



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Υποστήριξη της διάγνωσης της καρωτιδικής αθηρωμάτωσης με τη
βοήθεια προτύπων κινητικότητας και παραμόρφωσης του
αρτηριακού τοιχώματος**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Πρεβένιος Μαρίνος

Επιβλέπουσα: Κωνσταντίνα Σπ. Νικήτα
Καθηγήτρια Ε.Μ.Π

Αθήνα, Μάιος 2014

(Η σελίδα αυτή είναι σκόπιμα λευκή.)



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Υποστήριξη της διάγνωσης της καρωτιδικής αθηρωμάτωσης με τη
βοήθεια προτύπων κινητικότητας και παραμόρφωσης του
αρτηριακού τοιχώματος**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Πρεβένιος Μαρίνος

Επιβλέπουσα: Κωνσταντίνα Σπ. Νικήτα
Καθηγήτρια Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή, 16 Μαΐου 2014.

.....
Κ.Σπ. Νικήτα
Καθηγήτρια Ε.Μ.Π

.....
Δ.Δ. Κουτσούρης
Καθηγητής Ε.Μ.Π

.....
Γ. Ματσόπουλος
Επίκουρος Καθηγητής Ε.Μ.Π

Αθήνα, Μάιος 2014

.....
Πρεβένιος Μαρίνος

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Πρεβένιος Μαρίνος (2014)

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 τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Σε αυτή την εργασία προσεγγίστηκε μια σπουδαία κλινική πρόκληση, η υποστήριξη της διάγνωσης της καρωτιδικής αθηρωμάτωσης, που ευθύνεται για την πλειονότητα των εγκεφαλικών επεισοδίων. Σε αυτή τη κατεύθυνση εξετάστηκε η δυνατότητα διάκρισης μεταξύ ασταθών και μη, αθηρωματικών πλακών στη καρωτίδα αρτηρία, μέσω της ανάλυσης ακολουθιών εικόνων υπερήχου. Έχοντας στη διάθεση μας έναν εκτιμητή κίνησης, που είχε ήδη υλοποιηθεί, βελτιστοποιηθεί και αξιολογηθεί για την εκτίμηση της κίνησης του αρτηριακού τοιχώματος, εφαρμόστηκε σε ακολουθίες εικόνων υπερήχου B-σάρωσης για ένα σύνολο 96 ασθενών που πάσχουν από καρωτιδική αθηρωμάτωση. Τα αποτελέσματα από την εκτίμηση κίνησης για τον κάθε ασθενή χρησιμοποιήθηκαν για την παραγωγή 146 κατηγοριών κυματομορφών κίνησης, που αναπαριστούν μοτίβα κινήσεων και παραμορφώσεων σε σχέση με : (α) κινήσεις της αθηρωματικής πλάκας και των υγιών μερών του αρτηριακού τοιχώματος πλησίον της πλάκας, όπως και (β) τοπικές παραμορφώσεις και σχετικές κινήσεις του αρτηριακού τοιχώματος, αντίστοιχα. Αυτές οι κυματομορφές κίνησης τροφοδότησαν τα Κρυφά Μαρκοβιανά Μοντέλα (HMMs), που έχουν χρησιμοποιηθεί με μεγάλη επιτυχία στην αναγνώριση φωνής αλλά και σε πολλές άλλες εφαρμογές μηχανικής μάθησης. Στην περίπτωση μας, τα HMMs χρησιμοποιήθηκαν για να αναγνωρίσουν περιπτώσεις κυματομορφών συμπτωματικών και ασυμπτωματικών ασθενών, με την αναλογία του ότι ο εκάστοτε ασθενής συνεισφέρει με τη δική του καρωτιδική αρτηριακή "προφορά" στο μοντέλο. Η απόδοση (ακρίβεια κατηγοριοποίησης) στα HMMs κυμάνθηκε μεταξύ 57.05% και 81.44% για τις 146 κατηγορίες κυματομορφών, με μέση απόδοση 69.95%. Τα αποτελέσματα αυτά επαληθεύτηκαν με την χρήση Μηχανών Διανυσμάτων Υποστήριξης (SVMs) στο ίδιο σύνολο δεδομένων, που κατηγοριοποιήθηκαν με μέση απόδοση 59%. Στην περίπτωση που επιλέχθηκαν, μονάχα, αυτές οι κατηγορίες κυματομορφών που παρουσίασαν τα υψηλότερα ποσοστά απόδοσης, τα HMMs μπόρεσαν να διακρίνουν ασυμπτωματικούς και συμπτωματικούς ασθενείς με απόδοση 75.81% και 73.91% αντίστοιχα, ενώ τα SVMs έφεραν αποτελέσματα 78.66% και 69.84% αντίστοιχα. Με δεδομένη την υπεροχή των SVMs, επιλέχθηκαν από αυτό το μοντέλο, οι συγκεκριμένες κατηγορίες κυματομορφών και δημιουργήθηκε ένα σύστημα πλειοψηφίας, για την υποστήριξη της διάγνωσης της ασθένειας. Η πλήρης δυναμική του συστήματος μένει να αποδειχθεί στο πλαίσιο μελλοντικών μελετών, σε μεγάλο πλήθος ασθενών με καρωτιδική αθηρωμάτωση οι οποίοι θα υποβληθούν σε συχνούς επανελέγχους, ενώ το σύστημα μπορεί να εμπλουτιστεί με πρόσθετα χαρακτηριστικά με χρονική εξέλιξη, όπως η αρτηριακή πίεση και ο καρδιακός ρυθμός.

Λέξεις κλειδιά: *ανάλυση κίνησης, καρωτιδική αθηρωμάτωση, υπέρηχος, Κρυφά Μαρκοβιανά Μοντέλα, Μηχανές Διανυσμάτων Υποστήριξης*

(Η σελίδα αυτή είναι σκόπιμα λευκή.)

Abstract

This thesis addressed a major clinical challenge, namely *valid risk stratification for carotid atherosclerosis*, which constitutes the most common cause of stroke. Toward this direction, we investigated the potential of spatiotemporal ultrasound image analysis in discriminating between stable and vulnerable atherosclerotic plaques of the carotid artery. To this end, a motion estimator, which has previously been designed, optimized and evaluated for arterial wall motion estimation, was applied to B-mode ultrasound image sequences of 96 patients with carotid atherosclerosis. The results of motion estimation for each patient were used to produce 146 types of motion waveforms, which represented mobility and strain patterns that express: (a) the movements of the atherosclerotic lesion (plaque) and the healthy parts of the arterial wall adjacent to the plaque, and (b) local deformations and relative movements in the arterial wall, respectively. These sets of motion waveforms were fed to Hidden Markov Models (HMMs), which have been successfully used in voice recognition and several other tasks of machine learning. In this case, HMMs were used to recognize motion waveforms of symptomatic and asymptomatic cases, intuitively considering that, for each patient, a motion waveform represents the "accent" of the carotid-artery wall. The performance (classification accuracy) of HMMs ranged between 57.05% and 81.44% for the 146 types of motion waveforms, while the average performance was 69.95%. These results were cross-validated using Support Vector Machines (SVMs) on the same dataset, which reached, an average of 59% classification accuracy. In the case that only the motion waveforms with the highest discrimination power were used, the HMMs could identify asymptomatic and symptomatic patients with 75.81% and 73.91% accuracy, respectively; the corresponding results for SVMs were 78.66% and 69.84%, respectively. Given the high performance in the latter case, those motion waveforms and their incorporation in SVMs were used to create a majority voting system for risk stratification for the disease. The full potential of the system will be revealed on future prospective studies with frequent patient follow-ups, while the system could be further enriched with other temporal features, such as the arterial pressure and heart rate.

Keywords: *motion analysis, carotid atherosclerosis, ultrasound, Hidden Markov Models, Support Vector Machines*

Ευχαριστώ τους πολύ κοντινούς μου ανθρώπους για την υπομονή και την υποστήριξη.

Ευχαριστώ την καθηγήτρια μου Κ. Νικήτα που μου έδωσε το κίνητρο μέσα από τις πολύ ενδιαφέρουσες διαλέξεις της, να ασχοληθώ εμπράκτως με την βιοϊατρική.

Ευχαριστώ ακόμη την υποψήφια διδάκτορα Αιμιλία Γαστουνιώτη που με άμεσο τρόπο βοήθησε σε κάθε βήμα της διπλωματικής αυτής εργασίας.

Περιεχόμενα

I	Επισκόπηση - Βασικά στοιχεία μελέτης	8
1	Εισαγωγή	9
1.1	Εγκεφαλικό Επεισόδιο	10
1.2	Καρωτιδική αθηρωμάτωση και Υπερηχογραφική απεικόνιση	11
1.2.1	Ο ρόλος των μηχανικών ιδιοτήτων του αρτηριακού τοιχώματος	14
1.2.2	Επισκόπηση των εργασιών για την υποστήριξη της διάγνωσης	15
2	Αντικείμενο και δομή της εργασίας	17
2.1	Αντικείμενο	17
2.2	Συνεισφορά	18
2.3	Οργάνωση κειμένου	19
II	Θεωρητικό υπόβαθρο	20
3	Κρυφά Μαρκοβιανά Μοντέλα (Hidden Markov Models)	21
3.1	Θεωρία των Κρυφών Μαρκοβιανών Μοντέλων	21
3.1.1	Αλυσίδες Markov - Στοχαστικές ανελίξεις Markov	22
3.2	Αλγόριθμοι και Κρυφά Μαρκοβιανά Μοντέλα	23
3.3	Ορισμός των Κρυφών Μαρκοβιανών Μοντέλων	24
3.4	Στοχαστικά Αυτόματα - Αρχιτεκτονική Μοντέλου	26
3.5	Τα Τρία Βασικά Προβλήματα	27
3.5.1	Το πρόβλημα της εκτίμησης - Αλγόριθμος Forward	28
3.5.2	Το πρόβλημα της αποκωδικοποίησης - Αλγόριθμος Viterbi	29
3.5.3	Το πρόβλημα της εκπαίδευσης	29
3.5.3.1	Το κριτήριο Μέγιστης Πιθανοφάνειας (ML)	30
3.5.3.2	Βελτιστοποίηση παραμέτρων	31
3.5.3.3	Το κριτήριο Μέγιστης Αμοιβαίας Πληροφορίας (MMI)	32
4	Αναγνώριση φωνής με χρήση Κρυφών Μαρκοβιανών Μοντέλων	36
4.1	HTK (Hidden Markov Model Toolkit)	36
4.2	Γενικά Στοιχεία	37

4.3	Μεμονωμένη Αναγνώριση Λέξεων	38
4.4	Πιθανότητες Παρατήρησης (εξόδου)	41
4.5	Επανεκτίμηση Baum-Welch	42
4.6	Αναγνώριση και Αποκωδικοποίηση Viterbi	45
4.7	Αναγνώριση Συνεχούς Λόγου	46
5	Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines - SVMs)	50
5.1	Γενική περιγραφή	50
5.2	Πυρήνες του μοντέλου	52
5.2.1	Μη γραμμικές συναρτήσεις πυρήνα	52
III	Δεδομένα και μεθοδολογία	55
6	Πειραματικά δεδομένα	56
6.1	Διαδικασία εξαγωγής των κυματομορφών	56
6.2	Μορφή εξαγόμενων σημάτων	58
6.3	Κατηγορίες κυματομορφών	60
7	Ανάλογο του συστήματος αναγνώρισης φωνής για γενική αναγνώριση προτύπων	67
7.1	Προ-επεξεργασία Δεδομένων	67
7.1.1	Βήμα 1 - Η γραμματική	67
7.1.2	Βήμα 2 - Το λεξικό	68
7.1.3	Βήμα 3 - Δημιουργία αρχείων ετικέτας	69
7.1.4	Βήμα 4 - Κωδικοποίηση δεδομένων	70
7.2	Δημιουργία Μονοφωνικών Μοντέλων	71
7.2.1	Βήμα 5 - Δημιουργία μονοφωνικών μοντέλων	72
7.2.2	Βήμα 6 - Διόρθωση μέσω των μοντέλων "σιωπής"	73
7.2.3	Βήμα 7 - Εκ νέου ευθυγράμμιση των δεδομένων εκπαίδευσης	73
7.3	Κατασκευή Τριφώνων Συνδεδεμένων-Καταστάσεων	74
7.3.1	Βήμα 8 - Δημιουργία τριφώνων από μονόφωνα	74
7.3.2	Βήμα 9 - Δημιουργία τριφώνων συνδεδεμένων-καταστάσεων	76
7.4	Αξιολόγηση Επιπέδων Αναγνώρισης	78
7.4.1	Βήμα 10 - Αναγνώριση των δεδομένων αξιολόγησης	79
7.4.2	Σύνοψη διαδικασίας	80
8	Αξιολόγηση	82
8.1	Παράμετροι αξιολόγησης	82
8.2	Σύστημα αξιολόγησης	83
8.3	Οργάνωση πειραμάτων	84
8.3.1	Αλγόριθμος εύρεσης βέλτιστων τιμών κατωφλίων	84
8.3.2	Αλγόριθμος leave-one-out, cross validation	87

9	Επαλήθευση αποτελεσμάτων	89
9.1	Εκπαίδευση μοντελων SVM	89
9.1.1	Αλγόριθμος εκπαίδευσης	90
IV	Αποτελέσματα	92
10	Αποτελέσματα	93
10.1	Κρυφά Μαρκοβιανά Μοντέλα	93
10.2	Μικτά αποτελέσματα και σύγκλιση	94
10.3	Σύγκριση απόδοσης των δύο μοντέλων	98
10.4	Σύγκριση απόδοσης των δύο ομάδων ασθενών	99
10.5	Οπτική επιθεώρηση σε απόλυτα κατηγοριοποιημένες κυματομορφές	102
V	Συζήτηση και μελλοντικές προοπτικές	104
11	Συζήτηση	105
11.1	Αξιολόγηση Κρυφών Μαρκοβιανών Μοντέλων	105
11.2	Αξιολόγηση μικτών αποτελεσμάτων	106
11.3	Αξιολόγηση απόδοσης της κάθε ομάδας ασθενών	107
11.4	Τελικός ταξινομητής με χρήση συστήματος πλειοψηφίας	108
12	Επίλογος	111
12.1	Σύνοψη και συμπεράσματα	111
12.2	Μελλοντικές επεκτάσεις	112
	Βιβλιογραφία	113

Ευρετήριο εικόνων

1.1	Θρόμβος	10
1.2	Ισχαιμικό εγκεφαλικό επεισόδιο	11
1.3	Συσκευή υπερήχων	13
1.4	Υπέρηχος καρωτίδας	13
1.5	Υπέρηχος αθηρωματικής καρωτίδας	14
2.1	Επέκταση συστήματος για κατηγοριοποίηση ασθενών με καρωτιδική αθηρωμάτωση	18
3.1	Δομή Κρυφών Μαρκοβιανών Μοντέλων	25
3.2	Γενική μορφή λειτουργικής αρχιτεκτονικής ενός Κρυφού Μαρκοβιανού Μοντέλου	27
4.1	Στάδια επεξεργασίας στο HTK toolkit	36
4.2	Κωδικοποίηση/Αποκωδικοποίηση ηχητικού σήματος	38
4.3	Μεμονωμένο πρόβλημα αναγνώρισης λέξης	39
4.4	Παράδειγμα δημιουργίας Κρυφού Μαρκοβιανού Μοντέλου	40
4.5	Στάδιο εκπαίδευσης μοντέλου για μεμονωμένη αναγνώριση λέξης	41
4.6	Στάδιο αναγνώρισης για μεμονωμένες λέξεις	41
4.7	Αναπαράσταση μείγματος Gauss για τις συνιστώσες	42
4.8	Αλγόριθμος Viterbi για μεμονωμένη αναγνώριση λέξης	46
4.9	Δίκτυο μοντέλων για αναγνώριση συνεχούς λόγου	48
4.10	Καταγραφή πιθανών ορίων λέξεων	49
5.1	SVM και η ευθεία διαχωρισμού σε μία διάσταση	51
5.2	SVM και τα όρια γύρω από το επίπεδο διαχωρισμού	51
5.3	Πολυωνυμικός πυρήνας μετασχηματισμού SVM	53
5.4	Πυρήνας Ακτινικής Βάσης μετασχηματισμού SVM	53
6.1	Τυχαία κυματομορφή μέσα από τα πειραματικά δεδομένα	58
6.2	5 ασθενείς της ίδιας ομάδας για τη κατηγορία κυματομορφών (μέση γωνία ταχύτητας, ολόκληρη πλάκα)	59
6.3	5 ασθενείς της ίδιας ομάδας για τη κατηγορία κυματομορφών (διάμεση ολική ταχύτητα, οπίσθιο τοίχωμα)	59
6.4	Σχήμα κατασκευής των 146 κυματομορφών	66

7.1	Γραμματική του μοντέλου	68
7.2	Μοντέλο 3-καταστάσεων, αριστερά προς δεξιά	73
7.3	Μοντέλο σιωπής	74
7.4	Σύνδεση πινάκων μετάβασης καταστάσεων για τρίφωνα	76
7.5	Διάγραμμα με τα βασικά συστατικά και τη διασύνδεση τους για την κατασκευή του συστήματος	81
8.1	Διάγραμμα που περιγράφει τις κύριες διαδικασίες και επαναλήψεις του πειράματος εύρεσης των βέλτιστων κατωφλίων αποκοπής	85
8.2	Διάγραμμα που περιγράφει τις κύριες διαδικασίες και επαναλήψεις του cross validation	88
9.1	Διάγραμμα εκπαίδευσης και επαλήθευσης SVM	91
10.1	Διάγραμμα που δίνει τα αποτελέσματα του cross validation - HMM/SVM	98
10.2	Ακρίβεια HMM για τις 2 ομάδες ασθενών, για όλες τις κατηγορίες	100
10.3	Ακρίβεια SVM για τις 2 ομάδες ασθενών, για όλες τις κατηγορίες	101
10.4	Κυματομορφές της κατηγορίας 89 για τις 2 ομάδες ασθενών	102
10.5	Κυματομορφές της κατηγορίας 90 για τις 2 ομάδες ασθενών	103

Ευρετήριο πινάκων

1.1	Συγκεντρωτικός πίνακας χαρακτηριστικών υφής και ταξινομητών	16
10.1	Πίνακας που δείχνει τη δομή των αποτελεσμάτων μετά το cross validation	93
10.2	Παράδειγμα πίνακα μικτών αποτελεσμάτων όλων των μοντέλων	95
10.3	Πίνακας μικτών αποτελεσμάτων όλων των μοντέλων για όλες τις κατηγορίες	97
11.1	Πίνακας αποτελεσμάτων με βάση τις ομάδες των ασθενών	107
11.2	Υπόδειγμα πίνακα υπολογισμού ψήφων, συστήματος πλειοψηφίας	109
11.3	Πίνακας συστήματος πλειοψηφίας, ομάδα 1	109
11.4	Πίνακας συστήματος πλειοψηφίας, ομάδα 2	109

Πηγαίος κώδικας

12.1	exports.m	117
12.2	encoding_stage.m	125
12.3	timescale.m	130
12.4	findmin.m	132
12.5	findmax.m	132
12.6	err_rep.m	134
12.7	makedir.m	134
12.8	result.sh	136
12.9	cross.sh	138
12.10	svm.m	139
12.11	find_mean_val.m	146

Μέρος Ι

Επισκόπηση - Βασικά στοιχεία μελέτης

Κεφάλαιο 1

Εισαγωγή

Σε αυτό το εισαγωγικό κεφάλαιο παρουσιάζεται η πάθηση της καρωτιδικής αθηρωμάτωσης αλλά και διαδικασία που ενεργοποιείται πριν από ένα εγκεφαλικό επεισόδιο. Αναφέρονται στατιστικά δεδομένα ασθενών της Ευρώπης αλλά και των Η.Π.Α που πάσχουν από αρτηριακή νόσο και παρουσιάστηκαν τα ποσοστά θνησιμότητας αλλά και χρηματικών δαπανών. Στη συνέχεια γίνεται λόγος για την υπερηχογραφική απεικόνιση και η δυνατότητα υποστήριξης της διάγνωσης μέσω της σύστασης αλλά και των χαρακτηριστικών κίνησης της αθηρωματικής πλάκας. Αξίζει να αναφερθεί ότι η εργασία αυτή είναι μία από τις πρώτες προσπάθειες ταξινόμησης με χαρακτηριστικά κίνησης και μάλιστα η πρώτη που διερευνά ολόκληρα μοτίβα κίνησης.

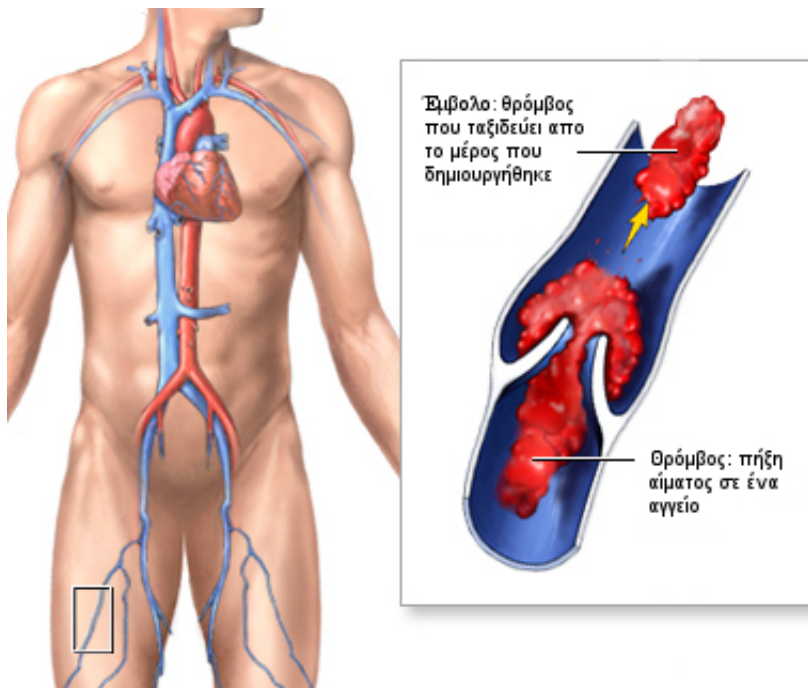
Μία από τις τρεις κυριότερες παθήσεις που οδηγούν στατιστικά σε θάνατο παγκοσμίως είναι η αγγειακή νόσος των καρωτίδων μαζί με την καρδιαγγειακή νόσο και τον καρκίνο. Κάθε χρόνο στις ΗΠΑ 795,000 άνθρωποι προσβάλλονται από ένα νέο ή επαναλαμβανόμενο εγκεφαλικό επεισόδιο [1]. Κατά προσέγγιση, 610.000 (76.7%) από αυτούς έχουν πρωτοεμφανιζόμενο επεισόδιο, ενώ για τους υπόλοιπους 185.000 (23.3%) πρόκειται για επαναλαμβανόμενο επεισόδιο. Το 2009 υπολογίστηκε πως στις Η.Π.Α κάθε 40 δευτερόλεπτα, συμβαίνει ένα εγκεφαλικό επεισόδιο και κάθε 4 λεπτά υπάρχει και ένας θάνατος λόγω αυτού. Το μέσο κόστος για άμεση ή έμμεση δαπάνη για την καρδιαγγειακή νόσο και τα εγκεφαλικά επεισόδια είναι περί τα \$312.6 εκατομμύρια (€226.7 εκατομμύρια), συμπεριλαμβάνοντας δαπάνες για ιατρούς, νοσοκομειακή περίθαλψη, φαρμακευτικές αγωγές, διαγνωστικές εξετάσεις, αλλά και έμμεσα κόστη από την απώλεια παραγωγικότητας στο εργατικό δυναμικό λόγω θανάτου.

Στην Ευρώπη μέχρι και το 2008 υπολογίστηκε πως η απώλεια ζωής από εγκεφαλικό επεισόδιο αυξήθηκε και έχει καταμετρηθεί ότι 200,000 άνδρες και σχεδόν 300.000 γυναίκες χάνουν την ζωή τους στην Ευρωπαϊκή Ένωση κάθε χρόνο [2]. Ετησίως στην Ε.Ε έχουμε 4.30 εκατομμύρια θανάτους από καρδιακές και αγγειακές παθήσεις που είναι ο κυρίαρχος λόγος θανάτου συνολικά στον πληθυσμό της. Σχεδόν οι μισοί (48%) από το σύνολο των θανάτων προέρχονται από καρδιαγγειακή νόσο (από τους οποίους το 54% είναι γυναίκες και το 43% άνδρες). Οι μισοί των θανάτων από καρδιαγγειακή νόσο τοποθετούνται στην στεφανιαία αρτηριακή νόσο ενώ το 1/3 αυτών στο

εγκεφαλικό επεισόδιο. Τα εγκεφαλικά επεισόδια καταλογίζονται ως δεύτερη αιτία θανάτων στην Ευρώπη με 1.24 εκατομμύρια ανθρώπων να χάνουν την ζωή τους ετησίως. Κοστίζει δε, στην οικονομία της Ε.Ε πάνω από €38 δισεκατομμύρια τον χρόνο, περίπου το 1/6 της συνολικής δαπάνης για καρδιαγγειακά προβλήματα. Σε ποσοστά κόστους έχουμε 49% σε δαπάνη υγείας, 23% σε παραγωγικά κόστη (απώλεια εργατικού δυναμικού) και 29% σε άτυπες δαπάνες για την φροντίδα ανθρώπων με εγκεφαλικά επεισόδια.

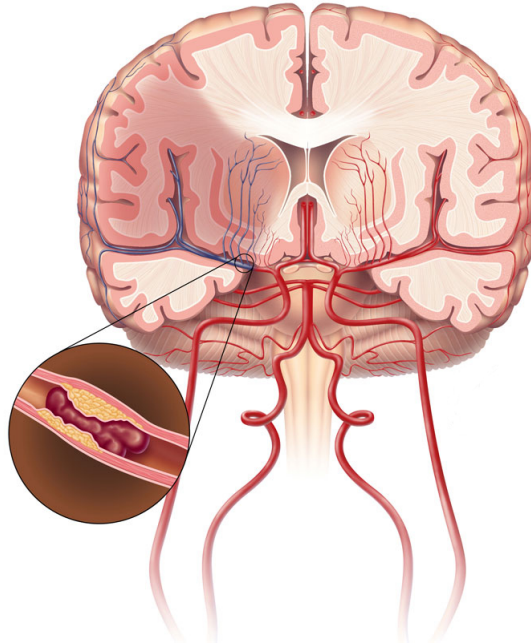
1.1 Εγκεφαλικό Επεισόδιο

Κάτω από κανονικές συνθήκες η γλυκόζη είναι το μόνο συστατικό που μεταβολίζεται από τον εγκέφαλο για να παρέχει την απαραίτητη ενέργεια επάρκεια. Δεδομένου ότι οι αποθηκευτικές ικανότητες του εγκεφάλου σε γλυκόζη είναι περιορισμένες, εξαρτάται άμεσα από την αδιάκοπη παροχή αίματος, δια της αρτηριακής οδού, οξυγόνου και γλυκόζης. Είναι γεγονός ότι η κύρια παρατηρούμενη εγκεφαλική βλάβη προέρχεται από μειωμένη παροχή αίματος σε κάποια περιοχή του εγκεφάλου. Όταν οι νευρώνες στην περιοχή αυτή μείνουν για κάποια λεπτά χωρίς τα απαραίτητα θρεπτικά συστατικά και οξυγόνο, σταματούν την λειτουργία τους και στην συνέχεια πεθαίνουν [3]. Αυτός ο νευρωνικός θάνατος αποκαλείται εγκεφαλικό επεισόδιο. Γνωρίζουμε φυσικά πως δεν υπάρχει η δυνατότητα επέκτασης/ανάπλασης/αναγέννησης του νευρωνικού δικτύου του εγκεφάλου, πράγμα που σημαίνει, μη αναστρέψιμη εγκεφαλική βλάβη ή και σε περιπτώσεις θάνατος του ασθενή. Υπάρχουν δύο βασικά είδη επεισοδίων, το ισχαιμικό και το εμβολικό εγκεφαλικό επεισόδιο.



Εικόνα 1.1: Θρόμβος που δημιουργείται και διαδικασία εμβολής από το σημείο δημιουργίας του. [Πηγή, U.S. National Library of Medicine, National Institutes of Health (<http://www.nlm.nih.gov/>)]

Ισχαιμικό εγκεφαλικό επεισόδιο συμβαίνει όταν η παροχή αίματος σε τμήμα του εγκεφάλου διακόπτεται, με αποτέλεσμα τα εγκεφαλικά κύτταρα να πεθάνουν. Η ροή του αίματος μπορεί να διακοπεί κυρίως με δύο μηχανισμούς: τη *θρόμβωση* και την *εμβολή* (εικόνα 1.1). Αντιστοιχεί στο 85% του συνόλου των επεισοδίων.



Εικόνα 1.2: Θρόμβος κοντά σε περιοχή στένωσης, λόγω αθηρωμάτωσης, που περιορίζει την ροή αίματος σε περιοχή εγκεφάλου (γκρίζα περιοχή). [Πηγή, The Internet Stroke Center (<http://www.strokecenter.org/>)]

Εμβολικό εγκεφαλικό επεισόδιο ορίζουμε την διαδικασία όπου ένας θρόμβος αίματος ή ένα αποκολλημένο κομμάτι αθηρωματικής πλάκας μεταφέρεται δια της αιματικής ροής λόγω καρδιακής αρρυθμίας, όπως μπορεί να συμβεί κατά την κολπική μαρμαρυγή. Συνήθως, αυτοί οι θρόμβοι παραμένουν στο εσωτερικό τοίχωμα της καρδιάς, αλλά περιστασιακά μπορεί να αποκολληθούν και ταξιδεύοντας με τη ροή του αίματος να αποφράξουν μια αρτηρία του εγκεφάλου (εικόνα 1.2). Το έμβολο μπορεί επίσης να προέρχεται από μια μεγάλη αρτηρία όπως στις κοινές καρωτίδες αρτηρίες, δύο σημαντικές αρτηρίες στον λαιμό που παρέχουν αίμα στον εγκέφαλο.

1.2 Καρωτιδική αθηρωμάτωση και Υπερηχογραφική απεικόνιση

Η καρωτιδική αθηρωμάτωση είναι μια αρτηριακή πάθηση που χαρακτηρίζεται από την σκλήρυνση μέρους του αρτηριακού τοιχώματος της καρωτίδας κοντά στο ενδοθήλιο από την δημιουργία πλακών και αποτελεί την κύρια αιτία πρόκλησης εγκεφαλικού επεισοδίου. Οι πλάκες αυτές αποτελούνται από (1) μεγάλο αριθμό ανώμαλων λείων μυικών κυττάρων, μακροφάγων και λεμφοκυττάρων, (2) συγκεντρώσεις χοληστερόλης και άλλων λιπιδίων, (3) παχιά στρώματα συνδετικού ιστού [3]. Με αυτό τον τρόπο, την στένωση δηλαδή του αυλού, αυξάνει η αντίσταση στην ροή

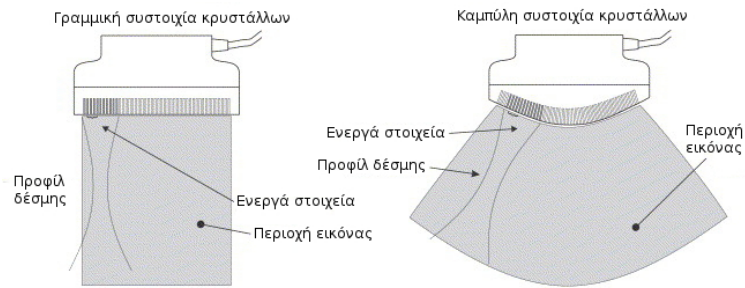
του αίματος, τα δυσλειτουργικά ενδοθηλιακά κύτταρα στην αθηροσκληρωτική περιοχή με την έκκριση ουσιών, όπως είναι το οξείδιο του αζώτου (NO, ουσία που παράγεται από την L-αργινίνη), η προστακυκλίνη και ο παρερχόμενος από το ενδοθήλιο υπερπολωτικός παράγοντας (EDHF), το ενδοθήλιο ασκεί χαλαρωτική δράση στα λεία μυϊκά κύτταρα των αγγείων, αυξάνει τη διάμετρό τους και άρα και την ποσότητα του αίματος που αυτά προωθούν στην περιφέρεια. Ταυτόχρονα όμως, παράγει και αγγειοσυσπαστικές ουσίες, όπως είναι η ενδοθηλίνη, η θρομβοξάνη και η αγγειοτενσίνη, ρυθμίζοντας έτσι ελαττωματικά τον τόνο και τη διάμετρο των αγγείων ανάλογα με τις εκάστοτε ανάγκες των ιστών.

Η ιστολογική ανάλυση της αθηρωματικής πλάκας και οι μηχανικές της ιδιότητες είναι μέγιστης σημασίας για την κατανόηση της διαδικασίας σχηματισμού αλλά και της μερικής ρήξης αυτής, με αποτέλεσμα την απόφραξη γειτονικών αρτηριών. Από τις διάφορες *παραμέτρους κατηγοριοποίησης αθηρωματικών πλακών* που έχουν προταθεί οι κύριες είναι μέσω των *λιπιδίων (lipid)*, του *ινώδους ιστού (fibrin)*, της *αιμορραγίας (haemorrhage)*, των *ασβεστομάτων (calcium)* που περιέχουν αλλά και οι συνδυασμοί των ανωτέρω, σε διάφορες περιεκτικότητες έκαστο. Έχει παρατηρηθεί σε δείγματα αμέσως μετά από χειρουργική αφαίρεση των πλακών από τις καρωτίδες ασθενή [4], ότι οι πλάκες που περιέχουν ασβεστόματα είναι οι πιο δύσπλαστες και πως οι υπόλοιπες κατηγορίες είναι τουλάχιστον 2 φορές πιο εύπλαστες. Γενικότερα, η συμπεριφορά των διαφορετικών δειγμάτων είναι σχετικά απρόβλεπτη, αλλά έχει παρατηρηθεί [5--7] σε συμπτωματικούς ασθενείς ότι οι πλάκες που είναι πλούσιες σε λιπίδια ή/και αίμα, (οι οποίες και εμφανίζονται διαφανείς στους υπερήχους) έχουν την υψηλότερη πιθανότητα για να δημιουργήσουν εγκεφαλικό επεισόδιο μελλοντικά. Είναι βεβαίως δεδομένο ότι το μέγεθος της αθηρωματικής πλάκας είναι ευθέως ανάλογο με την στένωση του αυλού που την περιέχει και ότι παράλληλα αυξάνονται σημαντικά οι πιθανότητες για ρήξη της αθηρωματικής πλάκας καθώς το μέγεθός της αυξάνει.

Ασυμπτωματικοί ασθενείς θεωρούνται συνήθως οι ασθενείς που δεν έχουν εμφανίσει επεισόδιο κατά τους τελευταίους 6 μήνες.

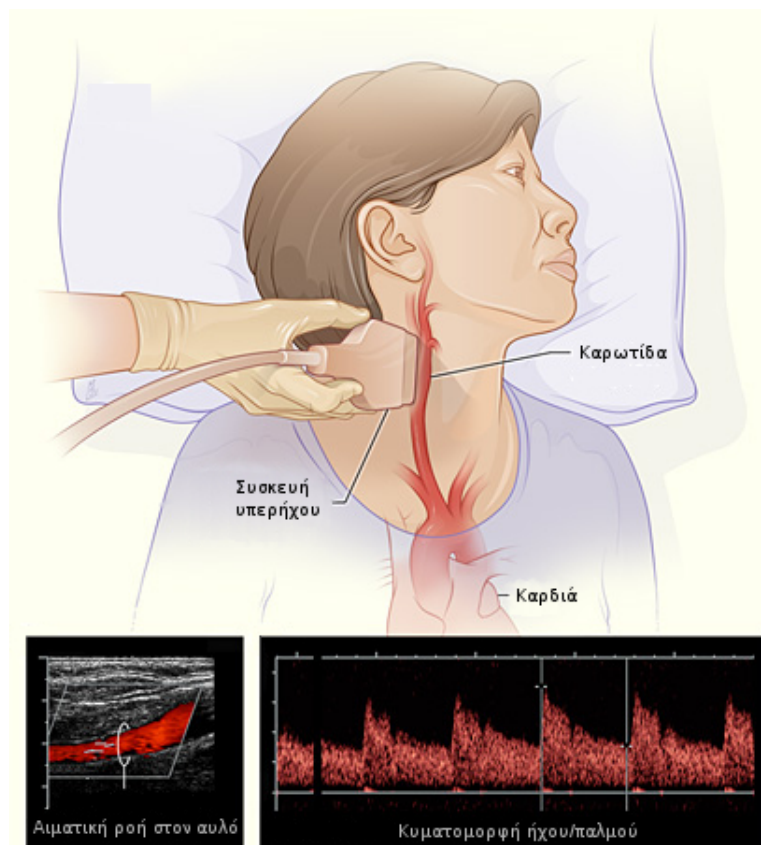
Η χρήση *υπερήχων β-σάρωσης* είναι η συνήθης μέθοδος απεικόνισης καρωτίδας. Είναι μια μη επεμβατική μέθοδος απεικόνισης η οποία λόγω της φύσεως της έχει τα θετικά της στοιχεία όπως και τα αρνητικά. Η ασφάλεια του ασθενή λόγω της μετάδοσης ήχου και όχι ηλεκτρομαγνητικής ακτινοβολίας από την συσκευή όπως και η άνεσή του, κατά την εξέταση είναι τα θετικά στοιχεία, απ' την άλλη πλευρά η υψηλή παραμόρφωση του σήματος αλλά και η περιορισμένη διακριτική ικανότητα δυσκολεύει την μελέτη ανθρωπίνου ιστού.

Η συσκευή εκμεταλλεύεται το πιεζοηλεκτρικό φαινόμενο που έχει παρατηρηθεί σε κρυστάλλους. Οι κρυστάλλοι αυτοί μπορούν να συντονίζονται σε υψηλές ηχητικές συχνότητες (πέραν του ακουστικού φάσματος) της τάξεως των 2-18 MHz.



Εικόνα 1.3: Απεικόνιση μέσω μετατροπέα υπερήχων.

Το αποτέλεσμα του συντονισμού τους είναι η μετατόπιση του εσωτερικού τους φορτίου, όπου με τη σειρά του δημιουργεί μια μεταβαλλόμενη τάση στις δύο πλευρές (δημιουργία δίπολου). Αυτή την τάση εκμεταλλεύεται η διάταξη για να λειτουργήσει. Σκοπός της είναι η απεικόνιση αντικειμένων μέσω ανάκλασης ηχητικών σημάτων (αντήχηση - echo) πάνω τους. Η διαφοροποίηση των ιστών κατά την απεικόνιση γίνεται σε σχέση με την αντήχηση του κάθε ιστού, η διαφορά αυτή λέγεται ακουστική εμπέδηση.

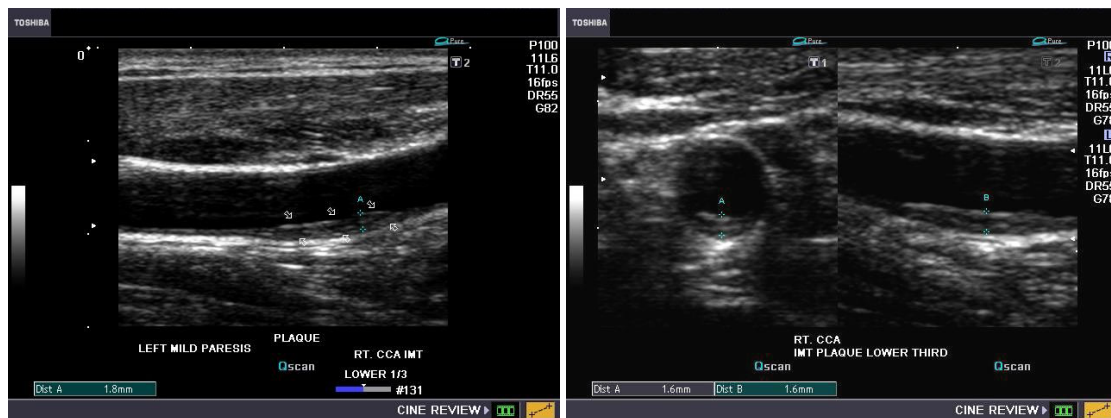


Εικόνα 1.4: Εικόνα υπερήχων (Doppler) καρωτίδας, η αιματική ροή στον αυλό φαίνεται με κόκκινο χρώμα. Δίπλα του απεικονίζεται ο καρδιακός παλμός μέσα από τον υπέρηχο. [Πηγή, The Worlds of David Darling - Encyclopedia of Science (<http://www.daviddarling.info/encyclopedia/>)]

Στην περίπτωση των διαγνωστικών υπερήχων για καρωτίδες, έχουμε χρήση κυρίως γραμμικής συστοιχίας κρυστάλλων (εικόνα 1.3). Κατά την εφαρμογή ομάδες στοιχείων διεγείρονται,

δημιουργώντας δέσμες με σειριακό τρόπο ώστε να σταλούν υπερηχητικά σήματα από όλα τα στοιχεία. Με την επιλεκτική χρονοκαθυστέρηση εστιάζεται η δέσμη και γι' αυτό εμφανίζεται το προφίλ κοίλο. Το βάθος συνδέεται άμεσα με την χρονική υστέρηση του εκπεμπόμενου παλμού, άρα επιλέγοντας συγκεκριμένες χρονοκαθυστερήσεις επιλέγουμε συγκεκριμένο βάθος αναπαραγωγής. Η ανακλώμενη, αλλοιωμένη, κατά φάση και πλάτος κυματομορφή μας διαχωρίζει το κάθε υλικό σύμφωνα με τη σύνθεση του. Η β-σάρωση (Brightness mode) απεικονίζει μια τομή επίπεδου σε δύο διαστάσεις (2D) και διαχωρίζει τα υλικά με διακριτή χρήση της φωτεινότητας, σε κλίμακα του γκρι (greyscale).

Η δομική ανάλυση αθηρωματικής πλάκας καρωτίδας μέσω υπερήχων (εικόνα 1.4) έχει χρησιμοποιηθεί κυρίως λόγω του μειωμένου κόστους και της μη επεμβατικής φύσης της. Η μετά-ανάλυση (post processing) των εικόνων (B-mode, εικόνα 1.5) για την κατηγοριοποίηση της σύστασης και του σχήματος έχει πραγματοποιηθεί με σχετική επιτυχία [8--12] με ποσοστά αναγνώρισης σύστασης περί του 55-78% για το ασβέστιο, για τον ινώδη ιστό 40-86% και 69-75% για τα λιπίδια που αποτελείται η πλάκα. Στοιχεία βέβαια που έχουν ακραίες διαφορές στο κάθε δείγμα αλλά και στην κάθε αλγοριθμική τεχνική αναγνώρισης κάθε μελέτης. Η ανάλυση της υφής και της σύστασης της πλάκας όπως και προαναφέρθηκε γίνεται σύμφωνα με την αντήχηση που επιστρέφει κάθε ουσία και είναι σκόπιμο εδώ να αναφερθεί πως οι αθηρωματικές πλάκες που περιέχουν λιπίδια ή αιμοφόρα αγγεία δείχνουν περισσότερο διαφανείς (διαπερατές) στην αντήχηση του υπερήχου, ενώ στην περίπτωση ασβεστομάτων ή ινωδών πλακών βλέπουμε φανερά την αντήχηση στον υπέρηχο.



Εικόνα 1.5: Μεσήλικας άνδρας ασθενής, με πρόσφατο μεταβατικό ισχαιμικό εγκεφαλικό επεισόδιο. Εικόνα β-σάρωσης δείχνει αθηρωματική πλάκα στο κατώτερο 1/3 της κοινής καρωτίδας, σε διαμήκη αλλά και ακτινική τομή του αυλού. Πλάκα που δεν περιέχει επί το πλείστον ασβεστόματα. [Πηγή, [Ultrasound image gallery gfgf\(http://www.ultrasound-images.com/\)](http://www.ultrasound-images.com/)]

1.2.1 Ο ρόλος των μηχανικών ιδιοτήτων του αρτηριακού τοιχώματος

Πέρα από την στατική αναγνώριση της σύστασης, μεγάλο ενδιαφέρον τα τελευταία χρόνια έχει παρατηρηθεί στην ανίχνευση κίνησης μέσω υπερήχων β-σάρωσης του αρτηριακού τοιχώματος. Μάλιστα η υποστήριξη της διάγνωσης της καρωτιδικής αθηρωμάτωσης, με τη βοήθεια χαρακτηριστικών που εξάγονται από εικόνες υπερήχων και αντιπροσωπεύουν τη σύσταση και τις μηχανο-

νικές ιδιότητες του αρτηριακού τοιχώματος, εντάχθηκε πρόσφατα στις μεγαλύτερες προκλήσεις στον τομέα της βιοϊατρικής τεχνολογίας [13].

Οι *παράμετροι κίνησης* εκφράζουν δύο κύριες μετατοπίσεις, την *διαμήκη μετατόπιση* και την *εγκάρσια ή ακτινική μετατόπιση*. Η καταγραφή γίνεται συνήθως για αρκετούς καρδιακούς κύκλους αλλά επιλέγονται για ανάλυση οι πιο πλήρεις κύκλοι και αναγνωρίζεται η συστολή (ή διαστολή) ως σημείο έναρξης [14]. Κύριοι παράγοντες είναι το *πλάτος και η ταχύτητα της μετατόπισης*, η *διαμηθικές τάσεις*, όπως και οι *συμπιέσεις* του αυλού, σε εγκάρσιες και διαμήκεις κατευθύνσεις.

Η ανάλυση κίνησης, μερών της κοινής καρωτίδας αρτηρίας, υγιών υποκειμένων, παρουσίασε μια σημαντική *διαμήκη συνιστώσα παραμόρφωσης* στα τοιχώματα συγκρίσιμη σε τάξη μεγέθους με την ακτινική συνιστώσα. Τα μοτίβα που προκύπτουν από τις *διαμήκεις μετατοπίσεις* του αρτηριακού τοιχώματος βρέθηκαν να είναι αρκετά διαφορετικά ανάμεσα σε υγιή υποκείμενα κοντινών ηλικιών και ίδιων φύλλων, αλλά σταθερά σε επανεξέταση μετά από τέσσερις μήνες.

Η επιθεώρηση της κίνησης της αθηρωματικής πλάκας είναι δυσκολότερη από αυτή του αρτηριακού τοιχώματος λόγω της ιδιαίτερης γεωμετρίας της και των ασύμμετρων ορίων της, αλλά ταυτόχρονα πολύ ενδιαφέρουσα κλινικά. Λίγες μονάχα μελέτες έχουν συσχετίσει μοτίβα κίνησης της αθηρωματικής πλάκας με το ρίσκο για εγκεφαλικό επεισόδιο. Οι ασυμπτωματικές πλάκες φάνηκε να έχουν *επιφανειακά διανύσματα κίνησης* με ίδια κατεύθυνση και μέτρο σε σχέση με αυτά της καρωτίδας αρτηρίας, σε αντίθεση με τις συμπτωματικές πλάκες όπου παρουσίασαν ενδείξεις αυτόνομων χαρακτηριστικών. Η *μέγιστη επιφανειακή ταχύτητα* δεν ήταν διαφορετική μεταξύ των δύο τύπων αθηρωματικών πλακών αλλά η *μέγιστη αποκλίνουσα ταχύτητα*, που ορίζεται ως η μέγιστη διαφορά μεταξύ μεγίστων και ελαχίστων ταχυτήτων, ήταν αισθητά μεγαλύτερη στις συμπτωματικές πλάκες.

1.2.2 Επισκόπηση των εργασιών για την υποστήριξη της διάγνωσης

Η ανάλυση που γίνεται σε ακτινική κατεύθυνση έχει αναλυθεί [8, 15, 16] και έχει εξάγει αποτελέσματα για την συσχέτισή της με την επικινδυνότητα και το στάδιο της νόσου. Η ανάλυση στην διαμήκη κατεύθυνση είναι λιγότερο αναπτυγμένη και εμφανίζει δυσκολότερες προκλήσεις λόγω της χαμηλής ποιότητας του σήματος κυρίως. Ακόμη δυσκολότερο πρόβλημα είναι η ανάλυση κίνησης της αθηρωματικής πλάκας, σε σχέση με την ανάλυση κίνησης του αρτηριακού τοιχώματος μέσω υπερήχου. Στοιχεία κίνησης της πλάκας όπως η μεταβολή της ταχύτητας, η σχετική κίνηση σε σχέση με τα τοιχώματα του αυλού (έσω και έξω χιτώνας), η κίνηση αλλά και η γωνία αυτής (ως διάνυσμα) σε συγκεκριμένα σημεία της πλάκας όπως η άνω και η κάτω επιφάνεια. Έτσι μπορεί έμμεσα να προβλεφθεί η πιθανότητα ρήξης (αποκόλλησης κομματιού) της πλάκας από την διαφοροποίηση των ίδιων των πλακών, σύμφωνα με το στάδιο ζωής τους (δηλαδή την εσωτερική και εξωτερική τους δομή) και την κίνηση που συσχετίζεται με το κάθε στάδιο. Εάν μπορέσει να παρατηρηθεί με ακριβή τρόπο η διαφορά μεταξύ συμπτωματικών και ασυμπτωματικών αθηρωμα-

τικών πλακών θα μπορέσει να συνεισφέρει φυσικά σε μεγάλο βαθμό στην πρόληψη μελλοντικών επεισοδίων.

Ο κύριος όγκος μελετών βασίζονται σε χαρακτηριστικά υφής (μέσω διαφορών αντήχησης στη σύσταση της πλάκας) και σε χαρακτηριστικά μορφολογίας με τους ταξινομητές που παρουσιάζονται αμέσως μετά (πίνακας 1.1). Η περίπτωση της κατηγοριοποίησης με παραμέτρους κίνησης έχει μελετηθεί μονάχα μερικώς.

Χαρακτηριστικά υφής	Ακρίβεια	Ταξινομητές	Ακρίβεια
Στατιστικές ν-τάξης [17--20]	72-90%	SOM (Self Organizing Map) [17]	73%
Μήτρες χωρικής εξάρτησης των τόνων του γκρι [17]	67-70%	kNN (k-nearest neighbors) [17]	69%
Στατιστικές παράμετροι διαφορών φωτεινότητας [17, 21]	64-67%	PNN (Probabilistic neural networks) [18, 19, 21]	67-93%
Πίνακας γειτονικών διαφορών επιπέδων του γκριζου [17]	63-68%	SVM (Support vector machines) [19--21]	73-90%
Πίνακες στατιστικών χαρακτηριστικών [17]	60-65%		
Μέθοδος της υφής της ενέργειας του Laws [17, 18]	72-99%		
Ανάλυση Fractal [17]	69-72%		
Φάσμα ισχύος Fourier [17]	59-63%		
Μορφολογία [17]	73-77%		
Πολυκλιμακωτή ανάλυση	73-90%		
Χαρακτηριστικά κίνησης			
Μέγιστη επιφανειακή ταχύτητα πλάκας [22]	-		
Μέγιστη σχετική επιφανειακή ταχύτητα [22]	-		

Πίνακας 1.1: Συγκεντρωτικός πίνακας χαρακτηριστικών υφής και ταξινομητών.

Είναι φανερό ότι δεν υπάρχουν επαρκή ερευνητικά αποτελέσματα για την κατηγοριοποίηση ασθενών με βάση τις κινηματικές και μηχανικές ιδιότητες της αθηρωματικής πλάκας και έτσι αυτή η εργασία έρχεται να συμπληρώσει ένα κενό σε αυτό το πεδίο έρευνας.

