



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

**Μέθοδοι βαθιάς μάθησης και άμβλυνσης προκαταλήψεων για
την ανίχνευση της νόσου COVID-19 με χρήση ηχητικών
καταγραφών βήχα**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
ΔΗΜΗΤΡΑΣ ΚΩΣΤΑΒΑΣΙΔΗ**

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

Αθήνα, Ιούλιος 2025



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

**Μέθοδοι βαθιάς μάθησης και άμβλυνσης προκαταλήψεων για
την ανίχνευση της νόσου COVID-19 με χρήση ηχητικών
καταγραφών βήχα**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
ΔΗΜΗΤΡΑΣ ΚΩΣΤΑΒΑΣΙΔΗ**

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

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 4^η Ιουλίου 2025.

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

.....
Δρ. Γεώργιος Στάμου,
Καθηγητής
Σ.Η.Μ.Μ.Υ Ε.Μ.Π.

.....
Δρ. Αθανάσιος Βουλόδημος,
Επίκουρος Καθηγητής
Σ.Η.Μ.Μ.Υ Ε.Μ.Π.

Αθήνα, Ιούλιος 2025

.....

Δήμητρα Κωσταβασίλη

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

Copyright © Δήμητρα Κωσταβασίλη, 2025.

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 της COVID-19 και οι σοβαρές επιπτώσεις της νόσου αποτελούν σημαντική πρόκληση για την επιστημονική κοινότητα, καθώς δημιουργούν την ανάγκη για ταχύτερα διαγνωστικά μέσα. Το γεγονός αυτό, ωθεί την ανάπτυξη καινοτόμων τεχνολογικών εργαλείων που στοχεύουν κυρίως στον περιορισμό του ιού και τη διερεύνηση της διασποράς του. Η μηχανική μάθηση αποτελεί βασικό πυλώνα της εξέλιξης των εργαλείων αυτών, συμβάλλοντας σημαντικά στη βελτίωση της διαδικασίας διάγνωσης. Παρά την υποχώρηση του πανδημικού χαρακτήρα της νόσου, η καινοτομία των συγκεκριμένων εφαρμογών δεν χάνει τη χρησιμότητά της καθώς αποτελεί πρόοδο στον κλάδο της βιοϊατρικής με σημαντικές προεκτάσεις στα πεδία της αυτοδιάγνωσης και της τηλεϊατρικής, αλλά και γενικότερα στην εξέλιξη των μοντέλων μηχανικής μάθησης. Σημαντικό κριτήριο στην αξιολόγηση της αποτελεσματικότητας των διαγνωστικών εργαλείων που βασίζονται στη μηχανική μάθηση πέρα από την ακρίβεια της πρόβλεψης, αποτελεί και η έννοια της αμεροληψίας. Η αμεροληψία συνοψίζεται στην δίκαιη εξαγωγή προβλέψεων από ένα μοντέλο μηχανικής μάθησης ανάμεσα σε άτομα ή υποομάδες του πληθυσμού, ανεξάρτητα από πιθανές διαφορές στα ευαίσθητα χαρακτηριστικά τους.

Στόχος της παρούσας διπλωματικής εργασίας, είναι η ανάπτυξη μοντέλων μηχανικής μάθησης για τη διάγνωση της νόσου COVID-19 με τη χρήση ηχητικών καταγραφών βήχα και τη βελτίωση της αμεροληψίας των μοντέλων αυτών με την εφαρμογή μεθόδων εξάλειψης των διακρίσεων. Ειδικότερα, αναπτύσσεται ένας απλός ταξινομητής τυχαίου δάσους και ένα τεχνητό νευρωνικό δίκτυο βασισμένο στο προεκπαιδευμένο μοντέλο βαθιάς μάθησης VGGish. Το σύνολο δεδομένων που χρησιμοποιήθηκε προήλθε από τη βάση δεδομένων COVID-19 Sounds του πανεπιστημίου Cambridge που περιλαμβάνει μια σειρά ηχογραφήσεων ανθρώπινων ήχων όπως ο βήχας, η αναπνοή και η ομιλία από νοσούντες και μη, συνοδευόμενων από πληροφορίες για τα ευαίσθητα χαρακτηριστικά του ατόμου στο οποίο αντιστοιχεί κάθε ηχογράφιση. Για τη βελτίωση του βαθμού αμεροληψίας των μοντέλων, εφαρμόστηκαν: ως μέθοδος προεπεξεργασίας η αφαίρεση συσχετίσεων, μέθοδοι κατά την επεξεργασία όπως η εκθετική κλίση και η ανταγωνιστική εκπαίδευση και ο βελτιστοποιητής κατωφλίου που αποτελεί μέθοδο μετά την επεξεργασία. Έπειτα από σύγκριση των αποτελεσμάτων των μοντέλων πριν και μετά την εφαρμογή κάθε μεθόδου, παρατηρείται ότι επιτυγχάνεται βελτίωση της αμεροληψίας των αποτελεσμάτων σε αρκετές κατηγορίες ευαίσθητων χαρακτηριστικών. Ενδεικτικά, οι καλύτερες τιμές της αναλογίας ίσης ευκαιρίας, που αποτελεί βασική μετρική της αμεροληψίας και έχει ιδανική τιμή 1, είναι: ως προς το φύλο 0,969, ως προς την ηλικία 0,998, ως προς το ιατρικό ιστορικό 0,972, ως προς τα συμπτώματα 0,997 και ως προς την πηγή συλλογής των δεδομένων 0,88. Αντίστοιχα, για τις συνδυαστικές κατηγορίες, τα βέλτιστα αποτελέσματα για την ίδια μετρική είναι: για το φύλο σε συνδυασμό με την ηλικία 0,943, για την ηλικία σε συνδυασμό με τα συμπτώματα 0,951, για το φύλο σε συνδυασμό με τα συμπτώματα 0,91, για την ηλικία σε συνδυασμό με το ιατρικό ιστορικό 0,911 και για το φύλο σε συνδυασμό με το ιατρικό ιστορικό 0,913. Ωστόσο, λαμβάνοντας υπόψη και την ακρίβεια των προβλέψεων, συχνά επιλέγονται μέθοδοι που επιτυγχάνουν καλύτερη ισορροπία μεταξύ ακρίβειας και αμεροληψίας, ακόμα και με κάποιο κόστος εξετάζοντας τα μεγέθη μεμονωμένα.

Λέξεις Κλειδιά

COVID-19, Μηχανική Μάθηση, Βαθιά Μάθηση, Επεξεργασία ήχου, Ηχητικά σήματα, Αμεροληψία, Εξάλειψη των διακρίσεων, VGGish

Abstract

The COVID-19 pandemic and the severe impact of the disease have imposed a significant challenge to the scientific community, creating the need for faster diagnostic means. This has led to the development of innovative technological tools, primarily aimed at containing the virus and investigating its spread. Machine learning serves as a key pillar in the advancement of these tools, significantly contributing to the improvement of the diagnostic process. Despite the decline of the pandemic status of the disease, the innovation of these applications remains valuable, representing progress in the field of biomedical science with important implications in self-diagnosis, telemedicine, and the broader development of machine learning models. A significant criterion in evaluating the effectiveness of machine learning-based diagnostic tools is, in addition to prediction accuracy, the concept of fairness. Fairness is the equitable generation of predictions by a machine learning model among individuals or subgroups of the population, regardless of potential differences in their sensitive attributes.

The aim of the present thesis is the development of machine learning models for diagnosing COVID-19 using cough audio recordings, as well as the improvement of these models' fairness through the application of bias mitigation methods. In particular, we develop a basic random forest classifier and an artificial neural network based on the pre-trained deep learning model VGGish. The dataset used comes from the COVID-19 Sounds database of the University of Cambridge, which includes a series of human sound recordings such as coughing, breathing, and speech, from both infected and non-infected individuals. The recordings are also accompanied by information on the sensitive attributes of the person corresponding to each recording. To improve the fairness of the models, the following debiasing methods were applied: as a pre-processing method, correlation remover; in-processing methods, such as exponentiated gradient and adversarial learning; and a post-processing method, the threshold optimizer. After comparing the models' performance before and after the application of each method, it was observed that fairness improved across several sensitive attribute categories. Indicatively, the best equal opportunity ratio values, which is a key fairness metric with an ideal value of 1, were: with respect to gender, 0.969, age, 0.998, medical history, 0.972, symptoms, 0.997, and data collection source, 0.88. Similarly, for combined attribute categories, the optimal results for the same metric were: gender and age, 0.943, age and symptoms, 0.951, gender and symptoms, 0.91, age and medical history, 0.911, and gender and medical history, 0.913. However, when also considering the accuracy of the predictions, methods that achieve a better balance between accuracy and fairness are often preferred, even at some cost when assessing these metrics individually.

Keywords

COVID-19, Machine Learning, Deep Learning, Audio Processing, Audio Signals, Fairness, Bias Mitigation, VGGish

Ευχαριστίες

Κατ'αρχάς, θα ήθελα να εκφάσω τις θερμές μου ευχαριστίες στην καθηγήτριά μου, Δρ. Κωνσταντίνα Νικήτα, για την ευκαιρία που μου έδωσε να εκπονήσω την διπλωματική μου εργασία στο εργαστήριο Βιοϊατρικών Προσομοιώσεων και Απεικονιστικής Τεχνολογίας (BIOSIM), την ανάθεση του ιδιαίτερα ενδιαφέροντος θέματος και την διαρκή της καθοδήγηση και έμπνευση κατά τη διάρκεια των σπουδών μου. Ευχαριστώ επίσης, ιδιαιτέρως, τον Υποψήφιο Διδάκτορα Θεοφάνη Γανιτίδη για τη συνεργασία και τις πολύτιμες συμβουλές του, οι οποίες συνέβαλαν καθοριστικά στην ολοκλήρωση της εργασίας με επιτυχία. Τέλος, ευχαριστώ την οικογένεια και τους φίλους μου για την υποστήριξη και την αγάπη τους σε όλη μου τη σταδιοδρομία.

Αθήνα, Ιούλιος 2025

Δήμητρα Κωσταβασίλη

Πίνακας περιεχομένων

Περίληψη.....	7
Abstract.....	8
Ευχαριστίες.....	10
1 Εισαγωγή.....	24
1.1 Εισαγωγικά Στοιχεία.....	24
1.2 Σκοπός της παρούσας διπλωματικής εργασίας.....	25
1.3 Δομή της διπλωματικής εργασίας.....	25
2 COVID-19.....	27
3 Μηχανική Μάθηση.....	30
3.1 Μηχανική Μάθηση και Βαθιά Μάθηση.....	30
3.2 Μεθοδολογίες Εκπαίδευσης Μοντέλων Μηχανικής Μάθησης.....	31
3.3 Αλγόριθμοι Μηχανικής Μάθησης για Ταξινόμηση.....	32
3.4 Τεχνητά Νευρωνικά Δίκτυα.....	33
3.5 Συνελκτικά Νευρωνικά Δίκτυα.....	36
3.5.1 Δομή και επίπεδα.....	36
3.5.2 Το Συνελκτικό Νευρωνικό Δίκτυο VGG.....	37
3.5.3 Το Συνελκτικό Νευρωνικό Δίκτυο VGGish.....	38
3.6 Εξαγωγή Χαρακτηριστικών από Ηχητικά Δεδομένα.....	38
3.6.1 Χαρακτηριστικά εισόδου σε Απλούς Ταξινομητές.....	39
3.6.2 Χαρακτηριστικά Εισόδου σε Συνελκτικά Νευρωνικά Δίκτυα.....	40
3.7 Αξιολόγηση Αλγορίθμων Μηχανικής Μάθησης.....	40
3.8 Μηχανική Μάθηση στην Ιατρική.....	43
4 Αμεροληψία στη Μηχανική Μάθηση.....	44
4.1 Η έννοια της Αμεροληψίας στη Μηχανική Μάθηση.....	44
4.2 Αμεροληψία στον τομέα της Ιατρικής.....	46
4.3 Αντιμετώπιση των Διακρίσεων.....	46
5 Εφαρμογές Μηχανικής Μάθησης κατά τη διάρκεια της Πανδημίας COVID-19.....	49
6 Μεθοδολογία.....	50
6.1 Δεδομένα.....	50
6.1.1 Περιγραφή Συνόλου Δεδομένων.....	50
6.1.2 Προεπεξεργασία Δεδομένων.....	51
6.1.3 Στατιστική Ανάλυση των Ευαίσθητων Χαρακτηριστικών.....	54
6.1.4 Εξαγωγή Χαρακτηριστικών.....	61
6.2 Ανάπτυξη Μοντέλου Μηχανικής Μάθησης.....	64
6.2.1 Εκπαίδευση Ταξινομητή Τυχάιου Δάσους.....	64
6.3 Ανάπτυξη Μοντέλου Βαθιάς Μάθησης.....	64
6.3.1 Εκπαίδευση Συνελκτικού Νευρωνικού Δικτύου.....	64
6.4 Αμεροληψία και Αντιμετώπιση των Διακρίσεων.....	66
6.4.1 Υλοποίηση Συναρτήσεων για τον υπολογισμό της Αμεροληψίας.....	66
6.4.2 Αντιμετώπιση των Διακρίσεων με Μεθόδους Προεπεξεργασίας.....	67
6.4.2.1 Αφαίρεση Συσχετίσεων (Correlation Remover).....	67
6.4.3 Αντιμετώπιση των Διακρίσεων με Μεθόδους κατά την Επεξεργασία.....	67
6.4.3.1 Εκθετική Κλίση (Exponentiated Gradient).....	67
6.4.3.2 Ανταγωνιστική Εκπαίδευση (Adversarial Learning).....	68
6.4.4 Αντιμετώπιση των Διακρίσεων με Μεθόδους μετά την Επεξεργασία.....	68
6.4.4.1 Βελτιστοποιητής Κατωφλίου (Threshold Optimizer).....	68
7 Αποτελέσματα.....	70
7.1 Ταξινόμηση.....	70

7.1.1 Ταξινόμηση με ενίσχυση της Αμεροληψίας Φύλου.....	70
7.1.2 Ταξινόμηση με ενίσχυση της Αμεροληψίας Ηλικίας.....	71
7.1.3 Ταξινόμηση με ενίσχυση της Αμεροληψίας του Ιατρικού Ιστορικού.....	72
7.1.4 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς τα Συμπτώματα.....	74
7.1.5 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς την Πηγή Συλλογής Δεδομένων...76	
7.1.6 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς την Ηλικία σε συνδυασμό με το Φύλο.....	77
7.1.7 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς την Ηλικία σε συνδυασμό με τα Συμπτώματα.....	79
7.1.8 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς το Φύλο σε συνδυασμό με τα Συμπτώματα.....	80
7.1.9 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς την Ηλικία σε συνδυασμό με το Ιατρικό Ιστορικό.....	82
7.1.10 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς το Φύλο σε συνδυασμό με το Ιατρικό Ιστορικό.....	83
7.1.11 Συνολική Αξιολόγηση Ταξινόμησης συγκριτικά για τις Μεθόδους Ενίσχυσης της Αμεροληψίας.....	85
7.2 Αμεροληψία.....	86
7.2.1 Αμεροληψία ως προς το Φύλο.....	86
7.2.2 Αμεροληψία ως προς την Ηλικία.....	91
7.2.3 Αμεροληψία ως προς το Ιατρικό Ιστορικό.....	96
7.2.4 Αμεροληψία ως προς τα Συμπτώματα.....	101
7.2.5 Αμεροληψία ως προς την Πηγή Συλλογής Δεδομένων.....	106
7.2.6 Αμεροληψία ως προς την Ηλικία σε συνδυασμό με το Φύλο.....	111
7.2.7 Αμεροληψία ως προς την Ηλικία σε συνδυασμό με τα Συμπτώματα.....	116
7.2.8 Αμεροληψία ως προς το Φύλο σε συνδυασμό με τα Συμπτώματα.....	122
7.2.9 Αμεροληψία ως προς την Ηλικία σε συνδυασμό με το Ιατρικό Ιστορικό.....	129
7.2.10 Αμεροληψία ως προς το Φύλο σε συνδυασμό με το Ιατρικό Ιστορικό.....	134
7.2.11 Μεταβολές Μετρικών Αμεροληψίας μετά την εφαρμογή Μεθόδων Εξάλειψης των Διακρίσεων.....	140
8 Συμπεράσματα – Μελλοντικές Προεκτάσεις.....	143
Βιβλιογραφία.....	145

Κατάλογος Εικόνων

Εικόνα 2.1: Εικόνα από ηλεκτρονικό μικροσκόπιο διέλευσης που δείχνει τον SARS-CoV-2, τον ιό που προκαλεί την COVID-19, απομονωμένο από ασθενή στις ΗΠΑ. Οι ακίδες στην εξωτερική επιφάνεια των σωματιδίων του ιού δίνουν στους κορωνοϊούς το όνομά τους, λόγω της κορώνας που σχηματίζουν [15].....	27
Εικόνα 2.2: Ενημερωτική αφίσα του ΠΟΥ την εποχή της πανδημίας με οδηγίες για τις μεθόδους διάγνωσης της COVID-19 [19].....	29
Εικόνα 3.1.1: Σχηματική απεικόνιση της Επιστήμης των Δεδομένων με τα διαδихικά υποσύνολά της: Τεχνητή Νοημοσύνη, Μηχανική Μάθηση, Βαθιά Μάθηση, Τεχνητά Νευρωνικά Δίκτυα [22].	30
Εικόνα 3.1.2: Διάγραμμα ενός απλού Τεχνητού Νευρωνικού Δικτύου που απεικονίζει τις μαθηματικές σχέσεις που περιγράφουν έναν απλό νευρώνα [23].....	31
Εικόνα 3.3.1: Σχηματική απεικόνιση Δυαδικού Ταξινομητή, που κατατάσσει τα δείγματα (samples) σε δύο κατηγορίες θετική (positive) και αρνητική (negative) και στη συνέχεια αξιολογεί τις προβλέψεις του [26].....	33
Εικόνα 3.5.1: Σχηματική αναπαράσταση της βασικής αρχιτεκτονικής ενός Συνελικτικού Νευρωνικού Δικτύου [35].....	37
Εικόνα 3.5.2: Σχηματική απεικόνιση της χρήσης του μοντέλου VGGish για αναγνώριση συναισθημάτων από ηχητικά σήματα.....	38
Εικόνα 3.6.1: Απεικόνιση Mel-φασματογραφήματος σήματος ήχου για είσοδο σε συνελικτικό νευρωνικό δίκτυο [41].....	40
Εικόνα 3.7.1: Απεικόνιση Πίνακα Σύγχυσης με βασικές μετρικές αξιολόγησης για την απόδοση μοντέλων ταξινόμησης [34].....	41
Εικόνα 3.7.2: Απεικόνιση καμπύλης ROC (Receiver Operating Characteristic Curve) [42].....	43
Εικόνα 4.1.1: Σχηματική απεικόνιση των μετρικών αξιολόγησης της Αμεροληψίας ως προς το φύλο, σε ένα πρόβλημα δυαδικής ταξινόμησης [48].....	45
Εικόνα 4.3.1: Σχηματική απεικόνιση των τεχνικών εξάλειψης των διακρίσεων στα μοντέλα Μηχανικής Μάθησης, οι οποίες μπορούν να εφαρμοστούν είτε μεμονωμένα είτε συνδυαστικά [53].	48
Εικόνα 6.1.1: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Φύλου.	54
Εικόνα 6.1.2: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Ηλικίας.....	55
Εικόνα 6.1.3: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Ιατρικού ιστορικού.....	56
Εικόνα 6.1.4: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Συμπτωμάτων.....	56
Εικόνα 6.1.5: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Πηγής Συλλογής Δεδομένων.....	57
Εικόνα 6.1.6: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Ηλικίας σε συνδυασμό με Φύλο.....	58
Εικόνα 6.1.7: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Ηλικίας σε συνδυασμό με Συμπτώματα.....	59
Εικόνα 6.1.8: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Φύλου σε συνδυασμό με Συμπτώματα.....	59
Εικόνα 6.1.9: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Ηλικίας σε συνδυασμό με Ιατρικό ιστορικό.....	60
Εικόνα 6.1.10: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Φύλου σε συνδυασμό με Ιατρικό ιστορικό.....	61

Εικόνα 6.1.11: Απεικόνιση των 68 τιμών χαρακτηριστικών επεξεργασίας ψηφιακού σήματος για το πρώτο κατά σειρά ηχητικό σήμα των δεδομένων μας.....	62
Εικόνα 6.1.12: Mel-φασματογράφημα της μετατροπής των δεδομένων εισόδου στο μοντέλο μας..	63
Εικόνα 6.3.1: Απεικόνιση των επιπέδων και του σχήματος εξόδου κάθε επιπέδου του μοντέλου που αναπτύξαμε.....	65
Εικόνα 7.1.1: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία του φύλου.....	71
Εικόνα 7.1.2: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία της ηλικίας.....	73
Εικόνα 7.1.3: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία ιατρικού ιστορικού.....	74
Εικόνα 7.1.4: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία των συμπτωμάτων.....	75
Εικόνα 7.1.5: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία πηγής συλλογής δεδομένων.....	77
Εικόνα 7.1.6: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία ηλικίας σε συνδυασμό με φύλο.....	78
Εικόνα 7.1.7: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία ηλικίας σε συνδυασμό με συμπτώματα.....	80
Εικόνα 7.1.8: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία φύλου σε συνδυασμό με συμπτώματα.....	81
Εικόνα 7.1.9: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία ηλικίας σε συνδυασμό με ιατρικό ιστορικό.....	83
Εικόνα 7.1.10: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία φύλου σε συνδυασμό με ιατρικό ιστορικό.....	84
Εικόνα 7.1.11: Σύγκριση μέσων τιμών μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για όλες τις κατηγορίες ευαίσθητων χαρακτηριστικών.	86
Εικόνα 7.2.1: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα του φύλου.....	87
Εικόνα 7.2.2: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα του φύλου.....	88
Εικόνα 7.2.3: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα του φύλου.....	89
Εικόνα 7.2.4: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα του φύλου.....	90
Εικόνα 7.2.5: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία του φύλου.....	91
Εικόνα 7.2.6: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα της ηλικίας.....	92
Εικόνα 7.2.7: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα της ηλικίας.....	93

[illegible]

[illegible]

εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα του φύλου σε συνδυασμό με τα συμπτώματα.....	126
Εικόνα 7.2.40: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία του φύλου σε συνδυασμό με τα συμπτώματα.....	128
Εικόνα 7.2.41: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.....	129
Εικόνα 7.2.42: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.....	130
Εικόνα 7.2.43: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.....	131
Εικόνα 7.2.44: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.....	133
Εικόνα 7.2.45: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.....	134
Εικόνα 7.2.46: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό.....	135
Εικόνα 7.2.47: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό.....	136
Εικόνα 7.2.48: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό.....	137
Εικόνα 7.2.49: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό.....	138
Εικόνα 7.2.50: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό.....	140

Κατάλογος Πινάκων

Πίνακας 6.1.1: Πίνακας κατηγοριοποίησης των ευαίσθητων χαρακτηριστικών των δεδομένων που χρησιμοποιήθηκαν στην παρούσα εργασία.....	51
Πίνακας 7.1.1: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία του φύλου.....	70
Πίνακας 7.1.2: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία του φύλου.....	70
Πίνακας 7.1.3: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία της ηλικίας.....	71
Πίνακας 7.1.4: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία της ηλικίας.....	72
Πίνακας 7.1.5: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία του ιατρικού ιστορικού.....	73
Πίνακας 7.1.6: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία του ιατρικού ιστορικού.....	73
Πίνακας 7.1.7: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία των συμπτωμάτων.....	74
Πίνακας 7.1.8: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία των συμπτωμάτων.....	75
Πίνακας 7.1.9: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία της πηγής συλλογής δεδομένων.....	76
Πίνακας 7.1.10: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία της πηγής συλλογής δεδομένων.....	76
Πίνακας 7.1.11: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία της ηλικίας σε συνδυασμό με το φύλο.....	77
Πίνακας 7.1.12: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία της ηλικίας σε συνδυασμό με το φύλο.....	78
Πίνακας 7.1.13: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα.....	79
Πίνακας 7.1.14: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα.....	79
Πίνακας 7.1.15: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία του φύλου σε συνδυασμό με τα συμπτώματα.....	80
Πίνακας 7.1.16: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία του φύλου σε συνδυασμό με τα συμπτώματα.....	81
Πίνακας 7.1.17: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.....	82
Πίνακας 7.1.18: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.....	82

Πίνακας 7.1.19: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό.....	83
Πίνακας 7.1.20: Πίνακας σύγκρισης μετρικών απόδοσης του συνελικτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό.....	84
Πίνακας 7.1.21: Πίνακας σύγκρισης μέσων τιμών των μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για όλες τις κατηγορίες.....	85
Πίνακας 7.1.22: Πίνακας σύγκρισης μέσων τιμών των μετρικών απόδοσης του συνελικτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων για όλες τις κατηγορίες.....	85
Πίνακας 7.2.1: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων του φύλου.....	87
Πίνακας 7.2.2: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων του φύλου.....	88
Πίνακας 7.2.3: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων του φύλου.....	89
Πίνακας 7.2.4: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων του φύλου.....	90
Πίνακας 7.2.5: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων της ηλικίας.....	92
Πίνακας 7.2.6: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων της ηλικίας.....	93
Πίνακας 7.2.7: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων της ηλικίας.....	94
Πίνακας 7.2.8: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων της ηλικίας.....	95
Πίνακας 7.2.9: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων του ιατρικού ιστορικού.....	97
Πίνακας 7.2.10: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων του ιατρικού ιστορικού.....	98
Πίνακας 7.2.11: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων του ιατρικού ιστορικού.....	99
Πίνακας 7.2.12: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων του ιατρικού ιστορικού.....	100
Πίνακας 7.2.13: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων των συμπτωμάτων.....	102
Πίνακας 7.2.14: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων των συμπτωμάτων.....	103
Πίνακας 7.2.15: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων των συμπτωμάτων.....	104
Πίνακας 7.2.16: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων των συμπτωμάτων.....	105
Πίνακας 7.2.17: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων της πηγής συλλογής δεδομένων.....	107
Πίνακας 7.2.18: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων της πηγής συλλογής δεδομένων.....	108
Πίνακας 7.2.19: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων της πηγής συλλογής δεδομένων.....	109

[illegible]

ιατρικό ιστορικό.....	138
Πίνακας 7.2.40: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων του φύλου σε συνδυασμό με το ιατρικό ιστορικό.....	139
Πίνακας 7.2.41: Πίνακας μεταβολών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή της αφαίρεσης συσχετίσεων. Με πράσινο χρώμα επισημαίνονται οι βελτιωμένες τιμές, ενώ με κόκκινο οι επιδεινωμένες.....	141
Πίνακας 7.2.42: Πίνακας μεταβολών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή της εκθετικής κλίσης. Με πράσινο χρώμα επισημαίνονται οι βελτιωμένες τιμές, ενώ με κόκκινο οι επιδεινωμένες.....	141
Πίνακας 7.2.43: Πίνακας μεταβολών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης. Με πράσινο χρώμα επισημαίνονται οι βελτιωμένες τιμές, ενώ με κόκκινο οι επιδεινωμένες.....	142
Πίνακας 7.2.44: Πίνακας μεταβολών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή του βελτιστοποιητή κατωφλίου. Με πράσινο χρώμα επισημαίνονται οι βελτιωμένες τιμές, ενώ με κόκκινο οι επιδεινωμένες.....	142

1 Εισαγωγή

1.1 Εισαγωγικά Στοιχεία

Η νόσος COVID-19 μπορεί να παρουσιαστεί μετά από μόλυνση με το νεότερο στέλεχος του κορονοϊού SARS-CoV-2. Αποτελεί μία σοβαρή ασθένεια που προσβάλλει κυρίως το αναπνευστικό σύστημα και για τους μη ασυμπτωματικούς ασθενείς προσομοιάζει σε κακοήγη πνευμονία, περιλαμβάνοντας όμως στο ευρύ φάσμα συμπτωμάτων και νευρολογικά συμπτώματα. Σε σοβαρές περιπτώσεις, η νόσος μπορεί να οδηγήσει μέχρι και σε θάνατο, με ομάδες υψηλότερου κινδύνου να αποτελούν οι ηλικιωμένοι και τα άτομα με υποκείμενα νοσήματα [1]. Παράγοντες όπως η ύπαρξη ασυμπτωματικών ασθενών, η σχετικά μεγάλη περίοδος επώασης του ιού και οι συνεχείς μεταλλάξεις του, οι οποίες δυσχεραίνουν τη ανάπτυξη αντισωμάτων, καθιστούν τον ιό ιδιαίτερα μεταδοτικό, γεγονός που τον Μάρτιο του 2020 οδήγησε τον ΠΟΥ στην ανακήρυξη την νόσου ως πανδημία [2]. Μέχρι και τον Ιανουάριο του 2025, η νόσος COVID-19 στοίχισε περισσότερες από 7,1 εκατομμύρια ανθρώπινες ζωές παγκοσμίως και παρά τη σημαντική πρόοδο στον περιορισμό της νόσου, ακόμα και σήμερα εξακολουθεί να είναι ιδιαίτερα ενεργή στον πληθυσμό [3]. Οι σοβαρότατες συνέπειες της εξάπλωσης της πανδημίας που οδήγησαν τον πλανήτη σε πρωτόγνωρες συνθήκες, ώθησαν την επιστημονική κοινότητα σε αναζήτηση μεθόδων έγκυρης και έγκαιρης διάγνωσης των κρουσμάτων.

Μεταξύ πολλών άλλων κλάδων της επιστήμης, η ιατρική μπορεί να ωφεληθεί σημαντικά από την ενσωμάτωση της τεχνητής νοημοσύνης τόσο στη διάγνωση ασθενειών όσο και στη βελτίωση των θεραπευτικών πρακτικών. Οι αλγόριθμοι μηχανική μάθησης, αξιοποιώντας την ευρεία ποικιλία των ψηφιακών ιατρικών δεδομένων όπως ηλεκτρονικοί φάκελοι υγείας (ΗΦΥ), βιολογικοί δείκτες, εικόνες, βίντεο και ηχητικά σήματα, που μας προσφέρουν οι σύγχρονες τεχνολογίες, μπορούν να συμβάλουν στην ταχύτερη και ακριβέστερη διάγνωση και υποβοήθηση των ιατρικών αποφάσεων [4], [5], [6]. Η Μηχανική Μάθηση, μπορεί επίσης να εξελίξει σε μεγάλο βαθμό τα εργαλεία αυτοδιάγνωσης, τα οποία μπορούν να αποτελέσουν βασικό παράγοντα στην προσπάθεια πρόληψης παθήσεων με μεγάλη ταχύτητα, προσβασιμότητα και χαμηλό κόστος [7], [8].

Η πανδημία COVID-19 και ως συνέπειές της, τα αυστηρά μέτρα περιορισμού, το υψηλό κόστος των διαγνωστικών ελέγχων και η γενικότερη αβεβαιότητα που προξένησε η εμφάνισή της, κατέστησαν επιτακτική την ανάγκη αξιοποίησης των εργαλείων που προσφέρει η μηχανική μάθηση για γρήγορη διάγνωση των ασθενών και διερεύνηση της διασποράς του ιού. Η φύση του ιού, ο οποίος προσβάλλει κυρίως το αναπνευστικό σύστημα αλλοιώνοντας το φυσιολογικό ηχητικό σήμα που παράγει ένα υγιές άτομο, στρέφει τις μεθόδους των μοντέλων διάγνωσης του στο πεδίο της επεξεργασίας ήχου. Παράλληλα, η συνεχής παρακολούθηση του ιού και των επιπτώσεών του οδήγησαν στη συσσώρευση μεγάλου όγκου νέων ιατρικών δεδομένων προς επεξεργασία για την εξαγωγή συμπερασμάτων σχετικά με τη νόσο. Σε αυτό το πλαίσιο, ευνοήθηκε η ανάπτυξη εφαρμογών μηχανικής μάθησης για την ανάλυση εικόνων και ήχων που σχετίζονταν με συμπτώματα της COVID-19. Έτσι, αναπτύχθηκαν εφαρμογές συλλογής ηχητικών δειγμάτων, προσβάσιμες από ηλεκτρονικές συσκευές με τη χρήση των μικροφώνων τους [9], [10], [11]. Αποτέλεσμα ήταν η δημιουργία μεγάλων συνόλων δεδομένων προς ανάλυση, τα οποία συνεχίζουν να χρησιμοποιούνται σε μοντέλα διάγνωσης. Αξίζει να σημειωθεί ότι η χρήση των εφαρμογών αυτών δεν περιορίζεται αποκλειστικά στην περίπτωση της COVID-19, αλλά μπορεί να διευρυνθεί και σε άλλες παθήσεις του αναπνευστικού συστήματος [12]. Επιπλέον, η πληθώρα και η διαθεσιμότητα των δεδομένων, ενθαρρύνει την περαιτέρω διερεύνηση και άλλων προεκτάσεων της μηχανικής μάθησης, όπως η εξασφάλιση αμεροληψίας. Συνολικά, οι εφαρμογές αυτές αποδείχθηκαν πολύτιμα εργαλεία για την αντιμετώπιση αυτής της παγκόσμιας υγειονομικής κρίσης και συνέβαλαν στην ανάπτυξη επιστημονικών μεθόδων στα πεδία της αυτοδιάγνωσης και της τηλεϊατρικής.

Η συνεχής εξέλιξη και ένταξη των εφαρμογών μηχανικής μάθησης στο πεδίο της ιατρικής έχει επαναπροσδιορίσει τη διάγνωση, την πρόγνωση και τη θεραπεία ασθενειών. Εγείρεται επομένως η ανάγκη διερεύνησης του κατά πόσο μπορούν οι εφαρμογές αυτές να είναι εξίσου ωφέλιμες στο σύνολο του πληθυσμού που εν δυνάμει θα μπορεί να τις χρησιμοποιήσει. Η αποφυγή των διακρίσεων (bias) σε αποφάσεις που αφορούν την ανθρώπινη υγεία κρίνεται απαραίτητη, ιδιαίτερα λόγω της σοβαρότητας των επιπτώσεων που μπορεί να έχουν [5], [8]. Έτσι, εκτός από την εξέταση της απόδοσης ενός μοντέλου, για την αξιολόγηση της αποτελεσματικότητάς του, εισάγεται η έννοια της Αμεροληψίας (Fairness), δηλαδή της ικανότητας ενός μοντέλου μηχανικής μάθησης να προβλέπει δίκαια, χωρίς να κάνει διακρίσεις που βασίζονται σε ευαίσθητα χαρακτηριστικά του πληθυσμού όπως το φύλο, η ηλικία, η φυλή, η γλώσσα ομιλίας, η οικονομική κατάσταση ή το ιατρικό ιστορικό ενός ατόμου [13]. Γίνεται λοιπόν αντιληπτή η ολοένα αυξανόμενη σημασία της αμεροληψίας ως κριτήριο στην επιλογή κατάλληλων μοντέλων σε ιατρικές εφαρμογές, καθώς προάγει την ηθική χρήση των τεχνολογιών και συγχρόνως ενισχύει την αξιοπιστία και την αποδοχή των συστημάτων από το κοινό. Επομένως, η διερεύνηση και ενίσχυση της αμεροληψίας αποτελεί σημαντική προτεραιότητα κατά την ενασχόλησή μας με την μηχανική μάθηση και ειδικότερα σε συσχέτιση με ιατρικές εφαρμογές.

1.2 Σκοπός της παρούσας διπλωματικής εργασίας

Το ζήτημα της αμεροληψίας αποτελεί σημαντική πρόκληση κατά την ανάπτυξη μοντέλων μηχανικής μάθησης και σχετίζεται με πολλούς παράγοντες όπως η επιλογή των ευαίσθητων χαρακτηριστικών που μελετάμε, η διαδικασία εκπαίδευσης των μοντέλων, αλλά και η ίδια η φύση των δεδομένων. Συχνά, υπάρχουσες ανισότητες ενδέχεται να ενισχύονται από τα μοντέλα μηχανικής μάθησης. Το γεγονός αυτό, μπορεί να έχει σοβαρές επιπτώσεις, ειδικότερα όταν πρόκειται για προβλήματα σε κρίσιμους τομείς όπως η υγεία.

Η παρούσα εργασία στοχεύει στην διερεύνηση του βαθμού αμεροληψίας μοντέλων μηχανικής μάθησης που αναπτύξαμε για την ανίχνευση της νόσου COVID-19 από ηχητικά σήματα βήχα και την αξιολόγηση διαφόρων μεθόδων ενίσχυσης της αμεροληψίας που εφαρμόσαμε στα μοντέλα. Πιο συγκεκριμένα, χρησιμοποιήσαμε έναν απλό δυαδικό ταξινομητή και ένα συνελκτικό νευρωνικό δίκτυο στα οποία εξετάσαμε πιθανές διακρίσεις σε πληθυσμιακές ομάδες με διαφορετικά ευαίσθητα χαρακτηριστικά, όπως το φύλο, η ηλικία, το ιατρικό ιστορικό και τα συμπτώματα, αλλά και συνδυαστικές κατηγορίες. Στη συνέχεια, εφαρμόσαμε και συγκρίναμε τέσσερις μεθόδους αντιμετώπισης των διακρίσεων στα μοντέλα μας, οι οποίες καλύπτουν το φάσμα των κατηγοριών εφαρμογής: πριν την επεξεργασία, κατά την επεξεργασία και μετά την επεξεργασία, επιτυγχάνοντας βελτίωση της αμεροληψίας σε αρκετές περιπτώσεις.

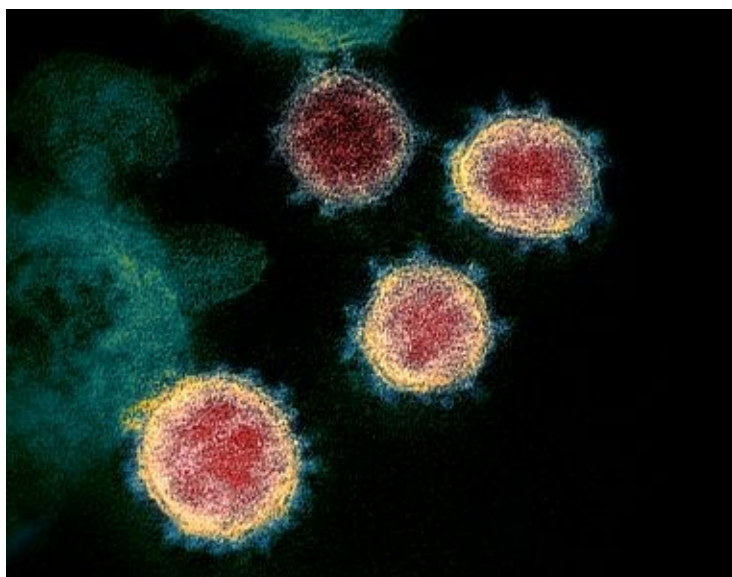
1.3 Δομή της διπλωματικής εργασίας

Η εργασία δομείται όπως παρουσιάζεται ακολούθως: Αρχικά, αναφέρονται οι απαραίτητες εισαγωγικές έννοιες για τη νόσο COVID-19, τη μηχανική μάθηση και την αμεροληψία στη μηχανική μάθηση. Ειδικότερα, παρουσιάζεται η σοβαρότητα της νόσου COVID-19 και η φύση της που επιτρέπει τη διάγνωση από ηχητικά σήματα και αναλύονται έννοιες της μηχανικής μάθησης, των νευρωνικών δικτύων και της επεξεργασίας ηχητικών δεδομένων. Αναλύεται επίσης η σημασία της αμεροληψίας στις ιατρικές εφαρμογές μηχανικής μάθησης και οι μέθοδοι ενίσχυσής της και αναφέρονται οι τρόποι αξιοποίησης της μηχανικής μάθησης στη διάγνωση της COVID-19. Στη συνέχεια, παρουσιάζεται το σύνολο δεδομένων που μελετήσαμε, η μεθοδολογία ανάπτυξης και εκπαίδευσης των μοντέλων που χρησιμοποιήσαμε και οι μέθοδοι ενίσχυσης της αμεροληψίας που χρησιμοποιήσαμε. Παρατίθενται, έπειτα, τα αποτελέσματα της αξιολόγησης των μοντέλων τόσο ως προς την επιτυχή ταξινόμηση αλλά και ως προς την αμεροληψία για διάφορες κατηγορίες ευαίσθητων χαρακτηριστικών πριν και μετά την εφαρμογή των μεθόδων αντιμετώπισης των

διακρίσεων. Τέλος, τα αποτελέσματα συνοψίζονται και σχολιάζονται με σκοπό την εξαγωγή των συμπερασμάτων που προκύπτουν από την σύγκρισή τους και τη διερεύνηση μελλοντικών προεκτάσεων του θέματος της εργασίας.

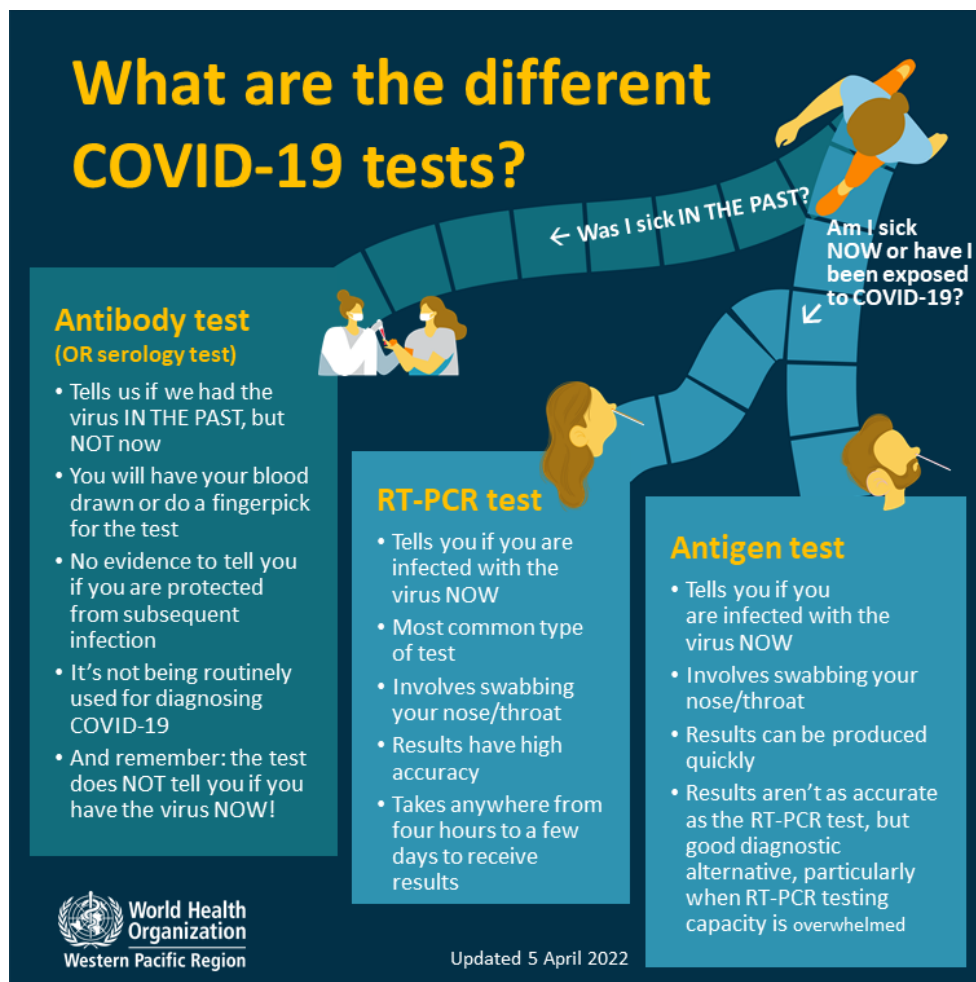
2 COVID-19

Το νεότερο στέλεχος του κορονοϊού SARS-CoV-2, εκδηλώθηκε ως μια νέα σοβαρή ασθένεια κυρίως του αναπνευστικού συστήματος, η οποία ονομάστηκε COVID-19 από τον Παγκόσμιο Οργανισμό Υγείας (ΠΟΥ) τον Φεβρουάριο του 2020 [14]. Ο ιός εμφανίστηκε πρώτη φορά στην πόλη Ουχάν της επαρχίας Χομπέι στην Κίνα τον Δεκέμβριο του 2019, όπου διαπιστώθηκε ότι αρκετά περιστατικά πνευμονίας άγνωστης προέλευσης, συνδέονταν με την τοπική αγορά θαλασσινών [15]. Η αποτυχία του αρχικού περιορισμού της νόσου οδήγησε σε ταχεία εξάπλωσή της και έτσι, στις 11 Μαρτίου 2020 ανακηρύχθηκε και επισήμως από τον ΠΟΥ ως πανδημία [2]. Δεν είναι ο πρώτος κορονοϊός, ο οποίος μολύνει τον άνθρωπο, καθώς οι παθογόνοι ιοί που προκαλούν ασθένειες στον ανθρώπινο οργανισμό (ανθρώπινοι κορονοϊοί, HCoV) περιλαμβάνουν 6 άλλα στελέχη που χαρακτηρίζονται ως SARS-CoV, αναπνευστικό σύνδρομο Μέσης Ανατολής (MERS)-CoV, HCoV-HKU1, HCoV-NL63, HCoV-OC43 και HCoV-229E [15]. Οι ανθρώπινοι κορονοϊοί ταυτοποιήθηκαν για πρώτη φορά στα μέσα της δεκαετίας του 1960 και ονομάστηκαν έτσι λόγω των αιχμών που διαθέτουν στην επιφάνειά τους και μοιάζουν με κορώνες. Ο ιός SARS-CoV-2 ο οποίος απεικονίζεται ακολούθως στην Εικόνα 2.1 ανήκει στους β-κορονοϊούς και τα κύρια κύτταρα-στόχοι του είναι τα επιθηλιακά κύτταρα του αναπνευστικού και του γαστρεντερικού συστήματος. Κλινικά και προκλινικά δεδομένα από μελέτες με άλλους κορονοϊούς υποδηλώνουν ευρύτερη διεισδυτικότητα στους ιστούς, γεγονός που μπορεί να οδηγήσει σε πιο σύνθετα κλινικά σενάρια. Έχει αποδειχθεί ότι η παρουσία των κορονοϊών στο ανθρώπινο σώμα, και ιδίως των β-κορονοϊών στους οποίους ανήκει ο SARS-CoV-2, δεν περιορίζεται στην αναπνευστική οδό, αλλά επεκτείνεται συχνά και στο Κεντρικό Νευρικό Σύστημα (ΚΝΣ) [16]. Η κλινική εικόνα εκδηλώνεται κυρίως ως κακοήθης πνευμονία, αν και πολλοί ασθενείς παρουσιάζουν νευρολογικά συμπτώματα. Μελέτες ανέφεραν ένα ευρύ φάσμα συμπτωμάτων που σχετίζονται με την COVID-19, όπως μη παραγωγικός βήχας, πυρετός, μυαλγία, κεφαλαλγία, κόπωση, απώλεια γεύσης ή/και όσφρησης, δύσπνοια, διάρροια και ναυτία ή έμετος, ενώ ορισμένοι ασθενείς παρουσιάζονται ασυμπτωματικοί [1].



Εικόνα 2.1: Εικόνα από ηλεκτρονικό μικροσκόπιο διέλευσης που δείχνει τον SARS-CoV-2, τον ιό που προκαλεί την COVID-19, απομονωμένο από ασθενή στις ΗΠΑ. Οι ακίδες στην εξωτερική επιφάνεια των σωματιδίων του ιού δίνουν στους κορονοϊούς το όνομά τους, λόγω της κορώνας που σχηματίζουν [15].

Σε σοβαρές περιπτώσεις, η νόσος μπορεί να οδηγήσει μέχρι και σε θάνατο, με ομάδες υψηλότερου κινδύνου να αποτελούν οι ηλικιωμένοι και τα άτομα με υποκείμενα νοσήματα. Ο χρόνος επώασης της νόσου, δηλαδή το χρονικό διάστημα που μεσολαβεί μεταξύ της μόλυνσης και της εκδήλωσης των συμπτωμάτων μπορεί να είναι από 2 έως 14 ημέρες, με τη μέση τιμή να είναι 5-6 ημέρες. Ένα άτομο που νοσεί θεωρείται περισσότερο μεταδοτικό μόλις εμφανίσει συμπτώματα, όμως δυνητικά μπορεί να μεταδίδει τη λοίμωξη μέχρι και 1-4 ημέρες πριν την εκδήλωση συμπτωμάτων, γεγονός που σε συνδυασμό με την ύπαρξη ασυμπτωματικών ασθενών, καθιστά τον ιό ιδιαίτερα μεταδοτικό. Ο μεγάλος βαθμός μεταδοτικότητας του ιού μπορεί επίσης να αποδοθεί και στις συνεχείς μεταλλάξεις του, οι οποίες καθιστούν δύσκολη τη ανάπτυξη αντισωμάτων. Για τους λόγους αυτούς, οι οδηγίες για τον περιορισμό της εξάπλωσης της COVID-19 την περίοδο της πανδημίας, περιλάμβαναν αυστηρά μέτρα κοινωνικής αποστασιοποίησης, υποχρεωτική χρήση μάσκας, καραντίνα για τους νοσούντες, περιορισμούς στις μετακινήσεις για όλο τον πληθυσμό, τηλεργασία και στα πρώτα στάδια ακόμη και ιχνηλάτηση των επαφών των κρουσμάτων [17]. Ο πολύ υψηλός ρυθμός μετάδοσης του SARS-CoV-2 οδήγησε την επιστημονική κοινότητα στην αναζήτηση μεθόδων έγκυρης και έγκαιρης διάγνωσης των κρουσμάτων. Στην προσπάθεια περιορισμού της εξάπλωσης της νόσου, χρησιμοποιήθηκαν δύο βασικές κατηγορίες ελέγχου: ο διαγνωστικός έλεγχος και ο έλεγχος αντισωμάτων, όπως καταδεικνύεται στην ενημερωτική αφίσα που απεικονίζεται στην Εικόνα 2.2. Στον πρώτο, στον οποίο ανιχνευόταν η παρουσία του ιού στον οργανισμό, μέσω της λήψης και ανάλυσης ρινοφαρυγγικού επιχρίσματος ή σιέλου, συγκαταλέγονταν: το μοριακό τεστ (PCR), το ταχύ τεστ αντιγόνου (rapid tests) και το τεστ αυτοδιάγνωσης (self-tests). Στον έλεγχο αντισωμάτων ανιχνευόταν, μέσω της λήψης δείγματος αίματος, η παρουσία αντισωμάτων που παρήχθησαν μετά από μόλυνση ή εμβολιασμό [18].



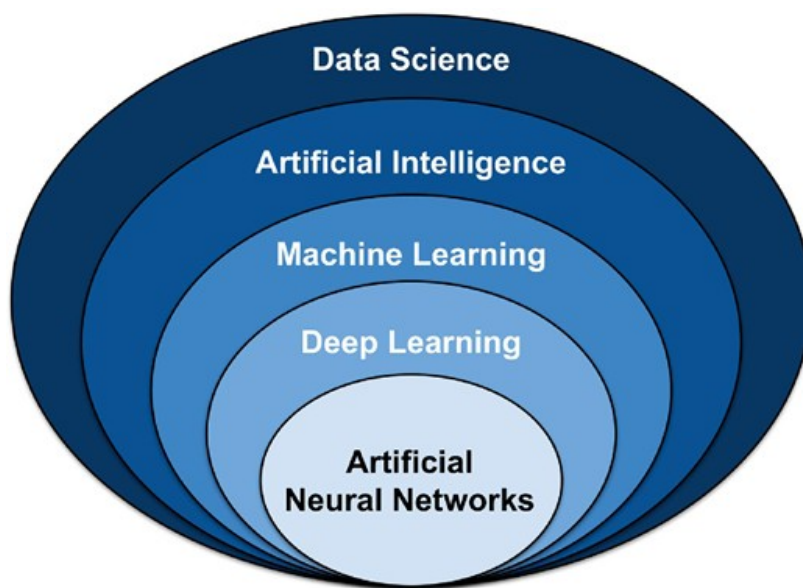
Εικόνα 2.2: Ενημερωτική αφίσα του ΠΟΥ την εποχή της πανδημίας με οδηγίες για τις μεθόδους διάγνωσης της COVID-19 [19].

Μέχρι και τον Ιανουάριο του 2025, η νόσος COVID-19 έχει στοιχίσει περισσότερες από 7,1 εκατομμύρια ανθρώπινες ζωές σε όλο τον πλανήτη και παρά τις σημαντικές προόδους στον τομέα της ιατρικής, ακόμα και σήμερα εξακολουθεί να είναι ιδιαίτερα ενεργή στον πληθυσμό. Το γεγονός αυτό, αποδεικνύουν και τα επίσημα στοιχεία του ΠΟΥ σύμφωνα με τα οποία, τις τέσσερις τελευταίες εβδομάδες του 2024 τα νέα κρούσματα ανήλθαν σε 161.000 παγκοσμίως [3], [20].

3 Μηχανική Μάθηση

3.1 Μηχανική Μάθηση και Βαθιά Μάθηση

Η Μηχανική Μάθηση (Machine Learning) αποτελεί κλάδο της Τεχνητής Νοημοσύνης (AI) που ανήκει στο ευρύτερο πεδίο της επιστήμης των υπολογιστών. Μπορεί να οριστεί ως μια διαδικασία ανάπτυξης υπολογιστικών συστημάτων με την ικανότητα να βελτιώνουν την απόδοσή τους κατά τη διάρκεια της εκτέλεσης, υλοποιώντας μια διαδικασία μάθησης χωρίς να χρειάζεται επαναπρογραμματισμός τους [21]. Στην παρούσα εργασία, εστιάζουμε στις μεθόδους Μηχανικής Μάθησης που ανήκουν στον τομέα της Βαθιάς Μάθησης και πιο συγκεκριμένα που αφορούν τα Τεχνητά Νευρωνικά Δίκτυα (ANN), όπως οπτικοποιείται και στην Εικόνα 3.1.1 .



Εικόνα 3.1.1: Σχηματική απεικόνιση της Επιστήμης των Δεδομένων με τα διαδοχικά υποσύνολά της: Τεχνητή Νοημοσύνη, Μηχανική Μάθηση, Βαθιά Μάθηση, Τεχνητά Νευρωνικά Δίκτυα [22].

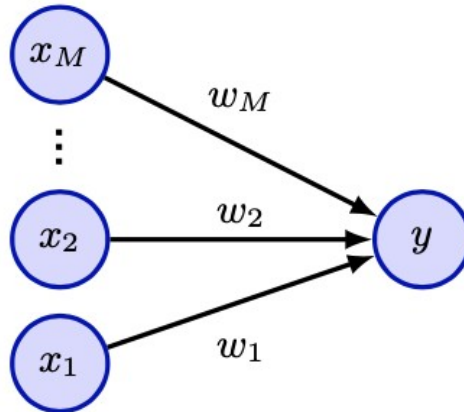
Τα Τεχνητά Νευρωνικά Δίκτυα αναπτύχθηκαν ως μια προσπάθεια προσομοίωσης της λειτουργίας του ανθρώπινου εγκεφάλου για την επεξεργασία πληροφορίας. Οι βασικές μονάδες επεξεργασίας στον εγκέφαλο είναι ηλεκτρικά ενεργά κύτταρα που ονομάζονται νευρώνες. Όταν ένας νευρώνας «πυροδοτείται», στέλνει ηλεκτρικό σήμα κατά μήκος του άξονά του το οποίο μέσω συνδέσεων που ονομάζονται συνάψεις, φτάνει σε άλλους νευρώνες. Αυτά τα ηλεκτρικά σήματα, τα οποία ονομάζονται νευροδιαβιβαστές μπορούν είτε να διεγείρουν είτε να αναστείλουν την «πυροδότηση» των άλλων νευρώνων, σε βαθμό που εξαρτάται από τη δύναμη κάθε σύναψης. Έτσι, και στα Τεχνητά Νευρωνικά Δίκτυα, οι βασικές μονάδες είναι οι τεχνητοί νευρώνες ή για ευκολία νευρώνες (neurons). Η απόκριση ενός νευρώνα i δίνεται μαθηματικά ως εξής:

$$a = \sum_{i=1}^M w_i x_i + b \quad (2.1)$$

$$y = f(a) \quad (2.2)$$

όπου τα $x_1, \dots, x_M$ αντιστοιχούν σε M εξόδους άλλων νευρώνων ή εξωτερικά σήματα που εισάγονται ως είσοδος, τα $w_1, \dots, w_M$ είναι συνεχείς μεταβλητές που ονομάζονται βάρη και εκφράζουν την

δύναμη κάθε σύναψης και η μεταβλητή b ονομάζεται πόλωση και χρησιμοποιείται για να αυξήσει ή να μειώσει το άθροισμα ανάλογα με το πρόσημό της. Επίσης, το μέγεθος a ονομάζεται προ-ενεργοποίηση, η μη γραμμική συνάρτηση f ονομάζεται συνάρτηση ενεργοποίησης και τέλος, η έξοδος y ονομάζεται ενεργοποίηση [23]. Η παραπάνω μαθηματική έκφραση μπορεί να απεικονιστεί διαγραμματικά όπως παρουσιάζεται ακολούθως στην Εικόνα 3.1.2.



Εικόνα 3.1.2: Διάγραμμα ενός απλού Τεχνητού Νευρωνικού Δικτύου που απεικονίζει τις μαθηματικές σχέσεις που περιγράφουν έναν απλό νευρώνα [23].

3.2 Μεθοδολογίες Εκπαίδευσης Μοντέλων Μηχανικής Μάθησης

Στον τομέα της Μηχανικής Μάθησης αναπτύσσονται αλγόριθμοι που μπορούν να διαχωριστούν ως προς τον τρόπο μάθησης σε τέσσερεις κατηγορίες: επιβλεπόμενης μάθησης (Supervised Learning), μη επιβλεπόμενης μάθησης (Unsupervised Learning), ημι-επιβλεπόμενης μάθησης (Semi-supervised Learning) και ενισχυτικής μάθησης (Reinforcement Learning).

- **Επιβλεπόμενη Μάθηση:** Επιβλεπόμενη μάθηση είναι η διαδικασία όπου ο αλγόριθμος εκπαιδεύεται στη δημιουργία μιας συνάρτησης με δοθείσες γνωστές επιθυμητές εξόδους που αντιστοιχίζονται στα δεδομένα εισόδου (σύνολο εκπαίδευσης), με απώτερο στόχο τη γενίκευση της συνάρτησης αυτής και για εισόδους με άγνωστη έξοδο. Χρησιμοποιείται συχνότερα σε προβλήματα ταξινόμησης (Classification), πρόγνωσης (Prediction) και διερμηνείας (Interpretation).
- **Μη Επιβλεπόμενη Μάθηση:** Κατά τη μη επιβλεπόμενη μάθηση, ο αλγόριθμος προσπαθεί να εντοπίσει συσχετισμούς και πρότυπα μεταξύ των δεδομένων εισόδου υπό μορφή παρατηρήσεων χωρίς να γνωρίζει τις επιθυμητές εξόδους. Ονομάζεται μη επιβλεπόμενη γιατί δεν χρειάζεται ανθρώπινη παρέμβαση με τη δημιουργία ετικετών για τα δεδομένα. Χρησιμοποιείται σε προβλήματα ανάλυσης συσχετισμών (Association Analysis) και ομαδοποίησης (Clustering).
- **Ημι-επιβλεπόμενη Μάθηση:** Η ημι-επιβλεπόμενη μάθηση, όπως προδίδει και η ονομασία της, προκύπτει από έναν συνδυασμό επιβλεπόμενης και μη επιβλεπόμενης μάθησης. Οι αλγόριθμοι της ημι-επιβλεπόμενης μάθησης χρησιμοποιούν για την εκπαίδευσή τους δεδομένα με γνωστές αλλά και με άγνωστες επιθυμητές εξόδους, τεχνική που αποδεικνύεται ιδιαίτερα χρήσιμη σε προβλήματα με πολύ μεγάλο όγκο δεδομένων εισόδου, που καθιστά δύσκολη και κοστοβόρα την πλήρη ετικετοποίησή τους. Ο λόγος για τον οποίο σε αυτή την τεχνική επιχειρείται η ανάμιξη των δύο προηγούμενων κατηγοριών είναι για να επιτευχθεί ακόμα καλύτερη απόδοση από αυτή που πετυχαίνουμε χρησιμοποιώντας μόνο τα δεδομένα

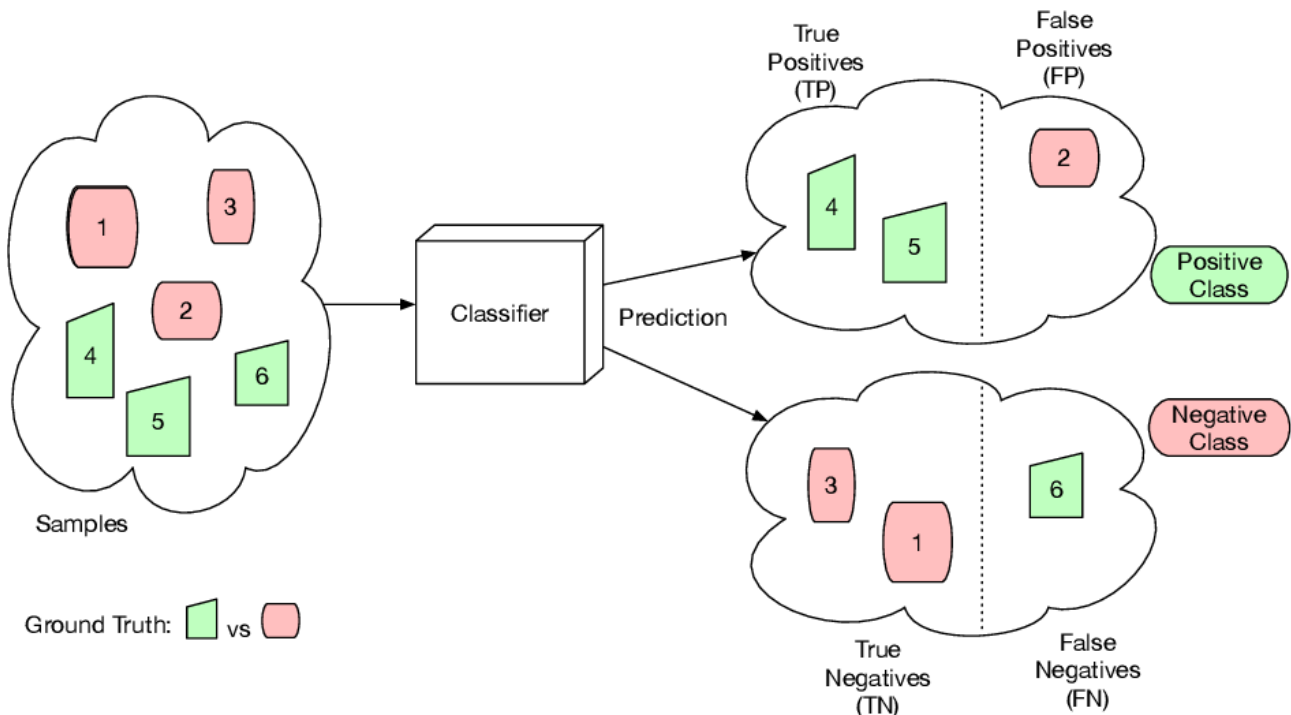
με ετικέτα. Μπορεί να χρησιμοποιηθεί επίσης σε προβλήματα ταξινόμησης (Classification) και ομαδοποίησης (Clustering).

- **Ενισχυτική Μάθηση:** Τέλος, κατά την ενισχυτική μάθηση, ο αλγόριθμος μαθαίνει μια στρατηγική ενεργειών μέσα από άμεση αλληλεπίδραση με το περιβάλλον, χωρίς να υπάρχει προκαθορισμένη έξοδος για κάθε είσοδο, αλλά με στόχο μία συνολική ανταμοιβή. Χρησιμοποιείται κυρίως σε προβλήματα σχεδιασμού (Planning), όπως για παράδειγμα ο έλεγχος κίνησης ρομπότ, αυτόνομα οχήματα, συστάσεις σε χρήστες εφαρμογών, βελτιστοποίηση εργασιών σε εργοστασιακούς χώρους [24] , [25].

3.3 Αλγόριθμοι Μηχανικής Μάθησης για Ταξινόμηση

Στο πεδίο της επιβλεπόμενης μηχανικής μάθησης που αναφέραμε προηγουμένως, μία βασική κατηγορία προβλημάτων που συχνά καλούμαστε να επιλύσουμε είναι αυτή των προβλημάτων ταξινόμησης. Οι αλγόριθμοι που χρησιμοποιούνται για αυτά τα προβλήματα, οι ταξινομητές (Classifiers), στοχεύουν στην κατάταξη κάθε εισόδου σε μία από ένα σύνολο διακριτών και προκαθορισμένων ετικετών ή κατηγοριών εξόδου. Τα κυριότερα είδη της ταξινόμησης διακρίνονται σε:

- **Διαδική Ταξινόμηση (Binary Classification):** Η είσοδος κατατάσσεται σε μία από δύο κατηγορίες (π.χ., “ναι” ή “όχι”, “θετικό” ή “αρνητικό”) όπως μπορούμε να παρατηρήσουμε στην Εικόνα 3.3.1.
- **Πολυκατηγορική Ταξινόμηση (Multiclass Classification):** Η είσοδος μπορεί να ανήκει σε μία από ένα σύνολο διαφορετικών κατηγοριών (π.χ., αναγνώριση χειρόγραφων ψηφίων από 0-9).
- **Πολυετικετική Ταξινόμηση (Multilabel Classification):** Η είσοδος μπορεί να ανήκει ταυτόχρονα σε παραπάνω από μία κατηγορίες (π.χ., κατάταξη μουσικών κομματιών σε περισσότερα από ένα είδη).



Εικόνα 3.3.1: Σχηματική απεικόνιση Δυναδικού Ταξινομητή, που κατατάσσει τα δειγματο (samples) σε δύο κατηγορίες θετική (positive) και αρνητική (negative) και στη συνέχεια αξιολογεί τις προβλέψεις του [26].

Για την επίλυση των απλούστερων προβλημάτων που ανήκουν στα είδη που αναφέρθηκαν παραπάνω, χρησιμοποιούνται συνηθέστερα κάποιοι από τους ακόλουθους βασικούς απλούς ταξινομητές (Simple Classifiers) [27], [28]:

- **Δέντρο Απόφασης (Decision Tree):** Είναι ταξινομητής βασιζόμενος σε απόφαση (Decision-Based) και αποτελείται από ένα σύνολο κανόνων απόφασης που βασίζονται στα χαρακτηριστικά των δεδομένων. Συχνά χρησιμοποιείται και ο αλγόριθμος για Τυχαίο Δάσος (Random Forest) που αποτελείται από ένα σύνολο Δέντρων Απόφασης.
- **Αφελής Μπεϋζιανός Ταξινομητής (Naive Bayesian Classifier):** Είναι ταξινομητής βασιζόμενος σε πιθανότητες (Probability-Based) και υποθέτει ισχυρή ανεξαρτησία μεταξύ των δεδομένων.
- **Μηχανή Διανυσμάτων Υποστήριξης (Support Vector Machine):** Είναι Γραμμικός Ταξινομητής (Linear Classifier) που η λειτουργία του βασίζεται στην εύρεση του βέλτιστου ορίου που διαχωρίζει τις κατηγορίες.
- **K – Πλησιέστερος Γείτονας (K-Nearest Neighbor):** Είναι Ταξινομητής Πλησιέστερου Γείτονα (Nearest Neighbor Classifier) που αποθηκεύει τα δεδομένα εκπαίδευσης και τα συγκρίνει με τα νέα δεδομένα που πρέπει να ταξινομήσει, χρησιμοποιώντας ένα μέτρο ομοιότητας.

