

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

Τμήμα Πληροφορικής



- Πτυχιακή Εργασία -

**Εφαρμογή τεχνικών μηχανικής μάθησης για την δημιουργία μοντέλων
κατηγοριοποίησης ασθενών**

Παπαθεοδώρου Μαρία Χριστίνα

Επιβλέπων: Έξαρχος Θεμιστοκλής

Κέρκυρα, 2022

Επιβλέπων

Θεμιστοκλής Έξαρχος

Βαθμίδα: Επίκουρος Καθηγητής

Ιόνιο Πανεπιστήμιο / Τμήμα Πληροφορικής

Τριμελής Επιτροπή

1.Θεμιστοκλής Έξαρχος

Βαθμίδα: Επίκουρος Καθηγητής

Ιόνιο Πανεπιστήμιο / Τμήμα Πληροφορικής

2.Σωτηροπούλου Άννα

Βαθμίδα: ΕΔΙΠ

Ιόνιο Πανεπιστήμιο / Τμήμα Πληροφορικής

3.Μυλωνάς Φοίβος

Βαθμίδα: Αναπληρωτής Καθηγητής

Ιόνιο Πανεπιστήμιο / Τμήμα Πληροφορικής

Πνευματική πρόοδος υπάρχει, εκεί που υπάρχει η αίσθηση ότι όλα είναι χάλια.
- Γέρων Παΐσιος-

Στην Οικογένεια μου.

Περίληψη | Abstract

Σκοπός: Οι καρδιαγγειακές παθήσεις είναι ένα μείζον πρόβλημα, που απασχολεί όλο τον κόσμο, με σοβαρές επιπτώσεις στην κοινωνία, την οικονομία και την ψυχολογία, του ατόμου και του κοινωνικού συνόλου. Αποτελεί την πρώτη αιτία θανάτου, μαζί με τον καρκίνο. Σκοπός αυτής της μελέτης είναι η εύρεση τεχνικών μηχανικής μάθησης, για την δημιουργία ενός μοντέλου πρόβλεψης των καρδιαγγειακών παθήσεων.

Μέθοδοι: Πλήθος ερευνών πραγματοποιούνται γύρω από την εύρεση όλων των καρδιαγγειακών παθήσεων σε πρώιμο στάδιο. Εμείς μελετήσαμε είκοσι βιβλιογραφικές αναφορές, των τελευταίων πέντε ετών, αναζητώντας τα πιο σύγχρονα συστήματα. Στην πορεία, βασισμένοι σε ένα αρχικό dataset 70.000 εγγραφών, δημιουργήσαμε έναν κανόνα, δοκιμάζοντας διάφορους αλγορίθμους ταξινόμησης, στο WEKA, αναζητώντας εκείνον με την καλύτερη απόδοση. Τα χαρακτηριστικά του dataset, αφορούσαν δεδομένα, προηγούμενων ασθενών, που επηρεάζουν τις καρδιαγγειακές παθήσεις, όπως η ηλικία, το φύλο, η πίεση και το κάπνισμα. Έπειτα από την επιλογή αλγορίθμου και την διεξαγωγή του κανόνα, δημιουργήσαμε εφαρμογή πρόβλεψης καρδιαγγειακών παθήσεων, χρησιμοποιώντας το την php, html και το XAMPP.

Αποτελέσματα: Τα αποτελέσματα ανέδειξαν πιο αποτελεσματικό αλγόριθμο τον J48 με min number of objects 200 και ποσοστό ακρίβειας 73.3576 %. Το τελικό σύστημα δοκιμάστηκε με δέκα έτοιμα δείγματα από το αρχικό dataset και ταξινόμησε σωστά τα 7 από τα 10.

Πρόλογος και Ευχαριστίες

Η προσπάθεια δημιουργίας μιας εργασίας, με θέμα την εφαρμογή τεχνικών μηχανικής μάθησης, για την δημιουργία μοντέλων κατηγοριοποίησης ασθενών, ξεκίνησε στο Ιόνιο Πανεπιστήμιο και συγκεκριμένα στο Τμήμα Πληροφορικής, το έτος 2021. Η εργασία αυτή έγινε στα πλαίσια της διπλωματικής μου, για την απόκτηση του πρώτου μου πτυχίου, πάνω στο αντικείμενο της πληροφορικής. Το πόνημα αυτό, πραγματοποιήθηκε από την Παπαθεοδώρου Μαρία Χριστίνα.

Σε αυτό το σημείο κρίνεται σκόπιμο, να ευχαριστήσω όλους όσους με βοήθησαν είτε με άμεσο είτε με έμμεσο τρόπο στην ολοκλήρωση της εργασίας μου. Ακόμα περισσότερο, του γονείς μου Παναγιώτη και Βασιλική και τα αδέρφια μου Εύη και Μιχάλη για την υπομονή και την συμπαράσταση, τους καθ' όλη τη διάρκεια της φοίτησης μου και ειδικότερα στην προσπάθεια ολοκλήρωσης τους. Κυρίως όμως θα ήθελα να ευχαριστήσω τον κύριο Θ. Έξαρχο που μου έδωσε την ευκαιρία να εργαστώ μαζί του και να αναλύσω ένα τόσο ενδιαφέρον θέμα, για τα εφόδια που μου πρόσφερε κατά την διάρκεια των μαθημάτων, αλλά και για όλη του την υποστήριξη και την καθοδήγηση του, στην συγγραφή της διπλωματικής μου εργασίας. Θέρμες ευχαριστίες οφείλω φυσικά και σε όλους τους καθηγητές με τους οποίους συνεργάστηκα όλα αυτά τα χρόνια καθώς και τους φίλους και συμφοιτητές μου, που πορευτήκαμε μαζί σε αυτό το ταξίδι.

Περιεχόμενα

Κεφάλαιο Α΄	8
Εισαγωγή	8
Κεφάλαιο Β΄	10
Περιγραφή Του Ιατρικού Προβλήματος	10
Κεφάλαιο Γ΄	14
Αντίκτυπος της καρδιακής ανεπάρκειας	14
Γ΄.1 Αριθμός Ασθενών	14
Γ΄.2 Κοινωνικό Κόστος	16
Γ΄.3 Οικονομικό Κόστος	17
Γ΄.4 Ψυχολογικό Κόστος	19
Κεφάλαιο Δ΄	22
Ανάγκη για τεχνικές DSS	22
Κεφάλαιο Ε΄	26
Βιβλιογραφική Ανασκόπηση	26
Ε΄.1 Βιβλιογραφικές Αναφορές	26
Ε΄.2 Πίνακας Σύγκρισης Βιβλιογραφικών Αναφορών	40
Ε΄.3 Αποτελέσματα Βιβλιογραφικής Ανασκόπησης	46
Κεφάλαιο ΣΤ΄	49
Επεξεργασία & Ανάλυση Δεδομένων Για Την Εφαρμογή Πρόβλεψης Καρδιακών Παθήσεων	49
ΣΤ΄.1 Συστήματα Υποστήριξης Αποφάσεων Στις Καρδιαγγειακές παθήσεις	49
ΣΤ΄.2 Περιγραφή Του Συνόλου Δεδομένων	50
ΣΤ΄.3 Data Analysis	51
Εικόνα 6.3: Εκκίνηση WEKA	53
ΣΤ΄ 3.1 Προεπεξεργασία	53
ΣΤ΄ 3.2 Επιλογή Δεδομένων	54
ΣΤ΄.4 Εφαρμογή Αλγορίθμων	60
ΣΤ΄ 4.1 Σύγκριση Αλγορίθμων	63
Κεφάλαιο Ζ΄	71
Δημιουργία Εφαρμογής	71
Κεφάλαιο Η΄	84
Συμπεράσματα	84
Κεφάλαιο Θ΄	85
Μελλοντικές Προεκτάσεις	85
Βιβλιογραφία / Δικτυογραφία	87
Συντμήσεις	91
Γλωσσάρι Ξενικών Όρων	92

Κατάλογος Σχημάτων

Σχήμα 1: Confusion Matrix.....	64
--------------------------------	----

Κατάλογος Πινάκων

B' .2 Πίνακας 1: Σύγκριση Βιβλιογραφικών Αναφορών.....	46
Στ' .2 Πίνακας 2: Περιγραφή του συνόλου δεδομένων.....	51
Στ' .4 Πίνακας 3: Αλγόριθμοι.....	62
Στ' .4.1 Πίνακας 4: Σωστών και λανθασμένων ταξινομήσεων.....	63
Στ' 4.1 Πίνακας 5: Χρόνος Εκτέλεσης Ταξινόμησης και root mean square error.....	63
ΣΤ' 4.1 Πίνακας 6: Μέτρα απόδοσης υπολογισμένα βάση confusion matrix με 10-fold cross validation.....	65
ΣΤ' 4.1 Πίνακας 7: Balance ταξινόμησης.....	65
ΣΤ' 4.1 Πίνακας 8: Number of objects j48 (Αξιολόγηση ταξινομήσεων).....	66
ΣΤ' 4.1 Πίνακας 9: Χρόνος Εκτέλεσης Ταξινόμησης & Root Mean Square Error J48.....	67
ΣΤ' 4.1 Πίνακας 10: Μέτρα απόδοσης υπολογισμένα βάση confusion matrix με 10-fold cross validation για διαφορετικά min number of objects στον J48.....	68
Στ' 4.1 Πίνακας 11: Balance ταξινόμησης για διαφορετικά min number of objects.....	68
Z' Πίνακας 12: Τυχαίες τιμές του dataset.....	82
Z' Πίνακας 13: Δέκα τυχαίες δοκιμές και αποτελέσματα	83

Κεφάλαιο Α΄

Εισαγωγή

Οι καρδιακές παθήσεις αποτελούν ένα μείζον πρόβλημα που απασχολεί εδώ και εκατοντάδες χρόνια, την ιατρική επιστήμη, σε παγκόσμια κλίμακα. Συγκεκριμένα τα ποσοστά των περιπτώσεων με καρδιακή ανεπάρκεια, αυξάνονται μέρα με την μέρα. Η πρόβλεψη της καρδιακής ανεπάρκειας είναι μια ιδιαίτερα δύσκολη διαδικασία και πολύπλοκη [2]. Ανά τακτά χρονικά διαστήματα, γίνονται ολοένα και περισσότερες προσπάθειες να μειωθούν ή ακόμα και να εξαλειφθούν τέτοιου είδους προβλήματα. Τον σκοπό αυτό εξυπηρετεί και αυτή η εργασία, που πραγματοποιήθηκε στα πλαίσια της απόκτησης του πρώτου μου διπλώματος, στον κλάδο της πληροφορικής. Με την ανάπτυξη της τεχνολογίας, η οποία εισέρχεται όλο και περισσότερο στην καθημερινότητα μας και στην ζωή, μας αναπτύσσονται πολλοί κλάδοι ένας εκ των οποίων είναι και αυτός της πληροφορικής. Σε αυτή την εργασία θα παρακολουθήσουμε τον τρόπο με τον οποίο, εφαρμογές τεχνικών μηχανικής μάθησης, κομμάτι της πληροφορικής, δημιουργούν μοντέλα για την κατηγοριοποίηση των ασθενών. Συγκεκριμένα δημιουργήσαμε μια εφαρμογή η οποία φανερώνει την ύπαρξη ή την απουσία καρδιαγγειακής νόσου. Η μηχανική μάθηση / Machine Learning (ML) αποτελεί ένα υποσύνολο της τεχνητής νοημοσύνης. Οι μηχανές αποκτούν μεγάλο όγκο δεδομένων, μέσα από βάσεις δεδομένων και εξάγουν μοτίβα, δηλαδή κανόνες. Η μέθοδος αυτή χρησιμοποιείται όλο και πιο πολύ στην ιατρική κοινότητα και ακόμα περισσότερο στον τομέα των καρδιαγγειακών παθήσεων [1].

Εμείς σε αυτή την εργασία, αρχικά ερευνήσαμε τα χαρακτηριστικά του προβλήματος και κάναμε μια σύντομη περιγραφή, γύρω από αυτό. Οι καρδιακές παθήσεις, αποτελούν μια ολόκληρη ομάδα διαταραχών, της καρδιάς και των αιμοφόρων αγγείων. Το πλήθος των καρδιαγγειακών παθήσεων είναι αρκετά εκτεταμένο και κάθε μια πάθηση ξεχωριστά έχει και τα δικά της χαρακτηριστικά. Αυτό που ισχύει αναμφισβήτητα όμως, για το πλήθος όλων των καρδιαγγειακών

παθήσεων, είναι ότι προκαλούν ιδιαίτερη ανησυχία, στον κλάδο της ιατρικής. Οι καρδιαγγειακές παθήσεις έχουν πολύπλοκη διαδικασία ανίχνευσης. Εκτός αυτού, χαρακτηριστικό των καρδιαγγειακών παθήσεων, είναι και η συννοσηρότητα, δηλαδή η εμφάνιση και άλλων παθήσεων. Βασική μελέτη, που δεν θα μπορούσε να παραληφθεί, πριν την δημιουργία ενός συστήματος πρόβλεψης, ήταν η αναζήτηση του αντίκτυπου, της καρδιαγγειακής νόσου. Για αυτόν τον λόγο αναζητήσαμε την έκταση που έχει το πρόβλημα, δηλαδή τον αριθμό των ασθενών. Στην συνέχεια μελετήσαμε τον αντίκτυπο που έχουν τα καρδιαγγειακά προβλήματα στην κοινωνία και στην οικονομία. Δεν θα μπορούσαμε φυσικά να παραλείψουμε, το ψυχολογικό κόστος, του ασθενή και των οικείων του. Έπειτα από την μελέτη του αντικτύπου, κρίθηκε σκόπιμο να επισημανθούν, ακόμα περισσότεροι λόγοι, για τους οποίους είναι απαραίτητα τα συστήματα υποστήριξης αποφάσεων, στην πρόβλεψη των καρδιαγγειακών παθήσεων. Σε επόμενο στάδιο, πραγματοποιήθηκε βιβλιογραφική ανασκόπηση. Η ανασκόπηση εξέτασε βιβλιογραφίες, οι οποίες δεν ξεπερνούσαν τα πέντε έτη δημοσίευσης. Ο λόγος για τον οποίο επιλέχθηκαν τέτοιου είδους βιβλιογραφίες, ήταν για να εντοπιστούν, οι σύγχρονοι μέθοδοι πρόβλεψης των καρδιαγγειακών ασθενειών, με πρόσφατα δεδομένα.

Αφού ολοκληρώθηκε το κομμάτι της έρευνας, ακολούθησε η υλοποίηση, του δικού μας συστήματος υποστήριξης αποφάσεων. Για την δημιουργία του, επιλέχθηκε αρχικά ένα dataset. Το dataset αυτό, θα βασιζόταν η εκπαίδευση του συστήματος, για την διεξαγωγή του μοντέλου. Περιέχει δεδομένα, από πραγματικούς ασθενείς, με χαρακτηριστικά που επηρεάζουν τα καρδιαγγειακά προβλήματα και στο τέλος είχε μια κλάση, η οποία φανέρωνε με βάση τα χαρακτηριστικά, όπως για παράδειγμα την ηλικία, το φύλο την αρτηριακή πίεση και την καπνιστική συνήθεια, αν ο ασθενής έπασχε από κάποια καρδιαγγειακή νόσο ή όχι. Πρίν όμως από αυτό το βήμα, θα έπρεπε να γίνουν ορισμένες βελτιώσεις στο dataset, για να επιτευχθεί η μέγιστη απόδοση πρόβλεψης. Για την διεξαγωγή του κανόνα, χρησιμοποιήθηκε το σύστημα WEKA. Σε αυτή την πλατφόρμα εισαγάγαμε το dataset, σε αρχείο μορφής arff. Έπειτα, δοκιμάσαμε διάφορους αλγόριθμους ταξινόμησης, όπως ο J48, ο Decision Table, ο Random Forest και ο JRip, για να βρούμε εκείνον, με την μέγιστη απόδοση. Μόλις τον εντοπίσαμε, προβήκαμε στις τελευταίες τροποποιήσεις, με στόχο την αύξηση της απόδοσης του. Τέλος, όταν φτάσαμε σε ένα επιθυμητό αποτέλεσμα, μετατρέψαμε τον κανόνα που είχαμε σε php, το ίδιο και το περιβάλλον της εφαρμογής μας και με την βοήθεια του XAMPP και της Html, δημιουργήσαμε την τελική μας εφαρμογή και την αξιολογήσαμε βασισμένοι στις δοκιμές που κάναμε.

Κεφάλαιο Β΄

Περιγραφή Του Ιατρικού Προβλήματος

Όπως προαναφέρθηκε η εργασία αυτή κάνει λόγο για την εφαρμογή τεχνικών μηχανικής μάθησης για την δημιουργία μοντέλων κατηγοριοποίησης ασθενών. Η εφαρμογή που αναπτύχθηκε και που θα αναλυθεί στην συνέχεια, αφορά κυρίως την ανίχνευση ύπαρξης καρδιακής νόσου. Οι καρδιακές παθήσεις αποτελούν μια ολόκληρη ομάδα διαταραχών της καρδιάς και των αιμοφόρων αγγείων. Στις καρδιακές παθήσεις συμπεριλαμβάνεται η στεφανιαία νόσος, όπου πρόκειται για μια ασθένεια των αιμοφόρων αγγείων που τροφοδοτούν τον καρδιακό μυ. Στις ασθένειες της καρδιάς, συμπεριλαμβάνονται και η εγκεφαλοαγγειακή νόσος και η περιφερική αρτηριακή νόσος που αφορούν τις επιπλοκές στις αρτηρίες που τροφοδοτούν τον εγκέφαλο ή τα χέρια και τα πόδια, αντίστοιχα [10]. Άλλη μια πάθηση της καρδιάς είναι η ρευματική καρδιοπάθεια, δηλαδή η βλάβη του καρδιακού μυ και των καρδιακών βαλβίδων έπειτα από ρευματικό πυρετό, που οφείλεται σε στρεπτοκοκκικά βακτήρια. Πολλές φορές όμως οι καρδιοπάθειες μπορεί να οφείλονται σε συγγενείς καρδιοπάθειες, γενετικές ανωμαλίες που εμποδίζουν την ομαλή ανάπτυξη και λειτουργία της καρδιάς, εξαιτίας κάποιας εκ γενετής δυσπλασίας της δομής της καρδιάς. Μια ακόμα πάθηση της καρδιάς είναι η εν τω βάθει φλεβική θρόμβωση και πνευμονική εμβολή. Σε αυτή την περίπτωση θρόμβοι αίματος από τις φλέβες των ποδιών, μπορεί να μετακινηθούν προς την καρδιά και τους πνεύμονες και να δημιουργήσουν μια μη αναστρέψιμη πάθηση [10].

Η καρδιά είναι ένα από τα πιο σημαντικά όργανα του σώματος και ενώ υπάρχουν αρκετές ασθένειες που μπορεί να απειλήσουν την ομαλή λειτουργία της, μια από τις πιο σημαντικές είναι η καρδιακή ανεπάρκεια. Η καρδιακή ανεπάρκεια είναι μια σοβαρή κατάσταση, κατά την οποία η καρδιά δεν αντλεί ομαλά το αίμα στο σώμα, όπως το αντλεί μια φυσιολογική καρδιά. Η συγκεκριμένη κατάσταση της καρδιάς, είναι μη αναστρέψιμη καθώς δεν υπάρχει θεραπεία, για αυτήν την πάθηση [2]. Κύρια αιτία εμφάνισης της, είναι ο κακός τρόπος ζωής. Η καρδιά είναι ένας μυς ο οποίος είναι

υπεύθυνος για τον “εφοδιασμό” του σώματος με αίμα. Λειτουργεί σαν αντλία και στέλνει αίμα, συνοδευόμενο με οξυγόνο και θρεπτικά συστατικά στο σώμα [3].

Με την αδυναμία της καρδιάς να στείλει την απαραίτητη ποσότητα αίματος στο σώμα, αυτό δηλαδή που συμβαίνει στην καρδιακή ανεπάρκεια, το σώμα δεν λαμβάνει την απαραίτητη ποσότητα αίματος και θρεπτικών ουσιών [3]. Αυτές οι θρεπτικές ουσίες μαζί με το οξυγόνο αποτελούν απαραίτητα στοιχεία, για την ομαλή λειτουργία του οργανισμού. Οι μύες του σώματος, με την ελλείψεις αυτές, δεν θρεφονται κατάλληλα, με αποτέλεσμα οι ασθενείς που πάσχουν από καρδιακή ανεπάρκεια, να εμφανίζουν πολύ πιο εύκολα κόπωση [3]. Η δυσλειτουργία της καρδιάς δεν σταματάει κατά την αποστολή του αίματος στο σώμα. Μια καρδιά η οποία πάσχει από καρδιακή ανεπάρκεια, όχι μόνο δεν μπορεί να στείλει την απαραίτητη ποσότητα αίματος στο σώμα, αλλά δεν μπορεί ούτε να λάβει αρκετό αίμα. Η καρδιά, δεν μπορεί να τραβήξει το αίμα πίσω. Το υπολειπόμενο αίμα, το οποίο δεν επιστρέφει πίσω, συσσωρεύεται στις φλέβες, αυξάνοντας την πίεση τους. Κατά την πίεση που δημιουργείται στις φλέβες, από την συσσώρευση, αποκρίνονται υγρά στους γύρω ιστούς, με κύρια και χαρακτηριστικά σημεία, τα πόδια και την κοιλιά. Φυσικά, η έκρηξη υγρών δεν είναι το μοναδικό σύμπτωμα της καρδιακής ανεπάρκειας. Ένα ακόμα σύμπτωμα της, που δεν εμφανίζεται λίγες φορές, είναι η δύσπνοια. Το υγρό, που εκκρίνεται από την συσσώρευση του αίματος στις φλέβες, μπορεί να συσσωρευτεί στους πνεύμονες. Κάτι τέτοιο δημιουργεί συμφόρηση στον πνεύμονα, προκαλώντας δύσπνοια στον ασθενή. Ο οργανισμός, αντιδρά σε αυτές τις μεταβολές του, κάνοντας ορισμένες προσαρμογές, αλλά για σύντομο χρονικό διάστημα. Η προσπάθεια αυτή του οργανισμού, όχι μόνο δεν αντιμετωπίζει το πρόβλημα, αλλά αντιθέτως επιδεινώνει και επιβαρύνει την κατάσταση [3]. Η δύσπνοια, στην αρχή εμφανίζεται κατά την έντονη σωματική άσκηση, ενώ στην πορεία, γίνεται μέρος της καθημερινότητας ενός ανθρώπου που πάσχει από καρδιακή ανεπάρκεια, τόσο στην άσκηση, όσο και σε καταστάσεις ηρεμίας και κατάκλισης. Ένα ακόμα σύμπτωμα της καρδιακής ανεπάρκειας, είναι η απότομη αύξηση βάρους, η εύκολη κόπωση, ακόμα και σε απλές συνήθειες για τον ασθενή δραστηριότητες, όπως προαναφέραμε. Στην καρδιακή ανεπάρκεια (ΚΑ) παρατηρούνται περιπτώσεις ασθενών με αρρυθμίες, με επεισόδια ζάλης και με πόνο στο στήθος ή αλλιώς στηθάγχη [3].

Η καρδιακή ανεπάρκεια μπορεί να δημιουργηθεί έπειτα από έμφραγμα του μυοκαρδίου, γνωστό και ως στεφανιαία νόσος, δηλαδή από την στένωση των αρτηριών, που αιματώνουν την καρδιά προσφέροντας της οξυγόνο και θρεπτικά συστατικά, παρεμποδίζοντας την διαδικασία αυτή [4]. Καρδιακή ανεπάρκεια μπορεί να προκύψει και από μυοκαρδίτιδα, μια φλεγμονώδη νόσο του καρδιακού μυ.

Ακόμα μια αιτία, είναι η αρρυθμιστη αρτηριακή πίεση. Η αρτηριακή πίεση, είναι η πίεση που ασκεί το αίμα στις φλέβες κατά την ροή του σε αυτές. Όταν υπάρχει υψηλή αρτηριακή πίεση, η καρδιά θα πρέπει να εργαστεί ακόμα πιο δυνατά, για να διατηρήσει επαρκή ποσότητα αίματος στο σώμα [5]. Συνεπώς όταν η αρτηριακή πίεση είναι αρρυθμιστη και παρατηρείται άνοδος, η καρδιά θα πρέπει να εργάζεται πιο σκληρά, προκαλώντας με τον καιρό την καρδιακή ανεπάρκεια.

Βασικός παράγοντας επιρροής της καρδιακής ανεπάρκειας, είναι και ο σακχαρώδης διαβήτης. Οι άνθρωποι με σακχαρώδη διαβήτη έχουν μια άνοδο της γλυκόζης του αίματος, που κατά κύριο λόγο οφείλεται στην ελλιπή παραγωγή ινσουλίνης, δηλαδή μιας ορμόνης που εκκρίνεται από το πάγκρεας. Τα αγγεία του σώματος προσβάλλονται και προκαλούνται στενώσεις στα στεφανιαία αγγεία. Ο σακχαρώδης διαβήτης επηρεάζει την καρδιακή ανεπάρκεια, τόσο μέσα από την δημιουργία αθηροσκλήρωσης όσο και από την διαταραχή του μεταβολισμού των μυοκαρδιακών κυττάρων και των εμφραγμάτων του μυοκαρδίου [3]. Ένας άνθρωπος με σακχαρώδη διαβήτη θα πρέπει να προσέχει ιδιαίτερα το βάρος του και κυρίως την παχυσαρκία, ακολουθώντας ισορροπημένη διατροφή, να προβεί σε κατάλληλη θεραπεία και να αποφεύγει την κατανάλωση άλατος [3].

Καρδιακή ανεπάρκεια μπορεί να δημιουργήσει και η βαλβιδική νόσος [3]. Η καρδιά έχει τέσσερις βαλβίδες, αριστερά και δεξιά. Αυτές που βρίσκονται αριστερά ονομάζονται αορτική και μιτροειδή βαλβίδα, ενώ εκείνες που υπάρχουν στη δεξιά πλευρά λέγονται πνευμονική και τριγλώχινα βαλβίδα. Στην Βαλβιδική Καρδιακή Νόσο (VHD), υπάρχει κάποια φθορά σε μια από τις προαναφερθείσες αρτηρίες. Οι βαλβίδες μπορούν να υποστούν φθορά είτε από φυσικά αίτια όπως είναι η γήρανση ή λόγω βακτηριακής λοίμωξης που έχει επιρροή στις βαλβίδες ή εξαιτίας ρευματικής καρδιοπάθειας ή ως συγγενούς ανωμαλίας. Κάτι τέτοιο έχει σαν αποτέλεσμα να επιβαρύνει την καρδιά και να προκαλεί καρδιακή ανεπάρκεια.

Άλλος ένας παράγοντας είναι τα χρόνια νοσήματα, όπως είναι η νεφρική ανεπάρκεια, η μεσογειακή αναιμία και οι κακοήθειες. Γενικά ασθένειες που επιβαρύνουν την ομαλή καρδιακή λειτουργία, μπορούν να οδηγήσουν σε βάθος χρόνων, σε καρδιακά προβλήματα ή αλλιώς καρδιακή ανεπάρκεια. Συγκεκριμένα για την νεφρική ανεπάρκεια, οι νεφροί κατά την φυσιολογική τους λειτουργία, παράγουν την ερυθροποιητίνη, μια ορμόνη η οποία βοηθάει τον μυελό των οστών να παράγει ερυθρά αιμοσφαίρια [6]. Όταν τα νεφρά δεν λειτουργούν ομαλά, η ορμόνη αυτή μειώνεται και προκαλεί αναιμία και κόπωση. Ο λόγος που συμβαίνει αυτό, είναι γιατί το αίμα δεν μεταφέρει αρκετό οξυγόνο, απαραίτητο για τις ανάγκες των μυών και του εγκεφάλου [6]. Εκτός αυτού όμως, στο αίμα συσσωρεύονται ουσίες νοσηρές, για τον οργανισμό, που κανονικά θα υποβάλλονταν από τα σύρα. Η καρδιά για να ισορροπήσει την κατάσταση αυτή, αυξάνει τους ρυθμούς της και το “έργο της” γίνεται πιο δύσκολο [6].

Κάτι ανάλογο συμβαίνει και με τις παθήσεις του θυρεοειδούς. Ο θυρεοειδής είναι ένας αδένας που βρίσκεται στο μπροστινό μέρος του λαιμού. Οι θυρεοειδικές ορμόνες ρυθμίζουν τον κυτταρικό μεταβολισμό και στέλνουν ένα μήνυμα στα κύτταρα κάνοντας τα είτε να επιταχύνουν, είτε να επιβραδύνουν. Ένας υπερδραστήριος θυρεοειδής προκαλεί ταχυπαλμία και η περίσσεια θυρεοειδική ορμόνη ασκεί πίεση στον καρδιακό μυ, αυξάνοντας τη λειτουργία άντλησης, προκαλώντας πολλές φορές διαταραχή του καρδιακού ρυθμού. Αντιθέτως όταν υπολειτουργεί ο θυρεοειδής αδένας, παράγεται πολύ λίγη ορμόνη, ο καρδιακός ρυθμός επιβραδύνεται, υπάρχει μειωμένη λειτουργία

άντλησης και αδυνατεί να αυξηθεί το επίπεδο χοληστερόλης, οδηγώντας πολλές φορές σε στεφανιαία νόσο [3].

Ένας ακόμα παράγοντας που επηρεάζει την καρδιακή ανεπάρκεια, είναι οι χρόνιες πνευμονοπάθειες. Η πνευμονοπάθεια εμποδίζει την ανταλλαγή αερίων, με αποτέλεσμα να μην μεταφέρεται αρκετό οξυγόνο από τους πνεύμονες στο αίμα και κατά συνέπεια από την κυκλοφορία του αίματος στα όργανα [3]. Σε αυτή την περίπτωση, η καρδιά θα προσπαθήσει να αντισταθμίσει την κατάσταση δουλεύοντας πιο σκληρά, προκειμένου να ανταποκριθεί στις ανάγκες οξυγόνου των οργάνων. Η μεγάλη πίεση της καρδιάς, οδηγεί με το καιρό σε καρδιακά προβλήματα [3].

Μεγάλες πιθανότητες εμφάνισης ενός καρδιακού προβλήματος, έχουμε και έπειτα από την χορήγηση θεραπείας κατά του καρκίνου. Στις ογκολογικές θεραπείες συμπεριλαμβάνονται, τα χημειοθεραπευτικά και οι ακτινοβολίες.

Όπως έχει αναφερθεί η καρδιακή ανεπάρκεια είναι μια χρόνια πάθηση που χειροτερεύει με το χρόνο. Η ταχύτητα με την οποία θα επιδεινωθεί, δεν είναι καθορισμένος και διαφέρει από άνθρωπο σε άνθρωπο [7]. Υπάρχουν περιπτώσεις που τα συμπτώματα παραμένουν σταθερά για μεγάλα διαστήματα (από μήνες έως χρόνια), ενώ σε άλλες χειροτερεύουν χρόνο με το χρόνο. Βέβαια υπάρχει και μια τρίτη περίπτωση, που οι ασθενείς παραμένουν σχετικά σταθεροί, με την βοήθεια φαρμάκων, και χειροτερεύουν ξαφνικά είτε λόγω εμφάνισης αρρυθμίας, είτε εξαιτίας κάποιας λοίμωξης, είτε μιας καρδιακής προσβολής / έμφραγμα [7]. Ένας ακόμα παράγοντας που απορυθμίζει μια χρόνια σταθερή καρδιακή ανεπάρκεια είναι η μη συμμόρφωση του ασθενούς στην θεραπεία ή στους περιορισμούς που πρέπει να ακολουθεί [7].

Φυσικά πρόκειται για μια ασθένεια η οποία μπορεί να προληφθεί, με την διατήρηση του σωματικού βάρους, με την λήψη συχνών μερίδων με φρούτα μέσα στην ημέρα, την σωματική άσκηση, την αποφυγή του καπνίσματος, την συστηματική παρακολούθηση της αρτηριακής πίεσης και την παρακολούθηση των λιπιδίων του αίματος. Σημαντικό είναι να ελέγχουμε και την διατροφή μας και να προτιμάμε την μεσογειακή διατροφή που περιλαμβάνει αντιοξειδωτικά στοιχεία, φυτικές ίνες, σύνθετους υδατάνθρακες και μονοακόρεστα λιπαρά οξέα [3].

Κεφάλαιο Γ΄

Αντίκτυπος της καρδιακής ανεπάρκειας

Κάνοντας λόγο για αντίκτυπο της καρδιακής ανεπάρκειας, εννοούμε το κοινωνικό, το οικονομικό και ψυχολογικό κόστος που έχει η καρδιακή ανεπάρκεια τόσο στον ίδιο τον ασθενή, όσο και στην οικογένεια του και κατ' επέκταση στο κοινωνικό σύνολό. Στο κεφάλαιο αυτό θα δούμε ακόμα, τον αριθμό των ασθενών που αντιμετωπίζουν καρδιακή ανεπάρκεια, σε παγκόσμια κλίμακα αλλά και σε Ευρώπη και Ελλάδα. Με αυτόν τον τρόπο, θα γνωρίζουμε και τον ακριβή αριθμό των ατόμων που αφορά άμεσα η συγκεκριμένη πάθηση.

Γ΄.1 Αριθμός Ασθενών

Ο αριθμός των ασθενών με καρδιακή ανεπάρκεια διερευνήθηκε τόσο σε παγκόσμια κλίμακα, όσο και σε Ελλάδα και Ευρώπη. Σκοπός της συγκεκριμένης ενότητας, ήταν η μελέτη της έκτασης της πάθησης αυτής. Με αυτόν τον τρόπο θα γνωρίζαμε πλέον, τον αριθμό των ατόμων που αφορά η ασθένεια καθώς και το τι εκτάσεις έχει πάρει το πρόβλημα μελλοντικά. Σύμφωνα με την Ελληνική Καρδιολογική Εταιρεία, η καρδιακή ανεπάρκεια σήμερα είναι πιο συχνή και από τον καρκίνο, με πάνω από 26 εκατομμύρια άτομα στον κόσμο να έρχονται αντιμέτωποι της [3]. Με βάση τα δεδομένα του 2019, η καρδιακή νόσος είναι η πρώτη στην λίστα των αιτιών θανάτου στην Ευρώπη αλλά και στην Ευρωπαϊκή Ένωση (Ε.Ε.). Στην Ευρώπη οι θάνατοι από καρδιακή νόσο, κάθε χρόνο, είναι 3,9 εκατομμύρια ενώ στην Ευρωπαϊκή Ένωση 1,8 εκατομμύρια. Με το πέρασ του χρόνου, στις περισσότερες ευρωπαϊκές χώρες η θνησιμότητα από καρδιακές παθήσεις μειώνεται σημαντικά, χωρίς όμως να υπάρχει μείωση και στον αριθμό της νοσηρότητας. Ο αριθμός των ανθρώπων που πάσχουν από κάποιο καρδιακό νόσημα παραμένει πολύ ανεβασμένο. Με βάση τα στοιχεία του 2019, μόνο στην

Ευρωπαϊκή Ένωση, περίπου 49 εκατομμύρια άνθρωποι ζουν με καρδιαγγειακή νόσο. Οι καρδιακές παθήσεις αποτελούν την πιο συχνή αιτία θανάτων, με ποσοστό 45% σε γυναίκες και 39% στους άντρες [9].

