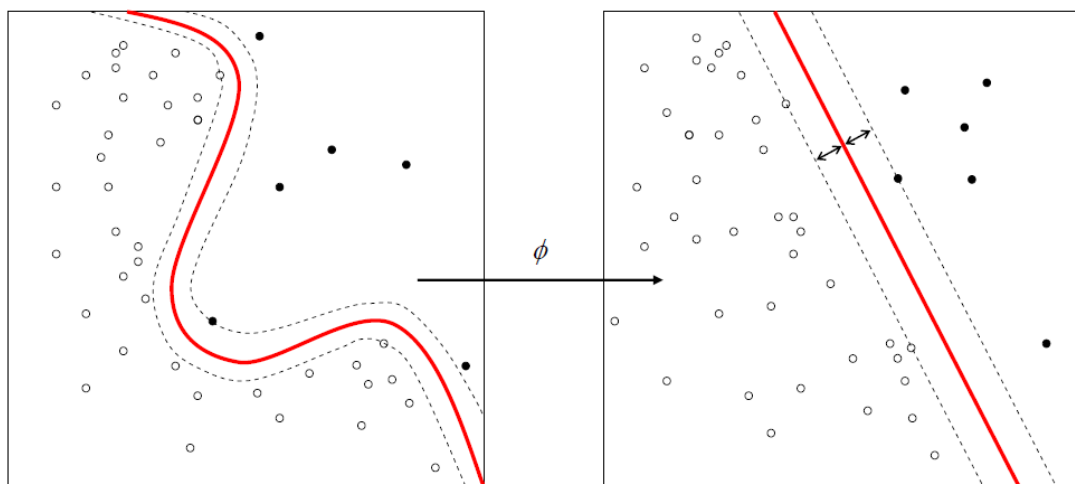
Οι μηχανές διανυσμάτων υποστήριξης [20] είναι ένα σύνολο μεθόδων που χρησιμοποιούνται για ταξινόμηση και αποτελούν εργαλείο πρόβλεψης για διάφορα προβλήματα. Μια τέτοια μηχανή διανυσμάτων δημιουργεί ένα υπερεπίπεδο ή ένα σύνολο υπερεπιπέδων σε ένα ψηλό χώρο διαστάσεων (διανύσματα). Μια μηχανή διανυσμάτων παίρνει ένα σύνολο δεδομένων εισόδου και προβλέπει για κάθε είσοδο ποια από τις δύο κλάσεις αποτελεί την έξοδο. Πιο συγκεκριμένα, παίρνοντας ένα σύνολο δεδομένων εκπαίδευσης, όπου το καθένα χαρακτηρίζεται σε ποια από τις δύο κατηγορίες ανήκει, μια τέτοια μηχανή δημιουργεί ένα μοντέλο το οποίο αναθέτει νέα δεδομένα στην μια ή στην άλλη κατηγορία. Το μοντέλο αυτό, είναι μια αναπαράσταση των δεδομένων ως σημεία στο χώρο, που σημειώθηκαν έτσι ώστε δεδομένα ξεχωριστών κατηγοριών να χωρίζονται από ένα κενό. Στη συνέχεια νέα δεδομένα αντιστοιχίζονται στον ίδιο χώρο και γίνεται η πρόβλεψη σε ποια από τις δύο κατηγορίες ανήκουν, ανάλογα με την πλευρά στην οποία θα πέσουν μέσα. Αυτός ο ταξινομητής ονομάζεται γραμμικός ταξινομητής (Σχήμα 2.8). Υπάρχει επίσης ο μη γραμμικός ταξινομητής (Σχήμα 2.9), όπου και πάλι εφαρμόζεται ο πιο πάνω τρόπος για τον διαχωρισμό των εξόδων αλλά με χρήση πυρήνων. Κάθε σημείο αντικαθίσταται από μια μη γραμμική συνάρτηση πυρήνα. Αυτό επιτρέπει στον αλγόριθμο να ταιριάζει το μέγιστο υπερεπίπεδο σε ένα μετασχηματισμένο χώρο χαρακτηριστικών.

Γιατί SVM;

Η ελαχιστοποίηση του σφάλματος στις μηχανές υποστήριξης, σε ορισμένες εφαρμογές, είναι ανώτερη από την ελαχιστοποίηση του σφάλματος των νευρωνικών δικτύων και αυτή είναι μια σημαντική διαφορά των δύο μεθόδων. Εντούτοις όμως, αποτελούν μια προέκταση των δικτύων Perceptron, και η εκπαίδευση τους είναι αρκετά εύκολη.



Σχήμα 2.8: Γραμμικός ταξινομητής όπου φαίνεται ότι η γραμμή απόφασης H1 δεν διαχωρίζει σωστά τις δύο κατηγορίες, ενώ η γραμμή απόφασης H2 και H3 χωρίζουν σωστά τις δύο κατηγορίες.



Σχήμα 2.9: Μη γραμμικός ταξινομητής που φαίνεται ότι δεν μπορεί να χωρίσει τις δύο κατηγορίες με γραμμικό τρόπο, ενώ με τον κατάλληλο μετασχηματισμό φαίνεται ότι χωρίζει σωστά τις δύο κατηγορίες.

2.4 Εξελικτικά Συστήματα

Τα εξελικτικά συστήματα είναι συστήματα υπολογιστών που μελετούν προγράμματα, τα οποία δίνουν λύσεις σε διάφορα προβλήματα που ενσωματώνουν εξελικτικές αρχές της φύσης. Τα συστήματα αυτά είναι εμπνευσμένα από τη βιολογία. Η περιοχή των εξελικτικών συστημάτων περιλαμβάνει 4 τεχνικές:

- Γενετικοί αλγόριθμοι
- Γενετικός προγραμματισμός
- Στρατηγικές εξέλιξης
- Εξελικτικός προγραμματισμός

Οι τεχνικές αυτές ανήκουν σε μια κατηγορία επίλυσης προβλημάτων που είναι ευρύτερα γνωστές και ως Εξελικτικοί Αλγόριθμοι. Τα μειονεκτήματα των κλασικών μεθόδων αναζήτησης και βελτιστοποίησης καθώς επίσης και η συνεχής αυξανόμενη ανάγκη παραγωγής λογισμικού που να μπορεί να εκμεταλλεύεται πιο αποδοτικά τις τεράστιες δυνατότητες του υλικού, ήταν οι βασικές αιτίες που ώθησαν τους επιστήμονες στην αναζήτηση νέων τεχνικών. Η πιο γνωστή τεχνική είναι η τεχνική των Γενετικών Αλγορίθμων.

2.4.1 Θεωρία της Εξέλιξης

Η θεωρία της εξέλιξης των ειδών που παρατάθηκε από το Δαρβίνο στα μέσα του περασμένου αιώνα, σκοπό είχε να δώσει μια εξήγηση για το φαινόμενο της ζωής, την προέλευση της και τις βασικές λειτουργίες της. Προκάλεσε μεγάλη αναστάτωση αφού ερχόταν σε σύγκρουση με τις επικρατέστερες θρησκευτικές αντιλήψεις περί προέλευσης της ζωής. Η θεωρία αυτή αναφέρει ότι οι οργανισμοί που δεν μπορούν να επιβιώσουν πεθαίνουν, ενώ οι υπόλοιποι πολλαπλασιάζονται μέσω της αναπαραγωγής. Επίσης, οι απόγονοι παρουσιάζουν μικρές διαφοροποιήσεις από τους προγόνους καθώς

επίσης, κάποιες τυχαίες μεταλλάξεις που συμβαίνουν, σπάνια οδηγούν στη δημιουργία νέων καλύτερων οργανισμών. Τέλος, αν το περιβάλλον αλλάζει με αργούς ρυθμούς, τα είδη εξελίσσονται έτσι ώστε να προσαρμοστούν σε αυτό. Τα κυριότερα σημεία της θεωρίας αυτής, ερμηνεύουν τον τρόπο λειτουργίας των Εξελικτικών Αλγορίθμων.

2.4.2 Γενετικοί Αλγόριθμοι

Οι γενετικοί αλγόριθμοι είναι μια μέθοδος η οποία προσφέρει εξέλιξη στον πληθυσμό των χρωμοσωμάτων, με σκοπό την εξεύρεση μιας καλής λύσης για ένα συγκεκριμένο πρόβλημα, μετά από μια σειρά γενεών. Ο βασικός μηχανισμός τους είναι εμπνευσμένος από τη Θεωρία της Εξέλιξης του Δαρβίνου. Είναι κυρίως αλγόριθμοι για βελτιστοποίηση και για αναζήτηση.

Οι γενετικοί αλγόριθμοι αναφέρονται σε άτομα-χρωμοσώματα μέσα σε ένα πληθυσμό. Τα χρωμοσώματα αποτελούνται από διάφορα στοιχεία-γονίδια (τμήματα του DNA) και κάθε γονίδιο επηρεάζει την κληρονομικότητα των χαρακτηριστικών. Κάθε άτομο αναπαριστά μια πιθανή λύση σε ένα πρόβλημα και η διαδικασία εξέλιξης που εφαρμόζεται στα χρωμοσώματα αντιστοιχεί σε ένα εκτενές ψάξιμο στο σύνολο των πιθανών λύσεων που ονομάζεται χώρος αναζήτησης ή χώρος καταστάσεων. Το αποκωδικοποιημένο περιεχόμενο ενός συγκεκριμένου χρωμοσώματος ονομάζεται φαινότυπος.

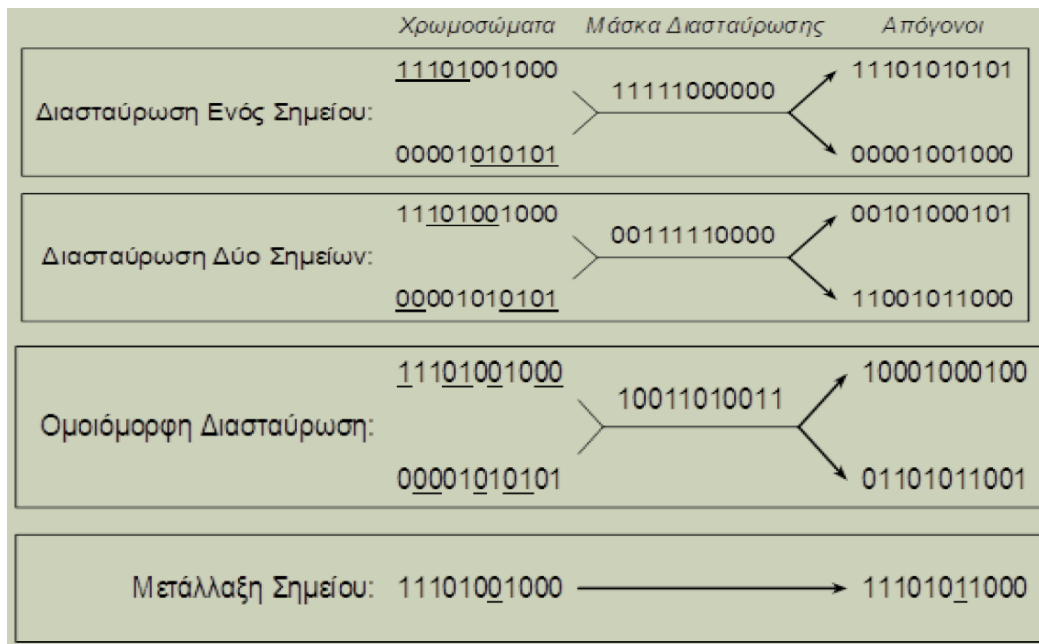
Αυτό που κάνει τους γενετικούς αλγόριθμους πιο εύρωστους από άλλες μεθόδους είναι το γεγονός ότι πραγματοποιούν αναζήτηση σε πολλές κατευθύνσεις διατηρώντας ένα πληθυσμό από πιθανές λύσεις και υποστηρίζουν ανταλλαγή πληροφοριών μεταξύ των κατευθύνσεων αυτών. Κάθε φορά, σε κάθε γενιά, οι καλές λύσεις αναπαράγονται ενώ οι κακές αφαιρούνται.

Βήματα Γενετικού Αλγορίθμου:

1. Κωδικοποίηση χρωμοσωμάτων
2. Αρχικοποίηση τιμών (δημιουργία αρχικού πληθυσμού)
3. Αποκωδικοποίηση χρωμοσωμάτων
4. Υπολογισμός ικανότητας ή αξιολόγησης (υπολογισμός Fitness)

5. Επιλογή χρωμοσωμάτων (επιλογή σύμφωνα με την συνάρτηση Fitness)
6. Αναπαραγωγή μεταξύ χρωμοσωμάτων
7. Διασταύρωση για δημιουργία νέου πληθυσμού-απογόνων
8. Μετάλλαξη (αν χρειάζεται)
9. Επανάληψη από το βήμα 2 μέχρι ένα σημείο τερματισμού

Ο πιο πάνω αλγόριθμος τερματίζει όταν βρεθεί μια τέλεια λύση με βάση τη συνάρτηση καταλληλότητας (Fitness), ή όταν γίνει σύγκλιση όλων των λύσεων σε μια. Η συνάρτηση καταλληλότητας παίρνει σαν είσοδο ένα χρωμόσωμα και δίνει ως έξοδο την ποιότητα του μοντέλου ή των ατόμων για ένα συγκεκριμένο σύνολο. Αφού επιλεγούν τα κατάλληλα άτομα για γονείς, γίνεται η αναπαραγωγή για τη δημιουργία απογόνων. Η επιλογή των ατόμων του πληθυσμού που θα περάσουν στον επόμενο πληθυσμό μπορεί να γίνει με τη χρήση της εξαναγκασμένης ρουλέτας, όπου κάθε μέλος του πληθυσμού έχει πιθανότητα επιλογής ίση με τη σχετική του απόδοση στον τρέχοντα πληθυσμό. Για την αναπαραγωγή στρίβεται η ρουλέτα και έτσι τα χρωμοσώματα που έχουν τη μεγαλύτερη απόδοση, θα έχουν μεγαλύτερο πλήθος απογόνων στην επόμενη γενιά, ενώ αυτές που έχουν χαμηλή απόδοση δεν θα υπάρχουν. Υπάρχουν πολλοί τελεστές που βοηθούν στην αναπαραγωγή όπως: διασταύρωση ενός σημείου, διασταύρωση δύο σημείων, ομοιόμορφη διασταύρωση, μετάλλαξη σημείου. Οι τελεστές αυτοί φαίνονται στο σχήμα 2.10. Μετά την αναπαραγωγή, γίνεται η διασταύρωση που στόχο έχει να προκύψουν άτομα στη νέα γενιά που θα διαφέρουν από τους γονείς τους και θα έχουν καλύτερα χαρακτηριστικά. Τέλος, γίνεται η μετάλλαξη (γίνεται αραιά). Ενεργεί σε ένα μόνο οργανισμό κάθε φορά και δρα βελτιωτικά για τους οργανισμούς. Στους βιολογικούς οργανισμούς, για την αναπαράσταση της λύσης, το αλφάβητο στο DNA αποτελείται από 4 γράμματα, ενώ στους γενετικούς αλγόριθμους η λύση αναπαρίσταται με μια δυαδική συμβολοσειρά.



