

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**ΙΣΧΑΙΜΙΑ ΤΟΥ ΜΥΟΚΑΡΔΙΟΥ ΚΑΙ ΠΡΟΒΛΕΨΗ ΕΠΙΠΛΟΚΩΝ
ΤΗΣ ΜΕ ΤΗΝ ΧΡΗΣΗ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ**

Γεώργιος Ντότσιος

Π2014148

Επιβλέπων Καθηγητής

Θεμιστοκλής Έξαρχος

Εξεταστική Επιτροπή

Κάτια-Λήδα Κερμανίδου

Αριστείδης Βραχάτης

Ευχαριστίες

Η παρούσα πτυχιακή εκπονήθηκε από τον φοιτητή Ντότσιο Γεώργιο του τμήματος πληροφορικής του Ιονίου Πανεπιστημίου, κατά το ακαδημαϊκό έτος 2022-2023 και με επιβλέποντα καθηγητή τον Θεμιστοκλή Έξαρχο. Θα ήθελα να ευχαριστήσω τον κ. Έξαρχο τόσο για την ανάθεση του θέματος, όσο και για την πλήρη εμπιστοσύνη που μου έδειξε σε όλη την διάρκεια εκπόνησης της εργασίας. Επιπλέον θα ήθελα να τον ευχαριστήσω για το ενδιαφέρον που έδειξε αλλά και για όλο τον χρόνο που διέθεσε.

Περιεχόμενα

Περίληψη.....	8
Abstract.....	8
Περιγραφή της καρδιάς.....	9
Ειδικά χαρακτηριστικά.....	9
Χαρακτηριστικά για ενήλικες.....	10
Φυσιολογία του καρδιαγγειακού συστήματος.....	10
Το φλεβοκομβικό τμήμα.....	11
Το κολποκοιλιακό τμήμα.....	11
Το μυοκάρδιο.....	12
Ισχαιμία.....	13
Μελέτες ανά περίπτωση.....	15
Αλγόριθμοι μηχανικής μάθησης που χρησιμοποιούνται στην πρόβλεψη, διάγνωση και θεραπεία καρδιακών παθήσεων.....	22
Decision Tree (DT)-Δέντρο αποφάσεων.....	22
Naïve bayes.....	22
Support Vector Machine (SVM)-Μηχανή διανύσματος υποστήριξης.....	22
K-Nearest Neighbor (K-κοντινότερος γείτονας).....	22
Random Forest (RF)-Τυχαίο δάσος.....	23
WEKA SOFTWARE.....	23
Περιγραφή του dataset.....	24
Κολπική μαρμαρυγή (FIBR_PREDS).....	24
Υπερκοιλιακή ταχυκαρδία (PREDS_TAH).....	24
Κοιλιακή ταχυκαρδία (JELUD_TAH).....	24
Κοιλιακή μαρμαρυγή (FIBR_JELUD).....	24
Αποκλεισμός AV τρίτου βαθμού (A_V_BLOK).....	24
Πνευμονικό οίδημα (OTEK_LANC).....	25
Ρήξη του μυοκαρδίου (RAZRIV).....	25
Σύνδρομο Dressler (DRESSLER).....	25
Χρόνια καρδιακή ανεπάρκεια (ZSN).....	25
Υποτροπή του εμφράγματος του μυοκαρδίου (REC_IM).....	25
Στηθάγχη μετά από έμφραγμα (P_IM_STEN).....	25
Φίλτρα.....	26
ΑΠΟΤΕΛΕΣΜΑΤΑ.....	27
Κολπική μαρμαρυγή (FIBRS_PREDS).....	30
Υπερκοιλιακή ταχυκαρδία (PREDS_TAH).....	34
Κοιλιακή ταχυκαρδία (JELUD_TAH).....	38
Κοιλιακή μαρμαρυγή (FIBR_JELUD).....	42

Αποκλεισμός AV τρίτου βαθμού (A_V_BLOK).....	46
Πνευμονικό οίδημα (OTEK_LANC).....	50
Ρήξη του μυοκαρδίου (RAZRIV).....	55
Σύνδρομο Dressler.....	59
Χρόνια καρδιακή ανεπάρκεια (ZSN).....	63
Υποτροπή του εμφράγματος του μυοκαρδίου (REC_IM).....	67
Στηθάγχη μετά το έμφραγμα (P_IM_STEN).....	71
ΣΥΜΠΕΡΑΣΜΑΤΑ.....	75
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	76

Εικόνες

Εικόνα 1 Χαρακτηριστικά-Μέρη της καρδιάς.....	10
---	----

Πίνακες

Πίνακας 1 Μελέτες σχετικά με την μυοκαρδιακή ισχαιμία και οι αλγόριθμοι μηχανικής μάθησης.....	20
Πίνακας 2 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes).....	30
Πίνακας 3 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes).....	30
Πίνακας 4 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest).....	31
Πίνακας 5 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest).....	31
Πίνακας 6 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48).....	31
Πίνακας 7 Σταθμισμένες μετρήσεις επιδόσεων (J48).....	32
Πίνακας 8 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3).....	32
Πίνακας 9 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3).....	32
Πίνακας 10 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4).....	32
Πίνακας 11 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4).....	33
Πίνακας 12 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5).....	33
Πίνακας 13 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5).....	33
Πίνακας 14 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes).....	34
Πίνακας 15 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes).....	34
Πίνακας 16 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest).....	35
Πίνακας 17 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest).....	35
Πίνακας 18 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48).....	35
Πίνακας 19 Σταθμισμένες μετρήσεις επιδόσεων (J48).....	35
Πίνακας 20 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3).....	36
Πίνακας 21 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3).....	36
Πίνακας 22 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4).....	36
Πίνακας 23 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3).....	37
Πίνακας 24 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5).....	37
Πίνακας 25 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5).....	37
Πίνακας 26 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes).....	38
Πίνακας 27 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes).....	38
Πίνακας 28 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest).....	39
Πίνακας 29 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest).....	39
Πίνακας 30 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48).....	39
Πίνακας 31 Σταθμισμένες μετρήσεις επιδόσεων (J48).....	39
Πίνακας 32 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3).....	40
Πίνακας 33 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3).....	40
Πίνακας 34 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4).....	40
Πίνακας 35 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4).....	41
Πίνακας 36 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5).....	41
Πίνακας 37 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5).....	41
Πίνακας 38 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes).....	42
Πίνακας 39 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes).....	42
Πίνακας 40 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest).....	42
Πίνακας 41 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest).....	43
Πίνακας 42 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48).....	43
Πίνακας 43 Σταθμισμένες μετρήσεις επιδόσεων (J48).....	43
Πίνακας 44 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3).....	44

[illegible]

Πίνακας 98 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes).....	63
Πίνακας 99 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes).....	63
Πίνακας 100 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest).....	63
Πίνακας 101 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest).....	64
Πίνακας 102 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48).....	64
Πίνακας 103 Σταθμισμένες μετρήσεις επιδόσεων (J48).....	64
Πίνακας 104 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3).....	65
Πίνακας 105 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3).....	65
Πίνακας 106 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4).....	65
Πίνακας 107 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4).....	65
Πίνακας 108 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5).....	66
Πίνακας 109 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5).....	66
Πίνακας 110 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes).....	67
Πίνακας 111 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5).....	67
Πίνακας 112 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest).....	67
Πίνακας 113 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest).....	68
Πίνακας 114 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48).....	68
Πίνακας 115 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest).....	68
Πίνακας 116 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3).....	68
Πίνακας 117 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3).....	69
Πίνακας 118 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4).....	69
Πίνακας 119 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4).....	69
Πίνακας 120 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5).....	70
Πίνακας 121 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5).....	70
Πίνακας 122 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5).....	71
Πίνακας 123 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5).....	71
Πίνακας 124 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest).....	72
Πίνακας 125 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest).....	72
Πίνακας 126 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48).....	72
Πίνακας 127 Σταθμισμένες μετρήσεις επιδόσεων (J48).....	72
Πίνακας 128 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3).....	73
Πίνακας 129 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3).....	73
Πίνακας 130 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4).....	73
Πίνακας 131 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4).....	73
Πίνακας 132 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5).....	74
Πίνακας 133 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5).....	74

Περίληψη

Η παρούσα διπλωματική εργασία έχει ως στόχο την πρόβλεψη της ισχαιμίας του μυοκαρδίου μέσω αλγορίθμων μηχανικής μάθησης. Τα χαρακτηριστικά που εξετάστηκαν ήταν κολπική μαρμαρυγή, κοιλιακή και υπερκοιλιακή τραχύτητα, αποκλεισμός AV τρίτου βαθμού, πνευμονικό οίδημα, κοιλιακή μαρμαρυγή, ρήξη του μυοκαρδίου, σύνδρομο Dressler, χρόνια καρδιακή ανεπάρκεια, υποτροπή του μυοκαρδίου, στηληράγχη από έμφραγμα. Το dataset παρείχε η ιστοσελίδα <https://archive.ics.uci.edu/dataset/579/myocardial+infarction> + complications του οποίου έγινε επεξεργασία με την εφαρμογή συγκεκριμένων φίλτρων και έπειτα των αλγορίθμων Naïve Bayes, Random Forest, J48 και KNN (N=3, 4, 5). Τα εξαγόμενα αποτελέσματα έδειξαν πως ο αλγόριθμος Naïve Bayes έχει την καλύτερη επίδοση για όλα τα χαρακτηριστικά. Τέλος, αξίζει να σημειωθεί ότι το Weka αποτελεί ένα χρήσιμο και ευέλικτο εργαλείο σε διάφορους τομείς δίνοντας αξιόπιστα αποτελέσματα.

Abstract

This thesis aims to predict myocardial ischemia through machine learning algorithms. The characteristics examined were atrial fibrillation, ventricular and supraventricular roughness, third degree AV block, pulmonary edema, myocardial rupture, myocardial rupture, myocardial rupture, Dressler's syndrome, chronic heart failure, myocardial recurrence, myocardial infarction stroke. The dataset was provided by <https://archive.ics.uci.edu/dataset/579/myocardial+infarction> + complications which was processed by applying specific filters and then Naïve Bayes, Random Forest, J48 and KNN algorithms (N=3, 4, 5). The extracted results showed that Naïve Bayes algorithm has the best performance for all the characteristics. Finally, it is worth noting that Weka is a useful and versatile tool in various domains giving reliable results.

Περιγραφή της καρδιάς

Η καρδιά βρίσκεται στη μέση της θωρακικής κοιλότητας και έχει ένα κυκλικό σχήμα, παρόμοιο με έναν υποστρόγγυλο κώνο που είναι γείρει στο πλάι. Το άνω μέρος της καρδιάς, γνωστό ως κορυφή, είναι στραμμένο προς τα εμπρός αριστερά και κατωτέρω προς το αριστερό ημιδιάφραγμα. Η βάση της καρδιάς είναι στραμμένη προς τα πίσω δεξιά και πάνω. Προσκολλημένη μπροστά από την καρδιά είναι το στέρνο, ενώ πίσω από αυτή βρίσκεται η θωρακική περιοχή της σπονδυλικής στήλης, και από τις δύο πλευρές της έχουμε τους πνεύμονες. Στο εσωτερικό της καρδιάς, το οποίο είναι ένα κοίλο μυώδες όργανο, το αίμα που προέρχεται από τις φλέβες εισέρχεται και στη συνέχεια ωθείται προς τις αρτηρίες. Αυτή η διαδικασία παίζει έναν ζωτικό ρόλο στη μεταφορά οξυγόνου στο σώμα. Η καρδιά βρίσκεται μεταξύ των δύο πνευμόνων στη θωρακική κοιλότητα. Περιβάλλεται από μια ειδική μεμβράνη που ονομάζεται περικάρδιο, ενώ οι εσωτερικές της κοιλότητες είναι επενδεδυμένες από μια λεπτή μεμβράνη που αποκαλείται ενδοκάρδιο. Ανάμεσα στο περικάρδιο και το ενδοκάρδιο, υπάρχει το παχύτερο τμήμα της καρδιάς, το μυοκάρδιο, το οποίο αποτελείται από ισχυρές μυϊκές ίνες. Η καρδιά είναι ουσιαστικά μια αντλία με τέσσερις κοιλότητες, γνωστές ως δεξιές και αριστερές κοιλότητες. Οι δεξιές κοιλότητες, που περιλαμβάνουν το δεξί κόλπο και τη δεξιά κοιλία, βρίσκονται μπροστά από τις αριστερές κοιλότητες, δηλαδή τον αριστερό κόλπο και την αριστερή κοιλία. Οι κοιλότητες βρίσκονται χαμηλότερα από τους κόλπους. Σε φυσιολογικές συνθήκες, δεν υπάρχει επικοινωνία μεταξύ των δεξιών και αριστερών κοιλοτήτων. Οι κόλποι επικοινωνούν με τις κοιλίες μέσω των κολποκοιλιακών βαλβίδων.

Ειδικά χαρακτηριστικά

Αριστερός κόλπος:

Ο αριστερός κόλπος είναι η πλατφόρμα που δέχεται τις πνευμονικές φλέβες και επικοινωνεί με την αριστερή κοιλία μέσω του αριστερού κολποκοιλιακού στομίου. Βρίσκεται στην πίσω-πάνω θέση της καρδιάς και σε αυτό το σημείο υπάρχει μια βαλβίδα, γνωστή και ως μιτροειδής βαλβίδα.

Αριστερή κοιλία:

Η αριστερή κοιλία έχει ωοειδές σχήμα και μια ομαλή επιφάνεια προς το διάφραγμα. Το αίμα ρέει προς την αορτή από την αριστερή κοιλία, και δέχεται αίμα από τον αριστερό κόλπο μέσω της μιτροειδούς βαλβίδας. Η κοιλία διαθέτει επίσης την αορτική βαλβίδα, η οποία εμποδίζει την ανεπιθύμητη επιστροφή του αίματος από την αορτή πίσω στην κοιλία.

Δεξιός κόλπος:

Ο δεξιός κόλπος είναι υπεύθυνος για την είσοδο της άνω και κάτω κοίλης φλέβας, καθώς και του στεφανιαίου κόλπου. Η άνω κοίλη φλέβα μεταφέρει αίμα από τα άνω άκρα προς την καρδιά, ενώ η κάτω κοίλη φλέβα το μεταφέρει από τα κάτω άκρα. Ο δεξιός κόλπος περιλαμβάνει την τριγλώχινα, μια βαλβίδα που αποτελείται από τρεις φύλλα και ελέγχει τη ροή αίματος από τον κόλπο στην κοιλία, εμποδίζοντας την επιστροφή του αίματος.

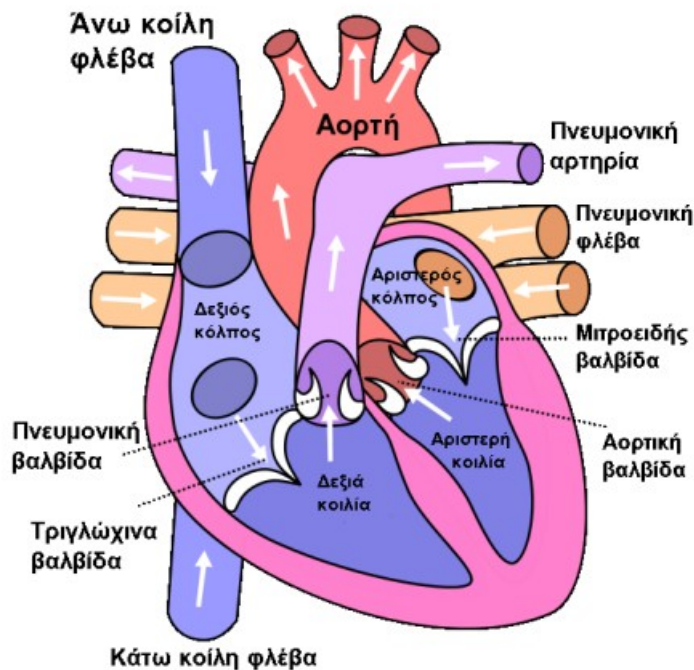
Δεξιά κοιλία:

Η δεξιά κοιλία έχει πυραμιδικό σχήμα και δεν έχει άμεση είσοδο-έξοδο αντίθετα με την αριστερή κοιλία. Το αίμα ρέει προς την πνευμονική αρτηρία από τη δεξιά κοιλία. Εδώ, η τριγλώχινα βαλβίδα διαχειρίζεται τη ροή αίματος από τον κόλπο στην κοιλία και αποτρέπει την επιστροφή του αίματος πίσω στον κόλπο. Το αίμα στη συνέχεια φτάνει στην πνευμονική αρτηρία για να φυλάξει το σώμα με οξυγόνο.

Χαρακτηριστικά για ενήλικες

Σαφώς οι διαστάσεις τις καρδιάς διαφέρουν από άνθρωπο σε άνθρωπο και εξαρτώνται από το φύλο, την ηλικία, την κληρονομικότητα και άλλους γενετικούς παράγοντες.

- Μήκος: 98 χιλιοστά
- Πλάτος: 105 χιλιοστά
- Περιφέρεια: 130 χιλιοστά
- Βάρος: 270 περίπου γραμμάρια



Εικόνα 1 Χαρακτηριστικά-Μέρη της καρδιάς

Φυσιολογία του καρδιαγγειακού συστήματος

Η καρδιά λειτουργεί αυτόματα, λαμβάνοντας παραγόμενα ερεθίσματα από διάφορες περιοχές του μυοκαρδίου. Αυτό σημαίνει ότι δεν χρειάζεται εξωτερικά νευρικά ερεθίσματα για να λειτουργήσει.

Όπως αναφέρθηκε προηγουμένως σχετικά με την ανατομία του μυοκαρδίου, τα κύτταρα του καρδιακού μυός διαθέτουν χαρακτηριστικά που τα καθιστούν αποτελεσματικά στην συστολική λειτουργία της καρδιάς. Για παράδειγμα, αποτελούνται από νήματα ακτίνης και μυοσίνης, τα οποία είναι διατεταγμένα σε συγκύτια. Ένα επιπλέον χαρακτηριστικό που είναι μοναδικό στο μυοκάρδιο, είναι η ικανότητά του να παράγει αυτόματα ηλεκτρικά ερεθίσματα που οδηγούν στη μηχανική συστολή (Κοντόπουλος, 2002, Μπαχαράκης, 2005).

Το φλεβοκομβικό τμήμα

Το τμήμα του φλεβοκόμβου περιλαμβάνει το ίδιο το κόμβο, γνωστό και ως "κόμβος Keith και Flack", και βρίσκεται εντός του μυοκαρδίου του δεξιού κόλπου, κοντά στο σημείο όπου εκβάλλεται η άνω κοίλη φλέβα. Ο φλεβοκόμβος αποτελείται από τρία διαφορετικά είδη κυττάρων, και συγκεκριμένα από κοινά κύτταρα, μεταβατικά κύτταρα και κύτταρα που είναι υπεύθυνα για την δημιουργία ρυθμού στην καρδιά.

Το κολποκοιλιακό τμήμα

Το τμήμα του κολποκοιλιακού συστήματος περιλαμβάνει τον κολποκοιλιακό κόμβο, γνωστό επίσης ως "κόμβος των Aschof-Tawara", που βρίσκεται στο πίσω τμήμα του μεσοκολπικού διαφράγματος κοντά στο σημείο εκβολής του στεφανιαίου κόλπου, καθώς και το δεμάτιο του His και τα δύο σκέλη του, τα οποία αποτελούν την ανατομική και λειτουργική επέκταση του κόμβου Aschof-Tawara. Το δεμάτιο του His διασχίζει τον ινώδη δακτύλιο, διατρέχει το κολποκοιλιακό διάφραγμα και διαιρείται, μετά από μια σύντομη πορεία, στα αριστερά και δεξιά σκέλη.

Είναι σημαντικό να σημειώσουμε ότι οι τελικές αποκλαδώσεις του συστήματος μέσα στο μυοκάρδιο των κοιλιών αποτελούνται από ίνες του Purkinje, οι οποίες δεν έχουν καθορισμένα όρια και τερματίζουν στις ίνες του συσταλτικού μυοκαρδίου (Ramrakha et al., 2008, Rosen et al., 2004).

Στην αρχική φάση, η διέγερση προκύπτει αυτόματα σε ίνες του φλεβοκόμβου, με το φαινόμενο της αναστροφής της ηλεκτρικής διπλοστοιβάδας της κυτταρικής μεμβράνης των μυϊκών ινών, ενεργοποιώντας τη διαδικασία της συστολής στα μικρά ινίδια. Συγκεκριμένα, κατά την αναπαραγωγή του δραστικού ερεθίσματος, προκαλείται προσωρινή αντιστροφή της ηλεκτρικής διπλοστοιβάδας, με αποτέλεσμα η εξωτερική πλευρά της μεμβράνης να αποκτήσει αρνητική φόρτιση σε σύγκριση με την εσωτερική (αντιστροφή) (Κοντόπουλος, 2002, Rosen et al., 2004). Άραγε, ο κόμβος των φλεβών, καθορίζει κατά κύριο λόγο, την συχνότητα παραγωγής των ερεθισμάτων, γι' αυτό και λειτουργεί ως ο φυσιολογικός ρυθμιστής του καρδιακού ρυθμού ή πρωτεύον κέντρο. Ωστόσο, σε ορισμένες περιπτώσεις όπου αυτός αναβληθεί, είναι εφικτή η παραγωγή ερεθισμάτων και από τον κόμβο των Aschof-Tawara, που θεωρείται ο εφεδρικός ρυθμιστής. Όμως ο πρωτεύον ρυθμιστής, παράγει κανονικά περίπου 60-80 ερεθίσματα κάθε λεπτό, ενώ από τον κόμβο Aschof-Tawara προέρχονται 50-55 ερεθίσματα το λεπτό ή ακόμα και λιγότερα (Βαρδής et al., 2005, Μπαχαράκης, 2005).