Οι καρδιακές παθήσεις επηρεάζονται με ένα ποσοστό, που ξεπερνά το 75%, από το περιβάλλον, συγκεκριμένα από την ατμοσφαιρική ρύπανση και τον θόρυβο. Σύμφωνα με μελέτη του 2019, παρατηρήθηκε μεγαλύτερη ατμοσφαιρική ρύπανση, σε χώρες μεσαίου εισοδήματος, συγκριτικά με άλλες χώρες με υψηλό εισόδημα [9]. Κατά μέσο όρο υπολογίζονται περίπου 48.000 νέα κρούσματα σε όλη την Ευρώπη, με στεφανιαία νόσο εξαιτίας της ηχορύπανσης του περιβάλλοντος. Οι καρδιακές παθήσεις απασχολούν κυρίως χώρες με μεσαίο εισόδημα. Η διαφορά σε σύγκριση με χώρες με υψηλό εισόδημα, ανέρχεται σε ένα ποσοστό 30%. Η έλλειψη βασικών πόρων, σε χώρες μεσαίου εισοδήματος, απαραίτητων για την στήριξη των ανθρώπων με καρδιακές παθήσεις, εμποδίζουν ορισμένες διαδικασίες όπως τη στεφανιαία παρέμβαση, τις καρδιοχειρουργικές επεμβάσεις, εμφύτευση συσκευών και δυσχεραίνουν την κατάσταση, συγκριτικά με χώρες υψηλού εισοδήματος [9]. Ένα χαρακτηριστικό των καρδιακών παθήσεων, όπως έχει προαναφερθεί, είναι ο πρόωρος θάνατος. Οι καρδιακές παθήσεις ευθύνονται για το 62% των πρόωγων θανάτων, δηλαδή πριν να συμπληρώσουν τα 70 έτη, στην Ευρώπη. Από αυτό το ποσοστό το 29% αφορά γυναίκες, ενώ το 33% είναι άντρες [9]. Παρά την συνεχόμενη μείωση των καρδιαγγειακών παθήσεων, σε πάρα πολλές χώρες, παραμένει η πιο κοινή αιτία θανάτων, ανεξαρτήτως γενετικού φύλου. Στις γυναίκες καταλαμβάνει ένα ποσοστό 47% ενώ στους άντρες ένα ποσοστό 39%, σύμφωνα με μελέτες των τελευταίων ετών. Ωστόσο, σε ορισμένες χώρες υψηλού εισοδήματος, οι εξελίξεις και η πρόοδος σχετικά με την πρόληψη τους και την θεραπεία, έχουν οδηγήσει στο να είναι ο καρκίνος η πιο κοινή αιτία θανάτου. Για τις γυναίκες αυτές οι χώρες είναι η Δανία, το Ισραήλ, η Ολλανδία, η Δημοκρατία του Αγίου Μαρίνου και το Ηνωμένο Βασίλειο. Για τους άντρες αυτές οι χώρες αποτελούνται από το Βέλγιο, τη Δανία, τη Γαλλία, την Ιρλανδία, το Ισραήλ, το Λουξεμβούργο, την Ολλανδία, τη Νορβηγία, τη Πορτογαλία, τη Δημοκρατία του Αγίου Μαρίνου, τη Σλοβενία, την Ισπανία, την Ελβετία και το Ηνωμένο Βασίλειο [9]. Γενικά η καρδιακές παθήσεις απασχολούν και αφορούν ένα μεγάλο μέρος του πληθυσμού, με αποτέλεσμα μια μελέτη ή μια έρευνα πάνω σε αυτό να μην αφορά μόνο μια μικρή ομάδα ανθρώπων, αλλά ένα μεγάλο αριθμό ασθενών και συγγενών αυτών.

Στις δυτικές κοινωνίες, με βάση τα δεδομένα του 2016, έπασχε από καρδιακή ανεπάρκεια, το 2% του πληθυσμού και υπολόγιζαν ότι μέχρι το 2025 θα αυξηθεί στο 3%, δηλαδή θα πάσχουν περισσότεροι από 20 εκατομμύρια άνθρωποι στην Ευρώπη. Το ποσοστό αυτό αυξάνεται και αφορά κατά 10% κυρίως άτομα ηλικίας 75 ετών και άνω ή ακόμη και 20% σε ηλικίες 85 χρονών και άνω. Αυτή η αύξηση, των ατόμων με καρδιακή ανεπάρκεια, οφείλεται ως έναν βαθμό σε κακή διατροφή και απουσία σωματικής άσκησης [11]. Η αύξηση των περιπτώσεων, καρδιακής ανεπάρκειας είναι αναπόφευκτη. Η θεραπεία μειώνει τον κίνδυνο θνησιμότητας, αλλά αφήνουν πίσω καρδιές που υπολειτουργούν [11].

Γ'.2 Κοινωνικό Κόστος

Η συγκεκριμένη υποενότητα, έχει σαν στόχο την μελέτη της επιρροής που έχουν τα καρδιακά νοσήματα, ως προς τον άνθρωπο και την κοινωνία. Η καρδιακή ανεπάρκεια αλλάζει δραματικά την ζωή των ανθρώπων. Οι απλές καθημερινές δραστηριότητες παύουν να είναι το ίδιο απλές, όπως είναι για έναν υγιή άνθρωπο. Δραστηριότητες όπως το ανέβασμα σκαλιών και η ενασχόληση με οικιακές δουλειές, όπως το καθάρισμα ή το σκούπισμα, το περπάτημα ή ακόμα και η αλλαγή ρούχων, δεν πραγματοποιούνται με την ίδια ευκολία. Η δυσκολία αυξάνεται ολοένα και περισσότερο με το πέρασμα του χρόνου, που δυσχεραίνει η κατάσταση [8]. Η κόπωση είναι ένα από τα βασικά και πιο σύνηθες χαρακτηριστικά, της καρδιακής ανεπάρκειας. Επηρεάζει αρνητικά την καθημερινή ζωή των ασθενών καθώς και την ποιότητα της ζωής τους. Οι ασθενείς, θα πρέπει να παρακολουθούν τις αλλαγές στην κόπωση τους, για να μπορούν να ανταποκρίνονται στα νέα δεδομένα. Για να αντιμετωπίσουν το αίσθημα της κόπωσης, μπορεί να χρειαστούν συμβουλές για στρατηγικές αυτοδιαχείρισης, από ειδικούς [12].

Τα στατιστικά στοιχεία της καρδιακής ανεπάρκειας, που καταγράφονται από την Ευρωπαϊκή Καρδιολογική Εταιρεία, αποτελούν μια ευρεία επισκόπηση του φόρτου της ασθένειας, στην κοινότητα της. Η επιβάρυνση των ασθενών είναι ιδιαίτερα υψηλή, κυρίως στις χώρες με μεσαίο εισόδημα, πράγμα το οποίο εν μέρη οδηγεί σε αύξηση της καπνιστικής συνήθειας, της υπέρτασης και του διαβήτη, συγκριτικά με χώρες υψηλού εισοδήματος. Η σχετική έλλειψη σύγχρονων εγκαταστάσεων θεραπείας και αποκατάστασης, στις περισσότερες από τις χώρες μεσαίου εισοδήματος, υποβαθμίζει την παροχή υγειονομικής περίθαλψης, ενώ η ανάγκη για αυτή, είναι συγκριτικά μεγαλύτερη [9].

Τα έτη ζωής προσαρμοσμένα στην αναπηρία ή αλλιώς DALYs εκφράζουν τον συνολικό αριθμό των ετών που επιβαρύνθηκαν εξαιτίας της κακής υγείας, του πρόωρου θανάτου ή αναπηρίας. Το DALYs το χρησιμοποιούμε για να μετρήσουμε συνοπτικά, την απώλεια υγείας, εξαιτίας της πάθησης που μελετάμε. Ένα DALYS εκφράζει έναν χαμένο χρόνο, από υγιή ζωή. Το άθροισμα τους σε έναν ολόκληρο πληθυσμό, είναι ένας τρόπος μέτρησης της διαφοράς μεταξύ της υπάρχουσας κατάστασης που επικρατεί με μια ιδανική κατάσταση, όπου όλοι ζουν μέχρι μια προχωρημένη ηλικία, χωρίς ασθένειες και αναπηρίες. Η ισχαιμική καρδιακή νόσος και το εγκεφαλικό αντιπροσωπεύουν ένα ποσοστό που ανέρχεται στο 82% των DAILYs, δηλαδή των ετών ζωής προσαρμοσμένων σε αναπηρία. Κατά την περίοδο 1990 με 2017, τα DAILYs που αποδόθηκαν σε καρδιαγγειακή νόσο, ανά 100.000 άτομα, ήταν υψηλότερα σε γυναίκες από ότι σε άντρες (3.219 έναντι 5.925). Σε χώρες με μεσαίο εισόδημα ήταν τρεις φορές υψηλότερα, σε σύγκριση με χώρες υψηλού εισοδήματος (7.160 έναντι 2.235) [9].

Όπως προαναφέρθηκε, οι χρόνιες παθήσεις αποτελούν την πιο συχνή αιτία θνησιμότητας και νοσηρότητας στην Ευρώπη. Εξαιτίας της αλλαγής του τρόπου ζωής και της αύξησης των προτύπων υγειονομικής περίθαλψης, οι χρόνιες ασθένειες αυξάνονται χωρίς όμως να μειώνεται η διάρκεια ζωής και αυτό έχει σαν αποτέλεσμα την γήρανση του πληθυσμού [11]. Στις περισσότερες περιπτώσεις όταν εμφανίζεται ένα χρόνιο νόσημα, σε σύντομο χρονικό διάστημα εμφανίζονται και άλλα. Σε αυτές τις περιπτώσεις λέμε ότι έχουμε συννοσηρότητες. Η συχνότητα εμφάνισης τους υπολογίζεται στο $\frac{1}{3}$ του ενήλικου πληθυσμού. Οι χρόνιες ασθένειες αντιμετωπίζονται ξεχωριστά, με αποτέλεσμα οι φροντιστές να μην είναι σε θέση να εξετάσουν την αλληλεπίδραση πολλών παραγόντων που αφορούν τα ατομικά σχέδια φροντίδας. Σε πολλές περιπτώσεις, πολλοί ασθενείς με συννοσηρότητα, απαιτούν την περίθαλψη από περισσότερους από έναν φροντιστή. Με βάση την ευρωπαϊκή καρδιολογική εταιρεία υπάρχει κατά μέσο όρο 85 καρδιολόγοι ανά εκατομμύριο άτομα, με μεγάλες εξαιρέσεις το Ηνωμένο Βασίλειο και την Ιρλανδία που έχουν < 30 ανά εκατομμύριο άτομα και την Ελλάδα και τη Δημοκρατία της Γεωργίας που έχουν > 250 καρδιολόγους ανά εκατομμύριο άτομα [8]. Ο μεγάλος αριθμός φροντιστών, δημιουργεί κενά στον συντονισμό και την επικοινωνία μεταξύ τους, εμποδίζοντας τη διαδικασία φροντίδας και οδηγώντας σε ιατρικά λάθη [11]. Συννοσηρή ασθένεια είναι και η καρδιακή ανεπάρκεια. Η φροντίδα των ατόμων με καρδιακή ανεπάρκεια, είναι ένα από τα πιο χαρακτηριστικά παραδείγματα ασθενειών, που απαιτούν οργανωμένη και ολοκληρωμένη φροντίδα. Η φροντίδα είναι εξαιρετικά πολύπλοκη και απαιτεί πολλούς ειδικούς γύρω από αυτήν.

Ο αριθμός των ασθενών, όπως αναφέρθηκε, είναι μεγάλος και αναμένεται να αυξηθεί ακόμα περισσότερο τα επόμενα χρόνια. Οι θεραπείες κατά των καρδιακών παθήσεων, μειώνουν σημαντικά τον κίνδυνο θανάτου, αλλά δεν μπορούν να θεραπεύσουν αποτελεσματικά. Κατά αυτόν τον τρόπο συμβάλλουν στην αύξηση του επιπολασμού, δημιουργώντας μια ακόμα μεγαλύτερη οικονομική επιβάρυνση [11]. Οι θεραπείες που λαμβάνουν οι ασθενείς με καρδιακή ανεπάρκεια δεν λαμβάνουν υπόψη τις υπόλοιπες ασθένειες και αυτό επιβαρύνει τους ασθενείς, αλλά και το σύστημα υγείας [13]. Οι ασθενείς με καρδιακές παθήσεις, σε αρκετές περιπτώσεις, αδυνατούν να συνεχίσουν να εργάζονται και παύουν να ανήκουν στο εργατικό δυναμικό [8].

Γ'.3 Οικονομικό Κόστος

Εκτός από τις επιπτώσεις στην ποιότητα της ζωής τους, οι ασθενείς με χρόνιες ασθένειες, έχουν σοβαρές οικονομικές επιβαρύνσεις, καθώς το 75% του συνολικού κόστους υγειονομικής περίθαλψης, δαπανάται για χρόνιες ασθένειες [11]. Οι καρδιαγγειακές ασθένειες έχουν σημαντική οικονομική επιβάρυνση που επηρεάζει τόσο τα άτομα, το σύστημα υγείας όσο και τις κοινωνίες σε παγκόσμιο εύρος. Οι χώρες υψηλού εισοδήματος, δαπανούν σε υγειονομική περίθαλψη, κατά μέσο όρο τέσσερις φορές περισσότερο, από εκείνες μεσαίου εισοδήματος. Με βάση τα δεδομένα του 2016, στην ομάδα υψηλού εισοδήματος, παρατηρούνται σημαντικές δαπάνες στην υγεία, με πρώτες την

Ελβετία, το Λουξεμβούργο και τη Νορβηγία να ξοδεύουν πάνω από 6.000\$ κατά κεφαλήν. Ακολουθεί η Γερμανία, η Σουηδία και η Ιρλανδία που διαθέτουν περισσότερα από 5.000\$ κατά κεφαλήν. Στις χώρες μεσαίου εισοδήματος, που ξοδεύουν περισσότερα στην υγειονομική περίθαλψη είναι η Βουλγαρία, η Ρωσική Ομοσπονδία και η Σερβία, με κόστος μεγαλύτερο από 1.300\$ κατά κεφαλήν. Τελευταίο στην λίστα είναι το Μαρόκο, με κατά κεφαλήν 466\$ και το Κιργιστάν, ξοδεύοντας 240\$ κατά κεφαλήν, αντίστοιχα [8]. Η καρδιαγγειακή υγειονομική περίθαλψη, στις χώρες μέλη της Ευρωπαϊκής Καρδιολογικής Εταιρείας, αποτελεί σημαντικό ποσοστό από τις συνολικές δαπάνες για την υγειονομική περίθαλψη. Με βάση τα στοιχεία ενός report, με θέμα τις εθνικές δαπάνες για την υγειονομική περίθαλψη, ανά πάθηση, ηλικία και φύλο, η καρδιακή ανεπάρκεια, αφορούσε παραπάνω από το 10%, αυτών των δαπανών. Εκτός όμως από το υψηλό ποσοστό κόστους, που υπήρξαν στην υγειονομική περίθαλψη, υψηλά ποσοστά παρατηρήθηκαν και στις δαπάνες για ενδονοσοκομειακή και φαρμακευτική φροντίδα. Μια ακόμα πιο πρόσφατη ανάλυση του 2013, στην οποία συμμετείχαν 11 ευρωπαϊκές χώρες, επιβεβαίωσε ότι η καρδιαγγειακή νόσος αντιπροσωπεύει περίπου το 16% των συνολικών δαπανών υγείας. Το κόστος επηρεάζεται και ανεβαίνει από την περίθαλψη που χρειάζονται οι ασθενείς με καρδιακή ανεπάρκεια, την εμφύτευση βηματοδότη, της μεταμοσχεύσεις καρδιάς, την φαρμακευτική αγωγή και διάφορες άλλες σχετικές επεμβάσεις. Συγκεκριμένα, ανάμεσα στις χώρες μέλη της Ευρωπαϊκής Καρδιολογικής εταιρείας, υπάρχουν μεγάλες διαφορές μεταξύ των μελών, ως προς τις μεταμοσχεύσεις καρδιάς. Στην Ελλάδα, τη Μάλτα, τη Βόρεια Μακεδονία, τη Ρουμανία και τη Σερβία τα ποσοστά κυμαίνονται < 2 μεταμοσχεύσεις ανά εκατομμύριο άτομα ετησίως, ενώ στο Βέλγιο, την Κροατία και τη Νορβηγία είναι > 7 μεταμοσχεύσεις [9]. Από την άλλη τα εμφυτεύματα βηματοδότη είναι < 40 εμφυτεύματα ανά εκατομμύριο ανθρώπους στην Αίγυπτο, την Κιργιζία και το Ουζμπεκιστάν, ενώ στην Γαλλία και τη Σουηδία γίνονται έως και > 1000 εμφυτεύματα ανά εκατομμύριο ανθρώπους [9].

Η World Health Federation (W.H.F) / παγκόσμια ομοσπονδία υγείας, προβλέπει πως μέχρι το 2030, το συνολικό παγκόσμιο κόστος που αφορά καρδιαγγειακές νόσους θα αυξηθεί από 863 δισεκατομμύρια δολάρια, που ήταν το 2010, σε 1044 δισεκατομμύρια δολάρια. Αυτή η αύξηση θα επηρεαστεί κυρίως από την γήρανση του πληθυσμού και την αύξηση του κόστους διάγνωσης και θεραπευτικών τεχνολογιών [9]. Με βάση τις αναλύσεις, για την οικονομική επιβάρυνση από καρδιακές παθήσεις, εξάγονται τα αποτελέσματα πως το κόστος στην Ευρωπαϊκή Ένωση, αγγίζει τα 210 δισεκατομμύρια ευρώ ετησίως, εκ των οποίων το 53%, δηλαδή 111 δισεκατομμύρια ευρώ, αφορούν την υγειονομική περίθαλψη, το 26%, δηλαδή 54 δισεκατομμύρια ευρώ οφείλονται σε απώλεια παραγωγικότητας, εξαιτίας των σχετικών δυσκολιών και αναπηριών που προκύπτουν και το 21%, δηλαδή 45 δισεκατομμύρια, για την άτυπη φροντίδα των ατόμων, που πάσχουν από καρδιακή ανεπάρκεια. Η ισχαιμική καρδιοπάθεια, συνέβαλε αυξάνοντας την οικονομική επιβάρυνση, της Ευρωπαϊκής Ένωσης, για καρδιακές ασθένειες, ανεβάζοντας το κόστος περίπου 59 δισεκατομμύρια ευρώ, για το έτος 2015. Σε εθνικό επίπεδο, η καρδιαγγειακή νόσος επιβαρύνει την οικονομία της

Ισπανίας κατά 7,7 δισεκατομμύρια ευρώ, το Ηνωμένο Βασίλειο 18,9 δισεκατομμύρια ευρώ και την Γερμανία 37,4 δισεκατομμύρια ευρώ. Γενικά η παγκόσμια επιβάρυνση που επιφέρει η καρδιακή ανεπάρκεια, δεν αποτελεί μόνο ένα ζήτημα υγείας, αλλά πρόκειται για μια οικονομική πρόκληση που αφορά κυρίως το κόστος των συστημάτων και την υγειονομική περίθαλψη, που αναμένεται να αυξηθούν εκθετικά τα επόμενα χρόνια [9].

Γ'.4 Ψυχολογικό Κόστος

Οι δυσκολίες που αντιμετωπίζουν οι ασθενείς με την υγεία τους και κυρίως το αίσθημα του κόπου, τους εμποδίζουν από τις καθημερινές τους συνήθειες. Οι βόλτες καθώς και η παρεύρεση τους σε κοινωνικές εκδηλώσεις, γίνονται ολοένα και πιο δύσκολες. Οι ασθενείς με το πέρας του χρόνου απομονώνονται όλο και πιο πολύ. Δεν μπορούν να ασχοληθούν με χόμπι ή να διατηρήσουν κοινωνικές επαφές. Τα συμπτώματα και οι αναπηρίες, σε ηλικία εργασίας, μπορούν να επηρεάσουν την ικανότητα τους να εργαστούν, την απασχόληση τους και την παραγωγικότητά τους [8]. Έτσι, παραμένουν περισσότερες ώρες στο σπίτι, με το αίσθημα της ανικανότητας ορισμένες φορές να τους κυριεύει. Σύμφωνα με μια μελέτη, βασισμένη στην πρωτοβάθμια φροντίδα, έπειτα από δεκαετή παρακολούθηση, στη Σκωτία, από το 2017, η ποιότητα της ζωής των ατόμων, με καρδιακή ανεπάρκεια είναι πολύ χαμηλότερη, συγκριτικά με την ποιότητα ζωής ατόμων με διάφορες μορφές καρκίνου [8].

Οι ασθενείς με καρδιαγγειακές παθήσεις, πολλές φορές, χρειάζεται να κάνουν αρκετές εισαγωγές σε νοσοκομεία, είτε εξαιτίας άλλων παθήσεων που προκύπτουν, είτε λόγω επιπλοκών, της ίδιας της πάθησης. Το ψυχολογικό φόρτο, αυτής της διαδικασίας, ερευνήσαμε μέσα από ένα σύνολο δεδομένων, που συμπεριλάμβανε συνεντεύξεις από έντεκα (11) ασθενείς, που πήραν μέρος σε μια εθνογραφική μελέτη, για την καρδιακή ανεπάρκεια και είχε ως επίκεντρο τις απρογραμμάτιστες εισαγωγές σε νοσοκομεία. Οι ασθενείς που ερωτήθηκαν, περιέγραψαν πως η καρδιακή ανεπάρκεια και τα συμπτώματα της επηρέασαν και άλλαξαν την καθημερινότητα τους και πως διέκοψαν τις καθημερινές δραστηριότητες ρουτίνας. Τα έντονα συμπτώματα τους οδήγησαν να αναζητήσουν άμεση ιατρική φροντίδα. Οι ίδιοι μάλιστα, φανέρωσαν πως αντιμετώπισαν και συνεχίζουν να αντιμετωπίζουν, μεγάλες δυσκολίες, στην συμβίωση της καρδιακής ανεπάρκειας, με άλλες ασθένειες. Οι ασθενείς καλούνται να αντιμετωπίσουν πολλαπλές επιπλοκές, εξαιτίας της συννοσηρότητας, ενώ παράλληλα λαμβάνουν πολλαπλές και πολύπλοκες θεραπευτικές αγωγές. Από την άλλη τα ραντεβού πληθύνουν, ενώ καλούνται παρακολουθούνται, από περισσότερες από μια ειδικότητες [13]. Οι περισσότεροι από τους ερωτηθέντες, επιθυμούσαν και ζητούσαν να επιστρέψουν στην προ-νοσηρή ζωή τους, την οποία χαρακτήριζαν πιο ανεξάρτητη. Ένας μεγάλος αριθμός, από τα άτομα που πήραν μέρος στην έρευνα, τόνισαν πως συνάντησαν μια σχετική αβεβαιότητα και σοβαρές καθυστερήσεις, στην διάγνωση της ασθένειας. Επίσης, σημείωσαν καθυστερήσεις στην ενημέρωση από τους

παρόχους υγειονομικής περίθαλψης, καθώς και δυσκολίες στην μεταξύ τους επικοινωνία. Για τους ασθενείς όμως, την μεγαλύτερη σημασία είχε η οικογένεια και η υποστήριξη της. Παρόλα αυτά συχνά εξέφραζαν ανησυχία, μήπως επιβαρύνουν τα αγαπημένα τους πρόσωπα, με την ασθένεια τους και τις ιδιαιτερότητες τους [13].

Τα νέα δεδομένα που έρχονται στις ζωές των ασθενών που πάσχουν από καρδιακή ανεπάρκεια, προκαλούν αναστάτωση στην καθημερινότητα τους. Η εύκολη πρόσβαση στην υγειονομική περίθαλψη, μέσω καλής επικοινωνίας, κυρίως μεταξύ των υπηρεσιών και η δημιουργία ενός ισχυρού δικτύου, υποστήριξης της οικογένειας αλλά και των κλινικών ιατρών, θα μείωνε τον αντίκτυπο της καρδιακής ανεπάρκειας στην ζωή των ασθενών και των γύρω τους [13].

Οι πάσχοντες, εκτός από την προσπάθεια να προσαρμοστούν στην “νέα” τους ζωή, σύμφωνα με την τρέχουσα πολιτική, που επικρατεί στην υγειονομική περίθαλψη, θα πρέπει να μάθουν να αυτοδιαχειρίζονται πολλαπλές χρόνιες ασθένειες, να λάβουν ένα είδος εκπαίδευσης έτσι ώστε να είναι ικανοί να παρακολουθούν τη δική τους υγεία και να αποκτήσουν την ικανότητα να αναγνωρίζουν την σοβαρότητα της νόσου [13]. Από την άλλη πλευρά η δυσκολία στην πρόβλεψη των συμπτωμάτων, κάνει τους πάσχοντες να νιωθούν αβοήθητοι και πλήρως εξαρτημένοι από τους οικείους τους και τον περίγυρο τους, οδηγώντας σε απουσία ελέγχου της πάθησης, επιβαρύνοντας επιπλέον την οικογένεια και το σύστημα υγειονομικής περίθαλψης. Η φροντίδα και η ευθύνη μπορεί να μετατοπιστεί, από τον ασθενή στη/ο σύζυγο του ή στην άμεση οικογένεια του προσώπου. Πολλές φορές, σημειώνεται ότι άτομα με καρδιακή ανεπάρκεια, τα οποία έχουν μεγάλη υποστήριξη από τα μέλη της οικογένειας τους, συσφίγγουν περισσότερο τις σχέσεις τους μαζί τους και οι οικογενειακοί δεσμοί ενισχύονται, ιδιαίτερα σε περιόδους ανάγκης, όπως τα προβλήματα ή η κακή υγεία. Η επιτυχημένη οικογενειακή λειτουργία τονώνει την ευημερία του πάσχοντος [13]. Κάποιοι από τους ερωτηθέντες ανέφεραν πως τα μέλη της οικογένειας τους ανέλαβαν ένα πολύ προστατευτικό και σημαντικό ρόλο, είτε προσπαθώντας να επικοινωνήσουν με τις υπηρεσίες υγείας, εκ μέρους τους, είτε προτείνοντας λύσεις για τον ασθενή ενάντια στις συστάσεις άλλων ιατρών. Από τους 11 που ερωτήθηκαν, μόνον ένας ερωτώμενος μίλησε για ελλιπή οικογενειακή υποστήριξη και περιέγραψε το αίσθημα της ανησυχίας και την έλλειψη εμπιστοσύνης, στη διαχείριση της καρδιακής ανεπάρκειας. Σε αντίθεση με όλους του υπόλοιπους, που δέχθηκαν αρκετά γρήγορα τη διάγνωση, της καρδιακής ανεπάρκειας και προσπάθησαν γρήγορα να επιστρέψουν στην καθημερινότητα τους, έπειτα από την υποστήριξη των φίλων τους και της οικογένειας τους, εκείνος αντιμετώπισε σοβαρές δυσκολίες στην αποδοχή της πάθησης και στην διαχείριση της. Γενικά η οικογένεια και οι φίλοι αποτέλεσαν το κλειδί για την ανάρρωση των ασθενών [13].

Εξίσου σημαντική είναι και η σχέση μεταξύ του ατόμου που πάσχει από κάποια καρδιαγγειακή νόσο και του γιατρού του. Πρόκειται για ένα θεμέλιο, στο οποίο θα βασιστεί η αποτελεσματική διαχείριση της νόσου. Για να υπάρξει αποτελεσματική σχέση, μεταξύ γιατρού και ασθενή, προϋποθέτει την ύπαρξη υψηλής ποιότητας επικοινωνίας, η οποία θα προσφέρει ένα πλήθος

πληροφοριών, που θα αφορούν κυρίως τις θεραπευτικές επιλογές και την υποστήριξη των ασθενών. Ένα οργανωμένο σύστημα υγειονομικής περίθαλψης μπορεί να προσφέρει στις κλινικές σχέσεις. Αντίθετα, μια κακή επικοινωνία, έχει την δυνατότητα να οδηγήσει σε αβεβαιότητα, σύγχυση και ανησυχία του ασθενούς, μειώνοντας ραγδαία την ποιότητα της ζωής του και εμποδίζει την ανάρρωση του από τα θέματα υγείας του. Για τους ακόμα πιο ευάλωτους ασθενείς, η σχέση τους με τον γιατρό, είναι ακόμα πιο σημαντική. Σε πολλές τέτοιες περιπτώσεις ατόμων, παρατηρούνται σχέσεις εξάρτησης, από τον γιατρό τους. Μέσα από την αποτελεσματική επικοινωνία, μπορεί να διασφαλιστεί ότι οι γιατροί έχουν επίγνωση των προσδοκιών των ασθενών. Έτσι μπορούν να ρυθμίσουν πιο εύκολα τα συναισθήματα τους, να τους βοηθήσουν να κατανοήσουν τις σχετικές πληροφορίες, ικανοποιώντας με αυτόν το τρόπο περισσότερο τους ασθενείς τους. Όταν οι κλινικοί γιατροί και οι νοσηλευτές είναι εύκολα διαθέσιμοι, οι ασθενείς νιώθουν την απαραίτητη ασφάλεια. Κάνει τον πάσχον να δείχνει μια κατανόηση στο ότι έχει αφιερωθεί χρόνος και προσπάθεια για την παροχή των πληροφοριών ή την διαβεβαίωση και την περίθαλψη του.

Κεφάλαιο Δ΄

Ανάγκη για τεχνικές DSS

Η καρδιακή ανεπάρκεια, όπως έχει περιγραφεί μέχρι τώρα, είναι μια περίπλοκη πάθηση, με εξίσου περίπλοκη διάγνωση. Οι πιθανοί ασθενείς, συνήθως επισκέπτονται τον γιατρό τους, αφού παρατηρήσουν κάποιες αλλαγές στην υγεία τους, οι οποίες αλλαγές, παραπέμπουν σε συμπτώματα καρδιακής ανεπάρκειας. Ο γιατρός, παίρνει το ιστορικό από τον επισκέπτη του και προβαίνει σε εξετάσεις. Τις περισσότερες φορές γίνεται φυσική εξέταση από τον γιατρό, έπειτα ακολουθεί ακτινογραφία θώρακα, ηλεκτροκαρδιογράφημα, υπερηχοκαρδιογράφημα, αιματολογικές εξετάσεις και μαγνητική τομογραφία καρδιάς [3]. Στόχος του γιατρού δεν είναι μόνο η διάγνωση της καρδιακής ανεπάρκειας, στοχεύει και στην εύρεση της αιτίας, που προκλήθηκε η ασθένεια. Τις περισσότερες φορές το υπερηχογράφημα είναι και αυτό που θα φανερώσει και το μέγεθος του προβλήματος. Αυτό όμως σε ορισμένες περιπτώσεις, δεν είναι αρκετό, για να μάθει όλες τις λεπτομέρειες της πάθησης ή για να διευκρινίσει την αιτία και μπορεί να χρειαστεί να προβεί σε άλλες εξετάσεις, όπως αέρια αίματος, δοκιμασία κόπωσης, σπινθηρογράφημα, στεφανιογραφία κ.α. [7]. Όπως καταλαβαίνουμε η διάγνωση της καρδιακής ανεπάρκειας είναι μια ιδιαίτερα χρονοβόρα διαδικασία, κουραστική και ψυχοφθόρα για τους πιθανούς ασθενείς. Απάντηση στο πρόβλημα αυτό έρχονται να δώσουν οι τεχνικές DSS.

Τα συστήματα DSS (Decision Support System) ή αλλιώς Συστήματα Υποστήριξης Αποφάσεων είναι ένα σύνολο από διάφορες μεθόδους. Συγκεκριμένα χρησιμοποιεί μοντέλα αποφάσεων, που στηρίζονται σε κανόνες, που αναπτύσσονται με βάση προηγούμενες βιβλιογραφικές αναφορές και συνεργασίες με ειδικούς ιατρούς. Ακόμα γίνεται χρήση από μοντέλα ταξινόμησης και αλγορίθμους βελτιστοποίησης. Τα συστήματα αυτά, στηρίζονται στην μηχανική μάθηση. Πρόκειται για μια κατηγορία πληροφοριακών συστημάτων, βασισμένων στα υπολογιστικά συστήματα, τα οποία

στο σύνολο τους προσφέρουν μια πλήρη, εξατομικευμένη και φιλική προς τον χρήστην πλατφόρμα διαχείρισης ασθενών [14].

Ένας ακόμα λόγος, ο οποίος καθιστά απαραίτητη την χρήση συστημάτων υποστήριξης αποφάσεων και συγκεκριμένα ένα σύστημα πρόβλεψης της καρδιακής ανεπάρκειας, είναι το γεγονός ότι η έγκαιρη διάγνωση και η έγκαιρη έναρξη της κατάλληλης θεραπείας, είναι το κλειδί για την ανακούφιση των συμπτωμάτων, η πρόληψη της πρόσθετης βλάβης στον οργανισμό ενώ παράλληλα υπάρχει βελτίωση στην πρόγνωση [7]. Μέχρι σήμερα, η πιο αποτελεσματική θεραπεία της καρδιακής ανεπάρκειας, γίνεται με την χορήγηση φαρμακευτικής αγωγής και αλλαγή του τρόπου ζωής. Οι ασθενείς, της καρδιακής ανεπάρκειας, θα πρέπει να εντάξουν στην καθημερινότητα τους, το περπάτημα, να ακολουθούν πρόγραμμα δίαιτας, με περιορισμό στο αλάτι και να αποφεύγουν τις επιβλαβείς συνήθειες για τον οργανισμό, όπως είναι το κάπνισμα και το αλκοόλ, προκειμένου να δούνε την υγεία τους να καλυτερεύει. Ένας πάσχον θα πρέπει να παρακολουθεί το βάρος του, με καθημερινό ζύγισμα, την πίεση και τις σφύξεις του. Οι μετρήσεις αυτές, καταγράφονται σε έναν πίνακα και το δείχνει στον γιατρό του. Αν παρατηρηθούν μεγάλες αλλαγές στις μετρήσεις, όπως αύξησεις μεγαλύτερες των 2 κιλών σε 3 μέρες, τότε πρέπει να ειδοποιήσει τον γιατρό του. Στο ίδιο διάγραμμα ο ασθενής καλό είναι να αναφέρει και άλλες αλλαγές που μπορεί να παρατηρήσει στο σώμα του και στα συμπτώματά του είτε προς το καλύτερο είτε προς το χειρότερο. Παράδειγμα αλλαγών αποτελεί το συχνό ξύπνημα κατά την διάρκεια της νύχτας, λόγω δυσκολίας στην αναπνοή. Επειδή λοιπόν, οι ενέργειες που πρέπει να κάνει ο ασθενής και οι δυσκολίες που υπάρχουν στην θεραπεία, όπως είδαμε είναι σημαντικές και η πρόωρη πρόβλεψη της καρδιακής ανεπάρκειας συμβάλλει θετικά στην εξομάλυνση τους, τίθενται σχεδόν απαραίτητη η παρέμβαση συστημάτων υποστήριξης αποφάσεων [7].

Οι σύγχρονοι μέθοδοι που χρησιμοποιούνται στην καρδιακή ανεπάρκεια μέχρι τώρα έχουν βοηθήσει στην πρόληψη της νόσου, στην μείωση της καρδιακής θνησιμότητας. Οι τεχνολογικές παρεμβάσεις, βοηθούν σημαντικά την διαχείριση της καρδιακής νόσου και βοηθούν στην εξομάλυνση της οικονομικής επιβάρυνσης πολλών χωρών που δεν μπορούν να αντέξουν το οικονομικό βάρος. Ο οικονομικός προϋπολογισμός, της υγειονομικής περίθαλψης, σε πολλές δυτικές χώρες, ήδη ξεπερνά το 10% του ΑΕΠ . Οι μέχρι τώρα προβλέψεις δείχνουν ότι, τα επόμενα 50 χρόνια, οι ελλείψεις στις ιατρικές τεχνολογίες θα είναι ο πιο σημαντικός παράγοντας γήρανσης του πληθυσμού και αύξησης του κόστους υγειονομικής περίθαλψης [9]. Οι χώρες ακόμα και τώρα, πρέπει να βρίσκουν τρόπους, να επιλέγουν πολιτικές, σχετικά με τα καρδιαγγειακά νοσήματα, που θα ανταποκρίνονται οικονομικά στις τοπικές ανάγκες υγείας [9].

Το 2013 ο Παγκόσμιος οργανισμός υγείας / World Health Organization (WHO) μαζί με τα μέλη, του συμφώνησε να τεθεί ένα παγκόσμιο σχέδιο δράσης, με όνομα "Global action plan for the prevention and control of NCDs 2013-2020". Το σχέδιο αυτό αφορούσε τον τρόπο με τον οποίο θα μειωθεί ο αριθμός των πρόωγων θανάτων κατά 25% έως το 2025. Το σχέδιο αποτελούνταν από εννέα

(9) στόχους, εκ των οποίων οι δύο αφορούσαν την πρόληψη και τον έλεγχο των καρδιαγγειακών παθήσεων [10]. Παρόλου που η καρδιακή ανεπάρκεια, είναι από τις πιο διαδεδομένες και σύνθετες χρόνιες ασθένειες καθώς τις περισσότερες φορές συνδέεται με άλλες παθήσεις, η μέχρι σήμερα προσέγγιση έχει ελλείψεις. Αυτές οι ελλείψεις αφορούν κυρίως την διάγνωση της πάθησης και τη θεραπεία και τη φροντίδα των ασθενών. Η κατάσταση αυτή επηρεάζεται από το γεγονός ότι οι ενδιαφερόμενοι έχουν περιορισμένη αλληλεπίδραση και οι διάφορες ασθένειες αντιμετωπίζονται μεμονωμένα [11].

