



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

**Ανάπτυξη μοντέλων μηχανικής μάθησης για την αναγνώριση
Covid-19 μέσω ηχητικών καταγραφών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χαράλαμπος Μ. Φρατζάκης

Επιβλέπων : Δρ. Κωνσταντίνα Σ. Νικήτα
Καθηγήτρια, Σ.Η.Μ.Μ.Υ. Ε.Μ.Π.

Αθήνα, Σεπτέμβριος 2024



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

**Ανάπτυξη μοντέλων μηχανικής μάθησης για την αναγνώριση
Covid-19 μέσω ηχητικών καταγραφών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χαράλαμπος Μ. Φρατζάκης

Επιβλέπων : Δρ. Κωνσταντίνα Σ. Νικήτα
Καθηγήτρια, Σ.Η.Μ.Μ.Υ. Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 14η Οκτωβρίου 2024.

Δρ. Κωνσταντίνα Σ. Νικήτα,
Καθηγήτρια, Σ.Η.Μ.Μ.Υ.
Ε.Μ.Π.

Δρ. Βουλοδημος Αθανασιος,
Επικουρος Καθηγητης,
Σ.Η.Μ.Μ.Υ Ε.Μ.Π.

Δρ. Γιώργος Στάμου
Καθηγητης, Σ.Η.Μ.Μ.Υ
Ε.Μ.Π.

Χαράλαμπος Μ. Φρατζάκης

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

Copyright © Χαράλαμπος Μ. Φρατζάκης, Εθνικό Μετσόβιο Πολυτεχνείο, 2024.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Η πανδημία COVID-19 έδωσε το έναυσμα για την ανάπτυξη καινοτόμων λύσεων σε πολλούς τεχνολογικούς τομείς, ανοίγοντας δρόμους στην ανάπτυξη ανασταλτικών μέσων για την καταπολέμηση της παγκόσμιας εξάπλωσης της. Η Τεχνητή Νοημοσύνη (AI) και η Μηχανική Μάθηση (ML) αναπόφευκτα παίζουν πρωτεύοντα ρόλο στην αντιμετώπιση των προκλήσεων που θέτει ο COVID-19, επιταχύνοντας την έρευνα, ενισχύοντας με διαγνωστικά εργαλεία τη διάγνωση και τελικά προσφέροντας λύσεις στην φαρέτρα των επιστημόνων.

Η παρούσα διπλωματική εργασία πραγματεύεται την ανάπτυξη μοντέλων μηχανικής μάθησης για τη διάγνωση της Covid-19 μέσω ηχητικών καταγραφών βήχα. Συγκεκριμένα, αναπτύσσεται πρόγραμμα εκπαίδευσης ταξινομητών με αρθρωτή δομή στη γλώσσα Python με στόχο την αναζήτηση της καλύτερης συνδυαστικής ρύθμισης υπερπαραμέτρων εξαγωγής χαρακτηριστικών για τους τύπους ταξινομητών Logistic Regression, K-Nearest Neighbors και Multi-layer Perceptron. Για την εξαγωγή χαρακτηριστικών χρησιμοποιούνται η μέθοδος MFCCs, Spectral Centroid, Spectral Roll-off και Zero-Crossing Rate. Για την εκπαίδευση και αξιολόγηση των ταξινομητών χρησιμοποιούνται οι βάσεις δεδομένων SmartyforCovid, και Coswara.

Τα αποτελέσματα καταδεικνύουν την ικανότητα των επιλεγθέντων μοντέλων να διακρίνουν αποτελεσματικά τα πάσχοντα με COVID-19 άτομα, καθώς και τη σημασία της σωστής επιλογής συνδυασμού χαρακτηριστικών για το έργο αυτό.

Λέξεις Κλειδιά

Covid-19, μηχανική και βαθιά μάθηση, ανάλυση αρχείων ήχου, MFCCs, Spectral Centroid, Spectral Roll-off, Zero-Crossing Rate (ZCR).

Abstract

The COVID-19 pandemic catalyzed innovations across numerous technological fields, paving the way for the development of preventive measures to combat global spread. Artificial Intelligence (AI) and Machine Learning (ML) inevitably play a central role in addressing the challenges posed by COVID-19, accelerating research, enhancing diagnostics through advanced tools, and ultimately offering solutions to scientists.

The present Thesis addresses the development of machine learning models for the diagnosis of COVID-19 through cough audio recordings. A modular classifier training program was developed in Python to determine the optimal hyperparameter settings to extract the data features to be used in training the Logistic Regression, K-Nearest Neighbors, and Multi-layer Perceptron classifiers. Feature extraction was based on the MFCCs method, Spectral Centroid, Spectral Roll-off and Zero-Crossing Rate. The classifiers were trained using the SmartyforCovid and Coswara databases.

Obtained results indicate the employed model's capacity to effectively distinguish between individuals with COVID-19, as well as the importance of correct feature combination towards this task.

Key words

Covid-19, Deep and Machine Learning, Speech and Signal Processing, MFCCs, Spectral Centroid, Spectral Roll-off, Zero-Crossing Rate (ZCR)

Ευχαριστίες

Θα ήθελα να ευχαριστήσω την καθηγήτριά μου Δρ. Κωνσταντίνα Νικήτα, που μου έδωσε την πολύτιμη ευκαιρία να εκπονήσω τη διπλωματική μου εργασία στο Εργαστήριο Βιοϊατρικών Προσομοιώσεων και Απεικόνισης της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσοβίου Πολυτεχνείου.

Επιπλέον, θα ήθελα να ευχαριστήσω θερμά τον Δρ. Κωνσταντίνο Μήτση και Δρ Μαρια Αθανασίου και τον Υποψήφιο Διδάκτορα Θεοφάνη Γανιτίδη για τη συνεργασία και τη βοήθεια που μου προσέφεραν.

Τέλος, θέλω να ευχαριστήσω από τα βάθη της καρδιάς μου την οικογένειά μου για την αδιάλειπτη υποστήριξη και αγάπη τους που μου προσφέρουν όλη μου τη ζωή.

Πίνακας Περιεχομένων

Περίληψη.....	5
Abstract	6
Ευχαριστίες.....	7
1. Εισαγωγή	21
1.1. Ιστορικό	21
1.2. Smarty4covid.....	22
1.3. Σκοπός της παρούσας διπλωματικής εργασίας	23
1.4. Δομή της Εργασίας.....	23
1.5. Κεφάλαια	24
2. Θεωρητικό μέρος.....	27
2.1. Covid-19.....	27
2.2. Μηχανική μάθηση	30
2.3. Βαθιά μάθηση.....	31
2.4. Διαφορές Τεχνητής Νοημοσύνης - Μηχανικής Μάθησης - Βαθιάς Μάθησης.....	32
2.5. Εφαρμογές μηχανικής μάθησης και βαθιάς μάθησης κατά την διάρκεια της πανδημίας COVID-19	33
2.6. Ανάλυση αρχείων ήχου	38
3. Μεθοδολογία	41
3.1. Εισαγωγή Δεδομένων.....	41
3.2. Εξαγωγή Χαρακτηριστικών	45
3.3. Προεπεξεργασία Δεδομένων	52
3.4. Μέθοδος Ensemble learning.....	56
3.5. Εκπαίδευση Ταξινομητών	58
3.6. Έλεγχος εκπαίδευσης ταξινομητών.....	64
3.7. Αξιολόγηση Ταξινομητών – Δείκτες αξιολόγησης.....	67
3.8. Ολοκλήρωση Προγράμματος	70
4. Αποτελέσματα εκπαίδευσης ταξινομητών	71
4.1. Αποτελέσματα εκπαίδευσης ταξινομητή Logistic Regression.....	71
4.2. Αποτελέσματα εκπαίδευσης ταξινομητή K-Nearest Neighbors	88
4.3. Αποτελέσματα εκπαίδευσης ταξινομητή Multi-layer Perceptron	105
4.4. Συγκριτικά στοιχεία.....	121
5. Συμπεράσματα - Προτάσεις	123
Βιβλιογραφία – Σημειώσεις – Πηγές	125

Κατάλογος Εικόνων

Εικόνα 1: Χάρτης Κρουσμάτων ανά Περιφερειακή Ενότητα στην Ελλάδα από τις 26 Φεβρουαρίου 2020 έως και τις 20 Δεκεμβρίου 2022[8]	28
Εικόνα 2: Συμπτώματα της COVID-19[13]	29
Εικόνα 3: Κατά κεφαλήν κρούσματα παγκοσμίως[14]	29
Εικόνα 4: AI vs ML vs DL [26]	33
Εικόνα 5: Αριθμός δημοσιεύσεων στα 10 μεγαλύτερα πεδία [33]	35
Εικόνα 6: Διάγραμμα ροής της επιλογής των δημοσιευμένων άρθρων στη μελέτη “A Comprehensive Overview of the COVID-19 Literature: Machine Learning–Based Bibliometric Analysis” [35]	36
Εικόνα 7: Αριθμός δημοσιεύσεων ανά μήνα (2020-2021) [39]	37
Εικόνα 8: Εφαρμογή της τεχνητής νοημοσύνης και της μηχανικής μάθησης στη μάχη κατά της COVID-19 [31]	38
Εικόνα 9: Σχηματική διάταξη αρχείων δεδομένων χρηστών smarty4covid	43
Εικόνα 10: Σχηματική διάταξη αρχείων smarty4covid ονοματισμένων ρόλων με καθορισμένη ιεραρχία	44
Εικόνα 11: Τμήμα MFCCs ενός segment	47
Εικόνα 12: Spectral Centroid ενός segment	48
Εικόνα 13: Spectral Roll-off ενός segment	49
Εικόνα 14: Zero-Crossing Rate ενός segment	50
Εικόνα 15: Τμήμα χαρακτηριστικών ενός δείγματος μετά από κλιμάκωση	53
Εικόνα 16: Τμήμα χαρακτηριστικών ενός δείγματος μετά από μέθοδο PCA	54
Εικόνα 17: Πλήθος δεδομένων στο σύνολο εκπαίδευσης και στο σύνολο ελέγχου	56
Εικόνα 18: Ανισορροπία των κλάσεων των δεδομένων	57
Εικόνα 19: Τα υποσύνολα εκπαίδευσης μετά το ensemble split	58
Εικόνα 20: Καμπύλη ROC για για “καλύτερους” και “χειρότερους” ταξινομητές[67]	66

Κατάλογος Πινάκων

Πίνακας 1: Τιμές υπερπαραμέτρων για την εξαγωγή χαρακτηριστικών	51
Πίνακας 2: Τιμές υπερπαραμέτρων για κάθε μοντέλο ταξινόμητη.....	60
Πίνακας 3: Σταθερά Στοιχεία Εκπαίδευσης	71
Πίνακας 4: Μεταβλητά Στοιχεία Εκπαίδευσης	71
Πίνακας 5: Αποτελέσματα εκπαίδευσης ταξινόμητη LR στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments.....	72
Πίνακας 6 Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινόμητη LR στο σύνολο δεδομένων smarty4covid.....	74
Πίνακας 7: Max F1 Score value per segment αποτελέσματος εκπαίδευσης ταξινόμητη LR στο σύνολο δεδομένων smarty4covid	75
Πίνακας 8: Max F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινόμητη LR στο σύνολο δεδομένων smarty4covid.....	76
Πίνακας 9: Best 20 values based on F1 Score αποτελέσματος εκπαίδευσης ταξινόμητη LR στο σύνολο δεδομένων smarty4covid	77
Πίνακας 10: Αποτελέσματα εκπαίδευσης ταξινόμητη LR στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments.....	80
Πίνακας 11: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινόμητη LR στο σύνολο δεδομένων coswara.....	81
Πίνακας 12: Max F1 Score value per segment αποτελέσματος εκπαίδευσης ταξινόμητη LR στο σύνολο δεδομένων coswara.....	82
Πίνακας 13: Max F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινόμητη LR στο σύνολο δεδομένων coswara	84
Πίνακας 14: Best 20 values based on F1 Score αποτελέσματος εκπαίδευσης ταξινόμητη LR στο σύνολο δεδομένων coswara.....	84
Πίνακας 15: Στατιστικά στοιχεία πίνακα σύγκρισης αποτελεσμάτων εκπαίδευσης ταξινόμητη LR.....	86
Πίνακας 16: Max F1 Score value per frame_size & n_segments αποτελεσμάτων εκπαίδευσης ταξινόμητη LR	87
Πίνακας 17: Μέσος Όρος F1 Score value per frame_size & n_segments αποτελεσμάτων εκπαίδευσης ταξινόμητη LR	88
Πίνακας 18: Αποτελέσματα εκπαίδευσης ταξινόμητη KNN στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments.....	89

Πίνακας 19: Στατιστικά στοιχεία πίνακα αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid	91
Πίνακας 20: Max F1 Score value per segment αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid	92
Πίνακας 21: Max F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid	93
Πίνακας 22: Best 20 values based on F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid	94
Πίνακας 23: Αποτελέσματα εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments.....	96
Πίνακας 24: Στατιστικά στοιχεία πίνακα αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara.....	97
Πίνακας 25: Max F1 Score value per segment αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara	98
Πίνακας 26: Max F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara.....	100
Πίνακας 27: Best 20 values based on F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara	100
Πίνακας 28: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινομητή KNN	102
Πίνακας 29: Max F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή KNN	103
Πίνακας 30: Μέσος Όρος F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή	104
Πίνακας 31: Αποτελέσματα εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments.....	106
Πίνακας 32: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid.....	107
Πίνακας 33: Max F1 Score value per segment αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid.....	108
Πίνακας 34: Max F1 Score value per frame_size & n_segments αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid	110
Πίνακας 35: Best 20 values based on F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid.....	110

Πίνακας 36: Αποτελέσματα εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments.....	113
Πίνακας 37: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara.....	114
Πίνακας 38: Max F1 Score value per segment αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara.....	115
Πίνακας 39: Max F1 Score value per frame_size & n_segments αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara	117
Πίνακας 40: Best 20 values based on F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara.....	117
Πίνακας 41: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινομητή MLP	119
Πίνακας 42: Max F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή MLP	120
Πίνακας 43: Μέσος Όρος F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή	120
Πίνακας 44: Max F1 Score, βάσεων δεδομένων & υπερπαραμέτρων ταξινομητών	121
Πίνακας 45: Max F1 Score value per frame_size & n_segments αποτελεσμάτων εκπαίδευσης ταξινομητών	122

Κατάλογος Διαγραμμάτων

Διάγραμμα 1: Παβδόγραμμα αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments	73
Διάγραμμα 2: Διάγραμμα αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments	73
Διάγραμμα 3: Παβδόγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size per n_segment.....	75
Διάγραμμα 4: Διάγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size per n_segment.....	76
Διάγραμμα 5: Παβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid (Best 20 values based on F1 Score).....	78
Διάγραμμα 6: Παβδόγραμμα αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments	80
Διάγραμμα 7: Διάγραμμα αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments	81
Διάγραμμα 8: Παβδόγραμμα MAX αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size per n_segment.....	83
Διάγραμμα 9: Διάγραμμα MAX αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size per n_segment	83
Διάγραμμα 10: Παβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara (Best 20 values based on F1 Score)	85
Διάγραμμα 11: Οριζόντιο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή LR	86
Διάγραμμα 12: Κάθετο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή LR.....	87
Διάγραμμα 13: Παβδόγραμμα αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments	90
Διάγραμμα 14: Διάγραμμα αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments	90

Διάγραμμα 15: Ραβδόγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments	92
Διάγραμμα 16: Διάγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments	93
Διάγραμμα 17: Ραβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid (Best 20 values based on F1 Score).....	94
Διάγραμμα 18: Ραβδόγραμμα αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments	96
Διάγραμμα 19: Διάγραμμα αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments	97
Διάγραμμα 20: Ραβδόγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments	99
Διάγραμμα 21: Διάγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments	99
Διάγραμμα 22: Ραβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara (Best 20 values based on F1 Score).....	101
Διάγραμμα 23: Οριζόντιο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή KNN	102
Διάγραμμα 24: Κάθετο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή KNN	103
Διάγραμμα 25: Ραβδόγραμμα αποτελέσματος εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments	106
Διάγραμμα 26: Διάγραμμα αποτελέσματος εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments	107
Διάγραμμα 27: Ραβδόγραμμα μέγιστων τιμών F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size per n_segment.....	109

Διάγραμμα 28: Διάγραμμα μέγιστων τιμών F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size per n_segment.....	109
Διάγραμμα 29: Ραβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid (Best 20 values based on F1 Score).....	111
Διάγραμμα 30: Ραβδόγραμμα αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments	113
Διάγραμμα 31: Διάγραμμα αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments	114
Διάγραμμα 32: Ραβδόγραμμα μέγιστων τιμών F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size per n_segment.....	116
Διάγραμμα 33: Διάγραμμα μέγιστων τιμών F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size per n_segment.....	116
Διάγραμμα 34: Ραβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara (Best 20 values based on F1 Score)	118
Διάγραμμα 35: Οριζόντιο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή MLP	119
Διάγραμμα 36: Κάθετο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή MLP.....	119

1. Εισαγωγή

1.1. Ιστορικό

Η πανδημία COVID-19 έδωσε το έναυσμα για την ανάπτυξη καινοτομιών σε πολλούς τεχνολογικούς τομείς, ανοίγοντας δρόμους στην ανάπτυξη ανασταλτικών μέσων για την καταπολέμηση της παγκόσμιας εξάπλωσης. Η Τεχνητή Νοημοσύνη (AI) και η Μηχανική Μάθηση (ML) αναπόφευκτα παίζουν πρωτεύοντα ρόλο στην αντιμετώπιση των προκλήσεων που θέτει ο COVID-19, επιταχύνοντας την έρευνα, ενισχύοντας με διαγνωστικά εργαλεία τη διάγνωση και τελικά προσφέροντας λύσεις στην φαρέτρα των επιστημόνων. Μια προσέγγιση κινητής υγείας που μπορεί να υποστηρίξει οικονομικά, γρήγορα, βιώσιμα και αποτελεσματικά τεστ, διευκολύνοντας πολλαπλές επαναλήψεις για την παρακολούθηση της εξέλιξης, θα μπορούσε να συμβάλει στον περιορισμό της εξάπλωσης της ασθένειας και στην καταστολή της αναζωπύρωσης της.

Μέσα σε αυτό το πλαίσιο, η ιδέα της αξιοποίησης της τεχνητής νοημοσύνης σε συνδυασμό με τις τρέχουσες κινητές τεχνολογίες έχει κινητοποιήσει την παγκόσμια επιστημονική κοινότητα για την υλοποίηση μιας εύχρηστης και ευρέως προσβάσιμης μεθόδου ανίχνευσης COVID-19. Μέσω της ανάλυσης σημάτων και της εφαρμογής τεχνητής νοημοσύνης πάνω σε ηχογραφήσεις βήχα, φωνής και αναπνοής ερευνάται η δυνατότητα ανίχνευσης καινοτόμων βιοδεικτών σχετικών με τον COVID-19.

Στη σύγχρονη βιβλιογραφία, οι περισσότερες προσεγγίσεις για την πρόβλεψη του κινδύνου COVID-19 από ηχογραφήσεις βασίζονται σε μοντέλα βαθιάς μάθησης, τα οποία τυπικά απαιτούν μεγάλες ποσότητες δεδομένων για να εκπαιδευτούν. Υπάρχουν πολλές μελέτες που έχουν προσανατολιστεί στη συλλογή ηχητικών εγγραφών από πολίτες, χρησιμοποιώντας μια διαδικτυακή διεπαφή για την συλλογή δεδομένων μέσω πλήθους. Η πρώτη προσπάθεια προς αυτή την κατεύθυνση ξεκίνησε στο πλαίσιο του έργου COVID-19 Sounds, η βάση δεδομένων του οποίου περιλαμβάνει 53.449 ηχητικά δείγματα. Το Coswara θεωρείται επίσης μία από τις μεγαλύτερες βάσεις δεδομένων συλλογής δεδομένων μέσω πλήθους, αν και περιλαμβάνει μόνο ήχους βήχα με 27.550 εγγραφές.

Τα δεδομένα συλλογής μέσω πλήθους όμως, περιλαμβάνουν τόσο ηχητικές εγγραφές που δεν σχετίζονται με το επιθυμητό περιεχόμενο της βάσης δεδομένων όσο και ηχητικές εγγραφές χαμηλής ποιότητας και αυξημένου θορύβου. Το πρόβλημα αυτό οδηγεί στην ανάγκη για μεθόδους επιμέλειας των δεδομένων, μια εξ' αυτών και το έργο smarty4covid.

1.2. Smarty4covid

Το Smarty4covid είναι ένα καινοτόμο έργο που αναπτύχθηκε για την αξιολόγηση και παρακολούθηση του κινδύνου της πανδημίας COVID-19 μέσω μιας πολυτροπικής προσέγγισης που συνδυάζει ηχητικά δεδομένα και μηχανική μάθηση. Το έργο ξεκίνησε τον Δεκέμβριο του 2021 και ολοκληρώθηκε τον Δεκέμβριο του 2022, με στόχο την ανάπτυξη ενός πλαισίου βαθιάς μάθησης που θα μπορούσε να εντοπίσει νέους βιοδείκτες της λοίμωξης και της προόδου της νόσου. Ο στόχος του έργου Smarty4covid ήταν η δημιουργία ενός έξυπνου πολυτροπικού πλαισίου για την αξιολόγηση και παρακολούθηση του κινδύνου COVID-19, καθώς τα ηχητικά σήματα ή τα ηχογραφήματα συνήθως αποτελούν άριστα δεδομένα εισόδου των μοντέλων βαθιάς μάθησης. Μετά από τις απαραίτητες εγκρίσεις από την Επιτροπή Ηθικής της Έρευνας του Εθνικού Μετσόβιου Πολυτεχνείου [1], μια διαδραστική διαδικτυακή εφαρμογή [2] υλοποιήθηκε και κυκλοφόρησε δημόσια ως μέσο συλλογής δεδομένων με ελεύθερη συγκατάθεσή των συμμετεχόντων στην έρευνα αποκλειστικά για ερευνητικούς σκοπούς, στοχεύοντας στην ενημέρωση της διάγνωσης και εξέλιξης της νόσου COVID-19 με την ανάπτυξη αλγορίθμων μηχανικής μάθησης.

Στην πράξη, το Smarty4covid συνδύασε ηχητικά δεδομένα όπως η αναπνοή, η φωνή και ο βήχας των ασθενών με αυτοαναφερόμενους παράγοντες κινδύνου και κλινικά δεδομένα. Αυτό το πολυτροπικό σύστημα εφαρμόστηκε τόσο σε κλινικές μελέτες με νοσηλευόμενους ασθενείς σε συνεργασία με το νοσοκομείο ΑΧΕΠΑ, όσο και σε ένα ευρύτερο πληθυσμό μέσω μιας διαδικτυακής εφαρμογής για τη συλλογή δεδομένων από το κοινό. Η συλλογή δεδομένων πραγματοποιήθηκε μέσω crowd-sourcing, όπου οι συμμετέχοντες κατέγραφαν φωνητικά δείγματα και παρείχαν πληροφορίες για την υγεία τους. Το Smarty4covid συντονίστηκε από το εργαστήριο BIOSIM της Εθνικής Τεχνικής Πανεπιστημίου Αθηνών, με τη συμμετοχή κλινικών και ερευνητικών εταίρων, και αποτελεί σημαντική εξέλιξη στην εφαρμογή της τεχνητής νοημοσύνης στη διάγνωση και παρακολούθηση ασθενειών.

Η επιστημονική προσέγγιση του έργου επέτρεψε την ανάπτυξη νέων μοντέλων που μπορούν να προβλέψουν τον κίνδυνο μόλυνσης από COVID-19 και να παρακολουθήσουν την εξέλιξη της νόσου σε μη νοσηλευόμενα άτομα. Οι πληροφορίες αυτές συμβάλλουν στην ανακάλυψη νέων βιοδεικτών που έχουν προγνωστική αξία, ενώ παράλληλα υποστηρίζουν την ανάπτυξη διαφανών και ερμηνεύσιμων μοντέλων μηχανικής μάθησης. Έκτοτε, πολλές μελέτες και εργασίες έχουν αντλήσει δεδομένα από την βάση δεδομένων smarty4covid στοχεύοντας στην ανίχνευση και αναγνώριση της νόσου COVID-19 μέσω σύγχρονων μοντέλων μάθησης.

1.3. Σκοπός της παρούσας διπλωματικής εργασίας

Η παρούσα διπλωματική εργασία πραγματεύεται την ανάπτυξη μοντέλων μηχανικής μάθησης για την αναγνώριση Covid-19 μέσω ηχητικών καταγραφών. Αναπτύχθηκε ένα πρόγραμμα εκπαίδευσης ταξινομητών με αρθρωτή δομή στη γλώσσα Python για να ευρεθεί η καλύτερη συνδυαστική ρύθμιση υπερπαραμέτρων εξαγωγής χαρακτηριστικών για τους τύπους ταξινομητών Logistic Regression, K-Nearest Neighbors και Multi-layer Perceptron. Για την εκπαίδευση των ταξινομητών χρησιμοποιήθηκε η βάση δεδομένων – εφαρμογή SmartyforCovid και συγκεκριμένα χρησιμοποιήθηκε το αρχείο ήχου βήχα “audio.cough.mp3” κάθε δείγματος και το αρχείο csv “smarty4covid_tabular_data.csv”, καθώς και η βάση δεδομένων Coswara, συγκεκριμένα το αρχείο ήχου βήχα με όνομα “cough-heavy.wav” κάθε δείγματος και το αρχείο csv “combined_data.csv”. Ο στόχος ήταν οι ταξινομητές να προβλέψουν αν ένα άτομο πάσχει από covid με βάση το υποβληθέν δείγμα ήχου βήχα τους.

1.4. Δομή της Εργασίας

Παρουσιάζεται η δομή της διπλωματικής εργασίας με συνοπτική αναφορά στο περιεχόμενο των κεφαλαίων της. Η εργασία μπορεί να χωριστεί σε τρία μέρη.

Στο πρώτο μέρος (Κεφάλαιο 2) μετά την εισαγωγή, παρουσιάζεται το απαραίτητο θεωρητικό υπόβαθρο σχετικά με την νόσο Covid-19, την μηχανική και την βαθιά μάθηση και την ανάλυση αρχείων ήχου.

Στο δεύτερο μέρος (Κεφάλαιο 3) παρουσιάζεται η μεθοδολογία που αναπτύχθηκε για την εκπαίδευση των ταξινομητών, τα project Smarty4covid & Coswara, οι μέθοδοι εξαγωγής χαρακτηριστικών, η εκπαίδευση και η αξιολόγηση ταξινομητών.

Αναδεικνύονται τεχνικές και μέθοδοι όπως:

- Οι μέθοδοι εξαγωγής χαρακτηριστικών MFCCs, Spectral Centroid, Spectral Roll-off και Zero-Crossing Rate (ZCR).
- Οι υπερπαραμέτροι "αριθμός των MFCCs" (n_mfcc), "μέγεθος του πλαισίου" (frame_size) και "αριθμός των τμημάτων" (n_segments).
- Οι μέθοδοι ταξινόμησης Logistic Regression, k-Nearest Neighbors (k-NN) και Multi-layer Perceptron (MLP).
- Οι μέθοδοι αξιολόγησης των ταξινομητών Grid Search 5-fold Cross Validation & ROC AUC Curve.

Στο τρίτο μέρος (Κεφάλαια 4-5), παρουσιάζονται τα αποτελέσματα που εξήχθησαν από το πρόγραμμα, τα συμπεράσματα της διπλωματικής, προτάσεις και προβληματισμοί και τονίζεται η ανάγκη συνέχισης παρόμοιων προσπαθειών με εμπλουτισμό στοιχείων, στοχεύοντας στην εύρεση πιο «έξυπνων» και αποδοτικών μεθόδων μηχανικής μάθησης για την αναγνώριση της Covid-19 μέσω ηχητικών καταγραφών και όχι μόνο.

1.5. Κεφάλαια

Στο **πρώτο** κεφάλαιο - εισαγωγή, παρουσιάζεται ο σκοπός και η δομή της εργασίας.

Στο **δεύτερο** κεφάλαιο, ορίζονται και περιγράφονται συνοπτικά οι έννοιες της νόσου Covid-19, της μηχανικής και της βαθιάς μάθησης και της ανάλυσης αρχείων ήχου.

Στο **τρίτο** κεφάλαιο, περιγράφονται τα projects Smarty4covid & Coswara και παρουσιάζεται η μεθοδολογία που αναπτύχθηκε για την εκπαίδευση των ταξινομητών. Παρουσιάζονται με εκτενέστερη ανάλυση οι μέθοδοι εξαγωγής χαρακτηριστικών όπως η μέθοδος MFCCs, η μέθοδος Spectral Centroid, η μέθοδος Spectral Roll-off και η μέθοδος Zero-Crossing Rate (ZCR). Παρουσιάζονται οι υπερπαράμετροι εξαγωγής χαρακτηριστικών όπως η υπερπαράμετρος "αριθμός των MFCCs" (`n_mfcc`), η υπερπαράμετρος "μέγεθος του πλαισίου" (`frame_size`) και η υπερπαράμετρος "αριθμός των τμημάτων" (`n_segments`).

Στο κεφάλαιο αυτό γίνεται επίσης αναφορά στην προεπεξεργασία των δεδομένων, αναλύεται ο κλιμακωτός min-max και γίνεται αναφορά στην μέθοδο της ανάλυσης κύριων συνιστωσών PCA. Επιπλέον, παρουσιάζεται η έννοια της μεθόδου Ensemble Learning και γίνεται κατηγοριοποίηση της μηχανικής μάθησης στις τρεις βασικές κατηγορίες, της επιβλεπόμενης μάθησης (Supervised Learning), της μη επιβλεπόμενης μάθησης (Unsupervised Learning) και της ενισχυτικής μάθησης (Reinforcement Learning).

Στα πλαίσια της Επιβλεπόμενης Μάθησης παρουσιάζονται οι μέθοδοι ταξινόμησης Logistic Regression, K-Nearest Neighbors (K-NN) και Multi-layer Perceptron (MLP) και γίνεται αναφορά στα τεχνητά νευρωνικά δίκτυα.

Κλείνοντας το κεφάλαιο αυτό γίνεται τόσο αναφορά στον έλεγχο εκπαίδευσης ταξινομητών και στις μεθόδους ελέγχου του, οι οποίες είναι η μέθοδος Grid Search 5-fold Cross Validation και η μέθοδος ROC AUC Curve, όσο και στους δείκτες αξιολόγησης Specificity, Sensitivity, Precision, Accuracy & F1 Score.

Στο **τέταρτο** κεφάλαιο παρουσιάζονται τα αποτελέσματα που εξήχθησαν από το πρόγραμμα.

Τέλος, στο **πέμπτο** κεφάλαιο, παρατίθενται τα συμπεράσματα, προτάσεις και προβληματισμοί και τονίζεται η ανάγκη συνέχισης παρόμοιων προσπαθειών με εμπλουτισμό στοιχείων, στοχεύοντας στην εύρεση πιο «έξυπνων» και αποδοτικών μεθόδων μηχανικής μάθησης για την αναγνώριση της Covid-19 μέσω ηχητικών καταγραφών και όχι μόνο.

2. Θεωρητικό μέρος

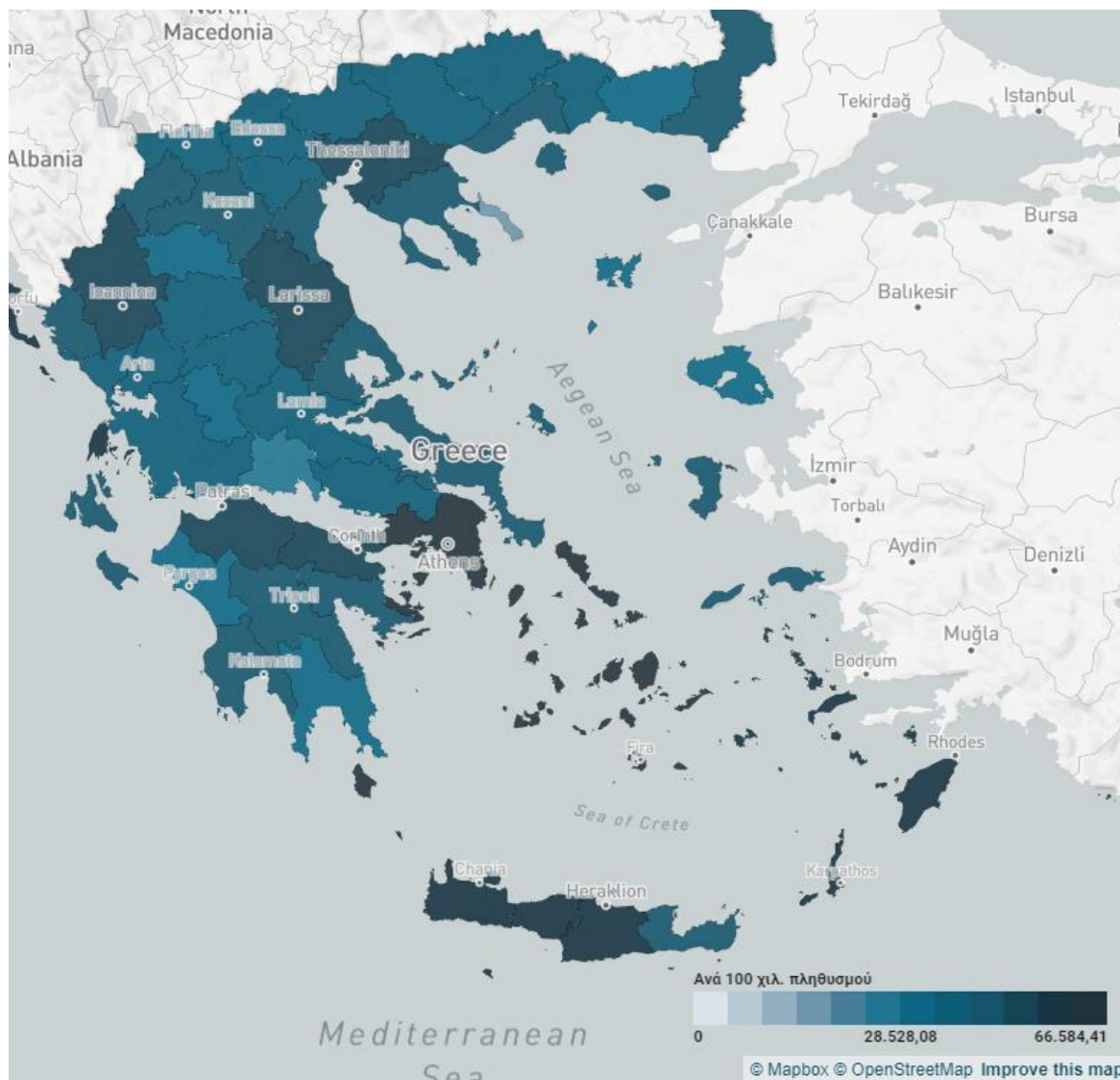
2.1. Covid-19

Η νόσος COVID-19 είναι μια μολυσματική ασθένεια που προκαλείται από τον ιό SARS-CoV-2, έναν νέο κορονοϊό που ανακαλύφθηκε τον Δεκέμβριο του 2019 στην πόλη Wuhan της Κίνας. Ο ιός ανήκει στην οικογένεια των κορονοϊών, που περιλαμβάνει και άλλους παθογόνους ιούς όπως ο SARS-CoV (υπεύθυνος για την επιδημία του SARS το 2003) και ο MERS-CoV (υπεύθυνος για την επιδημία του MERS το 2012). Η νόσος COVID-19 χαρακτηρίζεται από ένα ευρύ φάσμα συμπτωμάτων, που περιλαμβάνει πυρετό, βήχα, δύσπνοια, καθώς και πιο σοβαρά συμπτώματα όπως πνευμονία και αναπνευστική ανεπάρκεια. Σε σοβαρές περιπτώσεις, μπορεί να οδηγήσει σε θάνατο, ιδιαίτερα σε ηλικιωμένους και σε άτομα με υποκείμενα νοσήματα.

Η πρώτη αναφορά της νόσου έγινε στα τέλη Δεκεμβρίου του 2019 [3], όταν μια σειρά από περιπτώσεις πνευμονίας με άγνωστη αιτία αναφέρθηκαν στην πόλη Wuhan, στην επαρχία Hubei της Κίνας. Στις 7 Ιανουαρίου 2020, οι Κινέζοι επιστήμονες απομόνωσαν τον νέο κορονοϊό [4] και στις 11 Ιανουαρίου 2020 καταγράφηκε ο πρώτος θάνατος από τη νόσο.

Στις 30 Ιανουαρίου 2020, ο Παγκόσμιος Οργανισμός Υγείας (ΠΟΥ) κήρυξε την έξαρση της COVID-19 ως έκτακτη ανάγκη δημόσιας υγείας διεθνούς ενδιαφέροντος, και στις 11 Μαρτίου 2020 την χαρακτήρισε ως πανδημία [5]. Οι προσπάθειες για τον περιορισμό της εξάπλωσης του ιού περιλάμβαναν αυστηρά μέτρα κοινωνικής αποστασιοποίησης, καραντίνα, και περιορισμούς στη μετακίνηση.

Η επιστημονική κοινότητα ανταποκρίθηκε γρήγορα, με την ανάπτυξη εμβολίων να γίνεται σε πρωτοφανείς χρόνους. Τα πρώτα εμβόλια κατά της νόσου COVID-19 εγκρίθηκαν [6] για χρήση έκτακτης ανάγκης στα τέλη του 2020, και οι εκστρατείες εμβολιασμού ξεκίνησαν παγκοσμίως στις αρχές του 2021. Παρά τις δυσκολίες και τις προκλήσεις, τα εμβόλια αποδείχθηκαν αποτελεσματικά στη μείωση της σοβαρότητας της νόσου και στη μείωση των νοσηλειών και των θανάτων [7].

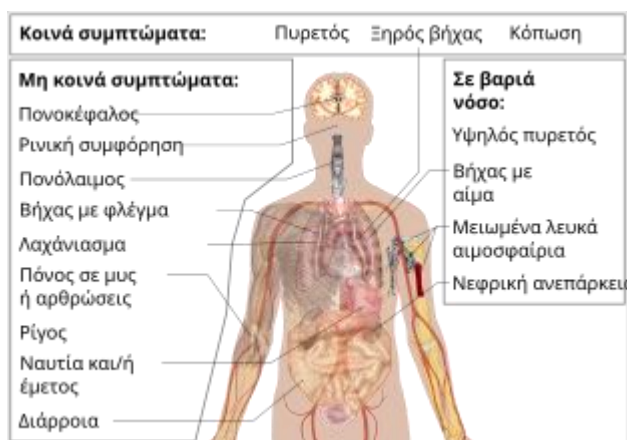


Εικόνα 1: Χάρτης Κρουσμάτων ανά Περιφερειακή Ενότητα στην Ελλάδα από τις 26 Φεβρουαρίου 2020 έως και τις 20 Δεκεμβρίου 2022[8]

Ως την 22 Ιουλίου 2022, περισσότερα από 567.000.000 κρούσματα έχουν επιβεβαιωθεί σε όλο τον κόσμο[9]. Θεωρείται ότι υπάρχει μια ελλιπής αναφορά των κρουσμάτων, ιδιαίτερα σε περιπτώσεις με ηπιότερα συμπτώματα.

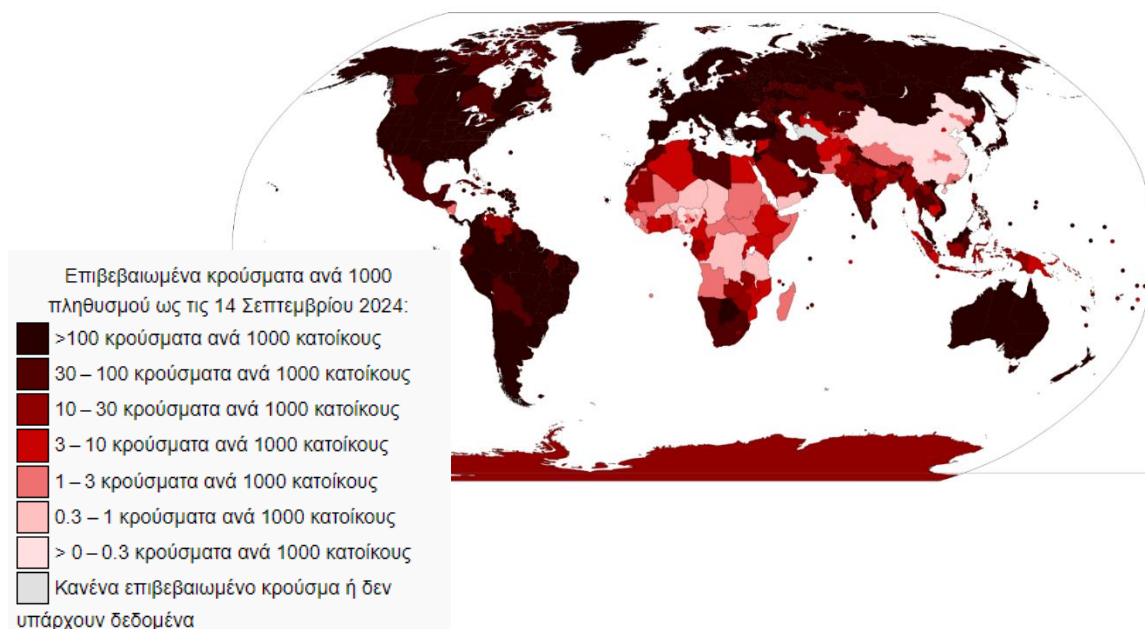
Η επιδημιολογική ανάλυση της εκδήλωσης έδειξε ένα πιθανό μοτίβο μιας «μικτής εστίας» - πιθανότατα υπήρχε μια συνεχής κοινή εστία στην Αγορά Θαλασσιών Χουανάν το Δεκέμβριο του 2019, ενδεχομένως από διάφορα ζωνοσογόνα γεγονότα. Μετά από αυτό, οι επιδημιολόγοι διαπίστωσαν ότι το ξέσπασμα πιθανότατα μετατράπηκε σε πηγή (που μεταδίδεται από άτομο σε άτομο), πιθανώς λόγω της ικανότητας μετάλλαξης του ιού. Κατά τα πρώτα στάδια, ο αριθμός των περιστατικών διπλασιαζόταν συνήθως κάθε 7,5 ημέρες[10]. Στις αρχές και στα μέσα Ιανουαρίου του 2020, ο ιός εξαπλώθηκε σε άλλες κινεζικές επαρχίες, με τη βοήθεια των μετακινήσεων της Κινεζικής Πρωτοχρονιάς. Στις 20 Ιανουαρίου, η Κίνα ανέφερε σχεδόν 140 νέους ασθενείς την ημέρα, συμπεριλαμβανομένων

δύο ατόμων στο Πεκίνο και ενός στη Σενζέν[11]. Μεταγενέστερα επίσημα στοιχεία δείχνουν ότι 6.174 ασθενείς που έχουν μολυνθεί από SARS-CoV-2 είχαν ήδη αναπτύξει συμπτώματα έως τις 20 Ιανουαρίου 2020[12].



Εικόνα 2: Συμπτώματα της COVID-19[13]

Στις 26 Φεβρουαρίου 2020, ο ΠΟΥ ανέφερε ότι, καθώς οι νέες περιπτώσεις που αναφέρθηκαν μειώθηκαν στην Κίνα, αλλά ξαφνικά αυξήθηκαν στην Ιταλία, το Ιράν και τη Νότια Κορέα, ο αριθμός νέων περιπτώσεων εκτός Κίνας ξεπέρασε τον αριθμό νέων περιπτώσεων στην Κίνα για πρώτη φορά στις 25 Φεβρουαρίου 2020. Τέλος, ο ιός εξαπλώθηκε γρήγορα και στις υπόλοιπες χώρες και περιοχές του κόσμου.



Εικόνα 3: Κατά κεφαλήν κρούσματα παγκοσμίως[14]

Οι αλλαγές που προκάλεσε η εξάπλωση του ιού βρίσκονται ακόμα και τώρα υπό μελέτη [15], και εργαλεία όπως η βαθιά μάθηση παρατηρήθηκε ότι χρησιμοποιήθηκαν εκτενώς για τη μελέτη του ιού και των αλλαγών αυτών.

2.2. Μηχανική μάθηση

Η μηχανική μάθηση (Machine Learning) είναι ένας κλάδος της τεχνητής νοημοσύνης (AI) που επικεντρώνεται στην ανάπτυξη αλγορίθμων και μοντέλων που επιτρέπουν στους υπολογιστές να μαθαίνουν από δεδομένα και να λαμβάνουν αποφάσεις ή προβλέψεις χωρίς να είναι ρητά προγραμματισμένοι για κάθε συγκεκριμένο έργο. Η μηχανική μάθηση χρησιμοποιεί στατιστικές μεθόδους για να αναλύσει μεγάλα σύνολα δεδομένων, αναγνωρίζοντας πρότυπα και σχέσεις σε αυτά τα δεδομένα, προσαρμόζοντας συνεχώς το μοντέλο της βάσει της εμπειρίας που αποκτά μέσω των αποτελεσμάτων που εξάγει και του βαθμού επιτυχίας τους.

Οι κύριες κατηγορίες της μηχανικής μάθησης περιλαμβάνουν τη **μάθηση με επίβλεψη** (supervised learning), όπου το μοντέλο εκπαιδεύεται με δεδομένα που περιλαμβάνουν και τις σωστές απαντήσεις, τη **μάθηση χωρίς επίβλεψη** (unsupervised learning), όπου το μοντέλο προσπαθεί να βρει κρυφά πρότυπα σε μη ετικετοποιημένα δεδομένα, και τη **μάθηση ενισχυτικού παράγοντα** (reinforcement learning), όπου το μοντέλο μαθαίνει μέσα από trial and error προκειμένου να μεγιστοποιήσει μια ανταμοιβή.

Η ιδέα της μηχανικής μάθησης χρονολογείται από τα πρώτα στάδια της τεχνητής νοημοσύνης, στις δεκαετίες του 1950 και του 1960. Ο Alan Turing [16], το 1950, πρότεινε την έννοια της "μηχανής που μαθαίνει", και αυτό το όραμα υλοποιήθηκε στα πρώτα προγράμματα που χρησιμοποιούσαν κανόνες για την προσομοίωση της ανθρώπινης νοημοσύνης. Το 1957, ο Frank Rosenblatt [17] ανέπτυξε το πρώτο μοντέλο τεχνητού νευρωνικού δικτύου, το perceptron, το οποίο αποτέλεσε ένα από τα πρώτα βήματα προς τη μάθηση από δεδομένα.

Κατά τη διάρκεια των δεκαετιών του 1970 και 1980, η μηχανική μάθηση επηρεάστηκε από την εξέλιξη των στατιστικών και των θεωριών της πληροφορίας και άρχισαν να εφαρμόζονται ευρέως μοντέλα όπως τα μοντέλα k-πλησιέστερων γειτόνων (k-nearest neighbors) και των δέντρων απόφασης (decision trees).

Στη δεκαετία του 1990, με την αύξηση της υπολογιστικής ισχύος και την πρόσβαση σε μεγαλύτερα σύνολα δεδομένων, η μηχανική μάθηση γνώρισε μια σημαντική αναζωπύρωση. Μοντέλα όπως τα βαθιά νευρωνικά δίκτυα (deep neural networks) ανέδειξαν τη δύναμη της μηχανικής μάθησης σε πολύπλοκα προβλήματα, όπως η αναγνώριση εικόνων και η επεξεργασία φυσικής γλώσσας [18], [19].