Σχήμα 2.10: Διάφοροι τελεστές για την αναπαραγωγή χρωμοσωμάτων

Πλεονεκτήματα Γενετικών Αλγορίθμων από άλλες μεθόδους αναζήτησης και βελτιστοποίησης

Οι γενετικοί αλγόριθμοι μπορούν να λύσουν δύσκολα προβλήματα εύκολα και γρήγορα, καθώς επίσης μπορούν να συνεργαστούν και με άλλα υπάρχοντα συστήματα αφού χρησιμοποιούν πληροφορίες μόνο της διαδικασίας που θα βελτιστοποιήσουν. Είναι επεκτάσιμοι και εξελίξιμοι και δεν απαιτούν περιορισμούς στις συναρτήσεις που επεξεργάζονται. Επίσης δεν ψάχνουν σε ένα μόνο σημείο, αλλά σε ένα πλήθος σημείων. Το σημαντικότερο τους πλεονέκτημα είναι ότι στηρίζονται στη θεωρία της εξέλιξης των ειδών, όπως εμφανίζεται στη ζωή.

2.4.3 Γενετικός προγραμματισμός

Ο Γενετικός προγραμματισμός αναπτύχθηκε από τον Stanford's John Koza το 1980 αντιμετωπίζοντας με αυτό τον τρόπο τους εξής περιορισμούς. Πρώτος περιορισμός ήταν το γεγονός ότι τα προγράμματα περιορίζονταν σε μέγεθος 64 εντολών, κάτι που οδηγούσε σε αποτυχία αν το πρόγραμμα δεν τερματιζόταν με επιτυχία με το τέλος της 64^{ης} εντολής. Δεύτερος περιορισμός ήταν ότι υπήρχε μόνο ένα πρόγραμμα, άρα υπήρχε μόνο ένας πληθυσμός που μόλις αναπτύχθηκε. Και τέλος, τρίτος περιορισμός ήταν το

γεγονός ότι η συνάρτηση που χρησιμοποιούταν δεν ήταν πολύ ξεκάθαρο ότι ήταν η κατάλληλη. Ο Γενετικός προγραμματισμός χρησιμοποιείται κυρίως για την εξέλιξη των προγραμμάτων. Τα προγράμματα αυτά αναπαρίστανται ως δομές δέντρων των οποίων οι εσωτερικοί κόμβοι παίρνουν στοιχεία από το σύνολο της συνάρτησης που χρησιμοποιείται και οι τερματικοί κόμβοι – φύλλα παίρνουν στοιχεία από το τερματικό σύνολο. Το σύνολο της συνάρτησης περιλαμβάνει διάφορες καταστάσεις και τελεστές, ενώ το τερματικό σύνολο περιλαμβάνει εισόδους, σταθερές και άλλες μεταβλητές που χρησιμοποιεί το πρόγραμμα. Σκοπός του γενετικού προγραμματισμού είναι να ανακαλύψει ένα πρόγραμμα μέσα στο χώρο των πιθανών προγραμμάτων που αναζητεί, το οποίο δίνει ένα επιθυμητό αποτέλεσμα για ένα σύνολο δεδομένων εισόδου.

Ο Γενετικός προγραμματισμός είναι ένα υποσύνολο των γενετικών αλγορίθμων που επεξεργάζεται εκτελέσιμα προγράμματα. Η διαφορά τους είναι ότι στον γενετικό προγραμματισμό τα μέλη του πληθυσμού είναι εκτελέσιμες δομές αντί για συμβολοσειρές από bits καθώς επίσης η καταλληλότητα του κάθε μέλους του πληθυσμού μετριέται εκτελώντας το, ενώ στους γενετικούς αλγόριθμους η καταλληλότητα εξαρτάται από το πρόβλημα.

Στην αρχή πρέπει να γίνουν τα πιο κάτω βήματα:

1. Καθορισμός του τερματικού συνόλου
2. Καθορισμός του συνόλου συνάρτησης
3. Καθορισμός μέτρου καταλληλότητας
4. Επιλογή παραμέτρων συστήματος
5. Καθορισμός τερματικών συνθηκών

Οι δύο κύριες παράμετροι του συστήματος είναι το μέγεθος του πληθυσμού και ο μέγιστος αριθμός γενιών που θα τρέξει το πρόγραμμα. Οι συνθήκες τερματισμού καθορίζονται κυρίως από τον αριθμό των γενιών που προσδιορίστηκαν και το πρόγραμμα που «κερδίζει» είναι το καλύτερο πρόγραμμα για την επίλυση του συγκεκριμένου προβλήματος.

Βήματα γενετικού προγραμματισμού:

1. Αρχικοποίηση πληθυσμού προγραμμάτων
2. Καθορισμός μέτρου καταλληλότητας για το κάθε πρόγραμμα
3. Αναπαραγωγή σύμφωνα με τις τιμές καταλληλότητας και τις πιθανότητες αναπαραγωγής
4. Εκτέλεση crossover
5. Πήγαινε στο βήμα 2 μέχρι την τερματική συνθήκη

2.4.4 Στρατηγικές Εξέλιξης

Η βασική αναπαράσταση των στρατηγικών εξέλιξης αναπαριστά το φαινότυπο ενός χρωμοσώματος αντί για το γονότυπο. Η αναπαράσταση αυτή είναι ένα διάνυσμα σταθερού μεγέθους από πραγματικές τιμές. Κάθε στοιχείου του διανύσματος αυτού έχει την πραγματική τιμή μιας μεταβλητής του προβλήματος. Οι στρατηγικές εξέλιξης βασίζονται στην εξέλιξη της εξέλιξης. Σκοπός είναι η μετακίνηση της μάζας του πληθυσμού προς την καλύτερη περιοχή του «τοπίου». Με βάση τον κανόνα «επιβίωση του ισχυρότερου» τα καλύτερα άτομα από κάθε γενιά μπορούν να αναπαράγουν.

Βήματα στρατηγικών εξέλιξης:

1. Αρχικοποίηση πληθυσμού
2. Εκτέλεση ανασυνδυασμού χρησιμοποιώντας μ γονείς για τη δημιουργία λ παιδιών
3. Εκτέλεση μετάλλαξης σε όλα τα παιδιά
4. Υπολογισμός νέου πληθυσμού
5. Επιλογή μ ατόμων από το νέο πληθυσμό
6. Πήγαινε στο βήμα 2 μέχρι την τερματική συνθήκη

2.4.5 Εξελικτικός προγραμματισμός

Ο εξελικτικός προγραμματισμός είναι ο ίδιος με τους γενετικούς αλγορίθμους στη χρήση ενός πληθυσμού από υποψήφιες λύσεις για να δώσουν μια απάντηση σε ένα συγκεκριμένο πρόβλημα. Σκοπός είναι να αναπτυχθούν μοντέλα συμπεριφοράς, μοντέλα που παρατηρούν αλληλεπιδράσεις του συστήματος με το περιβάλλον. Η

διαφορά με τους γενετικούς αλγόριθμους είναι ότι οι τελευταίοι εργάζονται στο χώρο του γονότυπου των πληροφοριών κωδικοποίησης ενώ ο εξελικτικός προγραμματισμός δίνει έμφαση στο χώρο του φαινοτύπου των παρατηρήσιμων συμπεριφορών.

Βήματα εξελικτικού προγραμματισμού:

1. Αρχικοποίηση πληθυσμού
2. Έκθεση τον πληθυσμό στο περιβάλλον
3. Υπολογισμός μέτρου καταλληλότητας για το κάθε μέλος
4. Τυχαία μετάλλαξη κάθε γονέα από τον πληθυσμό
5. Αξιολόγηση τους γονείς και τα παιδιά
6. Επιλογή μελών για τον νέο πληθυσμό
7. Πήγαινε στο βήμα 2 μέχρι την τερματική συνθήκη

2.3 Ασαφής Λογική

Η ασαφής λογική είναι ένα εργαλείο για την ενσωμάτωση δομημένης ανθρώπινης γνώσης σε εφαρμοσμένους αλγόριθμους (Kecman,2001). Είναι μια επέκταση της κλασσικής αριστοτέλειας λογικής, όπου οι λογικές προτάσεις δεν έχουν απόλυτες τιμές αλήθειας ή ψεύδους. Μια πρόταση μπορεί να είναι αληθής με κάποιο βαθμό αληθείας, και όχι απλά αληθής ή ψευδής.

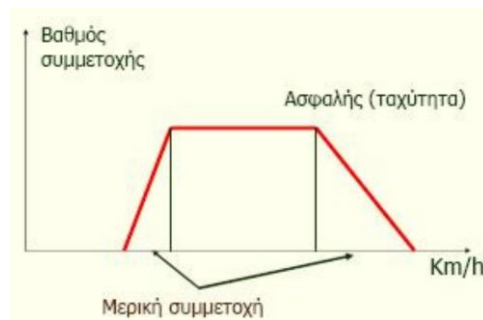
Είναι ένα υπερσύνολο της λογικής Boolean και επινοήθηκε για να λογικοποιήσει τις μισό-αλήθειες και τα παράδοξα. Είναι εύκολη στην κατανόηση, ταιριάζει στην ανθρώπινη γλώσσα και στον ανθρώπινο τρόπο σκέψης και είναι ανεκτική σε μη ακριβή δεδομένα. Δίνει τη δυνατότητα να υπάρξουν και άλλες τιμές σαν απάντηση στα προβλήματα. Όπως για παράδειγμα στο Σχήμα 2.11, στην ερώτηση πότε είναι κάποιος ψηλός, μπορεί να υπάρξουν περισσότερες από 2 απαντήσεις, σύμφωνα με την συνάρτηση συμμετοχής (**membership function**).

$$\mu(h) = \begin{cases} 0 & \rightarrow \text{εάν } h < 1,5\text{m} \\ \frac{(h-1,5)}{0,7} & \rightarrow \text{εάν } 1,5 \leq h \leq 2,2\text{m} \\ 1 & \rightarrow \text{εάν } h > 2,2\text{m} \end{cases}$$

Σχήμα 2.11: Παράδειγμα συνάρτησης συμμετοχής

Συνάρτηση Συμμετοχής (Membership function):

Η συνάρτηση συμμετοχής είναι η χαρακτηριστική «ταυτότητα» κάθε ασαφούς συνόλου. Εισάγεται η ιδέα της μερικής συμμετοχής ενός στοιχείου σε κάποια έννοια. Κάθε γλωσσική τιμή μιας γλωσσικής μεταβλητής έχει ανατεθεί σε μια συνάρτηση συμμετοχής γι' αυτό και τα κριτήρια επιλογής της συνάρτησης αυτής εξαρτώνται από το πρόβλημα. Όσο μεγαλύτερη είναι η τιμή συμμετοχής $A(x)$ τόσο περισσότερο τυπικά το x ανήκει στο A . Μια συνάρτηση συμμετοχής φαίνεται στο σχήμα 2.12.

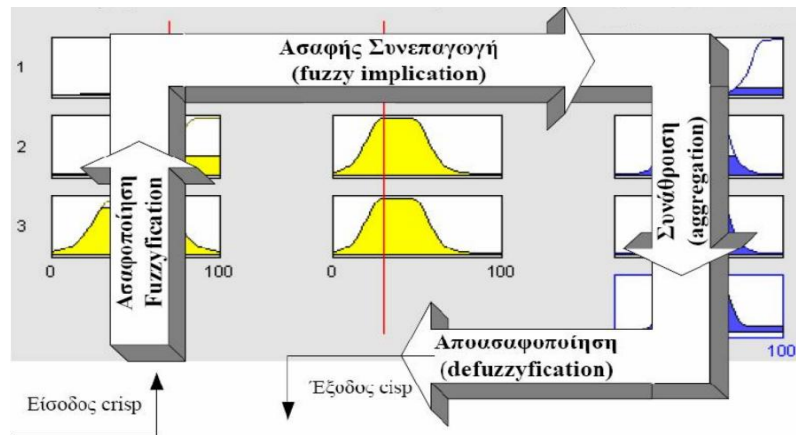


Σχήμα 2.12: Παράδειγμα συνάρτησης συμμετοχής

Ασαφείς κανόνες:

Ένας ασαφής κανόνας είναι ένας μηχανισμός αναπαράστασης της γνώσης, ο οποίος προσιδιάζει στον ανθρώπινο τρόπο σκέψης. Τα ασαφή σύνολα που εκφράζουν λεκτικούς όρους συνδυάζονται μεταξύ τους και δημιουργούν τους ασαφείς κανόνες. Ο τρόπος λειτουργίας του κανόνα διασπάται σε τρεις φάσεις. Τη φάση ασαφοποίησης, τη φάση ασαφής συνεπαγωγής και τη φάση αποασαφοποίησης. Οι τρεις φάσεις φαίνονται στο σχήμα 2.13. Ακολουθούμε τα εξής βήματα:

1. Λεκτικός διαμερισμός των εισόδων (Αναπαριστούμε τις μεταβλητές εισόδου και εξόδου με λεκτικούς όρους.)
2. Διατύπωση κανόνων
3. Καθορισμός τύπου τα ασαφούς συνεπαγωγής
4. Μέθοδος αποασαφοποίησης (Μια απεικόνιση που δείχνει ένα ασαφές σύνολο σε ένα πραγματικό αριθμό.)



