ΕΛΛΗΝΙΚΟ ΜΕΣΟΓΕΙΑΚΟ ΠΑΝΕΜΙΣΤΗΜΙΟ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

IDENTIFICATION OF PARKINSON´S DISEASE SPEECH CHARACTERISTICS BY UTILIZING DEEP LEARNING TECHNIQUES

ΛΑΠΠΑ ΘΕΟΔΩΡΑ
ΤΗ20026

Επιβλέπων: Καθ. Μανώλης Τσικνάκης

Ηράκλειο, Ιούλιος 2025

ΕΛΛΗΝΙΚΟ ΜΕΣΟΓΕΙΑΚΟ ΠΑΝΕΜΙΣΤΗΜΙΟ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΤΗΣ ΟΜΙΛΙΑΣ ΤΗΣ ΝΟΣΟΥ ΤΟΥ ΠΑΡΚΙΝΣΟΝ ΜΕ ΤΗ ΧΡΗΣΗ ΤΕΧΝΙΚΩΝ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ

ΛΑΠΠΑ ΘΕΟΔΩΡΑ
ΤΗ20026

Επιβλέπων: Καθ. Μανώλης Τσικνάκης

Ηράκλειο, Ιούλιος 2025

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας πτυχιακής εργασίας εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της πτυχιακής εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τη συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Τμήματος Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών Σχολή Μηχανικών του Ελληνικού Μεσογειακού Πανεπιστημίου.

Ευχαριστίες

Ο δρόμος για την ολοκλήρωση αυτής της διπλωματικής εργασίας δεν ήταν πάντα εύκολος, αλλά χάρη στους ανθρώπους που στάθηκαν δίπλα μου καθ' όλη την διάρκεια , έγινε εφικτός.. Όπως συμβαίνει και στη ζωή, έτσι και στον κλάδο της επιστήμης, η προσωπική προσπάθεια δεν αρκεί από μόνη της, χρειάζεται κατανόηση, πίστη και πρωτίστως στήριξη από τους ανθρώπους που έχουμε στο πλευρό μας. Πρώτα απ' όλα, οφείλω ένα τεράστιο ευχαριστώ στην οικογένειά μου, για τη συνεχή αγάπη, υποστήριξη και υπομονή που μου έδειξαν σε κάθε στάδιο της φοιτητικής μου ζωής. Η αγάπη τους υπήρξε το θεμέλιο πάνω στο οποίο μπόρεσα να χτίσω αυτή την προσπάθεια.

Θα ήθελα, επίσης, να ευχαριστήσω τους αγαπημένους μου φίλους, Ελένη, Γιώργο, Μαρία, Κωνσταντίνο, Μαρία, Δήμητρα, Ζαχαρένια, Σταυρούλα, Νάντια και Χριστίνα, που με τον τρόπο τους, ο καθένας και η καθεμιά, μου χάρισαν δύναμη, χαμόγελα και στιγμές ανακούφισης, όταν η πίεση γινόταν αφόρητη.

Εκφράζω την ειλικρινή μου ευγνωμοσύνη στον επιβλέποντα καθηγητή μου, κύριο Μανώλη Τσικνάκη, για την καθοδήγησή του και την εμπιστοσύνη που μου έδειξε καθ' όλη τη διάρκεια της εργασίας. Ιδιαίτερα θα ήθελα να ευχαριστήσω τον υποψήφιο διδάκτορα Βασίλειο Σκαραμαγκά, για τη συνεχή υποστήριξη, την απλόχερη βοήθεια και τη διαρκή παρουσία του από το πρώτο έως και το τελευταίο στάδιο της διπλωματικής. Χωρίς την καθοδήγησή και στήριξη του, η παρούσα εργασία δεν θα είχε λάβει την τελική της μορφή.

Όπως είπε ο Ίρβιν Γιάλομ:
«Η ελπίδα είναι το πιο ανθρώπινο συναίσθημα απ' όλα. Εκείνη σε τραβάει μπροστά, όταν όλα τα υπόλοιπα σε κρατάνε πίσω.». Με αυτή την εργασία κλείνει ένας κύκλος, και ανοίγει ένας καινούριος . Εύχομαι να συνεχίσω να πορεύομαι με τη ίδια θέληση και επιμονή που με οδήγησαν ως εδώ – και με ανθρώπους το ίδιο σπουδαίους, όπως αυτοί που ήταν δίπλα μου σ' αυτό το όμορφο ταξίδι.

Περίληψη

Η Νόσος του Πάρκινσον αποτελεί την δεύτερη πιο διαδεδομένη νευροεκφυλιστική διαταραχή που επηρεάζει την ομιλία των ασθενών στο 90% των περιπτώσεων, με το κύριο σύμπτωμά την υποκινητική δυσαρθρία. Η ανάγκη για έγκαιρη ανίχνευση και παρακολούθηση της νόσου, καθώς και η έλλειψη κατάλληλων διαγνωστικών εργαλείων, επικεντρωμένων στην ελληνική γλώσσα, πυροδοτεί την ανάγκη για την ανάπτυξη καινοτόμων εργαλείων για την διαγνωστική διαδικασία. Η ανάλυση ομιλίας, η οποία είναι μια μη επεμβατική και μη κοστοβόρα διαδικασία μπορεί να εντοπίσει χρήσιμους και αξιοποιήσιμους δείκτες, οι οποίοι θα συμβάλουν στον εντοπισμό της νόσου σε πρώιμο στάδιο, στην συνεχή παρακολούθηση και στην ενίσχυση της παραδοσιακής κλινικής εξέτασης.

Η παρούσα διπλωματική εργασία διερευνά την αναγνώριση χαρακτηριστικών ομιλίας της Νόσου του Πάρκινσον, μέσω τεχνικών βαθιάς μάθηση. Για τον σκοπό αυτό, αναλύθηκαν φωνητικά δεδομένα από 154 Έλληνες ασθενείς και από αυτούς εξήχθην 90 ακουστικά χαρακτηριστικά μέσω του τυποποιημένου συνόλου eGeMAPS.Για την αξιολόγηση χρησιμοποιήθηκαν έξι (6) ταξινομητικοί αλγόριθμοι: Logistic Regression, K-Nearest Neighbors, Random Forest, Support Vector Machine, XGBoost και Spiking Neural Networks, προκειμένου να γίνει η πρόβλεψη των κλινικά σημαντικών παραμέτρων της νόσου που είναι το στάδιο της νόσου μέσω της κλίμακας Hoehn & Yahr, η παρουσία δυσκινησίας και κινητικών διακυμάνσεων, καθώς και η ανταπόκριση στην θεραπεία με Levodopa σε ποσοτικό και ποιοτικό επίπεδο (UPDRS III dif class/UPDRS III dif per). Ακόλουθα, για να γίνει η επικύρωση των μοντέλων, αξιοποιήθηκε η αυστηρή μέθοδος Leave-One-Subject-Out (LOSO).

Τα μοντέλα συνόλου Random Forest και XGBoost παρουσίασαν της πιο ισορροπημένες επιδόσεις στο σύνολο των περιπτώσεων, αν και η διερεύνηση της ανταπόκρισης στην θεραπεία σε ποιοτικό και ποσοτικό επίπεδο, αποτέλεσε πρόκληση για όλα τα εξεταζόμενα μοντέλα. Το κύριο εύρημα της ανάλυσης αποτελεί ο εντοπισμός ενός προφίλ ασθενών, οι οποίοι παρουσιάζουν συστηματική χαμηλή ακρίβεια ταξινόμησης, και βρίσκονται σε ενδιάμεσα στάδια της νόσου, παρουσιάζουν δυσκινησία και η ανταπόκριση τους στην θεραπεία κυμαίνεται σε χαμηλά επίπεδα. Αυτά τα ευρήματα δείχνουν την δυσκολία που αντιμετωπίζουν τα υπολογιστικά μοντέλα να αποδίδουν σταθερά λόγω της ετερογένειας που εμφανίζει η Νόσος του Πάρκινσον.

Η περαιτέρω διερεύνηση των ευρημάτων αυτών μπορεί να συμβάλλει σε μελλοντικό σχεδιασμό πιο εξειδικευμένων, με επεμβατικών εργαλείων, για την υποστήριξη της κλινικής διάγνωσης, την διαρκή καταγραφή και έλεγχο της κατάστασης του ασθενούς, καθώς και στην ανάπτυξη λύσεων που θα είναι προσαρμοσμένες στις εκάστοτε υποομάδες ασθενών, ανάλογα με τις ανάγκες αυτών.

Λέξεις Κλειδιά:

Νόσος του Πάρκινσον, Βαθιά Μάθηση, Ανάλυση Ομιλίας, Φωνητικοί Βιοδείκτες, eGeMAPS, Φωνητικά Χαρακτηριστικά, Κλίμακα Hoehn & Yahr (H&Y), Δυσκινησία, Κινητικές Διακυμάνσεις

Abstract

Parkinson's disease is the second most common neurodegenerative disorder affecting speech in 90% of cases, with hypokinetic dysarthria being the main symptom. The need for early detection and monitoring of the disease, as well as the lack of appropriate diagnostic tools focused on the Greek language, triggers the need for the development of innovative tools for the diagnostic process. Speech analysis, which is a non-invasive and inexpensive procedure, can identify useful and exploitable indicators that will contribute to the early detection of the disease, continuous monitoring, and enhancement of traditional clinical examination.

This thesis investigates the recognition of speech characteristics of Parkinson's disease through deep learning techniques. To this end, voice data from 154 Greek patients were analyzed and 90 acoustic features were extracted from them using the standardized the Extended Geneva Minimalistic Acoustic Parameter Set (eGeMAPS)

Six (6) classification algorithms were used for the evaluation: Logistic Regression, K-Nearest Neighbors, Random Forest, Support Vector Machine, XGBoost, and Spiking Neural Networks, in order to predict the clinically important parameters of the disease, which are the stage of the disease using the Hoehn & Yahr scale, the presence of dyskinesia and motor fluctuations, as well as the response to Levodopa treatment at a quantitative and qualitative level (UPDRS III dif class/UPDRS III dif per). Subsequently, the rigorous Leave-One-Subject-Out (LOSO) method was used to validate the models.

The Random Forest and XGBoost ensemble models showed the most balanced performance across all cases, although investigating the response to treatment in qualitative and quantitative terms was a challenge for all models examined. The main finding of the analysis is the identification of a profile of patients who show consistently low classification accuracy, are in intermediate stages of the disease, have dyskinesia, and have low response to treatment. These findings demonstrate the difficulty computational models face in performing consistently due to the heterogeneity of Parkinson's disease.

Further investigation of these findings may contribute to the future design of more specialized, invasive tools to support clinical diagnosis, continuous recording and monitoring of the patient's condition, as well as the development of solutions tailored to specific patient subgroups, depending on their needs.

Keywords:

Parkinson’s Disease, Deep Learning, Speech Analysis, Vocal Biomarkers, eGeMAPS, Acoustic Features, Hoehn & Yahr Scale (H&Y), Dyskinesia, Motor Fluctuations.

Περιεχόμενα

Ευχαριστίες 2

Περίληψη 3

Abstract 4

Περιεχόμενα 5

Ευρετήριο Εικόνων 10

Ευρετήριο Πινάκων 16

1 Εισαγωγή 20

1.1 Η Νόσος του Πάρκινσον: Μια Γενική Επισκόπηση 20

1.2 Κλινικά Συμπτώματα της Νόσου 21

1.2.1 Κινητικά Συμπτώματα 21

1.2.2 Μη Κινητικά Συμπτώματα 21

1.2.3 Διαταραχές Ομιλίας και Υποκινητική Δυσαρθρία 22

1.3 Από την Κλινική Εκτίμηση στην Ανάλυση Ομιλίας 23

1.4 Η Συμβολή της Μηχανικής και Βαθιάς Μάθησης στην Ανίχνευση της Νόσου 24

1.5 Σκοπός και Συνεισφορά της Παρούσας Μελέτης στην Επιστημονική Έρευνα 24

1.6 Δομή 25

2 Βιβλιογραφική Ανασκόπηση / Θεωρητικό Υπόβαθρο 27

2.1 Γενικές Πληροφορίες 27

2.2 Ο Ρόλος της Ανάλυσης Ομιλίας στην Νόσο του Πάρκινσον 28

2.3 Χρήση Μοντέλων Βαθιάς Μάθησης (Deep Learning) για την Ανάλυση της Ομιλίας 28

2.4 Χαρακτηριστικά Ομιλίας στην Νόσο του Πάρκινσον: Μια Λεπτομερής Ανασκόπηση 29

2.4.1 Ο Δυσαρθρία στην Νόσο του Πάρκινσον (ΝΠ) 29

2.4.2 Άλλα Σημαντικά Χαρακτηριστικά Ομιλίας που αφορούν την ΝΠ 29

2.4.3 Τα Παραδοσιακά Χαρακτηριστικά που Χρησιμοποιούνται για την Ανάλυση Ομιλίας στην ΝΠ 30

2.5 Χρήση του eGeMAPS (extended Geneva Minimalistic Acoustic Parameter Set) στην Ανάλυση Ομιλίας 31

2.5.1 Το Σύνολο Χαρακτηριστικών eGeMAPS 31

2.5.2 Χρησιμότητα του Συνόλου eGeMAPS στην Ανάλυση Ομιλίας για την ΝΠ 32

2.5.3 Το eGeMAPS ως Εργαλείο Ακουστικής Ανάλυσης: Σύγχρονες Εφαρμογές 32
2.5.4 Ανίχνευση ΝΠ Μέσω του Συνόλου eGeMAPS 32
2.5.5 Ταξινόμηση Φαρμακευτικής Κατάστασης στην ΝΠ Μέσω eGeMAPS 33
2.6 Συστηματική Ανασκόπηση Εφαρμογών LM/LMM στην Ποσοτική Ανάλυση Ομιλίας στη ΝΠ 35
2.6.1 Ανασκόπηση Πρόσφατων Μελετών με Χρήση Γραμμικών Μοντέλων (LM/LMM) 35
2.7 Διερεύνηση της Μεθόδου Leave-One-Subject-Out (LOSO) στην Εκπαίδευση/Δοκιμή Μοντέλων Βαθιάς Μάθησης 37
2.8 Ταξινόμηση της ΝΠ Μέσω της Χρήσης Τεχνικών Βαθιάς Μάθησης. 38
2.8.1 Εισαγωγή 38
2.9 Εφαρμογή Μοντέλων Μάθησης για την Ανίχνευση της Νόσου του Πάρκινσον Μέσω της Ομιλίας 39
2.9.1 Ο Αλγόριθμος Spiking Neural Network (SNN) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας 39
2.9.2 Ο Αλγόριθμος Logistic Regression (LR) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας 40
2.9.3 Ο Αλγόριθμος Random Forest (RF) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας 41
2.9.4 Ο Αλγόριθμος Support Vector Machine (SVM) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας 43
2.9.5 Ο Αλγόριθμος K-Nearest Neighbors (K-NN) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας 44
2.9.6 Ο Αλγόριθμος Extreme Gradient Boosting (XGBoost) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας 46
3 Μεθοδολογία 56
3.1 Συλλογή Δεδομένων Ομιλίας και Περιγραφή Συνόλου Δεδομένων. 56
3.2 Διαδικασία Αποθορυβοποίησης (Denoising) με Χρήση του Λογισμικού Audacity 59
3.3 Εξαγωγή Ακουστικών ~~Χ~~αρακτηριστικών 60
3.4 Στατιστική Ανάλυση και Μοντελοποίηση των Ηχητικών Χαρακτηριστικών 62
3.4.1 Κλινικά Ερωτήματα και Παράμετροι Πρόβλεψης 62
3.4.2 Δομή και Οργάνωση του Δείγματος 63
3.4.3 Επιλογή Στατιστικών Μοντέλων για την Ανάλυση των Δεδομένων 63
3.4.4 Οργάνωση και Χωρισμός των Δεδομένων 63
3.4.5 Επιλογή και Προετοιμασία Ανεξάρτητων Χαρακτηριστικών 64
3.4.6 Υλοποίηση των Στατιστικών Μοντέλων 65
3.4.7 Επεξεργασία και Αξιολόγηση Αποτελεσμάτων 67

3.5 Ανάπτυξη και Αξιολόγηση Μοντέλων Ταξινόμησης με χρήση της μεθόδου Leave-One-Subject-Out (LOSO) σε Κλινικά Χαρακτηριστικά Ασθενών με Πάρκινσον. 68
3.5.1 Οργάνωση Δεδομένων και Επιλογή Στατιστικά Σημαντικών Χαρακτηριστικών. 68
3.5.2 Μεταβλητές-Στόχοι, Ταξινόμηση και Προ-επεξεργασία Δεδομένων 69
3.5.3 Εκπαίδευση και Αξιολόγηση Μοντέλων Ταξινόμησης με την Χρήση της Μεθόδου LOSO (Leave-One-Subject-Out). 70
3.5.4 Μοντέλα Μηχανικής Μάθησης για την Ταξινόμηση Κλινικών Δεδομένων 71
3.5.5 Αξιολόγηση της Απόδοσης των Μοντέλων Ταξινόμησης 73
4 Αποτελέσματα 75
4.1 Περιγραφική Ανάλυση του Δείγματος..... 76
4.2 Επεξεργασία και Εξαγωγή Χαρακτηριστικών 80
4.2.1 Περιγραφή του Αρχείου και του Συνόλου Χαρακτηριστικών eGeMAPS 80
4.2.2 Κατηγορίες Χαρακτηριστικών..... 80
4.3 Στατιστική Ανάλυση Σημαντικότητας Χαρακτηριστικών 81
4.3.1 Συσχέτιση Εξαγόμενων Φωνητικών Χαρακτηριστικών με το Κλινικό Στάδιο Hoehn & Yahr (H_Y_binary) 82
4.3.2 Συσχέτιση Εξαγόμενων Φωνητικών Χαρακτηριστικών με την Παρουσία Δυσκινησίας (Dyskinesia)..... 83
4.3.3 Συσχέτιση Εξαγόμενων Φωνητικών Χαρακτηριστικών με την Παρουσία Διακυμάνσεων (Fluctuatios) 84
4.3.4 Συσχέτιση Εξαγόμενων Φωνητικών Χαρακτηριστικών με Κατηγορίες Βαθμολογίας UPDRS-III (UPDRS_III_dif_class)..... 85
4.3.5 Συσχέτιση Εξαγόμενων Φωνητικών Χαρακτηριστικών με Ποσοστιαία Διαφορά Βαθμολογίας UPDRS-III (UPDRS_III_dif_per) 86
4.4 Συγκεντρωτική Ανάλυση της Συχνότητας Εμφάνισης των Χαρακτηριστικών Ομιλίας 87
4.5 Απόδοση Ταξινομητικών Μοντέλων στις Κατηγορίες Εκτίμησης της Νόσου..... 88
4.5.1 Απόδοση Μοντέλων για την Κατηγορία Hoehn & Yahr (H&Y) 89
4.5.2 Απόδοση Μοντέλων για την Κατηγορία Fluctuations 90
4.5.3 Απόδοση Μοντέλων για την Κατηγορία Dyskinesia 91
4.5.4 Απόδοση Μοντέλων για UPDRS_III_dif_class 92
4.5.5 Απόδοση Μοντέλων για UPDRS_III_dif_per 93
4.6 Εντοπισμός Ασθενών με Συστηματικά Χαμηλή Ακρίβεια Ταξινόμησης..... 95
4.6.1 Ασθενείς με Επαναλαμβανόμενα Χαμηλή Ακρίβεια – Βαθμολογική Κλίμακα Hoehn & Yahr (H&Y) 95

4.6.2 Ασθενείς με Επαναλαμβανόμενα Χαμηλή Ακρίβεια – Βαθμολογική Κλίμακα Dyskinesia 96
4.6.3 Ασθενείς με Επαναλαμβανόμενα Χαμηλή Ακρίβεια – Βαθμολογική Κλίμακα Fluctuations 97
4.6.4 Ασθενείς με Επαναλαμβανόμενα Χαμηλή Ακρίβεια – Βαθμολογική Κλίμακα UPDRS_III_dif_class 98
98
4.6.5 Ασθενείς με Επαναλαμβανόμενα Χαμηλή Ακρίβεια – Βαθμολογική Κλίμακα UPDRS_III_dif_per 99
99
4.7 Διερεύνηση Πιθανών Βιοδεικτών Συσχέτισης με την Απόδοση Προγνωστικών Μοντέλων 99
4.7.1 Πιθανοί Βιοδείκτες από την Κλίμακα Hoehn and Yahr (H&Y) 100
4.7.2 Παρουσία Δυσκινησίας ως Σημαντικός Πιθανός Βιοδείκτης 101
4.7.3 Διερεύνηση των Κινητικών Διακυμάνσεων και οι Περιορισμοί του Δείγματος 101
4.7.4 Ανάλυση της Ταξινόμησης Διαφορών στην Κλίμακα UPDRS III (UPDRS_III_dif_class) 102
4.7.5 Ποσοστιαία Μεταβολή της Κλίμακας UPDRS III (UPDRS_III_dif_per) και η Σχέση της με την Απόδοση των Μοντέλων 103
5 Συζήτηση 104
5.1 Ερμηνεία των Σημαντικότερων Φωνητικών Χαρακτηριστικών και η Κλινική τους Σημασία 105
5.2 Ερμηνεία της Απόδοσης των Ταξινομητικών Μοντέλων και η Κλινική τους Σημασία 106
5.3 Ανάλυση Ασθενών με Επαναλαμβανόμενη Χαμηλή Απόδοση Ταξινόμησης: Ετερογένεια και Περιορισμοί των Μοντέλων 108
5.4 Κλινική Σημασία και Προοπτικές στην Ιατρική Κοινότητα 109
5.5 Περιορισμοί της Μελέτης 110
5.6 Προτάσεις για Μελλοντική Ερευνητική Δραστηριότητα 111
5.7 Γενικά Συμπεράσματα 112
Βιβλιογραφία 113
Παράρτημα Α – Αναλυτική Περιγραφή των Λεκτικών Ασκήσεων ανά Σετ 121
Παράρτημα Β – Κατηγοριοποίηση Φωνητικών Χαρακτηριστικών του Συνόλου eGeMAPS 122
1. Φωνητικά Χαρακτηριστικά (38 χαρακτηριστικά) 122
2. Φασματικά Χαρακτηριστικά (24 χαρακτηριστικά) 123

3. Ενεργειακά Χαρακτηριστικά (11 χαρακτηριστικά) 123
4. Ρυθμικά/Χρονικά Χαρακτηριστικά (5 χαρακτηριστικά) 124
5. Μη Ταξινομημένα Χαρακτηριστικά (12 χαρακτηριστικά) 124
Παράρτημα Γ – Αναλυτικά Διαγράμματα Στατιστικής Σημαντικότητας Χαρακτηριστικών 125
Task 1 (A1): Εκφώνηση Φωνηέντων/Συμφώνων – Σταθερότητα Φωνής 125
Task 2 (A2): Εκφώνηση Λέξεων/Φράσεων – Αξιολόγηση Άρθρωσής 135
Task 3 (A3): Συνεχής Ανάγνωση – Συνολική Αξιολόγηση Ομιλίας 149
Task 4 (A4): Φωνητικές Ασκήσεις – Συντονισμός Αρθρώσεων 150
Task 5 (A5): Λεκτικό Ερέθισμα– Εστίαση σε Ομάδες Φωνηέντων 158
Task 6 (A6): Καταμέτρηση Αριθμών – Αξιολόγηση Ρυθμού Ομιλίας 187
Task 7 (A7) : Καταγραφή Αναπνοής – Φωνητική Συμπεριφορά 188
Task 8 (A8): Ηχογράφηση Βήχα – Ανάλυση Εκτός του Φάσματος Ομιλίας 189
Παράρτημα Δ – Αναλυτικά Διαγράμματα Αξιολόγησης Μοντέλων 190
Παράρτημα Δ – Κώδικες Υλοποίησης και Πειραματικών Διαδικασιών 192

§

Ευρετήριο Εικόνων

Εικόνα 1: Σύγκριση Εγκεφαλικής Λειτουργίας: Υγιές Άτομο και Άτομα με ΝΠ πριν και μετά την Εμφάνιση Συμπτωμάτων [4]..... 20
Εικόνα 2: Ενδεικτικά Συμπτώματα της Νόσου του Πάρκινσον ανά Σύστημα Οργάνων. [7]. 22
Εικόνα 3:Σύγκριση Ηχητικών Σημάτων Ασθενούς με ΝΠ και Υγιούς Ατόμου. [8] 23
Εικόνα 4:Οπτική απεικόνιση της επίδρασης της Νόσου του Πάρκινσον στους νευρώνες και τη μεταφορά κινητικών σημάτων [16]. 27
Εικόνα 5: Κατηγοριοποίηση των Ακουστικών Χαρακτηριστικών του eGeMAPS. 31
Εικόνα 6: Βιολογικός νευρώνας ως πρότυπο για τα Spiking Neural Networks (SNN). Τα ακουστικά χαρακτηριστικά της ομιλίας εισάγονται ως inputs, επεξεργάζονται στον πυρήνα και οδηγούν σε ταξινομητικά outputs. [56] 39
Εικόνα 7:Οριοθέτηση απόφασης σε γραμμικά και μη γραμμικά διαχωρίσιμα δεδομένα. Η Logistic Regression αποδίδει καλύτερα σε περιπτώσεις γραμμικής διαχωρισιμότητας (αριστερά), ενώ σε πιο σύνθετα δεδομένα (δεξιά) απαιτούνται πιο εξελιγμένοι ταξινομητές. [60] 40
Εικόνα 8:Αρχή λειτουργίας του Random Forest. Πολλαπλά decision trees εκπαιδεύονται ανεξάρτητα σε υποσύνολα των δεδομένων και η τελική απόφαση λαμβάνεται μέσω πλειοψηφικής ψήφου, ενισχύοντας τη σταθερότητα και την ακρίβεια του μοντέλου [73]..... 42
Εικόνα 9:Οπτική αναπαράσταση του αλγορίθμου Support Vector Machine. Το μοντέλο δημιουργεί ένα υπερεπίπεδο που διαχωρίζει τις κλάσεις με το μέγιστο δυνατό περιθώριο, βασιζόμενο σε σημεία-κλειδιά (support vectors) [77]. 43
Εικόνα 10: Παράδειγμα λειτουργίας του αλγορίθμου k-Nearest Neighbors (k-NN). Το νέο δείγμα (πορτοκαλί σημείο) ταξινομείται στην κατηγορία Α, καθώς οι περισσότεροι από τους k πλησιέστερους γείτονες ανήκουν σε αυτήν [86]. 45
Εικόνα 11: Αρχή Λειτουργίας του Gradient Boosting: Διαδοχικά Μοντέλα με Εστίαση στα Σφάλματα [93]..... 47
Εικόνα 12:Το μικρόφωνο Blue Snowball που χρησιμοποιήθηκε για την καταγραφή των δεδομένων ομιλίας..... 57
Εικόνα 13: Γραμμικά Μεικτά Μοντέλα (LMM) και Tasks. 65
Εικόνα 14:Απλά Γραμμικά Μοντέλα (LM) και Tasks. 66
Εικόνα 15:Διάγραμμα Ροής της Μεθόδου (LOSO) Leave-One-Subject-Out. 71
Εικόνα 16:Διαδικασία Ανάλυσης Ομιλίας της Νόσου του Πάρκινσον με Τεχνικές Μηχανικής Μάθησης. 75
Εικόνα 17:Κατανομή του Φύλου ανά Στάδιο Hoehn & Yahr. 77
Εικόνα 18:Κατανομή του Φύλου με Βάση την Παρουσία Κινητικών Διακυμάνσεων (Fluctuations). 78
Εικόνα 19:Κατανομή του Φύλου με Βάση την Παρουσία Δυσκινησίας (Dyskinesia). 78
Εικόνα 20:Κατανομή του Φύλου ως προς την Ποιότητα Ανταπόκρισης στη Θεραπεία με Levodopa..... 79
Εικόνα 21:Κατανομή του Φύλου ως προς την Ποσοστιαία Ανταπόκριση στη Θεραπεία με Levodopa..... 79
Εικόνα 22: Κατανομή Χαρακτηριστικών του eGeMAPS ανά Κατηγορία..... 81
Εικόνα 23:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση mom2.wav του Task 2..... 83

Εικόνα 24:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_ring.wav του Task 5 84
Εικόνα 25:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_spaceship.wav του Task 5 85
Εικόνα 26:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση school.wav του Task 2 86
Εικόνα 27:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση z2.wav του Task 1 87
Εικόνα 28: Τα 20 Χαρακτηριστικά Ομιλίας με τη Συχνότερη Εμφάνιση στα Σημαντικά Ευρήματα. 88
Εικόνα 29:Radar Plot – Συγκριτική Απόδοση Μοντέλων για την Κατηγορία Hoehn & Yahr (H&Y) 90
Εικόνα 30: Radar Plot – Συγκριτική Απόδοση Μοντέλων για την Κατηγορία Κινητικών Διακυμάνσεων (Fluctuations) 91
Εικόνα 31: Radar Plot – Συγκριτική Απόδοση Μοντέλων για την Κατηγορία Δυσκινησίας (Dyskinesia) 92
Εικόνα 32: Radar Plot – Συγκριτική Απόδοση Μοντέλων για τη Διακριτή Μεταβολή του UPDRS III (UPDRS_III_dif_class) 93
Εικόνα 33: Radar Plot – Συγκριτική Απόδοση Μοντέλων για την Ποσοστιαία Μεταβολή του UPDRS III (UPDRS_III_dif_per) 94
Εικόνα 34: Κατανομή Ασθενών με Χαμηλή Απόδοση Μοντέλου στην Κατηγορία H&Y. 95
Εικόνα 35:Κατανομή Ασθενών με Χαμηλή Απόδοση Μοντέλου στην Κατηγορία Dyskinesia. 96
Εικόνα 36:Κατανομή Ασθενών με Χαμηλή Απόδοση Μοντέλου στην Κατηγορία Fluctuations. 97
Εικόνα 37: Κατανομή Ασθενών με Χαμηλή Απόδοση Μοντέλου στην Κατηγορία UPDRS_III_dif_class. 98
Εικόνα 38: Κατανομή Ασθενών με Χαμηλή Απόδοση Μοντέλου στην Κατηγορία UPDRS_III_dif_per. 99
Εικόνα 39: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 1 (Κατηγορία H&Y). 100
Εικόνα 40: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 2 (Κατηγορία H&Y). 100
Εικόνα 41: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 1 (Κατηγορία Dyskinesia) 101
Εικόνα 42:Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 2 (Κατηγορία Dyskinesia) 101
Εικόνα 43: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 1 (Κατηγορία Fluctuations) 102
Εικόνα 44: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 2 (Κατηγορία Fluctuations) 102
Εικόνα 45: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 1 (Κατηγορία UPDRS_III_dif_class). 103
Εικόνα 46: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 2 (Κατηγορία UPDRS_III_dif_class). 103

Εικόνα 47: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 1 (Κατηγορία UPDRS_III_dif_per). 104
Εικόνα 48: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 2 (Κατηγορία UPDRS_III_dif_per). 104
Εικόνα 49: Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για το σύνολο του Task 1 126
Εικόνα 50:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση a1.wav του Task 1 127
Εικόνα 51:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση a2.wav του Task 1 128
Εικόνα 52:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την..... 129
Εικόνα 53:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση s1.wav του Task 1 130
Εικόνα 54:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση s2.wav του Task 1 131
Εικόνα 55:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση s3.wav του Task 1 132
Εικόνα 56:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση z1.wav του Task 1 133
Εικόνα 57:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση z2.wav του Task 1 134
Εικόνα 58:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση z3.wav του Task 1 135
Εικόνα 59:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για το σύνολο του Task 136
Εικόνα 60:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση ball1.wav του Task 2 137
Εικόνα 61:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση ball2.wav του Task 2 138
Εικόνα 62:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση earth.wav του Task 2 139
Εικόνα 63:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση illness.wav του Task 2 140
Εικόνα 64:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση lentils.wav του Task 2 141
Εικόνα 65:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση mom1.wav του Task 2 142
Εικόνα 66:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση mom2.wav του Task 2 143
Εικόνα 67:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση sakis.wav του Task 2 144
Εικόνα 68:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση say.wav του Task 2 145
Εικόνα 69:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση school.wav του Task 2 146

Εικόνα 70:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση wind.wav του Task 2 147
Εικόνα 71:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση yesno.wav του Task 2 148
Εικόνα 72:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση rainbow_passage.wav του Task 3 149
Εικόνα 73:σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για το σύνολο του Task 4 150
Εικόνα 74:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση aui.wav του Task 4 151
Εικόνα 75:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση ka.wav του Task 4 152
Εικόνα 76:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση kala.wav του Task 4 153
Εικόνα 77:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση oe.wav του Task 4 154
Εικόνα 78:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση pa.wav του Task 4 155
Εικόνα 79:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση pataka.wav του Task 4 156
Εικόνα 80:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση ta.wav του Task 4 157
Εικόνα 81:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για το σύνολο του Task 5 158
Εικόνα 82:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_autoind.wav του Task 5 159
Εικόνα 83:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_bat.wav του Task 5 160
Εικόνα 84:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_bicycle.wav του Task 5 161
Εικόνα 85:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_cinema.wav του Task 5 162
Εικόνα 86:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_cockroachkiller.wav του Task 5 163
Εικόνα 87:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_mailman.wav του Task 5 164
Εικόνα 88:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_pastrychef.wav του Task 5 165
Εικόνα 89:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_radio.wav του Task 5 166
Εικόνα 90:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_ring.wav του Task 5 167
Εικόνα 91:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_rose.wav του Task 5 168
Εικόνα 92:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_safe.wav του Task 5 169

Εικόνα 93:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_snail.wav του Task 5 170
Εικόνα 94:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_spaceship.wav του Task 5 171
Εικόνα 95:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_telecom.wav του Task 5 172
Εικόνα 96:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_autoind.wav του Task 5..... 173
Εικόνα 97:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_bat.wav του Task 5 174
Εικόνα 98:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_bicycle.wav του Task 5 175
Εικόνα 99:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_cinema.wav του Task 5 176
Εικόνα 100:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_cockroachkiller.wav του Task 5 177
Εικόνα 101:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_mailman.wav του Task 5..... 178
Εικόνα 102:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_pastrychef.wav του Task 5 179
Εικόνα 103:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_radio.wav του Task 5 180
Εικόνα 104:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_ring.wav του Task 5..... 181
Εικόνα 105:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_rose.wav του Task 5..... 182
Εικόνα 106:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_safe.wav του Task 5..... 183
Εικόνα 107:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_snail.wav του Task 5..... 184
Εικόνα 108:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_spaceship.wav του Task 5 185
Εικόνα 109:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_telecom.wav του Task 5..... 186
Εικόνα 110: Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση counting.wav του Task..... 187
Εικόνα 111:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση breathing.wav του Task..... 188
Εικόνα 112:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση coughing.wav του Task..... 189
Εικόνα 113:Σύγκριση μετρικών απόδοσης των μοντέλων για την πρόβλεψη της κλίμακας Hoehn & Yahr. 190
Εικόνα 114: Σύγκριση μετρικών απόδοσης των μοντέλων για την πρόβλεψη της Δυσκινησίας (Dyskinesia). 190
Εικόνα 115: Σύγκριση μετρικών απόδοσης των μοντέλων για την πρόβλεψη των Κινητικών Διακυμάνσεων (Fluctuations). 191

Εικόνα 116: Σύγκριση μετρικών απόδοσης των μοντέλων για την πρόβλεψη της κατηγορίας διαφοροποίησης στη κλίμακα UPDRS Part III (UPDRS_III_dif_class)..... 191
Εικόνα 117: Σύγκριση μετρικών απόδοσης των μοντέλων για την πρόβλεψη της ποσοστιαίας διαφοροποίησης στην κλίμακα UPDRS Part III (UPDRS_III_dif_per). 192

Ευρετήριο Πινάκων

Πίνακας 1:Συνοπτική Απεικόνιση των Διαταραχών Ομιλίας σε Ασθενείς με ΝΠ. 30
Πίνακας 2: Συγκριτικός Πίνακας Μελετών με Χρήση του eGeMAPS: Δεδομένα, Χαρακτηριστικά και Ευρήματα. 34
Πίνακας 3: Μελέτες που Χρησιμοποιούν Στατιστικά Μοντέλα για την Ανάλυση Φωνητικών Χαρακτηριστικών σε Ασθενείς με Νόσο του Πάρκινσον. 36
Πίνακας 4:Ενδεικτικές μελέτες που αξιοποίησαν επικύρωση τύπου LOSO για την ανάλυση φωνητικών χαρακτηριστικών στη Νόσο του Πάρκινσον. 38
Πίνακας 5: Συγκεντρωτικός πίνακας εφαρμογής των SNN, Logistic Regression, Random Forest, SVM,KNN και XGBoost στην ανάλυση ομιλίας για διάγνωση Νόσου του Πάρκινσον. 48
Πίνακας 6:Ασκήσεις Ομιλίας για την Ανάλυση Φωνητικών Χαρακτηριστικών στη Νόσο του Πάρκινσον. 58
Πίνακας 7:Φωνητικά Χαρακτηριστικά και η Συσχέτισή τους με την Ομιλία Ατόμων με Νόσο του Πάρκινσον. 60
Πίνακας 8: Εξεταζόμενα Κλινικά Ερωτήματα Σχετικά με τη Νόσο του Πάρκινσον. 62
Πίνακας 9:Κατηγοριοποίηση Ασκήσεων Ομιλίας για την Ανάλυση Φωνητικών Χαρακτηριστικών. 63
Πίνακας 10:Παρουσίαση των Εκπαιδευθέντων Μοντέλων Ταξινόμησης. 71
Πίνακας 11: Κατανομή του Δείγματος ανά Κλινική Μεταβλητή και Φύλο. 76
Πίνακας 12:Κατανομή των 90 φωνητικών χαρακτηριστικών του eGeMAPS ανά λειτουργική κατηγορία. 80
Πίνακας 13:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για την άσκηση mom2.wav του Task 2 (A2) 82
Πίνακας 14:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για την άσκηση s_ring.wav του Task 5 (A5) 83
Πίνακας 15:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για την άσκηση s_spaceship.wav του Task 5 (A5) 84
Πίνακας 16:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για την άσκηση school.wav του Task 2 (A2) 85
Πίνακας 17:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για την άσκηση z2.wav του Task 1 (A1) 86
Πίνακας 18: Συγκριτική Παρουσίαση των Μετρικών Απόδοσης των Μοντέλων για την Κατηγορία Hoehn & Yahr. 89
Πίνακας 19: Συγκριτική Παρουσίαση των Μετρικών Απόδοσης των Μοντέλων για την Κατηγορία Fluctuations. 90
Πίνακας 20: Συγκριτική Παρουσίαση των Μετρικών Απόδοσης των Μοντέλων για την Κατηγορία Dyskinesia. 92
Πίνακας 21: Συγκριτική Παρουσίαση των Μετρικών Απόδοσης των Μοντέλων για την Κατηγορία UPDRS_III_dif_class. 93
Πίνακας 22: Συγκριτική Παρουσίαση των Μετρικών Απόδοσης των Μοντέλων για την Κατηγορία UPDRS_III_dif_per. 94
Πίνακας 23: Αναλυτική Παρουσίαση των Σετ Ασκήσεων και των Περιλαμβανόμενων Ηχητικών Αρχείων. 121

Πίνακας 24:Συγκεντρωτικός Πίνακας Φωνητικών Χαρακτηριστικών. 122
Πίνακας 25: Συγκεντρωτικός Πίνακας Φασματικών Χαρακτηριστικών. 123
Πίνακας 26:Συγκεντρωτικός Πίνακας Ενεργειακών Χαρακτηριστικών. 123
Πίνακας 27: Συγκεντρωτικός Πίνακας Ρυθμικών/Χρονικών Χαρακτηριστικών. 124
Πίνακας 28:Συγκεντρωτικός Πίνακας Μη Ταξινομημένων Χαρακτηριστικών. 124
Πίνακας 29:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για το σύνολο του Task 1 (A1)..... 126
Πίνακας 30:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση a1.wav του Task 1 (A1) 127
Πίνακας 31: Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση a2.wav του Task 1 (A1) 128
Πίνακας 32:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση a3.wav του Task 1 (A1) 129
Πίνακας 33:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s1.wav του Task 1 (A1) 130
Πίνακας 34:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s2.wav του Task 1 (A1) 131
Πίνακας 35:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s3.wav του Task 1 (A1) 132
Πίνακας 36:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση z1.wav του Task 1 (A1) 133
Πίνακας 37:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση z2.wav του Task 1 (A1) 134
Πίνακας 38:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση z3.wav του Task 1 (A1) 135
Πίνακας 39:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για το σύνολο του Task 2 (A2)..... 136
Πίνακας 40:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση ball1.wav του Task 2 (A2)..... 137
Πίνακας 41:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση ball2.wav του Task 2 (A2)..... 138
Πίνακας 42:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση earth.wav του Task 2 (A2)..... 139
Πίνακας 43:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση illness.wav του Task 2 (A2)..... 140
Πίνακας 44:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση lentils.wav του Task 2 (A2)..... 141
Πίνακας 45:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση mom1.wav του Task 2 (A2)..... 142
Πίνακας 46:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση mom2.wav του Task 2 (A2)..... 143
Πίνακας 47:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση sakis.wav του Task 2 (A2)..... 144
Πίνακας 48:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση say.wav του Task 2 (A2) 145

Πίνακας 49:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση school.wav του Task 2 (A2) 146
Πίνακας 50:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση wind.wav του Task 2 (A2) 147
Πίνακας 51:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση yesno.wav του Task 2 (A2) 148
Πίνακας 52:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για το σύνολο του Task 3 (A3) 149
Πίνακας 53:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για το σύνολο του Task 4 (A4) 150
Πίνακας 54:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση aui.wav του Task 4 (A4) 151
Πίνακας 55:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση ka.wav του Task 4 (A4) 152
Πίνακας 56:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση kala.wav του Task 4 (A4) 153
Πίνακας 57:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση oe.wav του Task 4 (A4) 154
Πίνακας 58:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση pa.wav του Task 4 (A4) 155
Πίνακας 59:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση pataka.wav του Task 4 (A4) 156
Πίνακας 60:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση ta.wav του Task 4 (A4) 157
Πίνακας 61:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για το σύνολο του Task 5 (A5) 158
Πίνακας 62:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_autoind.wav του Task 5 (A5) 159
Πίνακας 63:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_bat.wav του Task 5 (A5) 160
Πίνακας 64:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_bicycle.wav του Task 5 (A5) 161
Πίνακας 65:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_cinema.wav του Task 5 (A5) 162
Πίνακας 66:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_cockroachkiller.wav του Task 5 (A5) 163
Πίνακας 67:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_mailman.wav του Task 5 (A5) 164
Πίνακας 68:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_pastrychef.wav του Task 5 (A5) 165
Πίνακας 69:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_radio.wav του Task 5 (A5) 166
Πίνακας 70:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_ring.wav του Task 5 (A5) 167
Πίνακας 71:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_rose.wav του Task 5 (A5) 168

Πίνακας 72:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_safe.wav του Task 5 (A5) 169
Πίνακας 73:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_snail.wav του Task 5 (A5) 170
Πίνακας 74:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_spaceship.wav του Task 5 (A5) 171
Πίνακας 75:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_telecom.wav του Task 5 (A5) 172
Πίνακας 76:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_autoind.wav του Task 5 (A5) 173
Πίνακας 77:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_bat.wav του Task 5 (A5) 174
Πίνακας 78:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_bicycle.wav του Task 5 (A5) 175
Πίνακας 79:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_cinema.wav του Task 5 (A5) 176
Πίνακας 80:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_cockroachkiller.wav του Task 5 (A5) 177
Πίνακας 81:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_mailman.wav του Task 5 (A5) 178
Πίνακας 82:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_pastrychef.wav του Task 5 (A5) 179
Πίνακας 83:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_radio.wav του Task 5 (A5) 180
Πίνακας 84:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_ring.wav του Task 5 (A5) 181
Πίνακας 85:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_rose.wav του Task 5 (A5) 182
Πίνακας 86:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_safe.wav του Task 5 (A5) 183
Πίνακας 87:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_snail.wav του Task 5 (A5) 184
Πίνακας 88:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_spaceship.wav του Task 5 (A5) 185
Πίνακας 89:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_telecom.wav του Task 5 (A5) 186
Πίνακας 90:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για το σύνολο του Task 6 (A6) 187
Πίνακας 91:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για το σύνολο του Task 7 (A7) 188
Πίνακας 92:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για το σύνολο του Task 8 (A8) 189

1 Εισαγωγή

1.1 Η Νόσος του Πάρκινσον: Μια Γενική Επισκόπηση

Η Νόσος του Πάρκινσον (ΝΠ) είναι μια νευροεκφυλιστική διαταραχή που εντοπίζεται στο κεντρικό νευρικό σύστημα του ατόμου [1]. Πρακτικά, η νόσος αυτή χαρακτηρίζεται από την προοδευτική απώλεια των ντοπαμινεργικών νευρώνων στην μέλαινα ουσία του ανθρωπίνου εγκεφάλου. Αυτό έχει ως αποτέλεσμα την σταδιακή μείωση της ντοπαμίνης, που αποτελεί έναν νευροδιαβιβαστή που είναι σημαντικός για τον έλεγχο της κίνησης [1]. Η έλλειψη ντοπαμίνης προκαλεί διαταραχές στη λειτουργία του εγκεφάλου, οι οποίες με τη σειρά τους οδηγούν στην εμφάνιση των χαρακτηριστικών κινητικών συμπτωμάτων της νόσου [2]. Σε επιδημιολογικό επίπεδο η ΝΠ, μετά το Alzheimer, είναι η δεύτερη πιο συχνή νευροεκφυλιστική νόσος που επηρεάζει περίπου το 1-2% του παγκόσμιου πληθυσμού άνω των 65 ετών και φέρεται ότι μέχρι το 2040 ο αριθμός των ατόμων που νοσούν με την νόσο θα έχει διπλασιαστεί, επιβαρύνοντας το παγκόσμιο σύστημα υγείας [2].Ο λόγος εμφάνισης της ΝΠ στην πλειοψηφία των περιπτώσεων παραμένει άγνωστη, ωστόσο η επιστημονική κοινότητα έχει καταλήξει στο συμπέρασμα ότι η συγκεκριμένη νόσος προκαλείται από έναν συνδυασμό κληρονομικών και περιβαλλοντικών παραγόντων [3]. Παρόλο που γενετικές μεταλλάξεις, κυριώς των γονιδίων LRRK2 (Leucine-Rich Repeat Kinase 2) και SNCA (Synuclein Alpha) συνδέονται με αυξημένο κίνδυνο εμφάνισης της νόσου, έχει αποδειχθεί ότι ευθύνονται για ένα μικρό ποσοστό των περιπτώσεων. Σε πολλές περιπτώσεις, χαρακτηριστικό κλινικό εύρημα αποτελεί η παρουσία των σωματίων Lewy, δηλαδή συσσωματώσεων της πρωτεΐνης α-συνουκλεΐνης εντός των νευρικών κυττάρων του εγκεφάλου και πιο συγκεκριμένα στην μέλαινα ουσία, γεγονός που οδηγεί στην σταδιακή εκφύλιση και τελικά θάνατο τους, μια διαδικασία που αποτελεί τον βασικό παθογενετικό μηχανισμό της ΝΠ [3].



Εικόνα 1: Σύγκριση Εγκεφαλικής Λειτουργίας: Υγιές Άτομο και Άτομα με ΝΠ πριν και μετά την Εμφάνιση Συμπτωμάτων [4].

Ο συνδυασμός όλων αυτών των παραγόντων – η σταδιακή επιδείνωση της νόσου, η έλλειψη μιας ολοκληρωμένης θεραπείας για την εξάλειψή αυτής και η αναμενόμενη αύξηση των περιστατικών - ενισχύει την ανάγκη για την ανάπτυξη νέων και καινοτόμων μεθόδων που θα συμβάλουν στην έγκαιρη και αξιόπιστη ανίχνευση της νόσου, καθώς και την διαρκή, κατ' οίκον παρακολούθηση (24/7) της κατάστασης των ασθενών, με σκοπό την βελτίωση της ποιότητας ζωής των ασθενών και την μείωση του ρυθμού επιδείνωσής, διότι δεν υπάρχει θεραπεία για την ΝΠ [4]. Η ανίχνευση και παρακολούθηση της νόσου του Πάρκινσον μέσω

της ανάλυσης της φωνής θεωρείται μία καινοτόμα, τεχνολογικά σύγχρονη μέθοδος και ιδιαίτερα υποσχόμενη προσέγγιση στον χώρο της νευροεκφυλιστικής διάγνωσης και παρακολούθησης. Δεν έχει ακόμα αντικαταστήσει πλήρως τις παραδοσιακές μεθόδους, αλλά συμπληρώνει αποτελεσματικά τις υπάρχουσες πρακτικές και βρίσκεται στο επίκεντρο της σύγχρονης έρευνας [5].

1.2 Κλινικά Συμπτώματα της Νόσου

1.2.1 Κινητικά Συμπτώματα

Η ΝΠ χαρακτηρίζεται από τέσσερα κύρια κινητικά συμπτώματα, τα οποία αποτελούν την βάση για την κλινική διάγνωση [6]. Αυτά τα συμπτώματα είναι:

- Ο τρόμος ηρεμίας, που αποτελεί ένα από τα πρώιμα συμπτώματα της νόσου και στο αρχικό στάδιο συχνά εμφανίζεται με την χαρακτηριστική κίνηση "pill-rolling" του χεριού [1].
- Η βραδυκινησία, που είναι πρακτικά η επιβράδυνση της εκούσιας κίνησης των άκρων, επηρεάζοντας της καθημερινότητα του ατόμου, καθώς και την εκφραστικότητα του προσώπου [1].
- Η μυϊκή δυσκαμψία, που αφορά την δυσκολία στην κίνηση των αρθρώσεων προκαλώντας πόνο και περιορισμούς στην κίνηση του [6].
- Η Αστάθεια/απώλεια ισορροπίας, που επηρεάζει την διατήρηση της ισορροπίας και των αντανακλαστικών στήριξης και εμφανίζεται σε πιο προχωρημένα στάδια της νόσου, αυξάνοντας τον κίνδυνο πτώσεων [6].

1.2.2 Μη Κινητικά Συμπτώματα

Πέρα από τα κινητικά συμπτώματα, η ΝΠ χαρακτηρίζεται από μια πληθώρα συμπτωμάτων που δεν αποτελούν κινητικά, ορισμένα από τα οποία μπορούν να εμφανιστούν ακόμα και χρόνια πριν την κλινική διάγνωση της νόσου και την εμφάνιση των πρώιμων συμπτωμάτων. Και αυτά τα συμπτώματα έχουν αντίκτυπο στην ποιότητα ζωής του ατόμου, δυσχεραίνοντας την καθημερινότητα του [1]. Αναμεσά σε αυτά τα συμπτώματα, περιλαμβάνονται:

- Οι διαταραχές ομιλίας.
- Οι διαταραχές ύπνου.
- Οι ψυχιατρικές διαταραχές (κατάθλιψη, άγχος).
- Οι αισθητηριακές διαταραχές (απώλεια όσφρησης-ανοσμία).
- Οι διαταραχές του αυτόνομου νευρικού συστήματος.
- Η κόπωση.
- Η σταδιακή εξασθένιση της μνήμης, που μπορεί να οδηγήσει σε άνοια.

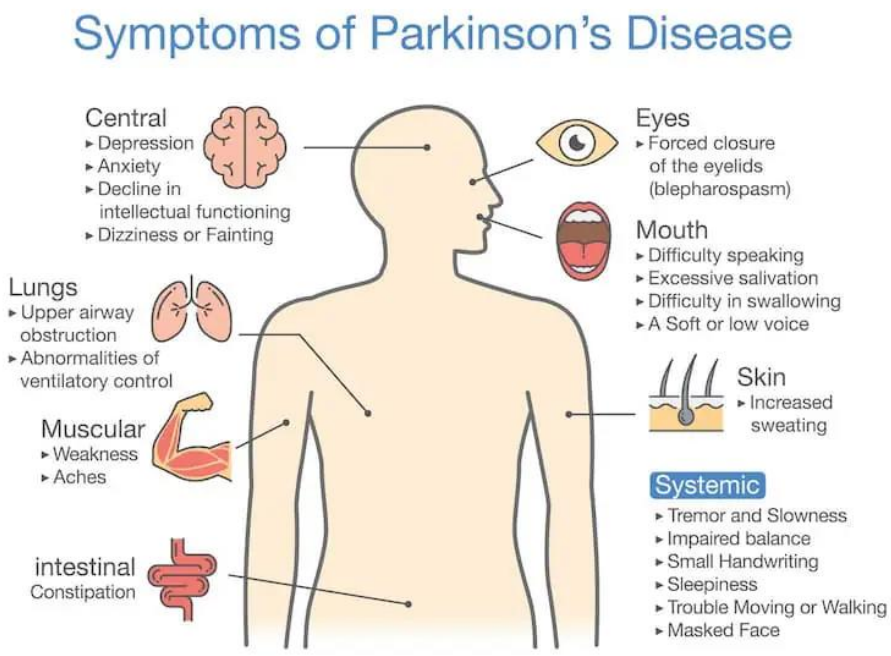
Αυτή η πληθώρα συμπτωμάτων της ΝΠ, τόσο σε κινητικά όσο και σε μη κινητικά, σε συνδυασμό με το γεγονός ότι η νόσος αυτή εξελίσσεται διαφορετικά από άτομο σε άτομο, δείχνουν ότι η ΝΠ δεν αποτελεί μια ομοιογενής ασθένεια αλλά εκδηλώνεται με διαφορετικούς τρόπους. Η ποικιλομορφία εμφάνιση της νόσου, καθιστά ακόμα πιο επιτακτική την ανάγκη για την ανάπτυξη καινοτόμων θεραπευτικών προσεγγίσεων, προσαρμοσμένων στον κάθε ασθενή ξεχωριστά. Αξίζει να αναφερθεί ότι ορισμένα μη κινητικά συμπτώματα, όπως οι διαταραχές ομιλίας, όπως προαναφέρθηκε, σε ορισμένες περιπτώσεις εκδηλώνονται αρκετά χρόνια πριν εμφανιστούν τα κύρια κινητικά σημεία, καθιστώντας την ανάλυση της ομιλίας και γενικότερα της φωνητικής λειτουργίας, έναν ευαίσθητο βιοδείκτη της ΝΠ [3]. Αυτή η αρκετά πρώιμη

εμφάνιση κάποιων χαρακτηριστικών, δίνει την δυνατότητα έγκαιρης ανίχνευσης της ΝΠ, πρωτού προκληθούν σοβαρές βλάβες στο νευρικό σύστημα του ατόμου. Αυτές οι αρχικές διαταραχές ομιλίας κρίνονται κλινικά σημαντικές και συμβάλλουν στην ανάπτυξη μεθόδων έγκαιρης διάγνωσης.

1.2.3 Διαταραχές Ομιλίας και Υποκινητική Δυσαρθρία

Οι διαταραχές ομιλίας είναι ένα από τα πιο σύνηθες μη κινητικά συμπτώματά της ΝΠ, επηρεάζοντας έως και το 90% των ασθενών [3]. Το κυρίαρχο φωνητικό χαρακτηριστικό είναι η υποκινητική δυσαρθρία η οποία επηρεάζει βασικές παραμέτρους της ομιλίας, όπως την αναπνοή, την φώνηση, την άρθρωση και την προσωδία. Η διαταραχή αυτή οφείλεται κυρίως στη μειωμένη κινητικότητα των μυών που συμμετέχουν στην παραγωγή του λόγου, με αποτέλεσμα τη φτωχή εκφορά, τον μονότονο ρυθμό και τη χαμηλή ένταση φωνής [3]. Επιπλέον σημαντικές αλλοιώσεις που αφορούν την φωνή του ατόμου είναι:

- Η υποφωνία (χαμηλή ένταση), που καθιστά δύσκολη την ομιλία.
- Η μονοτονία (απουσία τόνου και έντασης), καθιστώντας την ομιλία επίπεδη και χωρίς συναισθηματική χροιά.
- Ο ακανόνιστος ρυθμός στην ομιλία, που άλλοτε εκδηλώνεται ως βραδυφωνία και άλλοτε ως επιτάχυνση, περιλαμβάνοντας επίσης και ανεξήγητες παύσεις που δυσκολεύουν την σωστή ροή της ομιλίας .
- Η δυσχέρεια στην άρθρωση, με αποτέλεσμα ασαφής ομιλία
- Η αλλοιωμένη ποιότητα φωνής που μπορεί να εμφανίζεται με πιο βραχνή ή/ και τραχιά φωνή.



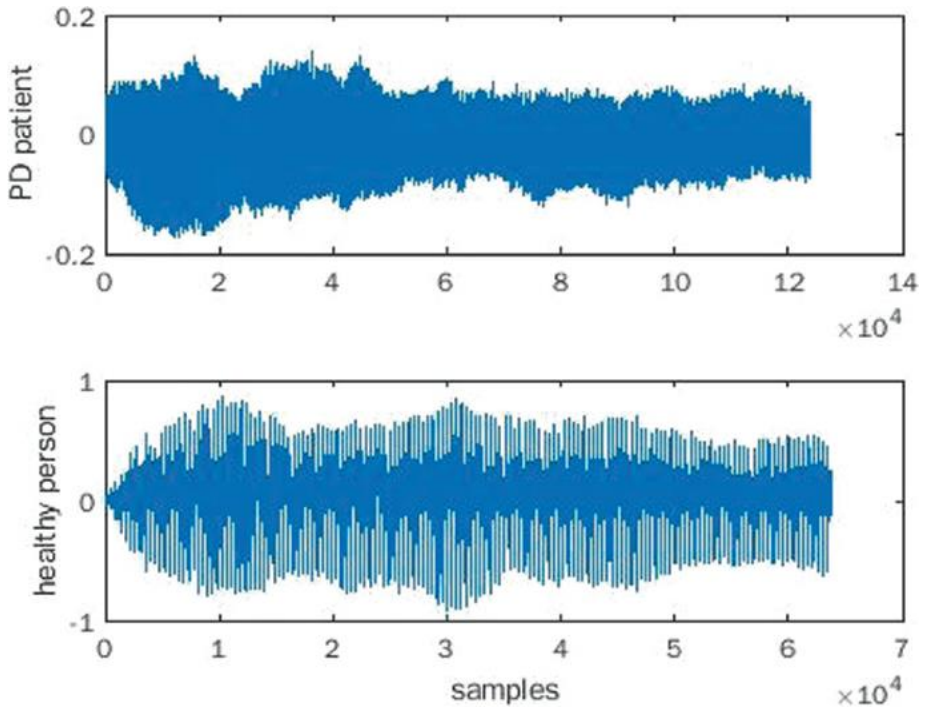
Εικόνα 2: Ενδεικτικά Συμπτώματα της Νόσου του Πάρκινσον ανά Σύστημα Οργάνων. [7]

Το γεγονός ότι επηρεάζονται ταυτόχρονα διαφορετικές πτυχές της ομιλίας, καταδεικνύει ότι η νόσος διαταράσσει την συνολική λειτουργία των συστημάτων που συνεργάζονται μεταξύ τους για την παραγωγή λόγου. Επομένως, η εμφάνιση όλων αυτών των αλλοιώσεων στην φωνή, ακόμη κα σε πρώιμα στάδια της ΝΠ, αποδεικνύουν ότι οι διαταραχές ομιλίας συγκαταλέγονται

στα κλινικά σημαντικά συμπτώματα της νόσου, αποτελώντας έναν αρχικό και αξιόπιστο δείκτη για την ανίχνευση της ΝΠ.
Για τον λόγο αυτό, η ανάλυση της ομιλίας με πιο σύνθετες μεθόδους, μπορεί να προφέρει μια αλληλουχία σημαντικών κλινικών πληροφοριών για το εκάστοτε στάδιο της νόσου, καθώς και για τον τρόπο και το χρονικό πλαίσιο εξέλιξης αυτής, συμβάλλοντας σε μια πιο έγκαιρη και ακριβή διάγνωση της ΝΠ.

1.3 Από την Κλινική Εκτίμηση στην Ανάλυση Ομιλίας

Η κλασσική διάγνωση της ΝΠ, βασίζεται παραδοσιακά στην κλινική εξέταση που γίνεται από ειδικούς γιατρούς-νευρολόγους, μέσω της προσεκτικής παρατήρησης των κινητικών συμπτωμάτων και της ανταπόκρισης του ασθενή στην χορήγηση λεβοντόπα [3]. Ωστόσο, στα αρχικά στάδια της νόσου, που τα συμπτώματα δεν είναι τόσο σοβαρά ακόμη και ίσως μοιάζουν με συμπτώματα που εμφανίζουν και άλλες παθήσεις, η ανίχνευση της ΝΠ ενδεχομένως ελλοχεύει αρκετές δυσκολίες. Έτσι, υπάρχει αυξημένος κίνδυνος εσφαλμένης ιατρικής διάγνωσης. Από την άλλη, οι απεικονιστικές τεχνικές που χρησιμοποιούνται για την διάγνωση της ΝΠ, όπως το Σπινθηρογράφημα Μεταφορέων Ντοπαμίνης (DaTscan-Dopamine Transporter Scan), η Τομογραφία Εκπομπής Ποζιτρονίων (PET-scan) και η Αξονική Τομογραφία (CT-scan) παρόλο που προσφέρουν σημαντικές πληροφορίες, συχνά είναι αρκετά δαπανηρές και ορισμένες από αυτές ακόμα και επεμβατικές [3]. Για τον λόγο αυτό η ιατρική κοινότητα ωθείται στην αναζήτηση νέων μεθόδων για την έγκαιρη και ακριβή διάγνωση της ΝΠ, για τον εντοπισμό νέων βιοδεικτών.



Εικόνα 3:Σύγκριση Ηχητικών Σημάτων Ασθενούς με ΝΠ και Υγιούς Ατόμου. [8]

Μέσα σε αυτό το διαγνωστικό πλαίσιο, η ανάλυση της ομιλίας έχει αποδειχθεί ένας υποσχόμενος τομέας προκειμένου να αναπτυχθούν νέες μέθοδοι για μη επεμβατική και χαμηλού κόστους διάγνωσης, καθώς τα δεδομένα ομιλίας μπορούν να συλλεχθούν εύκολα με απλές συσκευές (μικρόφωνο, έξυπνο τηλέφωνο) ακόμα και εξ αποστάσεως, διευκολύνοντας της διαγνωστική διαδικασία [3].
Η ανάλυση της ομιλίας και των εξαγόμενων ακουστικών χαρακτηριστικών, μπορούν να εντοπίσουν λεπτές αλλοιώσεις που προκαλεί η ΝΠ στους μηχανισμούς που είναι υπεύθυνοι για

την παραγωγή του λόγου, ακόμα και σε περιπτώσεις που δεν είναι ακόμα εύκολο να εντοπισθούν από την ανθρώπινη ακοή [3]. Αυτή η χρήση της ανάλυσης ομιλίας, ως διαγνωστικό εργαλείο, καταδεικνύει την ανάγκη της ιατρικής κοινότητας να στραφεί σε πιο φιλικές προς τον ασθενή μεθόδους, που μπορούν να εφαρμοστούν με ευκολία. Επίσης, εκτός από την ανίχνευση της νόσου και την διάκριση μεταξύ ασθενών και υγιών ατόμων, η ανάλυση ομιλίας μπορεί να αξιοποιηθεί και σε άλλες εφαρμογές που αφορούν την ΝΠ, όπως την παρακολούθηση της πορείας της νόσου, την επίδραση της φαρμακευτικής αγωγής, αλλά και στην γενικότερη κατανόηση των μηχανισμών που ευθύνονται για την εμφάνιση συμπτωμάτων της νόσου [3].

1.4 Η Συμβολή της Μηχανικής και Βαθιάς Μάθησης στην Ανίχνευση της Νόσου

Με δεδομένο ότι η ανάλυση ομιλίας ως βιοδείκτης για την ΝΠ, παρέχει σημαντικές πληροφορίες, χρήσιμες για την έγκαιρη διάγνωση της ΝΠ, η αξιοποίηση τεχνικών Μηχανικής (Machine Learning-ML) και Βαθιάς Μάθησής (Deep Learning-DL), σε συνδυασμό με αυτή την ανάλυση, προσφέρει μια πληθώρα ακόμα περισσότερων δυνατοτήτων για την επεξεργασία και την ερμηνεία των φωνητικών δεδομένων [3].
Πιο συγκεκριμένα, η ενσωμάτωση της Μηχανικής Μάθησης και ειδικότερα των τεχνικών Βαθιάς Μάθησης, συμβάλει στην βελτίωση της ανάλυσης των βιοϊατρικών σημάτων. Πιο ειδικά, τα μοντέλα Βαθιάς Μάθησης, που βασίζονται στην θεωρία των τεχνικών νευρωνικών δικτύων, είναι σε θέση να εντοπίζει αυτόματα μοτίβα και πληροφορίες από μη επεξεργασμένα δεδομένα, συμβάλλοντας στην μείωση της χειροκίνητης εξαγωγής χαρακτηριστικών [3]. Αυτή η ικανότητα καθιστά αυτές τις τεχνικές κατάλληλες για την ανάλυση της ομιλίας ασθενών με ΝΠ. Πληθώρα μελετών έχει δείξει ότι η χρήση παραδοσιακών μοντέλων Μηχανικής Μάθησης (Machine Learning – ML), όπως τα Support Vector Machines (SVM) και k-Nearest Neighbors (KNN), καθώς και των πιο σύγχρονων προσεγγίσεων της Βαθιάς Μάθησης (Deep Learning – DL), όπως τα Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks – RNNs) και τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks – CNNs), σε συνδυασμό με τα εξαγόμενα φωνητικά χαρακτηριστικά, παρουσιάζουν αρκετά καλές επιδόσεις στην ανίχνευση της νόσου, με υψηλή ακρίβεια διάγνωσης. Ωστόσο, παρά την υψηλή ακρίβεια των μοντέλων, παρουσιάζονται ακόμα προκλήσεις. Η ανάγκη ύπαρξης ποιοτικών βάσεων δεδομένων, στις οποίες προστατεύονται τα προσωπικά στοιχεία των ασθενών, με την απαίτηση απλής ερμηνείας των αποτελεσμάτων των μοντέλων από τους ιατρούς, αποτελούν βασικές προϋποθέσεις για την αξιοπιστία αυτών των εργαλείων στην κλινική πράξη [3]. Στο πλαίσιο αυτό, αξιοποιούνται πιο αυστηρές τεχνικές επικύρωσής, όπως η Leave-One-Subject-Out (LOSO), στις οποίες τα εκάστοτε μοντέλα εξετάζονται σε άγνωστα δεδομένα και έτσι αποφεύγεται το φαινόμενο της υπερεκπαίδευσης.
Αυτή η αυξανόμενη σημασία που δίνεται στην ερμηνευσιμότητα των αποτελεσμάτων των μοντέλων και την αξιοποίηση πιο αυστηρών μεθόδων αξιολόγησης, δείχνουν την πρόοδο και την βελτίωση των μεθόδων στον συγκεκριμένο κλάδο. Είναι κατανοητό ότι πλέον δεν αρκεί ένα μοντέλο να παρουσιάζει μόνο υψηλή ακρίβεια και απόδοση, θα πρέπει να είναι αξιόπιστο και σε θέση να αξιοποιηθεί σε μια εκτενής γκάμα εφαρμογών, καθώς και να είναι κατανοητό από την ιατρική κοινότητα με σκοπό την αποτελεσματική χρήση αυτού στην κλινική εξέταση.

1.5 Σκοπός και Συνεισφορά της Παρούσας Μελέτης στην Επιστημονική Έρευνα

Η παρούσα Διπλωματική Εργασία φιλοδοξεί να συνεισφέρει στην έρευνα για τις φωνητικές διαταραχές που εντοπίζονται στην ΝΠ. Αρχικά, προσπαθεί να ενισχύσει την αξιολόγηση της νόσου μέσω ποσοτικοποίησης των φωνητικών αλλοιώσεων και συσχέτισης αυτών με τις κλινικά σημαντικές παραμέτρους που εξετάζονται στην παρούσα μελέτη, όπως το στάδιο της

νόσου, η παρουσία ή όχι δυσκινησίας και κινητικών διακυμάνσεων, καθώς και η ανταπόκριση στην φαρμακευτική αγωγή. Επιδιώκεται δηλαδή να απαντηθούν τα κλινικά ερευνητικά ερωτήματα και πιο συγκεκριμένα να δοθεί απάντησή στο κατά πόσο τα εξαγόμενα φωνητικά χαρακτηριστικά είναι σε θέση να προβλέψουν με ακρίβεια σημαντικές παραμέτρους της ΝΠ.
Μέσω της εφαρμογής εξελιγμένων τεχνικών μηχανικής και βαθιάς μάθησης (συμπεριλαμβανομένων των Logistic Regression, KNN, Random Forest, SVM, SNNs και XGBoost), σε συνδυασμό με την ανάλυση των εξαγόμενων φωνητικών χαρακτηριστικών μέσω των δεδομένων ομιλίας των ασθενών, με έμφαση στο τυποποιημένο σύνολο χαρακτηριστικών eGeMAPS, η εργασία στοχεύει στον εντοπισμό χρήσιμων φωνητικών βιοδεικτών που θα συνεισφέρουν στην αντιμετώπιση της ΝΠ.
Οι εξαγόμενοι αυτοί βιοδείκτες ενδέχεται να συμβάλουν στον εντοπισμό και στην εκτίμηση της σοβαρότητας της ΝΠ αλλά και να παρατηρηθεί πως ο ασθενής αντιδρά στην εκάστοτε φαρμακευτική αγωγή. Έτσι, οι διαγνωστικές προσεγγίσεις μπορούν να διευκολυνθούν και να είναι πιο αποτελεσματικές.
Η εστίαση της παρούσας μελέτης σε Έλληνες ασθενείς (ερευνητικά κέντρα Ηρακλείου και Πάτρας), ενδέχεται να αναδείξει συγκεκριμένα χαρακτηριστικά ομιλίας, που συσχετίζονται με την ελληνική γλώσσα, συμβάλλοντας έτσι στην δημιουργία διαγνωστικών εργαλείων, προσαρμοσμένων στη συγκεκριμένη γλώσσα, προσδίδοντας ιδιαίτερη αξία στην μελέτη, ενισχύοντας την αξιοπιστία.
Στο πλαίσιο της ακολουθούμενης μεθοδολογίας, η αξιολόγηση διαφόρων αλγορίθμων μηχανικής και βαθιάς μάθησης, σε συνδυασμό με την αυστηρή μέθοδο επικύρωσής Leave-One-Subject-Out (LOSO), ενισχύει την εγκυρότητα των αποτελεσμάτων της έρευνας. Η ικανότητα των αξιοποιούμενων μοντέλων να εντοπίζουν χαρακτηριστικά ομιλίας, που μπορούν να συμβάλουν στον εντοπισμό του σταδίου της νόσου και των υπολοίπων παραμέτρων, υπερβαίνει την απλή ανίχνευση και αποσκοπεί σε μια πιο λεπτομερή καταγραφή της εξέλιξης την νόσου, συμβάλλοντας έτσι στην ανάπτυξη πιο εξειδικευμένων στρατηγικών παρακολούθησης. Επίσης , η ανάλυση της απόδοσης των μοντέλων σε επίπεδο μεμονωμένου ασθενούς και ο εντοπισμός ασθενών με συστηματικά χαμηλή ακρίβεια ταξινόμησης, έχει ιδιαίτερη σημασία για την μελέτη, διότι μπορεί να αναδείξει διαφοροποιήσεις στον τρόπο εμφάνισης των φωνητικών συμπτωμάτων ανά ασθενή ή να εντοπίσει αδυναμίες των τωρινών μεθόδων.
Τα αποτελέσματα της παρούσας μελέτης αποσκοπούν στον σχεδιασμό εξιδεικευμένων μη επεμβατικών και οικονομικά προσιτών εργαλείων διάγνωσης και παρακολούθησης της ΝΠ, με δυνατότητα ενσωμάτωσης σε τηλεϊατρικές εφαρμογές και συστήματα συνεχούς, 24/7 παρακολούθησης.

1.6 Δομή

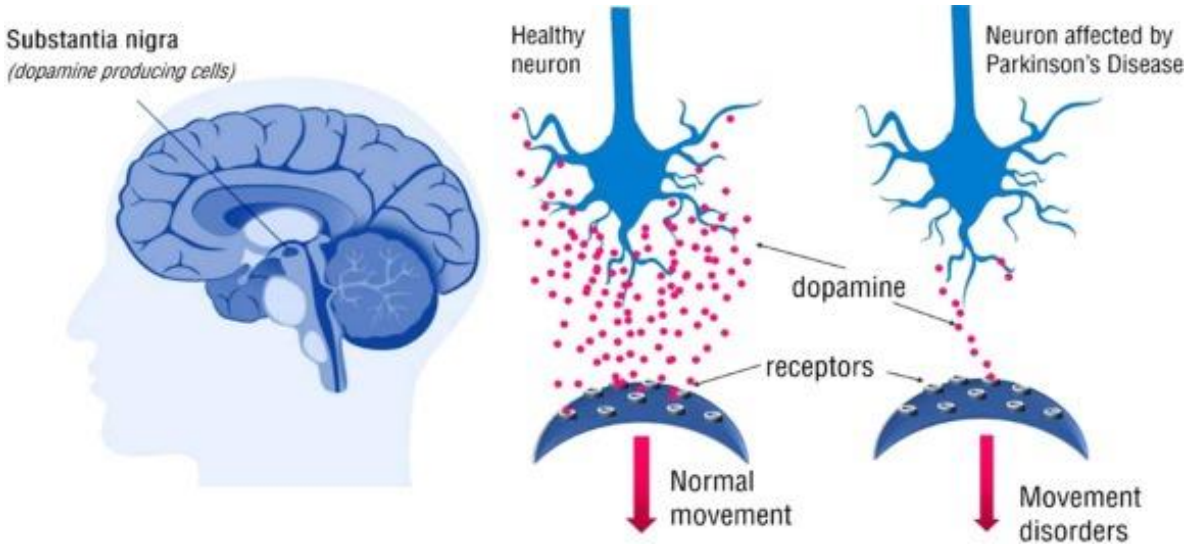
Το υπόλοιπο κείμενο της διπλωματικής εργασίας δομείται όπως ακολουθεί. Στο Κεφάλαιο 2 παρουσιάζεται το επιστημονικό υπόβαθρο της μελέτης. Γίνεται αναφορά στην ΝΠ, στα φωνητικά χαρακτηριστικά που συνδέονται άρρηκτα με αυτή , καθώς και με την σημασία της ομιλίας ως διαγνωστικό εργαλείο. Ταυτόχρονα, γίνεται αναφορά στις τεχνικές μηχανικής και βαθιάς μάθησης που έχουν εφαρμοστεί στο σύνολο των μελετών, με περισσότερη έμφαση στα σύνολα φωνητικών χαρακτηριστικών, στα στατιστικά και ταξινομητικά μοντέλα καθώς και στην συμβολή της αυστηρής μεθόδου επικύρωσης LOSO στην συνολική διαδικασία.
Στο Κεφάλαιο 3 περιλαμβάνεται η αναλυτική περιγραφή της συλλογής και επεξεργασίας δεδομένων, της εξαγωγής φωνητικών χαρακτηριστικών, της εφαρμογής στατιστικών αναλύσεων και την ανάπτυξη των μοντέλων ταξινόμησης.
Το Κεφάλαιο 4 παρουσιάζει τα αποτελέσματα της ανάλυσης, τόσο σε επίπεδο χαρακτηριστικών, όσο και σε επίπεδο απόδοσης των μοντέλων ταξινόμησης στις διάφορες κλινικές κατηγορίες. Στο κεφάλαιο αυτό περιλαμβάνεται η εκτενής ανάλυση των περιπτώσεων,

όπου παρατηρήθηκε χαμηλή απόδοση τόσο των μοντέλων όσο και σε επίπεδο ασθενών, με κύριο στόχο την κατανόηση πιθανών διαφοροποιήσεων μεταξύ των ασθενών ή αδυναμία στο τρόπο ανάλυσης των δεδομένων.
Στο Κεφάλαιο 5, γίνεται παράθεση των βασικών ευρημάτων της μελέτης και παρουσιάζονται τα κύρια συμπεράσματα που προκύπτουν από την ανάλυση των φωνητικών χαρακτηριστικών και της απόδοσης των ταξινομητικών μοντέλων, αναδεικνύοντας τα δυνατά σημεία της έρευνας αλλά και τα σημεία που ενδεχομένως χρήζουν βελτίωσης. Παράλληλα, εντοπίζονται αδυναμίες και περιορισμοί της έρευνας τα οποία αποτελούν αντικείμενο μελλοντικής έρευνας. Η εργασία ολοκληρώνεται με την Βιβλιογραφία και τα Παραρτήματα που περιλαμβάνουν αναλυτικές πληροφορίες που ενισχύουν την κατανόηση της ερευνητικής διαδικασίας.

2 Βιβλιογραφική Ανασκόπηση / Θεωρητικό Υπόβαθρο

2.1 Γενικές Πληροφορίες

Η Νόσος του Πάρκινσον αποτελεί μια από τις συχνότερες νευροεκφυλιστικές διαταραχές σε παγκόσμιο επίπεδο, που επηρεάζει τόσο τον κινητικό έλεγχο του ατόμου καθώς και μη κινητικές λειτουργίες, συμπεριλαμβανομένης και της ομιλίας, δημιουργώντας μια σημαντική επιβάρυνση στο σύστημα υγείας [9], [10]. Σύμφωνα με πιο πρόσφατες έρευνες, μέχρι το 2040 εκτιμάται ότι ο αριθμός των ατόμων που νοσούν από την ΝΠ θα έχει διπλασιαστεί, καθιστώντας την πρώιμη και ακριβή διάγνωση μείζονος σημασίας για την αποτελεσματικότερη διαχείριση της νόσου [2]. Η ΝΠ παρουσιάζει ένα ευρύ φάσμα συμπτωμάτων τα οποία διακρίνονται σε κινητικά και μη κινητικά. Τα κύρια κινητικά συμπτώματα περιλαμβάνουν τον τρόμο, την βραδυκινησία, την μυϊκή δυσκαμψία και διαταραχές τόσο στη στάση του σώματος, όσο και στην διατήρηση της ισορροπίας [11]. Πέρα από αυτά τα συμπτώματα, πολύ σημαντικά είναι και τα μη κινητικά όπως οι διαταραχές ύπνου, η κόπωση καθώς και οι διαταραχές που εντοπίζονται στην ομιλία του ατόμου. Παρά το γεγονός ότι τα κινητικά προβλήματα έχουν μελετηθεί εκτενέστερα, πολύ σημαντική είναι και η αξιολόγηση των μη κινητικών συμπτωμάτων διότι και αυτά επηρεάζουν σε υψηλό βαθμό την καθημερινότητα και την ποιότητα ζωής των ατόμων [12] [13]. Μεταξύ αυτών των μην κινητικών συμπτωμάτων, δίνεται μεγάλή βαρύτητα στις διαταραχές ομιλίας, οι οποίες εμφανίζονται ακόμα και από τα πρώιμα στάδια της νόσου. Υπολογίζεται ότι περίπου το 90% των ατόμων που πάσχουν από ΝΠ εμφανίζουν διαταραχές ομιλίας όπως η φωνητική άρθρωση, η φώνηση, η προσωδία και η ποιότητα φωνής, με τις πιο κύριες διαταραχές ομιλίας να θεωρούνται η δυσφωνία και η δυσαρθρία, καθώς και άλλα συμπτώματα όπως η μονότονη φωνή , η μειωμένη ένταση και η ασάφεια στην ομιλία [9], [14], [15].



Εικόνα 4:Οπτική απεικόνιση της επίδρασης της Νόσου του Πάρκινσον στους νευρώνες και τη μεταφορά κινητικών σημάτων [16].

Όλα αυτά τα συμπτώματα αποτελούν επίσης και πρώιμους δείκτες της ΝΠ, που μπορούν να εμφανιστούν μέχρι και χρόνια πριν την εκδήλωση των κύριων κινητικών συμπτωμάτων. Επομένως, οι διαταραχές που επηρεάζουν την ποιότητα της ομιλίας του ατόμου, όπως γίνεται αντιληπτό, έχουν αρνητικές επιπτώσεις στην ποιότητα ζωής των ασθενών. Οπότε, η έγκαιρη διάγνωση και διαχείριση της ΝΠ μέσω ανάλυσης της ομιλίας, λαμβάνοντας υπόψιν ποικίλα ακουστικά/γλωσσικά χαρακτηριστικά, κρίνεται σημαντική [17].

2.2 Ο Ρόλος της Ανάλυσης Ομιλίας στην Νόσο του Πάρκινσον

Η ανάλυση της ομιλίας σε ασθενείς με ΝΠ είναι μια μη επεμβατική και εύκολα εφαρμόσιμη μέθοδος για την ανίχνευση και διαχείριση της νόσου σε πρώιμο στάδιο, καθώς και άλλων χαρακτηριστικών που συνδέονται με την ΝΠ [18]. Η ευκολία εφαρμογής της μεθόδου βασίζεται στο γεγονός ότι τα δεδομένα μπορούν να συλλεχθούν με απλές και οικονομικές μεθόδους, μέσω κοινών μικροφώνων και άλλων έξυπνων συσκευών, χωρίς την επιτακτική ανάγκη για εξειδικευμένο εξοπλισμό. Επίσης, σαν μέθοδος δίνει την δυνατότητα απομακρυσμένης παρακολούθησης σε εφαρμογές τηλεϊατρικής [19]. Μέσω της ανάλυσης ομιλίας σε άτομα με ΝΠ, γίνεται αξιολόγηση των χαρακτηριστικών ομιλίας που συνδέονται με την νόσο και επηρεάζονται άρρηκτα από αυτή, καθιστώντας αυτή την τεχνική ένα πολύ χρήσιμο εργαλείο, που επιτρέπει την συνεχή παρακολούθηση [20]. Τα δεδομένα ομιλίας, τα οποία αποτελούν και τον πυρήνα της εφαρμογής αυτής της μεθόδου, μπορούν να συλλεχθούν με μεγάλη ευκολία μέσω κοινών μικροφώνων ή έξυπνων συσκευών, καθιστώντας την διαδικασία μια προσβάσιμη και μη επεμβατική μέθοδο [19]. Όλα αυτά τα χαρακτηριστικά καθιστούν την συγκεκριμένη μέθοδο ανάλυσης αρκετά πιο εύχρηστη, συγκριτικά με τις πιο ακριβές τεχνικές νευροαπεικόνισης, όπως η μαγνητική τομογραφία (MRI), η λειτουργική μαγνητική τομογραφία (fMRI), ή οι μέθοδοι PET και SPECT, οι οποίες απαιτούν εξειδικευμένο εξοπλισμό και υψηλό κόστος εφαρμογής [21]. Αυτή η ευκολία στην συλλογή των δεδομένων, καθώς και η δυνατότητα απομακρυσμένης ανάλυσης, μπορούν να συμβάλουν στην ανίχνευση και στην παρακολούθηση της ΝΠ στον ευρύ πληθυσμό και σε έναν πιο συστηματικό ρυθμό. Αυτό καθιστά την ανάλυση ομιλίας ένα πολύτιμο εργαλείο στα χέρια των επιστημόνων, για την έγκαιρη διάγνωση και παρακολούθηση της εξέλιξής της ΝΠ [22].

2.3 Χρήση Μοντέλων Βαθιάς Μάθησης (Deep Learning) για την Ανάλυση της Ομιλίας

Η ραγδαία εξέλιξη της Τεχνητής Νοημοσύνης και πιο συγκεκριμένα τα άλματα που έχουν γίνει στης εφαρμογές που βασίζονται στην Βαθιά Μάθηση (Deep Learning), προσφέρουν σημαντικές δυνατότητες στην ανάλυση δεδομένων ομιλίας και κατά συνέπεια στην ανίχνευση της ΝΠ και άλλων παραμέτρων, σημαντικών για την διερεύνηση της σημαντικότητας της νόσου [23].