Από το 2010 και έπειτα, η μηχανική μάθηση έχει γίνει ο κεντρικός άξονας της τεχνητής νοημοσύνης, με εφαρμογές που περιλαμβάνουν την ανάλυση μεγάλων δεδομένων, τα

αυτόνομα οχήματα, τις συστάσεις προϊόντων και τη διάγνωση ασθενειών. Η βαθιά μάθηση, μια υποκατηγορία της μηχανικής μάθησης που χρησιμοποιεί πολυεπίπεδα νευρωνικά δίκτυα, έχει πρωτοπορήσει στον τομέα, με επιτεύγματα όπως η αναγνώριση φωνής και η μηχανική μετάφραση, με ακρίβεια που πλησιάζει ή και ξεπερνά την ανθρώπινη απόδοση [20], [21].

2.3. Βαθιά μάθηση

Η βαθιά μάθηση (Deep Learning) είναι ένας εξειδικευμένος κλάδος της μηχανικής μάθησης που χρησιμοποιεί τεχνητά νευρωνικά δίκτυα πολλαπλών επιπέδων (deep neural networks) για την ανάλυση δεδομένων και την εκτέλεση σύνθετων υπολογιστικών εργασιών. Ο όρος βαθιά μάθηση, αναφέρεται στον μεγάλο αριθμό επιπέδων μάθησης με πολλά από αυτά τα επίπεδα κρυφά ανάμεσα στα δεδομένα εισόδου (εκπαίδευσης) και στα αποτελέσματα εξόδου εμπνευσμένη από τη δομή ενός ανθρώπινου εγκεφάλου. Τα μοντέλα βαθιάς μάθησης τείνουν να αυξάνουν την ακρίβειά τους, καθώς συνεχίζουν να βελτιώνονται με την αύξηση του μεγέθους των δεδομένων εκπαίδευσης όπως και ο ανθρώπινος εγκέφαλος. Οι πρόσφατες εξελίξεις στη βαθιά μάθηση έχουν βελτιωθεί σε σημείο που η βαθιά μάθηση έχει επιτύχει εντυπωσιακά αποτελέσματα σε τομείς όπως η αναγνώριση εικόνας, η επεξεργασία φυσικής γλώσσας, η ομιλία, και τα παιχνίδια στρατηγικής.

Οι ρίζες της βαθιάς μάθησης εντοπίζονται στις πρώτες μελέτες της τεχνητής νοημοσύνης και των νευρωνικών δικτύων τη δεκαετία του 1940 και του 1950. Το 1943, οι Warren McCulloch και Walter Pitts [22] πρότειναν το πρώτο μαθηματικό μοντέλο νευρώνα, το οποίο ενέπνευσε μελλοντικές έρευνες στην τεχνητή νοημοσύνη. Το 1958, ο Frank Rosenblatt ανέπτυξε το perceptron, ένα από τα πρώτα και πιο θεμελιώδη μοντέλα τεχνητών νευρωνικών δικτύων.

Ωστόσο, τα νευρωνικά δίκτυα αντιμετώπισαν περιορισμούς στη δεκαετία του 1970 και 1980, κυρίως λόγω του προβλήματος της εκπαίδευσης των βαθύτερων δικτύων. Η μέθοδος backpropagation, που προτάθηκε το 1986 από τους Rumelhart, Hinton & Williams [23], έδωσε λύση σε αυτό το πρόβλημα, επιτρέποντας την αποτελεσματική εκπαίδευση πολυεπίπεδων δικτύων. Παρόλα αυτά, τα βαθιά δίκτυα δεν είχαν μεγάλη επιτυχία εκείνη την εποχή λόγω περιορισμών στην υπολογιστική ισχύ και στην πρόσβαση σε μεγάλα σύνολα δεδομένων. Η κατάσταση άλλαξε στις αρχές της δεκαετίας του 2000, με την αύξηση της υπολογιστικής ισχύος και τη διαθεσιμότητα τεράστιων συνόλων δεδομένων. Το 2006, ο Geoffrey Hinton [24], ένας από τους πρωτοπόρους της βαθιάς μάθησης, εισήγαγε την έννοια της προεκπαίδευσης (pre-training) των βαθιών δικτύων, η οποία βελτίωσε σημαντικά τις επιδόσεις τους.

Η βαθιά μάθηση άρχισε να απογειώνεται το 2012, όταν μια ομάδα από το Πανεπιστήμιο του Τορόντο, με επικεφαλής τον Hinton, κέρδισε τον διαγωνισμό αναγνώρισης εικόνας ImageNet χρησιμοποιώντας ένα βαθύ συνελκτικό νευρωνικό δίκτυο [25]. Αυτή η επιτυχία ανέδειξε τη δυναμική της βαθιάς μάθησης και πυροδότησε την ταχεία υιοθέτηση της σε διάφορους τομείς. Σήμερα, η βαθιά μάθηση βρίσκεται στο επίκεντρο της τεχνητής νοημοσύνης, με ευρεία εφαρμογή στην αναγνώριση φωνής, στην επεξεργασία φυσικής γλώσσας, στην αυτόνομη οδήγηση, και σε πολλά άλλα πεδία.

2.4. Διαφορές Τεχνητής Νοημοσύνης - Μηχανικής Μάθησης - Βαθιάς Μάθησης

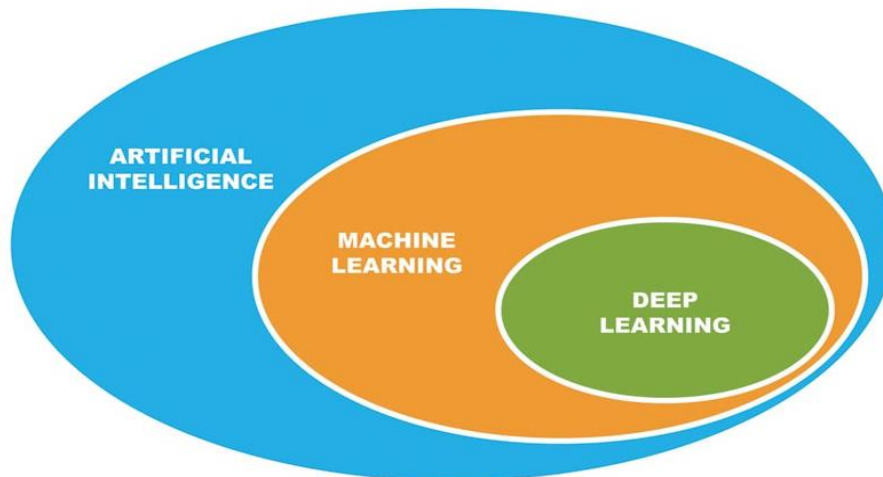
Η **Τεχνητή Νοημοσύνη - Artificial Intelligence (AI)** είναι ευρύτερος όρος και αναφέρεται στην ανάπτυξη συστημάτων που μπορούν να εκτελούν καθήκοντα που παραδοσιακά απαιτούν ανθρώπινη νοημοσύνη. Αυτά τα καθήκοντα μπορεί να περιλαμβάνουν την επίλυση προβλημάτων, τη λήψη αποφάσεων, την κατανόηση της φυσικής γλώσσας καθώς και τη μάθηση από δεδομένα.

Η **Μηχανική Μάθηση – Machine Learning (ML)**, ως υποσύνολο της Τεχνητής Νοημοσύνης, επικεντρώνεται στην ικανότητα των συστημάτων να μαθαίνουν από δεδομένα και να βελτιώνουν τις επιδόσεις τους σε συγκεκριμένες εργασίες χωρίς ρητές εντολές προγραμματισμού. Στην ουσία, χρησιμοποιεί αλγόριθμους που βασίζονται σε στατιστικά μοντέλα και προσαρμόζεται αυτόματα μέσω εκπαίδευσης σε δεδομένα.

Η **Βαθιά Μάθηση – Deep Learning (DL)**, ως υποσύνολο και αυτή της Τεχνητής Νοημοσύνης, είναι ένα υποσύνολο της Μηχανικής Μάθησης και επικεντρώνεται σε αλγόριθμους που εμπνέονται από τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου. Χρησιμοποιεί νευρωνικά δίκτυα με πολλές στρώσεις (layers) για την ανάλυση και την επεξεργασία μεγάλων ποσοτήτων δεδομένων. Τα δίκτυα αυτά, γνωστά ως βαθιά νευρωνικά δίκτυα, είναι ιδιαίτερα αποτελεσματικά σε πολύπλοκα προβλήματα όπως η αναγνώριση εικόνων, η ανάλυση βίντεο και η επεξεργασία φυσικής γλώσσας (NLP). Απαιτεί μεγάλες ποσότητες δεδομένων και υπολογιστική ισχύ, αλλά έχει τη δυνατότητα να επιλύει εξαιρετικά πολύπλοκα προβλήματα.

Συμπερασματικά, AI είναι η γενική έννοια της επιστήμης ανάπτυξης τέτοιων συστημάτων, **ML** είναι μια προσέγγιση για την υλοποίηση της Τεχνητής Νοημοσύνης μέσω της μάθησης από δεδομένα και **DL** είναι μια πιο εξειδικευμένη μορφή Μηχανικής Μάθησης που

χρησιμοποιεί βαθιά νευρωνικά δίκτυα για να αναλύει και να μαθαίνει από μεγάλα σύνολα δεδομένων με μεγαλύτερη ακρίβεια. Ουσιαστικά, η βαθιά μάθηση αποτελεί έναν τρόπο να επιτευχθεί η μηχανική μάθηση και η μηχανική μάθηση είναι ένας από τους κύριους τρόπους για την υλοποίηση της Τεχνητής Νοημοσύνης.



Εικόνα 4: AI vs ML vs DL [26]

2.5. Εφαρμογές μηχανικής μάθησης και βαθιάς μάθησης κατά την διάρκεια της πανδημίας COVID-19

Η **μηχανική μάθηση (ML)** και η **βαθιά μάθηση (DL)**, στη διάρκεια της πανδημίας COVID-19, έπαιξαν κρίσιμο ρόλο στην αντιμετώπιση προκλήσεων σε διάφορους τομείς με πολυδιάστατο τρόπο.

Υλοποιήθηκαν μοντέλα μάθησης τα οποία μπορούσαν να εκτιμήσουν τον ρυθμό εξάπλωσης και να εντοπίσουν περιοχές που θα αντιμετώπιζαν έξαρση. Έγινε **ανάλυση** μεγάλων ποσοτήτων **δεδομένων** από συστήματα υγείας, κοινωνικά δίκτυα και εφαρμογές ιχνηλάτησης και επεξεργάστηκαν με αλγορίθμους μηχανικής μάθησης για την πρόβλεψη της εξάπλωσης του ιού [27]. Η μηχανική μάθηση και η βαθιά μάθηση συνέβαλαν στην βελτίωση των **επιδημιολογικών μοντέλων** (π.χ. SIR models) με σκοπό την πρόβλεψη της πορείας της πανδημίας και την ανάλυση των επιπτώσεων από διάφορα μέτρα όπως lockdowns [28].

Η βαθιά μάθηση χρησιμοποιήθηκε ευρέως στην **επεξεργασία ιατρικών εικόνων** για τη διάγνωση της COVID-19 μέσω αξονικών τομογραφιών και ακτινογραφιών θώρακα [29], [30]. Τα μοντέλα που δημιουργήθηκαν μπορούσαν να αναγνωρίσουν σημάδια της νόσου και να βοηθήσουν τους γιατρούς στη γρήγορη και ακριβή διάγνωση, ειδικά όταν οι διαγνωστικές δοκιμές ήταν περιορισμένες.

Η μηχανική μάθηση επιτάχυνε τη διαδικασία **ανακάλυψης φαρμάκων** και τη δημιουργία **νέων εμβολίων** [27], [29]. Με την ανάλυση μεγάλων δεδομένων για πρωτεΐνες, γονιδιώματα και χημικές ενώσεις, τα μοντέλα μηχανικής μάθησης βοήθησαν τους επιστήμονες να εντοπίσουν ταχύτερα υποψήφια φάρμακα που θα μπορούσαν να χρησιμοποιηθούν για την αντιμετώπιση του ιού. Η τεχνολογία mRNA, που αξιοποιήθηκε στα εμβόλια της Pfizer και της Moderna, χρησιμοποιεί αλγόριθμους μηχανικής μάθησης για τη βελτιστοποίηση της παραγωγής και της σύνθεσης του mRNA [31].

Εφαρμογές **ιχνηλάτησης επαφών** βασισμένες στη μηχανική μάθηση ανέλυσαν την κίνηση των ατόμων και εντόπιζαν πιθανά κρούσματα με βάση τις επαφές και τις αλληλεπιδράσεις, συμβάλλοντας στον περιορισμό της εξάπλωσης. Αλγόριθμοι ML χρησιμοποιήθηκαν επίσης για την **παρακολούθηση συμπτωμάτων** μέσω εφαρμογών και διαδικτυακών πλατφορμών, εντοπίζοντας συσχετίσεις ανάμεσα σε συγκεκριμένα συμπτώματα και σοβαρότερες μορφές της νόσου.

Η βαθιά μάθηση συνέβαλε επίσης στην ανάπτυξη **συστημάτων τηλεϊατρικής** και **chatbots υγείας** που βοηθούσαν στη διάγνωση ασθενών σε απομακρυσμένες περιοχές, χωρίς να απαιτείται η φυσική παρουσία γιατρού. Αυτές οι πλατφόρμες παρείχαν οδηγίες και συμβουλές για τη διαχείριση των συμπτωμάτων.

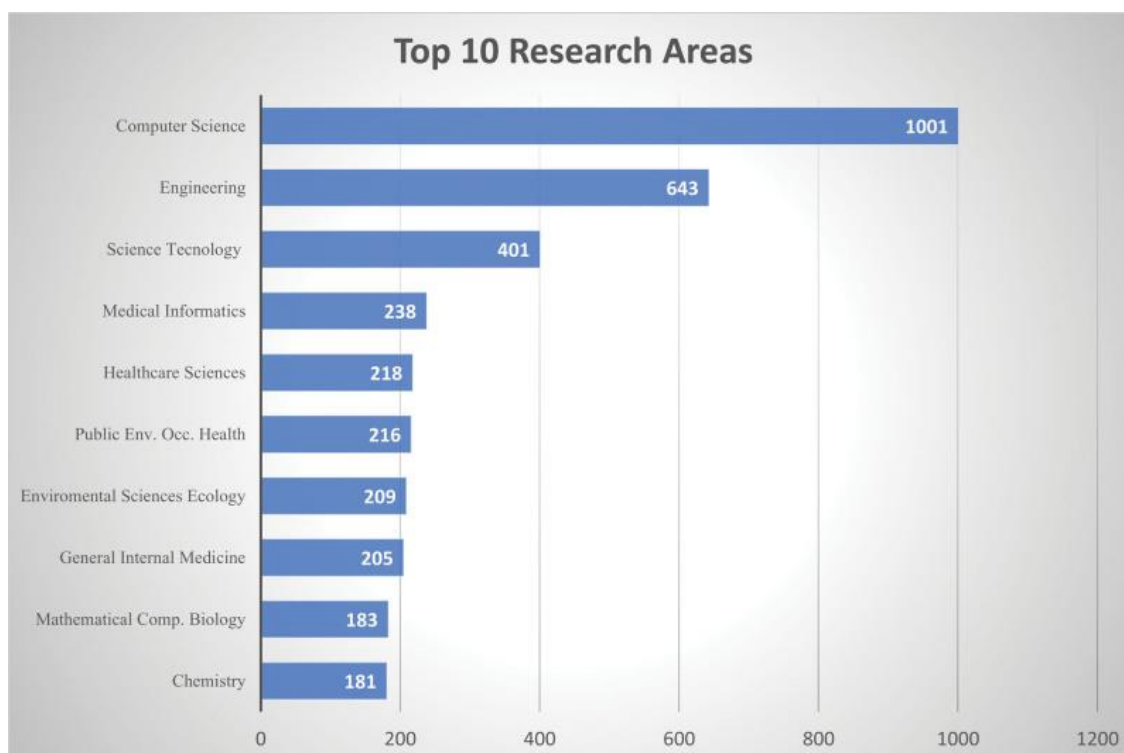
Η μηχανική μάθηση συνέβαλε στην **βελτιστοποίηση της εφοδιαστικής αλυσίδας** για υγειονομικό εξοπλισμό και εμβόλια. Αλγόριθμοι χρησιμοποιήθηκαν για να προβλέψουν τις ανάγκες σε συγκεκριμένες περιοχές και να διαχειριστούν την παραγωγή και διανομή εμβολίων και υλικών.

Τέλος, τα μοντέλα βαθιάς μάθησης αξιοποίησαν δεδομένα από τα **social media** για να αναλύσουν τις συμπεριφορές των πολιτών, τα επίπεδα άγχους και τις αντιδράσεις απέναντι στα μέτρα δημόσιας υγείας. Με αυτόν τον τρόπο, οι κυβερνήσεις μπορούσαν να προσαρμόσουν τις επικοινωνιακές στρατηγικές τους.

Συνολικά, η πανδημία COVID-19 αποτέλεσε μια μοναδική ευκαιρία για την εφαρμογή της μηχανικής και βαθιάς μάθησης σε πραγματικά σενάρια κρίσης, προωθώντας την ανάπτυξη και τη χρήση αυτών των τεχνολογιών στον τομέα της υγείας[32].

Επιστημονικές δημοσιεύσεις χρήσης μηχανικής και βαθιάς μάθησης κατά την διάρκεια της πανδημίας COVID-19.

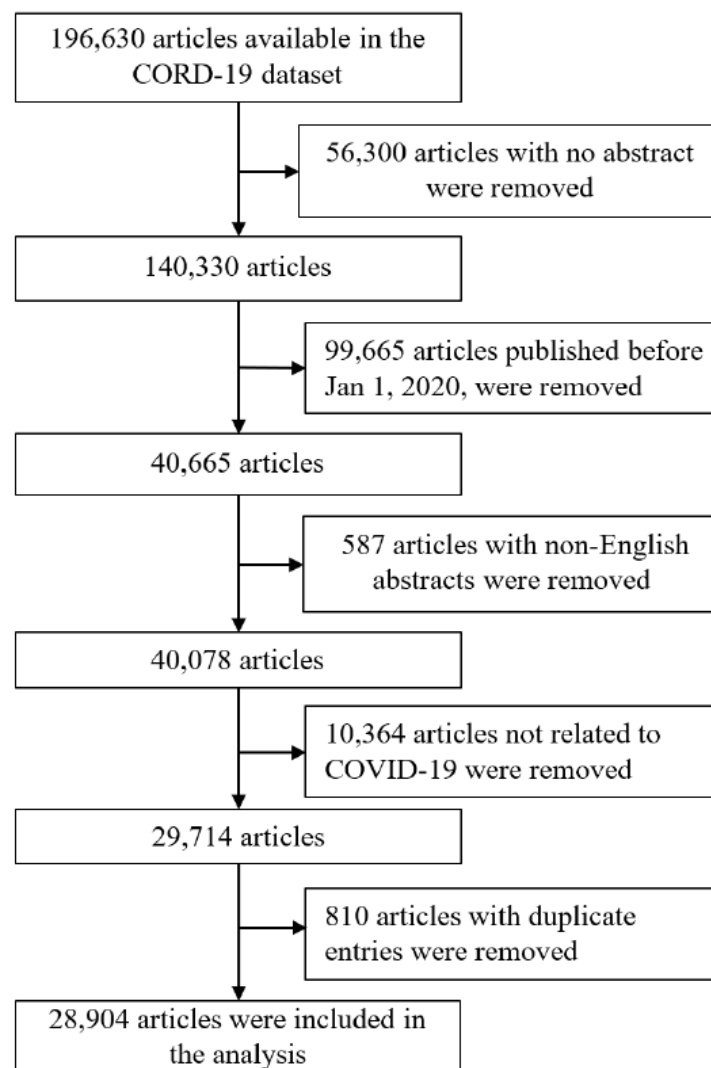
Κατά τη διάρκεια της πανδημίας COVID-19, υπήρξε μια **ραγδαία αύξηση** στις επιστημονικές δημοσιεύσεις που αφορούσαν τη χρήση της **μηχανικής μάθησης (ML)** και της **βαθιάς μάθησης (DL)** για την αντιμετώπιση της πανδημίας. Η αύξηση ήταν το αποτέλεσμα της επείγουσας ανάγκης για λύσεις και της δυνατότητας των τεχνολογιών ML/DL να συνεισφέρουν σε κρίσιμους τομείς, όπως η διάγνωση, η πρόβλεψη, και η ανάπτυξη φαρμάκων.



Εικόνα 5: Αριθμός δημοσιεύσεων στα 10 μεγαλύτερα πεδία [33]

Υπήρξε ανάγκη για ταχεία διάγνωση και ανάλυση μεγάλων δεδομένων καθώς η πανδημία δημιούργησε μια παγκόσμια κρίση όπου οι παραδοσιακές διαγνωστικές και προβλεπτικές μέθοδοι δεν μπορούσαν να ανταποκριθούν με επαρκή ταχύτητα και ακρίβεια. Τα εργαλεία ML/DL προσέφεραν τη δυνατότητα ανάλυσης μεγάλων ποσοτήτων δεδομένων υγείας και επιδημιολογικών πληροφοριών, επιταχύνοντας τη λήψη αποφάσεων. [34]

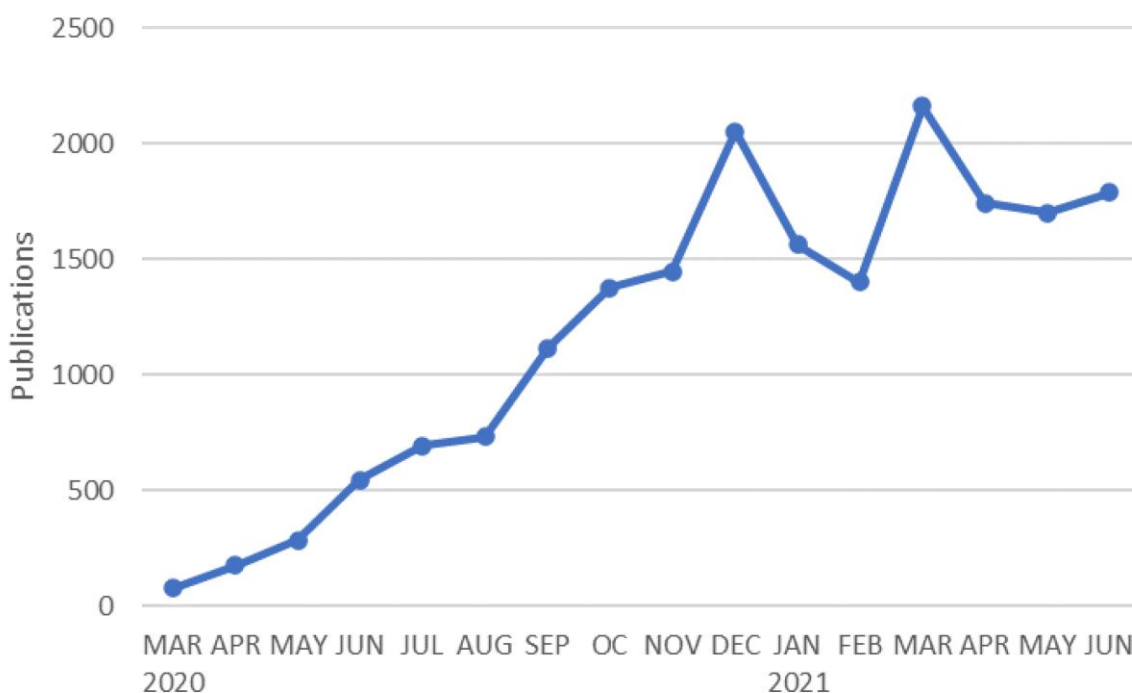
Παρατηρήθηκε εκρηκτική αύξηση των διαθέσιμων δεδομένων καθώς ο ιός εξαπλωνόταν, και οι κυβερνήσεις, τα νοσοκομεία και οι ερευνητικοί οργανισμοί παρήγαγαν τεράστιους όγκους δεδομένων, που περιλάμβαναν κλινικά δεδομένα από ασθενείς, εικόνες από ακτινογραφίες και CT σκαναρίσματα καθώς και γονιδιωματικά δεδομένα του ιού και των μεταλλάξεών του. Η πρόσβαση σε αυτά τα δεδομένα ώθησε την έρευνα για νέα μοντέλα ML που θα μπορούσαν να αναλύσουν τα δεδομένα γρήγορα και αποτελεσματικά. Τα ακαδημαϊκά και τα ερευνητικά κέντρα παγκοσμίως δημοσίευσαν ανοιχτές βάσεις δεδομένων, όπως την COVID-19 Open Research Dataset (CORD-19), που περιέχουν επιστημονικά άρθρα και δεδομένα σχετικά με τον ιό. Αυτές οι βάσεις δεδομένων χρησίμευσαν ως πηγή για πολλές έρευνες και δημοσιεύσεις.



Εικόνα 6: Διάγραμμα ροής της επιλογής των δημοσιευμένων άρθρων στη μελέτη “A Comprehensive Overview of the COVID-19 Literature: Machine Learning–Based Bibliometric Analysis” [35]

Οδηγηθήκαμε νομοτελειακά σε χρήση μοντέλων βαθιάς μάθησης σε ιατρικές εφαρμογές, καθώς η βαθιά μάθηση ήταν ήδη σε άνθηση στον τομέα της ιατρικής ανάλυσης και της διαγνωστικής πριν την πανδημία. Ωστόσο, με την έλευση της COVID-19, οι ερευνητές επιτάχυναν τις προσπάθειες για την εφαρμογή αυτών των τεχνολογιών σε διαγνωστικές διαδικασίες, όπως την ανάλυση ακτινογραφιών και την πρόβλεψη της σοβαρότητας των συμπτωμάτων.

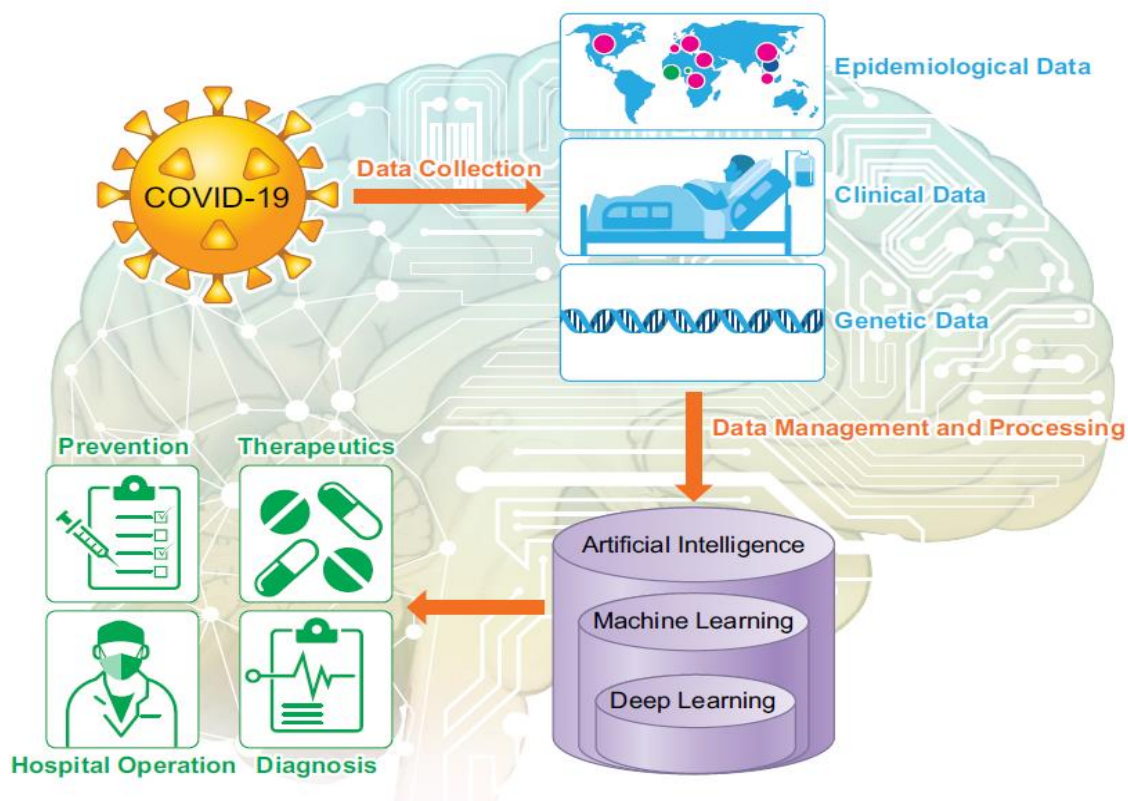
Σύμφωνα με δεδομένα από βάσεις όπως το **PubMed** και το **arXiv**, ο αριθμός των δημοσιεύσεων που αφορούσαν τη χρήση της μηχανικής μάθησης για την COVID-19 αυξήθηκε δραματικά από τις αρχές του 2020. Μέχρι το τέλος του 2020, υπήρχαν χιλιάδες νέες δημοσιεύσεις που επικεντρώνονταν στην εφαρμογή της ML και DL για την ανάλυση και πρόβλεψη της πανδημίας [36], [37]. Παρόλο που η ένταση των νέων δημοσιεύσεων για COVID-19 μειώθηκε σταδιακά από το 2021 καθώς προχωρούσε η διαχείριση της πανδημίας, η έρευνα σε ML και DL συνέχισε να αυξάνεται και να μετατρέπεται σε πιο **εξειδικευμένες μελέτες** που εστίαζαν σε συγκεκριμένα προβλήματα, όπως οι μακροχρόνιες επιπτώσεις της COVID-19 και οι αναλύσεις για τις μεταλλάξεις του ιού [38].



Εικόνα 7: Αριθμός δημοσιεύσεων ανά μήνα (2020-2021) [39]

Η πανδημία προώθησε τη **διατομεακή συνεργασία** μεταξύ ερευνητών από διαφορετικούς τομείς, όπως την βιοπληροφορική, την επιδημιολογία, την πληροφορική και την μηχανική μάθηση. Αυτή η συνεργασία οδήγησε σε αυξημένες δημοσιεύσεις σε περιοδικά τόσο στον τομέα της πληροφορικής όσο και της ιατρικής.

Οι πιο δημοφιλείς κατηγορίες δημοσιεύσεων ML/DL κατά την πανδημία περιλάμβαναν, διάγνωση μέσω εικόνων (π.χ. ακτινογραφίες, αξονικές τομογραφίες), **πρόβλεψη εξάπλωσης** του ιού και **μοντέλα επιδημιολογίας**, **ανακάλυψη φαρμάκων και πρωτεϊνικών δομών** μέσω αλγορίθμων ML και **ανάλυση κοινωνικών δεδομένων** για την κατανόηση της συμπεριφοράς του πληθυσμού και της επιρροής της πανδημίας.



Εικόνα 8: Εφαρμογή της τεχνητής νοημοσύνης και της μηχανικής μάθησης στη μάχη κατά της COVID-19 [31]

Πληθώρα αναφορών δημοσιεύτηκαν σε ιστοτόπους και βάσεις επιστημονικών δεδομένων όπως το CORD-19 Dataset [40] το οποίο παρείχε πρόσβαση σε χιλιάδες ακαδημαϊκά άρθρα σχετικά με την COVID-19 και συνέβαλε σε πολλές νέες δημοσιεύσεις που βασίζονταν σε τεχνολογίες μηχανικής μάθησης, το Google Scholar [41] και το PubMed [42] το οποίο έδειξε συνεχή άνοδο στις δημοσιεύσεις που σχετίζονται με την "COVID-19" και "Machine Learning" καθ' όλη τη διάρκεια της πανδημίας.

2.6. Ανάλυση αρχείων ήχου

Η ανάλυση αρχείων ήχου αναφέρεται στη διαδικασία επεξεργασίας και εξαγωγής πληροφοριών από ψηφιακά ηχητικά σήματα. Αυτή η διαδικασία περιλαμβάνει διάφορες τεχνικές, όπως τη μετατροπή του ηχητικού σήματος από το πεδίο του χρόνου στο πεδίο της

συχνότητας, την ανίχνευση χαρακτηριστικών (feature extraction), την ανάλυση φασματικής πυκνότητας, και την αναγνώριση προτύπων.

Οι βασικές μέθοδοι που χρησιμοποιούνται για την ανάλυση ήχου περιλαμβάνουν τον μετασχηματισμό Fourier (Fourier Transform) [43], που επιτρέπει την ανάλυση της συχνότητας ενός σήματος και τον μετασχηματισμό Short-Time Fourier Transform (STFT), που επιτρέπει την ανάλυση της περιεκτικότητας της συχνότητας ενός σήματος ήχου σε συνάρτηση με την ακρίβεια προσδιορισμού του στο πεδίο του χρόνου. Επιπλέον, η ανάλυση κύματος (wavelet analysis) χρησιμοποιείται για την επεξεργασία σημάτων που περιέχουν στοιχεία τόσο στο πεδίο της συχνότητας όσο και στο πεδίο του χρόνου.

Η ανάλυση αρχείων ήχου έχει ευρείες εφαρμογές, όπως την αναγνώριση φωνής, την αναγνώριση μουσικής και τον αυτόματο εντοπισμό και διάγνωση προβλημάτων σε μηχανές μέσω της ανάλυσης των ήχων που παράγουν.

Η ανάλυση ήχου ξεκίνησε με την ανάπτυξη της ακουστικής και της θεωρίας των σημάτων. Ο Joseph Fourier, στις αρχές του 19ου αιώνα, ανέπτυξε το θεωρητικό πλαίσιο για τον διαχωρισμό ενός σήματος σε συχνότητες με τον μετασχηματισμό Fourier. Αυτή η μέθοδος παραμένει θεμελιώδης στην ανάλυση των ηχητικών σημάτων.

Κατά τη διάρκεια του 20ού αιώνα, με την πρόοδο της τεχνολογίας ηχογράφησης και αναπαραγωγής, η ανάλυση ήχου εξελίχθηκε και έγινε πιο προσιτή. Σημαντική εξέλιξη ήταν η εφεύρεση του ψηφιακού ήχου στη δεκαετία του 1960, που επέτρεψε την αποθήκευση και επεξεργασία του ήχου σε ψηφιακή μορφή. Αυτή η εξέλιξη αποτέλεσε τη βάση για την ανάπτυξη σύγχρονων αλγορίθμων ανάλυσης.

Στη δεκαετία του 1980, οι πρώτοι αλγόριθμοι αναγνώρισης φωνής άρχισαν να αναπτύσσονται, με τις πρώτες εμπορικές εφαρμογές να εμφανίζονται στη δεκαετία του 1990 [44]. Ο αλγόριθμος κρυφών μοντέλων Markov (Hidden Markov Models, HMMs) έγινε ευρέως χρησιμοποιούμενος για την ανάλυση και αναγνώριση φωνής, επιτρέποντας την αντιστοίχιση ομιλίας με γραμματική και λεξιλόγιο [45].

Τα τελευταία χρόνια, η ανάλυση ήχου έχει ενισχυθεί από τις εξελίξεις στη βαθιά μάθηση, με νευρωνικά δίκτυα όπως τα συνελκτικά νευρωνικά δίκτυα (CNNs) και τα επαναλαμβανόμενα νευρωνικά δίκτυα (RNNs) να χρησιμοποιούνται για την αναγνώριση προτύπων ήχου και φωνής με εξαιρετικά αποτελέσματα [21], [46].

3. Μεθοδολογία

Αναπτύχθηκε ένα πρόγραμμα εκπαίδευσης ταξινομητών με αρθρωτή δομή στη γλώσσα Python για να ευρεθεί η καλύτερη συνδυαστική ρύθμιση υπερπαραμέτρων εξαγωγής χαρακτηριστικών για τους τύπους ταξινομητών Logistic Regression, K-Nearest Neighbors και Multi-layer Perceptron. Για την εκπαίδευση των ταξινομητών χρησιμοποιήθηκε η βάση δεδομένων – εφαρμογή SmartyforCovid και συγκεκριμένα χρησιμοποιήθηκε το αρχείο ήχου βήχα “audio.cough.mp3” κάθε δείγματος και το αρχείο csv “smarty4covid_tabular_data.csv”, καθώς και η βάση δεδομένων Coswara, συγκεκριμένα το αρχείο ήχου βήχα με όνομα “cough-heavy.wav” κάθε δείγματος και το αρχείο csv “combined_data.csv”. Ο στόχος ήταν οι ταξινομητές να προβλέψουν αν ένα άτομο πάσχει από covid με βάση το υποβληθέν δείγμα ήχου βήχα τους. Ο κώδικας του προγράμματος που αναπτύχθηκε βρίσκεται στην πλατφόρμα οργάνωσης κώδικα Github.

3.1. Εισαγωγή Δεδομένων

Αρχικά πραγματοποιείται προσπέλαση της τοποθεσίας των ηχητικών δειγμάτων του συνόλου δεδομένων και γίνεται εισαγωγή στο πρόγραμμα όλων των δειγμάτων που έχουν επιβεβαιωμένη ετικέτα covid (είτε η ετικέτα ήταν υγιής είτε μολυσμένη) για να διασφαλιστεί η ακρίβεια και η επαλήθευση των αποτελεσμάτων. Το σύνολο των δειγμάτων σε κάθε επανάληψη του προγράμματος προέρχεται αποκλειστικά από μία μόνο βάση τη φορά.

Smarty4covid

Περιγραφή

Ο στόχος του έργου smarty4covid ήταν η δημιουργία ενός έξυπνου πολυτροπικού πλαισίου για την αξιολόγηση και παρακολούθηση του κινδύνου της νοσού COVID-19 βασισμένου σε Εξηγήσιμη Βαθιά Μάθηση. Μετά από τις απαραίτητες εγκρίσεις από την Επιτροπή Ηθικής της Έρευνας του Εθνικού Μετσόβιου Πολυτεχνείου, μια διαδραστική διαδικτυακή εφαρμογή [2] υλοποιήθηκε και κυκλοφόρησε δημόσια ως μέσο συλλογής δεδομένων.

Το σύνολο δεδομένων smarty4covid περιέχει συνολικά 18.265 ηχητικές εγγραφές βήχα, αναπνοής (κανονικής, βαθιάς) και φωνής που αντιστοιχούν σε 4.673 χρήστες (Ελληνες και Κύπριους πολίτες). Λανθασμένα και θορυβώδη δείγματα αφαιρέθηκαν από το σύνολο δεδομένων και ένα υποσύνολο του (π.χ. 1.475 δείγματα) έχει επισημανθεί από ιατρικούς εμπειρογνώμονες. Επιπλέον, όλες οι διαθέσιμες πληροφορίες έχουν κωδικοποιηθεί σε μια καινοτόμο βάση γνώσεων οντολογίας ιστού (OWL) η οποία προσφέρει ένα ερμηνεύσιμο

πλαίσιο υψηλής εκφραστικότητας που μπορεί να χρησιμοποιηθεί για να εξηγήσει σύνθετα μοντέλα μηχανικής μάθησης μέσω της αναγνώρισης σημασιολογικών ερωτημάτων πάνω στη γνώση που μιμούνται το μοντέλο.

Η γενική προσέγγιση για την ανάπτυξη της βάσης δεδομένων smarty4covid περιλαμβάνει μια στρατηγική συλλογής δεδομένων μέσω crowdsourcing, ακολουθούμενη από μια διαδικασία διόρθωσης δεδομένων σε δύο στάδια, που περιλαμβάνει τον καθαρισμό και την επισημάνση των δεδομένων.

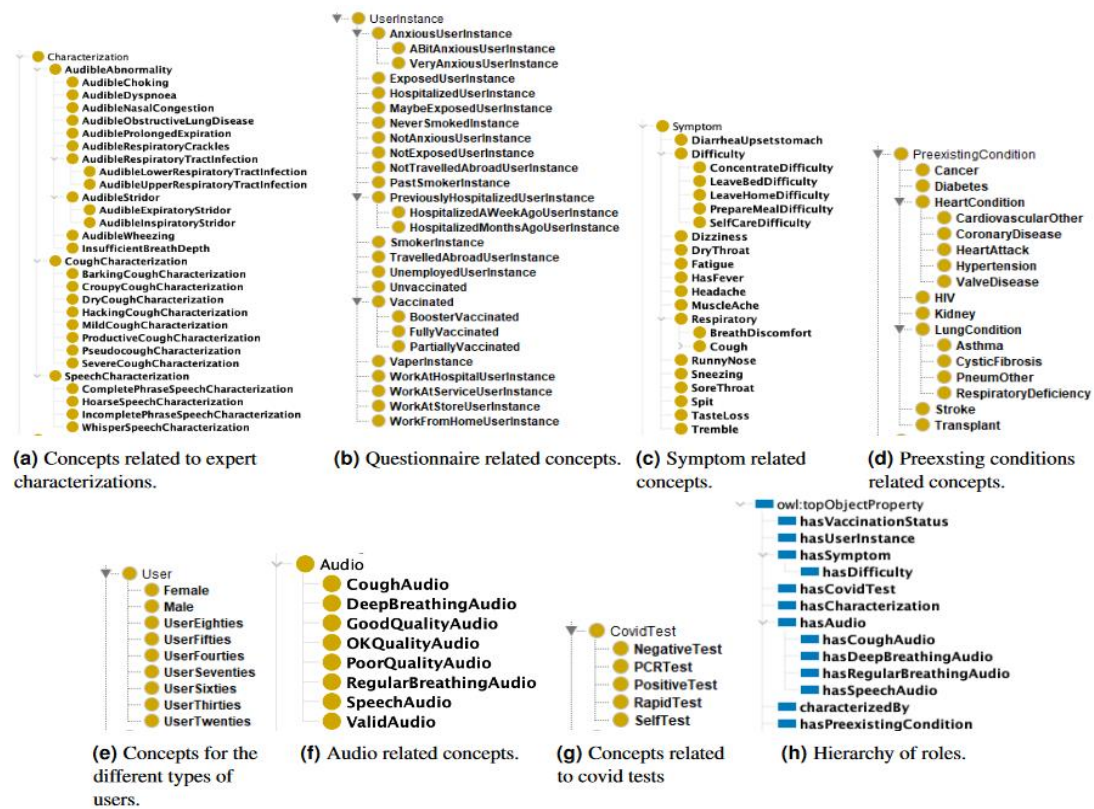
Η ευέλικτη και φιλική προς τον χρήστη διαδικτυακή εφαρμογή υλοποιήθηκε στοχεύοντας σε Έλληνες και Κύπριους πολίτες άνω των 18 ετών. Η διαδικτυακή εφαρμογή κυκλοφόρησε τον Ιανουάριο του 2022 κατά τη διάρκεια της εξάπλωσης του κύματος της μετάλλαξης όмикρον στην Ελλάδα, με αποτέλεσμα την υψηλή επικράτηση της νόσου COVID-19 (17,3% των χρηστών βρέθηκαν θετικοί στην νόσο). Η διαδικασία καθαρισμού δεδομένων πραγματοποιήθηκε μέσω μιας καμπάνιας crowdsourcing χρησιμοποιώντας το open-source εργαλείο επισημάνσης δεδομένων Label Studio. [47]

Δεδομένα Αρχείων

Μέρος του συνόλου δεδομένων smarty4covid που συλλέχθηκε μέσω crowdsourcing (4.303 υποβολές) οργανώθηκε σε αρχεία δεδομένων ώστε να είναι δημόσια διαθέσιμα. Τα αρχεία δεδομένων έχουν κατατεθεί στο αποθετήριο Zenodo (DOI: 10.5281/zenodo.7760170). Όπως φαίνεται στο ακόλουθο σχήμα 1 κάθε φάκελος περιέχει τις υποβολές συγκεκριμένου χρήστη. Ο φάκελος του χρήστη ονομάζεται με το id του χρήστη, το οποίο δημιουργείται σύμφωνα με το πρωτόκολλο UUID V4. Εκτός από τις υποβολές, περιλαμβάνεται επίσης ένα αρχείο json ("demographics_underlying_conditions.json") με πληροφορίες σχετικά με τα δημογραφικά στοιχεία (π.χ. ΔΜΣ, ηλικιακή ομάδα, φύλο) και πιθανές υποκείμενες καταστάσεις. Κάθε υποβολή αντιστοιχεί σε έναν ξεχωριστό υποφάκελο που ονομάζεται με το μοναδικό id της υποβολής και περιέχει:

1. Έγκυρες ηχογραφήσεις βήχα ("audio.cough.mp3"), βαθιάς αναπνοής ("audio.breath_deep.mp3") και κανονικής αναπνοής ("audio.breath_regular.mp3"). Κάθε ηχογράφιση έχει ρυθμό δειγματοληψίας 48 kHz και bitrate 64 kb/s.
2. Ένα αρχείο json ("main_questionnaire.json") με πληροφορίες που σχετίζονται με το τεστ για COVID-19 (αποτέλεσμα, τύπος και ημερομηνία), την κατάσταση εμβολιασμού κατά της νόσου, τα σχετικά συμπτώματα, τα ζωτικά σημεία και άλλα.

(π.χ. σύμπτωμα, τεστ COVID-19) και συμμετέχοντα (π.χ. υποκείμενη κατάσταση), αντίστοιχα. Αυτά τα άτομα συνδέονται μέσω κατάλληλα ονοματισμένων ρόλων με καθορισμένη ιεραρχία όπως απεικονίζεται στο επόμενο σχήμα 2.



Εικόνα 10: Σχηματική διάταξη αρχείων smarty4covid ονοματισμένων ρόλων με καθορισμένη ιεραρχία

Κάθε ρόλος συνδέεται με ένα πεδίο και μία εμβέλεια που υποδεικνύουν τους τύπους των ατόμων που μπορούν να συνδεθούν μέσω αυτού του ρόλου. Οι περιγραφείσες ιεραρχίες εννοιών και ρόλων παρέχονται σε μορφή OWL στο αρχείο [smarty-ontology.owl]. Χρησιμοποιώντας την ορολογία, οι πληροφορίες παρουσιάζονται στο σύνολο δεδομένων δηλώνονται με τη μορφή τριπλών στοιχείων, που παρέχονται στο αρχείο [smarty-triples.nt].

Coswara

Το Coswara είναι ένα καινοτόμο έργο που αναπτύχθηκε από το Indian Institute of Science (IISc) στη Βενγαλούρου της Ινδίας. Ο στόχος του Coswara είναι η ανάπτυξη ενός συστήματος για την ανίχνευση της νόσου COVID-19 μέσω της ανάλυσης ηχητικών δειγμάτων, όπως η φωνή, ο βήχας και η αναπνοή των χρηστών. Το έργο βασίζεται στην ιδέα ότι οι λοιμώξεις του αναπνευστικού συστήματος, όπως της νόσου COVID-19, μπορούν να προκαλέσουν χαρακτηριστικές αλλαγές στους ήχους που παράγονται κατά την αναπνοή, την ομιλία και τον βήχα, οι οποίες μπορούν να εντοπιστούν μέσω ανάλυσης ήχου με χρήση μηχανικής μάθησης.

Το Coswara ξεκίνησε στις αρχές του 2020 [48], όταν η πανδημία COVID-19 βρισκόταν σε πλήρη εξέλιξη και οι παραδοσιακές μέθοδοι διάγνωσης ήταν περιορισμένες λόγω του υψηλού κόστους και της ανάγκης για εξειδικευμένο εξοπλισμό. Το Coswara εισήγαγε μια προσέγγιση βασισμένη σε ηχητικά δεδομένα, η οποία είναι μη επεμβατική και μπορεί να εφαρμοστεί εύκολα μέσω διαδικτύου [49].

Η ομάδα ανάπτυξης κάλεσε άτομα από όλο τον κόσμο να συμμετάσχουν στη συλλογή δεδομένων, καταγράφοντας τα ηχητικά τους δείγματα μέσω μιας διαδικτυακής πλατφόρμας. Τα δεδομένα που συλλέχθηκαν χρησιμοποιήθηκαν για την εκπαίδευση αλγορίθμων μηχανικής μάθησης, οι οποίοι μπορούσαν να εντοπίσουν μοτίβα που σχετίζονται με την νόσο COVID-19.

Το έργο Coswara συνέβαλε σημαντικά στην κατανόηση του πώς μπορούν τα ηχητικά σήματα να χρησιμοποιηθούν ως βιοδείκτες για την ανίχνευση αναπνευστικών λοιμώξεων, και αποτέλεσε βάση για την ανάπτυξη φορητών εργαλείων διάγνωσης, τα οποία θα μπορούσαν να είναι προσβάσιμα από όλους, ανεξαρτήτως γεωγραφικής θέσης ή οικονομικών πόρων.

