



Τ.Ε.Ι. ΗΠΕΙΡΟΥ
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ ΤΗΛΕΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ

**ΑΥΤΟΜΑΤΗ ΔΙΑΓΝΩΣΗ ΕΠΙΛΗΨΙΑΣ ΜΕ
ΧΡΗΣΗ-ΑΝΑΛΥΣΗ
ΗΛΕΚΤΡΟΕΓΚΕΦΑΛΟΓΡΑΦΗΜΑΤΟΣ**

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΤΣΟΥΠΗ ΜΑΡΙΑ

ΑΜ: 9922

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ :

ΤΣΙΠΟΥΡΑΣ ΜΑΡΚΟΣ

Περιεχόμενα

1	Εισαγωγή.....	3
2	Ο ανθρώπινος εγκέφαλος.....	5
3	Επιληπτικές κρίσεις και τα αίτια τους.....	9
3.1	Γενικευμένες επιληπτικές κρίσεις.....	9
3.1.1	Ελαφράς μορφής κρίσεις ή μικρή επιληψία (petit mal):	9
3.1.2	Σοβαρής μορφής κρίσεις ή μεγάλη επιληψία (grand mal) :	10
3.2	Εστιακές επιληπτικές κρίσεις.....	11
3.2.1	Επιληψία κροταφικού λοβού.....	11
3.3	Αίτια επιληπτικών κρίσεων.....	12
4	Το Ηλεκτροεγκεφαλογράφημα και οι ρυθμοί του.....	13
4.1	Λειτουργία ΗΕΓ.....	15
4.2	Τοποθέτηση Ηλεκτροδίων.....	15
4.3	Ρυθμοί ΗΕΓ.....	16
4.4	Στάδια του ύπνου.....	19
5	Ηλεκτροεγκεφαλογράφημα και επιληψία.....	20
6	Λογισμικά που χρησιμοποιήθηκαν.....	23
6.1	WEKA (Waikato Environment for Knowledge Analysis)	23
6.1.1	<i>Naïve Bayes</i>	25
6.1.2	<i>k κοντινότερων γειτόνων (k Nearest Neighbor)</i>	26
6.1.3	<i>Δέντρα απόφασης (Decision Trees)</i>	28
6.1.4	<i>Τυχαίο Δάσος (Random Forest)</i>	29
6.1.5	<i>Πολυεπίπεδο Perceptron (Multiplayer Perceptron)</i>	31
6.2	MATLAB (matrix laboratory).....	35
7	Πειραματική Διάταξη.....	37
7.1	Διάγραμμα ροής.....	37
7.2	Βάση Δεδομένων.....	39
7.3	Σχεδιασμός και υλοποίηση φίλτρων.....	43
7.3.1	FIR (Finite Impulse Response) Filters-Πεπερασμένης Κρουστικής Απόκρισης Φίλτρα 46	
7.4	Εφαρμογή φίλτρων και εξαγωγή χαρακτηριστικών.....	47
7.5	Εφαρμογή classifiers και εξαγωγή αποτελεσμάτων στο weka	53
8	Αποτελέσματα-Συμπεράσματα.....	57
8.1	Αποτελέσματα.....	57
8.2	Συμπεράσματα.....	88
	ΒΙΒΛΙΟΓΡΑΦΙΑ	91

1 Εισαγωγή

Η παρούσα πτυχιακή εργασία αποτελεί μία προσπάθεια προσέγγισης της αυτόματης διάγνωσης μίας ασθένειας του εγκεφάλου, της επιληψίας, χρησιμοποιώντας το ηλεκτροεγκεφαλογράφημα (ΗΕΓ) για την καταγραφή και ανάλυση των σημάτων που παράγει ο εγκέφαλος.

Το ΗΕΓ αφορά την καταγραφή των διαφορών δυναμικών, τα οποία παρουσιάζονται στην εξωτερική δερματική επιφάνεια του ανθρώπινου κρανίου. Τα ηλεκτρικά σήματα προέρχονται από τους νευρώνες και είναι ασθενή, της τάξης των 10-300μV. Η μελέτη του ΗΕΓ βασίζεται στην ύπαρξη ή μη συγκεκριμένων κυματομορφών (των λεγόμενων ρυθμών) οι οποίοι διακρίνονται σύμφωνα με τις συχνότητες των αρμονικών από τις οποίες αποτελούνται. Οι κυριότεροι ρυθμοί που υπάρχουν στο ηλεκτροεγκεφαλογράφημα είναι: ο άλφα (8-13 Hz), ο βήτα (13-35 Hz), ο δέλτα (0.5-4 Hz) και ο θήτα (4-8 Hz). Το ηλεκτροεγκεφαλογράφημα συμβάλει σημαντικά στη μελέτη των νευρολογικών διαταραχών όπως στην πρόγνωση, διάγνωση και αντιμετώπιση της επιληψίας. Επίσης αποτελεί το κύριο μέσο στη μελέτη του ύπνου.

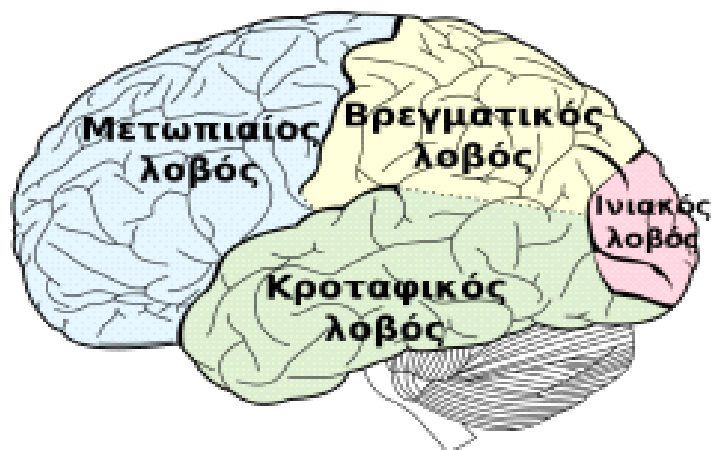
Η επιληψία, η οποία αποτελεί το αντικείμενο της πτυχιακής εργασίας, είναι η πιο συχνή διαταραχή του νευρικού συστήματος του ανθρώπινου εγκεφάλου. Αποτελεί μία οικογένεια νευρολογικών διαταραχών που χαρακτηρίζονται από επαναλαμβανόμενες επιληπτικές κρίσεις. Επιληπτική κρίση ορίζεται η προσωρινή διαταραχή του εγκεφάλου κατά την οποία μια ομάδα από νευρώνες αρχίζει να λειτουργεί ταυτόχρονα εκπέμποντας μη ομαλά ηλεκτρικά σήματα. Οι κρίσεις αυτές δυσκολεύουν σε μεγάλο βαθμό την καθημερινότητα του ανθρώπου και τον φθείρουν ψυχολογικά. Η καταγραφή των επιληπτικών κρίσεων δεν είναι εύκολη διαδικασία λόγω του ότι οι κρίσεις αυτές μπορεί να είναι σπάνιες και να εμφανίζονται απρόβλεπτα στις ΗΕΓ καταγραφές.

Στην παρούσα εργασία, αρχικά περιγράφεται η δομή του ανθρώπινου εγκεφάλου, ο ορισμός και τα χαρακτηριστικά της επιληψίας καθώς και μια ιστορική αναδρομή αντιμετώπισης της επιληψίας. Έπειτα, αναφέρονται οι μορφές των επιληπτικών κρίσεων, οι τύποι αυτών και τα αίτια που τις προκαλούν. Στο επόμενο κεφάλαιο γίνεται λεπτομερής περιγραφή του ηλεκτροεγκεφαλογράφου, η λειτουργία του, ο τρόπος που συνδέεται με την εξωτερική επιφάνεια του κρανίου και οι ρυθμοί του σύμφωνα με τους οποίους γίνεται η διάκριση. Στη συνέχεια αναφέρονται λίγα λόγια για τα λογισμικά που χρησιμοποιήθηκαν για την εξαγωγή χαρακτηριστικών. Στο έβδομο κεφάλαιο αναλύεται η πειραματική διαδικασία και παρουσιάζονται τα αποτελέσματα της, ενώ η πτυχιακή εργασία

ολοκληρώνεται με την παρουσίαση των αποτελεσμάτων από την εξόρυξη γνώσης, και των συμπερασμάτων που προκύπτουν.

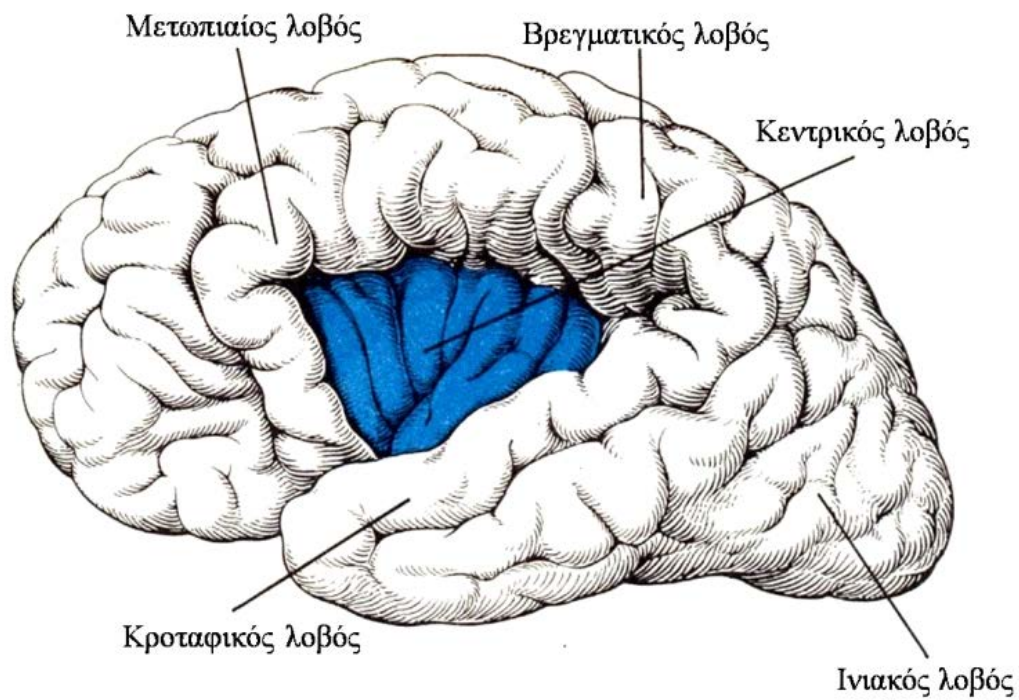
2 Ο ανθρώπινος εγκέφαλος

Ο εγκέφαλος αποτελεί το μεγαλύτερο και κυριότερο όργανο του κεντρικού νευρικού συστήματος του ανθρώπου. Αποτελείται από τέσσερις κοιλότητες που ονομάζονται κοιλίες. Βρίσκεται εντός του εγκεφαλικού κρανίου εξαιτίας της μεγάλης ευαισθησίας του. Έχει ωοειδές σχήμα με ένα εξόγκωμα προς την πίσω του μεριά. Περιβάλλεται από την φαία ουσία, που σχηματίζει τον φλοιό του εγκεφάλου και την λευκή ουσία που αποτελείται από νευρικές ίνες. Επίσης, περιβάλλεται από τρεις προστατευτικούς υμένες, τις μήνιγγες. Αυτές, από έξω προς τα μέσα, είναι η σκληρή μήνιγγα που βρίσκεται σε επαφή με την εσωτερική επιφάνεια του κρανίου και είναι ισχυρότερη, η αραχνοειδής μήνιγγα και η χοριοειδής μήνιγγα που περιβάλλει τον εγκέφαλο. Ο εγκέφαλος διαιρείται σε τρία κύρια μέρη: τα δύο ημισφαίρια, το στέλεχος και την παρεγκεφαλίτιδα. Κάθε ημισφαίριο αποτελείται από πέντε λοβούς: τον μετωπιαίο, τον βρεγματικό, τον ινιακό, τον κροταφικό και τον κεντρικό (ή νήσος του Reil). [1]



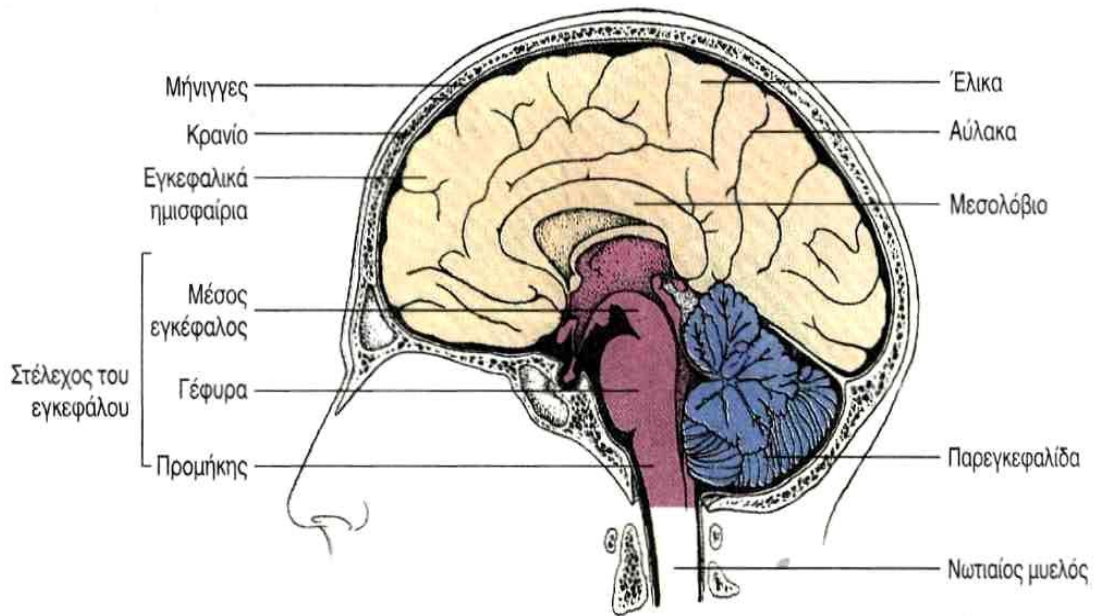
Εικόνα 2.1: Το αριστερό ημισφαίριο του εγκεφάλου χωρισμένο στους λοβούς.

Τα ημισφαίρια συνδέονται μεταξύ τους με τους συνδέσμους, οι οποίοι συνδέουν όμοια μέρη των δύο ημισφαιρίων. Οι σύνδεσμοι αυτοί είναι: το μεσολόβιο, που είναι και ο κυριότερος σύνδεσμος, ο πρόσθιος σύνδεσμος του εγκεφάλου, ο σύνδεσμος των υπόκαμπων, το διαφανές διάφραγμα και η ψαλίδα. Η εξωτερική επιφάνεια των ημισφαιρίων έχει πολλαπλές αυλακώσεις, τις αύλακες και ανάμεσά τους προεξοχές που ονομάζονται έλικες. Ανάλογα με το βάθος που έχουν οι ελικοειδείς αύλακες χωρίζουν τον εγκέφαλο σε λοβούς. [1]



Εικόνα 2.2: Ο κεντρικός λοβός του εγκεφάλου που φαίνεται αφαιρώντας μικρά τμήματα όλων των υπόλοιπων λοβών.

Το στέλεχος βρίσκεται στο οπίσθιο μέρος του εγκεφάλου, κάτω από τα δύο ημισφαίρια. Αποτελείται από τον διάμεσο και τον μέσο εγκέφαλο, την γέφυρα και τον προμήκη. Η παρεγκεφαλίτιδα βρίσκεται και αυτή στο οπίσθιο μέρος του εγκεφάλου, πίσω από το στέλεχος. Αποτελείται από δύο ημισφαίρια και τον σκώληκα που βρίσκεται ανάμεσά τους. [1]



Εικόνα 2.3: Δομή του εγκεφάλου, το δεξί ημισφαίριο όπως φαίνεται από την έσω επιφάνεια.

Στον εγκέφαλο εμφανίζονται διάφορα είδη μολύνσεων όπως η εγκεφαλίτιδα, απόστημα εγκεφάλου, υδροκεφαλισμός, όγκος στον εγκέφαλο, καρκίνος εγκεφάλου και παρεγκεφαλίτιδα.

Στην αρχαία Ελλάδα, η επιληψία θεωρούνταν ως μία υπερφυσική θεϊκή ή δαιμονική δύναμη. Δηλαδή, πίστευαν ότι οι διαταραχές από κάποια διανοητική ασθένεια ή φυσική παραμόρφωση οφείλονταν στην κατοχή του ανθρώπινου σώματος από κακά πνεύματα ή δαίμονες. Ως θεραπεία χρησιμοποιούσαν εξαγνιστικά και επικίνδυνα φάρμακα και θανάσιμα μέσα όπως η απελευθέρωση των δαιμόνων με διάτρηση του κρανίου. Πρώτος, ο Ιπποκράτης όρισε την ασθένεια ως ιατρικό πρόβλημα. Πίστευε πως δεν είναι ούτε κάποιο υπερφυσικό γεγονός, ούτε έχει κάτι το θεϊκό αλλά είναι μία ασθένεια σαν όλες τις άλλες και φυσικά υπάρχουν τρόποι για την αντιμετώπιση της. [2]

Στον Χριστιανισμό την αντιμετώπιζαν σαν δαιμονισμό και η θεραπεία της ήταν ο εξορκισμός. Υποστήριζαν ότι την επιληψία την "έστελνε" ο Θεός στους ανθρώπους, είτε για να τους τιμωρήσει εάν νοσούσαν σε μεγαλύτερη ηλικία, είτε για να τους δοκιμάσει εάν την είχαν εκ γενετής. Στο Βυζάντιο, οι απλοί άνθρωποι την αντιμετώπιζαν ως δαιμονισμό, αλλά οι γιατροί την θεωρούσαν ως ασθένεια με οργανική παθολογία του εγκεφάλου. Αργότερα, αποδείχθηκε πως η επιληψία δεν είναι ασθένεια αλλά σύμπτωμα. Η επιληψία σαν όρος σημαίνει αιφνίδια κατάληψη. Είναι μία χρόνια κατάσταση και η πιο συχνή διαταραχή της φυσιολογικής ηλεκτρικής δραστηριότητας των νευρικών κυττάρων του εγκεφάλου. Μία ομάδα από νευρικά κύτταρα (νευρώνες) αρχίζουν να λειτουργούν

ταυτόχρονα και να εκπέμπουν ξαφνικές εκρήξεις αυξημένης ενέργειας, δηλαδή ανώμαλα ηλεκτρικά σήματα. Αυτές είναι οι λεγόμενες επιληπτικές κρίσεις. Οι επιληπτικές κρίσεις εκδηλώνονται στον ασθενή με ασυνήθιστα αισθήματα και αντιλήψεις και ενίοτε σπασμούς στα άκρα, απώλεια συνείδησης, συσπάσεις μυών κ.α. [2,3]

Σήμερα, Βρετανοί επιστήμονες του πανεπιστημίου του Λίντς ανακάλυψαν το γονίδιο *Atp1a3* το οποίο συνδέεται με την επιληψία. Το γονίδιο αυτό παράγει ένα ένζυμο που ρυθμίζει τα επίπεδα του νατρίου και του καλίου στα εγκεφαλικά κύτταρα. Όταν υπάρχει περισσότερο νάτριο στον εγκέφαλο η αγωγιμότητα των νευρώνων αυξάνεται εκπέμποντας πολλά μη ομαλά ηλεκτρικά σήματα. Η ανισορροπία αυτή προκαλεί τις επιληπτικές κρίσεις. Οι επιστήμονες, σε εργαστηριακά πειράματα σε ποντίκια πάνω στο γονίδιο αυτό κατάφεραν να εμποδίσουν την ασθένεια. [4]

3 Επιληπτικές κρίσεις και τα αίτια τους

Οι επιληπτικές κρίσεις διακρίνονται σε δύο μεγάλες κατηγορίες στις γενικευμένες και στις εστιακές. Είναι γνωστοί πάνω από 40 τύποι επιληψίας που χωρίζονται στις δυο παραπάνω κατηγορίες. Στις γενικευμένες κρίσεις η αυξημένη ηλεκτρική δραστηριότητα διαδίδεται ταυτόχρονα και στα δύο ημισφαίρια του εγκεφάλου. Αυτή η ταυτόχρονη διάδοση της αυξημένης ηλεκτρικής δραστηριότητας σε όλον τον εγκέφαλο δημιουργεί είτε σοβαρής μορφής κρίσεις (ή μεγάλη επιληψία), είτε ελαφράς μορφής (ή μικρή επιληψία). [5]

3.1 Γενικευμένες επιληπτικές κρίσεις

3.1.1 Ελαφράς μορφής κρίσεις ή μικρή επιληψία (petit mal):

Είναι μια παροδική διαταραχή της συνείδησης. Στον ασθενή δεν φαίνεται κάποια κίνηση των μυών που να μην είναι φυσική και δεν πέφτει κάτω. Χωρίζεται σε τρεις τύπους κρίσεων :

α) Οι μυοκλονικές κρίσεις, που είναι ξαφνικές στιγμιαίες συσπάσεις μυών που αποτελούνται από απότομες δονήσεις ενός ή όλων των άκρων και ταυτόχρονη κάμψη του λαιμού και του κορμού.

β) Οι ακινητικές ή ατονικές κρίσεις, που χαρακτηρίζονται από ξαφνική απώλεια μυϊκού τόνου που οδηγεί το άτομο να πέσει στο έδαφος χωρίς να μπορεί να προστατευθεί με τις δικές του δυνάμεις από την πτώση. Σηκώνεται σχεδόν αμέσως αφού ανακτά γρήγορα του δυνάμεις του.

γ) Οι αφαιρέσεις, πρόκειται για στιγμιαίες απώλειες της συνείδησης, απότομη διακοπή κάθε είδους δραστηριότητα που πραγματοποιεί εκείνη την στιγμή, αδυναμία επαφής με το περιβάλλον και το πρόσωπο του ατόμου έχει χαμένη έκφραση και βλέμμα στο κενό. Αυτές οι κρίσεις σταματούν απότομα και το άτομο συνεχίζει την δραστηριότητα που έκανε χωρίς να αντιληφθεί κάτι από αυτό που συνέβη. Κάθε κρίση είναι μικρή σε διάρκεια (περίπου 1 δευτερόλεπτο) και συναντώνται συνήθως σε παιδιά. Αυτές οι μικρές κρίσεις μπορούν να επαναληφθούν πολλές φορές μέσα σε μία μέρα αλλά και διαδοχικά. Οι αφαιρέσεις χωρίζονται σε 6 τύπους : τις τυπικές, τις απλές, τις άτυπες, τις αφαιρέσεις με κλονικά στοιχεία, αφαιρέσεις με τονικά στοιχεία και αφαιρέσεις με ατονικά στοιχεία.

δ) Οι τονικοκλονικές κρίσεις, χαρακτηρίζονται από αιφνίδια έναρξη χωρίς πρόδρομα σημεία και στην συνέχεια εξελίσσονται σε τρεις φάσεις: την τονική,

ακολουθεί η κλονική και η φάση λύσης.

- Η τονική φάση ξεκινά με κραυγή, πλήρη κατάργηση της συνείδησης, γενικευμένο σφίξιμο του σώματος και το άτομο συνήθως πέφτει κάτω. Η τονική φάση διαρκεί περίπου 10-20 δευτερόλεπτα.

- Στην συνέχεια ακολουθεί η κλονική φάση κατά την οποία συμβαίνουν ρυθμικές, έντονες συσπάσεις όλων των άκρων οι οποίες αραιώνουν σιγά σιγά για να διακοπούν αυτόματα. Τα μάτια είναι ανοιχτά και γυρισμένα προς τα πάνω (βολβοστροφή), το στόμα σφιχτά κλειστό, η αναπνοή βγαίνει με δυσκολία, και τα χείλη και το πρόσωπο μελανιάζουν λόγω πτωχής οξυγόνωσης. Αυτή η φάση ονομάζεται και επεισόδιο σπασμών ή κρίση τύπου grand mal και διαρκεί πολύ λίγο, περίπου 30 δευτερόλεπτα.

- Αμέσως μετά την κλονική φάση επέρχεται η φάση λύσης. Διαρκεί από μερικά λεπτά έως μερικές δεκάδες λεπτών. Το άτομο αρχίζει να συνέρχεται αλλά είναι υποτονικό και ακίνητο και παρουσιάζει βαθιά θόλωση της συνείδησης και πλήρη μυϊκή χαλάρωση. Συνήθως θέλει να κοιμηθεί και όταν ξυπνά είναι καλά ή μπορεί να έχει πονοκέφαλο. Η αναπνοή επανέρχεται βαριά και θορυβώδης. Η φάση αυτή ονομάζεται μετακρισική φάση. [5]

3.1.2 Σοβαρές μορφές κρίσεις ή μεγάλη επιληψία (grand mal) :

Είναι μεγάλες κρίσεις που μπορούν να συμβούν απροειδοποίητα. Συνήθως όμως οι ασθενείς που προσβάλλονται από αυτές τις κρίσεις, νιώθουν πριν από την κρίση, μία αύρα, δηλαδή αισθήσεις όπως μυρωδιές, γεύσεις ή ακόμα και λάμψεις. Μετά από αυτό, επέρχεται ξαφνική απώλεια της συνείδησης και αν το άτομο στέκεται όρθιο πέφτει στο έδαφος. Η κρίση αυτή έχει τρεις φάσεις :

a) Η άκαμπτη τονική φάση κατά την οποία το σώμα είναι σε ακαμψία, ο ασθενής έχει τονικό μυϊκό σπασμό δηλαδή σπασμωδικές εκτινάξεις και τέντωμα του σώματος. Όλο το μυϊκό σύστημα βρίσκεται σε σύσπαση μπλοκάροντας τις αναπνευστικές κινήσεις, σταματώντας την αναπνοή και προκαλώντας μελάνιασμα (κυάνωση) του προσώπου.

b) Η κλονική φάση, που ακολουθεί αμέσως μετά την τονική φάση. Η αναπνοή γίνεται θορυβώδης και σπασμωδική. Οι σπασμωδικές εκτινάξεις των άκρων μειώνονται βαθμιαία και γίνονται δονήσεις και σε πολύ μικρό χρονικό διάστημα απομένει ατονία σε όλο το σώμα.

c) Η κωματώδης φάση από την οποία ο ασθενής βγαίνει σιγά σιγά από αυτήν αλλά

παραμένει αποπροσανατολισμένος για λίγη ώρα.

3.2 Εστιακές επιληπτικές κρίσεις

Στις εστιακές κρίσεις η αυξημένη ηλεκτρική δραστηριότητα περιορίζεται σε συγκεκριμένο τμήμα του εγκεφάλου. Ανάλογα με την περιοχή του εγκεφάλου στην οποία συμβαίνει η ηλεκτρική διέγερση των νευρώνων οι εκδηλώσεις της επιληπτικής κρίσης μπορεί να είναι κινητικές όπως ρυθμικές συσπάσεις των άκρων, αισθητικές όπως μούδιασμα ενός άκρου ή μιας πλευράς του σώματος, αισθητηριακές όπως αλλοιωμένη όραση, ακοή, ή όσφρηση, ή ψυχικές όπως παραισθήσεις, σύγχυση και περίεργα αισθήματα. Οι εστιακές κρίσεις αποτελούν το 60% όλων των περιπτώσεων της επιληψίας ενηλίκων και η πιο συχνή μορφή είναι η επιληψία κροταφικού λοβού. [5]

3.2.1 Επιληψία κροταφικού λοβού

Η επιληψία κροταφικού λοβού είναι μία χρόνια νευρολογική κατάσταση που χαρακτηρίζεται από επαναλαμβανόμενες κρίσεις που προκύπτουν από τον έναν ή και τους δύο κροταφικούς λοβούς του εγκεφάλου. Σύμφωνα με τη Διεθνή Ένωση κατά της επιληψίας έχουν αναγνωρισθεί διεθνώς δυο τύποι κροταφικής επιληψίας:

- 1) Η επιληψία μέσου κροταφικού λοβού, η οποία ελκύεται από τον ιππόκαμπο, την παραϊπποκαμποειδή έλικα και την αμυγδαλή της έσω επιφάνειας του κροταφικού λοβού. Κάποιες μελέτες έχουν αναδείξει ανωμαλίες του ιππόκαμπου σε επιληπτική κατάσταση που έχουν δημιουργηθεί από παρατεταμένες και επαναλαμβανόμενες επιληπτικές κρίσεις. Οι κρίσεις αυτές μπορούν να ξεκινήσουν από μία βλάβη στον αριστερό ή δεξί ιππόκαμπο. Παρουσιάζεται σε βρέφη από πυρετό λόγω κάποιας λοίμωξης πνευμονικής ή δερματικής και σε παιδιά με την μορφή πυρετικών σπασμών.
- 2) Η επιληψία πλάγιου κροταφικού λοβού, η οποία ελκύεται από τον νεοφλοιό της εξωτερικής επιφάνειας του κροταφικού λοβού του εγκεφάλου. Είναι λιγότερο συχνή και μπορεί να είναι κληρονομική. Η πλάγια κροταφική επιληψία με διαταραχές στην όραση ή την ακοή σχετίζεται με όγκους στον εγκέφαλο, μηνιγγίτιδα, εγκεφαλίτιδα, κάποιο τραύμα, αγγειακές δυσπλασίες ή συγγενείς δυσπλασίες του εγκεφάλου και σε άλλα αίτια που δεν μπορούν πάντα να προσδιοριστούν. [6]

Υπάρχουν τρεις τύποι κροταφικής επιληψίας :

- Απλές εστιακές επιληπτικές κρίσεις, αφορούν μικρές περιοχές του κροταφικού λοβού όπως η αμυγδαλή και ο ιππόκαμπος. Σε αυτές τις κρίσεις προκαλούνται παραισθήσεις στο άτομο που μπορεί να είναι ακουστικές, οσφρητικές,

γευστικές, οπτικές ή αφής. Μπορεί να νιώθει δυσφορία ή αισθήματα ευφορίας, φόβο, θυμό και άλλα αισθήματα. Αυτή είναι και η λεγόμενη αύρα δηλαδή το αίσθημα προειδοποίησης που έχει ο ασθενής πριν πάθει κάποια κρίση. Κατά την διάρκεια μίας απλής εστιακής κρίσης η ικανότητα του ατόμου να αντιλαμβάνεται τι γίνεται γύρω του και να επικοινωνεί με το περιβάλλον δεν επηρεάζεται καθόλου, οπότε έχει πλήρη συναίσθηση για το τι συμβαίνει και μπορεί να το περιγράψει αργότερα με ακρίβεια.

- Σύνθετες εστιακές επιληπτικές κρίσεις, το επίπεδο συνείδησης επηρεάζεται σε μικρότερο ή μεγαλύτερο βαθμό και έτσι ελαττώνεται η ικανότητα του ατόμου να αλληλεπιδρά με το περιβάλλον. Ξεκινούν με απλή επιληπτική κρίση και στην συνέχεια η ηλεκτρική εκφόρτιση εξαπλώνεται σε μεγαλύτερο τμήμα του κροταφικού λοβού με αποτέλεσμα την διαταραχή της συνείδησης. Το άτομο μπορεί να κινείται χωρίς κάποιο σκοπό ή να κάνει κάποιες στερεότυπες αυτόματες κινήσεις που γίνονται τελείως ασυνείδητα όπως κινήσεις ματιών, χεριών, στόματος, ασυνήθιστη συμπεριφορά ή ομιλία. Αντίθετα με τις απλές εστιακές κρίσεις, ο ασθενής έχει πολύ περιορισμένη αντίληψη των γεγονότων και δεν μπορεί αργότερα να περιγράψει τι του συνέβη.
- Κρίσεις που αρχίζουν στον κροταφικό λοβό, αλλά στη συνέχεια εξαπλώνονται σε ολόκληρο τον εγκέφαλο είναι ευρέως γνωστές ως Δευτερεύουσες Γενικευμένες Τονικοκλονικές κρίσεις (SGTCS). Αυτές αρχίζουν με μία απλή εστιακή κρίση και στην συνέχεια ο κορμός και τα πόδια παραμένουν σε κάμψη και ακολουθούν σπασμωδικές κινήσεις των άκρων. [6]

3.3 Αίτια επιληπτικών κρίσεων

Μία επιληπτική κρίση μπορεί να οφείλεται σε εγκεφαλική αιμορραγία, χτύπημα στο κεφάλι, ισχαιμικό αγγειακό εγκεφαλικό επεισόδιο, λήψη ναρκωτικών ουσιών, κατάχρηση αλκοολούχων ποτών, όγκο εγκεφάλου, μηνιγγίτιδα, νόσο Alzheimer, μεταβολικές διαταραχές όπως η υπογλυκαιμία κ.α. Τέτοιου είδους κρίσεις ονομάζονται δευτεροπαθείς ή συμπτωματικές. Οι κρίσεις όπου δεν ανευρίσκεται κάποιο αίτιο ονομάζονται ιδιοπαθείς ή κεντροεγκεφαλικές και φαίνεται να παίζει κάποιο ρόλο στην εκδήλωσή τους, η κληρονομικότητα.