3.4 Τεχνητά Νευρωνικά Δίκτυα

Όπως προαναφέρθηκε, τα Τεχνητά Νευρωνικά Δίκτυα είναι μοντέλα που προσπαθούν να προσομοιώσουν τη λειτουργία των νευρώνων του ανθρώπινου εγκεφάλου και αποτελούν βασικό πυλώνα της Βαθιάς Μάθησης. Η κεντρική ιδέα είναι ότι κάθε κόμβος-νευρώνας μπορεί να δεχθεί μια είσοδο, να αποθηκεύσει κάποια πληροφορία σχετικά με αυτήν και στη συνέχεια να μεταφέρει την πληροφορία στο επόμενο επίπεδο. Έτσι, σε κάθε επίπεδο γίνεται και πιο σύνθετη η πληροφορία που λαμβάνει και αντίστοιχα μεταδίδει, συγκριτικά με το προηγούμενο επίπεδο.

Ο διαχωρισμός των επιπέδων στη βασική δομή τους περιλαμβάνει τα:

- **Επίπεδο Εισόδου (Input layer):** Πρόκειται για το πρώτο επίπεδο σε ένα νευρωνικό δίκτυο στο οποίο εισάγονται τα αρχικά δεδομένα.
- **Κρυμμένα Επίπεδα (Hidden layers):** Τα κρυμμένα επίπεδα αποτελούν το ενδιάμεσο τμήμα ανάμεσα στο επίπεδο εισόδου και εξόδου και επιτελούν το μεγαλύτερο μέρος του υπολογισμού.
- **Επίπεδο Εξόδου (Output layer):** Το επίπεδο εξόδου είναι το τελικό επίπεδο σε ένα νευρωνικό δίκτυο και οδηγεί στην έξοδο του δικτύου. Στη περίπτωση που υπάρχει ένα πρόβλημα δυαδικής ταξινόμησης, θα έχει μόνο ένα νευρώνα στο επίπεδο εξόδου, ενώ στην περίπτωση που υπάρχει ένα πρόβλημα ταξινόμησης πολλαπλών κλάσεων, μπορεί να έχει πολλαπλούς νευρώνες στο επίπεδο εξόδου.

Τα τεχνητά νευρωνικά δίκτυα αποτελούνται από μεγάλο αριθμό κρυμμένων επιπέδων τα οποία αυξάνουν την πολυπλοκότητα, καθιστώντας έτσι δυνατή τη διαχείριση σημαντικά μεγάλου όγκου δεδομένων. Για τον λόγο αυτό, έχουν εφαρμογές σε ένα ευρύ φάσμα πεδίων, μεταξύ των οποίων τα

βασικότερα είναι: η αναγνώριση εικόνας και ήχου, η επεξεργασία φυσικής γλώσσας και η ρομποτική. Τα τεχνητά νευρωνικά δίκτυα έχουν την ικανότητα να αυτοματοποιούν τη διαδικασία εξαγωγής χαρακτηριστικών από τα δεδομένα, γεγονός που τα διαφοροποιεί από τα παραδοσιακά μοντέλα μηχανικής μάθησης, για τη χρήση των οποίων χρειάζεται να έχει προηγηθεί εξαγωγή χαρακτηριστικών των δεδομένων από τον άνθρωπο [29].

Προκειμένου να επιτύχουμε καλύτερα αποτελέσματα και να αποφύγουμε την υπερπροσαρμογή των μοντέλων βαθιάς μάθησης, συχνά χρησιμοποιούνται κάποια επιπλέον επίπεδα στα οποία εφαρμόζονται τεχνικές όπως [30]:

- **Απενεργοποίηση Νευρώνων (Dropout):** Σε κάθε εποχή της εκπαίδευσης απενεργοποιείται ένα ποσοστό νευρώνων τυχαία. Έτσι, το μοντέλο εξαναγκάζεται σε εκμάθηση διαφορετικών ανεξάρτητων χαρακτηριστικών, αφού η επιλογή τους ισοκατανέμεται μεταξύ όλων των νευρώνων.
- **Απενεργοποίηση Βαρών (Drop-Weights):** Αποτελεί τεχνική ανάλογη της απενεργοποίησης νευρώνων, αλλά αντί για νευρώνες σε κάθε εποχή απενεργοποιούνται οι συνδέσεις τους, δηλαδή τα βάρη.
- **Επαύξηση Δεδομένων (Data Augmentation):** Με τεχνητό τρόπο αυξάνουμε τον όγκο του συνόλου των δεδομένων εκπαίδευσης, ώστε το μοντέλο να εκπαιδευτεί σε φαινομενικά περισσότερα δεδομένα. Αυτό επιτυγχάνεται με διάφορους τρόπους επεξεργασίας όπως: περιστροφή, περικοπή, εισαγωγή θορύβου, κ.ά.
- **Κανονικοποίηση Παρτίδας (Batch Normalization):** Κανονικοποιούμε την έξοδο κάθε επιπέδου, αφαιρώντας τον μέσο όρο και διαιρώντας με την τυπική απόκλιση. Η τεχνική αυτή ομαλοποιεί τα δεδομένα σε κάθε επίπεδο, ώστε να επιταχύνει την εκπαίδευση και να σταθεροποιήσει το δίκτυο.

Η απόκτηση γνώσης για τα Τεχνητά Νευρωνικά Δίκτυα γίνεται μέσω της εκπαίδευσής τους. Ο ευρύτερα χρησιμοποιούμενος αλγόριθμος εκπαίδευσης πολυεπίπεδων Τεχνητών Νευρωνικών Δικτύων είναι ο αλγόριθμος της οπισθοδιάδοσης σφαλμάτων (back-propagation) [31] ο οποίος βασίζεται στη διάδοση πληροφορίας από την έξοδο στην είσοδο του δικτύου με σκοπό την ελαχιστοποίηση του σφάλματος μεταξύ πραγματικής και επιθυμητής εξόδου. Τα βασικά βήματα του αλγορίθμου για ΤΝΔ ενός ενδιάμεσου επιπέδου παρουσιάζονται ακολούθως [32] :

- Αρχικοποίηση των βαρών w και των πλώσεων b σε τυχαίες, μικρές τιμές και εισαγωγή του p διανύσματος εισόδου $x_{p1}, x_{p2}, \dots, x_{pN-1}$ στο επίπεδο εισόδου και της επιθυμητής εξόδου $d_1, d_2, \dots, d_{M-1}$.
- Υπολογισμός εξόδου κάθε νευρώνα του ενδιάμεσου επιπέδου μέσω της συνάρτησης ενεργοποίησης f ως εξής:

$$y_j = \sum_i w_{ji} x_{pi} + b_{pj} \quad (2.3)$$

$$O_{pj} = f(y_{pj}) \quad (2.4)$$

όπου y_{pj} είναι το άθροισμα των εισόδων του ΤΝΔ πολλαπλασιασμένων με το συντελεστή βαρύτητας w_{ji} από το νευρώνα i του προηγούμενου επιπέδου στο νευρώνα j και μετά την εφαρμογή της πόλωσης και O_{pj} είναι η έξοδος του νευρώνα j μετά την εφαρμογή της συνάρτησης ενεργοποίησης.

- Υπολογισμός πραγματικών τιμών εξόδου και συνολικού σφάλματος σύμφωνα με τις σχέσεις:

$$E_p = \frac{1}{2} \sum_{k=1}^M (d_{pk} - O_{pk})^2 \quad (2.5)$$

$$E = \sum_p E_p \quad (2.6)$$

όπου d_{pk} είναι η επιθυμητή έξοδος στον νευρώνα k για είσοδο p , M ο αριθμός των νευρώνων εξόδου και E το συνολικό σφάλμα.

- Προσαρμογή βαρών αναδρομικά, χρησιμοποιώντας τη σχέση:

$$w_{pji}(n+1) = w_{pji}(n) + \eta \delta_{pj}(n) O_{pi}(n) \quad (2.7)$$

όπου $w_{pji}(n)$ είναι το βάρος ανάμεσα στον νευρώνα j και τον νευρώνα i του προηγούμενου επιπέδου, O_{pi} η έξοδος του νευρώνα i , η είναι ο ρυθμός εκπαίδευσης (learning rate) και δ_{pj} είναι ο όρος σφάλματος για τον νευρώνα j .

- Αν ο j είναι νευρώνας εξόδου, τότε:

$$\delta_{pj} = (d_{pj} - O_{pj}) f'(y_{pj}) \Rightarrow \delta_{pj} = (d_{pj} - O_{pj}) O_{pj} (1 - O_{pj}) \quad (2.8)$$

όπου d_{pj} η επιθυμητή και O_{pj} η πραγματική έξοδος και f' η παράγωγος της συνάρτησης ενεργοποίησης.

- Αν ο j είναι νευρώνας ενδιάμεσου επιπέδου, τότε:

$$\delta_{pj} f'_{pj} \sum_k \delta_{pk} w_{kj} \Rightarrow \delta_{pj} = O_{pj} (1 - O_{pj}) \sum_k \delta_{pk} w_{kj} \quad (2.9)$$

- Οι πολώσεις των κόμβων των εσωτερικών επιπέδων προσαρμόζονται με αντίστοιχο τρόπο, υποθέτοντας ότι αποτελούν βάρη σε βοηθητικές εισόδους σταθερής τιμής.

- Κατά τον υπολογισμό των βαρών, προστίθεται ένας όρος α που ονομάζεται ορμή έτσι ώστε ο αλγόριθμος να συγκλίνει πιο γρήγορα:

$$w_{pji}(n+1) = w_{pji}(n) + \eta \delta_{pj} O_{pi} + \alpha (w_{pji}(n) - w_{pji}(n-1)), \quad 0 < \alpha < 1 \quad (2.10)$$

- Επανάληψη της διαδικασίας από το δεύτερο βήμα. Κάθε επανάληψη της προσαρμογής των συντελεστών βαρών ονομάζεται εποχή (epoch). Το πλήθος των εποχών εκπαίδευσης ενός ΤΝΔ εξαρτάται από το κριτήριο τερματισμού που επιλέγεται όπως η μικρή μεταβολή της συνάρτησης σφάλματος, η επίτευξη τιμής συνάρτησης σφάλματος μικρότερης από προκαθορισμένη τιμή, ή η συμπλήρωση προκαθορισμένου αριθμού επαναλήψεων.

Τελικός στόχος της εκπαίδευσης ενός τεχνητού νευρωνικού δικτύου είναι οι, κατά το δυνατόν, ορθότερες προβλέψεις, γεγονός που επιτυγχάνουμε με τη μείωση της συνάρτησης απώλειας (loss function). Η συνάρτηση απώλειας εκφράζει τη διαφορά μεταξύ της πραγματικής εξόδου (ground truth) και της πρόβλεψης του μοντέλου, παρέχοντας ένα αριθμητικό μέτρο του σφάλματος που χρησιμοποιείται για την βελτιστοποίηση των παραμέτρων στην εκπαίδευση (π.χ. βάρη και διακυμάνσεις). Η συνάρτηση απώλειας μαθηματικά εκφράζεται ως:

$$loss = |(y_{pred} - y_{actual})| \quad (2.11)$$

Ειδικότερα στα προβλήματα δυαδικής ταξινόμησης, χρησιμοποιείται συνηθέστερα η δυαδική συνάρτηση απώλειας διασταυρούμενης εντροπίας (Binary Cross Entropy Loss -BCE), η οποία συγκρίνει τις πιθανότητες εξόδου που προβλέπει το μοντέλο με την πραγματική κλάση της εξόδου που μπορεί να είναι 0 ή 1 και στη συνέχεια “τιμωρεί” τις πιθανότητες ανάλογα με την απόκλισή τους από την αναμενόμενη τιμή. Επομένως:

$$BCE = -\frac{1}{N} \sum_{i=1}^N y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \quad (2.12)$$

όπου N είναι ο αριθμός των δειγμάτων, y_i είναι οι πραγματικές ετικέτες και p_i είναι οι προβλεπόμενες.

Αντίστοιχα, σε προβλήματα πολυκατηγορικής ταξινόμησης, συνηθέστερα χρησιμοποιείται η κατηγορική συνάρτηση απώλειας διασταυρούμενης εντροπίας (Categorical Cross Entropy Loss -CCE), η οποία εκφράζεται ως:

$$CCE = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{ij} \log(p_{ij}) \quad (2.13)$$

όπου N είναι ο αριθμός των δειγμάτων, C είναι ο αριθμός των κατηγοριών, y_{ij} είναι η πραγματική ετικέτα για το δείγμα i και την κατηγορία j και p_{ij} είναι η προβλεπόμενη πιθανότητα το δείγμα i να ανήκει στην κατηγορία j [33].

3.5 Συνελικτικά Νευρωνικά Δίκτυα

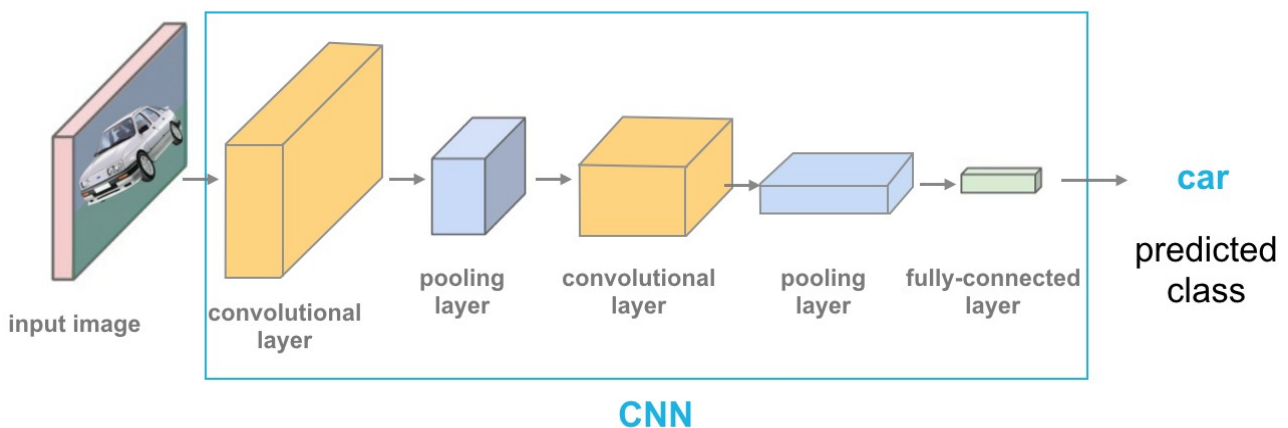
3.5.1 Δομή και επίπεδα

Τα Συνελικτικά Νευρωνικά Δίκτυα (CNNs) αποτελούν έναν από τους δημοφιλέστερους τύπους αρχιτεκτονικών που περιλαμβάνονται στο γενικό πλαίσιο των τεχνητών νευρωνικών δικτύων εξειδικεύονται στην επεξεργασία δεδομένων με χωρική δομή, όπως οι εικόνες και αποτελούνται από διαφορετικά είδη επιπέδων που αναπαρίστανται στην Εικόνα 3.5.1 και αναλύονται ακολούθως [34], [30]:

- **Επίπεδα Συνέλιξης (Convolutional Layers):** Πήραν την ονομασία τους λόγω της συνέλιξης που εφαρμόζεται ως βασική πράξη ανάμεσα στην συνάρτηση των τιμών εισόδου και στη συνάρτηση που ορίζεται ως ένας πυρήνας /φίλτρο (kernel/filter). Ο πυρήνας, είναι ένας πίνακας αριθμών που χρησιμοποιείται για την ανίχνευση χαρακτηριστικών σε μία εικόνα. Κάθε συνελικτικό επίπεδο περιέχει πολλαπλούς πυρήνες, καθένας από τους οποίους εκπαιδεύεται να εντοπίζει διαφορετικά χαρακτηριστικά. Με αυτό τον τρόπο ένα στοιχείο που εμφανίζεται πολλές φορές, μεταβάλλει τα βάρη του δικτύου με τέτοιο τρόπο, ώστε να το αναγνωρίζουν και τελικά να γίνεται η ανίχνευσή του, η οποία και αποτελεί ένα από τα μεγαλύτερα πλεονεκτήματα των συνελικτικών δικτύων (Feature Detection). Ως αποτέλεσμα, προκύπτει ένας χάρτης χαρακτηριστικών (feature map) που περιλαμβάνει πληροφορίες για συγκεκριμένα μοτίβα της εικόνας-εισόδου.
- **Επίπεδα Ενεργοποίησης (Activation Layers):** Σε αυτό το επίπεδο εφαρμόζεται μια μαθηματική συνάρτηση στην έξοδο κάθε νευρώνα, που καθορίζει αν θα μεταβιβάσει πληροφορία στο επόμενο επίπεδο. Το επίπεδο αυτό δεν μεταβάλλει την δομή του δικτύου, απλώς εισάγει μη-γραμμικότητα στο μοντέλο που επιτρέπει την εξαγωγή πιο περίπλοκων μοτίβων ανάμεσα στα δεδομένα. Οι συνηθέστερες συναρτήσεις που χρησιμοποιούνται για τα συνελικτικά δίκτυα είναι η Rectified Linear Unit (ReLU), που διατηρεί τις θετικές τιμές και θέτει τις αρνητικές ως 0 [$\max(0, z)$], και η Σιγμοειδής (Sigmoid) που περιορίζει την έξοδο μεταξύ 0 και 1 και είναι κατάλληλη για προβλήματα δυαδικής ταξινόμησης.
- **Επίπεδα Συγκέντρωσης (Pooling Layers):** Τα επίπεδα συγκέντρωσης εισάγονται πολλές φορές ανάμεσα σε διαδοχικά συνελικτικά επίπεδα. Σκοπός τους είναι η μείωση του μεγέθους του δικτύου μέσω της μείωσης των χωρικών διαστάσεων των χαρακτηριστικών, δηλαδή το πλάτος και το ύψος, χωρίς να μεταβάλλεται το βάθος. Συνηθίζεται να

εφαρμόζεται συγκέντρωση με μέγεθος $F=2$ και βήμα $S=2$. Το αποτέλεσμα του είναι ο υποτετραπλασιασμός του δικτύου, εφόσον μειώνεται στο μισό η κάθε πλευρά. Η επικρατέστερη μέθοδος είναι η μέθοδος του Μεγίστου (MaxPool), κατά την οποία επιλέγεται η μέγιστη τιμή από κάθε καθορισμένο παράθυρο της εικόνας και επικράτησε λόγω των καλύτερων αποτελεσμάτων της σε σχέση με την μέθοδο του Μέσου (AveragePool), κατά την οποία υπολογίζεται ο μέσος όρος των τιμών κάθε καθορισμένου παράθυρου της εικόνας.

- **Πλήρως Συνδεδεμένα Επίπεδα (Fully Connected Layers):** Τα πλήρως συνδεδεμένα επίπεδα τοποθετούνται προς την έξοδο του νευρωνικού δικτύου, αφού έχει προηγηθεί η μετατροπή των τριών διαστάσεων του δικτύου σε μία, με τη διαδικασία της επιπεδοποίησης (flattening). Η επιπεδοποίηση είναι απαραίτητη καθώς στα πλήρως συνδεδεμένα επίπεδα κάθε νευρώνας συνδέεται με όλους τους νευρώνες του προηγούμενου επιπέδου, κάτι που θα ήταν απαγορευτικό στις τρεις διαστάσεις. Από τα πλήρως συνδεδεμένα επίπεδα καταλήγουμε στη έξοδο, η οποία για ταξινομητικά προβλήματα λαμβάνει τη μορφή ενός διανύσματος πιθανότητας για κάθε κατηγορία.



Εικόνα 3.5.1: Σχηματική αναπαράσταση της βασικής αρχιτεκτονικής ενός Συνελικτικού Νευρωνικού Δικτύου [35].

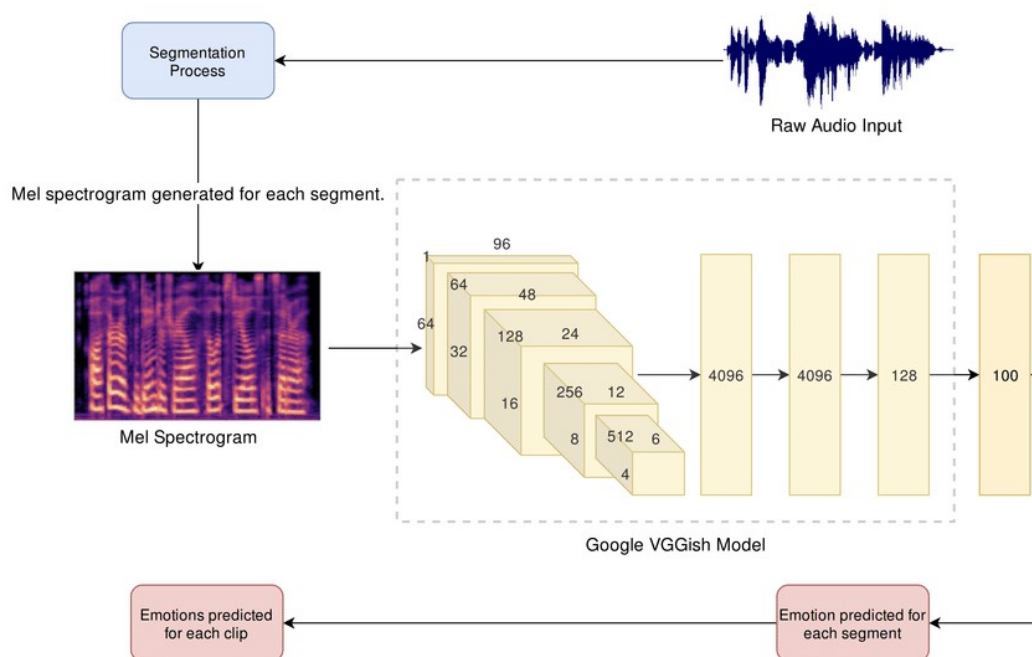
3.5.2 Το Συνελικτικό Νευρωνικό Δίκτυο VGG

Ένα από τα συνηθέστερα χρησιμοποιούμενα μοντέλα βαθιάς μάθησης στην επεξεργασία εικόνων λόγω της απλότητας και της αποτελεσματικότητάς του, είναι το VGG. Το μοντέλο VGG είναι ένα βαθύ συνελικτικό νευρωνικό δίκτυο που προτάθηκε για πρώτη φορά το 2013 από τους Andrew Zisserman και Karen Simonyan [36] για τον διαγωνισμό ImageNet Challenge, που είναι ένας από τους σημαντικότερους διαγωνισμούς στην επιστήμη της υπολογιστικής όρασης και της μηχανικής μάθησης. Το δίκτυο λαμβάνει ως είσοδο εικόνες προτύπου χρώματος RGB διαστάσεων 224×224 . Η βασική αρχιτεκτονική του στηρίζεται στην εφαρμογή πολλών συνελικτικών επιπέδων, στα οποία εφαρμόζονται με βήμα 1 εικονοστοιχείο φίλτρα 3×3 , που είναι οι ελάχιστες διαστάσεις για ανίχνευση κινήσεων σε όλες τις κατευθύνσεις και το καθιστούν πιο εύκολο στη διαχείριση. Χρησιμοποιούνται επίσης 1×1 συνελίξεις ως γραμμικός μετασχηματισμός της εισόδου. Στα επίπεδα ενεργοποίησης εφαρμόζεται μη γραμμική ενεργοποίηση, με τη συνάρτηση Rectified Linear Unit (ReLU). Μετά από ορισμένες ομάδες συνελικτικών επιπέδων, ακολουθεί ένα επίπεδο

συγκέντρωσης, στο οποίο πραγματοποιείται χωρική υποδειγματοληψία με βήμα 1 εικονοστοιχείο σε παράθυρο 2x2. Η διαδικασία αυτή μειώνει τον αριθμό των παραμέτρων των χαρτών χαρακτηριστικών που δημιουργούνται από κάθε βήμα συνέλιξης. Μετά την τελευταία ομάδα συνελκτικών επιπέδων, ακολουθούν τρία πλήρως συνδεδεμένα επίπεδα. Τα δύο πρώτα αποτελούνται από 4096 νευρώνες έκαστο, ενώ το τρίτο από 1000 νευρώνες, έναν για κάθε κλάση. Το τελευταίο πλήρως συνδεδεμένο επίπεδο χρησιμοποιεί συνάρτηση ενεργοποίησης softmax, κατάλληλη για πολυκατηγορική ταξινόμηση.

3.5.3 Το Συνελκτικό Νευρωνικό Δίκτυο VGGish

Παραλλαγή του VGG αποτελεί το μοντέλο VGGish [37], το οποίο χρησιμοποιείται για την επεξεργασία ηχητικών σημάτων, είτε ως εξαγωγέας χαρακτηριστικών (feature extractor) είτε ως αρχικοποίηση (warm start) για τα χαμηλότερα επίπεδα ενός μεγαλύτερου μοντέλου επεξεργασίας ήχου. Το μοντέλο VGGish περιλαμβάνει 11 επίπεδα με εκπαιδευσιμα βάρη, 8 συνελκτικά και 3 πλήρως συνδεδεμένα, όπως απεικονίζεται σχηματικά και στην Εικόνα 3.5.2. Οι κύριες διαφοροποιήσεις από το αρχικό μοντέλο VGG είναι τρεις: το μέγεθος της εισόδου, που ορίζεται σε διαστάσεις 96x64 ώστε να προσαρμοστεί στις διαστάσεις των Mel-φασματογραφημάτων που εισάγονται, η αφαίρεση τις τελευταίας ομάδας συνελκτικών επιπέδων και του τελευταίου επιπέδου συγκέντρωσης, και τέλος, η αντικατάσταση του τελευταίου πλήρως συνδεδεμένου επιπέδου 1000 νευρώνων από ένα επίπεδο που αντί για απευθείας ταξινόμηση, παράγει για κάθε είσοδο μαθηματικές αναπαραστάσεις (embeddings) 128 διαστάσεων.



Εικόνα 3.5.2: Σχηματική απεικόνιση της χρήσης του μοντέλου VGGish για αναγνώριση συναισθημάτων από ηχητικά σήματα.

3.6 Εξαγωγή Χαρακτηριστικών από Ηχητικά Δεδομένα

Για την εισαγωγή ηχητικών δεδομένων στα μοντέλα μηχανικής μάθησης, απαιτείται προηγουμένως η εξαγωγή χαρακτηριστικών των ηχητικών σημάτων, με τεχνικές όπως οι μετασχηματισμοί Fourier και αριθμητικές αναπαραστάσεις μεγεθών που σχετίζονται με το χρόνο ή τη συχνότητα του σήματος. Με τον τρόπο αυτό, μειώνεται ο όγκος των δεδομένων, βελτιώνεται η ακρίβεια των αλγόριθμων και επιταχύνεται η εκπαίδευση των μοντέλων. Για την εισαγωγή ηχητικών σημάτων σε

απλούς ταξινομητές, διενεργείται χειροκίνητη εξαγωγή των χαρακτηριστικών τους, εν αντιθέσει με τις περιπτώσεις χρήσης συνελκτικών νευρωνικών δικτύων στα οποία η εξαγωγή χαρακτηριστικών γίνεται απευθείας από το ίδιο το μοντέλο αφού πρώτα τα δεδομένα ήχου μετατραπούν σε φασματογραφήματα.

3.6.1 Χαρακτηριστικά εισόδου σε Απλούς Ταξινομητές

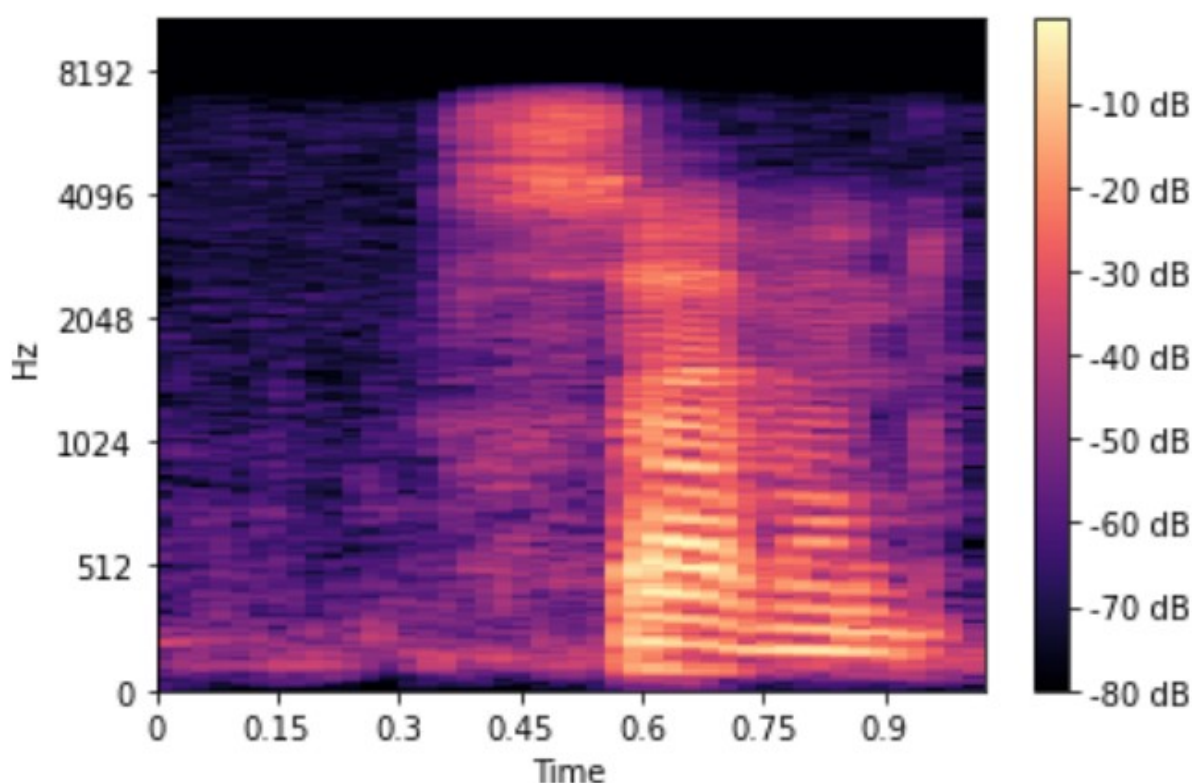
Τα χαρακτηριστικά Επεξεργασίας Ψηφιακού Σήματος (DSP) που συνήθως εξάγουμε από τα δεδομένα ήχου και χρησιμοποιούμε για την εκπαίδευση απλών ταξινομητών, είναι τα ακόλουθα:

- **Ανίχνευση Αιχμών Ενεργειακού Φακέλου (Energy Envelope Peak Detection - EEDP):** Ανιχνεύονται αιχμές στην ενέργεια του σήματος και έτσι εντοπίζει απότομες αυξομειώσεις της έντασης όπως κατά τον βήχα.
- **Ρυθμός Διαβάσεων από το Μηδέν (Zero Crossing Rate - ZCR):** Υπολογίζεται πόσες φορές το σήμα διέρχεται από τον άξονα x, δηλαδή πόσες φορές αλλάζει πρόσημο. Όσο υψηλότερος είναι ο ρυθμός αυτός τόσο πιο τραχύ είναι το σήμα.
- **Μέση Ισχύς Τετραγωνικής Ρίζας (Root Mean Square Power - RMSP):** Υπολογίζεται τη μέση τετραγωνική ισχύ του σήματος, ως ένδειξη του πόση ένταση έχει κατά μέσο όρο ο ήχος.
- **Δυναμικά Χαρακτηριστικά (Dynamic Features - DF):** Ανιχνεύονται δυναμικές μεταβολές στο σήμα, δηλαδή αλλαγές στην ένταση, τη συχνότητα ή τη φάση με την πάροδο του χρόνου.
- **Φασματικά Χαρακτηριστικά (Spectral Features):** Ανιχνεύονται χαρακτηριστικά του φάσματος του ήχου όπως το φασματικό εύρος.
- **Φασματική Επιπεδότητα και Βραχύχρονη Φασματική Απόκλιση (Spectral Flatness and Short-Time Spectral Deviation - SF_SSTD):** Ανιχνεύεται κατά πόσο η ενέργεια του σήματος κατανέμεται ομοιόμορφα ή παρουσιάζει διακυμάνσεις σε κατετημημένα χρονικά διαστήματα. Όσο υψηλότερη είναι η τιμή, έχουμε τόσο πιο επίπεδο σήμα που πλησιάζει τον λευκό θόρυβο [38].
- **Βραχύχρονη Φασματική Ένταση και Φασματική Πυκνότητα (Short-Time Spectral Loudness & Spectral Density - SSL_SD):** Υπολογίζεται πόσο δυνατός είναι ο ήχος και η κατανομή της πυκνότητας της ενέργειας σε κατετημημένα χρονικά διαστήματα.
- **Συντελεστές Λογαριθμικού Φάσματος Κλίμακας Mel (Mel-Frequency Cepstral Coefficients - MFCC):** Εφαρμόζεται μετασχηματισμός Fourier (FT) στο αρχικό χρονικό σήμα για να προκύψει το φάσμα, λογαριθμίζοντας υπολογίζεται το λογαριθμικό φάσμα του ηχητικού σήματος, εφαρμόζεται ένα φίλτρο Mel ώστε οι συχνότητες να προσεγγίζουν την ανθρώπινη ακοή και τέλος εφαρμόζεται ένας διακριτός συνημιτονικός μετασχηματισμός (DCT). Τα χαρακτηριστικά αυτά είναι ιδανικά για την αναγνώριση και ταξινόμηση του είδους το ήχου (π.χ., ομιλίας, βήχα, αναπνοής) και αποτελούν από τα σημαντικότερα χαρακτηριστικά στην επεξεργασία ηχητικών σημάτων [39].
- **Συχνότητα Κέντρου Βάρους (Centroid Frequency - CF):** Υπολογίζεται το κέντρο βάρους των συχνοτήτων ενός ηχητικού σήματος, δηλαδή μία μέση τιμή συχνότητας του ήχου. Όσο πιο κοντά είναι μία συχνότητα του σήματος στο κέντρο του τόσο μεγαλύτερη συμβολή έχει στο σήμα [40].
- **Μήκος Σήματος (Length - LGTH):** Υπολογίζεται η συνολική διάρκεια του ηχητικού σήματος.

- **Πυκνότητα Φασματικής Ισχύος (Power Spectral Density - PSD):** Ανιχνεύεται η κατανομή της ισχύος για τις διαφορετικές τιμές συχνοτήτων στο φάσμα του σήματος.

3.6.2 Χαρακτηριστικά Εισόδου σε Συνελκτικά Νευρωνικά Δίκτυα

Για να χρησιμοποιήσουμε τα δεδομένα ήχου ως είσοδο σε ένα συνελκτικό νευρωνικό δίκτυο, πρέπει προηγουμένως να τα μετατρέψουμε σε εικόνα. Οι απεικονίσεις των ηχητικών σημάτων και πιο συγκεκριμένα, οι οπτικοποιήσεις των συχνοτήτων του ήχου σε συνάρτηση με τον χρόνο είναι τα φασματογραφήματα (spectrograms). Για τη δημιουργία των φασματογραφημάτων εφαρμόζεται ο μετασχηματισμός Fourier μικρού χρόνου (STFT) που μετατρέπει το σήμα από τον χρόνο στο πεδίο της συχνότητας. Υποκατηγορία των φασματογραφημάτων αποτελούν τα Mel-φασματογραφήματα, που χρησιμοποιούν την κλίμακα συχνοτήτων Mel και απεικονίζονται στην Εικόνα 3.6.1. Η κλίμακα αυτή, έχει την ικανότητα να προσεγγίζει τον τρόπο που αντιλαμβάνεται τους ήχους το ανθρώπινο αυτί, γι' αυτό και προτιμάται με μοντέλα αναγνώρισης φωνής [41].



Εικόνα 3.6.1: Απεικόνιση Mel-φασματογραφήματος σήματος ήχου για είσοδο σε συνελκτικό νευρωνικό δίκτυο [41].

3.7 Αξιολόγηση Αλγορίθμων Μηχανικής Μάθησης

Η αξιολόγηση των αλγορίθμων μηχανικής μάθησης αποτελεί καίριο παράγοντα στην εξαγωγή συμπερασμάτων για το πόσο αποδοτικός είναι ο αλγόριθμος ως προς το πρόβλημα που στοχεύει να επιλύσει. Για τον λόγο αυτό χρησιμοποιούμε κάποιες μετρικές για την εξαγωγή αριθμητικών αποτελεσμάτων, που θα καθιστούν τα μοντέλα συγκρίσιμα ως προς την απόδοσή τους. Ο σκοπός της αξιολόγησης είναι, χρησιμοποιώντας μετρικές απόδοσης (performance metrics), να εκτιμήσει

πόσο καλά ένας αλγόριθμος γενικεύει από τα δεδομένα εκπαίδευσης στα δεδομένα δοκιμής ώστε να επιλεγεί ο καταλληλότερος αλγόριθμος για κάθε πρόβλημα.

Οι πιο συχνά χρησιμοποιούμενες μετρικές απόδοσης των μοντέλων ταξινόμησης που εξετάζουμε, είναι:

- **Πίνακας Σύγχυσης (Confusion Matrix):** Είναι ένας δισδιάστατος πίνακας, του οποίου οι γραμμές αποτελούνται από τις προβλεπόμενες κατηγορίες ταξινόμησης και οι στήλες του από τις πραγματικές. Στη δυαδική ταξινόμηση ειδικότερα, ο πίνακας λαμβάνει μορφή 2x2 και διακρίνει τα αποτελέσματα σε: Αληθώς Θετικά (True Positives - TP), Αληθώς Αρνητικά (True Negatives- TN), Ψευδώς Θετικά (False Positives - FP) και Ψευδώς Αρνητικά (False Negatives – FN). Πιο συγκεκριμένα:
 - Αληθώς Θετικά (TP): αληθώς θετικά, χαρακτηρίζονται τα αποτελέσματα που είναι στην πραγματικότητα θετικά και ταξινομήθηκαν ορθώς από το μοντέλο ως θετικά.
 - Αληθώς Αρνητικά (TN): Ως αληθώς αρνητικά χαρακτηρίζονται τα αποτελέσματα που είναι στην πραγματικότητα αρνητικά και ταξινομήθηκαν ορθώς από το μοντέλο ως αρνητικά.
 - Ψευδώς Θετικά (FP): ψευδώς θετικά, χαρακτηρίζονται τα αποτελέσματα που είναι στην πραγματικότητα αρνητικά και ταξινομήθηκαν λανθασμένα από το μοντέλο ως θετικά.
 - Ψευδώς Αρνητικά (FN): Ως ψευδώς αρνητικά χαρακτηρίζονται τα αποτελέσματα που είναι στην πραγματικότητα θετικά και ταξινομήθηκαν λανθασμένα από το μοντέλο ως αρνητικά.

		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	Sensitivity $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{(TN + FP)}$
		Precision $\frac{TP}{(TP + FP)}$	Negative Predictive Value $\frac{TN}{(TN + FN)}$	Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$

Εικόνα 3.7.1: Απεικόνιση Πίνακα Σύγχυσης με βασικές μετρικές αξιολόγησης για την απόδοση μοντέλων ταξινόμησης [34].

- **Ακρίβεια (Accuracy- Acc):** Η ακρίβεια αναπαριστά τον λόγο των σωστών προβλέψεων (TP+TN) προς το συνολικό πλήθος των αποτελεσμάτων (TP+TN+FN+FP). Οι τιμές που λαμβάνει κυμαίνονται μεταξύ 0,0-1,0 με ιδανική τιμή το 1,0.

$$\circ \frac{(TP+TN)}{(TP+TN+FN+FP)}$$

- **Πιστότητα (Precision – PR):** Η πιστότητα εκφράζει τον λόγο των αληθώς θετικών προβλέψεων (TP) προς το πλήθος του συνόλου των θετικών προβλέψεων (TP+FP). Οι τιμές που λαμβάνει κυμαίνονται επίσης μεταξύ 0,0-1,0 με ιδανική τιμή το 1,0.

$$\circ \frac{TP}{(TP+FP)}$$

- **Ευαισθησία (Sensitivity- SN /Recall- RE):** Η ευαισθησία εκφράζει τον λόγο των αληθώς θετικών προβλέψεων (TP) προς το πλήθος των πραγματικά θετικών δεδομένων (TP+FN). Οι τιμές που λαμβάνει κυμαίνονται και πάλι μεταξύ 0,0-1,0 με ιδανική τιμή το 1,0.

$$\circ \frac{TP}{(TP+FN)}$$

- **Ειδικευση (Specificity – SP):** Η ειδικευση ορίζεται ως ο λόγος των αληθώς αρνητικών προβλέψεων (TN) προς το πλήθος των πραγματικά αρνητικών δεδομένων (TN+FP). Οι τιμές που λαμβάνει κυμαίνονται, ομοίως, μεταξύ 0,0-1,0 με ιδανική τιμή το 1,0.

$$\circ \frac{TN}{(TN+FP)}$$

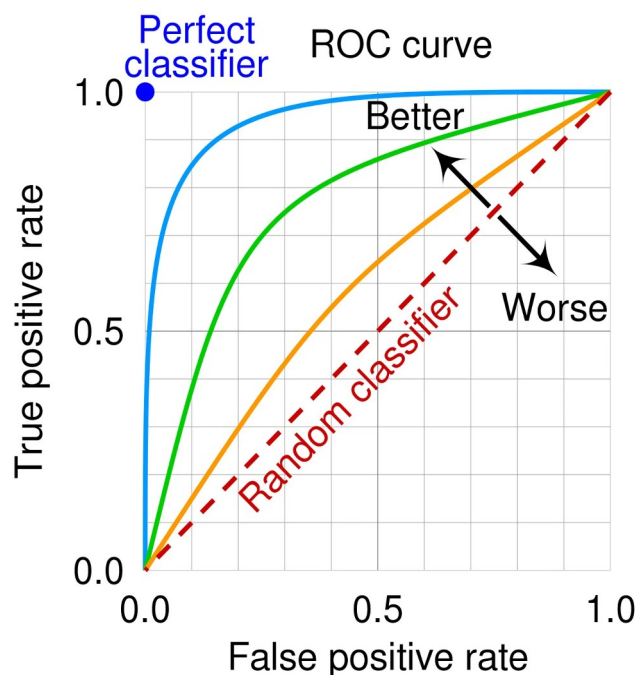
- **F1-Βαθμολογία (F1-Score):** Ως F1-Βαθμολογία ορίζεται ο αρμονικός μέσος της Πιστότητας και της Ευαισθησίας. Συνδυάζοντας τις δύο μετρικές, η F1-βαθμολογία παρέχει μία πιο ισορροπημένη μέτρηση της απόδοσης του αλγορίθμου και χρησιμοποιείται για να αξιολογήσει την απόδοση του μοντέλου σε σύνολα δεδομένων που είναι ανισομερώς κατανομημένα. Οι τιμές που λαμβάνει κυμαίνονται μεταξύ 0,0-1,0 με ιδανική τιμή το 1,0.

$$\circ F_1 = 2 \frac{(PR * SN)}{(PR + SN)} = \frac{2TP}{2TP + FP + FN}$$

- **Ισορροπημένη Ακρίβεια (Balanced Accuracy):** Η ισορροπημένη ακρίβεια αναπαριστά τη μέση τιμή της ευαισθησίας και της ειδικευσης και χρησιμοποιείται ως μια δικαιότερη εκτίμηση της απόδοσης του μοντέλου σε σχέση με τη συμβατική ακρίβεια. Οι τιμές που λαμβάνει κυμαίνονται μεταξύ 0,0-1,0 με ιδανική τιμή το 1,0.

$$\circ \text{Balanced Accuracy} = \frac{1}{2}(SN + SP)$$

- **ROC-AUC:** Η καμπύλη ROC (Receiver Operating Characteristic Curve) είναι μία γραφική παράσταση που κάθε σημείο της απεικονίζει τον TP δείκτη στον άξονα y και τον FP δείκτη στον άξονα x για τις διαφορετικές τιμές κατωφλίων. Επιθυμούμε όσο το δυνατόν υψηλότερη τιμή TP και χαμηλότερη FP και επομένως η καμπύλη να βρίσκεται αριστερά της διαγωνίου των αξόνων, όπως αναπαριστάται και στο ακόλουθο σχήμα. Η Περιοχή Κάτω Από Την Καμπύλη (Area Under Curve – AUC) παρέχει ένα συνολικό μέτρο απόδοσης σε όλα τα δυνατά κατώφλια ταξινόμησης και κυμαίνεται μεταξύ 0,0-1,0 με ιδανική τιμή το 1,0.



Εικόνα 3.7.2: Απεικόνιση καμπύλης ROC (Receiver Operating Characteristic Curve) [42].

3.8 Μηχανική Μάθηση στην Ιατρική

Όπως και σε πολλούς άλλους κλάδους της επιστήμης, έτσι και στην ιατρική, η μηχανική μάθηση μπορεί να προσφέρει καινοτόμες και αποτελεσματικές λύσεις, τόσο για τη διάγνωση ασθενειών όσο και στη βελτίωση των θεραπευτικών πρακτικών. Η ευρεία ποικιλία των ιατρικών δεδομένων όπως ηλεκτρονικοί φάκελοι υγείας (ΗΦΥ), βιολογικοί δείκτες, εικόνες, βίντεο και ηχητικά σήματα, που μας προσφέρουν οι σύγχρονες τεχνολογίες, επιτρέπουν την ανάπτυξη πολλών εφαρμογών μηχανικής μάθησης για την υποβοήθηση των ιατρικών αποφάσεων και την διευκόλυνση του έργου των επιστημόνων υγείας. Οι τομείς της ιατρικής τους οποίους μπορεί να ωφεληθεί η ένταξη της μηχανικής μάθησης ποικίλουν. Για παράδειγμα, η ανάπτυξη εφαρμογής που θα ανιχνεύει το στρες και τα επεισόδια άγχους αποδεικνύεται ιδιαίτερα χρήσιμη για την παρακολούθηση και τη βελτίωση της ψυχικής υγείας [4]. Στις μεταμοσχεύσεις επίσης, αναλύονται ιατρικά δεδομένα για να προβλεφθεί η απόρριψη μοσχεύματος ή η καταλληλότητα των δωρητών [5]. Μοντέλα μηχανικής μάθησης μπορούν ακόμη, να προβλέψουν τον κίνδυνο εμφάνισης καρδιαγγειακών νοσημάτων ως απόρροια μακροχρόνιας νόσησης από σακχαρώδη διαβήτη τύπου 2 [43], είτε, μέσω της επεξεργασίας φυσικής γλώσσας, να χρησιμοποιηθούν στην πρόβλεψη υποτροπής στην κατάχρηση εξαρτησιογόνων ουσιών [6]. Ειδικότερα, ως προς την αξιοποίηση των συνελκτικών νευρωνικών δικτύων, τα οποία όπως προαναφέραμε εξειδικεύονται στην επεξεργασία εικόνων για εξαγωγή συμπερασμάτων, αυτά μπορούν να τύχουν εφαρμογής τόσο από τους ίδιους τους ασθενείς, όσο και από την ιατρική κοινότητα. Κατ'οίκον, μπορούν να αξιοποιηθούν για την υποβοήθηση της αυτοδιάγνωσης, όπως για παράδειγμα αυτοδιάγνωση νόσησης από COVID-19 [7] ή για την εκτίμηση της σοβαρότητας μίας δερματικής πάθησης [8], ώστε να γίνει έγκαιρη απεύθυνση σε επιστήμονα υγείας. Αντίστοιχα, στον τομέα των ιατρικών εφαρμογών, η αξιοποίησή τους αφορά έναν μεγάλο αριθμό δραστηριοτήτων για την επεξεργασία εικόνων ιατρικής απεικόνισης, όπως η ανάλυση εικόνων απεικόνισης μαγνητικού συντονισμού (MRI) που υποβοηθά τη διάγνωση ατόμων με άνοια [44] και σε άλλη περίπτωση, τη διάγνωση ήπιας γνωστικής διαταραχής και της νόσου Αλτσχάιμερ [45].

4 Αμεροληψία στη Μηχανική Μάθηση

4.1 Η έννοια της Αμεροληψίας στη Μηχανική Μάθηση

Τα τελευταία χρόνια, η ραγδαία εξέλιξη και ένταξη των εφαρμογών τεχνητής νοημοσύνης στην καθημερινή ζωή, εγείρει την ανάγκη διερεύνησης του κατά πόσο μπορούν οι εφαρμογές αυτές να είναι εξίσου ωφέλιμες στο σύνολο του πληθυσμού που εν δυνάμει θα μπορεί να τις χρησιμοποιήσει. Από απλούστερα παραδείγματα προβλημάτων λήψης αποφάσεων, όπως προτάσεις για αγορές εμπορικών προϊόντων, επιλογής θεαμάτων και ανάγνωσης άρθρων στα μέσα κοινωνικής δικτύωσης, έως μέχρι πιο σύνθετα, όπως η ιατρική διάγνωση, η έγκριση δανείων και η υποβοήθηση λήψης δικαστικών αποφάσεων, είναι εμφανείς οι σοβαρές επιπτώσεις που μπορεί να έχει η λανθασμένη χρήση των εφαρμογών μηχανικής μάθησης, καθιστώντας επιτακτική την ανάγκη αποφυγής των διακρίσεων (bias). Έτσι, πέραν της καθεαυτής απόδοσης ενός μοντέλου, για την αξιολόγηση της αποτελεσματικότητάς του, εισάγονται έννοιες όπως η Αμεροληψία (Fairness). Ως αμεροληψία, ορίζεται η ικανότητα ενός μοντέλου μηχανικής μάθησης να πραγματοποιεί προβλέψεις με δίκαιο τρόπο ανάμεσα σε άτομα ή διαφορετικές υποομάδες του πληθυσμού, χωρίς να κάνει διακρίσεις που βασίζονται στα ευαίσθητα χαρακτηριστικά τους (sensitive features) [13]. Ευαίσθητα χαρακτηριστικά αποτελούν στοιχεία όπως το φύλο, η ηλικία, η φυλή, η γλώσσα ομιλίας, η οικονομική κατάσταση ή το ιατρικό ιστορικό ενός ατόμου. Η σημασία της αμεροληψίας είναι διττή: αφενός προάγει την ηθική χρήση των τεχνολογιών, αφετέρου ενισχύει την αξιοπιστία και την αποδοχή των συστημάτων από το κοινό.

Για την αξιολόγηση της αμεροληψίας έχουν αναπτυχθεί διάφορες μετρικές, που διευκολύνουν τη σύγκριση των προβλέψεων ενός μοντέλου μηχανικής μάθησης μεταξύ διαφορετικών υποομάδων του συνόλου δεδομένων. Ειδικότερα στην δυαδική ταξινόμηση, οι ευρύτερα χρησιμοποιούμενες μετρικές αμεροληψίας αναλύονται ακολούθως:

- **Δημογραφική Ισότητα (Demographic Parity):** Η δημογραφική ισότητα εκφράζει την ισότητα στην πιθανότητα θετικής πρόβλεψης, μεταξύ δημογραφικών ομάδων που διαχωρίζονται βάσει ενός ευαίσθητου χαρακτηριστικού A (όπως το φύλο ή η ηλικία). Πιο συγκεκριμένα, λέμε ότι ένας δυαδικός προβλεπτικός αλγόριθμος $\hat{Y}$ ικανοποιεί την δημογραφική ισότητα σε σχέση με το ευαίσθητο χαρακτηριστικό A όταν η πιθανότητα είναι ίση για όλες τις τιμές του ευαίσθητου χαρακτηριστικού A , δηλαδή [46]:

$$\circ P(\hat{Y}/A=0)=P(\hat{Y}/A=1)$$

Όπως γίνεται σαφές, η δημογραφική ισότητα εξασφαλίζει ίδια πρόβλεψη σε όλες τις υποομάδες, ανεξάρτητα από τα αληθή αποτελέσματα και επομένως μπορεί να υπολογιστεί συγκρίνοντας τον Ρυθμό Επιλογής (Selection Rate – SR) των υποομάδων, δηλαδή το ποσοστό των παραδειγμάτων που ταξινομήθηκαν ως θετικά: $SR = \frac{(TP+FP)}{Total}$

- **Ισότητα Πιθανότητας (Equalized Odds):** Η ισότητα πιθανότητας εξετάζει την ισότητα στην πιθανότητα πρόβλεψης μεταξύ δημογραφικών ομάδων που διαχωρίζονται βάσει ενός ευαίσθητου χαρακτηριστικού A , λαμβάνοντας όμως υπόψη και το αληθές αποτέλεσμα. Δηλαδή, ένας δυαδικός προβλεπτικός αλγόριθμος $\hat{Y}$ ικανοποιεί την δημογραφική ισότητα σε σχέση με το ευαίσθητο χαρακτηριστικό A αλλά και το αποτέλεσμα Y , αν $\hat{Y}$ και A είναι ανεξάρτητα για κοινό Y [47]:

$$\circ P(\hat{Y}/A=0, Y=y)=P(\hat{Y}/A=1, Y=y), y \in (0,1)$$

Για το αποτέλεσμα $y=1$, ο περιορισμός απαιτεί το $\hat{Y}$ να έχει ίσο ρυθμό αληθώς θετικών προβλέψεων (TPR) ανάμεσα στις δύο δημογραφικές ομάδες $A=0$ και $A=1$. Ο ρυθμός αληθώς θετικών προβλέψεων υπολογίζεται ως: $TPR = \frac{TP}{(TP+FP)}$

Για το $y=0$, ο περιορισμός εξισώνει τους ρυθμούς ψευδώς θετικών (FPR). Ο ρυθμός ψευδώς θετικών προβλέψεων, εκφράζει το ποσοστό λανθασμένα θετικών προβλέψεων προς το σύνολο των αρνητικών αποτελεσμάτων και υπολογίζεται ως: $FPR = \frac{FP}{(FP+TN)}$

Έτσι, οι ίσες πιθανότητες επιβάλλουν τόσο ίση μεροληψία όσο και ίση ακρίβεια σε όλες τις δημογραφικές ομάδες, αποφεύγοντας τόσο την υπερ-πρόβλεψη θετικών όσο και την υπό-πρόβλεψη θετικών [47].

- **Ίση Ευκαιρία (Equal Opportunity):** Στην περίπτωση δυαδικών αποτελεσμάτων, συχνά θεωρούμε το $Y=1$ ως το «πλεονεκτικό» αποτέλεσμα, όπως «μη χρεοκοπία σε δάνειο», «εισαγωγή σε πανεπιστήμιο» ή «προαγωγή σε εργασία». Μια διευρυμένη εκδοχή του περιορισμού της ισότητας πιθανότητας είναι να απαιτείται ίση πρόβλεψη μόνο για την ομάδα του «πλεονεκτικού» αποτελέσματος, χωρίς να καθορίζεται κάποιος περιορισμός για εκείνους που έχουν αρνητικό αποτέλεσμα. Έτσι, ένας δυαδικός προβλεπτικός αλγόριθμος $\hat{Y}$ ικανοποιεί την ίση ευκαιρία σε σχέση με τα A και Y αν:

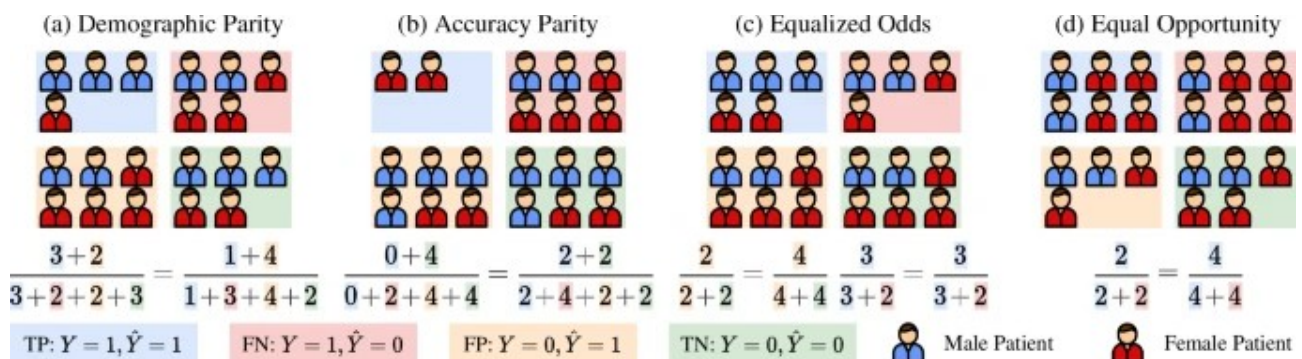
$$P(\hat{Y}/A=0, Y=1) = P(\hat{Y}/A=1, Y=1)$$

Επομένως, στην ίση ευκαιρία ο περιορισμός απαιτεί το $\hat{Y}$ να έχει μόνο ίσο ρυθμό αληθώς θετικών προβλέψεων (TPR) ανάμεσα στις δύο δημογραφικές ομάδες [47].

- **Ισότητα Ακρίβειας (Accuracy Parity):** Η ισότητα ακρίβειας εκφράζει την ισότητα στην πιθανότητα ορθής πρόβλεψης, μεταξύ δημογραφικών ομάδων που διαχωρίζονται βάσει του ευαίσθητου χαρακτηριστικού A . Λέμε ότι ένας δυαδικός προβλεπτικός αλγόριθμος $\hat{Y}$ ικανοποιεί την ισότητα ακρίβειας σε σχέση με το ευαίσθητο χαρακτηριστικό A , όταν η πιθανότητα ορθής πρόβλεψης είναι ίση για όλες τις τιμές του ευαίσθητου χαρακτηριστικού A , δηλαδή [48]:

$$P(\hat{Y}=Y/A=0) = P(\hat{Y}=Y/A=1)$$

Όπως γίνεται σαφές, η δημογραφική ισότητα εξασφαλίζει ίδια ακρίβεια πρόβλεψης σε όλες τις υποομάδες και επομένως υπολογίζεται συγκρίνοντας τις τιμές ακριβείας τους.



Εικόνα 4.1.1: Σχηματική απεικόνιση των μετρικών αξιολόγησης της Αμεροληψίας ως προς το φύλο, σε ένα πρόβλημα δυαδικής ταξινόμησης [48].

Είναι σημαντικό να έχουμε υπόψη ότι κάθε φορά πρέπει να αναλύουμε τις τιμές των μετρικών αμεροληψίας ανάλογα με το πρόβλημα που αντιμετωπίζουμε. Πιο συγκεκριμένα, η δημογραφική ισότητα, παρότι αποτελεί μια από τις βασικές μετρικές αμεροληψίας, μπορεί να παρουσιάσει περιορισμούς που πρέπει να λαμβάνονται υπόψη κατά την εφαρμογή της. Για παράδειγμα, η εξέταση μόνο του δείκτη δημογραφικής ισότητας στην επιλογή κατάλληλων υποψηφίων για μια θέση εργασίας, μπορεί να οδηγήσει μεν στην επιλογή των πιο καταρτισμένων ατόμων μιας υποομάδας, αλλά ταυτόχρονα να επιλέγονται τυχαία άτομα από άλλη υποομάδα λόγω της προσπάθειας για αμεροληψία. Αυτό μπορεί να συμβεί ιδιαίτερα όταν υπάρχει έλλειψη αντιπροσωπευτικών δεδομένων για όλες τις υποομάδες. Επιπλέον, εάν η πιθανότητα επιτυχίας μιας πρόβλεψης συνδέεται άμεσα με ένα ευαίσθητο χαρακτηριστικό, όπως για παράδειγμα η πρόβλεψη μη αποπληρωμής ενός δανείου συνδέεται άμεσα με την οικονομική κατάσταση του δανειολήπτη, η δημογραφική ισότητα μπορεί να επιβάλει μη ρεαλιστικούς περιορισμούς. Αυτό ενδέχεται να οδηγήσει σε χαμηλή ακρίβεια και λιγότερο αποτελεσματικές αποφάσεις [47]. Επομένως, κρίνεται ιδιαίτερα σημαντικός ο συνδυασμός των μετρικών αμεροληψίας για την εξαγωγή ποιοτικότερων συμπερασμάτων ανάλογα με το πρόβλημα που αντιμετωπίζουμε.

4.2 Αμεροληψία στον τομέα της Ιατρικής

Οι ολοένα και αυξανόμενες εφαρμογές των αλγορίθμων μηχανικής μάθησης στο πεδίο της ιατρικής έχει επαναπροσδιορίσει τη διάγνωση, την πρόγνωση και τη θεραπεία ασθενειών. Η εξάλειψη των διακρίσεων σε αποφάσεις που σχετίζονται με την ανθρώπινη υγεία κρίνεται απαραίτητη, καθώς τέτοιες αποφάσεις μπορεί να έχουν πολύ σοβαρές επιπτώσεις σε τομείς όπως:

- **Διάγνωση:** Μεροληψίες του μοντέλου μπορεί να οδηγήσουν σε άνιση ακρίβεια διάγνωσης μεταξύ διαφορετικών πληθυσμιακών ομάδων, όπως για παράδειγμα σε μία εφαρμογή ανάλυσης εικόνας για διάγνωση δερματικών παθήσεων, η οποία παρουσιάζει διακρίσεις ως προς το φύλο [8], αλλά και σε μοντέλο διάγνωσης ήπιας γνωστικής διαταραχής και νόσου Αλτσχάιμερ, το οποίο εμφανίζει διακρίσεις που σχετίζονται με τη φυλή και την ηλικία των ασθενών [49].
- **Θεραπεία και Κατανομή Πόρων:** Αλγόριθμοι που χρησιμοποιούνται για τη σύσταση θεραπευτικών πρωτοκόλλων ή κατανομής πόρων μπορεί να δίνουν προτεραιότητα σε συγκεκριμένες ομάδες, αφήνοντας άλλες χωρίς επαρκή υποστήριξη. Για παράδειγμα, από έναν αλγόριθμο προτεραιοποίησης ασθενών για μεταμόσχευση οργάνων, εάν υποεκπροσωπούνται ορισμένες ομάδες (π.χ., εθνότητες), μπορεί να εκτιμηθεί λανθασμένα ότι οι ομάδες αυτές έχουν χαμηλότερες πιθανότητες επιτυχίας θεραπείας με αποτέλεσμα να συμπεραίνεται χαμηλή προτεραιοποίηση [5].
- **Προσωποποιημένη Ιατρική:** Η προσωποποιημένη ιατρική βασίζεται σε αλγορίθμους που αναλύουν ατομικά γενετικά, βιολογικά και κλινικά δεδομένα για να προτείνουν θεραπείες προσαρμοσμένες στις ανάγκες του κάθε ασθενή. Ωστόσο, οι αλγόριθμοι αυτοί μπορεί να οδηγήσουν σε μεροληπτικά αποτελέσματα, όπως διακρίσεις φύλου σε εφαρμογή εξατομικευμένων προτάσεων διαχείρισης χρόνιου πόνου [50].

Δικαιολογημένα επομένως, η αμεροληψία αποτελεί ολοένα και σημαντικότερο κριτήριο στην επιλογή κατάλληλων μοντέλων σε ιατρικές εφαρμογές. Οι παραπάνω κίνδυνοι, καταδεικνύουν με σαφήνεια την επιτακτική ανάγκη διερεύνησης αλλά και ενίσχυσής της.

4.3 Αντιμετώπιση των Διακρίσεων

Όπως προκύπτει από την ανάλυση που προηγήθηκε, η εξάλειψη της μεροληψίας θα πρέπει να αποτελεί πρωταρχικό μέλημα κατά την ενασχόληση με εφαρμογές μηχανικής μάθησης και ειδικότερα στον τομέα της ιατρικής. Οι μέθοδοι αντιμετώπισης των διακρίσεων (bias mitigating

methods) κατηγοριοποιούνται συνήθως σε τρία είδη: μέθοδοι προεπεξεργασίας (pre-processing methods), μέθοδοι κατά την επεξεργασία (in-processing methods) και μέθοδοι μετά την επεξεργασία (post-processing methods). Πιο συγκεκριμένα [51],[52],[49]:

- **Μέθοδοι Προεπεξεργασίας:** Οι μέθοδοι προεπεξεργασίας στοχεύουν στην εξάλειψη των διακρίσεων με την κατάλληλη επεξεργασία των δεδομένων εισόδου πριν την εκπαίδευση του μοντέλου. Οι μέθοδοι αυτές είναι αποτελεσματικότερες σε περιπτώσεις που η μεροληψία προκύπτει από την ανισορροπία στα δεδομένα εκπαίδευσης ή από προκαταλήψεις που ενσωματώνονται σε αυτά. Μερικές βασικές τεχνικές αυτής της μεθόδου είναι :
 - **Δειγματοληψία (Sampling):** Κατάλληλη προσαρμογή του όγκου των δεδομένων κάθε υποομάδας με σκοπό την ισορροπημένη κατανομή των ευαίσθητων χαρακτηριστικών στο σύνολο των δεδομένων. Αυτό μπορεί να επιτευχθεί για παράδειγμα, με την αύξηση του δείγματος μειονοτικών ομάδων (upsampling) ή μείωση του δείγματος πλειονοτικών ομάδων (downsampling).
 - **Επαύξηση Δεδομένων (Data Augmentation):** Δημιουργία επιπλέον συνθετικών δεδομένων για την εξισορρόπηση των ανισοτήτων στις κατηγορίες ευαίσθητων χαρακτηριστικών. Χαρακτηριστικό παράδειγμα αποτελεί η εναλλαγή των στοιχείων μεταξύ ανδρών και γυναικών με σκοπό την μείωση των προκαταλήψεων φύλου.
 - **Αφαίρεση Συσχετίσεων (Correlation Remover):** Αφαιρούνται από τα χαρακτηριστικά των δεδομένων εκείνα που έχουν άμεση συσχέτιση με τα ευαίσθητα χαρακτηριστικά.
- **Μέθοδοι κατά την Επεξεργασία:** Οι μέθοδοι που εφαρμόζονται κατά την επεξεργασία του μοντέλου, επιδιώκουν την αντιμετώπιση των διακρίσεων κατά την διαδικασία της εκπαίδευσης. Ο σκοπός αυτός μπορεί να επιτευχθεί με:
 - **Περιορισμούς Μοντέλου (Model Constraints):** Αυτή η τεχνική αποτελεί προσθήκη όρων κανονικοποίησης με την επιβολή περιορισμών ή την προσθήκη επιπλέον συνθηκών στις παραμέτρους, ώστε να διασφαλιστεί ότι το μοντέλο εκπαιδεύεται με τρόπο που μειώνει τις διακρίσεις.
 - **Ανταγωνιστική Μάθηση (Adversarial Learning):** Κατά την ανταγωνιστική μάθηση εκπαιδεύουμε δύο μοντέλα ταυτόχρονα, όπου το ένα επιτελεί την λειτουργία ενός κλασικού προβλεπτικού μοντέλου για την κύρια εργασία, ενώ το ανταγωνιστικό του εκπαιδεύεται στην αναγνώριση των ευαίσθητων χαρακτηριστικών που στη συνέχεια εισάγει αντεστραμμένα πίσω στην είσοδο, ώστε να αναγκάσει το κύριο μοντέλο να κάνει ανθεκτικές προβλέψεις ως προς τα ευαίσθητα χαρακτηριστικά.
- **Μέθοδοι μετά την Επεξεργασία:** Οι μέθοδοι που εφαρμόζονται μετά την επεξεργασία προσαρμόζουν τις προβλέψεις του μοντέλου αφού έχει πραγματοποιηθεί η εκπαίδευσή του, έτσι ώστε να εξλειφθούν οι διακρίσεις στα τελικά αποτελέσματα. Ενδεικτικές τεχνικές αυτής της μεθόδου αποτελούν:
 - **Κατωφλίωση (Thresholding):** Θέτει διαφορετικά κατώφλια για κάθε υποομάδα μιας κατηγορίας ευαίσθητων χαρακτηριστικών ανάλογα με τις διακρίσεις που αντιμετωπίζει, ώστε να βελτιώσει την αμεροληψία του μοντέλου.
 - **Καλιμπράρισμα (Calibration):** Διορθώνει τις πιθανότητες πρόβλεψης του μοντέλου ώστε να πετύχει αναλογία των θετικών προβλέψεων με την πραγματική πιθανότητα εμφάνισης ενός γεγονότος για όλες τις υποομάδες.