Το Coswara αναγνωρίζεται ευρέως για την πρωτοποριακή του προσέγγιση και την επιστημονική του συμβολή στη διάγνωση αναπνευστικών νόσων με χρήση ηχητικών δειγμάτων [50].

3.2. Εξαγωγή Χαρακτηριστικών

Στην συνέχεια, ανάλογα με τον αριθμό των παρεχόμενων υπερπαραμέτρων εξαγωγής χαρακτηριστικών, εκτελείται εξαγωγή των χαρακτηριστικών στο σύνολο των δεδομένων. Οι υπερπαραμέτροι που επιλέχθηκαν είναι ο αριθμός των MFCCs, το μέγεθος του πλαισίου και ο αριθμός των τμημάτων. Όλες οι μέθοδοι και οι υπερπαραμέτροι που χρησιμοποιήθηκαν στην εξαγωγή χαρακτηριστικών βασίστηκαν στο επιστημονικό άρθρο “Preliminary diagnosis of COVID-19 based on cough sounds using machine learning algorithms” [51] και υλοποιήθηκαν με τη χρήση της βιβλιοθήκης librosa της Python με τις υπόλοιπες παράμετρους των μεθόδων να έχουν τις default τιμές τους. Το πρόγραμμα εξαγωγής χαρακτηριστικών κατασκευάστηκε ώστε να υποστηρίζει τρεις διαφορετικούς τύπους συναρτήσεων εξαγωγής για να δοκιμάζει σταδιακά την επίδραση των διαφορετικών υπερπαραμέτρων που αναφέρονται στο προαναφερόμενο έγγραφο.

Μέθοδος Εξαγωγής Χαρακτηριστικών MFCCs

Τα **Mel-Frequency Cepstral Coefficients (MFCCs)** αποτελούν μια από τις πιο διαδεδομένες και αποτελεσματικές μεθόδους για την εξαγωγή χαρακτηριστικών από σήματα ήχου, κυρίως στη φωνητική επεξεργασία. Τα MFCCs έχουν χρησιμοποιηθεί ευρέως σε εφαρμογές όπως η αναγνώριση ομιλίας, η ανάλυση συναισθημάτων από φωνή και η ταξινόμηση ήχων, λόγω της ικανότητάς τους να αποτυπώνουν τα ακουστικά χαρακτηριστικά του ανθρώπινου συστήματος ακοής [52]

Τα MFCCs προτάθηκαν τη δεκαετία του 1980 και έχουν γίνει το πρότυπο για την επεξεργασία ήχου και φωνής. Η αρχική τους ανάπτυξη σχετίζεται με την αναγνώριση ομιλίας, αλλά γρήγορα επεκτάθηκαν σε άλλους τομείς, όπως η αναγνώριση μουσικής και ηχητικών περιβαλλόντων.

Η διαδικασία εξαγωγής MFCCs περιλαμβάνει τα ακόλουθα στάδια:

- **Διαμόρφωση φάσματος ισχύος (Power Spectrum):** Το σήμα ήχου χωρίζεται σε μικρά τμήματα (frames), συνήθως 20-40ms, και για κάθε τμήμα υπολογίζεται το φάσμα ισχύος χρησιμοποιώντας τον Μετασχηματισμό Fourier (Fourier Transform).
- **Εφαρμογή φίλτρων Mel (Mel filterbank):** Το φάσμα ισχύος περνάει μέσα από μια σειρά φίλτρων που είναι κατανομημένα σε μια κλίμακα συχνοτήτων που προσεγγίζει την ανθρώπινη ακοή. Αυτή η κλίμακα συχνοτήτων είναι γνωστή ως **κλίμακα Mel**. Η κλίμακα Mel f_{melf} υπολογίζεται από τη συχνότητα f (σε Hz) με τον ακόλουθο τύπο:
$$f_{\text{mel}} = 2595 \cdot \log_{10} (1 + f / 700)$$
- **Λογαριθμική συμπίεση (Logarithmic Compression):** Το αποτέλεσμα των φίλτρων Mel υποβάλλεται σε λογαριθμική συμπίεση, προκειμένου να προσομοιωθεί η ανθρώπινη αντίληψη της έντασης του ήχου, καθώς η αντίληψη του ανθρώπινου αυτιού είναι λογαριθμική.
- **Διακριτός Μετασχηματισμός Συνημιτόνου (DCT):** Τέλος, η εφαρμογή του Διακριτού Μετασχηματισμού Συνημιτόνου (Discrete Cosine Transform, DCT) στα λογαριθμικά δεδομένα οδηγεί στην εξαγωγή των MFCCs. Οι πρώτοι συντελεστές από το αποτέλεσμα του DCT διατηρούνται, καθώς αυτοί περιέχουν τις πιο σημαντικές πληροφορίες για την αντιπροσώπευση του ήχου.

```

MFCCs in segment 7
[[-3.12329254e+02 -3.05374207e+02 -2.81030151e+02 -3.06605560e+02
  -3.10825500e+02 -3.05308289e+02 -3.12490356e+02 -3.21531067e+02
  -3.11526367e+02 -3.29572449e+02 -3.18971710e+02 -3.32260254e+02
  -3.25232208e+02 -3.02187805e+02 -2.91887878e+02]
[-2.04357662e+01 -3.44581299e+01 -2.29521370e+01 -2.96762848e+01
  -2.35890541e+01 -3.31315155e+01 -3.96271439e+01 -4.76014328e+01
  -3.24254227e+01 -2.88252754e+01 -3.32663994e+01 -3.49843216e+01
  -4.43811340e+01 -2.18815289e+01 -1.53551025e+01]
[-3.26410980e+01 -3.78234100e+01 -4.68593979e+01 -4.55143471e+01
  -4.40619049e+01 -5.08513947e+01 -3.24733696e+01 -3.16510544e+01
  -4.03091393e+01 -3.16650848e+01 -4.06993866e+01 -4.22794113e+01
  -3.55737915e+01 -4.35316162e+01 -4.42546997e+01]
[-2.34836769e+00 -1.22637539e+01 -1.15561581e+01 -7.93985367e-01
  -1.39538517e+01 -1.87320900e+01 6.20239115e+00 2.30330276e+00
  -1.33633957e+01 -1.27269707e+01 -1.61030960e+00 -5.21238375e+00
  4.84145498e+00 -1.41512508e+01 -8.23033714e+00]
[-3.25219650e+01 -1.80783157e+01 -3.41986542e+01 -9.10674191e+00

```

Εικόνα 11: Τμήμα MFCCs ενός segment

Μέθοδος Εξαγωγής Χαρακτηριστικών Spectral Centroid

Το **Spectral Centroid** είναι ένα από τα πιο διαδεδομένα ακουστικά χαρακτηριστικά που χρησιμοποιούνται στην ανάλυση ήχου και μουσικής. Αποτελεί ένα μέτρο που περιγράφει τη "φωτεινότητα" ή την "καθαρότητα" ενός ήχου, καθώς αντιπροσωπεύει το κέντρο μάζας του φάσματος συχνοτήτων ενός σήματος. Στην πράξη, το Spectral Centroid συνδέεται με την αντίληψη της συχνότητας από τον ακροατή, όπου υψηλότερες τιμές του χαρακτηριστικού υποδηλώνουν υψηλότερες συχνότητες που κυριαρχούν στο σήμα [53].

Η χρήση του Spectral Centroid ως χαρακτηριστικό ανάλυσης ήχου ξεκίνησε με την ανάγκη για καλύτερη περιγραφή των συχνοτικών χαρακτηριστικών των ηχητικών σημάτων, κυρίως στη μουσική ανάλυση και αναγνώριση ομιλίας. Με την εξέλιξη των ψηφιακών υπολογιστικών συστημάτων και της επεξεργασίας σημάτων, η μέθοδος αυτή έγινε βασικό εργαλείο σε πολλές εφαρμογές ακουστικής ανάλυσης.

Το Spectral Centroid υπολογίζεται ως το βαρυκεντρικό σημείο του φάσματος ισχύος ενός σήματος. Για να γίνει αυτό, το σήμα πρώτα μετασχηματίζεται από τον χρονικό στον συχνοτικό χώρο, συνήθως με τον **Μετασχηματισμό Fourier (Fourier Transform)**. Στη συνέχεια, το Spectral Centroid υπολογίζεται χρησιμοποιώντας τον ακόλουθο τύπο:

$$C = \frac{\sum_{k=0}^{N-1} f(k)|X(k)|}{\sum_{k=0}^{N-1} |X(k)|}$$

όπου:

- $f(k)$ είναι η συχνότητα που αντιστοιχεί στο k -οστό bin του φάσματος,
- $|X(k)|$ είναι το μέγεθος του φάσματος για το k -οστό bin,
- N είναι ο αριθμός των συχνοτήτων (bins) στο φάσμα.

Ο παραπάνω τύπος υπολογίζει το "μέσο" της κατανομής του φάσματος, με βάση τα βάρη που δίνει η ένταση κάθε συχνότητας. Έτσι, το Spectral Centroid εκφράζει την "μέση" συχνότητα που υπάρχει στο σήμα.

```
SC in segment 7
[[4423.47309584 4743.20135861 4089.28173142 4193.00551706 4170.38427476
 4404.71549932 4737.76252041 4951.01021224 4312.3274794 4785.2248203
 4431.44902451 4475.65411123 4761.44195495 4278.1082125 4084.80710295]]
```

Εικόνα 12: Spectral Centroid ενός segment

Μέθοδος Εξαγωγής Χαρακτηριστικών Spectral Roll-off

Το **Spectral Roll-off** είναι ένα ακουστικό χαρακτηριστικό που χρησιμοποιείται ευρέως στην επεξεργασία ήχου και σήματος, ειδικά στην ανάλυση μουσικής και φωνής. Αντιπροσωπεύει τη συχνότητα κάτω από την οποία βρίσκεται ένα συγκεκριμένο ποσοστό (συνήθως 85% ή 95%) της συνολικής φασματικής ενέργειας ενός σήματος. Το χαρακτηριστικό αυτό είναι ιδιαίτερα χρήσιμο για τον διαχωρισμό ήχων που έχουν διαφορετική φασματική κατανομή, όπως για παράδειγμα τους τόνους από τους αρμονικούς θορύβους.

Το Spectral Roll-off αναπτύχθηκε ως χαρακτηριστικό για την ανάλυση μουσικής και ήχων στις αρχές της δεκαετίας του 1990 και χρησιμοποιήθηκε αρχικά στην αναγνώριση μουσικών οργάνων και στη μουσική ταξινόμηση [53]. Με την εξέλιξη της υπολογιστικής επεξεργασίας και των τεχνικών ανάλυσης ήχου, το χαρακτηριστικό αυτό έγινε θεμελιώδες εργαλείο σε πολλούς τομείς της ακουστικής ανάλυσης.

Το Spectral Roll-off ορίζεται μαθηματικά ως η συχνότητα f_r που ικανοποιεί την ακόλουθη σχέση:

$$\sum_{k=0}^{f_r} |X(k)| = a \sum_{k=0}^{f_{max}} |X(k)|$$

όπου:

- $|X(k)|$ είναι το μέγεθος του φάσματος στην συχνότητα k ,
- f_{max} είναι η μέγιστη συχνότητα του φάσματος,
- a είναι το ποσοστό της φασματικής ενέργειας που είναι επιθυμητό να καλύψει το Spectral Roll-off, συχνά οριζόμενο ως 0.85 ή 0.95 (δηλαδή 85% ή 95% της φασματικής ενέργειας).

Η βασική ιδέα πίσω από το Spectral Roll-off είναι ότι αντικατοπτρίζει τη "κοπή" της φασματικής κατανομής σε ένα σημείο όπου η μεγαλύτερη συγκέντρωση ενέργειας βρίσκεται στις χαμηλότερες συχνότητες, αφήνοντας μόνο ένα μικρό ποσοστό ενέργειας στις υψηλότερες συχνότητες.

```
SR in segment 7
[[6890.625    7751.953125 6373.828125 7579.6875    7407.421875 7924.21875
  7751.953125 7924.21875  7407.421875 7924.21875  7062.890625 7579.6875
  7579.6875  7579.6875  7751.953125]]
```

Εικόνα 13: Spectral Roll-off ενός segment

Μέθοδος Εξαγωγής Χαρακτηριστικών Zero-Crossing Rate(ZCR)

Το **Zero-Crossing Rate (ZCR)** είναι ένα από τα πιο βασικά και ευρέως χρησιμοποιούμενα χαρακτηριστικά στην ανάλυση ηχητικών σημάτων, ειδικά σε εφαρμογές επεξεργασίας φωνής και μουσικής. Ουσιαστικά, το Zero-Crossing Rate μετρά τον αριθμό των φορές που το ηχητικό σήμα αλλάζει πρόσημο (δηλαδή διασχίζει τον άξονα του μηδενός) εντός μιας συγκεκριμένης χρονικής περιόδου. Το χαρακτηριστικό αυτό είναι ιδιαίτερα χρήσιμο για τη διάκριση μεταξύ διαφορετικών τύπων ηχητικών σημάτων, όπως μεταξύ φωνής και μουσικής, ή μεταξύ ομιλίας και θορύβου [54].

Το Zero-Crossing Rate έχει τις ρίζες του στην ανάλυση ηχητικών σημάτων ήδη από τα πρώτα χρόνια της ψηφιακής επεξεργασίας σήματος. Χρησιμοποιήθηκε αρχικά στην αναγνώριση ομιλίας και τη διάκριση μεταξύ φωνητικών και μη φωνητικών τμημάτων. Με την πρόοδο της

τεχνολογίας και την ανάπτυξη πιο πολύπλοκων μεθόδων επεξεργασίας σήματος, το ZCR διατηρεί τη σημασία του ως ένα απλό αλλά ισχυρό εργαλείο στην ανάλυση ήχου.

Το Zero-Crossing Rate ορίζεται ως ο αριθμός των αλλαγών πρόσημου του σήματος ανά μονάδα χρόνου. Για ένα διακριτό σήμα $x[n]$, το ZCR υπολογίζεται συνήθως ως:

$$ZCR = \frac{1}{N-1} \sum_{n=1}^{N-1} \mathbf{1}_{\{x[n] \cdot x[n-1] < 0\}}$$

όπου:

- N είναι ο αριθμός των δειγμάτων στο σήμα,
- $\mathbf{1}_{\{\cdot\}}$ είναι η ένδειξη της συνάρτησης που ισούται με 1 όταν η συνθήκη είναι αληθής και με 0 όταν είναι ψευδής.

Ουσιαστικά, το ZCR μετρά το πλήθος των σημείων στο σήμα όπου το πρόσημο αλλάζει από θετικό σε αρνητικό ή από αρνητικό σε θετικό. Αυτή η μέτρηση δίνει μια ένδειξη της συχνοτικής περιεκτικότητας του σήματος, με τα σήματα που περιέχουν υψηλότερες συχνότητες να έχουν συνήθως υψηλότερο ZCR.

```
ZCR in segment 7
[[0.2109375 0.3359375 0.359375 0.328125 0.375 0.296875 0.390625
 0.4140625 0.3828125 0.40625 0.3984375 0.3984375 0.359375 0.359375
 0.140625 ]]
```

Εικόνα 14: Zero-Crossing Rate ενός segment

Υπερπαράμετροι Εξαγωγής Χαρακτηριστικών

Οι υπερπαράμετροι που χρησιμοποιούνται στην εξαγωγή χαρακτηριστικών βασίζονται στο επιστημονικό άρθρο “Preliminary diagnosis of COVID-19 based on cough sounds using machine learning algorithms” [51] και είναι οι ακόλουθοι :

- Ο αριθμός των MFCCs (n_mfcc)
- Το μέγεθος του πλαισίου ($frame_size$)
- Ο αριθμός των τμημάτων ($n_segments$)

Ο **αριθμός των MFCCs** (n_mfcc) αναφέρεται στον αριθμό των χαρακτηριστικών συντελεστών που εξάγονται από τη διαδικασία εξαγωγής χαρακτηριστικών MFCCs.

Το **μέγεθος του πλαισίου** (frame_size) αναφέρεται στον αριθμό των διαδοχικών παρατηρήσεων ή σημείων δεδομένων που χρησιμοποιούνται σε κάθε βήμα ανάλυσης ή εκπαίδευσης σε αλγορίθμους μηχανικής μάθησης, ειδικά για χρονικές σειρές και αλληλουχίες δεδομένων και χρησιμοποιείται σε όλες τις μεθόδους εξαγωγής χαρακτηριστικών.

Ο **αριθμός των τμημάτων** (n_segments) χρησιμοποιείται για να καθορίσει τον αριθμό των διακριτών τμημάτων ή εννοιών στις οποίες χωρίζονται τα δεδομένα πριν την ανάλυση ή την εκπαίδευση του μοντέλου και χρησιμοποιείται σε όλες τις μεθόδους εξαγωγής χαρακτηριστικών.

Συναρτήσεις Εξαγωγής Χαρακτηριστικών

Συνοπτική Παρουσίαση Υπερπαραμέτρων Εξαγωγής Χαρακτηριστικών

Υπερπαραμέτρος	Περιγραφή	Εύρος τιμών
k_1	Αριθμός των MFCCs	$13 \times k_1$, όπου το k_1 παίρνει τιμές [1, 2, 3, 4, 5]
k_2	Μέγεθος του πλαισίου	2^{k_2} , όπου το k_2 παίρνει τιμές [8, 9, 10, 11, 12].
k_3	Αριθμός των τμημάτων	$10 \times k_3$, όπου το k_3 παίρνει τιμές [5, 7, 10, 12, 15].

Πίνακας 1: Τιμές υπερπαραμέτρων για την εξαγωγή χαρακτηριστικών

Πρώτη Συνάρτηση

Η πρώτη συνάρτηση, ονομάζεται `extract_features_simple` στον κώδικα, χρησιμοποιεί μία μέθοδο εξαγωγής χαρακτηριστικών - την μέθοδο MFCCs και μία υπερπαραμέτρο εξαγωγής χαρακτηριστικών - τον αριθμό των MFCCs. Ο σκοπός της πρώτης συνάρτησης είναι η εξαγωγή των MFCCs από το σύνολο κάθε ηχητικού δείγματος και η επιστροφή τους ως χαρακτηριστικά του δείγματος.

Η υπερπαραμέτρος “αριθμός των MFCCs”, που συμβολίζεται ως `n_mfcc` στον κώδικα, καθορίζει τον αριθμό των MFCCs που θα διατηρηθούν από το σύνολο των εξαγόμενων χαρακτηριστικών της μεθόδου εξαγωγής.

Δεύτερη Συνάρτηση

Η δεύτερη συνάρτηση, που ονομάζεται `extract_features` στον κώδικα, χρησιμοποιεί τέσσερις μεθόδους, τις μεθόδους MFCCs, Spectral Centroid, Spectral Roll-off και Zero-Crossing Rate και δύο υπερπαραμέτρους, τον αριθμό των MFCCs και το μέγεθος του πλαισίου. Ο σκοπός της δεύτερης συνάρτησης είναι η εξαγωγή των MFCCs και η εξαγωγή των Spectral Centroid, Spectral Roll-off και Zero-Crossing Rate καθενός πλαισίου (frame) από το σύνολο κάθε ηχητικού δείγματος και επιστρέφονται ως τα χαρακτηριστικά εκείνου του δείγματος.

Η υπερπαραμέτρος “αριθμός του πλαισίου”, που συμβολίζεται ως `frame_size` στον κώδικα, καθορίζει πόσα δεδομένα χρησιμοποιούνται για να δημιουργήσουν ένα πλαίσιο (frame) εισόδου στις μεθόδους εξαγωγής χαρακτηριστικών.

Τρίτη Συνάρτηση

Η τρίτη συνάρτηση, που ονομάζεται `extract_features_with_segments` στον κώδικα, χρησιμοποιεί τέσσερις μεθόδους, τις μεθόδους MFCCs, Spectral Centroid, Spectral Roll-off και Zero-Crossing Rate και τρεις υπερπαραμέτρους, τον αριθμό των MFCCs, το μέγεθος του πλαισίου και τον αριθμό των τμημάτων. Πρώτα γίνεται διαχωρισμός του ηχητικού αρχείου, ανάλογα με τον αριθμό των τμημάτων που καθορίζεται από την τρίτη υπερπαραμέτρο και στη συνέχεια γίνεται η εξαγωγή των MFCCs και η εξαγωγή των Spectral Centroid, Spectral Roll-off και Zero-Crossing Rate καθενός πλαισίου (frame) από το σύνολο κάθε ηχητικού δείγματος και επιστρέφονται ως τα χαρακτηριστικά εκείνου του δείγματος.

Η υπερπαραμέτρος “αριθμός των τμημάτων”, που συμβολίζεται ως `n_segments` στον κώδικα, χρησιμοποιείται για να καθορίσει τον αριθμό των διακριτών τμημάτων στα οποία χωρίζεται το ηχητικό αρχείο πριν περάσει κάθε τμήμα ξεχωριστά από όλες τις μεθόδους.

3.3. Προεπεξεργασία Δεδομένων

Το τελικό αποτέλεσμα της εξαγωγής χαρακτηριστικών είναι ένα πλέγμα από χαρακτηριστικά για κάθε δείγμα συναρτήσει του συνόλου δεδομένων και υπερπαραμέτρων. Με το πέρας της διαδικασίας αυτής, εκκινεί η προεπεξεργασία των δεδομένων αυτού του πλέγματος.

Έλεγχος Εγκυρότητας

Αρχικά, οποιαδήποτε μη έγκυρα αποτελέσματα από ατελή δεδομένα ή μη έγκυρες εξαγωγές αφαιρούνται από τα υπόλοιπα για να αποφευχθεί η εισαγωγή σφαλμάτων στην εκπαίδευση

των ταξινομητών. Αυτό επιτυγχάνεται με τον έλεγχο όλων των χαρακτηριστικών κάθε εξαγωγής για “NaN” τιμές, που υποδηλώνει ότι η εξαγωγή δεν πραγματοποιήθηκε σωστά.

Κλιμάκωση Χαρακτηριστικών

Στην συνέχεια, τα χαρακτηριστικά κλιμακώνονται, με τον τρέχοντα Scaler που χρησιμοποιείται να είναι ο Scaler Min-Max, ώστε να μειωθεί η επίδραση των υψηλών τιμών κάποιων χαρακτηριστικών.

```
Scaling Dataset Complete
Dataset after scaling: 3393
features_scaled[0]: [0.92210996 0.86667876 0.8648046 0.86984014 0.86130771 0.87089637
0.86542389 0.90182388 0.8706078 0.86007994 0.86036596 0.85548653
0.82678062 0.85777704 0.85420268 0.86465589 0.83087156 0.84719319
0.82592083 0.85557937 0.86103563 0.86971684 0.8523069 0.85364753
0.85963582 0.84154763 0.88341689 0.87655013 0.87050981 0.87752537
0.88633633 0.87964739 0.88777785 0.89499529 0.84608335 0.87091044
0.88205916 0.88862882 0.87488666 0.90294192 0.89525478 0.86520925
0.8974222 0.83175804 0.87008849 0.88458333 0.90135184 0.90755167
```

Εικόνα 15: Τμήμα χαρακτηριστικών ενός δείγματος μετά από κλιμάκωση

Κλιμακωτής Scaler Min-max

Ο **Scaler Min-Max** είναι μια μέθοδος κανονικοποίησης που μετατρέπει τα χαρακτηριστικά δεδομένων σε μια καθορισμένη κλίμακα, συνήθως μεταξύ 0 και 1. Αυτή η τεχνική είναι χρήσιμη για την προετοιμασία δεδομένων πριν την εφαρμογή αλγορίθμων μηχανικής μάθησης, ειδικά σε περιπτώσεις όπου οι διάφορες κλίμακες των χαρακτηριστικών μπορεί να επηρεάσουν την απόδοση των μοντέλων [55].

Διαδικασία Κανονικοποίησης

Η διαδικασία κανονικοποίησης με Scaler Min-Max περιλαμβάνει τα εξής βήματα:

1. Υπολογισμός Ελάχιστου και Μέγιστου:

- Εξάγονται οι ελάχιστες και μέγιστες τιμές κάθε χαρακτηριστικού (ή στήλης) στα δεδομένα.

2. Εφαρμογή Σχηματισμού:

- Κάθε τιμή x ενός χαρακτηριστικού μετατρέπεται στη νέα κλίμακα χρησιμοποιώντας τον τύπο:

$$x_{scaled} = \frac{x - \min(X)}{\max(X) - \min(X)}$$

όπου $\min(X)$ και $\max(X)$ είναι οι ελάχιστες και μέγιστες τιμές του χαρακτηριστικού X , αντίστοιχα.

3. Επιλογή Εύρους:

- Στον απλούστερο τύπο, το εύρος κανονικοποίησης είναι συνήθως $[0, 1]$. Ωστόσο, μπορεί να προσαρμοστεί σε οποιοδήποτε εύρος $[a, b]$ με τον τύπο:

$$x_{scaled} = a + \frac{x - \min(X)}{\max(X) - \min(X)} \times (b - a)$$

όπου a και b είναι οι επιθυμητές τιμές του εύρους.

Ανάλυση Κύριων Συνιστωσών PCA

Έπειτα, εφαρμόζεται η μέθοδος PCA (Ανάλυση Κύριων Συνιστωσών) για να μειωθεί η διαστατικότητα του συνόλου χαρακτηριστικών. Η μέθοδος εφαρμόζεται με επιλογή διατήρησης του 95% της διακύμανσης του συνόλου.

```
After PCA
features_pca: (3393, 460)
features_pca[0]]: [ 0.1433526  0.11987703 -0.09009818  0.06986078  0.01067694  0.04813412
 0.07439433 -0.03310576  0.03257658 -0.11963188 -0.14143108  0.1562283
 0.12926175  0.11259917 -0.05616885 -0.02883914 -0.07883544 -0.22934125
 0.18459285 -0.18840565  0.00174022 -0.04776802  0.02049348  0.19711342
 0.10313838 -0.08369575  0.16831461 -0.16585454  0.20213253  0.1504904
 0.00300962  0.07198024  0.16503779 -0.18869621  0.18057698  0.26393127
 -0.35640298  0.32167546  0.1604114  0.28074797  0.09154159  0.01723247]
```

Εικόνα 16: Τμήμα χαρακτηριστικών ενός δείγματος μετά από μέθοδο PCA

Η **Ανάλυση Κύριων Συνιστωσών (PCA)** είναι μια στατιστική τεχνική που χρησιμοποιείται για την μείωση της διάστασης των δεδομένων ενώ διατηρεί την περισσότερη δυνατή πληροφορία. Η μέθοδος αναλύει τα δεδομένα για να εντοπίσει τις κύριες συνιστώσες (principal components) που εξηγούν την περισσότερη διακύμανση των δεδομένων. Το PCA είναι ιδιαίτερα χρήσιμο όταν τα δεδομένα περιέχουν πολλές μεταβλητές και είναι επιθυμητό να ανακαλυφθούν οι κύριες τάσεις ή σχέσεις σε αυτά [56].

Βήματα Υλοποίησης PCA

1. Κεντριοποίηση των Δεδομένων:

- ο Το πρώτο βήμα είναι η αφαίρεση της μέσης τιμής από κάθε χαρακτηριστικό, ώστε τα δεδομένα να έχουν μηδενική μέση τιμή.

2. Υπολογισμός του Πίνακα Συνολοκληρωτικών Συσχετίσεων:

- ο Υπολογίζεται ο πίνακας συσχέτισης ή ο πίνακας διακύμανσης-συσχέτισης από τα κεντροποιημένα δεδομένα.

3. Εύρεση των Ιδιοτιμών και Ιδιοδιανυσμάτων:

- ο Υπολογίζονται οι ιδιοτιμές (eigenvalues) και τα ιδιοδιανύσματα (eigenvectors) του πίνακα διακύμανσης-συσχέτισης. Τα ιδιοδιανύσματα αντιπροσωπεύουν τις κύριες συνιστώσες του χώρου χαρακτηριστικών, ενώ οι ιδιοτιμές αντιπροσωπεύουν τη διακύμανση που εξηγεί κάθε συνιστώσα.

4. Επιλογή Κύριων Συνιστωσών:

- ο Επιλέγονται οι πρώτες k κύριες συνιστώσες (με βάση τις μεγαλύτερες ιδιοτιμές) που εξηγούν το μεγαλύτερο ποσοστό της συνολικής διακύμανσης των δεδομένων. Η επιλογή της μεταβλητής k εξαρτάται από την επιθυμητή ποσότητα διακύμανσης που είναι θεμιτό να διατηρηθεί.

5. Μετασχηματισμός Δεδομένων:

- ο Τα δεδομένα προβάλλονται στο νέο χώρο συνιστωσών με χρήση των επιλεγμένων ιδιοδιανυσμάτων. Αυτό οδηγεί σε έναν νέο πίνακα χαρακτηριστικών με μειωμένες διαστάσεις.

Διαχωρισμός Δεδομένων σε Σύνολο Εκπαίδευσης και σε Σύνολο Ελέγχου

Τέλος, τα δεδομένα χωρίζονται σε δύο σύνολα, ένα σύνολο εκπαίδευσης και ένα σύνολο ελέγχου χρησιμοποιώντας μια ειδικά κατασκευασμένη συνάρτηση διαχωρισμού. Το πρώτο σύνολο θα χρησιμοποιηθεί για την εκπαίδευση των ταξινομητών, ενώ το δεύτερο θα χρησιμοποιηθεί για να αξιολογήσει την απόδοση τους αφού ολοκληρωθεί η εκπαίδευση.

```
After Split  
x_train size: 2715  
x_test size: 678
```

Εικόνα 17: Πλήθος δεδομένων στο σύνολο εκπαίδευσης και στο σύνολο ελέγχου

3.4. Μέθοδος Ensemble learning

Όταν ολοκληρωθεί η προεπεξεργασία των δεδομένων, θα πραγματοποιηθεί ένας τελευταίος διαχωρισμός του συνόλου εκπαίδευσης με την μέθοδο ensemble learning και σύστημα ψηφοφορίας της μεθόδου το hard voting.

Περιγραφή μεθόδου Ensemble learning

Η μέθοδος **ensemble learning** είναι μια τεχνική στη μηχανική μάθηση που συνδυάζει τα αποτελέσματα πολλών μοντέλων (υπομοντέλα) για να δημιουργήσει ένα ισχυρότερο, πιο ακριβές και πιο γενικευμένο μοντέλο. Αντί να βασίζεται σε ένα μόνο μοντέλο για την πρόβλεψη, η μέθοδος ensemble learning συνδυάζει πολλαπλές προγνώσεις για να βελτιώσει την τελική απόδοση και να μειώσει τη γενική αβεβαιότητα [57].

Βασικές Μέθοδοι Ensemble Learning

1. Bagging (Bootstrap Aggregating):

- ο **Διαδικασία:** Δημιουργεί πολλαπλά σύνολα δεδομένων με επαναληπτική δειγματοληψία (bootstrap sampling) από το αρχικό σύνολο εκπαίδευσης. Κάθε υπομοντέλο εκπαιδεύεται σε ένα από αυτά τα υποσύνολα και οι τελικές προγνώσεις γίνονται με την πλειοψηφία ή τον μέσο όρο των προβλέψεων όλων των υπομοντέλων.

2. Boosting:

- ο **Διαδικασία:** Δημιουργεί ακολουθίες υπομοντέλων, όπου κάθε νέο μοντέλο προσπαθεί να διορθώσει τα λάθη των προηγούμενων μοντέλων. Το Boosting συνδυάζει τα αποτελέσματα όλων των μοντέλων με βάρη, ενισχύοντας τα πιο αδύναμα σημεία της πρόβλεψης.

3. Stacking:

- ο **Διαδικασία:** Εκπαιδεύει πολλαπλά μοντέλα (βασικά μοντέλα) σε ένα σύνολο δεδομένων και στη συνέχεια εκπαιδεύει ένα δεύτερο μοντέλο (μετα-μοντέλο ή συνδυαστής) χρησιμοποιώντας τις προβλέψεις των βασικών μοντέλων ως χαρακτηριστικά εισόδου. Ο συνδυαστής μαθαίνει πώς να συνδυάζει τις προβλέψεις των βασικών μοντέλων για να βελτιώσει την τελική απόδοση.

Σύστημα ψηφοφορίας “hard voting”.

Το σύστημα ψηφοφορίας “hard voting” είναι μία τεχνική συνδυασμού μοντέλων που χρησιμοποιείται για την ενίσχυση της απόδοσης σε προβλήματα μηχανικής μάθησης, ιδίως σε καταστάσεις ταξινόμησης. Ανήκει στην ευρύτερη κατηγορία του ensemble learning, που συνδυάζει πολλαπλά μοντέλα για να δημιουργήσει ένα ισχυρότερο και πιο γενικευμένο μοντέλο. Στο hard voting, οι προβλέψεις των μεμονωμένων μοντέλων συνδυάζονται και η τελική απόφαση λαμβάνεται με βάση την πλειοψηφία των ψήφων [19], [58].

Λήψη Προβλέψεων:

Κάθε μοντέλο (ή classifier) κάνει πρόβλεψη για τα δεδομένα εισόδου. Οι προβλέψεις είναι συνήθως κατηγορίες (ή class labels) και κάθε μοντέλο ψηφίζει για μία κατηγορία.

Συγκέντρωση Ψήφων:

Για κάθε περίπτωση, συγκεντρώνονται οι προβλέψεις από όλα τα μοντέλα. Η κατηγορία που λαμβάνει τις περισσότερες ψήφους επιλέγεται ως η τελική πρόβλεψη του συστήματος ψηφοφορίας.

Υλοποίηση

```
Total Samples Train: 2715
Class 0: 2316
Class 1: 399
Total Samples Test: 678
Class 0: 339
Class 1: 339
```

Εικόνα 18: Ανισορροπία των κλάσεων των δεδομένων

Στο πρόγραμμα που υλοποιήθηκε, το σύνολο εκπαίδευσης χωρίζεται σε 5 σύνολα δεδομένων, καθώς η κατανομή των κλάσεων στα δείγματα είναι μη ισορροπημένη, διότι υπάρχουν περίπου 5 φορές περισσότερα υγιή δείγματα από ό,τι μολυσμένα με covid στο dataset.

```
Ensemble Training Dataset 0 with: 862 Samples
Class 0: 463
Class 1: 399
Ensemble Training Dataset 1 with: 862 Samples
Class 0: 463
Class 1: 399
Ensemble Training Dataset 2 with: 862 Samples
Class 0: 463
Class 1: 399
Ensemble Training Dataset 3 with: 862 Samples
Class 0: 463
Class 1: 399
Ensemble Training Dataset 4 with: 863 Samples
Class 0: 464
Class 1: 399
10. Classification Start
```

Εικόνα 19: Τα υποσύνολα εκπαίδευσης μετά το ensemble split

Με την μέθοδο αυτή λοιπόν διασφαλίζεται ότι κάθε υποσύνολο εκπαίδευσης έχει το 1/5 της μείζονος κλάσης και ένα αντίγραφο ολόκληρης της ελάσσονος κλάσης.

3.5. Εκπαίδευση Ταξινομητών

Με την ολοκλήρωση του ensemble split, ξεκινά η εκπαίδευση των 4 τύπων ταξινομητών για κάθε σύνολο υπερπαραμέτρων, ώστε να μπορεί να βρεθεί ο καλύτερος συνδυασμός τόσο των υπερπαραμέτρων εξαγωγής χαρακτηριστικών όσο και των υπερπαραμέτρων εκπαίδευσης ταξινομητών.

Όλοι οι ταξινομητές και οι υπερπαραμέτροι που χρησιμοποιούνται στην εκπαίδευση των μοντέλων ταξινομητών βασίστηκαν στο επιστημονικό άρθρο “COVID-19 cough classification using machine learning and global smartphone recordings” [59] και υλοποιήθηκαν με τη χρήση της βιβλιοθήκης sklearn της Python.

Μηχανική Μάθηση

Η Μηχανική Μάθηση (Machine Learning) αποτελεί υποκλάδο της Τεχνητής Νοημοσύνης (Artificial Intelligence) και ασχολείται με την κατασκευή μοντέλων που υλοποιούν συγκεκριμένους αλγόριθμους, χρησιμοποιώντας πειραματικά δεδομένα με σκοπό την εξαγωγή χρήσιμων προβλέψεων ή συμπερασμάτων. Σύμφωνα με τον ορισμό του Mitchell [18]: «Ένα πρόγραμμα υπολογιστή λέγεται ότι μαθαίνει από την εμπειρία Ε ως προς κάποια

κλάση εργασιών T και μέτρο απόδοσης P , αν η απόδοσή του σε εργασίες από το T , όπως μετريέται από το P , βελτιώνεται μέσω της εμπειρίας E ».

Οι αλγόριθμοι μηχανικής μάθησης χωρίζονται σε τρεις μεγάλες κατηγορίες : Επιβλεπόμενη Μάθηση (Supervised Learning), Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning) και Ενισχυτική Μάθηση (Reinforcement Learning).

Επιβλεπόμενη Μάθηση (Supervised Learning)

Στην επιβλεπόμενη μάθηση χρησιμοποιούνται σύνολα δεδομένων εισόδου x και εξόδου y . Τα δεδομένα εισόδου x αποτελούνται, για κάθε δείγμα, από ένα σύνολο χαρακτηριστικών (features) και τα δεδομένα εξόδου y αποτελούνται από συγκεκριμένες ετικέτες (labels). Σκοπός της διαδικασίας εκπαίδευσης είναι το μοντέλο να μάθει, να υπολογίσει, τη συνάρτηση f που αντιστοιχίζει βέλτιστα τα δεδομένα εισόδου x με τις αντίστοιχες ετικέτες y .

$$x \rightarrow y \text{ ή } y = f(x)$$

Κατά τη διαδικασία αξιολόγησης, το μοντέλο βασιζόμενο στο σύνολο των χαρακτηριστικών εισόδου κάθε δείγματος αντιστοιχίζει μόνο του τις ετικέτες y με απώτερο σκοπό να αποκτήσει την ικανότητα να αναθέτει τις σωστές ετικέτες σε νέα, άγνωστα δεδομένα, χωρίς αυτές να έχουν καθοριστεί [21].

Παράδειγμα της επιβλεπόμενης μάθησης είναι το πρόβλημα της ταξινόμησης (classification), με τα δεδομένα ενός αρχικού συνόλου να ανήκουν σε δύο ή περισσότερες κλάσεις οι οποίες καθορίζονται από τις ετικέτες τους. Το μοντέλο ταξινομητή εκπαιδεύεται αρχικά σε ένα σύνολο δεδομένων (training set) και στη συνέχεια προβλέπει την κλάση ενός διαφορετικού συνόλου δεδομένων (σύνολο εξέτασης- test set) με βάση τα χαρακτηριστικά του. Το αρχικό σύνολο δεδομένων απαρτίζεται από το σύνολο εκπαίδευσης μαζί με το σύνολο εξέτασης.

Οι αλγόριθμοι της Επιβλεπόμενης μηχανικής μάθησης που χρησιμοποιούνται από το πρόγραμμα χωρίζονται σε τρεις κατηγορίες:

- Logistic Regression
- K-Nearest Neighbors (k-NN)
- Multi-layer Perceptron (MLP)

Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)

Σε αντίθεση με την προηγούμενη κατηγορία, η μη επιβλεπόμενη μάθηση δεν χρησιμοποιεί ζεύγη τιμών εισόδου-εξόδου. Ο στόχος της είναι να αναγνωριστούν μοτίβα (patterns) που ενδεχομένως κρύβονται στα δεδομένα εισόδου, χωρίς την ανατροφοδότηση από τιμές εξόδου.

Ενισχυτική Μάθηση (Reinforcement Learning)

Η ενισχυτική μάθηση διαφέρει από τις προηγούμενες κατηγορίες, καθώς εστιάζει στη λήψη βέλτιστων αποφάσεων. Τα μοντέλα ενισχυτικής μάθησης περιέχουν έναν πράκτορα (agent), ο οποίος αλληλεπιδρά με ένα περιβάλλον (environment). Ο πράκτορας λαμβάνει κάποια παρατήρηση (observation) από το περιβάλλον, εκτελεί μια ενέργεια και στη συνέχεια, λαμβάνει μια ανταμοιβή (reward) από το περιβάλλον. Αυτή η διαδικασία συνεχίζεται επαναληπτικά. Η συμπεριφορά του πράκτορα διέπεται από μια πολιτική (policy), η οποία είναι μια συνάρτηση που αντιστοιχίζει τις παρατηρήσεις του περιβάλλοντος σε ενέργειες. Ο στόχος του αλγορίθμου ενισχυτικής μάθησης είναι να παράγει μια καλή πολιτική [60].

Ταξινομητές

Συνοπτική Παρουσίαση

Υπερπαράμετρος	Περιγραφή	Ταξινομητής	Εύρος τιμών
Z ₁	Ισχύς κανονικοποίησης	LR	10^{i_1} όπου $i_1 = -7, -6, \dots, 6, 7$ (10^{-7} to 10^7)
Z ₂	Αριθμός των γειτόνων	KNN	Από το 10 μέχρι το 100 με βήμα 10
Z ₃	Μέγεθος του φύλλου	KNN	Από το 5 μέχρι το 30 με βήμα 5
Z ₄	Αριθμός των κρυφών νευρώνων	MLP	10 to 100 in steps of 10
Z ₅	Ισχύς της L2 κανονικοποίησης	MLP	10^{i_4} where $i_4 = -7, -6, \dots, 6, 7$ (10^{-7} to 10^7)
Z ₆	Αρχική τιμή του ρυθμού μάθησης	MLP	0 to 1 in steps of 0.05

Πίνακας 2: Τιμές υπερπαραμέτρων για κάθε μοντέλο ταξινομητή

Logistic Regression

Ο πρώτος ταξινομητής που υλοποιήθηκε είναι το μοντέλο Logistic Regression και η υλοποίηση του στον κώδικα χρησιμοποιεί τις default παραμέτρους.

Η υπερπαραμέτρος που εξετάστηκε είναι η ισχύς κανονικοποίησης (Z_1), που ελέγχει την συνάρτηση κόστους του μοντέλου και συμβολίζεται ως C στον κώδικα. Η ισχύς κανονικοποίησης χρησιμοποιείται για να μειωθεί η πιθανότητα overfitting του μοντέλου.

Περιγραφή Ταξινομητή

Η Logistic Regression είναι μια στατιστική μέθοδος που χρησιμοποιείται κυρίως για την ταξινόμηση δυαδικών ή πολυωνυμικών δεδομένων. Το βασικό της μοντέλο βασίζεται στην πιθανότητα ενός γεγονότος, η οποία εκφράζεται με τη χρήση της λογιστικής συνάρτησης (logistic function), μιας S-σχήματος συνάρτησης που ορίζει την πιθανότητα ενός γεγονότος ως:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n)}}$$

Η συνάρτηση μετατρέπει την έξοδο σε μια τιμή μεταξύ 0 και 1, που αντιπροσωπεύει την πιθανότητα του αποτελέσματος να ανήκει σε μία από τις δύο κατηγορίες.

Η Logistic Regression αναπτύχθηκε από τον David Cox το 1958 [61], αν και οι ρίζες της βρίσκονται στο έργο του Pierre-Simon Laplace και του Carl Friedrich Gauss για τις πιθανότητες και τη στατιστική εκτίμηση τον 18ο και 19ο αιώνα. Η χρήση της αυξήθηκε κυρίως στον τομέα της βιοστατιστικής και της κοινωνικής επιστήμης, όπου η ανάγκη για ανάλυση δυαδικών αποτελεσμάτων ήταν κρίσιμη. Με την ανάπτυξη των υπολογιστών, η Logistic Regression έγινε βασικό εργαλείο στη μηχανική μάθηση, ιδιαίτερα στην ταξινόμηση, την πρόβλεψη δυαδικών αποτελεσμάτων, και την ανάλυση επιβίωσης [62].

Στη Logistic Regression, οι βαρύτητες (weights) $\beta_0, \beta_1, \dots, \beta_n$ καθορίζονται μέσω της μέγιστης πιθανοφάνειας (Maximum Likelihood Estimation - MLE), μιας διαδικασίας που στοχεύει στην εύρεση των παραμέτρων που μεγιστοποιούν την πιθανότητα παρατήρησης των δεδομένων που χρησιμοποιούνται για την εκπαίδευση του μοντέλου.

Η συνάρτηση κόστους που χρησιμοποιείται για την εκπαίδευση του μοντέλου είναι η διασταυρούμενη εντροπία (cross-entropy loss), η οποία μετρά την απόσταση μεταξύ των προβλεπόμενων πιθανοτήτων και των πραγματικών ετικετών (labels) των δεδομένων.

K-Nearest Neighbors (k-NN)

Ο δεύτερος ταξινομητής που υλοποιήθηκε είναι το μοντέλο K-Nearest Neighbors και η υλοποίηση του στον κώδικα χρησιμοποιεί τις default παραμέτρους.

Οι υπερπαραμέτροι που εξετάστηκαν είναι ο αριθμός των γειτόνων (Z_2), που συμβολίζεται ως `n_neighbors` στον κώδικα και το μέγεθος του φύλλου (Z_3), που συμβολίζεται ως `leaf_size` στον κώδικα. Ο αριθμός των γειτόνων ελέγχει πόσα “γειτονικά” δεδομένα στο ελεγχόμενο στοιχείο, δηλαδή πόσα δεδομένα με τη μικρότερη απόσταση από αυτό, θα χρησιμοποιηθούν για την πρόβλεψη της ετικέτας του ελεγχόμενου στοιχείου. Το μέγεθος του φύλλου ελέγχει το πόσα nodes θα έχει το δέντρο που δημιουργείται από τον αλγόριθμο πρόβλεψης του μοντέλου.

Περιγραφή Ταξινομητή

Ο αλγόριθμος **k-Nearest Neighbors (k-NN)** είναι ένας από τους απλούστερους και πιο διαισθητικούς αλγόριθμους μηχανικής μάθησης για ταξινόμηση. Η βασική αρχή του k-NN είναι ότι ένα δεδομένο σημείο θα ταξινομηθεί με βάση την πλειοψηφία των κατηγοριών των k πιο κοντινών γειτόνων του.

Η διαδικασία ταξινόμησης περιλαμβάνει τα εξής βήματα:

1. **Υπολογισμός Αποστάσεων:** Για κάθε νέο δεδομένο σημείο, υπολογίζεται η απόσταση του από όλα τα σημεία στο σύνολο εκπαίδευσης. Συνήθως χρησιμοποιείται η ευκλείδεια απόσταση, αν και μπορούν να χρησιμοποιηθούν και άλλες μετρικές, όπως η απόσταση Manhattan ή η απόσταση Minkowski.
2. **Επιλογή των k Γειτόνων:** Τα σημεία του συνόλου εκπαίδευσης ταξινομούνται με βάση την απόσταση από το νέο δεδομένο σημείο, και επιλέγονται οι k πιο κοντινοί γείτονες.
3. **Ψήφισμα:** Το νέο δεδομένο σημείο ταξινομείται στην κατηγορία που αντιπροσωπεύεται από την πλειοψηφία των k γειτόνων. Στην περίπτωση ισοψηφίας, μπορούν να χρησιμοποιηθούν στρατηγικές όπως η χρήση των αποστάσεων για το βάρος των ψήφων.

Ο αλγόριθμος k-NN έχει τις ρίζες του στις εργασίες του Evelyn Fix και του Joseph Hodges το 1951 [63], οι οποίοι ανέπτυξαν τη βασική ιδέα του για χρήση σε στατιστική αναγνώριση προτύπων. Ο αλγόριθμος έγινε δημοφιλής στη δεκαετία του 1960 και 1970, κυρίως λόγω της απλότητας του και της αποτελεσματικότητάς του σε μικρά σύνολα δεδομένων.

Με την εξέλιξη της υπολογιστικής ισχύος, ο k-NN παρέμεινε δημοφιλής στον τομέα της μηχανικής μάθησης, καθώς είναι εύκολα κατανοητός, εύκολος στην υλοποίηση και δεν απαιτεί παραδοχές για την κατανομή των δεδομένων [64].