Με την εκκίνηση της διέγερσης, ξεκινά ο καρδιακός κύκλος ή η καρδιακή περίοδος, η οποία περιλαμβάνει τρεις βασικές φάσεις: τη συστολή των κολπικών κοιλιών, τη συστολή των κοιλιακών κοιλιών και τη διάσταση ή διακοπή της συστολής.

Όταν η διέγερση φτάσει στον κόμβο των κοιλιακών, διαδίδεται στις μυϊκές του ίνες και στη συνέχεια μεταβιβάζεται με ταχύτητα στο δέμα του His και στα δύο του σκέλη, προς τις κλαδικές τους άγκυρες, με τις οποίες μεταδίδεται σε πολλές περιοχές του μυοκαρδίου των κοιλιών, καταλήγοντας σε ένα χρονικό διάστημα περίπου 0,6 δευτερολέπτων. Έπεται μετά από αυτήν τη διέγερση η συντονισμένη συστολή των μυϊκών ινών του συνόλου του κοιλιακού.

Το μυοκάρδιο

Το καρδιακό μυϊκό ιστός αποτελεί ένα συγκεκριμένο τμήμα της καρδιάς, το οποίο διαθέτει την ικανότητα να συστέλλεται, επιτρέποντας την ροή του αίματος μέσω του κυκλοφορικού συστήματος. Το μυοκάρδιο περιλαμβάνει τους δύο κόλπους και τις δύο κοιλίες της καρδιάς, τα οποία αποτελούν το μυοκάρδιο των κόλπων και το μυοκάρδιο των κοιλιών. Αυτά τα δύο μέρη του μυοκαρδίου συνδέονται και αλληλεπιδρούν (Crawford, 2009).

Ο καρδιακός μυς, που είναι γνωστός επίσης και ως "μυοκάρδιο", ανήκει στην κατηγορία των τριών τύπων μυών που υπάρχουν στον ανθρώπινο οργανισμό. Συγκεκριμένα, στο ανθρώπινο σώμα υπάρχουν επίσης οι σκελετικοί και οι λείοι μύες. Ωστόσο, ο καρδιακός μυς είναι μοναδικός και δεν συναντάται ανάλογα σε άλλα σημεία του ανθρώπινου σώματος. Αναλυτικά, μια παρατήρηση κάτω από μικροσκόπιο αποκαλύπτει ότι το μυοκάρδιο αποτελείται από πολλές βαθιές κόκκινες ίνες, γνωστές και ως ινίδια. Κάθε μια από αυτές τις ίνες διασχίζει διακλαδώσεις, που έρχονται σε άμεση επαφή με τις ίνες των υπόλοιπων κυττάρων του καρδιακού μυοκαρδίου. Εξαιτίας αυτής της δομής, το μυοκάρδιο μπορεί να λειτουργεί σαν ένα δίκτυο (Crawford, 2009, Μπαχαράκης, 2005). Χάρη στο ιδιαίτερο ρόλο του, το μυοκάρδιο συχνά αναφέρεται στην αντίστοιχη βιβλιογραφία ως το "συσταλτικό μέρος της καρδιάς". Πραγματικά, κάθε κύτταρο του μυοκαρδίου είναι ικανό να συστολή ή χαλάρωση, αφού διαθέτει ένα είδος ενσωματωμένης ρυθμιζόμενης δύναμης. Σε αντίθεση με άλλους τύπους μυών, αυτά τα κύτταρα δεν είναι άμεσα συνδεδεμένα με το νευρικό σύστημα, αλλά λαμβάνουν ολόκληρη την εντολή για συστολή μέσω ενός αυτόνομου ηλεκτρικού σήματος, το οποίο παράγεται από τον κόμβο των φλεβών και τον κόμβο των κοιλιών. Παρόλα αυτά, τα νεύρα μπορούν να ρυθμίζουν τον ρυθμό και τη δύναμη της συστολής, χωρίς απαραίτητα να επηρεάζουν τον παλμό της καρδιάς.

Ο κόμβος των φλεβών βρίσκεται στο σημείο όπου το δεξί κόλπο συναντά την άνω υποδοχή της φλέβας. Αυτός αποτελείται από μια ομάδα κυττάρων, τα οποία σχηματίζουν τον πρωτογενή παραγωγό ηλεκτρικού σήματος της καρδιάς, γνωστός και ως "βηματοδότης". Η λειτουργία του έχει εξαιρετική σημασία, καθώς ο κάθε καρδιακός παλμός ξεκινά από αυτόν και ανάβει μια σειρά από ηλεκτρικές αντιδράσεις, που μεταφέρουν το σήμα στα δύο κόλπα και ενεργοποιούν τη σύσπασή τους. Έπειτα, το ηλεκτρικό σήμα μεταφέρεται στον κόμβο των κοιλιών. Στη συνέχεια, αυτό το ηλεκτρικό παλμό μεταβιβάζεται από τα κόλπα της καρδιάς σε μια περιοχή, η οποία βρίσκεται στο πίσω κάτω μέρος του μεσοκοιλιακού διαφράγματος. Αυτή η περιοχή είναι γνωστή ως "κόμβος των κοιλιών" και αποτελεί το μοναδικό σημείο σύνδεσης ανάμεσα στα κόλπα και τα κοίλα τμήματα. Συγκεκριμένα, ο κόμβος των κοιλιών λειτουργεί ως ηλεκτρική πύλη ελέγχου μεταξύ του κόμβου των φλεβών και των οδών που μεταφέρουν τα κοίλα τμήματα, εξασφαλίζοντας τη σωστή χρονική μετάδοση του ηλεκτρικού παλμού για την ενεργοποίηση της σύσπασης των κοιλιών (Crawford, 2009, Κρεμαστινός, 2005). Είναι σημαντικό να τονιστεί ότι το μυοκάρδιο αποτελεί μία από τις πιο ευάλωτες περιοχές του οργανισμού. Πιθανές διαταραχές στα κύτταρά του, ανωμαλίες στη ροή και ποιότητα του αίματος, προβλήματα οξυγόνωσης, διατροφικοί και γενετικοί παράγοντες μπορούν να οδηγήσουν σε ποικίλα προβλήματα του μυοκαρδίου, απειλώντας την λειτουργία του και την ανθρώπινη ζωή.

Οι πιο συχνές διαταραχές στη φυσιολογία του μυοκαρδίου είναι:

1. Ισχαιμική μυοκαρδιοπάθεια

2. Μυοκαρδιοπάθεια
3. Διατακτική μυοκαρδιοπάθεια
4. Μη ισχαιμική μυοκαρδιοπάθεια
5. Αποφρακτική μυοκαρδιοπάθεια (Κρεμαστινός 2005)

Ισχαιμία

Η αποκλεισμένη περίφραξη ορίζεται ως η υποβάθμιση ή η διακοπή της ροής του αίματος σε μέρος του σώματος, η οποία οφείλεται σε διαταραχή της αρτηρίας (Μπαμπινιώτης 1998). Ειδικότερα, η αιμοστάση του μυοκαρδίου προκύπτει από αναπηρία της ροής του αίματος στον καρδιακό μυ.

Ο λόγος της ισχαιμίας καρδιακής νόσου είναι πολυπληθής και ποικίλος. Η αθηροσκλήρωση είναι υπεύθυνη για ποσοστό μεγαλύτερο του 90% των περιπτώσεων. Άλλα αίτια που επηρεάζουν μεγάλες αγγειακές περιοχές ενδέχεται να είναι η ρήξη, η εμβολή, πρωτοπαθής διαχωριστικό ανεύρυσμα ή ανώμαλη εκκολπωτική απόκλιση αρτηρίας από την πνευμονική αρτηρία, υπερτροφική μυοκαρδιακή νόσος κ.α. (Τούτουζας, 1991). Εν προκειμένω, η αιτία της νόσου των καρδιακών αγγείων είναι η αθηροσκλήρωση. Κατά τη διάρκεια των ετών, αυτή η ασθένεια προκαλεί στενώσεις ή αποφράξεις των αγγείων της καρδιάς, περιορίζοντας ή αποκόπτοντας την παροχή του αίματος προς τμήματα του μυοκαρδίου, με αποτέλεσμα την κλινική εμφάνιση της νόσου των καρδιακών αγγείων (Τούτουζας, 1991). Η ζωτικότητα του νοσηρού μυοκαρδίου αναγνωρίζεται ως κύριος θεραπευτικός στόχος.

Έχει επιδοθεί ιδιαίτερη προσοχή σε αυτό το ζήτημα σε πειραματικό επίπεδο κατά τη δεκαετία του '70. Διάφορες επεμβάσεις κατά την αποφράξη των καρδιακών αγγείων είχαν ως αποτέλεσμα τη μείωση του καρδιακού επεισοδίου σε ανοικτή καρδιά σκύλων (Reimer et al., 1989). Έως τότε, τα πειραματικά μοντέλα που χρησιμοποιούνταν για την εκτίμηση του βαθμού του καρδιακού επεισοδίου ήταν ακατέργαστα και κανένα από τα ήδη προτεινόμενα μέτρα, όπως οι αναστολείς των β-αδρενεργικών υποδοχέων, δεν φαινόταν αποτελεσματικό. Μικρό χρονικό διάστημα αργότερα, παρατηρείται η επίδραση της επαναιμάτωσης στο ισχαιμικό μυοκάρδιο και σύντομα εισάγονται θεραπευτικές μέθοδοι για τη λύση θρόμβων. Στις αρχές της δεκαετίας του 1980, διερευνώνται αντιφλεγμονώδεις παράγοντες και αντιοξειδωτικοί για την εξέταση αποτελεσμάτων (Reimer et al., 1989). Ενδιαφέρον παρουσιάζει το γεγονός ότι έως το 1986 δεν ήταν γνωστό πόσο εφικτή ήταν η θεραπευτική μείωση της έκτασης του καρδιακού επεισοδίου. Παρότι προτείνονταν ορισμένα φάρμακα για τη μείωση της έκτασης του καρδιακού επεισοδίου, κανένα από τα μελετημένα μέτρα δεν επιβεβαιωνόταν πειραματικά. Η οξεία κατάσταση του καρδιακού επεισοδίου αποτελεί σημαντικό πρόβλημα υγείας παγκοσμίως. Πέρα από την ασθένεια και τη θνητότητα κατά την ακριβή φάση, η απώλεια μυοκαρδιακού ιστού μπορεί να οδηγήσει σε επιδεινούμενη καρδιακή ανεπάρκεια και θανάτωση. Ανάλογα με τα υποκείμενα αίτια που προκαλούν την αιμόσταση και κατ' επέκταση την ύστατη νεκρώση μυοκαρδίου, αλλά επίσης και τα χαρακτηριστικά της αποκάλυψής της, αυτό μπορεί να κατηγοριοποιηθεί σε έξι διακριτούς υποτύπους.

1. Υπότυπος 1

Συμπεριλαμβάνει την ακριβή νέκρωση μυοκαρδίου που συνδέεται με την παρουσία αιμόστασης. Στην πλειοψηφία των περιπτώσεων, η αιμόσταση αυτή αποδίδεται σε κάποιο προηγούμενο στεφανιαίο περιστατικό, όπως είναι η αποπλάκωση πλάκας.

2. Υπότυπος 2

Περιλαμβάνει την νέκρωση μυοκαρδίου που αναπτύσσεται δευτερογενώς από αιμόσταση. Στην περίπτωση αυτή, παρόλα αυτά, η αιμόσταση προκαλείται από την ενίσχυση της ανάγκης του οργανισμού για οξυγόνο και στη συνέχεια στην περιορισμένη παροχή, γεγονός που οδηγεί σε δυσλειτουργικά ιστολόγια, όπως είναι ο σπασμός της στεφανιαίας αρτηρίας, οι ακροδιαταραχές, η υπέρταση ή η υποτασία και η αναιμία.

3. Υπότυπος 3

Περιλαμβάνει την νέκρωση μυοκαρδίου που παρουσιάζει απότομη ληθαργία. Σε αυτή την περίπτωση, η ιδιότυπη ιατρική κατάσταση περιλαμβάνει την αποτυχία της καρδιάς του ασθενούς, ο οποίος εκ των προτέρων παρουσιάζει κατά κανόνα συμπτώματα που υποδεικνύουν μυοκαρδιακή αιμόσταση. Σε πολλές περιπτώσεις, η δυσλειτουργία συμπίπτει με τη νέα ανοδο ST ή νέα LBBB, καθώς επίσης και την εμφάνιση καινούργιου θρόμβου στη στεφανιαία αγγειογραφία ή κατά την νεκροψία.

4. Υπότυπος 4

Αναφέρεται στη νέκρωση μυοκαρδίου, αιμοστάσεως ή μη, που παρουσιάζεται μετά από επέμβαση διαπερατή αγγειοπλαστική (Percutaneous Coronary Intervention- PCI).

5. Υπότυπος 5

Αφορά στην κατηγοριοποίηση που συσχετίζεται στενά με την σωλήνωση. Σε πολλές περιπτώσεις, η παρουσία αίματος στην σωλήνωση αναγνωρίζεται κατά την ανατομοπαθολογική εξέταση, είτε στην νεκροψία, αν και ο ασθενής καταλήξει στον θάνατο.

6. Υπότυπος 6

Παραπέμπει στη νέκρωση μυοκαρδίου που σχετίζεται στενά με την αορτοστεφανιαία ανακατασκευή (Κρεμαστινός, 2005, Μανώλης, 2002, Μπαχαράκης, 2005).

Μελέτες ανά περίπτωση

Η ισχαιμία του μυοκαρδίου (MI) είναι μια ανώμαλη καρδιακή κατάσταση κατά την οποία το μυοκάρδιο στερείται της παροχής αίματος και οξυγόνου. Αποτελεί την πιο διαδεδομένη καρδιακή νόσο σε όλο τον κόσμο και η συχνότητά της αυξάνεται με τους ηλικιακούς παράγοντες. Ο ακριβής και έγκαιρος εντοπισμός και η θεραπεία της MI συμβάλλουν στην πρόληψη μιας σειράς τρομερών ισχαιμικών καρδιακών παθήσεων, όπως η καρδιακή προσβολή και η καρδιακή ανεπάρκεια. Ωστόσο, η χειροκίνητη ερμηνεία του ηλεκτροκαρδιογραφήματος (ΗΚΓ) για τη διάγνωση του MI είναι κουραστικό έργο για τους γιατρούς. Ως εκ τούτου, η αυτοματοποιημένη ταξινόμηση είναι πολύ χρήσιμη για τη διάγνωση της κατάστασης MI.

Στην παρούσα μελέτη, οι Lomsky et al., 2008 αξιολόγησαν το σύστημα σε ένα ξεχωριστό νοσοκομείο από εκείνο στο οποίο εκπαιδεύτηκε και να το συγκρίνουμε με ένα πακέτο λογισμικού ποσοτικοποίησης. Εκπαιδεύτηκε μια πλήρως αυτοματοποιημένη μέθοδος που βασίζεται σε νευρωνικά δίκτυα για την ερμηνεία των (σπινθηρογραφήματος μυοκαρδιακής αιμάτωσης) MPS όσον αφορά την ισχαιμία και το έμφραγμα του μυοκαρδίου, χρησιμοποιώντας 418 MPS από ένα νοσοκομείο. Χαρακτηριστικά από κάθε εξέταση που περιγράφουν την αιμάτωση ηρεμίας και καταπόνησης, την περιφερειακή και τη συνολική λειτουργία χρησιμοποιήθηκαν ως είσοδοι σε διαφορετικά νευρωνικά δίκτυα. Μετά την εκπαίδευση, το σύστημα αξιολογήθηκε χρησιμοποιώντας 532 MPS από άλλο νοσοκομείο. Οι εικόνες της δοκιμής υποβλήθηκαν επίσης σε επεξεργασία με το πακέτο λογισμικού ποσοτικοποίησης Emory Cardiac Toolbox (ECTb). Οι εικόνες ερμηνεύτηκαν από έμπειρους κλινικούς ιατρούς τόσο στο νοσοκομείο εκπαίδευσης όσο και στο νοσοκομείο δοκιμής, όσον αφορά την παρουσία ή την απουσία ισχαιμίας του μυοκαρδίου και/ή εμφράγματος και οι ερμηνείες αυτές χρησιμοποιήθηκαν ως χρυσός κανόνας. το νευρωνικό δίκτυο παρουσίασε ευαισθησία 90% και ειδικότητα 85% για την ισχαιμία του μυοκαρδίου. Η ειδικότητα για το ECTb ήταν 46% ($p < 0,001$), μετρούμενη με την ίδια ευαισθησία. Η ευαισθησία του νευρωνικού δικτύου για το έμφραγμα του μυοκαρδίου ήταν 89% και η ειδικότητα 96%. Η αντίστοιχη ειδικότητα για το ECTb ήταν 54% ($p < 0,001$). ένα σύστημα υποστήριξης αποφάσεων που βασίζεται σε νευρωνικά δίκτυα παρουσιάζει ερμηνείες πιο παρόμοιες με αυτές των έμπειρων κλινικών ιατρών σε σύγκριση με ένα συμβατικό πακέτο λογισμικού αυτοματοποιημένης ποσοτικοποίησης. Η μελέτη αυτή δείχνει τη δυνατότητα διάδοσης της τεχνογνωσίας έμπειρων κλινικών ιατρών σε λιγότερο έμπειρους ιατρούς με τη χρήση νευρωνικών δικτύων.

Οι Juárez-Orozco et al., 2018 διερεύνησαν την αποτελεσματικότητα της μηχανικής μάθησης (ML) στη χρήση εύκολα προσβάσιμων κλινικών και λειτουργικών δεδομένων για τον εντοπισμό ασθενών με ισχαιμία του μυοκαρδίου ή υψηλό κίνδυνο μείζονος ανεπιθύμητου καρδιακού συμβάντος (MACE) με τη χρήση ποσοτικών αποτελεσμάτων PET για το αποθεματικό αιμάτωσης του μυοκαρδίου (MPR). Οι ερευνητές ανέλυσαν δεδομένα από 1.234 ασθενείς που υποβλήθηκαν σε σαρώσεις PET αζώτου-13 αμμωνίας. Χρησιμοποίησαν δημογραφικές, κλινικές και λειτουργικές μεταβλητές για να εκπαιδεύσουν ένα μοντέλο ML που θα προέβλεπε δύο αποτελέσματα: την παρουσία ισχαιμίας του μυοκαρδίου (περιφερειακό MPR $< 2,0$) και αυξημένο κίνδυνο MACE (συνολικό MPR $< 2,0$). Το μοντέλο ML χρησιμοποίησε μια προσέγγιση ενισχυμένου συνόλου με 16 επιλεγμένα χαρακτηριστικά. Το μοντέλο πέτυχε περιοχή κάτω από την

καμπύλη (AUC) 0,72 για τον εντοπισμό ασθενών με μυοκαρδιακή ισχαιμία και AUC 0,71 για τον εντοπισμό ασθενών με αυξημένο κίνδυνο MACE. Η απόδοση του μοντέλου ML ήταν καλύτερη από εκείνη της λογιστικής παλινδρόμησης με τη χρήση παραδοσιακών μεταβλητών του μοντέλου κινδύνου από τις κατευθυντήριες γραμμές της ESC και για τις δύο εκβάσεις. Τα αποτελέσματα υποδηλώνουν ότι το ML είναι μια εφικτή και αποτελεσματική προσέγγιση για τη χρήση απλών κλινικών και λειτουργικών δεδομένων για τον εντοπισμό ασθενών με ισχαιμία του μυοκαρδίου και υψηλό κίνδυνο MACE στην ποσοτική απεικόνιση PET. Συνοπτικά, η μελέτη έδειξε ότι η μηχανική μάθηση μπορεί να χρησιμοποιηθεί για την πρόβλεψη της παρουσίας μυοκαρδιακής ισχαιμίας και του κινδύνου MACE χρησιμοποιώντας άμεσα διαθέσιμες κλινικές και λειτουργικές πληροφορίες από ποσοτικές σαρώσεις PET, υπερβαίνοντας τα παραδοσιακά μοντέλα λογιστικής παλινδρόμησης.