Τα ποσοστά κληρονομικότητας στην πρωτοπαθή γενικευμένη επιληψία είναι:

- 1-2% αν ο πατέρας είναι επιληπτικός,
- < 5% αν η μητέρα είναι επιληπτική και
- > 5% αν είναι και οι δύο γονείς επιληπτικοί.

4 Το Ηλεκτροεγκεφαλογράφημα και οι ρυθμοί του

Το κεντρικό νευρικό σύστημα αποτελείται από ένα σύνολο σχηματισμών που έχουν διαφορετικές λειτουργίες και διαφορετική ανατομική σύσταση. Σ' αυτές τις λειτουργικές και ανατομικές διαφορές αναλογούν επίσης και διαφορές στις διαγνωστικές μεθόδους, όσον αφορά την κλινική και εργαστηριακή έρευνα. Έτσι η ηλεκτρική ενέργεια του εγκεφάλου, που είναι το άθροισμα της ενέργειας των εκατομμυρίων νευρώνων, μελετάται διαμέσου της καταγραφής των βιοχημικών ηλεκτρικών δυναμικών με την μέθοδο του ηλεκτροεγκεφαλογραφήματος.

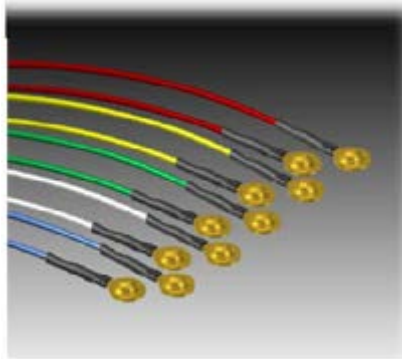
Το ηλεκτροεγκεφαλογράφημα αναφέρεται στην αυθόρμητη ηλεκτρική ρυθμική ταλάντωση του εγκεφάλου όταν βρίσκεται σε ηρεμία και δεν δέχεται εξωτερικά ερεθίσματα. Επινοήθηκε από τον Γερμανό νευρολόγο-ψυχίατρο H. Berger το 1929. Ήταν ο πρώτος που κατάφερε να καταγράψει την αυτόματη ηλεκτρική δραστηριότητα του εγκεφάλου σε χάρτινη ταινία. Είναι μια εξέταση η οποία χρησιμοποιείται για την διάγνωση επιληψίας, ημικρανιών, άνοιας και Alzheimer. Αποτελεί μια τεχνική καταγραφής, μέτρησης, επεξεργασίας και ερμηνείας της ηλεκτρικής δραστηριότητας του εγκεφάλου. Δεν έχει μόνο την άμεση διαγνωστική ικανότητα αλλά μπορεί ακόμα να προβλέψει την εμφάνιση εγκεφαλικών διαταραχών 5-10 χρόνια πριν από την εμφάνισή τους.

Η εξέταση αυτή γίνεται με την βοήθεια μίας ηλεκτρικής συσκευής, τον ηλεκτροεγκεφαλογράφο. Ο ηλεκτροεγκεφαλογράφος αποτελείται από έναν πίνακα συνδεσμολογίας ηλεκτροδίων που εφαρμόζονται στην επιφάνεια του εγκεφάλου και κατ' επέκταση στο δέρμα του τριχωτού της κεφαλής. Τα ηλεκτρόδια αυτά είναι επαναχρησιμοποιούμενα και συνήθως κατασκευάζονται από άργυρο και χλωριούχο άργυρο. Υπάρχουν τρεις κατηγορίες ηλεκτροδίων : [1]

- τα προσκεφαλοειδή, τα οποία έχουν δισκοειδή απόληξη που είναι καλυμμένη με σφουγγάρι ή βαμβάκι τυλιγμένο με γάζα και υποστηρίζεται από έναν πλαστικό αντίποδα



- τα επικολλώμενα, που και αυτά έχουν δισκοειδή απόληξη και στερεώνονται στο κεφάλι με μία αγώγιμη κρέμα ή κολλώδη ουσία παρόμοιας σύστασης με τα σωματικά υγρά- τους ηλεκτρολύτες



- τα βελλονοειδή, που έχουν σχήμα μακριάς βελόνας, χρησιμοποιούνται σε ειδικές περιοχές του εγκεφάλου και εφαρμόζονται είτε με τοπική αναισθησία είτε με ολική αναισθησία.



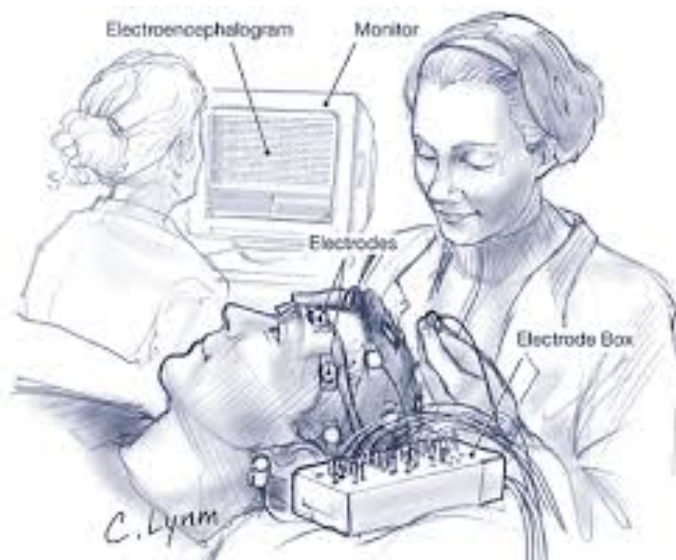
Για να τοποθετηθούν τα ηλεκτρόδια στην επιφάνεια του κρανίου αρχικά διαχωρίζονται οι τρίχες και καθαρίζεται το δέρμα με καθαρό οινόπνευμα για την απομάκρυνση των λιπαρών ουσιών του δέρματος. [7]

Επίσης, ο ΗΕΓράφος εκτός από τα ηλεκτρόδια αποτελείται από έναν επιλογέα απαγωγών, προενισχυτές, διατάξεις φιλτραρίσματος, την τελική διάταξη ενισχυτών, μία συσκευή πολυπλεξίας, έναν μετατροπέα αναλογικού σήματος σε ψηφιακό και σε έναν ηλεκτρονικό υπολογιστή αποδίδει την διαφορά δυναμικού της αυτόματης ηλεκτρικής δραστηριότητας του εγκεφάλου, υπό την μορφή κυμάτων. [1,8]

Για να εξασφαλιστεί καλή καταγραφή των κυμάτων για καλύτερη απεικόνιση είναι απαραίτητη η ενίσχυση των σημάτων λόγω του ότι τα ηλεκτρικά σήματα του εγκεφάλου εξασθενούν κατά 10 φορές τουλάχιστον από παρεμβολές μεταξύ των ηλεκτροδίων και του εγκεφάλου που έχει ένα στρώμα πάχους 2-3 εκατοστών. [7]

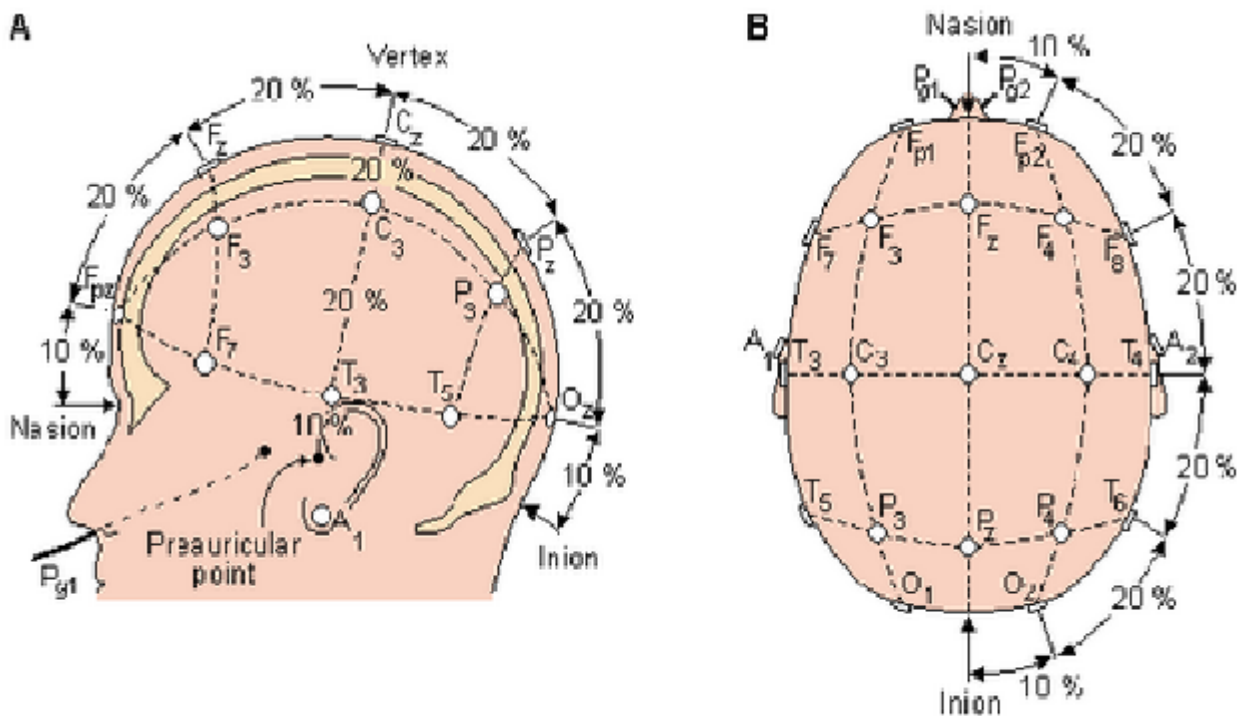
4.1 Λειτουργία ΗΕΓ

Τα ηλεκτρόδια μετατρέπουν το ρεύμα ιόντων των εγκεφαλικών κυττάρων σε ηλεκτρικό ρεύμα που οδηγείται στους αρχικούς ενισχυτές διάταξης (προενισχυτές). Στη συνέχεια οδηγείται στην διάταξη φιλτραρίσματος όπου υπάρχουν τα φίλτρα αποκοπής συχνοτήτων. Εκεί κατασκευάζονται ηλεκτρονικές διατάξεις που χρησιμοποιούνται για να επιλέξουν ή να περιορίσουν το φάσμα συχνοτήτων αφαιρώντας θορύβους που παρεμβάλλονται και παραμορφώνουν το σήμα. Μέσω της συσκευής της πολυπλεξίας οδηγείται στον μετατροπέα αναλογικού σήματος σε ψηφιακό. Το ψηφιακό αυτό σήμα που παράγεται από τον μετατροπέα καταμετράται σε ηλεκτρονικό βολτόμετρο. Ένας ηλεκτρονικός υπολογιστής λαμβάνει τα ψηφιακά δεδομένα των μετρήσεων και έτσι υπάρχει η δυνατότητα για επεξεργασία των δεδομένων αυτών και απεικόνιση του σήματος στην οθόνη ή και εκτύπωση αυτής σε χαρτί. Επίσης μπορεί να αποθηκευτεί το σήμα στον σκληρό δίσκο του υπολογιστή ή σε κάποια εξωτερική μονάδα αποθήκευσης για περαιτέρω μελέτη και σύγκριση άλλου σήματος του ίδιου ασθενή που βρίσκεται υπό παρακολούθηση. [8]



4.2 Τοποθέτηση Ηλεκτροδίων

Η τοποθέτηση του κάθε ηλεκτροδίου στο δέρμα του κεφαλιού δεν είναι τυχαία γιατί ανάλογα από την θέση εξαρτάται η συχνότητα, η φάση και το πλάτος του σήματος που θα καταγραφεί. Για την επιλογή των θέσεων καθενός από αυτά έχουν δημιουργηθεί διάφορα πρότυπα. Το πιο διαδεδομένο σύστημα τοποθέτησης είναι η διάταξη κατά Jaspers ή 10-20. Η ονομασία του συστήματος οφείλεται στην επιλογή του 20% της αποστάσεως μεταξύ των δύο αυτιών ως την απόσταση ανάμεσα σε δύο οποιαδήποτε ηλεκτρόδια και στην επιλογή του 10% της αποστάσεως μεταξύ των δύο αυτιών ως την απόσταση από το κάθε αυτί στο πλησιέστερο προς αυτό ηλεκτρόδιο του. Το πρότυπο αυτό καθιερώθηκε από την διεθνή ομοσπονδία των ενώσεων ΗΕΓ. [1]



Εικόνα 4.1: Τοποθέτηση ηλεκτροδίων στο τριχωτό της κεφαλής σύμφωνα με το Διεθνές σύστημα 10-20 όπως φαίνεται από Α) πλάγια αριστερά και Β) πάνω από το κεφάλι.

Μέχρι σήμερα έχουν προταθεί διάφορες επεκτάσεις του μοντέλου 10-20 για την αύξηση των παρεχόμενων καναλιών ούτως ώστε να έχουμε περισσότερες διαφορές δυναμικού μεταξύ δύο ηλεκτροδίων και άρα καλύτερο σήμα. Κάποιες από αυτές είναι τα πρότυπα 10-10 και 10-5 στα οποία παρεμβάλλονται ηλεκτρόδια ανάμεσα στις θέσεις του συστήματος 10-20. Τοποθετούνται επίσης και απαγωγές στους λοβούς των αυτιών και σε θέσεις κοντά στα μάτια. Το πρότυπο 10-5 βρίσκεται ακόμα σε πειραματικό στάδιο, ενώ το 10-10 έχει γίνει αποδεκτό και έχει τεκμηριωθεί από την Αμερικάνικη Ηλεκτροεγκεφαλική Κοινότητα (American Electroencephalographic Society) και χρησιμοποιείται ευρέως σήμερα. Η ευρεία χρησιμοποίησή του οφείλεται στο γεγονός ότι προσαρμόζεται σε διάφορες διαστάσεις κεφαλιών. [8]

Η μέθοδος του ηλεκτροεγκεφαλογράφηματος είναι φθηνή, απλή και ανώδυνη στην εφαρμογή της. προσφέρει μοναδική βοήθεια στον κλάδο της ιατρικής για διάγνωση και μελέτη επιληψίας, ψυχικών διαταραχών, άνοιας, οργανικών παθήσεων του εγκεφάλου και παρακολούθηση εγκεφαλοπαθειών ανεξαρτήτως αιτίας.

4.3 Ρυθμοί ΗΕΓ

Σύμφωνα με το φάσμα των συχνοτήτων των αρμονικών από τις οποίες αποτελούνται οι κυματομορφές του σήματος διακρίνονται οι λεγόμενοι ρυθμοί. Οι κυριότεροι ρυθμοί που

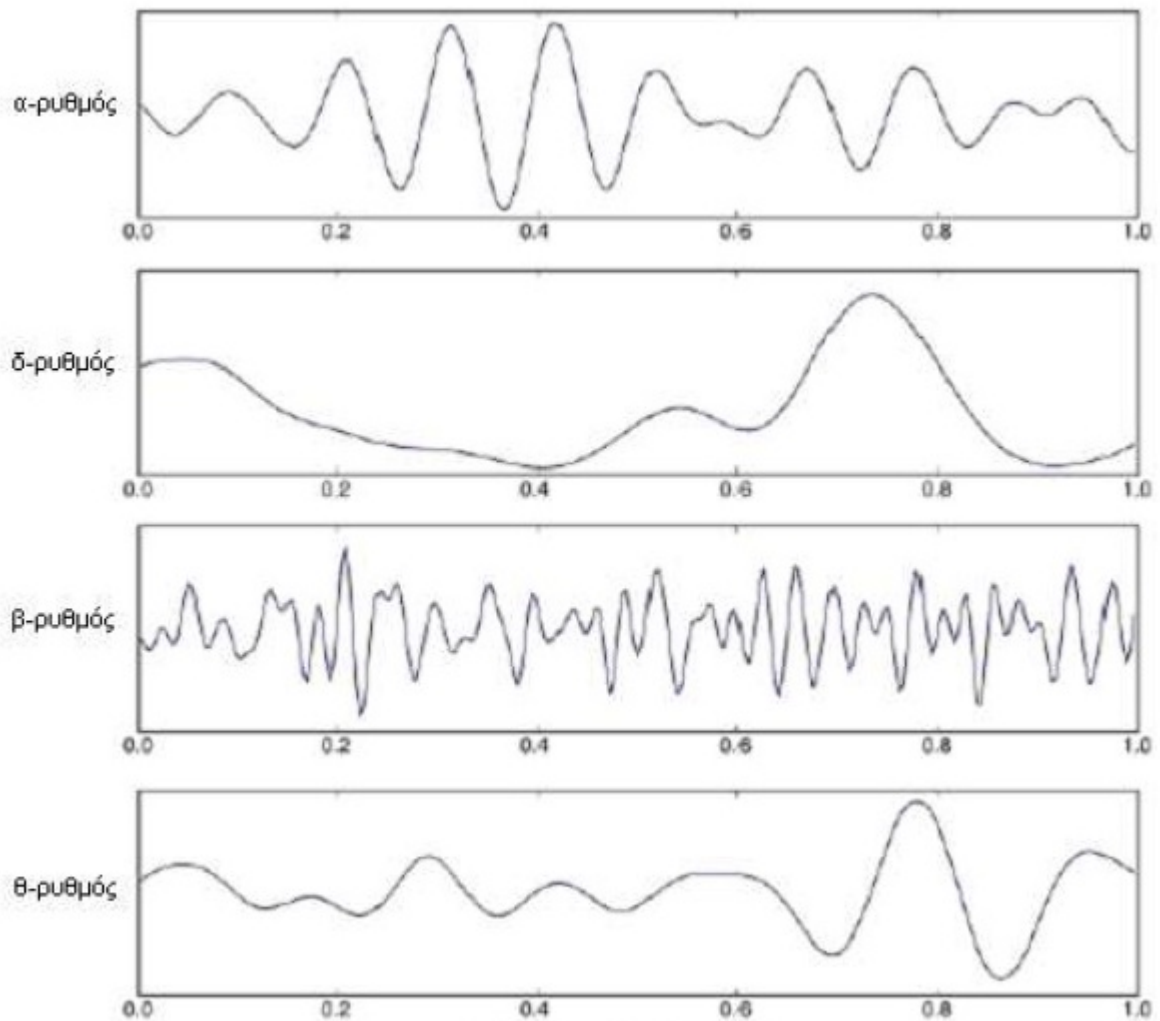
υπάρχουν και έχουν αναγνωρισθεί ανάλογα με το τμήμα του φάσματος συχνοτήτων είναι οι : άλφα, βήτα, δέλτα και θήτα.

Χαμηλής Συχνότητας ≤ 4 Hz – Ρυθμός Δέλτα: Καταγράφεται φυσιολογικά στους ενήλικες κατά τα στάδια του βαθύ ύπνου 3 και 4, ενώ σε κατάσταση εγρήγορσης υποδηλώνει βαριά εγκεφαλική βλάβη. Είναι φυσιολογικό να εντοπιστεί μόνο σε νεογέννητα έως και το δεύτερο έτος της ηλικίας τους, σε κατάσταση εγρήγορσης.

Συχνότητας 4 έως 8 Hz – Ρυθμός Θήτα: Θεωρείται παθολογικός ρυθμός και εμφανίζεται σε ενήλικες που βρίσκονται σε εγρήγορση. Παρατηρείται σε περιπτώσεις εγκεφαλικής βλάβης και σε παθολογικές καταστάσεις όπως η επιληψία. Στα παιδιά μέχρι 6 ετών θεωρείται φυσιολογικός.

Συχνότητας 8 έως 13 Hz – Ρυθμός Άλφα: Εντοπίζεται όταν είμαστε ξύπνιοι σε χαλάρωση και με τα μάτια κλειστά. Το άνοιγμα ή το κλείσιμο των ματιών προκαλεί μείωση ή αύξηση του ρυθμού, αντίστοιχα.

Συχνότητας 13 έως 35 Hz – Ρυθμός Βήτα: Συναντάται σε όλους τους ενήλικες και εμφανίζεται σε κατάσταση εγρήγορσης όταν είμαστε ξύπνιοι. Δεν εμφανίζεται κατά την διάρκεια του ύπνου. Ο ρυθμός αυτός συναντάται αυξημένος σε λειτουργίες όπως η νοητική συγκέντρωση, η διατήρηση της προσοχής αλλά και σε καταστάσεις άγχους. [1]



Εικόνα 4.2 : Ρυθμοί ηλεκτροεγκεφαλογράφηματος.

Οι ρυθμοί άλφα και βήτα θεωρούνται φυσιολογικοί ενώ οι ρυθμοί δέλτα και θήτα παθολογικοί στους ενήλικες. Οι ρυθμοί δέλτα και θήτα δεν θεωρούνται παθολογικοί στα παιδιά, στον ύπνο και κατά την υπέρπνοια.

Σε κάθε επιληπτικό ηλεκτροεγκεφαλογράφημα γίνεται διαχωρισμός των χρονικών περιόδων ανάλογα με την μορφή που έχουν οι κυματομορφές, δηλαδή τα αιχμηρά κύματα και οι αιχμές που παρατηρούνται στις καταγραφές του ηλεκτροεγκεφαλογράφηματος. Η χρονική περίοδος κατά την οποία εκδηλώνεται η επιληπτική κρίση ονομάζεται ictal (κριτική-επιληπτική περίοδος). Το χρονικό διάστημα πριν από την έναρξη της επιληπτικής κρίσης ή ανάμεσα από δυο επιληπτικές κρίσεις ονομάζεται interictal (μεσοκριτική περίοδος), δηλαδή ότι αυτή την χρονική περίοδο δεν υπάρχουν κλινικές εκδηλώσεις επιληπτικών κρίσεων. [1]

4.4 Στάδια του ύπνου

Ο ύπνος χαρακτηρίζεται από την εξαφάνιση του άλφα ρυθμού και εμφανίζεται ο βραδύτερος ρυθμός, δηλαδή ο ρυθμός δέλτα και ο θήτα.

Στάδιο 1 : Στο πρώτο στάδιο, που είναι το στάδιο του ελαφριού ύπνου, ξυπνάμε πολύ εύκολα. Βρισκόμαστε μεταξύ ύπνου και ξύπνιου. Οι κινήσεις των μυών μειώνονται και τα μάτια κινούνται πολύ αργά. Στο στάδιο αυτό νιώθουμε ξαφνικές συσπάσεις των μυών (μυοκλονίες του ύπνου) σαν αυτές που παθαίνουμε όταν τρομάξουμε από κάτι.

Στάδιο 2 : Στο στάδιο αυτό δεν υπάρχει μυϊκή δραστηριότητα και τα μάτια σταματούν να κινούνται. Στην καταγραφή του ΗΕΓ τα κύματα γίνονται πιο αργά σε σχέση με αυτά του σταδίου 1.

Στάδιο 3 : Μπαίνουμε στο στάδιο του βαθύ ύπνου. Τα ηλεκτρικά κύματα που καταγράφονται στο ΗΕΓ είναι πάρα πολύ αργά δηλαδή εμφανίζεται ο δέλτα ρυθμός.

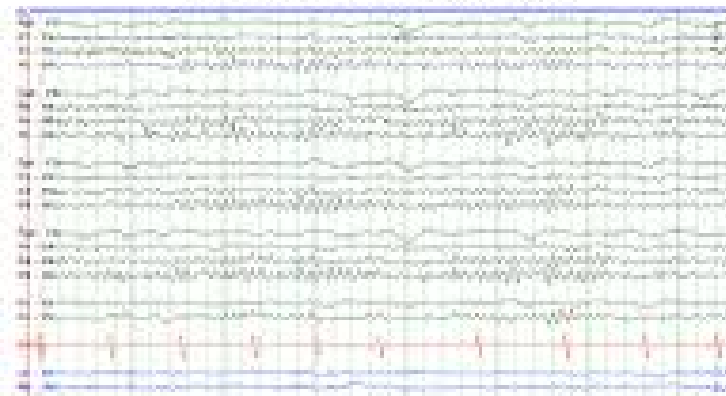
Στάδιο 4 : Και το στάδιο αυτό είναι το στάδιο του βαθύ ύπνου. Στον βαθύ ύπνο είναι πολύ δύσκολο να ξυπνήσει κάποιος. Αν κάποιος ξυπνήσει είναι αδύναμος και δεν μπορεί να προσαρμοστεί στην πραγματικότητα. Πολλοί άνθρωποι υπνοβατούν και βλέπουν εφιάλτες σ' αυτό το στάδιο.

Στάδιο 5 : Στο στάδιο αυτό η μυϊκή δραστηριότητα των άκρων παραλύει, τα μάτια κινούνται γρήγορα και απότομα, η αναπνοή γίνεται γρήγορη και ακανόνιστη, ο ρυθμός της καρδιάς και η αρτηριακή πίεση αυξάνονται. Το 5ο στάδιο είναι απαραίτητο για την ξεκούραση. Εδώ βλέπουμε και τα όνειρα.

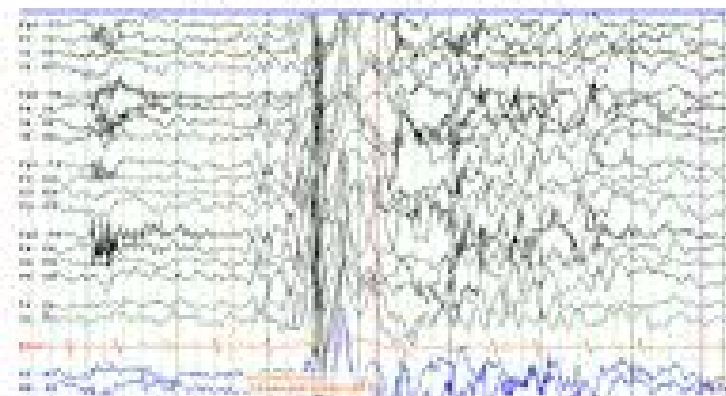
5 Ηλεκτροεγκεφαλογράφημα και επιληψία

Το ΗΕΓ είναι η πιο χρήσιμη και σημαντική εξέταση για τη διάγνωση της επιληψίας. Ωστόσο, ένα φυσιολογικό ηλεκτροεγκεφαλογράφημα δεν αποκλείει την πιθανότητα επιληψίας, γιατί ανάμεσα σε κρίσεις η εικόνα των κυμάτων της εγκεφαλικής δραστηριότητας μπορεί να φαίνεται φυσιολογική. Έτσι, κατά τη διάρκεια ενός ΗΕΓ μπορεί να μην εμφανιστεί μη φυσιολογική δραστηριότητα και να χρειαστεί η εξέταση να επαναληφθεί παραπάνω από μία φορές. Σε ορισμένες περιπτώσεις μπορεί να εμφανιστεί επιληπτική δραστηριότητα του εγκεφάλου η οποία δεν εκδηλώνεται με τη μορφή κρίσεων, ωστόσο ανιχνεύεται από το ηλεκτροεγκεφαλογράφημα. Το ΗΕΓ είναι ένα πολύ χρήσιμο εργαλείο για τη διάγνωση και τον εντοπισμό του είδους των κρίσεων που λαμβάνουν χώρα.

Normal EEG Awake



Myoclonic Seizure



Εικόνα 5.1 : Πάνω φυσιολογικό ΗΕΓ και κάτω επιληπτόμορφο ΗΕΓ.

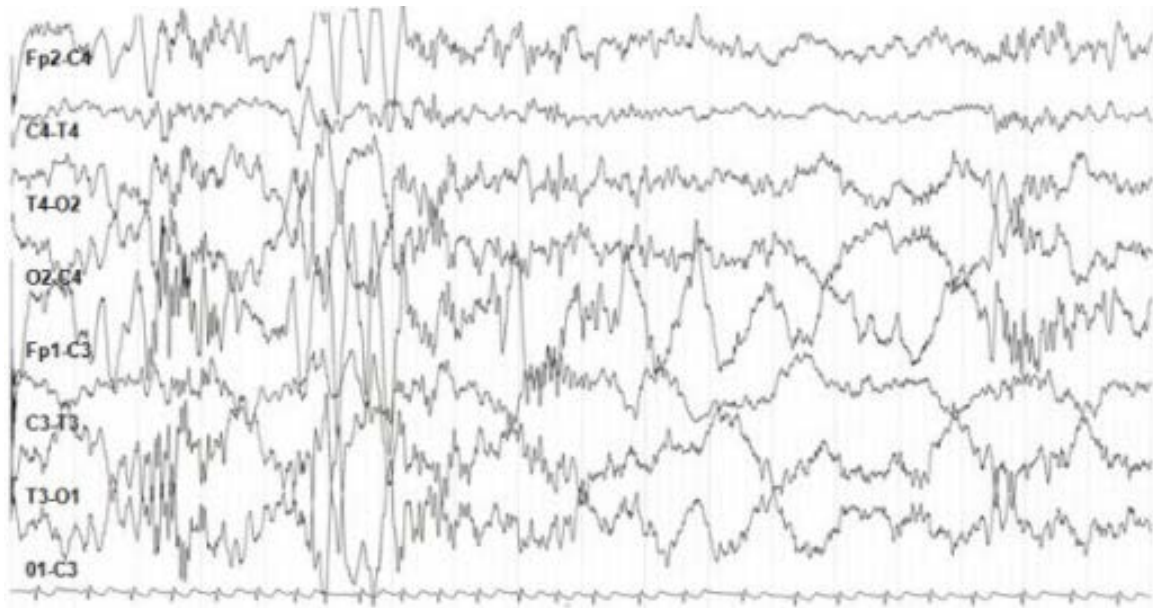
Η διάγνωση της επιληψίας με την βοήθεια του ΗΕΓ γίνεται με ανίχνευση ανωμαλιών σε σχέση με ένα φυσιολογικό ηλεκτροεγκεφαλογράφημα. Οι παρεκτροπές αυτές του παθολογικού ηλεκτροεγκεφαλογραφήματος έχουν κάποια δομή και επαφίεται στην

εξειδίκευση και εμπειρία του γιατρού να τις αναγνωρίσει και να τις ξεχωρίσει συγκριτικά με ένα φυσιολογικό ηλεκτροεγκεφαλογράφημα.



Εικόνα 5.2: Ηλεκτροεγκεφαλογράφημα επιφανείας επιληπτικού ασθενούς. Η κατακόρυφη γραμμή υποδεικνύει το δευτερόλεπτο έναρξης της κρίσης.

Κάποιες από αυτές τις δομές παρατηρούνται συνήθως σε περιόδους μεταξύ κρίσεων και τα χαρακτηριστικά τους ονόματα στην αγγλική βιβλιογραφία είναι «spikes» (αιχμές), «sharp waves» (αιχμηρά κύματα) και «spike and wave discharges» (αποφορτίσεις αιχμής κύματος). Για την καλύτερη μελέτη του ΗΕΓ ενός επιληπτικού ασθενή έχει γίνει από τους επιστήμονες ένας διαχωρισμός των χρονικών περιόδων στο επιληπτικό ΗΕΓ. Η περίοδος κατά την οποία εκδηλώνεται η κρίση ονομάζεται ictal (κριτική). Το χρονικό διάστημα που μεσολαβεί από την λήξη της επιληπτικής κρίσης μέχρι την έναρξη της επόμενης, ονομάζεται inter-ictal (μεσοκριτική). Χαρακτηρίζουμε ως περίοδο pre-ictal (προκριτική) τη μεταβατική περίοδο από την inter-ictal περίοδο στην ictal. Σύμφωνα με μελέτες κατά την περίοδο αυτή εμφανίζονται αποκλίσεις στη συμπεριφορά του ΗΕΓ σε σχέση με τις inter-ictal περιόδους και η χρονική της διάρκεια κυμαίνεται από μερικά λεπτά μέχρι ώρες.