Η Βαθιά Μάθηση αποτελεί έναν προηγμένο τομές της Μηχανικής Μάθησης (Machine Learning), ο οποίος βασίζει τις βασικές του αρχές στην χρήση τεχνητών νευρωνικών δικτύων πολλαπλών επιπέδων (Multi-Layered Artificial Neural Networks). Αυτά τα νευρωνικά δίκτυα βασίζουν την αρχή λειτουργίας τους στον τρόπο λειτουργίας του ανθρωπίνου εγκεφάλου, οπότε η χειροκίνητη εξαγωγή των χαρακτηριστικών δεν απαιτείται και τα μοντέλα μαθαίνουν αυτόματα από αυτά τα δεδομένα. Λόγω της φύσης των νευρωνικών δικτύων, αυτά είναι σε θέση να αναγνωρίζουν πιο περίπλοκες σχέσεις μεταξύ των δεδομένων, οπότε κρίνονται κατάλληλα για εφαρμογές που απαιτείται ανάλυση περίπλοκων δεδομένων [24]. Η ικανότητα αυτή έχει ουσιαστική συμβολή σε εφαρμογές επεξεργασίας φωνητικής γλώσσας, αναγνώριση εικόνας/ήχου, καθώς και την ανάλυση σημάτων ομιλίας που τα μοτίβα είναι αρκετά πολύπλοκα και ίσως η ανίχνευση τους από παραδοσιακές μεθόδους είναι δύσκολη. Τα μοντέλα βαθιάς μάθησης προσφέρουν την δυνατότητα εξαγωγής πληθώρας χαρακτηριστικών από δεδομένα, προσαρμοσμένα στις εκάστοτε ιδιότητες της ομιλίας των ασθενών που πάσχουν από ΝΠ. Επιπλέον, ένα μεγάλό πλεονέκτημα αυτών των μοντέλων είναι το ότι λόγω της φύσης τους, έχουν την δυνατότητα να αναγνωρίσουν και να καταλήξουν σε νέους βιοδείκτες,, σημαντικούς για την αξιολόγηση της ΝΠ [25].

Η παρούσα βιβλιογραφική ανασκόπηση διερευνά τις πρόσφατες ερευνητικές εργασίες που αφορούν την εξαγωγή και ερμηνεία των χαρακτηριστικών ομιλίας, που συσχετίζονται με την ΝΠ, μέσω της χρήσης τεχνικών Βαθιάς Μάθησης. Για την εξαγωγή των χαρακτηριστικών δίνεται έμφαση σε έρευνες που κάνουν χρήση του συνόλου ακουστικών παραμέτρων που

αφορούν την ομιλία, το eGeMAPS, το οποίο έχει σχεδιαστεί ειδικά για την ανάλυση παραγόντων που σχετίζονται με την φωνή και γενικά την ομιλία. Το σύνολο eGeMAPS αποτελείται από 88 ακουστικά χαρακτηριστικά τα οποία έχουν αποδειχθεί αξιόπιστα και άμεσα σχετιζόμενα με μεταβολές που αντανακλούν την ψυχοσωματική και νευρολογική κατάσταση του ατόμου, καθιστώντας το κατάλληλο για την μελέτη της ΝΠ. Επιπλέον, η χρήση του διευκολύνει την συγκρισιμότητα μεταξύ των μελετών, προσφέροντας κοινό σημείο αναφοράς [26], [27]. Σχετικά με την αξιολόγηση αυτών των ακουστικών χαρακτηριστικών, εξετάζεται η εφαρμογή μοντέλων Βαθιάς Μάθησης, καθώς και ο συνδυασμός αυτών με μοντέλα Μηχανικής Μάθησης (Machine Learning). Βασικός σκοπός, αποτελεί μια ολοκληρωμένη επισκόπηση της τρέχουσας βιβλιογραφίας στον τομέα ανίχνευσης και αξιολόγησης της ΝΠ και ο εντοπισμός πιθανών μελλοντικών κατευθύνσεων της έρευνας.

2.4 Χαρακτηριστικά Ομιλίας στην Νόσο του Πάρκινσον: Μια Λεπτομερής Ανασκόπηση

2.4.1 Ο Δυσαρθρία στην Νόσο του Πάρκινσον (ΝΠ)

Ένα από τα βασικά συμπτώματα της ΝΠ, που επηρεάζει την ομιλία του ασθενή, είναι η υποκινητική δυσαρθρία. Ουσιαστικά είναι μια διαταραχή που επιδρά αρνητικά στην κινητικότητα των μυών, κάνοντας τους να κινούνται με πιο αργό ρυθμό. Οι μυς αυτοί είναι υπεύθυνοι για την ορθή παραγωγή του λόγου [28]. Οι πτυχές της ομιλίας που εξαρτώνται σε μεγαλύτερο βαθμό από την διαταραχή αυτή είναι η φώνηση, η αναπνοή, η άρθρωση και η προσωδία. Σαν κύριο σύμπτωμα της νόσου , παρουσιάζει έναν πολυδιάστατο χαρακτήρα καθώς επιδρά σε πολλαπλά υποσυστήματα του λόγου. Τα υποσυστήματα του λόγου περιλαμβάνουν το αναπνευστικό, φωνητικό, αρθρωτικό και προσωδιακό τα οποία συνεργάζονται μεταξύ τους με σκοπό την φυσιολογική παραγωγή του λόγου. Αν παρουσιαστεί διαταραχή έστω και σε ένα από αυτά τα υποσυστήματα μπορεί να επηρεαστεί η ομιλία του ατόμου. Για τον λόγο αυτό η χρήση ολοκληρωμένων τεχνικών ανάλυσης, που θα είναι σε θέση να καταγράφουν ένα ευρύ φάσμα φωνητικών αλλαγών, κρίνεται απαραίτητη. Έτσι απαιτούνται μοντέλα που θα είναι ικανά να αναλύουν διάφορα ακουστικά χαρακτηριστικά που συνδέονται με τους μηχανισμούς παραγωγής ομιλίας [9].

2.4.2 Άλλα Σημαντικά Χαρακτηριστικά Ομιλίας που αφορούν την ΝΠ

Εκτός από την δυσαρθρία, ασθενείς με ΝΠ εμφανίζουν ανωμαλίες και σε άλλες πτυχές της ομιλίας. Αρκετοί από αυτούς έχουν μειωμένη ένταση φωνής (υποφωνία), η άρθρωση των συμφώνων γίνεται με ασαφή τρόπο και επίσης παρουσιάζουν και μειωμένη διακύμανση τόνου (μονοτονία). Ένα ακόμη χαρακτηριστικό είναι η εμφάνιση βραχνής ή τραχιάς ποιότητας φωνής, καθώς και τρόμο/αστάθεια κατά την διάρκεια της ομιλίας. Επιπλέον, σε ορισμένα άτομα παρουσιάζεται μεταβολή του ρυθμού ομιλίας, που συχνά συνδέεται με επιβράδυνση αυτής, όπως και παρατεταμένες παύσεις ή σιωπές που δεν ταιριάζουν στο εκάστοτε πλαίσιο. Τώρα το αν κάποιος ασθενής θα εμφανίσει συγκεκριμένα από αυτά τα συμπτώματα ή όχι, εξαρτάται αποκλειστικά στην προδιάθεση του ίδιου του ατόμου αλλά και στο στάδιο της νόσου στο οποίο βρίσκεται. Προφανώς και δεν θα παρουσιάσουν όλοι οι ασθενείς με ΝΠ την ίδια συμπτωματολογία, αλλά ακόμα και αν κάποιοι παρουσιάζουν ορισμένα κοινά συμπτώματα, αυτά δεν θα εμφανιστούν στον ίδιο βαθμό [9]. Προκειμένου να επιτευχθεί υψηλή διαγνωστική ακρίβεια αυτών των συμπτωμάτων, γίνεται χρήση μοντέλων βαθιάς μάθησης. Αυτά τα μοντέλα θα πρέπει να είναι σε θέση να αναγνωρίζουν τις μικρές αλλαγές στις διακυμάνσεις της φωνής. Πρακτικά, θα πρέπει τα μοντέλα που θα αξιοποιηθούν να προγραμματιστούν με τέτοιο τρόπο έτσι ώστε να μάθουν από τα μοτίβα που δημιουργούνται από την ανάλυση και ίσως να αναγνωρίσουν υποτυπους της ΝΠ, βασισμένους στα χαρακτηριστικά ομιλίας [9], [28].

2.4.3 Τα Παραδοσιακά Χαρακτηριστικά που Χρησιμοποιούνται για την Ανάλυση Ομιλίας στην ΝΠ

Η ανάλυση της ομιλίας για την ΝΠ επικεντρώνεται κυρίως στα φωνητικά χαρακτηριστικά που παρουσιάζουν ασυνήθιστες μετρήσεις, συγκριτικά με τις φυσιολογικές τιμές που παρουσιάζει ένα υγιές άτομο. Επιπλέον, από την μελέτη της συνεχούς ομιλίας εντοπίζονται οι κύριες αρθρωτικές διαταραχές όπως το μειωμένο εύρος και ταχύτητα κινήσεως της γλώσσας, των χειλιών και της γνάθου. Αυτές οι διαταραχές επηρεάζουν την προφορά κλειστών συμφώνων και φωνημάτων [2]. Άλλα σημαντικά χαρακτηριστικά που μελετώνται για την ανάλυση της ομιλίας σε άτομα με ΝΠ είναι τα προσωδιακά χαρακτηριστικά, τα οποία περιλαμβάνουν την διακύμανση του τόνου της φωνής, τον ρυθμό της ομιλίας καθώς και τις ανεξήγητες παύσεις. Όπως γίνεται αντιληπτό, η ανάλυση των χαρακτηριστικών ομιλίας πρέπει να γίνεται στο σύνολο αυτών και όχι μόνο σε μεμονωμένα χαρακτηριστικά, προκειμένου να καταγραφεί πλήρως το φάσμα των αλλαγών που επιφέρει η νόσος στην ομιλία [29]. Για να γίνει η ποσοτικοποίηση αυτών των διαταραχών λόγου, η ακουστική ανάλυση διαδραματίζει σημαντικό ρόλο, καθώς μπορεί να εντοπίσει τις λεπτές αποκλίσεις στην ομιλία, οι οποίες δεν είναι τόσο εύκολο να γίνουν αντιληπτές από τον απλό ακροατή. Αυτό καθιστά την ακουστική ανάλυση μια μέθοδο που προσδίδει μεγαλύτερη ακρίβεια και αξιοπιστία στην αξιολόγηση των χαρακτηριστικών, με σκοπό μια πιο σταθερή και μετρήσιμη μελέτη της επίδρασης της ΝΠ στην παραγωγή λόγου. Για μια εις βάθος και ακριβή ανάλυση αυτών των ηχητικών σημάτων, μέσω της εξαγωγής ποσοτικών χαρακτηριστικών γίνεται χρήση μοντέλων βαθιάς μάθησης. Η χρήση αυτών των τεχνικών βοηθούν στον διαχωρισμό μεταξύ ατόμων με ΝΠ και υγειών, καθώς και άλλων κατηγοριών που μπορεί να εντοπίζονται (σε ποιο στάδιο της νόσου βρίσκονται οι ασθενείς κ.τ.λ.) [30].

Για την καλύτερη κατανόηση των επιπτώσεων της ΝΠ στην ομιλία, ο Πίνακας 1 παρουσιάζει συνοπτικά τις κύριες κατηγορίες χαρακτηριστικών ομιλίας και τις τυπικές αλλοιώσεις που παρατηρούνται στους ασθενείς που εμφανίζουν συμπτώματα της νόσου.

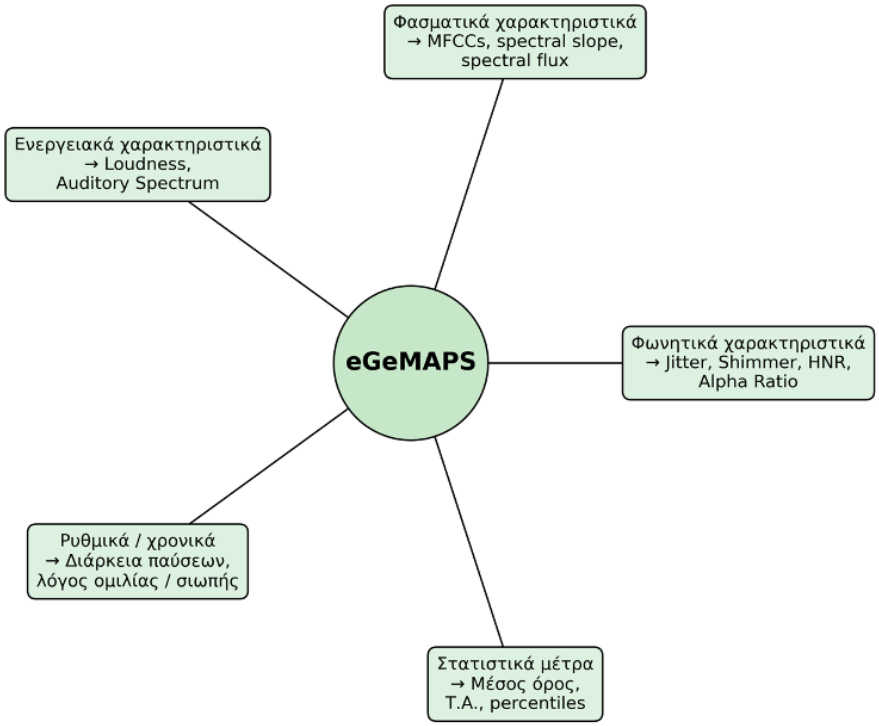
Πίνακας 1:Συνοπτική Απεικόνιση των Διαταραχών Ομιλίας σε Ασθενείς με ΝΠ.

Κατηγορία Χαρακτηριστικών Ομιλίας	**Τυπικές Αλλοιώσεις στη ΝΠ**
Ένταση φωνής	Υποφωνία(χαμηλή ένταση φωνής)
Τόνος / Προσωδία	Μονότονη φωνή/εκφραστικότητα
Ρυθμός Ομιλίας	Αργή/ακανόνιστη ομιλία
Αρθρωτική Ικανότητα	Ασαφή άρθρωση, μειωμένη ακρίβεια στην προφορά συμφώνων/φωνηέντων
Ποιότητα Φωνής	Βραχνάδα, τρόμος, φωνητική αστάθεια
Αναπνευστικός Συντονισμός	Αδυναμία ελέγχου αναπνοής κατά την ομιλία
Συνολική προσωδία	Έλλειψη εκφραστικότητας
Ταχύτητα Εκφοράς Λόγου	Απότομη επιτάχυνση ή επιβράδυνση της εκφοράς λόγου

Παύσεις/Διακοπές Ομιλίας	Παρατεταμένες/ακατάλληλες παύσεις που επηρεάζουν την ομαλή ροή του λόγου

2.5 Χρήση του eGeMAPS (extended Geneva Minimalistic Acoustic Parameter Set) στην Ανάλυση Ομιλίας

2.5.1 Το Σύνολο Χαρακτηριστικών eGeMAPS



Εικόνα 5: Κατηγοριοποίηση των Ακουστικών Χαρακτηριστικών του eGeMAPS.

Στο τομέα της ακουστικής ανάλυσης, ένα πρότυπο που χρησιμοποιείται ευρέως είναι το eGeMAPS (extended Geneva Minimalistic Acoustic Parameter Set). Ουσιαστικά είναι ένα τυποποιημένο σύνολο 88 ακουστικών χαρακτηριστικών που εφαρμόζονται στην ανάλυση ομιλίας [25], [26]. Το σύνολο αυτό έχει την ικανότητα να καταγράφει διάφορες διακυμάνσεις που εντοπίζονται στην φωνή, οι οποίες συσχετίζονται με την συναισθηματική κατάσταση του ατόμου και προφανώς και με υποκείμενες νόσους [26]. Αυτά τα 88 εξαγόμενα χαρακτηριστικά προκύπτουν μέσω υπολογιστικών συναρτήσεων (μέση τιμή, τυπική απόκλιση, jitter, shimmer κ.τ.λ.) σε περιγραφητές χαμηλού επιπέδου (LLD – Low Level Descriptors), που λαμβάνονται από την ακουστική καταγραφή της ομιλίας. Το σύνολο eGeMAPS, δίνει την δυνατότητα συγκρίσεων των αποτελεσμάτων μεταξύ διαφόρων ερευνών και συμβάλει συνδυασμό δεδομένων/μελετών, αυξάνοντάς έτσι την αποτελεσματικότητα των εκάστοτε μεθόδων που χρησιμοποιούνται [25]. Ουσιαστικά προσφέρει ένα κοινό σύνολο χαρακτηριστικών για τις έρευνες που διεξάγοντα, παίζοντας τον ρόλο μιας κοινής βάσης αναφοράς [26].

2.5.2 Χρησιμότητα του Συνόλου eGeMAPS στην Ανάλυση Ομιλίας για την ΝΠ

Όπως έγινε αναφορά και παραπάνω, εκτός από τα κινητικά προβλήματα που παρουσιάζονται στα άτομα με την ΝΠ, σημαντικά είναι και τα συμπτώματα που επηρεάζουν την παραγωγή της ομιλίας [31]. Το σύνολο eGeMAPS περιλαμβάνει πληθώρα ακουστικών χαρακτηριστικών που συμβάλουν στον εντοπισμό οποιασδήποτε αλλοίωση στην φωνή του ατόμου. Κάποια από τα κύρια χαρακτηριστικά που εντοπίζονται στο σύνολο και συσχετίζονται με ΝΠ είναι η θεμελιώδης συχνότητα (F0), η ενέργεια που έχει να κάνει με την ένταση της φωνής και τα φασματικά χαρακτηριστικά που αφορούν την ποιότητα της φωνής (τρόμος, βραχνάδα κ.τ.λ.) [25]. Η φύση του eGeMAPS να καταγράφει αλλαγές στις διακυμάνσεις της φωνής και να καλύπτει διάφορους τομείς που σχετίζονται με την παραγωγή της ομιλίας, το καθιστούν ένα πολύ χρήσιμο εργαλείο για την ανάλυση της ομιλίας σε άτομα που νοσούν από Πάρκινσον. Επιπλέον, επειδή η νόσος μπορεί να επηρεάσει αρνητικά την ομιλία του ατόμου με ποικίλους τρόπους, αυτό το σύνολο φωνητικών χαρακτηριστικών, διαθέτοντας ένα μεγάλο εύρος ακουστικών παραμέτρων, αυξάνει τις πιθανότητες να εντοπιστούν πιο αξιόπιστοι βιοδείκτες [26].

2.5.3 Το eGeMAPS ως Εργαλείο Ακουστικής Ανάλυσης: Σύγχρονες Εφαρμογές

Οι πρόσφατες έρευνες δείχνουν ένα ενδιαφέρον για την χρήση του συνόλου eGeMAPS στην ανάλυση της ομιλίας για την ΝΠ, καθώς και στην ανίχνευση της νόσου, στην κατηγοριοποίηση της κατάστασης του ασθενή και στην αναγνώριση της αποτελεσματικότητας της θεραπείας. Η χρήση αυτού του συνόλου ακουστικών χαρακτηριστικών χρησιμοποιούνται ευρέως, εφόσον θεωρείται ένα αρκετά αξιόπιστο εργαλείο για την διεξαγωγή των ερευνών. Παρά την ανοδική χρήση των διάφορων μοντέλων ταξινόμησης, η χρήση του eGeMAPS παραμένει ακόμα μια πολύ βοηθητική μέθοδος για τον εντοπισμό συγκεκριμένων ακουστικών χαρακτηριστικών, που ανάλογα με την αλλοίωση τους, συνδέονται με την ΝΠ.

2.5.4 Ανίχνευση ΝΠ Μέσω του Συνόλου eGeMAPS

Η χρήση του συνόλου eGeMAPS έχει συμβάλει στην ανίχνευση της ΝΠ στους ασθενείς, μέσω της ανάλυσης των ακουστικών χαρακτηριστικών που εξάγονται μέσω αυτού, από ηχητικά δεδομένα ατόμων. Πιο συγκεκριμένα, η μελέτη των N.Momeni et al [25]. Ασχολήθηκε με την ανίχνευση της ΝΠ μέσω ανάλυσης ηχητικών σημάτων που ηχογραφήθηκαν με την χρήση κινητού τηλεφώνου. Η εξαγωγή των ακουστικών χαρακτηριστικών πραγματοποιήθηκε με την χρήση του συνόλου eGeMAPS και οι ερευνητές προκειμένου να μειώσουν τις προκαταλήψεις λόγω ηλικίας και φύλου, έκαναν εφαρμογή μιας νέας τεχνικής. Η τεχνική αυτή περιλάμβανε την χρήση SHAP (SHapley Additive exPlanations) ανάλυσης για την εξαγωγή των πιο σημαντικών χαρακτηριστικών και τα κύρια ευρήματα αυτής της τεχνικής έδειξαν μονότονη φωνή και ανωμαλίες που εντοπιζόταν στον έλεγχο της αναπνοής, χαρακτηριστικά άρρηκτα συνδεδεμένα με την ΝΠ. Η ακρίβεια αυτού του μοντέλου έφτασε στο 80,3% και ανάκληση σε άγνωστα άτομο 89,4%, ποσοστά που υπερτερούν συγκριτικά με άλλες παρόμοιες μελέτες. Η συγκεκριμένη έρευνα ανέδειξε την σημασία της υψηλής ακρίβειας σε μοντέλα ανίχνευσης της ΝΠ, μέσω της ανάλυσης ομιλίας, με εξαγόμενα ακουστικά χαρακτηριστικά από το σύνολο eGeMAPS. Επίσης η εφαρμογή της μεθοδολογίας και σε άγνωστα άτομα, υπογραμμίζει την ανάγκη για δυνατότητα γενίκευσης των μοντέλων. Τώρα, σχετικά με την σημαντικότητα των φωνητικών χαρακτηριστικών, παραπάνω πληροφορίες παρέχει η μελέτη των A.Pompili et al [32], η οποία επικεντρώνεται στην ανάλυση ηχηρών/άφωνων τμημάτων ηχητικών σημάτων, μέσω δεδομένων που προέκυψαν από κινητό τηλέφωνο (δεδομένα mPower). Η εξαγωγή των χαρακτηριστικών έγινε μέσω eGeMAPS και η ανάλυση αυτών, όπως και στην προηγούμενη μελέτη, πραγματοποιήθηκε με την χρήση SHAP ανάλυσης. Σε αυτή την έρευνα, σημαντικότερα χαρακτηριστικά αναδείχθηκαν αυτά που σχετίζονται με τα ηχηρά/άφωνά

τμήματα της ομιλίας (Τυπική Απόκλιση, Ρυθμός), που είναι και ο κεντρικός πυλώνας της μελέτης. Η υψηλότερες τιμές αυτών των χαρακτηριστικών συνδέθηκαν με την ύπαρξη ΝΠ, γεγονός που συμβαδίζει με τις ήδη γνωστές διαταραχές ομιλίας για την ΝΠ. Η σύγκλιση με την πρώτη αναφερόμενη έρευνα ενισχύει τον ρόλο της εξαγωγής αυτών των χαρακτηριστικών, ως βιοδείκτες της νόσου.
Μια ακόμα ενδιαφέρουσα προσέγγιση αποτελεί η μελέτη των Niemelä et al. [33], οι οποίοι για την ταξινόμηση της ομιλίας έκαναν χρήση της ακουστικής βάσης δεδομένων Pitt Corpus. Η εξαγωγή των χαρακτηριστικών και σε αυτή την έρευνα πραγματοποιήθηκε με την συμβολή του eGeMAPS και το υψηλό ποσοστό ταξινόμησης (77,4%) προέκυψε μέσω της χρήσης 75 μη συσχετισμένων χαρακτηριστικών του συνόλου και ταξινομητή LDA (Linear Discriminant Analysis), τονίζοντας την σημαντικότητα της στρατηγικής επιλογής των κατάλληλων χαρακτηριστικών. Το γεγονός ότι δεν επιλέχθηκαν όλα τα χαρακτηριστικά για την ανάλυση, βελτίωσε την απόδοση, δείχνοντας ότι δεν συνεισφέρουν όλα τα χαρακτηριστικά εξίσου και δεν έχουν την ίδια βαρύτητα στο κάθε πρόβλημα/ εφαρμογή. Επιπλέον, η σύγκριση του συνόλου eGeMAPS με άλλα σύνολα ακουστικών χαρακτηριστικών (emobase, ComParE), κατέδειξε την υπεροχή του. Αξιοσημείωτη είναι και η μελέτη που διεξήχθη στο πλαίσιο μιας διπλωματικής εργασίας του Πανεπιστήμιου του Lund [34], η οποία εστιάζει στην χρήση της άρθρωσης των φωνηέντων, ως βιοδείκτες για την ανίχνευση την ΝΠ, συνδυάζοντας ακουστική ανάλυση και χρήση Μηχανικής Μάθησης (Machine Learning – ML/XGBoost) σε δεδομένα από το Voice Diagnostics. Αρχικά, επιλέχθηκαν 88 χαρακτηριστικά από το eGeMAPS, τα οποία στην συνέχεια μειώθηκαν στα 50. Η ακρίβεια του μοντέλου, παρόλο ήταν μέτρια/υψηλή, παρουσίαζε αυξημένη διακύμανση στην διασταυρούμενη επικύρωση (Cross – Validation), γεγονός που υπονομεύει την αξιοπιστία του μοντέλου. Τα ευρήματα της έρευνας αυτής, δείχνουν ότι η χρήση φωνηέντων από απλές προτάσεις δεν συμβάλει τόσο στην ανίχνευση της νόσου και επίσης αναδεικνύει και ορισμένους περιορισμούς του συνόλου eGeMAPS για ορισμένους τύπους ομιλίας, καθιστώντας σαφές ότι η περίπτωση αυτή χρήζει περαιτέρω έρευνας.

2.5.5 Ταξινόμηση Φαρμακευτικής Κατάστασης στην ΝΠ Μέσω eGeMAPS

Πέρα από την ανίχνευση της ΝΠ, το σύνολο eGeMAPS έχει χρησιμοποιηθεί και για την ταξινόμηση της φαρμακευτικής κατάστασης (ON/OFF) των ασθενών. Στην μελέτη των A. M.Pompili et al. [32], εξετάζεται η αναγνώριση της φαρμακευτικής κατάστασης ασθενών μέσω ανάλυσης ακουστικών χαρακτηριστικών από το σύνολο eGeMAPS, σε ηχητικά δεδομένα ομιλίας του FraLusoPark. Η ακρίβεια του μοντέλου άγγιξε το 95,2%, ποσοστό που υπερτερεί από άλλες μελέτες. Ωστόσο, η ακρίβεια μειώθηκε όταν μειώθηκε και η διαστατικότητα μέσω PCA (Principal Component Analysis), τονίζοντας την σημαντικότητα της πλήρους αναπαράστασης των εξαγομένων χαρακτηριστικών. Όλα αυτά επικυρώνουν την δυνατότητα του συνόλου να αναγνωρίζει λεπτές αλλαγές στην ομιλία, οι οποίες συσχετίζονται με την φαρμακευτική αγωγή του ατόμου, γεγονός αρκετά σημαντικό για την διεξαγωγή περαιτέρω μελετών στον τομέα αυτό. Η αξιοποίηση του eGeMAPS πέρα από τους τομείς που αναφέρθηκαν παραπάνω, επεκτείνεται και στην αξιολόγηση θεραπευτικών παρεμβάσεων. Πιο συγκεκριμένα, στην έρευνα των Van Bemmel et al. [35], γίνεται αξιολόγηση της επίδρασης μιας λογοθεραπευτικής προσέγγισης, μέσω παιχνιδιών που βασίζονται στην ομιλία ασθενών με ΝΠ. Μέσω της εξαγωγής ακουστικών χαρακτηριστικών από το eGeMAPS πριν την διεξαγωγή του παιχνιδιού, καθώς και μετά την περάτωση αυτού, εντοπίστηκαν χαρακτηριστικά που οι τιμές αυτών διέφεραν μετά την θεραπεία (αυξημένη ένταση ομιλίας και μείωση τονικού ύψους). Τα αποτελέσματα της έρευνας έδειξαν ότι το σύνολο ακουστικών χαρακτηριστικών, είναι ικανό να ανιχνεύει και να ποσοτικοποιεί βελτιώσεις που εμφανίζονται στην ομιλία του ασθενή, παρουσιάζοντας το ως μια αποτελεσματική λύση σε ορισμένες λογοθεραπευτικές προσεγγίσεις που αφορούν την ΝΠ.

Συγκεντρωτικά, το σύνολο eGeMAPS θεωρείται ένα αξιόπιστο σύνολο ακουστικών χαρακτηριστικών, που χρησιμοποιείται ευρέως για στρατηγικές ανάλυσης ομιλίας, που συσχετίζονται με την ΝΠ. Παρά την πρόοδο των πιο εξελιγμένων τεχνικών αξιολόγησης, το eGeMAPS παραμένει ένα πολύτιμο και αξιόλογο εργαλείο για την ανάλυση ηχητικών σημάτων που προέρχονται από ασθενείς με ΝΠ, λόγω της απλότητας αλλά και της υπολογιστικής του ικανότητας. Οι μελέτες που εξετάσθηκαν, δείχνουν ότι το συγκεκριμένο σύνολο χρησιμοποιείται σε ένα μεγάλο εύρος ιατρικών εφαρμογών και ερευνητικών πλαισίων, διευκολύνοντας την στατιστική ανάλυση και την εφαρμογή τεχνικών ταξινόμησης.

Οι μελέτες που παρουσιάζονται παρακάτω πίνακα (Πίνακα 2), συνοψίζουν πρόσφατες έρευνες που κάνουν χρήση του συνόλου eGeMAPS με σκοπό την ανάλυση της ομιλίας σε ασθενείς με ΝΠ, καλύπτοντας ένα ευρύ φάσμα εφαρμογών, όπως η ανίχνευση της νόσου, η ταξινόμηση της φαρμακευτικής κατάστασης και η αξιολόγηση θεραπευτικών παρεμβάσεων. Η ποικιλία των μεθόδων καθώς και τα αποτελέσματα των μελετών επιβεβαιώνουν την χρησιμότητα του eGeMAPS ως εργαλείο για την ανίχνευση της ΝΠ, ενώ παράλληλα υπογραμμίζει την ανάγκη για εκτενέστερη έρευνα και προσαρμογή αυτού στις εκάστοτε ερευνητικές απαιτήσεις.

Πίνακας 2: Συγκριτικός Πίνακας Μελετών με Χρήση του eGeMAPS: Δεδομένα, Χαρακτηριστικά και Ευρήματα.

Μελέτη	Σύνολο Δεδομένων	Συγκεκριμένα Χαρακτηριστικά eGeMAPS που Χρησιμοποιήθηκαν	Βασικά Ευρήματα σχετικά με το eGeMAPS
N.Momeni et al. (2024) [25]	mPower	Χαμηλή τυπική απόκλιση στην κλίση της F0 (πτώση και άνοδος), μεγαλύτερα άφωνα τμήματα, αυξημένη τυπική απόκλιση έντασης, μειωμένη τυπική απόκλιση F0, αυξημένος μέσος τοπικός τρόμος (jitter)	Αυτά τα χαρακτηριστικά συνδέονται με αυξημένη πιθανότητα ταξινόμησης ως ομιλία ατόμου με ΝΠ. Το μοντέλο πέτυχε 80,3% ακρίβεια σε άγνωστα άτομα.
Niemelä et al (2024) [33]	Pitt Corpus	75 μη συσχετισμένα χαρακτηριστικά από το σύνολο eGeMAPS.	Η χρήση μη συσχετισμένων χαρακτηριστικών eGeMAPS με ταξινομητή LDA πέτυχε αποτέλεσμα ταξινόμησης (77,4%).
Φοιτητική Διπλωματική Εργασία, Πανεπιστήμιο Lund (2024) [34]	Voice Diagnostics	50 χαρακτηριστικά από το σύνολο eGeMAPS (αρχικά 88).	Μέτρια ακρίβεια αλλά χαμηλή αξιοπιστία λόγω υψηλής διακύμανσης. Η ανάλυση φωνηέντων με eGeMAPS δεν βελτίωσε την έγκαιρη ανίχνευση σε σύγκριση με προηγούμενες εργασίες.
A.Pompili et al. (2023) [32]	mPower	Δεν αναφέρονται συγκεκριμένα χαρακτηριστικά.	Τα πιο σημαντικά χαρακτηριστικά για την ταξινόμηση της ΝΠ ήταν η τυπική απόκλιση, ο ρυθμός του ηχηρού τμήματος και η τυπική απόκλιση του άφωνου τμήματος.

Van Bemmel et al (2021) [35]	Δεδομένα ομιλίας πριν και μετά τη θεραπεία μέσω του παιχνιδιού "Treasure Hunters"	HammarbergIndexUV_sma3nz_amean, slopeV0-500_sma3nz_amean, F0semitoneFrom27.5Hz_sma3nz_percentile 50	Μετά τη θεραπεία, αυξήθηκαν οι τιμές των HammarbergIndex και slopeV0-500 (ένδειξη αύξησης της έντασης) και μειώθηκε η τιμή του F0semitone-Median (ένδειξη μείωσης του τόνου φωνής).
Pompili et al. (2020) [27]	FraLusoPark	26 LLDs και συναρτήσεις από το σύνολο eGeMAPS.	Υψηλή ακρίβεια στην ταξινόμηση της φαρμακευτικής κατάστασης: 90,54% σε μικτή ομιλία και 95,27% σε ημι-αυθόρμητη αφήγηση.

2.6 Συστηματική Ανασκόπηση Εφαρμογών LM/LMM στην Ποσοτική Ανάλυση Ομιλίας στη ΝΠ

Πριν την αξιολόγηση των ταξινομητικών αλγορίθμων είναι σημαντικό να εξετασθεί και η υπάρχουσα βιβλιογραφία στην ανάλυση της ομιλίας της ΝΠ μέσω χρήσης Απλών Γραμμικών Μοντέλων(LM) και Μεικτών Γραμμικών Μοντέλων(LMM).

Τα Απλά Γραμμικά Μοντέλα (LM) και Μεικτά Γραμμικά Μοντέλα (LMM) είναι χρήσιμα εργαλεία στην διαδικασία της στατιστική ανάλυση δεδομένων ομιλίας, ειδικά σε περιπτώσεις όπου έχουν καταγραφεί πολλαπλές ηχογραφήσεις για τον κάθε ασθενή. Ειδικότερα, τα LMM είναι κατάλληλα για τέτοιοι είδους δεδομένα, διότι λαμβάνουν σοβαρά υπόψιν τις σταθερές παραμέτρους, δηλαδή η ύπαρξη της νόσου ή η ανταπόκριση σε κάποια θεραπεία, καθώς και τις παραμέτρους που διαφοροποιούνται από ασθενή σε ασθενή, όπως τις ιδιαιτερότητες που εντοπίζονται στην ομιλία του ατόμου. Η ανάλυση των χαρακτηριστικών με την χρήση αυτών των μεθόδων καθώς και η αξιολόγηση της στατιστικής σημαντικότητας τους με βάση την τιμή p (p-value), συμβάλει στον εντοπισμό των κατάλληλων ακουστικών χαρακτηριστικών που έχουν άμεση σύνδεση με την ΝΠ [36].

Η κατανόησή αυτών των χαρακτηριστικών αλλά και της στατιστικής τους σημαντικότητας μέσω των μοντέλων LM και LMM, αποτελεί κομβικό σημείο στην σχεδίαση και αξιολόγηση των μοντέλων βαθιάς μάθησης [36], [23]. Τα χαρακτηριστικά αυτά μπορούν να αποτελέσουν ένα από τα κύρια σημεία στην επιλογή των πιο σημαντικών μεταβλητών εισόδου, καθώς και στην ερμηνεία των αποτελεσμάτων [24]. Επομένως, τα μοντέλα στατιστικής ανάλυσης LM/LMM και τα μοντέλα βαθιάς μάθησης δεν αποτελούν αντικρουόμενες ιδέες αλλά λειτουργούν συμπληρωματικά [36], [37]. Τα στατιστικά μοντέλα εντοπίζουν ποια χαρακτηριστικά ομιλίας σχετίζονται σε μεγαλύτερο βαθμό με την ΝΠ, ενώ τα μοντέλα βαθιάς μάθησης είναι σε θέση να αναγνωρίσουν πιο σύνθετα και πολύπλοκα μοτίβα τα οποία δεν ανιχνεύονται με τόση ευχέρεια μόνο μέσω συμβατικών στατιστικών προσεγγίσεων [23], [37].

2.6.1 Ανασκόπηση Πρόσφατων Μελετών με Χρήση Γραμμικών Μοντέλων (LM/LMM)

Η ανάλυση δεδομένων ομιλίας που σχετίζονται με την ΝΠ, έχει επωφεληθεί από την χρήση των Απλών Γραμμικών Μοντέλων και των Μεικτών Γραμμικών Μοντέλων, διότι συμβάλλουν στον εντοπισμό σημαντικών διαφοροποιήσεων, λαμβάνοντας υπόψιν αλλαγές στα δεδομένα των ατόμων [29]. Αρκετές μελέτες αξιοποιούν τα οφέλη αυτών των μοντέλων στατιστικής ανάλυσης για την αξιολόγηση των χαρακτηριστικών και την παρακολούθηση αλλαγών που εντοπίζονται στο σύνολο δεδομένων.

Η μελέτη των Thies et al. [36] έκανε χρήση LMM προκειμένου να αξιολογηθεί η επίδραση της λεβοντόπα στα χαρακτηριστικά ομιλίας σε άτομα με ΝΠ. Τα αποτελέσματα ανέδειξαν αρκετές στατιστικά σημαντικές μεταβλητές, ενώ η χρήση του LMM συνέβαλε στο έλεγχο των παραγόντων που θα μπορούσαν να επηρεάσουν την ακεραιότητα των αποτελεσμάτων και τις διακυμάνσεις που εντοπίζονται στο ίδιο το άτομο, εντοπίζοντας σημαντικές επιδράσεις της λεβοντόπα σε ακουστικά χαρακτηριστικά. Η παρέμβαση του στατιστικού μοντέλου στην όλη διαδικασία επέδρασσε κυρίως στο εντοπισμό βελτιώσεων στις φωνητικές λειτουργίες των ασθενών και στην διαμόρφωση της ομιλίας, χωρίς να φαίνεται να επηρεάζει σε τέτοιο βαθμό το αρθρωτικό σύστημα. Σε παρόμοιο πλαίσιο κινήθηκε και η μελέτη των Karimi et al. [38], που ερεύνησαν τις αλλαγές σε χαρακτηριστικά ομιλίας σε άτομα που νοσούσαν με ΝΠ και υγιή άτομα με την εφαρμογή του μοντέλου LMM. Διαπιστώθηκε στατιστική επίδραση του χρόνου στην παράμετρο F0SD (p-value = 0.0084), ωστόσο η επίδραση ομάδας/χρόνου στην ίδια παράμετρο δεν είχε στατιστική αξία. Από την άλλη, για το χαρακτηριστικό IntSD δεν υπήρξε κάποια επίδραση με την πάροδο του χρόνου.
Αξιοποιώντας πάλι τα οφέλη του μοντέλου LMM, η μελέτη των L. Brabenec et al. [39]. εξέτασε τις επιδράσεις της tDCS (transcranial Direct Current Stimulation**)** στην κινητική ικανότητα ασθενών με ΝΠ, μέσω στατιστικής ανάλυσης ομιλίας με την εφαρμογή LMM. Η ακουστική ανάλυση μέσω LMM έδειξε ότι αυτή η διέγερση αυτή επηρέασε την μέση διάρκεια των παύσεων (p-value = 0.015), γεγονός που δείχνει ότι ίσως βελτιώθηκε η εικόνα του ασθενή.
Τέλος, η μελέτη των H. Steurer et al. [40], ασχολήθηκε με την επίδραση της ομαδικής παρέμβασης HiCommunication στην ομιλία, σε άτομα με ΝΠ. Σημαντικές αλληλεπιδράσεις εντοπίσθηκαν για την ηχητική στάθμη της φωνής στην ανάγνωση (p-value = 0.003), στον μονόλογο (p-value = 0.009), αλλά και στον δείκτη ποιότητας φωνής (p-value = 0.016).
Οι μελέτες αυτές αναδεικνύουν τα οφέλη της εφαρμογής των στατιστικών μοντέλων ανάλυσης LM/LMM για την ανάλυση ομιλίας στην ΝΠ όταν υπάρχουν επαναλαμβανόμενες μετρήσεις. Τα μοντέλα αυτά συμβάλουν στην ανίχνευση των στατιστικά σημαντικών χαρακτηριστικών, βοηθώντας στην βαθύτερη κατανόηση της νόσου και συμβάλλοντας στην δημιουργία νέων διαγνωστικών εργαλείων.

Στον πίνακα που ακολουθεί (Πίνακας 3), συνοψίζονται οι πληροφορίες που αφορούν τις μελέτες που αναλύθηκαν παραπάνω.

Πίνακας 3: Μελέτες που Χρησιμοποιούν Στατιστικά Μοντέλα για την Ανάλυση Φωνητικών Χαρακτηριστικών σε Ασθενείς με Νόσο του Πάρκινσον.

Μελέτη	Στατιστικό Μοντέλο & Δείγμα	Είδος Ομιλίας & Χαρακτηριστικά	Ευρήματα & Τιμές p
Thies et al. (2025) [36]	LMM, 78 ΑΣΘ	DDK, μονόλογος, ανάγνωση – jitter, shimmer, stdF0, VOT	Βελτίωση σε med-ON, p=0.001–0.033
Karimi et al. (2025) [38]	LMM (ANOVA), 29 ΑΣΘ / 29 ΥΜ	Περιγραφή εικόνας – F0SD, IntSD	Επίδραση χρόνου στο F0SD (p=0.0084)
L. Brabenec et al. (2024) [39]	LMM, 14 ΑΣΘ	Συλλαβές, φωνήεντα – DurMED, relF0SD	DurMED μειώθηκε μετά από tDCS, p=0.014

H. Steurer et al. (2024) [40]	LMM, 95 ΑΣΘ	Ανάγνωση, μονόλογος – VSL, AVQI	Σημαντική βελτίωση στην παρέμβαση, p=0.003–0.016
Frankford, E. L., et al. (2024) [41]	Linear Regression, n/a	Αντιληπτική αξιολόγηση – αρθρωτική ακρίβεια	Καταληπτότητα σχετίζεται με ακρίβεια, p όχι δηλωμένο
Ghosal et al. (2024) [42]	Univariate Regression, 13 ΑΣΘ	n/a – CPPS, jitter, shimmer	CPPS ↔ UPDRS III, p=0.01–0.04

2.7 Διερεύνηση της Μεθόδου Leave-One-Subject-Out (LOSO) στην Εκπαίδευση/Δοκιμή Μοντέλων Βαθιάς Μάθησης

Η ανάπτυξη ισχυρών μοντέλων βαθιάς μάθησης προκειμένου να εντοπισθούν στατιστικά σημαντικά χαρακτηριστικά ομιλίας στην ΝΠ απαιτεί να χρησιμοποιηθούν αξιόπιστοι μέθοδοι όπως η Leave-One-Subject-Out (LOSO). Η εφαρμογή αυτής της μεθόδου βασίζεται στην λογική της προσομοίωσής ενός ρεαλιστικού κλινικού σεναρίου, όπου ο ασθενής και τα δεδομένα που συνδεόταν με αυτόν δεν αποτελούν μέρος της εκπαίδευσης [43]. Η προσέγγιση αυτή διασφαλίζει ότι δεν θα υπάρξει αλληλοεπικάλυψη μεταξύ των δεδομένων δοκιμής και εκπαίδευσης, γεγονός που συνεπάγεται ότι ο κίνδυνος υπερεκπαίδευσης μειώνεται αισθητά. Η μέθοδος LOSO ουσιαστικά διασφαλίζει ότι η απόδοση του εκάστοτε εξεταζόμενου μοντέλου βασίζεται αποκλειστικά στις ικανότητες του να αναγνωρίζει πρότυπα που ταυτίζονται με την ασθένεια και να εξετάζει την απόδοση του μοντέλου και κατά πόσο μπορεί να αποδώσει σε νέους ασθενείς και δεδομένα [44]. Η συγκεκριμένη μέθοδος αποτελεί την κατάλληλη επιλογή για σύνολα δεδομένων με πολλαπλές φωνητικές ηχογραφήσεις, καθώς και σε ιατρικές εφαρμογές που εξετάζονται παράμετροι που αφορούν την ανίχνευση και την γενικότερη έκβαση μιας νόσου διότι διέπεται από μειωμένη μεροληψία, αυξημένη συνάφεια και ικανοποιητική απόδοση ενισχύοντας της αξιοπιστία των ευρημάτων [45].

Η μέθοδος LOSO έχει εφαρμοστεί και αξιολογηθεί σε ένα μεγάλο φάσμα μοντέλων ταξινόμησης, από απλά τεχνητά νευρωνικά δίκτυα έως πιο σύνθετα μοντέλα σχετικά με φωνητικά χαρακτηριστικά ομιλίας. Οι επιδόσεις των μοντέλων που έκαναν χρήση της LOSO, αν και διέφεραν μεταξύ τους και δεν ακολουθούσαν το ίδιο πρότυπο ως προς την επίδοση, με άλλες μελέτες να έχουν πιο υψηλά ποσοστά ακρίβειας και άλλες πιο χαμηλά, ήταν σε θέση να παρέχουν μια πιο ρεαλιστική και αυστηρή προσέγγιση σε κλινικές εφαρμογές.

Όπως αναφέρθηκε και παραπάνω, η άμεση σύγκριση μεταξύ των μελετών κρίνεται δύσκολη εξαιτίας των διαφορετικών συνόλων δεδομένων, των στατιστικά σημαντικών χαρακτηριστικών που επιλέχθηκαν και την γενικότερη σύνθεση των αξιολογούμενων μοντέλων. Σε γενικές γραμμές, η μέθοδος LOSO, παρόλο που παρουσιάζει πιο συντηρητικές αποδόσεις, παραμένει αξιόπιστη λόγω του τρόπου λειτουργίας της. Ενδεικτικά, ένα ποσοστό μελετών έχει δείξει τιμές AUC κοντά στην τάξη του 0.68 και ακρίβεια περίπου στο 74-86% με χρήση της μεθόδου LOSO, ενώ άλλες μελέτες έχουν εξάγει πολύ καλύτερες επιδόσεις που πλησιάζουν ακόμα και το απόλυτο ποσοστό 100% σε ειδικά διαμορφωμένα σύνολα. Η μεθοδολογία αυτή διακατέχεται από ένα μεγαλύτερο ποσοστό αυστηρότητας συγκριτικά με άλλες μεθόδους, όμως αυτή η αυστηρότητα ταυτόχρονα προσδίδει μεγαλύτερη αξιοπιστία στα τελικά ευρήματα, αναλαμβάνοντας τον ρόλο μιας πιο ρεαλιστικής προσέγγισης που ενισχύει την εγκυρότητα των αποτελεσμάτων και την γενίκευση της μεθόδου σε ένα ευρύτερο φάσμα περιπτώσεων.

Στον παρακάτω πίνακα παρουσιάζεται μια λίστα μελετών που χρησιμοποίησαν την μέθοδο Leave-One-Subject-Out (LOSO) στην αξιολόγηση ταξινομητικών μοντέλων για την ανάλυση ηχητικών δεδομένων ομιλίας που αφορούν την ΝΠ. Ο πίνακας συνοψίζει βασικά στοιχεία της εκάστοτε μελέτης όπως το χρησιμοποιούμενο μοντέλο, τα χαρακτηριστικά που αξιοποιήθηκαν, οι μετρικές απόδοσης και όλα εκείνα τα δεδομένα που θα συμβάλουν στην κατανόηση των ερευνητικών τάσεων και της αποτελεσματικότητάς της μεθόδου.

Πίνακας 4:Ενδεικτικές μελέτες που αξιοποίησαν επικύρωση τύπου LOSO για την ανάλυση φωνητικών χαρακτηριστικών στη Νόσο του Πάρκινσον.

Μελέτη	Μοντέλο	Χαρακτηριστικά Ομιλίας	Μετρικές Απόδοσης
Quan et al.(2021) [46]	BiLSTM	Δυναμικά χαρακτηριστικά άρθρωσης (BBEs, MFCCs)	Ακρίβεια: 75.56%
Laganas et al. (2021) [47]	Μη καθορισμένο	Χαρακτηριστικά φωνής από τηλεφωνικές κλήσεις	AUC: 0.68
Berus et al. (2019) [48]	ANNs	26 χαρακτηριστικά Praat	Ακρίβεια: έως 86.47%
Ali et al. (2019) [49]	LDA-NN-GA	26 χαρακτηριστικά Praat	Ακρίβεια: 100% (δοκιμή, όλα τα χαρακτηριστικά), 82.14% (χαρακτηριστικά ανεξάρτητα φύλου)
Tougui et al. (2024) [50]	Audio Spectrogram Transformer (AST)	Φασματογραφήματα (από δεδομένα φωνής mPower)	AUC: 0.9589, Ακρίβεια: 91.35%
Siniukov et al. (2025) [51]	HuBERT + DAPT + DAT	Μαθημένα χαρακτηριστικά από ομιλία	Μέση Ειδικότητα:92.1%, Ευαισθησία: 91.2%

2.8 Ταξινόμηση της ΝΠ Μέσω της Χρήσης Τεχνικών Βαθιάς Μάθησης.

2.8.1 Εισαγωγή

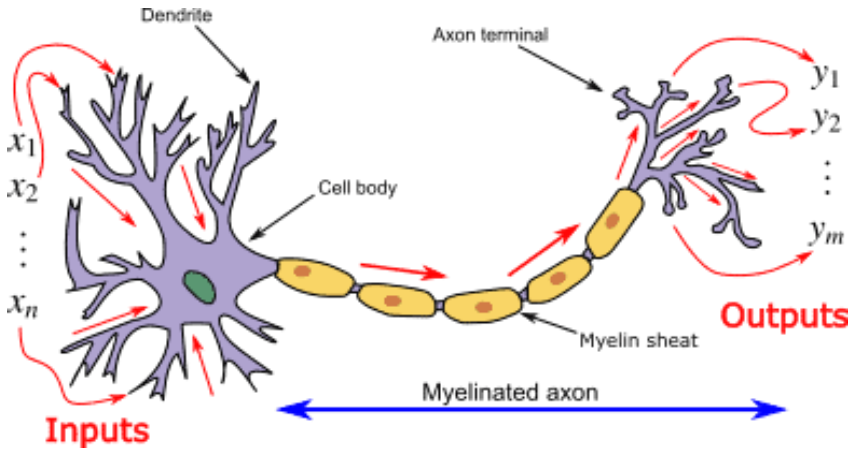
Τα τελευταία χρονιά η χρήση τεχνικών βαθιάς μάθησης, σε συνδυασμό με τεχνικών μηχανικής μάθησης έχει αποδειχθεί ένα πολύ χρήσιμο εργαλείο για έναν πιο αυτοματοποιημένο τρόπο ανίχνευσης της ΝΠ, μέσω της ανάλυσης ομιλίας και της εξαγωγής των ακουστικών χαρακτηριστικών. Καθώς η εμφάνιση αυτής της νευροεκφυλιστικής νόσου έχει αυξηθεί σε

παγκόσμιο επίπεδο, η χρήση αποτελεσματικών και εύκολων στην εφαρμογή διαγνωστικών τεχνικών, κρίνεται απαραίτητη. Οπότε η ανάλυση της ομιλίας σε συνδυασμό με τις διάφορες σύγχρονες τεχνικές μάθησης, μπορεί να συμβάλει ουσιαστικά στην ενίσχυση των ιατρικών διαδικασιών, προσφέροντας μια συμπληρωματική μέθοδο διάγνωσης [52].

2.9 Εφαρμογή Μοντέλων Μάθησης για την Ανίχνευση της Νόσου του Πάρκινσον Μέσω της Ομιλίας

2.9.1 Ο Αλγόριθμος Spiking Neural Network (SNN) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας

Τα SNN ουσιαστικά αποτελούν την 3η γενιά νευρωνικών δικτύων, τα οποία είναι γνωστά για την απόδοση τους και την μεγάλη ομοιότητα τους με τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου [53]. Αυτή η ιδιαιτερότητα τους τα διαφοροποιούν από τα παραδοσιακά νευρωνικά δίκτυα (ANN - Artificial Neural Network) και τα βαθιά νευρωνικά δίκτυα (DNN - Deep Neural Network). Σε αυτό το είδος νευρωνικών δικτύων, η οποιαδήποτε πληροφορία μεταδίδεται μεταξύ τεχνιτών νευρώνων μέσω ακίδων, εκπέμποντας 'αιχμές' (spikes), οι οποίες μιμούνται τα ηλεκτρικά σήματα που μεταδίδονται μεταξύ των νευρώνων του εγκεφάλου [17]. Αυτό που κάνει τα SNN ένα πολύ ενδιαφέρον μοντέλο ταξινόμησης, είναι η ικανότητα τους να αξιοποιούν την χρονική πληροφορία διότι κάθε φορά που εμφανίζεται μια αιχμή, μεταδίδονται σημαντικά δεδομένα [54]. Όλα αυτά καθιστούν το συγκεκριμένο μοντέλο μια έγκυρη μεθοδολογική προσέγγιση για ανάλυση δεδομένων που εξελίσσονται με την πάροδο του χρόνου, όπως και η ανάλυση της ανθρώπινης ομιλίας, η οποία εξετάζεται στην παρούσα μελέτη [55].



Εικόνα 6: Βιολογικός νευρώνας ως πρότυπο για τα Spiking Neural Networks (SNN). Τα ακουστικά χαρακτηριστικά της ομιλίας εισάγονται ως inputs, επεξεργάζονται στον πυρήνα και οδηγούν σε ταξινομητικά outputs. [56]

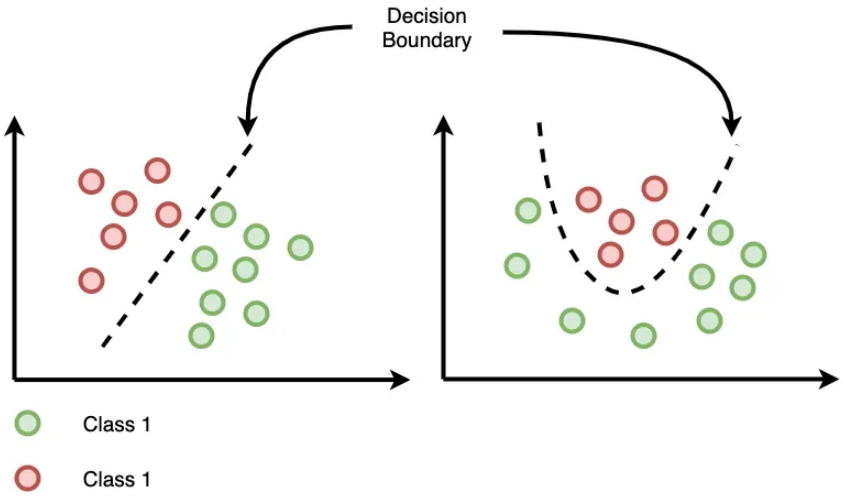
Μια πρόσφατη μελέτη των Das et al. [52] έκανε χρήση ενός καινοτόμου μοντέλου SNN,το SEFRON, προκειμένου να γίνει ανίχνευση της ΝΠ μέσω ανάλυσης της ομιλίας των ατόμων. Το SEFRON βασίζεται σε μια μεταβαλλόμενη συνάρτηση συναπτικής αποτελεσματικότητας και σε έναν νευρώνα leaky-intergrade-and-fire. Για την συγκεκριμένη έρευνα έγινε χρήση δύο συνόλων ηχητικών δεδομένων από το UCI Machine Learning Repository:

- Oxford Parkinson's Disease Detection Dataset (31 άτομα, 22 χαρακτηριστικά).
- Parkinson Dataset (80 άτομα, 44 χαρακτηριστικά).

Ακόλουθα, για την προ-επεξεργασία των δεδομένων, έγινε κανονικοποίηση των εισόδων και για την μετέπειτα μεθοδολογία αρχικά τα δεδομένα μετατράπηκαν σε χρόνους αιχμών, έγινε σχεδιασμός της αρχιτεκτονικής του SEFRON μοντέλου και της συναπτικής πλαστικότητας, μέσω STDP (Spike-Timing Dependent Plasticity) και σαφώς εκπαίδευση των συνάψεων και μετέπειτα η ταξινόμηση αυτών. Τα αποτελέσματα που προέκυψαν ήταν αρκετά ενθαρρυντικά, διότι για το πρώτο σύνολο δεδομένων η ακρίβεια προσέγγισε την μέγιστη τιμή (100%), ενώ για το δεύτερό σημειώθηκε ακρίβεια στο 94%, καθιστώντας αυτή την τεχνική εξίσου αποτελεσματική συγκριτικά με άλλες παραδοσιακές προσεγγίσεις. Επιπλέον, παρατηρήθηκε και υψηλή ενεργειακή απόδοση, ενισχύοντας την χρησιμότητα αυτού του μοντέλου ταξινόμησης για εφαρμογές τέτοιου τύπου.
Παρόλο που η χρήση των SNN για την ανάλυση ομιλίας σε ασθενείς με ΝΠ δεν είναι τόσο διαδεδομένη, συγκριτικά με άλλες μεθόδους νευρωνικών δικτύων, οι μέχρι τώρα μελέτες στον τομέα δείχνουν ότι το μοντέλο αυτό παρουσιάζει μια εξίσου υψηλή απόδοση στην διάγνωση της νόσου [52]. Η σημαντική ταξινομητική ικανότητα του SEFRON ως προς την ακρίβεια και ενεργειακή απόδοση, επιβεβαιώνουν την θετική εικόνα των SNN για το παρόν ερευνητικό αντικείμενο, συμβάλλοντας στην ανάπτυξη καινοτόμων πρακτικών στον ιατρικό τομέα.

2.9.2 Ο Αλγόριθμος Logistic Regression (LR) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας

Πέρα από τα SNN ως μοντέλο ταξινόμησης, αξίζει να εξετασθούν και άλλες απλούστερες προσεγγίσεις που έχουν εφαρμοστεί για την ανίχνευση της ΝΠ. Ένα χαρακτηριστικό παράδειγμα μοντέλου που έχει χρησιμοποιηθεί ουκ ολίγες φορές σε τέτοιες εφαρμογές αποτελεί το Logistic Regression μοντέλο (LR), που αποτελεί έναν αλγόριθμο δυαδικής ταξινόμησης [57]. Σαν μοντέλο κάνει χρήση μιας λογιστικής συνάρτησης (sigmoid), με σκοπό να προβλέψει την πιθανότητα να συμβεί ένα από τα δύο ενδεχόμενα που εξετάζονται. Πρόκειται για μια απλοϊκή μέθοδο η οποία είναι χρήσιμη και αποτελεσματική στην περίπτωση που τα δεδομένα εμφανίζουν σαφή διαχωρισμό μεταξύ των εκάστοτε κατηγοριών και μπορεί να χρησιμοποιηθεί ως βάση για την αξιολόγηση πιο σύνθετων μεθόδων ταξινόμησης [58], [59].



Εικόνα 7:Οριοθέτηση απόφασης σε γραμμικά και μη γραμμικά διαχωρίσιμα δεδομένα. Η Logistic Regression αποδίδει καλύτερα σε περιπτώσεις γραμμικής διαχωρισιμότητας (αριστερά), ενώ σε πιο σύνθετα δεδομένα (δεξιά) απαιτούνται πιο εξελιγμένοι ταξινομητές. [60]

Σε πρόσφατες μελέτες, και πιο συγκεκριμένα στην μελέτη των Hossain και Amenta *[61]*, έγινε χρήση και αξιολόγηση του μοντέλου LR σε σύνολο δεδομένων που αποτελούνταν από 252 ασθενείς και 754 χαρακτηριστικά ομιλίας. Το ποσοστό ακρίβειας της μεθόδου έφτασε στο

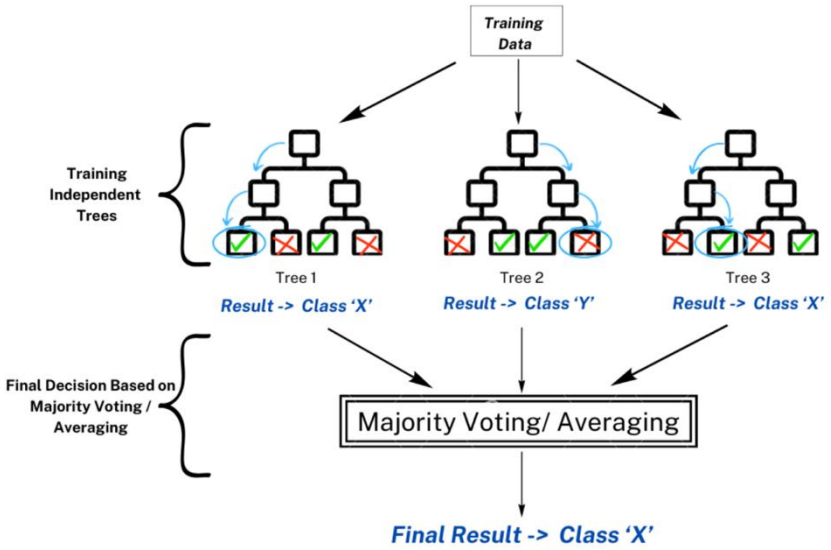
81,14% και όταν έγινε συνδυασμός της με LinearSVC για την επιλογή των σημαντικών χαρακτηριστικών , η ακρίβεια μειώθηκε ελάχιστα προσεγγίζοντας το 78,95%. Τα κύρια χαρακτηριστικά που αξιοποιήθηκαν ήταν τα: jitter, shimmer, ένταση φωνής, συχνότητες φωνηέντων και MFCCs. Σε παρόμοιο πλαίσιο κινήθηκαν και στην έρευνα τους οι Dhanalakshmi et al. *[62]*, οι οποίοι αξιολόγησαν την LR μαζί με άλλους ταξινομητές σε ένα μη ισορροπημένο σύνολο δεδομένων φωνητικών χαρακτηριστικών που αφορούσαν την ΝΠ. Προκειμένου να αντιμετωπισθεί το ζήτημα των μη ισορροπημένων δεδομένων, έγινε εφαρμογή της τεχνικής SMOTE-ENN (Synthetic Minority Over-sampling Technique - Edited Nearest Neighbors), καθώς και της PCA για την μείωση της διαστατικότητας. Η μελέτη αυτή έδειξε ότι η απόδοση της LR υστερούσε συγκριτικά με την μέθοδο SVM,η οποία εμφάνισε ακρίβεια της τάξεως του 96,5%. Ακόλουθα, στην μελέτη των Ratnakar Anand (International Journal of Advanced Research) *[63]*, που είχε ως σκοπό την αξιολόγηση των διάφορων τεχνικών ανάλυσης ομιλίας για την ανίχνευση της ΝΠ, έγινε σύγκριση τη LR με άλλους αλγορίθμους Μηχανικής και Βαθιάς Μάθησης, αναφέροντας ακρίβεια δοκιμής 84,62%, ένα αρκετό καλό ποσοστό. Μια ακόμα ενδιαφέρουσα προσέγγιση αποτελεί μια προδημοσίευση του 2024/2025 των Verma and Sinha *[64]*, η οποία είχε ως στόχο την διάκριση μεταξύ ατόμων με ΝΠ και υγιών ατόμων. Πρότεινε την χρήση μιας μεθόδου, βασισμένης στην LR, κάνοντας χρήση φωνητικών δεδομένων ομιλίας. Η έρευνα αυτή έδειξε ακρίβεια στην εκπαίδευση περίπου στο 88% και ακρίβεια δοκιμής στο 90%, ποσοστά που είναι αρκετά ενθαρρυντικά. Τέλος, σε μια πιλοτική μελέτη του 2025 των Oliveira et al. *[65]* , έγινε εφαρμογή της μεθόδου ταξινόμησης LR, με σκοπό την ανίχνευση της σοβαρότητας της ΝΠ. Αυτό επιτεύχθηκε μέσω αξιολόγησης της ομιλίας των ασθενών, εστιάζοντας σε ασκήσεις διαδοχοκινητικών φωνητικών δοκιμασιών (DDK) και σε χαρακτηριστικά φωνητικής, άρθρωσης και προσωδίας. Ο συνδυασμός όλων αυτών των χαρακτηριστικών με την χρήση LR, πέτυχε ακρίβεια 72%, γεγονός που υποδηλώνει ικανοποιητική απόδοση του μοντέλου.

Συγκεντρωτικά, σαν μέθοδος η LR παρουσίασε αρκετά ικανοποιητική απόδοση σε διαφορετικά σύνολα δεδομένων που αφορούσαν κυρίως την ταξινόμηση της ομιλίας στην ΝΠ, επικυρώνοντας την αξιοπιστία του μοντέλου. Επίσης, μπορεί να σημειωθεί και ότι η αποτελεσματικότητα της συγκεκριμένης μεθόδου μπορεί να επηρεαστεί από τις τεχνικές προεπεξεργασίας που θα χρησιμοποιηθούν σε μη κανονικοποιημένα σύνολα δεδομένων, αλλά και από μεθόδους επιλογής χαρακτηριστικών.

2.9.3 Ο Αλγόριθμος Random Forest (RF) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας

Στην συνέχεια αξίζει να γίνει αναφορά και σε πιο σύνθετους αλγορίθμους ταξινόμησης, οι οποίοι έχουν εφαρμοστεί εξίσου με επιτυχία στην ανίχνευση της ΝΠ, μέσω ανάλυσης ηχητικών δεδομένων. Μια ακόμη δημοφιλής μέθοδος ταξινόμησης αποτελεί η Random Forest (RF) [62]. Ουσιαστικά σε αυτόν τον αλγόριθμο γίνεται συνδυασμός δέντρων-αποφάσεων, τα οποία εκπαιδεύονται σε διαφορετικά υποσύνολα των διαθέσιμων εξεταζόμενων δεδομένων. Η τελική απόφαση προκύπτει από την συχνότητα της εμφάνισης ενός αποτελέσματος μεταξύ του συνόλου των δένδρων, με το αποτέλεσμα που εμφανίζεται συχνότερα να επικρατεί. Αυτός ο τρόπος λειτουργίας αποτρέπει την υπερεκπαίδευση (overfitting) και βελτιώνει την συνολική απόδοση του μοντέλου, καθιστώντας το RF ένα αξιόπιστο μοντέλο και ειδικά όταν υπάρχει ένα μεγάλο σύνολο δεδομένων [66] . Στην μελέτη των Dhanalakshmi et al. [62] που αναφέρθηκε και παραπάνω, το μοντέλο ταξινόμησης RF, παρουσίασε και αυτό όπως το μοντέλο LR, χαμηλότερη απόδοση σχετικά με το SVM σε μη ισορροπημένα δεδομένα ομιλίας. Αντίστοιχα, σε μια ανώνυμη μελέτη του 2024 αναφέρθηκε ακρίβεια του μοντέλου της τάξεως του 95%, προσεγγίζοντας το ποσοστό επιτυχίας του αλγορίθμου SVM.

Τώρα, στις μελέτες των Hossain και Amenta [61] & Ratnakar Anand (International Journal of Advanced Research) [63] που και οι δύο ασχολήθηκαν με την ανίχνευση της ΝΠ, η χρήση του μοντέλου RF, παρουσίασε ακρίβεια 89,74% και 80,26% αντίστοιχα, αναδεικνύοντας την αξιοπιστία του αλγορίθμου στην ταξινόμηση δεδομένων ομιλίας σχετιζόμενων με τη Νόσο του Πάρκινσον. Ακόλουθα, αξίζει να γίνει αναφορά και στην μελέτη των Jeong et al. [67], όπου έκαναν χρήση του συνόλου ακουστικών δεδομένων Parkinson's Disease Classification Dataset, εφαρμόζοντας τον αλγόριθμο ταξινόμησης RF με σκοπό την ανίχνευση της ΝΠ. Και εδώ η απόδοση εκπαίδευσης προσέγγισε το 88,16% και για δοκιμή 88,09%. Σε παρόμοιο πλαίσιο μελέτης κινήθηκε και μια παρεμφερή έρευνα που αξιοποίησε χαρακτηριστικά MFCCs σε συνδυασμό με το μοντέλο RF, με σκοπό την ταξινόμηση της ΝΠ. Έγινε αναλυτική περιγραφή της διαδικασίας συλλογής δεδομένων , της εξαγωγής των ακουστικών χαρακτηριστικών καθώς και της διαδικασίας εκπαίδευσης και αξιολόγησης του μοντέλου, εμφανίζοντας και πάλι ικανοποιητικά αποτελέσματα [68]. Η χρήση του RF αξιολογήθηκε και στην έρευνα των Baranwal et al. [69], που είχαν ως στόχο την πρόβλεψη της σοβαρότητας της ΝΠ στα άτομα 3ης ηλικίας. Και σε αυτή την περίπτωση το ποσοστό επιτυχίας κρίθηκε ικανοποιητικό. Υψηλό ποσοστό επιτυχίας σημειώθηκε εξίσου και σε άλλες πρόσφατες μελέτες. Στην μία από αυτές έγινε αξιολόγηση του μοντέλου RF, εκπαιδευμένο σε δεδομένα από την δημόσια πλατφόρμα Kaggle που περιλαμβάνει χαρακτηριστικά ομιλίας ασθενών με ΝΠ, σημειώνοντας ακρίβεια 91.83%, που θεωρείται αρκετά υψηλή [69]. Επίσης, σε ένα αντίστοιχο πείραμα που κινήθηκε στο ίδιο πλαίσιο, έγινε αξιολόγησε σε δεδομένα ομιλίας από 40 ασθενείς με ΝΠ, επιτυγχάνοντας απόδοση 72,50%, χαμηλότερη από παρόμοιες μελέτες αλλά και πάλι αρκετά ικανοποιητική [70]. Αντίστοιχη ικανοποιητική ακρίβεια (91,83%) παρουσίασε και μια ακόμα πρόσφατη μελέτη που αξιοποίησε τα σύνολα δεδομένων PPMI και UCI [71]. Τέλος, μια ακόμη έρευνα, που κρίνεται άξια αναφοράς, προκειμένου να αναγνωρίσει γονιδιοματικούς βιοδείκτες που συσχετίζονται με την ΝΠ, έκανε εφαρμογή του μοντέλου RF, σημειώνοντας και πάλι ένα καλό ποσοστό απόδοσης [72].



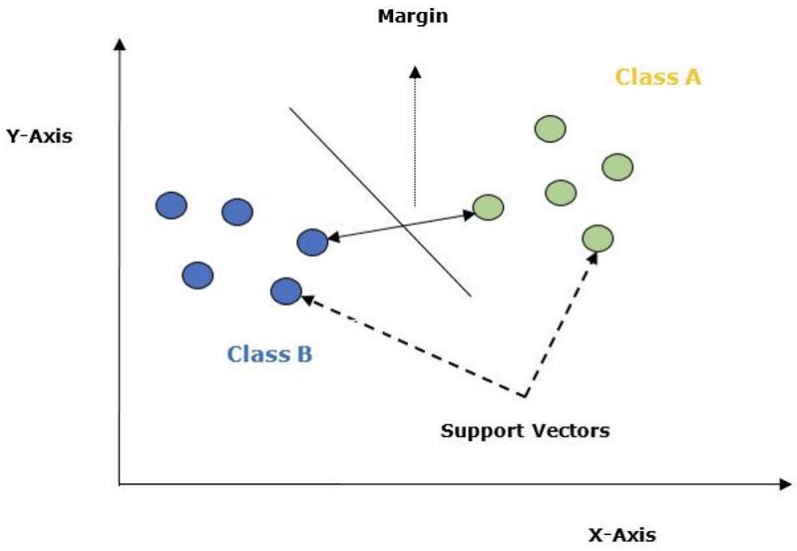
Εικόνα 8:Αρχή λειτουργίας του Random Forest. Πολλαπλά decision trees εκπαιδεύονται ανεξάρτητα σε υποσύνολα των δεδομένων και η τελική απόφαση λαμβάνεται μέσω πλειοψηφικής ψήφου, ενισχύοντας τη σταθερότητα και την ακρίβεια του μοντέλου [73].

Συνοψίζοντας, το μοντέλο ταξινόμησης Random Forest, αποτελεί και αυτό ένα αξιόπιστο αλγόριθμο ταξινόμησης, που χρησιμοποιείται στην ανάλυση ομιλίας στην ΝΠ, με τα ποσοστά απόδοσης και ακρίβειας να παρουσιάζονται συχνά υψηλά. Σαν μοντέλο είναι σε θέση να

διαχειρίζεται με ευκολία μεγάλα σύνολα δεδομένων, καθώς και να εντοπίζουν την σημαντικότητα των χαρακτηριστικών, στοιχεία που το μετατρέπουν σε ένα πολύτιμο εργαλείο στον εντοπισμό φωνητικών δεικτών, που αφορούν την παρουσία ή απουσία της ΝΠ, καθώς και την σοβαρότητα αυτής.

2.9.4 Ο Αλγόριθμος Support Vector Machine (SVM) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας

Ένα ακόμα μοντέλο που αξίζει να αναφερθεί είναι το Support Vector Machine (SVM). Πρόκειται για έναν πολύτιμο αλγόριθμο ταξινόμησης, με υψηλό ποσοστό επιτυχίας σε δυαδικά προβλήματα. Η βασική ιδέα του μοντέλου είναι η εύρεση ενός υπερεπιπέδου (hyperplane) , που διαχωρίζει αποτελεσματικά τις κλάσεις των δεδομένων με όσο το δυνατόν μεγαλύτερη απόσταση μεταξύ τους. Έτσι το μοντέλο είναι σε θέση να αποδώσει καλύτερα σε νέα άγνωστα δεδομένα [74]. Ένα ακόμη πλεονέκτημα του συγκεκριμένου μοντέλου είναι ότι μπορεί να επιφέρει υψηλό αποτέλεσμα ακόμα και σε δεδομένα που ο διαχωρισμός τους μέσω ειδικών μαθηματικών τεχνικών (Πυρηνικές Συναρτήσεις – Kernels) για απλοποίηση της μορφής των δεδομένων δεν είναι εύκολη υπόθεση [75]. Όλα αυτά τα χαρακτηριστικά καθιστούν και αυτό το μοντέλο κατάλληλο για τεχνικές δυσκολίες στην ανάλυση ηχητικών δεδομένων που αφορούν ασθενείς με ΝΠ [76]. Η απόδοση αυτού του μοντέλου έχει αξιολογηθεί εκτενώς στο εν λόγω πεδίο, στο πλαίσιο πλήθους μελετών.



Εικόνα 9:Οπτική αναπαράσταση του αλγορίθμου Support Vector Machine. Το μοντέλο δημιουργεί ένα υπερεπίπεδο που διαχωρίζει τις κλάσεις με το μέγιστο δυνατό περιθώριο, βασιζόμενο σε σημεία-κλειδιά (support vectors) [77].

Αρκετές μελέτες έχουν στρέψει το ενδιαφέρον της έρευνας τους προς τον αλγόριθμο ταξινόμησης Surrort Vector Machine (SVM), σε εφαρμογές που σχετίζονται στην ανάλυση ομιλίας σε ασθενείς με ΝΠ. Πιο συγκεκριμένα, στην έρευνα των Dhanalakshmi et al. [62] αξιολόγησαν τον αλγόριθμο SVM σε πρόβλημα ταξινόμησης ΝΠ και ο συνδυασμός του μοντέλου αυτού με την τεχνική SMOTE-END, έδειξε πολύ υψηλά ποσοστά ακρίβειας (96,5%). Σε παρόμοιο πλαίσιο στην έρευνα των Abdullah H. Al Nefaie et al. [74], η χρήση του SVM ανέφερε ακρίβεια 95%, πλησιάζοντας πολύ τα αποτελέσματα για Random Forest. Αξιοσημείωτα είναι και τα ποσοστά ακρίβειας του SVM μοντέλου που εμφανίζονται στην έρευνα των Ijaz et al. [63], φτάνοντας στο 87,18%. Αντίθετα, στην έρευνα των Hossain και

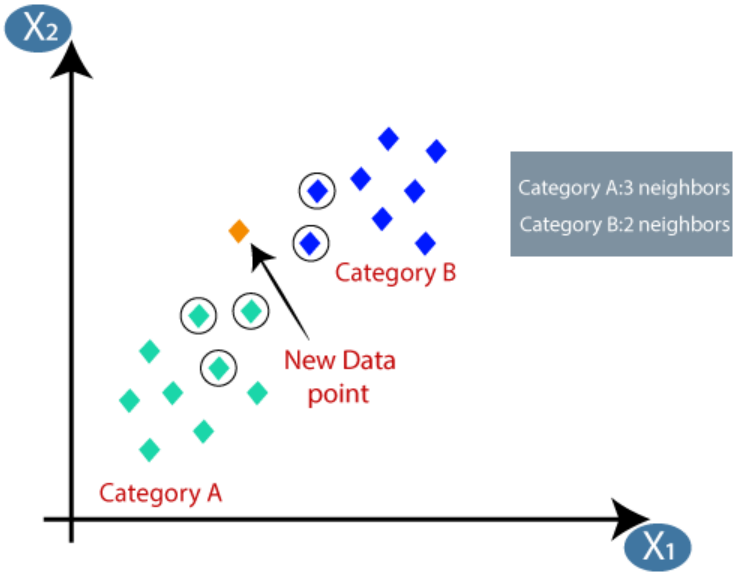
Amenta [61], όπου το μοντέλο δεν συνδικάστηκε με κάποια άλλη τεχνική, η ακρίβεια που σημειώθηκε ήταν χαμηλότερης τάξης (76,32%) αλλά εξίσου σε υψηλά επίπεδα. Τώρα η εφαρμογή του αλγορίθμου σε πιο στοχευμένα σύνολα δεδομένων, όπως στο σύνολο δεδομένων ομιλίας UCI, έδειξε ποσοστό επιτυχίας 91,25% [78], ωστόσο η εφαρμογή στο σύνολο δεδομένων Kaggle, έδειξε ότι το μοντέλο υστερούσε συγκριτικά με το άλλο εξεταζόμενο μοντέλο της έρευνας, το RF [69]. Προς την ίδια κατεύθυνση κινήθηκε και η έρευνα των Devi et al. οι οποίοι προσπάθησαν να βελτιώσουν την απόδοση του αλγορίθμου SVM σε συνδυασμό με την τεχνική PCA, προκειμένου να μειωθεί η διαστατικότητα. Τα αποτελέσματα του προβλήματος ταξινόμησης πολλαπλών κατηγοριών για την ανίχνευση της ΝΠ έδειξαν ενθαρρυντικά ποσοστά απόδοσης αγγίζοντας το 89%. Το ποσοστό της ακρίβειας αυξήθηκε στο 94,1%, όταν η εφαρμογή επικεντρώθηκε σε 34 επιλεγμένα χαρακτηριστικά ομιλίας, αποδεικνύοντας ότι πρωταρχικό ρόλο στην διακριτική ικανότητα του μοντέλου έχει η στρατηγική επιλογή των κατάλληλων χαρακτηριστικών (ποιότητα χαρακτηριστικών > ποσότητα χαρακτηριστικών) [79]. Επιπλέον, μια ακόμη σημαντική παράμετρος για την ακρίβεια του μοντέλου, αποτελεί η ορθή ρύθμιση των παραμέτρων του, προκειμένου να πετύχει στο μέγιστο τον εκάστοτε στόχο της έρευνας [79]. Αυτό επιβεβαιώνεται και από την μελέτη των Raya Alshammri et al. [80], όπου έγινε συνδυασμός του GridSearchCV με την τεχνική SMOTE. Το GridSearchCV είναι μια τεχνική επιλογής υπερπαραμέτρων, όπου δοκιμάζονται όλοι οι πιθανοί συνδυασμοί ενός πλέγματος τιμών και ο κάθε συνδυασμός αξιολογείται μέσω διασταυρούμενης επικύρωσης (Cross-Validation). Ο συνδυασμός με την καλύτερη απόδοση σε όλα τα folds είναι και αυτός που τελικά επιλέγεται [81]. Στην συγκεκριμένη μελέτη η αναλογία εκπαίδευσης/ δοκιμής ήταν στο 70/30 και το μοντέλο SVM απέδωσε με συνολική ακρίβεια 95%, που θεωρείται ένα αξιοσημείωτο ποσοστό απόδοσης. Σε μια ακόμη έρευνα, με στόχο την αξιολόγηση της διακριτικής ικανότητας ταξινόμησης στην ΝΠ έναντι του τρόμου, με βάση τα χαρακτηριστικά ομιλίας, έγινε χρήση στο αλγορίθμου SVM σε συνδυασμό με Gaussian πυρήνα. Και σε αυτή την μελέτη παρουσιάστηκαν ικανοποιητικά αποτελέσματα, ανεξάρτητα του αριθμού των κατηγοριών υπόκεινται στην ταξινόμηση [82]. Επίσης, η αξιοπιστία του μοντέλου επικυρώθηκε και από μια ακόμη πρόσφατη έρευνα που ανέδειξε ότι το SVM είναι σε θέση να διακρίνει με επιτυχία και άλλες φωνητικές νόσους, πέρα από την ΝΠ. Αυτό το εύρημα ενισχύει την αποτελεσματικότητα και την διακριτική ικανότητα του συγκεκριμένου μοντέλου στην ανάλυση φωνητικών διαταραχών, ανεξάρτητα αν αφορούν την ΝΠ [83].

Με γνώμονα όλο αυτό το ερευνητικό έργο που περιστρέφεται γύρω από τον αλγόριθμο SVM και τις εφαρμογές αυτού σε διάφορα ερευνητικά πλαίσια και πιο συγκεκριμένα στην ανίχνευση της ΝΠ, το μοντέλο αυτό αποδεικνύεται ως μια αποδοτική επιλογή ως προς την ανάλυση φωνητικών δεδομένων που συσχετίζονται με την ΝΠ. Η ακρίβεια του μοντέλου, στο πλήθος των περιπτώσεων παρουσιάζεται αρκετά ικανοποιητική, με ποσοστά να προσεγγίζουν πολλές φορές και το 90%. Πολύ ικανοποιητική ήταν και η συνολική απόδοση του μοντέλο, ειδικά όταν συνδυάστηκε και με άλλες τεχνικές προεπεξεργασίας. Τα ευρήματα που προέκυψαν μέσω της βιβλιογραφικής ανασκόπησης, καταδεικνύουν ότι η εφαρμογή του μοντέλου SVM μπορεί να φανεί εξίσου αποδοτική ή ακόμα και να ξεπερνά την απόδοση άλλων καθιερωμένων μεθόδων ταξινόμησης.

2.9.5 Ο Αλγόριθμος K-Nearest Neighbors (K-NN) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας

Έπειτα από την ανάλυση των πιο σύνθετων μοντέλων ταξινόμησης, αξίζει να εξετασθεί και ένας πιο απλός αλγόριθμος αλλά εξίσου σημαντικός στις εφαρμογές ταξινόμησης. Αυτός ο αλγόριθμος είναι ο K-Nearest Neighbors (K-NN) [74]. Το μοντέλο αυτό ουσιαστικά δεν εκπαιδεύεται από τα δεδομένα στην φάση της εκπαίδευσης, αλλά οι αποφάσεις που θα πάρει βασίζονται στα πλησιέστερα σημεία όταν του δίνεται ένα νέο δείγμα. Πιο αναλυτικά, όταν

δοθεί στο μοντέλο ένα νέο δείγμα, έστω ένα ηχητικό σήμα, το K-NN εστιάζει στους k πλησιέστερους γείτονες στο πλαίσιο των χαρακτηριστικών και το ταξινομεί στην κατηγορία στην οποία εμφανίζεται πιο συχνά. Σαν αλγόριθμος ταξινόμησης θεωρείται αρκετά απλός, τόσο στην υλοποίηση του όσο και στην κατανόηση του τρόπου λειτουργίας του [84]. Ωστόσο, το μοντέλο K-NN, αν το πλήθος των δεδομένων που χρησιμοποιεί είναι ογκώδες, μπορεί να μετατραπεί σε μια υπολογιστικά δαπανηρή διαδικασία. Επίσης, η απόδοση του συνδέεται αλληλένδετα με την ποιότητα των χαρακτηριστικών αλλά και στην στρατηγική επιλογή του k. Τώρα, όσον αφορά την χρήση αυτού του μοντέλου στην ανάλυση ομιλίας στην ΝΠ, χρησιμοποιείται ως μια βασική μέθοδο για τις διάφορες εφαρμογές και αυτό κυρίως για να γίνει η σύγκριση της συνολικής απόδοσης της μεθόδου έναντι άλλων πιο σύνθετων μοντέλων, για τον εντοπισμό της μέγιστης ακρίβειας στην εκάστοτε εφαρμογή [74], [85].



Εικόνα 10: Παράδειγμα λειτουργίας του αλγορίθμου k-Nearest Neighbors (k-NN). Το νέο δείγμα (πορτοκαλί σημείο) ταξινομείται στην κατηγορία Α, καθώς οι περισσότεροι από τους k πλησιέστερους γείτονες ανήκουν σε αυτήν [86].