Εικόνα 4.3.1: Σχηματική απεικόνιση των τεχνικών εξάλειψης των διακρίσεων στα μοντέλα Μηχανικής Μάθησης, οι οποίες μπορούν να εφαρμοστούν είτε μεμονωμένα είτε συνδυαστικά [53].

Η επιλογή της κατάλληλης μεθόδου αντιμετώπισης των διακρίσεων εξαρτάται από το πρόβλημα που καλούμαστε να επιλύσουμε, τον τύπο των δεδομένων και από το πού στοχεύουμε όσον αφορά την αμεροληψία.

5 Εφαρμογές Μηχανικής Μάθησης κατά τη διάρκεια της Πανδημίας COVID-19

Με το ξέσπασμα της πανδημίας COVID-19, τα αυστηρά μέτρα περιορισμού, το κόστος των διαγνωστικών ελέγχων και η γενικότερη αβεβαιότητα που προξένησε η καινούργια νόσος, δημιούργησε την ανάγκη για γρήγορη διάγνωση των ασθενών και διερεύνηση της διασποράς του ιού. Επιπλέον, η συνεχής παρακολούθηση του ιού και των επιπτώσεών του οδήγησαν στη δημιουργία μεγάλου όγκου νέων ιατρικών δεδομένων προς επεξεργασία για την εξαγωγή συμπερασμάτων που αφορούσαν τη νόσο. Στις συνθήκες αυτές, βρήκε πρόσφορο έδαφος η ανάπτυξη εφαρμογών μηχανικής μάθησης για την επεξεργασία εικόνων και ήχων που σχετίζονταν με συμπτώματα της COVID-19 [54], όπως οι ακτινογραφίες θώρακα [55] και τα ανθρώπινα ηχητικά σήματα [9]. Οι εφαρμογές αυτές αποτέλεσαν χρήσιμο εργαλείο για την αντιμετώπιση αυτής της παγκόσμιας κρίσης στην Υγεία και άνοιξαν νέους δρόμους στην ανάπτυξη επιστημονικών μεθόδων στα πεδία της αυτοδιάγνωσης και της τηλεϊατρικής.

Η ίδια η φύση του ιού (SARS)-CoV-2 στρέφει τις μεθόδους των μοντέλων διάγνωσης του στο πεδίο της επεξεργασίας ήχου, καθώς προσβάλλει κυρίως το αναπνευστικό σύστημα και προκαλεί συμπτώματα ξηρού βήχα, δύσπνοιας και βραχνής φωνής που αλλοιώνουν το φυσιολογικό ηχητικό σήμα που παράγει ένα υγιές άτομο. Για τη συγκέντρωση επαρκών δεδομένων που εξακολουθούν να χρησιμοποιούνται στα μοντέλα διάγνωσης, αναπτύχθηκαν εφαρμογές συλλογής ηχητικών δειγμάτων, προσβάσιμες από ηλεκτρονικές συσκευές με τη χρήση των μικροφώνων τους. Δημιουργήθηκαν έτσι μεγάλα σύνολα δεδομένων για επεξεργασία χαρακτηριστικών ήχου με κυριότερα παραδείγματα το “COVID-19 Sounds” [7], το “COUCHVID” [11], το “Coswara” [56], καθώς και το “smarty4covid” του Εθνικού Μετσόβιου Πολυτεχνείου [57]. Πολλά μοντέλα μηχανικής μάθησης που εκπαιδεύτηκαν σε αυτά τα δεδομένα, πέτυχαν υψηλά επίπεδα ακρίβειας, καθιστώντας τη διάγνωση γρήγορη, ανέξοδη και μη επεμβατική. Σημαντική πρόκληση ωστόσο, αποτέλεσε και η διατήρηση της ακρίβειας των μοντέλων στην πρόβλεψη, καθώς η δυναμική φύση των δεδομένων λόγω των εξελισσόμενων συνθηκών της πανδημίας, με την πάροδο του χρόνου οδηγούσε σε πτώση της απόδοσης [58].

Παρά την υποχώρηση του πανδημικού χαρακτήρα της νόσου, η καινοτομία των συγκεκριμένων εφαρμογών δεν χάνει τη χρησιμότητά της, καθώς η COVID-19 συνεχίζει να αποτελεί ενεργή νόσο με αρκετά κρούσματα παγκοσμίως. Αρκεί να λάβουμε υπόψη ότι ενδεχόμενη εσφαλμένη θετική διάγνωση COVID-19 λόγω των διακρίσεων στις προβλέψεις ενός μοντέλου, δεν συνιστά πρόβλημα μόνο για την ιατρική κοινότητα, αλλά προκαλεί και σημαντικές κοινωνικοοικονομικές συνέπειες, λόγω των μέτρων περιορισμού για σειρά δραστηριοτήτων του φερόμενου ως ασθενή (π.χ., απουσία από την εργασία). Παράλληλα, η χρήση των εφαρμογών αυτών δεν περιορίζεται αναγκαστικά στην περίπτωση της COVID-19, αλλά μπορεί να διευρυνθεί και σε άλλες νόσους του αναπνευστικού συστήματος [59]. Πέρα από το γεγονός αυτό, η πληθώρα και η διαθεσιμότητα των δεδομένων, δίνει την ευκαιρία για περαιτέρω διερεύνηση και άλλων προεκτάσεων της μηχανικής μάθησης, όπως η αμεροληψία και η ενίσχυσή της. Με την αξιοποίηση των προαναφερθέντων ηχητικών δεδομένων, μοντέλα διάγνωσης που εξετάστηκαν για την αμεροληψία τους και διαπιστώθηκε ότι παρουσίασαν διακρίσεις σε δημογραφικές κατηγορίες όπως το φύλο, η ηλικία, η συχνότητα καπνίσματος, κ.ά., κατέστη δυνατόν να βελτιωθούν ως προς την αμεροληψία τους.[60],[7],[59].

6 Μεθοδολογία

6.1 Δεδομένα

6.1.1 Περιγραφή Συνόλου Δεδομένων

Για τον πειραματισμό μας με μοντέλα πρόβλεψης νόσησης ή μη από COVID-19 αξιοποιήθηκαν ηχητικά σήματα από το σύνολο δεδομένων “COVID-19 Sounds” [9]. Το “COVID-19 Sounds” είναι ένα σύνολο δεδομένων που αναπτύχθηκε από ομάδα ερευνητών του πανεπιστημίου Cambridge της Αγγλίας την περίοδο της πανδημίας. Στόχος της προσπάθειας των ερευνητών, ήταν η αναζήτηση εργαλείων που θα επέτρεπαν μια νέα, οικονομική, μη επεμβατική και ευρέως διαθέσιμη μέθοδο ανίχνευσης κρουσμάτων COVID-19, ως μια χρήσιμη προσθήκη στις ήδη υπάρχουσες μεθόδους διάγνωσης, όπως οι μοριακές εξετάσεις PCR, τα τεστ αντιγόνων και τα τεστ αντισωμάτων, οι οποίες είναι επεμβατικές και σχετικά ακριβές, περιορίζοντας τη δυνατότητα διενέργειας μαζικών τεστ στο ευρύ κοινό. Το εγχείρημα βασίστηκε στην ιδέα ότι σε πολλές περιπτώσεις, άτομα που προσβάλλονται από λοίμωξη του αναπνευστικού συστήματος όπως αυτή που προκαλεί η COVID-19, παρουσιάζουν χαρακτηριστικές αλλαγές στους ήχους της αναπνοής, της ομιλίας ή του βήχα τους. Τα μοτίβα των αλλαγών αυτών, μπορούν να εντοπιστούν με την ανάλυση ήχου μέσω μοντέλων μηχανικής μάθησης. Για τον σκοπό αυτό η ομάδα ερευνητών του Cambridge μέσω μιας εφαρμογής κινητών συσκευών, συνέλεξε ήχους βήχα, βαθιάς αναπνοής και ομιλίας, που συνοδεύονταν από πληροφορίες των συμμετεχόντων σχετικά με δημογραφικά στοιχεία, ιατρικό ιστορικό, κατάσταση καπνίσματος, αποτελέσματα διαγνωστικού ελέγχου COVID-19, και συμπτώματα. Συλλέχθηκαν συνολικά 53.449 ηχητικά δείγματα από 36.116 συμμετέχοντες. Τα δεδομένα αυτά χρησιμοποιήθηκαν αρχικά σε δύο διαγνωστικά προβλήματα: την πρόβλεψη αναπνευστικών συμπτωμάτων και τη διάγνωση COVID-19.

Για τον σκοπό της παρούσας εργασίας, αξιοποιήθηκε μόνο το υποσύνολο του συνόλου δεδομένων που η αρχική μελέτη είχε διαμορφώσει για το δεύτερο σκέλος της έρευνας, δηλαδή τη διάγνωση COVID-19. Το υποσύνολο αυτό, περιλαμβάνει 1482 δείγματα, εκ των οποίων, 732 θετικά και 750 αρνητικά, για λόγους ισοκατανομής. Αποτελείται από ένα ψηφιακό αρχείο υπό μορφή csv και έναν ψηφιακό φάκελο που περιέχει υποφακέλους ξεχωριστά για κάθε συμμετέχοντα, με τα ατομικά αρχεία ήχου και ονομασία σύμφωνα με το όνομα χρήστη του (Uid). Στο αρχείο csv καταγράφεται, χωριστά για κάθε συμμετέχοντα, η διαδρομή (path) που οδηγεί στα αρχεία ήχου (wav) του, η ετικέτα (label) με τον χαρακτηρισμό του ηχητικού δείγματος (με τιμή 0 για αρνητικό και 1 για θετικό), καθώς και επιπλέον πληροφορίες που αφορούν άλλα ευαίσθητα χαρακτηριστικά (sensitive features). Από το σύνολο των πληροφοριών αυτών, στην παρούσα εργασία αξιοποιήθηκαν μόνο τα ηχητικά σήματα βήχα και τα ευαίσθητα χαρακτηριστικά που αφορούν στις κατηγορίες που παρουσιάζει ο Πίνακας 6.1.1, με αντίστοιχες πιθανές τιμές που περιελάμβανε η αρχική έρευνα στο συμπληρωματικό υλικό της [9]:

Πίνακας 6.1.1: Πίνακας κατηγοριοποίησης των ευαίσθητων χαρακτηριστικών των δεδομένων που χρησιμοποιήθηκαν στην παρούσα εργασία.

Κατηγορία	Όνομα Μεταβλητής	Πιθανές Τιμές
Φύλο	Sex	Άνδρας ("Male"), Γυναίκα ("Female"), Άλλο φύλο ("Other"), Άγνωστο φύλο ("pnts")
Ηλικιακή ομάδα	Age	16 έως 20 ("16- 20"), Δεκαετία των 20 ("20-29"), Δεκαετία των 30 ("30-39"), Δεκαετία των 40 ("40-49"), Δεκαετία των 50 ("50-59"), Δεκαετία των 60 ("60-69"), Δεκαετία των 70 ("70-79"), Δεκαετία των 80 ("80- 89"), 90 ετών και άνω ("90-"), Άγνωστη ηλικία ("pnts")
Ιατρικό ιστορικό	Medhistory	Άσθμα ("asthma"), Στενόκαρδο ("angina"), Καρκίνος ("cancer"), Χρόνια αποφρακτική πνευμονοπάθεια ("copd"), Κυστική ίνωση ("cystic"), Διαβήτης ("diabetes"), Υπέρταση ("hbp"), Έμφραγμα ("heart"), HIV ("hiv"), Άλλη χρόνια πάθηση ("long"), Άλλη πάθηση των πνευμόνων ("lung"), Άλλη καρδιακή πάθηση ("otherHeart"), Ιστορικό μεταμόσχευσης οργάνου ("organ"), Πνευμονική ίνωση ("pulmonary"), Ισχαιμικό επεισόδιο ("stroke"), Βαλβιδοπάθεια ("valvular"), Κανένα ("None"), Άγνωστο ("pnts")
Συμπτώματα	Symptoms	Ξηρός βήχας ("drycough"), Παραγωγικός βήχας ("wetcough"), Πονόλαιμος ("sorethroat"), Καταρροή ή ρινική συμφόρηση ("runnyblockednose"), Σφίξιμο στο στήθος ("tightness"), Απώλεια γεύσης ή και όσφρησης ("smelltasteloss"), Πυρετός ("fever"), Ρίγη ("chills"), Δύσπνοια ("shortbreath"), Ζάλη ("dizziness"), Καφαλαλγία ("headache"), Μυαλγία ("muscleache"), Κανένα ("None"), Άγνωστο ("pnts")

6.1.2 Προεπεξεργασία Δεδομένων

Επειδή ο διαχωρισμός των κατηγοριών των δεδομένων του "COVID-19 Sounds" ήταν ιδιαίτερα λεπτομερής στην αρχική έρευνα, στην περίπτωση του παρόντος πειράματος και προκειμένου να επιτύχουμε ασφαλέστερα συμπεράσματα ως προς την αμεροληψία των μοντέλων πρόβλεψης που είναι ο βασικός πυλώνας της παρούσας εργασίας, κρίθηκε σκόπιμο τόσο να αναδιατάξουμε τις ήδη υπάρχουσες υποκατηγορίες δεδομένων, όσο και να δημιουργήσουμε νέες συνδυαστικές κατηγορίες. Αναλυτικά:

- **Ηλικία:** από την κατηγοριοποίηση ανά δεκαετία, προχωρήσαμε σε ορισμό μόνο τεσσάρων υποκατηγοριών:
 - Νεαρότερη ηλικία ("16-39"),

- Μέση ηλικία (“40-59”),
- Άνω των 60 ετών (“above 60”)
- Όσοι δεν απάντησαν (“pnts”).
- **Υποκείμενα Νοσήματα:** Το ιατρικό ιστορικό διαχωρίστηκε σε τέσσερις υποκατηγορίες:
 - Χωρίς υποκείμενα νοσήματα (“None”),
 - Με τουλάχιστον ένα υποκείμενο νόσημα του αναπνευστικού (“Respiratory Disease”) όπως, άσθμα, χρόνια αποφρακτική πνευμονοπάθεια, κυστική ινώση, πνευμονική ίνωση ή άλλη ασθένεια των πνευμόνων, ακόμη και όταν ταυτόχρονα νοσεί και από άλλη ασθένεια.
 - Υποκείμενα νοσήματα που δεν σχετίζονται με το αναπνευστικό σύστημα (“Other Disease”).
 - Όσοι συμμετέχοντες δεν απάντησαν (“pnts”).
- **Συμπτώματα:** Ανάλογα με τον διαχωρισμό του ιατρικού ιστορικού, δημιουργήθηκαν επίσης τέσσερις υποκατηγορίες για τα συμπτώματα,
 - Καθόλου συμπτώματα (“None”),
 - Συμπτώματα του αναπνευστικού (“Respiratory Symptoms”) όπως, ξηρός βήχας, παραγωγικός βήχας, πονόλαιμος, ρινική συμφόρηση, δύσπνοια. Εδώ εντάσσονται και άτομα που παρουσιάζουν έστω και ένα σύμπτωμα του αναπνευστικού, ταυτόχρονα με άλλα συμπτώματα.
 - Όλα τα υπόλοιπα συμπτώματα που δεν σχετίζονται με το αναπνευστικό σύστημα (“Other Symptoms”).
 - Όσοι συμμετέχοντες δεν απάντησαν (“pnts”).
- **Πηγή Συλλογής Δεδομένων:** Εκτός από τον διαχωρισμό των παραπάνω στοιχείων σε νέες υποκατηγορίες, δημιουργήσαμε και μία νέα κατηγορία με όνομα “Source” που διαχωρίζει τα στοιχεία ανάλογα με την πηγή συλλογής δεδομένων η οποία καταδεικνύεται από το μέγεθος του ονόματος χρήστη και τη συχνότητα λήψης του δείγματος. Έτσι έχουμε τις εξής υποκατηγορίες:
 - Δεδομένα που συλλέχθηκαν από συσκευή με λογισμικό iOS (“ios”).
 - Δεδομένα που συλλέχθηκαν από συσκευή με λογισμικό Android (“android”).
 - Δεδομένα που συλλέχθηκαν από την αντίστοιχη σελίδα στο διαδίκτυο και όχι από την εφαρμογή (“web”).

Επιπλέον, δημιουργήσαμε και συνδυαστικές κατηγορίες με τη συνένωση ορισμένων βασικών κατηγοριών, με σκοπό να εξετάσουμε σε βάθος και πιο περίπλοκα σενάρια διακρίσεων. Οι νέες αυτές κατηγορίες είναι:

- **Ηλικία σε συνδυασμό με Φύλο,** με ονομασία “Age_Sex”: τα δεδομένα χωρίζονται σε κατηγορίες ανάλογα με το φύλο και την ηλικιακή ομάδα ως εξής: Γυναίκες νεαρότερης ηλικίας (“16-39-Female”), άνδρες νεαρότερης ηλικίας (“16-39-Male”), άλλο φύλο νεαρότερης ηλικίας (“16-39-Other”), άγνωστο φύλο νεαρότερης ηλικίας (“16-39-pnts”), γυναίκες μέσης ηλικίας (“40-59-Female”), άνδρες μέσης ηλικίας (“40-59-Male”), άλλο φύλο μέσης ηλικίας (“40-59-Other”), άγνωστο φύλο μέσης ηλικίας (“40-59-pnts”), γυναίκες άνω των 60 ετών (“above 60-Female”), άντρες άνω των 60 ετών (“above 60-Male”), άλλο

φύλο άνω των 60 ετών (“above 60-Other”), άγνωστο φύλο άνω των 60 ετών (“above 60-pnts”), γυναίκες άγνωστης ηλικίας (“pnts-Female”), άνδρες άγνωστης ηλικίας (“pnts-Male”), άλλο φύλο άγνωστης ηλικίας (“pnts-Other”), άγνωστο φύλο άγνωστης ηλικίας (“pnts-pnts”).

- **Ηλικία σε συνδυασμό με Συμπτώματα**, με ονομασία “Age_Symptoms”: τα δεδομένα χωρίζονται σε κατηγορίες ανάλογα με την ηλικία και τα συμπτώματα ως εξής: νεαρότερα άτομα χωρίς συμπτώματα (“16-39-none”), άτομα μέσης ηλικίας χωρίς συμπτώματα (“40-59-none”), άτομα άνω των 60 ετών χωρίς συμπτώματα (“above 60-none”), άγνωστη ηλικία χωρίς συμπτώματα (“pnts-none”), νεαρότερα άτομα με συμπτώματα του αναπνευστικού (“16-39-Respiratory Symptoms”), άτομα μέσης ηλικίας με συμπτώματα του αναπνευστικού (“40-59-Respiratory Symptoms”), άτομα άνω των 60 ετών με συμπτώματα του αναπνευστικού (“above 60-Respiratory Symptoms”), άγνωστη ηλικία με συμπτώματα του αναπνευστικού (“pnts-Respiratory Symptoms”), νεαρότερα άτομα με άλλα συμπτώματα (“16-39-Other Symptoms”), άτομα μέσης ηλικίας με άλλα συμπτώματα (“40-59-Other Symptoms”), άτομα άνω των 60 ετών με άλλα συμπτώματα (“above 60-Other Symptoms”), άγνωστη ηλικία με άλλα συμπτώματα (“pnts-Other Symptoms”), νεαρότερα άτομα με άγνωστη κατάσταση συμπτωμάτων (“16-39-pnts”), άτομα μέσης ηλικίας με άγνωστη κατάσταση συμπτωμάτων (“40-59-pnts”), άτομα άνω των 60 ετών με άγνωστη κατάσταση συμπτωμάτων (“above 60-pnts”), άγνωστη ηλικία με άγνωστη κατάσταση συμπτωμάτων (“pnts-pnts”).
- **Φύλο σε συνδυασμό με Συμπτώματα**, με ονομασία “Sex_Symptoms”: τα δεδομένα χωρίζονται σε κατηγορίες ανάλογα με το φύλο και τα συμπτώματα ως εξής: γυναίκες χωρίς συμπτώματα (“Female-none”), άνδρες χωρίς συμπτώματα (“Male-none”), άλλο φύλο χωρίς συμπτώματα (“Other-none”), άγνωστο φύλο χωρίς συμπτώματα (“pnts-none”), γυναίκες με συμπτώματα του αναπνευστικού (“Female-Respiratory Symptoms”), άνδρες με συμπτώματα του αναπνευστικού (“Male-Respiratory Symptoms”), άλλο φύλο με συμπτώματα του αναπνευστικού (“Other-Respiratory Symptoms”), άγνωστο φύλο με συμπτώματα του αναπνευστικού (“pnts-Respiratory Symptoms”), γυναίκες με άλλα συμπτώματα (“Female-Other Symptoms”), άνδρες με άλλα συμπτώματα (“Male-Other Symptoms”), άλλο φύλο με άλλα συμπτώματα (“Other-Other Symptoms”), άγνωστο φύλο με άλλα συμπτώματα (“pnts-Other Symptoms”), γυναίκες με άγνωστη κατάσταση συμπτωμάτων (“Female-pnts”), άνδρες με άγνωστη κατάσταση συμπτωμάτων (“Male-pnts”), άλλο φύλο με άγνωστη κατάσταση συμπτωμάτων (“Other-pnts”), άγνωστο φύλο με άγνωστη κατάσταση συμπτωμάτων (“pnts-pnts”).
- **Ηλικία σε συνδυασμό με ιατρικό ιστορικό**, με ονομασία “Age_Medhistory”: τα δεδομένα χωρίζονται σε κατηγορίες ανάλογα με την ηλικία και το ιατρικό ιστορικό των συμμετεχόντων ως εξής: νεαρότερα άτομα χωρίς υποκείμενα νοσήματα (“16-39-none”), άτομα μέσης ηλικίας χωρίς υποκείμενα νοσήματα (“40-59-none”), άτομα άνω των 60 ετών χωρίς υποκείμενα νοσήματα (“above 60-none”), άγνωστη ηλικία χωρίς υποκείμενα νοσήματα (“pnts-none”), νεαρότερα άτομα με νοσήματα του αναπνευστικού (“16-39-Respiratory Disease”), άτομα μέσης ηλικίας με νοσήματα του αναπνευστικού (“40-59-Respiratory Disease”), άτομα άνω των 60 ετών με νοσήματα του αναπνευστικού (“above 60-Respiratory Disease”), άγνωστη ηλικία με νοσήματα του αναπνευστικού “pnts-Respiratory Disease”, νεαρότερα άτομα με άλλα νοσήματα (“16-39-Other Disease”), άτομα μέσης ηλικίας με άλλα νοσήματα (“40-59-Other Disease”), άτομα άνω των 60 ετών με άλλα νοσήματα (“above 60-Other Disease”), άγνωστη ηλικία με άλλα νοσήματα (“pnts-Other Disease”), νεαρότερα άτομα με άγνωστη κατάσταση νοσημάτων (“16-39-pnts”), άτομα μέσης ηλικίας με άγνωστη κατάσταση νοσημάτων (“40-59-pnts”), άτομα άνω των 60 ετών

με άγνωστη κατάσταση νοσημάτων (“above 60-pnts”), άγνωστη ηλικία με άγνωστη κατάσταση νοσημάτων (“pnts-pnts”).

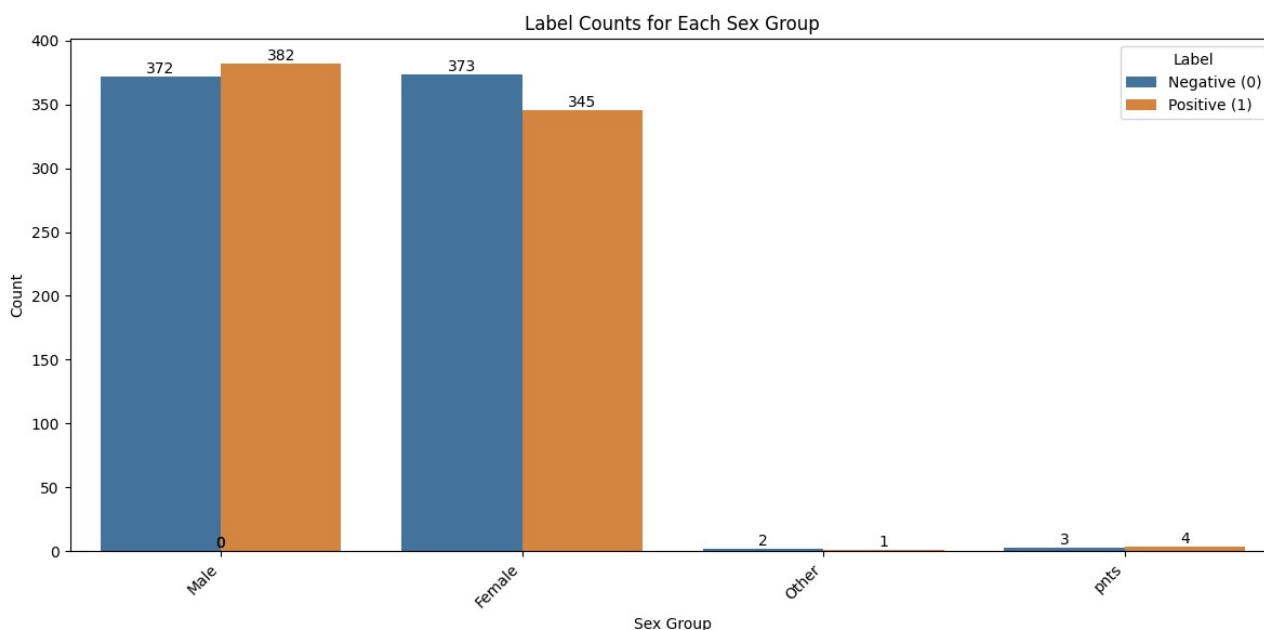
- **Φύλο σε συνδυασμό με ιατρικό ιστορικό**, με ονομασία “Sex_Medhistory”: τα δεδομένα χωρίζονται σε κατηγορίες ανάλογα με το φύλο και το ιατρικό ιστορικό ως εξής: γυναίκες χωρίς υποκείμενα νοσήματα (“Female-none”), άνδρες χωρίς υποκείμενα νοσήματα (“Male-none”), άλλο φύλο χωρίς υποκείμενα νοσήματα (“Other-none”), άγνωστο φύλο χωρίς υποκείμενα νοσήματα “pnts-none”, γυναίκες με νοσήματα του αναπνευστικού (“Female-Respiratory Disease”), άνδρες με νοσήματα του αναπνευστικού (“Male-Respiratory Disease”), άλλο φύλο με νοσήματα του αναπνευστικού (“Other-Respiratory Disease”), άγνωστο φύλο με νοσήματα του αναπνευστικού (“pnts-Respiratory Disease”), γυναίκες με άλλα νοσήματα (“Female-Other Disease”), άνδρες με άλλα νοσήματα (“Male-Other Disease”), άλλο φύλο με άλλα νοσήματα (“Other-Other Disease”), άγνωστο φύλο με άλλα νοσήματα (“pnts-Other Disease”), γυναίκες με άγνωστη κατάσταση νοσημάτων (“Female-pnts”), άνδρες με άγνωστη κατάσταση νοσημάτων (“Male-pnts”), άλλο φύλο με άγνωστη κατάσταση νοσημάτων (“Other-pnts”), άγνωστο φύλο με άγνωστη κατάσταση νοσημάτων (“pnts-pnts”).

6.1.3 Στατιστική Ανάλυση των Ευαίσθητων Χαρακτηριστικών

Πριν περάσουμε στην επεξεργασία των δεδομένων και στην εξαγωγή αποτελεσμάτων, είναι χρήσιμο να γνωρίζουμε την κατανομή του δείγματος και την κατάσταση νόσησης στις επιμέρους κατηγορίες, καθώς μπορεί να φανεί χρήσιμη στην εξαγωγή συμπερασμάτων. Αναλυτικά:

Φύλο:

Το σύνολο δεδομένων περιλαμβάνει 754 άνδρες, 718 γυναίκες, 3 άτομα άλλου φύλου και 7 άτομα αγνώστου φύλου, με την κατανομή σε νοσούντες και μη από COVID-19 να φαίνεται στο παρακάτω διάγραμμα:

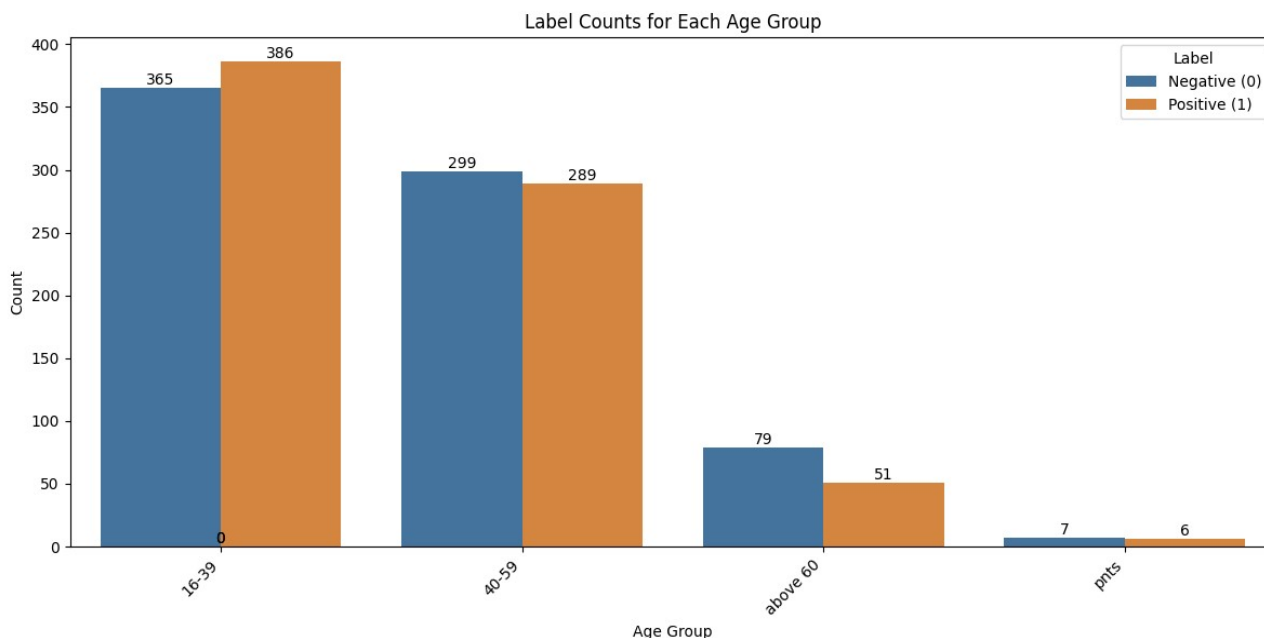


Εικόνα 6.1.1: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Φύλου.

Παρατηρούμε ότι μεταξύ των θετικών δειγμάτων οι περισσότεροι είναι άνδρες, ενώ στα αρνητικά δείγματα έχουμε ισορροπία ανάμεσα σε άνδρες και γυναίκες.

Ηλικία:

Παρατηρούμε υποεκπροσώπηση της ηλικιακής ομάδας άνω των 60 και στα θετικά και στα αρνητικά δείγματα, ενώ η πλειονότητα των συμμετεχόντων ανήκει στην ηλικιακή ομάδα 16-39 ετών. Ο διαχωρισμός των συμμετεχόντων σε θετικούς και αρνητικούς στην COVID-19 παρουσιάζεται στο ακόλουθο διάγραμμα:

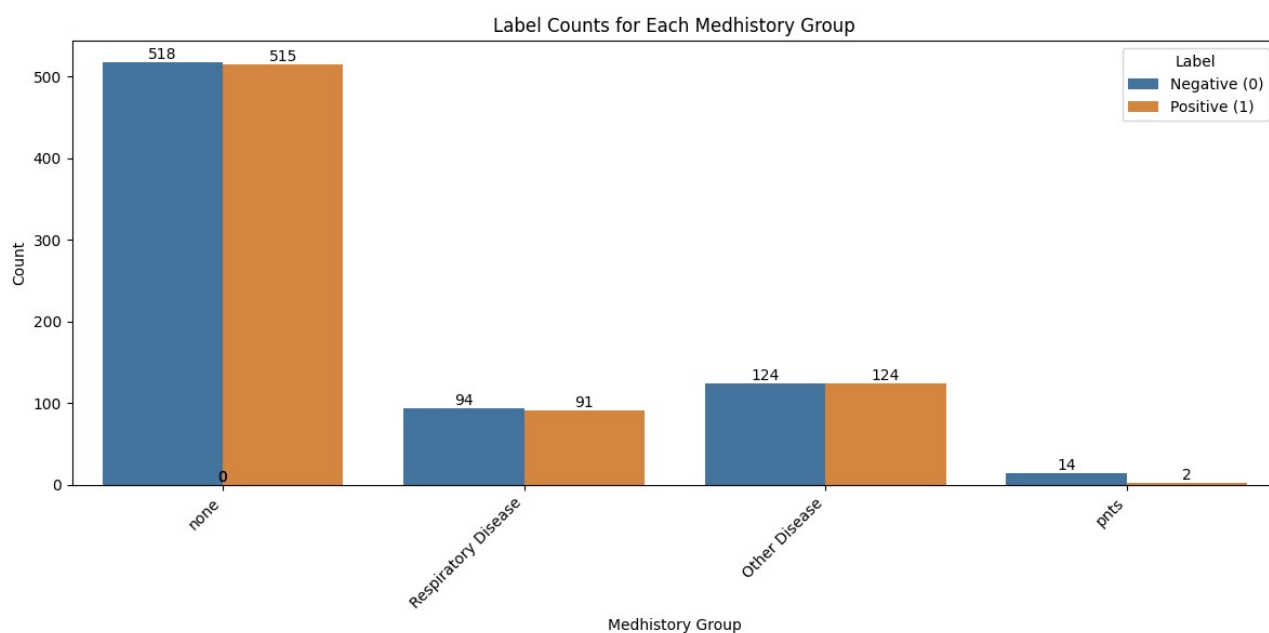


Εικόνα 6.1.2: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Ηλικίας.

Στο διάγραμμα διακρίνουμε ότι στη νεαρότερη ηλικιακή ομάδα εμφανίζονται περισσότερα θετικά δείγματα, ενώ στις δύο ηλικιακά μεγαλύτερες ομάδες είναι περισσότερα τα αρνητικά δείγματα.

Ιατρικό ιστορικό:

Οι περισσότεροι συμμετέχοντες δεν πάσχουν από κάποιο υποκείμενο νόσημα, ενώ μόλις 185 από τα 1482 δείγματα συνοδεύονται με στοιχεία πάθησης από υποκείμενο νόσημα του αναπνευστικού συστήματος.

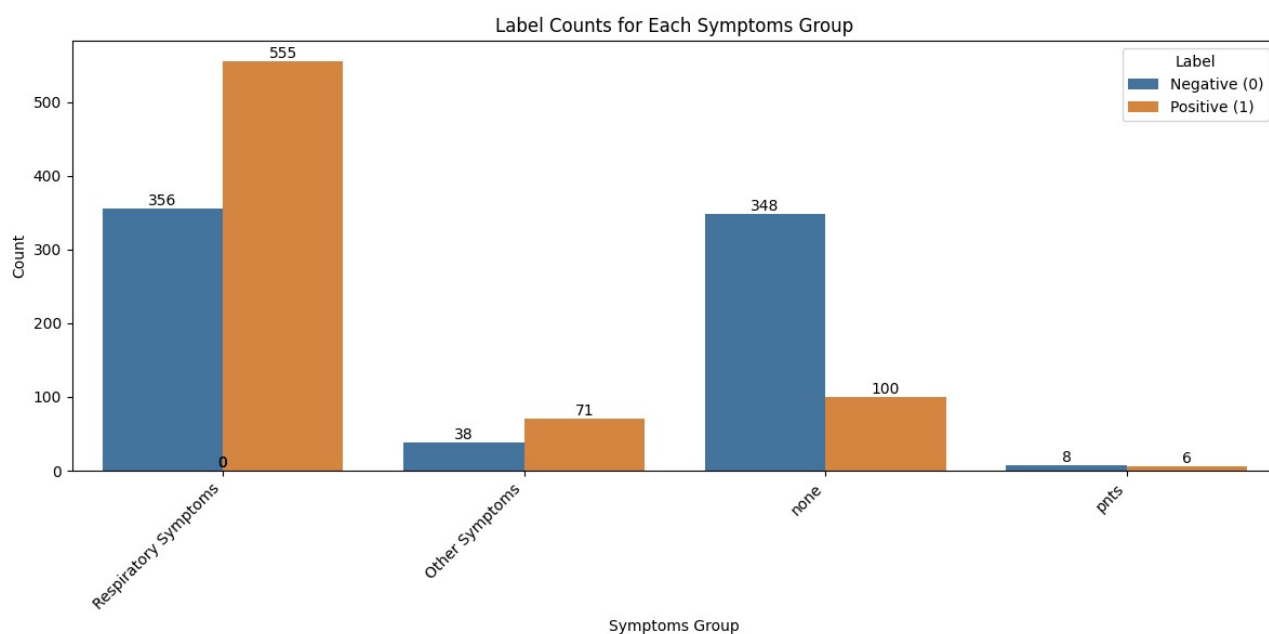


Εικόνα 6.1.3: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Ιατρικού ιστορικού.

Παρατηρώντας το παραπάνω διάγραμμα, διαπιστώνουμε ότι υπάρχει ισορροπία ανάμεσα στα θετικά και τα αρνητικά δείγματα νόσησης από COVID-19 στην κάθε υποκατηγορία.

Συμπτώματα:

Οι περισσότεροι συμμετέχοντες παρουσιάζουν συμπτώματα και ειδικότερα τουλάχιστον ένα σύμπτωμα που σχετίζεται με το αναπνευστικό σύστημα.

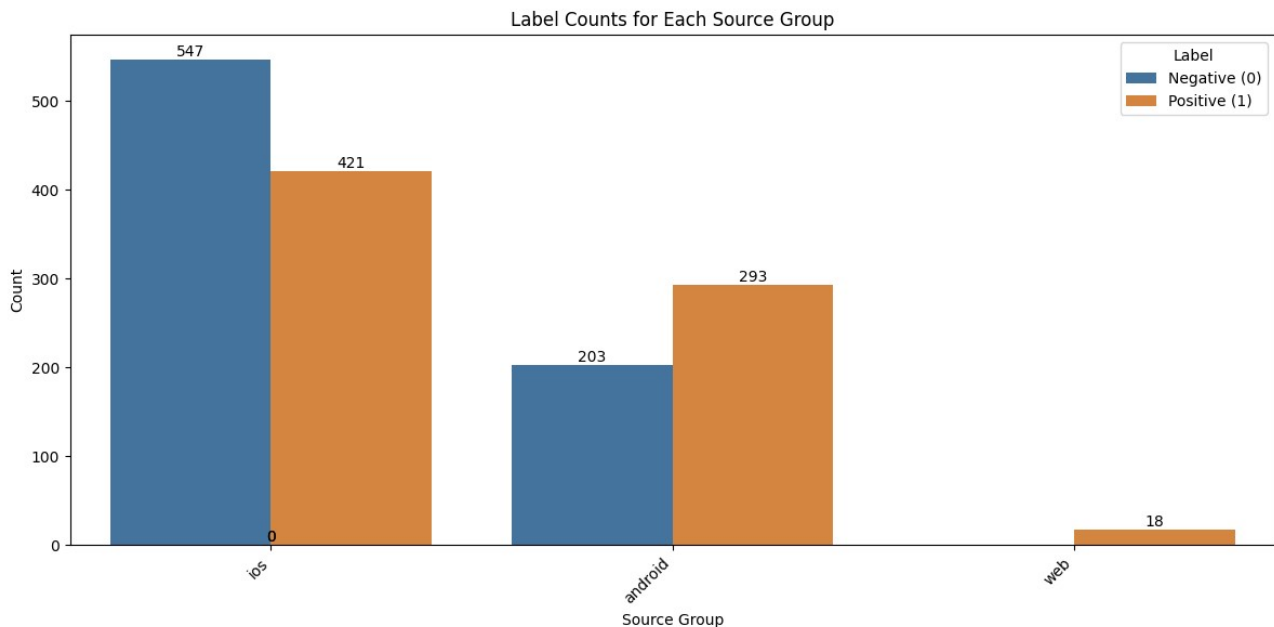


Εικόνα 6.1.4: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Συμπτωμάτων.

Παρατηρούμε ότι όσοι δεν εμφανίζουν συμπτώματα είναι ως επί το πλείστον αρνητικοί, ενώ από όσους εμφανίζουν συμπτώματα είτε στο αναπνευστικό σύστημα είτε άλλου είδους οι περισσότεροι είναι θετικοί.

Πηγή συλλογής δεδομένων:

Η πλειονότητα των δειγμάτων προέρχεται από λειτουργικό σύστημα iOS, ενώ η διαδικτυακή πλατφόρμα (web) παρουσιάζει εμφανώς μειωμένη συμμετοχή στη συλλογή δειγμάτων, με μόλις 18 σήματα να προέρχονται από εκεί.

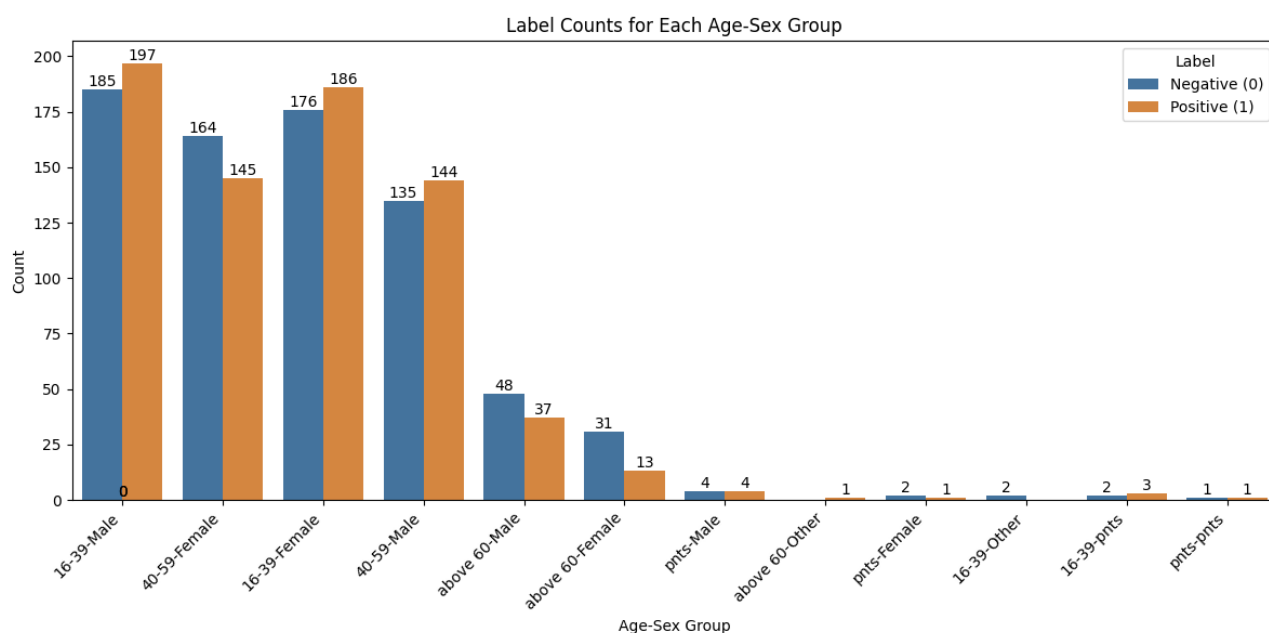


Εικόνα 6.1.5: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Πηγής Συλλογής Δεδομένων.

Παρατηρούμε ότι τα περισσότερα δείγματα που συλλέχθηκαν από συσκευή iOS είναι αρνητικά στην COVID-19, ενώ τα περισσότερα δείγματα που συλλέχθηκαν από συσκευή Android είναι θετικά. Επίσης, όλα τα δείγματα που συλλέχθηκαν από διαδικτυακή πλατφόρμα είναι θετικά.

Ηλικία σε συνδυασμό με φύλο:

Συνολικά στην νεαρότερη ηλικιακή ομάδα και στην ηλικιακή ομάδα άνω των 60, οι άντρες είναι περισσότεροι, ενώ στην μεσαία ηλικιακή ομάδα περισσότερες είναι οι γυναίκες.

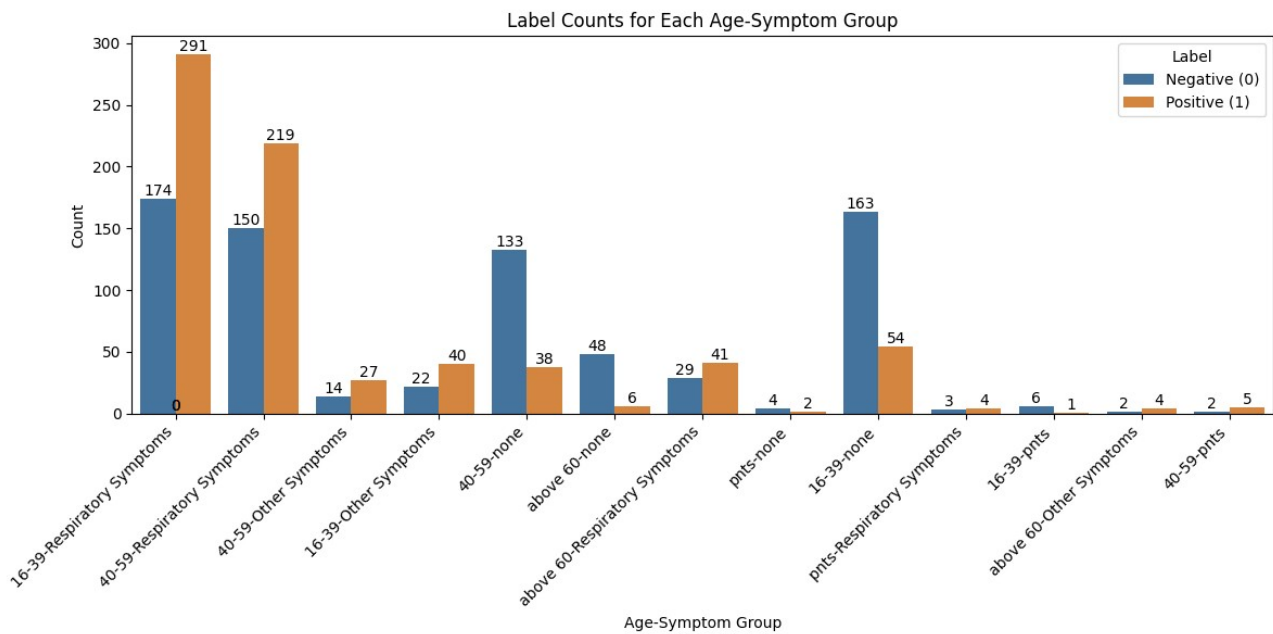


Εικόνα 6.1.6: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Ηλικίας σε συνδυασμό με Φύλο.

Παρατηρούμε ότι στην νεαρότερη ηλικιακή ομάδα και στους άνδρες και στις γυναίκες οι περισσότεροι είναι θετικοί, στην ηλικιακή ομάδα άνω των 60 ετών οι περισσότεροι συμμετέχοντες και των δύο φύλων είναι αρνητικοί, ενώ στην μεσαία ηλικιακή ομάδα 40-59 ετών έχουμε περισσότερους θετικούς άντρες και περισσότερες αρνητικές γυναίκες.

Ηλικία σε συνδυασμό με συμπτώματα:

Οι πολυπληθέστερες υποκατηγορίες είναι αυτές των ατόμων νεαρότερης ηλικίας με συμπτώματα του αναπνευστικού συστήματος και των ατόμων μέσης ηλικίας επίσης με συμπτώματα του αναπνευστικού συστήματος. Ακολουθούν οι κατηγορίες των ατόμων νεαρότερης ηλικίας χωρίς συμπτώματα και των ατόμων μέσης ηλικίας χωρίς συμπτώματα.

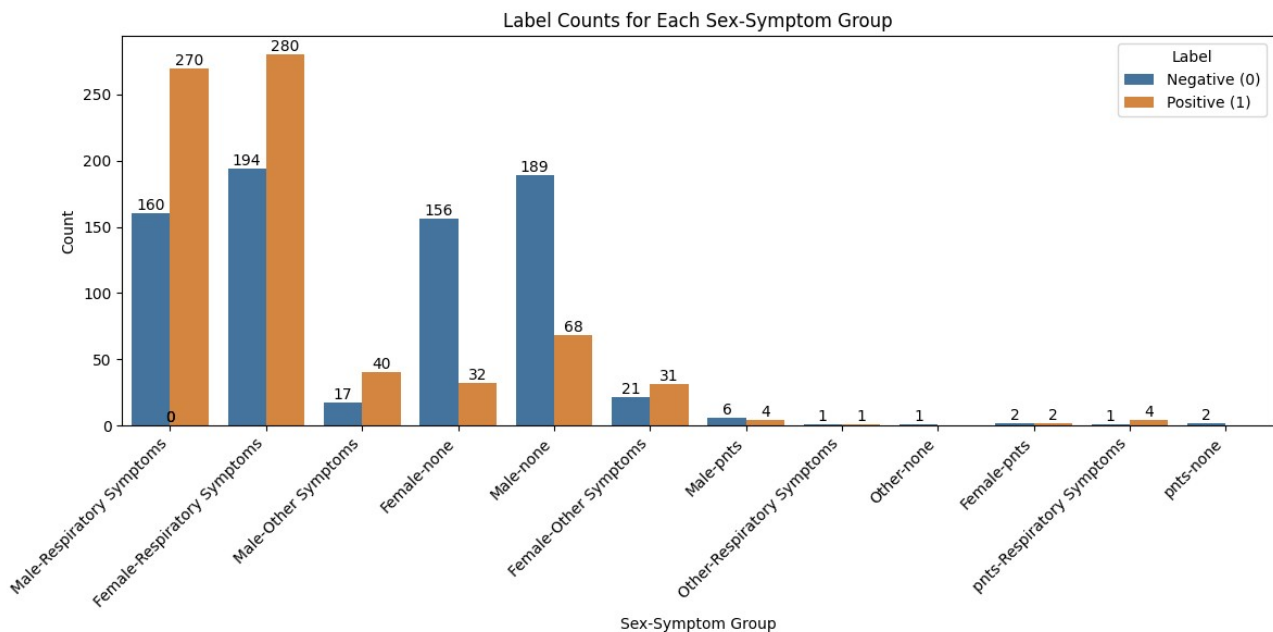


Εικόνα 6.1.7: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Ηλικίας σε συνδυασμό με Συμπτώματα.

Παρατηρούμε ιδιαίτερα έντονες διακυμάνσεις ανάμεσα στα θετικά και τα αρνητικά δείγματα για κάθε υποκατηγορία, με τις πολυπληθέστερες ομάδες να εμφανίζουν περισσότερα θετικά δείγματα όταν υπάρχουν συμπτώματα του αναπνευστικού και αρνητικά δείγματα όταν δεν υπάρχουν καθόλου συμπτώματα.

Φύλο σε συνδυασμό με συμπτώματα:

Υπάρχει ισορροπία στους άνδρες και τις γυναίκες ανά διαφορετική κατηγορία κατάστασης συμπτωμάτων. Μεγαλύτερες κατηγορίες αποτελούν οι γυναίκες με συμπτώματα του αναπνευστικού συστήματος και οι άνδρες επίσης με συμπτώματα του αναπνευστικού συστήματος.

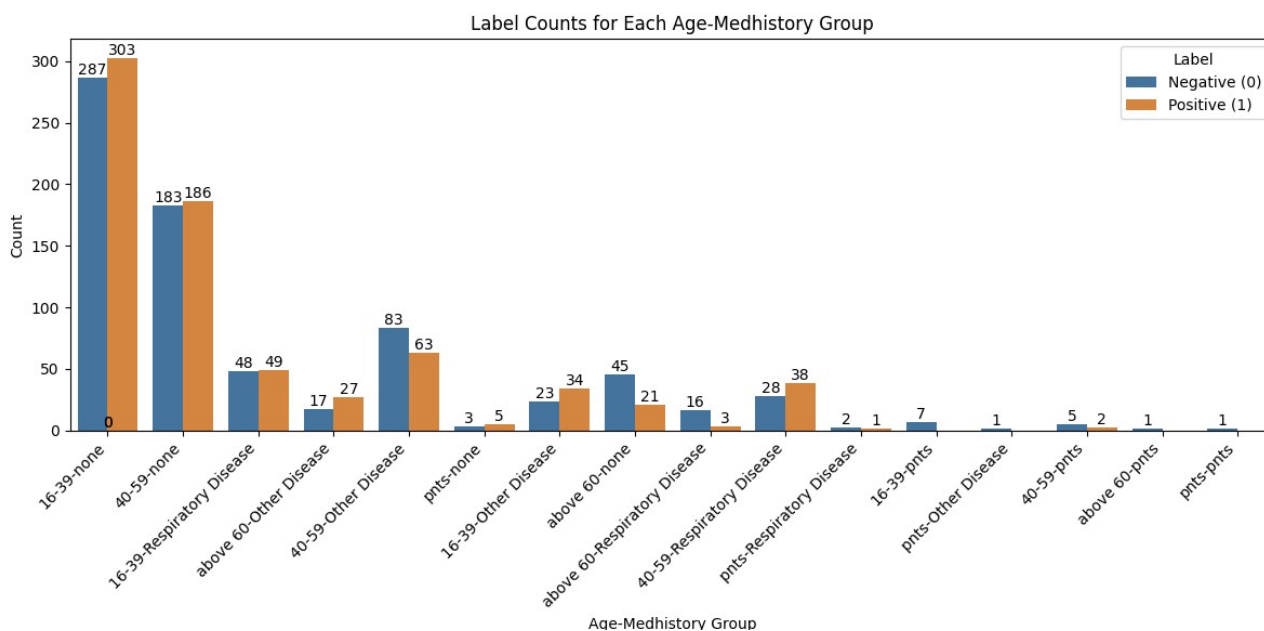


Εικόνα 6.1.8: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Φύλου σε συνδυασμό με Συμπτώματα.

Παρατηρούμε ότι στις υποκατηγορίες τα θετικά και αρνητικά δείγματα εμφανίζουν αποκλίσεις μεταξύ τους, όμως είναι ισορροπημένα για τα δύο φύλα στην ίδια κατηγορία συμπτωμάτων.

Ηλικία σε συνδυασμό με ιατρικό ιστορικό:

Και στις τρεις ηλικιακές ομάδες οι περισσότεροι συμμετέχοντες δεν πάσχουν από κάποιο υποκείμενο νόσημα. Από αυτούς που πάσχουν όμως από υποκείμενα νοσήματα, στη νεαρότερη ηλικιακή ομάδα, οι περισσότεροι έχουν κάποιο νόσημα του αναπνευστικού συστήματος, ενώ στις άλλες δύο ηλικιακές ομάδες οι περισσότεροι έχουν υποκείμενα νοσήματα που δε σχετίζονται με το αναπνευστικό.

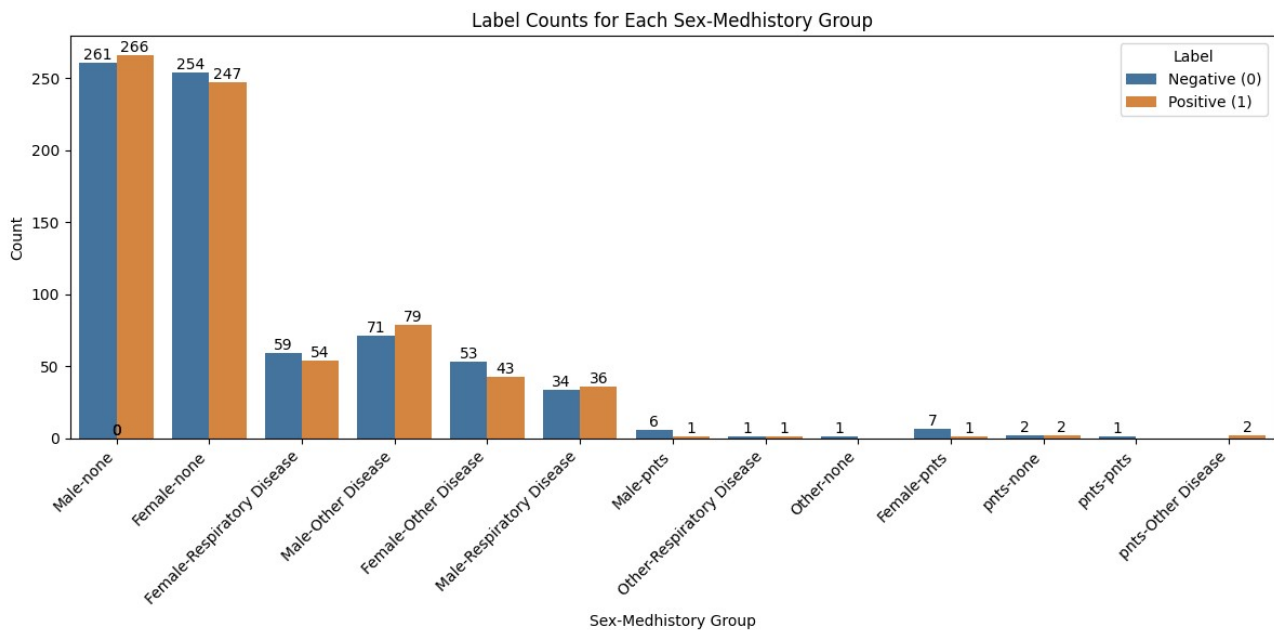


Εικόνα 6.1.9: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Ηλικίας σε συνδυασμό με Ιατρικό ιστορικό.

Παρατηρούνται γενικά μικρότερες αποκλίσεις μεταξύ θετικών και αρνητικών δειγμάτων σε κάθε υποκατηγορία.

Φύλο σε συνδυασμό με ιατρικό ιστορικό:

Τόσο οι περισσότεροι άντρες όσο και οι περισσότερες γυναίκες δεν πάσχουν από κάποιο υποκείμενο νόσημα. Από τους υπόλοιπους, οι περισσότεροι πάσχοντες από νοσήματα που σχετίζονται με το αναπνευστικό είναι γυναίκες, ενώ οι περισσότεροι πάσχοντες από άλλα νοσήματα είναι άνδρες.



Εικόνα 6.1.10: Διάγραμμα απεικόνισης θετικών και αρνητικών δειγμάτων ανά υποκατηγορία Φύλου σε συνδυασμό με Ιατρικό ιστορικό.

Παρατηρείται και εδώ σχετική ισορροπία μεταξύ θετικών και αρνητικών δειγμάτων στις περισσότερες υποκατηγορίες.

Από την κατανομή των δεδομένων σε όλες τις παραπάνω κατηγορίες, διαπιστώνουμε ότι η υποκατηγορία δειγμάτων για τα οποία δεν προσφέρεται κάποια απάντηση (“pnts”) έχει πολύ μικρό πλήθος και επιπλέον δεν παρέχει καμία πληροφορία, γι’ αυτό και δεν μπορεί να αξιολογηθεί ως προς τις διακρίσεις. Επομένως, στην εξαγωγή των αποτελεσμάτων εξαιρούνται οι υποκατηγορίες “pnts” ή στις συνδυαστικές κατηγορίες όλες οι υποκατηγορίες που περιέχεται το “pnts” ως απάντηση. Επιπλέον, στην κατηγορία του Φύλου, η επιλογή άλλου φύλου (“Other”), αποτελείται από 3 δείγματα και επομένως η υποκατηγορία αυτή επίσης εξαιρείται από τις μετρήσεις για την αμεροληψία, καθώς αλλοιώνει την ερμηνεία των αποτελεσμάτων λόγω του πολύ μικρού της πλήθους.

6.1.4 Εξαγωγή Χαρακτηριστικών

Τα δεδομένα που χρησιμοποιήθηκαν στην εργασία μας έχουν τη μορφή αρχείων ήχου wav. Οι διαδρομές για τα αρχεία αυτά μαζί με τα ευαίσθητα χαρακτηριστικά που αντιστοιχούν σε κάθε σήμα, αποτελούν ένα πλαίσιο δεδομένων (DataFrame) ως είσοδο προς επεξεργασία. Για να εισάγουμε τα ηχητικά σήματα στα μοντέλα που θα εκπαιδεύσουμε, πρέπει πρώτα να εξάγουμε τα χαρακτηριστικά τους. Ανάλογα με τον τύπο του μοντέλου που χρησιμοποιούμε κάθε φορά, δηλαδή στην περίπτωση μας τον Απλό Ταξινομητή Τυχαίου Δάσους και το Συνελικτικό Νευρωνικό Δίκτυο, εξάγουμε Χαρακτηριστικά Επεξεργασίας Ψηφιακού Σήματος (DSP Features) και Φασματογραφήματα (Spectrograms) αντίστοιχα.

Για την εξαγωγή των χαρακτηριστικών επεξεργασίας ψηφιακού σήματος, σε περιβάλλον γλώσσας προγραμματισμού python εισάγουμε το αρχείο .csv που προαναφέραμε και στη συνέχεια ορίζουμε οχτώ συχνοτικές ζώνες για την ανάλυση του σήματος και μία λίστα με έντεκα κατηγορίες χαρακτηριστικών που αναλύθηκαν στο θεωρητικό μέρος της παρούσας εργασίας (ενότητα 2.2.5) και συνοψίζονται ως εξής:

- Ανίχνευση Αιχμών Ενεργειακού Φακέλου (Energy Envelope Peak Detection – EEDP)
- Ρυθμός Διαβάσεων από το Μηδέν (Zero Crossing Rate – ZCR)

- Μέση Ισχύς Τετραγωνικής Πίζας (Root Mean Square Power – RMSP)
- Δυναμικά Χαρακτηριστικά (Dynamic Features – DF)
- Φασματικά Χαρακτηριστικά (Spectral Features)
- Φασματική Επιπεδότητα και Βραχύχρονη Φασματική Απόκλιση (Spectral Flatness and Short-Time Spectral Deviation – SF_SSTD)
- Βραχύχρονη Φασματική Ένταση και Φασματική Πυκνότητα (Short-Time Spectral Loudness & Spectral Density – SSL_SD)
- Συντελεστές Λογαριθμικού Φάσματος Κλίμακας Mel (Mel-Frequency Cepstral Coefficients – MFCC)
- Συχνότητα Κέντρου Βάρους (Centroid Frequency – CF)
- Μήκος Σήματος (Length – LGTH)
- Πυκνότητα Φασματικής Ισχύος (Power Spectral Density – PSD)

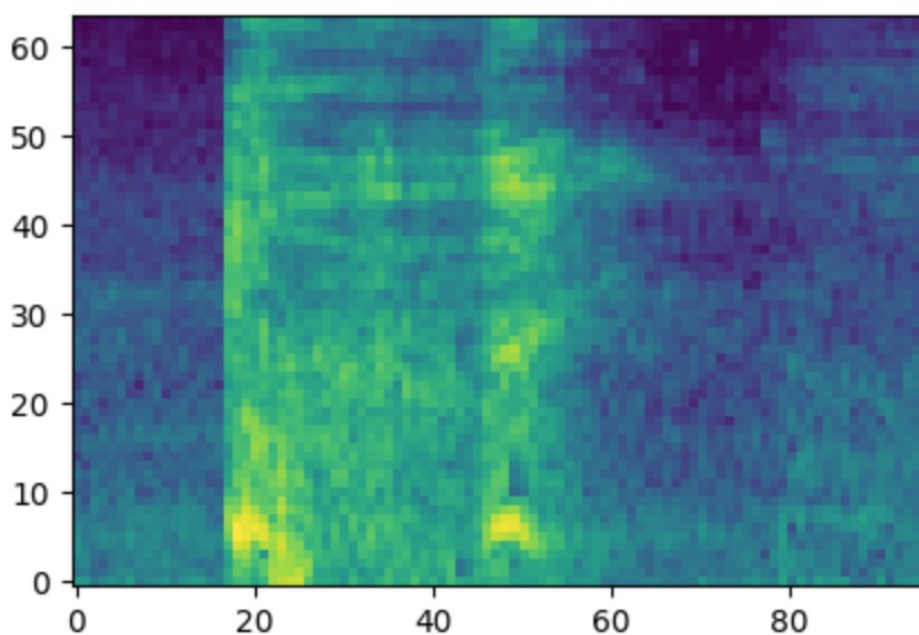
Ακολουθώντας, για κάθε αρχείο ήχου (βήχα) κατασκευάζεται το πλήρες μονοπάτι του, μέσα από το οποίο διαβάζεται το σήμα, ορίζεται συχνότητα δειγματοληψίας στα 48kHz και γίνεται κανονικοποίηση του σήματος, ώστε οι τιμές να κυμαίνονται στο ίδιο εύρος και εφαρμόζεται η αντίστοιχη μέθοδος από μία κλάση που δημιουργήθηκε με συναρτήσεις για την εξαγωγή κάθε χαρακτηριστικού. Αν η έξοδος κάποιας επιμέρους μεθόδου είναι πίνακας, διατηρείται μόνο η πρώτη τιμή του. Τελικά λαμβάνουμε για κάθε αρχείο ήχου ένα διάνυσμα με 68 τιμές χαρακτηριστικών που το περιγράφει. Τα διανύσματα αυτά αποθηκεύονται στο αρχικό πλαίσιο δεδομένων μαζί με τα υπόλοιπα στοιχεία κάθε δείγματος και το αρχείο είναι έτοιμο να εισαχθεί στο μοντέλο.

```
Row 0 (shape: (68,)):
[ 53.    51.    46.    52.    49.    53.
  48.    51.    49.    51.    48.    52.
  49.    55.    52.    47.    49.    52.
  49.    0.053  0.097  0.008  3471.384 70193.
 3624.476  1.273  3.819 6473701.42  0.001  0.
 -0.005  0.001 -293.561 128.155  -4.102 17.594
 -4.021 14.736  1.77  7.028  6.346 -5.688
  1.259  0.521  3.477 116.47 23.576 30.577
 19.382 22.532 11.364 14.348  8.521 10.81
 10.643  8.077  8.339  8.008 10.289  6.316
  0.122  0.091  0.028  0.027  0.04  0.005
  0.002  0.001]
```

Εικόνα 6.1.11: Απεικόνιση των 68 τιμών χαρακτηριστικών επεξεργασίας ψηφιακού σήματος για το πρώτο κατά σειρά ηχητικό σήμα των δεδομένων μας.