Merjulah et al., 2020 η καθυστερημένη ενίσχυση της μαγνητικής τομογραφίας στον βραχύ άξονα της μυοκαρδιακής ισχαιμίας (MI) εξετάζεται για την ταξινόμηση της MI. Χρησιμοποιώντας το Fuzzy c-means (FCM) όπου είναι μια τεχνική ομαδοποίησης δεδομένων στην οποία ένα σύνολο δεδομένων ομαδοποιείται σε N συστάδες με κάθε σημείο δεδομένων στο σύνολο δεδομένων να ανήκει σε κάθε συστάδα σε ορισμένο βαθμό. Η βασική ιδέα είναι να τμηματοποιηθεί ο ουλώδης ιστός στη μαγνητική τομογραφία του ασθενούς MI η οποία χωρίζει την εικόνα σε σκληρές και ασαφείς τεχνικές ομαδοποίησης. Το διαδικασία εξαγωγής χαρακτηριστικών ολοκληρώνεται μέσω του μορφολογικού φιλτραρίσματος τεχνική. Περαιτέρω, για την ταξινόμηση του MI, τα δεδομένα περνούν από το νευρωνικό δίκτυο τροφοδότησης προς τα εμπρός (FFNN) με τη χρήση Levenberg Marquardt Back διάδοσης (LMBP). Ωστόσο, στον ιατρικό τομέα υπάρχουν ζητήματα όπως η προστασία της ιδιωτικής ζωής, η ασφάλεια και η ανάλυση των δεδομένων. Επιπλέον, υπάρχουν άλλες προκλήσεις, όπως η ανίχνευση, η ταξινόμηση και η διάγνωση της MI. Για την επίλυση αυτών των πολύπλοκων προβλημάτων, η τεχνική της νοημοσύνης μπορεί να βοηθήσει τον γιατρό ως δευτερεύουσα γνώμη. Κίνητρο αποτελεί η ανίχνευση και η ταξινόμηση της MI με τη βοήθεια της προτεινόμενης μεθόδου με τη χρήση μαγνητικής τομογραφίας. Έρευνα μελέτες έχουν αφιερωθεί στην επεξεργασία και την ανάλυση του συνόλου δεδομένων MI για την εξαγωγή σημαντικών πληροφοριών, όπως η ανίχνευση και η ταξινόμηση του καρδιακού MI. Η ταξινόμηση της ισχαιμίας του μυοκαρδίου με βραχυπρόθεσμη ενίσχυση του άξονα στη μαγνητική τομογραφία με τη χρήση μιας τεχνικής μηχανικής μάθησης αποδίδει καλύτερη ακρίβεια από ό,τι άλλους αλγόριθμους ταξινόμησης. Το προτεινόμενο μοντέλο χρησιμοποιεί FCM για την κατάτμηση για τον εντοπισμό της περιοχής ουλής στην ισχαιμία του μυοκαρδίου του ασθενούς και η εξαγωγή χαρακτηριστικών γίνεται με τη βοήθεια της μαγνητικής τομογραφίας με καθυστερημένη ενίσχυση της μορφολογικής διαδικασίας. Τέλος, το μοντέλο εκπαιδεύεται χρησιμοποιώντας το Levenberg Marquardt back propagation στο FFNN το οποίο χρησιμοποιείται για την ταξινόμηση της ισχαιμίας του μυοκαρδίου. Το προτεινόμενο μοντέλο συγκρίνεται με τα άλλα αλγόριθμους ταξινόμησης και τα αποτελέσματα αποκαλύπτουν ότι το προτεινόμενο μοντέλο παράγει μεγαλύτερη ακρίβεια στην ταξινόμηση της ισχαιμίας του μυοκαρδίου. Επομένως, το προτεινόμενο μοντέλο είναι μια από τις πιο βιώσιμες τεχνικές για την ταξινόμηση της MI στη μαγνητική τομογραφία βραχέων αξόνων με καθυστερημένη ενίσχυση.

Η αξονική τομογραφία καρδιάς με τη χρήση μη ενισχυμένης βαθμολόγησης ασβεστίου στεφανιαίων αρτηριών (CACS) και η αξονική αγγειογραφία στεφανιαίων αρτηριών

(cCTA) έχει αποδειχθεί ότι παρέχουν εξαιρετική αξιολόγηση της στεφανιαίας νόσου (ΣΑΑ), συνδυάζοντας ανατομική και μορφολογική αξιολόγηση της ΣΑΑ για τη διαστρωμάτωση του καρδιαγγειακού κινδύνου και τη λήψη θεραπευτικών αποφάσεων, ενώ παράλληλα παρέχουν προγνωστική αξία για την εμφάνιση δυσμενούς καρδιακής έκβασης. Τα τελευταία χρόνια, η τεχνητή νοημοσύνη (AI) και, ειδικότερα, η εφαρμογή αλγορίθμων μηχανικής μάθησης (ML), έχουν προωθηθεί στην καρδιαγγειακή αξονική τομογραφία για τη βελτίωση των μονοπατιών λήψης αποφάσεων, τη διαστρωμάτωση του κινδύνου και την πρόβλεψη της έκβασης με πιο αντικειμενικό, αναπαραγώγιμο και ορθολογικό τρόπο. Η τεχνητή νοημοσύνη βασίζεται στην επιστήμη των υπολογιστών και τα μαθηματικά που βασίζονται σε μεγάλα δεδομένα, υπολογιστικές υποδομές υψηλής απόδοσης και εφαρμοσμένους αλγορίθμους. Η εφαρμογή της TN στην καθημερινή κλινική πρακτική ρουτίνας μπορεί να ενέχει δυνατότητες βελτίωσης της ροής εργασιών απεικόνισης και προώθησης της καλύτερης πρόβλεψης των αποτελεσμάτων και της αποτελεσματικότερης λήψης αποφάσεων στη διαχείριση των ασθενών. Επιπλέον, η αξονική τομογραφία αποτελεί έναν τομέα στον οποίο η ML μπορεί να είναι ιδιαίτερα χρήσιμη, όπως η CACS και η cCTA.

Η μελέτη των Shu et al., 2020 ανέπτυξε και επικύρωσε μια μέθοδο βασισμένη στη μηχανική μάθηση για την πρόβλεψη της χρόνιας ισχαιμίας του μυοκαρδίου χρησιμοποιώντας ένα νομόγραμμα ραδιομικής ακτινολογίας που προκύπτει από τη στεφανιαία αγγειογραφία με υπολογιστική τομογραφία (CCTA). Ανέλυσαν ασθενείς που υποβλήθηκαν τόσο σε απεικόνιση αιμάτωσης του μυοκαρδίου με υπολογιστική τομογραφία εκπομπής ενός φωτονίου (SPECTMPI) όσο και σε σαρώσεις CCTA. Οι ασθενείς χωρίστηκαν σε ομάδες εκπαίδευσης, δοκιμής και επικύρωσης. Χρησιμοποίησαν εικόνες CCTA για να υπολογίσουν πολυάριθμες ραδιομικρές παραμέτρους από το μυοκάρδιο, επιλέγοντας 8 βασικές παραμέτρους για να δημιουργήσουν μια εξατομικευμένη "ραδιομικρή υπογραφή" για κάθε ασθενή. Αυτή η υπογραφή, μαζί με τα κλινικά δεδομένα και τον βαθμό σοβαρότητας της στένωσης, τροφοδοτήθηκε σε ένα μοντέλο μηχανικής μάθησης για τη δημιουργία ενός "νομογράμματος ραδιομικής", το οποίο παρείχε μια εξατομικευμένη βαθμολογία κινδύνου για ισχαιμία του μυοκαρδίου. Η ακρίβεια του νομογράμματος μηχανικής μάθησης στην πρόβλεψη της ισχαιμίας που ανιχνεύθηκε με SPECTMPI ήταν υψηλή: 0,839 για την ομάδα εκπαίδευσης, 0,832 για την ομάδα δοκιμής και 0,816 για την ομάδα επικύρωσης. Η μελέτη κατέδειξε την κλινική χρησιμότητα του νομογράμματος στην πρόβλεψη της ισχαιμίας και στις τρεις κοορτές. Συνολικά, σε ολόκληρο τον πληθυσμό της μελέτης, το ραδιομικροσκοπικό νομόγραμμα υπερείχε τόσο της ραδιομικροσκοπικής υπογραφής όσο και του βαθμού σοβαρότητας της στένωσης για τη διάκριση της ισχαιμίας του μυοκαρδίου, με περιοχή κάτω από την καμπύλη χειριστή δέκτη (AUC) 0,824 για το νομόγραμμα, σε σύγκριση με 0,736 για την ραδιομικροσκοπική υπογραφή και 0,708 για τον βαθμό σοβαρότητας της στένωσης.

Το επεμβατικό αποθεματικό κλασματικής ροής (FFR) είναι ένα τυπικό εργαλείο για τον εντοπισμό στεφανιαίων στενώσεων που προκαλούν ισχαιμία. Ωστόσο, στην κλινική πρακτική, πάνω από το 70% των αποφάσεων θεραπείας εξακολουθούν να βασίζονται στην οπτική εκτίμηση της αγγειογραφικής στένωσης, η οποία έχει περιορισμένη ακρίβεια (περίπου 60%- 65%) για την πρόβλεψη $FFR < 0,80$. Ένας από τους λόγους για την οπτική-λειτουργική αναντιστοιχία είναι ότι η ισχαιμία του μυοκαρδίου μπορεί να επηρεάζεται από το μέγεθος του παρεχόμενου μυοκαρδίου, το οποίο δεν είναι πάντα εμφανές από τη στεφανιογραφία. Οι στόχοι της μελέτης των Hae et al., 2018 ήταν να

αναπτυχθεί ένας αλγόριθμος μηχανικής μάθησης (ML) με βάση την αγγειογραφία για την πρόβλεψη του παρεχόμενου όγκου του μυοκαρδίου για μια στένωση, όπως μετράται με τη χρήση στεφανιαίας αγγειογραφίας με υπολογιστική τομογραφία (CCTA), και στη συνέχεια να δημιουργηθεί ένας ταξινομητής με βάση την αγγειογραφία για τις βλάβες με $FFR < 0,80$ έναντι $\geq 0,80$. Διαπιστώσαμε ότι η ML με βάση την αγγειογραφία είναι χρήσιμη για την πρόβλεψη των υποτελών περιοχών του μυοκαρδίου και των βλαβών που προκαλούν ισχαιμία, αμβλύνοντας την οπτικο-λειτουργική αναντιστοιχία μεταξύ αγγειογραφικού και FFR . Η αξιολόγηση της κλινικής χρησιμότητας απαιτεί περαιτέρω επικύρωση σε μια μεγάλη, προοπτική μελέτη κοόρτης.

Οι Yang et al., 2020 ανέπτυξαν ένα μοντέλο μηχανικής μάθησης (ML) με χρήση εικόνων στεφανιαίας αγγειογραφίας με υπολογιστική τομογραφία (CTA) για την αξιολόγηση της ισχαιμίας του μυοκαρδίου σε ασθενείς με στεφανιαία αθηροσκλήρωση. Στην αναδρομική ανάλυση συμμετείχαν 110 ασθενείς, εκ των οποίων 58 είχαν μυοκαρδιακή ισχαιμία και 52 είχαν φυσιολογική αιμάτωση του μυοκαρδίου. Οι ασθενείς χωρίστηκαν σε σύνολα δεδομένων εκπαίδευσης και δοκιμής (αναλογία 7:3). Αυτοματοποιημένο λογισμικό τμηματοποίησε το μυοκάρδιο σε εικόνες CTA και εξήγαγε χαρακτηριστικά υψής χρησιμοποιώντας μια προσέγγιση που βασίζεται σε μοντέλο βαθιάς μάθησης και ονομάζεται CQK. Δημιουργήθηκαν επτά μοντέλα ML για την ταξινόμηση των περιπτώσεων είτε ως ισχαιμία του μυοκαρδίου είτε ως φυσιολογική παροχή αίματος. Οι επιδόσεις και η σταθερότητα των μοντέλων αξιολογήθηκαν με τη χρήση καμπυλών λειτουργικών χαρακτηριστικών δέκτη (ROC) με διασταυρούμενη επικύρωση. Το βέλτιστο μοντέλο επικυρώθηκε στη συνέχεια με ένα ανεξάρτητο σύνολο δεδομένων δοκιμής. Η μηχανή διανυσμάτων υποστήριξης με τη γραμμική μέθοδο extreme gradient boosting πέτυχε ακρίβεια 0,821 και περιοχή κάτω από την καμπύλη ROC (AUC) 0,777. Η μέθοδος νευρωνικού δικτύου (NN) πέτυχε ακρίβεια 0,818 και AUC 0,757. Το μοντέλο naive Bayes είχε την υψηλότερη ευαισθησία (0,942) και το μοντέλο random forest είχε την υψηλότερη ειδικότητα (0,85). Το μοντέλο k-nearest (κοντινότερων γειτόνων) είχε τη χαμηλότερη ακρίβεια (0,74). Το μοντέλο NN παρουσίασε τις χαμηλότερες σχετικές τυπικές αποκλίσεις (0,16 για την ακρίβεια και 0,08 για την AUC), υποδεικνύοντας υψηλή σταθερότητα. Η AUC του στο ανεξάρτητο σύνολο δεδομένων δοκιμής ήταν 0,72. Συμπερασματικά, το μοντέλο νευρωνικού δικτύου (NN) που χρησιμοποιεί ακτινολογικά χαρακτηριστικά από εικόνες CTA έδειξε την καλύτερη απόδοση στην πρόβλεψη της ισχαιμίας του μυοκαρδίου. Αυτό υποδηλώνει ότι το μοντέλο MN έχει πολλά υποσχόμενες δυνατότητες για την υποβοήθηση της λήψης κλινικών αποφάσεων για ασθενείς με στεφανιαία αθηροσκλήρωση.

Χρησιμοποιώντας έναν αυτοματοποιημένο αλγόριθμο (AUTOPLAQ) που βασίζεται στην αυτοματοποιημένη κατάτμηση και ταξινόμηση των στεφανιαίων για την ογκομετρική ποσοτικοποίηση της πλάκας, οι Dey et al., (2018) πρόσθεσαν χαρακτηριστικά της εικόνας σε μια προσέγγιση ML με αυτοματοποιημένο χαρακτηριστικό επιλογή και κατάταξη του κέρδους πληροφορίας για την πρόγνωση της ισχαιμίας με συγκεκριμένη βλάβη. Η βαθμολογία ML είχε ως αποτέλεσμα AUC 0,84, η οποία ήταν σημαντικά υψηλότερη σε σύγκριση με την σε σύγκριση με όλες τις μεμονωμένες μετρήσεις CT (AUC 0,63-0,76, $p < 0.05$). Σε μια πρόσφατη μελέτη, οι Xiong et al., (2015) εφάρμοσαν τρεις διαφορετικές προσεγγίσεις ML (Random Forest, Ada Boost, Naive Bayes) για την αυτοματοποιημένη κατάτμηση και οριοθέτηση της

αριστερής κοιλίας χρησιμοποιώντας τρία διαφορετικά μυοκαρδιακά χαρακτηριστικά που λαμβάνονται από εικόνες CTP αιμάτωσης του μυοκαρδίου με αξονική τομογραφία ηρεμίας, όπως η κανονικοποιημένη ένταση αιμάτωσης, ο διαδερμικός λόγος αιμάτωσης και το πάχος του τοιχώματος του μυοκαρδίου. Το διαπίστωσαν ότι ο αλγόριθμος Ada Boost είχε την καλύτερη απόδοση σε σύγκριση με τη χειροκίνητη τμηματοποίηση από έναν ειδικό αναγνώστη με AUC 0,73. Σύμφωνα με αυτή τη διερεύνηση, οι Han et al., (2018) χρησιμοποίησαν έναν ταξινομητή gradient boosting για επιβλεπόμενο ML για την πρόβλεψη φυσιολογικά σημαντικού CAD σε ηρεμία CTP του μυοκαρδίου σε ένα σύνολο δεδομένων 252 ασθενών. Οι συγγραφείς ανέφεραν διαγνωστική ακρίβεια, ευαισθησία και εξειδίκευση 68%, 53% και 85% της CTP που προστέθηκε στην αγγειογραφική στένωση cCTA>70%. για την πρόβλεψη της ισχαιμίας. Η προσθήκη της CTP ηρεμίας στην αξιολόγηση με cCTA οδήγησε σε αυξημένη AUC 0,75 έναντι 0,68 για αιμοδυναμικά σημαντική στεφανιαία νόσο CAD. Συνολικά, η εφαρμογή του ML στην CTP έχει ένα πιθανό μέλλον, ιδίως στην αυτοματοποιημένη κατάτμηση του μυοκαρδίου και στην αυτοματοποιημένη ανίχνευση των ελαττωμάτων αιμάτωσης τόσο στη στατική όσο και στη δυναμική CTP, ώστε να μειωθεί περαιτέρω η ανθρώπινη εισροή για τη μετεπεξεργασία της εικόνας. Επιπλέον, η χρήση της CTP σε μοντέλα κινδύνου που βασίζονται σε ML για την πρόβλεψη της ισχαιμίας θα έχει πιθανότατα ως αποτέλεσμα τη βελτίωση της διαγνωστικής ακρίβειας σύμφωνα με την CT-FFR. Η παρακάτω μελέτη επικεντρώνεται στην αντιμετώπιση της πρόκλησης της ανίχνευσης της ισχαιμίας του μυοκαρδίου, ιδίως σε περιπτώσεις με ανεπαίσθητες αλλαγές στο τυπικό ηλεκτροκαρδιογράφημα 12 απαγωγών (ΗΚΓ), γνωστό ως μη διαγνωστικό ΗΚΓ. Οι υπάρχουσες μέθοδοι που βασίζονται στις παραδοσιακές μεταβολές των κυμάτων ST και T στερούνται ακρίβειας. Οι Sun et al., (2022) εισάγει μια νέα προσέγγιση που συνδυάζει τη μάθηση συνόλου και τη ντετερμινιστική μάθηση για βελτιωμένη ανίχνευση της μυοκαρδιακής ισχαιμίας. Οι ερευνητές ανέπτυξαν μια μέθοδο που ονομάζεται καρδιοδυναμικό πρόγραμμα (CDG) χρησιμοποιώντας δυναμική μοντελοποίηση των σημάτων ΗΚΓ μέσω ντετερμινιστικής μάθησης. Εξήγαγαν συγκεκριμένα χαρακτηριστικά από το CDG, συμπεριλαμβανομένου του εκθέτη φασματικής προσαρμογής, του εκθέτη Lyapunov και της πολυπλοκότητας Lempel-Ziv. Στη συνέχεια, χρησιμοποίησαν έναν αλγόριθμο ετερογενών συνόλων με βάση το bagging, ο οποίος περιελάμβανε τη δημιουργία διαφορετικών βασικών ταξινομητών και τη συγκέντρωση των αποτελεσμάτων τους με τη χρήση σταθμισμένης ψηφοφορίας για να σχηματίσουν ένα μοντέλο συνόλου για την ανίχνευση της ισχαιμίας του μυοκαρδίου. Η μέθοδος εκπαιδεύτηκε και δοκιμάστηκε σε ένα κλινικό σύνολο δεδομένων που αποτελούνταν από 499 μη διαγνωστικά αρχεία ΗΚΓ 12 απαγωγών από 499 ασθενείς που συλλέχθηκαν σε τρία ανεξάρτητα ιατρικά κέντρα. Από αυτούς τους ασθενείς, 383 είχαν μυοκαρδιακή ισχαιμία και 116 δεν είχαν. Μέσω της διασταυρούμενης επικύρωσης, η προτεινόμενη προσέγγιση πέτυχε μέση ακρίβεια 89,10%, ευαισθησία 91,72% και ειδικότητα 82,69%. Η ακρίβεια του μοντέλου συνόλου ήταν πάνω από 82% για ασθενείς με μη διαγνωστικό ΗΚΓ σε όλα τα τρία ιατρικά κέντρα. Όταν δοκιμάστηκε σε ένα εξωτερικό σύνολο δεδομένων (σύνολο δεδομένων PTB), το μοντέλο επέδειξε ελπιδοφόρα αποτελέσματα με ακρίβεια 91,11%, ευαισθησία 90,49% και ειδικότητα 92,88%. Συμπερασματικά, η προσέγγιση της μελέτης που συνδυάζει τη μάθηση συνόλου και τη ντετερμινιστική μάθηση έδειξε υψηλή διαγνωστική ακρίβεια και γενίκευση στην κλινική πρακτική. Έχει τη δυνατότητα να συμπληρώσει τις συνήθεις μεθόδους ΗΚΓ για την κλινική διάγνωση της ισχαιμίας του μυοκαρδίου. Οι ερευνητές Kotu et al., (2015) πρότειναν ποσοτικά χαρακτηριστικά που εξάγονται από εικόνες

καρδιακού μαγνητικού συντονισμού ενισχυμένες με καθυστερημένο γαδολίνιο σε ασθενείς μετά από MI. Τα χαρακτηριστικά αυτά περιλάμβαναν το μέγεθος της ουλής, τη θέση και πληροφορίες υψής από το ουλωμένο μυοκάρδιο. Χρησιμοποίησαν διάφορους αλγορίθμους ταξινόμησης, συμπεριλαμβανομένων των αλγορίθμων k-κοντινότερου γείτονα (k-NN), μηχανή διανυσμάτων υποστήριξης (SVM), δέντρο απόφασης (DT) και Random Forest (RF)-τυχαίο δάσος, για να αξιολογήσουν την ικανότητα των χαρακτηριστικών να διακρίνουν μεταξύ των ομάδων υψηλού κινδύνου (20 ασθενείς) και χαμηλού κινδύνου (34 ασθενείς). Στο πείραμα 1 με βάση το MATLAB (leave-one-out cross-validation), ένας συνδυασμός χαρακτηριστικών που καταγράφει το μέγεθος, τη θέση και την ετερογένεια της ουλής πέτυχε ακρίβεια ταξινόμησης 94,44% και AUC 0,965. Το πείραμα 2, με χρήση του WEKA, χρησιμοποίησε εμφωλευμένη διασταυρούμενη επικύρωση με ταξινομητές k-NN, SVM, decision tree, random forest. Ο SVM απέδωσε μέση ακρίβεια 92,6% και AUC 0,921 για ένα συνδυασμό χαρακτηριστικών που αποτυπώνει τη θέση της ουλής και την ετερογένεια. Και τα δύο πειράματα έδειξαν ότι τα χαρακτηριστικά υψής από την ουλή ήταν σημαντικά για την ταξινόμηση και η προσθήκη χαρακτηριστικών εντοπισμού βελτίωσε την ακρίβεια. Τα προτεινόμενα διακριτικά χαρακτηριστικά, σε συνδυασμό με τους υπάρχοντες δείκτες κινδύνου, θα μπορούσαν να ενισχύσουν τη διαστρωμάτωση του αρρυθμικού κινδύνου σε ασθενείς μετά από MI. Τα χαρακτηριστικά αυτά έχουν πιθανές εφαρμογές στην κλινική λήψη αποφάσεων ή σε συστήματα αυτόματης υποστήριξης αποφάσεων, προσφέροντας πολλά υποσχόμενα αποτελέσματα για την καλύτερη εκτίμηση του κινδύνου. Στόχος της μελέτης των Mannil et al., (2019) ήταν η αξιολόγηση της ακρίβειας της ανάλυσης υψής (TA) στην ανίχνευση του εμφράγματος του μυοκαρδίου (MI) σε εικόνες καρδιακής υπολογιστικής τομογραφίας (CT) και η σύγκρισή της με την υποκειμενική οπτική διάγνωση. Δέκα ασθενείς με χρόνιο MI και 20 άτομα ελέγχου υποβλήθηκαν σε αξονικές τομογραφίες με διαφορετικές μεθόδους ανακατασκευής. Η υποκειμενική οπτική ανάλυση παρουσίασε καλή συμφωνία μεταξύ των αναγνωστών, αλλά φτωχή συμφωνία μεταξύ των αναγνωστών, με ποικίλα αποτελέσματα ανάλογα με την ανακατασκευή. Η TA επέδειξε σταθερά υψηλή ακρίβεια (94%-97%) και AUC (0,94-0,99) σε όλες τις ανακατασκευές. Όταν χωρίστηκε σε σύνολα εκπαίδευσης/δοκιμής, η συνολική απόδοση μειώθηκε, με τα καλύτερα αποτελέσματα με επαναληπτική ανακατασκευή στο επίπεδο 5 (σωστές ταξινομήσεις: 73%, AUC: 0,65). Η TA προσέφερε αντικειμενική και αναπαραγώγιμη διάγνωση του MI σε σύγκριση με την υποκειμενική οπτική ανάλυση, και η επίδραση της επαναληπτικής ανακατασκευής ήταν αξιοσημείωτη και στις δύο προσεγγίσεις.