Σχήμα 2.13: Φάσεις διάσπασης του κανόνα

Κεφάλαιο 3

Προηγούμενες Μελέτες

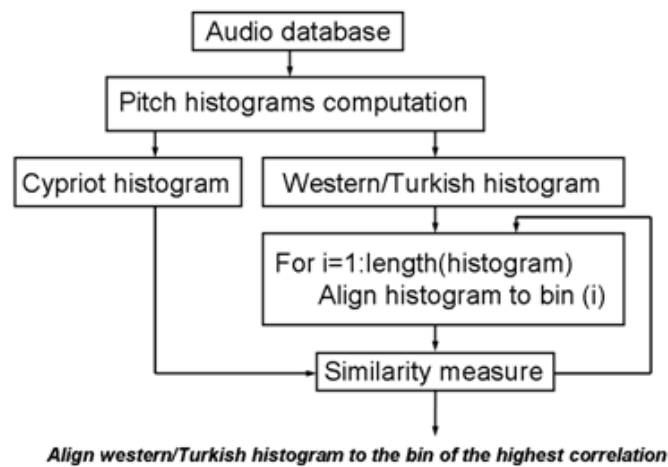
3.1 Εργασίες στο Πανεπιστήμιο Κύπρου	24
3.2 Εργασίες σε Διεθνές επίπεδο	26

3.1 Εργασίες στο Πανεπιστήμιο Κύπρου (papers)

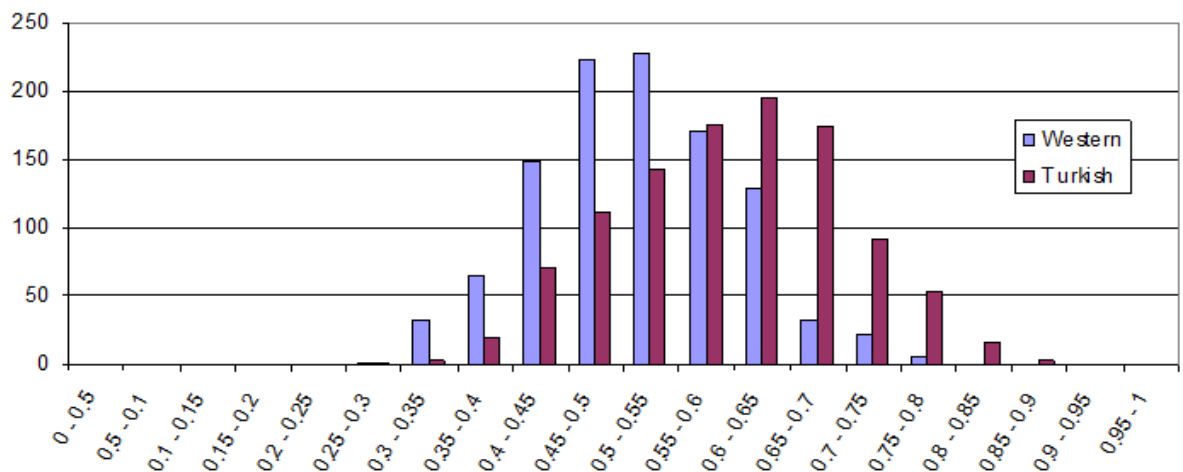
Στο Πανεπιστήμιο Κύπρου, έχουν γίνει πρόσφατα μελέτες σχετικές με το θέμα της εργασίας αυτής. Οι μελέτες αυτές πραγματοποιήθηκαν από τους Νεοκλέους Α., Παντελή Μ., Σχίζα Χ., και Petkov N. [12,13].

Στην πρώτη μελέτη [12], ο στόχος ήταν να διερευνήσουν το ενδεχόμενο ομοιότητας μεταξύ Κυπριακής παραδοσιακής μουσικής και της Τούρκικης μουσικής όσον αφορά την τονικότητα. Χρησιμοποίησαν μια βάση δεδομένων με 126 μονοφωνικά τραγούδια, εκ των οποίων τα 20 ήταν Κυπριακά, τα 53 ήταν Τούρκικα και τα υπόλοιπα 53 ήταν Δυτικά τραγούδια και σόλο αυτοσχεδιασμοί με διάφορα όργανα όπως φλάουτο, κλαρινέτο, σαξόφωνο και άλλα. Έγινε κατάτμηση των τραγουδιών σε πολλά πλαίσια και στο καθένα από αυτά εφαρμόστηκε ο Yin αλγόριθμος. Εξέτασαν ένα τμήμα των 120 πρώτων δευτερολέπτων κάθε τραγουδιού καθώς επίσης και 126 διανύσματα με τις εκτιμήσεις των συχνοτήτων τους. Στη συνέχεια, δημιουργήθηκαν τονικά ιστογράμματα και κάθε Κυπριακό ιστόγραμμα συγκρίθηκε με όλα τα Δυτικά και όλα τα Τούρκικα, δημιουργώντας έτσι ένα σύνολο από 2120 συγκρίσεις. Βρήκαν τους συντελεστές συσχέτισης μεταξύ των Κυπριακών ιστογραμμάτων και των Τούρκικων καθώς επίσης και μεταξύ των Κυπριακών ιστογραμμάτων και των Δυτικών. Δημιούργησαν τονικά ιστογράμματα αφού είναι ένα από τα κυριότερα χαρακτηριστικά για τονικές ομοιότητες μεταξύ μονοφωνικών τραγουδιών. Οι κορυφές ενός τέτοιου ιστογράμματος δείχνουν τις πιο συχνές νότες που παίζονται σε ένα τραγούδι. Η διαδικασία εφαρμογής του μέτρου ομοιότητας σε ιστογράμματα φαίνεται στο σχήμα 3.1. Τα αποτελέσματα έδειξαν ότι το

2.6% των συγκρίσεων μεταξύ Κυπριακών και Δυτικών τραγουδιών ήταν πολύ όμοια, ενώ το 9.3% πολύ ανόμοια, και το 55% όμοια. Το 80% των συγκρίσεων μεταξύ Κυπριακών και Τούρκικων τραγουδιών ήταν όμοια. Από τα πιο πάνω αποτελέσματα, παρατηρήθηκε ότι η υπόθεση ότι η Κυπριακή μουσική συμερίζεται ομοιότητες με την Τούρκικη μουσική, εν μέρει επιβεβαιώθηκε. Στο σχήμα 3.2 παρουσιάζονται τα αποτελέσματα όπου, τιμές στον άξονα X, που είναι πάνω από 0.5, θεωρούνται όμοιες, τιμές που είναι πάνω από 0.7 θεωρούνται πολύ όμοιες και τιμές κάτω από το 0.3 θεωρούνται πολύ ανόμοιες [12].



Σχήμα 3.1: Διαδικασία εφαρμογής μέτρου ομοιότητας μεταξύ ιστογραμμάτων



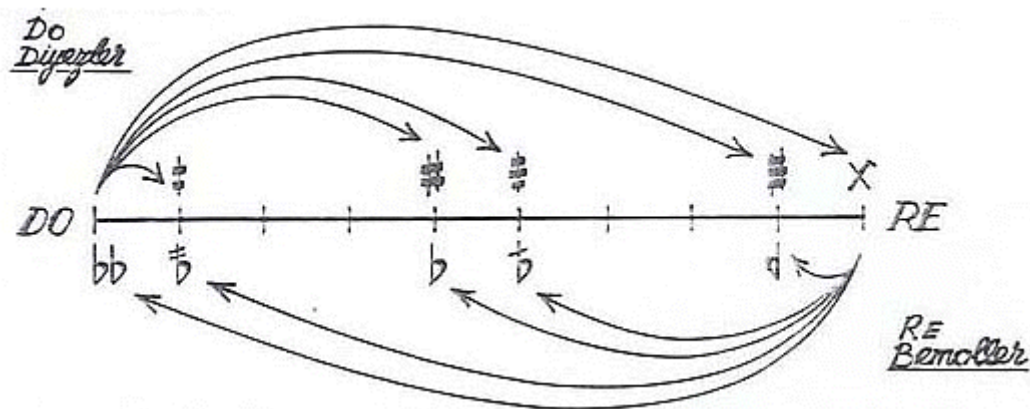
Σχήμα 3.2: Η κατανομή της εξόδου μεταξύ των κυπριακών τραγουδιών και των δυτικών τραγουδιών φαίνονται με μπλε χρώμα ενώ η κατανομή της εξόδου μεταξύ των κυπριακών τραγουδιών και των τούρκικων τραγουδιών φαίνονται με μωβ χρώμα.

Στην δεύτερη μελέτη [13], στόχος ήταν η μελέτη του τόνου διάφορων οργάνων πιθκιάβλι κάνοντας αναφορά στην κατασκευή τους. Για τους σκοπούς της μελέτης αυτής, ο Ανδρέας Γριστάκκος, ένας από τους λίγους κατασκευαστές πιθκιάβλι, ηχογραφήθηκε καθώς έπαιζε διάφορα είδη πιθκιάβλι. Ανάφερε 5 τύπους οργάνων πιθκιάβλι που χαρακτηρίζονται από μια συγκεκριμένη ποιότητα ήχου. Οι τύποι αυτοί ξεκινούν από βαθύ και τραχύ ήχο σε ψηλο-τονικούς ήχους με ονόματα hontro, kabatiko, mezikon, psilari, tsiripilli. Για υπολογιστική ανάλυση του τόνου όμως, οι μελωδίες αναπαρίστανται με τους πρώτους 4 τύπους πιθκιάβλι οι οποίες ηχογραφήθηκαν. Με την χρήση του αλγόριθμου Yin [6], υπολογίστηκε η θεμελιώδης συχνότητα και δημιουργήθηκαν ιστογράμματα με τη χρήση Γκαουσιανού πυρήνα. Από τα ιστογράμματα, εξάχθηκαν οι πιο σημαντικές νότες της μελωδίας και από αυτές υπολογίστηκαν τα διαστήματα μεταξύ τους. Από τα αποτελέσματα που πήραν, αποκάλυψαν ότι ο τόνος και το φάσμα συχνότητας διαφέρουν σημαντικά μεταξύ των διάφορων τύπων πιθκιάβλι. Αυτή η έρευνα είναι ένα προκαταρκτικό βήμα στην ανάλυση του τόνου του πιθκιάβλι με σκοπό την περεταίρω μελέτη του στο μέλλον [13].

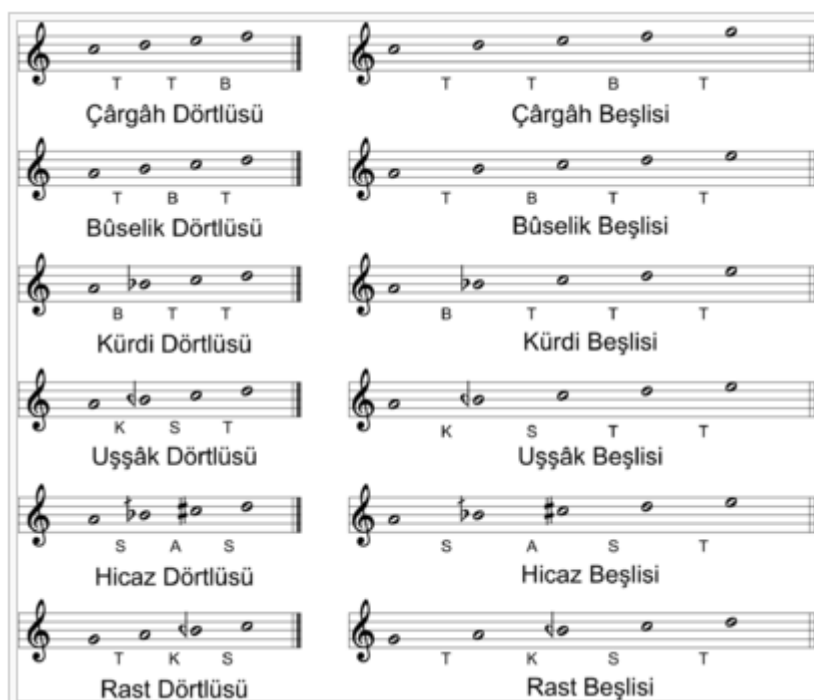
3.2 Εργασίες σε διεθνές επίπεδο

Σε διεθνές επίπεδο, έχουν γίνει αρκετές μελέτες σχετικές με το θέμα της εργασίας αυτής. Μερικές από τις μελέτες αυτές πραγματοποιήθηκαν από τους Adil Alpkocak, Alin Cenk Gedik, Baris Bozkurt, Sertan Senturk, Maria Panteli. [1,4,14,10]

Στην πρώτη μελέτη [1], στόχος των Adil Alpkocak και Alin Cenk Gedik ήταν η αναζήτηση χαρακτηριστικών των Τούρκικων makams τα οποία ορίζονται έμμεσα στη θεωρία. Στην Τούρκικη μουσική σύμφωνα με τη θεωρία Arel [10], η οκτάβα διαιρείται σε 53 ίσα τμήματα με ένα συνολικό τόνο που αναπαρίσταται από 9 τέτοια τμήματα. Χρησιμοποιούνται επίσης 8 τέτοια τμήματα, 5 ή και 4. Τα makam είναι διάφοροι συνδυασμοί από φθόγγους με συχνότητες παρμένες από τα 8 ή 9 τμήματα που είναι διερεμένος ένας τόνος. Ένας πιθανός παραλληλισμός στην δυτική μουσική θα μπορούσαν να ήταν οι κλίμακες. Στο σχήμα 3.3 φαίνονται οι πιο βασικές νότες και στο σχήμα 3.4 φαίνεται πως είναι χτισμένο ένα makam.



Σχήμα 3.3: Οι βασικές νότες σύμφωνα με το σύστημα Σολφέζ, για παράδειγμα, "Do" είναι C και το "Re" είναι D.



Σχήμα 3.4: Ένα makam στην τουρκική μουσική είναι χτισμένο από ένα τετράχορδο χτισμένο στην κορυφή ενός πεντάχορδου, ή το αντίστροφο. Φαίνονται τα 6 διαφορετικά τετράχορδα και πεντάχορδα έχουν το ίδιο όνομα.

Στην πρώτη απόπειρα αυτόματης ταξινόμησης των Τούρκικων makams, επιλέχθηκαν 10 makams σύμφωνα με την δημοτικότητα, την ομοιότητα και την ανομοιότητα τους για την επίτευξη μιας ισχυρής κατηγοριοποίησης. Το σύνολο των δεδομένων που χρησιμοποιήθηκαν αποτελείται από 200 MIDI αρχεία, 20 για κάθε 10 makam και η

κατηγοριοποίηση εφαρμόζεται μετά από απαιτούμενη προεπεξεργασία. Στην προεπεξεργασία, κάθε makam αναπαρίσταται ως συμβολοσειρά από νότες. Αφού γίνει η προεπεξεργασία, είναι εύκολο να χρησιμοποιηθεί ανάλυση N-gram για την εξαγωγή χαρακτηριστικών. Η ανάλυση N-gram είναι μια συνεχόμενη ακολουθία από n αντικείμενα μιας δεδομένης σειράς κειμένου ή ομιλίας. Μπορεί να είναι ένας οποιοσδήποτε συνδυασμός γραμμάτων τα οποία μπορεί να είναι φθόγγοι, συλλαβές, γράμματα, λέξεις ανάλογα με την κάθε εφαρμογή. Δημιούργησαν ένα ιεραρχικό ταξινομητή που αποτελείτο από ένα γραμμικό διαχωριστικό ταξινομητή και δέντρα απόφασης. Η μέση απόδοση του ταξινομητή τους ήταν 88% για ευαισθησία, 99% για επιλεκτικότητα και 98% για ακρίβεια το οποίο δείχνει υψηλή επιτυχία κατηγοριοποίησης των makams.