Προκειμένου γενικά να εισάγουμε τα ηχητικά σήματα σε ένα συνελκτικό νευρωνικό δίκτυο, πρέπει προηγουμένως να τα μετατρέψουμε σε Mel-Φασματογραφήματα ώστε στη συνέχεια το μοντέλο να είναι σε θέση να εξάγει μόνο του τα χαρακτηριστικά. Σε κάθε ηχητικό σήμα προσβάσιμο από το αντίστοιχο μονοπάτι του, ορίζεται συχνότητα δειγματοληψίας 16kHz σύμφωνα με τις απαιτήσεις του προεκπαιδευμένου μοντέλου VGGish (παραπομπή) στο οποίο θα εισαχθεί και γίνεται κανονικοποίηση για να έχουμε ίδιο εύρος τιμών. Στη συνέχεια, χρησιμοποιείται συνάρτηση μετατροπής του ήχου σε Mel-φασματογράφημα η οποία χωρίζει το σήμα σε μικρά χρονικά παράθυρα, εφαρμόζει Γρήγορο Μετασχηματισμό Fourier (FFT) και μετατρέπει τις συχνότητες σε κλίμακα Mel, σύμφωνα και πάλι με τις παραμέτρους που καθορίζονται από τις απαιτήσεις του προεκπαιδευμένου μοντέλου VGGish και οι οποίες έχουν γνωστοποιηθεί από τους δημιουργούς του. Πιο αναλυτικά, το μέγεθος των χρονικών παραθύρων στα οποία εφαρμόζεται ο Γρήγορος

Μετασχηματισμός Fourier ορίζεται ως 0,025 δευτερόλεπτα, το βήμα μεταξύ των χρονικών παραθύρων ορίζεται ως 0,010 δευτερόλεπτα, η ελάχιστη και η μέγιστη συχνότητα στα 125Hz και 7,5kHz αντίστοιχα και το πλήθος των ζωνών συχνοτήτων σε κλίμακα Mel ορίζεται ως 64. Η κλίμακα των Mel-φασματογραφημάτων μετατρέπεται από ισχύ σε ντεσιμπέλ (dB) και έπειτα το φασματογράφημα κανονικοποιείται σε εύρος [0, 255] που μετατρέπεται σε τύπο δεδομένων uint8, κατάλληλο για επεξεργασία ψηφιακής εικόνας. Οι άξονες του Mel-φασματογραφήματος αντιστρέφονται, όπως απαιτεί το μοντέλο στο οποίο θα εισαχθούν. Το τελικό φασματογράφημα μπορεί να θεωρηθεί ως μία εικόνα που απεικονίζει το ηχητικό σήμα, με οριζόντιο άξονα x τον χρόνο και κάθετο άξονα y τη συχνότητα. Τα χρώματα αντιπροσωπεύουν την ένταση σε ντεσιμπέλ σε κάθε σημείο της εικόνας, με χαμηλότερη ένταση το σκούρο μπλε και υψηλότερη το κίτρινο. Τα Mel-φασματογραφήματα αποθηκεύονται στο αρχικό πλαίσιο δεδομένων μαζί με τα υπόλοιπα στοιχεία κάθε δείγματος για να εισαχθεί στο μοντέλο. Τέλος, εφαρμόζονται δύο συναρτήσεις που προσαρμόζουν τα δεδομένα στις ίδιες διαστάσεις. Συγκεκριμένα, η πρώτη, μειώνει το μέγεθός τους με υποδειγματοληψία σε κάθε 48 τμήματα που ορίζονται με μήκος 96 χρονικά βήματα και προσθέτει μία επιπλέον διάσταση στην έξοδο ώστε να συμβαδίζει με τις διαστάσεις εισόδου του νευρωνικού δικτύου. Η δεύτερη, προσαρμόζει τα δεδομένα, κόβοντας ή αναδιπλώνοντάς τα, σε ίδιο συγκεκριμένο μήκος που καθορίζεται από τη μέση τιμή μήκους του 90% του συνόλου, προς αποφυγή των ακραίων τιμών. Ενδεικτικά, στην Εικόνα 6.1.12 απεικονίζεται ένα φασματογράφημα που προκύπτει από την επεξεργασία των δεδομένων με την παραπάνω διαδικασία.



Εικόνα 6.1.12: Mel-φασματογράφημα της μετατροπής των δεδομένων εισόδου στο μοντέλο μας.

Πριν την εισαγωγή του στο μοντέλο μηχανικής μάθησης, το σύνολο δεδομένων διαμοιράζεται, όπως προβλέπεται, σε τρία υποσύνολα: το σύνολο εκπαίδευσης (train set) που καταλαμβάνει το 70% του αρχικού συνόλου, το σύνολο επικύρωσης (validation set) που αποτελείται από ποσοστό 10% του αρχικού και το σύνολο ελέγχου (test set) που αποτελείται από το υπόλοιπο 20% του συνόλου δεδομένων. Ορίζεται σταθερός διαμοιρασμός, προκειμένου να λαμβάνουμε τα ίδια υποσύνολα κάθε φορά, εκτός αν επιλέξουμε το αντίθετο (όπως για παράδειγμα για να εξάγουμε ως αποτελέσματα τον μέσο όρο από 10 διαφορετικές εκτελέσεις του μοντέλου). Η διαστρωμάτωση (stratification) κατά τον διαχωρισμό γίνεται με βάση τις ετικέτες, ώστε να έχουμε ισορροπία στην αναλογία θετικών και αρνητικών περιπτώσεων και στα τρία υποσύνολα και λαμβάνοντας υπόψη μόνο τις μοναδικές ταυτότητες χρήστη (Uid). Σε περίπτωση που υπάρχουν πολλαπλά δείγματα από

τον ίδιο χρήστη, εισάγονται όλα στο ίδιο υποσύνολο. Έτσι διασφαλίζουμε ομοιόμορφη κατανομή των ετικετών στα σύνολα εκπαίδευσης, επικύρωσης και ελέγχου και επομένως καλύτερες προβλέψεις.

6.2 Ανάπτυξη Μοντέλου Μηχανικής Μάθησης

6.2.1 Εκπαίδευση Ταξινομητή Τυχαίου Δάσους

Κατά την ανάπτυξη του μοντέλου μηχανικής μάθησης για την ταξινόμηση των ηχητικών σημάτων, χρησιμοποιήσαμε τα χαρακτηριστικά επεξεργασίας ψηφιακού σήματος που εξήχθησαν όπως περιγράψαμε προηγουμένως. Παρότι η εκπαίδευση του ταξινομητή τυχαίου δάσους δεν απαιτεί υποσύνολο επικύρωσης, στον διαχωρισμό του συνόλου δεδομένων σε υποσύνολα εκπαίδευσης, επικύρωσης και ελέγχου, διατηρήθηκε αντίστοιχη αναλογία σε όλα τα μοντέλα του πειράματος, ακολουθώντας την μεθοδολογία που έχουμε προαναφέρει. Η εκπαίδευση του μοντέλου πραγματοποιήθηκε με τη χρήση ενός Απλού Ταξινομητή Τυχαίου Δάσους (Random Forest), ο οποίος εκπαιδεύεται με τα χαρακτηριστικά των δεδομένων του υποσυνόλου εκπαίδευσης και τις αντίστοιχες ετικέτες τους. Ο ταξινομητής τυχαίου δάσους αποτελεί ένα μοντέλο το οποίο κατασκευάζει πολλά δέντρα αποφάσεων, καθένα από τα οποία εκπαιδεύεται σε ένα τυχαία επιλεγμένο υποσύνολο των δεδομένων εκπαίδευσης. Όταν όλα τα δέντρα έχουν εκπαιδευτεί, η τελική πρόβλεψη του τυχαίου δάσους στο σύνολο ελέγχου προκύπτει από την πλειοψηφία των προβλέψεων του συνόλου των δέντρων. Μετά από την εξαγωγή των προβλέψεων του μοντέλου, για την αξιολόγηση της απόδοσης του, μέσω της σύγκρισης των προβλεπόμενων τιμών με τις πραγματικές ετικέτες των δεδομένων, υπολογίζονται: Ο Πίνακας Σύγχυσης (Confusion Matrix), η Ακρίβεια (Accuracy), η Ευαισθησία (Sensitivity) και η Ειδικευση (Specificity) του μοντέλου. Επίσης τα μεγέθη αυτά υπολογίζονται και ξεχωριστά για κάθε υποομάδα ενός ευαίσθητου χαρακτηριστικού που επιλέγουμε κάθε φορά, προκειμένου να χρησιμοποιηθούν σε συναρτήσεις, ώστε να εξάγουμε συμπεράσματα για την αμεροληψία του μοντέλου. Η διαδικασία αυτή επαναλαμβάνεται και από τις επαναλήψεις λαμβάνουμε τον μέσο όρο κάθε μετρικής που έχουμε επιλέξει, προκειμένου να έχουμε πιο σταθερά αποτελέσματα.

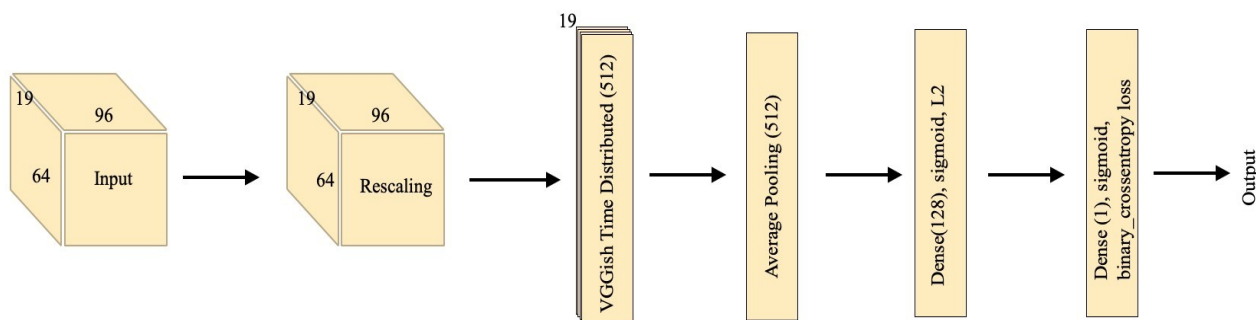
6.3 Ανάπτυξη Μοντέλου Βαθιάς Μάθησης

Για περαιτέρω διερεύνηση της αμεροληψίας και των τρόπων ενίσχυσης της στην δυαδική ταξινόμηση, αναπτύχθηκε μοντέλο βαθιάς μάθησης, καθώς λόγω της μεγαλύτερης πολυπλοκότητάς του προσφέρει καλύτερη ανάλυση των δεδομένων και ευκολότερη διαχείριση του όγκου τους. Το μοντέλο που δημιουργήθηκε, είναι ένα συνελκτικό νευρωνικό δίκτυο, το οποίο εξάγει πιο σύνθετα χαρακτηριστικά, λαμβάνοντας υπόψη και τη χρονική πληροφορία, γεγονός που καθιστά δυνατή την ανίχνευση πιο πολύπλοκων μοτίβων στα δεδομένα, σε σχέση με έναν απλό ταξινομητή.

6.3.1 Εκπαίδευση Συνελκτικού Νευρωνικού Δικτύου

Στην κατασκευή του μοντέλου βαθιάς μάθησης για δυαδική ταξινόμηση χρησιμοποιούμε το προεκπαιδευμένο συνελκτικό νευρωνικό δίκτυο VGGish [37], του οποίου η αρχιτεκτονική αναλύθηκε στο θεωρητικό μέρος της εργασίας και απεικονίζεται στην Εικόνα 6.3.1, σε συνδυασμό με ένα πολύ μικρότερο πλήρως συνδεδεμένο δίκτυο. Τα δεδομένα του συνόλου εκπαίδευσης, μετά την προσαρμογή τους στις επιθυμητές διαστάσεις, όπως περιγράφηκε προηγουμένως, εισάγονται και καθορίζουν το σχήμα της εισόδου που είναι (None, 19, 96, 64, 1), δηλαδή, χρονικές ακολουθίες αποτελούμενες από 19 χρονικά παράθυρα, όπου κάθε βήμα είναι μία εικόνα μεγέθους 96x64 που απεικονίζεται σε 1 κανάλι. Στη συνέχεια, μετατρέπονται σε δεδομένα μορφής float32 και οι τιμές τους κανονικοποιούνται στο διάστημα [0,1], χωρίς να αλλάζουν οι διαστάσεις. Από το VGGish

αφαιρούμε το τελικό πλήρως συνδεδεμένο επίπεδο, διατηρούμε τα προεκπαιδευμένα βάρη του, και το χρησιμοποιούμε για την εξαγωγή χαρακτηριστικών από τα δεδομένα που εισάγουμε. Η διαδικασία αυτή εφαρμόζεται σε καθένα από τα χρονικά παράθυρα του κάθε φασματογραφήματος, ώστε τα χαρακτηριστικά να εξάγονται για κάθε χρονικό παράθυρο ξεχωριστά (TimeDistributed), μετατρέποντας κάθε εικόνα διαστάσεων 96x64 σε διάνυσμα 512 χαρακτηριστικών. Μετά από αυτό το βήμα, τα δεδομένα έχουν διαστάσεις (None, 19, 512). Έπειτα, υπολογίζεται ο μέσος όρος της τιμής των χαρακτηριστικών για το σύνολο της χρονικής διάστασης, δηλαδή τα 19 χρονικά παράθυρα, με τα δεδομένα να λαμβάνουν τη μορφή (None, 512). Τα διανύσματα των 512 χαρακτηριστικών κάθε σήματος, διέρχονται από ένα πλήρες συνδεδεμένο επίπεδο με 128 νευρώνες. Για την αποφυγή της υπερεκπαίδευσης σε αυτό το επίπεδο, εφαρμόζεται η σιγμοειδής (sigmoid) συνάρτηση ενεργοποίησης, που περιορίζει τις τιμές των νευρώνων σε διάστημα [0,1], ομαλοποιώντας τις αποκλίσεις. Εφαρμόζεται επίσης, κανονικοποίηση L2, η οποία ενθαρρύνει τα μικρότερα βάρη, τιμωρώντας την συνάρτηση απώλειας για μεγάλα βάρη. Η είσοδος που προκύπτει, με διάσταση (None, 128) διέρχεται από το τελικό πλήρως συνδεδεμένο επίπεδο εξόδου, όπου γίνεται η δυαδική ταξινόμηση, με επίσης σιγμοειδής συνάρτηση ενεργοποίησης και έναν νευρώνα στην έξοδο που δίνει αποτέλεσμα (None, 1). Ως βελτιστοποιητής, ορίζεται ο Adam με χαμηλό ρυθμό εκμάθησης ($5e-5$) για πιο αργή αλλά σταθερή εκπαίδευση και ως συνάρτηση απώλειας ορίζεται η συνάρτηση δυαδικής διασταύρωσης εντροπίας (binary_crossentropy), κατάλληλη για τη δυαδική ταξινόμηση.



Εικόνα 6.3.1: Απεικόνιση των επιπέδων και του σχήματος εξόδου κάθε επιπέδου του μοντέλου που αναπτύξαμε.

Μετά τον ορισμό του μοντέλου, καλείται η συνάρτηση για τη δημιουργία του. Εκπαιδεύουμε το μοντέλο με το σύνολο εκπαίδευσης, αξιολογώντας την απόδοσή του στις προβλέψεις για το σύνολο επικύρωσης. Η εκπαίδευση επαναλαμβάνεται είτε μέχρι την συμπλήρωση του μέγιστου αριθμού εποχών εκπαίδευσης (epochs), είτε μέχρι η απόδοσή του να μην βελτιώνεται άλλο, έτσι ώστε να αποφευχθεί και η υπερεκπαίδευση. Μετά την εκπαίδευση του μοντέλου, εισάγουμε το σύνολο ελέγχου για να λάβουμε τις τελικές προβλέψεις. Ορίζουμε τις μετρικές με τις οποίες θέλουμε να αξιολογήσουμε την απόδοση του μοντέλου και οι οποίες είναι: ο Πίνακας Σύγχυσης (Confusion Matrix), η Ακρίβεια (Accuracy), η F1-Βαθμολογία (F1-Score), η Ισορροπημένη Ακρίβεια (Balanced Accuracy), η Περιοχή Κάτω από την Καμπύλη ROC (ROC-AUC), η Ευαισθησία (Sensitivity) και η Ειδικευση (Specificity). Για τον υπολογισμό των μετρικών αυτών χρησιμοποιείται το σύνηθες κατώφλι στη δυαδική ταξινόμηση 0,5 και επομένως οι προβλέψεις τιμής κάτω του 0,5 θεωρούνται αρνητικές και οι προβλέψεις άνω του 0,5 θετικές. Το πείραμα επαναλαμβάνεται με παραλλαγές στον διαμοιρασμό των υποσυνόλων δεδομένων και για την εξαγωγή των αποτελεσμάτων, υπολογίζεται ο μέσος όρος των τιμών που προκύπτουν από τις επαναλήψεις, για πιο αξιόπιστα συμπεράσματα.

6.4 Αμεροληψία και Αντιμετώπιση των Διακρίσεων

Μετά τη δημιουργία και την εκπαίδευση των μοντέλων ταξινόμησης, αναζητήσαμε εργαλεία ελέγχου της αμεροληψίας, ώστε αφού αξιολογήσουμε τον βαθμό αμεροληψίας πριν και μετά την αντιμετώπιση πιθανών διακρίσεων καθενός μοντέλου, να διαπιστώσουμε, συγκρίνοντάς τα, την αποτελεσματικότητά τους.

6.4.1 Υλοποίηση Συναρτήσεων για τον υπολογισμό της Αμεροληψίας

Για τον έλεγχο της αμεροληψίας των προβλέψεων των μοντέλων που κατασκευάσαμε, δημιουργήσαμε συναρτήσεις, οι οποίες χρησιμοποιούν τον ρυθμό αληθώς θετικών δειγμάτων (TPR), τον ρυθμό ψευδώς θετικών δειγμάτων (FPR), τον ρυθμό θετικών προβλέψεων (SR) και την ακρίβεια (Accuracy) που υπολογίσαμε βάσει των προβλέψεων κάθε μοντέλου μετά την εκπαίδευσή του. Έτσι, για τον υπολογισμό καθεμιάς από τις μετρικές αμεροληψίας, οι οποίες αναλύθηκαν στο θεωρητικό μέρος της παρούσας εργασίας, υλοποιήσαμε συναρτήσεις που εκφράζουν τα εξής μεγέθη:

- **Ισότητα Ακρίβειας (Accuracy Parity)**
 - **Διαφορά Ισότητας Ακρίβειας (Accuracy Parity Difference):** υπολογίζεται ως η διαφορά μεταξύ της μέγιστης και της ελάχιστης τιμής ακρίβειας ανάμεσα στις διαφορετικές υποομάδες που χαρακτηρίζονται από ένα ευαίσθητο χαρακτηριστικό. Η ιδανική τιμή της διαφοράς ισότητας ακρίβειας είναι 0.
 - **Αναλογία Ισότητας Ακρίβειας (Accuracy Parity Ratio):** υπολογίζεται ως η αναλογία της ελάχιστης προς την μέγιστη τιμή ακρίβειας ανάμεσα στις διαφορετικές υποομάδες που χαρακτηρίζονται από ένα ευαίσθητο χαρακτηριστικό. Η ιδανική τιμή της αναλογίας ισότητας ακρίβειας είναι 1.
- **Ισότητα πιθανότητας (Equalized Odds)**
 - **Διαφορά Ισότητας Πιθανότητας (Equalized Odds Difference):** υπολογίζεται ως η μέγιστη διαφορά που προκύπτει είτε από την διαφορά μεταξύ της μέγιστης και της ελάχιστης τιμής ρυθμού αληθώς θετικών δειγμάτων, είτε από την διαφορά μεταξύ της μέγιστης και της ελάχιστης τιμής ρυθμού ψευδώς θετικών δειγμάτων, μεταξύ των διαφορετικών υποομάδων μίας κατηγορίας. Η ιδανική τιμή της διαφοράς ισότητας πιθανότητας είναι 0.
 - **Αναλογία Ισότητας Πιθανότητας (Equalized Odds Ration):** υπολογίζεται ως η ελάχιστη αναλογία που προκύπτει είτε από την αναλογία της ελάχιστης προς την μέγιστη τιμή ρυθμού αληθώς θετικών δειγμάτων, είτε από την αναλογία της ελάχιστης προς την μέγιστη τιμή ρυθμού ψευδώς θετικών δειγμάτων, μεταξύ των διαφορετικών υποομάδων μίας κατηγορίας. Η ιδανική τιμή της αναλογίας ισότητας πιθανότητας είναι 1.
- **Ίση Ευκαιρία (Equal Opportunity)**
 - **Διαφορά Ίσης Ευκαιρίας (Equal Opportunity Difference):** υπολογίζεται ως η διαφορά μεταξύ της μέγιστης και της ελάχιστης τιμής ρυθμού αληθώς θετικών δειγμάτων, ανάμεσα στις διαφορετικές υποομάδες που χαρακτηρίζονται από ένα ευαίσθητο χαρακτηριστικό. Η ιδανική τιμή της διαφοράς ίσης ευκαιρίας είναι 0.
 - **Αναλογία Ίσης Ευκαιρίας (Equal Opportunity Ratio):** υπολογίζεται ως η αναλογία της ελάχιστης προς την μέγιστη τιμή ρυθμού αληθώς θετικών δειγμάτων, ανάμεσα στις διαφορετικές υποομάδες που χαρακτηρίζονται από ένα ευαίσθητο χαρακτηριστικό. Η ιδανική τιμή της αναλογίας ίσης ευκαιρίας είναι 1.

- **Δημογραφική Ισότητα (Demographic Parity)**

- **Διαφορά Δημογραφικής ισότητας (Demographic Parity Difference):** υπολογίζεται ως η διαφορά μεταξύ της μέγιστης και της ελάχιστης τιμής ρυθμού θετικών προβλέψεων, ανάμεσα στις διαφορετικές υποομάδες που χαρακτηρίζονται από ένα ευαίσθητο χαρακτηριστικό. Η ιδανική τιμή της διαφοράς δημογραφικής ισότητας είναι 0.
- **Αναλογία Δημογραφικής ισότητας (Demographic Parity Ratio):** υπολογίζεται ως η αναλογία της ελάχιστης προς την μέγιστη τιμή ρυθμού θετικών προβλέψεων, ανάμεσα στις διαφορετικές υποομάδες που χαρακτηρίζονται από ένα ευαίσθητο χαρακτηριστικό. Η ιδανική τιμή της αναλογίας δημογραφικής ισότητας είναι 1.

6.4.2 Αντιμετώπιση των Διακρίσεων με Μεθόδους Προεπεξεργασίας

6.4.2.1 Αφαίρεση Συσχετίσεων (Correlation Remover)

Κατά την εφαρμογή του αλγόριθμου αφαίρεσης συσχετίσεων, προεπεξεργαζόμαστε τα δεδομένα εισόδου για να αφαιρέσουμε ή να μειώσουμε τη συσχέτισή τους με τα ευαίσθητα χαρακτηριστικά, με σκοπό την εξάλειψη των διακρίσεων. Από το σύνολο δεδομένων διατηρούμε μόνο τις στήλες με τα χαρακτηριστικά επεξεργασίας ψηφιακού σήματος, τις ετικέτες, το όνομα χρήστη και το ευαίσθητο χαρακτηριστικό, για το οποίο επιδιώκουμε τη βελτίωση της αμεροληψίας. Έπειτα, μετατρέπουμε τη στήλη του ευαίσθητου χαρακτηριστικού σε αριθμητικά δεδομένα, όπως απαιτείται και στη συνέχεια εφαρμόζουμε την κλάση αφαίρεσης συσχετίσεων με παράμετρο την στήλη του ευαίσθητου χαρακτηριστικού, για το οποίο θέλουμε να εξαλείψουμε τις διακρίσεις. Αντικαθιστούμε τα χαρακτηριστικά επεξεργασίας ψηφιακού σήματος με τα νέα χαρακτηριστικά που προκύπτουν από τη εφαρμογή του αλγόριθμου, αφού προηγουμένως έχουν αφαιρεθεί τυχόν συσχετίσεις που υπάρχουν από τα ευαίσθητα χαρακτηριστικά. Τα νέα δεδομένα εισάγονται στον απλό ταξινομητή τυχαίου δάσους για δυαδική ταξινόμηση, αφού αντιστοιχήσουμε ξανά την κατηγορία των ευαίσθητων χαρακτηριστικών που έχουμε μετατρέψει σε αριθμητικές τιμές, με τις αρχικές τιμές της κατηγορίας. Εμφανίζουμε την συσχέτιση των χαρακτηριστικών επεξεργασίας ψηφιακού σήματος με το ευαίσθητο χαρακτηριστικό που έχουμε επιλέξει πριν και μετά την αφαίρεση των συσχετίσεων επιδιώκοντας οι νέες τιμές να είναι όσο γίνεται πιο κοντά στο μηδέν. Υπολογίζουμε τις μετρικές αξιολόγησης της αμεροληψίας εξάγοντας τον μέσο όρο κάθε μεγέθους από τις επαναλήψεις του πειράματος. Συγκρίνουμε, τέλος, τα αποτελέσματα αμεροληψίας με εκείνα που εξάγονται από τον απλό ταξινομητή τυχαίου δάσους, με είσοδο τα αρχικά δεδομένα χωρίς την αφαίρεση των συσχετίσεων.

6.4.3 Αντιμετώπιση των Διακρίσεων με Μεθόδους κατά την Επεξεργασία

6.4.3.1 Εκθετική Κλίση (Exponentiated Gradient)

Η Εκθετική Κλίση είναι ένας αλγόριθμος που χρησιμοποιείται για την καταπολέμηση των διακρίσεων σε μοντέλα απλών ταξινομητών. Παρεμβαίνει στην εκπαίδευση του μοντέλου μεταβάλλοντας με συνεχείς επαναλήψεις τα βάρη του, ώστε να μειώσει τη μεροληψία, διατηρώντας παράλληλα καλή απόδοση στις προβλέψεις. Για την εφαρμογή της μεθόδου αυτής στο μοντέλο απλού ταξινομητή που έχουμε αναπτύξει, καλούμε την κλάση της εκθετικής κλίσης, στην οποία θέτουμε ως εκτιμητή (estimator), που υλοποιεί τις μεθόδους της κλάσης, τον απλό ταξινομητή τυχαίου δάσους που περιγράψαμε προηγουμένως. Στη συνέχεια, επιλέγουμε την ισότητα πιθανότητας ως μετρική αμεροληψίας που απαιτείται ως περιορισμός (constraints), με βάση τον οποίο ο αλγόριθμος θα προσπαθήσει να ισορροπήσει τις προβλέψεις του ανάμεσα στις υποομάδες

του ευαίσθητου χαρακτηριστικού, καθώς λόγω της διπλής απαίτησης για ίσο ρυθμό αληθώς θετικών προβλέψεων και ίσο ρυθμό ψευδώς θετικών προβλέψεων, θέτει πιο αυστηρές προϋποθέσεις αμεροληψίας. Η εκπαίδευση του μοντέλου, γίνεται με είσοδο των αρχικών δεδομένων, προσθέτοντας όμως κάθε φορά και την κατηγορία ευαίσθητων χαρακτηριστικών στην οποία θέλουμε να εξαλείψουμε τις διακρίσεις. Η διαδικασία για καθεμία κατηγορία ευαίσθητων χαρακτηριστικών επαναλαμβάνεται, και η τελική τιμή κάθε μετρικής αμεροληψίας μετά την εφαρμογή της μεθόδου εκθετικής κλίσης, είναι ο μέσος όρος των επαναλήψεων.

6.4.3.2 Ανταγωνιστική Εκπαίδευση (Adversarial Learning)

Η μέθοδος ανταγωνιστικής εκπαίδευσης είναι μια μέθοδος καταπολέμησης των διακρίσεων που χρησιμοποιούμε στην διαδικασία εκπαίδευσης του μοντέλου. Εφαρμόζεται στο μοντέλο συνελκτικού νευρωνικού δικτύου που έχουμε αναπτύξει, για τον διαχωρισμό των ευαίσθητων χαρακτηριστικών από τα χαρακτηριστικά των δεδομένων εισόδου. Κατά τη μέθοδο αυτή, το μοντέλο λειτουργεί ανάλογα με το αρχικό, με τη διαφορά ότι θέτουμε ως είσοδο τα Mel-Φασματογραφήματα, όπως προβλέπεται, μαζί όμως με την κατηγορία ευαίσθητων χαρακτηριστικών ως προς την οποία θέλουμε να εξαλείψουμε τις διακρίσεις. Στη συνέχεια, προσθέτουμε στο μοντέλο ένα δεύτερο Επίπεδο Εξόδου (Output Layer), το οποίο ταξινομεί τα δείγματα με βάση το ευαίσθητο χαρακτηριστικό που ορίζουμε κάθε φορά. Πριν από αυτή την έξοδο, έχουμε εφαρμόσει ένα Επίπεδο Αντιστροφής Κλίσης (Gradient Reversal Layer), το οποίο λειτουργεί μέσω μίας συνάρτησης που ενώ κατά την εμπρόσθια διάδοση δεν μεταβάλλει τίποτα, κατά την οπισθοδιάδοση αντιστρέφει τα πρόσημα των παραγώγων των δεδομένων. Με αυτόν τον τρόπο, αντιστρέφεται η εκπαίδευση ως προς τα ευαίσθητα χαρακτηριστικά, αναγκάζοντας το μοντέλο να προβλέπει ανεξάρτητα από αυτά. Στο δεύτερο επίπεδο εξόδου που έχουμε προσθέσει, ορίζουμε τις κατηγορίες ταξινόμησης σε πλήθος αντίστοιχο με το πλήθος των υποκατηγοριών του ευαίσθητου χαρακτηριστικού, του οποίου προσπαθούμε να εξαλείψουμε τις διακρίσεις. Ως συνάρτηση ενεργοποίησης, επιλέγουμε τη softmax, η οποία είναι κατάλληλη για προβλήματα πολυκατηγορικής ταξινόμησης, όπως στην περίπτωση μας η ταξινόμηση των ευαίσθητων χαρακτηριστικών. Λόγω του μεγάλου όγκου δεδομένων και της πολυπλοκότητας του μοντέλου, εφαρμόσαμε ως μέθοδο κανονικοποίησης και για αποφυγή της υπερεκπαίδευσης, ένα Επίπεδο Απενεργοποίησης Νευρώνων (Dropout Layer) με παράμετρο 0,5. Η συνάρτηση απώλειας που ορίζουμε για την έξοδο που αφορά την εκπαίδευση του ευαίσθητου χαρακτηριστικού, είναι η κατηγορική διασταύρωση εντροπίας (categorical_crossentropy), ως την πλέον κατάλληλη για τις υποκατηγορίες του ευαίσθητου χαρακτηριστικού που είναι περισσότερες από δύο. Η απόδοση του μοντέλου καθορίζεται μόνο από τις προβλέψεις της δυαδικής ταξινόμησης για διάγνωση της COVID-19 και όχι από την διαδικασία ταξινόμησης των ευαίσθητων χαρακτηριστικών. Για κάθε κατηγορία ευαίσθητων χαρακτηριστικών, επαναλαμβάνουμε την εκτέλεση του μοντέλου και εξάγουμε ως τελικό αποτέλεσμα για κάθε μετρική αμεροληψίας, τον μέσο όρο των επαναλήψεων.

6.4.4 Αντιμετώπιση των Διακρίσεων με Μεθόδους μετά την Επεξεργασία

6.4.4.1 Βελτιστοποιητής Κατωφλίου (Threshold Optimizer)

Ο Βελτιστοποιητής Κατωφλίου είναι ένας αλγόριθμος καταπολέμησης διακρίσεων που εφαρμόζεται στο ήδη εκπαιδευμένο μοντέλο απλού ταξινομητή που έχουμε αναπτύξει. Μετά την εκπαίδευση του ταξινομητή τυχαίους με την διαδικασία που περιγράψαμε προηγουμένως, καλούμε την κλάση του Βελτιστοποιητή Κατωφλίου, στην οποία θέτουμε ως εκτιμητή (estimator) τον ταξινομητή αυτόν, ορίζοντάς τον ως ήδη εκπαιδευμένο. Στη συνέχεια, όπως πράξαμε και με την μέθοδο εκθετικής κλίσης του δυαδικού ταξινομητή, επιλέγουμε την ισότητα πιθανότητας ως περιορισμό, βάσει του οποίου ο αλγόριθμος θα προσπαθήσει να ισορροπήσει τις προβλέψεις του ανάμεσα στις υποομάδες της κατηγορίας των ευαίσθητων χαρακτηριστικών που επιλέγουμε.

Παράλληλα, θέτουμε ως στόχο μεγιστοποίησης την τιμή της ακρίβειας, προκειμένου να διατηρήσουμε την καλύτερη δυνατή ισορροπία μεταξύ της απόδοσης και της αμεροληψίας του μοντέλου. Εφαρμόζουμε τον βελτιστοποιητή κατωφλίου στα δεδομένα εκπαίδευσης, ορίζοντας και την κατηγορία ευαίσθητων χαρακτηριστικών στην οποία θέλουμε να εξαλείψουμε τις διακρίσεις και λαμβάνουμε τις προβλέψεις που παράγει το μοντέλο στο σύνολο ελέγχου, μέσω της προσαρμογής του κατωφλίου πρόβλεψης χωριστά για κάθε υποομάδα του ευαίσθητου χαρακτηριστικού. Υπολογίζουμε, τέλος, τις νέες τιμές των μετρικών αμεροληψίας που έχει επιτύχει το μοντέλο. Η διαδικασία επαναλαμβάνεται ορισμένες φορές για καθεμία κατηγορία ευαίσθητων χαρακτηριστικών. Οι τιμές των μετρικών αμεροληψίας για κάθε κατηγορία ευαίσθητων χαρακτηριστικών μετά την εφαρμογή του βελτιστοποιητή κατωφλίου προκύπτουν κατ'αναλογία με τα προηγούμενα μοντέλα ως οι μέσες τιμές των επαναλήψεων.

7 Αποτελέσματα

7.1 Ταξινόμηση

7.1.1 Ταξινόμηση με ενίσχυση της Αμεροληψίας Φύλου

Ακολουθως, παρουσιάζονται τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του απλού ταξινομητή, πριν και μετά την εφαρμογή μεθόδων καταπολέμησης των διακρίσεων ως προς το φύλο:

Μετρικές Απόδοσης Ταξινόμησης	Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων Ηλικίας	μετά από Εκθετική Κλίση για την Ηλικία	μετά από Βελτιστοποιητή Κατωφλίου για την Ηλικία
Ακρίβεια	0,5475	0,5538	0,5530	0,5528
Ευαισθησία	0,5141	0,4933	0,5657	0,5259
Ειδίκευση	0,5646	0,6040	0,5756	0,5793
Ισορροπημένη Ακρίβεια	0,5393	0,5486	0,5706	0,5526

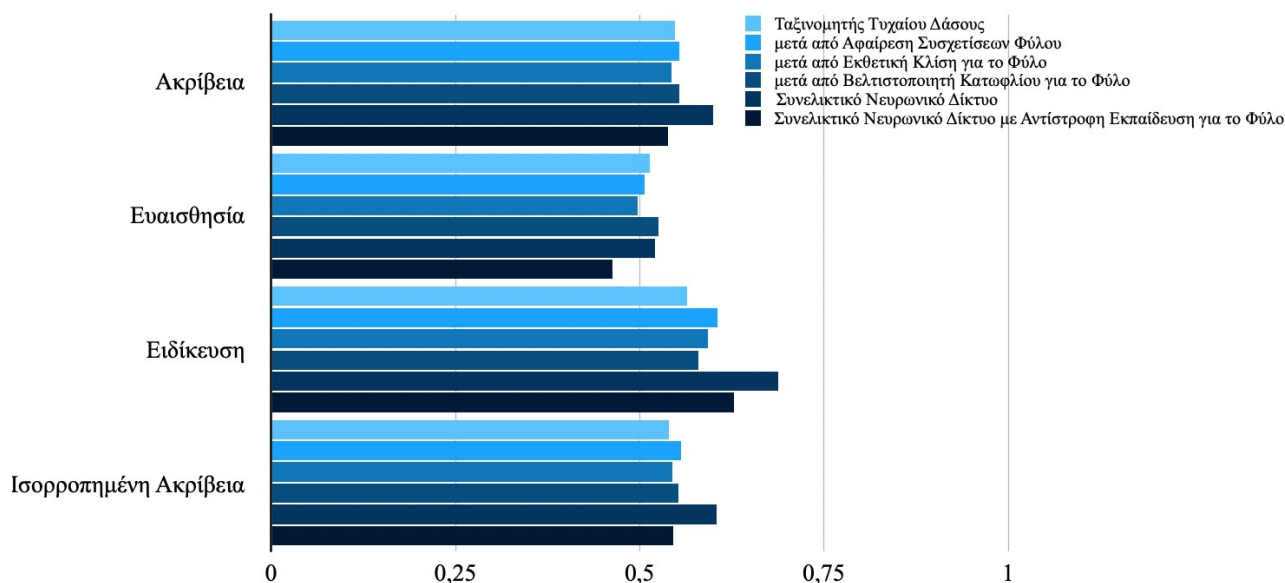
Πίνακας 7.1.1: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία του φύλου.

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του συνελκτικού νευρωνικού δικτύου, πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για καταπολέμηση των διακρίσεων ως προς το φύλο, είναι:

Πίνακας 7.1.2: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία του φύλου.

Μετρικές Απόδοσης Ταξινόμησης	Συνελκτικό Νευρωνικό Δίκτυο	Συνελκτικό Νευρωνικό Δίκτυο με Αντίστροφη Εκπαίδευση για την Ηλικία
Ακρίβεια	0,5998	0,5529
Ευαισθησία	0,5210	0,4798
Ειδίκευση	0,6878	0,6310
Ισορροπημένη Ακρίβεια	0,6044	0,5554
F1-Βαθμολογία	0,5519	0,4782
ROC-AUC	0,6439	0,5890

Συγκεντρωτικά, οι μετρικές απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία του φύλου, είναι:



Εικόνα 7.1.1: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία του φύλου.

7.1.2 Ταξινόμηση με ενίσχυση της Αμεροληψίας Ηλικίας

Ακολουθώντας, παρουσιάζονται τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων καταπολέμησης των διακρίσεων ως προς την ηλικία:

Πίνακας 7.1.3: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία της ηλικίας.

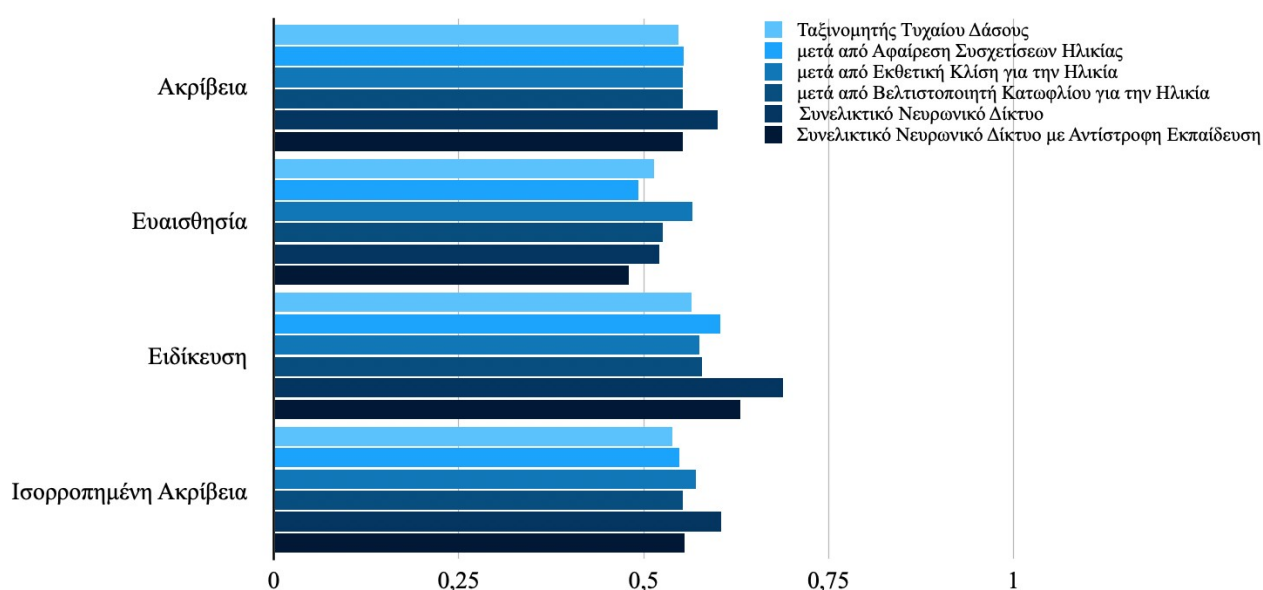
Μετρικές Απόδοσης Ταξινόμησης	Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων Ηλικίας	μετά από Εκθετική Κλίση για την Ηλικία	μετά από Βελτιστοποιητή Κατωφλίου για την Ηλικία
Ακρίβεια	0,5475	0,5538	0,5530	0,5528
Ευαισθησία	0,5141	0,4933	0,5657	0,5259
Ειδικευση	0,5646	0,6040	0,5756	0,5793
Ισορροπημένη Ακρίβεια	0,5393	0,5486	0,5706	0,5526

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για καταπολέμηση των διακρίσεων ως προς την ηλικία είναι:

Πίνακας 7.1.4: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία της ηλικίας.

Μετρικές Απόδοσης Ταξινόμησης	Συνελκτικό Νευρωνικό Δίκτυο	Συνελκτικό Νευρωνικό Δίκτυο με Αντίστροφη Εκπαίδευση για την Ηλικία
Ακρίβεια	0,5998	0,5529
Ευαισθησία	0,5210	0,4798
Ειδίκευση	0,6878	0,6310
Ισορροπημένη Ακρίβεια	0,6044	0,5554
F1-Βαθμολογία	0,5519	0,4782
ROC-AUC	0,6439	0,5890

Συγκεντρωτικά, οι μετρικές απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία της ηλικίας, είναι:



Εικόνα 7.1.2: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία της ηλικίας.

7.1.3 Ταξινόμηση με ενίσχυση της Αμεροληψίας του Ιατρικού Ιστορικού

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του απλού ταξινομητή, πριν και μετά την εφαρμογή μεθόδων καταπολέμησης των διακρίσεων ως προς το ιατρικό ιστορικό, παρουσιάζονται ακολούθως:

Πίνακας 7.1.5: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία του ιατρικού ιστορικού.

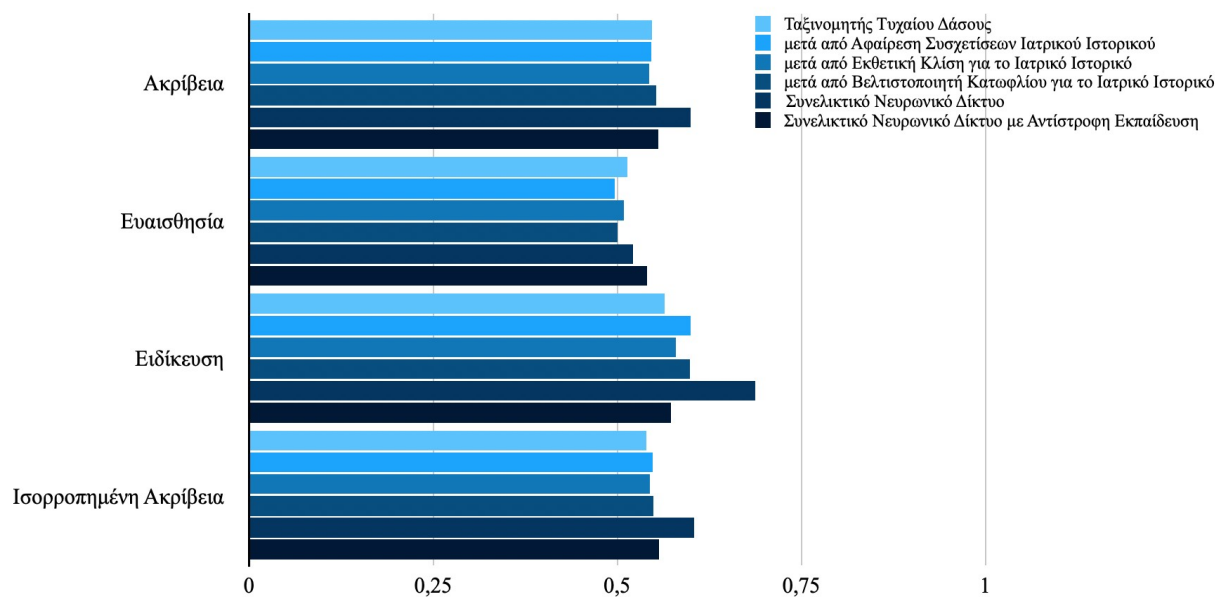
Μετρικές Απόδοσης Ταξινόμησης	Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων Ιατρικού Ιστορικού	μετά από Εκθετική Κλίση για το Ιατρικό Ιστορικό	μετά από Βελτιστοποιητή Κατωφλίου για το Ιατρικό Ιστορικό
Ακρίβεια	0,5475	0,5460	0,5437	0,5533
Ευαισθησία	0,5141	0,4963	0,5086	0,5002
Ειδίκευση	0,5646	0,5995	0,5800	0,5987
Ισορροπημένη Ακρίβεια	0,5393	0,5479	0,5443	0,5494

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του συνελικτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για καταπολέμηση των διακρίσεων ως προς το ιατρικό ιστορικό, είναι:

Πίνακας 7.1.6: Πίνακας σύγκρισης μετρικών απόδοσης του συνελικτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία του ιατρικού ιστορικού.

Μετρικές Απόδοσης Ταξινόμησης	Συνελικτικό Νευρωνικό Δίκτυο	Συνελικτικό Νευρωνικό Δίκτυο με Αντίστροφη Εκπαίδευση για το Ιατρικό Ιστορικό
Ακρίβεια	0,5998	0,5554
Ευαισθησία	0,5210	0,5403
Ειδίκευση	0,6878	0,5732
Ισορροπημένη Ακρίβεια	0,6044	0,5567
F1-Βαθμολογία	0,5519	0,5023
ROC-AUC	0,6439	0,5922

Οι μετρικές απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία της ηλικίας, συνολικά, είναι:



Εικόνα 7.1.3: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία ιατρικού ιστορικού.

7.1.4 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς τα Συμπτώματα

Στον παρακάτω πίνακα, παρουσιάζονται τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων καταπολέμησης των διακρίσεων ως προς τα συμπτώματα:

Πίνακας 7.1.7: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία των συμπτωμάτων.

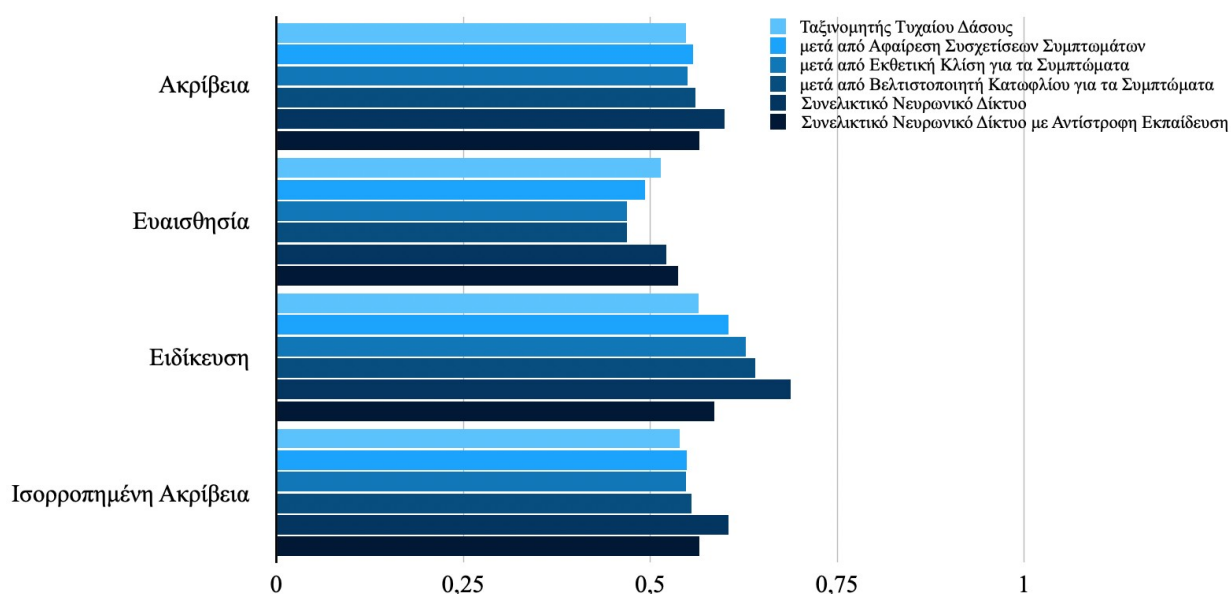
Μετρικές Απόδοσης Ταξινόμησης	Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων Συμπτωμάτων	μετά από Εκθετική Κλίση για τα Συμπτώματα	μετά από Βελτιστοποιητή Κατωφλίου για τα Συμπτώματα
Ακρίβεια	0,5475	0,5570	0,55	0,5606
Ευαισθησία	0,5141	0,4933	0,4689	0,4689
Ειδικευση	0,5646	0,6044	0,6275	0,6405
Ισορροπημένη Ακρίβεια	0,5393	0,5488	0,5482	0,5547

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του συνελκτικού νευρωνικού δικτύου, πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για καταπολέμηση των διακρίσεων ως προς τα συμπτώματα, είναι:

Πίνακας 7.1.8: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία των συμπτωμάτων.

Μετρικές Απόδοσης Ταξινόμησης	Συνελκτικό Νευρωνικό Δίκτυο	Συνελκτικό Νευρωνικό Δίκτυο με Αντίστροφη Εκπαίδευση για τα Συμπτώματα
Ακρίβεια	0,5998	0,5658
Ευαισθησία	0,5210	0,5374
Ειδίκευση	0,6878	0,5859
Ισορροπημένη Ακρίβεια	0,6044	0,5654
F1-Βαθμολογία	0,5519	0,5348
ROC-AUC	0,6439	0,5951

Συγκριτικά, οι μετρικές απόδοσης όλων των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία της ηλικίας, είναι:



Εικόνα 7.1.4: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία των συμπτωμάτων.

7.1.5 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς την Πηγή Συλλογής Δεδομένων

Ακολουθώντας, παρουσιάζονται τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του απλού ταξινομητή, πριν και μετά την εφαρμογή μεθόδων καταπολέμησης των διακρίσεων ως προς την πηγή συλλογής των δεδομένων:

Πίνακας 7.1.9: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία της πηγής συλλογής δεδομένων.

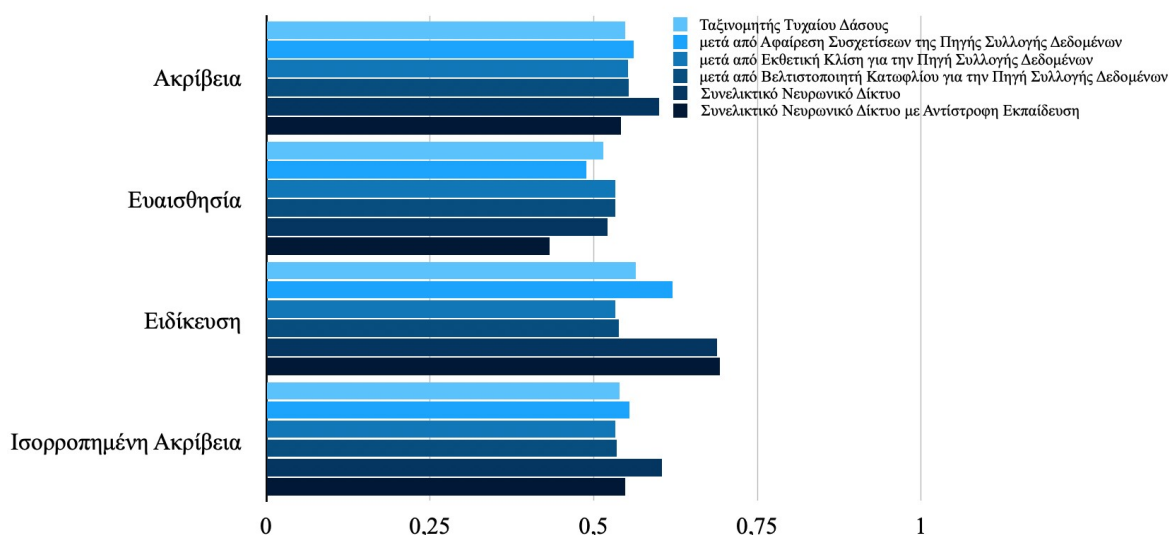
Μετρικές Απόδοσης Ταξινόμησης	Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων της Πηγής Συλλογής Δεδομένων	μετά από Εκθετική Κλίση για την Πηγή Συλλογής Δεδομένων	μετά από Βελτιστοποιητή Κατωφλίου για την Πηγή Συλλογής Δεδομένων
Ακρίβεια	0,5475	0,5612	0,5523	0,5538
Ευαισθησία	0,5141	0,4882	0,5330	0,5327
Ειδίκευση	0,5646	0,6199	0,5330	0,5384
Ισορροπημένη Ακρίβεια	0,5393	0,5540	0,5330	0,5355

Τα αποτελέσματα αξιολόγησης της απόδοσης, στις προβλέψεις του συνελκτικού νευρωνικού δικτύου, πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για καταπολέμηση των διακρίσεων ως προς την πηγή συλλογής των δεδομένων, είναι:

Πίνακας 7.1.10: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία της πηγής συλλογής δεδομένων.

Μετρικές Απόδοσης Ταξινόμησης	Συνελκτικό Νευρωνικό Δίκτυο	Συνελκτικό Νευρωνικό Δίκτυο με Αντίστροφη Εκπαίδευση για την Πηγή Συλλογής Δεδομένων
Ακρίβεια	0,5998	0,5417
Ευαισθησία	0,5210	0,4326
Ειδίκευση	0,6878	0,6927
Ισορροπημένη Ακρίβεια	0,6044	0,5483
F1-Βαθμολογία	0,5519	0,4375
ROC-AUC	0,6439	0,5883

Οι μετρικές απόδοσης των μοντέλων, πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία της πηγής συλλογής δεδομένων, συνολικά, είναι:



Εικόνα 7.1.5: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία πηγής συλλογής δεδομένων.

7.1.6 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς την Ηλικία σε συνδυασμό με το Φύλο

Στον παρακάτω πίνακα, παρουσιάζονται τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του απλού ταξινομητή, πριν και μετά την εφαρμογή μεθόδων καταπολέμησης των διακρίσεων ως προς την ηλικία σε συνδυασμό με το φύλο:

Πίνακας 7.1.11: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία της ηλικίας σε συνδυασμό με το φύλο.

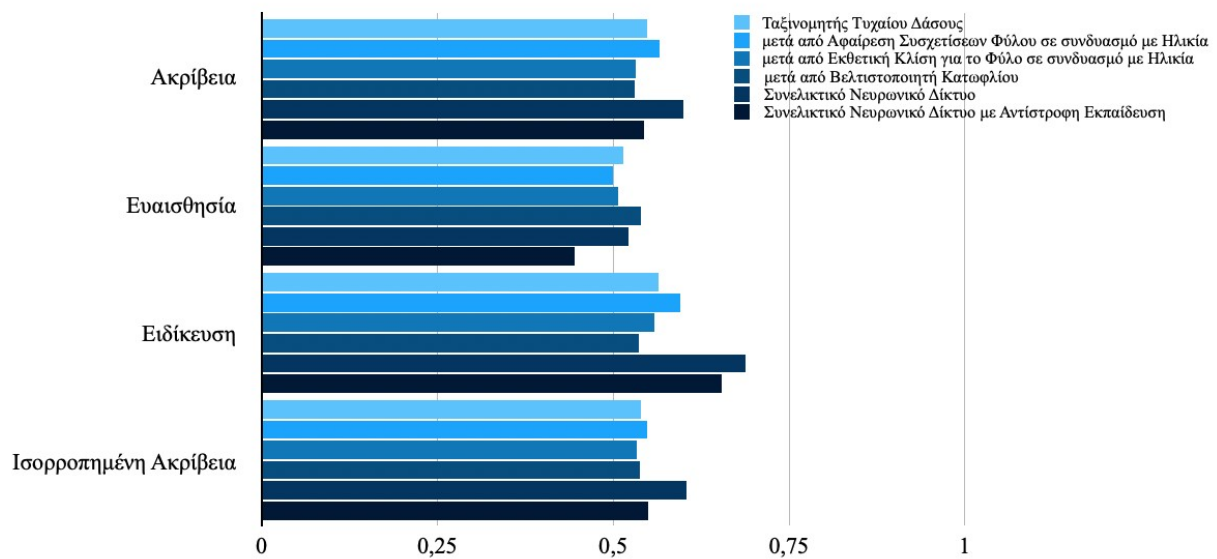
Μετρικές Απόδοσης Ταξινόμησης	Ταξινόμητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων Φύλου σε συνδυασμό με Ηλικία	μετά από Εκθετική Κλίση για το Φύλο σε συνδυασμό με Ηλικία	μετά από Βελτιστοποιητή Κατωφλίου για το Φύλο σε συνδυασμό με Ηλικία
Ακρίβεια	0,5475	0,5661	0,5313	0,5312
Εναισθησία	0,5141	0,5001	0,5070	0,5397
Ειδικευση	0,5646	0,5951	0,5580	0,5365
Ισορροπημένη Ακρίβεια	0,5393	0,5476	0,5329	0,5381

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για καταπολέμηση των διακρίσεων ως προς την ηλικία σε συνδυασμό με το φύλο, είναι:

Πίνακας 7.1.12: Πίνακας σύγκρισης μετρικών απόδοσης του συνελικτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία της ηλικίας σε συνδυασμό με το φύλο.

Μετρικές Απόδοσης Ταξινόμησης	Συνελικτικό Νευρωνικό Δίκτυο	Συνελικτικό Νευρωνικό Δίκτυο με Αντίστροφη Εκπαίδευση για το Φύλο σε συνδυασμό με Ηλικία
Ακρίβεια	0,5998	0,5435
Ευαισθησία	0,5210	0,4456
Ειδίκευση	0,6878	0,6545
Ισορροπημένη Ακρίβεια	0,6044	0,5492
F1-Βαθμολογία	0,5519	0,4346
ROC-AUC	0,6439	0,5768

Συγκριτικά, οι μετρικές απόδοσης όλων των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία της ηλικίας σε συνδυασμό με το φύλο, είναι:



Εικόνα 7.1.6: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία ηλικίας σε συνδυασμό με φύλο.

7.1.7 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς την Ηλικία σε συνδυασμό με τα Συμπτώματα

Ακολουθώς, παρουσιάζονται τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του απλού ταξινομητή, πριν και μετά την εφαρμογή μεθόδων καταπολέμησης των διακρίσεων ως προς την ηλικία σε συνδυασμό με τα συμπτώματα:

Πίνακας 7.1.13: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα.

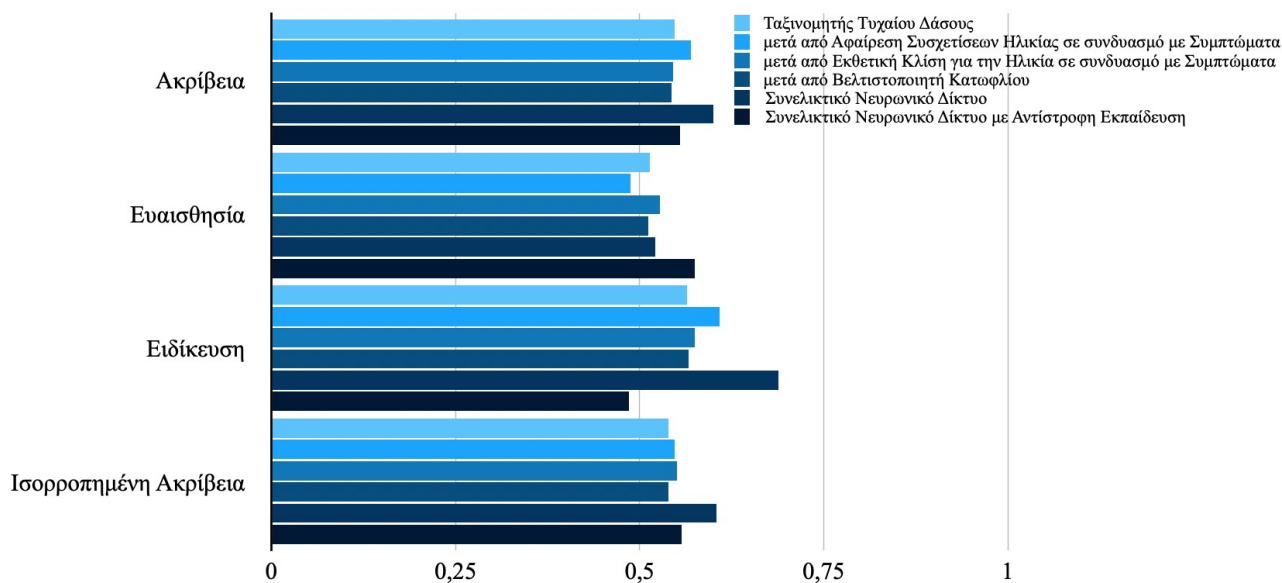
Μετρικές Απόδοσης Ταξινόμησης	Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων Ηλικίας σε συνδυασμό με Συμπτώματα	μετά από Εκθετική Κλίση για την Ηλικία σε συνδυασμό με Συμπτώματα	μετά από Βελτιστοποιητή Κατωφλίου για την Ηλικία σε συνδυασμό με Συμπτώματα
Ακρίβεια	0,5475	0,5691	0,5452	0,5435
Ευαισθησία	0,5141	0,4871	0,5272	0,5115
Ειδίκευση	0,5646	0,6086	0,5748	0,5661
Ισορροπημένη Ακρίβεια	0,5393	0,5478	0,551	0,5388

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για καταπολέμηση των διακρίσεων ως προς την ηλικία σε συνδυασμό με τα συμπτώματα, είναι:

Πίνακας 7.1.14: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα.

Μετρικές Απόδοσης Ταξινόμησης	Συνελκτικό Νευρωνικό Δίκτυο	Συνελκτικό Νευρωνικό Δίκτυο με Αντίστροφη Εκπαίδευση για την Ηλικία σε συνδυασμό με Συμπτώματα
Ακρίβεια	0,5998	0,5545
Ευαισθησία	0,5210	0,5745
Ειδίκευση	0,6878	0,4852
Ισορροπημένη Ακρίβεια	0,6044	0,5568
F1-Βαθμολογία	0,5519	0,5310
ROC-AUC	0,6439	0,5803

Οι μετρικές απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα, συνολικά, είναι:



Εικόνα 7.1.7: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία ηλικίας σε συνδυασμό με συμπτώματα.

7.1.8 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς το Φύλο σε συνδυασμό με τα Συμπτώματα

Ακολουθώς παρουσιάζονται τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων καταπολέμησης των διακρίσεων ως προς το φύλο σε συνδυασμό με τα συμπτώματα:

Πίνακας 7.1.15: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία του φύλου σε συνδυασμό με τα συμπτώματα.

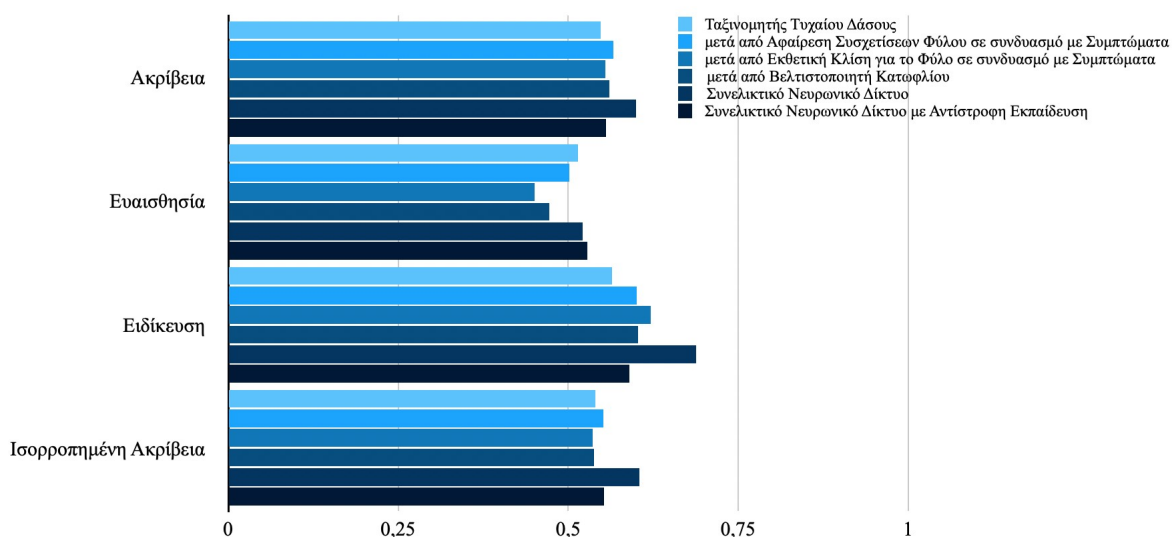
Μετρικές Απόδοσης Ταξινόμησης	Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων Φύλου σε συνδυασμό με Συμπτώματα	μετά από Εκθετική Κλίση για το Φύλο σε συνδυασμό με Συμπτώματα	μετά από Βελτιστοποιητή Κατωφλίου για το Φύλο σε συνδυασμό με Συμπτώματα
Ακρίβεια	0,5475	0,5659	0,5545	0,5605
Ευαισθησία	0,5141	0,5017	0,4509	0,4724
Ειδικευση	0,5646	0,6006	0,6216	0,6027
Ισορροπημένη Ακρίβεια	0,5393	0,5511	0,5362	0,5375

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του συνελκτικού νευρωνικού δικτύου, πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για καταπολέμηση των διακρίσεων ως προς το φύλο σε συνδυασμό με τα συμπτώματα, είναι:

Πίνακας 7.1.16: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία του φύλου σε συνδυασμό με τα συμπτώματα.