Multi-layer Perceptron (MLP)

Ο τέταρτος ταξινομητής που υλοποιήθηκε είναι το μοντέλο Multi-layer Perceptron και η υλοποίηση του στον κώδικα χρησιμοποιεί τις default παραμέτρους, εκτός από την παράμετρο `early_stopping`, όπου επίλεγχεται η τιμή: `True`.

Οι υπερπαραμέτροι που εξετάστηκαν είναι ο αριθμός των κρυφών νευρώνων (Z_4), που συμβολίζεται ως `hidden_layer_sizes` στον κώδικα, την ισχύς της L2 κανονικοποίησης (Z_5), που συμβολίζεται ως `alpha` στον κώδικα και την αρχική τιμή του ρυθμού μάθησης (Z_6), που συμβολίζεται ως `learning_rate_init` στον κώδικα. Ο αριθμός των κρυφών νευρώνων καθορίζει τον αριθμό των νευρώνων που βρίσκονται στο κάθε κρυφό επίπεδο του νευρωνικού δικτύου. Η ισχύς της L2 κανονικοποίησης χρησιμοποιείται για να μειωθεί η πιθανότητα *overfitting* του μοντέλου. Η αρχική τιμή του ρυθμού μάθησης καθορίζει τον αρχικό ρυθμό μάθησης που θα χρησιμοποιήσει ο αλγόριθμος βελτιστοποίησης.

Περιγραφή Ταξινομητή

Το **Multi-layer Perceptron (MLP)** είναι ένας τύπος νευρωνικού δικτύου που αποτελεί τη βάση για πολλά από τα σύγχρονα μοντέλα βαθιάς μάθησης. Είναι ένα *feedforward* νευρωνικό δίκτυο, που σημαίνει ότι η ροή των πληροφοριών κατευθύνεται μόνο από τα εισερχόμενα δεδομένα προς την έξοδο, χωρίς ανατροφοδότηση. Το MLP ανήκει στην κατηγορία των επιβλεπόμενων μοντέλων μάθησης και είναι ιδιαίτερα αποτελεσματικό για προβλήματα ταξινόμησης.

Το MLP αποτελείται από τουλάχιστον τρία επίπεδα νευρώνων: το **εισαγωγικό επίπεδο** (input layer), ένα ή περισσότερα **κρυφά επίπεδα** (hidden layers), και το **εξερχόμενο επίπεδο** (output layer). Κάθε νευρώνας σε ένα επίπεδο είναι συνδεδεμένος με όλους τους νευρώνες στο επόμενο επίπεδο, σχηματίζοντας ένα πλήρως συνδεδεμένο δίκτυο.

Οι νευρώνες σε κάθε επίπεδο εκτελούν μια γραμμική μετατροπή των εισερχόμενων σημάτων ακολουθούμενη από μια μη γραμμική συνάρτηση ενεργοποίησης, όπως η **συνάρτηση sigmoid**, η **hyperbolic tangent** ή η **ReLU (Rectified Linear Unit)**. Αυτές οι συναρτήσεις ενεργοποίησης επιτρέπουν στο MLP να μάθει και να αναπαραστήσει μη γραμμικές σχέσεις στα δεδομένα.

Η εκπαίδευση ενός MLP πραγματοποιείται με τη διαδικασία **backpropagation** [23] σε συνδυασμό με έναν αλγόριθμο βελτιστοποίησης, όπως η μέθοδος **Stochastic Gradient Descent (SGD)**. Η **backpropagation** περιλαμβάνει την υπολογιστική των παραγώγων των συναρτήσεων κόστους σε σχέση με τα βάρη του δικτύου και την ενημέρωση των βαρών με βάση αυτά τα παράγωγα, ώστε να μειωθεί το σφάλμα πρόβλεψης.

Τα MLPs αναπτύχθηκαν αρχικά στη δεκαετία του 1980, ωστόσο η πραγματική τους διάδοση έγινε τη δεκαετία του 1990, όταν βελτιώθηκαν οι μέθοδοι εκπαίδευσης και η υπολογιστική ισχύς. Παρά την εισαγωγή πιο σύνθετων μοντέλων, τα MLPs παραμένουν βασικό εργαλείο στην μηχανική μάθηση, και πολλές από τις ιδέες τους χρησιμοποιούνται σε σύγχρονα μοντέλα βαθιάς μάθησης.

3.6. Έλεγχος εκπαίδευσης ταξινομητών

Όλα τα υποσύνολα εκπαίδευσης του ensemble learning εκπαιδεύονται στον ίδιο τύπο ταξινομητή ώστε να μην υπάρχει ανάμιξη μεθόδων. Κάθε ταξινομητής χρησιμοποιεί Grid Search με 5-fold Cross Validation για να ανακαλύψει τις υπερπαραμέτρους που έχουν την καλύτερη απόδοση. Η βέλτιστη απόδοση βαθμολογείται ελέγχοντας την τιμή της ROC AUC Curve για κάθε μοντέλο. Ο καλύτερος ταξινομητής επιστρέφεται στο κυρίως πρόγραμμα, μαζί με τις υπερπαραμέτρους που χρησιμοποιήθηκαν για την εκπαίδευσή του.

Grid Search 5-fold Cross Validation

Η μέθοδος **Grid Search με 5-fold Cross Validation** είναι μια στρατηγική για την αναζήτηση της καλύτερης ρύθμισης υπερπαραμέτρων ενός μοντέλου μηχανικής μάθησης. Η Grid Search είναι μια μέθοδος αναζήτησης σε πλέγμα, όπου δοκιμάζονται όλες οι συνδυασμένες τιμές υπερπαραμέτρων σε έναν καθορισμένο χώρο. Η 5-fold Cross Validation είναι μια τεχνική διασταυρούμενης επικύρωσης που χρησιμοποιείται για την εκτίμηση της γενικευσιμότητας του μοντέλου και την αποφυγή υπερπροσαρμογής [65], [66].

Καθορισμός Υπερπαραμέτρων:

Επιλέγονται οι υπερπαραμέτροι που θα βελτιστοποιηθούν και καθορίζονται οι υποψήφιες τιμές για κάθε υπερπαραμέτρο.

Δημιουργία Πλέγματος:

Δημιουργείται ένα πλέγμα (grid) όλων των πιθανών συνδυασμών των υπερπαραμέτρων από τις καθορισμένες τιμές. Κάθε συνδυασμός είναι μια ξεχωριστή περίπτωση που θα δοκιμαστεί.

Αξιολόγηση Συνδυασμών:

Κάθε συνδυασμός υπερπαραμέτρων εκπαιδεύεται με το σύνολο δεδομένων και αξιολογείται με τη μέθοδο Cross Validation. Για κάθε συνδυασμό, υπολογίζεται μια αξιολόγηση απόδοσης (π.χ., ακρίβεια, F1-score) που βασίζεται στα αποτελέσματα των fold.

5-Fold Cross Validation

Διαχωρισμός Δεδομένων:

Το σύνολο δεδομένων διαχωρίζεται σε 5 τυχαία υποσύνολα (folds). Κάθε fold χρησιμεύει ως σύνολο επικύρωσης μια φορά, ενώ τα υπόλοιπα 4 folds χρησιμοποιούνται για την εκπαίδευση του μοντέλου.

Εκπαίδευση και Επικύρωση:

Το μοντέλο εκπαιδεύεται με τα 4 folds και επικυρώνεται με το 1 fold. Αυτή η διαδικασία επαναλαμβάνεται 5 φορές, με κάθε fold να χρησιμοποιείται ως σύνολο επικύρωσης ακριβώς μία φορά.

Συνδυασμός Αποτελεσμάτων:

Ο μέσος όρος των αξιολογήσεων απόδοσης σε όλα τα folds χρησιμοποιείται για την εκτίμηση της γενικής απόδοσης του μοντέλου με το συγκεκριμένο συνδυασμό υπερπαραμέτρων.

Επιλογή Καλύτερου Συνδυασμού:

Ο συνδυασμός υπερπαραμέτρων που έχει την καλύτερη μέση αξιολόγηση απόδοσης επιλέγεται ως η βέλτιστη ρύθμιση.

ROC AUC Curve

Μετρική Αξιολόγησης ταξινομητών AUC-ROC

Η μετρική Area Under the Receiver Operating Characteristic Curve (AUC-ROC) αποτελεί ένα μέτρο απόδοσης για προβλήματα ταξινόμησης για διάφορες επιλογές κατωφλίων, δείχνοντας την ικανότητα διάκρισης των κλάσεων από το μοντέλο. Ο όρος AUC αντιπροσωπεύει τον βαθμό διαχωρισιμότητας σε μια καμπύλη πιθανοτήτων (ROC) [19], [58]. Η καμπύλη ROC, μέσω διαγράμματος, την ικανότητα διαχωρισμού ενός συστήματος δυαδικής ταξινόμησης καθώς ποικίλλει το κατώφλι διαχωρισμού του. Η καμπύλη

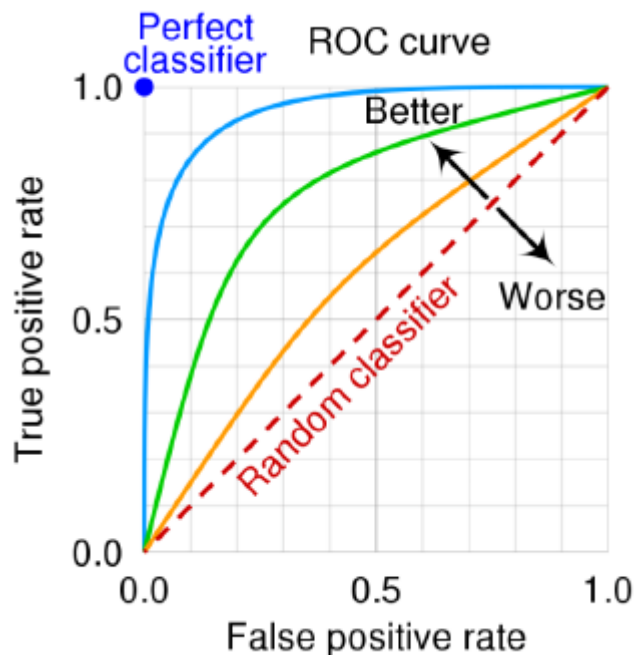
δημιουργείται με την αποτύπωση του ποσοστού αληθώς θετικών (True Positive Rate - TPR) έναντι του ποσοστού ψευδώς θετικών (False Positive Rate - FPR) σε διαφορετικές ρυθμίσεις κατωφλίων. Το ποσοστό αληθώς θετικών, αλλιώς ευαισθησία (recall), είναι ο λόγος των σωστά προβλεπόμενων θετικών παρατηρήσεων προς όλες τις παρατηρήσεις στην πραγματική κλάση και δίνεται από την ακόλουθη σχέση:

$$TPR = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

Ενώ, το ποσοστό των ψευδώς θετικών είναι ο λόγος των λανθασμένα προβλεπόμενων θετικών παρατηρήσεων προς όλες τις παρατηρήσεις στην πραγματική αρνητική κλάση και δίνεται από την ακόλουθη σχέση:

$$FPR = \frac{False\ Positives}{False\ Positives + True\ Negatives}$$

Η περιοχή κάτω από την καμπύλη (AUC) παρέχει ένα συνολικό μέτρο απόδοσης σε όλα τα δυνατά κατώφλια ταξινόμησης και κυμαίνεται μεταξύ των τιμών 0 και 1, όπως απεικονίζεται παρακάτω. Το μοντέλο επιτυγχάνει την ιδανική ταξινόμηση όταν AUC=1.



Εικόνα 20: Καμπύλη ROC για για “καλύτερους” και “χειρότερους” ταξινομητές[67]

Άξονας x: Ρυθμός Ψευδώς Θετικών (False Positive Rate, FPR).

Άξονας y: Ρυθμός Αληθώς Θετικών (True Positive Rate, TPR) ή Ευαισθησία (Sensitivity).

Δημιουργία ROC Curve:

Για κάθε πιθανό κατώφλι ταξινόμησης, υπολογίζονται οι τιμές FPR και TPR. Αυτές οι τιμές σχεδιάζονται ως σημεία στο γράφημα, και η καμπύλη σχηματίζεται συνδέοντας αυτά τα σημεία.

AUC (Area Under the Curve)

Η AUC είναι η επιφάνεια κάτω από την ROC Curve. Παρέχει μια συνολική μέτρηση της απόδοσης του ταξινομητή.

Ερμηνεία:

- ο **AUC = 1**: Τέλεια διάκριση (ο ταξινομητής δεν κάνει ποτέ λάθη).
- ο **$0.5 < \text{AUC} < 1$** : Ο ταξινομητής διαχωρίζει καλύτερα από την τυχαία πρόβλεψη. Μεγαλύτερη AUC υποδηλώνει καλύτερη απόδοση.
- ο **AUC = 0.5**: Ο ταξινομητής δεν διαχωρίζει καλύτερα από την τυχαία πρόβλεψη (τυχαία ταξινόμηση).
- ο **AUC < 0.5**: Ο ταξινομητής λειτουργεί χειρότερα από την τυχαία πρόβλεψη.

3.7. Αξιολόγηση Ταξινομητών – Δείκτες αξιολόγησης

Όταν η εκπαίδευση ολοκληρωθεί και για τα 4 μοντέλα, αυτά συνδυάζονται σε ένα ενιαίο ταξινομητή, ο οποίος με την σειρά του αξιολογείται στο σύνολο των δοκιμών. Ο ενιαίος ταξινομητής βασίζεται σε σύστημα ψηφοφορίας “hard voting”, που σημαίνει ότι η καθαρή πλειοψηφία αποφασίζει την έξοδο.

Οι δείκτες αξιολόγησης που χρησιμοποιούνται είναι οι δείκτες specificity, sensitivity, precision, accuracy και το F1 score [19], [58]. Ο ενιαίος ταξινομητής δεν χρησιμοποιεί την μετρική AUC-ROC, καθώς το σύστημα ψηφοφορίας δεν χρησιμοποιεί τις πιθανότητες κάθε κλάσης για την ψηφοφορία.

Δείκτης αξιολόγησης Specificity

Ο δείκτης **specificity** (ειδικότητα) είναι ένα σημαντικό μέτρο αξιολόγησης για τα μοντέλα ταξινόμησης. Η ειδικότητα είναι το αντίθετο της ευαισθησίας (sensitivity) και χρησιμοποιείται συνήθως σε συνδυασμό με άλλους δείκτες αξιολόγησης για να παρέχει μια πιο πλήρη εικόνα της απόδοσης του μοντέλου.

Η ειδικότητα μετρά την ικανότητα ενός μοντέλου να αναγνωρίζει σωστά τα αρνητικά παραδείγματα. Ο τύπος υπολογισμού της ειδικότητας είναι:

$$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives}$$

- True Negatives (TN): Αριθμός των αρνητικών παραδειγμάτων που ταξινομήθηκαν σωστά ως αρνητικά.
- False Positives (FP): Αριθμός των αρνητικών παραδειγμάτων που ταξινομήθηκαν εσφαλμένα ως θετικά.

Δείκτης αξιολόγησης Sensitivity

Ο δείκτης **sensitivity** (ευαισθησία) είναι ένα κρίσιμο μέτρο αξιολόγησης που χρησιμοποιείται για την εκτίμηση της ικανότητας ενός ταξινομητή να ανιχνεύει σωστά τα θετικά παραδείγματα. Ειδικότερα, μετρά την αναλογία των πραγματικών θετικών παραδειγμάτων που αναγνωρίζονται σωστά από το μοντέλο, και χρησιμοποιείται ευρέως σε καταστάσεις όπου η ανίχνευση των θετικών περιπτώσεων είναι ζωτικής σημασίας.

Η ευαισθησία ορίζεται ως η αναλογία των αληθών θετικών (True Positives, TP) προς το σύνολο των αληθών θετικών και των ψευδώς αρνητικών (False Negatives, FN). Ο τύπος υπολογισμού της ευαισθησίας είναι:

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

- True Positives (TP): Αριθμός των θετικών παραδειγμάτων που ταξινομήθηκαν σωστά ως θετικά.
- False Negatives (FN): Αριθμός των θετικών παραδειγμάτων που ταξινομήθηκαν εσφαλμένα ως αρνητικά.

Δείκτης αξιολόγησης Precision

Ο δείκτης **precision** (ακρίβεια) είναι ένα σημαντικό μέτρο αξιολόγησης σε μοντέλα ταξινόμησης που αναφέρεται στην ικανότητα του μοντέλου να κάνει σωστές θετικές προβλέψεις. Ειδικότερα, μετρά την αναλογία των αληθών θετικών (True Positives, TP) προς το σύνολο των παραδειγμάτων που το μοντέλο ταξινόμησε ως θετικά, συμπεριλαμβανομένων των ψευδώς θετικών (False Positives, FP). Ο δείκτης είναι ιδιαίτερα σημαντικός σε περιπτώσεις όπου η σωστή αναγνώριση των θετικών παραδειγμάτων έχει κρίσιμη σημασία.

Η ακρίβεια υπολογίζεται ως:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

- True Positives (TP): Ο αριθμός των θετικών παραδειγμάτων που ταξινομήθηκαν σωστά ως θετικά.
- False Positives (FP): Ο αριθμός των αρνητικών παραδειγμάτων που ταξινομήθηκαν εσφαλμένα ως θετικά.

Δείκτης αξιολόγησης Accuracy

Ο δείκτης **accuracy** (ακρίβεια) είναι ένα από τα πιο βασικά και ευρέως χρησιμοποιούμενα μέτρα αξιολόγησης της απόδοσης ενός μοντέλου ταξινόμησης. Αντιπροσωπεύει την αναλογία των σωστά ταξινομημένων παραδειγμάτων προς το συνολικό αριθμό παραδειγμάτων. Είναι ένα γενικό μέτρο της συνολικής απόδοσης του μοντέλου και χρησιμοποιείται ευρέως λόγω της απλότητας και της ευκολίας του στην ερμηνεία.

Η ακρίβεια υπολογίζεται ως:

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Number\ of\ Samples}$$

- True Positives (TP): Ο αριθμός των θετικών παραδειγμάτων που ταξινομήθηκαν σωστά ως θετικά.
- True Negatives (TN): Ο αριθμός των αρνητικών παραδειγμάτων που ταξινομήθηκαν σωστά ως αρνητικά.

Δείκτης αξιολόγησης F1 Score

Ο δείκτης **F1 Score** είναι ένα μέτρο αξιολόγησης που συνδυάζει την ακρίβεια (precision) και την ευαισθησία (recall) σε ένα ενιαίο αριθμητικό δείκτη. Ο δείκτης F1 Score χρησιμοποιείται ευρέως σε περιπτώσεις όπου υπάρχει ανισορροπία στις κλάσεις ή όταν η απόδοση του μοντέλου πρέπει να αξιολογηθεί με βάση την ισορροπία μεταξύ της ακρίβειας και της ευαισθησίας. Αποτελεί μια σημαντική μέτρηση για την απόδοση των μοντέλων ταξινόμησης, ιδίως σε προβλήματα με ανισόρροπα δεδομένα.

Ο δείκτης **F1 Score** είναι ο αρμονικός μέσος όρος της ακρίβειας (precision) και της ευαισθησίας (recall). Ορίζεται ως εξής:

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}$$

3.8. Ολοκλήρωση Προγράμματος

Αφού ολοκληρωθεί η αξιολόγηση του συνδυασμένου ταξινομητή ensemble learning, αποθηκεύονται σε ένα αρχείο csv οι παραπάνω δείκτες αξιολόγησης για κάθε τύπο ταξινομητή και η διαδικασία επαναλαμβάνεται μέχρι να εκπαιδευτούν όλοι οι τύποι ταξινομητών για όλους τους συνδυασμούς υπερπαραμέτρων.

4. Αποτελέσματα εκπαίδευσης ταξινομητών

Το πρόγραμμα εκπαίδευσης ταξινομητών, χρησιμοποίησε ως βάση δεδομένων την εφαρμογή **Smarty4Covid** καθώς και τη βάση δεδομένων **Coswara**, εφαρμόζοντας συνδυαστικές ρυθμίσεις υπερπαραμέτρων για τους τύπους ταξινομητών **Logistic Regression (LR)**, **K-Nearest Neighbors (KNN)** και **Multi-layer Perceptron (MLP)** και τα αποτελέσματα του παρουσιάζονται στους επόμενους πίνακες. Στόχος του προγράμματος ήταν η πρόβλεψη αν ένα άτομο πάσχει από COVID-19 με βάση το υποβληθέν δείγμα ήχου βήχα τους.

Στον επόμενο πίνακα παρατίθενται τα βασικά σταθερά στοιχεία τα οποία χρησιμοποιήθηκαν για την εξαγωγή των αποτελεσμάτων εκπαίδευσης των ταξινομητών.

Στοιχεία	Τιμές
Συνολικός αριθμός δεδομένων εκπαίδευσης	2715
Συνολικός αριθμός δεδομένων ελέγχου	678
Αριθμός Ensemble υποσυνόλων εκπαίδευσης	5

Πίνακας 3: Σταθερά Στοιχεία Εκπαίδευσης

Επίσης παρατίθενται στον επόμενο πίνακα τα βασικά μεταβλητά στοιχεία εκπαίδευσης των ταξινομητών.

Στοιχεία	Εύρος τιμών	Δυνατές τιμές
Ο αριθμός των MFCCs (n_mfcc)	14 - 70	14, 28, 42, 56, 70
Το μέγεθος του πλαισίου (frame_size)	256 - 4096	256, 512, 1024, 2048, 4096
Ο αριθμός των τμημάτων (n_segments)	50 - 150	50, 70, 100, 120, 150
Αποτελέσματα		
F1 Score (LR-KNN-MLP)	0 - 1	0% – 100%

Πίνακας 4: Μεταβλητά Στοιχεία Εκπαίδευσης

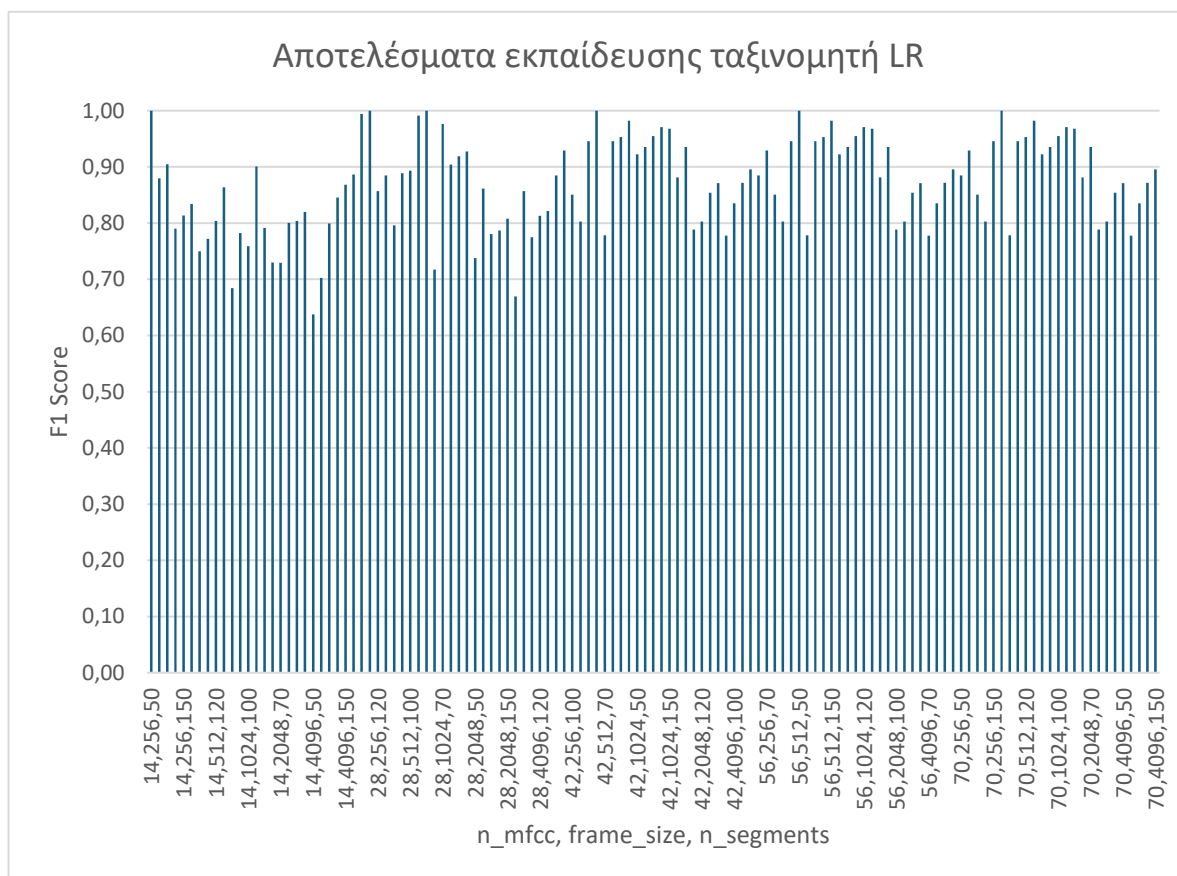
4.1. Αποτελέσματα εκπαίδευσης ταξινομητή Logistic Regression

4.1.1. Εκπαίδευση με βάση δεδομένων SmartyforCovid

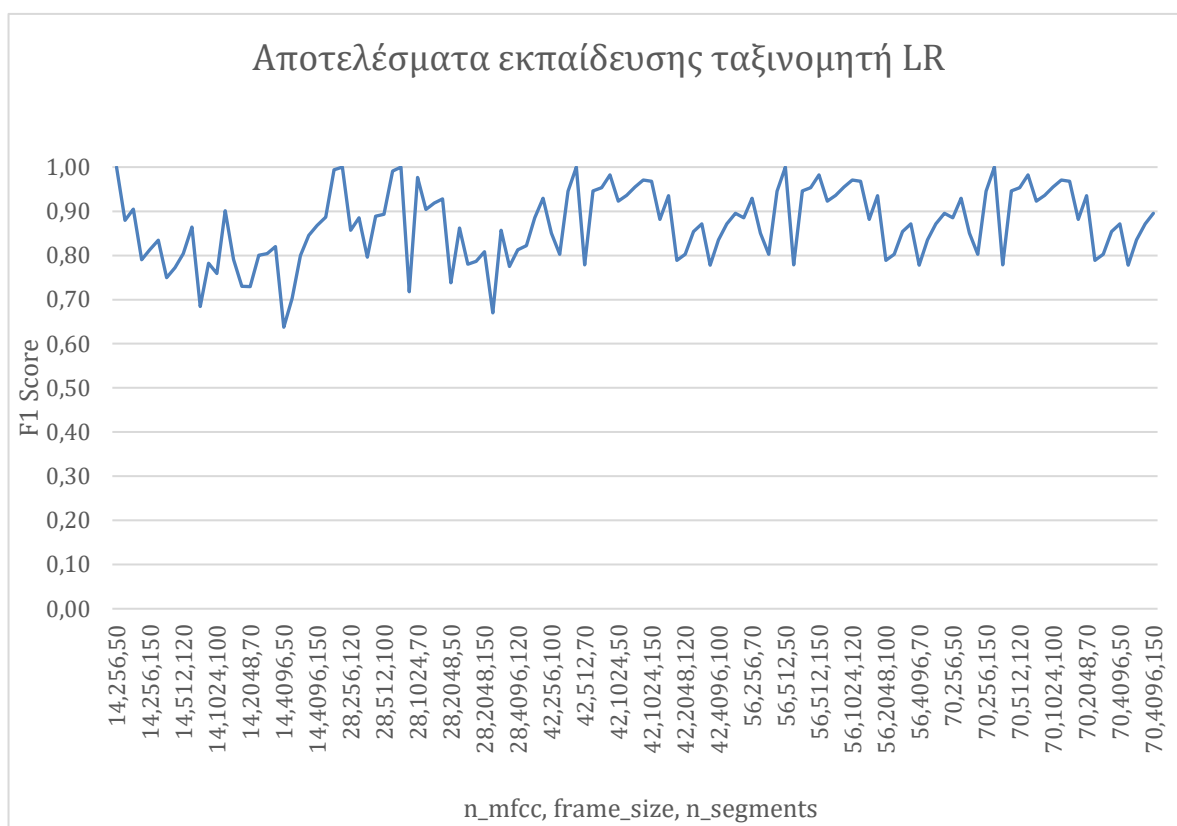
Το πρόγραμμα εκπαίδευσης ταξινομητών, χρησιμοποίησε την βάση δεδομένων της εφαρμογής **Smarty4Covid**, εφαρμόζοντας συνδυαστικές ρυθμίσεις υπερπαραμέτρων για τον τύπο ταξινομητή **Logistic Regression (LR)**. Τα αποτελέσματα του παρουσιάζονται στους επόμενους πίνακες και διαγράμματα και στη συνέχεια αναλύονται και συνάγονται χρήσιμα στατιστικά στοιχεία.

n_mfcc, frame_size, n_segments	F1 Score	n_mfcc, frame_size, n_segments	F1 Score	n_mfcc, frame_size, n_segments	F1 Score
14,256,50	1	28,2048,100	0,78041958	56,512,150	0,982404692
14,256,70	0,879765396	28,2048,120	0,786610879	56,1024,50	0,922636103
14,256,100	0,904899135	28,2048,150	0,808080808	56,1024,70	0,935672515
14,256,120	0,790490342	28,4096,50	0,669616519	56,1024,100	0,955007257
14,256,150	0,813411079	28,4096,70	0,857142857	56,1024,120	0,970845481
14,512,50	0,834319527	28,4096,100	0,774853801	56,1024,150	0,967930029
14,512,70	0,749629630	28,4096,120	0,813056380	56,2048,50	0,881405564
14,512,100	0,771929825	28,4096,150	0,821796760	56,2048,70	0,935483871
14,512,120	0,804093567	42,256,50	0,885007278	56,2048,100	0,788571429
14,512,150	0,863836018	42,256,70	0,929411765	56,2048,120	0,802877698
14,1024,50	0,684283727	42,256,100	0,850439883	56,2048,150	0,854197349
14,1024,70	0,782352941	42,256,120	0,802898551	56,4096,50	0,871201158
14,1024,100	0,758823529	42,256,150	0,945508100	56,4096,70	0,777618364
14,1024,120	0,900874636	42,512,50	1	56,4096,100	0,835260116
14,1024,150	0,791240876	42,512,70	0,778425656	56,4096,120	0,871491876
14,2048,50	0,729927007	42,512,100	0,945985401	56,4096,150	0,895434462
14,2048,70	0,729136164	42,512,120	0,953216374	70,256,50	0,885007278
14,2048,100	0,800569801	42,512,150	0,982404692	70,256,70	0,929411765
14,2048,120	0,804063861	42,1024,50	0,922636103	70,256,100	0,850439883
14,2048,150	0,820058997	42,1024,70	0,935672515	70,256,120	0,802898551
14,4096,50	0,637517630	42,1024,100	0,955007257	70,256,150	0,945508100
14,4096,70	0,702380952	42,1024,120	0,970845481	70,512,50	1
14,4096,100	0,799401198	42,1024,150	0,967930029	70,512,70	0,778425656
14,4096,120	0,845577211	42,2048,50	0,881405564	70,512,100	0,945985401
14,4096,150	0,868035191	42,2048,70	0,935483871	70,512,120	0,953216374
28,256,50	0,886567164	42,2048,100	0,788571429	70,512,150	0,982404692
28,256,70	0,994117647	42,2048,120	0,802877698	70,1024,50	0,922636103
28,256,100	1	42,2048,150	0,854197349	70,1024,70	0,935672515
28,256,120	0,856729378	42,4096,50	0,871201158	70,1024,100	0,955007257
28,256,150	0,885007278	42,4096,70	0,777618364	70,1024,120	0,970845481
28,512,50	0,795918367	42,4096,100	0,835260116	70,1024,150	0,967930029
28,512,70	0,888567294	42,4096,120	0,871491876	70,2048,50	0,881405564
28,512,100	0,893430657	42,4096,150	0,895434462	70,2048,70	0,935483871
28,512,120	0,991176471	56,256,50	0,885007278	70,2048,100	0,788571429
28,512,150	1	56,256,70	0,929411765	70,2048,120	0,802877698
28,1024,50	0,717579251	56,256,100	0,850439883	70,2048,150	0,854197349
28,1024,70	0,976470588	56,256,120	0,802898551	70,4096,50	0,871201158
28,1024,100	0,904069767	56,256,150	0,945508100	70,4096,70	0,777618364
28,1024,120	0,918840580	56,512,50	1	70,4096,100	0,835260116
28,1024,150	0,927745665	56,512,70	0,778425656	70,4096,120	0,871491876
28,2048,50	0,737727910	56,512,100	0,945985401	70,4096,150	0,895434462
28,2048,70	0,861717613	56,512,120	0,953216374		

Πίνακας 5: Αποτελέσματα εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments



Διάγραμμα 1: Ραβδόγραμμα αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc , $frame_size$, $n_segments$



Διάγραμμα 2: Διάγραμμα αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc , $frame_size$, $n_segments$

Εφαρμόζοντας τυπική στατιστική ανάλυση στα δεδομένα του ανωτέρω πίνακα συνάγονται τα επόμενα χρήσιμα στατιστικά στοιχεία.

Στοιχεία	Τιμές
Μέγιστη τιμή	1,000000000
Ελάχιστη τιμή	0,637517630
Διακύμανση	0,006752563
Μέσος Όρος	0,869045235
Διάμεσος	0,871491876
Επικρατούσα Τιμή	0,885007278

Πίνακας 6 Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid

Παρατηρείται ότι η μέγιστη τιμή είναι 1,0 δηλαδή 100% επιτυχία στην πρόβλεψη, η ελάχιστη τιμή είναι $\approx 0,6375$ δηλαδή 63,75 % επιτυχία στην πρόβλεψη, με διακύμανση τιμών $\approx 0,0068$. Ο μέσος όρος των αποτελεσμάτων είναι $\approx 0,8690$ δηλαδή 86,90 % επιτυχία στην πρόβλεψη, με διάμεση τιμή $\approx 0,8715$ (87,15 %) και επικρατούσα τιμή $\approx 0,8850$ (88,50 %). Συνάγεται από τις ανωτέρω τιμές της στατιστικής ανάλυσης, **με το σύνολο δεδομένων εκπαίδευσης smarty4covid**, ότι το ποσοστό επιτυχίας της πρόβλεψης για το αν ένα άτομο πάσχει από COVID-19 είναι αρκετά μεγάλο.

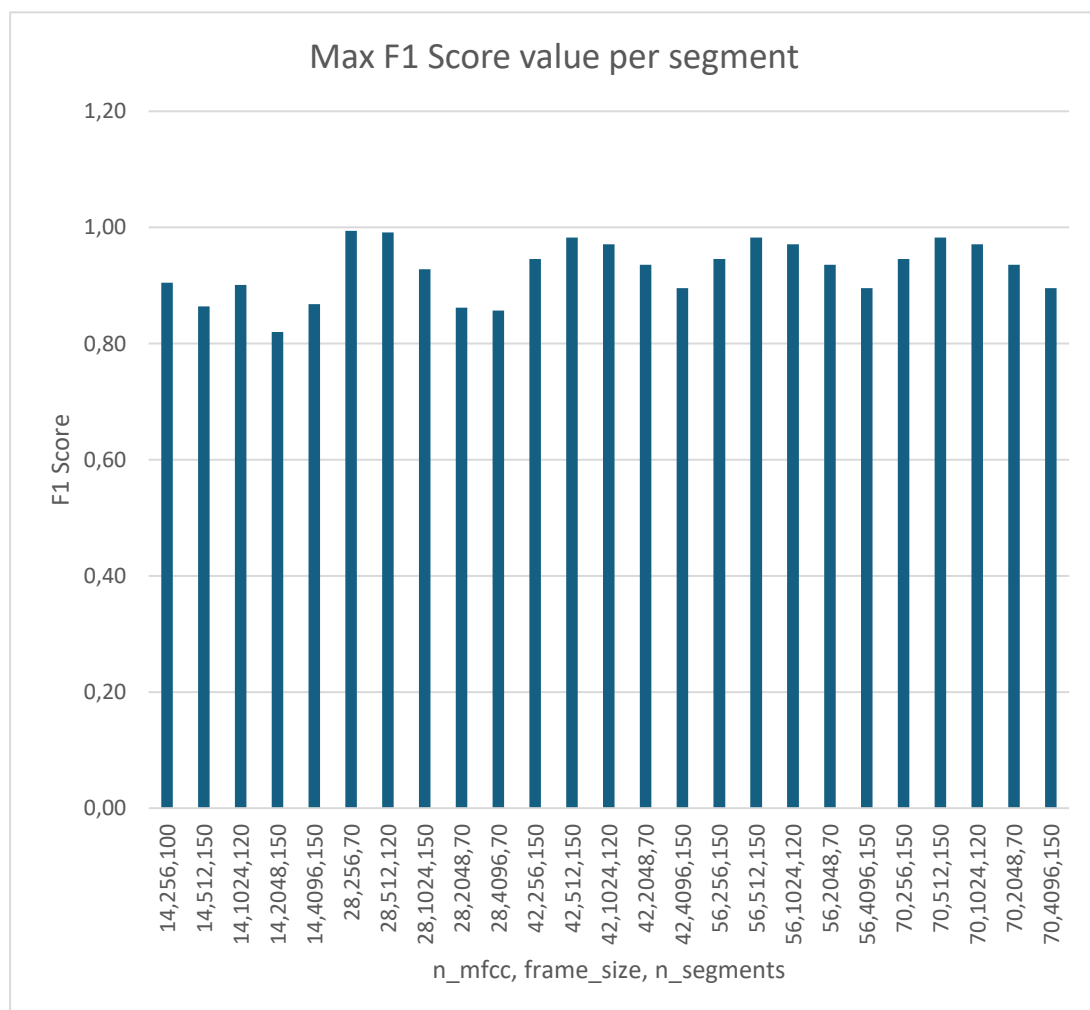
Συγκεκριμένα, παρατηρήθηκαν 6 αποτελέσματα F1 Score με τιμή 1, τα οποία θεωρούνται ως στατιστικές αιχμές, λόγω του χαμηλότερου μέσου όρου και του πλήθους ταξινομητών που εκπαιδεύτηκαν και η περαιτέρω στατιστική ανάλυση επικεντρώνεται στα υπόλοιπα αποτελέσματα.

Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού n_mfcc, frame_size & n_segments ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές n_segment ανά συνδυασμό n_mfcc & frame_size καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per segment).

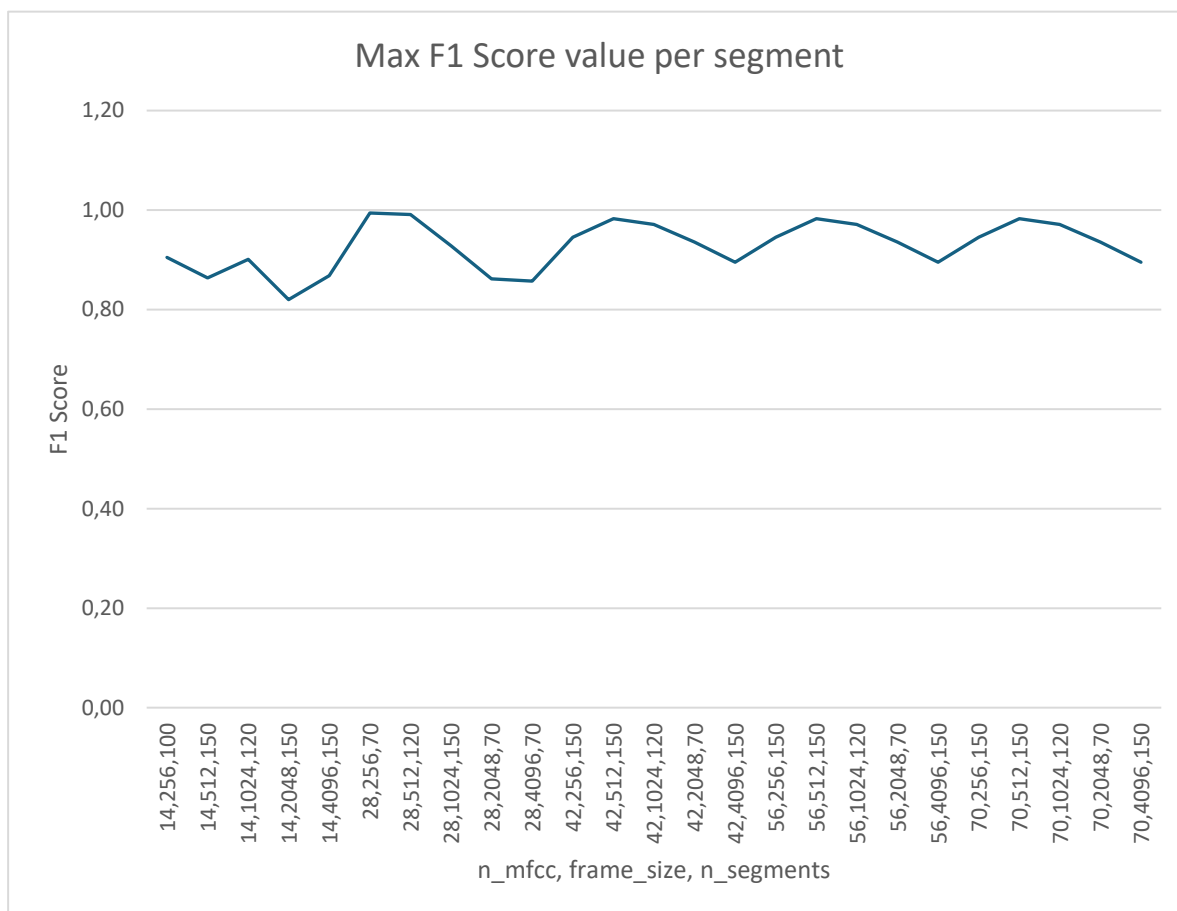
n_mfcc, frame_size, n_segments	Max F1 Score value per segment
14,256,100	0,904899135
14,512,150	0,863836018
14,1024,120	0,900874636
14,2048,150	0,820058997
14,4096,150	0,868035191
28,256,70	0,994117647
28,512,120	0,991176471
28,1024,150	0,927745665
28,2048,70	0,861717613
28,4096,70	0,857142857

42,256,150	0,945508100
42,512,150	0,982404692
42,1024,120	0,970845481
42,2048,70	0,935483871
42,4096,150	0,895434462
56,256,150	0,945508100
56,512,150	0,982404692
56,1024,120	0,970845481
56,2048,70	0,935483871
56,4096,150	0,895434462
70,256,150	0,945508100
70,512,150	0,982404692
70,1024,120	0,970845481
70,2048,70	0,935483871
70,4096,150	0,895434462

Πίνακας 7: Max F1 Score value per segment αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid



Διάγραμμα 3: Ραβδόγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size per n_segment



Διάγραμμα 4: Διάγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size per n_segment

Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού n_mfcc, frame_size & n_segments ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές frame_size & n_segments ανά n_mfcc καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per frame_size & n_segments).

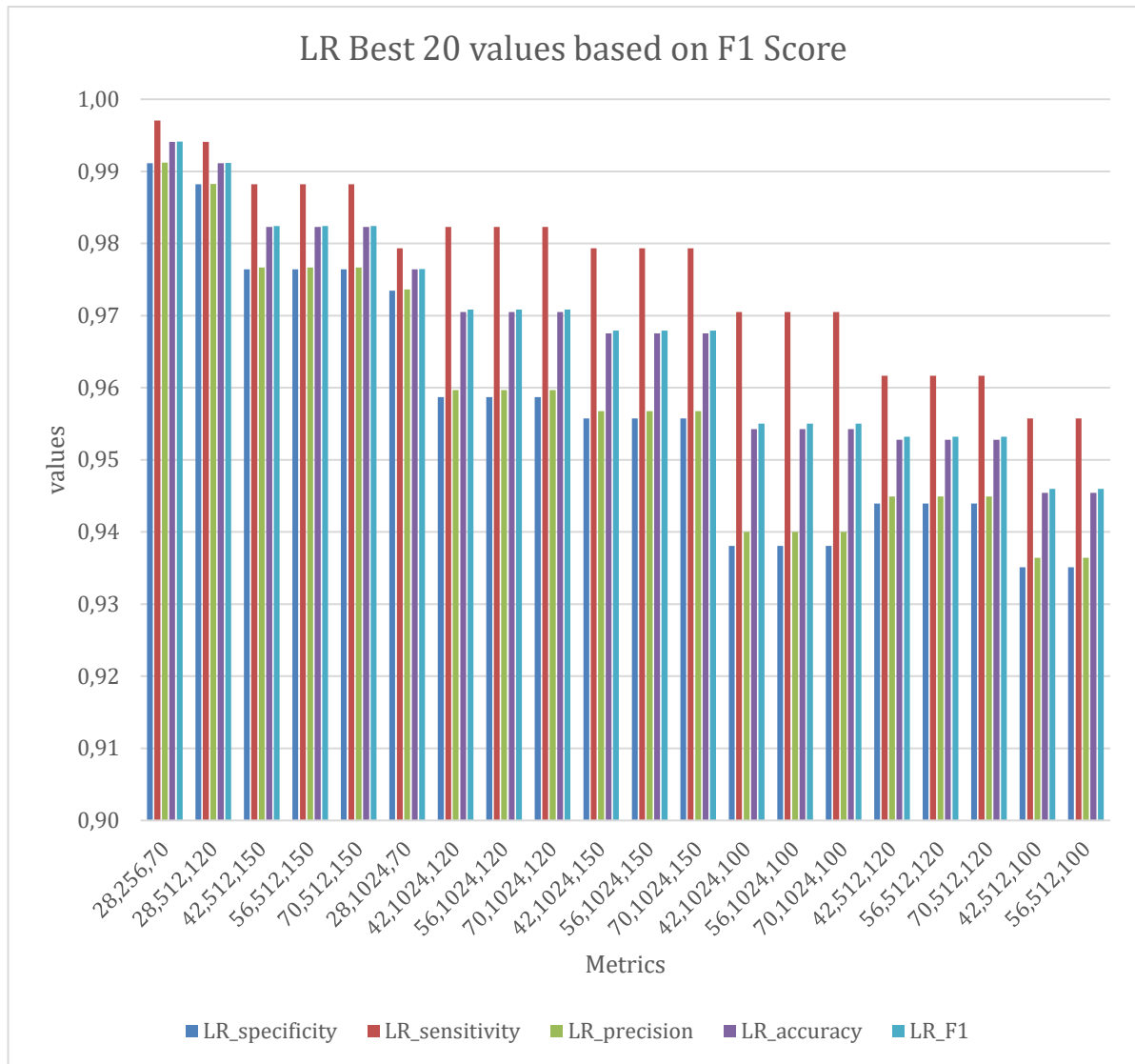
n_mfcc, frame_size, n_segments	Max F1 Score value per frame_size & n_segments
14,256,100	0,904899135
28,256,70	0,994117647
42,512,150	0,982404692
56,512,150	0,982404692
70,512,150	0,982404692

Πίνακας 8: Max F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid

Επιπρόσθετα στην ανάλυση παρατίθενται στον επόμενο πίνακα οι 20 μεγαλύτερες τιμές αποτελεσμάτων με βάση το F1 Score (Best 20 values based on F1 Score) με συνδυασμό n_mfcc, frame_size & n_segment. Στον πίνακα επιπλέον παρατίθενται και οι υπόλοιπες μετρικές για κάθε συνδυασμό.

n_mfcc, frame_size, n_segments	Specificity	Sensitivity	Precision	Accuracy	F1 Score
28,256,70	0,991150442	0,997050147	0,991202346	0,994100295	0,994117647
28,512,120	0,988200590	0,994100295	0,988269795	0,991150442	0,991176471
42,512,150	0,976401180	0,988200590	0,976676385	0,982300885	0,982404692
56,512,150	0,976401180	0,988200590	0,976676385	0,982300885	0,982404692
70,512,150	0,976401180	0,988200590	0,976676385	0,982300885	0,982404692
28,1024,70	0,973451327	0,979351032	0,973607038	0,976401180	0,976470588
42,1024,120	0,958702065	0,982300885	0,959654179	0,970501475	0,970845481
56,1024,120	0,958702065	0,982300885	0,959654179	0,970501475	0,970845481
70,1024,120	0,958702065	0,982300885	0,959654179	0,970501475	0,970845481
42,1024,150	0,955752212	0,979351032	0,956772334	0,967551622	0,967930029
56,1024,150	0,955752212	0,979351032	0,956772334	0,967551622	0,967930029
70,1024,150	0,955752212	0,979351032	0,956772334	0,967551622	0,967930029
42,1024,100	0,938053097	0,970501475	0,940000000	0,954277286	0,955007257
56,1024,100	0,938053097	0,970501475	0,940000000	0,954277286	0,955007257
70,1024,100	0,938053097	0,970501475	0,940000000	0,954277286	0,955007257
42,512,120	0,943952802	0,961651917	0,944927536	0,952802360	0,953216374
56,512,120	0,943952802	0,961651917	0,944927536	0,952802360	0,953216374
70,512,120	0,943952802	0,961651917	0,944927536	0,952802360	0,953216374
42,512,100	0,935103245	0,955752212	0,936416185	0,945427729	0,945985401
56,512,100	0,935103245	0,955752212	0,936416185	0,945427729	0,945985401

Πίνακας 9: Best 20 values based on F1 Score αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid



Διάγραμμα 5: Ραβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων smarty4covid (Best 20 values based on F1 Score)

Συμπέρασμα

Παρατηρείται ότι οι μεγαλύτερες τιμές (F1 Score) εμφανίζονται όταν υπάρχει συνδυασμός μικρών τιμών frame_size (256, 512 & 1024) ανεξάρτητα με τις τιμές τόσο του n_mfcc (εκτός της τιμής 14) όσο και του n_segments.