Η καθυστερημένη ή η ανακριβής διάγνωση της καρδιακής ανεπάρκειας, σε συνδυασμό με τις διάφορες επιπλοκές στην υγεία και τις συννοσηρότητες, οδηγούν στην χειροτέρευση της ποιότητας ζωής των ασθενών. Αυτό οφείλεται στο γεγονός, πως υπάρχει σχετική αδυναμία, στην πρόβλεψη εξατομικευμένων αναγκών, με αποτέλεσμα να μην επαρκούν και οι θεραπείες και να μην μπορούν να προβλέψουν την εξέλιξη της πάθησης. Όπως εύκολα διακρίνει κανείς, είναι απαραίτητο να καθοριστεί ο τρόπος με τον οποίον θα πραγματοποιούνται οι διαγνωστικές εξετάσεις και καθοδήγηση για την επαρκή ερμηνεία των αποτελεσμάτων [11]. Τα συμπτώματα της πάθησης, όπως η δύσπνοια, είναι κοινά με άλλες ασθένειες και δεν διακρίνουν την καρδιακή ανεπάρκεια από άλλες ασθένειες. Αυτό είναι ακόμα πιο έντονο στην πρωτοβάθμια περίθαλψη, όπου τα συμπτώματα είναι ήπια και το πρόβλημα είναι σε πρώιμο στάδιο. Σχεδόν τα μισά περιστατικά με καρδιακή ανεπάρκεια, διαγιγνώσκονται από τους γενικούς ιατρούς, ενώ άλλοι διαγιγνώσκονται κατά την πρώτη νοσηλεία. Αρκετές φορές μάλιστα, εξαιτίας της δυσκολίας που υπάρχει στην διάγνωση ασθενείς, λαμβάνουν λανθασμένα αποτελέσματα [11]. Πολλά αποτελέσματα αξιολόγησης, αναφέρουν πως τόσο στην διαγνωστική διαδικασία της πρωτοβάθμιας περίθαλψης, όσο και στην δευτεροβάθμια, δεν πραγματοποιούνται βασικές έρευνες που αναφέρονται στις οδηγίες. Όλες αυτές οι ελλείψεις που έχουν αναφερθεί, είναι πολύ σοβαρές, καθώς μπορεί να οδηγήσουν τους ασθενείς, σε νοσηλεία ή και σε θάνατο. Με έγκαιρη, ακριβή αναγνώριση ή/και παρακολούθηση των ασθενών, ένας μεγάλος αριθμός των νοσηλείων θα μπορούσε να αποφευχθεί. Ακόμα θα μπορούσε να “εμποδίσει” κατά κάποιον τρόπο την εξέλιξη της νόσου(ών) ή να συμβάλει στην πρόληψη της/τους [11]. Σε αυτό το σημείο, θα πρέπει να σημειωθεί ότι και η θεραπείες για τους ασθενείς με καρδιακή ανεπάρκεια, δεν είναι πλήρως εξατομικευμένες και δεν λαμβάνουν υπόψη όλες τις παθήσεις που μπορεί να υπάρχουν. Έτσι υπάρχει πιθανότητα να προκληθεί υπερβολική θεραπεία ή μη επαρκή, με συνέπεια την κακή έκβαση και την δημιουργία παρενεργειών. Η ανάγκη για δημιουργία εξατομικευμένων συστημάτων, λαμβάνοντας υπόψη όλα τα χαρακτηριστικά, είναι τεράστια [11].

Η πρόοδος της Τεχνητής Νοημοσύνης και της μηχανικής μάθησης, που στηρίζονται τα συστήματα υποστήριξης αποφάσεων, δεν δίνουν λύσεις μόνο ιατρικά ζητήματα και σε προβλήματα που σχετίζονται με καρδιαγγειακές παθήσεις. Μπορούν να χρησιμοποιηθούν σε αποφάσεις διαχείρισης, στη γενετική, στη διαχείριση του τρόπου ζωής, σε ζητήματα εθνικής ασφάλειας και στην πρόληψη περιβαλλοντικών καταστροφών. Δηλαδή, μπορούν να χρησιμοποιηθούν και να καλύψουν

τομείς υψηλού κινδύνου καθώς μειώνουν τον φόρτο εργασίας των χρηστών, αυξάνοντας την απόδοση εργασίας [15]. Διαγνωστικά συστήματα βασισμένα στην μηχανική μάθηση και στα συστήματα υποστήριξης αποφάσεων έχουν δημιουργηθεί στο παρελθόν και για άλλες παθήσεις, όπως για τη νόσο του Πάρκινσον, την ηπατίτιδα, τον καρκίνο. Πρόκειται για αυτοματοποιημένα συστήματα, με αλγορίθμους που βοηθούν στην αποτελεσματική ανίχνευση ασθενειών [16]. Η μηχανική μάθηση με τις μεθόδους που χρησιμοποιεί, είναι ειδικά διαμορφωμένη για να διαχειρίζεται πολύπλοκα δεδομένα και να παρέχει μοντέλα πρόβλεψης κινδύνου, με τα λιγότερα δυνατά σφάλματα σε ατομικό επίπεδο. Οι αλγόριθμοι που στηρίζονται στην μηχανική μάθηση είναι αποδεδειγμένα καλύτεροι, από τις στατιστικές μεθόδους, όσον αφορά την πρόβλεψη καρδιαγγειακών συμβάντων. Η μη γραμμική μοντελοποίηση σε συνδυασμό με τους πολύπλοκους υπολογισμούς και ο μοναδικός τρόπος ανάλυσης δεδομένων είναι τα βασικά χαρακτηριστικά που ξεχωρίζουν την μηχανική μάθηση από τη στατιστική ανάλυση [17].

Τα συστήματα υποστήριξης αποφάσεων έχουν την ικανότητα να συνδυάζουν τις διανοητικές ικανότητες των ανθρώπων με εκείνες των ηλεκτρονικών υπολογιστών. Στόχος τους είναι η βελτίωση της ποιότητας των λαμβανομένων αποφάσεων. Είναι συστήματα τα οποία επιτρέπουν την εύκολη αλληλεπίδραση (προγράμματα software), τα οποία χρησιμοποιούν μεθόδους, όπως η ανάλυση αποφάσεων για την ανάπτυξη μοντέλων. Τα μοντέλα αυτά, υποβοηθούν τους αποφασίζοντες να διαμορφώσουν εναλλακτικές λύσεις [21]. Σύμφωνα με τον Zeleny (1982) το χαρακτηριστικό από τα συστήματα υποστήριξης αποφάσεων, είναι να αυξάνουν τις ικανότητες των αποφασίζόντων, έτσι ώστε να προτείνουν στην πορεία καλύτερες λύσεις και όχι για να για να προτείνουν τα ίδια καλύτερες λύσεις, από εκείνες που μπορούν να αντιληφθούν από μόνοι τους οι αποφασίζοντες και στην περίπτωση μας γιατροί. [21] Με την χρήση ενός συστήματος υποστήριξης αποφάσεων, γίνεται προσπάθεια να υπάρξει βελτίωση στη ποιότητα της πληροφόρησης, στην οποία θα βασιστεί, στην συνέχεια μια απόφαση. Από το σύστημα δεν παρέχεται απλώς μια λύση, αλλά ένας αριθμός από εναλλακτικές λύσεις. Κατά αυτόν τον τρόπο οι αποφασίζοντες βοηθούνται, στο να κατανοήσουν εις βάθος τη φύση των διαφόρων προβλημάτων και κατά συνέπεια να βελτιώσουν τις αποφάσεις που θα λάβουν. [21]

Με βάση και τον αντίκτυπο που έχουν τα καρδιαγγειακά νοσήματα στον άνθρωπο και στην κοινωνία, τα οποία και είδαμε στο προηγούμενο κεφάλαιο, αλλά και με όσα μελετήσαμε σε αυτό, εύκολα οδηγούμαστε στο συμπέρασμα πως τα συστήματα υποστήριξης αποφάσεων σε συνδυασμό με την μηχανική μάθηση, πρόκειται να φέρουν επανάσταση στην ιατρική και συγκεκριμένα στην καρδιολογία.

Κεφάλαιο Ε΄

Βιβλιογραφική Ανασκόπηση

Στην συγκεκριμένη ενότητα πραγματοποιήθηκε βιβλιογραφική ανασκόπηση. Οι βιβλιογραφίες που μελετήθηκαν, αφορούσαν την πρόβλεψη των καρδιαγγειακών παθήσεων με την βοήθεια συστημάτων υποστήριξης αποφάσεων και μηχανικής μάθησης. Οι βιβλιογραφικές αναφορές, που χρησιμοποιήθηκαν στηρίχθηκαν πάνω σε έγκυρες έρευνες και μελέτες. Το έτος συγγραφής τους δεν ξεπερνάει τα πέντε χρόνια έτσι ώστε να είναι επίκαιρες, αφού η τεχνολογία εξελίσσεται μέρα με την μέρα. Σε πρώτο στάδιο, έγινε μια περίληψη από κάθε βιβλιογραφία ξεχωριστά. Η περίληψη αυτή αφορούσε την μέθοδο που ακολουθήθηκε στην έρευνα που μελετάμε το ακριβές πρόβλημα, το δείγμα πληθυσμού που έγινε και τέλος το ποσοστό επιτυχίας που είχε. Στην συνέχεια τα αποτελέσματα αυτά κατατάχθηκαν σε πίνακα και ακολούθησε η σύγκριση τους. Τέλος έγινε εκτενής ανάλυση των συμπερασμάτων της ανασκόπησης.

Ε΄.1 Βιβλιογραφικές Αναφορές

1. Πρόβλεψη Καρδιαγγειακών Συμβάντων Με Μηχανική Μάθηση [18]

Στην πρώτη βιβλιογραφική ανασκόπηση που μελετήσαμε είδαμε τον τρόπο με τον οποίο, η μηχανική μάθηση μπορεί να χρησιμοποιηθεί για την πρόβλεψη των καρδιαγγειακών παθήσεων. Στην συγκεκριμένη αναφορά, γίνεται χρήση του αλγορίθμου Random Forest, δηλαδή μιας τεχνικής μηχανικής μάθησης. Στην έρευνα αυτή πήραν μέρος άτομα από τη Πολυεθνική Μελέτη Αθηροσκλήρωσης / Multi-Ethnic Study of Atherosclerosis (MESA). Οι μετρήσεις χρησιμοποιήθηκαν για την πρόβλεψη των καρδιαγγειακών εκβάσεων για δωδεκάχρονη παρακολούθηση [18]. Η μελέτη παρακολούθησε την εξέλιξη της υποκλινικής νόσου, σε καρδιαγγειακά συμβάντα σε άτομα που

αρχικά ήταν απαλλαγμένα από καρδιαγγειακές παθήσεις. Συμμετείχαν 6814 άτομα, από το MESA, ηλικίας 45 έως 84 ετών, από 4 διαφορετικές εθνότητες και 6 κέντρα στις Ηνωμένες Πολιτείες. Ελήφθησαν 735 μεταβλητές, οι οποίες ήταν από μη επεμβατικές δοκιμές, ερωτηματολόγια, απεικονίσεις και πάνελ βιοδεικτών. Όπως αναφέρθηκε, χρησιμοποιήθηκε ο αλγόριθμος Random Forest, για να ξεχωρίσουν οι 20 κορυφαίοι προγνωστικοί παράγοντες κάθε αποτελέσματος. Το αποτέλεσμα έδειξε ότι οι απεικονίσεις, το ηλεκτροκαρδιογράφημα και οι βιοδείκτες ορού, ήταν συχνά στις λίστες των 20 κορυφαίων, σε αντίθεση με τους πιο παραδοσιακούς τρόπους διάγνωσης που συναντάμε συνήθως. Ο πιο σημαντικός προγνωστικός παράγοντας, που αφορούσε την θνησιμότητα ήταν η ηλικία. Από την άλλη πλευρά όμως, για την πρόγνωση του εγκεφαλικού επεισοδίου, οι πιο σημαντικοί παράγοντες ήταν τα επίπεδα γλυκόζης και οι μετρήσεις με υπερηχογράφημα καρωτίδων. Για την πρόγνωση της στεφανιαίας νόσου και για όλες τις εκβάσεις της αθηροσκληρωτικής καρδιαγγειακής νόσου, το ασβέστιο της στεφανιαίας αρτηρίας ήταν ο πιο σημαντικός παράγοντας. Αντίστοιχα για την καρδιακή ανεπάρκεια, οι πιο σημαντικοί παράγοντες για την πρόβλεψη, ήταν η δομή και η λειτουργία της αριστερής κοιλίας και η καρδιακή τροπονίνη-T. Η κρεατίνη και η ηλικία ήταν ανάμεσα στους κορυφαίους προγνωστικούς παράγοντες για την κολπική μαρμαρυγή. Ενώ το TNF-α (παράγοντας νέκρωσης όγκων) και το Νατριουρητικό Πεπτίδιο NT-proBNP ήταν σημαντικά για όλα τα αποτελέσματα. Το πόρισμά στο οποίο οδηγήθηκε η έρευνα αυτή, ήταν πως η μηχανική μάθηση, συνδυαστικά με το βαθύ φαινότυπο, βελτιώνουν την ακρίβεια των προβλέψεων, στις καρδιαγγειακές παθήσεις σε έναν ασυμπτωματικό αριθμό ατόμων.

2. Αυτοματοποιημένα Διαγνωστικά Συστήματα Βασισμένα σε Μηχανική Μάθηση Για Πρόβλεψη Καρδιακής Ανεπάρκειας Με Χρήση Διαφορετικών Τύπων Δεδομένων [16]

Η συγκεκριμένη βιβλιογραφία αναφέρεται, στα αυτοματοποιημένα διαγνωστικά συστήματα, τα οποία βασίζονται σε τεχνικές μηχανική μάθηση και εξόρυξης δεδομένων. Πρόκειται για μια βιβλιογραφική ανασκόπηση, η οποία στηρίζεται σε προηγούμενες τεχνικές μηχανικής μάθησης, εξόρυξης δεδομένων και μεθόδους δεδομένων που έχουν χρησιμοποιηθεί στο παρελθόν. Τα δεδομένα που συλλέχθηκαν ήταν από διάφορες βάσεις δεδομένων που δημιουργήθηκαν από το 1995 μέχρι και το 2021. Τα αυτοματοποιημένα συστήματα διάγνωσης, που βασίζονται σε Machine Learning, παρέχουν προσιτές αποτελεσματικές και αξιόπιστες λύσεις για την ανίχνευση των καρδιακών παθήσεων. Σε αντίθεση με προηγούμενες ανασκοπήσεις η εργασία που μελετάμε, στοχεύει στη συστηματική ανασκόπηση της αυτοματοποιημένης διάγνωσης για την πρόβλεψη καρδιακών παθήσεων, βασισμένη σε διαφορετικούς τύπους μεθόδων όπως για παράδειγμα μέθοδο δεδομένων βάσει κλινικών χαρακτηριστικών. Η εργασία που μελετάμε, αξιολογεί κριτικά τις προηγούμενες μεθόδους και παρουσιάζει τους περιορισμούς που εμφανίζουν αυτοί οι μέθοδοι. Αυτό το άρθρο παρουσιάζει ορισμένες μελλοντικές προεκτάσεις, στην ανίχνευση των καρδιαγγειακών παθήσεων, με

βάση την μηχανική μάθηση. Καταλήγει στο συμπέρασμα ότι υπάρχουν τρεις βασικοί παράγοντες, που συμμετέχουν για την αποτελεσματική ανίχνευση της καρδιακής ανεπάρκειας. Ο πρώτος είναι ότι τα δεδομένα, είναι πολύ σημαντικά για την αυτοματοποιημένη ανίχνευση, ο δεύτερος παράγοντας είναι η επικύρωση των δεδομένων. Η μηχανική μάθηση σύμφωνα με αυτήν την αναφορά, είναι ένας κλάδος που συνεχώς αναπτύσσεται και θα πρέπει να αναζητούνται νέες μέθοδοι, οι οποίες θα παρέχουν αποτελεσματική απόδοση [16].

3. Αποτελεσματική πρόβλεψη της καρδιακής ανεπάρκειας, μέσα από ένα έμπειρο, βελτιωμένο σύστημα, βασισμένο στην μηχανική υποστήριξη [19]

Σύμφωνα με αυτή την έρευνα, οι περισσότεροι ασθενείς με καρδιακή ανεπάρκεια, πεθαίνουν από την πάθηση τους τα επόμενα πέντε χρόνια, μετά την διάγνωση τους. Για αυτόν το λόγο επιδιώκει να δημιουργήσει ένα βελτιωμένο σύστημα πρόβλεψης της καρδιακής ανεπάρκειας. Πιο αναλυτικά, πρόκειται για ένα σύστημα, το οποίο συνδυάζει δύο μοντέλα μηχανών διανυσμάτων υποστήριξης (Support vector machine (SVM) models), για την αποτελεσματική πρόβλεψη της καρδιακής ανεπάρκειας. Το πρώτο από τα μοντέλα που χρησιμοποιήθηκαν, είναι γραμμικό και κανονικοποιημένο. Ένα από τα χαρακτηριστικά του είναι, πως έχει την δυνατότητα να εξαλείψει χαρακτηριστικά τα οποία δεν είναι απαραίτητα, συρρικνώνοντας τους συντελεστές τους. Το δεύτερο μοντέλο που χρησιμοποιήθηκε είναι ένα μοντέλο πρόβλεψης. Προκειμένου να γίνει βελτιστοποίηση των δύο μοντέλων, προτάθηκε η χρήση υβριδικού αλγορίθμου βαρυτικής αναζήτησης, που είναι ικανός να βελτιστοποιήσει τα δύο μοντέλα ταυτόχρονα. Τα αποτελέσματα αυτής της πρότασης, αξιολογήθηκαν μέσα από έξι διαφορετικές παραμέτρους. Οι παράμετροι αυτοί ήταν η ακρίβεια και η ευαισθησία, έγινε σύγκριση με τον συντελεστή συσχέτισης Matthews, ελέγχθηκε ως προς το γράφημα ROC, την περιοχή κάτω από την καμπύλη AUC και τέλος ως προς την ειδικότητα. Τα αποτελέσματα των δοκιμών έδειξαν, πως η μέθοδος βελτιώνει την απόδοση ενός συμβατικού μοντέλου κατά 3,3%. Η μέθοδος αυτή, δείχνει καλύτερη απόδοση σε σύγκριση με προηγούμενες μεθόδους που είχαν ακρίβεια 57,85% με 91,83%. Τέλος η μέθοδος αυτή, εμφάνισε καλύτερες επιδόσεις από τα άλλα υπερσύγχρονα μοντέλα, συνόλων μηχανικής μάθησης.

4. Ένα σύστημα υποστήριξης αποφάσεων, που χρησιμοποιεί στατιστικό μοντέλο και Gaussian Naive Bayes [20]

Σε αυτή την αναφορά, είδαμε μια ακόμα προσέγγιση, για την πρόβλεψη της καρδιακής ανεπάρκειας. Ανά διαστήματα προτίθενται διάφορα έξυπνα συστήματα υποστήριξης ιατρικών

αποφάσεων, για την ανίχνευση της καρδιακής ανεπάρκειας. Παρόλα αυτά, τα περισσότερα από αυτά τα συστήματα έχουν πολύ μικρά ποσοστά ακρίβειας. Η συγκεκριμένη βιβλιογραφική ανασκόπηση για να αναπτύξει την ακρίβεια πρόβλεψης, ανέπτυξε ένα σύστημα το οποίο βασίζεται σε δύο κύρια στάδια. Το πρώτο στάδιο, εφαρμόζει ένα στατιστικό μοντέλο, που κατατάσει δεκατρία σύνθετες χαρακτηριστικά, της καρδιακής ανεπάρκειας. Μέσα από διάφορες δοκιμές και τη χρήση του αλγορίθμου best-first search strategy, αναζητά ένα υποσύνολο, καλύτερων χαρακτηριστικών. Σε δεύτερο στάδιο, χρησιμοποιεί τον αλγόριθμο Gaussian Naive Bayes (GNB), ως μοντέλο πρόβλεψης. Η απόδοση του αλγορίθμου, αξιολογείται με βάση μια ηλεκτρονική βάση δεδομένων, με στοιχεία από άτομα με καρδιακή ανεπάρκεια. Η βάση αυτή διέθετε 298 άτομα, με καρδιακή ανεπάρκεια. Τα αποτελέσματα αυτής της έρευνας, έδειξαν πως θα μπορούσε να πετύχει ακρίβεια πρόβλεψης έως και 93,33%. Δηλαδή βελτιώνει την απόδοση πρόβλεψης, της ασθένειας κατά 3,33%. Η προτεινόμενη μέθοδος, εμφάνισε καλύτερη απόδοση από τις διαθέσιμες μεθόδους, που υπήρχαν μέχρι εκείνη την στιγμή στην βιβλιογραφία και είχαν ακρίβεια πρόβλεψης από 57,85% έως και 92,22%.

5. Μια προσέγγιση που βασίζεται σε αμοιβαία πληροφορημένα νευρωνικά δίκτυα, για την πρόβλεψη καρδιακής ανεπάρκειας [22]

Στην πέμπτη βιβλιογραφική αναφορά που μελετήθηκε, παρουσιάστηκε μια προσέγγιση, βασισμένη σε αμοιβαία πληροφορημένα νευρωνικά δίκτυα, για τη βελτίωση των δυνατοτήτων γενίκευσης, των συστημάτων υποστήριξης αποφάσεων, τα οποία αναπτύχθηκαν για την πρόβλεψη της καρδιακής ανεπάρκειας. Τα περισσότερα μοντέλα μηχανικής μάθησης, που έχουν αναπτυχθεί το τελευταίο διάστημα, για διάφορα μοντέλα μηχανικής μάθησης αλλά και για πρόβλεψη καρδιακής ανεπάρκειας, παρουσιάζουν προβλήματα υπερβολικής προσαρμογής. Όταν λέμε πρόβλημα προσαρμογής εννοούμε, ότι ορισμένα από αυτά, εμφανίζουν καλύτερη απόδοση στα δεδομένα εκπαίδευσης, αλλά δείχνουν κακή απόδοση στα δεδομένα δοκιμών ή το αντίστροφο, δείχνουν καλή απόδοση στα δεδομένα δοκιμών και κακή απόδοση στα δεδομένα εκπαίδευσης. Ένα μοντέλο το οποίο θα μπορούσε να παρουσιάσει καλύτερες επιδόσεις, στα δεδομένα εκπαίδευσης και στα δεδομένα δοκιμών, θα μείωνε σε μεγάλο βαθμό τα συνολικά σφάλματα πρόβλεψης. Τέτοια μοντέλα πρόβλεψης, θα μπορούσαν να αξιοποιηθούν από καρδιολόγους για την αποτελεσματική διάγνωση της καρδιακής ανεπάρκειας. Για μεγαλύτερη ακρίβεια, η συγκεκριμένη βιβλιογραφία, προτείνει ένα σύστημα υποστήριξης αποφάσεων, δύο σταδίων, αποφεύγοντας κατά αυτόν τον τρόπο την υπερβολική προσαρμογή. Το πρώτο στάδιο εφαρμόζει ένα στατικό μοντέλο, βασισμένο σε πληροφορίες, ενώ το δεύτερο στάδιο κάνει χρήση ενός νευρωνικού δικτύου. Η προσέγγιση αυτή, εφαρμόστηκε σε ένα υποσύνολο της δημόσιας, βάσης δεδομένων καρδιακών παθήσεων, του Κλίβελαντ, μια πόλης των ΗΠΑ. Τα αποτελέσματα, από τις πειραματικές δοκιμές έδειξαν, ότι το νέο εξελιγμένο σύστημα

υποστήριξης αποφάσεων, εμφανίζει πρόοδο στις δυνατότητες γενίκευσης και έχει μειωμένο ποσοστό σφάλματος στο 8,8%, το οποίο είναι αισθητά μικρότερο από άλλες πρόσφατες δημοσιευμένες μελέτες. Το μοντέλο, παρουσίασε ποσοστό ακρίβειας 93,33%, το οποίο ήταν υψηλότερο από άλλα εικοσιοκτώ πρόσφατα μοντέλα πρόβλεψης καρδιακής ανεπάρκειας, που η ακρίβεια τους κυμαίνονταν από το 57,85% μέχρι το 92,31%.

6. Αξιολόγηση διαφόρων τεχνικών μηχανικής μάθησης, για την πρόβλεψη της καρδιακής ανεπάρκειας [23]

Οι καρδιαγγειακές παθήσεις αποτελούν μια συχνή και όχι άδικα, ανησυχία για την δημόσια υγεία, σε παγκόσμιο επίπεδο. Επειδή οι καρδιαγγειακές παθήσεις αυξάνονται γρήγορα, εξαιτίας της ελλειπούς ευαισθητοποίησης, για την υγεία και του κακού τρόπου ζωής κρίνεται σχεδόν απαραίτητο, να υπάρχει ένα σύστημα το οποίο να μπορεί να διακρίνει, τον επιπολασμό της καρδιακής νόσου. Σε αυτό το άρθρο, εξετάζεται ο συνδυασμός από έξι τεχνικές μηχανικής μάθησης για την πρόβλεψη των καρδιακών παθήσεων. Η μέθοδος αξιολογήθηκε μέσα από οκτώ διαφορετικές ταξινομήσεις δεικτών απόδοσης (classification performance indices). Ακόμα η μέθοδος αυτή, αξιολογήθηκε με βάση τη Receiver operating characteristic ή αλλιώς καμπύλη ROC, δηλαδή μια γραφική παράσταση που παρουσιάζει τη διαγνωστική ικανότητα ενός δυαδικού συστήματος ταξινόμησης. Η πιο υψηλή ταξινόμηση είχε ακρίβεια 85%, χρησιμοποιούσε λογιστική παλινδρόμηση, είχε ευαισθησία 89% και specificity 81%.

7. Ένα σύστημα βασισμένο στον αλγόριθμο τυχαίας αναζήτησης και σε βελτιστοποιημένο μοντέλο random forest για την ανίχνευση καρδιακών παθήσεων [24]

Η βιβλιογραφία ασχολείται όλο και περισσότερο, με την ανίχνευση και την πρόβλεψη των καρδιαγγειακών παθήσεων. Η διάγνωση της καρδιακής ανεπάρκειας μέχρι και σήμερα αποτελεί ένα από τα πιο δύσκολα έργα, τόσο για τις αναπτυσσόμενες όσο και για τις υποανάπτυκτες χώρες που υπάρχει έλλειψη ειδικών και εξοπλισμού. Πολλοί ερευνητές, ανά διαστήματα έχουν αναπτύξει διαφορετικά έξυπνα συστήματα για την αυτοματοποιημένη ανίχνευση της καρδιακής ανεπάρκειας. Τα περισσότερα όμως από αυτά παρουσιάζουν σημαντικά προβλήματα υπερ προσαρμογής. Για παράδειγμα, όπως έχουμε σημειώσει ξανά, πολλές από τις μεθόδους ανίχνευσης, βελτιώνουν την ακρίβεια στα δεδομένα δοκιμών και υστερούν στα δεδομένα εκπαίδευσης. Σε αυτή την εργασία, έγινε ακόμα μια προσπάθεια, εξισορρόπησης της ακρίβειας των δεδομένων εκπαίδευσης και δοκιμών μέσα

από ένα νέο διαγνωστικό σύστημα. Το νέο αυτό σύστημα, εφάρμοσε τον αλγόριθμο τυχαίας αναζήτησης / random search algorithm (RSA), για την επιλογή χαρακτηριστικών και τον αλγόριθμο random forest, για την πρόβλεψη της καρδιακής ανεπάρκειας. Η βελτιστοποίηση επιτεύχθηκε με τον αλγόριθμο grid search.

Για την αξιολόγηση της ακρίβειας της μεθόδου, εκτελέστηκαν δύο είδη πειραμάτων. Στο πρώτο πείραμα αναπτύχθηκε μόνο ο random forest, ενώ στο δεύτερο πείραμα αναπτύχθηκε ο random forest βασισμένος στην τυχαία αναζήτηση. Τα πειράματα εκτελέστηκαν βασισμένα, σε μια ηλεκτρονική βάση δεδομένων του Cleveland. Η μέθοδος αυτή αποδείχθηκε αρκετά αποτελεσματική και λιγότερο περίπλοκη από την απλή εφαρμογή του random forest, καθώς έδωσε 3,3% μεγαλύτερη ακρίβεια χρησιμοποιώντας μόνο επτά χαρακτηριστικά. Σε αυτό το σημείο θα πρέπει να προσθέσουμε ότι, η εν λόγω μέθοδος, έφερε καλύτερες επιδόσεις από τις άλλες πέντε τεχνολογίες που στηρίχθηκαν στα μοντέλα μηχανικής μάθησης. Η προτεινόμενη μέθοδος πέτυχε ακρίβεια ταξινόμησης 93,33%, ενώ παράλληλα βελτίωσε και την ακρίβεια των δεδομένων εκπαίδευσης. Κατάφερε να παρουσιάσει καλύτερη απόδοση σε σύγκριση με άλλες έντεκα πρόσφατα προτεινόμενες μεθόδους, ανίχνευσης της καρδιακής ανεπάρκειας.

8. Βελτίωση της ακρίβειας πρόβλεψης, του κινδύνου καρδιοπάθειας, με βάση τις τεχνικές ταξινόμησης συνόλου [25]

Στην μηχανική μάθηση συμπεριλαμβάνεται και η τεχνητή νοημοσύνη. Ένα χαρακτηριστικό της μηχανικής μάθησης είναι, η πρόβλεψη ενός αποτελέσματος με βάση τα δεδομένα που διαθέτουμε. Το μηχάνημα εντοπίζει και αφομοιώνει μοτίβα, από το υπάρχον σύστημα δεδομένων και στην πορεία τα εφαρμόζει σε ένα άγνωστο σύνολο δεδομένων, για να προβλέψει το αποτέλεσμα. Μια τεχνική μηχανικής μάθησης που χρησιμοποιείται για προβλέψεις και μάλιστα ισχυρή, είναι η ταξινόμηση. Κάποιοι από τους αλγορίθμους ταξινόμησης, προβλέπουν με ικανοποιητική ακρίβεια, ενώ υπάρχουν άλλοι που φέρουν περιορισμένα αποτελέσματα. Αυτή η έρευνα εξετάζει μια μέθοδο, εν ονόματι ταξινόμηση συνόλου, η οποία χρησιμοποιείται για τη βελτίωση της ακρίβειας αλγορίθμων, που εμφανίζουν αδυναμίες, συνδυάζοντας πολλαπλούς ταξινομητές. Τα πειράματα με αυτό το εργαλείο, έγιναν χρησιμοποιώντας, ένα σύνολο από δεδομένα καρδιακών παθήσεων. Στην πορεία έγινε μια ανάλυση προκειμένου να προσδιοριστεί, ο τρόπος κατά τον οποίο θα μπορούσε να εφαρμοστεί η τεχνική του συνόλου, προκειμένου να βελτιωθεί η ακρίβεια στις προβλέψεις, των καρδιαγγειακών παθήσεων. Αυτή η βιβλιογραφική αναφορά, δεν εστιάζει μόνο στην αύξηση της ακρίβειας, των αδύναμων αλγορίθμων ταξινόμησης. Αντίθετα προσπαθεί να εφαρμόσει τον αλγόριθμο, σε ένα ιατρικό σύνολο δεδομένων, με στόχο την ανάδειξη της χρησιμότητας του, στην πρόβλεψη των καρδιαγγειακών παθήσεων σε πρώιμο στάδιο. Τα αποτελέσματα της έρευνας έδειξαν, πως οι τεχνικές

συνόλου όπως το bagging και το boosting, είναι αποτελεσματικές στη βελτίωση της ακρίβειας πρόβλεψης των αδύναμων ταξινομητών, ενώ παράλληλα παρουσίαζαν ικανοποιητικές επιδόσεις και στον εντοπισμό του κινδύνου για κάποια καρδιακή νόσο. Με την ταξινόμηση συνόλου, επιτεύχθηκε μια αύξηση ακρίβειας 7%, στην πρόβλεψη των καρδιακών παθήσεων. Η απόδοση της διαδικασίας βελτιώθηκε, με την εφαρμογή επιλογής χαρακτηριστικών, ενώ τα αποτελέσματα έδειξαν σημαντικές βελτιώσεις, στην ακρίβεια της πρόβλεψης.

9. Έγκαιρη πρόβλεψη καρδιακών προβλημάτων, μέσω προγνωστικής ανάλυσης σημάτων ECG [26]

Οι εξυπνες συσκευές για την παρακολούθηση της υγείας γίνονται ολοένα και περισσότερο τάση τα τελευταία χρόνια. Τα τελευταία έτη έχουν αναπτυχθεί συστήματα, που επιτρέπουν την απομακρυσμένη παρακολούθηση της καρδιάς, τα οποία αξιοποιώντας προηγμένες μεθόδους μηχανικής μάθησης εντοπίζουν καρδιακές διαταραχές, μέσω επεξεργασίας σημάτων ECG (ηλεκτροκαρδιογραφήματος). Παρόλα αυτά οι τρέχουσες τεχνολογίες υποφέρουν από δύο σημαντικά μειονεκτήματα: 1) υπάρχει έλλειψη ικανότητας έγκαιρης πρόβλεψης καρδιακών ανωμαλιών και 2) αποτυγχάνουν οι καταγραφές της μεταβλητότητας μεταξύ των ασθενών. Σε αυτή την βιβλιογραφική ανασκόπηση προτείνεται ένα νέο προγνωστικό πλαίσιο δύο σταδίων για την επεξεργασία του σήματος ηλεκτροκαρδιογραφήματος (ECG). Σε αυτή την περίπτωση ένας παγκόσμιος ταξινομητής, διακρίνει τις σοβαρές ανωμαλίες, δίνοντας έναν κόκκινο συναγερμό, έπειτα από σύγκριση του σήματος με ένα καθολικό μοντέλο αναφοράς. Τα δείγματα σήματος, που στέλνονται σαν φυσιολογικά, υποβάλλονται σε επακόλουθη ανάλυση απόκλισης. Οι κίτρινοι συναγερμοί ενεργοποιούνται όταν εντοπίζονται ήπιες παραμορφώσεις στο σήμα, σε σύγκριση με την βασική αναφορά που έχει τεθεί. Αυτό μπορεί να σημαίνει ότι ο συγκεκριμένος ασθενής, θα έρθει στο μέλλον αντιμέτωπος με κάποια καρδιακή πάθηση. Για να γίνει πιο εύκολα μια ακριβής ανάλυση απόκλισης, προτάθηκε να εφαρμοστεί ένας ελεγχόμενος μη γραμμικός μετασχηματισμός με βελτιστοποιημένες παραμέτρους. Ένας ακόμα στόχος είναι να υπάρχουν διαφορετικά σήματα ανάλογα με τις διαφορετικές κατηγορίες πάθησης. Η μέθοδος αυτή επιτυγχάνει ακρίβεια ταξινόμησης 96,6% και παρέχει ένα μοναδικό χαρακτηριστικό, αυτό της προγνωστικής ανάλυσης, δημιουργώντας προειδοποιητικά μηνύματα, σχετικά με τον κίνδυνο καρδιακών ανωμαλιών. Έτσι μπορούν να ληφθούν και προληπτικά μέτρα, σύμφωνα με τις οδηγίες που θα δώσει ο επιβλέπων γιατρός του ασθενή. Η πιθανότητα να παρατηρηθεί ένα σοβαρό πρόβλημα, δηλαδή να δοθεί κόκκινος συναγερμός, μετά από την παρατήρηση και σήμανση ενός κίτρινου συναγερμού, του ίδιου τύπου, αυξάνεται κατά 5% έως 10%. Ο βασικός στόχος αυτής της τεχνολογίας, είναι να βελτιωθεί και να γίνει ποιοτική η υγειονομική περίθαλψη, τόσο σε ηλικιωμένα άτομα, όσο και σε καρδιοπαθείς υψηλού κινδύνου.