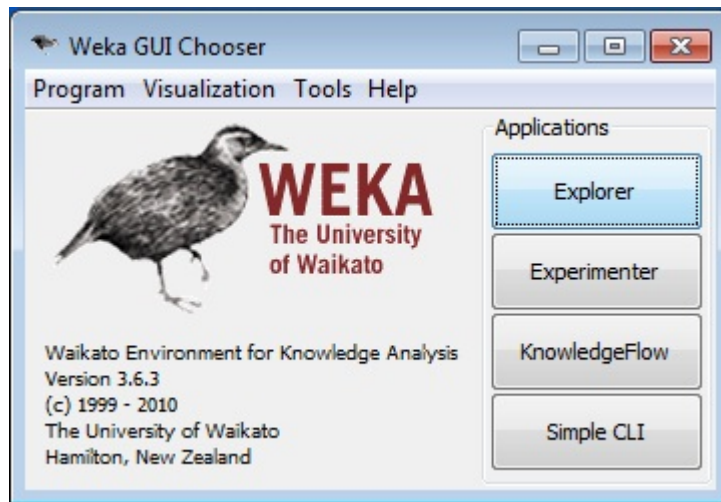


Εικόνα 5.3 : Χαρακτηριστικό ηλεκτροεγκεφαλογράφημα επιληπτικών κρίσεων.

Ένα τυπικό ΗΕΓ ρουτίνας διαρκεί 20-30 λεπτά, και περιλαμβάνει διαδικασίες ενεργοποίησης (υπέρπνοια, φωτεινό ερεθισμό) ή συγκεκριμένα ερεθίσματα στις σπάνιες περιπτώσεις αντακλαστικών συνδρόμων. Η ευαισθησία του ΗΕΓ ρουτίνας στην ανίχνευση της επιληψίας, καθώς και στην διάκρισή της από άλλα παροξυσμικά σύνδρομα, είναι περιορισμένη στο 25-55% (βάσει δημοσιευμένων στατιστικών). Αυτό διότι περισσότεροι από τους μισούς ασθενείς με επιληψία εμφανίζουν φυσιολογικό ΗΕΓ κατά τη ρουτίνα και σε ένα 10% των ασθενών με επιληψία δεν καταγράφονται ποτέ επιληπτόμορφα ΗΕΓ στοιχεία.

6 Λογισμικά που χρησιμοποιήθηκαν

6.1 WEKA (Waikato Environment for Knowledge Analysis)



Εικόνα 6.1: weka.

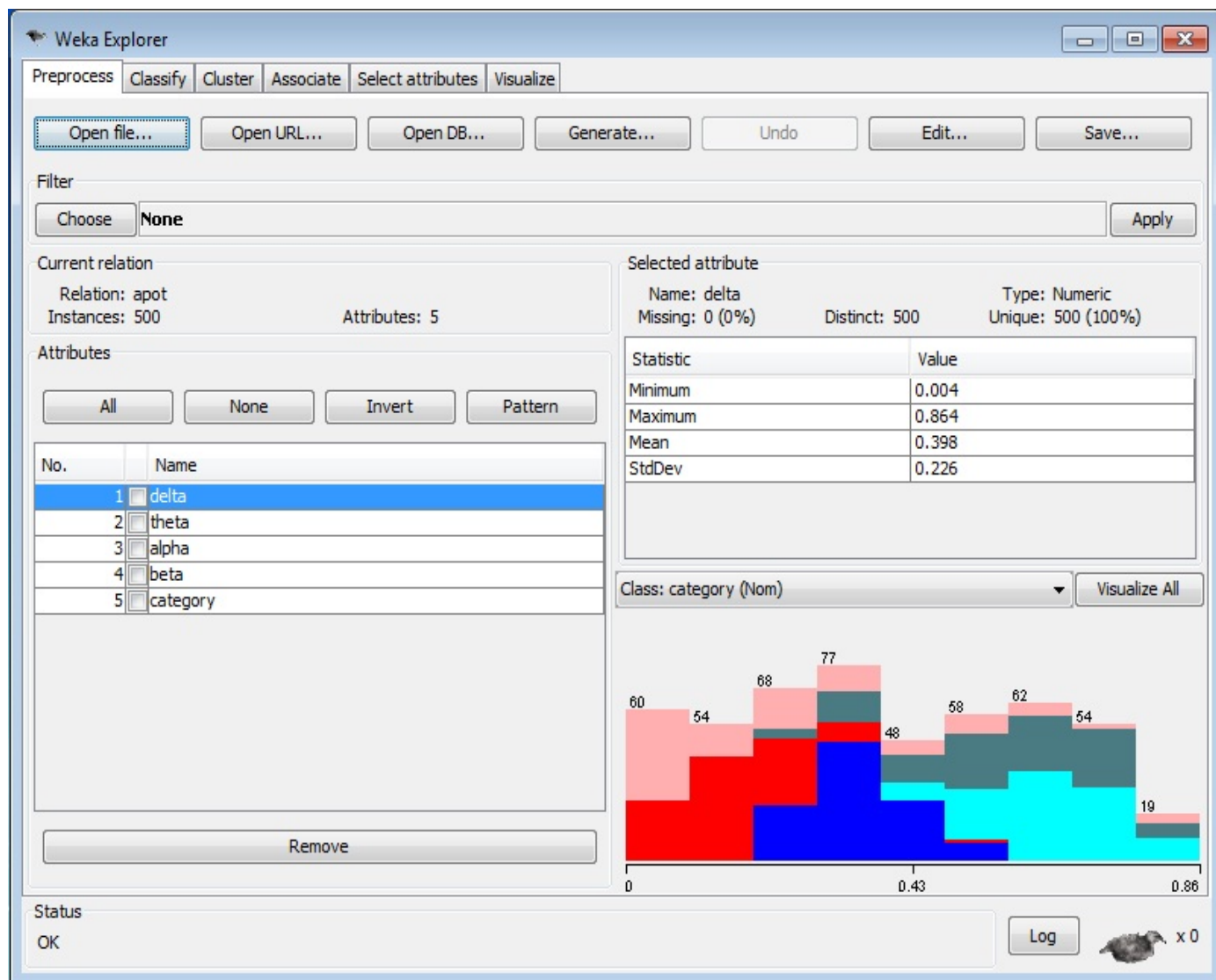
Το WEKA αποτελεί μία βιβλιοθήκη συναρτήσεων και κλάσεων σε γλώσσα προγραμματισμού JAVA. Είναι μια συλλογή από αλγόριθμους μηχανικής μάθησης για τις στοιχειώδεις εργασίες εξόρυξης δεδομένων. Οι αλγόριθμοι μπορούν να εφαρμοστούν άμεσα σε ένα σύνολο δεδομένων ή να κληθούν μέσα από τον κώδικα της JAVA. Παρέχει δυνατότητα για:

- Προεπεξεργασία των δεδομένων (τα εργαλεία για την προεπεξεργασία στο περιβάλλον του WEKA ονομάζονται filters).
- Δημιουργία «μοντέλων» από τα δεδομένα με κάποια διαδικασία εκπαίδευσης
- Χρησιμοποίηση στατιστικών μεγεθών για την αξιολόγηση των διαφόρων αλγορίθμων μάθησης.
- Απεικόνιση τόσο των αρχικών δεδομένων όσο και των αποτελεσμάτων μετά τη διαδικασία της εκπαίδευσης.

Παρέχει επίσης εργαλεία για ταξινόμηση και ομαδοποίηση, εξόρυξη κανόνων ένωσης και απεικόνιση. Είναι καλά προσαρμοσμένο για την ανάπτυξη νέων συστημάτων μηχανικής μάθησης. Μέσω της συγκεκριμένης βιβλιοθήκης δίνεται και η δυνατότητα εκπαίδευσης και ελέγχου αλγορίθμων που χρησιμοποιούνται για ένα σύνολο δεδομένων και ταξινόμησης αποτελεσμάτων. [9]

Η ταξινόμηση, ή αλλιώς κατηγοριοποίηση, αποτελεί μία από τις πιο βασικές και δημοφιλείς τεχνικές στην εξόρυξη δεδομένων. Η ταξινόμηση είναι η διαδικασία κατά την οποία ένα στοιχείο ανατίθεται σε κάποια από τις κλάσεις ενός προκαθορισμένου συνόλου κατηγοριών (κλάσεων). Οι τεχνικές κατηγοριοποίησης χρησιμοποιούν κατά κανόνα ένα

σύνολο εκπαίδευσης όπου όλα τα στοιχεία έχουν ήδη αντιστοιχηθεί στις κλάσεις. Ο αλγόριθμος ταξινόμησης θα μπορεί με αυτό το σύνολο δεδομένων εκπαίδευσης να αναγνωρίσει τα ιδιαίτερα χαρακτηριστικά κάθε κλάσης και έπειτα να ταξινομήσει τα νέα στοιχεία στις κλάσεις σύμφωνα με τις τιμές τους στα χαρακτηριστικά αυτά, συνεπώς η τεχνική της κατηγοριοποίησης χρησιμοποιεί μεθόδους μάθησης με επίβλεψη. [10]



Εικόνα 6.2: Περιβάλλον weka

Οι πιο γνωστές μέθοδοι κατηγοριοποίησης είναι τα Νευρωνικά Δίκτυα (MLP), τα Δέντρα Απόφασης (Decision Trees), η μάθηση κατά Bayes, η κατηγοριοποίηση κοντινότερων γειτόνων (Nearest Neighbor).

Στην παρούσα εργασία χρησιμοποιήθηκαν πέντε αλγόριθμοι ταξινόμησης, οι οποίοι είναι:

- 1) ο αφελής ταξινομητής Bayes (Naïve Bayes Classifier)
- 2) ο αλγόριθμος k κοντινότερων γειτόνων (k Nearest Neighbor)

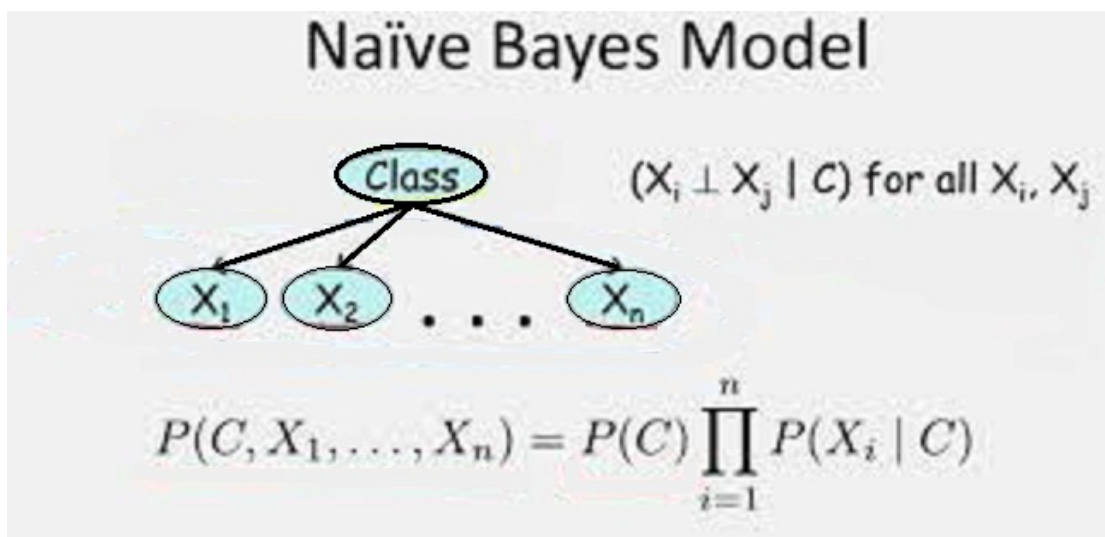
- 3) ο αλγόριθμος J48 για δημιουργία Δέντρου απόφασης (Decision Tree),
- 4) ο αλγόριθμος δημιουργίας Τυχαίου Δάσους (Random Forest),
- 5) το Νευρωνικό Δίκτυο πολυεπίπεδο Perceptron (Multiplayer Perceptron).

6.1.1 Naïve Bayes.

Το θεώρημα Bayes είναι μία πιθανοθεωρητική προσέγγιση της εξόρυξης δεδομένων που έχει ως στόχο την εύρεση της πιο πιθανής υπόθεσης από ένα σύνολο δεδομένων. [10]

Ο Naïve Bayes είναι ένας απλός πιθανοτικός ταξινομητής που βασίζεται στην εφαρμογή του θεωρήματος Bayes. Στόχος του είναι βρεθεί η μέγιστη πιθανότητα για κάθε κλάση και να ταξινομήσει ένα δείγμα x σε μια από τις προκαθορισμένες κατηγορίες. Η κύρια υπόθεση του Naïve Bayes είναι ότι όλα τα χαρακτηριστικά είναι ανεξάρτητες μεταβλητές. Έτσι με βάση το υπάρχον σύνολο δεδομένων ο ταξινομητής υπολογίζει την πιο πιθανή κατηγορία σύμφωνα με τον νόμο του Bayes. Θα εξετάσει χωριστά κάθε ένα από τα χαρακτηριστικά κάθε κατηγορίας και όχι την ταυτόχρονη συνύπαρξη των χαρακτηριστικών. Λειτουργεί υπό την παραδοχή ότι η παρουσία ή μη ενός χαρακτηριστικού ενεργεί ανεξάρτητα από την παρουσία ή μη ενός άλλου χαρακτηριστικού. [11]

Ένα δίκτυο Bayes είναι μια γραφική απεικόνιση που δείχνει τις πιθανοτικές συσχετίσεις μεταξύ των μεταβλητών (Εικόνα 6.3). Οι κόμβοι του δικτύου απεικονίζουν τυχαίες μεταβλητές. Χρησιμοποιεί άκυκλους γράφους για να μοντελοποιήσει τις εξαρτήσεις ανάμεσα στα χαρακτηριστικά των παραδειγμάτων και τις κατηγορίες. [12]



Εικόνα 6.3 : Παράδειγμα Αφελή ταξινομητή (NB)

6.1.2 *k* κοντινότερων γειτόνων (*k Nearest Neighbor*).

Ο αλγόριθμος ταξινόμησης *k* κοντινότερων γειτόνων είναι η πιο βασική μέθοδος μάθησης βασισμένη σε στιγμιότυπα. Η βασική ιδέα είναι ότι για να ταξινομηθεί ένα άγνωστο δεδομένο x βρίσκονται τα k κοντινότερα στοιχεία του συνόλου εκπαίδευσης ως προς το στοιχείο που θέλει να ταξινομηθεί. Το x ταξινομείται στην κατηγορία που πλειοψηφεί ανάμεσα στους k κοντινότερους γείτονες. Η τιμή του k καθορίζεται ανάλογα με το πρόβλημα. Κατά την ταξινόμηση άγνωστων δεδομένων η τιμή του k είναι σταθερή. Όταν υπάρχουν δύο κλάσεις το k επιλέγεται να είναι περιττός αριθμός. [13]

Υπάρχουν πολλοί τύποι για τον ορισμό της απόστασης μεταξύ των άγνωστων δεδομένων και των δεδομένων του συνόλου εκπαίδευσης. Μία από αυτές που είναι η πιο βασική είναι η συνάρτηση της ευκλείδειας απόστασης $d(x, x^i) = \sqrt{\sum_{j=1}^d (x_j - x_j^i)^2}$.

Κάποια πλεονεκτήματα του αλγόριθμου ταξινόμησης *k* κοντινότερων γειτόνων είναι τα εξής:

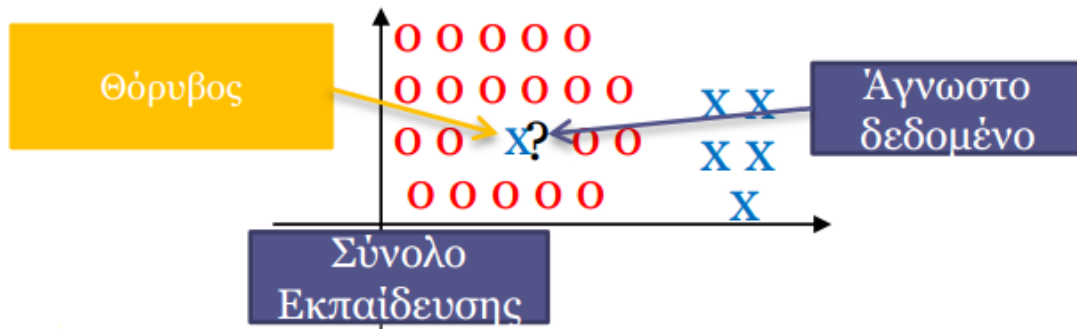
- δεν απαιτείται εκπαίδευση,
- επιλύει προβλήματα πολλών κατηγοριών,
- προσεγγίζει πολύ καλύτερα μια πολύπλοκη συνάρτηση-στόχο από άλλους μεθόδους,
- προγραμματίζεται εύκολα, και
- με μικρές αλλαγές στα δεδομένα εκπαίδευσης δεν παρατηρούνται μεγάλες αλλαγές στα αποτελέσματα ταξινόμησης.

Τα μειονεκτήματα του αλγορίθμου είναι τα παρακάτω:

- Υψηλό υπολογιστικό κόστος,
- δυσκολία στην εύρεση του καταλληλότερου k , και
- η αργή ταξινόμηση. [13]

Η τιμή του k , δηλαδή το πλήθος των κοντινότερων γειτόνων που ελέγχουμε για να ταξινομήσουμε άγνωστα δεδομένα καθορίζεται σύμφωνα με το πρόβλημα. Ανάλογα λοιπόν με την τιμή του k μπορεί να εμφανιστεί το φαινόμενο υπερεκπαίδευσης ή το φαινόμενο υποεκπαίδευσης.

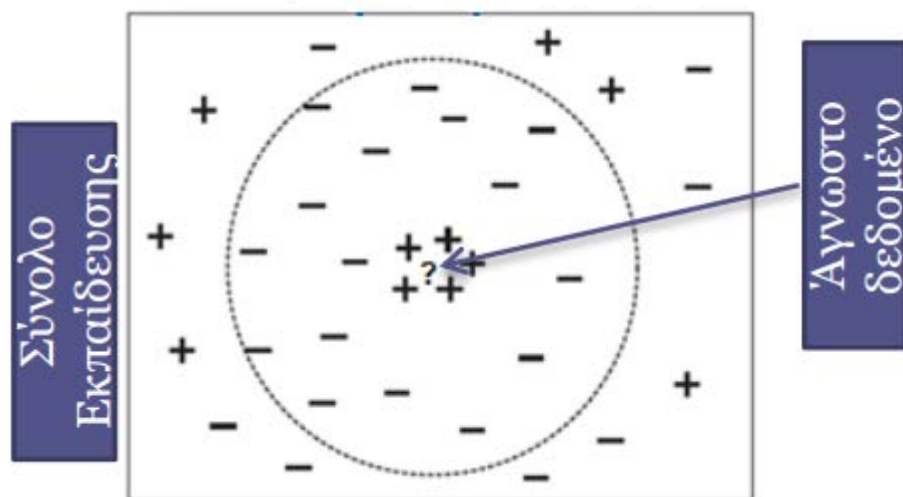
- Το φαινόμενο της υπερεκπαίδευσης εμφανίζεται όταν το k έχει μικρή τιμή.



Εικόνα 6.4: Παράδειγμα ταξινόμησης.

Θέλουμε να ταξινομήσουμε το στοιχείο «?» . Θεωρούμε $k=1$, δηλαδή 1-NN. Το στοιχείο «?» ταξινομείται ως X . Ωστόσο το X που βρίσκεται πολύ κοντά στο στοιχείο που θέλουμε να ταξινομήσουμε, είναι θόρυβος καθώς είναι απομονωμένο από τα υπόλοιπα X . Το πιθανότερο είναι ότι το στοιχείο «?» είναι κατηγορίας O καθώς βρίσκονται πολλά κοντά του και άρα μάλλον προέρχεται από αυτόν τον πληθυσμό.

- Το φαινόμενο της υποεκπαίδευσης εμφανίζεται όταν το k έχει πολύ μεγάλη τιμή. Αφού το k έχει μεγάλη τιμή, στην γειτονιά συμπεριλαμβάνονται και σημεία πολύ μακριά από το δεδομένο που θέλουμε να ταξινομήσουμε, οπότε και χάνεται η έννοια της γειτονιάς.



Εικόνα 6.5 : Παράδειγμα ταξινόμησης.

Θέλουμε να ταξινομήσουμε το στοιχείο «?» . Θεωρούμε $k=20$. Η γειτονιά του «?» βρίσκεται μέσα στον κύκλο. Οπότε το στοιχείο «?» ταξινομείται στην κατηγορία $-$. Το πιθανότερο είναι ότι ανήκει στην κατηγορία $+$ καθώς πιο κοντά στο στοιχείο που ταξινομείται υπάρχει πληθυσμός από στοιχεία κατηγορίας $+$. Άρα η μικρή γειτονιά από $+$

χάνεται λόγω του μεγάλου k ($k=20$). [13]

6.1.3 Δέντρα απόφασης (*Decision Trees*).

Τα δέντρα απόφασης είναι αλγόριθμοι που χρησιμοποιούν την τεχνική του «Διαίρει και Βασίλευε». Είναι μια δομή στην οποία κάθε κόμβος φύλλων αντιπροσωπεύει μία ταξινόμηση δηλαδή μία κλάση. Τα δέντρα απόφασης αποτελούνται από κανόνες AN-TOTE (IF-THEN) από την ρίζα του δέντρου προς τα φύλλα. Οι κόμβοι του δέντρου παίρνουν τα ονόματα των χαρακτηριστικών, τα φύλλα τις διάφορες πιθανές κλάσεις και οι ακμές δηλαδή τα κλαδιά παίρνουν τις δυνατές τιμές που μπορεί να πάρει ένα χαρακτηριστικό. Τα στοιχεία ταξινομούνται ακολουθώντας το μονοπάτι στο δέντρο, από την ρίζα προς τα φύλλα, και ελέγχονται κατά μήκος της διαδρομής σε διαδοχικούς κόμβους, λαμβάνοντας τις τιμές που αντιστοιχούν στις τιμές των χαρακτηριστικών ενός στοιχείου μέχρις ότου φτάσουμε σε ένα φύλλο, οπότε και ταξινομείται το αντίστοιχο στοιχείο σύμφωνα με την κλάση του φύλλου. [10]

Στην βιβλιογραφία έχει προταθεί ένα σύνολο από αλγορίθμους για την επαγωγή δέντρων απόφασης από δεδομένα. Ένας από αυτούς είναι ο J48 (ή C4.5) ο οποίος αποτελεί μία βελτίωση του αλγορίθμου ID3. Η επαγωγή δέντρου απόφασης από τα δεδομένα με βάση τον αλγόριθμο C4.5 περιλαμβάνει τα παρακάτω βήματα:

1. Ξεκίνα με έναν κόμβο που περιέχει όλες τις εγγραφές.
2. Ακολουθεί η διάσπαση του κόμβου (μοίρασμα των εγγραφών) με βάση μια συνθήκη-διαχωρισμού σε κάποιο από τα γνωρίσματα.
3. Αναδρομική κλήση του βήματος 2 σε κάθε κόμβο έως ότου οι εγγραφές ενός τελικού κόμβου να ανήκουν σε μια μόνο κλάση.
4. Αφού κατασκευαστεί το δέντρο, γίνονται κάποιες βελτιστοποιήσεις (tree pruning).

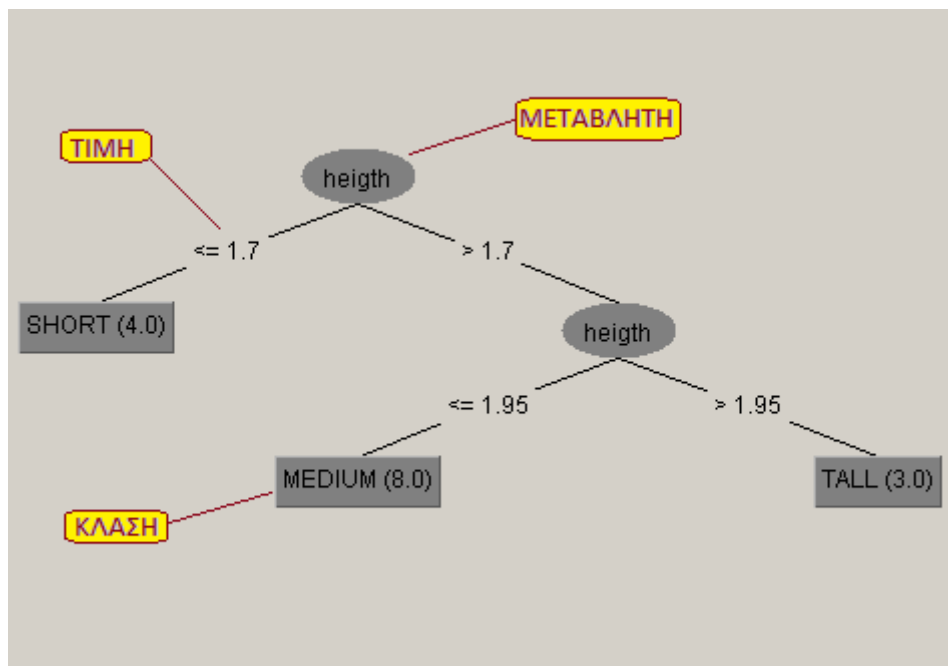
Κάποια πλεονεκτήματα των Δέντρων απόφασης είναι τα παρακάτω:

- Έχουν εύκολη υλοποίηση και γρήγορη εφαρμογή,
- μπορούν να χειριστούν μεγάλο αριθμό στοιχείων,
- η ταξινόμηση νέων εγγραφών γίνεται πολύ γρήγορα,
- δυνατότητα εξαγωγής ενός συνόλου κανόνων συμπερασμάτων από το δέντρο απόφασης, το οποίο είναι εύκολο στην κατανόηση και στην ερμηνεία, και
- η γραφική απεικόνισή τους. [10]

Τα μειονεκτήματα του ταξινομητή αυτού είναι :

- Εμφανίζουν το πρόβλημα του overfitting. (Το πρόβλημα του overfitting είναι το υπερβολικό ταίριασμα με τα δεδομένα εκπαίδευσης),
 - προβλήματα όταν λείπουν πολλά δεδομένα,
 - πραγματοποιούν διαίρεση του χώρου των δεδομένων σε υποσύνολα μόνο με «ευθείες γραμμές» για το εύρος τιμών ενός γνωρίσματος (απλά όρια απόφασης), και
 - δεν μπορούν να χειριστούν περίπλοκες σχέσεις ανάμεσα στα γνωρίσματα.
- [10]

Ένα παράδειγμα εφαρμογής δέντρου απόφασης σε εκπαιδευτικά δεδομένα είναι το παρακάτω (Εικόνα 6.6), όπου με βάση τα δεδομένα εκπαίδευσης που έχουμε, κατασκευάζουμε το δέντρο απόφασης.



Εικόνα 6.6: Δέντρο απόφασης

Με βάση το δέντρο απόφασης παράγονται οι παρακάτω κανόνες IF THEN :

- If height<=1.7 then class=SHORT
- If height>1.7 and height<=1.95 then class=MEDIUM
- If height>1.7 and height>1.95 then class=TALL

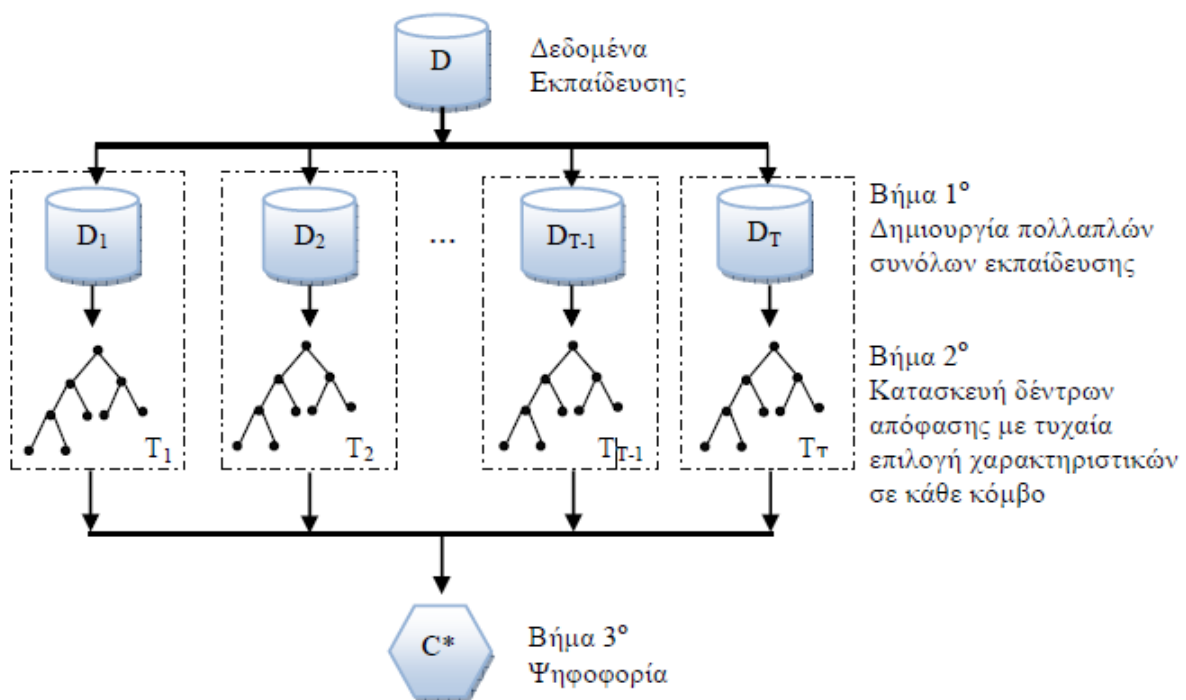
6.1.4 Τυχαίο Δάσος (Random Forest).

Το τυχαίο δάσος είναι ένας ταξινομητής ο οποίος αναπτύσσει πολλά δέντρα απόφασης

ταυτόχρονα στα οποία κάθε κόμβος χρησιμοποιεί ένα τυχαίο υποσύνολο των χαρακτηριστικών που χρησιμοποιούνται στο πρόβλημα ταξινόμησης. Κάθε δέντρο απόφασης δίνει μία ταξινόμηση και έχει την δυνατότητα να ψηφίζει την πιο δημοφιλή κλάση. Έτσι κάθε κλάση έχει κάποιες ψήφους (votes) και σύμφωνα με αυτές διαλέγεται η κλάση με τις περισσότερες ψήφους. Ο ταξινομητής αυτός χρησιμοποιεί τον μεγάλο αριθμό δέντρων απόφασης και αποφασίζει την κατηγορία με την επιλογή του τρόπου των κατηγοριών όπως καθορίζεται από τα μεμονωμένα δέντρα. [12]

Κάθε δέντρο κατασκευάζεται με βάση τον ακόλουθο αλγόριθμο:

- Αρχικά, ανατίθεται στην ρίζα του δέντρου το σύνολο δειγμάτων εκπαίδευσης, έστω N .
- Έστω M ο αριθμός των μεταβλητών-χαρακτηριστικών στην είσοδο του ταξινομητή.
- Στην συνέχεια, δημιουργούνται τυχαία πολλαπλά σύνολα εκπαίδευσης n , τα οποία είναι υποσύνολα από το σύνολο δειγμάτων N . Τα πολλαπλά αυτά σύνολα n αποτελούν κάθε ένα το σύνολο εκπαίδευσης για κάθε δέντρο.
- Για κάθε κόμβο στο δέντρο επιλέγονται τυχαία m μεταβλητές (από το σύνολο M μεταβλητών) στις οποίες στηρίζεται ο έλεγχος για την απόφαση του κόμβου. Η τιμή του m παραμένει σταθερή κατά την διάρκεια κατασκευής όλου του δάσους.
- Κάθε δέντρο αναπτύσσεται πλήρως χωρίς κλάδεμα. [14]



Εικόνα 6.7: Παράδειγμα κατασκευής δέντρων απόφασης (τυχαίου δάσους).