Στην δεύτερη μελέτη [4], οι Alin Cenk Gedik και Baris Bozkurt αξιολόγησαν την επικρατέστερη θεωρία της TTAM (Traditional Turkish Art Music) την Arel θεωρία, μέσω μιας διαδικασίας κατηγοριοποίησης makam, έτσι ώστε να καταλάβουν αν μπορεί να προσφέρει μια βάση για μελέτες MIR (Music Information Retrieval) στην TTAM, με παρόμοιο τρόπο όπως η θεωρία δυτικής μουσικής που παρέχει μια βάση για μελέτες MIR στην δυτική μουσική. Οι MIR μέθοδοι είναι ειδικά προσαρμοσμένες για τις ανάγκες της δυτικής μουσικής, εντούτοις όμως στην έρευνα αυτή η μελέτη της TTAM μπορεί να θεωρηθεί ως μια προσπάθεια παροχής συνεργασίας μεταξύ μεθόδων MIR και εθνομουσικολογίας. Η μελέτη αυτή βασίζεται στις πληροφορίες κατανομής του τονικού χώρου στην TTAM καθώς επίσης οι ερευνητές εδώ περιορίστηκαν σε 9 makams τα οποία αναπαριστούν το 50% του τρέχον ρεπερτορίου. Παρόλο των διαφορών στον τονικό χώρο μεταξύ δυτικής μουσικής και TTAM, εντούτοις και τα πρότυπα και οι ηχογραφήσεις αναπαρίστανται ως ιστογράμματα τονικό-συχνοτήτων, σε αντίθεση με την κατανομή τόνου στην δυτική μουσική. Μια νέα μέθοδος για την ανάλυση των τονικών συχνοτήτων, ήταν η ανάλυση των ηχογραφήσεων με τον αλγόριθμο YIN και οι θεμελιώδεις συχνότητες προεπεξεργάζονται με φίλτρα. Τα φίλτρα αυτά βασίζονται στα ακουστικά τους χαρακτηριστικά. Στη συνέχεια εφαρμόζεται ο αλγόριθμος τονικής ανίχνευσης και γίνεται η καταγραφή των αποτελεσμάτων. Υπολογίζονται οι κατανομές συχνότητας και στο τέλος κάθε ηχογράφισης αναπαρίσταται ως ένα ιστόγραμμα τονικό-συχνοτήτων. Στη συνέχεια σχεδίασαν ένα ταξινομητή, όπου κάθε δεδομένο ανήκει στην ίδια του την κλάση, άρα το makam κάθε ηχογράφησης είναι γνωστό. Ο ταξινομητής

μπορεί να κατηγοριοποιεί θετικά και αρνητικά δείγματα από το σωστό ή λάθος ρυθμό κατηγοριοποίησης. Η θεωρία Arel είναι επιτυχής όταν εφαρμόζεται σε πλαίσιο MIR για TTAM. Έτσι, οδηγήθηκαν στο συμπέρασμα ότι η θεωρία Arel με μερικές βελτιώσεις παρέχει έγκυρους «ορισμούς» για μελέτες MIR στην TTAM, όπως γίνεται και στη δυτική μουσική.

Σε μια άλλη μελέτη [14], στόχος του Sertan Senturk ήταν η πρόβλεψη της μελωδίας της μορφής *uzun hava*. Στη μελέτη αυτή, περιγράφεται μια βάση δεδομένων *uzun hava*, μιας μη μετρήσιμης δομημένης μορφής αυτοσχεδίασης και ενός συστήματος που χρησιμοποιεί Markov μοντέλα μεταβλητού μήκους (VLMM). Χρησιμοποίησε μια βάση δεδομένων με 77 τραγούδια, που περιλαμβάνει 10849 νότες και χρησιμοποιείται για την εκπαίδευση πολλών πολλαπλών σημείων. Στα πειράματα, τα αποτελέσματα από την τονική συσχέτιση που αντιστοιχίζει 12 κλίμακες τόνου δυτικής κλασσικής θεωρίας και 17 κλίμακες τόνου Τούρκικης μουσικής συγκρίνονται. Έδειξαν ότι τα VLMM είναι πολύ προβλέψιμα στην εξέλιξη των νότων.

Τέλος, σε μια άλλη μελέτη [10], στόχος της Μαρίας Παντελή ήταν η διερεύνηση της Βυζαντινής και Οθωμανικής μουσικής επιρροής, όπως παρατηρείται από τα τονικά μοτίβα. Η έννοια των τονικών μοτίβων εδώ περιλαμβάνει τον τόνο της κλίμακας, τη χρήση των διαστημάτων και τη σημασία της κλίμακας βαθμού. Εφαρμόστηκε μια υπολογιστική συγκριτική ανάλυση σε μελωδίες Κυπριακής μουσικής και κυρίως σε θρησκευτική Βυζαντινή και Οθωμανική μουσική. Η μεθοδολογία αυτή βασίστηκε στη χρήση των τονικών κλάσεων και είναι σχεδιασμένη για να ενσωματώνει τις ιδιαιτερότητες των παραδοσιακών αυτών μουσικών. Νέοι αλγόριθμοι χρησιμοποιήθηκαν για τον υπολογισμό του τονικού ιστογράμματος. Τα αποτελέσματα έδειξαν ότι η χρήση των διαστημάτων στην Κυπριακή μουσική ήταν εξίσου παρόμοια και με τις δύο άλλες μουσικές, ενώ η ανάδειξη της κλίμακας βαθμού έδειξε να μοιράζεται κάποιες τάσεις με την Βυζαντινή μουσική. Επίσης, η επιρροή στις τονικές κλίμακες έδειξε περισσότερη ομοιότητα με τη Βυζαντινή μουσική παρά με την Οθωμανική.

Κεφάλαιο 4

Ανάλυση θέματος και εξήγηση

4.1 Ορισμός θέματος / προβλήματος	30
4.2 Ερευνητική ερώτηση	31
4.3 Ανάλυση	31
4.4 Παραδοσιακές μέθοδοι αντιμετώπισης	39
4.5 Σύγκριση και συζήτηση μεθόδων	39

4.1 Ορισμός θέματος / προβλήματος

Σκοπός της διπλωματικής μου εργασίας είναι η ανάλυση της Κυπριακής παραδοσιακής μουσικής με σκοπό την μελέτη πιθανών διαφορών ή ομοιοτήτων με διάφορες μουσικές της Μεσογείου.

Από την ανάλυση εξάγονται σημαντικές πληροφορίες τόσο για τη μελωδία όσο και για το ρυθμό και άλλα χαρακτηριστικά της μουσικής. Οι πληροφορίες αυτές χρησιμοποιήθηκαν για τη δημιουργία ενός συστήματος το οποίο χαρακτηρίζει τη μουσική και την τοποθετεί σε διάφορες κατηγορίες. Το θέμα αυτό μπορεί να επηρεάσει και να βοηθήσει διάφορους κλάδους της επιστήμης όπως τη μουσικολογία, την εθνομουσικολογία, και την επιστήμη των ηλεκτρονικών υπολογιστών, αφού θα κατανοήσουν καλύτερα τα χαρακτηριστικά της Κυπριακής παραδοσιακής μουσικής και πιθανόν να καταφέρουν να προσδιορίσουν και τις πηγές που την επηρεάζουν.

Κύριος στόχος της εργασίας αυτής είναι να διερευνηθούν τονικές κυρίως ομοιότητες μεταξύ Κυπριακών δημοτικών τραγουδιών, σε σύγκριση με Τούρκικα και Δυτικά μονοφωνικά τραγούδια. Έγιναν αρκετές μελέτες και εντοπίστηκαν ομοιότητες μεταξύ Κυπριακής μουσικής και Βυζαντινής μουσικής [8,16]. Εντούτοις όμως πολύ λίγες

πληροφορίες υπάρχουν σχετικά με το αν επηρεάζει ή όχι η ανατολική μουσική, όπως αραβική ή τούρκικη, την κυπριακή λαϊκή μουσική.

Εργασίες στο θέμα αυτό έχουν γίνει από αρκετούς ερευνητές, αλλά με τη μεθοδολογία που ακολουθήθηκε στην παρούσα εργασία, δεν έχει γίνει άλλη εργασία για Κυπριακή μουσική. Έχουν γίνει όμως για άλλες μουσικές όπως για παράδειγμα Βυζαντινή μουσική, Ελληνική μουσική, Τούρκικη, και άλλες.

4.2 Ερευνητική ερώτηση

Ένα από τα κυριότερα ερωτήματα για αυτή την έρευνα είναι κατά πόσον η παραδοσιακή μουσική της Κύπρου είναι επηρεασμένη από άλλες μουσικές παραδόσεις όπως η Ελληνική, Τούρκικη, Βυζαντινή και άλλες μουσικές της Ανατολής. Υπάρχουν διάφοροι τρόποι για να διερευνηθεί αυτό το ερώτημα. Ένας τρόπος είναι να μελετηθούν όλες οι ηχογραφήσεις και οι μουσικές παρτιτούρες από ειδικούς εθνομουσικολόγους και να εξαχθούν συμπεράσματα. Ένα από τα μεγαλύτερα προβλήματα στην εθνομουσικολογία είναι το γεγονός ότι στις παραδοσιακές μουσικές – με κάποιες εξαιρέσεις – δεν υπάρχουν μουσικές παρτιτούρες, και η παράδοση μεταδίδεται από γενιά σε γενιά. Έτσι, ένας μουσικολόγος δεν μπορεί με ευκολία να βγάλει συμπεράσματα από τις μελωδίες, τους ρυθμούς κλπ και κατ' επέκταση, δεν μπορούν τα συμπεράσματα να αποδειχθούν και να καταγραφούν επιστημονικά χωρίς να υπάρχει μια συγκεκριμένη και αποδεκτή μεθοδολογία. Με την ανάλυση του ηχογραφημένου υλικού με υπολογιστικές μεθόδους, το ερώτημα αυτό μπορεί να απαντηθεί χρησιμοποιώντας και δημιουργώντας υπολογιστικά μοντέλα βασισμένα σε μουσικά χαρακτηριστικά όπως είναι η συχνότητα της νότας, η διάρκεια, η κλίμακα κλπ.

4.3 Ανάλυση

Χαρακτηριστικά χροιάς:

Για την ανάλυση της Κυπριακής παραδοσιακής μουσικής με στόχο την μελέτη πιθανών διαφορών ή ομοιοτήτων με διάφορες μουσικές της μεσογείου, είναι αναγκαία η εξαγωγή χαρακτηριστικών γνωρισμάτων των συγκεκριμένων κομματιών μουσικής. Με την εξαγωγή των χαρακτηριστικών αυτών, θα είναι πιο εφικτός ο χαρακτηρισμός και η

κατηγοριοποίηση της κάθε μουσικής σε διάφορες κατηγορίες οι οποίες θα έχουν μουσικό νόημα.

Mel Frequency Cepstral Coefficients

Τα Mel Frequency Cepstral Coefficients (mfcc) χρησιμοποιούνται ως χαρακτηριστικά σε συστήματα αναγνώρισης ομιλίας, όπως συστήματα τα οποία μπορούν να αναγνωρίσουν αυτόματα αριθμούς που λέγονται σε ένα τηλέφωνο, καθώς επίσης χρησιμοποιούνται και για την ανάκτηση πληροφοριών της μουσικής για την μετέπειτα κατηγοριοποίηση της σε διάφορες ομάδες.

Για να υπολογιστούν τα mfcc, πρώτα το ηχητικό σήμα χωρίζεται σε μικρά κομμάτια (παράθυρα) και σε κάθε κομμάτι πολλαπλασιάζεται μια συνάρτηση (συνήθως συνημιτονοειδής). Ο λόγος που πολλαπλασιάζεται η συνάρτηση είναι για την μείωση των αποτελεσμάτων ασυνέχειας στα άκρα του ηχητικού σήματος(σε κάθε κομμάτι). Με αυτό τον τρόπο μειώνεται ο θόρυβος στο ηχητικό σήμα και ισοδυναμεί με τη συνέλιξη στο πεδίο της συχνότητας του Fourier μετασχηματισμού των σημάτων.

Οι δύο συναρτήσεις που χρησιμοποιούνται πιο συχνά είναι οι Hanning και Hamming. Η συνάρτηση Hanning δίνεται από την σχέση (1) ενώ η συνάρτηση Hamming από τη σχέση (2).

$$w(n) = 0.5(1 - \cos(\frac{2\pi n}{N-1})) \quad (1)$$

$$w(n) = \alpha - \beta \cos(\frac{2\pi n}{N-1}) \quad (2)$$

Όπου:

n : είναι ο αριθμός δείγματος (χρόνος)

N : είναι το σύνολο των δειγμάτων του παραθύρου

$\alpha = 0.54$

$\beta = 1 - \alpha = 0.46$

Στη συνέχεια υπολογίζεται το συχνοτικό φάσμα με τη χρήση γρήγορου μετασχηματισμού Fourier (3) και ακολούθως μετατρέπεται το φάσμα των συχνοτήτων από γραμμικό σε μπάντες Mel. Το πλάτος συχνότητας και η φάση συχνότητας φαίνονται στη σχέση (4) και στη σχέση (5). Η σχέση που συνδέει το γραμμικό φάσμα και τις μπάντες Mel δείχνεται στην σχέση (6). Η μετατροπή αυτή γίνεται επειδή η απόκριση του ανθρώπινου ακουστικού συστήματος είναι διαφορετική και δεν είναι γραμμική όπως για παράδειγμα ο διαχωρισμός του συχνοτικού φάσματος σε οκτάβες.

$$\chi_k = \sum_{n=0}^{N-1} x_n e^{\frac{-2i\pi k}{Nn}} \quad (3)$$

$$\frac{|x_k|}{N} = \frac{\sqrt{\text{Re}(x_k)^2 + \text{Im}(x_k)^2}}{N} \quad (4)$$

$$\arg(x_k) = a * \tan 2(\text{Im}(x_k), \text{Re}(x_k)) \quad (5)$$

Όπου:

n : είναι ο αριθμός δείγματος χρόνου

k : είναι ο αριθμός δείγματος συχνότητας

N : είναι το σύνολο των δειγμάτων του παραθύρου

Re : το πραγματικό μέρος του μετασχηματισμού

Im : το φανταστικό μέρος του μετασχηματισμού

$$m = 2595 \log_{10} \left(1 + \frac{f}{700} \right) \quad (6)$$

Όπου:

f : γραμμικό φάσμα συχνότητας

Στην συνέχεια, κάθε μπάντα Mel πολλαπλασιάζεται με ένα σύνολο τριγωνικών φίλτρων που φαίνεται στη σχέση (7), και για κάθε μπάντα εφαρμόζεται διακριτός μετασχηματισμός συνημίτονου (DCT), που δείχνεται στη σχέση (8) και η τιμή που εξάγεται από τον μετασχηματισμό του συνημίτονου είναι ο συντελεστής MFCC (σχήμα 4.1).