10. Αξιολόγηση του επιπέδου κινδύνου της καρδιακής νόσου μέσα από ένα σύστημα βασισμένο σε κανόνες. [27]

Τα συστήματα βασισμένα στην γνώση, δεν είναι ιδιαίτερα αποτελεσματικά, για την λήψη αποφάσεων, για την παροχή διάγνωσης και θεραπείας των καρδιακών παθήσεων. Αυτό οφείλεται στις ασαφές και ανακριβείς πληροφορίες που παρέχουν. Για αυτόν τον λόγο τα αυτόματα διαγνωστικά συστήματα χρειάζονται πολύ χρόνο, για τη βελτίωση της διαγνωστικής τους ακρίβειας. Σε αυτήν την βιβλιογραφική αναφορά, αναπτύχθηκε ένα αυτόματο ασαφές διαγνωστικό σύστημα το οποίο βασίζεται στους γενετικούς αλγορίθμους, στην βελτιστοποίηση σμήνους σωματιδίων καθώς και στην βελτιστοποίηση πολλαπλών σωματιδίων σμήνων. Στόχος του ήταν η δημιουργία ενός συστήματος το οποίο θα μπορούσε να διαγνώσει και το επίπεδο του κινδύνου για καρδιακή νόσο. Το διαγνωστικό σύστημα, θα λειτουργούσε με τον ακόλουθο τρόπο.

Αρχικά θα γινόταν μια προεπεξεργασία του συνόλου δεδομένων, μετά θα γινόταν επιλογή χαρακτηριστικών, μέσα από στατιστικές μεθόδους όπως ο συντελεστής συσχέτισης, η μέθοδος R-Squared και Weighted Least Squared. Στην συνέχεια οι ασαφείς κανόνες επιλέγονταν με βάσει τα χαρακτηριστικά. Τέλος για να διαπιστωθεί το πόσο αποτελεσματικό ήταν το σύστημα και αν υπάρχει δυνατότητα εφαρμογής, έγινε μια ποσοτική, ποιοτική και συγκριτική ανάλυση. Η ανάλυση έγινε μέσα από διάφορα σύνολα δεδομένων, που υπάρχουν στην πραγματική ζωή. Μέσα από εμπειρική ανάλυση, παρατηρήθηκε ότι το υδριβικό μοντέλο, μπορεί να διαχειριστεί με ακρίβεια την γνώση και να αμβλύνει την αβεβαιότητα λήψης αποφάσεων. Φυσικά κάτι τέτοιο έχει επιφέρει μεγαλύτερη ακρίβεια, στα διαφορετικά διαθέσιμα δημόσια σύνολα δεδομένων, που αφορούν τις καρδιακές παθήσεις, συγκριτικά με άλλες υπάρχουσες μεθόδους.

11. Διεξαγωγή χαρακτηριστικών μέσω παράλληλης πιθανοτικής ανάλυσης κύριων συστατικών με στόχο την διάγνωση των καρδιακών παθήσεων [28]

Η αυτοματοποιημένη διάγνωση των ασθενειών, επιτυγχάνεται κυρίως, μέσω των συστημάτων υποστήριξης αποφάσεων. Το πόσο αποδοτικά είναι αυτά τα συστήματα εξαρτάται σε έναν πολύ μεγάλο βαθμό, από την επιλογή των σχετικών χαρακτηριστικών. Αυτή η διαδικασία μπορεί να γίνει πολύ δύσκολη, όταν στο σύνολο δεδομένων υπάρχουν ελλείψεις στα χαρακτηριστικά. Λέγεται ότι η πιθανοτική ανάλυση των κύριων συνιστωσών, μπορεί να αντιμετωπίσει με επιτυχία το πρόβλημα των χαμένων τιμών των χαρακτηριστικών. Σε αυτή την βιβλιογραφική αναφορά χρησιμοποιήθηκαν, τα αποτελέσματα ιατρικών δοκιμών ως είσοδος και σαν έξοδο έπαιρνε ένα υποσύνολο χαρακτηριστικών μειωμένων διαστάσεων, περέχοντας διάγνωση καρδιακών παθήσεων. Σαν μέθοδος εξήγαγε

χαρακτηριστικά με υψηλό αντίκτυπο σε ένα πρόβλημα, χρησιμοποιώντας την πιθανοτική ανάλυση των κύριων συνιστωσών, όπως προαναφέρθηκε και πιο πάνω. Η πιθανοτική ανάλυση των κύριων συνιστωσών, λειτούργησε με τον ακόλουθο τρόπο, εξήγαγε διανύσματα προβολής τα οποία συνέβαλαν στην υψηλότερη συνδιακύμανση. Τα διανύσματα προβολής χρησιμοποιήθηκαν για την μείωση της διάστασης του χαρακτηριστικού. Η παράλληλη ανάλυση ήταν ο τρόπος με τον οποίο έγινε, η επιλογή των διανυσμάτων προβολής. Σε μηχανές υποστήριξης διανυσμάτων, με βάση τη συνάρτηση ακτινικής βάσης, έμπαινε το υποσύνολο των χαρακτηριστικών με τη μειωμένη διάσταση. Οι μηχανές υποστήριξης διανυσμάτων, που βασίζονται στην συνάρτηση ακτινικής βάσης, έδωσαν την δυνατότητα ταξινόμησης, σε περισσότερες από μια κατηγορίες, όπως για παράδειγμα σε καρδιοπαθείς και σε φυσιολογικό υποκείμενο. Η μεθοδολογία αξιολογήθηκε μέσω της της ακρίβειας, της ευαισθησίας και του speciality, σε τρία σύνολα δεδομένων του UCI. Τα σύνολα δεδομένων ήταν του Κλίβελαντ, της Ελβετίας και της Ουγγαρίας. Έπειτα από την σύγκριση, αποδείχθηκε ότι η τεχνική αυτή πέτυχε, ακρίβεια 82,18% για τα δεδομένα του Κλίβελαντ, 85,82% για της Ουγγαρίας και 91,30% για τα δεδομένα της Ελβετίας.

12. Μια μελέτη για τα συστήματα πρόβλεψης της καρδιακής ανεπάρκειας [29]

Παρόλου που ένα ποσοστό νοσηρότητας και θνησιμότητας στην καρδιακή ανεπάρκεια, μπορεί να προληφθεί, παραμένει μια από τις πιο συχνές αιτίες θανάτου σε γυναίκες και άνδρες. Αυτό από μόνο του, αποδεικνύει την πολυπλοκότητα της καρδιακής νόσου. Η καρδιαγγειακή νόσος είναι ένα επίμονο πρόβλημα, το οποίο απαιτεί, την υποστήριξη από ένα σύστημα. Σε αυτή την έρευνα βλέπουμε με ποιον τρόπο, μια συστηματική προσέγγιση μπορεί να είναι ένα χρήσιμο εργαλείο, για την κατανόηση της δομής των σχέσεων, που χαρακτηρίζουν και καθιστούν την καρδιαγγειακή νόσο τόσο επίμονη και ανθεκτική στις αλλαγές. Ένας ακόμα στόχος της εργασίας που μελετάμε, ήταν ο εντοπισμός πιθανών σημείων μόχλευσης, της αποτελεσματικής πρόληψης και αντιμετώπισης των καρδιαγγειακών παθήσεων. Για την καθοδήγηση της διαδικασίας έρευνας, χρησιμοποιήθηκε η μεθοδολογία μαλακών συστημάτων. Μέσα από μια ανασκόπηση σε βιβλιογραφικές αναφορές καθοδηγήθηκε η κατασκευή ενός συστήματος καρδιαγγειακών παθήσεων και εντοπίστηκαν τα βασικά σημεία. Η έρευνα επικεντρώθηκε σε τρεις βασικούς μοχλούς. Ο πρώτος αφορά το άγχος, ο δεύτερος την εκπαίδευση και ο τρίτος την αγοραστική δύναμη. Οι τρεις αυτοί μοχλοί μπορούν να χρησιμοποιηθούν για την αντιμετώπιση των καρδιαγγειακών παθήσεων, δίνοντας μια οπτική αναπαράσταση, στο τρόπο με τον οποίο τροποποιήσεις, σε αυτές τις πτυχές, θα μπορούσαν να έχουν αντίκτυπο σε άλλες συνιστώσες, που επηρεάζουν τον καρδιακό κίνδυνο. Η έρευνα καταλήγει λοιπόν στο συμπέρασμα πως να δοθεί περισσότερη έμφαση σε αυτούς τους “μοχλούς”, οι παρεμβάσεις στην

δημόσια υγεία θα έχουν ένα καλύτερο αποτέλεσμα, όσον αφορά την μείωση της νοσηρότητας και της θνησιμότητας από καρδιαγγειακά νοσήματα.

13. Δίαιτα τύπου DASH και κίνδυνος για καρδιαγγειακές παθήσεις. [30]

Πρόκειται για μια 10 ετής παρακολούθηση μιας μελέτης. Η μελέτη ξεκίνησε όταν ευρήματα έδειξαν, ότι το διατροφικό πρότυπο DASH, παρέχει έναν προστατευτικό ρόλο απέναντι στην συχνότητα εμφάνισης και θνησιμότητας των καρδιαγγειακών παθήσεων. Η μελέτη έγινε σε ένα υπόδειγμα από την επιδημιολογική μελέτη ATTICA, αποτελούμενη από 669 ενήλικες με πλήρες διατροφικό προφίλ, κατά την έναρξη και επαρκές για την βαθμολογία της διαίτας DASH. Η παρακολούθηση όπως αναφέρθηκε, διήρκεσε δέκα έτη. Ξεκίνησε το 2002 και ολοκληρώθηκε το 2012. Ο τρόπος ζωής και παράμετροι όπως, δημογραφικοί και κλινικοί, αξιολογήθηκαν διεξοδικά κατά την έναρξη και κατά τις επιπτώσεις, οι οποίες καταγράφηκαν αναλυτικά στα ιατρικά αρχεία. Η τήρηση της δίαιτας τύπου DASH αξιολογήθηκε με βάση την αρχική δίαιτα τύπου DASH, η οποία αναπτύχθηκε για την μελέτη. Τα αποτελέσματα της έρευνας έδειξαν, πως η μέση τιμή της βαθμολογίας της δίαιτας DASH ήταν 27,1 (5,1), εύρος 13-41. Η τήρηση μια δίαιτας τέτοιου τύπου δεν σχετίζεται ούτε με τον δεκαετή κίνδυνο καρδιαγγειακής νόσου, ούτε με τις βασικές κλινικές παραμέτρους. Συνεπώς σε ένα μικρό δείγμα ενός μεσογειακού πληθυσμού, δεν ανιχνεύθηκε προστασία κατά των καρδιακών παθήσεων, μέσα από μια δίαιτα τύπου DASH.

14. Ένα σύστημα βασισμένο στον πλησιέστερο γείτονα για την πρόβλεψη της καρδιακής ανεπάρκειας. [31]

Αυτή η εργασία μελετά ένα σύστημα που βασίζεται σε αποστάσεις και στην μέθοδο του k-πλησιέστερου γείτονα. Με βάση τα αποτελέσματα από την εφαρμογή αυτών, γίνεται η διάγνωση των καρδιακών παθήσεων. Το σύστημα αυτό έχει υλοποιηθεί με δύο διαφορετικές διαμορφώσεις. Η πρώτη χρησιμοποιεί τρεις αποστάσεις, ενώ η δεύτερη χρησιμοποιεί πέντε αποστάσεις. Στην συνέχεια προστέθηκε μια σταθμισμένη απόδοση, η οποία βασιζόταν στην μέση ακρίβεια, που έδινε κάθε απόσταση όταν γινόταν εφαρμογή της μεθόδου k-πλησιέστερου γείτονα. Η συνολική απόδοση αυτού του συστήματος, έδωσε μέση ακρίβεια κοντά στο 85% . Η ακρίβεια αυτή, αφορούσε οποιαδήποτε από τις διαμορφώσεις και τις εκδόσεις που δοκιμάστηκαν, με το σύνολο δεδομένων των καρδιακών παθήσεων του Κλίβελαντ.

15. Ανάπτυξη ενός έξυπνου συστήματος πρόβλεψης των καρδιακών παθήσεων με την χρήση των αλγορίθμων ταξινόμησης [32]

Δεδομένου ότι στην πληροφορική, της υγειονομικής περίθαλψης, η προγνωστική ικανότητα για την εκτίμηση του καρδιακού κινδύνου αποτελεί μια από τις μεγαλύτερες προκλήσεις, η εργασία αυτή πραγματεύεται ένα τέτοιο ζήτημα. Η προγνωστική μοντελοποίηση, μέσω soft υπολογιστικών εργαλείων, αποτελεί μια από τις πιο οικονομικές επιλογές που διαθέτει η ιατρική. Συνεπώς τα soft υπολογιστικά εργαλεία είναι μια μεγάλη ανάγκη για την υγειονομική περίθαλψη. Έχουν την δυνατότητα να μοντελοποιούν και να αναλύουν δεδομένα, ενώ παράλληλα βοηθούν τον γιατρό να λάβει γρήγορα, σωστές και ορθές κλινικές αποφάσεις. Απαραίτητη για εξαγωγή προτύπων, που συνδέουν τις μεταβλητές του προγνωστικού, με μια βάση δεδομένων, είναι η ύπαρξη τεχνικών εξόρυξης δεδομένων. Οι τεχνικές που υπάρχουν μέχρι σήμερα, για την εξόρυξη δεδομένων είναι οι κατάλληλες, για την μοντελοποίηση πολύπλοκων δυναμικών διαδικασιών.

Σε αυτή την βιβλιογραφική αναφορά, που εξετάζουμε, γίνεται μια πρόταση σχετικά με την προσέγγιση του μοντέλου συνόλου, για τον συνδυασμό της προγνωστικής ικανότητας ενός μοντέλου με πολλούς ταξινομητές, για καλύτερη ακρίβεια πρόβλεψης. Συγκεκριμένα, σε αυτή την μελέτη συνδυάστηκαν πέντε προσεγγίσεις μοντέλων ταξινομητών, συμπεριλαμβανομένου του τεχνητού νευρωνικού δικτύου, του Naïve Bayesian, της μηχανής υποστήριξης, της ανάλυσης παλινδρόμησης και του random forest, για την πρόβλεψη και τη διάγνωση της καρδιαγγειακής νόσου. Τα δεδομένα, που αφορούσαν τις καρδιαγγειακές παθήσεις, του Cleveland και του Hungarian, εξήχθησαν μέσα από το αποθετήριο UCI. Τα αποτελέσματα από τις πειραματικές δοκιμές, έδειξαν πως το μοντέλο συνόλου, είναι μια καλή προσέγγιση, όσον αφορά την υψηλή προγνωστική ακρίβεια και την αξιοπιστία των επιδόσεων του, σε διαγνωστικό επίπεδο. Ωστόσο η έρευνα παρουσιάζει ένα έξυπνο σύστημα πρόβλεψης των καρδιακών παθήσεων, ως μια οικονομική και άμεση προγνωστική επιλογή, με εύχρηστη και φιλική γραφική διεπαφή για τον χρήστη.

16. Ανίχνευση συμφορητικής καρδιακής ανεπάρκειας με χρήση μηχανικής μάθησης [33]

Ο τρόπος κατά τον οποίο προσαρμόζεται η καρδιά, σε εσωτερικά και εξωτερικά ερεθίσματα, εξαρτάται από τη μεταβλητότητα του καρδιακού ρυθμού. Η μειωμένη μεταβλητότητα του καρδιακού ρυθμού, μπορεί να αποτελέσει έναν προγνωστικό παράγοντα, αρνητικών καρδιαγγειακών εκβάσεων. Σε αυτή την μελέτη, προτάθηκε ένα αυτοματοποιημένο σύστημα, το οποίο θα αναλύει τα σήματα της μεταβλητότητας του καρδιακού ρυθμού, εξάγοντας με αυτόν τον τρόπο, πολυτροπικά χαρακτηριστικά για την καταγραφή φασματικών, χρονικών και σύνθετων δυναμικών. Προκειμένου να αξιολογηθεί η απόδοση ανίχνευσης χρησιμοποιήθηκαν, ισχυρές τεχνικές μηχανικής μάθησης, όπως είναι η μηχανή

υποστήριξης διανυσμάτων με πυρήνα γραμμικό, γκαουσιανό, πολυωνύμου και ακτινική συνάρτησης βάσης. Ακόμα χρησιμοποιήθηκαν δέντρα απόφασης, η τεχνική του k-πλησιέστερου γείτονα και ταξινομητές συνόλου. Η απόδοση αξιολογήθηκε με βάση τη θετική προγνωστική τιμή, την αρνητική προγνωστική τιμή, το specificity, την ευαισθησία και τις τιμές που ήταν κάτω από την θετική προγνωστική καμπύλη λειτουργίας του δέκτη. Η υψηλότερη απόδοση σημειώθηκε, χρησιμοποιώντας γραμμικό πυρήνα, με συνολική ακρίβεια 93,1% , απόδοση χαρακτηριστικής καμπύλης δέκτη 0,97, διάστημα εμπιστοσύνης 95% [κάτω όριο ίσο με 0,04 και ανω όριο ίσο με 0,89]. Αμέσως μεγαλύτερη απόδοση, είχε το διαχωριστικό υποχώρου συνόλου, με συνολική ακρίβεια 91,4%, με χαρακτηριστική καμπύλη δέκτη 0,96, διάστημα εμπιστοσύνης 95% (κάτω όριο ίσο με 0,07 και άνω όριο ίσο με 0,81). Την χαμηλότερη απόδοση, είχε η μηχανή υποστήριξης διανύσματος, με γκαουσιανό πυρήνα με συνολική ακρίβεια 90,5%, καμπύλη δέκτη 0,96 , διάστημα εμπιστοσύνης 95% (κάτω όριο ίσο με 0,07 και άνω όριο ίσο με 0,86). Συνεπώς κατέληξε στο συμπέρασμα, ότι η προσέγγιση που πρότείνει, μπορεί να προσφέρει ένα αποτελεσματικό και συγχρόνως υπολογιστικά αποδοτικό εργαλείο, για την αυτόματη ανίχνευση ασθενών, με συμφορητική καρδιακή ανεπάρκεια.

17. Πρόβλεψη της καρδιακής ανεπάρκειας με δέντρα απόφασης [34]

Οι επιστήμονες, προκειμένου να καταφέρουν να προβλέψουν κάθε τύπο καρδιοπάθειας, σε πρώιμο στάδιο, αναζητούν διαρκώς διάφορες υπολογιστικές τεχνικές. Με την χρήση τεχνικών εξόρυξης δεδομένων, καταφέρνουν να μειώνουν τον αριθμό των εξετάσεων, που απαιτείται να κάνει ο ασθενής, για την ανίχνευση της καρδιακής ανεπάρκειας. Σε αυτή την μελέτη προτάθηκε μια τεχνική υβριδισμού, στην οποία χρησιμοποιούνται, ταξινομητές δέντρων αποφάσεων και τεχνητού νευρωνικού δικτύου, με στόχο την καλύτερη απόδοση στην πρόβλεψη των καρδιακών παθήσεων. Το σύστημα το οποίο χρησιμοποιήθηκε, για την εφαρμογή αυτών, ήταν το WEKA. Προκειμένου να αξιολογήσουν την απόδοση του αλγορίθμου, πραγματοποιήθηκε μια δοκιμή (tenfold validation test), στο σύνολο δεδομένων των ασθενών με καρδιοπάθεια, το οποίο λήφθηκε από το αποθετήριο UCI. Συγκεκριμένα αναλύθηκε η ακρίβεια, η ευαισθησία και το specificity τόσο του μεμονωμένου ταξινομητή όσο και της υβριδικής τεχνικής. Τα αποτελέσματα από την έρευνα έδειξαν ότι, το υβριδικό μοντέλο του αλγορίθμου των δέντρων απόφασης, όπως και τα περισσότερα συνδυαστικά μοντέλα, είναι ικανό να δημιουργήσει ένα σύστημα για την πρόβλεψη των καρδιακών παθήσεων, σε πρώιμο στάδιο, το οποίο θα βοηθάει τους γιατρούς.

18. Μοντέλα μηχανικής μάθησης για την πρόβλεψη-βαθμολογία του καρδιακού κινδύνου [35]

Οι έρευνες για την πρόβλεψη του καρδιακού κινδύνου στρέφονται όλο και πιο πολύ, στα συστήματα που στηρίζονται στην μηχανική μάθηση. Έτσι και αυτή η έρευνα έτεινε να προτείνει ένα πλαίσιο μοντελοποίησης, το οποίο βασίζεται κυρίως στη μηχανική μάθηση. Στην ανάλυση του καρδιακού κινδύνου, στην εργασία αυτή, έχουν ληφθεί υπόψη διάφοροι παράγοντες, όπως το ύψος, η αρτηριακή πίεση, η ηλικία, η καπνιστική συνήθεια και τα επίπεδα χοληστερόλης. Ανά διαστήματα έχουν εξεταστεί διεξοδικά πολλοί κλινικοί παράμετροι, σε συνδυασμό με τις διακυμάνσεις, στα σήματα ECG και σε άλλες παθήσεις, που σχετίζονται με την καρδιά, προκειμένου να αξιολογηθούν και να προσδιοριστούν, οι παράγοντες κινδύνου καρδιαγγειακής νόσου. Για την διαδικασία πρόβλεψης, χρησιμοποιήθηκαν λίγοι παράμετροι σημάτων ECG και σύνολα δεδομένων ζωτικών παραμέτρων, με στόχο την δημιουργία ενός καλά καθορισμένου μοντέλου, βαθμολογίας κινδύνου καρδιακής ανεπάρκειας με βάση την μηχανική μάθηση. Χρησιμοποιήθηκαν επίσης, καθιερωμένοι αλγόριθμοι μηχανικής μάθησης, ταξινομητές και αλγόριθμοι επιλογής. Στην συνέχεια έγινε μια σύγκριση της απόδοσης των μοντέλων αυτών. Τα άτομα που εξετάστηκαν, ανάλογα με τον κίνδυνο που διέτρεχαν, κατατάχθηκαν σε τρεις διαφορετικές κατηγορίες, την χαμηλή επικινδυνότητα, την μέτρια και την υψηλή. Τα χαρακτηριστικά που διεξήχθησαν από τα σήματα ECG, χρησιμοποιήθηκαν στην τροφοδοσία του συστήματος. Τα αποτελέσματα έδειξαν, τα νευρωνικά δίκτυα να ξεπερνούν τα άλλα μοντέλα, με ακρίβεια 97% . Ένα ακόμα πρόβλημα αυτής έρευνας, ήταν πως τα μοντέλα που προτάθηκαν, τα οποία βασίστηκαν στην μηχανική μάθηση, θα βοηθήσουν σημαντικά τους γιατρούς στην διάγνωση της καρδιαγγειακής νόσου.

19. Αναπαράσταση κίνησης πολλαπλών κλιμάκων για την πρόβλεψη της καρδιακής ανεπάρκειας [36]

Η καρδιά είναι ένα από τα πιο δύσκολα όργανα, που μπορεί κανείς να χαρακτηρίσει και πολλές φορές αποτελεί ένα challenge για την επιστήμη. Αυτό οφείλεται στην μη γραμμική απόδοση της και στις έντονες παραμορφώσεις που έχει το σχήμα της, κατά τη διάρκεια του καρδιακού κύκλου. Η δέκατη ένατη έρευνα που μελετήσαμε, παρουσίασε μια περιφερειακή αναπαράσταση, πολλαπλής κίνησης των καρδιακών δομών, που είναι σε θέση να αναγνωρίσουν παθολογίες, σε μαγνητικές τομογραφίες. Αρχικά χρησιμοποιήθηκε μια πυκνή οπτική ροή που θεωρεί μεγάλες μετατοπίσεις, για να ληφθεί μια αναπαράσταση του πεδίου ταχύτητας. Έπειτα, κωδικοποιήθηκαν τα περιφερειακά δυναμικά μοτίβα, σε ένα σχήμα πολλαπλών κλιμάκων, από χοντρό έως λεπτό, με στόχο την εύρεση και ανάδειξη, των πιο σχετικών καρδιακών μοτίβων, που παρέμειναν κατά μήκος των διαφορετικών

κλιμάκων. Μέσα από ένα σύνολο εμφανίσεων, προσανατολισμού ροής, που υπολογίστηκαν σε ολόκληρες περιοχές πολλαπλής κλίμακας, προέκυψε ένας περιγραφέας κίνησης. Αυτός ο περιγραφέας κίνησης, στην συνέχεια, αντιστοιχίζεται σε ένα ήδη εκπαιδευμένο ταξινομητή random forest, έτσι ώστε να ληφθεί μια πρόβλεψη, σχετικά με την καρδιακή κατάσταση. Προκειμένου να αξιολογηθεί το συγκεκριμένο μοντέλο, εφαρμόστηκαν και αξιολογήθηκαν δύο πειράματα. Τα αποτελέσματα φανερώνουν πως η στρατηγική αυτή είναι ισχυρή, λόγω του ότι οι πληροφορίες που συλλέγονται μέσω των πολλαπλών κλιμάκων, είναι σε θέση να αξιολογήσουν σωστά τις κινήσεις της καρδιάς κατά την διάρκεια του καρδιακού κύκλου, επιτυγχάνοντας με αυτόν τον τρόπο, καλή πρόβλεψη της καρδιακής νόσου.

20. Χρονική συχνότητα και υβριδικό νευρωνικό δίκτυο για την ανίχνευση και ταξινόμηση των καρδιακών ανωμαλιών [37]

Είναι γεγονός, πως μια κατάλληλη χρονική και φασματική επεξεργασία των σημάτων ηλεκτροκαρδιογραφήματος (ECG), μπορεί να διευκολύνει την οπτική ερμηνεία και να βοηθήσει στην διάκριση γνωστών προτύπων για ταξινόμηση. Αυτή η μελέτη πρότεινε μια μέθοδο, μη επεμβατική. Η μέθοδος αυτή, βασίστηκε σε ένα υβριδικό νευρωνικό δίκτυο και στη χρονική συχνότητα. Στόχος της ήταν, η ανίχνευση και η ταξινόμηση των καρδιακών ανωμαλιών, στα ηλεκτροκαρδιογραφήματα. Στις καρδιακές ανωμαλίες συμπεριλήφθηκε, η συμφορητική καρδιακή ανεπάρκεια, η κοιλιακή ταχυαρρυθμία, η αρρυθμία κακοήθους κολπικής γαστρεντερίτιδας, η ενδοκαρδιακή κολπική μαρμαρυγή, ο ρυθμός και οι ταλαντώσεις μετά τον παλμό της καρδιάς σε μερική επιληψία. Τα μη σταθερά, ακατέργαστα σήματα ECG κρίθηκαν και αξιολογήθηκαν από την ηλεκτρονική πηγή δεδομένων υγειονομικής περίθαλψης «PhysioBank», που περιείχε φυσιολογικά σήματα. Αυτά τα χρονικά σήματα, υποβάλλονταν σε επεξεργασία μέσω της διανομής Wigner-Ville, για να γίνει μια υψηλή ανάλυση. Στην συνέχεια οι χρονικές συχνότητες, χρησιμοποιούνταν για την εξαγωγή των παραμέτρων ανωμαλίας, με την βοήθεια των ειδικών ιατρών, που κατείχαν καλή διαγνωστική ακρίβεια. Για την μείωση χαρακτηριστικών και την επιλογή σημαντικών χαρακτηριστικών από τα σήματα ECG, χρησιμοποιούνταν η ανάλυση κύριου συστατικού. Τα χαρακτηριστικά που επιλέγονταν, χρησιμοποιούνταν για την εκπαίδευση του τεχνητού νευρωνικού δικτύου, πολλαπλών επιπέδων τροφοδοσίας, για ανίχνευση και ταξινόμηση κατά την φάση της εκπαίδευσης παραμέτρων, όπως οι συναρτήσεις εκπαίδευσης και ο αριθμός των εποχών. Το ποσοστό εκμάθησης επιλέγονταν, με την βοήθεια κατάλληλων κριτηρίων διακοπής. Τα αποτελέσματα των πειραμάτων που πραγματοποιήθηκαν, απέδειξαν ότι η προσέγγιση του υβριδικού νευρωνικού δικτύου, με χρονική συχνότητα και ανάλυση του κύριου συστατικού, είναι αποτελεσματική για την ανίχνευση και την ταξινόμηση καρδιαγγειακών ανωμαλιών.

Ε'2 Πίνακας Σύγκρισης Βιβλιογραφικών Αναφορών

ΑΝΑΦΟΡΑ	ΑΚΡΙΒΕΣ ΠΡΟΒΛΗΜΑ	ΜΕΘΟΔΟΛΟΓΙΑ	ΠΛΗΘΟΣ	ΑΠΟΤΕΛΕΣΜΑ	ΕΤΟΣ
(1)	Πρόβλεψη καρδιαγγειακών παθήσεων και εύρεση των πιο σημαντικών προγνωστικών παραγόντων για τις περισσότερες από αυτές ξεχωριστά.	Random Forest	Συμμετείχαν 6814, ασυμπτωματικά άτομα Ελήφθησαν 735 αποτελέσματα	Η μηχανική μάθηση μαζί με το βαθύ φαινότυπο, βελτιώνουν την ακρίβεια των προβλέψεων, στις καρδιαγγειακές παθήσεις.	2017
(2)	Συστηματική ανασκόπηση της αυτοματοποιημένης διάγνωσης για την πρόβλεψη καρδιακών παθήσεων με διαφορετικούς τύπους μεθόδων	Μηχανική Μάθηση Και Εξόρυξη Δεδομένων	Αξιολόγηση και επέκταση προηγούμενων μεθόδων	Τρεις βασικοί παράγοντες συμμετέχουν για την αποτελεσματική ανίχνευση της καρδιακής ανεπάρκειας.	2022
(3)	Αποτελεσματική πρόβλεψη της καρδιακής ανεπάρκειας	Συνδυασμός δύο μοντέλων μηχανών διανυσμάτων υποστήριξης (Support vector machine (SVM) models)	Δοκιμές και σύγκριση με προηγούμενες μεθόδους.	Η μέθοδος βελτιώνει την απόδοση ενός συμβατικού μοντέλου, κατά 3,3% Δείχνει καλύτερη απόδοση, σε σύγκριση με προηγούμενες μεθόδους που είχαν ακρίβεια 57,85% με 91,83%.	2019
(4)	Αποτελεσματική πρόβλεψη της	Χρήση δύο σταδίων: 1) στατιστικό μοντέλο, που κατατάσει 13 χαρακτηριστικά, της καρδιακής ανεπάρκειας και αναζητά τα καλύτερα	Βάση δεδομένων με 298 άτομα με καρδιακή	Είχε ακρίβεια πρόβλεψης, έως και 93,33% . Βελτιώνει την απόδοση πρόβλεψης, κατά 3,33%. Εμφάνισε καλύτερη απόδοση από τις διαθέσιμες μεθόδους της	2019

	καρδιακής ανεπάρκειας	αποτελέσματα με τον αλγόριθμο best-first search strategy 2) Gaussian Naive Bayes, ως μοντέλο πρόβλεψης	ανεπάρκεια.	βιβλιογραφίας και είχαν ακρίβεια πρόβλεψης από 57,85% έως και 92,22%.	
(5)	Αποτελεσματική πρόβλεψη της καρδιακής ανεπάρκειας	Εφαρμογή δύο σταδίων: 1) Εφαρμογή στατικού μοντέλου, βασισμένο σε πληροφορίες 2) χρήση νευρωνικού δικτύου	Πειραματικές δοκιμές σε υποσύνολο της δημόσιας, βάσης δεδομένων καρδιακών παθήσεων, του Κλίβελαντ	Μειωμένο ποσοστό σφάλματος στο 8,8% Παρουσίασε ποσοστό ακρίβειας 93,33% , υψηλότερο από άλλα 28 πρόσφατα μοντέλα πρόβλεψης καρδιακής ανεπάρκειας, με ακρίβεια 57,85%-92,31%.	2021
(6)	Πρόβλεψη επιπολασμού της καρδιακής νόσου	Συνδυασμός από έξι τεχνικές μηχανικής μάθησης,	Αξιολογήθηκε μέσα από 8 διαφορετικές ταξινομήσεις δεικτών απόδοσης (classification performance indices) και με βάση την καμπύλη Receiver operating characteristic (ROC)	Η πιο υψηλή ταξινόμηση είχε ακρίβεια 85%, χρησιμοποιούσε λογιστική παλινδρόμηση, είχε ευαισθησία 89% και specificity 81%	2018
(7)	Ανίχνευση των καρδιακών παθήσεων. Προσπάθεια, εξισορρόπησης της ακρίβειας των δεδομένων εκπαίδευσης και δοκιμών.	Εφάρμοσε τον αλγόριθμο τυχαίας αναζήτησης / random search algorithm, για την επιλογή χαρακτηριστικών και τον random forest, για την πρόβλεψη της καρδιακής ανεπάρκειας. Η βελτιστοποίηση επιτεύχθηκε με τον αλγόριθμο grid	Εκτελέστηκαν δύο είδη πειραμάτων. Στο πρώτο πείραμα αναπτύχθηκε μόνο ο random forest, ενώ στο δεύτερο πείραμα αναπτύχθηκε ο random forest βασισμένος στην τυχαία αναζήτηση. Τα πειράματα εκτελέστηκαν, βασισμένα σε ηλεκτρονική βάση	Η μέθοδος, έδωσε 3,3% μεγαλύτερη ακρίβεια, από την απλή εφαρμογή του random forest, χρησιμοποιώντας μόνο 7 χαρακτηριστικά. Εφερε καλύτερες επιδόσεις από άλλες 5 μεθόδους που στηρίχθηκαν στα μοντέλα μηχανικής μάθησης. Πέτυχε ακρίβεια ταξινόμησης 93,33%, ενώ παράλληλα βελτίωσε και την ακρίβεια των δεδομένων εκπαίδευσης. Παρουσίασε καλύτερη απόδοση, συγκριτικά με	2019

		search.	δεδομένων, του Cleveland.	άλλες 11 πρόσφατες μεθόδους, ανίχνευσης της καρδιακής ανεπάρκειας.	
(8)	Πρόβλεψη των καρδιαγγειακών παθήσεων σε πρώιμο στάδιο	Ταξινόμηση συνόλου	Χρήση ενός συνόλου από δεδομένα καρδιακών παθήσεων και στην συνέχεια ανάλυση.	Οι τεχνικές συνόλου, είναι αποτελεσματικές στη βελτίωση της ακρίβειας πρόβλεψης των αδύναμων ταξινομητών και παρουσίαζαν ικανοποιητικές επιδόσεις και στον εντοπισμό του κινδύνου για κάποια καρδιακή νόσο. Επιτεύχθηκε μια αύξηση ακρίβειας 7%, στην πρόβλεψη των καρδιακών παθήσεων. Η απόδοση, βελτιώθηκε με την εφαρμογή, επιλογής χαρακτηριστικών, ενώ τα αποτελέσματα έδειξαν σημαντικές βελτιώσεις, στην ακρίβεια της πρόβλεψης.	2019
(9)	Πρόβλεψη καρδιακών προβλημάτων, μέσω προγνωστικής ανάλυσης σημάτων ECG	Προγνωστικό πλαίσιο δύο σταδίων για την επεξεργασία του σήματος ηλεκτροκαρδιογραφήματος (ECG). Εμφάνιση κίτρινου σε περίπτωση κινδύνου για καρδιοπάθειες και κόκκινο σε υψηλό κίνδυνο.	Αξιολογείται ως προς την ακρίβεια ταξινόμησης σε σύγκριση με τα αποτελέσματα της τελευταίας τεχνολογίας.	Ακρίβεια ταξινόμησης 96,6% και παρέχει ένα μοναδικό χαρακτηριστικό, αυτό της προγνωστικής ανάλυσης, δημιουργώντας προειδοποιητικά μηνύματα, σχετικά με τον κίνδυνο καρδιακών ανωμαλιών.	2019
	Αξιολόγηση του επιπέδου κινδύνου της	Προεπεξεργασία του συνόλου δεδομένων, μετά επιλέγονται χαρακτηριστικά μέσα από στατιστικές μεθόδους όπως ο συντελεστής	Για να διαπιστωθεί το πόσο αποτελεσματικό είναι και αν υπάρχει δυνατότητα εφαρμογής, γίνεται μια ποσοτική,	Μέσα από εμπειρική ανάλυση, παρατηρήθηκε ότι το υδριβικό μοντέλο, μπορεί να διαχειριστεί με ακρίβεια την γνώση και να αμβλύνει την αβεβαιότητα λήψης αποφάσεων. Έχει	