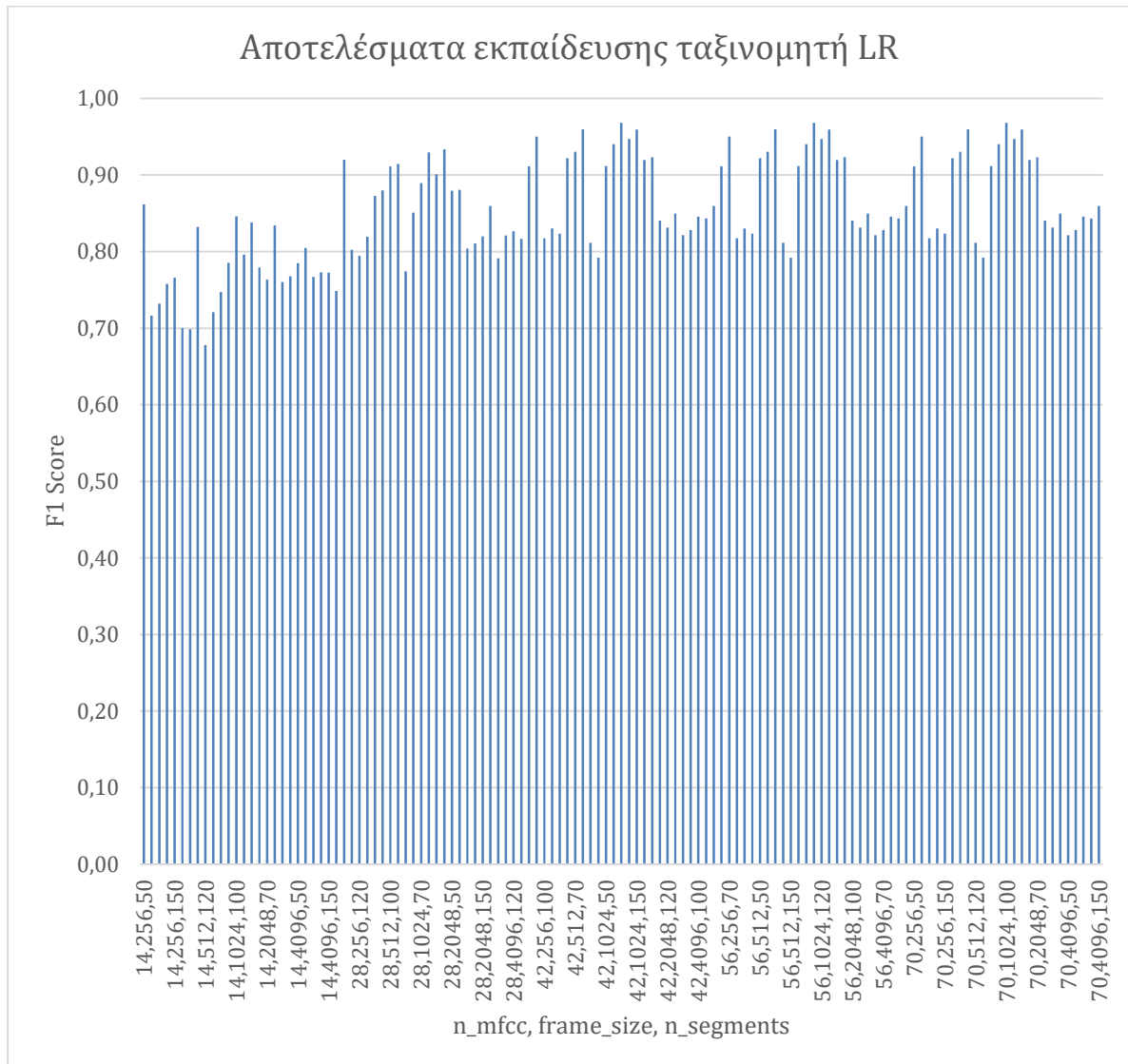
4.1.2. Εκπαίδευση με βάση δεδομένων Coswara

Το πρόγραμμα εκπαίδευσης ταξινομητών, χρησιμοποίησε την βάση δεδομένων της εφαρμογής **Coswara**, εφαρμόζοντας συνδυαστικές ρυθμίσεις υπερπαραμέτρων για τον τύπο ταξινομητή **Logistic Regression (LR)**. Τα αποτελέσματα του παρουσιάζονται στους επόμενους πίνακες και διαγράμματα και στη συνέχεια αναλύονται και συνάγονται χρήσιμα στατιστικά στοιχεία.

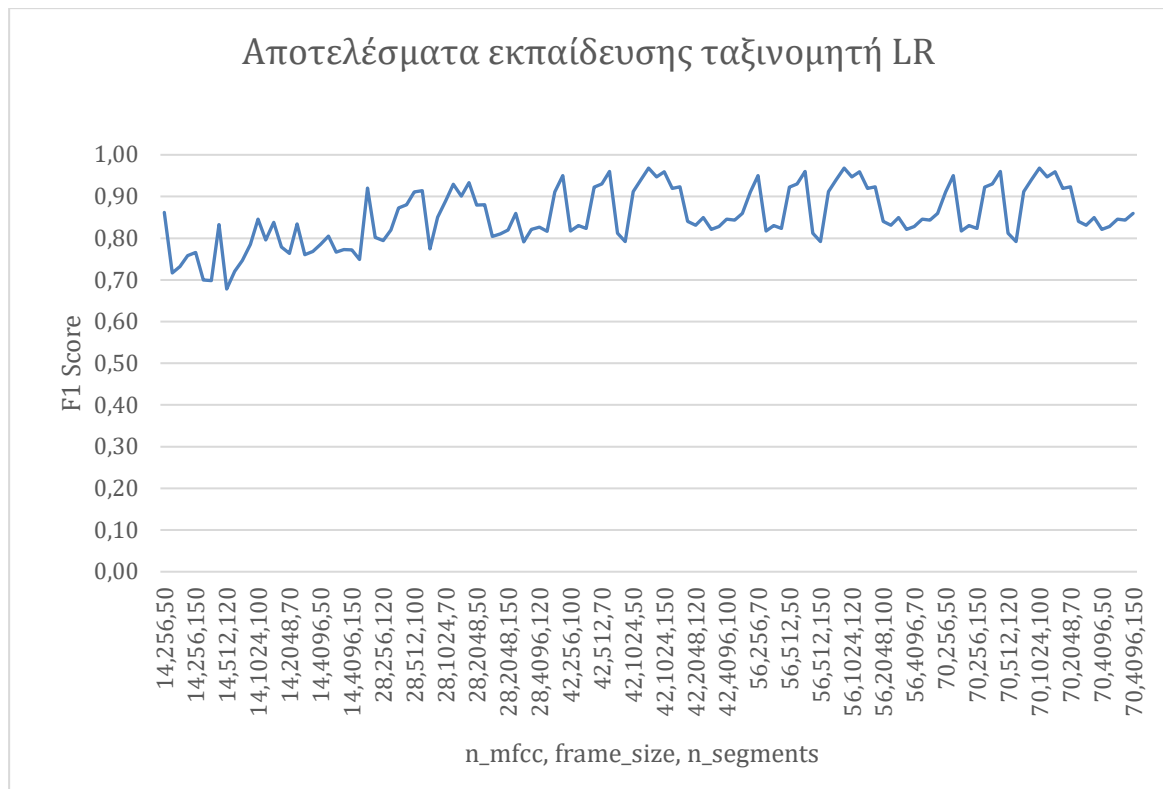
n_mfcc, frame_size, n_segments	F1 Score	n_mfcc, frame_size, n_segments	F1 Score	n_mfcc, frame_size, n_segments	F1 Score
14,256,50	0,861538462	28,2048,100	0,804000000	56,512,150	0,792000000
14,256,70	0,716636197	28,2048,120	0,810590631	56,1024,50	0,911646586
14,256,100	0,731884058	28,2048,150	0,819672131	56,1024,70	0,940239044
14,256,120	0,757973734	28,4096,50	0,859437751	56,1024,100	0,968253968
14,256,150	0,765799257	28,4096,70	0,791075051	56,1024,120	0,946938776
14,512,50	0,700173310	28,4096,100	0,821138211	56,1024,150	0,959514170
14,512,70	0,698656430	28,4096,120	0,826612903	56,2048,50	0,919678715
14,512,100	0,832323232	28,4096,150	0,816494845	56,2048,70	0,923076923
14,512,120	0,678111588	42,256,50	0,911242604	56,2048,100	0,840404040
14,512,150	0,720930233	42,256,70	0,950099800	56,2048,120	0,831275720
14,1024,50	0,746938776	42,256,100	0,817635271	56,2048,150	0,849484536
14,1024,70	0,785276074	42,256,120	0,830265849	56,4096,50	0,821355236
14,1024,100	0,845995893	42,256,150	0,823529412	56,4096,70	0,828157350
14,1024,120	0,795833333	42,512,50	0,922155689	56,4096,100	0,845528455
14,1024,150	0,838056680	42,512,70	0,930417495	56,4096,120	0,843177189
14,2048,50	0,779116466	42,512,100	0,959839357	56,4096,150	0,859470468
14,2048,70	0,763265306	42,512,120	0,811594203	70,256,50	0,911242604
14,2048,100	0,834008097	42,512,150	0,792000000	70,256,70	0,950099800
14,2048,120	0,760330579	42,1024,50	0,911646586	70,256,100	0,817635271
14,2048,150	0,767967146	42,1024,70	0,940239044	70,256,120	0,830265849
14,4096,50	0,784708249	42,1024,100	0,968253968	70,256,150	0,823529412
14,4096,70	0,804928131	42,1024,120	0,946938776	70,512,50	0,922155689
14,4096,100	0,767010309	42,1024,150	0,959514170	70,512,70	0,930417495
14,4096,120	0,773006135	42,2048,50	0,919678715	70,512,100	0,959839357
14,4096,150	0,772357724	42,2048,70	0,923076923	70,512,120	0,811594203
28,256,50	0,748623853	42,2048,100	0,840404040	70,512,150	0,792000000
28,256,70	0,920000000	42,2048,120	0,831275720	70,1024,50	0,911646586
28,256,100	0,802303263	42,2048,150	0,849484536	70,1024,70	0,940239044
28,256,120	0,794297352	42,4096,50	0,821355236	70,1024,100	0,968253968
28,256,150	0,819417476	42,4096,70	0,828157350	70,1024,120	0,946938776
28,512,50	0,872798434	42,4096,100	0,845528455	70,1024,150	0,959514170
28,512,70	0,880000000	42,4096,120	0,843177189	70,2048,50	0,919678715
28,512,100	0,911290323	42,4096,150	0,859470468	70,2048,70	0,923076923
28,512,120	0,914285714	56,256,50	0,911242604	70,2048,100	0,840404040
28,512,150	0,774193548	56,256,70	0,950099800	70,2048,120	0,831275720
28,1024,50	0,850715746	56,256,100	0,817635271	70,2048,150	0,849484536
28,1024,70	0,889336016	56,256,120	0,830265849	70,4096,50	0,821355236

28,1024,100	0,929577465	56,256,150	0,823529412	70,4096,70	0,828157350
28,1024,120	0,901010101	56,512,50	0,922155689	70,4096,100	0,845528455
28,1024,150	0,933333333	56,512,70	0,930417495	70,4096,120	0,843177189
28,2048,50	0,879518072	56,512,100	0,959839357	70,4096,150	0,859470468
28,2048,70	0,880478088	56,512,120	0,811594203		

Πίνακας 10: Αποτελέσματα εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc , $frame_size$, $n_segments$.



Διάγραμμα 6: Ραβδόγραμμα αποτελεσμάτων εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc , $frame_size$, $n_segments$



Διάγραμμα 7: Διάγραμμα αποτελεσμάτων εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc , $frame_size$, $n_segments$

Εφαρμόζοντας τυπική στατιστική ανάλυση στα δεδομένα του ανωτέρω πίνακα συνάγονται τα επόμενα χρήσιμα στατιστικά στοιχεία.

Στοιχεία	Τιμές
Μέγιστη τιμή	0,968253968
Ελάχιστη τιμή	0,678111588
Διακύμανση	0,004732127
Μέσος Όρος	0,853151746
Διάμεσος	0,843177189
Επικρατούσα Τιμή	0,911242604

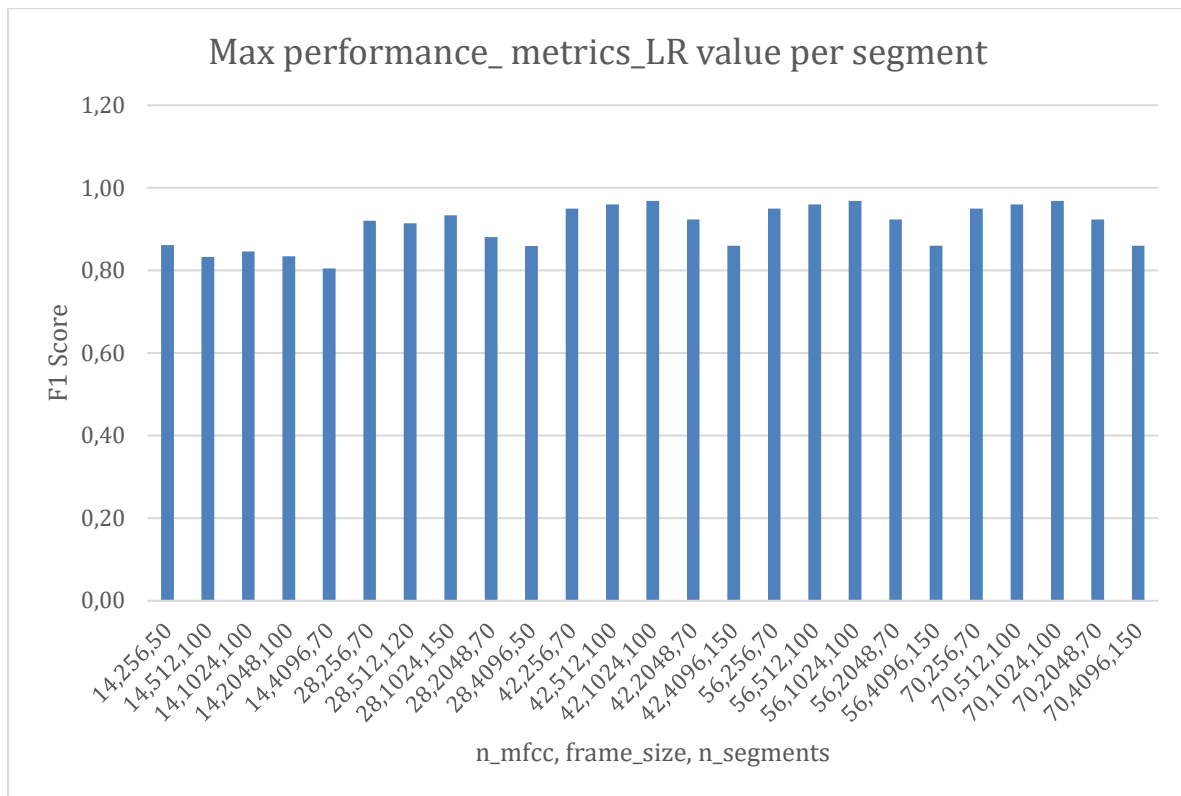
Πίνακας 11: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara

Παρατηρείται ότι η μέγιστη τιμή είναι 0,9683 δηλαδή 96,83 % επιτυχία στην πρόβλεψη, η ελάχιστη τιμή είναι $\approx 0,6781$ δηλαδή 67,81 % επιτυχία στην πρόβλεψη, με διακύμανση τιμών $\approx 0,0047$. Ο μέσος όρος των αποτελεσμάτων είναι $\approx 0,8532$ δηλαδή 85,32 % επιτυχία στην πρόβλεψη, με διάμεση τιμή $\approx 0,8432$ (84,32 %) και επικρατούσα τιμή $\approx 0,9112$ (91,12 %). Συνάγεται από τις ανωτέρω τιμές της στατιστικής ανάλυσης, **με το σύνολο δεδομένων εκπαίδευσης Coswara**, ότι το ποσοστό επιτυχίας της πρόβλεψης για το αν ένα άτομο πάσχει από COVID-19 είναι αρκετά μεγάλο.

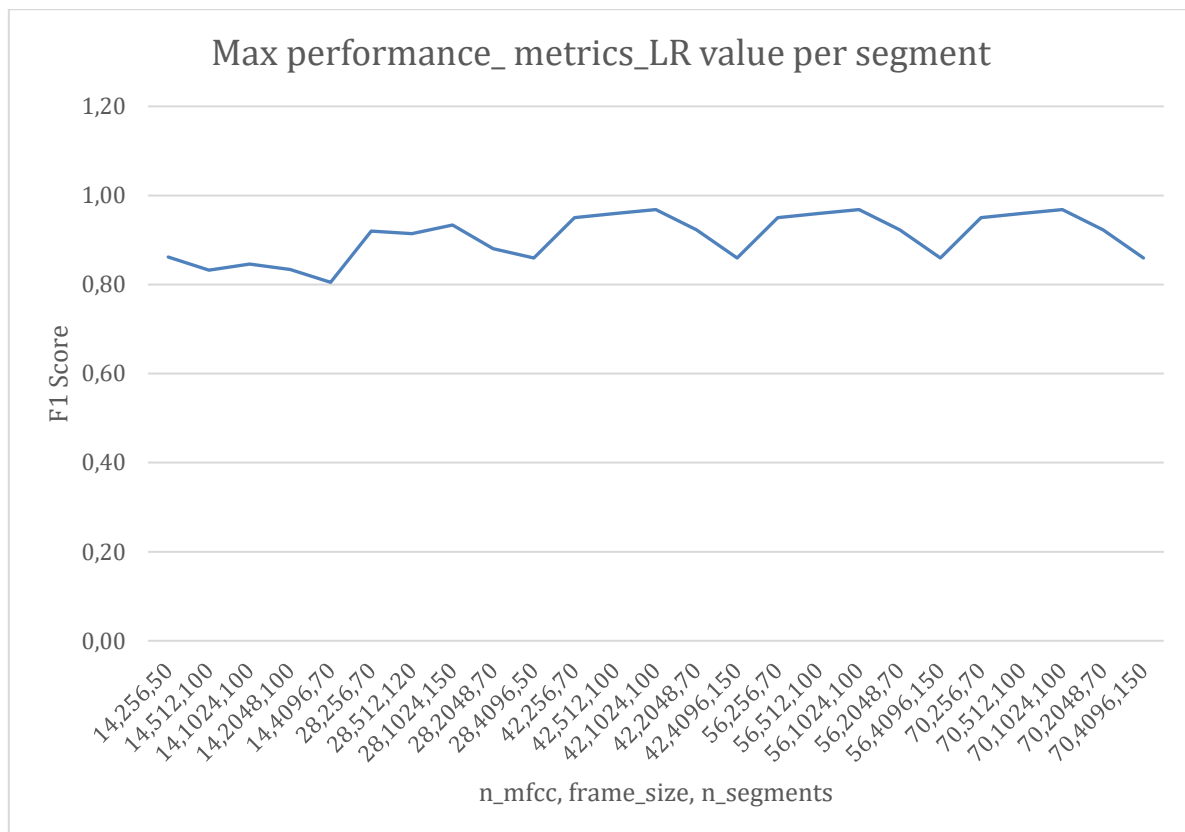
Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού n_mfcc, frame_size & n_segments ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές n_segment ανά συνδυασμό n_mfcc & frame_size καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per segment).

n_mfcc, frame_size, n_segments	Max F1 Score value per segment
14,256,50	0,861538462
14,512,100	0,832323232
14,1024,100	0,845995893
14,2048,100	0,834008097
14,4096,70	0,804928131
28,256,70	0,920000000
28,512,120	0,914285714
28,1024,150	0,933333333
28,2048,70	0,880478088
28,4096,50	0,859437751
42,256,70	0,950099800
42,512,100	0,959839357
42,1024,100	0,968253968
42,2048,70	0,923076923
42,4096,150	0,859470468
56,256,70	0,950099800
56,512,100	0,959839357
56,1024,100	0,968253968
56,2048,70	0,923076923
56,4096,150	0,859470468
70,256,70	0,950099800
70,512,100	0,959839357
70,1024,100	0,968253968
70,2048,70	0,923076923
70,4096,150	0,859470468

Πίνακας 12: Max F1 Score value per segment αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara



Διάγραμμα 8: Ραβδόγραμμα MAX αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size per n_segment



Διάγραμμα 9: Διάγραμμα MAX αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size per n_segment

Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού n_mfcc, frame_size & n_segments ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές frame_size & n_segments ανά n_mfcc καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per frame_size & n_segments).

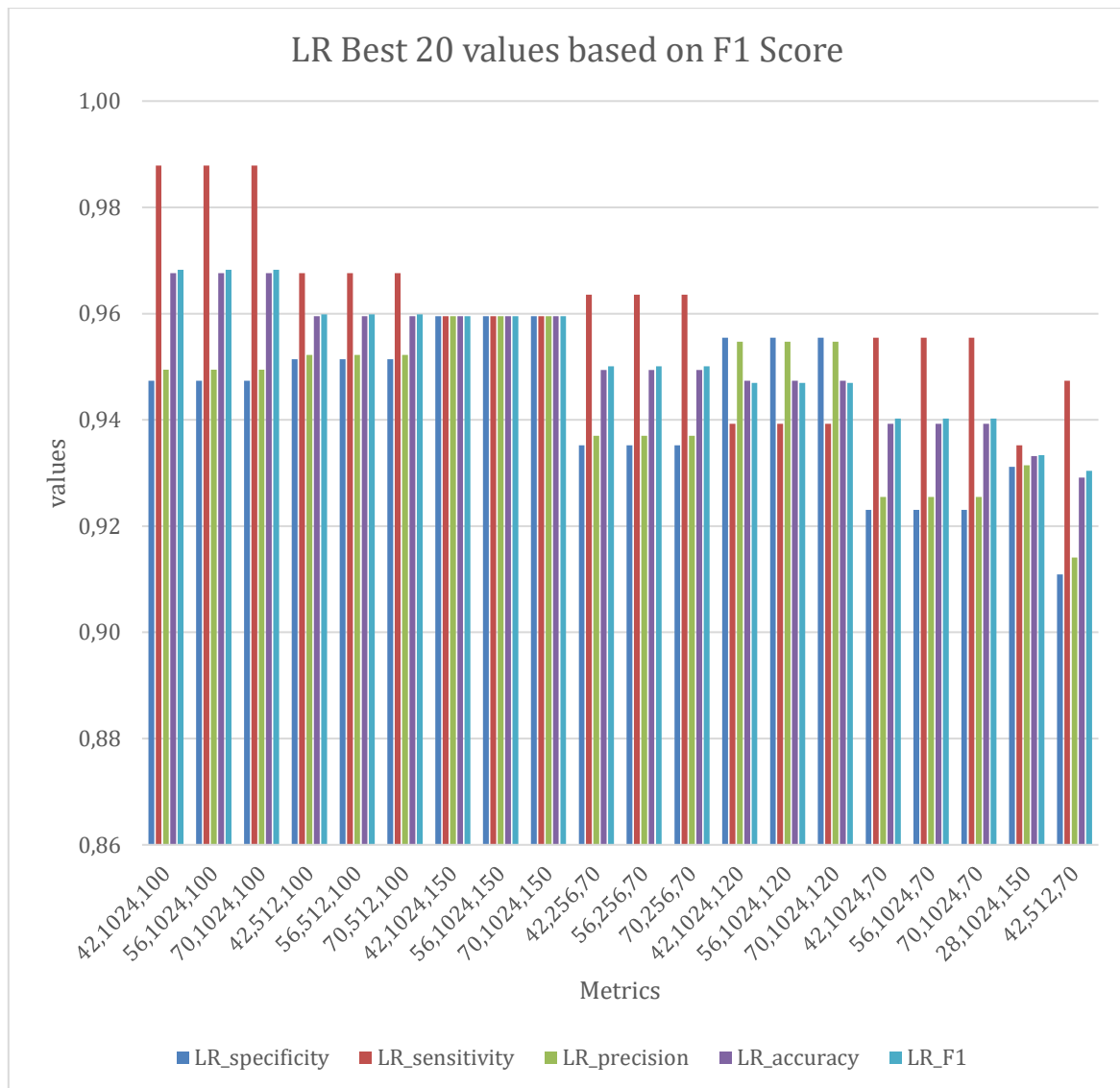
n_mfcc, frame_size, n_segments	Max F1 Score value per frame_size & n_segments
14,256,50	0,861538462
28,1024,150	0,933333333
42,1024,100	0,968253968
56,1024,100	0,968253968
70,1024,100	0,968253968

Πίνακας 13: Max F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara

Επιπρόσθετα στην ανάλυση παρατίθενται στον επόμενο πίνακα οι 20 μεγαλύτερες τιμές αποτελεσμάτων με βάση το F1 Score (Best 20 values based on F1 Score) με συνδυασμό n_mfcc, frame_size & n_segment. Στον πίνακα επιπλέον παρατίθενται και οι υπόλοιπες μετρικές για κάθε συνδυασμό.

n_mfcc, frame_size, n_segments	LR_specificity	LR_sensitivity	LR_precision	LR_accuracy	LR_F1
42,1024,100	0,947368421	0,987854251	0,949416342	0,967611336	0,968253968
56,1024,100	0,947368421	0,987854251	0,949416342	0,967611336	0,968253968
70,1024,100	0,947368421	0,987854251	0,949416342	0,967611336	0,968253968
42,512,100	0,951417004	0,967611336	0,952191235	0,959514170	0,959839357
56,512,100	0,951417004	0,967611336	0,952191235	0,959514170	0,959839357
70,512,100	0,951417004	0,967611336	0,952191235	0,959514170	0,959839357
42,1024,150	0,959514170	0,959514170	0,959514170	0,959514170	0,959514170
56,1024,150	0,959514170	0,959514170	0,959514170	0,959514170	0,959514170
70,1024,150	0,959514170	0,959514170	0,959514170	0,959514170	0,959514170
42,256,70	0,935222672	0,963562753	0,937007874	0,949392713	0,950099800
56,256,70	0,935222672	0,963562753	0,937007874	0,949392713	0,950099800
70,256,70	0,935222672	0,963562753	0,937007874	0,949392713	0,950099800
42,1024,120	0,955465587	0,939271255	0,954732510	0,947368421	0,946938776
56,1024,120	0,955465587	0,939271255	0,954732510	0,947368421	0,946938776
70,1024,120	0,955465587	0,939271255	0,954732510	0,947368421	0,946938776
42,1024,70	0,923076923	0,955465587	0,925490196	0,939271255	0,940239044
56,1024,70	0,923076923	0,955465587	0,925490196	0,939271255	0,940239044
70,1024,70	0,923076923	0,955465587	0,925490196	0,939271255	0,940239044
28,1024,150	0,931174089	0,935222672	0,931451613	0,933198381	0,933333333
42,512,70	0,910931174	0,947368421	0,914062500	0,929149798	0,930417495

Πίνακας 14: Best 20 values based on F1 Score αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara



Διάγραμμα 10: Ραβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελέσματος εκπαίδευσης ταξινομητή LR στο σύνολο δεδομένων coswara (Best 20 values based on F1 Score)

Συμπέρασμα

Παρατηρείται ότι οι μεγαλύτερες τιμές (F1 Score) εμφανίζονται όταν υπάρχει συνδυασμός μεγάλων τιμών `n_mfcc` (42, 56 & 70), μικρών τιμών `frame_size` (256, 512 & 1024) & μεγάλων τιμών `n_segments` (100 & 150).

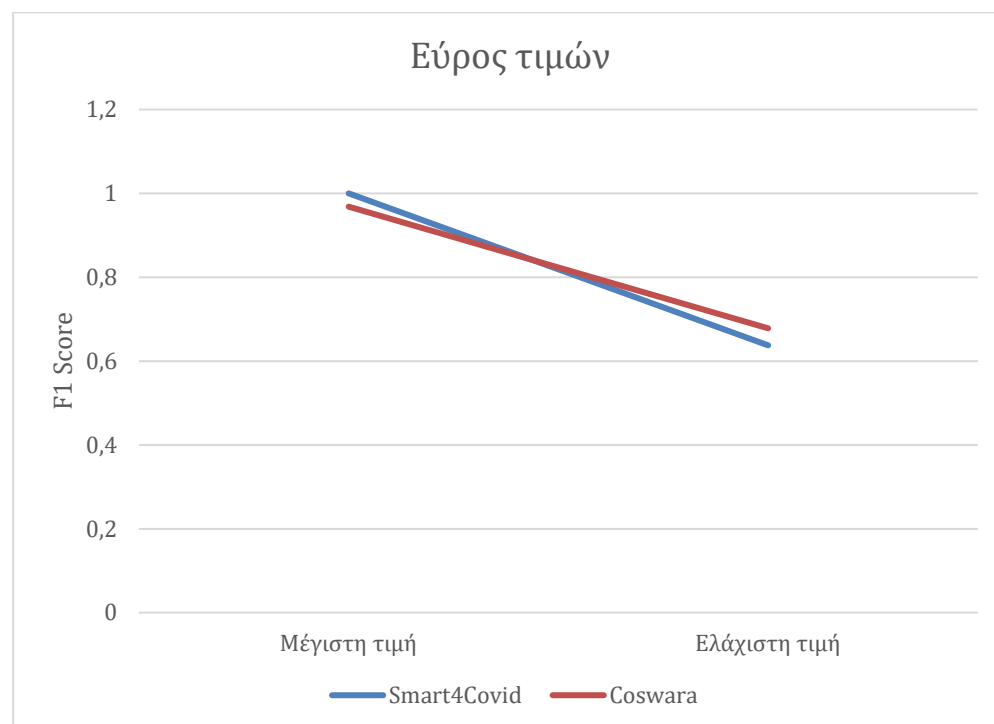
4.1.3. Σύγκριση αποτελεσμάτων από τις δυο βάσεις δεδομένων

Στον επόμενο πίνακα γίνεται σύγκριση των αποτελεσμάτων με βάση δεδομένων Smart4Covid & Coswara. Παρατηρείται τα αποτελέσματα συμπίπτουν σε όλους τους τομείς και παρουσιάζουν μια συνέπεια ανεξαρτήτως της βάσης των δεδομένων.

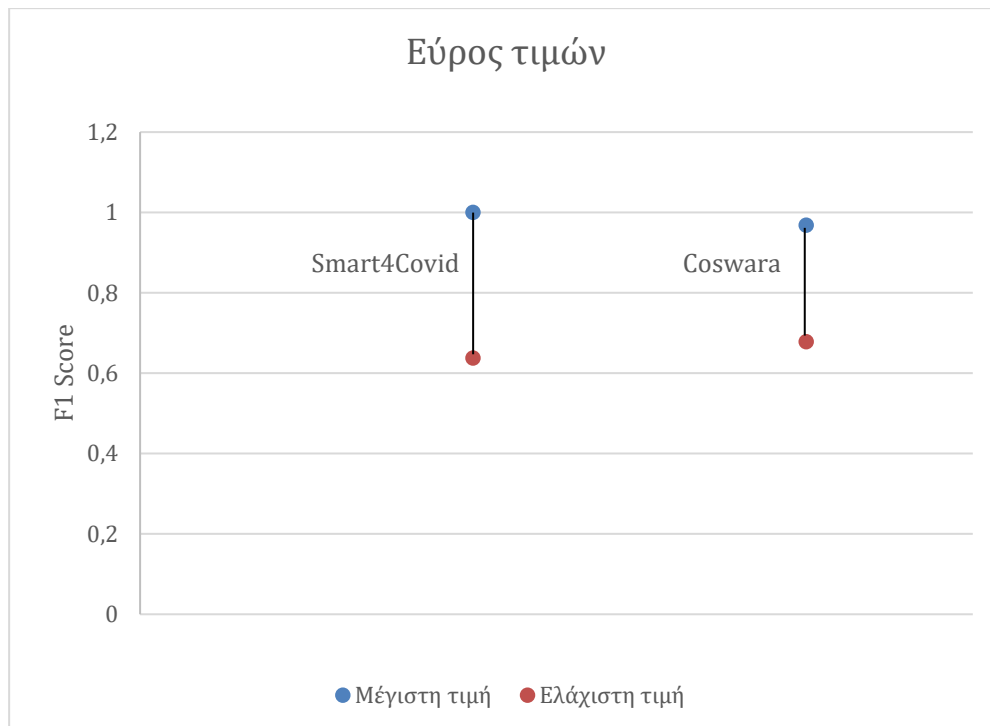
Στοιχεία	Αποτελέσματα	
	Smart4Covid	Coswara
Μέγιστη τιμή	1,000000000	0,968253968
Ελάχιστη τιμή	0,637517630	0,678111588
Διακύμανση	0,006752563	0,004732127
Μέσος Όρος	0,869045235	0,853151746
Διάμεσος	0,871491876	0,843177189
Επικρατούσα Τιμή	0,885007278	0,911242604

Πίνακας 15: Στατιστικά στοιχεία πίνακα σύγκρισης αποτελεσμάτων εκπαίδευσης ταξινομητή LR

Σχηματικά τα ανωτέρω συναγόμενα συμπεράσματα παρουσιάζονται καλύτερα με τα ακόλουθα συγκριτικά διαγράμματα.



Διάγραμμα 11: Οριζόντιο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή LR



Διάγραμμα 12: Κάθετο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή LR

Στον επόμενο πίνακα, με γνώμονα την εύρεση του βέλτιστου συνδυασμού `n_mfcc`, `frame_size` & `n_segments`, ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, γίνεται σύγκριση των τιμών `frame_size` & `n_segments` ανά `n_mfcc` καθώς και των μέγιστων τιμών αποτελεσμάτων (Max F1 Score value per `frame_size` & `n_segments`) και για τις δυο βάσεις δεδομένων (Smart4Covid & Coswara).

n_mfcc	Smart4Covid		Coswara	
	frame_size, n_segments	F1 Score	frame_size, n_segments	F1 Score
14	256,100	0,904899135	256,50	0,861538462
28	256,70	0,994117647	1024,150	0,933333333
42	512,150	0,982404692	1024,100	0,968253968
56	512,150	0,982404692	1024,100	0,968253968
70	512,150	0,982404692	1024,100	0,968253968

Πίνακας 16: Max F1 Score value per `frame_size` & `n_segments` αποτελεσμάτων εκπαίδευσης ταξινομητή LR

Για να γίνει αξιόπιστη σύγκριση δεδομένων, κρίθηκε σκόπιμη η εύρεση του μέσου όρου του συνδυασμού `n_mfcc`, `frame_size` & `n_segments` ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων ώστε να αποφευχθούν τυχόν παραπλανητικές κορυφές αποτελεσμάτων. Τα αποτελέσματα της σύγκρισης τους και για τις δυο βάσεις δεδομένων (Smart4Covid & Coswara), παρουσιάζονται στον επόμενο πίνακα.

<code>n_mfcc</code>	Smart4Covid	Coswara
14	0,802664730	0,771313016
28	0,861889729	0,850008012
42	0,893557239	0,881479234
56	0,893557239	0,881479234
70	0,893557239	0,881479234

Πίνακας 17: Μέσος Όρος F1 Score value per frame_size & n_segments αποτελεσμάτων εκπαίδευσης ταξινομητή LR

Συμπέρασμα

Από τους παραπάνω πίνακες (Πίνακας Max F1 Score & MO F1 Score) συνάγονται τα ακόλουθα συμπεράσματα:

Με την βάση δεδομένων Smart4Covid το καλύτερο F1 Score το δίνει ο συνδυασμός `frame_size = 512` και `n_segments = 150` ανεξαρτήτως τιμής `n_mfcc`.

Με την βάση δεδομένων Coswara το καλύτερο F1 Score το δίνει ο συνδυασμός `frame_size = 1024` και `n_segments = 100` ανεξαρτήτως τιμής `n_mfcc`.

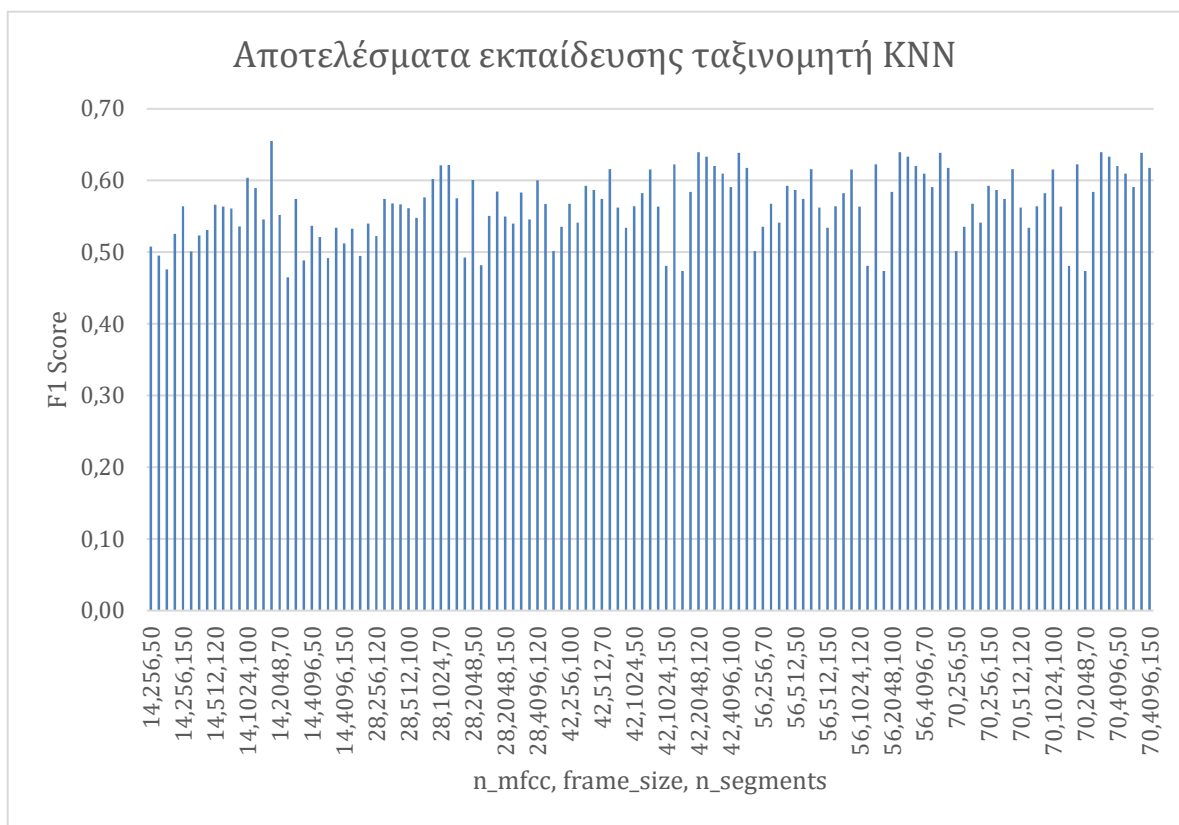
4.2. Αποτελέσματα εκπαίδευσης ταξινομητή K-Nearest Neighbors

4.2.1. Εκπαίδευση με βάση δεδομένων Smart4Covid

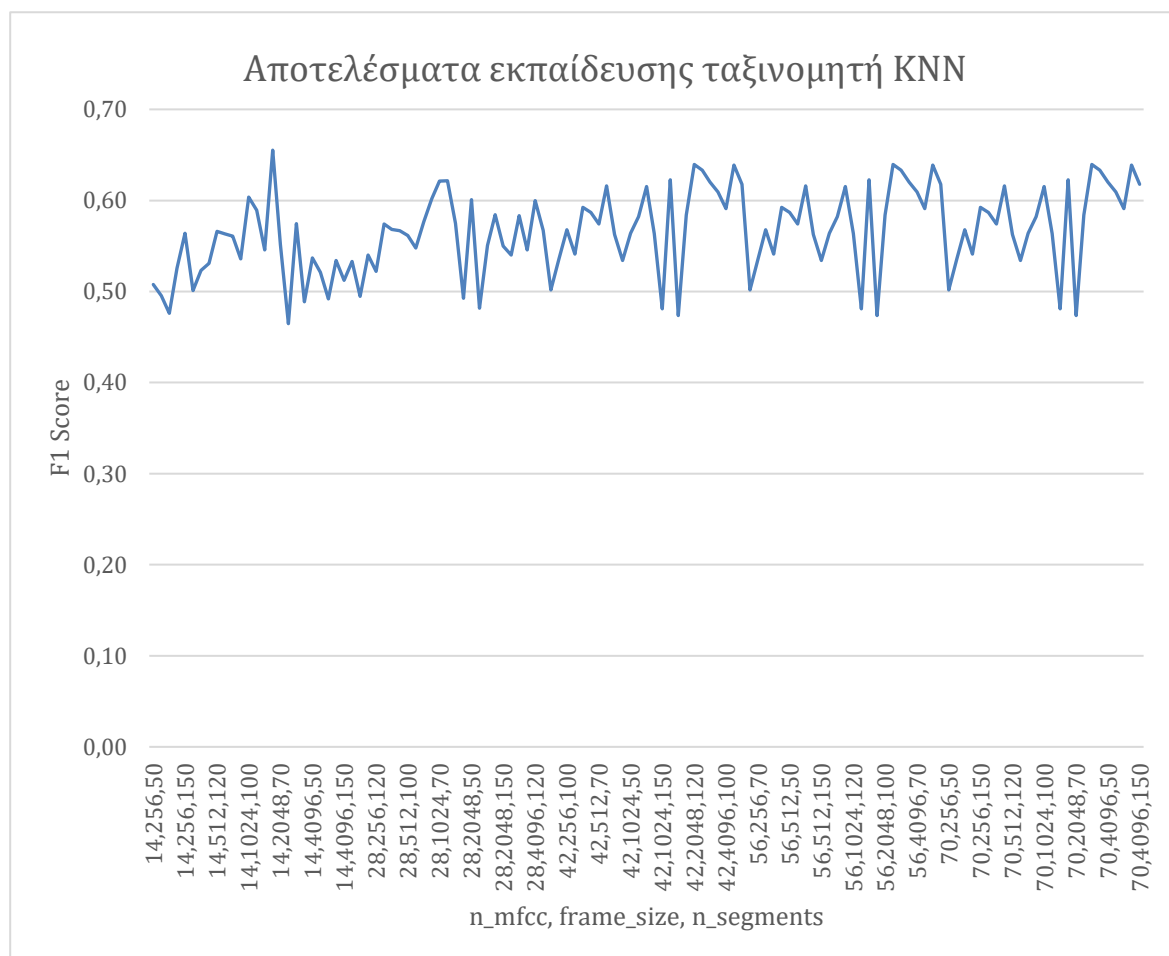
Το πρόγραμμα εκπαίδευσης ταξινομητών, χρησιμοποίησε την βάση δεδομένων της εφαρμογής **Smart4Covid**, εφαρμόζοντας συνδυαστικές ρυθμίσεις υπερπαραμέτρων για τον τύπο ταξινομητή **K-Nearest Neighbors (KNN)**. Τα αποτελέσματα του παρουσιάζονται στους επόμενους πίνακες & διαγράμματα, εν συνεχεία αναλύονται και συνάγονται χρήσιμα στατιστικά στοιχεία.

n_mfcc, frame_size, n_segments	F1 Score	n_mfcc, frame_size, n_segments	F1 Score	n_mfcc, frame_size, n_segments	F1 Score
14,256,50	0,507563025	28,2048,100	0,550544323	56,512,150	0,534201954
14,256,70	0,495274102	28,2048,120	0,584527221	56,1024,50	0,563953488
14,256,100	0,476021314	28,2048,150	0,549777117	56,1024,70	0,582352941
14,256,120	0,525806452	28,4096,50	0,540031397	56,1024,100	0,615384615
14,256,150	0,564102564	28,4096,70	0,583333333	56,1024,120	0,563517915
14,512,50	0,500881834	28,4096,100	0,545736434	56,1024,150	0,481054366
14,512,70	0,523333333	28,4096,120	0,600000000	56,2048,50	0,622641509
14,512,100	0,531001590	28,4096,150	0,567208272	56,2048,70	0,473684211
14,512,120	0,566096423	42,256,50	0,501694915	56,2048,100	0,583941606
14,512,150	0,563380282	42,256,70	0,535537190	56,2048,120	0,639490446
14,1024,50	0,560747664	42,256,100	0,567741935	56,2048,150	0,633288227
14,1024,70	0,535825545	42,256,120	0,541254125	56,4096,50	0,620320856
14,1024,100	0,603658537	42,256,150	0,592471358	56,4096,70	0,609395973
14,1024,120	0,589285714	42,512,50	0,586709887	56,4096,100	0,590909091
14,1024,150	0,545751634	42,512,70	0,574162679	56,4096,120	0,638820639
14,2048,50	0,655226209	42,512,100	0,616058394	56,4096,150	0,617842876
14,2048,70	0,551845343	42,512,120	0,562401264	70,256,50	0,501694915
14,2048,100	0,464716007	42,512,150	0,534201954	70,256,70	0,535537190
14,2048,120	0,574436090	42,1024,50	0,563953488	70,256,100	0,567741935
14,2048,150	0,488599349	42,1024,70	0,582352941	70,256,120	0,541254125
14,4096,50	0,536809816	42,1024,100	0,615384615	70,256,150	0,592471358
14,4096,70	0,521008403	42,1024,120	0,563517915	70,512,50	0,586709887
14,4096,100	0,491749175	42,1024,150	0,481054366	70,512,70	0,574162679
14,4096,120	0,534142640	42,2048,50	0,622641509	70,512,100	0,616058394
14,4096,150	0,512234910	42,2048,70	0,473684211	70,512,120	0,562401264
28,256,50	0,532883642	42,2048,100	0,583941606	70,512,150	0,534201954
28,256,70	0,494845361	42,2048,120	0,639490446	70,1024,50	0,563953488
28,256,100	0,540069686	42,2048,150	0,633288227	70,1024,70	0,582352941
28,256,120	0,522314050	42,4096,50	0,620320856	70,1024,100	0,615384615
28,256,150	0,574257426	42,4096,70	0,609395973	70,1024,120	0,563517915
28,512,50	0,568027211	42,4096,100	0,590909091	70,1024,150	0,481054366
28,512,70	0,566878981	42,4096,120	0,638820639	70,2048,50	0,622641509
28,512,100	0,561403509	42,4096,150	0,617842876	70,2048,70	0,473684211
28,512,120	0,547811994	56,256,50	0,501694915	70,2048,100	0,583941606
28,512,150	0,576489533	56,256,70	0,535537190	70,2048,120	0,639490446
28,1024,50	0,601851852	56,256,100	0,567741935	70,2048,150	0,633288227
28,1024,70	0,621329212	56,256,120	0,541254125	70,4096,50	0,620320856
28,1024,100	0,621621622	56,256,150	0,592471358	70,4096,70	0,609395973
28,1024,120	0,575000000	56,512,50	0,586709887	70,4096,100	0,590909091
28,1024,150	0,492659054	56,512,70	0,574162679	70,4096,120	0,638820639
28,2048,50	0,600870827	56,512,100	0,616058394	70,4096,150	0,617842876
28,2048,70	0,481802426	56,512,120	0,562401264		

Πίνακας 18: Αποτελέσματα εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments



Διάγραμμα 13: Ραβδόγραμμα αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments



Διάγραμμα 14: Διάγραμμα αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments

Εφαρμόζοντας τυπική στατιστική ανάλυση στα δεδομένα του ανωτέρω πίνακα συνάγονται τα επόμενα χρήσιμα στατιστικά στοιχεία.