Μετρικές Απόδοσης Ταξινόμησης	Συνελκτικό Νευρωνικό Δίκτυο	Συνελκτικό Νευρωνικό Δίκτυο με Αντίστροφη Εκπαίδευση για το Φύλο σε συνδυασμό με Συμπτώματα
Ακρίβεια	0,5998	0,5559
Ευαισθησία	0,5210	0,5276
Ειδίκευση	0,6878	0,5896
Ισορροπημένη Ακρίβεια	0,6044	0,5527
F1-Βαθμολογία	0,5519	0,5395
ROC-AUC	0,6439	0,6053

Συγκεντρωτικά, οι μετρικές απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία του φύλου σε συνδυασμό με τα συμπτώματα, είναι:



Εικόνα 7.1.8: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία φύλου σε συνδυασμό με συμπτώματα.

7.1.9 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς την Ηλικία σε συνδυασμό με το Ιατρικό Ιστορικό

Ακολουθώντας, παρουσιάζονται τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του απλού ταξινομητή, πριν και μετά την εφαρμογή μεθόδων καταπολέμησης των διακρίσεων ως προς την ηλικία σε συνδυασμό με το ιατρικό ιστορικό:

Πίνακας 7.1.17: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

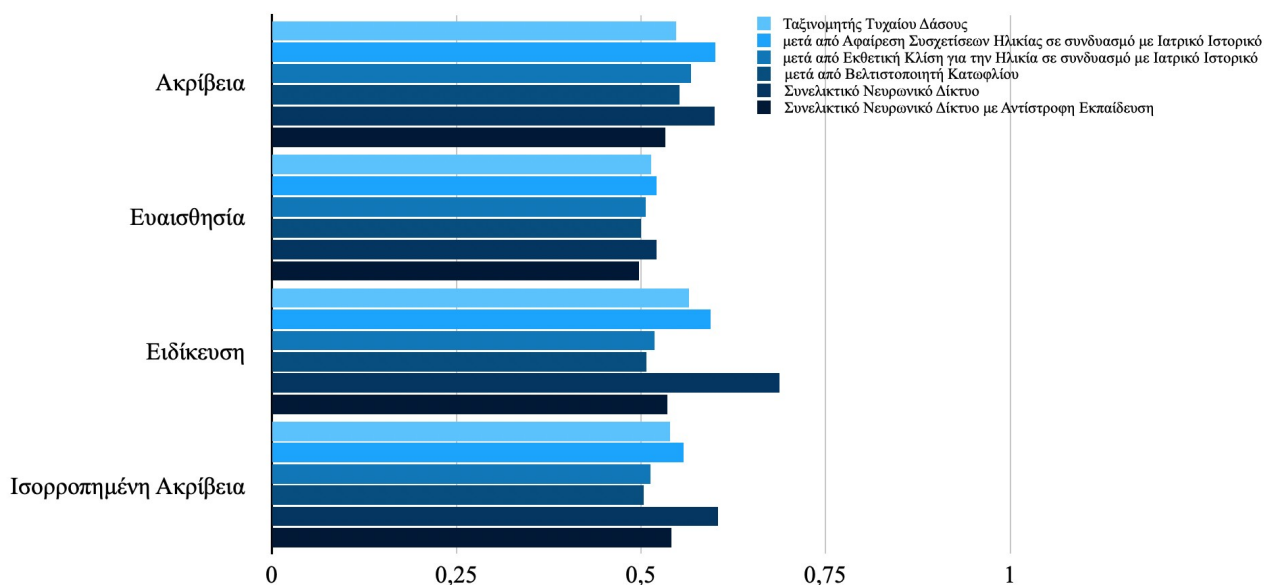
Μετρικές Απόδοσης Ταξινόμησης	Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων Ηλικίας σε συνδυασμό με Ιατρικό Ιστορικό	μετά από Εκθετική Κλίση για την Ηλικία σε συνδυασμό με Ιατρικό Ιστορικό	μετά από Βελτιστοποιητή Κατωφλίου για την Ηλικία σε συνδυασμό με Ιατρικό Ιστορικό
Ακρίβεια	0,5475	0,6006	0,5676	0,5523
Ευαισθησία	0,5141	0,5207	0,5068	0,5002
Ειδίκευση	0,5646	0,5945	0,5180	0,5069
Ισορροπημένη Ακρίβεια	0,5393	0,5576	0,5124	0,5035

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του συνελκτικού νευρωνικού δικτύου, πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για καταπολέμηση των διακρίσεων ως προς την ηλικία σε συνδυασμό με το ιατρικό ιστορικό, παρουσιάζονται ακολούθως:

Πίνακας 7.1.18: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

Μετρικές Απόδοσης Ταξινόμησης	Συνελκτικό Νευρωνικό Δίκτυο	Συνελκτικό Νευρωνικό Δίκτυο με Αντίστροφη Εκπαίδευση για την Ηλικία σε συνδυασμό με Ιατρικό Ιστορικό
Ακρίβεια	0,5998	0,5327
Ευαισθησία	0,5210	0,4972
Ειδίκευση	0,6878	0,5354
Ισορροπημένη Ακρίβεια	0,6044	0,5411
F1-Βαθμολογία	0,5519	0,4664
ROC-AUC	0,6439	0,5625

Οι μετρικές απόδοσης των μοντέλων, πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό, συνολικά, είναι:



Εικόνα 7.1.9: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία ηλικίας σε συνδυασμό με ιατρικό ιστορικό.

7.1.10 Ταξινόμηση με ενίσχυση της Αμεροληψίας ως προς το Φύλο σε συνδυασμό με το Ιατρικό Ιστορικό

Ακολούθως, παρουσιάζονται τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του απλού ταξινομητή , πριν και μετά την εφαρμογή μεθόδων καταπολέμησης των διακρίσεων ως προς το φύλο σε συνδυασμό με το ιατρικό ιστορικό:

Πίνακας 7.1.19: Πίνακας σύγκρισης μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

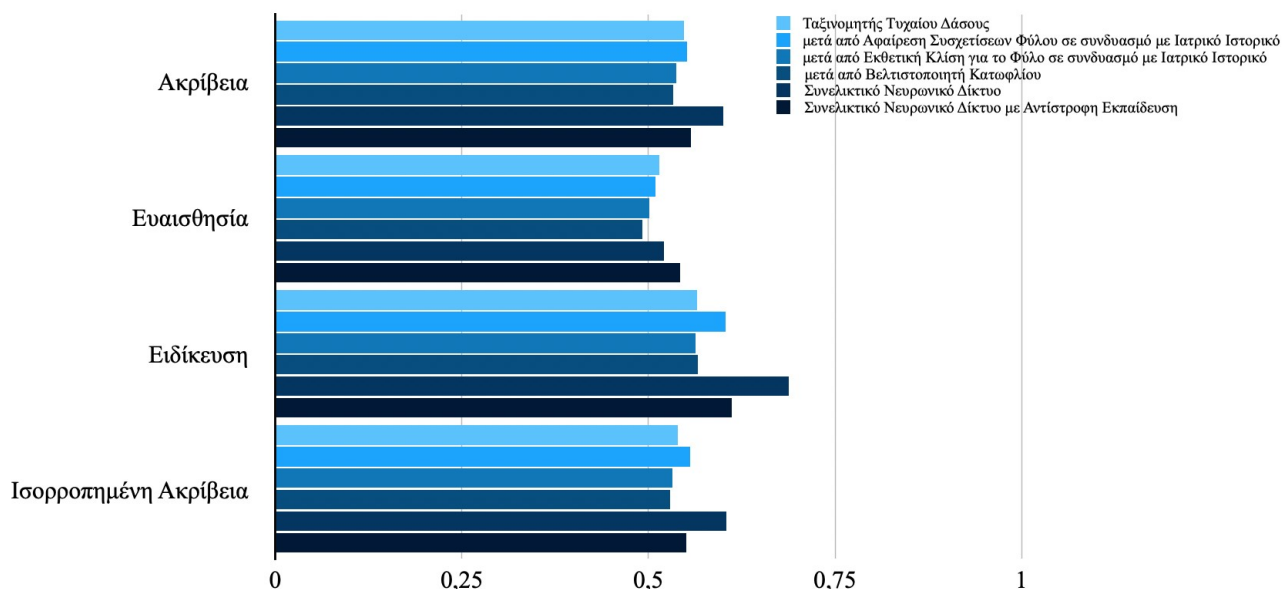
Μετρικές Απόδοσης Ταξινόμησης	Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων Φύλου σε συνδυασμό με Ιατρικό Ιστορικό	μετά από Εκθετική Κλίση για το Φύλο σε συνδυασμό με Ιατρικό Ιστορικό	μετά από Βελτιστοποιητή Κατωφλίου για το Φύλο σε συνδυασμό με Ιατρικό Ιστορικό
Ακρίβεια	0,5475	0,5515	0,5368	0,5336
Ευαισθησία	0,5141	0,5090	0,5012	0,4916
Ειδικευση	0,5646	0,6029	0,5630	0,5663
Ισορροπημένη Ακρίβεια	0,5393	0,5559	0,5321	0,5289

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του συνελκτικού νευρωνικού δικτύου, πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για καταπολέμηση των διακρίσεων ως προς το φύλο σε συνδυασμό με το ιατρικό ιστορικό, είναι:

Πίνακας 7.1.20: Πίνακας σύγκρισης μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

Μετρικές Απόδοσης Ταξινόμησης	Συνελκτικό Νευρωνικό Δίκτυο	Συνελκτικό Νευρωνικό Δίκτυο με Αντίστροφη Εκπαίδευση για το Φύλο σε συνδυασμό με Ιατρικό Ιστορικό
Ακρίβεια	0,5998	0,5564
Ευαισθησία	0,5210	0,5419
Ειδίκευση	0,6878	0,6116
Ισορροπημένη Ακρίβεια	0,6044	0,5505
F1-Βαθμολογία	0,5519	0,5308
ROC-AUC	0,6439	0,5842

Οι μετρικές απόδοσης των μοντέλων, πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό, συνολικά, είναι:



Εικόνα 7.1.10: Σύγκριση των μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για την κατηγορία φύλου σε συνδυασμό με ιατρικό ιστορικό.

7.1.11 Συνολική Αξιολόγηση Ταξινόμησης συγκριτικά για τις Μεθόδους Ενίσχυσης της Αμεροληψίας

Στον ακόλουθο πίνακα, παρουσιάζονται οι μέσες τιμές των μετρικών αξιολόγησης της απόδοσης όλων των κατηγοριών ευαίσθητων χαρακτηριστικών, για τις μεθόδους εξάλειψης των διακρίσεων που εφαρμόστηκαν στον απλό ταξινομητή τυχαίου δάσους, συγκριτικά με τις αρχικές τιμές που αφορούν τον ταξινομητή:

Πίνακας 7.1.21: Πίνακας σύγκρισης μέσων τιμών των μετρικών απόδοσης του απλού ταξινομητή πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων για όλες τις κατηγορίες.

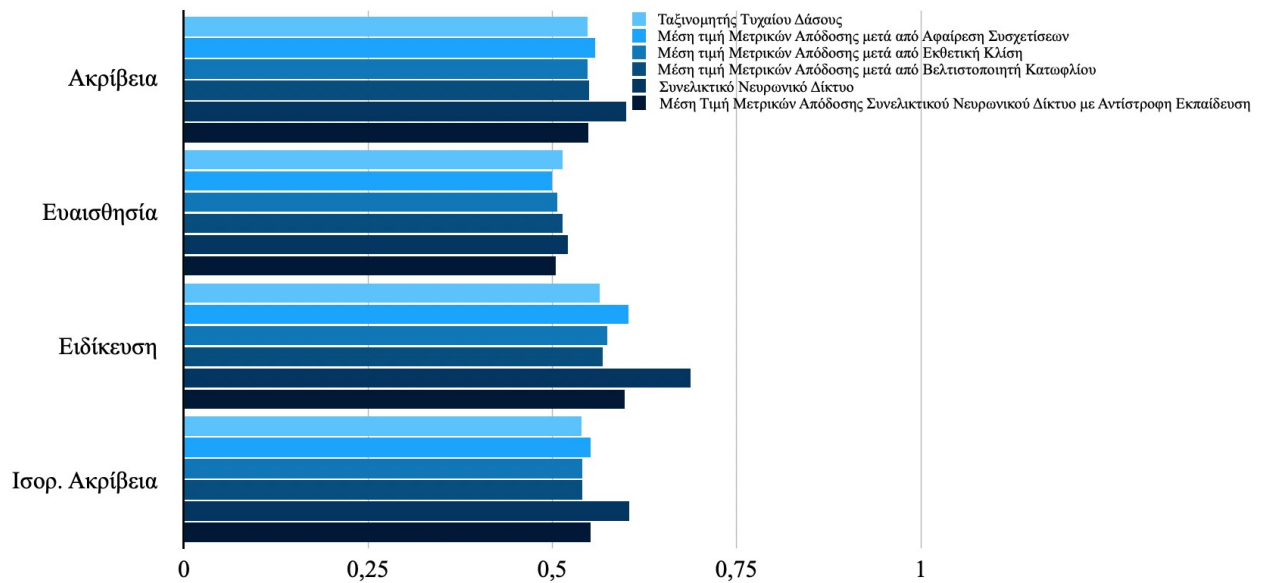
Μετρικές Απόδοσης Ταξινόμησης	Ταξινομητής Τυχαίου Δάσους	Μέση τιμή Μετρικών Απόδοσης μετά από Αφαίρεση Συσχετίσεων	Μέση τιμή Μετρικών Απόδοσης μετά από Εκθετική Κλίση	Μέση τιμή Μετρικών Απόδοσης μετά από Βελτιστοποιητή Κατωφλίου
Ακρίβεια	0,5475	0,5585	0,5478	0,5495
Ευαισθησία	0,5141	0,4996	0,5066	0,5140
Ειδικευση	0,5646	0,6035	0,5744	0,5681
Ισορροπημένη Ακρίβεια	0,5393	0,5515	0,5405	0,5410

Τα αποτελέσματα αξιολόγησης της απόδοσης στις προβλέψεις του συνελκτικού νευρωνικού δικτύου σε σύγκριση με τις μέσες τιμές των μετρικών απόδοσης από όλες τις κατηγορίες ευαίσθητων χαρακτηριστικών, μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης με στόχο την εξάλειψη των διακρίσεων και την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:

Πίνακας 7.1.22: Πίνακας σύγκρισης μέσων τιμών των μετρικών απόδοσης του συνελκτικού νευρωνικού δικτύου πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων για όλες τις κατηγορίες.

Μετρικές Απόδοσης Ταξινόμησης	Συνελκτικό Νευρωνικό Δίκτυο	Μέση Τιμή Μετρικών Απόδοσης Συνελκτικού Νευρωνικού Δίκτυο με Αντίστροφη Εκπαίδευση για όλες τις κατηγορίες ευαίσθητων χαρακτηριστικών
Ακρίβεια	0,5998	0,5493
Ευαισθησία	0,5210	0,5050
Ειδικευση	0,6878	0,5985
Ισορροπημένη Ακρίβεια	0,6044	0,5518
F1-Βαθμολογία	0,5519	0,4937
ROC-AUC	0,6439	0,5861

Οι μετρικές απόδοσης των μοντέλων κατά μέσο όρο, πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για όλες τις κατηγορίες, συνολικά, είναι:

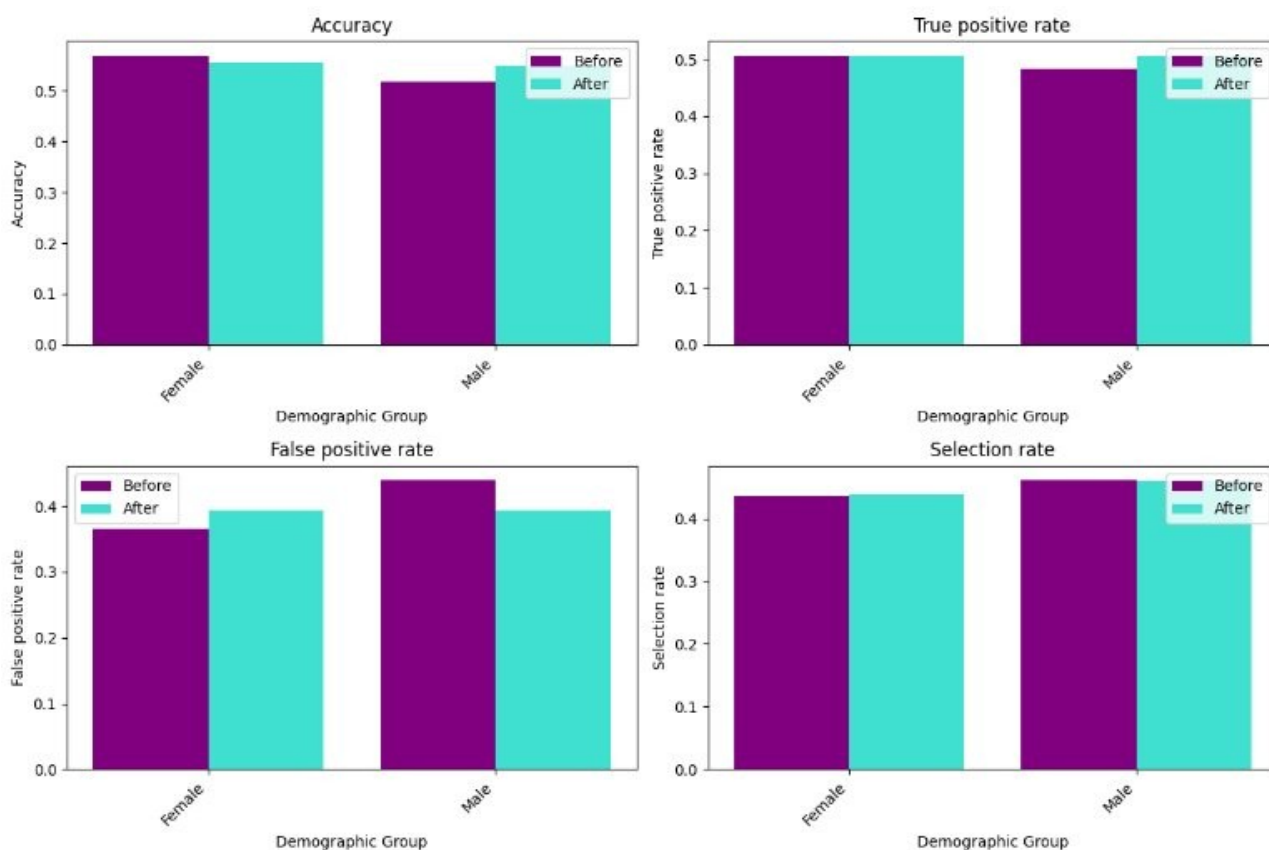


Εικόνα 7.1.11: Σύγκριση μέσων τιμών μετρικών απόδοσης των μοντέλων πριν και μετά την εφαρμογή μεθόδων εξάλειψης διακρίσεων για όλες τις κατηγορίες ευαίσθητων χαρακτηριστικών.

7.2 Αμεροληψία

7.2.1 Αμεροληψία ως προς το Φύλο

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



Εικόνα 7.2.1: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα του φύλου.

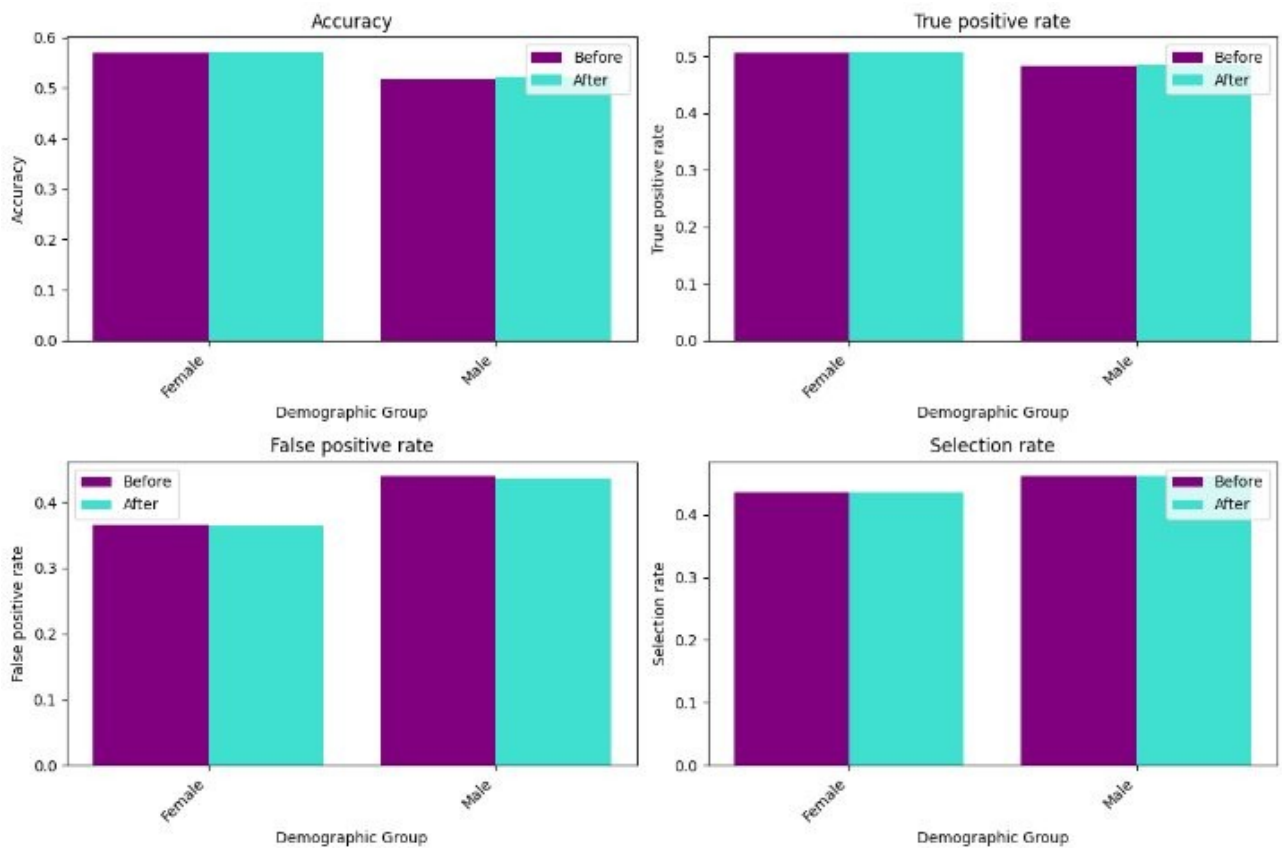
Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου πριν και μετά την αφαίρεση συσχετίσεων είναι:

Πίνακας 7.2.1: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων του φύλου.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων ως προς το Φύλο
Διαφορά Ισότητας Ακρίβειας	0,0512	0,3542
Αναλογία Ισότητας Ακρίβειας	0,9102	0,4848
Διαφορά Ισότητας Πιθανότητας	0,0737	0,0182
Αναλογία Ισότητας Πιθανότητας	0,8322	0,9558
Διαφορά Ίσης Ευκαιρίας	0,0226	0,0156
Αναλογία Ίσης Ευκαιρίας	0,9552	0,9698
Διαφορά Δημογραφικής Ισότητας	0,0261	0,3732
Αναλογία Δημογραφικής Ισότητας	0,9435	0,5406

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου ως ευαίσθητο

χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



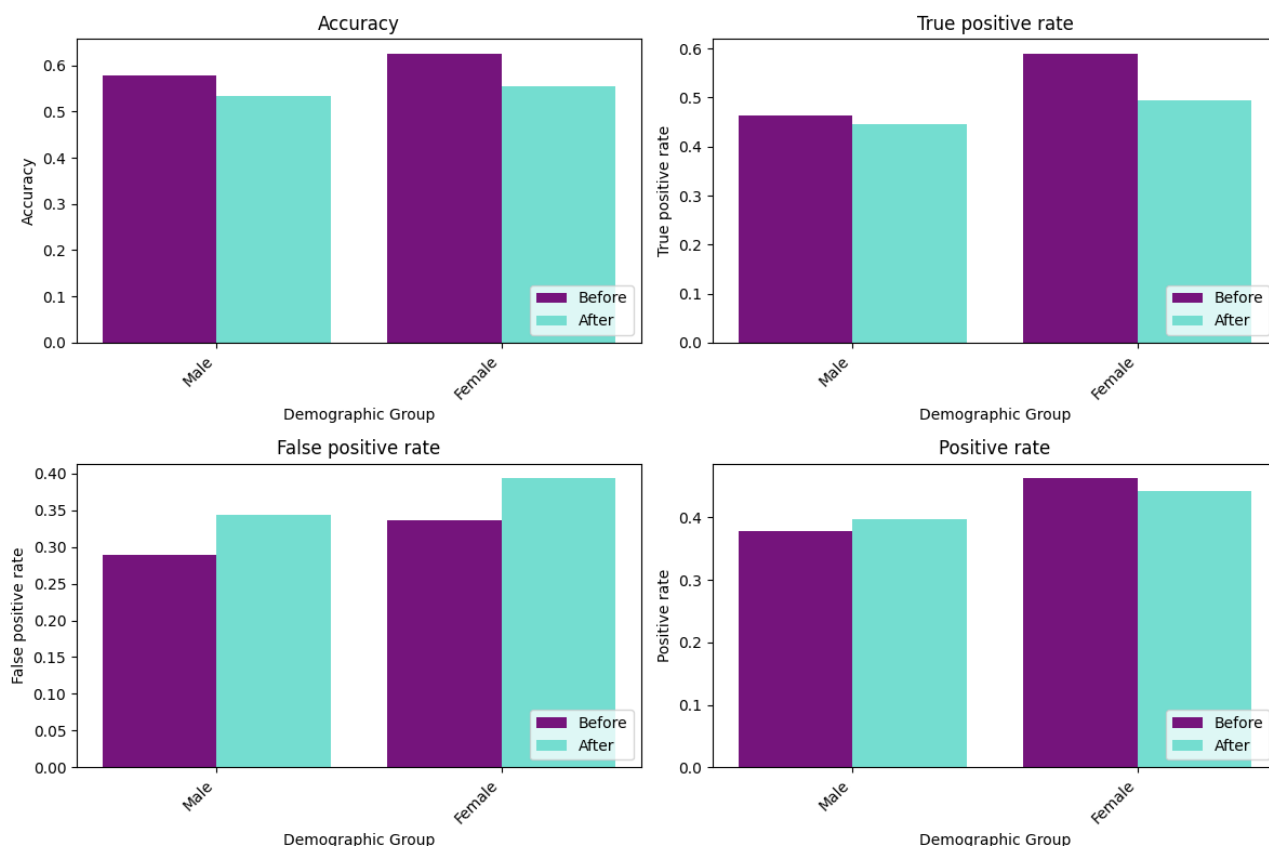
Εικόνα 7.2.2: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα του φύλου.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου πριν και μετά την μέθοδο εκθετικής κλίσης είναι:

Πίνακας 7.2.2: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων του φύλου.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Εκθετική Κλίση ως προς το Φύλο
Διαφορά Ισότητας Ακρίβειας	0,0512	0,05
Αναλογία Ισότητας Ακρίβειας	0,9102	0,911
Διαφορά Ισότητας Πιθανότητας	0,0737	0,0728
Αναλογία Ισότητας Πιθανότητας	0,8322	0,8335
Διαφορά Ίσης Ευκαιρίας	0,0226	0,023
Αναλογία Ίσης Ευκαιρίας	0,9552	0,954
Διαφορά Δημογραφικής Ισότητας	0,0261	0,025
Αναλογία Δημογραφικής Ισότητας	0,9435	0,945

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



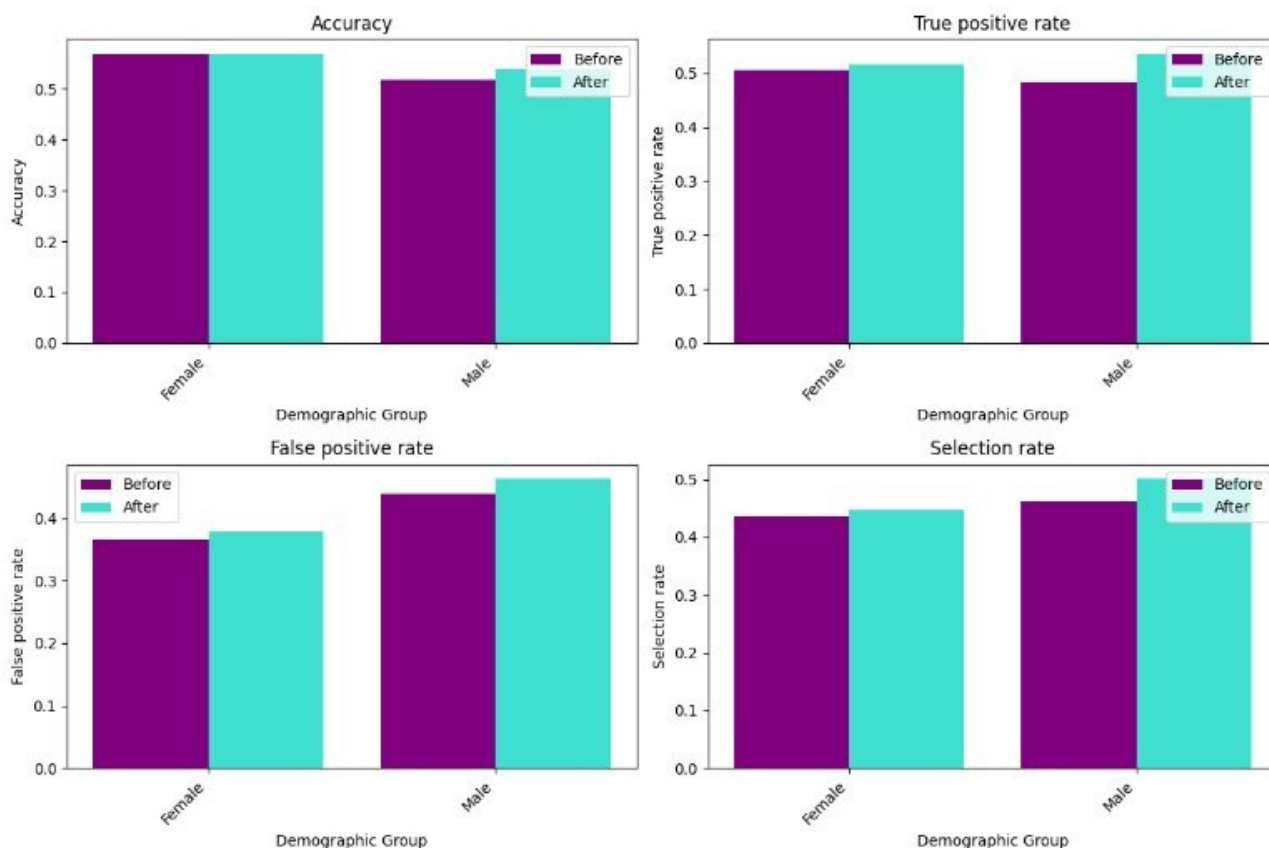
Εικόνα 7.2.3: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα του φύλου.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου πριν και μετά την μέθοδο ανταγωνιστικής εκπαίδευσης είναι:

Πίνακας 7.2.3: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων του φύλου.

Μετρικές Αμεροληψίας	Συνελκτικό Νευρωνικό Δίκτυο	μετά από Αντίστροφη Εκπαίδευση ως προς το Φύλο
Διαφορά Ισότητας Ακρίβειας	0,048	0,02
Αναλογία Ισότητας Ακρίβειας	0,92	0,96
Διαφορά Ισότητας Πιθανότητας	0,126	0,049
Αναλογία Ισότητας Πιθανότητας	0,785	0,874
Διαφορά Ίσης Ευκαιρίας	0,126	0,048
Αναλογία Ίσης Ευκαιρίας	0,785	0,903
Διαφορά Δημογραφικής Ισότητας	0,084	0,046
Αναλογία Δημογραφικής Ισότητας	0,817	0,895

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου του βελτιστοποιητή κατωφλίου για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



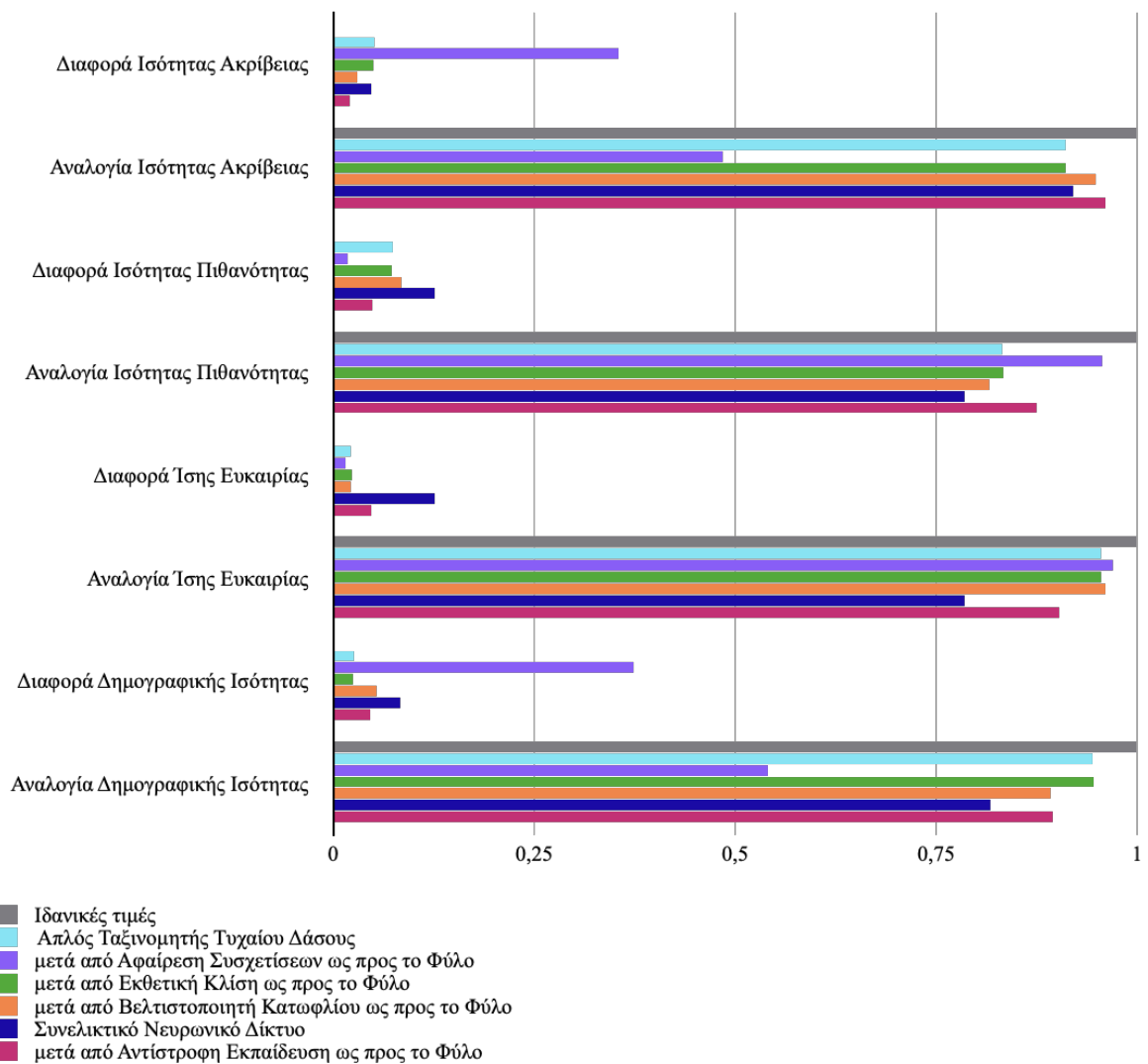
Εικόνα 7.2.4: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα του φύλου.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου πριν και μετά την μέθοδο βελτιστοποιητή κατωφλίου είναι:

Πίνακας 7.2.4: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων του φύλου.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Βελτιστοποιητή Κατωφλίου ως προς το Φύλο
Διαφορά Ισότητας Ακρίβειας	0,0512	0,0299
Αναλογία Ισότητας Ακρίβειας	0,9102	0,9474
Διαφορά Ισότητας Πιθανότητας	0,0737	0,0851
Αναλογία Ισότητας Πιθανότητας	0,8322	0,8163
Διαφορά Ίσης Ευκαιρίας	0,0226	0,0214
Αναλογία Ίσης Ευκαιρίας	0,9552	0,96
Διαφορά Δημογραφικής Ισότητας	0,0261	0,054
Αναλογία Δημογραφικής Ισότητας	0,9435	0,8921

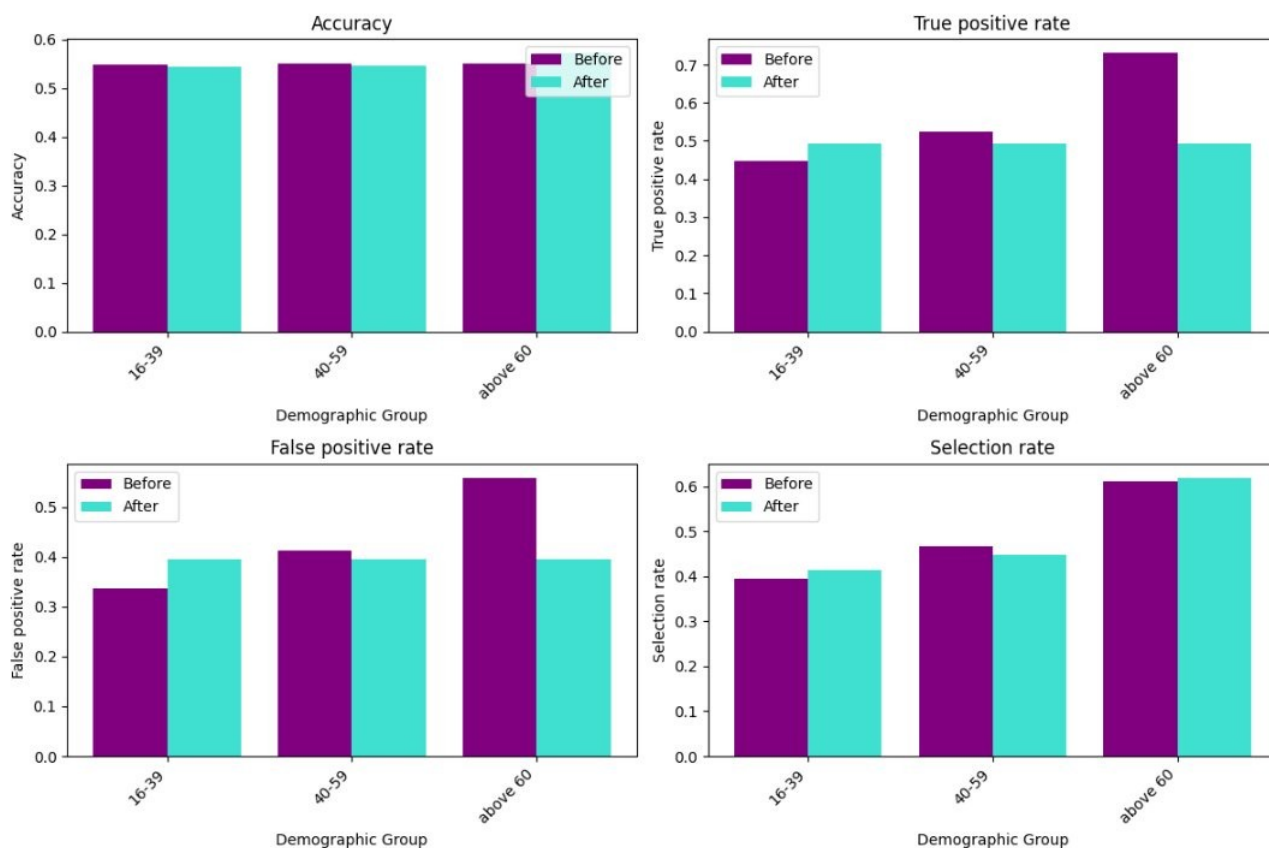
Συνολικά, οι τιμές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων του φύλου, συγκριτικά και με τις ιδανικές τιμές τους, παρουσιάζονται ακολούθως:



Εικόνα 7.2.5: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία του φύλου.

7.2.2 Αμεροληψία ως προς την Ηλικία

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



Εικόνα 7.2.6: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα της ηλικίας.

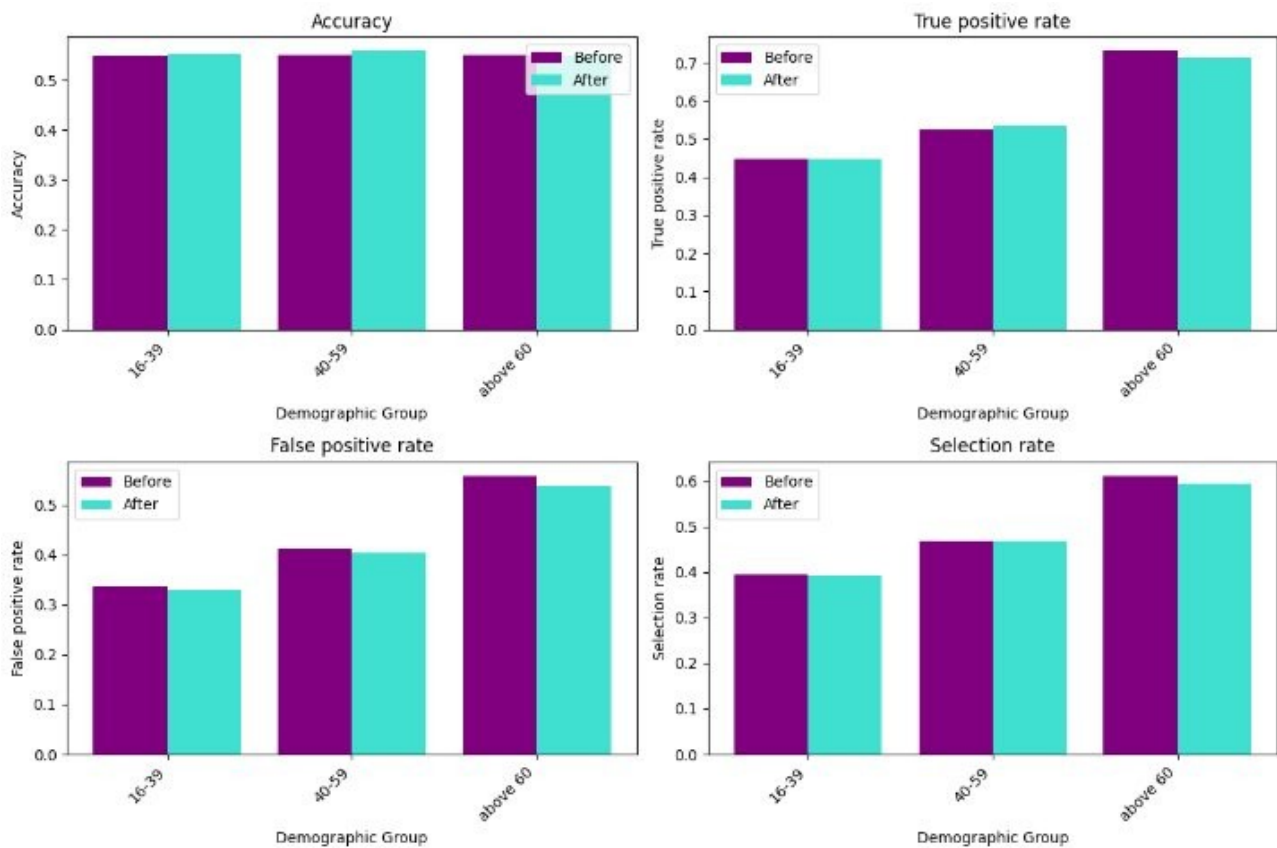
Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας πριν και μετά την αφαίρεση συσχετίσεων είναι:

Πίνακας 7.2.5: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων της ηλικίας.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων ως προς την Ηλικία
Διαφορά Ισότητας Ακρίβειας	0,0027	0,1638
Αναλογία Ισότητας Ακρίβειας	0,9951	0,7137
Διαφορά Ισότητας Πιθανότητας	0,2842	0,001
Αναλογία Ισότητας Πιθανότητας	0,6006	0,9979
Διαφορά Ίσης Ευκαιρίας	0,2842	0,001
Αναλογία Ίσης Ευκαιρίας	0,611	0,9979
Διαφορά Δημογραφικής Ισότητας	0,2179	0,2057
Αναλογία Δημογραφικής Ισότητας	0,6439	0,6679

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας ως ευαίσθητο

χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



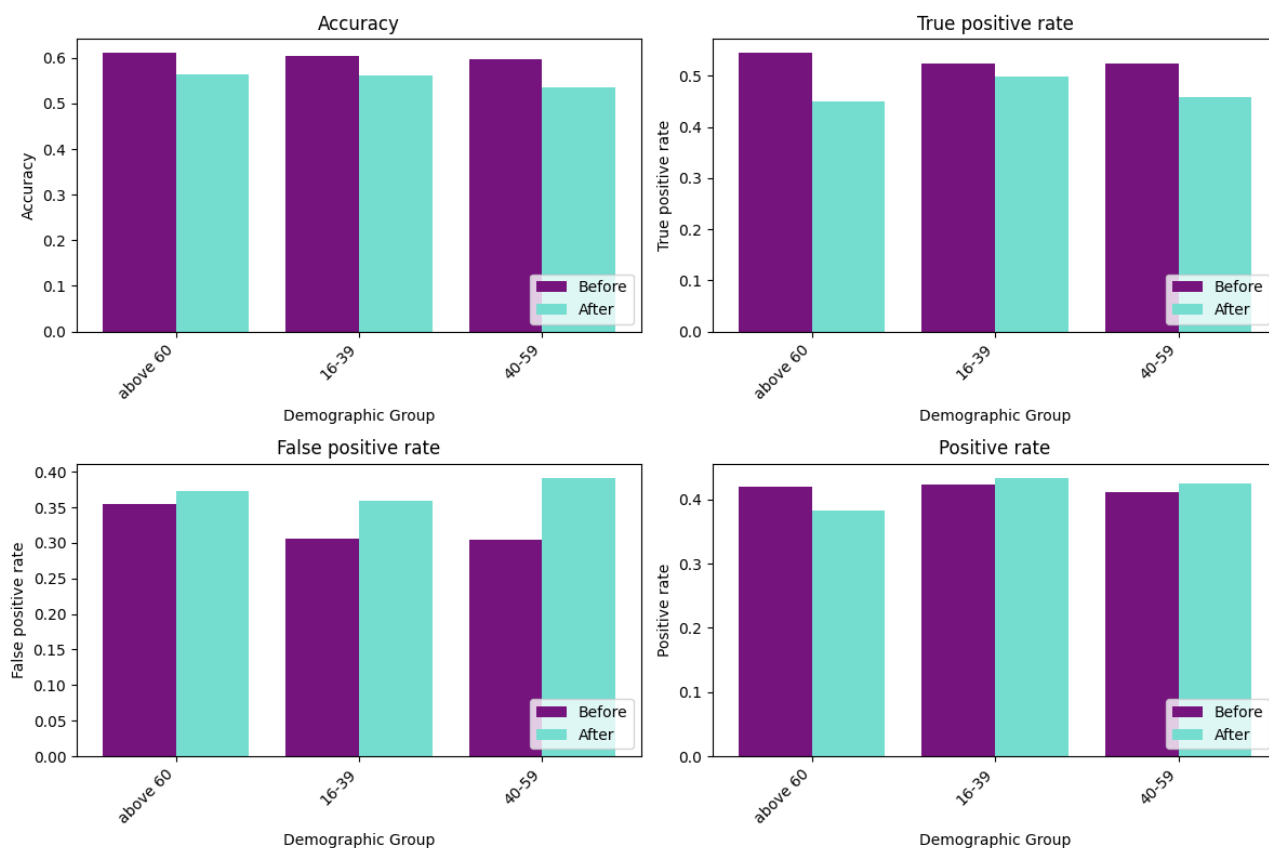
Εικόνα 7.2.7: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα της ηλικίας.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας πριν και μετά την μέθοδο εκθετικής κλίσης είναι:

Πίνακας 7.2.6: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων της ηλικίας.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Εκθετική Κλίση ως προς την Ηλικία
Διαφορά Ισότητας Ακρίβειας	0,0027	0,0103
Αναλογία Ισότητας Ακρίβειας	0,9951	0,9815
Διαφορά Ισότητας Πιθανότητας	0,2842	0,2654
Αναλογία Ισότητας Πιθανότητας	0,6006	0,6123
Διαφορά Ίσης Ευκαιρίας	0,2842	0,2654
Αναλογία Ίσης Ευκαιρίας	0,611	0,6278
Διαφορά Δημογραφικής Ισότητας	0,2179	0,2018
Αναλογία Δημογραφικής Ισότητας	0,6439	0,6604

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



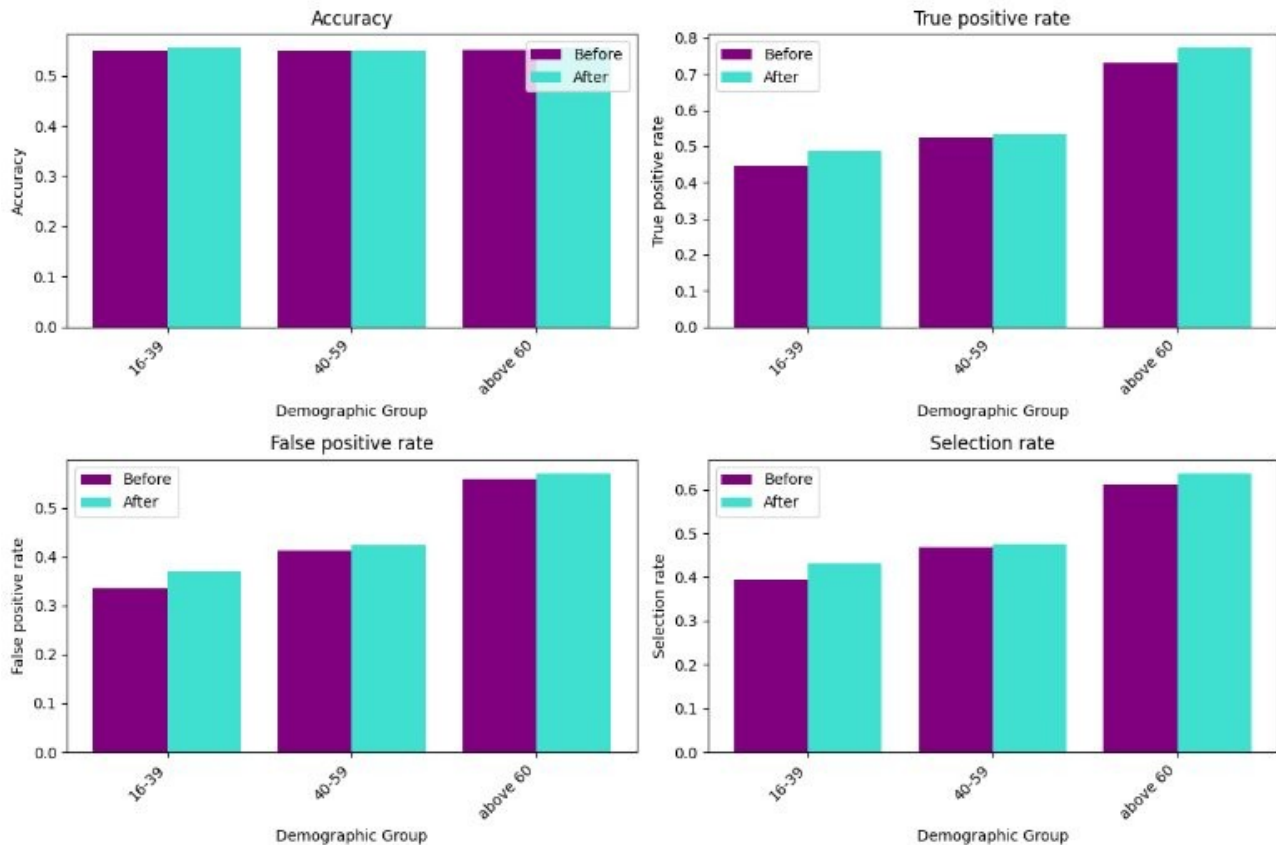
Εικόνα 7.2.8: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα της ηλικίας.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας πριν και μετά την μέθοδο ανταγωνιστικής εκπαίδευσης είναι:

Πίνακας 7.2.7: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων της ηλικίας.

Μετρικές Αμεροληψίας	Συνελκτικό Νευρωνικό Δίκτυο	μετά από Αντίστροφη Εκπαίδευση ως προς την Ηλικία
Διαφορά Ισότητας Ακρίβειας	0,012	0,029
Αναλογία Ισότητας Ακρίβειας	0,979	0,947
Διαφορά Ισότητας Πιθανότητας	0,05	0,049
Αναλογία Ισότητας Πιθανότητας	0,858	0,9
Διαφορά Ίσης Ευκαιρίας	0,02	0,049
Αναλογία Ίσης Ευκαιρίας	0,961	0,9
Διαφορά Δημογραφικής Ισότητας	0,012	0,051
Αναλογία Δημογραφικής Ισότητας	0,969	0,881

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου του βελτιστοποιητή κατωφλίου για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



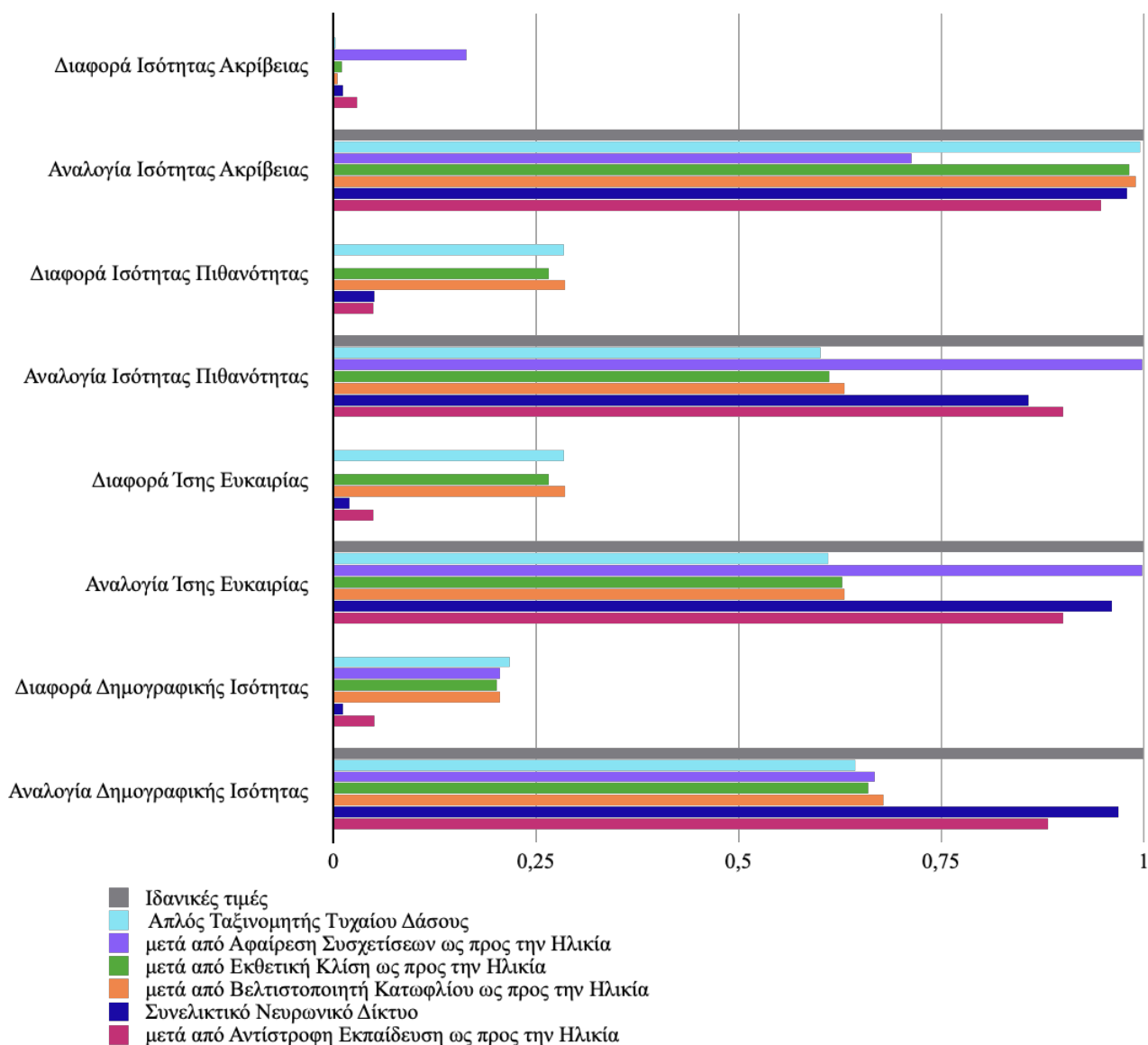
Εικόνα 7.2.9: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα της ηλικίας.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας πριν και μετά την μέθοδο βελτιστοποιητή κατωφλίου είναι:

Πίνακας 7.2.8: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων της ηλικίας.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Βελτιστοποιητή Κατωφλίου ως προς την Ηλικία
Διαφορά Ισότητας Ακρίβειας	0,0027	0,0057
Αναλογία Ισότητας Ακρίβειας	0,9951	0,9897
Διαφορά Ισότητας Πιθανότητας	0,2842	0,2858
Αναλογία Ισότητας Πιθανότητας	0,6006	0,63
Διαφορά Ίσης Ευκαιρίας	0,2842	0,2858
Αναλογία Ίσης Ευκαιρίας	0,611	0,63
Διαφορά Δημογραφικής Ισότητας	0,2179	0,2049
Αναλογία Δημογραφικής Ισότητας	0,6439	0,6782

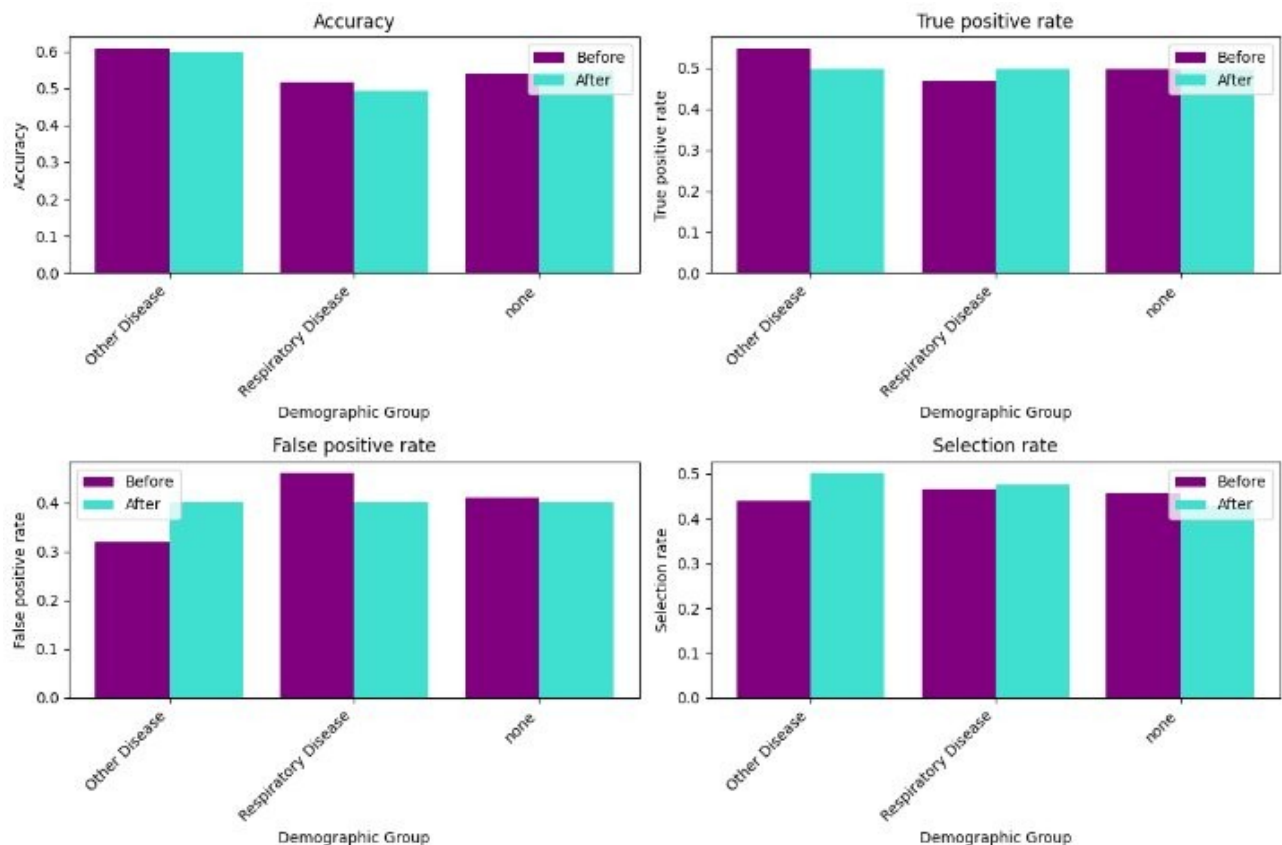
Συνολικά, οι τιμές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων της ηλικίας, συγκριτικά και με τις ιδανικές τιμές τους, παρουσιάζονται ακολούθως:



Εικόνα 7.2.10: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία της ηλικίας.

7.2.3 Αμεροληψία ως προς το Ιατρικό Ιστορικό

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του ιατρικού ιστορικού ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



Εικόνα 7.2.11: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα του ιατρικού ιστορικού.