(10)	καρδιακής νόσου μέσα από ένα σύστημα βασισμένο σε κανόνες	συσχέτισης, η μέθοδος R-Squared και Weighted Least Squared. Οι ασαφείς κανόνες επιλέγονται με βάση επιλεγμένων χαρακτηριστικών.	ποιοτική και συγκριτική ανάλυση, μέσα από διάφορα σύνολα δεδομένων της πραγματικής ζωής.	επιφέρει μεγαλύτερη ακρίβεια, στα διαφορετικά διαθέσιμα δημόσια σύνολα δεδομένων, που αφορούν τις καρδιακές παθήσεις, συγκριτικά με άλλες υπάρχουσες μεθόδους.	2018
(11)	Πρόβλεψη της καρδιακής ανεπάρκειας	Ως είσοδο πήραν αποτελέσματα ιατρικών δοκιμών και σαν έξοδο έπαιρνε ένα υποσύνολο χαρακτηριστικών, μειωμένων διαστάσεων, περέχοντας διάγνωση καρδιακών παθήσεων. Εξήγαγε χαρακτηριστικά με υψηλό αντίκτυπο, χρησιμοποιώντας την πιθανοτική ανάλυση των κύριων συνιστωσών.	Η μεθοδολογία αξιολογήθηκε μέσω της της ακρίβειας, της ευαισθησίας και του speciality, σε τρία σύνολα δεδομένων του UCI, του Κλίβελαντ, της Ελβετίας και της Ουγγαρίας	Η τεχνική αυτή, πέτυχε ακρίβεια 82,18% για τα δεδομένα του Κλίβελαντ, 85,82% για της Ουγγαρίας και 91,30% για τα δεδομένα της Ελβετίας.	2017
(12)	Τον τρόπο κατά τον οποίο, μια συστηματική προσέγγιση, μπορεί να είναι ένα χρήσιμο εργαλείο, για την κατανόηση της δομής των σχέσεων, που χαρακτηρίζουν και καθιστούν την καρδιαγγειακή νόσο τόσο επίμονη και ανθεκτική στις αλλαγές και ο εντοπισμός πιθανών σημείων μόχλευσης, της αποτελεσματικής πρόληψης και αντιμετώπισης των	Χρησιμοποιήθηκε η μεθοδολογία μαλακών συστημάτων και μέσα από ανασκόπηση σε βιβλιογραφικές αναφορές καθοδηγήθηκε η κατασκευή ενός συστήματος.	Ανασκόπηση βιβλιογραφικών αναφορών	Η έρευνα επικεντρώθηκε σε τρεις βασικούς μοχλούς, το άγχος, την εκπαίδευση και την αγοραστική δύναμη. Οι τρεις αυτοί μοχλοί μπορούν να χρησιμοποιηθούν για την αντιμετώπιση των καρδιαγγειακών παθήσεων, δίνοντας μια οπτική αναπαράσταση, στο τρόπο με τον οποίο τροποποιήσεις, σε αυτές τις πτυχές, θα μπορούσαν να έχουν αντίκτυπο σε άλλες συνιστώσες, που επηρεάζουν τον καρδιακό	2019

	καρδιαγγειακών παθήσεων.			κίνδυνο.	
(13)	Επιρροή της διαίτας τύπου DASH στις καρδιαγγειακές παθήσεις	Εφαρμογή σε ένα δείγμα πληθυσμού δίαιτα τύπου DASH και παρακολούθηση για δέκα έτη	Υπόδειγμα από την επιδημιολογική μελέτη ATTICA, αποτελούμενη από 669 ενήλικες με πλήρες διατροφικό προφίλ	Σε ένα μικρό δείγμα, ενός μεσογειακού πληθυσμού, δεν ανιχνεύθηκε προστασία κατά των καρδιακών παθήσεων, μέσα από μια δίαιτα τύπου DASH.	2019
(14)	Πρόβλεψη καρδιοπάθειας	Μέθοδος k-πλησιέστερου γείτονα	Δύο διαφορετικές διαμορφώσεις, μια με 3 αποστάσεις και μια με 5. Μια σταθμισμένη έκδοση βασισμένη στην μέση ακρίβεια. Οι εκδόσεις δοκιμάστηκαν με το σύνολο δεδομένων των καρδιακών παθήσεων του Κλίβελαντ.	Μέση ακρίβεια κοντά στο 85% . Η ακρίβεια αυτή, αφορούσε οποιαδήποτε από τις διαμορφώσεις και τις εκδόσεις που δοκιμάστηκαν	2018
(15)	Πρόβλεψη και διάγνωση καρδιαγγειακών παθήσεων	Προσέγγιση του μοντέλου συνόλου, για τον συνδυασμό της προγνωστικής ικανότητας ενός μοντέλου με πολλούς ταξινομητές, για καλύτερη ακρίβεια πρόβλεψης.	Τα δεδομένα, που αφορούσαν τις καρδιαγγειακές παθήσεις, του Cleveland και του Hungarian, εξήχθησαν μέσα από το αποθετήριο UCI.	Το μοντέλο συνόλου, είναι μια καλή προσέγγιση, όσον αφορά την υψηλή προγνωστική ακρίβεια και την αξιοπιστία των επιδόσεων του, σε διαγνωστικό επίπεδο.	2018
	Δημιουργία αυτοματοποιημένου συστήματος, το οποίο θα αναλύει τα σήματα	Σύστημα, το οποίο ανέληξε τα σήματα της μεταβλητότητας του καρδιακού ρυθμού, εξάγοντας	Για την αξιολόγηση ανίχνευσης χρησιμοποιήθηκαν ισχυρές τεχνικές μηχανικής μάθησης, όπως η μηχανή υποστήριξης διανυσμάτων με πυρήνα γραμμικό,	Κατέληξε στο συμπέρασμα, ότι η προσέγγιση που πρότεινε, μπορεί να προσφέρει ένα	

(16)	της μεταβλητότητας του καρδιακού ρυθμού με στόχο την ανίχνευση της καρδιακής ανεπάρκειας.	με αυτόν τον τρόπο πολυτροπικά χαρακτηριστικά για την καταγραφή φασματικών, χρονικών και σύνθετων δυναμικών	γκαουσιανό, πολυωνύμου και ακτινική συνάρτησης βάσης, δέντρα απόφασης, τεχνική k-πλησιέστερου γείτονα και ταξινομητές συνόλου. Η απόδοση αξιολογήθηκε, με βάση τη θετική προγνωστική τιμή, την αρνητική προγνωστική τιμή, το specificity, την ευαισθησία και τις τιμές που ήταν κάτω από την θετική προγνωστική καμπύλη λειτουργίας του δέκτη.	αποτελεσματικό και συγχρόνως υπολογιστικά αποδοτικό εργαλείο, για την αυτόματη ανίχνευση ασθενών, με συμφορητική καρδιακή ανεπάρκεια.	2020
(17)	Πρώιμη πρόβλεψη καρδιακών παθήσεων	Τεχνική υβριδισμού, στην οποία χρησιμοποιούνται, ταξινομητές δέντρων αποφάσεων και τεχνητού νευρωνικού δικτύου, Χρησιμοποιήθηκε, WEKA.	Προκειμένου να αξιολογήσουν την απόδοση του αλγορίθμου, πραγματοποιήθηκε μια δοκιμή (tenfold validation test), στο σύνολο δεδομένων των ασθενών με καρδιοπάθεια, το οποίο λήφθηκε από το αποθετήριο UCI.	Τα αποτελέσματα έδειξαν ότι, το υβριδικό μοντέλο του αλγορίθμου των δέντρων απόφασης, είναι ικανό να δημιουργήσει ένα σύστημα για την πρόβλεψη των καρδιακών παθήσεων, σε πρώιμο στάδιο, το οποίο θα βοηθάει τους γιατρούς.	2019
(18)	Πρόβλεψη-βαθμολογία του καρδιακού κινδύνου	Χρησιμοποιήθηκαν παράμετροι σημάτων ECG & σύνολα δεδομένων ζωτικών παραμέτρων, καθιερωμένοι αλγόριθμοι μηχανικής μάθησης,	Εξέταση διάφορων ατόμων	Τα αποτελέσματα έδειξαν, τα νευρωνικά δίκτυα να ξεπερνούν τα άλλα μοντέλα, με ακρίβεια 97%. Τα μοντέλα που προτάθηκαν, βασισμένα στην μηχανική μάθηση, θα βοηθήσουν σημαντικά τους	2021

		ταξινομητές και αλγόριθμοι επιλογής.		γιατρούς στην διάγνωση της καρδιαγγειακής νόσου.	
(19)	Πρόβλεψη της καρδιακής ανεπάρκειας	Αναπαράσταση κίνησης πολλαπλών κλιμάκων. Εύρεση των πιο συχνών μοτίβων, δημιουργία περιγραφέα που αντιστοιχίζεται με ένα ήδη εκπαιδευμένο ταξινομητή random forest	Αξιολόγηση μέσα από δύο πειράματα	Η στρατηγική αυτή είναι ισχυρή, γιατί οι πληροφορίες που συλλέγονται μέσω των πολλαπλών κλιμάκων, είναι σε θέση να αξιολογήσουν σωστά τις κινήσεις της καρδιάς, κατά την διάρκεια του καρδιακού κύκλου, επιτυγχάνοντας με αυτόν τον τρόπο, καλή πρόβλεψη της καρδιακής νόσου	2019
(20)	Ανίχνευση & ταξινόμηση των καρδιακών ανωμαλιών, στα ECG για την πρόβλεψη καρδιακών παθήσεων	Προσέγγιση του υβριδικού νευρωνικού δικτύου, με χρονική συχνότητα και ανάλυση του κύριου συστατικού	Αξιολόγηση μέσω πειραμάτων	Η προσέγγιση του υβριδικού νευρωνικού δικτύου, με χρονική συχνότητα και ανάλυση του κύριου συστατικού, είναι αποτελεσματική για την ανίχνευση και την ταξινόμηση καρδιαγγειακών ανωμαλιών.	2022

Πίνακας 1: Πίνακας Σύγκρισης Βιβλιογραφικών Αναφορών

Ε.3 Αποτελέσματα Βιβλιογραφικής Ανασκόπησης

Επειτα από την βιβλιογραφική ανασκόπηση που πραγματοποιήθηκε στο προηγούμενο κεφάλαιο, σε αυτό το κεφάλαιο, θα πραγματοποιηθεί μια σύγκριση, ποιοτική και ποσοτική με βάση τις βιβλιογραφικές αναφορές που συναντήσαμε. Στο σύνολο μελετήθηκαν, είκοσι βιβλιογραφικές αναφορές, οι χρονολογίες των οποίων δεν ξεπερνούσαν τα πέντε έτη δημοσίευσης τους. Ο λόγος για τον οποίο κρίθηκε απαραίτητο, να είναι αρκετά πρόσφατες οι βιβλιογραφικές αναφορές, είναι η ραγδαία εξέλιξη που έχει τόσο ο κλάδος της ιατρικής, όσο και της πληροφορικής που αναπτύσσεται μέρα με την μέρα. Όλες οι έρευνες που μελετήθηκαν αφορούν την πρόβλεψη των καρδιακών παθήσεων, ενώ ορισμένες από αυτές ασχολούνται παράλληλα και με την ταξινόμηση του κινδύνου για κάποια καρδιακή πάθηση. Πιο αναλυτικά η πρώτη σε σειρά βιβλιογραφική ανασκόπηση, αναζητά τους πιο σημαντικούς προγνωστικούς παράγοντες, η δεύτερη κάνει μια ανασκόπηση σχετικά με την

αυτοματοποιημένη διάγνωση, η έκτη βιβλιογραφική ανασκόπηση μελετά τον επιπολασμό της νόσου, ενώ η έβδομη συνδυάζει την ανίχνευση των καρδιακών παθήσεων με την προσπάθεια εξισορρόπησης της ακρίβειας των δεδομένων εκπαίδευσης και δοκιμών. Η όγδοη και η δέκατη έβδομη ανασκόπηση ασχολούνται με την πρόβλεψη των καρδιακών παθήσεων σε πρώιμο στάδιο, η δέκατη μιλάει για τον τρόπο αξιολόγησης του επιπέδου κινδύνου για καρδιακή νόσο. Η δωδέκατη σε σειρά ανασκόπηση ασχολείται με τον τρόπο κατά τον οποίο μια συστηματική προσέγγιση, μπορεί να είναι ένα χρήσιμο εργαλείο για την κατανόηση της δομής των σχέσεων, που χαρακτηρίζουν και καθιστούν την καρδιαγγειακή νόσο τόσο επίμονη και ανθεκτική στις αλλαγές και προσπαθεί να εντοπίσει τα πιθανά σημεία μόχλευσης, της αποτελεσματικής πρόληψης και αντιμετώπισης των καρδιαγγειακών παθήσεων. Η δέκατη τρίτη μελέτη πραγματοποιείται την επιρροή που έχει η δίαιτα τύπου DASH στις καρδιαγγειακές παθήσεις. Η εικοστή, ανιχνεύει και ταξινομεί τις καρδιακές ανωμαλίες που μέσω των σημάτων ECG που στέλνονται. Οι υπόλοιπες βιβλιογραφικές αναφορές, δηλαδή η τρίτη, η τέταρτη, η πέμπτη, η ένατη, η ενδέκατη, η δέκατη τέταρτη, η δέκατη πέμπτη, η δέκατη έκτη, η δέκατη όγδοη και η δέκατη ένατη, αναζητούν και ερευνούν ένα σύστημα πρόβλεψης των καρδιαγγειακών νόσων, με κύριο μέλημα τους, να υπάρχει υψηλή αξιοπιστία.

Ανάμεσα στις βιβλιογραφικές αναφορές, που πραγματεύονται το ίδιο θέμα, δηλαδή την δημιουργία ενός συστήματος πρόβλεψης των καρδιαγγειακών νόσων, οι πιο πρόσφατες βιβλιογραφικές ανασκοπήσεις είναι η δέκατη όγδοη και η πέμπτη, οι οποίες διεξήχθησαν το 2021. Η πιο παλιά, βιβλιογραφική ανασκόπηση, με το ίδιο πάλι θέμα, είναι η ενδέκατη, που διεξήχθη το 2017. Σε αυτό το σημείο θα πρέπει να πούμε ότι, ανάμεσα στις βιβλιογραφικές αναφορές, που πραγματεύονται αποκλειστικά, την δημιουργία ενός συστήματος πρόβλεψης, των καρδιαγγειακών παθήσεων, την μεγαλύτερη ακρίβεια την έχει η δέκατη όγδοη με ποσοστό 97%. Συνεπώς, η δέκατη όγδοη είναι από τις πιο πρόσφατες αναφορές και έχει την μεγαλύτερη ακρίβεια. Τα εργαλεία που χρησιμοποιεί αυτή η αναφορά, για την πρόβλεψη των καρδιαγγειακών παθήσεων, είναι τα σήματα ηλεκτροκαρδιογραφήματος (ECG), σύνολα δεδομένων ζωτικών παραμέτρων, αλγόριθμοι μηχανικής μάθησης, ταξινομητές και αλγόριθμοι επιλογής. Οι μέθοδοι αυτοί είναι ευρέως διαδεδομένοι και δεν είναι λίγοι εκείνοι που τις χρησιμοποιούν, για τις έρευνες τους.

Την τεχνική των σημάτων ηλεκτροκαρδιογραφήματος, την συναντήσαμε ξανά, στην εικοστή και στην ένατη αναφορά. Οι διαφορές ανάμεσα στις τρεις αυτές αναφορές, είναι το έτος διεξαγωγής τους, η ακρίβεια του αποτελέσματος, το πλήθος εξέτασης και ο σκοπός. Πιο αναλυτικά στην εικοστή αναφορά, σε αντίθεση με τις άλλες δύο, γίνεται και ταξινόμηση των ανωμαλιών, εκτός από ανίχνευση. Εκτός αυτού, η δέκατη όγδοη διεξήχθη το 2021, η ένατη το 2019 ενώ η εικοστή το 2022. Η ακρίβεια των αποτελεσμάτων της εικοστής, αγγίζει σχεδόν το 100%, αφού η πιθανότητα λάθος είναι σχεδόν μηδενική. Η ακρίβεια της δέκατης όγδοης είναι 97% δηλαδή απέχει περίπου 3% από την εικοστή και η ένατη έχει ακρίβεια 96,6%, δηλαδή μια πολύ μικρή απόκλιση συγκριτικά με δέκατη όγδοη. Η αξιολόγηση των τριών μεθόδων, έγινε για την εικοστή μέσω πειραματικών δοκιμών, για την

δέκατη όγδοη με εξέταση σε διαφορά άτομα και για την ένατη με σύγκριση των αποτελεσμάτων με της τελευταίας τεχνολογίας. Ανάμεσα σε αυτές τις τρεις, που μόλις συγκρίναμε, θα έλεγε κανείς ότι υπερτερεί, με βάση τους παράγοντες που αξιολογήσαμε, η εικοστή βιβλιογραφική αναφορά.

Σε μια γενική αξιολόγηση, η αναφορά που ο τρόπος που προσεγγίζει τα καρδιαγγειακές παθήσεις, είναι πιο κοντά, με αυτόν που θα χρησιμοποιηθεί σε αυτή την εργασία, είναι η δέκατη έβδομη. Χρησιμοποιεί παρόμοια και κοινά συστήματα και αλγορίθμους, με αυτά που χρησιμοποιήθηκαν και σε αυτή την εργασία και θα δούμε αναλυτικά στην πορεία. Πρόκειται για μια προσπάθεια, του 2019, η οποία χρησιμοποιεί το σύστημα WEKA, εφαρμόζει δέντρα αποφάσεων, αξιολογει με την μέθοδο tenfold validation test και τα αποτελέσματα της φαίνονται να είναι θετικά.

Η μέθοδος, που η αξιολόγηση της είχε μεγάλο εύρος, δηλαδή περισσότερο πλήθος, ήταν η πρώτη αναφορά που μελετήθηκε το 2017, η οποία έγινε σε 6814 ασυμπτωματικά άτομα, από τα οποία ελήφθησαν 735 αποτελέσματα. Στις περισσότερες βιβλιογραφίες, η αξιολόγηση δεν γινόταν με δείγματα ατόμων, αλλά με δεδομένα UCI ή με πειράματα και βάση προηγούμενων αναφορών. Η μέθοδος, που αποδείχθηκε, πως δεν είχε σχεδόν κανένα αποτέλεσμα, στις καρδιαγγειακές παθήσεις, ήταν η δίαιτα τύπου DASH, ειδικά σε χώρες, που ακολουθούν μεσογειακή διατροφή. Αυτή η μέθοδος, έρευνήθηκε το 2017 και αξιολογήθηκε με βάση ένα υπόδειγμα από την επιδημιολογική μελέτη ATTICA, αποτελούμενη από 669 ενήλικες με πλήρες διατροφικό προφίλ. Αυτή η εργασία, αποτέλεσε την δέκατη τρίτη βιβλιογραφική αναζήτηση.

Έπειτα λοιπόν, από μια ποιοτική και ποσοτική σύγκριση ανά κατηγορίες αλλά και ανά ομάδες αναφορών, μπορεί κανείς να καταλήξει σε διάφορα συμπεράσματα. Βασικό πόρισμα, κατά την αναζήτηση των αναφορών, είναι πως οι καρδιαγγειακές παθήσεις και ακόμα περισσότερο η πρόβλεψη τους, είναι ένα θέμα που απασχολεί ιδιαίτερα την επιστήμη και καταλαμβάνει μεγάλο μέρος της βιβλιογραφίας. Οι περισσότεροι ερευνητές, ψάχνουν την λύση βασισμένοι σε τεχνικές μηχανικής μάθησης, αν και η πιο πρόσφατη ανασκόπηση, έχει δείξει πως λύση μπορεί να δώσει και ένα σύστημα βασισμένο στο ηλεκτροκαρδιογράφημα.

Κεφάλαιο ΣΤ’

Επεξεργασία & Ανάλυση Δεδομένων Για Την Εφαρμογή Πρόβλεψης Καρδιακών Παθήσεων

Έπειτα από μια ερευνητική διαδικασία, που ακολουθήθηκε στα προηγούμενα κεφάλαια, σε αυτό το κεφάλαιο, θα παρακολουθήσουμε αναλυτικά, τον τρόπο δημιουργίας και λειτουργίας μιας εφαρμογής πρόβλεψης των καρδιαγγειακών παθήσεων. Βασικός σκοπός, αυτής της εργασίας, ήταν η δημιουργία μιας web εφαρμογής, η οποία θα βοηθούσε στην πρόβλεψη των καρδιαγγειακών προβλημάτων, φανερώνοντας ένα και μόνο μήνυμα. Η αυτοματοποιημένη διάγνωση των ασθενειών, όπως έχουμε δει, επιτυγχάνεται κυρίως, μέσω των συστημάτων υποστήριξης αποφάσεων. Το πόσο αποδοτικά είναι αυτά τα συστήματα εξαρτάται, κατά έναν μεγάλο βαθμό, από την επιλογή που θα γίνει στα σχετικά χαρακτηριστικά. Αυτή η διαδικασία γίνεται ακόμα πιο πολύπλοκη και δύσκολη όταν υπάρχουν χαρακτηριστικά τα οποία μπορεί να είναι λανθασμένα ή ελλιπείς [28]. Την αντιμετώπιση τέτοιου είδους προβλημάτων και τον τρόπο με τον οποίο αναλύουμε τα δεδομένα, σκοπεύουμε να αναλύσουμε σε αυτό το κεφάλαιο. Μια εφαρμογή σαν και αυτή κρίνεται σχεδόν απαραίτητη, κυρίως σαν σύμβουλος των ιατρών. Πρόκειται για ένα σύστημα υποστήριξης αποφάσεων, για την ανάπτυξη του οποίου, χρησιμοποιήθηκαν τεχνικές μηχανικής μάθησης, η εφαρμογή WEKA, αλγόριθμοι και φίλτρα της εφαρμογής.

ΣΤ’.1 Συστήματα Υποστήριξης Αποφάσεων Στις Καρδιαγγειακές παθήσεις

Ο αντίκτυπος των καρδιαγγειακών παθήσεων, τόσο στο κοινωνικό σύνολο όσο και στον άνθρωπο μεμονωμένα είναι μεγάλος. Αυτό οφείλεται κυρίως στα προβλήματα, που υπάρχουν στα συστήματα υγείας και στις μέχρι τώρα μεθόδους που υπάρχουν. Βασικός παράγοντας αυτών, είναι η συννοσηρότητα, που προκαλούν οι καρδιακές παθήσεις, καθώς και οι χρονοβόρες διαδικασίες, με μια

σειρά εξετάσεων που πρέπει να γίνουν, πριν οδηγηθεί ο γιατρός στο συμπέρασμα, πως ο ασθενής πάσχει ήδη από κάποια καρδιαγγειακή νόσο. Προκειμένου να οδηγηθούμε στην δημιουργία μιας εφαρμογής, που θα βοηθούσε στην πρόβλεψη, των καρδιαγγειακών παθήσεων, θα πρέπει να γίνει μια ανάλυση των δεδομένων που θα χρησιμοποιηθούν για αυτήν [38].

Τα συστήματα υποστήριξης λήψης αποφάσεων, τα οποία επιλέξαμε να βασιστούμε, εντάσσονται στα συστήματα ιατρικής υποστήριξης (Medical Support Systems). Βοηθούν τους κλινικούς ιατρούς, στην ανάλυση και την ερμηνεία των στοιχείων των ασθενών, για την λήψη της τελικής τους απόφασης, η οποία μπορεί να αφορά μια φαρμακευτική αγωγή, μια διάγνωση κ.λπ. Ένας ακόμα λόγος, που προτιμούνται τέτοια συστήματα, είναι ότι συμβάλλουν στον ποιοτικό έλεγχο, των ιατρικών αποφάσεων. Η διοικητική επιστήμη, μαζί με την επιχειρησιακή έρευνα, εργάζονται εδώ και πολλά χρόνια πάνω σε θέματα μοντελοποίησης αποφάσεων και ανάλυσης αποφάσεων με στόχο την δημιουργία συστημάτων, ικανά να λάβουν την καλύτερη απόφαση. Δηλαδή τα συστήματα υποστήριξης αποφάσεων, βοηθούν τους χρήστες τους να πάρουν αποφάσεις, δείχνοντας τους “καλές” αποφάσεις [38]. Χωρίς όμως αυτό να σημαίνει, πως για την λήψη της τελικής απόφασης, δεν είναι απαραίτητη και η ανθρώπινη κρίση και διαίσθηση. Αναπόσπαστο κομμάτι, ενός συστήματος υποστήριξης αποφάσεων, είναι η ύπαρξης μια βάσης δεδομένων, η οποία παρέχει τα δεδομένα, που θα αξιοποιήσει το σύστημα.

Τα συστήματα υποστήριξης αποφάσεων επιλέχθηκαν για αυτή την εργασία γιατί δίνουν την δυνατότητα άμεσης αλληλεπίδρασης με τον χρήστη, χρησιμοποιούν δεδομένα και μοντέλα, είναι βασισμένα σε υπολογιστές και μπορούν να αντιμετωπίσουν ακόμα και κακώς δομημένα συστήματα. Τα DSS που χρησιμοποιήθηκαν και που χρησιμοποιούνται για ιατρικούς λόγους, δηλαδή τα ιατρικά DSS, από τεχνικής άποψης, δεν διαφέρουν από εκείνα που συναντάμε σε άλλα επιστημονικά πεδία. Παρέχουν την δυνατότητα διαχείρισης πληροφορίας, που προέρχεται από άλλα συστήματα και επικοινωνίας μαζί τους. Τα συστήματα αυτά, επιλέχθηκαν για την εργασία μας, προκειμένου να αξιοποιηθούν οι δυνατότητες τους, όπως είναι η αύξηση των επιπέδων πρόβλεψης, η μείωση του χρόνου αναμονής και η ενημέρωση του ασθενή [38]. Φυσικά κατά την ανάπτυξη και την λειτουργία των DSS, θα πρέπει πάντα να λαμβάνουμε υπόψη μας, τις ιδιαιτερότητες του χώρου της υγείας, με ακόμα μεγαλύτερη έμφαση στον παράγοντα της διαχείρισης των προσωπικών και ευαίσθητων ιατρικών δεδομένων και πληροφοριών.

ΣΤ'.2 Περιγραφή Του Συνόλου Δεδομένων

Προκειμένου να αξιολογηθεί η απόδοση διαφόρων αλγορίθμων ταξινόμησης εξόρυξης δεδομένων, χρησιμοποιήθηκε ένα σύνολο δεδομένων, ανθρώπων που έχουν εξαστεί για καρδιαγγειακές παθήσεις. Αυτό το σύνολο δεδομένων αποτελείται από 11 χαρακτηριστικά και μια

κλάση και ο αριθμός των περιπτώσεων που περιλαμβάνει ανέρχεται στο 70.000. Ακολουθούν τα δεδομένα και τα χαρακτηριστικά τους.

Attribute	Description
Age / ηλικία	Objective Feature age int (days) μέρες
Gender / φύλο	Objective Feature gender categorical code (1,2) χαρακτηριστικός κωδικός 1- γυναίκα , 2- άντρας
Height / Ύψος	Objective Feature height int (cm)
Weight / Βάρος	Objective Feature weight float (kg) Κιλά
Systolic blood pressure / Συστολική Αρτηριακή Πίεση	Examination Feature ap_hi int
Diastolic blood pressure / Διαστολική Αρτηριακή Πίεση	Examination Feature ap_lo int
Cholesterol / Χοληστερίνη	Examination Feature cholesterol 1: φυσιολογικό επίπεδο, 2: πάνω από το κανονικό, 3: πολύ πάνω από το κανονικό
Glucose / Γλυκόζη	Examination Feature gluc 1: φυσιολογικό, 2: πάνω από το φυσιολογικό, 3: πολύ πάνω από το φυσιολογικό
Smoking / Κάπνισμα	Subjective Feature smoke binary 0: μη καπνιστής, 1: καπνιστής
Alcohol intake / Λήψη Αλκοόλ	Subjective Feature alco binary 0: οχι, 1: ναι
Physical activity / Φυσική Άσκηση	Subjective Feature active binary 0: οχι, 1:ναι
Class Cardio /	Presence or absence of cardiovascular disease Target Variable cardio binary Παρουσία ή απουσία καρδιαγγειακής νόσου

Πίνακας 2: Πίνακας Δεδομένων

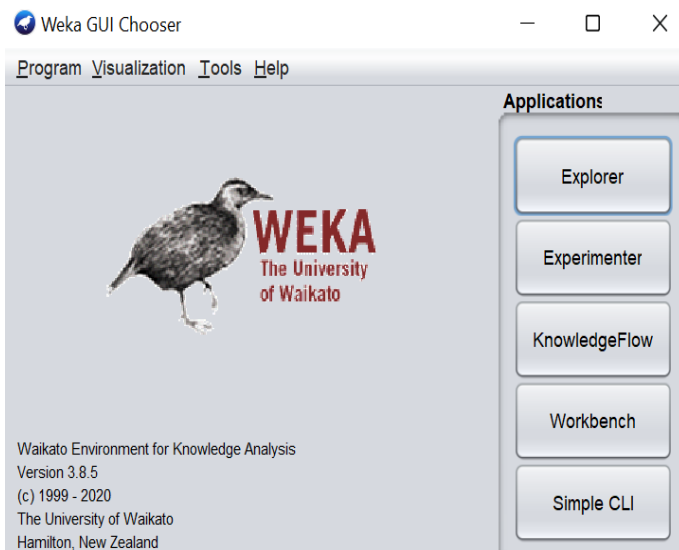
ΣΤ'.3 Data Analysis

Αρχικά για την δημιουργία της εφαρμογής, στηριχτήκαμε σε ένα σύνολο κλινικών δεδομένων. Το αρχικό μας αρχείο αποτελούνταν από 70.000 εγγραφές δεδομένων ασθενών, 11 κατηγορίες χαρακτηριστικών και μια κλάση, δηλαδή το πόρισμα της κάθε εγγραφής ασθενή. Οι κατηγορίες αφορούσαν την ηλικία του ατόμου, το ύψος και το βάρος, το φύλο, την συστολική αρτηριακή πίεση και την διαστολική αρτηριακή πίεση, το επίπεδο χοληστερίνης και γλυκόζης, την

ύπαρξη καπνιστικής συνήθειας ή μη, την φυσική δραστηριότητα και την τάση προς το αλκοόλ. Η τελευταία κατηγορία, δείχνει αν με βάση τα προηγούμενα, υπάρχει καρδιαγγειακή πάθηση. Για να καταφέρουμε να δημιουργήσουμε ένα σύστημα, το οποίο θα προέβλεπε την καρδιαγγειακή νόσο, έπρεπε με βάση τα δεδομένα που είχαμε να δημιουργήσουμε έναν κανόνα, για τον εντοπισμό του κινδύνου. Ο κανόνας αυτός θα υποδείκνυε τα κριτήρια με τα οποία το σύστημα, θα παίρνει αποφάσεις. Για να καταλήξουμε στην δημιουργία του κανόνα, κάναμε μια “εκπαίδευση” στα δεδομένα (training). Το εργαλείο που χρησιμοποιήσαμε για το training και για την διεξαγωγή του κανόνα ήταν το WEKA.

Το WEKA είναι ένα ελεύθερο λογισμικό, που το χρησιμοποιείται κυρίως για μηχανική μάθηση και εξόρυξη δεδομένων. Περιέχει μια αρκετά μεγάλη ποικιλία μεθόδων, για κατηγοριοποίηση, ανάλυση συστάδων, κανόνες συσχέτισης και παλινδρόμηση. Βοηθά επίσης, στην προεπεξεργασία των δεδομένων και διαθέτει εργαλεία οπτικοποίησης. Πρόκειται, όπως είπαμε, για ένα εργαλείο ανοικτού κώδικα, δηλαδή ο πηγαίος κώδικας είναι διαθέσιμος δημοσίως. Έτσι δίνει και την δυνατότητα σε χρήστες, με γνώσεις προγραμματισμού, να μπορούν να τροποποιήσουν και να εξελίσσουν τους αλγορίθμους του. Η γλώσσα που γράφτηκε το πρόγραμμα WEKA είναι η Java, προσφέροντας της με αυτό τον τρόπο, την δυνατότητα να εγκαθίσταται σε οποιοδήποτε λογισμικό. Το περιβάλλον εργασίας είναι γραφικό, επιτρέποντας την χρήση και σε χρήστες που δεν έχουν ιδιαίτερες γνώσεις προγραμματισμού.

Κατά την έναρξη του WEKA, εμφανίζεται ένα παράθυρο, από το οποίο ο χρήστης μπορεί να επιλέξει μια από τις πέντε κύριες εφαρμογές του WEKA, όπως βλέπουμε και στην εικόνα 6.3. Συγκεκριμένα, υπάρχει η επιλογή Explore, που είναι και η πιο γνωστή. Σε αυτή την διεπαφή ο χρήστης μπορεί να εκτελέσει τις κύριες εργασίες εξόρυξης δεδομένων όπως είναι η κατηγοριοποίηση, η ανάλυση συστάδων, η κατηγοριοποίηση, η ανακάλυψη κανόνων συσχέτισης, η οπτικοποίηση και η προεπεξεργασία των δεδομένων. Μια ακόμα επιλογή είναι ο Experimenter, ένα περιβάλλον για την διεξαγωγή πειραμάτων, τα οποία αξιολογούν μεθόδους παλινδρόμησης και κατηγοριοποίησης. Το εργαλείο αυτό διευκολύνει την σύγκριση, της επίδοσης διαφορετικών μοντέλων και διεξάγει τα αποτελέσματα σε μορφή πίνακα. Το περιβάλλον Knowledge, επιτρέπει την διεξαγωγή των ίδιων εργασιών, που επιτρέπει και το Explorer, με την διαφορά, ότι έχει διαφορετικό interface, δηλαδή διεπαφή. Εμείς, για την εργασία μας, επιλέξαμε το περιβάλλον Explorer, που όπως αναφέρθηκε, είναι και το πιο δημοφιλές.