Πίνακας 1 Μελέτες σχετικά με την μυοκαρδιακή ισχαιμία και οι αλγόριθμοι μηχανικής μάθησης.

Lomsky et al 2008	Merjula et al., 2020	Orozoco et al., 2018	Shu et al., 2020	Hae 2018	Yang 2022	Dey et al. 2018	Xiong et al., 2015	Han et al., 2018	Sun et al., 2022	Kotula et al., 2015	Mannil et al., 2019
NN	FCM	Severali ML models	(SVM), (KNN), RF	Severali ML models	NN, Naïve Bayes	AUTOPLAQ	RF, Ada Boost,	Gradient Boosting	CDG	k-NN, SVM, DT,	TA

					s, KNN, RF		Naïv e Baye s			RF	
--	--	--	--	--	------------------	--	------------------------	--	--	----	--

Αλγόριθμοι μηχανικής μάθησης που χρησιμοποιούνται στην πρόβλεψη, διάγνωση και θεραπεία καρδιακών παθήσεων

Decision Tree (DT)-Δέντρο αποφάσεων

Το δέντρο αποφάσεων (DT) είναι ένας αλγόριθμος μάθησης με επίβλεψη, που χρησιμοποιείται κυρίως για το πρόβλημα της ταξινόμησης. Μπορεί να χειριστεί τόσο κατηγορικά όσο και αριθμητικά δεδομένα. Είναι μια δομή που μοιάζει με δέντρο και αποτελείται από εσωτερικούς κόμβους, κόμβους φύλλων και κλάδους. Οι πιο συχνά χρησιμοποιούμενοι τύποι δέντρων απόφασης περιλαμβάνουν τα CART, C4.5, ID3, κ.λπ. Αρχικά υπολογίζει την εντροπία για κάθε χαρακτηριστικό/χαρακτηριστικό- στη συνέχεια διαιρεί το σύνολο δεδομένων σε δύο ή περισσότερα παρόμοια σύνολα χρησιμοποιώντας μεταβλητές ή τους προγνωστικούς παράγοντες (Ramalingam et al., 2018, Beyene et al., 2018).

Naïve bayes

Η τεχνική Naïve Bayes (NB) είναι μια αποτελεσματική τεχνική ταξινόμησης και βασίζεται στο θεώρημα του Bayes. Ο Naïve Bayes υποθέτει ότι υπάρχει ανεξαρτησία μεταξύ των προβλεπτικών παραγόντων- η εμφάνιση συγκεκριμένων χαρακτηριστικών σε μια κλάση είναι ανεξάρτητη από άλλα χαρακτηριστικά/χαρακτηριστικά. Ακόμη και αν υπάρχει εξάρτηση, τα χαρακτηριστικά θα συμβάλλουν ανεξάρτητα στην πιθανότητα. Χρησιμοποιείται για τον υπολογισμό της εκ των υστέρων πιθανότητας για κάθε κλάση με βάση την υπό συνθήκη πιθανότητα (Ramalingam et al., 2018, Beyene et al., 2018).

Support Vector Machine (SVM)-Μηχανή διανύσματος υποστήριξης

Είναι μια τεχνική μηχανικής μάθησης με επίβλεψη που χρησιμοποιείται για προβλήματα ταξινόμησης. Αντιμετωπίζει τόσο γραμμικά όσο και μη γραμμικά σύνολα δεδομένων. Για να γίνει η ταξινόμηση, η τεχνική της μηχανής διανυσμάτων υποστήριξης βρίσκει ένα υπερεπίπεδο και μεγιστοποιεί το περιθώριο για τη διάκριση μεταξύ των κλάσεων. Η μηχανή διανυσμάτων υποστήριξης εμφανίζει τα σημεία εκπαίδευσης στο χώρο χαρακτηριστικών ως σημεία έτσι ώστε τα σημεία που ανήκουν στις διαφορετικές κλάσεις να διαχωρίζονται από ένα όσο το δυνατόν μεγαλύτερο περιθώριο. Η SVM μπορεί να διαδραματίσει μεγάλο ρόλο όσον αφορά την πρόβλεψη καρδιακών παθήσεων (Ramalingam et al., 2018, Beyene et al., 2018).

K-Nearest Neighbor (K-κοντινότερος γείτονας)

Ο αλγόριθμος K-κοντινότερος γείτονας (KNN) είναι ένας πολύ απλός και εύκολος στην εφαρμογή αλγόριθμος μηχανικής μάθησης. Χρησιμοποιείται για την επίλυση προβλημάτων παλινδρόμησης και ταξινόμησης όπου δεν υπάρχει καμία ή πολύ λιγότερη γνώση σχετικά με την κατανομή των δεδομένων. Βρίσκει τα "k" πλησιέστερα σημεία δεδομένων στο σύνολο δεδομένων για τα οποία πρέπει να βρεθεί η τιμή-στόχος. Στη συνέχεια, αναθέτει το μέσο όρο αυτών των "k" σημείων δεδομένων στο συγκεκριμένο σημείο δεδομένων (Ramalingam et al., 2018, Beyene et al., 2018).

Random Forest (RF)-Τυχαίο δάσος

Τα τυχαία δάση (RF) είναι εποπτευόμενες μέθοδοι μηχανικής μάθησης που χρησιμοποιούνται για ταξινόμηση, παλινδρόμηση και άλλα προβλήματα μηχανικής μάθησης. Κατά τη διάρκεια της εκπαίδευσης, κατασκευάζουν πολλαπλά δέντρα αποφάσεων πριν δώσουν μια έξοδο. Στην περίπτωση της ταξινόμησης, χρησιμοποιείται σύστημα ψηφοφορίας για την απόφαση της κλάσης, ενώ στην παλινδρόμηση λαμβάνεται ο μέσος όρος ή ο μέσος όρος των εξόδων καθενός από τα μεμονωμένα δέντρα. Τα τυχαία δάση βασίζονται στην πεποίθηση ότι όσο μεγαλύτερος είναι ο αριθμός των δέντρων απόφασης, τόσο μεγαλύτερη είναι η πιθανότητα να ληφθεί η ακριβής έξοδος (Ramalingam et al., 2018, Beyene et al., 2018).

WEKA SOFTWARE

Το Weka περιλαμβάνει ένα σύνολο εργαλείων που χρησιμοποιούνται για την ανάλυση δεδομένων και την πρόβλεψη, καθώς και γραφικά περιβάλλοντα χρήστη που διευκολύνουν την πρόσβαση σε αυτά τα εργαλεία (Witten, Frank & Hall 2011). Η αρχική έκδοση του Weka ήταν ένα περιβάλλον Tcl/Tk για αλγορίθμους μοντελοποίησης που εφαρμόστηκαν σε άλλες γλώσσες προγραμματισμού, μαζί με βοηθητικά προγράμματα προεπεξεργασίας δεδομένων σε γλώσσα C και ένα σύστημα δημιουργίας Makefile για λειτουργίες μηχανικής μάθησης. Αρχικά σχεδιάστηκε κυρίως για την ανάλυση δεδομένων από τον αγροτικό τομέα, αλλά η πιο πρόσφατη έκδοση που χρησιμοποιεί τη γλώσσα Java εφαρμόζεται σε διάφορους τομείς, κυρίως στον τομέα της εκπαίδευσης. Παρακάτω παρουσιάζονται τα πλεονεκτήματα του Weka:

- Δωρεάν διαθεσιμότητα υπό την άδεια GNU
- Φορητότητα
- Πλήρης συλλογή τεχνικών προεπεξεργασίας και μοντελοποίησης δεδομένων
- Ευκολία χρήσης δεδομένων

Το Weka υποστηρίζει πληθώρα κλασσικών καθηκόντων στην εξόρυξη δεδομένων. Συγκεκριμένα, προσφέρει δυνατότητες προεπεξεργασίας δεδομένων, ομαδοποίησης, κατηγοριοποίησης, πρόβλεψης, απεικόνισης, και επιλογής χαρακτηριστικών. Σε όλες αυτές τις τεχνικές του Weka, η βάση τους είναι η υπόθεση ότι τα δεδομένα παρέχονται σε μορφή ενός ενιαίου αρχείου, όπου κάθε παρατήρηση περιγράφεται με ένα σταθερό αριθμό χαρακτηριστικών. Αυτά τα χαρακτηριστικά μπορεί να είναι κλασσικά αριθμητικά ή κατηγορικά, αλλά επιτρέπεται επίσης η χρήση και άλλων τύπων χαρακτηριστικών.

Περιγραφή του dataset

Κολπική μαρμαρυγή (FIBR_PREDS)

Η αρρυθμική (κολπική) μαρμαρυγή είναι μια κατάσταση κατά την οποία εμφανίζονται μη-φυσιολογικά ηλεκτρικά κυκλώματα, κυρίως εντός των δύο αιτίων της καρδιάς, εξαναγκάζοντας την καρδιά να παλμοκροτεί ανεξέλεγκτα και με ταχεία ρυθμότητα. Οι αιτίες και οι κοιλίες συστέλλονται ανορθόδοξα, με αποτέλεσμα να εμφανίζεται μια ακανόνιστη παλμοκρότηση της καρδιάς.

Υπερκοιλιακή ταχυκαρδία (PREDS_TAH)

Η υπερκοιλιακή ταχυκαρδία αποτελεί μία μορφή αρρυθμίας που προκαλείται από δυσλειτουργία των ηλεκτρικών μηχανισμών που ελέγχουν τον καρδιακό ρυθμό. Κατά τη διάρκεια των επεισοδίων, τα οποία μπορεί να διαρκούν μέχρι αρκετές ώρες, η καρδιά παλλόμενη ταχύτερα από το συνηθισμένο, ωστόσο, το κάνει με φυσιολογικό ρυθμό.\

Κοιλιακή ταχυκαρδία (JELUD_TAH)

Οι κοιλιακές ταχυκαρδίες είναι ανωμαλίες που πηγάζουν από τις κάτω κοιλότητες της καρδιάς, που ονομάζονται κοιλίες. Υπάρχουν δύο κατηγορίες κοιλιακών ταχυκαρδιών ανάλογα με την παρουσία ή την απουσία δομικών καρδιακών παθήσεων.

Κοιλιακή μαρμαρυγή (FIBR_JELUD)

Η κοιλιακή αρρυθμία εμφανίζεται όταν διακόπτεται ο καρδιακός ρυθμός, όπως κατά τη διάρκεια καρδιακής αποφρακτικής κρίσης. Αυτό επιπλέον προκαλεί μείωση των επιπέδων της αρτηριακής πίεσης, διακόπτοντας την κυκλοφορία αίματος σε ζωτικά όργανα.

Αποκλεισμός AV τρίτου βαθμού (A_V_BLOK)

Κολποκοιλιακός αποκλεισμός (Atrioventricular Block- AV block) ονομάζεται η μη διέλευση ηλεκτρικού ρεύματος από τους κόλπους προς τις κοιλίες είτε μερικές φορές (2ου βαθμού) είτε μόνιμα (3ου βαθμού).

Πνευμονικό οίδημα (OTEK_LANC)

Το πνευμονικό οίδημα είναι η συσσώρευση υγρών στον ιστό και στις κυψελίδες των πνευμόνων.

Ρήξη του μυοκαρδίου (RAZRIV)

Το απότομο καρδιακό επεισόδιο προκαλείται από την ξαφνική απόφραξη μιας αρτηρίας στεφανιαίας παροχής. Η απόφραξη, αντιστοίχως, προκαλείται από το γρήγορο δημιουργικό διάβρωμα επί μιας αθηροσκληρωτικής πλάκας που υπάρχει προϋπάρχουσα. Η θρομβογενής διαδικασία ενεργοποιείται από την τυχαία ρήξη (διάλυση) του εξωκυττάριου χιτώνα της αθηροσκληρωτικής πλάκας, προκαλώντας την ενεργοποίηση των αιμοπεταλίων και οδηγώντας σύντομα στον σχηματισμό εμπόδιου θρόμβου. Η διακοπή της αρτηριακής ροής και η έλλειψη, συνεπώς, της παροχής οξυγόνου στο τμήμα του μυοκαρδίου που τροφοδοτείται από την προσαχώρηση της αρτηρίας, οδηγεί σε διάστημα περίπου 2-3 ωρών στην εκδημιουργία της καρδιακής νέκρωσης. Επομένως, αυτό το μέρος του μυοκαρδίου χάνει την ικανότητά του να συστέλλεται και δεν συμβάλλει πλέον στην λειτουργία της καρδιάς, μετατρέπεται σε μια ουλή που διατηρεί μόνο την ανατομική ακεραιότητα της αριστερής κοιλίας ως κοιλότητας. Σε εξαιρετικές περιπτώσεις, το καρδιακό επεισόδιο μπορεί να προκαλείται και από την εμφάνιση έτοιμου θρομβογόνου υλικού σε μια αρτηρία του καρδιακού μυός.

Σύνδρομο Dressler (DRESSLER)

Αποτελεί μια φλεγμονώδη αντίδραση που επηρεάζει τους μεμβράνες που περιβάλλουν την καρδιά και τους πνεύμονες, δηλαδή το περίτερμο και τον πλεύρηκα. Συμβαίνει κυρίως σε ασθενείς που υποβάλλονται σε εγχειρήσεις της καρδιάς.

Χρόνια καρδιακή ανεπάρκεια (ZSN)

Η καρδιακή ανεπάρκεια αποτελεί μία νόσο κατά την οποία η καρδιά δεν είναι σε θέση να παράγει επαρκή ποσότητα αίματος για τα υπόλοιπα όργανα του οργανισμού και να διασφαλίσει την απαραίτητη οξυγονώση τους.

Υποτροπή του εμφράγματος του μυοκαρδίου (REC_IM)

Αυτό οφείλεται είτε σε υποκλινόμενη ισχαιμία είτε σε περικαρδίτιδα. Οι υποκλινόμενοι καρδιακοί πόνοι δεν είναι σπάνιοι κατά την πρώτη εικοσαετία και υποδεικνύουν νέο καρδιακό επεισόδιο εάν ακολουθούνται από ηλεκτροκαρδιογραφικές διαταραχές επί των ασθενών ήδη περιοχών και ειδικά αν υπάρχει νέα αύξηση των ενζύμων.

Στηθάγχη μετά από έμφραγμα (P_IM_STEN)

Η στηθάγχη εμφανίζεται όταν η καρδιά αναζητά επιπλέον ποσότητα αίματος πλούσια σε οξυγόνο από αυτήν που μπορούν να της παρέχουν οι αρτηρίες. Συνήθως εμφανίζεται μετά από φυσική άσκηση, σε στιγμές υψηλής άγχους ή μετά από ένα πλούσιο γεύμα.

Φίλτρα

Σε αρχικό στάδιο στο dataset εφαρμόστηκαν δύο φίλτρα και εν συνεχεία οι αλγόριθμοι μηχανικής μάθησης. Αυτά είναι τα παρακάτω:

- **SpreadSubSample:** Παράγει ένα τυχαίο υπόδειγμα ενός συνόλου δεδομένων. Το αρχικό σύνολο δεδομένων πρέπει να χωράει εξ ολοκλήρου στη μνήμη. Αυτό το φίλτρο σας επιτρέπει να καθορίσετε τη μέγιστη "διασπορά" μεταξύ της σπανιότερης και της πιο κοινής κλάσης. Για παράδειγμα, μπορείτε να καθορίσετε ότι πρέπει να υπάρχει το πολύ μια διαφορά 2:1 στις συχνότητες των κλάσεων. Όταν χρησιμοποιείται σε λειτουργία παρτίδας, οι επόμενες παρτίδες ΔΕΝ επαναδειγματοληφτούνται.
- **AttributeSelection:** Ένα φίλτρο χαρακτηριστικών με επίβλεψη που μπορεί να χρησιμοποιηθεί για την επιλογή χαρακτηριστικών. Είναι πολύ ευέλικτο και επιτρέπει το συνδυασμό διαφόρων μεθόδων αναζήτησης και αξιολόγησης.

ΑΠΟΤΕΛΕΣΜΑΤΑ

Τα παραπάνω χαρακτηριστικά εξετάστηκαν από τέσσερις (4) διαφορετικούς αλγορίθμους τον Naïve Bayes, τον Random Forest, τον J48 και τον KNN για N=3, 4, 5. Κάθε πίνακας απεικονίζει την τιμή μίας μέτρησης για κάθε αλγόριθμο ξεχωριστά. Αναλυτικότερα τα χαρακτηριστικά που χρειάζονται ενδιαφέρον είναι τα εξής:

- Σωστά ταξινομημένες περιπτώσεις (Correctly Classified Instances): Αυτό αναφέρεται στον αριθμό των σημείων δεδομένων ή των περιπτώσεων σε ένα σύνολο δεδομένων για τα οποία ένα μοντέλο μηχανικής μάθησης έχει προβλέψει σωστά τη μεταβλητή-στόχο. Σε ένα δυαδικό πρόβλημα ταξινόμησης (π.χ., spam vs. μη spam emails), μια σωστά ταξινομημένη περίπτωση σημαίνει ότι το μοντέλο προέβλεψε σωστά την κλάση (π.χ. προέβλεψε "spam" για ένα πραγματικό "spam" email).
- Εσφαλμένα ταξινομημένες περιπτώσεις (Incorrectly Classified Instances): Πρόκειται για τα σημεία δεδομένων ή τις περιπτώσεις που το μοντέλο προέβλεψε εσφαλμένα. Στο πλαίσιο της δυαδικής ταξινόμησης, σημαίνει ότι το μοντέλο προέβλεψε μια κλάση, αλλά η πραγματική κλάση είναι η άλλη (π.χ. προέβλεψε "spam" για ένα πραγματικό "μη spam" email).
- Στατιστικό Kappa (Kappa Statistic): Το στατιστικό Kappa (επίσης γνωστό ως Cohen's Kappa) μετρά τη συμφωνία μεταξύ των προβλεπόμενων ταξινομήσεων ενός μοντέλου και των πραγματικών ταξινομήσεων, λαμβάνοντας υπόψη την πιθανότητα η συμφωνία να προκύπτει τυχαία. Παρέχει ένα μέτρο της απόδοσης του μοντέλου που λαμβάνει υπόψη την τυχαία συμφωνία.
- Μέσο απόλυτο σφάλμα – Mean Absolute Error (MAE): Το MAE είναι μια μετρική που χρησιμοποιείται για τη μέτρηση της μέσης απόλυτης διαφοράς μεταξύ των προβλεπόμενων τιμών και των πραγματικών τιμών σε ένα πρόβλημα παλινδρόμησης. Σας δίνει μια ιδέα για το πόσο απέχουν οι προβλέψεις σας κατά μέσο όρο.
- Μέσο τετραγωνικό σφάλμα ρίζας – Root Mean Square Error (RMSE): Το RMSE είναι μια άλλη μετρική που χρησιμοποιείται σε προβλήματα παλινδρόμησης για τη μέτρηση της τετραγωνικής ρίζας του μέσου όρου των τετραγωνικών διαφορών μεταξύ των προβλεπόμενων και των πραγματικών τιμών. Χρησιμοποιείται συχνά όταν τα μεγαλύτερα σφάλματα πρέπει να τιμωρούνται περισσότερο.
- Σχετικό απόλυτο σφάλμα: Πρόκειται για ένα μέτρο του μέσου απόλυτου σφάλματος ως ποσοστό του μέσου όρου των πραγματικών τιμών. Παρέχει μια σχετική κατανόηση του σφάλματος.
- Ριζικό σχετικό τετραγωνικό σφάλμα – Relative Absolute Error (RMSE): Παρόμοια με το RMSE, αυτή η μετρική είναι ένα σχετικό μέτρο του τετραγωνικού σφάλματος, που συχνά εκφράζεται ως ποσοστό της τετραγωνικής διαφοράς μεταξύ του μέσου όρου των πραγματικών τιμών.
- Συνολικός αριθμός περιπτώσεων (Total Number of Cases): Αυτό αναφέρεται απλώς στον συνολικό αριθμό σημείων δεδομένων ή περιπτώσεων στο σύνολο δεδομένων σας που χρησιμοποιούνται για την εκπαίδευση και την αξιολόγηση του μοντέλου μηχανικής μάθησης. Πρόκειται για το μέγεθος του συνόλου δεδομένων σας.