Στοιχεία	Τιμές
Μέγιστη τιμή	0,655226209
Ελάχιστη τιμή	0,464716007
Διακύμανση	0,002150535
Μέσος Όρος	0,566138159
Διάμεσος	0,567741935
Επικρατούσα Τιμή	0,501694915

Πίνακας 19: Στατιστικά στοιχεία πίνακα αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid

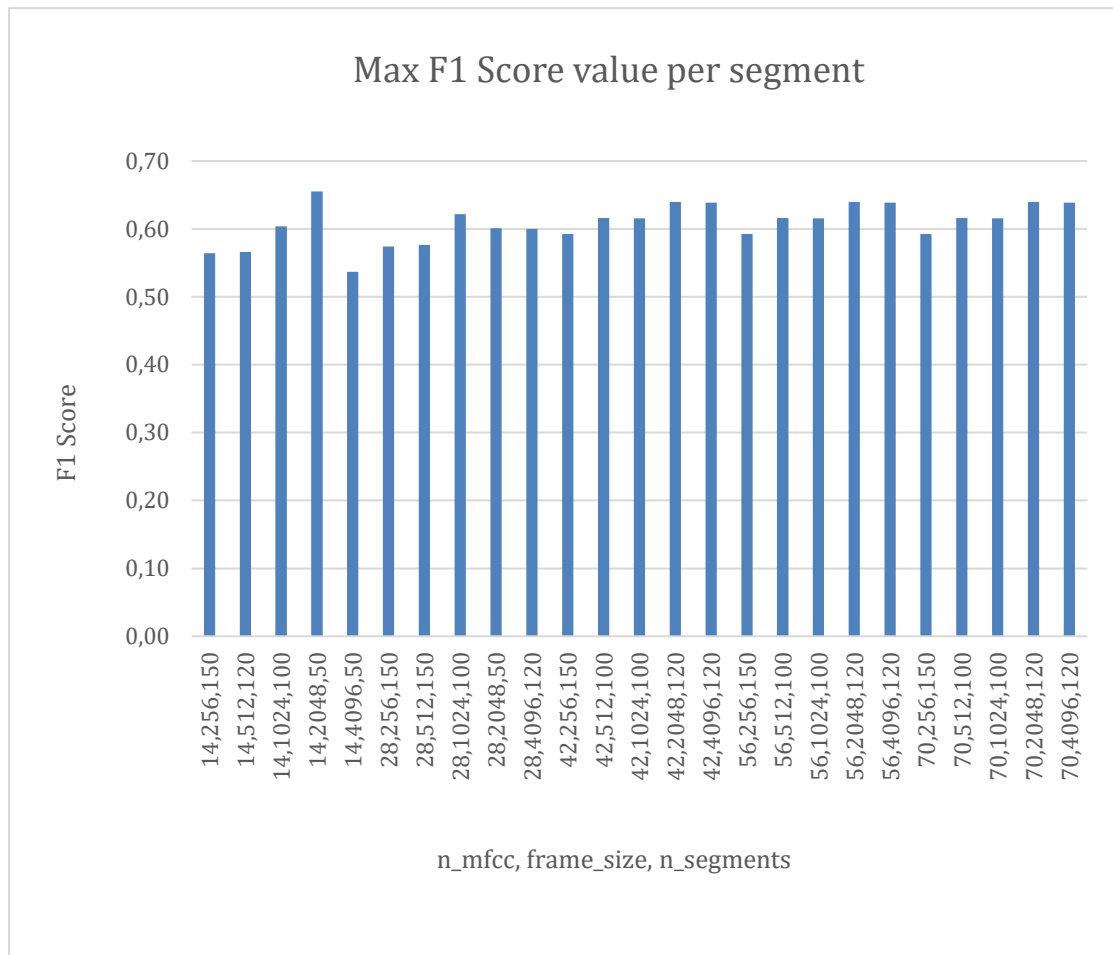
Παρατηρείται ότι η μέγιστη τιμή είναι $\approx 0,6552$ δηλαδή 65,52% επιτυχία στην πρόβλεψη, η ελάχιστη τιμή είναι $\approx 0,4646$ δηλαδή 46,46 % επιτυχία στην πρόβλεψη, με διακύμανση τιμών $\approx 0,0022$. Ο μέσος όρος των αποτελεσμάτων είναι $\approx 0,5661$ δηλαδή 56,61 % επιτυχία στην πρόβλεψη, με διάμεση τιμή $\approx 0,5677$ (56,77 %) και επικρατούσα τιμή $\approx 0,5017$ (50,17 %). Συνάγεται από τις ανωτέρω τιμές της στατιστικής ανάλυσης, **με το σύνολο δεδομένων εκπαίδευσης smarty4covid**, ότι το ποσοστό επιτυχίας της πρόβλεψης για το αν ένα άτομο πάσχει από COVID-19 είναι ικανοποιητικό.

Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού n_mfcc, frame_size & n_segments ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές n_segment ανά συνδυασμό n_mfcc & frame_size καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per segment).

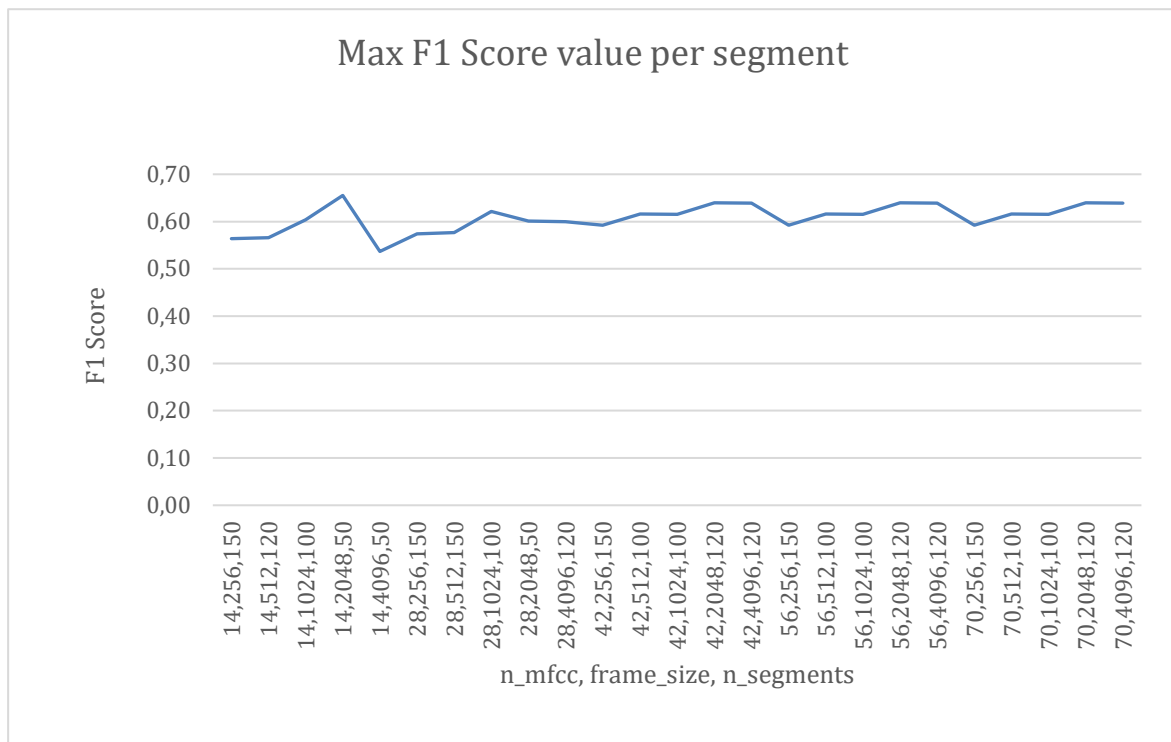
n_mfcc, frame_size, n_segments	Max F1 Score value per segment
14,256,150	0,564102564
14,512,120	0,566096423
14,1024,100	0,603658537
14,2048,50	0,655226209
14,4096,50	0,536809816
28,256,150	0,574257426
28,512,150	0,576489533
28,1024,100	0,621621622
28,2048,50	0,600870827
28,4096,120	0,600000000
42,256,150	0,592471358
42,512,100	0,616058394
42,1024,100	0,615384615
42,2048,120	0,639490446

42,4096,120	0,638820639
56,256,150	0,592471358
56,512,100	0,616058394
56,1024,100	0,615384615
56,2048,120	0,639490446
56,4096,120	0,638820639
70,256,150	0,592471358
70,512,100	0,616058394
70,1024,100	0,615384615
70,2048,120	0,639490446
70,4096,120	0,638820639

Πίνακας 20: Max F1 Score value per segment αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid



Διάγραμμα 15: Ραβδόγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments



Διάγραμμα 16: Διάγραμμα μέγιστων τιμών F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size, n_segments

Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού n_mfcc, frame_size & n_segments ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές frame_size & n_segments ανά n_mfcc καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per frame_size & n_segments).

n_mfcc, frame_size, n_segments	Max F1 Score value per frame_size & n_segments
14,2048,50	0,655226209
28,1024,100	0,621621622
42,2048,120	0,639490446
56,2048,120	0,639490446
70,2048,120	0,639490446

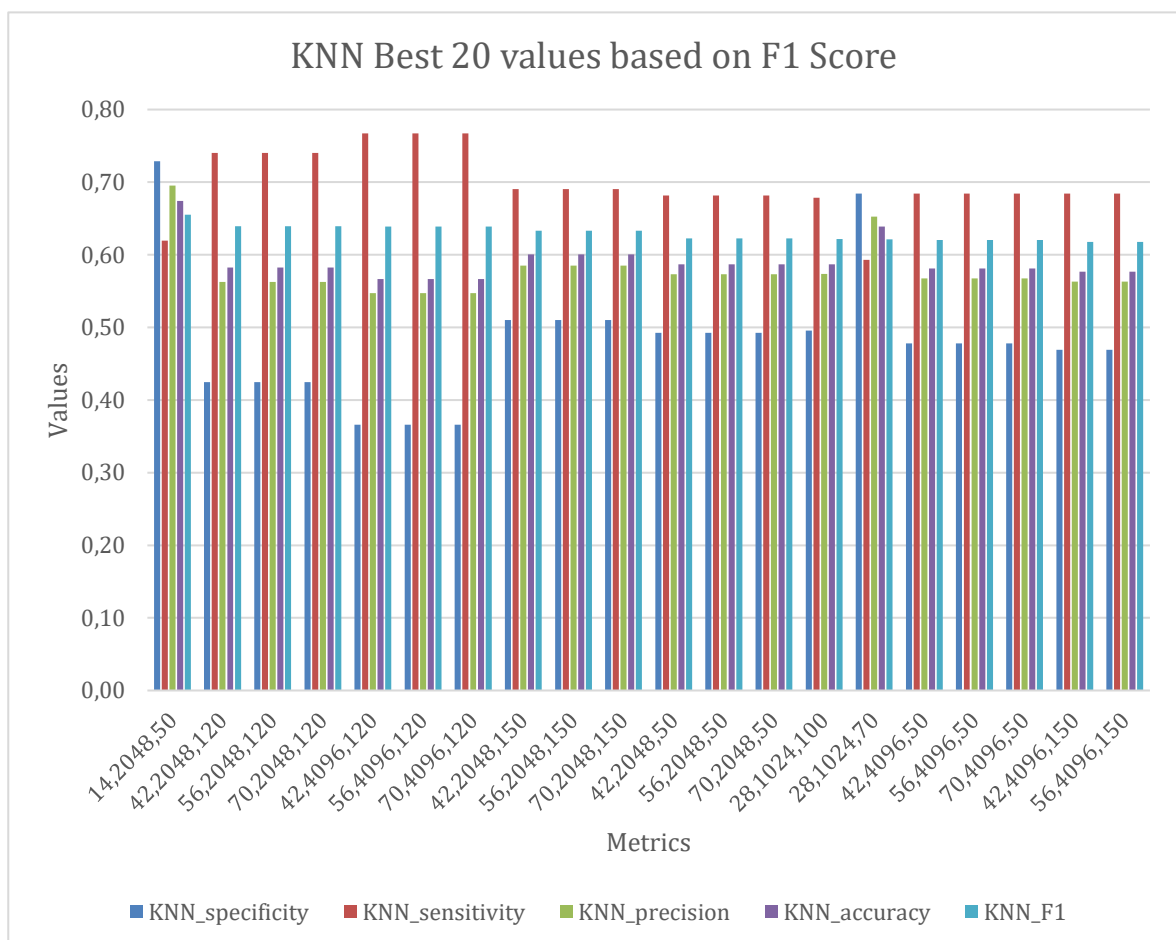
Πίνακας 21: Max F1 Score value per frame_size & n_segments αποτελεσμάτων εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid

Επιπρόσθετα στην ανάλυση παρατίθενται στον επόμενο πίνακα οι 20 μεγαλύτερες τιμές αποτελεσμάτων με βάση το F1 Score (Best 20 values based on F1 Score) με συνδυασμό n_mfcc, frame_size & n_segment. Στον πίνακα επιπλέον παρατίθενται και οι υπόλοιπες μετρικές για κάθε συνδυασμό.

n_mfcc, frame_size, n_segments	Specificity	Sensitivity	Precision	Accuracy	F1 Score
14,2048,50	0,728613569	0,619469027	0,695364238	0,674041298	0,655226209
42,2048,120	0,424778761	0,740412979	0,562780269	0,582595870	0,639490446
56,2048,120	0,424778761	0,740412979	0,562780269	0,582595870	0,639490446

70,2048,120	0,424778761	0,740412979	0,562780269	0,582595870	0,639490446
42,4096,120	0,365781711	0,766961652	0,547368421	0,566371681	0,638820639
56,4096,120	0,365781711	0,766961652	0,547368421	0,566371681	0,638820639
70,4096,120	0,365781711	0,766961652	0,547368421	0,566371681	0,638820639
42,2048,150	0,510324484	0,690265487	0,585000000	0,600294985	0,633288227
56,2048,150	0,510324484	0,690265487	0,585000000	0,600294985	0,633288227
70,2048,150	0,510324484	0,690265487	0,585000000	0,600294985	0,633288227
42,2048,50	0,492625369	0,681415929	0,573200993	0,587020649	0,622641509
56,2048,50	0,492625369	0,681415929	0,573200993	0,587020649	0,622641509
70,2048,50	0,492625369	0,681415929	0,573200993	0,587020649	0,622641509
28,1024,100	0,495575221	0,678466077	0,573566085	0,587020649	0,621621622
28,1024,70	0,684365782	0,592920354	0,652597403	0,638643068	0,621329212
42,4096,50	0,477876106	0,684365782	0,567237164	0,581120944	0,620320856
56,4096,50	0,477876106	0,684365782	0,567237164	0,581120944	0,620320856
70,4096,50	0,477876106	0,684365782	0,567237164	0,581120944	0,620320856
42,4096,150	0,469026549	0,684365782	0,563106796	0,576696165	0,617842876
56,4096,150	0,469026549	0,684365782	0,563106796	0,576696165	0,617842876

Πίνακας 22: Best 20 values based on F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid



Διάγραμμα 17: Ραβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων smarty4covid (Best 20 values based on F1 Score)

Συμπέρασμα

Παρατηρείται ότι οι μεγαλύτερες τιμές (F1 Score) εμφανίζονται όταν υπάρχει συνδυασμός μεγάλων τιμών `n_mfcc` (42, 56 & 70), μεγάλων τιμών `frame_size` (2048 & 4096) & μεγάλων τιμών `n_segments` (120 & 150) αν και η μεγαλύτερη τιμή (`performance_metrics_knn_F1`) προκύπτει από συνδυασμό μικρής τιμής `n_mfcc` (14), μεγάλης τιμής `frame_size` (2048) και μικρής τιμής `n_segments` (50).

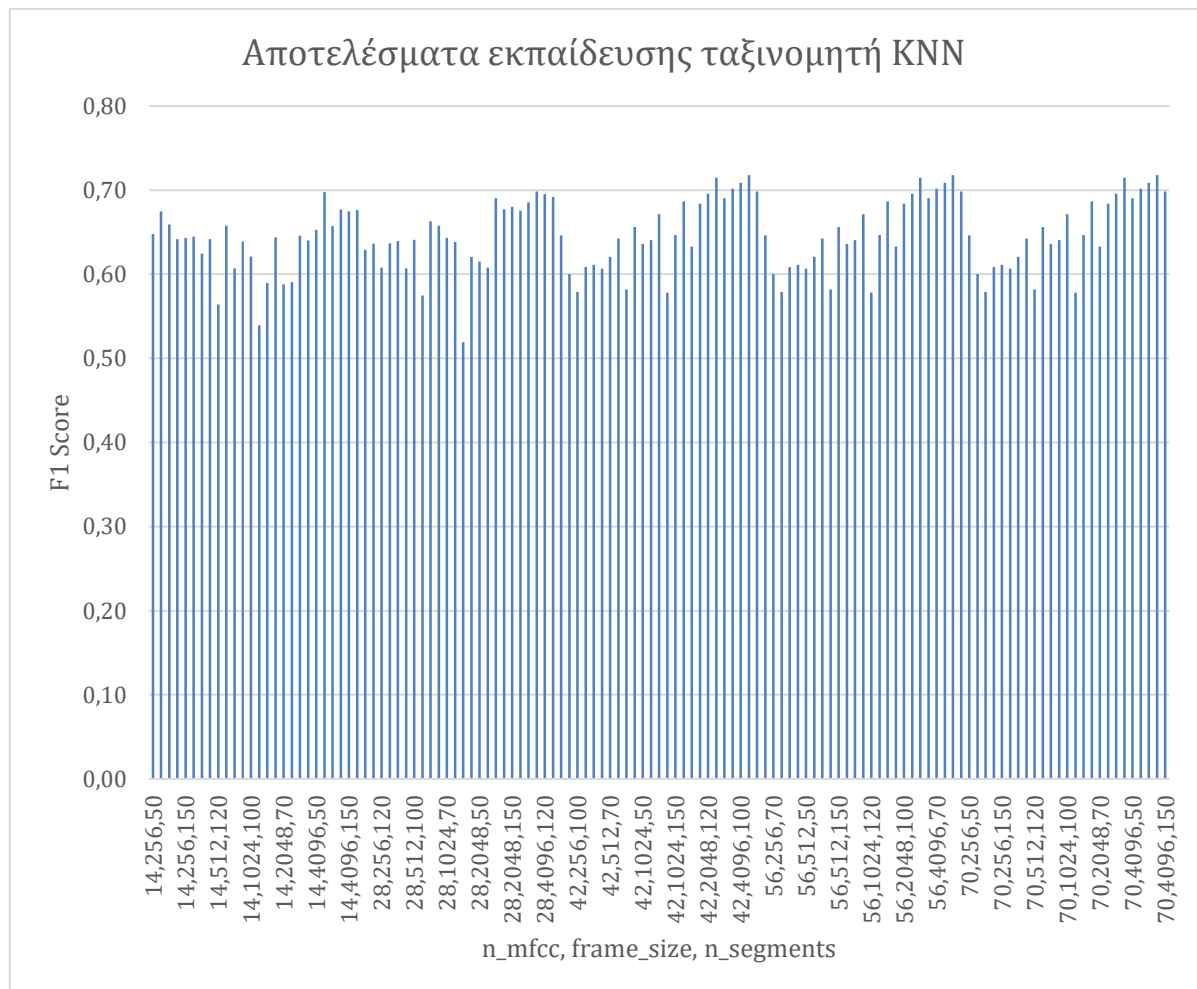
4.2.2. Εκπαίδευση με βάση δεδομένων Coswara

Το πρόγραμμα εκπαίδευσης ταξινομητών, χρησιμοποίησε την βάση δεδομένων της εφαρμογής **Coswara**, εφαρμόζοντας συνδυαστικές ρυθμίσεις υπερπαραμέτρων για τον τύπο ταξινομητή **K-Nearest Neighbors (KNN)**. Τα αποτελέσματα του παρουσιάζονται στους επόμενους πίνακες & διαγράμματα, εν συνεχεία αναλύονται και συνάγονται χρήσιμα στατιστικά στοιχεία.

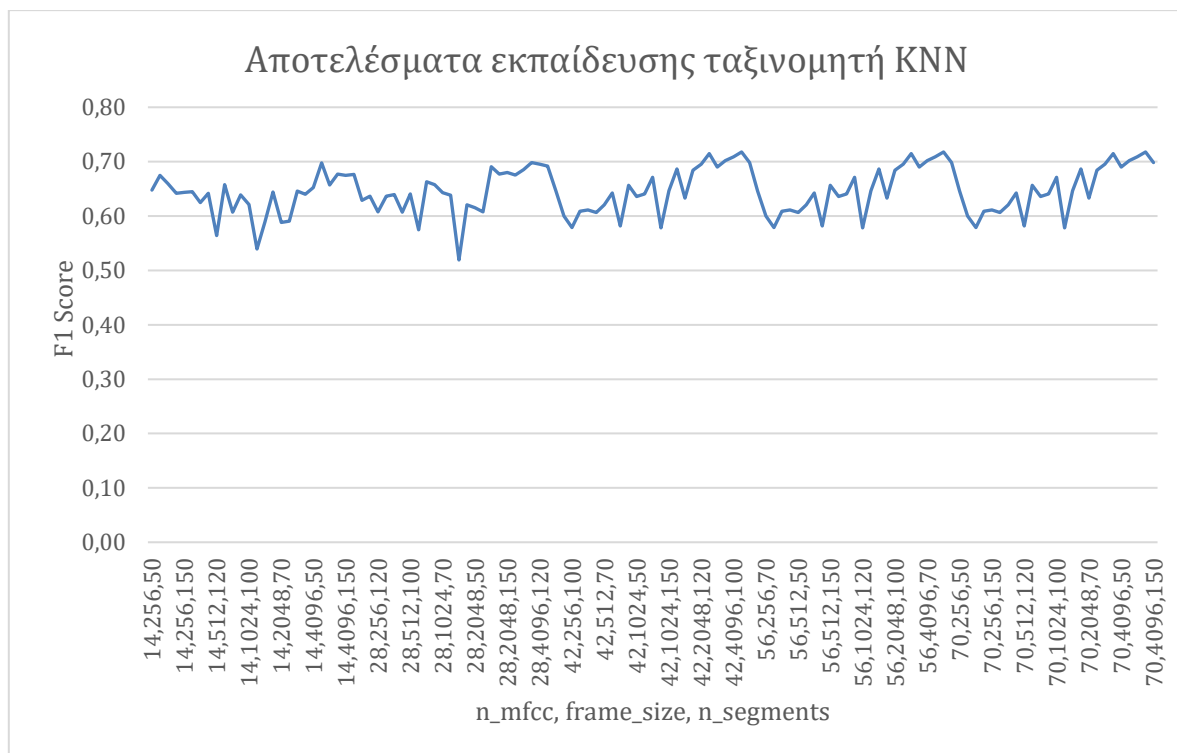
n_mfcc, frame_size, n_segments	F1 Score	n_mfcc, frame_size, n_segments	F1 Score	n_mfcc, frame_size, n_segments	F1 Score
14,256,50	0,647746244	28,2048,100	0,690518784	56,512,150	0,656303972
14,256,70	0,674735250	28,2048,120	0,677192982	56,1024,50	0,635838150
14,256,100	0,659200000	28,2048,150	0,680067002	56,1024,70	0,640449438
14,256,120	0,641711230	28,4096,50	0,675540765	56,1024,100	0,671428571
14,256,150	0,643356643	28,4096,70	0,685534591	56,1024,120	0,577962578
14,512,50	0,644628099	28,4096,100	0,698412698	56,1024,150	0,646536412
14,512,70	0,624561404	28,4096,120	0,695384615	56,2048,50	0,686514886
14,512,100	0,641732283	28,4096,150	0,691729323	56,2048,70	0,632812500
14,512,120	0,564102564	42,256,50	0,646209386	56,2048,100	0,683937824
14,512,150	0,657587549	42,256,70	0,600000000	56,2048,120	0,695652174
14,1024,50	0,606924644	42,256,100	0,578723404	56,2048,150	0,714723926
14,1024,70	0,638888889	42,256,120	0,608695652	56,4096,50	0,690184049
14,1024,100	0,620967742	42,256,150	0,611336032	56,4096,70	0,701966717
14,1024,120	0,539473684	42,512,50	0,606694561	56,4096,100	0,708782743
14,1024,150	0,589690722	42,512,70	0,620689655	56,4096,120	0,717791411
14,2048,50	0,643939394	42,512,100	0,642585551	56,4096,150	0,698224852
14,2048,70	0,587982833	42,512,120	0,581740977	70,256,50	0,646209386
14,2048,100	0,590631365	42,512,150	0,656303972	70,256,70	0,600000000
14,2048,120	0,645669291	42,1024,50	0,635838150	70,256,100	0,578723404
14,2048,150	0,640151515	42,1024,70	0,640449438	70,256,120	0,608695652
14,4096,50	0,652671756	42,1024,100	0,671428571	70,256,150	0,611336032
14,4096,70	0,697495183	42,1024,120	0,577962578	70,512,50	0,606694561
14,4096,100	0,657193606	42,1024,150	0,646536412	70,512,70	0,620689655
14,4096,120	0,677137871	42,2048,50	0,686514886	70,512,100	0,642585551

14,4096,150	0,674657534	42,2048,70	0,632812500	70,512,120	0,581740977
28,256,50	0,676384840	42,2048,100	0,683937824	70,512,150	0,656303972
28,256,70	0,628930818	42,2048,120	0,695652174	70,1024,50	0,635838150
28,256,100	0,636363636	42,2048,150	0,714723926	70,1024,70	0,640449438
28,256,120	0,607709751	42,4096,50	0,690184049	70,1024,100	0,671428571
28,256,150	0,636734694	42,4096,70	0,701966717	70,1024,120	0,577962578
28,512,50	0,639215686	42,4096,100	0,708782743	70,1024,150	0,646536412
28,512,70	0,606924644	42,4096,120	0,717791411	70,2048,50	0,686514886
28,512,100	0,640776699	42,4096,150	0,698224852	70,2048,70	0,632812500
28,512,120	0,574610245	56,256,50	0,646209386	70,2048,100	0,683937824
28,512,150	0,663043478	56,256,70	0,600000000	70,2048,120	0,695652174
28,1024,50	0,657587549	56,256,100	0,578723404	70,2048,150	0,714723926
28,1024,70	0,642998028	56,256,120	0,608695652	70,4096,50	0,690184049
28,1024,100	0,638297872	56,256,150	0,611336032	70,4096,70	0,701966717
28,1024,120	0,519187359	56,512,50	0,606694561	70,4096,100	0,708782743
28,1024,150	0,620825147	56,512,70	0,620689655	70,4096,120	0,717791411
28,2048,50	0,615087041	56,512,100	0,642585551	70,4096,150	0,698224852
28,2048,70	0,607843137	56,512,120	0,581740977		

Πίνακας 23: Αποτελέσματα εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments.



Διάγραμμα 18: Ραβδόγραμμα αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments



Διάγραμμα 19: Διάγραμμα αποτελεσμάτων εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc , $frame_size$, $n_segments$

Εφαρμόζοντας τυπική στατιστική ανάλυση στα δεδομένα του ανωτέρω πίνακα συνάγονται τα επόμενα χρήσιμα στατιστικά στοιχεία.

Στοιχεία	Τιμές
Μέγιστη τιμή	0,717791411
Ελάχιστη τιμή	0,519187359
Διακύμανση	0,001810029
Μέσος Όρος	0,645896760
Διάμεσος	0,643356643
Επικρατούσα Τιμή	0,646209386

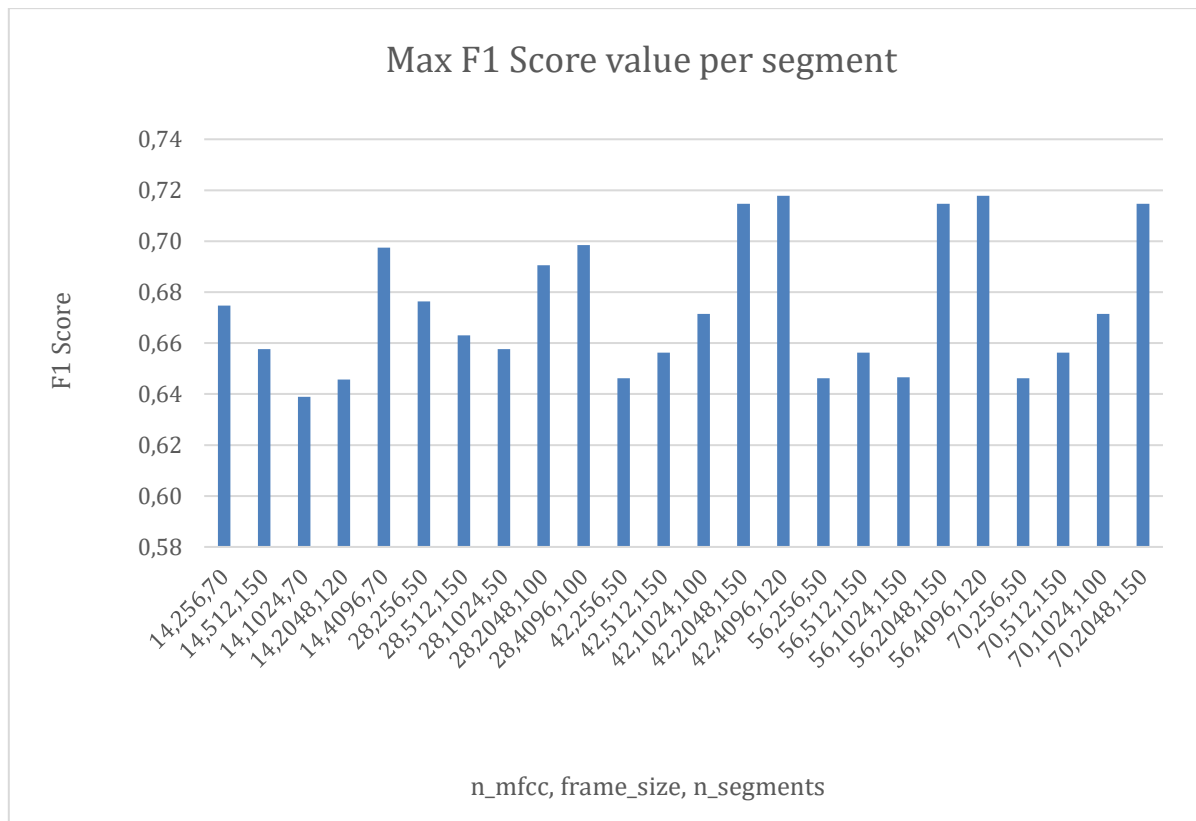
Πίνακας 24: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara

Παρατηρείται ότι η μέγιστη τιμή είναι $\approx 0,7178$ δηλαδή 71,78% επιτυχία στην πρόβλεψη, η ελάχιστη τιμή είναι $\approx 0,5192$ δηλαδή 51,92 % επιτυχία στην πρόβλεψη, με διακύμανση τιμών $\approx 0,0018$. Ο μέσος όρος των αποτελεσμάτων είναι $\approx 0,6459$ δηλαδή 64,59 % επιτυχία στην πρόβλεψη, με διάμεση τιμή $\approx 0,6434$ (64,34 %) και επικρατούσα τιμή $\approx 0,6462$ (64,62 %). Συνάγεται από τις ανωτέρω τιμές της στατιστικής ανάλυσης, **με το σύνολο δεδομένων εκπαίδευσης Coswara**, ότι το ποσοστό επιτυχίας της πρόβλεψης για το αν ένα άτομο πάσχει από COVID-19 είναι αρκετά ικανοποιητικό.

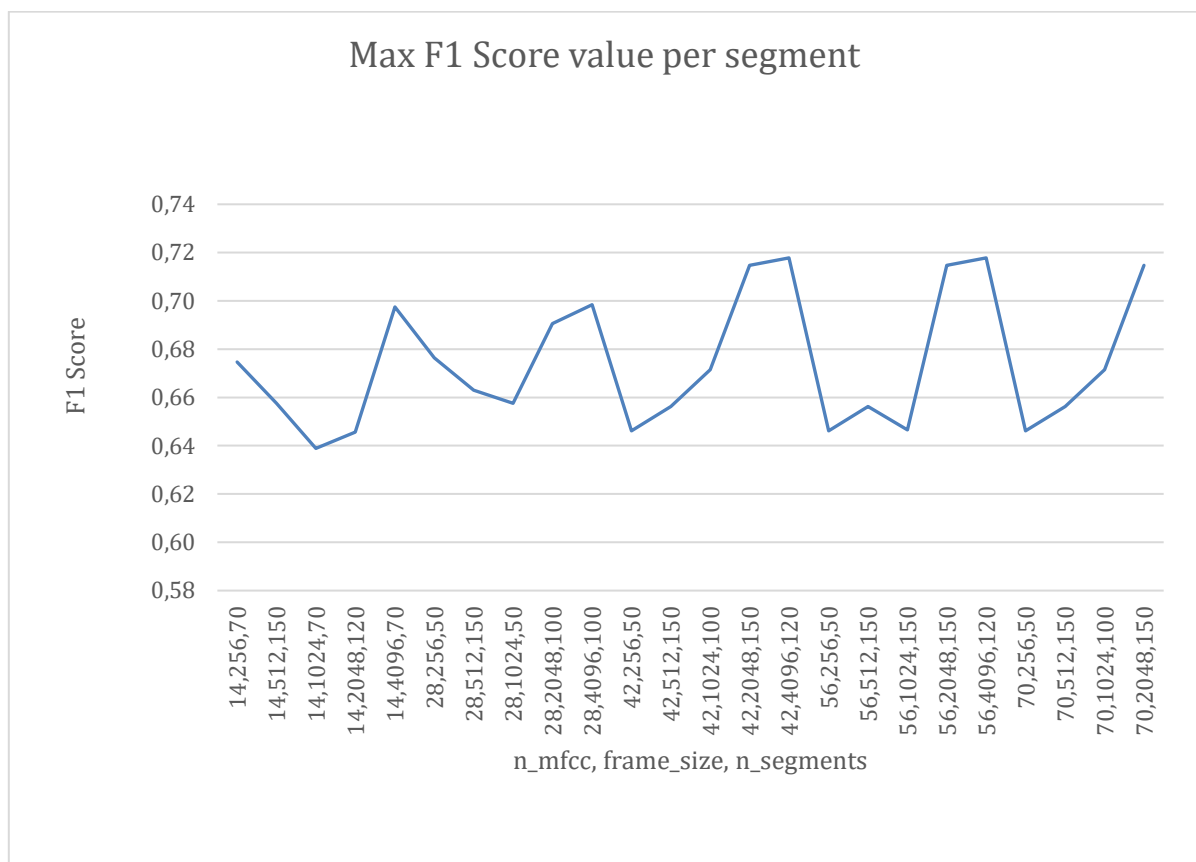
Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού n_mfcc, frame_size & n_segments ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές n_segment ανά συνδυασμό n_mfcc & frame_size καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per segment).

n_mfcc, frame_size, n_segments	Max F1 Score value per segment
14,256,70	0,674735250
14,512,150	0,657587549
14,1024,70	0,638888889
14,2048,120	0,645669291
14,4096,70	0,697495183
28,256,50	0,676384840
28,512,150	0,663043478
28,1024,50	0,657587549
28,2048,100	0,690518784
28,4096,100	0,698412698
42,256,50	0,646209386
42,512,150	0,656303972
42,1024,100	0,671428571
42,2048,150	0,714723926
42,4096,120	0,717791411
56,256,50	0,646209386
56,512,150	0,656303972
56,1024,150	0,646536412
56,2048,150	0,714723926
56,4096,120	0,717791411
70,256,50	0,646209386
70,512,150	0,656303972
70,1024,100	0,671428571
70,2048,150	0,714723926
70,4096,120	0,717791411

Πίνακας 25: Max F1 Score value per segment αποτελέσματος εκπαίδευσης ταξινόμητη KNN στο σύνολο δεδομένων coswara



Διάγραμμα 20: Ραβδόγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments



Διάγραμμα 21: Διάγραμμα μέγιστων τιμών F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size, n_segments

Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού n_mfcc, frame_size & n_segments ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές frame_size & n_segments ανά n_mfcc καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per frame_size & n_segments).

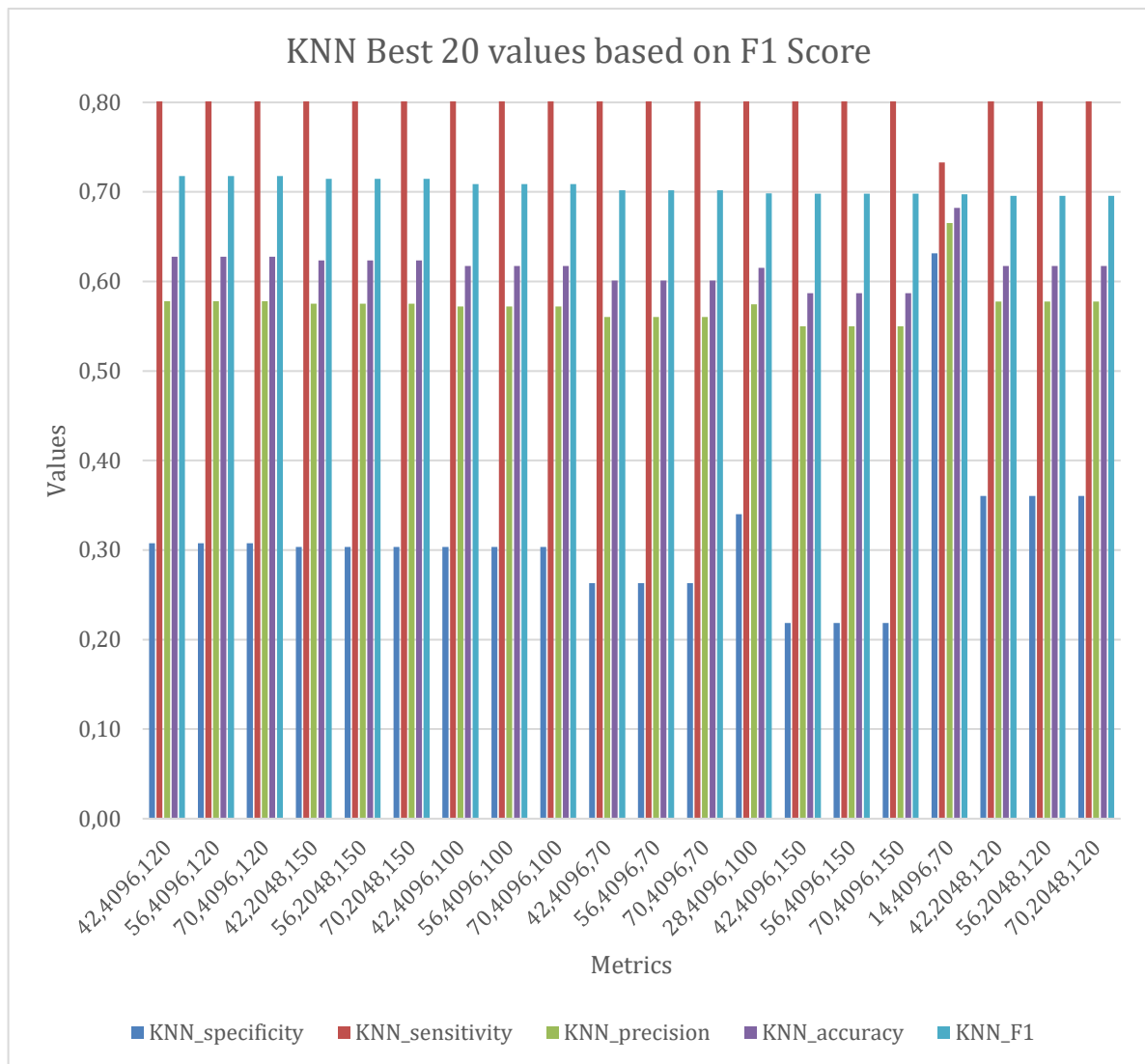
n_mfcc, frame_size, n_segments	Max F1 Score value per frame_size & n_segments
14,4096,70	0,697495183
28,4096,100	0,698412698
42,4096,120	0,717791411
56,4096,120	0,717791411
70,4096,120	0,717791411

Πίνακας 26: Max F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara

Επιπρόσθετα στην ανάλυση παρατίθενται στον επόμενο πίνακα οι 20 μεγαλύτερες τιμές αποτελεσμάτων με βάση το F1 Score (Best 20 values based on F1 Score) με συνδυασμό n_mfcc, frame_size & n_segment. Στον πίνακα επιπλέον παρατίθενται και οι υπόλοιπες μετρικές για κάθε συνδυασμό.

n_mfcc, frame_size, n_segments	Specificity	Sensitivity	Precision	Accuracy	F1 Score
42,4096,120	0,307692308	0,947368421	0,577777778	0,627530364	0,717791411
56,4096,120	0,307692308	0,947368421	0,577777778	0,627530364	0,717791411
70,4096,120	0,307692308	0,947368421	0,577777778	0,627530364	0,717791411
42,2048,150	0,303643725	0,943319838	0,575308642	0,623481781	0,714723926
56,2048,150	0,303643725	0,943319838	0,575308642	0,623481781	0,714723926
70,2048,150	0,303643725	0,943319838	0,575308642	0,623481781	0,714723926
42,4096,100	0,303643725	0,931174089	0,572139303	0,617408907	0,708782743
56,4096,100	0,303643725	0,931174089	0,572139303	0,617408907	0,708782743
70,4096,100	0,303643725	0,931174089	0,572139303	0,617408907	0,708782743
42,4096,70	0,263157895	0,939271255	0,560386473	0,601214575	0,701966717
56,4096,70	0,263157895	0,939271255	0,560386473	0,601214575	0,701966717
70,4096,70	0,263157895	0,939271255	0,560386473	0,601214575	0,701966717
28,4096,100	0,340080972	0,890688259	0,574412533	0,615384615	0,698412698
42,4096,150	0,218623482	0,955465587	0,550116550	0,587044534	0,698224852
56,4096,150	0,218623482	0,955465587	0,550116550	0,587044534	0,698224852
70,4096,150	0,218623482	0,955465587	0,550116550	0,587044534	0,698224852
14,4096,70	0,631578947	0,732793522	0,665441176	0,682186235	0,697495183
42,2048,120	0,360323887	0,874493927	0,577540107	0,617408907	0,695652174
56,2048,120	0,360323887	0,874493927	0,577540107	0,617408907	0,695652174
70,2048,120	0,360323887	0,874493927	0,577540107	0,617408907	0,695652174

Πίνακας 27: Best 20 values based on F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara



Διάγραμμα 22: Ραβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελέσματος εκπαίδευσης ταξινομητή KNN στο σύνολο δεδομένων coswara (Best 20 values based on F1 Score)

Συμπέρασμα

Παρατηρείται ότι οι μεγαλύτερες τιμές (F1 Score) εμφανίζονται όταν υπάρχει συνδυασμός μεγάλων τιμών n_mfcc (42, 56 & 70), μεγάλων τιμών $frame_size$ (2048 & 4096) & μεγάλων τιμών $n_segments$ (100, 20 & 150).

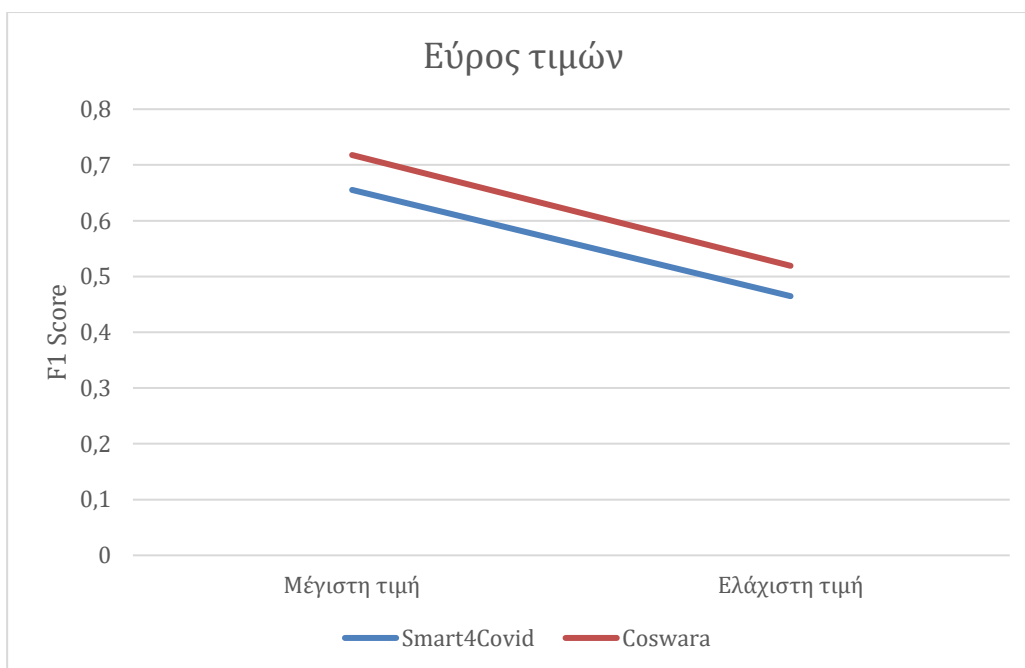
4.2.3. Σύγκριση αποτελεσμάτων από τις δυο βάσεις δεδομένων

Στον επόμενο πίνακα γίνεται σύγκριση των αποτελεσμάτων με βάση δεδομένων Smart4Covid & Coswara. Παρατηρείται τα αποτελέσματα συμπίπτουν σε όλους τους τομείς και παρουσιάζουν μια συνέπεια ανεξαρτήτως της βάσης των δεδομένων.