Εικόνα 6.3: Εκκίνηση WEKA

ΣΤ' 3.1 Προεπεξεργασία

Η χρήση του WEKA Explorer, ξεκίνησε εισάγοντας τα δεδομένα. Το WEKA δίνει την δυνατότητα στους χρήστες του, να εισάγουν δεδομένα από βάσεις SQL, μέσω του Java Database Connectivity, ή από διευθύνσεις URL. Όμως ο πιο συνηθισμένος τρόπος, φόρτωσης / εισαγωγής δεδομένων είναι μέσω αρχείων ARFF. Όταν κάνουμε λόγο για αρχεία ARFF εννοούμε απλά αρχεία κειμένου, οι τιμές των οποίων διαχωρίζονται μεταξύ τους με κόμμα. Ένα αρχείο ARFF περιέχει μια επικεφαλίδα, κατά την οποία ορίζονται τα πεδία και το όνομα της σχέσης, δηλαδή ο πίνακας δεδομένων. Το αρχείο αυτό, γράφεται αρχικά σε ένα ηλεκτρονικό σημειωματάριο, ή σε ένα word και στην συνέχεια αποθηκεύεται με την κατάληξη .arff, για να είναι έτοιμο να ανοίξει, στην εφαρμογή WEKA. Στην εικόνα 6.3.1.1, παρουσιάζεται το αρχείο ARFF, που χρησιμοποιήσαμε. Στην αρχή αυτού του αρχείου, αναγράφεται η λέξη “@relation”, δηλαδή “σχέση”, ακολουθούμενη από το όνομα του πίνακα δεδομένων (qualification), που στην περίπτωση μας είναι το cardio. Ακολουθεί η δήλωση των πεδίων. Κάθε πεδίο δηλώνεται με συγκεκριμένο τρόπο. Για καθένα πεδίο ξεχωριστά, απαιτείται μια γραμμή, στην αρχή της οποίας, υπάρχει η λέξη “@attribute” και στην συνέχεια το όνομα του. Στην δική μας περίπτωση έχουμε δώδεκα πεδία. Το πεδίο age (ηλικία), gender (το φύλο), height (το ύψος), weight (βάρος), ap_hi (συστολική αρτηριακή πίεση), ap_lo (διαστολική αρτηριακή πίεση), cholesterol (χοληστερίνη), gluc (γλυκόζη), smoke (ύπαρξη καπνιστικής συνήθειας), alco (λήψη αλκοόλ), active (φυσική άσκηση / γυμναστική), cardio (παρουσία ή απουσία καρδιαγγειακής νόσου). Κατόπιν, στην ίδια γραμμή, δηλώνεται ο τύπος του πεδίου. Αν το πεδίο είναι αριθμητικό, όπως στην περίπτωση της ηλικίας, του ύψους, του βάρους, της συστολικής αρτηριακής πίεσης και της διαστολικής αρτηριακής πίεσης, χρησιμοποιείται η λέξη “numeric”. Στην περίπτωση που κάποια πεδία, παίρνουν συγκεκριμένες τιμές, ανοίγει αγκύλη και αναγράφονται. Τέτοια περίπτωση ήταν το

φύλο, που δηλώνεται με 1 ή 2, η ύπαρξη της καπνιστικής συνήθειας, η λήψη αλκοόλ, η άσκηση και η κλάση που δηλώνονται με 0 ή 1 και τέλος η συστολική αρτηριακή πίεση και διαστολική αρτηριακή πίεση που ανάλογα με το επίπεδο, αναγράφεται ο αριθμός 1,2 ή 3. Μετά την δήλωση των πεδίων, ακολουθούν τα δεδομένα. Πριν από τα καθαυτό δεδομένα προηγείται η λέξη “@data”.

```
|@relation cardio

@attribute age numeric
@attribute gender {1,2}
@attribute height numeric
@attribute weight numeric
@attribute ap_hi numeric
@attribute ap_lo numeric
@attribute cholesterol {1,2,3}
@attribute gluc {1,2,3}
@attribute smoke {0,1}
@attribute alco {0,1}
@attribute active {0,1}
@attribute class {0,1}

@data
18393 2      168    62.0  110    80    1    1    0    0    1    0
20228 1      156    85.0  140    90    3    1    0    0    1    1
18857 1      165    64.0  130    70    3    1    0    0    0    1
17623 2      169    82.0  150   100    1    1    0    0    1    1
17474 1      156    56.0  100    60    1    1    0    0    0    0
21914 1      151    67.0  120    80    2    2    0    0    0    0
22113 1      157    93.0  130    80    3    1    0    0    1    0
22584 2      178    95.0  130    90    3    3    0    0    1    1
17668 1      158    71.0  110    70    1    1    0    0    1    0
19834 1      164    68.0  110    60    1    1    0    0    0    0
22520 1      160    66.0  120    80    1    1    0    0    1    0
```

Εικόνα 6.2.1.1: Αρχείο Arff

ΣΤ’ 3.2 Επιλογή Δεδομένων

Ολά τα δεδομένα από την βάση που χρησιμοποιήσαμε, συλλέχθηκαν κατά την στιγμή των ιατρικών εξετάσεων. Είναι πρόσφατα δεδομένα, με την τελευταία τους ενημέρωση να έχει γίνει πριν από 3 χρόνια. Υπάρχουν τρεις τύποι χαρακτηριστικών εισόδου. Συγκεκριμένα υπάρχουν οι πραγματικές πληροφορίες ή αλλιώς στόχος, τα αποτελέσματα των ιατρικών εξετάσεων και οι πληροφορίες που δίνονται από τον ίδιο τον ασθενή. Οι κατηγορίες των δεδομένων, είναι διαμορφωμένες, σύμφωνα με τους παράγοντες, που επηρεάζουν τις καρδιαγγειακές παθήσεις.

Πιο αναλυτικά η ηλικία, σχετίζεται άμεσα, αλλά όχι αποκλειστικά, με την καρδιαγγειακή νόσο. Η καρδιά, είναι ένας μυς, η λειτουργία του οποίου είναι καθημερινή και ασταμάτητη. Με τα χρόνια, η λειτουργία του μυ αυτού μπορεί να φέρει αλιώσεις. Η γήρανση σχετίζεται με τη σκλήρυνση της αορτής, την ίνωση του μυοκαρδίου, την καρδιακή αμυλοείδωση και την υπερτροφία του μυοκαρδίου. Αλληλεπιδράσεις, της γήρανσης, με άλλους παράγοντες, όπως το φύλο, την παχυσαρκία

και συννοσηρότητες όπως, τον σακχαρώδη διαβήτη και την χοληστερίνη, καθορίζουν την αναδιαμόρφωση της αριστερής κοιλίας. Η έκταση της καρδιακής βλάβης σε συνδυασμό με προηγούμενες και επακόλουθες συννοσηρότητες, είναι οι καθοριστικοί παράγοντες της έκβασης, η οποία είναι καλύτερη στις γυναίκες παρά στους άνδρες [39]. Παρά τη χαμηλότερη συχνότητα, εμφάνισης της καρδιακής ανεπάρκειας, σε νεότερες ηλικίες, συγκριτικά με την εμφάνιση στους ηλικιωμένους, κλινικοί παράγοντες για καρδιακή ανεπάρκεια, ενέχουν ισχυρότερο κίνδυνο σε νέους, από ότι σε ηλικιωμένους [40]. Για αυτόν τον λόγο στην έρευνα μας, δεν εξετάζεται μόνο η ηλικία, αλλά συνδυάζονται και άλλοι παράγοντες μαζί της, πριν την διεξαγωγή του συμπεράσματος.

Ένας παράγοντας, που συνδυάζεται μαζί με την ηλικία είναι το φύλο του εξεταζόμενου. Έχει παρατηρηθεί ότι η επίπτωση της καρδιαγγειακής νόσου στις γυναίκες είναι στις περισσότερες περιπτώσεις χαμηλότερη από ότι στους άνδρες [41]. Παρόλα αυτά, οι γυναίκες έχουν υψηλότερη θνησιμότητα, μετά από οξεία καρδιαγγειακά επεισόδια και χειρότερη πρόγνωση. Αυτές οι διαφορές ανάμεσα στις γυναίκες και στους άντρες, παρατηρούνται ανάμεσα σε διάφορα καρδιαγγειακά νοσήματα, συμπεριλαμβανομένης της στεφανιαίας νόσου, του εγκεφαλικού επεισοδίου, της καρδιακής ανεπάρκειας και των παθήσεων της αορτής [41]. Οι διαφορές αυτές μεταξύ των φύλων έχουν προκαλέσει εκτεταμένες ανησυχίες και η συνεκτίμηση των διαφορών των φύλων έχει μεγάλη σημασία για την πρόληψη, τη διάγνωση, τη διαχείριση και τη θεραπεία της καρδιαγγειακής νόσου. Αποτελέσματα ερευνών έχουν δείξει, πως οι διαφορές προκαλούνται κυρίως από εγγενή γονίδια και εξωτερικές περιβαλλοντικές επιρροές. Η διάκριση των διαφορών μεταξύ των φύλων στις καρδιαγγειακές παθήσεις, θα βοηθήσει στην πιο αποτελεσματική πρόβλεψη των καρδιαγγειακών συμβάντων [41].

Ένα μεγάλο ποσοστό, μεγαλύτερο από 70%, όλων των θανάτων παγκοσμίως, αποδίδονται σε μη μεταδοτικές ασθένειες, με πρώτες στην λίστα τις καρδιακές παθήσεις. Ο εντοπισμός των ευάλωτων πληθυσμών, είναι αναμφισβήτητα ένα κρίσιμο βήμα για την ανάπτυξη προγνωστικών συστημάτων και στρατηγικών, για τη μείωση του ποσοστού θανάτου, από αυτές τις ασθένειες. Εξαιτίας της απλότητας τους, οι ανθρωπομετρικοί δείκτες είναι πολύ ελκυστικοί, για την εκτίμηση της ευπάθειας ενός ατόμου σε μια ασθένεια. Το ύψος έχει πρόσφατα συνδεθεί, με πολλούς κινδύνους, τόσο για καρδιαγγειακές παθήσεις όσο και για καρκίνο [42]. Η κατεύθυνση του όμως είναι αντίθετη. Οι μελέτες δείχνουν, πως οι πιο κοντοί άνθρωποι διατρέχουν μεγαλύτερο κίνδυνο για καρδιαγγειακή νόσο. Πέρα από την απλή επιδημιολογική συσχέτιση και εξήγηση, πρόσφατες γενετικές μελέτες, κατέδειξαν μια αιτία αυτής της σχέσης, εξηγώντας τον περίπλοκο αυτό βιολογικό μηχανισμό μέσω συνδέσεων με μια τρίτη μεταβλητή, δηλαδή την πνευμονική λειτουργία και το μεταβολικό προφίλ. Σε αυτό το σημείο θα πρέπει να πούμε ότι στο ύψος συμβάλλουν γενετικοί και περιβαλλοντικοί παράγοντες [42]. Καθοριστικός παράγοντας του ύψους, σε διαφορετικούς πληθυσμούς, είναι η οικονομική ανάπτυξη και η διατροφή των ατόμων και όχι η γενετική. Με βάση τα τρέχοντα στοιχεία, που αφορούν την σχέση μεταξύ του ύψους και του κινδύνου για καρδιαγγειακή νόσο, τα άτομα με

χαμηλό ύψος θα πρέπει να λαμβάνουν πιο αυστηρές στρατηγικές για την πρόληψη της καρδιαγγειακής νόσου [42].

Φυσικά μεγάλη επιρροή στα καρδιαγγειακά νοσήματα έχει και το βάρος. Ο συνδυασμός βάρους και ύψους, και η αναλογία αυτών των δύο, είναι ο δείκτης μάζας σώματος. Όταν ο δείκτης μάζας σώματος είναι αρκετά υψηλός είτε σε έναν άντρα είτε σε μια γυναίκα, κάνουμε λόγο για παχυσαρκία [43]. Η παχυσαρκία είναι ένας καλά τεκμηριωμένος παράγοντας, που συμβάλλει στην αιτιολογία, της καρδιαγγειακής νόσου. Βοηθά επίσης την δημιουργία άλλων παραγόντων καρδιαγγειακού κινδύνου, όπως ο διαβήτης τύπου δύο, η υπέρταση και οι διαταραχές ύπνου. Οδηγεί επίσης στην ανάπτυξη καρδιαγγειακών παθήσεων και θνησιμότητας από καρδιαγγειακά νοσήματα, ανεξάρτητα από άλλους παράγοντες καρδιαγγειακού κινδύνου [43]. Τα πιο πρόσφατα δεδομένα, υπογραμμίζουν την κοιλιακή παχυσαρκία, όπως προσδιορίζεται από την περίμετρο της μέσης, ως δείκτη κινδύνου καρδιαγγειακής νόσου, ανεξάρτητο από τον δείκτη μάζας σώματος. Πρόοδοι στις μεθόδους απεικόνισης, για τον χαρακτηρισμό της σύστασης του σώματος, συμπεριλαμβανομένης της σπλαχνικής παχυσαρκίας, έχουν δείξει πως η σπλαχνική παχυσαρκία, είναι ένας ακόμη ανεξάρτητος δείκτης κακών καρδιακών εκβάσεων [43]. Με βάση τα δεδομένα της ευρωπαϊκής καρδιολογικής εταιρείας, 1 στους 5 ανθρώπους αντιμετωπίζει προβλήματα ευσαρκίας, δηλαδή πρόκειται για ένα ποσοστό μεγαλύτερο από 20% [9].

Για πολλά χρόνια η διαστολική πίεση ήταν το επίκεντρο των μελετών και των προγνωστικών διαδικασιών. Με το πέρας του χρόνου όμως μέσα από συγκρίσεις της συστολικής και της διαστολικής αρτηριακής πίεσης, βρήκαν ότι η συστολική πίεση είναι καλύτερος δείκτης του καρδιαγγειακού κινδύνου. Σε πολλές περιπτώσεις, καρδιαγγειακών παθήσεων, η συστολική αρτηριακή πίεση είναι πιο κατατοπιστική από την διαστολική πίεση, κυρίως στις προγνώσεις που αφορούσαν ηλικίες των 45 ετών και άνω. Η υψηλή αρτηριακή πίεση είναι ένας σημαντικός παράγοντας κινδύνου καρδιαγγειακών νοσημάτων. Ένα μέσο ποσοστό των ανδρών, που έχουν αυξημένη αρτηριακή πίεση είναι 27%, ενώ αντίστοιχα για τις γυναίκες είναι 22%. [9] Σε αυτό το σημείο, θα πρέπει να σημειωθεί πως, επειδή η συστολική και η διαστολική αρτηριακή πίεση, εκτός από ορισμένων περιπτώσεων, συχνά πηγαίνουν προς την ίδια κατεύθυνση, δηλαδή είναι σύμφωνες, είναι σύνηθες να υποδεικνύουν τον κίνδυνο με παρόμοιο τρόπο, όπως και να επιλέγεται για μια πρόβλεψη ένας συνδυασμός των δύο [44]. Εκτός αυτού, ένας ακόμη λόγος κατά τον οποίο χρησιμοποιούνται και οι δύο για προβλέψεις, είναι γιατί η σύγκριση τους αλλά και καθεμία ξεχωριστά οδηγούν σε διαφορετικά συμπεράσματα. Η μεμονωμένη συστολική υπέρταση, σε αρκετές μελέτες, σχετίζεται με τον αυξημένο κίνδυνο θνησιμότητας. Δεν πρέπει να χειριζόμαστε μεμονωμένα τα συστατικά της αρτηριακής πίεσης. Η αύξηση της συστολικής πίεσης, είναι επικίνδυνη είτε αν αυξάνεται ταυτόχρονα και η διαστολική πίεση είτε αν μειώνεται. Αντίθετα η διαστολική υπέρταση από μόνη της, μπορεί να είναι λιγότερο επιβλαβής. Φυσικά χρειάζονται περισσότεροι παράμετροι, για την καλύτερη και πιο αποτελεσματική πρόβλεψη του κινδύνου [44].

Οι διατροφικές συνήθειες, τα διαιτητικά λίπη και οι μεγάλες ποσότητες πρόσληψης χοληστερόλης, έχουν σημαντικό ρόλο στην αλλαγή των κινδύνων καρδιαγγειακής νόσου. Έρευνες έχουν αποδείξει, ότι ορισμένα θρεπτικά συστατικά, σε μεγάλες ποσότητες, αυξάνουν τους κινδύνους καρδιαγγειακής νόσου [48]. Οι απόψεις γύρω από την διατροφική χοληστερόλη, εξακολουθούν να είναι αντικρουόμενες και ασαφείς. Οι ανταποκρίσεις στη διατροφική χοληστερόλη, θα πρέπει να παρακολουθούνται, καθώς πολλές φορές είναι μεταβαλλόμενες. Παρόλα αυτά, για την πρόβλεψη των καρδιαγγειακών παθήσεων, δεν αρκεί μόνο η παρακολούθηση της [48].

Επιρροή στις καρδιακές παθήσεις, έχει και ο διαβήτης. Ασθενείς που πάσχουν από διαβήτη, παρουσιάζουν αυξημένο κίνδυνο καρδιαγγειακής νόσου και θνησιμότητας, που μειώνει το προσδόκιμο ζωής του κατά 5 με 15 χρόνια, ανάλογα με την ηλικία [47]. Μια μελέτη στην Φινλανδία έδειξε, ότι ο διαβήτης τύπου 1 και τύπου 2 έχουν κίνδυνο καρδιακής θνησιμότητας 5,2 και 4,9 φορές αντίστοιχα. Μια ολοκληρωμένη φροντίδα του διαβήτη, ανεξάρτητα από τον τύπο του, συνεπάγεται με μια ολοκληρωμένη διαχείριση του καρδιαγγειακού κινδύνου, ο οποίος περιλαμβάνει και άλλους παράγοντες πέρα από την γλυκόζη [47].

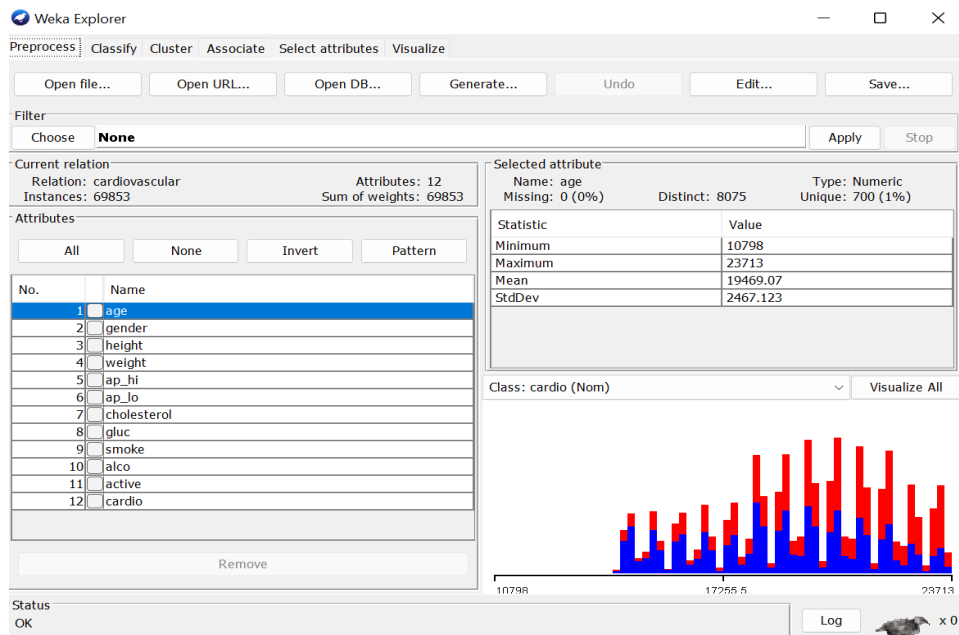
Ένας ακόμα παράγοντας, που εξετάστηκε, για την διεξαγωγή του μοντέλου πρόβλεψης, των καρδιαγγειακών παθήσεων, ήταν το κάπνισμα. Αποτελεί τον πιο διαδεδομένο παράγοντα, όταν γίνεται λόγος για παθήσεις της καρδιάς, καθώς την επηρεάζει άμεσα. Το κάπνισμα προκαλεί οξείες και χρόνιες ασθένειες και είναι μια από τις πιο σοβαρές απειλές για την υγεία του κοινού. Η συσχέτιση και η σύνδεση του καπνίσματος, με τον καρκίνο του πνεύμονα αναγνωρίστηκε νωρίς, αλλά η σχέση του καπνίσματος με την καρδιακή νόσο συζητήθηκε το 1960, χωρίς όμως να αναγνωρίζονται μεγάλες συσχετίσεις [45]. Αργότερα, το 1962, μια μελέτη διαπίστωσε πως το κάπνισμα προέβλεπε εμφράγματα του μυοκαρδίου, θνησιμότητα από στεφανιαία νόσο αλλά και από άλλες καρδιακές παθήσεις. Την ίδια χρονιά, οι ερευνητές του Framingham, σημείωσαν ότι, το κάπνισμα ήταν ένας παράγοντας καρδιακού κινδύνου, ανεξαρτήτως από άλλα χαρακτηριστικά, όπως η αρτηριακή πίεση ή η χοληστερόλη. Στην πορεία αποδείχθηκε όμως, ότι ενώ είναι ένας πολύ σημαντικός παράγοντας, δεν μπορεί να οδηγήσει σε συμπεράσματα από μόνος του [45]. Οι πρώτες σχετικές μελέτες, για τις επιπτώσεις του τσιγάρου στην καρδιαγγειακή υγεία, επικεντρώθηκαν στην νικοτίνη. Η νικοτίνη είναι μια βιοδραστική ένωση, που επηρεάζει τη συστολική και αρτηριακή πίεση, τον καρδιακό ρυθμό και τη καρδιακή παροχή. Πλέον είναι γνωστό, ότι το κάπνισμα αυξάνει τη φλεγμονή και τη θρόμβωση. Οδηγεί επίσης, σε οξειδωτικό στρες που εκδηλώνεται μέσω της αγγειοκινητικής δυσλειτουργίας, της ενεργοποίησης των αιμοπεταλίων και των λευκοκυττάρων, της υπεροξειδωσής των λιπιδίων και στον πολλαπλασιασμό των λείων μυών. Φυσικά σημαντικό ρόλο έχουν και οι 4000+ ενώσεις που βρίσκονται στον καπνό του τσιγάρου μαζί με το μονοξείδιο του άνθρακα [45]. Τα πιο πρόσφατα δεδομένα, από την Ευρωπαϊκή Καρδιολογική Εταιρεία, πως το κάπνισμα είναι μια βλαβερή συνήθεια, που συνεχίζει να επικρατεί, παρά τις εκστρατείες ενημέρωσης που έχουν γίνει. Σε όλες τις χώρες, μέλη της ESC, το 22,1% του

ενήλικου πληθυσμού, ηλικίας > 15 ετών, ήταν καθημερινά καπνιστές. Το ποσοστό των γυναικών είναι 2 στις 10 ενώ των ανδρών 3 στον 10. [9]

Όσον αφορά, τον παράγοντα του αλκοόλ, που μελετήσαμε για την διεξαγωγή του κανόνα, οι αποψεις διίστανται. Μελέτες, για την κατανάλωση του αλκοόλ σε ενήλικες, υποδεικνύουν ότι η κατανάλωση με μέτρο, δεν αυξάνει τον κίνδυνο καρδιακής ανεπάρκειας, υπέρτασης και άλλων καρδιαγγειακών ασθενειών. Στην πραγματικότητα μπορεί να βελτιώσει την αγγειακή λειτουργία και να μειώσει την πιθανότητα ισχαιμικού εγκεφαλικού [46]. Βέβαια, όπως και σε νεότερα άτομα, η μεγάλη κατανάλωση ή οι καταχρήσεις αλκοόλ, αναιρούν τις προστατευτικές καρδιακές επιδράσεις, ενώ παράλληλα αυξάνουν τα περιστατικά καρδιακής ανεπάρκειας και υπέρτασης. Για αυτόν το λόγο και δεδομένου, του κινδύνου που κρύβει το αλκοόλ, για άλλες ασθένειες, οι κλινικοί γιατροί δεν συνιστάται, να συμβουλεύουν τους ηλικιωμένους ασθενείς να ξεκινήσουν ή να αυξήσουν την κατανάλωση αλκοόλ, για καρδιο-προστατευτικούς λόγους [46]. Από την άλλη πλευρά, η μείωση της κατανάλωσης σε όσους πίνουν πολύ, είναι συνετή για όλες τις ηλικίες. Το ερώτημα το οποίο γεννάται, είναι το πόσο είναι το πολύ. Η απάντηση όσον αφορά την καρδιαγγειακή νόσο είναι πολύπλοκη. Ένας μεγάλος αριθμός από στοιχεία παρατηρήσεων, που υποστηρίζονται βιολογικά, υποδηλώνουν ότι λιγότερα από δύο ή δύο ποτήρια αλκοόλ την ημέρα, μπορεί να προστατεύουν έναν άνθρωπο από έμφραγμα του μυοκαρδίου και από την καρδιακή ανεπάρκεια. Αντίθετα η έντονη επεισοδιακή κατανάλωση αλκοόλ, δηλαδή περισσότερα από πέντε ποτά ανά ημέρα και η έντονη χρήση αλκοόλ, αυξάνουν τον κίνδυνο για καρδιακές παθήσεις [46].

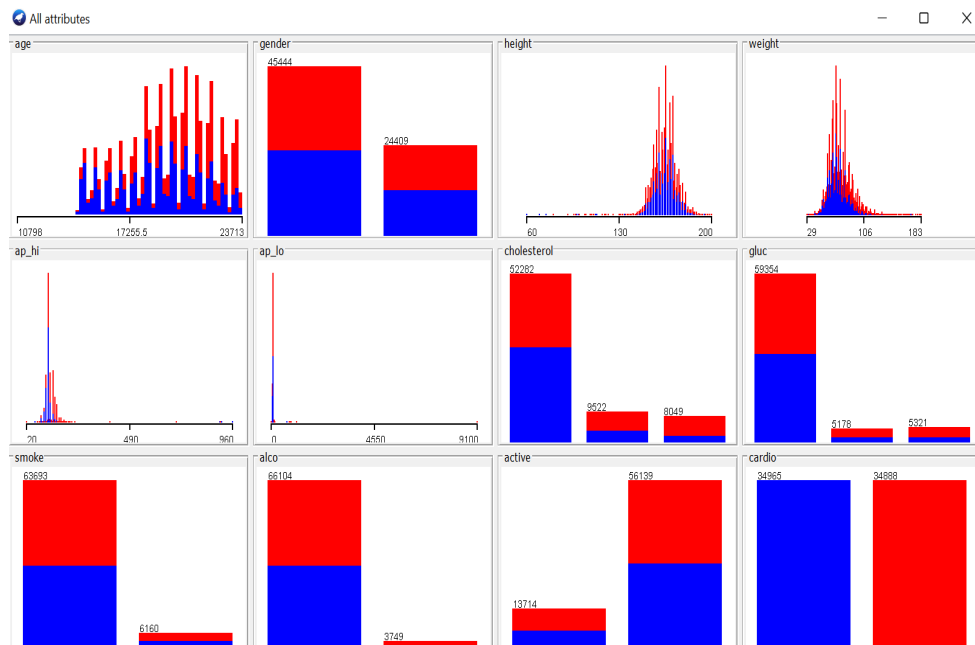
Τέλος ένας ακόμα παράγοντας που εξετάστηκε ήταν η φυσική άσκηση των ατόμων. Η φυσική άσκηση, δεν θα μπορούσε να έχει άλλες, παρά μόνο θετικές σχέσεις, με τα καρδιαγγειακά προβλήματα. Όσο πιο συχνά γυμνάζεται ένας άνθρωπος, τόσο πιο πολύ αποφεύγει τις καρδιακές παθήσεις. Αυτό οφείλεται κυρίως στην καλύτερη ροή του αίματος. Φυσικά και σε αυτή την περίπτωση θα πρέπει να αποφεύγονται οι υπερβολές, για να μην υπάρξουν αντίθετα αποτελέσματα.

Εκτός από την μεμονωμένη κατανόηση, κάθε κατηγορίας ξεχωριστά, βασικό ήταν, όπως έχει αναφερθεί ξανά, να βρεθεί ένα μοντέλο, στο οποίο θα συνδυάζεται κάθε κατηγορία και θα οδηγεί στο συμπέρασμα ύπαρξης των καρδιακών παθήσεων ή μη. Ο συνδυασμός αυτών των παραγόντων, θα μπορούσε να δώσει διαφορετικά αποτελέσματα, από εκείνα που θα μπορούσε να προβλέψει ένας άνθρωπος από μόνος του στο άκουσμα τους. Για την διεξαγωγή του κανόνα, χρησιμοποιήθηκε το πρόγραμμα WEKA. Όπως ήδη αναφέραμε, τα data, τα είχαμε ήδη μετατρέψει σε αρχείο arff και ήταν έτοιμα για επεξεργασία. Το άνοιγμα στην εφαρμογή έγινε με το explorer, που είναι και το πιο γνωστό. Με το αρχικό άνοιγμα εμφανίστηκε η οθόνη, που βλέπουμε και στην εικόνα 6.3.2.1. Κάτω αριστερά βλέπουμε όλες τις παραμέτρους, ενώ επιλέγοντας πάνω σε μια από αυτές τις κατηγορίες, στα δεξιά εμφανίζεται ένα διάγραμμα της, με βάσει τις τιμές που έχει.



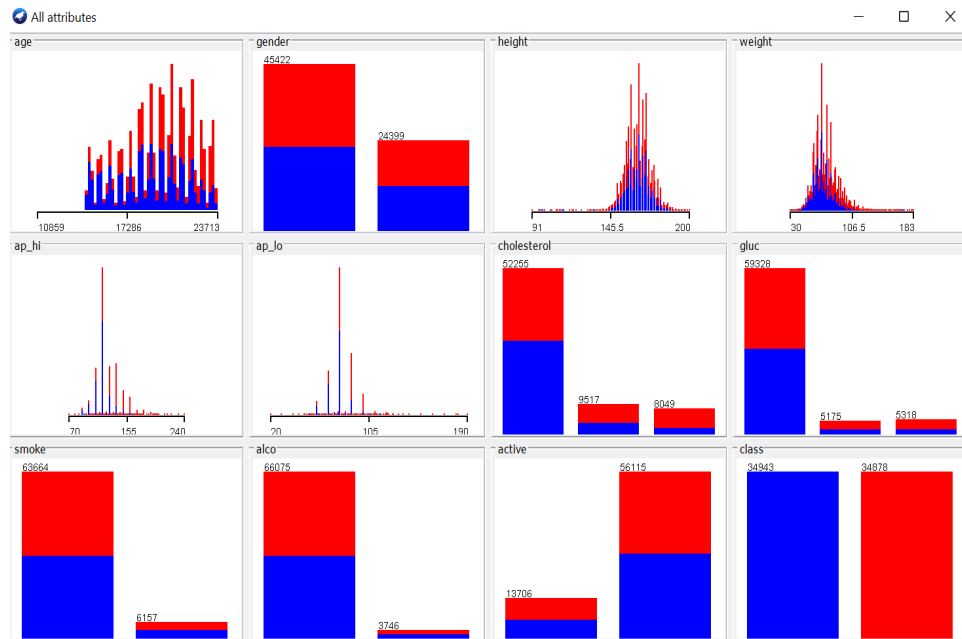
Εικόνα 6.2.2.1: Εισαγωγή στο WEKA

Πατώντας την επιλογή Visualize All, μπορούσαμε να έχουμε μαζική πρόσβαση, στα διαγράμματα από όλα τα δεδομένα. Το αποτέλεσμα αυτής της επιλογής, ήταν η εμφάνιση του παραθύρου της οθόνης, που φαίνεται στην Εικόνα 6.3.2.2. Σε περιπτώσεις, που τα δεδομένα μιας κατηγορίας ήταν nominal, όπως οι κατηγορίες gender, smoke, alco, active και η τελική μας κλάση, το διάγραμμα χωρίζονταν σε δύο στήλες. Αντίστοιχα όταν υπήρχαν τρεις κατηγορίες, οι στήλες που εξέφραζαν τις τιμές της κατηγορίας, γινόταν τρεις.



Εικόνα 6.3.2.2: Προβολή διαγραμμάτων

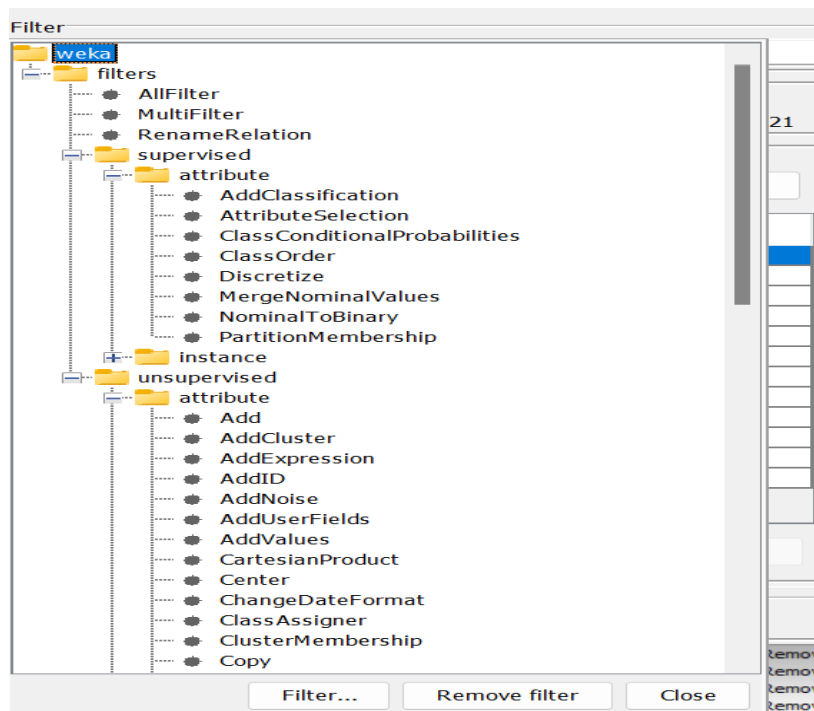
Σε αυτό το σημείο παρατηρήθηκε ότι το εύρος ορισμένων τιμών ήταν πολύ εκτεταμένο. Συγκεκριμένα στην κατηγορία height, weight, ap_hi και ap_lo, υπήρχαν τιμές οι οποίες ήταν σπάνιες περιπτώσεις. Παρόλα αυτά το πρόγραμμα μας, στην φάση της ταξινόμησης, θα προσπαθούσε να ταξινομήσει και αυτές τις περιπτώσεις, ταξινομώντας με λάθος τρόπο κάποιες άλλες. Για αυτό λοιπόν τον λόγο, προβήκαμε σε ορισμένες τροποποιήσεις, για την αποφυγή τέτοιων σφαλμάτων. Πιο αναλυτικά, ερευνήσαμε ποιες ήταν οι τιμές αυτές και τις διαγράψαμε. Τις αλλαγές στα διαγράμματα, μπορούμε να τις δούμε στην εικόνα 6.3.2.3.



Εικόνα 6.3.2.3: Διαγράμματα μετά την επεξεργασία

ΣΤ'.4 Εφαρμογή Αλγορίθμων

Επόμενο βήμα, αφού μικρήναμε το εύρος των δεδομένων, για την διεξαγωγή του κανόνα, ήταν η εφαρμογή διαφόρων φίλτρων και αλγορίθμων. Ένα παράδειγμα των φίλτρων που μπορεί να διαλέξει, ένας χρήστης του WEKA, βρίσκεται στην εικόνα 6.4.1. Τα φίλτρα αυτά, μπορεί κανείς να τα επιλέξει από την μπάρα filter, που βλέπουμε στην πρώτη οθόνη, που εμφανίζεται, όταν ανοίγουμε το αρχείο στο πρόγραμμα.



Εικόνα 6.4.1: Φίλτρα WEKA

Εμείς επιλέξαμε να δοκιμάσουμε το φίλτρο Attribute Selection, από την κατηγορία supervised και την υποκατηγορία attribute. Αυτό το φίλτρο επιλέγει από το dataset μας τις πιο σημαντικές κατηγορίες. Τα αποτελέσματα της εφαρμογής του φίλτρου, ήταν να έχουμε πολύ λίγες κατηγορίες, αφού από τις συνολικά δώδεκα, έμειναν μόνο οι μισές. Το φίλτρο διέγραφε τις κατηγορίες, gender, height, weight, gluc, smock και alco, αφήνοντας σαν πιο σημαντικά την ηλικία, την συστολική και διαστολική αρτηριακή πίεση, την χοληστερίνη, την σωματική άσκηση και την τελική κλάση. Οι πατηρήσεις μας ήταν, ότι το φίλτρο έκανε ελάχιστες τις κατηγορίες μας και δεν βοηθούσε στην διεξαγωγή, ενός κανόνα με όσο το δυνατόν λιγότερα σφάλματα και μεγαλύτερη ακρίβεια για αυτό και αφαιρέθηκε. Επόμενο βήμα, εφόσον αποφασίσαμε να μην επεξεργαστούμε με κάποιο φίλτρο τα δεδομένα, ήταν να περάσουμε από το preprocess, δηλαδή την προεπεξεργασία στην ταξινόμηση. Πατώντας στην επιλογή Classify, μπορούμε να επιλέξουμε οποιονδήποτε αλγόριθμο, για να τρέξει τα dataset. Εμείς επιλέξαμε να δοκιμάσουμε διάφορους αλγορίθμους, για να δούμε τις επιδόσεις τους. Στην συνέχεια έγινε μια σύγκριση μεταξύ αυτών, για να εντοπιστεί εκείνος με τα καλύτερα αποτελέσματα.

Ο πρώτος αλγόριθμος που τρέξαμε ήταν ο Naive Bayes. Πρόκειται για έναν πιθανολογικό ταξινομητή, που βασίζεται στην εφαρμογή του θεωρήματος του Bayes. Είναι από τα πιο απλά μοντέλα δικτύων Bayes, με δυνατότητα να επιτύχουν υψηλά επίπεδα ακρίβειας. Παρέχουν δυνατότητες επέκτασης, αφού πρόκειται για μια απλή τεχνική για την κατασκευή ταξινομητών.

Ο επόμενος αλγόριθμος βασισμένος στους κανόνες ήταν ο Jrip. Πιο αναλυτικά, ο αλγόριθμος αυτός, εφαρμόζει έναν εκπαιδευμένο κανόνα προτάσεων εν ονόματι “ Επαναλαμβανόμενο σταδιακό

κλάδεμα για την παραγωγή σφάλματος Reduction ”. Χρησιμοποιεί διαδοχικούς αλγορίθμους, για τη δημιουργία ταξινομημένων λιστών κανόνων. Ο αλγόριθμος περνάει από 4 στάδια. Το πρώτο στάδιο ονομάζεται ανάπτυξη κανόνα, το δεύτερο κλάδεμα, το τρίτο βελτιστοποίηση και το τέταρτο επιλογή.