Η πρώτη προσπάθεια ανάπτυξης ταξινομητή με χαρακτηριστικά κίνησης έγινε το 2013, που χρησιμοποιήθηκαν οι ακόλουθοι ταξινομητές, *SVM*, *kNN*, *BDT* - δυαδικά δέντρα αποφάσεων και *DA* - διακριτική ανάλυση [23]. Στην εργασία αυτή αναλύθηκαν τα δεδομένα 56 ασθενών, το οποίο ήταν χωρισμένο σε δύο ισάριθμα σύνολα συμπτωματικών/ασυμπτωματικών ασθενών. Για την κατηγοριοποίηση (διαφοροποίηση) των ασθενών μέσω χαρακτηριστικών κίνησης, υπήρξε μεγάλη επιτυχία στη χρήση *Support Vector Machines* (*Μηχανές Διανυσμάτων Υποστήριξης*) με ποσοστά αναγνώρισης ασθενών πάνω από 80% αλλά και αλγόριθμοι όπως *k*-πλησιέστεροι γείτονες, *Διακριτικής Ανάλυσης* και *Δέντρων Αποφάσεων* όπου δίνουν μεν χαμηλότερα ποσοστά αναγνώρισης αλλά όχι απογοητευτικά. Εδώ πρέπει να τονιστεί ότι η εργασία αυτή εξέταζε με στατικό τρόπο τα χαρακτηριστικά και όχι σαν κυματομορφές στο πεδίο του χρόνου.

Αυτή είναι και η διαφορά του αντικειμένου της παρούσας εργασίας σε σχέση με κάθε προηγούμενη. Διότι, για πρώτη φορά αναπτύσσεται ταξινομητής που δρα σε δεδομένα που αφορούν κινηματικές και μηχανικές ιδιότητες αθηρωματικής πλάκας, για την διαφοροποίηση συμπτωματικών και ασυμπτωματικών ασθενών που πάσχουν από καρωτιδική αθηρωμάτωση.

Κεφάλαιο 2

Αντικείμενο και δομή της εργασίας

2.1 Αντικείμενο

Σκοπός αυτής της εργασίας είναι η αξιολόγηση πειραματικών δεδομένων, με την χρήση Κρυφών Μαρκοβιανών Μοντέλων (Hidden Markov Models) για την ταύτιση συγκεκριμένων μοτίβων κίνησης και παραμόρφωσης, με την συμπτωματική ή ασυμπτωματική φύση των αθηρωματικών πλακών καρωτίδας.

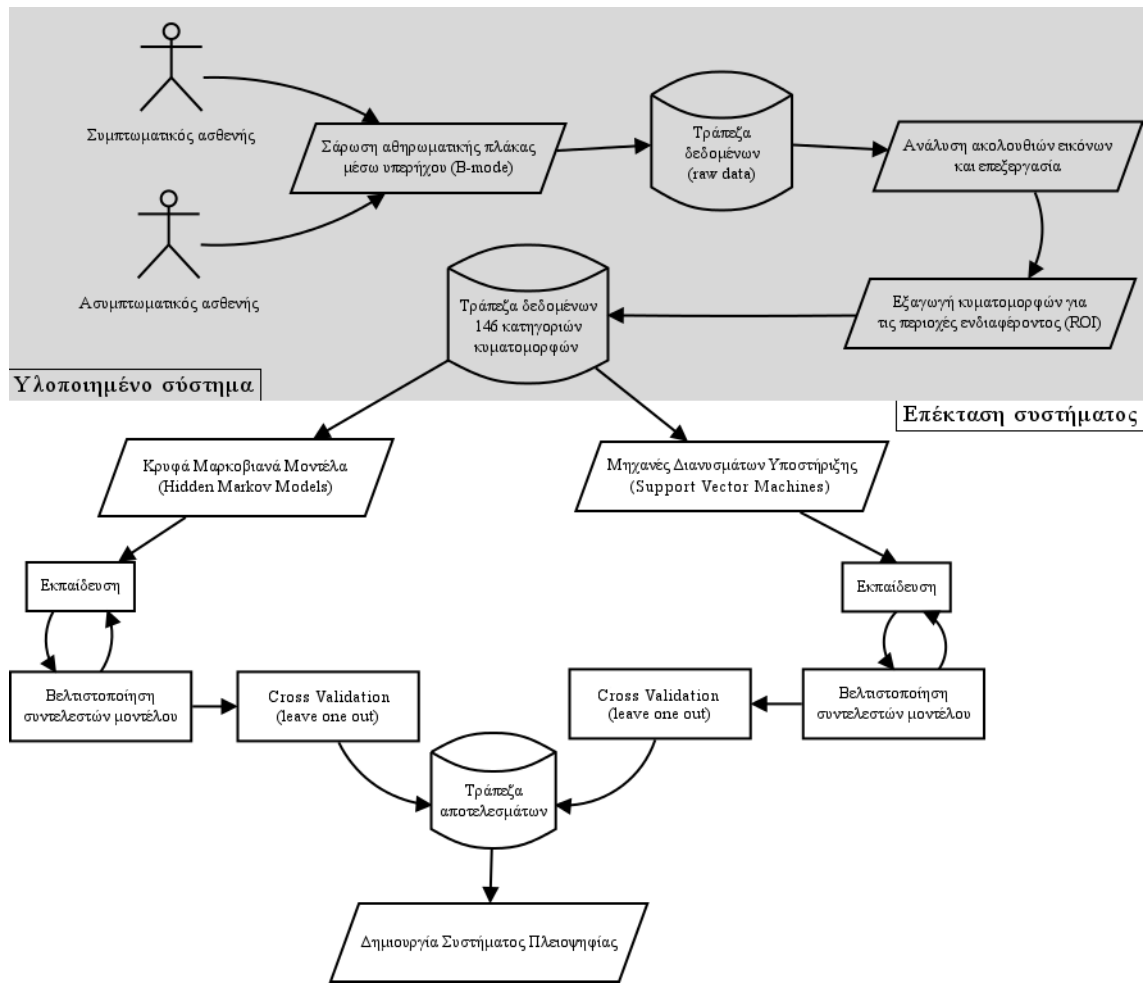
Τα διαθέσιμα δεδομένα αφορούν κυματομορφές που αντιπροσωπεύουν κινήσεις της αθηρωματικής πλάκας και του αρτηριακού τοιχώματος πλησίον αυτής, λαμβανομένων με υπερηχογράφημα β-σάρωσης σε ασθενείς που πάσχουν από αθηρωμάτωση σε καρωτίδα και είναι είτε συμπτωματικοί, είτε ασυμπτωματικοί. Προέκυψαν έτσι 146 κατηγορίες κυματομορφών για κάθε έναν από τους 96 ασθενείς, οι οποίες αναφέρονται σε κινήσεις συνολικά της αθηρωματικής πλάκας και του υγιούς τοιχώματος, σχετικές κινήσεις της πλάκας με αναφορά το υγιές τοίχωμα της αρτηρίας, αλλά και παράγωγα μεγέθη της κίνησης όπως η ταχύτητα, γωνία κτλ. Οι κυματομορφές διαφέρουν σε διάρκεια καθώς απεικονίζονται στον χρονικό άξονα με βάση τον καρδιακό παλμό, όπου είναι δεδομένο ότι διαφέρει από ασθενή σε ασθενή, έχουν όμως κοινή έναρξη την καρδιακή διαστολή (ή ισοδύναμα την συστολή του αγγείου).

Μέσα από την εμπειρία που έχει αποκτηθεί, μετά από τη χρήση των Κρυφών Μαρκοβιανών Μοντέλων στην αναγνώριση προτύπων και κυρίως στην αναγνώριση φωνής, μας δίνεται ένα μεγάλο θεωρητικό αλλά και προγραμματιστικό υπόβαθρο για τη δημιουργία ενός ευθέως ανάλογου συστήματος αναγνώρισης προτύπων με εφαρμογή στα 146 κατηγορίες κυματομορφών που προαναφέρθηκαν. Αυτό γίνεται με τη κατάστρωση διαφόρων σεναρίων προεπεξεργασίας των κυματομορφών ως προς το πλάτος αλλά και την διάρκεια τους, έτσι ώστε να υπάρξει η κατάλληλη κανονικοποίηση πριν την εισαγωγή των δεδομένων στο μοντέλο.

Μετά από μια σειρά βελτιστοποιήσεων των παραμέτρων του μοντέλου, ξεχωριστά για κάθε είδος, προκύπτουν τα αποτελέσματα τα οποία και διασταυρώνονται (cross validated) για να υπάρξει

μα αντικειμενικότερη απεικόνιση των δυνατοτήτων του μοντέλου. Στη συνέχεια για την επαλήθευση των αποτελεσμάτων αυτών, κατασκευάζεται ένα ακόμη μοντέλο κατηγοριοποίησης με την βοήθεια των Διανυσματικών Μηχανών Υποστήριξης (Support Vector Machines) και εξετάζεται η σύγκλιση των δύο.

Η συνολική δομή του συστήματος που επεκτείνεται μέσω αυτής της διπλωματικής εργασίας παρουσιάζεται στην εικόνα 2.1.



Εικόνα 2.1: Επέκταση συστήματος για κατηγοριοποίηση ασθενών με καρωτιδική αθηρωμάτωση.

2.2 Συνεισφορά

Η συνεισφορά της διπλωματικής συνοψίζεται ως εξής:

1. Μελετήθηκαν οι δυνατότητες των Κρυφών Μαρκοβιανών Μοντέλων.
2. Υλοποιήθηκαν 6 διαφορετικά σενάρια προεπεξεργασίας των δεδομένων.
3. Πραγματοποιήθηκε η εκπαίδευση των μοντέλων για κάθε σενάριο και προέκυψαν οι βέλτιστες παράμετροι για κάθε είδος κυματομορφών.

4. Για αυτές τις βέλτιστες παραμέτρους πραγματοποιήθηκε η διασταύρωση των αποτελεσμάτων μέσω της τεχνικής Cross Validation (leave-one-out).
5. Κατασκευή μοντέλου μέσω Support Vector Machines και εκπαίδευση αυτού.
6. Διασταύρωση αποτελεσμάτων όπως και στην προηγούμενη περίπτωση.
7. Επαλήθευση των διασταυρωμένων αποτελεσμάτων μέσω της συγκρίσεως των δύο μοντέλων.
8. Δημιουργία ενός συστήματος πλειοψηφίας για την κατηγοριοποίηση των ασθενών.

Οι αλγόριθμοι εκπαίδευσης, επανεκπαίδευσης, αξιολόγησης και διασταύρωσης έγιναν μέσω MATLAB (ο πηγαίος κώδικας υπάρχει στο τέλος του κειμένου) και η ταξινόμηση/επιλογή των καταλληλότερων παραμέτρων έγινε μέσω unix scripts από τα αρχεία εξόδου που είχαν προκύψει μέσα από τη διαδικασία εκπαίδευσης.

2.3 Οργάνωση κειμένου

Στο Μέρος II, παρουσιάζονται για λόγους πληρότητας τα σημαντικότερα κομμάτια της θεωρίας των Κρυφών Μαρκοβιανών Μοντέλων, όπως και οι κύριοι αλγόριθμοι που χρειάζονται για την επίλυση των τριών βασικών προβλημάτων (εκτίμηση, αποκωδικοποίηση, εκπαίδευση). Τέλος, παρουσιάζονται συνοπτικά οι Διανυσματικές Μηχανές Υποστήριξης. Στο Μέρος III, δίνεται το σύστημα αναγνώρισης φωνής (προτύπων) που θα χρησιμοποιηθεί σε αναλογία για την αναζήτηση προτύπων στην εργασία. Ακόμη παρουσιάζονται όλοι οι παράμετροι που χρειάστηκε να βελτιστοποιηθούν αλλά και οι αλγόριθμοι που κατασκευάστηκαν για το σκοπό αυτό. Η αξιολόγηση του μοντέλου, όπως και η επαλήθευση των αποτελεσμάτων μέσω ενός δεύτερου μοντέλου. Τα αποτελέσματα των μοντέλων και τα τελικά συμπεράσματα.

Μέρος II

Θεωρητικό υπόβαθρο

Κεφάλαιο 3

Κρυφά Μαρκοβιανά Μοντέλα (Hidden Markov Models)

Τα *Κρυφά Μαρκοβιανά Μοντέλα* (*Hidden Markov Models - HMMs*) αν και γνωστά για δεκαετίες [24] χαίρουν ιδιαίτερης προτίμησης τα τελευταία χρόνια και βρίσκονται μάλιστα σε στάδιο περαιτέρω ανάπτυξης. Οι εφαρμογές τους ποικίλουν και δεν εστιάζονται μονάχα στην αναγνώριση φωνής όπως είναι ευρέως γνωστό αλλά και σε γενικότερες κατηγορίες σημάτων όπως σήματα ηλεκτροεγκεφαλογραφημάτων [25], ηλεκτροκαρδιογραφημάτων [26] ακόμα και σε ανάλυση σεισμικών κυμάτων [27], με τελικό στόχο την συσχέτιση σήματος παρατήρησης (ή σημεία αυτού) με κάποια υπόθεση/αποτέλεσμα. Τέλος, δεν είναι και λίγα τα παραδείγματα χρήσης του μοντέλου σε βιολογία, ιατρική αλλά και σε χρηματοοικονομικά προβλεπτικά μοντέλα.

3.1 Θεωρία των Κρυφών Μαρκοβιανών Μοντέλων

Τα *Κρυφά Μαρκοβιανά Μοντέλα* είναι εκπαιδευσιμα πεπερασμένα στοχαστικά αυτόματα. Σήμερα θεωρούνται ως μια υποκατηγορία ή ως μια ειδική μορφή των δυναμικών Μπαεσιανών δικτύων (dynamic Bayesian networks).

Ένα *Κρυφό Μαρκοβιανό Μοντέλο* αποτελείται από δύο στοχαστικές ανελίξεις (ή διαδικασίες). Η πρώτη εξ αυτών, είναι μια αλυσίδα *Markov*, που χαρακτηρίζεται από καταστάσεις και πιθανότητες μετάβασης μεταξύ αυτών των καταστάσεων. Οι καταστάσεις της αλυσίδας είναι εξωτερικά μη παρατηρήσιμες, ή αλλιώς, "κρυφές". Η δεύτερη στοχαστική ανέλιξη παράγει εκπομπές (εξόδους) παρατηρήσιμες κάθε χρονική στιγμή, βασιζόμενη σε μια κατανομή πιθανοτήτων των καταστάσεων. Είναι σημαντικό να παρατηρήσουμε ότι ο όρος "κρυφές", όταν ορίζουμε ένα Κρυφό Μαρκοβιανό Μοντέλο αναφέρεται στις καταστάσεις της αλυσίδας *Markov* και όχι στις παραμέτρους του μοντέλου.

Στις επόμενες παραγράφους θα παρουσιαστούν συνοπτικά η ιστορία των *στοχαστικών ανελίξεων Markov* αλλά και των *αλυσίδων Markov* ως πρώτο μέρος και ως δεύτερο μέρος αναφέρονται

οι αναγκαίοι αλγόριθμοι για την κατασκευή των Κρυφών Μαρκοβιανών Μοντέλων με χρήση υπολογιστή.

3.1.1 Αλυσίδες Markov - Στοχαστικές ανελίξεις Markov

Ο Andrey Andreyevich Markov (1856 - 1922) ήταν ένας Ρώσος μαθηματικός, γνωστότερος για την δουλειά του στην θεωρία των στοχαστικών ανελίξεων Markov, η θεωρία του αργότερα έγινε γνωστή σαν αλυσίδες και στοχαστικές ανελίξεις Markov.

Ο Markov παρουσίασε τα πρώτα αποτελέσματα της θεωρίας του το 1906 όπου και χρησιμοποίησε για πρώτη φορά τον όρο "αλυσίδα". Το 1913 υπολόγιζε ακολουθίες γραμμάτων στα ρωσικά. Μία γενίκευση σε αριθμήσιμους, απείρων καταστάσεων, χώρους είχε δοθεί από τον Kolmogorov (1931). Οι αλυσίδες Markov συνδέονται αφαιρετικά με την τυχαία επίδραση της κίνησης σωματιδίων μέσα σε ένα αέριο σε κάποιο τρίτο αντικείμενο (Brownian motion) και την εργοδικότητα, δύο θέματα που απασχολούσαν τον κόσμο της φυσικής εκείνη την εποχή. Ο Markov είχε βέβαια κίνητρα μαθηματικά για τη μελέτη της επέκτασης του *νομού των μεγάλων αριθμών* σε εξαρτημένα ενδεχόμενα. Μέσα από αυτή τη προσέγγιση προέκυψε ένα γενικότερο στατιστικό εργαλείο οι στοχαστικές ανελίξεις Markov.

Στα μαθηματικά γενικότερα, στην θεωρία πιθανοτήτων και στατιστικής, μια ανέλιξη Markov μπορεί να θεωρηθεί ως ένα τυχαίο χρονομεταβλητό (στοχαστικό) φαινόμενο, για το οποίο ικανοποιούνται οι *ιδιότητες Markov*. Μια στοχαστική ανέλιξη μπορεί να θεωρηθεί ότι ικανοποιεί την *ιδιότητα Markov* εάν δεν έχει μνήμη. Η παρούσα κατάσταση του συστήματος δεν εξαρτάται από παρελθόντα και μελλοντικό χρόνο.

Οι στοχαστικές ανελίξεις Markov συναντώνται στις πιθανότητες και την στατιστική με τους ακόλουθους τρόπους. Είτε σαν στοχαστικές ανελίξεις που μέσα από μια απόδειξη έχουν την ιδιότητα Markov και με αυτό τον τρόπο κληρονομούν όλες τις υπόλοιπες ιδιότητες της στοχαστικής ανέλιξης Markov, είτε σαν αρχική παραδοχή ισχύος της ιδιότητας για μια στοχαστική ανέλιξη, έτσι, και την κατασκευή του μοντέλου. Εάν κοιτάζουμε την παραδοχή με όρους μοντελοποίησης, εισάγουμε μια στατιστική εξάρτηση σε ένα μοντέλο για μία στοχαστική ανέλιξη με τρόπο τέτοιο που επιτρέπει την μεταβολή αυτής της εξάρτησης με τρόπο δυναμικό.

Συνήθως, ο όρος αλυσίδα Markov, χρησιμοποιείται για να δείξει μια στοχαστική ανέλιξη Markov όπου έχει πεπερασμένο (ή μετρήσιμο) χώρο καταστάσεων. Η αλυσίδα Markov μπορεί ωστόσο να οριστεί σε συνεχή ή διακριτό χρόνο, αν και το σύννηθες αφορά σε διακριτό χρόνο.

3.2 Αλγόριθμοι και Κρυφά Μαρκοβιανά Μοντέλα

Η σημαντική πρόοδος της επιστήμης των υπολογιστών περί το 1940, με τα ερευνητικά αποτελέσματα των John Von Neuman, Alan Turing και άλλων, έδωσε την βάση για επιστήμονες σε όλο τον κόσμο να αναπτύξουν αλγορίθμους για να επιλύσουν σημαντικά προβλήματα με χρήση αυτομάτων. Την ίδια εποχή η γραμμική θεωρία στις επιστήμες αν και ελκυστικά εύχρηστη έδειχνε να μην αρκεί για να προσεγγίσει περίπλοκα φαινόμενα και έτσι στην θέση της σιγά σιγά άρχισε να αναπτύσσεται η μη-γραμμική θεωρία όπως και τα στοχαστικά μοντέλα. Μέχρι και το τέλος του 1960 έχουν ήδη αρχίσει να εφαρμόζονται τα Wiener filters, τα Kalman filters όπως και τα παράγωγα τους. Βασική ήταν όμως και η συμβολή του Claude Elwood Shannon που έδωσε ώθηση στην ενσωμάτωση των αυτομάτων, ντετερμινιστικών ή και στοχαστικών, στους ηλεκτρονικούς υπολογιστές και στις ηλεκτρικές συσκευές.

Ο αλγόριθμος *Πρόβλεψης μεγιστοποίησης - EM (expectation maximization)* έχει στενή σχέση με την *Μέγιστη πιθανοφάνεια (Maximum likelihood)*. Ιστορικά η Μέγιστη πιθανοφάνεια χρησιμοποιήθηκε από τους Gauss, Laplace κ.α και μεταγενέστερα έγινε πολύ γνωστή μέσω του Fisher. Αρκετά χρόνια μετά παρουσιάστηκε ο αλγόριθμος EM επίσημα από τους Arthur Dempster, Nan Laird και Donald Rubin ως ένας ενεργός αλγόριθμος που χρησιμοποιούταν ήδη ως μια επαναληπτική μέθοδος εύρεσης της Μέγιστης πιθανοφάνειας σε στατιστικά μοντέλα, που εξαρτώνται σε μη ευθέως παρατηρήσιμες λανθάνουσες (υποκείμενες) μεταβλητές, τώρα "βαπτίζεται", αλγόριθμος EM και ταυτόχρονα θεμελιώνεται θεωρητικά σε μια γενικότερη μορφή. Ο αλγόριθμος εναλλάσσεται μεταξύ του βήματος πρόβλεψης (E) που υπολογίζει μια πρόβλεψη για την πιθανοφάνεια, συμπεριλαμβάνοντας τις λανθάνουσες (άγνωστες) μεταβλητές ως δεδομένα του προβλήματος (γνωστές) και του βήματος μεγιστοποίησης (M), που υπολογίζει εκτιμήσεις για την Μέγιστη πιθανοφάνεια των παραμέτρων, μεγιστοποιώντας την προβλεπόμενη πιθανοφάνεια που υπολογίστηκε προηγουμένως στο βήμα (E). Στη συνέχεια φυσικά τα αποτελέσματα του βήματος (M) τροφοδοτούν την ανάδραση και γίνονται είσοδος σε ένα καινούργιο βήμα (E).

Ο αλγόριθμος EM χρησιμοποιείται συχνά στην εκπαίδευση αυτομάτων συστημάτων αποφάσεων, σε όραση υπολογιστή και ομαδοποίηση δεδομένων. Για την ανάλυση φυσικής γλώσσας δύο αλγόριθμοι της κατηγορίας EM επικρατούν, ο αλγόριθμος Baum-Welch (γνωστός και ως "forward-backward") και ο αλγόριθμος inside-outside για αυτόματα αναγωγή, πιθανοτικών, γραμματικών χωρίς συμφραζόμενα (context-free grammars).

Ο αλγόριθμος *Baum-Welch* είναι μια ειδική περίπτωση του γενικευμένου αλγορίθμου πρόβλεψης μεγιστοποίησης GEM (generalized expectation-maximization). Πήρε το όνομα του από τους Leonard E. Baum και Lloyd R. Welch. Ο αλγόριθμος χρησιμοποιείται για την εύρεση των αγνώστων παραμέτρων ενός Κρυφού Μαρκοβιανού Μοντέλου, για τον σκοπό αυτό κάνει χρήση του αλγορίθμου forward-backward. Μια από τις αρχικές δημοσιεύσεις για τον αλγόριθμο Baum-Welch ήταν αυτή του 1970 [28] και έκανε αναφορά σε μία επαναληπτική μέθοδο για την εύρεση Μέγιστης

πιθανοφάνειας, σε στατιστική εκτίμηση, που αφορούν αλυσίδες Markov.

Ο αλγόριθμος *Viterbi* ήταν το αποτέλεσμα της εργασίας του Andrew Viterbi το 1967 ως ένας αλγόριθμος αποκωδικοποίησης για κώδικες συνέλιξης μέσα από θορυβώδη γραμμές ψηφιακής επικοινωνίας και χρησιμοποιεί δυναμικό προγραμματισμό. Για την εύρεση της πιο πιθανής ακολουθίας των "κρυφών" καταστάσεων, ονομαζόμενο και *μονοπάτι-Viterbi*, που καταλήγει σε μία ακολουθία παρατηρούμενων συμβάντων. Τα τελευταία χρόνια έχει βρει εφαρμογή στην αποκωδικοποίηση σημάτων κυψελωτών δικτύων (CDMA, GSM), σε dial-up modems, σε πρωτόκολλα 802.11x ασύρματης μετάδοσης, όπως και σε δορυφορικές επικοινωνίες. Βρίσκει φυσικά και εφαρμογή σε πακέτα αναγνώρισης φωνής από μηχανή, ανίχνευσης λέξεων κλειδιών και στην βιοπληροφορική.

Για παράδειγμα, για κάποια συσκευή μετατροπής φωνής-κειμένου, το ακουστικό σήμα συμπεριφέρεται ως μια παρατηρούμενη ακολουθία συμβάντων και μία συμβολοσειρά (κείμενο σε φυσική γλώσσα) θεωρείται ως το "κρυφό αίτιο" του ακουστικού σήματος. Ο αλγόριθμος *Viterbi* θα βρει την πιθανότερη συμβολοσειρά για το ακουστικό σήμα.

3.3 Ορισμός των Κρυφών Μαρκοβιανών Μοντέλων

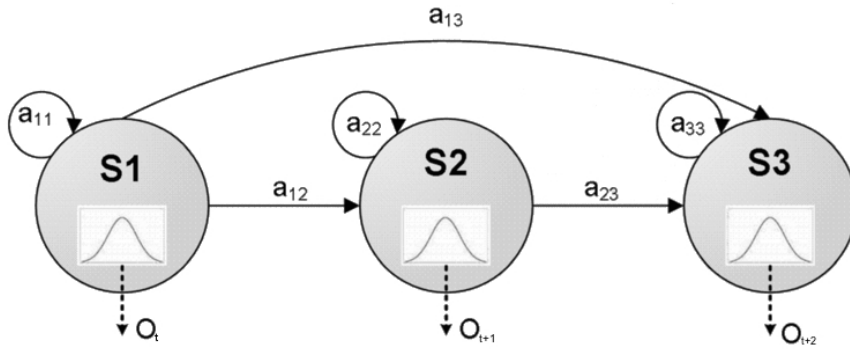
Ως *Κρυφό Μαρκοβιανό Μοντέλο* ορίζεται ένα πεπερασμένο εκπαιδευσιμο στοχαστικό αυτόματο, το οποίο μπορεί να εκφραστεί ως μια διπλή στοχαστική ανέλιξη που έχει τις ακόλουθες ιδιότητες:

- Η *πρώτη στοχαστική ανέλιξη* είναι ένα πεπερασμένο σύνολο καταστάσεων, όπου κάθε κατάσταση είναι συνήθως συνδεδεμένη με μία πολυδιάστατη κατανομή πιθανότητας. Οι μεταβάσεις μεταξύ των διαφόρων καταστάσεων είναι στατιστικά ορισμένες από ένα σύνολο πιθανοτήτων που ονομάζονται, *πιθανότητες μετάβασης (transition probabilities)*.
- Στη *δεύτερη στοχαστική ανέλιξη*, σε κάθε κατάσταση μπορούμε να παρατηρήσουμε ένα συμβάν. Εφόσον αναλύουμε αυτό που παρατηρούμε χωρίς να γνωρίζουμε σε ποιες καταστάσεις συνέβη το φαινόμενο, οι καταστάσεις είναι "κρυφές" ως προς τον παρατηρητή, εξού και η ονομασία Κρυφά Μαρκοβιανά Μοντέλα.

Κάθε Κρυφό Μαρκοβιανό Μοντέλο (εικόνα 3.1) ορίζεται από τις *καταστάσεις (states)*, τις *πιθανότητες κατάστασης (state probabilities)*, τις *πιθανότητες μετάβασης (transition probabilities)*, τις *πιθανότητες εκπομπών (emission probabilities)* και τις *αρχικές πιθανότητες (initial probabilities)*. Για τον πλήρη ορισμό του μοντέλου ορίζουμε τα παρακάτω.

1. Οι N καταστάσεις του μοντέλου,

$$S = \{S_1, \dots, S_N\} \quad (3.1)$$



Εικόνα 3.1: Δομή Κρυφών Μαρκοβιανών Μοντέλων. Όπου "S" οι καταστάσεις, "a" οι πιθανότητες μετάβασης, "O" οι παρατηρήσεις στον χρόνο.

2. Τα M σύμβολα παρατήρησης για κάθε κατάσταση,

$$V = \{u_1, \dots, u_M\}$$

Εάν οι παρατηρήσεις είναι σε συνεχή χρόνο, το $M \rightarrow \infty$.

3. Η κατανομή πιθανοτήτων μετάβασης κατάσταση, $A = \{a_{ij}\}$, όπου a_{ij} είναι η πιθανότητα μετάβασης του S_i κατά την χρονική στιγμή t , σε S_j την χρονική στιγμή $t+1$. Η δομή αυτού του στοχαστικού πίνακα ορίζει και την δομή (διασύνδεση) του συστήματος που κατασκευάζουμε μέσα από το μοντέλο. Εάν το στοιχείο a_{ij} είναι μηδέν, θα παραμείνει μηδέν ακόμη και μετά την διαδικασία εκπαίδευσης, άρα και δεν θα υπάρξει μετάβαση από την κατάσταση S_i στην S_j ,

$$a_{ij} = p\{q_{t+1} = j | q_t = i\}, \quad (1 \leq i, j \leq N) \quad (3.2)$$

όπου το q_t αναφέρεται στην παρούσα κατάσταση. Οι πιθανότητες μετάβασης θα πρέπει φυσικά να πληρούν τις κανονικές συνθήκες των στοχαστικών μεταβλητών,

$$a_{ij} \geq 0, \quad 1 \leq i, j \leq N, \quad \text{και} \quad \sum_{j=1}^N a_{ij} = 1 \quad (1 \leq i \leq N)$$

4. Η κατανομή πιθανότητας των συμβόλων παρατήρησης για κάθε κατάσταση, $B = \{b_j(k)\}$ όπου $b_j(k)$ η πιθανότητα για το σύμβολο u_k να εκπέμπεται στην κατάσταση S_j .

$$b_j(k) = p\{o_t = u_k | q_t = j\}, \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \quad (3.3)$$

όπου u_k το k -οστό σύμβολο παρατήρησης στην αλφάβητο του μοντέλου και o_t το τρέχον διάλυμα παραμέτρων. Εδώ πάλι πρέπει να τηρούνται τα ακόλουθα,

$$b_j(k) \geq 0, \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \quad \text{και} \quad \sum_{k=1}^M b_j(k) = 1 \quad (1 \leq j \leq N)$$

Εάν οι παρατηρήσεις είναι σε συνεχή χρόνο, τότε θα έπρεπε να χρησιμοποιηθεί συνεχής συνάρτηση πυκνότητας πιθανότητας, σε αντίθεση με τις διακριτές παρατηρήσεις. Συνήθως για την συνάρτηση πυκνότητας πιθανότητας γίνεται μια προσέγγιση από ένα άθροισμα με συντελεστές βάρους, M κανονικών κατανομών,

$$b_j(o_t) = \sum_{m=1}^M c_{jm} N(\mu_{jm}, \Sigma_{jm}, o_t) \quad (3.4)$$

όπου c_{jm} = συντελεστές βάρους, μ_{jm} = διανύσματα μέσων, και Σ_{jm} = πίνακες συνδιακύμανσης, τα c_{jm} θα πρέπει να πληρούν τις συνθήκες

$$c_{jm} \geq 0, \quad 1 \leq j \leq N, \quad 1 \leq m \leq M \quad \text{και} \quad \sum_{m=1}^M c_{jm} = 1, \quad (1 \leq j \leq N)$$

5. Η αρχική κατανομή πιθανοτήτων $\pi = \{\pi_i\}$, όπου π_i είναι η πιθανότητα ότι το μοντέλο βρίσκεται στην κατάσταση S_i την χρονική στιγμή $t = 0$ με

$$\pi_i = p\{q_1 = i\} \quad \text{και} \quad 1 \leq j \leq N \quad (3.5)$$

Κατά τον ορισμό του μοντέλου έμμεσα αλλά φανερά ορίζουμε και την διακριτή ή συνεχή του φύση. Υπάρχει ακόμη και η περίπτωση ενός μικτού μοντέλου με διακριτά αλλά και συνεχή στοιχεία. Συνήθως για τον ορισμό διακριτού μοντέλου χρησιμοποιείται μια παρεμφερής μορφή του

$$\lambda = (A, B, \pi) \quad (3.6)$$

και για τον ορισμό ενός συνεχούς μοντέλου

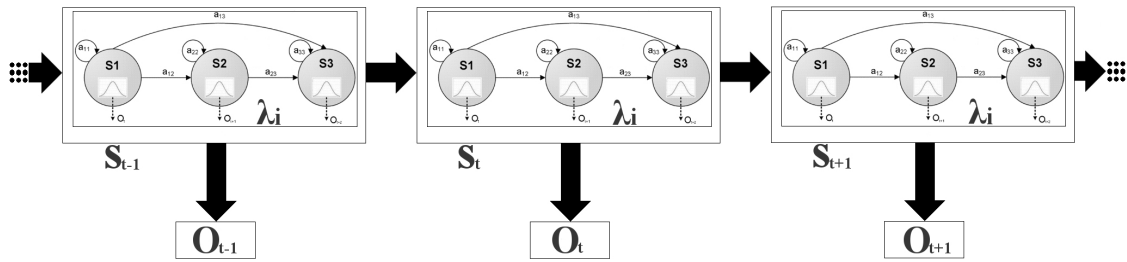
$$\lambda = (A, c_{jm}, \mu_{jm}, \Sigma_{jm}, \pi) \quad (3.7)$$

3.4 Στοχαστικά Αυτόματα - Αρχιτεκτονική Μοντέλου

Η μοντελοποίηση ενός (στοχαστικού αυτομάτου) Κρυφού Μαρκοβιανού Μοντέλου αποτελείται από δύο μέρη.

- Το πρώτο βήμα είναι να οριστεί η αρχιτεκτονική του μοντέλου
- Το δεύτερο βήμα είναι να οριστούν οι αλγόριθμοι εκπαίδευσης και λειτουργίας του μοντέλου.

Η εικόνα 3.2 δείχνει τη μορφή ενός γενικευμένου αυτομάτου που χρησιμοποιεί ένα Κρυφό Μαρκοβιανό Μοντέλο λ_i , με τις δύο παράλληλες στοχαστικές ανεξίτητες.



Εικόνα 3.2: Γενική μορφή λειτουργικής αρχιτεκτονικής ενός Κρυφού Μαρκοβιανού Μοντέλου.

Κάθε ένα από τα σχήματα αντιπροσωπεύει μια τυχαία μεταβλητή που μπορεί να περιέχει ένα σύνολο τιμών. Η τυχαία μεταβλητή $S_t = s(t)$ είναι η "κρυφή κατάσταση" την χρονική στιγμή t . Η τυχαία μεταβλητή $O_t = o(t)$ είναι η παρατήρηση την χρονική στιγμή t . Ο νόμος της δεσμευμένης πιθανότητας για την "κρυφή" μεταβλητή $s(t)$ την χρονική στιγμή t μας λέει ότι, για όλες τις μεταγενέστερες στιγμές, οι τιμές των "κρυφών" μεταβλητών $s(t)$ εξαρτώνται μονάχα από την τιμή της "κρυφής" μεταβλητής $s(t-1)$. Πράγμα που σημαίνει ότι το σύστημα λειτουργεί δίχως μνήμη, οι προγενέστερες τιμές δεν είναι απαραίτητες, έτσι η ιδιότητα Markov τηρείται. Για την δεύτερη στοχαστική ανέλιξη έχουμε ότι, η τιμή της παρατηρούμενης τυχαίας μεταβλητής $o(t)$ εξαρτάται από την τιμή της "κρυφής" μεταβλητής $s(t)$, για την ίδια χρονική τιμή t .

3.5 Τα Τρία Βασικά Προβλήματα

Ο αλγόριθμος εκπαίδευσης έχει σκοπό να βρει το καλύτερο σει πιθανοτήτων, για τις μεταβάσεις καταστάσεων και τις παρατηρήσεις του συστήματος. Έτσι παράγεται μια ακολουθία ή ακολουθίες ως έξοδος του αλγορίθμου. Τα τρία βασικά προβλήματα που προκύπτουν και επιλύονται στα Κρυφά Μαρκοβιανά Μοντέλα είναι τα ακόλουθα.

1. Το πρόβλημα της εκτίμησης (*The evaluation problem*)

Ποια η πιθανότητα ότι οι δοθείσες παρατηρήσεις $O = o_1, o_2, \dots, o_T$ να έχουν κατασκευαστεί από το μοντέλο $p\{O|\lambda\}$, όπου λ ένα Κρυφό Μαρκοβιανό Μοντέλο.

2. Το πρόβλημα της αποκωδικοποίησης (*The decoding problem*)

Ποια είναι η πιο πιθανή ακολουθία καταστάσεων για ένα δοσμένο μοντέλο λ που παρήγαγε τις παρατηρήσεις $O = o_1, o_2, \dots, o_T$.

3. Το πρόβλημα της εκπαίδευσης (*The learning problem*)

Πως πρέπει να ρυθμίσουμε τις παραμέτρους του μοντέλου $\{A, B, \pi\}$ για να μεγιστοποιήσουμε την πιθανότητα $p\{O|\lambda\}$, με δοσμένα ένα μοντέλο λ και μια ακολουθία παρατηρήσεων $O = o_1, o_2, \dots, o_T$.

Το πρόβλημα της εκτίμησης μπορεί να χρησιμοποιηθεί για μεμονωμένη αναγνώριση (λέξεων) προτύπων. Το πρόβλημα της αποκωδικοποίησης συνδέεται με την συνεχή αναγνώριση αλλά και

στην ομαδοποίηση. Το πρόβλημα της εκπαίδευσης, *πρέπει να λυθεί*, γιατί το μοντέλο πρέπει να λάβει τις κατάλληλες τιμές παραμέτρων ώστε να προχωρήσει στην αναγνώριση συγκεκριμένων προτύπων σε μία εφαρμογή. Η εκπαίδευση σε επόμενο στάδιο φυσικά προχωρά και στην ενθέρωση (επανεκπαίδευση) του μοντέλου με καινούργια στοιχεία ή και στην βελτιστοποίηση των παραμέτρων. Από εδώ και πέρα για λόγους συντομίας, όποτε γίνεται αναφορά σε μοντέλο, υπονοείται το Κρυφό Μαρκοβιανό Μοντέλο, σε αντίθετη περίπτωση θα αναφέρεται διαφορετικά.

3.5.1 Το πρόβλημα της εκτίμησης - Αλγόριθμος Forward

Για ένα δοσμένο μοντέλο $\lambda = \{A, B, \pi\}$ και μια ακολουθία παρατηρήσεων $O = o_1, o_2, \dots, o_T$ πρέπει να βρεθεί η πιθανότητα $p\{O|\lambda\}$. Παρόλο που αυτή η ποσότητα μπορεί να υπολογιστεί με συμβατική χρήση της θεωρίας πιθανοτήτων, αυτό θα είχε ως αποτέλεσμα μια πολυπλοκότητα της τάξεως του N^T . Η μέθοδος που μετριάζει την υπολογιστική πολυπλοκότητα είναι ο αλγόριθμος forward, που χρησιμοποιεί μια βοηθητική μεταβλητής

$$\alpha_t(i) = p\{o_1, o_2, \dots, o_t, q_t = i | \lambda\} \quad (3.8)$$

όπου $\alpha_t(i)$ είναι η *εμπρόσθια μεταβλητή* (*forward variable*) και o_1, o_2, \dots, o_t η μερική ακολουθία παρατήρησης. Από εκεί ο αναδρομικός τύπος είναι ο εξής

$$\alpha_{t+1}(j) = b_j(o_{t+1}) \sum_{i=1}^N \alpha_t(i) o_{ij}, \quad 1 \leq j \leq N, \quad 1 \leq t \leq T-1 \quad (3.9)$$

με $\alpha_1(j) = \pi_j b_j(o_1)$, $1 \leq j \leq N$. Με τον αναδρομικό τύπο μπορεί να υπολογιστεί το $\alpha_T(i)$, $1 \leq i \leq N$. Άρα η ζητούμενη πιθανότητα δίνεται

$$p\{O|\lambda\} = \sum_{i=1}^N \alpha_T(i) \quad (3.10)$$

Η *οπίσθια μεταβλητή* (*backward variable*) $\beta_t(i)$ ορίζεται ανάλογα.

$$\beta_t(i) = p\{o_{t+1}, o_{t+2}, \dots, o_T | q_t = i, \lambda\} \quad (3.11)$$

Δεδομένου ότι η παρούσα κατάσταση είναι η i , $\beta_t(i)$ είναι η πιθανότητα της μερικής ακολουθίας παρατηρήσεων $o_{t+1}, o_{t+2}, \dots, o_T$. Το $\beta_t(i)$ μπορεί να υπολογιστεί μέσω του αναδρομικού τύπου

$$\beta_t(i) = \sum_{j=1}^N \beta_{t+1}(j) a_{ij} b_j(o_{t+1}), \quad 1 \leq i \leq N, \quad 1 \leq t \leq T-1 \quad (3.12)$$

όπου $\beta_T(i) = 1$, $1 \leq i \leq N$. Επιπλέον, βλέπουμε ότι

$$\alpha_t(i) \beta_t(i) = p\{O, q_t = i | \lambda\}, \quad 1 \leq i \leq N, \quad 1 \leq t \leq T \quad (3.13)$$

Άρα, υπάρχουν δύο τρόποι για τον υπολογισμό της πιθανότητας $p\{O|\lambda\}$, είτε με την χρήση της εμπρόσθιας, είτε με την χρήση της οπίσθιας μεταβλητής.

$$p\{O|\lambda\} = \sum_{i=1}^N p\{O, q_t = i|\lambda\} = \sum_{i=1}^N \alpha_t(i)\beta_t(i) \quad (3.14)$$

3.5.2 Το πρόβλημα της αποκωδικοποίησης - Αλγόριθμος Viterbi

Δοσμένης μιας ακολουθίας παρατηρήσεων $O = o_1, o_2, \dots, o_T$ και ενός μοντέλου $\lambda = \{A, B, \pi\}$ αναζητούμε την πιο πιθανή ακολουθία καταστάσεων. Ο ορισμός της *πιο πιθανής ακολουθίας καταστάσεων* επηρεάζει άμεσα την λύση του προβλήματος. Κατά μια προσέγγιση, θέλουμε να βρούμε την πιθανότερη κατάσταση q_t και στη συνέχεια να ενώσουμε όλες αυτές τις q_t καταστάσεις που έχουν βρεθεί. Αυτό όμως συχνά οδηγεί σε ακολουθίες καταστάσεων που τελικά στερούνται χρησιμότητας και έτσι η προτιμότερη μέθοδος βρίσκεται στον αλγόριθμο Viterbi, με τον οποίο βρίσκεται η συνολική ακολουθία καταστάσεων με την μέγιστη πιθανοφάνεια.

Ορίζεται μια βοηθητική μεταβλητή που δίνει την καλύτερη πιθανότητα για τις ακολουθίες καταστάσεων και μερικών παρατηρήσεων μέχρι και τον χρόνο t , με δεδομένο ότι i είναι η τρέχουσα κατάσταση.

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} p\{q_1, q_2, \dots, q_{t-1}, q_t = i, o_1, o_2, \dots, o_{t-1}|\lambda\} \quad (3.15)$$

επάγεται ότι

$$\delta_{t+1}(j) = b_j(o_{t+1}) \left[\max_{1 \leq i \leq N} \delta_t(i) a_{ij} \right], \quad 1 \leq i \leq N, \quad 1 \leq t \leq T-1 \quad (3.16)$$

όπου $\delta_1(j) = \pi_j b_j(o_1)$, $1 \leq j \leq N$. Έτσι ο αλγόριθμος ξεκινάει από τον υπολογισμό του $\delta_T(j)$, $1 \leq j \leq N$ για την εύρεση της πιο πιθανής ακολουθίας καταστάσεων. Πάντοτε κρατείται ένας δείκτης στην "καλύτερη κατάσταση" κατά τον υπολογισμό του μεγίστου, καταλήγοντας στην κατάσταση $j^* = \arg \max_{1 \leq i \leq N} \delta_t(i)$. Αρχίζοντας από αυτή τη κατάσταση και ανατρέχοντας προς τα πίσω (backtracking) την συνολική ακολουθία καταστάσεων σύμφωνα με τον δείκτη που υπολογίζεται κάθε φορά. Έτσι έχουμε την προς αναζήτηση ακολουθία καταστάσεων. Ο αλγόριθμος μπορεί να συνοψιστεί ως μια αναζήτηση σε έναν γράφο, με κόμβους τις καταστάσεις του μοντέλου για κάθε χρονική στιγμή t , $1 \leq t \leq T$.

3.5.3 Το πρόβλημα της εκπαίδευσης

Πως μπορούμε να προσαρμόσουμε τις παραμέτρους του μοντέλου με τρόπο τέτοιο έτσι ώστε το δοσμένο σύνολο παρατηρήσεων (δεδομένα εκπαίδευσης - training set) να αναπαρίσταται με τον πιο ακριβή τρόπο. Ανάλογα με την εφαρμογή που τελικά θα έχει το μοντέλο, υπάρχει και διαφορετική παράμετρος που χρειάζεται να βελτιστοποιηθεί. Έτσι υπάρχουν διάφορα κριτήρια

βελτιστοποίησης της διαδικασίας εκπαίδευσης με κυριότερα, το κριτήριο *Μέγιστης Πιθανοφάνειας* (*Maximum likelihood*) και το κριτήριο *Μέγιστης Αμοιβαίας Πληροφορίας* (*Maximum Mutual Information*).

3.5.3.1 Το κριτήριο Μέγιστης Πιθανοφάνειας (ML)

Για μοντέλο λ_ω της κλάσης ω , προσπαθούμε να μεγιστοποιήσουμε την πιθανότητα, η ακολουθία παρατηρήσεων O^ω να ανήκει στην κλάση ω , σύμφωνα με τις παραμέτρους του μοντέλου. Μαθηματικά, η πιθανοφάνεια εκφράζεται

$$L_{tot} = p\{O^\omega | \lambda_\omega\} \quad (3.17)$$

Εάν θεωρήσουμε ότι μιλάμε για μία κλάση κάθε φορά, μπορούμε να εξαιρέσουμε τον δείκτη ω και να έχουμε

$$L_{tot} = p\{O | \lambda\} \quad (3.18)$$

Το μοντέλο $\lambda = \{A, B, \pi\}$ που μεγιστοποιεί την ποσότητα L_{tot} μπορεί να λυθεί μονάχα τοπικά. Για τον αναδρομικό υπολογισμό χρησιμοποιείται ο αλγόριθμος Baum-Welch ή μέθοδοι που βασίζονται στην κλίση (gradient).

Ο αλγόριθμος *Baum-Welch* [28] είναι γνωστός και ως αλγόριθμος *forward-backward*. Αυτός ο αλγόριθμος προκύπτει από την απλή άθροιση συμβάντων ή αναλυτικές μεθόδους για την μεγιστοποίηση της βοηθητικής ποσότητας

$$Q(\lambda, \bar{\lambda}) = \sum_q p\{q | O, \lambda\} \log [p\{O, q, \bar{\lambda}\}] \quad (3.19)$$

πάνω στο $\bar{\lambda}$. Πέραν της εμπρόσθιας και οπίσθιας πρέπει να οριστούν δύο ακόμη βοηθητικές μεταβλητές. Η πρώτη είναι η εξής

$$\xi_t(i, j) = p\{q_t = i, q_{t+1} = j | O, \lambda\} = \frac{p\{q_t = i, q_{t+1} = j, O | \lambda\}}{p\{O | \lambda\}} \quad (3.20)$$

Με την χρήση των εμπρόσθιων και οπίσθιων μεταβλητών έχουμε τελικά για την πρώτη βοηθητική μεταβλητή

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} \beta_{t+1}(j) b_j(o_{t+1})}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ij} \beta_{t+1}(j) b_j(o_{t+1})} \quad (3.21)$$

Η δεύτερη βοηθητική μεταβλητή είναι εκ των υστέρων πιθανότητα (ά posteriori)

$$\gamma_t(i) = p\{q_t = i | O, \lambda\} \quad (3.22)$$

και με τις εμπρόσθιες και οπίσθιες μεταβλητές προκύπτει

$$\gamma_t(i) = \frac{\alpha_t(i) \beta_t(i)}{\sum_{i=1}^N \alpha_t(i) \beta_t(i)} \quad (3.23)$$

Η συσχέτιση των $\xi_t(i, j)$ και $\gamma(i)$ δίνεται από την

$$\gamma(i) = \sum_{j=1}^N \xi_t(i, j), \quad 1 \leq i \leq N, \quad 1 \leq t \leq M \quad (3.24)$$

Για την μεγιστοποίηση της ποσότητας $p\{O|\lambda\}$ θα περιγραφεί διαδικασία εκπαίδευσης Baum-Welch. Από το μοντέλο $\lambda = \{A, B, \pi\}$ υπολογίζονται οι μεταβλητές α_t και β_t για να υπολογιστούν στη συνέχεια τα ξ_t και γ_t . Οι ακόλουθες παραστάσεις λέγονται *συναρτήσεις επανεκτίμησης (re-estimation formulas)* και χρησιμοποιούνται για την ανανέωση των παραμέτρων του μοντέλου.

$$\bar{\pi}_i = \gamma(i), \quad 1 \leq i \leq N \quad (3.25)$$

$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \quad 1 \leq i \leq N, \quad 1 \leq j \leq N \quad (3.26)$$

$$\bar{b}_j(k) = \frac{\sum_{t=1}^T \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)}, \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \quad (3.27)$$

3.5.3.2 Βελτιστοποίηση παραμέτρων

Η βελτιστοποίηση της παραμέτρου κ ενός μοντέλου λ γίνεται μέσω αλγορίθμων που εκμεταλλεύονται την κλίση όπως στην ακόλουθη εξίσωση

$$\kappa^t = \kappa^{t-1} - \zeta \left[\frac{\partial \psi}{\partial \kappa} \right]_{\kappa^{t-1}} \quad (3.28)$$

Εάν ορίσουμε

$$\psi = -\log(p\{O|\lambda\}) \quad (3.29)$$

για να υπολογίσουμε την μέγιστη πιθανοφάνεια, για κάθε παράμετρο κ η συνάρτηση $\frac{\partial \psi}{\partial \kappa}$, στο μοντέλο λ θα πρέπει να λυθεί για να λάβουμε το ελάχιστο ψ . Αυτό το ψ είναι η εκτίμηση της Μέγιστης Πιθανοφάνειας, μέσα από την μεγιστοποίηση του κ^t . Συνδέοντας το ψ με τις παραμέτρους του μοντέλου (4.14) λαμβάνουμε

$$L_{tot} = \sum_{i=1}^N p\{O, q_t = i|\lambda\} = \sum_{i=1}^N \alpha_t(i) \beta_t(i) \quad (3.30)$$

Η διαφορίση της τελευταίας σχέσης με τις (4.28, 4.29) ως προς τη μεταβλητή κ δίνει την

$$\frac{\partial \psi}{\partial \kappa} = -\frac{1}{L_{tot}} \frac{\partial L_{tot}}{\partial \kappa} \quad (3.31)$$

Η σχέση (4.31) υπολογίζει το $\frac{\partial \psi}{\partial \kappa}$ με την υπόθεση του ότι γνωρίζουμε πως το $\frac{\partial L_{tot}}{\partial \kappa}$ είναι επιλύσιμο. Αλλά αυτή η διαφορίση εξαρτάται από τις παραμέτρους του μοντέλου. Από την μία πλευρά δηλαδή, εξαρτάται από τις *πιθανότητες μετάβασης* $\alpha_{ij}, 1 \leq i, N \geq j$ και ταυτόχρονα από τις *πιθανότη-*

τες παρατήρησης $b_j(k)$, $j \in \{1, \dots, N\}$, $k \in \{1, \dots, M\}$. Για αυτό τον λόγο χρειάζεται ο υπολογισμός των διαφορικών και των δύο σετ πιθανοτήτων.