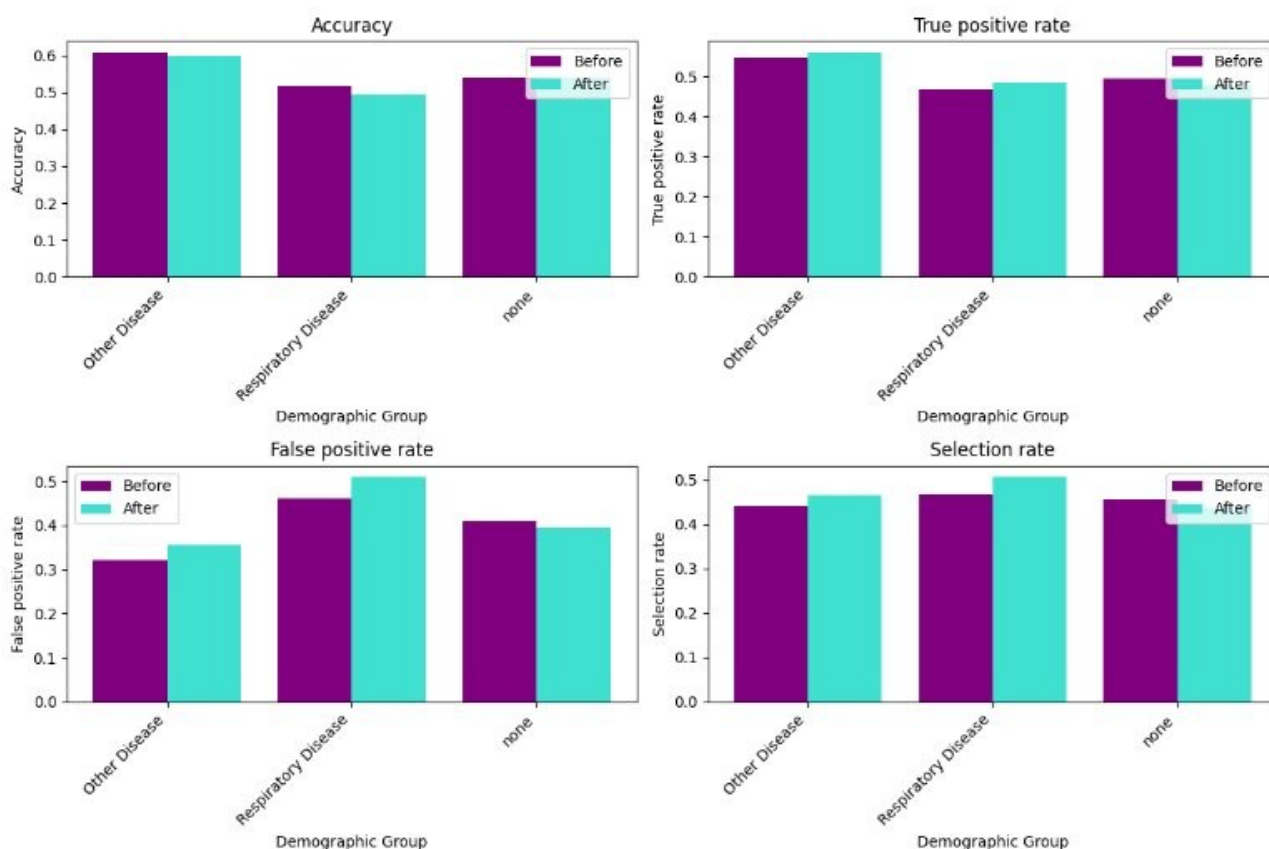
Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του ιατρικού ιστορικού πριν και μετά την αφαίρεση συσχετίσεων είναι:

Πίνακας 7.2.9: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων του ιατρικού ιστορικού.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων ως προς το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,0903	0,1036
Αναλογία Ισότητας Ακρίβειας	0,8517	0,8266
Διαφορά Ισότητας Πιθανότητας	0,1407	0,0158
Αναλογία Ισότητας Πιθανότητας	0,6948	0,962
Διαφορά Ίσης Ευκαιρίας	0,0804	0,014
Αναλογία Ίσης Ευκαιρίας	0,8533	0,9725
Διαφορά Δημογραφικής Ισότητας	0,0266	0,071
Αναλογία Δημογραφικής Ισότητας	0,9431	0,8585

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του ιατρικού ιστορικού ως

ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



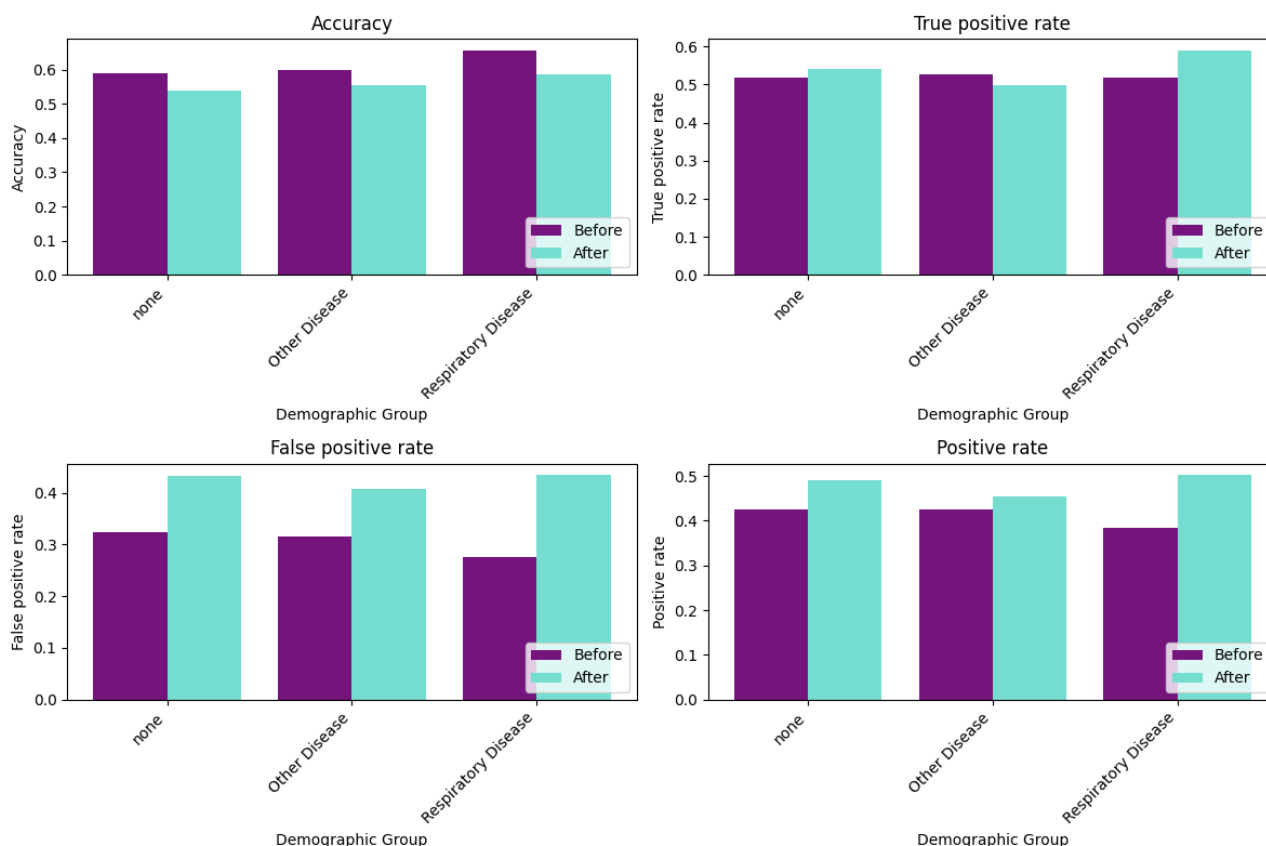
Εικόνα 7.2.12: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα του ιατρικού ιστορικού.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του ιατρικού ιστορικού πριν και μετά την μέθοδο εκθετικής κλίσης είναι:

Πίνακας 7.2.10: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων του ιατρικού ιστορικού.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Εκθετική Κλίση ως προς το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,0903	0,104
Αναλογία Ισότητας Ακρίβειας	0,8517	0,826
Διαφορά Ισότητας Πιθανότητας	0,1407	0,153
Αναλογία Ισότητας Πιθανότητας	0,6948	0,698
Διαφορά Ίσης Ευκαιρίας	0,0804	0,081
Αναλογία Ίσης Ευκαιρίας	0,8533	0,854
Διαφορά Δημογραφικής Ισότητας	0,0266	0,067
Αναλογία Δημογραφικής Ισότητας	0,9431	0,866

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του ιατρικού ιστορικού ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



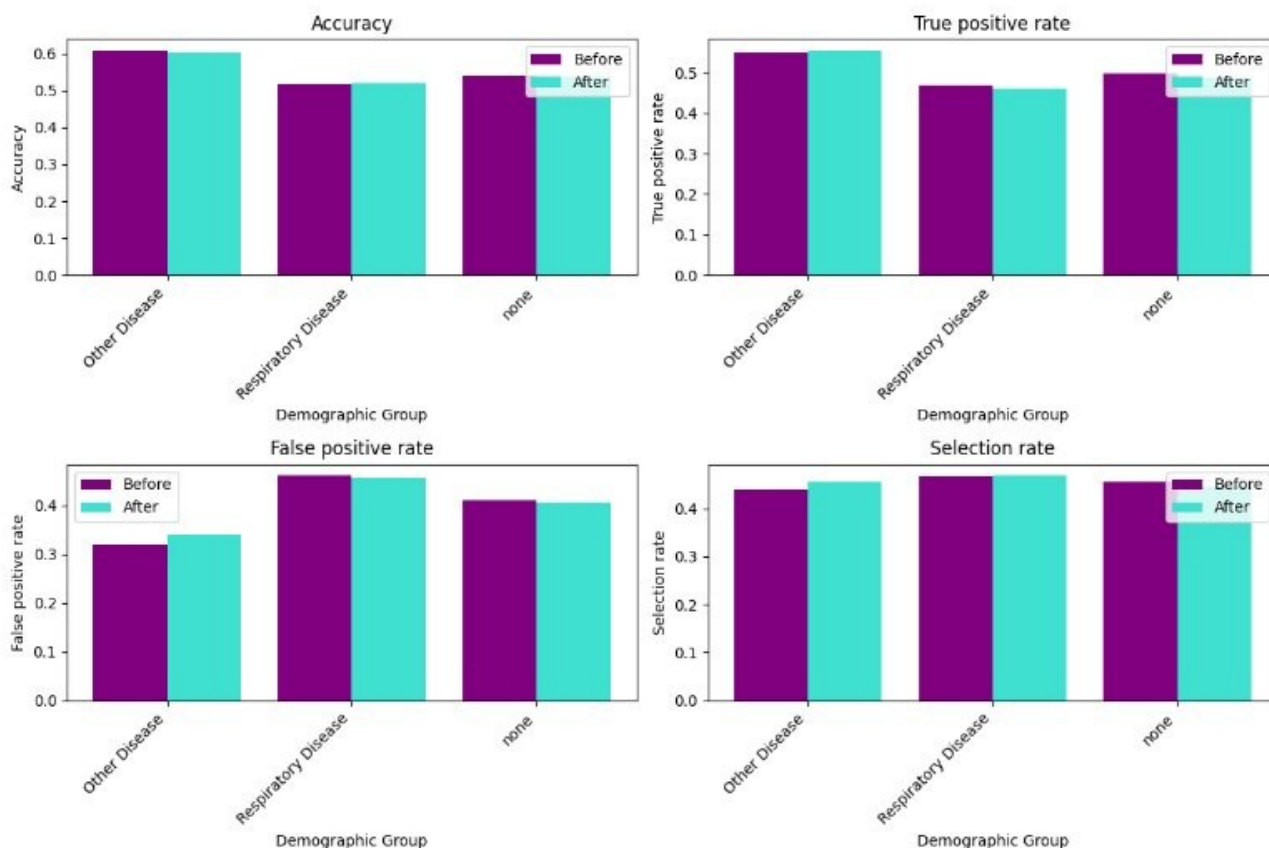
Εικόνα 7.2.13: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα του ιατρικού ιστορικού.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του ιατρικού ιστορικού πριν και μετά την μέθοδο ανταγωνιστικής εκπαίδευσης είναι:

Πίνακας 7.2.11: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων του ιατρικού ιστορικού.

Μετρικές Αμεροληψίας	Συνελικτικό Νευρωνικό Δίκτυο	μετά από Αντίστροφη Εκπαίδευση ως προς το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,067	0,045
Αναλογία Ισότητας Ακρίβειας	0,897	0,922
Διαφορά Ισότητας Πιθανότητας	0,046	0,092
Αναλογία Ισότητας Πιθανότητας	0,855	0,844
Διαφορά Ίσης Ευκαιρίας	0,011	0,092
Αναλογία Ίσης Ευκαιρίας	0,978	0,844
Διαφορά Δημογραφικής Ισότητας	0,041	0,049
Αναλογία Δημογραφικής Ισότητας	0,902	0,902

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του ιατρικού ιστορικού ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου του βελτιστοποιητή κατωφλίου για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



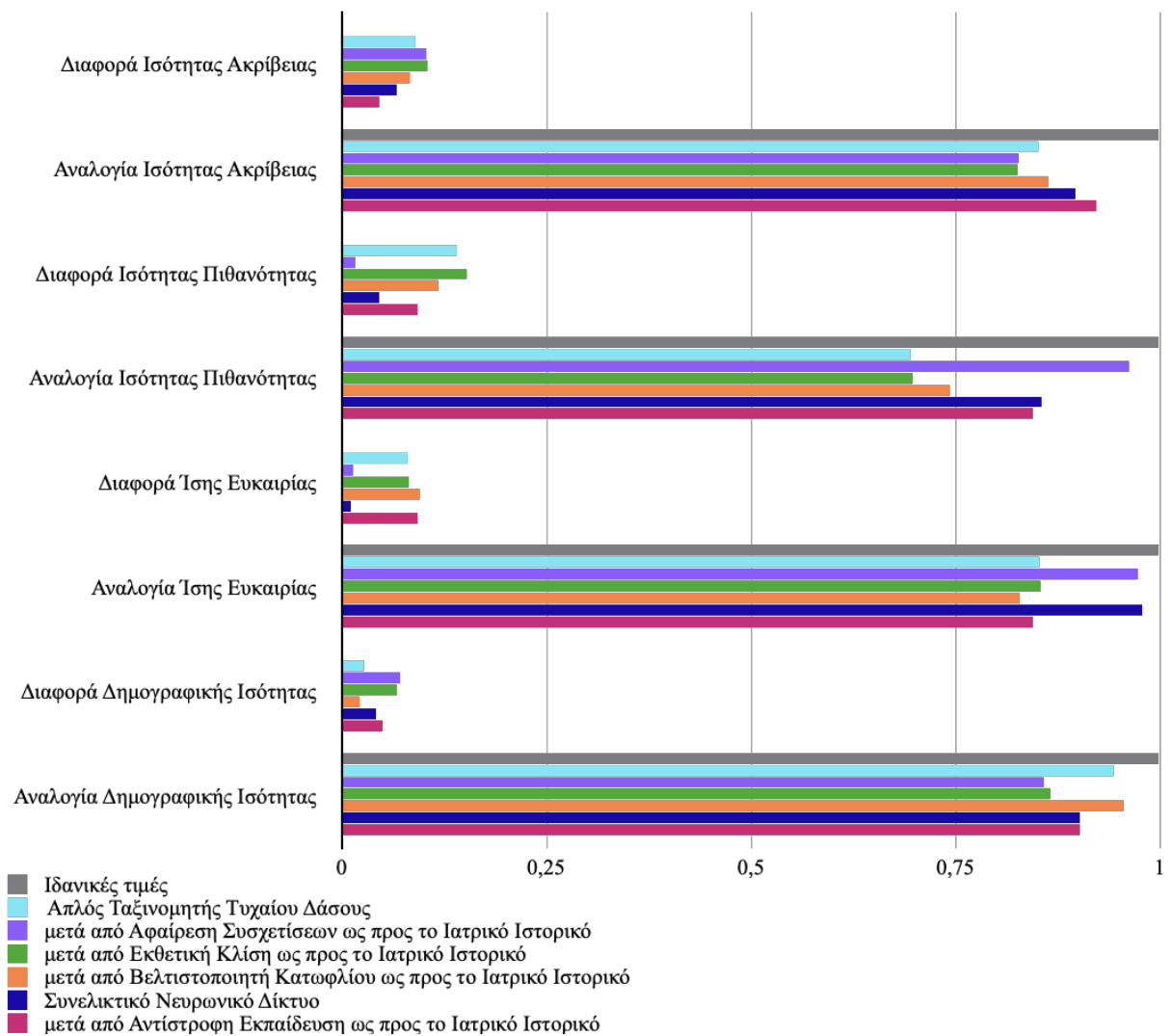
Εικόνα 7.2.14: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα του ιατρικού ιστορικού.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του ιατρικού ιστορικού πριν και μετά την μέθοδο βελτιστοποιητή κατωφλίου είναι:

Πίνακας 7.2.12: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων του ιατρικού ιστορικού.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Βελτιστοποιητή Κατωφλίου ως προς το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,0903	0,0824
Αναλογία Ισότητας Ακρίβειας	0,8517	0,8632
Διαφορά Ισότητας Πιθανότητας	0,1407	0,1178
Αναλογία Ισότητας Πιθανότητας	0,6948	0,7428
Διαφορά Ίσης Ευκαιρίας	0,0804	0,0953
Αναλογία Ίσης Ευκαιρίας	0,8533	0,8282
Διαφορά Δημογραφικής Ισότητας	0,0266	0,0211
Αναλογία Δημογραφικής Ισότητας	0,9431	0,955

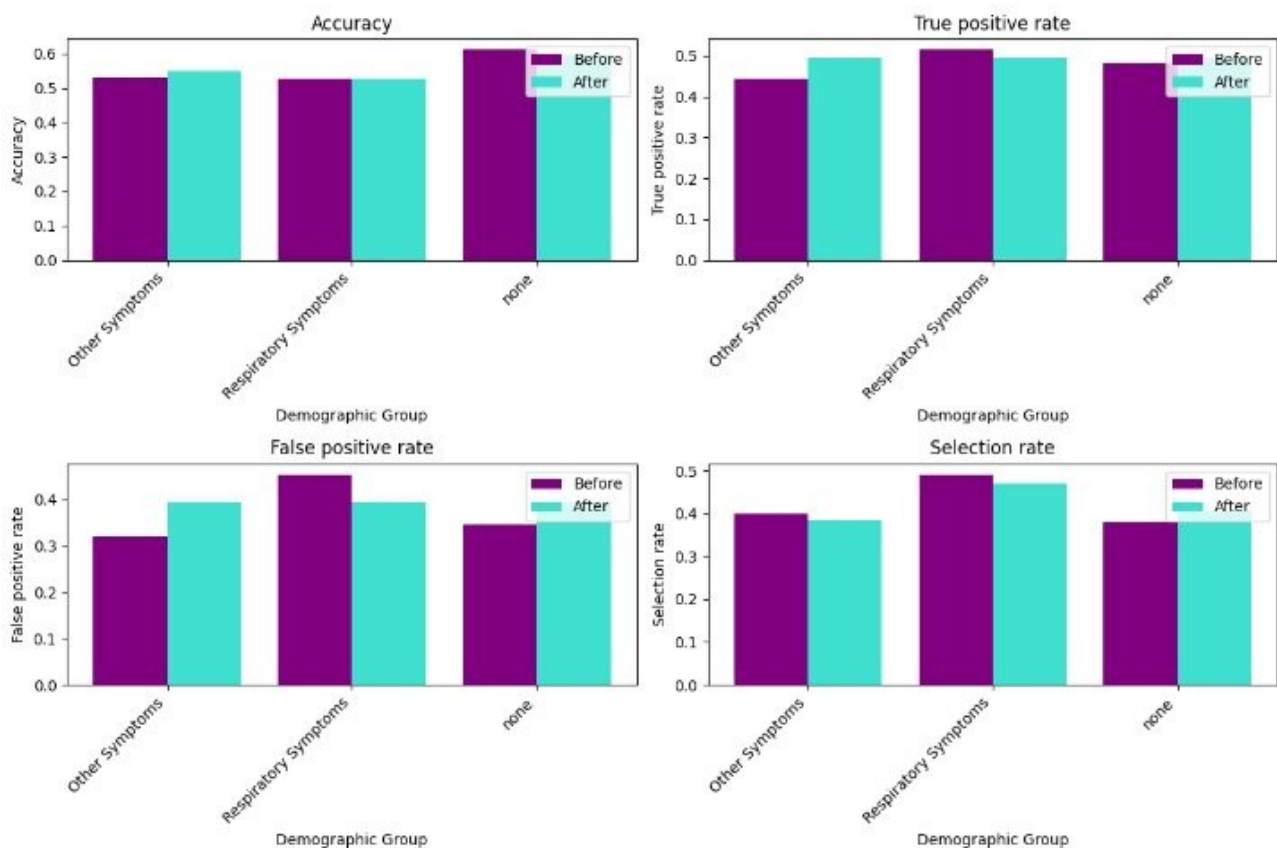
Συνολικά, οι τιμές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων του ιατρικού ιστορικού, συγκριτικά και με τις ιδανικές τιμές τους, παρουσιάζονται ακολούθως:



Εικόνα 7.2.15: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία του ιατρικού ιστορικού.

7.2.4 Αμεροληψία ως προς τα Συμπτώματα

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα των συμπτωμάτων ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



Εικόνα 7.2.16: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα των συμπτωμάτων.

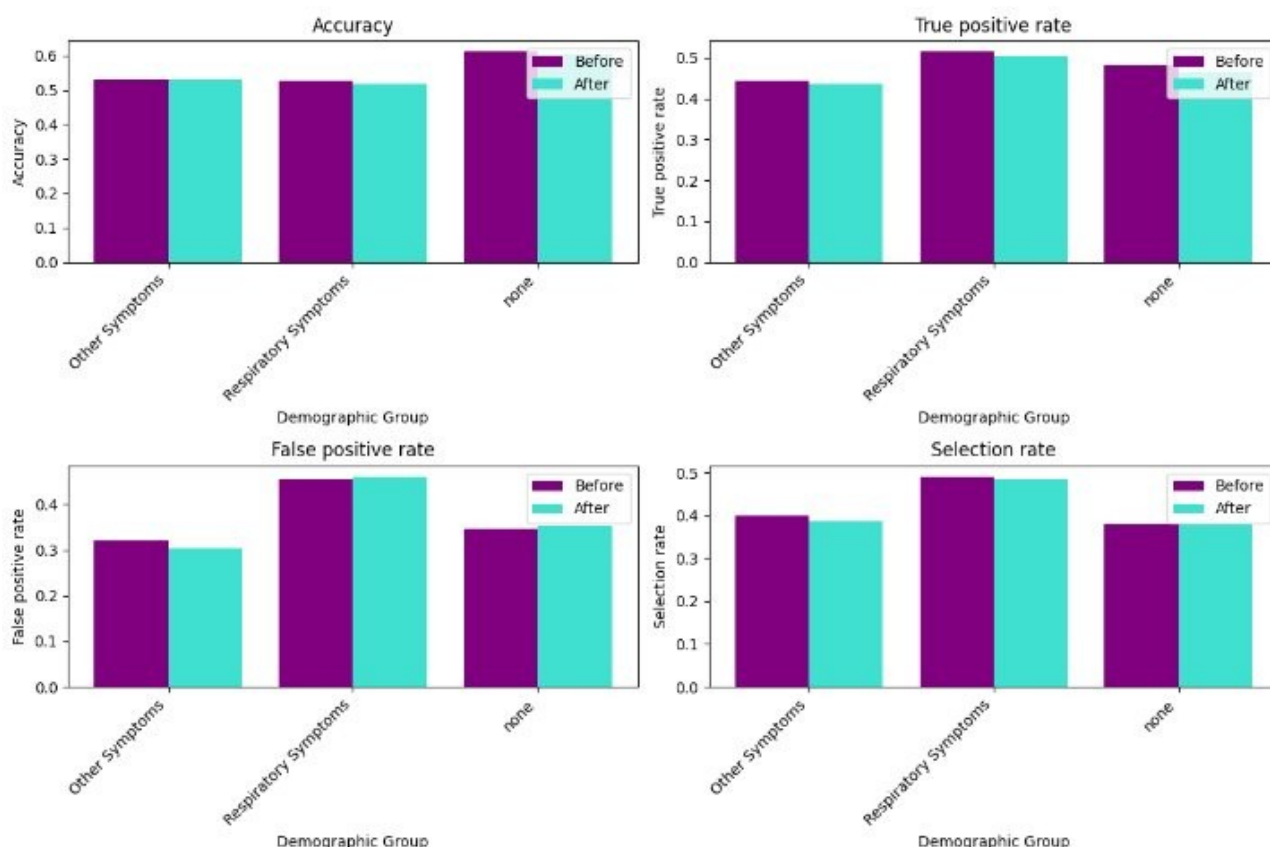
Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία των συμπτωμάτων πριν και μετά την αφαίρεση συσχετίσεων είναι:

Πίνακας 7.2.13: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων των συμπτωμάτων.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων ως προς τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,0852	0,0746
Αναλογία Ισότητας Ακρίβειας	0,8608	0,8748
Διαφορά Ισότητας Πιθανότητας	0,1347	0,0047
Αναλογία Ισότητας Πιθανότητας	0,7038	0,988
Διαφορά Ίσης Ευκαιρίας	0,0703	0,0013
Αναλογία Ίσης Ευκαιρίας	0,8634	0,9973
Διαφορά Δημογραφικής Ισότητας	0,1118	0,087
Αναλογία Δημογραφικής Ισότητας	0,7727	0,8156

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα των συμπτωμάτων ως

ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



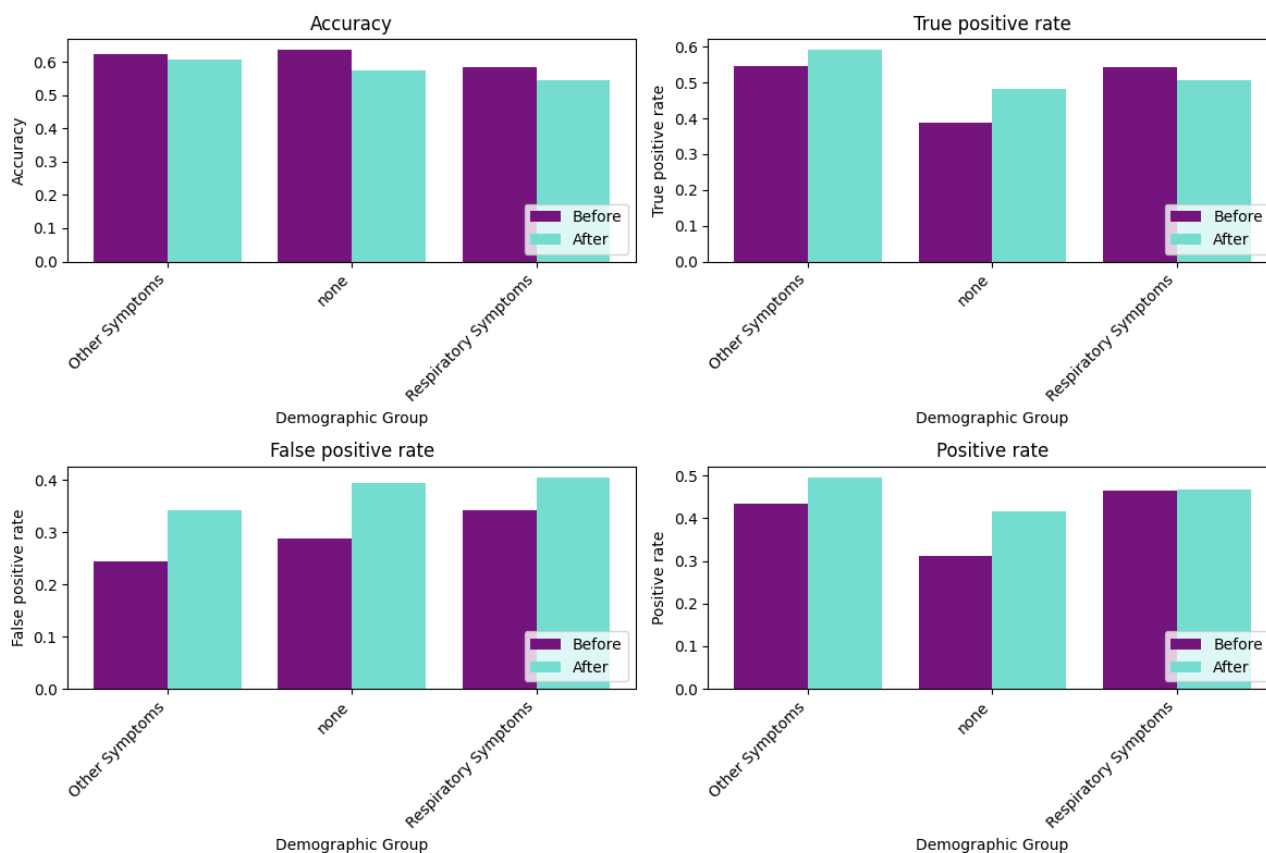
Εικόνα 7.2.17: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα των συμπτωμάτων.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία των συμπτωμάτων πριν και μετά την μέθοδο εκθετικής κλίσης είναι:

Πίνακας 7.2.14: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων των συμπτωμάτων.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Εκθετική Κλίση ως προς τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,0852	0,0853
Αναλογία Ισότητας Ακρίβειας	0,8608	0,8583
Διαφορά Ισότητας Πιθανότητας	0,1347	0,1583
Αναλογία Ισότητας Πιθανότητας	0,7038	0,6568
Διαφορά Ίσης Ευκαιρίας	0,0703	0,067
Αναλογία Ίσης Ευκαιρίας	0,8634	0,867
Διαφορά Δημογραφικής Ισότητας	0,1118	0,1063
Αναλογία Δημογραφικής Ισότητας	0,7727	0,7816

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα των συμπτωμάτων ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



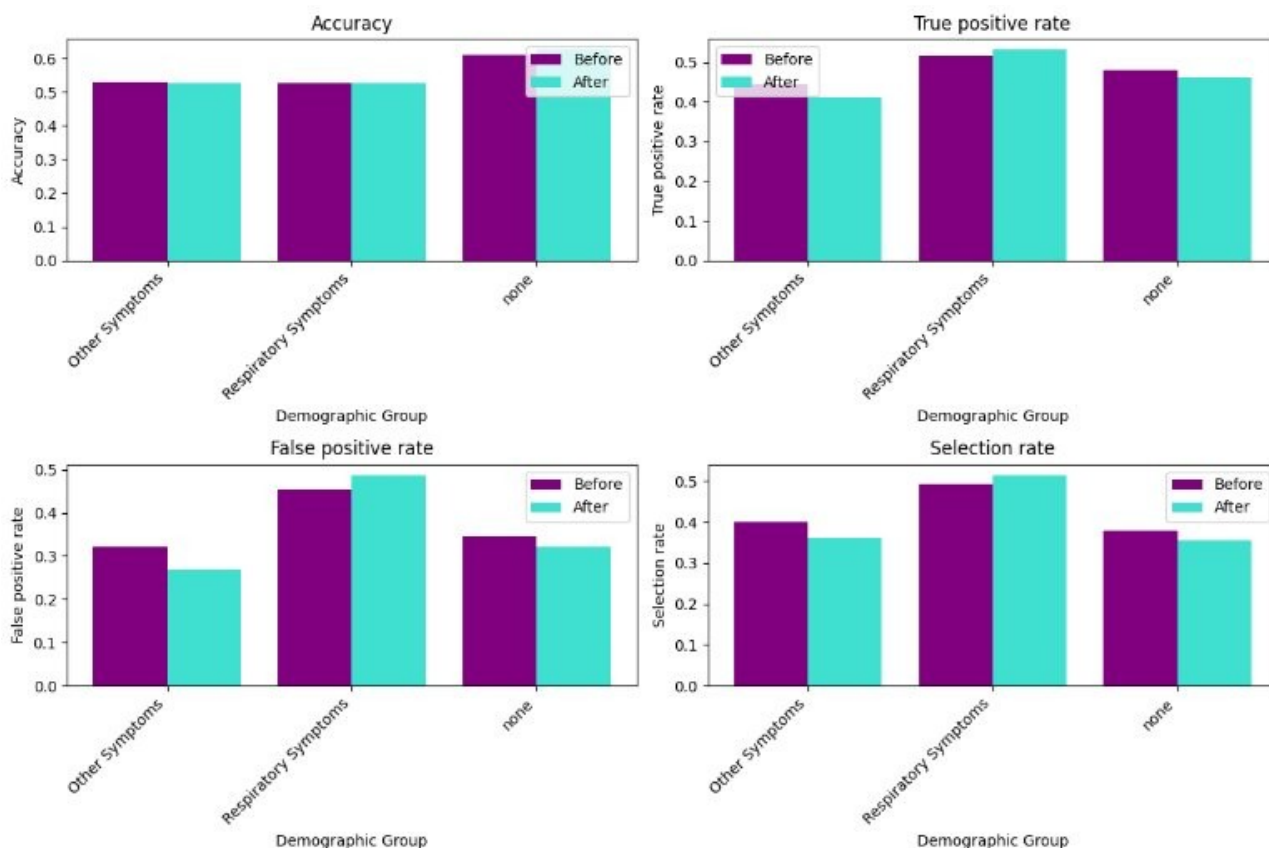
Εικόνα 7.2.18: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα των συμπτωμάτων.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία των συμπτωμάτων πριν και μετά την μέθοδο ανταγωνιστικής εκπαίδευσης είναι:

Πίνακας 7.2.15: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων των συμπτωμάτων.

Μετρικές Αμεροληψίας	Συνελικτικό Νευρωνικό Δίκτυο	μετά από Αντίστροφη Εκπαίδευση ως προς τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,053	0,064
Αναλογία Ισότητας Ακρίβειας	0,916	0,894
Διαφορά Ισότητας Πιθανότητας	0,156	0,109
Αναλογία Ισότητας Πιθανότητας	0,713	0,816
Διαφορά Ίσης Ευκαιρίας	0,1565	0,109
Αναλογία Ίσης Ευκαιρίας	0,713	0,816
Διαφορά Δημογραφικής Ισότητας	0,153	0,08
Αναλογία Δημογραφικής Ισότητας	0,916	0,838

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα των συμπτωμάτων ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου του βελτιστοποιητή κατωφλίου για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



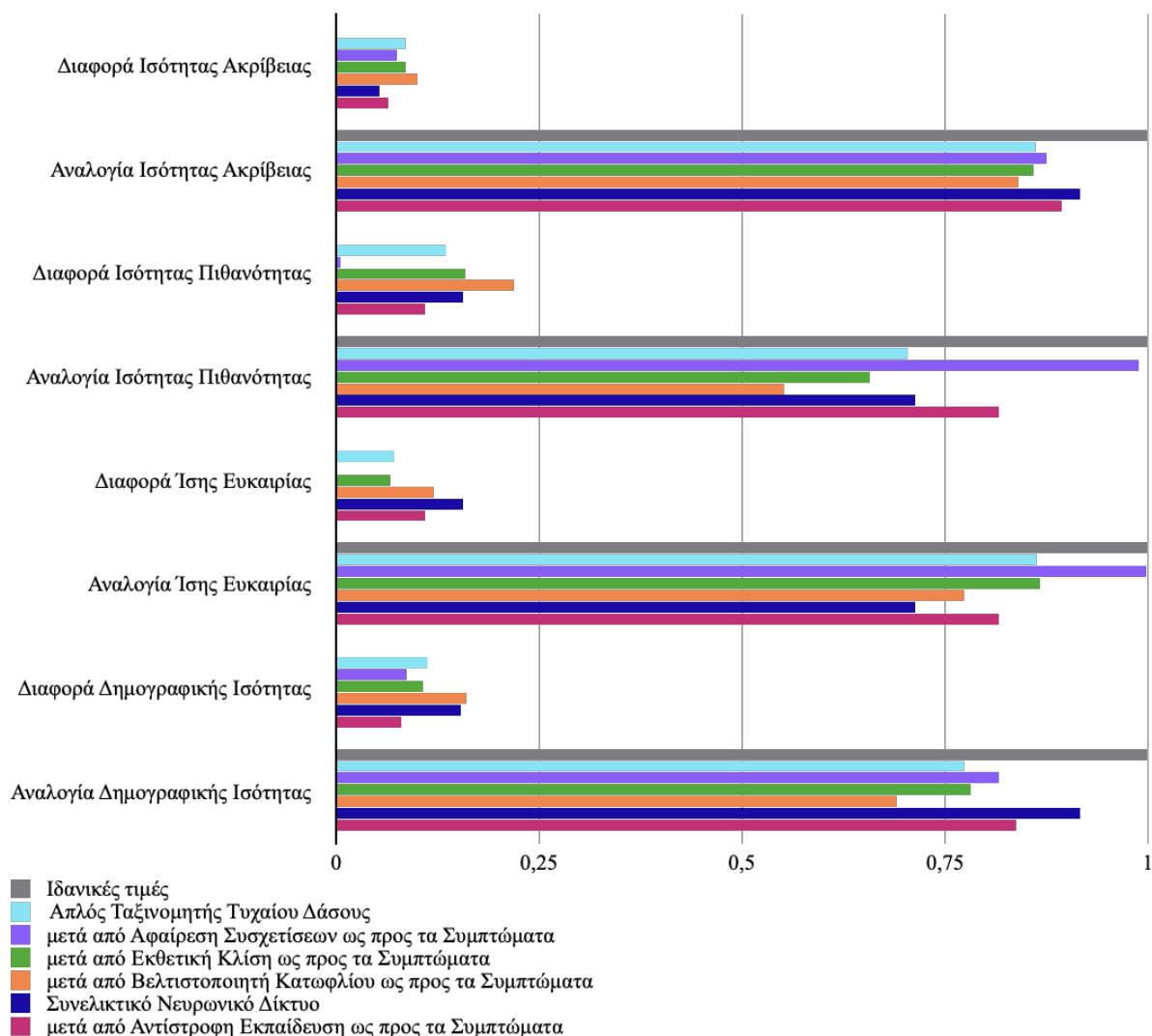
Εικόνα 7.2.19: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα των συμπτωμάτων.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία των συμπτωμάτων πριν και μετά την μέθοδο βελτιστοποιητή κατωφλίου είναι:

Πίνακας 7.2.16: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων των συμπτωμάτων.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Βελτιστοποιητή Κατωφλίου ως προς τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,0852	0,1002
Αναλογία Ισότητας Ακρίβειας	0,8608	0,8402
Διαφορά Ισότητας Πιθανότητας	0,1347	0,2185
Αναλογία Ισότητας Πιθανότητας	0,7038	0,5521
Διαφορά Ίσης Ευκαιρίας	0,0703	0,1204
Αναλογία Ίσης Ευκαιρίας	0,8634	0,7739
Διαφορά Δημογραφικής Ισότητας	0,1118	0,1594
Αναλογία Δημογραφικής Ισότητας	0,7727	0,6906

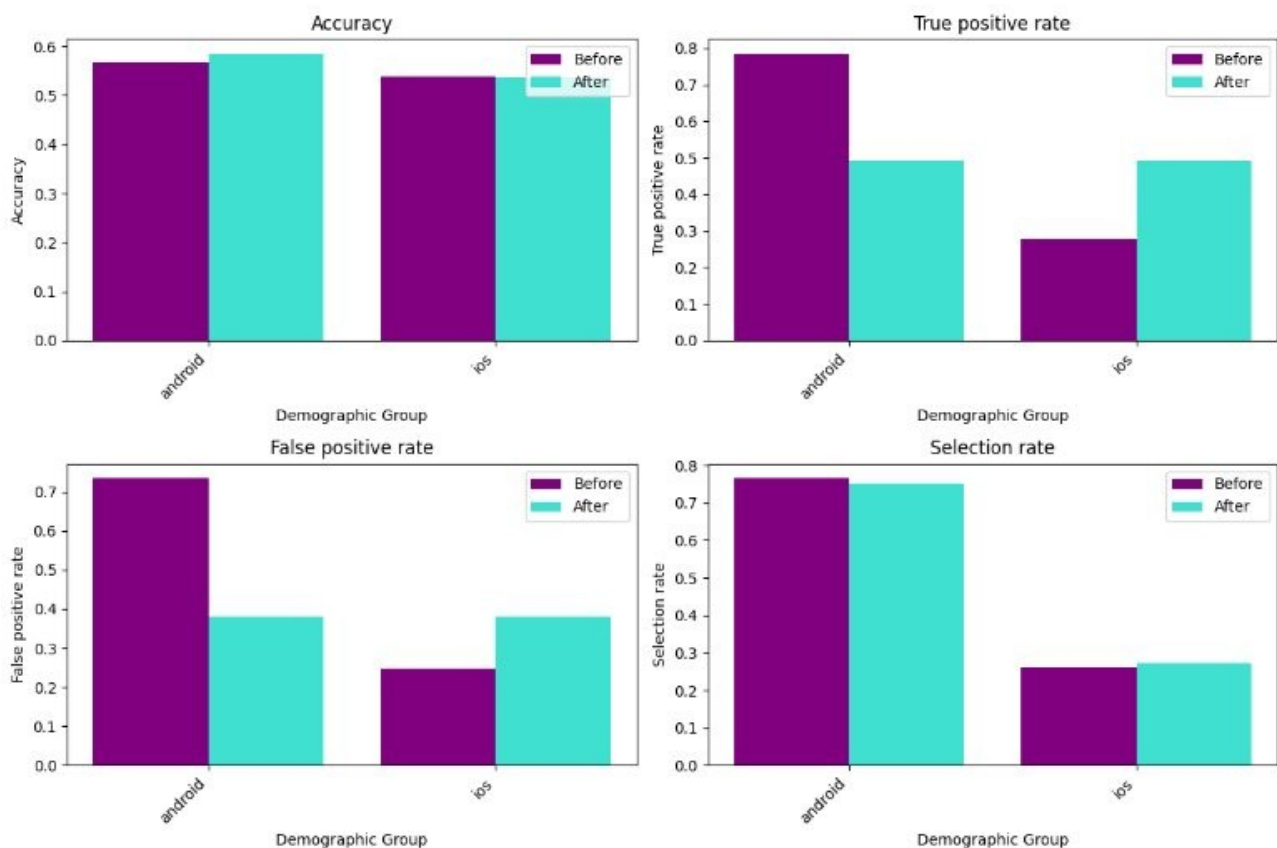
Συνολικά, οι τιμές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων των συμπτωμάτων, συγκριτικά και με τις ιδανικές τιμές τους, παρουσιάζονται ακολούθως:



Εικόνα 7.2.20: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία των συμπτωμάτων.

7.2.5 Αμεροληψία ως προς την Πηγή Συλλογής Δεδομένων

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της πηγής συλλογής δεδομένων ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



Εικόνα 7.2.21: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα της πηγής συλλογής δεδομένων.

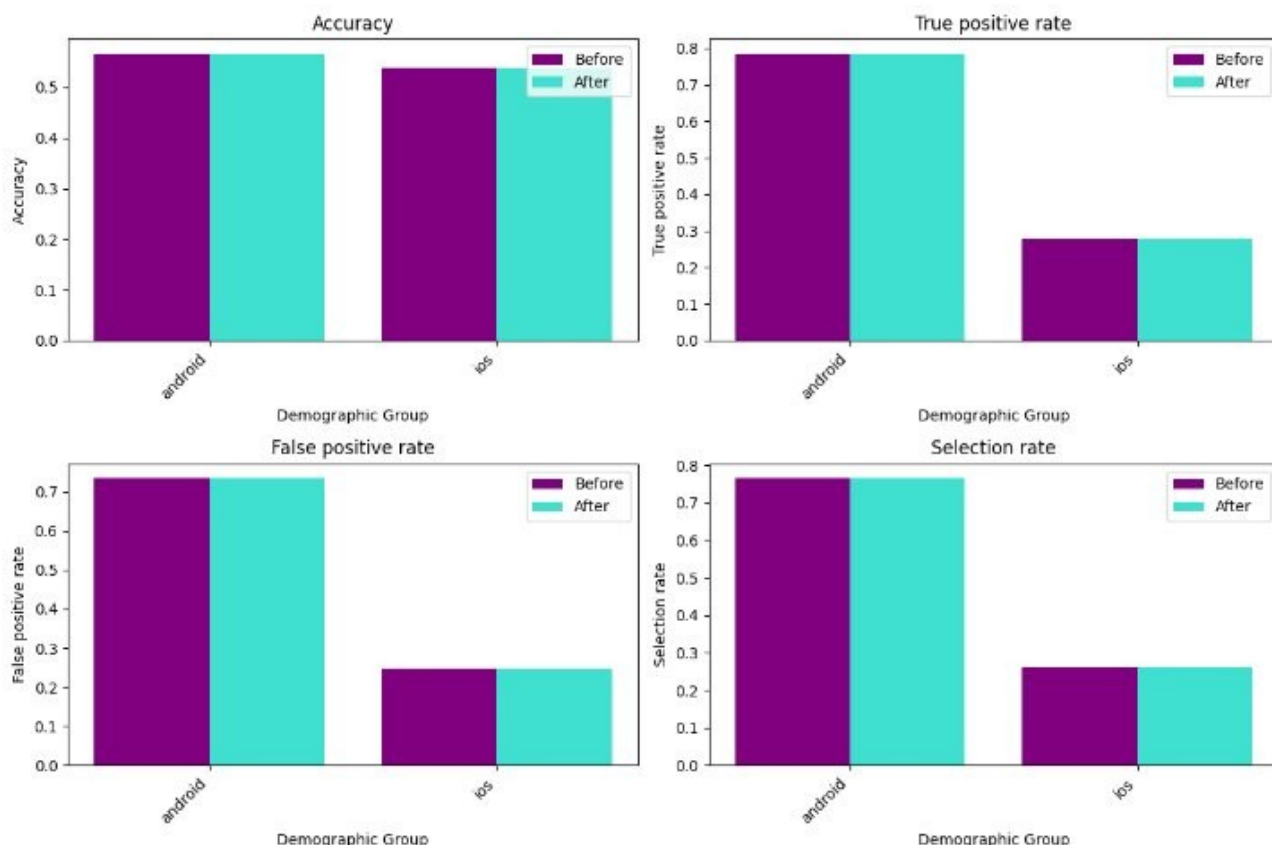
Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της πηγής συλλογής δεδομένων πριν και μετά την αφαίρεση συσχετίσεων είναι:

Πίνακας 7.2.17: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων της πηγής συλλογής δεδομένων.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων ως προς την Πηγή Συλλογής Δεδομένων
Διαφορά Ισότητας Ακρίβειας	0,0294	0,158
Αναλογία Ισότητας Ακρίβειας	0,9481	0,7725
Διαφορά Ισότητας Πιθανότητας	0,5067	0,0671
Αναλογία Ισότητας Πιθανότητας	0,3382	0,88
Διαφορά Ίσης Ευκαιρίας	0,5067	0,0671
Αναλογία Ίσης Ευκαιρίας	0,3556	0,88
Διαφορά Δημογραφικής Ισότητας	0,5032	0,4792
Αναλογία Δημογραφικής Ισότητας	0,3425	0,3626

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της πηγής συλλογής των

δεδομένων ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



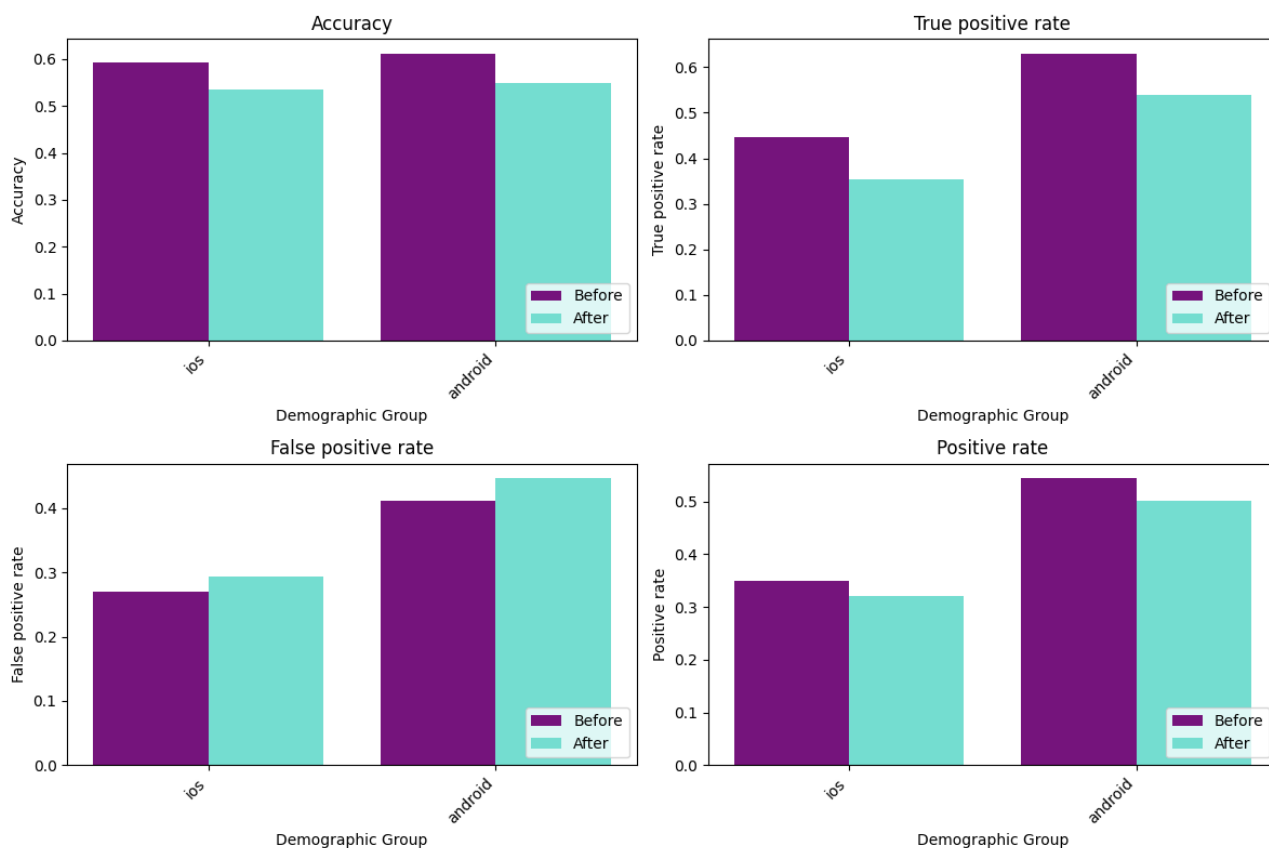
Εικόνα 7.2.22: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα της πηγής συλλογής δεδομένων.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της πηγής συλλογής δεδομένων πριν και μετά την μέθοδο εκθετικής κλίσης είναι:

Πίνακας 7.2.18: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων της πηγής συλλογής δεδομένων.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Εκθετική Κλίση ως προς την Πηγή Συλλογής Δεδομένων
Διαφορά Ισότητας Ακρίβειας	0,0294	0,0294
Αναλογία Ισότητας Ακρίβειας	0,9481	0,9481
Διαφορά Ισότητας Πιθανότητας	0,5067	0,5067
Αναλογία Ισότητας Πιθανότητας	0,3382	0,3382
Διαφορά Ίσης Ευκαιρίας	0,5067	0,5067
Αναλογία Ίσης Ευκαιρίας	0,3556	0,3556
Διαφορά Δημογραφικής Ισότητας	0,5032	0,5032
Αναλογία Δημογραφικής Ισότητας	0,3425	0,3425

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της πηγής συλλογής δεδομένων ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



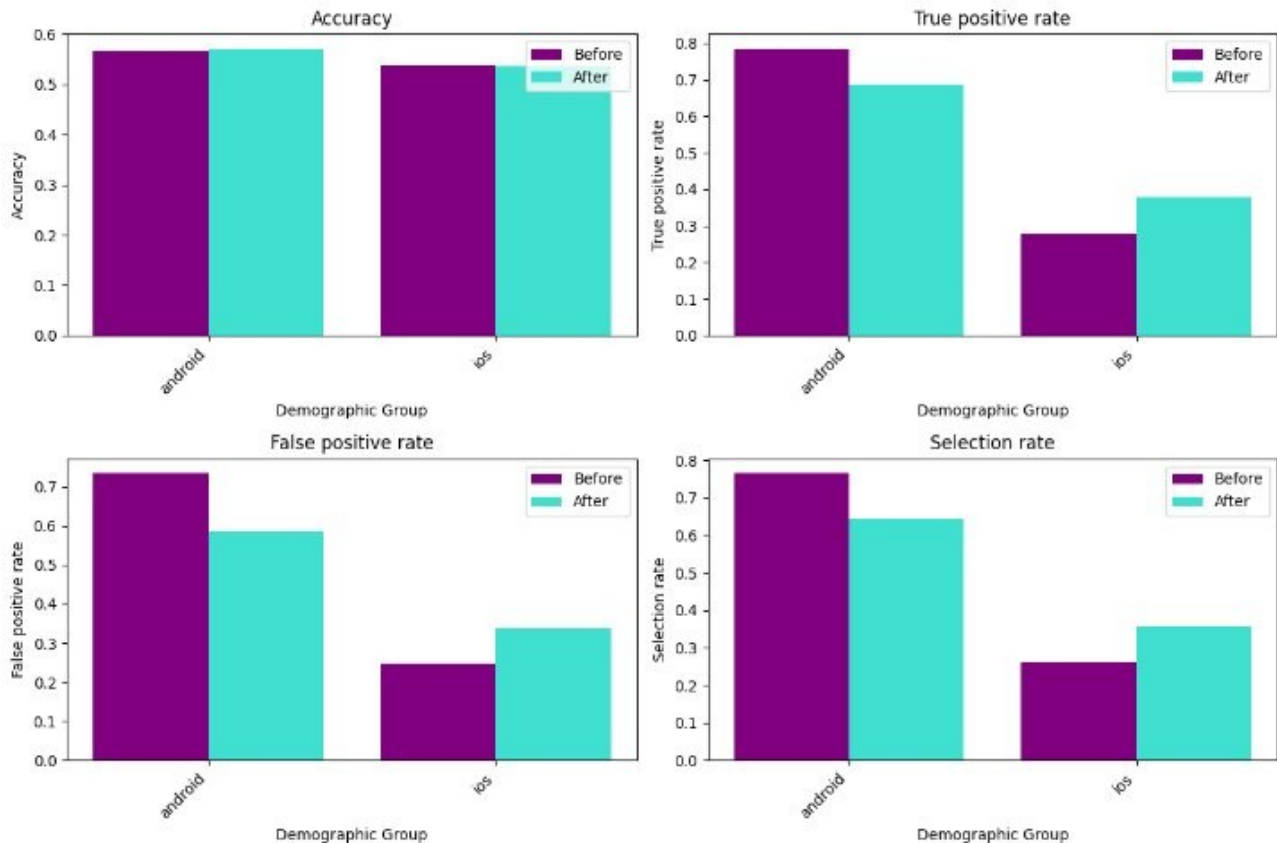
Εικόνα 7.2.23: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα της πηγής συλλογής δεδομένων.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της πηγής συλλογής δεδομένων πριν και μετά την μέθοδο ανταγωνιστικής εκπαίδευσης είναι:

Πίνακας 7.2.19: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων της πηγής συλλογής δεδομένων.

Μετρικές Αμεροληψίας	Συνελεκτικό Νευρωνικό Δίκτυο	μετά από Αντίστροφη Εκπαίδευση ως προς την Πηγή Συλλογής Δεδομένων
Διαφορά Ισότητας Ακρίβειας	0,0194	0,0140
Αναλογία Ισότητας Ακρίβειας	0,9683	0,9745
Διαφορά Ισότητας Πιθανότητας	0,1851	0,1852
Αναλογία Ισότητας Πιθανότητας	0,6569	0,6564
Διαφορά Ίσης Ευκαιρίας	0,1851	0,1852
Αναλογία Ίσης Ευκαιρίας	0,7064	0,6570
Διαφορά Δημογραφικής Ισότητας	0,1952	0,1825
Αναλογία Δημογραφικής Ισότητας	0,6415	0,6368

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της πηγής συλλογής δεδομένων ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου του βελτιστοποιητή κατωφλίου για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



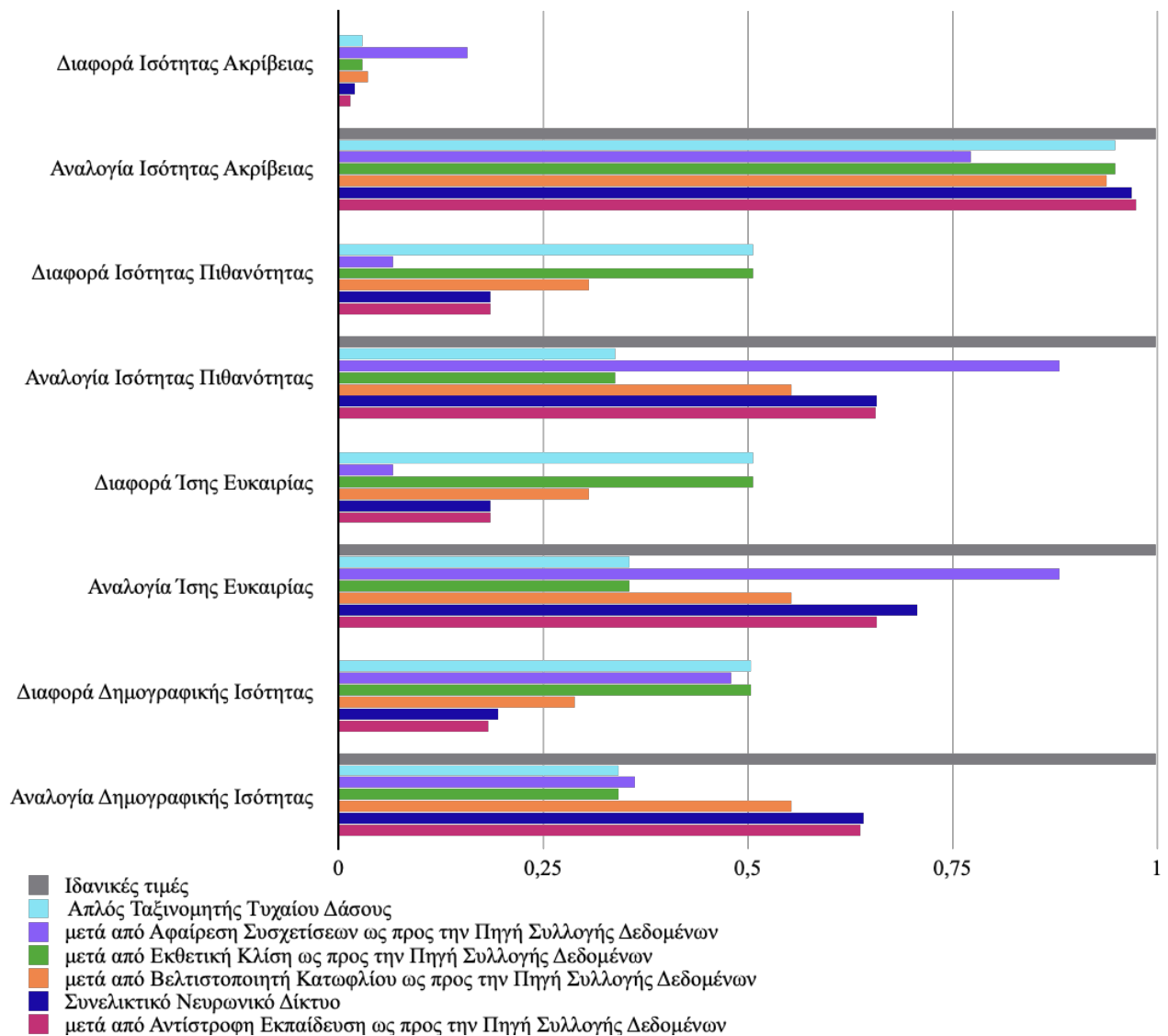
Εικόνα 7.2.24: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα της πηγής συλλογής δεδομένων.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της πηγής συλλογής δεδομένων πριν και μετά την μέθοδο βελτιστοποιητή κατωφλίου είναι:

Πίνακας 7.2.20: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων της πηγής συλλογής δεδομένων.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Βελτιστοποιητή Κατωφλίου ως προς την Πηγή Συλλογής Δεδομένων
Διαφορά Ισότητας Ακρίβειας	0,0294	0,0354
Αναλογία Ισότητας Ακρίβειας	0,9481	0,938
Διαφορά Ισότητας Πιθανότητας	0,5067	0,3065
Αναλογία Ισότητας Πιθανότητας	0,3382	0,5532
Διαφορά Ίσης Ευκαιρίας	0,5067	0,3065
Αναλογία Ίσης Ευκαιρίας	0,3556	0,5532
Διαφορά Δημογραφικής Ισότητας	0,5032	0,2884
Αναλογία Δημογραφικής Ισότητας	0,3425	0,5529

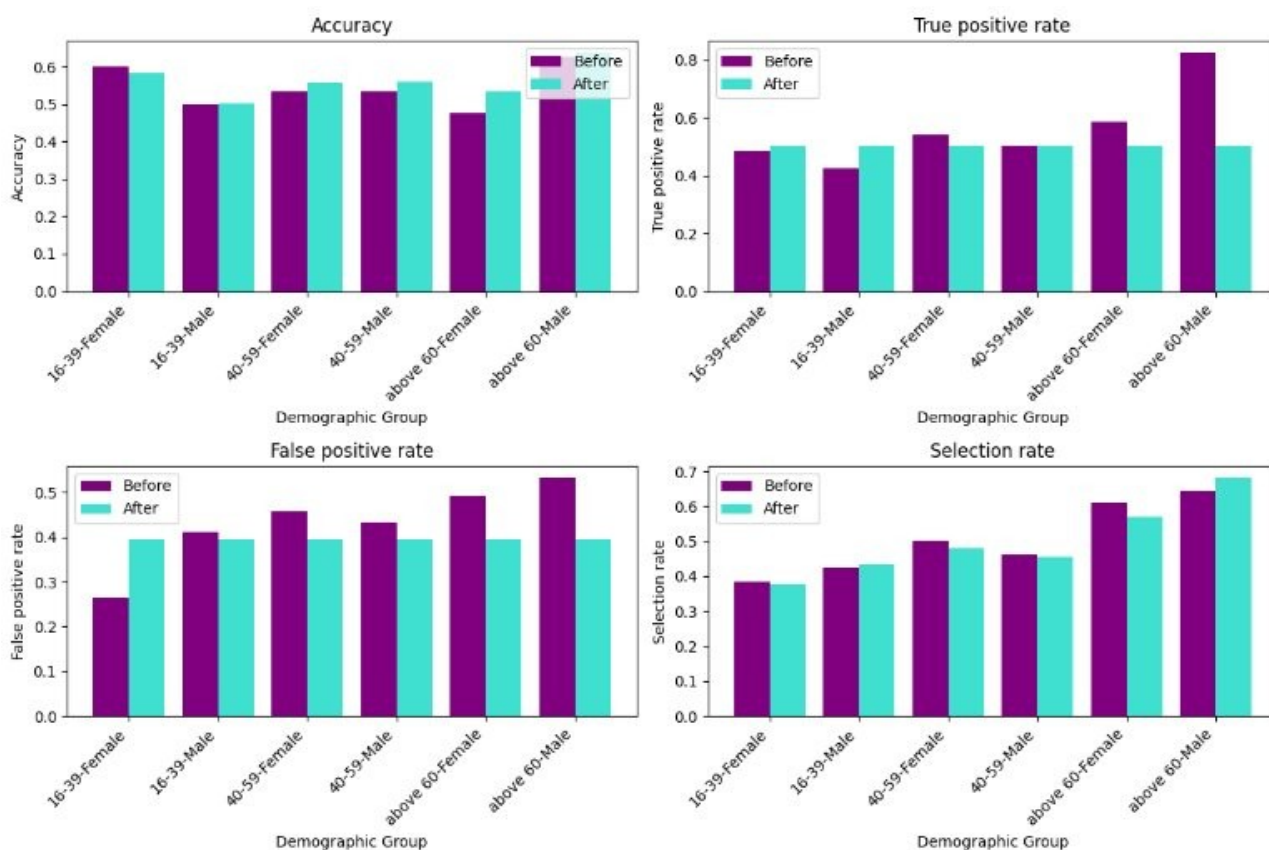
Συνολικά, οι τιμές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων ως προς την πηγή συλλογής δεδομένων, συγκριτικά και με τις ιδανικές τιμές τους, παρουσιάζονται ακολούθως:



Εικόνα 7.2.25: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία της πηγής συλλογής δεδομένων.

7.2.6 Αμεροληψία ως προς την Ηλικία σε συνδυασμό με το Φύλο

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με το φύλο ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



Εικόνα 7.2.26: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με το φύλο.

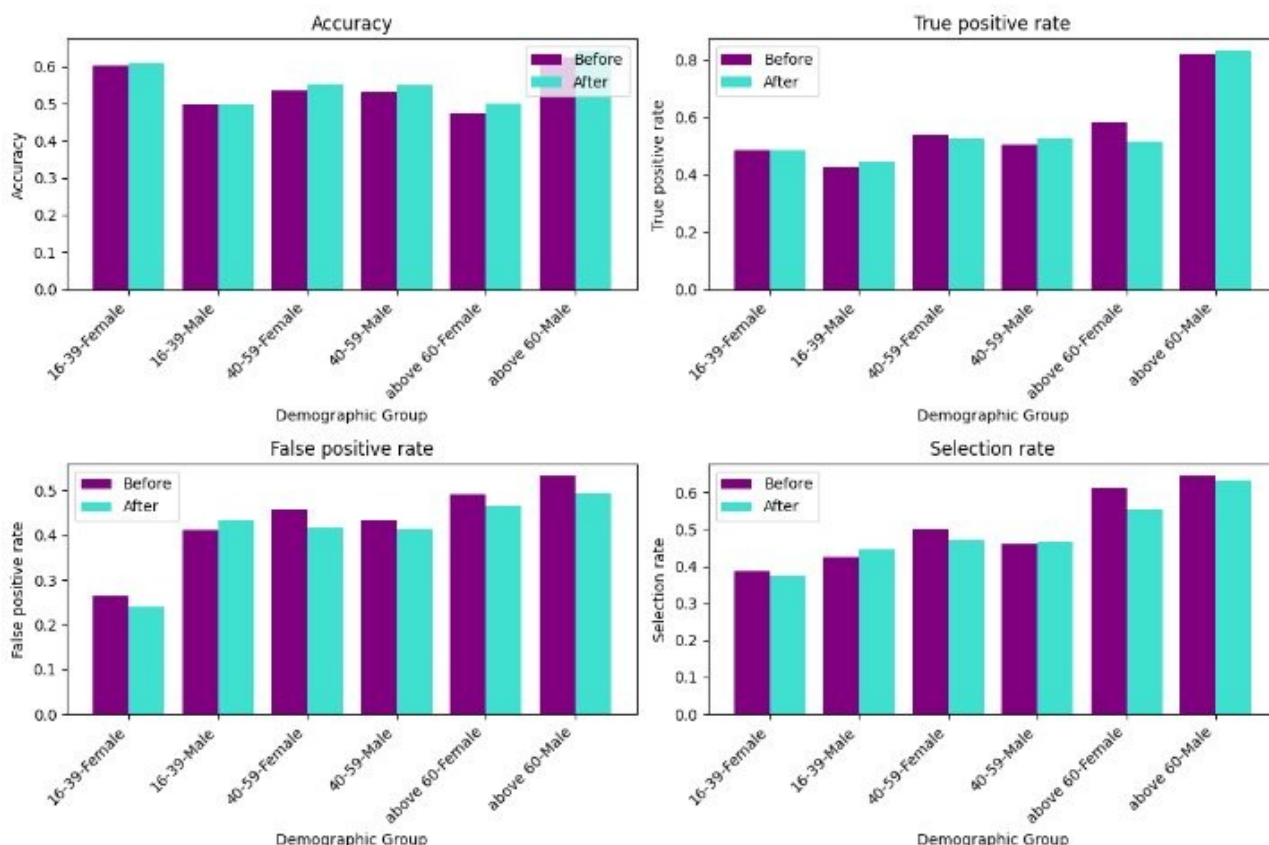
Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με το φύλο πριν και μετά την αφαίρεση συσχετίσεων είναι:

Πίνακας 7.2.21: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με το φύλο.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων ως προς την Ηλικία σε συνδυασμό με το Φύλο
Διαφορά Ισότητας Ακρίβειας	0,151	0,7143
Αναλογία Ισότητας Ακρίβειας	0,759	0
Διαφορά Ισότητας Πιθανότητας	0,3981	0,0837
Αναλογία Ισότητας Πιθανότητας	0,4953	0,8037
Διαφορά Ίσης Ευκαιρίας	0,3981	0,0283
Αναλογία Ίσης Ευκαιρίας	0,5164	0,9438
Διαφορά Δημογραφικής Ισότητας	0,2593	1
Αναλογία Δημογραφικής Ισότητας	0,5981	0

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με

το φύλο ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



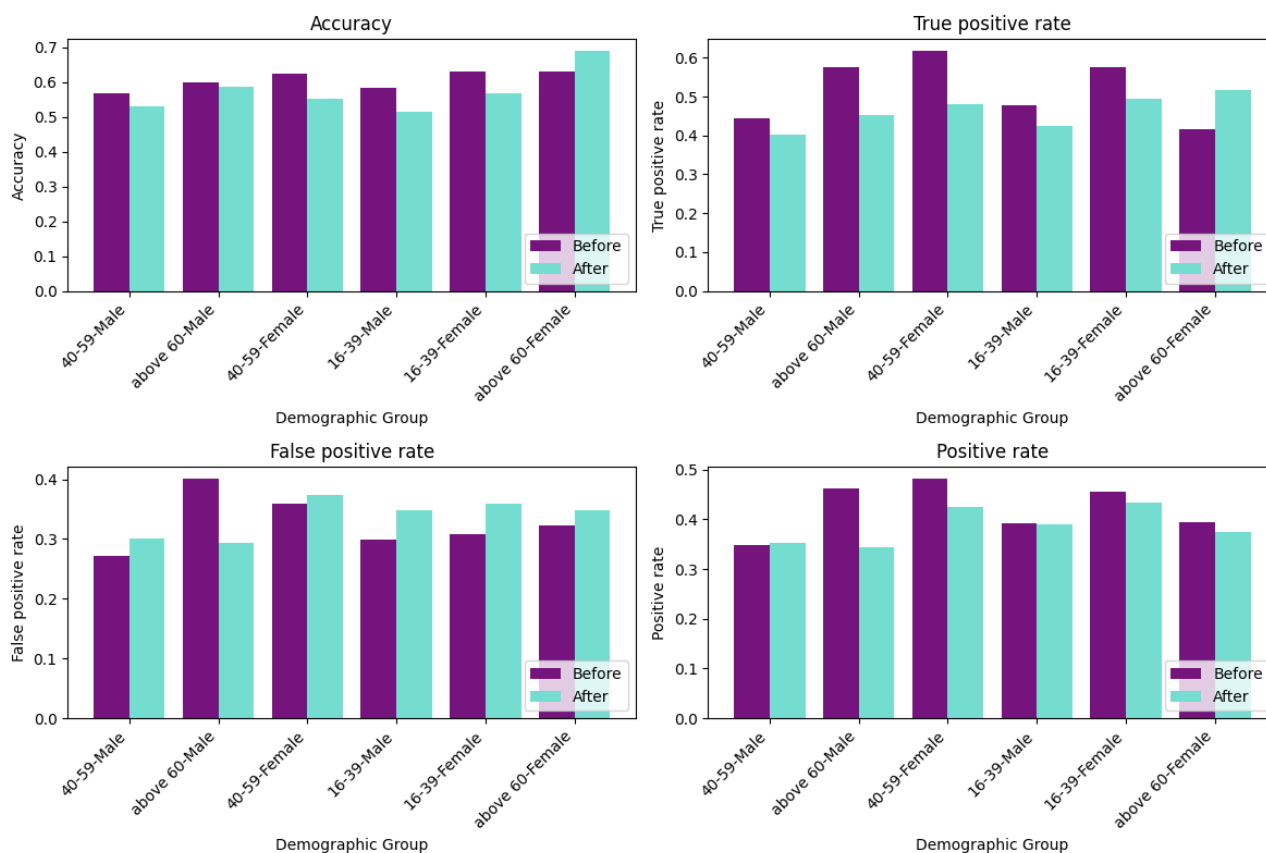
Εικόνα 7.2.27: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα της ηλικίας σε συνδυασμό με το φύλο.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με το φύλο πριν και μετά την μέθοδο εκθετικής κλίσης είναι:

Πίνακας 7.2.22: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με το φύλο.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Εκθετική Κλίση ως προς την Ηλικία σε συνδυασμό με το Φύλο
Διαφορά Ισότητας Ακρίβειας	0,151	0,1435
Αναλογία Ισότητας Ακρίβειας	0,759	0,7765
Διαφορά Ισότητας Πιθανότητας	0,3981	0,3892
Αναλογία Ισότητας Πιθανότητας	0,4953	0,4828
Διαφορά Ίσης Ευκαιρίας	0,3981	0,3892
Αναλογία Ίσης Ευκαιρίας	0,5164	0,5319
Διαφορά Δημογραφικής Ισότητας	0,2593	0,258
Αναλογία Δημογραφικής Ισότητας	0,5981	0,5925

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με το φύλο ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



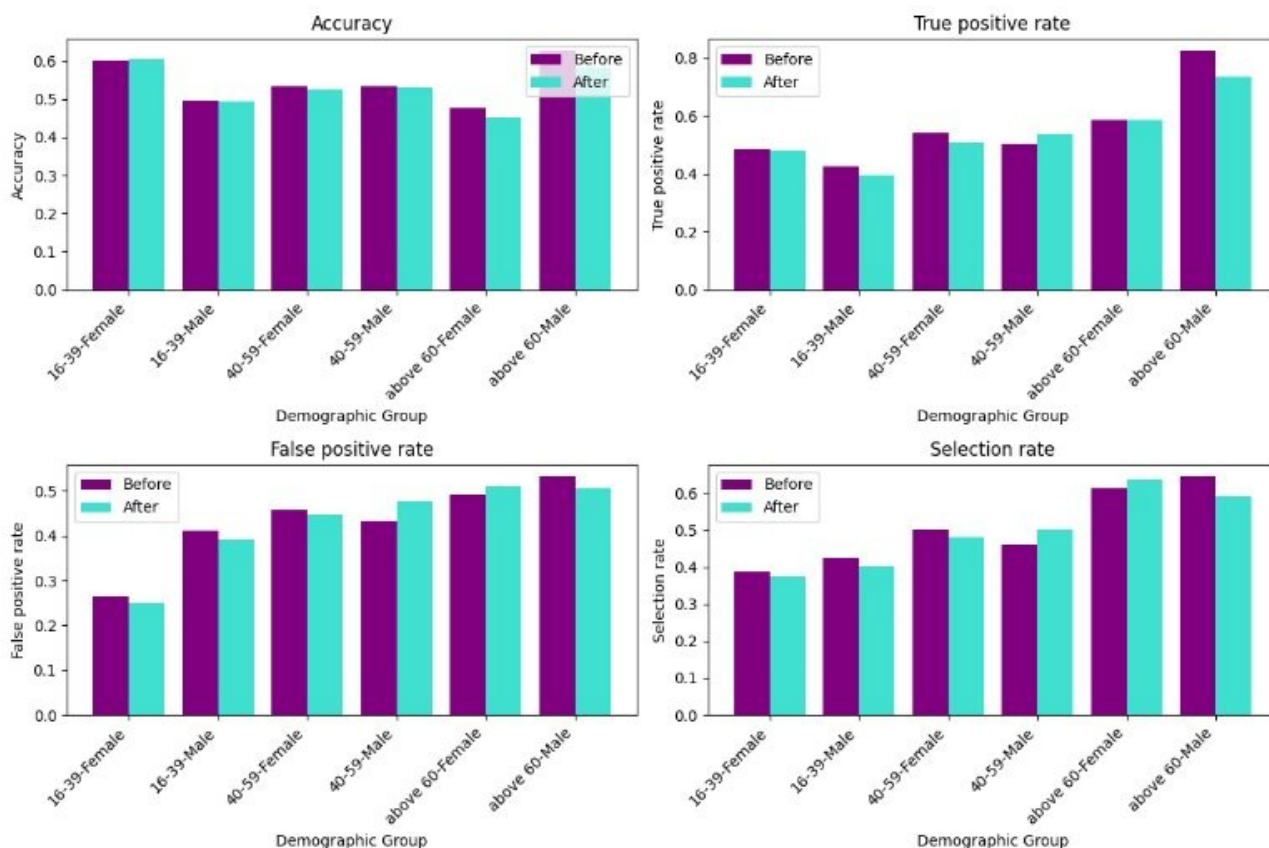
Εικόνα 7.2.28: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα της ηλικίας σε συνδυασμό με το φύλο.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με το φύλο πριν και μετά την μέθοδο ανταγωνιστικής εκπαίδευσης είναι:

Πίνακας 7.2.23: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με το φύλο.