Τα πλεονεκτήματα του ταξινομητή Τυχαίο Δάσος (Random Forest) είναι τα ακόλουθα:

- Υπάρχει η δυνατότητα παράλληλης επαγωγής των δέντρων,
- παρουσιάζουν ανεκτικότητα ως προς το θόρυβο και ως προς τα αριθμητικά σφάλματα στα δεδομένα εκπαίδευσης,
- παράγει τον ακριβέστερο ταξινομητή μεταξύ των τρεχόντων αλγορίθμων (από το 2004),
- περιλαμβάνει μια καλή μέθοδο για το ελλείπον στοιχείο και διατηρεί την ακρίβεια όταν λείπει ένα μεγάλο ποσοστό των στοιχείων,
- είναι μια μη επαναληπτική διαδικασία εκπαίδευσης, οπότε ο αλγόριθμος ολοκληρώνεται σε σταθερό αριθμό βημάτων,
- παράγει μια εσωτερική αμερόληπτη εκτίμηση του σφάλματος γενίκευσης καθώς εξελίσσεται η διαδικασία κατασκευής του δέντρου,
- εξαιτίας του μεγάλου πλήθους δέντρων στο δάσος, το σφάλμα γενίκευσης είναι περιορισμένο. Αυτό έχει ως αποτέλεσμα τη μη εμφάνιση φαινομένων υπερεκπαίδευσης,
- μπορούν να εκπαιδευτούν σε σύνολα δεδομένων υψηλής διάστασης όπως είναι τα κείμενα και οι εικόνες, χωρίς να εμφανίσουν σημαντικό βαθμό υπερεκπαίδευσης.

Μερικά μειονεκτήματα του ταξινομητή Τυχαίο Δάσος είναι:

- Υψηλό υπολογιστικό κόστος,
- για την επέκταση ενός μοντέλου με στόχο την εισαγωγή μιας ακόμα κατηγορίας απαιτείται η κατασκευή του μοντέλου από την αρχή,
- αναζητά το καλύτερο διαχωρισμό σε ένα μικρό υποσύνολο των χαρακτηριστικών και δεν κάνει εξαντλητική αναζήτηση,
- κάθε νέο παράδειγμα πρέπει να διασχίσει όλα τα δέντρα του δάσους για την εκτίμηση της κατηγορίας του. [12]

6.1.5 Πολυεπίπεδο Perceptron (Multiplayer Perceptron)

Ένας άλλος διαδομένος ταξινομητής είναι το πολυεπίπεδο perceptron, το οποίο είναι ένα νευρωνικό δίκτυο που αποτελείται από πολλούς τεχνητούς νευρώνες οργανωμένους σε επίπεδα και εκπαιδεύεται με επίβλεψη, με την χρήση του αλγορίθμου back propagation. [15] Οι αρχές και οι λειτουργίες των τεχνητών νευρωνικών δικτύων

βασίζονται στο νευρικό σύστημα των ζώντων οργανισμών και έχουν ως στόχο να προσομοιώσουν τους νευρώνες του ανθρώπινου εγκεφάλου. Η λειτουργία τους προσπαθεί να συνδυάσει τον τρόπο σκέψης του ανθρώπινου εγκεφάλου με τον αφηρημένο μαθηματικό τρόπο σκέψης. Τα νευρωνικά δίκτυα είναι συστήματα επεξεργασίας σύνθετων προβλημάτων πρόβλεψης με μεγάλο αριθμό εξαρτημένων μεταβλητών. Η δομή τους χαρακτηρίζεται από ένα γράφημα που αποτελείται από κόμβους και συνάψεις και σχηματίζουν ένα δίκτυο. Οι νευρώνες είναι το κύριο δομικό στοιχείο του δικτύου. Υπάρχουν τρία είδη νευρώνων, οι νευρώνες εισόδου, οι υπολογιστικοί ή ενδιάμεσοι νευρώνες και οι νευρώνες εξόδου. Οι νευρώνες εισόδου δέχονται τις πληροφορίες που θα υποστούν επεξεργασία και δεν υπολογίζουν τίποτα. Δηλαδή η έξοδος ενός νευρώνα εισόδου είναι η είσοδός του. Οι υπολογιστικοί νευρώνες, που είναι και το κρυφό επίπεδο και βρίσκονται μεταξύ των νευρώνων εισόδου-εξόδου, πολλαπλασιάζουν τις εισόδους τους με τα συναπτικά βάρη και υπολογίζουν το άθροισμα των γινομένων, το οποίο είναι και το όρισμα της συνάρτησης ενεργοποίησης. Τα αποτελέσματα αυτής της επεξεργασίας καταλήγουν στους νευρώνες εξόδου. Τα συναπτικά βάρη, δηλαδή τα βάρη των δεσμών-ακμών πρέπει πρώτα να προσδιοριστούν. Οι νευρώνες εισόδου είναι τόσοι όσες και οι ανεξάρτητες μεταβλητές και κάθε νευρώνας εισόδου συνδέεται με όλους τους νευρώνες στο πρώτο κρυμμένο επίπεδο με τους δεσμούς, οι οποίοι ονομάζονται συνάψεις και σε κάθε μία απ' αυτές αντιστοιχεί ένα βάρος που ονομάζεται συναπτικό βάρος. Οι νευρώνες του πρώτου κρυμμένου επιπέδου συνδέονται με την σειρά τους με νευρώνες ενός άλλου επιπέδου αν υπάρχει ή αλλιώς με τους νευρώνες στο επίπεδο εξόδου. Το MLP μπορεί να έχει οποιοδήποτε αριθμό από κρυμμένα επίπεδα και κάθε κρυμμένο επίπεδο οποιοδήποτε αριθμό από νευρώνες. Το πλήθος των κρυμμένων επιπέδων και νευρώνων σε αυτά καθορίζουν την αρχιτεκτονική του MLP. [15]

Οι νευρώνες εξόδου αναπαριστούν μία ή περισσότερες εξαρτημένες μεταβλητές. Το επίπεδο εξόδου έχει τόσους νευρώνες όσες και οι κατηγορίες των δεδομένων. Οι νευρώνες ονομάζονται κόμβοι. Κάθε κόμβος είναι στοιχείο επεξεργασίας που λειτουργεί ανεξάρτητα από τους άλλους κόμβους και χρησιμοποιεί μόνο τοπικά δεδομένα που καθοδηγούν την επεξεργασία.

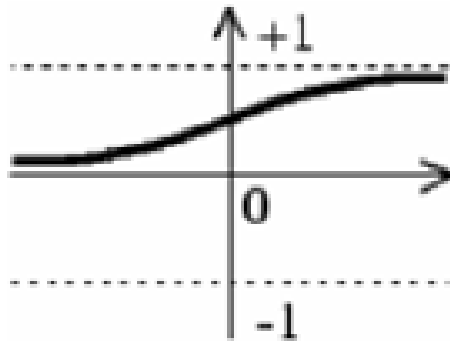
Νευρώνες του ίδιου επιπέδου αλλά και μη διαδοχικών επιπέδων δεν συνδέονται. Κάθε νευρώνας συνδέεται με όλους τους νευρώνες του αμέσως επόμενου επιπέδου. Όλες οι συνδέσεις έχουν φορά από το επίπεδο εισόδου προς το επίπεδο εξόδου.

Η συνάρτηση ενεργοποίησης των κρυμμένων επιπέδων μπορεί να είναι οποιαδήποτε αρκεί να παραγωγίζεται. Προτιμώνται μη γραμμικές συναρτήσεις όπως η σιγμοειδής

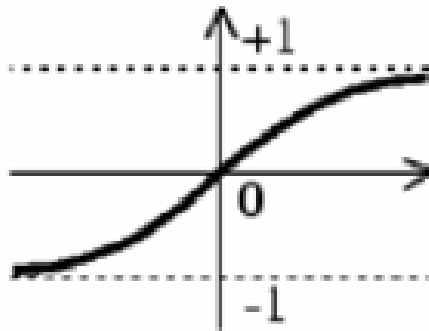
$$\varphi(u) = \sigma(u) = \frac{1}{1+e^{-u}} \quad \text{και η υπερβολική εφαπτόμενη} \quad \varphi(u) = \tanh(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}},$$

αλλά συνήθως επιλέγεται η σιγμοειδής συνάρτηση. Οι γραφικές παραστάσεις των

συναρτήσεων δίνονται στις εικόνες αντίστοιχα. [15]



Εικόνα 6.8: Γραφική παράσταση σιγμοειδούς συνάρτησης



Εικόνα 6.9: Γραφική παράσταση υπερβολικής εφαπτομένης

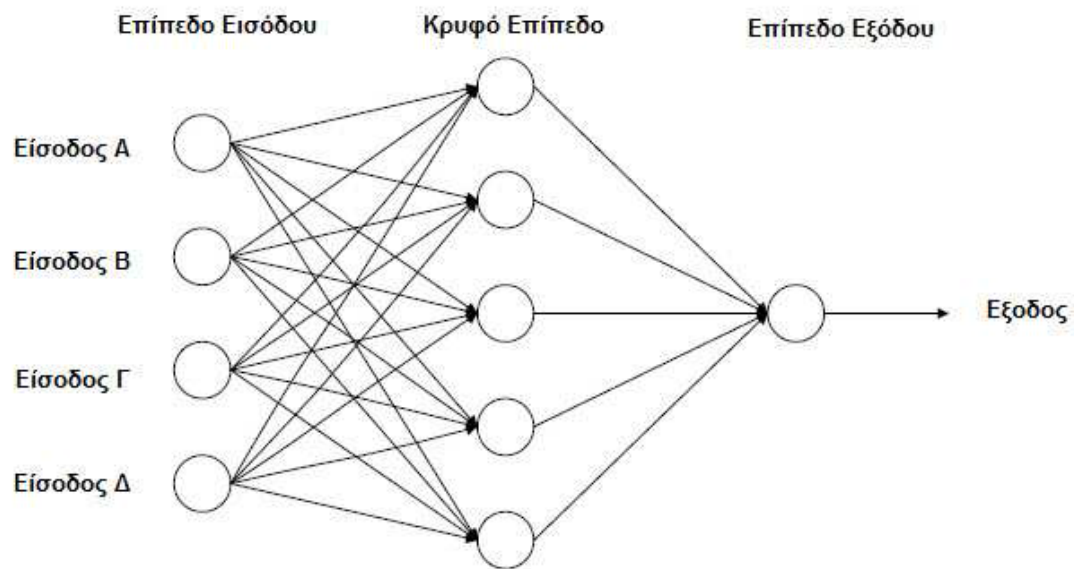
Τα πλεονεκτήματα των Νευρωνικών Δικτύων είναι μερικά από τα παρακάτω:

- Είναι εύκολα παραμετροποιήσιμα,
- η επίλυση προβλημάτων πολλών κατηγοριών,
- η επίλυση μη-γραμμικών διαχωρίσιμων προβλημάτων,
- αποδίδουν καλά σε πολύπλοκα προβλήματα εκεί όπου άλλες τεχνικές αποτυγχάνουν, και
- έχουν υψηλή ακρίβεια πρόγνωσης. [12]

Τα μειονεκτήματα τους είναι ότι:

- Είναι δύσκολα στην κατανόηση,
- η δημιουργία κανόνων από νευρωνικά δίκτυα δεν είναι απλή διαδικασία,
- χρειάζονται μεγάλο χρόνο εκπαίδευσης,
- οι τιμές του επιπέδου εισόδου πρέπει να είναι αριθμητικές,
- μπορεί να είναι πολύ ακριβά στην χρήση τους, και

- η εξήγηση και η ερμηνεία των βαρών είναι σχεδόν αδύνατη λόγω της μη γραμμικότητας. [12]

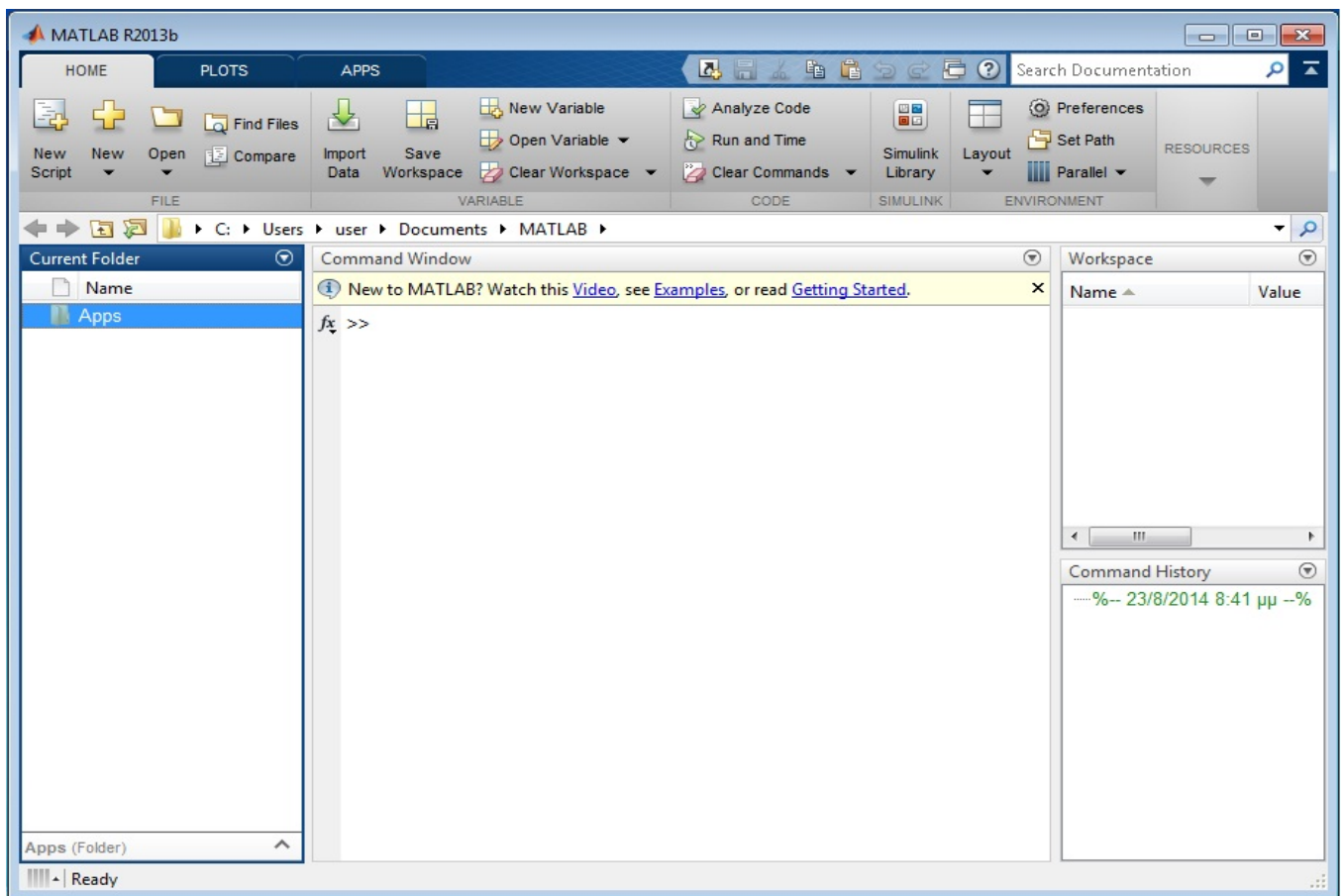


Εικόνα 6.10: Παράδειγμα νευρωνικού δικτύου.

6.2 MATLAB (*matrix laboratory*)

Το MATLAB (*matrix laboratory*) είναι ένα περιβάλλον αριθμητικής υπολογιστικής και μια γλώσσα προγραμματισμού υψηλού επιπέδου (*script language*), τέταρτης γενιάς, που επιτρέπει την ολοκλήρωση υπολογιστικά απαιτητικών εργασιών ταχύτερα και ευκολότερα απ' ό,τι οι παραδοσιακές γλώσσες προγραμματισμού, όπως C, C++, και Fortran. Χρησιμοποιείται κατά κύριο λόγο για την επίλυση μαθηματικών προβλημάτων, ωστόσο μπορεί να χρησιμοποιηθεί και για προγραμματισμό καθώς περιέχει εντολές από την C++ όπως την *if*, την *while* και την *switch*. [16]

Παρέχει ένα δυναμικό, εύχρηστο και ανοικτό υπολογιστικό περιβάλλον για την εκτέλεση απλών μαθηματικών υπολογισμών αλλά και πιο σύνθετων λειτουργιών πάνω σε εξειδικευμένες περιοχές εφαρμογών καθώς περιέχει ένα σύνολο συναρτήσεων και εξωτερικών βιβλιοθηκών (*Toolboxes*) για ανάπτυξη αλγορίθμων, επεξεργασία, ανάλυση και απεικόνιση δεδομένων, αριθμητικές μεθόδους επίλυσης προβλημάτων και μοντελοποίηση συστημάτων. Συγκεκριμένα εκτελεί απλές μαθηματικές πράξεις, αλλά εξίσου εύκολα χειρίζεται μιγαδικούς αριθμούς, δυνάμεις, ειδικές μαθηματικές συναρτήσεις, πίνακες, διανύσματα και πολυώνυμα. Μπορεί επίσης να αποθηκεύει και να ανακαλεί δεδομένα, να δημιουργεί και να εκτελεί ακολουθίες εντολών που αυτοματοποιούν διάφορους υπολογισμούς και να σχεδιάζει γραφικά. [17]



Εικόνα 6.11: Το περιβάλλον της Matlab.

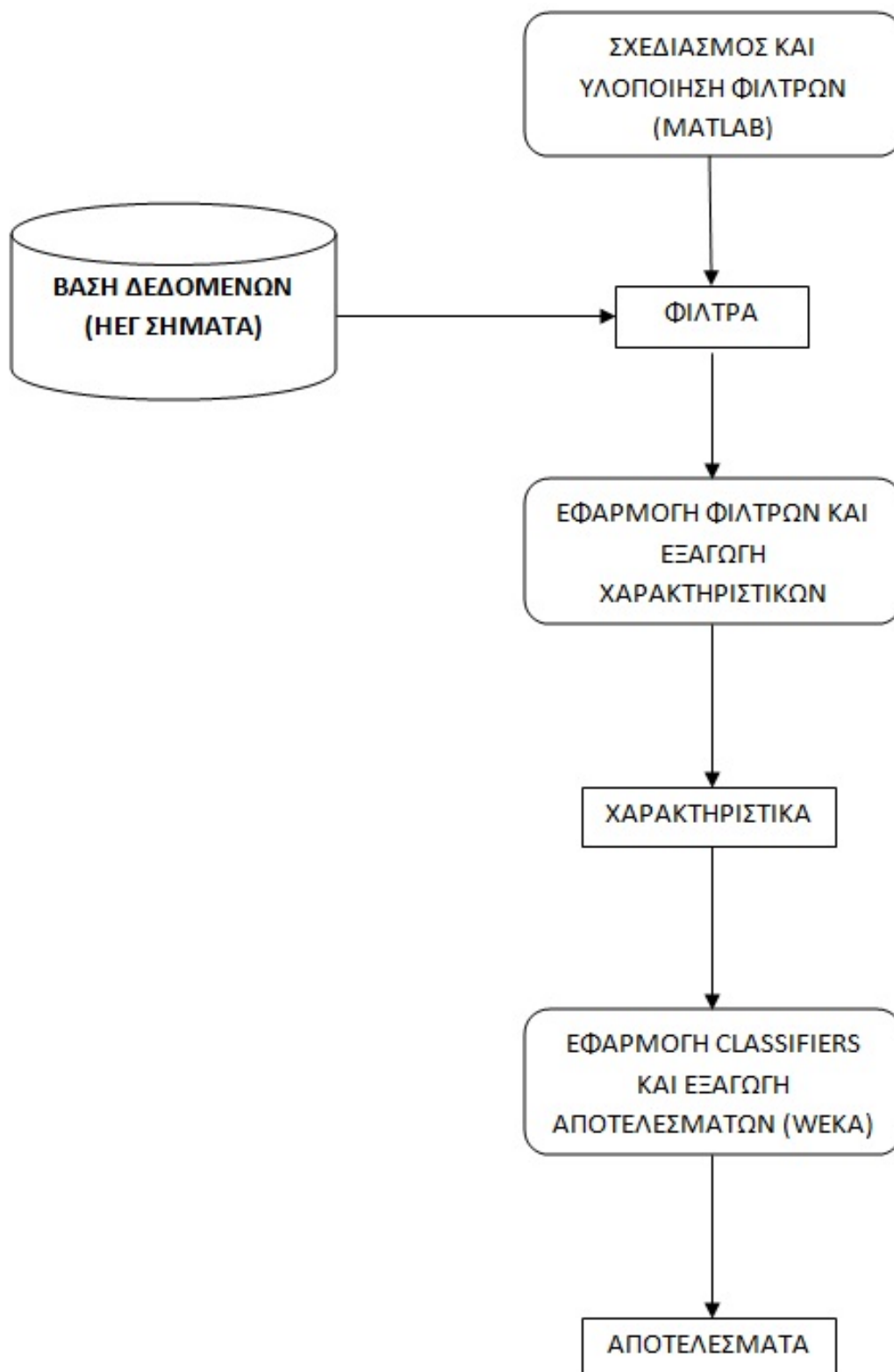
Το διαδραστικό περιβάλλον του προορίζεται για υλοποίηση επιστημονικών εφαρμογών σε ένα ευρύ φάσμα πεδίων, όπως στη Γραμμική Άλγεβρα, Στατιστική, Εφαρμοσμένα Μαθηματικά, Αριθμητική Ανάλυση και Επιστημονικό Υπολογισμό, Επεξεργασία Σημάτων και Εικόνας, Θεωρία Ελέγχου, Θεωρία Βελτιστοποίησης και Γραφικά. Έχει υλοποιηθεί σε πολλές λειτουργικές πλατφόρμες όπως Windows, Macintosh OS και Unix και έχει δύο βασικές εκδόσεις, την επαγγελματική και την εκπαιδευτική (student edition).

Οι λειτουργίες του Matlab διακρίνονται στις *τυποποιημένες*, δηλαδή σε αυτές που χειρίζονται αριθμητικά δεδομένα και εξάγουν αριθμητικά αποτελέσματα, και στις συναρτήσεις του *Symbolic Toolbox*, οι οποίες χειρίζονται και υπολογίζουν συμβολικές εκφράσεις, δηλαδή επεξεργάζονται μαθηματικά σύμβολα. Η γλώσσα του Matlab είναι α-τυπη δηλαδή δεν χρειάζεται δήλωση μεταβλητών και η κατάληξη των αρχείων του είναι *.m* (M files).

7 Πειραματική Διάταξη

7.1 Διάγραμμα ροής

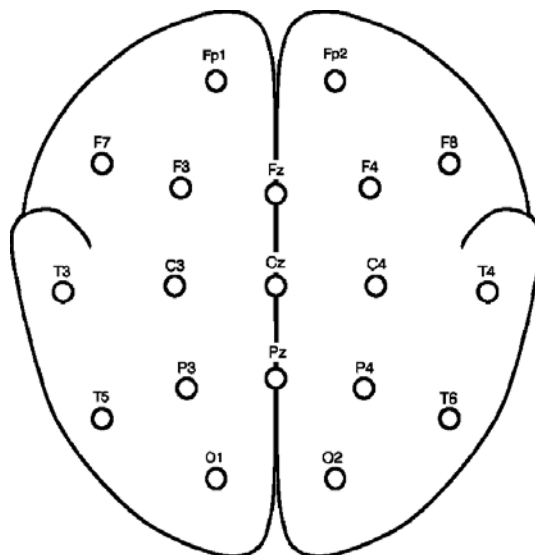
Η μεθοδολογία της πειραματικής διαδικασίας περιλαμβάνει τρία στάδια. Αρχικά χρησιμοποιείται μια βάση δεδομένων που αποτελείται από 500 ΗΕΓ σήματα ασθενών. Στη συνέχεια σχεδιάζονται και υλοποιούνται φίλτρα με την βοήθεια του προγραμματιστικού περιβάλλοντος Matlab, τα οποία βασίζονται στους τέσσερις βασικούς ρυθμούς ΗΕΓ. Τα φίλτρα αυτά εφαρμόζονται στην βάση δεδομένων και γίνεται εξαγωγή χαρακτηριστικών. Τα χαρακτηριστικά καταγράφονται σε αρχεία κατάλληλα για χρήση στο Weka, και εφαρμόζονται στα αρχεία αυτά, 5 ταξινομητές με διαφορετικές παραμέτρους (συνολικά 10 προσεγγίσεις ταξινόμησης), από τα οποία προκύπτουν τα αποτελέσματα. Στο Σχήμα 7.1 παρουσιάζεται το διάγραμμα ροής της μεθοδολογίας που χρησιμοποιήθηκε για την πειραματική διαδικασία.



Σχήμα 7.1: Διάγραμμα ροής πειραματικής διαδικασίας

7.2 Βάση Δεδομένων

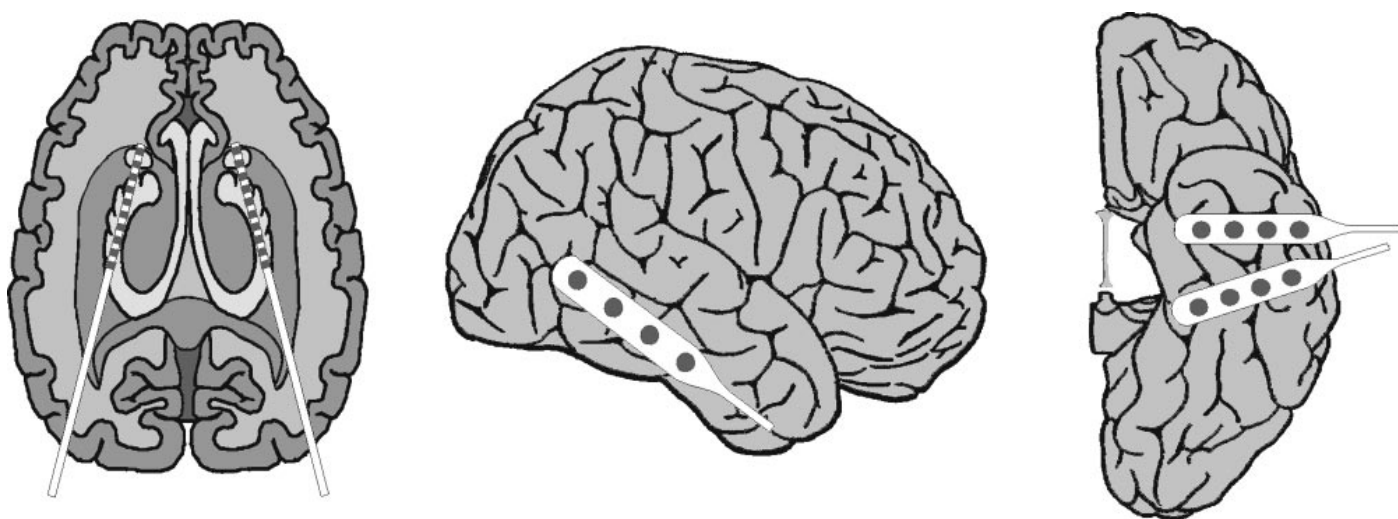
Η βάση δεδομένων που χρησιμοποιήθηκε για τις δοκιμές των μεθόδων και την εξαγωγή των αποτελεσμάτων είναι διαθέσιμη στο διαδίκτυο και έχει δημοσιευθεί από τους Andrzejak και τους συνεργάτες του. Η βάση δεδομένων περιλαμβάνει ΗΕΓ καταγραφές τόσο από υγιή άτομα όσο και από ασθενείς που υποφέρουν από επιληπτικές κρίσεις. Πιο συγκεκριμένα, η βάση δεδομένων περιλαμβάνει ΗΕΓ καταγραφές από 5 υποσύνολα δεδομένων (τα οποία δηλώνονται ως Z, O, N, F, και S) και το κάθε ένα από αυτά περιέχει 100 τμήματα μονοκαναλικών καταγραφών διάρκειας 23.6sec. Αυτά τα τμήματα επιλέχθηκαν και καταγράφηκαν σε συνεχείς πολύωρες πολυκαναλικές ΗΕΓ καταγραφές μετά από οπτική επιθεώρηση για θόρυβο (π.χ. εξαιτίας της μουσικής δραστηριότητας-ΗΜΓ και τις κινήσεις των ματιών-ΗΟΓ) και εφόσον, ικανοποιούσαν ένα κριτήριο στασιμότητας.



Σχήμα 7.2: Διάταξη της τοποθέτησης των επιφανειακών ηλεκτροδίων σύμφωνα με το Διεθνές 10-20 Σύστημα τοποθέτησης ηλεκτροδίων. Τα ονόματα των θέσεων των ηλεκτροδίων προέρχονται από την ανατομική τους τοποθέτηση. Τα ΗΕΓ τμήματα των υποσυνόλων Z και O ελήφθησαν από όλα τα ηλεκτρόδια που απεικονίζονται.

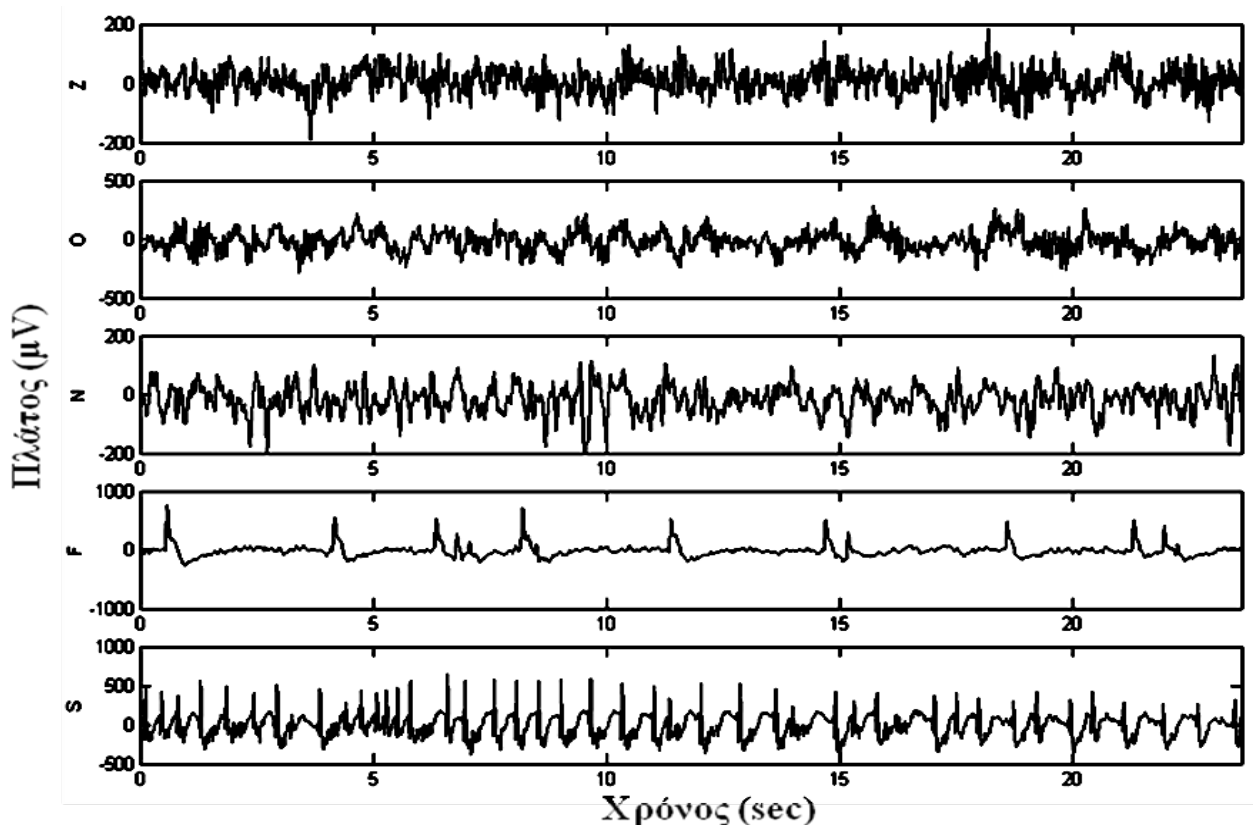
Τα υποσύνολα Z και O αποτελούνται από ΗΕΓ τμήματα που ελήφθησαν από πολύωρες ΗΕΓ καταγραφές 5 υγιών ατόμων, στις οποίες καταγραφές χρησιμοποιήθηκαν επιφανειακά ηλεκτρόδια που τοποθετήθηκαν στο τριχωτό της κεφαλής των ατόμων αυτών, σύμφωνα με το Διεθνές 10-20 Σύστημα τοποθέτησης ηλεκτροδίων (Σχήμα 7.2) και κάτω από τις εξής συνθήκες: 1) με μάτια ανοικτά (υποσύνολο Z) και 2) με μάτια κλειστά (υποσύνολο O). Αντίθετα, τα ΗΕΓ τμήματα των δύο υποσυνόλων N και F περιείχαν μόνο δραστηριότητα που μετρήθηκε σε διαστήματα 'ελεύθερα από επιληπτικές κρίσεις'

(seizure-free intervals) των πολύωρων ΗΕΓ καταγραφών 5 επιληπτικών ασθενών (δηλαδή διαστήματα τα οποία δεν περιείχαν καταγεγραμμένη επιληπτική κρίση). Τα διαστήματα αυτά καταγράφηκαν στην επιληπτογενή ζώνη (epileptogenic zone), στην περίπτωση του υποσύνολο N και στο σχηματισμό του ιππόκαμπου του αντίθετου ημισφαιρίου του εγκεφάλου, στην περίπτωση του υποσύνολου F (Σχήμα 7.3). Τέλος, τα ΗΕΓ τμήματα του υποσυνόλου S μετρήθηκαν σε διαστήματα των πολύωρων ΗΕΓ καταγραφών των ίδιων 5 επιληπτικών ασθενών, τα οποία περιείχαν καταγεγραμμένη δραστηριότητα επιληπτικής κρίσης (ictal δραστηριότητα).