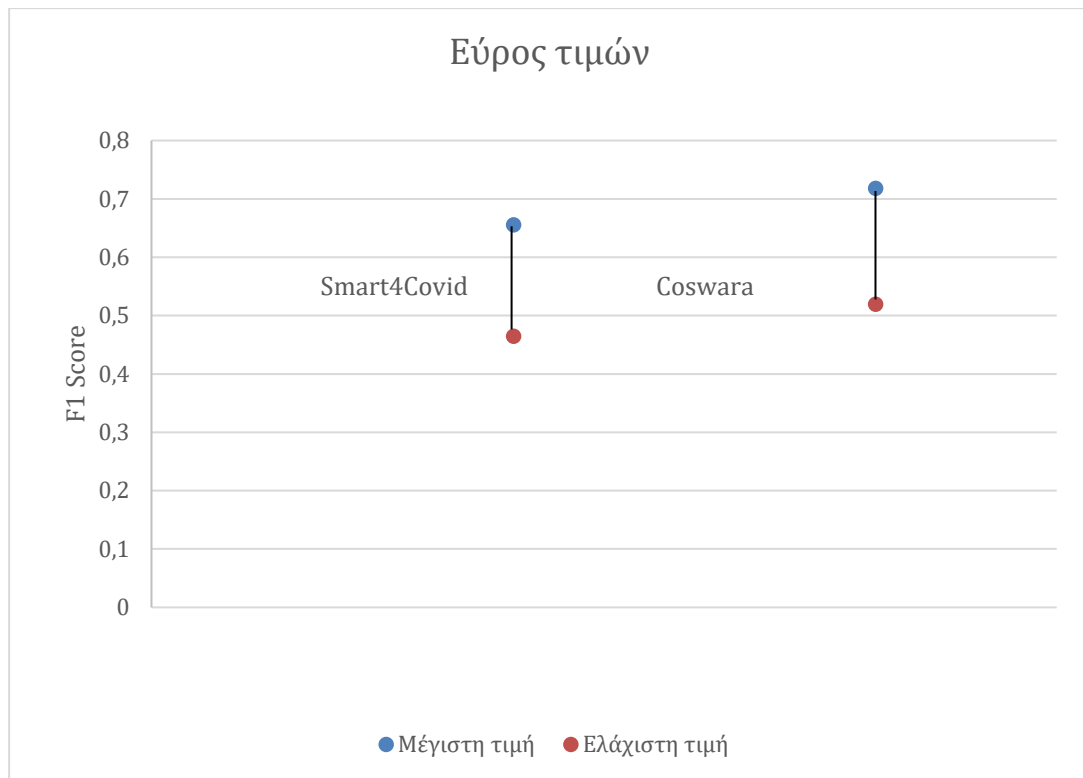
Στοιχεία	Αποτελέσματα	
	Smart4Covid	Coswara
Μέγιστη τιμή	0,655226209	0,717791411
Ελάχιστη τιμή	0,464716007	0,519187359
Διακύμανση	0,002150535	0,001810029
Μέσος Όρος	0,566138159	0,645896760
Διάμεσος	0,567741935	0,643356643
Επικρατούσα Τιμή	0,501694915	0,646209386

Πίνακας 28: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινόμητη KNN

Σχηματικά τα ανωτέρω συναγόμενα συμπεράσματα παρουσιάζονται καλύτερα με τα ακόλουθα συγκριτικά διαγράμματα.



Διάγραμμα 23: Οριζόντιο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινόμητη KNN



Διάγραμμα 24: Κάθετο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή KNN

Στον επόμενο πίνακα, με γνώμονα την εύρεση του βέλτιστου συνδυασμού `n_mfcc`, `frame_size` & `n_segments` ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, γίνεται σύγκριση των τιμών `frame_size` & `n_segments` ανά `n_mfcc` καθώς και των μέγιστων τιμών αποτελεσμάτων (Max F1 Score value per `frame_size` & `n_segments`) και για τις δυο βάσεις δεδομένων (Smart4Covid & Coswara).

n_mfcc	Smarty4Covid		Coswara	
	frame_size, n_segments	F1 Score	frame_size, n_segments	F1 Score
14	2048,50	0,655226209	4096,70	0,697495183
28	1024,100	0,621621622	4096,100	0,698412698
42	2048,120	0,639490446	4096,120	0,717791411
56	2048,120	0,639490446	4096,120	0,717791411
70	2048,120	0,639490446	4096,120	0,717791411

Πίνακας 29: Max F1 Score value per `frame_size` & `n_segments` αποτελέσματος εκπαίδευσης ταξινομητή KNN

Για να γίνει αξιόπιστη σύγκριση δεδομένων, κρίθηκε σκόπιμη η εύρεση του μέσου όρου του συνδυασμού `n_mfcc`, `frame_size` & `n_segments` ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων ώστε να αποφευχθούν τυχόν παραπλανητικές κορυφές αποτελεσμάτων. Τα αποτελέσματα της σύγκρισης τους και για τις δυο βάσεις δεδομένων (Smart4Covid & Coswara), παρουσιάζονται στον επόμενο πίνακα.

<code>n_mfcc</code>	Smart4Covid	Coswara
14	0,634513492	0,536779918
28	0,644276055	0,560050979
42	0,650231417	0,577953298
56	0,650231417	0,577953298
70	0,650231417	0,577953298

Πίνακας 30: Μέσος Όρος F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή

Συμπέρασμα

Από τους παραπάνω πίνακες (Πίνακας Max F1 Score & Μέσος Όρος F1 Score) συνάγονται τα ακόλουθα συμπεράσματα:

Με την βάση δεδομένων Smart4Covid το καλύτερο F1 Score το δίνει ο συνδυασμός `frame_size = 2048` και `n_segments = 120` ανεξαρτήτως τιμής `n_mfcc`.

Με την βάση δεδομένων Coswara το καλύτερο F1 Score το δίνει ο συνδυασμός `frame_size = 4096` και `n_segments = 120` ανεξαρτήτως τιμής `n_mfcc`.

4.3. Αποτελέσματα εκπαίδευσης ταξινομητή Multi-layer Perceptron

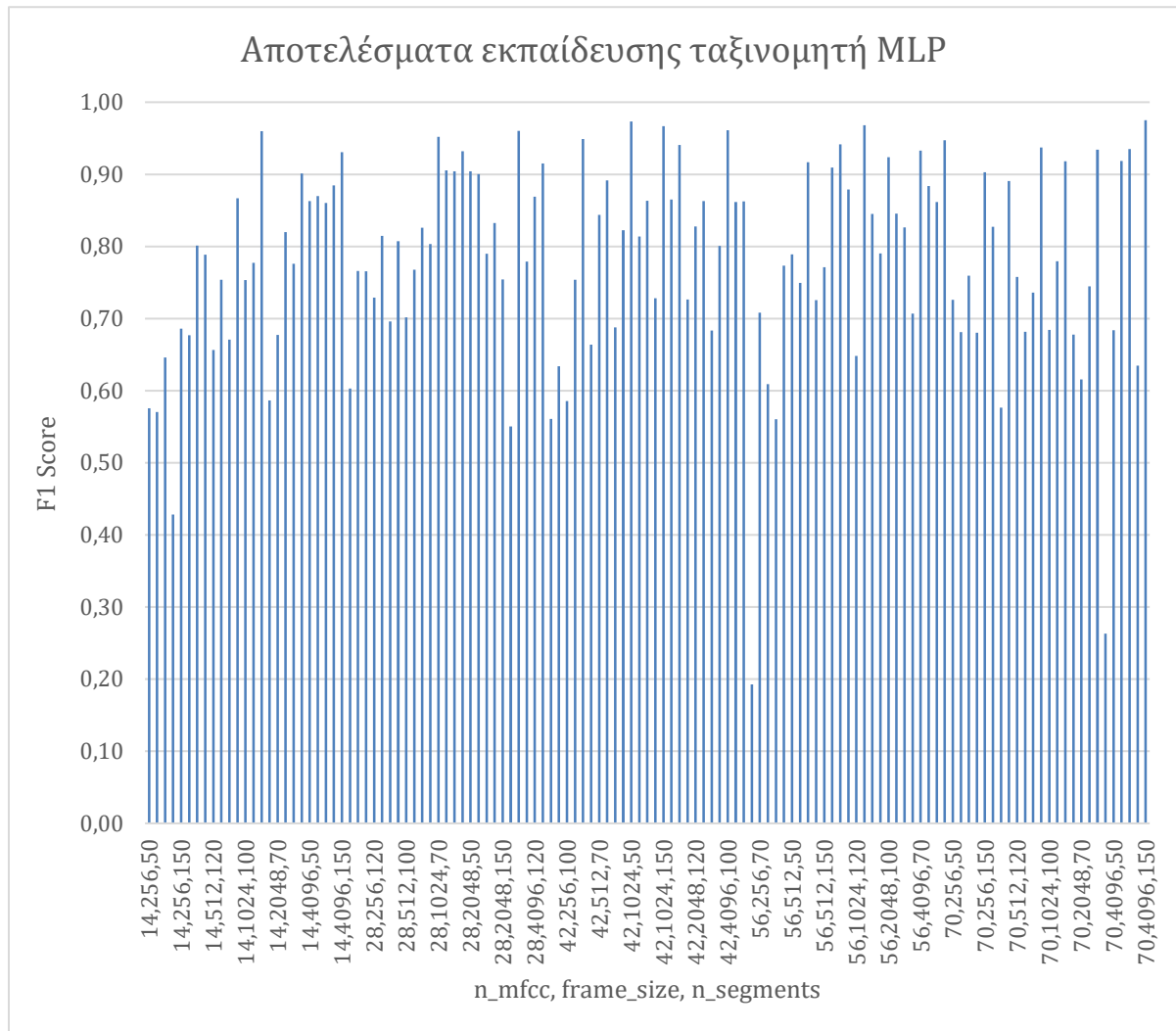
4.3.1. Εκπαίδευση με βάση δεδομένων Smarty4Covid

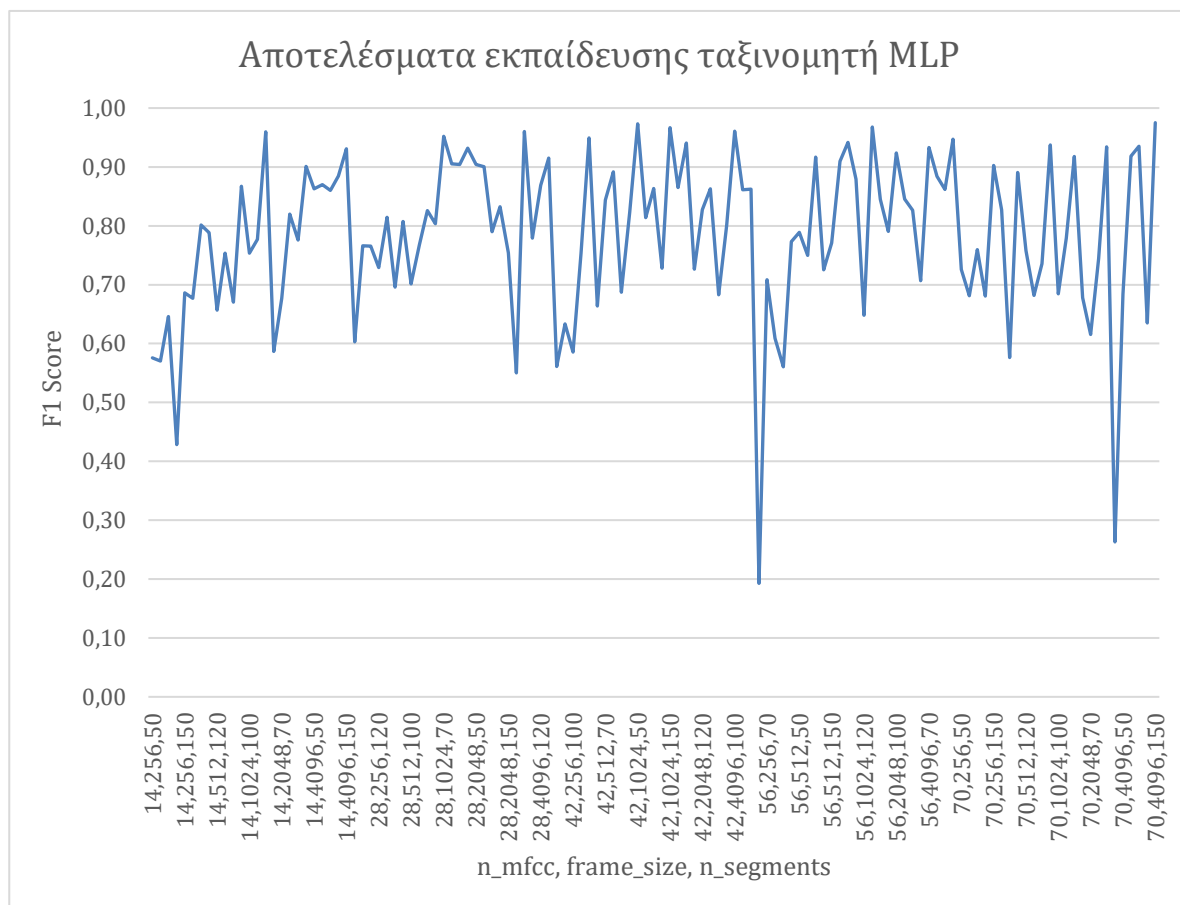
Το πρόγραμμα εκπαίδευσης ταξινομητών, χρησιμοποίησε την βάση δεδομένων της εφαρμογής **Smarty4Covid**, εφαρμόζοντας συνδυαστικές ρυθμίσεις υπερπαραμέτρων για τον τύπο ταξινομητή **Multi-layer Perceptron (MLP)**. Τα αποτελέσματα του παρουσιάζονται στους επόμενους πίνακες και διαγράμματα και στη συνέχεια αναλύονται και συνάγονται χρήσιμα στατιστικά στοιχεία.

n_mfcc, frame_size, n_segments	F1 Score		n_mfcc, frame_size, n_segments	F1 Score		n_mfcc, frame_size, n_segments	F1 Score
14,256,50	0,575624082		28,2048,100	0,790014684		56,512,150	0,771021992
14,256,70	0,570570571		28,2048,120	0,832417582		56,1024,50	0,909379968
14,256,100	0,646115906		28,2048,150	0,754266212		56,1024,70	0,941678521
14,256,120	0,428305400		28,4096,50	0,550495050		56,1024,100	0,879089616
14,256,150	0,685908319		28,4096,70	0,960365854		56,1024,120	0,648351648
14,512,50	0,676975945		28,4096,100	0,779014308		56,1024,150	0,967930029
14,512,70	0,801358234		28,4096,120	0,868741543		56,2048,50	0,845029240
14,512,100	0,788359788		28,4096,150	0,915204678		56,2048,70	0,790476190
14,512,120	0,656603774		42,256,50	0,560975610		56,2048,100	0,923794712
14,512,150	0,753741497		42,256,70	0,633663366		56,2048,120	0,845569620
14,1024,50	0,670640835		42,256,100	0,585600000		56,2048,150	0,826466916
14,1024,70	0,866965620		42,256,120	0,753623188		56,4096,50	0,706872370
14,1024,100	0,753363229		42,256,150	0,949101796		56,4096,70	0,932761087
14,1024,120	0,777173913		42,512,50	0,663608563		56,4096,100	0,883870968
14,1024,150	0,959654179		42,512,70	0,843575419		56,4096,120	0,861538462
14,2048,50	0,586520947		42,512,100	0,891461650		56,4096,150	0,947049924
14,2048,70	0,677098151		42,512,120	0,687500000		70,256,50	0,725988701
14,2048,100	0,819718310		42,512,150	0,822580645		70,256,70	0,681044268
14,2048,120	0,775910364		42,1024,50	0,973372781		70,256,100	0,759601707
14,2048,150	0,900976290		42,1024,70	0,813725490		70,256,120	0,680473373
14,4096,50	0,862800566		42,1024,100	0,863509749		70,256,150	0,902777778
14,4096,70	0,869692533		42,1024,120	0,728143713		70,512,50	0,827067669
14,4096,100	0,860335196		42,1024,150	0,966906475		70,512,70	0,576547231
14,4096,120	0,884462151		42,2048,50	0,864864865		70,512,100	0,890600924
14,4096,150	0,930875576		42,2048,70	0,940458015		70,512,120	0,757894737
28,256,50	0,602857143		42,2048,100	0,726231386		70,512,150	0,681664792
28,256,70	0,766101695		42,2048,120	0,827785818		70,1024,50	0,735785953
28,256,100	0,765562914		42,2048,150	0,862800566		70,1024,70	0,937125749
28,256,120	0,729194188		42,4096,50	0,683168317		70,1024,100	0,684310019
28,256,150	0,814492754		42,4096,70	0,800623053		70,1024,120	0,779220779
28,512,50	0,695757576		42,4096,100	0,960926194		70,1024,150	0,917933131
28,512,70	0,807262570		42,4096,120	0,861400894		70,2048,50	0,677918425
28,512,100	0,701438849		42,4096,150	0,862385321		70,2048,70	0,615595075
28,512,120	0,767857143		56,256,50	0,192771084		70,2048,100	0,744471744
28,512,150	0,825710754		56,256,70	0,708215297		70,2048,120	0,934114202

28,1024,50	0,803361345		56,256,100	0,608849558		70,2048,150	0,263291139
28,1024,70	0,952095808		56,256,120	0,560321716		70,4096,50	0,683840749
28,1024,100	0,905292479		56,256,150	0,773296245		70,4096,70	0,918309859
28,1024,120	0,904320988		56,512,50	0,789017341		70,4096,100	0,935185185
28,1024,150	0,931952663		56,512,70	0,749590835		70,4096,120	0,634831461
28,2048,50	0,903988183		56,512,100	0,916666667		70,4096,150	0,975182482
28,2048,70	0,900153610		56,512,120	0,725490196			

Πίνακας 31: Αποτελέσματα εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc , $frame_size$, $n_segments$





Διάγραμμα 26: Διάγραμμα αποτελέσματος εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc , $frame_size$, $n_segments$

Εφαρμόζοντας τυπική στατιστική ανάλυση στα δεδομένα του ανωτέρω πίνακα συνάγονται τα επόμενα χρήσιμα στατιστικά στοιχεία.

Στοιχεία	Τιμές
Μέγιστη τιμή	0,975182482
Ελάχιστη τιμή	0,192771084
Διακύμανση	0,018849843
Μέσος Όρος	0,782092337
Διάμεσος	0,790476190
Επικρατούσα Τιμή	0,862800566

Πίνακας 32: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid

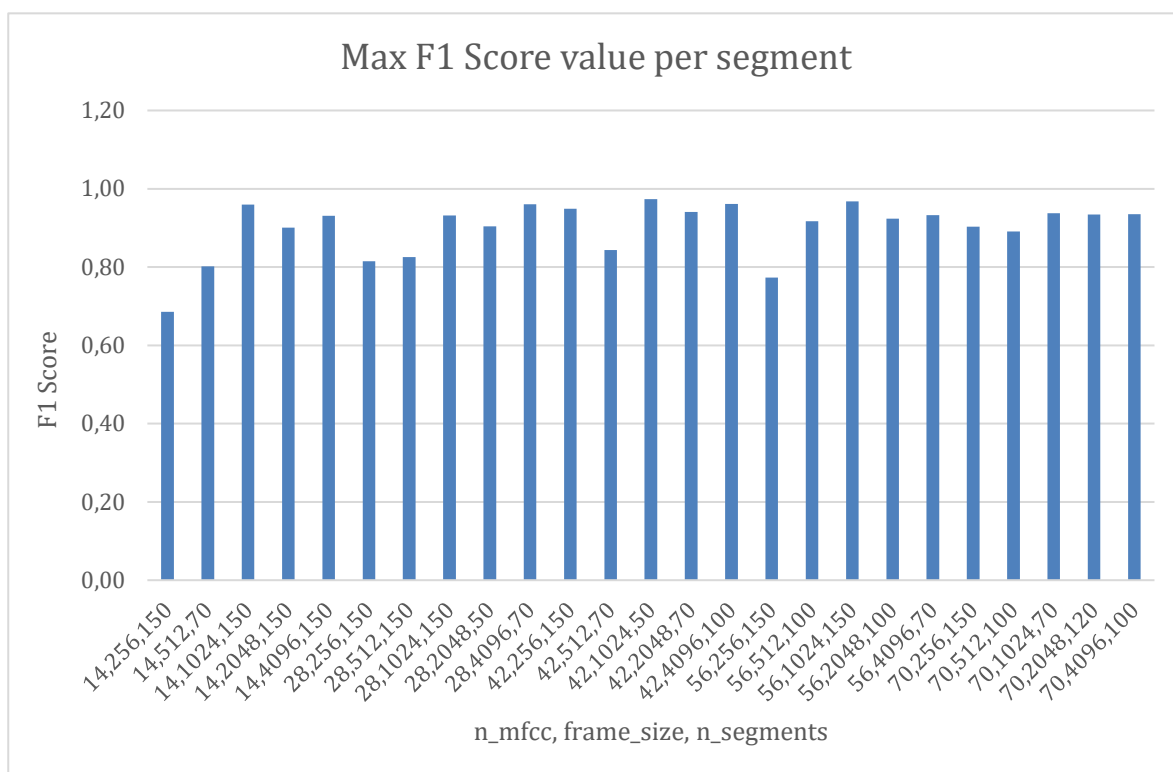
Παρατηρείται ότι η μέγιστη τιμή είναι 0,9752 δηλαδή 97,52% επιτυχία στην πρόβλεψη, η ελάχιστη τιμή είναι $\approx 0,1927$ δηλαδή 19,27 % επιτυχία στην πρόβλεψη (αν και υπάρχουν μόνο τρεις τιμές κάτω από 0,55), με διακύμανση τιμών $\approx 0,0189$. Ο μέσος όρος των αποτελεσμάτων είναι $\approx 0,7821$ δηλαδή 78,21 % επιτυχία στην πρόβλεψη, με διάμεση τιμή $\approx 0,7905$ (79,05 %) και επικρατούσα τιμή $\approx 0,8628$ (86,28 %). Συνάγεται από τις ανωτέρω τιμές

της στατιστικής ανάλυσης, με το σύνολο δεδομένων εκπαίδευσης **smarty4covid**, ότι το ποσοστό επιτυχίας της πρόβλεψης για το αν ένα άτομο πάσχει από COVID-19 είναι αρκετά μεγάλο.

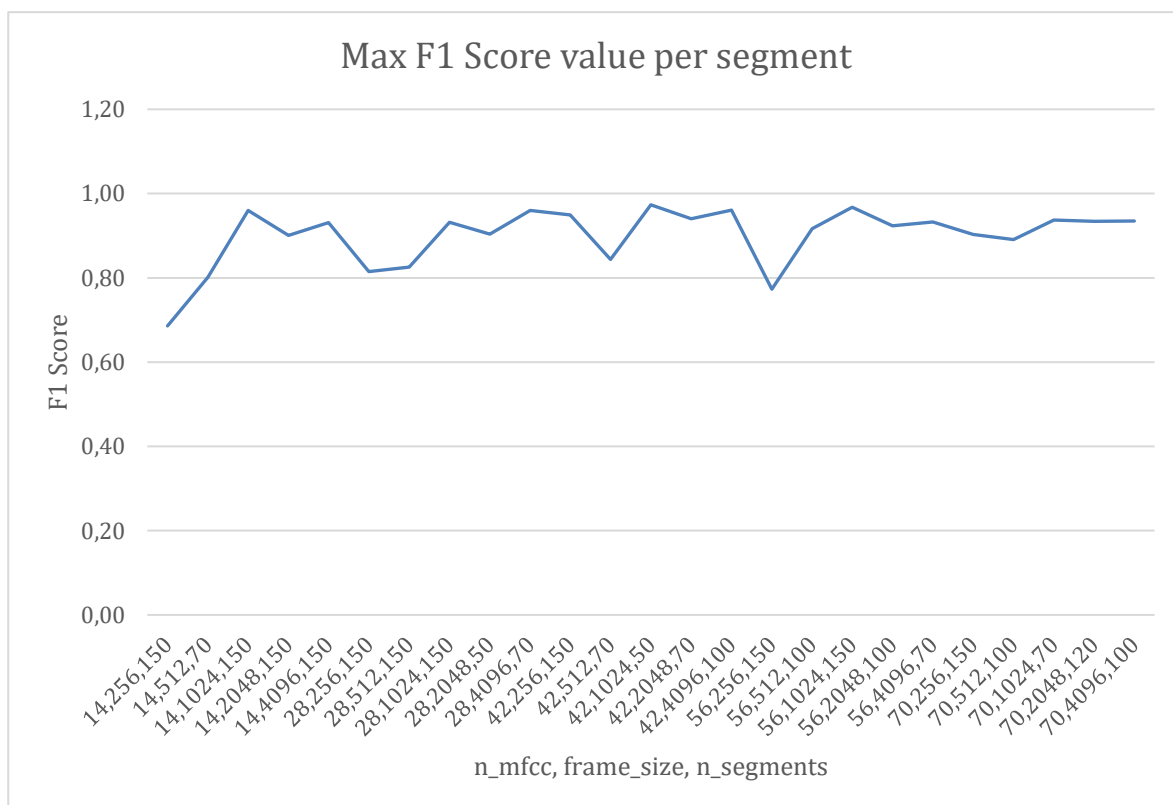
Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού **n_mfcc**, **frame_size** & **n_segments** ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές **n_segment** ανά συνδυασμό **n_mfcc** & **frame_size** καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per segment).

n_mfcc, frame_size, n_segments	Max F1 Score value per segment
14,256,150	0,685908319
14,512,70	0,801358234
14,1024,150	0,959654179
14,2048,150	0,900976290
14,4096,150	0,930875576
28,256,150	0,814492754
28,512,150	0,825710754
28,1024,150	0,931952663
28,2048,50	0,903988183
28,4096,70	0,960365854
42,256,150	0,949101796
42,512,70	0,843575419
42,1024,50	0,973372781
42,2048,70	0,940458015
42,4096,100	0,960926194
56,256,150	0,773296245
56,512,100	0,916666667
56,1024,150	0,967930029
56,2048,100	0,923794712
56,4096,70	0,932761087
70,256,150	0,902777778
70,512,100	0,890600924
70,1024,70	0,937125749
70,2048,120	0,934114202
70,4096,100	0,935185185

Πίνακας 33: Max F1 Score value per segment αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid



Διάγραμμα 27: Ραβδόγραμμα μέγιστων τιμών F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size per n_segment



Διάγραμμα 28: Διάγραμμα μέγιστων τιμών F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid με συνδυασμό n_mfcc, frame_size per n_segment

Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού n_mfcc , $frame_size$ & $n_segments$ ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές $frame_size$ & $n_segments$ ανά n_mfcc καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per $frame_size$ & $n_segments$).

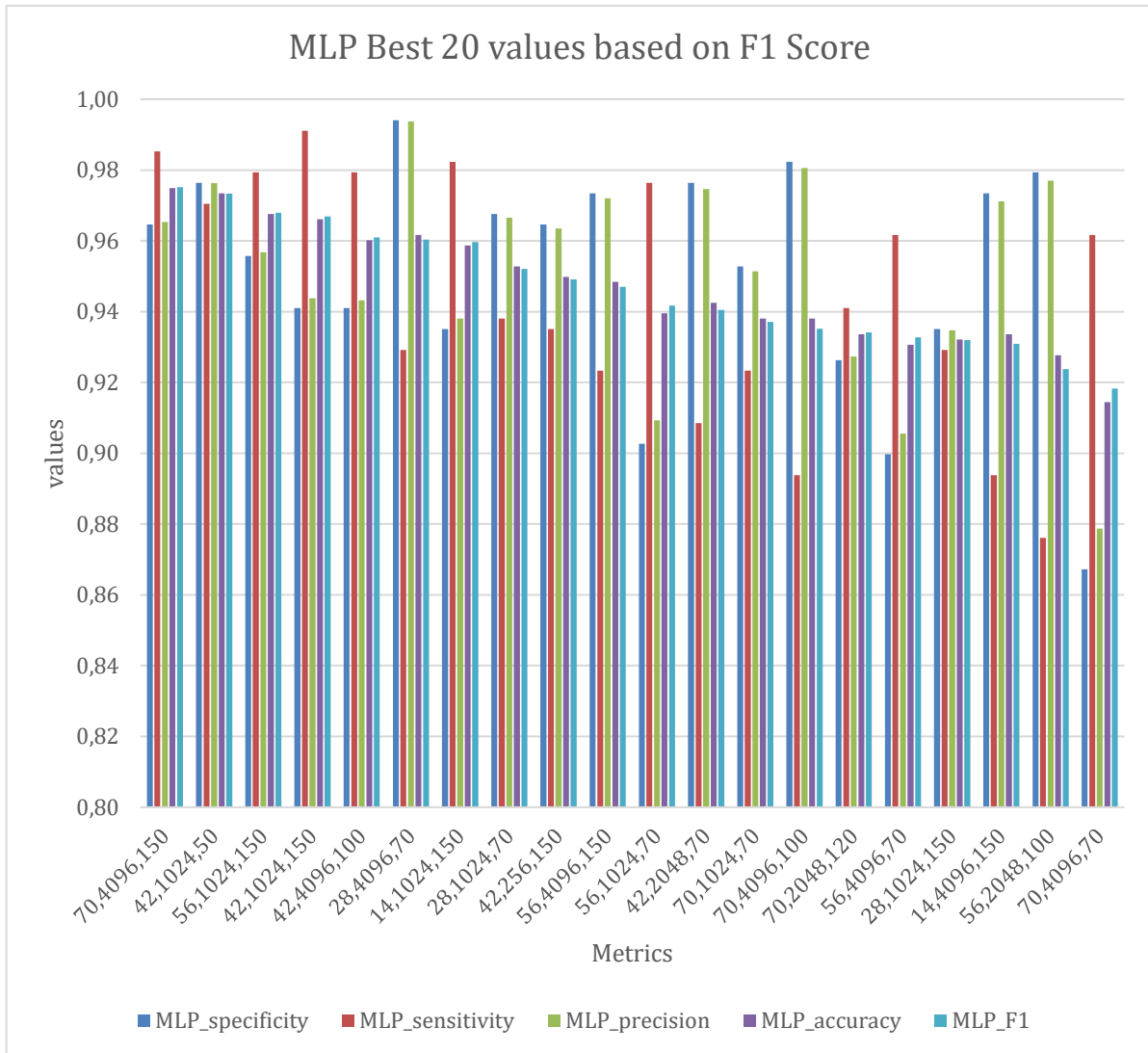
n_mfcc , $frame_size$, $n_segments$	Max F1 Score value per $frame_size$ & $n_segments$
14,1024,150	0,959654179
28,4096,70	0,960365854
42,1024,50	0,973372781
56,1024,150	0,967930029
70,1024,70	0,937125749

Πίνακας 34: Max F1 Score value per $frame_size$ & $n_segments$ αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid

Επιπρόσθετα στην ανάλυση παρατίθενται στον επόμενο πίνακα οι 20 μεγαλύτερες τιμές αποτελεσμάτων με βάση το F1 Score (Best 20 values based on F1 Score) με συνδυασμό n_mfcc , $frame_size$ & $n_segment$. Στον πίνακα επιπλέον παρατίθενται και οι υπόλοιπες μετρικές για κάθε συνδυασμό.

n_mfcc , $frame_size$, $n_segments$	Specificity	Sensitivity	Precision	Accuracy	F1 Score
70,4096,150	0,964601770	0,985250737	0,965317919	0,974926254	0,975182482
42,1024,50	0,976401180	0,970501475	0,976261128	0,973451327	0,973372781
56,1024,150	0,955752212	0,979351032	0,956772334	0,967551622	0,967930029
42,1024,150	0,941002950	0,991150442	0,943820225	0,966076696	0,966906475
42,4096,100	0,941002950	0,979351032	0,943181818	0,960176991	0,960926194
28,4096,70	0,994100295	0,929203540	0,993690852	0,961651917	0,960365854
14,1024,150	0,935103245	0,982300885	0,938028169	0,958702065	0,959654179
28,1024,70	0,967551622	0,938053097	0,966565350	0,952802360	0,952095808
42,256,150	0,964601770	0,935103245	0,963525836	0,949852507	0,949101796
56,4096,150	0,973451327	0,923303835	0,972049689	0,948377581	0,947049924
56,1024,70	0,902654867	0,976401180	0,909340659	0,939528024	0,941678521
42,2048,70	0,976401180	0,908554572	0,974683544	0,942477876	0,940458015
70,1024,70	0,952802360	0,923303835	0,951367781	0,938053097	0,937125749
70,4096,100	0,982300885	0,893805310	0,980582524	0,938053097	0,935185185
70,2048,120	0,926253687	0,941002950	0,927325581	0,933628319	0,934114202
56,4096,70	0,899705015	0,961651917	0,905555556	0,930678466	0,932761087
28,1024,150	0,935103245	0,929203540	0,934718101	0,932153392	0,931952663
14,4096,150	0,973451327	0,893805310	0,971153846	0,933628319	0,930875576
56,2048,100	0,979351032	0,876106195	0,976973684	0,927728614	0,923794712
70,4096,70	0,867256637	0,961651917	0,878706199	0,914454277	0,918309859

Πίνακας 35: Best 20 values based on F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid



Διάγραμμα 29: Ραβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελέσματος εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων smarty4covid (Best 20 values based on F1 Score)

Συμπέρασμα

Παρατηρείται ότι οι μεγαλύτερες τιμές (F1 Score) εμφανίζονται όταν υπάρχει συνδυασμός μεγάλων τιμών n_mfcc (42,56 & 70), μεσαίων και πολύ μεγάλων τιμών $frame_size$ (1024 & 4096) & μεγάλων τιμών $n_segments$ (150).

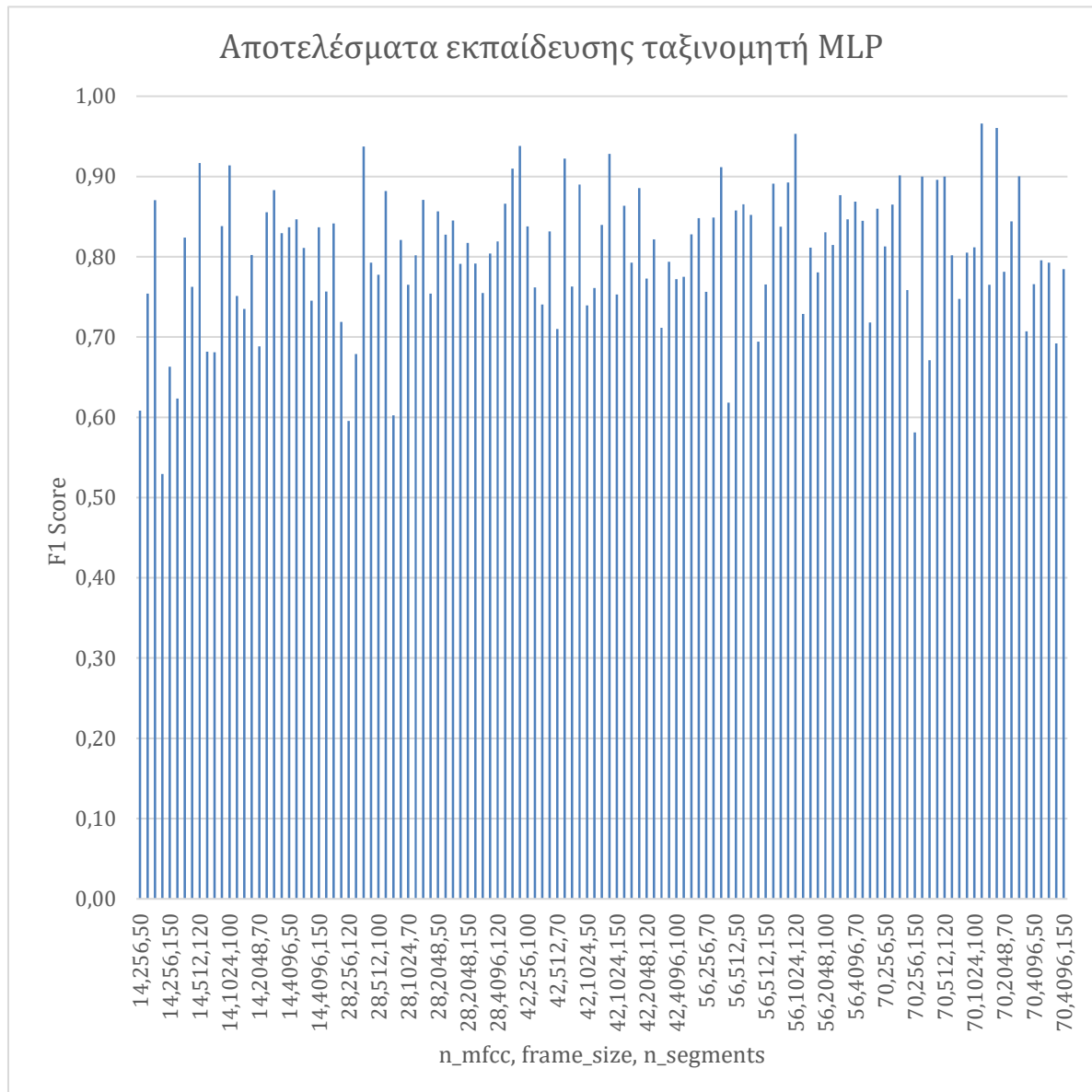
4.3.2. Εκπαίδευση με βάση δεδομένων Coswara

Το πρόγραμμα εκπαίδευσης ταξινομητών, χρησιμοποίησε την βάση δεδομένων της εφαρμογής **Coswara**, εφαρμόζοντας συνδυαστικές ρυθμίσεις υπερπαραμέτρων για τον τύπο ταξινομητή **Multi-layer Perceptron (MLP)**. Τα αποτελέσματα του παρουσιάζονται στους επόμενους πίνακες και διαγράμματα και στη συνέχεια αναλύονται και συνάγονται χρήσιμα στατιστικά στοιχεία.

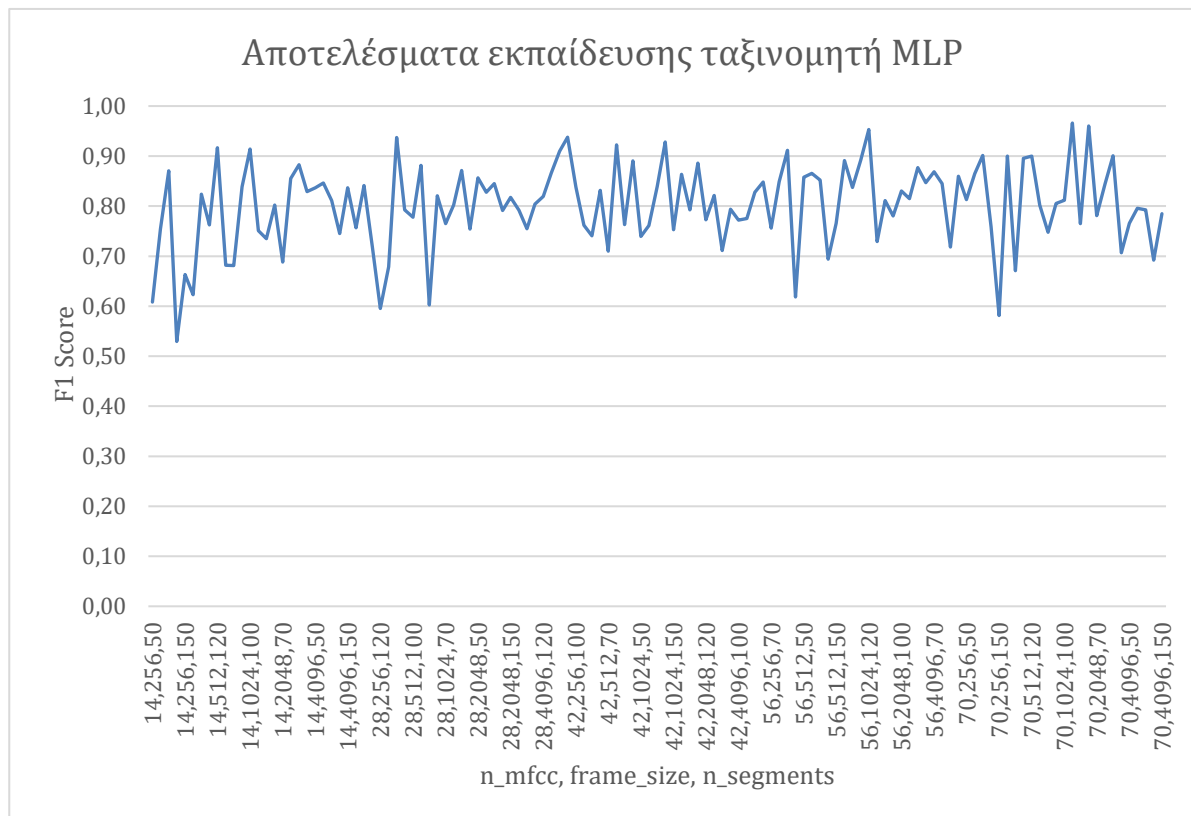
n_mfcc, frame_size, n_segments	F1 Score		n_mfcc, frame_size, n_segments	F1 Score		n_mfcc, frame_size, n_segments	F1 Score
14,256,50	0,608333333		28,2048,100	0,845070423		56,512,150	0,765384615
14,256,70	0,753926702		28,2048,120	0,791075051		56,1024,50	0,891170431
14,256,100	0,870406190		28,2048,150	0,817391304		56,1024,70	0,837500000
14,256,120	0,529411765		28,4096,50	0,791586998		56,1024,100	0,892430279
14,256,150	0,663194444		28,4096,70	0,754863813		56,1024,120	0,953156823
14,512,50	0,623239437		28,4096,100	0,804081633		56,1024,150	0,728850325
14,512,70	0,824000000		28,4096,120	0,819085487		56,2048,50	0,811244980
14,512,100	0,762677485		28,4096,150	0,866267465		56,2048,70	0,780487805
14,512,120	0,916835700		42,256,50	0,909836066		56,2048,100	0,830434783
14,512,150	0,681732580		42,256,70	0,937984496		56,2048,120	0,814814815
14,1024,50	0,681081081		42,256,100	0,837719298		56,2048,150	0,876923077
14,1024,70	0,838297872		42,256,120	0,761718750		56,4096,50	0,846601942
14,1024,100	0,913827655		42,256,150	0,740601504		56,4096,70	0,868852459
14,1024,120	0,751072961		42,512,50	0,831408776		56,4096,100	0,844720497
14,1024,150	0,734848485		42,512,70	0,709956710		56,4096,120	0,718279570
14,2048,50	0,802303263		42,512,100	0,922178988		56,4096,150	0,859960552
14,2048,70	0,688453159		42,512,120	0,762931034		70,256,50	0,812925170
14,2048,100	0,855291577		42,512,150	0,890151515		70,256,70	0,865079365
14,2048,120	0,882917466		42,1024,50	0,739420935		70,256,100	0,901287554
14,2048,150	0,829268293		42,1024,70	0,761099366		70,256,120	0,758620690
14,4096,50	0,836820084		42,1024,100	0,839552239		70,256,150	0,581235698
14,4096,70	0,846491228		42,1024,120	0,928301887		70,512,50	0,900000000
14,4096,100	0,811040340		42,1024,150	0,752941176		70,512,70	0,671081678
14,4096,120	0,745263158		42,2048,50	0,863543788		70,512,100	0,895705521
14,4096,150	0,836734694		42,2048,70	0,792531120		70,512,120	0,900000000
28,256,50	0,756660746		42,2048,100	0,885654886		70,512,150	0,801822323
28,256,70	0,841328413		42,2048,120	0,772908367		70,1024,50	0,747514911
28,256,100	0,718676123		42,2048,150	0,821497121		70,1024,70	0,805054152
28,256,120	0,595348837		42,4096,50	0,711316397		70,1024,100	0,811594203
28,256,150	0,678700361		42,4096,70	0,793926247		70,1024,120	0,966067864
28,512,50	0,937254902		42,4096,100	0,772009029		70,1024,150	0,764957265
28,512,70	0,792540793		42,4096,120	0,775000000		70,2048,50	0,960317460
28,512,100	0,777777778		42,4096,150	0,828000000		70,2048,70	0,781316348
28,512,120	0,881763527		56,256,50	0,848056537		70,2048,100	0,844181460
28,512,150	0,602531646		56,256,70	0,756264237		70,2048,120	0,900398406
28,1024,50	0,820960699		56,256,100	0,848948375		70,2048,150	0,707021792
28,1024,70	0,764957265		56,256,120	0,911591356		70,4096,50	0,765661253

28,1024,100	0,801843318		56,256,150	0,618395303		70,4096,70	0,795652174
28,1024,120	0,871028037		56,512,50	0,857664234		70,4096,100	0,792531120
28,1024,150	0,754166667		56,512,70	0,865217391		70,4096,120	0,691964286
28,2048,50	0,856589147		56,512,100	0,852248394		70,4096,150	0,784648188
28,2048,70	0,827433628		56,512,120	0,694174757			

Πίνακας 36: Αποτελέσματα εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc , $frame_size$, $n_segments$



Διάγραμμα 30: Ραβδόγραμμα αποτελέσματος εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc , $frame_size$, $n_segments$



Διάγραμμα 31: Διάγραμμα αποτελέσματος εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc , $frame_size$, $n_segments$

Εφαρμόζοντας τυπική στατιστική ανάλυση στα δεδομένα του ανωτέρω πίνακα συνάγονται τα επόμενα χρήσιμα στατιστικά στοιχεία.

Στοιχεία	Τιμές
Μέγιστη τιμή	0,966067864
Ελάχιστη τιμή	0,529411765
Διακύμανση	0,007143583
Μέσος Όρος	0,801429241
Διάμεσος	0,811244980
Επικρατούσα Τιμή	0,764957265

Πίνακας 37: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara

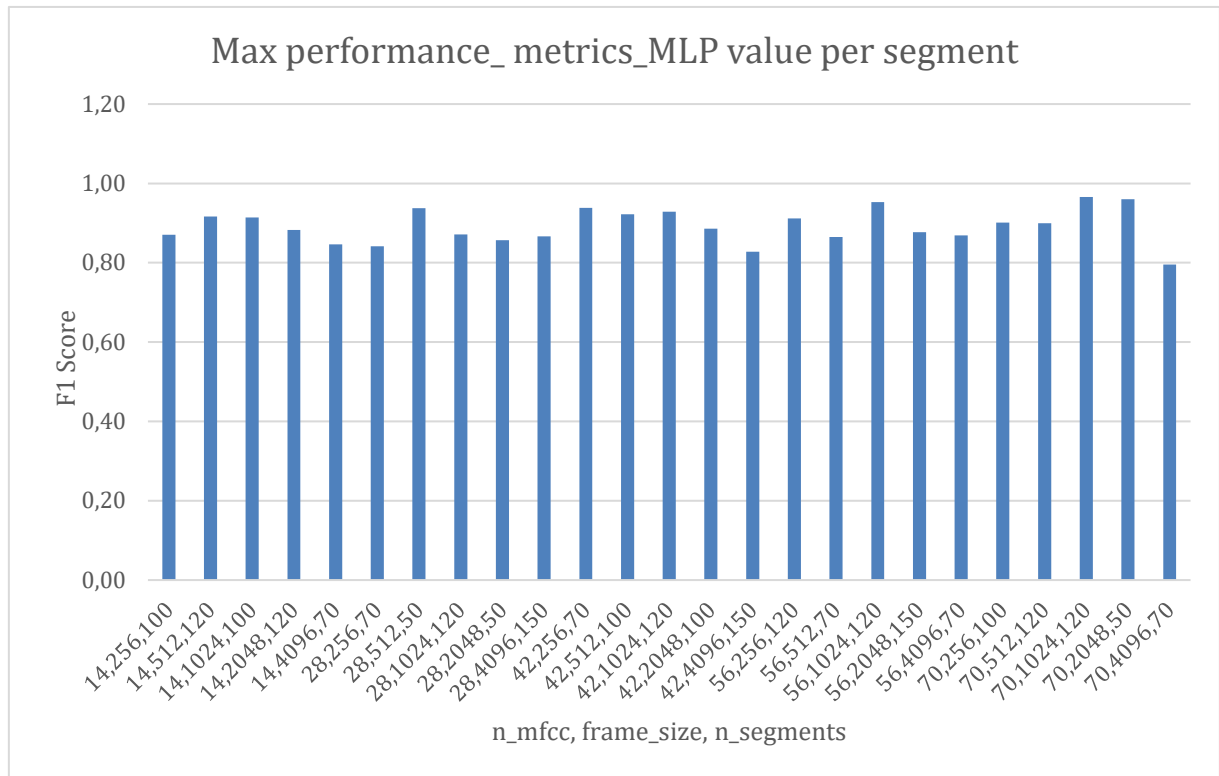
Παρατηρείται ότι η μέγιστη τιμή είναι 0,9661 δηλαδή 96,61% επιτυχία στην πρόβλεψη, η ελάχιστη τιμή είναι $\approx 0,5294$ δηλαδή 52,94 % επιτυχία στην πρόβλεψη (αν και υπάρχουν μόνο τρεις τιμές κάτω από 0,60), με διακύμανση τιμών $\approx 0,0071$. Ο μέσος όρος των αποτελεσμάτων είναι $\approx 0,8014$ δηλαδή 80,14 % επιτυχία στην πρόβλεψη, με διάμεση τιμή $\approx 0,8112$ (81,12 %) και επικρατούσα τιμή $\approx 0,7650$ (76,50 %). Συνάγεται από τις ανωτέρω τιμές

της στατιστικής ανάλυσης, με το σύνολο δεδομένων εκπαίδευσης *coswara*, ότι το ποσοστό επιτυχίας της πρόβλεψης για το αν ένα άτομο πάσχει από COVID-19 είναι αρκετά μεγάλο.