Αυτές οι μετρικές και οι όροι είναι ζωτικής σημασίας για την αξιολόγηση της απόδοσης και της ακρίβειας των μοντέλων μηχανικής μάθησης, τόσο σε εργασίες ταξινόμησης όσο και σε εργασίες παλινδρόμησης. Βοηθούν τους επιστήμονες δεδομένων και τους

επαγγελματίες της μηχανικής μάθησης να κατανοήσουν πόσο καλά αποδίδουν τα μοντέλα τους και πού μπορεί να χρειάζονται βελτιώσεις.

- **Τάξη (Class):** Στην δυαδική ταξινόμηση, υπάρχουν δύο κλάσεις, που συνήθως συμβολίζονται ως "θετικές" και "αρνητικές". Ο όρος αυτός προσδιορίζει ποια κλάση συζητείται ή αξιολογείται στις ακόλουθες μετρικές.
- **Ποσοστό TP (True Positive Rate ή Sensitivity):** Είναι ο λόγος των σωστά προβλεπόμενων θετικών περιπτώσεων (αληθώς θετικών) προς τον συνολικό αριθμό των πραγματικών θετικών περιπτώσεων. Μετρά την ικανότητα του μοντέλου να εντοπίζει σωστά θετικές περιπτώσεις.
- **Ποσοστό FP (False Positive Rate, ποσοστό ψευδώς θετικών αποτελεσμάτων):** Είναι ο λόγος των εσφαλμένα προβλεπόμενων θετικών περιπτώσεων (ψευδώς θετικές) προς τον συνολικό αριθμό των πραγματικών αρνητικών περιπτώσεων. Μετρά την τάση του μοντέλου να ταξινομεί αρνητικές περιπτώσεις ως θετικές.
- **Ακρίβεια (Precision):** Η ακρίβεια είναι ο λόγος των σωστά προβλεπόμενων θετικών περιπτώσεων (αληθώς θετικές) προς το συνολικό αριθμό των περιπτώσεων που προβλέπονται ως θετικές (αληθώς θετικές συν ψευδώς θετικές). Ποσοτικοποιεί την ακρίβεια των θετικών προβλέψεων που κάνει το μοντέλο.
- **Ανάκληση (ποσοστό αληθώς θετικών ή ευαισθησία):** Η ανάκληση είναι συνώνυμη με το ποσοστό TP. Είναι ο λόγος των σωστά προβλεπόμενων θετικών περιπτώσεων (αληθώς θετικών) προς το συνολικό αριθμό των πραγματικών θετικών περιπτώσεων. Μετρά πόσο καλά το μοντέλο καταγράφει όλες τις πραγματικές θετικές περιπτώσεις.
- **F-Measure:** Το F-Measure είναι ο αρμονικός μέσος όρος της ακρίβειας και της ανάκλησης. Συνδυάζει και τις δύο μετρικές σε μια ενιαία τιμή και είναι ιδιαίτερα χρήσιμος όταν υπάρχει ανισορροπία μεταξύ των θετικών και των αρνητικών κλάσεων.
- **MCC (Συντελεστής συσχέτισης Matthews):** Ο MCC είναι ένα μέτρο της ποιότητας των προβλέψεων δυαδικής ταξινόμησης. Λαμβάνει υπόψη τα αληθώς θετικά, τα αληθώς αρνητικά, τα ψευδώς θετικά και τα ψευδώς αρνητικά και παρέχει μια ισορροπημένη αξιολόγηση της απόδοσης του μοντέλου, ιδίως όταν πρόκειται για ανισόρροπα σύνολα δεδομένων.
- **Περιοχή ROC (Receiver Operating Characteristic Area):** Η περιοχή ROC (ROC Area) αντιπροσωπεύει την περιοχή κάτω από την καμπύλη ROC (Receiver Operating Characteristic). Η καμπύλη ROC απεικονίζει το ποσοστό TP έναντι του ποσοστού FP για διάφορα κατώτατα όρια ταξινόμησης. Η περιοχή ROC ποσοτικοποιεί την ικανότητα του μοντέλου να διακρίνει μεταξύ θετικών και αρνητικών κλάσεων.
- **Περιοχή PRC (Περιοχή καμπύλης ακρίβειας-ανάκλησης):** Το PRC Area (Περιοχή PRC) αντιπροσωπεύει την περιοχή κάτω από την καμπύλη ακρίβειας-ανάκλησης. Αυτή η καμπύλη απεικονίζει την ακρίβεια έναντι της ανάκλησης για διαφορετικά κατώτατα όρια ταξινόμησης. Το PRC Area είναι ιδιαίτερα χρήσιμο όταν πρόκειται για ανισόρροπα σύνολα δεδομένων και εστιάζει στην απόδοση της θετικής κλάσης.

Αυτές οι ποσότητες βοηθούν στην αξιολόγηση του πόσο καλά αποδίδει ένα δυαδικό μοντέλο ταξινόμησης όσον αφορά τον ορθό εντοπισμό θετικών περιπτώσεων, την ελαχιστοποίηση των ψευδώς θετικών και την εξισορρόπηση της ακρίβειας και της ανάκλησης. Είναι απαραίτητες για την κατανόηση των δυνατών και αδύνατων σημείων του μοντέλου όσον αφορά τη διαφοροποίηση μεταξύ των δύο κλάσεων.

Για την προβλεψη των καρδιοαγγειακών παθήσεων χρησιμοποιήθηκαν τα μοντέλα Naïve Bayes, Random Forest, J48 και KNN (N=3, 4, 5). Η επιλογή του καταλληλότερου μοντέλου έγινε σύμφωνα με τις τιμές του ROC καθώς αποτελεί σημαντικό μέτρο για την αξιολόγηση μοντέλων με διακριτές κλάσεις. Για την κολπική μαρμαρυγή ο αλγόριθμος Naïve Bayes παρουσιάζει την καλύτερη επίδοση (ROC=0.752). Για την υπερκοιλιακή ταχυκαρδία παρατηρούμε παρόμοια αποτελέσματα με τον Naïve Bayes να εξακολουθεί να έχει καλύτερη επίδοση έναντι των υπολοίπων αλγορίθμων με τιμή ROC=0.836. Ίδια εικόνα παραμένει όσον αφορά και την κοιλιακή μαρμαρυγή με τον Naïve Bayes να έχει τη μεγαλύτερη τιμή (ROC=0.825). Σύμφωνα με τους παραπάνω πίνακες ο αλγόριθμος Naïve Bayes παρουσιάζει την καλύτερη επίδοση (ROC=0.858) για τον αποκλεισμό τρίτου βαθμού. Για το πνευμονικό οίδημα η κατάσταση παραμένει ίδια όπου ο αλγόριθμος Naïve Bayes παρουσιάζει την καλύτερη επίδοση (ROC=0.806). Για τη ρήξη του μυοκαρδίου η επίδοση των μοντέλων ήταν εξαιρετική με τον Naïve bayes να παραμένει το καλύτερο μοντέλο (ROC=0.850). Για το σύνδρομο Dressler η απόδοση του Naïve bayes ήταν στο 0.850 ενώ τα υπόλοιπα μοντέλα παρουσίασαν μέτριες έως κακές επιδόσεις. Για την χρόνια καρδιακή ανεπάρκεια ο Naïve Bayes συνεχίζει τις εξαιρετικές επιδόσεις έναντι των υπολοίπων μοντέλων με τιμή ROC=0.719. Για την υποτροπή του εφράγματος του μυοκαρδίου ο Naïve bayes συνεχίζει να παραμένει το καταλληλότερο μοντέλο έστω και αν η τιμή του δείχνει μέτριες επιδόσεις (ROC=0.698). Σχετικά με την εκδήλωση στηθάγχης μετά το έμφραγμα ο Naïve Bayes επιλέγεται ως το πιο αποδοτικό μοντέλο με τιμή ROC=0.694 ενώ τα υπόλοιπα μοντέλα δεν κρίνονται κατάλληλα βάσει της απόδοσής τους. Τέλος για θανατηφόρο αίτιο τα μοντέλα Naïve Bayes και Random Forest είναι τα καταλληλότερα μεταξύ άλλων με ισάξιες επιδόσεις ROC=0.628 ενώ τα υπόλοιπα δεν κρίνονται κατάλληλα για τέτοιου είδους αξιολόγηση.

Στους παρακάτω πίνακες παρουσιάζονται αναλυτικά τα αποτελέσματα εξόδου για κάθε χαρακτηριστικό και οι επιδόσεις του εκάστοτε αλγορίθμου που χρησιμοποιήθηκε.

Κολπική μαρμαρυγή (FIBRS_PREDS)

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 340 (MaxCount: 170) από τις αρχικές 1530. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 19 χαρακτηριστικά εισόδου τα οποία όλα έχουν κάτω από 20% missing data και ένα εξόδου (FIBRS_PREDS). Στο κομμάτι των αποτελεσμάτων ο αλγόριθμος Naive Bayes παρουσιάζει τις περισσότερες σωστές προβλέψεις (234) ενώ ακολουθούν οι J48 (207), ο Random Forest (205) και ο KKN για N=3 (204). Πιο χαμηλές είναι οι επιδόσεις του KNN για N=4 και N=5 με 196 και 201 σωστές κατηγοριοποιήσεις αντίστοιχα. Στην περίπτωση του Naive Bayes επίσης παρατηρούμε μια αρκετά ικανοποιητική τιμή όσον αφορά το Kappa statistic (0.37), αρκετά μεγαλύτερη τιμή από τους άλλους αλγορίθμους που έχουν απόδοση κοντά στο 0,2, με την μικρότερη απόδοση να σημειώνει ο KNN για N=4 (0.1529). Οι τιμές για το Mean Absolute Error κυμαίνονται περίπου στα ίδια επίπεδα για όλους τους αλγορίθμους (μεταξύ 0.38 και 0.44), κάτι που συμβαίνει και για το Root mean squared error (από 0.45 έως 0.51). Ο Naive Bayes ξεχωρίζει και στις μετρήσεις για το Relative Absolute Error, σημειώνοντας 76.58% αισθητά μικρότερη τιμή σε σχέση με τους υπόλοιπους αλγορίθμους που οι τιμές τους είναι ανάμεσα στο 86% και το 88%. Το ίδιο παρατηρούμε και στις μετρήσεις για το Root Relative Squared Error με τον Naive Bayes να έχει απόδοση 90.44% και όλους τους άλλους να ακολουθούν με επιδόσεις άνω του 100%, με μόνη εξαίρεση τον KNN για N=5 όπου το Root Relative Squared Error έχει τιμή 97.92%.

NAÏVE BAYES

Πίνακας 2 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naive Bayes)

Metric	Value
Correctly Classified Instances	234 (68.8235%)
Incorrectly Classified Instances	106 (31.1765%)
Kappa statistic	0.3765
Mean Absolute Error	0.3829
Root Mean Squared Error	0.4522
Relative Absolute Error	76.5803%
Root Relative Squared Error	90.4433%
Total Number of Instances	340

Πίνακας 3 Σταθμισμένες μετρήσεις επιδόσεων (Naive Bayes)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.729	0.353	0.674	0.729	0.701	0.378	0.752	0.766
1	0.647	0.271	0.705	0.647	0.675	0.378	0.752	0.755
	0.688		0.690	0.688	0.688	0.378		

Weighted Avg.		0.312					0.752	
---------------	--	-------	--	--	--	--	-------	--

Random Forest

Πίνακας 4 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Correctly Classified Instances:	205
Incorrectly Classified Instances:	135
Kappa statistic:	0.2059
Mean Absolute Error:	0.4405
Root Mean Squared Error:	0.5025
Relative Absolute Error:	88.0915%
Root Relative Squared Error:	100.4978%
Total Number of Instances:	340

Πίνακας 5 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.676	0.471	0.590	0.676	0.630	0.208	0.640	0.640
1	0.529	0.324	0.621	0.529	0.571	0.208	0.640	0.626
Weighted Avg.	0.603	0.397	0.605	0.603	0.601	0.208	0.640	0.633

J48

Πίνακας 6 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	207 (60.8824%)
Incorrectly Classified Instances	133 (39.1176%)
Kappa statistic	0.2176
Mean Absolute Error	0.4424
Root Mean Squared Error	0.515
Relative Absolute Error	88.4748%
Root Relative Squared Error	103.0053%
Total Number of Instances	340

Πίνακας 7 Σταθμισμένες μετρήσεις επιδόσεων (J48)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.676	0.459	0.596	0.676	0.634	0.220	0.624	0.583
1	0.541	0.324	0.626	0.541	0.580	0.220	0.624	0.612
Weighted Avg.	0.609	0.391	0.611	0.609	0.607	0.220	0.624	0.598

KKN

K=3

Πίνακας 8 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value
Correctly Classified Instances	204 (60%)
Incorrectly Classified Instances	136 (40%)
Kappa statistic	0.2
Mean Absolute Error	0.4322
Root Mean Squared Error	0.5162
Relative Absolute Error	86.4308%
Root Relative Squared Error	103.2457%
Total Number of Instances	340

Πίνακας 9 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.676	0.459	0.596	0.676	0.634	0.220	0.624	0.583
1	0.541	0.324	0.626	0.541	0.580	0.220	0.624	0.612
Weighted Avg.	0.609	0.391	0.611	0.609	0.607	0.220	0.624	0.598

KNN

N=4

Πίνακας 10 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Correctly Classified Instances	196
Incorrectly Classified Instances	144
Kappa statistic	0.1529
Mean absolute error	0.434

Root mean squared error	0.5004
Relative absolute error	86.7906 %
Root relative squared error	100.0868 %
Total Number of Instances	340

Πίνακας 11 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.682	0.529	0.563	0.682	0.617	0.156	0.637	0.617
1	0.471	0.318	0.597	0.471	0.526	0.156	0.637	0.624
Weighted Avg.	0.576	0.424	0.580	0.576	0.572	0.156	0.637	0.621

KNN

N=5

Πίνακας 12 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Total Number of Instances	340
Root relative squared error	97.9206 %
Relative absolute error	86.1952 %
Root mean squared error	0.4896
Mean absolute error	0.431
Kappa statistic	0.1824
Incorrectly Classified Instances	139
Correctly Classified Instances	201

Πίνακας 13 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area		
0	0.624	0.441	0.586	0.624	0.604	0.183	0.646	0.625		
1	0.559	0.376	0.597	0.559	0.578	0.183	0.646	0.643		
Weighted Avg.	0.591	0.409	0.592	0.591	0.591	0.183	0.646	0.634		

Υπερκοιλιακή ταχυκαρδία (PREDS_TAH)

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 40 (MaxCount: 20) από τις αρχικές 1680. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 10 χαρακτηριστικά εισόδου και μετά από αξιολόγηση απομένουν 9 καθώς το χαρακτηριστικό D_AD_KBRIG έχει 55% missing data και δεν μπορεί να χρησιμοποιηθεί τα υπόλοιπα χαρακτηριστικά έχουν κάτω από 20% missing data. Αν και αριθμός των περιπτώσεων μοιάζει να είναι μικρός, από άποψη απόδοσης σωστής κατηγοριοποίησης, ο Random Forest έχει τις περισσότερες επιτυχημένες (28), με τον Naive Bayes να έχει μόλις μια λιγότερη(27). Από κοντά είναι και ο KNN για N=3 με 25 σωστές κατηγοριοποιήσεις, ενώ για N=4 και N=5 έχει 21 και 20 αντίστοιχα. Σε αυτό το κομμάτι την χειρότερη απόδοση την εμφανίζει ο J48 με μόλις 18. Για την μέτρηση Kappa statistic ο αλγόριθμος Random Forest εμφανίζει μια καλή απόδοση με την μέτρηση να φτάνει το 0.4. Εξίσου καλή είναι και η τιμή του Naive Bayes με 0.35 . Από εκεί και πέρα, αν εξαιρέσουμε την τιμή του KNN για N=3, όπου το Kappa statistic είναι 0.25 , όλες οι άλλες τιμές του στους αλγορίθμους είναι χαμηλές. Ο J48 έχει 0.1 , ο KNN για N=4 έχει 0.05 και για N=5 0 . Για το Mean Absolute Error, ο Naive Bayes εμφανίζει το μικρότερο με 0.34 και ακολουθούν οι Random Forest (0.36) και KNN για N=3 (0.38). Μεγαλύτεροι είναι οι αριθμοί για J48 (0.47) και για KNN όπου N=4 και N=5 (0.45 και 0.46 αντίστοιχα).Οι ίδιοι αλγόριθμοι ξεχωρίζουν και για το Root mean squared error. Η μικρότερη τιμή εμφανίζεται στον αλγόριθμο Naive Bayes (0.41) και η μεγαλύτερη στους J48 και KNN(K=4 και K=5) με 0.51. Για το Relative absolute error, πάλι ο Naive Bayes έχει την μικρότερη τιμή (69,61%), με τον με τον Random Forest να ακολουθεί(73.51%).Οι μεγαλύτερες τιμές εμφανίζεται στον J48 (95.76%). Ίδια εικόνα και για το Root relative squared error με τον Bayes να έχει 90.64% και τον J48 να εμφανίζει 103.04%.

NAÏVE BAYES

Πίνακας 14 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naive Bayes)

Metric	Value
Correctly Classified Instances	27
Incorrectly Classified Instances	13
Kappa statistic	0.35
Mean absolute error	0.3481
Root mean squared error	0.4123
Relative absolute error	69.6168 %
Root relative squared error	82.4649 %
Total Number of Instances	40

Πίνακας 15 Σταθμισμένες μετρήσεις επιδόσεων (Naive Bayes)

Class	TP Rate	FP Rate	Precisio	Recall	F-Measure	MCC	ROC Area	PRC Area
-------	---------	---------	----------	--------	-----------	-----	----------	----------

			n					
0	0.600	0.250	0.706	0.600	0.649	0.354	0.836	0.836
1	0.750	0.400	0.652	0.750	0.698	0.354	0.836	0.855
Weighted Avg.	0.675	0.325	0.679	0.675	0.673	0.354	0.836	0.846

RANDOM FOREST

Πίνακας 16 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Metric	Value
Correctly Classified Instances	28
Incorrectly Classified Instances	12
Kappa statistic	0.4
Mean absolute error	0.3676
Root mean squared error	0.4532
Relative absolute error	73.5152 %
Root relative squared error	90.6415 %
Total Number of Instances	40

Πίνακας 17 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.800	0.400	0.667	0.800	0.727	0.408	0.769	0.795
1	0.600	0.200	0.750	0.600	0.667	0.408	0.769	0.704
Weighted Avg.	0.700	0.300	0.708	0.700	0.697	0.408	0.769	0.749

J48

Πίνακας 18 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	22
Incorrectly Classified Instances	18
Kappa statistic	0.1
Mean absolute error	0.4788
Root mean squared error	0.5152
Relative absolute error	95.7619 %
Root relative squared error	103.047 %
Total Number of Instances	40

Πίνακας 19 Σταθμισμένες μετρήσεις επιδόσεων (J48)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.400	0.300	0.571	0.400	0.471	0.105	0.554	0.508
1	0.700	0.600	0.538	0.700	0.609	0.105	0.554	0.602
Weighted Avg.	0.550	0.450	0.555	0.550	0.540	0.105	0.554	0.555

KNN

N=3

Πίνακας 20 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value
Correctly Classified Instances	25
Incorrectly Classified Instances	15
Kappa statistic	0.25
Mean absolute error	0.3819
Root mean squared error	0.4948
Relative absolute error	76.3861 %
Root relative squared error	98.9574 %
Total Number of Instances	40

Πίνακας 21 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.800	0.550	0.593	0.800	0.681	0.267	0.754	0.766
1	0.450	0.200	0.692	0.450	0.545	0.267	0.754	0.687
Weighted Avg.	0.625	0.375	0.642	0.625	0.613	0.267	0.754	0.727

KNN

N=4

Πίνακας 22 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Metric	Value
Correctly Classified Instances	21
Incorrectly Classified Instances	19
Kappa statistic	0.05
Mean absolute error	0.4521
Root mean squared error	0.5167
Relative absolute error	90.4282 %

Root relative squared error	103.3323 %
Total Number of Instances	40

Πίνακας 23 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.650	0.600	0.520	0.650	0.578	0.052	0.649	0.649
1	0.400	0.350	0.533	0.400	0.457	0.052	0.649	0.664
Weighted Avg.	0.525	0.475	0.527	0.525	0.517	0.052	0.649	0.656

KNN

N=5

Πίνακας 24 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Correctly Classified Instances	20
Incorrectly Classified Instances	20
Kappa statistic	0
Mean absolute error	0.464
Root mean squared error	0.5148
Relative absolute error	92.792 %
Root relative squared error	102.9573 %
Total Number of Instances	40

Πίνακας 25 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.600	0.600	0.500	0.600	0.545	0.000	0.616	0.623
1	0.400	0.400	0.500	0.400	0.444	0.000	0.616	0.618
Weighted Avg.	0.500	0.500	0.500	0.500	0.495	0.000	0.616	0.621

Κοιλιακή ταχυκαρδία (JELUD_TAH)

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 84 (MaxCount: 42) από τις αρχικές 1658. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 17 χαρακτηριστικά εισόδου τα οποία πληρούν τις προϋποθέσεις ως χαρακτηριστικά εισόδου κάτω από 20% missing data. Όσον αφορά τα αποτελέσματα ο Naive Bayes(NB) έχει τις περισσότερες σωστές κατηγοριοποιήσεις με 69 ενώ ακολουθούν οι RandomForest(58), J48(55) KNN για N=3(53) και N=4(51) και N=5(54). Σχετικά με το Kappa statistic παρατηρούμε μερικές εξαιρετικές τιμές. Συγκεκριμένα ο NB έχει 0.64, ο RandomForest έχει 0.38 και ο J48 0.30. Χαμηλότερες τιμές έχει ο KNN όπου για N=3,4,5 παρουσιάζει τιμές 0.26, 0.21 και 0.28 αντίστοιχα. Για το Mean Absolute Error ο NB έχει το μικρότερο με 0.25, και ακολουθούν οι Random Forest (0.37), J48 (0.39) KNN για N=3(0.37). Ίδια παρατήρηση μπορούμε να κάνουμε και για το Root mean squared error, αφού η χαμηλότερη τιμή εμφανίζεται στον NB (0.37). Ο NB εμφανίζει και αρκετά χαμηλές τιμές για το Relative absolute error (50,93%) , αρκετά χαμηλότερο από τον αμέσως επόμενο αλγόριθμο που είναι ο RandomForest (73.98%). Για το Root relative squared error τη χαμηλότερη τιμή την έχει ο NB και πάλι (75.39%).