Μετρικές Αμεροληψίας	Συνελκτικό Νευρωνικό Δίκτυο	μετά από Αντίστροφη Εκπαίδευση ως προς την Ηλικία σε συνδυασμό με το Φύλο
Διαφορά Ισότητας Ακρίβειας	0,0650	0,1751
Αναλογία Ισότητας Ακρίβειας	0,8971	0,7456
Διαφορά Ισότητας Πιθανότητας	0,2002	0,1137
Αναλογία Ισότητας Πιθανότητας	0,6755	0,78
Διαφορά Ίσης Ευκαιρίας	0,2002	0,1137
Αναλογία Ίσης Ευκαιρίας	0,6755	0,78
Διαφορά Δημογραφικής Ισότητας	0,1344	0,0903
Αναλογία Δημογραφικής Ισότητας	0,7212	0,7916

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με το φύλο ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου του βελτιστοποιητή κατωφλίου για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



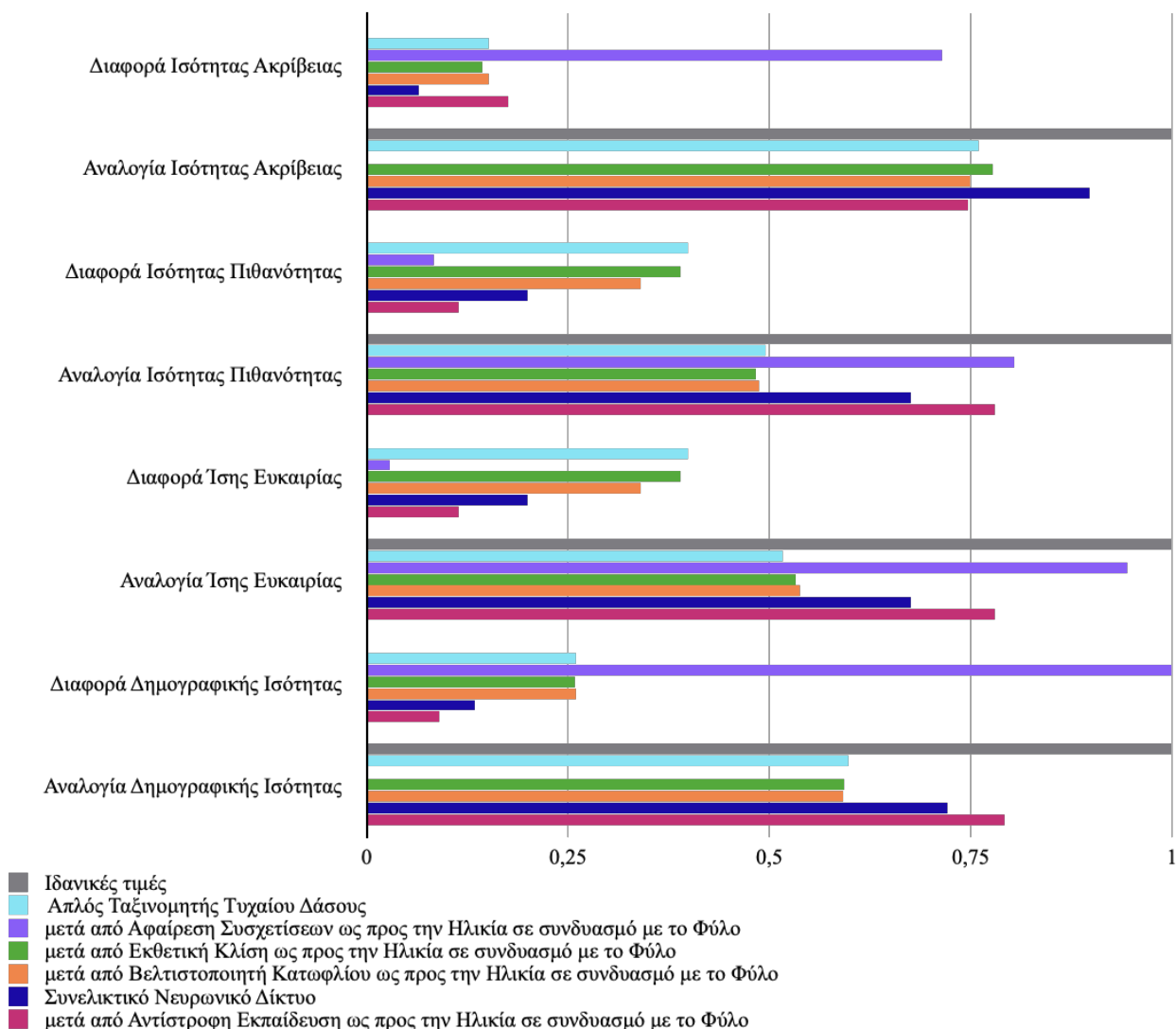
Εικόνα 7.2.29: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα της ηλικίας σε συνδυασμό με το φύλο.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με το φύλο πριν και μετά την μέθοδο βελτιστοποιητή κατωφλίου είναι:

Πίνακας 7.2.24: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με το φύλο.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Βελτιστοποιητή Κατωφλίου ως προς την Ηλικία σε συνδυασμό με το Φύλο
Διαφορά Ισότητας Ακρίβειας	0,151	0,1519
Αναλογία Ισότητας Ακρίβειας	0,759	0,7484
Διαφορά Ισότητας Πιθανότητας	0,3981	0,3398
Αναλογία Ισότητας Πιθανότητας	0,4953	0,4867
Διαφορά Ίσης Ευκαιρίας	0,3981	0,3398
Αναλογία Ίσης Ευκαιρίας	0,5164	0,5383
Διαφορά Δημογραφικής Ισότητας	0,2593	0,2593
Αναλογία Δημογραφικής Ισότητας	0,5981	0,5917

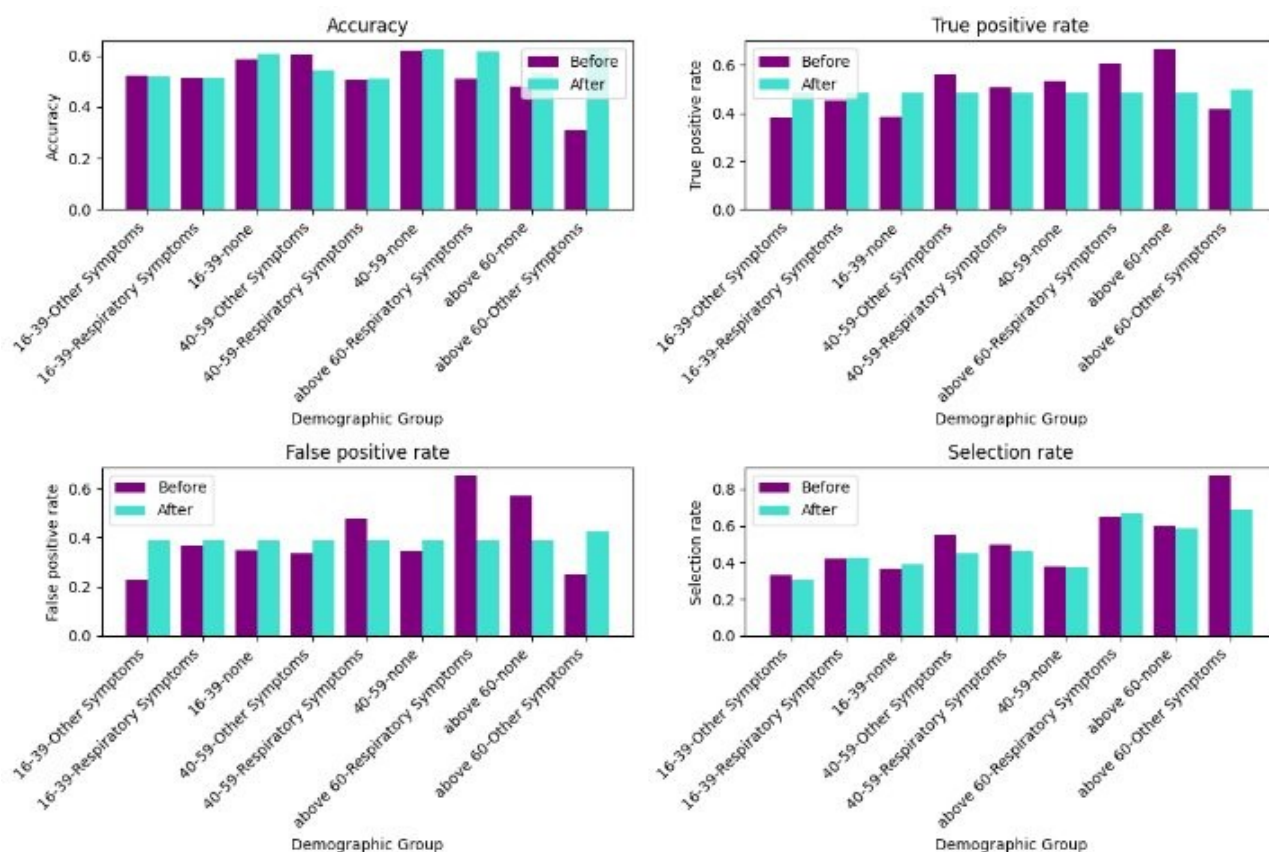
Συνολικά, οι τιμές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων ως προς την ηλικία σε συνδυασμό με το φύλο, συγκριτικά και με τις ιδανικές τιμές τους, παρουσιάζονται ακολούθως:



Εικόνα 7.2.30: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα.

7.2.7 Αμεροληψία ως προς την Ηλικία σε συνδυασμό με τα Συμπτώματα

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με τα συμπτώματα ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



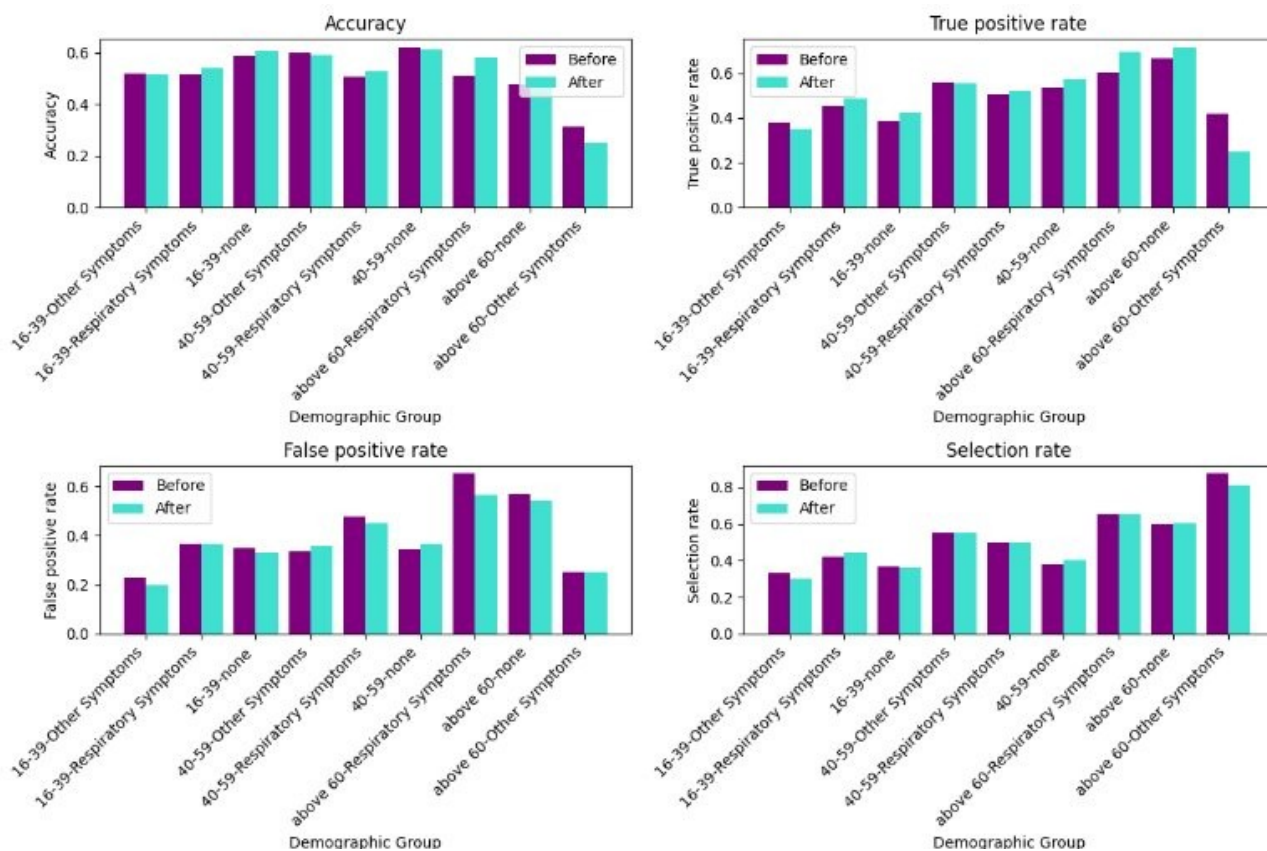
Εικόνα 7.2.31: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με τα συμπτώματα.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα πριν και μετά την αφαίρεση συσχετίσεων είναι:

Πίνακας 7.2.25: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με τα συμπτώματα.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων ως προς την Ηλικία σε συνδυασμό με τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,3088	0,3036
Αναλογία Ισότητας Ακρίβειας	0,503	0,5142
Διαφορά Ισότητας Πιθανότητας	0,4237	0,0544
Αναλογία Ισότητας Πιθανότητας	0,3504	0,8712
Διαφορά Ίσης Ευκαιρίας	0,285	0,0242
Αναλογία Ίσης Ευκαιρίας	0,5725	0,9516
Διαφορά Δημογραφικής Ισότητας	0,5443	0,6875
Αναλογία Δημογραφικής Ισότητας	0,3779	0

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με τα συμπτώματα ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



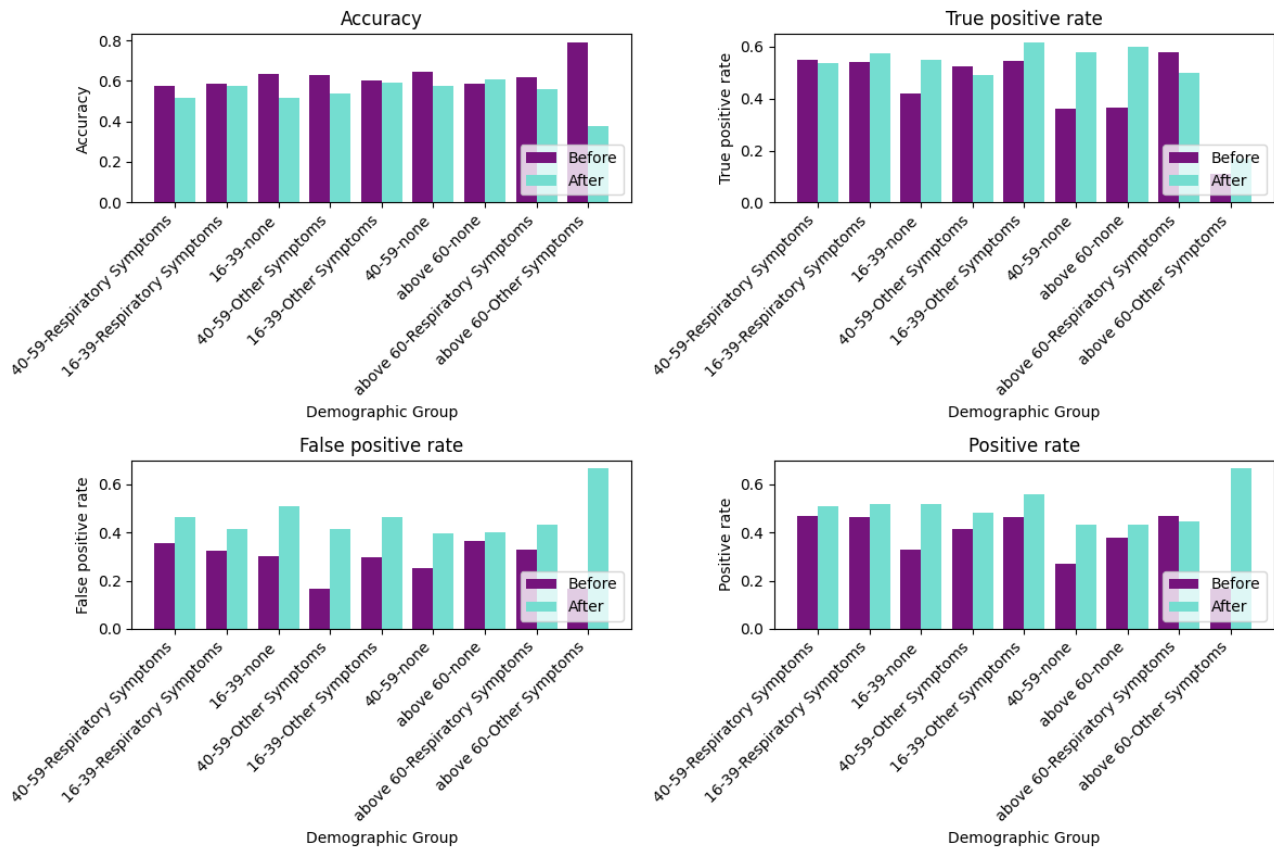
Εικόνα 7.2.32: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα της ηλικίας σε συνδυασμό με τα συμπτώματα.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με το φύλο πριν και μετά την μέθοδο εκθετικής κλίσης είναι:

Πίνακας 7.2.26: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με τα συμπτώματα.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Εκθετική Κλίση ως προς την Ηλικία σε συνδυασμό με τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,3088	0,3628
Αναλογία Ισότητας Ακρίβειας	0,503	0,408
Διαφορά Ισότητας Πιθανότητας	0,4237	0,4667
Αναλογία Ισότητας Πιθανότητας	0,3504	0,3488
Διαφορά Ίσης Ευκαιρίας	0,285	0,4667
Αναλογία Ίσης Ευκαιρίας	0,5725	0,3488
Διαφορά Δημογραφικής Ισότητας	0,5443	0,5139
Αναλογία Δημογραφικής Ισότητας	0,3779	0,3675

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με τα συμπτώματα ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



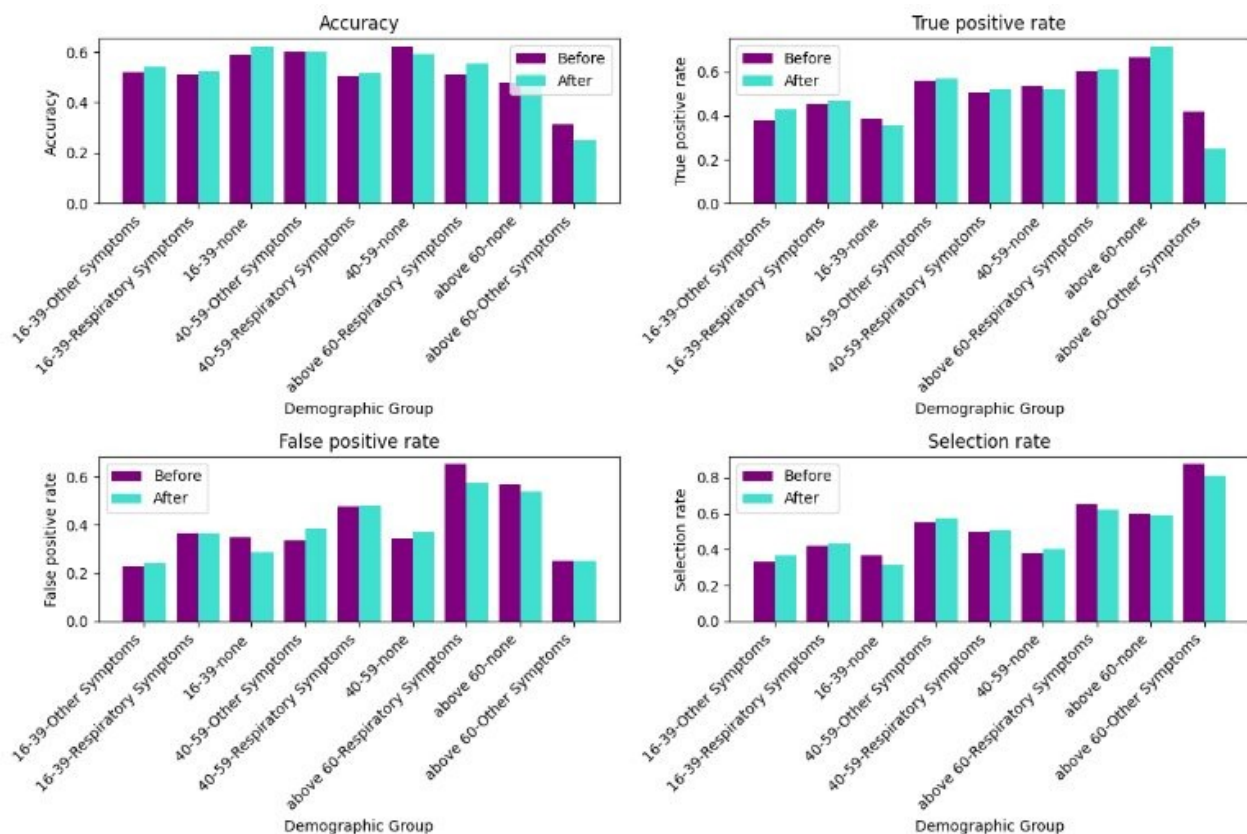
Εικόνα 7.2.33: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα της ηλικίας σε συνδυασμό με τα συμπτώματα.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα πριν και μετά την μέθοδο ανταγωνιστικής εκπαίδευσης είναι:

Πίνακας 7.2.27: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με τα συμπτώματα.

Μετρικές Αμεροληψίας	Συνελικτικό Νευρωνικό Δίκτυο	μετά από Αντίστροφη Εκπαίδευση ως προς την Ηλικία σε συνδυασμό με τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,2142	0,2339
Αναλογία Ισότητας Ακρίβειας	0,7295	0,6159
Διαφορά Ισότητας Πιθανότητας	0,4662	0,4512
Αναλογία Ισότητας Πιθανότητας	0,1924	0,2697
Διαφορά Ίσης Ευκαιρίας	0,4662	0,4512
Αναλογία Ίσης Ευκαιρίας	0,1924	0,2697
Διαφορά Δημογραφικής Ισότητας	0,3042	0,2331
Αναλογία Δημογραφικής Ισότητας	0,3540	0,6504

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με τα συμπτώματα ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου του βελτιστοποιητή κατωφλίου για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



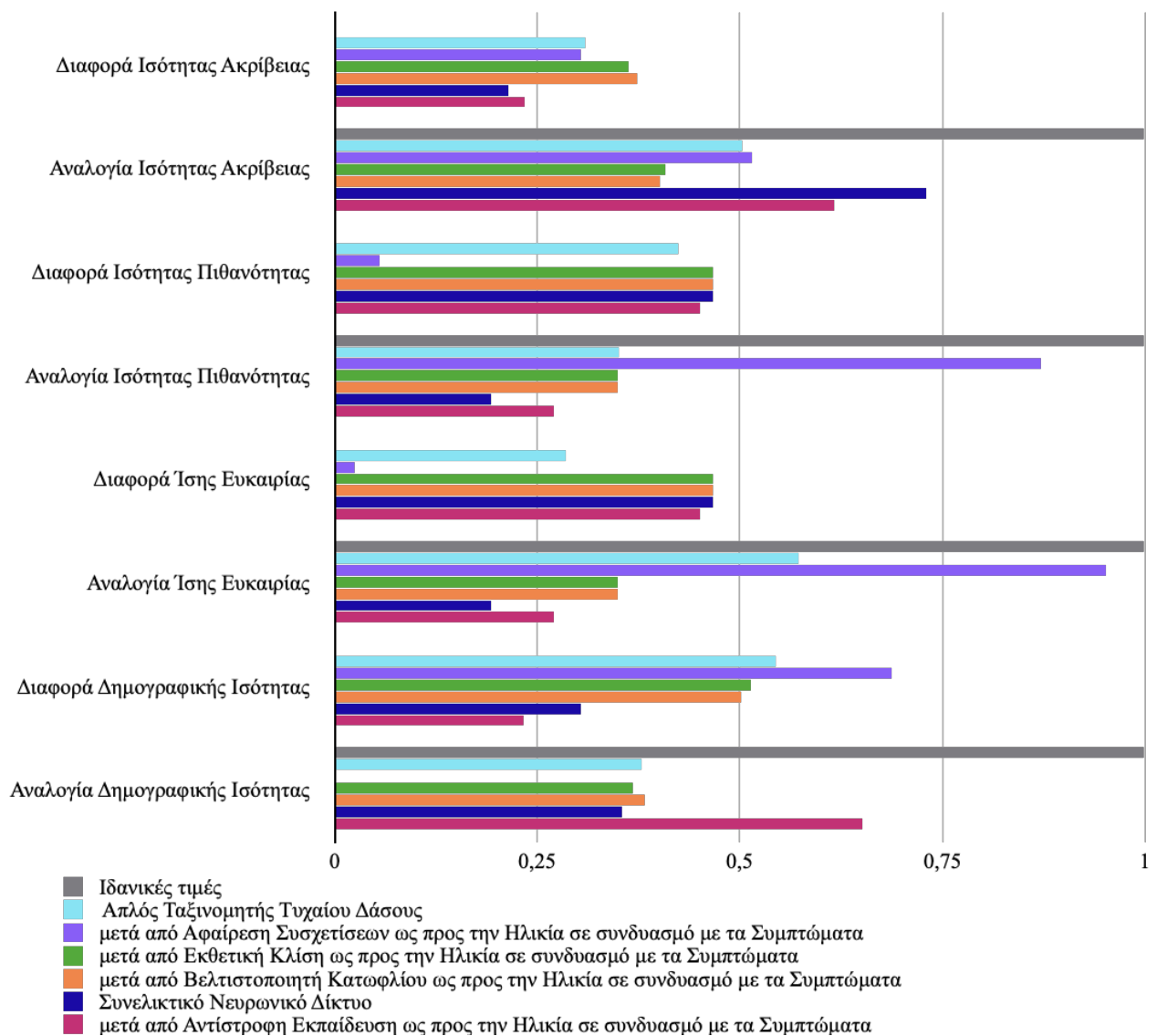
Εικόνα 7.2.34: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα της ηλικίας σε συνδυασμό με τα συμπτώματα.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα πριν και μετά την μέθοδο βελτιστοποιητή κατωφλίου είναι:

Πίνακας 7.2.28: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με τα συμπτώματα.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαιού Δάσους	μετά από Βελτιστοποιητή Κατωφλίου ως προς την Ηλικία σε συνδυασμό με τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,3088	0,3727
Αναλογία Ισότητας Ακρίβειας	0,503	0,4015
Διαφορά Ισότητας Πιθανότητας	0,4237	0,4667
Αναλογία Ισότητας Πιθανότητας	0,3504	0,3488
Διαφορά Ίσης Ευκαιρίας	0,285	0,4667
Αναλογία Ίσης Ευκαιρίας	0,5725	0,3488
Διαφορά Δημογραφικής Ισότητας	0,5443	0,5019
Αναλογία Δημογραφικής Ισότητας	0,3779	0,3823

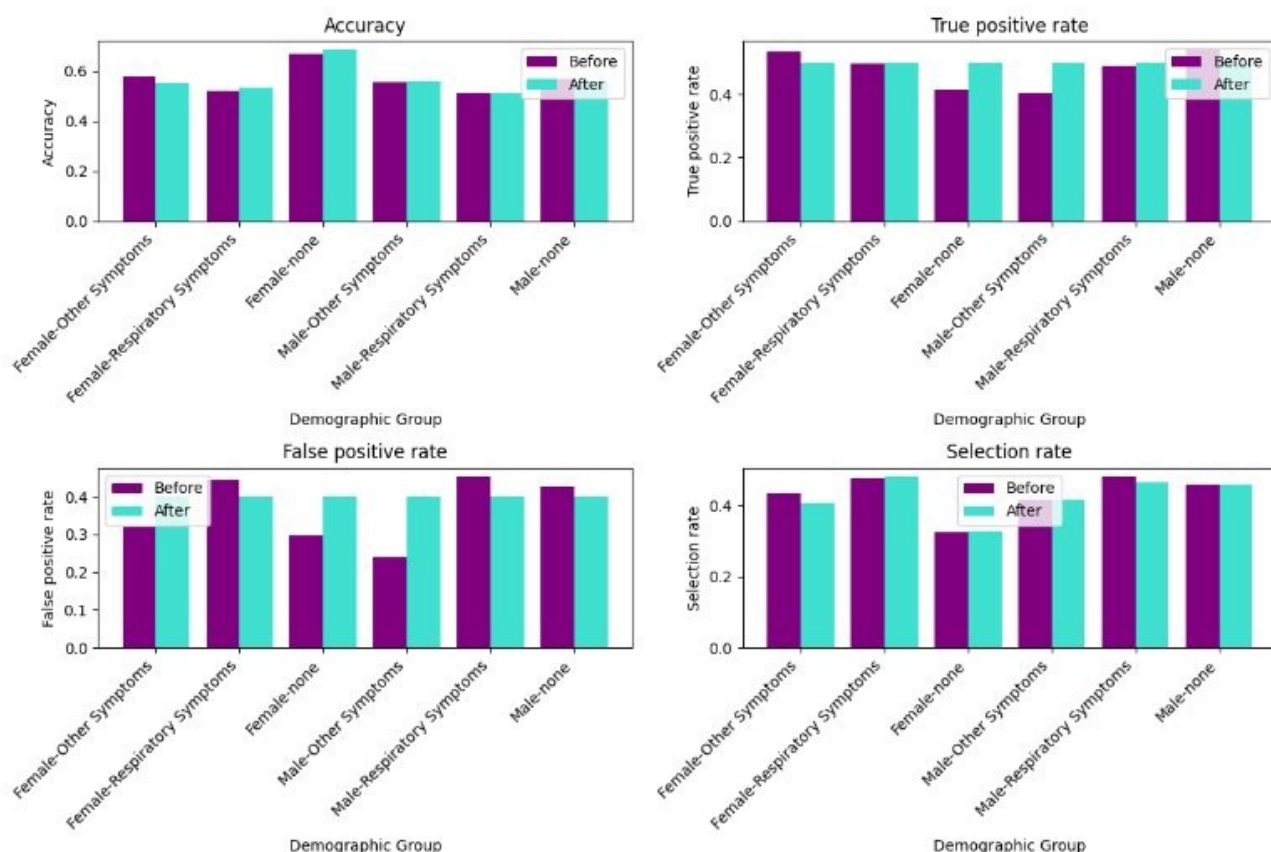
Συνολικά, οι τιμές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων ως προς την ηλικία σε συνδυασμό με τα συμπτώματα, συγκριτικά και με τις ιδανικές τιμές τους, παρουσιάζονται ακολούθως:



Εικόνα 7.2.35: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα.

7.2.8 Αμεροληψία ως προς το Φύλο σε συνδυασμό με τα Συμπτώματα

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου σε συνδυασμό με τα συμπτώματα ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



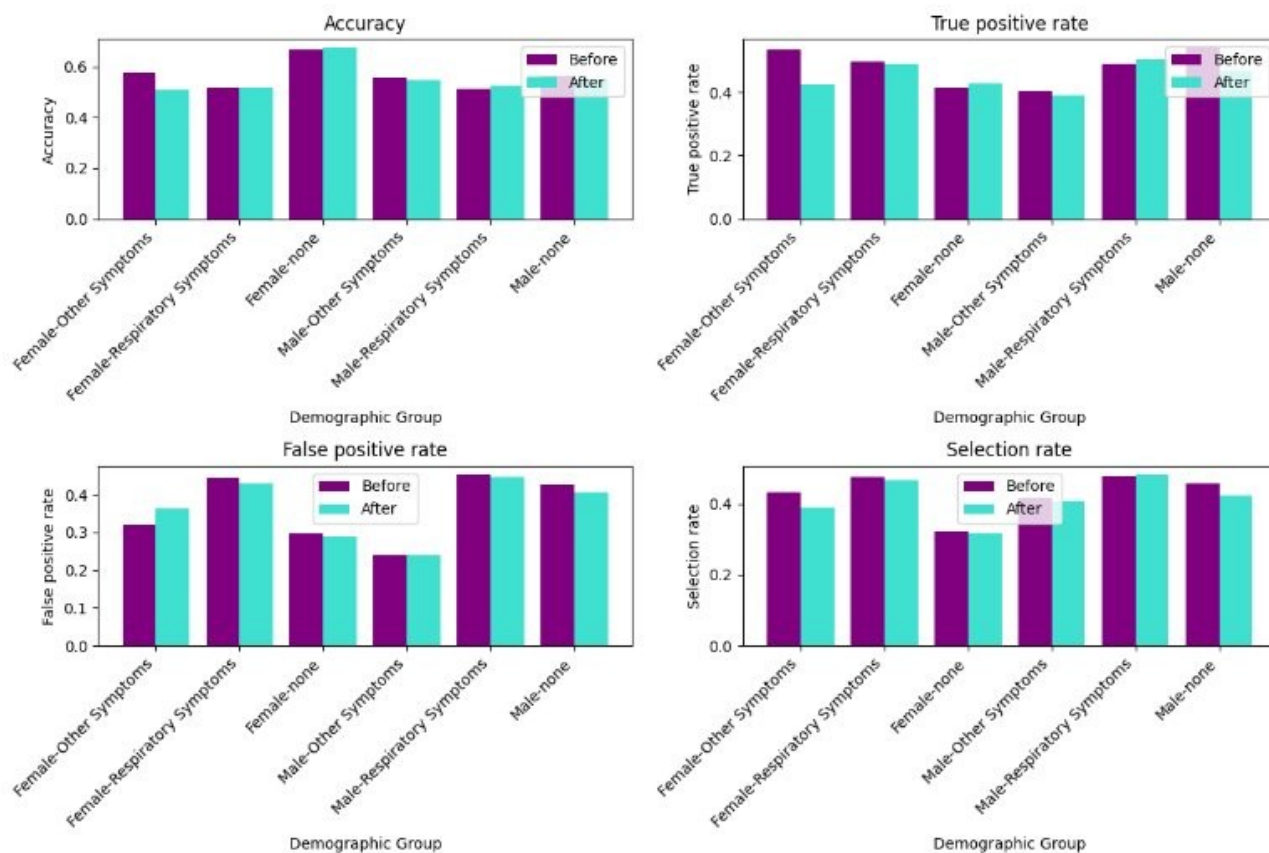
Εικόνα 7.2.36: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα του φύλου σε συνδυασμό με τα συμπτώματα.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με τα συμπτώματα πριν και μετά την αφαίρεση συσχετίσεων είναι:

Πίνακας 7.2.29: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων του φύλου σε συνδυασμό με τα συμπτώματα.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων ως προς το Φύλο σε συνδυασμό με τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,1536	1
Αναλογία Ισότητας Ακρίβειας	0,7693	0
Διαφορά Ισότητας Πιθανότητας	0,2146	0,1395
Αναλογία Ισότητας Πιθανότητας	0,5262	0,7117
Διαφορά Ίσης Ευκαιρίας	0,1381	0,0463
Αναλογία Ίσης Ευκαιρίας	0,7447	0,91
Διαφορά Δημογραφικής Ισότητας	0,1557	1
Αναλογία Δημογραφικής Ισότητας	0,6754	0

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου σε συνδυασμό με τα συμπτώματα ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



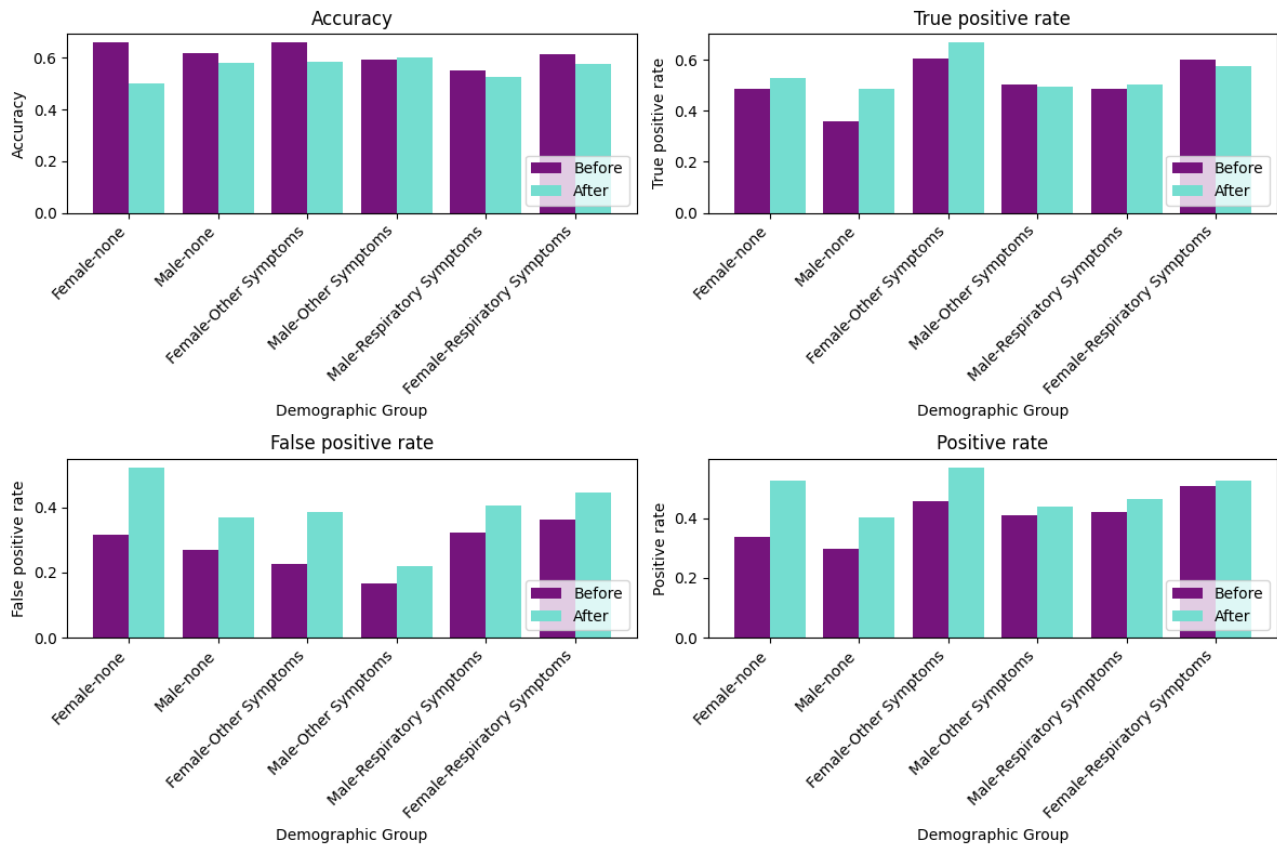
Εικόνα 7.2.37: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα του φύλου σε συνδυασμό με τα συμπτώματα.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου σε συνδυασμό με το φύλο πριν και μετά την μέθοδο εκθετικής κλίσης είναι:

Πίνακας 7.2.30: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων του φύλου σε συνδυασμό με τα συμπτώματα.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Εκθετική Κλίση ως προς το Φύλο σε συνδυασμό με τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,1536	0,1678
Αναλογία Ισότητας Ακρίβειας	0,7693	0,7517
Διαφορά Ισότητας Πιθανότητας	0,2146	0,2077
Αναλογία Ισότητας Πιθανότητας	0,5262	0,5344
Διαφορά Ίσης Ευκαιρίας	0,1381	0,1137
Αναλογία Ίσης Ευκαιρίας	0,7447	0,7749
Διαφορά Δημογραφικής Ισότητας	0,1557	0,1671
Αναλογία Δημογραφικής Ισότητας	0,6754	0,6546

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου σε συνδυασμό με τα συμπτώματα ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



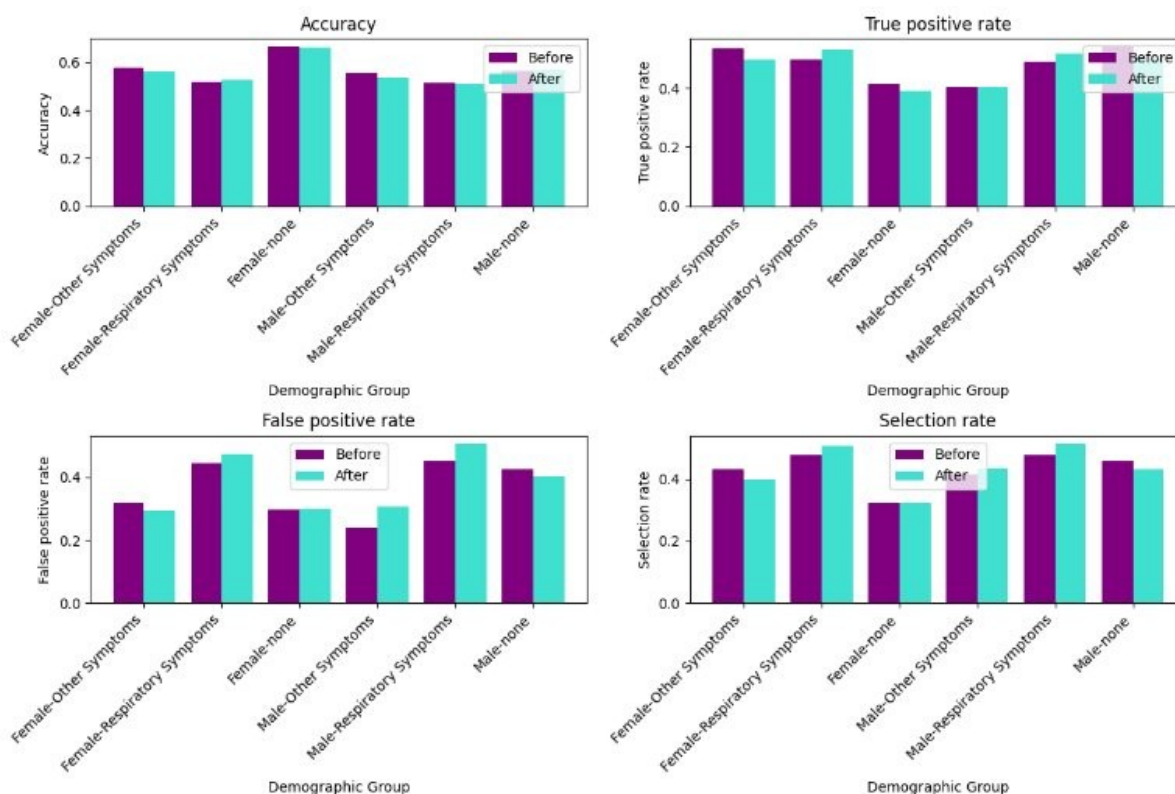
Εικόνα 7.2.38: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα του φύλου σε συνδυασμό με τα συμπτώματα.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου σε συνδυασμό με τα συμπτώματα πριν και μετά την μέθοδο ανταγωνιστικής εκπαίδευσης είναι:

Πίνακας 7.2.31: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων του φύλου σε συνδυασμό με τα συμπτώματα.

Μετρικές Αμεροληψίας	Συνελκτικό Νευρωνικό Δίκτυο	μετά από Αντίστροφη Εκπαίδευση ως προς το Φύλο σε συνδυασμό με τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,1076	0,0633
Αναλογία Ισότητας Ακρίβειας	0,8368	0,8944
Διαφορά Ισότητας Πιθανότητας	0,2448	0,1639
Αναλογία Ισότητας Πιθανότητας	0,4583	0,7020
Διαφορά Ίσης Ευκαιρίας	0,2448	0,1639
Αναλογία Ίσης Ευκαιρίας	0,5959	0,7423
Διαφορά Δημογραφικής Ισότητας	0,2122	0,1308
Αναλογία Δημογραφικής Ισότητας	0,5837	0,7635

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου σε συνδυασμό με τα συμπτώματα ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου του βελτιστοποιητή κατωφλίου για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



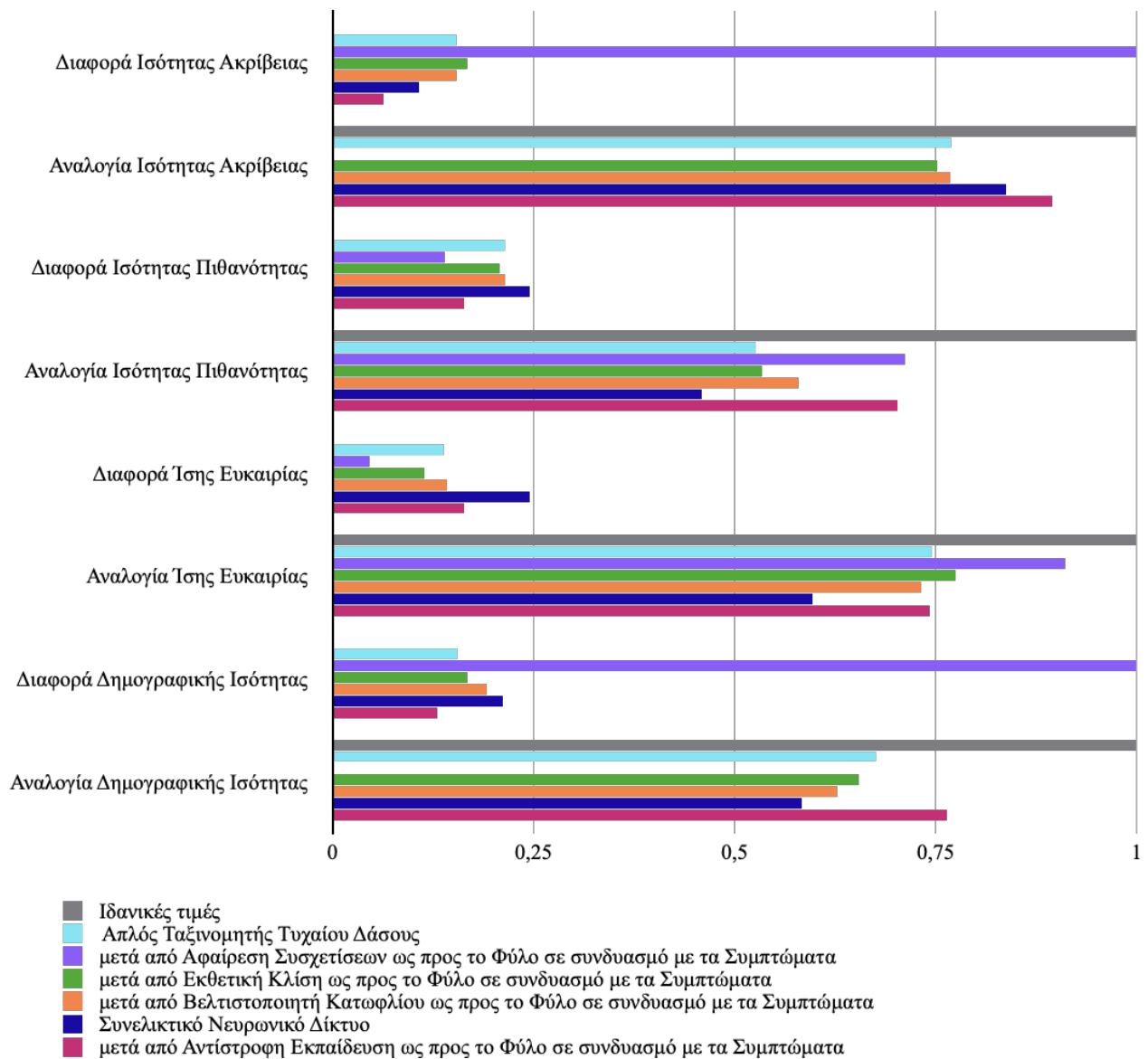
Εικόνα 7.2.39: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα του φύλου σε συνδυασμό με τα συμπτώματα.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου σε συνδυασμό με τα συμπτώματα πριν και μετά την μέθοδο βελτιστοποιητή κατωφλίου είναι:

Πίνακας 7.2.32: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων του φύλου σε συνδυασμό με τα συμπτώματα.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Βελτιστοποιητή Κατωφλίου ως προς το Φύλο σε συνδυασμό με τα Συμπτώματα
Διαφορά Ισότητας Ακρίβειας	0,1536	0,1537
Αναλογία Ισότητας Ακρίβειας	0,7693	0,7678
Διαφορά Ισότητας Πιθανότητας	0,2146	0,2139
Αναλογία Ισότητας Πιθανότητας	0,5262	0,5787
Διαφορά Ίσης Ευκαιρίας	0,1381	0,1425
Αναλογία Ίσης Ευκαιρίας	0,7447	0,7317
Διαφορά Δημογραφικής Ισότητας	0,1557	0,1922
Αναλογία Δημογραφικής Ισότητας	0,6754	0,627

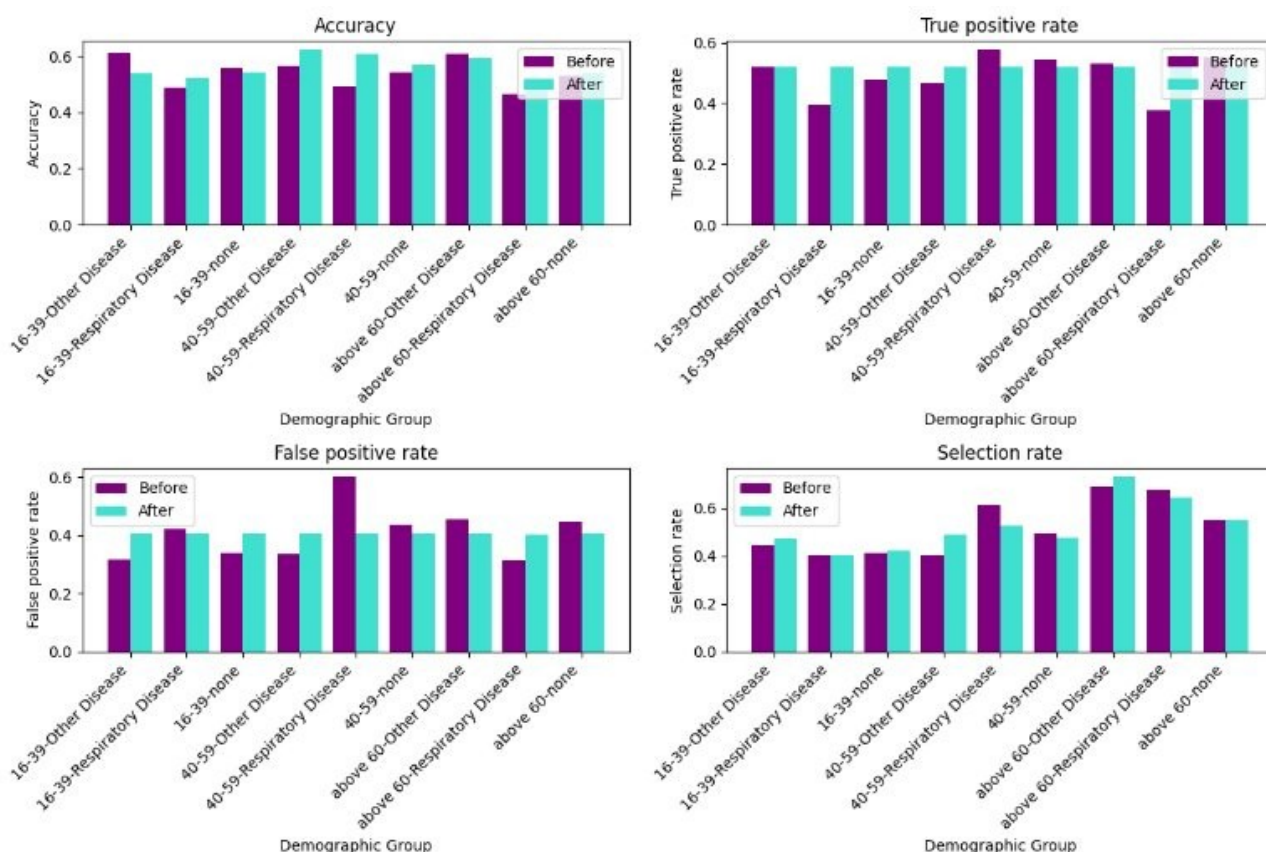
Συνολικά, οι τιμές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων ως προς το φύλο σε συνδυασμό με τα συμπτώματα, συγκριτικά και με τις ιδανικές τιμές τους, παρουσιάζονται ακολούθως:



Εικόνα 7.2.40: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία του φύλου σε συνδυασμό με τα συμπτώματα.