- *Κλίση Μέγιστης Πιθανοφάνειας εξαρτώμενη από πιθανότητες μετάβασης*

Για να υπολογίσουμε την κλίση σε σχέση με τις πιθανότητες μετάβασης, έχουμε

$$\frac{\partial L_{tot}}{\partial \alpha_{ij}} = \sum_{t=1}^T \frac{\partial L_{tot}}{\partial \alpha_{ij}} \frac{\partial \alpha_t(j)}{\partial \alpha_{ij}} \quad (3.32)$$

Η απλή παραγωγή

$$\frac{\partial L_{tot}}{\partial \alpha_{ij}} = \beta_t(j) \quad (3.33)$$

όπως και η διαφορίση (χρονική υστέρηση)

$$\frac{\partial \alpha_t(j)}{\partial \alpha_{ij}} = b_j(\alpha_t) \alpha_{t-1}(i) \quad (3.34)$$

Μέγιστης Πιθανοφάνειας (ML) και μετά από αντικαταστάσεις έχουμε

$$\frac{\partial \psi}{\partial \alpha_{ij}} = -\frac{1}{L_{tot}} \sum_{t=1}^T \beta_t(j) b_j(\alpha_t) \alpha_{t-1}(i) \quad (3.35)$$

- *Κλίση Μέγιστης Πιθανοφάνειας εξαρτώμενη από πιθανότητες παρατήρησης*

Με τον ίδιο τρόπο με τα ανωτέρω έχουμε

$$\frac{\partial L_{tot}}{\partial b_j(o_t)} = \frac{\partial L_{tot}}{\partial \alpha_t(j)} \frac{\partial \alpha_t(j)}{\partial b_j(o_t)} \quad (3.36)$$

και

$$\frac{\partial \alpha_t(j)}{\partial b_j(o_t)} = \frac{\alpha_t(j)}{b_j(o_t)} \quad (3.37)$$

τελικά το αποτέλεσμα της εκτίμησης πιθανότητας είναι

$$\frac{\partial \psi}{\partial b_j(o_t)} = -\frac{1}{L_{tot}} \frac{\alpha_t(j) \beta_t(j)}{b_j(o_t)} \quad (3.38)$$

3.5.3.3 Το κριτήριο Μέγιστης Αμοιβαίας Πληροφορίας (MMI)

Γενικότερα, για τη λύση προβλημάτων αναγνώρισης προτύπων με χρήση Κρυφών Μαρκοβιανών Μοντέλων, υπάρχουν δύο γενικοί τύποι στοχαστικής βελτιστοποίησης, η Μέγιστη Πιθανοφάνεια και η Μέγιστη Αμοιβαία Πληροφορία. Ο σκοπός της Μέγιστης Πιθανοφάνειας είναι να βελτιστοποιήσει διάφορες παραμέτρους μίας κλάσης ενός μοντέλου κάθε φορά, ανεξαρτήτως των παραμέτρων κάθε άλλης κλάσης. Η Μέγιστη Αμοιβαία Πληροφορία χρησιμοποιείται για την επίλυση προβλημάτων διαφοροποίησης σε εφαρμογές αναγνώρισης προτύπων μεταξύ κάποιον κλά-

σεων. Για τον υπολογισμό, μπορεί να θεωρηθεί ένα σύνολο Κρυφών Μαρκοβιανών Μοντέλων

$$\Lambda = \{\lambda_v, v \in \{1, \dots, V\}\} \quad (3.39)$$

ενός προβλήματος αναγνώρισης προτύπων. Ο σκοπός της βελτιστοποίησης είναι η ελαχιστοποίηση της *δεσμευμένης αβεβαιότητας* v ενός πλήρους μέρους μιας διεργασίας (στον πραγματικό κόσμο), με δοσμένη την ακολουθία παρατήρησης O^s κάθε κλάσης.

$$I(u|O^s, \Lambda) = -\log p\{u|O^s, \Lambda\} \quad (3.40)$$

Η *δεσμευμένη εντροπία* H , που μπορεί να εκφραστεί ως την εκτίμηση της δεσμευμένης πληροφορίας I .

$$H(V|O) = E[\{I(u|O^s, \Lambda)\}] \quad (3.41)$$

όπου V είναι το σύνολο όλων των κλάσεων και O είναι το σύνολο των ακολουθιών παρατηρήσεων. Άρα, η κοινή πληροφορία μεταξύ των κλάσεων και των παρατηρήσεων

$$H(V|O) = H(V) - H(V|O^s) \quad (3.42)$$

είναι μια μεγιστοποιημένη μεταβλητή $H(V)$, εξού και η ονομασία Μέγιστη Αμοιβαία Πληροφορία. Η μέθοδος αυτή αναφέρεται και ως *Maximum á Εμπρόσθιοι (MAP)*.

Περιγραφή της μέγιστης *á posteriori* εκτίμησης Για την εκτίμηση/εύρεση ενός Κρυφού Μαρκοβιανού Μοντέλου λ από τις παρατηρήσεις o , ορίζουμε μια συνάρτηση δειγματοληψίας f για τις παρατηρήσεις o , έτσι ώστε $f(o|\lambda)$ να είναι η πιθανότητα της ακολουθίας o και να ταυτίζεται με το μοντέλο. Η συνάρτηση $\lambda \mapsto f(o|\lambda)$ μπορεί να οριστεί ως τη συνάρτηση πιθανοφάνειας, έτσι ώστε η εκτίμηση

$$\hat{\lambda}_{ML}(o) = \arg \max_{\lambda} f(o|\lambda) \quad (3.43)$$

να είναι η εκτίμηση της μέγιστης πιθανοφάνειας για το μοντέλο λ . Εάν υποθέσουμε ότι προϋπάρχει μια κατανομή χ στα μοντέλα λ , μπορούμε να θεωρήσουμε το κάθε μοντέλο λ ως μία Τυχαία Μεταβλητή όπως στη κλασσική στατιστική του Bayes. Η εκ των υστέρων (posterior) κατανομή του λ είναι

$$\lambda \mapsto f(\lambda|o) = \frac{f(o|\lambda)\chi(\lambda)}{\int_{\lambda' \in \Lambda} f(o|\lambda')\chi(\lambda')\partial\lambda'} \quad (3.44)$$

όπου χ είναι η συνάρτηση πυκνότητας του λ και Λ το πεδίο στο οποίο ορίζεται το χ . Η μέθοδος της μέγιστης εκ των υστέρων εκτίμησης, εκτιμά τα Κρυφά Μαρκοβιανά Μοντέλα λ ως εξής

$$\hat{\lambda}_{ML}(o) = \arg \max_{\lambda} \frac{f(o|\lambda)\chi(\lambda)}{\int_{\lambda' \in \Lambda} f(o|\lambda')\chi(\lambda')\partial\lambda'} = \arg \max_{\lambda} f(o|\lambda)\chi(\lambda) \quad (3.45)$$

Ο παρονομαστής δεν εξαρτάται από το λ και έτσι δεν παίζει κανένα ρόλο στην βελτιστοποίηση. Για να ταυτίζεται η εκτίμηση MAP των μοντέλων λ με την εκτίμηση ML, πρέπει η εκ των υστέρων κατανομή χ να είναι ομοιόμορφη (δηλαδή σταθερή συνάρτηση).

Εφαρμογή στα Κρυφά Μαρκοβιανά Μοντέλα Σύμφωνα με την παραπάνω ανάλυση περί "μεγίστου εκ των υστέρων", η εκ των υστέρων πιθανότητα $p\{v|O^c, \Lambda\}$ μεγιστοποιείται όταν η εκτίμηση μέσω κριτηρίου MMI είναι

$$E_{MAP} = E_{MMI} = -\log p\{v|O^s, \Lambda\} = -\log \frac{p\{v, O^c|\Lambda\}}{p\{O^c|\Lambda\}} = -\log \frac{p\{v, O^c|\Lambda\}}{p\{\omega, O^c|\lambda\}} \quad (3.46)$$

όπου με ω αντιπροσωπεύεται κάθε πιθανή κλάση. Οι πιθανοφάνειες μπορούν να γραφούν ως

$$L_{tot}^{correct} = p\{v, O^c|\lambda\} \quad (3.47)$$

$$L_{tot}^{others} = \sum_{\omega} p\{\omega, O^c|\lambda\} \quad (3.48)$$

με τους δείκτες "correct" και "others" διαχωρίζονται οι "αληθινές" από τις "εσφαλμένες" κλάσεις. Από τις δύο συναρτήσεις έχουμε τις εκτιμήσεις

$$E_{MAP} = E_{MMI} = -\log \frac{L_{tot}^{correct}}{L_{tot}^{others}} \quad (3.49)$$

Σε αναλογία με την ML, για να ελαχιστοποιηθεί η εκτίμηση E_{MMI} , θέτουμε $\psi = E_{MMI}$ και υπολογίζουμε όπως πριν το διαφορικό

$$\frac{\partial \psi}{\partial \kappa} = \frac{1}{L_{tot}^{others}} \frac{\partial L_{tot}^{others}}{\partial \kappa} - \frac{1}{L_{tot}^{correct}} \frac{\partial L_{tot}^{correct}}{\partial \kappa} \quad (3.50)$$

$$\text{με } L_{tot}^{correct} = \sum_i \alpha_t(i) \beta_t(i) \quad \text{και} \quad L_{tot}^{others} = \sum_{\omega} \sum_{i \in \text{class } \omega} \alpha_t(i) \beta_t(i).$$

- *Κλίση Μέγιστης Αμοιβαίας Πληροφορίας εξαρτώμενη από πιθανότητες μετάβασης*

Με χρήση του δέλτα του Kronecker δ_{ku} , ο υπολογισμός είναι ο ακόλουθος

$$\frac{\partial L_{tot}}{\partial \alpha_{ij}} = \sum_{i=1}^T \frac{\partial L_{tot}}{\partial \alpha_t(j)} \frac{\partial \alpha_t(j)}{\partial \alpha_{ij}} \quad (3.51)$$

με

$$\left[\frac{\partial L_{tot}^{correct}}{\partial \alpha_{ij}} \right]_{i \in \text{class } k} = \delta_{ku} \sum_{i=1}^T \beta_t(j) b_j(o_t) \partial \alpha_{t-1}(i) \quad (3.52)$$

και

$$\frac{\partial L_{tot}^{others}}{\partial \alpha_{ij}} = \sum_{i=1}^T \beta_t(j) b_j(o_t) \alpha_{t-1}(i) \quad (3.53)$$

τελικώς

$$\left[\frac{\partial \psi}{\partial \alpha_{ij}} \right]_{i \in class k} = \left[\frac{1}{L_{tot}^{others}} - \frac{\delta_{ku}}{L_{tot}^{correct}} \right] \sum_{i=1}^T \beta_t(j) b_j(o_t) \alpha_{t-1}(i) \quad (3.54)$$

- *Κλίση Μέγιστης Αμοιβαίας Πληροφορίας εξαρτώμενη από πιθανότητες παρατήρησης*

Ο υπολογισμός της κλίσης για την Μέγιστη Αμοιβαία Πληροφορία είναι όμοιος με τους ανωτέρω υπολογισμούς

$$\frac{\partial L_{tot}}{\partial b_j(o_t)} = \frac{\partial L_{tot}}{\partial \alpha_t(j)} \frac{\partial \alpha_t(j)}{\partial b_j(o_t)} \quad (3.55)$$

με

$$\left[\frac{\partial L_{tot}^{correct}}{\partial b_j(o_t)} \right]_{j \in class k} = \delta_{ku} \frac{\alpha_t(j) \beta_t(j)}{b_j(o_t)} \quad (3.56)$$

και

$$\frac{\partial L_{tot}^{others}}{\partial b_j(o_t)} = \frac{\alpha_t(j) \beta_t(j)}{b_j(o_t)} \quad (3.57)$$

για να έχουμε τελικά

$$\left[\frac{\partial \psi}{\partial b_j(o_t)} \right]_{j \in class k} = \left[\frac{1}{L_{tot}^{others}} - \frac{\delta_{ku}}{L_{tot}^{correct}} \right] \frac{\alpha_t(j) \beta_t(j)}{b_j(o_t)} \quad (3.58)$$

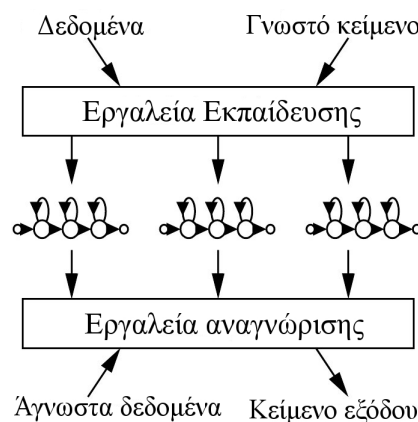
Κεφάλαιο 4

Αναγνώριση φωνής με χρήση Κρυφών Μαρκοβιανών Μοντέλων

Η χρήση Κρυφών Μαρκοβιανών μοντέλων είναι μια από τις κυρίαρχες μεθόδους που εφαρμόζονται τα τελευταία χρόνια, στις εφαρμογές αναγνώρισης φωνής. Το ότι υπήρξε μεγάλο ενδιαφέρον σε αυτό τον τομέα [29, 30] με αμέτρητες μελέτες, οδήγησε και στη δημιουργία πολλών προγραμματιστικών εργαλείων για τη δημιουργία, εκπαίδευση αλλά και αξιολόγηση Κρυφών Μαρκοβιανών Μοντέλων. Ένα από αυτά τα εργαλεία θα χρησιμοποιηθεί και σε αυτή την εργασία.

4.1 HTK (Hidden Markov Model Toolkit)

Το HTK είναι ένα πολυεπίπεδο εργαλείο για την κατασκευή Κρυφών Μαρκοβιανών Μοντέλων με χρήση ηλεκτρονικού υπολογιστή. Έχει διπλό copyright (1995-1999) Microsoft Corporation και (2001-2009) Cambridge University Engineering Department. Είναι υλοποιημένο για λειτουργικά συστήματα unix και μπορεί να χρησιμοποιηθεί για κάθε κυματομορφή στον χρόνο, αν και είναι σχεδιασμένο εξ ολοκλήρου για κατασκευή μοντέλων αναγνώρισης φωνής.



Εικόνα 4.1: Τα στάδια επεξεργασίας του φωνητικού σήματος στο HTK toolkit.

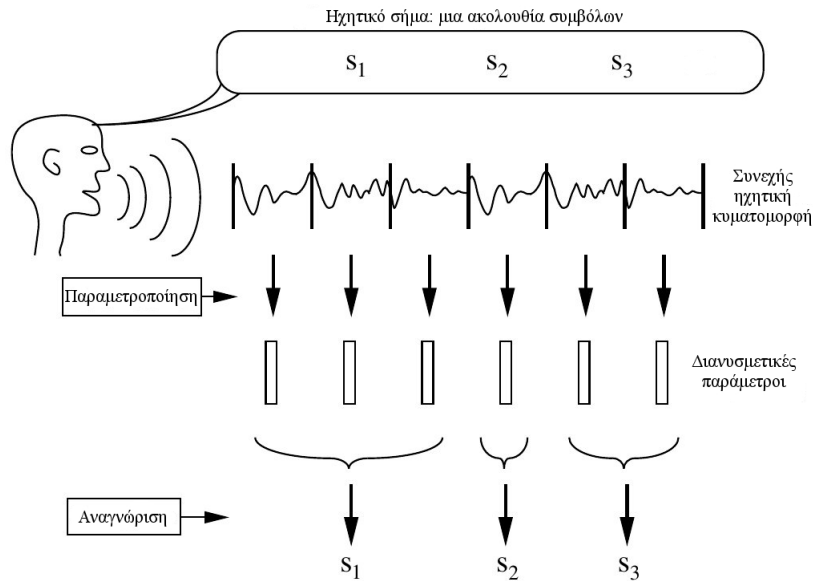
Όπως φαίνεται και στην εικόνα 4.1 υπάρχουν δύο κύρια στάδια επεξεργασίας. Το πρώτο στάδιο αναφέρεται σε αυτά τα εργαλεία που χρειάζονται για την δημιουργία του μοντέλου και κατ'επέκταση την εκπαίδευση του. Στάδιο στο οποίο είσοδος είναι οι ηχητικές κυματομορφές και τα αντίστοιχα φωνήματα (phonemes) τους, τα οποία αναφέρονται στην εκάστοτε κυματομορφή. Ως φωνήματα νοούνται οι ελάχιστες ηχητικές μονάδες που όταν ενωθούν σχηματίζουν ηχητικά σύνολα με γλωσσικό περιεχόμενο, όπως το μόρφημα και η λέξη. Το δεύτερο στάδιο έχει να κάνει με την αναγνώριση κυματομορφών και την ταύτιση τους με τα κατάλληλα φωνήματα, ώστε τελικώς να σχηματιστούν λέξεις. Αυτό το στάδιο που μπορεί να χρησιμοποιηθεί και ως ένα μέσο επαλήθευσης του μοντέλου άλλα και ως μέσο αναγνώρισης. Έχει ως είσοδο κυματομορφές και ως έξοδο την πιθανότερη ερμηνεία αυτής της κυματομορφής, σύμφωνα πάντα με την εκπαίδευση που έχει προηγηθεί. Ιδανικά στο μοντέλο θα ζητούσαμε για όλες τις κυματομορφές εκπαίδευσης (training data) να έχουμε απόλυτη ταύτιση/αναγνώριση, εάν οι ίδιες κυματομορφές χρησιμοποιηθούν ως άγνωστες κυματομορφές προς αναγνώριση (testing data).

4.2 Γενικά Στοιχεία

Τα συστήματα αναγνώρισης φωνής, γενικώς, υποθέτουν πως το ηχητικό σήμα είναι μια μορφή μηνύματος, κωδικοποιημένο σε μια ακολουθία ενός ή παραπάνω συμβόλων. Για το αντίστροφο πρόβλημα, το πρόβλημα της αναγνώρισης (αποκωδικοποίησης) μια ακολουθίας συμβόλων ενός ηχητικού σήματος, πρώτα η συνεχής κυματομορφή μεταφράζεται σε μια ακολουθία ισαπεχουσών διακριτών διανυσματικών παραμέτρων (εικόνα 4.2). Αυτή η ακολουθία των διανυσματικών παραμέτρων θεωρούμε πως σχηματίζει μια πιστή αναπαράσταση της ηχητικής κυματομορφής, λόγω της μικρής απόστασης μεταξύ των διανυσμάτων, μετά την διακριτοποίηση του συνεχούς σήματος (με διάρκεια διανύσματος 10ms περίπου), και θεωρούμε την *κυματομορφή στάσιμη*.

Ασχέτως εάν αυτή η υπόθεση περί στάσιμου σήματος, είναι μονάχα μέσα στα πλαίσια μιας προσέγγισης/γραμμικοποίησης σωστή, χρειάζεται ωστόσο να είναι η βασική μας υπόθεση για κάθε εφαρμογή των Κρυφών Μαρκοβιανών Μοντέλων. Ο ρόλος του συστήματος αναγνώρισης είναι η ταύτιση/απεικόνιση/σύνδεση μεταξύ των ακολουθιών διανυσματικών παραμέτρων με τις ζητούμενες ακολουθίες συμβόλων, για την δημιουργία λέξεων σε φυσική γλώσσα.

Δύο βασικές δυσκολίες προκύπτουν με φυσικό τρόπο. Η πρώτη έχει να κάνει με την *ένα-προς-ένα αντιστοίχιση συμβόλων-ομιλίας*, διότι διαφορετικά σύμβολα τελικά μπορεί να οδηγήσουν σε όμοιους ήχους, όπως γίνεται λόγω μεγάλων αποκλίσεων σε *προφορά*, σε *παραμόρφωσης* λόγω του περιβάλλοντος ή της ποιότητας δειγματοληψίας κτλ. Η δεύτερη έχει να κάνει κυρίως με την *συχνότητα*, δηλαδή την διάρκεια του κάθε συμβόλου, σε κάθε καταγεγραμμένη αναφορά του. Αυτό κάνει αδύνατη την αντιμετώπιση ηχητικών κυματομορφών σαν μια ακολουθία μεμονωμένων στατικών προτύπων. *Θα γίνει φανερό στην συνέχεια πως η περίπτωση των ηχητικών σημάτων, μεταφέρεται ως ένα ακριβές ανάλογο, στην περίπτωση της κατηγορίας κυματομορφών που είναι το αντικείμενο διερεύνησης αυτής της εργασίας.*



Εικόνα 4.2: Μεταφορά ηχητικού σήματος από τον συνεχή στο διακριτό χρόνο, μέσω διανυσματικών παραμέτρων.

Το πρόβλημα της επιλογής ορίων μεταξύ λέξεων μπορεί να αποφευχθεί εάν περιορίσουμε το ζητούμενο, σε μεμονωμένη αναγνώριση λέξεων. Όπως φαίνεται και στην εικόνα 4.3 η ηχητική κυματομορφή αντιστοιχεί σε ένα σύμβολο, το οποίο έχει οριστεί μέσα σε ένα λεξιλόγιο. Ακόμη και αν δείχνει η υποδιαίρεση του προβλήματος άωφελη, έχει μεγάλο φάσμα εφαρμογής σε διάφορα πρακτικά μοντέλα. Ακόμη, η λύση αυτού του πρωταρχικού βήματος είναι καίριας σημασίας για την επίλυση πολυσύνθετων προβλημάτων αναγνώρισης προτύπων.

4.3 Μεμονωμένη Αναγνώριση Λέξεων

Όπως ορίστηκε και στα προηγούμενα κεφάλαια, έχοντας μια ακολουθία παρατηρήσεων O , που ορίζονται ως

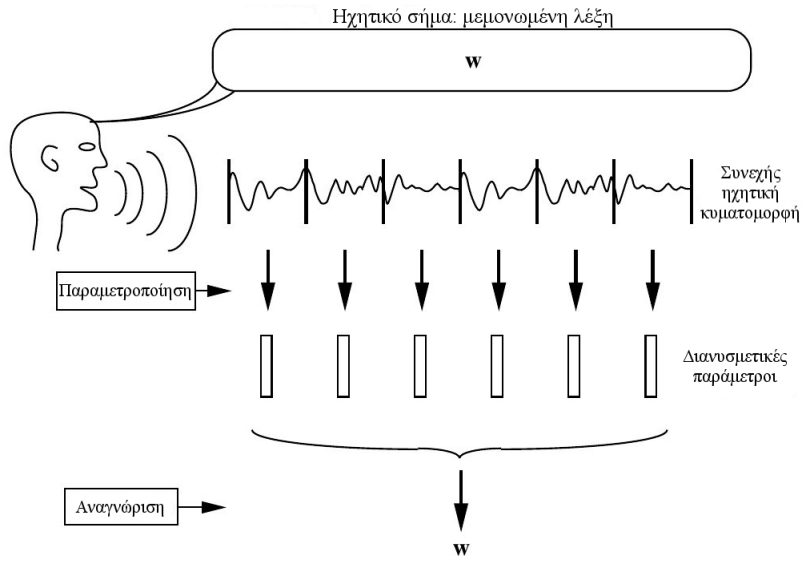
$$O = o_1, o_2, \dots, o_T \quad (4.1)$$

όπου o_t είναι το ηχητικό διάνυσμα παρατηρούμενο στον χρόνο t . Το πρόβλημα της μεμονωμένης αναγνώρισης λέξης μπορεί να θεωρηθεί ως τον υπολογισμό τουλάχιστον

$$\arg \max_i \{P(w_i|O)\} \quad (4.2)$$

με w_i να είναι η i -οστή λέξη του λεξιλογίου. Αυτή η πιθανότητα θα υπολογιστεί έμμεσα από τον κανόνα του Bayes

$$P(w_i|O) = \frac{P(O|w_i)P(w_i)}{P(O)} \quad (4.3)$$



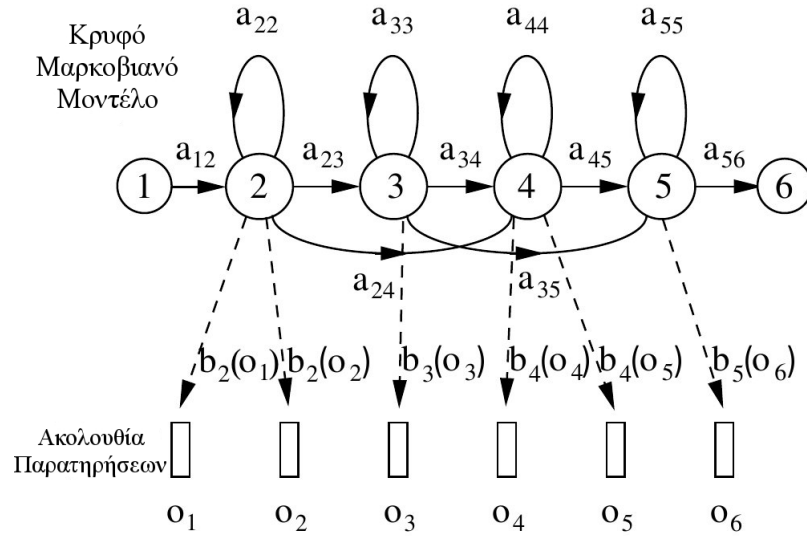
Εικόνα 4.3: Μεμονωμένο πρόβλημα αναγνώρισης λέξης.

Έτσι για ένα δεδομένο, εκ των προτέρων, σει πιθανοτήτων $P(w_i)$, η πιο πιθανή λέξη εξαρτάται μόνο από την πιθανοφάνεια $P(O|w_i)$. Δεδομένης της μεγάλης διαστάσεως της ακολουθίας παρατήρησης, ο απευθείας υπολογισμός της από κοινού δεσμευμένης πιθανότητας $P(o_1, o_2, \dots | w_i)$ τίθεται δύσκολος από πλευράς πολυπλοκότητας. Ωστόσο, σε ένα παραμετρικό μοντέλο παραγωγής λέξεων (όπως είναι το Κρυφό Μαρκοβιανό Μοντέλο), η εκτίμηση από τα δεδομένα είναι δυνατή αφού το πρόβλημα του υπολογισμού της $P(O|w_i)$ μπορεί να αντικατασταθεί από ένα απλούστερο πρόβλημα, αυτό του υπολογισμού των παραμέτρων ενός Κρυφού Μαρκοβιανού Μοντέλου.

Στη μέθοδο για την αναγνώριση φωνής, υποθέτουμε πως η ακολουθία διανυσμάτων παρατηρήσεων που αντιστοιχούν σε κάθε λέξη, παράγεται, για την κάθε λέξη, και από ένα Κρυφό Μαρκοβιανό Μοντέλο της ακόλουθης μορφής (εικόνα 4.4). Στο πεπερασμένο αυτόματο που προκύπτει, έχουμε για την αλλαγή κατάστασης (για κάποιο χρόνο t) σε μία κατάσταση j , την κατασκευή ενός ηχητικού διανύσματος o_t με πιθανότητα $b_j(o_t)$. Ακόμη, γνωρίζουμε πως και οι μεταβάσεις μεταξύ καταστάσεων i και j ορίζονται από έναν πίνακα διακριτών πιθανοτήτων με στοιχεία a_{ij} . Στο παράδειγμα υπάρχει ένα μοντέλο 6 καταστάσεων, που με την ακολουθία καταστάσεων $S = \{1, 2, 2, 3, 4, 4, 5, 6\}$ παράγεται η ακολουθία παρατήρησης $O = \{o_1, o_2, o_3, o_4, o_5, o_6\}$. Στο μοντέλο αυτό βλέπουμε ότι οι καταστάσεις 1 και 6 δεν δημιουργούν κάποια ανάδραση και θεωρούνται αρχική και τελική κατάσταση του μοντέλου, αυτό γίνεται για την ευκολία δημιουργίας σύνθετων μοντέλων.

Η από κοινού πιθανότητα ότι η ακολουθία O δημιουργείται από το μοντέλο λ μέσω της ακολουθίας S υπολογίζεται απλά από το γινόμενο των πιθανοτήτων μετάβασης και παρατήρησης (εξόδου).

$$P(O, S | \lambda) = a_{12}b_2(o_1)a_{22}b_2(o_2)a_{23}b_3(o_3)\dots \quad (4.4)$$



Εικόνα 4.4: Παράδειγμα δημιουργίας ενός Κρυφού Μαρκοβιανού Μοντέλου.

Στην πράξη όμως, η ακολουθία παρατήρησης O είναι το γνωστό μέρος του προβλήματος και η κρυφή στοχαστική διεργασία είναι η μεταβολή καταστάσεων S .

Με δεδομένο ότι η S είναι άγνωστη, η απαιτούμενη πιθανοφάνεια υπολογίζεται αθροίζοντας για όλες τις πιθανές μεταβάσεις $S = s(1), s(2), s(3), \dots, s(T)$, δηλαδή

$$P(O|\lambda) = \sum_S \alpha_{s(0)s(1)} \prod_{t=1}^T b_{s(t)}(o_t) \alpha_{s(t)s(t+1)} \quad (4.5)$$

όπου ορίζεται $s(0)$ για το συγκεκριμένο μοντέλο ως κατάσταση εισόδου και $s(T+1)$ ως κατάσταση εξόδου.

Ως ένα εναλλακτικό της ανωτέρω εξίσωσης για την πιθανοφάνεια είναι η επιλογή μονάχα της πιθανότερης ακολουθίας καταστάσεων, που είναι

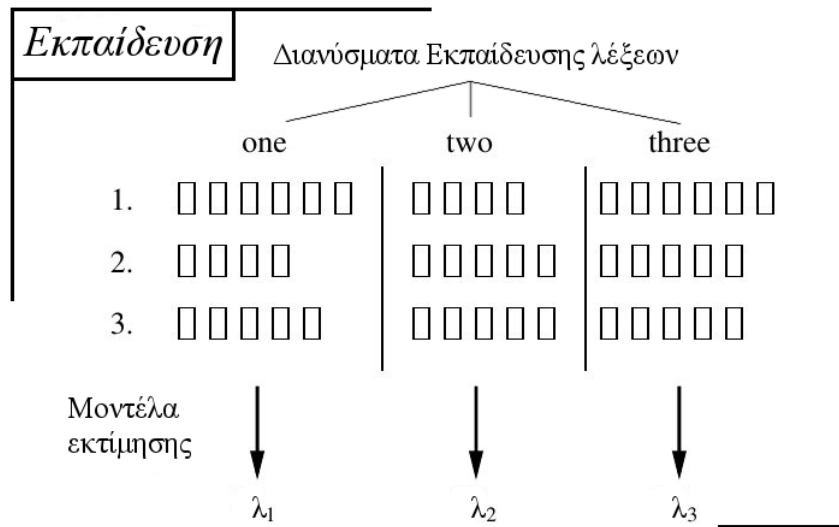
$$\hat{P}(O|\lambda) = \max_S \left\{ \alpha_{s(0)s(1)} \prod_{t=1}^T b_{s(t)}(o_t) \alpha_{s(t)s(t+1)} \right\} \quad (4.6)$$

όπου η αρχική θεώρηση μιλά για ένα σύνολο μοντέλων λ_i που αντιστοιχούν σε λέξεις w_i με

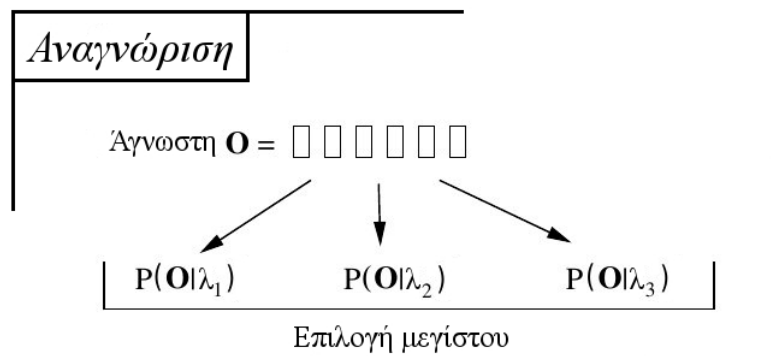
$$P(O|w_i) = P(O|\lambda_i) \quad (4.7)$$

και γνωστές τις παραμέτρους α_{ij} και $b_j(o_t)$ για κάθε μοντέλο λ_i . Εδώ είναι και όλη η γοητεία του μοντέλου, διότι, με δεδομένο ένα σύνολο κυματομορφών εκπαίδευσης για ένα μοντέλο λ , οι παράμετροι αυτού του μοντέλου μπορούν να προσδιοριστούν αυτόματα, μέσα από μια διαδικασία επανεκτίμησης. Έτσι εάν υπάρχει ικανό δείγμα αντιπροσωπευτικών παραδειγμάτων για κάθε λέξη, μπορεί να κατασκευαστεί ένα πλήρες μοντέλο, από πολλαπλές διαφορετικές πηγές για την κάλυψη αρκετών διαφορετικών ακουστικών (προφορών). Η μεμονωμένη αναγνώριση λέξεων συνοψίζεται

στις εικόνες 4.5, 4.6.



Εικόνα 4.5: Στάδιο εκπαίδευσης μοντέλου για μεμονωμένη αναγνώριση λέξης σχηματικά.



Εικόνα 4.6: Στάδιο αναγνώρισης για μεμονωμένες λέξεις.

Αρχικά εκπαιδεύεται ένα μοντέλο για κάθε λέξη, μέσα από έναν αριθμό παραδειγμάτων για την κάθε λέξη. Στην περίπτωση του σχήματος το λεξιλόγιο αποτελείται από τις τρεις λέξεις "one", "two" και "three". Για την αναγνώριση κάθε μιας από αυτές τις λέξεις, η πιθανοφάνεια του κάθε μοντέλου ως προς την παραγωγή της λέξης αυτής, υπολογίζεται, και η μέγιστη εξ αυτών θεωρείται πως αναγνωρίζει την λέξη.

4.4 Πιθανότητες Παρατήρησης (εξόδου)

Το HTK toolkit είναι κατασκευασμένο ώστε να αναπαριστά τις συνεχείς κατανομές παρατήρησης, με μείγμα πυκνοτήτων Gauss. Το κάθε διάνυσμα παρατήρησης σε χρόνο t κατακερματίζεται σε έναν αριθμό N ανεξαρτήτων ροών δεδομένων o_m . Η σχέση για τον υπολογισμό των $b_j(o_t)$ τότε

θα είναι

$$b_j(o_t) = \prod_{n=1}^N \left[\sum_{m=1}^{M_n} c_{jnm} \mathcal{N}(o_{nt}; \mu_{jnm}, \Sigma_{jnm}) \right]^{\gamma_n} \quad (4.8)$$

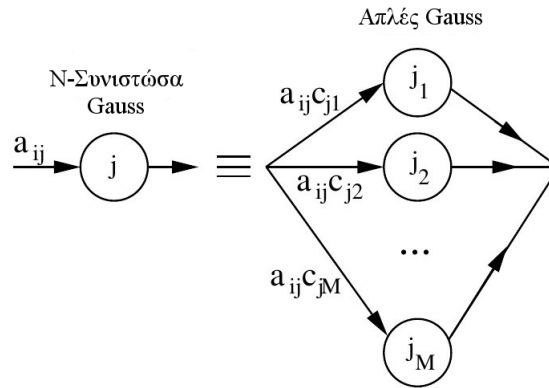
όπου n ο αριθμός των συνιστωσών Gauss στη ροή n , c_{jnm} το βάρος της m συνιστώσας, γ_n είναι το βάρος ροής και $\mathcal{N}(\cdot; \mu, \Sigma)$ η πολυμετάβλητη κατανομή Gauss

$$\mathcal{N}(o; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} e^{-\frac{1}{2}(o-\mu)^T \Sigma^{-1}(o-\mu)} \quad (4.9)$$

όπου k το μέγεθος της διάστασης του o .

4.5 Επανεκτίμηση Baum-Welch

Για να ορίσουμε τις παραμέτρους ενός Κρυφού Μαρκοβιανού Μοντέλου θα πρέπει να κάνουμε μια αρχική πρόβλεψη, έτσι ώστε στη συνέχεια, όπως γίνεται σε κάθε επαναληπτικό αλγόριθμο, να έχουμε βελτιώσεις τρέχοντας τον αλγόριθμο Baum-Welch. Οι συνιστώσες μπορούν να θεωρηθούν μια ειδική μορφή υπο-κατάστασης όπου οι πιθανότητες μετάβασης είναι και τα βάρη των συνιστωσών, όπως φαίνεται και στην εικόνα 4.7.



Εικόνα 4.7: Αναπαράσταση μείγματος Gauss για τις συνιστώσες.

Το κύριο πρόβλημα λοιπόν είναι ο υπολογισμός των μέσων, και των διασπορών του κάθε μοντέλου, όπου η κατανομή κάθε παρατήρησης (εξόδου), είναι μια συνιστώσα Gauss

$$b_j(o_t) = \frac{1}{\sqrt{(2\pi)^k |\Sigma_j|}} e^{-\frac{1}{2}(o_t - \mu_j)^T \Sigma_j^{-1}(o_t - \mu_j)} \quad (4.10)$$

Εάν υπήρχε μόνο μια κατάσταση j στο μοντέλο, η εκτίμηση παραμέτρων θα ήταν εύκολη. Η μέγιστη πιθανοφάνεια για κάθε μ_j και Σ_j θα ήταν οι μέσοι όροι

$$\hat{\mu}_j = \frac{1}{T} \sum_{t=1}^T o_t \quad (4.11)$$

και

$$\hat{\Sigma}_j = \frac{1}{T} \sum_{t=1}^T (o_t - \mu_j)(o_t - \mu_j)^T \quad (4.12)$$

Στην πράξη βέβαια υπάρχουν πολλές καταστάσεις και δεν υπάρχει (λόγω της κρυφής της φύσης) άμεση αντιστοίχιση των διανυσμάτων παρατήρησης με τις καταστάσεις. Ωστόσο, εάν υπάρξουν κάποιες προσεγγιστικές τιμές για τα διανύσματα των καταστάσεων, τότε οι εξισώσεις (5.11), (5.12) μπορούν να χρησιμοποιηθούν για την αρχικοποίηση των παραμέτρων. Με χρήση του αλγορίθμου Viterbi βρίσκεται η ακολουθία καταστάσεων με την μέγιστη πιθανοφάνεια και αντιστοιχίζονται τα διανύσματα παρατήρησης ξανά. Στη συνέχεια θα πρέπει να χρησιμοποιηθούν πάλι οι εξισώσεις (5.11), (5.12) γιατί επανεκτίμηση των τιμών. Τέλος, ολόκληρη η διαδικασία επαναλαμβάνεται μέχρι να συγκλίνουν όλες οι τιμές.

Αφού η πιθανοφάνεια της κάθε ακολουθίας παρατηρήσεων βασίζεται στην άθροιση όλων των πιθανών ακολουθιών καταστάσεων, κάθε διάνυσμα παρατήρησης o_t , προσφέρει στον υπολογισμό των παραμέτρων (μέγιστης πιθανοφάνειας) για κάθε κατάσταση j . Δηλαδή, αντί να συνδέουμε κάθε διάνυσμα παρατήρησης σε μια συγκεκριμένη κατάσταση, κάθε διάνυσμα παρατήρησης συνδέεται με κάθε κατάσταση, ανάλογα με το πόσο πιθανό είναι το μοντέλο να βρίσκεται σε αυτή τη κατάσταση όταν παρατηρείται το διάνυσμα.

Θεωρώντας πως με $L_j(t)$ συμβολίζουμε την πιθανότητα του να βρισκόμαστε στην κατάσταση j σε χρόνο t , τότε έχουμε τις εξής μετατροπές στις παραπάνω εξισώσεις, για τη δημιουργία σταθμισμένων μέσων

$$\hat{\mu}_j = \frac{\sum_{t=1}^T L_j(t) o_t}{\sum_{t=1}^T L_j(t)} \quad (4.13)$$

και

$$\hat{\Sigma}_j = \frac{\sum_{t=1}^T L_j(t) (o_t - \mu_j)(o_t - \mu_j)^T}{\sum_{t=1}^T L_j(t)} \quad (4.14)$$

Οι παραπάνω εξισώσεις είναι οι επανεκτιμήσεις Baum-Welch για το μοντέλο και βασίζονται στον υπολογισμό του $L_j(t)$, που γίνεται από αλγόριθμο Forward-Backward. Για την εμπρόσθια πιθανότητα έχουμε

$$\alpha_j(t) = P(o_1, \dots, o_t, x(t) = j | \lambda) \quad (4.15)$$

με τον αναδρομικό τύπο έχουμε

$$\alpha_j(t) = \left[\sum_{i=2}^{N-1} \alpha_i(t-1) \alpha_{ij} \right] b_j(o_t) \quad (4.16)$$

τα όρια προκύπτουν λόγω της πρώτης και της τελευταίας κατάστασης, όπου δεν εκπέμπουν (συνδέονται σε χρόνο $t + 1$) σε άλλες καταστάσεις. Οι αρχικές συνθήκες για την αναδρομή είναι

$$\alpha_1(1) = 1 \quad (4.17)$$

$$\alpha_j(1) = \alpha_{1j} b_j(o_1) \quad (4.18)$$

για $1 < j < N$ έχουμε

$$\alpha_N(T) = \sum_{i=2}^{N-1} \alpha_i(T) \alpha_{iN} = P(O|\lambda) \quad (4.19)$$

Η οπίσθια πιθανότητα

$$\beta_j(t) = P(o_{t+1}, \dots, o_T | x(t) = j, \lambda) \quad (4.20)$$

$$\beta_i(t) = \sum_{j=2}^{N-1} \alpha_{ij} b_j(o_{t+1}) \beta_j(t+1) \quad (4.21)$$

με αρχική συνθήκη

$$\beta_i(T) = \alpha_{iN} \quad (4.22)$$

για $1 < j < N$ έχουμε

$$\beta_1(1) = \sum_{j=2}^{N-1} \alpha_{1j} b_j(o_1) \beta_j(1) \quad (4.23)$$

Παρατηρούμε την ασυμμετρία μεταξύ εμπρόσθιας και οπίσθιας μεταβλητής πιθανότητας ως προς τον τύπο τους. Η μία είναι από κοινού πιθανότητα, ενώ η δεύτερη δεσμευμένη. Αυτό γίνεται με σκοπό τον υπολογισμό της $L_j(t)$ από το γινόμενο των πιθανοτήτων

$$\alpha_j(t) \beta_j(t) = P(O, x(t) = j | \lambda) \quad (4.24)$$

Έτσι

$$L_j(t) = P(x(t) = j | O, \lambda) = \frac{P(O, x(t) = j | \lambda)}{P(O | \lambda)} = \frac{1}{P(O | \lambda)} \alpha_j(t) \beta_j(t) \quad (4.25)$$

Συνοψίζοντας τα βήματα της επανεκτίμησης παραμέτρων ενός μοντέλου λ , έχουμε

1. Για κάθε διάνυσμα/πίνακα παράμετρο που επανεκτιμούμε, δημιουργείται ένας χώρος αποθήκευσης για τον παρονομαστή και τον αριθμητή των αθροίσεων (εξισώσεις 5.13, 5.14) και αναφέρονται ως *αθροιστές*.
2. Υπολογισμός των εμπρόσθιων και οπίσθιων μεταβλητών πιθανοτήτων για κάθε κατάσταση j και κάθε χρόνο t .
3. Για κάθε κατάσταση j και χρόνο t , με χρήση της πιθανότητας $L_j(t)$ και του τρέχοντος διανύσματος παρατήρησης o_t , γίνεται η ανανέωση των *αθροιστών* για την κάθε κατάσταση.
4. Χρήση των τιμών του τελικού *αθροιστή* για τον υπολογισμό των καινούργιων παραμέτρων.
5. Στην περίπτωση που η τιμή της πιθανότητας $P(O|\lambda)$ για αυτή την επανάληψη, δεν είναι μεγαλύτερη της προηγούμενης, θεωρούμε πως η μέθοδος έφτασε σε σύγκλιση, αλλιώς η επανάληψη συνεχίζεται με τις νέες τιμές, οι οποίες θα τροφοδοτήσουν την ανάδραση.

Για την εκπαίδευση του μοντέλου με παραπάνω από μια ακολουθία παρατήρησης (εκπαίδευσης), έχουμε την επανάληψη των βημάτων 2 και 3 για κάθε εναλλακτική ακολουθία παρατήρησης. Λόγω του μεγάλου αριθμού υπολογισμών (γινομένων) των εμπρόσθιων και οπίσθιων πιθανοτήτων, οι αριθμητικές τιμές γίνονται εξαιρετικά μικρές και έτσι γίνεται χρήση λογαρίθμων για περιορισμό των αριθμητικών σφαλμάτων.

4.6 Αναγνώριση και Αποκωδικοποίηση Viterbi

Προηγουμένως αναφέραμε τον αναδρομικό αλγόριθμο υπολογισμού της εμπρόσθιας πιθανότητας. Αυτός ο αλγόριθμος μπορεί να χρησιμοποιηθεί για την εύρεση του μοντέλου λ , που έχει την μέγιστη τιμή από τις πιθανότητες $P(O|\lambda_i)$ και έτσι να χρησιμοποιηθεί για αναγνώριση.

Στην πράξη το κομμάτι της αναγνώρισης, είναι προτιμότερο να βασίζεται στην ακολουθία καταστάσεων με την μέγιστη πιθανοφάνεια, αφού αυτό μπορεί εύκολα να γενικευτεί στην περίπτωση συνεχούς λόγου (ακολουθίας λέξεων), σε αντίθεση με την χρήση ολικής πιθανότητας που υστερεί. Για τον υπολογισμό της πιθανοφάνειας χρησιμοποιείται πρακτικά ο ίδιος αλγόριθμος με πριν, με την μόνη διαφορά ότι οι αθροίσεις αντικαθίσταται με τον τελεστή maximum. Για ένα μοντέλο λ , με $\varphi_j(t)$ να αναπαριστά τη μέγιστη πιθανοφάνεια των διανυσμάτων παρατήρησης o_1, \dots, o_t , για τρέχουσα κατάσταση j και χρόνο t . Η μερική πιθανοφάνεια μπορεί να υπολογιστεί από την αναδρομή

$$\varphi_j(t) = \max_i \{ \varphi_i(t-1) \alpha_{ij} \} b_j(o_t) \quad (4.26)$$

όπου

$$\varphi_1(1) = 1 \quad (4.27)$$

για $1 < j < N$

$$\varphi_j(1) = \alpha_{1j} b_j(o_1) \quad (4.28)$$

Η μέγιστη πιθανοφάνεια $\hat{P}(O|\lambda)$ δίνεται απόδειξη

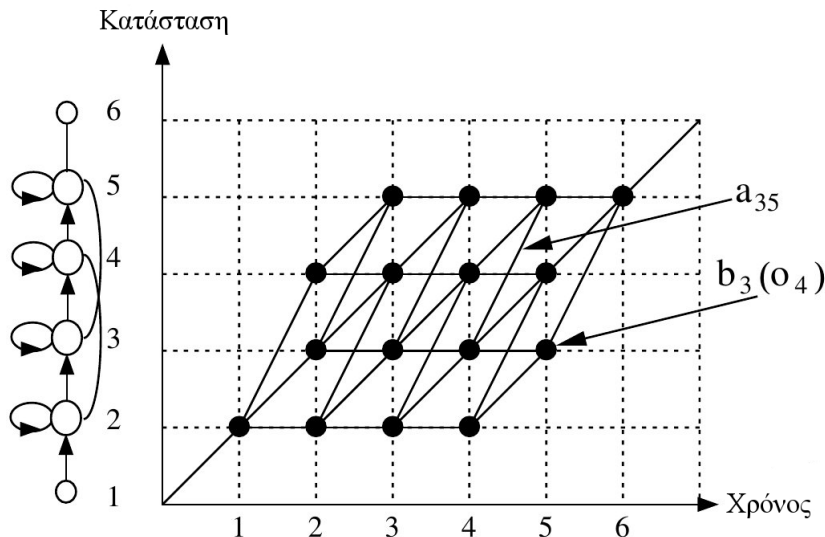
$$\varphi_N(T) = \max_i \{ \varphi_i(T) \alpha_{iN} \} \quad (4.29)$$

Για την επανεκτίμηση γίνεται χρήση λογαρίθμου και η ανάδραση της εξίσωσης 5.26 γίνεται

$$\psi_j(t) = \max_i \{ \psi_i(t-1) + \log(\alpha_{ij}) \} + \log(b_j(o_t)) \quad (4.30)$$

Αυτή η αναδρομή είναι κατά βάση και ο αλγόριθμος Viterbi. Όπως φαίνεται στην εικόνα 4.8, ο αλγόριθμος μπορεί να σχηματιστεί ως μια αναζήτηση του καλύτερου μονοπατιού σε ένα πίνακα όπου η οριζόντια διάσταση αναπαριστά τις καταστάσεις του μοντέλου και η κάθετη αναπαριστά

χρονικές στιγμές.



Εικόνα 4.8: Αλγόριθμος Viterbi για μεμονωμένη αναγνώριση λέξης.

Κάθε κύκλος στην εικόνα αναπαριστά την λογαριθμική πιθανότητα του να παρατηρηθεί κάποια κατάσταση σε κάποιο χρόνο, και με τις ακμές έχουμε τις *λογαριθμικές πιθανότητες μετάβασης* μεταξύ καταστάσεων. Η συνολική πιθανότητα για κάποιο μονοπάτι υπολογίζεται (λόγω λογαρίθμων) από το άθροισμα των λογαριθμικών πιθανοτήτων μετάβασης, διατρέχοντας αυτό το μονοπάτι. Τα μονοπάτια διατρέχονται από αριστερά στα δεξιά, για χρονική στιγμή t κάθε *μερικό μονοπάτι* $\psi_i(t-1)$ είναι γνωστό για όλες τις καταστάσεις i , έτσι με την εξίσωση (5.30) μπορεί να υπολογιστεί το $\psi_j(t)$, άρα και επεκτείνονται τα μερικά μονοπάτια στο επόμενο χρονικό βήμα.

4.7 Αναγνώριση Συνεχούς Λόγου

Όπως έχει ήδη παρουσιαστεί, ο συνεχής λόγος μπορεί να εκφραστεί ως ακολουθία διασυνδεδεμένων μοντέλων, δηλαδή ως ακολουθία λέξεων. Είναι θέμα ορισμού βέβαια το αν θα θεωρήσουμε διασυνδέσεις φωνημάτων ή διασυνδέσεις λέξεων, διότι αυτό έχει να κάνει με το πόσο λεπτομερές θέλουμε να είναι το μοντέλο μας. Οι δύο καταστάσεις, μια αρχική και μια τελική δρουν ως συνδετικοί κρίκοι για την διασύνδεση αυτή.

Υπάρχουν ωστόσο και πρακτικές δυσκολίες. Αυτές προκύπτουν κυρίως από την αδυναμία αναγνώρισης, σε συνεχή λόγο, να διαχωριστούν εμφανώς τα όρια μεταξύ των λέξεων/φωνημάτων (μοντέλων). Υπάρχει η δυνατότητα να υπάρξει ορισμός αυτών των ορίων με το χέρι, πράγμα όμως που καθιστά αυτή τη μέθοδο ανεφάρμοστη για κάποιο μεγάλο σύνολο δεδομένων. Τα μοντέλα με αυτό τον τρόπο μπορούν να εκπαιδευτούν με τη μέθοδο της μεμονωμένης λέξης που παρουσιάστηκε προηγουμένως. Τα αποτελέσματα αυτής της πρακτικής βέβαια οδηγούν σε φτωχά μοντέλα με χαμηλά ποσοστά αναγνώρισης. Η μέχρι τώρα ανάλυση δουλεύει σαν ένα καλό κομμάτι προ-

ετοιμασίας του μοντέλου (bootstrap operation). Για το κύριο μέρος γίνεται χρήση της λεγόμενης ενσωματωμένης εκπαίδευσης (embedded training).