$$w(n) = 1 - \left| \frac{n - \frac{N-1}{2}}{\frac{N+1}{2}} \right| \quad (7)$$

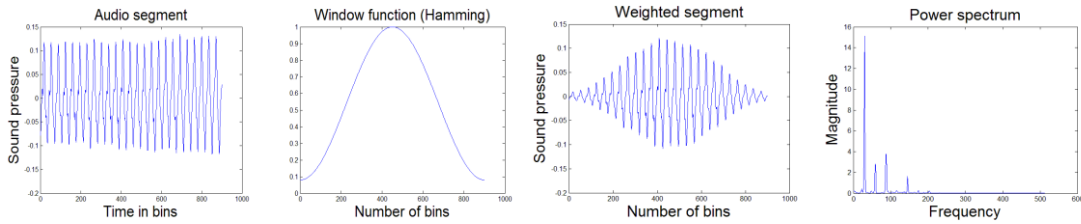
$$x_k = \frac{1}{2} (x_0 + (-1)^k x_{N-1}) + \sum_{n=1}^{N-2} x_n \cos \left[\frac{\pi n k}{N-1} \right] \quad (8)$$

Όπου:

n : είναι ο αριθμός δείγματος (χρόνος)

N : είναι το σύνολο των δειγμάτων του παραθύρου

$k = 0, 1, \dots, N-1$



Σχήμα 4.1: Στο πρώτο σχήμα φαίνεται ένα κομμάτι (παράθυρο) που παίρνουμε όταν το ηχητικό σήμα χωριστεί σε πολλά παράθυρα. Στο δεύτερο σχήμα, φαίνεται η συνάρτηση που θα πολλαπλασιαστεί με το κάθε κομμάτι και το αποτέλεσμα παρουσιάζεται στο τρίτο σχήμα. Τέλος, φαίνεται το συχνοτικό φάσμα που πάρθηκε με τη χρήση γρήγορου μετασχηματισμού Fourier.

Zero Crossing Rate

Το χαρακτηριστικό αυτό είναι ο μέσος αριθμός των φορών που το σήμα του ήχου διασχίζει τη γραμμή του μηδενικού πλάτους. Είναι δηλαδή ο ρυθμός με τον οποίο το σήμα του ήχου αλλάζει από θετικό σε αρνητικό και αντίθετα. Δίνεται από την σχέση (9).

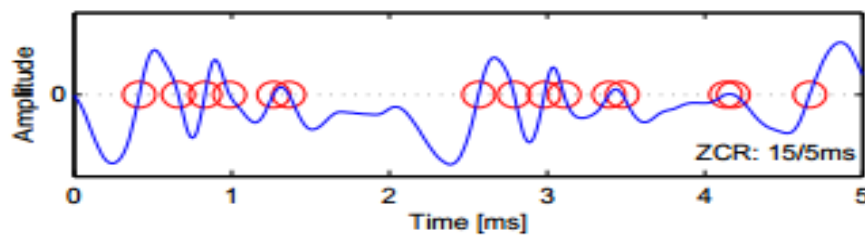
$$Zcr = \frac{1}{T-1} \sum_{t=1}^{T-1} \Pi \{s_t s_{t-1} < 0\} \quad (9)$$

Όπου:

S : είναι το ηχητικό σήμα μεγέθους T

$\Pi \{A\}$: η συνάρτηση, η οποία είναι 1 αν το A είναι true και 0 αν δεν είναι

Το zero crossing rate είναι ένα χαρακτηριστικό που χρησιμοποιείται για την εύρεση ομοιοτήτων σε διάφορα κομμάτια μουσικής. Εφαρμόζεται στην επεξεργασία ομιλίας για να διακριθούν τα κομμάτια ήχου από το θόρυβο. Επιπρόσθετα, χρησιμοποιείται για την κατηγοριοποίηση της μουσικής σε διάφορα είδη και ομάδες. Δεν μετريέται μόνο θόρυβος αλλά και τόνος (pitch). Όπως φαίνεται στο σχήμα 4.2, με μπλε χρώμα αναπαρίσταται το ηχητικό σήμα, και οι κόκκινοι κύκλοι αναπαριστούν το σημείο που το ηχητικό σήμα περνά τη γραμμή του μηδενικού πλάτους.

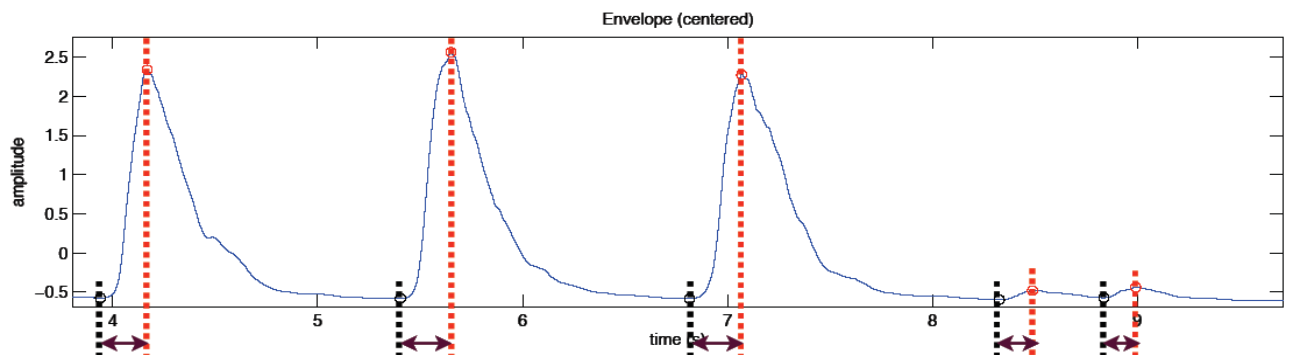


Σχήμα 4.2: Αναπαράσταση zero crossing rate με κομμάτι ήχου 5ms.

Attacktime

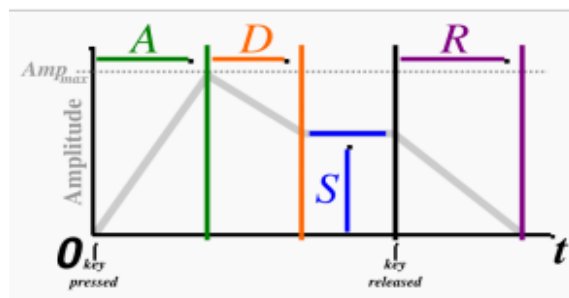
Attack-Decay-Sustain-Release (ADSR)

Το Attacktime είναι ένα χαρακτηριστικό το οποίο μπορεί να δώσει κάποιες πληροφορίες όσον αφορά τη χροιά. Ένας τρόπος για την περιγραφή της ατάκας είναι να υπολογιστεί η χρονική της διάρκεια. Η διάρκεια της ατάκας φαίνεται στο σχήμα 4.3, η απόσταση μεταξύ της μαύρης διακεκομμένης γραμμής και της κόκκινης.



Σχήμα 4. 3: Διάρκεια ατάκας

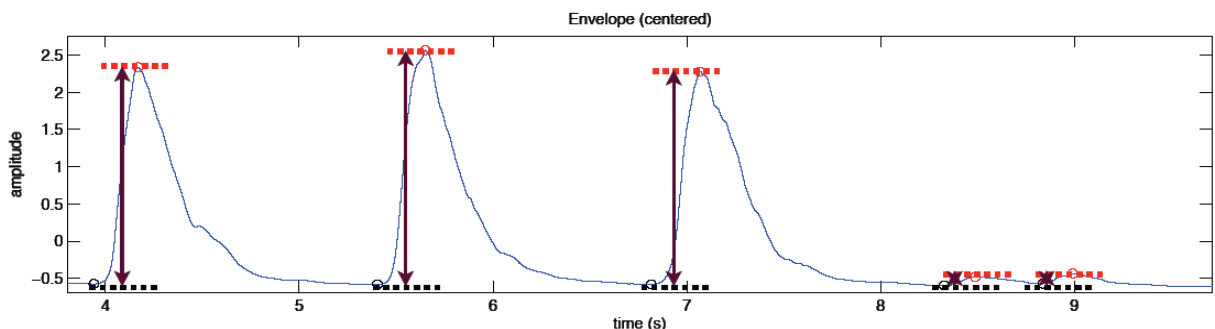
Για να γίνει πιο κατανοητός ο όρος της ατάκας, κρίνεται αναγκαίο να εξηγηθεί πως έχει μοντελοποιηθεί το περίβλημα του πλάτους μιας νότας. Το ADSR (σχήμα 4.4) περιγράφει χοντρικά το περίβλημα της έντασης μιας νότας και προέρχεται από τις λέξεις Attack-Decay-Sustain-Release. Η ατάκα (attack) δείχνει το πόσο γρήγορα φτάνει ο ήχος σε πλήρη ένταση μετά που ο ήχος είναι ενεργοποιημένος (για παράδειγμα μετά το πάτημα ενός πλήκτρου). Το decay δείχνει το πόσο γρήγορα πέφτει ο ήχος στο επίπεδο «διατήρησης» (sustain) μετά την αρχική κορυφή. Το sustain δείχνει την σταθερή ένταση που παίρνει ο ήχος μετά το decay και τέλος το Release δείχνει το πόσο γρήγορα ο ήχος εξασθενεί (όταν το πλήκτρο ελευθερώνεται).



Σχήμα 4.4: Attack-Decay-Sustain-Release (ADSR)

Attackleap

Ακόμα ένα χαρακτηριστικό που μπορεί να περιγράψει την ατάκα, είναι ο υπολογισμός της διαφοράς πλάτους μεταξύ της αρχής και του τέλους της φάσης ατάκας. Ο υπολογισμός αυτός φαίνεται στο σχήμα 4.5.

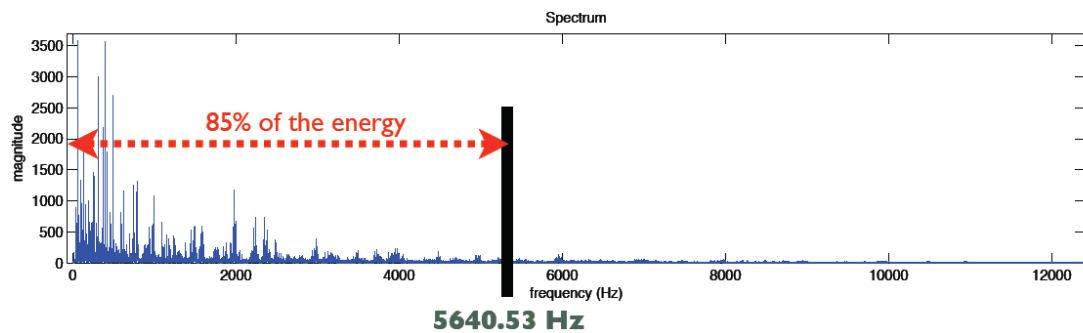


Σχήμα 4.5: Υπολογισμός διαφοράς πλάτους μεταξύ αρχής και τέλους της ατάκας

Rolloff

Για τον υπολογισμό του ποσού ψηλής συχνότητας, πρέπει να βρεθεί η συχνότητα κατά την οποία ένα κλάσμα της συνολικής ενέργειας περιέχεται κάτω από την

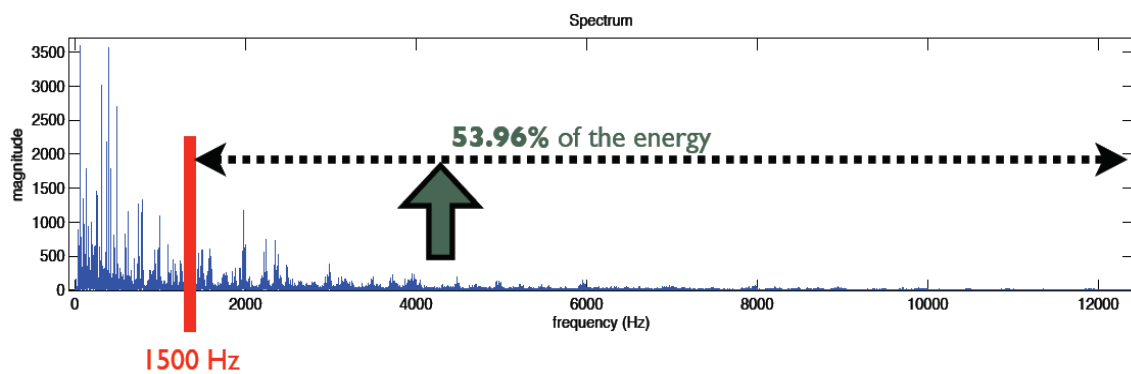
συχνότητα αυτή. Στο σχήμα 4.6, το 85% της συνολικής ενέργειας περιέχεται κάτω από την συχνότητα 5640.53 Hz.



Σχήμα 4.6: Υπολογισμός Rolloff

Brightness

Για τον υπολογισμό του ποσού ψηλής συχνότητας, μια διπλή μέθοδος είναι ο καθορισμός της cut-off συχνότητας και η μέτρηση του ποσού της ενέργειας που είναι πάνω από τη συχνότητα αυτή. Η καθορισμένη τιμή της συχνότητας αυτής είναι 1500Hz όπως φαίνεται και στο σχήμα 4.7.