Πέρα από την απλότητα του αλγορίθμου, το μοντέλο KNN έχει μελετηθεί εκτενώς στην ανάλυση φωνητικών δειγμάτων που επικεντρώνεται στην ΝΠ και λόγω της φύσης του συγκεκριμένου μοντέλου , μπορεί να εφαρμοστεί σε διαφορετικά σύνολα δεδομένων. Πιο συγκεκριμένα στην μελέτη των Abdullah H. Al Nefaie et al. [74], που είχε ως σκοπό την πρόβλεψη της εμφάνισης της ΝΠ με την χρήση διάφορων αλγορίθμων, το μοντέλο KNN πέτυχε ακρίβεια 92%. Σε παρόμοιο πλαίσιο κινήθηκε και η μελέτη των Dhanalakshmi et al. [62], που είχε ως κύριο στόχο την ταξινόμηση της ΝΠ χρησιμοποιώντας χαρακτηριστικά ομιλίας και συνδυασμό των τεχνικών SMOTE-ENN για αντιμετώπιση της ανισορροπίας και PCA για μείωση διαστατικότητα. Εδώ ο KNN πέτυχε χαμηλότερη ακρίβεια (δεν προσδιορίζεται) σε σύγκριση με την SVM που έδειξε 96,5% με την χρήση SMOTE-ENN. Σε μία ακόμη μελέτη, αυτή των Ratnakar Anand (International Journal of Advanced Research) [63], πέτυχαν ακρίβεια δοκιμής 82,05% με την χρήση του μοντέλου KNN. Η έρευνα αυτή αξιολόγησε διάφορες τεχνικές ανάλυσης ομιλίας για την ανίχνευση της ΝΠ. Από τη άλλη η μελέτη των Hossain και Amenta [61] που είχαν ως στόχο την ταξινόμηση ατόμων με ΝΠ και υγιών ατόμων με βάση ηχογραφήσεων φωνής σε διαφορετικές συχνότητες, ανέφεραν ακρίβεια δοκιμής 62,28% για το KNN, ποσοστό αρκετά χαμηλότερο από άλλες μελέτες. Σε μια ακόμη ενδιαφέρουσα έρευνα, αυτή των Mrs.A.G. Phakatkar et al. [78], η οποία εξέταζε την διάκριση

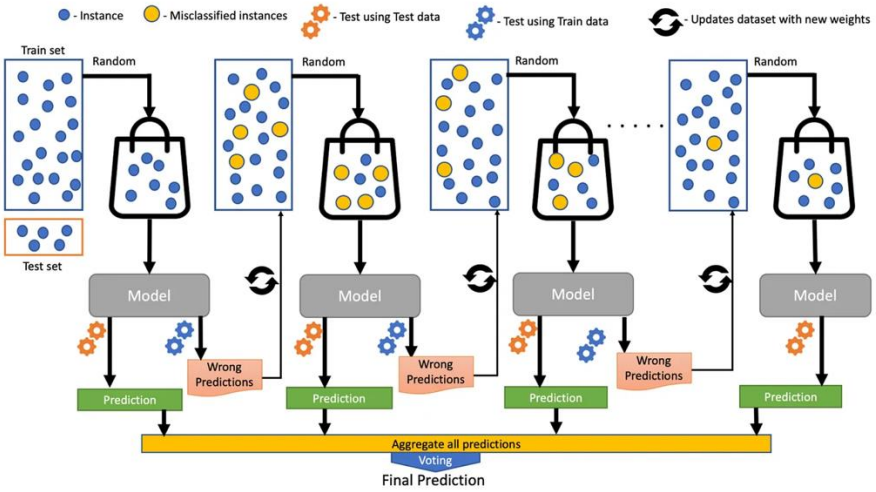
μεταξύ των διαφόρων διαταραχών φωνής, πέτυχε ακρίβεια 91,23% για το αλγόριθμο KNN στο σύνολο δεδομένων ομιλίας UCI, ποσοστό που πλησίαζε την απόδοση της SVM. Μια άλλη μελέτη του 2024 που είχε σαν κύριο μέλημα την διάγνωση της ΝΠ χρησιμοποιώντας ακουστικές μετρήσεις ομιλίας, έδειξε χαμηλότερη ακρίβεια για το KNN σε σύγκριση με το Random Forest σε πρώιμα αποτελέσματα που ήταν βασισμένα σε δεδομένα του συνόλου Kaggle [69]. Αντίθετα, μια μελέτη του 2022 των Boualoulou et al. [87], που ο κύριος πυλώνας της έρευνας ήταν η αυτόματη αναγνώριση του Πάρκινσον μέσω ανάλυσης φωνής, χρησιμοποίησε το KNN με Empirical Mode Decomposition (EMD) και Mel-Frequency Cepstral Coefficients (MFCCs) και παρουσίασε ακρίβεια της τάξης του 86,67%, που είναι ένα υψηλό ποσοστό. Τώρα, στην μελέτη των Arti Rana et al. [88], που ο στόχος ήταν η ανίχνευση της ΝΠ μέσω συγκριτικής ανάλυσης αλγορίθμων, πέτυχε ακρίβεια 87,17% για το KNN, παρόμοιο ποσοστό με το SVM. Τέλος, στην μελέτη Bali Devi et al., το μοντέλο KNN πέτυχε ακρίβεια 80% στα δεδομένα εκπαίδευσης και 82,14% στα δεδομένα δοκιμής, έπειτα από την εξάλειψη των χαρακτηριστικών που ήταν άμεσα εξαρτώμενα από το φύλο του ασθενή. Ο στόχος αυτής της μελέτης ήταν η ανάπτυξη ενός συστήματος ταξινόμησης για την νόσο του Πάρκινσον, βασιζόμενο στα τυπικά συμπτώματα της ΝΠ [79].

Η απόδοση του αλγορίθμου k-Nearest Neighbors (KNN) στην ταξινόμηση της ομιλίας στην Νόσο του Πάρκινσον (ΝΠ) παρουσιάζει σημαντική διακύμανση, με την συνολική ακρίβεια αυτού να κρίνεται μέτρια προς υψηλή. Η αποτελεσματικότητα του μοντέλου βασίζεται κυρίως στην στον τύπο και στην ποιότητα των χαρακτηριστικών, στον τρόπο προεπεξεργασίας των φωνητικών δεδομένων, καθώς και στην επιλογή του αριθμού των πλησιέστερων γειτόνων. Επιπλέον, η εφαρμογή του αλγορίθμου, σε συνδυασμό με άλλες εξειδίκευες τεχνικές εξαγωγής/αναπαράστασής χαρακτηριστικών, φαίνεται να βελτιώνει σε σημαντικό βαθμό την συνολική απόδοση της ταξινόμησης. Όλα τα παραπάνω δείχνουν πόσο σημαντικό είναι να ληφθούν υπόψιν οι παραπάνω παράγοντες στις εφαρμογές που γίνεται χρήση του KNN μοντέλου, για την ανάλυση ομιλίας σε άτομα με ΝΠ, με σκοπό να μεγιστοποιηθεί η αξιοπιστία και η ακρίβεια των αποτελεσμάτων.

2.9.6 Ο Αλγόριθμος Extreme Gradient Boosting (XGBoost) στην Ανίχνευση της Νόσου του Πάρκινσον Μέσω Ομιλίας

Ο αλγόριθμος Extreme Gradient Boosting (XGBoost) αποτελεί μια γνωστή τεχνική με υψηλή απόδοση και υψηλά ποσοστά αποτελεσματικότητας σε προβλήματα ταξινόμησης [89]. Ουσιαστικά αυτό που κάνει ο αλγόριθμος, είναι να σχηματίζει διαδοχικά αδύναμους ταξινομητές, σαν δένδρα αποφάσεων, που το κάθε δέντρο εκπαιδεύεται βασιζόμενο σε λάθη που είχαν προκύψει στο προηγούμενο. Με αυτόν τον τρόπο εκπαίδευσης η απόδοση του αλγορίθμου σταδιακά ενισχύεται και τα σφάλματα μειώνονται λόγω του παραγώγου της συνάρτησης κόστους (gradient) [82].
Ο συγκεκριμένος αλγόριθμος χρησιμοποιείται ευρέως σε εφαρμογές που αφορούν την ανάλυση βιοιατρικών δεδομένων [90]. Έτσι, γίνεται χρήση αυτού και στην ανάλυση φωνητικών δεδομένων από την ομιλία ατόμων που νοσούν με την ΝΠ, λόγω της ικανότητας του να διαχειρίζεται αποτελεσματικά πολύπλοκα ιατρικά δεδομένα με υψηλή απόδοση, καθώς και της ορθής διαχείρισης ελλιπών τιμών σε μεγάλα σύνολα δεδομένων[82], [90]. Η συχνή χρήση του αλγορίθμου σε πιο πρόσφατες μελέτες καθιστούν το συγκριμένο μοντέλο ένα αξιόπιστο εργαλείο σε τέτοιου είδους προβλήματα ακόμα και αν θεωρείται μια πιο παλιότερη μέθοδος συγκριτικά με νεότερα μοντέλα βαθιάς μάθησης [91].
Πιο πρόσφατες μελέτες δείχνουν ότι ο XGBoost, σε συνδυασμό με χαρακτηριστικά ομιλίας για την ενίσχυση της ΝΠ παρουσιάζει υψηλή ευελιξία [90], [92]. Ωστόσο η επιλογή των κατάλληλων χαρακτηριστικών και η διαδικασία της προεπεξεργασίας των δεδομένων

αποτελούν κομβικά σημεία για την επιτυχία της ταξινόμησης [90]. Στον αλγόριθμο αυτό οι τεχνικές που χρησιμοποιούνται για τον εντοπισμό και την τελική επιλογή των σημαντικότερων χαρακτηριστικών, αποτελεί η Αναδρομική Εξάλειψη Χαρακτηριστικών (RFE) και η ανάλυση της σημασίας αυτών μέσω π.χ. Shapley τιμών [90]. Σχετικά με την προεπεξεργασία και την κανονικοποίηση των χαρακτηριστικών, καθώς και την εξάλειψη της ανισορροπίας χρησιμοποιείται η τεχνική SMOTE (Synthetic Minority Over-sampling Technique), βελτιώνοντας την συνολική απόδοση του μοντέλου [92]. Όλα αυτά τα στοιχεία δείχνουν ότι αποτελεί έναν ισχυρό αλγόριθμο για την ανάλυση της ομιλίας στην ΝΠ διότι βασίζεται στην προσεκτική επιλογή των χαρακτηριστικών, σχετικών βιοδεικτών αλλά και της κατάλληλης προεπεξεργασίας των δεδομένων) [82], [90].



Εικόνα 11: Αρχή Λειτουργίας του Gradient Boosting: Διαδοχικά Μοντέλα με Εστίαση στα Σφάλματα [93].

Τα τελευταία χρόνια αρκετές μελέτες έχουν καταδείξει την αποτελεσματικότητα του αλγορίθμου XGBoost σε εφαρμογές ταξινόμησης μέσω ομιλίας που αφορούν νευροεκφυλιστικά νοσήματα, όπως είναι και η ΝΠ. Πιο συγκεκριμένα, η μελέτη των Adiga et al. [94] που πραγματεύθηκε την έγκαιρη μη επεμβατική διάγνωση της ΝΠ, με σκοπό την βελτίωση της διαγνωστικής διαδικασίας μέσω φωνητικών βιοδεικτών, κατέδειξε ακρίβεια του αλγορίθμου XGBoost της τάξεως του 98%, ένα πολύ υψηλό ποσοστό. Σε παρόμοιο πλαίσιο κινήθηκε και ο Yeboah [95], ο οποίος στην μελέτη του ασχολήθηκε με την έγκαιρη και ακριβή ανίχνευση της ΝΠ μέσω χαρακτηριστικών ομιλίας που αποτελούσαν βιοδείκτες της νόσου, με σκοπό τον εντοπισμό του καλύτερου ταξινομητικού μοντέλου. Διαπιστώθηκε ότι η ακρίβεια του XGBoost έφτασε στο 90-91% στο σύνολο δεδομένων UCI, υπερτερώντας από τα μοντέλα Logistic Regression, SVM,KNN καιGNB. Μια ακόμα ενδιαφέρουσα προσέγγιση αποτελεί η μελέτη του Santos [96] όπου είχε ως στόχο την αντιμετώπιση των δυσκολιών που αφορούν την έγκαιρη ανίχνευση της νόσου λόγω των ήπιων συμπτωμάτων στην αρχή της εμφάνισης της, δυσκολίες που προσπάθησαν να αντιμετωπιστούν με την εφαρμογή του αλγορίθμου XGBoost. Τα αποτελέσματα έδειξαν ποσοστό επιτυχίας 93.33% , F1-Score 95.65% και AUC 96.89%. Επιπλέον, οι Dudek et al. [97] πραγματοποίησαν ανάλυση βιοδεικτών ομιλίας με σκοπό την ανίχνευση της ΝΠ και την στατιστική αξιολόγηση αυτών των βιοδεικτών μέσω δεδομένων που συλλέχθηκαν σε μικτής πραγματικότητας περιβάλλον (MR). Και σε αυτή την μελέτη ο XGBoost εμφάνισε υψηλή απόδοση ειδικά για την μετρική F1-Score (0.90).
Μια ακόμη πρόσφατη μελέτη, αυτή των Van Gelderen Lisanne and Tejedor-García Cristian [98], πραγματεύτηκε την σημαντικότητα της πολύγλωσσής αξιολόγησης των βιοδεικτών που

προκύπτουν από ακουστικά και γλωσσικά δεδομένα για την ΝΠ, προκειμένου να αναπτυχθούν νέα εργαλεία, χρήσιμα για την διαγνωστική διαδικασία ανεξάρτητα από την γλώσσα των ασθενών. Σε αυτό το πλαίσιο, τα μη ερμηνεύσιμα μοντέλα (NIFMs) που κάνουν χρήση του μοντέλου βαθιάς μάθησης Wav2Vec2.0 υπερτερούσαν κατά 4% συγκριτικά με τα ερμηνευτικά μοντέλα χαρακτηριστικών (IFMs), στα οποία περιλαμβανόταν και ο XGBoost. Το μικρό αυτό ποσοστό διαφοράς μεταξύ των δυο διαφορετικών προσεγγίσεων επιβεβαιώνουν την ικανοποιητική απόδοση του αλγορίθμου XGBoost σε προβλήματα ανίχνευσης της ΝΠ μέσω φωνητικών δεδομένων.
Παρά την υπεροχή του XGBoost σε πληθώρα μελετών, σε ορισμένες εργασίες εντοπίζονται προκλήσεις ως προς την αξιοπιστία του αλγορίθμου, ειδικά σε εξειδικευμένα ερευνητικά πλαίσια, όπως στην μελέτη των Mattisson & Samuelsson Forsdik [34] στο πλαίσιο της διπλωματικής τους εργασίας. Η εργασία αυτή πραγματευόταν την χρήση εκφώνησης φωνηέντων ως βιοδείκτες για την ανίχνευση της ΝΠ, προκειμένου να δημιουργηθεί ένα προγνωστικό μοντέλο που κάνει χρήση ακουστικής ανάλυσης και μοντέλων ταξινόμησης. Τα αποτελέσματα έδειξαν υψηλή διακύμανση στο στάδιο της διασταυρούμενης επικύρωσης για το μοντέλο XGBoost, ενισχύοντας της ανησυχίες σχετικά με την αξιοπιστία του αλγορίθμου εντός συγκεκριμένων ερευνητικών πλαισίων. Οι προκλήσεις αυτές ενδέχεται να πηγάζουν από την φύση του μοντέλου, το οποίο εξαρτάται σε μεγάλό βαθμό από τις κατανομές των χαρακτηριστικών κατά την διάρκεια της εκπαίδευσης.
Ο XGBoost αποτελεί έναν ισχυρό αλγόριθμο, χρήσιμο για την ανίχνευση της ΝΠ μέσω ακουστικών δεδομένων ομιλίας , προσφέροντας υψηλή απόδοση. Καθιστά ιδανικό σημείο αναφορά λόγω της ευελιξίας του στην μεταχείριση των χαρακτηριστικών και την αξιόπιστη προβλεπτική του ικανότατα και στα πλαίσια μελετών που σχετίζονται με βαθιά μάθηση μπορεί με ευκολία να χρησιμοποιηθεί ως ταξινομητή σε χαρακτηριστικά ομιλίας.

Λαμβάνοντας υπόψιν την συμβολή του XGBoost στην ανίχνευση της ΝΠ μέσω δεδομένων ομιλίας και άλλων εφαρμογών που συσχετίζονται με την νόσο, αλλά και την βιβλιογραφικά τεκμηριωμένη ανάλυση των υπολοίπων ταξινομητικών μοντέλων, κρίνεται απαραίτητη η συγκριτική παρουσίαση αυτών. Στον παρακάτω πίνακα (Πίνακας 5) παρατίθενται συνοπτικά τα βασικά χαρακτηριστικά, ο σκοπός και η απόδοση των μοντέλων, παρέχοντας μια πιο ολοκληρωμένη εικόνα ως προς την χρησιμότητα τους στην ταξινόμηση φωνητικών δεδομένων ομιλίας από ασθενείς με ΝΠ.

Πίνακας 5: Συγκεντρωτικός πίνακας εφαρμογής των SNN, Logistic Regression, Random Forest, SVM,KNN και XGBoost στην ανάλυση ομιλίας για διάγνωση Νόσου του Πάρκινσον.

Τίτλος Μελέτης	Μοντέλα	Χαρακτηριστικά Ομιλίας	Ακρίβεια	Σύνολο Δεδομένων	Στόχος
A robust Parkinson's disease detection model based on time-varying synaptic efficacy function in spiking neural network (Das et al./2024) [52]	Spiking Neural Network (SEFRON)	Διάφορα ακουστικά χαρακτηριστικά (συχνότητα, jitter, shimmer, HNR, κ.α.)	100% (Dataset 1), 94% (Dataset 2)	UCI: Oxford Parkinson's Disease Detection Dataset, UCI: Parkinson Dataset with replicated acoustic features	Ανίχνευση ΝΠ

Machine Learning-Based Classification of Parkinson's Disease Patients Using Speech Biomarkers(Hossain and Amenta/2024) [57]	Logistic Regression, Random Forest, Support Vector Machine, K-Nearest Neighbors	754 χαρακτηριστικά φωνής (jitter, shimmer, ένταση, φωνήεντα, MFCCs κ.α.)	LR: 81.14%, RF: 80.26%, SVM: 76.32%, KNN: 62.28%	Δεν προσδιορίζεται	Ταξινόμηση ΝΠ έναντι υγιών ατόμων
Speech features-based Parkinson's disease classification using combined SMOTE-ENN and binary machine learning(Dhanalakshmi et al./2024) [62]	Logistic Regression, Random Forest, Support Vector Machine, K-Nearest Neighbours	Δεν προσδιορίζεται	SVM: 96.5% (με SMOTE-ENN)	Parkinson's disease speech features dataset (δημόσια διαθέσιμο)	Ταξινόμηση ΝΠ έναντι υγιών ατόμων
Evaluating speech analysis techniques for Parkinson's disease detection: A comparison of machine learning and deep learning algorithms (International Journal of Advanced Research/2024) [63]	Logistic Regression, Random Forest, Support Vector Machine, K-Nearest Neighbors	Δεν προσδιορίζεται	LR: 84.62%, RF: 89.74%, SVM: 87.18%, KNN: 82.05%	Δεν προσδιορίζεται	Ανίχνευση ΝΠ
Speech analysis for the detection of Parkinson's disease by combined use of empirical mode decomposition, Mel frequency cepstral coefficients and	K-Nearest Neighbor	Mel Frequency Cepstral Coefficients (MFCCs)	86.67% (με EMD)	Βάση δεδομένων 38 ηχογραφήσεων (20 με ΝΠ, 18 υγιείς)	Ανίχνευση ΝΠ

the K-nearest neighbor classifier(Boualoulou et al./2022) [87]					
Multiclass-Based Support Vector Machine for Parkinson's Disease Detection on Speech Data (Devi et al./2022) [79]	Support Vector Machine, Logistic Regression, Gaussian Naive Bayes, K-Nearest Neighbor's, Random Forest	Δεν προσδιορίζεται	SVM: 89% (με PCA), 94.1% (με 37 χαρακτηριστικά)	Δεν προσδιορίζεται	Ταξινόμηση ΝΠ
Parkinson's Disease Speech Classification Using 1D Convolutional Neural Networks(Yue and Wang/2024) [99]	Δεν αναφέρεται (χρησιμοποιήθηκε CNN)	Χαρακτηριστικά ομιλίας	100% και 92.86% (σε δύο σύνολα δεδομένων)	Δεν προσδιορίζεται	Ταξινόμηση ΝΠ
Speech Insight: Parkinson's Disease Diagnosis Using Random Forest Classifier (Venkatesan et al./2023) [71]	Random Forest	Δεν προσδιορίζεται	91.83%	PPMI και UCI	Διάγνωση ΝΠ
Early detection model of Parkinson's Disease using Random Forest Method on voice frequency data (Fahira et al./2023) [70]	Random Forest	26 τύποι ηχογραφήσεων φωνής (συχνότητα)	72.50%	Δεδομένα από 40 ασθενείς (20 με ΝΠ, 20 υγιείς)	Έγκαιρη ανίχνευση ΝΠ
Voice-based SVM Model Reliability for Identifying Parkinson's	Support Vector Machine	Jitter, Shimmer, SD(Pitch), HNR, NHR, Mean(F1)	Δεν προσδιορίζεται (εστίαση στην αξιοπιστία)	Δεδομένα από υγιείς, ασθενείς με ΝΠ, δυσφωνία	Ταξινόμηση ΝΠ έναντι άλλων παθολογικών καταστάσεων φωνής

Disease(Pah et al./2023) [83]				και λαρυγγίτιδα	
Machine learning-based classification of Parkinson's disease using acoustic features: Insights from multilingual speech tasks (Jeong et al./2024) [67]	K-Nearest Neighbor	Ακουστικά χαρακτηριστικά	84.73%	Δεδομένα από 100 Κορεάτες ασθενείς με ΝΠ και 100 υγιείς	Ταξινόμηση ΝΠ
Logistic Regression-Based Detection of Parkinson's Disease (Verma and Sinha/2025) [64]	Logistic Regression	Βασική συχνότητα (MDVP: Fo), jitter (MDVP: Jitter, MDVP: RAP), shimmer (MDVP: Shimmer, Shimmer: DDA), αναλογία θορύβου προς αρμονικές (NHR), RPDE, DFA, D2	~90% (ακρίβεια δοκιμής)	Δημόσια διαθέσιμο σύνολο δεδομένων	Ανίχνευση ΝΠ
Parkinson's disease detection from voice using artificial intelligence techniques: A review (Mujtaba H A et al./2024) [100]	Δεν αναφέρεται (ανασκόπηση)	Διάφορα ακουστικά χαρακτηριστικά (pitch, άρθρωση, ρυθμός ομιλίας)	Έως 99%	Διάφορα σύνολα δεδομένων	Ανίχνευση ΝΠ
A pilot study for speech assessment to detect the severity of Parkinson's disease: An ensemble approach	Logistic Regression (σε ensemble)	Χαρακτηριστικά φωνητικής, άρθρωσης, προσωδίας από εργασίες διαδοχοκινησίας (DDK)	72% (συνδυασμός χαρακτηριστικών με LR)	Μικρό δημόσιο σύνολο δεδομένων	Ανίχνευση σοβαρότητας ΝΠ

(Oliveira et al./ 2024) [65]					
Linguistic changes in spontaneous speech for detecting Parkinson's disease using large language models (Crawford J.L./2024) [101]	Δεν αναφέρεται (χρησιμοποιήθηκαν μεγάλα γλωσσικά μοντέλα)	Γλωσσικά χαρακτηριστικά (χρήση λέξεων, γραμματική, νόημα, πλαίσιο)	Έως 78%	Δεδομένα αυθόρμητης ομιλίας από ασθενείς με ΝΠ και υγιείς	Ανίχνευση ΝΠ
Comparative analysis of machine learning algorithms for accurate Parkinson's disease detection (Avula Venkata Sohan, Avula Venkata Meghana/2023) [102]	AdaBoostClassifier, GradientBoostingClassifier, KNeighborClassifier, LGBMClassifier, Random Forest Classifier	Δεν προσδιορίζεται	Random Forest: 90.26%, KNN: 82.5%, SVM: 88.9%, Logistic: 74.2%	Parkinson's Disease Voice Measurements dataset	Ανίχνευση ΝΠ
Machine learning-based classification of Parkinson's disease using acoustic features: Insights from multilingual speech tasks (Karim et al./2024) [72]	Random Forest, XGBoost, Light Gradient Boosting Machine Classifier, Extra Tree, Bagging	Γονιδιωματικά δεδομένα	Δεν προσδιορίζεται (ταξινόμηση βιοδεικτών)	Δεν προσδιορίζεται	Αναγνώριση βιοδεικτών ΝΠ
Classification of Parkinson's Disease Using Random Forest Classifier with 1D Patient Speech (Firdaus/2023) [103]	Random Forest	Διάφορα ακουστικά χαρακτηριστικά (περιλαμβανομένου του delta-delta log energy)	88.16%	Parkinson's Disease Classification Dataset	Ταξινόμηση ΝΠ

Speech-Based Classification of Parkinson's Disease and Essential Tremor: A Gaussian Mixture Models Approach (Chandrasekar, Silpaja/2024) [82]	Support Vector Machine (με GMMs)	Άρθρωση, φωνητική, προσωδία	Υποσχόμενη ακρίβεια (δυαδική και τριπλή ταξινόμηση)	Δεδομένα ομιλίας από ασθενείς με ΝΠ, βασικό τρόμο και υγιείς (Τσέχικα)	Ταξινόμηση ΝΠ και βασικού τρόμου
Machine Learning Based Early Detection of Parkinson's Disease using XGBoost and Random Forest (Adiga S. et al., 2024) [94]	XGBoost, Random Forest (RF), Logistic Regression (LR), Support Vector Machines (SVM), LightGBM, K-Nearest Neighbors (KNN)	Βιοϊατρικές μετρήσεις φωνής (jitter, shimmer, HNR)	XGBoost: 98%	UCI Parkinson's Disease Voice Dataset (195 ηχογραφήσεις, 31 άτομα, 23 με ΝΠ), PPMI. Ηλικίες 25-60.	Έγκαιρη και μη επεμβατική ανίχνευση της ΝΠ.
Analysis of Voice, Speech, and Language Biomarkers of Parkinson's Disease Collected in a Mixed Reality Setting (Dudek M. et al., 2025) [97]	XGBoost, Logistic Regression, SVM, Random Forests, AdaBoost	Ακουστικά (MFCCs, φασματικά, RASTA, F0, jitter, shimmer, σιωπές) και γλωσσικά (POS, NER, συχνότητα λέξεων) από 5 εργασίες ομιλίας.	F1-score (XGBoost): 0.90 ± 0.05 (επανεξιστόρηση ιστορίας, συνδυασμός ακουστικών/γλωσσικών)	57 Πολωνοί συμμετέχοντες (21 ΝΠ, 36 HC). Δεδομένα συλλεγμένα μέσω MR HMD.	Ανάλυση βιοδεικτών φωνής, ομιλίας και γλώσσας για έγκαιρη ανίχνευση ΝΠ.
Uncovering Acoustic Biomarkers to Classify Parkinson Disease through Machine Learning (Yeboah F., 2025) [95]	XGBoost, Logistic Regression, Gaussian Naive Bayes (GNB), KNN, SVM	Ακουστικά χαρακτηριστικά (βασικά, έντασης, μορφοποίησης, εύρος ζώνης, φωνητικών χορδών, MFCCs, κυματιδίων, TQWT) από παρατεταμένη φώνηση "α".	XGBoost: 90% (πλήρες σύνολο), 88% (μετά από επιλογή χαρακτηριστικών)	UCI ML Data Repository (252 ασθενείς: 188 ΝΠ, 64 HC). Ηλικίες 23-87.	Εντοπισμός του καλύτερου ταξινομητή για την ανίχνευση της ΝΠ.

Parkinson's Disease Detection using XGBoost and Machine Learning (Santos D.F., 2023) [96]	XGBoost	Μετρήσεις φωνητικής συχνότητας, αναλύσεις ήχου, δημογραφικά δεδομένα.	XGBoost: ~93.33%	Περιεκτικό σύνολο δεδομένων με φωνητικές μετρήσεις, αναλύσεις ήχου, δημογραφικά.	Εφαρμογή του XGBoost για την αντιμετώπιση των προκλήσεων της έγκαιρης και ακριβούς διάγνωσης της ΝΠ.
Machine learning-based classification of Parkinson's disease using acoustic features: Insights from multilingual speech tasks (Kim J. et al., 2024) [104]	Voting-based machine learning model (δεν διευκρινίζεται αν ο XGBoost είναι μέρος του)	eGeMAPSv2, νέα χαρακτηριστικά.	Συνολικού μοντέλου: 84.73%	100 Κορεάτες ασθενείς με ΝΠ και 100 υγιείς έλεγχοι.	Ανάπτυξη μοντέλου για την ταξινόμηση προτύπων ομιλίας σε ασθενείς με ΝΠ.
Multilingual evaluation of interpretable biomarkers to represent language and speech patterns in Parkinson's disease (Van Gelderen Lisanne &Tejedor-García Cristian 2023) [98]	Ερμηνεύσιμα Μοντέλα Βασισμένα σε Χαρακτηριστικά (IFMs) που περιλάμβαναν XGBoost, SVM, KNN, RF. Συγκρίθηκαν με Μη Ερμηνεύσιμα Μοντέλα Βασισμένα σε Χαρακτηριστικά (NIFMs) με DNN embeddings.	Προσωδιακά, γλωσσικά και γνωστικά χαρακτηριστικά.	IFMs (με XGBoost) πέτυχαν 4% χαμηλότερη ακρίβεια από NIFMs σε μονογλωσσικό περιβάλλον (δεν απομονώνεται η ακρίβεια μόνο του XGBoost).	NLS, Neurovoz, ItalianPVS, GITA, GermanPD, CzechPD.	Πολυγλωσσική αξιολόγηση βιοδεικτών (ακουστικών, γλωσσικών, γνωστικών) για την αναπαράσταση προτύπων γλώσσας και ομιλίας στη ΝΠ.

Η συγκριτική ανάλυση των μοντέλων ταξινόμησης (SNN, Logistic Regression, Random Forest, SVM, KNN, XGBoost), με σκοπό την ταξινόμηση της νόσου του Πάρκινσον (ΝΠ) μέσω ανάλυσης ομιλίας, έδειξε ότι κανένα από τα εξεταζόμενα μοντέλα δεν υπερτερεί σε όλα τα σενάρια. Η εφαρμογή των Spiking Neural Networks (SNN), αν και πολλά υποσχόμενη λόγω της υψηλής απόδοσης, η χρήση τους σε σενάρια που αφορούν την ταξινόμηση της ΝΠ παραμένει ακόμα περιορισμένη. Από την άλλη, η χρήση της Logistic Regression ως μιας απλής και αξιόπιστης προσέγγισης, που στην πληθώρα των περιπτώσεων αποδίδει υψηλή ακρίβεια και συνολική απόδοση, καθώς και η εφαρμογή Random Forest και Support Vector Machines (SVM) ως πιο ισχυροί ταξινομητές, με εξίσου υψηλές επιδόσεις, αποδεικνύουν ότι τα

παραδοσιακά μοντέλα εξακολουθούν να αποτελούν αξιόπιστες επιλογές για την ταξινόμηση της ομιλία στην ΝΠ. Πιο συγκεκριμένα, το Random Forest εκτός από την υψηλή του αποτελεσματικότητα, είναι σε θέση να δώσει πληροφορίες για την σημασία των χαρακτηριστικών του συνόλου δεδομένων, γεγονός πολύ σημαντικό για την πιο ορθή ερμηνεία των αποτελεσμάτων. Από την άλλη, οι Support Vector Machines παρουσιάζουν υψηλή ευελιξία και επίτευξη υψηλών επιδόσεως σε διαφορετικού τύπου συνόλου δεδομένων, ειδικά αν αυτό το μοντέλο συνδυαστεί με τις κατάλληλες τεχνικές προεπεξεργασίας. Τώρα, το K-Nearest Neighbors (KNN) παρουσιάζει μεταβλητή απόδοση ανάλογα με το πλαίσιο στο οποίο εφαρμόζεται, ωστόσο μπορεί να δείξει μια ικανοποιητική απόδοση, όταν γίνεται επιλογή των κατάλληλων χαρακτηριστικόν ομιλίας και μια πιο στρατηγική επιλογή της τιμής της παραμέτρου k. Η απλότητα αυτής της μεθόδου συχνά την καθιστούν ως ένα μέτρο σύγκρισης με τα παραπάνω πιο σύνθετα μοντέλα. Ωστόσο, ανεξάρτητα από το ποιο μοντέλο θα επιλεχθεί, υπάρχουν κάποιοι κοινοί παράγοντες που επηρεάζουν την τελική απόδοση αυτών των ταξινόμητών. Αυτοί οι κοινοί παράγοντες δεν είναι άλλοι από την ποιότητα των δεδομένων ομιλίας , την κατάλληλη επιλογή χαρακτηριστικών και την ορθή ρύθμιση των εκάστοτε παραμέτρων του κάθε μοντέλου.

3 Μεθοδολογία

3.1 Συλλογή Δεδομένων Ομιλίας και Περιγραφή Συνόλου Δεδομένων.

Στο παρόν κεφάλειο περιγράφονται αναλυτικά οι διαδικασίες , τα εργαλεία και οι τεχνικές που εφαρμόστηκαν, προκειμένου να διερευνηθούν και να απαντηθούν τα ερευνητικά ερωτήματα της μελέτης. Πιο συγκεκριμένα, αναλύονται η επιλογή και η εφαρμογή των διαφόρων μεθόδων έρευνας, η συλλογή και επεξεργασία του συνόλου δεδομένων που χρησιμοποιήθηκε, καθώς και ο τρόπος με τον οποίο αξιολογήθηκαν και ερμηνεύθηκαν τα αποτελέσματα. Κύριος στόχος αποτελεί η σαφής αποτύπωση των βημάτων της έρευνα και η αιτιολόγηση των εκάστοτε επιλογών που έγιναν σχετικά με την μεθοδολογία που ακολουθήθηκε. Έτσι επιδιώκεται η αξιοπιστία και η επιστημονική εγκυρότητα της έρευνας. Για την συλλογή των δεδομένων της παρούσας μελέτης, ακολουθήθηκε ένα πειραματικό πρωτόκολλο συλλογής πολυτροπικών δεδομένων από ασθενείς με την Νόσο του Πάρκινσον (ΝΠ). Τα δεδομένα που συλλέχτηκαν, αφορούσαν χαρακτηριστικά βάδισης, εκφράσεων προσώπου, ομιλίας και άνω άκρων. Στην εργασία αυτή επικεντρωθήκαμε στα χαρακτηριστικά που αφορούν την ομιλία, δεδομένου ότι αποτελούν το βασικό αντικείμενο της μελέτης, ώστε αυτά να χρησιμοποιηθούν για την δημιουργία μοντέλων μηχανικής και βαθιάς μάθησης για την διάγνωση και παρακολούθηση της εξέλιξης της νόσου. Η συλλογή των δεδομένων πραγματοποιήθηκε μέσω συνεργασίας διάφορων φορέων. Αυτοί οι φορείς είναι η Νευρολογική Κλινική του Πανεπιστημιακού Γενικού Νοσοκομείου Ηρακλείου (ΠΑΓΝΗ), η Νευρολογική Κλινική του Πανεπιστημιακού Νοσοκομείου Πατρών (ΠΓΝΠ), το Εργαστήριο Βιοϊατρικής Πληροφορικής και Υπηρεσιών Ηλεκτρονικής Υγείας (BMI) του Ελληνικού Μεσογειακού Πανεπιστημίου (ΕΛ.ΜΕ.ΠΑ), το Εργαστήριο Υπολογιστικής Βιοϊατρικής (CBML) του Ινστιτούτου Πληροφορικής του Ιδρύματος Τεχνολογίας και Έρευνας (ΙΤΕ) και η Μονάδα Ιατρικής Τεχνολογίας και Ευφυών Πληροφοριακών Συστημάτων του Πανεπιστημίου Ιωαννίνων. Για την συλλογή των δεδομένων ομιλίας έγινε χρήση καρδιοειδούς μικροφώνου υψηλής ευκρίνειας (Blue Microphone Snowball), το οποίο συνδέεται μέσω USB στον υπολογιστή, προκειμένου να ληφθούν τα ηχητικά δείγματα από τους συμμετέχοντες. Το μικρόφωνο που επιλέχθηκε είναι σχεδιασμένο για επαγγελματική ηχογράφηση φωνής και λόγω στης σταθερής ποιότητας και της καλής του απόδοσης στην καταγραφή μικρών διακυμάνσεων της φωνής (αλλαγή έντασης, συχνότητας κτλ.), που είναι κρίσιμες στην μελέτη των φωνητικών διαταραχών στην Νόσο Πάρκινσον, το καθιστούν την κατάλληλη επιλογή.

Η ηχογράφηση πραγματοποιήθηκε σε ελεγχόμενο εργαστηριακό περιβάλλον έτσι ώστε ο θόρυβος να είναι όσο το δυνατόν πιο αδύναμος, με σαφείς οδηγίες προς τους συμμετέχοντες και υπό την επίβλεψη εξειδικευμένου προσωπικού έτσι ώστε οι καταγραφές να είναι πιο αξιόπιστες. Σχετικά με την διαδικασία, το μικρόφωνο τοποθετήθηκε σε σταθερή απόσταση 20-30 εκ. από το στόμα των συμμετεχόντων και η διαδικασία ηχογράφησης πραγματοποιήθηκε με την βοήθεια του λογισμικού Praat (έκδοση 6.0.36) σε μονοφωνικό ήχο, δειγματοληψία 44.100 Hz.
Τα δεδομένα προέρχονται από συμμετέχοντες οι οποίοι επιλέχθηκαν με βάση αυστηρών πληθυσμιακών κριτήριων ένταξης και αποκλεισμού. Πιο αναλυτικά, προκειμένου οι συμμετέχοντες να είναι σε θέση να συμμετέχουν στην μελέτη, θα έπρεπε να πληρούν συγκεκριμένα κριτήρια ένταξης, τα οποία είναι:

- Να έχουν διαγνωστεί με ΝΠ βάσει των κλινικών κριτηρίων του Movement Disorder Society (MDS).
- Η ηλικία τους να κυμαίνεται από 20 έως 85 ετών.

- Να κατανοούν το Έντυπο Συγκατάθεσης και να είναι θετικοί στο να το υπογράψουν εκείνοι ή ο/η συνοδός τους, σε περίπτωση που εκείνοι αδυνατούν λόγω έντονων συμπτωμάτων.
- Η βαθμολογία τους στην κλίμακα Hoehn & Yahr (H&Y) είναι μικρότερη ή ίση του 3.

Πέρα από τα κριτήρια ένταξης, λαμβάνονται υπόψιν και τα κριτήρια αποκλεισμού. Εάν ένας ασθενής ικανοποιούσε έστω και ένα από τα παρακάτω κριτήρια, η συμμετοχή του στην μελέτη διακοπτότανε . Τα κριτήρια αυτά είναι:

- Αδυναμία συμμόρφωσης με τις διαδικασίες του πρωτοκόλλου.
- Διαταραχή ομιλίας η οποία δεν προέρχεται από ΝΠ.
- Αδυναμία ανεξάρτητης επικοινωνίας.
- Σοβαρή ψυχιατρική ή γνωσιακή διαταραχή.
- Παρουσία πάρεσης σε οπουδήποτε άκρο.
- Συνυπάρχουσα διάγνωση αγγειακού παρκινσονισμού.



Εικόνα 12:Το μικρόφωνο Blue Snowball που χρησιμοποιήθηκε για την καταγραφή των δεδομένων ομιλίας.

Για την εκτέλεση της παρούσας μελέτης, η συλλογή των πολυτροπικών δεδομένων πραγματοποιήθηκε σε ήσυχο, καλά φωτισμένο δωμάτιο, προκειμένου να ελαχιστοποιηθούν οι εξωτερικές παρεμβολές και να διασφαλιστεί η ακρίβεια των μετρήσεων.
Στο πλαίσιο της συλλογής των δεδομένων, οι συμμετέχοντες κλήθηκαν να ακολουθήσουν μια προκαθορισμένη σειρά ασκήσεων ομιλίας, οι οποίες επιλέχθηκαν κατάλληλα, με σκοπό την αξιολόγηση των κύριων χαρακτηριστικών της φωνής που συνδέονται με την Νόσο του Πάρκινσον. Οι ασθενείς, εκτελούσαν τις ασκήσεις υπό την καθοδήγηση των ερευνητών και σε ελεγχόμενο περιβάλλον. Η διαδικασία εκτελέστηκε σε δύο φάσεις. Στην πρώτη φάση ο ασθενής με ΝΠ βρισκόταν σε κατάσταση OFF, δηλαδή χωρίς να έχει πάρει την φαρμακευτική του αγωγή, ενώ στην δεύτερη φάση βρισκόταν σε κατάσταση ON, δηλαδή αφότου είχαν λάβει την φαρμακευτική αγωγή.
Οι ασκήσεις που εκτελέστηκαν κάλυπταν διάφορες πτυχές της φωνητικής λειτουργίας για να μπορούμε να αξιολογήσουμε καλύτερα της επιδράσεις της ΝΠ στην ομιλία. Ο Πίνακας 1 παρουσιάζει αναλυτικά τις ασκήσεις συνδυαστικά με τον στόχο της κάθε μιας.

Πίνακας 6:Ασκήσεις Ομιλίας για την Ανάλυση Φωνητικών Χαρακτηριστικών στη Νόσο του Πάρκινσον.

Άσκηση	Περιγραφή	Οδηγίες προς συμμετέχοντες	Σκοπός Άσκησης
Μέγιστος χρόνος Φώνησης (MPT) /a/	Εκφορά φωνήματος (/a/) για τη μέγιστη δυνατή διάρκεια.	"Πάρτε μία βαθιά ανάσα και ξεκινήστε να λέτε τον ήχο /ααα/ για όση ώρα μπορείτε." (3 επαναλήψεις)	Αξιολόγηση διάρκειας και σταθερότητας φώνησης.
Μέγιστος χρόνος Φώνησης (MPT) /s/, /z/	Εκφορά φωνημάτων (/s/, /z/) για τη μέγιστη δυνατή διάρκεια.	"Πάρτε μία βαθιά ανάσα και ξεκινήστε να λέτε τον ήχο /σσσ/ και /ζζζ/ για όσο μπορείτε." (3 επαναλήψεις)	Αξιολόγηση διάρκειας και σταθερότητας φώνησης.
Ανάγνωση Προτάσεων (CAPE-V Sentences)	Ανάγνωση προτάσεων με σταθερό ρυθμό.	"Θα σας διαβάσω προτάσεις και θα τις επαναλάβετε μετά από εμένα."	Αξιολόγηση σαφήνειας, άρθρωσης και σταθερότητας φωνής.
Rainbow Passage	Ανάγνωση προκαθορισμένης παραγράφου με σταθερό ρυθμό.	"Διαβάστε την παρακάτω παράγραφο."	Αξιολόγηση της ροής, ρυθμού και προσωδίας κατά την διάρκεια της ανάγνωσης.
Δυναμική Αρθρωτική Ταχύτητα (DDK Rate)	Γρήγορη εκφορά συγκεκριμένων συλλαβών.	"Πείτε τις παρακάτω συλλαβές όσο πιο γρήγορα μπορείτε: /pa/, /ta/, /ka/, /pataka/" (10 επαναλήψεις ανά συλλαβή)	Αξιολόγηση άρθρωσης, συντονισμού και κινητικότητας του μυός του στόματος.
Counting Task	Ταχεία καταμέτρηση αριθμών.	"Μετρήστε από το 1 έως το 10 όσο πιο γρήγορα μπορείτε."	Αξιολόγηση της ομιλίας και του ρυθμού εκφοράς.
Μη Λεκτικοί Ήχοι	Παραγωγή μη λεκτικών ήχων.	"Ροχαλίζετε όταν κοιμάστε; Γίνεται σχεδόν κάθε βράδυ; , Κλείστε το στόμα σας και εισπνεύστε βαθιά από τη μύτη 5 φορές, Βήξτε μια φορά"	Αξιολόγηση της ικανότητας παραγωγής μη λεκτικών ήχων.

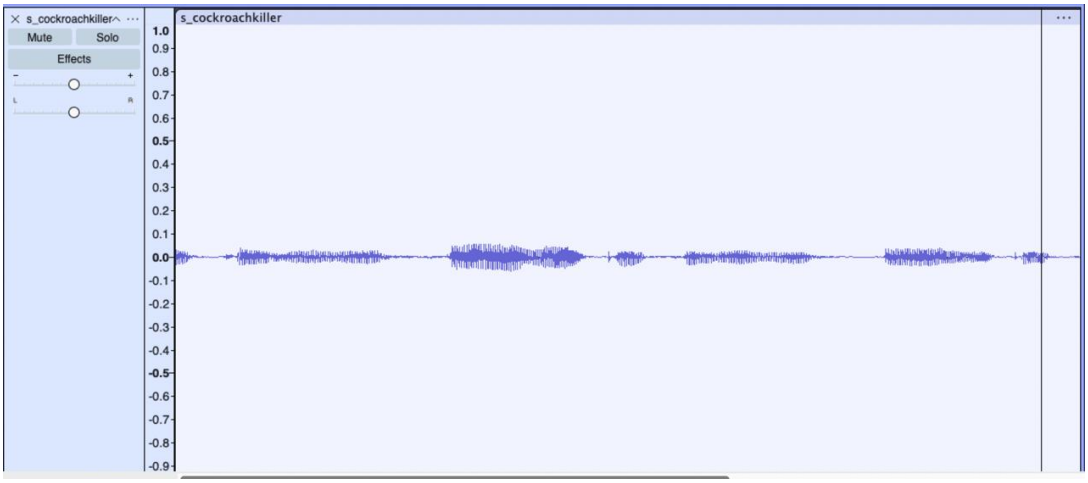
Συνολικά, η μελέτη περιλάμβανε 250 ασθενείς, ωστόσο για τους σκοπούς της παρούσας διπλωματικής εργασίας έγινε χρήση δεδομένων από 101 ασθενείς που προερχόταν από το Ηράκλειο και 53 ασθενείς από την Πάτρα.

3.2 Διαδικασία Αποθορυβοποίησης (Denoising) με Χρήση του Λογισμικού Audacity

Η προ-επεξεργασία (pre-processing) είναι ένα σημαντικό βήμα για την ανάλυση δεδομένων που αφορούν την ομιλία, γιατί σαν κύριο στόχο έχει να βελτιώσει την ποιότητα των δεδομένων ήχου πριν αυτά χρησιμοποιηθούν σε μοντέλα βαθιάς μάθησης. Ένα σημαντικό βήμα αυτής της διαδικασίας είναι η απομάκρυνση του θορύβου (denoising) από το ηχητικό σήμα έτσι ώστε το τελικό αρχείο ήχου να είναι καλύτερο ποιοτικά, το οποίο συνεπάγεται και με καλύτερα αποτελέσματα στην μετέπειτα ανάλυση.

Στην παρούσα μελέτη, η διαδικασία απομάκρυνσης του ανεπιθύμητου θορύβου πραγματοποιείται μέσω του λογισμικού Audacity, το οποίο διαθέτει τα κατάλληλα εργαλεία για την επεξεργασία των σημάτων ήχου. Το Audacity, είναι ένα αποτελεσματικό λογισμικό το οποίο είναι εύχρηστο και δίνει την δυνατότητα στον χρήστη, μέσω της λειτουργίας Noise Reduction που διαθέτει , να επιτύχει την αποθορυβοποίηση. Η διαδικασία που ακολουθήθηκε για την αποθορυβοποίηση περιγράφεται αναλυτικά παρακάτω. Αρχικά, έγινε επιλογή ενός αντιπροσωπευτικού τμήματος από το αρχείου ήχου, στο οποίο εντοπιζόταν ο ανεπιθύμητος θόρυβος χωρίς την παρουσία ομιλίας και αυτό το δείγμα που εμφανιζόταν ο θόρυβος χρησιμοποιήθηκε ως δείγμα αναφοράς. Ακόλουθα, μέσω των εργαλείων του Audacity, έγινε δημιουργία του προφίλ θορύβου (Noise Profile), το οποίο έχει να κάνει με τα χαρακτηριστικά του θορύβου που θέλουμε να απομακρύνουμε από τα αρχεία ήχου. Μετά την δημιουργία του προφίλ θορύβου, ακολούθησε η λειτουργία απομάκρυνσης θορύβου (Noise Reduction) σε όλα τα δείγματα που διαθέταμε. Σε κάθε δείγμα πραγματοποιήθηκαν οι κατάλληλες ρυθμίσεις στις παραμέτρους, προκειμένου να λάβουμε το καλύτερο δυνατό αποτέλεσμα, έχοντας ένα σήμα ήχου με όσο το δυνατόν λιγότερο θόρυβο και χωρίς να αλλοιωθεί η αρχική ποιότητα του δείγματος. Αφού ολοκληρώθηκε η διαδικασία, έγινε αξιολόγηση του τελικού αποτελέσματος. Η αξιολόγηση αυτή περιλάμβανέ, αρχικά την ακρόαση του επεξεργασμένου ήχου περικειμένου να εντοπίσουμε πιθανό θόρυβο που δεν είχε εξαλειφθεί και έπειτα την οπτική ανάλυση των φασματικών χαρακτηριστικών.

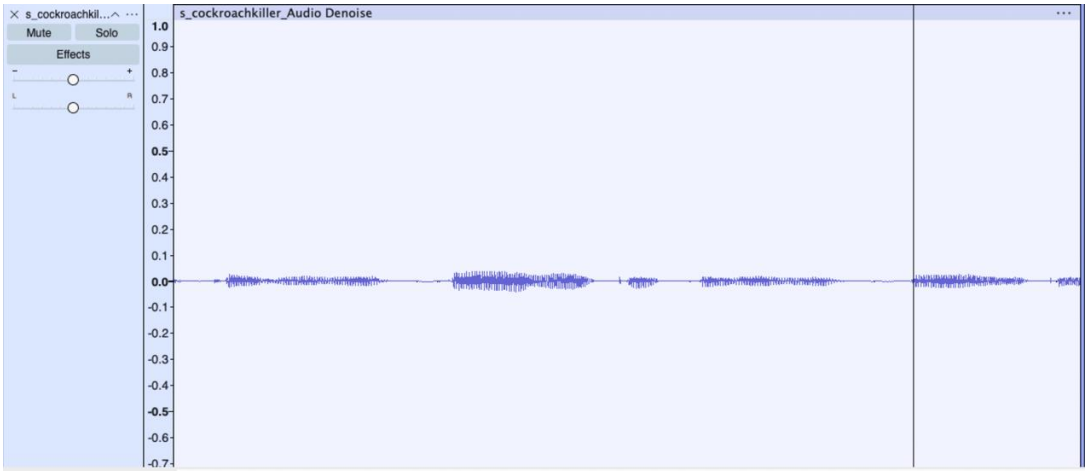
Για να γίνει πιο αντιληπτή η αποτελεσματικότητα της διαδικασίας αποθορυβοποίησης, παρακάτω παρουσιάζονται οι εικόνες που εμφανίζουν το αρχικό ηχητικό σήμα και το σήμα που προκύπτει μετά την διαδικασία, το οποίο είναι απαλλαγμένο από θόρυβο.



Εικόνα 1: Ηχητικό σήμα με παρουσία θορύβου πριν την διαδικασία αποθορυβοποίησης μέσω Audacity.

Η πρώτη εικόνα καταγράφει την αρχική κατάσταση του ηχητικού δείγματος, στην οποία παρατηρείται εμφανής παρουσία θορύβου. Αντίθετα, στη δεύτερη εικόνα παρουσιάζεται το ίδιο ηχητικό δείγμα μετά την εφαρμογή της διαδικασίας αποθορυβοποίησης μέσω του

εργαλείου Noise Reduction στο Audacity. Το σήμα εμφανίζεται αισθητά καθαρότερο, με μειωμένο υπόβαθρο.



Εικόνα 2: Ηχητικό σήμα μετά την διαδικασία αποθορυβοποίησης μέσω Audacity, με εμφανή μείωση θορύβου.

Μέσω αυτής της διαδικασίας εξάγονται καθαρότερα δείγματα ομιλίας, βελτιώνοντάς έτσι την συνολική ποιότητα των δεδομένων και κατά συνέπεια και την πιο ορθή ανάλυση των χαρακτηριστικών την ομιλίας σε άτομα που νοσούν από την Νόσο του Πάρκινσον.

3.3 Εξαγωγή Ακουστικών Χαρακτηριστικών

Προκειμένου να γίνει ανάλυση των ηχητικών δεδομένων και εξαγωγή των απαραίτητων χαρακτηριστικών, έγινε χρήση διαφόρων μεθόδων επεξεργασίας φωνής μέσω των βιβλιοθηκών Parselmouth(Praat), OpenSMILE, Librosa και eGeMAPS που προσφέρονται στην Python. Η χρήση αυτών των τεχνικών, δίνει την δυνατότητα να γίνει ποσοτικοποίηση των διαφόρων παραμέτρων παρέχοντας πληροφορίες για τις φωνητικές λειτουργίες και βοηθώντας στην διαφοροποίηση των ασθενών μεταξύ των κατηγοριών που αναφέρονται παρακάτω. Η διαδικασία που ακολουθήθηκε, περιλαμβάνει την ανάλυση και εξαγωγή των διαφόρων τύπων χαρακτηριστικών, που αναγράφονται αναλυτικά στον παρακάτω πίνακα.

Πίνακας 7:Φωνητικά Χαρακτηριστικά και η Συσχέτισή τους με την Ομιλία Ατόμων με Νόσο του Πάρκινσον.

Κατηγορία	Χαρακτηριστικό	Περιγραφή
Ακουστικά Χαρακτηριστικά	**Jitter**	Μεταβολή της θεμελιώδους συχνότητας (F0), που σχετίζεται με φωνητική αστάθεια.
	Shimmer	Διακύμανση στην ένταση της φωνής, που σχετίζεται με προβλήματα φωνητικής διαχείρισης.
	Harmonics-to-Noise Ratio (HNR)	Αναλογία αρμονικών προς θόρυβο (χαμηλές τιμές δείχνουν μειωμένη καθαρότητα).

Φασματικά Χαρακτηριστικά	**MFCCs (Mel-Frequency Cepstral Coefficients)**	Πληροφορίες για το ηχόχρωμα και φίλτρο φωνής, ανίχνευση αστάθειας και ανωμαλιών στην άρθρωση.
	Φασματική εντροπία(Spectral Entropy)	Μεταβολή φασματικής ενέργειας, που σχετίζεται με αδυναμία στον έλεγχο άρθρωσης.
	Formant Frequencies (F1, F2, F3)	Συχνότητες συντονισμού της φωνητικής οδού όπου επηρεάζουν την καθαρότητα της ομιλίας.
Ρυθμικά Χαρακτηριστικά	**Speech Rate**	Ρυθμός ομιλίας, αργός ή ακανόνιστος σε ασθενείς με Νόσο Πάρκινσον.
	Pause Duration	Μετρά τις παύσεις στην ομιλία, αυξημένες σε ασθενείς με Πάρκινσον.
Χαρακτηριστικά μέσω eGeMAPS	**Alpha Ratio & Hammarberg Index**	Κατανομή ενέργειας στις συχνότητες (ανισορροπία σε ασθενείς με Πάρκινσον).
	Pitch	Θεμελιώδης συχνότητα της φωνής (F0) (μείωση οδηγεί σε μονότονη ομιλία).
	Loudness (Intensities & Energy)	Συνολική αξιολόγηση δυναμικής φωνής, μείωση λόγω υποφωνίας.
	Soft Phonation Index (SPI)	Διακυμάνσεις στην παραγωγή φωνής, που σχετίζονται με ασταθή και χαμηλόφωνη ομιλία.

Στο πλαίσιο της παρούσας μελέτης, σχεδιάστηκε και υλοποιήθηκε μία πλήρως στοχευμένη διαδικασία εξαγωγής φωνητικών χαρακτηριστικών, βασισμένη σε σύγχρονες αλγοριθμικές τεχνικές. Για κάθε συμμετέχοντα και για κάθε κατάσταση (υπό την επήρεια φαρμακευτικής αγωγής – ON και χωρίς αγωγή – OFF), ακολουθήθηκε εξατομικευμένη ανάλυση, εξασφαλίζοντας την ακρίβεια και την αξιοπιστία των δεδομένων. Ιδιαίτερη έμφαση δόθηκε στη διατήρηση υψηλών προτύπων ποιότητας κατά τη συλλογή και την επεξεργασία των σημάτων, ώστε τα αποτελέσματα να αντανακλούν με τον πλέον αντιπροσωπευτικό τρόπο τα χαρακτηριστικά της ομιλίας. Η ροή εργασίας που ακολουθήθηκε, είναι αυτή που παρουσιάζεται παρακάτω. Αρχικά, κάθε ηχογράφηση που περιλαμβανόταν στους επιμέρους φακέλους συλλογής δεδομένων, οι οποίοι περιείχαν τα αρχεία όλων των συμμετεχόντων σε κατάσταση υπό φαρμακευτική αγωγή (ON) και εκτός φαρμακευτικής αγωγής (OFF), υποβλήθηκε σε προκαταρκτικό έλεγχο ποιότητας. Στη συνέχεια, ακολουθήθηκε η διαδικασία προεπεξεργασίας, με στόχο την ομοιογενή προετοιμασία των σημάτων για την ακριβή και αξιόπιστη εξαγωγή των φωνητικών χαρακτηριστικών.Από αυτά τα δεδομένα, εξήχθην τα χαρακτηριστικά που μας δίνουν τις απαραίτητες πληροφορίες για την σταθερότητα, την ένταση

και άλλων σημαντικών δεδομένων του φωνητικού σήματος, στοιχεία που συνδέονται άρρηκτα με την Νόσο του Πάρκινσον. Αφού ολοκληρωθεί η εξαγωγή των χαρακτηριστικών, όλα τα εξαγόμενα δεδομένα αποθηκεύονται σε ένα Pandas DataFrame, με την κάθε γραμμή να αντιστοιχεί σε μια επιμέρους άσκηση και οι στήλες στα χαρακτηριστικά και το τελικό σύνολο δεδομένων αποθηκεύεται σε ένα αρχείο .csv , το οποίο χρησιμοποιείται στην συνέχεια για περαιτέρω ανάλυση. Η διαδικασία που ακολουθήθηκε, διασφαλίζει ότι τα χαρακτηριστικά εξάγονται μεθοδικά χωρίς να επηρεάζεται η ακρίβεια και η ποιότητα των αποτελεσμάτων.

3.4 Στατιστική Ανάλυση και Μοντελοποίηση των Ηχητικών Χαρακτηριστικών

Αφού ολοκληρώθηκε η διαδικασία εξαγωγής και οργάνωσης των φωνητικών χαρακτηριστικών, το επόμενο στάδιο της μελέτης επικεντρώθηκε στη στατιστική ανάλυση των δεδομένων αυτών. Στην παρούσα ενότητα περιγράφεται η αναλυτική μεθοδολογία που ακολουθήθηκε για την στατιστική ανάλυση με σκοπό την αξιολόγηση της πρόβλεψης των κλινικών παραμέτρων που συνδέονται με την Νόσο του Πάρκινσον, μέσω των εξαχθέντων χαρακτηριστικών ομιλίας. Η διαδικασία αυτή πραγματοποιήθηκε με την χρήση σύγχρονων στατιστικών μοντέλων μέσω λεπτομερή κώδικα που υλοποιήθηκε στην γλώσσα προγραμματισμού R. Η ανάπτυξη της μεθόδου υλοποιήθηκε τόσο για τους 101 ασθενείς που προέρχονται από την περιοχή του Ηράκλειου, όσο και για τους 53 ασθενείς από της περιοχή της Πάτρας.

3.4.1 Κλινικά Ερωτήματα και Παράμετροι Πρόβλεψης

Στο πλαίσιο της παρούσας μελέτης, καθορίστηκαν συγκεκριμένα κλινικά ερωτήματα, με στόχο να αξιολογηθεί κατά πόσο τα χαρακτηριστικά της ομιλίας μπορούν να προβλέψουν σημαντικές παραμέτρους που σχετίζονται με τη Νόσο του Πάρκινσον. Τα ερωτήματα αφορούν τόσο τη σταδιοποίηση της νόσου όσο και την παρουσία κλινικών συμπτωμάτων και την ανταπόκριση στη θεραπευτική αγωγή. Στον ακόλουθο πίνακα παρουσιάζονται αναλυτικά οι κλινικές παράμετροι που εξετάστηκαν:

Πίνακας 8: Εξεταζόμενα Κλινικά Ερωτήματα Σχετικά με τη Νόσο του Πάρκινσον.

Κλινικό Ερώτημα	Τύπος Μεταβλητής	Περιγραφή
Hoehn & Yahr State (H&Y)	Δυαδική ή Πολυεπίπεδη	Εκτίμηση σταδίου βαρύτητας της νόσου: είτε ως H&Y ≥ 3 ή H&Y < 3 (δυαδική), είτε ως στάδιο 1, 2, 3 ή 4 (πολυεπίπεδη μεταβλητή).
Παρουσία Διακυμάνσεων (Fluctuations)	Δυαδική	Καταγραφή παρουσίας ή μη κινητικών διακυμάνσεων (Yes/No).
Παρουσία Δυσκινησίας (Dyskinesia)	Δυαδική	Καταγραφή παρουσίας ή μη δυσκινησίας (Yes/No).
Άμεση Ανταπόκριση στη Θεραπεία (Short-term Levodopa Response)	Ποσοτική και Ποιοτική	Ποσοτική αξιολόγηση μέσω της μεταβολής UPDRS_III_dif_% και ποιοτική ταξινόμηση σε κατηγορίες "Excellent", "Suboptimal" και "Poor".

3.4.2 Δομή και Οργάνωση του Δείγματος

Αφού παρουσιάστηκαν οι γενικές πληροφορίες που αφορούν την ΝΠ, η προεπισκόπηση των χαρακτηριστικών ομιλίας που είναι σημαντικά για την ανάλυση και αξιολόγηση της νόσου, καθώς και τα εξεταζόμενα κλινικά ερωτήματα που τίθενται να απαντηθούν στην παρούσα μελέτη, είναι σημαντικό να γίνει μια αναφορά στο δείγμα που χρησιμοποιήθηκε. Το δείγμα που αξιοποιήθηκε στην συγκεκριμένη μελέτη αποτελείται από άτομα που έχουν διαγνωστεί με την Νόσο του Πάρκινσον, τα οποία συμμετείχαν στις ηχογραφήσεις ομιλίας, συμβάλλοντας στην εξαγωγή των φωνητικών χαρακτηριστικών. Οι συμμετέχοντες προέρχονται από τα δύο ερευνητικά κέντρα, του Ηρακλείου και της Πάτρας, οι οποίοι εκτέλεσαν μια σειρά από διαφορετικές ασκήσεις ομιλίας, που έχουν περιγράφει αναλυτικότερα παραπάνω. Για καθένα από τους συμμετέχοντες συλλέχθηκαν δημογραφικά και σημαντικά κλινικά δεδομένα, όπως το φύλο, η ηλικία, το στάδιο της νόσου, παρουσία ή μη διακυμάνσεων και δυσκινησίας, καθώς και ανταπόκριση στην θεραπεία. Τα δεδομένα αυτά στην συνέχεια χρησιμοποιήθηκαν για την κατηγοριοποίηση των ασθενών και τη συσχέτιση τους με τα φωνητικά χαρακτηριστικά που εξήχθησαν..

3.4.3 Επιλογή Στατιστικών Μοντέλων για την Ανάλυση των Δεδομένων

Για να απαντηθούν αυτά τα κλινικά ερωτήματα έγινε χρήση στατιστικών μοντέλων και η επιλογή αυτών βασίστηκε στην φύση των δεδομένων. Για μεταβλητές που περιλάμβαναν επαναλαμβανόμενες μετρήσεις ανά ασθενή τα γραμμικά μεικτά μοντέλα (Linear Mixed Models - LMM) κρίθηκαν καταλληλότερα διότι μπορούν να λάβουν υπόψιν τις διαφορές που μπορεί να υπάρχουν σε μετρήσεις που αφορούν τον ίδιο ασθενή σε διαφορετικές χρονικές στιγμές ή συνθήκες. Αυτό σημαίνει ότι για τον ίδιο ασθενή, το μοντέλο αυτό αναγνωρίζει ότι οι τιμές που προκύπτουν ενδεχομένως να σχετίζονται και να επηρεάζονται από κοινούς παράγοντες. Αυτές οι ιδιότητες καθιστούν τα LMM πολύ χρήσιμα για την στατιστική ανάλυση, διότι μειώνεται το σφάλμα που ίσως πρόκυπτε αν θεωρούσαμε κάθε μέτρηση ανεξάρτητή, επομένως η ανάλυση γίνεται πιο αξιόπιστη και ακριβής. Από την άλλη, για μεταβλητές στις οποίες υπάρχει μόνο μια μέτρηση ανά ασθενή (απουσία επαναλαμβανόμενων μεταβλητών), η χρήση Απλών Γραμμικών Μοντέλων (Linear Models - LM) κρίθηκε πιο αποτελεσματική, διότι δεν απαιτείται πιο σύνθετη μέθοδος ανάλυσης. Τα μοντέλα αυτά ουσιαστικά μοντελοποιούν την σχέση μεταξύ της εξαρτημένης μεταβλητής και των ανεξάρτητων χωρίς να λαμβάνουν υπόψιν τυχόν αλλαγές στις μετρήσεις του ασθενή.

3.4.4 Οργάνωση και Χωρισμός των Δεδομένων

Τα δεδομένα που χρησιμοποιήθηκαν, διαχωρίστηκαν σε ομάδες ασκήσεων (tasks) με βάση την φωνητική δραστηριότητα προκειμένου να υπάρχει ομοιογένεια μεταξύ των δεδομένων για να υπάρχει σωστή κατηγοριοποίηση και η ακρίβεια των μοντέλων να είναι πιο υψηλή.
Στον Πίνακα 2 παρουσιάζονται αναλυτικά οι ομάδες ασκήσεων που δημιουργήθηκαν και χρησιμοποιήθηκαν για την διαδικασία της στατιστικής ανάλυσης. Στο Παράρτημα Α παρουσιάζεται μια πιο αναλυτική περιγραφή των ασκήσεων που εμπεριέχονται στο κάθε σετ ασκήσεων.

Πίνακας 9:Κατηγοριοποίηση Ασκήσεων Ομιλίας για την Ανάλυση Φωνητικών Χαρακτηριστικών.

Task	Τύπος Ασκήσεων

Task 1 (A1)	Συνεχής εκφώνηση φωνηέντων/συμφώνων για της αξιολόγηση σταθερότητας στην φωνή.
Task 2 (A2)	Λέξεις/φράσεις για αξιολόγηση σωστής άρθρωσης.
Task 3 (A3)	Συνεχής ανάγνωση κειμένου για συνολική αξιολόγηση ομιλίας.
Task 4 (A4)	Φωνητικές ασκήσεις για συντονισμό αρθρώσεων.
Task 5 (A5)	Εκφώνηση λέξεων που επικεντρώνονται σε συγκεκριμένες ομάδες φωνηέντων.
Task 6 (A6)	Καταμέτρηση αριθμών από 1-10 για αξιολόγηση ρυθμού ομιλίας.
Task 7 (A7)	Καταγραφή της αναπνοής.
Task 8 (A8)	Ηχογράφηση βήχα για της εξέταση των χαρακτηριστικών της φωνής σε συνθήκες εκτός ομιλίας.

3.4.5 Επιλογή και Προετοιμασία Ανεξάρτητων Χαρακτηριστικών

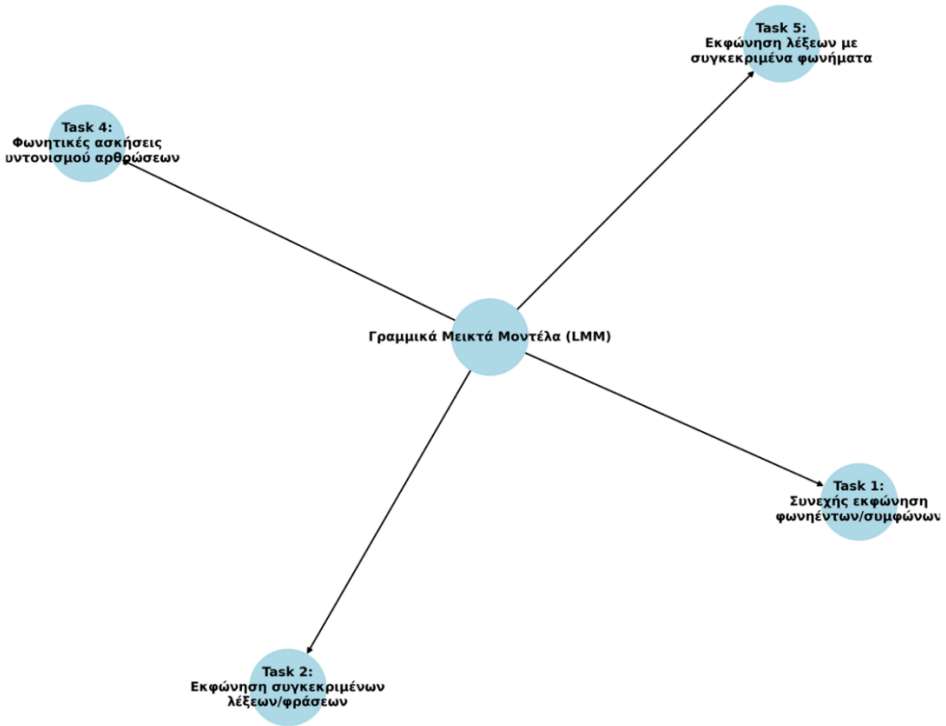
Αφότου ολοκληρώθηκε ο διαχωρισμός των δεδομένων σε συγκεκριμένες ομάδες ασκήσεων (Tasks), ακολούθησε η επιλογή και η προετοιμασία των ανεξάρτητων μεταβλητών (χαρακτηριστικά ομιλίας) πριν ξεκινήσει η διαδικασία της στατιστικής ανάλυσης. Για τα χαρακτηριστικά ομιλίας που χρησιμοποιήθηκαν, έγινε έλεγχος του αν ακολουθούν την κανονική κατανομή, που είναι μια βασική προϋπόθεση για την εγκυρότητα των αποτελεσμάτων.
Αρχικά, για κάθε χαρακτηριστικό έγινε ένας στατιστικός έλεγχος κανονικότητας με σκοπό να διακριβωθεί αν οι τιμές των χαρακτηριστικών ακολουθούν μια κανονική κατανομή ή αν παρατηρούνται αποκλίσεις. Σε περιπτώσεις που εντοπίστηκε απόκλιση, δηλαδή η τιμή του επιπέδου σημαντικότητας (p-value) ήταν μικρότερη της τάξης του 0.05, εφαρμόστηκαν στατιστικοί μετασχηματισμοί. Πιο συγκεκριμένα, εκεί που η κατανομή ακολουθούσε μια θετική ασυμμετρία, εφαρμόστηκε λογαριθμικός μετασχηματισμός (log-transformation) ενώ σε περιπτώσεις που δεν επαρκούσε αυτός ο μετασχηματισμός είτε λόγω μη θετικών δεδομένων, υψηλής ασυμμετρίας ή μη σταθερής διακύμανσης, χρησιμοποιήθηκε ο Box-Cox μετασχηματισμός που προσφέρει καλύτερη προσαρμογή στα δεδομένα και μεγαλύτερη ευελιξίας. Αφού έγινε ο εκάστοτε μετασχηματισμός, ακολούθησε η κανονικοποίση των τιμών με την μέθοδο Z-score για να βελτιωθεί η ακρίβεια της στατιστικής ανάλυσης και η απόδοση των μοντέλων πρόβλεψης .
Ακόλουθα, πραγματοποιήθηκε εκ νέου έλεγχος κανονικότητας έτσι ώστε να αξιολογηθεί αν τα τροποποιημένα χαρακτηριστικά πληρούν τις προϋποθέσεις για την χρήση των γραμμικών μοντέλων. Αυτή η διαδικασία είναι μείζονος σημασίας διότι διασφαλίζει την εγκυρότητα της στατιστικής ανάλυσης και κατά συνέπεια την ακρίβεια και την αξιοπιστία των αποτελεσμάτων.

3.4.6 Υλοποίηση των Στατιστικών Μοντέλων

Σε αυτό το στάδιο υλοποιήθηκαν τα στατιστικά μοντέλα προκειμένου να αξιολογηθούν οι σχέσεις μεταξύ των χαρακτηριστικών ομιλίας και των κλινικών παραμέτρων των ασθενών. Η επιλογή του εκάστοτε μοντέλου, όπως αναφέρθηκε και παραπάνω επιλέχθηκε με βάση την δομή των δεδομένων (αν υπήρχαν ή όχι επαναλαμβανόμενες μέτρησης ανά ασθενή). Η εφαρμογή των στατιστικών μοντέλων υλοποιήθηκε μέσω την γλώσσας προγραμματισμού R και πιο συγκεκριμένα μέσω της προσφερόμενης βιβλιοθήκης 'lmerTest'.
Τα μοντέλα που επιλέχθηκαν εφαρμόστηκαν ξεχωριστά για κάθε ομάδα ασκήσεων και για κάθε κλινικό ερώτημα που τέθηκε.
Τα Γραμμικά Μεικτά Μοντέλα (Linear Mixed Model-LMM) εφαρμόστηκαν στις περιπτώσεις που ο κάθε ασθενείς είχε επαναλαμβανόμενες μετρήσεις άρα υπήρχε συσχέτιση μεταξύ αυτών. Πιο αναλυτικά, η χρήση των LMM, κρίθηκε σημαντική για τις ομάδες ασκήσεων:

- Task 1: Συνεχής εκφώνηση φωνηέντων/συμφώνων.
- Task 2: Εκφώνηση συγκεκριμένων λέξεων/εκφράσεων.
- Task 4: Φωνητικές ασκήσεις για συντονισμό αρθρώσεων.
- Task 5: Εκφώνηση λέξεων που επικεντρώνονται σε συγκεκριμένες ομάδες φωνηέντων.

Στις συγκεκριμένες ομάδες, κάθε ασθενής είχε πολλές μετρήσεις σε διαφορετικές συνθήκες (ON/OFF κατάσταση), επομένως κάθε μοντέλο δημιουργήθηκε χρησιμοποιώντας έναν τυχαίο παράγοντα (random effect) για τον κάθε ασθενή με σκοπό την συμπερίληψη των διαφόρων μετρήσεων του και της σύνδεσης αυτών μεταξύ τους.

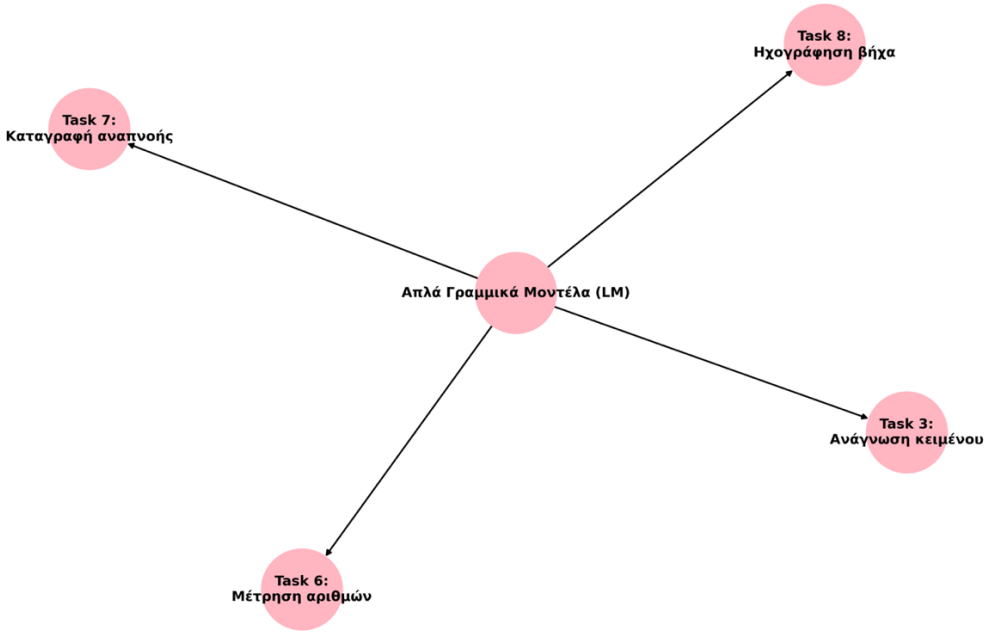


Εικόνα 13: Γραμμικά Μεικτά Μοντέλα (LMM) και Tasks.

Η γενική μορφή που χρησιμοποιήθηκε για τον ορισμό των LMM παρουσιάζεται παρακάτω:

$$Εξαρτημένη\ Μεταβλητή \sim Ανεξάρτητες\ Μεταβλητές + (1|Ασθενής)$$

Όπου εξαρτημένη μεταβλητή αποτελεί η εκάστοτε κλινική παράμετρος (H&Y, Fluctuations, Dyskinesia, UPDRS_III_dif_class, UPDRS_III_dif_%) που εξετάζεται κάθε φορά και ανεξάρτητες μεταβλητές τα χαρακτηριστικά ομιλίας που παίζουν το ρόλο του των προβλεπτικών παραγόντων. Ο όρος $(1|Ασθενής)$ χρησιμοποιείται προκειμένου να υποδηλώσει ότι το μοντέλο κάνει εκτίμηση διαφορετικής σταθεράς (intercept) για τον κάθε ασθενή, δηλαδή η κάθε μέτρηση δεν αντιμετωπίζεται ως ανεξάρτητη μεταβλητή αλλά λαμβάνεται υπόψιν η όποια συσχέτιση μεταξύ τους. Η σχέση μεταξύ κλινικών παραμέτρων και χαρακτηριστικών ομιλίας αναλύεται μέσω γραμμικού μοντέλου παλινδρόμησης, δηλαδή η εξαρτημένη μεταβλητή προκύπτει ως συνάρτηση των ανεξάρτητων. Παράλληλα, συνυπολογίζονται τυχαίου παράγοντες και με αυτόν τον τρόπο η ακρίβεια της μεθόδου βελτιώνεται.



Εικόνα 14:Απλά Γραμμικά Μοντέλα (LM) και Tasks.

Αντίθετα, τα Απλά Γραμμικά Μοντέλα (Linear Model -LM) εφαρμόστηκαν σε περιπτώσεις που ο κάθε ασθενής είχε μοναδική μέτρηση. Τα μοντέλα αυτά χρησιμοποιήθηκαν στις ακόλουθες ομάδες ασκήσεων:

- Task 3: Συνεχής ανάγνωση κειμένου.
- Task 6: Καταμέτρηση αριθμών από 1-10.
- Task 7: Καταγραφή της αναπνοής.
- Task 8: Ηχογράφηση βήχα.

Εδώ τα ο τύπος των δεδομένων που εξεταζόταν ήταν απλούστερος, επομένως η χρήση των LM ήταν επαρκής. Η γενική μορφή που χρησιμοποιήθηκε για τον ορισμό των LM είναι:

$$Εξαρτημένη\ Μεταβλητή \sim Ανεξάρτητες\ Μεταβλητές$$

Και εδώ εξαρτημένη μεταβλητή αποτελεί η εκάστοτε κλινική παράμετρος που εξετάζεται κάθε φορά και ανεξάρτητες μεταβλητές τα χαρακτηριστικά ομιλίας που παίζουν το ρόλο του των προβλεπτικών παραγόντων. Η απουσία επαναλαμβανόμενων μετρήσεων για κάθε ασθενή, οδήγησε στην μη ανάγκη ύπαρξης τυχαίων παραγόντων (random effects) για τη διαχείριση της ενδο-ατομικής μεταβλητότητας. Άρα η σχέση μεταξύ κλινικών παραμέτρων και χαρακτηριστικών ομιλίας αναλύεται μέσω γραμμικού μοντέλου παλινδρόμησης, δηλαδή η εξαρτημένη μεταβλητή προκύπτει ως συνάρτηση των ανεξάρτητων. Έτσι η ανάλυση της σχέσης μεταξύ χαρακτηριστικών ομιλίας και κλινικών παραμέτρων πραγματοποιήθηκε με άμεσο τρόπο και χωρίς την ανάγκη χρήσης κάποιου πιο απαιτητικού μοντέλου στατιστικής ανάλυσης.

3.4.7 Επεξεργασία και Αξιολόγηση Αποτελεσμάτων

Αφού εφαρμόστηκαν τα στατιστικά μοντέλα, έγινε αναλυτική αξιολόγηση των εξαγόμενων αποτελεσμάτων, μέσω των παρακάτω στατιστικών στοιχείων:

- Συντελεστής Εκτίμησης (Estimate): δείχνει το μέγεθος και την κατεύθυνση της επίδρασης του κάθε χαρακτηριστικού ομιλίας με την εκάστοτε κλινική παράμετρο που εξετάζεται. Πιο απλά, μας δίνει πληροφορίες σχετικά με τον βαθμό επίδρασης ενός συγκεκριμένου χαρακτηριστικού προς μια από τις κλινικές παραμέτρους, δείχνοντας αν και κατά πόσο επηρεάζει την εξέλιξη της νόσο στους ασθενείς. Αν ο συντελεστής εκτίμησης είναι θετικός, υποδηλώνει ότι η αύξηση της τιμής αυτού του χαρακτηριστικού, συνδέεται και με αύξηση της κλινικής παραμέτρου. Αν αυτή η αύξηση υποδηλώνει επιδείνωση ή βελτίωση της νόσου, εξαρτάται από την φύση της παραμέτρου που χρησιμοποιείται για ανάλυση. Στην παρούσα μελέτη, δεδομένου ότι όλες οι κλινικές παράμετροι που εξετάζονται έχουν να κάνουν με την σοβαρότητα της νόσου ή την παρουσία άλλων κλινικών μεταβλητών, μια θετική τιμή υποδηλώνει και επιδείνωση της κατάστασης του ασθενή. Από την άλλη, μια αρνητική τιμή συντελεστή εκτίμησης σχετίζεται με μείωση της κλινικής παραμέτρου, άρα βελτίωση της κλινικής εικόνας του ασθενή.
- Τυπικό Σφάλμα (Standard Error): δείχνει κατά πόσο η εκτίμηση των παραμέτρων είναι αξιόπιστη και ακριβής. Η ύπαρξη ενός μικρού τυπικού σφάλματος υποδηλώνει ότι η εκτίμηση είναι σταθερή και πιθανότατα κοντά στην πραγματική τιμή. Αντίθετα, ένα υψηλότερο τυπικό σφάλμα δείχνει απόκλιση από την πραγματική τιμή, άρα η επίδραση του χαρακτηριστικού στην εκάστοτε κλινική μεταβλητή κρίνεται λιγότερο ακριβή.
- Τιμή (t-values): αξιολογεί την στατιστική σημαντικότητα της εκτίμησης, δηλαδή αν η επίδραση ενός συγκεκριμένου χαρακτηριστικού σε μια κλινική παράμετρο είναι πραγματική και όχι απόρροια μιας τυχαίας διακύμανσης. Όσο πιο υψηλή η απόλυτη τιμή t , τόσο πιο ουσιαστική είναι η επίδραση ενός χαρακτηριστικού στην κλινική εικόνα.
- Στατιστική Σημαντικότητα – Τιμή p (p-values): υποδηλώνει την πιθανότητα ενός αποτελέσματος να έχει προκύψει τυχαία και να μην έχε βασιστεί στην πραγματική σχέση μεταξύ εξαρτημένης μεταβλητής και ανεξάρτητων μεταβλητών. Μια χαμηλή τιμή p, υποδηλώνει ότι το εξαγόμενο αποτέλεσμα έχει προκύψει βασιζόμενο στις πραγματικές σχέσεις ανάμεσα στις μεταβλητές, που συνεπάγεται με το ότι το χαρακτηριστικό επηρεάζει με ουσιαστικό τρόπο την κλινική μεταβλητή. Στην παρούσα μελέτη ως κριτήριο στατιστικής σημαντικότητας χρησιμοποιήσαμε το p-value < 0.05,

ώστε να εξετάσουμε αν η επίδραση ενός χαρακτηριστικού στην κλινική μεταβλητή δεν βασίζεται σε τυχαίες διακυμάνσεις.