Η ενσωματωμένη εκπαίδευση χρησιμοποιεί την ίδια διαδικασία Baum-Welch όπως και στην μεμονωμένη περίπτωση αλλά αντί να εκπαιδεύει το κάθε μοντέλο ξεχωριστά, εκπαιδεύει όλα τα μοντέλα παράλληλα. Αυτό γίνεται με τον ακόλουθο τρόπο

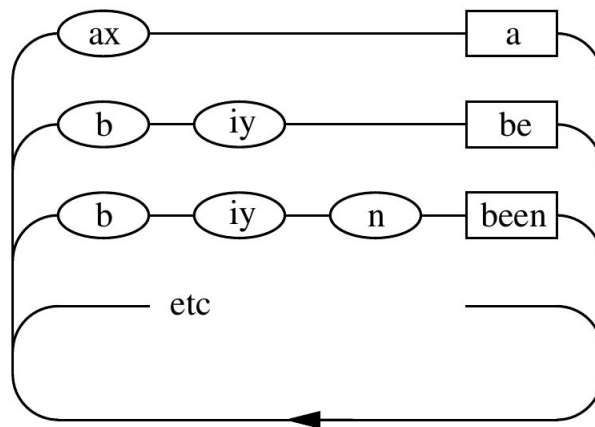
1. Δημιουργία και μηδενισμός των αθροιστών για όλες τις παραμέτρους του μοντέλου.
2. Ανάγνωση της επόμενης ακολουθίας εκπαίδευσης.
3. Κατασκευή σύνθετου μοντέλου από την συνένωση των μοντέλων με την σειρά που τα σύμβολα του λεξιλογίου αντιστοιχούν στα φωνήματα/λέξεις εκπαίδευσης.
4. Υπολογισμός των εμπρόσθιων και οπίσθιων πιθανοτήτων του μοντέλου.
5. Χρήση των εμπρόσθιων και οπίσθιων μεταβλητών, για τον υπολογισμό της πιθανότητας μια κατάσταση να βρίσκεται σε κάποια συγκεκριμένη χρονική στιγμή. Και ανανέωση των αθροιστών.
6. Επανάληψη από το βήμα 2 μέχρι να επεξεργαστούν όλες οι ακολουθίες εκπαίδευσης.
7. Χρήση των αθροιστών για τον υπολογισμό των νέων παραμέτρων για όλα τα μοντέλα.

Φυσικά όλα τα παραπάνω βήματα μπορούν να επαναληφθούν όσες φορές χρειαστεί, έτσι ώστε να επιτευχθεί η απαραίτητη σύγκλιση. Η ταυτοποίηση των ορίων για κάθε σύμβολο στις ακολουθίες εκπαίδευσης δεν πλέον είναι αναγκαία για την παραπάνω διαδικασία, αλλά αναγκαίες είναι οι ετικέτες (transcriptions) για κάθε σύμβολο που παρουσιάζεται στις ακολουθίες αυτές.

Η διαφορά που υπάρχει στον αλγόριθμο Viterbi για τον συνεχή λόγο εκφράζεται ως Token Passing Model. Εν συντομία, το μοντέλο αυτό κάνει την ευθυγράμμιση του μονοπατιού των καταστάσεων ρητή. Για κάθε κατάσταση j ενός μοντέλου, στη χρονική στιγμή t κρατείται ένα μετατόπισιμο όρισμα (token), που εκτός άλλων διατηρεί και την μερική λογαριθμική πιθανότητα $\psi_j(t)$. Αυτό το όρισμα αναπαριστά μια μερική ταύτιση μεταξύ της ακολουθίας παρατήρησης o_1, \dots, o_t και του μοντέλου, με την προϋπόθεση ότι το μοντέλο βρίσκεται στην κατάσταση j στον χρόνο t . Ο αλγόριθμος για την επέκταση του μονοπατιού που χρησιμοποιεί τον αναδρομικό τύπο (5.30) αντικαθίσταται από τον ανάλογο αλγόριθμο (token passing algorithm) που εκτελείται για κάθε χρονική στιγμή. Τα βασικά του βήματα είναι

1. Μετάφερε ένα αντίγραφο του κάθε ορίσματος της κατάστασης i σε όλες τις καταστάσεις j που συνδέονται με αυτή, αυξάνοντας την λογαριθμική πιθανότητα του αντιγράφου κατά $\log[a_{ij}] + \log[b_j(o_t)]$.
2. Για κάθε όρισμα που υπάρχει σε κάθε κατάσταση, απόρριψε όλα τα ορίσματα εκτός από αυτό με την μεγαλύτερη πιθανότητα.

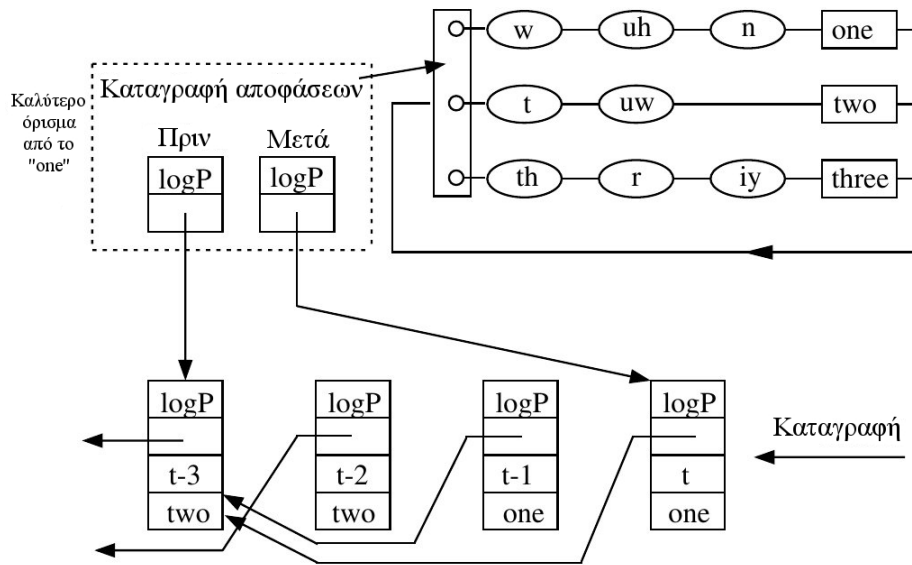
Υποθέτοντας πως η επιτρεπτή ακολουθία μοντέλων είναι ορισμένη με ένα δίκτυο διακριτών καταστάσεων, όπως για παράδειγμα στη εικόνα 4.9.



Εικόνα 4.9: Δίκτυο μοντέλων για αναγνώριση συνεχούς λόγου.

Όπου δείχνει ένα απλό δίκτυο στο οποίο η κάθε λέξη ορίζεται ως μια ακολουθία φωνημάτων (μοντέλων) και όλες οι λέξεις είναι τοποθετημένες σε ένα βρόχο. Σε αυτό το δίκτυο, τα οβάλ κουτιά είναι μοντέλα (φωνημάτων) και τα ορθογώνια είναι τελικοί κόμβοι λέξεων. Όταν το καλύτερο όρισμα φτάσει στο τέλος της φράσης (του λόγου), το μονοπάτι που ακολούθησε μέχρι εκεί συμπίπτει με την ακολουθία αναγνώρισης.

Το ιστορικό της διαδρομής ενός ορίσματος μέσα σε ένα δίκτυο μπορεί να καταγραφεί ως εξής. Κάθε όρισμα έχει έναν δείκτη που αποκαλείται κλάδος λέξης (*word end link*). Όταν ένα όρισμα προχωρά από την κατάσταση εξόδου κάποιας λέξης στην είσοδο κάποιας άλλης, αυτή η μετάβαση δείχνει ένα πιθανό *σύνоро μεταξύ λέξεων*. Έτσι υπάρχει μια καταγραφή (Word Link Record) που περιέχει την *ταυτότητα της λέξης* από την οποία το όρισμα έχει μόλις προκύψει και την τρέχουσα τιμή από τον *δείκτη του ορίσματος*. Ο πραγματικός δείκτης του ορίσματος αντικαθίσταται με ένα δείκτη στη καινούργια καταγραφή, όπως φαίνεται και στην εικόνα 4.10.



Εικόνα 4.10: Καταγραφή πιθανών ορίων λέξεων και περιεχόμενο ορισμάτων.

Όταν έχουν εξεταστεί όλες οι άγνωστες ακολουθίες λόγου, οι καταγραφές που έχουν δείκτη πάνω στο πιο ταιριαστό (πιθανό) όρισμα, μπορούν να ανατρεχθούν ώστε να προκύψει η πιο πιθανή ακολουθία λέξεων. Ταυτόχρονα μπορεί να εξαχθούν και οι θέσεις των ορίων.

Παρόλο που η εκπαίδευση και η αναγνώριση μοντέλων όπως περιγράφηκαν μπορούν να παράγουν υψηλής επίδοσης συστήματα αναγνώρισης, αυτά τα συστήματα μπορούν να βελτιωθούν περαιτέρω με την οικειοποίηση του μοντέλου στον ομιλητή. Αυτό γίνεται με επανεκπαίδευση του μοντέλου χρησιμοποιώντας καινούργια δεδομένα εκπαίδευσης (training data) τα οποία θα (επαν)ορίζουν κάποιες δεδομένες ετικέτες (λέξεις μέσα από το λεξιλόγιο) για την καλύτερη προσαρμογή του.

Κεφάλαιο 5

Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines - SVMs)

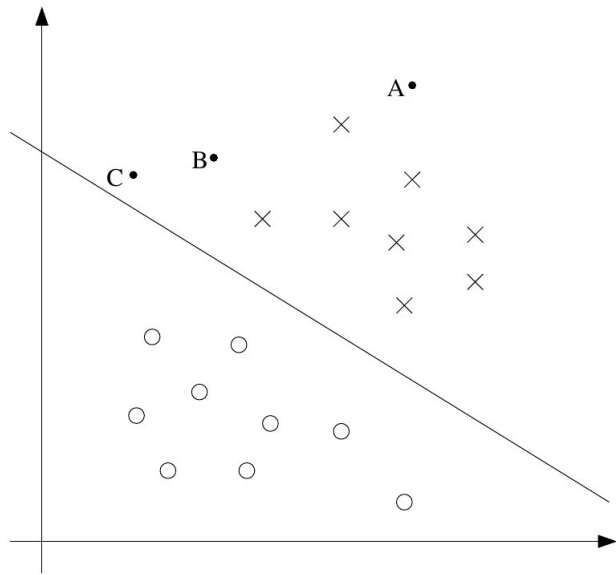
Οι Μηχανές Διανυσμάτων Υποστήριξης έχουν γνωρίσει μεγάλη επιτυχία στο επιστημονικό πεδίο της αναγνώρισης προτύπων και κατ' επέκταση στην κατηγοριοποίηση [31]. Λόγω της μεγάλης υπολογιστικής πολυπλοκότητας των συνηθισμένων στατιστικών μοντέλων, η χρήση SVM για εφαρμογές που βρίσκουν πρακτική στον πραγματικό κόσμο, είναι μια καλή επιλογή. Το μοντέλο δηλαδή, θα ανακαλύψει όχι την βέλτιστη αλλά μια ικανοποιητική λύση, σε πολύ λιγότερο χρόνο σε σχέση με την πλήρη λύση του εκάστοτε προβλήματος. Η πρακτική έχει αποδείξει γενικά ότι οι ταξινομητές που λειτουργούν με SVMs υπερτερούν των υπολοίπων (όχι όμως για πολύ μεγάλα σετ δεδομένων).

5.1 Γενική περιγραφή

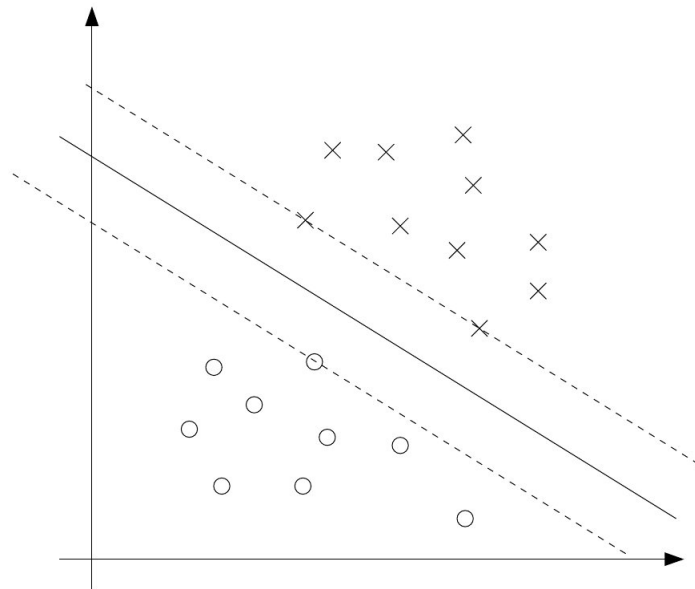
Τα SVMs [32] έχουν σαν βάση τους, στο κομμάτι της κατηγοριοποίησης την έννοια του *επίπεδου διαχωρισμού*. Με δεδομένες τις τιμές των προτύπων εκπαίδευσης για κάθε κατηγορία, το πρόβλημα που θα λυθεί είναι η εύρεση του βέλτιστου επιπέδου (οπτικά μπορεί να θεωρηθεί ως μία επιφάνεια gauss) που θα συμπεριλαμβάνει τα περισσότερα δυνατά στοιχεία της μίας κατηγορίας και το συμπληρωματικό σύνολο τα υπόλοιπα.

Στην εικόνα 5.1 παρατηρούμε ότι το σημείο A απέχει αρκετά από το επίπεδο διαχωρισμού, πολύ περισσότερο από ότι απέχει το σημείο C. Αυτό μας δίνει μια πολύ μεγαλύτερη βεβαιότητα για την κατηγορία του σημείου A και μεγάλη αβεβαιότητα για το σημείο C όπου και συνορεύει με το επίπεδο διαχωρισμού. Προφανώς κάτι ενδιάμεσο ισχύει για το σημείο B.

Στην εικόνα 5.2 με την συνεχή γραμμή φαίνεται το *επίπεδο μέγιστου ορίου διαχωρισμού*. Οι διακεκομμένες γραμμές ορίζουν τα όρια απόφασης ως προς την κατηγοριοποίηση των δεδομένων. Τα σημεία που εφάπτονται στα όρια απόφασης ορίζονται ως *support vectors*.



Εικόνα 5.1: SVM και το επίπεδο διαχωρισμού που ανάγεται σε ευθεία όταν έχουμε μία διάσταση.



Εικόνα 5.2: SVM και τα όρια γύρω από το επίπεδο διαχωρισμού.

Αυτή η έννοια μπορεί να επεκταθεί σε n -διαστάσεις και να προκύψει δια μέσω διαφορετικών πυρήνων (kernels) όπου και μετασχηματίζονται οι κατηγορίες σε κάποιο άλλο πεδίο/χώρο. Με χρήση διαφορετικών πυρήνων μετασχηματισμού έχουμε και διαφορετικά αποτελέσματα μετά τη διαδικασία εκπαίδευσης του μοντέλου. Για κάθε κατηγορία κυματομορφών λοιπόν θα έχουμε και τον κατάλληλο πυρήνα μετασχηματισμού που θα δίνει τα μεγαλύτερα ποσοστά αληθούς κατηγοριοποίησης.

5.2 Πυρήνες του μοντέλου

Η χρήση των SVMs γίνεται όταν υπάρχουν δεδομένα που περιέχουν ακριβώς δύο διαφορετικές κατηγορίες δεδομένων. Το μοντέλο κατηγοριοποιεί τα δεδομένα όπως περιγράφηκε παραπάνω, ψάχνοντας το καλύτερο (υπέρ-)επίπεδο διαχωρισμού, αυτό το επίπεδο που (βρίσκεται ανάμεσα και) έχει την μεγαλύτερη απόσταση από τις δύο κατηγορίες δεδομένων.

Τα δεδομένα εκπαίδευσης είναι ένα σύνολο σημείων (διανυσμάτων) x_i μαζί με ένα σύνολο κατηγοριών για το καθένα y_i . Για κάποια διάσταση d , τα $x_i \in \mathcal{R}^d$ και τα $y_i = \pm 1$. Η εξίσωση του επιπέδου διαχωρισμού είναι η

$$\langle w, x \rangle + b = 0 \quad (5.1)$$

όπου $w \in \mathcal{R}^d$, $b \in \mathcal{R}$, το $\langle w, x \rangle$ είναι το εσωτερικό γινόμενο των w και x .

Για να έχουμε το καλύτερο επίπεδο διαχωρισμού πρέπει να βρούμε τα w και b που ελαχιστοποιούν το μέτρο $\|w\|$ έτσι ώστε για όλα τα σημεία των δεδομένων (x_i, y_i) να ισχύει

$$y_i(\langle w, x_i \rangle + b) \geq 1 \quad (5.2)$$

τα διανύσματα υποστήριξης (support vectors) είναι τα x_i που βρίσκονται στο όριο, εκεί όπου ισχύει $y_i(\langle w, x_i \rangle + b) = 1$

5.2.1 Μη γραμμικές συναρτήσεις πυρήνα

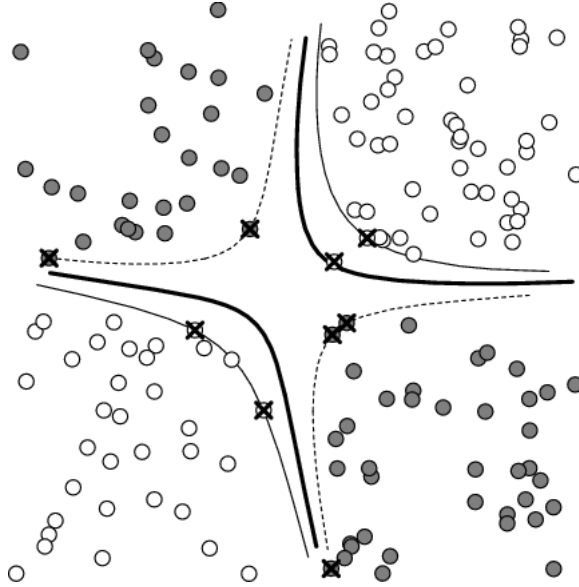
Κάποια προβλήματα κατηγοριοποίησης δεν έχουν ένα απλό επίπεδο που θα χρησιμοποιηθεί ως κριτήριο διαχωρισμού. Για αυτά τα προβλήματα χρησιμοποιείται μια εναλλακτική μαθηματική προσέγγιση που κρατάει σχεδόν ανέπαφη την απλότητα του επιπέδου διαχωρισμού. Ο *μη γραμμικός μετασχηματισμός* μέσω πυρήνα υποθέτει ότι υπάρχει μια κλάση συναρτήσεων $K(x, y)$ με την ακόλουθη ιδιότητα.

Υπάρχει ένας γραμμικός χώρος S και μία συνάρτηση ϕ που απεικονίζει το x στο S έτσι ώστε

$$K(x, y) = \langle \phi(x), \phi(y) \rangle \quad (5.3)$$

- Ο Πολυωνυμικός Πυρήνας (polynomial kernel), για κάποιο θετικό ακέραιο αριθμό d ,

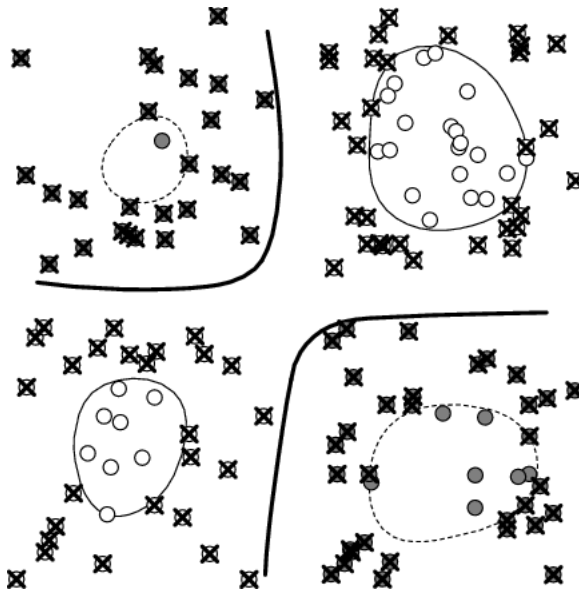
$$K(x, y) = (1 + \langle x, y \rangle)^d \quad (5.4)$$



Εικόνα 5.3: Πολυωνυμικός πυρήνας μετασχηματισμού SVM σε μία διάσταση. [Πηγή, *Machine Learning and Interactive Systems* (<http://malis.metz.supelec.fr/>)].

- Ο Πυρήνας Ακτινικής Βάσης (radial basis kernel), για κάποιο θετικό αριθμό σ ,

$$K(x, y) = \exp\left(-\frac{\langle (x - y), (x - y) \rangle}{2\sigma^2}\right) \quad (5.5)$$



Εικόνα 5.4: Πυρήνας Ακτινικής Βάσης μετασχηματισμού SVM σε μία διάσταση. [Πηγή, *Machine Learning and Interactive Systems* (<http://malis.metz.supelec.fr/>)].

- Ο Πολυεπίπεδος Νευρωνικός Πυρήνας (*multilayer perceptron kernel*), για κάποιο θετικό αριθμό p_1 και κάποιο αρνητικό αριθμό p_2 ,

$$K(x,y) = \tanh(p_1 \langle x,y \rangle + p_2) \quad (5.6)$$

Οι επιλογές για τον αλγόριθμο εύρεσης του επιπέδου διαχωρισμού μέσω του πολυεπίπεδου νευρωνικού πυρήνα είναι οι ακόλουθες.

1. Τετραγωνικός προγραμματισμός ('QP' - Quadratic programming).
2. Σειριακή ελάχιστη βελτιστοποίηση ('SMO' - Sequential Minimal Optimization).
3. Ελάχιστα τετράγωνα ('LS' - Least squares).

Μέρος ΙΙΙ

Δεδομένα και μεθοδολογία

Κεφάλαιο 6

Πειραματικά δεδομένα

Σε αυτό το κεφάλαιο γίνεται μια συνοπτική παρουσίαση των πειραματικών δεδομένων, με τη βοήθεια διαγραμμάτων αλλά και των περιγραφών για το κάθε είδος κυματομορφής. Τα δεδομένα προέκυψαν ύστερα από εξέταση 96 ηλικιωμένων ασθενών (ηλικίας 50-90 ετών) με ταυτοποιημένη καρωτιδική αθηρωμάτωση (διαγνωσμένη στένωση καρωτίδας >50%), οι οποίοι είχαν παραπεμφθεί στο Πανεπιστημιακό Γενικό Νοσοκομείο ΑΤΤΙΚΟΝ για εξέταση υπερήχου.

6.1 Διαδικασία εξαγωγής των κυματομορφών

Από τους παραπάνω ασθενείς σχηματίστηκαν δύο ομάδες. Στην ομάδα που θα ονομάσουμε, *συμπτωματική*, υπάρχουν ασθενείς που είχαν ένα κάποιο *ισχαιμικό εγκεφαλικό επεισόδιο* και στην ομάδα που θα ονομάσουμε, *ασυμπτωματική*, υπάρχουν οι ασθενείς που δεν είχαν νευρολογικά συμπτώματα, σε μια περίοδο 6 μηνών από την ημέρα της εξέτασης.

Τα απεικονιστικά δεδομένα πιστοποιήθηκαν από την επιτροπή του νοσοκομειακού ιδρύματος και όλοι οι ασθενείς έδωσαν εγγράφως τη συγκατάθεση τους για την επιστημονική χρήση των δεδομένων που προέκυψαν. Για κάθε δείγμα, η καρωτίδα αρτηρία σαρώθηκε στην διαμήκη κατεύθυνση σύμφωνα με τυποποιημένο πρωτόκολλο (dynamic range, 60 dB; persistence, low) και καταγράφηκε μια ακολουθία εικόνων Β-σάρωσης (B-mode) με ρυθμό υψηλότερο από 25καρέ/δευτ. για τουλάχιστον 3 δευτερόλεπτα (2-3 καρδιακοί κύκλοι). Η δυναμική Β-σάρωση στα διαμήκη τμήματα του αρτηριακού τοιχώματος επιτρέπει την εκτίμηση της κίνησης του ιστού σε δύο κατευθύνσεις, την *διαμήκη* αλλά και την *ακτινική*.

Για κάθε κλινική περίπτωση, ένας έμπειρος αγγειολόγος εντόπισε 4 περιοχές ενδιαφέροντος (ROI, region of interest), το πρόσθιο και οπίσθιο τοίχωμα του αγγείου και τις άνω/κάτω επιφάνειες της αθηρωματικής πλάκας. Τα τοιχώματα αναφέρονται σε υγιή μέρη του αρτηριακού τοιχώματος γειτονικά της αθηρωματικής πλάκας. Συμπεριλήφθηκαν στην ανάλυση κίνησης διότι πρόσφατες μελέτες έδειξαν ότι τα υγιή μέρη του αρτηριακού τοιχώματος πλησίον της αθηρωματικής πλάκας είναι πιθανοί δείκτες επικινδυνότητας για καρωτιδική αθηρωμάτωση [33].

Όλα τα εικονοστοιχεία που συνθέτουν τις 4 περιοχές ενδιαφέροντος, αλλά και την περιοχή όλης της πλάκας (περιοχή που περικλείεται από τις άνω/κάτω επιφάνειες που προαναφέρθηκαν), επιλέχθηκαν ως στόχοι για την ανάλυση κίνησης. Οι εγκάρσιες και διαμήκης θέσεις των στόχων σε σχέση με τον χρόνο εκτιμήθηκαν με χρήση ABM_{KF-K2} (*Area Based Matching*), έναν εκτιμητή κίνησης με αυξημένη ακρίβεια στην ανίχνευση κίνησης του αρτηριακού τοιχώματος [33, 34].

Από τις παραγόμενες κυματομορφές προέκυψε ένα σύνολο από 146 κατηγορίες κυματομορφών για την κινηματικές και μηχανικές ιδιότητες των περιοχών ενδιαφέροντος [33--35]. Ειδικότερα, παρήχθησαν κινηματικές κυματομορφές που αντιπροσωπεύουν *ταχύτητες* και *μετατοπίσεις* για

1. Ολόκληρη την πλάκα.
2. Για την άνω επιφάνεια της πλάκας (PTS - *Plaque Top Surface*).
3. Για την κάτω επιφάνεια της πλάκας (PBS - *Plaque Bottom Surface*).
4. Για το πρόσθιο τοίχωμα του αυλού.
5. Για το οπίσθιο τοίχωμα του αυλού.

όπως και δείκτες παραμόρφωσης, με χρήση κυματομορφών κίνησης για ζεύγη εικονοστοιχείων μιας ή δύο περιοχών ενδιαφέροντος για την *σχετική κίνηση* μεταξύ

1. Άνω και κάτω επιφανειών της πλάκας.
2. Της άνω επιφάνειας της πλάκας και των πρόσθιων/οπίσθιων τοιχωμάτων του αυλού (εάν η πλάκα βρίσκεται στο πρόσθιο ή στο οπίσθιο τοίχωμα αντίστοιχα).
3. Της κάτω επιφάνειας της πλάκας και των πρόσθιων/οπίσθιων τοιχωμάτων του αυλού (εάν η πλάκα βρίσκεται στο πρόσθιο ή στο οπίσθιο τοίχωμα αντίστοιχα).
4. Του πρόσθιου και οπίσθιου τοιχώματος του αυλού.

σε κάθε περίπτωση η τελικές κινηματικές/μηχανικές κυματομορφές για τον κάθε ασθενή αντιστοιχούν στις μέσες και στις διάμεσες κυματομορφές για τον εκάστοτε στόχο (ή στόχους, εάν έχουμε σχετικές κινήσεις/τάσεις). Οι παραμορφώσεις που υπολογίστηκαν

1. Ακτινική παραμόρφωση (RSI - *Radial strain*)

$$\max_t \left| \frac{(rP_{1(t)} - rP_{1(ED)}) - (rP_{2(t)} - rP_{2(ED)})}{rP_{1(ED)} - rP_{2(ED)}} \right| \text{ όπου } p_i = \text{εικονοστοιχείο}(i), r = \text{ακτινική}, ED = \text{τέλος διαστολής}, t = \text{χρόνος}$$

2. Διαμήκης παραμόρφωση (LSI - *Longitudinal strain*)

$$\max_t \left| \frac{(lP_{1(t)} - lP_{1(ED)}) - (lP_{2(t)} - lP_{2(ED)})}{lP_{1(ED)} - lP_{2(ED)}} \right| \text{ όπου } l = \text{διαμήκης}$$

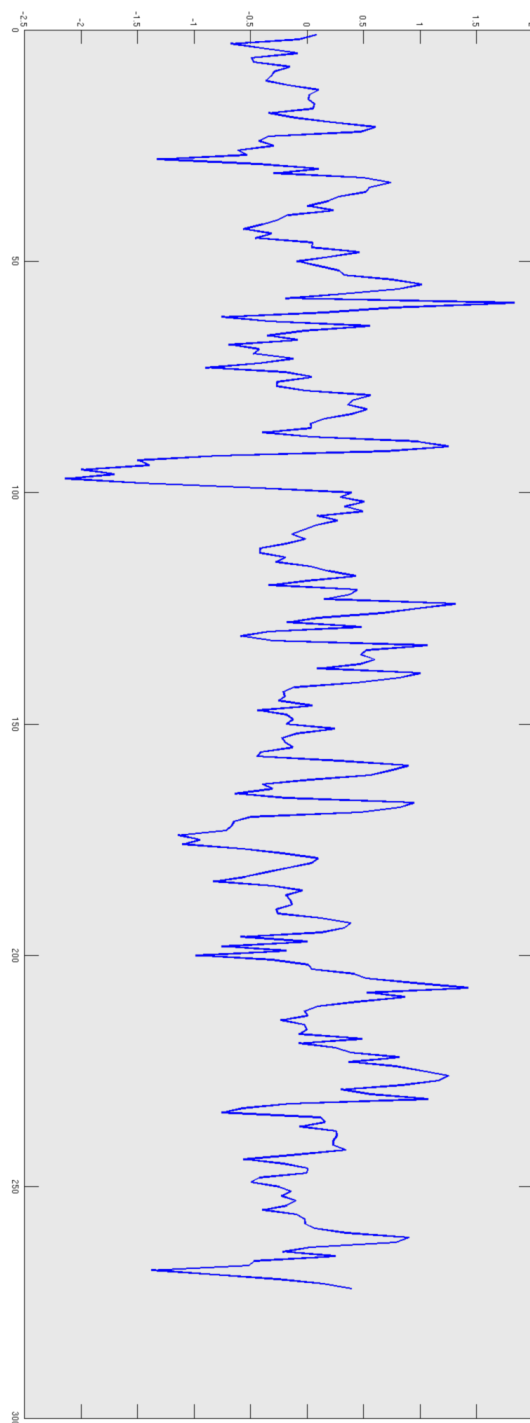
3. Ακτινική διατμητική παραμόρφωση (RSSI - *Radial shear strain*)

$$\max_t \left| \arctan \left(\frac{(rP_{1(t)} - rP_{1(ED)}) - (rP_{2(t)} - rP_{2(ED)})}{lP_{1(ED)} - lP_{2(ED)}} \right) \right|$$

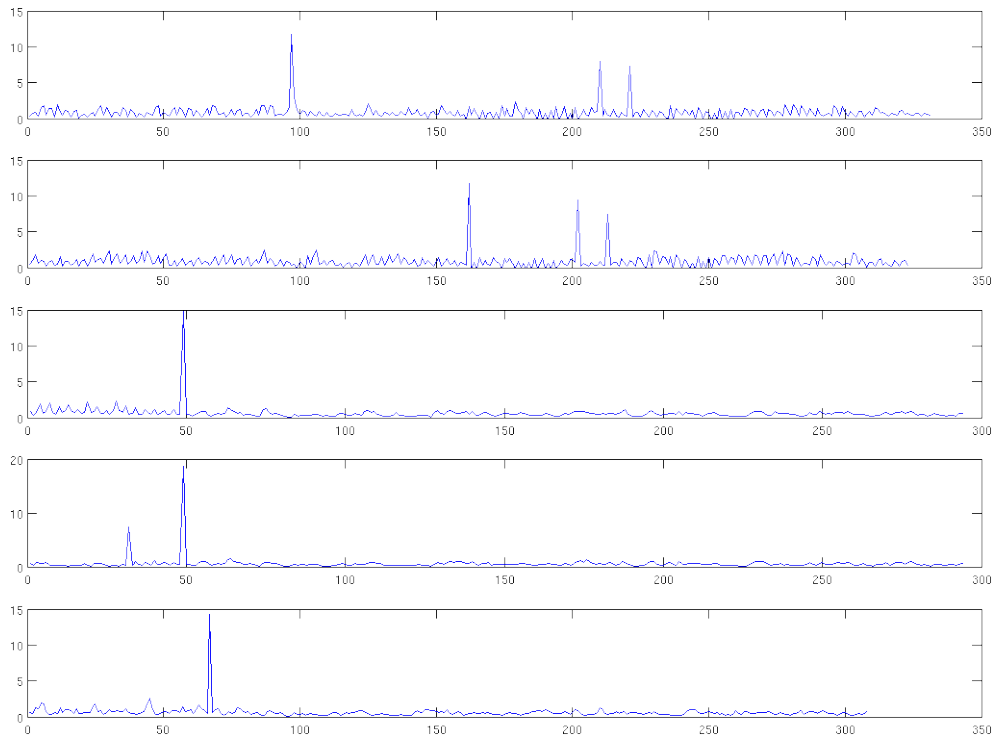
4. Διαμήκης διατμητική παραμόρφωση (LSSI - *Longitudinal shear strain*)

$$\max_t \left| \arctan \left(\frac{(lP_{1(t)} - lP_{1(ED)}) - (lP_{2(t)} - lP_{2(ED)})}{rP_{1(ED)} - rP_{2(ED)}} \right) \right|$$

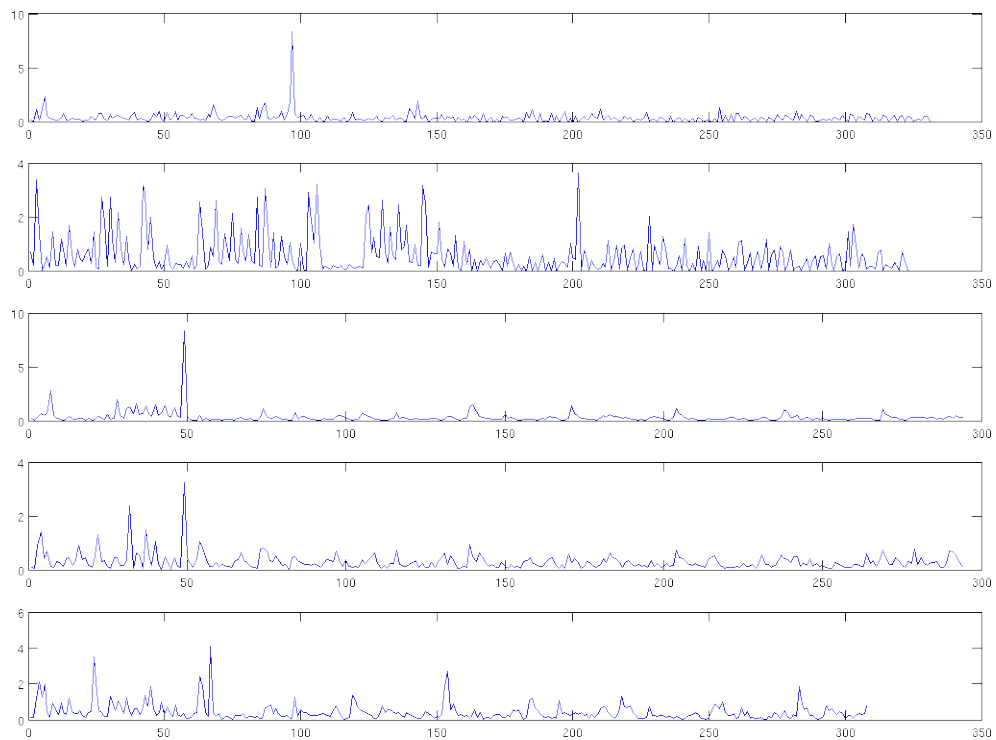
6.2 Μορφή εξαγόμενων σημάτων



Εικόνα 6.1: Τυχαία κυματομορφή μέσα από τα πειραματικά δεδομένα.



Εικόνα 6.2: 5 ασθενείς της ίδιας ομάδας για την κατηγορία κυματομορφών (μέση γωνία ταχύτητας, ολόκληρη πλάκα).



Εικόνα 6.3: 5 ασθενείς της ίδιας ομάδας για την κατηγορία κυματομορφών (διάμεση ολική ταχύτητα, οπίσθιο τοίχωμα).

Οι κυματομορφές και για τις 146 κατηγορίες είναι των παραπάνω μορφών, με αποτέλεσμα να καθίσταται αδύνατη η οπτική διάκριση μεταξύ των συμπτωματικών/ασυμπτωματικών ομάδων ασθενών, διότι τα χαρακτηριστικά αυτά διαφέρουν από κυματομορφή σε κυματομορφή. Χαρακτηριστικό παράδειγμα είναι οι δύο παραπάνω εικόνες 6.2 6.3 που για τις ίδιες κατηγορίες κυματομορφών και με ίδια ομάδα ασθενών, εξάγουμε διαφορετικές κυματομορφές για κάθε ασθενή. Ιδανικά θα θέλαμε μια αναπαράσταση όμοια του καρδιογραφήματος όπου για παράδειγμα το σύμπλεγμα QRS είναι ευδιάκριτο και μάλιστα μπορεί να αξιολογηθεί από καρδιολόγο άμεσα.

Στην περίπτωση των κατηγοριών των κυματομορφών που μας απασχολούν έχουμε ανάλογη μορφή με αυτή του ηλεκτροεγκεφαλογραφήματος. Η αναλογία έγκειται στην αδυναμία κατηγοριοποίησης του κάθε σήματος άμεσα με οπτική επιθεώρηση.

6.3 Κατηγορίες κυματομορφών

Οι 146 κατηγορίες κυματομορφών (χαρακτηριστικά κίνησης/παραμόρφωσης) που προέκυψαν είναι οι ακόλουθες και η διαδικασία φαίνεται στην εικόνα 6.4.

1. Ολόκληρη πλάκα [μέση] διαμήκης ταχύτητα
2. Ολόκληρη πλάκα [μέση] ακτινική ταχύτητα
3. Ολόκληρη πλάκα [μέση] απόλυτη διαμήκης ταχύτητα
4. Ολόκληρη πλάκα [μέση] απόλυτη ακτινική ταχύτητα
5. Ολόκληρη πλάκα [μέση] ολική ταχύτητα
6. Ολόκληρη πλάκα [μέση] γωνία ταχύτητας
7. Ολόκληρη πλάκα [διάμεση] διαμήκης ταχύτητα
8. Ολόκληρη πλάκα [διάμεση] ακτινική ταχύτητα
9. Ολόκληρη πλάκα [διάμεση] απόλυτη διαμήκης ταχύτητα
10. Ολόκληρη πλάκα [διάμεση] απόλυτη ακτινική ταχύτητα
11. Ολόκληρη πλάκα [διάμεση] ολική ταχύτητα
12. Ολόκληρη πλάκα [διάμεση] γωνία ταχύτητας
13. Ολόκληρη πλάκα [μέση] διαμήκης μετατόπιση
14. Ολόκληρη πλάκα [μέση] ακτινική μετατόπιση
15. Ολόκληρη πλάκα [μέση] απόλυτη διαμήκης μετατόπιση

16. Ολόκληρη πλάκα [μέση] απόλυτη ακτινική μετατόπιση
17. Ολόκληρη πλάκα [μέση] ολική μετατόπιση
18. Ολόκληρη πλάκα [μέση] γωνία μετατόπισης
19. Ολόκληρη πλάκα [διάμεση] διαμήκης μετατόπιση
20. Ολόκληρη πλάκα [διάμεση] ακτινική μετατόπιση
21. Ολόκληρη πλάκα [διάμεση] απόλυτη διαμήκης μετατόπιση
22. Ολόκληρη πλάκα [διάμεση] απόλυτη ακτινική μετατόπιση
23. Ολόκληρη πλάκα [διάμεση] ολική μετατόπιση
24. Ολόκληρη πλάκα [διάμεση] γωνία μετατόπισης
25. Σχετική κίνηση μέσα στην πλάκα [μέση] διαμήκης παραμόρφωση
26. Σχετική κίνηση μέσα στην πλάκα [διάμεση] διαμήκης παραμόρφωση
27. Σχετική κίνηση μέσα στην πλάκα [μέση] ακτινική παραμόρφωση
28. Σχετική κίνηση μέσα στην πλάκα [διάμεση] ακτινική παραμόρφωση
29. Σχετική κίνηση μέσα στην πλάκα [μέση] ακτινική διατμητική παραμόρφωση
30. Σχετική κίνηση μέσα στην πλάκα [διάμεση] ακτινική διατμητική παραμόρφωση
31. Σχετική κίνηση μέσα στην πλάκα [μέση] διαμήκης διατμητική παραμόρφωση
32. Σχετική κίνηση μέσα στην πλάκα [διάμεση] διαμήκης διατμητική παραμόρφωση
33. Άνω επιφάνεια πλάκας (PTS) [μέση] διαμήκης ταχύτητα
34. Άνω επιφάνεια πλάκας (PTS) [μέση] ακτινική ταχύτητα
35. Άνω επιφάνεια πλάκας (PTS) [μέση] απόλυτη διαμήκης ταχύτητα
36. Άνω επιφάνεια πλάκας (PTS) [μέση] απόλυτη ακτινική ταχύτητα
37. Άνω επιφάνεια πλάκας (PTS) [μέση] ολική ταχύτητα
38. Άνω επιφάνεια πλάκας (PTS) [μέση] γωνία ταχύτητας
39. Άνω επιφάνεια πλάκας (PTS) [διάμεση] διαμήκης ταχύτητα
40. Άνω επιφάνεια πλάκας (PTS) [διάμεση] ακτινική ταχύτητα
41. Άνω επιφάνεια πλάκας (PTS) [διάμεση] απόλυτη διαμήκης ταχύτητα

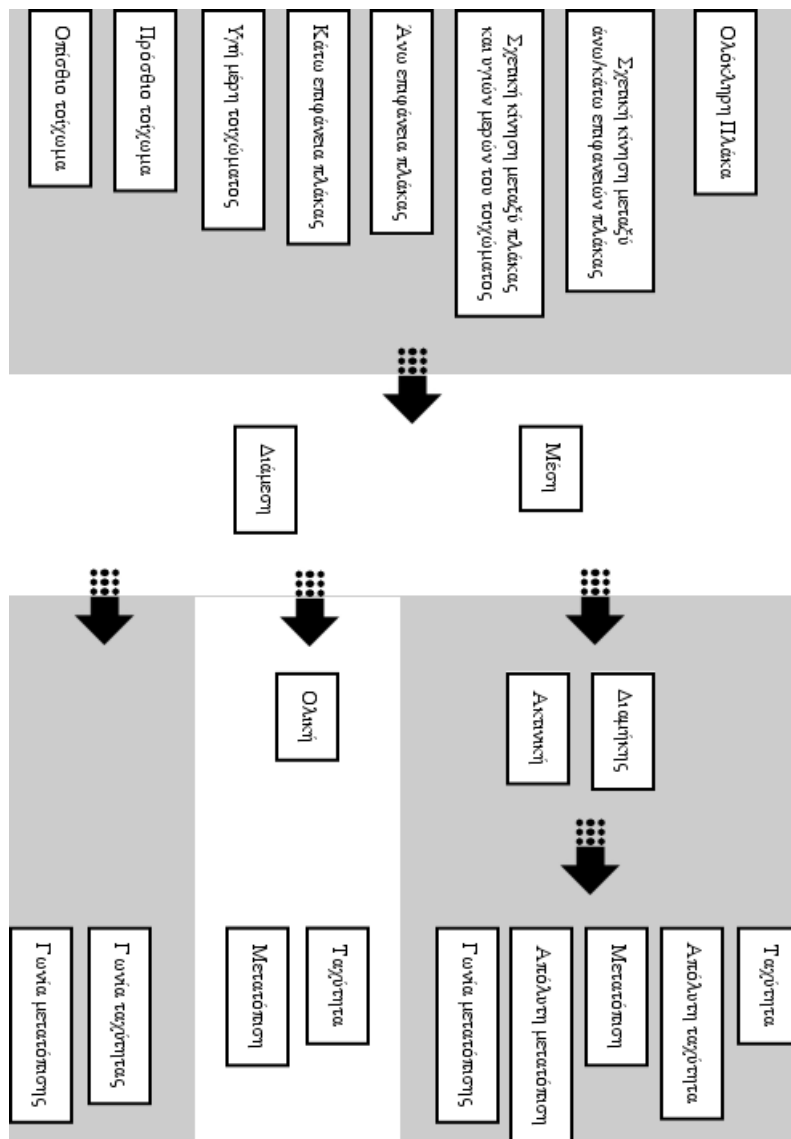
42. Άνω επιφάνεια πλάκας (PTS) [διάμεση] απόλυτη ακτινική ταχύτητα
43. Άνω επιφάνεια πλάκας (PTS) [διάμεση] ολική ταχύτητα
44. Άνω επιφάνεια πλάκας (PTS) [διάμεση] γωνία ταχύτητας
45. Άνω επιφάνεια πλάκας (PTS) [μέση] διαμήκης μετατόπιση
46. Άνω επιφάνεια πλάκας (PTS) [μέση] ακτινική μετατόπιση
47. Άνω επιφάνεια πλάκας (PTS) [μέση] απόλυτη διαμήκης μετατόπιση
48. Άνω επιφάνεια πλάκας (PTS) [μέση] απόλυτη ακτινική μετατόπιση
49. Άνω επιφάνεια πλάκας (PTS) [μέση] ολική μετατόπιση
50. Άνω επιφάνεια πλάκας (PTS) [μέση] γωνία μετατόπισης
51. Άνω επιφάνεια πλάκας (PTS) [διάμεση] διαμήκης μετατόπιση
52. Άνω επιφάνεια πλάκας (PTS) [διάμεση] ακτινική μετατόπιση
53. Άνω επιφάνεια πλάκας (PTS) [διάμεση] απόλυτη διαμήκης μετατόπιση
54. Άνω επιφάνεια πλάκας (PTS) [διάμεση] απόλυτη ακτινική μετατόπιση
55. Άνω επιφάνεια πλάκας (PTS) [διάμεση] ολική μετατόπιση
56. Άνω επιφάνεια πλάκας (PTS) [διάμεση] γωνία μετατόπισης
57. Κάτω επιφάνεια πλάκας (PBS) [μέση] διαμήκης ταχύτητα
58. Κάτω επιφάνεια πλάκας (PBS) [μέση] ακτινική ταχύτητα
59. Κάτω επιφάνεια πλάκας (PBS) [μέση] απόλυτη διαμήκης ταχύτητα
60. Κάτω επιφάνεια πλάκας (PBS) [μέση] απόλυτη ακτινική ταχύτητα
61. Κάτω επιφάνεια πλάκας (PBS) [μέση] ολική ταχύτητα
62. Κάτω επιφάνεια πλάκας (PBS) [μέση] γωνία ταχύτητας
63. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] διαμήκης ταχύτητα
64. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] ακτινική ταχύτητα
65. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] απόλυτη διαμήκης ταχύτητα
66. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] απόλυτη ακτινική ταχύτητα
67. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] ολική ταχύτητα

68. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] γωνία ταχύτητας
69. Κάτω επιφάνεια πλάκας (PBS) [μέση] διαμήκης μετατόπιση
70. Κάτω επιφάνεια πλάκας (PBS) [μέση] ακτινική μετατόπιση
71. Κάτω επιφάνεια πλάκας (PBS) [μέση] απόλυτη διαμήκης μετατόπιση
72. Κάτω επιφάνεια πλάκας (PBS) [μέση] απόλυτη ακτινική μετατόπιση
73. Κάτω επιφάνεια πλάκας (PBS) [μέση] ολική μετατόπιση
74. Κάτω επιφάνεια πλάκας (PBS) [μέση] γωνία μετατόπισης
75. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] διαμήκης μετατόπιση
76. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] ακτινική μετατόπιση
77. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] απόλυτη διαμήκης μετατόπιση
78. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] απόλυτη ακτινική μετατόπιση
79. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] ολική μετατόπιση
80. Κάτω επιφάνεια πλάκας (PBS) [διάμεση] γωνία μετατόπισης
81. Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [μέση] διαμήκης παραμόρφωση PTS/Υγιής
82. Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [διάμεση] διαμήκης παραμόρφωση PTS/Υγιής
83. Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [μέση] διαμήκης παραμόρφωση PBS/Υγιής
84. Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [διάμεση] διαμήκης παραμόρφωση PBS/Υγιής
85. Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [μέση] ακτινική παραμόρφωση PTS/Υγιής
86. Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [διάμεση] ακτινική παραμόρφωση PTS/Υγιής
87. Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [μέση] ακτινική παραμόρφωση PBS/Υγιής
88. Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [διάμεση] ακτινική παραμόρφωση PBS/Υγιής
89. Μέρη του υγιούς τοιχώματος [μέση] διαμήκης παραμόρφωση, οπίσθιο τοίχωμα
90. Μέρη του υγιούς τοιχώματος [διάμεση] διαμήκης παραμόρφωση, οπίσθιο τοίχωμα
91. Μέρη του υγιούς τοιχώματος [μέση] διαμήκης παραμόρφωση, πρόσθιο τοίχωμα
92. Μέρη του υγιούς τοιχώματος [διάμεση] διαμήκης παραμόρφωση, πρόσθιο τοίχωμα
93. Μέρη του υγιούς τοιχώματος [μέση] ακτινική παραμόρφωση οπίσθιο/πρόσθιο τοίχωμα

94. Μέρη του υγιούς τοιχώματος [διάμεση] ακτινική παραμόρφωση οπίσθιο/πρόσθιο τοίχωμα
95. Μέρη του υγιούς τοιχώματος [μέση] διαμήκης διατμητική παραμόρφωση, οπίσθιο τοίχ.
96. Μέρη του υγιούς τοιχώματος [διάμεση] διαμήκης διατμητική παραμόρφωση, οπίσθιο τοίχ.
97. Μέρη του υγιούς τοιχώματος [μέση] διαμήκης διατμητική παραμόρφωση, πρόσθιο τοίχ.
98. Μέρη του υγιούς τοιχώματος [διάμεση] διαμήκης διατμητική παραμόρφωση, πρόσθιο τοίχ.
99. Οπίσθιο τοίχωμα [μέση] διαμήκης ταχύτητα
100. Οπίσθιο τοίχωμα [μέση] ακτινική ταχύτητα
101. Οπίσθιο τοίχωμα [μέση] απόλυτη διαμήκης ταχύτητα
102. Οπίσθιο τοίχωμα [μέση] απόλυτη ακτινική ταχύτητα
103. Οπίσθιο τοίχωμα [μέση] ολική ταχύτητα
104. Οπίσθιο τοίχωμα [μέση] γωνία ταχύτητας
105. Οπίσθιο τοίχωμα [διάμεση] διαμήκης ταχύτητα
106. Οπίσθιο τοίχωμα [διάμεση] ακτινική ταχύτητα
107. Οπίσθιο τοίχωμα [διάμεση] απόλυτη διαμήκης ταχύτητα
108. Οπίσθιο τοίχωμα [διάμεση] απόλυτη ακτινική ταχύτητα
109. Οπίσθιο τοίχωμα [διάμεση] ολική ταχύτητα
110. Οπίσθιο τοίχωμα [διάμεση] γωνία ταχύτητας
111. Οπίσθιο τοίχωμα [μέση] διαμήκης μετατόπιση
112. Οπίσθιο τοίχωμα [μέση] ακτινική μετατόπιση
113. Οπίσθιο τοίχωμα [μέση] απόλυτη διαμήκης μετατόπιση
114. Οπίσθιο τοίχωμα [μέση] απόλυτη ακτινική μετατόπιση
115. Οπίσθιο τοίχωμα [μέση] ολική μετατόπιση
116. Οπίσθιο τοίχωμα [μέση] γωνία μετατόπισης
117. Οπίσθιο τοίχωμα [διάμεση] διαμήκης μετατόπιση
118. Οπίσθιο τοίχωμα [διάμεση] ακτινική μετατόπιση
119. Οπίσθιο τοίχωμα [διάμεση] απόλυτη διαμήκης μετατόπιση

120. Οπίσθιο τοίχωμα [διάμεση] απόλυτη ακτινική μετατόπιση
121. Οπίσθιο τοίχωμα [διάμεση] ολική μετατόπιση
122. Οπίσθιο τοίχωμα [διάμεση] γωνία μετατόπισης
123. Πρόσθιο τοίχωμα [μέση] διαμήκης ταχύτητα
124. Πρόσθιο τοίχωμα [μέση] ακτινική ταχύτητα
125. Πρόσθιο τοίχωμα [μέση] απόλυτη διαμήκης ταχύτητα
126. Πρόσθιο τοίχωμα [μέση] απόλυτη ακτινική ταχύτητα
127. Πρόσθιο τοίχωμα [μέση] ολική ταχύτητα
128. Πρόσθιο τοίχωμα [μέση] γωνία ταχύτητας
129. Πρόσθιο τοίχωμα [διάμεση] διαμήκης ταχύτητα
130. Πρόσθιο τοίχωμα [διάμεση] ακτινική ταχύτητα
131. Πρόσθιο τοίχωμα [διάμεση] απόλυτη διαμήκης ταχύτητα
132. Πρόσθιο τοίχωμα [διάμεση] απόλυτη ακτινική ταχύτητα
133. Πρόσθιο τοίχωμα [διάμεση] ολική ταχύτητα
134. Πρόσθιο τοίχωμα [διάμεση] γωνία ταχύτητας
135. Πρόσθιο τοίχωμα [μέση] διαμήκης μετατόπιση
136. Πρόσθιο τοίχωμα [μέση] ακτινική μετατόπιση
137. Πρόσθιο τοίχωμα [μέση] απόλυτη διαμήκης μετατόπιση
138. Πρόσθιο τοίχωμα [μέση] απόλυτη ακτινική μετατόπιση
139. Πρόσθιο τοίχωμα [μέση] ολική μετατόπιση
140. Πρόσθιο τοίχωμα [μέση] γωνία μετατόπισης
141. Πρόσθιο τοίχωμα [διάμεση] διαμήκης μετατόπιση
142. Πρόσθιο τοίχωμα [διάμεση] ακτινική μετατόπιση
143. Πρόσθιο τοίχωμα [διάμεση] απόλυτη διαμήκης μετατόπιση
144. Πρόσθιο τοίχωμα [διάμεση] απόλυτη ακτινική μετατόπιση
145. Πρόσθιο τοίχωμα [διάμεση] ολική μετατόπιση

146. Πρόσθιο τοίχωμα [διάμεση] γωνία μετατόπισης



Εικόνα 6.4: Σχήμα κατασκευής των 146 κυματομορφών.