NAÏVE BAYES

Πίνακας 26 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes)

Metric	Value
Correctly Classified Instances	69
Incorrectly Classified Instances	15
Kappa statistic	0.6429
Mean absolute error	0.2548
Root mean squared error	0.3772
Relative absolute error	50.934 %
Root relative squared error	75.396 %
Total Number of Instances	84

Πίνακας 27 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.810	0.167	0.829	0.810	0.819	0.643	0.871	0.888
1	0.833	0.190	0.814	0.833	0.824	0.643	0.871	0.850
Weighted Avg.	0.821	0.179	0.822	0.821	0.821	0.643	0.871	0.869

RANDOM FOREST

Πίνακας 28 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Metric	Value
Correctly Classified Instances	58
Incorrectly Classified Instances	26
Kappa statistic	0.381
Mean absolute error	0.3702
Root mean squared error	0.427
Relative absolute error	73.989%
Root relative squared error	85.3489%
Total Number of Instances	84

Πίνακας 29 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.690	0.310	0.690	0.690	0.690	0.381	0.800	0.814
1	0.690	0.310	0.690	0.690	0.690	0.381	0.800	0.760
Weighted Avg.	0.690	0.310	0.690	0.690	0.690	0.381	0.800	0.787

J48

Πίνακας 30 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	55
Incorrectly Classified Instances	29
Kappa statistic	0.3095
Mean absolute error	0.3977
Root mean squared error	0.4832
Relative absolute error	79.4912%
Root relative squared error	96.5819%
Total Number of Instances	84

Πίνακας 31 Σταθμισμένες μετρήσεις επιδόσεων (J48)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.667	0.357	0.651	0.667	0.659	0.310	0.674	0.626
1	0.643	0.333	0.659	0.643	0.651	0.310	0.674	0.642
Weighted Avg.	0.655	0.345	0.655	0.655	0.655	0.310	0.674	0.634

KNN

N=3

Πίνακας 32 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value	Percentage
Correctly Classified Instances	53	63.0952%
Incorrectly Classified Instances	31	36.9048%
Kappa statistic	0.2619	
Mean absolute error	0.3755	
Root mean squared error	0.473	
Relative absolute error	75.0505%	
Root relative squared error	94.5336%	
Total Number of Instances	84	

Πίνακας 33 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.810	0.548	0.596	0.810	0.687	0.280	0.751	0.790
1	0.452	0.190	0.704	0.452	0.551	0.280	0.750	0.723
Weighted Avg.	0.631	0.369	0.650	0.631	0.619	0.280	0.750	0.757

KNN

N=4

Πίνακας 34 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Metric	Value
Correctly Classified Instances	51
Incorrectly Classified Instances	33
Kappa statistic	0.2143
Mean absolute error	0.4022
Root mean squared error	0.4788
Relative absolute error	80.3911%
Root relative squared error	95.7021%

Total Number of Instances	84
---------------------------	----

Πίνακας 35 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.833	0.619	0.574	0.833	0.680	0.240	0.719	0.772
1	0.381	0.167	0.696	0.381	0.492	0.240	0.718	0.715
Weighted Avg.	0.607	0.393	0.635	0.607	0.586	0.240	0.718	0.743

KNN

N=5

Πίνακας 36 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Correctly Classified Instances	54
Incorrectly Classified Instances	30
Kappa statistic	0.2857
Mean absolute error	0.3983
Root mean squared error	0.4674
Relative absolute error	79.6092%
Root relative squared error	93.4152%
Total Number of Instances	84

Πίνακας 37 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.857	0.571	0.600	0.857	0.706	0.316	0.741	0.778
1	0.429	0.143	0.750	0.429	0.545	0.316	0.742	0.742
Weighted Avg.	0.643	0.357	0.675	0.643	0.626	0.316	0.741	0.760

Κοιλιακή μαρμαρυγή (FIBR_JELUD)

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 142 (MaxCount: 71) από τις αρχικές 1629. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 8 χαρακτηριστικά εισόδου τα οποία όλα έχουν κάτω από 20% missing data και ένα εξόδου (FIBR_JELUD). Με σειρά των περισσότερων correctly classified instances είναι οι αλγόριθμοι: Random Forest(111), Naive Bayes και KNN για N=3 (104), J48(103) και ακολουθεί πάλι ο KNN για N=4 και N=5. Παρατηρούμε ότι οι αλγόριθμοι είναι σχετικά κοντά στο θέμα σωστών κατηγοριοποιήσεων. Πολύ καλές είναι και οι τιμές και για το kappa statistic, με τον Random Forest(RF) να έχει τιμή 0.56 ενώ ικανοποιητικές είναι και οι τιμές που εμφανίζουν και οι NB(0.46), J48(0.45) και KNN, N=3(0.46). Για το Mean absolute error, οι τιμές είναι σχετικά κοντά, με την μικρότερη τιμή να έχει ο RF(0.3) και την μεγαλύτερη ο KNN, N=4 (0.39). Ο αλγόριθμος RF συνεχίζει να έχει καλύτερες τιμές και στις υπόλοιπες μετρήσεις, αφού για το Root mean squared error έχει τιμή 0.41, για το Relative absolute error 61.15% και για το Root relative squared error 83,74%.

Naïve bayes

Πίνακας 38 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes)

Metric	Value
Correctly Classified Instances	104
Incorrectly Classified Instances	38
Kappa statistic	0.4648
Mean absolute error	0.3236
Root mean squared error	0.4211
Relative absolute error	64.721%
Root relative squared error	84.2041%
Total Number of Instances	142

Πίνακας 39 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.761	0.296	0.720	0.761	0.740	0.466	0.825	0.822
1	0.704	0.239	0.746	0.704	0.725	0.466	0.825	0.825
Weighted Avg.	0.732	0.268	0.733	0.732	0.732	0.466	0.825	0.824

Random forest

Πίνακας 40 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Metric	Value
Correctly Classified Instances	111
Incorrectly Classified Instances	31
Kappa statistic	0.5634
Mean absolute error	0.3058
Root mean squared error	0.4188
Relative absolute error	61.1553%
Root relative squared error	83.7425%
Total Number of Instances	142

Πίνακας 41 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.732	0.169	0.813	0.732	0.770	0.566	0.816	0.823
1	0.831	0.268	0.756	0.831	0.792	0.566	0.816	0.789
Weighted Avg.	0.782	0.218	0.784	0.782	0.781	0.566	0.816	0.806

J48

Πίνακας 42 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	103
Incorrectly Classified Instances	39
Kappa statistic	0.4507
Mean absolute error	0.3793
Root mean squared error	0.4509
Relative absolute error	75.8604%
Root relative squared error	90.1656%
Total Number of Instances	142

Πίνακας 43 Σταθμισμένες μετρήσεις επιδόσεων (J48)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.690	0.239	0.742	0.690	0.715	0.452	0.689	0.666

1	0.761	0.310	0.711	0.761	0.735	0.452	0.689	0.618
Weighted Avg.	0.725	0.275	0.726	0.725	0.725	0.452	0.689	0.642

KKN

N=3

Πίνακας 44 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value
Correctly Classified Instances	104
Incorrectly Classified Instances	38
Kappa statistic	0.4648
Mean absolute error	0.3132
Root mean squared error	0.4318
Relative absolute error	62.6326%
Root relative squared error	86.353%
Total Number of Instances	142

Πίνακας 45 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.761	0.296	0.720	0.761	0.740	0.466	0.798	0.791
1	0.704	0.239	0.746	0.704	0.725	0.466	0.798	0.742
Weighted Avg.	0.732	0.268	0.733	0.732	0.732	0.466	0.798	0.767

KNN

N=4

Πίνακας 46 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Metric	Value
Correctly Classified Instances	99
Incorrectly Classified Instances	43
Kappa statistic	0.3944
Mean absolute error	0.3363
Root mean squared error	0.447

Relative absolute error	67.2473%
Root relative squared error	89.3903%
Total Number of Instances	142

Πίνακας 47 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.761	0.366	0.675	0.761	0.715	0.398	0.783	0.781
1	0.634	0.239	0.726	0.634	0.677	0.398	0.783	0.732
Weighted Avg.	0.697	0.303	0.700	0.697	0.696	0.398	0.783	0.756

KNN

N=5

Πίνακας 48 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Correctly Classified Instances	98
Incorrectly Classified Instances	44
Kappa statistic	0.3803
Mean absolute error	0.3363
Root mean squared error	0.4445
Relative absolute error	67.2465%
Root relative squared error	88.8959%
Total Number of Instances	142

Πίνακας 49 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.775	0.394	0.663	0.775	0.714	0.386	0.796	0.789
1	0.606	0.225	0.729	0.606	0.662	0.386	0.796	0.754
Weighted Avg.	0.690	0.310	0.696	0.690	0.688	0.386	0.796	0.772

Σύμφωνα με τους παραπάνω πίνακες ο αλγόριθμος Naïve Bayes παρουσιάζει την καλύτερη επίδοση (ROC=0.825).

Αποκλεισμός AV τρίτου βαθμού (A_V_BLOK)

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 114 (MaxCount: 57) από τις αρχικές 1643. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 12 χαρακτηριστικά εισόδου τα οποία όλα έχουν κάτω από 20% missing data και ένα εξόδου (A_V_BLOK). Τα αποτελέσματα για το συγκεκριμένο χαρακτηριστικό εξόδου είναι αρκετά ενθαρυντικά. Τα Correctly Classified Instances είναι 87 για τους αλγόριθμους NB, RF και KNN (N=4), ενώ για N=3 και N=5, ο KNN έχει μόλις ένα λιγότερο. Ο J48 έχει 81 σωστές κατηγοριοποιήσεις. Συνεπώς οι αλγόριθμοι που έχουν τα ίδια Correctly Classified Instances, έχουν τις ίδιες τιμές και για το kappa statistic. Για την μέτρηση Mean absolute error η μικρότερη τιμή συναντάται στον NB με τιμή 0.25, και ακολουθούν οι KNN, N=3 με τιμή 0.29 και N=4 με τιμή 0.30. Την μεγαλύτερη τιμή την βλέπουμε στον J48 με 0.36. Για την μέτρηση Root mean squared error όλοι οι αλγόριθμοι έχουν τιμή 0.41, με μόνη εξαίρεση τον J48 που έχει τιμή 0.45. Για το Relative absolute error η μικρότερη τιμή εμφανίζεται στον αλγόριθμο NB με 50.73% και η μεγαλύτερη στον J48(73.57%). Από την άλλη, για την μέτρηση Root relative squared error την καλύτερη απόδοση την έχει ο KNN, N=3 με 82.18%, τιμή ελάχιστα πιο μικρή από τον RF(82.42%).

NAÏVE BAYES

Πίνακας 50 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes)

Metric	Value
Correctly Classified Instances	87
Incorrectly Classified Instances	27
Kappa statistic	0.5263
Mean absolute error	0.2538
Root mean squared error	0.4173
Relative absolute error	50.7397 %
Root relative squared error	83.4113 %
Total Number of Instances	114

Πίνακας 51 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.772	0.246	0.759	0.772	0.765	0.526	0.858	0.809
1	0.754	0.228	0.768	0.754	0.761	0.526	0.858	0.886
Weighted Avg.	0.763	0.237	0.763	0.763	0.763	0.526	0.858	0.848

RANDOM FOREST

Πίνακας 52 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Metric	Value
Correctly Classified Instances	87
Incorrectly Classified Instances	27
Kappa statistic	0.5263
Mean absolute error	0.3267
Root mean squared error	0.4123
Relative absolute error	65.308 %
Root relative squared error	82.4226 %
Total Number of Instances	114

Πίνακας 53 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.807	0.281	0.742	0.807	0.773	0.528	0.815	0.737
1	0.719	0.193	0.788	0.719	0.752	0.528	0.815	0.850
Weighted Avg.	0.763	0.237	0.765	0.763	0.763	0.528	0.815	0.793

J48

Πίνακας 54 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	81
Incorrectly Classified Instances	33
Kappa statistic	0.4211
Mean absolute error	0.368
Root mean squared error	0.4573
Relative absolute error	73.5708 %
Root relative squared error	91.4128 %
Total Number of Instances	114

Πίνακας 55 Σταθμισμένες μετρήσεις επιδόσεων (j48)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.719	0.298	0.707	0.719	0.713	0.421	0.733	0.700
1	0.702	0.281	0.714	0.702	0.708	0.421	0.733	0.719
Weighted Avg.	0.711	0.289	0.711	0.711	0.711	0.421	0.733	0.709

KNN

N=3

Πίνακας 56 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value
Correctly Classified Instances	86
Incorrectly Classified Instances	28
Kappa statistic	0.5088
Mean absolute error	0.2951
Root mean squared error	0.4111
Relative absolute error	58.9862 %
Root relative squared error	82.1854 %
Total Number of Instances	114

Πίνακας 57 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.789	0.281	0.738	0.789	0.763	0.510	0.837	0.799
1	0.719	0.211	0.774	0.719	0.745	0.510	0.837	0.863
Weighted Avg.	0.754	0.246	0.756	0.754	0.754	0.510	0.837	0.831

KNN

N=4

Πίνακας 58 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Metric	Value
Correctly Classified Instances	87
Incorrectly Classified Instances	27

Kappa statistic	0.5263
Mean absolute error	0.3086
Root mean squared error	0.4131
Relative absolute error	61.6909 %
Root relative squared error	82.5791 %
Total Number of Instances	114

Πίνακας 59 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.807	0.281	0.742	0.807	0.773	0.528	0.826	0.789
1	0.719	0.193	0.788	0.719	0.752	0.528	0.826	0.851
Weighted Avg.	0.763	0.237	0.765	0.763	0.763	0.528	0.826	0.820

KNN

N=5

Πίνακας 60 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Correctly Classified Instances	86
Incorrectly Classified Instances	28
Kappa statistic	0.5088
Mean absolute error	0.3162
Root mean squared error	0.4198
Relative absolute error	63.2153 %
Root relative squared error	83.9227 %
Total Number of Instances	114

Πίνακας 61 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.789	0.281	0.738	0.789	0.763	0.510	0.817	0.782
1	0.719	0.211	0.774	0.719	0.745	0.510	0.817	0.845
Weighted Avg.	0.754	0.246	0.756	0.754	0.754	0.510	0.817	0.813

Πνευμονικό οίδημα (OTEK_LANC)

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 318 (MaxCount: 159) από τις αρχικές 1541. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 10 χαρακτηριστικά εισόδου μετά από εξέταση των χαρακτηριστικών αποκλείονται 2 τα οποία δεν πληρούν τις προϋποθέσεις το NA_KB και NOT_NA_KB με 38% και 41% missing data αντίστοιχα με τελικό αριθμό 8 χαρακτηριστικών τα οποία όλα έχουν κάτω από 20% missing data και ένα εξόδου (OTEK_LANC). Σχετικά με την σωστή κατηγοριοποίηση των περιπτώσεων, οι αλγόριθμοι έχουν σχεδόν την ίδια απόδοση. Τις περισσότερες σωστές κατηγοριοποιήσεις τις πέτυχε ο NB (233), μετά ακολουθούν οι RF και J48 (222) και τέλος από κοντά ο KNN με 215, 218 και 219 για N=3,4,5 αντίστοιχα. Λογικό είναι να επικρατεί και η ίδια εικόνα και στο kappa statistic με τον NB να έχει την μεγαλύτερη τιμή (0.4654) και μετά από κοντά όλοι οι υπόλοιποι αλγόριθμοι, με κατώτερο όριο τιμής αυτό την KNN για N=3 (0.35). Οι τιμές για το Mean absolute error κυμαίνονται από 0.33 (NB) έως 0.39 (KNN, N=4). Ίσες τιμές έχουν οι αλγόριθμοι και για το Root mean squared error με τιμές από 0.42(NB) έως 0.46. Αυτό όμως δεν συμβαίνει για το Relative absolute error αφού ο NB ξεχωρίζει αρκετά με τιμή 67.01% σε σχέση με την χειρότερη απόδοση 79.14% που έχει ο J48. Η δεύτερη καλύτερη απόδοση εμφανίζεται στον αλγόριθμο RF με 74.66%. Για το Root relative squared error η μικρότερη τιμή παρουσιάζεται πάλι στον NB και είναι 84.78%.

NAÏVE BAYES

Πίνακας 62 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes)

Metric	Value
Correctly Classified Instances	233
Incorrectly Classified Instances	85
Kappa statistic	0.4654
Mean absolute error	0.3351
Root mean squared error	0.4239
Relative absolute error	67.0114 %
Root relative squared error	84.7868 %
Total Number of Instances	318

Πίνακας 63 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.792	0.327	0.708	0.792	0.748	0.469	0.806	0.768
1	0.673	0.208	0.764	0.673	0.716	0.469	0.806	0.828
	0.733	0.267	0.736	0.733	0.732	0.469	0.806	0.798

Weighted Avg.								
---------------	--	--	--	--	--	--	--	--

RANDOM FOREST

Πίνακας 64 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Metric	Value
Correctly Classified Instances	222
Incorrectly Classified Instances	96
Kappa statistic	0.3962
Mean absolute error	0.3734
Root mean squared error	0.453
Relative absolute error	74.6685 %
Root relative squared error	90.5921 %
Total Number of Instances	318

Πίνακας 65 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.736	0.340	0.684	0.736	0.709	0.397	0.746	0.684
1	0.660	0.264	0.714	0.660	0.686	0.397	0.746	0.753
Weighted Avg.	0.698	0.302	0.699	0.698	0.698	0.397	0.746	0.719

J48

Πίνακας 66 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	222
Incorrectly Classified Instances	96
Kappa statistic	0.3962
Mean absolute error	0.3957
Root mean squared error	0.4624
Relative absolute error	79.1405 %
Root relative squared error	92.478 %

Total Number of Instances	318
---------------------------	-----

Πίνακας 67 Σταθμισμένες μετρήσεις επιδόσεων (J48)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.755	0.358	0.678	0.755	0.714	0.399	0.691	0.625
1	0.642	0.245	0.723	0.642	0.680	0.399	0.691	0.667
Weighted Avg.	0.698	0.302	0.701	0.698	0.697	0.399	0.691	0.646

KNN

N=3

Πίνακας 68 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value
Correctly Classified Instances	215
Incorrectly Classified Instances	103
Kappa statistic	0.3522
Mean absolute error	0.3888
Root mean squared error	0.4655
Relative absolute error	77.7504 %
Root relative squared error	93.0914 %
Total Number of Instances	318

Πίνακας 69 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.805	0.453	0.640	0.805	0.713	0.365	0.731	0.667
1	0.547	0.195	0.737	0.547	0.628	0.365	0.731	0.735
Weighted Avg.	0.676	0.324	0.689	0.676	0.671	0.365	0.731	0.701

KNN

N=4

Πίνακας 70 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Metric	Value
Correctly Classified	218

Instances	
Incorrectly Classified Instances	100
Kappa statistic	0.3711
Mean absolute error	0.3926
Root mean squared error	0.4632
Relative absolute error	78.5086 %
Root relative squared error	92.6382 %
Total Number of Instances	318

Πίνακας 71 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.824	0.453	0.645	0.824	0.724	0.386	0.734	0.661
1	0.547	0.176	0.757	0.547	0.635	0.386	0.734	0.741
Weighted Avg.	0.686	0.314	0.701	0.686	0.679	0.386	0.734	0.701

KNN

N=5

Πίνακας 72 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Correctly Classified Instances	219
Incorrectly Classified Instances	99
Kappa statistic	0.3774
Mean absolute error	0.392
Root mean squared error	0.4608
Relative absolute error	78.394 %
Root relative squared error	92.1573 %
Total Number of Instances	318

Πίνακας 73 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
-------	---------	---------	-----------	--------	-----------	-----	----------	----------

0	0.830	0.453	0.647	0.830	0.727	0.393	0.738	0.669
1	0.547	0.170	0.763	0.547	0.637	0.393	0.738	0.750
Weighted Avg.	0.689	0.311	0.705	0.689	0.682	0.393	0.738	0.709

Ρήξη του μυοκαρδίου (RAZRIV)

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 108 (MaxCount: 54) από τις αρχικές 1646. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 12 χαρακτηριστικά εισόδου τα οποία όλα έχουν κάτω από 20% missing data και ένα εξόδου (RAZVIR). Τις περισσότερες σωστές κατηγοριοποιήσεις τις εμφανίζει και πάλι ο αλγόριθμος NB με 86. Ακολουθούν οι RF, KNN, N=5 και J48 με 78 σωστές, και μετά ο KNN με 73 και 74 για N=3 και N=4. Πολύ καλή είναι η τιμή του NB για το kappa statistic με 0.59. Αρκετά μικρότερη αλλά εξίσου ικανοποιητική είναι η τιμή 0.44 που εμφανίζεται στους αλγορίθμους RF, J48 και KNN, N=5. Χαμηλότερη απόδοση ο KNN για N=3 (0.35) και N=4 (0.37). Οι τιμές για το Mean absolute error κυμαίνονται από 0.30 (NB) έως 0.38 (RF). Ίσες τιμές έχουν οι αλγόριθμοι και για το Root mean squared error με τιμές από 0.39 (NB) έως 0.46 (KNN, N=5). Χαμηλή τιμή έχει ο NB για την μέτρηση του Relative absolute error (60.64%), σε σχέση με τους άλλους αλγορίθμους, που έχουν τιμές 77.14% (RF), 71.5839 % (J48), 72.72% (KNN, N=3). Παρόμοια η κατάσταση και για το Root relative squared error, όπου ο NB έχει τιμή 84,78%, με την μεγαλύτερη τιμή να εμφανίζεται στον KNN για N=3 με 93.09 %.