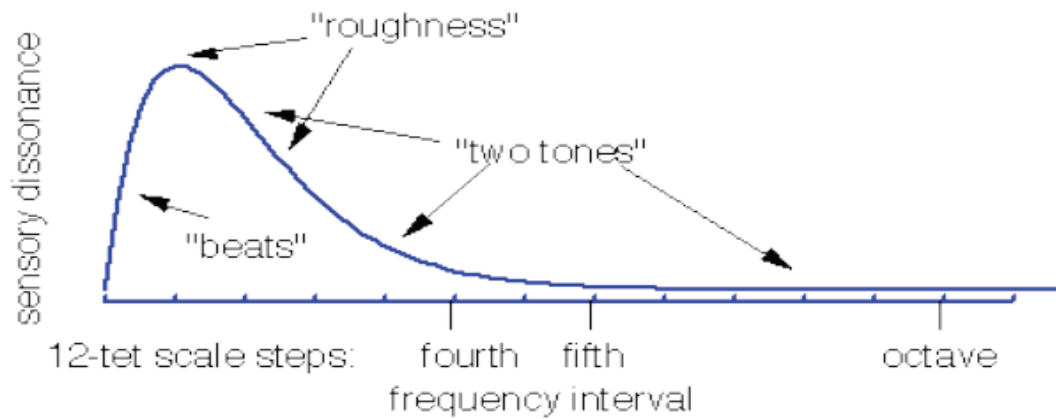


Σχήμα 4.7: Υπολογισμός Brightness

Roughness

Ακόμα ένα χαρακτηριστικό είναι η τραχύτητα ή αισθητηριακή παραφωνία. Για τον υπολογισμό της συνολικής τραχύτητας πρέπει να υπολογιστούν οι κορυφές του

φάσματος και παίρνοντας το μέσο όλων των παραφωνιών μεταξύ όλων των δυνατών ζευγαριών των κορυφών όπως φαίνεται στο σχήμα 4.8.



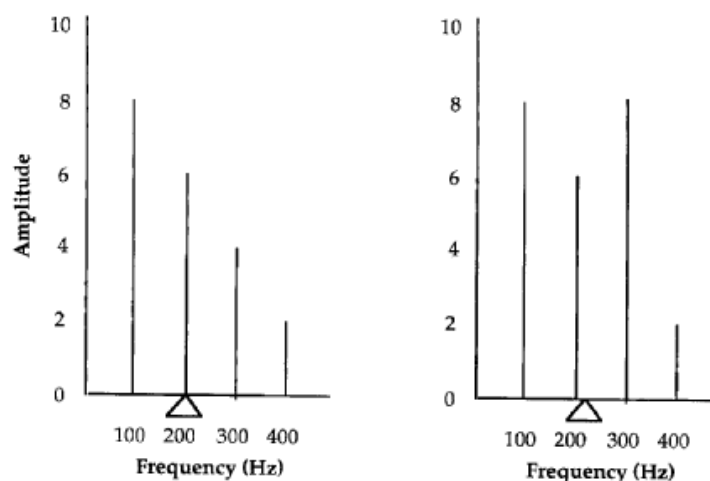
Σχήμα 4.8: Υπολογισμός Roughness

Regularity

Άλλο ένα γνώρισμα είναι η παρατυπία ενός φάσματος. Ο βαθμός δηλαδή μεταβολής των διαδοχικών κορυφών του.

Spectral Centroid

Το γνώρισμα αυτό είναι ένα μέτρο χαρακτηρισμού του φάσματος. Υποδεικνύει που βρίσκεται το κέντρο βάρους του φάσματος, όπως φαίνεται στο σχήμα 4.9, και είναι μια ένδειξη φωτεινότητας του ήχου.



Σχήμα 4.9: Spectral centroid

Entropy

Το γνώρισμα αυτό υποδεικνύει αν υπάρχουν κορυφές ή όχι στο φάσμα. Αν η καμπύλη του φάσματος είναι τελείως επίπεδη τότε η εντροπία είναι μέγιστη, αλλιώς αν η καμπύλη φάσματος εμφανίζει μόνο μια πολύ απότομη κορυφή τότε η εντροπία είναι ελάχιστη.

4.4 Παραδοσιακές μέθοδοι αντιμετώπισης

Στην εργασία αυτή χρησιμοποίησα low level features για την ανάλυση της μουσικής για το κάθε κομμάτι, με σκοπό την εύρεση συγκεκριμένων χαρακτηριστικών του. Τα χαμηλού επιπέδου χαρακτηριστικά που χρησιμοποίησα είναι το RMS (root mean square), το ZCR (zero crossing rate), το Spectral centroid, το Roll off, το Entropy καθώς επίσης και το Mfcc (Mel Frequency Cepstral Coefficients). Αφού έγινε η κατάλληλη επεξεργασία σε κάθε τραγούδι, πάρθηκαν τα πιο πάνω χαρακτηριστικά, και στη συνέχεια αξιολογήθηκαν για να βγουν συμπεράσματα και κατηγοριοποιήσεις τους.

Παρόλα αυτά, μπορούν να χρησιμοποιηθούν και υψηλού επιπέδου χαρακτηριστικά (high level features) για την ανάλυση της μουσικής. Τα υψηλού επιπέδου χαρακτηριστικά μπορεί να είναι η μελέτη της νότας, η διάρκειας της, η συγχορδίας, η τονικότητας της κλίμακας. Επίσης τα μουσικά όργανα, η χρονολογία που γράφτηκε η κάθε μουσική, ακόμα και ο εκτελεστής μπορεί να χρησιμοποιηθούν για την ανάλυση της. Τα χαρακτηριστικά αυτά μπορούν να περιγράψουν την μελωδία και τον ρυθμό με ένα τρόπο παρόμοιο με αυτόν που τα αντιλαμβάνεται ο άνθρωπος.

4.5 Σύγκριση και συζήτηση μεθόδων

Η πρώτη μέθοδος έδωσε αρκετά καλά αποτελέσματα. Κατηγοριοποιήθηκαν σωστά 92% της Δυτικής μουσικής και 100% της Τούρκικης μουσικής.

Η δεύτερη μέθοδος είναι μια πιο αξιόπιστη μέθοδος για μοντελοποίηση μελωδιών καθώς επίσης με την ανάλυση των υψηλού επιπέδου χαρακτηριστικών, γίνεται πιο καλή

κατηγοριοποίηση της μουσικής σε ομάδες [3,9,15,17]. Άρα πιθανόν να δίνει πιο καλά αποτελέσματα, με πιο συγκεκριμενοποιημένες κατηγοριοποιήσεις.

Κεφάλαιο 5

Πρόταση εργασίας

5.1 Προεπεξεργασία – Εξαγωγή Χαρακτηριστικών	42
5.2 Υπολογιστική Νοημοσύνη	44
5.3 Αποτελέσματα	45

Σαν πρώτη τεχνική, έγινε εξαγωγή κάποιων χαρακτηριστικών των τραγουδιών με τη χρήση του λογισμικού Matlab και στη συνέχεια με την κατάλληλη προεπεξεργασία στα δεδομένα, δημιούργησα διάφορα μοντέλα με τη χρήση του λογισμικού Weka [18]. Στόχος ήταν να διερευνήσω τονικές ομοιότητες της Κυπριακής δημοτικής μουσικής, Τούρκικης δημοτικής μουσικής και Δυτικής μουσικής χρησιμοποιώντας μια υπολογιστική προσέγγιση. Η βάση που χρησιμοποίησα περιείχε 127 μονοφωνικά τραγούδια, από τα οποία τα 37 ήταν Κυπριακά, τα 43 ήταν Τούρκικα και τα υπόλοιπα 47 ήταν Δυτικά. Στον πίνακα 5.1 φαίνονται τα Κυπριακά τραγούδια που χρησιμοποιήθηκαν, στον πίνακα 5.2 φαίνονται τα Δυτικά τραγούδια που χρησιμοποιήθηκαν και στον πίνακα 5.3 φαίνονται τα Τούρκικα τραγούδια που χρησιμοποιήθηκαν.

Πίνακας 5.1: Πίνακας Κυπριακών Τραγουδιών

Κυπριακά Τραγούδια
- 37 Τραγούδια, από Αντρέας Γριστάκκος, Γιάννης Ζαβρός

Πίνακας 5.2: Πίνακας Δυτικών Τραγουδιών

Δυτικά Τραγούδια
Με φλάουτο:
- 4 κινήσεις από partita για σόλο φλάουτο, από Bach
- 12 φαντασίας για σόλο φλάουτο, από Telemann
- «Syrinx», από Debussy
- «Soliloquy for Solo Flute Op. 44», από Lowell Liebermann
- «Image for solo flute», από Bozza
- «Danse de la chevre», από Arthur Honegger
- «Tango Etude», από Piazzolla
- «Daphnis et Chloe», από Ravel
- 9 σόλο αυτοσχεδιασμούς
Με φαγκάτο:
- 3 σόλο
Με κλαρινέτο:
- 3 σόλο, από Stravinsky
Με όμποε:
- 2 μονοφωνικά σόλο
Με σαξόφωνο:
- 8 μονοφωνικά σόλο

Πίνακας 5.3: Πίνακας Τούρκικων Τραγουδιών

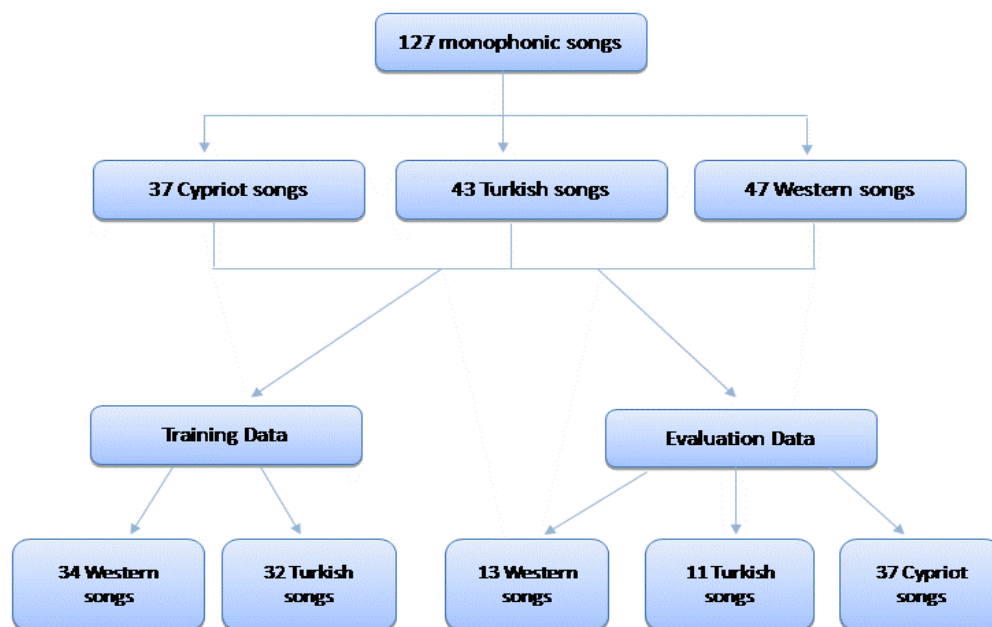
Τούρκικα Τραγούδια
*6 Makams με:
- 7 Hicaz
- 7 Huseiny
- 7 Huzzam
- 7 Nihavend
- 8 Saba
- 7 Ussak

5.1 Προεπεξεργασία - Εξαγωγή χαρακτηριστικών

Κάθε τραγούδι χωρίστηκε σε 13000 παράθυρα των 1024 πλαισίων μήκους. Για κάθε παράθυρο έγινε η εξαγωγή 16 χαμηλού επιπέδου χαρακτηριστικών και το μέσο και η τυπική απόκλιση κάθε χαρακτηριστικού αποθηκεύτηκαν σαν μια άλλη βάση

δεδομένων, δημιουργώντας έτσι ένα διάνυσμα 34 χαρακτηριστικών για κάθε τραγούδι. Τα χαρακτηριστικά χροιάς που χρησιμοποιήθηκαν ήταν: RMS-μέσο, RMS-τυπική απόκλιση, Zero crossing rate-μέσο, Zero crossing rate-τυπική απόκλιση, Spectral centroid-μέσο, Spectral centroid-τυπική απόκλιση, Roll off- μέσο, Roll off-τυπική απόκλιση, Entropy-μέσο, Entropy-τυπική απόκλιση, Mfcc (13 συντελεστές)-μέσο, Mfcc (13 συντελεστές)-τυπική απόκλιση. Στη συνέχεια για κάθε τραγούδι, εξάχθηκαν καμπύλες τονικού ύψους, χρησιμοποιώντας τον αλγόριθμο YIN [6], και δημιουργήθηκαν τονικά ιστογράμματα με ανάλυση 1200 πλαισίων ανά οκτάβα. Από τα ιστογράμματα αυτά, πάρθηκαν η θέση και το πλάτος των 7 πιο ψηλών κορυφών και έτσι δημιουργήθηκαν 14 μεσαίου επιπέδου τονικά χαρακτηριστικά.

Στη συνέχεια χώρισα τη βάση δεδομένων σε δεδομένα εκπαίδευσης και σε δεδομένα επαλήθευσης. Τα δεδομένα εκπαίδευσης αποτελούνταν από 34 Δυτικά τραγούδια και 32 Τούρκικα. Τα δεδομένα επαλήθευσης αποτελούνταν από 13 Δυτικά τραγούδια, 11 Τούρκικα και 37 Κυπριακά όπως φαίνεται στο σχήμα 5.1.

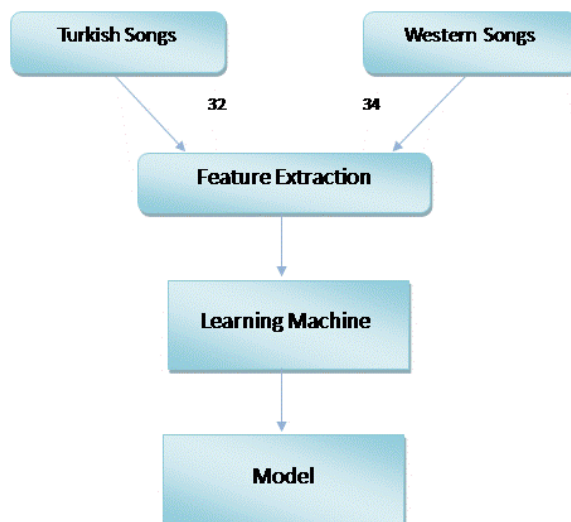


Σχήμα 5.1: Βάση δεδομένων με τα δεδομένα εκπαίδευσης και τα δεδομένα επαλήθευσης.