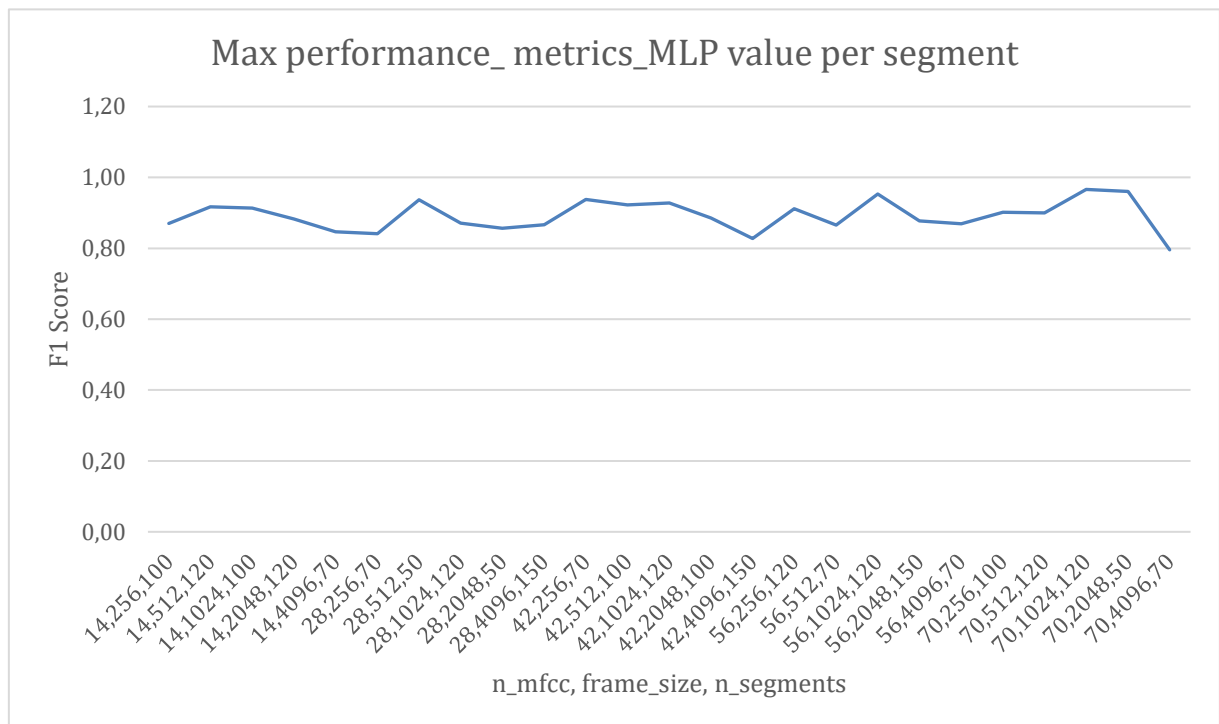
Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού *n_mfcc*, *frame_size* & *n_segments* ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές *n_segment* ανά συνδυασμό *n_mfcc* & *frame_size* καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per segment).

<i>n_mfcc</i> , <i>frame_size</i> , <i>n_segments</i>	Max F1 Score value per segment
14,256,100	0,870406190
14,512,120	0,916835700
14,1024,100	0,913827655
14,2048,120	0,882917466
14,4096,70	0,846491228
28,256,70	0,841328413
28,512,50	0,937254902
28,1024,120	0,871028037
28,2048,50	0,856589147
28,4096,150	0,866267465
42,256,70	0,937984496
42,512,100	0,922178988
42,1024,120	0,928301887
42,2048,100	0,885654886
42,4096,150	0,828000000
56,256,120	0,911591356
56,512,70	0,865217391
56,1024,120	0,953156823
56,2048,150	0,876923077
56,4096,70	0,868852459
70,256,100	0,901287554
70,512,120	0,900000000
70,1024,120	0,966067864
70,2048,50	0,960317460
70,4096,70	0,795652174

Πίνακας 38: Max F1 Score value per segment αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara



Διάγραμμα 32: Ραβδόγραμμα μέγιστων τιμών F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size per n_segment



Διάγραμμα 33: Διάγραμμα μέγιστων τιμών F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara με συνδυασμό n_mfcc, frame_size per n_segment

Προχωρώντας σε περαιτέρω ανάλυση, με γνώμονα την εύρεση του βέλτιστου συνδυασμού `n_mfcc`, `frame_size` & `n_segments` ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, παρατίθενται στον επόμενο πίνακα οι τιμές `frame_size` & `n_segments` ανά `n_mfcc` καθώς και οι μέγιστες τιμές αποτελεσμάτων (Max F1 Score value per `frame_size` & `n_segments`).

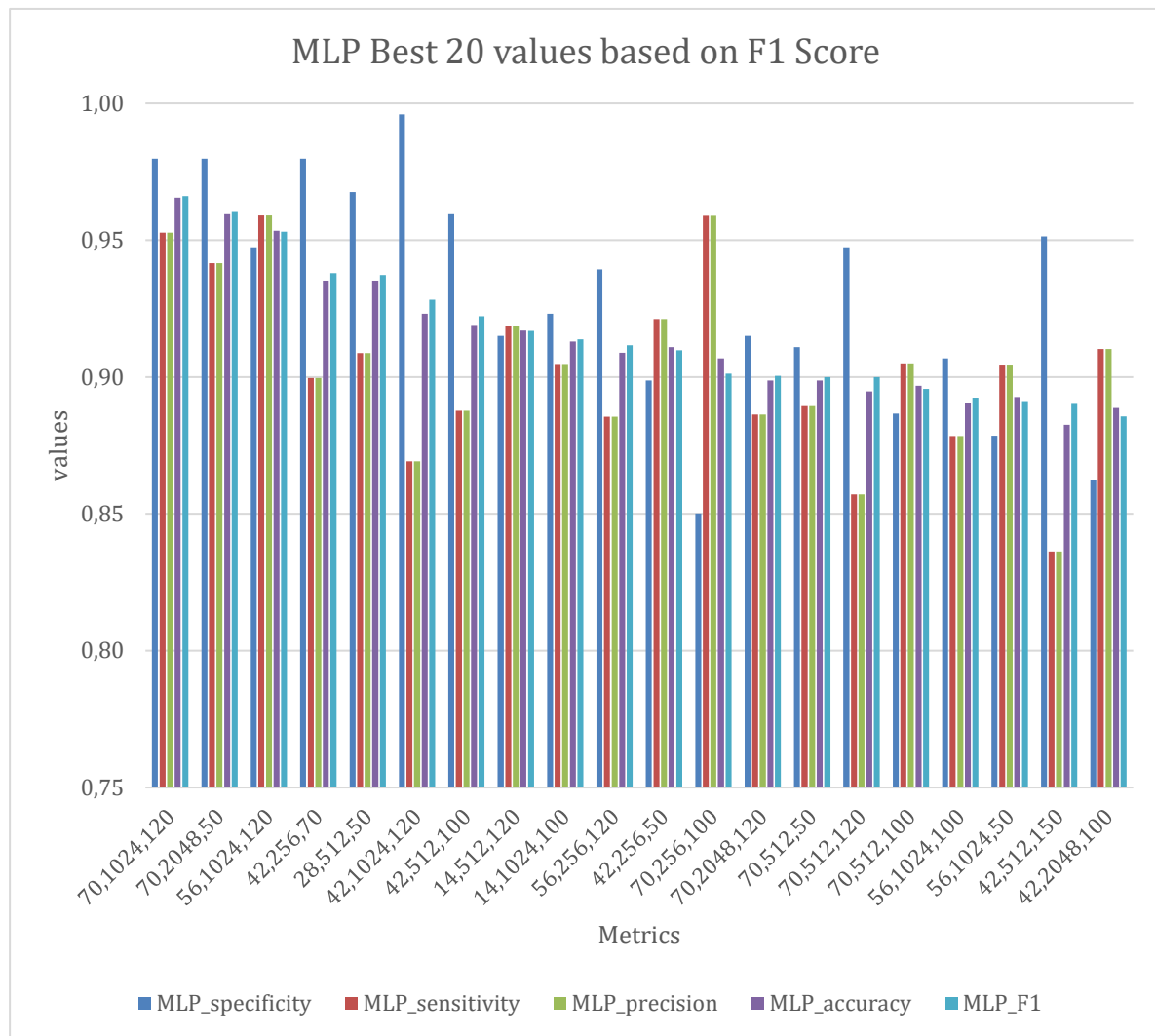
<code>n_mfcc</code> , <code>frame_size</code> , <code>n_segments</code>	Max F1 Score value per <code>frame_size</code> & <code>n_segments</code>
14,512,120	0,916835700
28,512,50	0,937254902
42,256,70	0,937984496
56,1024,120	0,953156823
70,1024,120	0,966067864

Πίνακας 39: Max F1 Score value per `frame_size` & `n_segments` αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων `coswara`

Επιπρόσθετα στην ανάλυση παρατίθενται στον επόμενο πίνακα οι 20 μεγαλύτερες τιμές αποτελεσμάτων με βάση το F1 Score (Best 20 values based on F1 Score) με συνδυασμό `n_mfcc`, `frame_size` & `n_segment`. Στον πίνακα επιπλέον παρατίθενται και οι υπόλοιπες μετρικές για κάθε συνδυασμό.

<code>n_mfcc</code> , <code>frame_size</code> , <code>n_segments</code>	Specificity	Sensitivity	Precision	Accuracy	F1 Score
70,1024,120	0,979757085	0,952755906	0,952755906	0,965587045	0,966067864
70,2048,50	0,979757085	0,941634241	0,941634241	0,959514170	0,960317460
56,1024,120	0,947368421	0,959016393	0,959016393	0,953441296	0,953156823
42,256,70	0,979757085	0,899628253	0,899628253	0,935222672	0,937984496
28,512,50	0,967611336	0,908745247	0,908745247	0,935222672	0,937254902
42,1024,120	0,995951417	0,869257951	0,869257951	0,923076923	0,928301887
42,512,100	0,959514170	0,887640449	0,887640449	0,919028340	0,922178988
14,512,120	0,914979757	0,918699187	0,918699187	0,917004049	0,916835700
14,1024,100	0,923076923	0,904761905	0,904761905	0,912955466	0,913827655
56,256,120	0,939271255	0,885496183	0,885496183	0,908906883	0,911591356
42,256,50	0,898785425	0,921161826	0,921161826	0,910931174	0,909836066
70,256,100	0,850202429	0,958904110	0,958904110	0,906882591	0,901287554
70,2048,120	0,914979757	0,886274510	0,886274510	0,898785425	0,900398406
70,512,50	0,910931174	0,889328063	0,889328063	0,898785425	0,900000000
70,512,120	0,947368421	0,857142857	0,857142857	0,894736842	0,900000000
70,512,100	0,886639676	0,904958678	0,904958678	0,896761134	0,895705521
56,1024,100	0,906882591	0,878431373	0,878431373	0,890688259	0,892430279
56,1024,50	0,878542510	0,904166667	0,904166667	0,892712551	0,891170431
42,512,150	0,951417004	0,836298932	0,836298932	0,882591093	0,890151515
42,2048,100	0,862348178	0,910256410	0,910256410	0,888663968	0,885654886

Πίνακας 40: Best 20 values based on F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων `coswara`



Διάγραμμα 34: Ραβδόγραμμα 20 καλύτερων τιμών μετρικών για κάθε συνδυασμό βάσει F1 Score αποτελεσμάτων εκπαίδευσης ταξινομητή MLP στο σύνολο δεδομένων coswara (Best 20 values based on F1 Score)

Συμπέρασμα

Παρατηρείται μια σχετική τυχαιότητα αποτελεσμάτων χωρίς εμφανή σχέση των υπερπαραμέτρων n_{mfcc} , $frame_size$ & $n_segments$, με μόνο τις υψηλότερες τιμές F1 score να έχουν υψηλές τιμές n_{mfcc} (70 & 56) και μέτριες τιμές $frame_size$ (1024 & 2048).

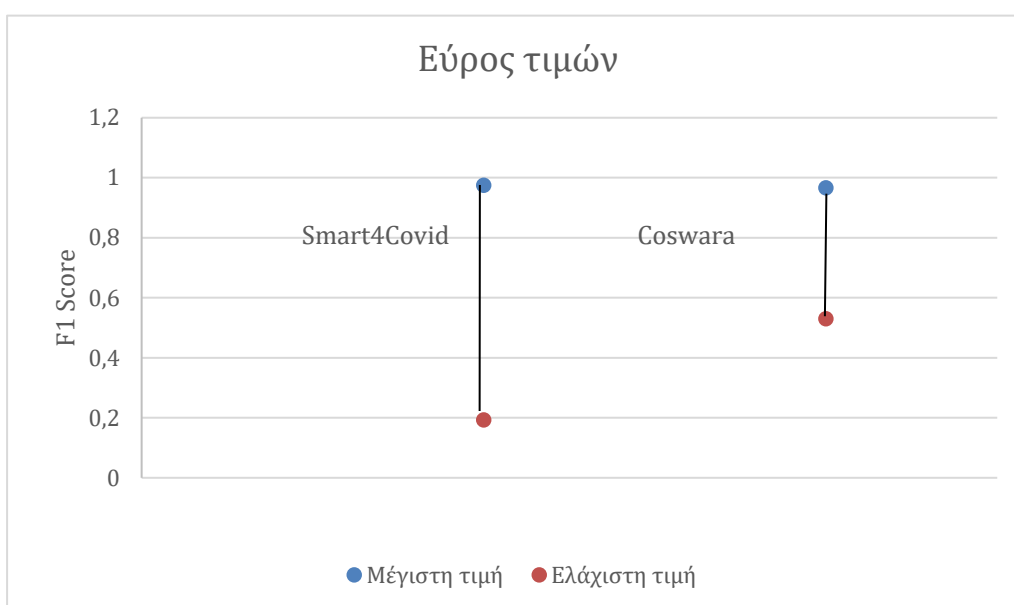
4.3.3. Σύγκριση αποτελεσμάτων από τις δυο βάσεις δεδομένων

Στον επόμενο πίνακα γίνεται σύγκριση των αποτελεσμάτων με βάση δεδομένων Smarty4Covid & Coswara. Παρατηρείται ότι τα αποτελέσματα συμπίπτουν σε όλους τους τομείς και παρουσιάζουν μια συνέπεια ανεξαρτήτως της βάσης των δεδομένων.

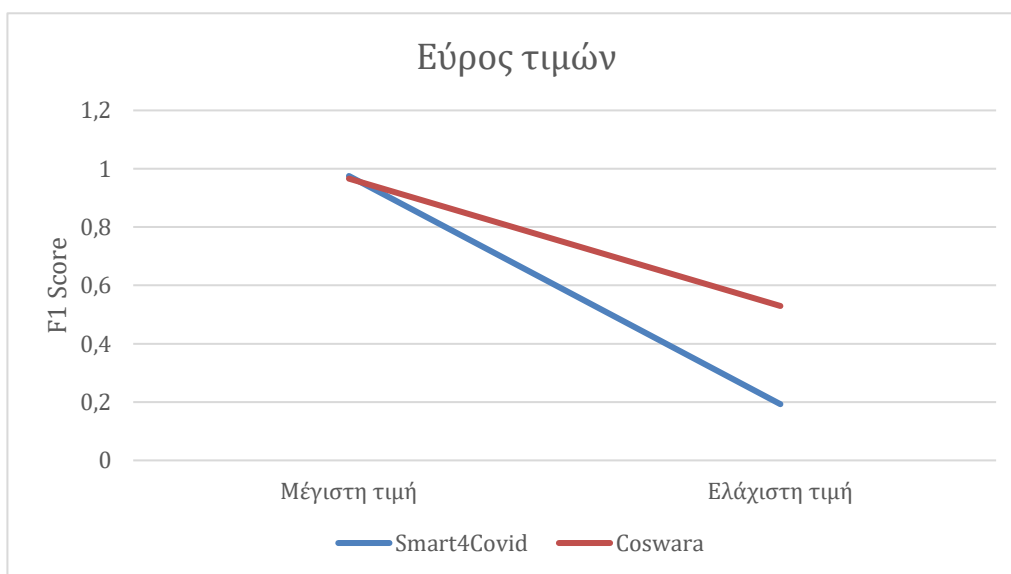
Στοιχεία	Αποτελέσματα	
	Smart4Covid	Coswara
Μέγιστη τιμή	0,975182482	0,966067864
Ελάχιστη τιμή	0,192771084	0,529411765
Διακύμανση	0,018849843	0,007143583
Μέσος Όρος	0,782092337	0,801429241
Διάμεσος	0,790476190	0,811244980
Επικρατούσα Τιμή	0,862800566	0,764957265

Πίνακας 41: Στατιστικά στοιχεία πίνακα αποτελεσμάτων εκπαίδευσης ταξινομητή MLP

Σχηματικά τα ανωτέρω συναγόμενα συμπεράσματα παρουσιάζονται καλύτερα με τα ακόλουθα συγκριτικά διαγράμματα.



Διάγραμμα 35: Οριζόντιο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή MLP



Διάγραμμα 36: Κάθετο συγκριτικό διάγραμμα εύρους τιμών αποτελεσμάτων εκπαίδευσης ταξινομητή MLP

Στον επόμενο πίνακα, με γνώμονα την εύρεση του βέλτιστου συνδυασμού `n_mfcc`, `frame_size` & `n_segments` ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, γίνεται σύγκριση των τιμών `frame_size` & `n_segments` ανά `n_mfcc` καθώς και των μέγιστων τιμών αποτελεσμάτων (Max F1 Score value per `frame_size` & `n_segments`) και για τις δυο βάσεις δεδομένων (Smarty4Covid & Coswara).

n_mfcc	Smarty4Covid		Coswara	
	frame_size, n_segments	F1 Score	frame_size, n_segments	F1 Score
14	1024,150	0,959654179	512,120	0,916835700
28	4096,70	0,960365854	512,50	0,937254902
42	1024,50	0,973372781	256,70	0,937984496
56	1024,150	0,967930029	1024,120	0,953156823
70	1024,70	0,937125749	1024,120	0,966067864

Πίνακας 42: Max F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή MLP

Για να γίνει αξιόπιστη σύγκριση δεδομένων, κρίθηκε σκόπιμη η εύρεση του μέσου όρου του συνδυασμού `n_mfcc`, `frame_size` & `n_segments` ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων ώστε να αποφευχθούν τυχόν παραπλανητικές κορυφές αποτελεσμάτων. Τα αποτελέσματα της σύγκρισης τους και για τις δυο βάσεις δεδομένων (Smarty4Covid & Coswara), παρουσιάζονται στον επόμενο πίνακα.

n_mfcc	Smarty4Covid	Coswara
14	0,751190055	0,771498758
28	0,809116823	0,790759362
42	0,805119715	0,813687588
56	0,788204008	0,822934941
70	0,756831085	0,808265555

Πίνακας 43: Μέσος Όρος F1 Score value per frame_size & n_segments αποτελέσματος εκπαίδευσης ταξινομητή

Συμπέρασμα

Από τους παραπάνω πίνακες (Πίνακας Max F1 Score & Μέσος Όρος F1 Score) συνάγονται τα ακόλουθα συμπεράσματα:

Με την βάση δεδομένων Smarty4Covid το καλύτερο F1 Score το δίνει ο συνδυασμός `n_mfcc` = 42 & 56, `frame_size` = 1024 και `n_segments` = 50 & 150.

Με την βάση δεδομένων Coswara το καλύτερο F1 Score το δίνει ο συνδυασμός `n_mfcc` = 56 & 70, `frame_size` = 1024 και `n_segments` = 120.

4.4. Συγκριτικά στοιχεία

Στον επόμενο πίνακα εμφανίζονται συγκεντρωτικά οι τιμές των υπερπαραμέτρων, των ταξινομητών και της μεγαλύτερης τιμής F1 Score που έχει επιτευχθεί χρησιμοποιώντας τις βάσεις δεδομένων Smart4Covid και Coswara.

Υπερπαραμέτροι	Ταξινομητές					
	LR		KNN		MLP	
	Smart4Covid	Coswara	Smart4Covid	Coswara	Smart4Covid	Coswara
n_segments n_mfcc, frame_size	28,256,70	42,1024,100	14,2048,50	42,4096,120	70,4096,150	70,1024,120
F1 Score	0,994117647	0,968253968	0,655226209	0,717791411	0,975182482	0,966067864
Ισχύς κανονικοποίησης (Hyper_C)	100	10	-	-	-	-
Αριθμός των γειτόνων (Hyper_ n_neighbors)	-	-	10	10	-	-
Μέγεθος του φύλλου Hyper_ leaf_size	-	-	5	5	-	-
Αριθμός των κρυφών νευρώνων (hidden_layer_sizes)	-	-	-	-	30	20
Ισχύς της L2 κανονικοποίησης (alpha)	-	-	-	-	1	0.001
Αρχική τιμή του ρυθμού μάθησης (learning_rate_init)	-	-	-	-	0.1	0.3

Πίνακας 44: Max F1 Score, βάσεων δεδομένων & υπερπαραμέτρων ταξινομητών

Στον επόμενο πίνακα, με γνώμονα την εύρεση του βέλτιστου συνδυασμού n_mfcc , $frame_size$ & $n_segments$, ο οποίος θα δώσει τις μεγαλύτερες τιμές αποτελεσμάτων, γίνεται σύγκριση των τιμών $frame_size$ & $n_segments$ ανά n_mfcc καθώς και των μέγιστων τιμών αποτελεσμάτων (Max F1 Score value per $frame_size$ & $n_segments$) και για τις δυο βάσεις δεδομένων (Smart4Covid & Coswara) και για τους τρεις ταξινομητές (LR, KNN & MLP).

			n_mfcc				
			14	28	42	56	70
Smart4Covid	$frame_size$, $n_segments$	LR	256,100	256,70	512,150	512,150	512,150
		KNN	2048,50	1024,100	2048,120	2048,120	2048,120
		MLP	1024,150	4096,70	1024,50	1024,150	1024,70
	F1 Score	LR	0,904899135	0,994117647	0,982404692	0,982404692	0,982404692
		KNN	0,655226209	0,621621622	0,639490446	0,639490446	0,639490446
		MLP	0,959654179	0,960365854	0,973372781	0,967930029	0,937125749
Coswara	$frame_size$, $n_segments$	LR	256,50	1024,150	1024,100	1024,100	1024,100
		KNN	4096,70	4096,100	4096,120	4096,120	4096,120
		MLP	512,120	512,50	256,70	1024,120	1024,120
	F1 Score	LR	0,861538462	0,933333333	0,968253968	0,968253968	0,968253968
		KNN	0,697495183	0,698412698	0,717791411	0,717791411	0,717791411
		MLP	0,916835700	0,937254902	0,937984496	0,953156823	0,966067864

Πίνακας 45: Max F1 Score value per $frame_size$ & $n_segments$ αποτελεσμάτων εκπαίδευσης ταξινομητών

Από τον ανωτέρω πίνακα συνάγεται ότι ο ταξινομητής LR δίνει τα καλύτερα αποτελέσματα (F1 Score) και στις δυο βάσεις δεδομένων (Smart4Covid & Coswara), έπεται με μικρή διαφορά ο ταξινομητής MLP, ενώ ο ταξινομητής KNN έχει σημαντικά μικρότερα αποτελέσματα, σε επίπεδο περίπου του 66% των αποτελεσμάτων των άλλων δυο. Επίσης η μεγαλύτερη τιμή F1 Score, ως αποτέλεσμα της εφαρμογής των ταξινομητών LR και KNN, στις δυο βάσεις δεδομένων εξαρτάται από τον συνδυασμό $frame_size$ & $n_segments$ και είναι ανεξάρτητη της τιμής n_mfcc . Όταν όμως εφαρμόζεται ο ταξινομητής MLP, η μεγαλύτερη τιμή F1 Score, και στις δυο βάσεις δεδομένων, εξαρτάται και από τις τρεις υπερπαραμέτρους (n_mfcc , $frame_size$ & $n_segments$).

5. Συμπεράσματα - Προτάσεις

Συζήτηση

Η παρούσα διπλωματική εργασία πραγματεύεται την ανάπτυξη μοντέλων μηχανικής μάθησης για την αναγνώριση Covid-19 μέσω ηχητικών καταγραφών. Αναπτύχθηκε ένα πρόγραμμα εκπαίδευσης ταξινομητών με αρθρωτή δομή στη γλώσσα Python για να ευρεθεί η καλύτερη συνδυαστική ρύθμιση υπερπαραμέτρων εξαγωγής χαρακτηριστικών για τους τύπους ταξινομητών Logistic Regression, K-Nearest Neighbors και Multi-layer Perceptron. Για την εκπαίδευση των ταξινομητών χρησιμοποιήθηκε η βάση δεδομένων – εφαρμογή SmartyforCovid, καθώς και η βάση δεδομένων Coswara. Ο στόχος ήταν οι ταξινομητές να προβλέψουν αν ένα άτομο πάσχει από COVID-19 με βάση το υποβληθέν δείγμα ήχου βήχα τους.

Αρχικά πραγματοποιήθηκε έρευνα μελετών στην εκπαίδευση μοντέλων μηχανικής μάθησης για την αναγνώριση της ασθένειας Covid-19, ώστε να γίνει επιλογή των κατάλληλων μεθόδων εξαγωγής χαρακτηριστικών και των επιθυμητών ταξινομητών. Στην συνέχεια αναπτύχθηκε το πρόγραμμα εκπαίδευσης με τη σταδιακή υλοποίηση των μεθόδων εξαγωγής και των ταξινομητών καθώς και όλων των διερευνούμενων υπερπαραμέτρων. Με την ολοκλήρωση της υλοποίησης τους, μελετήθηκαν όλοι οι επιθυμητοί συνδυασμοί των υπερπαραμέτρων εξαγωγής χαρακτηριστικών και εκπαιδεύτηκαν τα μοντέλα ταξινομητών σε κάθε συνδυασμό υπερπαραμέτρων και δεδομένων. Στο τέλος, έγινε σύγκριση των αποτελεσμάτων τους με βάσει τον δείκτη αξιολόγησης F1 Score και πραγματοποιήθηκε στατιστική ανάλυση πάνω σε αυτά.

Τα ευρήματα της στατιστικής ανάλυσης δείχνουν ότι υπάρχουν μοτίβα στα δεδομένα ήχου βήχα τα οποία επιτρέπουν στα μοντέλα ταξινομητών να διακρίνουν ικανοποιητικά τα άτομα που πάσχουν από COVID-19 από τα υγιή άτομα με βάση μόνο το υποβληθέν αρχείο ήχου τους. Επιπρόσθετα, παρατηρήθηκε ότι η επιλογή διαφορετικών συνδυασμών υπερπαραμέτρων εξαγωγής χαρακτηριστικών είχε σημαντικές επιπτώσεις στην επίδοση των μοντέλων, ενώ οι επιδόσεις των ταξινομητών που χρησιμοποιούσαν τις ίδιες υπερπαραμέτρους αλλά διαφορετικές βάσεις δεδομένων ήταν παραπλήσιες.

Συγκεκριμένα, παρατηρήθηκε ότι, με βάσει το F1 Score, ο ταξινομητής LR δίνει τα καλύτερα αποτελέσματα, με τον ταξινομητή MLP να έχει κοντινά αποτελέσματα. Ο ταξινομητής KNN έχει σημαντικά μικρότερα αποτελέσματα, σε επίπεδο περίπου του 66% των αποτελεσμάτων των άλλων δυο.

Μελλοντικές Προεκτάσεις

Η διπλωματική αυτή αποτελεί μια πρώτη υλοποίηση του προγράμματος εκπαίδευσης ταξινομητών με πολλές μελλοντικές προεκτάσεις. Αρχικά, θα μπορούσε να επεκταθεί το εύρος της έρευνας με την εξέταση περισσότερων συνόλων δεδομένων, διαφορετικών μεθόδων εξαγωγής χαρακτηριστικών, διαφορετικών υπερπαραμέτρων εξαγωγής καθώς και διαφορετικών τιμών για τις υπερπαραμέτρους. Παρομοίως θα μπορούσε να επεκταθεί το εύρος της έρευνας με τη χρήση διαφορετικών ταξινομητών, διαφορετικών υπερπαραμέτρων εκπαίδευσης και διαφορετικών τιμών για αυτές τις υπερπαραμέτρους.

Μελλοντικό ενδιαφέρον παρουσιάζει επίσης η επέκταση του ίδιου του προγράμματος εκπαίδευσης με διαφορετικό τρόπο προσέγγισης όπου ο διαχωρισμός των δεδομένων να γίνει με διαφορετικό τρόπο, ή ακόμα η μέθοδος ensemble να μην εφαρμοστεί και να χωριστούν τα τελικά σύνολα σε διαφορετική αναλογία συνόλων εκπαίδευσης-ελέγχου. Θα μπορούσε να εφαρμοστεί ακριβώς η ίδια μεθοδολογία σε μεγαλύτερο σύνολο δεδομένων, ίσως ακόμα και σε συνδυασμό δεδομένων από διαφορετικές βάσεις δεδομένων. Θα μπορούσαν επιπρόσθετα να διερευνηθούν κι άλλου τύπου αρχεία εκτός των αρχείων ήχου βήχα, όπως για παράδειγμα αρχεία ήχου της αναπνοής, που περιέχονται ήδη στις δύο βάσεις δεδομένων που ερευνηθήκαν, και να διερευνηθεί ο συσχετισμός μεταξύ των αποτελεσμάτων εκπαίδευσης ώστε να διερευνηθεί ποια δεδομένα ήχου κουβαλάνε «περισσότερη» σχετική πληροφορία για την ασθένεια. Σημαντικό θα ήταν να μελετηθούν και να εξεταστούν κι άλλες ασθένειες και να γίνει συσχετισμός μεταξύ εξαγόμενης πληροφορίας και ακρίβειας αποτελεσμάτων για κάθε υπό μελέτη ασθένεια.

Περαιτέρω μελλοντική προέκταση της παρούσας διπλωματικής συνιστά η διερεύνηση διαφορετικών τρεχόντων ζητημάτων της μηχανικής μάθησης, π.χ. οι μελλοντικές έρευνες θα μπορούσαν να διερευνήσουν το data shift των σχέσεων μεταξύ δεδομένων για να εξεταστεί η επίδοση των ταξινομητών σε βάθος χρόνου καθώς και το ζήτημα της εξηγησιμότητας, δηλαδή την προσπάθεια κατανόησης των προβλέψεων των ταξινομητών, ειδικά σε τομέα τόσο σημαντικό όσο ο τομέας υγείας.

Η μηχανική μάθηση, και ακόμα περισσότερο η βαθιά μάθηση, έχουν ακόμα πολλές πτυχές έρευνας στον τομέα της βιοϊατρικής και περαιτέρω συστηματικές έρευνες με μεγαλύτερη έκταση και σε ευρύτερο ερευνητικό πεδίο μπορούν να προσφέρουν σημαντικές νέες εφαρμογές αυτής της ραγδαίας αναπτυσσόμενης τεχνολογίας.

Βιβλιογραφία – Σημειώσεις – Πηγές

- [1] “Πολιτική Απορρήτου | SMARTY 4 COVID.”
<https://www.smarty4covid.org/privacy-policy> (accessed Mar. 07, 2024).
- [2] “SMARTY 4 COVID.” <https://www.smarty4covid.org/> (accessed Mar. 07, 2024).
- [3] “Coronavirus disease 2019 (COVID-19): situation report, 51.”
<https://iris.who.int/handle/10665/331475> (accessed Mar. 07, 2024).
- [4] N. Zhu *et al.*, “A Novel Coronavirus from Patients with Pneumonia in China, 2019,” *N. Engl. J. Med.*, vol. 382, no. 8, pp. 727–733, Feb. 2020, doi: 10.1056/nejmoa2001017.
- [5] “COVID-19 Overview and Infection Prevention and Control Priorities in non-U.S. Healthcare Settings | CDC.”
https://archive.cdc.gov/www_cdc_gov/coronavirus/2019-ncov/hcp/non-us-settings/overview/index.html (accessed Mar. 07, 2024).
- [6] F. P. Polack *et al.*, “Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine,” *N. Engl. J. Med.*, vol. 383, no. 27, pp. 2603–2615, Dec. 2020, doi: 10.1056/NEJMOA2034577/SUPPL_FILE/NEJMOA2034577_DATA-SHARING.PDF.
- [7] “COVID-19 Map - Johns Hopkins Coronavirus Resource Center.”
<https://coronavirus.jhu.edu/map.html> (accessed Mar. 07, 2024).
- [8] “COVID—19.” <https://lab.imedd.org/covid19/> (accessed Aug. 23, 2024).
- [9] E. Dong, H. Du, and L. Gardner, “An interactive web-based dashboard to track COVID-19 in real time,” *Lancet Infect. Dis.*, vol. 20, no. 5, pp. 533–534, May 2020, doi: 10.1016/S1473-3099(20)30120-1.
- [10] Q. Li *et al.*, “Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia,” *N. Engl. J. Med.*, vol. 382, no. 13, pp. 1199–1207, Mar. 2020, doi:

10.1056/NEJMOA2001316/SUPPL_FILE/NEJMOA2001316_DISCLOSURES.PDF.

- [11] “China confirms sharp rise in cases of SARS-like virus across the country.”
<https://web.archive.org/web/20200120055618/https://www.france24.com/en/20200120-china-confirms-sharp-rise-in-cases-of-sars-like-virus-across-the-country> (accessed Aug. 23, 2024).
- [12] Z. Feng *et al.*, “The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) — China, 2020,” *China CDC Weekly*, 2020, Vol. 2, Issue 8, Pages 113-122, vol. 2, no. 8, pp. 113–122, Feb. 2020, doi: 10.46234/CCDCW2020.032.
- [13] “COVID-19 - Βικιπαίδεια.” <https://el.wikipedia.org/wiki/COVID-19> (accessed Aug. 23, 2024).
- [14] “Πανδημία COVID-19 - Βικιπαίδεια.”
https://el.wikipedia.org/wiki/Πανδημία_COVID-19# (accessed Aug. 23, 2024).
- [15] E. Chatzidaki *et al.*, “Parenting Styles and Psychosocial Factors of Mother-Child Dyads Participating in the ENDORSE Digital Weight Management Program for Children and Adolescents during the COVID-19 Pandemic,” 2024, doi: 10.3390/children11010107.
- [16] A. M. Turing, “Computing Machinery and Intelligence,” *Parsing Turing Test Philos. Methodol. Issues Quest Think. Comput.*, pp. 23–65, 2009, doi: 10.1007/978-1-4020-6710-5_3.
- [17] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, Nov. 1958, doi: 10.1037/H0042519.
- [18] “Machine Learning textbook.” <http://www.cs.cmu.edu/~tom/mlbook.html> (accessed Mar. 22, 2024).
- [19] C. M. Bishop, *Prml*. 2006.
- [20] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nat.* 2015 5217553, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [21] I. Goodfellow, “Deep Learning,” *Prmu*, pp. 1–10, 2016, Accessed: Mar. 22, 2024. [Online]. Available:

https://books.google.com/books/about/Deep_Learning.html?hl=en&id=omivDQAAQBAJ.

- [22] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133, Dec. 1943, doi: 10.1007/BF02478259/METRICS.
- [23] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nat. 1986 3236088*, vol. 323, no. 6088, pp. 533–536, 1986, doi: 10.1038/323533a0.
- [24] G. E. Hinton, S. Osindero, and Y. W. Teh, “A Fast Learning Algorithm for Deep Belief Nets,” *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006, doi: 10.1162/NECO.2006.18.7.1527.
- [25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, Accessed: Sep. 07, 2024. [Online]. Available: <http://code.google.com/p/cuda-convnet/>.
- [26] “Machine Learning – nowmag.gr.” <https://nowmag.gr/machine-learning/> (accessed Aug. 29, 2024).
- [27] R. Vaishya, M. Javaid, I. H. Khan, and A. Haleem, “Artificial Intelligence (AI) applications for COVID-19 pandemic,” *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 14, no. 4, pp. 337–339, Jul. 2020, doi: 10.1016/J.DSX.2020.04.012.
- [28] P. Arora, H. Kumar, and B. K. Panigrahi, “Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India,” *Chaos, Solitons & Fractals*, vol. 139, p. 110017, Oct. 2020, doi: 10.1016/J.CHAOS.2020.110017.
- [29] T. T. Nguyen *et al.*, “Artificial Intelligence in the Battle against Coronavirus (COVID-19): A Survey and Future Research Directions,” Jul. 2020, Accessed: Sep. 22, 2024. [Online]. Available: <https://arxiv.org/abs/2008.07343v4>.
- [30] F. Shi *et al.*, “Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation, and Diagnosis for COVID-19,” *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 4–15, 2021, doi: 10.1109/RBME.2020.2987975.
- [31] A. Alimadadi, S. Aryal, I. Manandhar, P. B. Munroe, B. Joe, and X. Cheng,

- “Artificial intelligence and machine learning to fight covid-19,” *Physiol. Genomics*, vol. 52, no. 4, pp. 200–202, Apr. 2020, doi: 10.1152/PHYSIOLGENOMICS.00029.2020/ASSET/IMAGES/LARGE/ZH70042044120001.JPEG.
- [32] S. Latif *et al.*, “Leveraging Data Science to Combat COVID-19: A Comprehensive Review,” *IEEE Trans. Artif. Intell.*, vol. 1, no. 1, pp. 85–103, Aug. 2020, doi: 10.1109/TAI.2020.3020521.
- [33] A. Baygül Eden, A. Bakir Kayi, M. G. Erdem, and M. Demirci, “COVID-19 studies involving machine learning methods: A bibliometric study,” *Medicine (Baltimore)*, vol. 102, no. 43, p. E35564, Oct. 2023, doi: 10.1097/MD.00000000000035564.
- [34] E. S. Adamidi, K. Mitsis, and K. S. Nikita, “Artificial intelligence in clinical care amidst COVID-19 pandemic: A systematic review-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>),” *Comput. Struct. Biotechnol. J.*, vol. 19, pp. 2833–2851, 2021, doi: 10.1016/j.csbj.2021.05.010.
- [35] A. Abd-Alrazaq *et al.*, “A Comprehensive Overview of the COVID-19 Literature: Machine Learning–Based Bibliometric Analysis,” *J Med Internet Res* 2021;23(3)e23703 <https://www.jmir.org/2021/3/e23703>, vol. 23, no. 3, p. e23703, Mar. 2021, doi: 10.2196/23703.
- [36] L. L. Wang *et al.*, “CORD-19: The Covid-19 Open Research Dataset,” *ArXiv*, 2020, Accessed: Sep. 22, 2024. [Online]. Available: /pmc/articles/PMC7251955/.
- [37] H. Else, “How a torrent of COVID science changed research publishing - in seven charts,” *Nature*, vol. 588, no. 7839, p. 553, Dec. 2020, doi: 10.1038/D41586-020-03564-Y.
- [38] G. Colavizza, R. Costas, V. A. Traag, N. J. van Eck, T. van Leeuwen, and L. Waltman, “A scientometric overview of CORD-19,” *PLoS One*, vol. 16, no. 1, p. e0244839, Jan. 2021, doi: 10.1371/JOURNAL.PONE.0244839.
- [39] I. Rodríguez-Rodríguez *et al.*, “Applications of Artificial Intelligence, Machine Learning, Big Data and the Internet of Things to the COVID-19 Pandemic: A Scientometric Review Using Text Mining,” *Int. J. Environ. Res. Public Heal*. 2021, Vol. 18, Page 8578, vol. 18, no. 16, p. 8578, Aug. 2021, doi: 10.3390/IJERPH18168578.

- [40] “allenai/cord19: Get started with CORD-19.” <https://github.com/allenai/cord19> (accessed Aug. 10, 2024).
- [41] “Google Scholar.” <https://scholar.google.com/> (accessed Aug. 10, 2024).
- [42] “PubMed.” <https://pubmed.ncbi.nlm.nih.gov/> (accessed Aug. 10, 2024).
- [43] “Théorie analytique de la chaleur - Jean Baptiste Joseph Fourier - Βιβλία Google.” [https://books.google.gr/books?hl=el&lr=&id=Ez-dvn6anL8C&oi=fnd&pg=PR3&dq=Fourier,+J.+B.+J.+\(1822\).+Théorie+analytique+de+la+chaleur.+Paris:+Firmin+Didot+Père+et+Fils.&ots=YJ48jjXQrz&sig=PmKvu w7-QQH5qndPl15AEQ2elcQ&redir_esc=y#v=onepage&q&f=false](https://books.google.gr/books?hl=el&lr=&id=Ez-dvn6anL8C&oi=fnd&pg=PR3&dq=Fourier,+J.+B.+J.+(1822).+Théorie+analytique+de+la+chaleur.+Paris:+Firmin+Didot+Père+et+Fils.&ots=YJ48jjXQrz&sig=PmKvu w7-QQH5qndPl15AEQ2elcQ&redir_esc=y#v=onepage&q&f=false) (accessed Sep. 07, 2024).
- [44] S. G. Mallat, “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, 1989, doi: 10.1109/34.192463.
- [45] “(PDF) Fundamental of Speech Recognition - (Lawrence Rabiner - Biing Hwang Juang) | ziaul haq - Academia.edu.” https://www.academia.edu/4924307/Fundamental_of_Speech_Recognition_Lawrence_Rabiner_Biing_Hwang_Juang_ (accessed Mar. 22, 2024).
- [46] A. Graves, A. R. Mohamed, and G. Hinton, “Speech recognition with deep recurrent neural networks,” *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 6645–6649, Oct. 2013, doi: 10.1109/ICASSP.2013.6638947.
- [47] K. Zarkogianni *et al.*, “The smarty4covid dataset and knowledge base as a framework for interpretable physiological audio data analysis,” doi: 10.1038/s41597-023-02646-6.
- [48] N. Sharma *et al.*, “Coswara -- A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis,” *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2020-October, pp. 4811–4815, May 2020, doi: 10.21437/interspeech.2020-2768.
- [49] A. Muguli *et al.*, “DiCOVA Challenge: Dataset, task, and baseline system for COVID-19 diagnosis using acoustics,” *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 6, pp. 4241–4245, Mar. 2021, doi: 10.21437/Interspeech.2021-74.

- [50] J. Han *et al.*, “An Early Study on Intelligent Analysis of Speech under COVID-19: Severity, Sleep Quality, Fatigue, and Anxiety,” *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2020-October, pp. 4946–4950, Apr. 2020, doi: 10.21437/Interspeech.2020-2223.
- [51] A. Anupam, N. J. Mohan, S. Sahoo, and S. Chakraborty, “Preliminary diagnosis of COVID-19 based on cough sounds using machine learning algorithms,” *Proc. - 5th Int. Conf. Intell. Comput. Control Syst. ICICCS 2021*, pp. 1391–1397, May 2021, doi: 10.1109/ICICCS51141.2021.9432324.
- [52] S. B. Davis and P. Mermelstein, “Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences,” *IEEE Trans. Acoust.*, vol. 28, no. 4, pp. 357–366, 1980, doi: 10.1109/TASSP.1980.1163420.
- [53] G. Tzanetakis and P. Cook, “Musical genre classification of audio signals,” *IEEE Trans. Speech Audio Process.*, vol. 10, no. 5, pp. 293–302, Jul. 2002, doi: 10.1109/TSA.2002.800560.
- [54] T. Li and M. Ogihara, “Music genre classification with taxonomy,” *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. V, 2005, doi: 10.1109/ICASSP.2005.1416274.
- [55] V. Sharma, “A Study on Data Scaling Methods for Machine Learning,” *Int. J. Glob. Acad. Sci. Res.*, vol. 1, no. 1, pp. 31–42, Feb. 2022, doi: 10.55938/ijgasr.v1i1.4.
- [56] J. P. Bharadiya, “A Tutorial on Principal Component Analysis for Dimensionality Reduction in Machine Learning,” 2023. [Online]. Available: www.ijisrt.com.
- [57] N. Thomas Rincy and R. Gupta, “Ensemble learning techniques and its efficiency in machine learning: A survey,” *2nd Int. Conf. Data, Eng. Appl. IDEA 2020*, Feb. 2020, doi: 10.1109/IDEA49133.2020.9170675.
- [58] T. Hastie, J. Friedman, and R. Tibshirani, “The Elements of Statistical Learning,” 2001, doi: 10.1007/978-0-387-21606-5.
- [59] M. Pahar, M. Klopper, R. Warren, and T. Niesler, “COVID-19 cough classification using machine learning and global smartphone recordings,” *Comput. Biol. Med.*, vol. 135, p. 104572, Aug. 2021, doi: 10.1016/J.COMPBIOMED.2021.104572.
- [60] “Dive into Deep Learning — Dive into Deep Learning 1.0.3 documentation.” <https://d2l.ai/> (accessed Jul. 21, 2024).

- [61] D. R. Cox, “The Regression Analysis of Binary Sequences,” *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 20, no. 2, pp. 215–232, Jul. 1958, doi: 10.1111/J.2517-6161.1958.TB00292.X.
- [62] “Applied Logistic Regression - David W. Hosmer, Jr., Stanley Lemeshow, Rodney X. Sturdivant - Βιβλία Google.”
[https://books.google.gr/books?hl=el&lr=&id=bRoxQBIZRd4C&oi=fnd&pg=PR13&dq=Hosmer,+D.+W.,+Lemeshow,+S.,+%26+Sturdivant,+R.+X.+\(2013\).+Applied+Logistic+Regression+\(3rd+ed.\).+Wiley.+&ots=kM2PxncYc9&sig=yug4xt7HWglkbYuZ9WRCaB9JhEE&redir_esc=y#v=onepage&q&f=fa](https://books.google.gr/books?hl=el&lr=&id=bRoxQBIZRd4C&oi=fnd&pg=PR13&dq=Hosmer,+D.+W.,+Lemeshow,+S.,+%26+Sturdivant,+R.+X.+(2013).+Applied+Logistic+Regression+(3rd+ed.).+Wiley.+&ots=kM2PxncYc9&sig=yug4xt7HWglkbYuZ9WRCaB9JhEE&redir_esc=y#v=onepage&q&f=fa) (accessed Jul. 21, 2024).
- [63] “Discriminatory Analysis: Nonparametric Discrimination, Consistency Properties - Evelyn Fix - Βιβλία Google.”
[https://books.google.gr/books?hl=el&lr=&id=s85PAQAAMAAJ&oi=fnd&pg=PP1&dq=Fix,+E.,+%26+Hodges,+J.+L.+\(1951\).+Discriminatory+Analysis.+Nonparametric+Discrimination:+Consistency+Properties.+Report+Number+4,+USAF+School+of+Aviation+Medicine.+&ots=MKOcz-9E1d&s](https://books.google.gr/books?hl=el&lr=&id=s85PAQAAMAAJ&oi=fnd&pg=PP1&dq=Fix,+E.,+%26+Hodges,+J.+L.+(1951).+Discriminatory+Analysis.+Nonparametric+Discrimination:+Consistency+Properties.+Report+Number+4,+USAF+School+of+Aviation+Medicine.+&ots=MKOcz-9E1d&s) (accessed Jul. 21, 2024).
- [64] T. M. Cover and P. E. Hart, “Nearest Neighbor Pattern Classification,” *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967, doi: 10.1109/TIT.1967.1053964.
- [65] B. H. Shekar and G. Dagneu, “Grid search-based hyperparameter tuning and classification of microarray cancer data,” *2019 2nd Int. Conf. Adv. Comput. Commun. Paradig. ICACCP 2019*, Feb. 2019, doi: 10.1109/ICACCP.2019.8882943.
- [66] X. Jiang and C. Xu, “Deep Learning and Machine Learning with Grid Search to Predict Later Occurrence of Breast Cancer Metastasis Using Clinical Data,” *J. Clin. Med.*, vol. 11, no. 19, p. 5772, Oct. 2022, doi: 10.3390/JCM11195772/S1.
- [67] “Receiver operating characteristic - Wikipedia.”
https://en.wikipedia.org/wiki/Receiver_operating_characteristic (accessed Sep. 14, 2024).