Σχήμα 7.3: Διάταξη των ενδοκρανιακών που εμφυτεύονται για προεγχειρητική εκτίμηση της κατάστασης των επιληπτικών ασθενών. Εν τω βάθει ηλεκτρόδια εμφυτεύονται συμμετρικά στο σχηματισμό του ιππόκαμπου (πρώτο αριστερά σχήμα). Τα ΗΕΓ τμήματα των υποσυνόλων N και F ελήφθησαν από όλες τις επαφές των σχετικών εν των βάθει ηλεκτροδίων. Ηλεκτρόδια λωρίδας εμφυτεύονται στις πλευρικές και βασικές περιοχές του νεοφλοιού (μεσαίο και τελευταίο δεξιά σχήμα). Τα ΗΕΓ τμήματα του υποσυνόλου S ελήφθησαν από τις επαφές όλων των ηλεκτροδίων που απεικονίζονται.

Τα ΗΕΓ τμήματα των υποσυνόλων N, F και S προέρχονται από πολύωρες ΗΕΓ καταγραφές οι οποίες έχουν καταγραφεί με ενδοκρανιακά ηλεκτρόδια και οι οποίες χρησιμοποιήθηκαν για προεγχειρητική εκτίμηση της κατάστασης των επιληπτικών ασθενών. Πιο συγκεκριμένα, εν τω βάθει ηλεκτρόδια (depth electrodes) τοποθετήθηκαν συμμετρικά στο σχηματισμό του ιππόκαμπου. Επιπλέον, ηλεκτρόδια λωρίδας (strip electrodes) εμφυτεύθηκαν στις πλευρικές και βασικές περιοχές του νεοφλοιού. Τα ΗΕΓ τμήματα των υποσυνόλων N και F ελήφθησαν από όλες τις επαφές των σχετικών εν των βάθει ηλεκτροδίων, ενώ τα ΗΕΓ τμήματα του υποσυνόλου S ελήφθησαν από τις επαφές όλων των ηλεκτροδίων (εν τω βάθει ηλεκτρόδια και ηλεκτρόδια λωρίδας). Όλα τα ΗΕΓ σήματα καταγράφηκαν με το ίδιο σύστημα ενίσχυσης 128-καναλιών, χρησιμοποιώντας έναν κοινό μέσο όρο αναφοράς (average common reference).

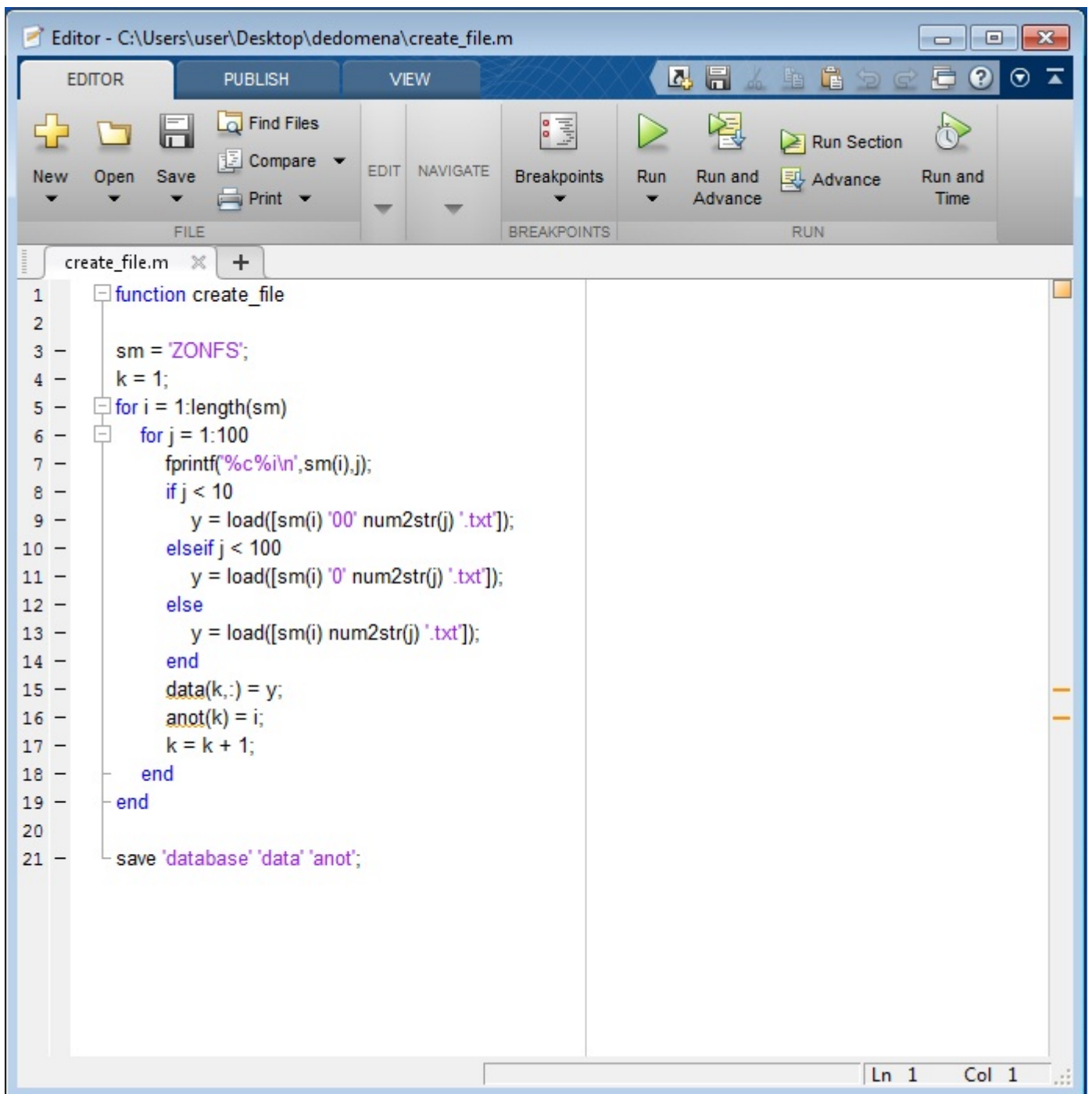


Σχήμα 7.4: Ενδεικτικά ΗΕΓ τμήματα για κάθε ένα από τα 5 υποσύνολα Z, O, N, F, και S αντίστοιχα. Το πλάτος των επιφανειακών ΗΕΓ καταγραφών (ΗΕΓ τμήματα των υποσυνόλων Z και O) είναι της τάξης των μερικών μV . Αντίθετα το πλάτος για τις ενδοκρανιακές καταγραφές (ΗΕΓ τμήματα των υποσυνόλων N και F) κυμαίνεται γύρω στα $100\mu\text{V}$. Τέλος, για την δραστηριότητα της επιληπτικής κρίσης αυτά τα δυναμικά μπορεί να υπερβαίνουν τα $1000\mu\text{V}$.

Μετά την 12-bit αναλογοψηφιακή μετατροπή, τα ΗΕΓ δεδομένα αποθηκεύτηκαν στο σκληρό δίσκο ενός ηλεκτρονικού υπολογιστή (H/Y) με συχνότητα δειγματοληψίας 173.61Hz. Επιπλέον, χρησιμοποιήθηκαν διατάξεις φίλτρων διέλευσης ζώνης συχνοτήτων από 0.5-85Hz (12dB/oct) από το σύστημα ενίσχυσης των 128-καναλιών. Παραδείγματα ΗΕΓ τμημάτων για κάθε ένα από τα υποσύνολα Z, O, N, F, και S φαίνονται στο Σχήμα 7.4. [18], [19]

Το παρακάτω script (Εικόνα 7.1) διαβάζει από το σκληρό δίσκο τα αρχεία των σημάτων (πρώτα διαβάζει το Z001.txt, στη συνέχεια το Z002.txt και μόλις διαβάσει το Z100.txt προχωράει στο O001.txt, ενώ η σειρά για τις κατηγορίες είναι πρώτα η Z στη συνέχεια η O και μετά η N,F,S) και δημιουργεί δύο πίνακες, τον data και τον anot, και ένα αρχείο .mat, το database.mat που είναι η βάση και ουσιαστικά περιέχει τους πίνακες data και anot. Ο πίνακας data αποτελείται από 500 γραμμές και η κάθε γραμμή περιλαμβάνει ένα σήμα (η πρώτη γραμμή το Z001.txt, η δεύτερη το Z001.txt κ.ο.κ) και ο πίνακας anot περιέχει 500 θέσεις οι οποίες υποδηλώνουν σε ποια κατηγορία ανήκει το κάθε σήμα (ο συμβολισμός των κατηγοριών στον πίνακα είναι αριθμητικός και γίνεται ως εξής : το Z συμβολίζεται με

1, το 0 συμβολίζεται με 2, το N συμβολίζεται με 3, το F συμβολίζεται με 4 και το S συμβολίζεται με 5).



Εικόνα 7.1: Δημιουργία βάσης δεδομένων.

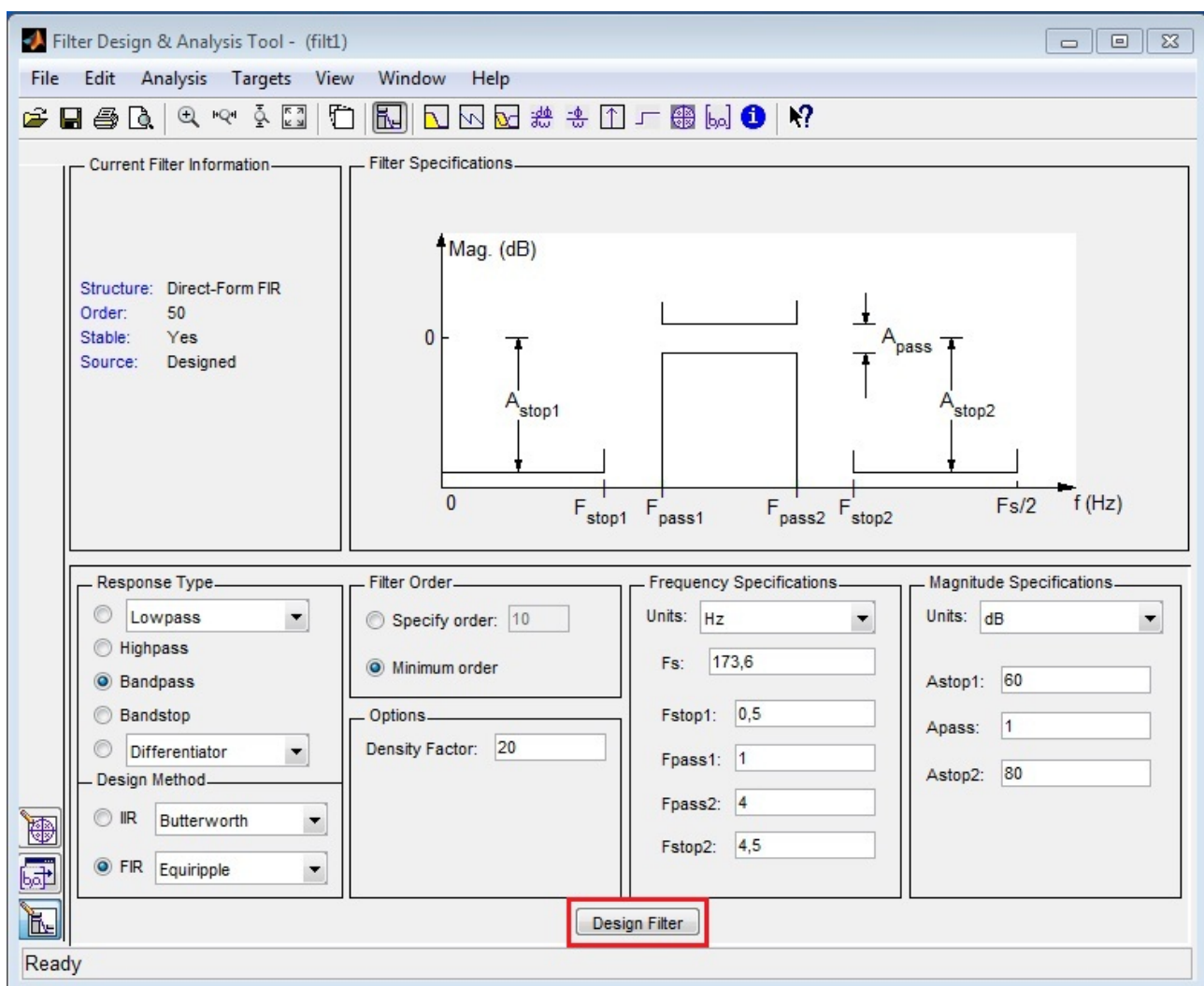
7.3 Σχεδιασμός και υλοποίηση φίλτρων

Η δημιουργία των φίλτρων πραγματοποιήθηκε με την εργαλειοθήκη επεξεργασίας σήματος SPTool (Signal Processing Toolbox) με τις εξής παραμέτρους:

Στο **Response Type** επιλέγεται το **Bandpass** και στο **Design Method** επιλέγεται **FIR** (το οποίο αναλύεται στην ενότητα 7.3.1).

Έπειτα, στο **Filter Order** επιλέγεται **Minimum Order** (προεπιλογή) και στο **Options** σαν Density Factor αφήνεται η τιμή 20 (προεπιλογή).

Εν συνεχεία, ρυθμίζονται τα χαρακτηριστικά των συχνοτήτων (Frequency Specifications) ανάλογα με τους ρυθμούς του ηλεκτροεγκεφαλογραφήματος οι οποίοι αναφέρονται αναλυτικά στην ενότητα 4.3. Για παράδειγμα, για να σχεδιαστεί το πρώτο φίλτρο χαμηλής συχνότητας (0-4 Hz), επιλέγεται σαν συχνότητα δειγματοληψίας $F_s=173.6$ Hz που ισχύει για όλα τα φίλτρα, $F_{stop1}=0.5$, $F_{pass1}=1$, $F_{pass2}=4$, $F_{stop2}=4.5$ και **Magnitude Specifications** αφήνεται η προεπιλογή. Η ίδια διαδικασία ακολουθείται και για τα υπόλοιπα φίλτρα.



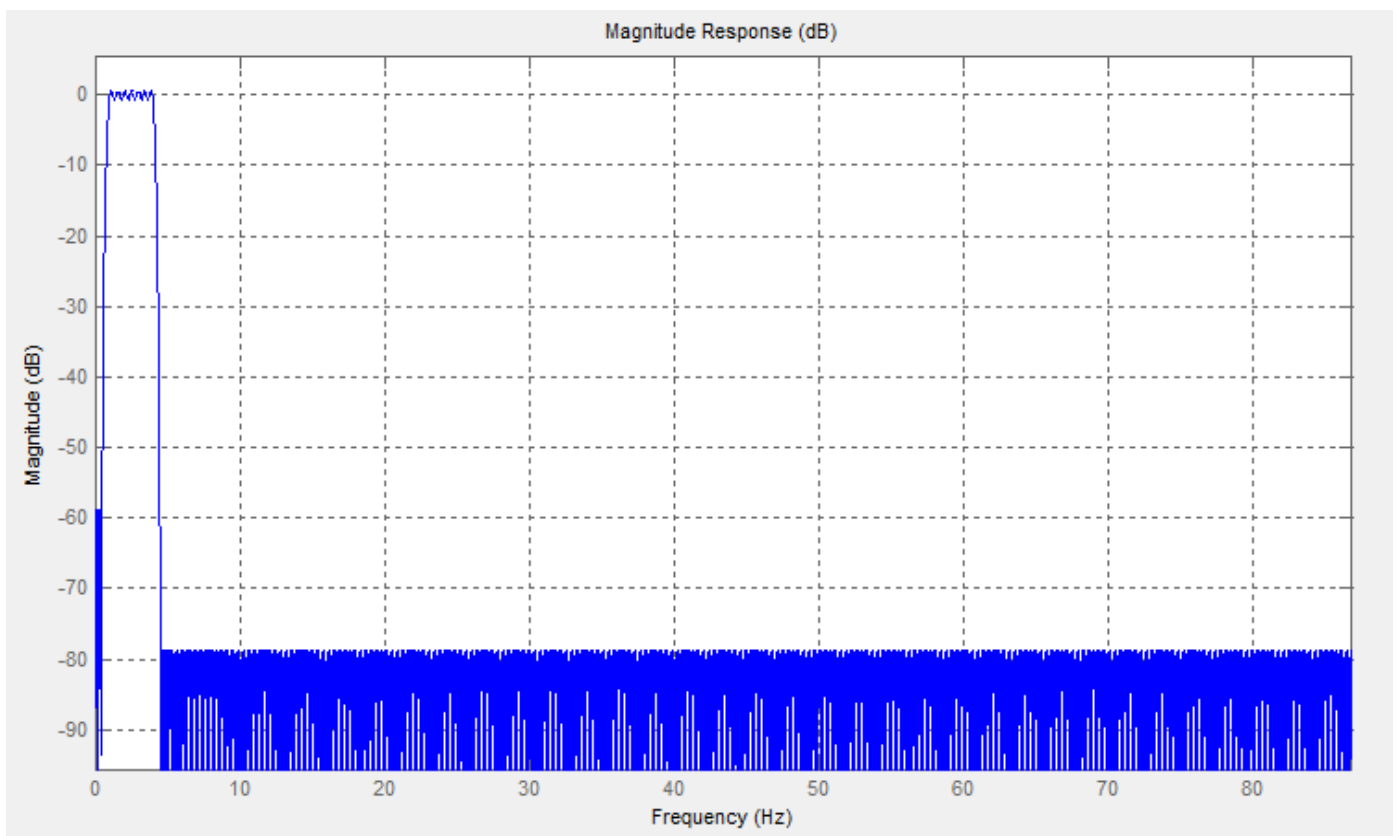
Εικόνα 7.2: Το περιβάλλον SPTool για δημιουργία φίλτρων.

Πατώντας το κουμπί **Design Filter** υπολογίζονται οι παράμετροι του κάθε φίλτρου. Στον πίνακα 7.1 παρουσιάζονται οι παράμετροι που χρησιμοποιήθηκαν για όλα τα φίλτρα.

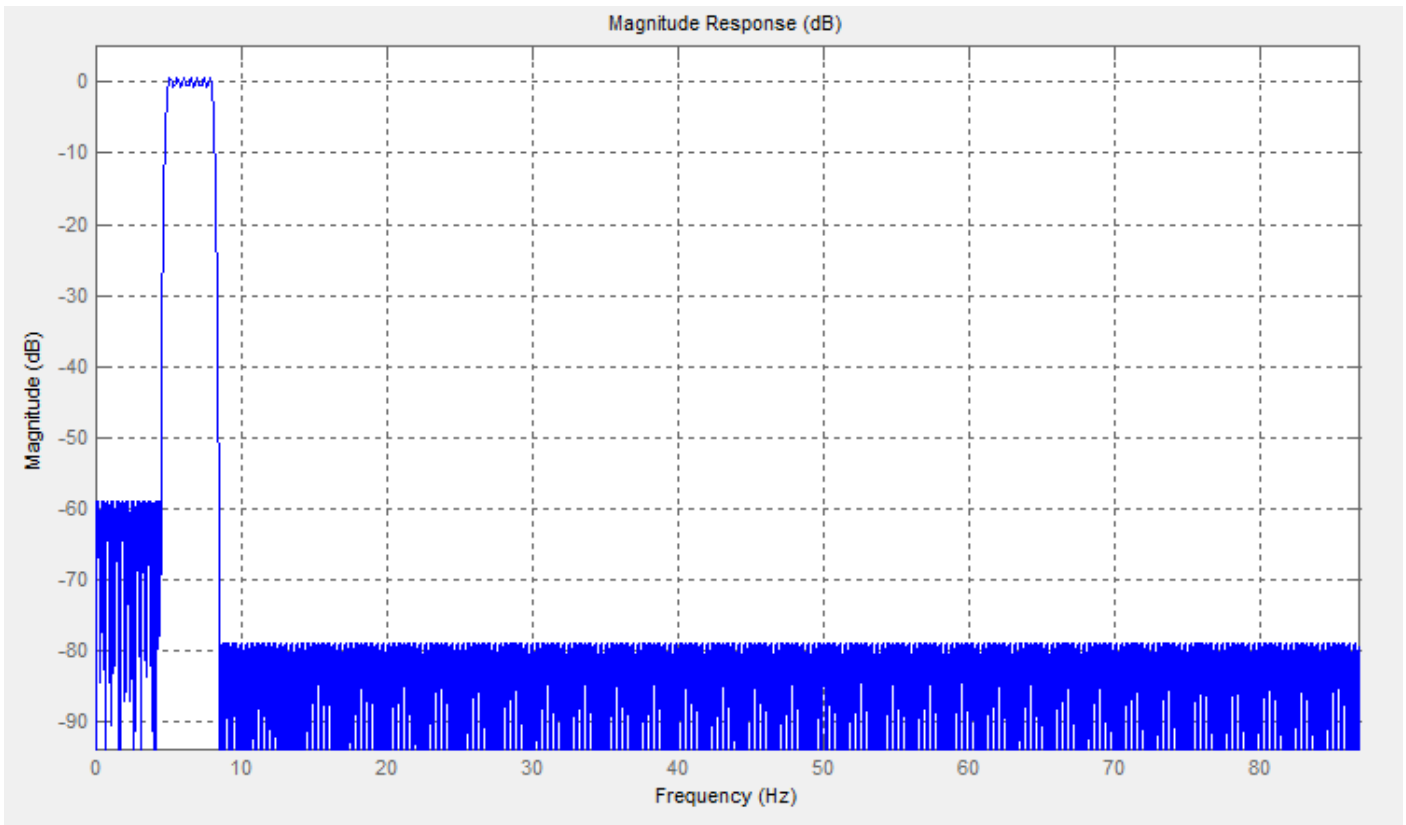
ΠΑΡΑΜΕΤΡΟΙ	ΦΙΛΤΡΑ			
	<i>filt1(0-4 Hz)</i>	<i>filt2(4-8 Hz)</i>	<i>filt3(8-13 Hz)</i>	<i>filt4(13-35 Hz)</i>
<i>F_s</i>	173.6	173.6	173.6	173.6
<i>F_{stop1}</i>	0.5	4.5	8.5	13.5
<i>F_{pass1}</i>	1	5	9	14
<i>F_{pass2}</i>	4	8	13	35
<i>F_{stop2}</i>	4.5	8.5	13.5	35.5

Πίνακας 7.1: Τιμές των παραμέτρων που χρησιμοποιήθηκαν για τον σχεδιασμό των φίλτρων.

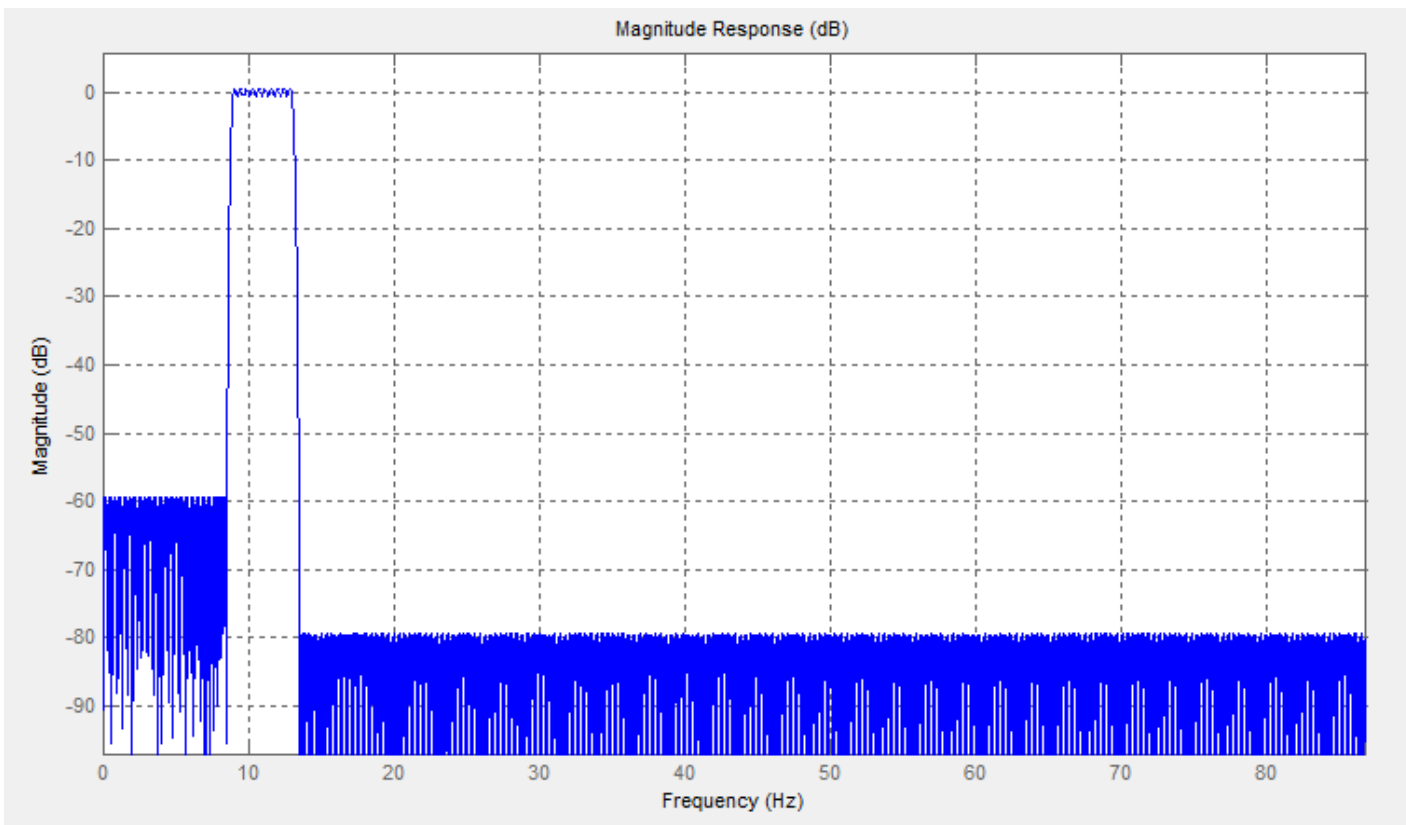
Στις εικόνες 7.3, 7.4, 7.5 και 7.6 παρουσιάζεται η απόκριση συχνότητων για κάθε ένα από τα 4 φίλτρα.



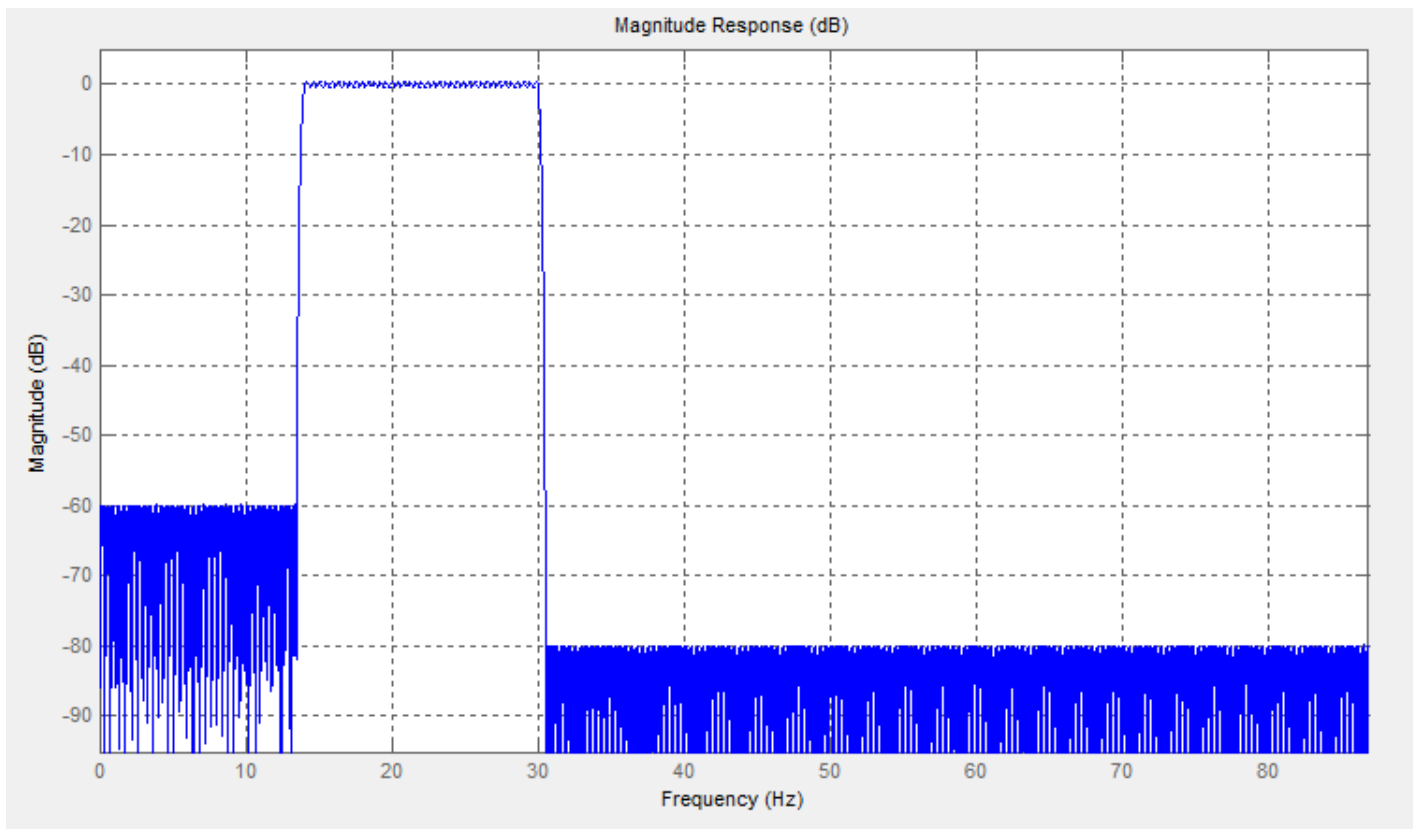
Εικόνα 7.3: Απόκριση συχνότητων πρώτου φίλτρου (*filt1*) χαμηλής συχνότητας 0-4 Hz (Ρυθμός Δέλτα).



Εικόνα 7.4: Απόκριση συχνότητων δεύτερου φίλτρου (filt2) συχνότητας 4-8 Hz (Ρυθμός Θήτα).



Εικόνα 7.5: Απόκριση συχνότητων τρίτου φίλτρου (filt3) συχνότητας 8-13 Hz (Ρυθμός Άλφα).



Εικόνα 7.6: Απόκριση συχνοτήτων τέταρτου φίλτρου (filt4) συχνότητας 13-35 Hz (Ρυθμός Βήτα).