Για την διευκόλυνση την ανάλυσης και της παρουσίασης των αποτελεσμάτων, όλα τα ευρήματα καταγράφηκαν σε πίνακές, σε αρχείο Excel, προκειμένου η αξιολόγηση των αποτελεσμάτων να γίνει με μεγαλύτερη ακρίβεια. Η μεθοδολογία που ακολουθήθηκε συνολικά αποσκοπούσε στην ελαχιστοποίηση των σφαλμάτων και στην ενίσχυση της αποτελεσματικότητας και αξιοπιστίας. Μέσω της διαδικασίας της στατιστικής ανάλυσης, κατέστη δυνατός ο εντοπισμός των χαρακτηριστικών φωνής που θα μπορούν να αποτελέσουν δείκτες για την εξέλιξη της Νόσου του Πάρκινσον, παρέχοντας πληροφορίες για την κλινική μελέτη των ασθενών. Τα αποτελέσματα αυτής της ανάλυσης αποτελούν την βάση για την συνέχεια της μελέτης.

3.5 Ανάπτυξη και Αξιολόγηση Μοντέλων Ταξινόμησης με χρήση της μεθόδου Leave-One-Subject-Out (LOSO) σε Κλινικά Χαρακτηριστικά Ασθενών με Πάρκινσον.

Αφού ολοκληρώθηκε το στάδιο της στατιστικής ανάλυσης και εξήχθην σημαντικές πληροφορίες σχετικά με την σχέση μεταξύ των χαρακτηριστικών ομιλίας και των κλινικών παραμέτρων των ασθενών , ακολούθησε η ανάπτυξη και η αξιολόγηση μοντέλων ταξινόμησης με την χρήση της μεθόδου Leave-One-Subject-Out (LOSO), αξιοποιώντας τα δεδομένα των ασθενών από την περιοχή του Ηρακλείου αλλά και της Πάτρας. Στόχος αυτού του βήματος είναι η ανάπτυξη αλγορίθμων ταξινόμησης προκειμένου να αξιολογηθεί κατά πόσο τα χαρακτηριστικά της ομιλίας μπορούν να προβλέψουν με ακρίβεια τις κλινικές μεταβλητές (μεταβλητές-στόχοι), οι οποίες αντιπροσωπεύουν κρίσιμες παραμέτρους της Νόσου του Πάρκινσον. Οι μεταβλητές αυτές είναι:

- Βαρύτητα ή στάδιο της Νόσου (H&Y).
- Παρουσία ή απουσία δυσκινησίας(Dyskinesia).
- Ύπαρξη διακυμάνσεων στα συμπτώματα της νόσου (Fluctuations).
- Ανταπόκριση στην θεραπεία, μέσω ποιοτικών (UPDRS_III_dif_class) και ποσοτικών (UPDRS_III_dif_%).

Ακολουθεί η αναλυτική περιγραφή της μεθοδολογίας του ακολουθήθηκε για την εκπαίδευση και αξιολόγηση των μοντέλων ταξινόμησης.

3.5.1 Οργάνωση Δεδομένων και Επιλογή Στατιστικά Σημαντικών Χαρακτηριστικών.

Η ανάλυση βασίστηκε στα δεδομένα ασθενών που προέκυψαν από την περιοχή του Ηράκλειου και την περιοχή της Πάτρας. Αυτά τα δεδομένα περιλάμβαναν τα χαρακτηριστικά ομιλίας, τις κλινικές μεταβλητές που αφορούσαν την Νόσο του Πάρκινσον καθώς και τα αποτελέσματα που προέκυψαν έπειτα από την εφαρμογή της στατιστικής ανάλυσης. Ο στόχος αυτής της διαδικασίας αποσκοπούσε στην διερεύνηση της αποτελεσματικότητας των μοντέλων ταξινόμησης, να προβλέψουν με ακρίβεια τις κλινικές μεταβλητές. Προκειμένου να μελετηθεί οποιαδήποτε διαφοροποίηση μεταξύ των δεδομένων των δύο περιοχών, η ανάλυση πραγματοποιήθηκε για τρείς διαφορετικές περιπτώσεις:

- Το σύνολο δεδομένων για τους ασθενείς της περιοχή του Ηρακλείου.
- Το σύνολο δεδομένων για τους ασθενείς της περιοχή της Πάτρας.
- Το σύνολο δεδομένων για όλους τους ασθενείς.

Κατά την εκτέλεση του αλγόριθμου, η επιλογή της εκάστοτε περίπτωσης γινόταν χειροκίνητα, επιλέγοντας κάθε φορά τα αντίστοιχα αρχεία.
Τα αρχεία που χρησιμοποιήθηκαν για την ανάλυση ήταν:

- Το αρχείο Excel που περιείχε όλα τα εξαγόμενα χαρακτηριστικά ομιλίας των ασθενών.
- Το αρχείο Excel με όλα τα στατιστικά στοιχεία και τις p-value τιμές.

Ακόλουθα, για να διασφαλιστεί ότι η ακρίβεια των προβλέψεων θα είναι όσο τον δυνατόν πιο αξιόπιστη και η απόδοση των μοντέλων θα βελτιωθεί, έγινε επιλογή των πιο σημαντικών χαρακτηριστικών ομιλίας. Η επιλογή αυτή βασίστηκε στα στατιστικά σημαντικά χαρακτηριστικά και πιο συγκεκριμένα στην τιμή p (p-value), η οποία θεωρήθηκε στατιστικά σημαντική όταν η τιμή της ορίστηκε μικρότερή του 0.05 ($p - value < 0.05$), προκειμένου ένα χαρακτηριστικό να θεωρηθεί σημαντικό. Η διαδικασία που ακολουθήθηκε περιλαμβάνει, αρχικά την φόρτωση των τιμών p μέσω του αρχείου excel, στην συνέχεια το φιλτράρισμα των χαρακτηριστικών με βάση το κριτήριο που ορίσθηκε και τέλος την επιλογή των σημαντικών χαρακτηριστικών, τα οποία είχαν και την ισχυρότερη συσχέτιση με τις κλινικές παραμέτρους.

3.5.2 Μεταβλητές-Στόχοι, Ταξινόμηση και Προ-επεξεργασία Δεδομένων

Με την ολοκλήρωση της επιλογής των σημαντικών χαρακτηριστικών, ακολούθησε η ανάπτυξη των μοντέλων ταξινόμησης, προκειμένου να γίνει πρόβλεψή των κλινικών μεταβλητών, που σχετίζονται με την ομιλία των ασθενών. Πριν ξεκινήσει η διαδικασία αυτή, κρίθηκε σημαντική η σωστή προετοιμασία των δεδομένων για τις μεταβλητές-στόχους που έχουν επιλεγεί. Όπως αναφέρθηκε και παραπάνω, οι μεταβλητές-στόχοι που έχουν επιλεγεί αντιπροσωπεύουν κρίσιμες παραμέτρους της Νόσου του Πάρκινσον και πιο συγκεκριμένα η ταξινόμηση των ασθενών βασίστηκε στα ακόλουθα:

Hoehn & Yahr (H&Y): η κλίμακα αυτή εκτιμά την σοβαρότητα της νόσου κατατάσσοντας τους ασθενείς στα στάδια 1 έως 5 και όσο μεγαλύτερο το στάδιο στο οποίο βρίσκεται ο ασθενής, τόσο πιο επιδεινωμένα εμφανίζονται τα φωνητικά χαρακτηριστικά. Προκειμένου να πραγματοποιηθεί η ταξινόμηση, η συνεχής αυτή μεταβλητή μετατράπηκε σε δυαδική, όπου τα στάδια 1,2 ήταν ήπιας σοβαρότητας, άρα η κωδικοποίηση αυτών ήταν 0, ενώ τα στάδια που ήταν μεγαλύτερα του 3 άρα προχωρημένης σοβαρότητας, είχαν κωδικοποίηση 1.

Fluctuations: η μεταβλητή αυτή υποδηλώνει την παρουσία ή όχι διακυμάνσεων στην ομιλία των ατόμων που νοσούν από Πάρκινσον. Τέτοιες διακυμάνσεις μπορεί να είναι αστάθεια στην φωνή και αλλαγές στην άρθρωση. Η μεταβλητή αυτή ήταν ήδη δυαδική (απουσία=0, παρουσία=1), οπότε δεν έγινε κάποια μετατροπή.

Dyskinesia: η μεταβλητή αυτή υποδηλώνει την παρουσία ή όχι δυσκινησιών και είναι ένα αρκετά σημαντικό χαρακτηριστικό της φωνής, διότι οι ακούσιες κινήσεις που παρατηρούνται στα άτομα με ΝΠ μπορούν να επηρεάσουν την σταθερότητα της ομιλίας (ένταση, ρυθμός, τόνος). Πρόκειται για δυαδική μεταβλητή (απουσία=0, παρουσία=1), οπότε δεν έγινε κάποια μετατροπή.

UPDRS_III_dif_per: ποσοστιαία διαφοροποίηση της φωνητικής κατάστασης του ασθενή. Η μεταβλητή απαρτίζονταν από συνεχείς τιμές οι οποίες δήλωναν το ποσοστό της μεταβολής της ομιλίας σύμφωνα με τα κριτήρια UPDRS, όμως για τις ανάγκες της ταξινόμησης η μεταβλητή μετατράπηκε σε κατηγορική οπότε:

- Μικρή αλλαγή : 0-10%
- Μεγάλη αλλαγή : 10-30%
- Υψηλή αλλαγή : > 30%

UPDRS_III_dif_class: ποιοτική διαφοροποίηση της φωνητικής κατάστασης του ασθενή. Η μεταβλητή ταξινομήθηκε σε τρείς κατηγορίες ανάλογα με τις αλλοιώσεις που παρατηρούνται στην φωνητική κατάσταση του ασθενή. Οι κατηγορίες αυτές είναι:

- Μικρή αλλαγή (κατηγορία 1-Excellent): η ομιλία του ατόμου είναι σχεδόν φυσιολογική, χωρίς να παρατηρούνται αλλοιώσεις.
- Μέτρια αλλαγή (κατηγορία 2-Suboptimal): παρατηρούνται αλλοιώσεις στην ομιλία του ατόμου, κυρίως στην άρθρωση και στον τόνο της φωνής.
- Υψηλή αλλαγή (κατηγορία 3-Poor): οι αλλοιώσεις που παρατηρούνται είναι σοβαρές και το άτομο δυσκολεύεται στην ομιλία.

Ακόλουθα, με σκοπό την ακρίβεια της ταξινόμησης, η εφαρμογή τεχνικών προ επεξεργασίας για την διόρθωση πιθανών ασυνεχειών και προβλημάτων στα δεδομένα, κρίθηκε απαραίτητη.

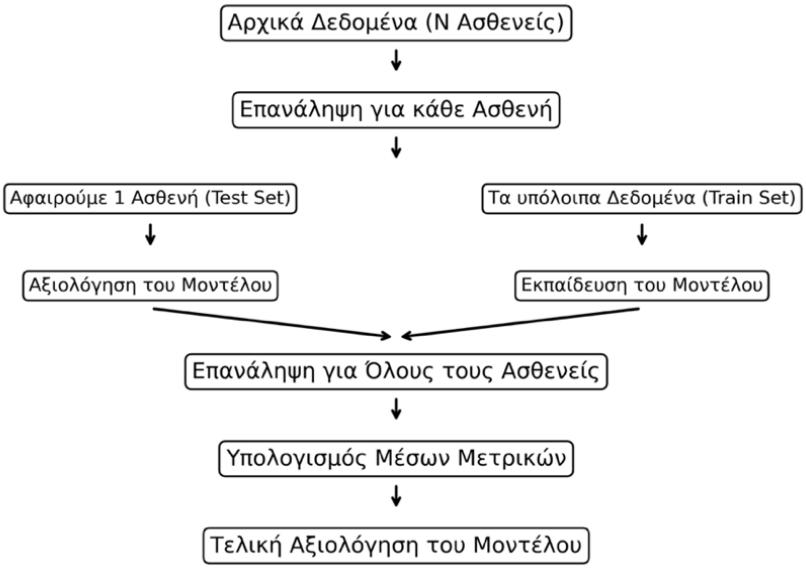
- Αντικατάσταση Κενών Τιμών (NaN): το πρώτο βήμα ήταν η αντικατάσταση κενών τιμών στο σύνολο των δεδομένων με την μέθοδο (SimpleImputer) , καθώς η ύπαρξη αυτών θα είχε αντίκτυπο στα αποτελέσματα. Για τις συνεχείς μεταβλητές , οι κενές τιμές αντικαταστάθηκαν με τον μέσο όρο των δεδομένων της ίδιας στήλης. Από την άλλη για τις κατηγορικές μεταβλητές οι κενές τιμές αντικαταστάθηκαν με την συχνότερη τιμή που εμφανίζεται στα δεδομένα για να υπάρχει συνοχή μεταξύ των δεδομένων.
- Κανονικοποίση Δεδομένων: προκειμένου όλα τα δεδομένα να βρίσκονται στην ίδια κλίμακα (μέσος όρος=1, τυπική απόκλιση=0), πραγματοποιήθηκε κανονικοποίση αυτών μέσω StandartScaler, έτσι ώστε όλα τα χαρακτηριστικά να συμβάλλουν ισότιμα στην διαδικασία της ταξινόμησης και να μην προκύπτουν λανθασμένα αποτελέσματα λόγω διαφορετικής κλίμακας μεταξύ δεδομένων.
- Κωδικοποίηση Μεταβλητών: για να μπορέσουν οι μεταβλητές να χρησιμοποιηθούν στα μοντέλα ταξινόμησης, η κωδικοποίηση τους κρίθηκε μείζονος σημασίας. Οι δυαδικές μεταβλητές μετασχηματίστηκαν σε αριθμητικές τιμές (0,1) , ενώ οι μεταβλητές οι οποίες είχαν παραπάνω από δυο κατηγορίες κωδικοποιήθηκαν κατάλληλα για να διατηρείται η σχέση με το υπόλοιπο σύνολο δεδομένων.
- Εξισορρόπηση Δεδομένων: κάποιες κατηγορίες είχαν συχνότερο ρυθμό εμφάνισης συγκριτικά με άλλες, οπότε έπρεπε να γίνει εξισορρόπηση δεδομένων. Η διαδικασία αυτή πραγματοποιήθηκε εντός των ταξινομικών μοντέλων που χρησιμοποιήθηκαν και έτσι το κάθε μοντέλο προσάρμοζε αυτόματα το βάρος των κατηγοριών και έτσι δεν χρειάστηκε να γίνει ξεχωριστά στα δεδομένα, εντός του αλγορίθμου, κάποια διαδικασία oversampling ή undersampling.

Αφού ολοκληρώθηκαν αυτά τα βήματα, τα δεδομένα ήταν πλέον έτοιμα να χρησιμοποιηθούν για την κατάλληλη εκπαίδευση των μοντέλων ταξινόμησης.

3.5.3 Εκπαίδευση και Αξιολόγηση Μοντέλων Ταξινόμησης με την Χρήση της Μεθόδου LOSO (Leave-One-Subject-Out).

Στα πλαίσια της μελέτης, για να γίνει η δημιουργία και η αξιολόγηση των μοντέλων ταξινόμησης , έγινε χρήση της μεθόδου LOSO (Leave-One-Subject-Out), η οποία είναι μια τεχνική διασταυρούμενης επικύρωσης (cross-validation) και χρησιμοποιείται στην συγκεκριμένη μελέτη, διότι έχουμε ιατρικά δεδομένα, τα οποία είναι ομαδοποιημένα ανά ασθενή. Η μέθοδος αυτή θεωρείται αρκετά αξιόλογή καθώς ο κάθε ασθενής αντιμετωπίζεται σαν να είναι νέος άγνωστος ασθενής και ταυτόχρονα διασφαλίζει ότι τα δεδομένα του, που θα χρησιμοποιηθούν για την αξιολόγηση (test set) θα βρίσκονται μόνο σε αυτό το σύνολο και όχι

και στο σύνολο εκπαίδευσης(train test), αποφεύγοντας έτσι την επικάλυψη δεδομένων (data leakage). Αυτά έχουν ως αποτέλεσμα η ακρίβεια του μοντέλου να είναι υψηλή, διότι η απόδοση εκτιμάται για τον κάθε ασθενή ξεχωριστά. Ωστόσο, σαν μέθοδος έχει το μειονέκτημα του υψηλού υπολογιστικού κόστους, επειδή η εκπαίδευση του μοντέλου γίνεται τόσες φορές, όσες και ο αριθμός των ασθενών.



Εικόνα 15:Διάγραμμα Ροής της Μεθόδου (LOSO) Leave-One-Subject-Out.

Η διαδικασία που ακολουθείται κατά την εκτέλεση αυτής της μεθόδου είναι αρκετά απλή, και συνοψίζεται στο παραπάνω διάγραμμα (Εικόνα 4) που δείχνει όλα τα βήματα που ακολουθούνται μέχρι να φτάσουμε στο τελικό στάδιο της μελέτης.
Σε κάθε επανάληψη του αλγορίθμου, ένας ασθενείς αφαιρείται από το σύνολο δεδομένων και χρησιμοποιείται μόνο για την αξιολόγηση (test set), ενώ το μοντέλο που εκπαιδεύεται κάθε φορά, εκπαιδεύεται στα δεδομένα όλων των υπόλοιπων ασθενών (training set). Το μοντέλο αξιολογείται με βάση τα δεδομένα του ασθενή που είχε αφαιρεθεί. Αυτή η διαδικασία επαναλαμβάνεται για τον κάθε ασθενή ξεχωριστά, μέχρι να εξεταστούν όλοι οι ασθενείς. Τα τελικά αποτελέσματα της μεθόδου λαμβάνουν υπόψιν τον μέσο όρο των μετρικών που χρησιμοποιήθηκαν για την αξιολόγησα, στα οποία θα γίνει πιο αναλυτική αναφορά παρακάτω.

3.5.4 Μοντέλα Μηχανικής Μάθησης για την Ταξινόμηση Κλινικών Δεδομένων

Σε αυτή την μελέτη, εφαρμόστηκαν έξι διαφορετικά μοντέλα ταξινόμησης, με στόχο να γίνει σύγκριση την απόδοσης τους σχετικά με το πόσο αποτελεσματικά είναι στην διάγνωση και αξιολόγηση της Νόσου του Πάρκινσον. Τα μοντέλα που εφαρμόστηκαν παρουσιάζονται στον παρακάτω πίνακα (Πίνακας 3).

Πίνακας 10:Παρουσίαση των Εκπαιδευθέντων Μοντέλων Ταξινόμησης.

Μοντέλο	Περιγραφή	Μειονεκτήματα	Πλεονεκτήματα
Logistic Regression	Γραμμικό μοντέλο που χρησιμοποιείται κυρίως για	Δεν είναι το κατάλληλο μοντέλο για μη	Απλό και αποδοτικό μοντέλο

	δυαδικά προβλήματα ταξινόμησής.	γραμμικά δεδομένα.	για γραμμικά δεδομένα.
K-Nearest Neighbors (KNN)	Μη γραμμικό μοντέλο που βασίζεται στην απόσταση από k πλησιέστερους γείτονες.	Όταν υπάρχει πλήθος δεδομένων έχει μεγάλο υπολογιστικό κόστος.	Απλό και αποδοτικό μοντέλο για μη γραμμικά δεδομένα.
Random Forest Classifier	Σύνολο από δέντρα αποφάσεων που συνεργάζονται μεταξύ τους προκειμένου οι προβλέψεις να γίνουν πιο ακριβής και να μην υπάρχει υπερεκπαίδευση.	Έχει υψηλό υπολογιστικό κόστος.	Ιδανικό για μεγάλο πλήθος δεδομένων, με υψηλή ακρίβεια.
Support Vector Machine (SVM)	Αλγόριθμος που προσπαθεί να βρει τον καλύτερο δυνατό διαχωρισμό μεταξύ των κατηγοριών και μπορεί να προσαρμοστεί σε πιο σύνθετα μοντέλα μέσω kernels.	Έχει υψηλό υπολογιστικό κόστος και η επιλογή του σωστού kernel είναι δύσκολη.	Αποτελεσματικό σε περίπλοκα δεδομένα αλλά όταν έχουμε μικρό σύνολο δεδομένων.
Spiking Neural Networks (SNNs) - LIF	Νευρωνικό δίκτυο που βασίζεται σε παλμούς για την επεξεργασία της πληροφορίας μιμούμενο την λειτουργία των νευρώνων του εγκέφαλου. Το LIF δρα σαν τους βιολογικούς νευρώνες, και ενεργοποιείται όταν το δυναμικό μεμβράνης ξεπεράσει ένα όριο.	Δύσκολη εκπαίδευση λόγω της ασυνεχούς φύσης των spikes.	Μιμείται πιο ρεαλιστικά τη λειτουργία των βιολογικών νευρώνων και είναι ενεργειακά αποδοτικό.
XGBoost	Αλγόριθμος που βασίζεται στον σταδιακό σχηματισμό δένδρων και προκειμένου να αποφύγει την υπερεκπαίδευση κάνει χρήση τεχνικών regularization.	Αν οι παράμετροι του αλγορίθμου δεν ρυθμιστούν με ιδανικό τρόπο, υπάρχει το ρίσκο της υπερεκπαίδευσης.	Υψηλή ακρίβεια, γρήγορη περάτωση της εκπαιδευτικής διαδικασίας και καλή διαχείριση των ελλείπων τιμών.

Η επιλογή των συγκεκριμένων αλγορίθμων βασίστηκε στην ανάγκη για την εύρεση της αποτελεσματικότερης μεθόδου για την επίλυση των ερωτημάτων που κλήθηκαν να απαντηθούν στην παρούσα μελέτη, με βάση τα ιδιαίτερα και πολύπλοκα χαρακτηριστικά των κλινικών δεδομένων που συνδέονται άρρηκτα με την Νόσο του Πάρκινσον. Αρχικά, έγινε επιλογή του αλγορίθμου Logistic Regression διότι είναι μία από τις πιο συνήθης μέθοδούς

ταξινόμησης που χρησιμοποιείται ευρέως σε προβλήματα δυαδικής ταξινόμησης. Τα αποτελέσματα που προκύπτουν μέσω αυτού του αλγόριθμου ερμηνεύονται με σαφή τρόπο, γεγονός το οποίο αποτελεί πολύ σημαντικό παράγοντα για την λήψη αποφάσεων σε κλινικά προβλήματα. Ο επόμενος αλγόριθμος που χρησιμοποιήθηκε είναι ο Random Forest κυρίως για την ικανότητα του να διαχειρίζεσαι αποτελεσματικά μεγάλο όγκο δεδομένων, όπως υπάρχει και σε αυτή την διπλωματική, με υψηλή ακρίβεια και χωρίς τον φόβο για μεγάλη πιθανότατα υπερεκπαίδευσης. Επίσης σαν μέθοδος έχει αποδειχθεί ότι είναι αποτελεσματική ακόμα και σε δεδομένα τα οποία περιέχουν ελλείψεις ή θόρυβο, παρέχοντας και σε αυτές τις περιπτώσεις σταθερά και αξιόπιστα αποτελέσματα. Στην συνέχεια, επιλέχθηκε ο αλγόριθμος K-Nearest Neighbors (KNN) λόγω της ικανοποιητικής του απόδοσης σε δεδομένα που δεν παρουσιάζουν γραμμικές σχέσεις μεταξύ του, κάτι που είναι απαραίτητο και χρήσιμο και για την συγκεκριμένη έρευνας. Τώρα για την αντιμετώπιση πιο περίπλοκων δεδομένων επιλέχθηκε η εφαρμογή του αλγορίθμου Support Vector Machine (SVM), ο οποίος έχει αποδειχθεί αποτελεσματικός στις περιπτώσεις που τα δεδομένα παρουσιάζουν πολυπλοκότητα και έχουν μη γραμμικές ιδιότητες. Άλλος ένας αλγόριθμος που επιλέχθηκε λόγω της ικανότητας του να ανταποκρίνεται με αποτελεσματικότητα σε πολύπλοκα και με μεγάλο πλήθος κλινικά δεδομένα είναι ο XGBoost. Ο αλγόριθμος αυτός είναι δημοφιλής στην εύστοχη διαχείριση ελλιπών δεδομένων και την αποφυγή υπερεκπαίδευσης, παρέχοντας υψηλή ακρίβεια στα τελικά αποτελέσματα, παράγοντες οι οποίοι εμφανίζουν μεγάλο ενδιαφέρον για τα ερωτήματα που πρέπει να απαντηθούν. Τέλος, ιδιαίτερο ενδιαφέρον αποτελεί η χρήση των Spiking Neural Networks (SNNs), λόγω της ικανότητας τους να προσομοιώνουν την λειτουργία των βιολογικών νευρώνων του ανθρώπινου εγκεφάλου. Η συγκεκριμένη επιλογή έγινε διότι η ΝΠ συνδέεται άρρηκτα με δυσλειτουργίες οι οποίες εντοπίζονται στην λειτουργία των νευρώνων ως προς την διάδοση των σημάτων εντός του εγκεφάλου του ασθενή. Επομένως η χρήση και η αξιολόγηση ενός τέτοιου μοντέλου αρχικά θα είναι σε θέση να προσεγγίσει με πιο ρεαλιστικό τρόπο το ζητούμενο της εργασίας αλλά ίσως προσφέρει και πολύτιμες πληροφορίες δίνοντας μια διαφορετική προσέγγιση στην αντιμετώπιση της νόσου.
Η συνδυαστική αξιολόγηση των προαναφερθέντων αλγορίθμων προσφέρει μια πολυδιάσπαση προσέγγιση στην ανάλυση των δεδομένων ομιλίας της παρούσας διπλωματικής εργασίας. Όλοι οι αλγόριθμοι συμβάλλουν στην κατανόηση των πολύπλοκων σχέσεων των κλινικών δεδομένων της ΝΠ, ο καθένας με την δική του διαφορετική προσέγγιση, αποδίδοντας αξιοπιστία και ακρίβεια στα αξιολογούμενα αποτελέσματα. Η εφαρμογή αυτών των μοντέλων στοχεύει στην απάντηση των κλινικών ερωτημάτων της μελέτης, συμβάλλοντας στην εξαγωγή χρήσιμων συμπερασμάτων που αφορούν την διάγνωση και την εξέλιξη της νόσου με την πάροδο του χρόνου. Επομένως η χρήση διαφορετικών αλγορίθμων συμβάλλει κυρίως στην ενίσχυση της αξιολόγησης των αποτελεσμάτων αποτελώντας μια αξιόπιστη βάση στην ερμηνευτική διαδικασία.

3.5.5 Αξιολόγηση της Απόδοσης των Μοντέλων Ταξινόμησης

Αφού πραγματοποιήθηκε η διαδικασία της εκπαίδευσης των μοντέλων, η αξιολόγηση της απόδοσης τους έγινε μέσω συγκεκριμένων μετρικών ταξινόμησης, με στόχο να συγκριθεί η ικανότητα πρόβλεψης των κλινικών παραμέτρων της Νόσου του Πάρκινσον ανάμεσα στα διάφορα μοντέλα. Αυτή η διαδικασία αποσκοπούσε στον εντοπισμό του πιο κατάλληλου μοντέλου ταξινόμησης με κλητήριο την ακρίβεια και την συνολική του απόδοση. Οι μετρικές που επιλέχθηκαν προκειμένου να ληφθεί μια ολοκληρωμένη εικόνα της απόδοσης των μοντέλων, είναι:

- Ακρίβεια (accuracy): είναι το ποσοστό σωστών προβλέψεων που πραγματοποίησε το μοντέλο, συγκριτικά με τα συνολικά δεδομένα. Υπολογίζεται μέσω του τύπου:

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

Όπου έχουμε:

- TP(True Positives) είναι οι περιπτώσεις όπου το μοντέλο προέβλεψε σωστά της παρουσία κλινικής παραμέτρου.
- TN (True Negatives) οι περιπτώσεις όπου το μοντέλο προέβλεψε σωστά την απουσία κλινικής παραμέτρου.
- FP (False Positives) οι περιπτώσεις όπου το μοντέλο προέβλεψε λάθος την παρουσία κλινικής παραμέτρου.
- FN (False Negatives) οι περιπτώσεις όπου το μοντέλο δεν εντόπισε την κλινική παράμετρο ενώ υπήρχε.

- Ακρίβεια Κατηγορίας (Precision): μας δίνει πληροφορίες για το πόσο ακριβείς είναι οι θετικές προβλέψεις του μοντέλου. Όσο μεγαλύτερη η τιμή της Precision, τόσο πιο αξιόπιστες οι προβλέψεις του μοντέλου. Υπολογίζεται μέσω του τύπου:

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

- Ευαισθησίας (Recall): μας δίνει πληροφορίες για το ποσοστό των πραγματικά θετικών περιπτώσεων που εντόπισε σωστά το μοντέλο. Όταν ένα μοντέλο έχει υψηλό ποσοστό ευαισθησίας , το ποσοστό των False Negatives (FN) είναι μικρό άρα δεν χάνει πολλές θετικές περιπτώσεις. Υπολογίζεται μέσω του τύπου:

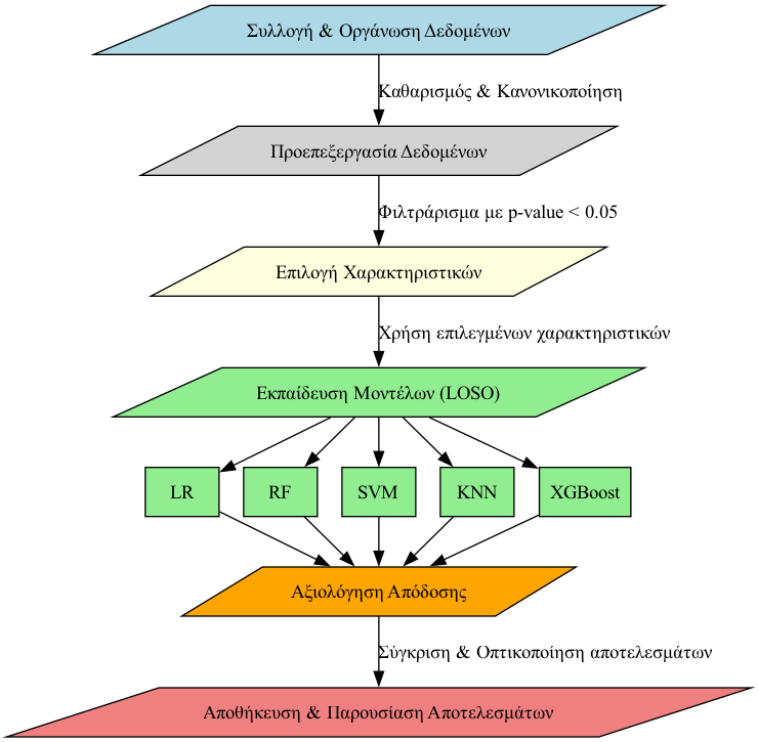
$$Recall = \frac{TP}{TP + FN}$$

- F1-Score: είναι η μέση τιμή μεταξύ Precision και Recall και μας βοηθάει να αξιολογήσουμε μοντέλα με μη ισορροπημένα δεδομένα. Ουσιαστικά, εξισορροπεί τις δύο μετρήσεις και μας παρέχει μια συνολική εικόνα της απόδοσης. Μια υψηλή τιμή F1-Score σημαίνει ότι η πρόβλεψη μιας κλινικής κατηγορίας γίνεται σωστά. Υπολογίζεται μέσω του τύπου:

$$F1 - Score = 2 * \left(\frac{Precision * Recall}{Precision + Recall}\right)$$

- ROC-AUC Score: η καμπύλη ROC ουσιαστικά είναι ένα γράφημα που μας παρέχει πληροφορίες για το πόσο καλά το μοντέλο διαχωρίζει θετικές από αρνητικές περιπτώσεις, συγκρίνοντας Recall με Specificity. Τώρα το ROC-AUC Score είναι μια τιμή που μετράει την περιοχή κάτω από την καμπύλη ROC και αυτή η τιμή δείχνει το κατά πόσο το μοντέλο διαχωρίζει με ορθό τρόπο τις διάφορες κατηγορίες. Η μετρική αυτή είναι πολύ σημαντική στις ιατρικές εφαρμογές διότι μας βοηθάει να αξιολογούμε αν το μοντέλο μπορεί να προβλέψει σωστά τις κλινικές μεταβλητές. Αν η τιμή της ROC-AUC είναι κοντά στην μονάδα, υποδηλώνει ότι το μοντέλο διαχωρίζει αποτελεσματικά τις κατηγορίες, ενώ όταν είναι κάτω από 0.5 σημαίνει ότι η ταξινόμηση είναι λανθασμένη.

Η συνολική διαδικασία και τα κύρια βήματα της μεθοδολογίας που ακολουθήθηκε για την ανάπτυξη και αξιολόγηση των μοντέλων ταξινόμησης συνοψίζεται στο παρακάτω διάγραμμα (Εικόνα 16).



Εικόνα 16:Διαδικασία Ανάλυσης Ομιλίας της Νόσου του Πάρκινσον με Τεχνικές Μηχανικής Μάθησης.

Η αξιολόγηση των μοντέλων μέσω των μετρικών, προσέφερε μια ολοκληρωμένη εικόνα σχετικά με την ικανότητα τους να προβλέπουν με επιτυχία ή όχι τις κλινικές μεταβλητές. Αυτές οι μετρικές, μας έδωσαν την δυνατότητα να συγκρίνουμε τα μοντέλα ταξινόμησης μεταξύ τους, προκειμένου να επιλέξουμε το πιο αποτελεσματικό για την ταξινόμηση των ασθενών. Επίσης βοήθησαν να σχηματίσουμε πιο εύκολα συσχετίσεις μεταξύ των χαρακτηριστικών ομιλίας και των κλινικών παραμέτρων. Δηλαδή , ποια χαρακτηριστικά σχετίζονται περισσότερο με τις κλινικές μεταβλητές και πως αυτά μπορούν να αξιοποιηθούν για περαιτέρω έρευνα. Η όλη διαδικασία που ακολουθήθηκε είναι χρήσιμη, διότι συμβάλει στην δημιουργία αξιόπιστων εργαλείων για την παρακολούθηση της Νόσου του Πάρκινσον με βάση ηχητικών δεδομένων ομιλίας.

Στο επόμενο κεφάλαιο παρουσιάζονται τα αποτελέσματα της ανάλυσης, εξετάζοντας την απόδοση των μοντέλων ταξινόμησης μέσω των μετρικών και την σχέση τους με τα κλινικά δεδομένα ώστε να αξιολογηθεί αν τα χαρακτηριστικά ομιλίας μπορούν να χρησιμοποιηθούν ως βιοδείκτες διάγνωσης.

4 Αποτελέσματα

Όπως αναφέρθηκε και προηγουμένως, στο παρόν κεφάλαιο γίνεται αναλυτική παρουσίαση των αποτελεσμάτων που προκύπτουν από την εφαρμογή της μεθοδολογίας που περιεγράφηκε διεξοδικά στο προηγούμενο κεφάλαιο. Αρχικά, γίνεται μια εκτενής περιγραφή και ανάλυση

των δημογραφικών και κλινικών χαρακτηριστικών του δείγματος που αξιοποιήθηκε για την παρούσα μελέτη, με σκοπό την κατανόηση την σύνθεσης και των γνωρισμάτων του πληθυσμού που συμμετείχε στην έρευνα. Στην συνέχεια ακολουθεί η παρουσίαση των αποτελεσμάτων που προέκυψαν μέσω της ανάλυσης των φωνητικών χαρακτηριστικών που εξήχθησαν με την βοήθεια του eGeMAPS. Η στατιστική επεξεργασία των φωνητικών χαρακτηριστικών πραγματοποιήθηκε με την χρήση Απλών Γραμμικών Μοντέλων(LM-Linear Model) και Γραμμικών Μεικτών Μοντέλων(LMM-Linear Mixed Model), προκειμένου να εντοπιστούν τα χαρακτηριστικά ομιλίας που αναδείχθηκαν ως στατιστικά σημαντικά σε σχέση με τις εξεταζόμενες κλινικές μεταβλητές που εξετάζονται. Τέλος, παρατίθενται τα αποτελέσματα που προέκυψαν μέσω των ταξινομητικών μοντέλων που αξιοποιήθηκαν με την χρήση της μέθόδου Leave-One-Subject-Out (LOSO), η αξιολόγηση των οποίων γίνεται μέσω μετρικών απόδοσης των μοντέλων (accuracy, F1-score, precision, recall, ROC-AUC). Όλα τα αποτελέσματα συνοδεύονται από πίνακες, διαγράμματα και άλλες οπτικές απεικονίσεις, με κύριο στόχο την διευκόλυνση της κατανόησης των ευρημάτων αλλά και την πληρέστερη παρουσίαση των ευρημάτων.

4.1 Περιγραφική Ανάλυση του Δείγματος

Στην παρούσα ενότητα γίνεται αναλυτική παρουσίαση της σύνθεσης του δείγματος που αξιοποιείται στην παρούσα μελέτη ως προς τα δημογραφικά και κλινικά χαρακτηριστικά. Πιο συγκεκριμένα αυτή η κατανομή βασίζεται στο φύλο των ασθενών, καθώς και στο στάδιο της Νόσου του Πάρκινσον με βάση την κλίμακα H&Y, την παρουσία ή μη κινητικών διακυμάνσεων και δυσκινησίας αλλά και την απόκριση στην θεραπεία με Levodopa. Στον παρακάτω πίνακα (Πίνακας 9) παρουσιάζεται αναλυτικά η κατανομή του κάθε δείγματος ανά φύλο και κατηγορία. Επιπλέον, γίνεται οπτική απεικόνισή αυτών, καθώς και μια συνοπτική περιγραφή, με σκοπό την καλύτερη κατανόηση της δομής του δείγματος που εξετάστηκε στην παρούσα διπλωματική εργασία.

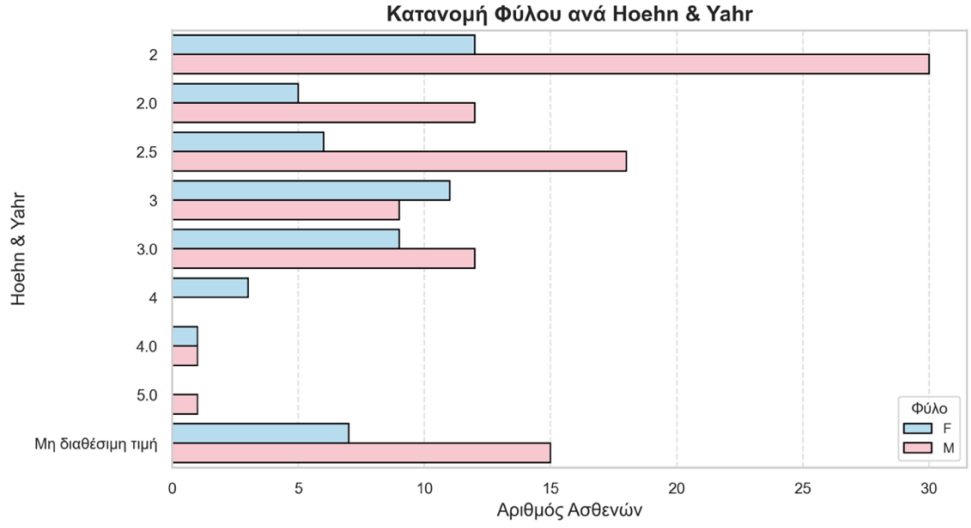
Πίνακας 11: Κατανομή του Δείγματος ανά Κλινική Μεταβλητή και Φύλο.

Κλινικό Ερώτημα	Κατηγορία	Άντρας	Γυναίκα	Σύνολο
Hoehn & Yah	**2**	42	17	59
	2.5	18	6	24
	3	21	20	41
	4	1	4	5
	5	1	0	1
	MD	15	7	22
Fluctuations	**Yes**	21	15	36
	No	21	43	64
	MD	-	-	-

Dyskinesia	**Yes**	25	21	46
	No	71	30	101
	MD	2	3	5
Levodopa Response Quality	**Excellent**	38	21	59
	Suboptimal	31	22	53
	Poor	13	3	16
	MD	8	16	24
Levodopa Response (%)	**0-15%**	10	5	15
	15-30%	32	19	51
	30-50%	27	14	41
	>50%	7	7	14
	Καμία/Αρνητική	1	6	7
	MD	16	8	24

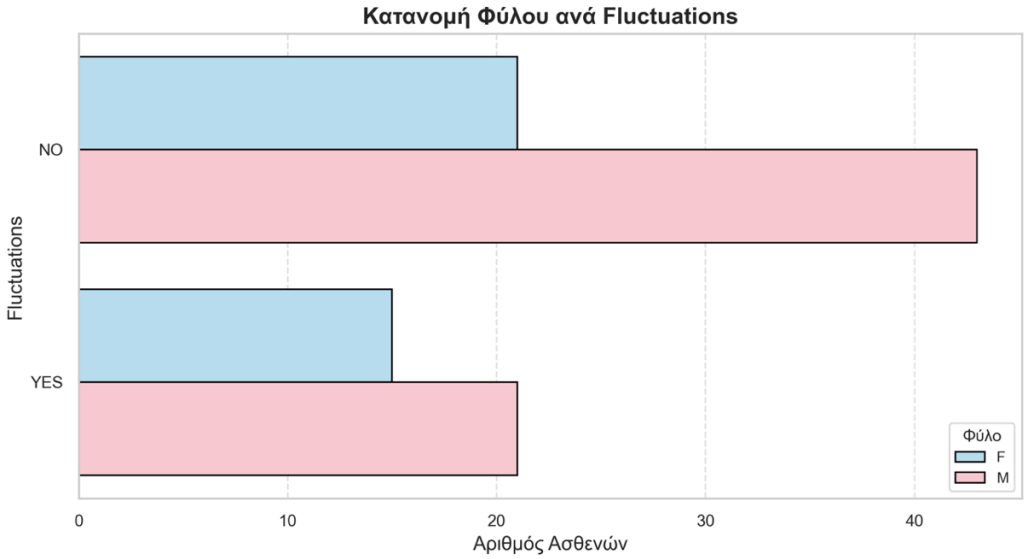
Πιο αναλυτικά, η κατανομή του δείγματος έχει ως εξής:

- Στάδιο νόσου (Hoehn & Yahr): η πλειοψηφία των συμμετεχόντων βρίσκεται στο Στάδιο 2 της νόσου (42 άντρες, 17 γυναίκες), καθώς και στο Στάδιο 3 (21 άντρες, 20 γυναίκες) και λιγότερα άτομα βρίσκονται σε πιο προχωρημένο στάδιο.



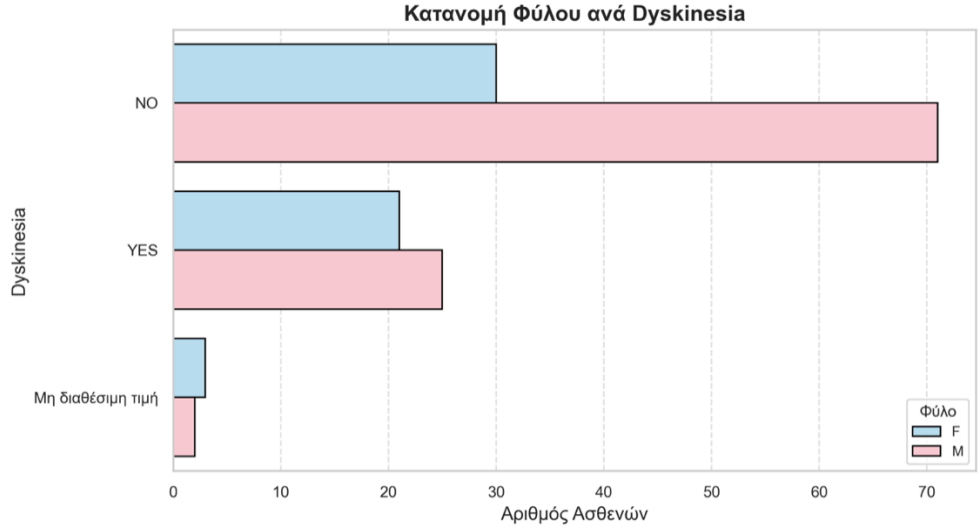
Εικόνα 17:Κατανομή του Φύλου ανά Στάδιο Hoehn & Yahr.

- Παρουσία Διακυμάνσεων (Fluctuations): η πλειοψηφία των ασθενών δεν παρουσίαζαν διακυμάνσεις (21 άντρες, 43 γυναίκες), με μόλις 36 άτομα να παρουσίαζαν (21 άντρες, 15 γυναίκες).



Εικόνα 18:Κατανομή του Φύλου με Βάση την Παρουσία Κινητικών Διακυμάνσεων (Fluctuations).

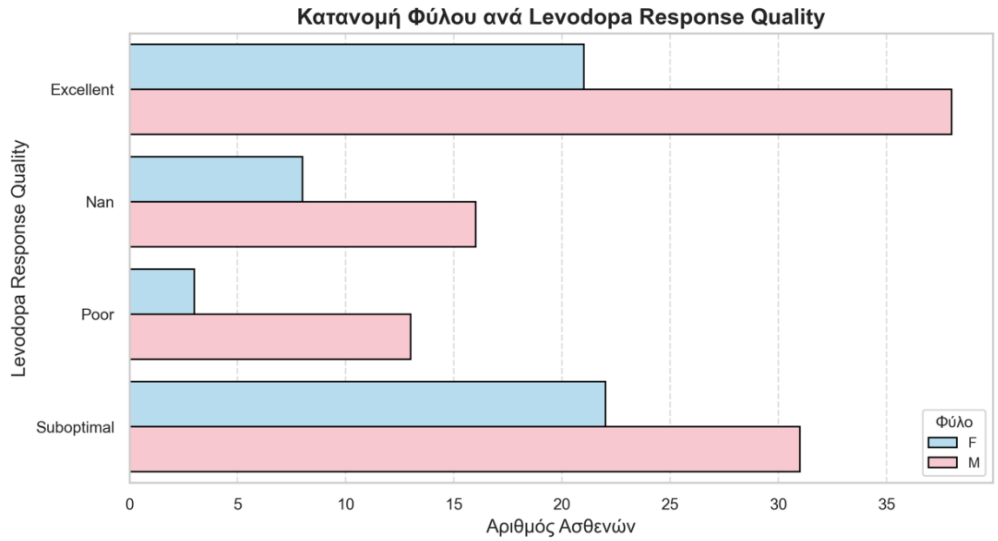
- Παρουσία Δυσκινησίας (Dyskinesia): η πλειοψηφία των ατόμων δεν παρουσίαζαν δυσκινησία (71 άντρες, 30 γυναίκες), ενώ ένα μικρό ποσοστό παρουσίαζε αυτό το σύμπτωμα (25 άντρες, 21 γυναίκες).



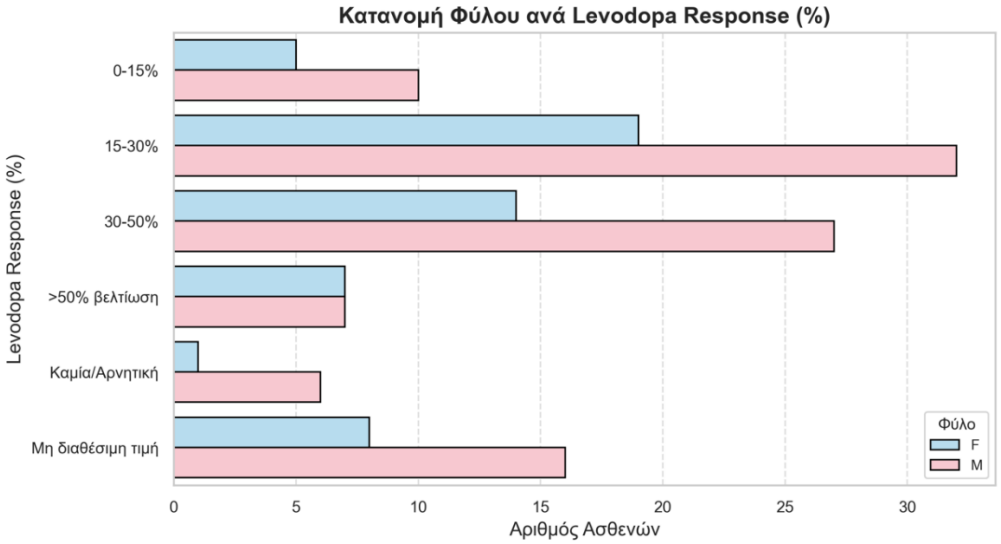
Εικόνα 19:Κατανομή του Φύλου με Βάση την Παρουσία Δυσκινησίας (Dyskinesia).

- Άμεση Ανταπόκριση στη Θεραπεία με Levodopa: τα δείγματα για την ποιοτική αξιολόγηση (UPDRS_III_dif_class) παρουσίαζαν κυρίως καλή ανταπόκριση στην θεραπεία (Excellent: 38 άντρες, 21 γυναίκες) αλλά υπήρχαν και αρκετά άτομα με υποβέλτιστη ανταπόκριση (Suboptimal: 31 άντρες, 22 γυναίκες). Πιο μικρό ποσοστό

παρουσιαζόταν στα άτομα με φτωχή ανταπόκριση στην θεραπεία (Poor: 13 άντρες, 3 γυναίκες). Τώρα για την ποσοστιαία μεταβολή (UPDRS_III_dif_%), το μεγαλύτερο ποσοστό των ασθενών παρουσίαζαν ποσοστό βελτίωσης 15-30% (32 άντρες, 19 γυναίκες), ενώ το μικρότερο ποσοστό εμφανίζεται στα άκρα, δηλαδή στα άτομα που είχαν βελτίωση υψηλότερη του 50%, καθώς και στα άτομα που εμφάνισαν καμία/αρνητική ανταπόκριση στην θεραπεία.



Εικόνα 20:Κατανομή του Φύλου ως προς την Ποιότητα Ανταπόκρισης στη Θεραπεία με Levodopa.



Εικόνα 21:Κατανομή του Φύλου ως προς την Ποσοστιαία Ανταπόκριση στη Θεραπεία με Levodopa.

Η αναλυτική παρουσίαση της κατανομής του δείγματος αποτελεί ένα βασικό υπόβαθρο για μια πιο σωστή κατανόηση και ερμηνεία των αποτελεσμάτων που ακολουθούν.

4.2 Επεξεργασία και Εξαγωγή Χαρακτηριστικών

4.2.1 Περιγραφή του Αρχείου και του Συνόλου Χαρακτηριστικών eGeMAPS

Τα φωνητικά χαρακτηριστικά, που αξιοποιήθηκαν για την διεξαγωγή της παρούσας μελέτης και εξήχθησαν βάσει του σετ eGeMAPS, προέκυψαν μέσω της επεξεργασίας των διαθέσιμων ηχητικών δεδομένων με την χρήση του εργαλείου openSMILE σε περιβάλλον Python. Το αρχείο που προέκυψε, περιλαμβάνει συνολικά 18.155 καταγραφές, που κάθε μια από αυτές αντιστοιχεί σε μια ηχογράφηση, καθώς και 113 στήλες. Από αυτές τις 113 στήλες, οι 90 αφορούν τα εξαγόμενα φωνητικά χαρακτηριστικά του eGeMAPS, τα οποία θα αξιοποιηθούν και στην μετέπειτα διαδικασία και οι υπόλοιπες 23 στήλες αφορούν δημογραφικά στοιχεία και κλινικά δεδομένα, που είναι εξίσου σημαντικά για την ανάλυση μας (φύλο, ηλικία, στάδιο Hoehn&Yahr κ.τ.λ.).

4.2.2 Κατηγορίες Χαρακτηριστικών

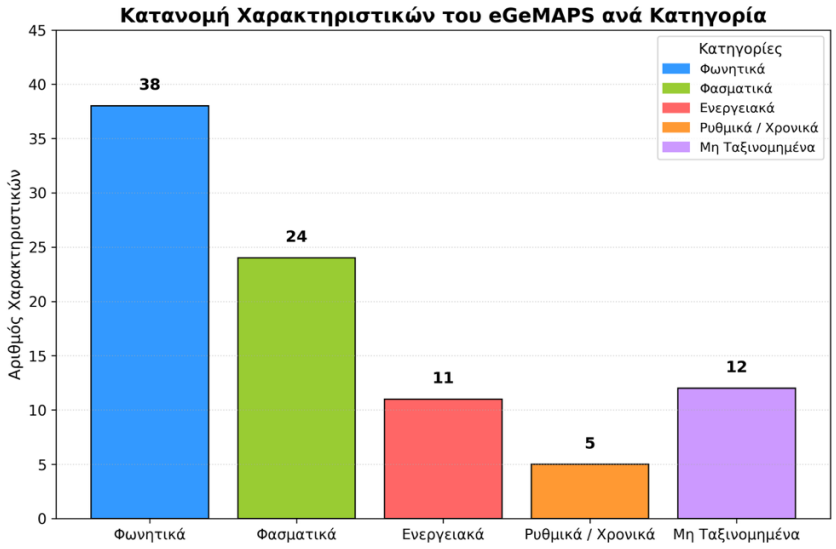
Το πλήθος των φωνητικών χαρακτηριστικών που εξήχθησαν, ανέρχονται στα 90. Αυτά τα 90 φωνητικά χαρακτηριστικά, που βασίζονται στο πρότυπο eGeMAPS, οργανώνονται σε τέσσερις βασικές κατηγορίες, ανάλογα με το είδος της πληροφορίας που παρέχουν. Προκειμένου να εξαχθούν οι χαρακτηριστικές τιμές, αρκετά από αυτά τα χαρακτηριστικά υπολογίζονται με την συμβολή στατιστικών συναρτήσεων και όχι απλών στιγμιαίων μετρήσεων. Γίνεται χρήση μετρικών όπως ο μέσος όρος, η τυπική απόκλιση, διάφορων ποσοστημόριων, καθώς και άλλων πιο εξειδικευμένων συναρτήσεων κλίσης (Slope-Related Functionals: στατιστικές συναρτήσεις που μετρούν άνοδο/κάθοδο μιας παραμέτρου σε ένα σύνολο τιμών με σκοπό την κατανόηση της δυναμικής μεταβολής ενός χαρακτηριστικού μέσα στον χρόνο). Με αυτόν το τρόπο, τα δεδομένα που εξάγονται είναι πιο αξιόπιστα και σταθερά και αποτυπώνουν με πιο σωστό τρόπο τις ιδιότητες τις φωνής (ένταση, ρυθμός, φασματική συμπεριφορά) [105].
Στον πίνακα που ακολουθεί (Πίνακας 12), παρουσιάζεται αναλυτικά η κατανομή των 90 χαρακτηριστικών του eGeMAPS ανά κατηγορία:

Πίνακας 12:Κατανομή των 90 φωνητικών χαρακτηριστικών του eGeMAPS ανά λειτουργική κατηγορία.

Κατηγορία Χαρακτηριστικών	Πλήθος	Περιγραφή
Φωνητικά	38	Παράμετροι που σχετίζονται με θεμελιώδη συχνότητα(F0), φορμαντικές συχνότητες(F1-F3), αστάθεια φωνής(jitter) και λόγο αρμονίας-θορύβου(HNR).
Φασματικά	24	Παράμετροι που σχετίζονται με την φασματική κατανομή και συμπεριφορά σήματος(MFCCs, spectral scope κ.τ.λ.).
Ενεργειακά	11	Παράμετροι που σχετίζονται με ένταση φωνής(loudness), μεταβλητότητα(shimmer) και συχνότητα ενεργειακών κορυφών(peaks per second).

Ρυθμικά/Χρονικά	5	Παράμετροι που σχετίζονται με χρονική ροή ομιλίας(ρυθμός άρθρωσης, διάρκεια φωνημάτων κ.τ.λ.).
Μη Ταξινομημένα	12	Βοηθητικά χαρακτηριστικά χωρίς άμεση αντιστοίχιση με μια από τις παραπάνω κατηγορίες(normalized energy values κ.τ.λ.).

Ακόλουθα, στο παρακάτω διάγραμμα (Εικόνα 22), παρουσιάζεται οπτικά η κατανομή των 90 φωνητικών χαρακτηριστικών ανά κατηγορία, με βάση την ταξινόμηση που έχει προηγηθεί



Εικόνα 22: Κατανομή Χαρακτηριστικών του eGeMAPS ανά Κατηγορία.

4.3 Στατιστική Ανάλυση Σημαντικότητας Χαρακτηριστικών

Μετά την εξαγωγή των 90 φωνητικών χαρακτηριστικών μέσω eGeMAPS, εφαρμόστηκε στοχευμένη στατιστική ανάλυση, με σκοπό τον εντοπισμό των χαρακτηριστικών που εμφανίζουν την υψηλότερη συσχέτιση με τις επιμέρους ασκήσεις ομιλίας-σετ ασκήσεων (Πίνακας 7) που πραγματοποίησαν οι συμμετέχοντες ασθενείς με ΝΠ. Για κάθε σετ ασκήσεων (A1,A2,A3,A4,A5,A6,A7,A8), αλλά και για την κάθε άσκηση ξεχωριστά υπολογίστηκαν αναλυτικά οι τιμές p-value για όλα τα διαθέσιμα χαρακτηριστικά φωνής.
Για την καλύτερη κατανόηση των αποτελεσμάτων, από τα συγκεντρωτικά αποτελέσματα, επιλέχθηκαν τα πέντε πιο σημαντικά χαρακτηριστικά με βάση την p-value τιμή, όπου το κατώφλι της στατιστικής σημαντικότητας θεωρήθηκε η τιμή 0.05. Τα χαρακτηριστικά που αναδείχθηκαν ως πιο σημαντικά παρουσιάζονται αναλυτικά στους παρακάτω πίνακες όπου αναγράφεται:

- Το όνομα του χαρακτηριστικού.
- Η τιμή p-value.
- Κατηγορία στην οποία ταξινομείται το εκάστοτε χαρακτηριστικό.

Παράλληλα, για το κάθε σετ ασκήσεων, καθώς και την κάθε άσκηση που εμπεριέχεται σε αυτό ξεχωριστά, παρουσιάζεται ένα ραβδόγραμμα που απεικονίζει οπτικά το πλήθος των σημαντικών χαρακτηριστικά ανά κατηγορία, τα οποία έχουν και την μεγαλύτερη στατιστική σημασία. Λόγω του μεγάλου αριθμού αποτελεσμάτων και γραφημάτων, η πλήρης παρουσίαση

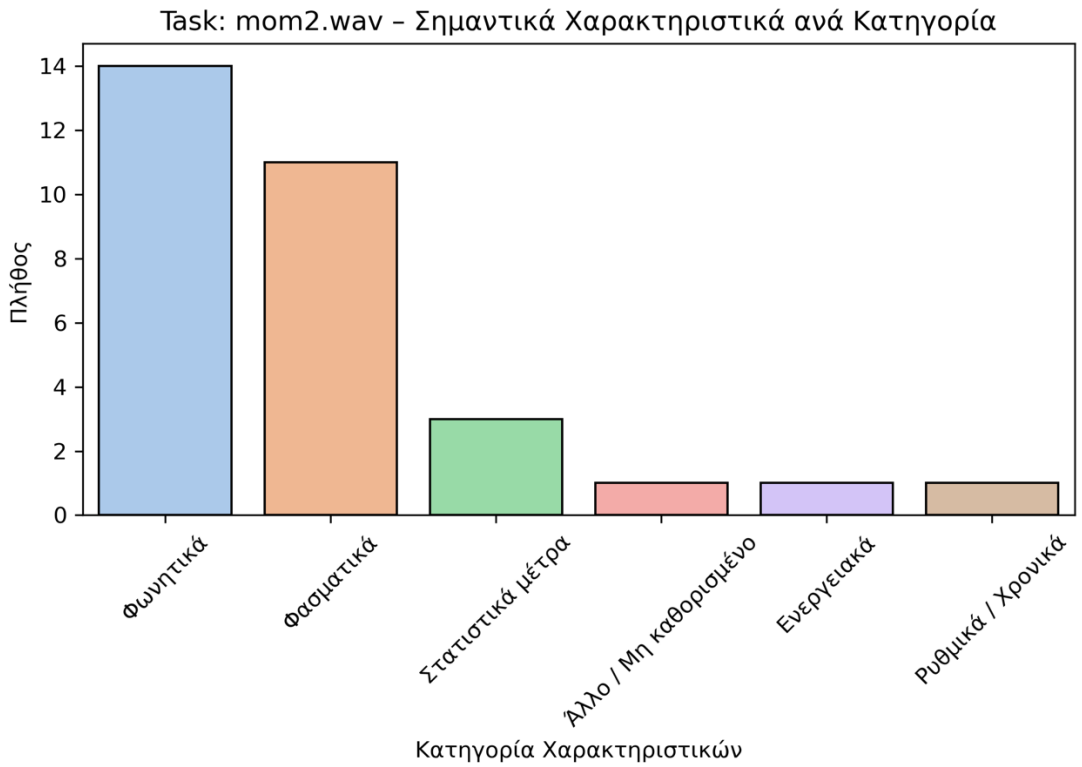
μεταφέρθηκε στο Παράρτημα Γ. Στο παρόν κεφάλαιο παρουσιάζονται μόνο τα αποτελέσματα που αφορούν τα πέντε βασικά κλινικά ερωτήματα, με αναφορά στην άσκηση που εμφάνισε τη σημαντικότερη στατιστική συσχέτιση ανά περίπτωση.

4.3.1 Συσχέτιση Εξαγόμενων Φωνητικών Χαρακτηριστικών με το Κλινικό Στάδιο Hoehn & Yahr (H_Y_binary)

Στην παρούσα ενότητα εξετάζεται η στατιστική συσχέτιση των φωνητικών χαρακτηριστικών με το στάδιο Hoehn & Yahr (H_Y_binary), το οποίο αποτελεί σημαντικό δείκτη εξέλιξης της ΝΠ. Προέκυψε ότι η άσκηση "mom2.wav" είχε την καλύτερη απόδοση με p-value κάτω του 0.05. Τα εν λόγω χαρακτηριστικά δείχνουν ότι ορισμένες φωνητικές διαφοροποιήσεις ίσως συσχετίζονται με το στάδιο της ΝΠ και η συγκεκριμένη φωνητική άσκηση ενδέχεται να είναι σημαντική για της διάκριση αυτής της κλινικής κατάστασης.

Πίνακας 13:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για την άσκηση mom2.wav του Task 2 (A2)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2bandwidth_sma3nz_amean	0.000046	Φωνητικά
F3frequency_sma3nz_amean	0.000076	Φωνητικά
F1bandwidth_sma3nz_amean	0.001174	Στατιστικά μέτρα
F0semitoneFrom27.5Hz_sma3nz_percentile20.0	0.005472	Φωνητικά
(Intercept)	0.005816	Άλλο / Μη καθορισμένο



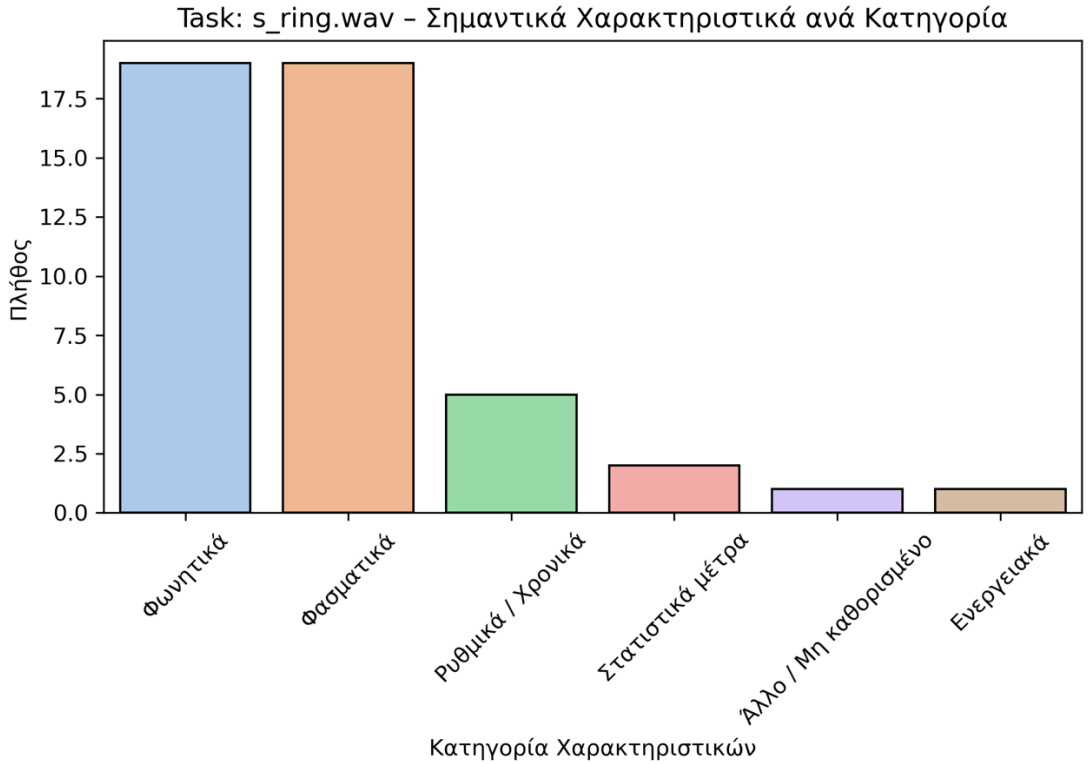
Εικόνα 23:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση mom2.wav του Task 2

4.3.2 Συσχέτιση Εξαγόμενων Φωνητικών Χαρακτηριστικών με την Παρουσία Δυσκινησίας (Dyskinesia)

Η δυσκινησία αποτελεί ένα από τα βασικά συμπτώματα της ΝΠ. Εδώ, η φωνητική άσκηση "s_ring.wav" παρουσίασε τα περισσότερα στατιστικά σημαντικά χαρακτηριστικά, υποδεικνύοντας ότι η συγκεκριμένη άσκηση συνδέεται με την ύπαρξη κινητικών διαταραχών, όπως οι ακούσιες κινήσεις που χαρακτηρίζουν τη δυσκινησία.

Πίνακας 14:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για την άσκηση s_ring.wav του Task 5 (A5)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F0semitoneFrom27.5Hz_sma3nz_percentile20.0	0.000093	Φωνητικά
F2bandwidth_sma3nz_amean	0.000235	Φωνητικά
mfcc2V_sma3nz_amean	0.002710	Φασματικά
jitterLocal_sma3nz_stddevNorm	0.003850	Φωνητικά
F3frequency_sma3nz_stddevNorm	0.005156	Φωνητικά



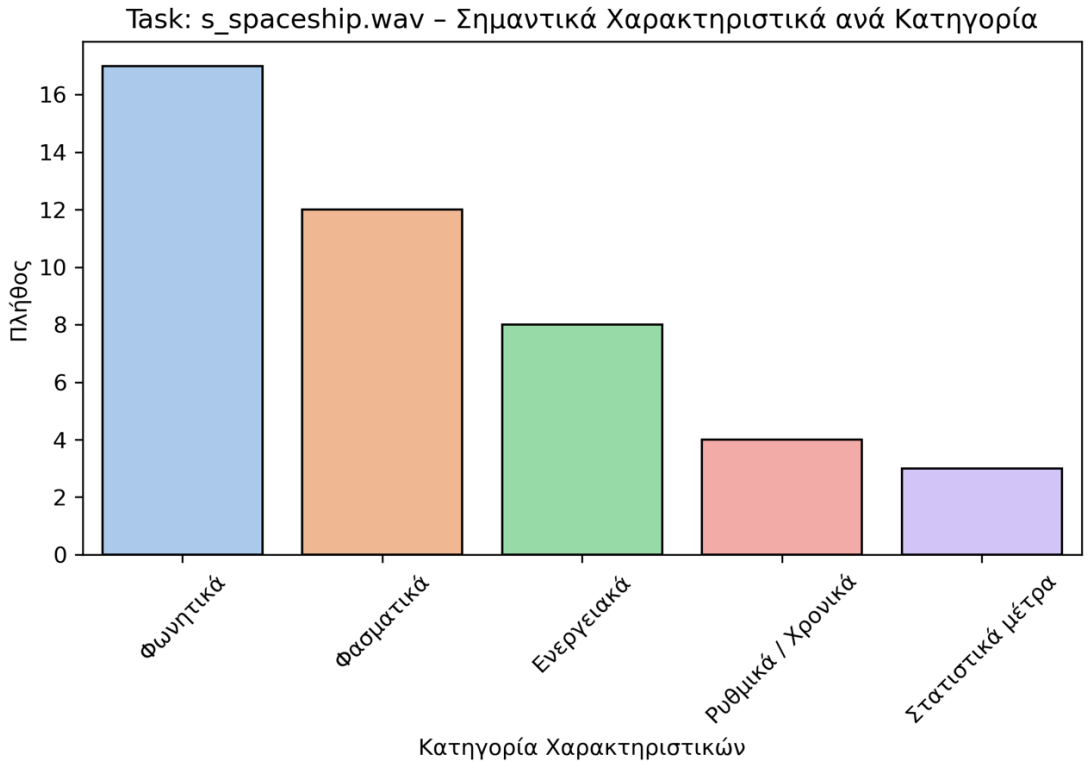
Εικόνα 24:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_ring.wav του Task 5

4.3.3 Συσχέτιση Εξαγόμενων Φωνητικών Χαρακτηριστικών με την Παρουσία Διακυμάνσεων (Fluctuatios)

Η παρουσία διακυμάνσεων (Fluctuations) αποτελεί σημαντικό στοιχείο για την διάγνωση της ΝΠ. Η άσκηση "s_spaceship.wav" φάνηκε να σχετίζεται στενά με αυτή την κλινική μεταβλητή, παρουσιάζοντας το μεγαλύτερο πλήθος χαρακτηριστικών με σημαντική p-value.

Πίνακας 15:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_spaceship.wav του Task 5 (A5)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2frequency_sma3nz_stddevNorm	0.000024	Φωνητικά
F1frequency_sma3nz_stddevNorm	0.000336	Φωνητικά
F0semitoneFrom27.5Hz_sma3nz_percentile20.0	0.001462	Φωνητικά
mfcc2_sma3_amean	0.003548	Φασματικά
mfcc1V_sma3nz_amean	0.004982	Φασματικά



Εικόνα 25:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση s_spaceship.wav του Task 5

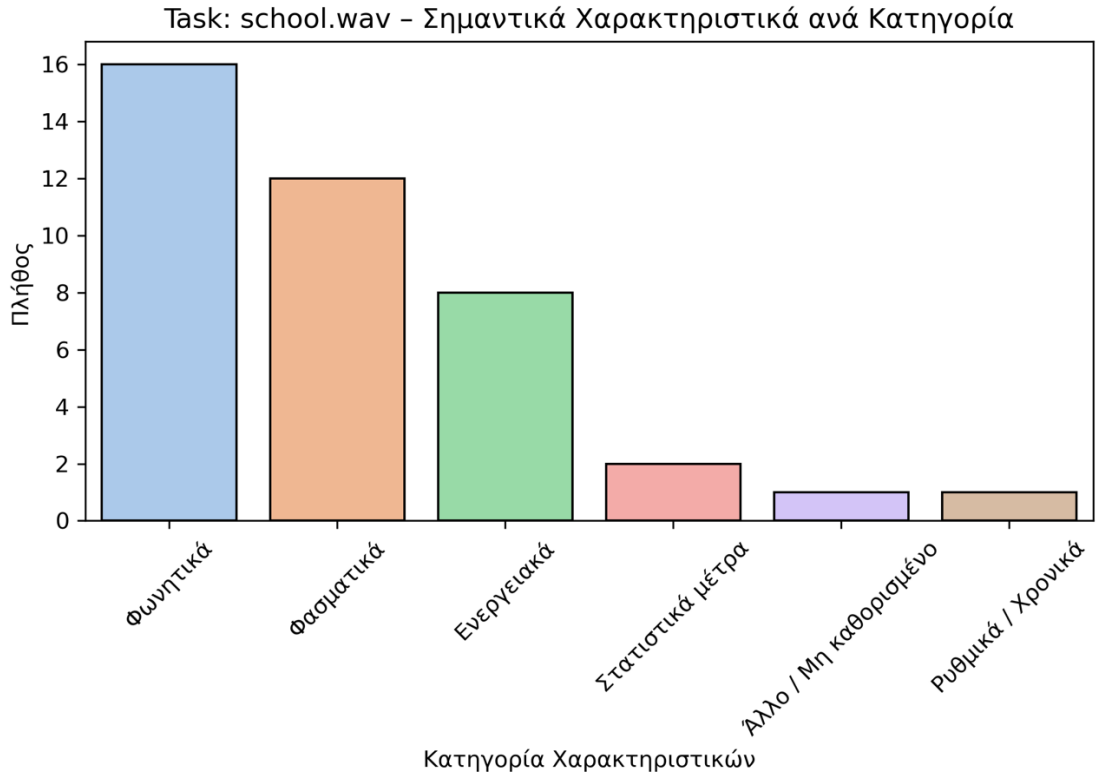
4.3.4 Συσχέτιση Εξαγόμενων Φωνητικών Χαρακτηριστικών με Κατηγορίες Βαθμολογίας UPDRS-III (UPDRS_III_dif_class)

Εδώ, διερευνάται η σχέση μεταξύ των φωνητικών χαρακτηριστικών και της κατηγοριοποιημένης διαφοράς στην κινητική κλίμακα UPDRS III. Φάνηκε ότι η άσκηση "school.wav" συγκέντρωσε τον μεγαλύτερο αριθμό χαρακτηριστικών με υψηλή στατιστική σημαντικότητα. Η διαφοροποίηση της φωνητικής απόδοσης για αυτή την άσκηση, σε συνδυασμό με την επιδείνωση της κατάστασης του ασθενή, υποδηλώνει ότι ίσως η παρούσα φωνητική άσκηση είναι χρήσιμη για την έγκαιρη εκτίμηση την πορείας της υγείας του ατόμου.

Πίνακας 16:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για την άσκηση school.wav του Task 2 (A2)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
loudness_sma3_percentile50.0	0.000016	Ενεργειακά
loudness_sma2_percentile50.0	0.000302	Ενεργειακά
F0semitoneFrom27.5Hz_sma3nz_percentile50.0	0.000677	Φωνητικά

F0semitoneFrom27.5Hz_sma3nz_stddevNorm	0.000988	Φωνητικά
F1amplitudeLogRelF0_sma3nz_stddevNorm	0.001044	Φωνητικά



Εικόνα 26:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση school.wav του Task 2

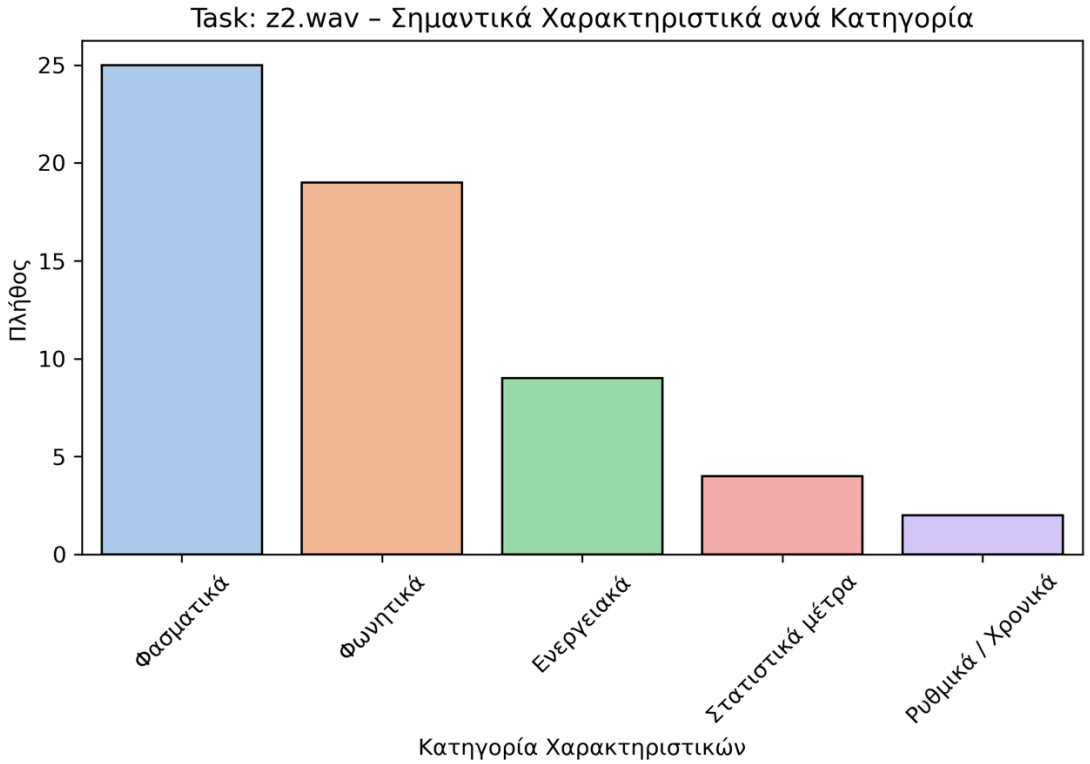
4.3.5 Συσχέτιση Εξαγόμενων Φωνητικών Χαρακτηριστικών με Ποσοστιαία Διαφορά Βαθμολογίας UPDRS-III (UPDRS_III_dif_per)

Τέλος, αναλύεται η ποσοστιαία διαφορά στην κλίμακα UPDRS III και η συσχέτιση της με της με την φωνητική απόδοση. Η άσκηση "z2.wav" αναδείχθηκε ως η πιο σημαντική εμφανίζοντας πολλά στατιστικά σημαντικά χαρακτηριστικά. Αυτό υποδηλώνει ότι ακόμα και μικρές ποσοτικές μεταβολές στην κινητική λειτουργία μπορούν να ανιχνευτούν μέσω διακριτών φωνητικές διαφοροποιήσεων.

Πίνακας 17:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση z2.wav του Task 1 (A1)

Χαρακτηριστικό (Predictor)	**P Value**	**Κατηγορία**

jitterLocal_sma3nz_amean	0.000005	Φωνητικά
jitterLocal_sma3nz_amean	0.000279	Φωνητικά
mfcc1_sma3_stddevNorm	0.001416	Φασματικά
StddevVoicedSegmentLengthSec	0.002607	Ρυθμικά/Χρονικά
loudness_sma3_meanRisingSlope	0.002778	Ενεργειακά



Εικόνα 27:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση z2.wav του Task 1

.

4.4 Συγκεντρωτική Ανάλυση της Συχνότητας Εμφάνισης των Χαρακτηριστικών Ομιλίας

Η γραφική απεικόνιση που ακολουθεί παρουσιάζει τα 20 χαρακτηριστικά ομιλία που εμφανίζονται με την μεγαλύτερη συχνότητα στην ανάλυση μας, με βάση τις τιμές του p-value (p-value < 0.05). Η καταγραφή της συχνότητας εμφάνισης προέκυψε μέσω της ανάλυσης της επαναληπτικότητας των επιμέρους χαρακτηριστικών για όλες τις ασκήσεις ομιλίας που πραγματοποιήθηκαν.

Η παρακάτω απεικόνιση μας προσφέρει σημαντικές πληροφορίες σχετικά με την κατανομή των χαρακτηριστικών στο σύνολο δεδομένων και για τον πιο εύκολο εντοπισμό δεικτών με την μεγαλύτερη σταθερότητα μεταξύ των επιμέρους ασκήσεων. Έτσι, το διάγραμμα αυτό θεωρείται ως ένα εργαλείο που εντοπίζει τα χαρακτηριστικά με την υψηλότερη διαγνωστική αξία και συμβάλει σε μια πιο εκτενής ποιοτική και ποσοτική ανάλυση των ευρημάτων, η οποία ακολουθεί στο επόμενο κεφάλαιο.



Εικόνα 28: Τα 20 Χαρακτηριστικά Ομιλίας με τη Συχνότερη Εμφάνιση στα Σημαντικά Ευρήματα.

4.5 Απόδοση Ταξινομητικών Μοντέλων στις Κατηγορίες Εκτίμησης της Νόσου

Μετά την εκτενής παρουσίαση των χαρακτηριστικών που εξήχθησαν μέσω της ανάλυσης των φωνητικών δεδομένων που προήρθαν από τα διαφορετικά σετ ασκήσεων που κλήθηκαν οι ασθενείς να εκτελέσουν, ακολουθεί η αξιολόγηση της ικανότητας των ταξινομητικών μοντέλων να προβλέπουν κλινικές παραμέτρους που συνδέονται με την ΝΠ. Η ενότητα που ακολουθεί επικεντρώνεται στην συγκεντρωτική ανάλυση της απόδοσης των μοντέλων που χρησιμοποιήθηκαν στα πλαίσια της παρούσας διπλωματικής εργασίας, προκειμένου να αξιολογηθούν οι πέντε βασικές κατηγορίες που συσχετίζονται με την ΝΠ όπως η βαρύτητα της νόσου βάση της κλίμακας H&Y, η παρουσία ή όχι διακυμάνσεων ή δυσκινησιών, καθώς και ποσοτική/ποιοτική μεταβολή της κλίμακας UPDRS III.
Για καθεμία από αυτές τις κατηγορίες γίνεται παρουσίαση των κύριων μετρικών απόδοσης, που είναι οι ακόλουθοι:

- Η ακρίβεια ταξινόμησης (Accuracy)
- Η ακρίβεια θετικής πρόβλεψης (Precision)
- Η ευαισθησία (Recall)
- Ο σταθμισμένος μέσος όρος (F1-Score)
- Όπου είναι εφικτό, η επιφάνεια κάτω από την καμπύλη ROC (ROC_AUC)

Στην μελέτη, προκειμένου να πραγματοποιηθεί η αξιολόγηση της εκάστοτε κατηγορίας, έγινε χρήση έξι ταξινομητικών αλγορίθμων, οι οποίοι είναι:

- K-Nearest Neighbors (KNN)
- Random Forest
- Support Vector Machine (SVM)
- Logistic Regression
- XGBoost
- Spiking Neural Network (SNN)

Η αξιολόγηση όλων των μοντέλων βασίστηκε στην χρήση της τεχνικής Leave-One-Subject-Out (LOSO), ως μέθοδος διασταυρούμενης επικύρωσης (cross-validation) προκειμένου τα εξαγόμενα αποτελέσματα να είναι αντικειμενικά και να αποφευχθεί το ενδεχόμενο υπερεκπαίδευσης(overfitting). Αυτό, διότι η αξιολόγηση του κάθε ασθενή προκύπτει αποκλειστικά σε δεδομένα που δεν συμμετείχαν στην εκπαίδευση του μοντέλου και χρησιμοποιούνταν μόνο για την αξιολόγηση.
Στις υποενότητες που ακολουθούν, παρουσιάζονται τα αποτελέσματα των συνολικών επιδόσεων για καθεμία από τις πέντε εξεταζόμενες κατηγορίες. Οι πίνακες περιλαμβάνουν τις αποδόσεις των έξι ταξινομητικών μοντέλων για κάθε κατηγορία, βάση των μετρικών απόδοσης. Οι πίνακες συνοδεύονται και με συγκριτικά γραφήματα που ενισχύουν την οπτικοποίηση και κατανόηση των αποτελεσμάτων, διευκολύνοντας την σύγκριση μεταξύ των μοντέλων για μια πιο σφαιρική αξιολόγηση των αποτελεσμάτων.

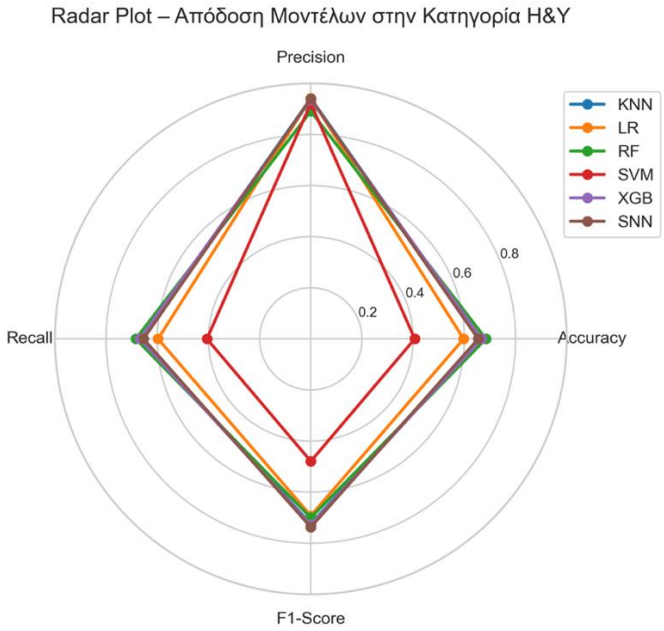
4.5.1 Απόδοση Μοντέλων για την Κατηγορία Hoehn & Yahr (H&Y)

Η συγκεκριμένη κατηγορία σχετίζεται με τα στάδια και κατά συνέπεια την βαρύτητα της νόσου με βάση την κλίμακα Hoehn & Yahr, η οποία ουσιαστικά αξιολογεί το στάδιο της κινητικής δυσλειτουργίας. Τα μοντέλα κλήθηκαν να διαχωρίσουν το πλήθος των ασθενών στα διάφορα επίπεδα βαρύτητας της νόσου, με βάση τα εξαχθείσα χαρακτηριστικά ομιλίας που προέκυψαν από τα ηχητικά δεδομένα. Η επίδοση του κάθε αλγόριθμου εξετάσθηκε με γνώμονα τους βασικούς δείκτες απόδοσης, που αναφέρθηκαν παραπάνω, οι οποίοι προέκυψαν έπειτα από την εφαρμογή της επαναληπτικής διαδικασίας LOSO.
Τα αποτελέσματα που προέκυψαν παρατίθενται παρακάτω, συνοδευόμενα από την αντίστοιχη οπτική αναπαράσταση αυτών.