Κεφάλαιο 7

Ανάλογο του συστήματος αναγνώρισης φωνής για γενική αναγνώριση προτύπων

Με φυσικό τρόπο σύμφωνα με την παραπάνω δομική ανάλυση των Κρυφών Μαρκοβιανών Μοντέλων για την αναγνώριση λέξεων (και κατ' επέκταση συνεχούς λόγου) καταλήγουμε στο ανάλογο μοντέλο για την αναγνώριση γενικών προτύπων, για οποιαδήποτε κυματομορφή που παρουσιάζει χρονική εξάρτηση. Στο κεφάλαιο αυτό παρουσιάζεται μέσα από βήματα η διαδικασία κατασκευής ολόκληρου του *συστήματος κατηγοριοποίησης κυματομορφών*, από το στάδιο του λεξιλογίου μέχρι και το στάδιο αναγνώρισης.

7.1 Προ-επεξεργασία Δεδομένων

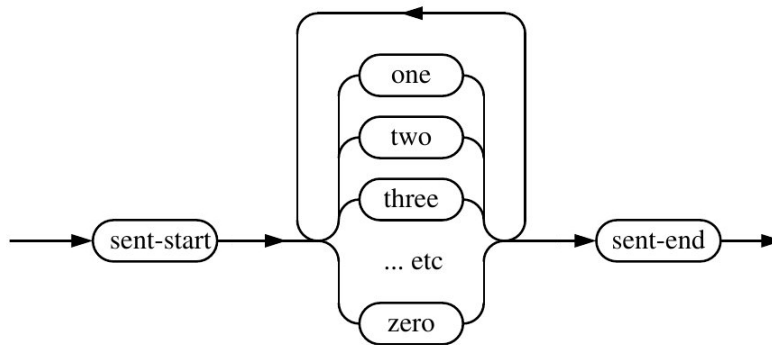
Το πρώτο στάδιο είναι η προετοιμασία των δεδομένων. Τα δεδομένα είναι οι 146 κυματομορφές για κάθε έναν από τους ασθενείς, όπου η *κάθε κυματομορφή θα ορίσει και μια κατηγορία*. Τα δεδομένα αυτά θα χρησιμοποιηθούν για την εκπαίδευση αλλά και για την αξιολόγηση της ακρίβειας του μοντέλου. Εάν μιλούμε για *δεδομένα εκπαίδευσης* (ή αλλιώς γνωστά δεδομένα στο μοντέλο) έχουμε την σύνδεση των κυματομορφών αυτών με το λεξιλόγιο (ετικέτες), έτσι κατηγοριοποιούμε τα δεδομένα με δυαδικό τρόπο, $1 - 0$, *συμπτωματικός - ασυμπτωματικός ασθενής*. Εάν μιλούμε για *δεδομένα αξιολόγησης* (ή αλλιώς άγνωστα δεδομένα στο μοντέλο) έχουμε πάλι την σύνδεση τους με το λεξιλόγιο για την κατηγοριοποίηση τους (*κατηγοριοποίηση όμως άγνωστη για το μοντέλο*) και στο στάδιο της αναγνώρισης βρίσκουμε τα ποσοστά επιτυχίας του μοντέλου για κάθε κυματομορφή.

7.1.1 Βήμα 1 - Η γραμματική

Στο μοντέλο μας, κομμάτι της γλώσσας πέρα από το λεξιλόγιο είναι και η γραμματική. Στο πακέτο HTK υπάρχει αυτή η δυνατότητα όπου βοηθά στη δημιουργία μοντέλων με *διακόπτες*. Λέξεις

κλειδιά δηλαδή που χρησιμεύουν για στην αλλαγή ροής κάποιας αυτοματοποιημένης διαδικασίας. Ένα από τα πιο γνωστά παραδείγματα είναι αυτό του αυτοματοποιημένου συστήματος κλήσης (εικόνα 7.1), που είναι μια σχετικά απλή επέκταση της γραμματικής που θα χρησιμοποιήσουμε εμείς.

```
$digit = ONE | TWO | THREE | FOUR | FIVE |
SIX | SEVEN | EIGHT | NINE | OH | ZERO;
( SENT-START <$digit> SENT-END )
```



Εικόνα 7.1: Γραμματική του μοντέλου για την κατηγοριοποίηση των κυματομορφών.

Οι αρχικές και τελικές καταστάσεις (send-start, send-end) είναι όπως παρουσιάστηκαν και σε προηγούμενο κεφάλαιο οι συνδετικοί κρίκοι μεταξύ διαφορετικών μοντέλων (λέξεων). Στην εικόνα φαίνεται καθαρά μια γραμματική που μπορεί να αναπαράγει κάθε αριθμητική ακολουθία 0-9, αυτό γίνεται δια μέσω ενός κλάδου ανάδρασης που επιτρέπει την επανάληψη των ψηφίων (το κάθε ψηφίο φυσικά είναι μια λέξη, σύνθεση φωνημάτων). Οι κανονικές εκφράσεις χωρίζονται από καθέτους "|" που είναι το λογικό OR δείχνοντας τις εναλλακτικές λέξεις και από "<", ">" όπου ορίζουμε την επαναληπτικότητα του αντικειμένου που εσωκλείεται.

7.1.2 Βήμα 2 - Το λεξικό

Το πρώτο βήμα για την κατασκευή ενός λεξιλογίου, είναι η δημιουργία μιας ταξινομημένης λίστας των απαραίτητων λέξεων. Στην περίπτωση μας το λεξιλόγιο χρησιμεύει για την δυαδική κατηγοριοποίηση, άρα και χρειαζόμαστε μονάχα 2 λέξεις. Εάν συμπεριλάβουμε παραπάνω από δύο λέξεις θα πρέπει να κατασκευάσουμε για το σύνολο των λέξεων και 2 υποσύνολα για την κατηγοριοποίηση των δεδομένων. Για ένα ισχυρό μοντέλο αναγνώρισης θα πρέπει να έχουμε και ένα μεγάλο αριθμό κυματομορφών (ακολουθιών) εκπαίδευσης για την κάθε λέξη, αυτό μας κάνει να περιορίσουμε τις λέξεις σε 2 και να αυξήσουμε στο μέγιστο την εκπαίδευση την δύο αυτών λέξεων μέσα από τους διαφορετικούς ασθενείς.

Η υπόθεση μας είναι ότι για κάθε είδος κυματομορφής υπάρχουν 2 λέξεις (ετικέτες) που μπορούν να ταυτιστούν. Η μία είναι για τον χαρακτηρισμό της κυματομορφής ως ασυμπτωματικής και η εναλλακτική για τον χαρακτηρισμό της ως συμπτωματική. Αυτό σημαίνει πως για το γνωστό μας

δείγμα ασθενών, θα συνδέσουμε αυτές τις κυματομορφές με την συγκεκριμένη λέξη (κατηγορία) εκ των προτέρων.

Η γενική μορφή των λεξικών στο πακέτο HTK έχει την εξής μορφή

```
WORD [outsym] p1 p2 p3 ....
```

αυτό σημαίνει πως η λέξη (WORD) προφέρεται ως η ακολουθία των φωνημάτων p1 p2 p3 Η συμβολοσειρά (string) μέσα στις αγκύλες ορίζει την συμβολοσειρά που θα εμφανιστεί ως έξοδος όταν αναγνωρίζεται αυτή η λέξη. Στην περίπτωση που μένει κενό τότε έξοδος ορίζεται αυτόματα η λέξη, ενώ στην περίπτωση που συμπεριλαμβάνεται αλλά είναι κενό ([]) τότε η έξοδος θα είναι κενή. Ένα παράδειγμα λεξιλογίου είναι το ακόλουθο

```
A                ah sp
A                ax sp
A                ey sp
CALL             k ao l sp
DIAL             d ay ax l sp
EIGHT           ey t sp
PHONE           f ow n sp
SENT-END        [ ] sil
SENT-START      [ ] sil
SEVEN           s eh v n sp
TO              t ax sp
TO              t uw sp
ZERO            z ia r ow sp
```

όπου παρατηρούμε διαφορετικές επιτρεπτές προφορές για τις ίδιες λέξεις, αλλά και μερικά καινούργια σύμβολα όπως το "sp" και "sil", που σημαίνουν *μικρή παύση* (sp - short pause) και *αφωνία* (sil - silence). Τα υπόλοιπα σύμβολα "s", "eh", "v", "n" κτλ, είναι μερικά από τα φωνήματα που σχηματίζουν την κάθε λέξη.

7.1.3 Βήμα 3 - Δημιουργία αρχείων ετικέτας

Για την εκπαίδευση του μοντέλου, κάθε *αρχείο εκπαίδευσης/πρότυπο* πρέπει να συσχετιστεί με κάποια ετικέτα. Αυτό γίνεται κατασκευάζοντας ένα αρχείο .mlf (Master Label File) το οποίο συσχετίζει/κατονομάζει κάθε αρχείο εκπαίδευσης με την χρήση λέξεων. Για παράδειγμα

```
#!MLF!#
"/cat2_file_147_1.lab"
pac
.
"/cat2_file_147_2.lab"
pac
```

```

.
"/cat1_file_147_3.lab"
one
.
"/cat1_file_147_4.lab"
one
.
.. ..

```

Αυτό δείχνει πως για το αρχείο *cat2_file_147_1.wav* υπάρχει η ετικέτα *pac* και ούτω καθεξής. Ο αστερίσκος πριν τη διαδρομή του αρχείου δείχνει την ανεξαρτησία του βάθους αναζήτησης για το αρχείο. Στη συνέχεια αυτές οι λέξεις θα μεταφραστούν σε *μορφή ετικέτας φωνημάτων* όπως για παράδειγμα

```

#!MLF!#
"/cat2_file_147_1.lab"
sil
p
ae
k
sil
.
"/cat2_file_147_2.lab"
sil
p
ae
k
sil
.. ..

```

όπου έχει προηγηθεί η εισαγωγή ενός *μοντέλου αφωνίας* (*silence model sil*) στην αρχή και στο τέλος κάθε φωνήματος και η διαγραφή όλων των *ετικετών μικρής παύσης* (*short-pause sp*).

7.1.4 Βήμα 4 - Κωδικοποίηση δεδομένων

Το τελικό στάδιο της προετοιμασίας των δεδομένων (εκπαίδευσης και αξιολόγησης) είναι η παραμετροποίηση τους από την πρωταρχική κατάσταση σε ακολουθίες *διανυσματικών παραμέτρων* (*feature vectors*). Η επιλογή των παραμέτρων βασίζεται είτε στην ανάλυση με *FFT* (*Fast Fourier transform*), είτε με *LPC* (*linear prediction coefficients*). Οι *συντελεστές Cepstrum στην Κλίμακα Συχνότητας Mel* (*MFCC*) που χρησιμοποιούνται προκύπτουν από την εξής διαδικασία

1. Μετασχηματισμός Fourier του σήματος.
2. Αντιστοίχιση της ισχύος του φάσματος που προέκυψε στην κλίμακα συχνότητας Mel, με χρήση τριγωνικών επικαλυπτόμενων παραθύρων.

3. Για όλες τις συχνότητες Mel δημιουργία μιας λίστας με τον λογάριθμο της ισχύος για κάθε μια από αυτές.
4. Διακριτός συνημιτονικός μετασχηματισμός της λίστας, σαν να ήταν ένα σήμα.
5. Τα πλάτη, του φάσματος που προκύπτει από το προηγούμενο βήμα, είναι οι συντελεστές MFCC.

Η διαδικασία της μετατροπής γίνεται μέσα από ένα αρχείο ρυθμίσεων από το οποίο θα επιλέξουμε τις παραμέτρους μετατροπής των κυματομορφών σε διανύσματα MFCC. Για παράδειγμα:

```
# Coding parameters
TARGETKIND = MFCC_0
TARGETRATE = 100000.0
SAVECOMPRESSED = T
SAVEWITHCRC = T
WINDOWSIZE = 250000.0
USEHAMMING = T
PREEMPCOEF = 0.97
NUMCHANS = 26
CEPLIFTER = 22
NUMCEPS = 12
ENORMALISE = F
```

Εν συντομία αυτό το αρχείο ορίζει πως κατά την μετατροπή θα καταλήξουμε σε MFCC διανύσματα όπου θα έχουν την μηδενική cepstral συνιστώσα C_0 , δειγματοληψία των 100ns, η έξοδος θα σωθεί σε συμπιεσμένη μορφή και θα προστεθεί το crc checksum. Ο μετασχηματισμός FFT θα έχει παράθυρο Hamming και το σήμα θα έχει προέμφαση πρώτης τάξης με έναν πολλαπλασιαστικό παράγοντα 0.97. Θα υπάρχουν 26 κανάλια στην τράπεζα φίλτρων (filterbank) και 12 MFCC συνιστώσες. Η επιλογή ENORMALISE είναι για την κανονικοποίηση της ενέργειας των αρχικών κυματομορφών. Άρα, τελικά θα υπάρξει μια μαζική μετατροπή των κυματομορφών μέσα από ένα script στην μορφή

```
training_wavs/cat2_file_2_1.wav mfcc_vect/training_wavs/cat2_file_2_1.mfc
training_wavs/cat2_file_2_2.wav mfcc_vect/training_wavs/cat2_file_2_2.mfc
training_wavs/cat1_file_2_3.wav mfcc_vect/training_wavs/cat1_file_2_3.mfc
...
...
```

που θα δώσει την *παραμετρική μορφή* των αρχικών κυματομορφών.

7.2 Δημιουργία Μονοφωνικών Μοντέλων

Σε αυτή την ενότητα, θα περιγραφεί η δημιουργία ισχυρών μονοφωνικών μοντέλων με μονομετάβλητη κατανομή Gauss. Το σημείο έναρξης θα είναι το σύνολο πανομοιότυπων μονοφωνικών

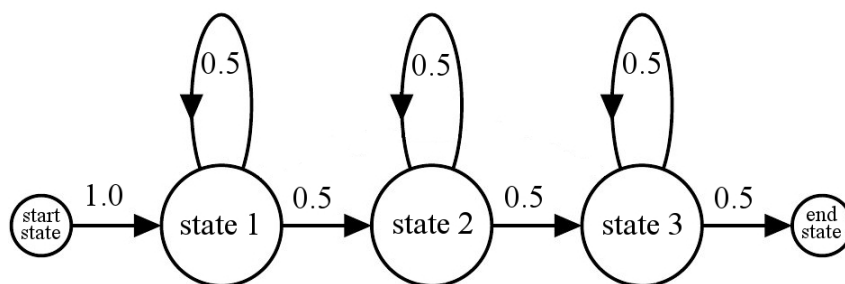
μοντέλων, όπου θα έχουν ίδιες μέσες τιμές και ίδιες διακυμάνσεις. Αυτά τα μοντέλα στη συνέχεια θα επανεκπαιδευτούν και θα προστεθούν μοντέλα μικρής παύσης ενδιάμεσα. Όταν υπάρξει η κατάλληλη επανεκπαίδευση μπορεί να γίνει μια *εθυγράμμιση των δεδομένων*, δηλαδή να δημιουργηθούν καινούργια αρχεία .mlf όπου η επιλογή των *προφορών* γίνεται σύμφωνα με την ακουστική ταύτιση. Με αυτά μπορεί να επανεκπαιδευτούν τα μοντέλα.

7.2.1 Βήμα 5 - Δημιουργία μονοφωνικών μοντέλων

Αρχικά ορίζουμε ένα πρωτότυπο μοντέλο, όπου θέλουμε να ορίσουμε την τοπολογία του. Για παράδειγμα όταν θέλουμε να δημιουργήσουμε ένα *μοντέλο 3-καταστάσεων από αριστερά στα δεξιά* 7.2, έχουμε

```
~o <VecSize> 39 <MFCC_0_D_A>
~h "proto"
<BeginHMM>
  <NumStates> 5
    <State> 2
      <Mean> 39
        0.0 0.0 0.0
      <Variance> 39 ...
        1.0 1.0 1.0 ...
    <State> 3
      <Mean> 39
        0.0 0.0 0.0 ...
      <Variance> 39
        1.0 1.0 1.0 ...
    <State> 4
      <Mean> 39
        0.0 0.0 0.0 ...
      <Variance> 39
        1.0 1.0 1.0 ...
  <TransP> 5
    0.0 1.0 0.0 0.0 0.0
    0.0 0.5 0.5 0.0 0.0
    0.0 0.0 0.5 0.5 0.0
    0.0 0.0 0.0 0.5 0.5
    0.0 0.0 0.0 0.0 0.0
<EndHMM>
```

με το κάθε διάνυσμα να έχει 39 στοιχεία. Αυτό προκύπτει ως άθροισμα των διανυσματικών παραμέτρων MFCC_0 (13 στοιχεία), MFCC_D (13 στοιχεία) και MFCC_A (13 στοιχεία). Όπου με _0 έχουμε την μηδενική cepstral συνιστώσα, με _D έχουμε τις delta συνιστώσες και με _A τις συνιστώσες επιτάχυνσης.



Εικόνα 7.2: Μοντέλο 3-καταστάσεων, αριστερά προς δεξιά, με τις πιθανότητες μετάβασης που ορίστηκαν με μορφή πίνακα παραπάνω.

Στη συνέχεια θα οριστούν εκ νέου οι μέσες τιμές και διακυμάνσεις, από το σύνολο των κυματομορφών που εκπαιδεύουν το μοντέλο. Έτσι όλες οι κατανομές Gauss του μοντέλου θα έχουν τις τιμές του δείγματος και όχι τις flat-start τιμές που είχαν πριν. Άρα θα έχει δημιουργηθεί τελικά ένα καινούργιο πρότυπο με πραγματικές παραμέτρους για τις κατανομές. *Εδώ πρέπει να τονιστεί ξανά, πως το κάθε μοντέλο απευθύνεται και σε ένα φώνημα (π.χ "ah"), έτσι υπάρχει ο παραπάνω ορισμός και επανεκπαίδευση για κάθε φώνημα.*

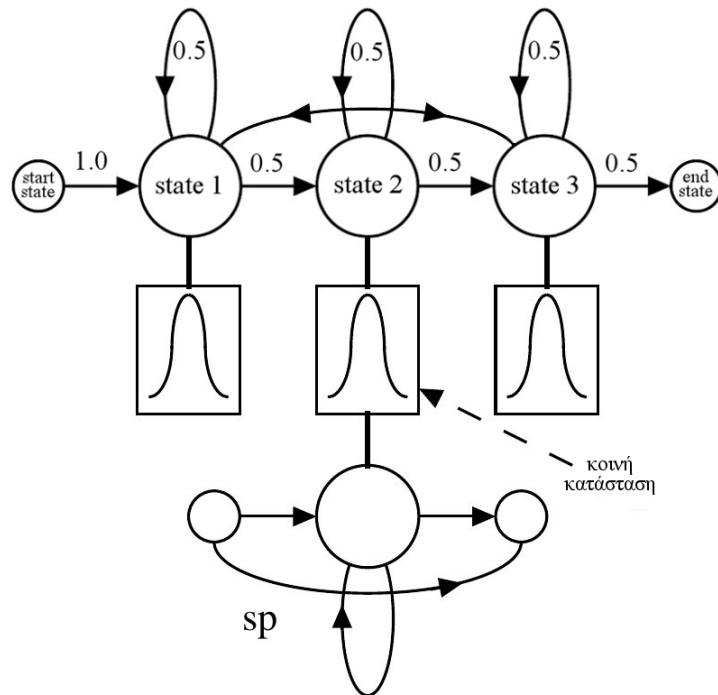
Κατά την επανεκτίμηση μπορούν να τεθούν κάποια όρια. Τα όρια αυτά αποσκοπούν στον περιορισμό των ευθυγραμμίσεων καταστάσεων που ο αλγόριθμος forward-backward συμπεριλαμβάνει στην άθροιση του. Έτσι υπάρχει κέρδος σε σχέση με το υπολογιστικό κόστος. Οι τιμές των ορίων θα αποδειχθούν και το σημαντικότερο κομμάτι της ανάλυσης των δεδομένων σε σχέση με το προς εκπαίδευση μοντέλο, μια προσεκτική επιλογή της τιμής αυτής θα δώσει και την καλύτερη δυνατή εκπαίδευση του μοντέλου.

7.2.2 Βήμα 6 - Διόρθωση μέσω των μοντέλων "σιωπής"

Σε προηγούμενο βήμα κατασκευάστηκε ένα μοντέλο 3-καταστάσεων από δεξιά στα αριστερά και ένα μοντέλο για την σιωπή. Το επόμενο βήμα είναι να προστεθούν μεταβάσεις από τις καταστάσεις 2 στην 4 και 4 στην 2. Η ιδέα είναι να γίνει το μοντέλο πιο ανθεκτικό επιτρέποντας στις επιμέρους καταστάσεις να απορροφούν τις όποιες *κρουστικές θορύβου* μπορεί να περιέχονται στο δείγμα. Επίσης, πρέπει να κατασκευαστεί το μοντέλο sr, που φαίνεται στην εικόνα 7.3, το μοντέλο αυτό λέγεται και T-μοντέλο. Έχει μια μόνο κατάσταση αλλά έχει και την δυνατότητα μετάβασης απευθείας από τον αρχικό στον τελικό κόμβο. Αυτό το μοντέλο έχει την κατάσταση εκπομπής του δεμένη με την κεντρική κατάσταση του μοντέλου *σιωπής*.

7.2.3 Βήμα 7 - Εκ νέου ευθυγράμμιση των δεδομένων εκπαίδευσης

Όπως αναφέρθηκε, για ίδιες λέξεις υπάρχει το ενδεχόμενο *πολλαπλών εκδόσεων (προφορών)* μέσω διαφορετικών φωνημάτων που την ορίζουν. Τα μοντέλα που έχουν δημιουργηθεί ως εδώ μπορούν να χρησιμοποιηθούν στην *εκ νέου ευθυγράμμιση* των δεδομένων εκπαίδευσης και την κατασκευή καινούργιων ετικετών. Η βασική διαφορά με την προηγούμενη μέθοδο συσχέτισης



Εικόνα 7.3: Μοντέλο σιωπής.

δεδομένων εκπαίδευσης με τις ετικέτες, είναι ότι επιλέγεται η καλύτερη ετικέτα (μέσα από τις διάφορες προφορές μιας λέξης) σύμφωνα με το καλύτερο ταίριασμα με τα δεδομένα εκπαίδευσης. Σε προηγούμενο βήμα είχαμε επιλέξει τυχαία την προφορά που χρησιμοποιείται για την κάθε λέξη. Με δεδομένο το ότι έχουν οριστεί οι ευθυγραμμίσεις, οι παράμετροι των μοντέλων μπορούν να επανεκτιμηθούν, για τις νέες προφορές.

7.3 Κατασκευή Τριφώνων Συνδεδεμένων-Καταστάσεων

Μετά τα μονοφωνικά Κρυφά Μαρκοβιανά Μοντέλα που κατασκευάστηκαν, το τελικό στάδιο βρίσκεται στη δημιουργία τριφωνικών μοντέλων, με καταστάσεις εξαρτώμενες από το περιεχόμενο των δεδομένων εκπαίδευσης (context-dependent). Αυτό θα πραγματοποιηθεί σε δύο στάδια. Αρχικά οι μονοφωνικές ετικέτες μετατρέπονται σε τριφωνικές και κατασκευάζεται ένα σύνολο τριφωνικών μοντέλων με αντιγραφή των ήδη κατασκευασμένων μονοφωνικών και επανεκτιμώντας. Έπειτα, οι όμοιες ακουστικά καταστάσεις των τριφωνικών μοντέλων συνδέονται, για να εξασφαλιστεί ότι όλες οι κατανομές πιθανοτήτων (καταστάσεων) μπορούν να εκτιμηθούν αξιόπιστα.

7.3.1 Βήμα 8 - Δημιουργία τριφώνων από μονόφωνα

Τα τριφωνικά μοντέλα μπορούν να δημιουργηθούν απλά από την κλωνοποίηση των προκατασκευασμένων μονοφωνικών μοντέλων και την εκ νέου εκτίμηση αυτών μέσω των τριφωνικών ετικετών. Βασικό είναι να κατασκευαστεί μια λίστα που θα περιέχει όλα τα τρίφωνα που υπάρχουν τουλάχιστον μια φορά στα δεδομένα εκπαίδευσης. Για κάθε λέξη θα προστεθούν τα "sil",

"sp" ως όρια. Για παράδειγμα οι λέξεις,

```
      This      man
sil th ih s sp m ae n sp ...
```

μετατρέπονται σε

```
sil th+ih th-ih+s ih-s sp m+ae m-ae+n ae-n sp ...
```

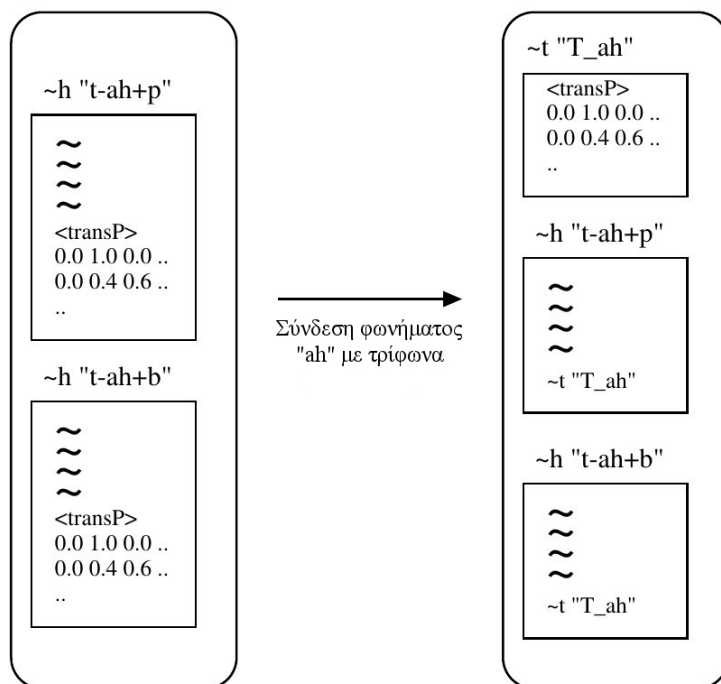
Αυτή η μορφή για τις τριφωνικές ετικέτες, αναφέρεται σαν *εσωτερικό λέξεως (word internal)*. Παρατηρούμε ότι παράγονται και *δίφωνα* από την ένωση φωνημάτων (μονοφώνων) των λέξεων, αλλά δεν θα παραχθούν δίφωνα που δεν γειτνιάζουν μέσα σε μία λέξη (πχ πρώτο και τρίτο φώνημα μίας λέξης). Η ένωση των φωνημάτων γίνεται μέσω δύο εντολών, της κλωνοποίησης (CL - clone) και της σύνδεσης (TI - tie) όλων των πινάκων πιθανοτήτων μετάβασης καταστάσεων σε κάθε τριφώνο

```
CL triphones1
TI T_ah {(*-ah+*, ah+*, *-ah).transP}
TI T_ax {(*-ax+*, ax+*, *-ax).transP}
TI T_ey {(*-ey+*, ey+*, *-ey).transP}
TI T_b {(*-b+*, b+*, *-b).transP}
TI T_ay {(*-ay+*, ay+*, *-ay).transP}
...
```

είναι προφανές ότι για κάθε φώνημα (πχ ah) κάποιας λέξης έχουμε την δημιουργία τριφώνου με σύνδεση μπρος-πίσω (*-ah+*), είτε δημιουργία διφώνου με σύνδεση μπρος (ah+*) και πίσω (*-ah). Ο αστερίσκος υποδεικνύει το "όποιο" φώνημα γειτνιάζει, τα "sil" και "sp" δεν θεωρούνται φωνήματα αλλά και δεν συνδέονται με κάποιο τρίφωνα αφού δεν είναι εξαρτώμενα (context-dependent). Τα T_ah, T_ax είναι μακρο-αναφορά (macros) που ορίζουν τα σύνθετα φωνήματα {(*-ah+*,...).transP}.

Κάθε εντολή (CL) δέχεται ως ορίσματα τρίφωνα και για κάθε ένα από αυτά τα μοντέλα της μορφής (a-b+c) παίρνει το φώνημα (b) και το αντιγράφει. Κάθε εντολή (TI) δέχεται ως όρισμα μακρο-αναφορές και τις παραμέτρους των μοντέλων που ορίζονται από αυτές. Έτσι για κάθε φώνημα που παρουσιάζεται στο κέντρο ενός τριφώνου προκύπτει (όπως φαίνεται στην εικόνα 7.4) η ακόλουθη σύνδεση.

Αυτός ο πίνακας μετάβασης αποτελεί κατά κάποιο τρόπο υπο-συνιστώσα για το κάθε μοντέλο λέξης. Με την σύνδεση των πινάκων μετάβασης, στην πραγματικότητα υποδεικνύουμε στο συνολικό σύστημα πως κάποια μοντέλα μοιράζονται το ίδιο σύνολο παραμέτρων. Στην παραπάνω εικόνα βλέπουμε αριστερά δύο ορισμούς μοντέλων όπου το καθένα έχει τον δικό του πίνακα μετάβασης καταστάσεων. Δεξιά φαίνεται η σύνδεση των μοντέλων, όπου τα μοντέλα πλέον δείχνουν στην μακρο-αναφορά που ορίζει και τις κοινές παραμέτρους αυτών.



Εικόνα 7.4: Παράδειγμα σύνδεσης πινάκων μετάβασης καταστάσεων για τρίφωνα με κεντρικό φώνημα το "ah".

Κατά την εκτίμηση παραμέτρων με αυτό τον τρόπο, προσπαθούμε να συλλέξουμε μοντέλα που έχουν ελάχιστες διαφορές μεταξύ τους μέσα από την εκπαίδευση. Έτσι είναι λογικό να μην συνδέσουμε μεταξύ τους, παραμέτρους μοντέλων που αλλάζουν αρκετά με βάση την δεδομένα εκπαίδευσης. Για κάποια τρίφωνα που θα εμφανιστούν μία ή δύο φορές στα δεδομένα εκπαίδευσης είναι βέβαιο πως θα υπάρχουν ελλείψεις εκτιμήσεις, άρα σε αυτή την περίπτωση θα ωφεληθούμε από την, από κοινού επανεκπαίδευση τους.

7.3.2 Βήμα 9 - Δημιουργία τριφώνων συνδεδεμένων-καταστάσεων

Το αποτέλεσμα του προηγούμενου βήματος είναι ένα σύνολο τριφωνικών μοντέλων, με κάθε ομάδα τριφώνων να έχει τον ίδιο πίνακα μετάβασης. Όταν εκτιμούνται οι παράμετροι αυτών των μοντέλων, οι περισσότερες από τις διακυμάνσεις των κατανομών εξόδου θα έχουν στρογγυλοποιηθεί προς τα κάτω, λόγω της έλλειψης δεδομένων για πολλές καταστάσεις. Το τελικό βήμα είναι η σύνδεση των καταστάσεων στο εσωτερικό των τριφώνων, έτσι ώστε να μοιράζονται δεδομένα και να βρεθούν καλύτερες εκτιμήσεις παραμέτρων.

Η επιλογή των καταστάσεων που θα πρέπει να συνδεθούν μεταξύ τους είναι μια διαδικασία πιο λεπτή, διότι θα καθορίσει σε μεγάλο βαθμό τα επίπεδα αναγνώρισης για το συνολικό σύστημα. Οι επιλογές για την ομαδοποίηση (clustering) και την μετ' έπειτα σύνδεση είναι δύο. Η πρώτη βασίζεται στα δεδομένα και χρησιμοποιεί μια μέτρηση ομοιότητας μεταξύ των καταστάσεων και η δεύτερη χρησιμοποιεί δέντρα αποφάσεων όπου υπάρχουν ερωτήματα σχετικά με το δεξί και το αριστερό μέρος κάθε τριφώνου. Με το δέντρο αποφάσεων γίνεται η προσπάθεια να βρεθούν τα

δεδομένα που θα οδηγήσουν στη μεγαλύτερη διαφορά, σε σχέση με την "ακουστική" του δείγματος και έτσι θα διαφοροποιηθούν για να σχηματιστούν ομάδες.

Για παράδειγμα μπορούμε να ορίζοντας το εξής

```
RO 100.0 stats

TR 0

QS "L_Class-Stop" {p-*,b-*,t-*,d-*,k-*,g-*}
QS "R_Class-Stop" {*+p,*+b,*+t,*+d,*+k,*+g}
...
QS "L_z" {z-*}
QS "R_z" {*+z}

TR 2

TB 350.0 "aa_s2" {(aa, *-aa, *-aa+*, aa+*).state[2]}
TB 350.0 "ae_s2" {(ae, *-ae, *-ae+*, ae+*).state[2]}
....
TB 350.0 "y_s4" {(y, *-y, *-y+*, y+*).state[4]}
TB 350.0 "z_s4" {(z, *-z, *-z+*, z+*).state[4]}
TB 350.0 "zh_s4" {(zh, *-zh, *-zh+*, zh+*).state[4]}

TR 1

AU "fulllist"
CO "tiedlist"

ST "trees"
```

όπου η εντολή (RO) χρησιμοποιείται για να θέσουμε το κατώφλι (outlier threshold) στο 100 και να φορτώσουμε το αρχείο στατιστικών stats. Το αρχείο αυτό προέκυψε στο προηγούμενο βήμα και αναφέρει στατιστικά σχετικά με την κατάληψη καταστάσεων από τρίφωνα και δίφωνα. Το κατώφλι αυτό ορίζει την ελάχιστη τιμή για την δημιουργία ομάδας, προς αποφυγή κάποιας κατάστασης που φαίνεται διαφορετική να αποτελέσει από μόνη της ομάδα. Η εντολή (QS) θέτει μια ερώτηση και κάθε ερώτηση ορίζεται από ένα σύνολο πιθανών φωνημάτων. Για παράδειγμα η πρώτη εντολή (QS) ορίζει μια ερώτηση με όνομα "L_Class-Stop" η οποία αληθεύει εάν το αριστερό μέρος του τριφώνου συμπίπτει με ένα από τα {p-*,b-*,t-*,d-*,k-*,g-*}.

Για ένα σύστημα τριφώνων, είναι απαραίτητο να συμπεριλαμβάνονται ερωτήσεις που αναφέρονται στο αριστερό αλλά και στο δεξί μέρος των τριφώνων. Ιδανικά, το σύνολο των ερωτήσεων θα πρέπει να είναι πλήρες, με την έννοια του ότι δεν αποκλείεται κανένας πιθανός συνδυασμός

που θα μπορούσε να παράξει αποτέλεσμα στην τελική αναγνώριση των προτύπων. Κάθε ερώτηση που ορίστηκε αλλά δεν μπορεί να απαντηθεί, απλά αγνοείται από το σύστημα.

Κάθε εντολή (TB) αναφέρεται και σε μία κατάσταση εξόδου, όπου εκεί, επιχειρεί να ομαδοποιήσει όλες τις καταστάσεις που περιέχουν κάποια χαρακτηριστικά. Για την κατάσταση 2 (πρώτη κατάσταση εξόδου), η εντολή (TB), εφαρμόζεται για το φώνημα "aa" με τα πιθανά τρίφωνα που προκύπτουν.

Η εντολή (TB) λειτουργεί ως εξής. Αρχικά κάθε σύνολο καταστάσεων που ορίζονται από το τελικό όρισμα της εντολής συγκεντρώνονται για να δημιουργήσουν μια ομάδα. Κάθε ερώτηση (QS) που φορτώνεται διχοτομεί την παραπάνω ομάδα. Η χρήση δύο συνόλων αντί ενός, επιτρέπει την λογαριθμική πιθανοφάνεια για τα δεδομένα εκπαίδευσης να αυξηθεί και τελικά ως πρώτος κλάδος του δέντρου επιλέγεται η ερώτηση που μεγιστοποιεί αυτή την πιθανοφάνεια. Η διαδικασία επαναλαμβάνεται μέχρις ότου η αύξηση της λογαριθμικής πιθανοφάνειας σε κάθε κόμβο να είναι μικρότερη από το κατώφλι (που είναι το πρώτο όρισμα της εντολής (TB), 350.0 στην περίπτωση μας).

Οι τιμές για τα κατώφλια στις εντολές (RO) και (TB) επηρεάζουν τον βαθμό σύνδεσης και έτσι τον αριθμό καταστάσεων εξόδου στο σύστημα, μετά την ομαδοποίηση. Οι τιμές αυτές πρέπει να σχετίζονται με το μέγεθος των δεδομένων εκπαίδευσης. Στο τέλος της διαδικασίας θα έχουμε την συγχώνευση για κάθε ζευγάρι ομάδων, με την προϋπόθεση ότι το αποτέλεσμα της συγχώνευσης δεν υπερβαίνει το κατώφλι. Ακόμη με την εντολή (AU) προσθέτουμε μια λίστα από τρίφωνα ώστε να συμπεριληφθούν και τα τρίφωνα που μπορεί να χρειαστούν κατά την διαδικασία αναγνώρισης. Αυτό επιτρέπει να υπάρχει η υποδομή για κάθε πιθανό τρίφωνο, εκτός των αναμενόμενων, που παρουσιάζονται στα δεδομένα εκπαίδευσης. Έτσι η εντολή τελικά παράγει τα κατάλληλα δέντρα αποφάσεων για την σύνθεση όλων των "καινούργιων" τριφώνων.

Όταν όλες οι συνδέσεις μεταξύ καταστάσεων έχουν δημιουργηθεί, κάποια από τα μοντέλα μπορεί να έχουν όλες τις καταστάσεις εξόδου τους κοινές και τους πίνακες μετάβασης, έτσι θα είναι όμοια. Με την εντολή (CO) αυτά τα μοντέλα που ταυτίζονται θα συνδεθούν μεταξύ τους και θα σωθούν σε ένα αρχείο "tiedlist". Η τελευταία εντολή (ST) αποθηκεύει τα δέντρα.

7.4 Αξιολόγηση Επιπέδων Αναγνώρισης

Το σύστημα σε αυτό το στάδιο είναι πλήρες και το μόνο που μένει να αποτιμηθεί είναι η αποτελεσματικότητά του. Τα μέρη του συστήματος έχουν οριστεί και τα δεδομένα εκπαίδευσης έχουν αξιοποιηθεί.

7.4.1 Βήμα 10 - Αναγνώριση των δεδομένων αξιολόγησης

Εάν θεωρήσουμε πως ένα αρχείο *test.scp* περιέχει την λίστα με την τοποθεσία των κωδικοποιημένων αρχείων κυματομορφών και με τον ορισμός δύο παραμέτρων αξιολόγησης. Της *ποινής εισαγωγής λέξης (word insertion penalty)* και της *κλίμακας γραμματικού παράγοντα (grammar scale factor)*. Το πρώτο αναφέρεται σε μια δεδομένη τιμή που θα προστίθεται σε κάθε αλλαγή λέξης κατά την αναγνώριση και το δεύτερο έχει να κάνει με τον πολλαπλασιασμό της λογαριθμικής πιθανότητας κατά την εναλλαγή λέξεων. Οι κυματομορφές θα αναγνωριστούν και στην συνέχεια θα γίνει ο απολογισμός της ακρίβειας των αποτελεσμάτων με βάση τις γνωστές ετικέτες για τα δεδομένα αυτά. Κάθε απόκλιση από την σωστή τιμή θα έχει επίπτωση στα ποσοστά αναγνώρισης σύμφωνα με τα παρακάτω.

Η αξιολόγηση της αναγνώρισης γίνεται με ταύτιση συμβολοσειρών. Το αποτέλεσμα της αναγνώρισης είναι ένα αρχείο εξόδου που περιέχει τις ετικέτες που "βλέπει" το μοντέλο για κάθε κυματομορφή. Αυτό το αρχείο θα συγκριθεί με το αρχείο των πραγματικών ετικετών. *Εδώ παρατηρούμε πως το μοντέλο μπορεί να παρουσιάσει αποκλίσεις στην αναγνώριση, των κυματομορφών εκπαίδευσης, οι οποίες κατά τα άλλα είναι γνωστές στο σύστημα.* Αυτό συμβαίνει λόγω στοχαστικής φύσης του μοντέλου, αφού η κάθε κυματομορφή δεν αποθηκεύεται αλλά επηρεάζει παραμέτρους του μοντέλου συνολικά και όχι τοπικά.

Όταν γίνει η αναγνώριση των συμβολοσειρών θα αξιολογηθεί η συνολική απόδοση μέσα από τα δεδομένα αναγνώρισης (προς αναγνώριση). Ο υπολογισμός του *Ποσοστού Αναγνώρισης (Percent Correct)* έχει ως εξής

$$\text{Percent Correct} = \frac{N - D - S}{N} \times 100\%$$

που μας δίνει το ποσοστό σωστών αναγνωρίσεων, με (N - number) το άθροισμα των κυματομορφών προς αναγνώριση, (S - substitution) τον αριθμό αντικαταστάσεων, (D - deletion) τον αριθμό διαγραφών και με (I - insertion) τις εισαγωγές λέξεων. Το ποσοστό αναγνώρισης δεν εμπεριέχει την ποινή που προκύπτει από την εισαγωγή λέξεων αλλά το *Ποσοστό Ακρίβειας (Percent Accuracy)* την συμπεριλαμβάνει

$$\text{Percent Accuracy} = \frac{N - D - S - I}{N} \times 100\%$$

και είναι μια πιο λεπτομερής περιγραφή του ποσοστού αναγνώρισης. Ένα παράδειγμα της μορφής εξόδου από την ρουτίνα αναγνώρισης είναι το εξής

Aligned transcription: cat2_file_3_1.lab vs cat2_file_3_1.rec

LAB: pac

REC: one

Aligned transcription: cat2_file_3_2.lab vs cat2_file_3_2.rec

LAB: pac

REC: one

```

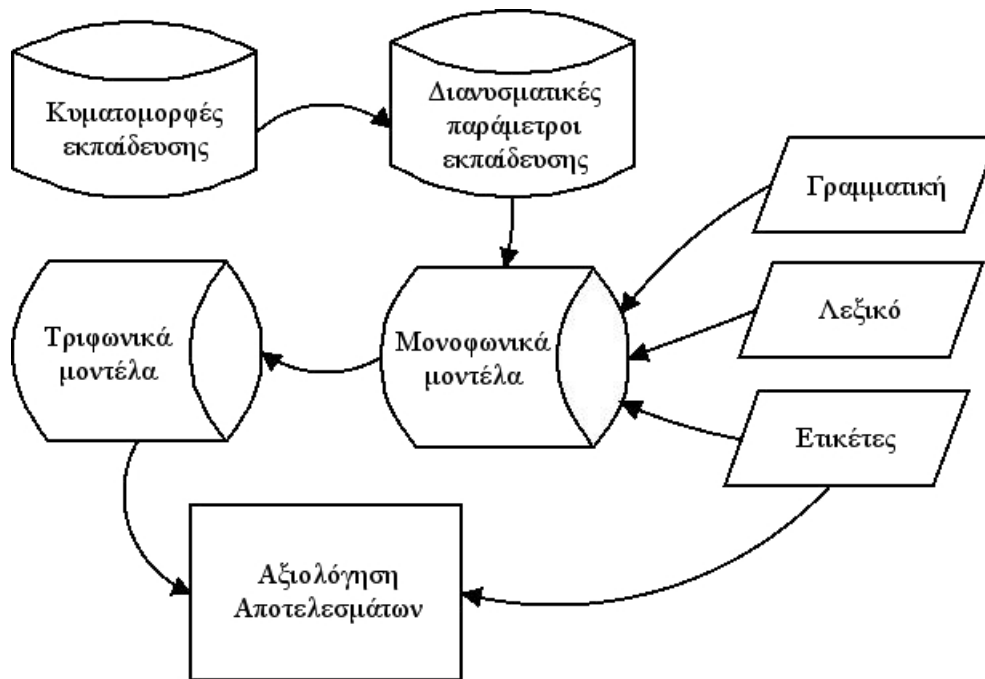
Aligned transcription: cat1_file_3_3.lab vs cat1_file_3_3.rec
LAB:    one
REC: one one
Aligned transcription: cat2_file_3_6.lab vs cat2_file_3_6.rec
LAB:    pac
REC: one one
Aligned transcription: cat2_file_3_13.lab vs cat2_file_3_13.rec
LAB:    pac
REC: one pac
Aligned transcription: cat2_file_3_26.lab vs cat2_file_3_26.rec
LAB:    pac
REC: one
Aligned transcription: cat1_file_3_27.lab vs cat1_file_3_27.rec
LAB:    one
REC: one one
===== HTK Results Analysis =====
Date: Sat Nov 23 02:19:43 2013
Ref : wordz/words-3.mlf
Rec : recout/recout-3.mlf
----- Overall Results -----
SENT: %Correct=73.08 [H=19, S=7, N=26]
WORD: %Corr=84.62, Acc=69.23 [H=22, D=0, S=4, I=4, N=26]
=====

```

Όπου έχουμε ένα δείγμα 26 κυματομορφών προς αναγνώριση, όπου υπήρξε η απόλυτη ταύτιση των συμβολοσειρών/ετικετών εξόδου (recout-3.mlf) με τις συμβολοσειρές αναφοράς (words-3.mlf), δεν υπήρξε διαγραφή λέξης, υπήρχαν 4 αντικαταστάσεις και 4 εισαγωγές συνολικά. Ακόμη υπάρχει και η ευθυγράμμιση των συμβολοσειρών (μέσω της εισαγωγής λέξεων) για διάφορες κυματομορφές, η οποία δεν φέρει επιπλέον βάρος στο ποσοστό αναγνώρισης, εφόσον η ευθυγράμμιση τελικά βοηθά στην αναγνώριση. Τέλος με την λέξη SENT ορίζουμε το ποσοστό αναγνώρισης προτάσεων και με την λέξη WORD το ποσοστό αναγνώρισης λέξεων. Όπως περιγράφηκε οι τιμές αναφέρονται σε *ποσοστό αναγνώρισης* (πχ 84.62%) και σε *ποσοστό ακρίβειας*(πχ 69.23%).

7.4.2 Σύνοψη διαδικασίας

Στην εικόνα 7.5 συνοψίζεται η διαδικασία κατασκευής του μοντέλου μέχρι εδώ.



Εικόνα 7.5: Διάγραμμα με τα βασικά συστατικά και τη σύνδεση τους για την κατασκευή του συστήματος.

Κεφάλαιο 8

Αξιολόγηση

Σε αυτό το κεφάλαιο θα γίνει η βελτιστοποίηση των εσωτερικών παραμέτρων του μοντέλου και βρίσκοντας τις βέλτιστες τιμές θα βρεθεί η *συνολική απόδοση αναγνώρισης του μοντέλου*, για κάθε ασθενή καθώς και για κάθε κυματομορφή. Αυτό θα γίνει κυρίως με την χρησιμοποίηση του *κατωφλιού αποκοπής* μέσα από την διαδικασία της επανεκπαίδευσης.

8.1 Παράμετροι αξιολόγησης

Κατά τον υπολογισμό της οπίσθιας μεταβλητής (όπως ορίστηκε στο κεφ. 4) για κάθε χρονική στιγμή πλέον αγνοούνται όλες οι τιμές που πέφτουν παρακάτω, από την μέγιστη τιμή μειωμένη κατά την τιμή του κατωφλιού αποκοπής. Κατά τον υπολογισμό της εμπρόσθιας μεταβλητής θα έχουμε σαν προϋπόθεση να ορίζουν οι αντίστοιχες εμπρόσθιες μεταβλητές. Η τιμή του κατωφλιού θα δοθεί σαν τριπλό όρισμα στον αλγόριθμο επανεκπαίδευσης με τη μορφή, *κατώφλι, βήμα, ανώτατο όριο* (prune, step, supremum). Το κατώφλι ορίζεται στο πρώτο όρισμα και στην περίπτωση που είναι πολύ μικρό και δεν προκύπτει υπολογισμός το κατώφλι αυτό θα αυξηθεί κατά "βήμα" (δεύτερο όρισμα). Στην περίπτωση που και πάλι δεν λαμβάνουμε αποτέλεσμα θα προστεθεί πάλι ένα "βήμα" στην προηγούμενη τιμή του κατωφλιού, επαναληπτικά μέχρι να υπάρξει αποτέλεσμα. Το τελευταίο όρισμα δίνει την μέγιστη τιμή που μπορεί να λάβει το κατώφλι μετά τις αυξήσεις.

Η δεύτερη παράμετρος που θα χρειαστεί βελτιστοποίηση είναι η ίδια η κυματομορφή ως προς την κανονικοποίηση της, προτού εισέλθει στο μοντέλο, ως ένα στάδιο προεπεξεργασίας, όπως θα φανεί και στη συνέχεια. Η κυματομορφή αυτή θα πρέπει να έχει συγκεκριμένα χαρακτηριστικά πριν μετασχηματιστεί σε ηχητική μορφή. Αυτό γίνεται διότι το εργαλείο που θα χρησιμοποιήσουμε έχει σαν κατεύθυνση την αναγνώριση φωνής άρα και θα μετατρέψουμε τις κυματομορφές των διαφόρων χαρακτηριστικών που έχουμε εξάγει από τους υπερήχους, σε αρχεία ήχου.