7.2.9 Αμεροληψία ως προς την Ηλικία σε συνδυασμό με το Ιατρικό Ιστορικό

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



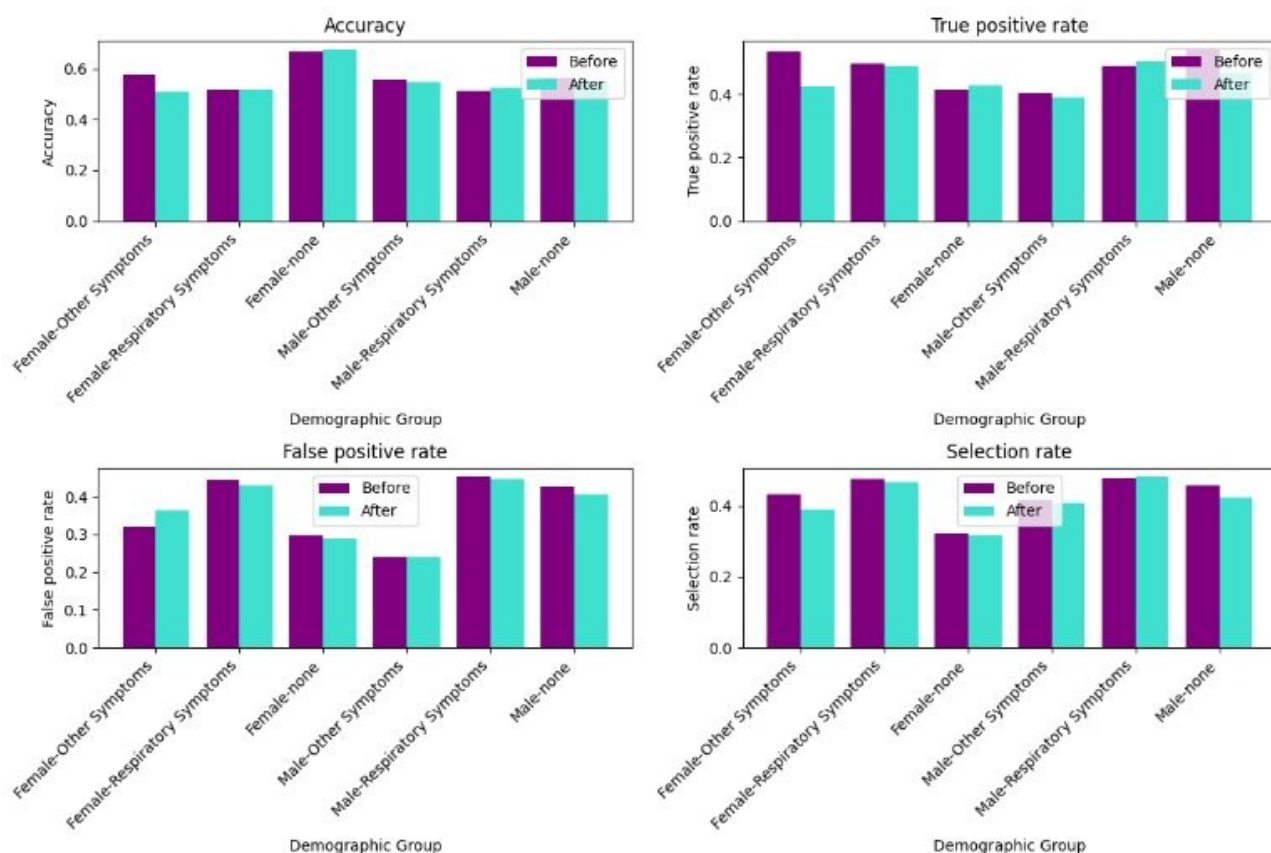
Εικόνα 7.2.41: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό πριν και μετά την αφαίρεση συσχετίσεων είναι:

Πίνακας 7.2.33: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων ως προς την Ηλικία σε συνδυασμό με το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,1478	0,6225
Αναλογία Ισότητας Ακρίβειας	0,7588	0
Διαφορά Ισότητας Πιθανότητας	0,2898	0,1127
Αναλογία Ισότητας Πιθανότητας	0,5189	0,7804
Διαφορά Ίσης Ευκαιρίας	0,2007	0,0468
Αναλογία Ίσης Ευκαιρίας	0,6514	0,9117
Διαφορά Δημογραφικής Ισότητας	0,2854	0,5951
Αναλογία Δημογραφικής Ισότητας	0,5858	0,4049

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



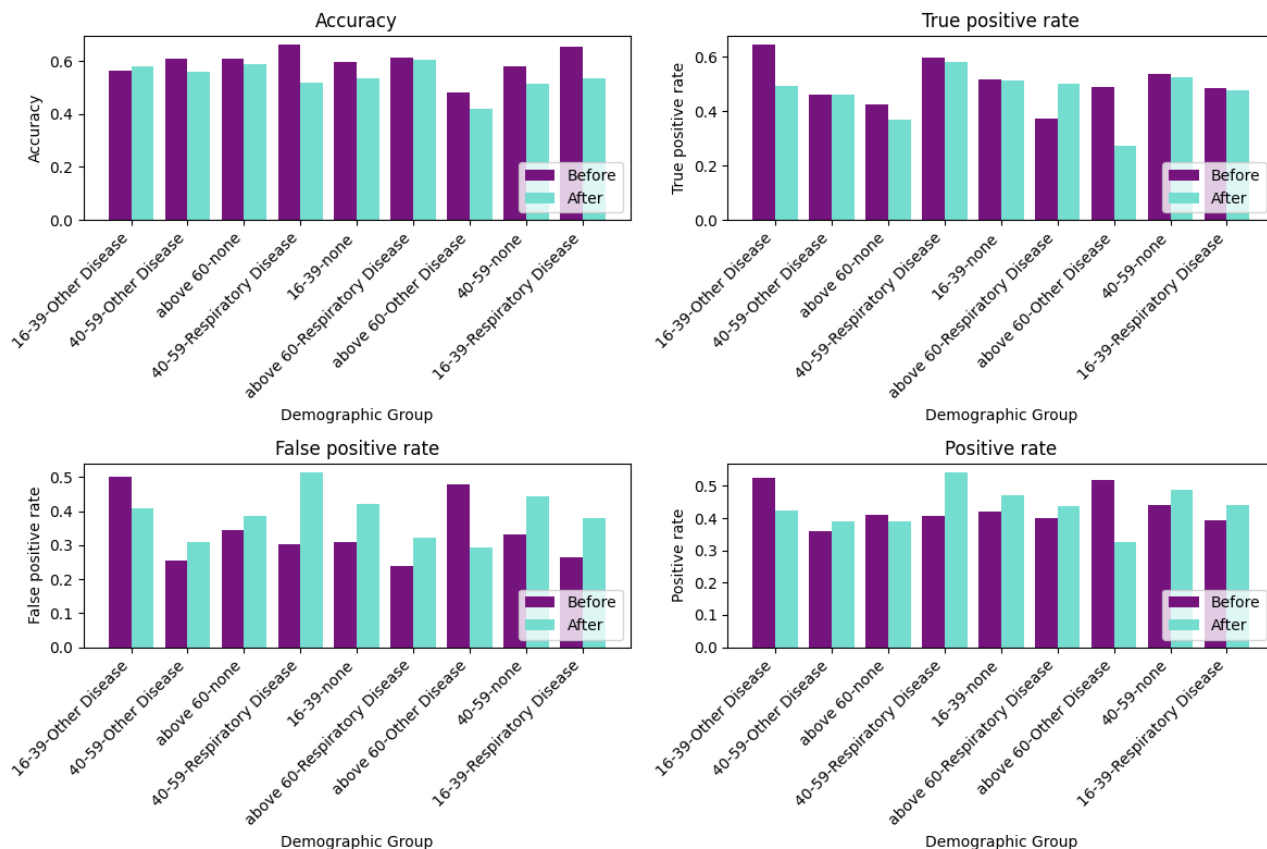
Εικόνα 7.2.42: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό πριν και μετά την μέθοδο εκθετικής κλίσης είναι:

Πίνακας 7.2.34: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Εκθετική Κλίση ως προς την Ηλικία σε συνδυασμό με το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,1478	0,1272
Αναλογία Ισότητας Ακρίβειας	0,7588	0,8014
Διαφορά Ισότητας Πιθανότητας	0,2898	0,3207
Αναλογία Ισότητας Πιθανότητας	0,5189	0,539
Διαφορά Ίσης Ευκαιρίας	0,2007	0,3207
Αναλογία Ίσης Ευκαιρίας	0,6514	0,539
Διαφορά Δημογραφικής Ισότητας	0,2854	0,2011
Αναλογία Δημογραφικής Ισότητας	0,5858	0,6646

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



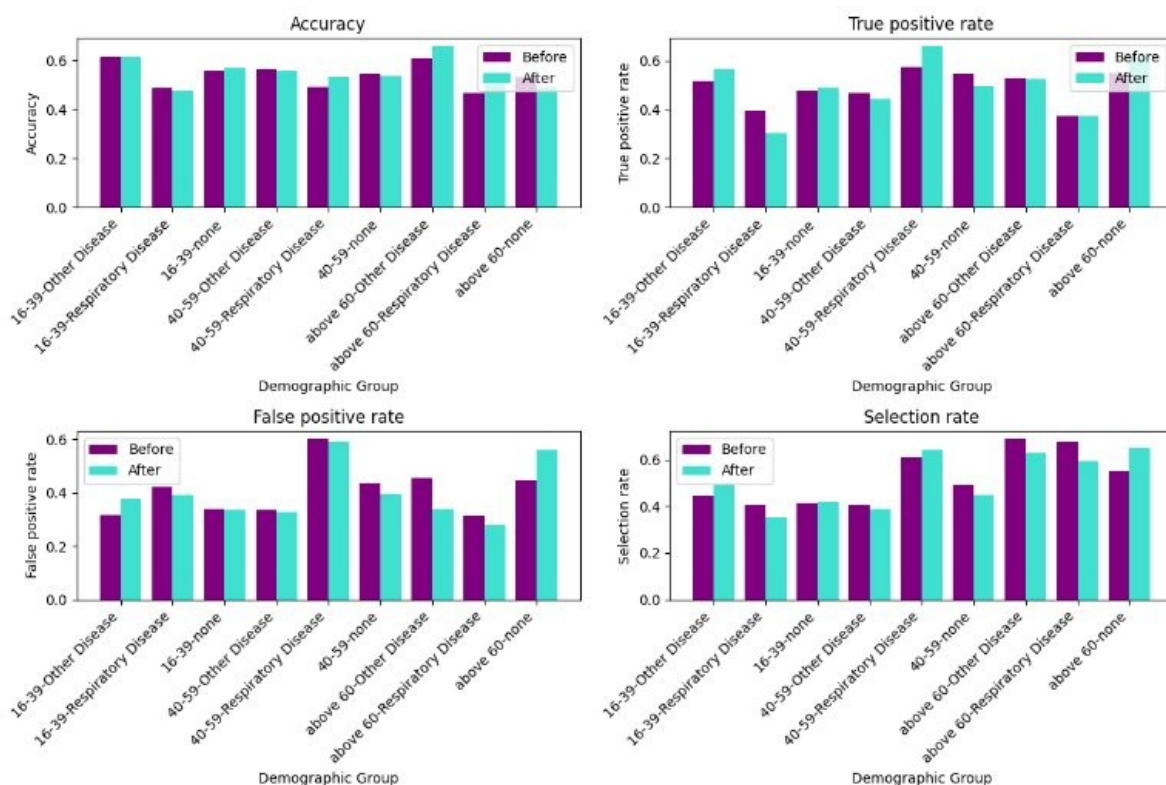
Εικόνα 7.2.43: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό πριν και μετά την μέθοδο ανταγωνιστικής εκπαίδευσης είναι:

Πίνακας 7.2.35: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

Μετρικές Αμεροληψίας	Συνελκτικό Νευρωνικό Δίκτυο	μετά από Αντίστροφη Εκπαίδευση ως προς την Ηλικία σε συνδυασμό με το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,1783	0,1865
Αναλογία Ισότητας Ακρίβειας	0,7301	0,6920
Διαφορά Ισότητας Πιθανότητας	0,2723	0,3083
Αναλογία Ισότητας Πιθανότητας	0,4796	0,4703
Διαφορά Ίσης Ευκαιρίας	0,2723	0,3083
Αναλογία Ίσης Ευκαιρίας	0,5763	0,4703
Διαφορά Δημογραφικής Ισότητας	0,1676	0,2167
Αναλογία Δημογραφικής Ισότητας	0,6811	0,6003

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου του βελτιστοποιητή κατωφλίου για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



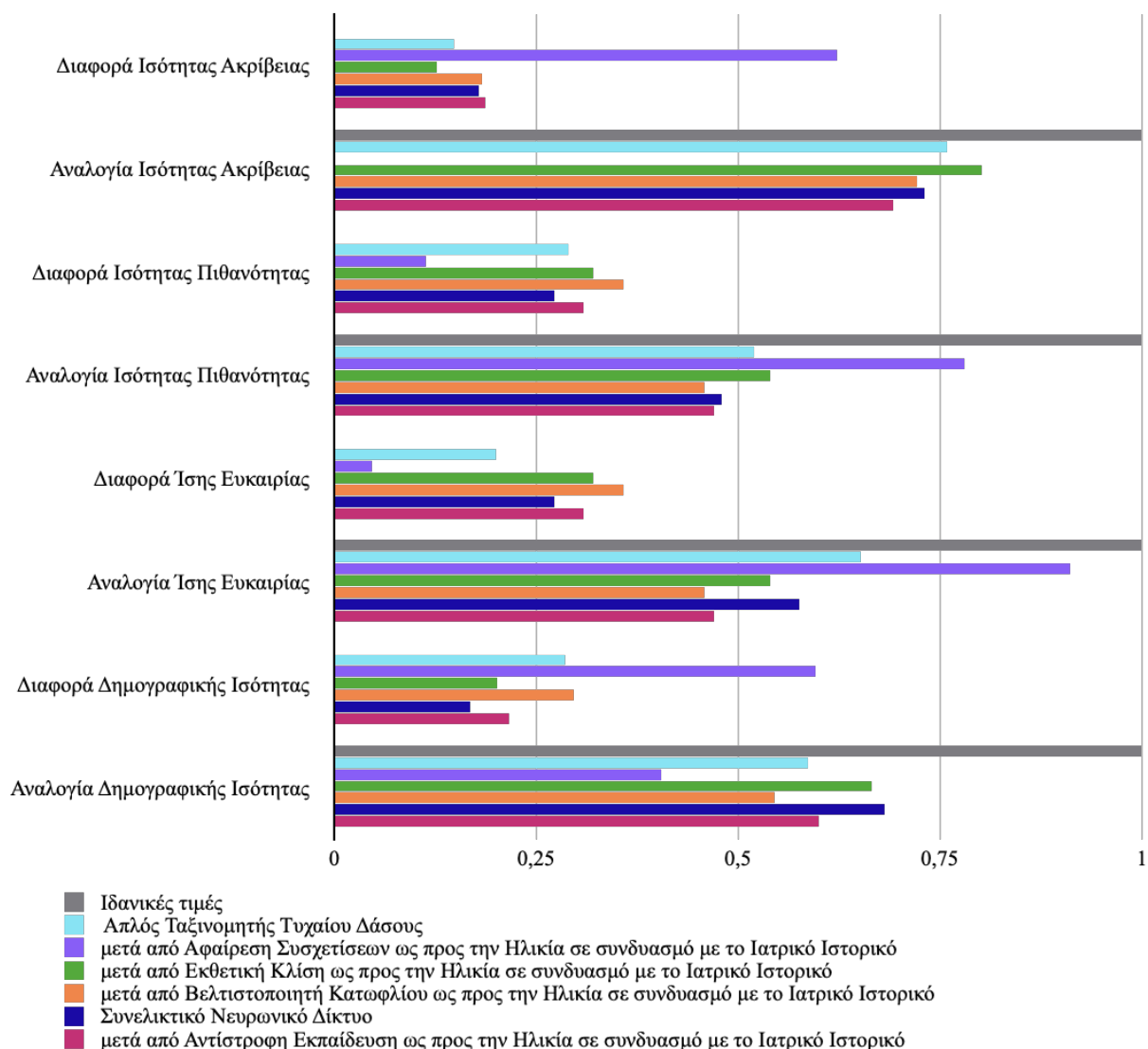
Εικόνα 7.2.44: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό πριν και μετά την μέθοδο βελτιστοποιητή κατωφλίου είναι:

Πίνακας 7.2.36: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Βελτιστοποιητή Κατωφλίου ως προς την Ηλικία σε συνδυασμό με το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,1478	0,1831
Αναλογία Ισότητας Ακρίβειας	0,7588	0,721
Διαφορά Ισότητας Πιθανότητας	0,2898	0,3573
Αναλογία Ισότητας Πιθανότητας	0,5189	0,4586
Διαφορά Ίσης Ευκαιρίας	0,2007	0,3573
Αναλογία Ίσης Ευκαιρίας	0,6514	0,4586
Διαφορά Δημογραφικής Ισότητας	0,2854	0,296
Αναλογία Δημογραφικής Ισότητας	0,5858	0,5449

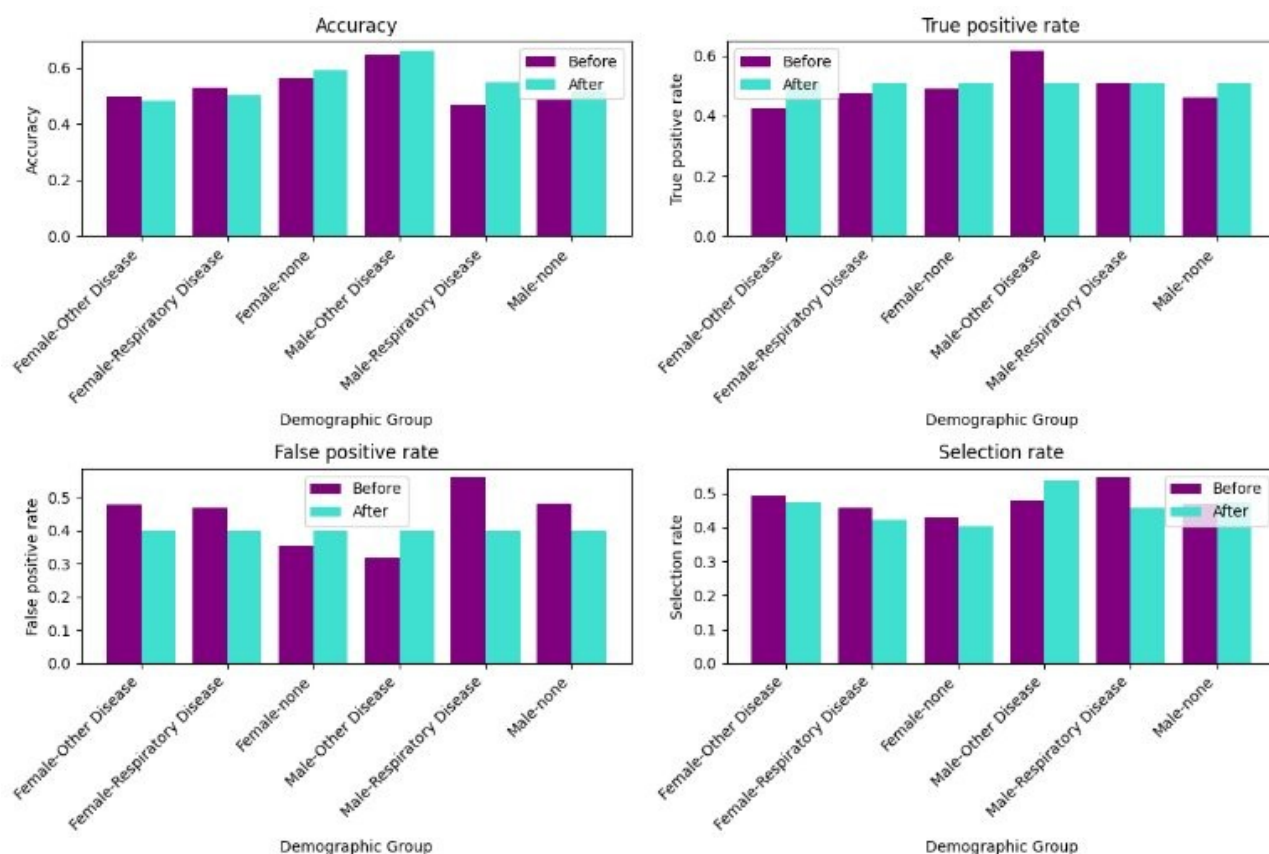
Οι τιμές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή όλων των μεθόδων εξάλειψης των διακρίσεων ως προς την ηλικία σε συνδυασμό με το ιατρικό ιστορικό, συγκριτικά και με τις ιδανικές τιμές τους, παρουσιάζονται ακολούθως:



Εικόνα 7.2.45: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό.

7.2.10 Αμεροληψία ως προς το Φύλο σε συνδυασμό με το Ιατρικό Ιστορικό

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



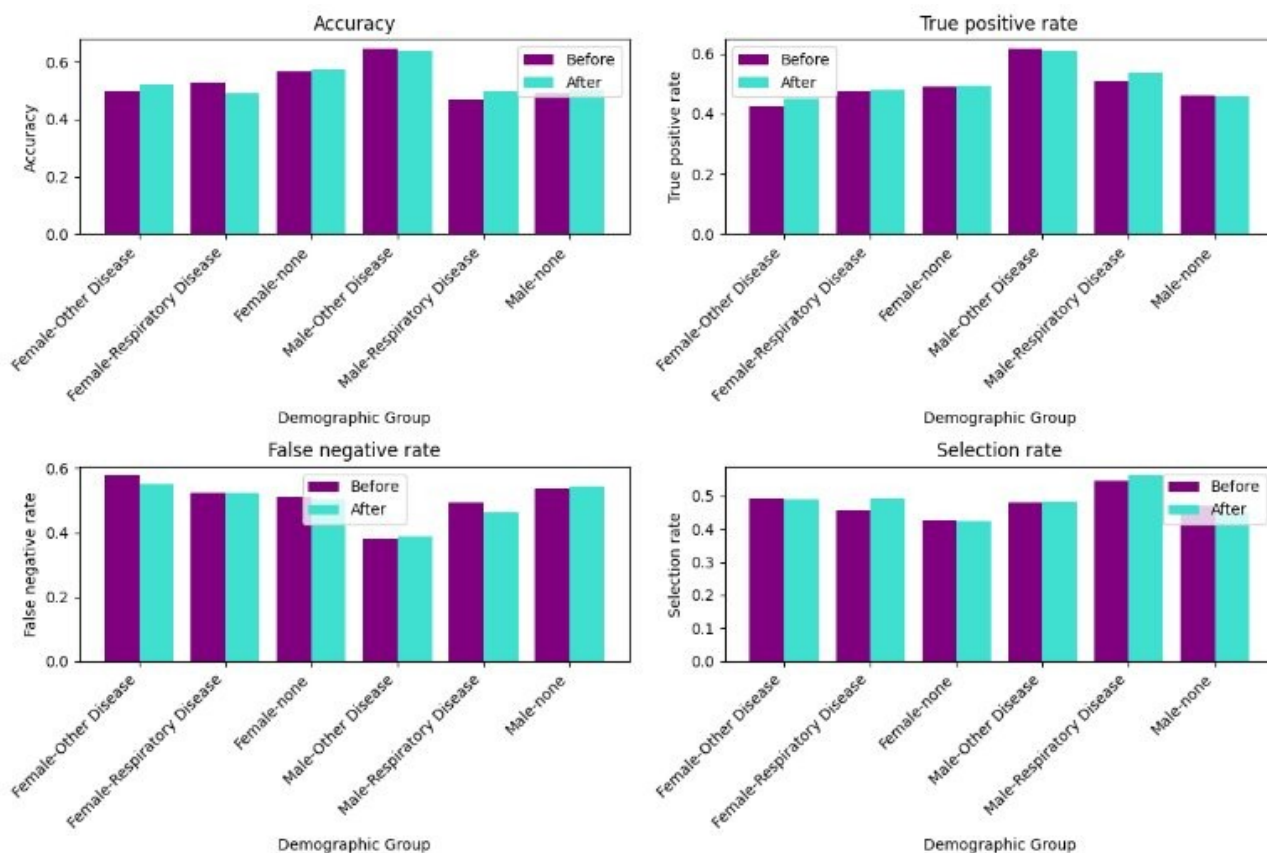
Εικόνα 7.2.46: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την αφαίρεση συσχετίσεων για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό πριν και μετά την αφαίρεση συσχετίσεων είναι:

Πίνακας 7.2.37: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την αφαίρεση συσχετίσεων για την εξάλειψη των διακρίσεων του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Αφαίρεση Συσχετίσεων ως προς το Φύλο σε συνδυασμό με το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,1792	1
Αναλογία Ισότητας Ακρίβειας	0,7231	0
Διαφορά Ισότητας Πιθανότητας	0,2397	0,1031
Αναλογία Ισότητας Πιθανότητας	0,5702	0,7826
Διαφορά Ίσης Ευκαιρίας	0,1936	0,0462
Αναλογία Ίσης Ευκαιρίας	0,6862	0,9133
Διαφορά Δημογραφικής Ισότητας	0,1186	1
Αναλογία Δημογραφικής Ισότητας	0,7826	0

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



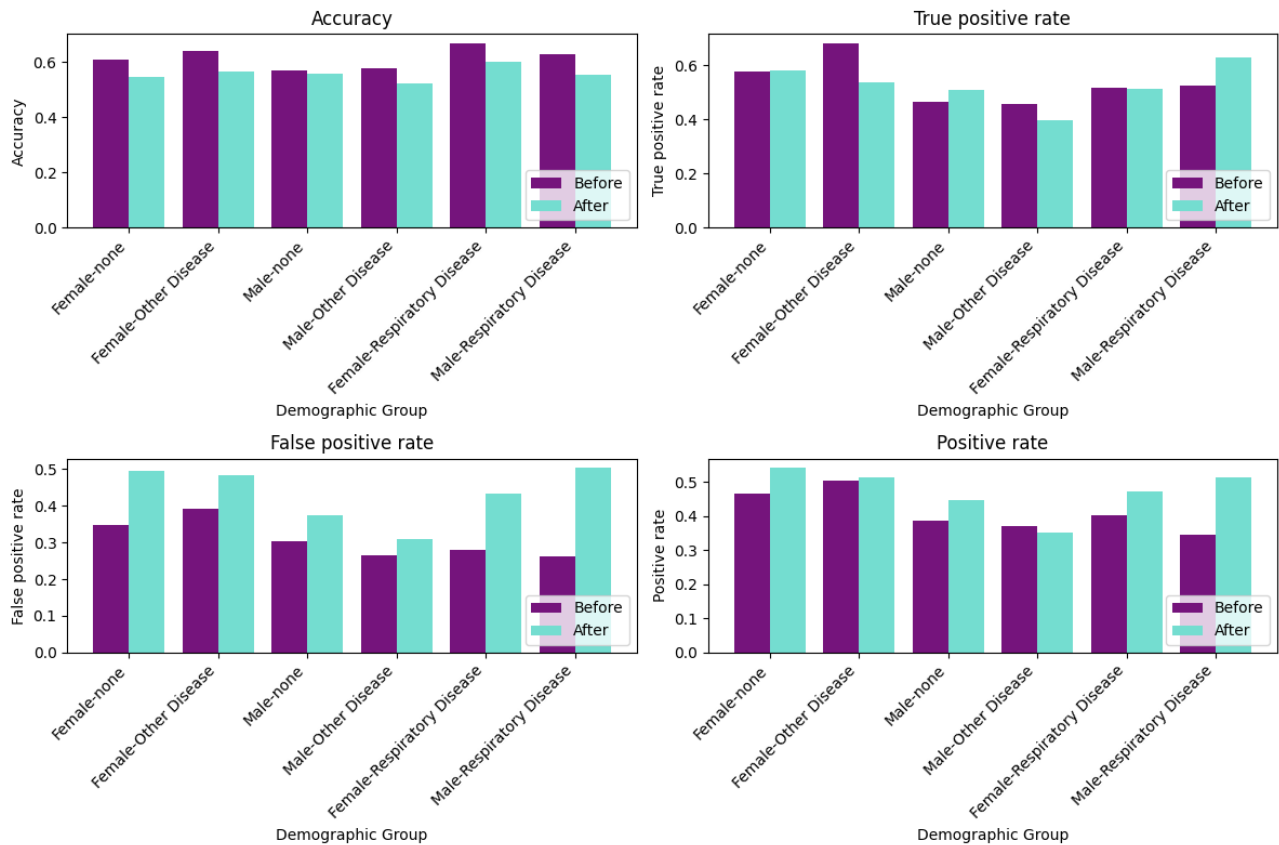
Εικόνα 7.2.47: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή εκθετικής κλίσης για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό πριν και μετά την μέθοδο εκθετικής κλίσης είναι:

Πίνακας 7.2.38: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή εκθετικής κλίσης για την εξάλειψη των διακρίσεων του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Εκθετική Κλίση ως προς το Φύλο σε συνδυασμό με το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,1792	0,1525
Αναλογία Ισότητας Ακρίβειας	0,7231	0,761
Διαφορά Ισότητας Πιθανότητας	0,2397	0,21
Αναλογία Ισότητας Πιθανότητας	0,5702	0,6136
Διαφορά Ίσης Ευκαιρίας	0,1936	0,1604
Αναλογία Ίσης Ευκαιρίας	0,6862	0,737
Διαφορά Δημογραφικής Ισότητας	0,1186	0,1271
Αναλογία Δημογραφικής Ισότητας	0,7826	0,7675

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



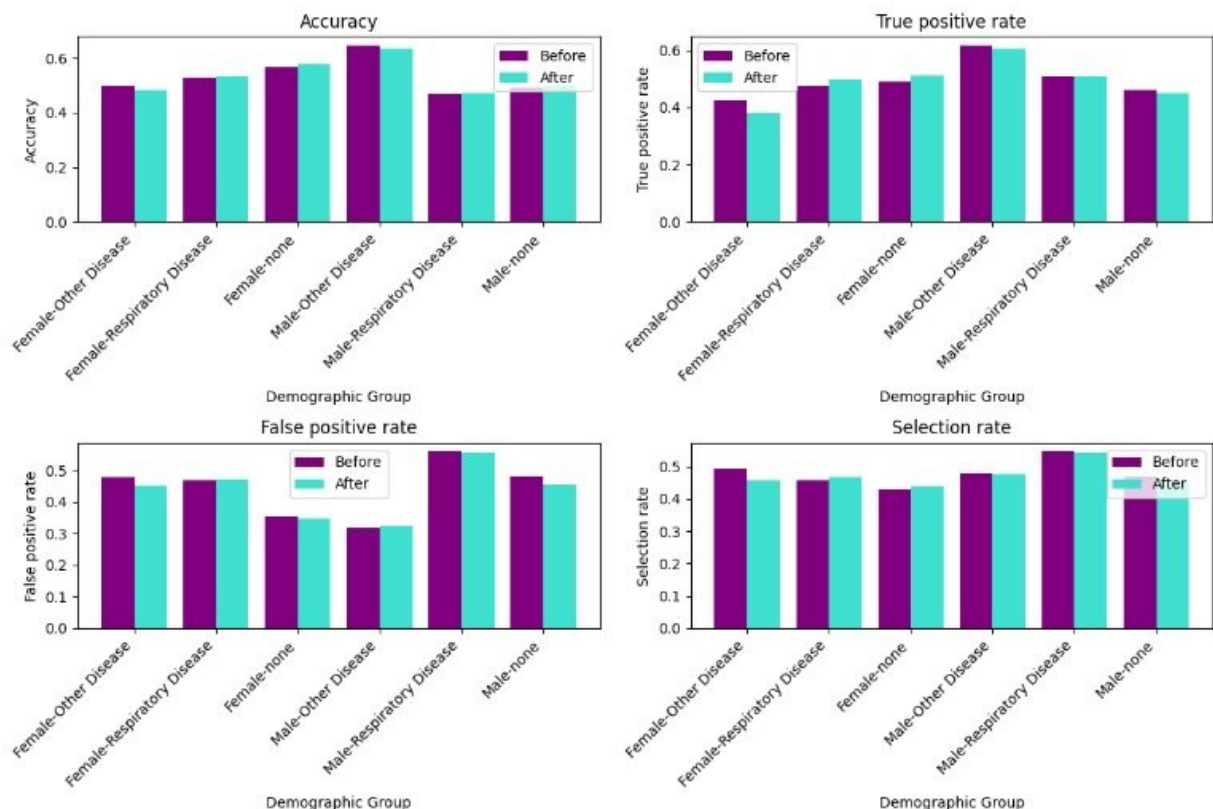
Εικόνα 7.2.48: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό πριν και μετά την μέθοδο ανταγωνιστικής εκπαίδευσης είναι:

Πίνακας 7.2.39: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

Μετρικές Αμεροληψίας	Συνελκτικό Νευρωνικό Δίκτυο	μετά από Αντίστροφη Εκπαίδευση ως προς το Φύλο σε συνδυασμό με το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,0984	0,0797
Αναλογία Ισότητας Ακρίβειας	0,8529	0,8678
Διαφορά Ισότητας Πιθανότητας	0,2225	0,2331
Αναλογία Ισότητας Πιθανότητας	0,6671	0,6137
Διαφορά Ίσης Ευκαιρίας	0,2225	0,2331
Αναλογία Ίσης Ευκαιρίας	0,6734	0,6292
Διαφορά Δημογραφικής Ισότητας	0,1582	0,1893
Αναλογία Δημογραφικής Ισότητας	0,6862	0,6502

Οι τιμές της ακρίβειας, του ρυθμού αληθώς θετικών δειγμάτων, του ρυθμού ψευδώς θετικών δειγμάτων και του ρυθμού θετικών προβλέψεων για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό ως ευαίσθητο χαρακτηριστικό, πριν και μετά την εφαρμογή της μεθόδου του βελτιστοποιητή κατωφλίου για την ενίσχυση της αμεροληψίας, παρουσιάζονται ακολούθως:



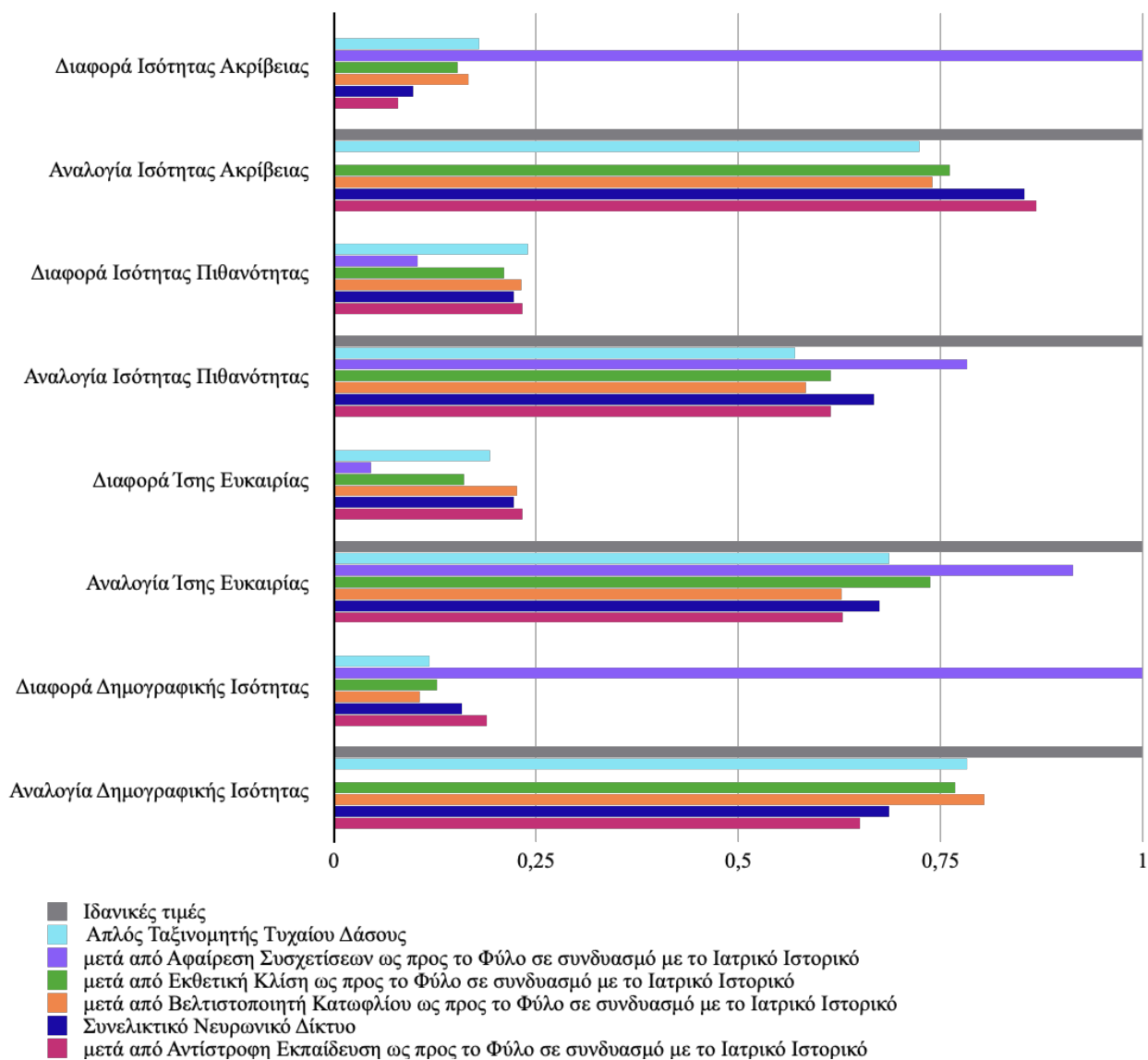
Εικόνα 7.2.49: Γραφική απεικόνιση της μεταβολής των τιμών ακρίβειας, ρυθμού αληθώς θετικών δειγμάτων, ρυθμού ψευδώς θετικών δειγμάτων και ρυθμού θετικών προβλέψεων πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για κάθε υποομάδα του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

Επομένως, οι τιμές των μετρικών αμεροληψίας για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό πριν και μετά την μέθοδο βελτιστοποιητή κατωφλίου είναι:

Πίνακας 7.2.40: Πίνακας τιμών των μετρικών αμεροληψίας πριν και μετά την εφαρμογή βελτιστοποιητή κατωφλίου για την εξάλειψη των διακρίσεων του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

Μετρικές Αμεροληψίας	Απλός Ταξινομητής Τυχαίου Δάσους	μετά από Βελτιστοποιητή Κατωφλίου ως προς το Φύλο σε συνδυασμό με το Ιατρικό Ιστορικό
Διαφορά Ισότητας Ακρίβειας	0,1792	0,1659
Αναλογία Ισότητας Ακρίβειας	0,7231	0,7396
Διαφορά Ισότητας Πιθανότητας	0,2397	0,2315
Αναλογία Ισότητας Πιθανότητας	0,5702	0,5827
Διαφορά Ίσης Ευκαιρίας	0,1936	0,2259
Αναλογία Ίσης Ευκαιρίας	0,6862	0,6271
Διαφορά Δημογραφικής Ισότητας	0,1186	0,1065
Αναλογία Δημογραφικής Ισότητας	0,7826	0,804

Οι τιμές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή όλων των μεθόδων εξάλειψης των διακρίσεων ως προς το φύλο σε συνδυασμό με το ιατρικό ιστορικό, συγκριτικά και με τις ιδανικές τιμές τους, παρουσιάζονται ακολούθως:



Εικόνα 7.2.50: Γραφική απεικόνιση των τιμών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή μεθόδων εξάλειψης των διακρίσεων και σε σύγκριση με τις ιδανικές τιμές, για την κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό.

7.2.11 Μεταβολές Μετρικών Αμεροληψίας μετά την εφαρμογή Μεθόδων Εξάλειψης των Διακρίσεων

Οι μεταβολές των τιμών κάθε μετρικής αμεροληψίας, πριν και μετά την εφαρμογή της μεθόδου αφαίρεσης συσχετίσεων, για όλες τις κατηγορίες ευαίσθητων χαρακτηριστικών, παρουσιάζονται ακολούθως:

Πίνακας 7.2.41: Πίνακας μεταβολών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή της αφαίρεσης συσχετίσεων. Με πράσινο χρώμα επισημαίνονται οι βελτιωμένες τιμές, ενώ με κόκκινο οι επιδεινωμένες.

ΑΦΑΙΡ. ΣΥΣΧΕΤ.	ΦΥΛΟ	ΗΛΙΚΙΑ	ΙΑΤΡ. ΙΣ.	ΣΥΜ.	ΠΗΓΗ ΔΕΔ.	ΗΛ.- ΦΥΛΟ	ΗΛ.- ΣΥΜ.	ΦΥΛΟ- ΣΥΜ.	ΗΛ.- ΙΑΤΡ. ΙΣ.	ΦΥΛΟ- ΙΑΤΡ. ΙΣ.
Δ. ΙΣΟΤ. ΑΚΡ.	591.82%	6002.76 %	14.75%	-12.44%	436.94%	372.97%	-1.67%	550.92%	321.14%	457.94%
Α. ΙΣΟΤ. ΑΚΡ.	-46.73%	-28.28%	-2.94%	1.62%	-18.52%	-100.00%	2.24%	-100.0%	-100.00%	-100.00%
Δ. ΙΣΟΤ. ΠΙΘ.	-75.28%	-99.64%	-88.76%	-96.47%	-86.77%	-78.97%	-87.16%	-35.02%	-61.10%	-57.00%
Α. ΙΣΟΤ. ΠΙΘ.	14.85%	66.16%	38.45%	40.38%	160.18%	62.25%	148.61%	35.25%	50.39%	37.25%
Δ. ΙΣΗΣ ΕΥΚ.	-31.27%	-99.64%	-82.56%	-98.11%	-86.77%	-92.89%	-91.52%	-66.44%	-76.70%	-76.13%
Α. ΙΣΗΣ ΕΥΚ.	1.52%	63.30%	13.97%	15.50%	147.42%	82.78%	66.21%	22.19%	39.97%	33.09%
Δ. ΔΗΜ. ΙΣΟΤ.	1332.5%	-5.61%	167.01%	-22.25%	-4.78%	285.62%	26.31%	542.20%	108.55%	743.32%
Α. ΔΗΜ. ΙΣΟΤ.	-42.70%	3.73%	-8.97%	5.55%	5.89%	-100.00%	-100.00%	-100.00%	-30.89%	-100.00%

Αντιστοίχως, οι μεταβολές των τιμών κάθε μετρικής αμεροληψίας, πριν και μετά την εφαρμογή της μεθόδου εκθετικής κλίσης, για όλες τις κατηγορίες ευαίσθητων χαρακτηριστικών, παρουσιάζονται στον ακόλουθο πίνακα:

Πίνακας 7.2.42: Πίνακας μεταβολών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή της εκθετικής κλίσης. Με πράσινο χρώμα επισημαίνονται οι βελτιωμένες τιμές, ενώ με κόκκινο οι επιδεινωμένες.

ΕΚΘ. ΚΛΙΣΗ	ΦΥΛΟ	ΗΛΙΚΙΑ	ΙΑΤΡ. ΙΣ.	ΣΥΜ.	ΠΗΓΗ ΔΕΔ.	ΗΛ.- ΦΥΛΟ	ΗΛ.- ΣΥΜ.	ΦΥΛΟ- ΣΥΜ.	ΗΛ.-ΙΑΤΡ. ΙΣ.	ΦΥΛΟ- ΙΑΤΡ. ΙΣ.
Δ. ΙΣΟΤ. ΑΚΡ.	-0.78%	285.51%	15.17%	0.19%	0%	-4.96%	17.49%	9.22%	-13.91%	-14.91%
Α. ΙΣΟΤ. ΑΚΡ.	0.11%	-1.37%	-3.00%	-0.29%	0%	2.31%	-18.89%	-2.29%	5.61%	5.24%
Δ. ΙΣΟΤ. ΠΙΘ.	-1.32%	-6.60%	9.31%	17.50%	0%	-2.23%	10.14%	-3.24%	10.68%	-12.41%
Α. ΙΣΟΤ. ΠΙΘ.	0.16%	1.96%	0.58%	-6.69%	0%	-2.53%	-0.45%	1.56%	3.88%	7.62%
Δ. ΙΣΗΣ ΕΥΚ.	2.02%	-6.60%	1.72%	-4.74%	0%	-2.23%	63.75%	-17.69%	59.79%	-17.15%
Α. ΙΣΗΣ ΕΥΚ.	-0.07%	2.74%	0.12%	0.41%	0%	3.00%	-39.07%	4.05%	-17.26%	7.40%
Δ. ΔΗΜ. ΙΣΟΤ.	-2.68%	-7.40%	154.21%	-4.95%	0%	-0.49%	-5.59%	7.30%	-29.54%	7.20%
Α. ΔΗΜ. ΙΣΟΤ.	0.15%	2.57%	-8.13%	1.15%	0%	-0.94%	-2.75%	-3.09%	13.46%	-1.94%

Οι μεταβολές των τιμών κάθε μετρικής αμεροληψίας, πριν και μετά την εφαρμογή της μεθόδου ανταγωνιστικής εκπαίδευσης, για όλες τις κατηγορίες ευαίσθητων χαρακτηριστικών, παρουσιάζονται επίσης στον ακόλουθο πίνακα:

Πίνακας 7.2.43: Πίνακας μεταβολών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης. Με πράσινο χρώμα επισημαίνονται οι βελτιωμένες τιμές, ενώ με κόκκινο οι επιδεινωμένες.

ΑΝΤ. ΕΚΠΑΙΔ.	ΦΥΛΟ	ΗΛΙΚΙΑ	ΙΑΤΡ. ΙΣ.	ΣΥΜ.	ΠΗΓΗ ΔΕΔ.	ΗΛ.-ΦΥΛΟ	ΗΛ.-ΣΥΜ.	ΦΥΛΟ-ΣΥΜ.	ΗΛ.-ΙΑΤΡ. ΙΣ.	ΦΥΛΟ-ΙΑΤΡ. ΙΣ.
Δ. ΙΣΟΤ. ΑΚΡ.	-57.26%	134.65%	-32.69%	20.97%	-27.84%	169.38%	9.20%	-41.17%	4.60%	-19.00%
Α. ΙΣΟΤ. ΑΚΡ.	+4.31%	-3.28%	2.81%	-2.43%	0.64%	-16.89%	-15.57%	6.88%	-5.22%	1.75%
Δ. ΙΣΟΤ. ΠΙΘ.	-60.97%	-1.19%	96.79%	-30.42%	0.05%	-43.21%	-3.22%	-33.05%	13.22%	4.76%
Α. ΙΣΟΤ. ΠΙΘ.	+11.3%	4.91%	-1.31%	14.45%	-0.08%	15.47%	40.18%	53.17%	-1.94%	-8.00%
Δ. ΙΣΗΣ ΕΥΚ.	-62.08%	139.42%	722.32%	-30.42%	0.05%	-43.21%	-3.22%	-33.05%	13.22%	4.76%
Α. ΙΣΗΣ ΕΥΚ.	+14.97 %	-6.39%	-13.78%	14.45%	-6.99%	15.47%	40.18%	24.57%	-18.39%	-6.56%
Δ. ΔΗΜ. ΙΣΟΤ.	-45.2%	303.9%	17.79%	-47.65%	-6.51%	-32.81%	-23.37%	-38.36%	29.30%	19.66%
Α. ΔΗΜ. ΙΣΟΤ.	+9.58%	-9.09%	0.01%	25.01%	-0.73%	9.76%	83.73%	30.80%	-11.86%	-5.25%

Τέλος, οι μεταβολές των τιμών κάθε μετρικής αμεροληψίας, πριν και μετά την εφαρμογή της μεθόδου βελτιστοποιητή κατωφλίου, για όλες τις κατηγορίες ευαίσθητων χαρακτηριστικών, παρουσιάζονται ακολούθως:

Πίνακας 7.2.44: Πίνακας μεταβολών των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή του βελτιστοποιητή κατωφλίου. Με πράσινο χρώμα επισημαίνονται οι βελτιωμένες τιμές, ενώ με κόκκινο οι επιδεινωμένες.

ΒΕΛΤ. ΚΑΤΩΦ.	ΦΥΛΟ	ΗΛΙΚΙΑ	ΙΑΤΡ. ΙΣ.	ΣΥΜ.	ΠΗΓΗ ΔΕΔ.	ΗΛ.-ΦΥΛΟ	ΗΛ.-ΣΥΜ.	ΦΥΛΟ-ΣΥΜ.	ΗΛ.-ΙΑΤΡ. ΙΣ.	ΦΥΛΟ-ΙΑΤΡ. ΙΣ.
Δ. ΙΣΟΤ. ΑΚΡ.	-41.56%	112.80%	-8.75%	17.65%	20.39%	0.58%	20.71%	0.06%	23.90%	-7.45%
Α. ΙΣΟΤ. ΑΚΡ.	4.09%	-0.54%	1.35%	-2.39%	-1.06%	-1.39%	-20.19%	-0.20%	-4.99%	2.29%
Δ. ΙΣΟΤ. ΠΙΘ.	15.41%	0.56%	-16.25%	62.19%	-39.51%	-14.64%	10.14%	-0.32%	-0.32%	-3.42%
Α. ΙΣΟΤ. ΠΙΘ.	-1.91%	4.90%	6.90%	-21.55%	63.56%	-1.75%	-0.45%	9.90%	9.90%	2.20%
Δ. ΙΣΗΣ ΕΥΚ.	-5.27%	0.56%	18.50%	71.20%	-39.51%	-14.64%	63.75%	3.23%	78.03%	16.69%
Α. ΙΣΗΣ ΕΥΚ.	0.50%	3.10%	-2.94%	-10.37%	55.54%	4.24%	-39.07%	-1.75%	-29.59%	-8.62%
Δ. ΔΗΜ. ΙΣΟΤ.	107.38 %	-5.95%	-20.68%	42.56%	-42.68%	0%	-7.79%	23.44%	3.71%	-10.23%
Α. ΔΗΜ. ΙΣΟΤ.	-5.45%	5.34%	1.26%	-10.63%	61.45%	-1.06%	1.15%	-7.17%	-6.97%	2.73%

8 Συμπεράσματα – Μελλοντικές Προεκτάσεις

Η σημασία της αμεροληψίας στην μηχανική μάθηση και ιδιαίτερα σε προβλήματα που σχετίζονται με την ιατρική, γίνεται ολοένα και πιο αντιληπτή στην επιστημονική κοινότητα. Όσο η τεχνητή νοημοσύνη εισβάλλει στην καθημερινότητά μας, το κοινό στο οποίο απευθύνονται οι εφαρμογές μηχανικής μάθησης διευρύνεται, με αποτέλεσμα να αυξάνεται και η ποικιλία των ευαίσθητων χαρακτηριστικών που το περιγράφουν, πολλές φορές επηρεάζοντας και τα αποτελέσματά τους.

Δημιουργείται, συνεπώς, μεγάλη ανάγκη για διερεύνηση του βαθμού αμεροληψίας των εφαρμογών αυτών, καθώς και της αποτελεσματικότητας των μεθόδων αντιμετώπισης τυχόν διακρίσεων που προκύπτουν μεταξύ διαφορετικών ομάδων του πληθυσμού, ώστε να επιτύχουμε δίκαιες προβλέψεις που δεν θα εξαρτώνται από τα ευαίσθητα χαρακτηριστικά. Ειδικότερα σε περιόδους παγκόσμιας κρίσης, όπως η πρόσφατη πανδημία που προκάλεσε η εξάπλωση της νόσου COVID-19, οι εφαρμογές μηχανικής μάθησης προσφέρονται ως γρήγορα, ανέξοδα και προσβάσιμα από μεγάλη μερίδα ανθρώπων εργαλεία διάγνωσης που μπορούν, λειτουργώντας συνδυαστικά με τις ιατρικές πρακτικές, να επιφέρουν μεγαλύτερη αποτελεσματικότητα και διευκόλυνση στο έργο των επιστημόνων υγείας. Στην παρούσα εργασία, εξετάστηκε ο βαθμός αμεροληψίας στη διάγνωση της νόσου COVID-19 από ηχητικά σήματα βήχα σε δύο μοντέλα, έναν ταξινομητή τυχαίου δάσους και ένα συνελκτικό νευρωνικό δίκτυο. Στη συνέχεια, εφαρμόστηκαν τέσσερις διαφορετικές μέθοδοι εξάλειψης των διακρίσεων, αντιπροσωπευτικές τόσο ως προς το είδος των μοντέλων μηχανικής μάθησης που χρησιμοποιήσαμε, όσο και ως προς το είδος των μεθόδων εξάλειψης των διακρίσεων ανάλογα με το στάδιο επεξεργασίας των δεδομένων κατά το οποίο εφαρμόζονται (μέθοδοι προεπεξεργασίας, κατά την επεξεργασία και μετά την επεξεργασία). Αναλυτικά, οι συνδυασμοί μεθόδων και μοντέλων στα οποία εφαρμόστηκαν είναι: αφαίρεση συσχετίσεων και εισαγωγή των νέων δεδομένων σε ταξινομητή τυχαίου δάσους, ταξινομητής τυχαίου δάσους με μέθοδο εκθετικής κλίσης, συνελκτικό νευρωνικό δίκτυο με ανταγωνιστική εκπαίδευση και ταξινομητής τυχαίου δάσους με βελτιστοποιητή κατωφλίου. Συγκρίνοντας τα αποτελέσματα τόσο της απόδοσης των προβλέψεων όσο και της αμεροληψίας που προέκυψαν μετά την εφαρμογή των παραπάνω μεθόδων, σε σχέση με τα αρχικά αποτελέσματα, οδηγούμαστε στις ακόλουθες βασικές παρατηρήσεις:

Η απόδοση των μοντέλων όσον αφορά τις προβλέψεις, δεν είναι ιδιαίτερα υψηλή και μάλιστα στη περίπτωση του συνελκτικού νευρωνικού δικτύου παρατηρείται πτώση στην απόδοση μετά την εφαρμογή της ανταγωνιστικής εκπαίδευσης για την εξάλειψη των διακρίσεων. Οι αποδόσεις στα μοντέλα του ταξινομητή τυχαίου δάσους, κυμαίνονται σε παρόμοιες τιμές, με μια μικρή αυξητική τάση μετά την εφαρμογή μεθόδων ενίσχυσης της αμεροληψίας. Συνολικά, ωστόσο, οι τιμές απόδοσης στις προβλέψεις των μοντέλων δεν παρουσιάζουν μεγάλες αποκλίσεις μεταξύ τους, ανεξάρτητα από το ποια μέθοδο εφαρμόσαμε κάθε φορά στο συγκεκριμένο πείραμα.

Παρατηρώντας τις μεταβολές των μετρικών αμεροληψίας, πριν και μετά την εφαρμογή της αφαίρεσης συσχετίσεων, μπορούμε να συμπεράνουμε ότι η μέθοδος αυτή πετυχαίνει γενικά σημαντική βελτίωση στην ισότητα πιθανότητας και στην ίση ευκαιρία, με τις τιμές των μετρικών να πλησιάζουν τις ιδανικές μετά την εφαρμογή της. Η μέθοδος αυτή όμως, φαίνεται να προκαλεί μεγάλη επιδείνωση της ισότητας ακρίβειας και της δημογραφικής ισότητας, προκαλώντας μεγάλες αποκλίσεις από τις ιδανικές τιμές τους.

Η μέθοδος της εκθετικής κλίσης πετυχαίνει βελτιώσεις των μετρικών αμεροληψίας σε αρκετές κατηγορίες ευαίσθητων χαρακτηριστικών, συμπεριλαμβανομένων και κατηγοριών που προέκυψαν από τη συνένωση απλούστερων, όπως για παράδειγμα, η ηλικία σε συνδυασμό με το φύλο και το φύλο σε συνδυασμό με το ιατρικό ιστορικό. Η βελτίωση που επιτυγχάνεται στις κατηγορίες αυτές κρίνεται ιδιαίτερα σημαντική, καθώς εκεί παρατηρούνται και εξ αρχής οι μεγαλύτερες αποκλίσεις από τις ιδανικές τιμές, λόγω της κατάτμησης του συνόλου σε περισσότερες υποομάδες.

Μετά την εφαρμογή της μεθόδου της ανταγωνιστικής εκπαίδευσης, επιτυγχάνουμε βελτίωση σε αντίστοιχο πλήθος μετρικών αμεροληψίας με τη μέθοδο της εκθετικής κλίσης, με μεγαλύτερα μάλιστα ποσοστά βελτίωσης. Ειδικότερα, αν λάβουμε υπόψη τις μετρικές εκτός της ισότητας ακρίβειας, η ανταγωνιστική εκπαίδευση παρουσιάζει μεγαλύτερη ομοιομορφία στις μεταβολές, καθώς η επιδείνωση είναι εστιασμένη κυρίως σε τέσσερις κατηγορίες: την ηλικία, το ιατρικό ιστορικό, την ηλικία σε συνδυασμό με το ιατρικό ιστορικό και το φύλο σε συνδυασμό με το ιατρικό ιστορικό.

Παρατηρείται, επίσης, ότι οι κατηγορίες της ηλικίας και του ιατρικού ιστορικού που δεν παρουσιάζουν βελτίωση στις διακρίσεις, συνδυαστικά επιδεινώνονται, καθώς η κατηγορία της ηλικίας σε συνδυασμό με το ιατρικό ιστορικό έχει τη μεγαλύτερη απόκλιση από τις ιδανικές τιμές των μετρικών αμεροληψίας. Ακόμη, η κατηγορία του φύλου σε συνδυασμό με το ιατρικό ιστορικό, που αποτελεί συνένωση της επιδεινωμένης κατηγορίας του ιατρικού ιστορικού και της βελτιωμένης κατηγορίας του φύλου, παρουσιάζει μεν σημαντική απόκλιση από τις ιδανικές τιμές, αλλά με μικρότερα ποσοστά μεταβολής σε σύγκριση με τις τιμές πριν την εφαρμογή της ανταγωνιστικής εκπαίδευσης.

Επιπλέον, επισημαίνεται, ότι όσον αφορά την εφαρμογή της ανταγωνιστικής εκπαίδευσης στο συνελκτικό νευρωνικό δίκτυο, παρά το γεγονός ότι στις κατηγορίες της ηλικίας και του ιατρικού ιστορικού παρατηρείται κατά βάση επιδείνωση των δεικτών αμεροληψίας, με σχετικά υψηλά ποσοστά μεταβολής, τελικώς, η επιδείνωση αυτή δεν συνεπάγεται και σημαντικά μεγάλες αποκλίσεις από τις ιδανικές τιμές, καθώς οι αποκλίσεις αυτές παρουσιάζονταν ήδη ιδιαίτερα μικρές και πριν την εφαρμογή της μεθόδου εξάλειψης των διακρίσεων.

Η μέθοδος του βελτιστοποιητή κατωφλίου, φαίνεται να είναι η λιγότερο αποτελεσματική μέθοδος στο παρόν πείραμα καθώς, ενώ πετυχαίνει βελτίωση σε ορισμένες τιμές των μετρικών αμεροληψίας, το πλήθος των τιμών που βελτιώνονται είναι μικρότερο σε σχέση με τις βελτιωμένες τιμές που παρατηρούμε στα άλλα μοντέλα που εφαρμόζονται οι υπόλοιπες μέθοδοι ενίσχυσης της αμεροληψίας. Επιπλέον, εκτός από την επιδείνωση των τιμών, σημαντική είναι και η ίδια η απόκλιση των τιμών των μετρικών αμεροληψίας από τις ιδανικές, η οποία στις περισσότερες κατηγορίες είναι μεγαλύτερη απ'ό,τι στα υπόλοιπα μοντέλα.

Γίνεται λοιπόν αντιληπτό, ότι οι μέθοδοι εξάλειψης των διακρίσεων επιτυγχάνουν σε ικανοποιητικό βαθμό την βελτίωση των μετρικών της αμεροληψίας σε αρκετές κατηγορίες ευαίσθητων χαρακτηριστικών, χωρίς ωστόσο να πετυχαίνουν ιδανικά αποτελέσματα ή καθολική βελτίωση των διαφορετικών κατηγοριών. Παρόλα αυτά, καθίσταται σαφές ότι η αμεροληψία είναι ένας σημαντικός παράγοντας της αποτελεσματικότητας ενός προβλεπτικού μοντέλου και άρα αποκτά ιδιαίτερη σημασία η περαιτέρω διερεύνηση τρόπων ενίσχυσής της. Ιδιαίτερη πρόκληση αποτελεί και η αναζήτηση του τρόπου με τον οποίο επηρεάζει η επιδίωξη της αμεροληψίας τον βαθμό ικανότητας πρόβλεψης των μοντέλων και η εύρεση της κατάλληλης ισορροπίας μεταξύ της καλύτερης δυνατής απόδοσης και της αμεροληψίας στις προβλέψεις. Ακόμη, οι τμηματικές βελτιώσεις στις μετρικές αμεροληψίας που παρατηρήσαμε μετά την εφαρμογή των διαφόρων μεθόδων με τις οποίες πειραματιστήκαμε, εγείρει το ενδιαφέρον για μελλοντική διερεύνηση της επίδρασης του συνδυασμού διαφορετικών μεθόδων εξάλειψης των διακρίσεων στην απόδοση και την αμεροληψία των μοντέλων. Τέλος, η συγκέντρωση μεγαλύτερου και πιο συμπεριληπτικού δείγματος, το οποίο θα περιλαμβάνει ευρύτερη ποικιλία ευαίσθητων χαρακτηριστικών σε μεγαλύτερη αριθμητική ισορροπία για κάθε κατηγορία, μπορεί να συμβάλει στην περαιτέρω βελτίωση της αμεροληψίας των μοντέλων πρόβλεψης σε πιο εξειδικευμένες υποκατηγορίες πληθυσμού.

Βιβλιογραφία

- [1] A. Çalica Utku, G. Budak, O. Karabay, E. Güçlü, H. D. Okan, and A. Vatan, “Main symptoms in patients presenting in the COVID-19 period,” *Scott. Med. J.*, vol. 65, no. 4, pp. 127–132, Nov. 2020, doi: 10.1177/0036933020949253.
- [2] W. H. Organization, “Coronavirus disease 2019 (COVID-19): situation report, 51,” Mar. 2020, Accessed: Jan. 21, 2025. [Online]. Available: <https://iris.who.int/handle/10665/331475>
- [3] “COVID-19 deaths | WHO COVID-19 dashboard,” datadot. Accessed: Jan. 21, 2025. [Online]. Available: <https://data.who.int/dashboards/covid19/cases>
- [4] A. C. Timmons *et al.*, “A Call to Action on Assessing and Mitigating Bias in Artificial Intelligence Applications for Mental Health,” *Perspect. Psychol. Sci. J. Assoc. Psychol. Sci.*, vol. 18, no. 5, pp. 1062–1096, Sep. 2023, doi: 10.1177/17456916221134490.
- [5] S. Ding *et al.*, “Fairly Predicting Graft Failure in Liver Transplant for Organ Assigning,” *AMIA. Annu. Symp. Proc.*, vol. 2022, p. 415, Apr. 2023.
- [6] H. M. Thompson *et al.*, “Bias and fairness assessment of a natural language processing opioid misuse classifier: detection and mitigation of electronic health record data disadvantages across racial subgroups,” *J. Am. Med. Inform. Assoc. JAMIA*, vol. 28, no. 11, pp. 2393–2403, Oct. 2021, doi: 10.1093/jamia/ocab148.
- [7] J. Han *et al.*, “Sounds of COVID-19: exploring realistic performance of audio-based digital testing,” *Npj Digit. Med.*, vol. 5, no. 1, pp. 1–9, Jan. 2022, doi: 10.1038/s41746-021-00553-x.
- [8] A. Kondaka, “Evaluating Gender Bias and Fairness in Skin Lesion Diagnoses using Convolutional Neural Networks,” 2024.
- [9] T. Xia *et al.*, “COVID-19 Sounds: A Large-Scale Audio Dataset for Digital Respiratory Screening”.
- [10] D. Bhattacharya *et al.*, “Coswara: A respiratory sounds and symptoms dataset for remote screening of SARS-CoV-2 infection,” *Sci. Data*, vol. 10, no. 1, p. 397, Jun. 2023, doi: 10.1038/s41597-023-02266-0.
- [11] L. Orlandic, T. Teijeiro, and D. Atienza, “The COUGHVID crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms,” *Sci. Data*, vol. 8, no. 1, p. 156, Jun. 2021, doi: 10.1038/s41597-021-00937-4.
- [12] R. Pfeifer, S. Vhaduri, and J. E. Dietz, “Mitigating Sex Bias in Audio Data-driven COPD and COVID-19 Breathing Pattern Detection Models,” Sep. 16, 2024, *arXiv*: arXiv:2409.10677. doi: 10.48550/arXiv.2409.10677.
- [13] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, “A Survey on Bias and Fairness in Machine Learning,” *ACM Comput Surv*, vol. 54, no. 6, p. 115:1–115:35, Apr. 2021, doi: 10.1145/3457607.
- [14] “Coronavirus.” Accessed: Jan. 21, 2025. [Online]. Available: <https://www.who.int/health-topics/coronavirus>
- [15] 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.
- [16] L. Steardo, L. Steardo, R. Zorec, and A. Verkhatsky, “Neuroinfection may contribute to pathophysiology and clinical manifestations of COVID-19,” *Acta Physiol. Oxf. Engl.*, vol. 229, no. 3, p. e13473, Jul. 2020, doi: 10.1111/apha.13473.
- [17] “Οδηγίες για την διαχείριση ύποπτου κρούσματος COVID-19 σε τουριστικό κατάλυμα - Εθνικός Οργανισμός Δημόσιας Υγείας.” Accessed: Jan. 21, 2025. [Online]. Available: <https://eody.gov.gr/odigies-gia-tin-diacheirisi-ypoptoy-kroysmatos-covid-19-se-toyristiko-katalyma/?print=print>

- [18] C. Li, C. Zhao, J. Bao, B. Tang, Y. Wang, and B. Gu, “Laboratory diagnosis of coronavirus disease-2019 (COVID-19),” *Clin. Chim. Acta Int. J. Clin. Chem.*, vol. 510, pp. 35–46, Nov. 2020, doi: 10.1016/j.cca.2020.06.045.
- [19] “Which type of COVID-19 test do I need.” Accessed: Jan. 25, 2025. [Online]. Available: <https://www.who.int/mongolia/multi-media/item/which-type-of-covid-19-test-do-i-need>
- [20] “COVID-19 cases | WHO COVID-19 dashboard,” datadot. Accessed: Jan. 21, 2025. [Online]. Available: <https://data.who.int/dashboards/covid19/cases>
- [21] Y. Zhang, *New Advances in Machine Learning*. BoD – Books on Demand, 2010.
- [22] R. Y. Choi, A. S. Coyner, J. Kalpathy-Cramer, M. F. Chiang, and J. P. Campbell, “Introduction to Machine Learning, Neural Networks, and Deep Learning,” *Transl. Vis. Sci. Technol.*, vol. 9, no. 2, p. 14, Feb. 2020, doi: 10.1167/tvst.9.2.14.
- [23] C. M. Bishop and H. Bishop, *Deep Learning*. Springer, 2024.
- [24] A. Georgouli and A. Γεωργούλη, “Μηχανική Μάθηση,” Feb. 2016, Accessed: Jan. 22, 2025. [Online]. Available: <http://beta.kallipos.gr/jspui/handle/11419/3382>
- [25] Y. C. A. P. Reddy, P. Viswanath, and B. E. Reddy, “Semi-supervised learning: a brief review,” *Int. J. Eng. Technol.*, vol. 7, no. 1.8, Art. no. 1.8, Feb. 2018, doi: 10.14419/ijet.v7i1.8.9977.
- [26] J. Dunkelau, “Fairness-Aware Machine Learning An Extensive Overview,” 2020. Accessed: Jan. 25, 2025. [Online]. Available: <https://www.semanticscholar.org/paper/Fairness-Aware-Machine-Learning-An-Extensive-Dunkelau/f4cea2cd1c29707d386da8bf0734934ea23bd488>
- [27] M. Kubat, *Εισαγωγή στη Μηχανική Μάθηση*. Εκδόσεις Foundas, 2021.
- [28] “(PDF) Classification Techniques in Machine Learning: Applications and Issues,” *ResearchGate*, Oct. 2024, Accessed: Jan. 22, 2025. [Online]. Available: https://www.researchgate.net/publication/319370844_Classification_Techniques_in_Machine_Learning_Applications_and_Issues
- [29] S. A. Bini, “Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care?,” *J. Arthroplasty*, vol. 33, no. 8, pp. 2358–2361, Aug. 2018, doi: 10.1016/j.arth.2018.02.067.
- [30] L. Alzubaidi *et al.*, “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” *J. Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.
- [31] “Comparative Study of First Order Optimizers for Image Classification Using Convolutional Neural Networks on Histopathology Images.” Accessed: Jan. 24, 2025. [Online]. Available: <https://www.mdpi.com/2313-433X/6/9/92>
- [32] K. Νικήτα, *Προσομοίωση Φυσιολογικών Συστημάτων*. Εκδόσεις TZIOΛΑ, 2011.
- [33] “What Is Cross-Entropy Loss Function? – 365 Data Science.” Accessed: May 18, 2025. [Online]. Available: <https://365datascience.com/tutorials/machine-learning-tutorials/cross-entropy-loss/>
- [34] “Convolutional Neural Networks: A Survey.” Accessed: Jan. 24, 2025. [Online]. Available: <https://www.mdpi.com/2073-431X/12/8/151>
- [35] C. Camacho, “Convolutional Neural Networks.” Accessed: Jan. 25, 2025. [Online]. Available: https://cezannec.github.io/Convolutional_Neural_Networks/
- [36] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” Apr. 10, 2015, *arXiv*: arXiv:1409.1556. doi: 10.48550/arXiv.1409.1556.
- [37] S. Hershey *et al.*, “CNN architectures for large-scale audio classification,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2017, pp. 131–135. doi: 10.1109/ICASSP.2017.7952132.
- [38] “Generalization of spectral flatness measure for non-Gaussian linear processes | IEEE Journals & Magazine | IEEE Xplore.” Accessed: Feb. 02, 2025. [Online]. Available: <https://ieeexplore.ieee.org/document/1316889?arnumber=1316889>

- [39] Md. A. Hossan, S. Memon, and M. A. Gregory, "A novel approach for MFCC feature extraction," in *2010 4th International Conference on Signal Processing and Communication Systems*, Sep. 2010, pp. 1–5. doi: 10.1109/ICSPCS.2010.5709752.
- [40] W. Flores-Fuentes *et al.*, "Combined application of Power Spectrum Centroid and Support Vector Machines for measurement improvement in Optical Scanning Systems," *Signal Process.*, vol. 98, pp. 37–51, Feb. 2014, doi: 10.1016/j.sigpro.2013.11.008.
- [41] M. Ashames and S. Ergin, *Mel-Spectrograms and Data Augmentation for Spoken Digit Classification*. 2021.
- [42] misun_song, "Understanding the ROC-AUC Curve," Medium. Accessed: Jan. 25, 2025. [Online]. Available: <https://medium.com/@msong507/understanding-the-roc-auc-curve-cc204f0b3441>
- [43] K. Zarkogianni, M. Athanasiou, A. C. Thanopoulou, and K. S. Nikita, "Comparison of Machine Learning Approaches Toward Assessing the Risk of Developing Cardiovascular Disease as a Long-Term Diabetes Complication," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 5, pp. 1637–1647, Sep. 2018, doi: 10.1109/JBHI.2017.2765639.
- [44] S. Moguilner, R. Whelan, H. Adams, V. Valcour, E. Tagliazucchi, and A. Ibáñez, "Visual deep learning of unprocessed neuroimaging characterises dementia subtypes and generalises across non-stereotypic samples," *eBioMedicine*, vol. 90, Apr. 2023, doi: 10.1016/j.ebiom.2023.104540.
- [45] M. E. Vrontzou, M. Athanasiou, K. V. Dalakleidi, I. Skampardoni, C. Davatzikos, and K. Nikita, "A comprehensive interpretable machine learning framework for mild cognitive impairment and Alzheimer's disease diagnosis," *Sci. Rep.*, vol. 15, no. 1, p. 8410, Mar. 2025, doi: 10.1038/s41598-025-92577-6.
- [46] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. Zemel, "Fairness through awareness," in *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, in ITCS '12. New York, NY, USA: Association for Computing Machinery, Jan. 2012, pp. 214–226. doi: 10.1145/2090236.2090255.
- [47] M. Hardt, E. Price, E. Price, and N. Srebro, "Equality of Opportunity in Supervised Learning," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2016. Accessed: Jan. 24, 2025. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2016/hash/9d2682367c3935defcb1f9e247a97c0d-Abstract.html
- [48] Z. Xu, J. Li, Q. Yao, H. Li, M. Zhao, and S. K. Zhou, "Addressing fairness issues in deep learning-based medical image analysis: a systematic review," *Npj Digit. Med.*, vol. 7, no. 1, pp. 1–16, Oct. 2024, doi: 10.1038/s41746-024-01276-5.
- [49] M. E. Vrontzou, M. Athanasiou, C. Davatzikos, and K. S. Nikita, "Comparative assessment of fairness definitions and bias mitigation strategies in machine learning-based diagnosis of Alzheimer's disease from MR images," May 29, 2025, *arXiv*: arXiv:2505.23528. doi: 10.48550/arXiv.2505.23528.
- [50] P. Gajane, S. Newman, M. Pechenizkiy, and J. D. Piette, "Investigating Gender Fairness in Machine Learning-driven Personalized Care for Chronic Pain," Jun. 14, 2024, *arXiv*: arXiv:2402.19226. doi: 10.48550/arXiv.2402.19226.
- [51] X. Han *et al.*, "FFB: A Fair Fairness Benchmark for In-Processing Group Fairness Methods," Jun. 11, 2024, *arXiv*: arXiv:2306.09468. doi: 10.48550/arXiv.2306.09468.
- [52] S. Caton and C. Haas, "Fairness in Machine Learning: A Survey," Oct. 04, 2020, *arXiv*: arXiv:2010.04053. doi: 10.48550/arXiv.2010.04053.
- [53] C. Saplicki, "Fairness in Machine Learning: Pre-Processing Algorithms," IBM Data Science in Practice. Accessed: Jan. 26, 2025. [Online]. Available: <https://medium.com/ibm-data-ai/fairness-in-machine-learning-pre-processing-algorithms-a670c031fba8>

- [54] E. S. Adamidi, K. Mitsis, and K. S. Nikita, "Artificial intelligence in clinical care amidst COVID-19 pandemic: A systematic review," *Comput. Struct. Biotechnol. J.*, vol. 19, pp. 2833–2850, Jan. 2021, doi: 10.1016/j.csbj.2021.05.010.
- [55] Md. Z. Islam, Md. M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Inform. Med. Unlocked*, vol. 20, p. 100412, Jan. 2020, doi: 10.1016/j.imu.2020.100412.
- [56] D. Bhattacharya *et al.*, "Coswara: A respiratory sounds and symptoms dataset for remote screening of SARS-CoV-2 infection," *Sci. Data*, vol. 10, no. 1, p. 397, Jun. 2023, doi: 10.1038/s41597-023-02266-0.
- [57] K. Zarkogianni *et al.*, "The smarty4covid dataset and knowledge base as a framework for interpretable physiological audio data analysis," *Sci. Data*, vol. 10, no. 1, p. 770, Nov. 2023, doi: 10.1038/s41597-023-02646-6.
- [58] T. Ganitidis, M. Athanasiou, K. Mitsis, K. Zarkogianni, and K. S. Nikita, "A Comprehensive Drift-Adaptive Framework for Sustaining Model Performance in COVID-19 Detection From Dynamic Cough Audio Data: Model Development and Validation," *J. Med. Internet Res.*, vol. 27, no. 1, p. e66919, Jun. 2025, doi: 10.2196/66919.
- [59] R. Pfeifer, S. Vhaduri, and J. E. Dietz, "Mitigating Sex Bias in Audio Data-driven COPD and COVID-19 Breathing Pattern Detection Models," Sep. 16, 2024, *arXiv*: arXiv:2409.10677. doi: 10.48550/arXiv.2409.10677.
- [60] T. Saeed, A. Ijaz, I. Sadiq, H. N. Qureshi, A. Rizwan, and A. Imran, "An AI-enabled Bias-Free Respiratory Disease Diagnosis Model using Cough Audio: A Case Study for COVID-19," Jan. 04, 2024, *arXiv*: arXiv:2401.02996. doi: 10.48550/arXiv.2401.02996.