7.3.1 FIR (Finite Impulse Response) Filters-Πεπερασμένης Κρουστικής Απόκρισης Φίλτρα

Τα FIR φίλτρα αποτελούν μερική περίπτωση των φίλτρων άπειρης κρουστικής απόκρισης (IIR). Από τα γραμμικά ψηφιακά φίλτρα, συνήθως προτιμούνται τα FIR φίλτρα έναντι των IIR λόγω της ευστάθειας στη λειτουργία τους και της γραμμικής απόκρισης που παρουσιάζουν στη φάση.

Οι βασικές κατηγορίες των FIR φίλτρων είναι: τα φίλτρα μέσης τιμής, η μέθοδος Μετασχηματισμός Fourier ή μέθοδος των παραθύρων, φίλτρα ισοκυματικά-βέλτιστα και τα φίλτρα με δειγματοληψία συχνότητας. Οι τύποι των ψηφιακών φίλτρων που μπορούν να σχεδιαστούν ως FIR φίλτρα ποικίλουν και μπορούν να έχουν όλες τις μορφές όπως βαθυπερατά (Low-pass), υψηπερατά (High-pass) και ζωνοπερατά (Band-pass). Επίσης οι προδιαγραφές τους μπορεί να είναι πολύ αυστηρές τόσο στη ζώνη

διέλευσης των συχνοτήτων όσο και στη ζώνη αποκοπής.

Η μορφή των FIR φίλτρων δίνεται από την συνάρτηση

$$y(n) = \sum_{k=0}^N b_k x(n - k),$$

στην οποία βλέπουμε ότι η έξοδος $y(n)$ είναι ο γραμμικός συνδυασμός των δειγμάτων εισόδου $x(n)$. Η γραμμική φάση αποτελεί τα βασικά χαρακτηριστικά των φίλτρων αυτών. Η ικανή και αναγκαία συνθήκη για γραμμική φάση είναι η συμμετρία των συντελεστών του FIR φίλτρου. Ένα άλλο πολύ σημαντικό χαρακτηριστικό είναι η ευστάθεια, η οποία οφείλεται στο γεγονός ότι τα FIR φίλτρα δεν έχουν κλάδο ανατροφοδότησης και έτσι η συνάρτηση μεταφοράς του φίλτρου μηδενίζεται. Στα FIR φίλτρα η έξοδος εξαρτάται μόνο από τις τιμές της εισόδου $x(n-k)$ με $k=0, N-1$.

Η κρουστική απόκριση είναι πεπερασμένη και η συνάρτηση μεταφοράς του φίλτρου $H(z)$ δεν έχει παρονομαστή, επομένως δεν έχει πόλους που δημιουργούν αστάθεια. Όλα αυτά έχουν ως αποτέλεσμα την απόλυτη ευστάθεια του φίλτρου που σημαίνει ότι για οποιοδήποτε πεπερασμένο σε μέγεθος και διάρκεια σήμα εισόδου, η έξοδος είναι επίσης πεπερασμένη σε μέγεθος και χρονική διάρκεια.

Η διαδικασία για τη σχεδίαση FIR φίλτρων περιλαμβάνει τα ακόλουθα στάδια:

- Καθορισμός των προδιαγραφών του φίλτρου.
- Υπολογισμός των συντελεστών του φίλτρου με μία από τις διαθέσιμες μεθόδους.
- Υλοποίηση με λογισμικό ή υλικό.

Οι προδιαγραφές των FIR φίλτρων αφορούν στο μέτρο και στη φάση της απόκρισης συχνότητας. Η απόκριση φάσης μπορεί να είναι γραμμική αν η κρουστική απόκριση είναι συμμετρική ή αντισυμμετρική. Η απόκριση μέτρου καθορίζεται με βάση τα παρακάτω χαρακτηριστικά για ένα βαθυπερατό FIR φίλτρο:

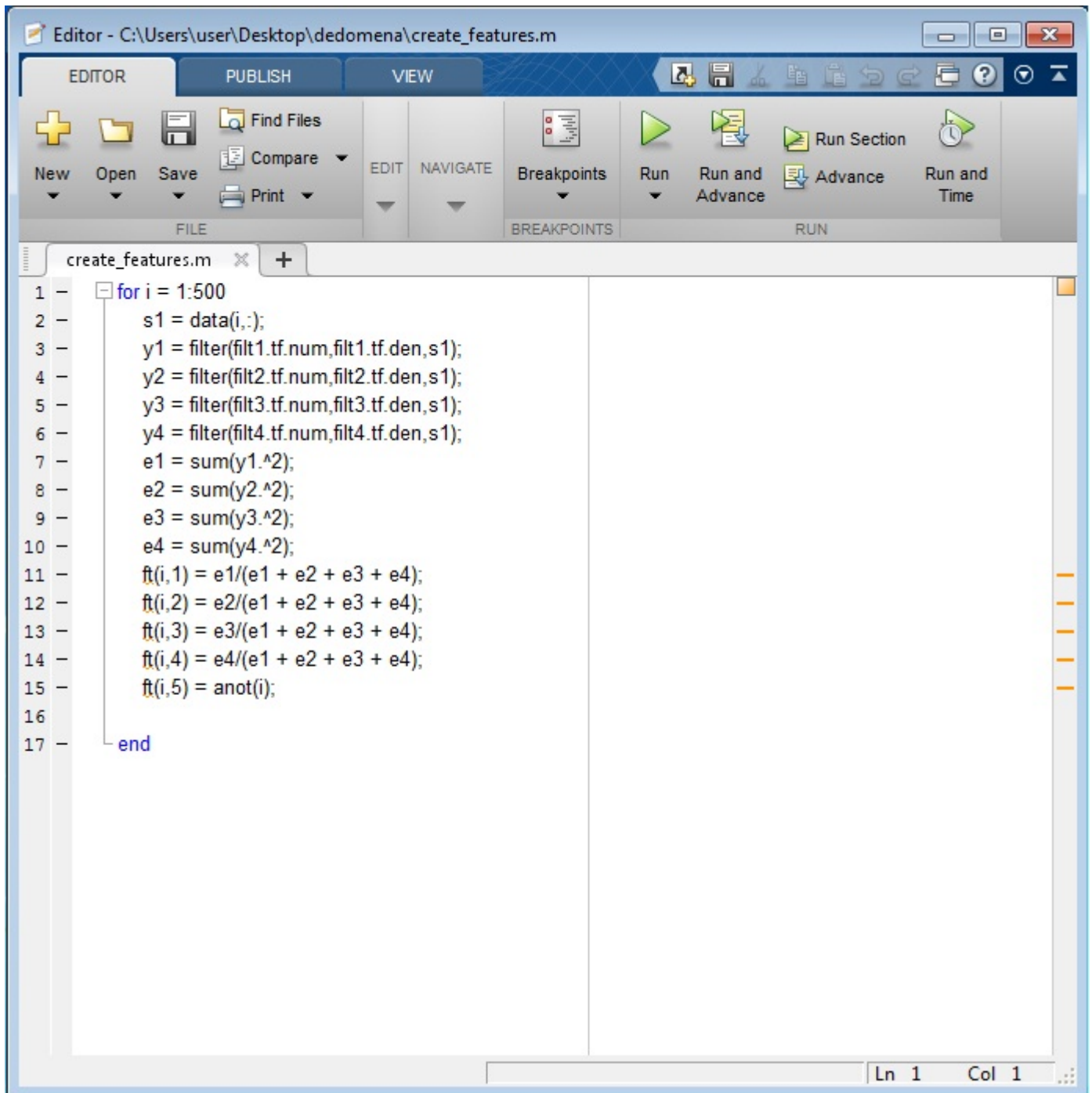
- δ_r απόκλιση στη ζώνη διέλευσης
- δ_s απόκλιση στη ζώνη αποκοπής
- f_r όριο της ζώνης διέλευσης
- f_s όριο της ζώνης αποκοπής
- F_s συχνότητα δειγματοληψίας.

Χρησιμοποιώντας τα φίλτρα πεπερασμένης κρουστικής απόκρισης καθίσταται δυνατή η υλοποίηση φίλτρων με τον επιθυμητό αριθμό ακριβείας. Η μέθοδος σχεδίασης βέλτιστων FIR φίλτρων είναι η καλύτερη επιλογή όταν είναι διαθέσιμο το αντίστοιχο λογισμικό. Η μέθοδος παρέχει πλήρη έλεγχο της κυμάτωσης στη ζώνη διέλευσης και αποκοπής. [20]

7.4 Εφαρμογή φίλτρων και εξαγωγή χαρακτηριστικών

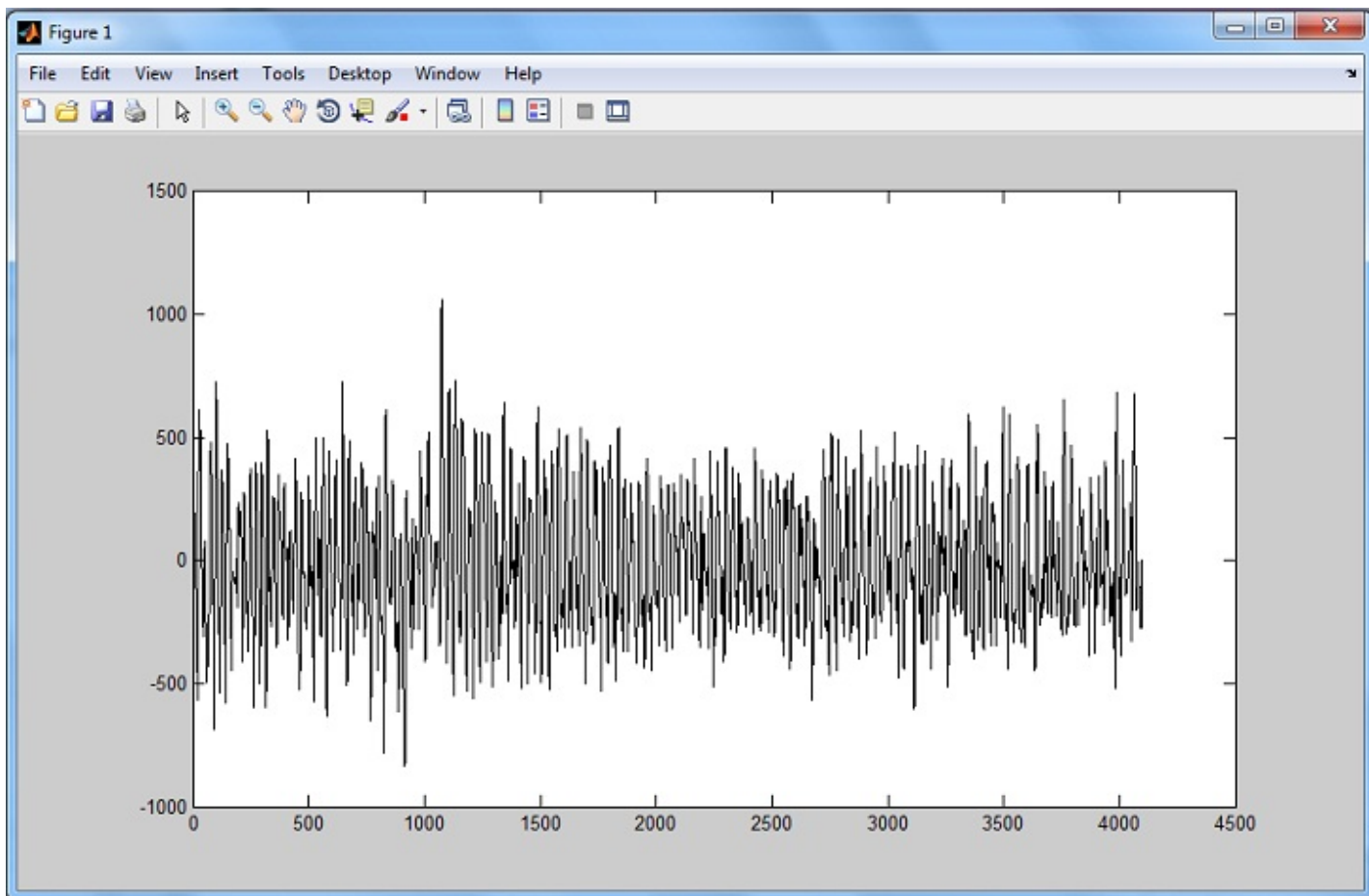
Στο παρακάτω εικόνα παρουσιάζεται ένα πρόγραμμα matlab, το οποίο υλοποιήθηκε για το φιλτράρισμα του αρχικού σήματος s_1 (πίνακας data), με τα 4 ζωνοπερατά φίλτρα διαφορετικών συχνοτήτων τα οποία έχουν εξαχθεί από το Signal Processing Toolbox (SPTool) σύμφωνα με τους ρυθμούς του ηλεκτροεγκεφαλογραφήματος.

Για κάθε φιλτραρισμένο σήμα (y_1, y_2, y_3, y_4) υπολογίζεται η ενέργεια (e_1, e_2, e_3, e_4) και τα ποσοστά ενέργειας σε κάθε περιοχή (πίνακας ft). Τέλος προστίθεται η κατηγορία ως πέμπτο χαρακτηριστικό, που θα χρησιμοποιηθεί για την διαδικασία εκπαίδευσης και αξιολόγησης των ταξινομητών.

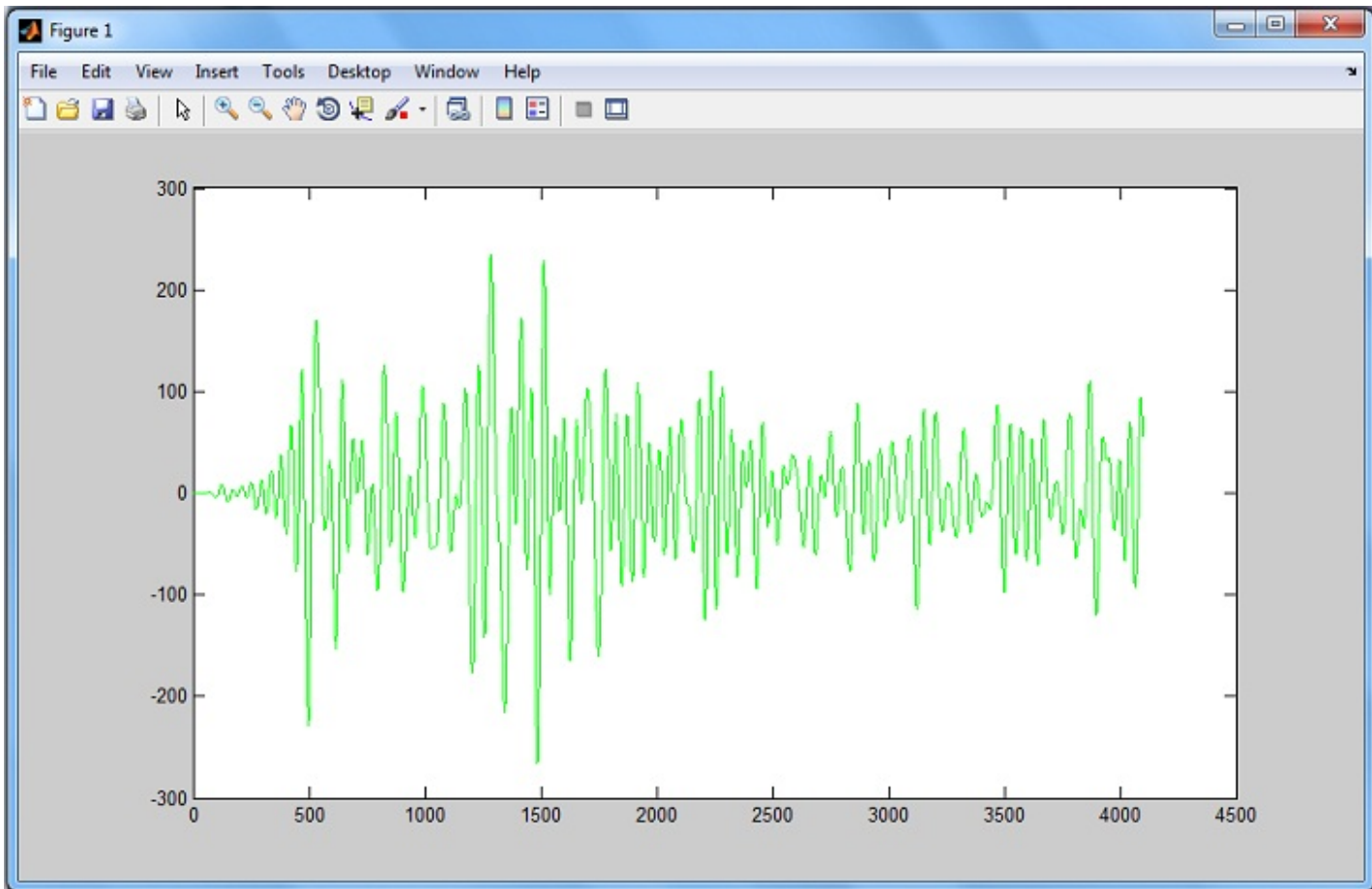


```
1 - for i = 1:500
2 -     s1 = data(i,:);
3 -     y1 = filter(filt1.tf.num,filt1.tf.den,s1);
4 -     y2 = filter(filt2.tf.num,filt2.tf.den,s1);
5 -     y3 = filter(filt3.tf.num,filt3.tf.den,s1);
6 -     y4 = filter(filt4.tf.num,filt4.tf.den,s1);
7 -     e1 = sum(y1.^2);
8 -     e2 = sum(y2.^2);
9 -     e3 = sum(y3.^2);
10 -    e4 = sum(y4.^2);
11 -    ft(i,1) = e1/(e1 + e2 + e3 + e4);
12 -    ft(i,2) = e2/(e1 + e2 + e3 + e4);
13 -    ft(i,3) = e3/(e1 + e2 + e3 + e4);
14 -    ft(i,4) = e4/(e1 + e2 + e3 + e4);
15 -    ft(i,5) = anot(i);
16 -
17 - end
```

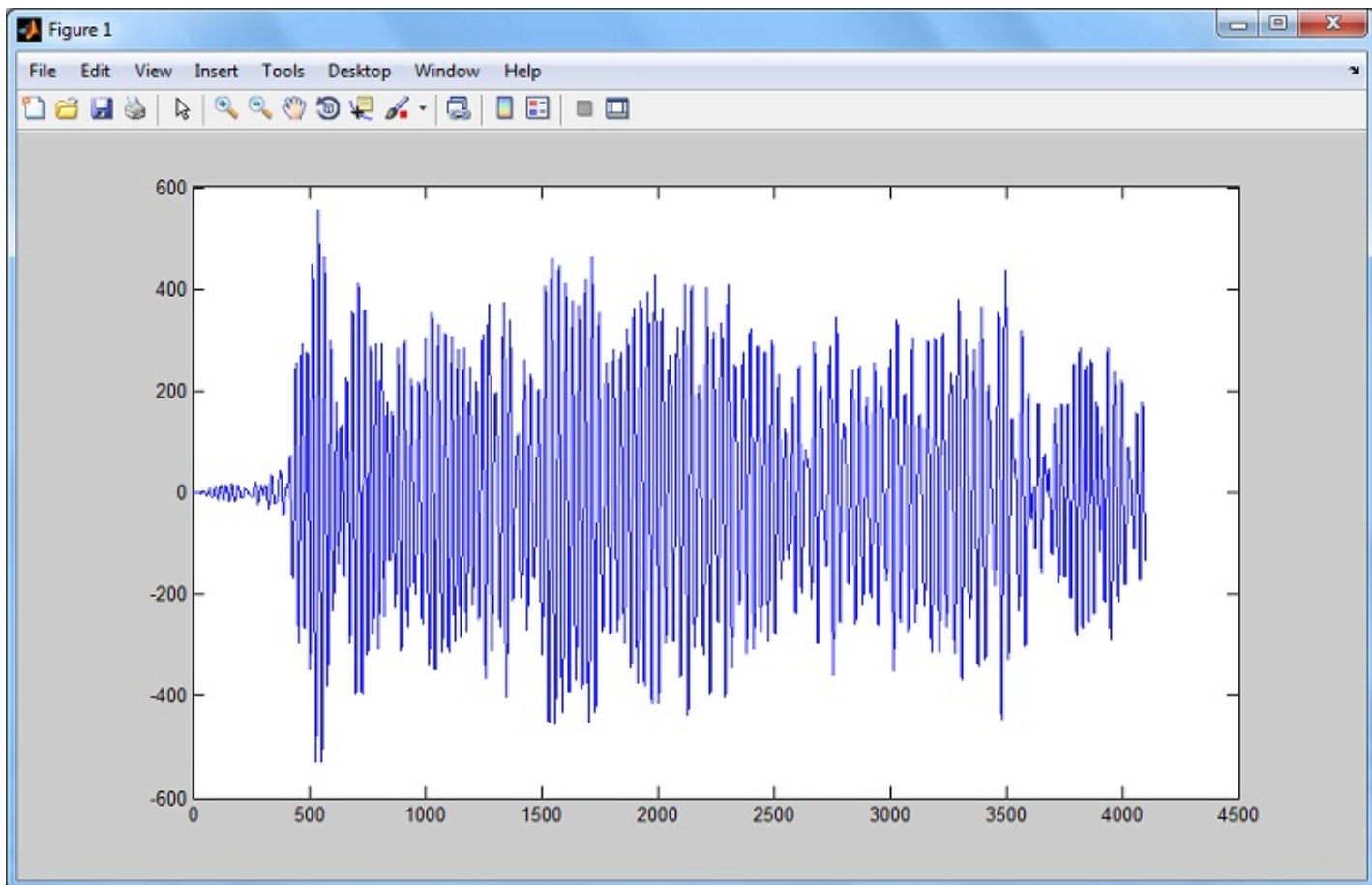
Εικόνα 7.7: Φιλτράρισμα σήματος με 4 ζωνοπερατά φίλτρα.



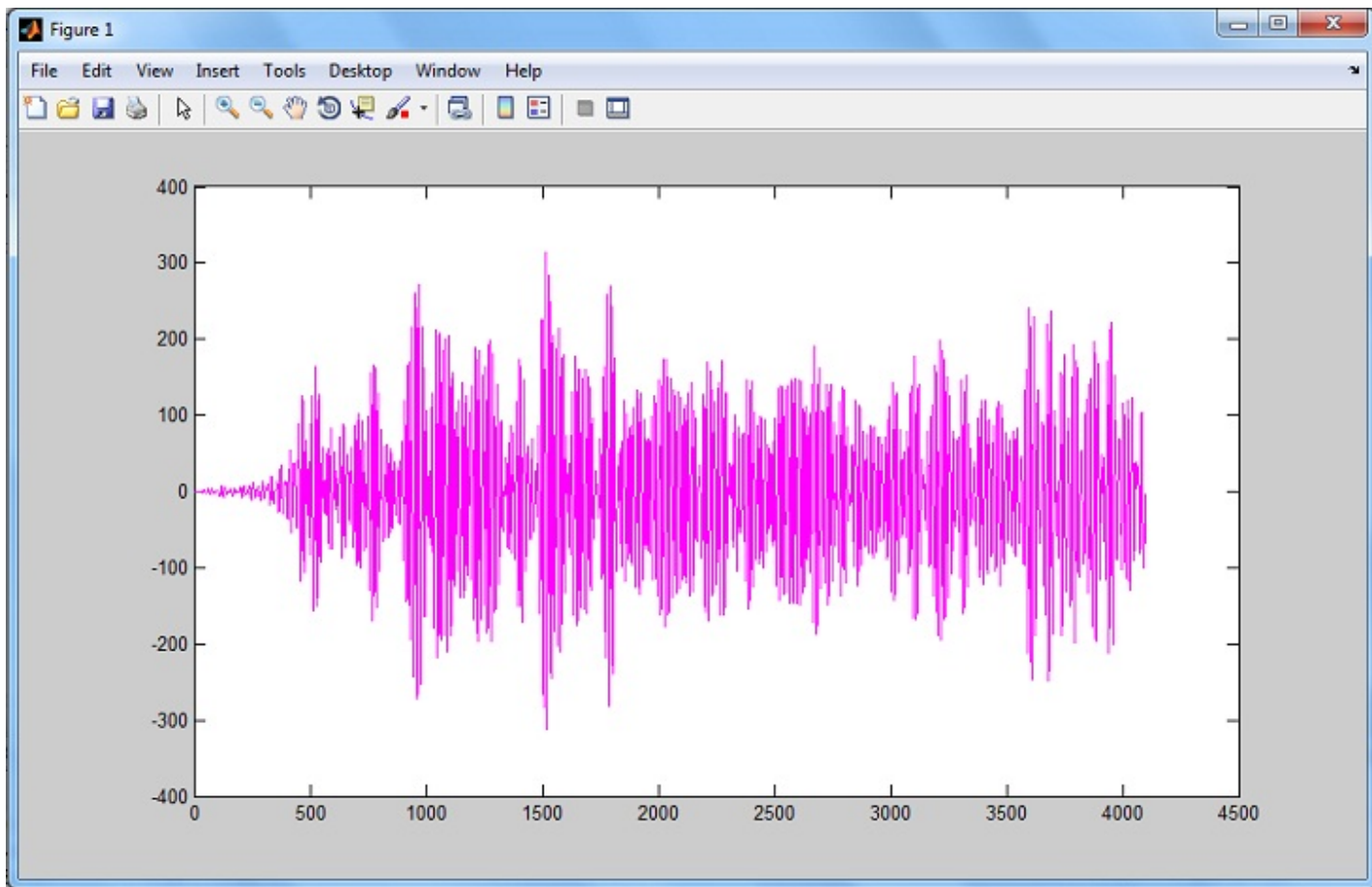
Εικόνα 7.8: Το αρχικό σήμα s_1 .



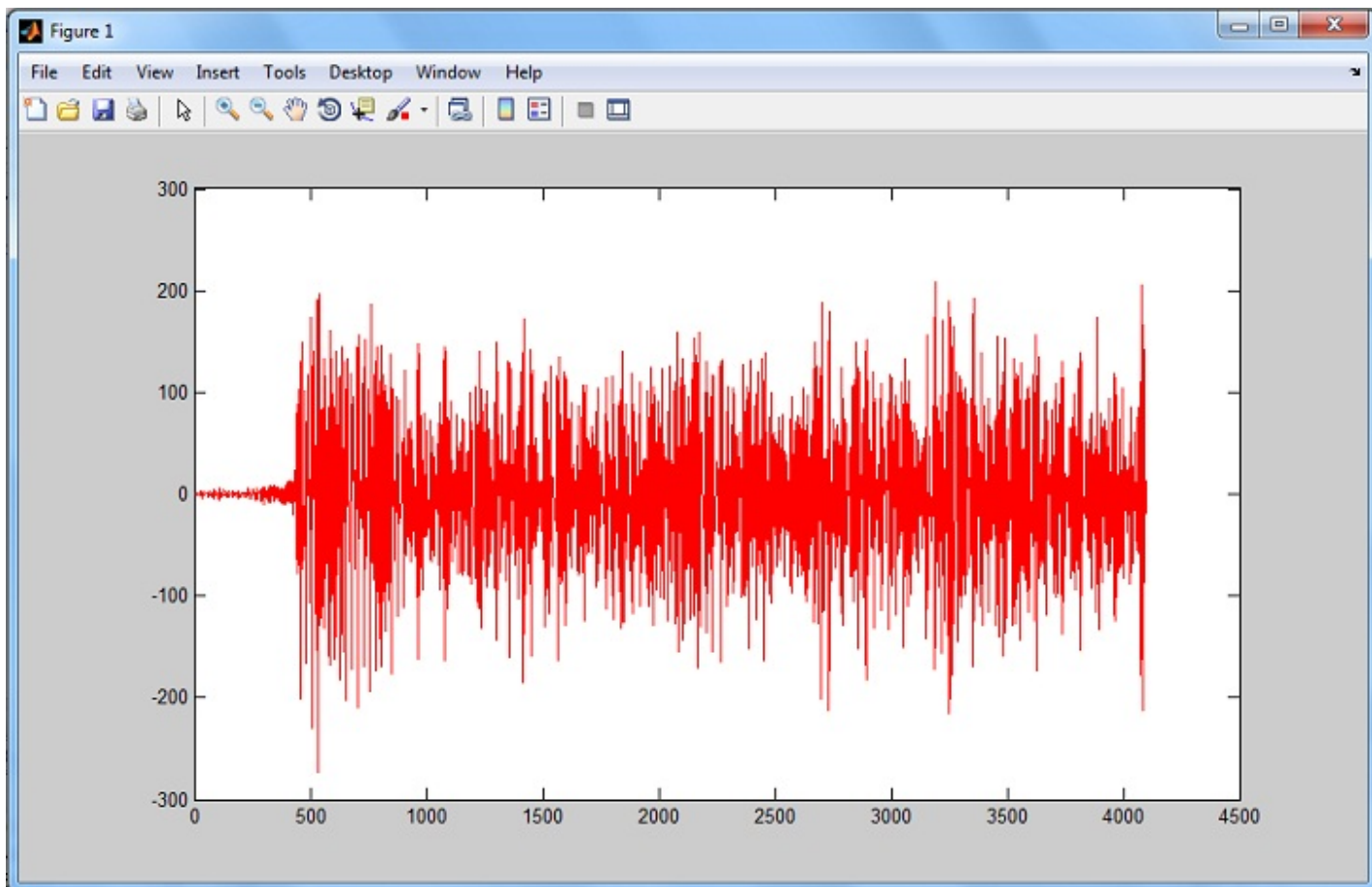
Εικόνα 7.9: Το αρχικό σήμα φιλτραρισμένο με ζωνοπερατό φίλτρο με κατώφλια 0.5-4 Hz.



Εικόνα 7.10: Το αρχικό σήμα φιλτραρισμένο με ζωνοπερατό φίλτρο με κατώφλια 4-8 Hz.



Εικόνα 7.11: Το αρχικό σήμα φιλτραρισμένο με ζωνοπερατό φίλτρο με κατώφλια 8-13 Hz.



Εικόνα 7.12: Το αρχικό σήμα φιλτραρισμένο με ζωνοπερατό φίλτρο με κατώφλια 13-35 Hz.

7.5 Εφαρμογή classifiers και εξαγωγή αποτελεσμάτων στο weka

Προκειμένου να χρησιμοποιηθούν τα χαρακτηριστικά της βάσης δεδομένων στο εργαλείο weka πρέπει να γίνει μια προετοιμασία στα δεδομένα. Στόχος αυτής της προετοιμασίας είναι να φέρουμε τα δεδομένα που έχουμε συλλέξει, σε μια τέτοια μορφή κατάλληλη ώστε να επιτρέπει η εφαρμογή αλγορίθμων εξόρυξης γνώσης. Έτσι λοιπόν δημιουργήθηκαν 5 διαφορετικά προβλήματα, άλλοτε με αυτούσιες κατηγορίες και άλλοτε χρησιμοποιώντας την συνένωση αυτών κάνοντας διάφορες παραλλαγές πειραμάτων. Για τον σχηματισμό των αντίστοιχων συνόλων δεδομένων που χρησιμοποιούνται σε κάθε πρόβλημα χρησιμοποιείται η βάση δεδομένων που περιγράφεται στην ενότητα 6.2. Κάθε ένα από τα σύνολα δεδομένων αποτελεί υποσύνολο της αρχικής βάσης δεδομένων. Στον πίνακα 7.2 παρουσιάζονται τα 5 διαφορετικά προβλήματα που δημιουργήθηκαν.

Πρόβλημα Ταξινόμησης	Κατηγορίες	Αριθμός σημάτων
1	Normal(Z)	100
	Normal(O)	100
	Non Seizure(N)	100
	Non Seizure(F)	100
	Seizure(S)	100
	Σύνολο	500
2	Normal(Z,O)	200
	Non Seizure(N,F)	200
	Seizure(S)	100
	Σύνολο	500
3	Non Seizure(Z,O,N,F)	400
	Seizure(S)	100
	Σύνολο	500
4	Normal(Z)	100
	Non Seizure(N)	100
	Seizure(S)	100
	Σύνολο	300
5	Normal(Z)	100
	Seizure(S)	100
	Σύνολο	200

Πίνακας 7.2: Οι κατηγορίες και ο αντίστοιχος αριθμός των ΗΕΓ καταγραφών για τα πέντε διαφορετικά προβλήματα.