8.2 Σύστημα αξιολόγησης

Η προεπεξεργασία αναφέρεται στην ιδιότητα που πρέπει να έχει μια κυματομορφή ώστε να μπορέσει να μετατραπεί σε αρχείο ήχου (Waveform Audio File Format - wav) συχνότητας δειγματοληψίας 8000Hz και διακριτοποίηση 16bit. Οι κυματομορφές που υπάρχουν στο δείγμα παρουσιάζουν απόλυτες τιμές αρκετά μεγαλύτερες της μονάδας. Οι τρόποι που αυτές οι κυματομορφές γίνονται κατάλληλες προς το μοντέλο είναι οι ακόλουθοι τρεις. Όπου θα συνδυαστούν με τα μονοφωνικά και τα τριφωνικά μοντέλα.

1. Για τις κυματομορφές που $\max|f(x)| \leq 1$ θα κρατήσουμε την κυματομορφή ως έχει. Και για τις υπόλοιπες κυματομορφές θα έχουμε $\frac{f(x)}{\max|f(x)|+0.05}$
2. Για τις κυματομορφές που $\max|f(x)| \leq 1$ θα έχουμε $\text{timescale}(f(x), \text{class_maximum})$. Και για τις υπόλοιπες κυματομορφές θα έχουμε $\text{timescale}(\frac{f(x)}{\max|f(x)|+0.05}, \text{class_maximum})$
3. Για όλες τις κυματομορφές θα έχουμε $\text{timescale}(\frac{f(x)}{\max|f(x)|+0.05}, \text{class_maximum})$.

Οι τρεις αυτές προσεγγίσεις κοιτάζουν με διαφορετικό τρόπο τα δεδομένα. Η τρίτη επιλογή δίνει μια ομοιογένεια στο δείγμα μέσω της κλιμάκωσης των επιτρεπτών τιμών όλων των κυματομορφών στο πεδίο $(-1, 1)$ αλλά και μέσω της χρονικής κλιμάκωσης των κυματομορφών. Αυτή η κλιμάκωση στον χρόνο γίνεται σύμφωνα με την μεγαλύτερη σε χρόνο κυματομορφή (class_maximum) για την συγκεκριμένη κατηγορία. Δηλαδή για κάθε κατηγορία οι κυματομορφές των ασθενών της, τελικά έχουν ένα σχεδόν ταυτόσημο μήκος περνώντας μέσα από την συνάρτηση $\text{timescale}()$.

Η συνάρτηση timescale δέχεται ένα σήμα και την επιθυμητή διάρκεια του τελικού σήματος. Η συνάρτηση προσθέτει ή αφαιρεί κατάλληλα δείγματα μέσα στην κυματομορφή ώστε να προκύψει η τελική ζητούμενη διάρκεια. Για την πρόσθεση καινούργιων τιμών σε κάποια παρούσα κυματομορφή ώστε να διευρυνθεί η χρονική της διάρκεια, χρησιμοποιείται η αριθμητική μέση τιμή σε σχέση με τις γειτονικές. Η μέθοδος γίνεται με αναδρομικό τρόπο, μέσω αύξησης και μείωσης των σημείων όπου υπάρχουν στην κυματομορφή, αυτή ανατροφοδοτεί την ανάδραση και έτσι προσεγγίζει την τελική διάρκεια από δεξιά και αριστερά ώστε να καταλήξει στην επιθυμητή διάρκεια.

Η ανάλυση έχει 2 βασικά μέρη. Το πρώτο μέρος αναφέρεται στην βελτιστοποίηση των συντελεστών για κάθε κατηγορία κυματομορφών, όπου πρακτικά είναι το μέρος που το μοντέλο εκπαιδεύεται με τον βέλτιστο τρόπο, σύμφωνα με τα υπάρχοντα δεδομένα. Το δεύτερο κομμάτι έχει να κάνει με την αξιολόγηση του μοντέλου μέσα από την τεχνική του leave-one-out (cross-validation), όπου επαληθεύει την ακρίβεια του μοντέλου μέσα από την εκπαίδευση του μοντέλου για όλους τους ασθενείς εκτός ενός. Και στην συνέχεια για αυτό τον ασθενή θα γίνει η αξιολόγηση. Φυσικά το ίδιο θα επαναληφθεί για κάθε ασθενή, όπου θα χρησιμοποιηθεί για την αναγνώριση όλων των κυματομορφών.

8.3 Οργάνωση πειραμάτων

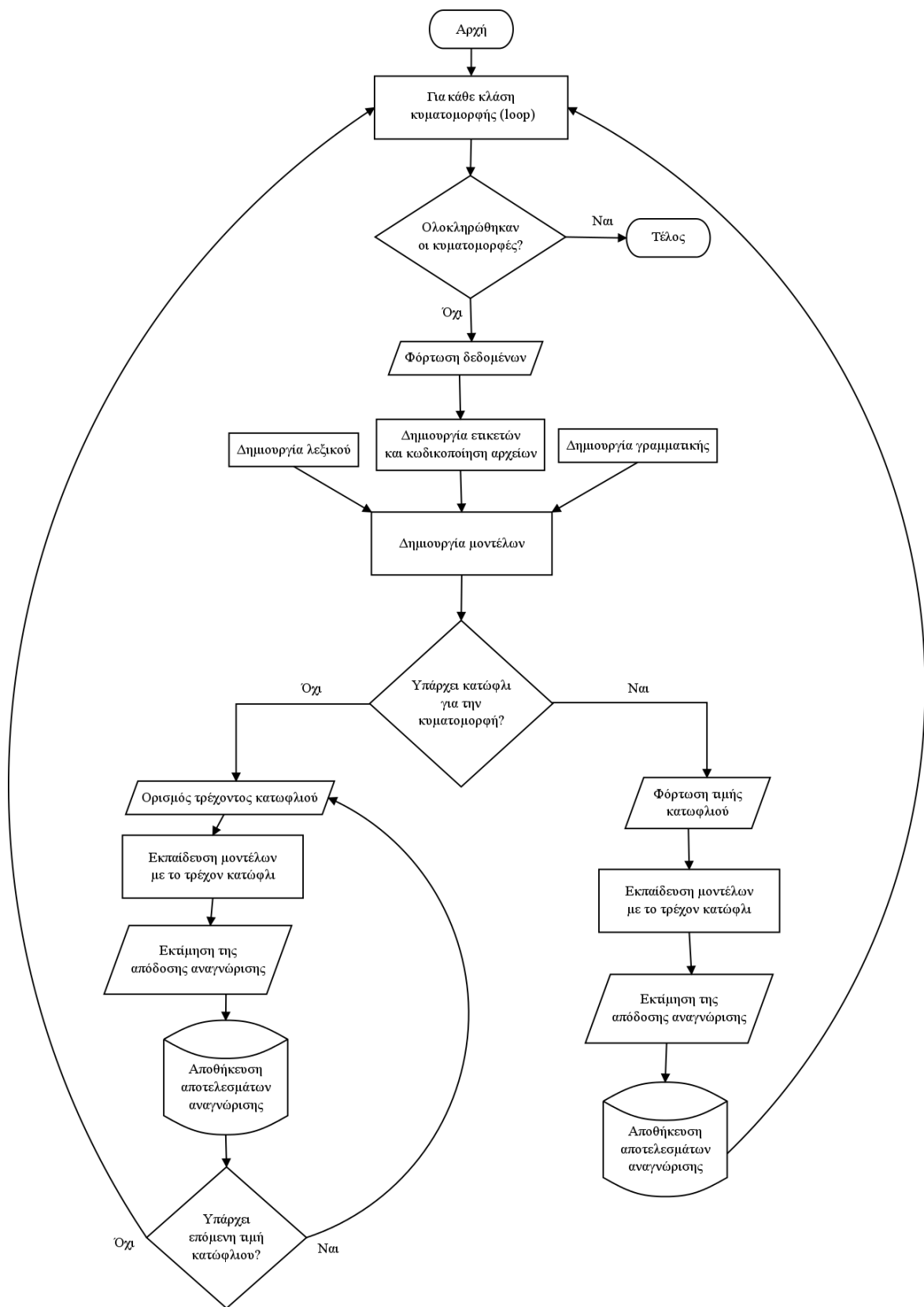
Η αρχική ανάλυση όπως θα φανεί παρακάτω είναι κοινή για όλες τις περιπτώσεις και έχει να κάνει με την εύρεση των βέλτιστων κατωφλιών αποκοπής, για κάθε μία από τις 146 κυματομορφές (features) που έχουν εξαχθεί από κάθε ασθενή. Οι τιμές των κατωφλιών θα βρεθούν μέσα από την εκπαίδευση του μοντέλου χρησιμοποιώντας όλα τα πειραματικά δεδομένα αλλά και η αξιολόγηση του μοντέλου θα χρησιμοποιήσει, επίσης, όλα τα διαθέσιμα δεδομένα. Με αυτό τον τρόπο βρίσκουμε κατά πόσο μπορεί η κάθε τιμή κατωφλιού στο μοντέλο να αναγνωρίσει τα δεδομένα εκπαίδευσης και όχι άγνωστα δεδομένα. Έτσι, βρίσκουμε την ικανότητα (ποσοστό) του μοντέλου να αφομοιώσει τα πειραματικά δεδομένα.

8.3.1 Αλγόριθμος εύρεσης βέλτιστων τιμών κατωφλιών

Τα βήματα που χρειάζονται για την διαδικασία εύρεσης των τιμών κατωφλιών, παρουσιάζονται στην εικόνα 8.1. Αρχικά γίνεται η φόρτωση των δεδομένων για την προεπεξεργασία που αναλύθηκε σε προηγούμενο κεφάλαιο. Αυτό το βήμα δηλαδή έχει να κάνει με τις αναγκαίες κατηγοριοποιήσεις των κυματομορφών, a priori, αλλά και την δημιουργία των αρχείων ήχου, ετικετών, λεξικού, γραμματικής και αρχείων λιστών που ορίζουν ποια δεδομένα θα χρησιμοποιηθούν στην εκπαίδευση και ποια στην αξιολόγηση του μοντέλου.

Κατά τη δημιουργία των μοντέλων, γίνεται μια επαναληπτική αξιολόγηση της εκπαίδευσης του κάθε μοντέλου για διαφορετικές τιμές κατωφλιού αποκοπής. Αυτό έχει ως αποτέλεσμα ένα αρχείο που περιέχει όλα τα αποτελέσματα (ποσοστά) αναγνώρισης για κάθε μια από τις 146 κατηγορίες κυματομορφών. Το κάθε αρχείο αναφέρεται σε μία κατηγορία και εμπεριέχει τα αποτελέσματα για όλες τις τιμές κατωφλιού. Άρα, από τη διαδικασία αυτή παράγονται 146 αρχεία, όπου περιέχουν περί τα 280 μεμονωμένα πειράματα εκπαίδευσης έκαστο. Τα αρχεία αυτά έχουν την εξής μορφή

```
----- Overall Results -----  
SENT: %Correct=61.54 [H=16, S=10, N=26]  
WORD: %Corr=92.31, Acc=46.15 [H=24, D=0, S=2, I=12, N=26]  
=====  
prune 0.100000 step 0.100000  
  
...  
...  
  
----- Overall Results -----  
SENT: %Correct=73.08 [H=19, S=7, N=26]  
WORD: %Corr=84.62, Acc=69.23 [H=22, D=0, S=4, I=4, N=26]  
=====  
prune 28.300000 step 0.100000
```



Εικόνα 8.1: Διάγραμμα που περιγράφει τις κύριες διαδικασίες και επαναλήψεις του πειράματος εύρεσης των βέλτιστων κατωφλιών αποκοπής.

Σκοπός αυτής της διαδικασίας είναι η ανάλυση των επιμέρους αρχείων και η δημιουργία ενός μονάχα αρχείου, το οποίο θα περιέχει την τιμή του κατωφλιού, που ορίζει το μέγιστο επίπεδο αναγνώρισης (accuracy) για την κάθε κατηγορία. Αυτό θα γίνει μέσα από ένα bash script στο unix shell που χρησιμοποιεί έναν αλγόριθμο, που

1. Για κάθε ένα από τα 146 αρχεία αποτελεσμάτων.
2. Αναζητά γραμμικά, από την πρώτη ως την τελευταία γραμμή, αυτή που περιέχει την έκφραση "acc=100". Δηλαδή την μεγαλύτερη δυνατή ακρίβεια.
3. Σε περίπτωση που δεν βρεθεί αυτή η ακρίβεια, τότε η γραμμική διαδικασία αναζήτησης στο έγγραφο επαναλαμβάνεται με βήμα μείωσης του ποσοστού αναγνώρισης -0.01, δηλαδή "acc=99.99".
4. Στην περίπτωση που θα βρεθεί η έκφραση για την ακρίβεια που ορίστηκε παραπάνω. Τότε, εξάγει την γραμμή που ακολουθεί το ποσοστό αυτό και περιέχει την λέξη prune. Αυτή η γραμμή αντιγράφεται ως έχει.
5. Και τελικά θα εγγραφεί σε ένα αρχείο εξόδου με τη μορφή class-"1-146" prune "0.1-28.3" step "0.1".

Η ενδεικτική μορφή του αρχείου (class.txt) που προκύπτει για τις βέλτιστες τιμές κατωφλιού είναι η εξής:

```
class-1 prune 27.700000 step 0.100000
class-2 prune 0.100000 step 0.100000
class-3 prune 27.900000 step 0.100000
class-4 prune 28.100000 step 0.100000
class-5 prune 4.300000 step 0.100000
class-6 prune 3.100000 step 0.100000
...
...
...
class-143 prune 20.100000 step 0.100000
class-144 prune 6.300000 step 0.100000
class-145 prune 5.100000 step 0.100000
class-146 prune 0.100000 step 0.100000
```

Η διαδικασία που περιγράφηκε θα επαναληφθεί για 6 περιπτώσεις συνολικά, 3 για μονοφωνικά μοντέλα και 3 για τριφωνικά. Το σύνολο δηλαδή των πειραμάτων μέχρι και αυτό το στάδιο είναι $6 * 146 * 283 = 247908$ σε αριθμό. Μετά το πέρας αυτής της διαδικασίας εκπαίδευσης και αξιολόγησης των μοντέλων με τα γνωστά (ίδια) δεδομένα, θα πρέπει να ακολουθήσει, χρησιμοποιώντας τα βέλτιστα κατώφλια, η αξιολόγηση με άγνωστα δεδομένα για το μοντέλο. Δηλαδή, με τεχνική του leave-one-out (cross-validation) όπου αφήνουμε κάθε φορά έναν ασθενή άγνωστο προς το μοντέλο (κατά τη διαδικασία εκπαίδευσης) για να τον χρησιμοποιήσουμε στη συνέχεια σαν test case

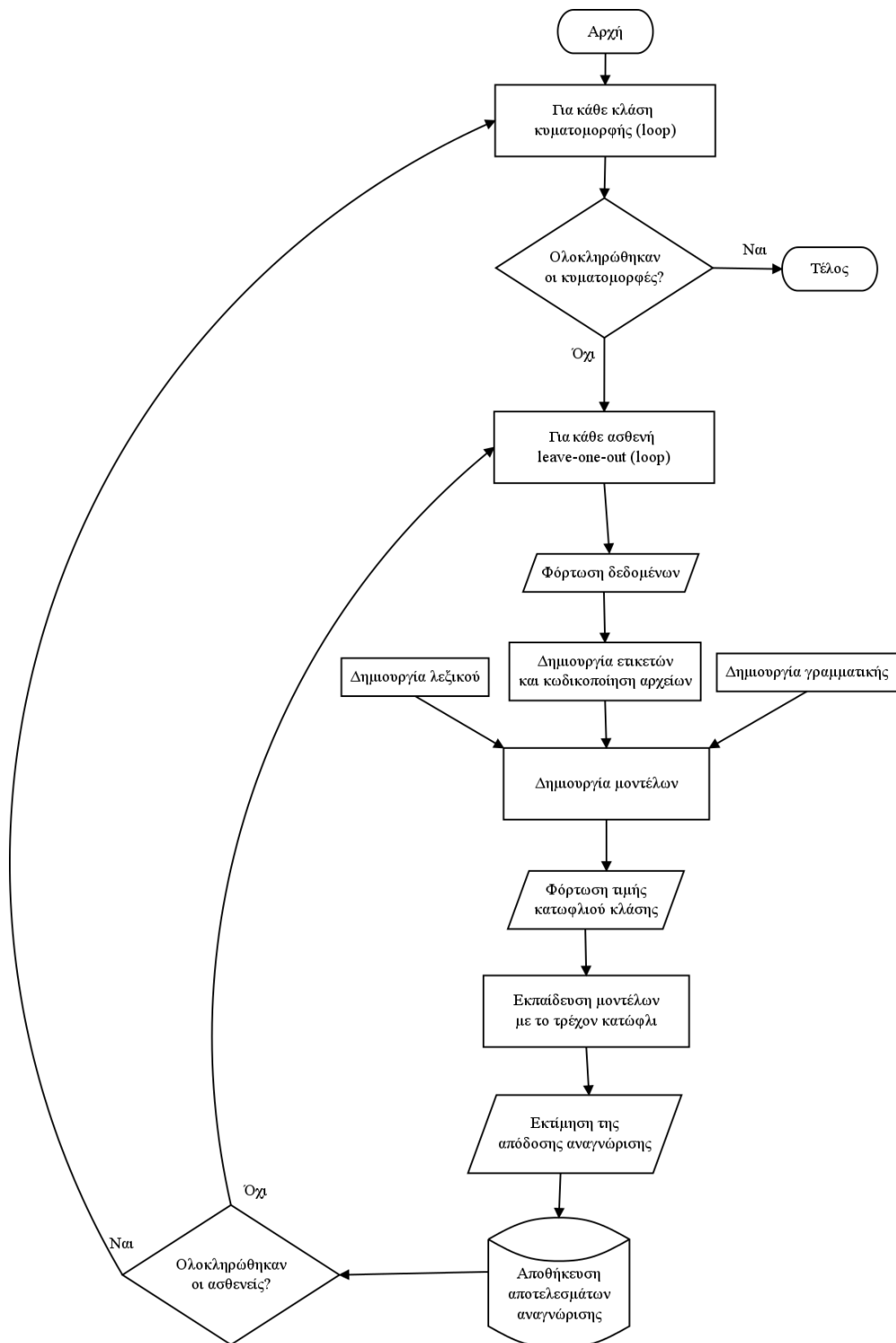
για την αξιολόγηση της ικανότητας του μοντέλου να κατηγοριοποιήσει τις επιμέρους κατηγορίες κυματομορφών του.

8.3.2 Αλγόριθμος *leave-one-out*, *cross validation*

Έχοντας ως πλέον καταγεγραμμένα σε αρχείο τα αποτελέσματα της ανωτέρω διαδικασίας, θα τα χρησιμοποιήσουμε στην *αξιολόγηση του μοντέλου με άγνωστα δεδομένα*. Ο παρακάτω αλγόριθμος (εικόνα 8.2) συνοψίζει την διαδικασία.

Σε αυτή τη περίπτωση έχουμε την *αναγνώριση της κάθε κυματομορφής ξεχωριστά για τον άγνωστο ασθενή*. Αυτό σημαίνει πως έχουμε την εκπαίδευση (με τις βέλτιστες τιμές καταφυλιών) του μοντέλου κυκλικά, αγνοώντας κάθε φορά έναν ασθενή, όπου αυτός ο ασθενής θα αναγνωριστεί τελικά (ως άγνωστο δείγμα). Στην προσέγγιση αυτή θα έχουμε μετά από κάθε πέρασμα, για κάθε κυματομορφή δηλαδή, ένα αρχείο εξόδου (αναγνώρισης). Σε αυτή τη περίπτωση κάθε αρχείο θα απαντά σε δυαδική μορφή, είτε 0% είτε 100% ως προς την ταύτιση με την αλήθεια (α priori γνώση μέσω των ετικετών). Εδώ έχουμε στην πραγματικότητα ένα πείραμα για την κάθε κυματομορφή, σύνολο $(\text{άθροισμα_ασθενών}) * (146 \text{ κατηγορίες})$.

Τα αποτελέσματα θα εξαχθούν από αυτά τα αρχεία πάλι μέσω ενός bash script και θα παραχθεί ως έξοδος ένα αρχείο *.csv. Με παρόμοιο τρόπο θα εξαχθούν τα δεδομένα μέσω αναγνώρισης εκφράσεων στο κείμενο των αρχείων. Το *.csv αρχείο χρησιμεύει διότι η δομή του βοηθά την απευθείας απεικόνιση σε προγράμματα που χρησιμοποιούν *λογιστικά φύλλα* (spreadsheet), έτσι η μετα-ανάλυση των αποτελεσμάτων μπορεί να συνεχιστεί με αυτών των προγραμμάτων.



Εικόνα 8.2: Διάγραμμα που περιγράφει τις κύριες διαδικασίες και επαναλήψεις του cross validation.

Κεφάλαιο 9

Επαλήθευση αποτελεσμάτων

Τα αποτελέσματα της ανωτέρω πειραματικής διαδικασίας, παρόλο που μας ικανοποιούν, ως μηχανικοί δεν μπορούμε να τα εμπιστευθούμε χωρίς μια παραπάνω διερεύνηση. Η διαδικασία που θα ακολουθήσει μπορεί να παρομοιαστεί με την επαλήθευση των αστρονόμων ως προς την ύπαρξη της σκοτεινής ενέργειας που γίνεται μέσω 3 ευθειών που σχηματίζονται από τις ενέργειες της κοσμικής μικροκυματικής ακτινοβολίας υποβάθρου (cosmic microwave background radiation), της ενέργειας από τις εκρήξεις υπερnovα και την ενέργεια που καταναλώνεται για την διαστολή του σύμπαντος. Το σημείο τομής αυτών τριών των ευθειών έδωσε και τη βάση για την απόδειξη της θεωρίας.

Έτσι, κατασκευάζοντας ένα νέο μοντέλο το οποίο θα προσεγγίσει το ίδιο πρόβλημα, θα έχουμε δύο σημεία για τη λύση του ίδιου προβλήματος.

9.1 Εκπαίδευση μοντέλων SVM

Η εκπαίδευση θα γίνει με τρόπο όμοιο των Κρυφών Μαρκοβιανών Μοντέλων. *Η προσέγγιση χρησιμοποιεί 4 διαφορετικούς πυρήνες για τη περίπτωση της αυτόματης κλιμάκωσης των κυματομορφών (autoscaling) και την περίπτωση που δεν πραγματοποιείται καμία κλιμάκωση. Άρα, συνολικά 8 διαφορετικές περιπτώσεις μοντέλων SVM.*

Για κάθε κυματομορφή θα κατασκευαστούν τα αντίστοιχα διανύσματα για όλες τις (διακριτές) χρονικές στιγμές, όπου αυτές οι τιμές αντιστοιχούν στα πλάτη τους. Με αυτό τον τρόπο θα δημιουργήσουμε για κάθε κατηγορία δύο σύνολα, όπου το ένα θα αποτελείται από διανύσματα της πρώτης κατηγορίας (συμπτωματικοί ασθενείς) και το δεύτερο σύνολο από τα υπόλοιπα (ασυμπτωματικοί ασθενείς).

Κατά τη δημιουργία των μοντέλων θα χρειαστεί να υπολογίσουμε τον αριθμητικό μέσο για το πλάτος κάθε κυματομορφής διότι υπάρχουν περιπτώσεις κυματομορφών που παρουσιάζουν στοιχεία NaN (Not a Number) μη αριθμητικά και έτσι προκύπτει πρόβλημα ως προς τις αριθμητικές

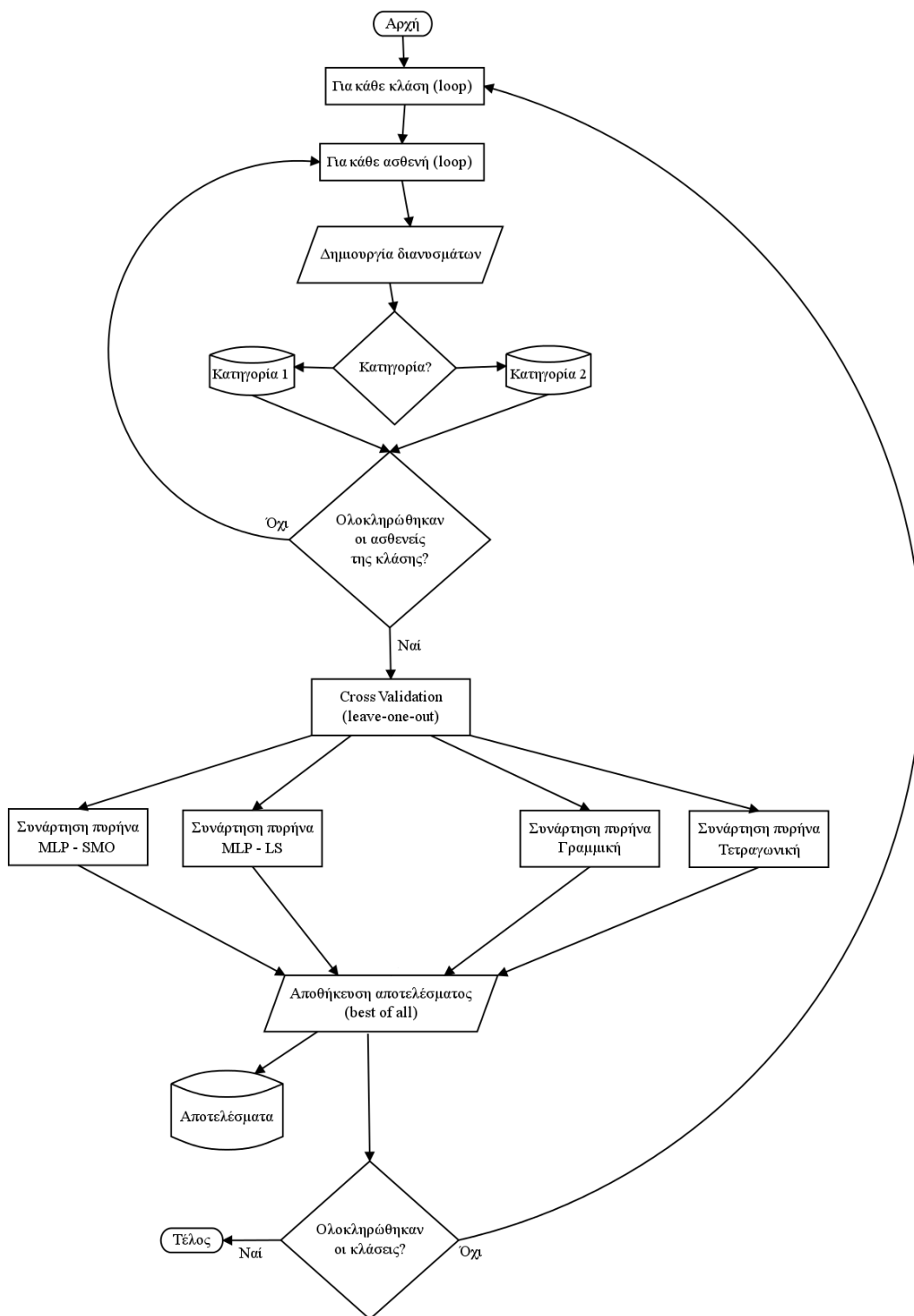
πράξεις που απαιτεί το μοντέλο. Για αυτή τη περίπτωση υπάρχει η επιλογή είτε να μην θεωρηθεί έγκυρη η κυματομορφή, είτε να θεωρηθεί μονάχα ένα κομμάτι της έγκυρο, είτε να προστεθούν κάποιες τιμές έναντι των τιμών που απουσιάζουν. Για να προσθέσουμε τιμές θα μπορούσαμε να παρεμβάλλουμε μεταξύ δύο γνωστών τιμών με κάποιο τρόπο, πράγμα που θα οδηγούσε στην αλλοίωση της αρχικής κυματομορφής ως προς το περιεχόμενο (άρα θα κάναμε έμμεσα μια πρόβλεψη). Η επιλογή του να προσθέσουμε την λιγότερο επιβλαβή τιμή οδηγεί στο να τοποθετήσουμε μια τιμή που να μην εμπλέκεται τουλάχιστον στα στατιστικά δεδομένα με βίαιο τρόπο. Έτσι επιλέγουμε να προσθέσουμε τον αριθμητικό μέσο του σήματος για κάθε NaN που συναντούμε. *Γνωρίζοντας πως τα δεδομένα που λείπουν είναι λίγα σε αριθμό, θεωρούμε πως το τελικό αποτέλεσμα δεν επηρεάζεται στο σύνολό του.*

9.1.1 Αλγόριθμος εκπαίδευσης

Όπως φαίνεται στην εικόνα 9.1, η διαδικασία εκπαίδευσης έχει ως σκοπό την επαλήθευση των αποτελεσμάτων μέσω της τεχνικής διασταύρωσης leave-one-out και μέσω της σύγκρισης των αποτελεσμάτων για κάθε πυρήνα μετασχηματισμού, βρίσκουμε τον καταλληλότερο πυρήνα για κάθε κατηγορία κυματομορφών. Έτσι, ως έξοδο έχουμε τα μέγιστα ποσοστά αναγνώρισης για κάθε κατηγορία, όπου τα αποτελέσματα της κάθε κατηγορίας προέκυψαν από κάποιο πυρήνα εκπαίδευσης.

Εδώ θα πάρουμε τα τελικά δεδομένα του cross validation για τις δύο περιπτώσεις που μας ενδιαφέρουν. Ο διαχωρισμός γίνεται ως προς την παράμετρο του μοντέλου για την κανονικοποίηση των δεδομένων. Αυτό θα μας δώσει δηλαδή παρόμοια γραφήματα με αυτά των Κρυφών Μαρκοβιανών Μοντέλων, με τη διαφορά ότι για κάθε κατηγορία θα περιέχονται 2 σημεία.

Η διαδικασία που ακολουθήθηκε στα Κρυφά Μαρκοβιανά Μοντέλα ως προς την προεπεξεργασία των κυματομορφών εδώ κρίθηκε περιττή, καθώς δεν υπήρχε η δέσμευση για εύρος τιμών $(-1, 1)$ όπως και η ανάγκη για αλγοριθμική επεξεργασία του χρονικού εύρους των κυματομορφών, διότι η παράμετρος της κλιμάκωσης (scaling) που διαθέτει το μοντέλο SVM είναι ευθέως ανάλογη της διαδικασίας που ακολουθήθηκε στα Μαρκοβιανά μοντέλα.



Εικόνα 9.1: Διάγραμμα εκπαίδευσης και επαλήθευσης SVM.

Μέρος IV

Αποτελέσματα

Κεφάλαιο 10

Αποτελέσματα

10.1 Κρυφά Μαρκοβιανά Μοντέλα

Τα αποτελέσματα θα παρουσιαστούν επί το πλείστον σε γραφικές παραστάσεις, λόγω της τεράστιας πληροφορίας που υπάρχει σε μορφή πινάκων. Ωστόσο, θα παρουσιαστεί με σαφή αλλά σύντομο τρόπο η μετα-ανάλυση των δεδομένων αυτών. Η δομή του πίνακα αποτελεσμάτων είναι η ακόλουθη (πίνακας 10.1).

	Κατηγορία				145	146	Ακρίβεια
	1	2	.	.			
Ασθενής							
1	100	0	.	.	0	100	71.34%
2	0	0	.	.	100	100	45.63%
3	100	100	.	.	0	100	75.44%
.
.
27	0	100	.	.	0	100	64.38%
28	100	100	.	.	100	0	33.80%
.
.
Cross validation	62.96%	66.67%	.	.	59.26%	44.44%	57.62%
Ακρίβεια εκπαίδευσης	62.96%	88.89%	.	.	85.19%	92.59%	85.35%

Πίνακας 10.1: Πίνακας που δείχνει τη δομή των αποτελεσμάτων μετά το cross validation.

Υπάρχουν 6 τέτοιοι πίνακες από τους οποίους επιλέγονται τα καλύτερα αποτελέσματα αναγνώρισης μετά από το cross validation για την κάθε κυματομορφή. Με άλλα λόγια, επιλέγεται η καταλληλότερη μέθοδος εκπαίδευσης για κάθε κατηγορία κυματομορφών.

Στον παραπάνω πίνακα παρουσιάζονται στις γραμμές οι ασθενείς που αναγνωρίστηκαν αληθώς "100" ή ψευδώς "0" στις στήλες που αντιπροσωπεύουν και από μία κατηγορία κυματομορφών. Άρα, στο πάνω μέρος του πίνακα παρουσιάζονται οι επιτυχίες και τα σφάλματα αναγνώρισης του

μοντέλου για κάθε κυματομορφή και για κάθε ασθενή, μετά από την αξιολόγηση του leave-one-out. Στην αξιολόγηση αυτή, ο ασθενής που έμεινε εκτός εκπαίδευσης είναι και αυτός που φαίνεται στον πίνακα. Στο κάτω μέρος του πίνακα, βλέπουμε τις δύο διαφορετικές εκπαιδεύσεις του μοντέλου.

Στην περίπτωση που αναφερόμαστε στην ακρίβεια εκπαίδευσης πρακτικά βλέπουμε πόσο ακρίβες είναι το μοντέλο όταν του ζητούμε να αναγνωρίσει τα δεδομένα που εκπαιδεύτηκε. Όπως έχει ειπωθεί νωρίτερα, το μοντέλο όντας στατιστικό δεν έχει μνήμη για τις κυματομορφές εκπαίδευσης του, αυτό που κάνουν αυτές οι κυματομορφές είναι να του προσαρμόζουν τις παραμέτρους. Έτσι παρουσιάζονται αποκλίσεις από το 100% που θα περίμενε κάποιος σκεπτόμενος αφελώς, χωρίς να γνωρίζει τη φύση του μοντέλου. Στην δεύτερη περίπτωση που αναφερόμαστε σε *cross validation*, βλέπουμε το ποσοστό επιτυχίας του μοντέλου στην αναγνώριση των άγνωστων δεδομένων που του δώσαμε, αφήνοντας κυκλικά όλους τους ασθενείς για κάθε κατηγορία, εκτός διαδικασίας εκπαίδευσης. Είναι λογικό να παρατηρήσουμε ποσοστά επιτυχίας ίσα ή μικρότερα από εκείνα της αυτο-αξιολόγησης (ακρίβεια εκπαίδευσης) καθώς στη μία περίπτωση έχουμε το πλεονέκτημα της εκ των προτέρων "γνώσης" των δεδομένων που αναγνωρίζονται.

Τέλος, η τελευταία στήλη ορίζει το ποσοστό ακρίβειας για την συνολική κατηγοριοποίηση του ασθενή ως συμπτωματικό ή ασυμπτωματικό. Σύμφωνα με όλες τις κυματομορφές που έχουν εξαχθεί από τον ασθενή, λαμβάνουμε το ποσοστό επιτυχούς αναγνώρισης για τον κάθε ασθενή. Το αποτέλεσμα αυτό θα μπορούσε να υποδείξει την *αντιστροφή της υπάρχουσας γνώσης* για τη φύση του ασθενή, εφόσον κατατάσσουμε, α priori, τον ασθενή σε μια κατηγορία και το μοντέλο καταλήγει στο αντίθετο αποτέλεσμα. Σε επόμενη ενότητα θα εξηγηθεί γιατί αυτό το ποσοστό δεν είναι το βέλτιστο και μπορεί να βελτιωθεί αισθητά. Ακόμη με το ίδιο σκεπτικό και για όλες τις κατηγορίες φαίνεται στην τελευταία στήλη του κάτω μέρους, το ποσοστό αναγνώρισης συνολικά πλέον, για το cross validation και την ακρίβεια εκπαίδευσης.

Επιλέγοντας τα υψηλότερα ποσοστά για τις κυματομορφές μέσω cross validation καταλήγουμε στην επόμενη γραφική παράσταση (εικόνα 10.1). Οι γραμμές που ενώνουν τα ποσοστά δεν έχουν φυσική σημασία καθώς οι κυματομορφές δεν έχουν καμία αλληλεξάρτηση, αλλά προσφέρουν μία καλύτερη οπτική ως προς το μέγεθός τους. Τα ποσοστά αυτά παρουσιάζουν ως ελάχιστο το 60.47% και ως μέγιστο το 84.88%, ο μέσος όρος αναγνώρισης βρίσκεται στο 73.51%.

10.2 Μικτά αποτελέσματα και σύγκλιση

Στον πίνακα 10.2 που ακολουθεί θα παρουσιαστούν, ως παράδειγμα, σε οργανωμένη μορφή μερικές από τις κατηγορίες όπου δείχνουν τα ποσοστά αναγνώρισης για κάθε μοντέλο/μέθοδο και τη σχέση μεταξύ τους. Εδώ αξίζει να αναφερθεί πως η κύρια διαφορά στις τιμές προκύπτει από τα δύο διαφορετικά μοντέλα και όχι από την διαφορά ποσοστών αναγνώρισης άλλης μεθόδου στο ίδιο μοντέλο (svm scaled - svm unscaled).

Αντικείμενο	Δείκτης	HMM	SVM	Αστάθεια
Ολόκληρη πλάκα [μέση]	διαμήκης ταχύτητα	76.74%	82.19%	5.45
Ολόκληρη πλάκα [μέση]	γωνία ταχύτητας	79.07%	82.11%	3.04
Κάτω επιφάνεια πλάκας (PBS) [μέση]	διαμήκης μετατόπιση	77.91%	81.58%	13.42
Κάτω επιφάνεια πλάκας (PBS) [μέση]	ολική μετατόπιση	79.07%	85.26%	6.19
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	ακτινική μετατόπιση	70.93%	83.28%	12.75
Οπίσθιο τοίχωμα [διάμεση]	ακτινική μετατόπιση	80.23%	78.77%	1.84

Πίνακας 10.2: Παράδειγμα πίνακα μικτών αποτελεσμάτων όλων των μοντέλων/μεθόδων για μερικές από τις κατηγορίες. Τυχαίες τιμές.

Τα ποσοστά του πίνακα αναφέρονται στην επιτυχία αναγνώρισης (μέσω cross validation) για κάθε ένα από τα μοντέλα/μεθόδους, όπου το αντικείμενο σε συνδυασμό με τον δείκτη ορίζουν την κάθε κατηγορία. Η αστάθεια αναφέρεται στη μέγιστη διαφορά μεταξύ των ποσοστών επιτυχίας. Το ζητούμενο εδώ είναι να υπάρχουν ποσοστά αναγνώρισης μεγαλύτερα ή ίσα του 65% και ταυτόχρονα αναζητώντας την μικρότερη αστάθεια. Με αυτό τον τρόπο θα έχουμε μεγαλύτερη ασφάλεια ως προς τη δυνατότητα του μοντέλου, για συγκεκριμένες κατηγορίες και τα ποσοστά αναγνώρισης τους.

Στον πίνακα εμφανίζονται με έντονα γράμματα οι τιμές που δίνουν μια ισχυρή δυνατότητα να κατηγοριοποιούν τους ασθενείς σύμφωνα με τα πειραματικά δεδομένα. Άρα, προκύπτουν ερωτήματα σχετικά με τις κατηγορίες όπου τα ποσοστά τους είναι μεν υψηλά αλλά δεν παρουσιάζουν μεγάλη ομοιογένεια ως προς τα ποσοστά αυτά και στα 2 μοντέλα.

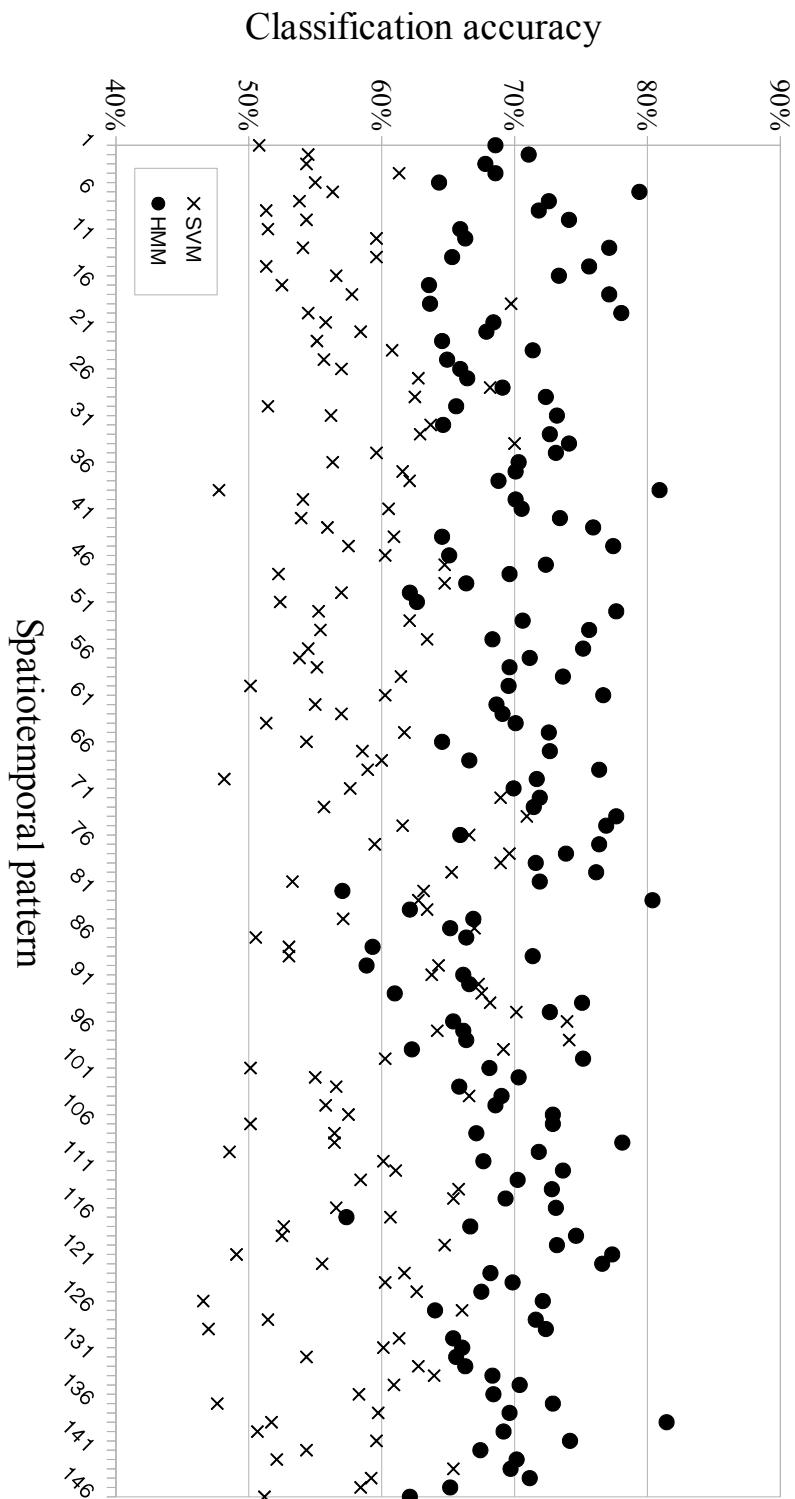
Ακολουθεί ο πίνακας 10.3, αποτελεσμάτων για όλα τα μοντέλα και όλες τις κατηγορίες κυματομορφών.

Αντικείμενο	Δείκτης	HMM	SVM	Ασάθεια
Ολόκληρη πλάκα [μέση]	διαμήκης ταχύτητα	68.56 %	50.78 %	17.78
Ολόκληρη πλάκα [μέση]	ακτινική ταχύτητα	71.06 %	54.47 %	16.59
Ολόκληρη πλάκα [μέση]	απόλυτη διαμήκης ταχύτητα	67.80 %	54.34 %	13.46
Ολόκληρη πλάκα [μέση]	απόλυτη ακτινική ταχύτητα	68.56 %	61.31 %	7.25
Ολόκληρη πλάκα [μέση]	ολική ταχύτητα	64.31 %	55.00 %	9.31
Ολόκληρη πλάκα [μέση]	γωνία ταχύτητας	79.39 %	56.31 %	23.08
Ολόκληρη πλάκα [διάμεση]	διαμήκης ταχύτητα	72.57 %	53.81 %	18.76
Ολόκληρη πλάκα [διάμεση]	ακτινική ταχύτητα	71.81 %	51.31 %	20.5
Ολόκληρη πλάκα [διάμεση]	απόλυτη διαμήκης ταχύτητα	74.09 %	54.34 %	19.75
Ολόκληρη πλάκα [διάμεση]	απόλυτη ακτινική ταχύτητα	65.90 %	51.44 %	14.46
Ολόκληρη πλάκα [διάμεση]	ολική ταχύτητα	66.28 %	59.60 %	6.68
Ολόκληρη πλάκα [διάμεση]	γωνία ταχύτητας	77.12 %	54.07 %	23.05
Ολόκληρη πλάκα [μέση]	διαμήκης μετατόπιση	65.30 %	59.60 %	5.7
Ολόκληρη πλάκα [μέση]	ακτινική μετατόπιση	75.60 %	51.31 %	24.29
Ολόκληρη πλάκα [μέση]	απόλυτη διαμήκης μετατόπιση	73.33 %	56.57 %	16.76
Ολόκληρη πλάκα [μέση]	απόλυτη ακτινική μετατόπιση	63.56 %	52.5 %	11.06
Ολόκληρη πλάκα [μέση]	ολική μετατόπιση	77.12 %	57.76 %	19.36
Ολόκληρη πλάκα [μέση]	γωνία μετατόπισης	63.63 %	69.73 %	6.1
Ολόκληρη πλάκα [διάμεση]	διαμήκης μετατόπιση	78.03 %	54.47 %	23.56
Ολόκληρη πλάκα [διάμεση]	ακτινική μετατόπιση	68.40 %	55.78 %	12.62
Ολόκληρη πλάκα [διάμεση]	απόλυτη διαμήκης μετατόπιση	67.87 %	58.42 %	9.45
Ολόκληρη πλάκα [διάμεση]	απόλυτη ακτινική μετατόπιση	64.54 %	55.13 %	9.41
Ολόκληρη πλάκα [διάμεση]	ολική μετατόπιση	71.36 %	60.78 %	10.58
Ολόκληρη πλάκα [διάμεση]	γωνία μετατόπισης	64.92 %	55.65 %	9.27
Σχετική κίνηση μέσα στην πλάκα [μέση]	διαμήκης παραμόρφωση	65.90 %	56.97 %	8.93
Σχετική κίνηση μέσα στην πλάκα [διάμεση]	διαμήκης παραμόρφωση	66.43 %	62.76 %	3.67
Σχετική κίνηση μέσα στην πλάκα [μέση]	ακτινική παραμόρφωση	69.09 %	68.15 %	0.94
Σχετική κίνηση μέσα στην πλάκα [διάμεση]	ακτινική παραμόρφωση	72.34 %	62.50 %	9.84
Σχετική κίνηση μέσα στην πλάκα [μέση]	ακτινική διατμητική παραμόρφωση	65.60 %	51.44 %	14.16
Σχετική κίνηση μέσα στην πλάκα [διάμεση]	ακτινική διατμητική παραμόρφωση	73.18 %	56.18 %	17
Σχετική κίνηση μέσα στην πλάκα [μέση]	διαμήκης διατμητική παραμόρφωση	64.62 %	63.68 %	0.94
Σχετική κίνηση μέσα στην πλάκα [διάμεση]	διαμήκης διατμητική παραμόρφωση	72.65 %	62.89 %	9.76
Άνω επιφάνεια πλάκας (PTS) [μέση]	διαμήκης ταχύτητα	74.09 %	70.00 %	4.09
Άνω επιφάνεια πλάκας (PTS) [μέση]	ακτινική ταχύτητα	73.10 %	59.60 %	13.5
Άνω επιφάνεια πλάκας (PTS) [μέση]	απόλυτη διαμήκης ταχύτητα	70.30 %	56.31 %	13.99
Άνω επιφάνεια πλάκας (PTS) [μέση]	απόλυτη ακτινική ταχύτητα	70.07 %	61.57 %	8.5
Άνω επιφάνεια πλάκας (PTS) [μέση]	ολική ταχύτητα	68.78 %	62.10 %	6.68
Άνω επιφάνεια πλάκας (PTS) [μέση]	γωνία ταχύτητας	80.90 %	47.76 %	33.14
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	διαμήκης ταχύτητα	70.07 %	54.07 %	16
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	ακτινική ταχύτητα	70.53 %	60.52 %	10.01
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	απόλυτη διαμήκης ταχύτητα	73.40 %	53.94 %	19.46
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	απόλυτη ακτινική ταχύτητα	75.90 %	55.92 %	19.98
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	ολική ταχύτητα	64.54 %	60.92 %	3.62
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	γωνία ταχύτητας	77.42 %	57.50 %	19.92
Άνω επιφάνεια πλάκας (PTS) [μέση]	διαμήκης μετατόπιση	65.07 %	60.26 %	4.81
Άνω επιφάνεια πλάκας (PTS) [μέση]	ακτινική μετατόπιση	72.34 %	64.73 %	7.61
Άνω επιφάνεια πλάκας (PTS) [μέση]	απόλυτη διαμήκης μετατόπιση	69.62 %	52.23 %	17.39
Άνω επιφάνεια πλάκας (PTS) [μέση]	απόλυτη ακτινική μετατόπιση	66.36 %	64.73 %	1.63
Άνω επιφάνεια πλάκας (PTS) [μέση]	ολική μετατόπιση	62.12 %	56.97 %	5.15
Άνω επιφάνεια πλάκας (PTS) [μέση]	γωνία μετατόπισης	62.65 %	52.36 %	10.29
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	διαμήκης μετατόπιση	77.65 %	55.26 %	22.39
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	ακτινική μετατόπιση	70.60 %	62.10 %	8.5
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	απόλυτη διαμήκης μετατόπιση	75.60 %	55.39 %	20.21
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	απόλυτη ακτινική μετατόπιση	68.33 %	63.42 %	4.91
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	ολική μετατόπιση	75.15 %	54.47 %	20.68
Άνω επιφάνεια πλάκας (PTS) [διάμεση]	γωνία μετατόπισης	71.13 %	53.81 %	17.32
Κάτω επιφάνεια πλάκας (PBS) [μέση]	διαμήκης ταχύτητα	69.62 %	55.13 %	14.49
Κάτω επιφάνεια πλάκας (PBS) [μέση]	ακτινική ταχύτητα	73.63 %	61.44 %	12.19
Κάτω επιφάνεια πλάκας (PBS) [μέση]	απόλυτη διαμήκης ταχύτητα	69.54 %	50.13 %	19.41
Κάτω επιφάνεια πλάκας (PBS) [μέση]	απόλυτη ακτινική ταχύτητα	76.66 %	60.26 %	16.4
Κάτω επιφάνεια πλάκας (PBS) [μέση]	ολική ταχύτητα	68.63 %	55.00 %	13.63
Κάτω επιφάνεια πλάκας (PBS) [μέση]	γωνία ταχύτητας	69.09 %	56.97 %	12.12
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	διαμήκης ταχύτητα	70.07 %	51.31 %	18.76
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	ακτινική ταχύτητα	72.57 %	61.71 %	10.86
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	απόλυτη διαμήκης ταχύτητα	64.54 %	54.34 %	10.2
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	απόλυτη ακτινική ταχύτητα	72.65 %	58.55 %	14.1
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	ολική ταχύτητα	66.59 %	60.00 %	6.59
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	γωνία ταχύτητας	76.36 %	58.94 %	17.42
Κάτω επιφάνεια πλάκας (PBS) [μέση]	διαμήκης μετατόπιση	71.66 %	48.15 %	23.51
Κάτω επιφάνεια πλάκας (PBS) [μέση]	ακτινική μετατόπιση	69.92 %	57.63 %	12.29
Κάτω επιφάνεια πλάκας (PBS) [μέση]	απόλυτη διαμήκης μετατόπιση	71.89 %	68.94 %	2.95
Κάτω επιφάνεια πλάκας (PBS) [μέση]	απόλυτη ακτινική μετατόπιση	71.43 %	55.65 %	15.78
Κάτω επιφάνεια πλάκας (PBS) [μέση]	ολική μετατόπιση	77.65 %	70.92 %	6.73
Κάτω επιφάνεια πλάκας (PBS) [μέση]	γωνία μετατόπισης	76.89 %	61.57 %	15.32
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	διαμήκης μετατόπιση	65.90 %	66.57 %	0.67
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	ακτινική μετατόπιση	76.36 %	59.47 %	16.89