Πίνακας 18: Συγκριτική Παρουσίαση των Μετρικών Απόδοσης των Μοντέλων για την Κατηγορία Hoehn & Yahr.

Μοντέλο	Accuracy	Precision	Recall	F1-Score	ROC-AUC
K-Nearest Neighbors	0.663	0.941	0.663	0.724	0.655
Logistic Regression	0.597	0.942	0.597	0.692	0.543
Random Forest	**0.684**	0.891	0.684	0.700	0.538

Support Vector Machine	0.406	0.921	0.406	0.479	0.546
XGBoost	0.667	0.936	0.668	0.728	0.538
Spiking Neural Network	0.654	0.940	0.654	0.737	-



Εικόνα 29:Radar Plot – Συγκριτική Απόδοση Μοντέλων για την Κατηγορία Hoehn & Yahr (H&Y)

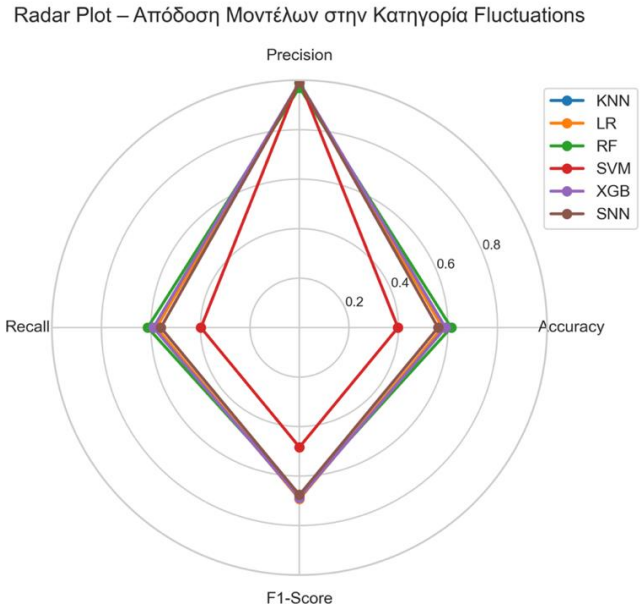
4.5.2 Απόδοση Μοντέλων για την Κατηγορία Fluctuations

Η κατηγορία Fluctuations αφορά την παρουσία ή όχι κινητικών διακυμάνσεων που συνδέονται με την ομιλία, σε άτομα που νοσούν από την ΝΠ. Οι αξιοποιούμενοι αλγόριθμοι ως στόχο έχουν να εντοπίσουν, μέσω των χαρακτηριστικών ομιλίας, διάφορα μοτίβα φωνής που να αποδεικνύουν μεταβολές στην συνολική κινητική λειτουργία του ασθενή. Όπως και στην πρώτη περίπτωση, έτσι και εδώ η αξιολόγηση στηρίζεται στα ηχητικά δεδομένα και υλοποιείται με την χρήση της μεθόδου LOSO.Ακολουθεί η συγκεντρωτική απεικόνιση των επιδόσεων των αξιολογούμενων μοντέλων πλαισιωμένη από τα σχετικά γραφήματα.

Πίνακας 19: Συγκριτική Παρουσίαση των Μετρικών Απόδοσης των Μοντέλων για την Κατηγορία Fluctuations.

Μοντέλο	Accuracy	Precision	Recall	F1-Score	ROC-AUC
K-Nearest Neighbors	0.587	0.990	0.587	0.688	-

Logistic Regression	0.578	0.990	0.578	0.693	-
Random Forest	**0.613**	0.970	0.613	0.675	-
Support Vector Machine	0.398	1.000	0.398	0.483	-
XGBoost	0.590	0.990	0.590	0.688	-
Spiking Neural Network	0.560	0.990	0.560	0.676	-



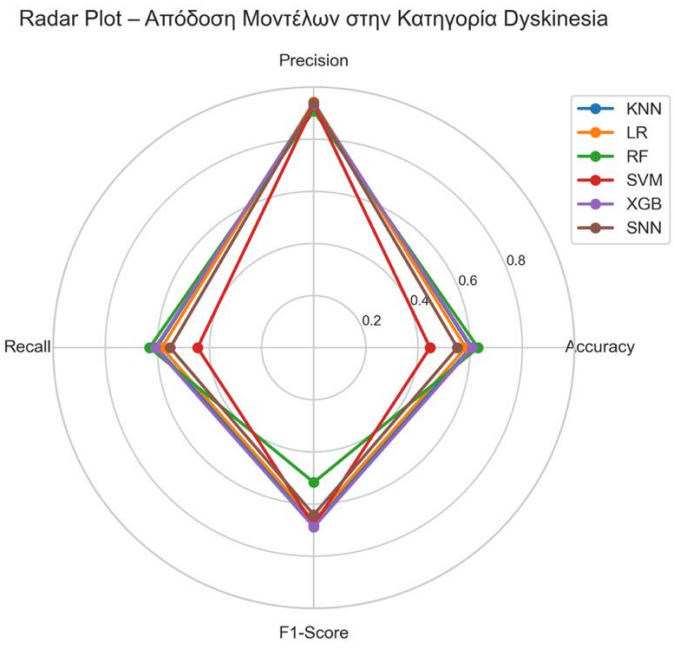
Εικόνα 30: Radar Plot – Συγκριτική Απόδοση Μοντέλων για την Κατηγορία Κινητικών Διακυμάνσεων (Fluctuations)

4.5.3 Απόδοση Μοντέλων για την Κατηγορία Dyskinesia

Η κατηγορία αυτή εστιάζει κυρίως στην αναγνώριση της παρουσίας ή όχι δυσκινησίας, που σαν σύμπτωμα της νόσου εμφανίζεται σε πιο προχωρημένο στάδιο αυτής. Όπως και παραπάνω, η ταξινόμηση πραγματοποιείται βασιζόμενη στα ηχητικά δείγματα και η επίδοση των μοντέλων αξιοποιείται ως προς την ικανότητα τους να διαχωρίζουν με επιτυχία αυτές τις δυο καταστάσεις στο πλήθος των ασθενών.
Τα αναλυτικά αποτελέσματα της αξιολόγησης παρατίθενται παρακάτω, συνοδευόμενα από διαγράμματα που έχουν ως σκοπό την ενίσχυση της κατανόησης των επιδόσεων των μοντέλων.

Πίνακας 20: Συγκριτική Παρουσίαση των Μετρικών Απόδοσης των Μοντέλων για την Κατηγορία Dyskinesia.

Μοντέλο	Accuracy	Precision	Recall	F1-Score	ROC-AUC
K-Nearest Neighbors	0.603	0.938	0.603	0.668	0.611
Logistic Regression	0.583	0.943	0.583	0.664	0.555
Random Forest	**0.630**	0.908	0.630	0.517	0.526
Support Vector Machine	0.446	0.930	0.446	0.682	0.478
XGBoost	0.608	0.934	0.608	0.688	0.535
Spiking Neural Network	0.551	0.939	0.551	0.641	0.570



Εικόνα 31: Radar Plot – Συγκριτική Απόδοση Μοντέλων για την Κατηγορία Δυσκινησίας (Dyskinesia)

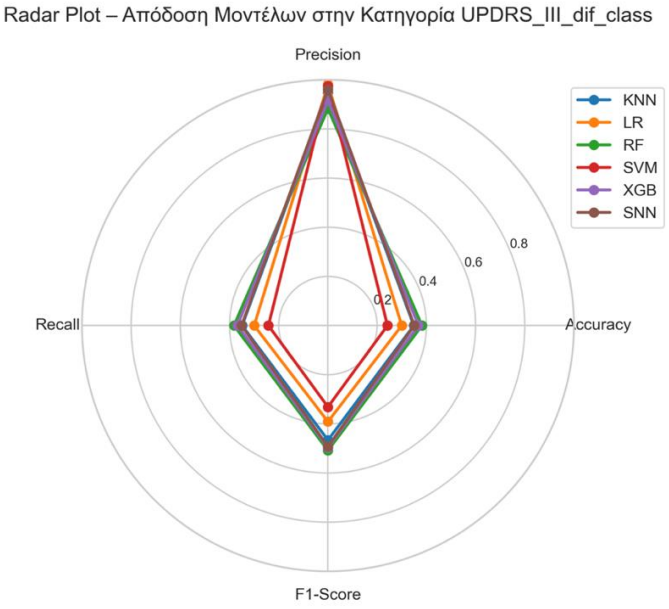
4.5.4 Απόδοση Μοντέλων για UPDRS_III_dif_class

Η τέταρτη κατηγορία αφορά την ικανότητα των μοντέλων να είναι σε θέση να ταξινομήσουν τους ασθενείς στις εκάστοτε κατηγορίες με βάση την διαφορά της βαθμολογίας UPDRS_III πριν και μετά την λήψη της αγωγής. Η πρόβλεψη αυτή αφορά κυρίως τον εκτοπισμό της βελτίωσης ή επιδείνωσής της συνολικής εικόνας των ασθενών.

Ακολουθεί η παρουσίαση των αντίστοιχων πινάκων με τις επιδόσεις των διαφόρων μοντέλων, συνοδευόμενο από γραφικές απεικονίσεις για ενίσχυση της κατανόησης των αποτελεσμάτων.

Πίνακας 21: Συγκριτική Παρουσίαση των Μετρικών Απόδοσης των Μοντέλων για την Κατηγορία UPDRS_III_dif_class.

Μοντέλο	Accuracy	Precision	Recall	F1-Score	ROC-AUC
K-Nearest Neighbors	0.349	0.948	0.349	0.466	-
Logistic Regression	0.300	0.981	0.300	0.391	-
Random Forest	**0.382**	0.883	0.382	0.508	-
Support Vector Machine	0.242	0.974	0.242	0.331	-
XGBoost	0.368	0.916	0.368	0.496	-
Spiking Neural Network	0.351	0.961	0.351	0.492	-



Εικόνα 32: Radar Plot – Συγκριτική Απόδοση Μοντέλων για τη Διακριτή Μεταβολή του UPDRS III (UPDRS_III_dif_class)

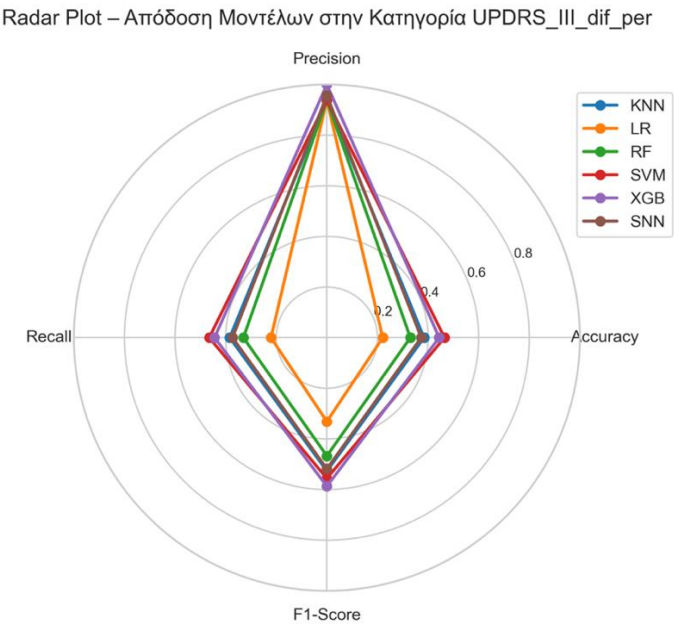
4.5.5 Απόδοση Μοντέλων για UPDRS_III_dif_per

Η τελευταία κατηγορία επικεντρώνεται στην πρόβλεψη της ποσοστιαίας μεταβολής της βαθμολογίας UPDRS_III, η οποία πρώτα έχει μετατραπεί σε διακριτή κλίμακα με βάση το

ποσοστό μείωσης ή αύξησης του ποσοστού. Μέσω αυτής της προσέγγισης, επιτρέπεται η συσχέτιση μεταξύ των μεταβολών στα ποσοστά βελτίωσης ή επιδείνωσής της νόσου, με βάση τα χαρακτηριστικά ομιλίας.
Τα αποτελέσματα που προέκυψαν καταγράφονται παρακάτω, συνοδευόμενα από τα αντίστοιχα διαγράμματα σύγκρισης μεταξύ των μοντέλων ταξινόμησης.

Πίνακας 22: Συγκριτική Παρουσίαση των Μετρικών Απόδοσης των Μοντέλων για την Κατηγορία UPDRS_III_dif_per.

Μοντέλο	Accuracy	Precision	Recall	F1-Score	ROC-AUC
K-Nearest Neighbors	0.384	0.938	0.384	0.527	-
Logistic Regression	0.221	0.938	0.221	0.331	-
Random Forest	0.330	0.938	0.330	0.467	-
Support Vector Machine	**0.465**	0.938	0.465	0.556	-
XGBoost	0.444	1.000	0.444	0.587	-
Spiking Neural Network	0.373	0.957	0.373	0.518	-



Εικόνα 33: Radar Plot – Συγκριτική Απόδοση Μοντέλων για την Ποσοστιαία Μεταβολή του UPDRS III (UPDRS_III_dif_per)

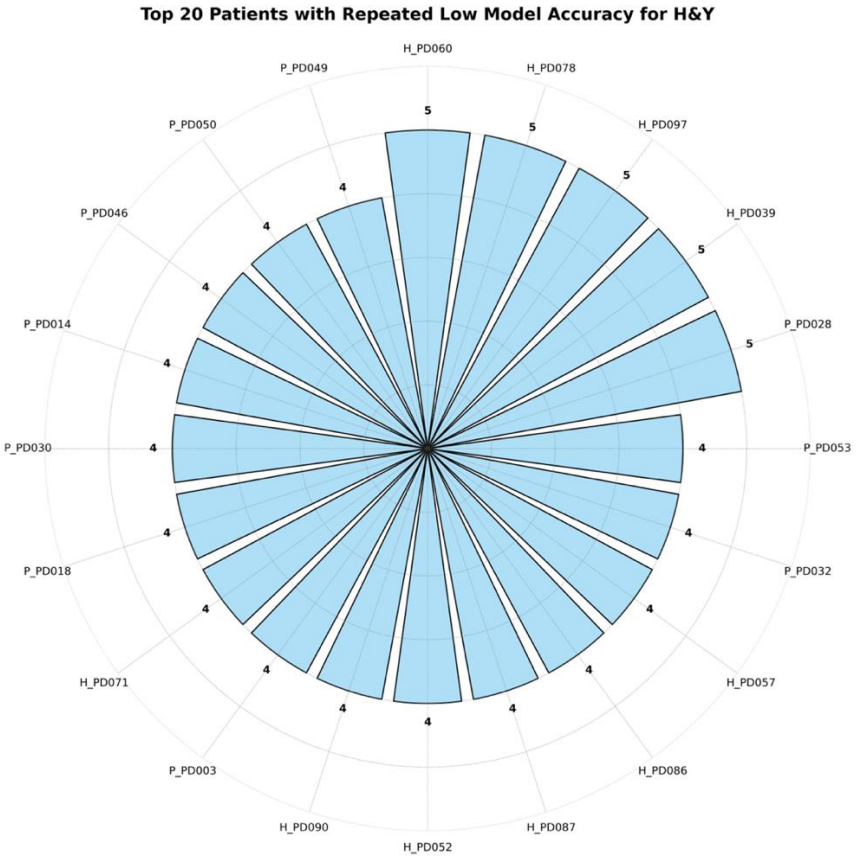
4.6 Εντοπισμός Ασθενών με Συστηματικά Χαμηλή Ακρίβεια Ταξινόμησης

Στην παρούσα ενότητα εξετάζεται η απόδοση του μοντέλου σε επίπεδο ασθενών με έμφαση στις περιπτώσεις των ασθενών όπου παρατηρήθηκε συστηματική χαμηλή ακρίβεια. Η αξιολόγηση βασίζεται στις πέντε διαφορετικές βαθμολογικές κλίμακες που σχετίζονται με την ΝΠ:

- Hoehn & Yahr (H&Y)
- Dyskinesia
- Fluctuations
- UPDRS_III_dif_class
- UPDRS_III_dif_per

Στην συνέχεια, παρατίθενται οι 20 ασθενείς με τις περισσότερες αποτυχίες ταξινόμησης ανά κατηγορία , προκειμένου να γίνει αναγνώριση πιθανών υποομάδων ή μεμονωμένων περιπτώσεων που αποτελούν πρόκληση για το εκάστοτε μοντέλο. Η απεικόνιση των αποτελεσμάτων επιλέχθηκε να γίνει μέσω κυκλικών ραβδογραμμάτων (radial bar charts), ώστε η σύγκριση να είναι ευκολότερη.

4.6.1 Ασθενείς με Επαναλαμβανόμενα Χαμηλή Ακρίβεια – Βαθμολογική Κλίμακα Hoehn & Yahr (H&Y)



Εικόνα 34: Κατανομή Ασθενών με Χαμηλή Απόδοση Μοντέλου στην Κατηγορία H&Y.

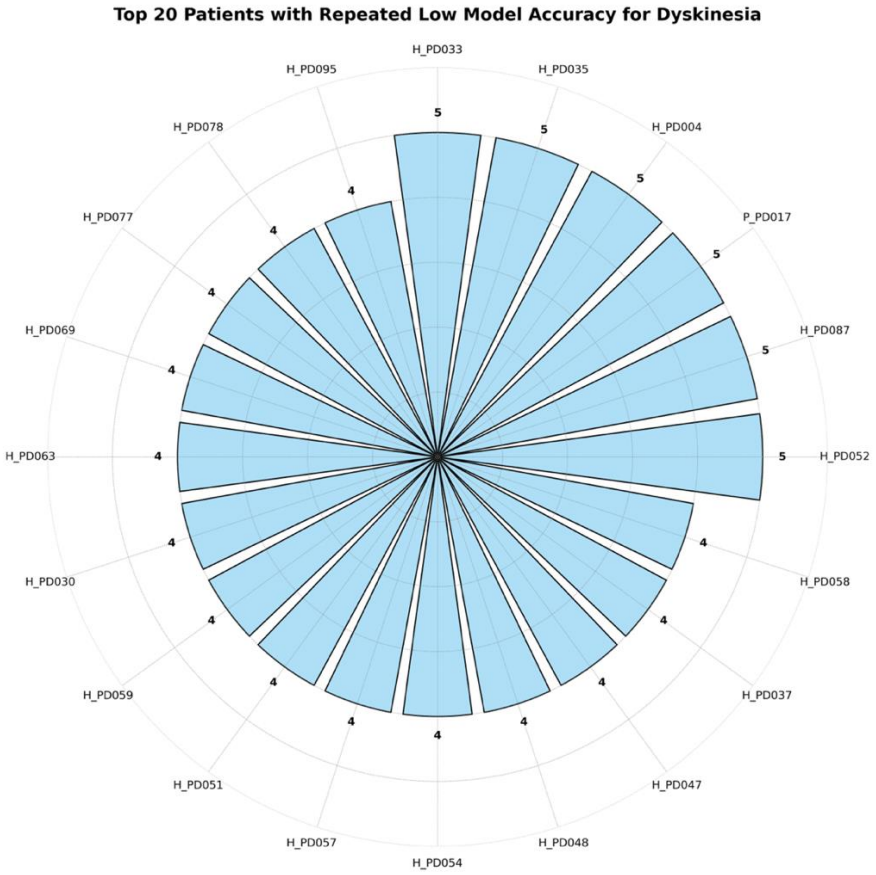
Για την κλίμακα Hoehn & Yahr, η οποία χρησιμοποιείται για την εκτίμηση της γενικής σοβαρότητας της νόσου, παρατηρήθηκε ότι οι ασθενείς:

- H_PD060

- H_PD078
- H_PD097
- P_PD028
- H_PD039

συγκέντρωσαν συνολικά πέντε αποτυχημένες προβλέψεις, γεγονός το οποίο υποδεικνύει μια αδυναμία του μοντέλου να διακρίνει τα διαφορετικά στάδια εξέλιξης για αρκετούς ασθενείς με βάση φωνητικά χαρακτηριστικά. Αποτυχία ταξινόμησης παρουσίασαν και άλλοι ασθενείς, που φαίνονται στο διάγραμμα. Η συχνότητα εμφάνισης αποτυχιών σε ασθενείς τόσο προχωρημένων όσο και πρώιμων σταδίων ενδεχομένως σχετίζεται με αλληλοεπικαλυπτόμενα φωνητικά μοτίβα ή ελλιπή διάκριση των χαρακτηριστικών στη φάση προεπεξεργασίας.

4.6.2 Ασθενείς με Επαναλαμβανόμενα Χαμηλή Ακρίβεια – Βαθμολογική Κλίμακα Dyskinesia



Εικόνα 35:Κατανομή Ασθενών με Χαμηλή Απόδοση Μοντέλου στην Κατηγορία Dyskinesia.

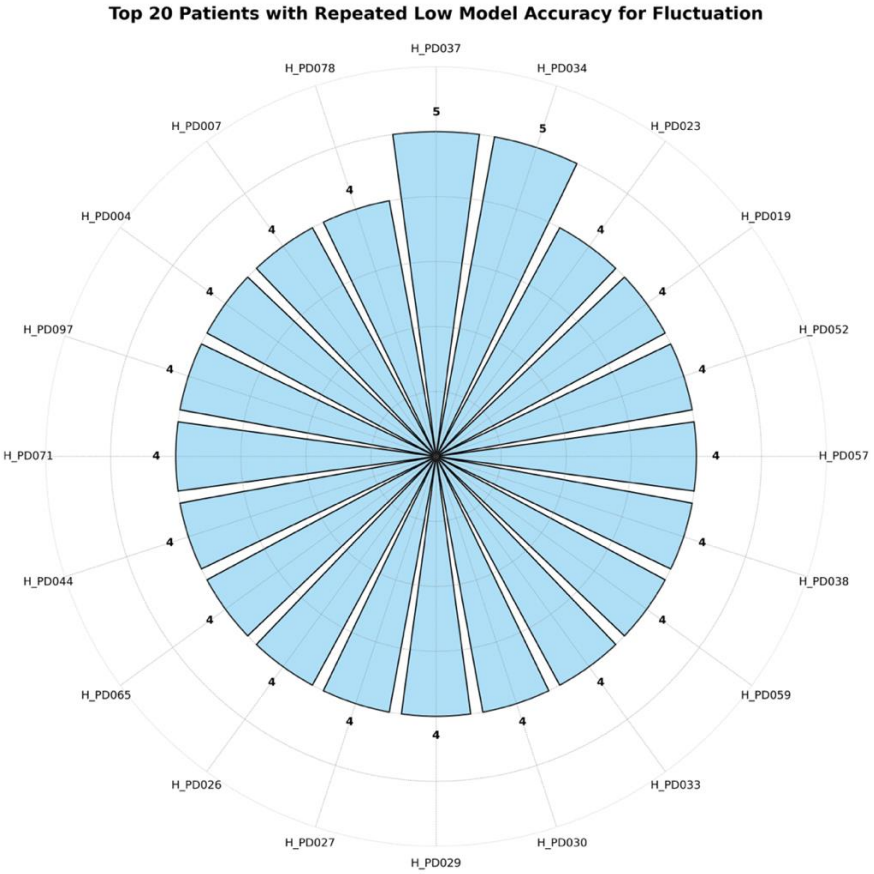
Στην συγκεκριμένη κατηγορία εξετάζεται η παρουσία ή απουσία δυσκινησίας. Δηλαδή ακούσιων κινήσεων. Σε αυτήν την περίπτωση οι ασθενείς που εμφάνισαν την χαμηλότερη επιτυχία ταξινόμησης, έως και πέντε φορές είναι:

- H_PD033
- H_PD035
- H_PD004
- P_PD017

- H_PD087

Πολλοί από αυτούς τους ασθενείς παρουσίασαν αποτυχία ταξινόμησης και σε άλλες κατηγορίες, οπότε υπάρχει μια σταθερή δυσκολία του μοντέλου να προβλέψει επιτυχώς περιπτώσεις ορισμένων ασθενών.

4.6.3 Ασθενείς με Επαναλαμβανόμενα Χαμηλή Ακρίβεια – Βαθμολογική Κλίμακα Fluctuations



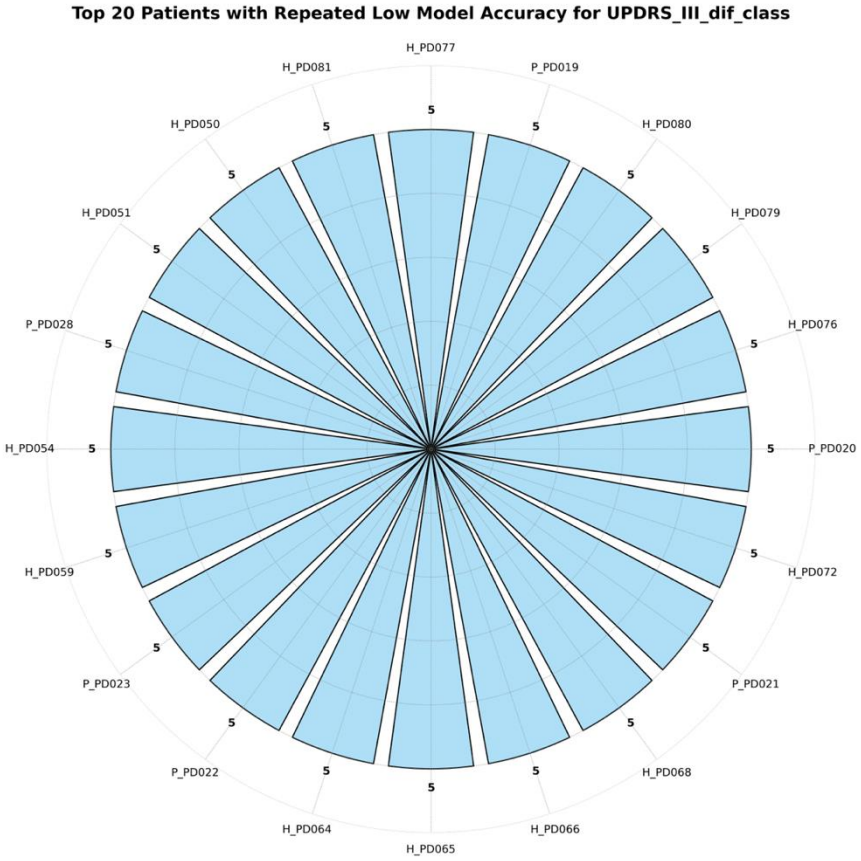
Εικόνα 36:Κατανομή Ασθενών με Χαμηλή Απόδοση Μοντέλου στην Κατηγορία Fluctuations.

Η κατηγορία Fluctuations αξιολογεί την παρουσία ή απουσία διακυμάνσεων στους ασθενείς που νοσούν από ΝΠ. Όπως και στις προηγούμενες περιπτώσεις, έτσι και εδώ, αρκετοί ασθενείς εμφάνισαν χαμηλή επιτυχία ταξινόμησης, με αυτούς που είχαν τις περισσότερες αποτυχίες να είναι:

- H_PD037
- H_PD034
- H_PD023
- H_PD019
- H_PD052

Παρατηρείται ότι αρκετοί ασθενείς εμφανίζονται και στις παραπάνω κατηγορίες και πιο συγκεκριμένα ο ασθενής H_PD052 και στις τρεις, στοιχείο που δείχνει γενικό πρόβλημα πρόβλεψης για αυτόν τον ασθενή.

4.6.4 Ασθενείς με Επαναλαμβανόμενα Χαμηλή Ακρίβεια – Βαθμολογική Κλίμακα UPDRS_III_dif_class



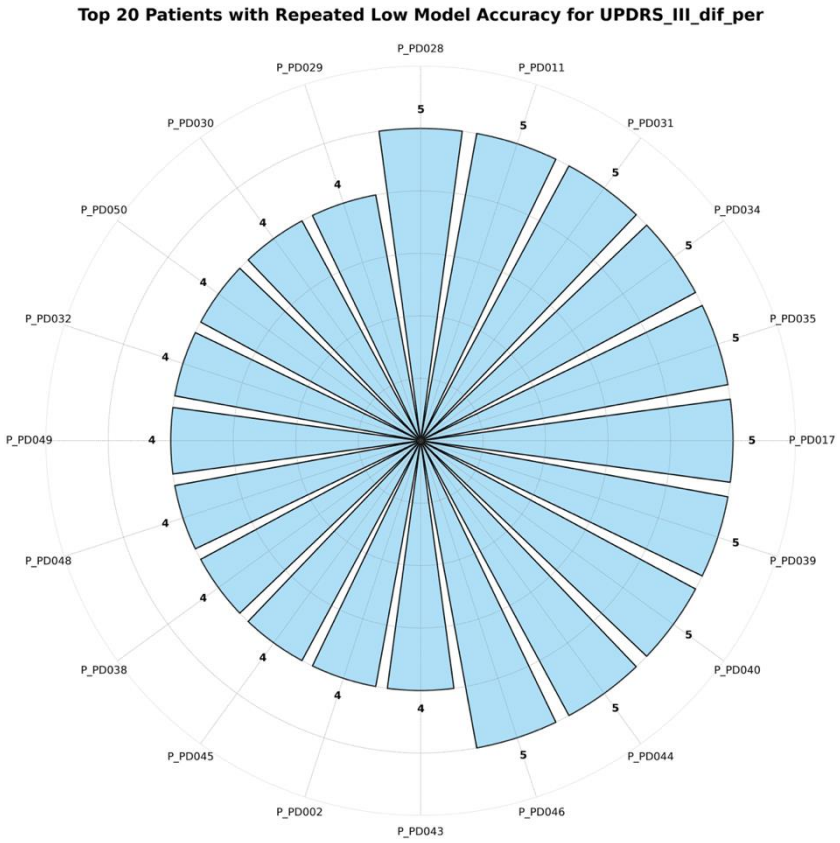
Εικόνα 37: Κατανομή Ασθενών με Χαμηλή Απόδοση Μοντέλου στην Κατηγορία UPDRS_III_dif_class.

Στην κατηγορία αυτή εξετάζεται η κατανομή των ασθενών με βάση της μεταβολή της κινητικής βαθμολογίας UPDRS III σε κλάσεις (Excellent, Suboptimal, Poor). Οι ασθενείς που εμφάνισαν τις χαμηλότερες επιδόσεις είναι:

- P_PD028
- P_PD011
- P_PD031
- P_PD034
- P_PD035

Το μοντέλο φαίνεται να αποτυγχάνει επανειλημμένα σε ασθενείς με μικρές ή δύσκολες στη φωνητική αποτύπωση μεταβολές κινητικότητας.

4.6.5 Ασθενείς με Επαναλαμβανόμενα Χαμηλή Ακρίβεια – Βαθμολογική Κλίμακα UPDRS_III_dif_per



Εικόνα 38: Κατανομή Ασθενών με Χαμηλή Απόδοση Μοντέλου στην Κατηγορία UPDRS_III_dif_per.

Τέλος, η κατηγορία UPDRS_III_dif_per, εξετάζει την ποσοστιαία διαφορά στην κινητική βαθμολογία UPDRS III μεταξύ μετρήσεων. Εδώ, οι ασθενείς με την χαμηλότερη επιτυχία ταξινόμησης είναι:

- P_PD077
- P_PD019
- H_PD080
- P_PD079
- P_PD076

Οι περισσότεροι από αυτούς τους ασθενείς δεν εμφανίζονται στις υπόλοιπες κατηγορίες, άρα ενδεχομένως η ποσοστιαία διαφορά επηρεάζει τη φωνή με τρόπο που δεν εντοπίζεται συστηματικά από το μοντέλο.

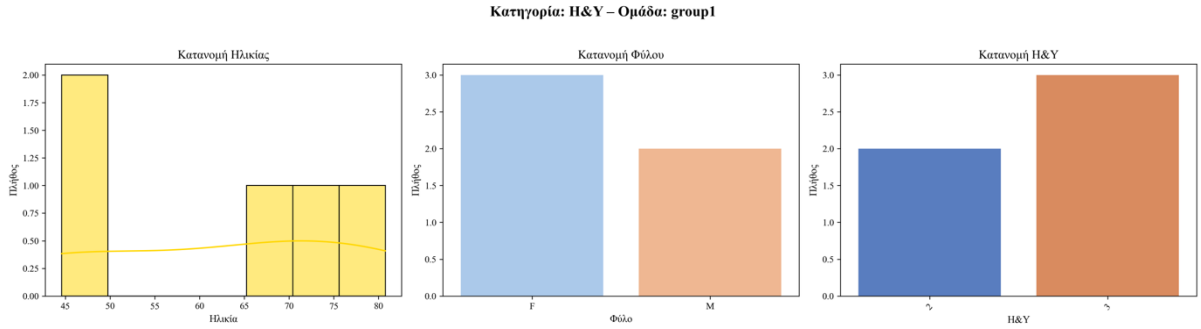
4.7 Διερεύνηση Πιθανών Βιοδεικτών Συσχέτισης με την Απόδοση Προγνωστικών Μοντέλων

Η παρούσα υποενότητα επικεντρώνεται στην διερεύνηση πιθανών κλινικών και δημογραφικών εξαγόμενων χαρακτηριστικών που ίσως αποτελούν βιοδείκτες. Η ανάλυση γίνεται διαφοροποιώντας τους ασθενείς σε δύο εξεταζόμενες ομάδες. Η πρώτη ομάδα περιλαμβάνει

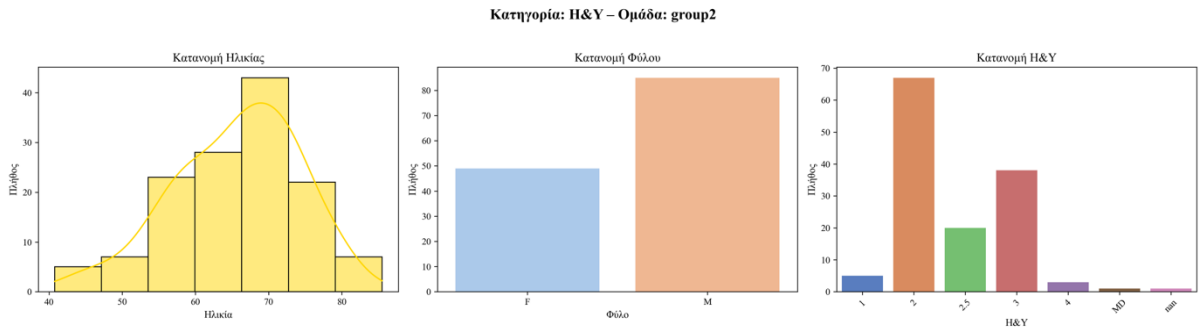
τους ασθενείς οι οποίοι παρουσιάζουν συστηματική αποτυχία στα ταξινομητικά μοντέλα (Ομάδα 1: αποτυχία σε 5 μοντέλα), και η δεύτερη ομάδα περιλαμβάνει τους υπόλοιπους ασθενείς (Ομάδα 2: < 5 αποτυχίες). Ο κύριος στόχος αυτής της ανάλυσης αποτελεί ο εντοπισμός πιθανών παραγόντων που είναι σε θέση να επηρεάσουν την προβλεψιμότητα ενός παράγοντα που συνδέεται άμεσα με την ΝΠ, προσφέροντας έτσι ευκαιρίες για περαιτέρω έρευνα και βελτίωση των μοντέλων. Η ανάλυση βασίζεται στην οπτικοποίηση των κατανομών των εξαγόμενων αποτελεσμάτων, στις διάφορες εξεταζόμενες κατηγορίες, με μεγαλύτερη έμφαση στην σύγκριση μεταξύ των δύο υποκείμενων ομάδων. Ωστόσο, είναι αρκετά σημαντικό να τονισθεί, ότι σε ορισμένες περιπτώσεις, λόγω του μικρού δείγματος ασθενών στην Ομάδα 1, η ερμηνεία γίνεται με πιο επιφυλακτικό τρόπο και η διένεξη περαιτέρω έρευνας στις συγκεκριμένες περιπτώσεις, κρίνεται απαραίτητη.

4.7.1 Πιθανοί Βιοδείκτες από την Κλίμακα Hoehn and Yahr (H&Y)

Η ανάλυση για την κλίμακα H&Y ανέδειξε διαφορές ανάμεσα στις δύο ομάδες. Αρχικά, η Ομάδα 1, που αποτελούνταν από μόλις πέντε ασθενείς, χαρακτηρίζεται κυρίως από την παρουσία αυτών αποκλειστικά στα στάδια 2 (40%) και 3 (60%) της κλίμακας H&Y της νόσου. Τώρα σε δημογραφικό πλαίσιο, οι περισσότεροι ασθενείς είναι γυναίκες και η ηλικία τους κατανέμεται σε διάφορα έτη. Δεδομένου του μικρού δείγματος ασθενών στην Ομάδα 1, τα ευρήματα αυτά χρήζουν περαιτέρω διερεύνησης, για να επιβεβαιωθεί αν το στάδιο στην κλίμακα H&Y ή το φύλο αποτελούν όντως βιοδείκτες και επηρεάζουν την απόδοση των μοντέλων. Από την άλλη, στην Ομάδα 2, που αποτελούνταν από πολύ μεγαλύτερο δείγμα (N=133), τα αποτελέσματα είναι διάσπαρτα μεταξύ των σταδίων, ωστόσο το πλήθος αυτών συγκεντρώνεται στο στάδιο 2 (~50%) και στάδιο 3 (~28%) της νόσου. Σε αντίθεση με την Ομάδα 1, στην Ομάδα 2, η πλειοψηφία των ασθενών φαίνεται να είναι άντρες και να συγκεντρώνονται στην ηλικιακή ομάδα των 70 ετών.



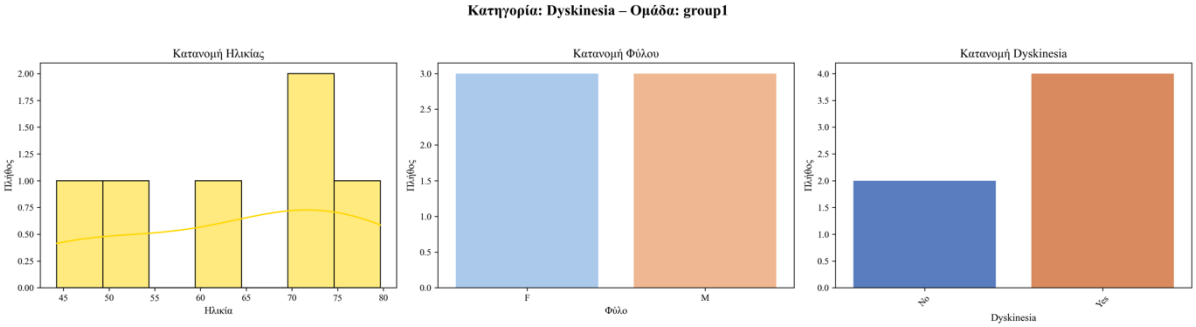
Εικόνα 39: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 1 (Κατηγορία H&Y).



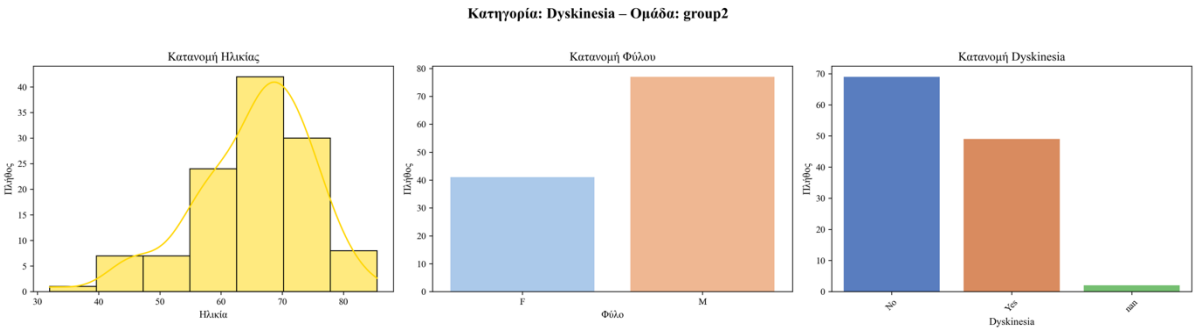
Εικόνα 40: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 2 (Κατηγορία H&Y).

4.7.2 Παρουσία Δυσκινησίας ως Σημαντικός Πιθανός Βιοδείκτης

Στην Ομάδα 1 (N=6), το ~80% των ασθενών εμφάνιζε δυσκινησία, ποσοστό που έρχεται σε αντίθεση με την Ομάδα 2 (N=121), όπου το ποσοστό αυτό πέφτει στο ~40.5%.Τα άτομα με δυσκινησία φαίνεται ότι καθιστούν δύσκολη την πρόγνωση στα μοντέλα, επομένως η περαιτέρω ανάλυση αυτής της παραμέτρου κρίνεται επιτακτική. Επιπλέον, και στις δύο ομάδες παρατηρείται ότι η ηλικία των ασθενών κυμαίνεται στα ~70 έτη, ενώ σχετικά με το φύλο, στην Ομάδα 1 η κατανομή είναι ισάξια, ενώ στην Ομάδα 2 το πλήθος των ασθενώ είναι γυναίκες.



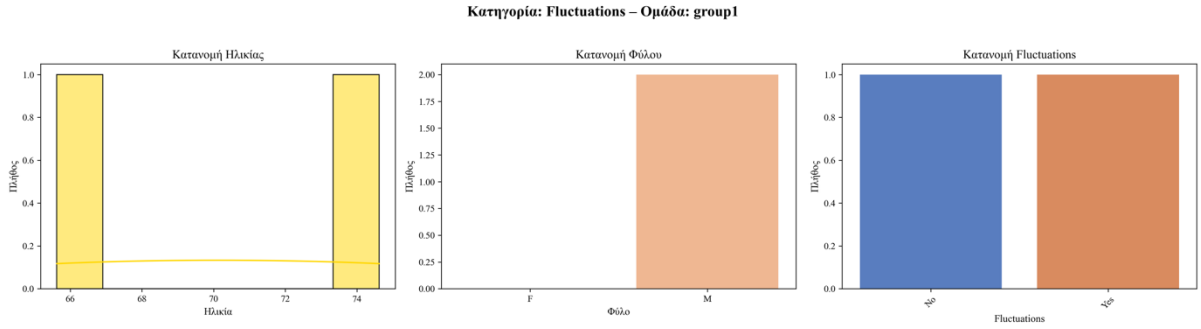
Εικόνα 41: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 1 (Κατηγορία Dyskinesia).



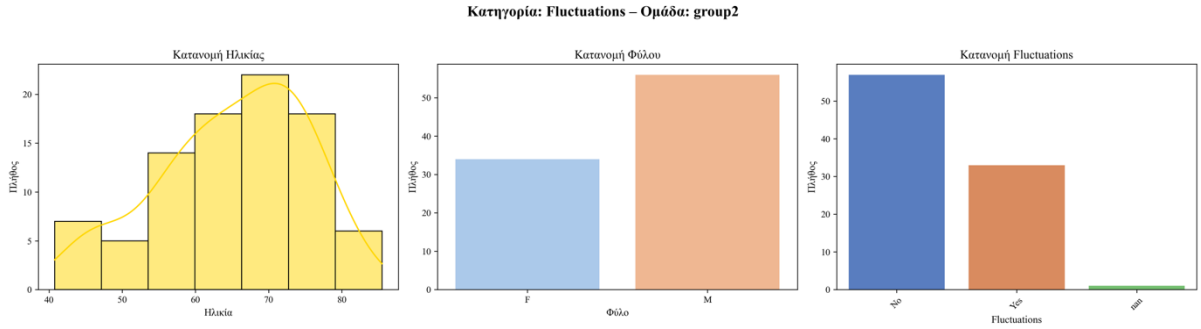
Εικόνα 42:Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 2 (Κατηγορία Dyskinesia).

4.7.3 Διερεύνηση των Κινητικών Διακυμάνσεων και οι Περιορισμοί του Δείγματος

Η ανάλυση των κινητικών διακυμάνσεων περιορίστηκε αρκετά, καθώς το δείγμα για την Ομάδα 1 ήταν πολύ μικρό (N=2), όπου σε αυτούς τους ασθενείς ο ένας παρουσίαζε κινητικές διακυμάνσεις (50%), ενώ ο άλλος όχι (50%), οπότε δεν μπορούν να εξαχθούν αξιόπιστα συμπεράσματα σχετικά με το αν η παρουσία διακυμάνσεων αποτελούν βιοδείκτη της ΝΠ. Όσο αφορά την Ομάδα 2 (N=90), το μεγαλύτερο ποσοστό των ασθενών δεν παρουσίαζαν δυσκινησίας (~62.2%). Τώρα, σχετικά με τα δημογραφικά στοιχεία, και στις δύο ομάδες το πλήθος των ασθενών είναι άντρες, και ότι αφορά την ηλικία αυτών, στην Ομάδα 1 είναι μεταξύ 60-70 ετών, ενώ στην Ομάδα 2 κυμαίνεται στα ~70 έτη, όπως και στις παραπάνω περιπτώσεις



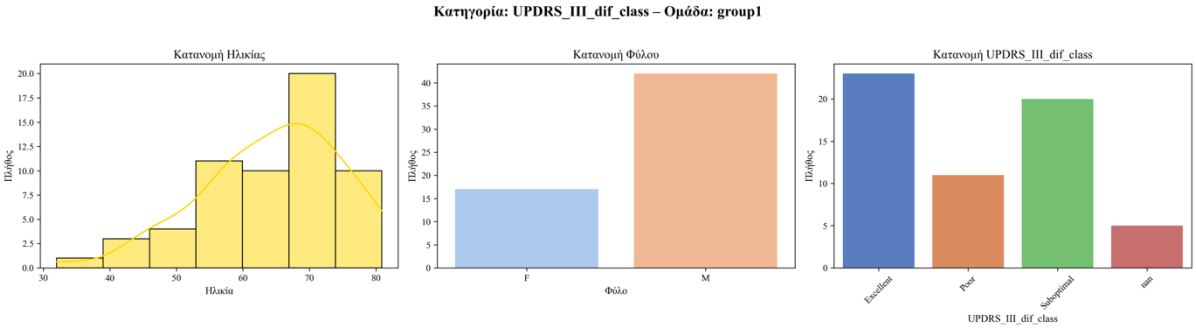
Εικόνα 43: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 1 (Κατηγορία Fluctuations).



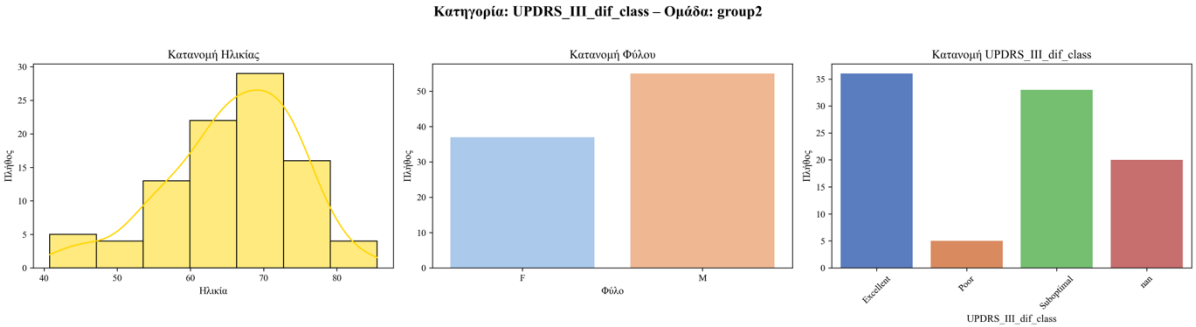
Εικόνα 44: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 2 (Κατηγορία Fluctuations).

4.7.4 Ανάλυση της Ταξινόμησης Διαφορών στην Κλίμακα UPDRS III (UPDRS_III_dif_class)

Η κατηγοριοποίηση των διαφορών της κλίμακας UPDRS_III, που αντικαθρεπτίζει την απόκριση στην θεραπεία, ανέδειξε σημαντικά ευρήματα. Στην Ομάδα 1 (N=58), ένα μεγάλο ποσοστό ασθενών (~20.4%), ταξινομούνται στην κατηγορία “Poor”, ενώ στην Ομάδα 2 (N=93), το ποσοστό αυτό κυμαίνεται σε πολύ μικρότερο ποσοστό (~8%). Σχετικά με τις άλλες δύο περιπτώσεις (Excellent, Suboptimal), τα ποσοστά μεταξύ των δύο ομάδων δεν παρουσίαζαν μεγάλη ποσοστιαία διαφορά, με ελαφρώς υψηλότερα ποσοστά στην Ομάδα 2. Επίσης, είναι σημαντικό να αναφερθεί ότι η συγκεκριμένη κατηγορία περιείχε ένα αξιοσημείωτο πλήθος απροσδιόριστων τιμών, που θα πρέπει να ληφθεί υπόψιν σε περίπτωση περαιτέρω ανάλυσης. Όσον αφορά το φύλο και την ηλικιακή ταξινόμηση, στην Ομάδα 1 παρατηρήθηκε χαμηλότερο ποσοστό γυναικών (~29.3%), έναντι της Ομάδας 2 που το ποσοστό αυτό ήταν υψηλότερο (~40.8%). Και στις δύο ομάδες, η ηλικία προσέγγιζε και εδώ στα ~70 έτη.



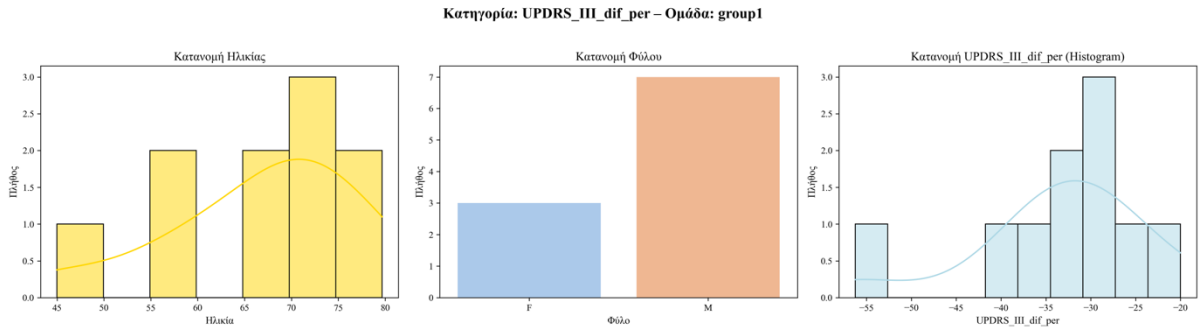
Εικόνα 45: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 1 (Κατηγορία UPDRS_III_dif_class).



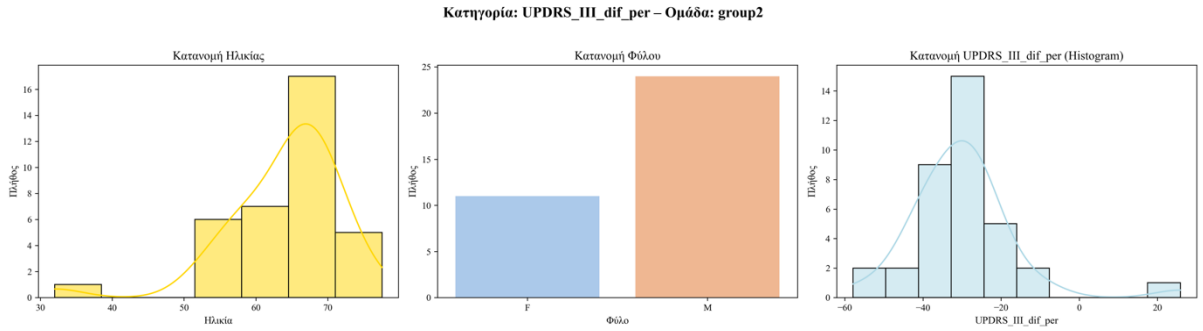
Εικόνα 46: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 2 (Κατηγορία UPDRS_III_dif_class).

4.7.5 Ποσοστιαία Μεταβολή της Κλίμακας UPDRS III (UPDRS_III_dif_per) και η Σχέση της με την Απόδοση των Μοντέλων

Η ανάλυση της ποσοστιαίας μεταβολής της κλίμακας UPDRS_III, ενισχύει τα ευρήματα της παραπάνω κατηγορίας. Εδώ , το πλήθος των ασθενών της Ομάδας 1 (N=10) παρουσίαζε τιμές που ήταν αποκλειστικά αρνητικές (~30-40%), υποδηλώνοντάς μικρότερη βελτίωση με την λήψη της αγωγής. Αντίθετα, η Ομάδα 2 (N=35), παρόλο που εμφάνιζε και αυτή αρνητικές τιμές, εδώ η συνολική κατανομή φαίνεται να μην είναι τόσο κοντά στις αρνητικές τιμές. Ακόλουθα, όσο αφορά την κατανομή του φύλου, και στις δύο ομάδες οι ασθενείς ήταν άντρες, ενώ η ηλικία τους, και σε αυτή την περίπτωση συγκεντρώνεται ~70-75 έτη.



Εικόνα 47: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 1 (Κατηγορία UPDRS_III_dif_per).



Εικόνα 48: Κατανομή Δημογραφικών και Κλινικών Δεικτών – Ομάδα 2 (Κατηγορία UPDRS_III_dif_per).

5 Συζήτηση

Η παρούσα διπλωματική εργασία αποσκοπούσε στην διερεύνηση των φωνητικών αλλοιώσεων της ομιλίας σε Έλληνες ασθενείς, πάσχοντες από την Νόσο του Πάρκινσον (ΝΠ), κάνοντας χρήση πληθώρας φωνητικών χαρακτηριστικών, που προέρχονται από το σύνολο eGeMAPS, και εφαρμόζοντας σύγχρονες τεχνικές μηχανικής και βαθιάς μάθησης. Κύριος στόχος ήταν η αξιολόγηση των ταξινομητικών μοντέλων και η διερεύνηση των δυνατοτήτων τους όταν αυτά αξιοποιούνται στην διαγνωστική διαδικασία, καθώς και γενικά στην παρακολούθηση της νόσου. Τα ευρήματα που προέκυψαν από την εξαγωγή και ανάλυση των χαρακτηριστικών, καθώς και η μετέπειτα αξιολόγηση της απόδοσης των ταξινομητικών μοντέλων αποτελούν σημαντικές βάσεις για μελλοντική έρευνα στο κλάδο της ΝΠ.

Η μεθοδολογία που ακολουθήθηκε περιλάμβανε την εφαρμογή και την συγκριτική ανάλυση μεταξύ έξι (6) διαφορετικών ταξινομητικών αλγορίθμων: K-Nearest Neighbors (KNN), Random Forest, Support Vector Machine (SVM), Logistic Regression, XGBoost, καθώς και τα Spiking Neural Networks (SNNs). Η επιλογή των συγκεκριμένων αλγορίθμων βασίστηκε κυρίως στην ποικιλομορφία αυτών ως προς τον τρόπο λειτουργίας τους, αξιοποιώντας απλά γραμμικά μοντέλα, μέχρι και πιο σύνθετες αρχιτεκτονικές όπως τα τεχνητά νευρωνικά δίκτυα. Βασικός πυλώνας της διαδικασίας αποτέλεσε η εφαρμογή της μεθόδου Leave-One-Subject-Out (LOSO), ως στρατηγική διασταυρούμενης επικύρωσης, που διασφαλίζει πιο αντικειμενική αποτίμηση της απόδοσης των μοντέλων, αντανακλώντας την λειτουργία αυτών σε ρεαλιστικά ιατρικά περιστατικά και συμβάλλοντας στην αποφυγή φαινομένων διαρροής πληροφοριών.

Για την ορθότερη ερμηνεία των αποτελεσμάτων κρίνεται απαραίτητη η πλήρης κατανόηση της σύνθεσης του δείγματος, όπως αυτή παρουσιάστηκε αναλυτικά στην ενότητα 4.1 των αποτελεσμάτων, όπου γινόταν αναφορά στα δημογραφικά και κλινικά χαρακτηριστικά, στοιχεία τα οποία είναι θεμελιώδους αξίας. Η κατανομή των ασθενών ως προς το φύλο, το στάδιο της νόσου με την κλίμακα H&Y, την παρουσία δυσκινησίας και κινητικών διακυμάνσεων, καθώς και η ανταπόκριση στην θεραπεία με Levodopa αποτελεί το θεμέλιο για την αξιολόγηση της βαρύτητας των φωνητικών δεικτών και της αξιολόγησης των εφαρμοζόμενων μοντέλων, που αναλύονται διεξοδικά στις επερχόμενές υποενότητες.

5.1 Ερμηνεία των Σημαντικότερων Φωνητικών Χαρακτηριστικών και η Κλινική τους Σημασία

Η διαδικασία της στατιστικής ανάλυσης εφαρμόστηκε σε οκτώ (8) διαφορετικά σετ ασκήσεων (Task1-8) και μέσω αυτής εξήχθη μια πληθώρα φωνητικών χαρακτηριστικών του συνόλου eGeMAPS, τα οποία κρίθηκαν ως στατιστικά σημαντικά ($p < 0.05$) για την διαφοροποίηση των κλινικών καταστάσεων στην ΝΠ. Τα χαρακτηριστικά τα οποία είχαν πυκνότερη συχνότητα εμφάνισης αφορούν κυρίως την θεμελιώδη συχνότητα F0, καθώς και τις διάφορους φασματικούς δείκτες, όπως φαίνεται και στην Εικόνα 87, όπου χαρακτηριστικά σαν τα F2frequency_sma3nz_amean και F3frequency_sma3nz_amean δείχνουν διαταραχές στην τονικότητα και στην ποιότητα της φωνής ως βασικά στοιχεία της παρκινσονικής ομιλίας. Αυτά τα ευρήματα συνδέονται άρρηκτα με την κλινική εικόνα της υποκινητικής δυσαρθρίας, της πιο συχνής διαταραχής λόγου στην ΝΠ, η οποία οδηγεί σε φωνή χωρίς έντονες μεταβολές, μονότονη, χαμηλής έντασης και με αλλοιωμένο ηχόχρωμα.
Περαιτέρω ανάλυση των εξαγόμενων φωνητικών χαρακτηριστικών ανέδειξε συγκεκριμένους παραμέτρους που έχουν ιδιαίτερη σημασία στην διαδικασία της διάγνωσης και συνδέονται άμεσα με φωνητικές αλλοιώσεις της ΝΠ. Πιο συγκεκριμένα, φωνητικά χαρακτηριστικά όπως jitter, shimmer, HNR και οι συχνότητες F1, F2, F3, εμφάνισαν σημαντική διαγνωστική αξία. Η σημασία αυτών των χαρακτηριστικών έγινε εμφανής σε ασκήσεις συνεχούς φώνησης (Task 1, Εικόνα 23), καθώς και στην εκφώνηση μεμονωμένων λέξεων/φράσεων (Task 2, Task 5 και Εικόνα 33,Εικόνα 55). Αυτά τα ευρήματα επιβεβαιώνουν την υπάρχουσα γνώση, διότι η ΝΠ επηρεάζει κυρίως τον νευρομυϊκό έλεγχο του λάρυγγα και κατά συνέπεια προκαλεί δυσκολίες και γενικότερα αστάθεια στην παραγωγή λόγου. Αυτό εκφράζεται με αυξημένες τιμές για jitter, shimmer και HNR, καθώς και μεταβολές στης φορμαντικές συχνότητες. Όλες αυτές οι μεταβολές αποτυπώνουν τις δυσκολίες που αντιμετωπίζουν οι ασθενείς με ΝΠ να παράγουν φωνή που να διέπετε από σταθερότητα και ελεγχόμενη φώνηση.
Πέρα από τα φωνητικά χαρακτηριστικά, ιδιαίτερη σημασία έδειξαν και τα φασματικά χαρακτηριστικά όπως τα: Mel-Frequency Cepstral Coefficients (MFCCs), η κλίση του φάσματος (spectral slope) και η φασματική ροή (spectral flux). Οι παράμετροι αυτοί εμφανίστηκαν κυρίως σε ασκήσεις που αποτελούνταν από πιο απαιτητικές λεκτικές ασκήσεις όπως η συνεχής ανάγνωση (Task 3, Εικόνα 46), καθώς και οι ασκήσεις που αφορούσαν τον συντονισμό των φωνητικών αρθρώσεων (Task 4, Εικόνα 47). Μεταξύ αυτών των χαρακτηριστικών, ξεχωρίζουν οι MFCCs που συσχετίζονται με την ιδιότητες του φωνητικού σωλήνα και τις κινήσεις των αρθρώσεων. Οι μεταβολές αυτών των παραμέτρων αντανακλούν την δυσκολία ασθενών με ΝΠ να εκτελούν με ακρίβεια και συνέπεια συντονισμένες αρθρωτικές κινήσεις, γεγονός που αποδίδεται στην βραδυκινησία και μυϊκή δυσκαμψία, δύο εκ των σημαντικότερων συμπτωμάτων της νόσου.
Σημαντικά ευρήματα προέκυψαν και από την ανάλυση των ενεργειακών και χρονικών χαρακτηριστικών την ομιλίας. Αν και η εμφάνιση τους ήταν λιγότερο συχνή, έδειξαν και αυτά στατιστική σημαντικότητα. Η μειωμένη ένταση (Loudness-υποφωνία) είναι ένα συχνό γνώρισμα της παρκινσονικής ομιλίας, καθώς και αλλοιώσεις στον ρυθμό (βραδυλαλία ή

ασυντόνιστες παύσεις), παρατηρούνται συχνά σε άτομα με ΝΠ. Πιο συγκεκριμένα, κατά την διάρκεια της αριθμητικής καταμέτρησης (Task 6, Εικόνα 84), ο δείκτης loudnessPeaksPerSec αναδείχθηκε στατιστικά σημαντικός, ενδεχομένως λόγω της δυσκολίας των ασθενών να διατηρήσουν σταθερές και ομαλές μεταβάσεις έντασης κατά την παραγωγή λόγου.
Η συστηματική εμφάνιση αυτών των κατηγοριών φωνητικών χαρακτηριστικών στα διάφορα είδη λεκτικών ασκήσεων καταδεικνύει την επιρροή της ΝΠ στην ομιλία των ασθενών σε διάφορα επίπεδα. Οι αλλοιώσεις που προκύπτουν δεν αφορούν αποκλειστικά την φωνή ή το ρυθμό της ομιλίας, αλλά γενικότερα στην παραγωγή του λόγου, από τον έλεγχο της έντασης μέχρι τον ορθό συντονισμό των εκφράσεων του προσώπου. Επομένως, όπως γίνεται αντιληπτό, δεν αρκεί η ανάλυση μόνο μίας παράμετρού για την περιγραφή των φωνητικών δυσκολιών που αντιμετωπίζουν οι ασθενείς με ΝΠ. Για τον λόγο αυτό, η αξιοποίηση διαφόρων χαρακτηριστικών ομιλίας προσφέρει μια πιο πλήρη και αξιόπιστη εικόνα των φωνητικών δυσλειτουργιών, αποτελώντας πολύτιμο εργαλείο τόσο για την έγκαιρη διάγνωση όσο και για την παρακολούθηση της εξέλιξης της νόσου.

5.2 Ερμηνεία της Απόδοσης των Ταξινομητικών Μοντέλων και η Κλινική τους Σημασία

Η συγκριτική αξιολόγηση των έξι ταξινομητικών μοντέλων (KNN, Logistic Regression, Random Forest, SVM, XGBoost, SNN), μεσώ της εφαρμογής της αυστηρής μεθόδου LOSO, επέτρεψε σε μια εις βάθος διερεύνηση των δυνατοτήτων τους στην πρόβλεψή κλινικών δεικτών για την ΝΠ, στις διάφορες εξεταζόμενες κατηγορίες (H&Y, Fluctuations, Dyskinesia, UPDRS_III_dif_class, UPDRS_III_dif_per).
Σχετικά με την ταξινόμηση του σταδίου της ΝΠ, με βάση την κλίμακα Hoehn & Yahr**,** τα μοντέλα συνόλου όπως Random Forest (Accuracy: 0.684, F1-Score: 0.700) και XGBoost (Accuracy: 0.667, F1-Score: 0.728) παρουσίασαν τις πιο ισορροπημένες αποδόσεις. Η σχετικά υψηλή ακρίβεια θετικής πρόβλεψης (Precision) συνδυαστικά με την μέτρια απόδοση (Accuracy) και ευαισθησία (Recall) στο πλήθος των μοντέλων δείχνει ότι ακόμα και αν οι θετικές προβλέψεις ήταν ορθές, τα μοντέλα δεν ανίχνευαν ορισμένες περιπτώσεις ή σύγχεαν στάδια της νόσου με παρόμοια φωνητικά χαρακτηριστικά. Η σύγχυση αυτή παρατηρείται ενδεχομένως λόγω των λεπτών φωνητικών διαφορών που εντοπίζονται στα διαδοχικά στάδια της νόσου ή σε πιθανές ανισορροπίες στο δείγμα εκπαίδευσης.
Παρόμοια συμπεριφορά παρατηρήθηκε και στην ταξινόμηση της παρουσίας ή όχι διακυμάνσεων (Fluctuations) και δυσκινησίας (Dyskinesia), όπου η τιμή του Precision ήταν υψηλότερη συγκριτικά με Accuracy και Recall. Για την κατηγορία των Fluctuations, τα περισσότερα μοντέλα συμπεριλαμβανομένων των KNN (0.990), Logistic Regression (0.990), XGBoost (0.990) και SNN (0.990) έδειξαν πολύ καλή απόδοση για Precision και τελευταίο ακολούθησε ο Random Forest (0.970). Παρόλη την καλή απόδοση σε Precision, η υψηλότερη ακρίβεια σημειώθηκε για το Random Forest σε μόλις 0.613, ενώ για τα υπόλοιπα μοντέλα εμφανίστηκε χαμηλότερη. Τώρα σχετικά με τις τιμές για f1-score, έδειξαν να είναι πιο συγκρατημένες, με το Logistic Regression να παρουσιάζει την υψηλότερη τιμή (0.693). Τα ευρήματα δείχνουν ότι ενώ τα μοντέλα έχουν υψηλή απόδοση στην πρόβλεψη ύπαρξης διακυμάνσεων, η συνολική τους ικανότητα να εντοπίσουν όλες τις περιπτώσεις ασθενών με αυτό το σύμπτωμα, είναι περιορισμένη. Επομένως, και σε αυτή την περίπτωση, η χρήση των φωνητικών χαρακτηριστικών για την ανίχνευση των διακυμάνσεων εξακολουθούν να αποτελούν πρόκληση.
Ακόλουθα, στην περίπτωση της παρουσίας ή όχι Dyskinesia, η τάση αυτή συνεχίστηκε, με τιμές για Precision να παρουσιάζονται για άλλη μια φορά υψηλές, δηλαδή KNN (0.938), Logistic Regression (0.943), SVM (0.930), XGBoost (0.934) και SNN (0.939). Και σε αυτή την κατηγορία την υψηλότερη απόδοση την εμφάνισε ο αλγόριθμος Random Forest (Accuracy: 0.630), αλλά με την χαμηλότερη Precision (0.908) και F1-Score (0.517) σε σχέση με τα υπόλοιπα μοντέλα. Ένας πιθανός λόγος για αυτή την απόκλιση σε όλα τα μοντέλα ενδεχομένως

είναι η συντηρητική συμπεριφορά των μοντέλων στην πρόβλεψη της παρουσίας δυσκινησίας. Δηλαδή τα μοντέλα προβλέπουν την ύπαρξη δυσκινησίας μόνο όταν είναι βέβαια για την επιλογή αυτή, άρα μειώνουν τα ψευδώς θετικά και κατά συνέπεια αυξάνουν τα ψευδώς αρνητικά. Ωστόσο, στη κατηγορία αυτή το πιο ανησυχητικό εύρημα αποτελούν οι χαμηλές τιμές ROC-AUC για όλα τα μοντέλα (KNN: 0.611, LR: 0.555, RF: 0.526, XGBoost: 0.535, SNN: 0.570, SVM: 0.478). Αυτό δείχνει ότι τα αξιοποιούμενα φωνητικά χαρακτηριστικά δεν διαθέτουν επαρκή διακριτική ικανότητα προκειμένου να προβλέψουν ορθά την παρουσία δυσκινησίας. Στο πλαίσιο αυτό, φαίνεται ότι για τα συγκεκριμένα φωνητικά χαρακτηριστικά, η ανάλυση ομιλίας δεν είναι σε θέση να αποδώσει αξιόπιστα αποτελέσματα για την ανίχνευση της δυσκινησίας σε άτομα με ΝΠ. Αυτό πιθανώς οφείλεται στην γενικότερη δύσκολη φύση της δυσκινησίας, όπου οι ακούσιες κινήσεις που μπορεί να εκδηλωθούν, να μην επηρεάζουν την ομιλία με προβλέψιμο τρόπο, καθιστώντας την ανίχνευση της απαιτητική.
Η ταξινόμηση φάνηκε να είναι δυσκολότερη στην πρόβλεψη της ανταπόκρισης των ασθενών με ΝΠ στην θεραπεία με Levodopa. Η αξιολόγηση έγινε ποιοτικά μέσω της κατηγορικής μεταβλητής UPDRS_III_dif_class και ποσοτικά μέσω της μεταβλητής UPDRS_III_dif_per. Παρατηρήθηκε ότι και στις δύο περιπτώσεις η συνολική απόδοση των μοντέλων ήταν πολύ χαμηλότερη συγκριτικά με την ταξινόμηση των υπόλοιπων κλινικών εξεταζόμενων χαρακτηριστικών. Στην κατηγορία UPDRS_III_dif_class, η οποία ταξινομεί την ανταπόκριση στην θεραπεία μέσω των τριών βαθμίδων "Excellent", "Suboptimal", "Poor", η συνολική απόδοση κρίνεται απογοητευτική. Ο αλγόριθμος Random Forest έδειξε την υψηλότερη ακρίβεια (Accuracy: 0.382), η οποία κυμαίνεται και πάλι σε πολύ χαμηλά ποσοστά, και ακολουθούν τα υπόλοιπα μοντέλα με πολύ χαμηλότερες τιμές. Αρκετά χαμηλές ήταν και οι τιμές για F1-Score σε όλα τα μοντέλα, ωστόσο η τιμή για Precision στα μοντέλα SVM (0.974) και LR (ο.981) φάνηκε να δημιουργεί ερωτήματα, συγκριτικά με την συνολική εικόνα που επέδειξε αυτή η κατηγορία. Αυτή η αντίθεση δείχνει ότι πιθανώς τα μοντέλα να προβλέπουν ορθά την κατηγορία που έχει την υψηλότερη συχνότητα εμφάνισης και αποτυγχάνουν στις υπόλοιπες. Συμπεραίνουμε ότι το να προβλεφθεί κατά πόσο ανταποκρίνεται ένας ασθενής στην θεραπείας, με βάση τα φωνητικά χαρακτηριστικά, είναι αρκετά απαιτητικό, διότι οι αλλαγές που παρατηρούνται στην φωνή και σχετίζονται με την απόκριση στην θεραπεία ίσως είναι πολύ μικρές για να ανιχνευθούν.
Τέλος, για την κατηγορία UPDRS_III_dif_per παρατηρήθηκε παρόμοια συμπεριφορά με την παραπάνω κλινική κατάσταση. Εδώ, τα αποτελέσματα ταξινομήθηκαν με βάση συγκεκριμένες ποσοστιαίες ομάδες (0–15%, 15–30% κ.λπ.), ανάλογα με το ποσοστό βελτίωσής ή επιδείνωσής της ΝΠ. Εδώ, το SVM (Accuracy: 0.465, F1-Score: 0.556) και το XGBoost (Accuracy: 0.444, F1-Score: 0.587) παρουσίασαν τις καλύτερες επιδόσεις αλλά και πάλι χαμηλές. Βέβαια, για το μοντέλο XGBoost, παρά την χαμηλή του ακρίβεια, παρουσίασε υψηλό Precision (1.000), άρα και εδώ το μοντέλο πιθανώς προέβλεψε με επιτυχία ένα μικρό αριθμό μιας κατηγορίας, χωρίς όμως να είναι σε θέση να ταξινομήσει το σύνολο δεδομένων. Γενικά, στην θεραπεία με Levodopa, οι αλλαγές που παρατηρούνται στην ομιλία μετά την λήψη του φαρμάκου από τον ασθενή δεν είναι εύκολα ανιχνεύσιμες διότι διαφέρουν από άτομο σε άτομο. Επίσης, η διακριτοποίηση μιας συνεχούς μεταβλητής και μετατροπή της σε ποσοστιαία ενδεχομένως να εισάγει τεχνητά όρια και να περιπλέκει την συνολική διαδικασία ταξινόμησης. Επομένως, βασιζόμενοί σε όλα αυτά, μπορούμε να πούμε ότι η προσπάθεια ποιοτικής ή ποσοτικής ταξινόμησης της ανταπόκρισης στην θεραπείας απαιτεί περαιτέρω έρευνα για την εξαγωγή πιο αξιόπιστων αποτελεσμάτων.
Συγκεντρωτικά, τα μοντέλα Random Forest και XGBoost παρουσίασαν πιο σταθερή και υψηλότερη απόδοση στο πλήθος των κατηγοριών και μετρικών. Αυτό ίσως οφείλεται στο γεγονός ότι αυτοί οι αλγόριθμοι (αλγόριθμοι συνόλου - ensemble models) χαρακτηρζίονται από την αποτελεσματικότητα τους να εντοπίζουν με μεγαλύτερη ευκολία περίπλοκα μοτίβα στα δεδομένα, να ανταποκρίνονται ακόμα και αν εντοπίζεται θόρυβος και να αποφεύγουν την

υπερβολική προσαρμογή στα δεδομένα εκπαίδευσης. Επίσης, η χρήση της μεθόδου LOSO προσδίδει βαρύτητα και μεγαλύτερη αξιοπιστία στα εξαγόμενα ευρήματα, διότι δείχνει την ικανότητα των μοντέλων να γενικεύουν την απόδοση τους σε άγνωστους ασθενείς, προσεγγίζοντας με μεγαλύτερο ρεαλισμό τις συνθήκες εφαρμογής σε ένα υπαρκτό κλινικό περιβάλλον.

5.3 Ανάλυση Ασθενών με Επαναλαμβανόμενη Χαμηλή Απόδοση Ταξινόμησης: Ετερογένεια και Περιορισμοί των Μοντέλων

Η εκτενής ανάλυση των περιπτώσεων όπου οι ασθενείς με ΝΠ εμφάνισαν συστηματικά χαμηλή απόδοση κρίνεται μείζονος σημασίας για την πληρέστερη κατανόηση της φύσης της νόσου. Από την διερεύνηση του προφίλ των ασθενών αυτών, εξήχθη σημαντικά αποτελέσματα που αφορούν τόσο την κλινική εικόνα, όσο και δημογραφικά στοιχεία. Η σύγκριση, για την εξαγωγή πιθανών βιοδεικτών της νόσου πραγματοποιήθηκε μεταξύ της "Ομάδας 1" (ασθενείς με αποτυχίες > = 5) και της "Ομάδας 2" (ασθενείς με λιγότερες από 5 αποτυχίες).
Ένα από τα βασικά ευρήματα αποτελεί η δυσκολία των μοντέλων να προβλέπουν με επιτυχία το στάδιο της νόσου στην κλίμακα H&Y, όταν αυτό κυμαίνεται μεταξύ των ενδιάμεσων σταδίων (στάδιο 2, στάδιο 3). Για την κατηγορία H&Y, στην Ομάδα 1, εντοπίστηκε ότι το 40% των ασθενών που παρουσίαζαν συστηματική αποτυχία βρισκόταν στα στάδιο 2 της νόσου, ενώ το 60% στο στάδιο 3. Παρόμοιο μοτίβο ακολούθησαν και οι ασθενείς της Ομάδας 2, όπου το 50% των ασθενών βρισκόταν στο στάδιο 2 και το 28% στο στάδιο 3. Στα συγκεκριμένα στάδια, τα συμπτώματά της ΝΠ ίσως είναι λιγότερο ευδιάκριτα, διότι δεν διαφέρουν σε μεγάλο βαθμό μεταξύ τους, όσο θα διέφεραν σε πιο πρώιμα ή προχωρημένα στάδια, όπου οι αλλοιώσεις ενδέχεται να παρουσιάζονται πιο έντονες. Αυτή η απροσδιόριστη περιοχή μεταξύ των ενδιάμεσων σταδίων ενδεχομένως να οδηγεί σε αλληλοεπικάλυψη των φωνητικών χαρακτηριστικών, δυσχεραίνοντας την ταξινόμηση των μοντέλων. Το στάδιο της νόσου στην κλίμακα H&Y, θα μπορούσε να αποτελεί βιοδείκτη της ΝΠ, ωστόσο το μικρό δείγμα ασθενών στο οποίο βασίζεται αυτή η παρατήρηση, αποτελεί την βάση για περισσότερη έρευνα ως προς αυτό τον δείκτη.
Ακόμα πιο σημαντική φάνηκε να είναι η συσχέτισή της παρουσίας δυσκινησίας (Dyskinesia) με την επαναλαμβανόμενη αστοχία των μοντέλων στην ταξινόμηση. Στην Ομάδα 1 το πλήθος των ασθενών (~80%) παρουσίαζαν δυσκινησία, ποσοστό που έρχεται σε αντίθεση με την Ομάδα 2, όπου το ποσοστό αυτό πέφτει στο ~40.5%, δηλαδή σχεδόν στο μισό. Αυτή η αξιοσημείωτη διαφορά, υποδηλώνει έντονα ότι η παρουσία δυσκινησίας ενδέχεται να συνδέεται άρρηκτα με την συστηματική αποτυχία των μοντέλων, και να αποτελεί έναν πιθανό βιοδείκτη της ΝΠ. Αυτή η παρατήρηση ενισχύεται και από το γεγονός ότι τα συμπτώματα συγκινησίας που εμφανίζονται στους ασθενείς επηρεάζουν σημαντικά και τους μύες που είναι υπεύθυνοι για την παραγωγή ομιλίας και αναπνοής. Έτσι, τα φωνητικά σήματα περιλαμβάνουν περισσότερο θόρυβο, δυσκολεύοντας την ανίχνευση των χαρακτηριστικών που αφορούν αποκλειστικά την ομιλία στους ασθενείς με Πάρκινσον. Επίσης, είναι σημαντικό να αναφερθεί ότι αρκετοί ασθενείς που παρουσίαζαν χαμηλή απόδοση ταξινόμησης για δυσκινησία, απέτυχαν εξίσου και στις υπόλοιπες εξεταζόμενες κατηγορίες.
Τώρα, όσον αφορά την συσχέτιση της παρουσίας κινητικών διακυμάνσεων (Fluctuatios), με την συστηματική χαμηλή απόδοση ταξινόμησης, δεν μπορούν να εξαχθούν αξιόπιστα συμπεράσματα, λόγω του περιορισμένου αριθμού δείγματος. Επομένως, δεν είναι εφικτός ο χαρακτηριστικός αυτής της κατηγορίας ως ένας πιθανός βιοδείκτης της ΝΠ και χρήζει περεταίρω διερεύνησης. Βέβαια, ορισμένοι ασθενείς, όπως ο H_PD052, ο οποίος απέτυχε σε όλες τις κατηγορίες, υποδηλώνει ότι ένα πλήθος αυτών χαρακτηρίζεται γενικά από πιο περίπλοκο φωνητικό προφίλ, γεγονός που καθιστά την ταξινόμηση ιδιαίτερα δύσκολο σε αυτές τις περιπτώσεις.

Σε παρόμοιο μοτίβο κυμάνθηκαν και οι ασθενείς με χαμηλή ανταπόκριση στην θεραπεία με Levodopa είτε σε ποιοτικό(UPDRS_III_dif_class)επίπεδο, είτε σε ποσοτικό (UPDRS_III_dif_per). Και στις δύο αυτές περιπτώσεις η απόδοση των μοντέλων κρίθηκε αρκετά χαμηλή. Στην Ομάδα 1, το ποσοστό των ασθενών που ταξινομήθηκαν στην κατηγορία "Poor" έφτασε στο ~20.4 %, ενώ στην Ομάδα 2, μόλις στο ~8%. Αυτή η μεγάλη διαφορά μεταξύ των δύο ομάδων, υποδηλώνει ότι μια κακή απόκριση στην θεραπεία, θα μπορούσε να αποτελεί έναν σημαντικό βιοδείκτη, που θα ήταν σε θέση να συσχετιστεί με την αποτυχία των μοντέλων. Αυτό σημαίνει ότι ενδεχομένως στους ασθενείς αυτούς οι μηχανισμοί που αφορούν την ομιλία και επηρεάζονται από την νόσο, δεν ανταποκρίνονται πλήρως στην θεραπεία ή φωνή αυτών παρουσιάζει πιο περίπλοκα μοτίβα τα οποία τα μοντέλα δεν μπορούν να ταξινομήσουν σωστά.
Τέλος, σχετικά με τα δημογραφικά στοιχεία, δεν εντοπίστηκαν τόσο ξεκάθαρα μοτίβα που θα μπορούσαν να συσχετιστούν άμεσα με την συστηματική αποτυχία ορισμένων ασθενών στην ταξινόμηση. Αν και η πλειονότητα των ασθενών με χαμηλή απόδοση κυμαινόταν ~70 έτη, αυτό δεν μπορεί να θεωρηθεί ένας ισχυρός βιοδείκτης της ΝΠ, διότι αντανακλά την τυπική ηλικιακή ομάδα όπου η νόσος εμφανίζεται συχνότερα, ενώ είναι πιθανό οι φωνητικές αλλοιώσεις να οφείλονται στην γήρανση και όχι αποκλειστικά στις επιδράσεις της ΝΠ. Τώρα σχετικά με την επίδραση του φύλου, φάνηκε να μην παίζει καθοριστικό ρόλο στην αποτυχία αυτή διότι τα αποτελέσματα έδειξαν αντικρουόμενα στοιχεία. Επομένως, το φύλο των ασθενών δεν μπορεί να αποτελέσει έναν από τους βασικούς βιοδείκτες της ΝΠ, αλλά ίσως αν συνδυαστεί με άλλους κλινικούς παράγοντες να επηρεάσει έμμεσα τα αποτελέσματα. Ανεξαρτήτως των περιπτώσεων, η περεταίρω διερεύνηση κρίνεται απαραίτητη.
Τελικά, η αποτυχία των μοντέλων να ταξινομήσουν ορθά ορισμένους ασθενείς συνδέεται πιθανώς με κλινικά χαρακτηριστικά που αντικατοπτρίζουν την φύση και την περίπλοκη δομή της ΝΠ. Ασθενείς οι οποίοι βρίσκονται σε ενδιάμεσα στάδια της νόσου με βάση την κλίμακα Hoehn & Yahr, παρουσιάζουν δυσκινησία ή έχουν μειωμένη ανταπόκριση στης θεραπεία με Levodopa, φαίνεται να αποτελούν τους ισχυρότερους βιοδείκτες για την ΝΠ και χρήζουν περαιτέρω έρευνας. Η κατανόηση αυτών των δεικτών και της σημασίας τους, θα μπορούσαν να βοηθήσουν στην ουσιαστική βελτίωση των μελλοντικών διαγνωστικών διαδικασιών και να εδραιώσουν την ανάλυση ομιλίας ένα πολύτιμο εργαλείο για την ανίχνευση, αξιολόγηση και παρακολούθησής της Νόσου του Πάρκινσον.