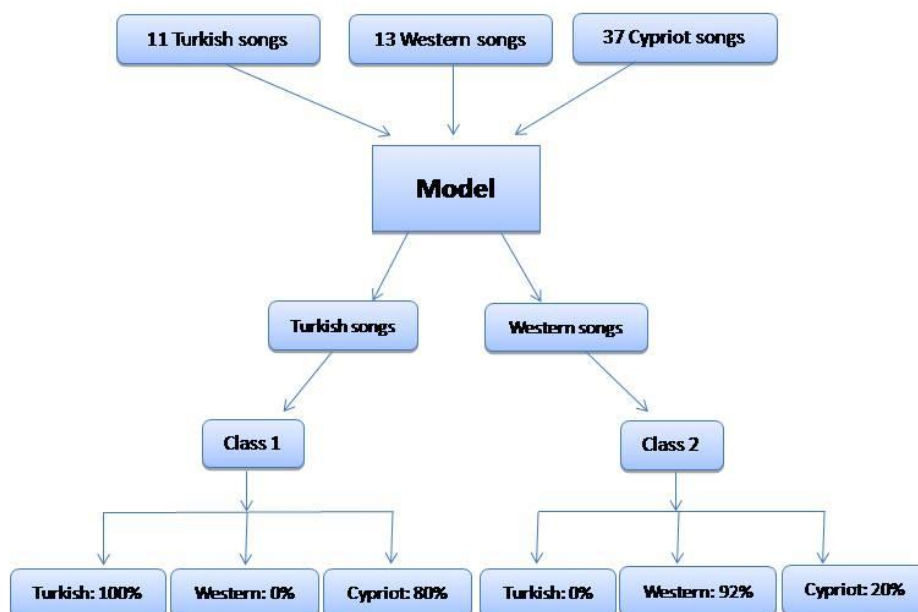
5.2 Υπολογιστική Νοημοσύνη

Δημιούργησα μοντέλα με επιβλεπόμενη μάθηση με τη χρήση νευρωνικών δικτύων Multilayer Perceptron ενός κρυφού επιπέδου, K κοντινότερου γείτονα, και μηχανές διανυσμάτων (support vector machines) με πυρήνες 1,2,3. Έκανα 3 πειράματα προκειμένου να γίνει καλύτερη κατανόηση των τονικών ομοιοτήτων και ομοιοτήτων χροιάς καθώς επίσης και των διαφορών μεταξύ Δυτικής και Κυπριακής μουσικής και μεταξύ Τούρκικης και Κυπριακής μουσικής. Η διαδικασία εκπαίδευσης και επαλήθευσης φαίνονται στο σχήμα 5.5 και 5.6.

Για την κατηγοριοποίηση των τραγουδιών έθεσα ένα κατώφλι με τιμή 0.5 έτσι ώστε τραγούδια με τιμή εξόδου κάτω από το κατώφλι αυτό να μπαίνουν στην ομάδα των Δυτικών τραγουδιών και τραγούδια με τιμή εξόδου πάνω από το κατώφλι να μπαίνουν στην ομάδα των Τούρκικων τραγουδιών.



Σχήμα 5.5: Φάση εκπαίδευσης συστήματος, όπου έγινε χρήση 32 Τούρκικων τραγουδιών και 34 Δυτικών τραγουδιών, για την κατηγοριοποίησή τους.



Σχήμα 5.6: Φάση επαλήθευσης συστήματος, όπου έγινε χρήση 11 Τούρκικων τραγουδιών, 13 Δυτικών τραγουδιών και 37 Κυπριακών τραγουδιών.

5.3 Αποτελέσματα

Όλα τα μοντέλα μπόρεσαν να μάθουν με επιτυχία όλα τα δεδομένα εκπαίδευσης (Τούρκικα, Δυτικά τραγούδια) με πολύ καλά ποσοστά επιτυχίας, σχεδόν 100%.

Στο πρώτο πείραμα, που χρησιμοποιήθηκαν μόνο χαρακτηριστικά χροιάς, το μοντέλο των νευρωνικών δικτύων μπόρεσε και κατηγοριοποίησε σωστά 92% των Δυτικών τραγουδιών (1 τραγούδι λανθασμένα), και 91% των Τούρκικων τραγουδιών (1 τραγούδι λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 84% των Κυπριακών τραγουδιών σε Δυτικά και το 16% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.4. Το μοντέλο με 1 κοντινότερο γείτονα μπόρεσε και κατηγοριοποίησε σωστά 100% των Δυτικών τραγουδιών, και 82% των Τούρκικων τραγουδιών (2 τραγούδια λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 38% των Κυπριακών τραγουδιών σε Δυτικά και το 62% σε Τούρκικα. Τα αποτελέσματα από το

μοντέλο αυτό φαίνονται στον πίνακα 5.5. Το μοντέλο με 2 κοντινότερους γείτονες μπόρεσε και κατηγοριοποίησε σωστά 100% των Δυτικών τραγουδιών, και 91% των Τούρκικων τραγουδιών (1 τραγούδι λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 22% των Κυπριακών τραγουδιών σε Δυτικά και το 78% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.6. Το μοντέλο μηχανής διανυσμάτων με πυρήνα 1 μπόρεσε και κατηγοριοποίησε σωστά 77% των Δυτικών τραγουδιών (3 τραγούδια λανθασμένα), και 73% των Τούρκικων τραγουδιών (3 τραγούδια λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 73% των Κυπριακών τραγουδιών σε Δυτικά και το 27% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.7. Το μοντέλο μηχανής διανυσμάτων με πυρήνα 2 μπόρεσε και κατηγοριοποίησε σωστά 92% των Δυτικών τραγουδιών (1 τραγούδι λανθασμένα), και 91% των Τούρκικων τραγουδιών (1 τραγούδι λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 60% των Κυπριακών τραγουδιών σε Δυτικά και το 40% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.8. Το μοντέλο μηχανής διανυσμάτων με πυρήνα 3 μπόρεσε και κατηγοριοποίησε σωστά 54% των Δυτικών τραγουδιών (6 τραγούδια λανθασμένα), και 91% των Τούρκικων τραγουδιών (1 τραγούδι λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 38% των Κυπριακών τραγουδιών σε Δυτικά και το 60% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.9.

Πίνακας 5.4: Μοντέλο Νευρωνικών Διχτύων

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	12	1
	Turkish songs	1	10
	Cypriot songs	31	6

Πίνακας 5.5: Μοντέλο με 1 Κοντινότερο Γείτονα

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	13	0
	Turkish songs	2	9
	Cypriot songs	14	23

Πίνακας 5.6: Μοντέλο με 2 Κοντινότερους Γείτονες

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	13	0
	Turkish songs	1	10
	Cypriot songs	8	29

Πίνακας 5.7: Μοντέλο μηχανής διανυσμάτων με 1 πυρήνα

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	10	3
	Turkish songs	3	8
	Cypriot songs	27	10

Πίνακας 5.8: Μοντέλο μηχανής διανυσμάτων με 2 πυρήνες

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	12	1
	Turkish songs	1	10
	Cypriot songs	22	15

Πίνακας 5.9: Μοντέλο μηχανής διανυσμάτων με 3 πυρήνες

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	7	6
	Turkish songs	1	10
	Cypriot songs	14	22

Στο δεύτερο πείραμα, που χρησιμοποιήθηκαν μόνο τονικά χαρακτηριστικά, το μοντέλο των νευρωνικών δικτύων μπόρεσε και κατηγοριοποίησε σωστά 54% των Δυτικών τραγουδιών (6 τραγούδια λανθασμένα), και 82% των Τούρκικων τραγουδιών (2 τραγούδια λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 51% των Κυπριακών τραγουδιών σε Δυτικά και το 46% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.10. Το μοντέλο με 1 κοντινότερο γείτονα μπόρεσε και κατηγοριοποίησε σωστά 77% των Δυτικών τραγουδιών (3 τραγούδια λανθασμένα), και 73% των Τούρκικων τραγουδιών (3 τραγούδια λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 54% των Κυπριακών τραγουδιών σε Δυτικά και το 46% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.11. Το μοντέλο με 2 κοντινότερους γείτονες μπόρεσε και κατηγοριοποίησε σωστά 84% των Δυτικών τραγουδιών (2 τραγούδια λανθασμένα), και 91% των Τούρκικων τραγουδιών (1 τραγούδι λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 51% των Κυπριακών τραγουδιών σε Δυτικά και το 35% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.12. Το μοντέλο μηχανής διανυσμάτων με πυρήνα 1

μπόρεσε και κατηγοριοποίησε σωστά 77% των Δυτικών τραγουδιών (3 τραγούδια λανθασμένα), και 100% των Τούρκικων τραγουδιών. Το μοντέλο αυτό κατηγοριοποίησε το 41% των Κυπριακών τραγουδιών σε Δυτικά και το 59% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.13. Το μοντέλο μηχανής διανυσμάτων με πυρήνα 2 μπόρεσε και κατηγοριοποίησε σωστά 69% των Δυτικών τραγουδιών (4 τραγούδια λανθασμένα) , και 91% των Τούρκικων τραγουδιών (1 τραγούδι λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 46% των Κυπριακών τραγουδιών σε Δυτικά και το 54% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.14. Το μοντέλο μηχανής διανυσμάτων με πυρήνα 3 μπόρεσε και κατηγοριοποίησε σωστά 61% των Δυτικών τραγουδιών (5 τραγούδια λανθασμένα), και 100% των Τούρκικων τραγουδιών. Το μοντέλο αυτό κατηγοριοποίησε το 41% των Κυπριακών τραγουδιών σε Δυτικά και το 57% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.15.

Πίνακας 5.10: Μοντέλο Νευρωνικών Διχτύων

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	7	6
	Turkish songs	2	9
	Cypriot songs	19	17

Πίνακας 5.11: Μοντέλο με 1 Κοντινότερο Γείτονα

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	10	3
	Turkish songs	3	8
	Cypriot songs	20	17

Πίνακας 5.12: Μοντέλο με 2 Κοντινότερους Γείτονες

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	11	2
	Turkish songs	1	10
	Cypriot songs	19	13

Πίνακας 5.13: Μοντέλο μηχανής διανυσμάτων με 1 πυρήνα

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	10	3
	Turkish songs	0	11
	Cypriot songs	15	22

Πίνακας 5.14: Μοντέλο μηχανής διανυσμάτων με 2 πυρήνες

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	9	4
	Turkish songs	1	10
	Cypriot songs	17	20

Πίνακας 5.15: Μοντέλο μηχανής διανυσμάτων με 3 πυρήνες

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	8	5
	Turkish songs	0	11
	Cypriot songs	15	21

Στο τρίτο πείραμα, που χρησιμοποιήθηκαν τονικά χαρακτηριστικά και χαρακτηριστικά χροιάς, το μοντέλο των νευρωνικών δικτύων μπόρεσε και κατηγοριοποίησε σωστά 92% των Δυτικών τραγουδιών (1 τραγούδι λανθασμένα), και 91% των Τούρκικων τραγουδιών (1 τραγούδι λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 84% των Κυπριακών τραγουδιών σε Δυτικά και το 16% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.16. Το μοντέλο με 1 κοντινότερο γείτονα μπόρεσε και κατηγοριοποίησε σωστά 100% των Δυτικών τραγουδιών, και 69% των Τούρκικων τραγουδιών (2 τραγούδια λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 38% των Κυπριακών τραγουδιών σε Δυτικά και το 62% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.17. Το μοντέλο με 2 κοντινότερους γείτονες μπόρεσε και κατηγοριοποίησε σωστά 100% των Δυτικών τραγουδιών, και 91% των Τούρκικων τραγουδιών (1 τραγούδι λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 22% των Κυπριακών τραγουδιών σε Δυτικά και το 78% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.18. Το μοντέλο μηχανής διανυσμάτων με πυρήνα 1 μπόρεσε και κατηγοριοποίησε σωστά 77% των Δυτικών τραγουδιών (3 τραγούδια λανθασμένα), και 73% των Τούρκικων τραγουδιών (3 τραγούδια λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 73% των Κυπριακών τραγουδιών σε Δυτικά και το 27% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.19. Το μοντέλο μηχανής διανυσμάτων με 2 πυρήνες μπόρεσε και κατηγοριοποίησε σωστά 92% των Δυτικών τραγουδιών (1 τραγούδι λανθασμένα), και 91% των Τούρκικων τραγουδιών (1 τραγούδι λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 59% των Κυπριακών τραγουδιών σε Δυτικά και το 41% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.20. Το μοντέλο μηχανής διανυσμάτων με 3 πυρήνες μπόρεσε και κατηγοριοποίησε

σωστά 54% των Δυτικών τραγουδιών (6 τραγούδια λανθασμένα), και 91% των Τούρκικων τραγουδιών (1 τραγούδι λανθασμένα). Το μοντέλο αυτό κατηγοριοποίησε το 38% των Κυπριακών τραγουδιών σε Δυτικά και το 59% σε Τούρκικα. Τα αποτελέσματα από το μοντέλο αυτό φαίνονται στον πίνακα 5.21.

Πίνακας 5.16: Μοντέλο Νευρωνικών Διχτύων

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	13	0
	Turkish songs	0	11
	Cypriot songs	30	7

Πίνακας 5.17: Μοντέλο με 1 Κοντινότερο Γείτονα

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	13	0
	Turkish songs	3	8
	Cypriot songs	26	11

Πίνακας 5.18: Μοντέλο με 2 Κοντινότερους Γείτονες

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	13	0
	Turkish songs	1	10
	Cypriot songs	16	21

Πίνακας 5.19: Μοντέλο μηχανής διανυσμάτων με 1 πυρήνα

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	11	2
	Turkish songs	0	11
	Cypriot songs	26	11

Πίνακας 5.20: Μοντέλο μηχανής διανυσμάτων με 2 πυρήνες

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	13	0
	Turkish songs	1	10
	Cypriot songs	32	5

Πίνακας 5.21: Μοντέλο μηχανής διανυσμάτων με 3 πυρήνες

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	13	0
	Turkish songs	0	11
	Cypriot songs	32	5

Κεφάλαιο 6

Συμπεράσματα

6.1 Γενικά συμπεράσματα

54

6.1 Γενικά συμπεράσματα

Μετά την ανάλυση των αποτελεσμάτων που λήφθηκαν από την όλη διαδικασία που ακολουθήθηκε, καταλήξαμε σε ορισμένα συμπεράσματα. Κάποια από τα συμπεράσματα ήταν αναμενόμενα από την αρχή, προέκυψαν όμως και άλλα στην πορεία. Πετύχαμε αρκετά καλά αποτελέσματα θέτοντας έτσι βάσεις για τυχόν μελλοντικές εργασίες.

Καταρχάς κάποια από τα αποτελέσματα ίσως να φαίνονται αντιφατικά, δηλαδή να μην συμφωνούν πλήρως με την απόδοση των συγκεκριμένων μοντέλων. Η χρήση περισσότερων ή διαφορετικών χαρακτηριστικών για την εκπαίδευση των μοντέλων ή ακόμα και η χρήση υψηλού επιπέδου χαρακτηριστικών, ίσως να έδινε λύση στο πιο πάνω πρόβλημα.