Αντικείμενο	Δείκτης	HMM	SVM	Ασάθια
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	απόλυτη διαμήκης μετατόπιση	73.86 %	69.60 %	4.26
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	απόλυτη ακτινική μετατόπιση	71.59 %	68.94 %	2.65
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	ολική μετατόπιση	76.13 %	65.26 %	10.87
Κάτω επιφάνεια πλάκας (PBS) [διάμεση]	γωνία μετατόπισης	71.89 %	53.28 %	18.61
Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [μέση]	διαμήκης παραμόρφωση PTS/Υγιής	57.04 %	63.15 %	6.11
Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [διάμεση]	διαμήκης παραμόρφωση PTS/Υγιής	80.37 %	62.76 %	17.61
Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [μέση]	διαμήκης παραμόρφωση PBS/Υγιής	62.12 %	63.42 %	1.3
Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [διάμεση]	διαμήκης παραμόρφωση PBS/Υγιής	66.89 %	57.1 %	9.79
Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [μέση]	ακτινική παραμόρφωση PTS/Υγιής	65.15 %	66.97 %	1.82
Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [διάμεση]	ακτινική παραμόρφωση PTS/Υγιής	66.36 %	50.52 %	15.84
Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [μέση]	ακτινική παραμόρφωση PBS/Υγιής	59.31 %	53.02 %	6.29
Σχετ. κίν. πλάκας κ υγιούς τοιχ. πλησίον αυτής [διάμεση]	ακτινική παραμόρφωση PBS/Υγιής	71.36 %	53.02 %	18.34
Μέρη του υγιούς τοιχώματος [μέση]	διαμήκης παραμόρφωση, οπίσθιο τοίχωμα	58.86 %	64.27 %	5.41
Μέρη του υγιούς τοιχώματος [διάμεση]	διαμήκης παραμόρφωση, οπίσθιο τοίχωμα	66.13 %	63.76 %	2.37
Μέρη του υγιούς τοιχώματος [μέση]	διαμήκης παραμόρφωση, πρόσθιο τοίχωμα	66.59 %	67.28 %	0.69
Μέρη του υγιούς τοιχώματος [διάμεση]	διαμήκης παραμόρφωση, πρόσθιο τοίχωμα	60.98 %	67.52 %	6.54
Μέρη του υγιούς τοιχώματος [μέση]	ακτινική παραμόρφωση οπίσθιο/πρόσθιο τοίχωμα	75.07 %	68.15 %	6.92
Μέρη του υγιούς τοιχώματος [διάμεση]	ακτινική παραμόρφωση οπίσθιο/πρόσθιο τοίχωμα	72.65 %	70.13 %	2.52
Μέρη του υγιούς τοιχώματος [μέση]	διαμήκης διατμητική παραμόρφωση, οπίσθιο τοίχ.	65.37 %	73.94 %	8.57
Μέρη του υγιούς τοιχώματος [διάμεση]	διαμήκης διατμητική παραμόρφωση, οπίσθιο τοίχ.	66.13 %	64.18 %	1.95
Μέρη του υγιούς τοιχώματος [μέση]	διαμήκης διατμητική παραμόρφωση, πρόσθιο τοίχ.	66.36 %	74.10 %	7.74
Μέρη του υγιούς τοιχώματος [διάμεση]	διαμήκης διατμητική παραμόρφωση, πρόσθιο τοίχ.	62.27 %	69.15 %	6.88
Οπίσθιο τοίχωμα [μέση]	διαμήκης ταχύτητα	75.15 %	60.26 %	14.89
Οπίσθιο τοίχωμα [μέση]	ακτινική ταχύτητα	68.10 %	50.13 %	17.97
Οπίσθιο τοίχωμα [μέση]	απόλυτη διαμήκης ταχύτητα	70.30 %	55.00 %	15.3
Οπίσθιο τοίχωμα [μέση]	απόλυτη ακτινική ταχύτητα	65.83 %	56.57 %	9.26
Οπίσθιο τοίχωμα [μέση]	ολική ταχύτητα	69.01 %	66.57 %	2.44
Οπίσθιο τοίχωμα [μέση]	γωνία ταχύτητας	68.56 %	55.78 %	12.78
Οπίσθιο τοίχωμα [διάμεση]	διαμήκης ταχύτητα	72.87 %	57.5 %	15.37
Οπίσθιο τοίχωμα [διάμεση]	ακτινική ταχύτητα	72.87 %	50.13 %	22.74
Οπίσθιο τοίχωμα [διάμεση]	απόλυτη διαμήκης ταχύτητα	67.12 %	56.44 %	10.68
Οπίσθιο τοίχωμα [διάμεση]	απόλυτη ακτινική ταχύτητα	78.10 %	56.44 %	21.66
Οπίσθιο τοίχωμα [διάμεση]	ολική ταχύτητα	71.81 %	48.55 %	23.26
Οπίσθιο τοίχωμα [διάμεση]	γωνία ταχύτητας	67.65 %	60.13 %	7.52
Οπίσθιο τοίχωμα [μέση]	διαμήκης μετατόπιση	73.63 %	61.05 %	12.58
Οπίσθιο τοίχωμα [μέση]	ακτινική μετατόπιση	70.22 %	58.42 %	11.8
Οπίσθιο τοίχωμα [μέση]	απόλυτη διαμήκης μετατόπιση	72.80 %	65.78 %	7.02
Οπίσθιο τοίχωμα [μέση]	απόλυτη ακτινική μετατόπιση	69.31 %	65.39 %	3.92
Οπίσθιο τοίχωμα [μέση]	ολική μετατόπιση	73.10 %	56.57 %	16.53
Οπίσθιο τοίχωμα [μέση]	γωνία μετατόπισης	57.34 %	60.65 %	3.31
Οπίσθιο τοίχωμα [διάμεση]	διαμήκης μετατόπιση	66.66 %	52.63 %	14.03
Οπίσθιο τοίχωμα [διάμεση]	ακτινική μετατόπιση	74.62 %	52.5 %	22.12
Οπίσθιο τοίχωμα [διάμεση]	απόλυτη διαμήκης μετατόπιση	73.18 %	64.73 %	8.45
Οπίσθιο τοίχωμα [διάμεση]	απόλυτη ακτινική μετατόπιση	77.34 %	49.07 %	28.27
Οπίσθιο τοίχωμα [διάμεση]	ολική μετατόπιση	76.59 %	55.52 %	21.07
Οπίσθιο τοίχωμα [διάμεση]	γωνία μετατόπισης	68.18 %	61.71 %	6.47
Πρόσθιο τοίχωμα [μέση]	διαμήκης ταχύτητα	69.84 %	60.26 %	9.58
Πρόσθιο τοίχωμα [μέση]	ακτινική ταχύτητα	67.50 %	62.63 %	4.87
Πρόσθιο τοίχωμα [μέση]	απόλυτη διαμήκης ταχύτητα	72.12 %	46.57 %	25.55
Πρόσθιο τοίχωμα [μέση]	απόλυτη ακτινική ταχύτητα	64.01 %	66.05 %	2.04
Πρόσθιο τοίχωμα [μέση]	ολική ταχύτητα	71.59 %	51.44 %	20.15
Πρόσθιο τοίχωμα [μέση]	γωνία ταχύτητας	72.34 %	46.97 %	25.37
Πρόσθιο τοίχωμα [διάμεση]	διαμήκης ταχύτητα	65.37 %	61.31 %	4.06
Πρόσθιο τοίχωμα [διάμεση]	ακτινική ταχύτητα	66.06 %	60.13 %	5.93
Πρόσθιο τοίχωμα [διάμεση]	απόλυτη διαμήκης ταχύτητα	65.60 %	54.34 %	11.26
Πρόσθιο τοίχωμα [διάμεση]	απόλυτη ακτινική ταχύτητα	66.28 %	62.76 %	3.52
Πρόσθιο τοίχωμα [διάμεση]	ολική ταχύτητα	68.33 %	63.94 %	4.39
Πρόσθιο τοίχωμα [διάμεση]	γωνία ταχύτητας	70.37 %	60.92 %	9.45
Πρόσθιο τοίχωμα [μέση]	διαμήκης μετατόπιση	68.40 %	58.28 %	10.12
Πρόσθιο τοίχωμα [μέση]	ακτινική μετατόπιση	72.87 %	47.63 %	25.24
Πρόσθιο τοίχωμα [μέση]	απόλυτη διαμήκης μετατόπιση	69.62 %	59.73 %	9.89
Πρόσθιο τοίχωμα [μέση]	απόλυτη ακτινική μετατόπιση	81.43 %	51.71 %	29.72
Πρόσθιο τοίχωμα [μέση]	ολική μετατόπιση	69.16 %	50.65 %	18.51
Πρόσθιο τοίχωμα [μέση]	γωνία μετατόπισης	74.16 %	59.6 %	14.56
Πρόσθιο τοίχωμα [διάμεση]	διαμήκης μετατόπιση	67.42 %	54.34 %	13.08
Πρόσθιο τοίχωμα [διάμεση]	ακτινική μετατόπιση	70.15 %	52.1 %	18.05
Πρόσθιο τοίχωμα [διάμεση]	απόλυτη διαμήκης μετατόπιση	69.69 %	65.39 %	4.3
Πρόσθιο τοίχωμα [διάμεση]	απόλυτη ακτινική μετατόπιση	71.13 %	59.21 %	11.92
Πρόσθιο τοίχωμα [διάμεση]	ολική μετατόπιση	65.15 %	58.42 %	6.73
Πρόσθιο τοίχωμα [διάμεση]	γωνία μετατόπισης	62.12 %	51.18 %	10.94
Μέσος όρος		69.95%	58.57%	12.16

Πίνακας 10.3: Πίνακας μικτών αποτελεσμάτων όλων των μοντέλων για όλες τις κατηγορίες.

10.3 Σύγκριση απόδοσης των δύο μοντέλων



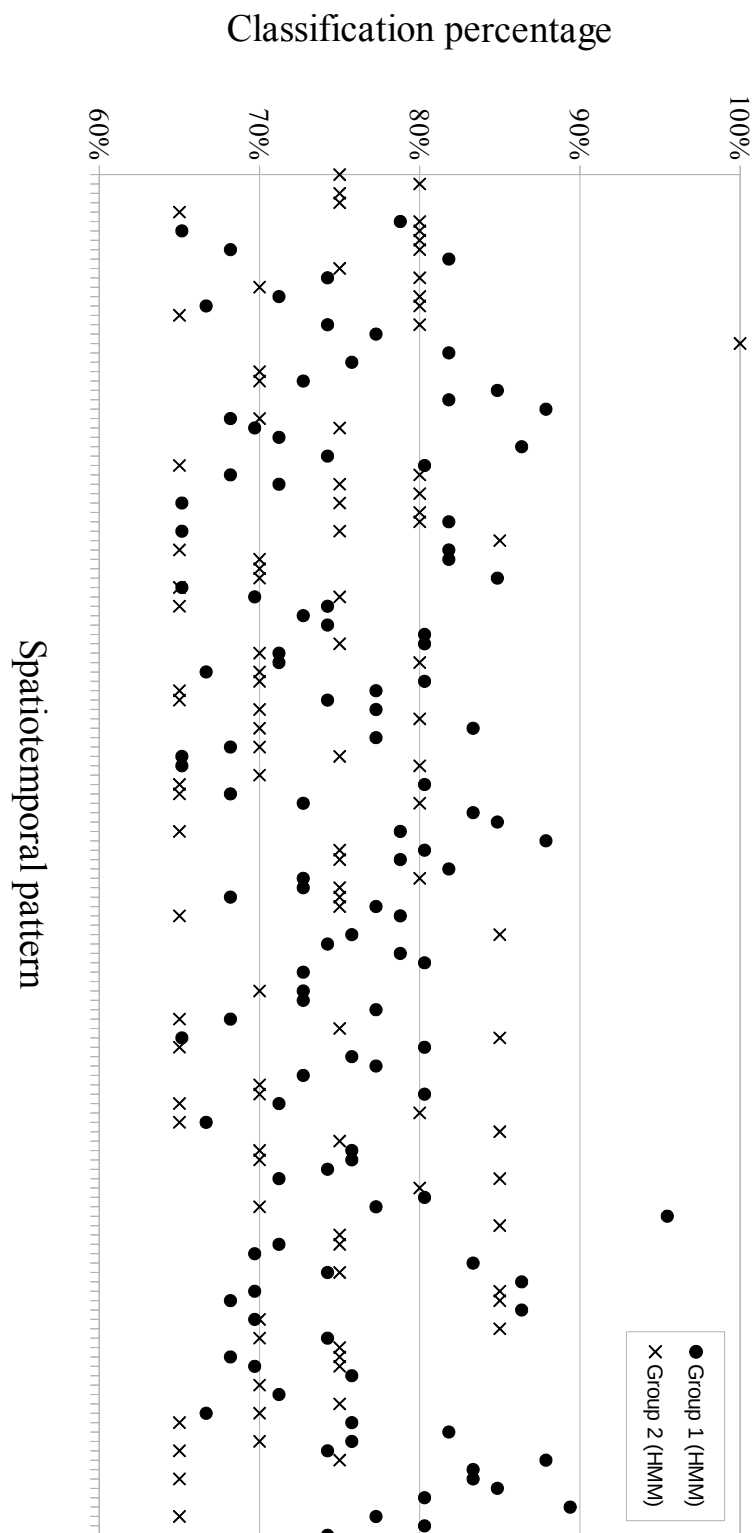
Εικόνα 10.1: Διάγραμμα που δίνει τα αποτελέσματα του cross validation για κάθε κατηγορία. Όπου HMM είναι το Κρυφό Μαρκοβιανό Μοντέλο και όπου SVM οι Μηχανές Διανυσμάτων Υποστήριξης.

Το παραπάνω διάγραμμα (εικόνα 10.1) έχοντας σκοπό την σύγκριση των δύο μοντέλων, έχει συγχωνεύσει τις δύο περιπτώσεις των SVM (με ή χωρίς κλιμάκωση) σε μία, όπου και περιέχει τα μεγαλύτερα ποσοστά από τις δύο. Είναι φανερό πως τα Κρυφά Μαρκοβιανά Μοντέλα συνολικά υπερτερούν των Support Vector Machines στη συγκεκριμένη εφαρμογή. Η απόδοση των πρώτων ως προς το ποσοστό αναγνώρισης κυμαίνεται περί το 69.95% ενώ των δεύτερων στα 58.57%.

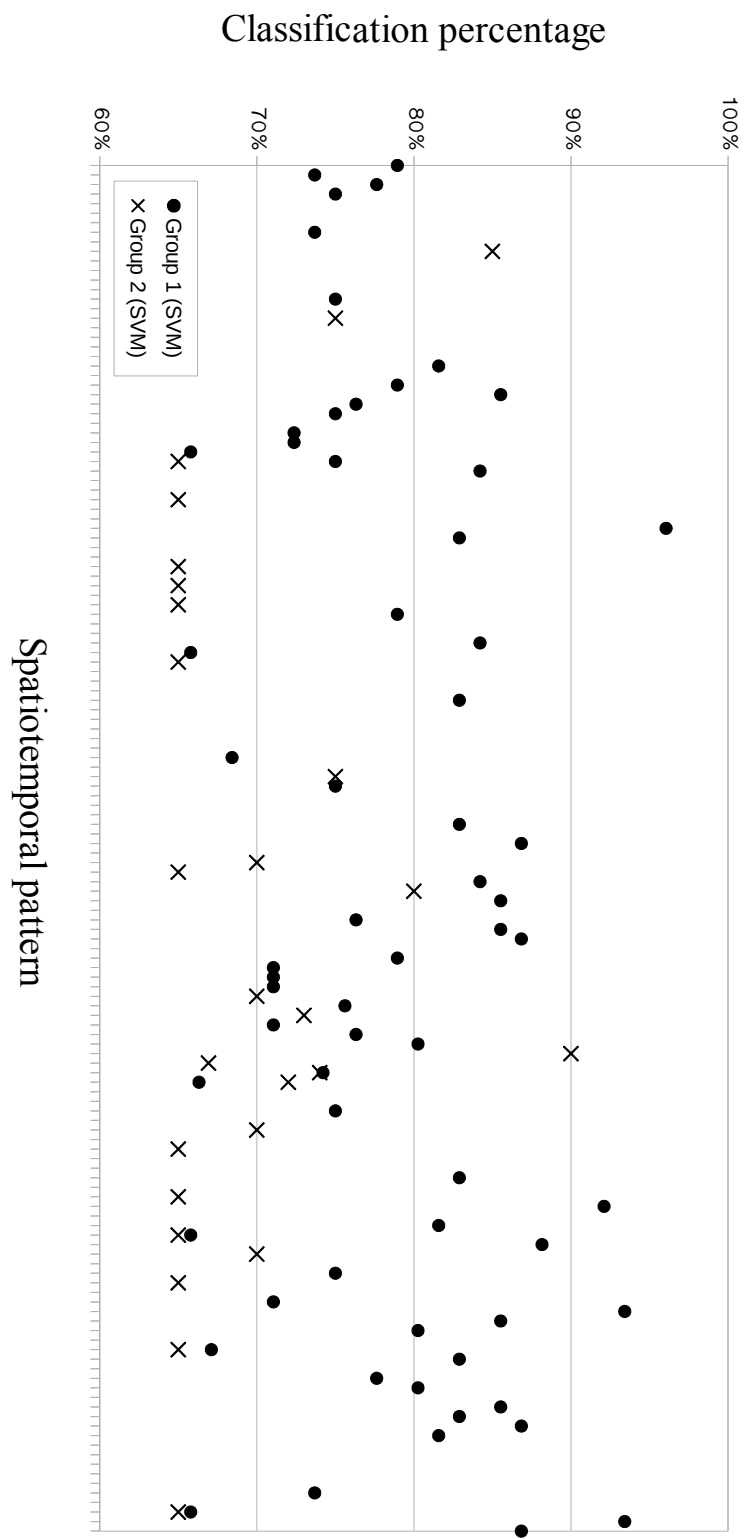
10.4 Σύγκριση απόδοσης των δύο ομάδων ασθενών

Ιδιαίτερο ενδιαφέρον έχουν τα επόμενα διαγράμματα που δείχνουν και για τα δύο μοντέλα τα αποτελέσματα για την ακρίβεια της κάθε κατηγορίας κυματομορφών αλλά με βάση τις ομάδες (group) των ασθενών (συμπτωματικοί/ασυμπτωματικοί). Έτσι θα μπορέσουμε να δούμε ποιες κατηγορίες κυματομορφών είναι καταλληλότερες για τη μία ομάδα και ποιες για την άλλη. Ακόμη, θα δούμε το συνολικό επίπεδο ακρίβειας για κάθε ομάδα ασθενών στα μοντέλα αλλά και το βέλτιστο επίπεδο αναγνώρισης που θα προκύψει από την επιλογή των "δυνατότερων" κατηγοριών για κάθε ομάδα.

Τα διαγράμματα που ακολουθούν (εικόνες 10.2, 10.3) παρουσιάζουν τα ποσοστά αναγνώρισης με τον τρόπο που παρουσιάστηκαν και νωρίτερα αλλά για κάθε διάγραμμα υπάρχουν 2 ομάδες ασθενών. Αυτό γίνεται και για τα δύο μοντέλα.



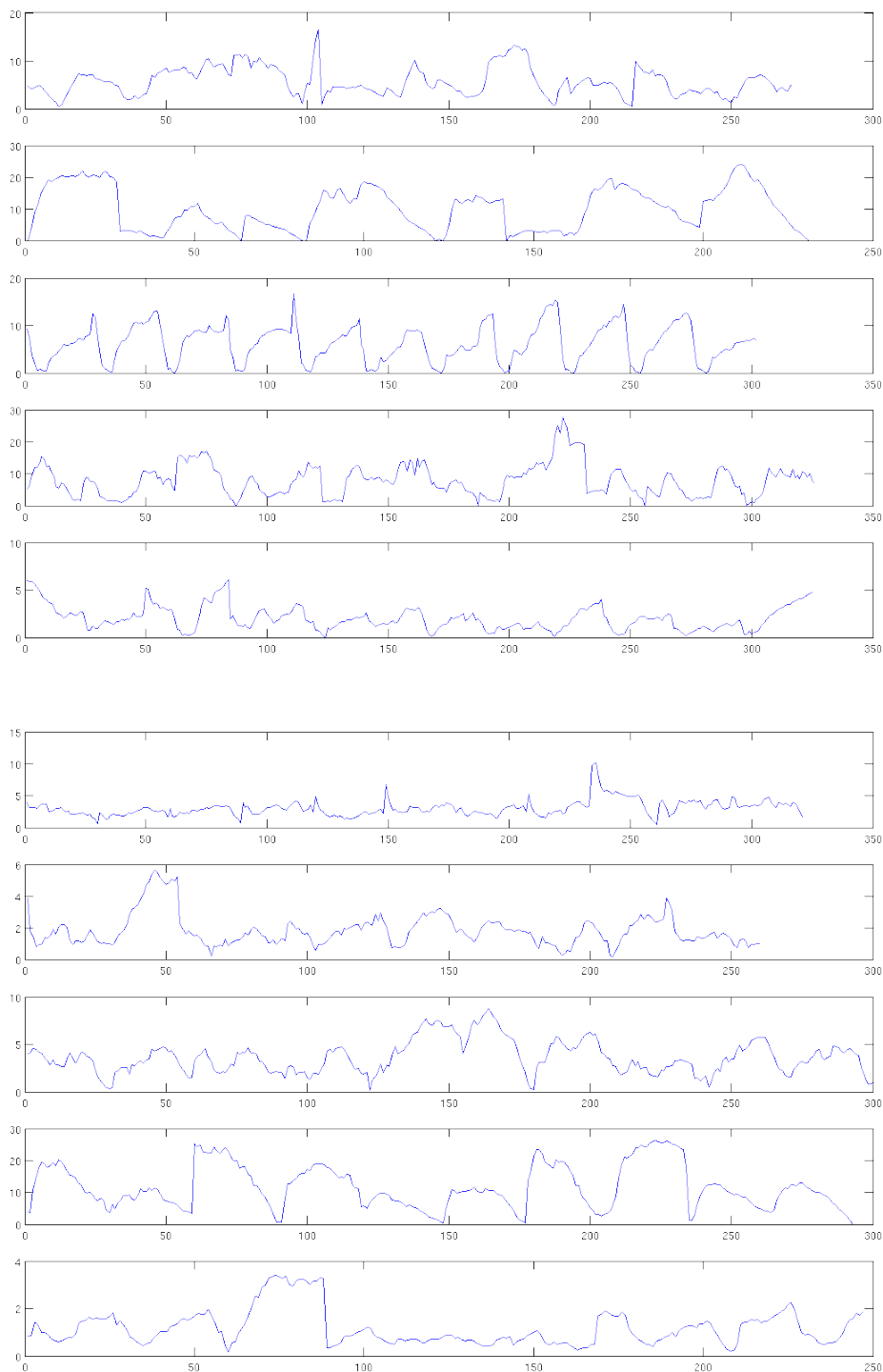
Εικόνα 10.2: Ακρίβεια HMM για τις 2 ομάδες ασθενών, για επιλεγμένες κατηγορίες.



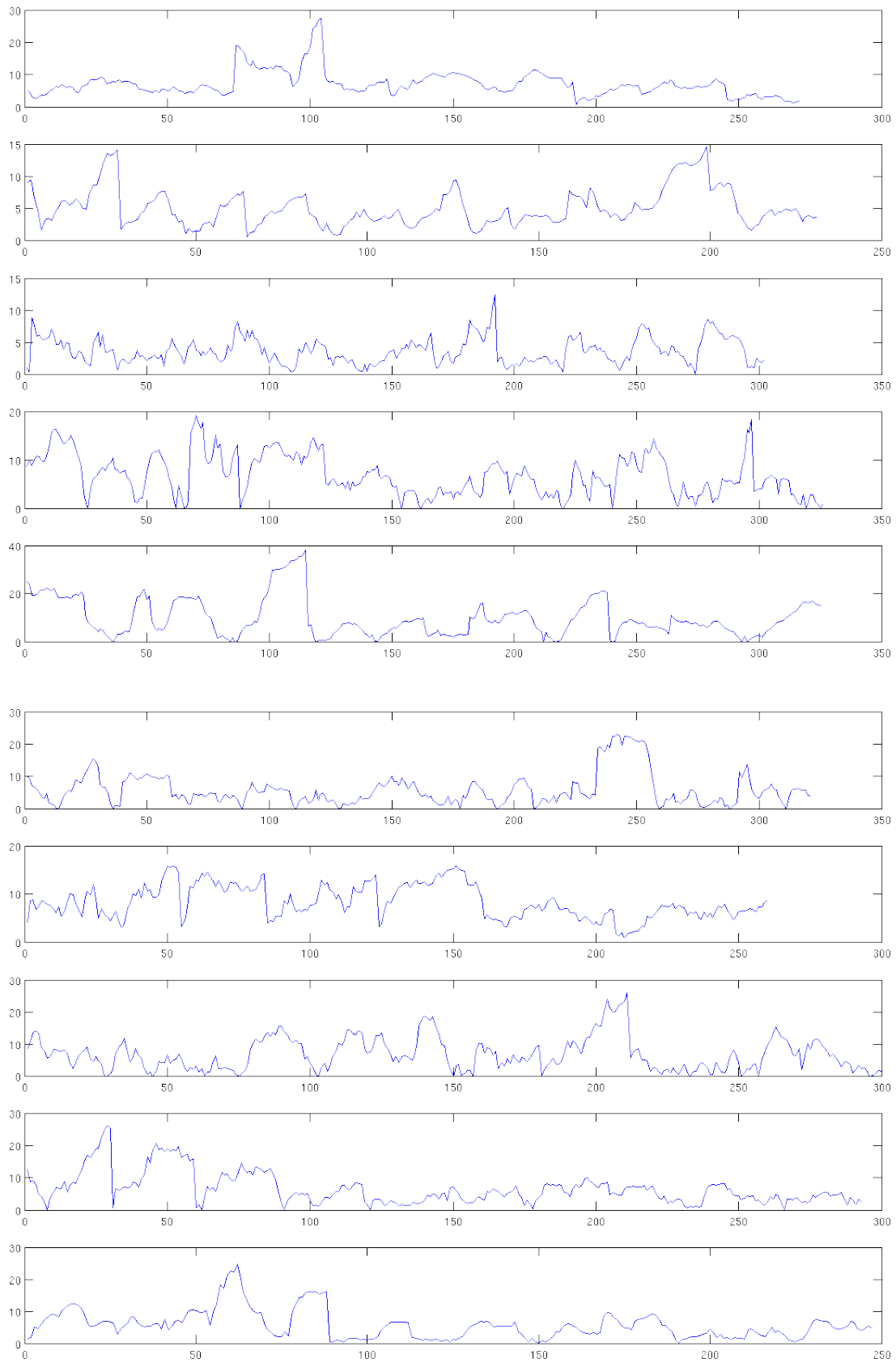
Εικόνα 10.3: Ακρίβεια SVM για τις 2 ομάδες ασθενών, για επιλεγμένες κατηγορίες.

10.5 Οπτική επιθεώρηση σε απόλυτα κατηγοριοποιημένες κυματομορφές

Η οπτική επιθεώρηση για χαρακτηριστικά ομοιότητας σε κυματομορφές που έχουν κατηγοριοποιηθεί με υψηλή ακρίβεια από το μοντέλο, είναι αδύνατη (εικόνες 10.4, 10.5).



Εικόνα 10.4: Κυματομορφές της κατηγορίας 89 για τις 2 ομάδες ασθενών. Πάνω, 5 ασθενείς της ομάδας 1 και κάτω 5 ασθενείς της ομάδας 2.



Εικόνα 10.5: Κυματομορφές της κατηγορίας 90 για τις 2 ομάδες ασθενών. Πάνω, 5 ασθενείς της ομάδας 1 και κάτω 5 ασθενείς της ομάδας 2.

Μέρος V

Συζήτηση και μελλοντικές προοπτικές

Κεφάλαιο 11

Συζήτηση

Στο κεφάλαιο αυτό θα γίνει ο σχολιασμός των αποτελεσμάτων του προηγούμενου κεφαλαίου αλλά και θα παρουσιαστεί η συλλογιστική πορεία που οδήγησε στην κατασκευή τους.

11.1 Αξιολόγηση Κρυφών Μαρκοβιανών Μοντέλων

Τα Κρυφά Μαρκοβιανά μοντέλα δείχνουν σε βασικές γραμμές ότι μπορούν να προσεγγίσουν τα δεδομένα εκπαίδευσης σε αρκετά καλό επίπεδο. Ωστόσο, για να θεωρηθεί κάποια κατηγορία *αξιόπιστη* θα θέλαμε ποσοστά αναγνώρισης πάνω από 65%, αυτό παρουσιάζεται σε 116/45 κατηγορίες (group1/group2). Άρα, για κάθε κατηγορία με υψηλά ποσοστά (της τάξεως 70-80%) θα πρέπει να βρεθεί κάποιος τρόπος για να μας δοθεί μια ακόμη απόδειξη ότι αυτά τα ποσοστά είναι αξιόπιστα. Αυτό θα γίνει με μια τελειώς διαφορετική μέθοδο αναγνώρισης προτύπων με σκοπό την *παράτηρηση συγκλίσεων μεταξύ των δύο μοντέλων*.

Ακόμη, αναφέραμε την τελευταία στήλη των αποτελεσμάτων (πίνακας 10.1) ως ένα μέγεθος αμφιλεγόμενο, καθώς δέχεται την κάθε κατηγορία ομότιμα, ασχέτως εάν αυτή η κατηγορία παρουσιάζει πολύ χαμηλά ποσοστά αναγνώρισης. Εδώ προκύπτουν δύο θέματα, το ένα έχει να κάνει με την αξιολόγηση της κυματομορφής ως προς το περιεχόμενό της, με την έννοια του ότι δεν αποδίδει σαν κατηγορία, δεν σχετίζεται δηλαδή με το αποτέλεσμα. Ενώ το δεύτερο θέμα έχει να κάνει με το κατά πόσο μπορεί το μοντέλο να κατανοήσει την κυματομορφή και να εκπαιδευτεί από τα δεδομένα της. Και στις δύο περιπτώσεις η κυματομορφές αυτές θα μπορούσαν να χαρακτηριστούν ως ακατάλληλες και ως εκ τούτου μη χρήσιμες για την συνολική ακρίβεια του μοντέλου.

Μέσα από την ανάλυση αυτή αφενός μπορούμε να αποκλείσουμε κάποιες κατηγορίες, αφετέρου να βελτιστοποιήσουμε την συνολική πιθανότητα αναγνώρισης του μοντέλου, για τον κάθε ασθενή, αγνοώντας τις ακατάλληλες κατηγορίες. Στη συνέχεια θα γίνει φανερό ποια ποσοστά αναγνώρισης κατηγοριών είναι αξιόπιστα και ποια όχι.

11.2 Αξιολόγηση μικτών αποτελεσμάτων

Η εύρεση των μικτών αποτελεσμάτων μας βοηθάει να συγκρίνουμε τα δύο μοντέλα όχι τόσο ως προς την απόδοση αλλά ως προς τη σύγκλιση, πράγμα που είναι το ζητούμενο, διότι ακόμη και η πληροφορία (εκ των προτέρων γνώση) για την κατάσταση του κάθε ασθενή (συμπτωματικός ή μη), δεν μπορεί να εγγυηθεί την ακρίβεια της διάγνωσης (επαλήθευσης) μέσω κάποιας συγκεκριμένης κατηγορίας. Άρα, ο σκοπός της σύγκρισης έχει να κάνει με το ποσοστό εμπιστοσύνης που θα δώσουμε στην κάθε κατηγορία κυματομορφών ξεχωριστά, σε σχέση με τον ρόλο της ως παράμετρο (δείκτη) για το πρόβλημα της κατηγοριοποίησης των ασθενών.

Αυθαίρετα αλλά όχι καταχρηστικά δόθηκε μεγάλη βαρύτητα στα αποτελέσματα των κατηγοριών που παρουσιάζουν μεταξύ τους διαφορές στα ποσοστά αναγνώρισης, μικρότερη του 5%. Έτσι μπορέσαμε μέσα από το σύνολο των 146 κατηγοριών να διακρίνουμε τις σταθερές (stable) κατηγορίες, τα ποσοστά των οποίων είναι σχεδόν ταυτόσημα αλλά έχουν προκύψει όπως περιγράφηκε από 3 διαφορετικές μεθόδους (2 διαφορετικά μοντέλα). Φυσικά δεν έχει μεγάλη αξία η ταύτιση χαμηλών ποσοστών αναγνώρισης αλλά η ταύτιση των υψηλότερων. Με αυτό τον τρόπο δείχνουμε πως υπάρχουν στο δείγμα κατηγορίες που εν μέρει σε ένα μοντέλο αποδίδουν, αλλά σε ένα άλλο υπολείπονται.

Κοιτάζοντας το συνδυαστικό διάγραμμα αποτελεσμάτων (εικόνα 10.1) μπορούμε να εξάγουμε το εξής συμπέρασμα. Υπάρχουν σημαντικές συγκλίσεις μεταξύ των Κρυφών Μαρκοβιανών Μοντέλων και των Support Vector Machines ως προς την κατηγοριοποίηση των ασθενών. Πράγμα που μας κάνει να θεωρούμε ότι για αρκετές κυματομορφές (κατηγορίες) υπάρχει ένδειξη και μάλιστα ισχυρή ως προς τη συσχέτιση τους με την συμπτωματική ή ασυμπτωματική φύση του ασθενή.

Ακόμη βλέπουμε με τη βοήθεια μιας καμπύλης (η οποία πάλι δεν έχει φυσική σημασία) τα μέγιστα αποτελέσματα αναγνώρισης για την κάθε κατηγορία και έτσι έχουμε μια διπλή άποψη σχετικά με τα αποτελέσματα. Από τη μία πλευρά παρατηρούμε τις μέγιστες τιμές ως ένα μέτρο άνω ορίου αλλά από την άλλη αξιολογούμε αυτό το όριο ώστε να χαρακτηρίσουμε τα αποτελέσματα για κάποιες κατηγορίες ανεπαρκή ή τυχαία. Στην περίπτωση που το άνω όριο βρίσκεται σε ποσοστό 50-60% είναι φυσιολογικό να θεωρήσουμε την συγκεκριμένη κατηγορία ως ακατάλληλη. Αυτό μπορεί να σημαίνει είτε πως τα 2 μοντέλα δεν μπορούν να καλύψουν τα συγκεκριμένα πρότυπα μέσα από την εκπαίδευση, είτε πως δεν υπάρχει ουσιαστικό περιεχόμενο σε σχέση με αυτή τη κατηγορία. Όσο μεγαλώνει το δείγμα εκπαίδευσης τόσο πιο εύκολο γίνεται να ξεχωρίσουμε τα τυχαία δεδομένα (κατηγορίες) από το ποσοστό αναγνώρισης που θα προκύψει. Εάν κάποιο χαρακτηριστικό (κατηγορία) ανήκει στην παραπάνω κατηγορία, σύμφωνα με τη θεωρία των μεγάλων αριθμών το ποσοστό αναγνώρισης θα πλησιάσει το 50%.

Άρα μια μέθοδος είναι να κοιτάξουμε την τάση (παράγωγο) που έχουν τα ποσοστά των κατηγοριών, καθώς αυξάνουμε το δείγμα και αν αυτή η τάση παρουσιάζει μια σταθερή τάση προς το 50%

μιλάμε για τυχαίο σήμα και έτσι το αποσυνδέουμε από τη μελέτη ως ανεπαρκές. Σε διαφορετική τάση, με την αύξηση του δείγματος εκπαίδευσης, μπορούμε να παρατηρήσουμε αύξηση ποσοστών και έτσι να θεωρήσουμε λογικά πως αυτή η τάση συνδέεται άμεσα με τη μελέτη μας.

Ακόμη βλέπουμε πως οι περισσότερες κατηγορίες κυματομορφών (έστω για μια από τις μεθόδους των μοντέλων), που περνούν το κατώφλι του 65% ασχέτως από το ποσοστό αστάθειας, όπως το ορίσαμε προηγουμένως. Για αυτές τις κατηγορίες είναι απλά θέμα μεγέθους δείγματος, για να αποδειχθεί στατιστικά η ικανότητά τους. Το ίδιο θα μπορούσε εύκολα να ισχύσει και για κατηγορίες κοντά στο 65%. Το ερώτημα που τίθεται πλέον είναι, γνωρίζοντας τα άνω όρια αναγνώρισης για το δείγμα μας, με ποίο τρόπο θα εξετάσουμε την ακρίβεια αυτής της πρόβλεψης συνδυαστικά.

Κοιτάζοντας (στον πίνακα 10.3) την στήλη που ορίζει τη μέγιστη διαφορά μεταξύ των μεθόδων, παρατηρούμε ότι μονάχα 64 κατηγορίες βρίσκονται σε συμφωνία και στα 3 μοντέλα (απόκλιση μικρότερη των 10 ποσοστιαίων μονάδων μεταξύ των ποσοστών αναγνώρισης τους) με αυτές να παρουσιάζουν ποσοστά αναγνώρισης άνω του 65%. *Πράγμα που σημαίνει ότι υπάρχει πλέον ισχυρή ένδειξη ότι αυτές οι 64 κατηγορίες μπορούν αξιόπιστα να αναγνωρίσουν την κατηγορία ενός ασθενή με ακρίβεια τάξεως άνω του 65%.* Στον ακόλουθο πίνακα παρουσιάζονται οι κατηγορίες αυτές που μέσα από την ανάλυση των αποτελεσμάτων προέκυψαν ως δυνητικά αξιόπιστες, σε σχέση με την συνολική δυνατότητα διάκρισης ασθενών, ασχέτως ομάδας ασθενών.

11.3 Αξιολόγηση απόδοσης της κάθε ομάδας ασθενών

Σημαντικό είναι να δείξουμε και τα ποσοστά ακρίβειας για την κάθε ομάδα ασθενών καθώς και λόγω του μεγέθους του δείγματος αλλά και πιθανόν, άλλων παραγόντων, υπάρχουν κατηγορίες κυματομορφών που ευνοούν την αναγνώριση μιας ομάδας ασθενών έναντι της άλλης. Τα αναλυτικά διαγράμματα που απεικονίζουν με γραφικό τρόπο τα ποσοστά αναγνώρισης για την κάθε ομάδα ασθενών παρουσιάστηκαν στις εικόνες 10.2 - 10.3, για τα μοντέλα HMM και SVM αντίστοιχα.

Στον πίνακα 11.1 παρουσιάζονται τα αποτελέσματα (ακρίβεια) αναγνώρισης για κάθε ένα μοντέλο και για κάθε ομάδα ασθενών.

Μοντέλο / Κατηγορίες	Ομάδα 1 (N=76)			Ομάδα 2 (N=20)		
	(όλες)	(>65%)	(>65%)	(όλες)	(>65%)	(>65%)
HMM	72.26%	75.81%	-	67.64%	73.91%	-
SVM	66.23%	-	78.66%	50.90%	-	69.84%
Σύνολο κατηγοριών	146	116	64	146	45	26

Πίνακας 11.1: Πίνακας αποτελεσμάτων με βάση τις ομάδες των ασθενών.

Οι ομάδες περιέχουν 76 και 20 ασθενείς αντίστοιχα και εξετάζεται το ποσοστό αναγνώρισης για όλες (146) αλλά και για υποσύνολα των κατηγοριών κυματομορφών. Αυτό μας επιτρέπει να επιλέξουμε κατηγορίες με υψηλά ποσοστά αναγνώρισης και να παρακάμψουμε αυτές που υστερούν. Έτσι, για κάθε ομάδα και για κάθε μοντέλο επιλέχθηκαν κατηγορίες που βρίσκονται πάνω από ένα κατώφλι (πχ >80%) με το κατώφλι να είναι τέτοιο ώστε να συμπεριλαμβάνει μεγάλο αριθμό κατηγοριών. Αυτό γίνεται γιατί όταν χρειαστεί να εκτιμηθεί η κατάσταση ενός ασθενή χρειάζεται να συνυπολογιστούν όσο περισσότερες κυματομορφές είναι δυνατόν, ώστε το αποτέλεσμα της κατηγοριοποίησης να είναι στατιστικά ισχυρό.

11.4 Τελικός ταξινομητής με χρήση συστήματος πλειοψηφίας

Η τελική αποτίμηση της κατάστασης του ασθενή μπορεί να γίνει, με δεδομένα τα ποσοστά επιτυχίας που παρουσιάστηκαν προηγουμένως, μέσω ενός συστήματος πλειοψηφίας. Το σύστημα αυτό θα κατασκευαστεί με συντελεστές βάρους για κάθε μία από τις καλύτερες κατηγορίες κυματομορφών. Άρα, τα ποσοστά άνω του 50% δεν θα καταμετρούνται άμεσα σαν μια θετική ψήφος αλλά σαν μέρος θετικής ψήφου. Το συνολικό αποτέλεσμα μετά τον πολλαπλασιασμό του συντελεστή βάρους άλλα και την άθροιση θα δώσει την θετική ή αρνητική ψήφο.

$$Vote(p_1, p_2, \dots, p_n, w_1, w_2, \dots, w_n) = \left\lfloor \frac{1}{2} + \frac{\sum_{i=1}^n (p_i w_i) - n/2}{n} \right\rfloor \quad (11.1)$$

όπου p_i είναι το ποσοστό ακρίβειας για την κατηγορία κυματομορφών i και w_i ο συντελεστής βάρους για την κατηγορία κυματομορφών i .

Το ποσοστό ακρίβειας προκύπτει από το μοντέλο και δείχνει το μέγεθος ταύτισης της κυματομορφής με την εκ των προτέρων γνώση. Ο συντελεστής βάρους είναι το συνολικό ποσοστό ακρίβειας για κάθε κατηγορία κυματομορφών και προκύπτει μέσα από το cross validation από ολόκληρο το δείγμα εκπαίδευσης.

	Κατηγορία					Ψήφος
	1	2	.	.		
				145	146	
Ασθενής						
1	1	0.78	.	0.54	1	1
2	0	0.45	.	1	0.88	1
3	0.98	0	.	0	1	0
.
.
27	0	1	.	0	0.33	1
28	0.44	1	.	1	0	0
.	1
.	1
Cross validation	62.96%	66.67%	.	59.26%	44.44%	

Πίνακας 11.2: Υπόδειγμα πίνακα υπολογισμού ψήφων, συστήματος πλειοψηφίας.

Στον πίνακα 11.2 παρουσιάζεται η τελική μορφή της κατηγοριοποίησης των ασθενών, όπου η τελευταία στήλη (ψήφος) υπολογίζεται για κάθε ασθενή σύμφωνα με την εξίσωση 11.1. Στην συνέχεια θα παρουσιαστούν δύο πλειοψηφικά συστήματα ταξινόμησης όπου θα αναφέρονται στις δύο ομάδες ασθενών και θα κατηγοριοποιούν τους ασθενείς με βάση την ομάδα. Δημιουργήθηκε ένα σύστημα με high sensitivity και ένα σύστημα με high specificity. Η τομή των δύο κατασκευάζει τον τελικό ταξινομητή όπου έχει αυξημένο sensitivity/specificity αλλά παράλληλα δεν μπορεί να αποφανθεί σε περίπτωση που δεν συμφωνούν τα επιμέρους συστήματα. Ο τρόπος επιλογής κατηγοριών κυματομορφών για την κάθε ομάδα αναφέρεται στον πίνακα 11.1.

Υπόθεση "ανήκει στην ομάδα 1"						
	Κατηγορία					Ψήφος
	1	2	.	.		
				67	68	
Ασθενής						
Άγνωστος	p_1	p_2	.	p_{63}	p_{64}	$Vote()$
Cross validation	w_1	w_2	.	w_{63}	w_{64}	

Πίνακας 11.3: Πίνακας συστήματος πλειοψηφίας, για την ομάδα ασθενών 1. Μοντέλο SVM.

Ο πίνακας 11.3 αναφέρεται στα αποτελέσματα του μοντέλου SVM (low risk group) για τις 64 κατηγορίες κυματομορφών που παρουσίασαν ποσοστά αναγνώρισης μεγαλύτερα του 65%.

Υπόθεση "ανήκει στην ομάδα 2"						
	Κατηγορία					Ψήφος
	1	2	.	.		
				14	15	
Ασθενής						
Άγνωστος	p_1	p_2	.	p_{25}	p_{26}	$Vote()$
Cross validation	w_1	w_2	.	w_{25}	w_{26}	

Πίνακας 11.4: Πίνακας συστήματος πλειοψηφίας, για την ομάδα ασθενών 2. Μοντέλο SVM.

Ο πίνακας 11.4 αναφέρεται στα αποτελέσματα του μοντέλου SVM (high risk group) για τις 26 κατηγορίες κυματομορφών που παρουσίασαν ποσοστά αναγνώρισης μεγαλύτερα του 65%.

Η ολική (μέση) ικανότητα του ταξινομητή όπως προαναφέρθηκε κυμαίνεται περί το 78% για την ομάδα 1 ($N=76$), που μέσω του συστήματος πλειοψηφίας καταλήξαμε σε μια ακρίβεια περίπου 80% (*specificity - low risk*). Για την ομάδα 2 ($N=20$) έχουμε ολική (μέση) ικανότητα του ταξινομητή περί του 70%, που μέσω του συστήματος πλειοψηφίας καταλήξαμε σε μια ακρίβεια περίπου 69% (*sensitivity - high risk*).

Κεφάλαιο 12

Επίλογος

Κατά την επισκόπηση των αποτελεσμάτων προέκυψαν ερωτήματα αλλά και επιλογές ως προς τη λύση αυτών. Παρουσιάστηκαν με σύντομο τρόπο τα ποσοστά αναγνώρισης για την κάθε κατηγορία, κυρίως μέσω γραφημάτων και περιγράφηκε η σχέση που βρίσκουμε να χαρακτηρίζει αυτά τα ποσοστά με το πρόβλημα προς λύση.

Σκοπός αυτής της μελέτης ήταν η διασταύρωση πειραματικών δεδομένων, κινήσεις αθηρωματικής πλάκας, σε συμπτωματικούς και συμπτωματικούς ασθενείς. Τα δεδομένα προέκυψαν από εικόνα υπερήχου μετά από μια σειρά μελέτης που είχε προηγηθεί της παρούσας εργασίας. Ως αποτέλεσμα της μελέτης θεωρείται λοιπόν ένα σύστημα ικανό να διαχωρίζει/κατηγοριοποιεί ασθενείς με δεδομένα εισόδου αυτά τα πειραματικά δεδομένα.

12.1 Σύνοψη και συμπεράσματα

Λόγω του συνολικού μεγέθους δείγματος που είχαμε στην κατοχή μας, τα αποτελέσματα είναι ακριβή στο επίπεδο που επιτρέπει το μέγεθος αυτό. Για ακριβώς αυτό το λόγο χρησιμοποιήθηκε κατά την αξιολόγηση των μοντέλων που δημιουργήθηκαν, η μέθοδος διασταύρωσης αποτελεσμάτων (cross validation), εκπαιδεύοντας κάθε φορά τα μοντέλα με έναν ασθενή λιγότερο, δοκιμάζουμε τα μοντέλα κυκλικά (round robin) για τον άγνωστο ασθενή. Έτσι, έχουμε μια καλύτερη εικόνα για τις δυνατότητες του μοντέλου ως προς την κατηγοριοποίηση των ασθενών.

Αποδείχθηκε ότι τα Κρυφά Μαρκοβιανά Μοντέλα είναι ικανά να κατηγοριοποιήσουν τους ασθενείς μέσα από τα δεδομένα δεικτών κίνησης αθηρωματικής πλάκας καρωτίδας σε ποσοστό 74.86%. Το ποσοστό αυτό αναφέρεται στα διασταυρωμένα (cross-validated) αποτελέσματα θα μπορούσε να κυμανθεί μεταξύ του 70-77% αυξάνοντας τα δεδομένα εκπαίδευσης. Αυτό σημαίνει πως έχουμε την ικανότητα να αναγνωρίσουμε την συμπτωματική φύση ενός ασθενή με αθηρωμάτωση καρωτίδας σε ένα ικανοποιητικό ποσοστό.

Η απόδειξη της ικανότητα βέβαια δεν αποδεικνύει και την βέλτιστη λύση για το πρόβλημα. Καθώς η διαδικασία εκπαίδευσης και βελτιστοποίησης των παραμέτρων του συστήματος, για 96 ασθενείς με 146 κατηγορίες κυματομορφών διαρκεί σε ένα σύστημα με ταχύτητα μετάδοσης σκληρού δίσκου στα 150 MB/s περίπου 1 μήνα. Η διαδικασία αναγνώρισης ενός μόνο καινούργιου ασθενή φυσικά διαρκεί μερικά δευτερόλεπτα. Σε αντίθεση με τα Support Vector Machines που για το ίδιο πλήθος ασθενών εκπαιδεύεται μέσα σε μερικά λεπτά. Αυτό δείχνει τη χρονική δυσκολία αλλά και την πολυπλοκότητα του μοντέλου, άρα το συμπέρασμα είναι ότι η χρήση τους είναι ανώφελη για αυτή την εφαρμογή καθώς υπάρχει η περίπτωση χρήσης άλλου μοντέλου με καλύτερα ποσοστά κατηγοριοποίησης και μικρότερης πολυπλοκότητας.

12.2 Μελλοντικές επεκτάσεις

Μια σημαντική επέκταση στην πειραματική διαδικασία είναι μια επαυξημένη κυματομορφή, που θα περιέχει και πληροφορία ως προς τον *καρδιακό παλμό* αλλά και την *αρτηριακή πίεση*. *Ο διαφορετικός καρδιακός ρυθμός όπως και η διαφορά πίεσης του κάθε ασθενή θα μπορούσε να αλλοιώσει αισθητά την εξαγόμενη κυματομορφή*. Αυτό μπορεί να γίνει διότι ο αυλός συμπεριφέρεται διαφορετικά στο χρόνο, σε σχέση με την σύσπαση της καρωτίδας αλλά και σε σχέση με τις τάσεις που αναπτύσσονται από την αιματική ροή στην αθηρωματική πλάκα. Φαίνεται λογικό, να κυριαρχούν διαφορετικές συχνότητες ταλάντωσης για την πλάκα, έχοντας διαφορετικές πιέσεις στο κυκλοφορικό σύστημα. Ακόμη, στοιχεία σύστασης της αθηρωματικής πλάκας θα μπορούσαν να κατασκευάσουν υπο-κατηγορίες για ένα μοντέλο που θα περιέχει σημαντικό αριθμό ασθενών, ώστε να είναι δυνατή η εκπαίδευση για κάθε μια από τις υπο-κατηγορίες που αφορούν τη σύσταση της πλάκας.

Ενδιαφέρον επίσης παρουσιάζει το ποσοστό αναγνώρισης του κάθε ασθενή, σε σχέση με την εκ των προτέρων γνώση. Το γεγονός δηλαδή του να υπάρχει ένας συμπτωματικός (ή ασυμπτωματικός) ασθενής στο σύστημα και στη συνέχεια αυτός ο ασθενής να παρουσιάζει ποσοστά αναγνώρισης της τάξης του 30%. Αυτό σημαίνει είτε ότι ο ασθενής δεν μπορεί να κατηγοριοποιηθεί από τα μοντέλα ή ότι η *αρχική εκτίμηση της συμπτωματικής (ή ασυμπτωματικής) φύσης του έχει μεταβληθεί*. Αυτή η παρατήρηση μπορεί να αποδειχθεί σωτήρια για ασθενείς που μετέχουν σε κάποια επανεξέταση ή προϋπάρχουν στο σύστημα. Για τους νέους ασθενείς όπου και δεν υπάρχει εκ των προτέρων γνώση η παραπάνω διαπίστωση δεν βρίσκει εφαρμογή.