Στην συνέχεια δοκιμάσαμε ακόμα έναν αλγόριθμο βασισμένο σε κανόνες, τον Decision Table. Ο αλγόριθμος αυτός δημιουργεί έναν απλό ταξινομητή, πλειοψηφίας, με πίνακα αποφάσεων. Στην πορεία, αξιολογεί τα υποσύνολα των χαρακτηριστικών, με βάση την καλύτερη αναζήτηση. Τέλος χρησιμοποιεί διασταυρωμένη επικύρωση για αξιολόγηση.

Ο τέταρτος αλγόριθμος που χρησιμοποιήσαμε ήταν ο J48. Ο αλγόριθμος αυτός, δημιουργεί δέντρα αποφάσεων, χρησιμοποιώντας τον αλγόριθμο C4.5 και είναι επέκταση του αλγορίθμου ID3. Πρόκειται για ακόμα έναν αλγόριθμο ταξινόμησης, τα αποτελέσματα του οποίου θα τα δούμε στην πορεία.

Τέλος εφαρμόστηκε ο αλγόριθμος Random Forest και Random Tree. Τα τυχαία δάση και τα τυχαία δέντρα απόφασης ανήκουν στην ίδια κατηγορία αλγορίθμων. Πρόκειται για μεθόδους εκμάθησης συνόλου ταξινόμησης και παλινδρόμησης. Τα τυχαία δάση τις περισσότερες φορές υπερτερούν από τα τυχαία δέντρα.

Αλγόριθμος	Description
Naive Bayes	Πιθανολογικός αλγόριθμος ταξινόμησης. Χρησιμοποιεί την στατιστική μέθοδο για ταξινόμηση.
Jrip	Αλγόριθμος “κλαδέματος” Χρήση τεσσάρων σταδίων: 1) ανάπτυξη κανόνα 2) κλάδεμα 3) βελτιστοποίηση 4) επιλογή
Decision Table	Αλγόριθμος βασισμένος σε πίνακες. Δημιουργεί πίνακες αποφάσεων και αξιολογεί τα υποσύνολα των χαρακτηριστικών, με βάση την καλύτερη αναζήτηση, διασταυρωμένη επικύρωση για αξιολόγηση.
J48	Δημιουργεί δέντρα αποφάσεων, χρησιμοποιώντας τον αλγόριθμο C4.5
Random Tree	Μέθοδος εκμάθησης συνόλου ταξινόμησης και παλινδρόμησης
Random Forest	Μέθοδος εκμάθησης συνόλου ταξινόμησης και παλινδρόμησης

Πίνακας 3: Αλγορίθμων

ΣΤ' 4.1 Σύγκριση Αλγορίθμων

Πίνακας Σύγκρισης Σωστών & Λανθασμένων Ταξινομήσεων

Τα αποτελέσματα αυτής της μελέτης, βασίζονται στα διάφορα τεστ που εφαρμόζονται και σε επιλογές που υπάρχουν. Γίνεται διασταυρωμένη επικύρωση και εφαρμόζονται ποσοστιαία κριτήρια. Για την πρόβλεψη, η επικύρωση γίνεται με το k-fold cross validation με k=10. Αυτό σημαίνει ότι η μέθοδος εκτελεί την δοκιμή 10 φορές. Τις πρώτες 9 γίνεται μια προπόνηση (training) και η τελευταία επανάληψη χρησιμοποιείται για δοκιμή. Χρησιμοποιήθηκε επίσης μια προσέγγιση διαίρεσης ποσοστών για σύγκριση των αποτελεσμάτων και για την απόδοση των αλγορίθμων που χρησιμοποιήθηκαν. Οι επιδόσεις και τα αποτελέσματα που διεξήχθησαν, παρουσιάζονται και συγκρίνονται στον πίνακα 4.

Αλγόριθμος	Σωστές Ταξινομήσεις	Λανθασμένες Ταξινομήσεις	Ποσοστό Λάθους	Ποσοστό Σωστών
Naive Bayes	50.306	19.515	27,95%	72,05%
JRip	50.880	18.941	27,1279%	72,8721%
Decision Table	51.080	18.741	26,8415%	73,1585%
J48	51.047	18.774	26,8888%	73,1112%
Random Tree	44.664	25.157	36,0307%	63,9693%
Random Forest	49.921	19.900	28,5015%	71,4985%

Πίνακας 4: Σωστών και λανθασμένων ταξινομήσεων.

Τα αποτελέσματα της σύγκρισης δείχνουν, ότι ο J48 μαζί με τον αλγόριθμο Decision Table, έχουν τις πιο υψηλές αποδόσεις και τα λιγότερα σφάλματα στις ταξινομήσεις τους. Ο J48 καταφέρνει να ταξινομήσει σωστά ένα ποσοστό 73,1112%, ενώ ο Decision Table ταξινομεί σωστά το 73,8721%. Βέβαια αυτό δεν αρκεί για να βγει, ένα συμπέρασμα, σχετικά με την καλύτερο αλγόριθμο. Μετά από τον J48 και τον Decision Table, ακολουθεί ο JRip με ποσοστό ακρίβειας 72,8721%. Την χειρότερη απόδοση παρατηρούμε να έχει ο αλγόριθμος Random Tree, με απόδοση σωστών ταξινομήσεων 63,9693%.

Χρόνος Εκτέλεσης Ταξινόμησης και root mean square error

Αλγόριθμος	Time to Build the Model (Seconds)	Root Mean Squared Error
------------	-----------------------------------	-------------------------

Naive Bayes	0.2 sec	0.4448
JRip	9.27 sec	0.4425
Decision Table	8 sec	0.4276
J48	4.78 sec	0.4348
Random Tree	0.86 sec	0.6001
Random Forest	36.31 sec	0.442

Πίνακας 5: Χρόνος Εκτέλεσης Ταξινόμησης και root mean square error

Τον μεγαλύτερο χρόνο, για να χτύσει το μοντέλο, μεταξύ αυτών των αλγορίθμων, τον έκανε ο Random Forest. Συγκεκριμένα έκανε 36.27 δευτερόλεπτα να χτύσει το μοντέλο. Ο δεύτερος μεγαλύτερος χρόνος ήταν του JRip με 9.27 δευτερόλεπτα. Ακολούθησε ο αλγόριθμος Decision Table με 8 δευτερόλεπτα, ο J48 με 4.78 δευτερόλεπτα. Τους καλύτερους χρόνους, δόμησης του μοντέλου, τους έκαναν οι Naive Bayes και Random Tree με 0.2 και 0.86 δευτερόλεπτα αντίστοιχα.

Υπάρχουν πολλοί τρόποι κατά τους οποίους υπολογίζονται οι διαφορετικές επιδόσεις των αλγορίθμων. Το σχήμα 1 που ακολουθεί, απεικονίζει επίσης ένα μοντέλο πίνακα σύγχυσης, γνωστό και ως πίνακα έκτακτης ανάγκης. Σε αυτόν τον πίνακα, κάθε σειρά εμφανίζει την πραγματική κλάση, ενώ κάθε στήλη δείχνει την προβλεπόμενη κλάση.

TP	FN
FP	TN

Σχήμα 1: Confusion Matrix

Τα TP (True Positive) και TN (True Negative) είναι περιπτώσεις που έχουν ταξινομηθεί σωστά ενώ τα FP (False Positive) και FN (False Negative) είναι οι περιπτώσεις που έχουν ταξινομηθεί λάθος. Ένα ακόμα μέτρο αξιολόγησης είναι η ακρίβεια της (%) των σωστών στοιχείων που έχουν επιλεγεί και υπολογίζεται ως εξής: $\text{Ανάκλαση}(R) = TP / (TP+FN)$. Με την βοήθεια της ακρίβειας και της ανάκλασης υπολογίζεται το F-Measure(F), δηλαδή ο αρμονικός μέσος όρος ακρίβειας και ανάκλασης, που υπολογίζεται ως $F=2*R*P/(R+P)$

Μέτρα απόδοσης υπολογισμένα βάση confusion matrix με 10-fold cross validation

Αλγόριθμος	TP Rate	FP Rate	Precision (P)	Recall (R)	F Measure(F)
Naive Bayes	0.720	0.280	0.729	0.720	0.718
JRip	0.729	0.271	0.729	0.729	0.729
Decision Table	0.732	0.268	0.733	0.732	0.731
J48	0.731	0.269	0.732	0.731	0.731
Random Tree	0.640	0.360	0.640	0.640	0.640
Random Forest	0.715	0.285	0.715	0.715	0.715

Πίνακας 6: Μέτρα απόδοσης υπολογισμένα βάση confusion matrix με 10-fold cross validation

Ο παραπάνω πίνακας δείχνει τον ρυθμό TP και FP κάθε ταξινομητή. Ο μέσος όρος Precision, Recall και F-Measure υπολογίζονται με τον confusion matrix με 10-fold cross validation. Το Decision Table έχει το υψηλότερο ποσοστό TP. Το TP rate του είναι 0.732, ενώ το αμέσως επόμενο του J48 είναι 0.731. Αμέσως μετά ακολουθεί ο αλγόριθμος JRip με 0.729 και ο Naive Bayes με 0.720. Το χαμηλότερο TP rate το έχουν οι αλγόριθμοι των δέντρων και δασών.

Ένα ακόμα κριτήριο για την επιλογή του κατάλληλου αλγορίθμου ήταν η κατανομή των σωστών και των λάθος ταξινομήσεων. Ήταν σημαντικό να υπάρχει ισορροπία στα δεδομένα. Δεν θα έπρεπε να επιλεγεί ένας αλγόριθμος, ο οποίος να έχει ελάχιστα λάθη στις ταξινομήσεις του (a), με $a=0$ και το αντικείμενο (b), με $b=1$ να ταξινομείται λάθος. Την ισορροπία κάθε αλγορίθμου θα την δούμε στον πίνακα 7 που ακολουθεί.

Balance ταξινόμησης

Αλγόριθμος	a b
Naive Bayes	28498 6445 13070 21808
JRip	26181 8762 10179 24699
Decision Table	26851 8092 10649 24229
J48	26776 8167

	10607 24271
Random Tree	22421 12522 12635 22243
Random Forest	25583 9360 10540 24338

Πίνακας 7: Balance ταξινόμησης

Το καλύτερο balance, το έχει ο αλγόριθμος Random Tree, έχει την μικρότερη διαφορά ανάμεσα στο a και b. Ακολουθεί ο αλγόριθμος Random Forest και ο Jrip. Το χειρότερο balance το έχει ο αλγόριθμος Naive Bayes, ενώ το λιγότερο κακό balance, ο αλγόριθμος J48.

Οι συγκρίσεις από όλα τα στοιχεία, των δοκιμών που κάναμε, έδειξαν πως καλύτερη απόδοση από όλους τους αλγορίθμους έχει ο αλγόριθμος J48. Ο αλγόριθμος ταξινόμησης J48 ήταν πολύ κοντά στις σωστές ταξινομήσεις με τον αλγόριθμο Decision Table, αλλά υπερετρούσε σε μεγάλο βαθμό στον χρόνο εκτέλεσης. Επίσης ο J48 είχε πολύ μικρή διαφορά με τον Decision Table σε ποσοστό TP. Στο μόνο που υστερούσε ήταν το balance, χωρίς όμως να είναι μη αποδεκτό. Συνεπώς εστιάσαμε περισσότερο στον αλγόριθμο J48 και εφαρμόσαμε ορισμένες αλλαγές, για να δούμε πως θα γίνει καλύτερο το μοντέλο.

Το τελευταίο βήμα πριν, την δημιουργία του αλγορίθμου J48, ήταν οι αλλαγές στον αλγόριθμο, προκειμένου να βελτιώσουμε ακόμα περισσότερο την απόδοση του μοντέλου, που επιλέξαμε να βασιστούμε. Αρχικά αλλάξαμε το min number of object και δοκιμάσαμε διάφορες τιμές, για να δούμε τις αλλαγές, που είχε ο αλγόριθμος. Τα βασικά στοιχεία που άλλαζαν σε κάθε τιμή και μας ενδιέφεραν, φαίνονται στους ακόλουθους πίνακες. Ο ελάχιστος αριθμός παρουσιών ανά φύλλο, δηλαδή το min number of objects, έχει την τάση να αυξάνει την ακρίβεια στο αλγόριθμο J48.

Αξιολόγηση ταξινόμησης με διαφορετικά Number of objects για τον j48

Min Number Of Objects	Σωστές Ταξινομήσεις	Λανθασμένες Ταξινομήσεις	Ποσοστό Λάθους	Ποσοστό Σωστών
2 / αρχικό	51047	18774	26.8888 %	73.1112 %
50	51167	18654	26.7169 %	73.2831 %
100	51156	18665	26.7326 %	73.2674 %
150	51195	18626	26.6768 %	73.3232 %
200	51219	18602	26.6424 %	73.3576 %

250	51156	18665	26.7326 %	73.2674 %
300	51185	18636	26.6911 %	73.3089 %
350	51218	18603	26.6438 %	73.3562 %
400	51147	18674	26.7455 %	73.2545 %

Πίνακας 8: Number of objects j48 (Αξιολόγηση ταξινομήσεων)

Ο ελάχιστος αριθμός παρουσιών ανά φύλλο, που έδωσε το μεγαλύτερο ποσοστό σωστών ταξινομήσεων ήταν το 200, με ποσοστό 73.3576 %. Ο επόμενος αριθμός ήταν το 350, με ποσοστό σωστών ταξινομήσεων 73.3562 %. Το μικρότερο ποσοστό, το συναντήσαμε με ελάχιστο αριθμό παρουσιών ανά φύλλο 2, που ήταν η αρχική μας δοκιμή. Οι διαφορές ανάμεσα στα ποσοστά ταξινομήσεων, δεν θα αρκούσαν για να οδηγηθούμε σε νέο συμπέρασμα, αν δεν πληρούνταν και τα υπόλοιπα κριτήρια αξιολόγησης ενός αλγορίθμου. Συνέχεια συγκρίθηκε ο χρόνος εκτέλεσης ταξινόμησης και root mean square error.

Χρόνος Εκτέλεσης Ταξινόμησης & Root Mean Square Error J48

Min Number Of Objects	Time to Build the Model (Seconds)	Root Mean Squared Error
2 / αρχικό	4.78 sec	0.4348
50	1.72 sec	0.4294
100	1.68 sec	0.4294
150	1.37 sec	0.4292
200	1.44 sec	0.4291
250	1.37 sec	0.4293
300	1.38 sec	0.4292
350	1.29 sec	0.4291
400	1.35 sec	0.4296

Πίνακας 9: Χρόνος Εκτέλεσης Ταξινόμησης & Root Mean Square Error J48

Ο καλύτερος χρόνος εκτέλεσης παρατηρήθηκε για min number of objects 350, με χρόνο εκτέλεσης 1.29 sec. Ο δεύτερος καλύτερος χρόνος εκτέλεσης, ήταν για ελάχιστο αριθμός παρουσιών

ανά φύλλο 400, με χρόνο εκτέλεσης 1.35 sec. Ο χειρότερος χρόνος ήταν για min number of objects 2 με χρόνο 4.78 sec και ο επόμενος χειρότερος ήταν 1.72 sec για ελάχιστο αριθμό 50.

Μέτρα απόδοσης υπολογισμένα βάση confusion matrix με 10-fold cross validation για διαφορετικά min number of objects στον J48

Min Number Of Objects	TP Rate	FP Rate	Precision (P)	Recall (R)	F Measure(F)
2 / αρχικό	0.731	0.269	0.732	0.731	0.731
50	0.733	0.267	0.734	0.733	0.732
100	0.733	0.267	0.734	0.733	0.732
150	0.733	0.267	0.734	0.733	0.733
200	0.734	0.266	0.735	0.734	0.733
250	0.733	0.267	0.735	0.733	0.732
300	0.733	0.267	0.735	0.733	0.732
350	0.734	0.267	0.736	0.734	0.733
400	0.733	0.268	0.734	0.733	0.732

Πίνακας 10: Μέτρα απόδοσης υπολογισμένα βάση confusion matrix με 10-fold cross validation για διαφορετικά min number of objects στον J48

Ο παραπάνω πίνακας δείχνει τον ρυθμό TP και FP, για διαφορετικά min number of objects στον αλγόριθμο J48. Όπως έχει αναφερθεί, ο μέσος όρος Precision Recall και F-Measure υπολογίζονται με τον confusion matrix με 10-fold cross validation. Για min number of object 350 και 200 έχουμε το υψηλότερο TP rate, το οποίο είναι 0.734 και για τα δύο. Το αμέσως επόμενο υψηλό TP rate, είναι το 0.733, που ισχύει για min number of objects 50, 100, 150, 250, 300, 400. Το χαμηλότερο TP rate, το είχε το αρχικό, δηλαδή το 2 με 0.731.

Τελευταία παρατήρηση που έπρεπε να κάνουμε, για να επιλέξουμε το min number of object, που θα έδινε στον αλγόριθμο καλύτερη απόδοση, ήταν να δούμε, ποια ταξινόμηση είχε καλύτερο balance, συγκριτικά με την απόδοση που έδινε. Για να συγκρίνουμε την ισορροπία, κάθε δοκιμής, δημιουργήσαμε ακόμα έναν πίνακα, με τα στοιχεία που φανερώνουν, το αν υπάρχει balance.

Balance ταξινόμησης για διαφορετικά min number of objects

Min Number Of Objects	a b
2 / αρχικό	26776 8167 10607 24271
50	26869 8074 10580 24298
100	26807 8136 10529 24349
150	26790 8153 10473 24405
200	26902 8041 10561 24317
250	27173 7770 10895 23983
300	27305 7638 10998 23880
350	27519 7424 11179 23699
400	27116 7827 10847 24031

Πίνακας 11: Balance ταξινόμησης για διαφορετικά min number of objects

Το καλύτερο balance, σύμφωνα με τα αποτελέσματα το έχει το min number of object 50. Οι διαφορές ανάμεσα στην ισορροπία των αλγορίθμων, μέχρι το min number of object 200 είναι μικρές. Μετά όμως παρατηρείται, μια αρνητική αλλαγή. Οι επόμενοι αριθμοί, δεν εμφανίζουν καλή ισορροπία ταξινόμησης, με χειρότερο το min number of object 350.

Κατά την διάρκεια της επεξεργασίας παρατηρήθηκε ακόμα μια αλλαγή. Όσο το min number of objects αυξάνεται, τόσο μειώνεται ο αριθμός των φύλων, δηλαδή μικραίνει ο όγκος του δέντρου. Τα δύο πόσα min number of object και αριθμός των φύλων, φαίνεται να έχουν αντιστρόφως αναλογική σχέση μεταξύ τους. Στο min number of objects 2, ο αριθμός των φύλων ήταν 522, το οποίο αύξανε την πολυπλοκότητα του δέντρου. Μετά όταν ο ελάχιστος αριθμός έγινε 50, ο αριθμός των φύλων μειώθηκε σημαντικά στο 57. Στην πορεία που το min number of objects, αυξήθηκε σε 100, ο αριθμός των φύλων έγινε 46. Το ίδιο ίσχυε και για min number of objects 150, 200, 300, 350

και 400 που ο αριθμός των φύλλων άλλαξε σε 41, 32, 30, 26, 24 και 25 αντίστοιχα. Φυσικά όσο πιο μικρός ήταν ο αριθμός των φύλλων, τόσο πιο μικρό ήταν και το μέγεθος του δέντρου.

Έπειτα από την ανάλυση και την παρατήρηση των αλλαγών που φέρνει το min number of objects προβήκαμε στο ακόλουθο συμπέρασμα. Το min number of objects, που φέρνει την καλύτερη απόδοση στον αλγόριθμο J48, ήταν ο αριθμός 200. Είχε το μεγαλύτερο ποσοστό σωστών ταξινομήσεων, (73,3576%) με 512119 σωστές ταξινομήσεις και 18602 λάθος ταξινομήσεις ποσοστό 26.6424%. Συγκριτικά με τα αποτελέσματα άλλων αριθμών, έχτισε το μοντέλο του σε πολύ καλό χρόνο (1.44 sec), ο οποίος δεν διέφερε, παρά μόνο κάτι δευτερόλεπτα από τον καλύτερο χρόνο που συναντήσαμε, που ήταν 1.29 sec. Στην σύγκριση με τις τιμές, του TP rate, είχε το πιο υψηλό ποσό, ενώ παράλληλα είχε καλό balance στις ταξινομήσεις του. Ο αριθμός των φύλλων του δέντρου του ήταν 32 και το μέγεθος του δέντρου ήταν 59. Συνεπώς, πρόκειται για ένα αρκετά διαχειρίσιμο δέντρο, το οποίο διαθέτει έναν αρκετά υψηλό ποσοστό ακρίβειας στον κανόνα του.

Κεφάλαιο Ζ'

Δημιουργία Εφαρμογής

Έπειτα από την δοκιμή διαφόρων αλγορίθμων, ακολούθησε μια ανάλυση και σύγκριση μεταξύ τους. Ο αλγόριθμος που έφερε συγκριτικά τα καλύτερα αποτελέσματα ήταν, ο J48. Για την βελτίωση του αλγορίθμου J48, προβήκαμε σε ορισμένες δοκιμές όπως στην αλλαγή του min number of objects. Ακολούθησε σύγκριση, των νέων δοκιμών και καταλήξαμε στο συμπέρασμα, πως ο min number of objects, που είναι κατάλληλος για να βελτιώσει την απόδοση, του μοντέλου, είναι το 200.

Ο κανόνας που διεξήχθη ήταν ο ακόλουθος:

```
ap_hi <= 129
| age <= 19939
| | cholesterol = 1: 0 (21658.0/4573.0)
| | cholesterol = 2: 0 (2605.0/838.0)
| | cholesterol = 3
| | | gluc = 1: 1 (424.0/110.0)
| | | gluc = 2: 1 (37.0/8.0)
| | | gluc = 3
| | | | weight <= 73.5: 0 (409.0/166.0)
| | | | weight > 73.5: 1 (209.0/96.0)
| age > 19939
| | cholesterol = 1
| | | age <= 22217
| | | | ap_hi <= 118: 0 (2195.0/632.0)
| | | | ap_hi > 118
| | | | | active = 0
```

```

| | | | | weight <= 83.5
| | | | | | age <= 21172: 0 (615.0/245.0)
| | | | | | age > 21172: 1 (550.0/261.0)
| | | | | | weight > 83.5: 1 (227.0/98.0)
| | | | | active = 1: 0 (5695.0/2166.0)
| | | age > 22217
| | | | active = 0: 1 (710.0/274.0)
| | | | active = 1
| | | | | ap_hi <= 119: 0 (576.0/255.0)
| | | | | ap_hi > 119
| | | | | | gender = 1: 1 (1136.0/516.0)
| | | | | | gender = 2
| | | | | | weight <= 76.5: 0 (560.0/241.0)
| | | | | | weight > 76.5: 1 (275.0/120.0)
| | | cholesterol = 2
| | | | ap_hi <= 115: 0 (578.0/205.0)
| | | | ap_hi > 115
| | | | | active = 0: 1 (236.0/98.0)
| | | | | active = 1
| | | | | age <= 22057: 0 (599.0/247.0)
| | | | | age > 22057: 1 (260.0/114.0)
| | | cholesterol = 3: 1 (1643.0/490.0)
ap_hi > 129
| | | | ap_hi <= 138
| | | | cholesterol = 1
| | | | | smoke = 0
| | | | | age <= 21739
| | | | | | ap_lo <= 89
| | | | | | height <= 172
| | | | | | | ap_lo <= 79: 0 (308.0/129.0)
| | | | | | | ap_lo > 79
| | | | | | | | active = 0: 1 (365.0/166.0)
| | | | | | | | active = 1
| | | | | | | | | gender = 1: 1 (1080.0/530.0)
| | | | | | | | | gender = 2: 0 (325.0/136.0)
| | | | | | | height > 172: 0 (340.0/132.0)

```

```
| | | | ap_lo > 89: 1 (1645.0/661.0)
| | | | age > 21739: 1 (1491.0/540.0)
| | | smoke = 1: 0 (616.0/260.0)
| | cholesterol = 2: 1 (1346.0/528.0)
| | cholesterol = 3: 1 (1739.0/353.0)
| ap_hi > 138: 1 (19369.0/3126.0)
```

Με βάση αυτόν τον κανόνα, που έβγαλε το πρόγραμμα WEKA, γράψαμε το μοντέλο σε php και στην συνέχεια κάναμε χρήση της html για να εμφανίσουμε μια τοπική εφαρμογή. Η ιδιότητα αυτής της εφαρμογής θα ήταν να προβλέπει τις καρδιακές παθήσεις σε έναν άνθρωπο. Τα στοιχεία που απαιτούνται για την πρόβλεψη είναι το φύλο, η ηλικία, το ύψος, το βάρος, η συστολική αρτηριακή πίεση, η διαστολική αρτηριακή πίεση, το επίπεδο χοληστερίνης, το επίπεδο γλυκόζης, η ύπαρξη καπνιστικής συνήθειας και η σωματική άσκηση. Η λήψη αλκοόλ, δεν υπολογίζεται και δεν επηρεάζει με κάποιον τρόπο τον κανόνα, καθώς όπως παρατηρούμε, δεν υπάρχει σε κάποια συνθήκη του. Η δημιουργία της πλατφόρμας, για την πρόβλεψη των καρδιακών παθήσεων, έγινε με το ακόλουθο τρόπο. Αρχικά γράψαμε σε html και php ένα κομμάτι κώδικα, το οποίο περιέγραφε το πως θα είναι η πλατφόρμα, τον τίτλο τα στοιχεία, που θα πρέπει να εισάγει ο χρήστης και τα πλήκτρα για την λειτουργία του προγράμματος. Σε αυτό το σημείο, θα έπρεπε να καθορίσουμε και το ποιές τιμές είναι αποδεκτές, για το κάθε στοιχείο, που εισάγει ο χρήστης. Ακολουθεί ο κώδικας ο οποίο προέκυψε από αυτή την διαδικασία.

```
<html>
<head>
  <title>Test Data | Cardiovascular Disease</title>
  <meta name="viewport" content="width=device-width, initial-scale=1">
  <script src="https://code.jquery.com/jquery-3.5.1.slim.min.js"
integrity="sha384-DfXdz2htPH0lsSSs5nCTpuj/zy4C+OGpamoFVy38MVBnE+IbbVYUew+OrCXaR
kfj" crossorigin="anonymous"></script>
  <script src="https://cdn.jsdelivr.net/npm/bootstrap@4.6.0/dist/js/bootstrap.bundle.min.js"
integrity="sha384-Piv4xVNRyMGpQkS2by6br4gNJ7DXjqk09RmUpJ8jgGtD7zP9yug3goQfGII0yAn
s" crossorigin="anonymous"></script>
  <link rel="stylesheet"
href="https://cdn.jsdelivr.net/npm/bootstrap@4.6.0/dist/css/bootstrap.min.css"
integrity="sha384-B0vP5xmATw1+K9KRQjQERJvTumQW0nPEzvf6L/Z6nronJ3oUOFUFpCjEUQo
uq2+I" crossorigin="anonymous">
<style>
@media (min-width: 800px) {
  .container{
    max-width: 800px;
  }
}
```

```

</style>
</head>
<body>
  <nav class="navbar navbar-expand-lg bg-dark navbar-dark">
    <a class="navbar-brand" href="index.php">Cardiovascular Disease</a>
  </nav>
  <br />
  <div class="container">
    <form class="border p-5" method="post" action="results.php">
      <p class="h4 mb-4 text-center">Test Data</p>

      <label for="textInput">age</label>
      <input class="form-control" type="text" name="age" placeholder="Number input" required>
      <small class="form-text text-muted mb-4">Input must be positive integer number.</small>

      <label for="textInput">gender</label>
      <input class="form-control" type="text" name="gender" placeholder="Number input"
required>
      <small class="form-text text-muted mb-4">Input must be positive integer number.</small>

      <label for="textInput">height</label>
      <input class="form-control" type="text" name="height" placeholder="Number input"
required>
      <small class="form-text text-muted mb-4">Input must be positive integer number.</small>

      <label for="textInput">weight</label>
      <input class="form-control" type="text" name="weight" placeholder="Number input"
required>
      <small class="form-text text-muted mb-4">Input must be positive integer number.</small>

      <label for="textInput">active</label>
      <input class="form-control" type="text" name="active" placeholder="Number input"
required>
      <small class="form-text text-muted mb-4">Input must be positive integer number.</small>

      <label for="textInput">smoke</label>
      <input class="form-control" type="text" name="smoke" placeholder="Number input"
required>
      <small class="form-text text-muted mb-4">Input must be positive integer number.</small>

      <label for="textInput">cholesterol</label>
      <input class="form-control" type="text" name="cholesterol" placeholder="Number input"
required>
      <small class="form-text text-muted mb-4">Input must be positive integer number.</small>

      <label for="textInput">gluc</label>
      <input class="form-control" type="text" name="gluc" placeholder="Number input" required>

```

```

<small class="form-text text-muted mb-4">Input must be positive integer number.</small>

<label for="textInput">ap_hi</label>
<input class="form-control" type="text" name="ap_hi" placeholder="Number input"
required>
<small class="form-text text-muted mb-4">Input must be positive integer number.</small>

<label for="textInput">ap_lo</label>
<input class="form-control" type="text" name="ap_lo" placeholder="Number input"
required>
<small class="form-text text-muted mb-4">Input must be positive integer number.</small>

<button class="btn btn-info btn-block" type="submit" name="submit">Submit</button>
</form>
</div>
<br />
<br />
</body>
</html>

```

Στην συνέχεια, δημιουργήσαμε ένα δεύτερο κομμάτι κώδικα, σε ξεχωριστό αρχείο. Οι γλώσσες που χρησιμοποιήσαμε, ήταν για ακόμα μια φορά η php και η html. Αυτή την φορά θέλαμε να μοντελοποιήσουμε τον κανόνα, τον οποίο θα ακολουθεί το πρόγραμμα, πριν διεξάγει το αποτέλεσμα. Το αποτέλεσμα το οποίο προκύπτει είναι αν ο εξεταζόμενος θα πάσχει από κάποια καρδιαγγειακή ασθένεια ή όχι. Ο κανόνας διεξήχθη από το σύστημα WEKA και μετατράπηκε, σύμφωνα με προαναφερθείσες γλώσσες προγραμματισμού. Ακολουθεί ο κώδικας που χρησιμοποιήθηκε.

```

<?php
    if (isset($_POST['submit'])) {
    } else {
        header("Location: index.php");
    }
?>
<html>
<head>
    <title>Results | Cardiovascular Disease</title>
    <meta name="viewport" content="width=device-width, initial-scale=1">
    <script src="https://code.jquery.com/jquery-3.5.1.slim.min.js"
integrity="sha384-DfXdz2htPH0lsSSs5nCTpuj/zy4C+OGpamoFVy38MVBnE+IbbVYUew+OrCXaR
kfj" crossorigin="anonymous"></script>
    <script src="https://cdn.jsdelivr.net/npm/bootstrap@4.6.0/dist/js/bootstrap.bundle.min.js"
integrity="sha384-Piv4xVNRyMGpqsS2by6br4gNJ7DXjqk09RmUpJ8jgGtD7zP9yug3goQfGII0yAn
s" crossorigin="anonymous"></script>
    <link rel="stylesheet"
href="https://cdn.jsdelivr.net/npm/bootstrap@4.6.0/dist/css/bootstrap.min.css"

```

```

integrity="sha384-B0vP5xmATw1+K9KRQjQERJvTumQW0nPEzvF6L/Z6nronJ3oUOFUFpCjEUQo
uq2+1" crossorigin="anonymous">
<style>
@media (min-width: 800px) {
    .container{
        max-width: 800px;
    }
}
</style>
</head>
<body>
<nav class="navbar navbar-expand-lg bg-dark navbar-dark">
    <a class="navbar-brand" href="index.php">Cardiovascular Disease</a>
</nav>
<br />
<div class="container">
    <?php
        if (isset($_POST['submit'])) {

            $age = $_POST['age'];
            $gender = $_POST['gender'];
            $height = $_POST['height'];
            $weight = $_POST['weight'];
            $active = $_POST['active'];
            $smoke = $_POST['smoke'];
            $cholesterol = $_POST['cholesterol'];
            $gluc = $_POST['gluc'];
            $ap_hi = $_POST['ap_hi'];
            $ap_lo = $_POST['ap_lo'];

            $msg0 = "<div class='alert alert-info alert-dismissible'>Absence of cardiovascular
disease.</div> <a class='btn btn-info btn-lg btn-block' href='index.php' role='button'>Go Back</a>";

            $msg1 = "<div class='alert alert-info alert-dismissible'>Presence of cardiovascular
disease.</div> <a class='btn btn-info btn-lg btn-block' href='index.php' role='button'>Go Back</a>";

            if ($ap_hi <= 129) {
                if ($age <= 19939) {
                    if ($cholesterol = 1 | $cholesterol = 2) {
                        echo $msg0;
                    } elseif ($cholesterol = 3) {
                        if ($gluc = 1 | $gluc = 2) {
                            echo $msg1;
                        } elseif ($gluc = 3) {
                            if ($weight <= 73.5) {
                                echo $msg0;

```

```

        } elseif ($weight > 73.5) {
            echo $msg1;
        }
    }
}
} elseif ($age > 19939) {
    if ($cholesterol = 1) {
        if ($age <= 22217) {
            if ($ap_hi <= 118) {
                echo $msg0;
            } elseif ($ap_hi > 118) {
                if ($active = 0) {
                    if ($weight <= 83.5)
{
                                if ($age <=
21172) {

echo $msg0;

                                } elseif
($age > 21172) {

echo $msg1;

                                }
                            } elseif ($weight >
83.5) {

                                echo $msg1;
                            }
                        } elseif ($active = 1) {
                            echo $msg0;
                        }
                    }
                } elseif ($age > 22217) {
                    if ($active = 0) {
                        echo $msg1;
                    } elseif ($active = 1) {
                        if ($ap_hi <= 119) {
                            echo $msg0;
                        } elseif ($ap_hi > 119) {
                            if ($gender = 1) {
                                echo $msg1;
                            } elseif ($gender =
2) {

                                if ($weight
<= 76.5) {

echo $msg0;

```

```

    } elseif
($weight > 76.5) {
echo $msg1;
}
}
}
}
}
} elseif ($cholesterol = 2) {
    if ($ap_hi <= 115) {
        echo $msg0;
    } elseif ($ap_hi > 115) {
        if ($active = 0) {
            echo $msg1;
        } elseif ($active = 1) {
            if ($age <= 22057) {
                echo $msg0;
            } elseif ($age > 22057) {
                echo $msg1;
            }
        }
    }
}
} elseif ($cholesterol = 3) {
    echo $msg1;
}
}
} elseif ($ap_hi > 129) {
    if ($ap_hi <= 138) {
        if ($cholesterol = 1) {
            if ($smoke = 0) {
                if ($age <= 21739) {
                    if ($ap_lo <= 89) {
                        if ($height <= 172) {
                            if ($ap_lo <= 79) {
                                echo $msg0;
                            } elseif
($ap_lo > 79) {
                                if
($active = 0) {
                                    echo $msg0;
                                }
                                elseif ($active = 1) {
                                    if ($gender = 1) {

```

```

        echo $msg1;

    } elseif ($active = 2) {

        echo $msg0;

    }

}

}

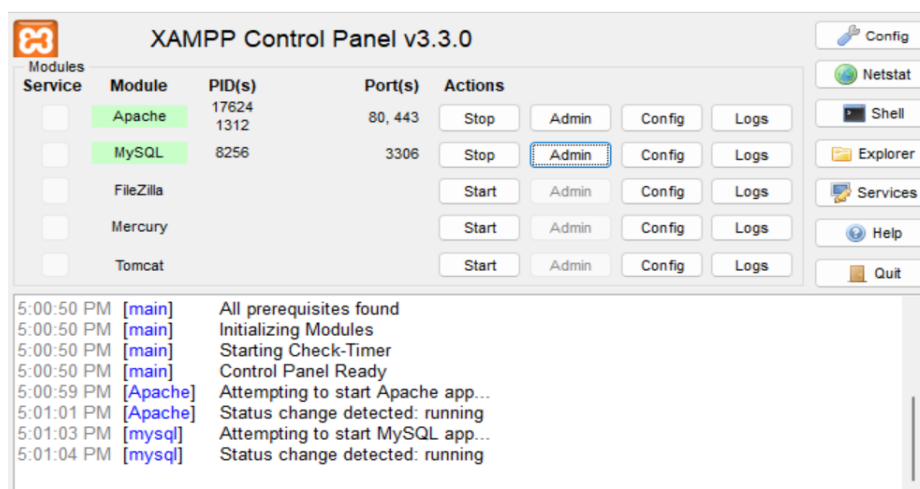
} elseif ($height > 172) {
    echo $msg0;
}
} elseif ($ap_lo > 89) {
    echo $msg1;
}
} elseif ($age > 21739) {
    echo $msg1;
}
} elseif ($smoke = 1) {
    echo $msg0;
}
} elseif ($cholesterol = 2) {
    echo $msg1;
} elseif ($cholesterol = 3) {
    echo $msg1;
}
} elseif ($ap_hi > 138) {
    echo $msg1;
}
}

}

?>
</div>
</body>
</html>

```

Τα δύο αυτά αρχεία που είδαμε, γράφτηκαν σε notepad και αποθηκεύτηκαν σε ένα κοινό φάκελο. Για να τρέξει η εφαρμογή μας, χρειάστηκε να χρησιμοποιηθεί και το πρόγραμμα XAMPP. Το XAMPP είναι ένα ελεύθερο λογισμικό, ανοικτού κώδικα και περιέχει ένα πακέτο από προγράμματα. Έπειτα από την εγκατάσταση του προγράμματος, βάλαμε στον φάκελο htdocs, που περιέχει ήδη το πρόγραμμα, τον δικό μας φάκελο με τα δύο αρχεία που είχαμε δημιουργήσει. Επόμενο μας βήμα ήταν, να ενεργοποιήσουμε το Apache και την MySQL από το πρόγραμμα, όπως φαίνεται στην εικόνα 7.1.