ΝΑΪΒΕ BAYES

Πίνακας 74 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes)

Metric	Value
Correctly Classified Instances	86
Incorrectly Classified Instances	22
Kappa statistic	0.5926
Mean absolute error	0.3035
Root mean squared error	0.3936
Relative absolute error	60.6453 %
Root relative squared error	78.6545 %
Total Number of Instances	108

Πίνακας 75 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.759	0.167	0.820	0.759	0.788	0.594	0.850	0.824
1	0.833	0.241	0.776	0.833	0.804	0.594	0.850	0.821
Weighted Avg.	0.796	0.204	0.798	0.796	0.796	0.594	0.850	0.822

RANDOM FOREST

Πίνακας 76 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Metric	Value
Correctly Classified Instances	78
Incorrectly Classified Instances	30
Kappa statistic	0.4444
Mean absolute error	0.386
Root mean squared error	0.437
Relative absolute error	77.14 %
Root relative squared error	87.3384 %
Total Number of Instances	108

Πίνακας 77 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.759	0.315	0.707	0.759	0.732	0.446	0.783	0.736
1	0.685	0.241	0.740	0.685	0.712	0.446	0.783	0.818
Weighted Avg.	0.722	0.278	0.723	0.722	0.722	0.446	0.783	0.777

J48

Πίνακας 78 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	78
Incorrectly Classified Instances	30
Kappa statistic	0.4444
Mean absolute error	0.3582
Root mean squared error	0.4588
Relative absolute error	71.5839 %
Root relative squared error	91.688 %
Total Number of Instances	108

Πίνακας 79 Σταθμισμένες μετρήσεις επιδόσεων (J48)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.611	0.167	0.786	0.611	0.688	0.456	0.727	0.678
1	0.833	0.389	0.682	0.833	0.750	0.456	0.727	0.703
Weighted Avg.	0.722	0.278	0.734	0.722	0.719	0.456	0.727	0.690

KNN

N=3

Πίνακας 80 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value
Correctly Classified Instances	73
Incorrectly Classified Instances	35
Kappa statistic	0.3519
Mean absolute error	0.3639
Root mean squared error	0.4657
Relative absolute error	72.7245 %
Root relative squared error	93.06 %
Total Number of Instances	108

Πίνακας 81 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.704	0.352	0.667	0.704	0.685	0.352	0.725	0.685
1	0.648	0.296	0.686	0.648	0.667	0.352	0.725	0.707
Weighted Avg.	0.676	0.324	0.676	0.676	0.676	0.352	0.725	0.696

KNN

N=4

Πίνακας 82 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Metric	Value
Correctly Classified Instances	74
Incorrectly Classified Instances	34

Kappa statistic	0.3704
Mean absolute error	0.3692
Root mean squared error	0.4642
Relative absolute error	73.7829 %
Root relative squared error	92.7743 %
Total Number of Instances	108

Πίνακας 83 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.778	0.407	0.656	0.778	0.712	0.377	0.731	0.683
1	0.593	0.222	0.727	0.593	0.653	0.377	0.731	0.701
Weighted Avg.	0.685	0.315	0.692	0.685	0.682	0.377	0.731	0.692

KNN

N=5

Πίνακας 84 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Correctly Classified Instances	78
Incorrectly Classified Instances	30
Kappa statistic	0.4444
Mean absolute error	0.3628
Root mean squared error	0.4438
Relative absolute error	72.5089 %
Root relative squared error	88.696 %
Total Number of Instances	108

Πίνακας 85 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.722	0.278	0.722	0.722	0.722	0.444	0.765	0.724
1	0.722	0.278	0.722	0.722	0.722	0.444	0.765	0.735
Weighted Avg.	0.722	0.278	0.722	0.722	0.722	0.444	0.765	0.730

Σύνδρομο Dressler

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 150 (MaxCount: 75) από τις αρχικές 1625. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 17 χαρακτηριστικά εισόδου τα οποία όλα έχουν κάτω από 20% missing data και ένα εξόδου (Dressler). Στο συγκεκριμένο χαρακτηριστικό εξόδου, παρατηρούμε αρκετά μεγάλη διαφορά της απόδοσης του αλγορίθμου NB σε σχέση με τους υπολοίπους. Συγκεκριμένα, τα Correctly Classified Instances είναι 103, με τον RF να έχει 92, τον J48 81 και τον KNN να έχει 97, 94, 93 για N=3,4,5 αντίστοιχα. Συνεπώς, και οι τιμές του NB είναι αρκετά πιο ικανοποιητικές από τους άλλους αλγορίθμους, αφού έχει τιμές για το Kappa statistic 0.37, για το Mean absolute error 0.36, 0.4597 για το Root mean squared error και 73.80 για το Relative absolute error. Πολύ κακή απόδοση για το συγκεκριμένο χαρακτηριστικό, έχει ο J48 αφού ενδεικτικά έχει τιμές για το kappa statistic μόλις 0.08, και τιμές για το realtive absolute error 93.43% και για το Root relative squared error 104.90%.

Naïve Bayes

Πίνακας 86 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes)

Metric	Value
Correctly Classified Instances	103
Incorrectly Classified Instances	47
Kappa statistic	0.3733
Mean absolute error	0.3692
Root mean squared error	0.4597
Relative absolute error	73.8027%
Root relative squared error	91.8945%
Total Number of Instances	150

Πίνακας 87 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.789	0.281	0.738	0.789	0.763	0.510	0.817	0.782
1	0.719	0.211	0.774	0.719	0.745	0.510	0.817	0.845
Weighted Avg.	0.754	0.246	0.756	0.754	0.754	0.510	0.817	0.813

RANDOM FOREST

Πίνακας 88 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Metric	Value
Correctly Classified Instances	92
Incorrectly Classified Instances	58
Kappa statistic	0.2267
Mean absolute error	0.4532
Root mean squared error	0.4992
Relative absolute error	90.5872 %
Root relative squared error	99.7873 %
Total Number of Instances	150

Πίνακας 89 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.640	0.413	0.608	0.640	0.623	0.227	0.632	0.629
1	0.587	0.360	0.620	0.587	0.603	0.227	0.632	0.594
Weighted Avg.	0.613	0.387	0.614	0.613	0.613	0.227	0.632	0.611

J48

Πίνακας 90 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	81
Incorrectly Classified Instances	69
Kappa statistic	0.08
Mean absolute error	0.4674
Root mean squared error	0.5248
Relative absolute error	93.4358%
Root relative squared error	104.9073%
Total Number of Instances	150

Πίνακας 91 Σταθμισμένες μετρήσεις επιδόσεων (J48)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.467	0.387	0.547	0.467	0.504	0.081	0.562	0.529
1	0.613	0.533	0.535	0.613	0.571	0.081	0.562	0.567
Weighted Avg.	0.540	0.460	0.541	0.540	0.538	0.081	0.562	0.548

KNN

N=3

Πίνακας 92 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value
Correctly Classified Instances	97
Incorrectly Classified Instances	53
Kappa statistic	0.2933
Mean absolute error	0.4562
Root mean squared error	0.5094
Relative absolute error	91.2009 %
Root relative squared error	101.8296 %
Total Number of Instances	150

Πίνακας 93 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.600	0.307	0.662	0.600	0.629	0.295	0.618	0.600
1	0.693	0.400	0.634	0.693	0.662	0.295	0.618	0.571
Weighted Avg.	0.647	0.353	0.648	0.647	0.646	0.295	0.618	0.585

KNN

N=4

Πίνακας 94 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Metric	Value
Correctly Classified Instances	94
Incorrectly Classified Instances	56
Kappa statistic	0.2533
Mean absolute error	0.4652
Root mean squared error	0.5138
Relative absolute error	93.0035 %
Root relative squared error	102.7157%
Total Number of Instances	150

Πίνακας 95 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.627	0.373	0.627	0.627	0.627	0.253	0.612	0.592
1	0.627	0.373	0.627	0.627	0.627	0.253	0.613	0.567
	0.627	0.373	0.627	0.627	0.627	0.253	0.613	0.579

Weighted Avg.								
---------------	--	--	--	--	--	--	--	--

KNN

N=5

Πίνακας 96 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Correctly Classified Instances	93
Incorrectly Classified Instances	57
Kappa statistic	0.24
Mean absolute error	0.464
Root mean squared error	0.507
Relative absolute error	92.7555%
Root relative squared error	101.3387%
Total Number of Instances	150

Πίνακας 97 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.573	0.333	0.632	0.573	0.601	0.241	0.616	0.600
1	0.667	0.427	0.610	0.667	0.637	0.241	0.616	0.578
Weighted Avg.	0.620	0.380	0.621	0.620	0.619	0.241	0.616	0.589

Χρόνια καρδιακή ανεπάρκεια (ZSN)

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 788 (MaxCount: 394) από τις αρχικές 1306. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 11 χαρακτηριστικά εισόδου τα οποία όλα έχουν κάτω από 20% missing data και ένα εξόδου (ZSN). Σε ένα αρκετά μεγάλο σύνολο περιπτώσεων (788), παρατηρούμε ότι ο αλγόριθμος NB έχει αρκετές σωστές κατηγοριοποιήσεις (512). Αρκετά καλή είναι και η απόδοση του στο kappa statistic 0.29 ενώ σχετικά καλές του είναι και οι τιμές για τα Root mean squared error (0.41) και Root mean squared error (0.45). Ο συγκεκριμένος αλγόριθμος όμως έχει αρκετά μεγάλες τιμές για το Relative absolute error (82.63%) και για Relative absolute error (91.27%). Από εκεί και πέρα, παρόμοιες είναι οι επιδόσεις των αλγορίθμων RF και J48, όπου έχουν ίδιους αριθμούς για τα Correctly Classified Instances (482). Αρκετά κοντά είναι και οι αριθμοί και για τις υπόλοιπες μετρήσεις, με εξαίρεση το Root relative squared error όπου ο J48 έχει τιμή 95.92%, αισθητό πιο μικρό από το 107.92% που εμφανίζει ο RF. Λίγο πιο υψηλές είναι οι μετρήσεις για τον KNN που έχει Correctly Classified Instances 497, 496, 494 για N=3,4,5 αντίστοιχα. Για τις υπόλοιπες μετρήσεις του KNN και για τις 3 τιμές του N είναι αρκετά όμοιες.

Naïve Bayes

Πίνακας 98 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes)

Metric	Value
Correctly Classified Instances	512
Incorrectly Classified Instances	276
Kappa statistic	0.2995
Mean absolute error	0.4132
Root mean squared error	0.4564
Relative absolute error	82.634%
Relative absolute error	91.277%
Total Number of Instances	788

Πίνακας 99 Σταθμισμένες μετρήσεις επιδόσεων (Naïve Bayes)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.769	0.470	0.621	0.769	0.687	0.308	0.719	0.712
1	0.530	0.231	0.697	0.530	0.602	0.308	0.719	0.748
Weighted Avg.	0.650	0.350	0.659	0.650	0.645	0.308	0.719	0.730

RANDOM FOREST

Πίνακας 100 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Metric	Value
Correctly Classified Instances	482
Incorrectly Classified Instances	306
Kappa statistic	0.2234
Mean absolute error	0.4067
Root mean squared error	0.5397
Relative absolute error	81.344%
Root relative squared error	107.9294%
Total Number of Instances	788

Πίνακας 101 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.690	0.467	0.596	0.690	0.640	0.226	0.648	0.606
1	0.533	0.310	0.633	0.533	0.579	0.226	0.648	0.661
Weighted Avg.	0.612	0.388	0.615	0.612	0.609	0.226	0.648	0.634

J48

Πίνακας 102 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	482
Incorrectly Classified Instances	306
Kappa statistic	0.2234
Mean absolute error	0.436
Root mean squared error	0.4796
Relative absolute error	87.2063%
Root relative squared error	95.9214%
Total Number of Instances	788

Πίνακας 103 Σταθμισμένες μετρήσεις επιδόσεων (J48)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.736	0.513	0.589	0.736	0.655	0.231	0.627	0.575
1	0.487	0.264	0.649	0.487	0.557	0.231	0.627	0.669
Weighted Avg.	0.612	0.388	0.619	0.612	0.606	0.231	0.627	0.622

KNN

N=3

Πίνακας 104 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value
Correctly Classified Instances	497
Incorrectly Classified Instances	291
Kappa statistic	0.2614
Mean absolute error	0.4007
Root mean squared error	0.5006
Relative absolute error	80.143%
Root relative squared error	100.1207%
Total Number of Instances	788

Πίνακας 105 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.683	0.421	0.618	0.683	0.649	0.263	0.679	0.653
1	0.579	0.317	0.646	0.579	0.610	0.263	0.679	0.658
Weighted Avg.	0.631	0.369	0.632	0.631	0.630	0.263	0.679	0.655

KNN

N=4

Πίνακας 106 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Metric	Value
Correctly Classified Instances	496
Incorrectly Classified Instances	292
Kappa statistic	0.2589
Mean absolute error	0.4064
Root mean squared error	0.4903
Relative absolute error	81.2843%
Root relative squared error	98.0495%
Total Number of Instances	788

Πίνακας 107 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.708	0.449	0.612	0.708	0.656	0.262	0.681	0.646
1	0.551	0.292	0.654	0.551	0.598	0.262	0.681	0.671
	0.629	0.371	0.633	0.629	0.627	0.262	0.681	0.659

Weighted Avg.								
---------------	--	--	--	--	--	--	--	--

KNN

N=5

Πίνακας 108 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Correctly Classified Instances	494
Incorrectly Classified Instances	294
Kappa statistic	0.2538
Mean absolute error	0.4078
Root mean squared error	0.486
Relative absolute error	81.5554%
Root relative squared error	97.2004%
Total Number of Instances	788

Πίνακας 109 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.660	0.406	0.619	0.660	0.639	0.254	0.683	0.663
1	0.594	0.340	0.636	0.594	0.614	0.254	0.683	0.675
Weighted Avg.	0.627	0.373	0.627	0.627	0.626	0.254	0.683	0.669

Υποτροπή του εμφράγματος του μυοκαρδίου (REC_IM)

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 318 (MaxCount: 159) από τις αρχικές 1541. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 19 χαρακτηριστικά εισόδου τα οποία όλα έχουν κάτω από 20% missing data και ένα εξόδου (REC_IM). Τις περισσότερες σωστές ταξινομήσεις τις έχει ο NB με 208, και στην συνέχεια είναι ο RF (196), ο KNN για N=5 και N=4 με 197 και 193, ενώ για N=3 έχει 189 σωστές κατηγοριοποιήσεις. Ο J48 έχει τις λιγότερες με 186. Συνεπώς ίδια είναι και η εικόνα για το kappa statistic με τον NB να έχει 0.3, ο RF 0.23, ενώ ταυτόχρονα ο KNN εμφανίζει την μεγαλύτερη τιμή του για N=5 (0.23). Για την μέτρηση του Mean absolute error, η μικρότερη τιμή εμφανίζεται στον NB 0.39, και έπειτα όλοι οι αλγόριθμοι έχουν παρόμοια μέτρηση με τιμές μεταξύ 0.45-0.47. Εμφανή διαφορά υπάρχει και για την μέτρηση Relative absolute error, που ο NB έχει τιμή 79.84% με την αμέσως επόμενη να είναι η 91.13 του j48.

Naïve Bayes

Πίνακας 110 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naïve Bayes)

Metric	Value
Correctly Classified Instances	208
Incorrectly Classified Instances	110
Kappa statistic	0.3082
Mean absolute error	0.3992
Root mean squared error	0.4837
Relative absolute error	79.8432%
Root relative squared error	96.7438%
Total Number of Instances	318

Πίνακας 111 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.704	0.396	0.640	0.704	0.671	0.310	0.698	0.681
1	0.604	0.296	0.671	0.604	0.636	0.310	0.698	0.705
Weighted Avg.	0.654	0.346	0.656	0.654	0.653	0.310	0.698	0.693

RANDOM FOREST

Πίνακας 112 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Metric	Value
Correctly Classified Instances	196
Incorrectly Classified	122

Instances	
Kappa statistic	0.2327
Mean absolute error	0.4587
Root mean squared error	0.491
Relative absolute error	91.735%
Root relative squared error	98.1878%
Total Number of Instances	318

Πίνακας 113 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.623	0.390	0.615	0.623	0.619	0.233	0.640	0.587
1	0.610	0.377	0.618	0.610	0.614	0.233	0.640	0.654
Weighted Avg.	0.616	0.384	0.616	0.616	0.616	0.233	0.640	0.621

J48

Πίνακας 114 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	186
Incorrectly Classified Instances	132
Kappa statistic	0.1698
Mean absolute error	0.4557
Root mean squared error	0.5165
Relative absolute error	91.1383%
Root relative squared error	103.2984%
Total Number of Instances	318

Πίνακας 115 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.629	0.459	0.578	0.629	0.602	0.170	0.593	0.550
1	0.541	0.371	0.593	0.541	0.566	0.170	0.593	0.611
Weighted Avg.	0.585	0.415	0.586	0.585	0.584	0.170	0.593	0.580

KNN

N=3

Πίνακας 116 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value
Correctly Classified Instances	189
Incorrectly Classified Instances	129
Kappa statistic	0.1887
Mean absolute error	0.4669
Root mean squared error	0.5116
Relative absolute error	93.384%
Root relative squared error	102.3146%
Total Number of Instances	318

Πίνακας 117 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.698	0.509	0.578	0.698	0.632	0.193	0.576	0.532
1	0.491	0.302	0.619	0.491	0.547	0.193	0.576	0.620
Weighted Avg.	0.594	0.406	0.599	0.594	0.590	0.193	0.576	0.576

KNN

N=4

Πίνακας 118 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Metric	Value
Correctly Classified Instances	193
Incorrectly Classified Instances	125
Kappa statistic	0.2138
Mean Absolute Error	0.4707
Root Mean Squared Error	0.5072
Relative Absolute Error	94.1358 %
Root Relative Squared Error	101.4451 %
Total Number of Instances	318

Πίνακας 119 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.736	0.522	0.585	0.736	0.652	0.221	0.579	0.536
1	0.478	0.264	0.644	0.478	0.549	0.221	0.579	0.612
Weighted Avg.	0.607	0.393	0.615	0.607	0.600	0.221	0.579	0.574

KNN

N=5

Πίνακας 120 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Correctly Classified Instances	197
Incorrectly Classified Instances	121
Kappa statistic	0.239
Mean Absolute Error	0.4689
Root Mean Squared Error	0.5014
Relative Absolute Error	93.7818 %
Root Relative Squared Error	100.2844 %
Total Number of Instances	318

Πίνακας 121 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.742	0.503	0.596	0.742	0.661	0.247	0.591	0.537
1	0.497	0.258	0.658	0.497	0.566	0.247	0.591	0.628
Weighted Avg.	0.619	0.381	0.627	0.619	0.614	0.247	0.591	0.583

Στηθάγχη μετά το έμφραγμα (P_IM_STEN)

Οι περιπτώσεις μετά την εφαρμογή του φίλτρου SpreadSubSample είναι στις 296 (MaxCount: 148) από τις αρχικές 1552. Έπειτα μετά την εφαρμογή του φίλτρου AttributeSelection απομένουν 23 χαρακτηριστικά εισόδου τα οποία όλα έχουν κάτω από 20% missing data και ένα εξόδο (P_IM_STEN). Στο κομμάτι των αποτελεσμάτων, οι πιο πολλές σωστές κατηγοριοποιήσεις εμφανίζονται και πάλι στον NB με 184. Ο KNN έρχεται δεύτερος για N=5 και N=4 με 160. Για N=3 η τιμή είναι 158. Ένα λιγότερο έχει ο J48(157). Παράλληλα τα λιγότερα τα έχει ο RF με 151, ο οποίος έχει το μικρότερο kappa statistic 0.02. Η μεγαλύτερη τιμή για την συγκεκριμένη μέτρηση είναι 0.24 και την πετυχαίνει ο MB. Οι τιμές 0.06 και 0.08 που έχουν οι KNN και J48 δεν είναι ικανοποιητικές. Για το Mean Absolute Error η μικρότερη τιμή είναι η 0.41 του NB και έπειτα οι τιμές κυμαίνονται από 0.46(KNN,N=3) έως 0.49(RF). Το μικρότερο Root Mean Squared Error το συναντάμε στον NB (0.48). Ακολουθεί ο KNN για N=5 και N=4 με τιμές 0.52 και 0.53. Με την ίδια απόδοση βλέπουμε και τον J48(0.53) ενώ το μεγαλύτερο Root Mean Squared Error το έχει ο RF με μέτρηση 0.62. Ο Naive Bayes έχει τις μικρότερες τιμές και στις υπόλοιπες μετρήσεις σφαλμάτων. Συγκεκριμένα για το Relative Absolute Error έχει τιμή 82.31% και για το Root Relative Squared Error έχει τιμή 96.02%.