Τέλος, πρέπει να παρατηρηθεί ξανά ότι τα Support Vector Machines είναι μια πάρα πολύ αποδοτική μέθοδος κατηγοριοποίησης των δεδομένων μας, ταυτόχρονα σε χρόνο και πολυπλοκότητα, αλλά όπως φαίνεται και στον πηγαίο κώδικα, έχει μια πολύ πιο συμπαγή και κατανοητή γραφή. Δεν βασίζεται σε συγκεκριμένο λειτουργικό σύστημα (platform independent) και μπορεί να κατασκευαστεί με ελάχιστες γραμμές κώδικα και δεν βασίζεται στη ταχύτητα των περιφερειακών συστημάτων (σκληρός δίσκος) αλλά μονάχα στην επεξεργαστική ισχύ και τη μνήμη του υπολογιστικού συστήματος.

Βιβλιογραφία

- [1] Alan S Go, Dariush Mozaffarian, Véronique L Roger, Emelia J Benjamin, Jarett D Berry, William B Borden, Dawn M Bravata, Shifan Dai, Earl S Ford, Caroline S Fox, et al. Heart disease and stroke statistics—2013 update a report from the american heart association. *Circulation*, 127(1):e6--e245, 2013.
- [2] Steven Allender, Peter Scarborough, Viv Peto, Mike Rayner, Jose Leal, Ramon Luengo-Fernandez, and Alastair Gray. European cardiovascular disease statistics. *European Heart Network*, 3:11--35, 2008.
- [3] Eric P Widmaier, Hershel Raff, and Kevin T Strang. *Vander's human physiology*. McGraw-Hill Higher Education, 2006.
- [4] Eoghan Maher, Arthur Creane, Sherif Sultan, Niamh Hynes, Caitriona Lally, and Daniel J Kelly. Tensile and compressive properties of fresh human carotid atherosclerotic plaques. *Journal of biomechanics*, 42(16):2760--2767, 2009.
- [5] Marie-Louise M Grønholdt, Børge G Nordestgaard, Torben V Schroeder, Sissel Vorstrup, and Henrik Sillesen. Ultrasonic echolucent carotid plaques predict future strokes. *Circulation*, 104(1):68--73, 2001.
- [6] Osamu Honda, Seigo Sugiyama, Kiyotaka Kugiyama, Hironobu Fukushima, Shinichi Nakamura, Shunichi Koide, Sunao Kojima, Nobutaka Hirai, Hiroaki Kawano, Hirofumi Soejima, et al. Echolucent carotid plaques predict future coronary events in patients with coronary artery disease. *Journal of the American College of Cardiology*, 43(7):1177--1184, 2004.
- [7] M-L M Grønholdt, Britt M Wiebe, H Laursen, Tina G Nielsen, TV Schroeder, and H Sillesen. Lipid-rich carotid artery plaques appear echolucent on ultrasound b-mode images and may be associated with intraplaque haemorrhage. *European journal of vascular and endovascular surgery*, 14(6):439--445, 1997.
- [8] Hairong Shi, Tomy Varghese, Robert J Dempsey, Mohammed S Salamat, and James A Zagzebski. Relationship between ultrasonic attenuation, size and axial strain parameters for ex vivo atherosclerotic carotid plaque. *Ultrasound in medicine & biology*, 34(10):1666--1677, 2008.

- [9] F Rakebrandt, DC Crawford, D Havard, D Coleman, and JP Woodcock. Relationship between ultrasound texture classification images and histology of atherosclerotic plaque. *Ultrasound in medicine & biology*, 26(9):1393--1402, 2000.
- [10] PM Shankar, Flemming Forsberg, and Lauren Lown. Statistical modeling of atherosclerotic plaque in carotid b mode images—a feasibility study. *Ultrasound in medicine & biology*, 29(9):1305--1309, 2003.
- [11] Robin J Watson, Calum C McLean, M Pauliina Moore, Timothy Spencer, Donald M Salter, Tom Anderson, Keith AA Fox, and W Norman McDicken. Classification of arterial plaque by spectral analysis of in vitro radio frequency intravascular ultrasound data. *Ultrasound in medicine & biology*, 26(1):73--80, 2000.
- [12] G Alastair Lammie, Joanna Wardlaw, Paul Allan, CV Ruckley, Russell Peek, and David F Signorini. What pathological components indicate carotid atheroma activity and can these be identified reliably using ultrasound? *European Journal of Ultrasound*, 11(2):77--86, 2000.
- [13] Spyretta Golemati, Aimilia Gastounioti, and Konstantina S. Nikita. Toward novel noninvasive and low-cost markers for predicting strokes in asymptomatic carotid atherosclerosis: The role of ultrasound image analysis. *IEEE Trans. Biomed. Engineering*, 60(3):652--658, 2013.
- [14] Torbjørn Dahl, Jon Bang, Alexandra Ushakova, Stian Lydersen, and Hans Olav Myhre. Parameters describing motion in carotid artery plaques from ultrasound examination: A reproducibility study. *Ultrasound in medicine & biology*, 30(9):1133--1143, 2004.
- [15] Cristina Giannattasio and Giuseppe Mancina. Arterial distensibility in humans. modulating mechanisms, alterations in diseases and effects of treatment. *Journal of hypertension*, 20(10):1889--1899, 2002.
- [16] Guillaume Zahnd, Loïc Boussel, Adrien Marion, Marion Durand, Philippe Moulin, André Sérusclat, and Didier Vray. Measurement of two-dimensional movement parameters of the carotid artery wall for early detection of arteriosclerosis: a preliminary clinical study. *Ultrasound in medicine & biology*, 37(9):1421--1429, 2011.
- [17] Christina I Christodoulou, Constantinos S. Pattichis, Marios Pantziaris, and Andrew Nicolaidis. Texture-based classification of atherosclerotic carotid plaques. *Medical Imaging, IEEE Transactions on*, 22(7):902--912, 2003.
- [18] Stavroula Gr Mougiakakou, Spyretta Golemati, Ioannis Gousias, Andrew N Nicolaidis, and Konstantina S Nikita. Computer-aided diagnosis of carotid atherosclerosis based on ultrasound image statistics, laws' texture and neural networks. *Ultrasound in medicine & biology*, 33(1):26--36, 2007.

- [19] Nikolaos Tsiaparas, Spyretta Golemati, Ioannis Andreadis, John S Stoitsis, Ioannis Valavanis, and Konstantina S Nikita. Comparison of multiresolution features for texture classification of carotid atherosclerosis from b-mode ultrasound. *Information Technology in Biomedicine, IEEE Transactions on*, 15(1):130--137, 2011.
- [20] NN Tsiaparas, S Golemati, I Andreadis, J Stoitsis, I Valavanis, and KS Nikita. Assessment of carotid atherosclerosis from b-mode ultrasound images using directional multiscale texture features. *Measurement Science and Technology*, 23(11):114004, 2012.
- [21] Edward Kyriacou, Marios S Pattichis, Constantinos S Pattichis, A Mavrommatis, Christina I Christodoulou, S Kakkos, and Andrew Nicolaides. Classification of atherosclerotic carotid plaques using morphological analysis on ultrasound images. *Applied Intelligence*, 30(1): 3--23, 2009.
- [22] J Stoitsis, S Golemati, KS Nikita, and AN Nicolaides. Characterization of carotid atherosclerosis based on motion and texture features and clustering using fuzzy c-means. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, volume 1, pages 1407--1410. IEEE, 2004.
- [23] Σταύρος Μακροδημήτρης. Διερεύνηση της δυνατότητας διάκρισης επικίνδυνων αθηρωματικών πλακών με χρήση χαρακτηριστικών κίνησης. *Τομέας Συστημάτων Μετάδοσης Πληροφορίας και Τεχνολογίας Υλικών, ΣΗΜΜΥ, Ε.Μ.Π, DT2013-0162*, 2013.
- [24] Lawrence Rabiner and B Juang. An introduction to hidden markov models. *ASSP Magazine, IEEE*, 3(1):4--16, 1986.
- [25] Bernhard Obermaier, Christoph Guger, Christa Neuper, and Gert Pfurtscheller. Hidden markov models for online classification of single trial eeg data. *Pattern recognition letters*, 22(12):1299--1309, 2001.
- [26] Julien Thomas, Cédric Rose, and François Charpillet. A support system for eeg segmentation based on hidden markov models. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 3228--3231. IEEE, 2007.
- [27] Ligdamis Gutiérrez, J Ibaez, Guillermo Cortés, Javier Ramírez, C Benítez, Virginia Tenorio, and A Isaac. Volcano-seismic signal detection and classification processing using hidden markov models. application to san cristóbal volcano, nicaragua. In *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*, volume 4, pages IV--522. IEEE, 2009.
- [28] Leonard E Baum, Ted Petrie, George Soules, and Norman Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *The annals of mathematical statistics*, 41(1):164--171, 1970.

- [29] Bhupinder Singh, Neha Kapur, and Puneet Kaur. Speech recognition with hidden markov model: A review. *International Journal*, 2(3), 2012.
- [30] Yong Zhao and Biing-Hwang Juang. Modeling heterogeneous data sources for speech recognition using synchronous hidden markov models. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 7403--7407. IEEE, 2013.
- [31] Marti A. Hearst, ST Dumais, E Osman, John Platt, and Bernhard Scholkopf. Support vector machines. *Intelligent Systems and their Applications, IEEE*, 13(4):18--28, 1998.
- [32] Christopher JC Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121--167, 1998.
- [33] A Gastounioti, S Golemati, JS Stoitsis, and KS Nikita. Carotid artery wall motion analysis from b-mode ultrasound using adaptive block matching: in silico evaluation and in vivo application. *Physics in medicine and biology*, 58(24):8647, 2013.
- [34] A Gastounioti, S Golemati, J Stoitsis, and KS Nikita. Comparison of kalman-filter-based approaches for block matching in arterial wall motion analysis from b-mode ultrasound. *Measurement Science and Technology*, 22(11):114008, 2011.
- [35] Spyretta Golemati, J Stoitsis, Aimilia Gastounioti, Alexandros C Dimopoulos, Vassiliki Korpouli, and Konstantina S Nikita. Comparison of block matching and differential methods for motion analysis of the carotid artery wall from ultrasound images. 2012.

Πηγαίος κώδικας

Αρχείο 12.1: exports.m

```
1 % Copyright © Prevenios Marinos (2014). All rights reserved.
2
3 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
5 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Data Preparation %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
6 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
7 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
8
9 clearvars;
10 load('HMM_2ndDataset.mat');
11
12
13 %unix('rm -R mfcc_vect');
14 %unix('rm -R training_wavs');
15
16 mkdir('mfcc_vect',0);
17 mkdir('mfcc_vect/training_wavs/',0); mkdir('phones0',0); mkdir('
    aligned',0); mkdir('recout',0); mkdir('cross',0); mkdir('
    crossout',0);
18 mkdir('training_wavs',0); mkdir('trains',0); mkdir('codetrains'
    ,0); mkdir('tests',0); mkdir('results',0); mkdir('wordz',0);
    mkdir('crresults',0);
19 mkdir('hmm0s',0); mkdir('hmm1s',0); mkdir('hmm2s',0); mkdir('
    hmm3s',0); mkdir('hmm4s',0); mkdir('hmm5s',0); mkdir('hmm6s'
    ,0);
20 mkdir('hmm7s',0); mkdir('hmm8s',0); mkdir('hmm9s',0); mkdir('
    wintri',0);
21
22 [pat_end,~]=size(motion); %taking the data array sizes
23 [pat_end_2,~]=size(motion); %taking the data array sizes
24
25 %DEBUG = str2num(argDEBUG);
```

```

26 %wav_start = str2num(argwav_start);
27 %wav_end = str2num(argwav_end);
28
29
30 %if (DEBUG~=0 && DEBUG~=1)
31     TRI=1;          %bypass value for (triphones - 1) or (monophones - 0)
32     DEBUG=0;       %bypass value for (prune val - 1) or (cross val - 0)
33     wav_start=138; %bypass value
34     [~,wav_end]=size(motion); %taking the data array sizes
35     %wav_end=84; %bypass value
36 %end
37
38
39 if (DEBUG==1)
40     pat_end=1;
41 else
42     disp(' '); disp('_____'); disp('Cross
         validation , leave one out'); disp('
         _____');
43 end
44
45
46
47 for j=wav_start:wav_end % waveforms loop
48
49 for leave_one_out=1:pat_end
50
51     if (~isempty(motion{leave_one_out , j}) && leave_one_out~=77 &&
        leave_one_out~=46 && leave_one_out~=51 && leave_one_out~=52 &&
        leave_one_out~=53 && leave_one_out~=54 && leave_one_out~=55 &&
        leave_one_out~=56 && leave_one_out~=84 && leave_one_out~=93)
52
53         if (DEBUG==1)
54             disp(' '); disp('_____'); disp('Step , prune
                 finder'); disp('_____');
55             encoding_stage(1 , pat_end_2 , j , 0); % (step , prune finder)
56         else
57             encoding_stage(1 , pat_end , j , leave_one_out); % encoding the
                 raw data into proper vector form
58         end
59
60
61 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Creating monophone HMMs %%%%%%%%%%%%%%
62
63

```

```

64 mkdir( 'hmm0s/hmm0-%u',j);
65
66 clash=sprintf( 'HCompV -C config -f 0.01 -m -S trains/train-%u.scp -M
    hmm0s/hmm0-%u proto',j,j); unix(clash); % Creating monophone HMMs
67
68 unix( 'cp monophones1 monophones0');
69 unix( 'sed ''/sp/d'' monophones0 >> tmp'); %Create "monophones0" file
    with the same content as "monophones1" except the "sp" entry.
70 unix( 'cp tmp monophones0');
71 unix( 'rm tmp');
72
73 unix( 'java MakeHMMDefs proto monophones0'); %Create "hmmdefs" file
74 unix( 'java GenerateMacros vFloors'); % Create the "macros" file by
    opening "vFloors"
75
76
77
78 mkdir( 'hmmls/hmml-%u',j);
79
80 clash=sprintf( 'mv macros hmm0s/hmm0-%u/macros',j); unix(clash);
81
82 clash=sprintf( 'mv hmmdefs hmm0s/hmm0-%u/hmmdefs',j); unix(clash);
83
84
85
86 %Taking from file the best (pre-computed) pruning value
87 bash=sprintf( 'awk ''/class-%u /{print $3}'' class.txt',j);
88 [a,pruned]=unix(bash); clearvars bash;
89 err_rep(pruned);
90
91 bash=sprintf( 'awk ''/class-%u /{print $5}'' class.txt',j);
92 [a,stepd]=unix(bash); clearvars bash;
93 err_rep(stepd);
94
95 result2='';
96 crresult2='';
97
98 prune=str2num(pruned);
99 step=str2num(stepd); %takes from file the best stepping value
100
101 if (DEBUG==1)
102     prune_start = 0.1; % (step, prune finder)
103     step_start = 0.2; % (step, prune finder)
104
105     prune_end = 28.3; % (step, prune finder)

```

```

106     step_end = 0.2; % (step , prune finder)
107 else
108     prune_start=prune;
109     prune_end=prune;
110     step_start=step;
111     step_end=step;
112 end
113
114 for prune = prune_start:0.2:prune_end % (step , prune finder)
115     for step = step_start:0.1:step_end % (step , prune finder)
116
117         bash = sprintf('bash HERest.sh hmm0s/hmm0-%u hmm1s/hmm1-%u 7
118             config_phones0/phones0-%u.mlf trains/train-%u.scp
119             monophones0 %f %f',j,j,j,j,prune,step);
120         [a,b]=unix(bash); clearvars bash;
121         err_rep(b);
122         % HERest runned #n times and saved in hmm1
123
124 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Fixing the silence models %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
125
126         mkdir('hmm2s/hmm2-%u',j);
127
128         clash=sprintf('java Add_SP hmm1s/hmm1-%u/hmmdefs hmm2s/hmm2-%u/
129             hmmdefs',j,j); [a,b]=unix(clash); err_rep(b);
130
131         mkdir('hmm3s/hmm3-%u',j);
132
133         clash=sprintf('HHed -H hmm1s/hmm1-%u/macros -H hmm2s/hmm2-%u/
134             hmmdefs -M hmm3s/hmm3-%u sil.hed monophones1',j,j,j); [a,b
135             ]=unix(clash);
136         err_rep(b);
137
138         mkdir('hmm4s/hmm4-%u',j);
139
140         bash=sprintf('bash HERest.sh hmm3s/hmm3-%u hmm4s/hmm4-%u 7
141             config_phones0/phones0-%u.mlf trains/train-%u.scp
142             monophones1 %f %f',j,j,j,j,prune,step); [a,b]=unix(bash);
143         clearvars bash;
144         err_rep(b);
145
146 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Realigning the data %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
147
148         clash=sprintf('HVite -l '*' -o SWT -b silence -C config -a -H
149             hmm4s/hmm4-%u/macros -H hmm4s/hmm4-%u/hmmdefs -i aligned/
150             aligned-%u.mlf -m -y lab -I wordz/words-%u.mlf -S trains /

```



```

142         train-%u.scp dict monophones1',j,j,j,j,j); [a,b]=unix(clash
143     );
144     err_rep(b);
145
146     mkdir('hmm5s/hmm5-%u',j);
147
148     bash = sprintf('bash HERest.sh hmm4s/hmm4-%u hmm5s/hmm5-%u 7
149         config phones0/phones0-%u.mlf trains/train-%u.scp
150         monophones1 %3.3f %3.3f',j,j,j,j,prune,step); [a,b]=unix(
151         bash); clearvars bash;
152     err_rep(b);
153
154 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Making Triphones from Monophones %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
155
156 if (TRI==1)
157     % Generating triphones transcriptions for training data
158     % mktri.led: edit script explaining how to handle pauses (p.
159     38)
160     % wintri.mlf: word-internal triphone transcriptions (output)
161     % triphones1: list of triphones (output)
162     clash=sprintf('HLEd -n triphones1 -l '*' -i wintri/wintri-%u.
163         mlf mktri.led aligned/aligned-%u.mlf',j,j); [a,b]=unix(
164         clash);
165     err_rep(b);
166
167     clash=sprintf('perl maketrihed monophones1 triphones1'); [a,b]=
168         unix(clash);
169     err_rep(b);
170
171     mkdir('hmm6s/hmm6-%u',j);
172
173     % Generating context-dependent triphones by cloning monophones
174     clash=sprintf('HHEd -H hmm5s/hmm5-%u/macros -H hmm5s/hmm5-%u/
175         hmmdefs -M hmm6s/hmm6-%u mktri.hed monophones1',j,j,j); [a,
176         b]=unix(clash);
177     err_rep(b);
178
179     mkdir('hmm7s/hmm7-%u',j);
180
181     % Reestimating the triphone set with HERest
182     clash=sprintf('bash HERest.sh hmm6s/hmm6-%u hmm7s/hmm7-%u 7
183         config wintri/wintri-%u.mlf trains/train-%u.scp triphones1
184         %3.3f %3.3f',j,j,j,j,prune,step); [a,b]=unix(clash);
185     err_rep(b);
186
187

```

```

174 % Generate a list of all possible triphones on the entire
      dictionary, not just those on the training set
175 clash=sprintf('HDMan -A -D -T 1 -b sp -n fulllist -g global.ded
      -l flog dict-tri lexicon.lex'); [a,b]=unix(clash); err_rep
      (b);
176
177 clash=sprintf('cp fulllist fulllist1; cat triphones1 >>
      fulllist1; perl fixfullist.pl fulllist1 fulllist'); [a,b]=
      unix(clash); err_rep(b);
178
179 clash=sprintf('cp tree.hed trees.hed'); [a,b]=unix(clash);
      err_rep(b);
180 clash=sprintf('perl mkelscript.prl TB 350 monophones0 >> trees.
      hed'); [a,b]=unix(clash); err_rep(b);
181 clash=sprintf('printf "\nTR 1\n\nAU \"fulllist\"\n\nCO \"tiedlist
      \"\n\nST \"trees\"\" >> trees.hed'); [a,b]=unix(clash);
      err_rep(b);
182
183 mkdir('hmm8s/hmm8-%u',j);
184
185 % Here we use a method based on decision trees, which is based
      on asking questions about the left and right context of
      each triphone
186 clash=sprintf('HHEd -H hmm7s/hmm7-%u/macros -H hmm7s/hmm7-%u/
      hmmdefs -M hmm8s/hmm8-%u trees.hed triphones1 > log',j,j,j)
      ;[a,b]=unix(clash); err_rep(b);
187
188
189 mkdir('hmm9s/hmm9-%u',j);
190
191
192 clash=sprintf('bash HERest.sh hmm8s/hmm8-%u hmm9s/hmm9-%u 7
      config wintri/wintri-%u.mlf trains/train-%u.scp tiedlist
      %3.3f %3.3f',j,j,j,j,prune,step); [a,b]=unix(clash);
193 err_rep(b);
194 end
195
196 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Recognizer Evaluation %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
197
198 clash=sprintf('cp trains/train-%u.scp tests/test-%u.scp',j,j);
      unix(clash);
199
200 if(TRI==1)
201 % ----- Triphones -----

```

```

202 clash=sprintf('bash HVite.sh hmm9s/hmm9-%u tests/test-%u.scp
      recout/recout-%u.mlf wdnet dict tiedlist wordz/words-%u.mlf
      ',j,j,j,j);
203
204 else
205
206 % ----- Monophones -----
207 clash=sprintf('bash HVite.sh hmm5s/hmm5-%u tests/test-%u.scp
      recout/recout-%u.mlf wdnet dict monophones1 wordz/words-%u.
      mlf',j,j,j,j);
208 end
209 [status,result]=unix(clash);
210
211
212
213 %%%%%%%%%% ----- (Leave-one-out cross-validation) ----- %%%%%%%%%%
214 if (DEBUG~=1)
215     if (TRI==1)
216         % ----- Triphones -----
217         clash=sprintf('bash HVite.sh hmm9s/hmm9-%u cross/cross-%u.scp
              crossout/crossout-%u.mlf wdnet dict tiedlist wordz/words-%u.
              mlf',j,leave_one_out,leave_one_out,j);
218     else
219         % ----- Monophones -----
220         clash=sprintf('bash HVite.sh hmm5s/hmm5-%u cross/cross-%u.scp
              crossout/crossout-%u.mlf wdnet dict monophones1 wordz/words
              -%u.mlf',j,leave_one_out,leave_one_out,j);
221     end
222     [crstatus,crresult]=unix(clash);
223 end
224 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Saving results %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
225
226 where=sprintf('prune %f step %f\n',prune,step);
227 result2=sprintf('%s\n%s%s\n',result2,result,where);
228 re=sprintf('results/results%u.txt',j);
229
230
231
232 if (DEBUG==1) % (step, prune finder)
233
234     if (prune<=prune_start && step<=step_start)
235         results = fopen(re,'w');
236     end
237
238 else

```

```

239     crresult2=sprintf( '%s\n%s%s\n', crresult2 , crresult , where);
240     ress=sprintf( 'crresults / results%u-%u.txt ', j , leave_one_out);
241     crresults = fopen(ress , 'w');
242     fprintf( crresults , '%s ' , crresult2);
243     results = fopen(re , 'w');
244 end
245
246 fprintf( results , '%s ' , result2);
247
248 if( DEBUG~=1)
249     disp( re);
250     disp( ress);
251     disp( 'written ');
252
253     fclose( results);
254     fclose( crresults);
255 else
256     result2='';
257 end
258
259 end % (step , prune finder)
260
261     if( DEBUG==1)
262         disp( prune); % (step , prune finder)
263     end
264
265 end % (step , prune finder)
266
267     end
268 end
269
270     if( DEBUG==1)
271         fclose( results); % (step , prune finder)
272     end
273
274 end

```

Αρχείο 12.2: encoding_stage.m

```

1 % Copyright © Prevenios Marinos (2014). All rights reserved.
2
3 function [] = encoding_stage(pat_start , pat_end , wav_start , leave_one_out)
4 % pat_start = first patient index
5 % pat_end   = last patient index
6 % wav_start = first waveform index
7 % wav_end   = last waveform index
8
9 load( 'HMM_2ndDataset.mat' );
10
11 class_maximum = findmax(motion);
12
13
14 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
15 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
16 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Data Preparation %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
17 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
18 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
19
20
21 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Lexicon %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
22
23 lexicon = fopen('lexicon.lex','wt');
24 fprintf(lexicon, '!ENTER          sil\n!EXIT          sil\nnone
25           w ah n\npac          p ae k\n');
26 fclose(lexicon);
27
28 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Task Grammar %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
29
30 gram = fopen('gram','wt');
31 fprintf(gram, '$word = one | pac;\n(SENT-START <$word> SENT-END)\n');
32 fclose(gram);
33
34 unix('HParse gram wdnet'); % Creating word network "wdnet" using HParse
35
36
37 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Dictionary %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
38
39 wlist = fopen('wlist','wt');
40 fprintf(wlist, 'one\npac\n');
41 fclose(wlist);
42
43 unix('HDMan -m -w wlist -n monophonesl -l dlog dict lexicon.lex');

```

```

44
45 unix('printf "SENT-END [] sil\nSENT-START [] sil\nsilence sil\n" >>
    dict'); % Adding .. with the alphabetical order of the file.
46 unix('printf "sil\n" >> monophones1'); % Adding "sil sp" at the end of
    the file.
47
48
49 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Transcriptions and Coding data %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
50
51 mkphones = fopen('mkphones0.led','wt');
52 fprintf(mkphones,'EX\nIS sil sil\nDE sp\n');
53 fclose(mkphones);
54
55
56 wav_config = fopen('wav_config','wt');
57 fprintf(wav_config,'SOURCEFORMAT = WAV\nTARGETKIND = MFCC_0_D_A\n\
    nTARGETRATE = 8000.0\nSAVECOMPRESSED = T\nSAVEWITHCRC = T\
    nWINDOWSIZE = 8000.0\nUSEHAMMING = T\nPREEMCOEF = 0.97\nNUMCHANS =
    26\nCEPLIFTER = 22\nNUMCEPS = 12\nENORMALISE = F');
58 fclose(wav_config);
59
60 config = fopen('config','wt');
61 fprintf(config,'TARGETKIND = MFCC_0_D_A\nTARGETRATE = 8000.0\
    nSAVECOMPRESSED = T\nSAVEWITHCRC = T\nWINDOWSIZE = 8000.0\
    nUSEHAMMING = T\nPREEMCOEF = 0.97\nNUMCHANS = 26\nCEPLIFTER = 22\
    nNUMCEPS = 12\nENORMALISE = F');
62 fclose(config);
63
64
65
66 j=wav_start; % waveforms
67
68 clash=sprintf('wordz/words-%u.mlf',j);
69 words = fopen(clash,'wt');
70 fprintf(words,'#!MLF!\n');
71
72 clash=sprintf('trains/train-%u.scp',j);
73 train = fopen(clash,'wt');
74
75 clash=sprintf('cross/cross-%u.scp',leave_one_out);
76 crossfile = fopen(clash,'wt');
77
78 clash=sprintf('codetrains/codetrain-%u.scp',j);
79 codetrain = fopen(clash,'wt'); % Note the 'wt' for writing in text
    mode

```

```

80
81
82 for i=pat_start:pat_end % patients loop
83
84     if (motion{i,1}==1) % patient's classification 1 for ???, 2
85         for ?????
86             cat='cat1_';
87         elseif (motion{i,1}==2)
88             cat='cat2_';
89         end
90
91         if (~isempty(motion{i,j}) && i~=77 && i~=46 && i~=51 && i~=52
92             && i~=53 && i~=54 && i~=55 && i~=56 && i~=84 && i~=93) %
93             exclude empty waveforms from dataset, AND patient 18,
94
95             filename = strcat('training_wavs/',cat,'file_',int2str(j),'
96                 _',int2str(i)); % filename counter, waveform j,
97                 patient i
98
99             if (max(abs(motion{i,j}))<=1)
100                 sig_normalized=motion{i,j};
101                 sig_normalized=motion{i,j}/(max(abs(motion{i,j}))
102                 +0.05);
103
104                 sig_normalized=timescale(sig_normalized,
105                 class_maximum(j));
106             else
107                 sig_normalized=motion{i,j}/(max(abs(motion{i,j}))
108                 +0.05); % normalize waveform to (-1,1)
109                 sig_normalized=timescale(sig_normalized,
110                 class_maximum(j));
111             end
112
113             wavwrite(sig_normalized, filename); % export
114                 waveforms, FS=8000, 16bit, *.wav
115
116             fprintf(codetrain, '%s.wav mfcc_vect/%s.mfc\n', filename,
117                 filename); % export the mfc script (codetrain)
118
119             if (motion{i,1}==1)
120                 fprintf(words, '*/%sfile_%s_%s.lab\n', cat,
121                     int2str(j), int2str(i)); %
122                     transcriptions
123                 fprintf(words, 'one\n.\n');

```

```

112
113         if (i~=leave_one_out)
114             fprintf(train, 'mfcc_vect/%s.mfc\n',
115                     filename);
116         else
117             fprintf(crossfile, 'mfcc_vect/%s.mfc\n',
118                     filename);
119         end
120     elseif (motion{i,1}==2)
121         fprintf(words, '*/%sfile_%s_%s.lab "\n', cat,
122                 int2str(j), int2str(i)); %
123                 transcriptions
124         fprintf(words, 'pac\n.\n');
125     if (i~=leave_one_out)
126         fprintf(train, 'mfcc_vect/%s.mfc\n',
127                 filename);
128     else
129         fprintf(crossfile, 'mfcc_vect/%s.mfc\n',
130                 filename);
131     end
132     end
133     else
134         disp(' '); disp('empty waveform or banned patient');
135         filename = strcat('training_wavs/', cat, 'file_', int2str(j), '
136             _', int2str(i)); % filename counter, waveform j,
137             patient i
138         disp(filename); disp(' ');
139     end
140 end
141
142 clash=sprintf('HLEd -l ''*' -d dict -i phones0/phones0-%u.mlf
143             mkphones0.led wordz/words-%u.mlf', j, j);
144 [~,b]=unix(clash); % Generate the phone level MLF "phones0.mlf" via
145             HLEd
146 err_rep(b);
147
148 clash=sprintf('HCopy -A -D -T 1 -C wav_config -S codetrains/codetrain-%
149             u.scp', j);
150 [~,b]=unix(clash); % Feature (vectors) extraction (.mfc files)
151 err_rep(b);
152

```



```
146  
147  
148 fclose(crossfile);  
149 fclose(words);  
150 fclose(codetrain);  
151 fclose(train);  
152  
153 end
```

Αρχείο 12.3: timescale.m

```
1 % Copyright © Prevenios Marinos (2014). All rights reserved.
2
3 function output = timescale(sig , outlength)
4 % sig = input signal
5 % outlength = desired output length after scaling
6
7
8 n = length(sig);
9
10 if ((n/2) > outlength)
11     sig=timescale(sig , outlength*2);
12     n = length(sig);
13     timediv = floor(n/abs(n-outlength));
14 elseif ((2*n) < outlength)
15     sig=timescale(sig , outlength/2);
16     n = length(sig);
17     timediv = floor(n/abs(n-outlength));
18 else
19     timediv = floor(n/abs(n-outlength));
20 end
21
22
23
24 % Remove samples or add samples (shrink or expand)
25 j=1;
26
27 if (n>=outlength)
28
29     for i = 1:n
30
31         if (mod(i , timediv)~=0)
32
33             output(j ,1) = sig(i ,1);
34             j=j+1;
35
36         end
37
38     end
39
40 else
41
42     for i = 1:n
43
44         if (mod(i , timediv)==0)
```

```
45         output(j,1) = sig(i,1);
46         if (i<n)
47             output(j+1,1) = (sig(i,1)+sig(i+1,1))/2;
48         else
49             output(j+1,1) = sig(i,1);
50         end
51         j=j+2;
52     else
53         output(j,1) = sig(i,1);
54         j=j+1;
55     end
56 end
57
58 end
59
60 end
61
62
63
64 end
```

Αρχείο 12.4: findmin.m

```
1 % Copyright © Prevenios Marinos (2014). All rights reserved.
2
3 function output = findmin(motion)
4
5 output(1) = 0;
6 localmin=9999;
7 [patients ,~] = size(motion);
8
9 for i=2:(length(motion))
10
11     for j=1:patients
12
13         if (localmin > length(motion{j , i}))
14             localmin = length(motion{j , i});
15         end
16
17     end
18
19     output(i) = localmin;
20     localmin=9999;
21
22 end
```

Αρχείο 12.5: findmax.m

```
1 % Copyright © Prevenios Marinos (2014). All rights reserved.
2
3 function output = findmax(motion)
4
5 output(1) = 0;
6 localmax=0;
7 [patients ,~] = size(motion);
8
9 for i=2:(length(motion))
10
11     for j=1:patients
12
13         if (localmax < length(motion{j , i}))
14             localmax = length(motion{j , i});
15         end
16
17     end
18
```

```
19 |     output(i) = localmax;  
20 |     localmax=0;  
21 |  
22 | end
```

Αρχείο 12.6: err_rep.m

```
1 % Copyright © Prevenios Marinos (2014). All rights reserved.
2
3 function [] = err_rep(b)
4
5     if (strfind(b, 'bad data'))
6         disp(b);
7
8         % err = fopen('tmp.txt','wt');
9         %fprintf(err,'%s',b);
10        %fclose(err);
11
12        %unix('awk '/bad data/{print $5}' tmp.txt >> error2.txt');
13        %unix('sed '$!N; /^\(.*\) \n\1$/!P; D' error2.txt >> error3.txt
14            ');
15        %unix('sed '$!N; /^\(.*\) \n\1$/!P; D' error.txt >> error4.txt')
16            ;
17        %unix('cp error4.txt error.txt');
18        %unix('sed -r 's|.|(.*) .mfc.*|\1|' error3.txt >> error.txt');
19
20        % unix('rm tmp.txt');
21        % unix('rm error2.txt');
22        % unix('rm error3.txt');
23        % unix('rm error4.txt');
24
25    elseif (strfind(b, 'ERROR'))
26        disp(b);
27    elseif (strfind(b, 'error'))
28        disp(b);
29    elseif (strfind(b, 'Error'))
30        disp(b);
31    end
```

Αρχείο 12.7: mkdir.m

```
1 % Copyright © Prevenios Marinos (2014). All rights reserved.
2
3 function [] = mkdir(sprintf_dir, class)
4
5 h=sprintf(sprintf_dir, class);
6 a=exist(h, 'dir');
7
8 if (a~=7)
9     mkdir(h);
```

10 || end
11 ||
12 || end

Αρχείο 12.8: result.sh

```
1 # Copyright © Prevenios Marinos (2014). All rights reserved.
2
3 #!/bin/sh
4 # Script to find the best pruning for each waveform from result files
5 # Usage :$ bash result.sh [start class] [end class]
6
7 float_scale=2
8
9 function float_eval()
10 {
11     local stat=0
12     local result=0.0
13     if [[ $# -gt 0 ]]; then
14         result=$(echo "scale=$float_scale; $*" | bc -q 2>/dev/null)
15         stat=$?
16         if [[ $stat -eq 0 && -z "$result" ]]; then stat=1; fi
17     fi
18     echo $result
19     return $stat
20 }
21
22 function float_cond()
23 {
24     local cond=0
25     if [[ $# -gt 0 ]]; then
26         cond=$(echo "$*" | bc -q 2>/dev/null)
27         if [[ -z "$cond" ]]; then cond=0; fi
28         if [[ "$cond" != 0 && "$cond" != 1 ]]; then cond=0; fi
29     fi
30     local stat=$((cond == 0))
31     return $stat
32 }
33
34 rm final.txt
35 if [ -f "results/results140.txt" ];
36 then
37     rm class.txt
38 else
39     echo " "
40 fi
41
42
43 num=$1
44
```



```

45 while (( num <= $2 ))
46 do
47
48 cnt=100.0
49 step=0.01
50
51 res=$(sed -n -e '/Acc='$cnt '/, /step/p' results/results"$num".txt)
52
53 while [ "$res" == "" ]
54 do
55     cnt=$(float_eval "$cnt - $step")
56     res=$(sed -n -e '/Acc='$cnt '/, /step/p' results/results"$num".txt)
57     # echo $cnt
58 done
59
60 echo "$res" >> "tempy.txt"
61 res2=$(tail -1 tempy.txt)
62
63 res3=$(sed -e './.{H;$!d;}' -e 'x;/' "$(tail -1 tempy.txt)" '!d' results
64 /results"$num".txt)
65 rm tempy.txt
66 echo "Results from class "$num"" >> "final.txt"
67
68 if [ -f "results/results140.txt" ];
69 then
70     echo "class-"$num" "$res2"" >> "class.txt"
71 else
72     echo "Testing mode"
73 fi
74
75 echo "$res3" >> "final.txt"
76 echo "" >> "final.txt"
77 echo "" >> "final.txt"
78 echo $num
79 (( num++ ))
80
81 done
82
83 rm accuracy.txt
84 cat final.txt | grep Acc= >> accuracy.txt

```

Αρχείο 12.9: cross.sh

```
1 #!/bin/sh
2 # Copyright © Prevenios Marinos (2014). All rights reserved.
3
4 # Script to export the cross-validated results
5 # for
6 # Usage :$ bash cross.sh [start class] [end class]
7
8 patient=1
9 class=$1
10 classdef=$class
11 res=''
12 space=' '
13 newline=$'\n'
14 rm corr.csv
15
16 while (( patient <= 96 ))
17 do
18
19     while (( class <= $2 ))
20     do
21         corr=$corr$(cat crresults/results"$class"-"$patient".txt | grep Acc |
22             cut -d \ -f 2 | sed 's/^\.\{6\}//' | sed s '/. $//')$space
23         (( class++ ))
24     done
25
26     corr=$corr$newline
27     # acc=$acc$newline
28     (( patient++ ))
29     class=$classdef
30
31 done
32
33 echo "$corr" >> "corr.csv"
```

```

1 % Copyright © Prevenios Marinos (2014). All rights reserved.
2
3 clearvars;
4
5 load('HMM_2ndDataset.mat');
6
7 %autoscale=1; % svm autoscaling in training
8
9 %%finding mean value for every waveform
10 motion_mean=find_mean_val();
11
12
13 for j=2:147
14
15     disp({'Class',j});
16     string='';
17     valid=0;
18
19     for i=1:96
20
21         if (~isempty(motion{i,j})) % exclude empty waveforms from dataset
22
23             valid=valid+1; % valid patients for each
24             waveform (class)
25             classification = motion{valid,1};
26
27             if (classification==2)
28                 classification=0;
29             end
30
31             n = length(motion{i,j});
32
33             group(valid,1)=classification; % patient's classification (
34             known value)
35
36             for k=1:n % for every time unit, make an svm vector
37
38                 if (isnan(motion{i,j}(k,1))) % avoiding NaN in training
39                     data
40                     xdata(valid,k) = motion_mean{i,j}; %case of NaN is
41                     accounted as
42
43                     %the mean value
44                     of the
45                     waveform

```

```

39         else
40             xdata(valid,k) = motion{i,j}(k,1);
41         end
42
43     end
44
45 end
46
47 end
48
49 clearvars ress;
50
51 for (p=1:96) % leave one out – cross validation, all valid patient's
52             % in a round robin cross validation, each time one
53             % patient
54             % is the test patient, and the others are the training
55             % data
56
57     for (pol=1:valid)
58         train(pol,1)=1;
59         test(pol,1)=0;
60     end
61
62     train(p,1)=0;
63     train = logical(train);
64
65     test(p,1)=1;
66     test = logical(test);
67
68     group=logical(group);
69
70     %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Here we choose the best classification
71     %%%%%%%%%%%
72     %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% percentage from the svm methods
73     %%%%%%%%%%%
74
75     cp = classperf(group);
76     svmStruct = svmtrain(xdata(train,:),group(train),'Kernel_Function',
77         'mlp','Method','SMO','AutoScale',0);
78     classes = svmclassify(svmStruct,xdata(test,:));
79     classperf(cp,classes,test);
80     ress_1(p)=cp;
81
82     cp = classperf(group);

```

```

82 svmStruct = svmtrain(xdata(train,:),group(train),'Kernel_Function',
83     'mlp','Method','SMO','AutoScale',1);
84 classes = svmclassify(svmStruct,xdata(test,:));
85 classperf(cp,classes,test);
86 ress_2(p)=cp;
87
88 cp = classperf(group);
89 svmStruct = svmtrain(xdata(train,:),group(train),'Kernel_Function',
90     'mlp','Method','LS','AutoScale',0);
91 classes = svmclassify(svmStruct,xdata(test,:));
92 classperf(cp,classes,test);
93 ress_3(p)=cp;
94
95 cp = classperf(group);
96 svmStruct = svmtrain(xdata(train,:),group(train),'Kernel_Function',
97     'mlp','Method','LS','AutoScale',1);
98 classes = svmclassify(svmStruct,xdata(test,:));
99 classperf(cp,classes,test);
100 ress_4(p)=cp;
101
102 cp = classperf(group);
103 svmStruct = svmtrain(xdata(train,:),group(train),'Kernel_Function',
104     'linear','AutoScale',0);
105 classes = svmclassify(svmStruct,xdata(test,:));
106 classperf(cp,classes,test);
107 ress_5(p)=cp;
108
109 cp = classperf(group);
110 svmStruct = svmtrain(xdata(train,:),group(train),'Kernel_Function',
111     'linear','AutoScale',1);
112 classes = svmclassify(svmStruct,xdata(test,:));
113 classperf(cp,classes,test);
114 ress_6(p)=cp;
115
116 cp = classperf(group);
117 svmStruct = svmtrain(xdata(train,:),group(train),'Kernel_Function',
118     'quadratic','AutoScale',0);
119 classes = svmclassify(svmStruct,xdata(test,:));
120 classperf(cp,classes,test);
121 ress_7(p)=cp;
122
123 cp = classperf(group);
124 svmStruct = svmtrain(xdata(train,:),group(train),'Kernel_Function',
125     'quadratic','AutoScale',1);
126 classes = svmclassify(svmStruct,xdata(test,:));

```

```

116 classperf(cp, classes, test);
117 ress_8(p)=cp;
118
119
120 % saving cross validation results for each waveform (class)
121 if (p==valid)
122
123     rnsn=0;
124     for (dd=1:valid) bou(dd)=ress_1(dd).CorrectRate; rnsn=rnsn+
        ress_1(dd).CorrectRate; end
125     corr1(:,j)=bou';
126     asc1 = (bou(2)+bou(3)+bou(9)+bou(11)+bou(14)+bou(15)+bou(25)+
        bou(37)+bou(41)+bou(43)+bou(63)+bou(73)+bou(74)+bou(76)+bou
        (81)+bou(86)+bou(87)+bou(91)+bou(95)+bou(96))/20;
127     asc2 = (rnsn-(bou(2)+bou(3)+bou(9)+bou(11)+bou(14)+bou(15)+bou
        (25)+bou(37)+bou(41)+bou(43)+bou(63)+bou(73)+bou(74)+bou
        (76)+bou(81)+bou(86)+bou(87)+bou(91)+bou(95)+bou(96)))/76;
128     asc = (asc1+asc2)/2;
129     average(j-1) = asc; % contains the cross
        validation (leave one out), for each waveform,
130 % small group accuracy
131     class(j-1)=1; % chosen scenario
132     corr(:,j)=bou'; % cross validation results
        for waveform j
133
134     rnsn=0;
135     for (dd=1:valid) bou(dd)=ress_2(dd).CorrectRate; rnsn=rnsn+
        ress_2(dd).CorrectRate; end
136     corr2(:,j)=bou';
137     bsc1 = (bou(2)+bou(3)+bou(9)+bou(11)+bou(14)+bou(15)+bou(25)+
        bou(37)+bou(41)+bou(43)+bou(63)+bou(73)+bou(74)+bou(76)+bou
        (81)+bou(86)+bou(87)+bou(91)+bou(95)+bou(96))/20;
138     bsc2 = (rnsn-(bou(2)+bou(3)+bou(9)+bou(11)+bou(14)+bou(15)+bou
        (25)+bou(37)+bou(41)+bou(43)+bou(63)+bou(73)+bou(74)+bou
        (76)+bou(81)+bou(86)+bou(87)+bou(91)+bou(95)+bou(96)))/76;
139     bsc = (bsc1+bsc2)/2;
140     if (average(j-1) < bsc)
141         average(j-1) = bsc;
142         class(j-1)=2;
143         corr(:,j)=bou';
144     end
145
146     rnsn=0;
147     for (dd=1:valid) bou(dd)=ress_3(dd).CorrectRate; rnsn=rnsn+
        ress_3(dd).CorrectRate; end

```

```

148 corr3 (: , j)=bou ' ;
149 csc1 = ( bou (2)+bou (3)+bou (9)+bou (11)+bou (14)+bou (15)+bou (25)+
      bou (37)+bou (41)+bou (43)+bou (63)+bou (73)+bou (74)+bou (76)+bou
      (81)+bou (86)+bou (87)+bou (91)+bou (95)+bou (96) ) /20 ;
150 csc2 = ( rnsn -(bou (2)+bou (3)+bou (9)+bou (11)+bou (14)+bou (15)+bou
      (25)+bou (37)+bou (41)+bou (43)+bou (63)+bou (73)+bou (74)+bou
      (76)+bou (81)+bou (86)+bou (87)+bou (91)+bou (95)+bou (96) ) ) /76 ;
151 csc = ( csc1+csc2 ) /2 ;
152 if ( average (j-1) < csc )
153     average (j-1) = csc ;
154     clss (j-1)=3 ;
155     corr (: , j)=bou ' ;
156 end
157
158 rnsn=0 ;
159 for (dd=1: valid) bou (dd)=ress_4 (dd) . CorrectRate ; rnsn=rnsn+
      ress_4 (dd) . CorrectRate ; end
160 corr4 (: , j)=bou ' ;
161 dsc1 = ( bou (2)+bou (3)+bou (9)+bou (11)+bou (14)+bou (15)+bou (25)+
      bou (37)+bou (41)+bou (43)+bou (63)+bou (73)+bou (74)+bou (76)+bou
      (81)+bou (86)+bou (87)+bou (91)+bou (95)+bou (96) ) /20 ;
162 dsc2 = ( rnsn -(bou (2)+bou (3)+bou (9)+bou (11)+bou (14)+bou (15)+bou
      (25)+bou (37)+bou (41)+bou (43)+bou (63)+bou (73)+bou (74)+bou
      (76)+bou (81)+bou (86)+bou (87)+bou (91)+bou (95)+bou (96) ) ) /76 ;
163 dsc = ( dsc1+dsc2 ) /2 ;
164 if ( average (j-1) < dsc )
165     average (j-1) = dsc ;
166     clss (j-1)=4 ;
167     corr (: , j)=bou ' ;
168 end
169
170 rnsn=0 ;
171 for (dd=1: valid) bou (dd)=ress_5 (dd) . CorrectRate ; rnsn=rnsn+
      ress_5 (dd) . CorrectRate ; end
172 corr5 (: , j)=bou ' ;
173 esc1 = ( bou (2)+bou (3)+bou (9)+bou (11)+bou (14)+bou (15)+bou (25)+
      bou (37)+bou (41)+bou (43)+bou (63)+bou (73)+bou (74)+bou (76)+bou
      (81)+bou (86)+bou (87)+bou (91)+bou (95)+bou (96) ) /20 ;
174 esc2 = ( rnsn -(bou (2)+bou (3)+bou (9)+bou (11)+bou (14)+bou (15)+bou
      (25)+bou (37)+bou (41)+bou (43)+bou (63)+bou (73)+bou (74)+bou
      (76)+bou (81)+bou (86)+bou (87)+bou (91)+bou (95)+bou (96) ) ) /76 ;
175 esc = ( esc1+esc2 ) /2 ;
176 if ( average (j-1) < esc )
177     average (j-1) = esc ;
178     clss (j-1)=5 ;

```

```

179         corr(:,j)=bou';
180     end
181
182     rnsn=0;
183     for (dd=1:valid) bou(dd)=ress_6(dd).CorrectRate; rnsn=rnsn+
184         ress_6(dd).CorrectRate; end
185     corr6(:,j)=bou';
186     fsc1 = (bou(2)+bou(3)+bou(9)+bou(11)+bou(14)+bou(15)+bou(25)+
187         bou(37)+bou(41)+bou(43)+bou(63)+bou(73)+bou(74)+bou(76)+bou
188         (81)+bou(86)+bou(87)+bou(91)+bou(95)+bou(96))/20;
189     fsc2 = (rnsn-(bou(2)+bou(3)+bou(9)+bou(11)+bou(14)+bou(15)+bou
190         (25)+bou(37)+bou(41)+bou(43)+bou(63)+bou(73)+bou(74)+bou
191         (76)+bou(81)+bou(86)+bou(87)+bou(91)+bou(95)+bou(96)))/76;
192     fsc = (fsc1+fsc2)/2;
193     if (average(j-1) < fsc)
194         average(j-1) = fsc;
195         clss(j-1)=6;
196         corr(:,j)=bou';
197     end
198
199     rnsn=0;
200     for (dd=1:valid) bou(dd)=ress_7(dd).CorrectRate; rnsn=rnsn+
201         ress_7(dd).CorrectRate; end
202     corr7(:,j)=bou';
203     gsc1 = (bou(2)+bou(3)+bou(9)+bou(11)+bou(14)+bou(15)+bou(25)+
204         bou(37)+bou(41)+bou(43)+bou(63)+bou(73)+bou(74)+bou(76)+bou
205         (81)+bou(86)+bou(87)+bou(91)+bou(95)+bou(96))/20;
206     gsc2 = (rnsn-(bou(2)+bou(3)+bou(9)+bou(11)+bou(14)+bou(15)+bou
207         (25)+bou(37)+bou(41)+bou(43)+bou(63)+bou(73)+bou(74)+bou
208         (76)+bou(81)+bou(86)+bou(87)+bou(91)+bou(95)+bou(96)))/76;
209     gsc = (gsc1+gsc2)/2;
210     if (average(j-1) < gsc)
211         average(j-1) = gsc;
212         clss(j-1)=7;
213         corr(:,j)=bou';
214     end
215
216     rnsn=0;
217     for (dd=1:valid) bou(dd)=ress_8(dd).CorrectRate; rnsn=rnsn+
218         ress_8(dd).CorrectRate; end
219     corr8(:,j)=bou';
220     hsc1 = (bou(2)+bou(3)+bou(9)+bou(11)+bou(14)+bou(15)+bou(25)+
221         bou(37)+bou(41)+bou(43)+bou(63)+bou(73)+bou(74)+bou(76)+bou
222         (81)+bou(86)+bou(87)+bou(91)+bou(95)+bou(96))/20;

```



```

210     hsc2 = (rns - (bou(2)+bou(3)+bou(9)+bou(11)+bou(14)+bou(15)+bou
          (25)+bou(37)+bou(41)+bou(43)+bou(63)+bou(73)+bou(74)+bou
          (76)+bou(81)+bou(86)+bou(87)+bou(91)+bou(95)+bou(96)))/76;
211     hsc = (hsc1+hsc2)/2;
212     if (average(j-1) < hsc)
213         average(j-1) = hsc;
214         class(j-1)=8;
215         corr(:,j)=bou';
216     end
217
218
219     end
220
221 end
222
223
224 end

```

```
1 % Copyright © Prevenios Marinos (2014). All rights reserved.
2
3 function [motion_mean] = find_mean_val()
4
5 load('HMM_2ndDataset.mat');
6
7 motion_mean=motion;
8
9 %%finding mean value for every waveform
10 for j=2:147
11
12     for i=1:96
13
14         sum=0;
15         n = length(motion{i,j});
16
17         for k=1:n % for every time unit
18
19             if (~isnan(motion{i,j}(k,1))) % avoiding NaN in training
20                 data
21
22                 sum = sum + motion{i,j}(k,1);
23
24             end
25         end
26
27         motion_mean{i,j}=sum/n;
28
29     end
30
31 end
32
33 end
```