Πρώτα από όλα, ήταν αναμενόμενο να ξεχωρίζει με πολύ καλή ακρίβεια στη φάση εκπαίδευσης λόγω του ότι οι δύο μουσικές, Τούρκικη και Δυτική, είναι πολύ διαφορετικές αλλά μεταξύ τους μοιάζουν πολύ. Από κάποια αποτελέσματα φαίνεται ότι η Κυπριακή μουσική ίσως να μοιάζει περισσότερο με την Τούρκικη μουσική.

Επίσης, η κατηγοριοποίηση της μουσικής έγινε με χαμηλού επιπέδου χαρακτηριστικά και δεν μοντελοποιήθηκαν οι μελωδίες τους, ούτε κάποιο ρυθμικό χαρακτηριστικό. Έγινε μοντελοποίηση μόνο χαρακτηριστικών χροιάς και τονικών χαρακτηριστικών, που

είναι μια ειδική περίπτωση και δεν καλύπτει γενικά όλα τα χαρακτηριστικά της μουσικής.

Επίσης, από τα πιο πάνω αποτελέσματα, μόνο τα μοντέλα που χρησιμοποιήθηκαν τονικά χαρακτηριστικά και χαρακτηριστικά χροιάς μπόρεσαν να διακρίνουν τη Δυτική μουσική από την Τούρκικη μουσική. Τα μοντέλα που χτίστηκαν μόνο με χαρακτηριστικά χροιάς, έδειξαν ότι 78% των Κυπριακών τραγουδιών κατηγοριοποιούνται ως Τούρκικα τραγούδια, και το υπόλοιπο 22% κατηγοριοποιούνται ως Δυτικά τραγούδια. Τα αποτελέσματα αυτά φαίνονται στον πίνακα 6.1.

Πίνακας 6.1: Μοντέλο με 2 Κοντινότερους Γείτονες

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	13	0
	Turkish songs	1	10
	Cypriot songs	8	29

Τα μοντέλα που χτίστηκαν μόνο με τονικά χαρακτηριστικά, δεν μπόρεσαν να διακρίνουν εντελώς τα Δυτικά τραγούδια από τα Τούρκικα. Έδειξαν ότι 60% των Κυπριακών τραγουδιών κατηγοριοποιούνται ως Τούρκικα τραγούδια, και το υπόλοιπο 40% κατηγοριοποιούνται ως Δυτικά τραγούδια. Τα αποτελέσματα αυτά φαίνονται στον πίνακα 6.2.

Πίνακας 6.2: Μοντέλο μηχανής διανυσμάτων με 1 πυρήνα

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	10	3
	Turkish songs	0	11
	Cypriot songs	15	22

Τα μοντέλα που χτίστηκαν με τονικά χαρακτηριστικά και χαρακτηριστικά χροιάς, μπόρεσαν να διακρίνουν τα Δυτικά τραγούδια από τα Τούρκικα. Έδειξαν ότι 19% των Κυπριακών τραγουδιών κατηγοριοποιούνται ως Τούρκικα τραγούδια, και το υπόλοιπο 81% κατηγοριοποιούνται ως Δυτικά τραγούδια. Τα αποτελέσματα αυτά φαίνονται στον πίνακα 6.3.

Πίνακας 6.3: Μοντέλο Νευρωνικών Διχτύων

confusion matrix		Predicted class	
		Western songs	Turkish songs
Actual class	Western songs	13	0
	Turkish songs	0	11
	Cypriot songs	30	7

Τέλος, ίσως θα πρέπει η βάση δεδομένων να μεγαλώσει και να υποστεί κάποια περισσότερη επεξεργασία, έτσι ώστε να είναι πιο γενική, να είναι πιο αντιπροσωπευτικά τα μοντέλα και να γίνει πιο διακριτή η κατηγοριοποίηση των μουσικών.

Κεφάλαιο 7

Θέματα μελλοντικής μελέτης

7.1 Μελλοντική Εργασία

57

7.1 Μελλοντική Εργασία

Όπως έγινε αντιληπτό, η υλοποίηση τέτοιων συστημάτων μάθησης και η ανάλυση της Κυπριακής μουσικής για την εύρεση τονικών ομοιοτήτων με άλλες μουσικές, βοηθά τόσο τον τομέα της Μουσικής όσο και τον τομέα της Επιστήμης των Υπολογιστών καθώς επίσης και άλλων. Γι' αυτό και πιστεύω είναι πολύ σημαντικό να συνεχιστούν οι μελέτες αυτές. Λαμβάνοντας υπόψη τις εμπειρίες και τα όσα έμαθα και αποκόμισα από την όλη προσπάθεια που έκανα θα αναφέρω κάποιες προτάσεις που πιστεύω ότι αξίζει να μελετηθούν έτσι ώστε να δοθεί μια συνέχεια στην δουλειά που έχει γίνει.

Μια πρώτη προσπάθεια για μελλοντική εργασία που μπορεί να γίνει είναι η εξαγωγή χαρακτηριστικών υψηλού επιπέδου (high-level) για μοντελοποίηση των μελωδιών και των ρυθμών. Τα χαρακτηριστικά αυτά μπορεί να είναι η μελέτη της νότας, η διάρκεια της, η συγχορδία, η τονικότητα της κλίμακας. Επίσης τα μουσικά όργανα, η χρονολογία που γράφτηκε η κάθε μουσική, ακόμα και ο εκτελεστής μπορούν να χρησιμοποιηθούν για την ανάλυση αυτή. Με τη μέθοδο αυτή, γίνεται πιο καλή κατηγοριοποίηση της μουσικής σε ομάδες και δίνει καλύτερα αποτελέσματα με πιο συγκεκριμενοποιημένες κατηγοριοποιήσεις, κάτι που εξακρίβωσε και ο C.McKay στο άρθρο του [9].

Αναφορές

- [1] A.Alpkocak, A.Cenk Gedik, “Classification of Turkish Songs according to Makams by using n-grams”, Proceedings of the 15. Turkish Symposium on Artificial Intelligence and Neural Networks, TAINN, Mugla, 2006
- [2] R.J. Babaian, H.Fritsche, A.Ayala, V.Bhadkamkar, D.A.Johnston, W.Naccarato, Z.Zhang, “Performance of a neural network in detecting prostate cancer in the prostate-specific antigen reflex range of 2.5 to 4.0 ng/mL”, Urology 2000 Vol: 56:1000-1006, 2000
- [3] Brown, J. C, “Determination of meter of musical scores by Autocorrelation”, Journal of the Acoustical Society of America 94 (4): 1953-1957, 1993
- [4] A.Cenk Gedik, B.Bozkurt, “Evaluation of the Makam Scale Theory of Arel for Music Information Retrieval on Traditional Turkish Art Music”, Journal of New Music Research, 38(2): 103-116, 2009
- [5] T.Colin, A.Iollo, D.Lombard, O.Saut, “Prediction of the Evolution of Thyroidal Lung Nodules Using a Mathematical Model”, Ercim News 82, 2010
- [6] de Cheveigne, A., & Kawahara, H., “YIN a fundamental frequency estimator for speech and music.” The Journal of the Acoustical Society of America, 111 (4), 2002
- [7] R.C.Eberhart, Y.Shi, “Computational Intelligence Concepts to Implementations”, 2007
- [8] T.Kallinikos, “Kypriaki laiki mousa”, Nicosia, Cyprus - Michaelides 1951.

- [9] C.McKay, “Automatic Genre Classification as a Study of the Viability of High-Level Features for Music Classification”, Faculty of Music, McGill University , 2004
- [10] M.Panteli, “Pitch Patterns of Cypriot Folk Music between Byzantine and Ottoman Influence”, Master Thesis UPF, Universitat Pompeu Fabra, Barcelona, 2011
- [11] H.A.Rowley, S.Baluja, T.Kanade, “Neural Network-Based Face Detection”, IEEE Transactions on pattern analysis and machine Intelligence, Vol.20, No.1, 1998
- [12] C.Schizas, A.Neocleous, M.Panteli, N.Petkov, “Identification of Similarities between the Turkish Makam Scales and the Cypriot Folk Music”, HELINA's 5th National Conference, 2012
- [13] C.Schizas, A.Neocleous, M.Panteli, N.Petkov, “The tuning of pithkiavli instrument”, Internal report, University of Cyprus, 2012
- [14] S.Senturk, “Computational modeling of Improvisation in Turkish Folk music using variable-length Markov models”, Master Thesis in Music Technology in the School of Music, Georgia Institute of Technology, 2011
- [15] Shan, M. K., F. F. Kuo, “Music style mining and classification by melody”, IEICE Transactions on Information and Systems E86-D (3): 655–659, 2003
- [16] S.Tobolis, “Traditional Cyprian songs and dances”, Nicosia Cyprus, 1980
- [17] B.Whitman, P.Smaragdis, “Combining musical and cultural features for intelligent style detection”, Proceedings of the International Symposium on Music Information Retrieval 47–52, 2002

- [18] Weka Machine Learning
http://www.google.com.cy/#hl=el&gs_rn=9&gs_ri=psy-ab&cp=4&gs_id=e&xhr=t&q=weka&es_nrs=true&pf=p&output=search&sclien_t=psy-ab&oq=weka&gs_l=&pbx=1&bav=on.2,or.r_qf.&bvm=bv.45512109,d.ZWU&fp=22231b652fd6db86&biw=1440&bih=799
- [19] K-nearest neighbors algorithm
https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm
- [20] Support Vector Machine
http://en.wikipedia.org/wiki/Support_vector_machine

Παράρτημα Α

Ο πιο κάτω κώδικας υλοποιεί τη διαδικασία εξαγωγής του σχήματος 4.1, όπου το ηχητικό σήμα χωρίζεται σε πολλά παράθυρα, εφαρμόζεται η συνάρτηση που θα πολλαπλασιαστεί με το συγκεκριμένο παράθυρο και στη συνέχεια με τη χρήση γρήγορου μετασχηματισμού Fourier υπολογίζεται το συχνοτικό φάσμα.

```
sig=wavread('peloponisiako-disk6_1_Sarki_(Organiko).wav');
sig2=sig(44100:45000,1);
w=hamming(length(sig2));
ws=sig2.*w;
a = miraudio(ws,44100);
mirspectrum(a)
figure(1)
plot(sig2)
title('Audio segment')
xlabel('Time in bins')
ylabel('Sound pressure')
figure(2)
plot(w)
title('Window function (Hamming)')
xlabel('Number of bins')
figure(3)
plot(ws)
title('Weighted segment')
xlabel('Number of bins')
ylabel('Sound pressure')
```

Ο πιο κάτω κώδικας εξάγει τα χαρακτηριστικά που χρησιμοποιήθηκαν (zero cross, centroid, brightness, spread, skewness, kurtosis, roll off, spectrum entropy, flatness, roughness, irregularity, inharmonicity, mfcc) για την ανάλυση των διάφορων τραγουδιών.

```
starttime=5;
endtime=40*44100;
w=1024;
wu='sp';
h=128;
hu='sp';
%%% Select features %%%
%%% Dynamics features
irms=1;
%%% Timbre features
izerocross=1;
icentroid=1;
ibrightness=0;
ispread=0;
iskewness=0;
ikurtosis=0;
irolloff=1;
ispectentropy=1;
iflatness=0;
iroughness=0;
iirregularity=0;
iinharmonicity=0;
imfcc=1;
%%% Pitch features
ichromagram=0;
n=1;
folders={'Western','Turkey','Cyprus'};
for i=1:length(folders)
```

```

eval(['cd ' folders{i}])
files = dir('*.wav');
for k_files=1:length(files)
    audio = miraudio (files(k_files).name,'Extract',starttime,endtime,'sp');
    frame=mirframe(audio,w,wu,h,hu);
    myflow = mirstruct;
    %%% Dynamics features
    if irms
        rms=mirrms(frame);
        table_features{n}='rms';
        dataset(n,:)=mirgetdata(rms);
        n=n+1;
    end
    %%% Timbre features
    if izerocross
        zero_crossing=mirzerocross(frame);
        table_features{n}='zerocross';
        dataset(n,:)=mirgetdata(zero_crossing);
        n=n+1;
    end
    if icentroid
        centroid=mircentroid(frame);
        table_features{n}='centroid';
        dataset(n,:)=mirgetdata(centroid);
        n=n+1;
    end
    if ibrightness
        brightness=mirbrightness(frame);
        table_features{n}='brightness';
        dataset(n,:)=mirgetdata(brightness);
        n=n+1;
    end
    if ispread

```

```

        spread=mirspread(frame);
        table_features{n}='spread';
        dataset(n,:)=mirgetdata(spread);
        n=n+1;
    end
    if iskewness
        skewness=mirskewness(frame);
        table_features{n}='skewness';
        dataset(n,:)=mirgetdata(skewness);
        n=n+1;
    end
    if ikurtosis
        kurtosis=mirkurtosis(frame);
        table_features{n}='kurtosis';
        dataset(n,:)=mirgetdata(kurtosis);
        n=n+1;
    end
    if irolloff
        roll_off=mirrolloff(frame);
        table_features{n}='roll_off';
        dataset(n,:)=mirgetdata(roll_off);
        n=n+1;
    end
    if ispectentropy
        entropy=mirentropy(frame);
        table_features{n}='entropy';
        dataset(n,:)=mirgetdata(entropy);
        n=n+1;
    end
    if iflatness
        flatness=mirflatness(frame);
        table_features{n}='flatness';
        dataset(n,:)=mirgetdata(flatness);
    end

```

```

        n=n+1;
    end
    if iroughness
        roughness=mirroughness(frame);
        table_features{n}='roughness';
        dataset(n,:)=mirgetdata(roughness);
        n=n+1;
    end
    if iirregularity
        regularity=mirregularity(frame);
        table_features{n}='regularity';
        dataset(n,:)=mirgetdata(regularity);
        n=n+1;
    end
    if iinharmoniccity
        inharmoniccity=mirinharmoniccity(frame);
        table_features{n}='inharmoniccity';
        dataset(n,:)=mirgetdata(inharmoniccity);
        n=n+1;
    end
    if imfcc
        spec=mirspectrum(frame,'Max',3000);
        cepstrum=mirmfcc(spec);
        for icoeff=0:length(cepstrum(:,1))-1
            name={['mfcc' int2str(icoeff+1)]};
            table_features{n+icoeff}=name;
        end
        mfcc=mirgetdata(cepstrum);
        dataset(n:n+length(mfcc(:,1))-1,:)=mfcc;
        n=n+length(cepstrum(:,1));
    end
    %%%% Pitch features
%    if ichromagram

```

```
%      chromagram=mirchromagram(audio);
%      table_features{n}='inharmonicity';
%      dataset(n,:)=mirgetdata(inharmonicity);
%      n=n+1;
%      end
    end
end
```