5.4 Κλινική Σημασία και Προοπτικές στην Ιατρική Κοινότητα

Τα ευρήματα της παρούσας διπλωματικής εργασίας, παρά τους περιορισμούς που εμφανίζουν, δείχνουν ότι η ανάλυση ομιλίας μπορεί να έχει σημαντική αξία στην αντιμετώπιση της ΝΠ, με έμφαση στον ελληνικό πληθυσμό. Κάνοντας χρήση σύγχρονων υπολογιστικών τεχνικών για την ανάλυση των φωνητικών χαρακτηριστικών, φαίνεται ότι είναι εφικτό να δημιουργηθούν εύχρηστα, οικονομικά και μη επεμβατικά εργαλεία που θα ενισχύσουν την διαγνωστική διαδικασία και θα έχουν ένα θετικό αντίκτυπο στην συνεχή παρακολούθηση της νόσου και την αξιολόγησης της πορείας της θεραπείας.
Η δυνατότητα καταγραφής της φωνής ατόμων με ΝΠ και η συσχέτιση των εξαγόμενων φωνητικών δεικτών με τις κλινικές παραμέτρους, όπως π.χ. η κλίμακα H&Y, μέσω της χρήσης τεχνικών βαθιάς μάθησης, προσφέρει την δυνατότητα ανάπτυξής συμπληρωματικών διαγνωστικών εργαλείων. Στην ιατρική κοινότητα, η διάγνωση της ΝΠ βασίζεται στην κλινική εξέταση, ωστόσο οι τεχνικές που χρησιμοποιούνται (π.χ. DaTSCAN) είναι αρκετά δαπανηρές και σε ορισμένες περιπτώσεις όχι εύκολα προσβάσιμες. Οπότε η ανάλυση ομιλίας φαίνεται να είναι μια πρακτική λύση που θα μπορούσε να συμβάλει στην αντιμετώπιση αυτών των μειονεκτημάτων της κλινικής εξέτασης. Η ομιλία των ασθενών μπορεί να καταγραφεί μέσω κοινών συσκευών (π.χ. μικρόφωνο) και έτσι η ανίχνευση της νόσου μπορεί να επιτευχθεί ακόμα και σε πρώιμο στάδιο. Επίσης, η εστίαση της έρευνας αποκλειστικά σε ελληνόφωνους

είναι πολύ σημαντική, διότι η γλώσσα είναι ένας δείκτης που μπορεί να επηρεάσει τον τρόπο εμφάνισης των συμπτωμάτων, οπότε είναι θετικό να αναπτυχθεί ένα εργαλείο προσαρμοσμένο σε αυτά τα δεδομένα. Επιπλέον, η αναγνώριση συγκεκριμένων χαρακτηριστικών ομιλίας, όπως ο τόνος της φωνής (F0) και οι συντελεστές MFCCs θεωρούνται σημαντικοί για την διάγνωση και συμβάλλουν στην έρευνα στον κλάδο της ΝΠ.
Εκτός από την συμβολή της ανάλυσης ομιλίας στην διαγνωστική διαδικασία, αυτή μπορεί να συμβάλει και στην βελτίωση της παρακολούθησης της εξέλιξης της ΝΠ με την πάροδο του χρόνου, καθώς και στην αξιολόγηση της ανταπόκρισης των ασθενών στην θεραπεία. Η συσχέτιση των εξαγόμενων φωνητικών χαρακτηριστικών με τις κλινικές εξεταζόμενες παραμέτρους, όπως η H&Y ή UPDRS-III, καθώς και με την παρουσία συμπτωμάτων, όπως οι κινητικές διακυμάνσεις ή η δυσκινησία, παρέχουν στους ιατρούς σημαντικούς δείκτες. Επίσης, λόγω του ότι εντοπίζονται ομάδες ασθενών που είναι πιο δύσκολο να ταξινομηθούν σωστά (όπως όσοι έχουν έντονες δυσκινησίες ή δεν ανταποκρίνονται στην θεραπεία) δείχνει πόσο σημαντική είναι η διαφοροποίηση ανά περίπτωση. Επιπρόσθετα, παρέχονται πολύτιμες πληροφορίες που βοηθούν στην κατανόηση της λειτουργίας του εγκεφάλου και των αλλαγών που εντοπίζονται στην ομιλία λόγω παρουσίας της νόσου. Οι πληροφορίες αυτές, καθώς και οι δείκτες που προκύπτουν, μπορούν να βοηθήσουν τους ιατρούς στην έγκαιρη διάγνωση και σχεδιασμό πιο εξατομικευμένων θεραπειών, προσαρμοσμένων στις ανάγκες του κάθε ασθενή ξεχωριστά βελτιώνοντας την καθημερινότητα τους.
Τέλος, αλλά εξίσου σημαντικό, είναι το ότι η μέθοδος της ανάλυσης ομιλίας είναι απλή και χαμηλή σε κόστος, οπότε ενδείκνυται για ενσωμάτωση σε εφαρμογές τηλεϊατρικής και απομακρυσμένης παρακολούθησης (Remote Patient Monitoring - RPM). Έτσι ασθενείς που διαμένουν σε απομακρυσμένες περιοχές μπορούν να ωφεληθούν από αυτή την πρακτική και η πορεία της υγείας τους να παρακολουθείται συνεχώς, επιτρέποντας έγκυρη ανίχνευση της επιδείνωσής και ανταπόκρισης από τους ιατρούς. Με αυτόν τον τρόπο, περιορίζονται οι ανισότητες που υπάρχουν ως προς της πρόσβαση στην υγεία, το κόστος αλλά και οι ασθενείς ταλαιπωρούνται λιγότερο.

5.5 Περιορισμοί της Μελέτης

Παρά την προσπάθεια για μια σωστή μεθοδολογική προσέγγιση, η παρούσα διπλωματική εργασία παρουσιάζει μια ακολουθία περιορισμών που είναι σημαντικό να αναφερθούν, διότι συμβάλουν στην ερμηνεία των αποτελεσμάτων και στην ανάγκη για μελλοντική έρευνα.
Ένας από τους βασικούς περιορισμούς αποτέλεσε το μέγεθος του δείγματος. Παρόλο που το συνολικό πλήθος των 154 ασθενών είναι επαρκές για την μελέτη, όταν η ανάλυση αφορούσε εξειδικευμένες ομάδες αυτού του δείγματος, όπως το πλήθος των ασθενών με συνεχόμενη αποτυχία ταξινόμησης, το δείγμα περιορίστηκε σε αρκετά μικρό αριθμό, που σε ορισμένες κατηγορίες ήταν μόλις δέκα (10) ασθενείς, γεγονός που μειώνει την στατιστική ισχύ των εξαγόμενων αποτελεσμάτων. Ένας επιπλέον περιορισμός θεωρείται η γενικότερη ετερογένειά της ΝΠ, διότι η φωνή των ασθενών μπορεί να επηρεαστεί και από άλλους παράγοντες πέρα από την ύπαρξη της νόσου. Καταστάσεις όπως η διάρκεια της νόσου, η ταυτόχρονη ύπαρξη άλλων παθήσεων, λεπτομέρειες που αφορούν την λήψη της φαρμακευτικής αγωγής αλλά και ίσως γνωσιακές δυσκολίες, ήταν πτυχές που δεν μπορούσαν να ληφθούν υπόψη στην υπάρχουσα ανάλυση. Ίσως διαφοροποιούσαν τα τελικά ευρήματά, διότι ενδεχομένως προκαλούσαν αλλοιώσεις στην ομιλία του ατόμου, οι οποίες δεν συσχετίζονται άμεσα με τις παραμέτρους της νόσου που εξετάσθηκαν.
Ένα σημείο που αξίζει εξίσου να αναφερθεί, είναι η διαδικασία της αποθορυβοποίησης των αρχικών ακατέργαστων ηχητικών σημάτων, που πραγματοποιήθηκε μέσω του λογισμικού Audacity. Αν και η χρήση αυτού του λογισμικού αποτελεί μια ευρέως γνωστή τεχνική που χρησιμοποιείται για την εξαγωγή θορύβου από ηχητικά δεδομένα, με σκοπό την βελτίωση την

ποιότητας, ίσως επιδρά αρνητικά σε πολύ μικρές και δύσκολα αντιληπτές διαφορές της φωνής, που στην παρουσία εργασία θα είχαν σημαντική διαγνωστική αξία.
Τέλος, ένας περιορισμός ενδεχομένως αποτελεί η επιλογή των φωνητικών χαρακτηριστικών με βάση την στατιστική τους σημαντικότητα και χρήση κατωφλίου $p < 0.05$. Αν και αυτή η τιμή κατωφλιού είναι ευρέως χρησιμοποιούμε για τέτοιες περιπτώσεις, στην συγκεκριμένη έρευνα ίσως να αγνοήθηκαν χαρακτηριστικά που ως μονάδα δεν είχαν τεράστια στατιστική αξία, αλλά σε συνδυασμό με άλλα να φαινόταν χρήσιμα. Από την άλλη, ίσως αναδείχθηκαν και χαρακτηριστικά τα ήταν σημαντικά μόνο για το συγκεκριμένο δείγμα, χωρίς να αντιπροσωπεύουν το γενικό σύνολο.
Οι περιορισμοί αυτοί δεν υπονομεύουν την σημαντικότητα των αποτελεσμάτων της παρούσας μελέτης, αλλά συμβάλλουν στην κατανόησή των ορίων και σε ποια σημεία θα πρέπει να δοθεί περαιτέρω προσοχή. Επίσης, συμβάλλουν στην βελτίωσή της μελλοντικής έρευνας στην φωνητική ανάλυση για την Νόσο του Πάρκινσον, προκειμένου η ανάλυση ομιλίας να μπορέσει να είναι ένα αξιοποιήσιμο και έμπιστο εργαλείο στην διάγνωση και παρακολούθηση της πορείας της νόσου.

5.6 Προτάσεις για Μελλοντική Ερευνητική Δραστηριότητα

Τα συμπεράσματα που προέκυψαν από την ανάλυση, καθώς και η περιορισμοί της παρούσας έρευνας αναδεικνύουν αρκετές προοπτικές για μελλοντική έρευνα, η οποία θα συμβάλει σε βαθύτερη κατανόηση της ΝΠ και της σημαντικότητάς της ανάλυσης ομιλίας στον τομέα αυτό.
Ένα αρχικό βήμα θα μπορούσε να είναι η συλλογή ακόμα πιο ετερογενούς δείγματος ασθενώς, όπου θα συμπεριλάμβανε περισσότερα άτομα στα διάφορα στάδια της νόσου, ασθενείς με διαφορετικούς φαινοτύπους, που θα παρουσίαζαν και άλλα συμπτώματα. Επίσης, πολύ σημαντική είναι η κατεύθυνση της μελέτης σε πιο διαχρονικές τεχνικές, όπου η παρακολούθηση της πορείας των ασθενών πραγματοποιείται σε ένα μεγαλύτερο εύρος χρόνου, με απώτερο σκοπό την κατανόηση των φωνητικών αλλοιώσεων με την πάροδο του χρόνου και πως αυτές συσχετίζονται με την εξέλιξη της ΝΠ και την απόκριση στην θεραπεία.
Μια ακόμα ενδιαφέρουσα ιδέα, αποτελεί ο συνδυασμός των φωνητικών χαρακτηριστικών με άλλα πολυτροπικά δεδομένα όπως οι εκφράσεις του προσώπου, η γραφή ή άλλων γενετικών χαρακτηριστικών. Αυτός ο σύνδεσμός μπορεί να βελτιώσει την ακρίβεια και την αξιοπιστία των διάφορων κλινικών προσεγγίσεων ενισχύοντάς την αξία των αποτελεσμάτων. Επίσης θα μπορούσαν να πραγματοποιηθούν ορισμένες βελτιώσεις στον τομέα την ανάλυσης των δεδομένων. Θα μπορούσαν να αξιοποιηθούν πιο προηγμένες τεχνικές για την επεξεργασία των φωνητικών χαρακτηριστικών, οπού αυτά αντί να επιλέγονται μέσω παραδοσιακών μεθόδων, όπως η στατιστική ανάλυση, να γίνεται χρήση τεχνικών όπου κατά την διαδικασία της εκπαίδευσης να γίνεται η ταυτόχρονη επιλογή των στατιστικά σημαντικότερων τεχνικών, εξοικονομώντας πόρους του συστήματος. Επιπλέον, η πιο εντατική χρήση τεχνικών βαθιάς μάθησης, που επιτρέπουν στο μοντέλο να αναγνωρίζει μοτίβα απευθειας από τα δεδομένα χωρίς ανθρώπινη παρέμβαση, θα προσφέρει νέες δυνατότητες για μια πιο ακριβή και αυτοματοποιημένη προσέγγιση για την αξιολόγησή των φωνητικών δεδομένων για την ΝΠ.
Πιο συγκεκριμένα η χρήση Spiking Neural Networks (SNNs), αποτελεί μια πολλά υποσχόμενη κατηγορία που σίγουρα χρήζει περαιτέρω έρευνας. Αυτό διότι το συγκεκριμένο είδος τεχνητού νευρωσικού δικτύου έχει βασίσει την αρχιτεκτονική του στην λειτουργία του ανθρωπίνου εγκεφάλου, προσπαθώντας να μιμηθεί την διαδικασία μετάδοσης των πληροφορίων μέσω χρονικών ‘αιχμών’, καθιστώντας τα κατάλληλα για την ανάλυση της ομιλίας. Βέβαια, η έρευνα στο συγκεκριμένο κλάδο είναι ακόμα σε πρώιμο στάδιο, οπότε η πιο εντατική διερεύνηση αυτών ως προς την απόδοση και την αξιοπιστία σε ιατρικά δεδομένα, χρίζεται σημαντική.
Τέλος, ένα ενδιαφέρον πεδίο έρευνας αποτελεί η εξατομικευμένη ιατρική, μελετώντας την φωνή κάθε ασθενή ξεχωριστά προκειμένου να εξαχθούν πληροφορίες σχετικά με το ποιοι ασθενείς βελτιώνονται περισσότερο και ποιοι λιγότερο, έπειτα από την έκθεση τους σε

συγκεκριμένη φαρμακευτική αγωγή. Δηλαδή, να εξετασθεί αν η ανάλυση της φωνής είναι σε θέση από μόνη της να δείξει ποιος ασθενής ανταποκρίνεται καλύτερα στην θεραπεία.

5.7 Γενικά Συμπεράσματα

Η παρούσα διπλωματική εργασία συνέβαλε στην καλύτερη κατανόηση της επιρροής που έχει η Νόσος του Πάρκινσον στην ομιλία των ασθενών. Μέσω της ανάλυσης των φωνητικών χαρακτηριστικών, που εξήχθησαν από το σύνολο eGeMAPS και της αξιοποίησης διάφορων ταξινομητικών μοντέλων μέσω της αυστηρής μεθόδου επικύρωσης LOSO, αποδείχθηκε ότι πληθώρα χαρακτηριστικών σχετίζονται άμεσα με την κλινική εικόνα της νόσου. Πιο συγκεκριμένα, τα μοντέλα Random Forest και XGBoost παρουσίασαν την καλύτερη απόδοση αναμεσά στα ταξινομητικά μοντέλα, ωστόσο η πρόβλεψη της ανταπόκριση στην θεραπεία ή η παρουσία διακυμάνσεων αποτελεί μεγάλη πρόκληση για την ανάλυση ομιλίας ασθενών με ΝΠ. Τα εξαγόμενα ευρήματα τονίζουν την ανάγκη για συνεχή έρευνα, κάνοντας χρήση περισσότερων και πιο ποικιλόμορφων φωνητικών δεδομένων, με απώτερο σκοπό την γενίκευση των αποτελεσμάτων. Επίσης, η χρήση πιο εξελιγμένων τεχνικών, έναντι των απλών στατιστικών εργαλείων κρίνεται επιτακτική, για την εύρεση ακόμα περισσότερων μοτίβων που θα ενισχύσουν την ερευνητική διαδικασία. Πολύ σημαντικό είναι επίσης να κατανοήσουμε ότι ένα μοντέλο, δεν αρκεί να είναι απλά ακριβές, θα πρέπει να είναι ξεκάθαρη και η αιτία που αυτό συμπεριφέρεται έτσι και αν τελικά τα αποτελέσματα αυτού είναι ρεαλιστικά, ειδικά σε εφαρμογές που συνδέονται με την ιατρική κοινότητα.
Για τον σκοπό αυτό, είναι σημαντική η δημιουργία διαγνωστικών, μη επεμβατικών εργαλείων που θα συνδυάζονται με την κλασική ιατρική εξέταση και θα έχουν ως πυρήνα την ανάλυση ομιλία ατόμων με ΝΠ, συμβάλλοντας έτσι στην έγκαιρη διάγνωση και παρακολούθηση της νόσου. Για να πραγματοποιηθεί αυτό, είναι σημαντική η συνεργασία μεταξύ των επιστημόνων του ιατρικού και μηχανικού κλάδου έτσι ώστε τα διαγνωστικά εργαλεία που θα δημιουργηθούν να είναι πρακτικά, ασφαλή και εύκολα στην χρήση.

Βιβλιογραφία

[1] M. C. Staff, «Parkinson’s disease – Symptoms and causes,» 25 May 2023. [Ηλεκτρονικό]. Available: https://www.mayoclinic.org/diseases-conditions/parkinsons-disease/symptoms-causes/syc-20376055.

[2] R. Ileșan, S. Ștefănigă, R. Fleșar και e. al., «In Silico Decoding of Parkinson’s: Speech & Writing Analysis,» *Journal of Clinical Medicine,* τόμ. 13, αρ. 18, p. 5573, 2024.

[3] M. Clinic, «Parkinson’s Disease – Overview,» May 2024. [Ηλεκτρονικό]. Available: https://www.mayo.edu/research/labs/parkinsons-disease/overview.

[4] O. A. Teixeira, «Cientistas descobrem sinais da Doença de Parkinson muitos anos antes dos sintomas,» Estadão, 9 April 2024 . [Ηλεκτρονικό]. Available: https://www.estadao.com.br/ciencia/cientistas-descobrem-sinais-da-doenca-de-parkinson-muitos-anos-antes-dos-sintomas/.

[5] D. F. V. F. B. Z. H. P. M. S.-P. L. &. G. N. M. Xavier, «Voice analysis in Parkinson's disease - a systematic literature review,» *Artificial Intelligence in Medicine,* τόμ. 163, αρ. DOI: 10.1016/j.artmed.2025.103109, 2025.

[6] N. I. o. N. D. a. Stroke, «Parkinson’s Disease,» 16 May 2024. [Ηλεκτρονικό]. Available: https://www.ninds.nih.gov/health-information/disorders/parkinsons-disease.

[7] StartStemCells, «Parkinson’s Disease Treatment,» 2024. [Ηλεκτρονικό]. Available: https://www.startstemcells.com/parkinson-disease-treatment.html .

[8] C. Yannakakis και G. A. A. M. Papadopoulos, «Speech biomarkers of Parkinson’s disease: Comparing shallow and deep learning models,» σε *Artificial Intelligence Applications and Innovations. AIAI 2023 IFIP WG 12.5 International Workshops*, τόμ. 521, Cham, MacIntyre H.,Maglogiannis I., 2023, p. 456–468.

[9] L. v. G. a. C. Tejedor-García, "Innovative Speech-Based Deep Learning Approaches for Parkinson's Disease Classification: A Systematic Review,", Appl. Sci., vol. 14, no. 17, p. 7873, Sep. 2024.

[10] H. R. a. M. A. Akhloufi, «A review of machine learning and deep learning for Parkinson's disease detection,» *Discover Artificial Intelligence,* τόμ. 5, αρ. 1, p. 24, 2025.

[11] P. Foundation, «Non-Movement Symptoms,» Parkinson's Foundation. [Ηλεκτρονικό]. Available: https://www.parkinson.org/understanding-parkinsons/non-movement-symptoms. [Πρόσβαση 16 March 2025].

[12] C. Clinic, «Parkinson's Disease: What It Is, Causes, Symptoms & Treatment,» Cleveland Clinic, [Ηλεκτρονικό]. Available: https://my.clevelandclinic.org/health/diseases/8525-parkinsons-disease-an-overview. [Πρόσβαση 15 March 2025].

[13] K. R. Chaudhuri και A. H. Schapira, «Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment,» *The Lancet Neurology,* τόμ. 8, αρ. 5, pp. 464-474, 2009.

[14] N. F. Niloofar Fadavi, "Early Recognition of Parkinson's Disease Through Acoustic Analysis and Machine Learning," *arXiv,* 2024.

[15] K. Tjaden, «Speech and Swallowing in Parkinson's Disease,» *Topics in Geriatric Rehabilitation,* τόμ. 24, αρ. 2, pp. 115-126, 2008.

[16] H. Ç. A. A. a. K. P. O. Karaman, «Robust automated Parkinson disease detection based on voice signals with transfer learning,» *Expert Systems with Applications,* τόμ. 178, αρ. 115013, 2021.

[17] P. D. e. al, «A robust Parkinson’s disease detection model based on time-varying synaptic efficacy function in spiking neural network,» *BMC Neurology,* τόμ. 24, αρ. 492, 2024.

[18] M. M. K. M. I. O. A. H. D. T. K. Awais Mahmood, «End-to-End Deep Learning Method for Detection of Invasive Parkinson's Disease,» *Diagnostics,* τόμ. 13, αρ. 6, p. 1088, 2023.

[19] V. Reddy, «PPINtonus: Early Detection of Parkinson's Disease Using Deep-Learning Tonal Analysis,» arXiv.org, 2024.

[20] K. U. o. Technology, «Altered speech may be the first sign of Parkinson's disease,» *ScienceDaily,* 2023.

[21] A. Mahmood, M. Mehroz Khan, M. Imran, O. Alhajlah, H. Dhahri και T. Karamat, «End-to-End Deep Learning Method for Detection of Invasive Parkinson's Disease,» *Diagnostics,* τόμ. 13, αρ. 6, p. 1088, 2023.

[22] B. Çelik και A. Akbal, «Distinguishing Parkinson's Patients Using Voice-Based Feature Extraction and Classification,» *arXiv preprint,* p. arXiv:2501.14390, 2025.

[23] L. Simone, M. G. Camporeale, V. M. Rubino, V. Gervasi και G. Dimauro, «Interpretable Early Detection of Parkinson's Disease through Speech Analysis,» *arXiv preprint,* p. arXiv:2504.17739, 2025.

[24] J. L. Crawford, «Linguistic changes in spontaneous speech for detecting Parkinson’s disease using large language models,» *PLOS Digital Health,* τόμ. 4, αρ. 2, p. e0000757 , 2025.

[25] N. Momeni, S. Whitling και A. Jakobsson, «Detecting Parkinson's Disease Using Voice Recordings from Mobile Devices,» σε *32nd European Signal Processing Conference (EUSIPCO 2024)*, Lyon, 2024.

[26] F. Eyben, K. R. Scherer, B. W. Schuller, J. Sundberg, E. André, C. Busso, L. Y. Devillers, J. Epps, P. Laukka, S. S. Narayanan και K. P. Truong, «The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing,» *IEEE Transactions on Affective Computing,* τόμ. 7, αρ. 2, pp. 190-202, 2016.

[27] A. Pompili, D. Gimeno-Gómez, C. Botelho, A. Abad και C.-D. Martínez-Hinarejos, «Unveiling Interpretability in Self-Supervised Speech Representations for Parkinson's Diagnosis,» *arXiv preprint,* p. arXiv:2412.02006, 2024.

[28] X. Pan, B. Liang και T. Cao, «A bibliometric analysis of speech and language impairments in Parkinson’s disease based on Web of Science,» *Frontiers in Psychology,* τόμ. 15, αρ. 1374924, 2024.

[29] T. e. al., «Acoustic Analysis of Voice and Speech Characteristics in Early Untreated Parkinson's Disease,» σε *Proceedings of the 6th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications (MAVEBA)*, 2011.

[30] D. Escobar-Grisales, C. D. Ríos-Urrego και J. R. Orozco-Arroyave, «Deep Learning and Artificial Intelligence Applied to Model Speech and Language in Parkinson's Disease,» *Diagnostics,* τόμ. 13, αρ. 13, p. 2163, 2023.

[31] D. Gimeno-Gómez, C. Botelho, A. M. Pompili, A. Abad και C.-D. Martínez-Hinarejos, «Unveiling Interpretability in Self-Supervised Speech Representations for Parkinson's

Diagnosis,» *IEEE Journal of Selected Topics in Signal Processing,* τόμ. 99, αρ. 1, pp. 1-14, 2025.

[32] A. M. Pompili, R. Solera-Ureña, A. Abad, R. Cardoso, I. Guimarães, M. Fabbri, I. P. Martins και J. Ferreira, «Assessment of Parkinson's Disease Medication State Through Automatic Speech Analysis,» σε *Interspeech 2020,* 2020.

[33] M. Niemelä, M. von Bonsdorff, S. Äyrämö και T. Kärkkäinen, «Identification of Cognitive Decline from Spoken Language through Feature Selection and the Bag of Acoustic Words Model,» *arXiv preprint,* 2024.

[34] P. Mattisson και F. Samuelsson Forsdik, A Speech Signal Processing Approach for Early Detection of Parkinson's Disease, Lund: Lund University, 2024.

[35] L. van Bemmel, C. Cucchiarini και H. Strik, «Using Feature Selection to Evaluate Pathological Speech After Training with a Serious Game,» σε *Proceedings of the 12th International Conference of Experimental Linguistics (ExLing 2021)*, 2021.

[36] J. Rusz, J. Procházková, J. Novotný, R. Čmejla και E. Růžička, «Automatic speech analysis combined with machine learning reliably predicts the motor state in people with Parkinson's disease,» *NPJ Parkinson's Disease,* τόμ. 10 , αρ. 1, p. 31, 2024.

[37] G. Di Cesare, R. Gharib, G. Carta και e. al., «A Machine Learning Approach for Early Detection of Parkinson's Disease through Speech Feature Analysis,» *Biomedical Signal Processing and Control,* τόμ. 90, 2024.

[38] K. e. al., «A Pilot Study Comparing Speech Characteristics in People with Parkinson's Disease across Motor States,» medRxiv, 2025. [Ηλεκτρονικό]. Available: https://www.medrxiv.org/content/10.1101/2025.03.25.25324640v1.

[39] D. K. J. M. L. R. V. K. a. I. R. L. Brabenec, «Short-term effects of transcranial direct current stimulation on motor and speech outcomes in Parkinson's disease,» *Brain Sciences* , τόμ. 14, αρ. 3, p. 276, 2024.

[40] F. A. J. K. G. A. R. E. F. a. E. S. H. Steurer, «Speech and neuroimaging effects following HiCommunication: a novel speech treatment for Parkinson's disease,» *Brain Communications,* τόμ. 6, αρ. 4, 2024.

[41] M. L. R. A. M. H. M. a. C. A. T.-F. E. L. Frankford, «Contributions of speech timing and articulatory precision to listener perceptions of intelligibility and naturalness in Parkinson's disease,» *Journal of Speech, Language, and Hearing Research,* τόμ. 67, p. 2951–2963, 2024.

[42] R. A. R. A. H. H. M. Z. A. a. A. V. J. M. Diaz-Cadena, «Associations of Voice Metrics with Postural Function in Parkinson's Disease,» *Life,* τόμ. 15, αρ. 1, 2024.

[43] S. Kunjan, T. S. Grummett, K. J. Pope και T. W. Lewis, «The Necessity of Leave One Subject Out (LOSO) Cross Validation for EEG Disease Diagnosis,» σε *Brain Informatics,* Springer, Cham, 2021.

[44] R. e. al., «Voice disorder detection using machine learning algorithms: An application in speech and language pathology,» *Engineering Applications of Artificial Intelligence,* τόμ. 133, 2024.

[45] Q. Yang, X. Li, X. Ding, F. Xu και Z. Ling, «Deep learning-based speech analysis for Alzheimer's disease detection: a literature review,» *Alzheimer's Research & Therapy,* τόμ. 14, p. 186, 2022.

[46] C. Quan, K. Ren και Z. Luo, «A Deep Learning Based Method for Parkinson's Disease Detection Using Dynamic Features of Speech,» *IEEE Access,* τόμ. 9, p. 10239–10252, 2021.

[47] C. Laganas, D. Iakovakis, S. Hadjidimitriou, V. Charisis, S. B. Dias, S. Bostantzopoulou, Z. Katsarou, L. Klingelhoefer, H. Reichmann, D. Trivedi, K. R. Chaudhuri και L. J. Hadjileontiadis, «Parkinson's Disease Detection Based on Running Speech Data From Phone Calls,» *IEEE Transactions on Biomedical Engineering,* pp. 1-12, 2021.

[48] L. Berus, S. Klancnik, M. Brezocnik και M. Ficko, «Classifying Parkinson's Disease Based on Acoustic Measures Using Artificial Neural Networks,» *Sensors,* τόμ. 19, p. 16, 2019.

[49] L. Ali, C. Zhu, Z. Zhang και Y. Liu, «Automated Detection of Parkinson's Disease Based on Multiple Types of Sustained Phonations Using Linear Discriminant Analysis and Genetically Optimized Neural Network,» *IEEE Journal of Translational Engineering in Health and Medicine,* τόμ. 7, pp. 1-11, 2019.

[50] I. Tougui, M. Zakroum, O. Karrakchou και M. Ghogho, «Transformer-based transfer learning on self-reported voice recordings for Parkinson's disease diagnosis,» *Scientific Reports,* τόμ. 14, p. Article number: 30131, 2024 .

[51] M. Siniukov, E. Xing, S. Attaripour Isfahani και M. Soleymani, «Towards a Generalizable Speech Marker for Parkinson's Disease Diagnosis,» *arXiv preprint,* 2025.

[52] D. e. al., «A robust Parkinson's disease detection model based on time-varying synaptic efficacy function in spiking neural network,» *BMC Neurology,* τόμ. 24, p. 492, 2024.

[53] W. Maass, «Networks of Spiking Neurons: The Third Generation of Neural Network Models,» *Neural Networks,* τόμ. 10, αρ. 9, p. 1659–1671, 1997.

[54] M. Yao, H. Gao, G. Zhao, D. Wang, Y. Lin, Z. Yang και G. Li, «Temporal-wise Attention Spiking Neural Networks for Event Streams Classification,» *arXiv preprint,* 2021.

[55] P. G. Malettira, S. Negi, W. Ponghiran και K. Roy, «TSkips: Efficiency Through Explicit Temporal Delay Connections in Spiking Neural Networks,» *arXiv preprint,* 2024.

[56] G. W. De Silva, «Exploring the World of Artificial Neural Networks – A Beginner's Overview,» University of Colombo, Colombo, Sri Lanka, 2020.

[57] M. A. Hossain και F. Amenta, «Machine Learning-Based Classification of Parkinson's Disease Patients Using Speech Biomarkers,» *Journal of Parkinson's Disease,* τόμ. 14, pp. 95-109, 2024.

[58] U. S. C. Group, «What is complete or quasi-complete separation in logistic regression and what are some strategies to deal with the issue?,» [Ηλεκτρονικό]. Available: https://stats.oarc.ucla.edu/other/mult-pkg/faq/general/faqwhat-is-complete-or-quasi-complete-separation-in-logistic-regression-and-what-are-some-strategies-to-deal-with-the-issue/.

[59] D. D. o. D. Science, «Logistic Regression Cannot Perfectly Model Well-separated Classes,» [Ηλεκτρονικό]. Available: https://blog.dailydoseofds.com/p/logistic-regression-cannot-perfectly.

[60] Abhigyan, «Understanding Logistic Regression,» 26 July 2020. [Ηλεκτρονικό]. Available: https://medium.com/analytics-vidhya/understanding-logistic-regression-b3c672deac04.

[61] F. A. Mohammad Amran Hossain, «Machine Learning-Based Classification of Parkinson's Disease Patients Using Speech Biomarkers,» *Journal of Parkinson's Disease,* τόμ. 14, pp. 95-109, 2024.

[62] S. D. R. S. Samiappan Dhanalakshmi, «Speech features-based Parkinson's disease classification using combined SMOTE-ENN and binary machine learning,» *Health Technology,* τόμ. 14, αρ. 2, p. 393–406, 2023.

[63] R. Anand, «Evaluating Speech Analysis Techniques for Parkinson's Disease Detection: A Comparison of Machine Learning and Deep Learning Algorithms,» *International Journal of Advanced Research,* τόμ. 12, αρ. 5, pp. 1118-1137, 2024.

[64] S. B. K. Verma Vanshika, «Logistic Regression-Based Detection of Parkinson's Disease,» *Preprints.org (preprint),* 2025.

[65] G. C. Oliveira, N. D. Pah, Q. C. Ngo, A. Yoshida, N. B. Gomes, J. P. Papa και D. Kumar, «A pilot study for speech assessment to detect the severity of Parkinson's disease: An ensemble approach,» *Computers in Biology and Medicine,* τόμ. 185, p. 109565, 2025.

[66] B. Leo, « Random Forests,» *Machine Learning,* τόμ. 45, pp. 5-32, 2001.

[67] S. M. Jeong, Y. D. Song, C. L. Seok, J. Y. Lee, E. C. Lee και H. J. Kim, «Machine learning-based classification of Parkinson's disease using acoustic features: Insights from multilingual speech tasks,» *Computers in Biology and Medicine ,* τόμ. 182, 2024.

[68] S. Lakkakula, D. LakshmiSaranya, J. Swathi και A. V. A. N. Sarvani, «Detecting Parkinson's Disease Through Voice Analysis,» *International Advanced Research Journal in Science, Engineering and Technology (IARJSET),* τόμ. 11, αρ. 3, pp. 96-100, 2024.

[69] R. Baranwal, U. Gupta, S. Singh, P. Mishra και N. Agarwal, «Parkinson Disease Detection Comparison Using Machine Learning Techniques,» *International Journal of Research Publication and Reviews,* τόμ. 5, αρ. 5, p. 10808–10813, 2024.

[70] N. R. Fahira, A. Lawi και M. Aqsha, «Early Detection Model of Parkinson's Disease Using Random Forest Method on Voice Frequency Data,» *Journal of Natural Sciences and Mathematics Research,* τόμ. 9, αρ. 1, pp. 30-38, 2023.

[71] S. K. Venkatesan S, «Speech Insight: Parkinson's Disease Diagnosis Using Random Forest Classifier,» *International Journal of Creative Research Thoughts (IJCRT),* τόμ. 11, αρ. 7, pp. 419-424, 2023.

[72] A. T. A. F. Y. D. Karim Ayesha, «Machine learning-based classification of Parkinson's disease using acoustic features: Insights from multilingual speech tasks,» *Computers in Biology and Medicine,* τόμ. 182, 2024.

[73] S. Team, «Random Forest Algorithm in Machine Learning,» 27 January 2025. [Ηλεκτρονικό]. Available: https://www.sitepoint.com/random-forest-algorithm-in-machine-learning/.

[74] A. H. Al-Nefaie, T. H. H. Aldhyani και D. Koundal, «Developing System-based Voice Features for Detecting Parkinson's Disease Using Machine Learning Algorithms,» *Journal of Disability Research,* τόμ. 3, pp. 1-10, 2024.

[75] M. Shahbakhi, D. T. Far και E. Tahami, «peech Analysis for Diagnosis of Parkinson's Disease Using Genetic Algorithm and Support Vector Machine,» *Journal of Biomedical Science and Engineering,* τόμ. 7, pp. 147-156, 2014.

[76] S. R. Kadiri, M. Kodali και P. Alku, «Severity Classification of Parkinson's Disease from Speech using Single Frequency Filtering-based Features,» *arXiv preprint,* pp. 1-9, 2023.

[77] Tutorialspoint, « Support Vector Machine (SVM) in Machine Learning,» 2025. [Ηλεκτρονικό]. Available:

https://www.tutorialspoint.com/machine_learning/machine_learning_support_vector_machine.htm.

[78] T. V. Phakatkar A. G, «Parkinson's Disease Detection on Unbalanced Speech Data using Convolutional Neural Networks,» *Communications on Applied Nonlinear Analysis,* τόμ. 32, αρ. 2, pp. 202-214, 2025.

[79] D. e. al., «An Efficient Method to Diagnose Parkinson's Disease Using Machine Learning Algorithms,» *Journal of Engineering Research and Reports,* τόμ. 24, αρ. 11, pp. 54-61, 2023.

[80] R. Alshammri, G. Alharbi, E. Alharbi και I. Almubark, «Machine learning approaches to identify Parkinson's disease using voice signal features,» *rontiers in Artificial Intelligence,* τόμ. 6, p. Article 1084001, 2023.

[81] J. Jumanto, «Optimizing Support Vector Machine Performance for Parkinson's Disease Diagnosis Using GridSearchCV and PCA-Based Feature Extraction,» *Journal of Information Systems Engineering and Business Intelligence,* τόμ. 10, αρ. 1, pp. 38-50, 2024.

[82] C. Silpaja, «Speech-Based Classification of Parkinson's Disease and Essential Tremor: A Gaussian Mixture Models Approach,» AZoAi, 20 February 2024. [Ηλεκτρονικό]. Available: https://www.azoai.com/news/20240220/Speech-Based-Classification-of-Parkinsons-Disease-and-Essential-Tremor-A-Gaussian-Mixture-Models-Approach.aspx.

[83] N. D. Pah, V. Indrawati και D. K. Kumar, «Voice-based SVM Model Reliability for Identifying Parkinson's Disease,» *IEEE Access,* τόμ. 11, pp. 1-10, 2023.

[84] D. S. C. P. Indu Richa, «A modified kNN algorithm to detect Parkinson's disease,» *Network Modeling Analysis in Health Informatics and Bioinformatics,* τόμ. 12, αρ. 24, 2023.

[85] A. A. Çelik Burak, « Distinguishing Parkinson's Patients Using Voice-Based Feature Extraction and Classification,» 2024, https://arxiv.org/abs/2501.14390.

[86] T. Tech, «K-Nearest Neighbor Algorithm for Machine Learning,» 2025. [Ηλεκτρονικό]. Available: https://www.tpointtech.com/k-nearest-neighbor-algorithm-for-machine-learning.

[87] N. Boualoulou, B. Nsiri, T. Belhoussine Drissi και S. Zayrit, «Speech Analysis for the Detection of Parkinson's Disease by Combined Use of Empirical Mode Decomposition, Mel Frequency Cepstral Coefficients, and the K-Nearest Neighbor Classifier,» *ITM Web of Conferences,* τόμ. 43, 2022.

[88] A. Rana, A. Dumka, R. Singh, M. Rashid, N. Ahmad και M. K. Panda, «An efficient machine learning approach for diagnosing Parkinson's disease by utilizing voice features,» *Electronics,* τόμ. 11, αρ. 22, p. 3782, 2022.

[89] D. F. Santos, «Parkinson's Disease Detection using XGBoost and Machine Learning,» *medRxiv,* pp. 1-15, 2023.

[90] P. Ghaheri, H. Nasiri, A. Shateri και A. Homafar, «Diagnosis of Parkinson's Disease Based on Voice Signals Using SHAP and Hard Voting Ensemble Method,» *medRxiv,* 2022.

[91] B. S. S. N. Sudha Laxmaiah, «Parkinson Disease Prediction Using XGBoost,» *International Journal of Engineering & Science Research,* τόμ. 13, αρ. 4, pp. 1-12, 2023.

[92] D. C. S. Wang Chen, «Imbalance-XGBoost: Leveraging weighted and focal losses for binary label-imbalanced classification with XGBoost,» *arXiv preprint arXiv:1908.01672,* 2021.

[93] R. Harode, «A Deep Dive into Boosting,» SFU Professional Computer Science, 4 February 2020. [Ηλεκτρονικό]. Available: https://medium.com/sfu-cspmp/xgboost-a-deep-dive-into-boosting-f06c9c41349.

[94] S. Adiga, S. P., V. Pai, A. Shetty, D. L. Devadiga, S. Patil και Sharadhi, «Machine Learning Based Early Detection of Parkinson's Disease using XGBoost and Random Forest,» *Cuestiones de Fisioterapia,* τόμ. 53, αρ. 3, pp. 1175-1184, 2024.

[95] Y. Felix, «Uncovering acoustic biomarkers to classify Parkinson disease through machine learning,» University of Central Florida, Orlando, FL, 2025.

[96] S. D. Francisco, «Parkinson's Disease Detection using XGBoost and Machine Learning,» *medRxiv preprint,* 2023.

[97] M. Dudek, D. Hemmerling, M. Kaczmarska, J. Stepien, M. Daniol, M. Wodzinski και M. Wojcik-Pedziwiatr, «Analysis of Voice, Speech, and Language Biomarkers of Parkinson's Disease Collected in a Mixed Reality Setting,» *Sensors,* τόμ. 25, αρ. 8, p. 2405, 2025.

[98] T.-G. C. Van Gelderen Lisanne, « Innovative Speech-Based Deep Learning Approaches for Parkinson's Disease Classification: A Systematic Review,» *Applied Sciences,* τόμ. 14, αρ. 17, p. 7873, 2024.

[99] W. H. Yue Zhen, «Parkinson's Disease Speech Classification Using 1D Convolutional Neural Networks,» *Frontiers in Computing and Intelligent Systems,* τόμ. 10, αρ. 3, pp. 15-17, 2024.

[100] M. H. Ali, S. L. Mohammed και A. Al-Naji, «Parkinson's disease detection from voice using artificial intelligence techniques: A review,» *AIP Conference Proceedings,* τόμ. 3232, αρ. 1, 2024.

[101] J. L. Crawford, «inguistic changes in spontaneous speech for detecting Parkinson's disease using large language models,» *PLOS Digital Health,* τόμ. 4, αρ. 2, 2025.

[102] A. V. Sohan και A. V. Meghana, «Comparative Analysis of Machine Learning Algorithms for Parkinson's Disease Prediction,» *International Research Journal of Engineering and Technology (IRJET),* τόμ. 10, αρ. 12, pp. 733-740, 2023.

[103] M. Z. Firdaus, «Classification of Parkinson's Disease Using Random Forest Classifier with 1D Patient Speech Recordings,» Medium, 27 March 2023. [Ηλεκτρονικό]. Available: https://medium.com/@muhammad.zaky.firdaus/classification-of-parkinsons-disease-using-random-forest-classifier-with-1d-patient-speech-d42e85b45188.

[104] Y.-D. S. C.-L. S. J.-Y. L. E. C. L. a. H.-J. K. S.-M. Jeong, «Machine learning-based classification of Parkinson's disease using acoustic features: Insights from multilingual speech tasks,» *Computers in Biology and Medicine,* τόμ. 182, 2024.

[105] F. yben, K. R. Scherer, B. W. Schuller, J. Sundberg, E. André, C. Busso, L. Y. Devillers, J. Epps, P. Laukka, S. S. Narayanan και K. P. Truong, «The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing,» *IEEE Transactions on Affective Computing,* τόμ. 7, αρ. 2, p. 190–202, 2016.

[106] L. v. G. a. C. Tejedor-García, «Innovative Speech-Based Deep Learning Approaches for Parkinson's Disease Classification: A Systematic Review,» *Applied Sciences,* τόμ. 14, αρ. 17, pp. 73-78, 2024.

[107] L. Simone, M. G. Camporeale, V. M. Rubino, V. Gervasi και G. Dimauro, «Interpretable Early Detection of Parkinson's Disease through Speech Analysis,» *arXiv Preprint*, 2025.

[108] R. G. E. I. Alshammri.

[109] T. M. W. J. W. L. e. a. M. D. Rizzo, «Voice as a biomarker in Parkinson’s disease: A systematic review and guidelines for future research,» *Frontiers in Neurology*, τόμ. 12, αρ. 640587, p. DOI: 10.3389/fnagi.2021.640587, 2021.

Παράρτημα Α – Αναλυτική Περιγραφή των Λεκτικών Ασκήσεων ανά Σετ

Ο παρακάτω πίνακας παρουσιάζει αναλυτικά τα ηχητικά αρχεία που περιλαμβάνονται σε κάθε σετ ασκήσεων, όπως αυτές ορίζονται στο πειραματικό πρωτόκολλο. Στο κάθε σετ ασκήσεων εντοπίζονται ασκήσεις προφορικού λόγου που στοχεύουν στην αξιολόγηση των διαφορετικών πτυχών της ομιλίας των ασθενών.

Πίνακας 23: Αναλυτική Παρουσίαση των Σετ Ασκήσεων και των Περιλαμβανόμενων Ηχητικών Αρχείων.

Task	Ηχητικά Αρχεία
Task 1 (A1)	a1.wav, a2.wav, a3.wav, z1.wav, z2.wav, z3.wav, s1.wav, s2.wav, s3.wav
Task 2 (A2)	sakis.wav, say.wav, school.wav, wind.wav, yesno.wav, lentils.wav, mom1.wav, mom2.wav, illness.wav, earth.wav, ball1.wav, ball2.wav
Task 3 (A3)	rainbow_passage.wav
Task 4 (A4)	oe.wav, pa.wav, pataka.wav, ta.wav, ka.wav, kala.wav, aui.wav
Task 5 (A5)	f_autoind.wav, f_bat.wav, f_bicycle.wav, f_cinema.wav, f_cockroachkiller.wav, f_mailman.wav, f_pastrychef.wav, f_radio.wav, f_ring.wav, f_rose.wav, f_safe.wav, f_snail.wav, f_spaceship.wav, f_telecom.wav, s_autoind.wav, s_bat.wav, s_bicycle.wav, s_cinema.wav, s_cockroachkiller.wav, s_mailman.wav, s_pastrychef.wav, s_radio.wav, s_ring.wav, s_rose.wav, s_safe.wav, s_snail.wav, s_spaceship.wav, s_telecom.wav
Task 6 (A6)	counting.wav
Task 7 (A7)	breathing.wav
Task 8 (A8)	coughing.wav

Παράρτημα Β – Κατηγοριοποίηση Φωνητικών Χαρακτηριστικών του Συνόλου eGeMAPS

Στο παρόν παράρτημα γίνεται αναλυτική περιγραφή των ακουστικών παράμετρων του συνόλου eGeMAPS, ταξινομημένων στις πέντε βασικές κατηγορίες: φωνητικά, φασματικά, ενεργειακά, ρυθμικά/χρονικά και μη ταξινομημένα χαρακτηριστικά.
Οι πίνακες παρακάτω παραθέτουν αναλυτικά τα ονόματα των χαρακτηριστικών ανά κατηγορία

1. Φωνητικά Χαρακτηριστικά (38 χαρακτηριστικά)

Πίνακας 24:Συγκεντρωτικός Πίνακας Φωνητικών Χαρακτηριστικών.

Λίστα Φωνητικών Χαρακτηριστικών	
1. F0semitoneFrom27.5Hz_sma3nz_amean	2. F0semitoneFrom27.5Hz_sma3nz_stddevNorm
3. F0semitoneFrom27.5Hz_sma3nz_percentile80.0	4. F0semitoneFrom27.5Hz_sma3nz_percentile20.0
5. F0semitoneFrom27.5Hz_sma3nz_range	6. F0semitoneFrom27.5Hz_sma3nz_stddevRisingSlope
7. F0semitoneFrom27.5Hz_sma3nz_meanFallingSlope	8. F1frequency_sma3nz_amean
9. F1frequency_sma3nz_stddevNorm	10. F2frequency_sma3nz_amean
11. F2frequency_sma3nz_stddevNorm	12. F3frequency_sma3nz_amean
13. F3frequency_sma3nz_stddevNorm	14. F1bandwidth_sma3nz_amean
15. F1bandwidth_sma3nz_stddevNorm	16. F2bandwidth_sma3nz_amean
17. F2bandwidth_sma3nz_stddevNorm	18. F3bandwidth_sma3nz_amean
19. F3bandwidth_sma3nz_stddevNorm	20. shimmerLocaldB_sma3nz_amean
21. shimmerLocaldB_sma3nz_stddevNorm	22. HNRdBACF_sma3nz_amean
23. HNRdBACF_sma3nz_stddevNorm	24. jitterLocal_sma3nz_amean
25. jitterLocal_sma3nz_stddevNorm	26. F1amplitudeLogRelF0_sma3nz_amean
27. F2amplitudeLogRelF0_sma3nz_amean	28. F3amplitudeLogRelF0_sma3nz_amean
29. F1amplitudeLogRelF0_sma3nz_stddevNorm	30. F2amplitudeLogRelF0_sma3nz_stddevNorm
31. F3amplitudeLogRelF0_sma3nz_stddevNorm	32. VoicedSegmentsPerSec

33. MeanVoicedSegmentLengthSec	34. StddevVoicedSegmentLengthSec
35. MeanUnvoicedSegmentLength	36. StddevUnvoicedSegmentLength
37. voicingFinalUnclipped_sma3nz_amean	38. voicingFinalUnclipped_sma3nz_stddevNorm

2. Φασματικά Χαρακτηριστικά (24 χαρακτηριστικά)

Πίνακας 25: Συγκεντρωτικός Πίνακας Φασματικών Χαρακτηριστικών.

Λίστα Φασματικών Χαρακτηριστικών	
1. alphaRatioV_sma3nz_amean	2. alphaRatioV_sma3nz_stddevNorm
3. hammarbergIndexV_sma3nz_amean	4. hammarbergIndexV_sma3nz_stddevNorm
5. slopeV0-500_sma3nz_amean	6. slopeV0-500_sma3nz_stddevNorm
7. slopeV500-1500_sma3nz_amean	8. slopeV500-1500_sma3nz_stddevNorm
9. spectralFlux_sma3_amean	10. spectralFlux_sma3_stddevNorm
11. mfcc1_sma3_amean	12. mfcc1_sma3_stddevNorm
13. mfcc2_sma3_amean	14. mfcc2_sma3_stddevNorm
15. mfcc3_sma3_amean	16. mfcc3_sma3_stddevNorm
17. mfcc4_sma3_amean	18. mfcc4_sma3_stddevNorm
19. mfcc1V_sma3nz_amean	20. mfcc2V_sma3nz_amean
21. mfcc3V_sma3nz_amean	22. mfcc4V_sma3nz_amean
23. alphaRatioUV_sma3nz_amean	24. spectralFluxUV_sma3nz_amean

3. Ενεργειακά Χαρακτηριστικά (11 χαρακτηριστικά)

Πίνακας 26:Συγκεντρωτικός Πίνακας Ενεργειακών Χαρακτηριστικών.

Λίστα Ενεργειακών Χαρακτηριστικών	
1. loudness_sma3_amean	2. loudness_sma3_stddevNorm
3. loudness_sma3_percentile20.0	4. loudness_sma3_percentile50.0
5. loudness_sma3_percentile80.0	6. loudness_sma3_meanRisingSlope

7. loudness_sma3_stddevRisingSlope	8. loudness_sma3_meanFallingSlope
9. loudness_sma3_stddevFallingSlope	10. loudnessPeaksPerSec
11. pcm_RMSenergy_sma_amean	

4. Ρυθμικά/Χρονικά Χαρακτηριστικά (5 χαρακτηριστικά)

Πίνακας 27: Συγκεντρωτικός Πίνακας Ρυθμικών/Χρονικών Χαρακτηριστικών.

Λίστα Ρυθμικών/Χρονικών Χαρακτηριστικών	
1. F0semitoneFrom27.5Hz_sma3nz_stddevRisingSlope	2. F0semitoneFrom27.5Hz_sma3nz_meanFallingSlope
3. VoicedSegmentsPerSec	4. MeanVoicedSegmentLengthSec
5. StddevUnvoicedSegmentLength	

5. Μη Ταξινομημένα Χαρακτηριστικά (12 χαρακτηριστικά)

Πίνακας 28:Συγκεντρωτικός Πίνακας Μη Ταξινομημένων Χαρακτηριστικών.

Λίστα Μη Ταξινομημένων Χαρακτηριστικών	
1. pcm_loudness_sma_amean	2. pcm_loudness_sma_stddev
3. F0_sma_amean	4. F0_sma_stddev
5. intensity_sma_amean	6. intensity_sma_stddev
7. F1_sma_amean	8. F2_sma_amean
9. F3_sma_amean	10. mfcc0_sma_amean
11. mfcc0_sma_stddev	12. voicingProbability_sma_amean

Παράρτημα Γ – Αναλυτικά Διαγράμματα Στατιστικής Σημαντικότητας Χαρακτηριστικών

Στο παρόν παράρτημα παρουσιάζονται τα αναλυτικά διαγράμματα που απεικονίζουν τη στατιστική σημαντικότητα των εξαγόμενων χαρακτηριστικών για κάθε επιμέρους φωνητική άσκηση. Η στατιστική ανάλυση πραγματοποιήθηκε με στόχο τον εντοπισμό των παραμέτρων ομιλίας που παρουσιάζουν την μεγαλύτερη στατιστική σημαντικότητα ως προς τα κλινικά ερωτήματα που εξετάστηκαν (Hoehn & Yahr, Fluctuations, Dyskinesia, UPDRS_III_dif_class, UPDRS_III_dif_per).

Τα παρακάτω διαγράμματα αποτυπώνουν την κατανομή των p-value για κάθε χαρακτηριστικό, επιτρέποντας την οπτική ερμηνεία της σχετικής σημαντικότητασαυτών. Μέσω αυτής της διαδικασίας καθίσταται εφικτή η αναγνώριση των ασκήσεων και των χαρακτηριστικών που συμβάλλουν ουσιαστικά στη διαφοροποίηση των επιπέδων της νόσου του Πάρκινσον, παρέχοντας πολύτιμες ενδείξεις για την αξιοπιστία και διαγνωστική αξία των φωνητικών δεδομένων. Επίσης, τα διαγράμματα συνοδεύονται από σχετικούς πίνακες που περιλαμβάνουν τα πέντε (5) επικρατέστερα φωνητικά χαρακτηριστικά με την χαμηλότερη τιμή για την μεταβλητή p-value.Η παρουσίαση αυτή συμπληρώνει τη βασική ανάλυση που περιλαμβάνεται στο κυρίως σώμα της εργασίας και αποτελεί αναπόσπαστο μέρος της ερμηνείας των αποτελεσμάτων της.

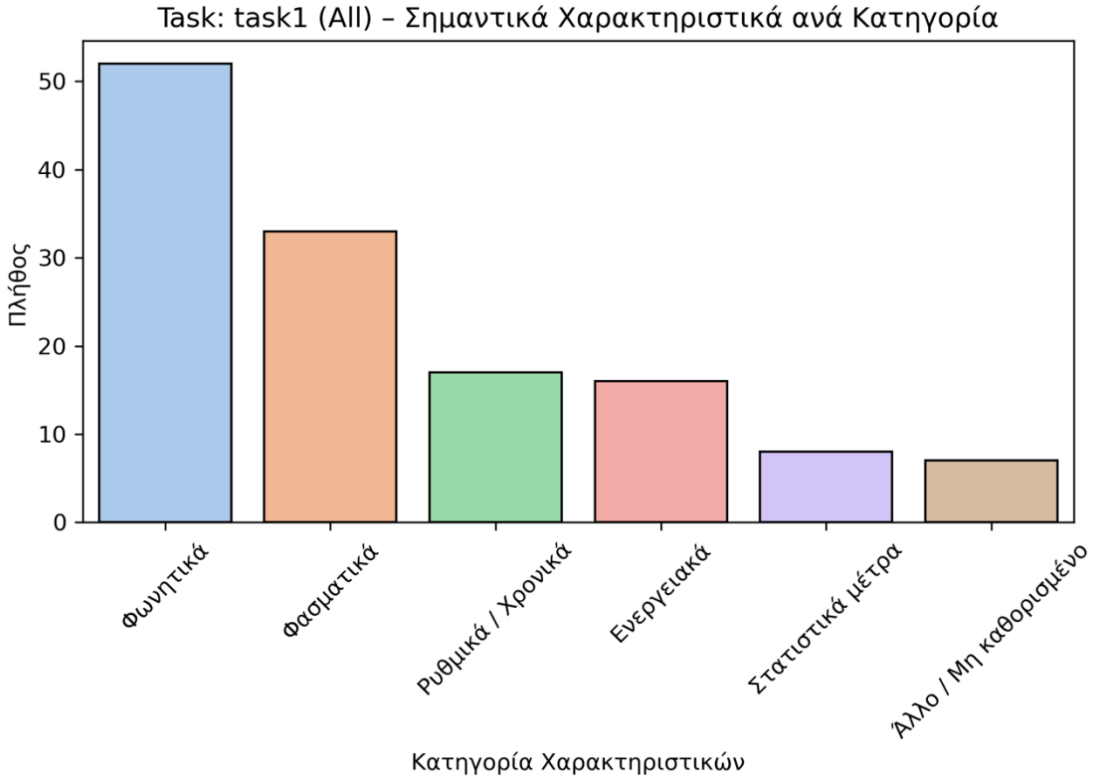
Task 1 (A1): Εκφώνηση Φωνηέντων/Συμφώνων – Σταθερότητα Φωνής

Το συγκεκριμένο Σετ Ασκήσεων περιλαμβάνει συνεχή εκφώνηση φωνηέντων και συμφώνων, με κύριο στόχο την αξιολόγηση της σταθερότητας της φωνής σε άτομα που νοσούν από την ΝΠ. Μέσω αυτής της επαναλαμβανόμενης φώνησης των συγκεκριμένων ασκήσεων μπορεί να πραγματοποιηθεί ανάλυση παραμέτρων, σημαντικών για την μετέπειτα αξιολόγηση. Αυτοί οι παράμετροι συνδέονται με την μυϊκή σταθερότητα των φωνητικών χορδών. Ο εντοπισμός διακυμάνσεων στις φωνητικές παραμέτρους, συνδέεται με πρώιμα στάδια της νόσου όπου εντοπίζεται συχνά απώλεια ελέγχου της φωνητικής παραγωγής.

Για την εκφώνηση φωνηέντων/συμφώνων παρακάτω παρουσιάζονται τα 5 σημαντικότερα ακουστικά χαρακτηριστικά (με p value < 0.05) για κάθε επιμέρους άσκηση που περιλαμβάνεται στο σετ ασκήσεων, καθώς και ένα συγκεντρωτικός πίνακας, που παρουσιάζει τα 5 κορυφαία χαρακτηριστικά για το σύνολο των ασκήσεων. Επιπλέον, παρουσιάζονται και τα διαγράμματα που απεικονίζουν το πλήθος των στατιστικά σημαντικών χαρακτηριστικών ανά κατηγορία.

Πίνακας 29:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για το σύνολο του Task 1 (A1)

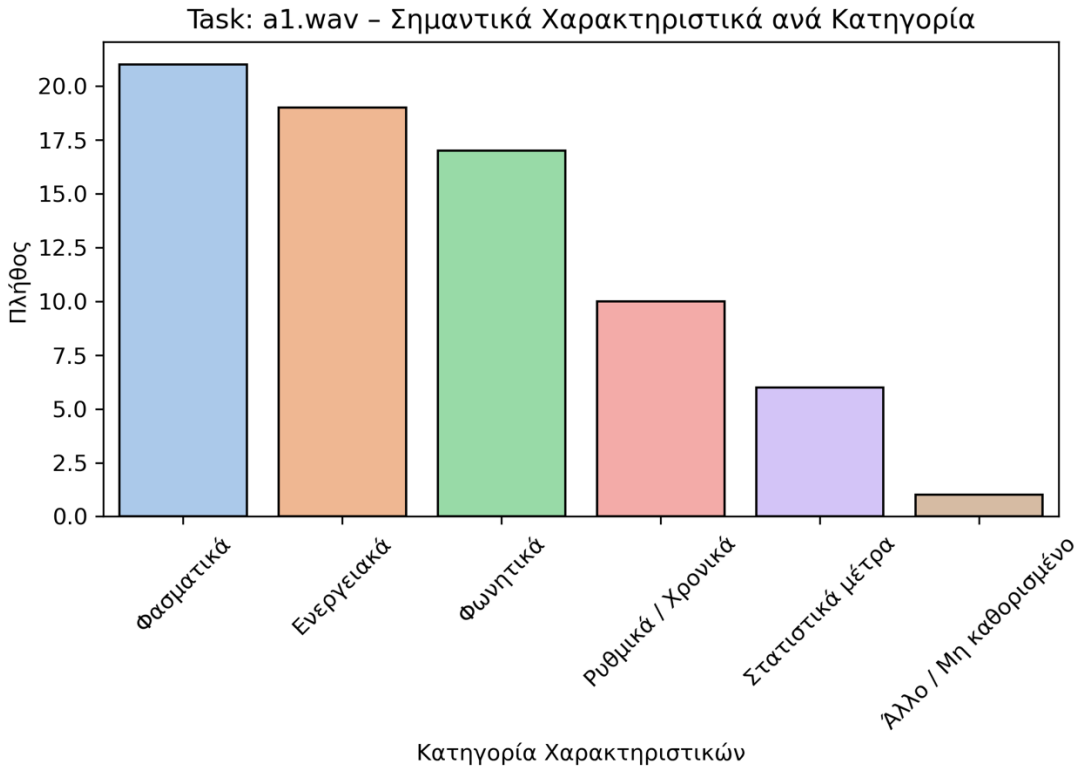
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
(Intercept)	1.98×10^{-15}	Άλλο / Μη καθορισμένο
(Intercept)	6.83×10^{-9}	Άλλο / Μη καθορισμένο
MeanVoicedSegmentLengthSec	3.79×10^{-7}	Ρυθμικά / Χρονικά
loudness_sma3_stddevNorm	8.01×10^{-6}	Ενεργειακά
(Intercept)	2.28×10^{-5}	Άλλο / Μη καθορισμένο



Εικόνα 49: Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για το σύνολο του Task 1

Πίνακας 30:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση a1.wav του Task 1 (A1)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2frequency_sma3nz_stddevNorm	0.001086	Φωνητικά
F1frequency_sma3nz_stddevNorm	0.001191	Φωνητικά
hammarbergIndexUV_sma3nz_amean	0.001372	Στατιστικά Μέτρα
spectralFluxUV_sma3nz_amean	0.001505	Φασματικά
loudness_sma3_percentile80.0	0.001708	Ενεργειακά



Εικόνα 50:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση a1.wav του Task 1

Πίνακας 31: Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση a2.wav του Task 1 (A1)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F1amplitudeLogRelF0_sma3nz_amean	0.006382	Φωνητικά
jitterLocal_sma3nz_stddevNorm	0.016111	Φωνητικά
(Intercept)	0.018215	Άλλο / Μη καθορισμένο
mfcc4_sma3_amean	0.019201	Φασματικά
F1bandwidth_sma3nz_stddevNorm	0.020987	Στατιστικά Μέτρα



Εικόνα 51:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση a2.wav του Task 1

Πίνακας 32:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση a3.wav του Task 1 (A1)

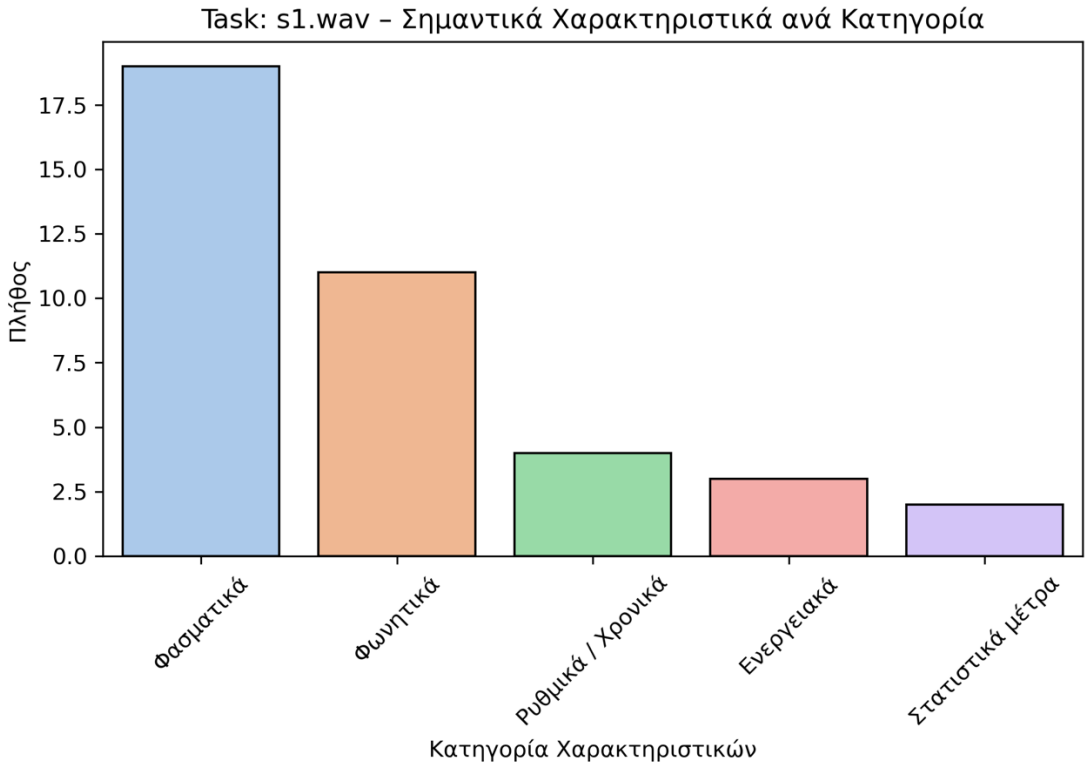
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
MeanVoicedSegmentLengthSec	0.000075	Ρυθμικά/Χρονικά
alphaRatioV_sma3nz_amean	0.000206	Φασματικά
loudness_sma3_stddevNorm	0.000394	Ενεργειακά
F0semitoneFrom27.5Hz_sma3nz_stddevRisingSlope	0.000595	Φωνητικά
mfcc3_sma3_amean	0.000950	Φασματικά



Εικόνα 52:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση a3.wav του Task 1

Πίνακας 33:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s1.wav του Task 1 (A1)

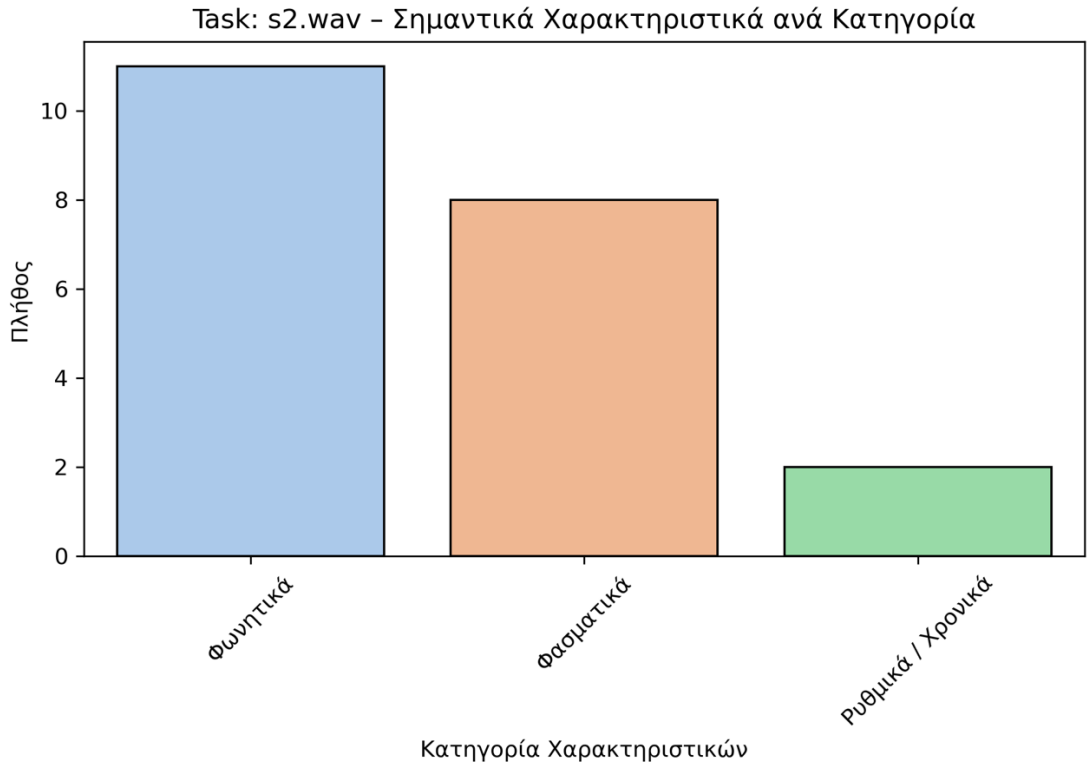
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc2_sma2_amean	0.001530	Φασματικά
mfcc2_sma3_amean	0.002404	Φασματικά
StddevUnvoicedSegmentLength	0.004831	Ρυθμικά/Χρονικά
alphaRatioUV_sma3nz_amean	0.005594	Φασματικά
alphaRatioV_sma3nz_stddevNorm	0.005842	Φασματικά



Εικόνα 53:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s1.wav του Task 1

Πίνακας 34:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s2.wav του Task 1 (A1)

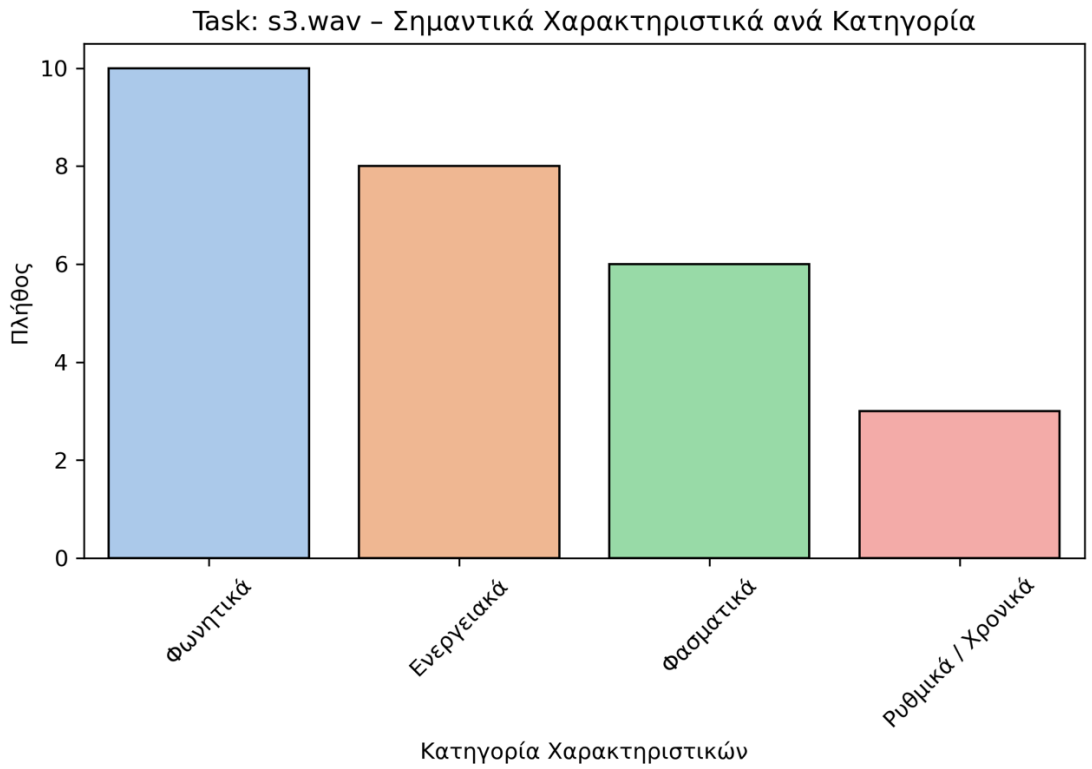
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2bandwidth_sma3nz_amean	0.000193	Φωνητικά
F3frequency_sma3nz_amean	0.000385	Φωνητικά
MeanUnvoicedSegmentLength	0.001675	Ρυθμικά/Χρονικά
mfcc2_sma3_amean	0.002289	Φασματικά
F1frequency_sma3nz_amean	0.004785	Φωνητικά



Εικόνα 54:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s2.wav του Task 1

Πίνακας 35:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s3.wav του Task 1 (A1)

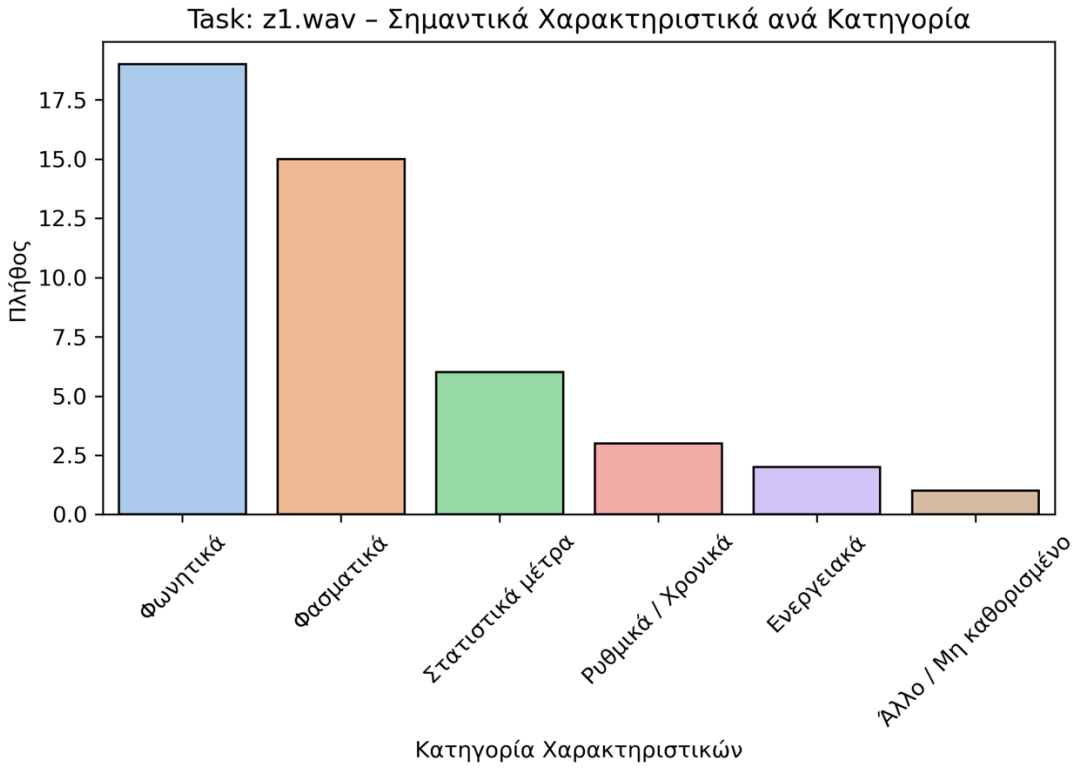
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc4V_sma3nz_amean	0.002650	Φασματικά
mfcc4_sma3_stddevNorm	0.003582	Φασματικά
loudness_sma3_stddevNorm	0.010086	Ενεργειακά
loudness_sma3_percentile20.0	0.016941	Ενεργειακά
HNRdBACF_sma3nz_stddevNorm	0.018640	Φωνητικά



Εικόνα 55:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s3.wav του Task 1

Πίνακας 36:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση z1.wav του Task 1 (A1)

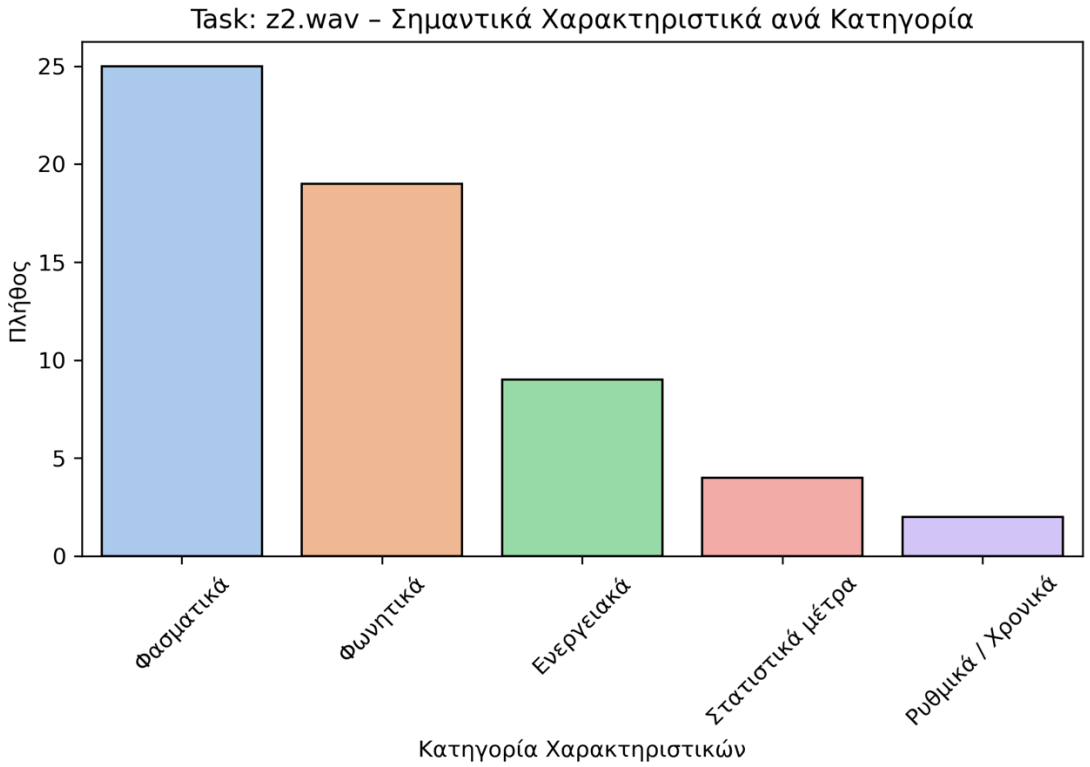
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F1amplitudeLogRelF0_sma3nz_stddevNorm	0.000233	Φωνητικά
F2amplitudeLogRelF0_sma3nz_stddevNorm	0.000392	Φωνητικά
F3amplitudeLogRelF0_sma3nz_amean	0.000498	Φωνητικά
F1bandwidth_sma3nz_stddevNorm	0.000666	Στατιστικά Μέτρα
mfcc2V_sma3nz_stddevNorm	0.000902	Φασματικά



Εικόνα 56:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση z1.wav του Task 1

Πίνακας 37:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση z2.wav του Task 1 (A1)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
jitterLocal_sma3nz_amean	0.000005	Φωνητικά
jitterLocal_sma3nz_amean	0.000279	Φωνητικά
mfcc1_sma3_stddevNorm	0.001416	Φασματικά
StddevVoicedSegmentLengthSec	0.002607	Ρυθμικά/Χρονικά
loudness_sma3_meanRisingSlope	0.002778	Ενεργειακά



Εικόνα 57:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση z2.wav του Task 1

Πίνακας 38:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση z3.wav του Task 1 (A1)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc3V_sma3nz_stddevNorm	0.000095	Φασματικά
mfcc3_sma3_stddevNorm	0.000102	Φασματικά
mfcc2V_sma3nz_stddevNorm	0.001565	Φασματικά
mfcc2_sma3_stddevNorm	0.005469	Φασματικά
F2frequency_sma3nz_amean	0.007168	Φωνητικά



Εικόνα 58:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση z3.wav του Task 1

Task 2 (A2): Εκφώνηση Λέξεων/Φράσεων – Αξιολόγηση Άρθρωσής

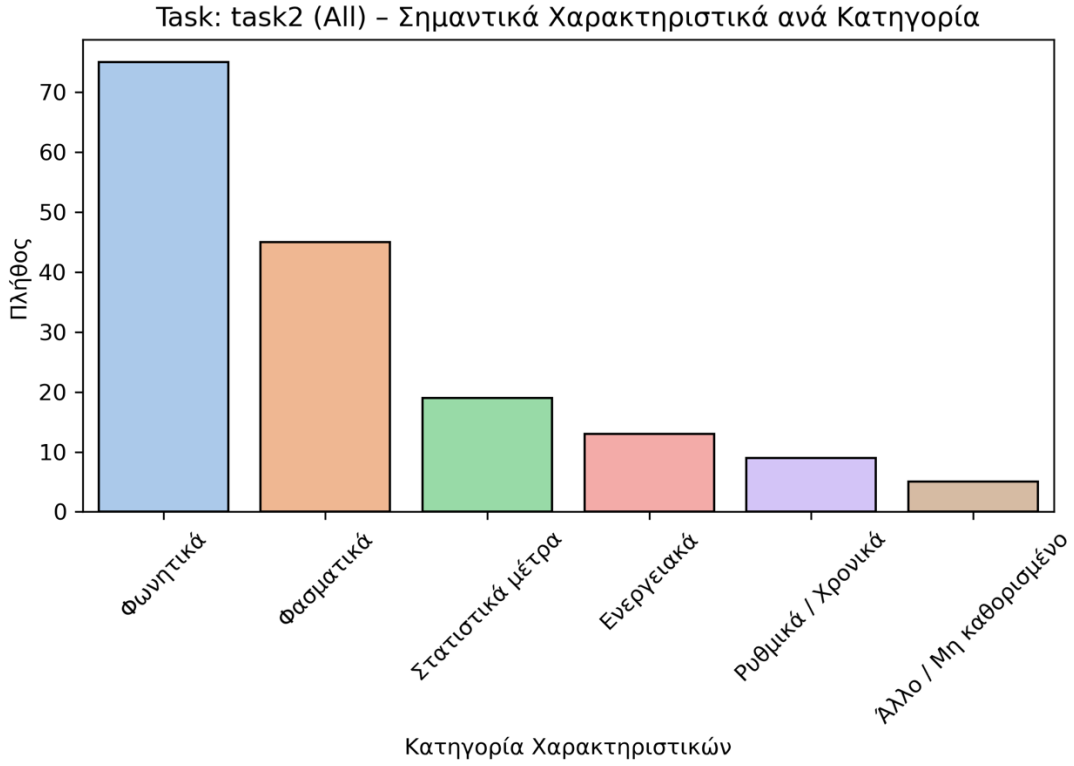
Η δεύτερη ενότητα περιλαμβάνει της εκφορά διαφόρων λέξεων και φράσεων προκειμένου να γίνει αξιολόγηση της σταθερότητας και της ευχέρειας της άρθρωσης στην ομιλία του ατόμου. Αποκλίσεις που εντοπίζονται στις παραμέτρους που συνδέονται με τις συγκεκριμένες ασκήσεις, υποδεικνύουν την παρουσία δυσαρθρίας, ενός κοινού συμπτώματος στην ΝΠ. Οι

ασκήσεις αυτές δίνουν την δυνατότητα συστηματικής αξιολόγησης της κινητικότητας των αρθρωτών μηχανισμών μέσω ανάλυσης χαρακτηριστικών όπως η φασματική ενεργειακή κατανομή.

Παρακάτω παρουσιάζονται αναλυτικά τα εξαγόμενα αποτελέσματα που επικεντρώνονται στην ανάλυση του συγκεκριμένου σετ ασκήσεων.