Αναλυτικότερα:

1. Η πρώτη περίπτωση αφορά τις πέντε κατηγορίες στις οποίες είναι χωρισμένες οι ΗΕΓ καταγραφές στη βάση δεδομένων, οι οποίες δηλώνονται ως Z, O, N, F, και S. Το υποσύνολο Z είναι καταγραφή φυσιολογικού ΗΕΓ με μάτια ανοικτά και το υποσύνολο O είναι καταγραφή φυσιολογικού ΗΕΓ με μάτια κλειστά. Οι ΗΕΓ καταγραφές των δύο υποσυνόλων N και F περιείχαν μόνο δραστηριότητα που μετρήθηκε σε διαστήματα τα οποία δεν περιείχαν καταγεγραμμένη επιληπτική κρίση. Στην περίπτωση του υποσύνολου N η καταγραφή έγινε στην επιληπτογενή ζώνη, ενώ στην περίπτωση του υποσύνολου F, στο σχηματισμό του ιππόκαμπου του, αντίθετου ημισφαιρίου του εγκεφάλου. Το υποσύνολο S περιέχει καταγεγραμμένη δραστηριότητα

επιληπτικής κρίσης (ictal δραστηριότητα).

2. Στη δεύτερη περίπτωση γίνεται συνένωση κατηγοριών με αποτέλεσμα να έχουμε τρεις κατηγορίες. Την κατηγορία με φυσιολογικό ΗΕΓ (ΗΕΓ καταγραφές που προέρχονται από υγιή άτομα) στην οποία ανήκουν τα υποσύνολα Z και O, την κατηγορία ΗΕΓ χωρίς επιληπτική κρίση (ΗΕΓ καταγραφές που προέρχονται από ασθενείς με επιληψία που όμως δεν παρουσιάζουν επιληπτική κρίση κατά τη διάρκεια της καταγραφής) στην οποία συνενώθηκαν τα υποσύνολα N και F και την κατηγορία ΗΕΓ με επιληπτική κρίση που ανήκει το υποσύνολο S (ΗΕΓ καταγραφές που προέρχονται από ασθενείς με επιληψία που παρουσιάζουν επιληπτική κρίση κατά τη διάρκεια της καταγραφής).
3. Στην τρίτη περίπτωση συναντώνται δύο κατηγορίες. Η κατηγορία με φυσιολογικό ΗΕΓ (ΗΕΓ καταγραφές που προέρχονται από υγιή άτομα και καταγραφές που προέρχονται από ασθενείς με επιληψία που όμως δεν παρουσιάζουν επιληπτική κρίση κατά τη διάρκεια της καταγραφής) στην οποία συνενώθηκαν τα υποσύνολα Z, O, N και F και την κατηγορία ΗΕΓ με επιληπτική κρίση που ανήκει το υποσύνολο S (ΗΕΓ καταγραφές που προέρχονται από ασθενείς με επιληψία που παρουσιάζουν επιληπτική κρίση κατά τη διάρκεια της καταγραφής).
4. Στην τέταρτη περίπτωση δεν χρησιμοποιούνται όλες οι ΗΕΓ καταγραφές από την βάση δεδομένων. Επιλέχθηκαν μόνο οι καταγραφές οι οποίες ανήκουν στις κατηγορίες Z (με φυσιολογικό ΗΕΓ), N (ΗΕΓ χωρίς επιληπτική κρίση) και S (ΗΕΓ με επιληπτική κρίση).
5. Και στην τελευταία περίπτωση επιλέγονται από την βάση δεδομένων δύο κατηγορίες και δεν χρησιμοποιούνται όλες οι ΗΕΓ καταγραφές. Οι κατηγορίες αυτές είναι: η Z (με φυσιολογικό ΗΕΓ) και η S (ΗΕΓ με επιληπτική κρίση).

Στα παραπάνω προβλήματα θα εφαρμοστούν οι τεχνικές εξόρυξης γνώσης που αναφέρθηκαν αναλυτικά στο προηγούμενο κεφάλαιο (κεφάλαιο 5). Οι αλγόριθμοι ταξινόμησης που εφαρμόστηκαν για την εξαγωγή αποτελεσμάτων φαίνονται στον παρακάτω πίνακα (Πίνακας 7.3) με τις παραμέτρους που χρησιμοποιήθηκαν.

Ταξινομητής	Συντόμευση	Παράμετρος
Naïve Bayes	NB	
k Nearest Neighbor	KNN5	k=5
	KNN10	k=10
Decision Tree (C4.5)	J48_M2	M=2
	J48_M5	M=5
	J48_M9	M=9
Random Forest	RF	
Multiplayer Perceptron	MLP5	H=5
	MLP10	H=10
	MLP20	H=20

Πίνακας 7.3: Οι αλγόριθμοι ταξινόμησης με τις παραμέτρους τους.

8 Αποτελέσματα-Συμπεράσματα

Προκειμένου να διαπιστωθεί η αποτελεσματικότητα των διαφόρων αλγορίθμων που χρησιμοποιήθηκαν, γίνεται σύγκριση μεταξύ αυτών ώστε να διαπιστωθεί ποιος είναι ο καλύτερος. Αυτό που μας ενδιαφέρει στη σύγκριση αυτή είναι ποιος αλγόριθμος δίνει την καλύτερη ακρίβεια.

Για να πραγματοποιηθούν οι μετρήσεις ως προς την ακρίβεια ταξινόμησης των προγνωστικών μοντέλων χρησιμοποιήθηκε η τεχνική της δεκαπλής στρωματοποιημένης αντεπικύρωσης (10-fold stratified cross validation), η οποία είναι μια ευρέως χρησιμοποιούμενη τεχνική ελέγχου. Στην k-fold τεχνική το σύνολο δεδομένων χωρίζεται τυχαία σε k ανεξάρτητα υποσύνολα ίδιου μεγέθους. Ύστερα κατασκευάζεται το μοντέλο πρόγνωσης με τον αλγόριθμο εξόρυξης γνώσης χρησιμοποιώντας τα k-1 σύνολα. Το σύνολο που απομένει το χρησιμοποιείται για να ελεγχθεί η απόδοση του αλγορίθμου. Αυτή η διαδικασία επαναλαμβάνεται για κάθε ένα από τα k σύνολα που έχουμε. Έτσι γίνεται μια μέτρηση της ακρίβειας της κατηγοριοποίησης για κάθε ένα από τα k σύνολα. Η τελική ακρίβεια του αλγορίθμου είναι ο μέσος όρος των ακριβειών των επιμέρους μετρήσεων.

8.1 Αποτελέσματα

Ένα χαρακτηριστικό που είναι χρήσιμο να χρησιμοποιηθεί για την αξιολόγηση των αλγορίθμων που μελετήθηκαν είναι ο πίνακας σύγχυσης ή αλλιώς πίνακας ενδεχομένων (confusion matrix), ο οποίος συνοψίζει τα αποτελέσματα μετά την εξέταση του συνόλου δοκιμής στον αλγόριθμο. Στην διαγώνιο του παρουσιάζονται τα πρότυπα που έχουν ταξινομηθεί σωστά ανά κλάση και στις υπόλοιπες θέσεις τα πρότυπα που έχουν ταξινομηθεί λανθασμένα. Από τον πίνακα σύγχυσης ταξινόμησης υπολογίζονται τα εξής:

- Η ακριβής ταξινόμηση
- Η ευαισθησία και
- Η ακριβής θετική πρόβλεψη.

Η ακριβής ταξινόμηση ορίζεται από το άθροισμα των στοιχείων της κυρίας διαγωνίου δια το άθροισμα όλων των υπολοίπων στοιχείων του πίνακα. Η ευαισθησία ορίζεται από το άθροισμα των στοιχείων της κυρίας διαγωνίου δια το άθροισμα όλων των στοιχείων της κάθε γραμμής. Η ακριβής θετική πρόβλεψη ορίζεται από το άθροισμα των στοιχείων της κυρίας διαγωνίου δια το άθροισμα όλων των στοιχείων της κάθε στήλης. Επιπλέον υπολογίστηκε ο σταθμισμένος μέσος όρος για την ευαισθησία και την ακριβή θετική πρόβλεψη.

Πίνακας 8.1: Πίνακας σύγκρισης ταξινόμησης για το πρώτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον Αφελή ταξινομητή Bayes.

		Ταξινόμηση				
		<i>Z</i>	<i>O</i>	<i>N</i>	<i>F</i>	<i>S</i>
Βάση Δεδομένων	<i>Z</i>	88	1	3	3	5
	<i>O</i>	7	89	0	1	3
	<i>N</i>	5	0	79	14	2
	<i>F</i>	7	0	58	30	5
	<i>S</i>	5	9	8	8	70

Πίνακας 8.2: Αποτελέσματα για το πρώτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον Αφελή ταξινομητή Bayes. Η ακρίβεια ταξινόμησης είναι 71.2%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>Z</i>	0.88	0.786
<i>O</i>	0.89	0.899
<i>N</i>	0.79	0.534
<i>F</i>	0.3	0.536
<i>S</i>	0.7	0.824
Σταθμισμένος μέσος όρος	0.712	0.716

Πίνακας 8.3: Πίνακας σύγκρισης ταξινόμησης για το πρώτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με k=5.

		Ταξινόμηση				
		<i>Z</i>	<i>O</i>	<i>N</i>	<i>F</i>	<i>S</i>
Βάση Δεδομένων	<i>Z</i>	92	0	2	3	3
	<i>O</i>	6	94	0	0	0
	<i>N</i>	4	0	62	33	1
	<i>F</i>	6	0	41	52	1
	<i>S</i>	7	6	3	5	79

Πίνακας 8.4: Αποτελέσματα για το πρώτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με $k=5$. Η ακρίβεια ταξινόμησης είναι 75.8%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>Z</i>	0.92	0.8
<i>O</i>	0.94	0.94
<i>N</i>	0.62	0.574
<i>F</i>	0.52	0.559
<i>S</i>	0.79	0.94
Σταθμισμένος μέσος όρος	0.758	0.763

Πίνακας 8.5: Πίνακας σύγχυσης ταξινόμησης για το πρώτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με $k=10$.

		Ταξινόμηση				
		<i>Z</i>	<i>O</i>	<i>N</i>	<i>F</i>	<i>S</i>
Βάση Δεδομένων	<i>Z</i>	91	0	2	2	5
	<i>O</i>	7	93	0	0	0
	<i>N</i>	5	0	70	25	0
	<i>F</i>	8	0	48	43	1
	<i>S</i>	11	12	6	6	65

Πίνακας 8.6: Αποτελέσματα για το πρώτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με $k=10$. Η ακρίβεια ταξινόμησης είναι 72.4%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>Z</i>	0.91	0.746
<i>O</i>	0.93	0.886
<i>N</i>	0.7	0.556
<i>F</i>	0.43	0.566
<i>S</i>	0.65	0.915
Σταθμισμένος μέσος όρος	0.724	0.734

Πίνακας 8.7 : Πίνακας σύγχυσης ταξινόμησης για το πρώτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=2$.

		Ταξινόμηση				
		<i>Z</i>	<i>O</i>	<i>N</i>	<i>F</i>	<i>S</i>
Βάση Δεδομένων	<i>Z</i>	84	2	5	4	5
	<i>O</i>	5	90	0	2	3
	<i>N</i>	4	0	78	14	4
	<i>F</i>	1	0	59	38	2
	<i>S</i>	9	3	4	3	81

Πίνακας 8.8: Αποτελέσματα για το πρώτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=2$. Η ακρίβεια ταξινόμησης είναι 74.2%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>Z</i>	0.84	0.816
<i>O</i>	0.9	0.947
<i>N</i>	0.78	0.534
<i>F</i>	0.38	0.623
<i>S</i>	0.81	0.853
Σταθμισμένος μέσος όρος	0.742	0.755

Πίνακας 8.9 : Πίνακας σύγχυσης ταξινόμησης για το πρώτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=5$.

		Ταξινόμηση				
		<i>Z</i>	<i>O</i>	<i>N</i>	<i>F</i>	<i>S</i>
Βάση Δεδομένων	<i>Z</i>	80	2	4	4	10
	<i>O</i>	5	88	0	2	5
	<i>N</i>	4	0	77	16	3
	<i>F</i>	3	1	58	36	2
	<i>S</i>	8	5	6	1	80

Πίνακας 8.10: Αποτελέσματα για το πρώτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=5$. Η ακρίβεια ταξινόμησης είναι 72.2%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>Z</i>	0.8	0.8
<i>O</i>	0.88	0.917
<i>N</i>	0.77	0.531
<i>F</i>	0.36	0.61
<i>S</i>	0.8	0.8
Σταθμισμένος μέσος όρος	0.722	0.732

Πίνακας 8.11: Πίνακας σύγχυσης ταξινόμησης για το πρώτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=9$.

		Ταξινόμηση				
		<i>Z</i>	<i>O</i>	<i>N</i>	<i>F</i>	<i>S</i>
Βάση Δεδομένων	<i>Z</i>	82	1	5	5	7
	<i>O</i>	6	85	0	4	5
	<i>N</i>	4	0	82	11	3
	<i>F</i>	2	0	55	38	5
	<i>S</i>	9	7	7	4	73

Πίνακας 8.12: Αποτελέσματα για το πρώτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=9$. Η ακρίβεια ταξινόμησης είναι 72%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>Z</i>	0.82	0.796
<i>O</i>	0.85	0.914
<i>N</i>	0.82	0.55
<i>F</i>	0.38	0.613
<i>S</i>	0.73	0.785
Σταθμισμένος μέσος όρος	0.72	0.732

Πίνακας 8.13: Πίνακας σύγχυσης ταξινόμησης για το πρώτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή Τυχαίου Δάσους RF.

		Ταξινόμηση				
		<i>Z</i>	<i>O</i>	<i>N</i>	<i>F</i>	<i>S</i>
Βάση Δεδομένων	<i>Z</i>	86	3	2	1	8
	<i>O</i>	5	90	0	1	4
	<i>N</i>	5	0	57	34	4
	<i>F</i>	2	0	39	57	2
	<i>S</i>	6	3	2	4	85

Πίνακας 8.14: Αποτελέσματα για το πρώτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή Τυχαίου Δάσους RF. Η ακρίβεια ταξινόμησης είναι 75%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>Z</i>	0.86	0.827
<i>O</i>	0.9	0.938
<i>N</i>	0.57	0.57
<i>F</i>	0.57	0.588
<i>S</i>	0.85	0.825
Σταθμισμένος μέσος όρος	0.75	0.749

Πίνακας 8.15: Πίνακας σύγχυσης ταξινόμησης για το πρώτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 5 κρυφούς νευρώνες.

		Ταξινόμηση				
		<i>Z</i>	<i>O</i>	<i>N</i>	<i>F</i>	<i>S</i>
Βάση Δεδομένων	<i>Z</i>	86	1	1	3	9
	<i>O</i>	6	90	0	1	3
	<i>N</i>	5	0	74	16	5
	<i>F</i>	3	0	62	31	4
	<i>S</i>	9	6	6	6	73

Πίνακας 8.16: Αποτελέσματα για το πρώτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 5 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 70.8%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>Z</i>	0.86	0.789
<i>O</i>	0.9	0.928
<i>N</i>	0.74	0.517
<i>F</i>	0.31	0.544
<i>S</i>	0.73	0.777
Σταθμισμένος μέσος όρος	0.708	0.711

Πίνακας 8.17: Πίνακας σύγχυσης ταξινόμησης για το πρώτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 10 κρυφούς νευρώνες.

		Ταξινόμηση				
		<i>Z</i>	<i>O</i>	<i>N</i>	<i>F</i>	<i>S</i>
Βάση Δεδομένων	<i>Z</i>	87	2	3	1	7
	<i>O</i>	4	90	0	1	5
	<i>N</i>	4	0	77	16	3
	<i>F</i>	3	0	64	30	3
	<i>S</i>	5	2	4	5	84

Πίνακας 8.18: Αποτελέσματα για το πρώτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 10 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 73.6%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>Z</i>	0.87	0.845
<i>O</i>	0.9	0.957
<i>N</i>	0.77	0.52
<i>F</i>	0.3	0.566
<i>S</i>	0.84	0.824
Σταθμισμένος μέσος όρος	0.736	0.742

Πίνακας 8.19: Πίνακας σύγχυσης ταξινόμησης για το πρώτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 20 κρυφούς νευρώνες.

		Ταξινόμηση				
		<i>Z</i>	<i>O</i>	<i>N</i>	<i>F</i>	<i>S</i>
Βάση Δεδομένων	<i>Z</i>	88	2	3	2	5
	<i>O</i>	5	92	0	1	2
	<i>N</i>	5	0	70	20	5
	<i>F</i>	3	1	59	34	3
	<i>S</i>	4	3	3	3	87

Πίνακας 8.20: Αποτελέσματα για το πρώτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 20 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 74.2%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>Z</i>	0.88	0.838
<i>O</i>	0.92	0.939
<i>N</i>	0.7	0.519
<i>F</i>	0.34	0.567
<i>S</i>	0.87	0.853
Σταθμισμένος μέσος όρος	0.742	0.743

Πίνακας 8.21: Πίνακας σύγχυσης ταξινόμησης για το δεύτερο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον Αφελή ταξινομητή Bayes.

		Ταξινόμηση		
		<i>ZO</i>	<i>NF</i>	<i>S</i>
Βάση Δεδομένων	<i>ZO</i>	188	6	6
	<i>NF</i>	10	177	13
	<i>S</i>	12	17	71

Πίνακας 8.22: Αποτελέσματα για το δεύτερο πρόβλημα ταξινόμησης χρησιμοποιώντας τον Αφελή ταξινομητή Bayes. Η ακρίβεια ταξινόμησης είναι 87.2%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>ZO</i>	0.94	0.895
<i>NF</i>	0.885	0.885
<i>S</i>	0.71	0.789
Σταθμισμένος μέσος όρος	0.872	0.87

Πίνακας 8.23: Πίνακας σύγκρισης ταξινόμησης για το δεύτερο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή *k* κοντινότερων γειτόνων KNN με *k*=5.

		Ταξινόμηση		
		<i>ZO</i>	<i>NF</i>	<i>S</i>
Βάση Δεδομένων	<i>ZO</i>	192	7	1
	<i>NF</i>	10	189	1
	<i>S</i>	13	9	78

Πίνακας 8.24: Αποτελέσματα για το δεύτερο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή *k* κοντινότερων γειτόνων KNN με *k*=5. Η ακρίβεια ταξινόμησης είναι 91.8%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>ZO</i>	0.96	0.893
<i>NF</i>	0.945	0.922
<i>S</i>	0.78	0.975
Σταθμισμένος μέσος όρος	0.918	0.921

Πίνακας 8.25 : Πίνακας σύγκρισης ταξινόμησης για το δεύτερο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή *k* κοντινότερων γειτόνων KNN με *k*=10.

		Ταξινόμηση		
		<i>ZO</i>	<i>NF</i>	<i>S</i>
Βάση Δεδομένων	<i>ZO</i>	191	5	4
	<i>NF</i>	8	191	1
	<i>S</i>	21	19	60

Πίνακας 8.26: Αποτελέσματα για το δεύτερο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με $k=10$. Η ακρίβεια ταξινόμησης είναι 88.4%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>ZO</i>	0.955	0.868
<i>NF</i>	0.955	0.888
<i>S</i>	0.6	0.923
Σταθμισμένος μέσος όρος	0.884	0.887

Πίνακας 8.27: Πίνακας σύγχυσης ταξινόμησης για το δεύτερο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=2$.

		Ταξινόμηση		
		<i>ZO</i>	<i>NF</i>	<i>S</i>
Βάση Δεδομένων	<i>ZO</i>	184	4	12
	<i>NF</i>	7	186	7
	<i>S</i>	13	8	79

Πίνακας 8.28: Αποτελέσματα για το δεύτερο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=2$. Η ακρίβεια ταξινόμησης είναι 89.8%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>ZO</i>	0.92	0.902
<i>NF</i>	0.93	0.939
<i>S</i>	0.79	0.806
Σταθμισμένος μέσος όρος	0.898	0.898

Πίνακας 8.29 : Πίνακας σύγχυσης ταξινόμησης για το δεύτερο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=5$.

		Ταξινόμηση		
		<i>ZO</i>	<i>NF</i>	<i>S</i>
Βάση Δεδομένων	<i>ZO</i>	178	5	17
	<i>NF</i>	5	186	9
	<i>S</i>	13	7	80

Πίνακας 8.30: Αποτελέσματα για το δεύτερο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=5$. Η ακρίβεια ταξινόμησης είναι 88.8%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>ZO</i>	0.89	0.908
<i>NF</i>	0.93	0.939
<i>S</i>	0.8	0.755
Σταθμισμένος μέσος όρος	0.888	0.89

Πίνακας 8.31: Πίνακας σύγχυσης ταξινόμησης για το δεύτερο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=9$.

		Ταξινόμηση		
		<i>ZO</i>	<i>NF</i>	<i>S</i>
Βάση Δεδομένων	<i>ZO</i>	174	5	21
	<i>NF</i>	4	187	9
	<i>S</i>	13	6	81

Πίνακας 8.32: Αποτελέσματα για το δεύτερο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=9$. Η ακρίβεια ταξινόμησης είναι 88.4%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>ZO</i>	0.87	0.911
<i>NF</i>	0.935	0.944
<i>S</i>	0.81	0.73
Σταθμισμένος μέσος όρος	0.884	0.888

Πίνακας 8.33: Πίνακας σύγχυσης ταξινόμησης για το δεύτερο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή Τυχαίου Δάσους RF.

		Ταξινόμηση		
		<i>ZO</i>	<i>NF</i>	<i>S</i>
Βάση Δεδομένων	<i>ZO</i>	187	6	7
	<i>NF</i>	9	186	5
	<i>S</i>	8	8	84

Πίνακας 8.34: Αποτελέσματα για το δεύτερο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή Τυχαίου Δάσους RF. Η ακρίβεια ταξινόμησης είναι 91.4%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>ZO</i>	0.935	0.917
<i>NF</i>	0.93	0.93
<i>S</i>	0.84	0.875
Σταθμισμένος μέσος όρος	0.914	0.914

Πίνακας 8.35: Πίνακας σύγχυσης ταξινόμησης για το δεύτερο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 5 κρυφούς νευρώνες.

		Ταξινόμηση		
		<i>ZO</i>	<i>NF</i>	<i>S</i>
Βάση Δεδομένων	<i>ZO</i>	189	5	6
	<i>NF</i>	6	187	7
	<i>S</i>	13	13	74

Πίνακας 8.36: Αποτελέσματα για το δεύτερο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 5 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 90%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>ZO</i>	0.945	0.909
<i>NF</i>	0.935	0.912
<i>S</i>	0.74	0.851
Σταθμισμένος μέσος όρος	0.9	0.898

Πίνακας 8.37: Πίνακας σύγκυσης ταξινόμησης για το δεύτερο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 10 κρυφούς νευρώνες.

		Ταξινόμηση		
		<i>ZO</i>	<i>NF</i>	<i>S</i>
Βάση Δεδομένων	<i>ZO</i>	188	6	6
	<i>NF</i>	5	193	2
	<i>S</i>	8	8	84

Πίνακας 8.38: Αποτελέσματα για το δεύτερο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 10 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 93%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>ZO</i>	0.94	0.935
<i>NF</i>	0.965	0.932
<i>S</i>	0.84	0.913
Σταθμισμένος μέσος όρος	0.93	0.93

Πίνακας 8.39: Πίνακας σύγκυσης ταξινόμησης για το δεύτερο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 20 κρυφούς νευρώνες.

		Ταξινόμηση		
		<i>ZO</i>	<i>NF</i>	<i>S</i>
Βάση Δεδομένων	<i>ZO</i>	188	6	6
	<i>NF</i>	4	194	2
	<i>S</i>	7	7	86

Πίνακας 8.40: Αποτελέσματα για το δεύτερο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 20 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 93.6%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
<i>ZO</i>	0.94	0.945
<i>NF</i>	0.97	0.937
<i>S</i>	0.86	0.915
Σταθμισμένος μέσος όρος	0.936	0.936

Πίνακας 8.41: Πίνακας σύγχυσης ταξινόμησης για το τρίτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον Αφελή ταξινομητή Bayes.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	364	36
	S	37	63

Πίνακας 8.42: Αποτελέσματα για το τρίτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον Αφελή ταξινομητή Bayes. Η ακρίβεια ταξινόμησης είναι 85.4%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.91	0.908
S	0.63	0.636
Σταθμισμένος μέσος όρος	0.854	0.853

Πίνακας 8.42: Πίνακας σύγχυσης ταξινόμησης για το τρίτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με k=5.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	395	5
	S	23	77

Πίνακας 8.43: Αποτελέσματα για το τρίτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με k=5. Η ακρίβεια ταξινόμησης είναι 94.4%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.988	0.945
S	0.77	0.939
Σταθμισμένος μέσος όρος	0.944	0.944

Πίνακας 8.45: Πίνακας σύγκρισης ταξινόμησης για το τρίτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με $k=10$.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	395	5
	S	43	57

Πίνακας 8.46: Αποτελέσματα για το τρίτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με $k=10$. Η ακρίβεια ταξινόμησης είναι 90.4%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.988	0.902
S	0.57	0.919
Σταθμισμένος μέσος όρος	0.904	0.905

Πίνακας 8.47: Πίνακας σύγκρισης ταξινόμησης για το τρίτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=2$.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	378	22
	S	17	83

Πίνακας 8.48: Αποτελέσματα για το τρίτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=2$. Η ακρίβεια ταξινόμησης είναι 89.6667%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.945	0.957
S	0.83	0.79
Σταθμισμένος μέσος όρος	0.922	0.924

Πίνακας 8.49: Πίνακας σύγκρισης ταξινόμησης για το τρίτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=5$.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	380	20
	S	16	84

Πίνακας 8.50: Αποτελέσματα για το τρίτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=5$. Η ακρίβεια ταξινόμησης είναι 92.8 %.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.95	0.96
S	0.84	0.808
Σταθμισμένος μέσος όρος	0.928	0.929

Πίνακας 8.51: Πίνακας σύγκρισης ταξινόμησης για το τρίτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=9$.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	379	21
	S	20	80

Πίνακας 8.52: Αποτελέσματα για το τρίτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=9$. Η ακρίβεια ταξινόμησης είναι 91.8%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.948	0.95
S	0.8	0.792
Σταθμισμένος μέσος όρος	0.918	0.918

Πίνακας 8.53: Πίνακας σύγκρισης ταξινόμησης για το τρίτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή Τυχαίου Δάσους RF.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	390	10
	S	20	80

Πίνακας 8.54: Αποτελέσματα για το τρίτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή Τυχαίου Δάσους RF. Η ακρίβεια ταξινόμησης είναι 94 %.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.975	0.951
S	0.8	0.889
Σταθμισμένος μέσος όρος	0.94	0.939

Πίνακας 8.55: Πίνακας σύγκρισης ταξινόμησης για το τρίτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 5 κρυφούς νευρώνες.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	396	4
	S	34	66

Πίνακας 8.56: Αποτελέσματα για το τρίτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 5 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 92.4%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.99	0.921
S	0.66	0.943
Σταθμισμένος μέσος όρος	0.924	0.925

Πίνακας 8.57: Πίνακας σύγκρισης ταξινόμησης για το τρίτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 10 κρυφούς νευρώνες.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	393	7
	S	19	81

Πίνακας 8.58: Αποτελέσματα για το τρίτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 10 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 94.8%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.983	0.954
S	0.81	0.92
Σταθμισμένος μέσος όρος	0.948	0.947

Πίνακας 8.59: Πίνακας σύγκρισης ταξινόμησης για το τρίτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 20 κρυφούς νευρώνες.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	389	11
	S	15	85

Πίνακας 8.60: Αποτελέσματα για το τρίτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 20 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 94.8%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.973	0.963
S	0.85	0.885
Σταθμισμένος μέσος όρος	0.948	0.947

Πίνακας 8.61: Πίνακας σύγχυσης ταξινόμησης για το τέταρτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον Αφελή ταξινομητή Bayes.

		Ταξινόμηση		
		Z	N	S
Βάση Δεδομένων	Z	89	4	7
	N	6	91	3
	S	8	9	83

Πίνακας 8.62: Αποτελέσματα για το τέταρτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον Αφελή ταξινομητή Bayes. Η ακρίβεια ταξινόμησης είναι 87.6667%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
Z	0.89	0.864
N	0.91	0.875
S	0.83	0.892
Σταθμισμένος μέσος όρος	0.877	0.877

Πίνακας 8.63: Πίνακας σύγχυσης ταξινόμησης για το τέταρτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με k=5.

		Ταξινόμηση		
		Z	N	S
Βάση Δεδομένων	Z	95	2	3
	N	5	93	2
	S	5	7	88

Πίνακας 8.64: Αποτελέσματα για το τέταρτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με k=5. Η ακρίβεια ταξινόμησης είναι 92%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
Z	0.95	0.905
N	0.93	0.912
S	0.88	0.946
Σταθμισμένος μέσος όρος	0.92	0.921

Πίνακας 8.65: Πίνακας σύγχυσης ταξινόμησης για το τέταρτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με $k=10$.

		Ταξινόμηση		
		Z	N	S
Βάση Δεδομένων	Z	93	1	6
	N	4	95	1
	S	10	13	77

Πίνακας 8.66: Αποτελέσματα για το τέταρτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με $k=10$. Η ακρίβεια ταξινόμησης είναι 88.3333%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
Z	0.93	0.869
N	0.95	0.872
S	0.77	0.917
Σταθμισμένος μέσος όρος	0.883	0.886

Πίνακας 8.67: Πίνακας σύγχυσης ταξινόμησης για το τέταρτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=2$.

		Ταξινόμηση		
		Z	N	S
Βάση Δεδομένων	Z	89	2	5
	N	5	92	3
	S	9	3	88

Πίνακας 8.68: Αποτελέσματα για το τέταρτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=2$. Η ακρίβεια ταξινόμησης είναι 89.6667%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
Z	0.89	0.864
N	0.92	0.948
S	0.88	0.88
Σταθμισμένος μέσος όρος	0.897	0.898

Πίνακας 8.69: Πίνακας σύγχυσης ταξινόμησης για το τέταρτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=5$.

		Ταξινόμηση		
		Z	N	S
Βάση Δεδομένων	Z	89	2	9
	N	3	95	2
	S	10	3	87

Πίνακας 8.70: Αποτελέσματα για το τέταρτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=5$. Η ακρίβεια ταξινόμησης είναι 90.3333 %.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
Z	0.89	0.873
N	0.95	0.95
S	0.87	0.888
Σταθμισμένος μέσος όρος	0.903	0.903

Πίνακας 8.71: Πίνακας σύγχυσης ταξινόμησης για το τέταρτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=9$.

		Ταξινόμηση		
		Z	N	S
Βάση Δεδομένων	Z	87	4	9
	N	3	92	5
	S	6	7	87

Πίνακας 8.72: Αποτελέσματα για το τέταρτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=9$. Η ακρίβεια ταξινόμησης είναι 88.6667%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
Z	0.87	0.906
N	0.92	0.893
S	0.87	0.861
Σταθμισμένος μέσος όρος	0.887	0.887

Πίνακας 8.73: Πίνακας σύγχυσης ταξινόμησης για το τέταρτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή Τυχαίου Δάσους RF.

		Ταξινόμηση		
		Z	N	S
Βάση Δεδομένων	Z	90	2	8
	N	4	93	3
	S	6	6	88

Πίνακας 8.74: Αποτελέσματα για το τέταρτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή Τυχαίου Δάσους RF. Η ακρίβεια ταξινόμησης είναι 90.3333 %.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
Z	0.9	0.9
N	0.93	0.921
S	0.88	0.889
Σταθμισμένος μέσος όρος	0.903	0.903

Πίνακας 8.75: Πίνακας σύγχυσης ταξινόμησης για το τέταρτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 5 κρυφούς νευρώνες.

		Ταξινόμηση		
		Z	N	S
Βάση Δεδομένων	Z	91	1	8
	N	3	92	5
	S	4	5	91

Πίνακας 8.76: Αποτελέσματα για το τέταρτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 5 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 91.3333%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
Z	0.91	0.929
N	0.92	0.939
S	0.91	0.875
Σταθμισμένος μέσος όρος	0.913	0.914

Πίνακας 8.77: Πίνακας σύγχυσης ταξινόμησης για το τέταρτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 10 κρυφούς νευρώνες.