Εικόνα 7.1: XAMPP program

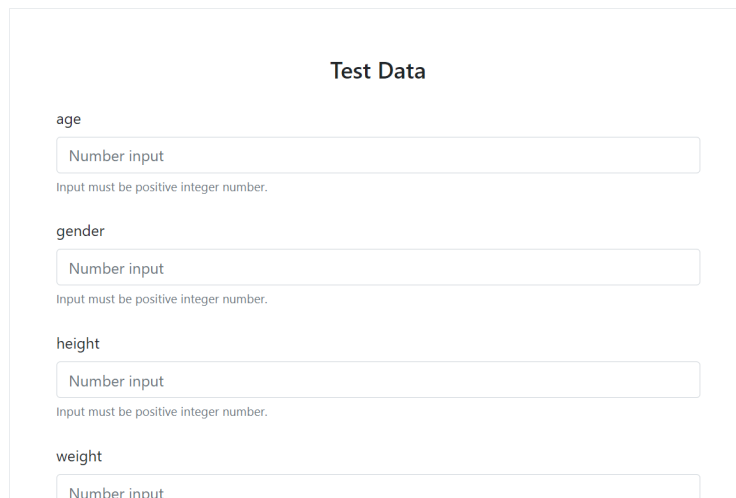
Το πρόγραμμα θα έτρεχε τοπικά, για αυτό και το αναζητήσαμε στον browser που διαθέτει ο υπολογιστής μας, απλά και μόνο γράφοντας την λέξη localhost. Με την επιλογή του enter, άνοιξαν τα αρχεία που είχαμε, στον φάκελο htdocs, του xampp προγράμματος, στον υπολογιστή μας, όπως μπορούμε να δούμε και στην εικόνα 7.2.

Index of /

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 applications.html	2019-08-27 17:02	3.5K	
 bitnami.css	2019-08-27 17:02	177	
 cardio/	2022-06-06 02:54	-	
 dashboard/	2021-12-21 22:04	-	
 favicon.ico	2015-07-16 18:32	30K	
 htdocs.zip	2022-06-06 03:09	2.5K	
 img/	2021-12-23 13:46	-	
 webalizer/	2021-12-23 13:46	-	
 xampp/	2021-12-23 13:46	-	

Εικόνα 7.2: Τοπικά αρχεία XAMPP

Εμείς από αυτά τα αρχεία επιλέξαμε το cardio, που ήταν ο φάκελος που είχαμε δημιουργήσει, με τα δύο αρχεία php και html. Με την επιλογή του φακέλου η εφαρμογή ξεκινάει να τρέχει. Το αποτέλεσμα το βλέπουμε στις εικόνες 7.3 , 7.4 και 7.5 που ακολουθούν.



Test Data

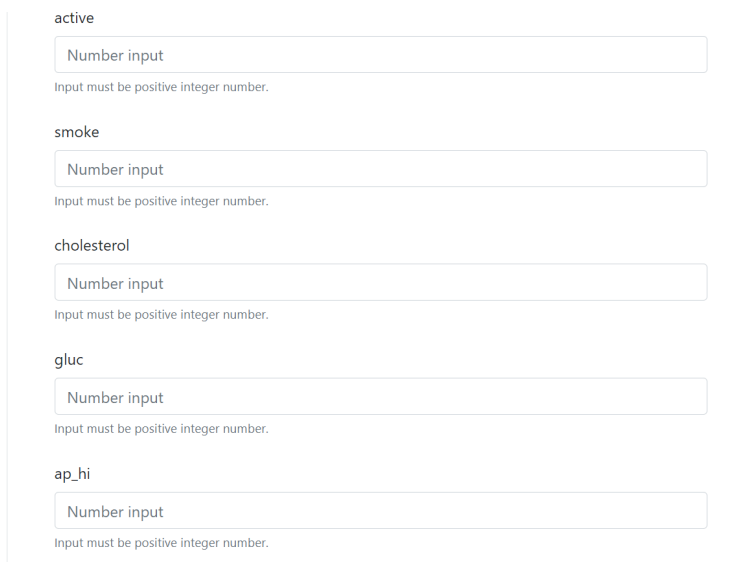
age
Number input
Input must be positive integer number.

gender
Number input
Input must be positive integer number.

height
Number input
Input must be positive integer number.

weight
Number input

Εικόνα 7.3: Screenshot εφαρμογής (1)



active
Number input
Input must be positive integer number.

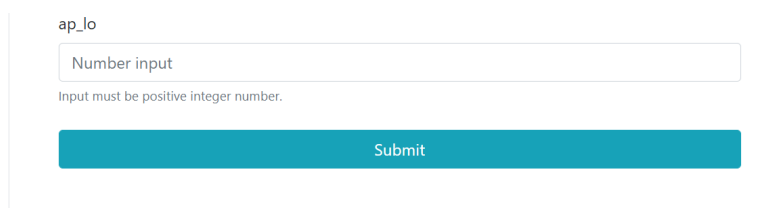
smoke
Number input
Input must be positive integer number.

cholesterol
Number input
Input must be positive integer number.

gluc
Number input
Input must be positive integer number.

ap_hi
Number input
Input must be positive integer number.

Εικόνα 7.4: Screenshot εφαρμογής (2)



ap_lo
Number input
Input must be positive integer number.

Submit

Εικόνα 7.5: Screenshot εφαρμογής (3)

Σε αυτό το σημείο δοκιμάσαμε την ποιότητα της εφαρμογής. Προκειμένου να εξετάσουμε το πόσο αποτελεσματική είναι. Πήραμε έτοιμες περιπτώσεις, από το αρχικό dataset, που χρησιμοποιήσαμε για την εκπαίδευση του συστήματος και εισάγαμε τις τιμές τους, για να επαληθεύσουμε πως το αποτέλεσμα θα ήταν ίδιο με αυτό της κλάσης. Αν το αποτέλεσμα ήταν ίδιο, θα

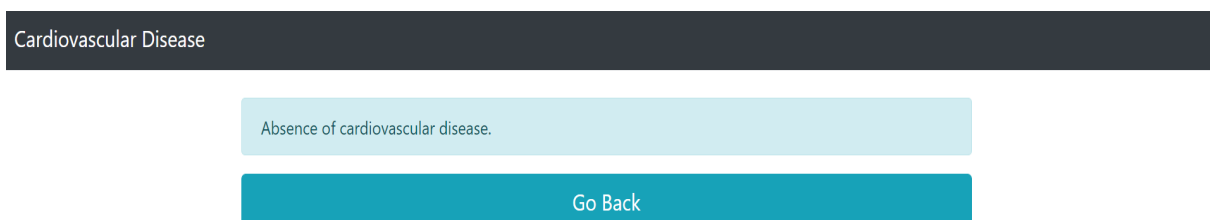
επιβεβαίωσε την εγκυρότητα της εφαρμογής μας, αλλιώς θα έπρεπε να επανεξεταστεί και πιθανόν να δημιουργηθεί εκ νέου μια.

Πήραμε τυχαία, τις ακόλουθες τιμές:

Age	16038
Gender	1
Height	171
Weight	71
Systolic blood pressure	120
Diastolic blood pressure	80
Cholesterol	1
Glucose	1
Smoking	0
Physical activity	0
Presence or absence of cardiovascular disease	0

Πίνακας 12: Τυχαίες τιμές του dataset

Το σύστημα, σε περιπτώσεις μη ύπαρξης καρδιακής νόσου, εμφανίζει το μήνυμα absence of cardiovascular disease, δηλαδή απουσία καρδιαγγειακής νόσου, ενώ σε περιπτώσεις ύπαρξης καρδιαγγειακής νόσου εμφανίζει το μήνυμα Presence of cardiovascular disease, που σημαίνει παρουσία καρδιαγγειακών παθήσεων. Βάζοντας τις τιμές στην εφαρμογή, αναμέναμε το αποτέλεσμα να δείξει, ότι δεν υπάρχει κάποια καρδιαγγειακή ασθένεια και πράγματι το μήνυμα που μας εμφάνισε ήταν το παρακάτω.



Εικόνα 7.6: Αποτέλεσμα δοκιμής

Την διαδικασία αυτή ακολουθήσαμε για άλλες 10 περιπτώσεις, με σκοπό να μετρήσουμε τις φορές που θα κάνει λάθος πρόβλεψη. Από τις τυχαίες δοκιμές που κάναμε, βασισμένοι στα αρχικά δεδομένα που είχαμε διεξήγαμε τα ακόλουθα αποτελέσματα. Κάναμε 10 δοκιμές, εκ των οποίων οι 7 ταξινομήθηκαν σωστά, ενώ οι 3 λάθος, με βάση την τελική κλάση που μας έδινε το αρχικό dataset. Τα στοιχεία που εξετάστηκαν και τα αποτελέσματα που λάβαμε βρίσκονται στον πίνακα 12.

Age	16038	20228	18393	20425	22447	18392	17623	22584	19809	15243
Gender	1	1	2	2	1	1	2	2	1	2
Height	171	156	168	175	154	165	169	178	158	171
Weight	71	85	62	75	69	65	82	95	78	67
Physical activity	0	1	1	1	1	0	1	1	1	1
Smoking	0	0	0	0	0	1	0	0	0	0
Cholesterol	1	3	1	1	2	2	1	3	1	1
Glucose	1	1	1	1	1	1	1	3	1	1
Systolic blood pressure	120	140	110	140	160	120	150	130	110	110
Diastolic blood pressure	80	90	80	90	100	80	100	90	70	60
Class	0	1	0	1	0	1	1	1	0	0
Αποτέλεσμα συστήματος	0	1	0	1	1	0	1	0	0	0

Πίνακας 13: Δέκα τυχαίες δοκιμές και αποτελέσματα

Βασικό μειονέκτημα του συστήματος είναι πως, σε περιπτώσεις, όπου κάποιος πληκτρολογήσει λανθασμένα αποτελέσματα, είτε εσκεμμένα, είτε από αμέλεια, το σύστημα δεν μπορεί να το αναγνωρίσει και διεξάγει αποτέλεσμα, με βάση τα στοιχεία που έχει λάβει. Εκτελεί κανονικά τον κανόνα και δίνει απάντηση στον χρήστη, χωρίς να ελέγξει την εγκυρότητα των όσων έχει λάβει.

Κεφάλαιο Η'

Συμπεράσματα

Οι καρδιαγγειακές παθήσεις είναι ένα μείζον πρόβλημα που απασχολεί ιδιαίτερα την επιστήμη της ιατρικής τα τελευταία χρόνια. Εξαιτίας του γεγονότος ότι πρόκειται, για τις πιο σύνθετες αιτίες θανάτου και επιφέρει μεγάλο κόστος στην κοινωνία, στην οικονομία, αλλά και στην ψυχολογία του ατόμου και του κοινωνικού συνόλου, έχει πάρει σαν πρόβλημα τεράστιες εκτάσεις παγκοσμίως. Λύση στο πρόβλημα έρχεται να δώσει η πληροφορική με τα συστήματα υποστήριξης αποφάσεων και με την μηχανική μάθηση. Τα στατιστικά μοντέλα, σταματάνε να επιλέγονται, για την πρόβλεψη των καρδιαγγειακών ασθενειών και την θέση τους παίρνουν ολοένα και περισσότερο μοντέλα που βασίζονται στην μηχανική μάθηση. Ο λόγος για τον οποίο συμβαίνει αυτό, είναι ότι τα στατιστικά μοντέλα, δεν μπορούν να αποθηκεύσουν και να κατηγοριοποιήσουν μεγάλο όγκο τιμών. Στην μηχανές μηχανικής μάθησης, γίνεται ένας συνδυασμός από τεχνικές εξόρυξης δεδομένων, για την ταξινόμηση διαφορετικών ειδών, με βάση προηγούμενα αρχεία και δεδομένα ασθενών. Ο πιο σημαντικός παράγοντας, στην εξόρυξη δεδομένων και αυτό που προσπαθούν να πετύχουν οι περισσότερες βιβλιογραφικές αναφορές και οι ιατρικές βιομηχανίες, είναι η ακρίβεια. Είναι ένα από τα πιο δύσκολα μέρη στην πρόβλεψη των καρδιαγγειακών παθήσεων. Η αναζήτηση του αλγορίθμου με την καλύτερη απόδοση είναι πολύ σημαντική και η δοκιμή αρκετών αλγορίθμων κρίνεται σχεδόν απαραίτητη. Στην δική μας περίπτωση, έπειτα από αρκετές δοκιμές, με διαφορετικούς αλγορίθμους και ύστερα από ανάλυση μεταξύ τους, καταλήξαμε στο συμπέρασμα, πως ο αλγόριθμος που κατάφερε να ανταπεξέλθει καλύτερα, στην δική μας περίπτωση, ήταν ο J48. Οι ταξινομήσεις κατά την φάση του training, ήταν 73.3576 % σωστές και από τις 10 δοκιμές που κάναμε σε τυχαία δείγματα από το αρχικό dataset, προέβλεψε σωστά τις 7 ενώ στις 3 έδειξε αντίθετο αποτέλεσμα. Συνεπώς ένα σύστημα σαν αυτό με περισσότερο training θα μπορούσε να αποτελέσει μια μηχανή “σύμβουλο” για τους γιατρούς. Προβλέποντας τις καρδιαγγειακές παθήσεις, θα βοηθάει τον γιατρό να προλαμβάνει τον εξεταζόμενο, πριν γίνει ασθενής, μειώνοντας τις περιπτώσεις που οι άνθρωποι πάσχουν από κάποια καρδιαγγειακή νόσο, ενώ παράλληλα θα μειώνονται και οι επιπτώσεις.

Κεφάλαιο Θ’

Μελλοντικές Προεκτάσεις

Ένας πολύ σημαντικός παράγοντας για την ομαλή λειτουργία των συστημάτων, που βασίζονται στην μηχανική μάθηση, είναι η κατανόηση της λειτουργίας τους. Η έλλειψη ξεκάθαρης κατανόησης, του τρόπου με τον οποίο, η πληροφορία από τις μεταβλητές εισόδου, συσχετίζεται με τους προγνωστικούς παράγοντες εκβάσεων, δημιουργούν αβεβαιότητα και έλλειψη εμπιστοσύνης, στην ενσωμάτωση των μηχανών λήψης αποφάσεων, που βασίζονται στην μηχανική μάθηση, στην επιστήμη της ιατρικής και όχι μόνο. Ένα υποδοκιμασμένο μοντέλο, το οποίο πιθανόν να είναι ακατάλληλο, μπορεί να αποτελέσει τεράστιο κίνδυνο, για την δημόσια υγεία. Αντίθετα οι δοκιμές με σύνολα δεδομένων, με στόχο την επικύρωση και την επίτευξη της μέγιστης ακρίβειας, εμπνέει εμπιστοσύνη και πολύ πιθανόν να οδηγήσει, σε μελλοντική εφαρμογή του μοντέλου. Η δυνατότητα σωστής ερμηνεία των δεδομένων δεν είναι ευρέως διαδεδομένη για τα μοντέλα που βασίζονται στην μηχανική μάθηση, όσο είναι για τις στατιστικές έρευνες μεταξύ των κλινικών ιατρών. Ως εκ τούτου, η ενσωμάτωση προγνωστικών μοντέλων, σε όλο το σύστημα, δεν είναι μια πρακτική των νοσοκομείων ή των κλινικών, επί του παρόντος. Η μοντελοποίηση, η οποία βασίζεται στην τεχνητή νοημοσύνη, καθοδηγείται σε έναν μεγάλο βαθμό από επιστήμονες δεδομένων. Κάτι τέτοιο φυσικά κάνει την διαχείριση των δεδομένων, τον προγραμματισμό και την εφαρμογή των κατάλληλων αλγορίθμων ιδιαίτερα πολύπλοκα, με αποτέλεσμα να γίνεται αποθαρρυντική εργασία για του κλινικούς γιατρούς. Για αυτόν λοιπόν τον λόγο, συνιστούμε να καθιερωθεί μια ομαδική προσέγγιση κατά την οποία θα περιλαμβάνονται στατιστικολόγοι, επιστήμονες δεδομένων και κλινικοί γιατροί.

Ένα ακόμα ζήτημα, που εμποδίζει ιδιαίτερα το κοινό, στο να συμμετέχει με τον δικό του τρόπο, είναι η ασφάλεια του συστήματος. Πολλοί αμφισβητούν την ασφάλεια των πληροφοριών τους και θεωρούν ότι από ένα σύστημα σαν αυτό, απουσιάζει το ιατρικό απόρρητο. Για αυτόν λοιπόν το λόγο, θα πρέπει το σύστημα να είναι πλήρως ελεγχμένο, από πλευράς ασφάλειας για να κερδίσει την εμπιστοσύνη των ασθενών.

Σε ένα ακόμα πιο εξελιγμένο σύστημα, θα μπορούσαν να υπάρχουν προεκτάσεις. Συγκεκριμένα θα μπορούσε εκτός από την ένδειξη κινδύνου καρδιαγγειακής πάθησης, να εμφανίζεται και ποια θα είναι αυτή την πάθηση. Με αυτό τον τρόπο οι γιατροί, θα μπορούν να εστιάσουν ακόμα περισσότερο, στο ποιά είναι η κατάλληλη θεραπεία πρόληψης.

Αν και η τεχνητή νοημοσύνη, είναι ευρέως διαδεδομένη, στις εμπορευματοποιημένες βιομηχανίες, εξακολουθεί να υστερεί στον κλάδο της ιατρικής. Προσωπική μας άποψη είναι όμως, ότι η τεχνητή νοημοσύνη και η μηχανική μάθηση, μπορεί να μεταμορφώσει την ιατρική βιομηχανία. Έχει την δυνατότητα να προσφέρει αποτελέσματα με μεγάλη ιατρική ακρίβεια, τα οποία θα γίνουν σύμβουλοι των ειδικών, δίνοντας λύση σε μεγάλα ερωτήματα της ιατρικής, ένα εκ των οποίων είναι και οι καρδιαγγειακές παθήσεις.

Βιβλιογραφία / Δικτυογραφία

- [1]Al'Aref SJ, Anchouche K, Singh G, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J*. 2019;40(24):1975-1986. doi:10.1093/eurheartj/ehy404
- [2]Rachana R Sanni, H.S. Guruprasad, Analysis of performance metrics of heart failed patients using Python and machine learning algorithms, *Global Transitions Proceedings*, Volume 2, Issue 2, 2021, Pages233-237, ISSN2666-285X, (<https://www.sciencedirect.com/science/article/pii/S2666285X2100056X>)
- [3]Ελληνική Καρδιολογική Εταιρεία, “Μάθε για την Καρδιακή Ανεπάρκεια”, 1 Μαρτίου 2022, pages 8, (<https://www.hcs.gr/astheneis-koino/enimerotiko-yliko/kardiaki-aneparkeia-2/>)
- [4] Λεμπιδάκης Δημήτρης, Λουλακάκης Ματθαίος, Χαμηλός Μιχάλης, Σκαλίδης Εμμανουήλ, “Στεφανιαία νόσος: Ένας κατά συρροή δολοφόνος ζει δίπλα(;) μας”, 25 Νοεμβρίου 2021, page 1, (<https://www.hcs.gr/astheneis-koino/stefaniaia-nosos/>)
- [5]Χριστίνα Χρυσοχόου, Κωνσταντίνος Τσιούφης, “Μάθε για την αρτηριακή σου πίεση! Έχω αρτηριακή πίεση αυξημένη Τι πρέπει να γνωρίζω” , 25 Νοεμβρίου 2021, page 1, (<https://www.hcs.gr/astheneis-koino/artiriaki-piesi/>)
- [6] Παγκόσμιας Ημέρας Νεφρού 2022, Ελληνική Νεφρολογική Εταιρεία, Διαδικτυακή Ημερίδα Ενημέρωσης, 10 Μαρτίου 2022 (<http://www.worldkidneyday.gr/>)
- [7]ΚΑΡΔΙΑΚΗ ΑΝΕΠΑΡΚΕΙΑ ΑΠΟ ΤΟΝ κ. Σ.ΚΑΚΟΥΡΟ , Ελληνική Καρδιολογική Εταιρεία, 25 Νοεμβρίου 2021, (<https://www.hcs.gr/astheneis-koino/kardiaki-aneparkeia/>)
- [8]Heart Failure and Cardiovascular Diseases - A European Heart Network Paper, cehn fighting heart disease and stroke, 22 April 2019, (<https://ehnheart.org/publications-and-papers/publications/1202:heart-failure-and-cardiovascular-diseases.html>)
- [9]ESC Atlas of Cardiology (<https://www.escardio.org/Research/ESC-Atlas-of-cardiology>)
- [10]Καρδιαγγειακές παθήσεις (CVD), World Health Organization, 11/ Ιουνίου/ 2021, (<https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-cvds>)
- [11]Brunner-La Rocca HP, Fleischhacker L, Golubnitschaja O, et al. Challenges in personalised management of chronic diseases-heart failure as prominent example to advance the care process. *EPMA J*. 2016;7(1):2. Published 2016 Jan 30. doi:10.1186/s13167-016-0051-9
- [12]Schjoedt I, Sommer I, Bjerrum MB. Experiences and management of fatigue in everyday life among adult patients living with heart failure: a systematic review of qualitative evidence. *JBIS Database System Rev Implement Rep*. 2016;14(3):68-115. doi:10.11124/JBISRIR-2016-2441
- [13]Fry M, McLachlan S, Purdy S, Sanders T, Kadam UT, Chew-Graham CA. The implications of living with heart failure; the impact on everyday life, family support, co-morbidities and access to healthcare: a secondary qualitative analysis. *BMC Fam Pract*. 2016;17(1):139. Published 2016 Sep 26. doi:10.1186/s12875-016-0537-5
- [14]Marko Bohanec, Gennaro Tartarisco, Flavia Marino, Giovanni Pioggia, Paolo Emilio Puddu, Michele Salvatore Schiariti, Anneleen Baert, Sofie Pardeens, Els Clays, Aljoša Vodopija, Mitja Luštrek, HeartMan DSS: A decision support system for self-management of congestive heart failure,

Expert Systems with Applications, Volume 186, 2021, 115688, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2021.115688>.

[15] Saiyed Faiyaz Waris, S. Koteeswaran, Heart disease early prediction using a novel machine learning method called improved K-means neighbor classifier in python, Materials Today: Proceedings, 2021, ISSN 2214-7853, <https://doi.org/10.1016/j.matpr.2021.01.570>.

[16] Javeed A, Khan SU, Ali L, Ali S, Imrana Y, Rahman A. Machine Learning-Based Automated Diagnostic Systems Developed for Heart Failure Prediction Using Different Types of Data Modalities: A Systematic Review and Future Directions. Comput Math Methods Med. 2022;2022:9288452. Published 2022 Feb 3. doi:10.1155/2022/9288452

[17] Patel B, Sengupta P. Machine learning for predicting cardiac events: what does the future hold?. Expert Rev Cardiovasc Ther. 2020;18(2):77-84. doi:10.1080/14779072.2020.1732208

[18] Ambale-Venkatesh B, Yang X, Wu CO, et al. Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis. Circ Res. 2017;121(9):1092-1101. doi:10.1161/CIRCRESAHA.117.311312

[19] L. Ali et al., "An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart Failure," in IEEE Access, vol. 7, pp. 54007-54014, 2019, doi: 10.1109/ACCESS.2019.2909969.

[20] Liaqat Ali, Shafqat Ullah Khan, Noorbakhsh Amiri Golilarz, Imrana Yakubu, Iqbal Qasim, Adeeb Noor, Redhwan Nour, "A Feature-Driven Decision Support System for Heart Failure Prediction Based on Statistical Model and Gaussian Naive Bayes", Computational and Mathematical Methods in Medicine, vol. 2019, Article ID 6314328, 8 pages, 2019. <https://doi.org/10.1155/2019/6314328>

[21] Νικόλαος Ματσατσίνης “ Συστήματα Υποστήριξης Αποφάσεων” 978-960-6759-44-4, ΕΚΔΟΣΕΙΣ ΝΕΩΝ ΤΕΧΝΟΛΟΓΙΩΝ, Αριθμός Σελίδων: 1008, κεφ 6ο pg 301-330

[22] L. Ali, S.A.C. Bukhari, “ An Approach Based on Mutually Informed Neural Networks to Optimize the Generalization Capabilities of Decision Support Systems Developed for Heart Failure Prediction”, IRBM, Volume 42, Issue 5, 2021, Pages 345-352, ISSN 1959-0318, <https://doi.org/10.1016/j.irbm.2020.04.003>.

[23] Dwivedi, A.K. Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural Comput&Applic* 29, 685–693 (2018). <https://doi.org/10.1007/s00521-016-2604-1>

[24] A. Javeed, S. Zhou, L. Yongjian, I. Qasim, A. Noor and R. Nour, "An Intelligent Learning System Based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection," in IEEE Access, vol. 7, pp. 180235-180243, 2019, doi: 10.1109/ACCESS.2019.2952107.

[25] C. Beulah Christalin Latha, S. Carolin Jeeva, Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques, Informatics in Medicine Unlocked, Volume 16, 2019, 100203, ISSN 2352-9148, <https://doi.org/10.1016/j.imu.2019.100203>.

[26] J. Chen, A. Valehi and A. Razi, "Smart Heart Monitoring: Early Prediction of Heart Problems Through Predictive Analysis of ECG Signals," in IEEE Access, vol. 7, pp. 120831-120839, 2019, doi: 10.1109/ACCESS.2019.2937875.

[27] Paul, A., Shill, P., Rabin, M. et al. Adaptive weighted fuzzy rule-based system for the risk level assessment of heart disease. Appl Intell 48, 1739–1756 (2018). <https://doi.org/10.1007/s10489-017-1037-6>

[28] Syed Muhammad Saqlain Shah, Safeera Batool, Imran Khan, Muhammad Usman Ashraf, Syed Hussnain Abbas, Syed Adnan Hussain, Feature extraction through parallel Probabilistic Principal Component Analysis for heart disease diagnosis, Physica A: Statistical Mechanics and its Applications, Volume 482, 2017, Pages 796-807, ISSN 0378-4371, <https://doi.org/10.1016/j.physa.2017.04.113>.

- [29] A Systems View of Cardiovascular Disease, Madisen Cotter, Portland State University Follow, First Advisor Leslie McBride, Date of Award 5-24-2019, Document Type Thesis, Degree Name Bachelor of Science (B.S.) in Health Studies: Health Sciences and University Honors, Department Health Studies, Subjects Cardiovascular system -- Diseases -- Risk factors, System theory, DOI 10.15760/honors.793, <https://pdxscholar.library.pdx.edu/honorsthesis/775/>
- [30] Volume: 25 issue: 3, page(s): 225-230 Article first published online: July 18, 2019; Issue published: September , 2019 Eirini Bathrellou, Meropi D Kontogianni, Evaggelia Chrysanthopoulou, Ekavi Georgousopoulou, Christina Chrysohoou, Christos Pitsavos, Demosthenes Panagiotakos Department of Nutrition & Dietetics, School of Health Science and Education, Harokopio University, Athens, Greece First Cardiology Clinic, School of Medicine, University of Athens, Greece
- [31] A. P. Pawlovsky, "An ensemble based on distances for a kNN method for heart disease diagnosis," 2018 International Conference on Electronics, Information, and Communication (ICEIC), 2018, pp. 1-4, doi: 10.23919/ELINFOCOM.2018.8330570.
- [32] Jan M, Awan AA, Khalid MS, Nisar S. Ensemble approach for developing a smart heart disease prediction system using classification algorithms. Research Reports in Clinical Cardiology. 2018;9:33-45 <https://doi.org/10.2147/RRCC.S172035>
- [33] Lal Hussain, Imtiaz Ahmed Awan, Wajid Aziz, Sharjil Saeed, Amjad Ali, Farukh Zeeshan, Kyung Sup Kwak, "Detecting Congestive Heart Failure by Extracting Multimodal Features and Employing Machine Learning Techniques", BioMed Research International, vol. 2020, Article ID 4281243, 19 pages, 2020. <https://doi.org/10.1155/2020/4281243>
- [34] Maji, S., Arora, S. (2019). Decision Tree Algorithms for Prediction of Heart Disease. In: Fong, S., Akashe, S., Mahalle, P. (eds) Information and Communication Technology for Competitive Strategies. Lecture Notes in Networks and Systems, vol 40. Springer, Singapore. https://doi.org/10.1007/978-981-13-0586-3_45
- [35] J. B. S K, P. N. Sudha, R. S, S. A. A Kashyap and A. L, "Machine Learning Techniques for Cardiovascular Risk Score -Prediction," 2021 IEEE Mysore Sub Section International Conference (MysuruCon), 2021, pp. 505-509, doi: 10.1109/MysuruCon52639.2021.9641581.
- [36] A. Moreno, J. Rodriguez and F. Martínez, "Regional Multiscale Motion Representation for Cardiac Disease Prediction," 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA), 2019, pp. 1-5, doi: 10.1109/STSIVA.2019.8730231.
- [37] Imran Shafi, Aesar Aziz, Sadia Din, Imran Ashraf, Reduced features set neural network approach based on high-resolution time-frequency images for cardiac abnormality detection, Computers in Biology and Medicine, Volume 145, 2022, 105425, ISSN 0010-4825, <https://doi.org/10.1016/j.compbiomed.2022.105425>.
- [38] Ιωαννης Αποστολάκης, Ηρακλής Βαρλαμης “Πληροφοριακά Συστήματα Υγείας” Ηλεκτρονικές υπηρεσίες υγείας , τρίτη βελτιωμένη έκδοση, εκδόσεις Παπαζήση, σειρά: Κοινωνικές επιστήμες & υγεια , κεφάλαιο 5, σελ: 115-135
- [39] Age dependent associations of risk factors with heart failure: pooled population based cohort study BMJ 2021; 372 doi: <https://doi.org/10.1136/bmj.n461> (Published 23 March 2021) BMJ 2021;372:n461
- [40] Filippou Tripodskiadis, Andrew Xanthopoulos, Javed Butler, Cardiovascular Aging and Heart Failure: JACC Review Topic of the Week, Journal of the American College of Cardiology, Volume 74, Issue 6, 2019, Pages 804-813, ISSN 0735-1097, <https://doi.org/10.1016/j.jacc.2019.06.053>.
- [41] Zujie Gao, Zengsheng Chen, Anqiang Sun, Xiaoyan Deng, Gender differences in cardiovascular disease, Medicine in Novel Technology and Devices, Volume 4, 2019, 100025, ISSN 2590-0935, <https://doi.org/10.1016/j.medntd.2019.100025>.

- [42]Moon J, Hwang I, C: The Link between Height and Cardiovascular Disease: To Be Deciphered. *Cardiology* 2019;143:114-115. doi: 10.1159/000502032
- [43]Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association Published in *Circulation*, May 2021 DOI 10.1161/cir.0000000000000973 Pubmed ID 33882682 Authors Tiffany M. Powell-Wiley, Paul Poirier, Lora E. Burke, Jean-Pierre Després, Penny Gordon-Larsen, Carl J. Lavie, Scott A. Lear, Chiadi E. Ndumele, Ian J. Neeland, Prashanthan Sanders, Marie-Pierre St-Onge
- [44]Strandberg, Timo E.; Pitkala, Kaisu What is the most important component of blood pressure: systolic, diastolic or pulse pressure?, *Current Opinion in Nephrology and Hypertension*: May 2003 - Volume 12 - Issue 3 - p 293-297
- [45]Kristian B. Filion, Russell V. Luepker, Cigarette Smoking and Cardiovascular Disease: Lessons from Framingham, *Global Heart*, Volume 8, Issue 1, 2013, Pages 35-41, ISSN 2211-8160, <https://doi.org/10.1016/j.gheart.2012.12.005>.
- [46]Toma, A., Paré, G. & Leong, D.P. Alcohol and Cardiovascular Disease: How Much is Too Much?. *Curr Atheroscler Rep* 19, 13 (2017). <https://doi.org/10.1007/s11883-017-0647-0>
- [47]Edoardo Mannucci, Ilaria Dicembrini, Angelo Lauria, Paolo Pozzilli; Is Glucose Control Important for Prevention of Cardiovascular Disease in Diabetes?. *Diabetes Care* 1 August 2013; 36 (Supplement_2): S259–S263. <https://doi.org/10.2337/dcS13-2018>
- [48]Wallimann-Annema W. Aktueller Stellenwert der HDL-Cholesterin-Bestimmung für die kardiovaskuläre Risikoabschätzung [The Current Significance of Measuring HDL-Cholesterol in Cardiovascular Risk Assessment]. *Praxis* (Bern 1994). 2021;110(7):383-390. doi:10.1024/1661-8157/a003677
- [49] Αρχικό dataset, Cardiovascular Disease dataset, The dataset consists of 70 000 records of patients data, 11 features + target, SVETLANA ULIANOVA· UPDATED 3 YEARS AGO <https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset>

Συντμήσεις

ML	Machine Learning
KA	Καρδιακή Ανεπάρκεια
VHD	Βαλβιδική Καρδιακή Νόσο
E.E.	Ευρωπαϊκή Ένωση
WHO	World Health Organization
DALYS	Disability-adjusted life year
W.H.F	World Health Federation
DSS	Decision Support System
ΑΕΠ	Ακαθάριστο Εγχωριο Προϊον
MESA	Multi-Ethnic Study of Atherosclerosis
NT-proBNP	Ventricular natriuretic peptide
SVMs	Support-vector machines
GNB	Gaussian Naive Bayes
ROC	Receiver operating characteristic
RSA	Random search algorithm
ECG	Σήματα ηλεκτροκαρδιογραφήματος
Ap_hi	Systolic blood pressure
Ap_lo	Diastolic blood pressure
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative

Γλωσσάρι Ξενικών Όρων

Machine Learning	Μηχανική Μάθηση
World Health Organization	Παγκόσμιος Οργανισμός Υγείας
Disability-adjusted life year	Έτη ζωής προσαρμοσμένα στην αναπηρία
World Health Federation (W.H.F)	Παγκόσμια Ομοσπονδία Υγείας
Decision Support System (DSS)	Συστήματα Υποστήριξης Αποφάσεων
Multi-Ethnic Study of Atherosclerosis	Πολυεθνική Μελέτη Αθηροσκλήρωσης
Random Forest	Τυχαία Δάση
Ventricular natriuretic peptide	Νατριουρητικό Πεπτίδιο
Support-vector machines	Μηχανές διανυσμάτων υποστήριξης
Best-first search strategy	Αλγόριθμος καλύτερη πρώτη αναζήτηση
Random search algorithm	Αλγόριθμο τυχαίας αναζήτησης
Challenge	Πρόκληση
Interface	Διεπαφή
Gender	Φύλο
Age	Ηλικία
Height	Ύψος
Weight	Βάρος
Smoking	Κάπνισμα
Alcohol	Αλκοόλ
Cholesterol	Χοληστερίνη
Glucose	Γλυκόζη

Physical activity	Φυσική Άσκηση
Systolic blood pressure	Συστολική αρτηριακή πίεση
Diastolic blood pressure	Διαστολική αρτηριακή πίεση
Preprocess	Προεπεξεργασία
Classify	Ταξινόμηση
Dataset	Σύνολο δεδομένων
Balance	Ισορροπία