Πίνακας 39:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για το σύνολο του Task 2 (A2)

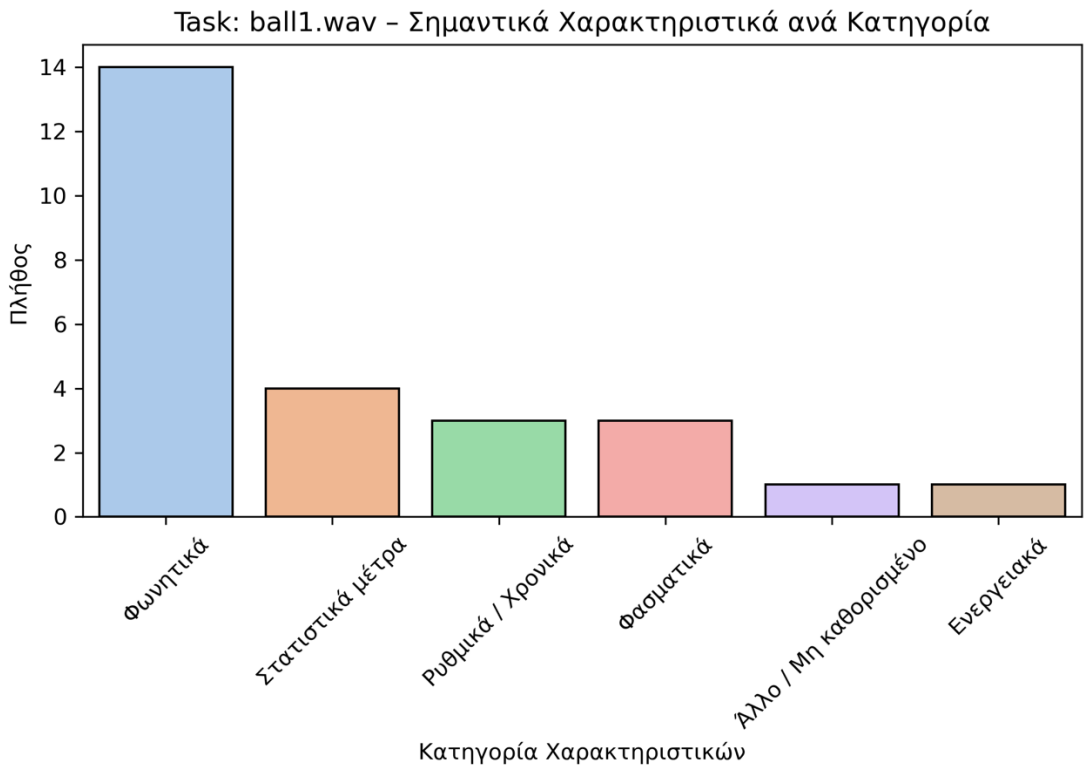
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F3frequency_sma3nz_amean	4.96×10^{-14}	Φωνητικά
F2amplitudeLogRelF0_sma3nz_amean	1.71×10^{-11}	Φωνητικά
F1amplitudeLogRelF0_sma3nz_amean	1.45×10^{-9}	Φωνητικά
alphaRatioV_sma3nz_stddevNorm	5.01×10^{-9}	Φασματικά
F2amplitudeLogRelF1_sma3nz_amean	5.93×10^{-9}	Φωνητικά



Εικόνα 59:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για το σύνολο του Task

Πίνακας 40:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση ball1.wav του Task 2 (A2)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2amplitudeLogRelF0_sma3nz_amean	0.004541	Φωνητικά
spectralFlux_sma3_stddevNorm	0.007483	Φασματικά
F3amplitudeLogRelF0_sma3nz_amean	0.007732	Φωνητικά
StddevUnvoicedSegmentLength	0.012240	Ρυθμικά/Χρονικά
mfcc4V_sma3nz_stddevNorm	0.013687	Φασματικά



Εικόνα 60:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση ball1.wav του Task 2

Πίνακας 41:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση ball2.wav του Task 2 (A2)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F3frequency_sma3nz_amean	0.004541	Φωνητικά
loudnessPeaksPerSec	0.007483	Ενεργειακά
loudnessPeaksPerSec	0.007732	Ενεργειακά
F3frequency_sma3nz_amean	0.012240	Φωνητικά
shimmerLocaldB_sma3nz_amean	0.013687	Φωνητικά



Εικόνα 61:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση ball2.wav του Task 2

Πίνακας 42:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση earth.wav του Task 2 (A2)

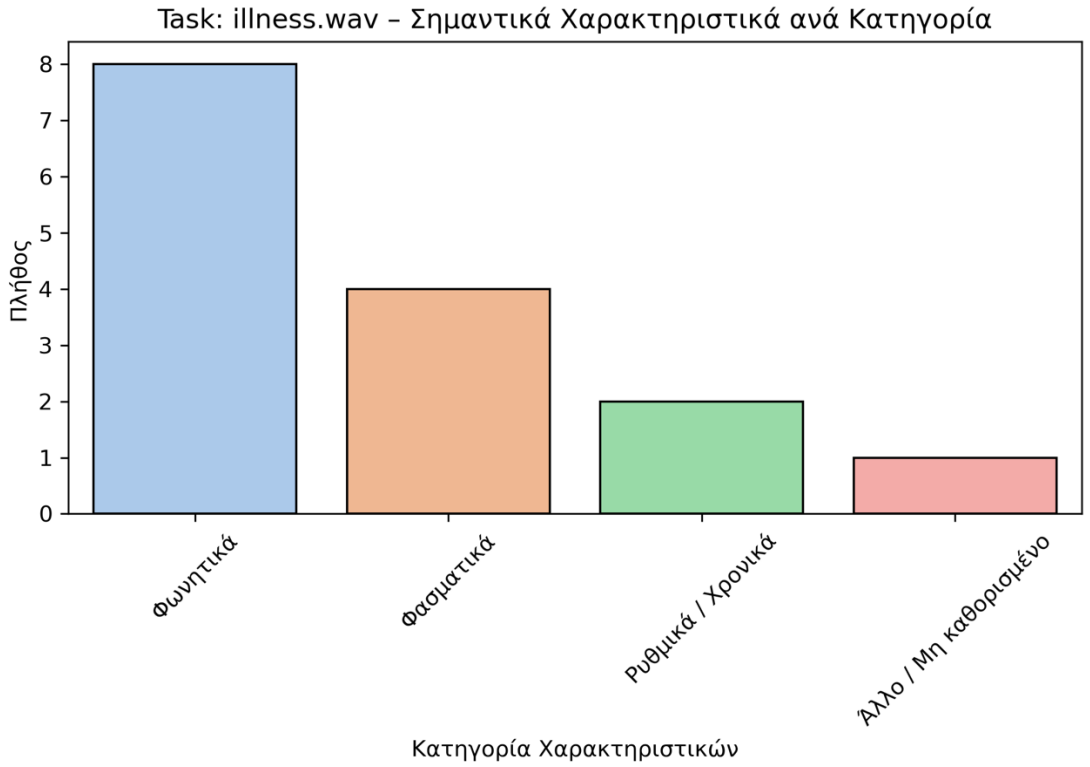
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
loudnessPeaksPerSec	0.004541	Ενεργειακά
MeanUnvoicedSegmentLength	0.007483	Ρυθμικά/Χρονικά
MeanUnvoicedSegmentLength	0.007732	Ρυθμικά/Χρονικά
jitterLocal_sma3nz_amean	0.012240	Φωνητικά
loudness_sma3_stddevRisingSlope	0.013687	Ενεργειακά



Εικόνα 62:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση earth.wav του Task 2

Πίνακας 43:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση illness.wav του Task 2 (A2)

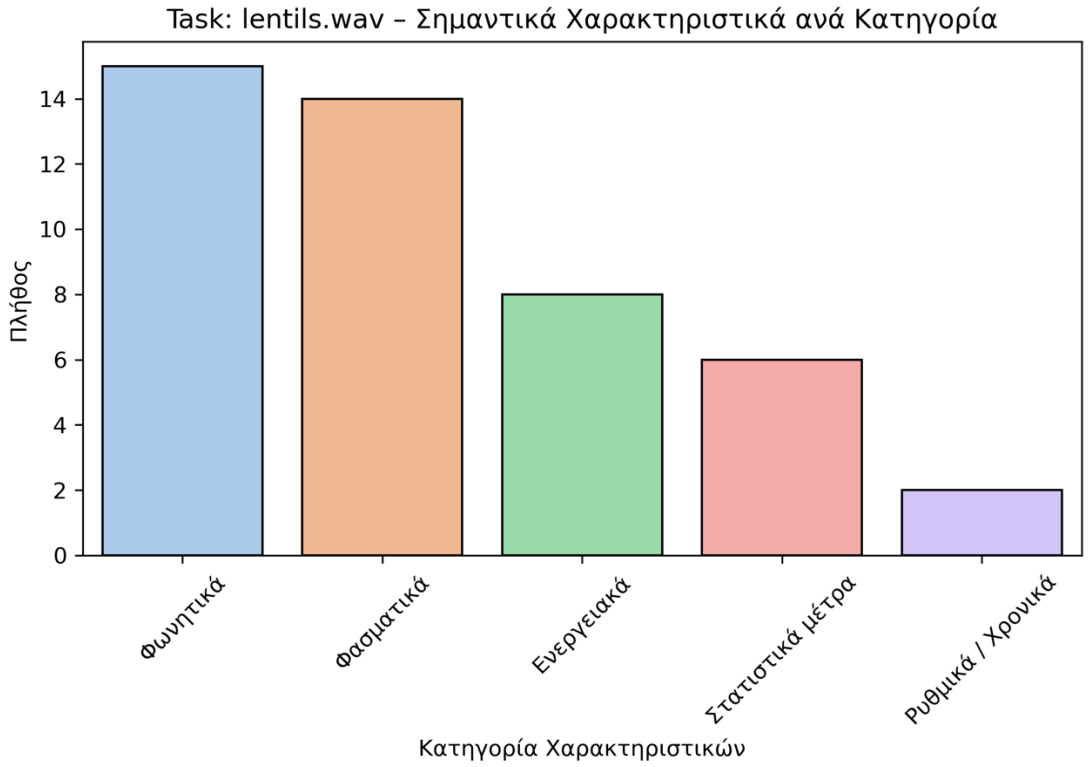
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc3V_sma3nz_stddevNorm	0.004541	Φασματικά
F3frequency_sma3nz_amean	0.007483	Φωνητικά
(Intercept)	0.007732	Άλλο / Μη καθορισμένο
F1amplitudeLogRelF0_sma3nz_stddevNorm	0.012240	Φωνητικά
StddevUnvoicedSegmentLength	0.013687	Ρυθμικά/Χρονικά



Εικόνα 63:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση illness.wav του Task 2

Πίνακας 44:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση lentils.wav του Task 2 (A2)

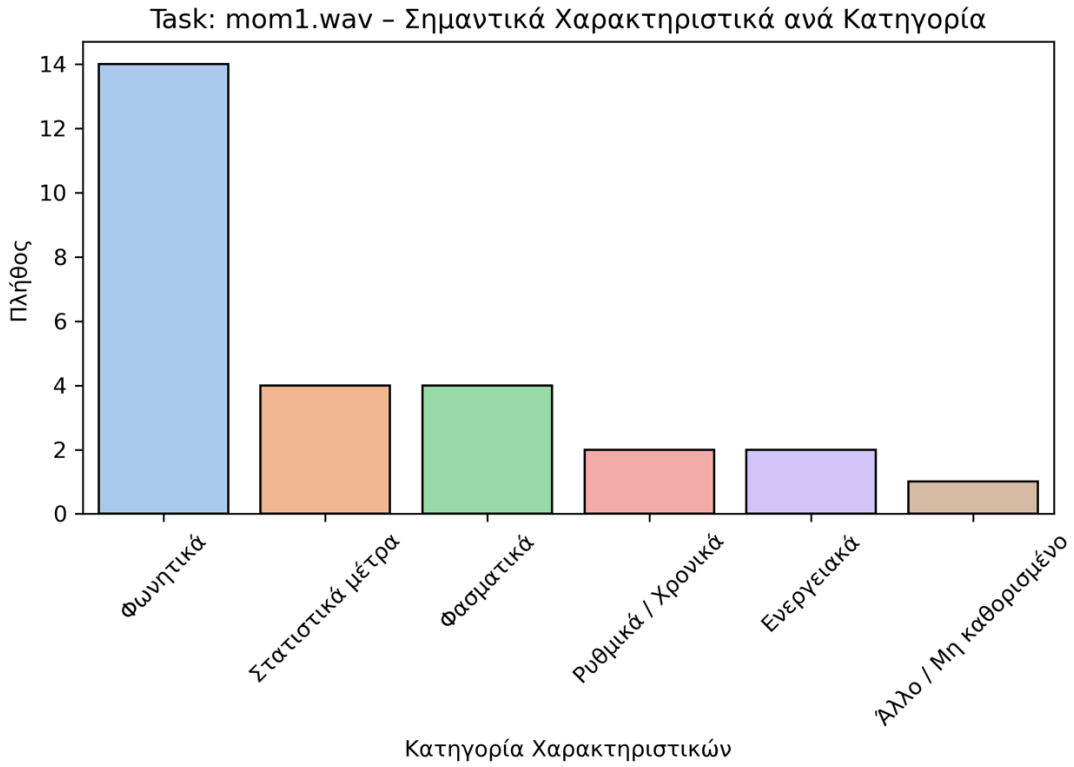
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc3_sma3_amean	0.000543	Φασματικά
hammarbergIndexV_sma3nz_amean	0.003441	Στατιστικά μέτρα
mfcc2V_sma3nz_stddevNorm	0.004041	Φασματικά
alphaRatioUV_sma3nz_amean	0.004560	Φασματικά
shimmerLocaldB_sma3nz_amean	0.006300	Φωνητικά



Εικόνα 64:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση lentils.wav του Task 2

Πίνακας 45:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση mom1.wav του Task 2 (A2)

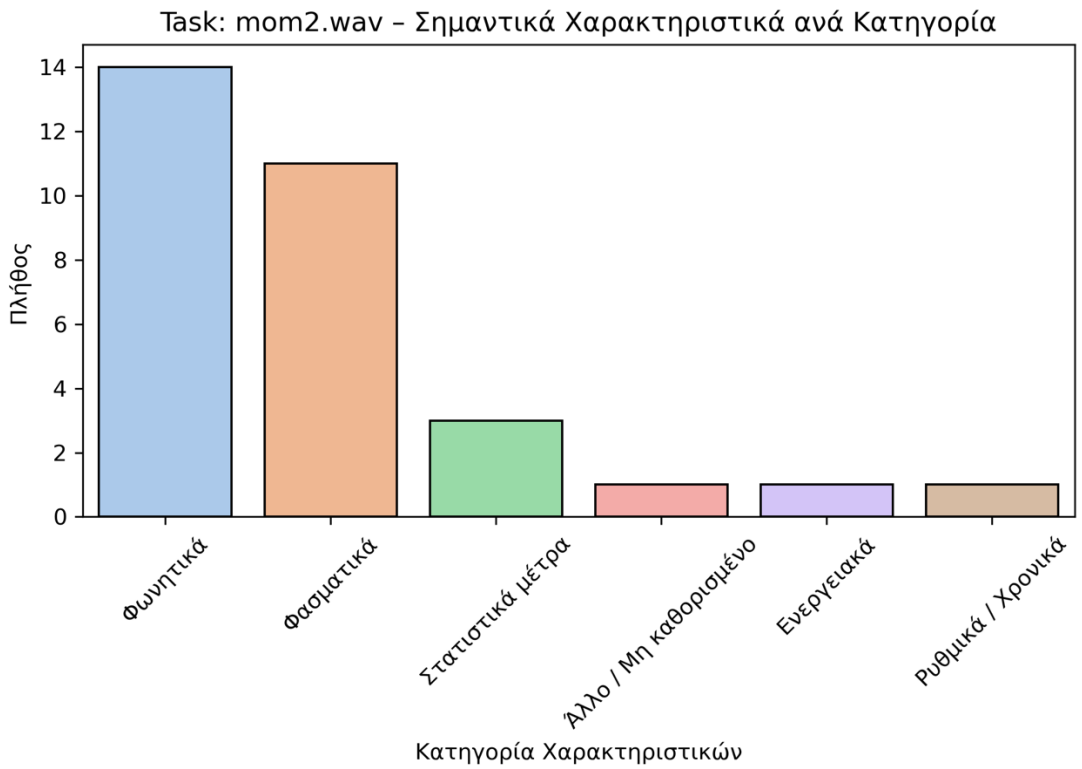
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F1bandwidth_sma3nz_stddevNorm	0.000176	Στατιστικά μέτρα
F3frequency_sma3nz_amean	0.000275	Φωνητικά
F2bandwidth_sma3nz_amean	0.000679	Φωνητικά
F1bandwidth_sma3nz_amean	0.001489	Στατιστικά μέτρα
F0semitoneFrom27.5Hz_sma3nz_percentile80.0	0.001626	Φωνητικά



Εικόνα 65:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση mom1.wav του Task 2

Πίνακας 46:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση mom2.wav του Task 2 (A2)

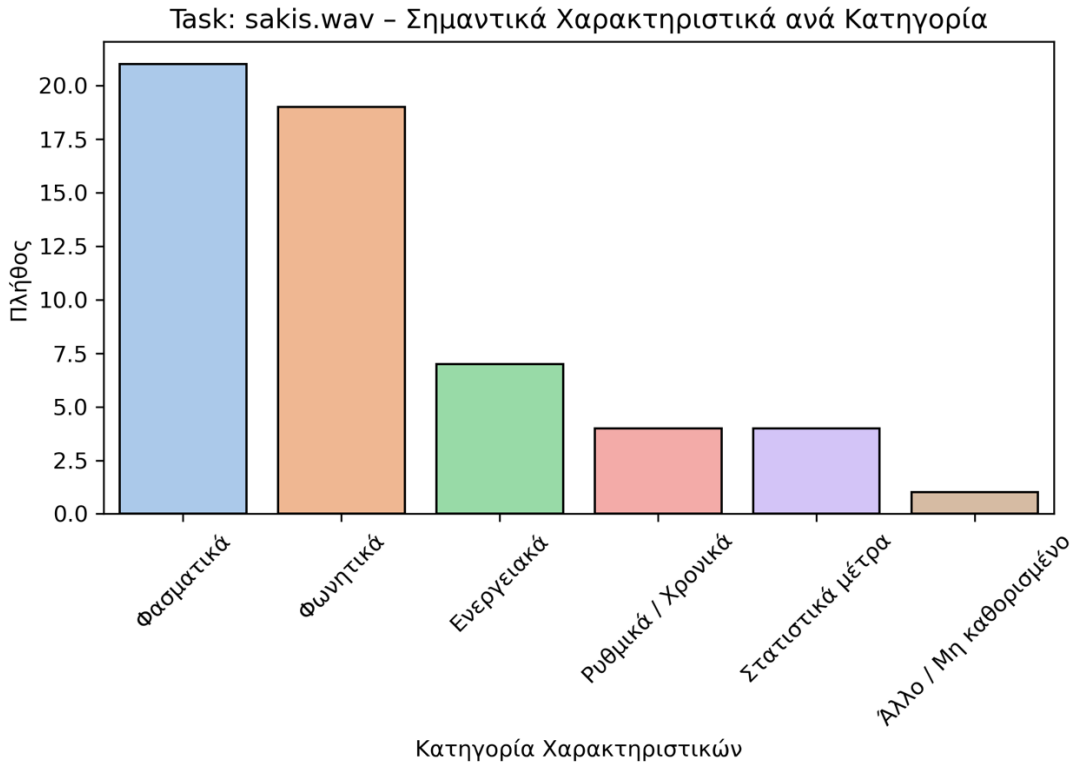
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2bandwidth_sma3nz_amean	0.000046	Φωνητικά
F3frequency_sma3nz_amean	0.000076	Φωνητικά
F1bandwidth_sma3nz_amean	0.001174	Στατιστικά μέτρα
F0semitoneFrom27.5Hz_sma3nz_percentile20.0	0.005472	Φωνητικά
(Intercept)	0.005816	Άλλο / Μη καθορισμένο



Εικόνα 66:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση mom2.wav του Task 2

Πίνακας 47:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση sakis.wav του Task 2 (A2)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
hammarbergIndexUV_sma3nz_amean	0.000256	Στατιστικά μέτρα
(Intercept)	0.000272	Άλλο / Μη καθορισμένο
F3frequency_sma3nz_amean	0.000274	Φωνητικά
alphaRatioUV_sma3nz_amean	0.000295	Φασματικά
mfcc1_sma3_amean	0.000318	Φασματικά



Εικόνα 67:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση sakis.wav του Task 2

Πίνακας 48:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση say.wav του Task 2 (A2)

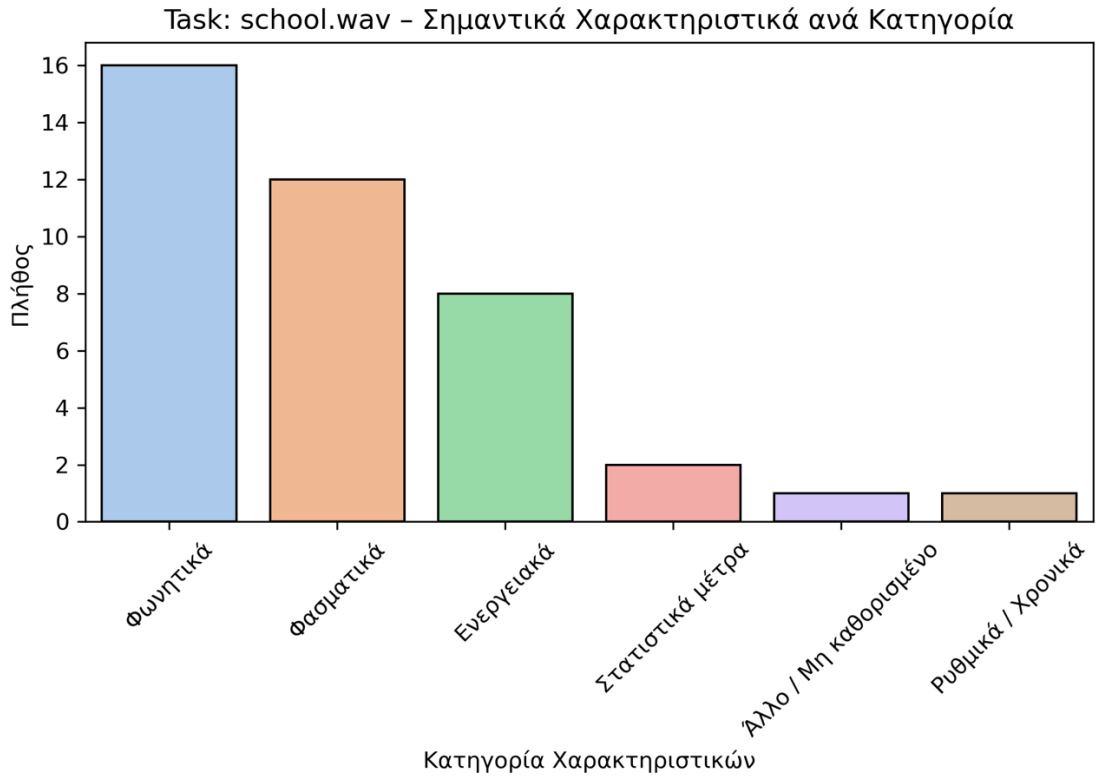
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F1amplitudeLogRelF0_sma3nz_amean	0.000569	Φωνητικά
mfcc4V_sma3nz_amean	0.000587	Φασματικά
mfcc4_sma3_amean	0.000786	Φασματικά
alphaRatioV_sma3nz_stddevNorm	0.001390	Φασματικά
spectralFlux_sma3_amean	0.002820	Φασματικά



Εικόνα 68:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση say.wav του Task 2

Πίνακας 49:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση school.wav του Task 2 (A2)

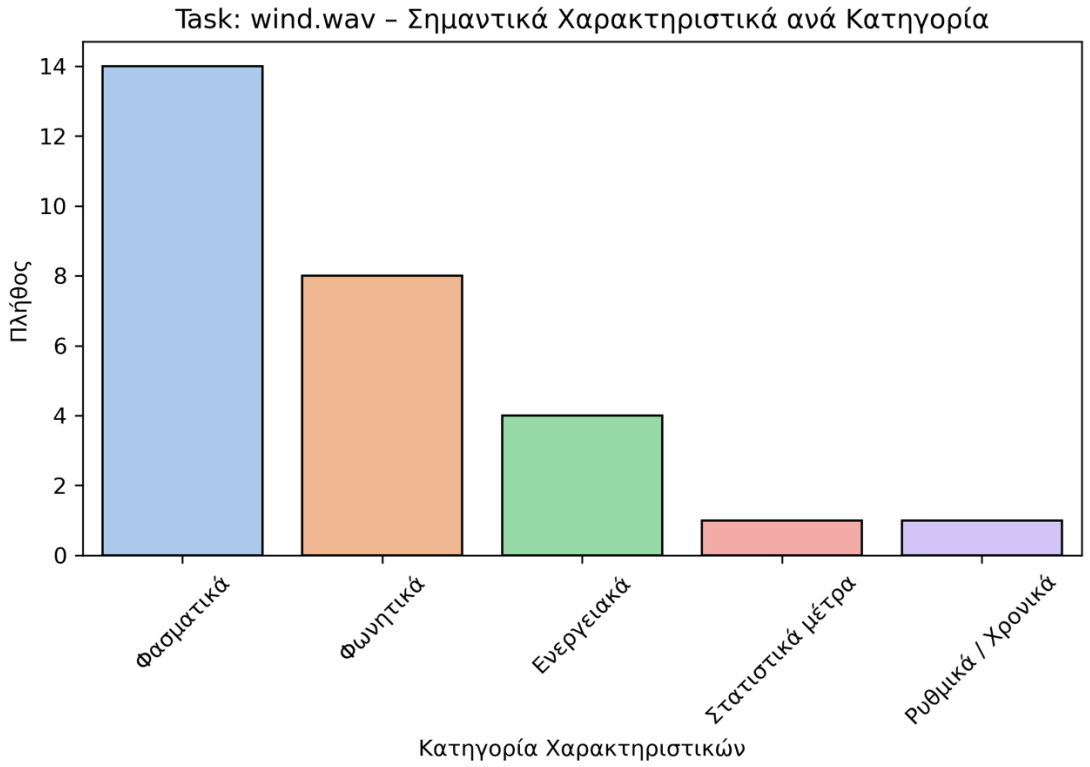
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
loudness_sma3_percentile50.0	0.000016	Ενεργειακά
loudness_sma2_percentile50.0	0.000302	Ενεργειακά
F0semitoneFrom27.5Hz_sma3nz_percentile50.0	0.000677	Φωνητικά
F0semitoneFrom27.5Hz_sma3nz_stddevNorm	0.000988	Φωνητικά
F1amplitudeLogRelF0_sma3nz_stddevNorm	0.001044	Φωνητικά



Εικόνα 69:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση school.wav του Task 2

Πίνακας 50:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση wind.wav του Task 2 (A2)

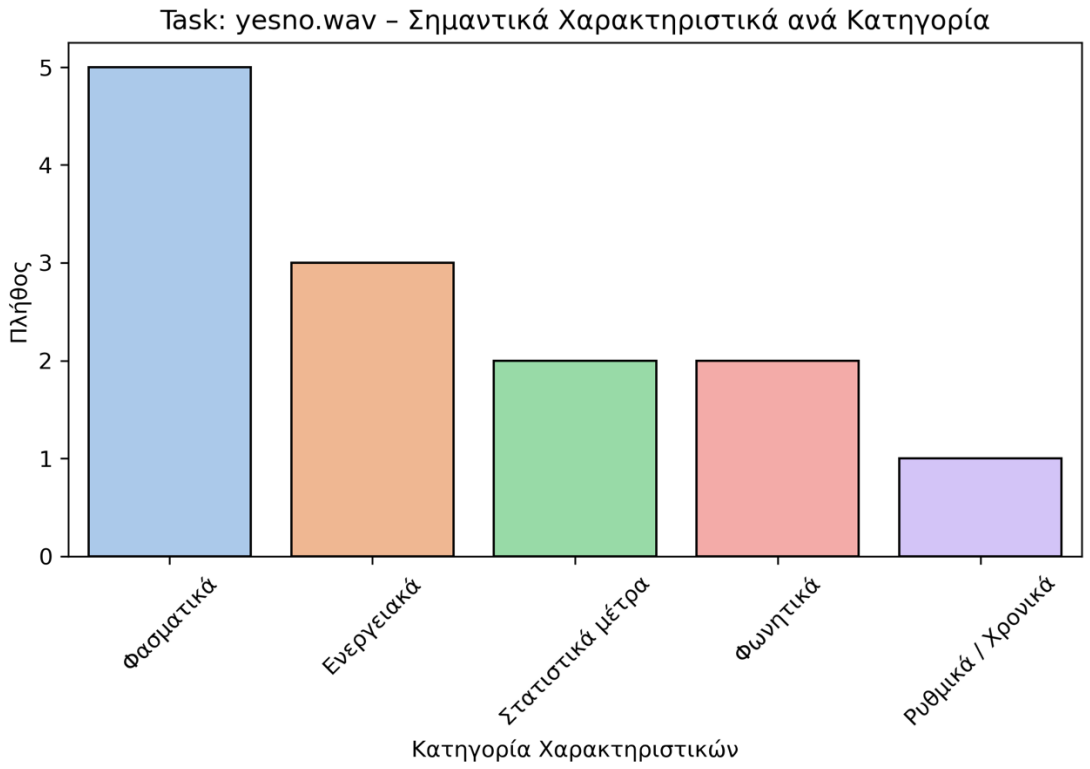
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
alphaRatioV_sma3nz_stddevNorm	0.006644	Φασματικά
mfcc4V_sma3nz_amean	0.008425	Φασματικά
mfcc2_sma3_amean	0.008858	Φασματικά
loudness_sma3_percentile20.0	0.009155	Ενεργειακά
mfcc4_sma3_amean	0.009892	Φασματικά



Εικόνα 70:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση wind.wav του Task 2

Πίνακας 51:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση yesno.wav του Task 2 (A2)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
loudness_sma3_percentile80.0	0.009952	Ενεργειακά
hammarbergIndexV_sma3nz_stddevNorm	0.020101	Στατιστικά μέτρα
loudness_sma3_amean	0.023446	Ενεργειακά
mfcc3_sma3_stddevNorm	0.025667	Φασματικά
F0semitoneFrom27.5Hz_sma3nz_stddevFallingSlope	0.027452	Φωνητικά



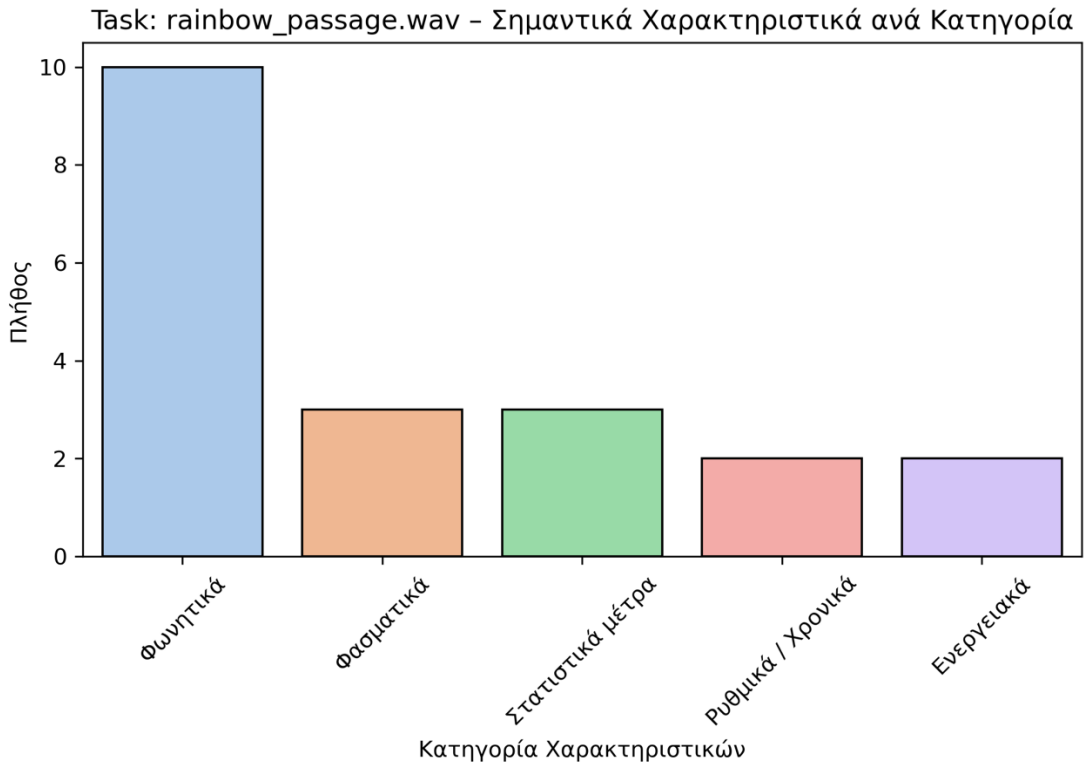
Εικόνα 71:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση yesno.wav του Task 2

Task 3 (A3): Συνεχής Ανάγνωση – Συνολική Αξιολόγηση Ομιλίας

Η ανάγνωση ενός συγκεκριμένου κειμένου που έχει συνταχθεί, δίνει την δυνατότητα να γίνει μια συνολική εκτίμηση των χαρακτηριστικών ομιλίας του ατόμου που νοσεί, χαρακτηριστικά τα οποία μπορούν να συνδεθούν με την παρουσία και βαρύτητα της νόσου. Αυτά τα χαρακτηριστικά είναι η προσωδία, ο ρυθμός ομιλίας και οι εκτενής παύσεις κατά την διάρκεια της ομιλίας. Η ανάλυση αυτής της άσκησης συνεισφέρει στην εξαγωγή χαρακτηριστικών με υψηλή διαγνωστική αξία , που θα βοηθήσουν στην μετέπειτα ανάλυση. Τα αποτελέσματα που προέκυψαν αναλύονται στον πίνακα που ακολουθεί, με τα χαρακτηριστικά με την υψηλότερη διακριτική ικανότητα.

Πίνακας 52:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για το σύνολο του Task 3 (A3)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
loudness_sma3_stddevFallingSlope	0.000330	Ενεργειακά
HNRdBACF_sma3nz_amean	0.001982	Φωνητικά
shimmerLocaldB_sma3nz_stddevNorm	0.004306	Φωνητικά
F0semitoneFrom27.5Hz_sma3nz_percentile80.0	0.007715	Φωνητικά
loudness_sma3_meanFallingSlope	0.008551	Ενεργειακά



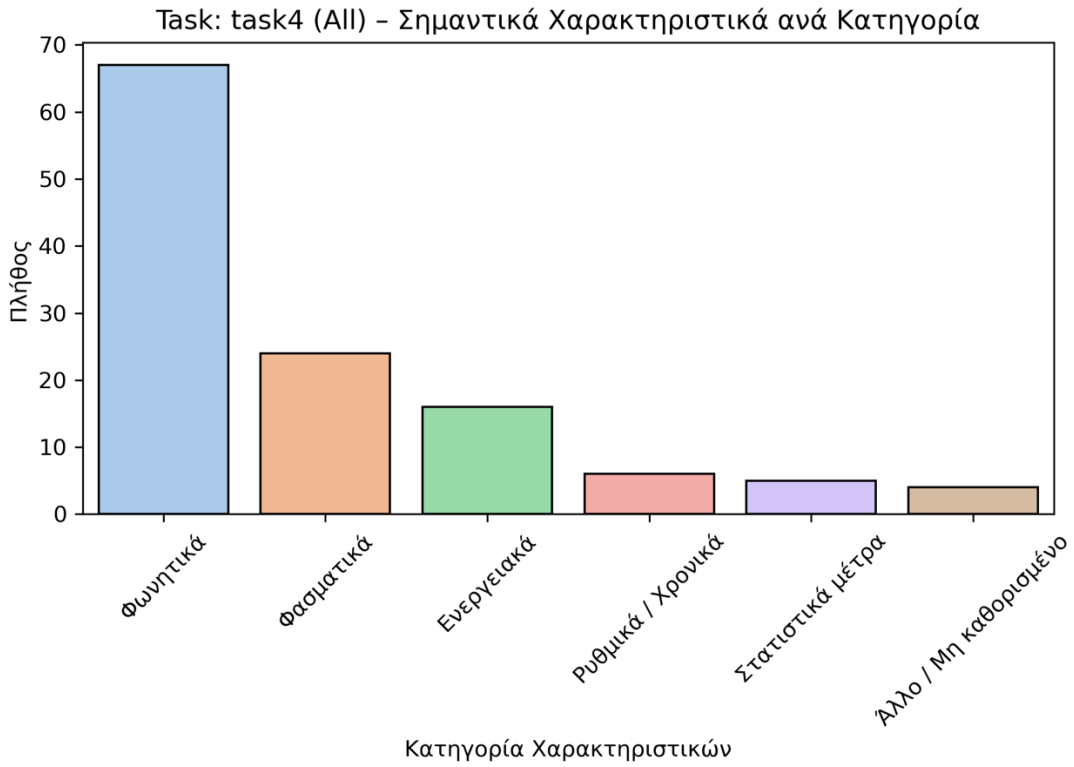
Εικόνα 72:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση rainbow_passage.wav του Task 3

Task 4 (A4): Φωνητικές Ασκήσεις – Συντονισμός Αρθρώσεων

Το τέταρτο σετ ασκήσεων προλαμβάνει επαναληπτική εκφορά συλλάβων προκειμένου να εκτιμηθεί ο ρυθμός της ομιλίας και η ευχέρεια της λειτουργίας των μυϊκών ομάδων που συνεργάζονται για την δημιουργία της φωνής. Αυτοί οι δύο παράμετροι είναι πολύ σημαντικοί για την εξέταση της ΝΠ. Η αξιολόγηση των παραμέτρων παρουσιάζονται παρακάτω, αποτυπώνοντας τα χαρακτηριστικά με την υψηλότερη στατιστική σημασία.

Πίνακας 53:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για το σύνολο του Task 4 (A4)

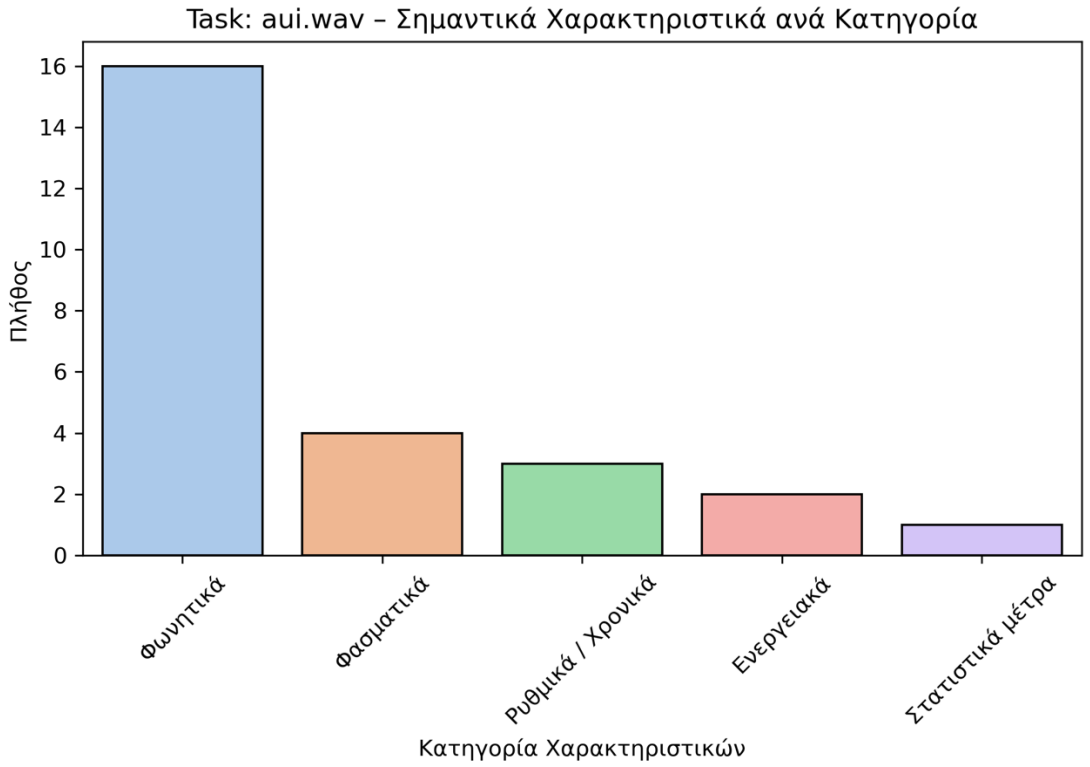
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F1amplitudeLogRelF0_sma3nz_amean	3.15×10^{-9}	Φωνητικά
shimmerLocaldB_sma3nz_amean	4.72×10^{-8}	Φωνητικά
F3frequency_sma3nz_amean	1.13×10^{-6}	Φωνητικά
loudness_sma3_amean	1.80×10^{-6}	Ενεργειακά
loudness_sma3_percentile80.0	8.52×10^{-6}	Ενεργειακά



Εικόνα 73:σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για το σύνολο του Task 4

Πίνακας 54:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση aui.wav του Task 4 (A4)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F1amplitudeLogRelF0_sma3nz_stddevNorm	0.000020	Φωνητικά
jitterLocal_sma3nz_stddevNorm	0.004483	Φωνητικά
F0semitoneFrom27.5Hz_sma3nz_percentile50.0	0.007406	Φωνητικά
F0semitoneFrom27.5Hz_sma3nz_amean	0.009050	Φωνητικά
F2bandwidth_sma3nz_stddevNorm	0.011871	Φωνητικά



Εικόνα 74:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση aui.wav του Task 4

Πίνακας 55:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση ka.wav του Task 4 (A4)

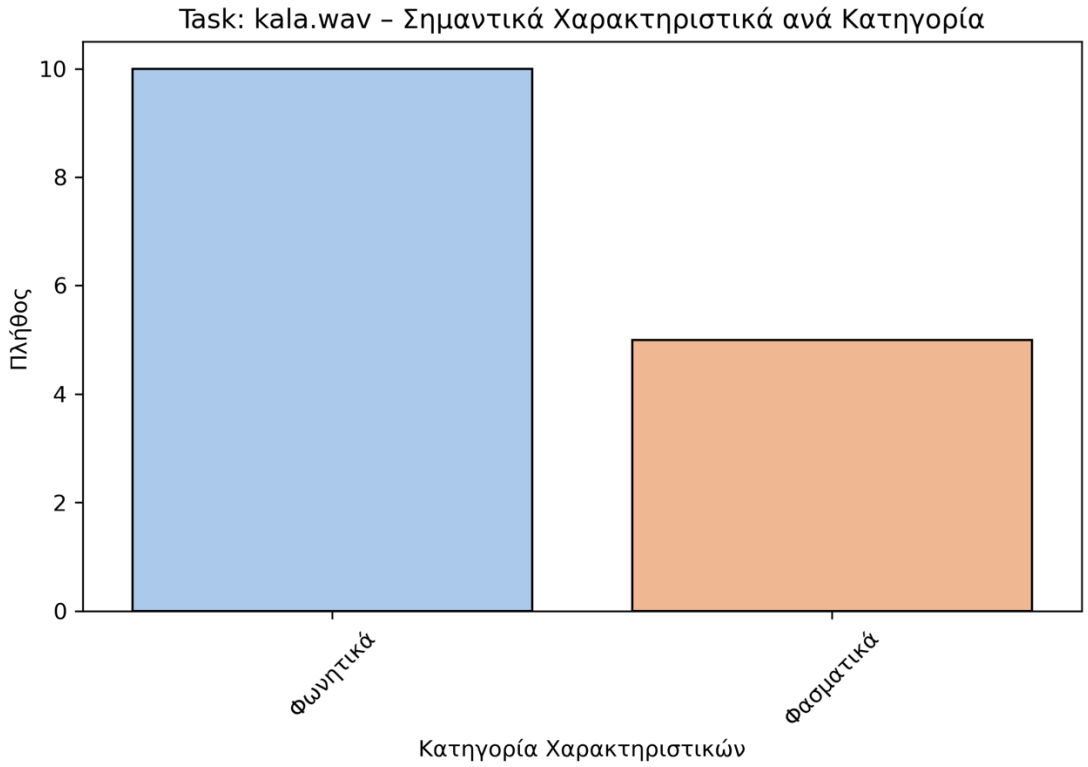
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc2V_sma3nz_amean	0.000241	Φασματικά
F0semitoneFrom27.5Hz_sma3nz_meanRisingSlope	0.000482	Φωνητικά
mfcc2V_sma3nz_amean	0.000510	Φασματικά
mfcc2_sma3_amean	0.000590	Φασματικά
loudness_sma3_percentile20.0	0.001471	Ενεργειακά



Εικόνα 75:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση ka.wav του Task 4

Πίνακας 56:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση kala.wav του Task 4 (A4)

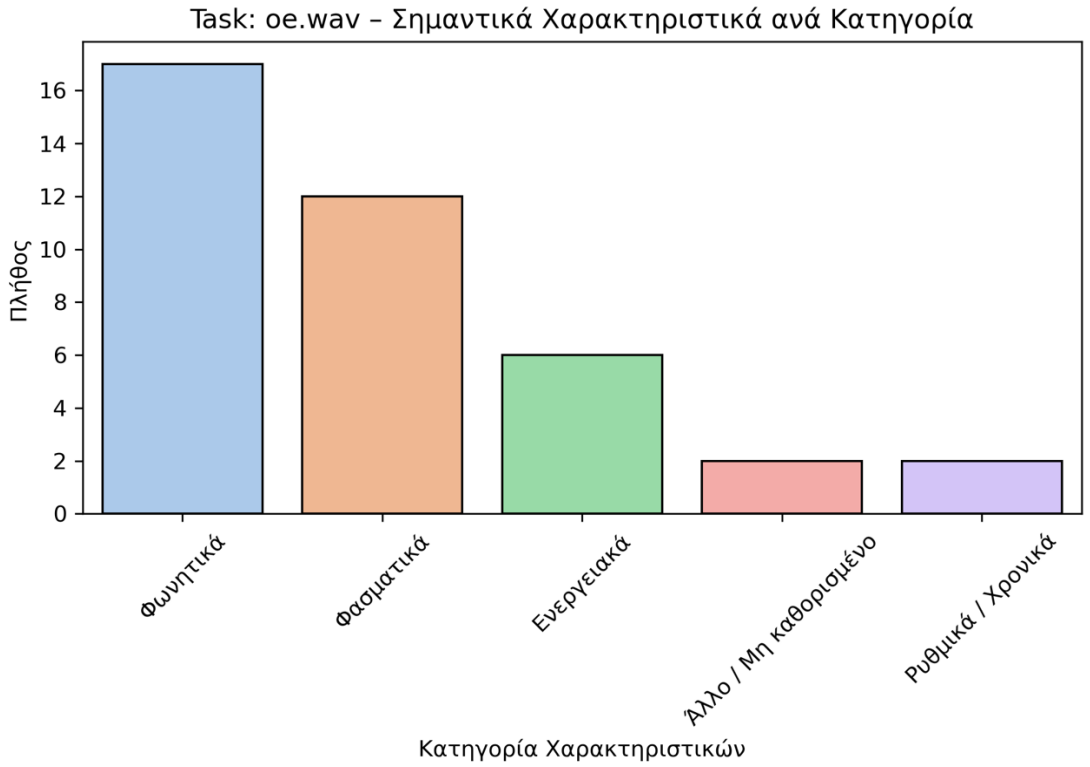
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F0semitoneFrom27.5Hz_sma3nz_percentile50.0	0.003947	Φωνητικά
spectralFlux_sma3_stddevNorm	0.007487	Φασματικά
F3frequency_sma3nz_amean	0.008146	Φωνητικά
F2frequency_sma3nz_amean	0.008783	Φωνητικά
mfcc2_sma3_stddevNorm	0.009897	Φασματικά



Εικόνα 76:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση kala.wav του Task 4

Πίνακας 57:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση oe.wav του Task 4 (A4)

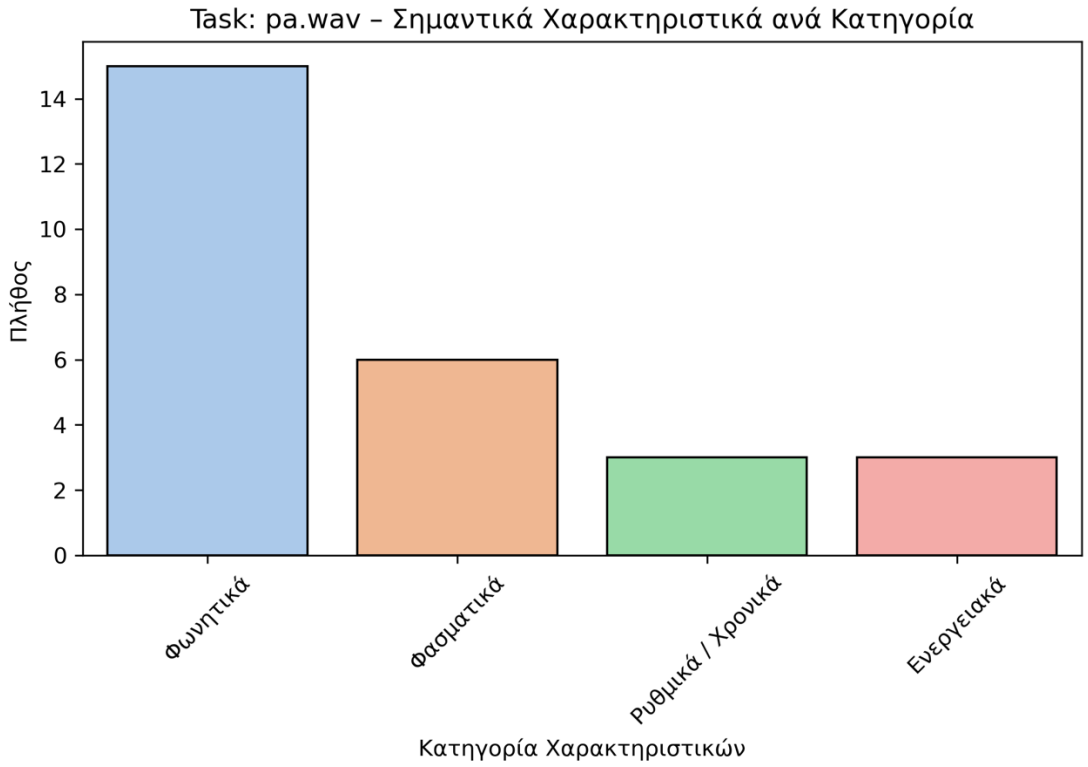
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F1frequency_sma3nz_amean	0.005649	Φωνητικά
mfcc3V_sma3nz_amean	0.006464	Φασματικά
F2bandwidth_sma3nz_amean	0.006573	Φωνητικά
StddevUnvoicedSegmentLength	0.010601	Ρυθμικά / Χρονικά
loudness_sma3_stddevNorm	0.012743	Ενεργειακά



Εικόνα 77:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση oe.wav του Task 4

Πίνακας 58:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση pa.wav του Task 4 (A4)

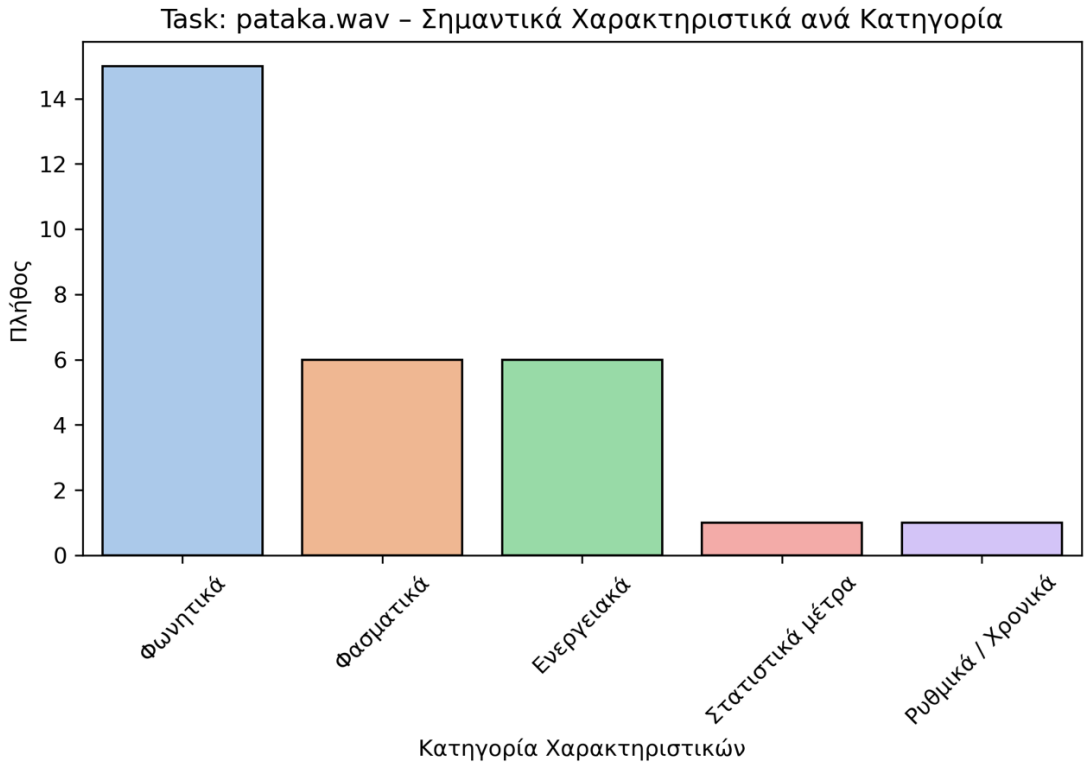
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
MeanVoicedSegmentLengthSec	0.003947	Ρυθμικά / Χρονικά
F0semitoneFrom27.5Hz_sma3nz_percentile80.0	0.007487	Φωνητικά
mfcc1V_sma3nz_amean	0.008146	Φασματικά
mfcc1_sma3_amean	0.008783	Φασματικά
shimmerLocaldB_sma3nz_stddevNorm	0.009897	Φωνητικά



Εικόνα 78:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση pa.wav του Task 4

Πίνακας 59:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση pataka.wav του Task 4 (A4)

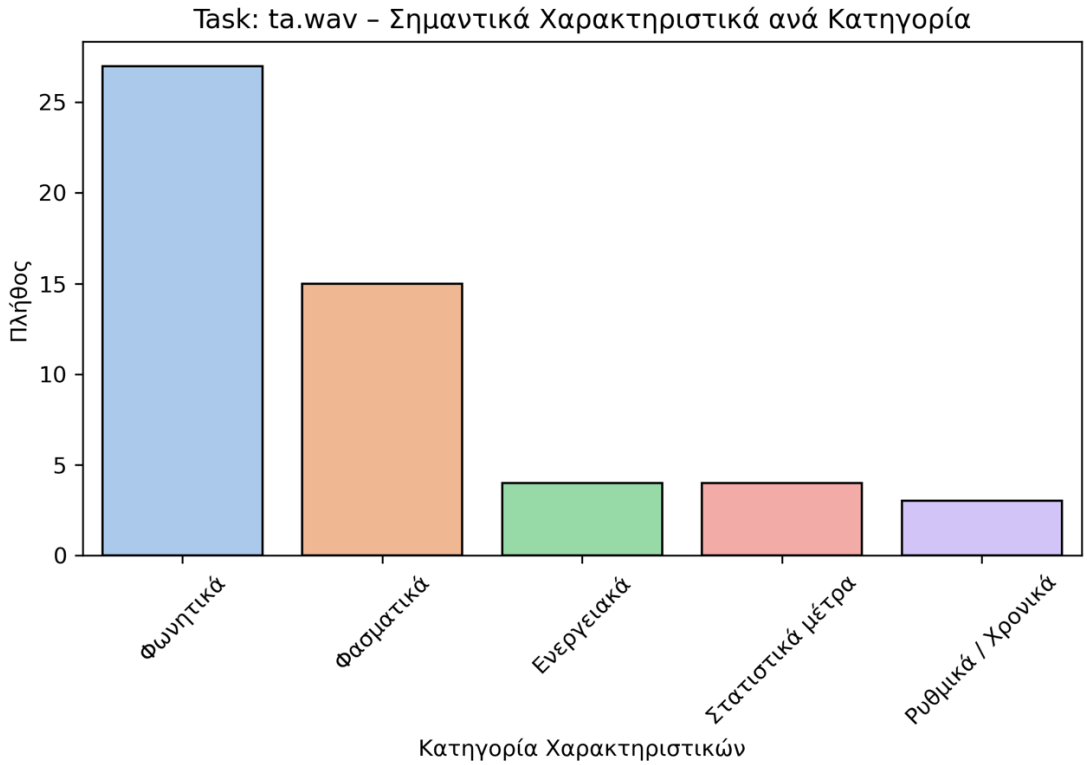
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
shimmerLocaldB_sma3nz_stddevNorm	0.001471	Φωνητικά
HNRdBACF_sma3nz_amean	0.003574	Φωνητικά
F1frequency_sma3nz_amean	0.003800	Φωνητικά
shimmerLocaldB_sma3nz_amean	0.005010	Φωνητικά
loudness_sma3_stddevFallingSlope	0.011488	Ενεργειακά



Εικόνα 79:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση pataka.wav του Task 4

Πίνακας 60:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση ta.wav του Task 4 (A4)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
spectralFlux_sma3_stddevNorm	0.000371	Φασματικά
spectralFluxV_sma3nz_stddevNorm	0.000371	Φασματικά
F3amplitudeLogRelF0_sma3nz_amean	0.000371	Φωνητικά
F1amplitudeLogRelF0_sma3nz_amean	0.000371	Φωνητικά
F0semitoneFrom27.5Hz_sma3nz_stddevRisingSlope	0.000371	Φωνητικά



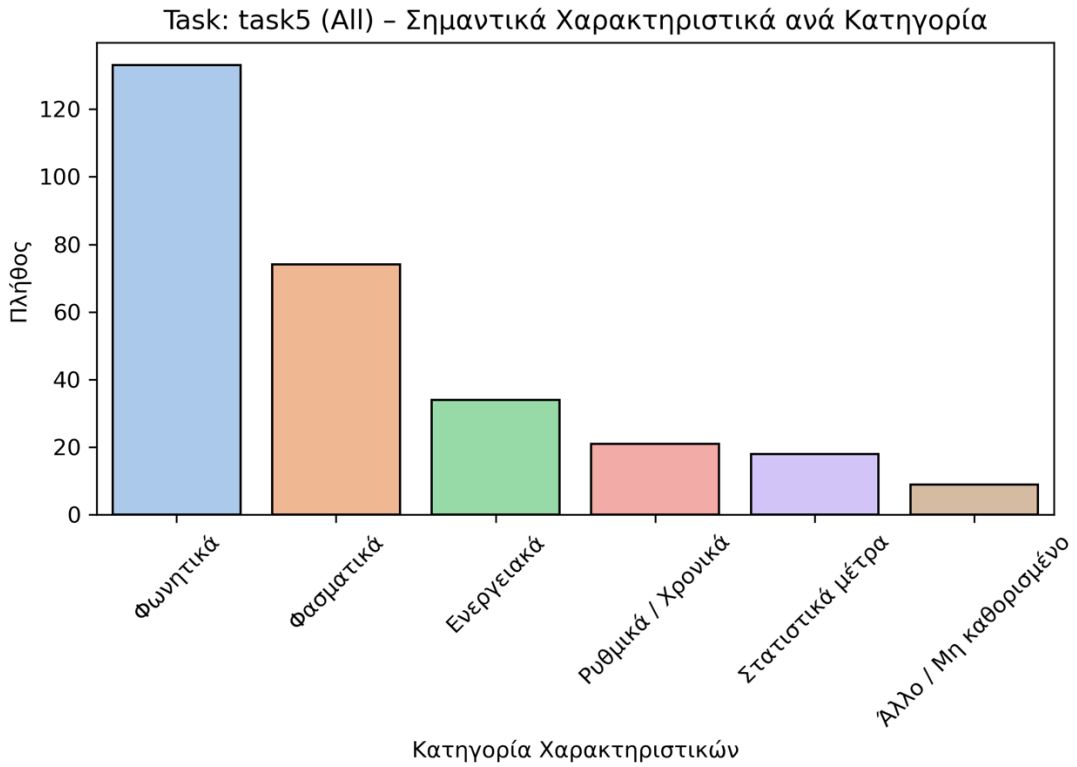
Εικόνα 80:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση ta.wav του Task 4

Task 5 (A5): Λεκτικό Ερέθισμα– Εστίαση σε Ομάδες Φωνηέντων

Το σύνολο των παρακάτω ασκήσεων βασίζεται στην εκφορά συγκεκριμένων λέξεων, που εμπεριέχουν στοχευμένα φωνήεντα, με κύριο στόχο την εξαγωγή χαρακτηριστικών μέσω των οποίων θα γίνει πιο κατανοητή η εκτίμηση των φωνητικών στοιχείων και πως αυτά επηρεάζονται από την παρουσία της νόσου, καθώς και την βαρύτητά αυτής στον κάθε ασθενή. Η αναλυτική παρουσίαση των πιο σημαντικών χαρακτηριστικών για αυτό το σετ ασκήσεων, παρατίθεται παρακάτω.

Πίνακας 61:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για το σύνολο του Task 5 (A5)

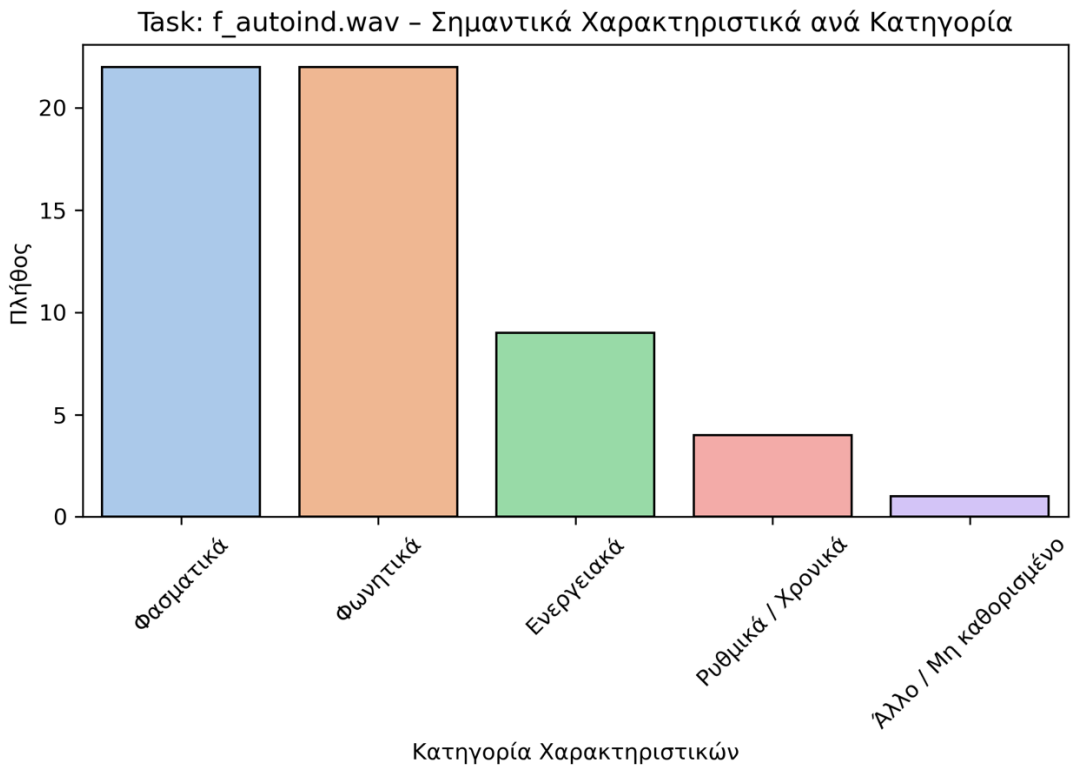
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2bandwidth_sma3nz_amean	6.93×10^{-25}	Φωνητικά
F1amplitudeLogRelF0_sma3nz_amean	2.83×10^{-23}	Φωνητικά
loudness_sma3_stddevNorm	3.98×10^{-23}	Ενεργειακά
F1bandwidth_sma3nz_amean	9.64×10^{-20}	Φωνητικά
F3frequency_sma3nz_amean	1.12×10^{-17}	Φωνητικά



Εικόνα 81:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για το σύνολο του Task 5

Πίνακας 62:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_autoind.wav του Task 5 (A5)

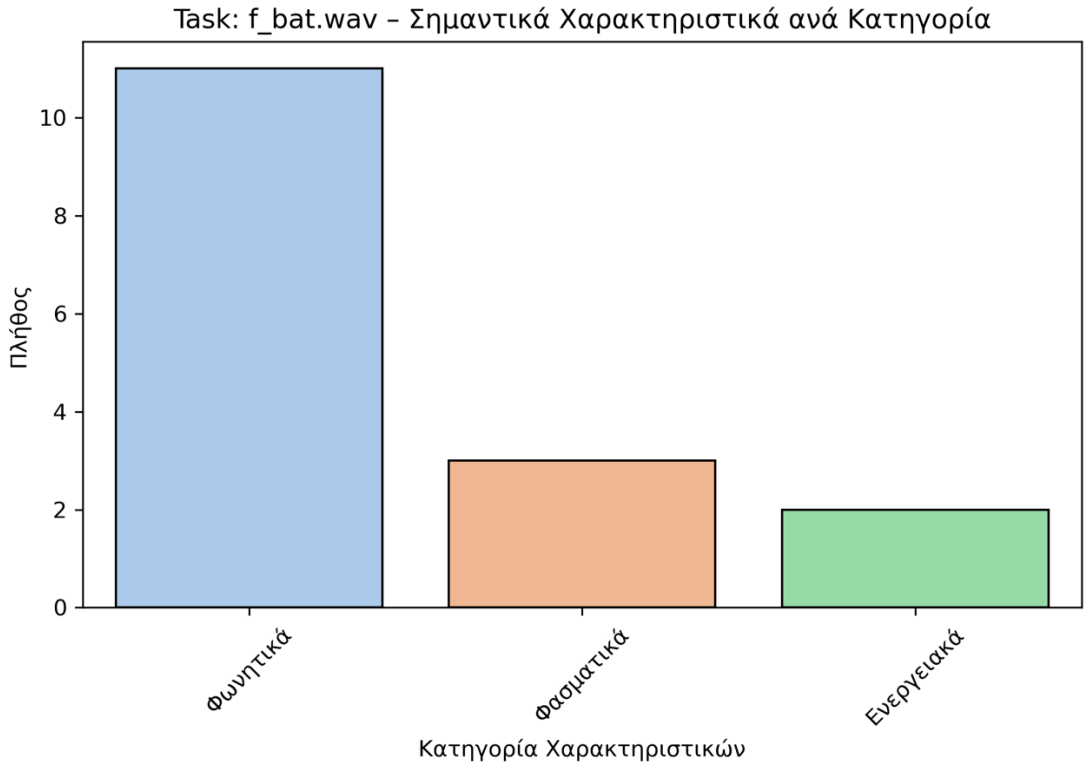
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2bandwidth_sma3nz_amean	6.93×10^{-25}	Φωνητικά
F1amplitudeLogRelF0_sma3nz_amean	2.83×10^{-23}	Φωνητικά
loudness_sma3_stddevNorm	3.98×10^{-23}	Ενεργειακά
F1bandwidth_sma3nz_amean	9.64×10^{-20}	Φωνητικά
F3frequency_sma3nz_amean	1.12×10^{-17}	Φωνητικά



Εικόνα 82:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_autoind.wav του Task 5

Πίνακας 63:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_bat.wav του Task 5 (A5)

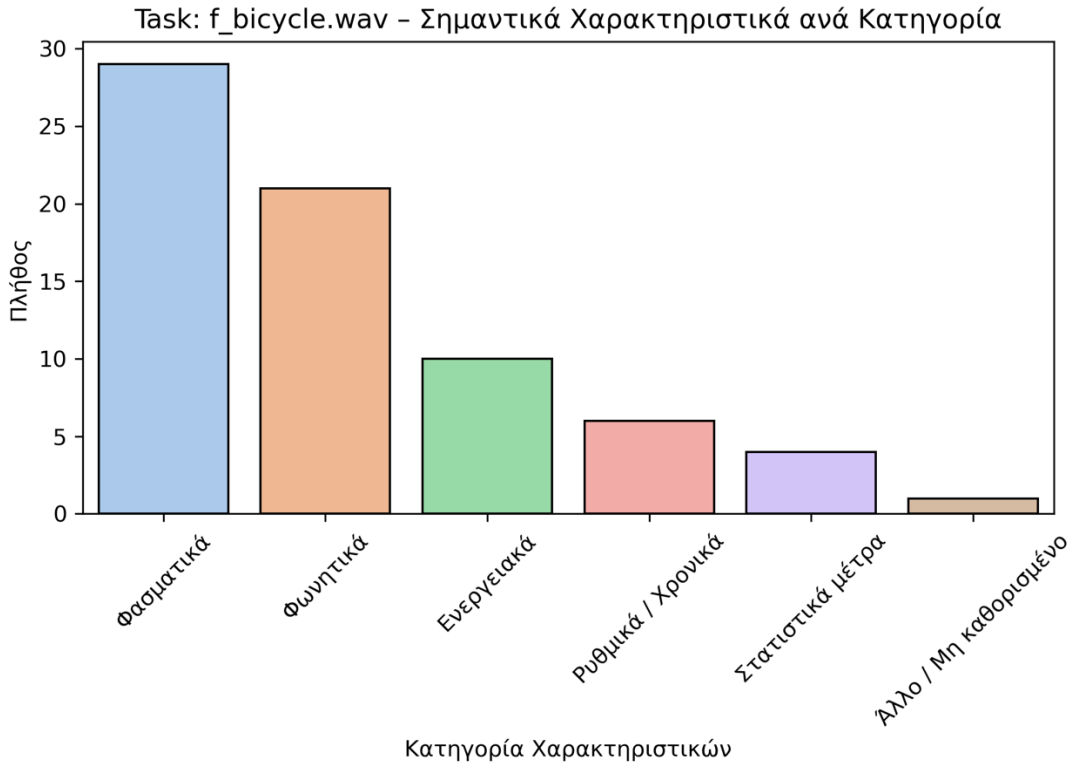
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F1amplitudeLogRelF0_sma3nz_amean	0.000206	Φωνητικά
F3amplitudeLogRelF0_sma3nz_stddevNorm	0.000890	Φωνητικά
F2amplitudeLogRelF0_sma3nz_stddevNorm	0.002359	Φωνητικά
F2amplitudeLogRelF0_sma3nz_amean	0.003419	Φωνητικά
shimmerLocaldB_sma3nz_amean	0.006950	Φωνητικά



Εικόνα 83:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_bat.wav του Task 5

Πίνακας 64:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_bicycle.wav του Task 5 (A5)

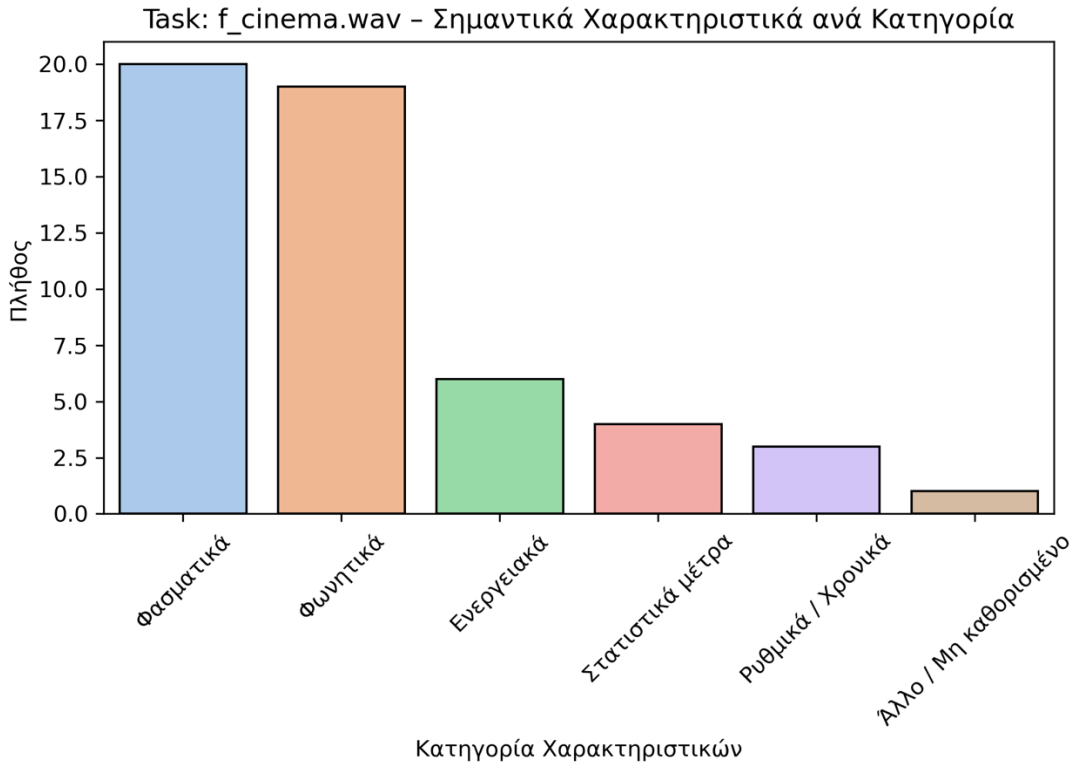
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
spectralFlux_sma3_stddevNorm	0.000852	Φασματικά
F3frequency_sma3nz_amean	0.001483	Φωνητικά
mfcc1V_sma3nz_amean	0.001557	Φασματικά
mfcc1_sma3_amean	0.001678	Φασματικά
mfcc2_sma3_amean	0.002059	Φασματικά



Εικόνα 84:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_bicycle.wav του Task 5

Πίνακας 65:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_cinema.wav του Task 5 (A5)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F0semitoneFrom27.5Hz_sma3nz_stddevNorm	0.001171	Φασματικά
jitterLocal_sma3nz_amean	0.001951	Φωνητικά
mfcc2V_sma3nz_amean	0.002359	Φασματικά
F3bandwidth_sma3nz_stddevNorm	0.002524	Φωνητικά
VoicedSegmentsPerSec	0.003195	Ρυθμικά / Χρονικά



Εικόνα 85:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_cinema.wav του Task 5

Πίνακας 66:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_cockroachkiller.wav του Task 5 (A5)

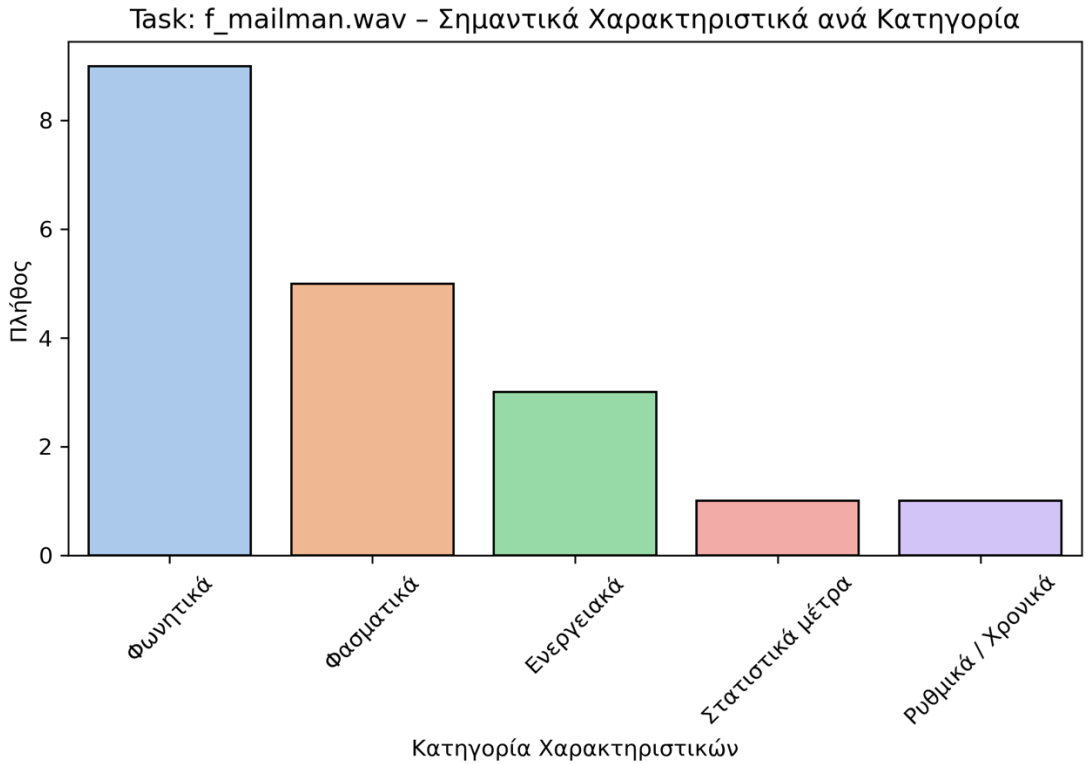
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
jitterLocal_sma3nz_amean	0.003020	Φωνητικά
F1frequency_sma3nz_amean	0.005464	Φωνητικά
HNRdBACF_sma3nz_amean	0.006293	Φωνητικά
mfcc4_sma3_stddevNorm	0.006847	Φασματικά
hammarbergIndexUV_sma3nz_amean	0.010837	Στατιστικά μέτρα



Εικόνα 86:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_cockroachkiller.wav του Task 5

Πίνακας 67:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_mailman.wav του Task 5 (A5)

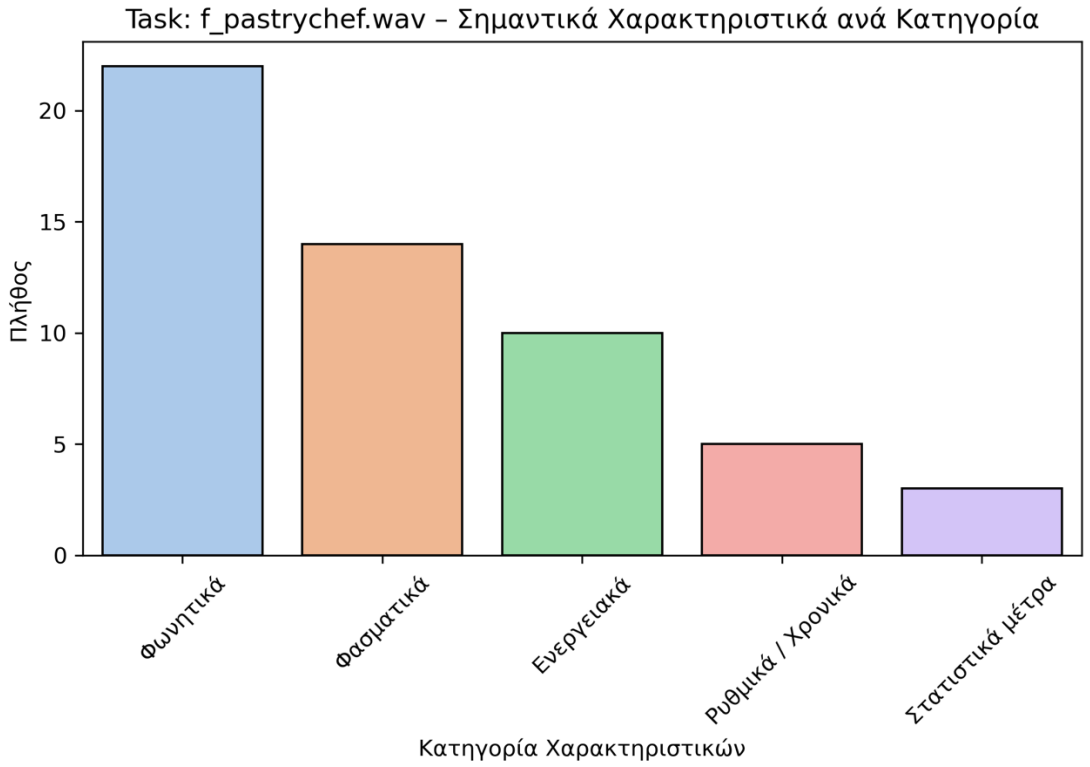
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2bandwidth_sma3nz_stddevNorm	0.013569	Φωνητικά
loudness_sma3_stddevFallingSlope	0.014318	Ενεργειακά
spectralFluxUV_sma3nz_amean	0.016378	Φασματικά
spectralFluxUV_sma3nz_amean	0.016836	Φασματικά
alphaRatioUV_sma3nz_amean	0.017905	Φασματικά



Εικόνα 87:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_mailman.wav του Task 5

Πίνακας 68:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_pastrychef.wav του Task 5 (A5)

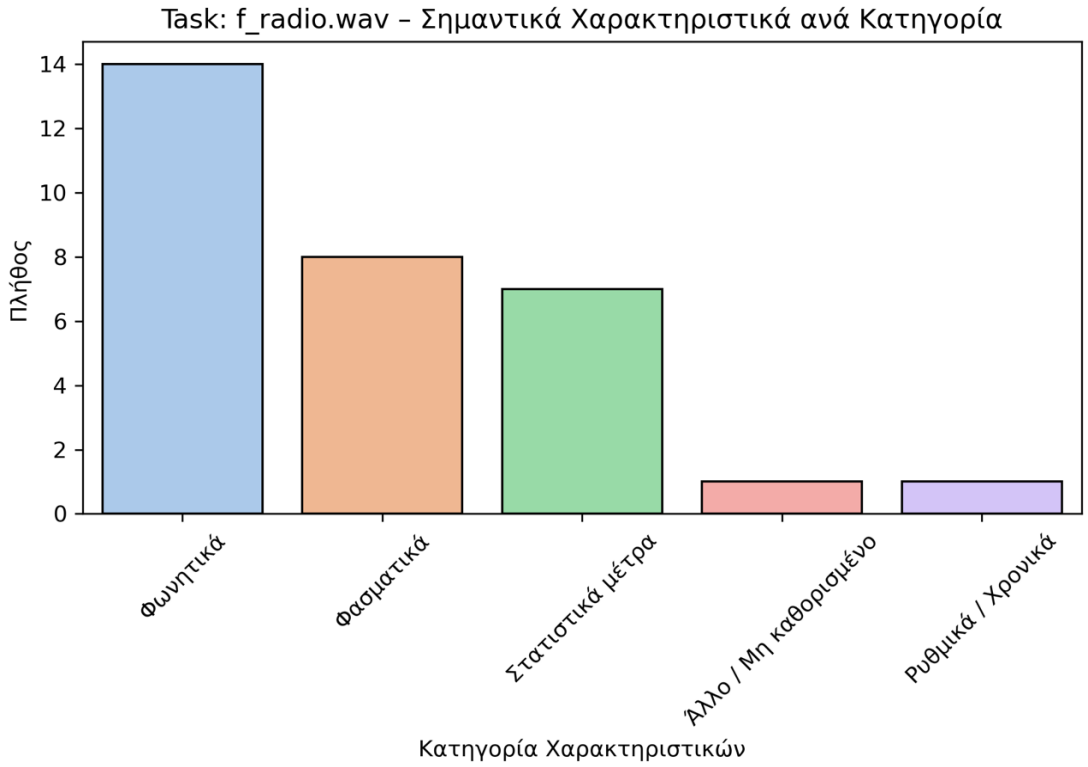
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
loudness_sma3_stddevRisingSlope	0.000173	Ενεργειακά
VoicedSegmentsPerSec	0.000557	Ρυθμικά / Χρονικά
MeanVoicedSegmentLengthSec	0.000739	Ρυθμικά / Χρονικά
spectralFluxV_sma3nz_stddevNorm	0.001196	Φασματικά
F0semitoneFrom27.5Hz_sma3nz_stddevRisingSlope	0.003914	Φωνητικά



Εικόνα 88:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_pastrychef.wav του Task 5

Πίνακας 69:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_radio.wav του Task 5 (A5)

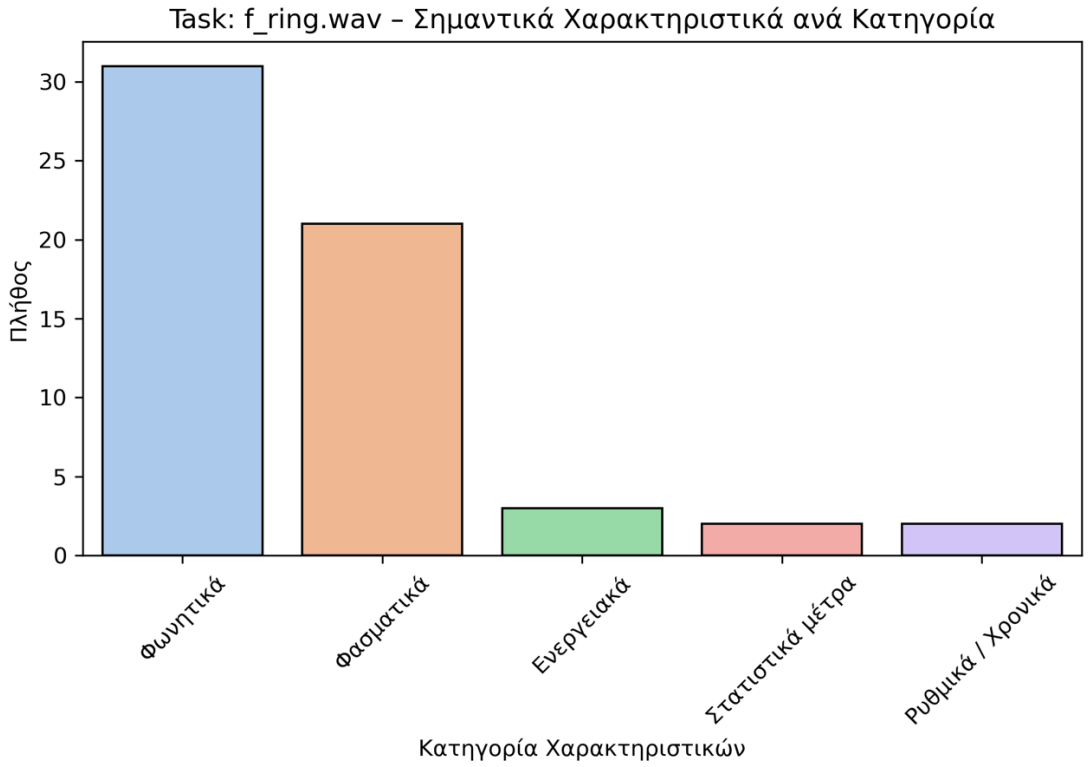
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc2V_sma3nz_stddevNorm	0.000067	Φασματικά
HNRdBACF_sma3nz_stddevNorm	0.000695	Φωνητικά
mfcc2V_sma3nz_stddevNorm	0.001204	Φασματικά
alphaRatioUV_sma3nz_amean	0.002982	Φασματικά
F3amplitudeLogRelF0_sma3nz_amean	0.003012	Φωνητικά