		Ταξινόμηση		
		Z	N	S
Βάση Δεδομένων	Z	90	6	4
	N	3	94	3
	S	6	5	89

Πίνακας 8.78: Αποτελέσματα για το τέταρτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 10 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 91%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
Z	0.9	0.909
N	0.94	0.895
S	0.89	0.927
Σταθμισμένος μέσος όρος	0.91	0.91

Πίνακας 8.79: Πίνακας σύγχυσης ταξινόμησης για το τέταρτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 20 κρυφούς νευρώνες.

		Ταξινόμηση		
		Z	N	S
Βάση Δεδομένων	Z	92	5	3
	N	3	95	2
	S	6	4	90

Πίνακας 8.80: Αποτελέσματα για το τέταρτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 20 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 92.3333%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
Z	0.92	0.911
N	0.95	0.913
S	0.9	0.947
Σταθμισμένος μέσος όρος	0.923	0.924

Πίνακας 8.81: Πίνακας σύγκρισης ταξινόμησης για το πέμπτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον Αφελή ταξινομητή Bayes.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	91	9
	S	11	89

Πίνακας 8.82: Αποτελέσματα για το πέμπτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον Αφελή ταξινομητή Bayes. Η ακρίβεια ταξινόμησης είναι 90%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.91	0.892
S	0.89	0.908
Σταθμισμένος μέσος όρος	0.9	0.9

Πίνακας 8.83: Πίνακας σύγκρισης ταξινόμησης για το πέμπτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με k=5.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	95	5
	S	8	92

Πίνακας 8.84: Αποτελέσματα για το πέμπτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με k=5. Η ακρίβεια ταξινόμησης είναι 93.5%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.95	0.922
S	0.92	0.948
Σταθμισμένος μέσος όρος	0.935	0.935

Πίνακας 8.85: Πίνακας σύγκυσης ταξινόμησης για το πέμπτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με $k=10$.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	95	5
	S	14	86

Πίνακας 8.86: Αποτελέσματα για το πέμπτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή k κοντινότερων γειτόνων KNN με $k=10$. Η ακρίβεια ταξινόμησης είναι 90.5%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.95	0.872
S	0.86	0.945
Σταθμισμένος μέσος όρος	0.905	0.905

Πίνακας 8.87: Πίνακας σύγκυσης ταξινόμησης για το πέμπτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=2$.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	93	7
	S	14	86

Πίνακας 8.88: Αποτελέσματα για το πέμπτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=2$. Η ακρίβεια ταξινόμησης είναι 89.5%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.93	0.869
S	0.86	0.925
Σταθμισμένος μέσος όρος	0.895	0.895

Πίνακας 8.89: Πίνακας σύγκυσης ταξινόμησης για το πέμπτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=5$.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	90	10
	S	13	84

Πίνακας 8.90: Αποτελέσματα για το πέμπτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=5$. Η ακρίβεια ταξινόμησης είναι 88.5 %.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.9	0.874
S	0.87	0.897
Σταθμισμένος μέσος όρος	0.885	0.885

Πίνακας 8.91: Πίνακας σύγκυσης ταξινόμησης για το πέμπτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=9$.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	91	9
	S	12	88

Πίνακας 8.92: Αποτελέσματα για το πέμπτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή δέντρου απόφασης J48 με $M=9$. Η ακρίβεια ταξινόμησης είναι 89.5%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.91	0.883
S	0.88	0.907
Σταθμισμένος μέσος όρος	0.895	0.895

Πίνακας 8.93: Πίνακας σύγκυσης ταξινόμησης για το πέμπτο πρόβλημα ταξινόμησης, χρησιμοποιώντας τον ταξινομητή Τυχαίου Δάσους RF.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	92	8
	S	13	87

Πίνακας 8.94: Αποτελέσματα για το πέμπτο πρόβλημα ταξινόμησης χρησιμοποιώντας τον ταξινομητή Τυχαίου Δάσους RF. Η ακρίβεια ταξινόμησης είναι 89.5 %.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.92	0.876
S	0.87	0.916
Σταθμισμένος μέσος όρος	0.895	0.896

Πίνακας 8.95: Πίνακας σύγκυσης ταξινόμησης για το πέμπτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 5 κρυφούς νευρώνες.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	93	7
	S	11	89

Πίνακας 8.96: Αποτελέσματα για το πέμπτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 5 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 91%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.93	0.894
S	0.89	0.927
Σταθμισμένος μέσος όρος	0.91	0.911

Πίνακας 8.97: Πίνακας σύγκυσης ταξινόμησης για το πέμπτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 10 κρυφούς νευρώνες.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	94	6
	S	10	90

Πίνακας 8.98: Αποτελέσματα για το πέμπτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 10 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 92%.

Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.94	0.904
S	0.9	0.938
Σταθμισμένος μέσος όρος	0.92	0.921

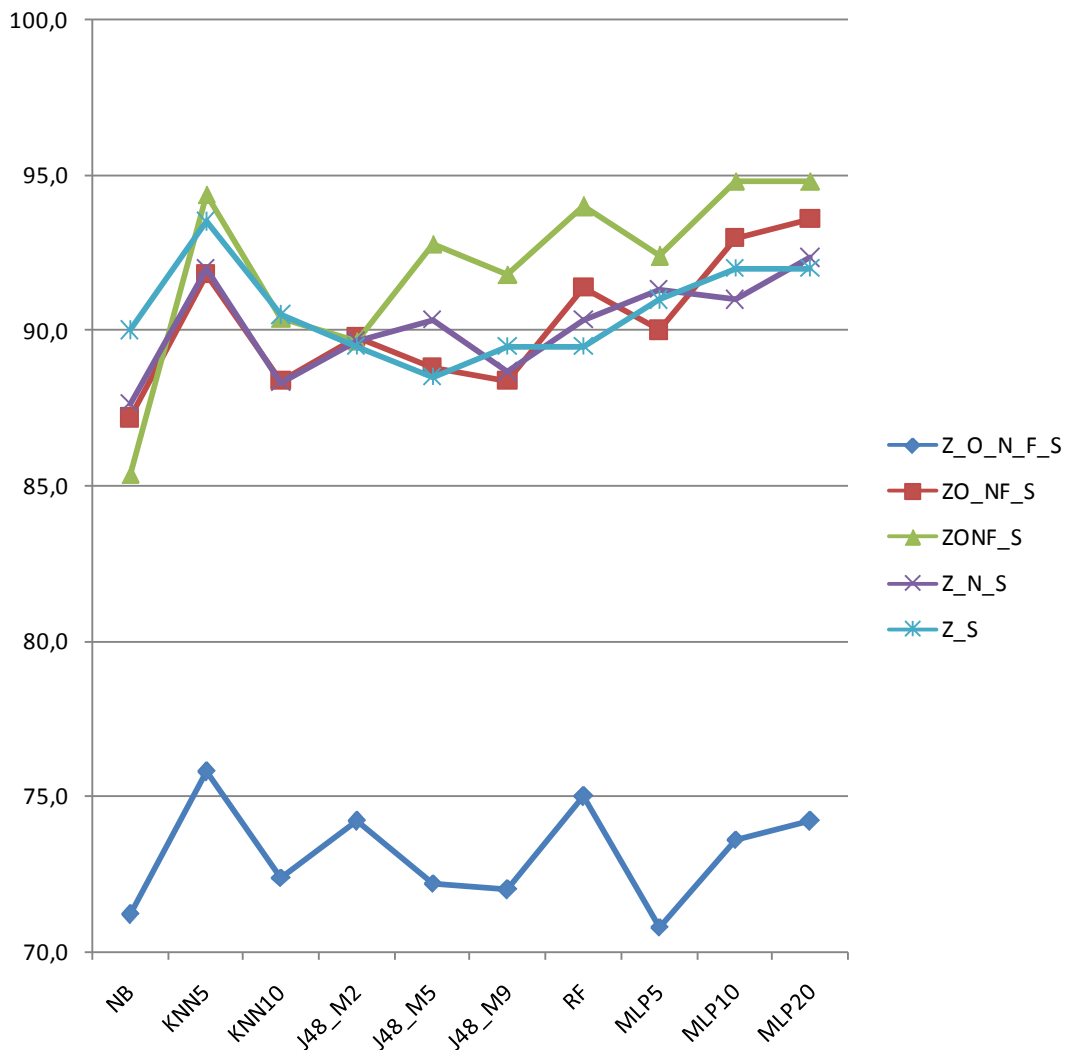
Πίνακας 8.99: Πίνακας σύγκυσης ταξινόμησης για το πέμπτο πρόβλημα ταξινόμησης, χρησιμοποιώντας το Νευρωνικό Δίκτυο με 20 κρυφούς νευρώνες.

		Ταξινόμηση	
		ZONF	S
Βάση Δεδομένων	ZONF	93	7
	S	9	91

Πίνακας 8.100: Αποτελέσματα για το πέμπτο πρόβλημα ταξινόμησης χρησιμοποιώντας το Νευρωνικό Δίκτυο με 20 κρυφούς νευρώνες. Η ακρίβεια ταξινόμησης είναι 92%.

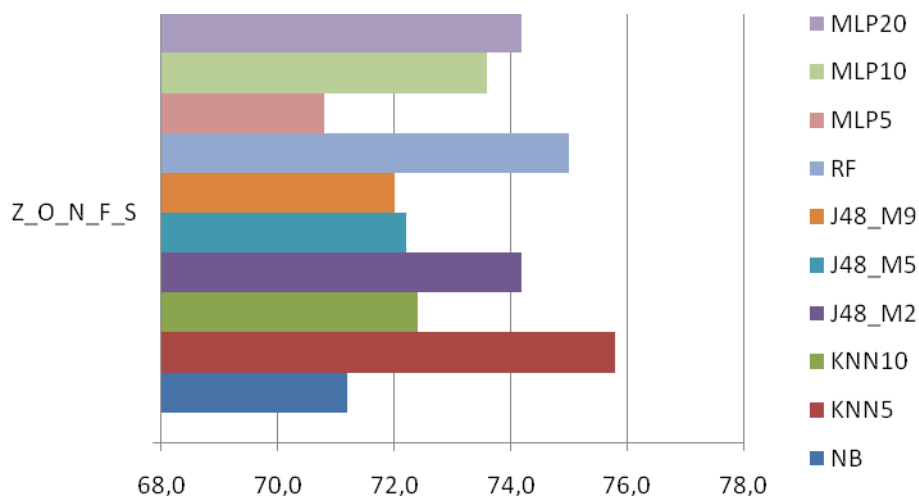
Κατηγορία	Ευαισθησία	Ακρίβεια θετικής πρόβλεψης
ZONF	0.93	0.912
S	0.91	0.929
Σταθμισμένος μέσος όρος	0.92	0.92

Στη συνέχεια, έγινε μια γραφική αναπαράσταση των παραπάνω αποτελεσμάτων έτσι ώστε να μπορούν να εξαχθούν κάποια συμπεράσματα σε σχέση με τη διαδικασία ταξινόμησης. Αρχικά, έγινε μια γραφική αναπαράσταση ως προς την Ακρίβεια Ταξινόμησης κάθε ταξινομητή (με όλες τις διαφορετικές παραμέτρους) σε σχέση με κάθε πρόβλημα, το οποίο παρουσιάζεται στην Εικόνα 8.1.

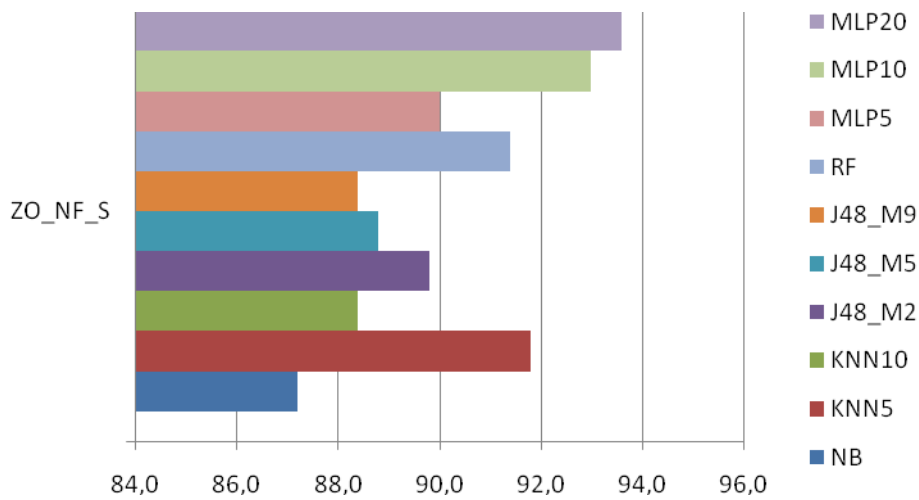


Εικόνα 8.1: Συνολική ακρίβεια ταξινόμησης για τους 5 ταξινομητές (με όλες τις διαφορετικές παραμέτρους) για τα 5 προβλήματα ταξινόμησης.

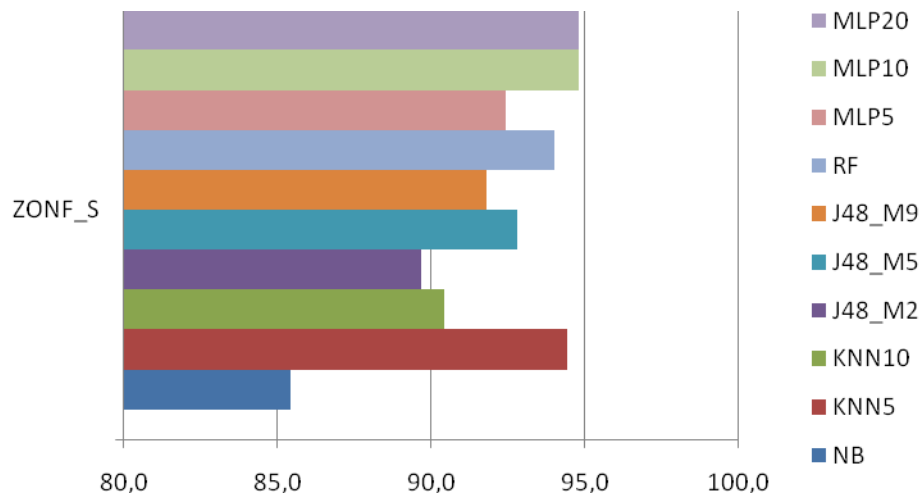
Έπειτα γίνεται η αντίστροφη διαδικασία, δηλαδή η παρουσίαση των αποτελεσμάτων για κάθε ένα από τα 5 προβλήματα ταξινόμησης, σε σχέση με τους ταξινομητές. Τα αποτελέσματα παρουσιάζονται στις εικόνες 8.2.a, 8.2.b, 8.2.c, 8.2.d, 8.2.e (σε κάθε εικόνα παρουσιάζονται τα αποτελέσματα που σχετίζονται με ένα συγκεκριμένο πρόβλημα που με τη σειρά του σχετίζεται με όλους τους ταξινομητές).



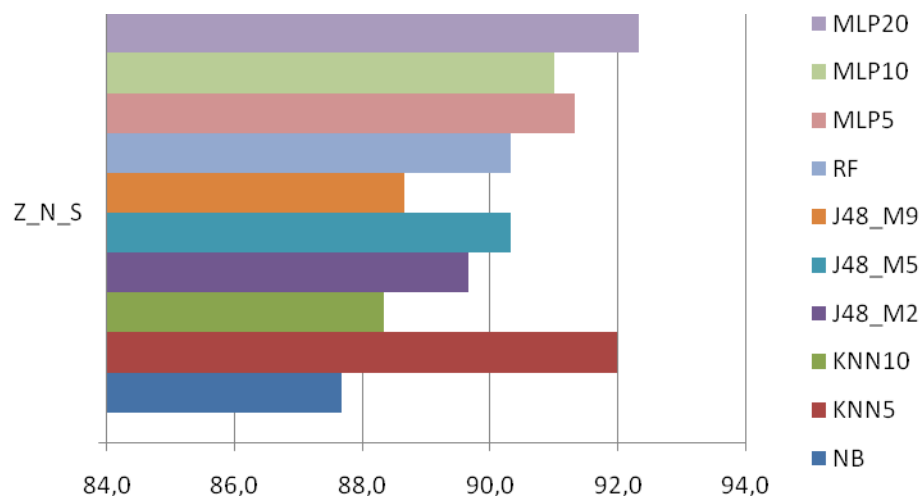
Εικόνα 8.2.a: Αποτελέσματα ακρίβειας ταξινόμησης για το πρόβλημα Z_O_N_F_S από τους 5 ταξινομητές (με όλες τις διαφορετικές παραμέτρους).



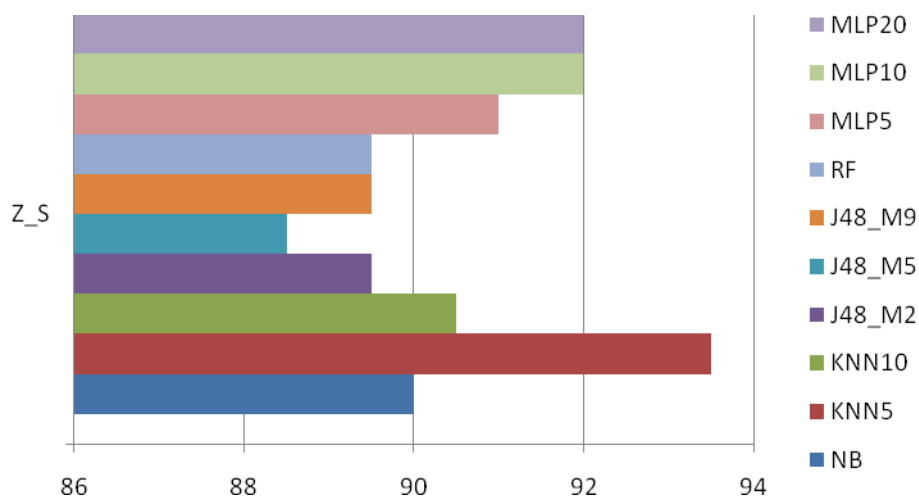
Εικόνα 8.2.b: Αποτελέσματα ακρίβειας ταξινόμησης για το πρόβλημα ZO_NF_S από τους 5 ταξινομητές (με όλες τις διαφορετικές παραμέτρους).



Εικόνα 8.2.c: Αποτελέσματα ακρίβειας ταξινόμησης για το πρόβλημα ZONF_S από τους 5 ταξινομητές (με όλες τις διαφορετικές παραμέτρους).



Εικόνα 8.2.d: Αποτελέσματα ακρίβειας ταξινόμησης για το πρόβλημα Z_N_S από τους 5 ταξινομητές (με όλες τις διαφορετικές παραμέτρους).



Εικόνα 8.2.e: Αποτελέσματα ακρίβειας ταξινόμησης για το πρόβλημα Z_S από τους 5 ταξινομητές (με όλες τις διαφορετικές παραμέτρους).

8.2 Συμπεράσματα

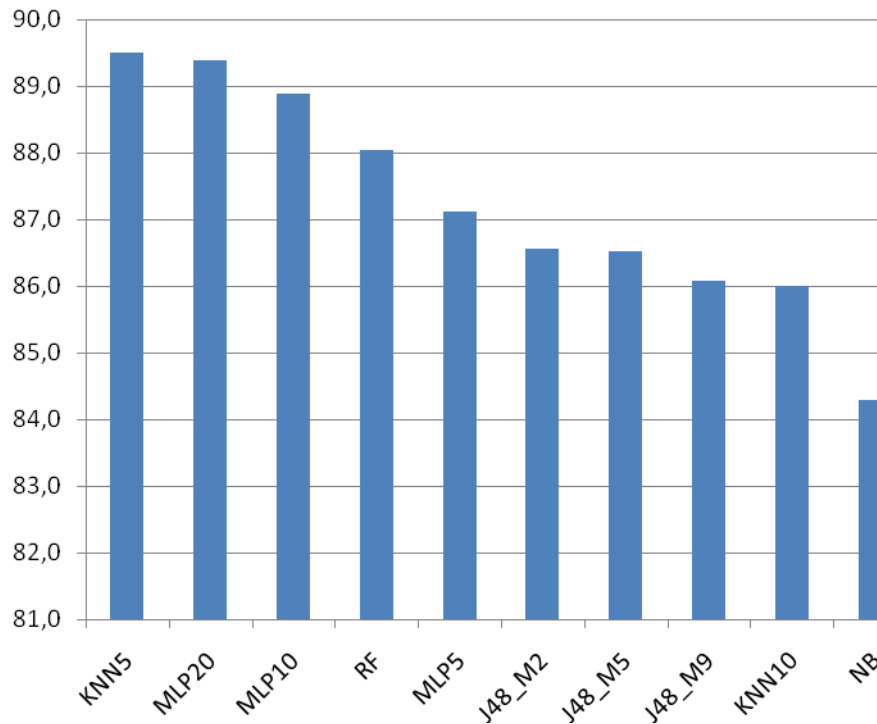
Από τα αποτελέσματα που παρουσιάστηκαν στους παραπάνω πίνακες και στα γραφήματα, ένα γενικό συμπέρασμα το οποίο μπορεί να εξαχθεί είναι ότι η μελέτη του παραπάνω προβλήματος με βάση την προτεινόμενη μεθοδολογία είναι εφικτή και μπορεί να οδηγήσει σε πολύ υψηλά ποσοστά ταξινόμησης, που κυμαίνονται από 75% (για το πρόβλημα των πέντε κατηγοριών) έως και 95% (για τα προβλήματα των δύο και τριών κατηγοριών). Συνεπώς, η μελέτη των σημάτων ΗΕΓ, χρησιμοποιώντας χαρακτηριστικά που αποτυπώνουν την κατανομή ενέργειας σε διαφορετικές συχνοτικές μπάντες, μπορεί να οδηγήσει σε πολύ καλό διαχωρισμό ως προς τις κατηγορίες που υπάρχουν στη συγκεκριμένη Βάση Δεδομένων. Οι συχνοτικές μπάντες που χρησιμοποιούνται στην προτεινόμενη μεθοδολογία είναι αυτές που παρουσιάζονται στη διεθνή ιατρική ορολογία με αποτέλεσμα η συνολική μεθοδολογία ανάλυσης να είναι κατανοητή από τους γιατρούς που σχετίζονται με το συγκεκριμένο πρόβλημα (κυρίως νευρολόγοι).

Τα προβλήματα που μελετήθηκαν αποτυπώνουν διαφορετικές ιατρικές καταστάσεις αλλά και ταυτόχρονα παρουσιάζουν διαφορετική δυσκολία σε σχέση με την ταξινόμηση. Στην εικόνα είναι φανερό ότι τα αποτελέσματα για το πρόβλημα Z_O_N_F_S κυμαίνονται από 70% έως 76%, ανεξάρτητα του ταξινομητή που χρησιμοποιείται, τα οποία είναι πολύ χαμηλότερα όλων των αποτελεσμάτων που προκύπτουν για όλα τα υπόλοιπα προβλήματα (ZO_NF_S, ZONF_S, Z_N_S, Z_S) από οποιονδήποτε ταξινομητή. Αυτό είναι σε εναρμόνιση σε σχέση με τη δυσκολία ταξινόμησης που παρουσιάζει το κάθε πρόβλημα, καθώς το Z_O_N_F_S είναι πρόβλημα πέντε κατηγοριών ενώ όλα τα υπόλοιπα είναι προβλήματα

δύο και τριών κατηγοριών. Επιπλέον, τα προβλήματα δύο κατηγοριών (ZONF_S, Z_S) έχουν ελαφρώς καλύτερα αποτελέσματα από τα προβλήματα τριών κατηγοριών (ZO_NF_S, Z_N_S): ο μέσος όρος Ακρίβειας Ταξινόμησης για το πρόβλημα Z_O_N_F_S, για όλους τους ταξινομητές, είναι 73.1%, για τα προβλήματα ZO_NF_S και Z_N_S (προβλήματα τριών κατηγοριών) είναι 90.2% και για τα δύο προβλήματα, ενώ για τα προβλήματα Z_S και ZONF_S (προβλήματα δύο κατηγοριών) είναι 90.6% και 92% αντίστοιχα, αποτυπώνοντας έτσι την αυξανόμενη δυσκολία ταξινόμησης σε σχέση με τον αριθμό των κατηγοριών.

Παρόλα αυτά τα πολύ χαμηλά αποτελέσματα που παρουσιάζονται στο πρόβλημα των πέντε κατηγοριών (Z_O_N_F_S) μπορούν να εξηγηθούν από τα υψηλά ποσοστά λανθασμένης ταξινόμησης που παρατηρούνται μεταξύ των κατηγοριών N και F. Ο μέσος όρος, ως προς όλους τους ταξινομητές, ποσοστού λανθασμένης ταξινόμησης δεδομένων της κατηγορίας N στην κατηγορία F είναι 20%, ενώ της κατηγορίας F στην κατηγορία N είναι 54%, με αποτέλεσμα η λανθασμένη ταξινόμηση, μόνο μεταξύ αυτών των δύο κατηγοριών, να μειώνει τη συνολική ακρίβεια ταξινόμησης κατά 14.84%. Το συγκεκριμένο αποτέλεσμα μπορεί να γίνει καλύτερα κατανοητό με βάση την ιατρική περιγραφή των κατηγοριών N και F οι οποίες περιείχαν μόνο δραστηριότητα που μετρήθηκε σε διαστήματα 'ελεύθερα από επιληπτικές κρίσεις' των πολύωρων ΗΕΓ καταγραφών 5 επιληπτικών ασθενών, τα οποία καταγράφηκαν στην επιληπτογενή ζώνη, στην περίπτωση του υποσύνολο N και στο σχηματισμό του ιππόκαμπου του αντίθετου ημισφαιρίου, στην περίπτωση του υποσύνολο F. Συνεπώς από τα αποτελέσματα προκύπτει ότι η καταγραφή με διαφορετικά ηλεκτρόδια στις δύο αυτές ζώνες παρουσιάζει πολύ μεγάλες ομοιότητες και τα χαρακτηριστικά ενέργειας, σε συχνοτικές μπάντες που χρησιμοποιούνται στην παρούσα μελέτη, δεν μεταφέρουν αρκετή πληροφορία ώστε να είναι εφικτός ο διαχωρισμός των δύο αυτών κατηγοριών.

Στις εικόνες 8.2.a, 8.2.b, 8.2.c, 8.2.d και 8.2.e παρουσιάζονται τα αποτελέσματα Ακρίβειας Ταξινόμησης από όλους τους ταξινομητές για κάθε πρόβλημα ξεχωριστά. Για το πρόβλημα πέντε κατηγοριών (Z_O_N_F_S) τα καλύτερα αποτελέσματα προκύπτουν από τον KNN5 και από τον RF. Για τα προβλήματα τριών κατηγοριών (ZO_NF_S και Z_N_S) τα καλύτερα αποτελέσματα προκύπτουν από τον MLP20 (και στις δύο περιπτώσεις), ενώ για τα προβλήματα δύο κατηγοριών (ZONF_S και Z_S) τα καλύτερα αποτελέσματα προκύπτουν από τον MLP20 και από τον KNN5. Στην εικόνα 8.3 παρουσιάζεται ο μέσος όρος Ακρίβειας Ταξινόμησης για κάθε ταξινομητή ως προς τα πέντε προβλήματα.



Εικόνα 8.3: Μέσος Όρος Ακρίβειας Ταξινόμησης για κάθε ταξινομητή.

Τα καλύτερα αποτελέσματα επιτυγχάνονται από τον KNN5 και τον MLP20 (Ακρίβεια Ταξινόμησης 89.5% και 89.4% αντίστοιχα) ενώ η χαμηλότερη τιμή παρατηρείται από τον NB. Από την εικόνα 8.3 προκύπτει ότι όλοι οι ταξινομητές παρουσιάζουν υψηλά αποτελέσματα με σχετικά μικρή μεταξύ τους διακύμανση, καθώς η μέση Ακρίβεια Ταξινόμησης κάθε ταξινομητή για όλα τα προβλήματα κυμαίνεται από 84.3% έως 89.5% με εύρος 5.2%, μέση τιμή 87.2% και τυπική απόκλιση 1.68%. Τα παραπάνω συνάδουν με το γενικότερο συμπέρασμα της μελέτης, ότι η μεθοδολογία ανάλυσης που χρησιμοποιήθηκε, καθώς και τα χαρακτηριστικά που αποτυπώνουν την ποσοστιαία κατανομή ενέργειας στις διάφορες συχνοτικές μπάντες, μπορούν να χρησιμοποιηθούν για την ταξινόμηση των κατηγοριών που υπάρχουν στη βάση.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Α. Τ. Τζάλλας, «Αυτόματη διάγνωση ηλεκτροεγκεφαλογραφήματος», Διδακτορική Διατριβή.
- [2] <http://el.wikipedia.org/wiki/Επιληψία>
- [3] Νευρικό Σύστημα, Τόμος 4, Εγκυκλοπαίδεια ΥΓΕΙΑ, Εκδόσεις Δομική
- [4] http://www.ygeiaonline.gr/index.php?option=com_content&view=article&id=1921:2009-09-10-08-08-11&catid=108:neurolo
- [5] http://iatreion.gr/newsdesk_info.php?newsPath=42&newsdesk_id=391&osCsid=21423d65c8de52a04e253b9cc176de61
- [6] <http://www.emedi.gr/κλασική-ιατρική/νευρολογία/item/1036-επιληψία-του-κροταφικού-λοβού.html#.UrMzvelltdg>
- [7] Μ. Ε. Διαμαντάκη, «Καταγραφή και επεξεργασία εγκεφαλογραφήματος που προκύπτει από υποσυνείδητα ερεθίσματα», Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο
- [8] Ε. Ι. Τσιανάκα, «Καταγραφή και επεξεργασία εγκεφαλικών προκλητών δυναμικών σε πειραματικές συνθήκες με υποσυνείδητα ερεθίσματα», Μεταπτυχιακή Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο
- [9] Γεωργούλας, «WEKA», Σημειώσεις Τμήματος Τεχνολογίας Πληροφορικής και Τηλεπικοινωνιών, ΤΕΙ Ηπείρου
- [10] Δ. Παπανικολάου, «Εφαρμογή Τεχνικών Εξόρυξης Γνώσης στην Εκπαίδευση», Μεταπτυχιακή Εργασία, Πανεπιστήμιο Πατρών
- [11] Α. Δήμας, «Τεχνικές για την Εξαγωγή Γνώσης από την Πλατφόρμα του Twitter», Διπλωματική Εργασία, Πανεπιστήμιο Πατρών
- [12] Π.Κολοσούσας, «Ανάπτυξη συστήματος ταξινόμησης ωτορινολαρυγγολογικών παθήσεων με σκοπό την αξιολόγηση και το σχεδιασμό του πλάνου αποκατάστασης», Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο
- [13] Μ. Τσίπουρας, «Κ-Κοντινότεροι γείτονες», Σημειώσεις Τμήματος Τεχνολογίας Πληροφορικής και Τηλεπικοινωνιών, ΤΕΙ Ηπείρου

- [14] Σ. Ολύμπιος, «*Bagging και Boosting μέθοδοι για την κατασκευή μοντέλων με εφαρμογές στα χρηματοοικονομικά*», Πτυχιακή Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο
- [15] Μ. Τσίπουρας, Γ. Τζώρτζης, «*Τεχνητά Νευρωνικά Δίκτυα*», Σημειώσεις Τμήματος Τεχνολογίας Πληροφορικής και Τηλεπικοινωνιών, ΤΕΙ Ηπείρου
- [16] <http://el.wikipedia.org/wiki/MATLAB>
- [17] <http://www.cc.uoc.gr/matlab>
- [18] EEG data from Bonn. [Online] website at http://www.epileptologie-bonn.de/front_content.php?idcat=193&lang=3&changelang=3
- [19] Andrzejak RG, Lehnertz K, Rieke C, Mormann F, David P and Elger CE. Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, Phys Rev E 2001;64:061907(1-8)
- [20] Β. Αναστασόπουλος, «*Ανάλυση και Επεξεργασία Ψηφιακών σημάτων*», Σημειώσεις Πανεπιστημίου Πατρών, 1998.