Naïve Bayes

Πίνακας 122 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Naive Bayes)

Metric	Value
Correctly Classified Instances	184
Incorrectly Classified Instances	112
Kappa statistic	0.2432
Mean Absolute Error	0.4116
Root Mean Squared Error	0.4802
Relative Absolute Error	82.3104 %
Root Relative Squared Error	96.0295 %
Total Number of Instances	296

Πίνακας 123 Σταθμισμένες μετρήσεις επιδόσεων (Naive Bayes)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.574	0.331	0.634	0.574	0.603	0.244	0.694	0.699
1	0.669	0.426	0.611	0.669	0.639	0.244	0.694	0.691
Weighted Avg.	0.622	0.378	0.623	0.622	0.621	0.244	0.694	0.695

RANDOM FOREST

Πίνακας 124 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (Random Forest)

Metric	Value
Correctly Classified Instances	151
Incorrectly Classified Instances	145
Kappa statistic	0.0203
Mean Absolute Error	0.4923
Root Mean Squared Error	0.627
Relative Absolute Error	98.4572 %
Root Relative Squared Error	125.3908 %
Total Number of Instances	296

Πίνακας 125 Σταθμισμένες μετρήσεις επιδόσεων (Random Forest)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.514	0.493	0.510	0.514	0.512	0.020	0.518	0.493
1	0.507	0.486	0.510	0.507	0.508	0.020	0.518	0.552
Weighted Avg.	0.510	0.490	0.510	0.510	0.510	0.020	0.518	0.523

J48

Πίνακας 126 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (J48)

Metric	Value
Correctly Classified Instances	157
Incorrectly Classified Instances	139
Kappa statistic	0.0608
Mean Absolute Error	0.4844
Root Mean Squared Error	0.5396
Relative Absolute Error	96.8675 %
Root Relative Squared Error	107.9149 %
Total Number of Instances	296

Πίνακας 127 Σταθμισμένες μετρήσεις επιδόσεων (J48)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.446	0.385	0.537	0.446	0.487	0.062	0.546	0.543
1	0.615	0.554	0.526	0.615	0.567	0.062	0.546	0.522
Weighted Avg.	0.530	0.470	0.531	0.530	0.527	0.062	0.546	0.533

KNN

N=3

Πίνακας 128 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=3)

Metric	Value
Correctly Classified Instances	158
Incorrectly Classified Instances	138
Kappa statistic	0.0676
Mean Absolute Error	0.4611
Root Mean Squared Error	0.5454
Relative Absolute Error	92.2155 %
Root Relative Squared Error	109.0835 %
Total Number of Instances	296

Πίνακας 129 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=3)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.446	0.378	0.541	0.446	0.489	0.069	0.562	0.583
1	0.622	0.554	0.529	0.622	0.571	0.069	0.562	0.540
Weighted Avg.	0.534	0.466	0.535	0.534	0.530	0.069	0.562	0.561

KNN

N=4

Πίνακας 130 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=4)

Metric	Value
Correctly Classified Instances	160
Incorrectly Classified Instances	136
Kappa statistic	0.0811
Mean Absolute Error	0.4712
Root Mean Squared Error	0.5307
Relative Absolute Error	94.2279 %
Root Relative Squared Error	106.1348 %
Total Number of Instances	296

Πίνακας 131 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=4)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
-------	---------	---------	-----------	--------	-----------	-----	----------	----------

0	0.520	0.439	0.542	0.520	0.531	0.081	0.569	0.574
1	0.561	0.480	0.539	0.561	0.550	0.081	0.569	0.561
Weighted Avg.	0.541	0.459	0.541	0.541	0.540	0.081	0.569	0.568

KNN

N=5

Πίνακας 132 Σύνοψη μετρήσεων αξιολόγησης μοντέλου (KNN, N=5)

Metric	Value
Correctly Classified Instances	160
Incorrectly Classified Instances	136
Kappa statistic	0.0811
Mean Absolute Error	0.4705
Root Mean Squared Error	0.5209
Relative Absolute Error	94.0925 %
Root Relative Squared Error	104.1706 %
Total Number of Instances	296

Πίνακας 133 Σταθμισμένες μετρήσεις επιδόσεων (KNN, N=5)

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
0	0.405	0.324	0.556	0.405	0.469	0.084	0.563	0.573
1	0.676	0.595	0.532	0.676	0.595	0.084	0.563	0.556
Weighted Avg.	0.541	0.459	0.544	0.541	0.532	0.084	0.563	0.564

ΣΥΜΠΕΡΑΣΜΑΤΑ

Συνοψίζοντας η παρούσα διπλωματική εργασία είχε ως σκοπό τη μελέτη της πρόβλεψης της ισχαιμίας του μυοκάρδιου χρησιμοποιώντας διάφορους αλγόριθμους μηχανικής μάθησης μέσω του εργαλείου Weka. Το Weka αποτελεί ένα εργαλείο για την ανάλυση δεδομένων και την πρόβλεψη, παρέχοντας γραφικά περιβάλλοντα χρήστη για την ευκολία πρόσβασης. Η μηχανική μάθηση αποτελεί ένα αναπόσπαστο εργαλείο για την πρόβλεψη παθήσεων στον ιατρικό τομέα και αποτελεί βασικό γρανάτζι για μελέτες με μεγάλο όγκο δεδομένων. Στη σύγχρονη βιβλιογραφία υπάρχουν ποικίλες μελέτες οι οποίες χρησιμοποιούν πολύπλοκους αλγόριθμους για την εξέταση καρδιοαγγειακών παθήσεων (Shu et al., 2020; Yang et al 2022).

Σε αυτή την εργασία εξετάστηκαν η κολπική μαρμαρυγή, η υπερκοιλιακή ταχυκαρδία, η κοιλιακή ταχυκαρδία και μαρμαρυγή, ο αποκλεισμός AV τρίτου βαθμού, το πνευμονικό οίδημα, η ρήξη του μυοκαρδίου, το σύνδρομο Dressler, η χρόνια καρδιακή ανεπάρκεια, η υποτροπή του εμφράγματος του μυοκαρδίου, η στηθάγχη μετά το έμφραγμα και το θανατηφόρο αποτέλεσμα. Στα πλαίσια της αξιολόγησης των δυαδικών μοντέλων ταξινόμησης για την πρόβλεψη καρδιοαγγειακών παθήσεων, εξετάστηκαν τα μοντέλα Naïve Bayes, Random Forest, J48, και KNN (N=3, 4, 5). Η αξιολόγηση βασίστηκε σε ποσότητες όπως το ROC (Receiver Operating Characteristic), προκειμένου να εκτιμηθεί η ικανότητα των μοντέλων να ορθώς εντοπίζουν θετικές περιπτώσεις και να ελαχιστοποιούν τα ψευδώς θετικά, επιδιώκοντας ένα ισορροπημένο συνδυασμό ακρίβειας και ανάκλησης.

Τα αποτελέσματα δείχνουν ότι, για πολλές καρδιακές καταστάσεις, το μοντέλο Naïve Bayes επιδεικνύει υψηλή απόδοση σύμφωνα με τις τιμές ROC. Συγκεκριμένα, για παθήσεις όπως η κολπική μαρμαρυγή, η υπερκοιλιακή ταχυκαρδία, και η ρήξη του μυοκαρδίου, το Naïve Bayes ξεχωρίζει με υψηλή αξιοπιστία (ROC=0.752, 0.836, 0.850 αντίστοιχα). Επιπλέον, η εφαρμογή διάφορων μοντέλων επισημαίνει την ευελιξία της μηχανικής μάθησης, αναδεικνύοντας την αποτελεσματικότητα του Naïve Bayes σε πολλαπλές παθήσεις, ενώ το Random Forest και το KNN φαίνεται να παρουσιάζουν αξιολογες επιδόσεις σε κάποιες περιπτώσεις. Μπορούμε εύκολα να διαπιστώσουμε ότι ο Naïve Bayes έχει την καλύτερη απόδοση συγκριτικά με τους άλλους αλγόριθμους που εξετάστηκαν. Από άποψη σωστών κατηγοριοποιήσεων, σταθερά ο Naïve Bayes εμφανίζει τις περισσότερες. Εξάιρεση αποτελεί το χαρακτηριστικό εξόδου της υπερκοιλιακής ταχυκαρδίας, όπου ο Random Forest εμφανίζει μόλις μια παραπάνω σωστή κατηγοριοποίηση, αλλά και πάλι ο αριθμός των συνολικών περιπτώσεων που εξετάστηκαν στο συγκεκριμένο χαρακτηριστικό είναι αρκετά μικρός (40). Ένα δεύτερο χαρακτηριστικό που ο Random Forest εμφανίζει περισσότερες σωστές κατηγοριοποιήσεις είναι αυτό της κοιλιακής μαρμαρυγής, εμφανίζοντας 7 παραπάνω σε ένα συνολικό δείγμα 142 περιπτώσεων. Παράλληλα παρατηρούμε ότι ο Naïve Bayes επιτυγχάνει αρκετές φορές και υψηλά επίπεδα στην μέτρηση του kappa statistic, που ξεπερνούν μερικές φορές την τιμή 0.4. Συγκεκριμένα για την κοιλιακή ταχυκαρδία εμφανίζει την τιμή 0.64, για την κοιλιακή μαρμαρυγή 0.46, για τον αποκλεισμό AV τρίτου βαθμού 0.52, για το πνευμονικό οίδημα 0.46, για την ρήξη του μυοκαρδίου 0.59. Μια ακόμα αξιολογή παρατήρηση σχετικά με την καλή επίδοση του Naïve Bayes είναι πως για το χαρακτηριστικό εξόδου Dressler, ενώ οι υπόλοιποι αλγόριθμοι εμφανίζουν αρκετά χαμηλά αποτελέσματα (με χειρότερη επίδοση τον J48 που αγγίζει τα όρια της τυχαιότητας με kappa statistic 0.08), ο Naïve Bayes έχει επίδοση 0.3733 για το kappa statistic ενώ το ROC area είναι 0.817. Ένα ακόμα γεγονός που αξίζει να σημειωθεί, είναι πως ενώ για τον αλγόριθμο KNN, όσο αυξάνεται το N, οι επιδόσεις του είτε

μειώνονται, είτε παραμένουν σταθερές, αυτό δεν ισχύει για το χαρακτηριστικό εξόδου της υποτροπή εμφράγματος του μυοκαρδίου. Για $N=3$ οι σωστές κατηγοριοποιήσεις είναι 189, όμως ο αριθμός αυτός αυξάνεται για $N=4$ (193) και έχει ακόμα μεγαλύτερη τιμή για $N=5$ (197). Υπάρχουν βέβαια και οι συνολικές κακές επιδόσεις για ένα συγκεκριμένο χαρακτηριστικό. Για το χαρακτηριστικό της στηθάγχης μετά το έμφραγμα, κανένας αλγόριθμός δεν κατάφερε να έχει κάποια ικανοποιητικά αποτελέσματα. Όλοι τους εμφανίζουν τιμές για το kappa statistic σταθερά < 0.1 και τιμές για το Root Relative Squared Error $> 100\%$, με μόνη εξαίρεση τον Naïve Bayes, που με τιμές kappa statistic = 0.2432, Relative Absolute Error = 82.31 %, Root Relative Squared Error = 96.02 % και ROC area = 0.6, δεν μπορεί να θεωρηθεί ότι έχει ικανοποιητική απόδοση. Τέλος από την άλλη πλευρά, υπάρχουν και τα χαρακτηριστικά εξόδου για τα οποία συνολικά όλοι οι αλγόριθμοι είχαν μια πολύ καλή απόδοση. Συγκεκριμένα, υπάρχει το χαρακτηριστικό του πνευμονικού οιδήματος (OTEK_LANC), όπου όλοι τους παρουσίασαν αρκετά Correctly Classified Instances και ταυτόχρονα εμφάνισαν καλές τιμές στις μετρήσεις για το kappa statistic και για τη ROC area. Και αυτό μάλιστα συμβαίνει εξετάζοντας αρκετά μεγάλο δείγμα (318 instances). Υπάρχουν βέβαια και τα χαρακτηριστικά εξόδου του αποκλεισμού AV 3ου βαθμού και αυτό της ρήξης του μυοκαρδίου, όπου και εκεί όλοι οι αλγόριθμοι που εξετάστηκαν στην παρούσα εργασία, εμφάνισαν αρκετά ικανοποιητικά αποτελέσματα, με μόνη διαφορά όμως πως έτρεχαν για αρκετές λιγότερες περιπτώσεις. (A_V_Block Total Number of Instances=114, ρήξη του μυοκαρδίου (RAZRIV) Total Number of Instances=108).

Τα συμπεράσματα αυτής της διπλωματικής εργασίας παρέχουν σημαντικές πληροφορίες σχετικά με τη χρήση διαφόρων αλγορίθμων μηχανικής μάθησης για την πρόβλεψη της ισχαιμίας του μυοκαρδίου. Ενώ ο αλγόριθμος Naïve Bayes φαίνεται να εμφανίζει την καλύτερη επίδοση σε όλους τους προβλεπτικούς παράγοντες που εξετάστηκαν, δεν μπορούμε να καταλήξουμε σε συμπέρασμα ότι είναι επικρατέστερος για όλες τις μελέτες του μυοκαρδίου, καθώς η επίδοση κάθε αλγορίθμου εξαρτάται από τα χαρακτηριστικά των δεδομένων και τους παράγοντες που εξετάζονται. Σημαντικό είναι επίσης το συμπέρασμα που αναφέρει ότι οι Kotu et al. και οι Desai et al. έχουν παρουσιάσει σε προηγούμενες μελέτες ότι άλλοι αλγόριθμοι, όπως ο SVM και ο Rotation Forest, αντίστοιχα, εμφανίζουν καλύτερη ακρίβεια για συγκεκριμένες εφαρμογές στον τομέα της καρδιολογίας. Αυτό υπογραμμίζει τη σημασία της προσαρμογής του αλγορίθμου στις συγκεκριμένες ανάγκες και χαρακτηριστικά κάθε μελέτης. Επιπλέον, η αναφορά στη συλλογή και επεξεργασία των δεδομένων ως κομβικό ρόλο στην απόδοση του κάθε αλγορίθμου είναι σημαντική. Αυτό υπογραμμίζει τη σημασία της προσεκτικής επιλογής και προεπεξεργασίας των δεδομένων πριν από την εφαρμογή των αλγορίθμων. Τέλος, η εργασία αναδεικνύει τη σημασία της κατανόησης και χρήσης διαφόρων εργαλείων και αλγορίθμων στον τομέα της επιστήμης των δεδομένων και της μηχανικής μάθησης. Επίσης, επισημαίνει τις ιδιαιτερότητες που μπορεί να επηρεάσουν τα αποτελέσματα της μελέτης, όπως οι διαφορετικές εφαρμογές, οι διαφορετικές παθήσεις και τα διαφορετικά σενάρια. Συνολικά, η διπλωματική αυτή εργασία προσφέρει σημαντικές συνεισφορές στον τομέα της καρδιολογίας και της μηχανικής μάθησης, υπογραμμίζοντας τη σημασία της προσαρμογής των αλγορίθμων στις συγκεκριμένες απαιτήσεις κάθε μελέτης και τη σημασία της προεπεξεργασίας των δεδομένων.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Crawford, M. (2009). Σύγχρονη καρδιολογία. Αθήνα: Πασχαλίδης.

Dey, D., Gaur, S., Ovrehus, K. A., Slomka, P. J., Betancur, J., Goeller, M., ... & Norgaard, B. L. (2018). Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: a multicentre study. *European radiology*, 28, 2655-2664.

Golovenkin, S.E., Shulman, V.A., Rossiev, D.A., Shesternya, P.A., Nikulina, S.Yu., Orlova, Yu.V., and Voino-Yasenetsky, V.F.. (2020). Myocardial infarction complications. UCI Machine Learning Repository. <https://doi.org/10.24432/C53P5M>.

Hae, H., Kang, S. J., Kim, W. J., Choi, S. Y., Lee, J. G., Bae, Y., ... & Park, S. J. (2018). Machine learning assessment of myocardial ischemia using angiography: Development and retrospective validation. *PLoS Medicine*, 15(11), e1002693.

Han, D., Lee, J. H., Rizvi, A., Gransar, H., Baskaran, L., Schulman-Marcus, J., ... & Min, J. K. (2018). Incremental role of resting myocardial computed tomography perfusion for predicting physiologically significant coronary artery disease: a machine learning approach. *Journal of Nuclear Cardiology*, 25, 223-233.

Kotu, L. P., Engan, K., Borhani, R., Katsaggelos, A. K., Ørn, S., Woie, L., & Eftestøl, T. (2015). Cardiac magnetic resonance image-based classification of the risk of arrhythmias in post-myocardial infarction patients. *Artificial intelligence in medicine*, 64(3), 205-215.

Lomsky, M., Gjertsson, P., Johansson, L., Richter, J., Ohlsson, M., Tout, D., ... & Edenbrandt, L. (2008). Evaluation of a decision support system for interpretation of myocardial perfusion gated SPECT. *European Journal of Nuclear Medicine and Molecular Imaging*, 35, 1523-1529.

M. Chala Beyene, Survey on prediction and analysis the occurrence of heart disease using data mining techniques **118**(8), 165–174 (2018) [Online]. Available <http://www.ijpam.eu>

Mannil, M., von Spiczak, J., Muehlematter, U. J., Thanabalasingam, A., Keller, D. I., Manka, R., & Alkadhi, H. (2019). Texture analysis of myocardial infarction in CT: comparison with visual analysis and impact of iterative reconstruction. *European journal of radiology*, 113, 245-250.

Maroko PR., Kjekshus, JK., Sobel BE.; Watanabe T., Covell, J. W., Ross J. Jr., Braunwald E. Factors influencing infarct size following experimental coronary artery occlusions. *Circulation* 1971; 43: 67-82.

Merjulah, R., & Chandra, J. (2020). An Integrated Segmentation Techniques for Myocardial Ischemia. *Pattern Recognition and Image Analysis*, 30, 530-540.

Ramalingam, V. V., Dandapath, A., & Raja, M. K. (2018). Heart disease prediction using machine learning techniques: a survey. *International Journal of Engineering & Technology*, 7(2.8), 684-687.

Ramrakha, P. and Hill, J. (2008). Oxford Handbook of Cardiology. London: Oxford University Press.

Reimer KA., Murry CE., Richard VJ. The role of neutrophils and free radicals in the ischemic-reperfused heart: why the confusion and controversy? J. Moll. Cell. Cardiol. 1989; 21: 1225-1239.

Rosen, S.D., Sharma, S. and Oakley, S.M. (2004). Cardiology. New York: Thieme editions.

Shu Z, Cui S, Zhang Y, Xu Y, Huang S, Fu L, et al. Predicting chronic myocardial ischemia using CCTA-based radiomics machine learning nomogram. J Nucl Cardiol. 2020. [https://doi.org/ 10.1007/s12350-020-02204-2](https://doi.org/10.1007/s12350-020-02204-2).

Sun, Q., Liang, C., Chen, T., Ji, B., Liu, R., Wang, L., ... & Wang, C. (2022). Early detection of myocardial ischemia in 12-lead ECG using deterministic learning and ensemble learning. *Computer Methods and Programs in Biomedicine*, 226, 107124.

Xiong, G., Kola, D., Heo, R., Elmore, K., Cho, I., & Min, J. K. (2015). Myocardial perfusion analysis in cardiac computed tomography angiographic images at rest. *Medical image analysis*, 24(1), 77-89.

Βαρδής, Π., Γουδέβενος, Ι., και Ελισάφ, Μ. (επ.) (2005). Βασική Κλινική Καρδιολογία. Αθήνα: Πασχαλίδης.

Γ.Μπαμπινιώτης: Λεξικό της νέας ελληνικής γλώσσας. Κέντρο λεξικολογίας 1998:796

Κοντόπουλος, Α. (2002). Επίτομη καρδιολογία. Θεσσαλονίκη: University Studio Press.

Κρεμαστινός, Δ. (2005). Επίτομη Κλινική Καρδιολογία. Αθήνα: Πασχαλίδης.

Μπαχαράκας, Σ. (2005). Μαθήματα στην καρδιολογία. Θεσσαλονίκη: University Studio Press.

Π.Τούτουζας: Θεραπεία της στεφανιαίας νόσου, 2η έκδοση. Παρισιάνος 1991:1-5