Εικόνα 89:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_radio.wav του Task 5

Πίνακας 70:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_ring.wav του Task 5 (A5)

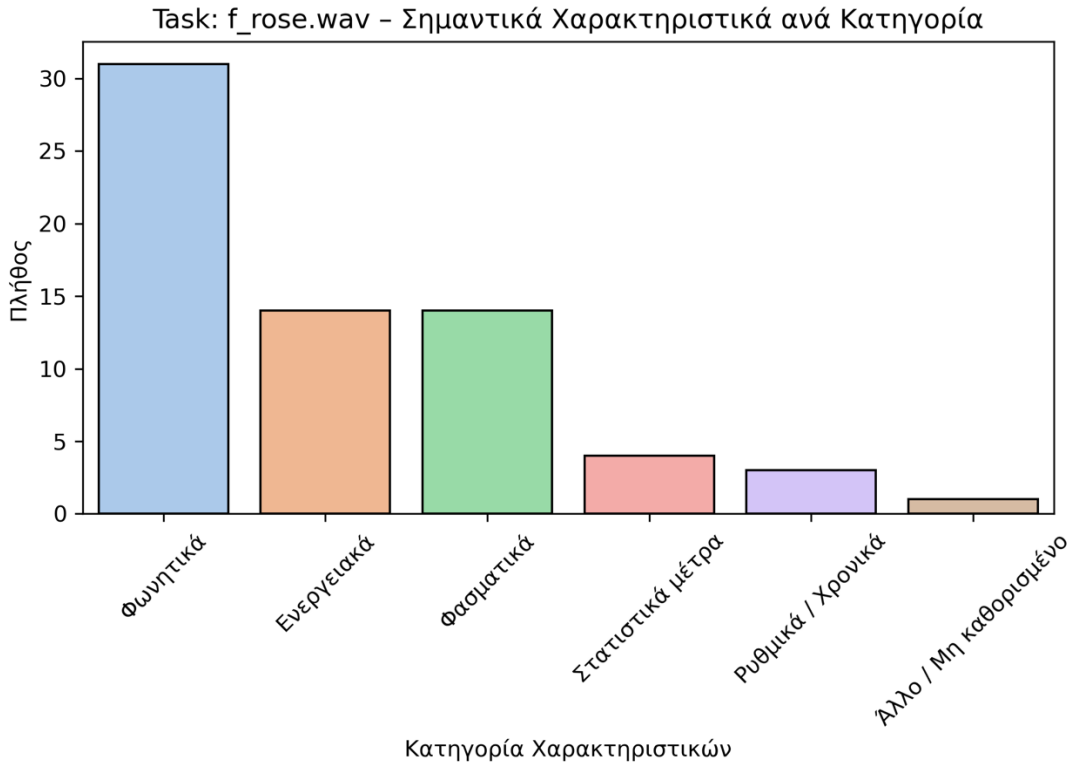
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc1V_sma3nz_stddevNorm	0.001226	Φασματικά
mfcc2_sma3_amean	0.001297	Φασματικά
shimmerLocaldB_sma3nz_stddevNorm	0.001308	Φωνητικά
F2frequency_sma3nz_stddevNorm	0.001429	Φωνητικά
shimmerLocaldB_sma3nz_amean	0.001460	Φωνητικά



Εικόνα 90:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_ring.wav του Task 5

Πίνακας 71:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_rose.wav του Task 5 (A5)

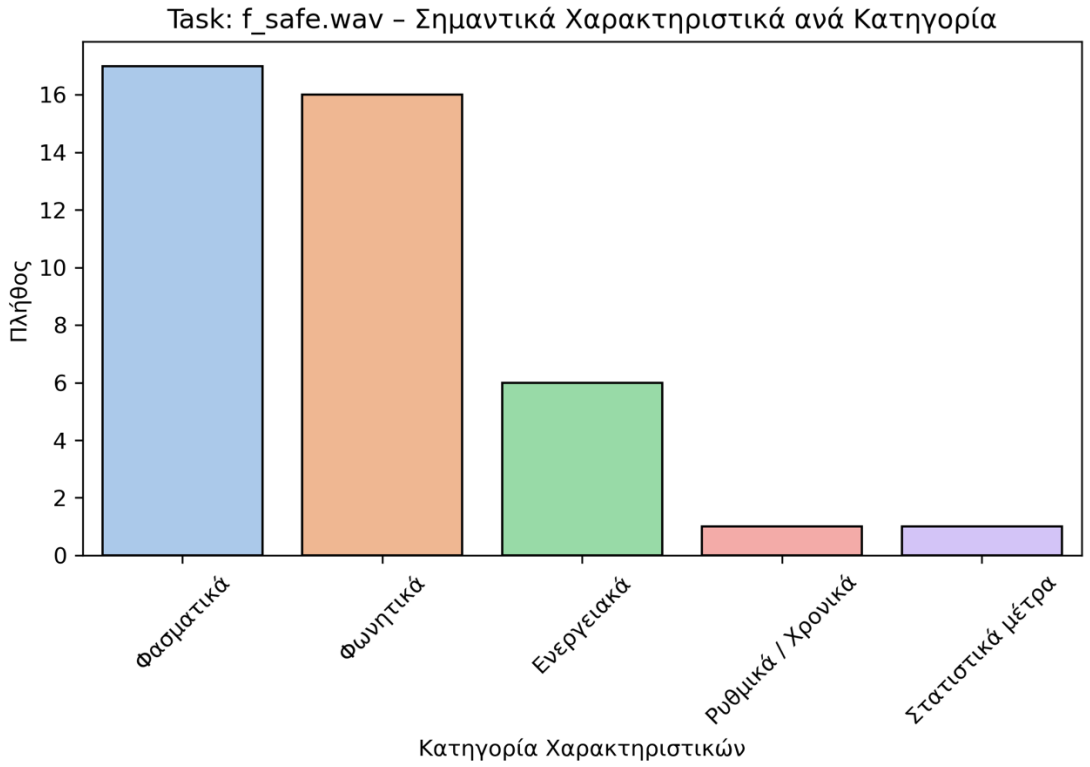
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F0semitoneFrom27.5Hz_sma3nz_meanFallingSlope	0.000607	Φωνητικά
F0semitoneFrom27.5Hz_sma3nz_stddevFallingSlope	0.003498	Φωνητικά
F2frequency_sma3nz_stddevNorm	0.003920	Φωνητικά
loudness_sma3_stddevNorm	0.004057	Ενεργειακά
loudness_sma3_percentile20.0	0.004629	Ενεργειακά



Εικόνα 91:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_rose.wav του Task 5

Πίνακας 72:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_safe.wav του Task 5 (A5)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
loudness_sma3_amean	0.001558	Ενεργειακά
F0semitoneFrom27.5Hz_sma3nz_stddevNorm	0.002980	Φωνητικά
mfcc3_sma3_amean	0.004106	Φασματικά
mfcc3V_sma3nz_amean	0.005278	Φασματικά
F3bandwidth_sma3nz_amean	0.005470	Φωνητικά



Εικόνα 92:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_safe.wav του Task 5

Πίνακας 73:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_snail.wav του Task 5 (A5)

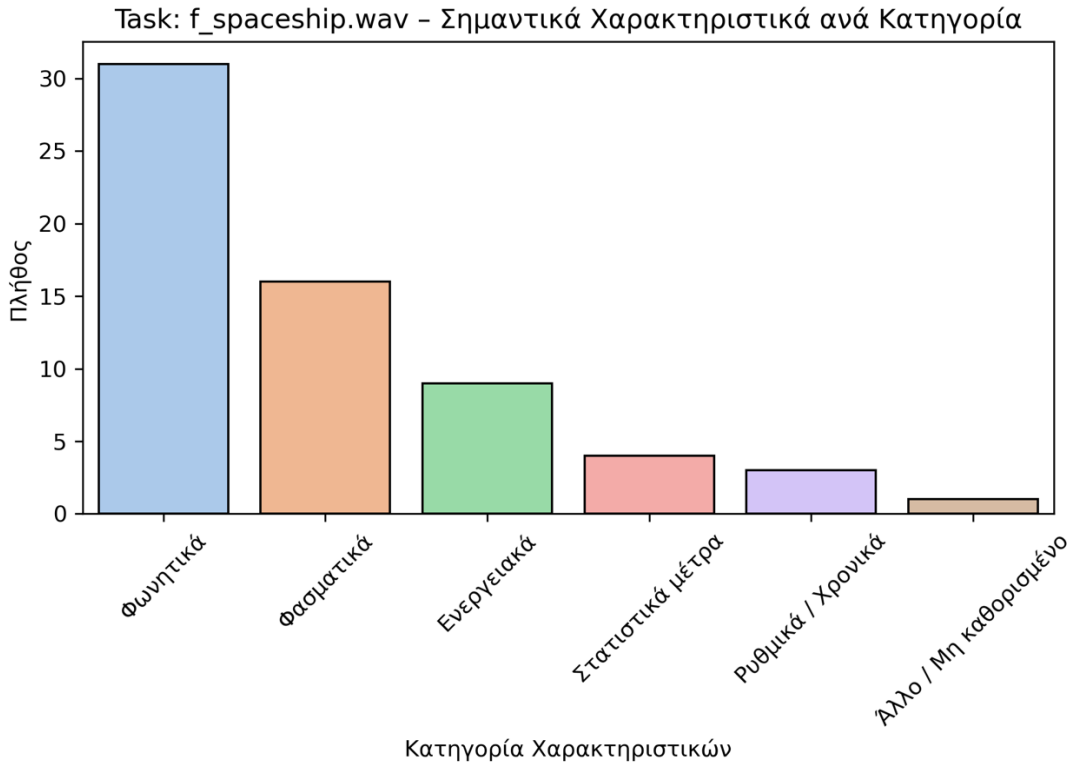
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F1frequency_sma3nz_amean	0.002297	Φωνητικά
F1amplitudeLogRelF0_sma3nz_amean	0.003445	Φωνητικά
mfcc4_sma3_stddevNorm	0.004347	Φασματικά
mfcc2_sma3_amean	0.004368	Φασματικά
loudness_sma3_percentile80.0	0.006818	Ενεργειακά



Εικόνα 93:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_snail.wav του Task 5

Πίνακας 74:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_spaceship.wav του Task 5 (A5)

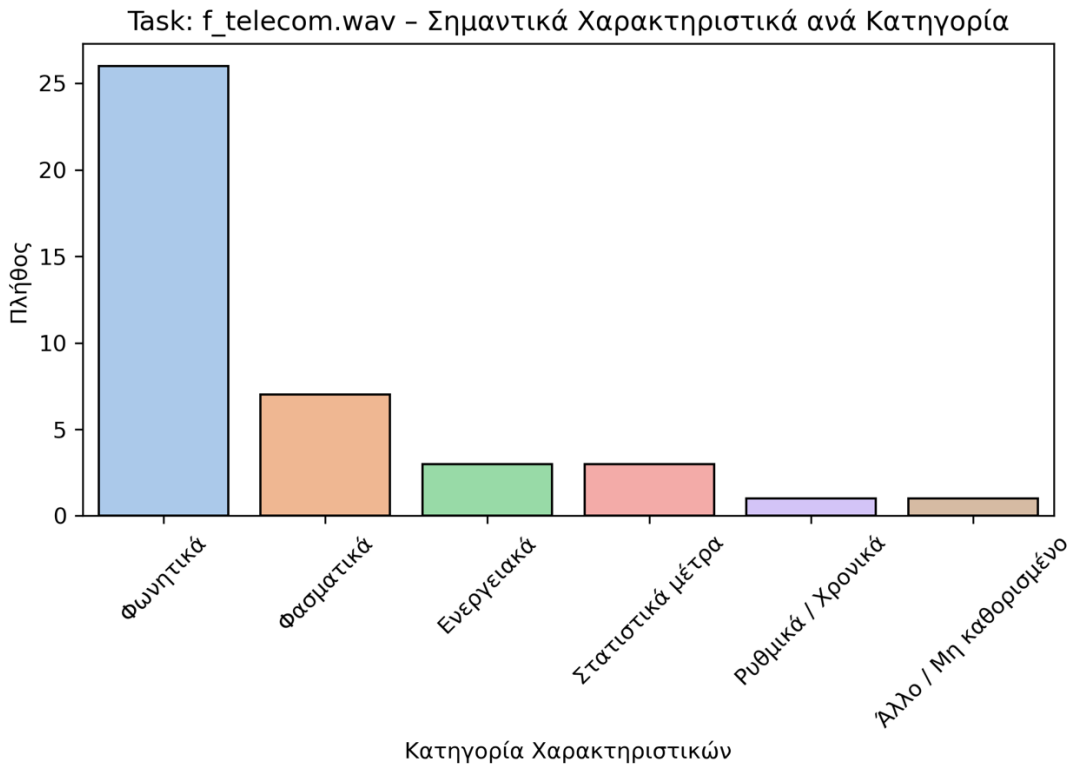
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc4_sma3_amean	0.000539	Φασματικά
HNRdBACF_sma3nz_stddevNorm	0.000618	Φωνητικά
mfcc4V_sma3nz_amean	0.000714	Φασματικά
MeanVoicedSegmentLengthSec	0.002464	Ρυθμικά / Χρονικά
F3amplitudeLogRelF0_sma3nz_amean	0.004072	Φωνητικά



Εικόνα 94:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_spaceship.wav του Task 5

Πίνακας 75:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση f_telecom.wav του Task 5 (A5)

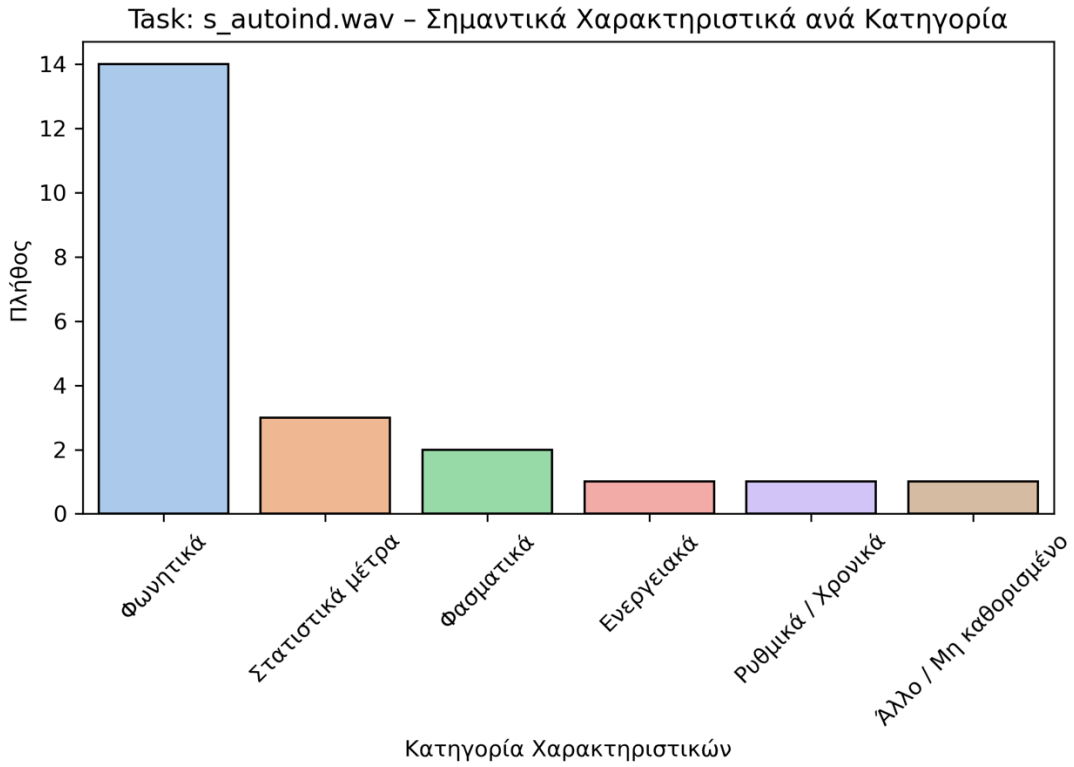
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
HNRdBACF_sma3nz_amean	0.000508	Φωνητικά
F3frequency_sma3nz_stddevNorm	0.000721	Φωνητικά
F0semitoneFrom27.5Hz_sma3nz_stddevNorm	0.000729	Φωνητικά
F1amplitudeLogRelF0_sma3nz_amean	0.001056	Φωνητικά
F2semitoneFrom27.5Hz_sma3nz_stddevNorm	0.001273	Φωνητικά



Εικόνα 95:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση f_telecom.wav του Task 5

Πίνακας 76:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_autoind.wav του Task 5 (A5)

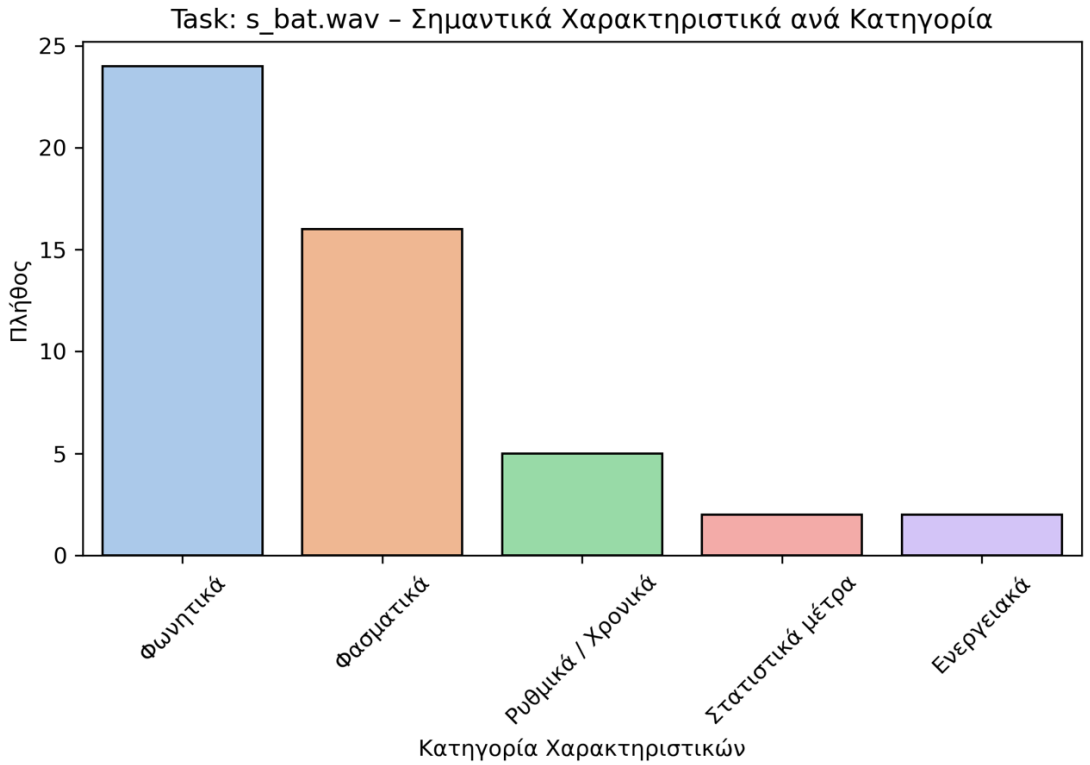
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F3frequency_sma3nz_amean	0.000585	Φωνητικά
F2bandwidth_sma3nz_amean	0.000891	Φωνητικά
F1frequency_sma3nz_amean	0.001141	Φωνητικά
mfcc4V_sma3nz_stddevNorm	0.005987	Φασματικά
hammarbergIndexV_sma3nz_amean	0.007724	Στατιστικά μέτρα



Εικόνα 96:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_autoind.wav του Task 5

Πίνακας 77:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_bat.wav του Task 5 (A5)

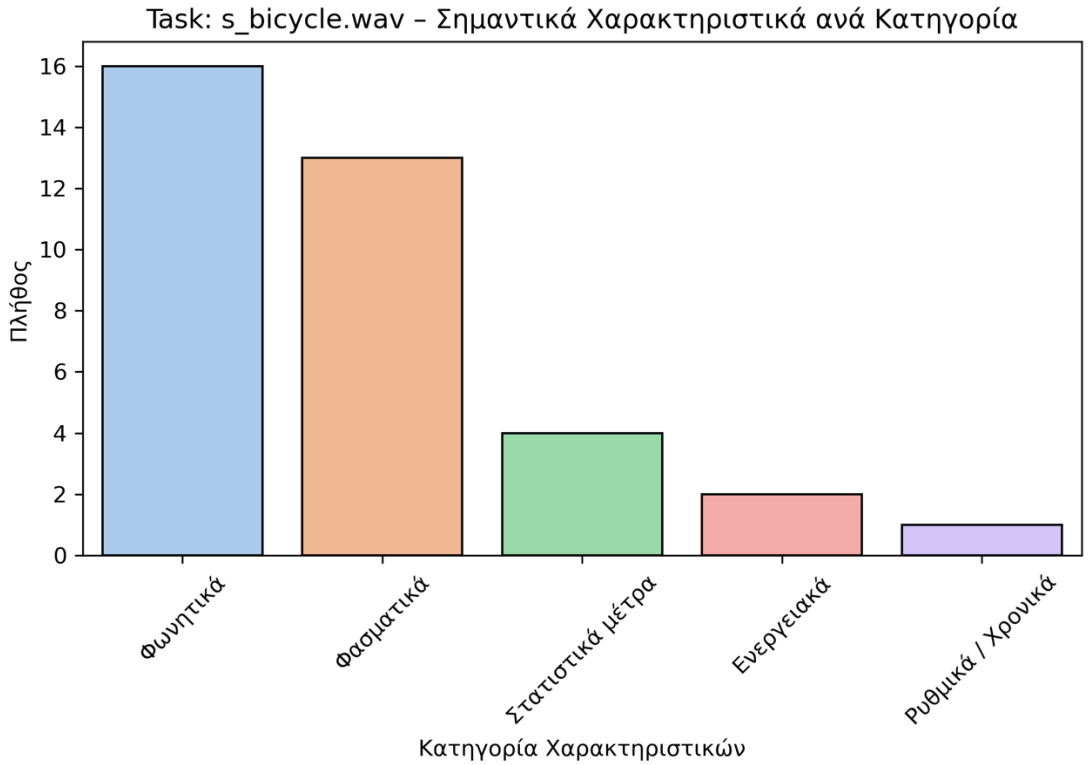
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F3frequency_sma3nz_stddevNorm	0.002246	Φωνητικά
mfcc3_sma3_stddevNorm	0.003054	Φασματικά
mfcc3V_sma3nz_stddevNorm	0.004388	Φασματικά
F0semitoneFrom27.5Hz_sma3nz_meanRisingSlope	0.004503	Φωνητικά
MeanUnvoicedSegmentLength	0.004597	Ρυθμικά / Χρονικά



Εικόνα 97:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_bat.wav του Task 5

Πίνακας 78:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_bicycle.wav του Task 5 (A5)

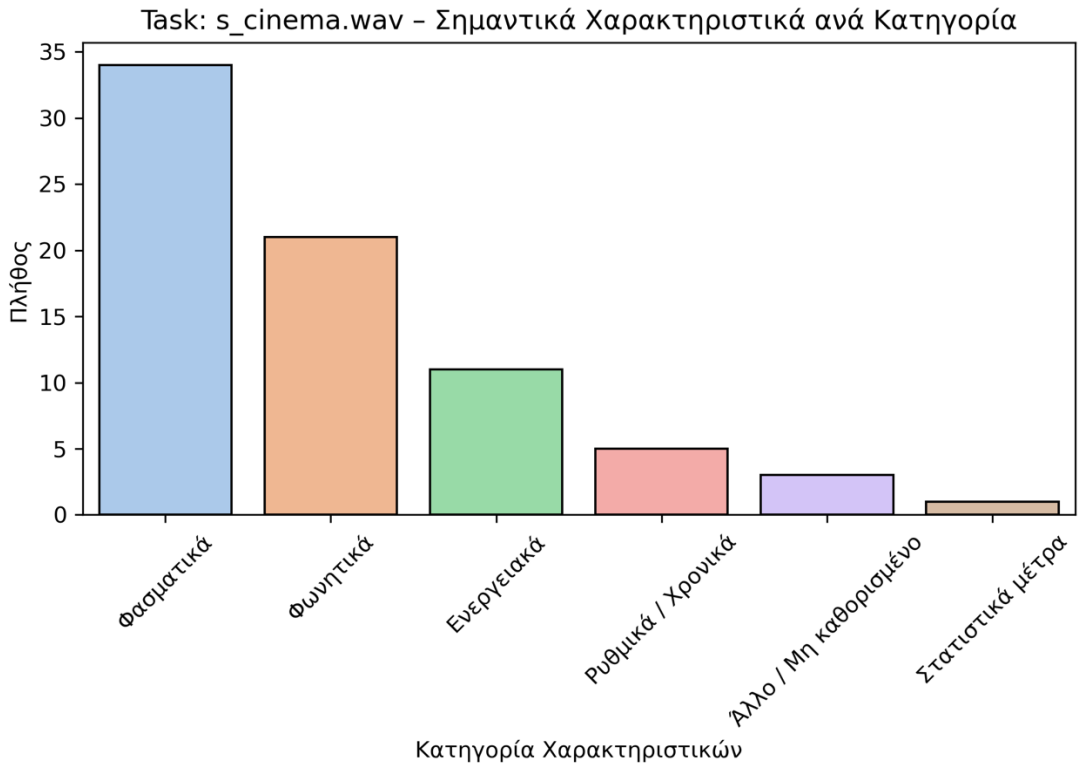
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F1frequency_sma3nz_stddevNorm	0.003479	Φωνητικά
mfcc2_sma3_amean	0.006804	Φασματικά
F0semitoneFrom27.5Hz_sma3nz_stddevFallingSlope	0.007236	Φωνητικά
F2bandwidth_sma3nz_amean	0.007741	Φωνητικά
F1semitoneFrom27.5Hz_sma3nz_meanFallingSlope	0.008476	Φωνητικά



Εικόνα 98:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_bicycle.wav του Task 5

Πίνακας 79:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_cinema.wav του Task 5 (A5)

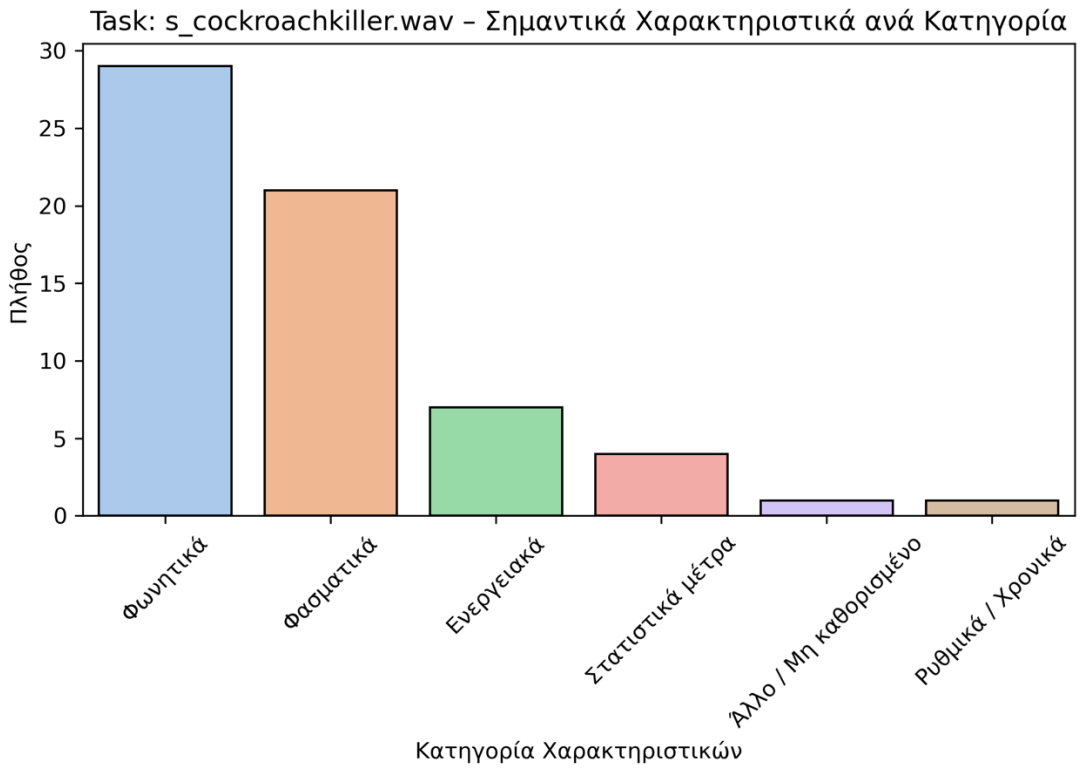
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F0semitoneFrom27.5Hz_sma3nz_meanRisingSlope	0.000653	Φωνητικά
(Intercept)	0.001516	Άλλο / Μη καθορισμένο
mfcc1_sma3_amean	0.001565	Φασματικά
F1semitoneFrom27.5Hz_sma3nz_stddevRisingSlope	0.001704	Φωνητικά
hammarbergIndexV_sma3nz_stddevNorm	0.002040	Στατιστικά μέτρα



Εικόνα 99:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_cinema.wav του Task 5

Πίνακας 80:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_cockroachkiller.wav του Task 5 (A5)

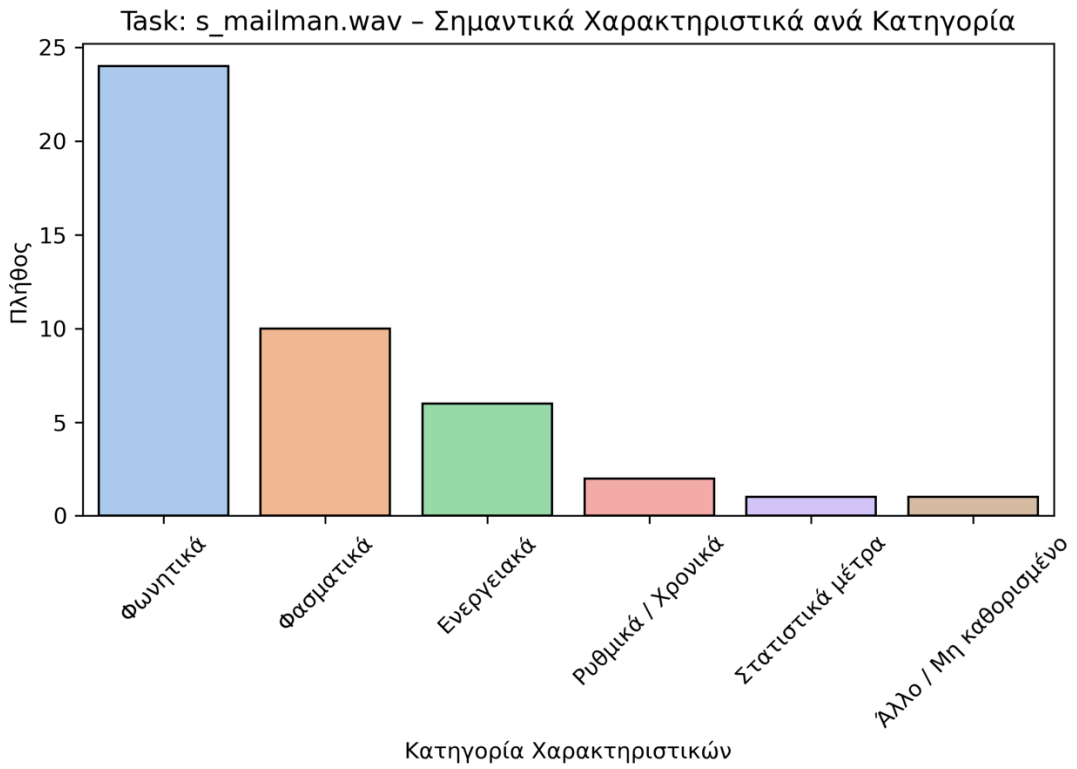
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F3frequency_sma3nz_amean	0.000587	Φωνητικά
F1amplitudeLogRelF0_sma3nz_stddevNorm	0.000707	Φωνητικά
mfcc3_sma3_amean	0.002158	Φασματικά
F1bandwidth_sma3nz_amean	0.004872	Στατιστικά μέτρα
HNRdBACF_sma3nz_stddevNorm	0.005536	Φωνητικά



Εικόνα 100:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_cockroachkiller.wav του Task 5

Πίνακας 81:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_mailman.wav του Task 5 (A5)

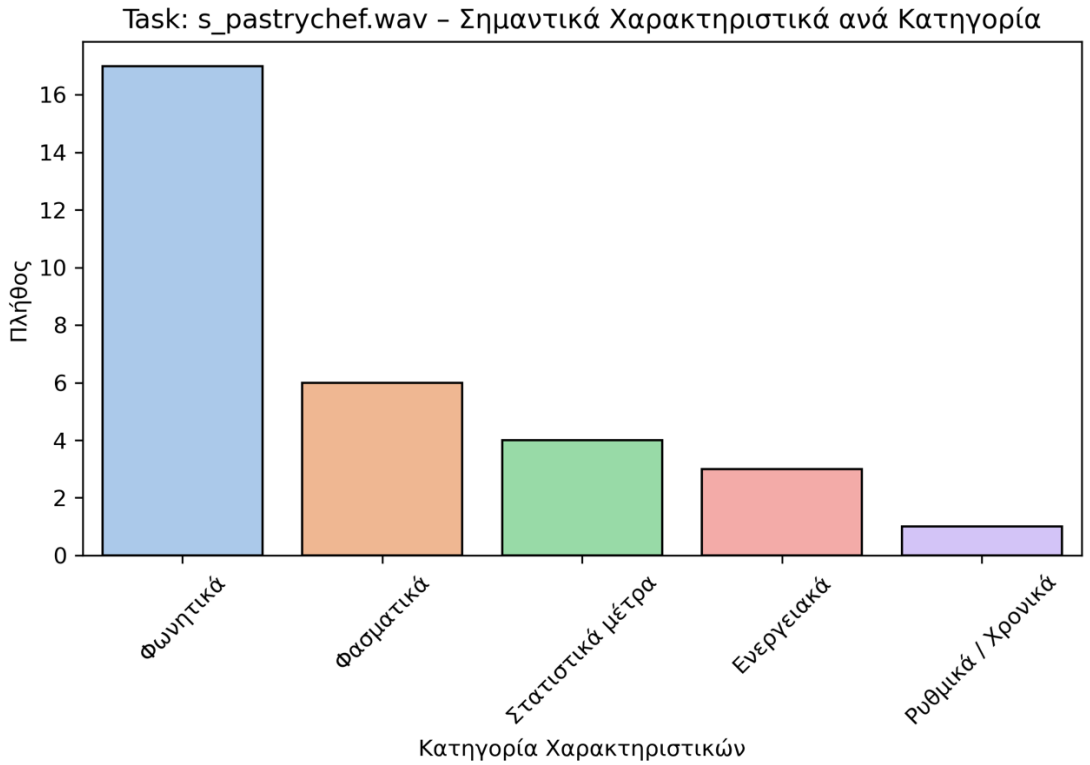
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F0semitoneFrom27.5Hz_sma3nz_meanFallingSlope	0.001073	Φωνητικά
F1semitoneFrom27.5Hz_sma3nz_meanFallingSlope	0.001190	Φωνητικά
F2semitoneFrom27.5Hz_sma3nz_meanFallingSlope	0.002100	Φωνητικά
F1amplitudeLogRelF0_sma3nz_stddevNorm	0.002379	Φωνητικά
F3amplitudeLogRelF0_sma3nz_stddevNorm	0.003091	Φωνητικά



Εικόνα 101:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_mailman.wav του Task 5

Πίνακας 82:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_pastrychef.wav του Task 5 (A5)

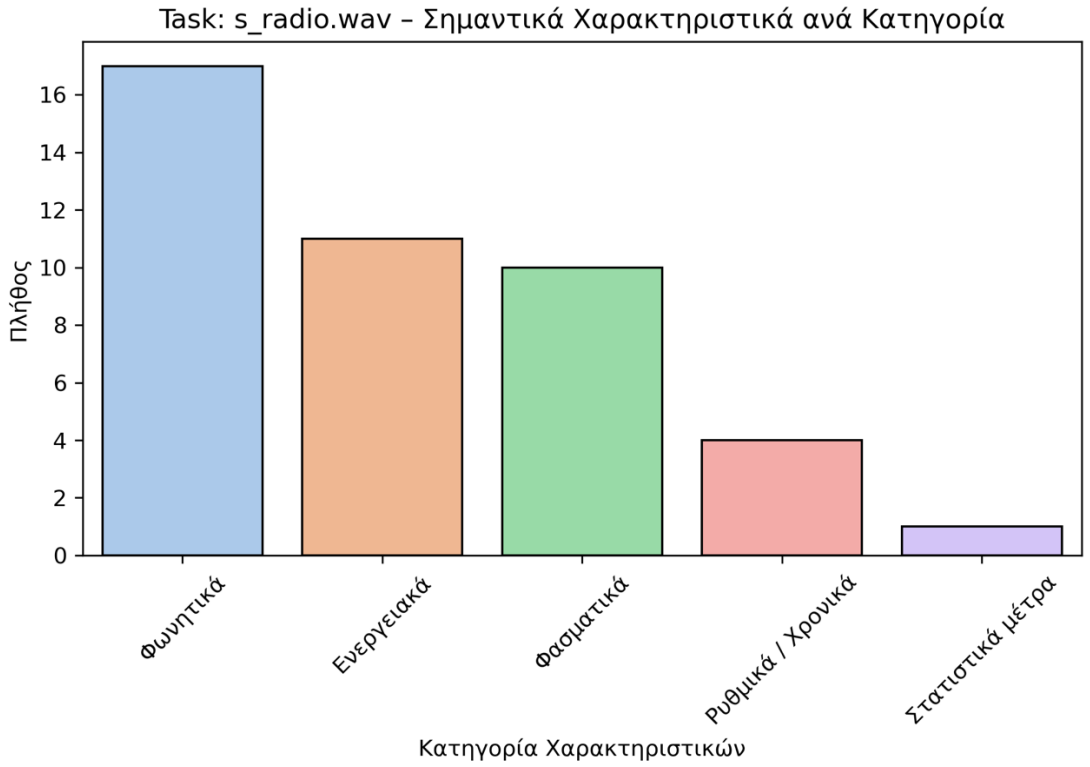
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2amplitudeLogRelF0_sma3nz_stddevNorm	0.000746	Φωνητικά
F2amplitudeLogRelF0_sma3nz_amean	0.003904	Φωνητικά
hammarbergIndexUV_sma3nz_amean	0.005472	Στατιστικά μέτρα
F1amplitudeLogRelF0_sma3nz_stddevNorm	0.007263	Φωνητικά
F3frequency_sma3nz_amean	0.008398	Φωνητικά



Εικόνα 102:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_pastrychef.wav του Task 5

Πίνακας 83:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_radio.wav του Task 5 (A5)

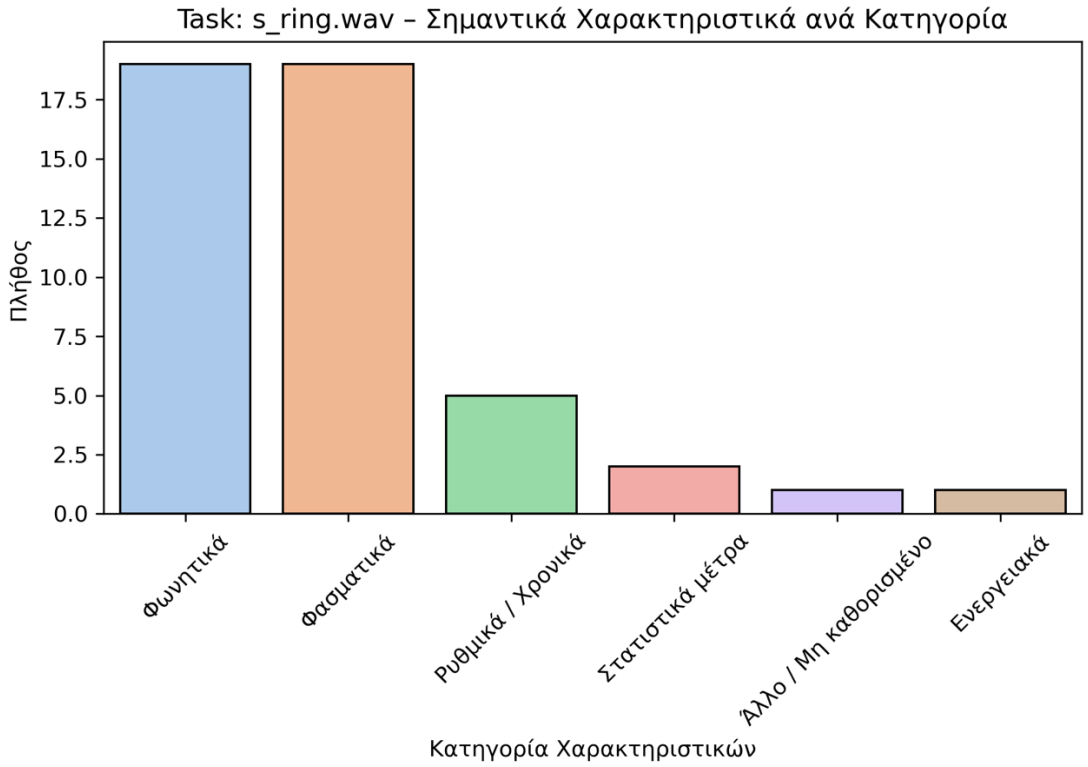
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
loudness_sma3_percentile20.0	0.001227	Ενεργειακά
loudness_sma3_percentile30.0	0.005389	Ενεργειακά
mfcc1V_sma3nz_stddevNorm	0.005425	Φασματικά
F3amplitudeLogRelF0_sma3nz_stddevNorm	0.005623	Φωνητικά
MeanVoicedSegmentLengthSec	0.010271	Ρυθμικά / Χρονικά



Εικόνα 103:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_radio.wav του Task 5

Πίνακας 84:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_ring.wav του Task 5 (A5)

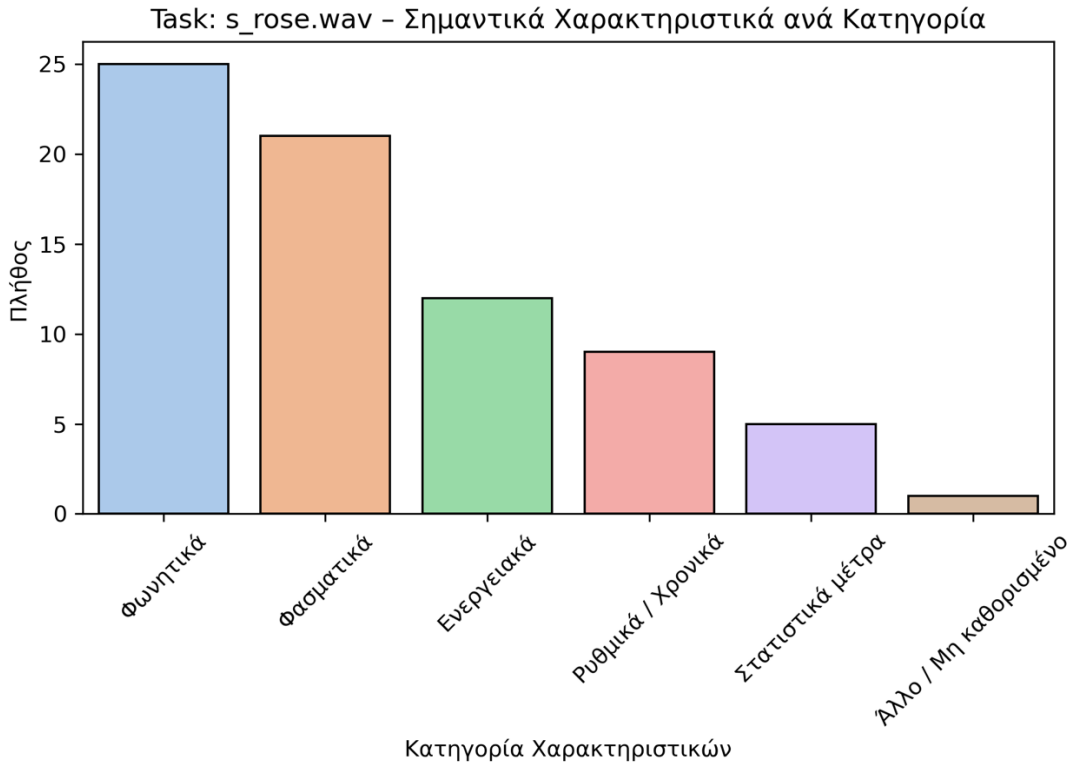
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F0semitoneFrom27.5Hz_sma3nz_percentile20.0	0.000093	Φωνητικά
F2bandwidth_sma3nz_amean	0.000235	Φωνητικά
mfcc2V_sma3nz_amean	0.002710	Φασματικά
jitterLocal_sma3nz_stddevNorm	0.003850	Φωνητικά
F3frequency_sma3nz_stddevNorm	0.005156	Φωνητικά



Εικόνα 104:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_ring.wav του Task 5

Πίνακας 85:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_rose.wav του Task 5 (A5)

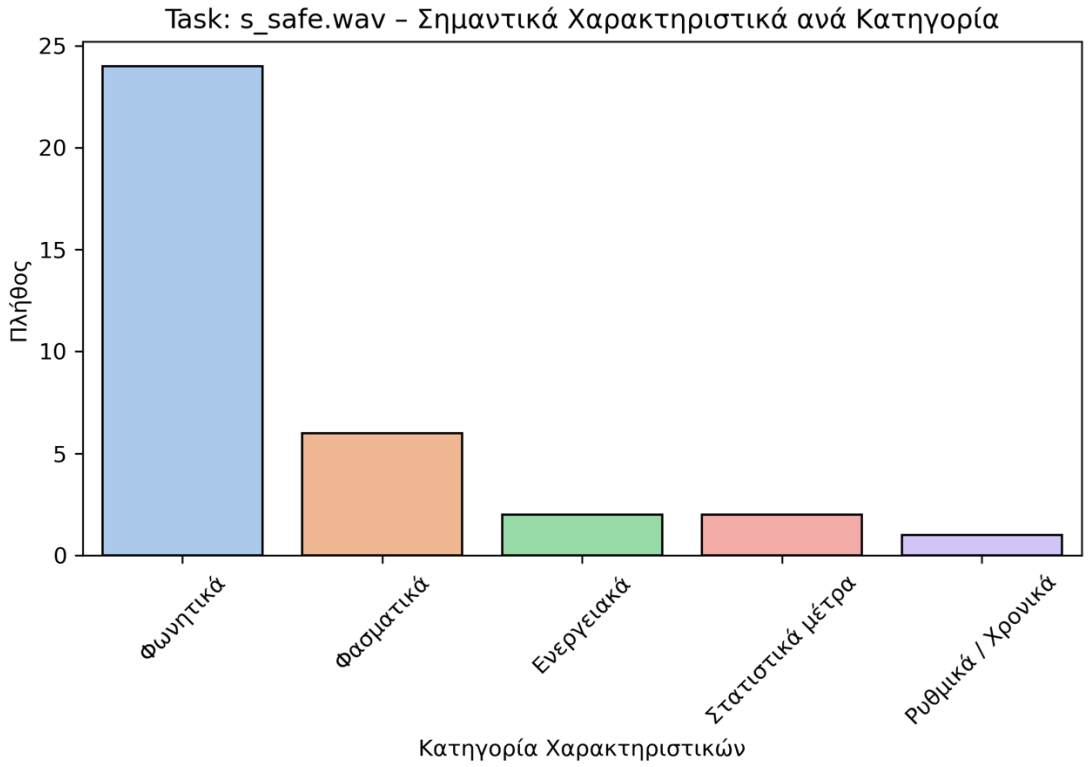
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F3frequency_sma3nz_stddevNorm	0.000112	Φωνητικά
F2bandwidth_sma3nz_amean	0.000256	Φωνητικά
MeanVoicedSegmentLengthSec	0.001270	Ρυθμικά / Χρονικά
mfcc1_sma3_stddevNorm	0.001340	Φασματικά
mfcc2V_sma3nz_amean	0.001393	Φασματικά



Εικόνα 105:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_rose.wav του Task 5

Πίνακας 86:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_safe.wav του Task 5 (A5)

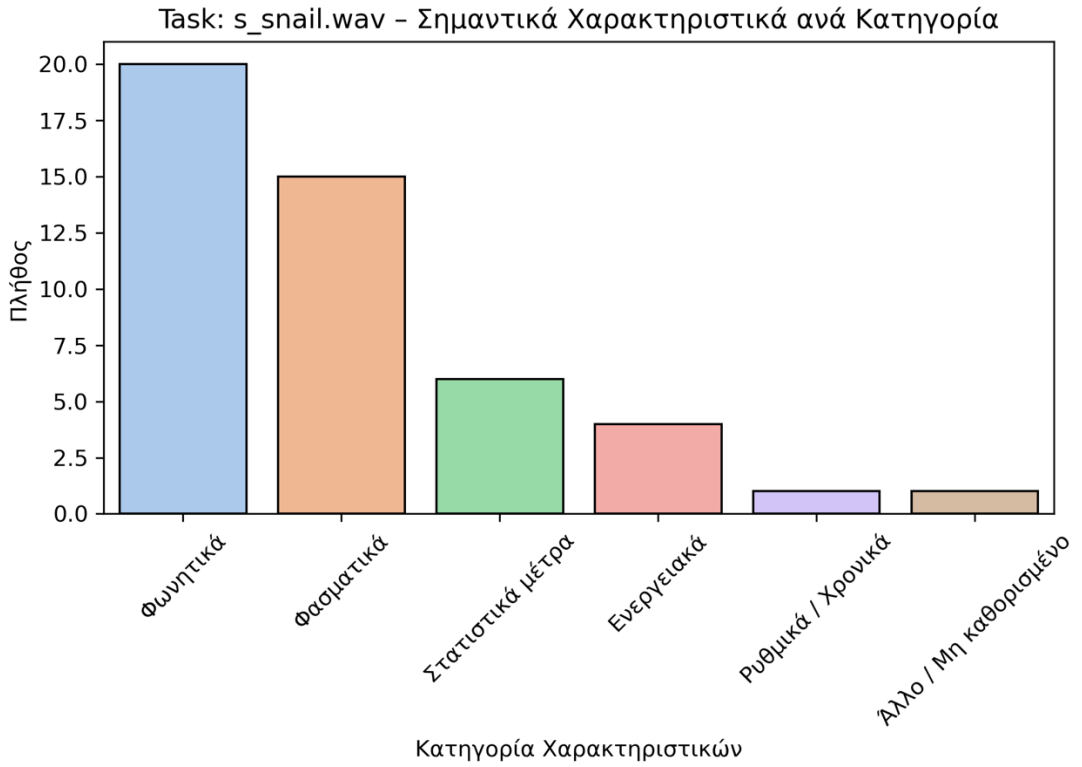
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2amplitudeLogRelF0_sma3nz_stddevNorm	0.000146	Φωνητικά
F2amplitudeLogRelF0_sma3nz_amean	0.000776	Φωνητικά
F1amplitudeLogRelF0_sma3nz_stddevNorm	0.000911	Φωνητικά
hammarbergIndexV_sma3nz_amean	0.001372	Στατιστικά μέτρα
F3amplitudeLogRelF0_sma3nz_stddevNorm	0.003348	Φωνητικά



Εικόνα 106:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_safe.wav του Task 5

Πίνακας 87:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_snail.wav του Task 5 (A5)

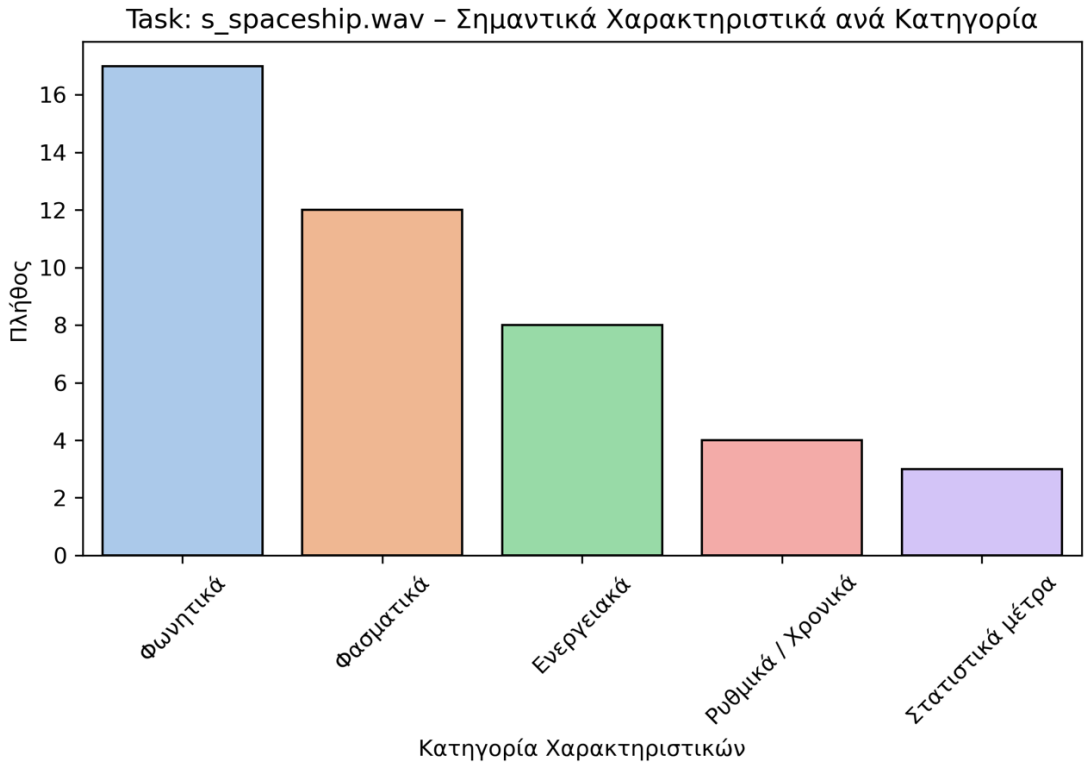
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F0semitoneFrom27.5Hz_sma3nz_stddevRisingSlope	0.001431	Φωνητικά
F1frequency_sma3nz_amean	0.001504	Φωνητικά
F2bandwidth_sma3nz_amean	0.001536	Φωνητικά
F1bandwidth_sma3nz_amean	0.002587	Στατιστικά μέτρα
F2frequency_sma3nz_amean	0.002591	Φωνητικά



Εικόνα 107:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_snail.wav του Task 5

Πίνακας 88:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_spaceship.wav του Task 5 (A5)

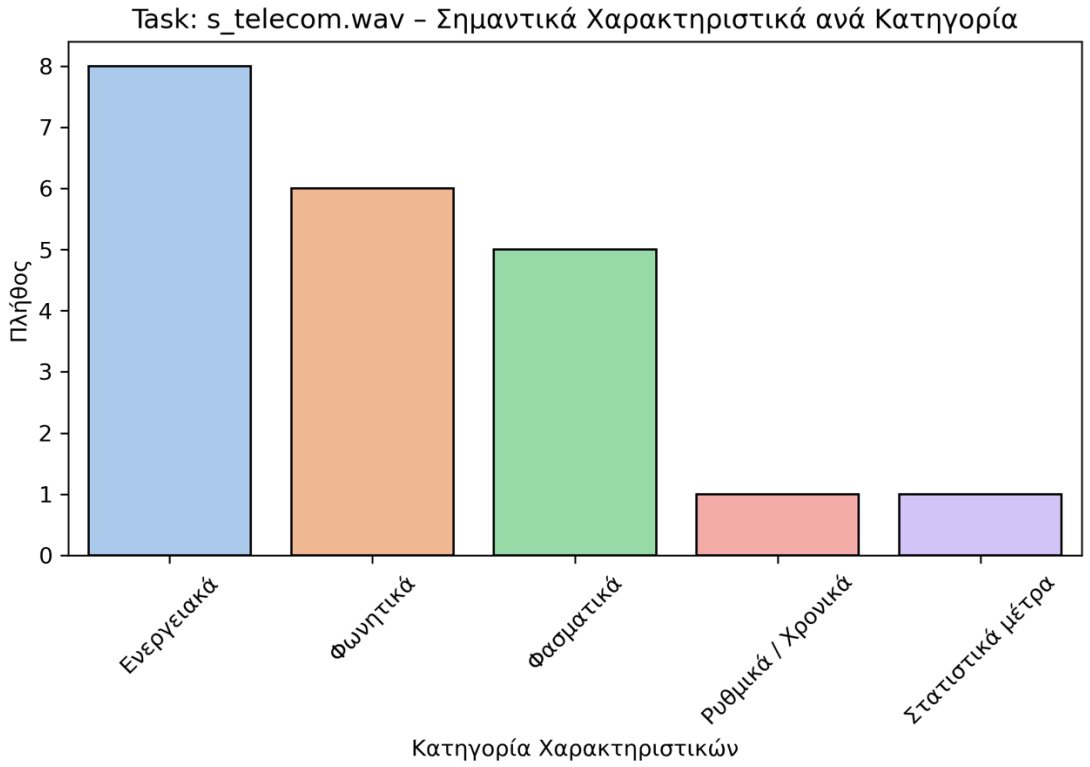
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
F2frequency_sma3nz_stddevNorm	0.000024	Φωνητικά
F1frequency_sma3nz_stddevNorm	0.000336	Φωνητικά
F0semitoneFrom27.5Hz_sma3nz_percentile20.0	0.001462	Φωνητικά
mfcc2_sma3_amean	0.003548	Φασματικά
mfcc1V_sma3nz_amean	0.004982	Φασματικά



Εικόνα 108:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_spaceship.wav του Task 5

Πίνακας 89:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για την άσκηση s_telecom.wav του Task 5 (A5)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc1V_sma3nz_stddevNorm	0.009549	Φασματικά
loudness_sma3_amean	0.010642	Ενεργειακά
loudness_sma3_amean	0.015093	Ενεργειακά
StddevVoicedSegmentLengthSec	0.017497	Ρυθμικά / Χρονικά
F0semitoneFrom27.5Hz_sma3nz_percentile20.0	0.018154	Φωνητικά



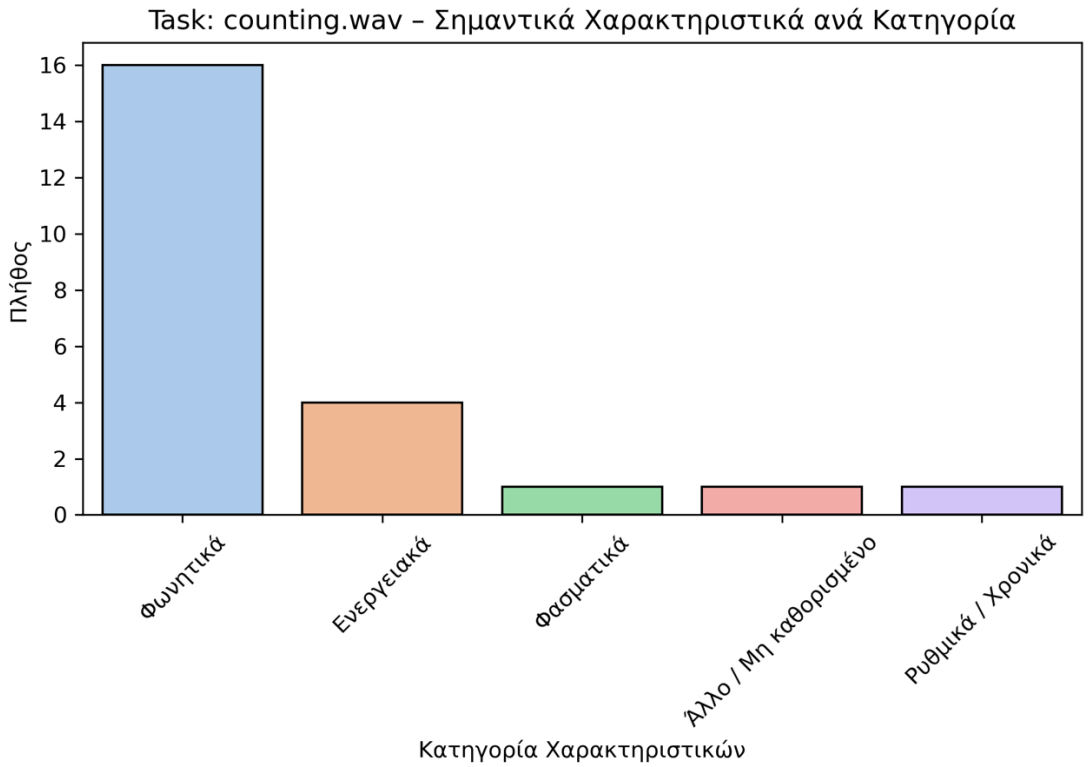
Εικόνα 109:Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση s_telecom.wav του Task 5

Task 6 (A6): Καταμέτρηση Αριθμών – Αξιολόγηση Ρυθμού Ομιλίας

Στην συγκεκριμένη άσκηση ζητείται από τον ασθενή να μετρήσει από το 1 έως το 10 με σκοπό την καταγραφή της χρονικής αλληλουχίας και της διάρκειας των παύσεων μεταξύ της καταμέτρησης. Με αυτόν τον τρόπο μπορεί να γίνει αξιολόγηση της ρυθμικότητας και της συνοχής της ομιλίας του ατόμου, χαρακτηριστικά που συνήθως αλλοιώνονται σε άτομα που νοσούν από την Νόσο του Πάρκινσον. Η ανάλυση που ακολουθεί συνοψίζει τα ευρήματα που σχετίζονται με την συγκεκριμένη άσκηση.

Πίνακας 90:Τα πέντε σημαντικότερα χαρακτηριστικά (p < 0.05) για το σύνολο του Task 6 (A6)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
loudnessPeaksPerSec	0.000057	Ενεργειακά
loudnessPeaksPerSec	0.010551	Ενεργειακά
HNRdBACF_sma3nz_stddevNorm	0.011570	Φωνητικά
F3frequency_sma3nz_amean	0.012083	Φωνητικά
HNRdBACF_sma2nz_stddevNorm	0.012686	Φωνητικά



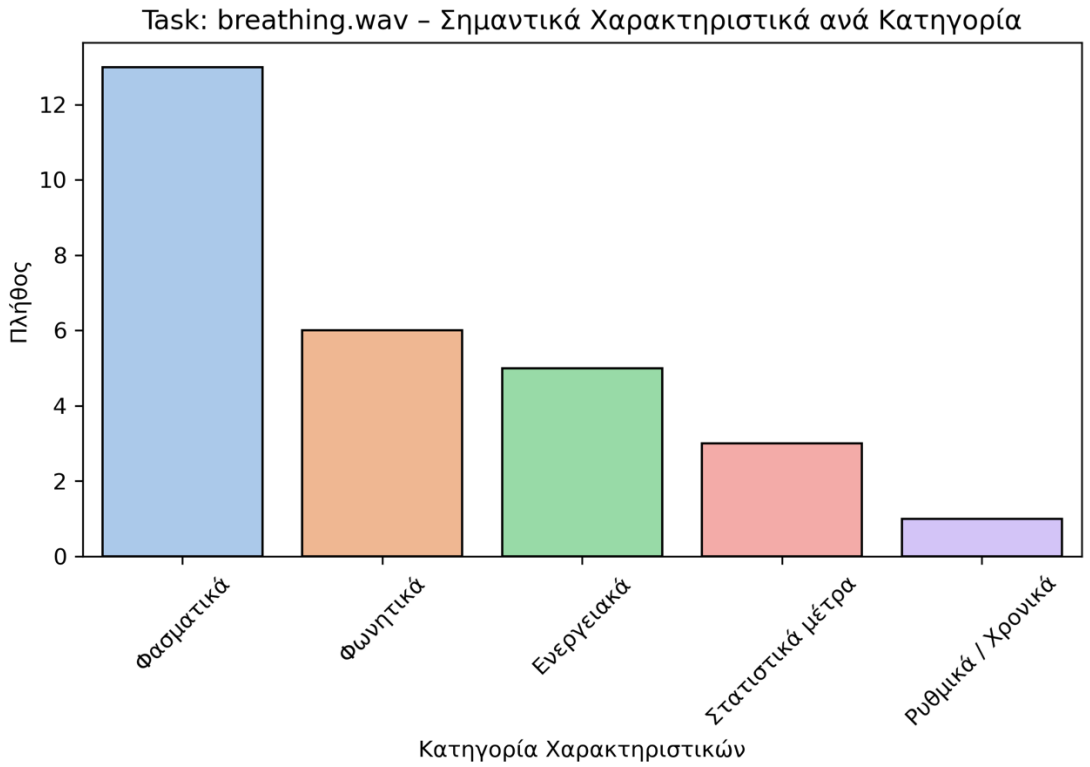
Εικόνα 110: Πλήθος σημαντικών χαρακτηριστικών (p < 0.05) ανά λειτουργική κατηγορία για την άσκηση counting.wav του Task

Task 7 (A7) : Καταγραφή Αναπνοής – Φωνητική Συμπεριφορά

Ακόλουθα, σε αυτή την άσκηση αξιολογείται η αναπνοή του ατόμου πριν αλλά και μετά την εκφορά των λέξεων προκειμένου να γίνει αξιολόγηση της αναπνευστικής ικανότητας του ατόμου και για να παρθούν πληροφορίες για το κατά πόσο το άτομο είναι σε θέση να συντονίσει την αναπνοή του και κατά συνέπεια γενικότερα την φωνή, παράμετροι αρκετά σημαντικοί για την αξιολόγηση. Στον πίνακα που παρατίθεται καταγράφονται τα χαρακτηριστικά που κρίθηκαν ως στατιστικά σημαντικά καθώς και το αντίστοιχο σχεδιάγραμμα με την συνολική κατανομή των χαρακτηριστικών.

Πίνακας 91:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για το σύνολο του Task 7 (A7)

Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
mfcc3V_sma3nz_stddevNorm	0.000030	Φασματικά
mfcc1_sma3_amean	0.000548	Φασματικά
MeanVoicedSegmentLengthSec	0.000943	Ρυθμικά / Χρονικά
mfcc3_sma3_stddevNorm	0.001132	Φασματικά
mfcc1V_sma3nz_amean	0.001528	Φασματικά



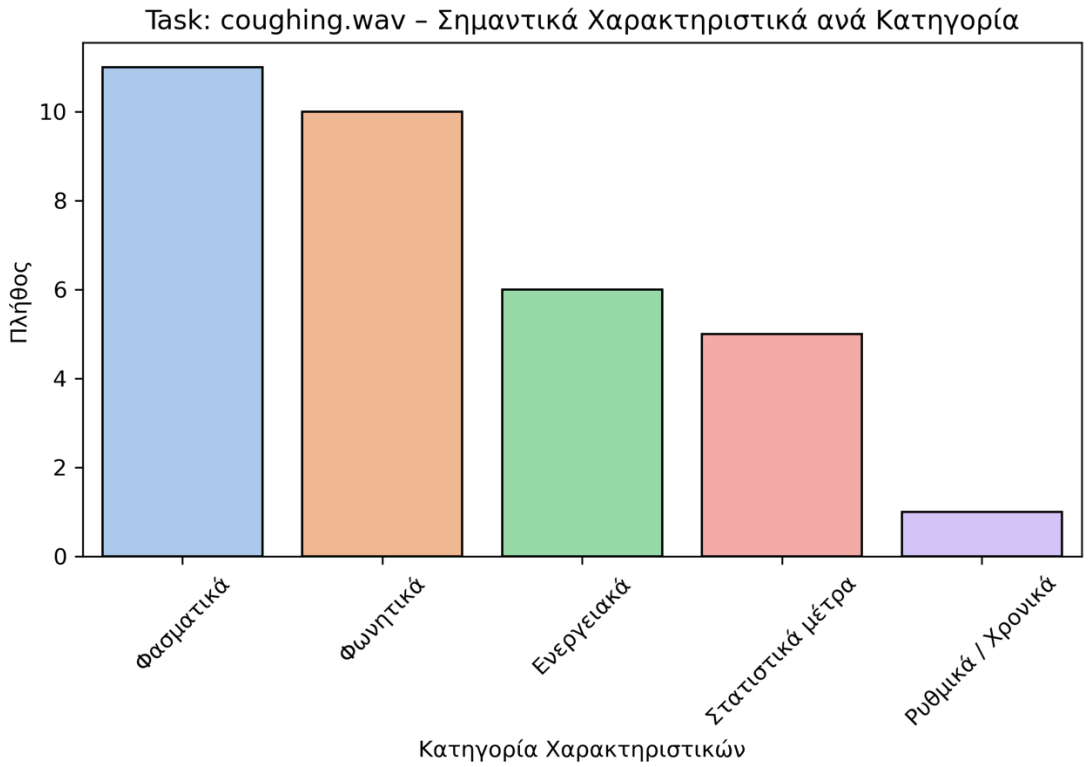
Εικόνα 111:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση breathing.wav του Task

Task 8 (A8): Ηχογράφηση Βήχα – Ανάλυση Εκτός του Φάσματος Ομιλίας

Τέλος το τελευταίο σετ ασκήσεων που μελετήθηκε αφορούσε την καταγραφή του βήχα, που αν και σαν άσκηση δεν εντάσσεται άμεσα στην παραγωγή του λόγου, συνδέεται με την αποτύπωσή της λειτουργίας του αναπνευστικού συστήματος και μέσω της περαιτέρω ανάλυσης μπορεί να εντοπισθούν αλλοιώσεις που συνδέονται άρρηκτα με φωνητικές δυσλειτουργίες οι οποίες ενδέχεται να έχουν επηρεαστεί από την ύπαρξη της ΝΠ. Στον ακόλουθο πίνακα, καθώς και στον συγκεντρωτικό γράφημα φαίνονται τα σημαντικότερα εξαγόμενα χαρακτηριστικά αλλά και η γενικότερη εικόνα των συνολικών χαρακτηριστικών.

Πίνακας 92:Τα πέντε σημαντικότερα χαρακτηριστικά ($p < 0.05$) για το σύνολο του Task 8 (A8)

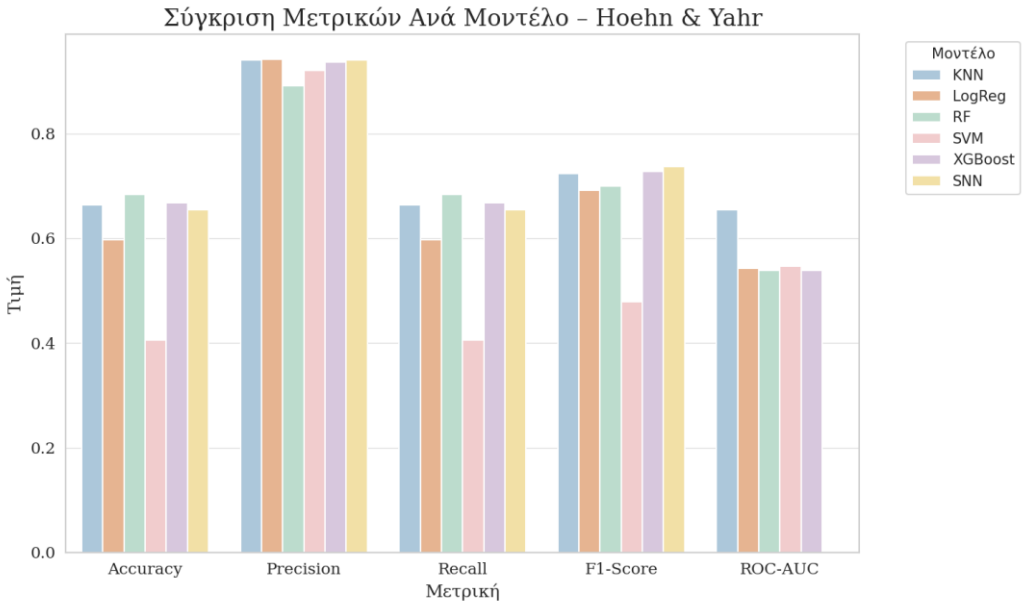
Χαρακτηριστικό (Predictor)	P Value	Κατηγορία
alphaRatioUV_sma3nz_amean	0.002189	Φασματικά
F0semitoneFrom27.5Hz_sma3nz_meanFallingSlope	0.003595	Φωνητικά
mfcc3V_sma3nz_amean	0.004034	Φασματικά
loudness_sma3_stddevNorm	0.006310	Ενεργειακά
mfcc1V_sma3nz_amean	0.007026	Φασματικά



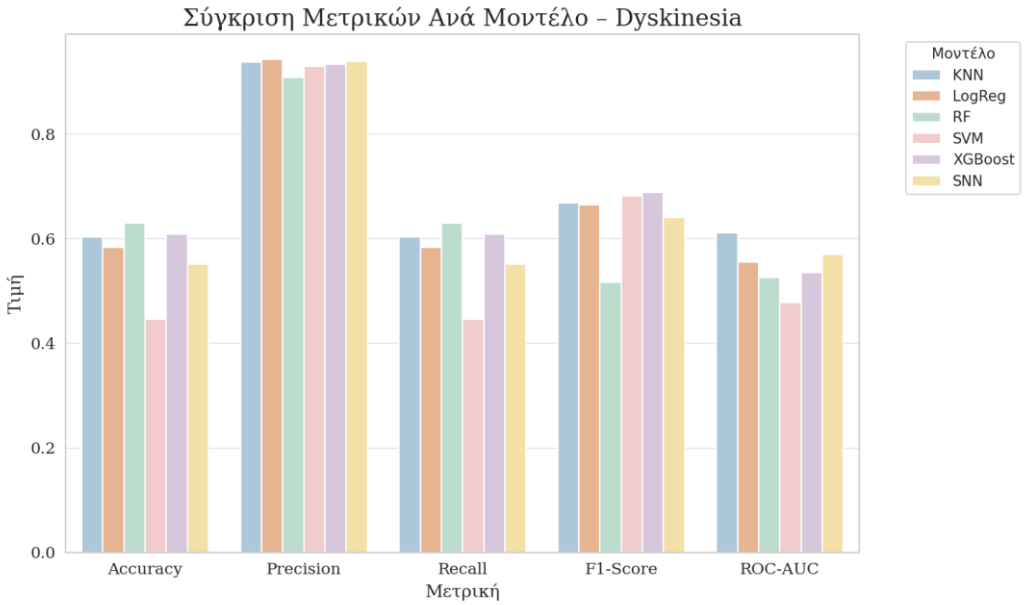
Εικόνα 112:Πλήθος σημαντικών χαρακτηριστικών ($p < 0.05$) ανά λειτουργική κατηγορία για την άσκηση coughing.wav του Task

Παράρτημα Δ – Αναλυτικά Διαγράμματα Αξιολόγησης Μοντέλων

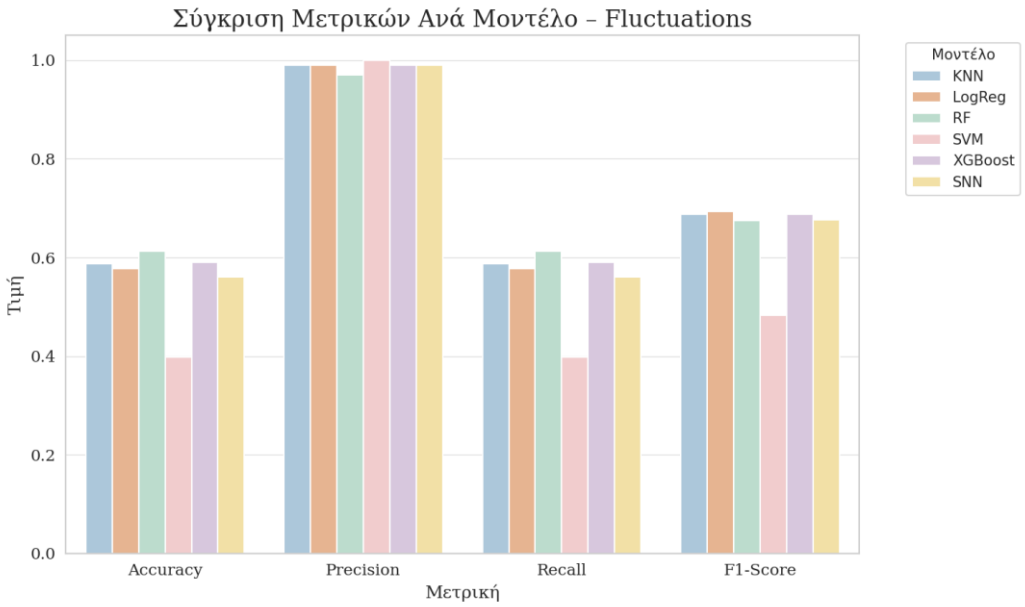
Στο Παράρτημα Δ παρατίθενται όλα τα αναλυτικά διαγράμματα που αφορούν την συγκριτική απόδοση των έξι (6) ταξινομητικών μοντέλων που εξετάζονται στην παρούσα μελέτη. Το κάθε διάγραμμα αντιστοιχεί και σε ένα από τα προβλήματα πρόβλεψης.
Η αξιολόγηση του κάθε μοντέλου πραγματοποιείται με την χρήση των καθιερωμένων μετρικών, δηλαδή η Ακρίβεια (Accuracy) , η Ευστοχία (Precision), η Ανάκληση (Recall), το F1-Score και όπου ήταν εφικτό η περιοχή κάτω της καμπύλης ROC (ROC-AUC).



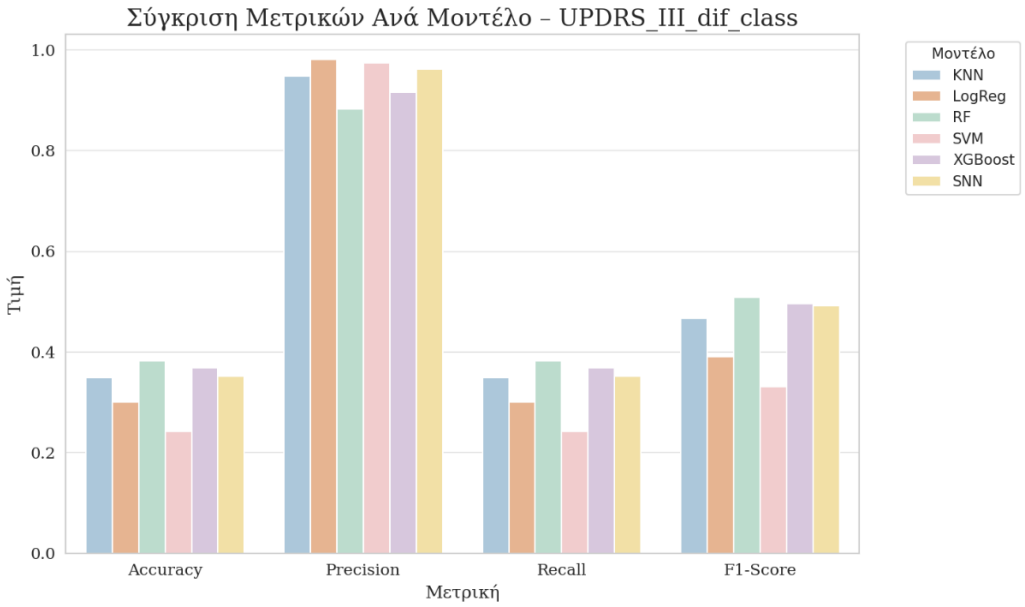
Εικόνα 113:Σύγκριση μετρικών απόδοσης των μοντέλων για την πρόβλεψη της κλίμακας Hoehn & Yahr.



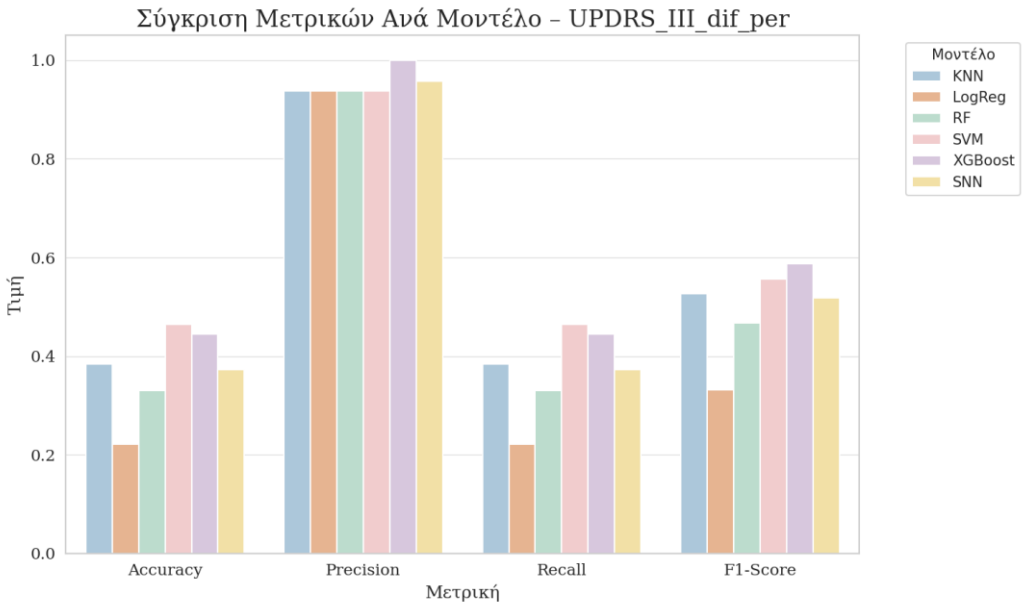
Εικόνα 114: Σύγκριση μετρικών απόδοσης των μοντέλων για την πρόβλεψη της Δυσκινησίας (Dyskinesia).



Εικόνα 115: Σύγκριση μετρικών απόδοσης των μοντέλων για την πρόβλεψη των Κινητικών Διακυμάνσεων (Fluctuations).



Εικόνα 116: Σύγκριση μετρικών απόδοσης των μοντέλων για την πρόβλεψη της κατηγορίας διαφοροποίησης στη κλίμακα UPDRS Part III (UPDRS_III_dif_class).



Εικόνα 117: Σύγκριση μετρικών απόδοσης των μοντέλων για την πρόβλεψη της ποσοστιαίας διαφοροποίησης στην κλίμακα UPDRS Part III (UPDRS_III_dif_per).

Παράρτημα Δ – Κώδικες Υλοποίησης και Πειραματικών Διαδικασιών

Στο παρόν παράρτημα περιλαμβάνει τον πλήρη κώδικα που αναπτύχθηκε για τις ανάγκες της παρούσας διπλωματικής εργασίας, με στόχο την διασφάλιση της διαφάνειάς, της επιστημονικής εγκυρότητας και την δυνατότητα αναπαραγωγής των αποτελεσμάτων. Ο κώδικας έχει ως σκοπό την πλήρη τεκμηρίωσή των σταδίων της πειραματικής διαδικασίας, όπως αυτή παρουσιάζεται αναλυτικά στην ενότητα της μεθοδολογίας (Κεφάλαιο 3).

Η υλοποίηση του κώδικα έχει πραγματοποιηθεί σε γλώσσα προγραμματισμού Python (έκδοση 3.11.5) για το πλήθος του κώδικα και R (έκδοση 4.4.1) για την στατιστική ανάλυση και οργανώνεται σε αρχεία Jupyter Notebook (.ipynb). Για την υλοποίηση χρησιμοποιήθηκαν οι καθιερωμένες βιβλιοθήκες, αποδεκτές από την ερευνητική κοινότητα, όπως Pandas για χειρισμό και ανάλυση δεδομένων, NumPy για αριθμητικούς υπολογισμούς, Scikit-learn για την εκπαίδευση και αξιολόγηση μοντέλων, Matplotlib και Seaborn για τη δημιουργία διαγραμμάτων και στατιστικών απεικονίσεων, καθώς και άλλες βιβλιοθήκες χρήσιμές για την περάτωση της διαδικασίας. Ο κώδικας καλύπτει όλα τα στάδια του ερευνητικού πειράματος: Καθαρισμός/Προεπεξεργασία Δεδομένων, Στατιστική ανάλυση, Σχηματισμός/Εκπαίδευση Μοντέλων, Αξιολόγηση Απόδοσης και Οπτικοποίηση Αποτελεσμάτων.

Όλο το υλικό είναι αναρτημένο σε δημόσιο αποθετήριο GitHub, το οποίο περιλαμβάνει και αναλυτικό αρχείο README.md με τις απαραίτητες οδηγίες εγκατάστασης και εκτέλεσης.

Σύνδεσμος αποθετηρίου GitHub: https://github.com/doralappa/thesis-parkinson